UC Berkeley
UC Berkeley Previously Published Works

Authors

Montoya, Vanessa R
Ready, Trine M
Felton, Abby
et al.

A Virus-Packageable CRISPR System Identifies Host Dependency Factors Co-Opted by Multiple HIV-1 Strains

Vanessa R. Montoya,[a] Trine M. Ready,[b] Abby Felton,[b] Sydney R. Fine,[b] Molly OhAinle,[b*] Michael Emerman[b]

[a]Molecular and Cellular Biology Graduate Program, University of Washington, Seattle, Washington, USA
[b]Divisions of Human Biology and Basic Sciences, Fred Hutchinson Cancer Center, Seattle, Washington, USA

ABSTRACT At each stage of the HIV life cycle, host cellular proteins are hijacked by the virus to establish and enhance infection. We adapted the virus packageable HIV-CRISPR screening technology at a genome-wide scale to comprehensively identify host factors that affect HIV replication in a human T cell line. Using a smaller, targeted HIV Dependency Factor (HIVDEP) sublibrary, we then performed screens across HIV strains representing different clades and with different biological properties to define which T cell host factors are important across multiple HIV strains. Nearly 90% of the genes selected across various host pathways validated in subsequent assays as bona fide host dependency factors, including numerous proteins not previously reported to play roles in HIV biology, such as UBE2M, MBNL1, FBXW7, PELP1, SLC39A7, and others. Our ranked list of screen hits across diverse HIV-1 strains form a resource of HIV dependency factors for future investigation of host proteins involved in HIV biology.

IMPORTANCE With a small genome of ~9.2 kb that encodes 14 major proteins, HIV must hijack host cellular machinery to successfully establish infection. These host proteins necessary for HIV replication are called "dependency factors." Whole-genome, and then targeted screens were done to try to comprehensively identify all dependency factors acting throughout the HIV replication cycle. Many host processes were identified and validated as critical for HIV replication across multiple HIV strains.

KEYWORDS CRISPR screen, T cells, dependency factor, human immunodeficiency virus, transcription factors, virus replication

With a small genome of ~9.2 kb that encodes 14 major proteins, HIV must hijack host cellular machinery to successfully establish infection. These host proteins, called dependency factors, are needed for entry, transit into the nucleus, uncoating, viral integration, as well as subsequent steps of transcription of viral RNA, RNA splicing and export, protein translation, and virion assembly and budding.

While many of these important host factors have been identified one by one, high-throughput methods have also attempted to create comprehensive maps of host dependency and/or restriction factors. For example, four genome-wide RNA-interference screens (1–4) as well as a genome-wide CRISPR/Cas9 screen (5) sought to identify dependency factors for HIV infection. While these screens identified genes important for cellular attachment and entry, nuclear entry, integration, transcription, and nuclear export (1–5), there was very poor overlap across screens (6). Other approaches have sought to identify novel dependency factors through protein-protein interaction screens and identify hundreds of human proteins that are physical interactors with the 18 HIV-1 proteins and polyproteins in Jurkat T-lymphocyte cells, and 293T cells (7). A subset of these interacting proteins were functionally validated in primary CD4$^+$ T cells as HIV dependency or restriction factors (8). In addition, interaction-based and gene network approaches have been used to computationally predict dependency factor genes (9) and to develop a viral-host dependency epistasis map

(Ve-MAP) (10). However, none of these approaches have used an assay with a functional readout to examine multiple viral strains across the entire HIV life cycle in T cells.

We had previously developed a high-throughput CRISPR screening method, called HIV-CRISPR, in which lentiviral genomes encoding sgRNAs are incorporated into budding virions, serving as a readout for genes important for HIV infection (11–13). Although this work focused on finding restriction factors against HIV-1, we demonstrated that this method could also identify dependency factors, although as the guide library used in those studies specifically targeted interferon-stimulated genes, the number of dependency factors identified was relatively limited. Here, we use a whole-genome guide library to look more globally at HIV dependency factors in a T-cell line. We used these data to inform the design of a smaller, custom HIV Dependency Factor (HIVDEP) CRISPR sublibrary to identify HIV dependency factors across multiple HIV strains. Our screens identified many previously reported HIV dependency factors across multiple parts of the viral life cycle, including the HIV-1 receptor CD4, coreceptor CXCR4, LEDGF/p75, NF-κB, and many genes encoding components of the mediator complex. Further, we identify genes not previously reported to play a role in HIV biology involved in transcriptional regulation, protein degradation, RNA regulation, as well as epigenetic factors affecting both early and late events of viral replication. Genes not previously identified as being important for HIV replication that were validated here include genes involved in Cullin-ring ligase mediated protein degradation, such as UBE2M and FBXW7, pre-mRNA alternative splicing regulator MBNL1, transcription factor PELP1, and zinc transporter/tyrosine kinase activator SLC39A7. Our expanded catalog of ~200 host dependency factors required at different stages of the life cycle has potential to inform therapeutic and cure design.

RESULTS

A whole-genome CRISPR screen for HIV dependency genes. The HIV-CRISPR screening method has previously identified restriction factors in THP-1 cells (a CD4$^+$ monocytic leukemia cell line) involved in the interferon response against HIV using a focused CRISPR guide library of interferon-stimulated genes (ISGs) (11). To identify HIV dependency factors more comprehensively, we adapted this approach to a whole-genome strategy utilizing the Toronto Knockout version 3 whole-genome (TKOv3) library. This library targets 18,053 protein-coding genes with 4 guides per gene and includes 142 nontargeting control guides (14).

Our screening strategy is outlined in Fig. 1A. The HIV-CRISPR vector is a lentiviral vector that contains the sequences for a guide RNA (from an sgRNA library) targeting a single gene, the Cas9 enzyme, a viral packaging signal (psi), and a repaired long terminal repeat region which allows transcription of a genomic RNA (11). In addition to transcribing Cas9 and the encoded sgRNA, cells that are transduced with the HIV-CRISPR vector are also capable of making mRNA corresponding to full-length HIV-CRISPR genomes after infection with HIV-1. Thus, upon infection with an HIV strain of interest, newly produced virions will also encapsidate the HIV-CRISPR genome in *trans*. Therefore, if a dependency factor has been knocked out, these cells will support less productive infection and fewer HIV-CRISPR genomes containing the sgRNA specific for the dependency factor will be packaged and released from infected cells. This relative depletion of sgRNAs in the viral supernatant, compared to the genomic DNA (13) is the readout in which we infer knockout of a dependency factor. The advantage of this system is that CRISPR guides can be evaluated in bulk rather than arrayed in single wells, which makes a whole-genome approach feasible.

We transduced Jurkat-CCR5 cells, a T cell line susceptible to HIV infection and engineered to express the CCR5 coreceptor, with the HIV-CRISPR vector containing the TKOv3 whole-genome library. The cells were then infected with the LAI strain of HIV-1 at an MOI of 0.5, and genomic DNA and viral RNA was harvested from cell pellets and viral supernatants, respectively, 3 days after infection. These samples were then deep sequenced to quantify enrichment or depletion of guide RNAs in the viral supernatants compared to guide representation in the cellular genomic DNA (13). Using the Model-based Analysis of Genome wide CRISPR/Cas9 Knockout (MAGeCK) screens algorithm and the log 2-fold change in gene
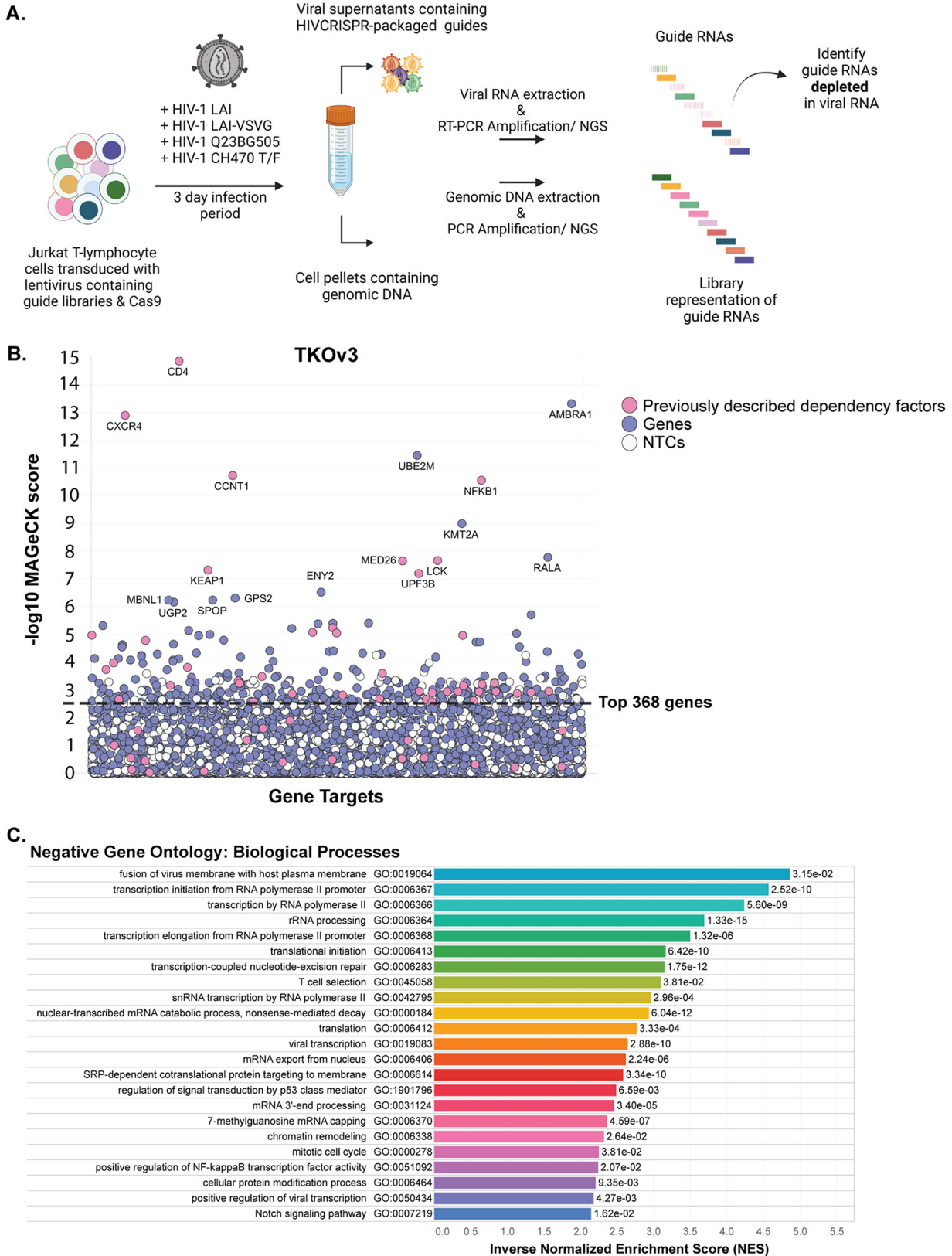
FIG 1 Genome-wide HIV-CRISPR screening to identify dependency factor candidates for the HIV Dependency Factor guide library. (A) HIV-CRISPR screening process. Jurkat-CCR5 cells containing a CRISPR knockout library are infected with HIV-1 in duplicate infections. Strains used in this study are listed above the arrow. Viral RNA and genomic DNA were collected 3 days postinfection and sequences corresponding to sgRNAs present in virions (vRNA) and genomic DNA (gDNA) were quantified through deep sequencing. (B) MAGeCK gene analysis of the genome-wide screen

depletion based on all four guide RNAs, statistical scores were assigned (15, 16) to delineate the most depleted genes, indicating factors important for infection (Fig. 1B; Table S1).

