CRNKL1 Is a Highly Selective Regulator of Intron-Retaining HIV-1 and Cellular mRNAs

Han Xiao,[a] Emanuel Wyler,[b] Miha Milek,[b] Bastian Grewe,[c] Philipp Kirchner,[d] Arif Ekici,[d] Ana Beatriz Oliveira Villela Silva,[a] Doris Jungnickl,[a] Florian Full,[a] Marco Thomas,[a] Markus Landthaler,[b,e] Armin Ensser,[a] Klaus Überla[a]

[a]Institute of Clinical and Molecular Virology, University Hospital Erlangen, Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany
[b]Berlin Institute for Medical Systems Biology, Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany
[c]Department of Molecular and Medical Virology, Ruhr-University, Bochum, Germany
[d]Institute of Human Genetics, University Hospital Erlangen, Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany
[e]IRI Life Sciences, Institute für Biologie, Humboldt Universität zu Berlin, Berlin, Germany

Emanuel Wyler and Miha Milek contributed equally to this article.

ABSTRACT The HIV-1 Rev protein is a nuclear export factor for unspliced and incompletely spliced HIV-1 RNAs. Without Rev, these intron-retaining RNAs are trapped in the nucleus. A genome-wide screen identified nine proteins of the spliceosome, which all enhanced expression from the HIV-1 unspliced RNA after CRISPR/Cas knockdown. Depletion of DHX38, WDR70, and four proteins of the Prp19-associated complex (ISY1, BUD31, XAB2, and CRNKL1) resulted in a more than 20-fold enhancement of unspliced HIV-1 RNA levels in the cytoplasm. Targeting of CRNKL1, DHX38, and BUD31 affected nuclear export efficiencies of the HIV-1 unspliced RNA to a much larger extent than splicing. Transcriptomic analyses further revealed that CRNKL1 also suppresses cytoplasmic levels of a subset of cellular mRNAs, including some with selectively retained introns. Thus, CRNKL1-dependent nuclear retention is a novel cellular mechanism for the regulation of cytoplasmic levels of intron-retaining HIV-1 mRNAs, which HIV-1 may have harnessed to direct its complex splicing pattern.

IMPORTANCE To regulate its complex splicing pattern, HIV-1 uses the adaptor protein Rev to shuttle unspliced or partially spliced mRNA from the nucleus to the cytoplasm. In the absence of Rev, these RNAs are retained in the nucleus, but it is unclear why. Here we identify cellular proteins whose depletion enhances cytoplasmic levels of the HIV-1 unspliced RNA. Depletion of one of them, CRNKL1, also increases cytoplasmic levels of a subset of intron-retaining cellular mRNA, suggesting that CRNKL1-dependent nuclear retention may be a basic cellular mechanism exploited by HIV-1.

KEYWORDS CRNKL1, RNA splicing, association, human immunodeficiency virus, intron-retaining RNA, nuclear retention

Intron retention is increasingly recognized as a regulatory mechanism of gene expression and diversification of the cellular proteome (1). Transcriptomic analyses indicate that a large percentage of eukaryotic genes encode transcripts retaining introns (2, 3). Intron-retaining transcripts may be trapped in the nucleus, leading to their degradation. Alternatively, specialized nuclear export pathways may enable certain intron-retaining transcripts to reach the cytoplasm, where they can be translated or undergo nonsense-mediated decay. Retroviruses use intron retention to encode multiple proteins from one primary transcript. RNA secondary structures or cellular or viral adaptor proteins link intron-retaining retroviral transcripts in the nucleus to the TAP/NXF1 or the CRM-1 nuclear export factors (4), leading to the export of intron-retaining RNAs.

While a number of studies have explored how nuclear retention can be overcome, the factors mediating nuclear retention of intron-containing mRNAs are mostly unknown but are likely linked to partial assembly of splicing complexes on intron-containing RNA (5).

The first evidence for a regulated export and expression of intron-retaining RNAs in mammalian cells was obtained in studies investigating the replication of human immunodeficiency virus type 1 (HIV-1). Alternative splicing and intron retention leads to three types of transcripts: fully spliced (FS) mRNAs, incompletely spliced (IS) mRNAs, and the unspliced (US) genomic RNA (6). The process of viral gene expression is temporally regulated, resulting in early expression of regulatory proteins and late expression of structural proteins (7–9). In the early phase, the US and IS transcripts are retained in the nucleus and are subjected to the nuclear splicing or degradation machinery (10). Only FS transcripts are transported to the cytoplasm via the default mRNA export pathway, TAP/NXF1, leading to expression of Tat, Rev, and Nef proteins (11, 12). The regulatory proteins Tat and Rev are nuclear cytoplasmic shuttle proteins and regulate viral gene expression. While Tat transactivates transcription by enhancing RNA elongation, Rev fine-tunes posttranscriptional processes and acts as a switch to initiate the late phase of the viral replication cycle (reviewed in references 13 and 14). It binds to the Rev-responsive element (RRE), which is present in the HIV-1 US and IS transcripts but not in the FS transcripts, and multimerizes on the RRE. The protein exporting factor CRM1 and associated factors are recruited by the Rev-RRE complex, leading to the nuclear-cytoplasmic export of these two classes of RNA (15).

Nuclear retention of intron-retaining HIV-1 RNAs has been hypothesized to be due to a number of mechanisms that are not mutually exclusive. One of them is the lack of cis-acting elements necessary for access to the TAP/NXF1 nuclear export pathway. Cell biological and virological studies have discovered that RNAs normally gain access to the exportin NXF1 by two means: either by recruitment of the transcription-export (TREX) complex via splicing (16) or by evolving particular regulatory sequences, e.g., the constitutive transport element (CTE) from type D retrovirus Mason-Pfizer monkey virus (MPMV), which has direct binding activity to NXF1 (17). A second explanation is the suboptimal splice sites of HIV that lead to the trapping of the intron-retaining RNAs in the spliceosome and subsequent degradation (18, 19). In addition, multiple inhibitory sequences (INS) were identified in the retained introns (20–24). The presumed inactivation of these inhibitory sequences by codon optimizing the AT-rich HIV-1 sequence abolished nuclear retention (25) independent of the presence or absence of HIV-1 splice donor (SD) and splice acceptor (SA) sites (26, 27). Why AT-rich intron-containing HIV-1 sequences, but not the codon-optimized variants, are retained remains elusive. It has been reported that knockdown of the paraspeckle long-noncoding RNA (lncRNA) NEAT1 could promote HIV Gag production through increased nucleocytoplasmic export of INS-containing RNAs, implicating the nuclear paraspeckles in the nuclear retention of HIV US and IS RNAs (28). However, this has been questioned by a study demonstrating that HIV-1 US RNA was not actually colocalized with paraspeckles (29), despite that several paraspeckle proteins (PSF, Martin3, and RBM14) have been reported to associate with HIV-1 INS elements and/or promote the Rev-dependent RNA export (30–34, 98).

In addition to these viral determinants and lncRNA contributing to nuclear mRNA retention, depletion of hnRNPA2/B1 has been shown to increase cytoplasmic levels of HIV-1 genomic RNA, but this increase did not enhance Gag expression levels and the precise mechanism remains unknown (35). With regard to nuclear retention of mRNA in general, a cellular complex named pre-mRNA retention and splicing complex (RES) has been identified in Saccharomyces cerevisiae (36). The RES is composed of Bud13, Snu17p, and Pml1P, and inactivation of any of the genes encoding these proteins in yeast reduces splicing of an intron with a weak 5′ splice donor. In addition, Pml1P inactivation enhanced expression of a pre-mRNA reporter construct without substantial reduction of the spliced mRNA, suggesting a role of Pml1P and the RES in nuclear

retention. More recently, BUD13 was identified to bind to the poorly spliced mRNA of mammalian IRF7, a master regulator of the interferon response (37). Knockdown of BUD13 specifically enhanced retention of intron 4 of IRF7 in stimulated cells, thus reducing IRF7 protein levels and dampening the interferon response. However, intron-containing IRF7 mRNA could not be detected in the cytoplasm (37), indicating additional nuclear trapping mechanisms or rapid cytoplasmic degradation of intron-containing IRF7 mRNA. HIV-1 introns also differ from typical RES-dependent introns (37–39) by their low GC content, challenging the hypothesis that HIV-1 splicing may be regulated by the RES.

An alternative nuclear trapping mechanism could be a cellular quality control mechanism for mRNAs located at the nuclear pore complex (NPC). Experimental or stress-induced interference with this complex results in leakage of intron-retaining mRNAs into the cytoplasm (40–43). This finding suggests that the NPC-associated quality control complex also has a role in nuclear retention of intron-containing mRNAs, probably as a backup mechanism not influencing the initial splicing events. Here, we identified cellular proteins involved in nuclear trapping of intron-retaining HIV-1 mRNAs by performing a genome-wide screen for cellular factors, whose depletion resulted in enhanced expression from HIV-1 unspliced transcripts in the absence of its export factor Rev. Furthermore, we examined the effect of identified cellular factors on HIV US and FS RNA localization and expression. This revealed that CRNKL1 is a nuclear retention factor of the HIV-1 unspliced RNA and a selective regulator of cytoplasmic levels of a subset of intron-retaining cellular mRNAs.

RESULTS

Reporter cell lines for HIV-1 Rev-independent expression from the unspliced RNA. The overall strategy to identify cellular proteins involved in the nuclear trapping of intron-retaining HIV-1 mRNAs was to inactivate cellular genes by a genome-wide lentiviral CRISPR-Cas knockout library and to select knockout cells that show enhanced Gag expression from a proviral Rev-deficient HIV-1 reporter construct. To easily quantify HIV Gag expression in living cells, the Gag open reading frame (ORF) was fused to the blue fluorescent protein (BFP) gene. The *rev* gene was inactivated by two point mutations, and the DsRed reporter gene was expressed in place of *nef* from a fully spliced HIV-1 transcript in order to mark HIV-infected cells in the absence of Rev (Fig. 1A). Since we aimed to establish stable cell clones containing such proviral reporter constructs, expression of *pol*, *vif*, *vpr*, and *vpu* was also blocked by point mutations to avoid difficulties due to potential cytotoxic or cytostatic effects of these viral proteins (44–46). The *env* gene was inactivated by a 4-bp deletion, resulting in a frameshift. Although this reporter construct, designated HIV-dual-GT encodes only the G̲ag-BPF fusion protein, DsRed, and T̲at (for sufficient transcriptional activation), all known *cis*-acting sequences potentially interacting with cellular components should be maintained.

To characterize the expression pattern of HIV-dual-GT, HEK293T cells were transfected with plasmids containing HIV-dual-GT or a *rev*-deficient *env* deletion mutant of HIV in the presence or absence of a Rev expression plasmid. Western blot analyses revealed Rev-dependent expression of the Gag-BFP fusion protein (Fig. 1B). Flow cytometric analysis of HEK293T cells transfected with HIV-dual-GT revealed a major DsRed-positive but BFP-negative cell population, which shifted to a double-positive population by cotransfection of a Rev expression plasmid (Fig. 1C). DsRed expression seemed to decrease in the presence of Rev, consistent with reduced splicing and consequently reduced generation of fully-spliced DsRed mRNA due to rapid export of unspliced transcripts (7).

To generate reporter cell lines harboring a properly integrated proviral DNA of HIV-dual-GT, it was encapsidated and pseudotyped by cotransfection with HIV-gag-pol and VSV-G expression plasmids. The Rev deficiency was also complemented by cotransfection of a Rev expression plasmid. Titers of the HIV-dual-GT vector were in the range of 5×10^5 transforming units (TU)/ml. Flow cytometric analysis of Jurkat cells transduced with these VSV-G-pseudotyped HIV-dual-GT vector particles revealed a distinct
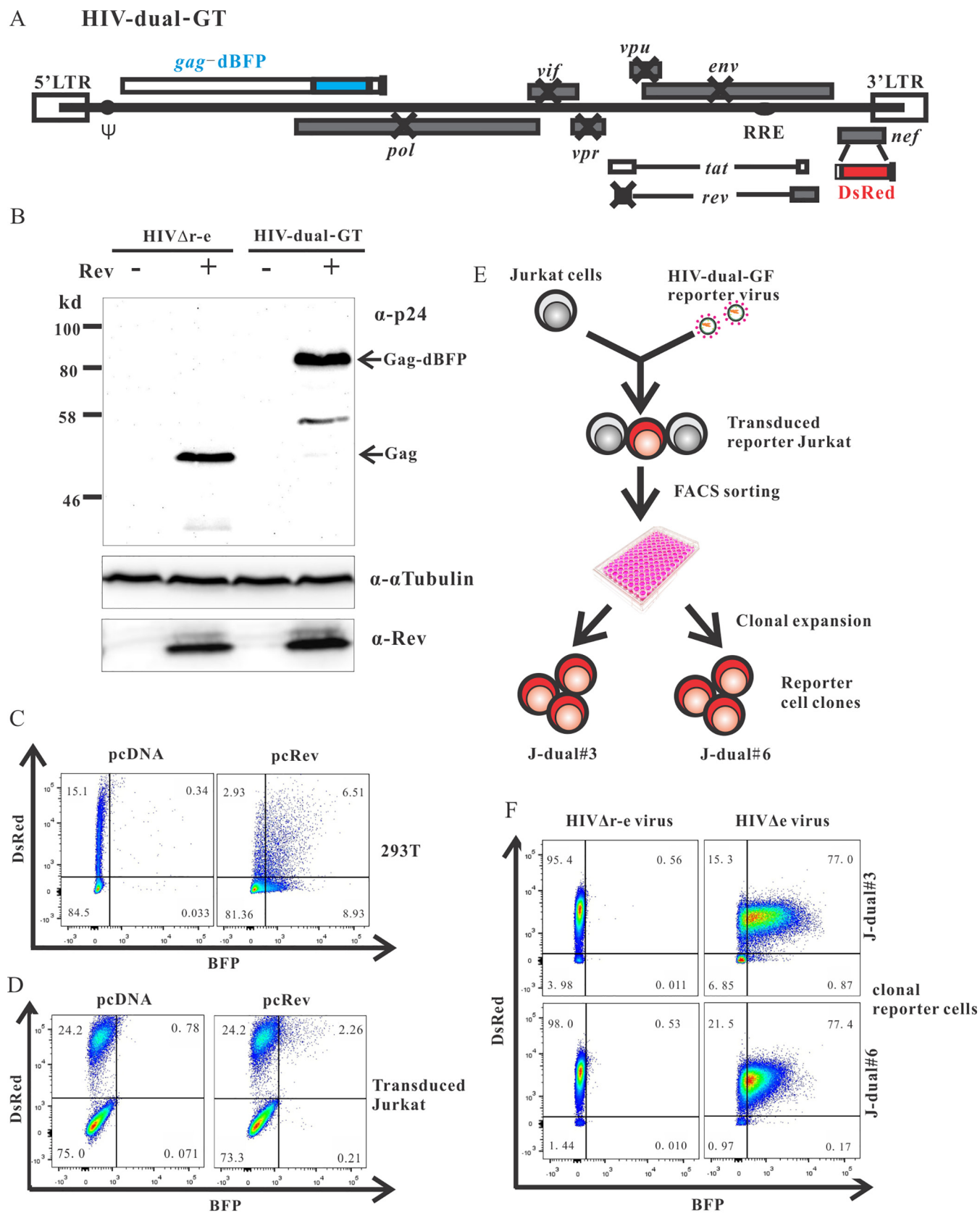
FIG 1 Establishment of HIV-dual-GT reporter cell clones. (A) Map of the HIV-dual-GT proviral reporter construct. (B) Western blot analysis of lysates of HEK293T cells transfected with HIV-dual-GT or HIV.r⁻e⁻ in the presence or absence of a Rev expression plasmid (pcRev) with antibodies to the p24 capsid

(Continued on next page)

population of cells that were DsRed positive but BFP negative (Fig. 1D). Transient transfection of this transduced bulk culture with a Rev expression plasmid shifted approximately 10% of the DsRed-positive population to a DsRed and BFP double-positive population. The magnitude of this effect was lower than in the cotransfection experiments in 293T cells (Fig. 1C), consistent with the lower transfection efficiency for Jurkat cells.

