HIV-induced membraneless organelles orchestrate post-nuclear entry steps.

Viviana Scoca[1, 2], Marion Louveaux*[3, 4], Renaud Morin*[5], Dmitry Ershov[4, 6], Jean-Yves Tinevez[4], Francesca Di Nunzio[1]

[1]Department of Virology, Advanced Molecular Virology and Retroviral Dynamics Group, Pasteur Institute, Paris, France
[2] BioSPC Doctoral School, Université de Paris, Paris, France
[3] Department of Cell Biology and Infection, Biological Image Analysis, Pasteur Institute, Paris, France
[4] Image Analysis Hub/ C2RT, Pasteur Institute, Paris, France
[5] Imactiv-3D, 1 Place Pierre Potier, 31106 Toulouse, France
[6] Department of Computational Biology, Hub of Bioinformatics and Biostatistics, Pasteur Institute, USR 3756 CNRS, Paris, France.

*****equal contribution

Running title: Deciphering the host nucleus subversion by HIV-1 to replicate.

Abstract

HIV integration occurs in chromatin sites that favor the release of high levels of viral progeny, alternatively the virus is also able to coexist discreetly with the host. To uncover the HIV-1 DNA fate in the nuclear landscape, we directly tracked the viral DNA and the viral RNA by coupling HIV-1 ANCHOR technology with RNA FISH or MCP-MS2 RNA-tagging bacterial system. Our computational imaging analysis revealed that proviral forms are early located in proximity of the nuclear periphery of mitotic and non-mitotic cells. HIV-1 triggers the nuclear clustering of the host factor CPSF6, but yet its role is poorly understood. Our data show that CPSF6 clusters are part of HIV-induced membraneless organelles (HIV-1 MLOs). Interestingly, we observed that viral genomic RNAs and the integrase proteins are sequestered in CPSF6 clusters, in which the virus can reverse transcribe, while the late retrotranscribed DNA is excluded from these structures. HIV-1 MLOs are formed in the interchromatin space, indicating that they are not proviral sites, but orchestrate viral events prior to the integration step. We found transcription-competent proviruses localize outside HIV-1 MLOs, in LEDGF-abundant regions, known to be active chromatin sites. This study highlights single functional host-proviral complexes in their nuclear landscape, which is markedly restructured by HIV-1 to favor viral replication.

Introduction

Immediately after fusion at the plasma membrane, HIV-1 cores are released into the cytoplasm and move towards the nucleus, while the viral RNA (vRNA) genome begins the process of reverse transcription (RT) into double-stranded DNA (dsDNA) (Campbell and Hope, 2015; Di Nunzio, 2013; Scoca and Di Nunzio, 2021a). Once imported in the nucleus, the interplay between the HIV-1 DNA genome and the host chromatin compartment is crucial for the fate of the virus-host coexistence. HIV-1 can adopt either an episomal or a proviral form (Butler et al., 2001). The fate of these viral DNA (vDNA) forms and their surrounding chromatin landscape dictate the evolution of HIV infection (Maldarelli, 2016; Sharkey et al., 2011; Sharkey et al., 2000). The mechanisms determining the fate of the proviral DNA in the nucleus are still under investigation (Liu et al., 2020; Olson et al., 2019) also because of the limits imposed by available technologies. Real-time imaging approaches provide new insight into unprecedented spatial information of individual HIV-1 infected cells. Several strategies have enabled the visualization of the viral components during nuclear entry to indirectly track the HIV-1 pre-integration complex (PIC) : Integrase (IN)-TC/FlAsH, APOBEC 3F (A3F)-YFP, IN-YFP, Gag-iGFP, Cyclophilin A (CypA)-DsRed/ Capsid (CA), INsfGFP, and CA-GFP (Burdick et al., 2017; Hubner et al., 2009; Lelek et al., 2012) (Francis et al., 2016; Francis and Melikyan, 2018) (Mamede et al., 2017) (Albanese et al., 2008; Francis et al., 2014). Recent new discoveries in the field shed light on the central role of the viral capsid in leading the viral genome in the host nucleus (Blanco-Rodriguez et al., 2020; Burdick et al., 2020), contrary to what was widely believed (Farnet and Haseltine, 1991; Suzuki and Craigie, 2007). The viral nuclear translocation is dictated by the capsid interplay with host factors, such as Cleavage and polyadenylation specificity factor subunit 6 (CPSF6)(Buffone et al., 2018; Lee et al., 2010; Price et al., 2014), which is a paraspeckle factor (Fox et al., 2002; Naganuma and Hirose, 2013) and Nup153 (Di Nunzio et al., 2013; Lelek et al., 2015). CPSF6, vDNA, vRNA and transcription factors have been found in nuclear speckles (NSs), suggesting that their co-appearance is consistent with productive integration in these nuclear bodies (Francis et al., 2020; Stultz et al., 2017). Interestingly, CPSF6 contains intrinsically disordered mixed-charge domains which are responsible for the formation of liquid condensates *in vitro* (Greig et al., 2020) and they may account for the remodeling of membraneless organelles (MLOs) in infected cells (Scoca and Di Nunzio, 2021b).

However, although NSs contain viral components they are by definition interchromatin granules (Lamond and Spector, 2003), thus they cannot be HIV-1 integration sites. Thus, the specific role of these host/viral nuclear organelles remain unclear. Of note, mechanisms underlying viral post-nuclear entry steps and the role of nuclear structures induced by HIV-1 during viral replication could be revealed only by the direct co-labeling of vDNA forms and host factors. Thus far, all vDNA forms from the early to the late RT products can be indistinguishably visualized in fixed cells, either by supplying 5-Ethynyl-2´-deoxyuridine (EdU) during infection (Peng et al., 2014; Rensen et al., 2020), which is limited to non-dividing cells (De Wit et al., 2019) or by DNA-FISH (Marini et al., 2015) or bDNA (Chin et al., 2015). However, the exclusive labeling of late reverse transcribed DNA products in fixed and in live cells is challenging but important, because they are the only vDNA forms included in the mature PIC.

Here, we were able to specifically live track the late reverse transcribed DNA as episomal or as proviral forms, showing their divergent intranuclear behavior. In parallel, we detect the location of functional proviruses in the nuclear space of the main target cells of the virus, such as $CD4^{+}$T cells and macrophages, by using HIV-1 ANCHOR methodology (Blanco-Rodriguez et al., 2020) to label the vDNA and RNA FISH or the MCP-MS2 system (Tantale et al., 2016) to identify the vRNA foci. Next, we deepened the study of the nuclear landscape that surrounds proviruses. We observed that viral infection reprogramed the nuclear location of CPSF6, building HIV-1 membraneless organelles (HIV-1 MLOs). Our data emphasize post nuclear entry steps finely regulated by the interplay between viral and host components. In particular, we investigated whether HIV-1 MLOs could be physiological sites of nuclear reverse transcription, as well as maturation centers of HIV-1 PIC. We found that HIV-1 MLOs are enlarged NSs and constitute microenvironments favorable for HIV-1 PIC maturation. In addition, we observed foci of viral replication excluded but in the vicinity of HIV-1 MLOs and of the nuclear envelope (NE). Of note, our results indicate that HIV-1 DNA localization is key to identifying partners of HIV-1 that aid the virus to hijack cellular mechanisms to persist in the host cells and to highlight the nuclear landscape surrounding single viral genomes. Our study provides new insights into how HIV reprograms and markedly restructures the nuclear environment to orchestrate viral replication steps.

