A two-color haploid genetic screen identifies novel host factors involved in HIV latency

Running Title: Identification of novel HIV latency host factors.

Authors: Michael D Röling¹, Mahsa Mollapour Sisakht^{1#}, Enrico Ne^{1#}, Panagiotis Moulos^{2#}, Mateusz Stoszko¹, Elisa De Crignis¹, Helen Bodmer¹, Tsung Wai Kan¹, Maryam Akbarzadeh¹, Vaggelis Harokopos², Pantelis Hatzis², Robert-Jan Palstra^{1*}, Tokameh Mahmoudi^{1*}

authors contributed equally

Affiliations:

1 Department of Biochemistry, Erasmus University Medical Center, Ee634 PO Box 2040 3000CA Rotterdam, the Netherlands

2 Institute for Fundamental Biomedical Research, Biomedical Sciences Research Center, Alexander Fleming', Fleming, Fleming Street 34, 16672, Vari, Greece

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

Phone N: +31 (0)107043324, Fax N: +31(0)10704747

Robert-Jan Palstra: r.palstra@erasmusmc.nl

Phone N: +31 (0)107043323, Fax N: +31(0)10704747

Abstract

To identify novel host factors as putative targets to reverse HIV latency, we performed an insertional mutagenesis genetic screen in a latently HIV-1-infected pseudo-haploid KBM7 cell line (Hap-Lat). Following mutagenesis, insertions were mapped to the genome and bioinformatic analysis resulted in the identification of 69 candidate host genes involved in maintaining HIV-1 latency. A select set of candidate genes was functionally validated using shRNA mediated depletion in latent HIV-1 infected J-Lat A2 and 11.1 T cell lines. We confirmed ADK, CHD9, CMSS1, EVI2B, EXOSC8, FAM19A, GRIK5, IRF2BP2, NF1, and USP15 as novel host factors involved in the maintenance of HIV latency. Chromatin immunoprecipitation assays indicated that CHD9, a Chromodomain Helicase DNA-binding protein, maintains HIV latency via direct association with the HIV 5'LTR, and its depletion results in increased histone acetylation at the HIV-1 promoter, concomitant with HIV-1 latency reversal. FDA-approved inhibitors 5-Iodotubercidin, Trametinib, and Topiramate, targeting ADK, NF1, and GRIK5, respectively were characterized for their latency reversal potential. While 5-Iodotubercidin exhibited significant cytotoxicity in both J-Lat and primary CD4+ T cells, Trametinib reversed latency in J-Lat cells but not in latently HIV-1-infected primary CD4+ T cells. Crucially, Topiramate reversed latency in cell-line models and latently infected primary CD4+ T cells, without inducing T cell activation or significant toxicity. Thus, using an adaptation of a haploid forward genetic screen, we identified novel and druggable host factors contributing to HIV-1 latency.

Importance

A reservoir of latent HIV-1-infected cells persists in the presence of combination antiretroviral therapy (cART), representing a major obstacle for viral eradication. Reactivation of the latent HIV-1 provirus is part of curative strategies which aim to promote clearance of the infected cells. Using a two-color haploid screen, we identified 69 candidate genes as latency maintaining host factors and functionally validated a subset of 10 of those in additional T-cell based cell line models of HIV-1 latency. We further demonstrated that CHD9 is associated with HIV-1's promoter, the 5'LTR while this association is lost upon reactivation. Additionally, we characterized the latency reversal potential of FDA compounds targeting ADK, NF1, and GRIK5 and identify the GRIK5 inhibitor Topiramate as a viable latency reversal agent with clinical potential.

Introduction

Combination antiretroviral therapy (cART) has proven to effectively abrogate viral replication in HIV-1-infected patients and has substantially reduced AIDS-related mortality. However, cART is not curative and patients must remain on life-long medication regimens, as interruption of the therapy leads to rapid rebound of viral replication (1). This is due to the persistence of a reservoir of latently infected cells, harboring replication-competent provirus blocked at the level of gene expression, that escape clearance by the immune system (2). Therapeutic strategies toward HIV-1 cures aim to inactivate, reduce or completely eradicate the latent reservoir, such that, upon cessation of cART, the patient's immune system can effectively control the infection or fully clear it (2). Strategies aiming to reduce or eliminate the reservoir rely on drugs, termed latency-reversing agents (LRAs), capable of inducing the latent HIV infected cells to express viral genes to render infected cells susceptible to cytopathic effects and/or recognition and clearance by the immune system (3). Much focus has therefore been placed on finding small molecules for activating transcription of the latent HIV-1 provirus, which are not cytotoxic and do not induce harmful T cell activation or proliferation.

The identification of molecules capable of inducing HIV gene expression has been largely accomplished using candidate approaches, which build on existing knowledge of transcription factors and co-factors that bind to and regulate transcription at the HIV LTR (4–6). One of the most clinically studied classes of molecules in HIV latency reversal, HDAC inhibitors (HDACis) such as SAHA (7), valproic acid (8), rhomidepsin (9) and M344 (10), target HDACs, which have been shown to be recruited to the HIV LTR by multiple transcription factors (11–13) to deacetylate histones and repress transcription. Similarly, agonists of the PKC pathway, such as prostratin and bryostratin, have been studied as activators of the HIV LTR, as they induce nuclear localization and LTR binding of the NFκB (p65/p50) heterodimer, a potent activator of HIV transcription (14, 15). While this candidate approach has led to the identification of druggable targets or potential candidate LRAs, none of the LRAs currently under clinical investigation are capable of strong latency reversal in patients or lead to a reduction in the size of the latent HIV reservoir (16, 17). Thus, to identify more potent and clinically relevant LRAs, it is critical to identify the full repertoire of functionally relevant host factors and pathways that play a role in the maintenance of latency.

Complementary to the approach of targeting candidate LTR-bound transcription factors and complexes for inhibition or activation, other studies have embarked on unbiased screens of small molecule libraries to identify compounds capable of reversing HIV latency (18). Alternatively, recent large-scale unbiased gene knock-out/knock-down screens have been employed to unravel the molecular mechanisms of latent HIV (19–28). RNA interference methods, which rely on the reduction of mRNA expression levels, have been widely used as screening platforms in mammalian cells (29). However, this method has been shown to suffer from serious limitations, including the presence of false-positives due to off-target effects and the persistence of residual gene expression, which results in false negatives (30, 31). For example, several RNAi screens have been performed to identify host cell factors essential for HIV infection and replication (21, 32–34); surprisingly, very little overlap was found between the lists of genes identified in these screens, pointing to the need for other screening methods that achieve complete inactivation of genes. Another large-scale unbiased gene disruption approach uses the precision of CRISPR-Cas9 targeting for complete gene knock-outs (35). A recent screen using a lentiviral sgRNA sub-library targeting nuclear proteins identified MINA53 as a possible latency-promoting gene (LPG) (28). In mammalian cells, functional analysis via forward genetic screens and mutational analysis has largely been hampered due to diploidy, as, in somatic cells, when one copy of a gene essential for a cellular process is inactivated, the second copy often remains active and compensates for that partial loss. Pseudo-haploid screens are based on KBM7s, a chronic myelogenous leukemia (CML) cell line, which is haploid for all chromosomes except for chromosome 8 and a 30Mb stretch of chromosome 15 (36), allowing for forward genetics in mammalian cells (37). Using Gene-Trap (GT) retrovirus-mediated mutagenesis for generating a library of gene knockouts in KBM7 cells enables unbiased loss-of-function screening in mutant cells for the identification of host genes essential to a specific cell function. Haploid screens have proven to be a powerful method to identify genes involved in drug import (38, 39), druggable target genes to treat cancer (40–42), key components of cellular pathways (43–47), and genes involved in the pathogenesis of various viruses (48–55) and susceptibility to toxins (37, 56, 57).

To identify novel host genes that could potentially serve as molecular targets for HIV latency reversal, we performed insertional mutagenesis in latently HIV-1-infected KBM7 cells. First, using a flow cytometry activated cell sorting-based strategy described previously (58, 59), we generated a latent pseudo-haploid KBM7/(Hap-Lat) cell line that harbors an integrated

transcriptionally silent HIV-1 5'LTR controlling the expression of a GFP reporter gene. Hap-Lat cells were then subjected to GT insertional mutagenesis, using a mCherry reporter GT virus, and Hap-Lat cells expressing both GFP (as a reporter of HIV-1 expression) and mCherry (confirming the presence of a GT integration within an active gene) were then sorted using fluorescence-activated cell sorting (FACS). GT integration sites within these two populations were then amplified by inverse PCR and mapped to the genome by high-throughput sequencing. Computational candidate identification resulted in the identification of 69 candidate genes, whose function is required for maintenance of HIV latency, but which are not essential to cell survival. For functional follow-up experiments, we selected a subset of 16 GT-identified putative HIV latency re-enforcing target genes, which are also expressed in primary CD4⁺ T cells, and examined the effect of their depletion via shRNA knock-down in maintaining latency in J-Lat T cell lines A2 and 11.1. Functional validation identified 10 putative novel latency promoting genes: ADK, CHD9, CMSS1, GRIK5, USP15, IRF2BP2, EXOSC8, NF1, EVI2B, and FAM19A. Because of the obvious mechanistic potential of CHD9, a Chromodomain Helicase DNA-binding protein, we explored its association with the 5'LTR using ChIP-qPCR analysis. We found CHD9 to be enriched at the 5'LTR in the latent state but to dissociate from the HIV-1 promoter after PMA treatment. Loss of CHD9 enrichment upon re-activation was accompanied by a relative increase in H3 acetylation at the HIV-1 LTR, indicating a functional shift in chromatin organization from repressed to transcriptionally active. We also found three genes in our candidate list, ADK, GRIK5, and NF1 to be targetable by existing small molecule inhibitors. The latency reversal potential of these compounds was evaluated in J-Lat T cell lines A2 and 11.1, and in a primary cell model of HIV latency. We found 5-Iodotubercidin, which inhibits ADK, not to be a viable LRA due to its toxicity in both A2, 11.1, and primary CD4⁺ T-cells. While the NF1 inhibitor Trametinib reactivated latent HIV in the J-Lat cell lines, it was not effective in the primary cell model of latency. On the other hand, the GRIK5 inhibitor Topiramate reversed HIV-1 latency in both J-Lat T-cell lines as well as in primary CD4⁺ T cells harboring latent HIV-1, without significant associated cytotoxicity or T cell activation, and thus presents an interesting potential novel LRA.

Results

Establishment of a latent pseudo-haploid cell line

To identify potentially druggable host genes as putative molecular targets for HIV latency reversal, we generated pseudo-haploid latent HIV-1-infected cells in which we performed insertional mutagenesis according to a strategy schematically depicted in Figure 1A. A latent HIV-1-infected pseudo-haploid KBM7 cell line was generated according to a previously described strategy used in Jurkat cells (58, 59) (Figure 1B). Subsequently, haploid latent (Hap-Lat) cells harboring a latent integrated HIV-1-derived virus containing a GFP reporter, were subjected to insertional mutagenesis using gene trap virus carrying an mCherry reporter. Instead of using a lethality-selecting scheme, our system relies on fluorescence activated cell sorting (FACS) to select for reactivated cells, marked by elevated GFP expression resulting from insertional mutagenesis of genes essential for maintenance of HIV latency. Cells expressing both GFP and mCherry were FACS-sorted and expanded for multiple rounds, after which GT insertion sites were mapped and identified by inverse PCR and high-throughput sequencing. To establish a latent HIV infection in the pseudo haploid KBM7 system, near-haploid KBM7 cells were infected at low MOI with the single infectious cycle HIV-derived virus LTR-GFP-LTR, in which GFP reporter expression is controlled by the activity of the HIV-1 promoter or 5'LTR (Figure 1A). After infection, the population of GFP negative cells comprising mainly uninfected cells and putative latently infected cells were sorted by FACS and subsequently stimulated with 5ng/μl TNF-α and 350μM Vorinostat. In response, a small percentage of existing latent HIV-infected cells were transcriptionally reactivated and expressed GFP. These GFP positive cells were -sorted by FACS as single cells and expanded (Figure 1B). The resulting clonal latent KBM7 haploid lines were characterized by flow cytometry to determine GFP expression at basal and stimulated states. From the clonal lines established, Hap-Lat#1 was selected for low basal activity and significant reactivation upon stimulation (Figure 1C). To ensure maintenance of haploidy, Haplat#1 cells were periodically sorted to enrich for the 5% of smallest cells (Supplemental figure 1A).

Gene-trap mutagenesis of Hap-Lat cells

Hap-Lat#1 cells were mutagenized by infection with a murine stem cell virus (MSCV)-derived viral gene-trap (GT) vector containing an inactivated 3' LTR, an adenoviral splice-acceptor site, an mCherry reporter cassette and a polyA terminator tail (37). Dendritic and myeloid cells are

notoriously refractory to retroviral infection (60, 61). We indeed found that infectivity of KBM7 is poor compared to T-cell-derived cell lines SupT1 and Jurkat, as well as other myeloid-derived cell lines (Supplemental Figure 1B). SAMHD1, a nucleotide scavenger, has been identified as a causative restricting host factor that limits the free available pool of nucleotides for reverse transcription (62). To bypass SAMHD1-mediated restriction, we supplemented cells with 2 μ M nucleosides (dNs), which increased GT infectivity by approximately two-fold (Supplemental figure 1C). GT preferentially inserts in the 5' regions of genes (63), effectively knocking out gene expression by truncating the native transcript (Supplemental figure 1D). For a full-scale mutagenesis experiment, approximately 200 million Hap-Lat #1 cells were mutagenized using two rounds of infection with GT-mCherry in the presence of exogenously supplied nucleosides. Infection of Hap-Lat#1 with GT-mCherry effectively caused reactivation of a subpopulation of latent KBM7 cells (Figure 1D-E). We reasoned that insertional mutagenesis in genes essential for maintenance of HIV latency would result in Hap-Lat#1 cells expressing the GFP reporter. We determined GT integrations in individual GFP/mCherry double-positive clones and estimated that the sequential infection resulted in 1 to 4 GT integrations per cell, with the majority of cells containing one integration (data not shown). By gating conservatively, approximately 1-4% GFP/mCherry double-positive cells were then sorted and expanded (Figure 1E). After expansion, reactivated cells tend to revert to a latent state (Supplemental figure 1E). To enrich for a more stable GFP-expressing, mCherry GT-containing double-positive cell population, cell sorting was repeated for multiple rounds (Figure 1F). Sequential rounds of sorting led to the appearance of a stable subpopulation within the total double-positive population expressing high levels of mCherry (Figure 1F). To examine any potential biological differences between the two, we separately sorted the total double-positive population and the mCherry-high subpopulation, which we designated GFP^{Total} and GFP^{Sub} respectively, for a final (5th) round (Figure 1F). Genomic DNA extracted from GFP^{Total} and GFP^{Sub} obtained in the 5th round of sorting was used to determine GT integrations, while that of a pool of unsorted GT-infected cells was used as a reference.

