© Copyright 2022

Emily Hsieh

Using a CRISPR-based screen to study individual and combination molecular mechanisms of HIV-1 latency

Emily Hsieh

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Michael Emerman, Chair

Julie Overbaugh

Daphne C. Avgousti

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

Abstract

Using a CRISPR-based screen to study individual and combination molecular mechanisms of HIV-1 latency

Emily Hsieh

Chair of the Supervisory Committee:

Michael Emerman

Department of Microbiology and Department of Global Health

Long-lived immune cells harboring transcriptionally silenced, latent Human Immunodeficiency Virus Type 1 (HIV-1) proviruses continues to pose a barrier to HIV-1 cure. HIV-1 latency involves a breadth of host and viral factors that influence transcription initiation, transcription elongation, and chromatin modifications of the long terminal repeat (LTR) of the HIV-1 provirus. These encompass both positive and negative regulators of transcription including factors that influence the activity of transcription factors such as NF-κB, factors that influence the ability of the viral transactivator, Tat, to recruit the elongation complex P-TEFb to the viral promoter, and factors that influence both activating and repressive modifications to histones surrounding the HIV-1 LTR. Thus, to examine the interplay of overlapping mechanisms involved in HIV-1 latency, I have developed a new HIV-1 latency CRISPR-based screening

strategy that combines screening for multiple pathways by targeting one pathway while simultaneously activating another transcriptional mechanism. In this thesis, I established a high-throughput CRISPR-based screen named Latency HIV-CRISPR that uses the packaging of guideRNAs into the viral supernatant as a direct readout of factors involved in the maintenance of HIV-1 latency. I devised a strategy for latency reversal that uses a custom generated guideRNA library targeting epigenetic regulatory genes combined with a treatment with or without AZD5582, an activator of the non-canonical NF-κB pathway and a latency reversal agent. This screen identified novel individual and combination pathways that contribute to HIV-1 latency. These studies provide progress to the goal of clearance and ultimately elimination of the latent reservoir of HIV-1 infected cells using pathways that increase both the potency and specificity of latency reversal agents.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

This work would not have been possible without the support, guidance, and collaboration of many individuals. The work in chapter 2 was a result of an exciting collaboration with Molly OhAinle and Derek Janssens. I also want to thank additional collaborators including Steve Henikoff, Ed Browne, and Patrick Paddison for their insight. Additionally, I am grateful for input from Steve Hahn and B. Matija Peterlin, latency discussions from the CARE Collaboratory, primary cell latency model advice from Jackson Peterson, and cell lines from Warner Greene and Darell Bigner. This work was financially supported by the NSF Graduate Research Fellowship (DGE-1256082, DGE-1762114) and the Viral Pathogenesis and Evolution Training Grant (T32 AI 083203).

To the Emerman lab, past and present, I am deeply grateful for the supportive training environment. I want to thank Michael Emerman, my advisor, for his thoughtful mentorship and eternal optimism that has challenged and encouraged me to be a better scientist and more open-minded to follow the data. I also want to thank Molly OhAinle for her enthusiastic mentorship and creative perspectives to science that have been helpful for my growth as a scientist. I am also grateful for the Latency Reversal Agents team of Carley Gray, Terry Hafer, and Yennifer Delgado for helping to build a vibrant HIV-1 latency community. Thank you to the many members of the Hutch community that foster an incredibly exciting learning environment.

Specifically, thank you to Jasmine Gonzalez for administrative support; Daryl Humes and Hannah Itell for primary cell discussions; Lucas Carter, Rafal Donczew, and Annabel Olson for experimental advice; Jorja Henikoff, Matt Fitzgibbon, and Pritha Chanana for bioinformatics support; Fred Hutch flow cytometry folks for their support; Andy Marty, Alyssa Dawson, Dolores Covarrubias, and Elizbeth Jensen for their genomics help; Luna Yu, Pat Heath, and David Chambers for resolving all things IT related; and Maia Low and Andrea Brocato for MCB related support.

Thank you to my thesis committee members, Daphne Avgousti, Steve Hahn, Emily Hatch, and Julie Overbaugh for their insight and emphasis on thinking about the project from a big picture view. A special thank you to Daphne and Julie for reading and providing helpful feedback on my thesis. I would also like to thank my previous scientific mentors who have all contributed to my scientific career including Harmit Malik, Ben Kerr, Ben tenOever, Nitin Phadnis, Mia Levine, and Ines Anna Drinnenberg.

I am also very grateful for my amazing support group of friends who are always there for me, no matter where they are in the world. To Johnny, my husband, I am so thankful for your support every step of the way. To Rebecca, my sister, thank you for always being candid and finding humor in all situations. Last of all, to my parents, thank you for your countless sacrifices and continued support to make all of this possible.

Chapter 1. Introduction

Over 37 million people across the globe are living with Human Immunodeficiency Virus Type 1 (HIV-1) [8], making it one of the world's most serious public health challenges. Current formulations of antiretroviral therapy (ART) are highly effective in suppressing replication of HIV-1. However, ART is not a cure, since, except for rare exceptional cases, virus replication resumes within a few weeks if ART is interrupted [9]. This clinical observation indicates the presence of a latent HIV-1 reservoir that cannot be cleared by ART. A small subset of such infected cells continues to exist in a silent state resulting in the persistence of the viral genome in the host, which defines the latent reservoir. Using a novel CRISPR screening technique incorporating small molecule activators of HIV-1 latent proviruses, the investigation of the molecular pathways that maintain HIV-1 latency and the interactions of these pathways are the subject of this thesis.

1.1 Integration of the HIV-1 provirus and viral latency

Human Immunodeficiency Virus Type 1 (HIV-1) is a member of the lentivirus genus in the family of *Retroviridae* (retroviruses). Retroviruses infect most, if not all, vertebrate species, and related sequences are found throughout the eukaryotic tree. These viruses have a life cycle with defining features including the conversion of viral nucleic material from RNA to DNA through a process of reverse transcription and the ability to efficiently integrate into host cell genomes. Integration of proviral DNA into host DNA is an obligate step of the viral lifecycle, and it is an irreversible process. As a result, integration confers the virus the ability to persist in the host and to proliferate as cells divide. It is also this important step that allows for the virus to

go into a process of latency and generate a latent reservoir in which the virus makes little or no viral proteins and thus escapes elimination by host immune responses.

Early in the study of HIV-1 latency, it was assumed that after HIV-1 transmission, a small proportion of the virus infected cells transition into a long-term resting state and seed a stable, latent reservoir. It is expected that these events are rare, at least in CD4+ T cells where this has been studied the most intensively (reviewed in [10]). HIV-1 integration is dependent on T-cell activation [11, 12], but this also leads to active replication of the virus. So, only if a cell can evade cytopathic immune clearance and transition into a resting state successfully is this possible. Nonetheless, when it does occur, this has a critical impact on latency and HIV-1 cure research as resting memory T cells have a long lifespan. The quiescent cells maintaining the latent provirus decay very slowly with a half-life of about 44 months in adults [13, 14], which presents a great challenge for HIV-1 eradication.

The first suggestion of the presence of the latent reservoir in HIV-1 infected individuals was prior to the availability of ART [15]. However, the usage of ART allowed for the more definitive identification of replication-competent provirus from resting CD4+ T cells that could be reactivated [16, 17]. Moreover, the latent reservoir did not substantially decrease in size as the length of time on ART increased [18], which established its longevity. These studies also established foundational tools now known as the quantitative viral outgrowth assay (QVOA): a limiting dilution assay using CD4+ T cells from an ART-suppressed person living with HIV-1 with one round of T cell activation to determine the frequency of cells with replication-competent proviruses to define the latent reservoir [18]. Later studies observed that once there is clinical evidence of HIV-1 infection, even in the early stages, a latent reservoir is established [19-21]. Additionally, latent reservoirs are also established *in utero* or at birth [22-24].

1.2 Dynamics of the HIV-1 latent reservoir

Contrary to some initial assumptions, the long-lived latent reservoir is not static, and in fact, is actively maintained and shaped by various forces that expand and retract it. One major source of reservoir expansion is clonal expansion or clonal proliferation of cells that all contain the same integrated provirus. Multiple mechanisms likely collectively drive the process of expansion of cells harboring HIV-1 including homeostatic proliferation, antigen-driven proliferation, and integration of the HIV-1 genome at a site promoting cellular proliferation (**Figure 1.1**). Homeostatic proliferation is a process that induces the proliferation of memory T cells without cellular differentiation or activation and the process is maintained by cytokines such as interleukin-7 (IL-7). IL-7 has been demonstrated to promote the persistence of latently infected cells [25, 26]. Repeated exposure to microbial peptides can also drive proliferation in an antigen-specific manner. Examples of antigens include Cytomegalovirus (CMV), Epstein-Barr virus (EBV), and Influenza, which are viruses that can cause chronic or recurrent co-infections and enrichment of HIV-infected CD4+ T cells responsive to these antigens have been identified [27, 28]. Regardless of the mechanism, sequencing data supports the overall process of clonal proliferation as low residual levels of viremia are found to be dominated by identical viral sequences [29, 30].

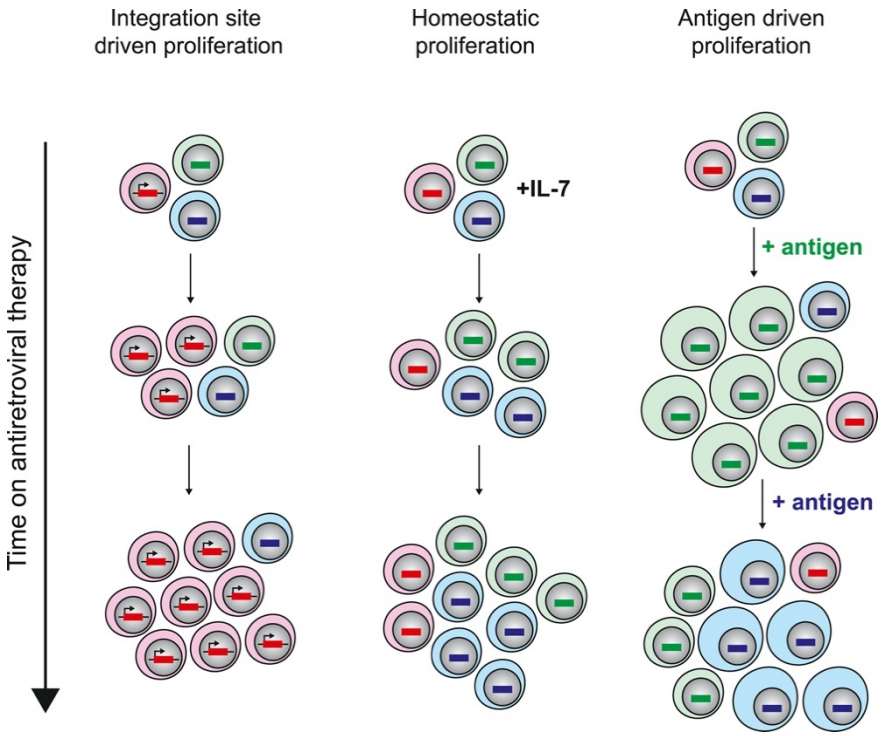


Figure 1.1. Multiple mechanisms underlie the expansion of cells harboring HIV-1 provirus.

A schematic representation of the three main mechanisms (antigen-driven proliferation, homeostatic proliferation, and integration site of the HIV-1 genome) suggested to drive the process of expansion of cells with an integrated HIV-1 provirus. **Figure 1.1** was reproduced from [1] with permission, Elsevier license number 5383861215084

One of the difficulties in identifying the source of latent proviruses is that a large portion of the persistent, latent reservoir for ART-suppressed people living with HIV-1 is composed of defective proviruses [31]. The process of reverse transcription can generate high levels of defective proviruses. While the replication-competent counterparts that upon reactivation can be eliminated through immune clearance and cytopathic effects, the defective proviruses produce reduced to no viral proteins, which result in lower likelihood of targeting by the immune system and thereby accumulate over time. Indeed, the rate of decay of cells harboring intact proviruses is significantly more rapid than the decay rate of cells with defective proviruses [32]. Defining an accurate method to measure the HIV-1 reservoir has been a continuing challenge as defective proviruses need to be excluded from this measurement. While QVOA has been used as a gold standard in the field, one limiting aspect is that the assay uses a single round of activation to estimate the reservoir size. However, not all proviruses are reactivated in one or even after a few rounds of activation and this can result in an underestimate of the reservoir size [33].

Development of a sequence-based method for estimating replication-competent latent proviruses called intact proviral DNA assay (IPDA) has provided another tool that can largely distinguish between defective and replication-competent proviruses [31]. Even after acute infection followed by early ART treatment, only 7% of the proviruses in CD4+ T cells of people living with HIV-1 were found to be intact [31]. Nonetheless, expanded clones of replication-competent HIV-1 have been detected after treatment for long periods of time and therefore maintain the stability of the latent reservoir [34].

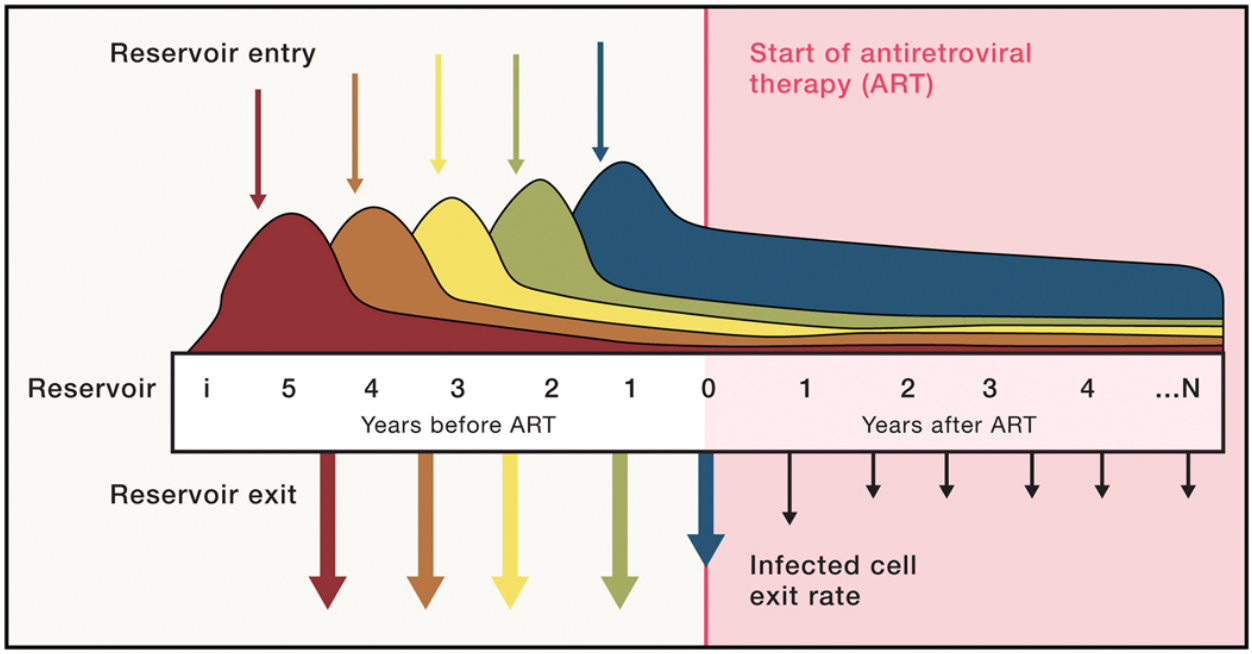


Figure 1.2. The latent reservoir is dynamic.

A model describing the latent reservoir as constant overall but dynamic in terms of the composition of the reservoir. Prior to ART initiation, new HIV-1 infected cells enter the reservoir as other HIV-1 infected cells exit because of reactivation. ART initiation alters this dynamic by stabilizing the latent reservoir, such that the reservoir is most similar to the composition just before ART initiation. **Figure 1.2** was reproduced from [2] with permission, Elsevier license number 5387440137799.

Advances in sequencing technologies have also allowed groups to further study the impact of ART initiation on the HIV-1 latent reservoir. Prior to ART initiation, one model suggests that the latent reservoir is very dynamic (**Figure 1.2**). While the overall size of the

latent reservoir stays relatively constant, there is a continuous turnover of HIV-1 infected cells going into latency and HIV-1 infected cells exiting latency upon reactivation, and thus the viral genomes in the reservoir are constantly decaying and reseeding [2, 35]. Several studies have also demonstrated that ART initiation itself likely changes this dynamic as the latent reservoir that stabilizes during ART is most similar to the reservoir in existence just before ART initiation [35-37]. In other words, in the presence of ART, there are no longer new viral genomes to enter the latent reservoir and simultaneously, there is a reduction in the exit of the reservoir as a byproduct of ART treatment. Initiation of ART results in viral suppression, which lowers viral loads and decreases antigen circulation and has other immunological effects. These processes can lead to decreases in CD4+ T cell turnover rate [38], increases in CD4+ T cells counts, and increases in CD127+ CD4+ memory T cell frequencies [39], which altogether may create an environment that further stabilizes the reservoir and allows for viral persistence. Thus, the latent reservoir is dynamic to conditions of cell proliferation, virus seeding, and other aspects of immune control and these factors need to be considered in any effective cure strategy for HIV-1.

1.3 Cell types of the HIV-1 latent reservoir

Although cells containing latent HIV-1 were first identified in resting CD4+ T cells from peripheral blood [15], the ultimate source of the HIV-1 latent reservoir in blood and tissues has been better characterized as memory CD4+ T cell subset populations. This includes the central memory T cells [40] and additional other subsets including naïve [41], stem-cell memory [42], central memory [43], transitional memory [44], and effector memory cells. As these cells differentiate, transition and turnover, defining the persistence of HIV-1 in one or another of these cell types becomes further complicated. Additionally, the proviral genome can remain quiescent

during these transition periods, which makes quantification further challenging. With the advent of additional phenotypic markers to define CD4+ T cell subpopulations, detailed analysis of multiple different CD4+ T cell subpopulations from ART-suppressed people living with HIV-1 demonstrate that while naïve CD4+ T cells harbor less integrated provirus compared to more mature populations [45], they have longer half-lives and slower turnover [46]. This suggests that as the cells mature, their turnover may increase, but are replaced by proliferation of less differentiated cells.

While many of the studies on the cell types that serve as sources of the latent reservoir focus on peripheral blood for its ease in accessibility, non-human primate studies, human biopsy studies, and human autopsy studies have been critical to better map the tissue reservoirs of HIV-1. Indeed, studies have demonstrated that gut mucosa, lymph node, and reproductive tracts are more likely the major reservoirs of latently infected cells. In particular, the gut is a large reservoir for HIV-1 with high frequencies of infection [47, 48] and especially CCR6+ memory CD4+ T cells that account for the majority of infected cells [49, 50]. The lymph nodes are also a large tissue reservoir [51]. One important, but understudied tissue source of the latent reservoir is the reproductive tracts. A study in penile tissues of ART-suppressed people living with HIV-1 identified inducible and replication-competent HIV proviruses in urethral macrophages [52]. HIV-1 DNA and RNA was also identified in cervical tissue from ART-suppressed people living with HIV-1 [53]. Another tissue reservoir of interest is the central nervous system (CNS), which is high in density of macrophages and microglial cells derived from resident macrophages. Although there are few studies on HIV-1 latency in the CNS, HIV-1 DNA has been detected in CNS tissue from ART-suppressed people living with HIV-1 [54]. Additionally, replication-competent virus from a quantitative viral outgrowth assay was detected in CNS brain tissue from

SIV-infected pig-tailed macaques that were ART-suppressed [55]. Ongoing research is in progress to perform a similar study in human brain tissues with a particular interest in microglia cells. Finally, a study examining HIV-1 across the whole human body was performed in an autopsy study of people living with HIV-1 and found that HIV-1 DNA was identified in blood cells and 28 tissues, but the main sources appeared to be from the blood and lymphoid tissues [56]. Different cell and tissue types will have different transcriptional programs, and thus, targeting of future HIV-1 cure strategies may require a combination of different approaches for each major reservoir source.

1.4 Models of HIV-1 latency

It has been challenging to study HIV-1 latency *in vivo* due to the low cell count of latently infected cells and the difficulty in distinguishing these cells given the absence of any obvious phenotypic markers. As a result, although each is imperfect, different *in vitro* models have been established to, at least, partly, mimic different aspects of HIV-1 latency. Importantly, no single model completely recapitulates the biological properties of HIV-1 latency as it occurs *in vivo*, but each model has its strengths and weaknesses. Cell line models are useful in that they are genetically tractable and are a powerful tool to use to conduct genetic screens. However, each cell line model is also flawed in that they are often derived from leukemic cells with infinite life spans and altered transcriptional and activation profiles. Various latently infected T-cell lines [57-60] have been generated including the different clonal Jurkat cell lines called J-Lat cells [61]. One particular clone of J-Lat cells, the J-Lat 5A8 cells [62], clustered closely to primary models in an extensive comparison of reactivation profiles of multiple HIV-1 latency model systems [63]. However, as these cell lines are clonal in their virus integration sites, unlike the

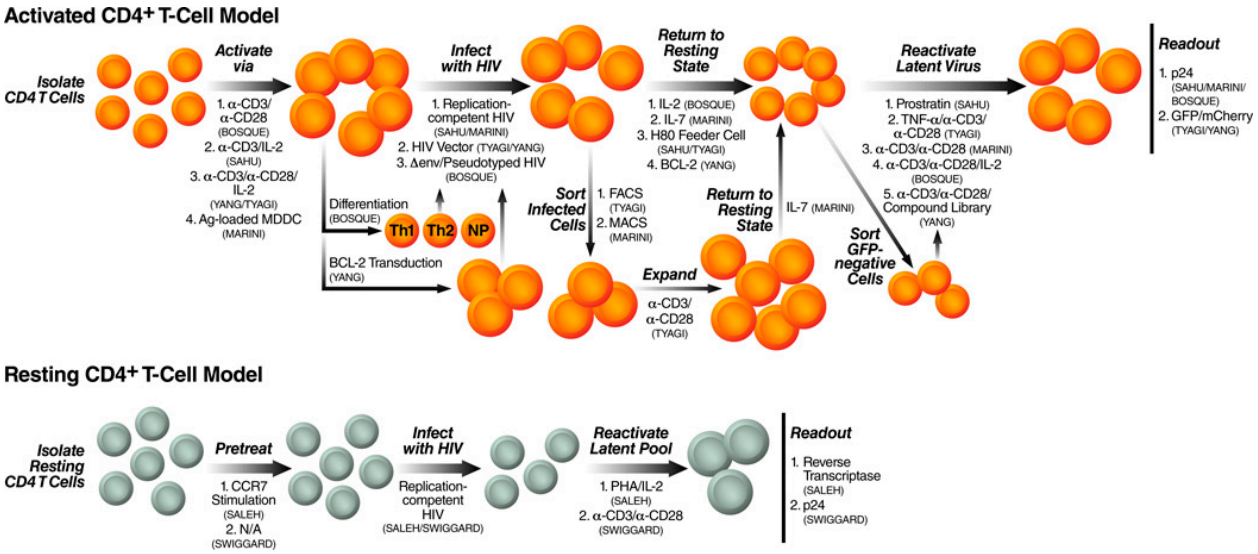


Figure 1.3. Primary CD4+ T cell models of HIV-1 latency.

This diagram outlines the different models in which HIV-1 latency can be established in primary human CD4+ T cells. The main two groups of models either involve infection of activated CD4+ T cells or infection of resting CD4+ T cells. The types of activation, treatment, expansion, reactivation, or readout for each specific model is described. The first author of the publication associated with each model is included. **Figure 1.3** was reproduced from [6] with permission, Elsevier license number 5387441313778.

polyclonal nature of HIV-1 latency *in vivo*, latency studies that use multiple different models are preferable.

Primary cell models of HIV-1 latency have also been established to address some of the concerns with cell line models. There are two major groups of primary cellular models of HIV-1 latency – one category of models involves infection of activated CD4+ T cells and the other category involves infection of resting CD4+ T cells (**Figure 1.3**). The models based on infecting activated CD4+ T cells are an attempt to emulate how HIV-1 latency after infection might be established *in vivo* and allows the cells to undergo the transition into a resting state to establish HIV-1 latency [64-70]. One challenge of this model is acquiring enough cells to use for further experiments since the life span of the activated CD4+ T cells is not long without environmental stimuli and the number of cells harboring latent HIV-1 is often small. To overcome these

challenges, these models have adjusted conditions such as culturing the CD4+ T cells in the presence of cytokines such as IL-2 or IL-7 or other strategies without cytokines in order to generate enough relevant cells to study HIV-1 latency.

One primary CD4+ T cell HIV-1 latency model was established by Sahu et al [64] which involves using CD4+ T cells isolated from healthy donors, T-cell activated by anti-CD3 antibody, infected with a replication-competent virus, and then co-cultured with a brain tumor-derived cell line, H80 cells, to allow the cells to more efficiently transition and be maintained in a resting state. The mechanism of the feeder cell line is still unclear, but it allows for the production of primarily central memory CD4+ T cells in the absence of ectopic cytokines. It is worth noting that more than half of the CD4+ T cells in this model constitutively express the early activation marker, CD69, which suggests that not all of the cells are in a completely resting state. Another model using the H80 feeder cells was developed by Tyagi et al [65], which uses a reporter HIV-1 vector for infection and includes an enrichment step via FACS for infected cells followed by an expansion step using an anti-CD3 and anti-CD28 antibody and IL-2 cocktail. Similar to the previous model, a subset of these resting population cells continues to express the reporter, suggesting some continued viral expression. A further extension and refinement of this model (and a variant of which is used for the studies in this thesis) is to use HIV-1 reporter viruses that discriminate between the latent and actively replicating viral populations [71, 72].

Another significant model development allowed for the study of central and effector memory CD4+ T cell populations [67]. CD4+ T cells are activated and primed and cultured in conditions to produce polarized and non-polarized CD4+ T cells. These cells are then infected with a pseudotyped HIV-1 virus and subsequently cultured to transition back to a resting state, although in the study, no T cell activation marks, such as CD69, were assessed to determine the

final state. Additionally, a new model, called QUECEL has been developed by Dobrowolski et al to consistently generate four major T cell subsets (Th1, Th2, Th17, and Treg) by defining cocktails of cytokines that can be used to polarize naïve CD4+ T cells and subsequently infected with a reporter HIV-1 virus to establish latency [69]. Establishment of these major T cell subsets aims to generate viral reactivation profiles with better correlation to viral induction in cells from ART-suppressed people living with HIV-1. A strategy to generate effector memory cells without the use of cytokines or feeder cells is by using a system which involves transducing the activated CD4+ T cells with a lentiviral vector carrying the Bcl-2 gene, a downstream target of IL-7 [68]. Expression of Bcl-2 allows for the CD4+ T cells to become long-lived as it serves as an anti-apoptotic protein. After generating the resting cells, they are transduced with an GFP reporter HIV-1 vector with premature stop codons in most HIV-1 genes in order to be minimally cytopathic. Enrichment steps can be taken afterwards to select for cells that are latent based on the reporter and activation markers, CD69 and CD25.

Each of the models discussed above use infection of activated cells that are then transitioned to a resting (e.g. memory-like) state. However, it is not yet clear if HIV-1 latency *in vivo* is achieved from the transition of infected, proliferating cells to a resting state or rather from direct infection of resting cells. Thus, HIV-1 latency models have also been developed that directly infect resting CD4+ T cells in the absence of any stimulation [73]. Because expression of CCR7, which is a marker for defining memory T cell subsets [74], is correlated with a greater efficiency of infection of resting CD4+ T cells, additional strategies to boost CCR7 expression have also been developed by isolating resting CD4+ T cells and stimulating with the CCR7 ligands, CCL19 and CCL21, followed by infection of replication-competent HIV-1 to generate the latent model [75]. Nonetheless, these strategies are especially challenging due to a lack of

dNTPs leading to inefficient reverse transcription and unsuccessful integration of HIV-1 in resting CD4+ T cells [76].