Results

Live tracking of HIV-1 DNA forms in the nucleus of target cells.

With the aim of studying HIV-1 DNA nuclear fate in different cell types, we benefited of the HIV-1 ANCHOR DNA labeling system (Fig. 1A) (Blanco-Rodriguez et al., 2020). Once fully retrotranscribed, the HIV-1 genome carrying the bacterial sequence ANCH3 is specifically recognized by OR-GFP protein, which is expressed in the target cells, genetically modified by the lentiviral vector (LV) OR-GFP (Blanco-Rodriguez et al., 2020). The accumulation of OR-GFP, which is a modified version of the bacterial ParB protein (Graham et al., 2014; Sanchez et al., 2015), on the ANCH3 sequence generates bright nuclear signals (Fig.1A). Importantly, OR protein binds exclusively the double stranded DNA (Saad et al., 2014). We cloned ANCH3 in the nef gene, thus, HIV-1 ANCHOR system exclusively detects late reverse transcripts, because the ANCH3 sequence is one of the last sequences to be converted into double stranded DNA (Fig. 1A). Therefore, the HIV-1 ANCHOR approach allows only the visualization of potential functional vDNA. Since the HIV-1 genome dissociates from viral proteins to become a proviral DNA or an extrachromosomal DNA, the track of these viral forms is the only direct measurement to investigate the fate of the viral genome into the host nuclear space. Unlike previous studies (Burdick et al., 2017; Burdick et al., 2020; Chin et al., 2015; Francis et al., 2020; Francis and Melikyan, 2018), we are now able to directly visualize by live imaging the nuclear vDNA as GFP puncta appearing in the host nucleus once retrotranscribed (Fig. 1A). HIV-1 ANCHOR is able to specifically label the vDNA as shown by the correlation between the nuclear GFP puncta and the MOI (multiplicity of infection) used (Fig. S1A). This specificity was corroborated using the reverse transcription inhibitor, Nevirapine (NEV), which completely prevents the detection of vDNA (Fig. 1B, Fig.S1B). We then investigated whether HIV-1 ANCHOR allows the detection of different intranuclear vDNAs, such as episomal and integrated forms. To generate episomal forms, either we infected cells with HIV-1 ANCH3 in the presence of the integration inhibitor Raltegravir (RAL), or using an integration-deficient virus, which is mutated at the catalytic site of the integrase (IN) enzyme (HIV-1 ANCH3 IN_{D116A}). We detected intranuclear GFP puncta under both conditions (Fig. 1B), meaning that HIV-1 ANCHOR permits the detection of nuclear episomal vDNA forms. The experiments were validated by qPCR, which confirmed the extrachromosomal forms generated during infection when the integration step is impeded (Fig. 1B). We

next assessed whether HIV-1 ANCHOR is sufficiently sensitive to allow the detection of integrated vDNA at the level of a single provirus. We selected single HeLa P4R5 clones carrying the HIV-1 ANCH3 genome. The imaging data correlates with results obtained by ALU PCR (Fig. 1C), demonstrating that this technology is powerful enough to detect a single provirus. Thus, we used this system to live track both proviruses and episomal forms to investigate their behaviour in the nucleus (video 1, 2, and 3). We performed live tracking of vDNAs in cells at six days post infection (p.i.) (video 1), which contain a negligible amount of episomal forms, likely lost during cell division (Fig. S1C), in the single-cell clone (video 2) or in cells containing only the episomal vDNA (video 3). Then, we compared vDNA trajectories in the aforementioned experimental conditions (Fig. 1D). The stable clone and the cells at six days p.i. show vDNA with similar trajectories and diffusion coefficient (D (nm^2/s) < 5). Instead, infected cells carrying unintegrated viruses harbored two populations composed of episomal forms with divergent behaviors, one with a low D (D (nm^2/s) < 5) and the other with a high D (5 < D (nm^2/s) < 20) (Fig. 1D). Our data hint that most, but not all, episomal forms are retained in particular nuclear regions, most likely interacting with host factors or chromatin compartments, consistent with the findings of other studies (Geis and Goff, 2019; Zhu et al., 2018), whereas some episomal forms are free to move in the nuclear space. On the other hand, the integrated forms showed a uniform, low-diffusive behavior.

Detection of actively transcribing proviruses at single-cell level in HIV-1 target cells.

The identification of the functional vDNA in the nucleus of infected cells has been one of our goals. For this purpose, we combined HIV-1 ANCHOR technology with RNA FISH (single molecule inexpensive FISH, smiFISH) against HIV-1 polymerase transcripts, to detect vRNA foci of transcription (Tsanov et al., 2016), using ethanol treatment and 5h of probe hybridization (Fig. S4). As expected, in cells treated with RAL, where the integration step is blocked, vRNA signal is highly depleted. In untreated cells it is possible to observe vDNA-vRNA associations, highlighting the proviruses location (Fig. 2A), in fact proviruses are the only transcribing forms. We strengthen our results on the ability to visualize proviruses by showing the vDNA-vRNA association by live imaging. HIV-1 transcriptional activity was tracked by MCP-GFP system (kind gift from E. Bertrand) to label the vRNA (Tantale et al., 2016) coupled

with HIV-1 ANCHOR technology (Fig. S2A and B). Next, we compared the intensity of ANCHOR signals of the episomal forms (+ RAL) with those of the integrated forms (-RAL) and observed that the signal of unintegrated DNA was two-fold brighter than the proviral DNA spot (Fig. 2A) in HeLa cells. Interestingly, we also remarked, during live imaging, that the vDNA of the integration-deficient virus (HIV-1 ANCH3 IN_{D116A}) clusters in the nucleus (Fig. S2C, video 4). Aggregation of extrachromosomal forms could prompt a brighter appearance of viral episomal DNA compared to a dimmer signal of the individual proviral DNA (Fig. 2A), at least in HeLa cells. Next, we pinpointed transcriptionally active proviruses in the main HIV-1 target cells, $CD4^+$ T cells and macrophages, by co-detecting vRNA and vDNA (Fig. 2 B-D). We observed that not all vDNAs were associated to a transcriptional focus, most likely because of the presence of silent episomal forms. On the other hand, an abundant part of vRNA foci colocalized with vDNA (Fig. 2B and C), meaning that we are able to detect single transcribing proviruses. Taken together, our results suggest that we are able to track down transcribing proviruses in the nuclear space of different cells and distinguish them from silent forms.

Nuclear localization of actively transcribing proviruses at single-cell level in dividing and non-dividing cells.

