Bimodal Expression Patterns, and Not Viral Burst Sizes, Predict the Effects of Vpr on HIV-1 Proviral Populations in Jurkat Cells

Edmond Atindaana,[a] Abena Kissi-Twum,[b,c] Sarah Emery,[d] Cleo Burnett,[a] Jake Pitcher,[a] Myra Visser,[a] Jeffrey M. Kidd,[d] Alice Telesnitsky[a]

[a]Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan, USA
[b]West African Centre for Cell Biology of Infectious Pathogens (WACCBIP), University of Ghana, Legon, Greater Accra Region, Ghana
[c]Department of Biochemistry, Cell and Molecular Biology, University of Ghana, Legon, Greater Accra Region, Ghana
[d]Human Genetics, University of Michigan Medical School, Ann Arbor, Michigan, USA

ABSTRACT Integration site landscapes, clonal dynamics, and latency reversal with or without *vpr* were compared in HIV-1-infected Jurkat cell populations, and the properties of individual clones were defined. Clones differed in fractions of long terminal repeat (LTR)-active daughter cells, with some clones containing few to no LTR-active cells, while almost all cells were LTR active for others. Clones varied over 4 orders of magnitude in virus release per active cell. Proviruses in largely LTR-active clones were closer to preexisting enhancers and promoters than low-LTR-active clones. Unsurprisingly, major *vpr*+ clones contained fewer LTR-active cells than *vpr*− clones, and predominant *vpr*+ proviruses were farther from enhancers and promoters than those in *vpr*− pools. Distances to these marks among intact proviruses previously reported for antiretroviral therapy (ART)-suppressed patients revealed that patient integration sites were more similar to those in the *vpr*+ pool than to *vpr*− integrants. Complementing *vpr*-defective proviruses with *vpr* led to the rapid loss of highly LTR-active clones, indicating that the effect of Vpr on proviral populations occurred after integration. However, major clones in the complemented pool and its *vpr*− parent population did not differ in burst sizes. When the latency reactivation agents prostratin and JQ1 were applied separately or in combination, *vpr*+ and *vpr*− population-wide trends were similar, with dual-treatment enhancement being due in part to reactivated clones that did not respond to either drug applied separately. However, the expression signatures of individual clones differed between populations. These observations highlight how Vpr, exerting selective pressure on proviral epigenetic variation, can shape integration site landscapes, proviral expression patterns, and reactivation properties.

IMPORTANCE A bedrock assumption in HIV-1 population modeling is that all active cells release the same amount of virus. However, the findings here revealed that when HIV-infected cells expand into clones, each clone differs in virus production. Reasoning that this variation in expression patterns constituted a population of clones from which differing subsets would prevail under differing environmental conditions, the cytotoxic HIV-1 protein Vpr was introduced, and population dynamics and expression properties were compared in the presence and absence of Vpr. The results showed that whereas most clones produced fairly continuous levels of virus in the absence of Vpr, its presence selected for a distinct subset of clones with properties reminiscent of persistent populations in patients, suggesting the possibility that the interclonal variation in expression patterns observed in culture may contribute to proviral persistence *in vivo*.

KEYWORDS HIV-1 expression properties, HIV-1 persistence, HIV-1 population dynamics, integration sites, latency reactivation, latency-reversing agents, LRAs

HIV-1 establishes stable reservoirs in patients treated with antiretroviral therapy (ART), which consist of cells containing replication-competent proviruses that are not cleared by the immune system and that can rekindle spreading infection (1). It is generally assumed that the latent reservoir consists of proviruses that are transcriptionally silent, but what causes this is unclear. Reservoir establishment and maintenance are multifaceted and may involve the intracellular depletion of transcription factors, viral regulatory protein deficiencies, provirus integration position effects, and epigenetic variation (2–12). The oligoclonal nature of ART-suppressed patients' proviruses suggests that the reservoir's long-lived nature is at least partially due to infected cells' proliferation, which may be either homeostatic or driven by T cell receptor engagement (13–15).

Fundamental aspects of latency in patients, such as the size of the latent reservoir, remain poorly defined. Method improvement for reservoir quantification is ongoing, as recent studies have demonstrated that quantitative viral outgrowth assays (QVOAs) underestimate the latent reservoir, while PCR-based quantification can overestimate it due to the predominance of defective proviruses (16–18). Experimental approaches for reactivation must sometimes be applied multiple times to achieve reactivation *ex vivo* (19). Moreover, reported genetic dissimilarities between *ex vivo* outgrowth virus and reemergent viremia suggest that existing reactivation approaches may not accurately detect rebound-competent virus (20, 21). Large interclonal differences in virion release per T cell and variability in virion release upon latency disruption may further complicate reservoir size determination (22, 23).

The rarity of latently infected cells in virally suppressed individuals makes the study of *in vitro* latency models necessary. Some tissue culture models for HIV-1 latency and reactivation, such as the widely used J-Lat clones, assess long terminal repeat (LTR) promoter activity by reporter gene expression but lack genetic elements believed to be dispensable for HIV-1 latency (24–26). One such element is the *vpr* gene, whose product plays roles in viral infectivity *in vivo* (27–30) but also causes cell cycle arrest and can induce widespread changes in host gene expression (31–33). Many latency models use *vpr*-defective proviruses, which, when cultured over time to allow proviral silencing, can be used for reactivation studies (34–37). However, in such systems, the extent to which proviral quiescence represents the silencing and outgrowth of a subset of integrant clones versus global proviral silencing is unknown.

Cell-based latency models have been used to pilot candidate cure strategies, including "shock-and-kill" and "block-and-lock" approaches. In contrast to ART, which prevents viral spread, the shock-and-kill method involves inducing virus expression with the intention that this will lead to either cytopathic death of reactivated cells or host immune recognition and infected cell clearance (38). Candidate latency reactivation approaches that perturb cellular pathways or complement intracellular deficiencies in experimental models include the use of prostratin, which stimulates T cells without inducing cellular proliferation and increases the level of NF-κB (39, 40). Other latency-reversing agents (LRAs) include those that act to increase the level of pTEFb, including the BET bromodomain inhibitor JQ1 (41), as well as treatments that modify the chromatin environment, such as histone deacetylase inhibitors like suberoyl anilide hydroxamic acid (SAHA) and entinostat (42–46). In contrast, the block-and-lock approach seeks to permanently silence proviruses to prevent their reactivation (47).

Molecules that are effective at reversing latency in various tissue culture models have been identified, but evidence thus far suggests either that these are too toxic to be therapeutically useful or that they fail to reduce the size of the latent reservoir in patients (46, 48–51). This discordance may be due in part to differences among cell-based latency models, as some use infectious virus while others use viral reporters, and some use clonal integrants while others use heterogeneous proviral populations (25). Thus, the inconsistent results in LRA reactivation studies using these models may reflect that each model captures at best a subset of the expression properties that contribute to the latent reservoir (52), and whereas the effects on heterogeneous
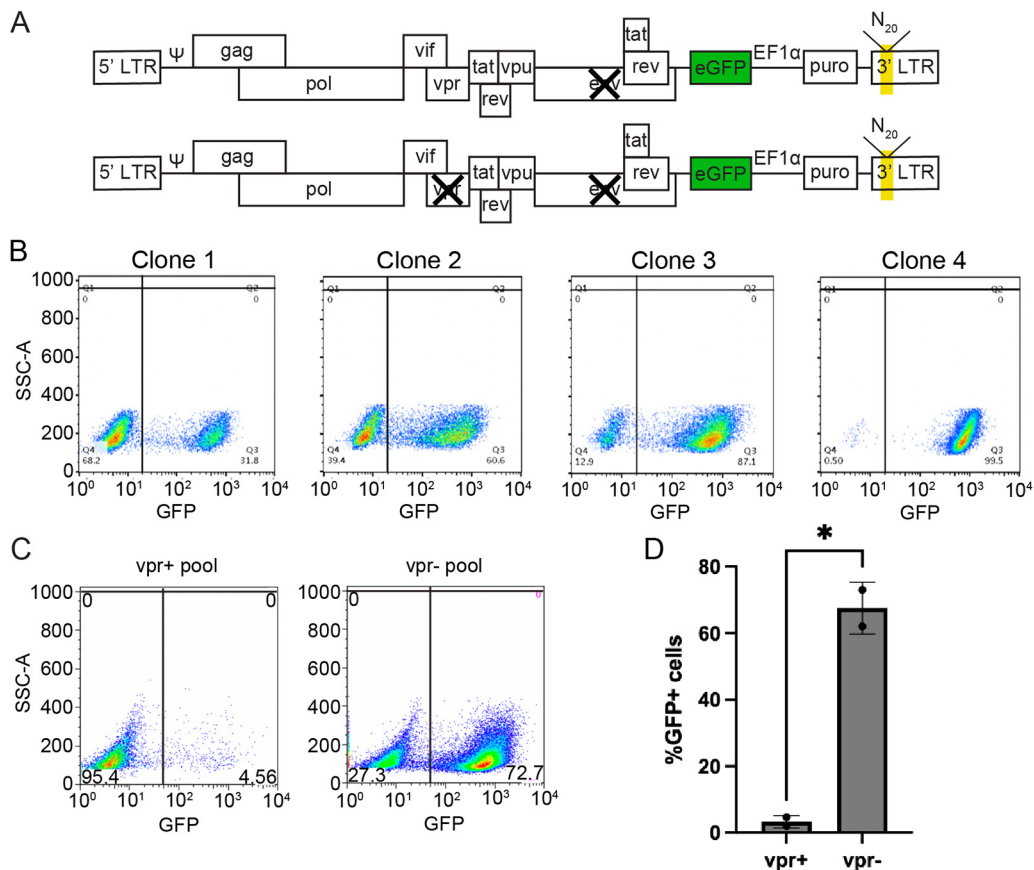
FIG 1 Generation of zip-coded HIV-1 proviral pools. Polyclonal cell pools containing *vpr*+ and *vpr*− proviruses were established in Jurkat T cells and compared. (A) Schematic of vectors used to produce zip-coded virus (not to scale). The yellow shading in the 3′ LTR indicates the position of the 20-bp randomized barcode in U3. eGFP, enhanced green fluorescent protein. (B) *vpr*− clone expression properties. Three clones were isolated by limiting dilution of *vpr*− pool cells, expanded for 14 days, and subjected to flow cytometry. The flow plots show that clonal pools 1 to 3 have different percentages of GFP+ cells. The *x* axis indicates the GFP signal detected in the FITC channel; numbers in each gate indicate the percentages of cells gated GFP− or GFP+. SSC, side scatter. (C) Representative flow cytometry plots of polyclonal Jurkat T cells generated by infection with barcoded *vpr*+ (left) or *vpr*− (right) viruses that were puromycin selected for 4 days and expanded for an additional 10 days. (D) Bar graph comparing the percentages of GFP+ puromycin-resistant cells in polyclonal *vpr*+ and *vpr*− populations. Data represent results for two *vpr*+ pools and two *vpr*− pools (* indicates a *P* value of <0.05 by a paired *t* test).

populations of integrants have been examined extensively, the contributions of individual clones' behavior to aggregate population responses largely have not been explored.

In the current study, the HIV-1 expression properties of hundreds of individual integrant clones were compared within polyclonal populations of barcoded proviruses. The influences of integration sites and the presence or absence of *vpr* on the populations' clonal structures and their reactivation dynamics were investigated. Together, these findings suggest that Vpr's cytotoxic activity disproportionately affects a distinct subset of infected cells and plays critical roles in both shaping transcriptionally inactive proviral populations and defining their reactivation potential.

RESULTS

***vpr*− and *vpr*+ proviral pools differed in their numbers of LTR-active cells.** To study the effects of Vpr on the responsiveness of individual proviruses in the absence of virus spread, Jurkat cells were infected with distinguishable *vpr*-positive (*vpr*+) or *vpr*-negative (*vpr*−) versions of the NL4-3 derivative shown in Fig. 1A (53). The use of the EF1α promoter to drive the constitutive expression of the puromycin resistance gene allowed the selection of infected cells independent of LTR expression. Each

genomic RNA in an infecting virus contained a unique 20b randomized "barcode" inserted into U3, which was duplicated in both LTRs after reverse transcription and enabled the tracking of individual proviruses. We refer to the barcodes as "zip codes" because in the context of proviruses, they report the genetic neighborhoods of each integrant. Infected cell populations were passaged without cell cloning, thus generating polyclonal integrant populations within which transcriptionally active cells were identified using LTR-driven green fluorescent protein (GFP) expression or by progeny genomic RNA released in Env$^-$ virions.

It has long been recognized that at least some HIV-1 integrants establish clones in which a fraction of the total daughter cells possess active LTRs while proviruses are silenced in other sibling cells (54). Previous work using zip-coded vpr^- derivatives of the vector used here has shown that for each integrant, the clone gives rise to a mixture of cellular progeny that includes some GFP$^+$ cells and some GFP$^-$ cells and that over time sibling cells switch between LTR-active (GFP$^+$) and -inactive (GFP$^-$) expression phenotypes while maintaining LTR-active and -inactive cells in stable equilibrium proportions (53). To confirm that individual integrant clones in the current study also contained clone-specific mixtures of GFP$^+$ and GFP$^-$ cells, single cells were isolated by limiting dilution from the vpr^- pool and expanded, and cells from each clone were then subjected to flow cytometry (Fig. 1B). Consistent with previously reported differences in bimodal expression patterns among proviral clones (53), the results showed that each clonal pool was comprised of both GFP$^+$ and GFP$^-$ cells, with LTR-active proportions being discernible by GFP$^+$ percentages (%GFP$^+$) that differed among the clones: mostly GFP$^+$ cells for some clones (high-%GFP$^+$ clones) and distinctly different GFP$^+$ percentages for others (Fig. 1B). To confirm that these cells were clonal and contained single proviruses and, thus, that the observed mixed cell phenotypes did not reflect the presence of more than one clone, PCR amplicons from these clones were Sanger sequenced without subcloning and determined to display unique sequences, indicating that at least a large majority of each clone's cells contained a unique zip code (see Materials and Methods).