While the models described thus far involve an *in vitro* HIV-1 infection, additional models bridge *in vivo* and *ex vivo* by isolating memory CD4+ T cells from ART-suppressed people living with HIV-1 followed by *in vitro* expansion in the presence of ART and reestablishment of the resting state [70]. This model has a higher level of variability as the cells are derived from people living with HIV-1 with varying endogenous latent reservoirs. Overall, this highlights that there are notable differences between each model and in a comprehensive examination of many of these models by HIV-1 reactivation using a panel of agents, no single model could perfectly recapitulate the *ex vivo* responses of T cells from a person living with HIV-1 [63].

More complex *in vivo* models also exist to study HIV-1 latency including animal models of non-human primates (NHPs) and humanized bone marrow/liver/thymus (BLT) mice. Both models can provide critical information on the mechanistic understanding of HIV-1 latency and also the safety of the strategies for humans. NHPs can be infected with Simian Immunodeficiency Virus (SIV) where an initial high viral load is achieved as well as the rapid establishment of a long-lived reservoir. NHPs are responsive to some forms of ART and important studies have been performed in NHPs to better understand the timing of the initial seeding of the viral reservoir and how the timing of ART initiation has an impact [77-79]. While mouse models are refractory to HIV-1 infection, this challenge has been overcome by the establishment of humanized BLT mice. Nonobese diabetic/severe combined immune deficiency (NOD/SCID) mice are transplanted with CD34+ cells to repopulate the mouse with human cells and implanted with human fetal thymic and liver tissues to develop functional T cells [80].

Similar to the NHP models, humanized BLT mice can be infected by HIV-1, are responsive to ART, and establish a latent reservoir [81]. While neither model is the perfect representation of a human HIV-1 infection, the two models complement each other to contribute to HIV-1 Cure research. This has been successfully observed in recent studies evaluating two different approaches to activate latent HIV [82, 83].

There are a great variety of models that have been established to study HIV-1 latency from *in vitro* cell lines to primary cell models to animal models. While no single model is a perfect representation of HIV-1 latency *in vivo*, each model is well-equipped to address a different set of questions regarding HIV-1 latency. As HIV-1 latency *in vivo* is a complex system, the benefit of this plethora of models is that many different aspects of HIV-1 latency can be examined piecewise to contribute to our overall understanding of the mechanisms of HIV-1 latency.

1.5 Transcriptional control of the HIV-1 provirus

HIV-1 latency is largely a transcriptionally silent state of the provirus that is controlled at multiple steps including the site of integration, transcription initiation, and transcription elongation. The following mechanisms described are mostly based on studies performed in T cells, so it is important to note that the mechanisms controlling HIV-1 transcription in other cell types, such as microglial cells, may differ or have differing importance. HIV-1 preferentially integrates into the host genome into regions of active transcription [84, 85] and this has been demonstrated in cell lines and in CD4+ T cells of people living with HIV-1 [86, 87]. This integration step is influenced by nuclear topography with preference for integration at open chromatin that is proximal to the nuclear pore [88]. However, this presents a paradox as these

proviruses, despite being integrated in active genes, have the potential for contributing to the latent reservoir and remaining transcriptionally silent. One proposed theory to explain this phenomenon is transcriptional interference, which in this case refers to a host gene negatively impacting the transcription of the HIV-1 promoter in *cis*. In Jurkat models of latency, chromatin reassembly factors have been identified to be recruited after RNA Polymerase II processing resulting in repression of the HIV-1 promoter [89]. On the other hand, HIV-1 integration into chromosomal sites less permissive to active viral transcription also occurs and it is possible that these rare integration sites play an outsized role in the establishment of the latent reservoir [90]. Furthermore, the integration site itself may facilitate the survival and/or proliferation of the latently infected cell [91].

Regardless of the site of integration, a major contributing element to HIV-1 transcriptional control are the host cellular factors that silence or activate the HIV-1 LTR. The HIV-1 promoter is divided into three regions: the U3, R, and U5 (**Figure 1.4B**) and can also be functionally divided into four regions: the negative regulatory element (NRE), the enhancer region, the core promoter, and the 5' Untranslated Region (5' UTR) [92]. Many of these host factors act by binding to regions of the HIV-1 LTR to influence transcription initiation at the HIV-1 promoter (**Figure 1.4B**). The NRE contains many transcription factor binding sites such as activator protein-1 (AP-1), which regulate the transcription of the HIV-1 provirus. The enhancer region induces high levels of viral transcription and contains binding sites for transcription factors such as nuclear factor-κB (NF-κB) and nuclear factor of activated T-cells (NF-AT). The core promoter contains the TATA box and Sp1 binding sites that help to position RNA Polymerase II to initiate transcription. The 5' UTR contains additional transcription binding sites such as Yin Yang 1 (YY1) and importantly, encodes the transactivation response

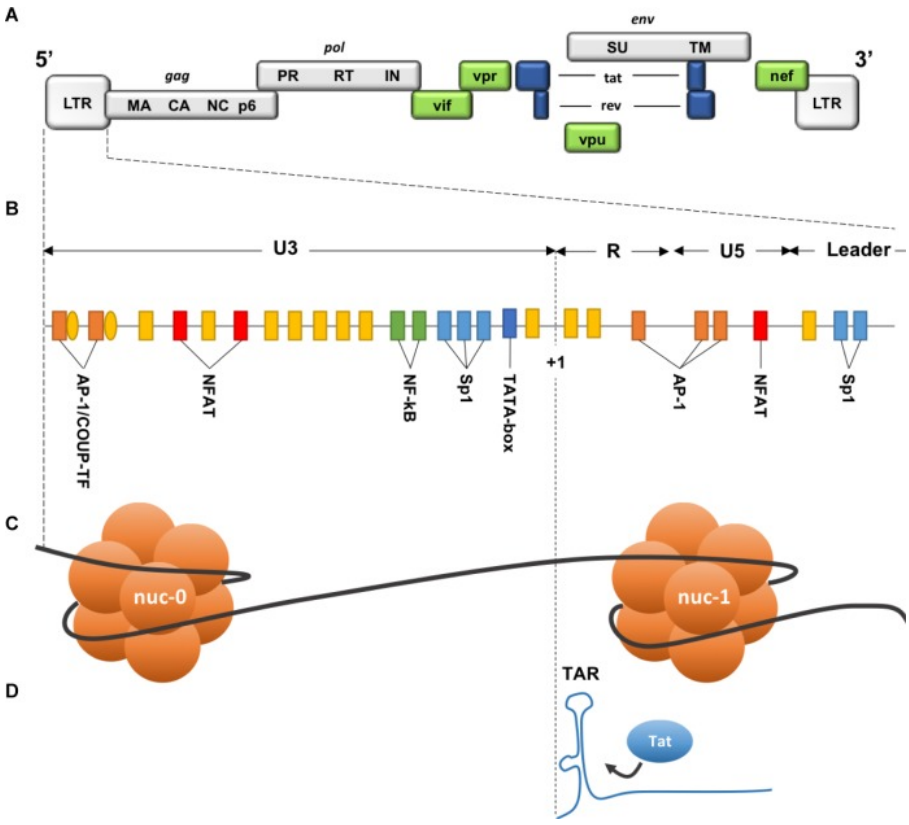


Figure 1.4. HIV-1 LTR encodes many sites for host cellular factor regulation.

This schematic describes the HIV-1 LTR. **(A)** HIV-1 genome. **(B)** Breakdown of the HIV-1 LTR into three regions including U3, R, and U5 and the transcription factor binding sites encompassed by each region. **(C)** The sites of the two nucleosomes (nuc-0 and nuc-1) situated at the promoter region of the HIV-1. **(D)** The location of the transactivation response (TAR) region to which viral protein transactivator (Tat) binds for efficient transcription elongation. **Figure 1.4** was reproduced from [5] with permission, Creative Commons Attribution License 4.0 International (CC BY 4.0).

(TAR) region, an RNA stem-loop structure, which is the binding site of the viral protein transactivator (Tat) and critical for transcription elongation (**Figure 1.4B, D**). Altogether, the transcription factors introduced are also coupled to the activation status of the cell and serve a role in the molecular mechanism of latency. Specifically, NF-κB and NF-AT are effectors tightly linked to cell activation and also play critical roles in HIV-1 transcription. Both transcription factors are sequestered in the cytoplasm in resting cells and cellular activation by T-cell receptor (TCR) engagement or tumor necrosis factor (TNF)-α signaling initiates a signaling cascade

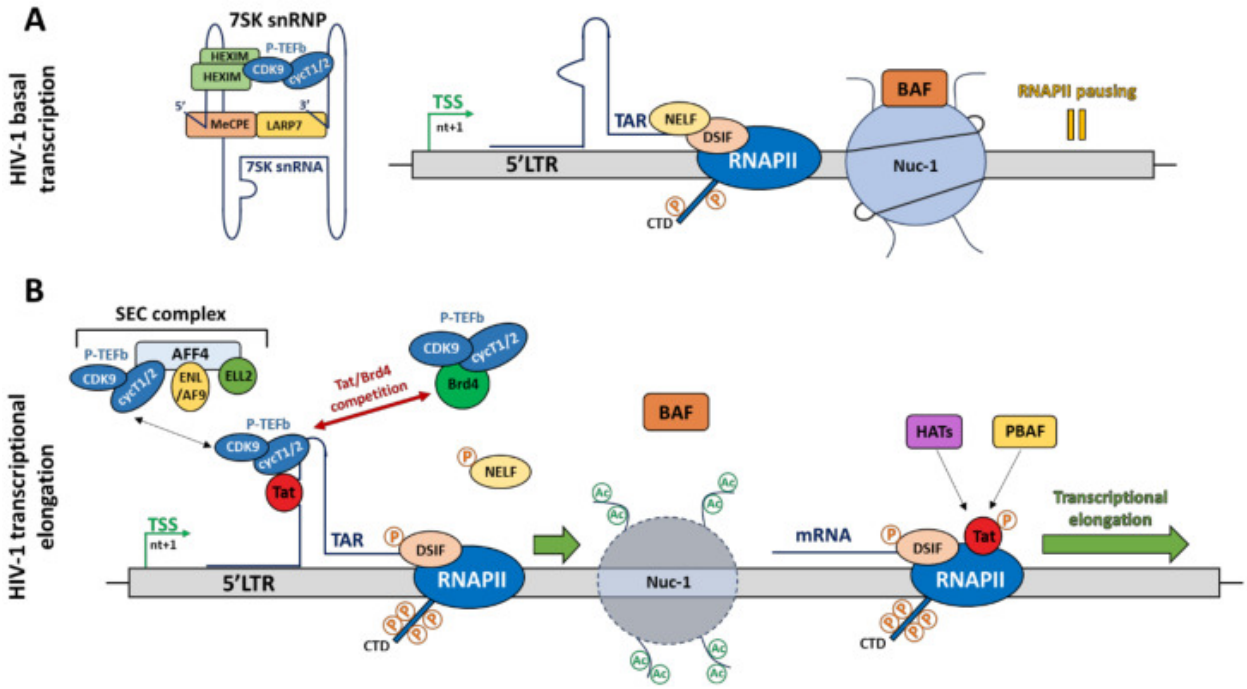


Figure 1.5. Regulation of HIV-1 transcriptional mechanisms.

Schematic of the molecular mechanisms involved in HIV-1 transcription. **(A)** After transcription initiation, RNA Polymerase II (RNA Pol II) pauses due to multiple layers of regulation including recruitment of two host pausing factors named negative elongation factor (NELF) and DRB sensitivity-inducing factor (DSIF), positioning of repressive nuc-1 just downstream of the HIV-1 transcription start site by BAF, and sequestration of P-TEFb in the inactive form of the 7SK snRNA complex. **(B)** The accumulation of viral protein Tat initiates efficient transcription elongation with the release of the pausing factors, recruitment of active P-TEFb to phosphorylate RNA Pol II, and remodeling of nuc-1 to upstream of the HIV-1 transcription start site. Additional epigenetic regulation occurs on nuc-1 and Tat. **Figure 1.5** was reproduced from [7] with permission, Creative Commons CC BY license.

leading to the translocation of the both NF-κB and NF-AT into the nucleus where they mediate the activation HIV-1 transcription [93].

A distinguishing feature of HIV-1 transcription compared to cellular transcription is the regulation by the viral protein Tat. Tat is a part of the positive regulatory feedback circuit and together with the co-factor positive transcription elongation factor b (P-TEFb), a host protein complex, are essential players for efficient transcription elongation of the HIV-1 provirus. In the absence of Tat, transcription initiation occurs normally, but RNA Pol II elongation activity is mostly inhibited by negative factors including negative elongation factor (NELF) and DRB

sensitivity inducing factor (DSIF) that bind 3' to the transcription initiation site [94] resulting in promoter proximal pausing of RNA Pol II (**Figure 1.5A**). As a result, the majority of the time, only short abortive transcripts are generated, but occasional full-length Tat-encoding transcripts are produced, which ultimately allows for the accumulation of Tat and transition into the Tat-mediated phase of HIV-1 transcription [95].

In the presence of Tat, the host cellular P-TEFb, which is composed of a Cyclin T1 (CycT1) subunit and a catalytic CDK9 subunit, is directed to the paused RNA Pol II at the HIV-1 promoter to activate transcription elongation (**Figure 1.5B**). CycT1 of P-TEFb interacts with Tat, which subsequently binds to the 5'UTR of the HIV-1 promoter [96]. CDK9 undergoes conformational changes induced by Tat [97] such that CDK9 phosphorylates NELF to relieve the transcription elongation block [98] and CDK9 also phosphorylates RNA Pol II [99] and Spt5 [100, 101] to activate the polymerase activity. This establishes the positive feedback loop resulting in amplified, efficient HIV-1 transcription and virion production. The vital role of Tat in active transcription also implies its importance in latency. Reducing functional Tat availability has been demonstrated to affect latency. In a Jurkat model of latency, the frequency of provirus silencing could be dramatically increased by introducing a mutation to attenuate Tat [102]. Additionally, in CD4+ T cells from people living with HIV-1, an increase in Tat variants has been observed in the latently-infected cells after ART initiation compared to the mostly productively-infected cells prior to ART initiation [103].

As the HIV-1 LTR is largely silent in latently infected cells, the pathway described above to activate transcription initiation and/or elongation must be absent or inhibited in HIV-1 latency. This idea is the basis for an HIV-1 Cure strategy known as "shock and kill" (alternatively known as "kick and kill") in which the latent proviruses are induced to reactivate transcription and

followed by a step to eliminate those cells after viral proteins are produced [1]. The "shock" aspect of this strategy largely focuses on pathways by which the transcription initiation and elongation factors are induced by small molecules known as latency reversal agents (LRAs). Many LRAs have been identified to target the NF-κB pathway as it is a critical inducer of HIV-1 transcription. These LRAs are generally protein kinase C (PKC) pathway activators that ultimately activate the NF-κB pathway and include bryostatin and prostratin, but potentially also induce a resistance to apoptosis [104]. However, as the NF-κB pathway is also involved in many diverse functions, a concern is that activating this pathway could result in unwanted host consequences. As an alternative for a more selective and longer lasting NF-κB activation, recent studies have targeted the non-canonical NF-κB pathway using second mitochondrial-derived activator of caspases (SMAC) mimetics that degrade inhibitor of apoptosis (IAP) proteins and relieve an inhibitor of the non-canonical NF-κB pathway. This approach identified LRAs with a greater specificity for the HIV-1 LTR over other host promoters [83, 105].

Another target for LRAs is the activation of P-TEFb to activate HIV-1 transcription through the transcription elongation process, which is facilitated by release from the inhibitory complex called 7SK small nuclear RNA (7SK snRNP). One example is usage of the small molecule hexamethylene bisacetamide (HMBA), which induces the phosphorylation of HEXIM1 of the 7SK snRNP and leads to the release of P-TEFb [106]. An additional avenue to activate P-TEFb for HIV-1 transcription is by removing the P-TEFb binding competitor of Tat, BRD4, a bromodomain containing protein [107] (**Figure 1.5B**). Bromodomain inhibitors such as JQ1 have been demonstrated to induce viral reactivation, likely by removing BRD4 and increasing P-TEFb levels for Tat binding, but the specifics of the mechanism remain under investigation [108-110].

Despite the wealth of LRAs that have been identified, activation with each individual LRA is not sufficient for activation of the complete, relevant latent reservoir. Resting CD4+ T cells from people living with HIV-1 underwent maximum *in vitro* T cell activation and two rounds of stimulation and still did not reactivate all inducible proviruses [33]. Thus, alternative approaches such as using combinations of LRAs that target different transcriptional and epigenetic regulators have been proposed and tested. Combination LRA studies, such as combining AZD5582, a transcription initiation activator, and I-BET151, a transcription elongation activator, have demonstrated that there is synergy between the two LRAs in Jurkat models of latency, but it is not adequate for induction in CD4+ T cells of people living with HIV-1 [111]. Further research on the transcriptional regulation of HIV-1 latency is warranted to dictate the direction of future combination LRA approaches for inducing the latent reservoir.

1.6 Epigenetics and HIV-1 latency

The transcription initiation and elongation state of the HIV-1 LTR is also highly influenced by the chromatin landscape at and around the HIV-1 provirus. In addition to chromatin-binding factors, chromatin accessibility is determined by the organization of DNA in nucleosomes. Each nucleosome is composed of 146 base pairs of DNA wrapped around an octamer of two copies of each of the four core histones – H2A, H2B, H3, and H4 [112]. Each histone has an unstructured tail that can be modified by post-translational modifications, such as histone acetylation and histone methylation, which reveal a histone code [113] and can affect chromatin condensation to change accessibility of the DNA and recruitment of regulatory proteins. Two nucleosomes, called nucleosome-0 (nuc-0) and nucleosome-1 (nuc-1), are situated at the promoter region of HIV-1 with nuc-1 positioned just downstream of the HIV-1 promoter

transcription start site and a critical block to transcription elongation [114] (**Figure 1.4C**). Upon activation of HIV-1 transcription, chromatin remodeling to nuc-1 occurs resulting in changes in DNA accessibility [115] (**Figure 1.5B**).

Remodeling of the chromatin can be facilitated by multiple mechanisms including post-translational covalent modification of histone tails and DNA methylation [116]. Histone modifications have been demonstrated to have a functional role in HIV-1 latency (**Figure 1.5B**). For example, a global analysis of histone modifications in wildtype CD4+ T cells demonstrated that a broad range of lysine acetylation marks are associated with transcriptionally active euchromatin while a few lysine methylation marks are associated with transcriptionally repressed chromatin [117]. Histone acetylation marks at the HIV-1 LTR are most often associated with active transcription and histone acetyltransferases (HAT) are recruited to the HIV-1 LTR upon activation. HAT complexes such as p300/CBP and P/CAF are associated with stimulation of HIV-1 transcription [118, 119]. p300/CBP are coactivators of NF-κB, which suggests that binding of NF-κB to the HIV-1 LTR also facilitates chromatin remodeling of the HIV-1 provirus [119, 120]. Additionally, Tat activity is also regulated by acetylation as acetylation of Tat by p300/CBP is critical for Tat transactivation [121-123]. Further evidence to demonstrate the importance of acetylation in HIV-1 transcription is the identification of histone deacetylases (HDACs), proteins that facilitate the removal of histone acetylation marks, as transcriptional repressors. HDAC proteins are divided into four classes and Class I HDACs are primarily associated with deacetylation of the histones at the LTR [124]. One example of Class I HDACs is histone deacetylase 1 (HDAC1), which has been demonstrated to inhibit HIV-1 transcription and is recruited to the HIV-1 LTR by the cellular host factors YY1 and LSF [125, 126]. Furthermore, inhibitors of HDACs were identified as candidate LRAs such as Vorinostat

(SAHA) and Romidepsin [127]. These HDAC inhibitors prevent the removal of acetylation marks, enhance latency reversal, and continue to be tested in the clinic [128, 129]. Despite the importance of acetylation, it cannot be the sole or major mark that regulates HIV-1 latency *in vivo* since HDAC inhibitors tested clinically (such as SAHA) did not result in significant reductions in the frequency of replication-competent HIV-1 in resting CD4+ T cells in human trials [128].

Another prominent histone mark that has been associated with HIV-1 transcription is H3 methylation in the context of transcriptional repression. One mark is H3K27 methylation, which is usually associated with constitutive heterochromatin and the other mark is H3K9 methylation, which is usually associated with facultative heterochromatin [130]. The Polycomb Repressive Complex 2 (PRC2) mediates H3K27 methylation and is composed of the core proteins of SUZ12, the chromatin reader EED, and the catalytic methyltransferase EZH2. EZH2 and H3K27me3 have been demonstrated as important for maintaining HIV-1 latency in Jurkat-based and primary T cell models of latency [131-133]. In fact, treatment of latently infected cells with the EZH2 methyltransferase inhibitors GSK-343 and EPZ-6438 have resulted in viral reactivation in both Jurkat-based and primary T cell models of latency [131, 133]. Additionally, PRC2 interacts with the Polycomb Repressive 1 (PRC1) complex, but the recruitment mechanisms remain complicated [134]. Regardless, PRC1 complex members have also been identified at the HIV-1 LTR during latency [133].

H3K9 methylation is a marker for repression and can be mediated by the histone methyltransferase, G9a or EHMT2. G9a/EHMT2 inhibitors (BIX01294 and UNC-0638) have been used to show that H3K9 methylation is important to HIV-1 latency maintenance [131, 135]. Moreover, UNC-0638 treatment of cells from ART-suppressed people living with HIV-1 has

been shown to modestly stimulate viral reactivation on its own and more effectively when in synergy with another LRA such as SAHA [131]. Further evidence for the role of H3K9 methylation in HIV-1 latency is the identification of the Human Silencing Hub (HUSH) as a potential latency maintenance factor. HUSH is composed of the H3K9me3 reader MPP8, TASOR, and periphilin and the knockdown of TASOR and MPP8 in a Jurkat based model of latency resulted in a decrease of H3K9me3 along with reactivation of the HIV-1 based viral vector [136]. A similar result confirmed that knockout of each of the HUSH subunits resulted in viral reactivation in most, but not all J-Lat cell lines tested [136]. Further evidence for a role of the HUSH complex in primate lentiviruses is the fact that it is targeted for degradation by the SIV and HIV-2 Vpx protein [137]. As HIV-1 does not encode a Vpx protein and does not degrade HUSH, there are potential important differences in latency control between SIVs in non-human primates and HIV-1 in humans. The role of HUSH in the context of HIV-1 transcriptional silencing requires continued investigation.

In addition to post-translational modifications, factors that facilitate nucleosome remodeling and histone chaperones also serve a role in HIV-1 latency maintenance (**Figure 1.5A, B**). For example, SLTM is a chromatin regulator that increases chromatin accessibility of the HIV-1 provirus, including at the LTR in a Jurkat model of latency [138]. Another example is the BRG-1 associated factor (BAF) and Polybromo-associated BAF (PBAF) mammalian SWI/SNF chromatin remodeling complexes. BAF has been demonstrated to position the repressive nuc-1 just downstream of the transcription start site and upon HIV-1 transcription activation, BAF is removed from the LTR and nuc-1 is repositioned [139] (**Figure 1.5A, B**). PBAF binds to acetylated Tat and is involved in optimal Tat activity, especially in facilitating transcription elongation [140]. In a study of chromatin reassembly factors including Facilitates

Chromatin Transcription (FACT) complex, CHD1, and the histone chaperones ASF1a and HIRA, knockdown of each of these factors promoted latency reversal in a Jurkat model of HIV-1 latency [89]. These FACT complex proteins are usually associated with active mammalian host transcription and their contradictory role in repressing HIV-1 transcription suggests when HIV-1 provirus integrates into the intron of an actively transcribing gene, transcriptional interference from the active host promoter may suppress HIV-1 transcription, and reversal from latency occurs when the neighboring host gene is silenced [89].

1.7 Dissertation overview

The ultimate goal for HIV-1 Cure studies is for people living with HIV-1 to stop ART altogether without experiencing any viral rebound. HIV-1 has the unique biological characteristic of integration, which results in the enigmatic phase of latency. The recent developments on models to study HIV-1 latency has provided a greater understanding of the dynamics of HIV-1 latency including the establishment of the reservoir in the presence of ART, the cell types in which HIV-1 reservoirs are found, and the complex interaction of epigenetic and transcriptional factors that regulate HIV-1 transcription. My graduate thesis focused on the development of a novel CRISPR-based screening strategy to begin to build a comprehensive understanding of the molecular mechanisms of HIV-1 latency in order to discover targets that would increase the potency and specificity of latency reversal. The strategy is focused on exploring the interplay of epigenetic regulators with other transcriptional mechanisms through the usage of LRAs. The hope is for this study to contribute to building a better strategy to control HIV-1 viral rebound effectively and efficiently.

Chapter 2. A modular CRISPR screen identifies individual and combination pathways contributing to HIV-1 latency

This manuscript is currently submitted to bioRxiv (https://doi.org/10.1101/2022.08.23.504195).

2.1 Abstract

Transcriptional silencing of latent HIV-1 proviruses entails complex and overlapping mechanisms and are a major barrier to *in vivo* elimination of HIV-1. We developed a new latency CRISPR screening strategy, called Latency HIV-CRISPR, which uses the packaging of guideRNA-encoding lentiviral vector genomes into the supernatant of budding virions as a direct readout of factors involved in the maintenance of HIV-1 latency. We developed a custom guideRNA library targeting epigenetic regulatory genes and paired the screen with and without a latency reversal agent – AZD5582, an activator of the non-canonical NFκB pathway – to examine a combination of mechanisms controlling HIV-1 latency. A component of the Nucleosome Acetyltransferase of H4 histone acetylation (NuA4 HAT) complex, ING3, acts in concert with AZD5582 to activate proviruses in J-Lat cell lines and in a primary CD4+ T cell model of HIV-1 latency. We found that the knockout of ING3 reduces acetylation of the H4 histone tail and BRD4 occupancy on the HIV-1 LTR, and the combination of ING3 knockout with the activation of non-canonical NFκB via AZD5582 act together to dramatically increase initiation and elongation of RNA Polymerase II on the HIV-1 provirus in a manner that is nearly unique among all cellular promoters.

2.2 Introduction

Effective antiretroviral therapy (ART) can drive HIV-1 viral loads to undetectable levels [141]. However, ART does not eliminate the virus from people living with HIV-1. Upon

interruption of ART, there is a rapid rebound of virus replication from a long-lived latent reservoir primarily found in memory CD4+ T cells among other cell types [142]. This latent reservoir of replication-competent HIV-1 proviruses is a significant obstacle to the complete clearance of HIV-1 [14], and purging or managing this latent reservoir is essential to achieve a functional cure.

The latent HIV-1 reservoir is often transcriptionally silent but can produce infectious virus upon T-cell activation [143]. The transcriptional silencing of the provirus in HIV-1 latency is dependent on integration site [90, 144] and a breadth of host factors that limit transcriptional initiation and elongation of the HIV-1 proviral genome [145]. Small molecule inhibitors or activators, known as latency reversal agents (LRAs), have been shown to activate HIV-1 viral transcription in cell line models of HIV-1 latency [146] and in primary cell cultures derived from people living with HIV-1 [131, 147]. AZD5582 is a second mitochondria-derived activator of caspases (SMAC) mimetic that is an example of an LRA that provides some specificity for latency reversal by activating the non-canonical NFκB pathway [83]. Many LRAs have shown success at reactivating latent HIV-1 in HIV-1 latency models, but even the most well-established LRAs have modest or no effect in clinical trials on their own [2]. This highlights the complexity of mechanisms involved in maintaining the HIV-1 latent state, and the need to identify novel epigenetic factors and transcriptional mechanisms controlling HIV-1 latency.