As expected, given their essential role in viral entry, the receptor CD4 and the coreceptor CXCR4 (required for the CXCR4-tropic strain HIV-1$_{LAI}$ used in this screen), were among the most depleted genes, ranking as the top scoring dependency factors (Fig. 1B). Additionally, many other previously reported dependency factors were identified in the 300 top scoring dependency factors (shown as pink dots) that include all different parts of the HIV life cycle, including entry (CD4, CXCR4, LCK), integration and uncoating (PSIP1, TNPO1), transcription (POLR2A, NFKB1, SP1, CyclinT1 [CCNT1]) mediator complex genes (MED7, MED10, MED16, MED18, MED23, MED26), and budding (TSG101, CHMP4B). Importantly, the nontargeting controls randomly binned to synthetic "genes" cluster near the bottom.

As a global approach to identify the most negatively enriched biological processes of the genome-wide screen (candidate dependency factor pathways), we used the Gene Set Enrichment Analysis (GSEA) as part of the MAGeCK-Flute pipeline (15). Of the top 20 negatively enriched Gene Ontology pathways (Fig. 1C), three are classified as important for viral processes, including fusion of virus membrane with host plasma membrane (GO:0019064), viral transcription (GO:0019082), and positive regulation of viral transcription (GO:0050434). Notably, there were enriched pathways identified that may play roles for multiple stages of the viral life cycle, including entry/egress, transcription, mRNA capping and processing, translation, protein modification/trafficking, and T cell selection. Notwithstanding the caveat that the screen missed some expected dependency factors that did not cluster away from the nontargeting controls (NTCs) (pink dots below the line in Fig. 1B), this screening approach can identify dependency factors across the entire viral life cycle and provides ample opportunity to investigate novel gene roles in the HIV life cycle.

An HIV dependency sublibrary screened with multiple HIV-1 strains. Smaller, targeted libraries are more powerful, as it is easier to maintain suitable coverage, more guides can be utilized per gene, and the smaller screens can be done in much higher throughput in more replicates. Therefore, before doing any validation of the hits in the whole-genome screen, we used data from our whole genome screen to inform the construction of a targeted CRISPR guide library called the HIVDEP (HIV Dependency Factor) library. We initiated this library by including guides targeting genes in the top scoring 500 from the TKOv3 genome-wide screen. After subtracting nontargeting controls from the rankings, this amounted to 368 genes. We also added back the known dependency factors that fell below this threshold in the screen, including CCR5 and PAPSS1, which are essential for CCR5-dependent strains. Because transcription and chromatin modification/regulation were major categories of negatively enriched genes in the whole-genome screen (Fig. 1C), we also performed an additional targeted screen using a custom-designed Human Epigenome (HuEpi) library consisting of 838 human epigenome and epigenetic regulator genes (17) and included any gene with <10% false discovery rate (FDR) from the HuEpi screen that were not already included in the top TKOv3 gene list (Fig. S1A) as well as smaller number of negatively enriched genes in an ISG-related library (11) (Fig. S1B), resulting in an additional 131 genes. Finally, 11 genes previously called "nonessential" (18) that were neither depleted nor enriched in our genome-wide screen were included as negative-control genes. In total, the HIVDEP library consists of 525 genes targeted by 8 guides and 210 NTC guides for a total of 4,401 guides (Table S2).

FIG 1 Legend (Continued)
showing the most depleted (dependency) genes in Jurkat T cells infected with HIV-1$_{LAI}$ in duplicate infections. The x axis shows randomly arrayed target genes. The y axis shows the -log$_{10}$ MAGeCK score for each gene. Host factors previously reported as dependency factors are shown in pink. Other genes in the library are shown in purple. "Synthetic" nontargeting control (NTC) genes generated *in silico* by iterate random binning of the 142 NTC sgRNA sequences to generate a negative-control set are shown in white. Gene names are shown for all hits with a -log$_{10}$ MAGeCK score greater than 6 and the entire list is in Table S1. The top 368 most depleted candidate genes (excluding the synthetic NTC genes) are indicated by the dashed black line. (C) Gene Set Enrichment Analysis (GSEA) of the genome-wide screen. The top 20 most enriched Negative Gene Ontologies (most enriched pathways of the depleted/dependency factor candidate genes) are shown here in ranked order by inverse Normalized Enrichment Score (NES). Adjusted P values are displayed next to each gene ontology.

TABLE 1 Viruses used in the HIVDEP screens reported in this study[a]

HIV-1 strain	Clade	Coreceptor tropism	Traits	% p24 3 dpi per biological replicate
LAI	B	CXCR4	Lab-adapted	LAI: R1 = 10.57; R2 = 10.33
				LAI redo: R1 = 7.05; R2 = 7.5
LAIΔenv-VSV-G	B	NA	Lab-adapted/alternate entry mechanism	R1 = 23.93; R2 = 23.9
Q23.BG505	A	CCR5	Primary isolate	R1 = 6.19; R2 = 6.32
CH470 TF	B	CCR5	Primary isolate/transmitted founder virus	R1 = 20.6; R2 = 22.9

[a]LAI was used in replicate in two independent experiments to assess reproducibility of the screen (referred to as "LAI redo" in this paper). Percent infection of each biological replicate of each HIVDEP screen (as measured through flow cytometry quantification of cytoplasmic gag 3 days postinfection) is shown in the last column.

To identify specific host factors that are either important for all HIV-1 strains, or ones that are strain-specific, we used the HIVDEP library screened against genetically distinct HIV-1 strains (Table 1). Specifically, these include HIV-1$_{LAI}$, which is in Clade B and uses coreceptor CXCR4, HIV-1$_{Q23.BG505}$ which is a clade A CCR5-tropic strain (19), and a transmitted/founder virus from clade B that is R5-tropic, HIV-1$_{CH470\ TF}$ (20). We also included an HIV-1 strain (LAI) that was deleted for its own envelope gene and pseudotyped with VSV-G to delineate entry-specific factors. Thus, we used HIV-1 strains that represent two clades (A and B), utilize different entry mechanisms, (-X4, -R5, or VSV-G), and are either lab-adapted versus primary strains. Each screen was done in duplicate and the number of infected cells in each screen ranged from 6% to 23% (Table 1). In addition, we also performed screens with the HIV-1$_{LAI}$ strain on two different occasions each with a separate transduction of the library followed by infection with different stocks of HIV-1$_{LAI}$ (screens referred to as LAI and LAI redo in Fig. 2).

We compared the MAGeCK scores of genes versus the synthetic nontargeting controls (NTC) in the genome-wide screen versus each of the screens done with the dependency factor-focused HIVDEP library using box and whisker plots (Fig. 2A). While the mean score of the synthetic nontargeting controls in each screen are similar, the mean of the MAGeCK scores for the genes in each HIVDEP screen was significantly higher than the mean of the genes in the genome-wide TKOv3 screen (Fig. 2A; Welch's *t* test). Moreover, the scores of the 95th percentile of genes (top horizontal line in each plot) were also increased in the HIVDEP library relative to the whole-genome screen for each of the HIV strains tested (Fig. 2A). This indicates that the sublibrary approach was successful in increasing the distance of noise to signal (NTCs to genes) relative to the whole-genome screen. Moreover, waterfall plots of the top 20 hits from each screen (Fig. 2B to F) show that known dependency factors are identified, such as the receptor CD4, CCNT1 (CycT1), NFKB1, and members of the mediator complex. Moreover, several novel genes score in the top 20 for multiple strains, including UBE2M, MBNL1, FBXW7, PCGF1, and PPP2CA, implicating those gene products in processes important for viral infection across HIV-1 strains.

Validation of hits from the combined HIV dependency factor screens across different cellular processes. While hits across two screens done with the same strain are similar, the absolute MAGeCK scores between the two are not identical (LAI and LAIredo in Fig. 2B and C), Therefore, to compare hits across screens, we normalized our data with a z score analysis, as previously described (21). We thus created a ranked order list of both the average of all screen results by z score as well as the replicates from screens with each strain (Table S3). We created pathway-focused heatmaps using the z scores for each replicate and the GSEA pathways established from the whole-genome screen (Fig. 3). As an initial control to determine if the screen could distinguish different strains, we looked at known entry factors that would be expected to be different between CXCR4-using, CCR5-using, and the VSV-G pseudotyped viruses in the Binding and Viral Entry Gene Set (Fig. 3A). The HIV-1$_{LAI}$ screens correctly identify CXCR4, but not CCR5 nor PAPSS1, while the opposite is true of the CCR5-tropic viruses, Q23BG505 and CH470TF. Each of the strains that uses wild-type HIV-1 envelope require CD4, but the VSV-G pseudotyped HIV-1 does not. Lymphoid specific Src kinase (LCK), which is important for T cell signaling downstream of CD4 engagement and reported to be important for viral core transit from the plasma membrane to the nucleus (22), tracks with CD4 in our screens (Fig. 3A).

The pathway-focused heatmaps also show enriched genes for each screen with a large number involved in transcription, as well as rRNA processing, protein modification, chromatin
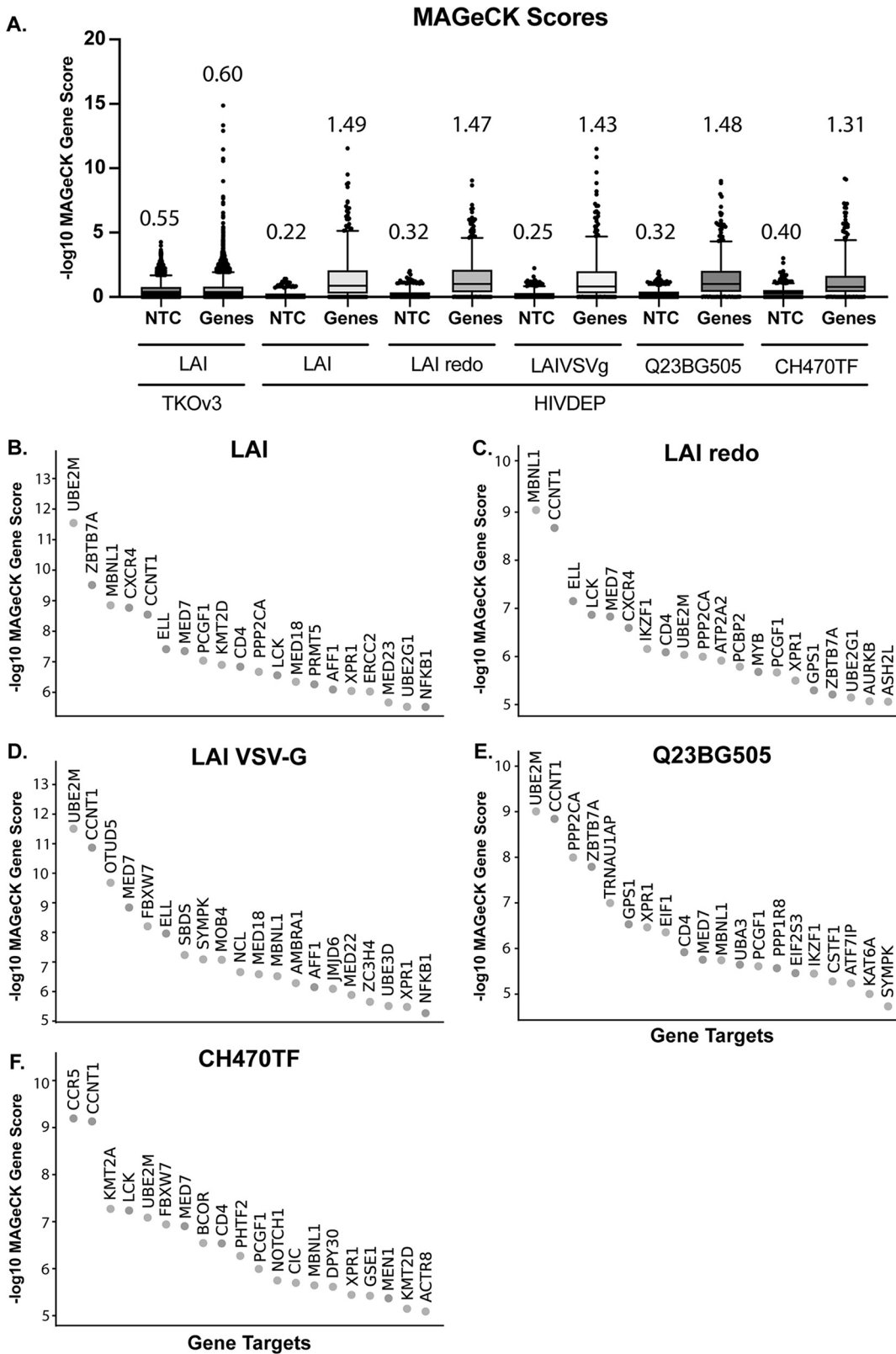
FIG 2 Iterative screening with the HIVDEP sublibrary enriches previously reported and candidate dependency factors. (A) MAGeCK score comparison of each CRISPR screen. Nontargeting control (NTC) sgRNA scores were randomly binned (four NTC guides per gene for TKOv3 or 8 NTC guides per gene for HIVDEP) to recapitulate the same number of genes in the respective libraries. The y axis shows -log₁₀ MAGeCK scores. The mean MAGeCK scores for the synthetic NTC versus the genes are shown above each graph and are represented by the line within each box. The top 95th percentile of NTC/genes