The location of actively transcribing proviruses in the nuclear space of dividing cells is still a subject of debate (Achuthan et al., 2018; Burdick et al., 2017; Chin et al., 2015; Di Primio et al., 2013; Lelek et al., 2015; Marini et al., 2015), whereas their nuclear location in terminally differentiated cells, as far as we know, has never been studied before. In dividing cells, the 3D analysis of the nuclear location of the vDNA showed the vDNA signals at 48 h p.i. to be more randomly distributed throughout the nucleus compared to the vDNA spots at seven days p.i. (avg distances from the nuclear envelope (NE): ~2.5 µm at 48 h p.i. vs ~1.4 µm at 7 days p.i.) (Fig. 3A). The difference in the average distance of the vDNA spots from the NE between the early and late time points p.i. may be due to the co-existence of heterogeneous viral forms at 48 h p.i. (episomal and proviral forms), whereas the vDNA population is more homogeneous at seven days p.i. (Fig. 1D, Fig. 3A). The episomal forms are distributed at a slightly greater distance from the NE than the mixed vDNAs population at 48 hour p.i. when compared to the vDNA at 7 d p.i. (Fig. 3A). Interestingly, we did not detect differences between the distances of the vRNA foci from the NE at both time points (48h p.i. and

7d p.i.) and the vDNA spots at seven days p.i. (< 2 µm in avg), indicating that in dividing cells the integration of the proviral DNA occurs early during infection near to the NE (Fig. 3A). We performed similar 3D analysis to calculate the distances of vDNAs and vRNAs from the NE in non-dividing cells at three days p.i.. The lack of mitosis hinders the loss of the episomal forms in non-dividing cells (Fig. 3B). However, when we compared the integration-defective viruses (only episomal forms) with the integration-efficient viruses (episomal and proviral forms), the vDNAs of the latter were found slightly closer to the NE (Fig. 3B). While the vRNAs (viral transcription foci) were closer to the NE with respect to episomal forms (~2.5 µm for the vDNA IN_{D116A} vs ~1.8 µm for the vRNA) (Fig.3B). Finally, our data show a similar nuclear spatial distance of viral transcriptional foci from the NE at single-cell level in both dividing and non-dividing cells (~1.8 µm in HeLa cells and in THP-1 cells) (Fig.3C).

Late reverse transcribed DNAs separate from IN proteins sequestered in CPSF6 clusters.

A long-standing question in the early steps of HIV life cycle is related to the study of IN-vDNA dynamics after nuclear entry. How long does the association last and can we detect it? There are, indeed, multiple difficulties due to the likely short-time nature of the event or due to the technical limitations of vDNA and IN visualization. Therefore, we exploited our vDNA tracking system to investigate about IN proteins localization in primary $CD4^+$T cell and THP-1 cells, challenging both cell types with HIV-1 ANCH3. We found in both cell types that the vDNA/IN association event was very rare (Fig. 4A). Nuclear IN proteins were found in the nucleus separated from the late reverse transcripts (Fig. 4A). To confirm our data we performed live-imaging, asking whether we could pinpoint the dynamic evolution of vDNA/IN association in the nucleus of infected cells by coupling the HIV-1 ANCHOR system with the GIR virus (generous gift from Edward Campbell) (Dharan et al., 2016; Hulme et al., 2015). GIR virus is produced using a viral genome carrying the IN mutated in the catalytic site (HIV-1 ANCH3 IN_{D116A}) together with Gag-integrase-Ruby (GIR) plasmid. As consequence, the only active IN in these viral particles is the IN-Ruby (Fig. 4B, Fig. S3). During a time lapse of 4 hours in HeLa cells (Fig. S3), we have observed the signals of the IN and of the vDNA always separated (~ 0.6 to 0.8 µm) (Fig. S3, video 5) corroborating the results in fixed cells (Fig. 4A). Importantly, we did not synchronize the infection to have the opportunity to visualize in a single cell different viral dynamics. We found several

IN foci alone and then colocalizing with the vDNA before the separation (Fig. 4B, movie 6, 7). We analysed the association of the two signals that rapidly (~10 min) separated (Fig. 4B, video 6,7). These data suggest that the IN proteins that are part of the mature PIC cannot be detected by a conventional light microscopy. Next, we investigated whether the viral IN proteins are retained in particular nuclear regions by some nuclear host factors. We found that the viral IN occupies, in nearly all cases analyzed, the same positions of the host Cleavage and Polyadenylation Specificity Factor subunit 6 (CPSF6), in both dividing and non-dividing cells (IN to CPSF6 ~ 94.4% in CD4+T cells and ~92.5% in THP-1 cells) (Figure 4C). Of note, HIV-1 reprograms the nuclear localization of CPSF6 from being randomly dispersed in the nucleoplasm to form particular aggregates in both primary lymphocytes and macrophages (Fig. 4C). Taken together our results indicate that IN proteins after viral nuclear entry are retained in CPSF6 clusters, while the late reverse transcribed DNA separates from these sites (Fig. 4A-C). We next wondered if other viral components are also hosted in CPSF6 MLOs together with the IN. Adapting our RNA FISH protocol with a longer probe hybridization and coupling it with the immunostaining of CPSF6 (Immuno-RNA FISH) (Fig. S4), we uncovered the presence of incoming vRNA inside the MLOs (Fig. 4D). Upon NEV treatment, HIV-1 genomes locate in CPSF6 clusters, both in $CD4^{+}$T and THP-1 cells. In physiological conditions it seems that vRNA/CPSF6 association events are more abundant in non-dividing cells than in dividing cells (~87% vs. ~33%), at 2 days post infection (Fig. 4D). Overall, the sequestration of viral IN proteins and viral genomes in CPSF6 clusters hints that CPSF6 MLOs could serve as sites to release a mature PIC that contains only few IN proteins needed for the integration step (Ballandras-Colas et al., 2017; Hare et al., 2009). We called these particular nuclear niches HIV-induced membraneless organelles (HIV-1 MLOs).

HIV-induced membraneless organelles (HIV-1 MLOs) are sites of nuclear reverse transcription.

To better characterize the role of HIV-1 MLOs, we performed Immuno-RNA FISH at different time post infection in THP-1 cells. We observed the majority of vRNAs retained in CPSF6 clusters at 2 days p.i., while at 3 days p.i. and even more at 7 days p.i. almost half of the total vRNA foci were excluded from CPSF6 clusters (Figure 5A). Genomic vRNAs localize 100% with CPSF6 upon NEV treatment at all times p.i. (Fig. 5A), this suggests that the vRNA transcription foci locate outside CPSF6 clusters. What

is the functional meaning of the incoming vRNA sequestration in CPSF6 clusters? To answer to this question we performed Chromatin Immunoprecipitation (ChIP) of the endogenous CPSF6 protein in infected and uninfected THP-1 cells at 2 days p.i., followed by real-time PCR (Fig. 5B). Early, intermediate and late RT products were amplified using appropriate primers. Of note, early transcripts were more abundant than others, revealing an “ongoing” reverse transcription inside CPSF6 clusters. (Fig. 5B). Our results explain previous data showing EdU-labeled viral episomal forms clustering in nuclear niches of macrophages (Francis et al., 2020; Rensen et al., 2020). Importantly, the presence of the IN inside HIV-1 MLO is independent of the presence of the reverse transcription, highlighting that the retrotranscribed DNA is not the driving force of the HIV-1 MLOs formation (Fig 5C). Since CPSF6 participates to the viral nuclear translocation and post-nuclear entry steps by interacting with the viral CA (Bejarano et al., 2019; Burdick et al., 2017; Lee et al., 2010; Price et al., 2014) (Francis et al., 2020), it might be that the viruses inside HIV-1 MLOs have not yet completed the uncoating process.

The integration of the transcriptionally active proviruses occurs outside HIV-1 MLOs.