Studies on the functional role of host epigenetic regulation of HIV-1 transcription (reviewed in [127, 145, 148]) has resulted in the identification of some specific host epigenetic regulatory genes that are involved in maintaining HIV-1 latency. For example, KAT5 of the Nucleosome Acetyltransferase of H4 histone acetylation (NuA4 HAT) complex deposits a uniquely high profile of acetylated lysine residues on the H4 histone tail (H4Ac) found at the

HIV-1 provirus LTR [147]. This results in the recruitment of the long isoform of the *Bromodomain Containing 4* (*BRD4*) gene [107, 109, 147]. Additionally, the long BRD4 isoform serves as a competitor for positive transcription elongation factor b (P-TEFb) binding to the HIV-1 transactivator, Tat [107]. *BRD4* also encodes a short isoform that recruits the repressive BAF complex to the HIV-1 provirus [149]. Another set of genes implicated in HIV-1 latency encode components of the Polycomb Repressive Complex 2 (PRC2) which catalyzes the methylation of histone H3 Lysine 27 to maintain transcriptional repression of genes, including throughout development [131-133].

The use of genetic screens has been effective in revealing pathways involved in HIV-1 latency in a single-gene manner [105, 131, 138, 150-155]. Here, we sought to create a novel CRISPR-based screening platform with the capacity to allow for rapid, parallel assessment of multiple pathways that contribute to the maintenance of HIV-1 latency both with and without the presence of additional LRAs. We reasoned that by combining LRAs of one mechanism with a CRISPR screen that targets a different mechanism, we could uncover latency reversal pathways with increased potency and specificity for the HIV-1 LTR. We established a novel HIV-based latency CRISPR screening strategy that uses J-Lat cells [61, 62] in combination with our recently developed HIV-CRISPR screening methodology [156] to use reactivation of the latent provirus as the reporter for the screen. To identify epigenetic pathways that act with and without LRAs to activate latent proviruses, we generated a custom CRISPR library targeting epigenetic regulatory genes and performed the screen in the absence and presence of a low dose of AZD5582. In the absence of AZD5582, we found *CUL3*, a scaffold for an E3 ubiquitin ligase, as a novel HIV-1 latency maintenance and establishment factor in both J-Lat cells and a primary CD4+ T cell model of HIV-1 latency. Moreover, we identified the NuA4 HAT complex to be of importance

to HIV-1 latency, specifically two complex subunit members – ACTL6A and ING3. In particular, *Inhibitor of Growth Family Member (ING3)* knockout combined with AZD5582 treatment resulted in enhanced viral reactivation in the J-Lat cells and a primary CD4+ T cell model of HIV-1 latency. By using automated CUT&Tag to investigate genome-wide chromatin occupancy [157], we found that *ING3* knockout alone and in combination with AZD5582 treatment causes H4Ac and BRD4 levels to be reduced at the HIV-1 LTR. However, only in the presence of AZD5582 in the *ING3* knockout cells, do we see a substantial increase in RNA Polymerase II Serine 5 phosphorylation (RNA-Pol2-S5p) occupancy at the LTR and this dramatic increase is nearly unique among all promoters in the human genome. Simultaneously, the combination of *ING3* knockout and AZD5582 treatment results in increases in RNA-Pol2-S5p and RNA-Pol2-S2p within the body of the provirus, which indicates increased transcription initiation and elongation and correlates with the increases of viral reactivation. Our novel HIV-CRISPR screening approach provides an avenue to explore factors that act in combination to promote HIV-1 latency.

2.3 Results

2.3.1 Latency HIV-CRISPR identifies novel epigenetic factors involved in maintaining HIV-1 latency

Our aim was to establish a high throughput CRISPR-Cas9 knockout strategy using the biology of HIV-1 packaging and budding to facilitate a screen for investigating the complexities of epigenetic regulation of HIV-1 latency. This approach uses a previously described HIV-CRISPR vector [156] which is a lentiviral based vector with two intact LTRs, a packaging signal, Cas9, and a library of single guide RNAs (sgRNAs). The unique feature of the HIV-

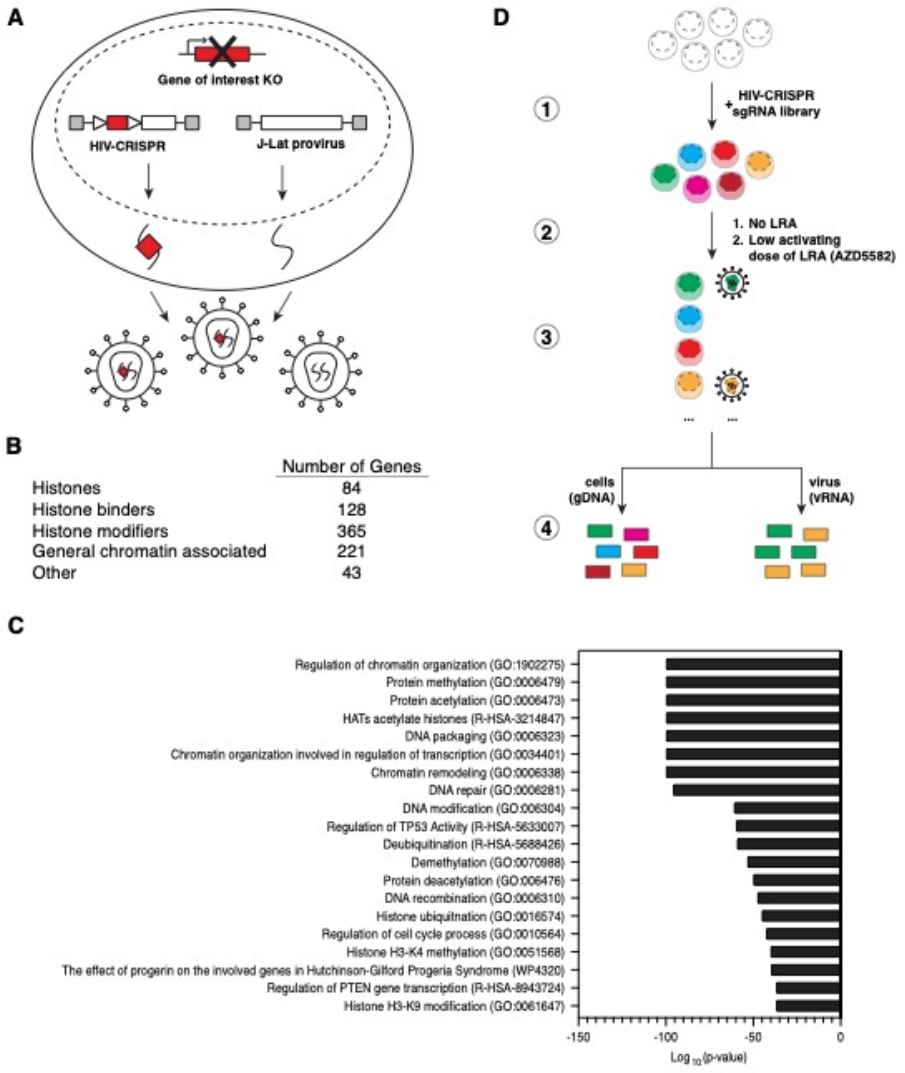


Figure 2.1. The Latency HIV-CRISPR screen.

(A) Schematic summarizing the Latency HIV-CRISPR screen in J-Lat cell lines showing the latent, integrated provirus and the HIV-CRISPR vector which delivers Cas9 and a sgRNA and produces a packageable genomic RNA. Red boxes represent the gene or sgRNA targeting the gene of interest; gray boxes represent functional HIV-1 LTRs at both sides of the vector and provirus; triangles represent internal promoters for sgRNA and Cas9 transcription. **(B)** Categorical distribution of the genes targeted by the Human Epigenome (HuEpi) sgRNA library. **(C)** Metascape Gene Ontology (GO) analysis [3] of the genes targeted by the HuEpi sgRNA library. **(D)** Summary of workflow from the generation of the pool of J-Lat HuEpi knockout cells to the sequencing and comparison of the abundance of guideRNAs found in the viral RNA (vRNA) pool versus the genomic DNA (gDNA) pool. GuideRNAs enriched in the viral supernatant RNA relative to the genomic DNA represent the gene(s) that upon knockout result in latency reversal. **Figure 2.1** was reproduced from [4] with permission, Creative Commons Attribution-Noncommerical-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

CRISPR vector is that in the presence of replication-competent HIV-1, the genome of the HIV-CRISPR vector is transcribed into RNA and packaged *in trans* into HIV-1 virions. We reasoned that in J-Lat cells, an *in vitro* T-lymphocyte HIV-1 latency model which harbors a near full-length HIV-1 provirus, any sgRNAs encoded in HIV-CRISPR genomes that target candidate HIV-1 latency genes will be packaged into the reactivated virus particle, secreted, and enriched in the viral supernatant (**Figure 2.1A**). Consequently, the RNA of the viral supernatant from the HIV-CRISPR latency screen can undergo next-generation sequencing to determine sgRNAs that are enriched in the viral particles and thereby identify candidate genes involved in HIV-1 latency maintenance. This results in a direct readout for the gene knockout which incorporates functional, reactivation activity.

To specifically investigate host epigenetic regulators involved in the maintenance of HIV-1 latency, we generated a custom human epigenome specific sgRNA CRISPR library (HuEpi). This library contains sgRNAs targeting epigenome factors such as histones, histone binders (e.g., histone readers and chaperones), histone modifiers (e.g., histone writers and erasers), and general chromatin associated factors (e.g., RNA and DNA modifiers) (**Figure 2.1B, C**). Most of the genes targeted in this library are derived from the EpiFactor database [158]. To this set we also added histones and other hand-selected gene regulatory complexes. The total library contains 5,309 sgRNAs targeting 841 genes (6 sgRNAs per gene with a few exceptions) and 252 non-targeting controls (NTCs) [159] (**Table S1,** see https://www.biorxiv.org/content/10.1101/2022.08.23.504195v1.supplementary-material). The small size of the library improves the likelihood of high coverage of all the sgRNAs throughout the CRISPR screening process relative to a whole-genome screen.

The sgRNA sequences from the HuEpi library were cloned in bulk into the HIV-CRISPR vector and the lentiviral library was subsequently transduced into J-Lat cells at a low MOI (MOI = 0.4) (**Figure 2.1D-1**). We transduced two independent J-Lat cell clones, the J-Lat 10.6 [61] and J-Lat 5A8 [62] cells, which each have different proviral integration sites and the latter of which responds to LRAs more similarly to primary CD4+ T cell HIV-1 latency models compared to other cell-line models [63]. Transduced cells were then subjected to 10-14 days of puromycin selection after which the cells and viral supernatant were harvested and processed by next-generation sequencing of the sgRNAs to identify the candidate HIV-latency genes (**Figure 2.1D-3**). Genomic DNA was extracted from the cell pellet to determine the baseline representation of each sgRNA in the transduced sgRNA library. Viral RNA was extracted from the concentrated viral supernatant to identify the sgRNA population that facilitated latent HIV-1 provirus transcription reactivation (**Figure 2.1D-4**). Using the MAGeCK pipeline [160], we identified candidate HIV-1 latency maintenance factors by selecting genes in which the sgRNA counts of the viral RNA were enriched compared to the sgRNA counts in the genomic DNA. Genes with an increased sgRNA count in the viral RNA, which equates to an increase in the fold change, are predicted to play a role in maintaining HIV-1 latency.

Overall, the Latency HIV-CRISPR screen results for both J-Lat 10.6 and J-Lat 5A8 cell lines demonstrated a significant enrichment in specific sgRNAs compared to the NTCs. We compared the distribution of the log fold change (LFC) of NTCs against the gene targeting sgRNAs. While the LFC of most gene targeting sgRNAs clustered around the median LFC of the NTCs, we observed that a subset of the gene targeting sgRNAs were significantly enriched compared to the NTC population. These sgRNAs are the sgRNAs of interest that target potential HIV-1 latency maintenance factor genes (**Figure 2.2A**).

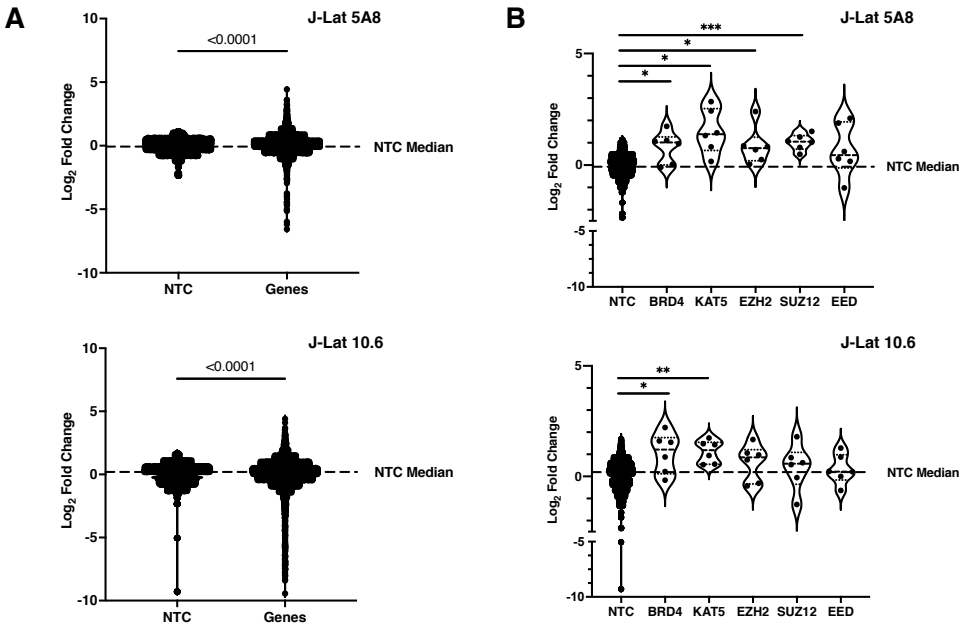
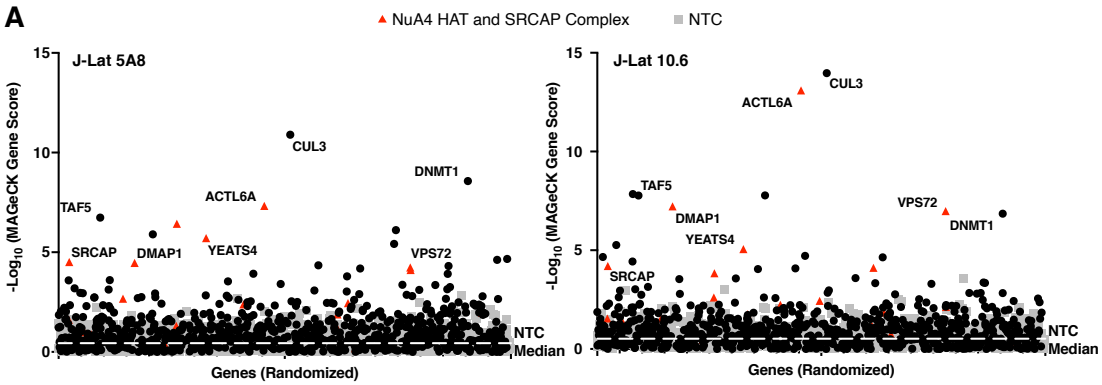
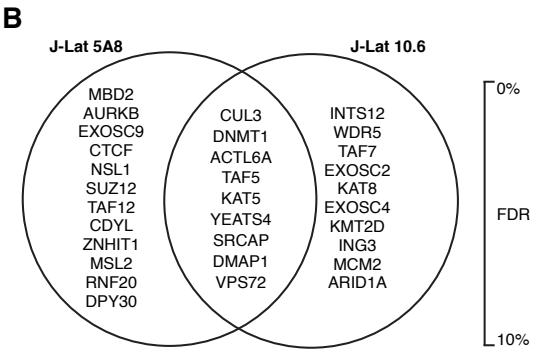


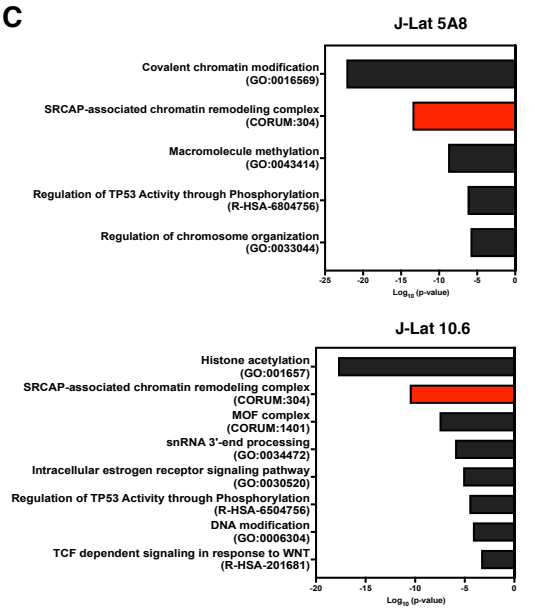
Figure 2.2. GuideRNA level enrichment for known and novel genes of interest in the Latency HIV-CRISPR screen.

(A) GuideRNA enrichment (log2 of fold change) in both J-Lat cell lines for the genes targeted by the HuEpi sgRNA library compared to the Non-Targeting Control (NTC) guides. For statistical analysis, the HuEpi gene sgRNAs are compared to the NTC control. **(B)** GuideRNA level enrichment of known HIV-1 latency maintenance factors BRD4, KAT5, and PRC2 complex members (EZH2, SUZ12, and EED) in the Latency HIV-CRISPR screen. For statistical analysis, all conditions are compared to the NTC control. Welch's t-test, p-value =<0.05 = *, =<0.05 = **, =<0.001 = ***. **Figure 2.2** was reproduced from [4] with permission, Creative Commons Attribution-Noncommerical-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

To determine top candidate genes from the Latency HIV-CRISPR screen, we ranked the genes by MAGeCK gene score. For both J-Lat cell lines, we observe that there are many genes that are significantly enriched compared to the NTCs. We identified eight previously unreported epigenetic HIV-1 latency maintenance genes, defined by a false discovery rate (FDR) less than 10% in both J-Lat 10.6 and J-Lat 5A8 cell lines, as our top hits and they are labelled in the figure (**Figure 2.3A**). BRD4 and KAT5 are two host factors previously reported to silence HIV-1 transcription ([107, 149, 161] and [147], respectively), and in our screen, the sgRNAs targeting *BRD4* and *KAT5* are significantly enriched compared to the sgRNAs targeting NTCs in both J-







D

		-log10 (MAGeCK Gene Score) J-Lat 5A8	J-Lat 10.6
NuA4 HAT complex	TRRAP	2.45	1.85
	EP400	2.38	2.24
	EPC2	0.99	1.33
	EPC1	0.38	0.47
	KAT5	6.43	3.83
	ING3	1.88	4.11
	MRGBP	0.88	0.26
	ACTB	0.19	0.75
	MORF4L1	1.30	2.61
	MORF4L2	1.34	1.18
	MEAF6	2.67	1.54
	MBTD1	1.96	1.17
SRCAP complex	BRD8	1.62	1.54
Shared	DMAP1	4.46	7.22
Shared	RUVBL2	1.25	1.76
Shared	RUVBL1	1.96	0.32
Shared	ACTL6A	7.32	13.10
Shared	VPS72	4.24	6.98
Shared	YEATS4	5.71	5.06
Shared	SRCAP	4.51	4.20
	ACTR6	0.38	2.44
	ZNHIT1	4.11	2.13

Figure 2.3. The Latency HIV-CRISPR screen in J-Lat 10.6 and J-Lat 5A8 cells identifies a set of mutual, novel hits.

(A) The -log10 of the MAGeCK scores for each gene targeted by the HuEpi sgRNA library (black circles and red triangles) and NTCs (gray squares) are calculated and displayed. Gene names are labelled for hits that have a <10% false discovery rate (FDR) in both J-Lat cell lines (center of Venn diagram in (B)); red triangles represent members of the NuA4 HAT complex and SRCAP complex. NTCs are artificial NTC genes designed by iterative binning of NTC sgRNA sequences (see Methods). Genes are randomized on the x-axis, but the same order is used for both right and left panels. The y-axis is the inverse log10 of the MAGeCK score. **(B)** Top hit genes (cut off of <10% FDR) in common and unique to each J-Lat cell line are ordered by significance and FDR with the top of the list having the highest significance and lowest FDR. **(C)** Metascape GO analysis [3] of the gene hits with a <10% FDR in each J-Lat cell line. **(D)** Analysis of the -log10 of the MAGeCK scores of the genes overlapping and unique to the NuA4 HAT and SRCAP complex compared to the NTCs. The higher the number, the more statistically significant it is of a hit. Red font is for genes that score higher than the average NTC score. **Figure 2.3** was reproduced from [4] with permission, Creative Commons Attribution-Noncommerical-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

Lat cell lines (**Figure 2.2B; Table S2** see https://www.biorxiv.org/content/10.1101/2022.08.23.504195v1.supplementary-material). In addition to individual genes, the Latency HIV-CRISPR screen also identified genes encoding subunits of protein complexes important to HIV-1 latency. PRC2 has also been previously reported to silence HIV-1 transcription [131, 146] and members of the complex including *EZH2*, the catalytic component, and *SUZ12* and *EED* showed a trend of enrichment in our screen (**Figure 2.2B; Table S2** see https://www.biorxiv.org/content/10.1101/2022.08.23.504195v1.supplementary-material). Thus, our Latency HIV-CRISPR screen system identifies bona fide HIV-1 latency maintenance factors.

2.3.2 Latency HIV-CRISPR identifies multiple members of the NuA4 complex and CUL3

To identify the top candidate HIV-1 latency genes, we integrated the guideRNA data to a gene level using MAGeCK analysis [160] (**Figure 2.3A; Table S3** see https://www.biorxiv.org/content/10.1101/2022.08.23.504195v1.supplementary-material). As we were primarily interested in hits that are independent of integration site, we considered the top

gene hits as those that that have an FDR < 10% and also are shared between both J-Lat 10.6 and J-Lat 5A8 cell lines (**Figure 2.3B**, middle of Venn diagram). We observed that multiple members of the Nucleosome Acetyltransferase of H4 (NuA4) complex scored highly in the screen (**Figure 2.3A**). The NuA4 complex is highly conserved in eukaryotes [162] and acetylates histone H4 [163] and has been implicated in many genomic processes including DNA damage repair and transcription [164]. Many proteins that are part of the NuA4 complex also overlap with the SNF2-Related CBP Activator Protein (SRCAP) complex which exchanges the H2A of the canonical nucleosome with the H2A.Z variant [165-167]. A gene ontology (GO) enrichment analysis [3] of the top gene hits (FDR <10% for each J-Lat cell line) demonstrated that the SRCAP-associated chromatin remodeling complex is highly enriched in both J-Lat cell lines (**Figure 2.3C**). Based on the MAGeCK gene scores, six of the seven genes (*BRD8*, *DMAP1*, *RUVBL1*, *RUVBL2*, *ACTL6A*, *VPS72*, and *YEATS4*) that encode proteins overlapping between the NuA4 HAT and SRCAP complexes scored higher than the median MAGeCK score for the NTCs (**Figure 2.3D**). The only exception is *RUVBL1* in the J-Lat 10.6 screen, which scored below the NTC median MAGeCK score (RUVBL1 = 0.318 vs NTC median = 0.518). Additionally, most genes unique to each complex also had MAGeCK scores above the NTC median in both cell lines. Genes unique to the NuA4 HAT complex included *TRRAP*, *EP400*, *EPC2*, *KAT5*, *ING3*, *MORF4L1*, *MORF4L2*, *MEAF6*, and *MBTD1* and genes unique to the SRCAP complex included *SRCAP* and *ZNHIT1* (**Figure 2.3D**). This analysis suggests that the Latency HIV-CRISPR screen strategy can extend beyond identification of single genes of interest to groups of genes encoding complexes of relevance, which can further shed light on the mechanisms involved in maintaining HIV-1 latency.

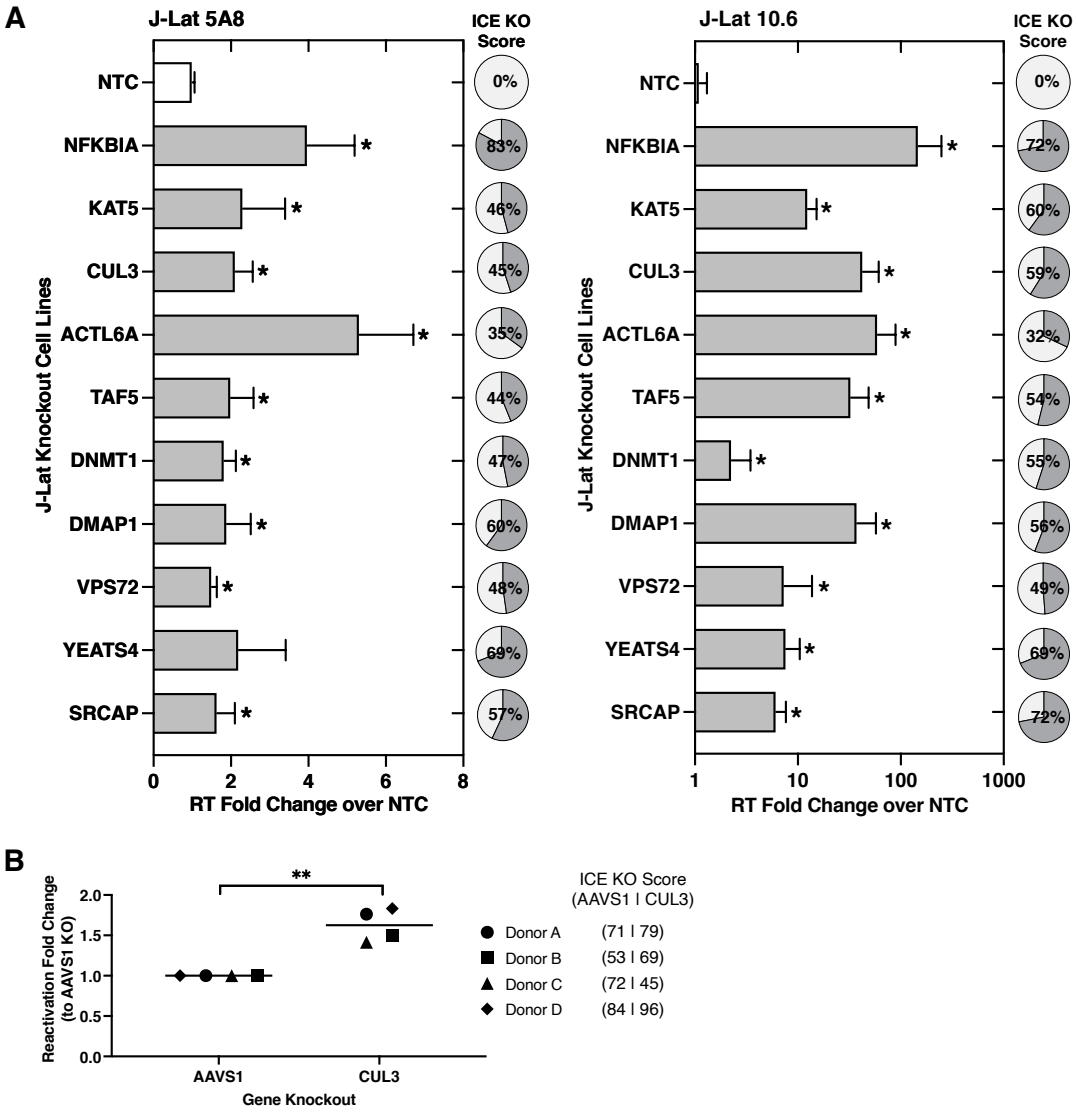


Figure 2.4. Validation of top hits from Latency HIV-CRISPR screen.