To obtain more homogenous reporter cells, Jurkat cell clones containing HIV-dual-GT were established by transducing Jurkat cells at a multiplicity of infection (MOI) of 0.1, followed by flow cytometric sorting of single DsRed-positive, BFP-negative cells into wells of a 96-well plate (Fig. 1E). Two of the expanded DsRed-positive cell clones, showing a rather homogenous population of DsRed-positive and BPF-negative cells and designated J-dual#3 and J-dual#6, were transduced at an MOI of 2 with VSV-G-peudotyped env deletion mutants of HIV containing or lacking a functional rev gene. In the presence of a functional rev gene, the BFP fluorescence was strongly induced in both reporter cell clones (Fig. 1F). Therefore, both clones were selected for the screening of cellular genes that suppress BFP expression from the HIV-dual-GT reporter virus.

Genome-wide screening for cellular factors suppressing expression from intron-containing HIV-1 genomic RNA. Poor expression levels of HIV-1 structural genes in the absence of Rev that can be overcome by modifying the codon usage suggest that there is an active cellular suppressive mechanism acting on the level of the viral transcript. To identify cellular proteins necessary for this postulated suppressive mechanism, we used a GeCKOv2 genome-wide CRISPR-Cas knockout library kindly provided by the laboratory of Feng Zhang (47, 48). The plasmid library contains 122,411 different single guide RNA (sgRNA) genes targeting 19,050 human genes (6 sgRNAs per gene) and 1,864 sgRNA genes targeting microRNA (miRNA) genes (4 sgRNAs per miRNA gene) and includes 1,000 nontargeting (NT) sgRNA genes. Cas9, sgRNA, and puromycin resistance genes are expressed from a single lentiviral vector construct, lentiCRISPRv2. After amplification of the plasmid library and confirmation of its diversity by next-generation sequencing (NGS; data not shown), a stock of a VSV-G-pseudotyped lentiviral vector library was prepared by transient transfection of 293T cells and titrated on the Jurkat reporter cell clones, J-dual#3 and J-dual#6, and parental Jurkat cells. All three cell lines were then transduced with the lentiviral CRISPR library at an MOI of around 2. Four days after transduction, the 0.01 to 0.02% DsRed-positive cells showing the highest BFP expression levels were sorted for NGS of the sgRNA genes delivered to these cells by the lentiviral vector library. To be able to determine enrichment of sgRNA genes enhancing Gag-BFP expression levels, the representation of each sgRNA gene from nonselected Jurkat cells transduced with the lentiviral CRISPR library was also determined. Selective enrichment of sgRNA sequences in sorted J-dual#3 and J-dual#6 cells compared to nonsorted Jurkat cells was then determined by MAGeCK, a computational tool developed for the analysis of CRISPR screens (49, 50). The calculated robust ranking aggregation (RRA) score reflects enrichment of sgRNA sequences targeting the same gene in the selected cells. The distribution of the RRA score for all targeted genes is plotted in Fig. 2C, with the top 12 candidate genes highlighted. For 11 of these, at least two different sgRNA sequences targeting the same gene were enriched in the selected cells, arguing against confounding off-target effects of single sgRNA (Fig. 2D). The probability of false discovery of the top 12 candidate genes ranged from 0.0012 to 0.15.

Analyzing interaction of validated hits shows a network of pre-mRNA splicing factors. Based on the enrichment of the individual sgRNAs targeting each of the top 12 candidate genes in the Gag-BFP-positive cells selected by flow cytometry (see Fig. S1 in the supplemental material), two sgRNA sequences were chosen for each of the candidate genes (Table S1A) for further validation. Gene-specific sgRNA sequences and a

FIG 1 Legend (Continued)
protein, α-tubulin, or Rev. (C) 293T cells transfected with HIV-dual-GT and pcRev or pcDNA3.1 were analyzed by flow cytometry for Gag-BFP and DsRed expression. (D) Jurkat cells transduced with the HIV-dual-GT reporter vector (MOI, 0.3) and transiently transfected with pcRev or pcDNA control plasmid were analyzed by flow cytometry. (E) Scheme for the generation of reporter cell clones. (F) HIV-dual-GT-transduced Jurkat reporter cell clones J-dual#3 and J-dual#6 were superinfected with VSV-G-pseudotyped HIV-1 viruses differing only in the expression of Rev and analyzed by flow cytometry. All assays were performed at least twice, and results of one representative experiment are shown.
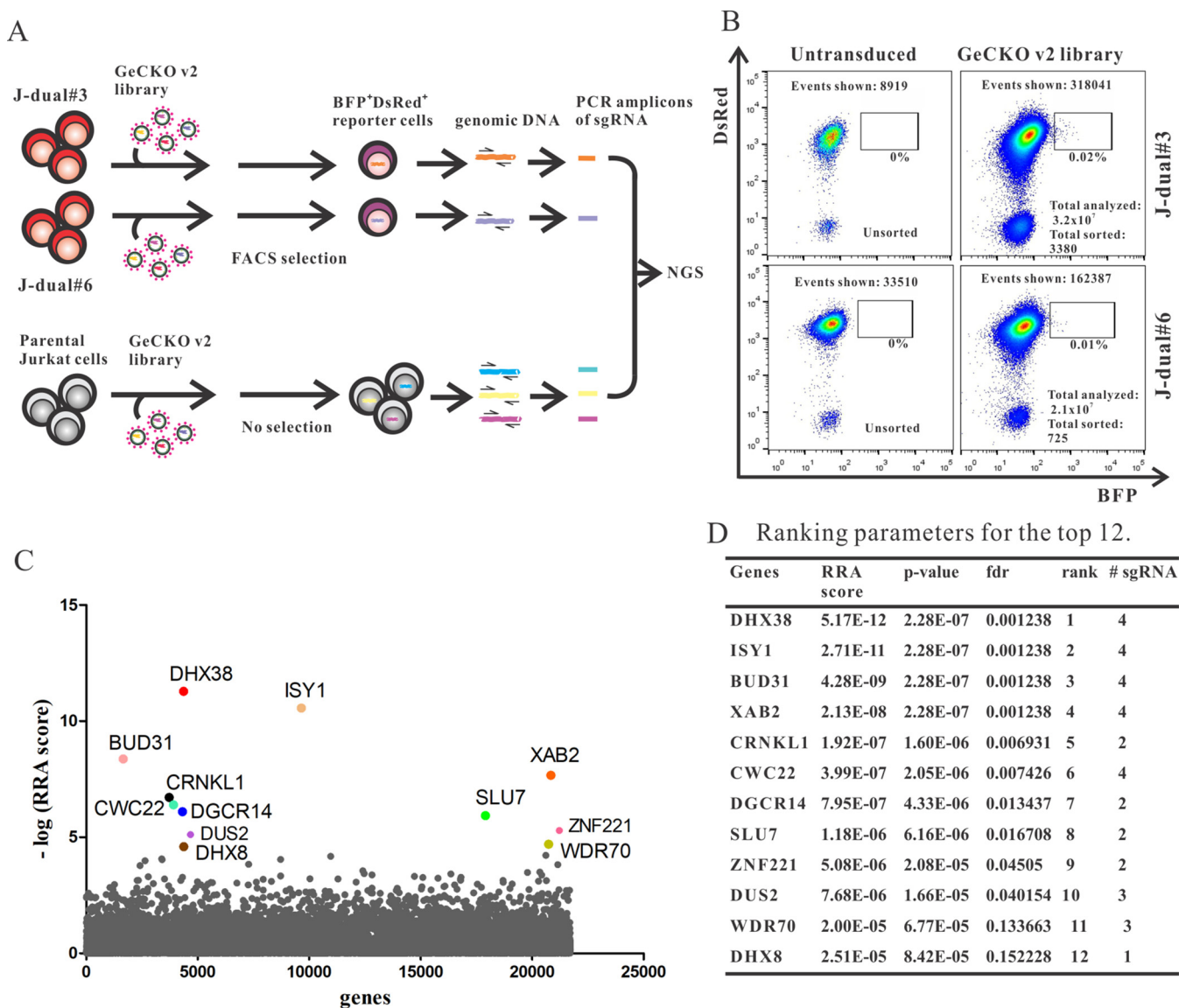
FIG 2 Genome-wide CRISPR/Cas screen. (A) Outline of screening strategy. (B) Flow cytometric analysis of J-dual#3 and J-dual#6 reporter cells without transduction and 4 days after transduction by the GeCKO v2 lentiviral vector library (MOI, around 2) immediately prior to the sorting process. The total number of events shown in the graph, total number of cells analyzed in the screen, and total numbers of cells sorted are indicated. (C) Enrichment analysis for guide RNAs in transduced and selected J-dual#3 and J-dual#6 reporter cells compared to transduced but unselected Jurkat cells by the MAGeCK program. The robust ranking aggregation (RRA) score reflects the enrichment of guide RNA targeting the same gene in cells selected from both reporter cell clones. The distribution of the negative $\log_{10}$ of the RRA score for all targeted genes in alphabetical order on the x axis is shown, with the top 12 candidate genes highlighted. (D) Additional MAGeCK ranking parameters of the top 12 candidate genes. fdr, false discovery rate; # sgRNA, number of enriched guide RNAs targeting the same gene.

nontargeting sequence, named NT1, were cloned individually into the lentiCRISPRv2 vector plasmid. J-dual#3 and J-dual#6 reporter cells were then transduced with lentiviral vector particles transferring these sgRNA sequences. Upregulation of BFP expression was observed by flow cytometry for 10 of the 12 candidate genes. For nine of them, transfer of both sgRNA sequences upregulated BFP expression in both reporter cell clones (Fig. S2A). For DUS2 and ZNF221, the screening results could not be confirmed. Since long half-lives of mRNAs and proteins may prolong the time from inactivation of the targeted gene to the decline of protein levels, we continued to monitor the reporter cells transduced with sgRNA genes targeting DUS2 and ZNF221 for up to 2 weeks postinfection. However, upregulation of Gag-BFP remained undetectable (data not shown). In addition, we confirmed sufficient transduction rates by the lentiviral DUS2 and ZNF221 targeting vectors, excluding the possibility that reduced lentiviral vector titers are responsible for
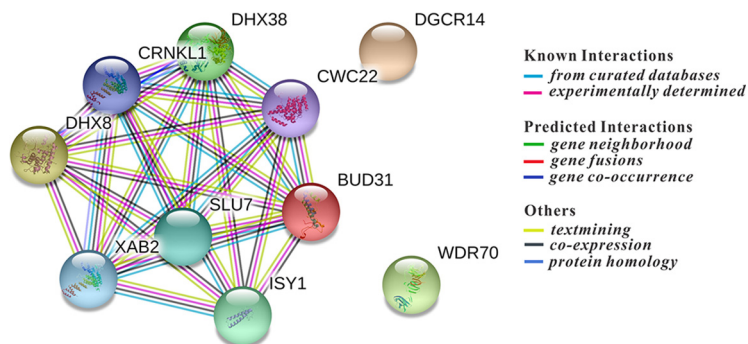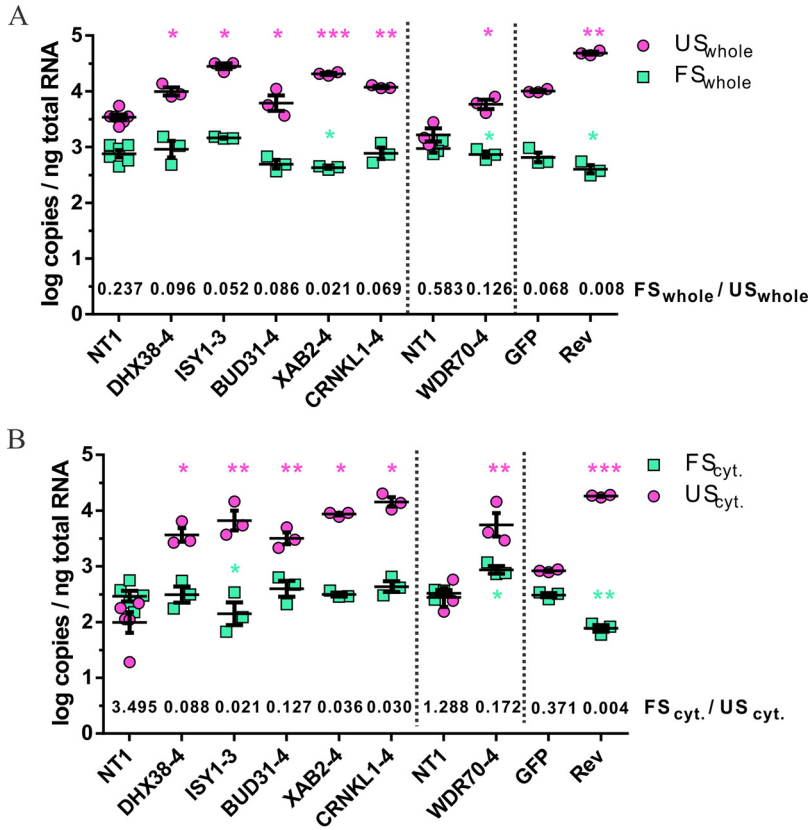
FIG 3 Protein interaction network of confirmed screening hits. The 10 validated hits were analyzed by the online database STRING for protein-protein interactions. The graph shows known and predicted protein-protein interactions for the most enriched pathway, the "mRNA splicing" pathway (pathway ID, GO 0000398; false discovery rate, 5.07e−14; minimum required interaction score, 0.400 [medium confidence]).

their failure to upregulate BFP expression (Fig. S2B). Despite this, the entire screening approach was highly specific, showing a false discovery rate (FDR) of only 0.166.

A first hint of potential mechanisms by which inactivation of the identified cellular genes could enhance Gag-BFP expression in the reporter cell lines was obtained by searching for functional protein association networks between the identified genes using the STRING online database (https://string-db.org/). Strikingly, 9 of the 10 confirmed hits are involved in pre-mRNA splicing (Fig. 3). Among them is ESS2/DGCR14, for which no interactions with the other identified factors have been reported but which is associated with U6 snRNA as well as U1 and U4 snRNAs (51). WDR70 is not linked to splicing; however, it contains a WD40 repeat domain which is shared by a number of splicing factors (52). The pre-mRNA splicing process includes two excision steps catalyzed by a large and dynamic RNP machine. The confirmed hits DHX38, DHX8, SLU7, and DGCR14 are typical second-step factors. In contrast, XAB2, ISY1, CRNKL1, CWC22, and BUD31 should be recruited at the first step, and intriguingly, all belong to the nuclear Prp19-associated complex (53–56). This complex is involved in diverse nuclear functions, such as DNA double-strand break (DSB) repair, transcriptional elongation, and splicing (57). The core factor of the complex, PRPF19 (human Prp19), however, was not identified in the screen. Consistently, targeting PRPF19 by two sgRNAs in J-dual#3 did not affect Gag-BFP expression (Fig. S3A). Since most of the confirmed hits are components of the spliceosome, we analyzed whether PRPF8, an ATPase/helicase at the core of the spliceosome (58), could also regulate the HIV Gag expression. Moreover, BUD13 from the RES complex and Tpr from the NPC, which were reported to have a nuclear retention function, were tested as well by using two different CRISPR/Cas sgRNAs targeting each of the genes. None of these targeting sgRNAs tested induced BFP expression (Fig. S3B to D).