Mapping of insertion sites to identify host factors maintaining HIV latency

To determine the host sequences flanking the GT insertion sites, inverse PCR with primers annealing to internal sequences in the gene trap vector followed by amplification was performed. The amplified products were processed for high-throughput sequencing (Figure 1F). For GFP^{Sub} 2

biological replicates, samples A and B, were generated. For GFP^{Total} 3 biological replicates were generated, samples C, D and E. To estimate the sampling depth of our GT, we re-sequenced GFP^{Sub} sample B and GFP^{Total} sample D at greater depth. The resulting NGS datasets were processed for candidate gene identification. A previously described method to analyze GT data, HaSAPPY (Haploid Screen Analysis Package in Python), was rigorously re-implemented, appended with additional steps for quality control, library normalization, and optimized resolution for the selection of integration sites (64). HaSAPPY assigns a Local Outlier Factor (LOF) score to each gene in a sample based on a triplet score derived from the number of putative GT integrations in the sample compared to the reference. For each population, we compiled all genes with an LOF score >3 from each replicate and obtained 686 hits for GFP^{Total} and 382 hits for GFP^{Sub}. 183 genes were common to both populations (Supplementary figure S2A). Next, we investigated any potential biological basis for the difference between GFP^{Total} and GFP^{Sub}. Since expression levels of the integrated mCherry reporter appear to be on average higher in the GFP^{Sub} population than in the GFP^{Total} population, we wondered if expression levels of the targeted genes were higher in GFP^{Sub}. We obtained recently published KBM7 gene expression data (65) and found no substantial difference in the average level of expression between the two populations (Supplementary figure S2B). To determine if there was any difference in the functionality of the GT target genes found in GFP^{Total} and GFP^{Sub}, we performed enrichment analysis using GO terms and found no substantial differences in enrichment for biological process and molecular function ontologies (Supplementary figure S2C and D). Similarly, both populations contain a comparable fraction (32.7%) of integrations within non-coding or anti-sense genes (Supplementary figure S2E). Finally, we cross-referenced the GT target genes found in GFP^{Total} and GFP^{Sub} populations to the HIV interaction database (<https://www.ncbi.nlm.nih.gov/genome/viruses/retroviruses/hiv-1/interactions/>) and found that both the GFP^{Total} and the GFP^{Sub} populations contain similar fractions (21.9% and 20.2%, respectively) of genes previously reported to be involved in HIV biology, fractions which are substantially higher than the 7.4% found for the complete list of ENSEMBL genes (Supplementary figure S2F). In order to limit our extensive list of candidate genes for follow up functional validation, we applied more stringent thresholds for each population and defined candidate genes as having an LOF score equal to or greater than 3 in at least 2 biological replicate samples. We thus identified 19 candidate genes in the GFP^{Sub} population and 55 in the GFP^{Total} population (Figures 2A and B and Table 1).

Candidate list validation

Since our bioinformatics analysis did not reveal a defining difference between the GFP^{Total} and the GFP^{Sub} populations, we decided to proceed with functional assays using shRNA-mediated depletion of candidate genes obtained from both populations. To prioritize the candidate genes found in the KBM7 haploid screen for functional validation in the more biologically relevant Jurkat T-cell-based HIV-1 latency models J-Lat A2 and 11.1, we focused on protein-coding genes and took into account LOF scores as well as gene expression in white blood cells (as extracted from the GTEx portal (<https://www.gtexportal.org/home/>) or Illumina's Human BodyMap 2.0 project (<http://www.ensembl.info/2011/05/24/human-bodymap-2-0-data-from-illumina/>)) and the Jurkat T-cell model (66) (Table 1). Eight genes from GFP^{Sub} and nine genes from GFP^{Total} were selected as candidates for validation using shRNA-mediated depletion in J-Lat A2 and J-Lat 11.1 cells. These cells contain a latent HIV-derived GFP reporter driven by the 5'HIV-LTR and are well-established model systems for HIV latency (58, 67). While J-Lat A2 cells contain an integrated latent LTR-Tat-IRES-GFP virus, J-Lat 11.1 cells contain an envelope defective full-length HIV-1 genome expressing GFP in place of Nef. We used flow cytometry to assess reactivation as measured by GFP expression and extracted RNA to assess knockdown of the targeted candidate genes and expression of the HIV genes GAG, POL, and TAT by RT-qPCR (figure 2C-E and Supplemental figure S3 and S4). We observed significant latency reversal upon candidate gene knockdown in 10 of 15 genes depleted by shRNA (ADK, CHD9, CMSS1, EVI2B, EXOSC8, FAM19A, GRIK5, IRF2BP2, NF1, and USP15) in both J-Lat A2 and J-Lat 11.1 cells (Figure 2C-E and Supplemental figure S3 and Table 1). Knockdown of SCN9A, RHOF, SPN, COPS5 and EVL did not result in significant latency reversal in one or both J-Lat models (Supplemental figure S4). These results demonstrate that a significant proportion of the candidate genes found in our myeloid-derived KBM7 haploid screen play a role in maintenance of HIV-1 latency in the more relevant T cell-derived J-Lat A2 and J-Lat 11.1 HIV latency models.

CHD9 is an LTR-associated repressor of HIV-1 transcription

Interestingly, two genes, CIITA and CHD9, from our candidate list are associated with the GO-term DNA binding (GO:0003677). CIITA is a well-established factor involved in HIV expression and has been previously shown to inhibit Tat function and hence viral replication (68, 69). The

Chromodomain helicase DNA binding protein 9 (CHD9) is a member of an ATP-dependent chromatin remodeler family, the members of which modulate DNA-histone interactions and positioning of nucleosomes and play key roles in stem cell regulation, development, and disease (70). Previously, we have shown that chromatin remodeling by another ATP-dependent remodeler, the BAF complex, plays a crucial role in maintenance of HIV-1 latency and its re-activation (59). We therefore decided to further characterize the role of CHD9 in regulating HIV-1 gene expression. We knocked down CHD9 using a lentivirally transduced shRNA and verified its depletion in both J-Lat A2 and J-Lat 11.1 cells at the protein level by Western blotting (Figure 3A). Depletion of CHD9 led to a significant reversal of latency, as shown by an increase in the percentages of GFP positive cells (Figure 3B and C). Latency reversal was also confirmed by increased expression of viral genes Gag, Pol, and Tat in J-Lat 11.1 (Figure 3D). To characterize a potential direct association of CHD9 with the latent HIV-1 5'LTR, we performed chromatin immuno-precipitation (ChIP) in latent and PMA-activated 11.1 J-Lat cells. The positions of nucleosomes within the latent HIV-1 LTR are rigid, with positioned nucleosomes Nuc-0, Nuc-1, and Nuc-2 separated by DNase I-sensitive regions HSS1 and HSS2, respectively, visually summarized in Figure 3E. CHD9 was found to be enriched throughout the HIV-1 LTR in latent 11.1 J-Lat cells, predominantly present over the Nuc0-HSS1 region, and this association was significantly decreased after LTR activation by PMA treatment (Figure 3F and Supplemental figure S5). No significant enrichment was observed at the HIV-1 internal control region (the GFP reporter) or the unrelated gene locus (HK2) (Figure 3F and Supplemental figure S5). As expected, PMA treatment led to an increase in histone H3 acetylation, a mark of active chromatin, over the HIV-1 5'LTR (Figure 3G and Supplemental figure S5). Our data demonstrate a role for CHD9 in the repression of HIV-1 gene expression and maintenance of latency via direct recruitment to and association with the HIV 5'LTR.

Pharmacological targeting of ADK, GRIK5 and NF1

With the aim of identifying potential latency reversing agents (LRAs), we performed a literature search and identified three candidate genes, ADK, GRIK5, and NF1, all present in the GFP^{Sub} list, for which FDA-approved small molecule inhibitors are available. We therefore examined the latency reversal potential of the ADK inhibitor 5-Iodotubercidin, the GRIK5 inhibitor Topiramate, and the NF1 inhibitor Trametinib. Adenosine kinase (ADK) is a phosphotransferase that converts

adenosine into 5'-adenosine-monophosphate and thus plays a major role in regulating the intracellular and extracellular concentrations of adenosine, activation of specific signaling pathways, and bioenergetic and epigenetic functions (71, 72). 5-Iodotubercidin is a purine derivative that inhibits adenosine kinase by competing with adenosine for binding to the enzyme (73). Glutamate ionotropic receptor kainate type subunit 5 (GRIK5) is a subunit of the tetrameric kainate receptor (KAR), a subgroup of ionotropic glutamate receptors. GRIK5, together with GRIK4, binds glutamate, whereas subunits GRIK1-3 form functional ion-channels (74). Topiramate is an FDA-approved GRIK5 inhibitor employed as an anti-epileptic drug and is used to manage seizures and prevent migraines (75). Neurofibromin 1(NF1) is ubiquitously expressed; however, its highest levels are found in cells of the central nervous system and it has been described to function as a negative regulator of the Ras signal transduction pathway (76, 77). Trametinib is an NF1 inhibitor which is also known as a mitogen-activated protein kinase (MAPK) kinase (MEK) inhibitor with anticancer activity and is FDA-approved for use in metastatic malignant melanoma (78).

We examined the effects of treatment with 5-Iodotubercidin, Topiramate, and Trametinib in J-Lat 11.1, J-Lat A2 cells, and a primary CD4⁺ T cell model of latency, on latency reversal. Treatment with the ADK inhibitor 5-Iodotubercidin resulted in only a moderate induction of latency reversal in both A2 and 11.1 J-Lat cells and was accompanied by significant toxicity, especially at high concentrations [16μM] (Figure 4A & B). Only at a 5-Iodotubercidin concentration of 4μM in 11.1 J-Lat cells was a significant increase in GFP positive cells with acceptable viability observed (Figure 4A & B). Treatment of J-Lat A2 and 11.1 cells with the GRIK5 inhibitor Topiramate or the NF1 inhibitor Trametinib resulted in significant, concentration-dependent latency reversal (Figure 4A & B). Importantly, neither Topiramate nor Trametinib displayed significant toxicity in comparison with the DMSO vehicle control group, as measured by gating for live cells. In particular, treatment of J-Lat 11.1 cells with Trametinib resulted in significant latency reversal at all evaluated concentrations [250nM, 1μM, 4μM and 16μM], with minimal associated toxicity (Figure 4B).

Next, we evaluated the latency reversal potential of the inhibitors in a more clinically relevant primary ex vivo infection latency model, in which CD4⁺ T cells are infected with a full-length non-replication competent HIV-1 virus driving expression of a luciferase reporter (Figure 4C) (79). After spin infection, cells are allowed to rest for 4 days and are treated with the different

compounds, as indicated, for 48 hours, followed by measurement of HIV-1 LTR-driven luciferase activity and staining for Annexin positive cells and T-cell activation markers CD69 and CD25 (Figure 4D- G and Supplemental figure S6). Similar to what was observed for the A2 and 11.1 J-Lat cell-lines, treatment of HIV-1-infected latent primary CD4⁺ T cells with the ADK-inhibitor 5-Iodotubercidin did not result in significant latency reversal (Figure 4D), while it produced significant toxicity at the highest concentration used, as indicated by Annexin V staining (Figure 4E). Treatment with the GRIK5 inhibitor Topiramate, at a concentration of 16 and 20 μ M, resulted in significant ($p < 0.05$) reversal of latency, as measured by a 3-6 fold increases in mean luciferase activity compared to untreated controls (Figure 4D). Topiramate treatment did not significantly affect CD4⁺ T cell viability after 48 hours, while control treatment with a toxic (200nM) concentration of Gliotoxin (GTX) caused apoptosis of primary CD4⁺ T-cells, as evidenced by an increase in the percentage of Annexin V-positive cells, as was observed previously (Figure 4E) (80). In contrast to what was observed in the A2 and 11.1 J-Lat cell-lines, treatment of latently infected CD4⁺ T-cells with the NF1 inhibitor, Trametinib did not lead to a significant reversal of latency (Figure 4D), while viability was only moderately affected (figure 4E).

For LRA candidates to be viable in a clinical setting, it is crucial that they reactivate latent HIV without inducing activation of CD4⁺ T-cells. Therefore, we examined the potential of our candidate LRAs to induce expression of the activation markers CD25 and CD69 in treated CD4⁺T cells. While the ADK inhibitor 5-Iodotubercidin resulted in markedly elevated expression of the activation markers, the GRIK5 inhibitor Topiramate and NF1 inhibitor Trametinib did not induce significant induction of the T cell activation markers CD69 and CD25 (Figure 4F & G and Supplemental figure S6). As expected, PMA/Ionomycin treatment substantially activated T cells (Figure 4F & G and Supplemental figure S6). Taken together, our data indicate that the FDA-approved GRIK5 inhibitor Topiramate reverses HIV-1 latency in a variety of T cell models of latency, without induction of T cell activation and with limited cytotoxicity, and, therefore, can be considered an attractive LRA for further mechanistic and pre-clinical investigation.

Discussion

In search of potentially novel host factors and pathways that play a role in the maintenance of HIV-1 latency, we performed a two-color haploid genetic screen in latent HIV-1-infected KBM7 cells. An important advantage presented by this approach is that identification of putative functionally relevant candidate latency-promoting host target genes does not require a-priori knowledge of the molecular determinants of latency and is thus completely unbiased. Additionally, gene-trap insertional mutagenesis has enabled identification of previously unappreciated latency genes and cellular pathways, which likely modulate latency not only via direct physical association with the HIV-1 promoter, but also indirectly, through involvement in cellular signaling. We produced a list of 69 candidate genes and proceeded to validate ten candidates, the depletion of which in various in vitro models of latency, including primary CD4+ T-cell models, led to latency reversal.

The haploid KBM7 cell line, while a powerful system for GT-mediated forward genetic screens, is a myeloid cell line (36). HIV can infect and establish latent infection in monocytes and macrophages (81, 82), but these are not considered to be the prime source of the latent reservoir; this begs the question how relevant the myeloid nuclear environment may be for HIV latency in lymphocytic cells. Nevertheless, when we tested a selected subset of the candidate genes by shRNA-mediated knockdown in a T-cell derived model of HIV latency, we found that knockdown of 67% of candidates (10 out of 15) led to latency reversal, which demonstrates the validity of our approach. Moreover, several of the genes identified in our gene trap screen have previously been implicated in HIV susceptibility: 15 out of 69 genes are listed in the HIV-1 human interaction database (<https://www.ncbi.nlm.nih.gov/genome/viruses/retroviruses/hiv-1/interactions/>), a detailed database of all known interactions between HIV-1 and the human host. Furthermore, one of the candidates, IRF2BP2, is a potential target of HIV-associated nucleotide polymorphisms within a cluster of regulatory DNA elements (66) which loop to and potentially regulate the IRF2BP2 promoter in CD4+ T-cells (83, 84). In the current study, we demonstrate that knock-down of IRF2BP2 results in latency reversal. IRF2BP2 has been shown to interact with NFAT1 and to repress transcriptional activity (85), providing a plausible mechanism for its role in latency maintenance.