(A) Validation of the top 9 gene hits of the Latency HIV-CRISPR screen was performed by individually knocking out each J-Lat cell line with two different guide RNAs and measuring viral reactivation by quantifying HIV-1 reverse transcriptase activity of the viral supernatant. *NFκBIA* knockout is a positive control. The reverse transcriptase activity from released virions was normalized to basal activity from the NTC knockout. For each for the knockout cell lines, ICE analysis was performed and the average knockout is shown in pie charts. Multiple unpaired t-tests, p-value =<0.05 = *. **(B)** *CUL3* knockout resulted in significant viral reactivation in a primary CD4+ T cell model of HIV-1 latency using cells from four healthy donors. Viral reactivation was measured by flow cytometry and normalized to the *AAVS1* knockout cells. Paired t-test, p-value =<0.05 = *, =<0.05 = **. **Figure 2.4** was reproduced from [4] with permission, Creative Commons Attribution-Noncommerical-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

To validate the top screen hits, we generated four individual J-Lat 10.6 and 5A8 knockout cell lines for each gene (two sgRNAs per gene). As positive controls, we used two sgRNAs targeting *NFκBIA*, an inhibitor for NFκB, and a gene that was not included in the HuEpi library. Upon knockout of *NFκBIA* in the J-Lat cells, we expect an increase in HIV-1 transcriptional activity, which leads to virus production. As negative controls, we used two NTC sgRNAs that do not target human loci. To assess latency reversal in the J-Lat cells, we measured HIV-1 reverse transcriptase activity in the supernatant, which is a readout after the viral transcription, translation, and budding steps have occurred. Because the transduced cells were assayed as a pool of gene knockouts, we simultaneously measured CRISPR editing efficiency in bulk using the program Inference of CRISPR Edits (ICE) [168]. The knockout efficiency for the heterogenous cell lines we generated varied between 32% to 83% (**Figure 2.4A**). A recent study demonstrated that a median knockout of 35% was less efficient, but still had a meaningful effect [169].

Compared to the cell lines transduced with NTC sgRNAs, we found that individually knocking out all eight novel candidate gene hits resulted in significant latency reversal in at least one of the J-Lat cell lines we tested (**Figure 2.4A**). In fact, some of the reverse transcriptase measurements from the gene knockout cell lines are comparable to the positive control knockout of *NFκBIA*. In addition, we confirmed that knockout of a top scoring gene hit in our screen, *KAT5*, a previously reported HIV-1 latency maintenance factor [147], results in viral reactivation in both J-Lat cell lines (**Figure 2.4A**). Moreover, genes encoding protein components of the NuA4 HAT and SRCAP complexes described above, *ACTL6A*, *DMAP1*, *VPS72*, and *YEATS4* also all validated as latency maintenance factors in J-Lat cells (**Figure 2.4A**). *ACTL6A* knockout, in particular, results in a high level of HIV-1 reactivation, which may be a result of *ACTL6A*

encoding a protein that is a subunit of both NuA4 HAT and SRCAP complexes (**Figure 2.3D**). Other genes validated in this functional investigation include *DMAP1* which encodes a protein that forms a complex with DNMT1 to mediate transcriptional repression [170] and has been described to play a role in HIV-1 latency in some studies [148] and *TAF5* which functions in scaffold formation and is a critical subunit of the general transcription factor TFIID.

As *CUL3* was the top gene hit in our Latency HIV-CRISPR screen in both J-Lat 10.6 and J-Lat 5A8 cell lines and validated as a latency maintenance factor in J-Lat cells, we further characterized *CUL3* in the context of a primary CD4+ T cell HIV-1 latency model system. CUL3 is part of the cullin family and serves as the scaffold for Cullin-RING E3 ligase complexes, which ubiquitinate proteins [171]. For our primary cell HIV-1 latency model system, we used a dual HIV-1 reporter virus (pNL4-3-Δ6-dreGFP-CD90) that is derived from previously established systems [68, 172]. Here, the reporter virus has been modified to encode a destabilized eGFP gene and a mouse Thy1.2 (CD90.2) gene, which are separated by an IRES and both reporters are under control of the HIV-1 LTR [173]. As the mouse Thy1.2 reporter has a long half-life and is slow to turnover, it serves as a marker to determine the cells that have been infected by the reporter virus (Thy1.2+). Amongst the infected cells (Thy1.2+), the eGFP reporter, which has a short half-life, serves as a marker to determine the cells that are actively transcribing or actively infected (Thy1.2+, GFP+) versus minimally transcribing or latently infected (Thy1.2+, GFP-). We knocked out *CUL3* in the primary CD4+ T cell HIV-1 latency model using Cas9-gRNA ribonucleoproteins (RNPs) including 3 different sgRNAs. As a negative control, the *AAVS1* gene is targeted for knockout, which is in a “safe harbor” locus in which disruption of this gene does not have adverse effects on the cell [174].

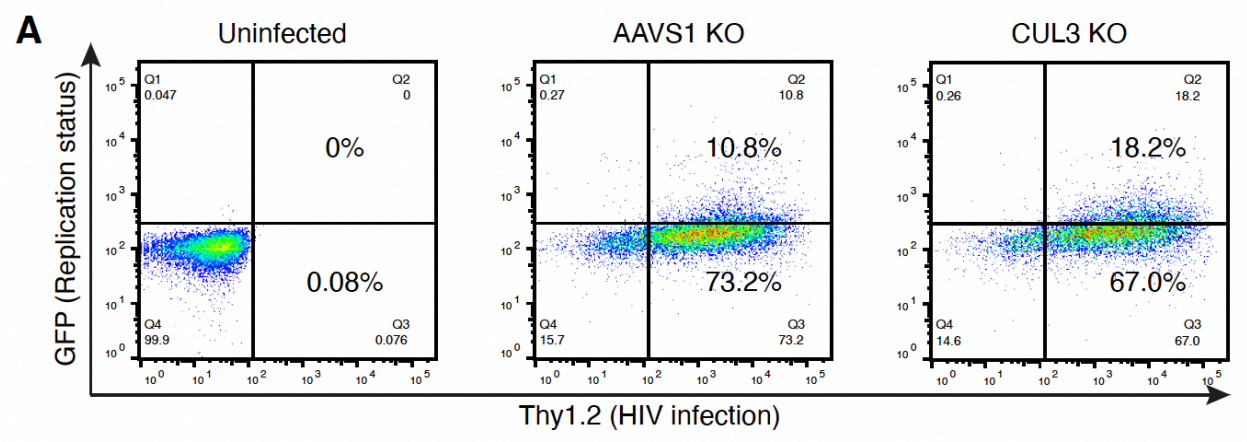


Figure 2.5. Validation of the top hit from the Latency HIV-CRISPR screen in primary CD4+ T cells.

(A) Representative flow cytometry plots of viral reactivation levels in wildtype primary CD4+ T cells and primary CD4+ T cell model of HIV latency cells upon knockout of *AAVS1* and *CUL3*. **Figure 2.5** was reproduced from [4] with permission, Creative Commons Attribution-Noncommerical-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

We expected that if CUL3 plays a role in HIV-1 latency establishment, there will be a higher reactivated population in the *CUL3* knockout cells compared to the negative control (*AAVS1*) knockout cells. We acquired CD4+ T cells from four healthy donors and infected the cells with the dual reporter virus. Indeed, upon performing the described CRISPR/Cas9-mediated knockout and measuring the functional output of the reporter virus, we observed that *CUL3* knockout in all four donors resulted in significant HIV-1 reactivation and subsequent decrease in the latent population compared to *AAVS1* knockout (**Figure 2.4B; 2.5A**). This suggests that post-translational ubiquitin modification by CUL3 of a substrate may facilitate the establishment and maintenance of HIV-1 latency. Our data is consistent with previous work as CUL3 has been demonstrated to negatively regulate HIV-1 transcription during productive infection through the NFκB pathway [175]. Moreover, these data demonstrate that the Latency HIV-CRISPR screening strategy is effective in identifying novel epigenetic HIV-1 latency maintenance genes that have been validated in multiple systems.

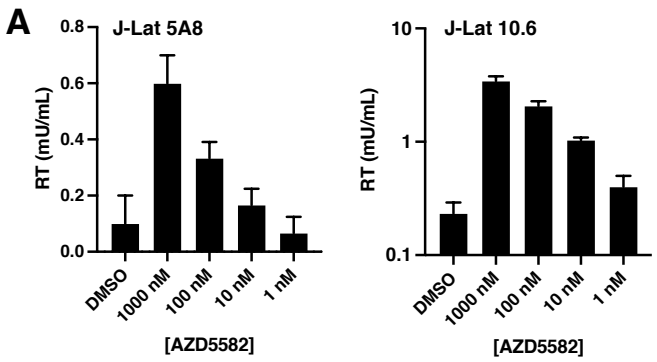


Figure 2.6. AZD5582 dose curve in J-Lat cells.

(A) AZD5582 dose curve performed on both J-Lat 10.6 and 5A8 cell lines to determine viral reactivation levels. **Figure 2.6** was reproduced from [4] with permission, Creative Commons Attribution-Noncommerical-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

2.3.3 An LRA Latency HIV-CRISPR screen reveals regulatory changes that act in combination during reactivation of HIV-1 latency

A critical extension of this HIV-CRISPR latency screen methodology is the ability to examine other transcriptional mechanisms in conjunction with epigenetic regulation factors. Thus, to address HIV-1 latency as a state of multiple, parallel mechanisms, we modified the screen by treating the pool of human epigenome knockout cells with a low-activating dose of LRA. The goal was to identify instances of combinations between the gene knockout and LRA resulting in a significant increase in viral reactivation (i.e. a top hit in the screen). As a proof of principle, we used the recently identified LRA, AZD5582, which is a SMAC mimetic and non-canonical NFκB activator [83]. We first determined 10 nM AZD5582 as the low-activating dose of LRA to use by performing an HIV-1 viral reactivation dose curve using both J-Lat cell lines (**Figure 2.6**). The same HuEpi knockout pool of cells used in the Latency HIV-CRISPR screen described in **Figure 2.3** were in parallel treated for 24 hours with 10 nM of AZD5582 (**Figure 2.1D-4**) to identify epigenetic regulatory genes that upon knockout, combine with the

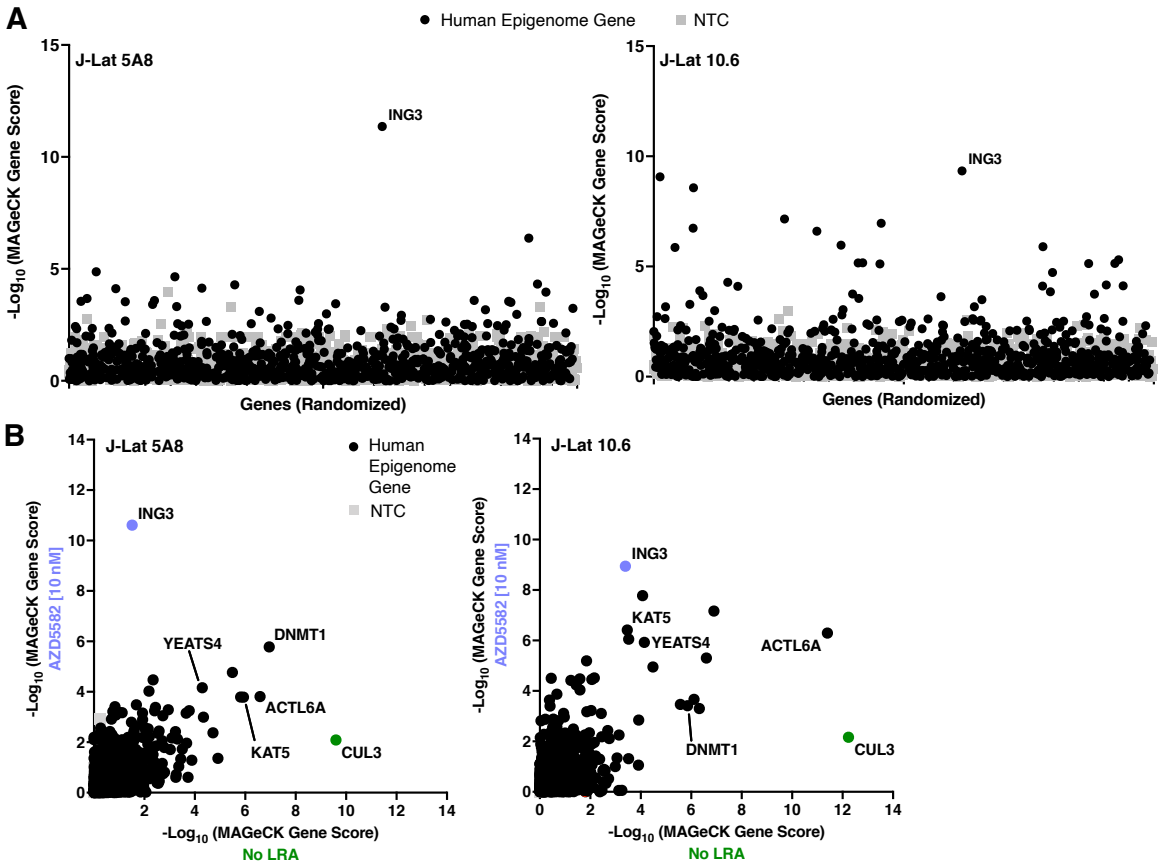


Figure 2.7. LRA Latency HIV-CRISPR screen identifies *ING3* in combination with AZD5582 as a HIV-1 latency maintenance factor.

(A) ING3 is the top hit of the LRA Latency HIV-CRISPR screen. The Latency HIV-CRISPR screen HuEpi knockout cells were treated with a low activating dose (10 nM) of AZD5582. The -log10 of the MAGeCK score (on the y-axis) for each gene targeted by the HuEpi sgRNA library (black circles) and NTCs (gray squares) are calculated and displayed. NTCs are artificial NTC genes designed by iterative binning of NTC sgRNA sequences (see Methods). Genes are randomized on the x-axis, but the same order is used for both right and left panels. **(B)** Comparison of the Latency HIV-CRISPR screen by MAGeCK score in the presence (y-axis) and absence of AZD5582 (x-axis) with HuEpi genes (circles) and NTCs (gray squares). The data for the screen without an LRA (x-axis) is from Figure 2.3 as these two screens were performed in parallel. The genes unique to each screen are closest to the respective axis and the genes that are in common to both screens are at the center. *ING3* is highlighted in periwinkle and *CUL3* in green. **Figure 2.7** was reproduced from [4] with permission, Creative Commons Attribution-Noncommerical-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

transcriptional initiation pathway targeted by AZD5582, to result in a significant increase in viral reactivation.

The most prominent hit of the AZD5582 LRA Latency HIV-CRISPR screen in both J-Lat cell lines was *Inhibitor of Growth Family Member 3* (*ING3*), a novel HIV-1 latency factor (**Figure 2.7A; Table S4, S5** see https://www.biorxiv.org/content/10.1101/2022.08.23.504195v1.supplementary-material). To confirm that *ING3* was a top gene hit unique to the LRA condition, we compared the LRA Latency HIV-CRISPR screen results to the Latency HIV-CRISPR screen results in the absence of LRA. Overall, many gene hits overlapped between the Latency HIV-CRISPR screen in the presence and absence of AZD5582 (**Figure 2.7B**). For example, the previously identified *KAT5* gene and genes that we validated in **Figure 2.4** including *DNMT1*, *YEATS4*, and *ACTL6A* were still amongst the top gene hits, which suggests that these genes broadly play a role in HIV-1 latency maintenance regardless of activation of the non-canonical NFκB pathway. We verified that *ING3* was the top gene hit unique to the screen in the presence of AZD5582 and sought to perform additional validation for this gene. We also observed that *CUL3*, the top hit unique to the screen in the absence of AZD5582, was relatively depleted in the screen conducted in the presence of AZD5582 (**Figure 2.7B**). This suggests a potential overlap between AZD5582 non-canonical NFκB function and *CUL3* function in the context of HIV-1 latency reversal.

To validate *ING3*, the top hit enriched in the AZD5582 LRA Latency HIV-CRISPR arm of the screen, we generated *ING3* knockout cell lines in both J-Lat 10.6 and 5A8 cells. Simultaneously, cell lines transduced with the NTC sgRNAs were generated as a negative control. The pools of knockout cells were treated with or without a low reactivating dose (10 nM) of AZD5582 for 24 hours and then viral reactivation was measured by HIV-1 reverse

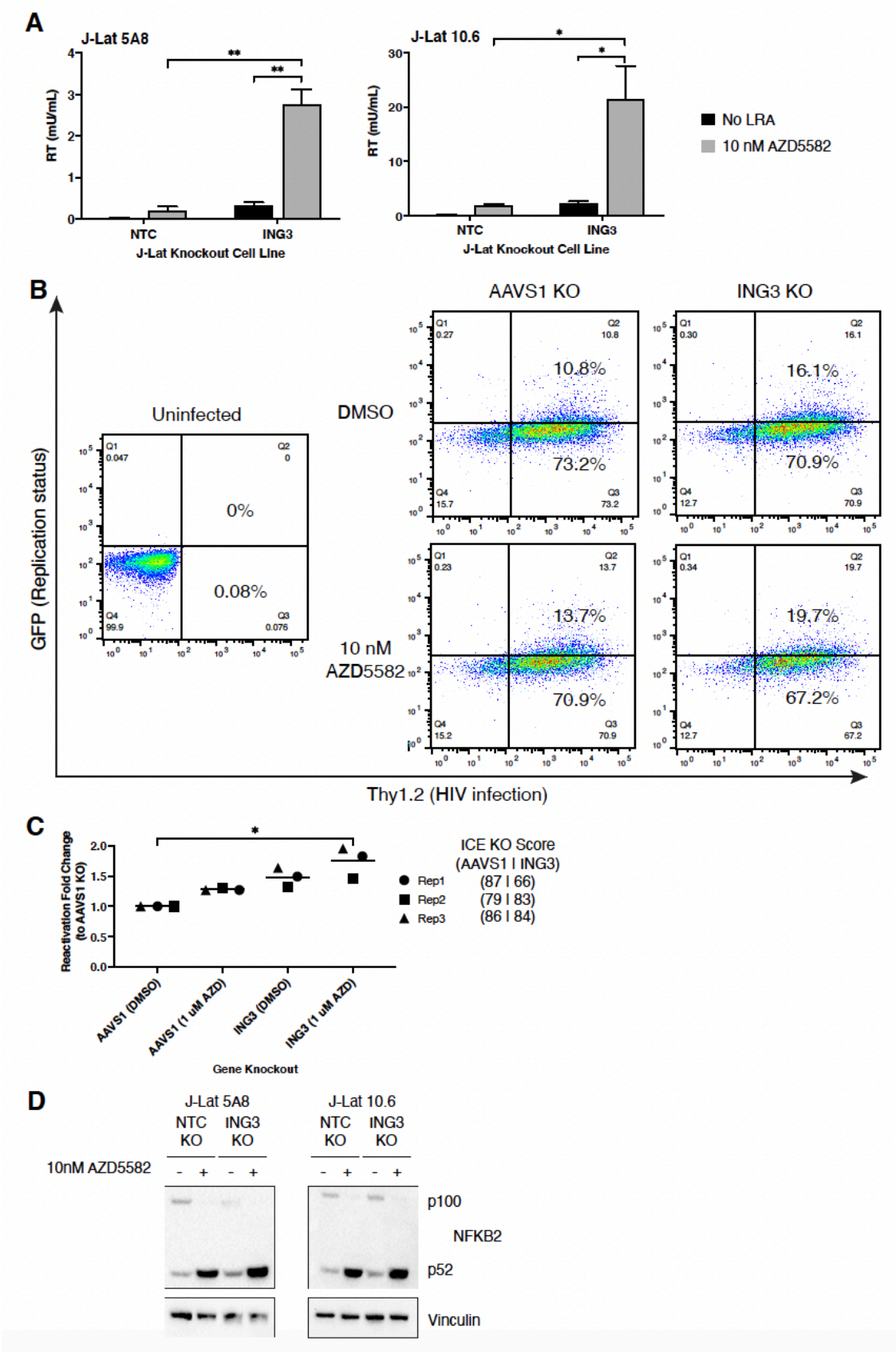

A
J-Lat 5A8
RT (mU/mL)
NTC
ING3
J-Lat Knockout Cell Line
J-Lat 10.6
No LRA
10 nM AZD5582
B
AAVS1 KO
ING3 KO
Uninfected
DMSO
10 nM AZD5582
0%
0.08%
10.8%
73.2%
16.1%
70.9%
13.7%
70.9%
19.7%
67.2%
GFP (Replication status)
Thy1.2 (HIV infection)
C
Reactivation Fold Change (to AAVS1 KO)
AAVS1 (DMSO)
AAVS1 (1 uM AZD)
ING3 (DMSO)
ING3 (1 uM AZD)
Gene Knockout
ICE KO Score (AAVS1 I ING3)
Rep1 (87 I 66)
Rep2 (79 I 83)
Rep3 (86 I 84)
D
J-Lat 5A8
J-Lat 10.6
NTC KO
ING3 KO
10nM AZD5582
p100
NFKB2
p52
Vinculin

Figure 2.8. Validation of *ING3* in combination with AZD5582 as a HIV-1 latency maintenance factor.

(A) HIV-1 reverse transcriptase activity (y-axis) of the viral supernatant of NTC sgRNA transduction or *ING3* knockout in J-Lat 10.6 and 5A8 treated with 10 nM AZD5582 or an equivalent volume of DMSO. *ING3* knockout and AZD5582 treatment combine to result in a significant increase in viral reactivation. Paired t-test, p-value =<0.05 = *, =<0.05 = **. **(B)** Representative flow cytometry plots of primary CD4+ T cell HIV-1 latency model cells that are *AAVS1* or *ING3* knockouts treated with DMSO or 1 μM AZD5582 treatment. Thy1.2-, GFP- cells (quadrant 4) are uninfected; Thy1.2+, GFP- (quadrant 3) cells are infected with the dual reporter HIV-1 virus and latent; Thy1.2+, GFP+ cells (quadrant 2) are infected and reactivated. **(C)** Three independent knockouts of *AAVS1* and *ING3* in primary CD4+ T cell HIV-1 latency model cells were performed in one healthy donor and each pool of knockout cells were treated with DMSO or 1 μM AZD5582 (AZD). Knockout of *ING3* and AZD5582 treatment combined resulted in significant reactivation compared to knockout of AAVS1. Paired t-test, p-value =<0.05 = *. **(D)** Western blot showing similar p52 levels are detected in the control and ING3 knockout J-Lat 10.6 and 5A8 cell lines upon treatment of 10 nM AZD5582. Activation of the non-canonical NFκB (NFκB2) pathway is marked by a decrease in p100 and an increase in the cleaved product of p52. **Figure 2.8** was reproduced from [4] with permission, Creative Commons Attribution-Noncommerical-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

transcriptase activity. Based on screen results, we hypothesized that knockout of *ING3* alone would result in some level of viral reactivation compared to the transduction of NTC sgRNAs, but upon treatment with a low reactivating dose of AZD5582, viral reactivation would become even more pronounced. Consistent with our hypothesis, we observed that in both J-Lat cell lines, low reactivating dose AZD5582 treatment of the cell lines transduced with the NTC sgRNAs results in minimal viral reactivation, but treatment of the *ING3* knockout cell lines results in significant, increased viral reactivation (**Figure 2.8A**). These results suggest a mechanistic interplay between *ING3* inhibition and AZD5582 treatment during the reactivation of HIV-1 transcription. Moreover, we observed a similar trend in the primary CD4+ T cell HIV-1 latency model described previously. In one healthy donor sample of CD4+ T cells in which the latency model was established, we performed three independent *AAVS1* and *ING3* knockouts, followed by 1 μM AZD5582 treatment, and observed significant viral reactivation in the combination condition of *ING3* knockout and AZD5582 treatment (**Figure 2.8B, C**). We observed similar results in a total of two healthy donor samples of CD4+ T cells (**Figure 2.9A**). This enhancement

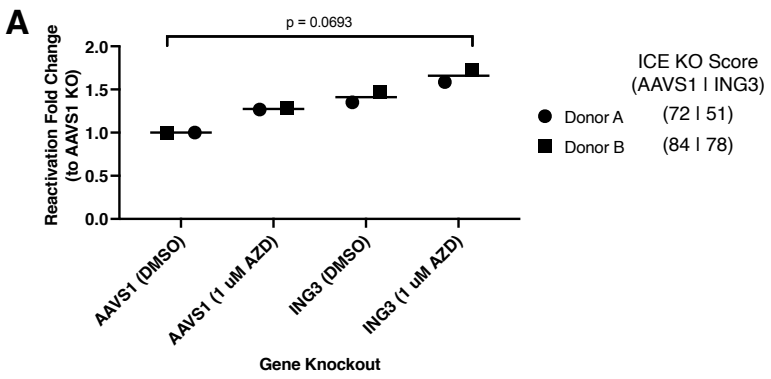


Figure 2.9. Validation of *ING3* in combination with AZD5582 as a HIV-1 latency maintenance factor in two healthy donors.

(A) *AAVS1* and *ING3* knockouts in a primary CD4+ T cell HIV-1 latency model were performed in two healthy donors and each pool of knockout cells were treated with DMSO or 1 μM AZD5582. *ING3* knockout and AZD5582 treatment combined resulted in significant reactivation compared to *AAVS1* knockout. **Figure 2.9** was reproduced from [4] with permission, Creative Commons Attribution-Noncommerical-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

of viral reactivation in two latency models suggests that there may be interplay between the function of *ING3* and the non-canonical NFκB pathway, which is targeted by AZD5582, both in cell lines and in primary CD4+ T cells.

As AZD5582 activates the non-canonical NFκB pathway, it is possible that *ING3* knockout could either enhance this activity or act through an independent pathway in the context of HIV-1 viral reactivation. To distinguish between these two possibilities, we measured the protein levels of NFκB2 products which includes p100, the cytoplasmic NFκB2 protein, and p52, the cleaved and active subunit that translocates to the nucleus. As expected, activation of the non-canonical NFκB upon AZD5582 treatment results in an increase in the cleaved p52 product and a decrease in the p100 product in the control (NTC) cells (**Figure 2.8D**). We find that p52 levels are similar in the control and *ING3* knockout J-Lat cell lines (**Figure 2.8D**) suggesting that another pathway besides the non-canonical NFκB is the main driver resulting in HIV-1 reactivation. Altogether the LRA Latency HIV-CRISPR screen demonstrates the powerful

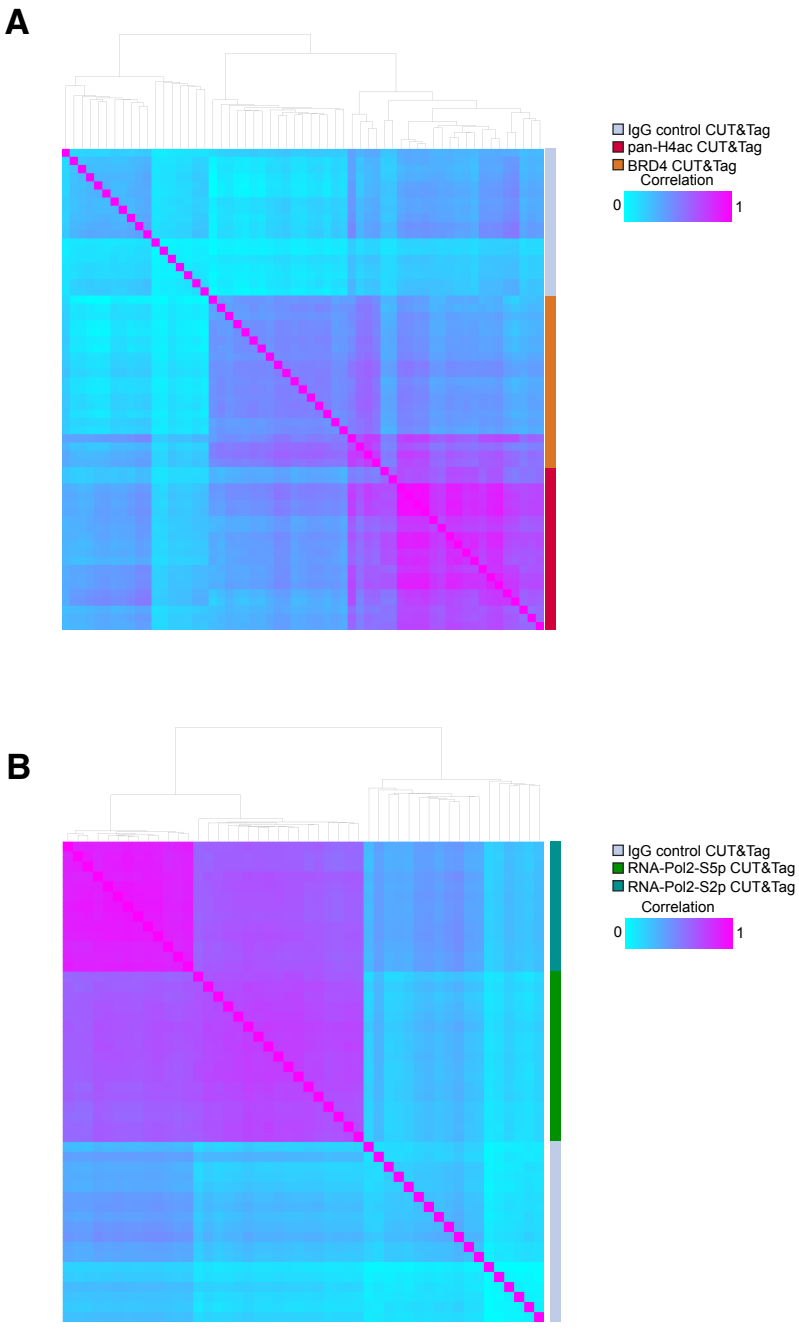


Figure 2.10. The genome-wide signal of automated CUT&Tag replicates is highly correlated.