CRNKL1 and other cellular factors impact HIV-1 unspliced and fully spliced mRNA levels. To follow up on the reduction of splicing as the potential mechanism of enhanced Gag-BFP expression, the top 5 hits and WDR70 were analyzed for their effect on HIV-1 mRNA splicing. Attempts to generate J-dual#3 cells with a stable knockout of the six selected target genes were not successful, presumably due to the fact that the selected genes are essential for cell viability (59). We therefore analyzed HIV RNA splicing and nuclear export after knockdown of these genes in short-term cultures. J-dual#3 cells were transduced with the lentiviral vectors encoding the different guide RNA genes, and Gag-BFP-positive cells were sorted to enrich for cells with the desired gene knockdown. Total RNA was then extracted from the sorted cells, and the copy numbers of unspliced HIV RNA per ng extracted total RNA were determined and compared to the copy numbers of unspliced RNA in J-dual#3 cells transduced with a negative control vector (NT1). Consistent with the stronger BFP expression, unspliced HIV RNA copy numbers were also enhanced 3- to 9-fold (Fig. 4A). Fully spliced HIV-1 RNAs

FIG 4 Role of confirmed screening hits in HIV-1 RNA splicing and export. (A) J-dual#3 cells were coinfected at an MOI of 3 with a lentiviral vector encoding Cas9 and a lentiviral vector expressing the guide RNAs indicated. In addition, J-dual#3 cells were transduced by a lentiviral vector encoding GFP or Rev at an MOI of 1. At 4 days after infection, BFP-positive cells were sorted by flow cytometry from cells transduced with lentiviral vectors encoding Rev or the guide RNAs targeting the confirmed hits. Matched controls expressing NT1 or GFP were used without sorting. Total RNA was purified from whole-cell extracts. HIV-1 US and FS RNA copy numbers in these samples were quantified by RT-qPCR and are shown as log copies per ng total RNA extracted from whole cells. Log copies from at least three experiments, as well as their mean and standard error of the mean (SEM), are shown. The mean of the FS_{whole}/US_{whole} ratio for each experiment is shown above the x axis for each of the treatment groups. (B) J-dual#3 knockout cells and control cells were prepared as described for panel A, but total RNA was extracted from the cytoplasmic fraction of each cell sample. Log copies per ng total cytoplasmic RNA and mean $FS_{cyt.}/US_{cyt.}$ ratios. were calculated and presented as in panel A. (C) The fold changes in the different treatment groups for the splicing efficacy (FS_{whole}/US_{whole}), the nuclear

(Continued on next page)

were also quantified, revealing a marginal decrease in cells transduced with XAB2 and WDR70 targeting sgRNAs but not with the other gene targeting sgRNA (Fig. 4A). Differential effects of the gene targeting HIV-1 FS and HIV-1 US RNA also lead to decreased ratios of FS to US transcripts (Fig. 4A). For comparative reasons, we also transduced J-dual#3 cells in an independent experiment with a lentiviral vector expressing Rev, sorted Gag-BFP-positive cells, and quantified HIV-1 US and FS transcripts. Rev enhanced the US HIV-1 RNA levels approximately 5-fold, which coincided with a minor reduction in FS transcripts (Fig. 4A).

The splicing process is closely linked to nuclear mRNA trafficking (60, 61). Therefore, we also determined the effect of targeting the six selected candidate genes on the cytoplasmic levels of HIV-1 US and FS RNA (Fig. 4B). Proper separation of the cytoplasm from nuclear components was controlled by examining GAPDH (glyceraldehyde-3-phosphate dehydrogenase) pre-mRNA levels and a nuclear marker protein in the cytoplasmic fraction from Jurkat cells (62, 63). GAPDH pre-mRNA levels per ng total RNA extracted from the cytoplasm were approximately 50-fold lower than GAPDH pre-mRNA levels per ng total RNA extracted from whole cells. Consistently, the nuclear LaminB protein could not be detected in the cytoplasmic fractions (Fig. S4).

After transduction with the different targeting vectors, HIV-1 US RNA in the cytoplasm of Gag-BFP-positive cells was enhanced approximately 22- to 116-fold, while the cytoplasmic FS RNA levels were affected only 0.6- to 2.7-fold (Fig. 4B). The magnitude of the enhancement of cytoplasmic levels of HIV-1 US RNA exceeded the enhancement of HIV-1 US RNA in whole-cell extracts, indicating that knocking down the selected target genes not only reduced splicing but also enhanced nuclear export of the HIV-1 US RNA. Strong enhancement of cytoplasmic HIV-1 US RNA export was also observed in Rev-expressing cells. However, in contrast to the cells transduced with the targeting vectors, cytoplasmic FS RNA levels in Rev-transduced cells were significantly reduced, consistent with the previously reported suppression of the TAP/NXF1-mediated RNA export by Rev (64).

To dissect the relative contributions of inhibition of splicing and enhancement of nuclear export of HIV-1 US RNA, we compared the ratio of HIV-1 FS RNA to HIV-1 US RNA in whole-cell extracts as a measure of splicing efficiency and the ratio of HIV-1 US RNA in cytoplasmic extracts to the HIV-1 US RNA in whole-cell extracts as a measure of nuclear export efficiency. This revealed two different patterns of responses. The fold reduction in the splicing efficiencies in cells transduced with ISY1, XAB2, and WDR70 targeting vectors mirrored the fold enhancement of the respective nuclear export efficiencies. Expression of Rev induced a similar pattern of response. In contrast, transduction with DHX38, BUD31, or CRNKL1 targeting vectors predominantly enhanced the nuclear export efficiency with at least 3-fold-lower effects on splicing efficiency (Fig. 4C). In particular, knockdown of CRNKL1 enhanced the ratio of HIV-1 US RNA in cytoplasmic extracts to HIV-1 US RNA in total cell extracts more than 34-fold, while the ratio of HIV-1 FS RNA to HIV-1 US RNA in whole-cell extracts was decreased only 4-fold. Targeting of CRNKL1 also led to the strongest enhancement of cytoplasmic levels of HIV-1 US RNAs without reducing cytoplasmic HIV-1 FS RNA levels (Fig. 4C), indicating that CRNKL1 is a major nuclear retention factor of the genomic HIV-1 US RNA.

CRNKL1 associates with HIV-1 US RNA in the nucleus. As a nuclear retention factor of HIV-1 US RNA, CRNKL1 should either directly bind to the HIV-1 US RNA or be in the same RNP complex. Due to lack of a putative RNA binding motif, it was suggested

FIG 4 Legend (Continued)
export efficacy ($US_{cyt.}/US_{whole}$), and cytoplasmic levels of unspliced (US_{cyt}) and fully spliced (FS_{cyt}) transcripts were calculated by dividing the values in the treatment groups by the ones obtained from the matched control cells (NT1, GFP) as described in Materials and Methods. A paired t test was performed on the log copy numbers/ng total RNA between each of the treatment groups and the matched control. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. Unless otherwise indicated, differences were not significant. Vertical dotted lines separate experiments that were not performed in parallel and therefore contain independent controls. WDR70.KD cells were sorted on day 7 after transduction, while other cells were sorted on day 4.
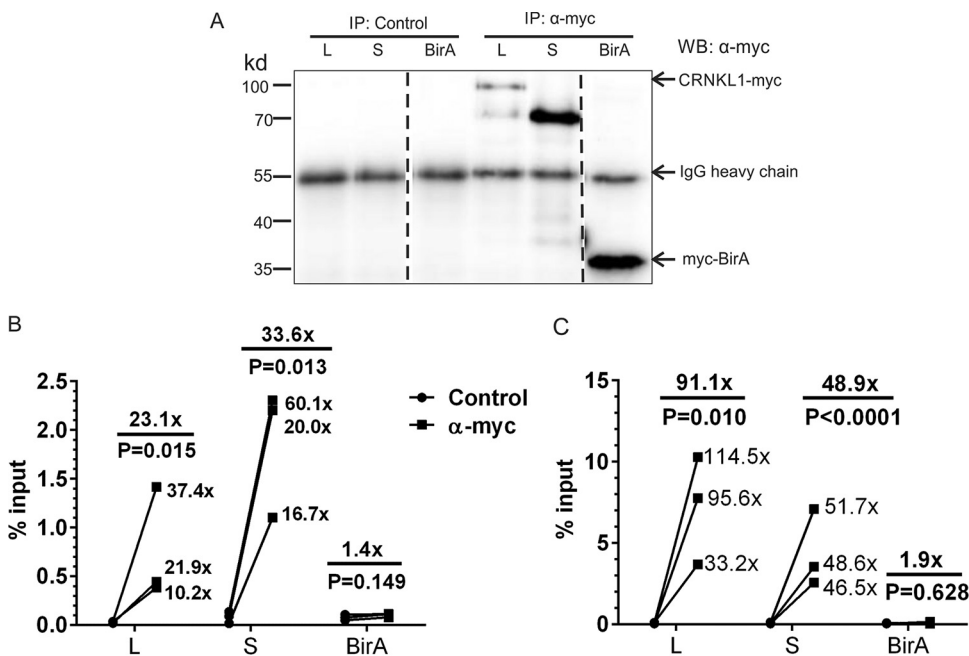
FIG 5 CRNKL1 proteins associate with HIV US RNA. (A) 293T cells transfected with expression plasmids encoding the myc-tagged CRNKL1 or an irrelevant control protein (BirA) were infected with VSV-G-pseudotyped HIV containing inactivating mutations in *rev* and *env* (HIV.r⁻e⁻). Nuclear fractions of these cells were immunoprecipitated with beads coated with anti-myc antibody or an isotype-matched control antibody (IgG2a). The precipitates were characterized by Western blot analysis with an anti-myc antibody. Dotted lines indicate excision of irrelevant lanes from the Western blot image. The precipitates described in panel A were also analyzed by RT-qPCR for HIV-US RNA and HIV-FS RNA levels. Results are expressed as percentages of HIV-1 US RNA (B) and HIV-1 FS RNA (C) copy numbers in the precipitate relative to the copy numbers in the nuclear fraction (input) prior to the immunoprecipitation. Fold changes of each replicate and mean fold changes are indicated. The ratio paired *t* test was used for statistical analysis. L, CRNKL1 large isoform; S, CRNKL1 small isoform.

that CRNKL1 binds to RNA via interactions with RNA binding proteins (65). However, CRNKL1 also contains a half-a-tetratricopeptide repeat (HAT) motif that has been implicated with direct RNA binding activity (66). To test whether CRNKL1 is indeed linked to nuclear HIV-1 US RNA, the two major isoforms of CRNKL1 (100 and 83 kDa [67]) were expressed in 293T cells with a myc tag. Pulldown experiments of nuclear extracts of these cells with an anti-myc antibody and an isotype-matched control antibody confirmed specific immunoprecipitation of both isoforms (Fig. 5A). Precipitates of the small and large isoforms of CRNKL1 also contained 34- and 23-fold higher levels of HIV-1 US RNA (Fig. 5B), confirming that at least 1% to 2% of the nuclear HIV-1 US RNA is associated with CRNKL1. No interaction was observed between HIV-1 US RNA and an irrelevant myc-tagged protein, BirA, revealing the specificity of HIV-1 US RNA interaction with CRNKL1. Comparable analyses for HIV FS RNA also revealed strong and specific associations with both isoforms of CRNKL1 (Fig. 5C). Since CRNKL1 is a member of the postcatalytic spliceosome (68), association of spliced transcripts with CRNKL1 complexes is expected. The different effects of the CRNKL1 knockdown on cytoplasmic RNA levels of HIV-1 US and FS RNA can therefore be best explained by the association of CRNKL1 with two or more RNP complexes differing in function.

Knockdown of CRNKL1 enhances Gag expression in acutely infected 293T cells. To exclude potential off-target effects of CRISPR/Cas-mediated genome editing, CRNKL1 expression levels were also reduced by small interfering RNAs (siRNAs). Cotransfection of 293T cells with a pool of four siRNAs targeting CRNKL1 and expression plasmids encoding the large and short forms of CRNKL1 led to strongly reduced CRNKL1 expression levels (Fig. 6A), confirming the functionality of the siRNAs. A simple model to test the effect of CRNKL1 knockdown on HIV-1 Gag expression might be to cotransfect
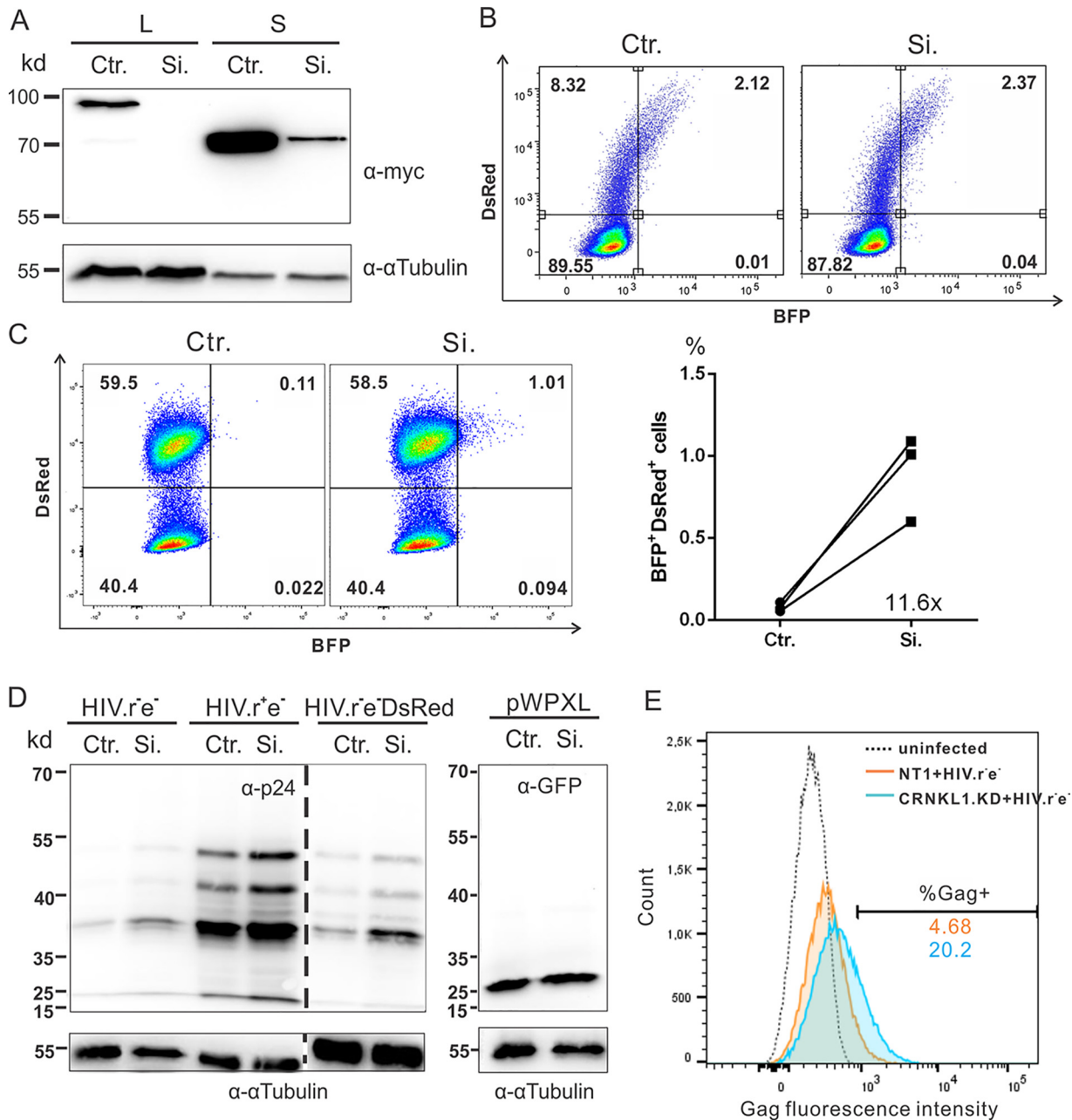
FIG 6 CRNKL1 knockdown increases expression of Gag after acute HIV infection. (A) Knockdown of CRNKL1 expression by siRNAs. CRNKL1-targeting siRNAs (Si.) and nontargeting control siRNAs (Ctr.) were cotransfected with myc-tagged CRNKL1 expression plasmids to test their efficacy and specificity. Exogenous CRNKL1 expression and endogenous α-tubulin expression in transfected 293T cells were monitored by anti-myc and anti-α-tubulin Western blot analyses, respectively. (B) Effect of CRNKL1 knockdown on Gag-BFP expression in transiently transfected cells. Targeting or nontargeting siRNAs were cotransfected with the HIV-dual-GT plasmid into 293T cells. Gag-BFP reporter expression was monitored by FACS analysis. (C) Effect of CRNKL1 knockdown on Gag-BFP expression in stably transduced cells. 293T cells stably transduced with HIV-dual-GT were transfected with the siRNAs and analyzed by flow cytometry. Shown are representative results from one experiment (left panel) and the percentage of BFP$^+$ DsRed$^+$ cells plotted from three independent experiments (right panel). (D) Effect of CRNKL1 knockdown on Gag expression in cells acutely infected with HIV-1. 293T cells were transfected twice within 4 days with siRNAs and then infected with VSV-G-pseudotyped HIV-1 env mutants with (HIV.r$^-$e$^-$) or without (HIV.R$^+$e$^-$) inactivating mutations of Rev. A reporter virus expressing wild-type Gag (HIV.r$^-$e$^-$DsRed) and a lentiviral vector expressing GFP from an internal promoter (pWPXL) were also included. Gag, GFP, and α-tubulin expression was analyzed by Western blotting using antibodies targeting the proteins indicated. The dotted line indicates excision of blank lanes from the Western blot image. (E) Effect of CRNKL1 knockdown on Gag expression in cells acutely infected with HIV-1 by flow cytometry. 293T cells were infected with CRNKL1 targeting or nontargeting CRISPR lentivirus and VSV-G-pseudotyped HIV.r$^-$e$^-$. Gag expression was analyzed on the single-cell level by intracellular staining for Gag and flow cytometry.