Among the candidate gene list, we also identified CHD9, a member of the chromodomain helicase DNA-binding (CHD) family of the ATP dependent chromatin remodelers. Members of this family are involved in various cellular processes and in normal development and disease; however, CHD9

is one of the least-studied members. We found that CHD9 is associated with the latent HIV-1 5'LTR and is displaced upon promoter activation by PMA stimulation, suggesting that it acts as a repressor of HIV transcription. Indeed, depletion of CHD9 by shRNA-mediated knockdown led to de-repression of HIV, as observed by increase in the expression of the HIV-1 LTR-driven reporter GFP as well as the HIV-1 genes Gag, Pol and Tat. Future studies will determine the mode of recruitment of CHD9 to the HIV-1 LTR and the molecular interplay between CHD9 and other chromatin remodeler and modifying complexes associated with the latent HIV-1 5'LTR (4).

A large subset of the list of 69 candidate genes are non-coding RNAs (N=22); upon closer inspection we found that 16 of those are at least in part oriented in an anti-sense direction with respect to known protein-coding genes. We currently do not know if these non-coding transcripts fulfill a biological function in maintenance of HIV-1 latency, directly through their transcripts, through regulatory effects on other (protein-coding) genes in cis or in trans, through other effects, or if they represent mapping artefacts. Further investigation into these possibilities is ongoing.

An interesting observation emerging from our experimental set-up is the appearance of a sub-population of cells after multiple rounds of sorting that is more stable in its GFP expression. This sub-population is also higher in mCherry expression, as compared to the total double-positive population. The reversal of double-positive cells after sorting may reflect the intrinsically stochastic transcription of HIV-1(86–88). Latent cell lines are notoriously sensitive to cellular stresses, which cause reactivation (89–91). It is, therefore, possible that the less stable GFP-positive cells represent cells that are temporarily activated, and which slowly revert back to a latent state. Therefore, we focused our analysis and validation experiments mainly on the stable GFP-high population. Nevertheless, the few candidate genes unique to the total population that we tested in our shRNA knockdown validation experiment had a similar false positive ratio (i.e. 33.3%) as we found for the candidate genes from the stable GFP-positive population.

For our analysis, we set a strict criterion of the LOF score being ≥ 3 in at least two biological replicates, although a cut-off at lower LOF scores of 2 or above has been used previously (92). We chose this stringent criteria to confidently identify 69 candidate genes. We believe however that our dataset likely contains additional candidate genes that we miss due to the stringency applied. This is exemplified by our observation that STRING analysis indicates that many of the 598 protein-coding genes with an LOF score of 3 and higher functionally interact (Supplemental Figure

S7). Moreover, 32% of these genes (195) appear in the HIV-1 human interaction database, pointing to their potential roles in HIV-1 biology. Importantly, proteins with well-described functions in HIV biology, such as, e.g., IL32 (LOF=9.5; (66), RUNX1 (LOF=12.6; (93), and SMARCA4/BRG1 (LOF=8.7; (59), are detected with high LOF scores in at least one of our samples.

Identification of functionally relevant molecular targets and novel compounds effective in latency reversal has proven to be an outstanding challenge in the field (94). Most clinically investigated LRAs have thus far failed to significantly impact the size of the latent reservoir in patients. From our list of validated host factors we identified three druggable targets, ADK, NF1, and GRIK5. The ADK inhibitor 5-Iodotubercidin displayed cytotoxicity in both J-Lat and primary models of HIV latency. Since knockdown of ADK did not lead to significant toxicity, off target, pleiotropic effects of 5-Iodotubercidin may underlie the observed T cell toxicity. 5-Iodotubercidin is the prototype nucleoside ADK inhibitor, with a relatively high IC₅₀. Novel nucleoside and non-nucleoside ADK inhibitors have been developed with more favorable pharmacological properties that would be attractive to pursue in the context of HIV-1 latency reversal in future studies (72). Trametinib, a small molecule inhibitor of NF1, reactivated HIV-1 in J-Lat cell models with modest effects on viability but displayed significant toxicity in the primary cell models. These results highlight the importance of interrogating the potential effectiveness of candidate LRAs in multiple models of HIV-1 latency to circumvent potential integration-site mediated biases present in cell line models of latency, and, most importantly, to confirm effectiveness in the more relevant primary CD4⁺ T cells harboring latent HIV-1 in vivo.

Our data point to the GRIK5 inhibitor Topiramate as a potentially promising compound for latency reversal. Topiramate, effectively reversed latency in primary HIV-1-infected CD4⁺ T cells without inducing significant T cell activation or cytotoxicity; this makes Topiramate a potentially clinically promising LRA and a target for further investigation. GRIK5 (glutamate ionotropic receptor kainate type subunit 5), primarily studied in neurons, is a subunit of the tetrameric kainate receptor (KAR), a subgroup of ionotropic glutamate receptors. GRIK5, together with GRIK4, bind glutamate, whereas subunits GRIK1-3 form functional ion-channels (74). In B-cells, KAR activation by glutamate increases ADAM10 levels, leading to increased B cell proliferation and immunoglobulin production (95). Topiramate is primarily used as an anticonvulsant or antiepileptic drug. While the exact mechanism by which Topiramate exerts anticonvulsant or

antiepileptic properties is unclear, it has been shown to block voltage-dependent sodium and calcium channels (73, 96), to inhibit the excitatory glutamate pathway and enhance inhibition by GABA (97). Interestingly, Topiramate can induce cytochrome P450 family member CYP3A4 activity and potentially negatively affect the metabolism of many drugs (98), which should be taken into account when considering a potential therapeutic combination of LRAs in future studies.

Acknowledgments

TM received funding from the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC STG 337116 Trxn-PURGE, Dutch AIDS Fonds grant 2014021, and Erasmus MC mRACE research grant. RJP received funding from Dutch AIDS Fonds grant 2016014. We thank Thijn Brummelkamp for providing materials and expertise.

Conflicts of interest

The authors have no conflict of interest

Author Contributions

MR, MMS, EN, MS, EDC, HB, TWK, MA, RJP and TM carried out experiments and performed data analysis. PM, VH, and PH performed Nextgen sequencing and data analysis. MR, EN, MMS, RJP and TM conceived the study and wrote the manuscript. All authors read and approved the final manuscript.

Materials and Methods

Cell culture

KBM7 (a kind gift from Thijn Brummelkamp) and Hap-Lat (latent HIV infected KBM7-derived cell lines) were cultured in IMDM media (ThermoFisher Scientific) supplemented with 10% FCS and 2% Pen/Strep. Haploidy of KBM7 cells was maintained by periodically sorting cells for size (5% smallest). Ploidy of cells was determined by propidium iodide staining

Establishment of haploid latent (Hap-Lat) HIV infected cell lines

We used a strategy described previously (58, 59) to generate latently HIV infected KBM7 haploid cell lines. Minimal HIV (LTR-GFP, HIV-658 (99) virus was produced by transfection of 293T cells in 15cm culture dishes using polyethylenimine (PEI) Sigma Aldrich) with a mixture of 6.8µg p658, 2µg VSVG and 4.5µg Gag-pol plasmids. Haploid KBM7 cells were infected with HIV-derived virus at low MOI such that approximately 5-10% of cells became productively infected as determined by GFP expression. 5 days after infection, the GFP negative cell population harboring uninfected and potentially latently infected cells were sorted by flow cytometry activated cell sorting (FACS) and stimulated with 350nM SAHA (Selleck Chem), and 5ng/µl TNFα (Sigma-Aldrich). Twenty-four hours post stimulation, GFP positive cells were single cell sorted by FACS into 96 well plates. Clones were expanded and characterized for their basal GFP expression and their potential for re-activation. From the clonal cell lines generated Hap-Lat #1 was selected for low background and relative high reactivation upon stimulation (GFP negative under basal conditions but inducible to express GFP upon activation).

Gene Trap virus production and mutagenesis of Hap-Lat cell lines

We adapted a strategy described previously (37) to mutagenize Hap-Lat cells. Briefly, Gene Trap virus was produced by transfection of HEK293T cells in 15cm culture dishes using 180 µl polyethylenimine (PEI) with a mixture of 8 µg pGT-mCherry, 2.1 µg pAdvantage, 3.1 µg CMV-VSVG and 4.8 µg Gag-pol plasmids in a total volume of 1ml serum free RPMI. After 12 hours the growth media was changed to fresh FCS supplemented RPMI. The virus was collected at 12, 24, 36 and 48 hours. For the GT mutagenesis, 192 million Hap-Lat #1 cells were pre-incubated with 100µM dNTPs (Invitrogen) 1hr before transduction. Cells are resuspended in 192ml of undiluted GT virus. Spin infection was performed for 90minutes @ 1500rpm in 24 well plates

containing 2 million cells per well and 8µl protamine sulfate per ml. After 24 hours cells were subjected to a second round of infection after which virus was removed, and cells were left to recover in supplemented IMDM. We estimate that this sequential infection resulted in 1 to 4 GT integrations per cell, with the majority of cells containing one integration (data not shown). A first round of sorting was performed 8 days after the second infection. Subsequently, between 500.000 and 1 million GFP-mCherry double positive cells were sorted each day for 5 consecutive days using two cell sorters (BD FACSAriaII SORP and BD FACSAriaIII). Sorted cells were pooled, left to recover and expanded for the next round of sorting. A second round of sorting was performed 17 days after the second infection. Between 2.6 and 3 million cells were collected each day for 3 consecutive days using two sorters. These cells were pooled and left to recover and expand for a third round of sorting on day 22 after the second infection, yielding 6.1 million cells from 1-day sorting on two machines. On day 29 after the second infection a fourth round of sorting was performed, yielding 7.2 million cells from 1-day sorting on two machines. After the cells of the fourth round of sorting were left to recover and expand a sub-population of highly mCherry and GFP double positive cells were apparent. In a final fifth round of sorting the total GFP-mCherry positive population (10.1 million cells) and the GFP-mCherry positive high sub-population (5.9 million cells) were respectively sorted on day 33 and 35 after the second infection. After recovery, genomic DNA was isolated from the sorted cell populations, as well as from a population of mutagenized unsorted cells.

Mapping insertion sites through inverse PCR

A previously described inverse PCR protocol was adapted to determine host sequences flanking the proviral insertion sites (37). Briefly, genomic DNA was isolated from 5 million cells using the DNAeasy kit (Qiagen). 4µg gDNA was digested with NlaIII or MseI. After PCR spin column purification (Qiagen), 50µl of eluted digested DNA was ligated using T4 DNA ligase (Roche) in a volume of 2ml. The reaction mix was purified using spin columns and used in as template in a PCR reaction with primers annealing to internal sequences in the gene trap vector (5'-CTGCAGCATCGTTCTGTGTT-3' and 5'-TCTCCAAATCTCGGTGGAAC-3'). The resulting PCR products were used for library preparation.

High-throughput sequencing and identification of integration sites

Sequencing libraries were created using the Ion Plus Fragment Library Kit (ThermoFisher Scientific) according to manufacturer's instructions, with minor modifications: briefly, 15ng of the PCR products were diluted with ddH₂O to a final volume of 39 µl. The samples were end-repaired, adaptor ligated and amplified; this was followed by 2 rounds of purification using Agencourt AMPure XP beads. Library qualities and quantities were assessed by Bioanalyzer, using the DNA High Sensitivity Kit (Agilent Technologies). The quantified libraries were pooled together in 10plex, at a final concentration of 40pM. Templating, enrichment and chip loading was performed on an Ion Proton Chef system using the Ion PI™ Hi-Q™ Chef Kits (ThermoFisher Scientific); sequencing was performed on an Ion Proton PI™ V3 chip, with the Ion PI™ Hi-Q™ Sequencing 200 Kit, on an Ion Proton instrument (ThermoFisher Scientific), according to manufacturer's instructions.

The resulting FASTQ files were processed to remove the gene trap vector primers. To this end, we used R/Bioconductor packages ShortRead (100) and Biostrings to match primer sequences within the sequenced reads, allowing for two mismatches at maximum for each primer sequence. The identified primers were trimmed and reads shorter than 15bp were discarded. The remaining reads were aligned to the human reference genome version hg19 retrieved from the Illumina iGenomes repository, using bwa mem. The produced aligned reads (BAM files) were then subjected to integration site identification. To this end, we re-implemented the HaSAPPy algorithm (64) in R. The re-implemented version operates on BAM instead of SAM files and was enriched with additional noise filtering steps, which improved the overall process of detecting integration sites using the Local Outlier Method (LOF), as described in the original HaSAPPy implementation. Specifically, we added attributes for i) filtering out read artifacts from introns, based on the overall intronic read distribution for each sample, ii) selecting the fraction of neighboring (Independent Insertions, Disrupting Insertions, Bias) triplets for LOF analysis and iii) library normalization across samples. After the quality control steps, our HaSAPPY re-implementation assigns a Local Outlier Factor (LOF) score to each gene in a sample based on a triplet score derived from the number of putative GT integrations in the sample compared to the reference, additionally taking into account the potential bias arising from unbalanced sense and anti-sense read numbers. LOF is a metric to indicate to what extent a vector measurement deviates from a population of vectors with the same properties. In our case, a triplet vector corresponds to a gene and LOF measures the

deviation from the respective population, i.e. a gene is an outlier based on a distribution of vectorized scores, such as the aforementioned triplet. The re-implemented version can be found in <https://github.com/pmoulos/ngs-stone-age/blob/master/R/hasar.R>.

ShRNA knock-down of candidate genes in A2 and 11.1 cell lines

Pre-designed shRNA sequences (MISSION® shRNA library (Sigma); Table 2) were purchased as bacterial glycerol stocks from the Erasmus Center for Biomics. Virus was produced as follows, 3×10^5 HEK293T cells/ml were plated in a 10 cm dish cells one day before transfection and transfected with 4,5 µg of pCMVΔR8.9 (envelope), 2 µg of pCMV-VSV-G (packaging) and 6ug of shRNA vector. The transfection mixture was removed after 12 hours and replaced with fresh RPMI medium contain 10% FBS. Virus containing supernatant was collected at 36, 48, 60- and 72-hours post-transfection. Jurkat, J-Lat A2 (LTR-Tat-IRES-GFP) and J-Lat 11.1 (integrated full-length HIV-1 genome mutated in env gene and GFP replacing Nef) cells were cultured in RPMI-1640 medium (Sigma) supplemented with 10% FBS and 100 µg/ml penicillin-streptomycin at 37 °C in a humidified 95% air-5% CO₂ atmosphere. A2 and 11.1 cell lines were infected by adding 500 µl of filtered lentivirus to 2.5 ml of cell culture. After 48 hours, medium was refreshed and cells were puromycin selected. 14 days after infection, RNA was harvested to determine knock-down of genes by qPCR and reactivation of latent HIV was measured by flow-cytometry and RT-qPCR for Gag, Pol and Tat.