CPSF6 phase separation does not randomly occur in the nuclear space. CPSF6 clusters fully colocalize with the marker of nuclear speckles (NSs), SC35 (Fig. 6 A-C). Importantly, we remarked that CPSF6 and SC35 signals mainly occupy the interchromatin space (Fig. 6A). NSs are, indeed, interchomatin granules (Spector and Lamond, 2011). SC35 strongly associates to another speckle factor, SON, independently of viral infection (Pearson’s r ~0.995), indicating that CPSF6 clusters in the host NS (Fig. 6B). Contrary to EdU labeling (Rensen et al., 2020), the HIV-1 ANCHOR system allows the visualization of a single provirus (Fig.1C) and does not interfere with viral transcription (Blanco-Rodriguez et al., 2020). Thus, we asked where in the nuclear space the viral expression takes place. We first observed that 3 days p.i. almost all late reverse transcripts are excluded from CPSF6 clusters (Fig.6B), as expected from the vDNA/IN separation data (Fig. 4A and B). We computed, as well, the distances of the vDNAs from the boundary of the closest NS (avg ~ 0.6 µm) (Fig. 6B). Similar data were obtained upon integration inhibition (+RAL), where the majority of episomal DNA forms locate outside CSPF6 clusters and in average ~ 0.4 µm from SC35 (Fig. S5A). Taking into account that we found the late reverse transcribed DNA

located outside HIV-1 MLO, to understand the exact location of transcriptionally active proviruses we used MCP-MS2 RNA labelling system. We observed that MCP-GFP is unable to label the viral incoming RNA (Fig.6C). Under NEV treatment, MCP-GFP is filtered out of HIV-1 MLOs as shown by the intensity profiles, whether in physiological conditions it labels the viral transcription foci (Fig. 6C). Active proviruses are excluded from CPSF6 clusters (Fig. 6D) and located in average at a distance from the closest CS35 NS ~ 0.6 µm). These results are independent of the presence of vpx (Fig. S5B) that we used to increase viral infectivity in THP-1 cells (Hrecka et al., 2011; Laguette et al., 2011). Taken together, we display that both vDNA (episomal and proviral forms) and vRNA foci (active proviruses) are excluded from HIV-1 MLOs (Fig. 6E), clearly indicating that these structures cannot be sites for viral integration or replication. More in detail, we found foci of viral transcription located in avg at a distance ~ 0.9 µm from the closest HIV-1 MLO (CPSF6+SC35), while total vDNAs were found ~1.8 µm from the nearest HIV-1 MLO (Fig. 6F). These data were corroborated by another vRNA labelling approach, immuno-RNA FISH, coupled to HIV-1 ANCHOR to detect the transcribing proviruses, indicating that viral integration and transcription occur outside CPSF6 clusters (Fig. 6G). Of note, we confirmed the results obtained in THP-1 cells (Fig. 5A) in primary monocyte-derived macrophages (MDMs), using an HIV-1 GFP reporter virus (Fig. 7). Also in the MDMs we could appreciate the blockade of the incoming vRNA in CPSF6, while the transcription foci locate outside these organelles (Fig. 7). Finally, we were able to correlate the percentage of vRNA foci that locate outside the MLOs per nucleus and the viral GFP expression (Pearson's r ~0.63), showing that those are integration and transcription sites of HIV-1 (Fig.7).

HIV-1 integration sites are marked by LEDGF.

We extended our study on THP-1 cells at 3 days p.i. to investigate whether Lens Epithelium-Derived Growth Factor (LEDGF), which is a partner of the viral IN and a HIV-1 integration sites selection factor (Cherepanov et al., 2003; Ciuffi et al., 2005; Emiliani et al., 2005; Ferris et al., 2010; Hare et al., 2009; Lelek et al., 2015; Llano et al., 2006; Shun et al., 2008; Shun et al., 2007; Wang et al., 2014), is associated to the proviruses forming a complex visible by imaging. We observed that CPSF6 proteins cluster in proximity of LEDGF sites (~ 43%), without co-localizing (Fig. 8A). There is a similar number of LEDGF clusters in both infected and uninfected cells (Fig. 8A), while CPSF6 clusters are mainly formed upon infection, suggesting that CSPF6 is prompted

by the virus to localize near the pre-existing LEDGF regions (Fig.8A). Of note, the PWWP domain of LEDGF has been shown to interact with H3K36me3 (Pradeepa et al., 2014; Pradeepa et al., 2012), which is a marker of euchromatin. In particular, this post-translational modification is enriched in HIV-1 integration sites (Ciuffi et al., 2005; Lelek et al., 2015; Vansant et al., 2020; Wang et al., 2009). Therefore, we analysed the location of proviral foci of transcription in relation to LEDGF. In average ~64% vDNA and ~73% vRNA foci per nucleus were found in complex with LEDGF (Fig. 8B). Importantly, LEDGF is also known to interact with splicing factors (Singh et al., 2015). This is in agreement with our data showing vDNA and vRNA localized in avg ~ 0.6 µm from the closest NS (Fig. 6B and D). Next we have characterized in more detail the viral genome associations with LEDGF with respect to CPSF6. We calculated that ~ 87.9% of the vDNA/LEDGF complexes and ~80% of vRNA foci/LEDGF complexes are uncoupled to CPSF6 (Figure 8B). We found only few vDNA (~12.1%) and vRNA (~20%) foci associated to LEDGF clusters near to CPSF6 (Fig. 8B). Our results suggest that the two cellular factors, CPSF6 and LEDGF, orchestrate two sequential steps of viral replication, the maturation of the PIC and its integration, respectively (Fig. 9). Functional proviruses are not located inside HIV-1 MLOs, but in external active chromatin domains, enriched in LEDGF (Fig. 8B) and in proximity of SC35 NS (Fig. 6D), likely to benefit of a favorable environment for viral replication.

Discussion

Studies of parasites that permanently integrate into the host genome show that they have evolved to target chromatin in such a way as to optimize their coexistence with the host and favor the release of new viral progeny (Bushman, 2003; Craig and Marszalek, 2002). HIV-1 needs to rapidly release high levels of newly generated viruses after infection, intuitively explaining the preference of active genes as integration targets (Ciuffi et al., 2005; Di Nunzio et al., 2013; Lelek et al., 2015; Maldarelli, 2016; Schroder et al., 2002). However, the nuclear location of actively transcribing individual proviruses in dividing or non-dividing cells has never been visualized before. Thus far, the direct detection of HIV-1 genomes after reverse transcription and the distribution of proviral transcriptional foci has been technically challenging. HIV-1 ANCHOR system allows the live tracking of different nuclear late retrotranscribed DNA forms (Fig.1A-C). We observed a homogenous population