these siRNAs targeting CRNKL1 with the HIV-dual-GT plasmid into 293T cells. However, by doing this, we could not observe a rise in the Gag-BFP reporter protein expression (Fig. 6B). In contrast, transfection of CRNKL1-targeting siRNAs into 293T cells stably transduced with the same HIV-dual-GT reporter virus revealed an 11.6-fold increase in the percentage of transduced cells expressing Gag-BFP (Fig. 6C). Given the complex and dynamic chromosomal architecture (69), the nuclear metabolism of mRNA expressed from a circular plasmid can differ from mRNA expressed from a proviral DNA integrated into a host chromosome. This may explain why the effect of CRNKL1 knockdown is detectable in transduced but not transfected cells.

To further explore the effect of the CRNKL1 knockdown in acutely infected cells in the presence and absence of Rev, 293T cells were first transfected with CRNKL1-targeting siRNAs or controls prior to infection with different VSV-G-pseudotyped HIV constructs. CRNKL1 knockdown enhanced Gag expression from HIV *env* mutants independent of Rev (Fig. 6D, left panel). This enhancement was not due to an increased susceptibility to virus infection after CRNKL1 siRNA treatment, since green fluorescent protein (GFP) expression levels after transduction with a lentiviral vector encoding the reporter gene from an internal promoter remained unchanged (Fig. 6D, right panel). The magnitude of the enhancing effect of CRNKL1 knockdown by Western blot analyses in infected 293T cells was lower than expected, based on the analyses of cytoplasmic RNA levels in transduced Jurkat cells enriched for CRNKL1 knockdown (Fig. 4B). As the enhancing effect of the CRNKL1 knockdown may be masked by background expression of Gag from cells expressing unreduced levels of CRNKL1, we also analyzed Gag expression after CRNKL1 knockdown on the single-cell level by flow cytometry. 293T cells were transduced with lentiviral vectors targeting CRNKL1 or a control vector and then infected with a VSV-G-pseudotyped HIV containing only inactivating mutations in *env* and *rev*. Intracellular staining for Gag revealed a 4-fold increase in the percentage of Gag-positive cells after CRNKL1 knockdown (Fig. 6E), indicating that the enhancing effect is not restricted to the HIV-dual-GT reporter virus or Jurkat cells.

CRNKL1 knockdown shifts cytoplasmic levels in HIV-1 splice variants. To confirm the role of CRNKL1 in the regulation of cytoplasmic HIV RNA levels and to explore potential effects of CRNKL1 on cytoplasmic levels of cellular RNAs, a transcriptomic analysis was performed after a short-term knockdown of CRNKL1. J-dual#3 reporter cells were again transduced with the CRNKL1 targeting vector, and Gag-BFP-positive cells were sorted by flow cytometry. RNA was extracted from the cytoplasmic fraction of sorted cells and J-dual#3 cells transduced with a control vector. The extracted RNAs from four biological replicates each were then enriched for mRNAs and sequenced by the Illumina HiSeq procedure. In addition to this, cytoplasmic mRNA was extracted from J-dual#3 cells expressing Gag-BFP after transduction with the Rev-encoding lentiviral vector and sequenced after mRNA enrichment.

To confirm the effectiveness of Cas9/sgRNA-mediated knockdown by transcriptomic analysis, a core target region in exon 3 of the CRNKL1 gene was defined as nucleotides (nt) −3 to +3 relative to the SpCas9 cleavage site. The number of transcriptome sequencing (RNA-seq) reads that contained 10 nucleotides upstream and downstream of the core target region were then determined (Fig. S5). While 55 reads on average mapped to the region flanking the core target region in the control cells, only a mean of 5.8 reads was detected in cells transduced with the CRNKL1 knockdown vector (Fig. S5). Larger deletions and substitutions outside the core target region probably reduce the number of reads that can be mapped to the flanking region. A comparison of the reads from CRNKL1 knockdown cells that can be mapped to the flanking region with the wild-type CRNKL1 core target region revealed insertions, deletions, and point mutations in 22 out of 23 reads. In control cells, only 2 of 219 reads contained single point mutations, clearly indicating efficient inhibition of CRNKL1 expression after Cas9/sgRNA targeting (Fig. S5).

HIV reads were mapped to the proviral DNA of HIV-dual-GT, revealing more reads mapping to intronic regions after CRNKL1 knockdown and Rev expression than in the
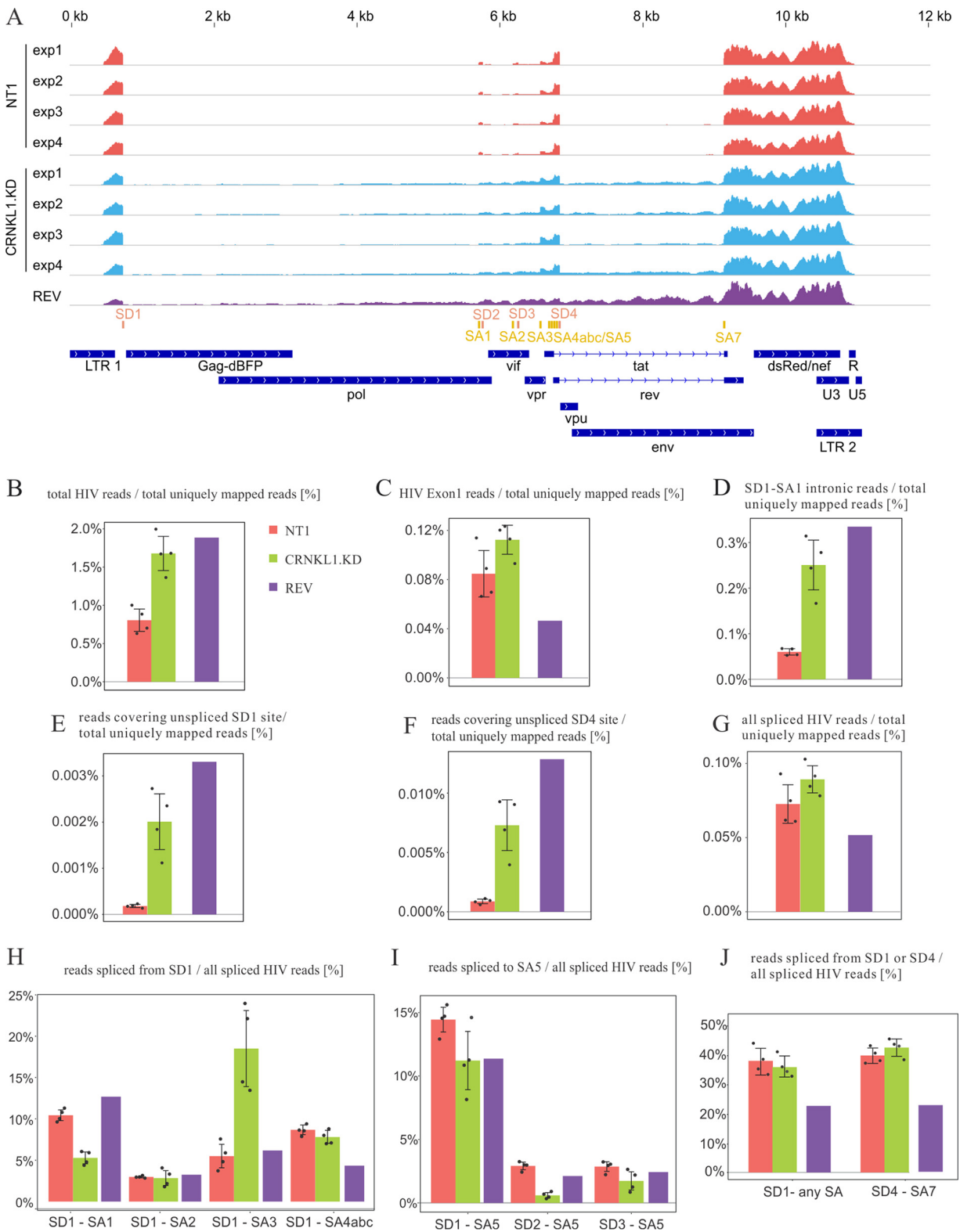
FIG 7 Effect of CRNKL1 depletion on HIV RNA splicing efficiency, accuracy, and export. (A) Alignments of reads from the transcriptomic analysis to HIV-dual-GT. Shown are coverage profiles for the following individual RNA-seq samples, indicated by color (from top to bottom): the nontargeting control (NT1, red), CRNKL1 knockdown (cyan), and *rev* overexpression (purple). Splice sites are shown in orange (donors) and dark yellow (acceptors). Coding regions and long terminal repeats (LTRs) are indicated in dark blue. (B) Percentage of all HIV reads among the total (HIV plus human) uniquely mapped reads. (C) Percentage of reads mapping to HIV exon 1 (from the transcription start site to SD1) among the total uniquely mapped reads. (D) Percentage of reads in the first intron between SD1 and SA1 among the total uniquely

(Continued on next page)

controls (Fig. 7A). The percentage of reads mapping to HIV among all the reads uniquely mapping to either the human or the viral genome (total uniquely mapped reads) was enhanced 2-fold in CRNKL1 knockdown cells (Fig. 7B). This was not due to an enhanced transcriptional activity, since the percentage of reads mapping to exon 1, which is shared by all HIV transcripts, was affected only marginally (Fig. 7C). Changes in HIV-1 US RNA were assessed by determining the percentage of reads mapping to the first intron of HIV-1 between splice donor 1 (SD1) and splice acceptor 1 (SA1) (Fig. 7D) or mapping to the unspliced SD1 site (Fig. 7E), revealing 4- and 12-fold increases under CRNKL1 knockdown conditions, respectively. The percentage of reads mapping to the unspliced SD4 site and therefore containing the *env* intron also increased by a factor of 9 (Fig. 7F). The change in the percentage of all spliced HIV reads was negligible (Fig. 7G), indicating that knockdown of CRNKL1 does not reduce splicing of HIV RNAs in general.

To further explore differences in cytoplasmic levels of differentially spliced HIV-1 transcripts, the precise splice site usage was also systematically analyzed by calculating the percentage of reads covering specific SD-SA pairs among all spliced HIV reads (Fig. S6A and B). This revealed that knockdown of CRNKL1 reduced cytoplasmic HIV transcripts spliced from SD1 to SA1 approximately 2-fold, while transcripts spliced from SD1 to SA3 were enhanced 3-fold (Fig. 7H). To further assess the effect of the CRNKL1 knockdown on the major Env-encoding transcript, the percentage of reads spliced to SA5 was determined (Fig. 7I). Independent of the SD sites, usage of SA5 tended to be reduced, which is similar to an effect recently reported for the splicing factor DDX17 (70). Moreover, the level of fully spliced transcripts seemed unaffected by the CRNKL1 knockdown, since the percentages of reads spliced either from SD1 to any splice acceptor site or from SD4 to SA7 were nearly the same (Fig. 7J).

CRNKL1 is a regulator of cytoplasmic cellular mRNA levels. The influence of the CRNKL1 knockdown on cytoplasmic mRNA levels was also determined (Fig. S7A and B). To display the results, the $\log_2$-transformed expression differences were plotted against the expression level for each gene (Fig. 8A). To reduce noise and arbitrarily large fold changes from low-expression genes, the $\log_2$ fold changes are shrunk, putting less weight on low-expression genes. Significant differences after correction for multiple testing revealed the upregulation of 1,249 and downregulation of 2,569 mRNAs, respectively. Two hundred ten mRNAs even showed a more than 4-fold ($\log_2$ fold change, >2) increase after knockdown of CRNKL1. Interestingly, upregulated mRNAs were enriched for the functional categories involved in transcription (Fig. 8B) while downregulated mRNAs were more frequently associated with functions in cell cycle and organelle organization (Fig. 8B).

Given the close link between RNA export and splicing, a more general role of CRNKL1 in cytoplasmic levels of splice variants was analyzed upon CRNKL1 knockdown. Individual splicing events (including skipped exons, alternative splice sites, and retained introns) were quantified using the percent spliced-in (PSI) metric (97), and differential analysis was carried out by SUPPA (71). Of all quantified splicing events in cytoplasmic mRNAs upon CRNKL1 depletion, 2.4% were upregulated and 1.6% were downregulated (Fig. S7C and D). Similarly, alternative 3' and 5' splice sites were more frequently detected upon CRNKL1 knockdown, suggesting that CRNKL1 inhibits usage of alternative splice sites (Fig. 8C).