In the current study, four independent polyclonal integrant pools were established by infecting roughly 5×10^6 Jurkat T cells at a low multiplicity of infection (<0.0005) to ensure that puromycin-resistant infected cells contained only one provirus per cell. Two of the pools contained vpr^+ proviruses, and two had integrants lacking vpr. Flow cytometry analysis after 14 days of expansion showed significantly fewer GFP$^+$ cells in the two vpr^+ infected cell pools than in the two vpr^- provirus pools ($<5\%$ versus 73% GFP$^+$, respectively; $P = 0.042$ by a paired t test) (Fig. 1C and D). To address the possibility that uninfected cell outgrowth might have contributed to GFP$^-$ cell populations, an uninfected Jurkat cell control was cultured in parallel and subjected to the same schedule of puromycin treatment and selection-free medium exposure as that for the infected cells. After a total of 14 days, these uninfected control cultures were analyzed by flow cytometry to ensure that no surviving uninfected cells were detectable that might confound observations. In an additional test of the possible presence of uninfected cells, dual aliquots of each pool were analyzed about 30 days after pool establishment. One aliquot was cultured with puromycin, and the other was left untreated. When analyzed by propidium iodide (PI) staining for cell death, no difference was observed between treated and untreated pool controls.

At 14 days postinfection, vpr^- and vpr^+ cell pools were each sorted into GFP$^+$ and GFP$^-$ subpools, and cellular DNA was harvested from an aliquot of each subpool immediately after sorting. The integrants' zip codes were amplified from the cellular DNA samples in at least two separate PCRs per cellular DNA sample, and the subpools' zip code contents were catalogued by high-throughput sequencing. Correlation analysis for the fractional abundances of specific zip codes in replicate reactions showed reproducible zip code detection (see Fig. S1 in the supplemental material).

After analyzing roughly 3 million sequencing reads per pool, zip codes were ordered by read abundance. Determining how many unique barcodes were present in
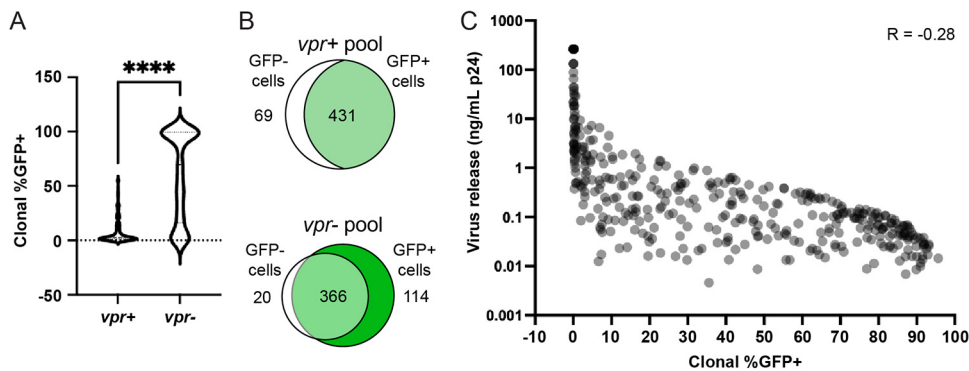
FIG 2 Expression properties of individual clones within the vpr^+ and vpr^- pools. (A) Comparison of the numbers of high- and low-LTR-active clones in vpr^+ and vpr^- pools. The fractions of total cells in each clone that sorted GFP$^+$ (LTR active) were calculated for the 500 most abundant clones in each pool by quantifying the fraction of each zip code (integrated barcode) in the GFP$^+$ sorted cells' genomic DNA to its quantity in the unsorted pool by high-throughput sequencing and normalizing for the percentage of the unsorted pool that was GFP$^+$ (see Materials and Methods). Violin plots compare the %GFP$^+$ values (indicative of the clones' fraction of LTR-active daughter cells) for the top 500 most abundant zip codes in each pool (**** indicates a P value of <0.0001 by a Mann-Whitney U two-tailed test). (B) Venn diagrams of the numbers of independent zip codes in GFP$^+$ and GFP$^-$ sorted subpools for vpr^+ and vpr^- pools. The number in the white region indicates the number of different zip codes observed only in GFP$^-$ sorted cells, that in the light-green region of the intersection represents the number of zip codes that were present in both GFP$^+$ and GFP$^-$ sorted subpools, and the number in the dark-green region represents zip codes present in only the green sorted cells. The 500 reported zip codes in each pool represent those present in the top 78% and 88% of reads for vpr^+ and vpr^- pools, respectively, when clones were ordered by read abundance. (C) Scatterplot displaying each clone's % GFP$^+$ value on the x axis and virus release per GFP$^+$ cell on the y axis (Pearson's $R = -0.28$).

each GFP$^+$ pool revealed that similar numbers of zip codes (approximately 2,000), each indicative of a single integration event, were detected in the GFP$^+$ sorted cells from both the vpr^+ and vpr^- pools, even though the GFP$^+$ cell fraction of the vpr^+ pool was very small. Because all four pools had been infected at the same multiplicity of infection, this finding suggested that the numbers of GFP$^+$ integrants initially established did not differ markedly between pools, which is consistent with previous work in dendritic cells that indicates that the extent of proviral integration does not differ depending on the presence or absence of Vpr (55).

Combining population-wide percentages of GFP$^+$ cells in the unsorted cells, as determined by flow cytometry, with zip code read counts within sorted subpool libraries allowed calculating the percentage of LTR-active cells (%GFP$^+$ values) within each cell clone (Fig. 2A). The clones' %GFP$^+$ values were calculated using zip code abundances in fluorescence-activated cell sorter (FACS)-sorted GFP$^+$ and GFP$^-$ cells' DNA and normalizing the values to population-wide GFP positivity levels at sort time, as described in Materials and Methods. Notably, for the top 500 most abundant clones in the unfractionated vpr^+ and vpr^- pools, the median %GFP$^+$ value was significantly lower in the vpr^+ pool (P = 0.0001 by a Mann-Whitney U two-tailed test), indicating that most cells in the vpr^+ pool were members of low-LTR-active clones (Fig. 2A). When GFP$^+$ sorted cells from vpr^+ and vpr^- pools were cultured further, no viable vpr^+ GFP$^+$ cells were detected after 3 days, and thus, although integrated zip codes in the freshly sorted cells were determined, the expression properties and clonal structures of vpr^+ GFP$^+$ cells could not be analyzed further. These results indicated that integration events that resulted in GFP expression were equally likely in vpr^+ and vpr^- proviral pools when the cells were examined early after infection. However, the depletion of the vpr^+ GFP$^+$ cell subpopulations suggested that whenever cells with vpr^+ proviruses switched phenotypes from LTR inactive to LTR active, the resulting GFP$^+$ cells died upon further cell culturing.

Vpr shaped the clonal structure of infected pools. The observation of fewer GFP$^+$ cells in the vpr^+ pool than in cells with vpr^- proviruses was not surprising due to Vpr's well-known cytotoxic effects. Along with *env*, *tat*, *nef*, and *vpu*, *vpr* is one of the HIV-1

genes reported to be cytotoxic in at least some contexts (56–58). However, despite Vpr's cytotoxicity, a small fraction of vpr^+ GFP$^+$ cells was observed among unsorted vpr^+ cells, even though sorted vpr^+ GFP$^+$ cells did not survive 3 days of culture.

One plausible reason why unsorted vpr^+ populations contained rare GFP$^+$ cells was that these cells arose via recent phenotypic switches from clones that were largely inactive. To test this possibility, zip codes in GFP$^+$ and GFP$^-$ sorted subpools were compared for both vpr^+ and vpr^- integrants (Fig. 2B). Ordering zip codes by their abundance in the unsorted pools and analyzing those that comprised the top 500 revealed that among vpr^- cells, about 73.2% of all the zip codes were found in both subpools, while 4% were observed only in GFP$^-$ cells and 22.8% were observed only in the GFP$^+$ subpool. In contrast, 86.2% of the vpr^+ cells' zip codes were found in both subpools, with 13.8% being observed only in the GFP$^-$ subpool and none of the zip codes being found exclusively in the GFP$^+$ subpool. This suggested that the small fraction of GFP$^+$ cells in the vpr^+ pool (Fig. 1C and D) resulted from the recent acquisition of LTR expression by cells from the larger GFP$^-$ cell pool. If the vpr^- pool is assumed to be relatively representative of a population that results when all clones are equally viable, this suggests that when initially integrated, proviruses whose daughter cells were largely or always GFP$^+$ were the dominant class of clones.

Clones differed in LTR-active cell burst sizes. Because the proviral vectors used in this study contained most HIV-1 genes and expressed packageable RNAs as well as a GFP reporter, expression properties could be measured by both GFP expression and virion release. To determine the amount of virus released per LTR-active cell in each clone, total virus release from the vpr^- GFP$^+$-fractionated cell pool was quantified by p24 release; the limited viability of the vpr^+ GFP$^+$ cell fraction prevented this analysis for the vpr^+ pool. Relative zip code amounts in virion RNA and in genomic DNA (gDNA) extracted from the GFP$^+$ cells were then determined and used to calculate the amount of virus released per LTR-active cell for each clone. These results are indicated in Fig. 2C, with each clone's release per LTR-active cell presented on the y axis and clones ordered by their member cells' %GFP$^+$ values on the x axis. All clones in Fig. 2C are represented in light gray; regions that appear to display darker shading indicate the presence of multiple clones at the same coordinates. We and others have previously reported significant differences among HIV-1-infected cell clones in the amount of virus released per active cell (23, 53), and the results here indicated that burst sizes for the integrant clones in the current study ranged over 4 orders of magnitude.

Notably, burst sizes were slightly negatively correlated with %GFP$^+$ values (Pearson's $R = -0.28$) (Fig. 2C), indicating that clones that displayed higher %GFP$^+$ values tended to produce fewer virions per LTR-active cell than low-%GFP$^+$ clones. Because the intactness of these proviruses was not addressed directly, it is possible that some of the observed variation in virus release reflected provirus defects. However, because this work quantified encapsidated viral RNAs, even low-virus-expressing proviruses must have remained largely intact, as proviruses with large internal deletions would have lost the ability to assemble viral particles. Another possible contributor to these differences may be that large burst sizes conferred some selection against high-LTR-active clones during the weeks of infected cell passage that preceded these measurements. Although Vpr, which was not present in these cells, is the HIV-1 protein best recognized as cytotoxic, the expression of the retained HIV-1 proteins or other elements may attenuate high-%GFP$^+$ clones in more subtle ways.

High-LTR-active clones' proviruses were in closer proximity to genome marks associated with active gene expression. The finding that LTR-active vpr^+ cells were rapidly lost suggested that proviruses that had integrated into more active regions of the host genome might be selected against within polyclonal vpr^+ populations. To examine this notion, we compared the proximities of vpr^+ and vpr^- integration sites to genomic features associated with active gene expression. First, zip code integration sites were determined using cellular DNA harvested 10 days after the establishment of each pool and mapped to 1,171 and 1,121 unique sites in the human genome for vpr^+ and vpr^- pool members, respectively (see Materials and Methods). Next, the
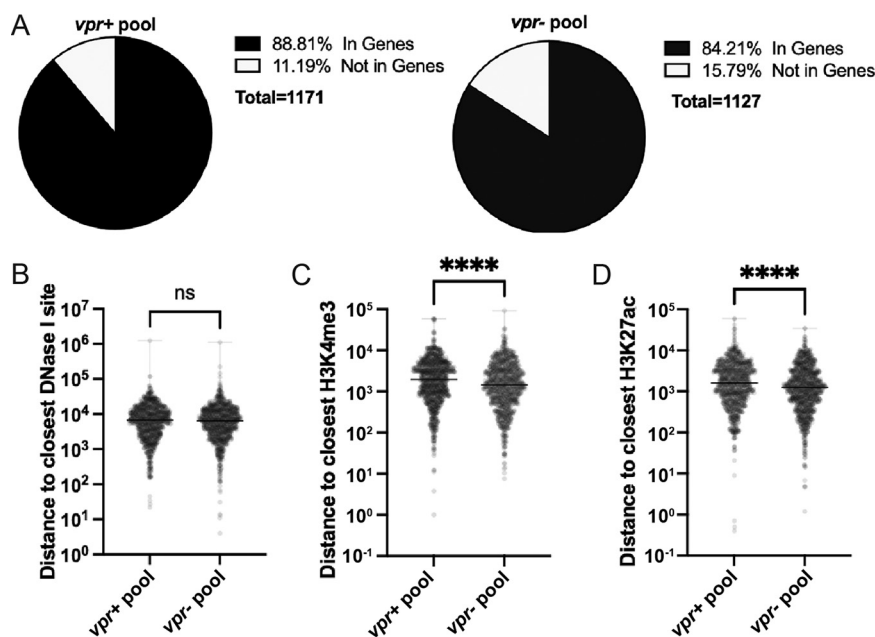
FIG 3 Comparison of integration sites in vpr^+ and vpr^- pools. Integration sites for both vpr^+ and vpr^- pool cells harvested at 10 days postinfection were determined, and distances to the nearest indicated genome features were mapped and compared. (A) Pie charts showing the percentage of integration sites found within genes versus those not found in genes for vpr^+ and vpr^- pools. (B to D) Box plots showing pairwise comparisons between vpr^+ and vpr^- integration sites' distances to the closest DNase I hypersensitivity site (means = 10,687 and 10,236, respectively) (B), H3K4me3 (means = 32,989 and 27,047, respectively) (C), and H3K27ac (means = 28,485 and 22,228, respectively) (D). A Mann-Whitney U two-tailed test was conducted for pairwise comparisons for panels B to D (ns [nonsignificant] and **** indicate P values of >0.05 and <0.0001, respectively) ($n = 1,171$ and 1,127 for vpr^+ and vpr^-, respectively).

proportions of zip codes located in genes versus those not in genes (as defined by ENCODE for Jurkat cells [59]) were compared (see Fig. 3A and Fig. S2 for the first and duplicate pools, respectively). Consistent with previous reports (60–62), the results indicated that similar majorities of integrants were established within genes regardless of whether or not vpr was present. Next, vpr^+ and vpr^- integrants were compared for their proximities to specific genome marks associated with active gene expression that have been reported to preexist in Jurkat cells (63). No differences were found in distances to the closest DNase I sensitivity sites, which are associated with open chromatin ($P = 0.1854$ by a Mann-Whitney U test) (Fig. 3B). However, the proximities to H3K27ac (associated with enhancers) ($P < 0.0001$ by a Mann-Whitney U two-tailed test) and H3K4me3 (associated with active promoters) ($P < 0.0001$ by a Mann-Whitney U two-tailed test) marks differed significantly, with proviruses from vpr^+ pools being somewhat farther from these marks (Fig. 3C and D).