(A) Correlation Matrix colored according to the pair-wise Pearson correlation of pan-H4Ac, BRD4, and IgG negative control samples across the merged pan-H4Ac and BRD4 peak sets. All pan-H4Ac samples group together by hierarchical clustering as do all of the BRD4 samples. **(B)** Same as (A) but showing the pair-wise Pearson correlation of RNA-Pol2-S5p and RNA-Pol2-S2p over the merged peak sets of these marks. **Figure 2.10** was reproduced from [4] with permission, Creative Commons Attribution-Noncommerical-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

potential for examining multiple mechanisms simultaneously to gain a more comprehensive understanding of the mechanisms underpinning HIV-1 latency and to further improve existing LRAs.

2.3.4 ING3 knockout in combination with AZD5582 treatment decreases pan-H4Ac and BRD4 levels and increases RNA-Pol2-S5p levels at the HIV-1 LTR

To better understand how AZD5582 treatment and *ING3* knockout act in combination to promote reactivation of latent HIV-1 proviruses, we examined the chromatin-related changes to the HIV-1 provirus by performing automated CUT&Tag [157]. Specifically, we performed genome-wide profiling of the J-Lat 10.6 cell line under four different conditions: (i) Transduction of NTC sgRNA with a treatment of DMSO (NTC + DMSO), the negative control (ii) Transduction of NTC sgRNA with a treatment of 10 nM AZD5582 (NTC + AZD5582) (iii) *ING3* knockout with a treatment of DMSO (ING3 KO + DMSO) (iv) *ING3* knockout with a treatment of 10 nM AZD5582 (ING3 KO + AZD5582).

ING3 is a known member of the NuA4 HAT complex, which functions to acetylate the histone H4 tail [176]. BRD4 interacts with acetylated H4 [147] and both isoforms of BRD4 are known to negatively regulate HIV-1 transcription [107, 149]. To examine whether the reduction in *ING3* promotes activation of HIV-1 by modulating histone H4 acetylation and BRD4 occupancy on the HIV-1 LTR, we performed automated CUT&Tag using antibodies that recognized acetylated histone H4 (pan-H4Ac), BRD4, or a non-specific IgG negative control. As quality control, we determined the signal levels of pan-H4Ac and BRD4 from the CUT&Tag data and confirmed the replicates of each antibody were most highly correlated amongst each individual antibody (**Figure 2.10**). We first examined pan-H4Ac marks at and around the HIV-1

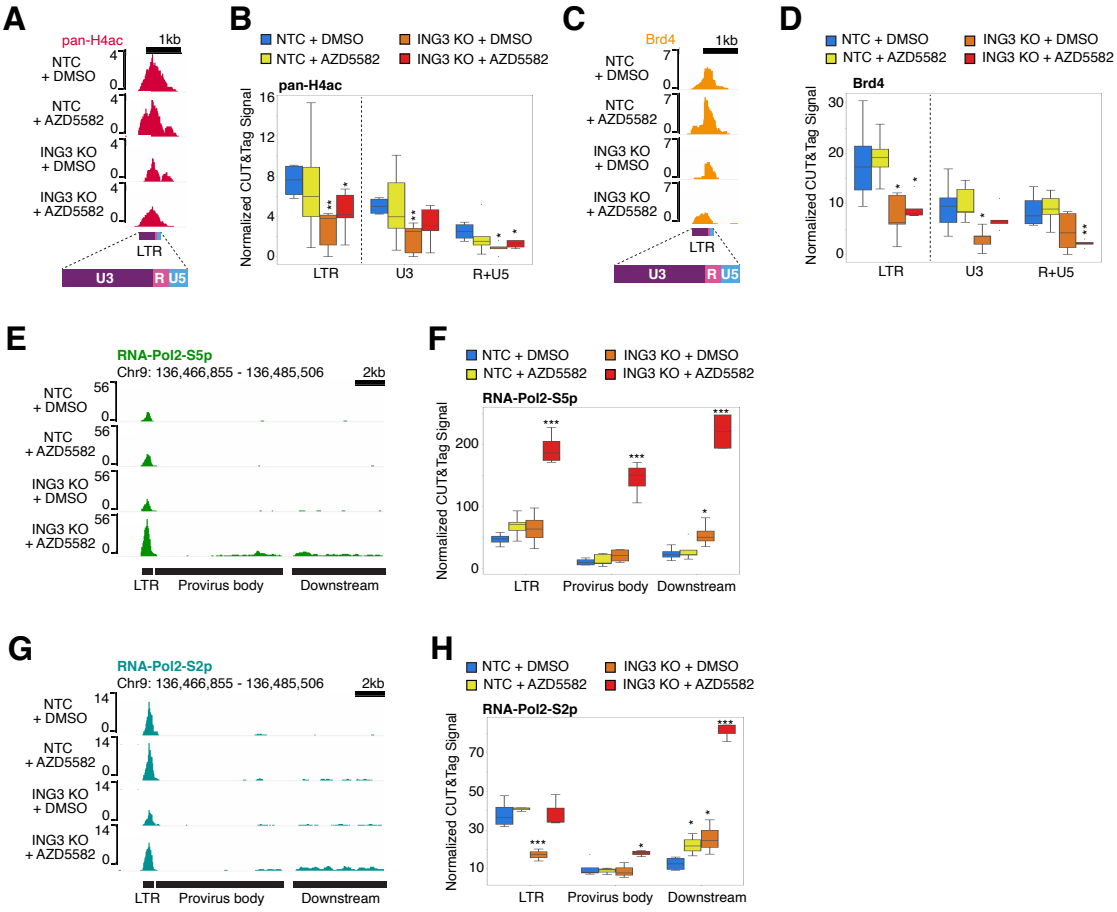

A
pan-H4ac
1kb
NTC + DMSO
NTC + AZD5582
ING3 KO + DMSO
ING3 KO + AZD5582
LTR
U3
R U5
B
NTC + DMSO
NTC + AZD5582
ING3 KO + DMSO
ING3 KO + AZD5582
pan-H4ac
Normalized CUT&Tag Signal
LTR
U3
R+U5
C
Brd4
1kb
NTC + DMSO
NTC + AZD5582
ING3 KO + DMSO
ING3 KO + AZD5582
LTR
U3
R U5
D
NTC + DMSO
NTC + AZD5582
ING3 KO + DMSO
ING3 KO + AZD5582
Brd4
Normalized CUT&Tag Signal
LTR
U3
R+U5
E
RNA-Pol2-S5p
Chr9: 136,466,855 - 136,485,506
2kb
NTC + DMSO
NTC + AZD5582
ING3 KO + DMSO
ING3 KO + AZD5582
LTR
Provirus body
Downstream
F
NTC + DMSO
NTC + AZD5582
ING3 KO + DMSO
ING3 KO + AZD5582
RNA-Pol2-S5p
Normalized CUT&Tag Signal
LTR
Provirus body
Downstream
G
RNA-Pol2-S2p
Chr9: 136,466,855 - 136,485,506
2kb
NTC + DMSO
NTC + AZD5582
ING3 KO + DMSO
ING3 KO + AZD5582
LTR
Provirus body
Downstream
H
NTC + DMSO
NTC + AZD5582
ING3 KO + DMSO
ING3 KO + AZD5582
RNA-Pol2-S2p
Normalized CUT&Tag Signal
LTR
Provirus body
Downstream

Figure 2.11. *ING3* knockout decreases pan-H4Ac and BRD4 levels and stimulates HIV-1 transcriptional initiation and elongation upon addition of AZD5582.

(A) Genome browser tracks centered over the HIV-1 LTR showing the pan-H4Ac signal decreases in the *ING3* knockout alone and in combination with AZD5582 conditions. The y-axis represents read count. Because the HIV-1 LTR sequence is identical between 5' LTR vs. the 3' LTR, the CUT&Tag reads are combined onto one LTR. The LTR is subdivided into three regions: U3, R (containing the transcription start site), and U5. **(B)** Box plot showing the pan-H4Ac levels quantified over the full LTR including the U3 region and the R+U5 regions that include the transcriptional start site. The y-axis is the pan-H4Ac base pair coverage normalized to the total base pair coverage across the genome. Blue represents the transduction of NTC sgRNA with treatment of DMSO; yellow represents the transduction of NTC sgRNA with a treatment of 10 nM AZD5582; orange represents *ING3* knockout with treatment of DMSO; red represents *ING3* knockout with treatment of 10 nM AZD5582. Replicates for IgG n=18 and pan-H4Ac n=24. For statistical analysis, all conditions are compared to the NTC knockout and DMSO treatment control. P-value <0.05 = *, <0.005=**. **(C)** Genome browser tracks showing BRD4 levels decrease in the *ING3* knockout alone and in combination with AZD5582 conditions. **(D)** Same as (B) but quantifying BRD4 levels. Replicates for BRD4 n=16. **(E)** Genome browser tracks showing RNA-Pol2-S5p levels increase at the HIV-1 LTR, as well as the body of the provirus downstream of the 5' LTR ("provirus body"), and the region of the host genome downstream of the 3' LTR of the integrated provirus ("downstream") upon *ING3* knockout and AZD5582 treatment combined. (F) Box plot showing the quantification of the RNA-Pol2-S5p CUT&Tag signal over the HIV-1 LTR, as well as the body of the provirus, and the region of the host genome downstream of the provirus. The y-axis is the RNA-Pol2-S5p base pair coverage normalized to the total base pair coverage across the genome. Blue represents the transduction of NTC sgRNA with treatment of DMSO; yellow represents the transduction of NTC sgRNA with a treatment of 10 nM AZD5582; orange represents *ING3* knockout with treatment of DMSO; red represents *ING3* knockout with treatment of 10 nM AZD5582. Replicates for RNA-Pol2-S5p n=13. For statistical analysis, all conditions are compared to the NTC knockout and DMSO treatment control. P-value <0.05 = *, <0.005=**, <0.0005=***. **(G)** Genome browser tracks showing RNA-Pol2-S2p levels increase over the body of the HIV-1 provirus as well as the host genome downstream of the provirus upon ING3 knockout and AZD5582 treatment combined. **(H)** Same as (F) but quantifying RNA-Pol2-S2p levels. Replicates for RNA-Pol2-S2p n=17. **Figure 2.11** was reproduced from [4] with permission, Creative Commons Attribution-Noncommerical-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

provirus. We observed no change in pan-H4Ac signal over the HIV-1 LTR upon 10 nM AZD5582 treatment, but consistent with the known function of ING3, upon *ING3* knockout, there was a significant decrease in pan-H4Ac signal at the HIV-1 LTR (p=0.0061) (**Figure 2.11A, B**). Additionally, the combination of *ING3* knockout with AZD5582 treatment, resulted in a significant decrease in pan-H4Ac signal at the HIV-1 LTR (p=0.0415). Notably, in the *ING3* knockout condition alone, we observed a significant decrease in pan-H4Ac signal over both the U3 region of the HIV-1 LTR (p=0.0070) as well as a significant decrease in the pan-H4Ac signal

in the R and U5 region of the HIV-1 LTR (p=0.0126) containing the HIV-1 provirus transcription start site (**Figure 2.11A, B**). When the *ING3* knockout is combined with AZD5582 treatment, there is only a significant decrease in the R and U5 region of the HIV-1 LTR (p=0.0016). Thus, *ING3* knockout appears to be the main driver of the change in pan-H4Ac levels on the HIV-1 LTR in the presence and absence of AZD5582.

We hypothesized that the reduction in pan-H4Ac levels we observed on the HIV-1 LTR in the *ING3* knockout condition would correlate with reduced recruitment of BRD4. Indeed, as compared to NTC + DMSO, AZD5582 treatment alone did not affect BRD4 levels over the HIV-1 LTR; however, BRD4 levels were significantly decreased over the LTR both in the condition of *ING3* knockout alone and in combination with AZD5582 treatment (p=0.0235, p=0.0349, respectively) (**Figure 2.11C, D**). In the absence of AZD5582, the *ING3* knockout results in a significant decrease in the occupancy of BRD4 only over the U3 region of the LTR (p=0.0219), while in the presence of AZD5582 combined with the *ING3* knockout, the decrease in BRD4 levels was only significant over the R and U5 region of the HIV-1 LTR (p=0.0159) (**Figure 2.11C, D**). Thus, our pan-H4Ac and BRD4 CUT&Tag results suggest that in the combination condition, the R and U5 region of the HIV-1 LTR containing the HIV-1 transcription start site has a reduction of inhibitory signals that likely contributes to the reactivation of the latent HIV-1 provirus.

Next, we examined the consequences of the decreased pan-H4Ac and BRD4 levels at the HIV-1 LTR by interrogating markers for transcription initiation and elongation as mediated by the phosphorylation of RNA Polymerase II C-terminal domain of Rpb1 (RNA-Pol2-S5p and RNA-Pol2-S2p, respectively) [177] at the HIV-1 LTR, body of the provirus ("provirus body"), and the region just downstream of the HIV-1 provirus integration site ("downstream"). Again, as

quality control, we found the CUT&Tag signal of the RNA-Pol2-S5p replicates were most highly correlated with one another and were also correlated with the CUT&Tag signal of the RNA-Pol2-S2p replicates (**Figure 2.10**). We find that there is a striking, significant increase in RNA-Pol2-S5p signal at the HIV-1 LTR (p=4.2827E-5), provirus body (p=8.943E-5), and downstream (p=2.03E-5) only in the condition where cells that are knocked out for *ING3* and additionally treated with AZD5582. In contrast, for the RNA-Pol2-S2p signal in cells with a combination of *ING3* knockout and AZD5582 treatment, we observed a significant increase in the RNA-Pol2-S2p signal in the provirus body (p=0.0458) and downstream region (p=3.89E-5), but not over the LTR (**Figure 2.11E, F**). Interestingly, in the intermediate condition of only *ING3* knockout, the signal of RNA-Pol2-S2p is significantly reduced at the HIV-1 LTR (p=0.00583) (**Figure 2.11G, H**). This reduction in RNA-Pol2-S2p signal might be partially explained by the concomitant reduction we see in pan-H4Ac signal at the HIV-1 LTR in the *ING3* knockout condition alone. The striking increase in HIV-1 reverse transcriptase activity in the combination condition is concordant with our CUT&Tag results and suggest that the knockout of *ING3* combines with stimulation of the non-canonical NFκB pathway by AZD5582 to promote a potent increase in the RNA-Pol2-S5p and RNA-Pol2-S2p on the HIV-1 proviral genome.

As we observed the most significant RNA-Pol2-S5p level changes in the combined *ING3* knockout and AZD5582 treatment condition, we wondered whether increased RNA-Pol2-S5p levels was occurring genome-wide or if this was specific to the HIV-1 provirus. We compared the RNA-Pol2-S5p signal in the combined *ING3* knockout cells treated with AZD5582 to the control (NTC + DMSO) condition. We found the RNA-Pol2-S5p peaks over the HIV-1 LTR, provirus body, and downstream region were amongst the peaks with the most significant increase

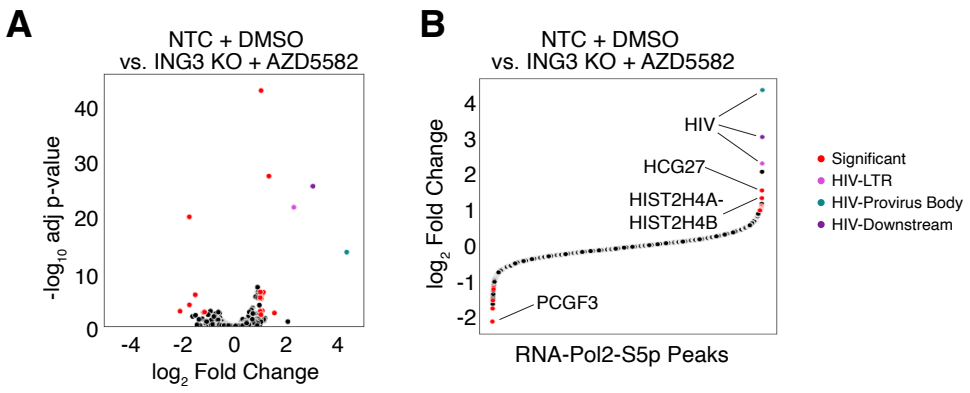


Figure 2.12. Changes in RNA-Pol2-S5p levels upon *ING3* knockout combined with AZD5582 treatment are nearly unique to the HIV-1 provirus.

(A) Volcano plot comparing RNA-Pol2-S5p peaks between the NTC KO + DMSO (negative control) and ING3 KO + AZD5582 conditions. Highlighted peaks over the HIV-1 LTR (magenta circle), the body of the provirus (purple circle), and downstream region of HIV-1 provirus (teal circle) are the regions of the highest RNA-Pol2-S5p fold change and are highly significant. Other regions with an absolute log2 fold change greater than 1 and a -log10 adjusted p-value > 2 are highlighted in red. **(B)** Scatter plot showing all of the RNA-Pol2-S5p peaks rank ordered by fold change. Peaks are colored as in (A). The top three regions are the HIV-1 LTR, body of the HIV-1 provirus, and downstream region of the HIV-1 provirus. The next two regions overlap HCG27 and the histone cluster 2 spanning HIST2H4A-HIST2H4B. The region with the greatest reduction in RNA-Pol2-S5p signal overlaps PCGF3. **Figure 2.12** was reproduced from [4] with permission, Creative Commons Attribution-Noncommerical-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

in RNA-Pol2-S5p signal (**Figure 2.12A**). We then rank ordered all the RNA-Pol2-S5p peaks by Log_2-fold change and found that out of the 2,547 total RNA-Pol2-S5p peaks, the top three peaks arranged from highest to lowest fold change were the HIV-1 provirus body, the region downstream of the HIV-1 integration site, and the HIV-1 LTR (**Figure 2.12B**). The two RNA-Pol2-S5p peaks with a significant increase in signal and the next highest fold change cover HCG27, a long non-coding RNA, and HIST2H4A – HIST2H4B, part of histone cluster 2 in the human genome. The RNA-Pol2-S5p peak with the strongest reduction in RNA-Pol2-S5p signal in *ING3* knockout cells treated with AZD5582 covers PCGF3, a part of the Polycomb group PRC1-like complex (**Figure 2.12B**). Overall, this suggests the increase in transcription of the HIV-1 provirus caused by *ING3* knockout in combination with AZD5582 treatment, as

correlated to the RNA-Pol2-S5p mark, is highly specific and nearly unique in the human genome. Thus, the CUT&Tag data reveal that *ING3* knockout results in a reduction in H4Ac and BRD4 levels over the HIV-1 LTR and that *ING3* knockout acts in combination with AZD5582 treatment to dramatically increase the transcriptional initiation and transcription elongation of the HIV-1 provirus leading to latency reversal.

2.4 Discussion

The hypothesis underlying this work is that modulating multiple overlapping mechanisms that control HIV-1 LTR-driven transcription will increase the potency and specificity of HIV-1 latency reactivation. That is, the goal is to find pathways that target potent transcriptional activation of latent HIV-1 proviruses with minimal global effects. We designed and validated a modular CRISPR screening approach which uses the incorporation of HIV-CRISPR genomes encoding sgRNAs into budding virions as a direct readout of activation of the latent provirus. The Latency HIV-CRISPR screen uses a sgRNA library of epigenetic regulatory genes and is paired with and without an LRA, AZD5582, to identify HIV-1 latency factors that act independently and in combination with AZD5582. We identified *ING3* as a gene whose knockout resulted in an enhanced increase of viral reactivation only in the presence of AZD5582. Using automated CUT&Tag, we observed that this enhancement is associated with active transcription marks as demonstrated by increased levels of RNA-Pol2-Ser5p and RNA-Pol2-Ser2p at the HIV-1 genome and mechanistically may be dependent on a reduction in histone H4 acetylation and BRD4 occupancy over the HIV-1 LTR. We also identified CUL3 and multiple members of the NuA4 HAT complex as important for HIV-1 latency maintenance.

2.4.1 Combinations of pathways with specificity for activating HIV-1 transcription

We find that two independent mechanisms - *ING3* knockout and low-activating dose of AZD5582 – which when combined, have unique, specific effects on the HIV-1 LTR and function together in order to efficiently activate HIV-1 transcription by increasing the presence of initiated and elongating RNA-Pol2 at the LTR and within the body of the HIV-1 provirus (**Figure 2.11**). Under these conditions used in our study, the RNA-Pol2-S5p peaks with the highest fold change throughout the entire genome were at the HIV-1 LTR and the provirus body (**Figure 2.12**). By CUT&Tag, RNA-Pol2-S5p appears as the primary marker for HIV-1 transcription in the J-Lat 10.6 cell line and this may be a result of P-TEFb, the major elongation factor for HIV-1 transcription, is a CTD kinase that can phosphorylate Ser5 and Ser2 individually, but not simultaneously [178]. Our model suggests that a reduction in NuA4 acetylation and BRD4-dependent occupancy specifically primes the HIV-1 LTR for enhanced transcription in response to *ING3* knockout and activation of the non-canonical NFκB pathway.

Our observed effect of a reduction in histone acetylation and BRD4 recruitment to the HIV-1 LTR that is concomitant with enhanced RNA-Pol2 initiation and elongation suggests the HIV-1 provirus may be subject to unexpected transcriptional control. In most mammalian transcriptional units driven by RNA-Pol2, NuA4 HAT is predominantly localized to promoters of active genes, specifically around the transcription start sites [179-181]. NuA4 HAT preferentially acetylates histone H4 [182] and this activity can impact gene expression regulation to often result in activation of transcription. Moreover, recruitment of other HAT complexes such as p300/CBP and P/CAF to the HIV-1 LTR followed by subsequent acetylation is associated with stimulation, rather than silencing, of HIV-1 transcription [118, 119, 121-123, 183, 184]. Additionally, using *in vitro* nucleosome-assembled templates in a Tat-independent

system, the NuA4 HAT activates HIV-1 transcription [185]. Conversely, in the presence of Tat, the NuA4 HAT complex has been implicated in silencing of HIV-1 transcription. For example, in both a primary CD4+ T cell model and cells from ART-treated people living with HIV-1, the absence of KAT5, the catalytic component of the NuA4 HAT complex, promotes activation of HIV-1 transcription as a result of reduced H4 acetylation and BRD4 long isoform occupancy, and increased levels of the Super Elongation Complex (SEC) members at the HIV-1 LTR [147]. Our findings are consistent with this observation as our Latency HIV-CRISPR screen identified multiple members of the NuA4 HAT complex, including KAT5, as top candidates and all the candidates were validated as latency maintenance factors through individual knockouts. In fact, our screen suggests that not only the catalytic subunit, but the NuA4 HAT complex altogether regulates HIV-1 transcription negatively.

We observed a reduction in H4 acetylation across the entire HIV-1 LTR in response to *ING3* knockout, but in the presence of AZD5582 we see the most significant reduction of H4Ac signal and BRD4 occupancy at the R and U5 region which includes the HIV-1 transcription start site. This observation suggests, the increase in accessibility of the HIV-1 transcription start site when *ING3* knockout combined with AZD5582 treatment likely promotes both transcription initiation and elongation, culminating in the enhanced viral reactivation we observe in both J-Lat cells and primary CD4+ T cell HIV-1 latency models. Additionally, the dosage of AZD5582 used was a low-reactivating dose, suggesting that by inhibiting the NuA4 complex in parallel, it may be possible to efficiently activate HIV-1 transcription while causing minimal disruptions to a host system. HIV-1 is subject to complex transcriptional regulation, and these results highlight the importance of simultaneously examining multiple mechanisms to identify opportunities to specifically stimulate transcription of latent HIV-1 proviruses. These results are similar in

concept to a recent study that found that the combination of AZD5582 and I-BET151 uniquely activated HIV-1 transcription in a Jurkat model of latency [111].

In support of the hypothesis that these mechanisms are nearly unique to the HIV-1 LTR, relatively few loci in the host genome showed a significant increase in RNA-Pol2-S5p occupancy when *ING3* knockout is combined with a low-activating dose treatment of AZD5582. One of these regions is downstream of the HIV-1 provirus integration site in J-Lat 10.6 cells and part of the *SEC16A* gene (**Figure 2.12**). Activation of human genes near or at the integrated provirus due to active transcription from an integrated provirus is a likely scenario as HIV-1 preferentially integrates into highly transcribed genes [84], and would be an unavoidable consequence of any reactivation strategy used for eradicating the latent HIV-1 reservoir. In the combined condition of *ING3* knockout and AZD5582 treatment, two other loci, a region of histone cluster 2 spanning HIST2H4A – HIST2H4B and a region overlapping with the long-noncoding RNA, HCG27, also showed statistically significant, large fold changes in RNA-Pol2-S5p levels, but these changes were not as prominent as those observed over the HIV-1 provirus (**Figure 2.12**). In addition, we observed the gene PCGF3 had significantly reduced RNA-Pol2-S5p occupancy which could potentially to contribute to latency reversal as it is part of the non-canonical Polycomb group RING finger 3/5-PRC1 complex with silencing function [186] and is involved in the recruitment of PRC1 and PRC2 which have been shown to contribute to HIV-1 proviral silencing [131, 146, 187].

2.4.2 Comparison to other HIV-1 latency screens

Our Latency HIV-CRISPR screening strategy is a unique approach to identify factors involved in HIV-1 latency. However, previous screens have also explored HIV-1 latency factors

using gene knockout, knockdown, or silencing screening [105, 131, 138, 150-155]. Generally, these previous screens used whole genome sgRNA libraries which make them unbiased for specific pathways, but this also limits the strength in statistical power to determine gene hits. Nonetheless, these previous screens have uncovered an exciting breadth of HIV-1 latency factors from the mTOR pathway [150] to proteasome related genes [153, 155] to the non-canonical NFκB pathway [105]. As epigenetics plays an important role in HIV-1 latency and it is still under investigation, we focused our Latency HIV-CRISPR screen for epigenetic regulatory factors by using a small, custom designed sgRNA library. Additionally, a smaller library can increase the dynamic range of a screen [188]. Our screen also relies on the HIV-CRISPR vector [156] to perform a high-throughput screen, which uses HIV-1 replication, rather than another reporter, as the direct readout of the screen. The difference in readout from other screens including screens with epigenetic hits [131, 138, 152] can result in different gene hits being highlighted. Another challenge of studying mechanisms of HIV-1 latency is the interaction of complex mechanisms in parallel and to address this using a CRISPR screen approach, we combined the screen with a low-activating dose of LRA treatment. By focusing our sgRNA library on epigenetic regulators and performing our screen with and without AZD5582 treatment, we identified a unique combination of modifications that synergize to stimulate HIV-1 transcriptional initiation and elongation. In the future, by exploring the combination of additional LRAs with Latency HIV-CRISPR based screens, we will improve the specificity and potency of novel LRAs and continue to work towards a curative treatment strategy that eradicates the HIV-1 latent pool.