Flow cytometry for GFP expression in the J-Lat cell lines

Cells were collected in PBS. GFP fluorescent signal indicating latency reversal was monitored using a LSRFortessa (BD Biosciences). Viability was determined using the forward scatter area versus side scatter area profiles (FSC-A, SSC-A). Data was analyzed using FlowJo software (version 9.7.4, Tree Star).

Total RNA Isolation and Quantitative RT-PCR (RT-qPCR)

Total RNA was isolated from transduced A2 and 11.1 cells using Trizol (Thermo Fisher) on day 14 after infection using Total RNA Zol out kit (A&A Biotechnology), residual genomic DNA was digested with DNaseI (Life technologies). cDNA was synthesized using superscript II Reverse Transcriptase (Life Technologies) using oligo (dT) primers or random primers (for Gag, Pol and

Tat). RT-qPCR reactions were performed on a CFX Connect Real-Time PCR Detection System thermocycler (BioRad) using GoTaq qPCR Master Mix (Promega) (3 min at 95 °C, followed by 40 cycles of 95 °C for 10 s and 60 °C for 30 s). Melting curve analysis was performed to assess specificity of RT-qPCR products. Primers used for real-time PCR are listed in Table 3. Expression data was calculated using 2- $\Delta\Delta C_t$ method (101). β -2-microglobulin (B2M) and GAPDH were used as housekeeping genes for the analysis.

Isolation and ex vivo infection of primary CD4⁺ T cells

HIV-1 latency *ex vivo* model was generated based on Lassen and Greene method by spinoculation (90 min at 1200 g), washed in PBS and cultured in RPMI 1640 containing 10% FBS, 100 μ g/ml penicillin-streptomycin, IL2 (5ng/ml) and the antiretroviral drug Saquinavir Mesylate (5 μ M) (79). HEK 293T cells were transfected with HXB2 Env and pNL4.3.Luc.R-E- plasmids using PEI (Polyethylenimine) as described above. Supernatants containing pseudovirus particles are collected at 24, 48 and 72 h post-transfection. Peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll density gradient sedimentation of buffy coats from healthy donors. Total CD4⁺ T cells were separated by negative selection using RosetteSep™ Human CD4⁺ T Cell Enrichment Cocktail (StemCell Technologies). Primary CD4⁺ T cells were plated at a concentration of 1.5x10⁶ cells/mL left overnight for recovery and infected. Two days after infection, cells were collected, washed once with PBS and treated with 5-Iodotubercidin, Topiramate or Trametinib. Cells were collected in 48 hr after treating, washed once in PBS and luciferase activity was measured using Luciferase Assay System (Promega).

Flow cytometry for T cells activation and toxicity assay

Primary CD4⁺ T cells isolated from buffy coats of healthy volunteers were treated with different concentration of 5-Iodotubercidin, Topiramate and Trametinib and PMA/ionomycin as a positive control.

After 24 and 48 hours, cells were stained with Annexin V to examine the percentage of cells undergoing apoptosis and with the surface receptor CD69 and CD25 to measure T cell activation. For staining, 10⁶ cells were washed with PBS supplemented with 3% FBS and 2.5 mM CaCl₂ followed by staining with Annexin V-PE (Becton and Dickinson), CD69-FITC (eBioscience) and CD25-APC (eBioscience) for 20 minutes at 4°C in the presence of 2.5 mM CaCl₂. Cells were then

washed with PBS/FBS/CaCl₂ and analyzed by flow cytometry. Between 2 - 4x10⁵ events were collected per sample within 3 hours after staining on a LSRFortessa (BD Biosciences, 4 lasers, 18 parameters) and analyzed using FlowJo software (version 9.7.4, Tree Star).

Western Blotting

10x10⁶ cells were lysed for 30 minutes on ice with 200 ul lysis buffer (150 mM NaCl, 30 mM Tris (pH 7.5), 1 mM EDTA, 1% Triton X-100, 10% Glycerol, 0.5 mM DTT, protease inhibitor cocktail tablets (EDTA-free) (Roche). Cell lysates were clarified by centrifugation (14,000 rpm for 30 min at 4°C), mixed with 4x sodium dodecyl sulfate (SDS)-loading buffer containing 0.1M DTT and boiled at 95°C for 5 min. Samples were run in a 10% SDS-polyacrylamide gel at 100V. The proteins were transferred to polyvinylidene difluoride (PVDF) membranes and the membranes were probed with anti-CHD9 (13402-1-AP, Proteintech) and anti- α -tubulin (ab6160, Abcam). The next day, blots were incubated with HRP conjugated secondary antibody and proteins were imaged by chemical luminescence using Super Signal West Pico (Thermo Scientific).

Chromatin immunoprecipitation (ChIP) and qPCR

Approximately 50 to 100 million J-Lat 11.1 cells per condition were collected and crosslinked with 1% formaldehyde (Polysciences inc.) in 40 ml PBS, supplemented with 1mM MgCl₂ and 1mM CaCl₂, at RT for 30 min with 15rpm vertical rotation. The reaction was quenched with 125 mM Glycine. Cross-linked cells were pelleted at 800rcf for 5 minutes RT and subjected to chromatin enrichment as described previously (102). Briefly: the pellets were re-suspended in 2 ml ChIP incubation buffer (1% SDS, 1% Triton x-100, 150mM NaCl, 1mM EDTA, pH 8.0, 0.5 mM EGTA, 20mM HEPES pH 7.6, protease inhibitor cocktail tablets (EDTA-free, Roche) containing 1% SDS and sonicated (30' ON/30' OFF intervals Diagenode Bioruptor plus) to obtain chromatin fragments between 300 and 500bp. Sonicated chromatin was spun at 20817 rcf at 4°C for 30 minutes and diluted 10 times in ChIP incubation buffer without SDS. Diluted chromatin corresponding to 30-40 million cells was precleared overnight at 4°C with vertical rotation at 15rpm with 80uls of Protein A Sepharose™ 4 Fast flow (GE Healthcare) beads and immunoprecipitated overnight with 40uls of beads and 5ug of antibody (13402-1-AP, Proteintech). Samples were washed two times per each buffer (5minutes, 15rpm vertical rotation) with buffer 1 (0.1% SDS, 0.1% DOC, 1% Triton x-100, 150mM NaCl, 1 mM EDTA pH 8.0, 0.5 mM EGTA,

20mM HEPES pH8.0.), buffer 2 (500mM NaCl: 0.1% SDS, 0.1% DOC, 1% Triton x-100, 500mM NaCl, 1 mM EDTA pH 8.0, 0.5mM EGTA, 20mM HEPES pH8.0), buffer 3 (0.25M LiCL, 0.5%DOC, 0.5% NP-40, 1mM EDTA pH 8.0, 0.5mM EGTA pH 8.0, 20mM HEPES pH 8.0), buffer 4 (1mM EDTA pH8.0, 0.5mM EGTA pH 8.0, 20mM HEPES pH 8.0) to remove unspecific binding. Finally, each sample was eluted in 400ul elution buffer (1% SDS, 0.1M NaHCO₃) de-crosslinked overnight at 65°C in presence of 200mM NaCl, phenol-chloroform extracted, ethanol precipitated and subjected to qPCR analysis using the primer sets summarized in Table 3.

Figure legends

Figure 1 Schematic representation of the two-color haploid screening strategy for identification of novel host factors and cellular pathways involved in the maintenance of HIV latency.

(A) Scheme depicting the generation of a clonal latent haploid KBM7 cell line. To this end, haploid KBM7 cells were infected with a minimal HIV virus harboring GFP (HIV-1 658). After stimulation, reactivated cells were sorted and left to expand and revert to a latent state.

(B) FACS plots depicting the establishment of haploid HIV-1 latently infected KBM7 cell-line (Hap Lat #1). Parental KBM7 cells were infected with a minimal HIV virus carrying a GFP reporter (HIV-1 658). GFP negative cells, consisting of uninfected and latently infected cells, were sorted by FACS. The polyclonal cell pool was stimulated with a cocktail of latency reversal agents (LRAs) and reactivated cells were clonally sorted by FACS and expanded to generate haploid latent cell lines.

(C) Hap-Lat #1 displays low basal activity (GFP expression) but is effectively reactivated using LRAs.

(D) Hap-Lat#1 cells were mutagenized by infection with a Gene Trap (GT) virus infection harboring a mCherry reporter. Cells infected with the GT reporter will be mCherry positive (red asterisk). Latently infected KBM7 cells that reactivate following GT mutagenesis will be double positive for GFP and mCherry (Green cells, red asterisk).

(E) Representative FACS plots demonstrating gating strategy for sorting double positive cells (GFP, mCherry).

(F) Double positive cells (GFP, mCherry) are sorted in multiple rounds to eliminate cells stochastically reverting back to a GFP negative state. During these rounds of sorting, a stable and distinct double positive sub-population (GFP Sub) appears which was sorted separately.

Figure 2 Identification and validation of candidate host factors.

(A) LOF scores of genes in the GFP^{sub} population (samples A, B1 and B2), validated candidates are indicated. Genes with LOF>3 are in large font, underlined genes comply to our candidate gene selection criteria and have a LOF>3 in at least two biological replicates within either the GFP^{sub} or GFP^{total} population while genes with LOF<3 but complying to our selection criteria based on other samples are depicted in small font and italics.

(B) LOF scores of genes in the GFP^{total} population (samples C,D1,D2 and E), Markings as in Figure 2A.

(C) Functional validation of candidate hits ADK, GRIK5, NF1 by shRNA mediated depletion followed by determination of latency reversal by flow-cytometry and RT-PCR in latently infected J-Lat A2 (left panels) and 11.1 cells (right panels). Flow-cytometry bar plots: green bars show the percent of GFP positive cells after knockdown over control (black bar), left y-axis, whereas gray bars show cell viability, right y-axis. Viral reactivation is confirmed by RT-qPCR for viral genes Tat, Gag and Pol in J-Lat 11.1 cells. Statistical significance was calculated using ratio-paired t-test and multiple comparison t-test on Log2 transformed fold changes * – $p < 0,05$, ** – $p < 0,01$, *** – $p < 0,001$,

Figure 3 CHD9 regulates HIV-1latency in J-Lat A2s and J-Lat 11.1 cells.

(A) Western blot for CHD9 in control and CHD9 shRNA depleted J-Lat 11.1 and A2 cells, α -Tubulin as a loading control.

(B) Flow-cytometry bar plots demonstrating latency reversal after CHD9 shRNA depletion in J-Lat 11.1 cells: green bars indicate the percent of GFP positive cells after knockdown over control (black bar), left y-axis, whereas gray bars show cell viability, right y-axis.

(C) Latency reversal after CHD9 shRNA depletion in J-Lat A2 cells.

(D) Viral reactivation is confirmed by RT-qPCR in J-Lat 11.1 cells for the viral genes Tat, Gag and Pol. Data are normalized to GAPDH and represented as fold increase over sh Control. Statistical significance was calculated using ratio-paired t-test and multiple comparison t-test on Log2 transformed fold changes * – $p < 0,05$, ** – $p < 0,01$, *** – $p < 0,001$.

(E) Schematic of HIV genome. 5' LTR region further segmented into the U3, R, and U5 regions. Amplicons used in ChIP-qPCR experiments are indicated.

(F) ChIP-qPCR analysis of CHD9 binding to the HIV-1 5'LTR in untreated and PMA stimulated J-Lat 11.1 cells. Data are represented as percentage of the input.

(G) ChIP-qPCR analysis of Histone H3 occupancy at the HIV-1 5'LTR untreated and PMA stimulated J-Lat 11.1 cells. Data are represented as percentage of the input.

(H) ChIP-qPCR analysis of Histone H3 acetylation at the HIV-1 5'LTR in untreated and PMA stimulated J-Lat 11.1cells. (E-H) Data represent the average ($\pm$ SD) of two technical replicates.

Figure 4 HIV latency reversal by small molecule inhibitors of three candidate genes ADK, GRIK5, NF1.

(A) Latency reversal potential upon 48 h treatment of J-Lat A2 cells with increasing concentrations of 5-iodotubercidin (ADK inhibitor), Topiramate (GRIK5 inhibitor) and Trametinib (NF1 inhibitor) was evaluated by flow-cytometry. Treatment with DMSO (black bar) is used as a negative control. Percent of GFP positive cells is indicated by green bars (left y-axes), cell viability is indicated by grey bars (right y-axis).

(B) Latency reversal potential upon 48 h treatment of J-Lat 11.1 cells.

(C) Schematic representation of candidate LRAs treatment in a primary cell model of latent HIV-1 infection. CD4⁺ T-cells are isolated on day 0 and spin infected on day 1. On day 2 virus is removed by media change in presence of saquinavir. Latently infected cells are stimulated with candidate LRAs on day 5 and HIV reactivation is evaluated at day 7.

(D) Latency reversal as measured by Luciferase activity in a primary cell model of HIV latency after treatment with 5-Iodotubercidin, Trametinib and Topiramate in different concentrations. Plots show the fold increase in luciferase activity, measured in relative light units (RLU), after treatment with different concentrations of 5-iodotubercidin (ADK inhibitor), Topiramate (GRIK5 inhibitor) and Trametinib (NF1 inhibitor). Each dot represents a single measurement, black horizontal lines show the average fold increase for each treatment. Averaged data of at least 3 independent experiments performed using each time two different donors (totaling at least 6 different donors). PMA was used as a positive control. Statistical significance was calculated using t test, * – $p < 0,05$; ** – $p < 0,005$; *** – $p < 0,0005$.

(E) Percentage of cells expressing apoptosis marker Annexin V in primary CD4⁺ T cells upon treatment with candidate LRAs for 48 hours. Treatment with a toxic concentration of Gliotoxin (GTX) 200nM was used as a positive control. Experiments were performed in uninfected cells obtained from 6 healthy donors. Data are presented as mean $\pm$ SD from three independent experiments. The (†) symbol indicates low viability.

(F) Percentage of cells expressing marker of cell activation CD69 in primary CD4⁺ T cells from 6 healthy donors, data are presented as mean $\pm$ SD from three independent experiments of 2 different healthy donors upon treatment with candidate LRAs for 48 hours. Treatment with PMA/Ionomycin PMA is used as a positive control.

(G) Percentage of cells expressing marker of cell activation CD25 in primary CD4⁺ T cells from 6 healthy donors, data are presented as Mean $\pm$ SD from three independent experiments of 2 different healthy donors upon treatment with candidate LRAs for 48 hours. Treatment with PMA/Ionomycin is used as a positive control. Statistical significance was calculated using one-way Anova, multiple comparison test. Asterisks indicate the level of significance. (**p < 0.01, ***p < 0.001).