FIG 7 Legend (Continued)
mapped reads. (E) Percentage of reads across the unspliced SD1 site among the total uniquely mapped reads. (F) Percentage of reads across the unspliced SD4 site among the total uniquely mapped reads. (G) Percentage of reads spanning SD and SA sites (spliced reads) among the total uniquely mapped reads. (H) Percentages of reads spanning SD1 and the indicated splice acceptor sites among all spliced HIV reads. For SD1-SA4abc, reads spliced from SD1 to SA4a, SA4b, and SA4c are summed up. (I) Percentages of reads spanning the indicated splice donor sites and SA5 among all spliced HIV reads. (J) Percentage of reads spliced from SD1 to any SA site or from SD4 to SA7 among all spliced HIV reads. (B to J) Shown are averages and standard deviations for the control and CRNKL1 knockdown samples. Individual values are shown as black dots. For the *rev* overexpression sample (purple), the individual value is shown.
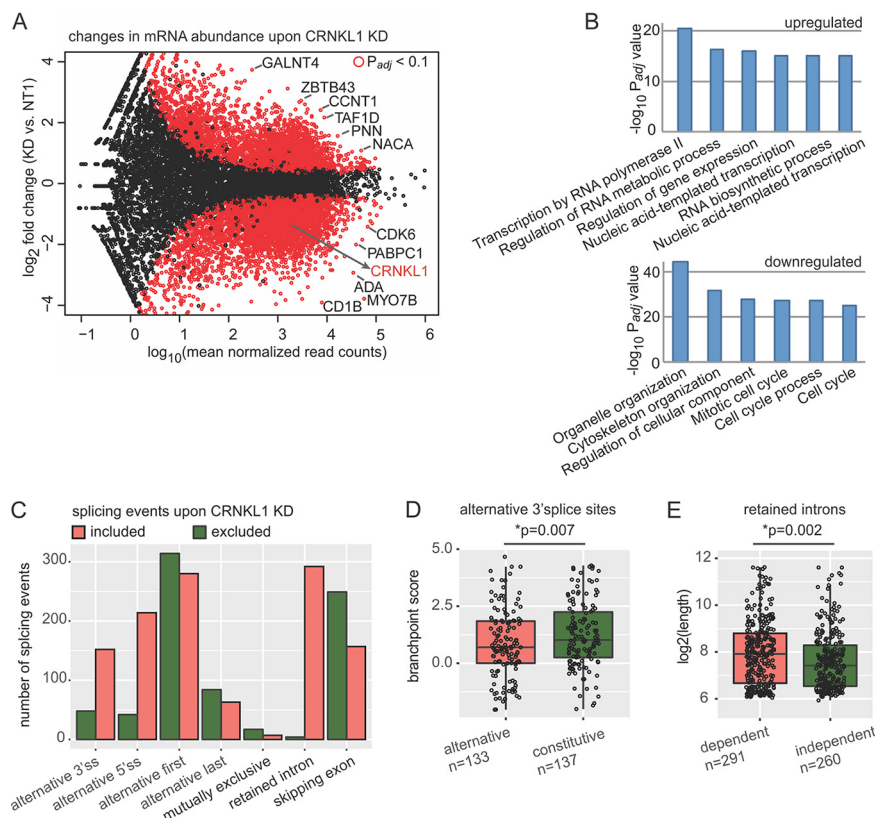
FIG 8 Changes in the host transcriptome upon CRNKL1 knockdown (KD). (A) RNA-seq differential expression analysis of mRNA abundance upon CRNKL1 KD. An MA plot shows differences in the abundance of protein coding transcripts in CRNKL1 KD versus control (NT1) conditions. mRNAs with significant differences are depicted in red ($P_{adj} < 0.1$). (B) Gene Ontology enrichment analysis of mRNAs with upregulated and downregulated expression upon CRNKL1 KD. The top enriched GO Biological Process terms are shown. (C) Upregulated (included) (ΔPSI [KD − NT1] > 0.25, $P_{adj} < 0.01$) and downregulated (excluded) (ΔPSI [KD − NT1] < −0.25, $P_{adj} < 0.01$) splicing events were quantified from the RNA-seq CRNKL1 KD experiment. The bar plot shows absolute numbers of splicing events that are up- or downregulated upon CRNKL1 knockdown. (D) Branch point scores were calculated by BPfinder, and alternative 3′ splice sites with higher usage in CRNKL1 KD condition were compared to 3′ splice sites of the same genes with unchanged (constitutive) usage. Box plots and data points are shown. The Wilcoxon rank sum test was used to test for significance. Absolute numbers of 3′ splice sites included in the analysis for each group are indicated. (E) Length of retained introns with higher inclusion upon KD (CRNKL1 dependent) was compared to the length of unchanged introns upon KD (CRNKL1 independent), which were located in the same transcripts as CRNKL1-dependent ones. Box plots and data points are shown. The Wilcoxon rank sum test was used to test for significance. Absolute numbers of retained introns analyzed in both groups are indicated. PSI, percent spliced-in for splicing events.

A total of 295 introns were more frequently retained in cytoplasmic mRNAs after CRNKL1 knockdown (Fig. 8C; Table S2). Further analyses revealed that not all introns of the same transcript were affected in the same way (Fig. S8) but rather that intron retention was highly specific for one or a few introns of the same gene. Enhanced cytoplasmic levels of intron-retaining mRNAs could be confirmed by real-time PCR using total cytoplasmic RNAs extracted for RNA-seq and primers spanning the respective exon-intron junctions (Fig. S8).

Genes containing introns or alternative splice sites that were upregulated after CRNKL1 knockdown were not enriched for particular functional categories (data not shown). Enhanced detection of introns or alternative splice sites in cytoplasmic mRNAs after CRNKL1 knockdown indicate that CRNKL1 suppresses either these splicing events or the nuclear export of the respective splice variants. To gain insight into the principles of regulation of these CRNKL1 knockdown-dependent introns and alternative splice sites, we analyzed common features of these splicing events, including splice
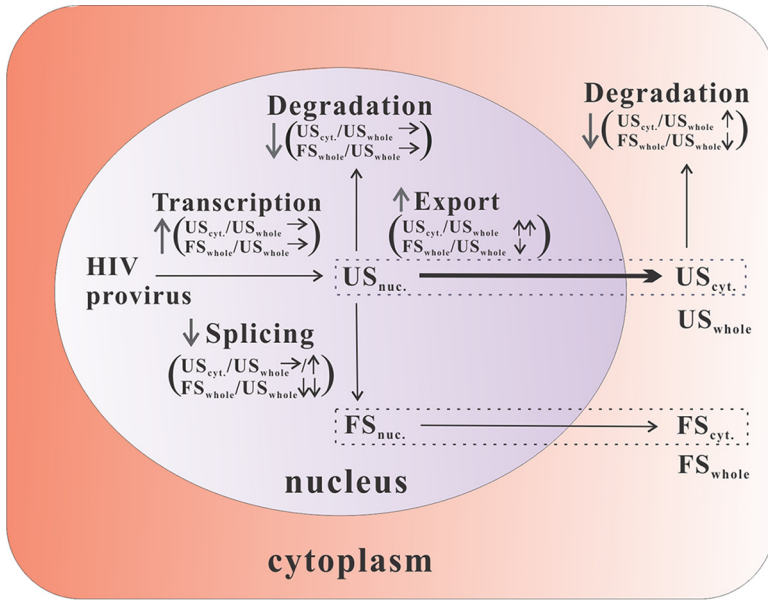
FIG 9 Potential mechanisms leading to increased cytoplasmic levels of unspliced HIV-1 RNA. Graphical representation of HIV-1 RNA metabolism and expected consequences of upregulation (↑) or downregulation (↓) regulation or no change (→) of the respective steps on the nuclear export efficacy (US_{cyt}/US_{whole}) and splicing efficacy (FS_{whole}/US_{whole}). The nuclear export pathway is highlighted in bold.

site strength, branch point and polypyrimidine tract score, GC content, and intron length. CRNKL1 knockdown-dependent alternative 3′ splice sites showed significantly lower branch point scores than constitutive 3′ splice sites (Fig. 8D) but a similar length of polypyrimidine tracts (Fig. S7E). We also found that CRNKL1-dependent introns were significantly longer than CRNKL1-independent introns within the same transcript (Fig. 8E) and tended to have longer polypyrimidine tracts (Fig. S7F) but showed no difference in GC content (Fig. S7G).

DISCUSSION

Using a genome-wide screen for cellular factors repressing expression of an HIV-1 structural protein from the unspliced RNA, we identified 10 target proteins linked to mRNA metabolism, with nine of them associated with the spliceosome. Interestingly, five of them (ISY1, BUD31, XAB2, CRNKL1, and CWC22) are found in the Prp19-associated complex. This complex is important for stabilizing the spliceosome, which is also involved in DNA repair, transcriptional elongation, and RNA export (57). Given that the yeast homolog of WDR70 was reported to interact with Prp19 (72) and is now shown to affect RNA splicing, it could be a new member of the Prp19-associated complex. Due to the close linkage of splicing with RNA stability and nuclear RNA export, it is difficult to dissect the precise mechanism by which deletion of the identified nuclear proteins enhance cytoplasmic RNA levels of HIV-1 US RNA.

Theoretically, enhanced transcription, reduced nuclear or cytoplasmic degradation, decreased splicing, and/or enhanced nuclear export can increase cytoplasmic levels of HIV-1 US RNA (US_{cyt}) (Fig. 9). Enhanced transcription and decreased nuclear degradation seem unlikely, since the percentage of reads mapping to the first HIV-1 exon present on all viral transcripts was not significantly enhanced (Fig. 7C). The absence of enhanced cytoplasmic levels of FS RNAs and DsRed expression levels from the dual reporter construct (Fig. 4B; Fig. S2A) also argues against such a mechanism. Assuming a single pool of nuclear HIV-1 US RNA that could follow either of the two remaining pathways, differential effects on the nuclear export efficiency (US_{cyt}/US_{whole}) and splicing efficacy (FS_{whole}/US_{whole}) are expected (Fig. 9). Enhancing nuclear export of HIV-1 US RNA should strongly enhance the US_{cyt}/US_{whole} ratio. Since rapid export may reduce

splicing and degradation, the FS_{whole}/US_{whole} ratio may decrease responsively, although to a lower extent. Reduced splicing should primarily decrease the FS_{whole}/US_{whole} ratio but not the $US_{cyt.}/US_{whole}$ ratio (Fig. 9). Similar to an enhanced nuclear export, reduced degradation of cytoplasmic HIV-1 US RNA should also lead to enhanced $US_{cyt.}/US_{whole}$ and decreased FS_{whole}/US_{whole} ratios (Fig. 9), but the association of the identified proteins with the spliceosome localized in the nucleus argues against such a cytoplasmic effector mechanism.

Clearly, ablation of all six target genes tested reduced the splicing efficacy (FS_{whole}/US_{whole}), indicating that the targeted genes are enrolled in splicing and/or nuclear retention. The relative contribution of reduced splicing versus enhanced nuclear export of the HIV-1 US RNA differed between the different target genes. The XAB2 knockdown decreased the splicing efficacy 12.5-fold and the nuclear export efficacy of the HIV-1 US RNA 11.7-fold. Since splicing and nuclear export efficacies were affected to a similar degree, it is not possible to conclude whether XAB2 acts predominantly as a splicing or a nuclear retention factor. This is consistent with previous observations that depletion of splicing factors (e.g., SF1, U2AF65, and Prp19) can cause pre-mRNA leakage into the cytoplasm (36, 73–75). In contrast to the effect of the XAB2 knockdown, the splicing efficacy after knockdown of CRNKL1 is decreased only 4.0-fold, while the nuclear export efficiency based on the HIV-1 US RNA increases 34.6-fold. Since the ratio of HIV-1 FS to US transcripts in the cytoplasmic RNA is also reduced to a larger extent than in the whole-cell RNA, enhanced nuclear export of the HIV US RNA after inactivation of the target gene seems to play a dominant role. This indicates that CRNKL1 is required primarily for retention of the HIV-1 US RNA in the nucleus and not for its splicing. The precise molecular mechanism by which the CRNKL1 knockdown leads to a more than 100-fold enhancement of cytoplasmic HIV US RNA levels remains to be further investigated. Our observation that targeting of CRNKL1 did not enhance cytoplasmic levels of HIV-1 FS RNA, which uses the default TAP/NXF1 export pathway, indicates that CRNKL1 is not a global nuclear mRNA retention factor. Our transcriptomic analysis also indicates a highly selective differential regulation of cytoplasmic levels of HIV-1 splice variants. The only spliced HIV-1 transcripts that were clearly upregulated by CRNKL1 knockdown have spliced SD1 to the SA3 site (Fig. 7G) and may encode Tat. Whether this represents an upregulation of *env* intron-retaining transcripts or the Tat-encoding fully spliced transcript is unclear. The enhanced percentage of cytoplasmic HIV transcripts harboring an unspliced SD4 site (Fig. 7F) and the unchanged percentage of transcripts harboring a spliced SD4 site (Fig. 7J) argue for the former.

The transcriptomic analyses after CRNKL1 knockdown also revealed a selective regulation of cytoplasmic levels of a subset of cellular mRNAs. We found that 11.0% and 20.1% of expressed mRNAs in the cytoplasm were up- and downregulated, respectively. Functionally, the upregulated mRNAs are associated with transcription, while downregulated mRNAs are more frequently associated with the cell cycle and organelle organization, suggesting that CRNKL1 participates as a master regulator in these cellular programs. The large percentage of cellular genes affected by CRNKL1 knockdown may be due not only to a direct consequence of reduced CRNKL1 levels but also to secondary effects requiring further investigations. Depletion of an essential gene such as the CRNKL1 gene may induce cell death. However, at the time of the analyses, the pattern of forward and side scatters during flow cytometric analysis of BFP-positive CRNKl1 knockout cells and BFP-negative cells were indistinguishable, while dead cells clearly differed (data not shown). Consistently, the percentage of BFP-positive cells remained unchanged for several days after the day of the analyses. In addition, the transcriptomic pathway analyses did not provide any evidence for upregulation of stress response genes or apoptosis-related genes.

Since altered nuclear export after CRNKL1 depletion could result in differences in the abundance of different cytoplasmic mRNA isoforms, we investigated alternatively spliced cytoplasmic isoforms upon CRNKL1 knockdown. Upon CRNKL1 depletion, a much higher number of included introns and alternative 3′ and 5′ splice sites than the

excluded ones were detected, suggesting that CRNKL1 promotes splicing of these events or highly selective nuclear retention. Intron retention in the absence of CRNKL1 was highly restricted and even varied among introns of the same gene (see Fig. S8 in the supplemental material). Searching for common properties of introns that were retained revealed that their mean length was significantly greater than that of introns with unchanged inclusion upon CRNKL1 depletion. In addition, usage of 3′ splice sites with lower branch point scores seems to be blocked in the presence of CRNKL1. Therefore, the sequence features of CRNKL1-dependent splicing events seem to reside at the 3′ end rather than at the 5′ end of the intron.

As a component of the core spliceosome, CRNKL1 is well conserved from yeast to human, is ubiquitously expressed in various tissues, and has been shown to affect splicing reactions *in vitro* (53, 54, 67). Since CRNKL1 has a HAT motif composed of alpha-helix pair repeats and belongs to the Prp19-associated complex, it has been implicated primarily in DNA repair and cell cycle regulation (76, 77). Noticeably, this motif may also confer direct RNA-binding activity. Beyond that, CRNKL1 has been reported to be involved in RNA elongation and RNA export (78, 79). The role of CRNKL1 in RNA extension and export is especially interesting, demonstrating that CRNKL1 has interactions with the TREX exporting machinery at an early stage of RNA synthesis, which might determine the "export" or "retain" status of an RNA.

Mechanistically, one plausible hypothesis based on the nuclear retention observed for the HIV-1 unspliced RNA is that a nuclear RNP complex containing CRNKL1 binds to RNA motifs present in a selective subset of cellular and viral introns, leading to their nuclear retention and splicing. In support of this hypothesis, we observed that CRNKL1 is associated with nuclear HIV-1 unspliced RNA (Fig. 5C). The precise binding motif that is bound by CRNKL1 or CRNKL1-containing RNP complexes needs to be further investigated. Binding of CRNKL1-containing RNPs to binding motifs on cellular transcripts leading to nuclear retention and degradation may also explain the increase in cytoplasmic levels of cellular mRNAs upon CRNKL1 knockdown. However, it is not possible to exclude that changes in cellular mRNA levels are secondary effects of the CRNKL1 knockdown.

In summary, our study identifies three spliceosomal proteins that are required for nuclear retention of the HIV-1 unspliced RNA and indicates that CRNKL1-dependent nuclear retention is a novel mechanism for the regulation of cytoplasmic levels of intron-retaining HIV-1 mRNAs that HIV-1 may have hijacked to regulate its complex splicing pattern.