To correlate the clones' LTR-active proportions with their proximities to active chromatin marks, distances from these marks were compared for high- and low-LTR-active clones. Few clones displayed intermediate %GFP$^+$ levels (Fig. 2A), and all major clones in the vpr^+ pool had clonal %GFP$^+$ proportions of 30% or lower. Therefore, only integrants from the vpr^- pool were examined, and low-LTR-active clones (%GFP$^+$ values of $<30\%$) were compared to high-LTR-active clones (%GFP$^+$ values of $\geq60\%$) (Fig. 4A). This analysis showed that the distances to both H3K27ac and H3K4me3 marks were shorter among high-LTR-active clones than among low-LTR-active clones, albeit not dramatically, suggesting that the difference in integration site proximities observed between vpr^+ and vpr^- proviruses may be due to the survival of high-LTR-active clones in the absence of Vpr.

Because the vpr^+ pool was devoid of high-LTR-active clones, its high- and low-LTR-active integrants could not readily be compared. However, with the reasoning that the
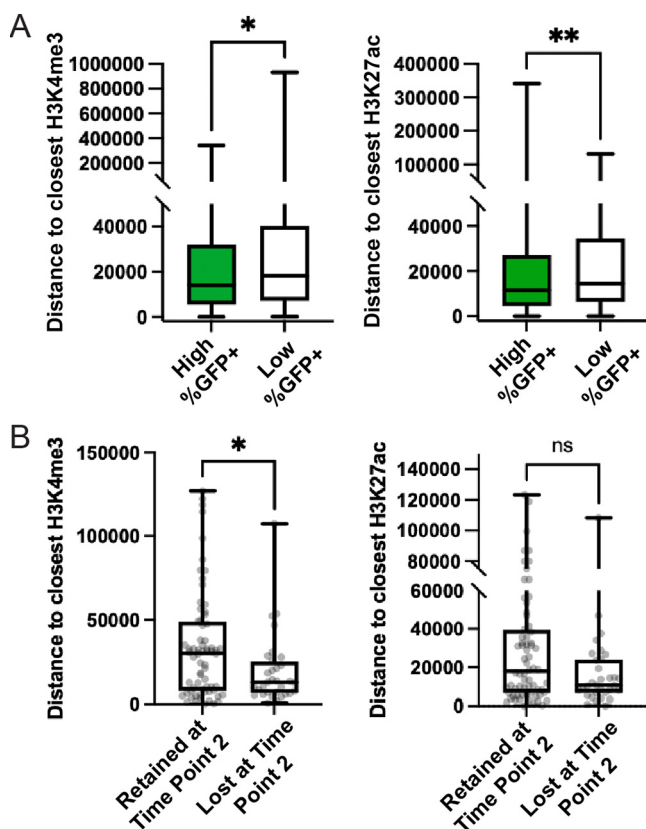
FIG 4 Integration site proximities to chromatin marks and their changes over time in *vpr*⁺ pools. (A) Distances to the indicated marks for high- and low-%GFP⁺ clones. The %GFP⁺ values for each zip code from the *vpr*⁻ pool were determined as described in Materials and Methods. Zip codes with % GFP⁺ values of ≥60% were binned as "high-LTR-active clones" (595 zip codes), and zip codes with % GFP⁺ values of <30% were binned as "low-LTR-active clones" (222 zip codes). Pairwise comparisons of the distances to the closest H3K4me3 (left) and H3K27ac (right) marks are shown. (B) Changes in predominant integration site features for the *vpr*⁺ pools over time. Among the top 100 most abundant zip codes in the unsorted *vpr*⁺ pool at time point 1 (10 days postinfection), 68 were retained among the top 100 at time point 2 (24 days postinfection), while the remaining 32 time point 1 top 100 clones were lost at time point 2 and supplanted by other, formerly less-abundant zip codes. Box plots compare the closest H3K4me3 or H3K27ac mark for the 68 time point 1 zip codes that were retained in the top 100 at time point 2 to the 32 time point 1 zip codes that had been lost at time point 2 (ns indicates a *P* value of >0.05 by a Mann-Whitney U two-tailed test).

abundances of residual high-LTR-active clones would gradually diminish over time, the dynamics of the *vpr*⁺ pool were examined by comparing samples harvested 2 weeks apart. Examination of the 100 most abundant zip codes in the unsorted *vpr*⁺ cell pool showed that only 68 of the top 100 zip codes present at day 10 postinfection (time point 1) were observed among the 100 most abundant zip codes on day 24. When integration site proximities to H3K4me3 and H3K27ac marks were compared, the 32 zip codes that were absent from the top 100 on day 24 (time point 2) (Fig. 4B) were significantly closer to H3K4me3 (*P* = 0.0193 by a Mann-Whitney U two-tailed test) and tended to be closer to H3K27ac marks (although not with statistical significance [*P* = 0.1581 by a Mann-Whitney U two-tailed test]) than the 68 time point 1 clones that remained within the top 100 most abundant clones at time point 2.

The spectra of LRA responses were similar for vpr⁺ and vpr⁻ populations, but clonal behaviors differed. To address possible functional consequences of *vpr*, the reactivation properties of LTR-inactive *vpr*⁻ and *vpr*⁺ subpools were compared. The LRAs prostratin, a protein kinase C agonist, and JQ1, a bromodomain inhibitor, were applied separately or in combination to the GFP⁻ subpopulations of each pool for 24 h. Reactivation for each treated GFP⁻ population was monitored by both determining the changes in the frequency of GFP⁺ cells, as measured by flow cytometry, and
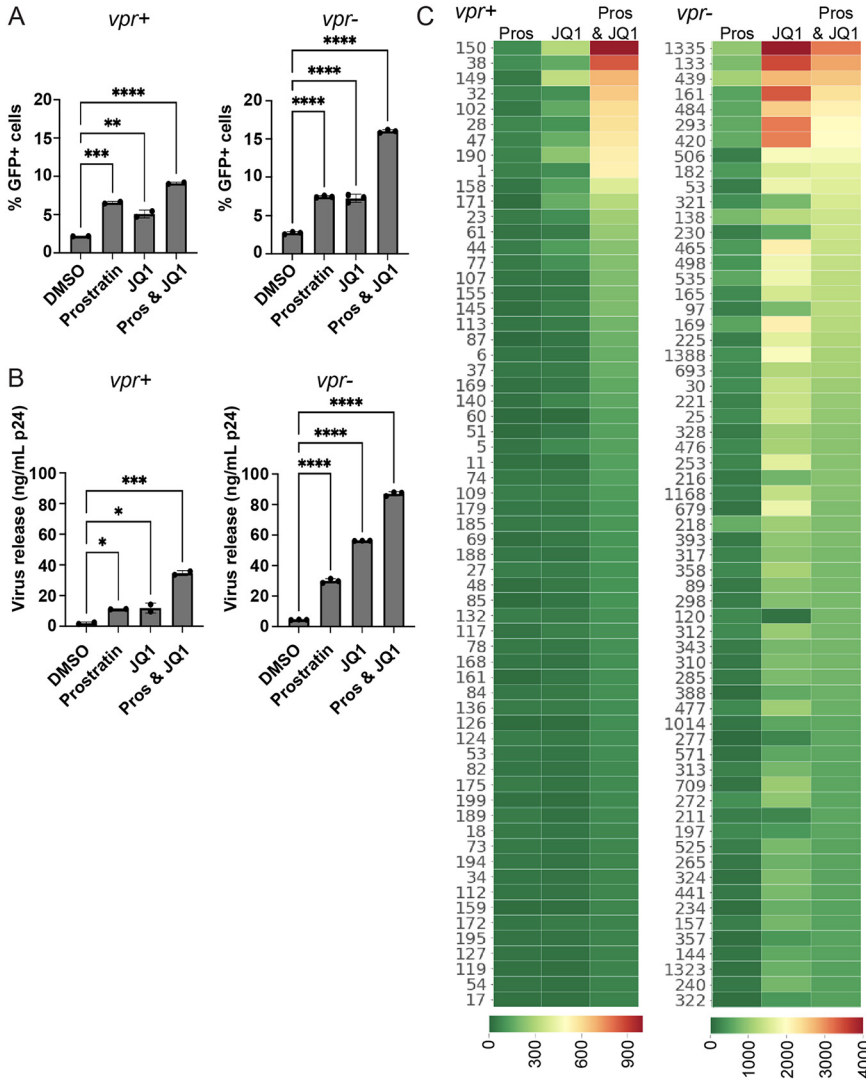
FIG 5 Effects of LRAs on zip-coded pools and clones. *vpr*⁺ and *vpr*⁻ GFP⁻ cell fractions were exposed to 0.1% DMSO, 2 μM JQ1, 2 μM prostratin, or a combination of prostratin and JQ1. Reactivation to GFP⁺ was measured at 24 h postinfection by flow cytometry, and virus release was quantified using a reverse transcriptase assay and normalized to define p24 levels. (A and B) Bar graphs showing the frequency of GFP⁺ cells after LRA treatment (left panel for each pool) (A) and the amount of virus released into the culture supernatant (right panel for each pool) (B) for the indicated polyclonal pools. The error bars show the means and standard deviations from two experimental treatment repetitions. (C) Heat maps of the clonal (zip code) virus release per treated GFP⁻ cell (left, *vpr*⁺ pool; right, *vpr*⁻ pool). Numbers at the left of each panel are clone identifiers generated by ordering proviral zip codes in decreasing relative abundances, as determined for the unsorted pools. Every row represents a unique cell clone's response. The clones were ordered from top to bottom by diminishing virus release per treated cell upon dual-LRA treatment. The color bar indicates the extent of release per treated cell based on p24 values in arbitrary units. Note that the same unit values were used for both *vpr*⁺ and *vpr*⁻ pools but that the color scales, as presented at the bottom of the panels, differ between pools.

quantifying virus release (Fig. 5A and B; Fig. S3A and B). The results indicated that compared to single prostratin and JQ1 treatments, dual treatment resulted in additive levels of reactivation in both *vpr*⁺ and *vpr*⁻ populations by both criteria (Fig. 5A shows reactivation monitored by GFP⁺ cells, and Fig. 5B shows virus release [the left panel indicates reactivation for *vpr*⁺, and the right panel shows reactivation for *vpr*⁻]). In dual-LRA treatments for both pools, there was an ~4-fold increase in GFP⁺ cells, while virus release increased by ~30-fold relative to the dimethyl sulfoxide (DMSO) control. The most significant difference between the *vpr*⁻ and *vpr*⁺ pools was that the absolute amount of virus release upon treatment was 3-fold higher in the *vpr*⁻ pools, and the

responsiveness to JQ1 was lower in the vpr^+ pools. These differences between pools in their extents of reactivation were not due to differences in cell viability (Fig. S4C). The observation that reactivation was enhanced by dual prostratin and JQ1 treatment is consistent with previous works by Boehm et al. and Darcis et al. using the same drugs in cell culture models of latency and *ex vivo*-treated cells from HIV-1 patients, respectively (24, 41).

Next, the behaviors of individual proviral clones within the populations were determined. Virus release from the treated cells was quantified by p24 equivalents, and cDNA was generated using virion genomic RNA upon LRA treatment. Zip codes were amplified from the viral cDNA and also from an untreated aliquot of the GFP$^-$ cells' DNA, and zip code libraries were high-throughput sequenced. The results were normalized to calculate the average virus release per treated cell for each clone (Fig. 5C; Fig. S3D).

This analysis revealed that many clones were not detectably reactivated. For example, of the 500 most abundant clones in the vpr^- pool, sequencing of virion RNA after reactivation treatment revealed no evidence of virus production for 102 of these abundant clones in the dually prostratin- and JQ1-induced virion cDNA pools. When cells were treated with both drugs, some clones in both the vpr^+ and vpr^- pools displayed enhanced virus release per treated cell compared to single-treatment conditions. Interestingly, both proviral pools included a subset of clones that were not detectably reactivated by either prostratin or JQ1 when the LRAs were applied alone but that were reactivated upon dual-LRA treatment.

Surprisingly, and in only the vpr^- pool, the reactivation levels observed under dual-LRA conditions were lower than those observed with single-LRA use for a subset of clones (Fig. 5C, right, compare, for example, the top dozen rows in the vpr^- columns to those for the vpr^+ clones). These same patterns of reactivation were observed when experiments were repeated using a second set of independently established vpr^- and vpr^+ pools (Fig. S3D). A possibility suggested by this pattern of reactivation is that Vpr degrades a negative regulator of HIV-1 gene expression that is induced in Jurkat cells by prostratin treatment. Indeed, Vpr has been reported to cause the depletion or mislocalization of several factors that can repress HIV-1 expression, including class I histone deadenylases, the transcription factor ZBTB2, and the negative regulator of pTEFb, CTIP2 (64–67).

Complementation of the *vpr*-defective pool led to depletion of high-LTR-active clones but did not discriminate against high-burst-size clones. As an additional test of the effects of *vpr* on polyclonal population composition, a functional copy of the *vpr* gene was added to cells harboring vpr^- proviruses at 16 days postinfection. This allowed a comparison of the behaviors of each of several hundred specific individual integrants in the presence or absence of *vpr*. This was achieved by transducing the vpr^- pool with a *tat*-deficient lentiviral vector containing LTR-driven *vpr*, which also contained the fluorescent marker *mKO* expressed from the constitutive simian virus 40 (SV40) promoter. This resulted in the expression of Vpr only when a preexisting vpr^- provirus was transcriptionally active. mKO (A monomeric version of the bioluminescent protein Kusabira Orange, which is derived from a Lithophyllon concinna gene) expression was used as a proxy for *vpr* vector transduction and enabled cells that were successfully transduced to be sorted in the phycoerythrin (PE) channel (Fig. 6A). At 48 h posttransduction, an unsorted aliquot was saved, while the remaining cells were sorted to identify GFP$^-$ PE$^+$ cells (Fig. S4A and B) for the work described below.