Supplemental Figure 1 Generation of haploid latent cell lines.

(A) Ploidy of Hap-Lat #1 compared to Hap-Lat #1 cells cultured over a prolonged period and Jurkat cells. Haploidy of HapLat #1 cells can be maintained over time by periodic cell sorting of small cells.

(B) Infectivity to HIV-1 of KBM7 and other myeloid lineage cell lines, U937 and MV4, is poor compared to lymphocyte derived cell lines (SupT1 and Jurkat) as shown by percentage GFP positive cells after 4 days of infections with minimal HIV (658).

(C) Nucleoside pre-incubation before and after supplementation during infection improves infectivity of KBM7 cells.

(D) Schematic representing the mechanism of GT insertional mutagenesis. The inactive LTRs of MSCV flank a splice acceptor (SA) site, a mCherry reporter and a poly-adenylation (pA) terminator,

(E) After sorting and expansion, sorted double-positive cells tend to revert to a latent state as indicated by a loss of GFP signal. FACS plots obtained 1day (left panel) and 5 days (right panel) after sorting.

Supplemental Figure 2 Characterization of gene trap integration candidate genes.

(A) Venn diagram depicting all genes with an LOF score >3 in one replicate from the GFP^{Total} (686 hits) and the GFP^{Sub} (382 hits) population. 183 genes are common to both populations.

(B) Expression levels of genes in the GFP^{Total} and the GFP^{Sub} populations. No significant differences in the average expression levels of genes between the two populations is found (Mann-Whitney test).

(C) GO-slim term analysis for biological process of genes in the GFP^{Total} and the GFP^{Sub} populations. No substantial differences in GO-slim term enrichment are found.

(D) GO-slim term analysis for molecular function of genes from both populations.

(E) Percentage of non-coding genes present among genes in the GFP^{Total} and the GFP^{Sub} populations are similar.

(F) Percentage of genes present in the HIV interaction database (<https://www.ncbi.nlm.nih.gov/genome/viruses/retroviruses/hiv-1/interactions/>) among genes in the GFP^{Total} and the GFP^{Sub} populations are similar but higher compared to the total ENSEMBL gene list.

Supplemental Figure 3 Functional validation of positive candidates.

Re-activation of HIV-1 in J-Lat A2 and 11.1 cell lines was assessed by measuring the percentage of cells expressing GFP (green bars) and cell viability (gray bars) using flow-cytometry. Efficacy of knockdown by shRNA was quantitated in J-Lat 11.1 cells by RT-qPCR (left panel) as well as expression of viral genes Tat, Pol and Gag. RT-qPCR data (right panel) are presented as mean $\pm$ SD normalized to the control. Statistical significance was calculated using ratio-paired t-test and multiple comparison t-test on Log2 transformed fold changes (* – $p < 0,05$, ** – $p < 0,01$, *** – $p < 0,001$).

Supplemental Figure 4 Functional validation of false positive candidates.

Efficacy of knockdown by shRNA was quantitated by RT-qPCR. Re-activation of HIV-1 was assessed by measuring by the percentage of cells expressing GFP (green bars) and cell viability (gray bars). RT-qPCR data are presented as mean $\pm$ SD normalized to the control. Statistical significance was calculated using ratio-paired t-test and multiple comparison t-test on Log2 transformed fold changes (* – $p < 0,05$, ** – $p < 0,01$, *** – $p < 0,001$).

Supplemental Figure 5 Additional data on the functional characterization of CHD9.

(A) Efficacy of CHD9 knockdown in J-Lat A2 cells and J-Lat 11.1 cells by shRNA was quantitated by RT qPCR

(B) Biological replicate of a ChIP-qPCR analysis of CHD9 binding to the HIV-1 5'LTR in untreated J-Lat 11.1 cells and PMA stimulated cells (Similar to data from figure 3G-H). Data are represented as percentage of the input and represent the average ($\pm$ SD) of two technical replicates.

Supplemental Figure 6 Characterization of T-cell activation by candidate LRAs.

(A) FACS plot of staining for T-cell activation marker CD69 after treatment of primary CD4+ T cells from 6 healthy donors with candidate LRAs (5-Iodotubercidin, Topiramate and Trametinib). Treatment with PMA/Ionomycin PMA is used as a positive control (Plots correspond to data from figure 4F).

(B) FACS plot of staining for T-cell activation marker CD25 after treatment with candidate LRAs (Plots correspond to data from figure 4G).

Supplemental Figure 7 STRING analysis of protein coding candidate genes.

STRING analysis indicates that many of the 598 protein coding candidate genes with an LOF score of 3 and higher from the GFP^{Total} and the GFP^{Sub} populations functionally interact. Validated candidate genes are indicated in red. Circle size and green shade indicate the number of samples the Gene Trap target gene is found in. Circled genes are part of our candidate genes list. Background color indicates if the Gene Trap target gene is found in the GFP^{Total}, in the GFP^{Sub} or in both populations.

Table 1 Prioritized candidate list for the GFP^{Total} and the GFP^{Sub} populations.

Prioritized candidate genes are defined as having an LOF equal to or greater than 3 in at least 2 independent samples. Overlapping genes are underlined in both lists. Results of shRNA mediated knockdown validation experiments of selected candidate genes in A2 and 11.1 J-Lat cell models are indicated.

Table 2 List of Mission shRNA clones.

Recombinant DNA		
sh ADK	MISSION® shRNA library (Sigma)	TRCN000011535
sh GRIK5	MISSION® shRNA library (Sigma)	TRCN000011537
sh SPN	MISSION® shRNA library (Sigma)	TRCN000011539
sh Control	MISSION® shRNA library (Sigma)	SHC002
sh CMSS1	MISSION® shRNA library (Sigma)	TRCN000011541
sh EVL	MISSION® shRNA library (Sigma)	TRCN000011542
sh COPS5	MISSION® shRNA library (Sigma)	TRCN000011543
sh EXOSC8	MISSION® shRNA library (Sigma)	TRCN000011544
sh SLC7A6	MISSION® shRNA library (Sigma)	TRCN000011545
sh USP15	MISSION® shRNA library (Sigma)	TRCN000011547
sh NF1	MISSION® shRNA library (Sigma)	TRCN000011638
sh IRF2BP2	MISSION® shRNA library (Sigma)	TRCN000011635
sh RHOF	MISSION® shRNA library (Sigma)	TRCN000011642
sh EVI2B	MISSION® shRNA library (Sigma)	TRCN000011644
sh CHD9	MISSION® shRNA library (Sigma)	TRCN0000230236
sh FAM19A	MISSION® shRNA library (Sigma)	TRCN000011641

Table 3 List of qPCR primers.

List of RT-qPCR primers		
Name of gene	Forward Primer	Reverse Primer
ADK	GACACAAGCCCTGCCAAAGAT	TCAGAGACCAGTTGAGACAGAA
CMSS1	GCCAATGATTTGACTCACAGTCT	CTGAATGCTGTCATCGACCTAAT
IRF2BP2	CCCATGACTCCTACATCCTCTT	GAGGGCGGACTGTTGCTATTC
NF1	AGATGAAACGATGCTGGTCAAA	CCTGTAACCTGGTAGAAATGCCA
CHD9	GAAATGCTAGAAAGGTTGGAGG	CGGGACCAGTGAGAATACGTT
FAM19A	TGTTAAAACGGGAACCTTGTGAGG	AAGCATCCACACATGATGGAG
SCN9A	ATTCGTGGCTCCTTGTCTTCTG	CTACTGGCTTGGCTGATGTTAC
RHOF	CCCCATCGGTGTTTCGAGAAG	GGCCGTGTCGTAGAGGTTT
GRIK5	CCACCGTGAGCCATATCTGTG	CGCGAAGCGAAGGTACTGAA
USP15	CGACGCTGCTCAAAACCTC	TCCCATCTGGTATTTGTCCCAA
EVL	CTTCCGTGATGGTCTACGATG	TGCAACTTGACTCCAACGACT
EXOSC8	CGGTTCAATTAGTACCGCAGAT	ACGTATCCTTTATCAGGGGCGAT
COPS5	TGGGTCTGATGCTAGGAAAGG	CTATGATACCACCCGATTGCATT
SLC7A6	CGGGCTTTCAGTTGTTAGACC	ACACAGTCAATCGCTCACGG
EVI2B	ACCAACACAATTCAGCGACAC	GTTGTAGGCAAGTGGTTGTCC
SPN	GCTGGTGGTAAGCCCAGAC	GGCTCGCTAGTAGAGACCAAA
POL	GGTTTATTACAGGGACAGCAGAGA	ACCTGCCATCTGTTTTCCATA
GAG	AGTAGTGTGTGCCCGTCTGT	TCGCTTTCAGGTCCCTGTTTCG
TAT	CAAAAGCCTTAGGCATCTCCT	CCACCTTCTTCTTCGATTCTT
β2Microglobulin	AGCGTACTCCAAAGATTCAGGTT	ATGATGCTGCTTACATGTCTCGAT
GAPDH	CAAGAAGGTGGTGAAGCAG	GCCAAATTCGTTGTCATACC
List of ChIP-qPCR primers		
amplicon	Forward Primer	Reverse Primer
HIV-1 Nuc 0	CCACACACAAGGCTACTTCC	AACTGGTACTAGCTTGTAGCAC
HIV-1 HSS1	TGTGAGCCTGCATGGGATGG	GAAAGTCCCCAGCGGAAAGT
HIV-1 Nuc1/HSS2	CGTCTGTTGTGTGACTCTGGT	TCGAGAGAGCTCCTCTGGTT
HIV-1 HSS2/Nuc2	GCCCCGAACAGGGACTTGAAA	TTGGCGTACTCACCAGTCG
HIV-1 Nuc2/Gag	GGTGCGAGAGCGTCAGTAT	AGCTCCCTGCTTGCCCATA
Hk2	GCCGACTCTTGTATTGCCTG	TATTGTAGCACGGCCGGAAA
GFP (HIV-1)	GGAGTGGTCCCAGTTCTTGTTG	ACAGGTAGCTTCCCAGTAGTGC

References:

1. Siliciano JM, Siliciano RF. 2015. The Remarkable Stability of the Latent Reservoir for HIV-1 in Resting Memory CD4+ T Cells. *The Journal of Infectious Diseases* 212:1345–1347.
2. Sengupta S, Siliciano RF. 2018. Targeting the latent reservoir for HIV-1. *Immunity* 48:872–895.
3. Deeks SG. 2012. HIV: Shock and kill. *Nature* 487:439–440.
4. Ne E, Palstra R-J, Mahmoudi T. 2018. Transcription: Insights From the HIV-1 Promoter. *Int Rev Cell Mol Biol* 335:191–243.
5. Van Lint C, Bouchat S, Marcello A. 2013. HIV-1 transcription and latency: an update. *Retrovirology* 10:67.
6. Pereira LA, Bentley K, Peeters A, Churchill MJ, Deacon NJ. 2000. A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. *Nucleic Acids Res* 28:663–668.
7. Archin NM, Espeseth A, Parker D, Cheema M, Hazuda D, Margolis DM. 2009. Expression of latent HIV induced by the potent HDAC inhibitor suberoylanilide hydroxamic acid. *AIDS Research and Human Retroviruses* 25:207–212.
8. Ylisastigui L, Archin NM, Lehrman G, Bosch RJ, Margolis DM. 2004. Coaxing HIV-1 from resting CD4 T cells: histone deacetylase inhibition allows latent viral expression. *AIDS* 18:1101–1108.
9. Wei DG, Chiang V, Fyne E, Balakrishnan M, Barnes T, Graupe M, Hesselgesser J, Irrinki A, Murry JP, Stepan G, Stray KM, Tsai A, Yu H, Spindler J, Kearney M, Spina CA, McMahon D, Lalezari J, Sloan D, Mellors J, Geleziunas R, Cihlar T, DeJesus E, Rockstroh JK, Henry K, Molina JM, Gathe J, Raffi F, Rachlis A, Stellbrink HJ, Hardy WD, Torti C, Palmer S, Maldarelli F, Wiegand A, Bernstein B, Hanna GJ, Chun TW, Jr RTD, Ostrowski M, Justement JS, Engel D, Papasavvas E, Kostman JR,

Mounzer K, Grant RM, Gross R, Siliciano JD, Siliciano RF, Siliciano JD, Kajdas J, Finzi D, Quinn TC, Chadwick K, Chomont N, El-Far M, Ancuta P, Trautmann L, Procopio FA, Eisele E, Siliciano RF, Eriksson S, Graf EH, Dahl V, Strain MC, Yukl SA, Siliciano JD, Siliciano RF, Richman DD, Margolis DM, Delaney M, Greene WC, Hazuda DJ, Cohort CAT, Smit C, Geskus R, Walker S, Sabin CA, Coutinho R, Friis-Moller N, Sabin CA, Weber R, d'Arminio Monforte A, El-Sadr WM, Coiras M, Lopez-Huertas MR, Perez-Olmeda M, Alcamí J, Colin L, Lint C Van, Williams SA, Chen LF, Kwon H, Ruiz-Jarabo CM, Verdin E, Jiang G, Espeseth A, Hazuda DJ, Margolis DM, Wightman F, Ellenberg P, Churchill M, Lewin SR, Archin NM, Espeseth A, Parker D, Cheema M, Hazuda DJ, Edelstein LC, Micheva-Viteva S, Phelan BD, Dougherty JP, Matalon S, Rasmussen TA, Dinarello CA, Beliakova-Bethell N, Zhang JX, Singhania A, Lee V, Terry VH, Archin NM, Liberty AL, Kashuba AD, Choudhary SK, Kuruc JD, Rasmussen TA, Sogaard OS, Brinkmann C, Wightman F, Lewin SR, Ueda H, Nakajima H, Hori Y, Fujita T, Nishimura M, Bertino EM, Otterson GA, Bosque A, Planelles V, Bosque A, Planelles V, Piekarczyk RL, Frye R, Prince HM, Kirschbaum MH, Zain J, Sandor V, Bakke S, Robey RW, Kang MH, Blagosklonny M V, Palmer S, Kearney M, Maldarelli F, Halvas EK, Bixby CJ, Kearney M, Palmer S, Maldarelli F, Shao W, Polis MA, Kearney M, Spindler J, Shao W, Maldarelli F, Palmer S, Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Furumai R, Matsuyama A, Kobashi N, Lee KH, Nishiyama M, Burnett JC, Lim KI, Calafi A, Rossi JJ, Schaffer D V, Friedman J, Cho WK, Chu CK, Keedy KS, Archin NM, Bouchat S, Gatot JS, Kabeya K, Cardona C, Colin L, Xing S, Bullen CK, Shroff NS, Shan L, Yang HC, Xing S, Bhat S, Shroff NS, Zhang H, Lopez JA, Shan L, Deng K, Shroff NS, Durand CM, Rabi SA, Badley AD, Sainski A, Wightman F, Lewin SR, Smalls-Mantey A, Doria-Rose N, Klein R, Patamawenu A, Migueles SA, Bhutani D, Vaishampayan UN, Iglesias-Ussel M, Vandergeeten C, Marchionni L, Chomont N, Romerio F, Sweet MJ, Shakespear MR, Kamal NA, Fairlie DP, Licciardi P V, Ververis K, Tang ML, El-Osta A, Karagiannis TC, Roger T, Lugin J, Roy D Le, Goy G, Mombelli M, Doherty R, O'Farrelly C, Meade KG, Hara N, Alkanani AK, Dinarello CA, Zipris