2.5 Methods

2.5.1 J-Lat wildtype, clonal knockout, and pooled knockout cells

The HIV-1 latency cell line J-Lat 5A8 [62] and J-Lat 10.6 [61] were grown in RPMI-1640 medium (Thermo Fisher Scientific) supplemented with 10% Fetal Bovine Serum (FBS), Penicillin-Streptomycin (Pen/Strep), and 10 mM HEPES. In order to use these cells for screening with HIV-CRISPR, we performed CRISPR/Cas9-mediated knockout of Zinc Antiviral Protein (ZAP) [156] by electroporation of the J-Lat 10.6 and 5A8 cells with a Gene Knockout v2 kit (GKOv2) for ZAP (Synthego, Redwood City, CA) complexed with 1 μL of 20 μM Cas9-NLS (UC Berkeley Macro Lab). 5 days post electroporation, the cells were single cell sorted into a 96-well U-bottom plate (Sony MA900 Multi-Application Cell Sorter – Fred Hutch Flow Cytometry shared resource) and individual clones with biallelic knockouts of ZAP were used for subsequent Latency HIV-CRISPR screens. J-Lat cells with HIV-CRISPR KO pools targeting NFKBIA, KAT5, CUL3, ACTL6A, VPS72, DNMT1, DMAP1, SRCAP, YEATS4, and two non-targeting controls (**Supplementary File 1**, see https://www.biorxiv.org/content/10.1101/2022.08.23.504195v1.supplementary-material) were generated by transduction of lentivirus and subsequent 0.4 μg/mL puromycin selection for 10-14 days. To generate the CRISPR/Cas9-edited knockout pools, the sgRNAs that were the most overrepresented in the viral supernatant (highest combined p-value) of the Latency HIV-CRISPR screen were selected. Cell lines were determined to be mycoplasma free by the Fred Hutch Specimen Processing/Research Cell Bank shared resource.

2.5.2 Plasmids

HIV-CRISPR plasmid was previously described [156]. HIV-CRISPR constructs targeting genes of interest were cloned by annealing complementary oligos (**Supplementary File 1**, see https://www.biorxiv.org/content/10.1101/2022.08.23.504195v1.supplementary-material) with overhangs that allow directional cloning into HIV-CRISPR using the BsmBI restriction sites. pMD2.G and psPAX2 plasmids were gifts from Didier Trono (Addgene #12259 and #12260, respectively). pMD2.Cocal plasmid was a gift from Hans-Peter Kiem [189]. lentiCRISPRv2 plasmid was a gift from Feng Zhang (Addgene #52961). The pNL4-3-Δ6-dreGFP-CD90 was previously described [173].

2.5.3 Human Epigenome CRISPR/Cas9 sgRNA library construction

The Human Epigenome (HuEpi) sgRNA library is composed of 841 genes of which 778 genes derive from the database Epifactor [158] and 63 genes were hand selected. For most genes, six sgRNA sequences were generated using GUIDES (Graphical User Interface for DNA Editing Screens) [190]. Two genes (HEATR1 and MCM2) have fourteen guides and one gene (UBE2N) has one guide. 252 non-targeting sgRNAs sourced from GeCKO v2.0 library [159] were also included resulting in a generated library of a total of 5,309 guides. The HuEpi sgRNA was synthesized (Twist Biosciences, San Francisco, CA) and cloned into HIV-CRISPR. Oligo pools were amplified using Phusion High-Fidelity DNA Polymerase (NEB) combined with 1 ng of pooled oligo template, primers ArrayF and ArrayR (ArrayF primer: TAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG and ArrayR primer: ACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCT

AGCTCTAAAAC), an annealing temperature of 59°C, an extension time of 20 s, and 25 cycles. Following PCR amplification, a 140 bp amplicon was gel-purified and cloned into BsmBI (NEB; R0580) digested HIVCRISPR using Gibson Assembly (NEB; E2611S). Each Gibson reaction was carried out at 50°C for 60 min. Drop dialysis was performed on each Gibson reaction according to the manufacturer's protocol using a Type-VS Millipore membrane (VSWP 02500). 5 µl of the reaction was used to transform 25 µl of Endura electrocompetent cells (Lucigen; 60242-2) according to the manufacturer's protocol using a Gene Pulser (BioRad). To ensure adequate representation, sufficient parallel transformations were performed and plated onto carbenicillin containing LB agarose 245 mm x 245 mm plates (Thermo Fisher) at 300-times the total number of oligos of each library pool. After overnight growth at 37°C, colonies were scraped off, pelleted, and used for plasmid DNA preps using the Endotoxin-Free Nucleobond Plasmid Midiprep kit (Takara Bio; 740422.10). The HuEpi library was sequenced and contains all 5,309 sgRNAs included in the synthesis (GSE215430).

2.5.4 Lentivirus production

293T cells (ATCC;CRL-3216) cultured in DMEM (Thermo Fisher Scientific) supplemented with 10% FBS and PenStrep were plated at 2E5 cells/mL in 2 mL in 6-well plates one day prior to transfection. Transfection is performed using TransIT-LT1 reagent (Mirus Bio LLC; MIR2300) with 3 uL of transfection reagent per µg of DNA. For lentiviral preps, 293Ts were transfected with 667 ng lentiviral plasmid, 500 ng psPAX2, and 333 ng pMD2.G. One day post-transfection, media was replaced. Two- or three- days post-transfection, viral supernatants were filtered through a 0.2 µm filter (Thermo Scientific; 720-1320). For HuEpi library lentiviral preps, the same transfection of 293Ts was performed and supernatants from forty 6-well plates

were combined and concentrated by ultracentrifugation. About 30 mL of supernatant are aliquoted into a polypropylene tube (Beckman Coulter; 326823) and underlaid with sterile-filtered 20% sucrose (20% sucrose, 1 mM EDTA, 20 mM HEPES, 100 mM NaCl, distilled water). Each of the polypropylene tubes are placed in a swinging bucket and spun in a SW 28 rotor at 23,000 rpm for 1 hour at 4°C in a Beckman Coulter Optima L-90K Ultracentrifuge. Supernatants were decanted and pellets are resuspended in DMEM or RPMI over several hours at 4°C. Concentrated lentivirus was used immediately or aliquots were made and stored at -80°C. All lentiviral transductions were performed in the presence of 20 μg/mL DEAE-Dextran (Sigma-Alrdich; D9885). For generating pNL4-3-Δ6-dreGFP-CD90 stocks, the procedure was the same as the lentiviral preps except that 293Ts were transfected with 900 ng lentiviral plasmid, 450 ng psPAX2, and 150 ng pMD2.Cocal.

2.5.5 LRA and no LRA Latency HIV-CRISPR screening

The HuEpi library lentiviral preps were titered by a colony-forming assay in TZM-bl cells (NIH AIDS Reagent Program; ARP-8129) and used to transduce J-Lat 10.6 and J-Lat 5A8 cells at an MOI of 0.4. For the transduction, 3E6 cells per replicate (>500x coverage of the HuEpi library) was used and the spinoculation was performed at 1100xg for 30 min in the presence of 20 μg/mL DEAE-Dextran. Cells were selected in puromycin (0.4 μg/mL) for 10-14 days. In preparation for the LRA screen, AZD5582 dihydrochloride (Tocris; 5141) was resuspended in DMSO to 1 mM stocks. Subsequent dilutions are performed in RPMI and used immediately. Upon completion of selection, for the LRA screen, 3E6 cells per replicate per cell line were treated with 10 nM AZD5582 for 24 hours. Cells and supernatants were collected post selection or post treatment. Cells were washed once with DPBS (Gibco; 14190144) and cell

pellets were stored at -20°C. Genomic DNA was extracted from cell pellets with the QIAamp DNA Blood Midi Kit (Qiagen; 51183) and genomic DNA was eluted in distilled water. Viral supernatants were spun at 1100xg to remove cell debris, filtered through a 0.22 μm filter (Millipore Sigma, SE1M179M6), overlaid on a 20% sucrose cushion, and concentrated in a SW 28 rotor for 1 hour at 4°C. The pellet is resuspended in RPMI and stored at -80°C. Viral RNA was extracted from the concentrated virus with the QIAamp Viral RNA Mini Kit (Qiagen, 52904). The sgRNA sequences found in the genomic DNA and viral RNA samples were amplified by PCR (Agilent; 600677) and RT-PCR (Invitrogen; 18064014), respectively, using HIV-CRISPR specific primers. A second round of PCR is performed to barcode and prepare the libraries for Illumina sequencing (**Supplementary File 1**, see https://www.biorxiv.org/content/10.1101/2022.08.23.504195v1.supplementary-material). Each amplicon was cleaned up using double-sided bead clean-up (Beckman Coulter; A63880), quantified with a Qubit dsDNA HS Assay Kit (Invitrogen; Q32854), and pooled to 10 nM for each library. Library pools are sequenced on a single lane of an Illumina HiSeq 2500 in Rapid Run mode (Fred Hutch Genomics and Bioinformatics shared resource).

2.5.6 Screen analysis

Raw sequencing data is available as a GEO DataSet (GSE215430). Library pools are demultiplexed, reads are assigned to respective samples, trimmed, and aligned to the HuEpi library via Bowtie [191]. An artificial NTC sgRNA gene set the same size as the HuEpi library was generated by iteratively binning the NTC sgRNA sequences. Analysis of the screen to determine relative enrichment or depletion of the sgRNAs and genes were performed using the MAGeCK statistical package [160].

2.5.7 Western blotting

Cells were harvested and washed once with ice cold phosphate-buffered saline (PBS) and pelleted by centrifugation (300 x g for 3 min). Whole cell lysate was extracted on ice by first resuspending the cell pellet with Pierce IP Lysis Buffer (ThermoFisher Scientific; 87787) supplemented with cOmplete protease inhibitor cocktail (Roche; 11697498001). The samples were then incubated on ice for 10 min with brief vortexing every 2-3 min and followed by a centrifugation at 16,000 x g for 20 min at 4°C. To prepare the samples, add 4x NuPAGE LDS Sample Buffer (ThermoFisher Scientific; NP0008) containing 5% 2-Mercaptoethanol (Millipore Sigma; M3148) and boil the samples at 95°C for 5 min. Lysates were resolved on a NuPAGE 4-12% Bis-Tris pre-cast gel (ThermoFisher Scientific; NP0336) and transferred to nitrocellulose membranes (Biorad; 1620115). Blocking was performed for 1 hr at room temperature using 5% milk/0.1% Tween-20 added to Tris-buffered saline (TBS) (20 mM Tris base and 150 mM NaCl at pH 7.6). Immunoblotting was performed using the primary antibodies CUL3 (Cell Signaling; 2759) at 1:1000, NFkB2 (Cell Signaling; 4882) at 1:1000, and viniculin (Santa Cruz; sc-25336) at 1:5000. Membranes were washed with TBST 6 times for 5 min each. The following secondary antibodies were used at a 1:5000 dilution: goat anti-rabbit IgG-HRP (R&D Systems; HAF008) and goat anti-mouse IgG-HRP (R&D Systems; HAF007). Membranes were developed with SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher; 34095) and visualized on a BioRad Chemidoc MP Imaging System.

2.5.8 Genomic editing analysis

Knockout cells were harvested and DNA was extracted using the QIAamp DNA Blood Mini Kit (Qiagen; 51104). Primers from each targeted locus (Supplementary File 1) were used to

amplify the edited loci and the PCR was performed with either Platinum Taq DNA Polymerase High Fidelity (ThermoFisher Scientific; 11304011) or Q5 High-Fidelity DNA Polymerase (NEB; M0491S). Sanger sequencing was performed on PCR amplicons (Fred Hutch Genomics shared resource) using a sequencing primer (Supplementary File 1) and results were analyzed by Inference of CRISPR Edits (ICE) [168] to determine gene editing outcome.

2.5.9 Virus release assay (RT assay)

Clarified viral supernatants are harvested and reverse transcriptase activity was measured using the HIV-1 reverse transcriptase (RT) activity assay as previously described [192, 193]. A standard curve is generated for all assays using a titered stock of HIV-1_{LAI} that was aliquoted at and stored at -80°C.

2.5.10 Ethics statement

All primary cell data is from anonymous blood donors and is classified as "human subjects exempt" research by the Fred Hutchinson Cancer Center Institutional Review Boards, according to National Institutes of Health (NIH) guidelines (http://grants.nih.gov/grants/policy/hs/faqs_aps_definitions.htm). No animal work was done.

2.5.11 Primary CD4+ T cell latency model and knockout

Used leukocyte filters from healthy donors were obtained from Bloodworks Northwest and total peripheral blood mononuclear cells (PBMCs) were isolated by purification over Ficoll (Millipore Sigma; GE17-1440-02). CD4+ T cells were isolated using magnetic negative selection (StemCell Technologies; 17952) and used immediately or frozen for storage in liquid nitrogen.

CD4+ T cells were activated using anti-CD2, CD3, and CD28 beads (Miltenyi Biotec; 130-091-441) for 2 days and grown in RPMI-1640 medium (Thermo Fisher Scientific) supplemented with 10% Fetal Bovine Serum (FBS), Penicillin-Streptomycin (Pen/Strep), 1x GlutaMAX (ThermoFisher Scientific; 35050061), 10 mM HEPES, 100 U/mL human interleukin-2 (Millipore Sigma; 11011456001), 2 ng/mL human interleukin-7 (Peptrotech; 200-07), and 2 ng/mL human interleukin-15 (Peprotech; 200-15). The beads were then removed by magnetic separation and cells were spinoculated with HIV-GFP-Thy1.2 (pNL4-3-Δ6-dreGFP-CD90) and 8 ug/mL polybrene (Millipore Sigma; TR-1003-G) for 2 hrs at 1100 x g. Spinoculated cells were then resuspended in fresh media with IL-2, IL-7, and IL-15 and incubated for 3 days before actively infected cells (Thy1.2+) were isolated by magnetic positive selection (Stem Cell Technologies; 18951). Infected cells (Thy1.2+) are maintained for 2 days before knockout of AAVS1, CUL3, and ING3 using GKOv2 kits from Synthego. crRNPs for each gene of interest were generated as described previously in the Knockout Cell Clones and Pools section except that supplemented P3 buffer (Lonza; V4SP-3096) is used instead of SE buffer. For each electroporation, 1.5E6 cells were pelleted by centrifugation at 100 x g for 10 min at 25°C, washed once with PBS, pelleted again by centrifugation, PBS was removed, and resuspended with a crRNP complex. The resuspended cells were immediately transferred into the cuvette of the P3 Primary Cell Nucleofector Kit (Lonza; V4SP-3096) and electroporated using code EH-100 on the Lonza 4D-Nucleofector. 80 µL of prewarmed supplemented RPMI media with IL-2, IL-7, and IL-15 was added and cells were allowed to recover for 10 min in the 37°C incubator. 300 µL of fresh media was added and cells are transferred to a 96-well plate. 200 µL of additional supplemented media was added 2 days later. The knockout cells are maintained at 1E6 cells/mL with fresh media supplemented with IL-2, IL-7, and IL-15. At 14 days post infection,

cells are co-cultured with H80 feeder cells and maintained at 2E6 cells/mL in media supplemented with 20 U/mL IL-2. If AZD5582 treatment applies, cells are treated 1 uM AZD5582 media for 24 hours. At 19 days post infection, cells were stained with Thy1.2 antibody (Biolegend; 140323), fixed with 4% PFA, and underwent flow cytometry analysis (CD FACS Celesta – Fred Hutch Flow Cytometry shared resource).

2.5.12 Automated CUT&Tag profiling

We prepared nuclei from J-Lat 10.6 cells under four conditions: (i) NTC knockout with a treatment of DMSO (negative control) (ii) NTC knockout with a treatment of 10 nM AZD5582 (iii) ING3 knockout with a treatment of DMSO (iv) ING3 knockout with a treatment of 10 nM AZD5582. Up to 10 million cells were pelleted in 1.5 mL microfuge tubes spun at 300 x g for 10 min. Cells were then resuspended in 1 mL of ice cold NE1 Buffer (20 mM HEPES-KOH pH 7.9, 10 mM KCl, 0.5 mM Spermidine, 0.1% TritonX-100, 20% Glycerol, with Roche complete EDTA-free protease inhibitor tablet), and incubated on ice for 10 min. Nuclei were then centrifuged at 4°C at 1,300 x g for 4 min. Nuclei were then resuspended in 1 mL of Wash Buffer (20 mM HEPES pH7.5, 150 mM NaCl, 0.5 mM spermidine, supplemented with Roche complete EDTA-free protease inhibitor tablet). 10 μL was used to determine the concentration of native J-Lat 10.6 nuclei, and nuclei were diluted to a concentration of 1 million nuclei/900 μL of Wash Buffer and 900 μL aliquots of this suspension was mixed with 100 μL of DMSO in Cryovials, which were then sealed and placed inside a Mr. Frosty isopropanol chamber for slow freezing at -80°C. Nuclei were then stored at -80°C until use. For automated CUT&Tag processing, nuclei were thawed at room temperature, washed in wash buffer, and bound to concanvalin-A (ConA) paramagnetic beads (Bangs Laboratories; BP531) for magnetic separation as described on the

protocols.io website (https://doi.org/10.17504/protocols.io.bgztjx6n).Samples were then suspended in antibody binding buffer and split for overnight incubation with antibodies specific to panH4Ac (Active Motif; 39925), BRD4 (Cell Signaling; 13440), RNA-Pol2-S5p (Cell Signaling; 13523), RNA-Pol2-S2p (Cell Signaling; 13499), and IgG control(Abcam; 172730). Sample processing was performed in a 96 well plate using 100K Con-A bound nuclei per reaction on a Beckman Coulter Biomek liquid handling robot according to the AutoCUT&Tag protocol available from the protocols.io website (https://doi.org/10.17504/protocols.io.bgztjx6n) and described previously [157] (Fred Hutch Genomics shared resource).

2.5.13 CUT&Tag sequencing data processing and analysis

For AutoCUT&Tag sample pooling and sequencing, the size distribution and molar concentration of libraries were determined using an Agilent 4200 TapeStation, and up to 96 barcoded CUT&Tag libraries were pooled at approximately equimolar concentration for sequencing. Paired-end 2 × 50 bp sequencing was performed on the NextSeq 2000 platform by the Fred Hutchinson Cancer Research Center Genomics Shared Resources. This yielded 5–10 million reads per antibody. Sequences the extended into the 3' adapter were first removed using the adapter clipping tool by cutadapt 2.9 with parameters:

-j 8 --nextseq-trim 20 -m 20 -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA -A AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT -Z.

To construct a reference reflecting integration of HIV-1 into J-Lat cells, we started with the UCSC hg38 human reference sequence from the Illumina iGenomes collection (https://support.illumina.com/sequencing/sequencing_software/igenome.html). Sequence for the integrated HIV-1 genome, including flanking human sequence, was obtained from accession

MN989412.1 [194]. Accounting for flanking human sequence, we excised chr9:136468439-136468594 from the reference Fasta file and inserted MN989412.1 in its place. We also rewrote gene & exon annotations downstream of the integration site, shifting all features in the associated GTF file by 10206bp to account for newly inserted HIV-1 sequence. Indexes for Bowtie2 v2.4.1 [195] and STAR v2.7.7a [196] were built using these modified Fasta and GTF files. The HIV-1 insertion is bounded by a pair of LTRs that, because they comprise identical sequence, are difficult for some alignment programs to interpret. To avoid these difficulties when needed, another version of the reference sequence was prepared with the second copy of the LTR sequence (634bp) masked by 'N's.

Fastq files were aligned to the custom hg38 genome with the HIV-1 genome inserted in chromosome 9 using Bowtie2 version 2.4.2 with the following parameters: --very-sensitive-local --soft-clipped-unmapped-tlen --dovetail --no-mixed --no-discordant -q --phred33 -I 10 -X 1000. Raw sequencing data is available as a GEO DataSet (GSE215430). Because the read depth and fraction of PCR duplicates varied considerably between replicates, we removed all duplicate reads from the Bed files. Peak calling was performed for the BRD4, pan-H4ac, RNA-Pol2-S5p, and RNA-Pol2-S2p data sets on the pooled reads from all replicates and all conditions using SEACR version 1.3 [197]. We called peaks using two settings, (1) in which using the IgG control data with stringent settings, and (2) using a FDR of 0.01, and then used the FDR 0.01 peaks that overlapped with IgG peaks as our final peak set. For correlation analysis between replicates, the BRD4 peak set was merged with the pan-H4ac peak set, and the RNA-Pol2-S5p peak set was merged with the RNA-Pol2-S2p peak set. For statistical comparisons presented in Figure 7, we quantified the base pair coverage of pan-H4ac or BRD4 over the LTR, the U3 region, or the R + U5 region and these values were normalized to the total base pair coverage

across the genome for each replicate. Similarly, for comparisons of the RNA-Pol2-S5p data and RNA-Pol2-S2p data, we quantified the base pair coverage over the LTR, Provirus Body, and Downstream Region of the HIV provirus these values were normalized to the total base pair coverage across the genome for each replicate. For comparisons in Figure 7, p-values were calculated using a two-sample t tests (two sided) with the SciPy.stats.ttest_ind() function in Python; P values were not corrected for multiple-hypothesis testing. For global comparison of RNA-Pol2-S5p data between the NTC + DMSO condition and the ING KO + AZD5582 condition, presented in Figure 8, reads that overlapped RNA-Pol2-S5p peaks were counted for each replicate and the log 2 Fold Change and adjusted p-values were calculated using DESeq2 version 1.32.0 using the Wald test. The adjusted p-values are corrected for multiple-hypothesis testing in a manner that is proportional to the number of RNA-Pol2-S5p peaks. DESeq2 assigns peaks with extremely sparse data are assigned an adjusted p-value of NaN, and these peaks were excluded form downstream analysis.

Chapter 3. Conclusions and Future Directions

Transcriptional initiation and elongation, and epigenetic regulation are all molecular mechanisms that simultaneously contribute to establishing and maintaining HIV-1 latency and importantly, the goal is to identify the interactions that are specific to HIV-1 transcription rather than general host transcription. In Chapter 2, I described a CRISPR-based screening approach called Latency HIV-CRISPR that was used as a key tool to expand our knowledge of the interplay amongst the molecular mechanisms at work in the context of HIV-1 latency. Unique to this screening strategy is the use of HIV-1 replication as the readout instead of a reporter. I then demonstrated that the screening strategy is able to identify novel interactions maintaining HIV-1 latency in Jurkat and primary CD4+ T cell models of latency – specifically between the inhibition of *ING3* and AZD5582. The CUT&Tag analysis has demonstrated that this interaction is likely acetylation- and BRD4-dependent and the transcriptional activation is nearly unique to the HIV-1 LTR. The results of the Latency HIV-CRISPR screen provide many avenues of future directions for extensions of this work.

3.1 Further exploration of *ING3* inhibition combined with AZD5582 treatment as an effective mode of HIV-1 latency reversal

The identification of *ING3* as a latency maintenance factor in combination with AZD5582 is intriguing as *ING3* encodes a protein that is a subunit of the NuA4 HAT complex. In future work, it will be important to determine if the role of ING3 in latency reversal is NuA4 HAT-independent or NuA4 HAT-dependent manner. My data demonstrating that *ING3* knockout alone changes histone H4 acetylation at the HIV-1 LTR (**Figure 2.11A, B**) suggests that ING3 is acting as part of an acetyltransferase complex, but it is possible that ING3 is

functioning through another complex, or that the effect of *ING3* knockout is more indirect. If ING3 does act as part of the NuA4 complex in latency reversal in combination with AZD5582, then this information would lend itself to further exploration to see if targeting the catalytic subunit of the NuA4 HAT complex, *KAT5*, could itself be targeted for a combination latency reversal strategy. Indeed, I did find that knockout of *KAT5* by CRISPR/Cas9-mediated knockout does lead to latency reversal (**Figure 2.4A**), but I did not test the combination of *KAT5* knockout in combination with AZD5582 to determine if the phenotype is similar or different from *ING3* inhibition in J-Lat cells and a primary CD4+ T cell model of HIV-1 latency [173]. Furthermore, there is a small molecule inhibitor of KAT5, MG-149, [147] that could also be explored as a latency reversal agent in combination with AZD5582.

If studies with a KAT5 inhibitor combined with other latency reversal agents do show promise in primary cell models of HIV-1 latency, then the next step would be to test these small molecules on CD4+ T cells isolated from ART-suppressed people living with HIV-1 in which the integrated provirus populations more closely model *in vivo* latency. If these experiments also recapitulate the original result, it would be intriguing to apply this drug combination on *in vivo* models of HIV-1 latency such as humanized BLT mice or in SHIV or SIV models of latency in rhesus macaques. In fact, AZD5582 has already been tested in both models [83] with minimal toxicity and marginal efficacy. Based on my results thus far, the efficacy of AZD5582 or related molecules that activate the non-canonical NFκB pathway could be augmented by inhibition of the pathways involved in the acetylation of the HIV-1 LTR. The pharmacodynamics and safety of the KAT5 inhibitor in humanized BLT mice and rhesus macaques will have to be determined; nonetheless, it will be interesting to determine if latency reversal can be detected upon treatment with the KAT5 inhibitor combined with AZD5582 in both animal models by measuring HIV-1

or SIV viral RNA levels in resting CD4+ T cells isolated from immune tissues such as bone marrow and lymph nodes. Additionally, it will be important to determine if there are any changes in cell-associated HIV-1 or SIV DNA levels between the combination drug treatment and a negative control as this would be indicative of a reduction in the latent viral reservoir, which would be an exciting step towards the ultimate goal of latent reservoir elimination. Furthermore, these results might also spur development of additional inhibitors of KAT5.

If the phenotype of AZD5582 treatment combined with KAT5 inhibition or ING3 inhibition differs, then it is likely that ING3 is playing a NuA4 HAT complex-independent role in the context of viral reactivation. *ING3* has been recently identified as an endogenous retrovirus transcriptional repressor mediating H3K27me3 modifications [198], which could be an alternative role *ING3* plays in the context of HIV-1 latency. As there currently is not a small molecule inhibitor of ING3, to address this hypothesis, one approach could be to generate a bivalent chemical protein degrader, also known as PROTACs, which uses the E3 ubiquitin ligase pathway to degrade an endogenous protein of interest [199]. PROTACs are being used in clinical trials to target cancers, including but not limited to, the estrogen receptor for breast cancer, so they already have the potential to be used as therapeutics in humans [200]. Additionally, PROTACs have been successfully generated for a wide variety of targets including epigenetic regulators such as the PRC2 complex and result in efficient degradation with high specificity [201]. Once a PROTAC is developed for ING3, latency reversal experiments can be performed to test the combination of ING3 PROTAC and AZD5582 in J-Lat cells, primary CD4+ T cell models of HIV-1 latency, isolated cells from ART-suppressed people living with HIV-1, and *in vivo* animal models, to determine the feasibility of these reagents as therapeutics to eliminate the latent reservoir.

3.2 Exploration of additional latency maintenance factors from Latency HIV-CRISPR screen

The Latency HIV-CRISPR screen produced several other HIV-1 latency maintenance genes of interest that I validated using two J-Lat models including *ACTL6A*, *SRCAP*, and *CUL3* (**Figure 2.4A**). Investigating the role in which each of these genes play a role in HIV-1 latency would be of great interest. Knockout of *ACTL6A* in both J-Lat 5A8 and 10.6 cells produced the highest levels of viral reactivation as measured by HIV-1 reverse transcriptase activity (**Figure 2.4A**). ACTL6A, also known as BAF53A, is a subunit of the NuA4 HAT complex, but is also a subunit of the BAF complex that has been shown to reposition nucleosome-1 resulting in lower DNA accessibility of the HIV-1 LTR which likely contributes to maintenance of a repressive state [139]. ACTL6A is also a subunit of the SRCAP complex that functions to exchange the histone H2A with histone variant H2A.Z [202] and the NuA4 HAT complex may enhance SCRAP activity [203]. Interestingly, my screen identified hits unique to the SRCAP complex, like SRCAP (**Figure 2.3D**). I hypothesize that the role of ACTL6A in two relevant complexes likely accounts for the high level of HIV-1 reactivation in response to *ACTL6A* knockout.