A comparison of the proportions of GFP$^+$ cells in the parental vpr^- pool to the proportions in the same pool "reverted" to vpr^+ by the introduction of the *vpr* vector (referred to here as the vpr^{rev} pool) showed a marked decrease in GFP$^+$ cells between day 2 and day 14 posttransduction (Fig. 6B). Although it cannot be ruled out that some feature of the transduction vector other than *vpr* influenced the cells' properties, the selective loss of GFP$^+$ transduced cells was consistent with predictions based on observations with the vpr^+ pool, which showed that LTR-active cells were depleted in the presence of Vpr.

The proviral zip code abundances in unsorted vpr^- and vpr^{rev} pools were then compared. Of the 500 most abundant clones in the vpr^- pool, zip codes were split into two
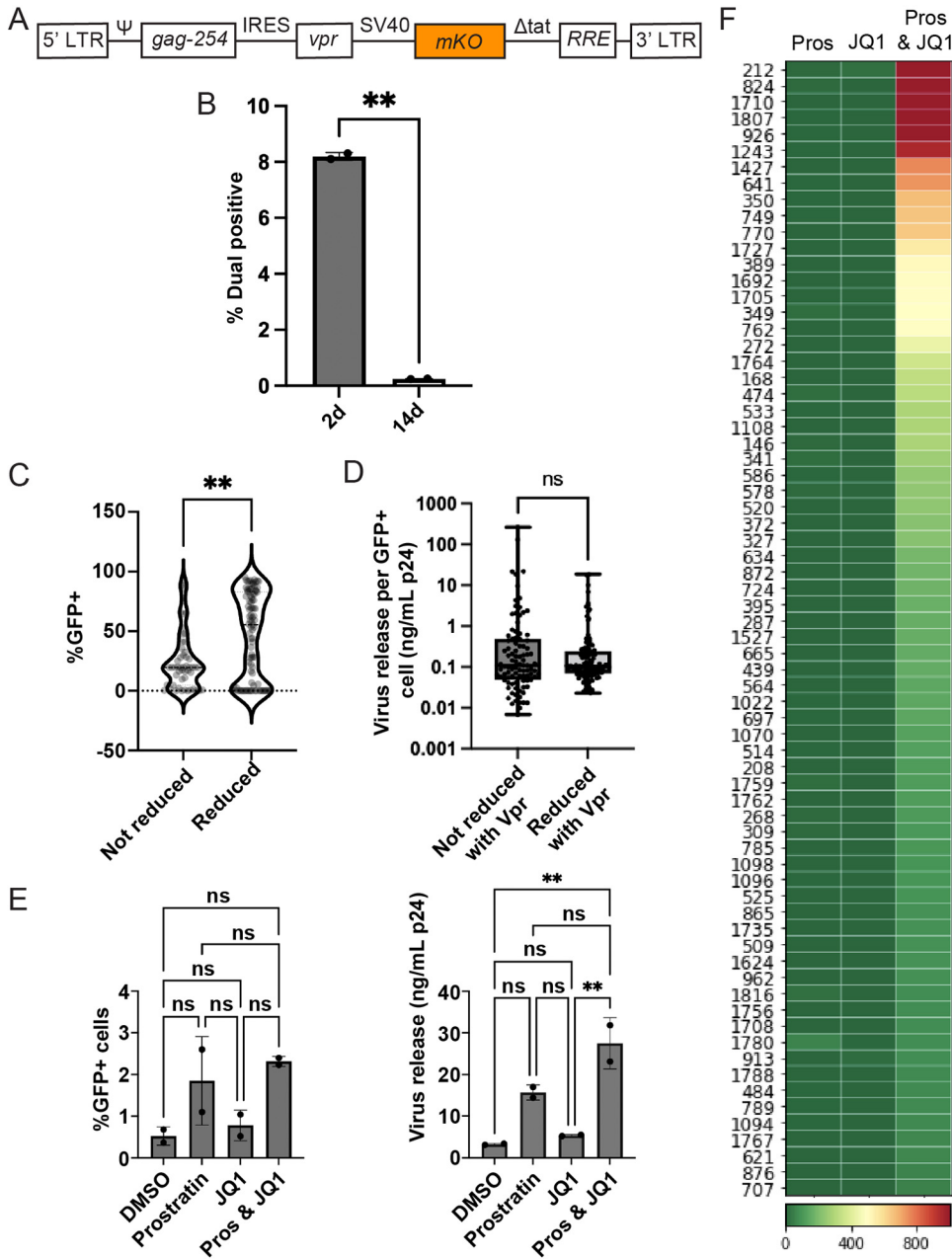
FIG 6 Complementation of vpr⁻ pools with a Vpr expression vector. (A) Schematic diagram of the Tat-deficient Vpr expression vector (not to scale). This vector was used to add vpr to vpr⁻ pools, and effects were determined by high-throughput analysis of the indicated viruses and cell samples. The vector contains IRES-vpr under the control of the HIV-1 LTR promoter and an SV40 promoter-driven mKO fluorescent reporter. (B) Population changes after vpr addition. Shown is a bar graph of the percentages of GFP⁺ cells for two polyclonal vpr^rev pools established in parallel at day 2 (2d) and day 14 (14d) posttransduction ($P = 0.0076$ [significant pairwise comparison by a paired t test]). (C) Changes in clone sizes. Pairwise comparisons of %GFP⁺ values between the clones that were reduced (≥10-fold reduction in the relative abundance) and those that were not reduced (<1.5-fold reduction in the relative abundance) after vpr addition (significant pairwise comparison by a Mann-Whitney U two-tailed test). (D) Changes in clone sizes are not affected by viral burst sizes. Shown are pairwise comparisons of virus release per GFP⁺ cell between the clones that were reduced (≥10-fold reduction in the relative abundance) and those that were not reduced (<1.5-fold reduction in the relative abundance) after vpr addition (significant pairwise comparison by a Mann-Whitney U two-tailed test). (E) Bar graphs showing the percentages of GFP⁺ cells 24 h after reactivation with JQ1, prostratin, and dual treatment with prostratin-JQ1 for the vpr^rev pool (left) and the vpr⁻ pool (right). (F) Heat map of the clonal (zip code) virus release per treated GFP⁻ cell of the vpr^rev pool. Numbers at the left of each panel are clone identifiers generated by ordering proviral zip codes in decreasing relative abundances, as determined for the unsorted pools. Every row represents a unique cell clone's response. The clones were ordered from top to bottom by diminishing virus release per treated cell upon dual-LRA treatment. The color bar indicates the extent of release per treated cell in arbitrary units. Cell density and other culture and assay conditions are the same as the ones described in the legend of Fig. 5.

groups based on the observed fold changes in relative abundance: those that were reduced by 10-fold or more in vpr^{rev} (reduced clones) and those with no observed reduction in their relative abundance upon vpr addition (not reduced). A comparison between the "reduced" and "not-reduced" groups revealed that nearly all vpr^- pool clones that displayed high %GFP$^+$ values, that is, clones in which most member cells displayed LTR activity, were reduced in the vpr^{rev} pool (Fig. 6C).

In contrast, when the burst sizes of individual clones present in the vpr^{rev} population, determined as described above for the parental vpr^- pool (Fig. 6D), were compared to those of the clones that were reduced versus those that were not reduced upon complementation, no differences in the amounts of virus released per active cell were detectable between groups.

The vpr^{rev} pool was then used to address whether vpr^- pools rendered vpr^+ by complementation displayed reactivation patterns similar to those of the original vpr^+ pools. Transduced GFP$^-$ vpr^{rev} subpools were subjected to prostratin, JQ1, or dual treatment and analyzed by flow cytometry and virus release (Fig. 6E). In general, the magnitude of reactivation as measured by flow cytometry did not differ significantly compared to the no-drug control in the transduced cells (Fig. 6E, left). However, despite the significant increase in virus release for dually treated samples (Fig. 6E, right), the magnitude was diminished relative to that for the parental vpr^- GFP$^-$ cells, with cells responding less to JQ1 treatment than they did before Vpr addition. Consistent with expectations if the phenotypes described above reflected Vpr, transduced cells showed enhanced reactivation upon dual-LRA treatment (Fig. 6E; Fig. S4C). Notably, clones (such as clones 53, 133, 161, 293, and 420, etc., in Fig. 6F) that were highly responsive to JQ1 treatment in the parental untransduced vpr^- pool and that differentiated the LRA responsiveness of the vpr^- pool from that of the vpr^+ pool were highly represented among those lost upon vpr addition.

Genomic features of persistent intact proviruses in patients on ART are more similar to those of vpr^+ than to those of vpr^- pool members. Although patients' vpr genes are not routinely sequenced, it seems likely that most proviruses in HIV-1 patients on ART contain intact Vpr and have survived its selective pressures. Having observed differences between vpr^+ and vpr^- viruses in cultured cells in terms of their reactivation patterns and proximity to genome features, we sought to determine which class of these proviruses more closely resembled those in persistent clinical isolates.

To do this, previously published data on patients' intact provirus integration sites (68) were analyzed for their proximity to H3K4me3, H3K27ac, and DNase I hypersensitivity sites that have been reported for Jurkat T cells and primary CD4$^+$ T cells (see Materials and Methods) (63). When these proximities were compared to those in the vpr^- and vpr^+ pools established here, prominent proviruses in the vpr^- pool, and not those in the vpr^+ pool, were found to significantly differ from the patient proviruses in their proximities to H3K4me3 and H3K27ac marks (Fig. 7A and B; Fig. S5) but not DNase I hypersensitivity sites (Fig. 7C). These results suggested that the integration site distribution of persistent proviruses in patients across the tested genome marks was similar to those observed in *in vitro*-established vpr^+ proviral populations but significantly different from what was observed within vpr^- proviral pools.

DISCUSSION

Here, we describe how Vpr disrupts polyclonal provirus population structures and alters the expression properties and latency-reversing agent responsiveness of residual proviral populations in cultured cells. The results illustrate the functional significance of HIV-1's bimodal expression phenotypes in shaping proviral populations and show that whereas Vpr's cytotoxicity will lead to the rapid depletion of clones constitutively or frequently expressing HIV-1 genes, vpr^+ proviruses that are capable of supporting brief robust bursts of virion production can readily persist in proliferating infected cell populations *in vitro*.

Cell-based models are critical for HIV-1 persistence work, but they differ by cell type, the form of HIV-1 used, whether they use clones or polyclonal populations, and other parameters. Many experimental latency systems include reporter genes and/or
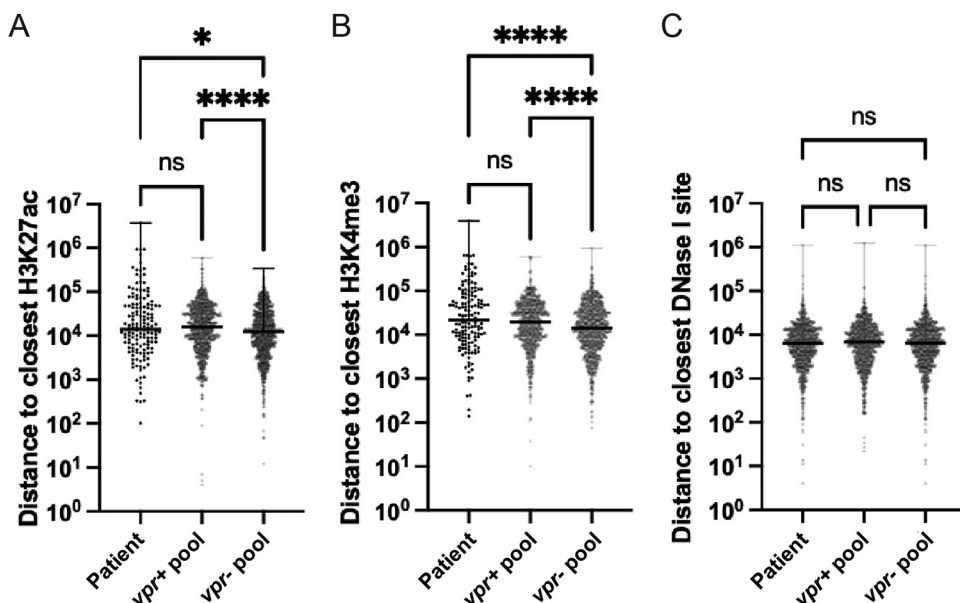
FIG 7 Comparison of patients' integration site features with vpr^+ and vpr^- pools. Previously published HIV integration sites from three patients were used to determine base distances to the closest H3K27ac, H3K4me3, and DNase I sensitivity sites. Distances were compared to those of the vpr^+ and vpr^- pools. Box plots compare the base distances of vpr^+, vpr^-, and patients' proviral integration sites to H3K27ac (means = 28,485, 22,228, and 89,771, respectively) (A), H3K4me3 (means = 32,989, 27,047, and 97,863, respectively) (B), and DNase I (means = 10,687, 10,236, and 10,191, respectively) (C) sensitivity sites (ns, *, **, ***, and **** indicate P values of >0.05, <0.05, <0.01, 0.001, and 0.0001, respectively, by a Kruskal-Wallis test).

delete viral genes believed to be unnecessary for silencing and reactivation (36, 37, 69). For example, two prominent studies that used barcoded proviruses to track virus dissemination in animals used vpr^- proviruses (36, 69). To address the consequences of such exclusions, the current study compared proviral populations with and without *vpr*. Note, however, that all vectors here also lacked *env* and *nef*, which are cytotoxic under some conditions and affect cell physiology in many ways, and thus, these omissions may have biased outcomes (70).

We have previously studied interclonal variation in HIV-1 gene expression using barcoded proviruses (53). That work revealed a broad range of expression variation among integrant clones. When virions produced by barcoded integrants were pseudotyped and polyclonal progeny were generated, the expression properties associated with specific barcodes in the first generation were lost, demonstrating that expression variation was nondeterministic and may be influenced by position effects. If similar variation exists *in vivo*, proviruses during suppressive ART may exist in epigenetic quasispecies upon which selection can act (71). Vpr was used as a source of selective pressure on infected cell populations here.