remodeling, and translation (Fig. 3). We then chose a subset of genes across different host cellular processes (Fig. 3) to functionally validate using infectious virus in multiround spreading infections with both LAI and with Q23BG505. We also picked some genes that scored lower in the screens in order to determine a cutoff from which we would be more confident of remaining hits that were not directly validated, as well as some genes that scored more highly for one strain versus another (Fig. 4A for the list ranked in descending order of average z score as well as the z score for each replicate of each gene). Our validation strategy used two guides per gene to knock out candidate dependency factors in Jurkat-CCR5 cells. Positive controls included knocking out CD4, and negative controls included NTC guides, as well as guides targeting genes that are not involved in HIV or T cell biology, CD19 (B cell marker), and AAVS1 (the AAV integration site often used as a safe harbor locus [23–25]). The knockout (KO) cells were maintained as pools rather than clones to avoid clone-to-clone differences and were tested for gene editing as well as HIV infection within 2 weeks of generation.

Each knockout pool was infected with HIV-1 strains at an MOI of 0.15. All infections were done in triplicate and virus growth was measured by assaying reverse transcriptase activity released into the supernatant 0, 3, 5, and 7 days after infection. Replication curves for the majority of the genes of interest (blue) lines were in-between the CD4-KO lines and negative-control lines (black), exhibiting inhibition of infection or slower growth kinetics for both LAI infection and Q23BG505 infections (Fig. 4B and Fig. S2). Infections were done in two batches with similar controls in both and grouped as such (Fig. 4C is batch 1 and Fig. 4D is batch 2). We used all the data points in the spreading infection to calculate an area under the curve (AUC) for each infection (Fig. 4C and D). We find that knockout of UBE2M which was one of the top hits in each of the screens (Fig. 2B to F) had the strongest phenotype in the spreading infections for both viruses tested (Fig. 4C and D). In order of their effects, the strongest hits in this set of genes (aside from the positive control of CD4) were UBE2M, PELP1, MBNL1, ZBTB7A, and FBXW7. In addition, KMT2D, KMT2A, KEAP1, CNOT3, AFF1, AMBRA1, SLC39A7, PUF60, RALA, and COPS4 validated for both viruses. MOB4 (Fig. 4C; Fig. S3C and D) and WDR36 at day 7 (Fig. S3C and D) validated for Q23BG505, but not for LAI. This was predicted for WDR36 since it's z score was lower for Q23BG505 (Fig. 4A), although not expected for MOB4, which had similar z scores across the two viral strains in the HIVDEP screen (Fig. 4A). We found similar results if we measured the RT activity in the supernatant at day 5 or at day 7 as the primary readout instead of AUC (Fig. S3C and D), and find that at both time points, knockout of UBE2M which was one of the top hits in each of the screens (Fig. 2C and 3F) had the strongest phenotype in the spreading infections. At the time of infection, we also quantified the gene editing efficiency for each pool which were generally between 50% and 80% (pie charts in Fig. 4). Therefore, as these are less than complete knockouts, the degree of decreased virus replication in the absence of the candidate genes is likely an underestimate. However, a lack of gene editing would not explain the failure of ARID1A nor MEN1 to validate (Fig. 4D). There were not notable differences in the cell growth of the knockout pools, other than the pooled PELP1 KO cells early in the selection process. PELP1 had lower editing efficiency, 27% as determined by Inference of CRISPR Edits (ICE, see Materials and Methods), which may indicate a larger outgrowth of the WT population with a smaller population of KO in the pool. Nonetheless, these results show that about 90% of the selected screen hits validated as HIV dependency factors to some degree.

To test the hypothesis that the strength of effect of the knockouts on HIV infectivity is correlated with its z score in the screen, we graphed the z score against the respective amount of virus replication represented by the AUC (Fig. 4C and D) in spreading infections. Indeed, there was a negative correlation between the strain specific z scores and the amount

FIG 2 Legend (Continued)
is represented as the top horizontal line. For statistical analyses, the MAGeCK scores of the synthetic nontargeting controls (shown here as NTCs) were compared to the MAGeCK scores of the genes in each screen. The gene MAGeCK scores per each screen were also compared across screens. Each comparison resulted in significant P values (****, $P < 0.0001$; Welch's t test). Waterfall plots of the top 20 genes or all genes in each HIVDEP screen in descending order are shown for the following HIV-1 strains (B) LAI, (C) LAIredo, (D) LAI-VSV-G, (E) Q23BG505, and (F) CH470TF.
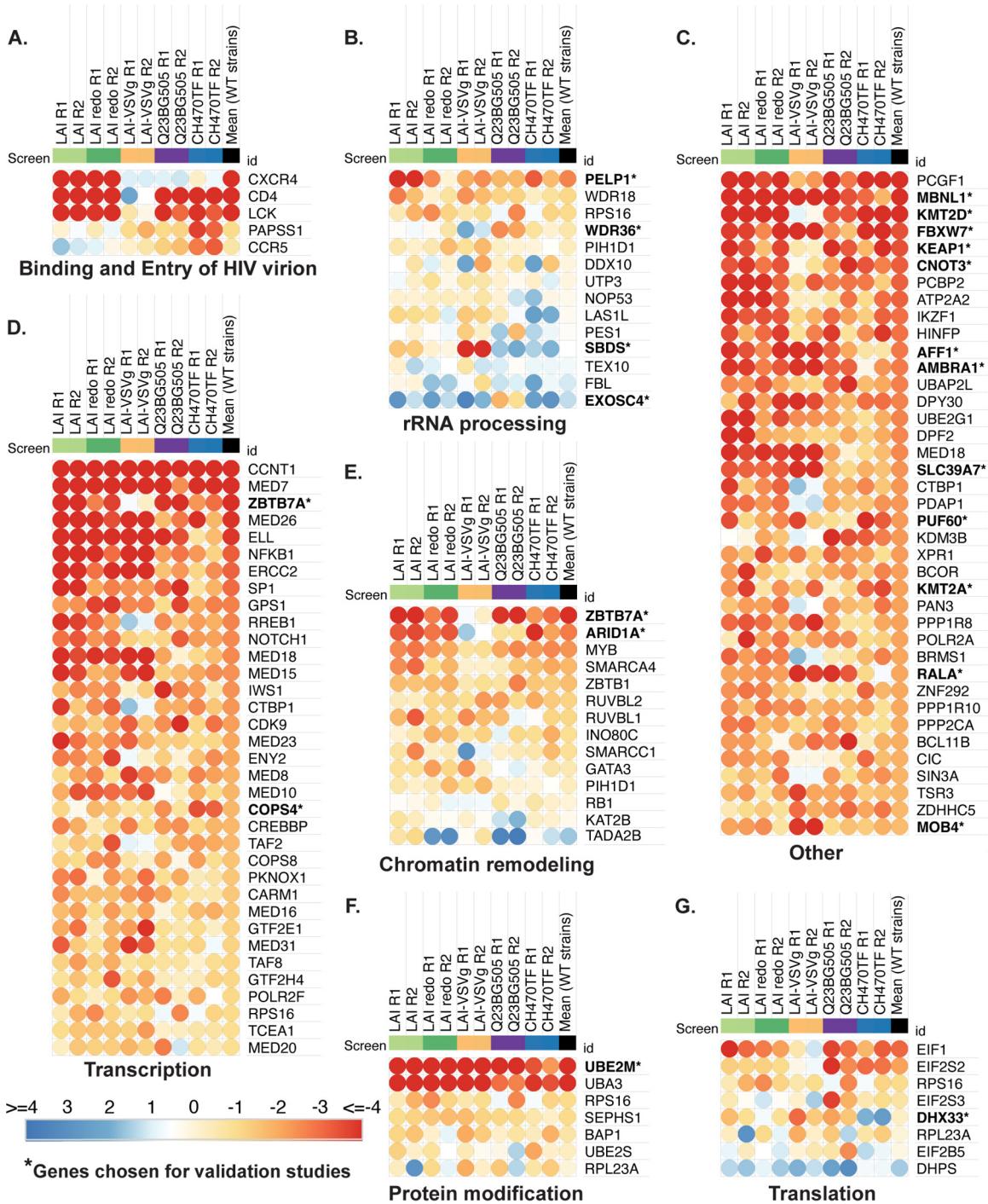
FIG 3 Common and differential use of host cellular pathways by HIV-1 strains. Comparative pathway-focused heatmaps showing enriched or depleted sgRNAs across each HIVDEP screen. The pathways shown are derived from the top 20 most enriched Negative Gene Ontologies of the genome-wide screen (Fig. 1C). Any genes not included in the HIVDEP library were excluded. The z scores were calculated as previously described (21). Z scores on each heatmap are colored from red (lowest/most depleted genes, i.e., dependency factors) to blue (highest/most enriched genes, i.e., negative or restriction factors). The median NTC z score was 0.5 and marks the inflection in the color scale. (A to G). Each biological replicate is represented as a separate column showing the mean scores across wild-type (non-VSV-G pseudotyped) strains. The "Transcription" and "Other" heatmaps were truncated to the top 40 hits each heatmap. Gene names that are bolded with an asterisk indicate they were chosen for validation studies in Fig. 4 and 5. "Other" shows genes that were not assigned to any of the top Gene Ontology categories from Fig. 1C.