- D, Alvarez-Breckenridge CA, Yu J, Price R, Wei M, Wang Y, Mombelli M, Lugrin J, Rubino I, Chanson AL, Giddey M, Kelly-Sell MJ, Kim YH, Straus S, Benoit B, Harrison C, Messi M, Giacchetto I, Nagata K, Lanzavecchia A, Natoli G. 2014. Histone Deacetylase Inhibitor Romidepsin Induces HIV Expression in CD4 T Cells from Patients on Suppressive Antiretroviral Therapy at Concentrations Achieved by Clinical Dosing. *PLoS Pathogens* 10:e1004071.
10. Ying H, Zhang Y, Zhou X, Qu X, Wang P, Liu S, Lu D, Zhu H. 2012. Selective Histone deacetylase Inhibitor M344 Intervenes in HIV-1 Latency through Increasing Histone Acetylation and Activation of NF-kappaB. *PLoS One* 7.
 11. Williams SA, Chen L-F, Kwon H, Ruiz-Jarabo CM, Verdin E, Greene WC. 2006. NF- κ B p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. *EMBO J* 25:139–149.
 12. Coull JJ, Romerio F, Sun JM, Volker JL, Galvin KM, Davie JR, Shi Y, Hansen U, Margolis DM. 2000. The human factors YY1 and LSF repress the human immunodeficiency virus type 1 long terminal repeat via recruitment of histone deacetylase 1. *J Virol* 74:6790–6799.
 13. Jiang G, Espeseth A, Hazuda DJ, Margolis DM. 2007. c-Myc and Sp1 Contribute to Proviral Latency by Recruiting Histone Deacetylase 1 to the Human Immunodeficiency Virus Type 1 Promoter. *Journal of Virology* 81:10914–10923.
 14. Williams SA, Chen L-F, Kwon H, Fenard D, Bisgrove D, Verdin E, Greene WC. 2004. Prostratin antagonizes HIV latency by activating NF-kappaB. *J Biol Chem* 279:42008–42017.

15. Pérez M, de Vinuesa AG, Sanchez-Duffhues G, Marquez N, Bellido ML, Muñoz-Fernandez MA, Moreno S, Castor TP, Calzado MA, Muñoz E. 2010. Bryostatins synergize with histone deacetylase inhibitors to reactivate HIV-1 from latency. *Curr HIV Res* 8:418–429.
16. Archin NM, Liberty AL, Kashuba AD, Choudhary SK, Kuruc JD, Crooks AM, Parker DC, Anderson EM, Kearney MF, Strain MC, Richman DD, Hudgens MG, Bosch RJ, Coffin JM, Eron JJ, Hazuda DJ, Margolis DM. 2012. Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. *Nature* 487:482–485.
17. Elliott JH, Wightman F, Solomon A, Ghneim K, Ahlers J, Cameron MJ, Smith MZ, Spelman T, McMahon J, Velayudham P, Brown G, Roney J, Watson J, Prince MH, Hoy JF, Chomont N, Fromentin R, Procopio FA, Zeidan J, Palmer S, Odeval L, Johnstone RW, Martin BP, Sinclair E, Deeks SG, Hazuda DJ, Cameron PU, Sékaly R-P, Lewin SR. 2014. Activation of HIV Transcription with Short-Course Vorinostat in HIV-Infected Patients on Suppressive Antiretroviral Therapy. *PLOS Pathogens* 10:e1004473.
18. Sadowski I, Hashemi FB. 2019. Strategies to eradicate HIV from infected patients: elimination of latent provirus reservoirs. *Cell Mol Life Sci* 76:3583–3600.
19. Das B, Dobrowolski C, Luttge B, Valadkhan S, Chomont N, Johnston R, Bacchetti P, Hoh R, Gandhi M, Deeks SG, Scully E, Karn J. 2018. 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* 115:E7795–E7804.
20. Jiang W-M, Zhang X-Y, Zhang Y-Z, Liu L, Lu H-Z. 2014. A high throughput RNAi screen reveals determinants of HIV-1 activity in host kinases. *Int J Clin Exp Pathol* 7:2229–2237.

21. Besnard E, Hakre S, Kampmann M, Lim HW, Hosmane NN, Martin A, Bassik MC, Verschueren E, Battivelli E, Chan J, Svensson JP, Gramatica A, Conrad RJ, Ott M, Greene WC, Krogan NJ, Siliciano RF, Weissman JS, Verdin E. 2016. The mTOR Complex Controls HIV Latency. *Cell Host Microbe* 20:785–797.
22. Zhu J, Davoli T, Perriera JM, Chin CR, Gaiha GD, John SP, Sigiollot FD, Gao G, Xu Q, Qu H, Pertel T, Sims JS, Smith JA, Baker RE, Maranda L, Ng A, Elledge SJ, Brass AL. 2014. Comprehensive identification of host modulators of HIV-1 replication using multiple orthologous RNAi reagents. *Cell Rep* 9:752–766.
23. Li Z, Hajian C, Greene WC. 2020. Identification of unrecognized host factors promoting HIV-1 latency. *PLOS Pathogens* 16:e1009055.
24. Yang X, Wang Y, Lu P, Shen Y, Zhao X, Zhu Y, Jiang Z, Yang H, Pan H, Zhao L, Zhong Y, Wang J, Liang Z, Shen X, Lu D, Jiang S, Xu J, Wu H, Lu H, Jiang G, Zhu H. 2020. PEBP1 suppresses HIV transcription and induces latency by inactivating MAPK/NF- κ B signaling. *EMBO reports* 21:e49305.
25. Goff SP. 2008. Knockdown Screens to Knockout HIV-1. *Cell* 135:417–420.
26. Rathore A, Iketani S, Wang P, Jia M, Sahi V, Ho DD. 2020. CRISPR-based gene knockout screens reveal deubiquitinases involved in HIV-1 latency in two Jurkat cell models. 1. *Scientific Reports* 10:5350.
27. Krasnopolsky S, Kuzmina A, Taube R. 2020. Genome-wide CRISPR knockout screen identifies ZNF304 as a silencer of HIV transcription that promotes viral latency. *PLoS Pathog* 16:e1008834.

28. Huang H, Kong W, Jean M, Fiches G, Zhou D, Hayashi T, Que J, Santoso N, Zhu J. 2019. A CRISPR/Cas9 screen identifies the histone demethylase MINA53 as a novel HIV-1 latency-promoting gene (LPG). *Nucleic Acids Res* 47:7333–7347.
29. Martin SE, Caplen NJ. 2007. Applications of RNA Interference in Mammalian Systems. *Annual Review of Genomics and Human Genetics* 8:81–108.
30. Yamamoto-Hino M, Goto S. 2013. In Vivo RNAi-Based Screens: Studies in Model Organisms. 4. *Genes* 4:646–665.
31. Hart T, Brown KR, Sircoulomb F, Rottapel R, Moffat J. 2014. Measuring error rates in genomic perturbation screens: gold standards for human functional genomics. *Molecular Systems Biology* 10.
32. Das B, Dobrowolski C, Luttge B, Valadkhan S, Chomont N, Johnston R, Bacchetti P, Hoh R, Gandhi M, Deeks SG, Scully E, Karn J. 2018. 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* 115:E7795–E7804.
33. Zhu J, Davoli T, Perriera JM, Chin CR, Gaiha GD, John SP, Sigiollot FD, Gao G, Xu Q, Qu H, Pertel T, Sims JS, Smith JA, Baker RE, Maranda L, Ng A, Elledge SJ, Brass AL. 2014. Comprehensive identification of host modulators of HIV-1 replication using multiple orthologous RNAi reagents. *Cell Reports* 9:752–766.
34. Boehm D, Jeng M, Camus G, Gramatica A, Schwarzer R, Johnson JR, Hull PA, Montano M, Sakane N, Pagans S, Godin R, Deeks SG, Krogan NJ, Greene WC, Ott M. 2017. SMYD2-Mediated Histone Methylation Contributes to HIV-1 Latency. *Cell Host and Microbe* 21:569-579.e6.

35. Yu JSL, Yusa K. 2019. Genome-wide CRISPR-Cas9 screening in mammalian cells. *Methods* 164–165:29–35.
36. Kotecki M, Reddy PS, Cochran BH. 1999. Isolation and characterization of a near-haploid human cell line. *Exp Cell Res* 252:273–280.
37. Carette JE, Guimaraes CP, Varadarajan M, Park AS, Wuethrich I, Godarova A, Kotecki M, Cochran BH, Spooner E, Ploegh HL, Brummelkamp TR, Muller HJ, Brass AL, Philips JA, Rubin EJ, Perrimon N, Hao L, Salomon R, Webster RG, Moscona A, Kotecki M, Reddy PS, Cochran BH, Nagata S, Luo B, Cong F, Lara-Tejero M, Galán JE, Nesic D, Hsu Y, Stebbins CE, Wistrand M, Kall L, Sonnhhammer ELL, Miyaji M, Guerra L, Wollscheid B, Sprong H, Collier RJ, Liu S., Milne GT, Kuremsky JG, Fink GR, Leppla SH, Milne JC, Blanke SR, Hanna PC, Collier RJ, Scobie HM, Rainey GJA, Bradley KA, Young JAT, Naglich JG, Metherall JE, Russell DW, Eidels L, Mattheakis LC, Shen WH, Collier RJ, Liu S., Leppla SH, Chen JY, Bodley JW, Hillenmeyer ME, Forsburg SL, Ma Y, Creanga A, Lum L, Beachy PA, Echeverri CJ, Goff SP, Bushman FD. 2009. Haploid genetic screens in human cells identify host factors used by pathogens. *Science (New York, NY)* 326:1231–1235.
38. Reiling JH, Clish CB, Carette JE, Varadarajan M, Brummelkamp TR, Sabatini DM. 2011. A haploid genetic screen identifies the major facilitator domain containing 2A (MFSD2A) transporter as a key mediator in the response to tunicamycin. *PNAS* 108:11756–11765.
39. Birsoy K, Wang T, Possemato R, Yilmaz OH, Koch CE, Chen WW, Hutchins AW, Gultekin Y, Peterson TR, Carette JE, Brummelkamp TR, Clish CB, Sabatini DM. 2013. MCT1-mediated transport of a toxic molecule is an effective strategy for targeting glycolytic tumors. *Nat Genet* 45:104–108.

40. Logtenberg MEW, Jansen JHM, Raaben M, Toebe M, Franke K, Brandsma AM, Matlung HL, Fauster A, Gomez-Eerland R, Bakker NAM, van der Schot S, Marijt KA, Verdoes M, Haanen JBAG, van den Berg JH, Neefjes J, van den Berg TK, Brummelkamp TR, Leusen JHW, Scheeren FA, Schumacher TN. 2019. Glutaminyl cyclase is an enzymatic modifier of the CD47- SIRP α axis and a target for cancer immunotherapy. *Nature Medicine*. Nature Publishing Group.
41. Mezzadra R, De Bruijn M, Jae LT, Gomez-Eerland R, Duursma A, Scheeren FA, Brummelkamp TR, Schumacher TN. 2019. SLFN11 can sensitize tumor cells towards IFN- γ -mediated T cell killing. *PLoS ONE* 14.
42. Heijink AM, Talens F, Jae LT, van Gijn SE, Fehrmann RSN, Brummelkamp TR, van Vugt MATM. 2019. BRCA2 deficiency instigates cGAS-mediated inflammatory signaling and confers sensitivity to tumor necrosis factor- α -mediated cytotoxicity. *Nature Communications* 10.
43. Mezzadra R, Sun C, Jae LT, Gomez-Eerland R, De Vries E, Wu W, Logtenberg MEW, Slagter M, Rozeman EA, Hofland I, Broeks A, Horlings HM, Wessels LFA, Blank CU, Xiao Y, Heck AJR, Borst J, Brummelkamp TR, Schumacher TNM. 2017. Identification of CMTM6 and CMTM4 as PD-L1 protein regulators. *Nature* 549:106–110.
44. Janssen LME, Averink T V., Blomen VA, Brummelkamp TR, Medema RH, Raaijmakers JA. 2018. Loss of Kif18A Results in Spindle Assembly Checkpoint Activation at Microtubule-Attached Kinetochores. *Current Biology* 28:2685-2696.e4.
45. Raaijmakers JA, van Heesbeen RGHP, Blomen VA, Janssen LME, van Diemen F, Brummelkamp TR, Medema RH. 2018. BUB1 Is Essential for the Viability of Human Cells in which the Spindle Assembly Checkpoint Is Compromised. *Cell Reports* 22:1424–1438.

46. Blomen VA, Májek P, Jae LT, Bigenzahn JW, Nieuwenhuis J, Staring J, Sacco R, Van Diemen FR, Olk N, Stukalov A, Marceau C, Janssen H, Carette JE, Bennett KL, Colinge J, Superti-Furga G, Brummelkamp TR. 2015. Gene essentiality and synthetic lethality in haploid human cells. *Science* 350:1092–1096.
47. Lee CC, Carette JE, Brummelkamp TR, Ploegh HL. 2013. A reporter screen in a human haploid cell line identifies CYLD as a constitutive inhibitor of NF-κB. *PLoS one* 8:e70339.
48. Carette JE, Raaben M, Wong AC, Herbert AS, Obernosterer G, Mulherkar N, Kuehne AI, Kranzusch PJ, Griffin AM, Ruthel G, Cin PD, Dye JM, Whelan SP, Chandran K, Brummelkamp TR. 2011. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature* 477:340–343.
49. Jae LT, Raaben M, Riemersma M, Van Beusekom E, Blomen VA, Velds A, Kerkhoven RM, Carette JE, Topaloglu H, Meinecke P, Wessels MW, Lefeber DJ, Whelan SP, Van Bokhoven H, Brummelkamp TR. 2013. Deciphering the glycosylome of dystroglycanopathies using haploid screens for Lassa virus entry. *Science* 340:479–483.
50. Staring J, Von Castelmuur E, Blomen VA, Van Den Hengel LG, Brockmann M, Baggen J, Thibaut HJ, Nieuwenhuis J, Janssen H, Van Kuppeveld FJM, Perrakis A, Carette JE, Brummelkamp TR. 2017. PLA2G16 represents a switch between entry and clearance of Picornaviridae. *Nature* 541:412–416.
51. Hoffmann HH, Schneider WM, Blomen VA, Scull MA, Hovnanian A, Brummelkamp TR, Rice CM. 2017. Diverse Viruses Require the Calcium Transporter SPCA1 for Maturation and Spread. *Cell Host and Microbe* 22:460-470.e5.