Furthermore, preliminary work from the Strahl Lab from UNC Chapel Hill used chromatin fractionation to demonstrate that *SRCAP* knockout J-Lat cell lines I generated have a decrease in modified H2A.Z. In this experiment, *SRCAP* knockout J-Lat cells underwent a round of fractionation followed by a round of digestion to generate a soluble and chromatin fraction. Western blotting was then performed for each of these samples to determine the levels of modified and unmodified H2A.Z in each fraction. In both the soluble and chromatin fraction, unmodified H2A.Z levels were similar for NTC sgRNA transduced and *SRCAP* knockout J-Lat cell lines. However, for the modified H2A.Z, which is potentially phosphorylated or

ubiquitinated, in the soluble fraction there was no signal for the negative control and *SRCAP* knockout J-Lat cell lines and in the chromatin fraction, the modified H2A.Z levels were reduced in the *SRCAP* knockout J-Lat cell line compared to the negative control. These results suggest that interplay of the SRCAP complex, H2A.Z variant post-translational modification, and histone modification by acetylation is critical for HIV-1 latency. Performing additional CUT&Tag experiments on these gene knockout cell lines would provide insight into the consequence of each gene function on the HIV-1 LTR. For example, for the *ACTL6A* knockout J-Lat cell lines, in addition to examining pan-H4Ac, BRD4, RNA-Pol2-S5p, and RNA-Pol2-S2p levels, as previously performed with the *ING3* knockout cell lines treated with AZD5582, investigation of H2A.Z and KAT5 levels at the HIV-1 LTR would be informative. Based on the known function of *ACTL6A* and the validation data I have generated (**Figure 2.4A**), I expect the viral reactivation to be H4Ac and BRD4-dependent. In other words, I expect a reduction of pan-H4Ac and BRD4 levels at the HIV-1 LTR compared to a negative control and increased RNA-Pol2-S5p and RNA-Pol2-S2p levels at the HIV-1 LTR and provirus body. Additionally, if ACTL6A is simultaneously playing a role in both the NuA4 HAT complex and SRCAP complex, I expect reduced levels of both H2A.Z and KAT5, and potentially a shift in signal localization for H2A.Z. I expect these effects to be most pronounced at the HIV-1 LTR, but given the expansive role of ACTL6A, it is possible that the described effects could also be observed at other host promoters. Through optimization of more efficient gene editing and fine-tuning to use less input material for CUT&Tag protocols, these experiments could be expanded to explore marks at the HIV-1 LTR in primary T cell models of HIV-1 latency in which the integration site is heterogenous.

3.3 Additional unpublished work on CUL3

The top hit from my Latency HIV-CRISPR screen in the absence of LRA treatment was *CUL3* which I also validated in a primary CD4+ T cell model of HIV-1 latency (**Figure 2.4B**). CUL3 has been demonstrated to negatively regulate HIV-1 transcription during productive infection through the NFκB pathway [175]. My data is consistent with this model in that when I compared gene hits from Latency HIV-CRISPR screens performed in the presence and absence of AZD5582, I observed that CUL3 was preferentially decreased as a gene hit in the presence of AZD5582, which suggests pathway redundancy. Alternatively, ubiquitination has been shown to be involved in many aspects of HIV-1 transcriptional elongation such as ubiquitination of HEXIM-1 to release P-TEFb from 7SK [204], ubiquitination of HIV-1 Tat protein by PJA2 to activate transcription [205], and ubiquitination of factors (HCF1/2) to stabilize ELL2, a subunit of the Super Elongation Complex that stimulates transcriptional elongation [206]. CUL3 could play a novel role in the ubiquitination of one or more of these substrates.

In an effort to narrow down the targeted substrate, I performed an additional set of Latency HIV-CRISPR screens in both J-Lat 10.6 and J-Lat 5A8 cells. Cullin proteins function as a scaffold and often through an adaptor protein, the scaffold is linked to a substrate through a substrate recognition protein [207]. CUL3 does not use an adaptor subunit and instead the substrate recognition protein incorporates a BTB (Bric-a-brac, Tramtrack, Broad-complex) fold that recognizes CUL3 and substrates [207]. In order to potentially define the substrate CUL3 is targeting in the context of HIV-1 latency, I designed a custom guideRNA library targeting a breadth of proteins with a BTB fold and were predicted to be CUL3 substrates using a dataset produced by Emanuele et al from a mass spectrometry-based proteomics screen and protein stability genetic screen [208]. Given the possibility that CUL3 could be targeting an unexpected

target in the context of HIV-1 latency, I also expanded the library to include other cullins, predicted adaptors of other cullins, and predicted substrates of other cullins, which are also based off the dataset generated by Emanuele et al [208]. Altogether, the cullin associated guideRNA library targets 679 genes with 8 guideRNAs targeting each gene and an additional 272 non-targeting controls for a total of 5,704 guideRNAs.

This cullin associated CRISPR guideRNA library was then used in the Latency HIV-CRISPR screen, which had to be adapted to examine both adaptors and substrates. In the context of the Latency HIV-CRISPR screen, a cullin adaptor is predicted to have a phenotype similar to a cullin such that a gene knockout would result in J-Lat viral reactivation and thereby enrichment of the HIV-CRISPR vector encoding the respective single guide RNA (sgRNA) in the viral supernatant. However, a cullin substrate, which I predict is likely to be targeted for degradation after ubiquitination, is considered a transcription activator and would have the opposite effect. In other words, gene knockout of a cullin substrate would result in the maintenance of silenced viral transcription, which equates to a depletion of HIV-CRISPR vector encoding the respective sgRNA in the viral supernatant. Unfortunately, in this circumstance, there is not an easy way to distinguish between a candidate substrate and a non-targeting control (NTC) since they will both produce low signal in the screen. My solution was to introduce an activation step after the generation of the cullin associated library gene knockout pool of cells was generated in order to create a dynamic range allowing for the examination of a depletion of guideRNAs. By treating these cells with a CD3 and CD28 antibody, which is a high-level *in vitro* T cell activator, most J-Lat cells will reactivate and the J-Lat cells that contain a knockout for a cullin substrate will likely remain latent, produce little to no HIV-CRISPR vector encoding the respective sgRNA in the viral supernatant, and will appear as depleted in the screen results.

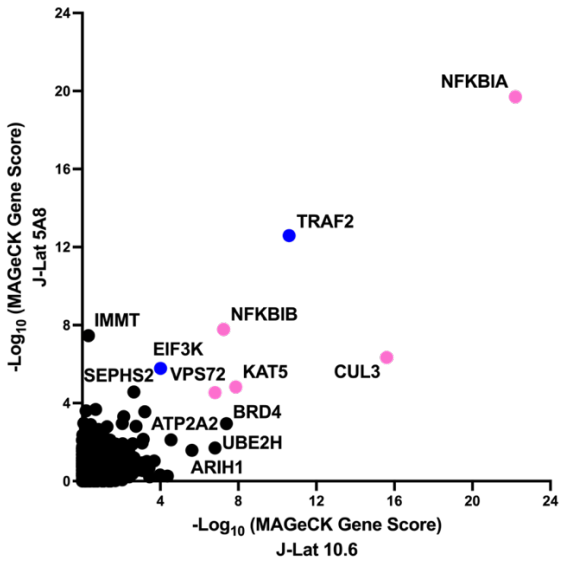


Figure 3.1. Latency HIV-CRISPR screen for cullin adaptors using the cullin associated guideRNA library.

Comparison of the Latency HIV-CRISPR screen by MAGeCK score using the cullin associated guideRNA library in the J-Lat 5A8 cell line (y-axis) and in the J-Lat 10.6 cell line (x-axis). Gene hits in these screens are enriched upon knockout and are candidate adaptors or latency maintenance factors. The genes targeted by the cullin associated guide RNA library are represented by circles and the NTCs are represented by gray squares. The genes unique to each screen are closest to the respective axis and the genes that are in common to both screens are at the center. Gene hits that are <5% FDR in both J-Lat 5A8 and 10.6 cell lines are indicated in pink if it is an expected latency factor and in blue if it is a new candidate factor. Gene hits that are <5% FDR in one cell line is indicated in black with a label.

Each of these screens resulted in the identification of new candidate genes to investigate further at a future time. In both screens, top gene hits were defined as genes with a false discovery rate (FDR) of less than 5%. For the adaptor screen, I did not have any expectations for an obvious gene hit as *CUL3* does not utilize adaptor proteins. Previously validated gene hits from the original Latency HIV-CRISPR screen including *NFKIBA*, *KAT5*, *VPS72* and *CUL3* (**Figure 2.4A**) were top hits in this version of the screen in both J-Lat 5A8 and J-Lat 10.6 cell lines (**Figure 3.1**). Additionally, another top hit, *NFKBIB*, similar to *NFKBIA*, is also involved in NFκB activation [209] and would also be expected to be a latency maintenance factor (**Figure 3.1**). Two novel candidate genes that were top hits in both J-Lat cell lines were *TRAF2* and

EIF3K (**Figure 3.1**). *Eukaryotic translation initiation factor 3 subunit K* (*EIF3K*) has been identified to interact with CUL3 in a mass spectrometry screen [208], but not examined in the context of HIV-1 latency, which makes this gene of great interest to investigate further. *TNF receptor associated factor 2* (*TRAF2*) was a surprising top hit as the gene function of *TRAF2* is to activate NFκB activation [210]. This function is the opposite of what is expected for a latency maintenance factor and it is possible that the knockout of *TRAF2* has indirect effects to result in the phenotype observed in this screen. This screen also identified an additional set of genes including *IMMT, SEPHS2, ATP2A2, BRD4, UBE2H,* and *ARIH1*, which were considered top hits in only one of the two J-Lat cell lines and may also be of interest to validate (**Figure 3.1**).

For the substrate screen, the positive controls *NFKB1* and *MED26*, which are genes that function to enhance HIV-1 transcription ([93] and [211], respectively), scored highly. Additionally, two candidate genes –*Lysine acetyltransferase 7* (*KAT7*) and *F-box and WD repeat domain containing 7* (*FBXW7*) – were also top hits in both J-Lat 10.6 and 5A8 cell lines (**Figure 3.2**) and ongoing work in the lab will seek to validate this finding. This screen had better success in the J-Lat 10.6 cell line compared to the J-Lat 5A8 cell line based on the overall higher MAGeCK gene scores and thereby many candidate genes had an FDR <5% in the J-Lat 10.6 cell line, but very few met the FDR <5% cutoff in the J-Lat 5A8 cell lines (**Figure 3.2**). In order to narrow down the top candidates to validate, I filtered for gene hits that had an FDR <5% in the J-Lat 10.6 cell line and were in the top 40 gene hits in the J-Lat 5A8 cell line and this resulted in a list of the following candidates – *SMG8*, *IKBKB*, *AFF4*, and *APBB1IP* (**Figure 3.2**). Of particular interest is *Inhibitor of Nuclear Factor Kappa-B Kinase Subunit Beta (IKBKB)* as this

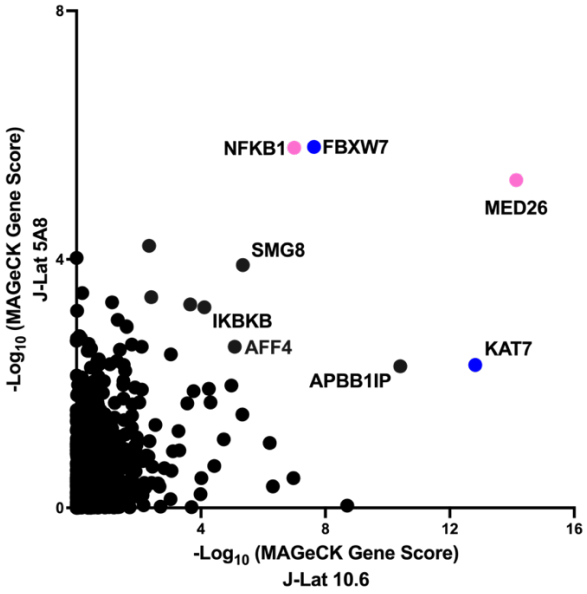


Figure 3.2. Latency HIV-CRISPR screen for cullin substrates using the cullin associated guideRNA library.

Comparison of the Latency HIV-CRISPR screen by MAGeCK score using the cullin associated guideRNA library in the J-Lat 5A8 cell line (y-axis) and in the J-Lat 10.6 cell line (x-axis). Gene hits in these screens are depleted upon knockout and are candidate substrates or transcription activating factors. The genes targeted by the cullin associated guide RNA library are represented by circles and the NTCs are represented by gray squares. The genes unique to each screen are closest to the respective axis and the genes that are in common to both screens are at the center. Gene hits that are <5% FDR in both J-Lat 5A8 and 10.6 cell lines are indicated in pink if it is a known latency factor and in blue if it is a new candidate factor. Gene hits that are <5% FDR in the J-Lat 10.6 cell line and a top 40 hit in the J-Lat 5A8 cell line are indicated in black with a label.

gene has been demonstrated to be a substrate of CUL3-mediated ubiquitination and activates the NFκB pathway [212]. Many candidate genes have been identified through this additional Latency HIV-CRISPR screen using the new cullin associated guideRNA library and will require further investigation in future work.

3.4 Additional extensions of the Latency HIV-CRISPR screen

The Latency HIV-CRISPR screen I developed is versatile to many different formats. One important application is to use the screen in other cell types and directly in primary cells. For

example, there is some data suggesting that a latent reservoir exists in myeloid cells [54, 55]. Immortalized primary microglia cells have been established [213] to study HIV-1 and after establishing a clonal, HIV-1 integrated cell line, these cells could be utilized for performing a Latency HIV-CRISPR screen to examine transcriptional and epigenetic regulators in microglia cells. The similarities and differences between peripheral blood and microglia cells would be critical to identify as most targets for HIV-1 Cure strategies such as the identification of LRAs for "shock and kill" are focused on circulating CD4+ T cells rather than microglia cells [214].

Another extension of the Latency HIV-CRISPR screen is to apply the screen in the context of primary cells. Utilizing primary cells presents many challenges including accumulating sufficient cell numbers, longevity of cells, and more, but this would be an important and relevant context to screen for HIV-1 latency factors. The T cell Optimized for Packaging (TOP) vector, which is a lentiviral based vector, was recently designed to effectively deliver guideRNAs into primary T cells [215] and in combination with electroporation of Cas9, has the potential to be a pivotal tool for primary T cell HIV-1 latency CRIPSR screening. Additionally, while likely not feasible to use for screening purposes, rapid and efficient gene editing methods have recently been established in resting CD4+ T cells [216], which would be an important system to validate gene hits from the Latency HIV-CRISPR screen.

A final extension of the Latency HIV-CRISPR screen would be to perform the screen with other LRAs or combinations of LRAs. In Chapter 2, I described the screen using AZD5582 as the LRA, which resulted in gene hits that were validated in Jurkat and primary CD4+ T cell models of HIV-1 latency. There are many categories of LRAs that could be used in the Latency HIV-CRISPR screen including PKC agonists, HDAC inhibitors, BET inhibitors, and more [214]. Importantly, a low-activating dose of AZD5582 was used for the Latency HIV-CRISPR screen,

which suggests that potentially sub-optimal LRAs can also be used in the screen to extend the mechanistic understanding of why the LRA may not be highly effective in latency reversal. Furthermore, combinations of LRAs can also be tested that target different aspects of transcription to build a comprehensive view of the transcriptional mechanisms involved in HIV-1 silencing.

Altogether, the work described in this thesis establishes a novel CRISPR-based approach to studying HIV-1 latency factors with a focus on understanding how multiple mechanisms interact and identifying factors with specificity and potency for the HIV-1 LTR in order to contribute to HIV-1 Cure. The long-term goal is for the factors and combinations of pathways described here and in future screens to be targeted by small molecules or other inhibitor reagents and tested in various models of HIV-1 latency including cells isolated from ART-suppressed people living with HIV-1 and *in vivo* animal models. The process will require determining the optimized combination of targets, dosing schema, and number of rounds of treatment that maximizes viral reactivation, safety, and efficacy. Additionally, HIV-1 induction is only half of the strategy to eliminate the latent reservoir. Equally important is the process of clearance of the virus that remains. As the viral reactivation may take place in anatomical regions widely distributed across the human body and the HIV-1-specific immune response may be limited, it will be critical to develop strategies to improve the immune response for clearance such as boosting HIV-1-specific CD8+ T cells and engineering of antibodies to target multiple HIV-1 antigens or simultaneously recruit effector cells (reviewed in [2]). The goal of my thesis work is to set a foundation to expand the knowledge of molecular mechanisms of HIV-1 latency to support the development of strategies for HIV-1 Cure.

Bibliography

1. Cohn, L.B., N. Chomont, and S.G. Deeks, *The Biology of the HIV-1 Latent Reservoir and Implications for Cure Strategies.* Cell Host Microbe, 2020. **27**(4): p. 519-530.
2. Margolis, D.M., et al., *Curing HIV: Seeking to Target and Clear Persistent Infection.* Cell, 2020. **181**(1): p. 189-206.
3. Zhou, Y., et al., *Metascape provides a biologist-oriented resource for the analysis of systems-level datasets.* Nat Commun, 2019. **10**(1): p. 1523.
4. Hsieh, E., et al., *A modular CRISPR screen identifies individual and combination pathways contributing to HIV-1 latency.* bioRxiv, 2022: p. 2022.08.23.504195.
5. Chougui, G. and F. Margottin-Goguet, *HUSH, a Link Between Intrinsic Immunity and HIV Latency.* Front Microbiol, 2019. **10**: p. 224.
6. Hakre, S., et al., *HIV latency: experimental systems and molecular models.* FEMS Microbiol Rev, 2012. **36**(3): p. 706-16.
7. Dutilleul, A., A. Rodari, and C. Van Lint, *Depicting HIV-1 Transcriptional Mechanisms: A Summary of What We Know.* Viruses, 2020. **12**(12).
8. UNAIDS. *UNAIDS DATA 2020*. 2020 August 1, 2022]; Available from: https://www.unaids.org/en/resources/documents/2020/unaids-data.
9. Davey, R.T., Jr., et al., *HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression.* Proc Natl Acad Sci U S A, 1999. **96**(26): p. 15109-14.
10. Dufour, C., et al., *The multifaceted nature of HIV latency.* J Clin Invest, 2020. **130**(7): p. 3381-3390.
11. Tong-Starksen, S.E., P.A. Luciw, and B.M. Peterlin, *Human immunodeficiency virus long terminal repeat responds to T-cell activation signals.* Proc Natl Acad Sci U S A, 1987. **84**(19): p. 6845-9.
12. Nabel, G. and D. Baltimore, *An inducible transcription factor activates expression of human immunodeficiency virus in T cells.* Nature, 1987. **326**(6114): p. 711-3.
13. Crooks, A.M., et al., *Precise Quantitation of the Latent HIV-1 Reservoir: Implications for Eradication Strategies.* J Infect Dis, 2015. **212**(9): p. 1361-5.
14. Finzi, D., et al., *Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy.* Nat Med, 1999. **5**(5): p. 512-7.
15. Chun, T.W., et al., *In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency.* Nat Med, 1995. **1**(12): p. 1284-90.
16. Wong, J.K., et al., *Recovery of replication-competent HIV despite prolonged suppression of plasma viremia.* Science, 1997. **278**(5341): p. 1291-5.
17. Chun, T.W., et al., *Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy.* Proc Natl Acad Sci U S A, 1997. **94**(24): p. 13193-7.
18. Finzi, D., et al., *Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy.* Science, 1997. **278**(5341): p. 1295-300.

19. Hocqueloux, L., et al., *Long-term immunovirologic control following antiretroviral therapy interruption in patients treated at the time of primary HIV-1 infection.* Aids, 2010. **24**(10): p. 1598-601.
20. Sáez-Cirión, A., et al., *Post-treatment HIV-1 controllers with a long-term virological remission after the interruption of early initiated antiretroviral therapy ANRS VISCONTI Study.* PLoS Pathog, 2013. **9**(3): p. e1003211.
21. Colby, D.J., et al., *Rapid HIV RNA rebound after antiretroviral treatment interruption in persons durably suppressed in Fiebig I acute HIV infection.* Nat Med, 2018. **24**(7): p. 923-926.
22. Luzuriaga, K., et al., *A trial of three antiretroviral regimens in HIV-1-infected children.* N Engl J Med, 2004. **350**(24): p. 2471-80.
23. Persaud, D., et al., *Slow human immunodeficiency virus type 1 evolution in viral reservoirs in infants treated with effective antiretroviral therapy.* AIDS Res Hum Retroviruses, 2007. **23**(3): p. 381-90.
24. Garcia-Broncano, P., et al., *Early antiretroviral therapy in neonates with HIV-1 infection restricts viral reservoir size and induces a distinct innate immune profile.* Sci Transl Med, 2019. **11**(520).
25. Bosque, A., et al., *Homeostatic proliferation fails to efficiently reactivate HIV-1 latently infected central memory CD4+ T cells.* PLoS Pathog, 2011. **7**(10): p. e1002288.
26. Vandergeeten, C., et al., *Interleukin-7 promotes HIV persistence during antiretroviral therapy.* Blood, 2013. **121**(21): p. 4321-9.
27. Henrich, T.J., et al., *Human Immunodeficiency Virus Type 1 Persistence Following Systemic Chemotherapy for Malignancy.* J Infect Dis, 2017. **216**(2): p. 254-262.
28. Mendoza, P., et al., *Antigen-responsive CD4+ T cell clones contribute to the HIV-1 latent reservoir.* J Exp Med, 2020. **217**(7).
29. Tobin, N.H., et al., *Evidence that low-level viremias during effective highly active antiretroviral therapy result from two processes: expression of archival virus and replication of virus.* J Virol, 2005. **79**(15): p. 9625-34.
30. Bailey, J.R., et al., *Residual human immunodeficiency virus type 1 viremia in some patients on antiretroviral therapy is dominated by a small number of invariant clones rarely found in circulating CD4+ T cells.* J Virol, 2006. **80**(13): p. 6441-57.
31. Bruner, K.M., et al., *A quantitative approach for measuring the reservoir of latent HIV-1 proviruses.* Nature, 2019. **566**(7742): p. 120-125.
32. Peluso, M.J., et al., *Differential decay of intact and defective proviral DNA in HIV-1-infected individuals on suppressive antiretroviral therapy.* JCI Insight, 2020. **5**(4).
33. Ho, Y.C., et al., *Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure.* Cell, 2013. **155**(3): p. 540-51.
34. Wang, Z., et al., *Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane.* Proc Natl Acad Sci U S A, 2018. **115**(11): p. E2575-e2584.
35. Pankau, M.D., et al., *Dynamics of HIV DNA reservoir seeding in a cohort of superinfected Kenyan women.* PLoS Pathog, 2020. **16**(2): p. e1008286.
36. Abrahams, M.R., et al., *The replication-competent HIV-1 latent reservoir is primarily established near the time of therapy initiation.* Sci Transl Med, 2019. **11**(513).
37. Brodin, J., et al., *Establishment and stability of the latent HIV-1 DNA reservoir.* Elife, 2016. **5**.

38. Mohri, H., et al., *Increased turnover of T lymphocytes in HIV-1 infection and its reduction by antiretroviral therapy.* J Exp Med, 2001. **194**(9): p. 1277-87.
39. Xu, W., et al., *CD127 Expression in Naive and Memory T Cells in HIV Patients Who Have Undergone Long-Term HAART.* Lab Med, 2017. **48**(1): p. 57-64.
40. Siliciano, J.D., et al., *Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells.* Nat Med, 2003. **9**(6): p. 727-8.
41. Zerbato, J.M., et al., *Naive CD4+ T Cells Harbor a Large Inducible Reservoir of Latent, Replication-competent Human Immunodeficiency Virus Type 1.* Clin Infect Dis, 2019. **69**(11): p. 1919-1925.
42. Buzon, M.J., et al., *HIV-1 persistence in CD4+ T cells with stem cell-like properties.* Nat Med, 2014. **20**(2): p. 139-42.
43. Chomont, N., et al., *HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation.* Nat Med, 2009. **15**(8): p. 893-900.
44. Soriano-Sarabia, N., et al., *Quantitation of replication-competent HIV-1 in populations of resting CD4+ T cells.* J Virol, 2014. **88**(24): p. 14070-7.
45. Ostrowski, M.A., et al., *Both memory and CD45RA+/CD62L+ naive CD4(+) T cells are infected in human immunodeficiency virus type 1-infected individuals.* J Virol, 1999. **73**(8): p. 6430-5.
46. Bacchus-Souffan, C., et al., *Relationship between CD4 T cell turnover, cellular differentiation and HIV persistence during ART.* PLoS Pathog, 2021. **17**(1): p. e1009214.
47. Chun, T.W., et al., *Persistence of HIV in gut-associated lymphoid tissue despite long-term antiretroviral therapy.* J Infect Dis, 2008. **197**(5): p. 714-20.
48. Estes, J.D., et al., *Defining total-body AIDS-virus burden with implications for curative strategies.* Nat Med, 2017. **23**(11): p. 1271-1276.
49. Gosselin, A., et al., *HIV persists in CCR6+CD4+ T cells from colon and blood during antiretroviral therapy.* Aids, 2017. **31**(1): p. 35-48.
50. Anderson, J.L., et al., *Human Immunodeficiency Virus (HIV)-Infected CCR6+ Rectal CD4+ T Cells and HIV Persistence On Antiretroviral Therapy.* J Infect Dis, 2020. **221**(5): p. 744-755.
51. Banga, R., et al., *Blood CXCR3(+) CD4 T Cells Are Enriched in Inducible Replication Competent HIV in Aviremic Antiretroviral Therapy-Treated Individuals.* Front Immunol, 2018. **9**: p. 144.
52. Ganor, Y., et al., *HIV-1 reservoirs in urethral macrophages of patients under suppressive antiretroviral therapy.* Nat Microbiol, 2019. **4**(4): p. 633-644.
53. Cantero-Pérez, J., et al., *Resident memory T cells are a cellular reservoir for HIV in the cervical mucosa.* Nat Commun, 2019. **10**(1): p. 4739.
54. Lamers, S.L., et al., *HIV DNA Is Frequently Present within Pathologic Tissues Evaluated at Autopsy from Combined Antiretroviral Therapy-Treated Patients with Undetectable Viral Loads.* J Virol, 2016. **90**(20): p. 8968-83.
55. Avalos, C.R., et al., *Brain Macrophages in Simian Immunodeficiency Virus-Infected, Antiretroviral-Suppressed Macaques: a Functional Latent Reservoir.* mBio, 2017. **8**(4).
56. Chaillon, A., et al., *HIV persists throughout deep tissues with repopulation from multiple anatomical sources.* J Clin Invest, 2020. **130**(4): p. 1699-1712.
57. Folks, T.M., et al., *Tumor necrosis factor alpha induces expression of human immunodeficiency virus in a chronically infected T-cell clone.* Proc Natl Acad Sci U S A, 1989. **86**(7): p. 2365-8.