of virus replication produced in each knockout pool (Fig. 5A); LAI ($R^2 = 0.39$, $P = 0.001$) and Q23BG505 ($R^2 = 0.41$, $P = 0.001$). We also find that this correlation holds if we use the mean z score of all WT strains tested in the screen (Fig. 5B; LAI [$R^2 = 0.30$, $P = 0.007$] and Q23BG505 [$R^2 = 0.46$, $P = 0.0004$]), which argues that despite strain differences, the list of hits (Table S3)
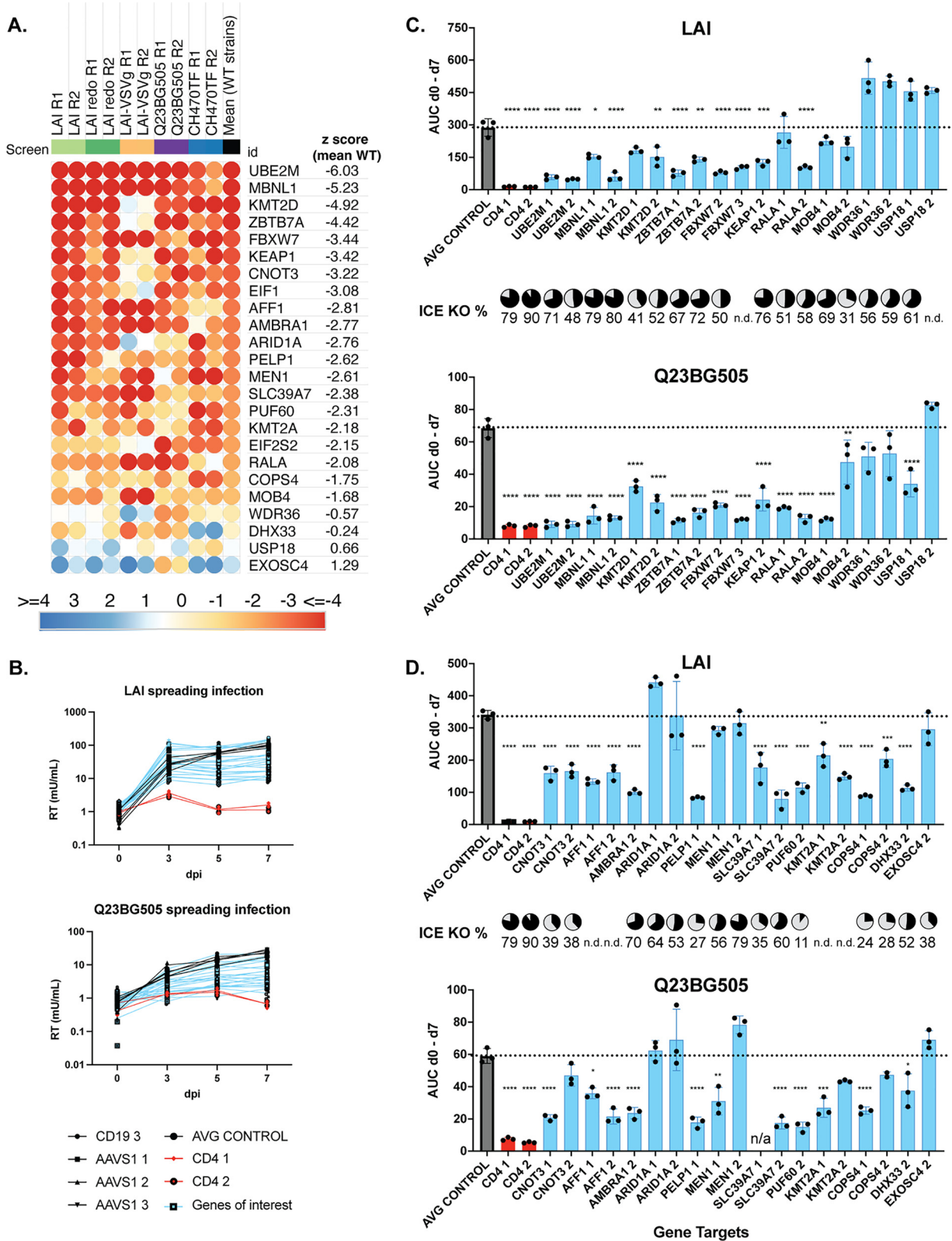
FIG 4 Validation of curated top hits list. (A) Heatmap of the candidate genes used for validation studies, ordered by the mean z score of WT strains (B) Pooled knockout Jurkat-CCR5 generated by transducing with lentiviral vectors encoding sgRNAs, including positive-control gene CD4, negative controls CD19 or AAVS1, or candidate dependency factor genes. Two knockout lines per gene were generated using the highest scoring sgRNAs selected from across each HIVDEP screen. EIF1 KO cells were not used in the infection assays because of poor viability. Viral supernatants were collected at days 0, 3, 5, and 7 to assess overall effect on replication kinetics via reverse transcriptase activity
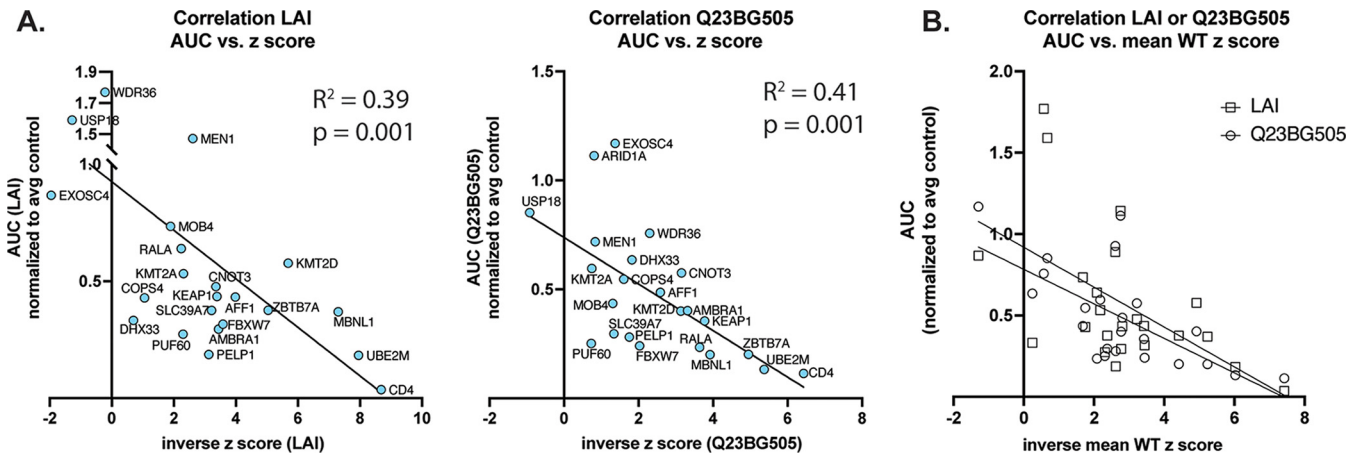
(Continued on next page)

FIG 5 Correlation of gene z scores with area under the curve (AUC) (A) AUC was compared to the inverse z score of either LAI (left) or Q23BG505 (right). For statistical analysis, the mean biological replicate inverse z score per each gene for either LAI or Q23BG505 are compared to the mean AUC for infection of both guide knockouts per gene from Fig. 4. Simple linear regression: LAI z score versus AUC ($R^2 = 0.39$, $P = 0.001$); Q23BG505 z score versus AUC ($R^2 = 0.41$, $P = 0.001$); (B) same as panel A, but the mean z score for all wild-type strains (Table S3) was used rather than strain-specific z score. Mean WT z score versus LAI AUC ($R^2 = 0.30$, $P = 0.007$); Mean WT inverse z score versus Q23BG505 AUC ($R^2 = 0.46$, $P = 0.0004$). Squares and top line represent LAI; circles and lower line represent Q23BG505.

can be used as a guide to identify HIV dependency factors. Thus, hits that scored higher in our screen are more likely to be authentic HIV dependency factors than those with lower z scores.

At the lower end of hits in our validation experiments, MOB4 (mean WT z score, -1.683) validated for Q23BG505, though not LAI while each gene with z scores greater than MOB4 did not validate except for DHX33. WDR36, the next highest z score gene we chose for validation, did validate for Q23BG505 at day 7, but not for LAI (Fig. S3C and D), as predicted from the screen, though this phenotype was weak as the significance was lost when calculating AUC (Fig. 4C). We therefore set two tiers of z score cutoffs with different levels of stringency to best identify hits for future validation studies. The most conservative tier, tier 1, was set to the MOB4 z score cutoff (Table S3). The tier 2 cutoff was set to the WDR36 z score. This is especially stringent as there are well studied and previously described dependency factors that score outside tier 1, but within tier 2, including PSIP1/LEDGF (ranking no. 102 with mean WT z score of -1.37), which was also validated as a dependency factor in our system (Fig. 6). This cutoff is likely still conservative as DHX33, which fell below tier 2 (ranking no. 280), did validate for both viruses (Fig. 4D), highlighting that there are likely other false negatives below this cutoff. We then conducted a literature review to identify how many of the 198 genes within tiers 1 and 2 had already been identified in screens as HIV dependency factors, as well as which ones have been experimentally validated (Table S3). We find that 35 genes in this list had been previously identified in one or more previous screens (1–3, 5, 8). Also, 57 of the 198 genes in our tier 1 and 2 groups had been validated as HIV dependency factors through functional studies found in the literature (Table S3). Therefore, we report here an estimated over 140 novel HIV dependency factors, including the genes validated in Fig. 4 and 6, with more confidence in hits with lower z scores, such as UBE2M and MBNL1 (Fig. 5).

Combined HIV dependency screens highlight host entry factor preferences. As many experiments with HIV are done with VSV-G pseudotyped viruses to increase titers in single round infections, we were interested in further exploring hits with high differential z scores between the strains with WT HIV-1 envelope and VSV-G pseudotyped virus. Thus,

FIG 4 Legend (Continued)
at each time point. The spreading infections were performed over two batches. The y axis shows reverse transcriptase milliUnits/mL. Batch 1 is shown in panel B and batch 2 is shown in Fig. S2. (C and D) Area under the curve (AUC) was calculated for each cell line after 7 days of infection in batch 1 (C) and batch 2 (D) with either LAI or Q23BG505. Infection data of each guide is shown separately. For statistical analysis, all conditions are compared to the mean of the control cell lines (CD19 and AAVS1). One-way Anova; Tukey's multiple corrections test: ns, $P > 0.05$; *, $P < 0.05$; **, $P < 0.05$; ***, $P < 0.001$; ****, $P < 0.0001$. For each knockout line, Synthego ICE analysis was performed and knockout scores are displayed as pie charts in line with the corresponding gene target. ND, could not be determined.
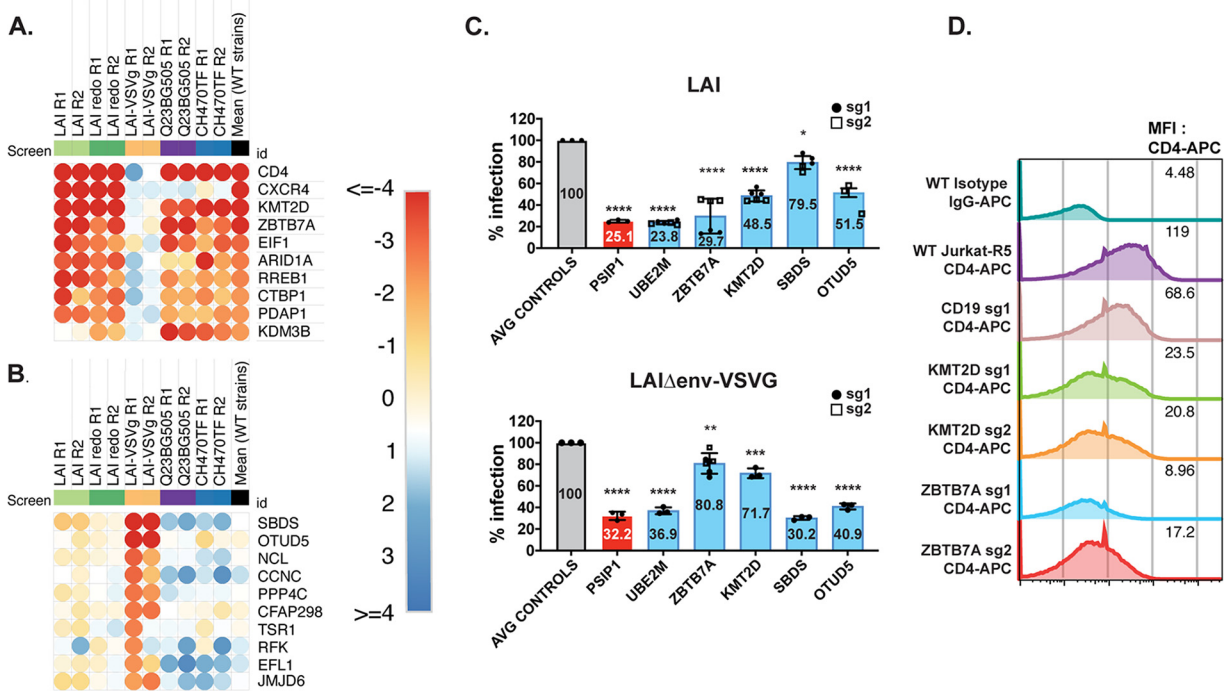
FIG 6 Entry-specific host factors. (A) Heatmap of the top 10 most depleted sgRNAs based on the mean of the wild-type strains, but not for VSV-G pseudotyped HIV-1. (B) Heatmap of the top 10 most depleted sgRNAs for VSV-G-pseudotyped HIV-1, but not for HIV-1 with WT HIV. Matrix for determination of panels A and B is in Fig. S3 with the heatmap z score values in Table S3. (C) Jurkat-CCR5 cell pools edited for gene targets of interest were created by transducing wild-type Jurkat-CCR5 cells with lentiCRISPRv2 sgRNA constructs using sgRNAs for CD19 (B-cell marker used as a negative control), PSIP1 encoding p75/LEDGF (positive control), and genes of interest (UBE2M, ZBTB7A, KMT2D, SBDS, and OTUD5), and selected with puromycin for at least 10 days. Knockout pools were infected with HIV 1$_{LAI}$ or VSV-G-pseudotyped HIV-1 which both encode luciferase in place of the nef gene. Luciferase expression in infected WT or knockout cells was quantified 2 days postinfection. All infections were done in triplicate using two pools of knockouts with one or two different sgRNA guides per gene. Data for the two different guides are shown as black circles (sg1) and open squares (sg2) for each pooled knockout cell line. The mean percentage of luciferase activity of all replicates relative to the control cells is displayed on each bar. (D) Surface CD4 expression of WT or KO Jurkat-ZsGreen/CCR5 cells was quantified using flow cytometry for CD4-APC. Mean fluorescence intensity (MFI) for each cell line is shown in each respective row. Two different knockout pools corresponding to two different guides per gene (except for CD19) are shown.