52. Raaben M, Jae LT, Herbert AS, Kuehne AI, Stubbs SH, Chou Y ying, Blomen VA, Kirchhausen T, Dye JM, Brummelkamp TR, Whelan SP. 2017. NRP2 and CD63 Are Host Factors for Lujo Virus Cell Entry. *Cell Host and Microbe* 22:688-696.e5.
53. Staring J, van den Hengel LG, Raaben M, Blomen VA, Carette JE, Brummelkamp TR. 2018. KREMEN1 Is a Host Entry Receptor for a Major Group of Enteroviruses. *Cell Host and Microbe* 23:636-643.e5.
54. Jangra RK, Herbert AS, Li R, Jae LT, Kleinfelter LM, Slough MM, Barker SL, Guardado-Calvo P, Román-Sosa G, Dieterle ME, Kuehne AI, Muenz NA, Wirchnianski AS, Nyakatura EK, Fels JM, Ng M, Mittler E, Pan J, Bharrhan S, Wec AZ, Lai JR, Sidhu SS, Tischler ND, Rey FA, Moffat J, Brummelkamp TR, Wang Z, Dye JM, Chandran K. 2018. Protocadherin-1 is essential for cell entry by New World hantaviruses. *Nature*. Nature Publishing Group.
55. Qi W, Wang R, Chen H, Wang X, Xiao T, Boldogh I, Ba X, Han L, Zeng X, Agger K, Santoni-Rugiu E, Holmberg C, Karlström O, Helin K, Bennardo N, Cheng A, Huang N, Stark JM, Bultman S, Gebuhr T, Yee D, Mantia C La, Nicholson J, Gilliam A, Randazzo F, Metzger D, Chambon P, Crabtree G, Chai B, Huang J, Cairns BR, Laurent BC, Chen X, Cui D, Papusha A, Zhang X, Chu CD, Tang J, Chen K, Pan X, Ira G, Costelloe T, Louge R, Tomimatsu N, Mukherjee B, Martini E, Khadaroo B, Dubois K, Wiegant WW, Thierry A, Burma S, Cruz LA, Guecheva TN, Bonato D, Henriques JA, Eapen V V, Sugawara N, Tsabar M, Wu WH, Haber JE, Falck J, Coates J, Jackson SP, Feng Z, Scott SP, Bussen W, Sharma GG, Guo G, Pandita TK, Powell SN, Franken NA, Rodermond HM, Stap J, Haveman J, van Bree C, Helfricht A, Wiegant WW, Thijssen PE, Vertegaal AC, Luijsterburg MS, van Attikum H, Iacovoni JS, Caron P, Lassadi I, Nicolas E, Massip L, Trouche D, Legube G, Johnson RD, Jasin M, Khanna KK, Jackson SP, Kruhlak MJ, Celeste A, Dellaire G, Fernandez-Capetillo O, Müller WG, McNally JG, Bazett-Jones DP, Nussenzweig A, Kwon SJ, Park JH, Park EJ, Lee SA, Lee HS, Kang SW,

- Kwon J, Lee HS, Park JH, Kim SJ, Kwon SJ, Kwon J, Lim DS, Vogel H, Willerford DM, Sands AT, Platt KA, Hasty P, Lisby M, Barlow JH, Burgess RC, Rothstein R, Littlewood TD, Hancock DC, Danielian PS, Parker MG, Evan GI, Liu J, Heyer WD, Liu J, Doty T, Gibson B, Heyer WD, Lok BH, Carley AC, Tchang B, Powell SN, Martens JA, Winston F, Morshead KB, Ciccone DN, Taverna SD, Allis CD, Oettinger MA, New JH, Sugiyama T, Zaitseva E, Kowalczykowski SC, Nimonkar A V, Genschel J, Kinoshita E, Polaczek P, Campbell JL, Wyman C, Modrich P, Kowalczykowski SC, O'Driscoll M, Osley MA, Tsukuda T, Nickoloff JA, Park JH, Park EJ, Lee HS, Kim SJ, Hur SK, Imbalzano AN, Kwon J, Patenge N, Elkin SK, Oettinger MA, Peng G, Yim EK, Dai H, Jackson AP, Burgt I, Pan MR, Hu R, Li K, Lin SY, Peterson CL, Almouzni G, Rai R, Dai H, Multani AS, Li K, Chin K, Gray J, Lahad JP, Liang J, Mills GB, Meric-Bernstam F, Ray A, Mir SN, Wani G, Zhao Q, Battu A, Zhu Q, Wang QE, Wani AA, Rich T, Allen RL, Wyllie AH, Roberts CW, Orkin SH, Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM, Filippo JS, Sung PA, Klein H, Shinohara A, Ogawa T, Smith CL, Peterson CL, Sugiyama T, Kowalczykowski SC, Sung PA, Libura J, Richardson C, Symington LS, Tong H, Zhao B, Shi H, Ba X, Wang X, Jiang Y, Zeng X, Tyteca S, Vandromme M, Legube G, Chevillard-Briet M, Trouche D, van Attikum H, Gasser SM, van Gent DC, Hoeijmakers JH, Kanaar R, Wang X, Haber JE, Weinstock DM, Nakanishi K, Helgadottir HR, Jasin M, Wolner B, van Komen S, Sung PA, Peterson CL, Wong AK, Shanahan F, Chen Y, Lian L, Ha P, Hendricks K, Ghaffari S, Iliev D, Penn B, Woodland AM, Wray J, Liu J, Nickoloff JA, Shen Z, Xu Y, Sun Y, Jiang X, Ayrappetov MK, Moskwa P, Yang S, Weinstock DM, Price BD, Yun MH, Hiom K, Zhang J, Ma Z, Treszezamsky A, Powell SN. 2015. BRG1 promotes the repair of DNA double-strand breaks by facilitating the replacement of RPA with RAD51. *Journal of cell science* 128:317–330.
56. Guimaraes CP, Carette JE, Varadarajan M, Antos J, Popp MW, Spooner E, Brummelkamp TR, Ploegh HL. 2011. Identification of host cell factors required for intoxication through use of modified cholera toxin. *Journal of Cell Biology* 195:751–764.

57. Tafesse FG, Guimaraes CP, Maruyama T, Carette JE, Lory S, Brummelkamp TR, Ploegh HL. 2014. GPR107, a G-protein-coupled receptor essential for intoxication by *Pseudomonas aeruginosa* exotoxin a, localizes to the Golgi and is cleaved by furin. *Journal of Biological Chemistry* 289:24005–24018.
58. Jordan A, Bisgrove D, Verdin E. 2003. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *The EMBO journal* 22:1868–1877.
59. Rafati H, Parra M, Hakre S, Moshkin Y, Verdin E, Mahmoudi T. 2011. Repressive LTR nucleosome positioning by the BAF complex is required for HIV latency. *PLoS Biology* 9:e1001206.
60. Goujon C, Arfi V, Pertel T, Luban J, Lienard J, Rigal D, Darlix J-L, Cimorelli A. 2008. Characterization of Simian Immunodeficiency Virus SIVSM/Human Immunodeficiency Virus Type 2 Vpx Function in Human Myeloid Cells. *Journal of Virology* 82:12335–12345.
61. Kaushik R, Zhu X, Stranska R, Wu Y, Stevenson M. 2009. A Cellular Restriction Dictates the Permissivity of Nondividing Monocytes/Macrophages to Lentivirus and Gammaretrovirus Infection. *Cell Host and Microbe* 6:68–80.
62. Laguette N, Sobhian B, Casartelli N, Ringard M, Chable-Bessia C, Ségéral E, Yatim A, Emiliani S, Schwartz O, Benkirane M. 2011. SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* 474:654–657.
63. Hansen J, Floss T, Van Sloun P, Füchtbauer EM, Vauti F, Arnold HH, Schnütgen F, Wurst W, Von Melchert H, Ruiz P. 2003. A large-scale, gene-driven mutagenesis approach for the functional analysis of the mouse genome. *Proceedings of the National Academy of Sciences of the United States of America* 100:9918–9922.

64. Minin GD, Postlmayr A, Wutz A. 2018. HaSAPPy: A tool for candidate identification in pooled forward genetic screens of haploid mammalian cells. *PLOS Computational Biology* 14:e1005950.
65. Mayor-Ruiz C, Bauer S, Brand M, Kozicka Z, Siklos M, Imrichova H, Kaltheuner IH, Hahn E, Seiler K, Koren A, Petzold G, Fellner M, Bock C, Müller AC, Zuber J, Geyer M, Thomä NH, Kubicek S, Winter GE. 2020. Rational discovery of molecular glue degraders via scalable chemical profiling. 11. *Nature Chemical Biology* 16:1199–1207.
66. Palstra R-J, Crignis E de, Röling MD, Staveren T van, Kan TW, Ijcken W van, Mueller YM, Katsikis PD, Mahmoudi T. 2018. Allele-specific long-distance regulation dictates IL-32 isoform switching and mediates susceptibility to HIV-1. *Science Advances* 4:e1701729.
67. Jordan A, Defechereux P, Verdin E. 2001. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *EMBO J* 20:1726–1738.
68. Forlani G, Turrini F, Ghezzi S, Tedeschi A, Poli G, Accolla RS, Tosi G. 2016. The MHC-II transactivator CIITA inhibits Tat function and HIV-1 replication in human myeloid cells. *J Transl Med* 14:94.
69. Accolla RS, Mazza S, De Lerma Barbaro A, De Maria A, Tosi G. 2002. The HLA class II transcriptional activator blocks the function of HIV-1 Tat and inhibits viral replication. *Eur J Immunol* 32:2783–2791.
70. Mills AA. 2017. The Chromodomain Helicase DNA-Binding Chromatin Remodelers: Family Traits that Protect from and Promote Cancer. *Cold Spring Harb Perspect Med* 7.

71. Park J, Gupta RS. 2013. Adenosine Metabolism, Adenosine Kinase, and Evolution, p. 23–54. *In* Masino, S, Boison, D (eds.), Adenosine: A Key Link between Metabolism and Brain Activity. Springer, New York, NY.
72. Boison D. 2013. Adenosine Kinase: Exploitation for Therapeutic Gain. *Pharmacol Rev* 65:906–943.
73. Zhang X, Jia D, Liu H, Zhu N, Zhang W, Feng J, Yin J, Hao B, Cui D, Deng Y, Xie D, He L, Li B. 2013. Identification of 5-Iodotubercidin as a genotoxic drug with anti-cancer potential. *PLoS One* 8:e62527.
74. Rojas A, Dingledine R. 2013. Ionotropic Glutamate Receptors: Regulation by G-Protein-Coupled Receptors. *Mol Pharmacol* 83:746–752.
75. Shank RP, Gardocki JF, Streeter AJ, Maryanoff BE. 2000. An overview of the preclinical aspects of topiramate: pharmacology, pharmacokinetics, and mechanism of action. *Epilepsia* 41:3–9.
76. Walker JA, Upadhyaya M. 2018. Emerging therapeutic targets for neurofibromatosis type 1. *Expert Opin Ther Targets* 22:419–437.
77. Harrell Stewart DR, Clark GJ. 2020. Pumping the brakes on RAS - negative regulators and death effectors of RAS. *J Cell Sci* 133.
78. Wright CJM, McCormack PL. 2013. Trametinib: first global approval. *Drugs* 73:1245–1254.
79. Lassen KG, Hebbeler AM, Bhattacharyya D, Lobritz MA, Greene WC. 2012. A flexible model of HIV-1 latency permitting evaluation of many primary CD4 T-cell reservoirs. *PLoS One* 7:e30176.
80. Stoszko M, Al-Hatmi AMS, Skriba A, Roling M, Ne E, Crespo R, Mueller YM, Najafzadeh MJ, Kang J, Ptackova R, LeMasters E, Biswas P, Bertoldi A, Kan TW, Crignis E de, Sulc M, Lebbink JHG, Rokx C,

Verbon A, Ijcken W van, Katsikis PD, Palstra R-J, Havlicek V, Hoog S de, Mahmoudi T. 2020.

Gliotoxin, identified from a screen of fungal metabolites, disrupts 7SK snRNP, releases P-TEFb, and reverses HIV-1 latency. *Science Advances* 6:eaba6617.

81. Kumar A, Abbas W, Herbein G. 2014. HIV-1 Latency in Monocytes/Macrophages. *Viruses* 6:1837–1860.
82. Wong ME, Jaworowski A, Hearps AC. 2019. The HIV Reservoir in Monocytes and Macrophages. *Front Immunol* 10.
83. Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T, Rosen N, Kohn A, Twik M, Safran M, Lancet D, Cohen D. 2017. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)* 2017.
84. Lai B, Tang Q, Jin W, Hu G, Wangsa D, Cui K, Stanton BZ, Ren G, Ding Y, Zhao M, Liu S, Song J, Ried T, Zhao K. 2018. Trac-looping measures genome structure and chromatin accessibility. *Nat Methods* 15:741–747.
85. Carneiro FRG, Ramalho-Oliveira R, Mognol GP, Viola JPB. 2011. Interferon Regulatory Factor 2 Binding Protein 2 Is a New NFAT1 Partner and Represses Its Transcriptional Activity ▽ . *Mol Cell Biol* 31:2889–2901.
86. Razooky BS, Pai A, Aull K, Rouzine IM, Weinberger LS. 2015. A hardwired HIV latency program. *Cell* 160:990–1001.
87. Hansen MMK, Martin B, Weinberger LS. 2019. HIV Latency: Stochastic across Multiple Scales. *Cell Host & Microbe* 26:703–705.