58. Pomerantz, R.J., et al., *Cells nonproductively infected with HIV-1 exhibit an aberrant pattern of viral RNA expression: a molecular model for latency.* Cell, 1990. **61**(7): p. 1271-6.
59. Antoni, B.A., et al., *NF-kappa B-dependent and -independent pathways of HIV activation in a chronically infected T cell line.* Virology, 1994. **202**(2): p. 684-94.
60. Micheva-Viteva, S., et al., *High-throughput screening uncovers a compound that activates latent HIV-1 and acts cooperatively with a histone deacetylase (HDAC) inhibitor.* J Biol Chem, 2011. **286**(24): p. 21083-91.
61. Jordan, A.B., D.; Verdin, E., *HIV reproducibly establishes a latent infection after acute infection of T cells in vitro.* The EMBO Journal, 2003. **22**: p. 1868-1877.
62. Chan, J.K., et al., *Calcium/calcineurin synergizes with prostratin to promote NF-kappaB dependent activation of latent HIV.* PLoS One, 2013. **8**(10): p. e77749.
63. Spina, C.A., et al., *An in-depth comparison of latent HIV-1 reactivation in multiple cell model systems and resting CD4+ T cells from aviremic patients.* PLoS Pathog, 2013. **9**(12): p. e1003834.
64. Sahu, G.K., et al., *A novel in vitro system to generate and study latently HIV-infected long-lived normal CD4+ T-lymphocytes.* Virology, 2006. **355**(2): p. 127-37.
65. Tyagi, M., R.J. Pearson, and J. Karn, *Establishment of HIV latency in primary CD4+ cells is due to epigenetic transcriptional silencing and P-TEFb restriction.* J Virol, 2010. **84**(13): p. 6425-37.
66. Marini, A., J.M. Harper, and F. Romerio, *An in vitro system to model the establishment and reactivation of HIV-1 latency.* J Immunol, 2008. **181**(11): p. 7713-20.
67. Bosque, A. and V. Planelles, *Induction of HIV-1 latency and reactivation in primary memory CD4+ T cells.* Blood, 2009. **113**(1): p. 58-65.
68. Yang, H.C., et al., *Small-molecule screening using a human primary cell model of HIV latency identifies compounds that reverse latency without cellular activation.* J Clin Invest, 2009. **119**(11): p. 3473-86.
69. Dobrowolski, C., et al., *Entry of Polarized Effector Cells into Quiescence Forces HIV Latency.* MBio, 2019. **10**(2).
70. Takata, H., et al., *Modeling HIV-1 Latency Using Primary CD4(+) T Cells from Virally Suppressed HIV-1-Infected Individuals on Antiretroviral Therapy.* J Virol, 2019. **93**(11).
71. Calvanese, V., et al., *Dual-color HIV reporters trace a population of latently infected cells and enable their purification.* Virology, 2013. **446**(1-2): p. 283-92.
72. Battivelli, E., et al., *Distinct chromatin functional states correlate with HIV latency reactivation in infected primary CD4(+) T cells.* Elife, 2018. **7**.
73. Swiggard, W.J., et al., *Human immunodeficiency virus type 1 can establish latent infection in resting CD4+ T cells in the absence of activating stimuli.* J Virol, 2005. **79**(22): p. 14179-88.
74. Sallusto, F., et al., *Two subsets of memory T lymphocytes with distinct homing potentials and effector functions.* Nature, 1999. **401**(6754): p. 708-12.
75. Saleh, S., et al., *CCR7 ligands CCL19 and CCL21 increase permissiveness of resting memory CD4+ T cells to HIV-1 infection: a novel model of HIV-1 latency.* Blood, 2007. **110**(13): p. 4161-4.
76. Pierson, T.C., et al., *Molecular characterization of preintegration latency in human immunodeficiency virus type 1 infection.* J Virol, 2002. **76**(17): p. 8518-31.

77. Whitney, J.B., et al., *Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys.* Nature, 2014. **512**(7512): p. 74-7.
78. Whitney, J.B., et al., *Prevention of SIVmac251 reservoir seeding in rhesus monkeys by early antiretroviral therapy.* Nat Commun, 2018. **9**(1): p. 5429.
79. Okoye, A.A., et al., *Early antiretroviral therapy limits SIV reservoir establishment to delay or prevent post-treatment viral rebound.* Nat Med, 2018. **24**(9): p. 1430-1440.
80. Melkus, M.W., et al., *Humanized mice mount specific adaptive and innate immune responses to EBV and TSST-1.* Nat Med, 2006. **12**(11): p. 1316-22.
81. Denton, P.W., et al., *Generation of HIV latency in humanized BLT mice.* J Virol, 2012. **86**(1): p. 630-4.
82. McBrien, J.B., et al., *Robust and persistent reactivation of SIV and HIV by N-803 and depletion of CD8(+) cells.* Nature, 2020. **578**(7793): p. 154-159.
83. Nixon, C.C., et al., *Systemic HIV and SIV latency reversal via non-canonical NF-kappaB signalling in vivo.* Nature, 2020. **578**(7793): p. 160-165.
84. Schröder, A.R., et al., *HIV-1 integration in the human genome favors active genes and local hotspots.* Cell, 2002. **110**(4): p. 521-9.
85. Lewinski, M.K., et al., *Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription.* J Virol, 2005. **79**(11): p. 6610-9.
86. Han, Y., et al., *Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes.* J Virol, 2004. **78**(12): p. 6122-33.
87. Liu, H., et al., *Integration of human immunodeficiency virus type 1 in untreated infection occurs preferentially within genes.* J Virol, 2006. **80**(15): p. 7765-8.
88. Marini, B., et al., *Nuclear architecture dictates HIV-1 integration site selection.* Nature, 2015. **521**(7551): p. 227-31.
89. Gallastegui, E., et al., *Chromatin reassembly factors are involved in transcriptional interference promoting HIV latency.* J Virol, 2011. **85**(7): p. 3187-202.
90. Jiang, C., et al., *Distinct viral reservoirs in individuals with spontaneous control of HIV-1.* Nature, 2020. **585**(7824): p. 261-267.
91. Wagner, T.A., et al., *HIV latency. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection.* Science, 2014. **345**(6196): p. 570-3.
92. Pereira, L.A., et al., *A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter.* Nucleic Acids Res, 2000. **28**(3): p. 663-8.
93. Chen, L.F. and W.C. Greene, *Shaping the nuclear action of NF-kappaB.* Nat Rev Mol Cell Biol, 2004. **5**(5): p. 392-401.
94. Ping, Y.H. and T.M. Rana, *DSIF and NELF interact with RNA polymerase II elongation complex and HIV-1 Tat stimulates P-TEFb-mediated phosphorylation of RNA polymerase II and DSIF during transcription elongation.* J Biol Chem, 2001. **276**(16): p. 12951-8.
95. Bisgrove, D., et al., *Molecular mechanisms of HIV-1 proviral latency.* Expert Rev Anti Infect Ther, 2005. **3**(5): p. 805-14.
96. Wei, P., et al., *A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA.* Cell, 1998. **92**(4): p. 451-62.
97. Tahirov, T.H., et al., *Crystal structure of HIV-1 Tat complexed with human P-TEFb.* Nature, 2010. **465**(7299): p. 747-51.

98. Fujinaga, K., et al., *Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element.* Mol Cell Biol, 2004. **24**(2): p. 787-95.
99. Kim, Y.K., et al., *Phosphorylation of the RNA polymerase II carboxyl-terminal domain by CDK9 is directly responsible for human immunodeficiency virus type 1 Tat-activated transcriptional elongation.* Mol Cell Biol, 2002. **22**(13): p. 4622-37.
100. Ivanov, D., et al., *Domains in the SPT5 protein that modulate its transcriptional regulatory properties.* Mol Cell Biol, 2000. **20**(9): p. 2970-83.
101. Bourgeois, C.F., et al., *Spt5 cooperates with human immunodeficiency virus type 1 Tat by preventing premature RNA release at terminator sequences.* Mol Cell Biol, 2002. **22**(4): p. 1079-93.
102. Pearson, R., et al., *Epigenetic silencing of human immunodeficiency virus (HIV) transcription by formation of restrictive chromatin structures at the viral long terminal repeat drives the progressive entry of HIV into latency.* J Virol, 2008. **82**(24): p. 12291-303.
103. Yukl, S., et al., *Latently-infected CD4+ T cells are enriched for HIV-1 Tat variants with impaired transactivation activity.* Virology, 2009. **387**(1): p. 98-108.
104. French, A.J., et al., *Reactivating latent HIV with PKC agonists induces resistance to apoptosis and is associated with phosphorylation and activation of BCL2.* PLoS Pathog, 2020. **16**(10): p. e1008906.
105. Pache, L., et al., *BIRC2/cIAP1 Is a Negative Regulator of HIV-1 Transcription and Can Be Targeted by Smac Mimetics to Promote Reversal of Viral Latency.* Cell Host Microbe, 2015. **18**(3): p. 345-53.
106. Contreras, X., et al., *HMBA releases P-TEFb from HEXIM1 and 7SK snRNA via PI3K/Akt and activates HIV transcription.* PLoS Pathog, 2007. **3**(10): p. 1459-69.
107. Bisgrove, D.A., et al., *Conserved P-TEFb-interacting domain of BRD4 inhibits HIV transcription.* Proc Natl Acad Sci U S A, 2007. **104**(34): p. 13690-5.
108. Zhu, J., et al., *Reactivation of latent HIV-1 by inhibition of BRD4.* Cell Rep, 2012. **2**(4): p. 807-16.
109. Li, Z., et al., *The BET bromodomain inhibitor JQ1 activates HIV latency through antagonizing Brd4 inhibition of Tat-transactivation.* Nucleic Acids Res, 2013. **41**(1): p. 277-87.
110. Bartholomeeusen, K., et al., *Bromodomain and extra-terminal (BET) bromodomain inhibition activate transcription via transient release of positive transcription elongation factor b (P-TEFb) from 7SK small nuclear ribonucleoprotein.* J Biol Chem, 2012. **287**(43): p. 36609-16.
111. Falcinelli, S.D., et al., *Combined noncanonical NF-κB agonism and targeted BET bromodomain inhibition reverse HIV latency ex vivo.* J Clin Invest, 2022. **132**(8).
112. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 A resolution.* Nature, 1997. **389**(6648): p. 251-60.
113. Jenuwein, T. and C.D. Allis, *Translating the histone code.* Science, 2001. **293**(5532): p. 1074-80.
114. Verdin, E., *DNase I-hypersensitive sites are associated with both long terminal repeats and with the intragenic enhancer of integrated human immunodeficiency virus type 1.* J Virol, 1991. **65**(12): p. 6790-9.

115. Verdin, E., P. Paras, Jr., and C. Van Lint, *Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation.* Embo j, 1993. **12**(8): p. 3249-59.
116. Berger, S.L., *The complex language of chromatin regulation during transcription.* Nature, 2007. **447**(7143): p. 407-12.
117. Wang, Z., et al., *Combinatorial patterns of histone acetylations and methylations in the human genome.* Nat Genet, 2008. **40**(7): p. 897-903.
118. Van Lint, C., et al., *Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation.* Embo j, 1996. **15**(5): p. 1112-20.
119. Lusic, M., et al., *Regulation of HIV-1 gene expression by histone acetylation and factor recruitment at the LTR promoter.* Embo j, 2003. **22**(24): p. 6550-61.
120. Gerritsen, M.E., et al., *CREB-binding protein/p300 are transcriptional coactivators of p65.* Proc Natl Acad Sci U S A, 1997. **94**(7): p. 2927-32.
121. Benkirane, M., et al., *Activation of integrated provirus requires histone acetyltransferase. p300 and P/CAF are coactivators for HIV-1 Tat.* J Biol Chem, 1998. **273**(38): p. 24898-905.
122. Ott, M., et al., *Acetylation of the HIV-1 Tat protein by p300 is important for its transcriptional activity.* Curr Biol, 1999. **9**(24): p. 1489-92.
123. Hottiger, M.O. and G.J. Nabel, *Interaction of human immunodeficiency virus type 1 Tat with the transcriptional coactivators p300 and CREB binding protein.* J Virol, 1998. **72**(10): p. 8252-6.
124. Keedy, K.S., et al., *A limited group of class I histone deacetylases acts to repress human immunodeficiency virus type 1 expression.* J Virol, 2009. **83**(10): p. 4749-56.
125. Coull, J.J., et al., *The human factors YY1 and LSF repress the human immunodeficiency virus type 1 long terminal repeat via recruitment of histone deacetylase 1.* J Virol, 2000. **74**(15): p. 6790-9.
126. He, G. and D.M. Margolis, *Counterregulation of chromatin deacetylation and histone deacetylase occupancy at the integrated promoter of human immunodeficiency virus type 1 (HIV-1) by the HIV-1 repressor YY1 and HIV-1 activator Tat.* Mol Cell Biol, 2002. **22**(9): p. 2965-73.
127. Turner, A.W. and D.M. Margolis, *Chromatin Regulation and the Histone Code in HIV Latency.* Yale J Biol Med, 2017. **90**(2): p. 229-243.
128. Archin, N.M., et al., *Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy.* Nature, 2012. **487**(7408): p. 482-5.
129. Archin, N.M., et al., *Interval dosing with the HDAC inhibitor vorinostat effectively reverses HIV latency.* J Clin Invest, 2017. **127**(8): p. 3126-3135.
130. Barth, T.K. and A. Imhof, *Fast signals and slow marks: the dynamics of histone modifications.* Trends Biochem Sci, 2010. **35**(11): p. 618-26.
131. Nguyen, K., et al., *Multiple Histone Lysine Methyltransferases Are Required for the Establishment and Maintenance of HIV-1 Latency.* mBio, 2017. **8**(1).
132. Friedman, J., et al., *Epigenetic silencing of HIV-1 by the histone H3 lysine 27 methyltransferase enhancer of Zeste 2.* J Virol, 2011. **85**(17): p. 9078-89.
133. Tripathy, M.K., et al., *H3K27 Demethylation at the Proviral Promoter Sensitizes Latent HIV to the Effects of Vorinostat in Ex Vivo Cultures of Resting CD4+ T Cells.* J Virol, 2015. **89**(16): p. 8392-405.

134. Blackledge, N.P., N.R. Rose, and R.J. Klose, *Targeting Polycomb systems to regulate gene expression: modifications to a complex story.* Nat Rev Mol Cell Biol, 2015. **16**(11): p. 643-649.
135. Imai, K., H. Togami, and T. Okamoto, *Involvement of histone H3 lysine 9 (H3K9) methyltransferase G9a in the maintenance of HIV-1 latency and its reactivation by BIX01294.* J Biol Chem, 2010. **285**(22): p. 16538-45.
136. Tchasovnikarova, I.A., et al., *GENE SILENCING. Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells.* Science, 2015. **348**(6242): p. 1481-1485.
137. Chougui, G., et al., *HIV-2/SIV viral protein X counteracts HUSH repressor complex.* Nat Microbiol, 2018. **3**(8): p. 891-897.
138. Pedersen, S.F., et al., *Inhibition of a Chromatin and Transcription Modulator, SLTM, Increases HIV-1 Reactivation Identified by a CRISPR Inhibition Screen.* J Virol, 2022: p. e0057722.
139. Rafati, H., et al., *Repressive LTR nucleosome positioning by the BAF complex is required for HIV latency.* PLoS Biol, 2011. **9**(11): p. e1001206.
140. Easley, R., et al., *Transcription through the HIV-1 nucleosomes: effects of the PBAF complex in Tat activated transcription.* Virology, 2010. **405**(2): p. 322-33.
141. Deeks, S.G., et al., *HIV infection.* Nat Rev Dis Primers, 2015. **1**: p. 15035.
142. Chun, T.W., et al., *Re-emergence of HIV after stopping therapy.* Nature, 1999. **401**(6756): p. 874-5.
143. Sengupta, S. and R.F. Siliciano, *Targeting the Latent Reservoir for HIV-1.* Immunity, 2018. **48**(5): p. 872-895.
144. Einkauf, K.B., et al., *Intact HIV-1 proviruses accumulate at distinct chromosomal positions during prolonged antiretroviral therapy.* J Clin Invest, 2019. **129**(3): p. 988-998.
145. Mbonye, U. and J. Karn, *The Molecular Basis for Human Immunodeficiency Virus Latency.* Annu Rev Virol, 2017. **4**(1): p. 261-285.
146. Turner, A.W., et al., *Evaluation of EED Inhibitors as a Class of PRC2-Targeted Small Molecules for HIV Latency Reversal.* ACS Infect Dis, 2020. **6**(7): p. 1719-1733.
147. Li, Z., et al., *The KAT5-Acetyl-Histone4-Brd4 axis silences HIV-1 transcription and promotes viral latency.* PLoS Pathog, 2018. **14**(4): p. e1007012.
148. Boehm, D. and M. Ott, *Host Methyltransferases and Demethylases: Potential New Epigenetic Targets for HIV Cure Strategies and Beyond.* AIDS Res Hum Retroviruses, 2017. **33**(S1): p. S8-s22.
149. Conrad, R.J., et al., *The Short Isoform of BRD4 Promotes HIV-1 Latency by Engaging Repressive SWI/SNF Chromatin-Remodeling Complexes.* Mol Cell, 2017. **67**(6): p. 1001-1012.e6.
150. Besnard, E., et al., *The mTOR Complex Controls HIV Latency.* Cell Host Microbe, 2016. **20**(6): p. 785-797.
151. Das, B., et al., *Estrogen receptor-1 is a key regulator of HIV-1 latency that imparts gender-specific restrictions on the latent reservoir.* Proc Natl Acad Sci U S A, 2018. **115**(33): p. E7795-e7804.
152. Boehm, D., et al., *SMYD2-Mediated Histone Methylation Contributes to HIV-1 Latency.* Cell Host Microbe, 2017. **21**(5): p. 569-579.e6.

153. Li, Z., et al., *Reiterative Enrichment and Authentication of CRISPRi Targets (REACT) identifies the proteasome as a key contributor to HIV-1 latency.* PLoS Pathog, 2019. **15**(1): p. e1007498.
154. Huang, H., et al., *A CRISPR/Cas9 screen identifies the histone demethylase MINA53 as a novel HIV-1 latency-promoting gene (LPG).* Nucleic Acids Res, 2019. **47**(14): p. 7333-7347.
155. Rathore, A., et al., *CRISPR-based gene knockout screens reveal deubiquitinases involved in HIV-1 latency in two Jurkat cell models.* Sci Rep, 2020. **10**(1): p. 5350.
156. OhAinle, M., et al., *A virus-packageable CRISPR screen identifies host factors mediating interferon inhibition of HIV.* Elife, 2018. **7**.
157. Janssens, D.H., et al., *Automated CUT&Tag profiling of chromatin heterogeneity in mixed-lineage leukemia.* Nat Genet, 2021. **53**(11): p. 1586-1596.
158. Medvedeva, Y.A., et al., *EpiFactors: a comprehensive database of human epigenetic factors and complexes.* Database, 2015. **2015**.
159. Sanjana, N.E., O. Shalem, and F. Zhang, *Improved vectors and genome-wide libraries for CRISPR screening.* Nat Methods, 2014. **11**(8): p. 783-784.
160. Li, W., et al., *MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens.* Genome Biology, 2014. **15**(12): p. 554.
161. Yang, Z., et al., *Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4.* Mol Cell, 2005. **19**(4): p. 535-45.
162. Doyon, Y., et al., *Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans.* Mol Cell Biol, 2004. **24**(5): p. 1884-96.
163. Ginsburg, D.S., C.K. Govind, and A.G. Hinnebusch, *NuA4 lysine acetyltransferase Esa1 is targeted to coding regions and stimulates transcription elongation with Gcn5.* Mol Cell Biol, 2009. **29**(24): p. 6473-87.
164. Doyon, Y. and J. Côté, *The highly conserved and multifunctional NuA4 HAT complex.* Curr Opin Genet Dev, 2004. **14**(2): p. 147-54.
165. Ruhl, D.D., et al., *Purification of a human SRCAP complex that remodels chromatin by incorporating the histone variant H2A.Z into nucleosomes.* Biochemistry, 2006. **45**(17): p. 5671-7.
166. Mizuguchi, G., et al., *ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex.* Science, 2004. **303**(5656): p. 343-8.
167. Luk, E., et al., *Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome.* Cell, 2010. **143**(5): p. 725-36.
168. Hsiau, T., et al., *Inference of CRISPR Edits from Sanger Trace Data.* bioRxiv, 2019: p. 251082.
169. Hiatt, J., et al., *A functional map of HIV-host interactions in primary human T cells.* Nat Commun, 2022. **13**(1): p. 1752.
170. Rountree, M.R., K.E. Bachman, and S.B. Baylin, *DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci.* Nat Genet, 2000. **25**(3): p. 269-77.
171. Pintard, L., A. Willems, and M. Peter, *Cullin-based ubiquitin ligases: Cul3-BTB complexes join the family.* Embo j, 2004. **23**(8): p. 1681-7.
172. Jefferys, S.R., et al., *Epigenomic characterization of latent HIV infection identifies latency regulating transcription factors.* PLoS Pathog, 2021. **17**(2): p. e1009346.

173. Peterson, J.J., et al., *A histone deacetylase network regulates epigenetic reprogramming and viral silencing in HIV infected cells.* bioRxiv, 2022: p. 2022.05.09.491199.
174. Chu, V.T., et al., *Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells.* Nat Biotechnol, 2015. **33**(5): p. 543-8.
175. Langer, S., et al., *The E3 Ubiquitin-Protein Ligase Cullin 3 Regulates HIV-1 Transcription.* Cells, 2020. **9**(9).
176. Doyon, Y., et al., *ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation.* Mol Cell, 2006. **21**(1): p. 51-64.
177. Hsin, J.P. and J.L. Manley, *The RNA polymerase II CTD coordinates transcription and RNA processing.* Genes Dev, 2012. **26**(19): p. 2119-37.
178. Czudnochowski, N., C.A. Bösken, and M. Geyer, *Serine-7 but not serine-5 phosphorylation primes RNA polymerase II CTD for P-TEFb recognition.* Nat Commun, 2012. **3**: p. 842.
179. Ravens, S., et al., *Tip60 complex binds to active Pol II promoters and a subset of enhancers and co-regulates the c-Myc network in mouse embryonic stem cells.* Epigenetics Chromatin, 2015. **8**: p. 45.
180. Devoucoux, M., et al., *Oncogenic ZMYND11-MBTD1 fusion protein anchors the NuA4/TIP60 histone acetyltransferase complex to the coding region of active genes.* Cell Rep, 2022. **39**(11): p. 110947.
181. Li, J., et al., *ZMYND11-MBTD1 induces leukemogenesis through hijacking NuA4/TIP60 acetyltransferase complex and a PWWP-mediated chromatin association mechanism.* Nat Commun, 2021. **12**(1): p. 1045.
182. Auger, A., et al., *Eaf1 is the platform for NuA4 molecular assembly that evolutionarily links chromatin acetylation to ATP-dependent exchange of histone H2A variants.* Mol Cell Biol, 2008. **28**(7): p. 2257-70.
183. Marzio, G., et al., *HIV-1 tat transactivator recruits p300 and CREB-binding protein histone acetyltransferases to the viral promoter.* Proc Natl Acad Sci U S A, 1998. **95**(23): p. 13519-24.
184. Col, E., et al., *The histone acetyltransferase, hGCN5, interacts with and acetylates the HIV transactivator, Tat.* J Biol Chem, 2001. **276**(30): p. 28179-84.
185. Steger, D.J., et al., *Purified histone acetyltransferase complexes stimulate HIV-1 transcription from preassembled nucleosomal arrays.* Proc Natl Acad Sci U S A, 1998. **95**(22): p. 12924-9.
186. Almeida, M., et al., *PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation.* Science, 2017. **356**(6342): p. 1081-1084.
187. Sharma, A.L., et al., *CBF-1 Promotes the Establishment and Maintenance of HIV Latency by Recruiting Polycomb Repressive Complexes, PRC1 and PRC2, at HIV LTR.* Viruses, 2020. **12**(9).
188. Gonçalves, E., et al., *Minimal genome-wide human CRISPR-Cas9 library.* Genome Biol, 2021. **22**(1): p. 40.
189. Humbert, O., et al., *Development of Third-generation Cocal Envelope Producer Cell Lines for Robust Lentiviral Gene Transfer into Hematopoietic Stem Cells and T-cells.* Mol Ther, 2016. **24**(7): p. 1237-46.

190. Meier, J.A., F. Zhang, and N.E. Sanjana, *GUIDES: sgRNA design for loss-of-function screens.* Nat Methods, 2017. **14**(9): p. 831-832.
191. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome Biol, 2009. **10**(3): p. R25.
192. Roesch, F., M. OhAinle, and M. Emerman, *A CRISPR screen for factors regulating SAMHD1 degradation identifies IFITMs as potent inhibitors of lentiviral particle delivery.* Retrovirology, 2018. **15**(1): p. 26.
193. Vermeire, J., et al., *Quantification of reverse transcriptase activity by real-time PCR as a fast and accurate method for titration of HIV, lenti- and retroviral vectors.* PLoS One, 2012. **7**(12): p. e50859.
194. Chung, C.H., et al., *Integrated Human Immunodeficiency Virus Type 1 Sequence in J-Lat 10.6.* Microbiol Resour Announc, 2020. **9**(18).
195. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nat Methods, 2012. **9**(4): p. 357-9.
196. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
197. Meers, M.P., D. Tenenbaum, and S. Henikoff, *Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling.* Epigenetics Chromatin, 2019. **12**(1): p. 42.
198. Song, Y., et al., *Inhibitor of growth protein 3 epigenetically silences endogenous retroviral elements and prevents innate immune activation.* Nucleic Acids Res, 2021. **49**(22): p. 12706-12715.
199. Cromm, P.M. and C.M. Crews, *Targeted Protein Degradation: from Chemical Biology to Drug Discovery.* Cell Chem Biol, 2017. **24**(9): p. 1181-1190.
200. Qin, H., et al., *Overview of PROTACs Targeting the Estrogen Receptor: Achievements for Biological and Drug Discovery.* Curr Med Chem, 2022. **29**(22): p. 3922-3944.
201. Potjewyd, F., et al., *Degradation of Polycomb Repressive Complex 2 with an EED-Targeted Bivalent Chemical Degrader.* Cell Chem Biol, 2020. **27**(1): p. 47-56.e15.
202. Feng, Y., et al., *Cryo-EM structure of human SRCAP complex.* Cell Res, 2018. **28**(11): p. 1121-1123.
203. Altaf, M., et al., *NuA4-dependent acetylation of nucleosomal histones H4 and H2A directly stimulates incorporation of H2A.Z by the SWR1 complex.* J Biol Chem, 2010. **285**(21): p. 15966-77.
204. Faust, T.B., et al., *The HIV-1 Tat protein recruits a ubiquitin ligase to reorganize the 7SK snRNP for transcriptional activation.* Elife, 2018. **7**.
205. Faust, T.B., et al., *PJA2 ubiquitinates the HIV-1 Tat protein with atypical chain linkages to activate viral transcription.* Sci Rep, 2017. **7**: p. 45394.
206. Wu, J., et al., *Host cell factors stimulate HIV-1 transcription by antagonizing substrate-binding function of Siah1 ubiquitin ligase to stabilize transcription elongation factor ELL2.* Nucleic Acids Res, 2020. **48**(13): p. 7321-7332.
207. Sarikas, A., T. Hartmann, and Z.Q. Pan, *The cullin protein family.* Genome Biol, 2011. **12**(4): p. 220.
208. Emanuele, M.J., et al., *Global identification of modular cullin-RING ligase substrates.* Cell, 2011. **147**(2): p. 459-74.
209. Thompson, J.E., et al., *I kappa B-beta regulates the persistent response in a biphasic activation of NF-kappa B.* Cell, 1995. **80**(4): p. 573-82.

210. Pasquereau, S., A. Kumar, and G. Herbein, *Targeting TNF and TNF Receptor Pathway in HIV-1 Infection: from Immune Activation to Viral Reservoirs.* Viruses, 2017. **9**(4).
211. Ruiz, A., et al., *Characterization of the influence of mediator complex in HIV-1 transcription.* J Biol Chem, 2014. **289**(40): p. 27665-76.
212. Lee, D.F., et al., *KEAP1 E3 ligase-mediated downregulation of NF-kappaB signaling by targeting IKKbeta.* Mol Cell, 2009. **36**(1): p. 131-40.
213. Garcia-Mesa, Y., et al., *Immortalization of primary microglia: a new platform to study HIV regulation in the central nervous system.* J Neurovirol, 2017. **23**(1): p. 47-66.
214. Spivak, A.M. and V. Planelles, *Novel Latency Reversal Agents for HIV-1 Cure.* Annu Rev Med, 2018. **69**: p. 421-436.
215. Humes, D., S. Rainwater, and J. Overbaugh, *The TOP vector: a new high-titer lentiviral construct for delivery of sgRNAs and transgenes to primary T cells.* Mol Ther Methods Clin Dev, 2021. **20**: p. 30-38.
216. Albanese, M., et al., *Rapid, efficient and activation-neutral gene editing of polyclonal primary human resting CD4(+) T cells allows complex functional analyses.* Nat Methods, 2022. **19**(1): p. 81-89.