we averaged the z scores for each gene across HIV-1 strains with wild-type envelopes and compared these scores with the z scores for the VSV-G pseudotyped virus (Fig. S3). Using this correlation matrix, we ranked the 10 most highly ranked genes scoring as dependency factors for the HIV-1 strains with WT envelopes, but not LAI-VSV-G (Fig. 6A) and separately ranked the 10 most highly ranked genes for VSV-G pseudotyped HIV, but not for any of the WT HIV-1 stains (Fig. 6B). We then functionally tested examples of each of these by knocking out the genes in Jurkat T cells and infecting them in a single round of infection with either an HIV-1 that either used VSV-G for entry or authentic HIV-1 envelope for entry. As positive controls, we used PSIP1 (also known as LEDGF), which should be needed for integration regardless of entry mechanism (26), and UBE2M which scored as one of the highest hits for all strains used in our screens regardless of entry mechanism or coreceptor (Fig. 4). Negative controls included both Jurkat cells in which sgRNAs were used to target CD19, AAVS1, the safe harbor locus described above, and untransduced Jurkat cells.

We find that the decrease in infectivity of the cells knocked out for both PSIP1 and UBE2M was similar regardless of whether entry was through HIV-1 env or VSV-G (Fig. 6C). Moreover, as predicted from the screen data, ZBTB7A and KMT2D were important for infection with HIV-1 envelope but had a minimal effect on infection with the HIV-1 pseudotyped with VSV-G (Fig. 6C). In contrast, knockout of SBDS affected the VSV-G pseudotyped virus, but not HIV-1 envelope-mediated entry while OTUD5 knockout had an effect on both viruses but decreased the VSV-G pseudotype to a greater extent. Examination of CD4 levels on the pools of knockout cells show that cells with ZBTB7A and KMT2D have reduced levels of cell surface CD4 relative to the control knockout cells (Fig. 6D). These results indicate

some HIV dependency factors act indirectly on the virus by affecting receptor availability on the cell surface, and other factors that affect pseudotyped viruses may not be relevant for WT HIV. Thus, our screens with whole genome and more focused libraries identify both shared and strain-specific HIV dependency factors across a broad range of cellular processes.

DISCUSSION

We first used HIV-CRISPR screening at the genome-wide scale and then at a smaller, targeted scale to uncover novel genes that act as HIV-1 dependency factors. By using the HIV-CRISPR screening technique, which assesses enrichment or depletion of guide-encoding genomes based on HIV-1 release from infected cells, we can establish a powerful functional assay to identify factors that promote HIV infection across the entire viral life cycle in a T lymphocyte cell line. Additionally, we were able to uncover dependency factors across multiple strains, from different clades, and with different coreceptor tropisms. In addition to the genes that were validated, we also conducted GSEA pathway analyses to investigate which host pathways to which HIV-1 is most dependent in these screens, with transcription-related pathways being the most represented.

Transcription and chromatin remodeling factors as a major axis of HIV host dependency. Of the enriched pathways we identified, there were six transcription-related pathways (Fig. 1C) which comprised a list of 58 genes that were included in the HIVDEP library, as well as an enriched pathway that includes chromatin remodeling genes. These are the largest pathway-focused heatmaps, demonstrating an emphasis on transcription as a major focal point for host factors that promote HIV replication (Fig. 3D). Some of the most enriched genes included previously reported dependency factors (positive controls) such as pTEFb components, CCNTI and CDK9, which are important for transcriptional elongation of HIV transcripts, (mean WT z scores: -7.97 and -2.18), transcription factors NF-κB and SP1 (mean WT z scores: -3.24 and -3.02), and several mediator complex genes (MED7, MED26, MED18, MED15, MED23, MED8, MED10, MED16, MED20) (Table S3). In addition to Tat-recruited pTEFb members, the Super Elongation Complex (SEC) has also been reported to be required for efficient Tat transactivation and elongation (27, 28). ELL is a member of the SEC and scored highly for all strains, except for CH470TF. Similarly, AFF1, which encodes a scaffold protein of the SEC, also scored highly for each virus except CH470TF. AFF1 was also identified as a dependency factor in primary T cells (8). This is consistent with our data, as we see a reduction of infection for both LAI and Q23BG505 in AFF1 KO Jurkat cells (Fig. 4D). CREBBP, which encodes CREB-binding protein (CBP), is recruited by Tat to the viral LTR (29), and acetylation of Tat by CBP/p300 has been shown to be important for transcriptional activation at the LTR (30, 31). Here, we have shown that CREBBP is important for four different viruses as it scored highly in our screens for LAI, LAI-VSV-G, Q23BG505, and CH470TF (mean z score: -1.767).

Host dependency factors could impact HIV replication through either direct interaction with viral proteins or viral RNA/DNA, or by indirectly affecting other host dependency factors. There are likely several implicated factors in this study that act indirectly on viral replication. For example, ZBTB7A (also known as FBI-1 and LRF) is a zinc-finger protein involved in a diverse array of activities involving transcriptional corepressors (32). It has also previously been identified as a protein that binds the HIV-1 LTR and associates with HIV-1 Tat (33, 34). Notably, ZBTB7A was recently identified to affect human coronavirus 229E through modulation of oxidative stress (35). However, we find that ZBTB7A was a hit for each virus except for the VSV-G pseudotyped virus (mean replicate z score of 0.2) and resulted in lower cell surface CD4 levels in the knockout pools, implying ZBTB7A transcriptionally regulates pathways important for CD4 levels, rather than HIV transcription itself. Similarly, KMT2D, also known as MLL2, is a lysine methyltransferase, but known for posttranslational modification of histones (36) appears to act indirectly on HIV through effects on CD4 levels.

Another top hit of this study, PELP1, was previously identified in an siRNA screen in HeLa-CD4 cells as a restriction factor (37). However, our screens and validation experiments suggest that it acts as a dependency factor in Jurkat cells (mean replicate WT z score: -2.617) (Fig. 3B and 4D). PELP1 has been shown to interact with SETDB1, a methyltransferase, oncogene, and restriction factor that effectively inhibits Tat activity by methylation (38). Therefore,

it is possible that PELP1 promotion of SETDB1-Akt activation (39) may draw SETDB1 away from Tat, thus allowing transactivation. Therefore, PELP1 coregulation of HIV-1 transcription factors and SETDB1 activity could explain the dramatic reduction we see upon gene knockout in Jurkat T cells.

Cellular protein modification as HIV dependency factors. Our screens identified multiple genes involving cullin-mediated ubiquitin ligase complexes. Cullin ring ligases are multisubunit E3 ubiquitin ligase complexes responsible for ubiquitylating ~20% of cellular proteins targeted for degradation (40). Cullin ring ligases (CRLs) must undergo neddylation to be activated. Neddylation involves the transfer of ubiquitin-like molecule NEDD8 to a lysine residue of a substrate, subsequently affecting activity, conformation, and/or subcellular localization. Notably, UBE2M, which neddylates CRL-1, CRL-2, CRL-3, and CRL-4 via the substrate specific E3 ubiquitin ligase Rbx1 (mean WT z score: –0.54), was one of the top scoring hits across all screens, including the genome-wide LAI screen and the HIVDEP LAI-VSV-G screen, following essential dependency factors CD4 and CyclinT1 (mean WT z score CCNT1: –7.97; mean WT z score CD4: –7.42; mean WT z score UBE2M: –6.03). Knockout of UBE2M in Jurkat cells exhibited the most potent inhibition of infection of all the validated genes except for the positive control, CD4 (Fig. 4). Neddylation, and specifically UBE2M, has been shown previously to be important for activation of CRL-4 and CRL-5 for efficient degradation of restriction factors SAMHD1 by HIV-2 Vpx and APOBEC3G by Vif, respectively (41). However, as neither of these factors are important in Jurkat T cells infected with wild type HIV, we hypothesize we have identified other effects of UBE2M neddylation for HIV-1 replication. UBA3 is one of two components of the NEDD8 E1 Activating Enzyme (NAE) heterodimer, just upstream of UBE2M in the neddylation process (42). UBA3 scored extremely well across all screens, including the LAI-VSV-G pseudotyped virus (Fig. 3F; mean WT zscore: –4.258; mean z score of all: –4.61). Notably, inhibitors targeting the NAE complex have been effective in inhibiting HIV-1 infection (41), further validating these degradation pathways as important for the virus to replicate.

FBXW7, a substrate receptor for degradation of proteins dependent on phosphorylation status has a large array of transcriptional effects in cells (43) and was also validated in our screens (Fig. 3C; Fig. 4) as an HIV dependency factor. FBXW7 has many substrates through which there are downstream effects. Interestingly, FBXW7 substrate CCNE2 encodes protein Cyclin E1, and Cdk2 and Cyclin E have previously been shown to induce HIV Tat activity by phosphorylation of the Ser16 residue, thus acting as a dependency factor (44–47). It is possible that HIV-1 hijacks FBXW7 to sequester it, thus preventing or co-opting degradation of CCNE2 or one of its other substrates that have no previous implication for HIV-1 infection.