formed by proviral DNA with low diffusion coefficient. Episomal forms showed a heterogeneous behavior, characterized by two different populations, implying that these two viral forms can interact with different factors in the nuclear space (Fig. 1D). Moreover, imaging analysis showed that viral episomal forms identified by the accumulation of OR-GFP on ANCH3 sequences gave rise to brighter signals than proviral forms in HeLa cells (Fig. 2A). This phenomenon could explain the differences in behavior between integrated and non-integrated forms (Fig.1D). Intriguingly, we found that episomal vDNAs form clusters, which could result in a brighter signal than integrated forms (Fig. S2C, video 4). Next, to track the nuclear location of functional proviruses, we co-labeled vDNA and vRNA by combining HIV-1 ANCHOR technology with RNA FISH or with MCP-MS2 system (Tantale et al., 2016) in the main target cells of HIV-1 (Fig.2A-D, Fig.S2A and B). Therefore, we investigated the spatial location of proviruses in dividing and non-dividing cells. Three-dimensional analysis of the distribution of vDNAs during early (48 h p.i.) infection in mitotic cells displays that they are more randomly distributed than the proviral DNA from cells at 7days p.i. (Fig. 3A). We also observed that the location of viral transcription sites (vRNA) is highly similar between early and late time points of infection, indicating that in HeLa cells the proviral DNAs, the only transcribing forms (Fig. 2A), immediately localize to near the NE early after infection (Fig. 3A). This vicinity to the NE was detected as well in non-dividing cells (Fig. 3B), highlighting that in both mitotic and non-mitotic cells the proviruses locate in the nuclear periphery (Fig.3C). We next investigated the interplay among viral components and their association with host factors during post-nuclear entry steps. IN proteins strongly co-localize with CPSF6 clusters in macrophage-like cells and in $CD4^+$ T lymphocytes at early time of infection (Fig. 4C), in agreement with another study (Francis et al., 2020). Contrary to CPSF6, the late vDNA products of retrotranscription rarely co-localize with IN in both types of cells (Fig. 4A). However, in live imaging we pinpointed the budding of newly synthesized late reverse transcribed DNA from the IN clusters, which occurs within minutes (Fig.4B, video 6,7).These results are in line with the dynamic coexistence of the IN and vDNA, which remain separated but in the proximity for several hours, as it has been shown by real-time imaging in HeLa cells (Fig. S3, Table S1, Table S2, video 5). This result can appear in contrast with recent data from other groups and from us obtained with another DNA labeling technology, EdU (Francis et al., 2020; Rensen et al., 2020), showing the vDNA and the IN strongly colocalized with CPSF6. This apparent discrepancy can be explained by the intrinsic

features of vDNA labeling tools. While EdU labeling cannot provide direct information on viral replication, as EdU reduces viral transcription (Rensen et al., 2020), HIV-1 ANCHOR system allows the detection of late reverse transcripts (Fig.1A) that can generate functional proviral DNA, as we observed in HeLa cells, human lymphocytes, and macrophages (Fig. 2A-D). Thus, we asked whether CPSF6 clusters are physiological sites of nuclear RT and of viral replication. First, we performed ChIP of the endogenous CPSF6 followed by real time PCR in THP-1 cells infected or uninfected (48 h p.i.). The PCR results show an abundance of early and intermediate RT products in comparison to late RT products, indicating an ongoing RT process in CPSF6 sites. These results are in agreement with a detection of the nascent vDNA (Fig. 4B) and with the labeling of the vDNA in these sites (Rensen et al., 2020). Second, we were able to correlate the progression of the viral infection with the nuclear location of the vRNA foci (Fig. 5A). At early time p.i. almost all vRNA are sequestered in CPSF6 clusters to be reverse transcribed (Fig.5A). Concurrently with the advancement of the viral infection we observed an increase in vRNA foci outside CPSF6 clusters, these foci are never present in cells treated with NEV where the viral replication is completely blocked (Fig. 5A). Importantly, our results show a positive correlation between the number of vRNA foci excluded from CPSF6 clusters and the GFP expression (HIV-1 reporter gene) (Fig. 7). Thus, HIV-1 induces the formation of novel membraneless organelles (MLOs) retaining CPSF6 proteins and viral components, such as IN and viral RNA genomes to host viral reverse transcription (Fig. 4C and D, 5A and B). HIV-1 MLOs form in the interchromatin space, remodeling the host nuclear speckles, characterized by the presence of SC35 and SON (Fig.6A and B). Likely, CPSF6 relocates to speckles prompted by the virus to usurp functions linked to these MLOs. We show that the formation of CPSF6 cluster is independent of viral DNA integration or viral DNA synthesis (Fig. S5A, Fig. 5C) and that late reverse transcribed products are mainly excluded from them (Fig. 6B and E, Fig.7). These results are in agreement with a rapid exclusion of late retrotranscribed DNAs from the IN proteins that are retained in CPSF6 clusters (Fig. 4A-C). Next, we calculated the distribution of the actively transcribing proviruses in the nuclear landscape. Interestingly, we observed that the MCP-MS2 system detects only the vRNA associated with viral replication sites, as MCP-GFP is unable to penetrate HIV-1 MLOs that shield the incoming vRNA genome, acting as selective filters (Fig. 6C). When we compared the distances from the boundary of the closest HIV-1 MLOs (CPSF6 + SC35), we observed that the vRNA

transcription foci were closer than the vDNAs, indicating that the HIV-1 MLOs-neighbouring regions are the hub of viral replication (Fig. 6 F-G). We also observed that CPSF6 relocates to the vicinity of LEDGF during HIV infection (Fig. 8A). LEDGF associates to highly-spliced transcriptional units through its direct interaction with several splicing factors (Singh et al., 2015) and determines HIV-1 integration sites distribution in active chromatin sites (Ciuffi et al., 2005; Ferris et al., 2010; Shun et al., 2007). However, the visualization of a complex formed by the transcribing provirus and LEDGF remained utopian for a long time. Here, we were able to visualize by fluorescence microscopy a complex formed by the vDNA or vRNA and LEDGF. Almost 80% of those complexes are formed by LEDGF uncoupled to CPSF6 (Fig. 8B). Nonetheless, we observed that less than 20% of them can include CPSF6 associated to LEDGF. Whether there is a functional difference between these two complexes it can be further investigated in future studies. Overall, our study highlights the proviral location in the nuclear landscape remodelled by the viral infection (Fig. 9). We found that HIV-1-MLOs sequester the viral IN (Fig. 4C, Fig.5C), as well viral RNA genomes (Fig. 5A) and rapidly release the late retrotranscribed DNA, acting as HIV-1 reverse transcription and PIC maturation sites (Fig.4B, Fig. 5A and B). Thus, HIV-1 MLOs play a role prior to the integration step. Taken together our results show that functional proviruses establish near to the NE (Fig. 3C), excluded from CPSF6 clusters, but locate in a favorable transcriptional environment (Fig. 6D, Fig.8B) at less than 1 µm from HIV-1 MLOs (Fig.6F). Our study supports how new single-cell level approaches are pivotal in the study of functional viral nuclear dynamics (Fig.9), to discriminate the cells susceptible to fuel viremia or that concur to viral persistence.

Materials and Methods

Cell lines, primary cells and generation of genetically modified single-cell clones.

HEK 293T cells (ATCC) are human embryonic kidney cells used to produce lentiviral vectors and HIV-1. HeLa P4R5 reporter cells (Charneau et al., 1994). HeLa MCP-GFP is a HeLa clone stably expressing MCP-GFP bacterial fusion protein (kind gift from E. Bertrand) (Tantale et al., 2016). Jurkat cells (ATCC) are human immortalized T lymphocytes derived from acute T cells leukemia. THP-1 cells (ATCC) are human