In our previous study, we compared the expression patterns of barcoded vpr^- proviruses in primary CD4$^+$ T cells to those in Jurkat T cells and observed indistinguishable spectra of expression patterns (53). Although primary CD4$^+$ T cells are a system closer to natural infection, these cells survive only 2 to 4 weeks *ex vivo* unless they are treated with antiapoptotic agents (48, 72) and do not proliferate unless they are stimulated in ways that can affect proviral expression patterns. Thus, because extended cell passaging was required and retaining expression patterns associated with initial integration events was desirable, we chose to work with Jurkat cells here, which limits the physiological relevance of our findings. Note, however, that Vpr reportedly exerts similar effects on primary CD4$^+$ and Jurkat T cells (73).

Our and others' works have shown that individual proviral clones can contain mixtures of LTR-active and -inactive cells (53, 74–76). Furthermore, the clonal progeny of individual infected cells can shift from being silenced to expressing their proviruses while

maintaining overall proportions of cells with active proviruses (53, 74, 75). For example, in polyclonal vpr^- populations, individual member cells within each clone either express HIV-1 genes, as monitored by a GFP reporter, p24 staining, and virion release, or do not. Although some diminution of LTR expression is observed over time, for the most part, each clone adopts a stable, heritable pattern of bimodal gene expression (53).

Previous reports suggest that HIV-1's bifurcating expression profiles exist in a minority of clones (54) and that stochastic fluctuations in gene expression enable probabilistic LTR-on/off fate decisions that are initially unstable but become stabilized by posttranscriptional feedback mechanisms (77). Our observations of bimodal HIV-1 expression patterns likely describe a similar phenomenon, although the majority of our clones display mixed expression phenotypes, and these phenotypes interchange over time in replicating cell populations, suggesting that they may exist in an oscillating circuit (53, 54, 78).

The zip-coded provirus approach used here enabled comparisons of vpr^+ and vpr^- provirus-containing cell pools. Not surprisingly, vpr^+ pools contained significantly fewer GFP$^+$ cells (indicative of transcriptionally active proviruses) than did vpr^- pools. Dominant vpr^+ clones included very few member cells that expressed HIV-1 genes (i.e., they were low-LTR-active clones), whereas high-LTR-active clones dominated the vpr^- pool.

These findings may help explain the inconsistencies in previous estimates of the fractions of HIV-1-infected cells that are transcriptionally active. Using polyclonal vpr^+ proviruses in vitro, Dahabieh et al. suggested that most integrated proviruses are silent (79), even though it has been estimated that only a small fraction of infected patient cells are latent (32, 80). Consistent with the results of Dahabieh et al., we found that HIV-1 is transcriptionally inactive in most cells that persisted in vpr^+ populations, with this reflecting the selective survival and relative amplification of low-LTR-active clones. However, the results showed that low-LTR-active clones with large burst sizes were not appreciably depleted. This finding is striking considering that GFP$^+$ cells, cells that, in a binary sense, would score LTR active, displayed interclonal variation in burst size that spanned 4 orders of magnitude.

These observations support the following model (Fig. 8). The initial integration site distribution will be Vpr agnostic, as previously demonstrated (55). As infected cells proliferate to form clones (13–15), some daughter cells will be LTR active, while other cells will be LTR inactive (53, 54, 78). Each clone will maintain a clone-specific equilibrium population of LTR-active versus -inactive cells over time, which is determined at least in part by integration site features. However, within clones, individual member cells' phenotypes are transitory, with cells alternating between LTR-on and -off states (53, 54, 78). For high-LTR-active clones, most cells will be LTR active and subject to Vpr's cytopathic effects. In contrast, for low-LTR-active clones, even clones that display very large burst sizes when cells are LTR active, clonal proliferation will proceed largely unimpeded.

The vpr^+ and vpr^- proviral populations and their constituent clones here were compared for latency reactivation agent responsiveness, using LRAs selected based on their reported ability to induce HIV-1 to levels comparable to maximum reactivation by T cell activation (81). The overall pool trends for the vpr^+ and vpr^- proviral populations were similar, although the responses differed in magnitude. In contrast, zip code analysis revealed that the behaviors of individual clones within the vpr^+ and vpr^- populations differed. What caused the observed reduced levels of reactivation in the vpr^+ pools was not determined. The widespread transcriptomic changes that Vpr induces in CD4 T cells may have played a role (33), and surviving LTR-inactive clones may have included defective proviruses. Another possibility is that the selective pressure exerted by Vpr may have depleted clones prone to high levels of reactivation. Consistent with this notion, proviruses farther from genome marks associated with active gene expression, a feature associated with vpr^+ clones in the present study, are reportedly relatively resistant to reactivation (82). The adoption by the vpr^{rev} pools of the vpr^+ pool's patterns of LRA responsiveness further corroborated the impact of Vpr on reactivation properties.

A comparison of integrants that dominated passaged vpr^+ and vpr^- pools suggested that Vpr selects against proviruses proximal to genome marks associated with active gene
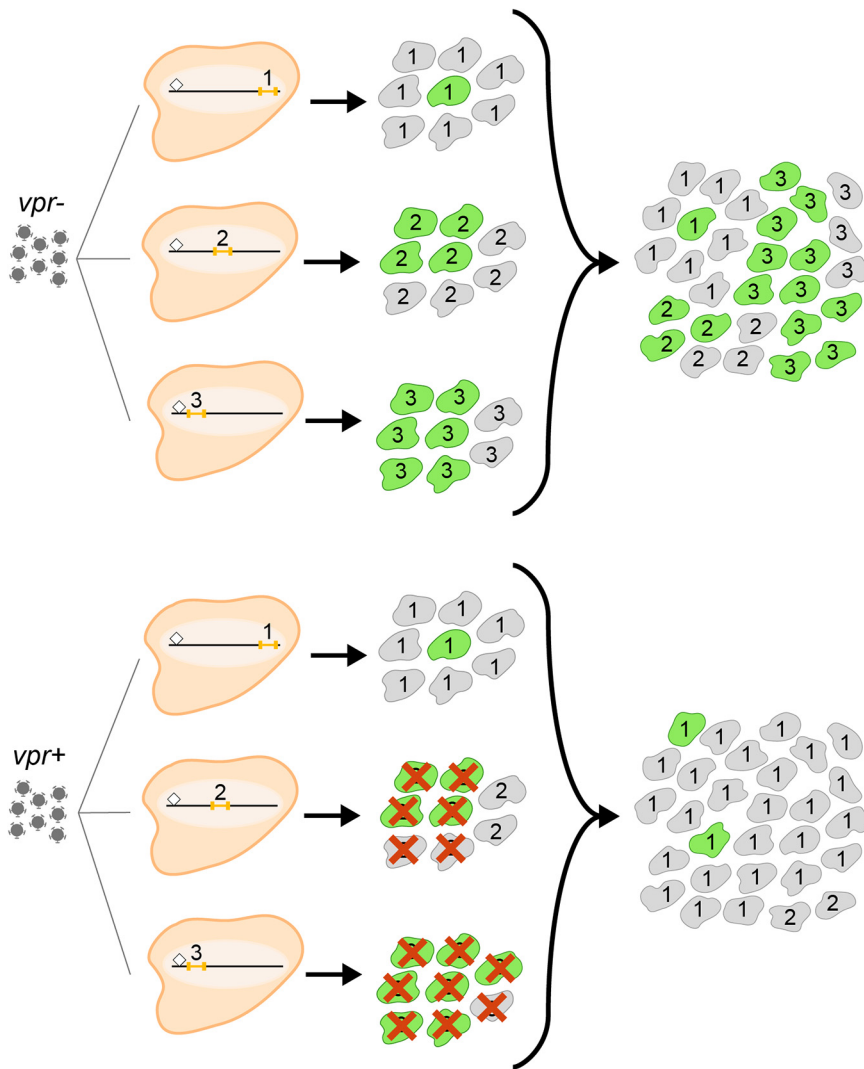
FIG 8 Schematic model of the effect of Vpr on the integrant populations' proviral landscapes. From the left, proviruses integrate at indistinguishable genomic positions (designated 1, 2, and 3 above the small orange bars, which indicate proviruses) regardless of whether the infecting virus was Vpr$^+$ or Vpr$^-$. Note that integrants 1, 2, and 3 differ in their distances to the closest chromatin mark, which is indicated by a small diamond, which may affect their expression characteristics and responses to LRAs. Next, infected cells divide to form cell clones. For clone 1, cells that express HIV genes (indicated by green cells) are rare (low-LTR-active clone), and for other clones, more cellular members of each clone express HIV (e.g., clone 3 [high-LTR-active clone]). This pattern of bimodal gene expression is an intrinsic property of each clone (53). The red X's indicate that high-LTR-active clone members are selectively depleted by Vpr. As a result, in the polyclonal populations shown at the right, low-LTR-active clones are enriched in the vpr$^+$ pool.

expression. To address the relationship of these observations to persistent populations *in vivo*, we mapped the distances to preexisting marks for previously published integration sites of intact, inferred to be intact, or large proviruses in three patients on antiretroviral therapy (ART) (68). This analysis indicated that the proximities observed in patients were significantly more similar to those of vpr$^+$ proviruses than to those of vpr$^-$ pool members.

While proviral expression and latency have traditionally been viewed as mutually exclusive states, recent evidence, such as the selective depletion of patients' intact proviruses over time and detectable HIV-1 RNAs in unstimulated latent cells, suggests that both basal expression and proviral inducibility may be heterogeneous, even within individual cell types (83, 84). Based on their study of replication-competent noninduced proviruses, Ho and colleagues noted that the expression of certain proviruses

was infrequent despite T cell activation, suggesting that the induction of latent proviruses might be a stochastic process (19). Stochastic modeling coupled with laboratory experimentation in a separate study provided evidence that the transition from HIV-1 latency to viral outgrowth is a stochastic property (22). These examples of unanticipated expression patterns *ex vivo*, and their conclusions of HIV-1 expression stochasticity, were based in part on a failure to establish causal links between conditions believed to promote activation and its occurrence.

A possibly related but little-discussed problem observed with transcriptionally silenced *in vitro* latency models is that they tend to "spontaneously revert" during passage and generate small proportions of cells with transcriptionally active proviruses (75, 85). Our previous work and the findings here suggest an explanation slightly different from "spontaneous reversion." Previous quantification of phenotypic switching within expression-sorted polyclonal populations demonstrated that the phenotypes of active and inactive subpopulations of proviral clones are not static (53). Instead, these phenotypes equilibrate such that most LTR-inactive cells are not stably silenced but rather represent the transient off-state of an on-off continuum. Thus, it seems conceivable that spontaneous reversion by a small subset of latency model cells, like the low proportion of GFP$^+$ cells in the current study's *vpr*$^+$ pool, might more accurately be viewed as the programmed equilibration of a low-LTR-active clone than the outlier behavior of a subset of silenced cells.

Burst size variation of a magnitude similar to that described here is observed among patient samples reactivated *ex vivo* (23), and it seems likely that at least vestiges of bimodal expression are retained during natural infection. Weinberger and colleagues have proposed that HIV-1's bifurcating expression phenotypes may represent a bet-hedging strategy, wherein some daughter cells support robust virus production while others remain quiescent, thus allowing clonal persistence whether or not the environment is conducive for viral expression (54). The selective pressures exerted by Vpr that were revealed in the present work may constitute one such determinant that specifies which proviruses will survive and contribute to persistent HIV-1 populations.

MATERIALS AND METHODS

Tissue culture and cell lines. Two immortalized mammalian cell lines were used in this study. Human embryonic kidney HEK 293T and Jurkat cells were purchased from the ATCC (ATCC CRL3216 and ATCC TIB-152, respectively). Both cell lines were preserved as frozen stocks. HEK 293T cells were cultured in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS) (Gemini) and 125 μM gentamicin. Jurkat cells were cultured in RPMI 1640 medium supplemented with 10% FBS, 100 U/mL penicillin, 100 μg/mL streptomycin, 2 mM glutamine, and 55 μM β-mercaptoethanol. To propagate cells, frozen 1-mL aliquots of each cell line were thawed rapidly in a 37°C water bath and added to 9 mL of their respective prewarmed media. After mixing gently, cells were centrifuged at 2,500 rpm for 5 min. The supernatant was discarded, and the cell pellet was resuspended gently in prewarmed medium and plated in an appropriately sized tissue culture plate. HEK 293T cells were subcultured when confluence was between 75 and 100%. Jurkat cells were passaged 1/10 when the cell concentration reached 1 $\times$ 10^6 cells/mL.

Zip-coded vector and virus production. Zip-coded vector DNA templates were generated by the digestion of previously described modified "inside-out" forms of GKO, which are called GPV$^+$ and GPV$^-$ here (53, 75), with ClaI and MluI. The resulting 11.4-kb DNA fragment, which was devoid of a plasmid backbone, was gel purified using a QIAquick gel extraction kit (catalogue no. 28706; Qiagen, Germantown, MD). A 304-bp insert fragment was generated by PCR with GPV$^+$ or GPV$^-$ as the template for *vpr*$^+$ and *vpr*$^-$, respectively; Phusion high-fidelity (HF) DNA polymerase (New England BioLabs, Inc., Ipswich, MA); and primers 5'-GACAAGATATCCTTGATCTGNNNNNNNNNNNNNNNNNNNNNGGCCATCGATGTGGATCTACCACACACAAGGC-3' and 5'-CGGTGCCTGATTAATTAAACGCGTGCTCGAGACCTGGAAAAAC-3'. The 11.4-kb and 304-bp degenerate barcode-containing fragments were joined by Gibson assembly using HiFi DNA assembly mix (New England BioLabs) in a molar ratio of 1:5. Three micrograms of the resulting covalently closed circular DNA was directly cotransfected with 330 ng of pHEF-VSV-G (Addgene plasmid 22501) into a 70% confluent monolayer of HEK 293T cells in a 10-cm dish using polyethylenimine (Polysciences, Inc., Warrington, PA) at 4$\times$ total transfected DNA in 800 μL 150 mM NaCl. DMEM was replaced with 4 mL RPMI 1640 medium with 10% FBS and 1% penicillin-streptomycin (Pen/Strep) at 24 h posttransfection, and culture medium was harvested by filtering through a 0.22-μm filter (catalogue no. 09-720-511; Fisher Scientific) at 48 h posttransfection.