Additional HIV dependency factors. Of the genes chosen for validation studies, we validated 88% of these as dependency factors for at least one virus (Fig. 4). Thus, as there are 72 genes in tier 1 and 198 genes in tier 2, if we use this percentage of validation rate as a rough estimate across the genes in tier 1 and tier 2, we estimate our CRISPR screens have, at minimum, identified at least 156 HIV dependency factors in addition to the 18 validated here (Table S3). However, this is likely a minimum number as there are known HIV dependency factors that are below no. 198 (Table S3) on the list. For example, known dependency factors for budding, TSG101 (ranking at no. 299 with mean WT z score of −0.12) and CHMP4B (ranking at no. 147 with mean WT z score of −0.91), did not meet this cutoff. Thus, the absence of a gene in our top list is not evidence of its lack of effect. We find a negative correlation between z score and the magnitude of effect in spreading infections (Fig. 6), such that generally the higher the z score, the less likely that gene product will affect HIV replication.

There are certain caveats to our screen results. First, the screen is done in Jurkat T cells, therefore host genes that are more important in primary cells could be missed in our system. There was only a small overlap between our HIVDEP guide library and the library of interactors that were screened and validated in primary cells (8), however, we did identify some of the same genes, (including AFF1, which was validated in both studies) suggesting that our screens have identified factors relevant to primary CD4$^+$ T cell infection (Table S3). Of the previously mentioned RNAi screens and CRISPR screens, only one used the same cell type used in this study (4). Yeung et al. (4), used an shRNA whole-genome screening approach; however, there were only 3 overlapping genes, implicating that there are likely

differences in identification based on viruses used for each screen and/or significance criteria. Future screening with HIVDEP in other HIV-permissive cell types, most importantly primary CD4$^+$ T cells, may identify other host factors important for HIV-1 infection. Also, HIVDEP was designed based on a genome-wide screen using HIV$_{LAI}$, a CXCR4-tropic strain. Therefore, some genes that may be important for another strain, but not HIV$_{LAI}$ (or important for HIV$_{LAI}$ at a lower degree than our cutoff), could have been excluded from our design for the sublibrary HIVDEP, resulting in missing some relevant host factors. In the future, a new HIVDEP library can be generated based on other HIV-1 strains to potentially include more candidates based on strain-specific dependencies. Finally, as the differences in the two screens using HIV$_{LAI}$ in two different instances suggest (Fig. 3 and Tables S1 and 3), there is some inherent noise in large scale screens, therefore, some differences between different screens and different platforms, are likely stochastic. Finding hits among screens with multiple viral strains increases the confidence of hits, as well as independent validation of hits listed in Table S3. Despite these limitations, we have identified and validated many new dependency factors for HIV-1, including factors that are important across clades, as well as some that are strain-specific. This list of host dependency factors can be used to further explore which pathways HIV-1 must hijack and has substantial implications for anti-HIV-1 therapeutic design and HIV cure strategies.

MATERIALS AND METHODS

Cell culture. The Jurkat (ATCC) cell line was cultured in RPMI 1640 medium (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS), penicillin-streptomycin (Pen/Strep), and 10 mM HEPES. 293T and TZM-bl cells (ATCC) were cultured in Dulbecco modified Eagle medium (DMEM) (Thermo Fisher Scientific) with 10% FBS and Pen/Strep. Mycoplasma testing was done prior to screens and validation studies by the Fred Hutch Specimen Processing/Research Cell Bank Shared Resource and was not detected in any cell lines used in these experiments.

Plasmids. The HIV-CRISPR vector was previously described (11). HIV-CRISPR constructs targeting genes of interest were cloned by annealing complementary oligonucleotides with overhangs that allow directional cloning into HIV-CRISPR using the BsmBI restriction sites (Table S2). lentiCRISPRv2 plasmid was obtained from Feng Zhang via Addgene no. 52961. pMD2.G and psPAX2 were obtained from Didier Trono via Addgene no. 12259/12260. sgRNAs designed to target genes of interest were cloned into lentiCRISPRv2 using methods previously published (11). The wild type (HIV-1$_{LAI}$) and env-deleted (HIV-1$_{LAI-VSV-G}$) HIV-1$_{LAI}$ proviruses were previously described (48, 49). The Clade An HIV-1$_{Q23.BG505}$ molecular clone was previously described (19, 50). CH470TF was provided by Beatrice Hahn (51). The lentiviral pHIV-dTomato (Addgene plasmid; 21374; RRID: Addgene_21374) and pHIV-Zsgreen (Addgene plasmid; 18121; RRID:Addgene_18121) expression vectors were deposited by Bryan Welm. The pHIV-zsGreen/CCR5 and pHIV-dTomato/CCR5 constructs were created by cloning the human CCR5 CDS into pHIV-zsGreen and pHIV-dTomato backbones using the BamHI and NotI sites. The Toronto human knockout pooled library (TKOv3) was obtained from Jason Moffat via Addgene no. 90294 (14).

HIV dependency factor (HIVDEP) CRISPR/Cas9 sgRNA library construction. The HIVDEP sgRNA library is composed of 525 genes (4,191 sgRNAs). The top scoring 368 genes (as determined by -log$_{10}$ MAGeCK score) of the genome-wide screen were included in the HIVDEP library. Additional unique top scoring genes with <10%FDR from HuEpi and PIKA screens were added to the HIVDEP library that contained an additional 131 genes. NR5A1, NHLRC4, SFTPA2, ZNF768, MYL10, GIMAP5, SPG21, CHSY3, ZNF25, REG1A, and ATXN3 were manually selected as nonessential genes (18) that were neither enriched nor depleted in the genome-wide screen. Six new sgRNAs were designed using two algorithms, GUIDES (52) and CHOPCHOP (53). We also manually included other known dependency factors, including CCR5 and PAPSS1. A total of 212 nontargeting controls were designed using GUIDES and included. The HIVDEP sgRNA library was synthesized (Twist Biosciences) and cloned into HIV-CRISPR. Oligo pools were amplified using Herculase II Fusion DNA polymerase (Agilent; 600677) combined with 1 ng of pooled oligonucleotide template, primers ArrayF and ArrayR (ArrayF primer: TAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG and ArrayR primer: ACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTC TAGCTCTAAAAC), an annealing temperature of 59℃, an extension time of 20 s, and 25 cycles. Following PCR amplification, a 140-bp amplicon was gel-purified and cloned into BsmBI (NEB; R0580) digested HIVCRISPR using Gibson Assembly (NEB; E2611S). Each Gibson reaction was carried out at 50℃ for 60 min. Drop dialysis was performed on each Gibson reaction according to the manufacturer's protocol using a Type-VS Millipore membrane (VSWP 02500). Then, 5 μL of the reaction was used to transform 25 μL of Endura electrocompetent cells (Lucigen; 60242-2) according to the manufacturer's protocol using a Gene Pulser (Bio-Rad). To ensure adequate representation, sufficient parallel transformations were performed and plated onto carbenicillin containing LB agarose 245 × 245 mm plates (Thermo Fisher) at 492-times the total number of oligonucleotides of each library pool. After overnight growth at 37℃, colonies were scraped off, pelleted, and used for plasmid DNA preps using the endotoxin-free Nucleobond plasmid midiprep kit (TaKaRa Bio; 740422.10). The HIVDEP library was sequenced and contains all 4,191 sgRNAs included in the synthesis (GEO Data set, accession numbers GSM6945645 and GSM6945646).

Virus and lentivirus production. A total of 293 T cells (ATCC; CRL-3216) were plated at 1.5 × 10^5 cells/mL in 6-well plates 1 day prior to transfection. Next, 3 μL of TransIT-LT1 reagent (Mirus Bio LLC) transfection agent was used per μg of DNA. For lentiviral preps, 293Ts were transfected with 667 ng lentiviral plasmid, 500 ng psPAX2, and 333 ng MD2.G. For HIV-1 production, 293Ts were transfected with 1 μg/well proviral DNA.

One day posttransfection, medium was replaced. Two or three days posttransfection lentiviral supernatants of the same type were combined and filtered through a 0.2 μm filter (Thermo Scientific; 720-1320). For lentiviral preps used for the creation of CRISPR/Cas9 knockout lines, supernatants were harvested and clarified from three wells per each lentiviral prep and were concentrated in microcentrifuge tubes for 1 h at 4°C at 16,100 $\times$ g. The volume was reduced to equal 4$\times$ concentration before vortexing vigorously and resuspending at 4°C for 2 days. Each tube was combined per lentiviral stock before using for transduction. For HIVDEP library preps, supernatants from 40$\times$6-well plates were pooled and concentrated by ultracentrifugation as previously described (13). Concentrated lentivirus was used immediately, or aliquots were made and stored at −80°C. To increase infectivity of the CH470 TF stock, we concentrated ∼75 mL of virus to ∼3 mL (25$\times$ concentration) using an Amicon Ultra-15 centrifugal filter unit (Millipore Sigma; UFC905008). All viral and lentiviral infections and transductions were done in the presence of 20 μg/mL DEAE-Dextran (Sigma; D9885).

HIV-CRISPR screening. Prior to the HIV-CRISPR screening or generation of knockout cell lines for validation studies, we performed CRISPR/Cas9-mediated knockout of Zinc Antiviral Protein ZC3HAV1 (ZAP) to increase efficiency of the HIV-CRISPR vector (11). We used gene Knockout v2 kit (GKOv2) for ZAP, including the following sgRNA sequences: GTGGTGTTGGAGACCGG, CCTGGAGCAGCGCGTCC, and TGAAGCAGCACACCTCC (Synthego, Redwood City, CA) complexed with 1 μL of 20 μM Cas9-NLS (UC Berkeley Macro Lab) and single cell sorted into a 96-well U-bottom plate to make clonal knockouts. Clonal KO lines were identified and selected using the ICE editing analysis software (Synthego). An individual clone with biallelic knockouts were used to create both ZsGreen/CCR5 and dTomato/CCR5 stably expressing lines by transduction with pHIV-ZsGreen/CCR5 or pHIV-dTomato/CCR5 lentiviruses followed by cell cytometry sorting for expression of CCR5 via ZsGreen or dTomato fluorescence using the BD FACSCANTO II or Sony MA900 (Fred Hutch Flow Cytometry Core) and analyzed with FlowJo software. These Jurkat cells were transduced with HIV-CRISPR vectors containing the TKOv3, HuEpi, PIKA, or HIVDEP library at an MOI of <1 and selected for puromycin resistance at 0.4 μg/mL. Ten days post puromycin selection, cells were infected with HIV-1 strains and levels of infection were measured by intracellular p24gag staining (Table 1) to obtain at least an MOI of 0.1. To maintain ∼500$\times$ coverage of sgRNAs for the HIV-1$_{Q23BG505}$ screen, which was more difficult to achieve an MOI of 0.1, we increased the number of cells infected proportionally to the lower MOI (Table 1). Genomic DNA and viral RNA was harvested and amplified 3 days postinfection, sequenced, and analyzed using the MAGeCK-FLUTE algorithm (15).

Statistical analysis of HIV-CRISPR screen data. Library pools were demultiplexed, then reads were assigned to libraries to assign sequences to each sample (allowing no mismatches), trimmed, and aligned to the TKOv3 or HIVDEP sgRNA libraries using Bowtie (54). An artificial or "Synthetic" NTC sgRNA set the same size of TKOv3, PIKA, HuEpi, or HIVDEP was created by iteratively binning NTC sgRNA sequences (4 NTC sgRNAs/gene, 8 NTC sgRNAs/gene, 6 NTC sgRNAs/gene, or 8 NTC sgRNAs/gene for each library, respectively). Relative enrichment of sgRNAs and genes were analyzed using the MAGeCK-Flute statistical package (15). Enriched gene ontologies of screen data were determined by GSEA (55) as a part of the MAGeCK-Flute package. For each biological HIVDEP screen replicate, z scores were calculated for each gene for comparison across screens performed with different viruses at different points in time. These were calculated as previously described (21) with additional data preprocessing modifications as can be found here: https://github.com/amcolash/hiv-crispr-zscore-analysis. These z scores and enriched pathway data of the TKOv3 screen were used to generate pathway-focused heatmaps across each HIVDEP screen using Morpheus (https://software.broadinstitute.org/morpheus).