immortalized monocyte cells derived from acute monocytic leukemia, which, once seeded, were differentiated into macrophage-like cells under Phorbol 12-myristate 13-acetate (PMA) treatment (167nM). SupT1R5 cells are T lymphoblastic lymphoma cells which were transduced to stably express CCR5 (kind gift of N. Casartelli). HEK293T cells and HeLa cells were cultivated in DMEM medium supplemented with 10% Fetal Bovine Serum (FBS) and 1% penicillin-streptomycin. Jurkat, SupT1R5 and THP-1 cells were cultivated in RPMI 1640 medium supplemented with 10% Fetal Bovine Serum (FBS) and 1% penicillin-streptomycin solution. Primary $CD4^+$T and MDM cells were purified from healthy donors' blood obtained via EFS (Etablissement Français du Sang, Paris), through density gradient centrifugation with Ficoll 400 and isolation of PBMCs buffy coat. $CD4^+$T cells positive selection was performed with human CD4 Microbeads (Miltenyi Biotec #130-045-101). The day after $CD4^+$T cells were activated with T cell Activation/Expansion kit (Miltenyi Biotec #130-091-441) according to the datasheet, complete RPMI 1640 medium was supplemented with interleuykin-2 (IL-2) (400 IU/mL). For MDMs the PBMCs were incubated at 37°C for 2 hours, then, non-adherent cells were washed away and complete RPMI 1640 medium was supplemented with M-CSF (10 ng/mL). After 3 days, the medium was changed with RPMI without M-CSF and cells were left to differentiate for other 4 days before proceeding with infection. HIV-1 ANCH3 single-cell clone was generated starting from HeLa P4R5 cells which were infected with HIV-1 ANCH3 virus at MOI 1. Cells were diluted to 1 cell per well, in 96-well plates. Cell-clone colonies were tested for β-galactosidase expression to check viral expression (kit Roche #11758241001). Positive clones were transduced with lentiviral vector (LV) CMV-OR-GFP for imaging analyses of HIV-1 ANCH3 provirus. All cells were kept in incubator at 37°C and 5% CO_2.

Plasmids, lentiviral vectors and viral productions

HIV-1ΔEnvIN$_{HA}$ΔNef plasmid encodes for the ΔEnvHIV-1 LAI (BRU) viral genome where the Integrase (IN) protein is fused to the hemagglutinin (HA) tag (kindly provided by F. Mammano)(Petit et al., 1999; Petit et al., 2000). HIV-1ΔEnvIN$_{HA}$ΔNef ANCH3 (HIV-1 ANCH3) was obtained through ANCH3 insertion: ANCH3 sequence was cloned by PCR using the template plasmid pANCH3 as we previously described in Blanco et al. (Blanco-Rodriguez et al., 2020). The ANCHOR™ technology and sequences are exclusive property of NeoVirTech (Germier et al., 2017; Mariame et al., 2018). HIV-

1ΔEnv IN_{HA}(D116A) ΔNef ANCH3 was obtained by insertional mutagenesis using the QuikChange II XL Site-Directed Mutagenesis kit (Agilent #200522), for integrase (IN) mutation at D116. The lentiviral vector plasmids CMV OR-GFP and CMV OR-SANTAKA were obtained by cloning OR-GFP/OR-SANTAKA (plamids from NeoVirTech) in a pTripCMV (ΔU3) plasmid through PCR and cloned using restriction site, AgeI and SgrDI**.** HIV-1ΔEnvIN_{HA}ΔNef ANCH3 MS2 plasmid was obtained by inserting the MS2x64 sequence amplified by PCR from pMK123**-**MS2x64 (Tantale et al., 2016) in HIV-1ΔEnvIN_{HA}ΔNef ANCH3. Lentiviral vectors and HIV-1 viruses were produced by transient transfection of HEK293T cells through calcium chloride coprecipitation. Co-transfection was performed as following, LV carrying OR-GFP/OR-SANTAKA: 10 µg of OR-GFP/OR-SANTAKA LV, 10 µg of NDK plasmid (gag-pol-tat-rev) and 2.5 µg of pHCMV-VSV-G envelope expression plasmid; LV carrying MCP-GFP (kindly provided by E. Bertand): 10 µg of MCP-GFP, 10 µg of NDK plasmid (gag-pol-tat-rev) and 2.5 µg of pHCMV-VSV-G envelope expression plasmid; HIV-1ΔEnv pseudotyped with VSV-G: 10 µg HIV-1ΔEnvIN_{HA}ΔNef (IN WT or IN(D116A)) plasmid and 2.5 ug of pHCMV-VSV-G plasmid. VSV-G/HIV-1ΔEnv IN_{HA}(D116A) ΔNef ANCH3 has been also produced in combination with GIR (Gag-IN-Ruby plasmid, kindly provided by E. Campbell) (Dharan et al., 2016; Hulme et al., 2015). For THP-1 infection, HIV-1 ANCH3 or HIV-1 ANCH3 MS2 viruses and LV OR-GFP or LV MCP-GFP were all produced adding 3 µg of SIV_{MAC} Vpx (Durand et al., 2013) at the moment of HEK293T transfection. MDMs were infected with a virus HIV-1 ΔEnv Δnef GFP complemented with Vpx (Rensen et al., 2021). After the collection of the supernatant 48h post-transfection, lentiviral particles were concentrated by ultracentrifugation for 1 h at 22000 rpm at 4°C and stored at −80°C. Lentiviral vectors and viruses were tittered through late reverse transcripts (LTR) qPCR in HEK293T cells.

Quantitative PCR and primers

Total DNA of 10^6 infected cells was extracted at 6 hours post infection for Late Reverse Transcripts (LRTs) qPCR and 24 hours post infection for 2LTR-circles and ALU-PCR, through QIAmp DNA micro kit (QIAGEN #56304). Real-Time PCR of LRTs was performed to assess DNA synthesis and used to normalize qPCR data on viral input, the reactions were carried on in 20 µL, in iTaqUniversal SYBR Green Supermix (Bio-Rad #1725124) using primers for U3 sequence: U3 FX: 5'-TTCCGCTGGGGACTTTCCAGGG-3', U3 RX: 5'-AGGCTCAGATCTGGTCTAACC-3'.

Real-Time PCR of 2LTR-circles was used to asses nuclear import efficiency. Reactions were performed in 20 µL, in Maxima Probe/ROX qPCR Mastermix (ThermoFisher #K0232) using primers for 2LTR-circle junction: 2LTR FX: 5'-AACTAGGGAACCCACTGCTTAAG-3', 2LTR RX: 5′-TCCACAGATCAAGGATATCTTGTC-3′, 2-LTR probe: 5'-(FAM)-ACACTACTTGAAGCACTCAAG-GCAAGCTTT-(TAMRA)-3'. Real-Time PCR of proviral integrations consisting in a first non-kinetic PCR in 50 µL in Platinum SuperFi DNA Polymerase (ThermoFisher #12351250) and in a second step of qPCR reaction in 20 µL in Maxima Probe/ROX qPCR Mastermix (ThermoFisher #K0232). Primer Alu 166: 5'-TCCCAGCTACTCGGGAGGCTGAGG-3', Alu 2: 5'-GCCTCCCAAAGTGCTGGGATTACAG-3', LambdaU3: 5′-ATGCCACGTAAGCGAAACTTTCCGCTGGGGACTTTCCAGGG-3′ for ALU PCR; Lambda: 5'-ATGCCACGTAAGCGAAACT-3', U5: 5'-CTGACTAAAAGGGTCTGAGG-3', Probe: 5'-(FAM)- TTAAGCCTCAATAAAGCTTGCCTTGAGTGC-(TAMRA) for ALU qPCR (Di Nunzio et al., 2013). In all experiments β-actin detection was used for normalization. β-actin FX: 5'-AACACCCCAGCCATGTACGT-3', β-actin RX: 5-CGGTGAGGATCTTCATGAGGTAGT-3', β-actin probe: (FAM)-CCAGCCAGGTCCAGACGCAGGA-(BHQ1).