MATERIALS AND METHODS

Plasmids. Plasmids expressing viral proteins, HIV-1 Rev (pcRev), HIV-1 Tat (pcTat), HIV-1 codon-optimized Gag-Pol (Hgp^{syn}), and VSV-G (pHiT-G), were described previously (27, 62). Lentiviral packaging plasmid pPAX2 and pMD2.G were obtained from Addgene (no. 12260 and 12259). HIV-1 proviral plasmid HIV.R$^+$e$^-$ is based on the HIV-1 molecular clone NL4-3 and contains a 4-nucleotide deletion at its NheI site, leading to inactivation of *env*. The plasmid HIV.r$^-$e$^-$ contains an additional mutation of the start codon of *rev* and a premature stop codon in the first *rev* exon (63). The plasmid HIV.r$^-$e$^-$ DsRed contains the mutations that HIV.r$^-$e$^-$ has and is further mutated with *nef* by replacing it with DsRed as described for HIV-dual-GT below. Lentiviral vector pWPXL delivering a GFP gene driven by an internal EF1α promoter was obtained from Addgene (no. 12257). The Rev coding lentivector pWPXL-Rev was cloned by replacing the GFP sequence between the MluI and EcoRI site in the pWPXL vector with the Rev coding sequence from pcRev. The large isoform of myc-tagged CRNKL1 coding plasmid CRNKL1(L)-myc was obtained from Origene (no. RC220334). The coding sequence of large CRNKL1 relative to amino acids (aa) 1 to 169 was deleted from CRNKL1(L)-myc, generating the expression construct for small isoform CRNKL1 [CRNKL1(S)-myc].

Proviral reporter plasmid HIV-dual-GT was constructed based on HIV.r$^-$e$^-$. A cassette containing mTagBFP2, a PEST coding sequence (destabilized BFP), and two stop codons were ligated in frame to the 3′ end of *gag*, forming a new ORF for Gag-BFP. The slippery sequence TTTTTTA (nt 2085 to nt 2091 of GenBank entry AF324493.2) was mutated to CTTCCTG to prevent Gag-Pol translation (80). The coding sequence of Nef close to its 5¨end (nt 8796 to nt 8880 of GenBank entry AF324493.2) was replaced by a cassette consisting of a 3×Ala linker sequence and the DsRed max coding region followed by two stop codons. The sequence (nt 5131 to nt 6297 of GenBank entry AF324493.2) ranging from *vif* to *vpu* was replaced by the corresponding region of NL4-3-Δ6-drEGFP (81) harboring inactivating point mutations in *vif*, *vpr*, and *vpu*. The *rev* ORF was inactivated as described above for HIV.r$^-$e$^-$ by PCR mutagenesis.

The plasmid HIV-dual-GT including the final sequence is available from AddGene (plasmid no. 122696). Individual CRISPR knockout plasmids were constructed based on lentiCRISPRv2 (no. 52961; Addgene) and GeCKO v2 sgRNA target sequences. All sgRNA sequences of the GeCKO v2 library can be found at www.addgene.org/pooled-library/zhang-human-gecko-v2/. The different sgRNA sequences selected for validation are provided in Table S1A in the supplemental material. Each of the sgRNA sequences was cloned into the vector by following the instructions on the Addgene webpage (www.addgene.org/52961/).

GeCKO v2 library. The pooled human GeCKO v2 CRISPR knockout plasmid library was a gift from Feng Zhang (47) and was obtained from Addgene (no. 1000000048 and 1000000049). It is supplied as two sublibraries, each containing three different sgRNA constructs targeting each gene in the genome. Amplification of the lentiviral vector plasmid library was performed essentially as described by the contributors' protocol (47). The complexity of the library was confirmed by HiSeq NGS of the pooled plasmid library (data not shown).

Cell culture. HEK293T (DSMZ) and TZM-bl (82) (NIH AIDS reagent program) cell lines were maintained in Dulbecco's Modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS) and 1% penicillin/streptomycin (D10). Jurkat, clone E61 (ATCC), and reporter cell lines J-dual#3 and J-dual#6 were maintained in Roswell Park Memorial Institute (RPMI) 1640 supplemented with 10% FCS and 1% penicillin/streptomycin (R10). After expansion, J-dual#3 and J-dual#6 stocks were stored frozen in aliquots at $-80°C$. A fresh aliquot was thawed and expanded in each assay.

Cell transfection. HEK293T cells were routinely transfected by the polyethylenimine (PEI) precipitation method (62). In general, 3.5 million cells per T25 flask (CellStar) were seeded 24 h before transfection. For each transfection, plasmid DNA was mixed with calf thymus carrier DNA (Thermo Fisher Scientific) to give a total of $10\,\mu g$ DNA. The $10\,\mu g$ DNA was mixed with $15\,\mu l$ PEI ($1\,\mu g/\mu l$; Sigma-Aldrich) and added to the cells. After 8 h of incubation with the cells, the transfection medium was replaced with fresh D10 medium. Supernatants or cells were harvested 48 h posttransfection. To produce pseudotyped HIV.R$^+$e$^-$, HIV.r$^-$e$^-$, HIV.r$^-$e$^-$DsRed, and HIV-dual-GT vector particles, $3\,\mu g$ proviral plasmid, $3\,\mu g$ Hgp^{syn}, $2\,\mu g$ pHiT-G, $1\,\mu g$ pcTat, and $1\,\mu g$ pcRev were used. To produce individual CRISPR knockout (CRISPR-KO) lentiviral vector particles, $5\,\mu g$ lentiCRISPRv2 plasmid, $3.75\,\mu g$ pPAX2, and $1.25\,\mu g$ pMD2.G were used. To produce the pooled human GeCKO v2 lentiviral vector library, 24 million 293T cells were seeded into a T175 tissue culture flask. At 24 h later, $20\,\mu g$ GeCKO v2 library plasmids, $15\,\mu g$ pPAX2, $5\,\mu g$ pMD2.G, and $120\,\mu l$ GenJet transfection reagent (SignaGen) were used for transfection. HIV-dual-GT-transduced Jurkat cells were transfected by the electroporation method described elsewhere (83). In brief, 5.0 million stably transduced cells were electroporated with $50\,\mu g$ plasmids by the Gene Pulser Xcell electroporation system (Bio-Rad) at 250 V and 1,500 μF. At 48 h posttransfection, cells were harvested for flow cytometry analysis. The transfection efficiency after expression of a GFP reporter plasmid was estimated to be between 10% and 20%.

siRNA transfection. A pool of four siRNA duplexes targeting CRNKL1 (Dharmacon; sequences shown in Table S1E) or the control, siGENOME nontargeting siRNA control pool no. 2 (D-001206-14-50; Dharmacon), was transfected into 293T or 293T reporter cells using the DharmaFECT 1 transfection reagent (Dharmacon). For 2×10^5 cells, 10 pmol siRNA with or without 100 ng plasmid was transfected essentially as described by the manufacturer. CRNKL1-myc expression by siRNA knockdown was examined by Western blotting at 2 days posttransfection; Gag-BFP expression influenced by CRNKL1 knockdown was measured by fluorescence-activated cell sorting (FACS) at 4 days posttransfection. To enhance the CRNKL1 knockdown efficiency, 4×10^5 cells were given two times 60 pmol siRNAs in 4 days and then infected with VSV-G-pseudotyped HIV variants or a control lentivector at an MOI of 1. At 2 days postinfection, the cells were lysed using lysis buffer, whose volume was adjusted by cell number.

Western blot analyses. Immunoblotting for HIV-1 Gag/CA p24 was described previously (63). Immunoblotting for Rev was performed with primary antibody sheep anti-Rev (1:4,500, no. H6006; US Biological Life Sciences) and secondary antibody rabbit anti-goat Ig/HRP (1:5,000, no. P0160; Dako). Immunoblotting for myc-tagged proteins and for GFP was performed using primary antibodies monoclonal mouse anti-myc tag (1:8,000, no. 2276; Cell Signaling Technology) and monoclonal mouse anti-GFP (1:500, no. sc-9996; Santa Cruz) in combination with the secondary antibody goat anti-mouse IgG (1:5,000, no. 115-035-146; Dianova). Immunoblotting for α-tubulin was performed using primary antibody polyclonal rabbit anti-α-tubulin antiserum (1:4,000, no. 600-401-880; Rockland), and secondary swine anti-rabbit Ig/horseradish peroxidase (HRP) (1:5,000, no. P0217; Dako). In experiments shown in Fig. 1B, 6A, and 6C, anti-α-tubulin blotting was conducted after stripping the blotted anti-CA and anti-myc membranes with stripping buffer (no. 2504; Millipore). The membranes were then reblocked and subsequently reblotted with antibodies as described above. Immunoblotting for LaminB was performed with primary antibody goat anti-LaminB (C-20) (1:500, no. sc-6216; Santa Cruz Biotechnology) and secondary antibody rabbit anti-goat Ig/HRP (1:5,000, no. P0160; Dako).

Intracellular staining. 293T cells were transduced by the CRNKL1 targeting or nontargeting CRISPR lentivectors at an MOI of 4. Two days later, the transduced cells were superinfected with VSV-G-pseudotyped HIV.r$^-$e$^-$ at an MOI of 1. For intracellular staining, 1.0×10^6 cells were fixed at 2 days postsuperinfection using 2% paraformaldehyde (PFA) (room temperature, 30 min) and then permeabilized with 0.5% saponin in FACS buffer (room temperature, 10 min). Cells were subsequently stained with $100\,\mu l$ anti-p24 primary antibody ($7\,ng/\mu l$ in permeabilization buffer; 4°C, 1 h), followed by washing with 1 ml permeabilization buffer two times. Staining with secondary antibody was done with $100\,\mu l$ rat anti-mouse IgG1/allophycocyanin (APC) conjugated (no. 17-4015-82; eBioscience) (1:500 dilution in permeabilization buffer; 4°C, 1 h), followed by washing two times.

Lentiviral vector preparation, infection, and titration. Lentiviral vector particles were produced by transfection of 293T cells with the plasmids mentioned above. HIV-1 proviruses were packaged with a third-generation packaging system (Hgp^{syn}, pcRev, pcTat, and pHiT-G) (27, 84, 85). pWPXL, pWPXL-Rev, CRISPR-KO constructs, and a pool of the two GeCKO v2 plasmid A and B sublibraries were packaged with a second-generation packaging system (pPAX2 and pMD2.G) (48, 86) for higher infectious titers. At 48 h after transfection, the supernatants from transfected 293T cells were cleared and then filtered through 0.45-μm filters. Viral stocks were aliquoted and stored under −80°C. Two batches of the GeCKO v2 lentiviral vector library were prepared. HiSeq NGS analyses of sgRNA representation in cells transduced with both batches showed an almost 100% identification of the sgRNAs encoded by the two GeCKO v2 plasmid sublibraries.

Jurkat, J-dual#3, and J-dual#6 cells were infected with the lentiviral vectors or the lentiviral vector library by spinoculation. A total of 1.5 million cells were infected with 1 ml vector preparation, in the presence of 8 μg/ml Polybrene (Sigma-Aldrich), at 2,000 rpm, for 3 h, at 33°C. After spinoculation, the supernatants were removed and the cells were continued to be cultured.

VSV-G-pseudotyped HIV.R$^+$e$^−$, HIV.r$^−$e$^−$, HIV.r$^−$e$^−$DsRed, and HIV-dual-GT vector particles were titrated on TZM-bl cells as previously described (63). The titers were 5.8×10^6 and 5.1×10^6 TU/ml for two independent titrations of HIV.R$^+$e$^−$, 4.3×10^6 and 5.7×10^6 TU/ml for two independent titrations of HIV.r$^−$e$^−$, 6.1×10^6 and 5.4×10^6 TU/ml for two independent titrations of HIV.r$^−$e$^−$DsRed, and 6.4×10^5 and 5.0×10^5 TU/ml for two independent titrations of HIV-dual-GT. VSV-G-pseudotyped HIV-dual-GT vector particles were also titrated on Jurkat cells. The lentiviral reporter vector was serially diluted to 1:1, 1:2, 1:4, 1:8, and 1:16. The infection was done by the spinoculation method. At 48 h postinfection, cells were analyzed by flow cytometry to determine the percentage of DsRed-positive cells in each dilution. Vector titers on Jurkat cells were calculated as follows for all dilutions: (initial cell number) × (percentage of DsRed-positive cells) × (dilution factor). The highest value resulting from these dilutions was taken as the vector titer on Jurkat cells. The determined titers for two independent titrations of HIV-dual-GT were 4.9×10^5 and 6.9×10^5 TU/ml. The individual CRISPR-KO vectors and pWPLX-Rev, as well as the GeCKO v2 lentiviral vector library, were titrated on J-dual reporter cells and parental Jurkat cells. The cells infected by the serially diluted virus were grown for 24 h and then divided into two equal aliquots. One aliquot was then cultured in 2 μg/ml puromycin, the other without puromycin. After 30 h of growth, viable cells were counted by trypan blue staining. The transduction percentage was calculated in the following way: [(number of living cells in treated population) × 100]/(number of living cells in untreated population). For the GeCKO v2 lentiviral vector library, the transduction percentage was multiplied by the number of cells exposed to calculate the titer for the different dilutions of the library. The highest value obtained from the different dilutions of the same batch of the lentiviral vector library was taken as its titer. The titers were 1.3×10^7 TU/ml for the first batch and 2.2×10^6 TU/ml for the second batch.

Flow cytometry analysis and cell sorting. Flow cytometry analyses were done using the BDLSRII flow cytometer (BD Biosciences). FACS data were analyzed by the FlowJo software. Cell sorting was done on a MoFlo XDP (Beckman Coulter) at the FACS core facility of the medical faculty.

Next-generation sequencing of sgRNA libraries. Genomic DNA was extracted from the sorted or control cells using either the Quick-DNA minikit or the Quick-DNA midi plus kit (Zymo Research). The integrated lentiCRISPRv2 vector sgRNA sequences were amplified through three rounds of PCR using the *Taq* DNA polymerase S (high specificity) (Genaxxon Bioscience). All obtained genomic DNA was used as a template in parallel reactions in PCR1, with the following cycling conditions: 94°C for 2 min, then 18 cycles of 95, 55, and 70°C for 15, 20, and 45 s, respectively, and a final extension at 70°C for 3 min. PCR1 reaction mixtures of respective cell samples were then combined, and 10 μl of the pooled PCR1 product was used as a template in PCR2, consisting of 95°C for 2 min, then 15 cycles of 95, 60, and 72°C for 15, 20, and 30 s, respectively, and a final extension at 72°C for 3 min. In the indexing PCR3, 10 ng of purified PCR2 product was subjected to 95°C for 2 min, then 8 cycles of 95, 59, and 72°C for 15, 20, and 30 s, respectively, and a final extension at 72°C for 3 min. The primer sequences and primer combinations used for each sample at each step are presented in Tables S1B and C. After PCR2 and PCR3, PCR products were purified by the AMPure XP PCR-cleanup kit (Beckman Coulter); DNA concentrations were measured by the Qubit Quant-iT dsDNA HS assay kit (Thermo Fischer Scientific). Manipulations followed the manufacturers' instructions. DNA libraries were sequenced on Illumina MiSeq (for enriched samples) or Ilumina HiSeq 2500 (for unselected control) instruments. NGS readouts were subjected to MAGeCK analysis (49).

Cell fractionation, RNA isolation, and RNA quantification. The cell fractionation protocol has been described previously (87). The plasma membrane was lysed with chilled NP-40 buffer consisting of 10 mM HEPES (pH 7.8 adjusted by KOH), 10 mM KCl, 20% glycerol, and 0.25% NP-40. Dithiothreitol (DTT) was added to a final concentration of 1 mM just before use. After incubation on ice for a maximum of 2 min, the supernatant consisting of the cytoplasmic fractions was collected after centrifugation at 400 × *g*, for 5 min, at 4°C. Cytoplasmic RNA was extracted by the TRIzol (Thermo Fisher Scientific)-chloroform (Sigma-Aldrich) method (88) or by the Direct-zol RNA miniprep plus kit (Zymo Research). RNA from the whole-cell extracts was purified using the RNeasy kit (Qiagen). RNA extracted by TRIzol-chloroform or the RNeasy kit was treated with DNase I (NEB) to remove DNA contamination. Under the protection of 5 mM EDTA, DNase I was inactivated at 75°C for 10 min. For the Direct-zol RNA miniprep plus kit-purified RNA, DNA contamination was removed by the in-column DNA digestion reagents supplied. The total RNA concentration was measured by the Qubit Quant-iT RNA HS assay kit (Thermo Fisher Scientific) according to the manufacturer's instructions.