Dependency factor validation through spreading infections and luciferase quantification. Twenty-four candidate genes were knocked out in Jurkat-CCR5 cells using the two most efficient guides per gene from the sgRNA library (as calculated by log$_2$ fold change sgRNA enrichment). We also generated NTC, CD19-, and AAVS1-, and CD4-CRISPR/Cas9 knockout Jurkat-CCR5 cells as negative and positive controls. Knockout cell pools were created via transduction with lentiCRISPRv2 containing gene targeting constructs. Less than 24 h posttransduction, medium was replaced with RPMI containing 0.4 μg/mL puromycin. Transduced cell pools were selected with puromycin for 10 days prior to HIV infection. gDNA was harvested for editing analysis after 10 to 12 days under selection. Knockout cells were maintained as pools rather than individual clones to remove artifacts of clone-to-clone heterogeneity in infection. We measured infection of the knockout pools by both luciferase luminescence at 2 days postinfection and by measuring virus release 3 days after infection by measuring reverse transcriptase activity in viral supernatants as previously described (56). The titer of a stock of HIV-1$_{LAI}$ was determined, then the stock was aliquoted at −80°C, and used as the standard curve in all assays.

Genomic editing analysis. Knockout cells were harvested, and genomic DNA was extracted using the QIAamp DNA blood minikit (Qiagen; 51185). Sites of editing were amplified using primers specific to each targeted locus (Table S3). Primers were designed to amplify a 500- or 1,000-bp amplicon of the targeted locus using either Q5 High-Fidelity DNA polymerase (NEB; M0491S) or Platinum *Taq* DNA polymerase High Fidelity (ThermoFisher Scientific; 11304011). PCR amplicons were sequenced (Fred Hutch Shared Resources Genomics Core – Sanger sequencing) and analyzed by ICE (Synthego) to determine the percentage of alleles edited at each locus in the cell population (57).

Flow cytometry/p24gag analysis. Intracellular p24gag staining was conducted on cells 3 days postinfection to determine viral titer of each stock before using these for CRISPR screens or infection experiments with LAI, LAI-VSV-G, Q23BG505, and CH470TF. Cells were harvested and fixed in 4% paraformaldehyde for 10 min and diluted to 1% in Phosphate Buffered Saline (PBS). Cells were permeabilized in 0.5% Triton-X for 10 min and stained with 1:300 KC57-FITC (Beckman Coulter 6604665; RRID: AB_1575987) or 1:300 KC57-RD1 (Beckman Coulter 6604667; RRID: AB_1575989). Cells were analyzed on a Celesta or Fortessa flow cytometer (Fred Hutch Flow Cytometry Core). To assess infectivity of viral stocks used for the spreading infections, cells were fixed in BD Fix/Perm for 20 min, washed, and permeabilized in BD Perm/wash buffer. The cells were stained with 1:300 KC57-FITC (Beckman Coulter 6604665; RRID: AB_1575987) or 1:300 KC57-RD1 Beckman Coulter 6604667 in BD Perm/wash buffer for 30 min, washed, and then resuspended in PBS. Cells were read

on a Celesta 2 or 3 (Fred Hutch Flow Cytometry Core) and analyzed in FlowJo. For CD4 cell surface marker staining, cells were washed twice in PBS, stained in PBS/1% bovine serum albumin (BSA), incubated at 4°C for 1 h, washed twice in PBS, and resuspended in PBS. The cells were stained with 1:50 CD4-APC (BD Biosciences 555349; AB_398593) and analyzed on Celesta 2 flow cytometer (Fred Hutch Flow Cytometry Core).

Data availability. Sequence data for CRISPR screens generated for this study at the NCBI Gene Expression Omnibus (GEO) at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE223326. All data generated are included in the manuscript and supporting files. The computational pipeline used to analyze the sequencing data for the z score analysis and data processing for generating Fig. 1B and Fig. S1 is available on GitHub (https://github.com/amcolash/hiv-crispr-zscore-analysis).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, PDF file, 0.2 MB.
FIG S2, PDF file, 0.1 MB.
FIG S3, PDF file, 0.3 MB.
TABLE S1, XLSX file, 3.3 MB.
TABLE S2, XLSX file, 0.2 MB.
TABLE S3, XLSX file, 0.3 MB.

REFERENCES

1. Konig R, Zhou Y, Elleder D, Diamond TL, Bonamy GM, Irelan JT, Chiang CY, Tu BP, De Jesus PD, Lilley CE, Seidel S, Opaluch AM, Caldwell JS, Weitzman MD, Kuhen KL, Bandyopadhyay S, Ideker T, Orth AP, Miraglia LJ, Bushman FD, Young JA, Chanda SK. 2008. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. Cell 135:49–60. https://doi.org/10.1016/j.cell.2008.07.032.
2. Zhou H, Xu M, Huang Q, Gates AT, Zhang XD, Castle JC, Stec E, Ferrer M, Strulovici B, Hazuda DJ, Espeseth AS. 2008. Genome-scale RNAi screen for host factors required for HIV replication. Cell Host Microbe 4:495–504. https://doi.org/10.1016/j.chom.2008.10.004.
3. Brass AL, Dykxhoorn DM, Benita Y, Yan N, Engelman A, Xavier RJ, Lieberman J, Elledge SJ. 2008. Identification of host proteins required for HIV infection through a functional genomic screen. Science 319:921–926. https://doi.org/10.1126/science.1152725.
4. Yeung ML, Houzet L, Yedavalli VS, Jeang KT. 2009. A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication. J Biol Chem 284:19463–19473. https://doi.org/10.1074/jbc.M109.010033.
5. Park RJ, Wang T, Koundakjian D, Hultquist JF, Lamothe-Molina P, Monel B, Schumann K, Yu H, Krupzcak KM, Garcia-Beltran W, Piechocka-Trocha A, Krogan NJ, Marson A, Sabatini DM, Lander ES, Hacohen N, Walker BD. 2017. A genome-wide CRISPR screen identifies a restricted set of HIV host dependency factors. Nat Genet 49:193–203. https://doi.org/10.1038/ng.3741.
6. Bushman FD, Malani N, Fernandes J, D'Orso I, Cagney G, Diamond TL, Zhou H, Hazuda DJ, Espeseth AS, Konig R, Bandyopadhyay S, Ideker T, Goff SP, Krogan NJ, Frankel AD, Young JA, Chanda SK. 2009. Host cell factors in HIV replication: meta-analysis of genome-wide studies. PLoS Pathog 5: e1000437. https://doi.org/10.1371/journal.ppat.1000437.
7. Jager S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, Clarke SC, Shales M, Mercenne G, Pache L, Li K, Hernandez H, Jang GM, Roth SL, Akiva E, Marlett J, Stephens M, D'Orso I, Fernandes J, Fahey M, Mahon C, O'Donoghue AJ, Todorovic A, Morris JH, Maltby DA, Alber T, Cagney G, Bushman FD, Young JA, Chanda SK, Sundquist WI, Kortemme T, Hernandez RD, Craik CS, Burlingame A, Sali A, Frankel AD, Krogan NJ. 2011. Global landscape of HIV-human protein complexes. Nature 481:365–370.
8. Hiatt J, Hultquist JF, McGregor MJ, Bouhaddou M, Leenay RT, Simons LM, Young JM, Haas P, Roth TL, Tobin V, Wojcechowskyj JA, Woo JM, Rathore U, Cavero DA, Shifrut E, Nguyen TT, Haas KM, Malik HS, Doudna JA, May AP, Marson A, Krogan NJ. 2022. A functional map of HIV-host interactions in primary human T cells. Nat Commun 13:1752. https://doi.org/10.1038/s41467-022-29346-w.
9. Fu C, Yang S, Yang X, Lian X, Huang Y, Dong X, Zhang Z. 2020. Human gene functional network-informed prediction of HIV-1 host dependency factors. mSystems 5. https://doi.org/10.1128/mSystems.00960-20.
10. Gordon DE, Watson A, Roguev A, Zheng S, Jang GM, Kane J, Xu J, Guo JZ, Stevenson E, Swaney DL, Franks-Skiba K, Verschueren E, Shales M, Crosby DC, Frankel AD, Marson A, Marazzi I, Cagney G, Krogan NJ. 2020. A quantitative genetic interaction map of HIV infection. Mol Cell 78:197–209.e7. https://doi.org/10.1016/j.molcel.2020.02.004.
11. OhAinle M, Helms L, Vermeire J, Roesch F, Humes D, Basom R, Delrow JJ, Overbaugh J, Emerman M. 2018. A virus-packageable CRISPR screen identifies host factors mediating interferon inhibition of HIV. Elife 7. https://doi.org/10.7554/eLife.39823.

12. Ohainle M, Kim K, Komurlu Keceli S, Felton A, Campbell E, Luban J, Emerman M. 2020. TRIM34 restricts HIV-1 and SIV capsids in a TRIM5alpha-dependent manner. PLoS Pathog 16:e1008507. https://doi.org/10.1371/journal.ppat.1008507.

13. Roesch F, OhAinle M. 2020. HIV-CRISPR: a CRISPR/Cas9 screening method to identify genes affecting HIV replication. Bio Protoc 10:e3614. https://doi.org/10.21769/BioProtoc.3614.

14. Hart T, Tong AHY, Chan K, Van Leeuwen J, Seetharaman A, Aregger M, Chandrashekhar M, Hustedt N, Seth S, Noonan A, Habsid A, Sizova O, Nedyalkova L, Climie R, Tworzyanski L, Lawson K, Sartori MA, Alibeh S, Tieu D, Masud S, Mero P, Weiss A, Brown KR, Usaj M, Billmann M, Rahman M, Constanzo M, Myers CL, Andrews BJ, Boone C, Durocher D, Moffat J. 2017. Evaluation and design of genome-wide CRISPR/SpCas9 knockout screens. G3 (Bethesda) 7:2719–2727. https://doi.org/10.1534/g3.117.041277.

15. Wang B, Wang M, Zhang W, Xiao T, Chen CH, Wu A, Wu F, Traugh N, Wang X, Li Z, Mei S, Cui Y, Shi S, Lipp JJ, Hinterndorfer M, Zuber J, Brown M, Li W, Liu XS. 2019. Integrative analysis of pooled CRISPR genetic screens using MAGeCKFlute. Nat Protoc 14:756–780. https://doi.org/10.1038/s41596-018-0113-7.

16. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, Irizarry RA, Liu JS, Brown M, Liu XS. 2014. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol 15:554. https://doi.org/10.1186/s13059-014-0554-4.

17. Hsieh E, Janssens DH, Paddison PJ, Browne EP, Henikoff S, OhAinle M, Emerman M. 2022. A modular CRISPR screen identifies individual and combination pathways contributing to HIV-1 latency. PLoS Pathog 19:e1011101. https://doi.org/10.1371/journal.ppat.1011101.

18. Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, Lander ES, Sabatini DM. 2015. Identification and characterization of essential genes in the human genome. Science 350:1096–1101. https://doi.org/10.1126/science.aac7041.

19. Poss M, Overbaugh J. 1999. Variants from the diverse virus population identified at seroconversion of a clade A human immunodeficiency virus type 1-infected woman have distinct biological properties. J Virol 73:5255–5264. https://doi.org/10.1128/JVI.73.7.5255-5264.1999.