Epifluorescence and Immunofluorescence of HIV-1 infected cells

For imaging studies in adherent cell-lines, the cells were plated on coverslips (12mm diameter in 24 well plates or 18 mm diameter in 12 well plate, #1, ThermoFisher) at least 24h before fixation. For HIV-1 ANCH3 imaging, HeLa P4R5 cells were transduced with OR-GFP LV (MOI 0.2) and then infected usually 24 hours later with HIV-1 ANCH3 or HIV-1 IN_{D116A} ANCH3 at different MOIs (from 5 to 30). NEVIRAPINE (NEV) 10 µM or RALTEGRAVIR (RAL) 20 µM were used to block respectively DNA synthesis or integration. All THP-1 cells studies were carried on after 48h of differentiation with PMA (167 nM), then, they were transduced with LV OR-GFP Vpx (MOI 0.5) or LV MCP-GFP Vpx (MOI 2) and 48 h later they were infected with HIV-1 ANCH3 Vpx (MOI 10-20) or HIV-1 ANCH3 MS2 ± Vpx (MOI 2.5). The medium was always supplemented with PMA. For HIV-1 ANCH3 imaging in primary lymphocytes, $CD4^+$T cells, 3 days after activation, were transduced with OR-GFP LV (MOI 0.5) then infected with HIV-1 ANCH3 (MOI 100). The day of fixation cells in suspension were

seeded on poly-L-lysine (Sigma #P4707) coated coverslips and shortly centrifuged to allow attachment. The day of fixation all cells were washed with PBS and fixed with 4% PFA for 15 minutes. For protein staining, cells were treated with glycine 0.15% for 10 min, permeabilized with 0.5% Triton X-100 for 30 min and blocked with 1% bovine serum albumin (BSA) for 30 min. All incubations were carried out at room temperature, in the dark, in humid chamber, 1h with primary antibodies and 45 min with secondary antibodies. Washes between antibody incubations and antibodies dilution were done in 1% BSA. Primary antibodies were diluted as follows: anti-HA 1:500 (Roche #11867423001), anti-HIV-1 p24 (CA) 1:400 (NIH #3537), anti-CPSF6 1:400 (Novus Biologicals #NBP1-85676), anti-SC35 1:200 (Abcam #ab11826), anti-LEDGF 1:200 (BD Bioscience #611715). Secondary antibodies used were the following: Goat anti-Mouse Alexa-647 1:300 (1:100 for CA) (Invitrogen #A21235); Donkey anti-Rat Alexa-488 1:100 (Invitrogen# A21208) for IN-HA or Goat anti-Rat Alexa-647 1:300 (Invitrogen #A21247), 1:100 for IN-HA; Donkey anti-Rabbit Cy3 1:500 (Jackson Lab #711-165-152). Finally, cells were stained with Hoechst 33342 1:5000 (Invitrogen #H3570) for 5 minutes. Coverslips were mounted on glass slides (Star Frost) with Prolong Diamond Antifade Mountant (Life Technologies #P36970). Confocal microscopy was carried out with a Zeiss LSM700 inverted microscope, with a 63X objective (Plan Apochromat, oil immersion, NA=1.4).

RNA FISH

For RNA FISH studies in adherent cell-lines, the cells were seeded on coverslips. HeLa P4R5 cells expressing OR-GFP were infected with HIV-1 ANCH3 (MOI 30) with or without RAL (20 µM) or infected with HIV-1 ANCH3 (MOI 5). Jurkat cells expressing OR-GFP were infected with HIV-1 ANCH3 (MOI 5). Primary activated $CD4^{+}$T cells, 3 days after activation, were together transduced with OR-GFP LV (MOI 10) and infected with HIV-1 ANCH3 (MOI 50). THP-1 cells were differentiated with PMA for 48h and, after 48h transduction with LV OR-GFP Vpx (MOI 0.5), the cells were infected with HIV-1 ANCH3 Vpx (MOI 20). Primary MDMs were seeded before differentiation on coverslips, and after 7 days, transduced with LV OR-GFP Vpx (MOI 10) and infected with HIV-1 ANCH3 Vpx (MOI 40). The day of fixation suspension cells were seeded on poly-L-lysine (Sigma #P4707) coated coverslips and centrifuged to allow attachment. Then, cells were washed with PBS and fixed with 4% PFA for 15 minutes and

incubated in 70% ethanol at -20°C at least for one night. Primary smiFISH probes have been designed against HIV-1 *pol* transcript and a shared readout sequence for secondary probe alignment. Twenty-four smiFISH probes (Table S3) against HIV pol were designed with Oligostan (Tsanov et al., 2016) and purchased from Integrated DNA Technologies (IDT). Primary probes were pre-hybridized with a secondary FLAP probe conjugated to Cy5 fluorophore through pairing with the readout sequence (Eurofins). Washes and hybridization were performed with Stellaris Buffers (WASH buffer A, WASH buffer B, Hybridization Buffer; LGC Biosearch Technologies), following the manufacturer protocol. Hybridization with the probe was carried out at 37°C in a dark humid chamber for 5 hours. Finally, cells were stained with Hoechst 33342 1:5000 (Invitrogen #H3570) for 5 minutes. Coverslips were mounted on glass slides (Star Frost) with Prolong Diamond Antifade Mountant (Life Technologies #P36970). Confocal microscopy was carried out with a Zeiss inverted LSM700 microscope, with a 63X objective (Plan Apochromat, oil immersion, NA=1.4).

IMMUNO-RNA FISH

THP-1 and primary MDM cells were seeded on coverslips and infected with HIV-1 ANCH3 vpx (MOI 20). SupT1R5 cells were infected with HIV-1 ANCH3 (MOI 20) and the day of fixation they were seeded on poly-L-lysine (Sigma #P4707) coated coverslips and centrifuged to allow attachment. The day of fixation the coverslips with the cells were washed with PBS and fixed with 4% PFA for 15 minutes to proceed with CPSF6 immunostaining. Coverslips where incubated with a blocking/permeabilization solution (PBS-BSA 1%, Vanadyl ribonucleoside complexes (VRCs) solution 2mM (Sigma #94742), TRITON-X 0.3%) for 1 h. Then, the coverslips were incubated with anti-CPSF6 1:400 (#NBP1-85676) in blocking/permeabilization solution for 1h, washed and incubated for 45 min with anti-Rabbit secondary antibody, at room temperature, in the dark, in humid chamber. The coverslips were fixed again in PFA 4% for 10 min. Subsequent RNA FISH and mounting protocol was carried out as mentioned in the paragraph before, but the incubation with smiFISH probes was carried on overnight to better allow the detection of incoming genomic viral RNA in CPSF6 clusters (Fig.S4). Confocal microscopy was carried out with a Zeiss LSM700 inverted microscope, with a 63X objective (Plan Apochromat, oil immersion, NA=1.4).