88. Tantale K, Garcia-Oliver E, L'Hostis A, Yang Y, Robert M-C, Gostan T, Basu M, Kozulic-Pirher A, Andrau J-C, Muller F, Basyuk E, Radulescu O, Bertrand E. 2020. Stochastic pausing at latent HIV-1 promoters generates transcriptional bursting. *bioRxiv* 2020.08.25.265413.
89. Piette J, Legrand-Poels S. 1994. HIV-1 reactivation after an oxidative stress mediated by different reactive oxygen species. *Chemico-Biological Interactions* 91:79–89.
90. Vallejo-Gracia A, Chen IP, Perrone R, Besnard E, Boehm D, Battivelli E, Tezil T, Krey K, Raymond KA, Hull PA, Walter M, Habrylo I, Cruz A, Deeks S, Pillai S, Verdin E, Ott M. 2020. FOXO1 promotes HIV latency by suppressing ER stress in T cells. *Nat Microbiol* 5:1144–1157.
91. Timmons A, Fray E, Kumar M, Wu F, Dai W, Bullen CK, Kim P, Hetzel C, Yang C, Beg S, Lai J, Pomerantz JL, Yuki SA, Siliciano JD, Siliciano RF. 2020. HSF1 inhibition attenuates HIV-1 latency reversal mediated by several candidate LRAs In Vitro and Ex Vivo. *PNAS* 117:15763–15771.
92. Kriegel H-P, Kröger P, Schubert E, Zimek A. 2009. LoOP: local outlier probabilities, p. 1649–1652. *In Proceedings of the 18th ACM conference on Information and knowledge management*. Association for Computing Machinery, New York, NY, USA.
93. Klase Z, Yedavalli VSRK, Houzet L, Perkins M, Maldarelli F, Brenchley J, Strebel K, Liu P, Jeang K-T. 2014. Activation of HIV-1 from latent infection via synergy of RUNX1 inhibitor Ro5-3335 and SAHA. *PLoS Pathog* 10:e1003997.
94. Stoszek M, Ne E, Abner E, Mahmoudi T. 2019. A broad drug arsenal to attack a strenuous latent HIV reservoir. *Curr Opin Virol* 38:37–53.
95. Sturgill JL, Mathews J, Scherle P, Conrad DH. 2011. Glutamate signaling through the kainate receptor enhances human immunoglobulin production. *J Neuroimmunol* 233:80–89.

96. Zona C, Ciotti MT, Avoli M. 1997. Topiramate attenuates voltage-gated sodium currents in rat cerebellar granule cells. *Neurosci Lett* 231:123–126.
97. White HS, Brown SD, Woodhead JH, Skeen GA, Wolf HH. 2000. Topiramate modulates GABA-evoked currents in murine cortical neurons by a nonbenzodiazepine mechanism. *Epilepsia* 41:17–20.
98. Nallani SC, Glauser TA, Hariparsad N, Setchell K, Buckley DJ, Buckley AR, Desai PB. 2003. Dose-dependent induction of cytochrome P450 (CYP) 3A4 and activation of pregnane X receptor by topiramate. *Epilepsia* 44:1521–1528.
99. Verdin E, Paras P, Van Lint C. 1993. Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. *EMBO J* 12:3249–3259.
100. Morgan M, Anders S, Lawrence M, Aboyoun P, Pagès H, Gentleman R. 2009. ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics* 25:2607–2608.
101. Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* 3:1101–1108.
102. Ne E, Crespo R, Izquierdo-Lara R, Kocer S, Gorszka A, van Staveren T, Kan TW, Dekkers D, Rokx C, Palstra R-J, Demmers J, Mahmoudi T. 2020. dCas9 Targeted Chromatin and Histone Enrichment for Mass Spectrometry (Catchet-MS) Identifies IKZF1 as a Novel Drug-Able Target for HIV-1 Latency Reversal. ID 3553972. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Figure 1

A 1. Latency establishment

B Uninfected KBM7 cells → Infection and GFP- sorting → Stimulation (+LRAs) and GFP+ sorting

C

E

D

2. Gene trap mutagenesis

F

3. double positive GFP/mCherry cells sorting and sequencing

Figure 2

Figure 3

Figure 4

S5**A** **B**

A bioRxiv preprint doi: <https://doi.org/10.1101/2021.01.20.427543>; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 1

bioRxiv preprint doi: <https://doi.org/10.1101/2021.01.20.427543>; this version posted January 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Candidate gene list GFP^{Sub}

Gene name	Sub population			Total population				Expr (100*FPKM) 1/2		Knock Down		GFP-FACS		HIV genes expression			Full name	HIV-1 Interactions
	A	B1	B2	C	D1	D2	E	WBC	Jurkat	A2	11.1	A2	11.1	GAG	POL	TAT		
AC005616.1	3,4	3,7	3,4	0,0	1,9	0,0	3,4	0.58	Na								Non-coding mRNA	v
AC079466.1	3,4	3,5	2,2	0,0	1,6	4,3	9,7	NA	NA								Non-coding mRNA	
ADK	3,3	32,2	1,0	0,0	1,7	1,7	1,0	16	469.5	v	v	v	v	v	v	v	Adenosine Kinase	
CAPN9	3,3	0,0	4,0	0,0	0,0	0,0	0,0	0.8	3								Calpain 9	
CHD9	13,3	6,8	3,9	0,0	16,7	1,0	1,3	15	672	v	v	v	v	v	v	v	Chromodomain Helicase DNA Binding Protein 9	
CTD-2066L21.3	3,7	4,1	74,0	3,6	0,0	27,7	5,1	NA	NA								Non-coding mRNA	
ESYT1	3,5	3,5	0,0	0,0	4,5	6,1	0,0	81	2328								Extended Synaptotagmin 1	
FAM19A2	8,4	8,7	9,0	0,0	0,0	28,9	5,8	NA	24.5	v	v	v	v	v	v	v	Family With Sequence Similarity 19 Member A2	
FGF1	3,7	4,1	1,2	0,0	4,3	4,3	0,0	3	0								Fibroblast Growth Factor 1	
GRIK5	3,3	4,8	2,8	0,0	0,0	0,0	0,0	4	9	v	v	v	v	v	v	v	Glutamate Ionotropic Receptor Kainate Type Subunit 5	
IRF2BP2	3,4	3,6	3,5	0,0	0,0	0,0	0,0	9	2205	v	v	v	v	v	v	v	Interferon Regulatory Factor 2 Binding Protein 2	v
NF1	14,0	2,5	4,1	1,1	2,8	3,6	4,1	10	1452	v	v	v	v	v	v	v	Neurofibromin 1	
RP11-355N15.1	19,1	9,2	9,2	4,7	3,2	6,4	0,0	NA	NA								Non-coding mRNA	
RP11-755B10.4	3,3	1,0	6,8	0,0	4,4	1,0	0,0	NA	NA								Non-coding mRNA	
RP4-569M23.2	3,0	0,0	4,8	0,0	0,0	3,3	0,0	NA	NA								Non-coding mRNA	
RP4-613B23.1	6,8	4,6	0,0	0,0	4,2	0,0	0,0	NA	NA								Non-coding mRNA	
SCN9A	3,5	3,1	1,7	0,0	0,0	0,0	0,0	4	2.5	v	v	X	X				Sodium Voltage-Gated Channel Alpha Subunit 9	
SPN	11,5	2,8	6,0	0,0	0,0	1,0	0,0	60	8743	v	v	X	X				Sialophorin	
ZNF385D	3,9	4,1	1,0	0,0	0,0	1,1	0,0	7	3								Zinc Finger Protein 385D	

Candidate gene list GFP^{Total}

Gene name	Sub population			Total population				Expr (100*FPKM) 1/2		Knock Down		GFP-FACS		HIV genes expression			FullName	HIV-1 Interactions
	A	B1	B2	C	D1	D2	E	WBC	Jurkat	A2	11.1	A2	11.1	GAG	POL	TAT		
AC004381.7	0,0	1,2	0,0	0,0	0,0	3,0	5,8	NA	NA								Non-coding mRNA	v
AC018890.6	0,0	6,8	4,3	0,0	4,0	9,4	3,6	NA	NA								Non-coding mRNA	
AC079466.1	3,4	3,5	2,2	0,0	1,6	4,3	9,7	NA	NA								Non-coding mRNA	
AC084018.1	0,0	2,5	2,8	0,0	0,0	5,4	4,1	NA	NA								Non-coding mRNA	
ACSM3	1,1	3,2	8,5	0,0	1,3	11,5	3,4	5	9								Acyl-CoA Synthetase Medium Chain Family Member 3	
ALDH1A2	9,0	1,2	2,2	4,6	3,7	3,3	0,0	6	833								Aldehyde Dehydrogenase 1 Family Member A2	
AVL9	0,0	3,5	1,3	3,4	3,1	2,3	0,0	15	420								AVL9 Cell Migration Associated	
C17orf64	0,0	8,6	16,7	0,0	15,3	13,8	5,4	2	0								Chromosome 17 Open Reading Frame 64	
C9orf139	1,1	2,4	2,1	8,3	3,0	10,6	1,0	NA	12								Chromosome 9 Open Reading Frame 139	
CCDC26	1,0	1,6	1,2	0,0	4,1	2,0	9,6	6	NA								CCDC26 Long Non-Coding RNA	v
CDH23	0,0	12,1	3,5	2,2	4,1	3,4	4,2	15	57								Cadherin Related 23	
CIITA	0,0	1,8	6,0	3,3	2,7	2,9	4,2	26	1								Class II Major Histocompatibility Complex Transactivator	
CMSS1	0,0	3,1	20,0	6,4	12,5	7,8	7,1	8	1066	v	v	v	v	v	v	v	Cms1 Ribosomal Small Subunit Homolog	
COP55	0,0	0,0	1,0	0,0	5,5	1,1	6,6	52	1337	v	v	X	X				COP9 Signalosome Subunit 5	
CTD-2066L21.3	3,7	4,1	74,0	3,6	0,0	27,7	5,1	NA	NA								Non-coding mRNA	
CTD-2260A17.2	3,7	0,0	1,7	0,0	0,0	5,2	4,5	NA	NA								Non-coding mRNA	
CTD-2370N5.3	2,7	13,2	3,7	2,3	6,0	8,8	3,0	NA	NA								Non-coding mRNA	
CTD-2376I20.1	0,0	3,5	0,0	0,0	4,5	0,0	5,0	NA	NA								Non-coding mRNA	
CTD-3187F8.14	0,0	0,0	1,1	3,5	4,6	0,0	0,0	NA	NA								Non-coding mRNA	
CYP2F2P	0,0	5,3	13,0	1,3	0,0	8,2	5,2	4	NA								Cytochrome P450 Family 2 Subfamily F Member 2, Pseudogene	v
DEF6	0,0	0,0	1,0	5,8	4,2	0,0	0,0	78	1336								DEF6 Guanine Nucleotide Exchange Factor	
DOCK11	0,0	0,0	1,0	0,0	1,6	14,9	5,8	53	635.5								Dedicator Of Cytokinesis 11	
EGFL7	0,0	0,0	3,9	3,4	5,4	1,5	0,0	11	79								EGF Like Domain Multiple 7	
EVI2B	1,9	8,0	2,5	1,4	2,5	6,3	3,3	234	207.5	v	v	v	v	v	v	v	Ecotropic Viral Integration Site 2B	
EVL	0,0	3,1	8,1	0,0	7,1	8,7	4,9	42	2656.5	v	v	v	X				Enah/Vasp-like	
EXOSC8	0,0	1,1	0,0	0,0	0,0	8,7	3,9	37	1068.5	v	v	v	v	v	v	v	Exosome component 8	
FAM19A2	8,4	8,7	9,0	0,0	0,0	28,9	5,8	NA	24.5	v	v	v	v	v	v	v	Family With Sequence Similarity 19 Member A2	
GSG1	0,0	0,0	3,8	0,0	1,5	3,1	4,2	NA	20								Germ Cell Associated 1	
LINC00922	1,2	0,0	1,0	3,2	0,0	1,1	5,6	5	NA								Non-coding mRNA	
LY86-AS1	0,0	8,9	6,8	0,0	6,6	9,5	8,2	2	8.5								LY86 Antisense RNA 1	v
NF1	14,0	2,5	4,1	1,1	2,8	3,6	4,1	10	1452	v	v	v	v	v	v	v	Neurofibromin 1	
NINJ2	0,0	2,6	4,5	0,0	0,0	9,1	5,2	35	9.5								Ninjurin 2	
NPHP3	0,0	3,6	1,0	3,4	0,0	6,4	0,0	13	64.5								Nephrocystin 3	
OXR1	0,0	2,3	1,0	0,0	1,1	3,6	5,8	38	563.5								Oxidation Resistance 1	
PAIP1	0,0	0,0	1,0	0,0	2,7	7,2	5,5	39	809								Poly(A) Binding Protein Interacting Protein 1	
RHOF	0,0	9,7	41,2	0,0	0,0	14,8	5,5	21	343.5	v	v	X	X				Ras Homolog Family Member F, Filopodia Associated	
RP11-297K8.2	0,0	3,2	0,0	10,5	22,9	0,0	0,0	NA	NA								Non-coding mRNA	
RP11-336K24.12	0,0	1,3	2,4	0,0	4,4	5,6	3,8	NA	NA								Non-coding mRNA	
RP11-355N15.1	19,1	9,2	9,2	4,7	3,2	6,4	0,0	NA	NA								Non-coding mRNA	
RP11-68I3.2	0,0	4,5	2,5	0,0	5,0	0,0	3,6	NA	NA								Non-coding mRNA	v
RP11-691H4.4	0,0	3,5	6,7	3,2	0,0	3,7	0,0	NA	NA								Non-coding mRNA	
RPN2	0,0	0,0	1,5	3,1	0,0	4,3	0,0	94	5373.5								Ribophorin II	
SDCCAG8	0,0	0,0	1,0	0,0	1,0	5,6	3,4	35	210.5								Serologically Defined Colon Cancer Antigen 8	
SERTAD2	0,0	1,1	11,6	0,0	1,6	4,3	6,4	39	837								SERTA Domain Containing 2	
SLC7A6	0,0	1,2	1,3	0,0	3,9	6,5	7,4	28	1480								Solute Carrier Family 7 Member 6	
SPESP1	0,0	1,4	1,5	0,0	0,0	3,1	4,9	10	0								Sperm Equatorial Segment Protein 1	
ST3GAL3	0,0	0,0	1,2	1,0	1,0	4,6	3,2	24	294								ST3 Beta-Galactoside Alpha-2,3-Sialyltransferase 3	
STRADB	0,0	2,1	7,6	0,0	3,2	3,8	6,1	28	659.5								STE20 Related Adaptor Beta	
TBC1D1	0,0	0,0	1,0	0,0	0,0	4,7	14,3	74	1437								TBC1 Domain Family Member 1	
TRAK2	0,0	4,7	1,7	0,0	4,1	2,9	5,9	39	672								Trafficking Kinesin Protein 2	v
TRAPPC6A	0,0	0,0	0,0	0,0	3,9	3,8	9,8	41	218								Trafficking Protein Particle Complex 6A	
UFC1	0,0	1,3	0,0	0,0	2,8	15,9	3,2	88	1380.5								Ubiquitin-Fold Modifier Conjugating Enzyme 1	
UNK	2,9	2,1	0,0	0,0	3,8	4,0	6,1	17	633.5								Unkempt homolog(drosophila)	
USP15	0,0	0,0	3,4	0,0	1,1	7,2	5,0	81	816.5	v	v	v	v	v	v	v	Ubiquitin Specific Peptidase 15	
WIPF1	0,0	1,8	3,1	0,0	3,0	3,4	5,5	146	1356								WAS/WASL Interacting Protein Family Member 1	