20. Parrish NF, Gao F, Li H, Giorgi EE, Barbian HJ, Parrish EH, Zajic L, Iyer SS, Decker JM, Kumar A, Hora B, Berg A, Cai F, Hopper J, Denny TN, Ding H, Ochsenbauer C, Kappes JC, Galimidi RP, West AP, Jr, Bjorkman PJ, Wilen CB, Doms RW, O'Brien M, Bhardwaj N, Borrow P, Haynes BF, Muldoon M, Theiler JP, Korber B, Shaw GM, Hahn BH. 2013. Phenotypic properties of transmitted founder HIV-1. Proc Natl Acad Sci U S A 110:6626–6633. https://doi.org/10.1073/pnas.1304288110.

21. Schneider WM, Luna JM, Hoffmann HH, Sanchez-Rivera FJ, Leal AA, Ashbrook AW, Le Pen J, Ricardo-Lax I, Michailidis E, Peace A, Stenzel AF, Lowe SW, MacDonald MR, Rice CM, Poirier JT. 2021. Genome-scale identification of SARS-CoV-2 and pan-coronavirus host factor networks. Cell 184:120–132.e14. https://doi.org/10.1016/j.cell.2020.12.006.

22. Vasiliver-Shamis G, Cho MW, Hioe CE, Dustin ML. 2009. Human immunodeficiency virus type 1 envelope gp120-induced partial T-cell receptor signaling creates an F-actin-depleted zone in the virological synapse. J Virol 83:11341–11355. https://doi.org/10.1128/JVI.01440-09.

23. Smith JR, Maguire S, Davis LA, Alexander M, Yang F, Chandran S, Ffrench-Constant C, Pedersen RA. 2008. Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. Stem Cells 26:496–504. https://doi.org/10.1634/stemcells.2007-0039.

24. Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKelver RC, Katibah GE, Amora R, Boydston EA, Zeitler B, Meng X, Miller JC, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R. 2009. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol 27:851–857. https://doi.org/10.1038/nbt.1562.

25. Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, Kuhn R. 2015. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotechnol 33:543–548. https://doi.org/10.1038/nbt.3198.

26. Llano M, Saenz DT, Meehan A, Wongthida P, Peretz M, Walker WH, Teo W, Poeschla EM. 2006. An essential role for LEDGF/p75 in HIV integration. Science 314:461–464. https://doi.org/10.1126/science.1132319.

27. Liu R, Chen C, Li Y, Huang Q, Xue Y. 2020. ELL-associated factors EAF1/2 negatively regulate HIV-1 transcription through inhibition of Super Elongation Complex formation. Biochim Biophys Acta Gene Regul Mech 1863:194508. https://doi.org/10.1016/j.bbagrm.2020.194508.

28. He N, Liu M, Hsu J, Xue Y, Chou S, Burlingame A, Krogan NJ, Alber T, Zhou Q. 2010. HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. Mol Cell 38:428–438. https://doi.org/10.1016/j.molcel.2010.04.013.

29. Marzio G, Tyagi M, Gutierrez MI, Giacca M. 1998. HIV-1 tat transactivator recruits p300 and CREB-binding protein histone acetyltransferases to the viral promoter. Proc Natl Acad Sci U S A 95:13519–13524. https://doi.org/10.1073/pnas.95.23.13519.

30. Deng L, de la Fuente C, Fu P, Wang L, Donnelly R, Wade JD, Lambert P, Li H, Lee CG, Kashanchi F. 2000. Acetylation of HIV-1 Tat by CBP/P300 increases transcription of integrated HIV-1 genome and enhances binding to core histones. Virology 277:278–295. https://doi.org/10.1006/viro.2000.0593.

31. Ott M, Schnolzer M, Garnica J, Fischle W, Emiliani S, Rackwitz HR, Verdin E. 1999. Acetylation of the HIV-1 Tat protein by p300 is important for its transcriptional activity. Curr Biol 9:1489–1492. https://doi.org/10.1016/s0960-9822(00)80120-7.

32. Gupta S, Singh AK, Prajapati KS, Kushwaha PP, Shuaib M, Kumar S. 2020. Emerging role of ZBTB7A as an oncogenic driver and transcriptional repressor. Cancer Lett 483:22–34. https://doi.org/10.1016/j.canlet.2020.04.015.

33. Pendergrast PS, Wang C, Hernandez N, Huang S. 2002. FBI-1 can stimulate HIV-1 Tat activity and is targeted to a novel subnuclear domain that includes the Tat-P-TEFb-containing nuclear speckles. Mol Biol Cell 13:915–929. https://doi.org/10.1091/mbc.01-08-0383.

34. Pessler F, Hernandez N. 2003. Flexible DNA binding of the BTB/POZ-domain protein FBI-1. J Biol Chem 278:29327–29335. https://doi.org/10.1074/jbc.M302980200.

35. Zhu X, Trimarco JD, Williams CA, Barrera A, Reddy TE, Heaton NS. 2022. ZBTB7A promotes virus-host homeostasis during human coronavirus 229E infection. Cell Rep 41:111540. https://doi.org/10.1016/j.celrep.2022.111540.

36. Wang LH, Aberin MAE, Wu S, Wang SP. 2021. The MLL3/4 H3K4 methyltransferase complex in establishing an active enhancer landscape. Biochem Soc Trans 49:1041–1054. https://doi.org/10.1042/BST20191164.

37. Liu L, Oliveira NM, Cheney KM, Pade C, Dreja H, Bergin AM, Borgdorff V, Beach DH, Bishop CL, Dittmar MT, McKnight A. 2011. A whole genome screen for HIV restriction factors. Retrovirology 8:94. https://doi.org/10.1186/1742-4690-8-94.

38. Van Duyne R, Easley R, Wu W, Berro R, Pedati C, Klase Z, Kehn-Hall K, Flynn EK, Symer DE, Kashanchi F. 2008. Lysine methylation of HIV-1 Tat regulates transcriptional activity of the viral LTR. Retrovirology 5:40. https://doi.org/10.1186/1742-4690-5-40.

39. Liu Z, Liu J, Ebrahimi B, Pratap UP, He Y, Altwegg KA, Tang W, Li X, Lai Z, Chen Y, Shen L, Sareddy GR, Viswanadhapalli S, Tekmal RR, Rao MK, Vadlamudi RK. 2022. SETDB1 interactions with PELP1 contributes to breast cancer endocrine therapy resistance. Breast Cancer Res 24:26. https://doi.org/10.1186/s13058-022-01520-4.

40. Soucy TA, Smith PG, Milhollen MA, Berger AJ, Gavin JM, Adhikari S, Brownell JE, Burke KE, Cardin DP, Critchley S, Cullis CA, Doucette A, Garnsey JJ, Gaulin JL, Gershman RE, Lublinsky AR, McDonald A, Mizutani H, Narayanan U, Olhava EJ, Peluso S, Rezaei M, Sintchak MD, Talreja T, Thomas MP, Traore T, Vyskocil S, Weatherhead GS, Yu J, Zhang J, Dick LR, Claiborne CF, Rolfe M, Bolen JB, Langston SP. 2009. An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. Nature 458:732–736. https://doi.org/10.1038/nature07884.

41. Nekorchuk MD, Sharifi HJ, Furuya AK, Jellinger R, de Noronha CM. 2013. HIV relies on neddylation for ubiquitin ligase-mediated functions. Retrovirology 10:138. https://doi.org/10.1186/1742-4690-10-138.

42. Walden H, Podgorski MS, Huang DT, Miller DW, Howard RJ, Minor DL, Jr, Holton JM, Schulman BA. 2003. The structure of the APPBP1-UBA3-NEDD8-ATP complex reveals the basis for selective ubiquitin-like protein activation by an E1. Mol Cell 12:1427–1437. https://doi.org/10.1016/s1097-2765(03)00452-0.

43. Thirimanne HN, Wu F, Janssens DH, Swanger J, Diab A, Feldman HM, Amezquita RA, Gottardo R, Paddison PJ, Henikoff S, Clurman BE. 2022. Global and context-specific transcriptional consequences of oncogenic Fbw7 mutations. Elife 11. https://doi.org/10.7554/eLife.74338.

44. Strohmaier H, Spruck CH, Kaiser P, Won KA, Sangfelt O, Reed SI. 2001. Human F-box protein hCdc4 targets cyclin E for proteolysis and is mutated in a breast cancer cell line. Nature 413:316–322. https://doi.org/10.1038/35095076.

45. Ivanov A, Lin X, Ammosova T, Ilatovskiy AV, Kumari N, Lassiter H, Afangbedji N, Niu X, Petukhov MG, Nekhai S. 2018. HIV-1 Tat phosphorylation on Ser-16 residue modulates HIV-1 transcription. Retrovirology 15:39. https://doi.org/10.1186/s12977-018-0422-5.

46. Deng L, Ammosova T, Pumfery A, Kashanchi F, Nekhai S. 2002. HIV-1 Tat interaction with RNA polymerase II C-terminal domain (CTD) and a dynamic association with CDK2 induce CTD phosphorylation and transcription from HIV-1 promoter. J Biol Chem 277:33922–33929. https://doi.org/10.1074/jbc.M111349200.

47. Nekhai S, Zhou M, Fernandez A, Lane WS, Lamb NJ, Brady J, Kumar A. 2002. HIV-1 Tat-associated RNA polymerase C-terminal domain kinase, CDK2, phosphorylates CDK7 and stimulates Tat-mediated transcription. Biochem J 364:649–657. https://doi.org/10.1042/BJ20011191.

48. Gummuluru S, Kinsey CM, Emerman M. 2000. An in vitro rapid-turnover assay for human immunodeficiency virus type 1 replication selects for cell-to-cell spread of virus. J Virol 74:10882–10891. https://doi.org/10.1128/JVI.74.23.10882-10891.2000.

49. Bartz SR, Vodicka MA. 1997. Production of high-titer human immunodeficiency virus type 1 pseudotyped with vesicular stomatitis virus glycoprotein. Methods 12:337–342. https://doi.org/10.1006/meth.1997.0487.

50. Haddox HK, Dingens AS, Hilton SK, Overbaugh J, Bloom JD. 2018. Mapping mutational effects along the evolutionary landscape of HIV envelope. Elife 7:e34420. https://doi.org/10.7554/eLife.34420.

51. Fenton-May AE, Dibben O, Emmerich T, Ding H, Pfafferott K, Aasa-Chapman MM, Pellegrino P, Williams I, Cohen MS, Gao F, Shaw GM, Hahn BH, Ochsenbauer C, Kappes JC, Borrow P. 2013. Relative resistance of HIV-1 founder viruses to control by interferon-alpha. Retrovirology 10:146. https://doi.org/10.1186/1742-4690-10-146.

52. Meier JA, Zhang F, Sanjana NE. 2017. GUIDES: sgRNA design for loss-of-function screens. Nat Methods 14:831–832. https://doi.org/10.1038/nmeth.4423.

53. Labun K, Montague TG, Krause M, Torres Cleuren YN, Tjeldnes H, Valen E. 2019. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. Nucleic Acids Res 47:W171–W174. https://doi.org/10.1093/nar/gkz365.

54. Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25. https://doi.org/10.1186/gb-2009-10-3-r25.

55. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102:15545–15550. https://doi.org/10.1073/pnas.0506580102.

56. Vermeire J, Naessens E, Vanderstraeten H, Landi A, Iannucci V, Van Nuffel A, Taghon T, Pizzato M, Verhasselt B. 2012. Quantification of reverse transcriptase activity by real-time PCR as a fast and accurate method for titration of HIV, lenti- and retroviral vectors. PLoS One 7:e50859. https://doi.org/10.1371/journal.pone.0050859.

57. Conant D, Hsiau T, Rossi N, Oki J, Maures T, Waite K, Yang J, Joshi S, Kelso R, Holden K, Enzmann BL, Stoner R. 2022. Inference of CRISPR Edits from Sanger Trace Data. Crispr J 5:123–130. https://doi.org/10.1089/crispr.2021.0113.