Generation of zip-coded proviral pools. A total of 1,000 μL of virus-containing medium was mixed with Polybrene at a final concentration of 0.5 μg/mL and brought to a total volume of 2,000 μL by the addition of RPMI 1640. This infection mixture was added to 5 $\times$ 10^6 Jurkat cells and incubated in two wells of a 12-well plate at 37°C with 5% CO$_2$ for 5 h. Infected cells were then transferred to 10-mL Falcon tubes and centrifuged for 5 min at 2,500 rpm at 4°C. Following centrifugation, the supernatants were replaced with

fresh medium, and cell pellets were resuspended and cultured in two wells of a 6-well plate at 37°C with 5% CO_2. At 24 h postinfection, puromycin was added to a final concentration of 0.5 μg/mL. The infected cells were expanded into 6-cm culture plates without puromycin on day 5. At 10 days postinfection, the culture supernatant was replaced with fresh medium, and the cultures were divided into aliquots to be either frozen, prepared for integration site sequencing, or further expanded for four additional days and sorted into GFP$^+$ and GFP$^-$ subpools for subsequent experiments. Puromycin-resistant colony-forming titers of pseudotyped barcoded virus stocks were determined by infection of HEK 293T cells, and equivalent infectious units were used to generate vpr^+ and vpr^- Jurkat cell pools.

Isolation of clonal populations. Single cells were isolated from the above-described infected Jurkat cell libraries by limiting dilution and expanded to generate clonal populations. The purity of clones was verified by zip code. Briefly, a 555-bp U3-R PCR product was amplified from cellular DNA extracted from each clonal population using the primers 5′-ACGAAGACAAGATATCCTTGATC-3′ and 5′-GCACTCAAGGCAAGCTT-3′, which flank the zip-coded region. PCRs used Q5 Hot Start high-fidelity 2× master mix (New England BioLabs) according to the manufacturer's protocol, for 30 cycles with a 1-min extension step at 72°C and a 60°C annealing temperature. PCR amplicons were gel purified using a QIAquick gel extraction kit (catalogue no. 28706; Qiagen, Germantown, MD) and submitted to Genewiz (South Plainfield, NJ) for Sanger sequencing with 5′-GCACTCAAGGCAAGCTT-3′. The results revealed that clones 1, 2, and 3 (Fig. 1B) contained the zip codes AATACAAGTCGGACCACCTG, GTAACCTTGGCGTCAGGAG, and GTGATGGTAGCGACAGCGTG, respectively.

Construction of the vpr^+ lentiviral vector and its use. A 2,933-bp *in vitro* DNA-synthesized *IRES-vpr mKO* fragment was ordered from Genewiz (South Plainfield, NJ) using an internal ribosome entry site (IRES) sequence from pTRIPZ-hDDX5/7 (Addgene plasmid 71307) (86) and *vpr* and *mKO* from GPV$^+$. An HIV-1 lentiviral vector fragment was generated by XbaI and MfeI digestion of pWA18puro (87) to remove its puromycin resistance cassette and was gel purified using a QIAquick gel extraction kit (catalogue no. 28706; Qiagen, Germantown, MD). The resulting DNA fragments were joined by Gibson assembly using HiFi DNA assembly mix (New England BioLabs) according to the manufacturer's protocol to generate plasmid pEA216-1. Ten micrograms of pEA216-1 was then cotransfected with 5 μg of the pCMVΔR8.2 helper plasmid (Addgene plasmid 122263) and 1 μg of pHEF-VSV-G into 70% confluent monolayers of HEK 293T cells in a 10-cm dish using polyethylenimine (Polysciences, Inc., Warrington, PA) at 4× total transfected DNA in 800 μL 150 mM NaCl. DMEM was replaced with 4 mL RPMI 1640 medium with 10% FBS and 1% Pen/Strep at 24 h posttransfection, and culture medium was harvested by filtering through a 0.22-μm filter (catalogue no. 09-720-511; Fisher Scientific) at 48 h posttransfection. Parental vpr^- pools were infected with this filtered medium and sorted for mKO and GFP expression at 48 h posttransduction to generate vpr^{rev} pools. Vpr expression was evaluated by cell death. All mKO$^+$ cells were assumed to contain the Tat-inducible *vpr* vector. Therefore, the survival of mKO$^+$ cells (PE) compared to dually positive cells (see Fig. S4A and B, left panels, in the supplemental material) suggested that in the absence of Tat from a preexisting provirus, uninduced Vpr expression was not pronounced enough to kill at least most of the transduced cells. However, the possibility that low levels of uninduced Vpr had led to the depletion of some singly positive cells cannot be ruled out.

In determining the effect of Vpr addition on parental vpr^- pools, zip code abundances in sequences amplified from genomic DNA of parental vpr^- unsorted cells and from Vpr expression vector-transduced vpr^{rev} mKO$^+$ cells were compared. Two categories of zip codes were defined: reduced zip codes were those that were reduced 10-fold or more after Vpr transduction, whereas those with a change of 1.5-fold or less in their relative abundances were regarded as not being affected by *vpr* addition (not reduced).

Flow cytometry and cell sorting. For fluorescence-activated cell sorter (FACS) analysis by flow cytometry, Jurkat cells were suspended in phosphate-buffered saline (PBS) containing 1% FBS (FACS buffer). Dead cells were excluded from all analyses and sorting experiments using propidium iodide (PI). Acquisition was carried out on the fluorescein isothiocyanate (FITC) channel for GFP and on the PE channel for PI. Cell fluorescence was assessed using a BD LSR Fortessa instrument (BD Biosciences), and data were analyzed using FlowJo software, version 10.6 (FlowJo, LLC, Ashland, OR). Infected cells were sorted into GFP$^+$ and GFP$^-$ subpopulations by flow cytometry using a FACSAria II (BD Biosciences, Franklin Lakes, NJ) or an iCyt Synergy SY3200 (Sony Biotechnology, San Jose, CA) cell sorter at the flow cytometry core of the University of Michigan.

Latency-reversing agents and reactivation. JQ1 and prostratin were purchased from Sigma-Aldrich. Each LRA was dissolved in DMSO (Thermo Fisher) to produce stocks. For each experiment, stocks were added to culture medium to achieve final concentrations of 2 μM JQ1 and 10 μM prostratin. Dual-LRA treatment was performed by adding the two LRAs to the same culture medium to achieve their respective single-LRA concentrations. For reactivation experiments, GFP$^-$ cells sorted on day 14 postinfection were cocultured with the appropriate LRA for 24 h. Cells were then centrifuged at 2,000 rpm at 4°C for 5 min, and cells pellets were washed twice with ice-cold FACS buffer after being stained with PI for 5 min at room temperature. The resulting cells were then washed and assessed by flow cytometry, and the p24 concentration in the culture supernatants was determined by a reverse transcriptase (RT) assay.

Zip code sequencing libraries. Zip codes were amplified from the genomic DNA of infected cells as well as from the RNA of virus released into cell media. The generation of zip-coded sequencing libraries from infected cells was initiated by harvesting DNA from an aliquot of 2 × 10^6 infected cells. Genomic DNA extraction was carried out using the Qiagen (Germantown, MD) DNeasy blood and tissue kit. Zip codes were then amplified by PCR from 200 ng of the DNA template using Phusion high-fidelity DNA polymerase (New England BioLabs) in HF buffer. Primers were designed to flank the zip code region (primer sequences 5′-NNACGAAGACAAGATATCCTTGATC-3′ and 5′-NNTGTGTGGTAGATCCACATCG-3′). Multiple copies of these primers were created, each with a unique pair of known, randomized

nucleotides at the 5′ end, to confirm that no cross-contamination had occurred between samples. Reaction mixtures were cycled 29 times with a 30-s extension step at 72°C and a 59°C annealing temperature. Zip code amplicons were purified using the DNA Clean and Concentrator-5 kit (catalogue no. D4013; Zymo Research, CA) and eluted in 15 μL of MilliQ H_2O. To amplify zip codes from virus, a tissue culture plate of infected cells was decanted into a conical tube and centrifuged at 2,500 rpm for 5 min. The virus-containing medium was separated from the cell pellet and passed through a 0.22-μm filter. To concentrate the virus, medium was subjected to ultracentrifugation (25,000 rpm) for 120 min through a 20% sucrose cushion. Viral pellets were then resuspended in 200 μL PBS, and viral RNA was extracted using the Quick-RNA viral kit (catalogue no. R1034 and R1035; Zymo Research, CA) according to the manufacturer's protocol and eluted in 10 μL RNase-free water. cDNA was synthesized using 5 μL of the eluent as the template, using the U3 antisense primer 5′-TGTGTGGTAGATCCACATCG-3′ and Moloney murine leukemia virus (M-MLV) RT RNase (H$^-$) (catalogue no. MR3681; Promega, WI) according to the manufacturer's protocol. Zip codes were amplified from this cDNA under the conditions described above. The zip code amplicons were then used to generate MiSeq libraries for sequencing as described previously (53).

Integration site determination. Genomic DNA was extracted from vpr^+ and vpr^- cells at 10 days postinfection using the Qiagen DNeasy blood and tissue kit (Qiagen), and 200 ng of DNA was sheared to 1-kb fragments using an M220 instrument and microTUBE according to the manufacturer's recommended settings (Covaris, Woburn, MA). HIV-1 insertion site libraries were prepared and sequenced using methods described previously (53).

Quantification of virus release. Virion production was quantified using a real-time reverse transcription-PCR assay developed previously by Pizzato et al., as modified by Kharytonchyk et al. (87, 88). Briefly, viral lysates were prepared by adding 5 μL of the culture supernatant to 5 μL of lysis buffer. Using MS2 RNA as the template, MS2 cDNA was synthesized with viral lysates and quantified by real-time PCR in one reaction. Released virus was quantified and normalized for p24 levels based on values determined in parallel for reference samples.

Zip code quantification and analysis. Zip codes were identified and quantified from Illumina sequencing reads using a previously described custom suite of tools implemented in Python (https://github.com/KiddLab/hiv-zipcode-tools). Briefly, 2- by 75-bp paired reads were merged using flash v1.2.11 (73). Zip codes were identified by searching for known flanking sequences (with up to 1 mismatch). Only candidate zip codes with a length of 17 to 23 nucleotides were considered, and the read count for each unique zip code was tabulated. To identify sets of zip codes for further analysis, zip code families, which account for PCR and sequencing errors, were determined by clustering together the observed unique zip codes. Abundances for the zip codes were then determined by assigning unique zip codes to the most abundant family whose sequence was within 2 mismatches and summing their associated read counts. Only zip code families with corresponding data in the integration site data were selected for further analysis.

For each latency reversal treatment condition, clonal virus release was determined by multiplying the fractional abundances of zip codes from the cDNA sequencing libraries of each treatment by their corresponding sample's pool p24 concentration, as measured 24 h after LRA treatment. The resulting clonal p24 values for each zip code in prostratin-, JQ1-, and combination prostratin-JQ1-treated samples were divided by clonal p24 values defined for cell samples exposed to only DMSO to determine fold changes after LRA treatment.

The burst sizes of clones were determined by multiplying the fractional abundances of zip codes from the cDNA library of the unstimulated sample's pool p24 concentration (basal viral release). The resulting p24 values for each zip code were then divided by their corresponding fraction of abundances determined in the gDNA sequencing library of GFP$^+$ sorted cells from the same unstimulated sample from which the cDNA library was made.

The %GFP$^+$ value for each zip code was determined as $F_i = [(G_i \times P)/(G_i \times P + W_i \times Q)] \times 100$ (53), where F_i is the GFP$^+$ percentage of zip code i, G_i is the fraction abundance of zip code i in the GFP$^+$ sorted pool, W_i is the fraction abundance of zip code i in the GFP$^-$ sorted pool, P is the fraction of cells that were sorted into the GFP$^+$ pool, and Q is the fraction of cells that were sorted into the GFP$^-$ pool.

Determination of chromatin marks. H3K27ac marks annotated for the Jurkat E-6-1 clone were sourced from ChIP-Atlas (http://chip-atlas.org/view?id=SRX1041803), H3K4me3 and DNase I sensitivity site data sets were downloaded from the ENCODE Project (https://www.encodeproject.org/) with Sequence Read Archive accession no. SRX1041803 and ENCODE Project identifiers ENCFF304GVP (https://www.encodeproject.org/experiments/ENCSR000EOS/) for genome marks preexisting in Jurkat cells, iENCFF341XUX (https://www.encodeproject.org/experiments/ENCSR807WEO/), and ENCFF053LHH (https://www.encodeproject.org/experiments/ENCSR724GUS/) (63). Bedtools was then used to map the distance to the closest known annotated marks. Analysis of the proximity to these genome marks was done using the matplotlib and scipy.stats packages in Python, and the results were exported into GraphPad Prism version 9.1.2 to plot graphs.

Data availability. All sequence data have been deposited to the SRA under accession no. SRX1041803.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 0.4 MB.

FIG S2, TIF file, 0.8 MB.

FIG S3, TIF file, 1.3 MB.

FIG S4, TIF file, 1.6 MB.
FIG S5, TIF file, 0.3 MB.
TABLE S1, XLSX file, 0.2 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Bailey J, Blankson JN, Wind-Rotolo M, Siliciano RF. 2004. Mechanisms of HIV-1 escape from immune responses and antiretroviral drugs. Curr Opin Immunol 16:470–476. https://doi.org/10.1016/j.coi.2004.05.005.

2. Lenasi T, Contreras X, Peterlin BM. 2008. Transcriptional interference antagonizes proviral gene expression to promote HIV latency. Cell Host Microbe 4:123–133. https://doi.org/10.1016/j.chom.2008.05.016.

3. Han Y, Lin YB, An W, Xu J, Yang H-C, O'Connell K, Dordai D, Boeke JD, Siliciano JD, Siliciano RF. 2008. Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. Cell Host Microbe 4:134–146. https://doi.org/10.1016/j.chom.2008.06.008.

4. Lassen K, Han Y, Zhou Y, Siliciano J, Siliciano RF. 2004. The multifactorial nature of HIV-1 latency. Trends Mol Med 10:525–531. https://doi.org/10.1016/j.molmed.2004.09.006.