Reverse transcription-PCRs (RT-qPCRs) for quantification of HIV-1 US and FS RNAs and cellular GAPDH pre-mRNAs were done with the QuantiTect SYBR green RT-PCR kit (Qiagen) and performed

essentially as described previously (62, 89). Using the Applied Biosystems 7500 real-time PCR machine, cycling conditions were as follows: for HIV-1 US RNA, 95°C for 10 s for denaturation, 65°C for 60 s for annealing, and 72°C for 30 s for elongation and fluorescence detection; for HIV-1 FS RNA, 95°C for 10 s for denaturation, 64°C for 15 s for annealing, 72°C for 15 s for elongation, and 81°C for 30 s for fluorescence detection.

RT-qPCRs for relative quantification of intron 6-retaining RPL10 mRNAs and intron 1-retaining C19orf53 mRNAs were done with primers given in Table S1D. Cycling conditions for both were as follows: 95°C for 10 s for denaturation, 50°C for 60 s for annealing, and 72°C for 60 s for elongation and fluorescence detection.

Calculation of the ratios of HIV FS to US RNA levels and fold changes. To measure the effect of candidate genes and Rev on RNA splicing efficacy, we first determined the ratio of FS_{whole} to US_{whole} by dividing the FS copy numbers/ng extracted RNA by the US copy numbers/ng extracted RNA for each sample of each treatment group (knockdown or +Rev) and then calculated the mean of the ratios of replicates of each treatment group. The mean of the ratios of each treatment group was then divided by the mean of the ratios observed in matched control cells transduced with the NT1 or the GFP expression vector to calculate the fold changes given in Fig. 4C.

To measure the effect of candidate genes and Rev on the nuclear export efficacy (41) for HIV-1 US RNA for each treatment group, the mean US copy numbers/ng RNA extracted from the replicates of the cytoplasmic fractions were divided by the mean US copy numbers/ng RNA extracted from replicates of the whole-cell lysates. The ratio of each treatment group was then again divided by the ratio observed in controls cells transduced with the NT1 or the GFP expression vector to calculate the fold changes given in Fig. 4C.

Fold changes for cytoplasmic levels of US or FS RNA were calculated by dividing the mean US or FS copy numbers per ng of extracted RNA of each treatment group by the mean derived from the control groups.

RNA immunoprecipitation. Nuclear fractions from 2×10^7 transfected and infected 293T cells were lysed by 600 μl of prechilled NP-40 lysis buffer (50 mM HEPES-KOH at pH 7.4, 150 mM KCl, 2 mM EDTA, 0.5% [vol/vol] NP-40, 0.5 mM DTT, Roche complete EDTA-free protease inhibitor cocktail, 40 U/ml Promega RNasin RNase inhibitor) while incubated on ice for 5 min. Lysates were cleared by centrifugation (16,000 relative centrifugal force [RCF], 4°C, 15 min). A 200-μl aliquot of the supernatant was used to quantify the amount of input RNA. Another 200-μl aliquot of the supernatant was incubated with 20 μl protein G Dynabeads (Invitrogen) precoated with 1 μg of anti-myc tag antibody or the same amount of an IgG2a isotype control antibody (Southern Biotech) respectively. RNP binding was performed overnight at 4°C. Washing was done with prechilled immunoprecipitation (IP) wash buffer (50 mM HEPES-KOH at pH 7.4, 150 mM KCl, 2 mM EDTA, 0.05% [vol/vol] NP-40, 0.5 mM DTT, complete EDTA-free protease inhibitor cocktail, 40 U/ml Promega RNasin RNase inhibitor) for six times with the aid of a magnetic rack (Invitrogen). Proteins bound to the beads were recovered from 10% of the beads by heat denaturation and characterized by Western blot analyses. RNA associated with the remaining 90% of the beads was purified by the TRIzol method. Input RNA and bead-associated RNA were analyzed by RT-qPCR.

RNA-seq. Cytoplasmic RNAs were extracted from the sorted BFP-positive or unsorted control J-dual#3 cells with the Direct-zol RNA miniprep plus kit. RNA qualities were controlled by the 2100 Bioanalyzer instrument (Agilent) (data not shown). Poly(A)-containing RNAs were enriched, fragmented, and reverse transcribed using the Illumina TruSeq stranded mRNA kit. Work flows for generating DNA libraries from each sample were performed essentially as described by the Illumina standard protocol (https://support.illumina.com/downloads/truseq-stranded-mrna-reference-guide-1000000040498.html). The libraries were sequenced on an Illumina HiSeq 2500 platform, and the reads were saved together with a quality score. Raw data were converted to reads including a quality score by Software bcl2fastq v2.17. Sequences matching Illumina TruSeq adapters were masked. Low-quality bases, poly(A) or poly(T) stretches, and masked bases were trimmed by Software fqtrim v0.9.5. Read quality was checked after sequencing and after base trimming by Software fastqc v0.11.8. For the viral genome, reads were aligned to HIV-dual-GT using hisat2 (90). Read counting for the different features was performed using samtools (91) and regtools (92).

For splicing and differential expression analysis, reads were aligned to the human transcriptome using RSEM v1.2.20 (93) and Bowtie v1.1.2 (93) as the read alignment program using default parameters. Differential expression analysis was performed by DESeq2 v1.18.1 (93) using the RSEM output for read counts per gene. We considered genes that showed an absolute $\log_2$-transformed fold change greater than 1 and an adjusted P value (P_{adj}) of less than 0.1 as differentially expressed. To exclude low-expression genes, we considered only those with a mean number of transcripts per kilobase million (TPM) greater than 1 in either untreated or knockdown samples. Gencode v19 annotation was used in all steps of the analysis, including in the definition of protein coding genes. The read mapping profiles were exhibited via the Integrative Genomics Viewer (IGV) v2.3 (94). Differential splicing analysis was performed with SUPPA (71) using TPM values from the RSEM transcript isoform output. For PSI calculation, events were required to pass the expression filter of TPM greater than 1 (-f 1 parameter). We considered splicing events with P_{adj} values lower than 0.01 and absolute ΔPSI values greater than 0.25 as differentially included or excluded. Polypyrimidine track length and branch point scores were calculated using BPfinder (95). Splice site strength was calculated using MaxENT (96) (https://hollywood.mit.edu/burgelab/software.html).

Data availability. The raw and processed data of RNA-seq are accessible via GEO (Gene Expression Omnibus) accession number GSE144347.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 1.2 MB.

FIG S2, TIF file, 1.9 MB.

FIG S3, TIF file, 0.6 MB.

FIG S4, TIF file, 0.2 MB.

FIG S5, TIF file, 0.2 MB.

FIG S6, TIF file, 0.4 MB.

FIG S7, TIF file, 0.8 MB.

FIG S8, TIF file, 0.7 MB.

TABLE S1, TIF file, 1.5 MB.

TABLE S2, XLSX file, 0.2 MB.

REFERENCES

1. Jacob AG, Smith CWJ. 2017. Intron retention as a component of regulated gene expression programs. Hum Genet 136:1043–1057. https://doi.org/10.1007/s00439-017-1791-x.

2. Braunschweig U, Barbosa-Morais NL, Pan Q, Nachman EN, Alipanahi B, Gonatopoulos-Pournatzis T, Frey B, Irimia M, Blencowe BJ. 2014. Widespread intron retention in mammals functionally tunes transcriptomes. Genome Res 24:1774–1786. https://doi.org/10.1101/gr.177790.114.

3. Zhang S-J, Wang C, Yan S, Fu A, Luan X, Li Y, Shen QS, Zhong X, Chen J-Y, Wang X, Tan BC-M, He A, Li C-Y. 2017. Isoform evolution in primates through independent combination of alternative RNA processing events. Mol Biol Evol 34:2453–2468. https://doi.org/10.1093/molbev/msx212.

4. Hammarskjold M-L, Rekosh D. 2017. SR proteins: to shuttle or not to shuttle, that is the question. J Cell Biol 216:1875–1877. https://doi.org/10.1083/jcb.201705009.

5. Galganski L, Urbanek MO, Krzyzosiak WJ. 2017. Nuclear speckles: molecular organization, biological function and role in disease. Nucleic Acids Res 45:10350–10368. https://doi.org/10.1093/nar/gkx759.

6. Cochrane AW, McNally MT, Mouland AJ. 2006. The retrovirus RNA trafficking granule: from birth to maturity. Retrovirology 3:18. https://doi.org/10.1186/1742-4690-3-18.

7. Ahmad N, Maitra RK, Venkatesan S. 1989. Rev-induced modulation of Nef protein underlies temporal regulation of human immunodeficiency virus replication. Proc Natl Acad Sci U S A 86:6111–6115. https://doi.org/10.1073/pnas.86.16.6111.

8. Karn J, Stoltzfus CM. 2012. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. Cold Spring Harb Perspect Med 2:a006916. https://doi.org/10.1101/cshperspect.a006916.

9. Sauter D, Hotter D, Van Driessche B, Stürzel CM, Kluge SF, Wildum S, Yu H, Baumann B, Wirth T, Plantier J-C, Leoz M, Hahn BH, Van Lint C, Kirchhoff F. 2015. Differential regulation of NF-κB-mediated proviral and antiviral host gene expression by primate lentiviral Nef and Vpu proteins. Cell Rep 10:586–599. https://doi.org/10.1016/j.celrep.2014.12.047.

10. Lemieux C, Marguerat S, Lafontaine J, Barbezier N, Bähler J, Bachand F. 2011. A pre-mRNA degradation pathway that selectively targets intron-containing genes requires the nuclear poly(A)-binding protein. Mol Cell 44:108–119. https://doi.org/10.1016/j.molcel.2011.06.035.

11. Sandri-Goldin RM. 2004. Viral regulation of mRNA export. J Virol 78:4389–4396. https://doi.org/10.1128/jvi.78.9.4389-4396.2004.

12. Cullen BR. 2000. Nuclear RNA export pathways. Mol Cell Biol 20:4181–4187. https://doi.org/10.1128/mcb.20.12.4181-4187.2000.

13. Grewe B, Uberla K. 2010. The human immunodeficiency virus type 1 Rev protein: menage a trois during the early phase of the lentiviral replication cycle. J Gen Virol 91:1893–1897. https://doi.org/10.1099/vir.0.022509-0.

14. Mousseau G, Valente ST. 2017. Role of host factors on the regulation of Tat-mediated HIV-1 transcription. Curr Pharm Des 23:4079–4090. https://doi.org/10.2174/1381612823666170622104355.

15. Suhasini M, Reddy TR. 2009. Cellular proteins and HIV-1 Rev function. Curr HIV Res 7:91–100. https://doi.org/10.2174/157016209787048474.

16. Katahira J. 2012. mRNA export and the TREX complex. Biochim Biophys Acta 1819:507–513. https://doi.org/10.1016/j.bbagrm.2011.12.001.

17. Bray M, Prasad S, Dubay JW, Hunter E, Jeang KT, Rekosh D, Hammarskjöld ML. 1994. A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. Proc Natl Acad Sci U S A 91:1256–1260. https://doi.org/10.1073/pnas.91.4.1256.

18. Malim MH, Cullen BR. 1993. Rev and the fate of pre-mRNA in the nucleus: implications for the regulation of RNA processing in eukaryotes. Mol Cell Biol 13:6180–6189. https://doi.org/10.1128/mcb.13.10.6180.

19. Pollard VW, Malim MH. 1998. The HIV-1 Rev protein. Annu Rev Microbiol 52:491–532. https://doi.org/10.1146/annurev.micro.52.1.491.

20. Cochrane AW, Jones KS, Beidas S, Dillon PJ, Skalka AM, Rosen CA. 1991. Identification and characterization of intragenic sequences which repress human immunodeficiency virus structural gene expression. J Virol 65:5305–5313. https://doi.org/10.1128/JVI.65.10.5305-5313.1991.

21. Mikaélian I, Krieg M, Gait MJ, Karn J. 1996. Interactions of INS (CRS) elements and the splicing machinery regulate the production of Rev-responsive mRNAs. J Mol Biol 257:246–264. https://doi.org/10.1006/jmbi.1996.0160.

22. Brighty DW, Rosenberg M, Steitz JA. 1994. A cis-acting repressive sequence that overlaps the Rev-responsive element of human immunodeficiency virus type 1 regulates nuclear retention of env mRNAs independently of known splice signals. Proc Natl Acad Sci U S A 91:8314–8318. https://doi.org/10.1073/pnas.91.18.8314.

23. Black A, Luo J, Chun S, Bakker A, Fraser J, Rosenblatt J. 1996. Specific binding of polypyrimidine tract binding protein and hnRNP A1 to HIV-1 CRS elements. Virus Genes 12:275–285. https://doi.org/10.1007/BF00284648.

24. Maldarelli F, Martin MA, Strebel K. 1991. Identification of posttranscriptionally active inhibitory sequences in human immunodeficiency virus type 1 RNA: novel level of gene regulation. J Virol 65:5732–5743. https://doi.org/10.1128/JVI.65.11.5732-5743.1991.

25. Haas D, Park E-C, Seed B. 1996. Codon usage limitation in the expression of HIV-1 envelope glycoprotein. Curr Biol 6:315–324. https://doi.org/10.1016/s0960-9822(02)00482-7.

26. Graf M, Bojak A, Deml L, Bieler K, Wolf H, Wagner R. 2000. Concerted action of multiple cis-acting sequences is required for Rev dependence of late human immunodeficiency virus type 1 gene expression. J Virol 74:10822–10826. https://doi.org/10.1128/jvi.74.22.10822-10826.2000.

27. Wagner R, Graf M, Bieler K, Wolf H, Grunwald T, Foley P, Überla K. 2000. Rev-independent expression of synthetic *gag-pol* genes of human immunodeficiency virus type 1 and simian immunodeficiency virus: implications for the safety of lentiviral vectors. Hum Gene Ther 11:2403–2413. https://doi.org/10.1089/104303400750038507.

28. Zhang Q, Chen C-Y, Yedavalli VSRK, Jeang K-T. 2013. NEAT1 long noncoding RNA and paraspeckle bodies modulate HIV-1 posttranscriptional expression. mBio 4:e00596-12. https://doi.org/10.1128/mBio.00596-12.

29. Kula A, Gharu L, Marcello A. 2013. HIV-1 pre-mRNA commitment to Rev mediated export through PSF and Matrin 3. Virology 435:329–340. https://doi.org/10.1016/j.virol.2012.10.032.

30. Zolotukhin AS, Michalowski D, Bear J, Smulevitch SV, Traish AM, Peng R, Patton J, Shatsky IN, Felber BK. 2003. PSF acts through the human immunodeficiency virus type 1 mRNA instability elements to regulate virus expression. Mol Cell Biol 23:6618–6630. https://doi.org/10.1128/mcb.23.18.6618-6630.2003.

31. Yedavalli VSRK, Jeang K-T. 2011. Matrin 3 is a co-factor for HIV-1 Rev in regulating post-transcriptional viral gene expression. Retrovirology 8:61. https://doi.org/10.1186/1742-4690-8-61.

32. Fong K-W, Li Y, Wang W, Ma W, Li K, Qi RZ, Liu D, Songyang Z, Chen J. 2013. Whole-genome screening identifies proteins localized to distinct nuclear bodies. J Cell Biol 203:149–164. https://doi.org/10.1083/jcb.201303145.

33. Kula A, Guerra J, Knezevich A, Kleva D, Myers MP, Marcello A. 2011. Characterization of the HIV-1 RNA associated proteome identifies Matrin 3 as a nuclear cofactor of Rev function. Retrovirology 8:60. https://doi.org/10.1186/1742-4690-8-60.

34. Sarracino A, Gharu L, Kula A, Pasternak AO, Avettand-Fenoel V, Rouzioux C, Bardina M, De Wit S, Benkirane M, Berkhout B, Van Lint C, Marcello A. 2018. Posttranscriptional regulation of HIV-1 gene expression during replication and reactivation from latency by nuclear matrix protein MATR3. mBio 9:e02158-18. https://doi.org/10.1128/mBio.02158-18.

35. Gordon H, Ajamian L, Valiente-Echeverrìa F, Lévesque K, Rigby WF, Mouland AJ. 2013. Depletion of hnRNP A2/B1 overrides the nuclear retention of the HIV-1 genomic RNA. RNA Biol 10:1714–1725. https://doi.org/10.4161/rna.26542.

36. Dziembowski A, Ventura A-P, Rutz B, Caspary F, Faux C, Halgand F, Laprévote O, Séraphin B. 2004. Proteomic analysis identifies a new complex required for nuclear pre-mRNA retention and splicing. EMBO J 23:4847–4856. https://doi.org/10.1038/sj.emboj.7600482.

37. Frankiw L, Majumdar D, Burns C, Vlach L, Moradian A, Sweredoski MJ, Baltimore D. 2019. BUD13 promotes a type I interferon response by countering intron retention in Irf7. Mol Cell 73:803–814.e6. https://doi.org/10.1016/j.molcel.2018.11.038.

38. Amit M, Donyo M, Hollander D, Goren A, Kim E, Gelfman S, Lev-Maor G, Burstein D, Schwartz S, Postolsky B, Pupko T, Ast G. 2012. Differential GC content between exons and introns establishes distinct strategies of splice-site recognition. Cell Rep 1:543–556. https://doi.org/10.1016/j.celrep.2012.03.013.

39. Fernandez JP, Moreno-Mateos MA, Gohr A, Miao L, Chan SH, Irimia M, Giraldez AJ. 2018. RES complex is associated with intron definition and required for zebrafish early embryogenesis. PLoS Genet 14:e1007473. https://doi.org/10.1371/journal.pgen.1007473.

40. Galy V, Gadal O, Fromont-Racine M, Romano A, Jacquier A, Nehrbass U. 2004. Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. Cell 116:63–73. https://doi.org/10.1016/s0092-8674(03)01026-2.

41. Coyle JH, Bor Y-C, Rekosh D, Hammarskjold M-L. 2011. The Tpr protein regulates export of mRNAs with retained introns that traffic through the Nxf1 pathway. RNA 17:1344–1356. https://doi.org/10.1261/rna.2616111.

42. Bonnet A, Bretes H, Palancade B. 2015. Nuclear pore components affect distinct stages of intron-containing gene expression. Nucleic Acids Res 43:4249–4261. https://doi.org/10.1093/nar/gkv280.

43. Zander G, Hackmann A, Bender L, Becker D, Lingner T, Salinas G, Krebber H. 2016. mRNA quality control is bypassed for immediate export of stress-responsive transcripts. Nature 540:593–596. https://doi.org/10.1038/nature20572.

44. Blanco R, Carrasco L, Ventoso I. 2003. Cell killing by HIV-1 protease. J Biol Chem 278:1086–1093. https://doi.org/10.1074/jbc.M205636200.

45. Malim MH, Emerman M. 2008. HIV-1 accessory proteins—ensuring viral survival in a hostile environment. Cell Host Microbe 3:388–398. https://doi.org/10.1016/j.chom.2008.04.008.

46. Stewart SA, Poon B, Jowett JB, Xie Y, Chen IS. 1999. Lentiviral delivery of HIV-1 Vpr protein induces apoptosis in transformed cells. Proc Natl Acad Sci U S A 96:12039–12043. https://doi.org/10.1073/pnas.96.21.12039.

47. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, Heckl D, Ebert BL, Root DE, Doench JG, Zhang F. 2014. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343:84–87. https://doi.org/10.1126/science.1247005.

48. Sanjana NE, Shalem O, Zhang F. 2014. Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods 11:783–784. https://doi.org/10.1038/nmeth.3047.

49. 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.

50. Shifrut E, Carnevale J, Tobin V, Roth TL, Woo JM, Bui CT, Li PJ, Diolaiti ME, Ashworth A, Marson A. 2018. Genome-wide CRISPR screens in primary human T cells reveal key regulators of immune function. Cell 175:1958–1971.e15. https://doi.org/10.1016/j.cell.2018.10.024.

51. Takada I, Tsuchiya H, Hidano S, Takahashi S, Kobayashi T, Ogawa H, Nakagawa S, Makishima M. 2018. Ess2 bridges transcriptional regulators and spliceosomal complexes via distinct interacting domains. Biochem Biophys Res Commun 497:597–604. https://doi.org/10.1016/j.bbrc.2018.02.110.

52. Schapira M, Tyers M, Torrent M, Arrowsmith CH. 2017. WD40 repeat domain proteins: a novel target class? Nat Rev Drug Discov 16:773–786. https://doi.org/10.1038/nrd.2017.179.

53. Yan C, Wan R, Bai R, Huang G, Shi Y. 2017. Structure of a yeast step II catalytically activated spliceosome. Science 355:149–155. https://doi.org/10.1126/science.aak9979.

54. Zhan X, Yan C, Zhang X, Lei J, Shi Y. 2018. Structure of a human catalytic step I spliceosome. Science 359:537–545. https://doi.org/10.1126/science.aar6401.

55. Hou S, Qu D, Li Y, Zhu B, Liang D, Wei X, Tang W, Zhang Q, Hao J, Guo W, Wang W, Zhao S, Wang Q, Azam S, Khan M, Zhao H, Zhang L, Lei H. 2019. XAB2 depletion induces intron retention in POLR2A to impair global transcription and promote cellular senescence. Nucleic Acids Res 47:8239–8254. https://doi.org/10.1093/nar/gkz532.

56. Zhang X, Yan C, Hang J, Finci LI, Lei J, Shi Y. 2017. An atomic structure of the human spliceosome. Cell 169:918–929.e14. https://doi.org/10.1016/j.cell.2017.04.033.

57. Chanarat S, Sträßer K. 2013. Splicing and beyond: the many faces of the Prp19 complex. Biochim Biophys Acta 1833:2126–2134. https://doi.org/10.1016/j.bbamcr.2013.05.023.

58. Grainger RJ, Beggs JD. 2005. Prp8 protein: at the heart of the spliceosome. RNA 11:533–557. https://doi.org/10.1261/rna.2220705.

59. 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.

60. Reed R. 2003. Coupling transcription, splicing and mRNA export. Curr Opin Cell Biol 15:326–331. https://doi.org/10.1016/s0955-0674(03)00048-6.

61. Valencia P, Dias AP, Reed R. 2008. Splicing promotes rapid and efficient mRNA export in mammalian cells. Proc Natl Acad Sci U S A 105:3386–3391. https://doi.org/10.1073/pnas.0800250105.

62. Blissenbach M, Grewe B, Hoffmann B, Brandt S, Uberla K. 2010. Nuclear RNA export and packaging functions of HIV-1 Rev revisited. J Virol 84:6598–6604. https://doi.org/10.1128/JVI.02264-09.

63. Grewe B, Hoffmann B, Ohs I, Blissenbach M, Brandt S, Tippler B, Grunwald T, Uberla K. 2012. Cytoplasmic utilization of human immunodeficiency virus type 1 genomic RNA is not dependent on a nuclear interaction with Gag. J Virol 86:2990–3002. https://doi.org/10.1128/JVI.06874-11.

64. Taniguchi I, Mabuchi N, Ohno M. 2014. HIV-1 Rev protein specifies the viral RNA export pathway by suppressing TAP/NXF1 recruitment. Nucleic Acids Res 42:6645–6658. https://doi.org/10.1093/nar/gku304.

65. Edenfeld G, Volohonsky G, Krukkert K, Naffin E, Lammel U, Grimm A, Engelen D, Reuveny A, Volk T, Klämbt C. 2006. The splicing factor crooked neck associates with the RNA-binding protein HOW to control glial cell maturation in Drosophila. Neuron 52:969–980. https://doi.org/10.1016/j.neuron.2006.10.029.

66. Hammani K, Cook WB, Barkan A. 2012. RNA binding and RNA remodeling activities of the half-a-tetratricopeptide (HAT) protein HCF107 underlie its effects on gene expression. Proc Natl Acad Sci U S A 109:5651–5656. https://doi.org/10.1073/pnas.1200318109.

67. Chung S, Zhou Z, Huddleston KA, Harrison DA, Reed R, Coleman TA, Rymond BC. 2002. Crooked neck is a component of the human spliceosome and implicated in the splicing process. Biochim Biophys Acta 1576:287–297. https://doi.org/10.1016/S0167-4781(02)00368-8.

68. Fica SM, Oubridge C, Wilkinson ME, Newman AJ, Nagai K. 2019. A human postcatalytic spliceosome structure reveals essential roles of metazoan factors for exon ligation. Science 363:710–714. https://doi.org/10.1126/science.aaw5569.

69. Liu M, Lu Y, Yang B, Chen Y, Radda JSD, Hu M, Katz SG, Wang S. 2020. Multiplexed imaging of nucleome architectures in single cells of mammalian tissue. Nat Commun 11:2907. https://doi.org/10.1038/s41467-020-16732-5.

70. Sithole N, Williams CA, Vaughan AM, Kenyon JC, Lever AML. 2018. DDX17 specifically, and independently of DDX5, controls use of the HIV A4/5 splice acceptor cluster and is essential for efficient replication of HIV. J Mol Biol 430:3111–3128. https://doi.org/10.1016/j.jmb.2018.06.052.

71. Alamancos GP, Pagès A, Trincado JL, Bellora N, Eyras E. 2015. Leveraging transcript quantification for fast computation of alternative splicing profiles. RNA 21:1521–1531. https://doi.org/10.1261/rna.051557.115.

72. Ren L, McLean JR, Hazbun TR, Fields S, Vander Kooi C, Ohi MD, Gould KL. 2011. Systematic two-hybrid and comparative proteomic analyses reveal novel yeast pre-mRNA splicing factors connected to Prp19. PLoS One 6: e16719. https://doi.org/10.1371/journal.pone.0016719.

73. Katahira J, Yoneda Y. 2009. Roles of the TREX complex in nuclear export of mRNA. RNA Biol 6:149–152. https://doi.org/10.4161/rna.6.2.8046.

74. Takemura R, Takeiwa T, Taniguchi I, McCloskey A, Ohno M. 2011. Multiple factors in the early splicing complex are involved in the nuclear retention of pre-mRNAs in mammalian cells. Genes Cells 16:1035–1049. https://doi.org/10.1111/j.1365-2443.2011.01548.x.

75. Wysoczanski P, Zweckstetter M. 2016. Retention and splicing complex (RES)—the importance of cooperativity. RNA Biol 13:128–133. https://doi.org/10.1080/15476286.2015.1096484.

76. Karamysheva Z, Díaz-Martínez LA, Warrington R, Yu H. 2015. Graded requirement for the spliceosome in cell cycle progression. Cell Cycle 14:1873–1883. https://doi.org/10.1080/15384101.2015.1039209.

77. Zhang K, Smouse D, Perrimon N. 1991. The crooked neck gene of Drosophila contains a motif found in a family of yeast cell cycle genes. Genes Dev 5:1080–1091. https://doi.org/10.1101/gad.5.6.1080.

78. Chanarat S, Seizl M, Strässer K. 2011. The Prp19 complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes. Genes Dev 25:1147–1158. https://doi.org/10.1101/gad.623411.

79. Lei H, Zhai B, Yin S, Gygi S, Reed R. 2013. Evidence that a consensus element found in naturally intronless mRNAs promotes mRNA export. Nucleic Acids Res 41:2517–2525. https://doi.org/10.1093/nar/gks1314.

80. Biswas P, Jiang X, Pacchia AL, Dougherty JP, Peltz SW. 2004. The human immunodeficiency virus type 1 ribosomal frameshifting site is an invariant sequence determinant and an important target for antiviral therapy. J Virol 78:2082–2087. https://doi.org/10.1128/jvi.78.4.2082-2087.2004.

81. 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.

82. Platt EJ, Wehrly K, Kuhmann SE, Chesebro B, Kabat D. 1998. Effects of CCR5 and CD4 cell surface concentrations on infections by macrophage-tropic isolates of human immunodeficiency virus type 1. J Virol 72:2855–2864. https://doi.org/10.1128/JVI.72.4.2855-2864.1998.

83. de Jong SJ, Albrecht J-C, Schmidt M, Müller-Fleckenstein I, Biesinger B. 2010. Activation of noncanonical NF-kappaB signaling by the oncoprotein Tio. J Biol Chem 285:16495–16503. https://doi.org/10.1074/jbc.M110.102848.

84. Malim MH, Hauber J, Fenrick R, Cullen BR. 1988. Immunodeficiency virus rev trans-activator modulates the expression of the viral regulatory genes. Nature 335:181–183. https://doi.org/10.1038/335181a0.

85. Fouchier RA, Meyer BE, Simon JH, Fischer U, Malim MH. 1997. HIV-1 infection of non-dividing cells: evidence that the amino-terminal basic region of the viral matrix protein is important for Gag processing but not for post-entry nuclear import. EMBO J 16:4531–4539. https://doi.org/10.1093/emboj/16.15.4531.

86. Naldini L, Trono D, Verma IM. 2016. Lentiviral vectors, two decades later. Science 353:1101–1102. https://doi.org/10.1126/science.aah6192.

87. Thomas M, Sonntag E, Müller R, Schmidt S, Zielke B, Fossen T, Stamminger T. 2015. pUL69 of human cytomegalovirus recruits the cellular protein arginine methyltransferase 6 via a domain that is crucial for mRNA export and efficient viral replication. J Virol 89:9601–9615. https://doi.org/10.1128/JVI.01399-15.

88. Rio DC, Ares M, Hannon GJ, Nilsen TW. 2010. Purification of RNA using TRIzol (TRI reagent). Cold Spring Harb Protoc 2010:pdb.prot5439. https://doi.org/10.1101/pdb.prot5439.

89. Brandt S, Blissenbach M, Grewe B, Konietzny R, Grunwald T, Uberla K. 2007. Rev proteins of human and simian immunodeficiency virus enhance RNA encapsidation. PLoS Pathog 3:e54. https://doi.org/10.1371/journal.ppat.0030054.

90. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37:907–915. https://doi.org/10.1038/s41587-019-0201-4.

91. Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987–2993. https://doi.org/10.1093/bioinformatics/btr509.

92. Feng Y-Y, Ramu A, Cotto KC, Skidmore ZL, Kunisaki J, Conrad DF, Lin Y, Chapman WC, Uppaluri R, Govindan R, Griffith OL, Griffith M. 2018. RegTools: integrated analysis of genomic and transcriptomic data for discovery of splicing variants in cancer. bioRxiv https://doi.org/10.1101/436634.

93. Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323. https://doi.org/10.1186/1471-2105-12-323.

94. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. Nat Biotechnol 29:24–26. https://doi.org/10.1038/nbt.1754.

95. Corvelo A, Hallegger M, Smith CWJ, Eyras E. 2010. Genome-wide association between branch point properties and alternative splicing. PLoS Comput Biol 6:e1001016. https://doi.org/10.1371/journal.pcbi.1001016.

96. Yeo G, Burge CB. 2004. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol 11:377–394. https://doi.org/10.1089/1066527041410418.

97. Schafer S, Miao K, Benson CC, Heinig M, Cook SA, Hubner N. 2015. Alternative splicing signatures in RNA-seq data: percent spliced in (PSI). Curr Protoc Hum Genet 87:11.16.1–11.16.14. https://doi.org/10.1002/0471142905.hg1116s87.

98. Budhiraja S, Liu H, Couturier J, Malovannaya A, Qin J, Lewis DE, Rice AP. 2015. Mining the human complexome database identifies RBM14 as an XPO1-associated protein involved in HIV-1 Rev function. J Virol 89:3557–3567. https://doi.org/10.1128/JVI.03232-14.