5. Williams SA, Greene WC. 2007. Regulation of HIV-1 latency by T-cell activation. Cytokine 39:63–74. https://doi.org/10.1016/j.cyto.2007.05.017.

6. Kulkosky J, Sullivan J, Xu Y, Souder E, Hamer DH, Pomerantz RJ. 2004. Expression of latent HAART-persistent HIV type 1 induced by novel cellular activating agents. AIDS Res Hum Retroviruses 20:497–505. https://doi.org/10.1089/088922204323087741.

7. Mbonye U, Wang B, Gokulrangan G, Shi W, Yang S, Karn J. 2018. Cyclin-dependent kinase 7 (CDK7)-mediated phosphorylation of the CDK9 activation loop promotes P-TEFb assembly with Tat and proviral HIV reactivation. J Biol Chem 293:10009–10025. https://doi.org/10.1074/jbc.RA117.001347.

8. Archin NM, Sung JM, Garrido C, Soriano-Sarabia N, Margolis DM. 2014. Eradicating HIV-1 infection: seeking to clear a persistent pathogen. Nature reviews. Microbiology 12(11):750–764. https://doi.org/10.1038/nrmicro3352.

9. Verdin E, Paras P, Jr, Van Lint C. 1993. Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. EMBO J 12:3249–3259. https://doi.org/10.1002/j.1460-2075.1993.tb05994.x.

10. Blazkova J, Trejbalova K, Gondois-Rey F, Halfon P, Philibert P, Guiguen A, Verdin E, Olive D, Van Lint C, Hejnar J, Hirsch I. 2009. CpG methylation controls reactivation of HIV from latency. PLoS Pathog 5:e1000554. https://doi.org/10.1371/journal.ppat.1000554.

11. Emiliani S, Fischle W, Ott M, Van Lint C, Amella CA, Verdin E. 1998. Mutations in the tat gene are responsible for human immunodeficiency virus type 1 postintegration latency in the U1 cell line. J Virol 72:1666–1670. https://doi.org/10.1128/JVI.72.2.1666-1670.1998.

12. Emiliani S, Van Lint C, Fischle W, Paras P, Jr, Ott M, Brady J, Verdin E. 1996. A point mutation in the HIV-1 Tat responsive element is associated with postintegration latency. Proc Natl Acad Sci U S A 93:6377–6381. https://doi.org/10.1073/pnas.93.13.6377.

13. Mendoza P, Jackson JR, Oliveira TY, Gaebler C, Ramos V, Caskey M, Jankovic M, Nussenzweig MC, Cohn LB. 2020. Antigen-responsive CD4+ T cell clones contribute to the HIV-1 latent reservoir. J Exp Med 217:e20200051. https://doi.org/10.1084/jem.20200051.

14. Simonetti FR, Zhang H, Soroosh GP, Duan J, Rhodehouse K, Hill AL, Beg SA, McCormick K, Raymond HE, Nobles CL, Everett JK, Kwon KJ, White JA, Lai J, Margolick JB, Hoh R, Deeks SG, Bushman FD, Siliciano JD, Siliciano RF. 2021. Antigen-driven clonal selection shapes the persistence of HIV-1-infected CD4+ T cells in vivo. J Clin Invest 131:e145254. https://doi.org/10.1172/JCI145254.

15. Wang Z, Gurule EE, Brennan TP, Gerold JM, Kwon KJ, Hosmane NN, Kumar MR, Beg SA, Capoferri AA, Ray SC, Ho Y-C, Hill AL, Siliciano JD, Siliciano RF. 2018. Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane. Proc Natl Acad Sci U S A 115:E2575–E2584. https://doi.org/10.1073/pnas.1720665115.

16. Eriksson S, Graf EH, Dahl V, Strain MC, Yukl SA, Lysenko ES, Bosch RJ, Lai J, Chioma S, Emad F, Abdel-Mohsen M, Hoh R, Hecht F, Hunt P, Somsouk M, Wong J, Johnston R, Siliciano RF, Richman DD, O'Doherty U, Palmer S, Deeks SG, Siliciano JD. 2013. Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies. PLoS Pathog 9:e1003174. https://doi.org/10.1371/journal.ppat.1003174.

17. Bruner KM, Wang Z, Simonetti FR, Bender AM, Kwon KJ, Sengupta S, Fray EJ, Beg SA, Antar AAR, Jenike KM, Bertagnolli LN, Capoferri AA, Kufera JT, Timmons A, Nobles C, Gregg J, Wada N, Ho Y-C, Zhang H, Margolick JB, Blankson JN, Deeks SG, Bushman FD, Siliciano JD, Laird GM, Siliciano RF. 2019. A quantitative approach for measuring the reservoir of latent HIV-1 proviruses. Nature 566:120–125. https://doi.org/10.1038/s41586-019-0898-8.

18. Wonderlich ER, Subramanian K, Cox B, Wiegand A, Lackman-Smith C, Bale MJ, Stone M, Hoh R, Kearney MF, Maldarelli F, Deeks SG, Busch MP, Ptak RG, Kulpa DA. 2019. Effector memory differentiation increases detection of replication-competent HIV-l in resting CD4+ T cells from virally suppressed individuals. PLoS Pathog 15:e1008074. https://doi.org/10.1371/journal.ppat.1008074.

19. Ho Y-C, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DI, Lai J, Blankson JN, Siliciano JD, Siliciano RF. 2013. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. Cell 155:540–551. https://doi.org/10.1016/j.cell.2013.09.020.

20. Lu C-L, Pai JA, Nogueira L, Mendoza P, Gruell H, Oliveira TY, Barton J, Lorenzi JCC, Cohen YZ, Cohn LB, Klein F, Caskey M, Nussenzweig MC, Jankovic M. 2018. Relationship between intact HIV-1 proviruses in circulating CD4+ T cells and rebound viruses emerging during treatment interruption. Proc Natl Acad Sci U S A 115:E11341–E11348. https://doi.org/10.1073/pnas.1813512115.

21. Cohen YZ, Lorenzi JCC, Krassnig L, Barton JP, Burke L, Pai J, Lu C-L, Mendoza P, Oliveira TY, Sleckman C, Millard K, Butler AL, Dizon JP, Belblidia SA, Witmer-Pack M, Shimeliovich I, Gulick RM, Seaman MS, Jankovic M, Caskey M, Nussenzweig MC. 2018. Relationship between latent and rebound viruses in a clinical trial of anti-HIV-1 antibody 3BNC117. J Exp Med 215:2311–2324. https://doi.org/10.1084/jem.20180936.

22. Hataye JM, Casazza JP, Best K, Liang CJ, Immonen TT, Ambrozak DR, Darko S, Henry AR, Laboune F, Maldarelli F, Douek DC, Hengartner NW, Yamamoto T, Keele BF, Perelson AS, Koup RA. 2019. Principles governing establishment versus collapse of HIV-1 cellular spread. Cell Host Microbe 26:748–763.e20. https://doi.org/10.1016/j.chom.2019.10.006.

23. Bui JK, Mellors JW, Cillo AR. 2016. HIV-1 virion production from single inducible proviruses following T-cell activation ex vivo. J Virol 90:1673–1676. https://doi.org/10.1128/JVI.02520-15.

24. Darcis G, Kula A, Bouchat S, Fujinaga K, Corazza F, Ait-Ammar A, Delacourt N, Melard A, Kabeya K, Vanhulle C, Van Driessche B, Gatot J-S, Cherrier T, Pianowski LF, Gama L, Schwartz C, Vila J, Burny A, Clumeck N, Moutschen M, De Wit S, Peterlin BM, Rouzioux C, Rohr O, Van Lint C. 2015. An in-depth comparison of latency-reversing agent combinations in various in vitro and ex vivo HIV-1 latency models identified bryostatin-1+ JQ1 and ingenol-B+ JQ1 to potently reactivate viral gene expression. PLoS Pathog 11:e1005063. https://doi.org/10.1371/journal.ppat.1005063.

25. Pace MJ, Agosto L, Graf EH, O'Doherty U. 2011. HIV reservoirs and latency models. Virology 411:344–354. https://doi.org/10.1016/j.virol.2010.12.041.

26. Jordan A, Bisgrove D, Verdin E. 2003. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. EMBO J 22:1868–1877. https://doi.org/10.1093/emboj/cdg188.

27. Eckstein DA, Sherman MP, Penn ML, Chin PS, De Noronha CM, Greene WC, Goldsmith MA. 2001. HIV-1 Vpr enhances viral burden by facilitating infection of tissue macrophages but not nondividing CD4+ T cells. J Exp Med 194:1407–1419. https://doi.org/10.1084/jem.194.10.1407.

28. Kino T, Gragerov A, Slobodskaya O, Tsopanomichalou M, Chrousos GP, Pavlakis GN. 2002. Human immunodeficiency virus type 1 (HIV-1) accessory protein Vpr induces transcription of the HIV-1 and glucocorticoid-responsive promoters by binding directly to p300/CBP coactivators. J Virol 76:9724–9734. https://doi.org/10.1128/jvi.76.19.9724-9734.2002.

29. Sato K, Misawa N, Iwami S, Satou Y, Matsuoka M, Ishizaka Y, Ito M, Aihara K, An DS, Koyanagi Y. 2013. HIV-1 Vpr accelerates viral replication during acute infection by exploitation of proliferating CD4+ T cells in vivo. PLoS Pathog 9:e1003812. https://doi.org/10.1371/journal.ppat.1003812.

30. Zhang Q, Kang Y, Wang S, Gonzalez GM, Li W, Hui H, Wang Y, Rana TM. 2021. HIV reprograms host m6Am RNA methylome by viral Vpr protein-mediated degradation of PCIF1. Nat Commun 12:5543. https://doi.org/10.1038/s41467-021-25683-4.

31. Zhang F, Bieniasz PD. 2020. HIV-1 Vpr induces cell cycle arrest and enhances viral gene expression by depleting CCDC137. Elife 9:e55806. https://doi.org/10.7554/eLife.55806.

32. Yao X-J, Mouland AJ, Subbramanian RA, Forget J, Rougeau N, Bergeron D, Cohen EA. 1998. Vpr stimulates viral expression and induces cell killing in human immunodeficiency virus type 1-infected dividing Jurkat T cells. J Virol 72:4686–4693. https://doi.org/10.1128/JVI.72.6.4686-4693.1998.

33. Bauby H, Ward CC, Hugh-White R, Swanson CM, Schulz R, Goujon C, Malim MH. 2021. HIV-1 Vpr induces widespread transcriptomic changes in CD4+ T cells early postinfection. mBio 12:e01369-21. https://doi.org/10.1128/mBio.01369-21.

34. Chen H-C, Martinez JP, Zorita E, Meyerhans A, Filion GJ. 2017. Position effects influence HIV latency reversal. Nat Struct Mol Biol 24:47–54. https://doi.org/10.1038/nsmb.3328.

35. 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. https://doi.org/10.1093/emboj/20.7.1726.

36. Fennessey CM, Pinkevych M, Immonen TT, Reynaldi A, Venturi V, Nadella P, Reid C, Newman L, Lipkey L, Oswald K, Bosche WJ, Trivett MT, Ohlen C, Ott DE, Estes JD, Del Prete GQ, Lifson JD, Davenport MP, Keele BF. 2017. Genetically-barcoded SIV facilitates enumeration of rebound variants and estimation of reactivation rates in nonhuman primates following interruption of suppressive antiretroviral therapy. PLoS Pathog 13:e1006359. https://doi.org/10.1371/journal.ppat.1006359.

37. Jefferys SR, Burgos SD, Peterson JJ, Selitsky SR, Turner A-MW, James LI, Tsai Y-H, Coffey AR, Margolis DM, Parker J, Browne EP. 2021. Epigenomic characterization of latent HIV infection identifies latency regulating transcription factors. PLoS Pathog 17:e1009346. https://doi.org/10.1371/journal.ppat.1009346.

38. Deeks SG. 2012. HIV: shock and kill. Nature 487:439–440. https://doi.org/10.1038/487439a.

39. Korin YD, Brooks DG, Brown S, Korotzer A, Zack JA. 2002. Effects of prostratin on T-cell activation and human immunodeficiency virus latency. J Virol 76:8118–8123. https://doi.org/10.1128/jvi.76.16.8118-8123.2002.

40. Williams SA, Chen L-F, Kwon H, Fenard D, Bisgrove D, Verdin E, Greene WC. 2004. Prostratin antagonizes HIV latency by activating NF-κB. J Biol Chem 279:42008–42017. https://doi.org/10.1074/jbc.M402124200.

41. Boehm D, Calvanese V, Dar RD, Xing S, Schroeder S, Martins L, Aull K, Li P-C, Planelles V, Bradner JE, Zhou M-M, Siliciano RF, Weinberger L, Verdin E, Ott M. 2013. BET bromodomain-targeting compounds reactivate HIV from latency via a Tat-independent mechanism. Cell Cycle 12:452–462. https://doi.org/10.4161/cc.23309.

42. Zhu J, Gaiha GD, John SP, Pertel T, Chin CR, Gao G, Qu H, Walker BD, Elledge SJ, Brass AL. 2012. Reactivation of latent HIV-1 by inhibition of BRD4. Cell Rep 2:807–816. https://doi.org/10.1016/j.celrep.2012.09.008.

43. Contreras X, Schweneker M, Chen C-S, McCune JM, Deeks SG, Martin J, Peterlin BM. 2009. Suberoylanilide hydroxamic acid reactivates HIV from latently infected cells. J Biol Chem 284:6782–6789. https://doi.org/10.1074/jbc.M807898200.

44. Edelstein LC, Micheva-Viteva S, Phelan BD, Dougherty JP. 2009. Activation of latent HIV type 1 gene expression by suberoylanilide hydroxamic acid (SAHA), an HDAC inhibitor approved for use to treat cutaneous T cell lymphoma. AIDS Res Hum Retroviruses 25:883–887. https://doi.org/10.1089/aid.2008.0294.

45. Jiang G, Mendes EA, Kaiser P, Sankaran-Walters S, Tang Y, Weber MG, Melcher GP, Thompson GR, III, Tanuri A, Pianowski LF, Wong JK, Dandekar S. 2014. Reactivation of HIV latency by a newly modified ingenol derivative via protein kinase Cδ-NF-κB signaling. AIDS 28:1555–1566. https://doi.org/10.1097/QAD.0000000000000289.

46. 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 Res Hum Retroviruses 25:207–212. https://doi.org/10.1089/aid.2008.0191.

47. Li C, Mori L, Valente ST. 2021. The block-and-lock strategy for human immunodeficiency virus cure: lessons learned from didehydro-cortistatin A. J Infect Dis 223:S46–S53. https://doi.org/10.1093/infdis/jiaa681.

48. Xing S, Bullen CK, Shroff NS, Shan L, Yang H-C, Manucci JL, Bhat S, Zhang H, Margolick JB, Quinn TC, Margolis DM, Siliciano JD, Siliciano RF. 2011. Disulfiram reactivates latent HIV-1 in a Bcl-2-transduced primary CD4+ T cell model without inducing global T cell activation. J Virol 85:6060–6064. https://doi.org/10.1128/JVI.02033-10.

49. Yang H-C, Xing S, Shan L, O'Connell K, Dinoso J, Shen A, Zhou Y, Shrum CK, Han Y, Liu JO, Zhang H, Margolick JB, Siliciano RF. 2009. Small-molecule screening using a human primary cell model of HIV latency identifies compounds that reverse latency without cellular activation. J Clin Invest 119:3473–3486. https://doi.org/10.1172/JCI39199.

50. 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. https://doi.org/10.1038/nature11286.

51. Spivak AM, Andrade A, Eisele E, Hoh R, Bacchetti P, Bumpus NN, Emad F, Buckheit R, III, McCance-Katz EF, Lai J, Kennedy M, Chander G, Siliciano RF, Siliciano JD, Deeks SG. 2014. A pilot study assessing the safety and latency-reversing activity of disulfiram in HIV-1-infected adults on antiretroviral therapy. Clin Infect Dis 58:883–890. https://doi.org/10.1093/cid/cit813.

52. Spina CA, Anderson J, Archin NM, Bosque A, Chan J, Famiglietti M, Greene WC, Kashuba A, Lewin SR, Margolis DM, Mau M, Ruelas D, Saleh S, Shirakawa K, Siliciano RF, Singhania A, Soto PC, Terry VH, Verdin E, Woelk C, Wooden S, Xing S, Planelles V. 2013. An in-depth comparison of latent HIV-1 reactivation in multiple cell model systems and resting CD4+ T cells from aviremic patients. PLoS Pathog 9:e1003834. https://doi.org/10.1371/journal.ppat.1003834.

53. Read DF, Atindaana E, Pyaram K, Yang F, Emery S, Cheong A, Nakama KR, Burnett C, Larragoite ET, Battivelli E, Verdin E, Planelles V, Chang C-H, Telesnitsky A, Kidd JM. 2019. Stable integrant-specific differences in bimodal HIV-1 expression patterns revealed by high-throughput analysis. PLoS Pathog 15:e1007903. https://doi.org/10.1371/journal.ppat.1007903.

54. Weinberger LS, Burnett JC, Toettcher JE, Arkin AP, Schaffer DV. 2005. Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. Cell 122:169–182. https://doi.org/10.1016/j.cell.2005.06.006.

55. Miller CM, Akiyama H, Agosto LM, Emery A, Ettinger CR, Swanstrom RI, Henderson AJ, Gummuluru S. 2017. Virion-associated Vpr alleviates a postintegration block to HIV-1 infection of dendritic cells. J Virol 91:e00051-17. https://doi.org/10.1128/JVI.00051-17.

56. Ohagen A, Ghosh S, He J, Huang K, Chen Y, Yuan M, Osathanondh R, Gartner S, Shi B, Shaw G, Gabuzda D. 1999. Apoptosis induced by infection of primary brain cultures with diverse human immunodeficiency virus type 1 isolates: evidence for a role of the envelope. J Virol 73:897–906. https://doi.org/10.1128/JVI.73.2.897-906.1999.

57. Bartz SR, Emerman M. 1999. Human immunodeficiency virus type 1 Tat induces apoptosis and increases sensitivity to apoptotic signals by up-regulating FLICE/caspase-8. J Virol 73:1956–1963. https://doi.org/10.1128/JVI.73.3.1956-1963.1999.

58. Xu X-N, Laffert B, Screaton GR, Kraft M, Wolf D, Kolanus W, Mongkolsapay J, McMichael AJ, Baur AS. 1999. Induction of Fas ligand expression by HIV involves the interaction of Nef with the T cell receptor ζ chain. J Exp Med 189:1489–1496. https://doi.org/10.1084/jem.189.9.1489.

59. ENCODE. 2012. Experiment summary for ENCSR000BXX. https://www.encodeproject.org/experiments/ENCSR000BXX/.

60. Sherrill-Mix S, Lewinski MK, Famiglietti M, Bosque A, Malani N, Ocwieja KE, Berry CC, Looney D, Shan L, Agosto LM, Pace MJ, Siliciano RF, O'Doherty U, Guatelli J, Planelles V, Bushman FD. 2013. HIV latency and integration site placement in five cell-based models. Retrovirology 10:90. https://doi.org/10.1186/1742-4690-10-90.

61. Poon B, Chen IS. 2003. Human immunodeficiency virus type 1 (HIV-1) Vpr enhances expression from unintegrated HIV-1 DNA. J Virol 77:3962–3972. https://doi.org/10.1128/jvi.77.7.3962-3972.2003.

62. Vansant G, Chen H-C, Zorita E, Trejbalová K, Miklík D, Filion G, Debyser Z. 2020. The chromatin landscape at the HIV-1 provirus integration site determines viral expression. Nucleic Acids Res 48:7801–7817. https://doi.org/10.1093/nar/gkaa536.

63. ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 489:57–74. https://doi.org/10.1038/nature11247.

64. Bruce JW, Bracken M, Evans E, Sherer N, Ahlquist P. 2021. ZBTB2 represses HIV-1 transcription and is regulated by HIV-1 Vpr and cellular DNA damage responses. PLoS Pathog 17:e1009364. https://doi.org/10.1371/journal.ppat.1009364.

65. Le Douce V, Herbein G, Rohr O, Schwartz C. 2010. Molecular mechanisms of HIV-1 persistence in the monocyte-macrophage lineage. Retrovirology 7:32. https://doi.org/10.1186/1742-4690-7-32.

66. Cherrier T, Le Douce V, Eilebrecht S, Riclet R, Marban C, Dequiedt F, Goumon Y, Paillart J-C, Mericskay M, Parlakian A, Bausero P, Abbas W, Herbein G, Kurdistani SK, Grana X, Van Driessche B, Schwartz C, Candolfi E, Benecke AG, Van Lint C, Rohr O. 2013. CTIP2 is a negative regulator of P-TEFb. Proc Natl Acad Sci U S A 110:12655–12660. https://doi.org/10.1073/pnas.1220136110.

67. Romani B, Kamali Jamil R, Hamidi-Fard M, Rahimi P, Momen SB, Aghasadeghi MR, Allahbakhshi E. 2016. HIV-1 Vpr reactivates latent HIV-1 provirus by inducing depletion of class I HDACs on chromatin. Sci Rep 6: 31924. https://doi.org/10.1038/srep31924.

68. Einkauf KB, Lee GQ, Gao C, Sharaf R, Sun X, Hua S, Chen SM, Jiang C, Lian X, Chowdhury FZ, Rosenberg ES, Chun T-W, Li JZ, Yu XG, Lichterfeld M. 2019. Intact HIV-1 proviruses accumulate at distinct chromosomal positions during prolonged antiretroviral therapy. J Clin Invest 129:988–998. https://doi.org/10.1172/JCI124291.

69. Marsden MD, Zhang T-H, Du Y, Dimapasoc M, Soliman MSA, Wu X, Kim JT, Shimizu A, Schrier A, Wender PA, Sun R, Zack JA. 2020. Tracking HIV rebound following latency reversal using barcoded HIV. Cell Rep Med 1: 100162. https://doi.org/10.1016/j.xcrm.2020.100162.

70. Fujii Y, Otake K, Tashiro M, Adachi A. 1996. Soluble Nef antigen of HIV-1 is cytotoxic for human CD4+ T cells. FEBS Lett 393:93–96. https://doi.org/10.1016/0014-5793(96)00859-9.

71. Domingo E, Sheldon J, Perales C. 2012. Viral quasispecies evolution. Microbiol Mol Biol Rev 76:159–216. https://doi.org/10.1128/MMBR.05023-11.

72. Kim M, Hosmane NN, Bullen CK, Capoferri A, Yang H-C, Siliciano JD, Siliciano RF. 2014. A primary CD4+ T cell model of HIV-1 latency established after activation through the T cell receptor and subsequent return to quiescence. Nat Protoc 9:2755–2770. https://doi.org/10.1038/nprot.2014.188.

73. Gummuluru S, Emerman M. 1999. Cell cycle- and Vpr-mediated regulation of human immunodeficiency virus type 1 expression in primary and transformed T-cell lines. J Virol 73:5422–5430. https://doi.org/10.1128/JVI.73.7.5422-5430.1999.

74. Weinberger AD, Weinberger LS. 2013. Stochastic fate selection in HIV-infected patients. Cell 155:497–499. https://doi.org/10.1016/j.cell.2013.09.039.

75. Battivelli E, Dahabieh MS, Abdel-Mohsen M, Svensson JP, Da Silva IT, Cohn LB, Gramatica A, Deeks S, Greene WC, Pillai SK, Verdin E. 2018. Distinct chromatin functional states correlate with HIV latency reactivation in infected primary CD4+ T cells. Elife 7:e34655. https://doi.org/10.7554/eLife.34655.

76. Telwatte S, Morón-López S, Aran D, Kim P, Hsieh C, Joshi S, Montano M, Greene WC, Butte AJ, Wong JK, Yukl SA. 2019. Heterogeneity in HIV and cellular transcription profiles in cell line models of latent and productive infection: implications for HIV latency. Retrovirology 16:32. https://doi.org/10.1186/s12977-019-0494-x.

77. Hansen MMK, Wen WY, Ingerman E, Razooky BS, Thompson CE, Dar RD, Chin CW, Simpson ML, Weinberger LS. 2018. A post-transcriptional feedback mechanism for noise suppression and fate stabilization. Cell 173: 1609–1621.15. https://doi.org/10.1016/j.cell.2018.04.005.

78. Rosier BJ, De Greef TF. 2015. Synthetic biology: how to make an oscillator. Elife 4:e12260. https://doi.org/10.7554/eLife.12260.

79. Dahabieh MS, Ooms M, Simon V, Sadowski I. 2013. A doubly fluorescent HIV-1 reporter shows that the majority of integrated HIV-1 is latent shortly after infection. J Virol 87:4716–4727. https://doi.org/10.1128/JVI.03478-12.

80. Stewart SA, Poon B, Jowett J, Chen I. 1997. Human immunodeficiency virus type 1 Vpr induces apoptosis following cell cycle arrest. J Virol 71: 5579–5592. https://doi.org/10.1128/JVI.71.7.5579-5592.1997.

81. Laird GM, Bullen CK, Rosenbloom DI, Martin AR, Hill AL, Durand CM, Siliciano JD, Siliciano RF. 2015. Ex vivo analysis identifies effective HIV-1 latency-reversing drug combinations. J Clin Invest 125:1901–1912. https://doi.org/10.1172/JCI80142.

82. Einkauf KB, Osborn MR, Gao C, Sun W, Sun X, Lian X, Parsons EM, Gladkov GT, Seiger KW, Blackmer JE, Jiang C, Yukl SA, Rosenberg ES, Yu XG, Lichterfeld M. 2022. Parallel analysis of transcription, integration, and sequence of single HIV-1 proviruses. Cell 185:266–282.e15. https://doi.org/10.1016/j.cell.2021.12.011.

83. Peluso MJ, Bacchetti P, Ritter KD, Beg S, Lai J, Martin JN, Hunt PW, Henrich TJ, Siliciano JD, Siliciano RF, Laird GM, Deeks SG. 2020. Differential decay of intact and defective proviral DNA in HIV-1-infected individuals on suppressive antiretroviral therapy. JCI Insight 5:e132997. https://doi.org/10.1172/jci.insight.132997.

84. Neidleman J, Luo X, Frouard J, Xie G, Hsiao F, Ma T, Morcilla V, Lee A, Telwatte S, Thomas R, Tamaki W, Wheeler B, Hoh R, Somsouk M, Vohra P, Milush J, James KS, Archin NM, Hunt PW, Deeks SG, Yukl SA, Palmer S, Greene WC, Roan NR. 2020. Phenotypic analysis of the unstimulated in vivo HIV CD4 T cell reservoir. Elife 9:e60933. https://doi.org/10.7554/eLife.60933.

85. McNamara LA, Ganesh JA, Collins KL. 2012. Latent HIV-1 infection occurs in multiple subsets of hematopoietic progenitor cells and is reversed by NF-κB activation. J Virol 86:9337–9350. https://doi.org/10.1128/JVI.00895-12.

86. Yang J, Zhao Y, Kalita M, Li X, Jamaluddin M, Tian B, Edeh CB, Wiktorowicz JE, Kudlicki A, Brasier AR. 2015. Systematic determination of human cyclin dependent kinase (CDK)-9 interactome identifies novel functions in RNA splicing mediated by the DEAD box (DDX)-5/17 RNA helicases. Mol Cell Proteomics 14:2701–2721. https://doi.org/10.1074/mcp.M115.049221.

87. Kharytonchyk S, King SR, Ndongmo CB, Stilger KL, An W, Telesnitsky A. 2016. Resolution of specific nucleotide mismatches by wild-type and AZT-resistant reverse transcriptases during HIV-1 replication. J Mol Biol 428:2275–2288. https://doi.org/10.1016/j.jmb.2016.04.005.

88. Pizzato M, Erlwein O, Bonsall D, Kaye S, Muir D, McClure MO. 2009. A one-step SYBR green I-based product-enhanced reverse transcriptase assay for the quantitation of retroviruses in cell culture supernatants. J Virol Methods 156:1–7. https://doi.org/10.1016/j.jviromet.2008.10.012.