Time-lapse microscopy

For all live imaging studies, 10^5 cells were plated on a polymer-coverslip bottom µ-Dish 35 mm (ibidi #81156) at least 24 hours before 2D or 3D video acquisition with an UltraView VOX inverted Microscope (Perkin-Elmer), based on a CSU-X spinning-disk (Yokogawa), and using a 63X objective (Plan Apochromat, oil immersion, NA=1.4). For integrated provirus imaging, HIV-1 ANCH3 single-cell clone was transduced with OR-GFP LV (MOI 0.2) for 3 days and then plated on the polymer-bottom dish the day before acquisition. On the other hand, HeLa P4R5 cells were infected with HIV-1 ANCH3 (MOI 10) and five days post infection cells were plated on the polymer-coverslip bottom dish and transduced with LV OR-GFP for 24h. For the acquisitions of the episomal viral DNA, HeLa P4R5 were firstly transduced with OR-GFP LV (MOI 0.2) and the day after cells were plated on the polymer-bottom dish and infected with HIV-1 ANCH3 (MOI 10) with RAL 20 µM or with the HIV-1 ANCH3 IN_{D116A}. Cells were imaged in 3D in continue for 10 minutes. Experiments of co-live tracking of IN-Ruby (GIR) and vDNA (OR-GFP) were performed in HeLa P4R5 cells transduced with LV OR-GFP (MOI 0.2). The following day, cells were infected with HIV-1 ANCH3 IN_{D116A} complemented with the GIR plasmid using an MOI 50. Frames were acquired every 30 seconds in 2D video (Fig. S3, video 5) and every minute for 3D videos (Fig. 4B, video 6,7), at different times post infection. For co-live-imaging of vDNA (OR-SANTAKA) and vRNA foci (MCP-GFP), HeLa MCP-GFP cells were transduced with OR-SANTAKA LV (MOI 1) and 24 h later infected with HIV-1 ANCH3 MS2 (MOI 50).

Chromatin Immunoprecipitation with qPCR

Four hundred million of THP-1 cells infected with HIV-1 (MOI 5) or uninfected were fixed with 1% formaldehyde for 15 min and quenched with 0.125 M glycine, then sent to Active Motif Services (Carlsbad, CA) to be processed for ChIP-qPCR. In brief, chromatin was isolated by the addition of lysis buffer, followed by disruption with a Dounce homogenizer. Lysates were sonicated and the DNA sheared to an average length of 300-500 bp. Genomic DNA (Input) was prepared by treating aliquots of chromatin with RNase, proteinase K and heat for de-crosslinking, followed by ethanol precipitation. Pellets were resuspended and the resulting DNA was quantified on a NanoDrop spectrophotometer. Extrapolation to the original chromatin volume allowed quantitation of the total chromatin yield. An aliquot of chromatin (40 µg) was precleared with protein A agarose beads (Invitrogen). Genomic DNA regions of interest were isolated using 4 µg of antibody against CPSF6 (#NBP1-85676). Complexes were

washed, eluted from the beads with SDS buffer, and subjected to RNase and proteinase K treatment. Crosslinks were reversed by incubation overnight at 65°C, and ChIP DNA was purified by phenol-chloroform extraction and ethanol precipitation. Quantitative PCR (qPCR) reactions were carried out in triplicate on specific genomic regions using SYBR Green SuperMix (BioRad). The resulting signals were normalized for primer efficiency by carrying out qPCR for each primer pair using Input DNA. Different RT products were amplified using the following primers, early RT products: 5′-GCCTCAATAAAGCTTGCCTTGA-3′ and 5′-TGACTAAAAGGGTCTGAGGGATCT-3′; intermediate RT products: 5′ -CTAGAACGATTCGCAGTTAATCCT-3′ and 5′ -CTAT CCTTTGAT GCACACAATAGAG-3′; late RT products: 5′-TGTGTGCCCGTCTGTTGTGT-3′ and 5′-GAGTCCTGCGTCGAGAGAGC-3'. De-crosslinked ChIP products were checked by western blot against CPSF6, loading 37.5 ng of DNA per sample.

Imaging Analysis and Statistics

All data were analyzed in GraphPad Prism 8 (GraphPad Software, La Jolla California USA, www.graphpad.com). All images and videos were analyzed in Fiji (Schindelin et al., 2012) and Icy version 2.0.3.0 (de Chaumont et al., 2012) software. Here below the custom-made protocols:

Particle tracking and MSD measurement

3D videos were analyzed in Fiji (Schindelin et al., 2012) and viral nuclear particles were tracked along time in 3D using TrackMate plugin (Tinevez et al., 2017). The resulting tracks were exported in MATLAB (The MathWorks, Natick, USA) and then used for mean-squared displacement (MSD) analysis with @msdanalyzer(Tarantino et al., 2014). MSD curves were fitted by a straight line from which we derived the value of the diffusion coefficient: D = a / (2*nDim), where « a » is the slope of the MSD curve and « nDim » the dimensionality of the image. MSD curves for which the R^2 value was lower than 0.8 were discarded from analysis. Only the initial 25% of the MSD curves were used for the linear fit.

vDNA spot intensity analysis

The 2D confocal images were analysed with Icy software version 2.0.3.0 (de Chaumont et al., 2012) using a custom made protocol (http://icy.bioimageanalysis.org/protocol/batch-spots-detections-and-statistics and doi: 10.5281/zenodo.3925262, example of HeLa confocal image available at doi:

10.5281/zenodo.3925213). vDNA spots were detected extracting the green channel and applying first a median filter of half size = 1 and then 3, the Spot Detector with a scale of 1 (sensitivity to 55) on the nuclear ROI only. Data from the spot detection (e.g. mean fluorescence intensity) were exported for analysis. vRNA mean fluorescence intensity signal was computed from the nuclei ROI.

vDNA-vRNA association analysis

For vDNA - vRNA association studies in Jurkat cells, 3D confocal images were analyzed in Icy version 2.0.3.0 using two successive custom made protocols (first protocol:http://icy.bioimageanalysis.org/protocol/nuclei-segmentation-from-maximum-intensity-projections/ with doi: 10.5281/zenodo.3925181) and second protocol: http://icy.bioimageanalysis.org/protocol/spots-detection-and-colocalisation-analysis-with-soda/ with doi: 10.5281/zenodo.3925191, example of Jurkat confocal image available at doi: 10.5281/zenodo.3925213). Nuclei were segmented on a maximum intensity projection of the Hoechst channel using the Icy block Gaussian filter with a radius of sigma x=5, y=5 and z=5 pixels, the Thresholder block with a manual threshold of 5000 and the Label Extractor block, which uses connected components to separate the nuclei. Cy5 spots were detected with the HK Means block (Dufour et al., doi.org/10.1109/ICPR.2008.4761748).The logical operation block was used to keep only the spots detected in the nucleus. These spots were added as input parameter “List detection 1” of the SODA block. EGFP spots were detected using the log3D block to enhance the spots, inverting the log transformed image using the Math Operation Expression block with the following expression: A*B with A scalar=-1 and B the log transformed image, detecting the spots with the Local Extrema block and segmenting the spots with a dilation and a connected components operations. The logical operation block was used to keep only the spots detected in the nucleus. These spots were added as input parameter “List detection 2” of the SODA block. The nucleus ROI was kept as input parameter for “ROI of Analysis (Cell’s shape)”. The Cy5 and GFP spots ROIs were then imported in the SODA block (Lagache et al., 2018) to analyse colocalization of spots within the nucleus only with a maximum radius of 10, a step of 1 and no fixed search distance.

vDNA/vRNA nuclear envelope distance analysis

The nuclei were segmented on the Hoechst channel using background subtraction followed by a multilevel Otsu thresholding. The resulting nuclei masks were then directly applied on the vDNA and vRNA channels to select the relevant 3D information

inside each labelled region. Inside these regions, vRNA spots were identified using 3D non-local means denoising (Buades et al., "Non-local means denoising." *Image Processing On Line* 1 (2011): 208-212) and automated segmentation, while vDNA spots were extracted with the image analysis pipeline previously described (Komatsu et al., 2018). For each vDNA and vRNA spot in was computed the distance to the nuclear envelope in the 2D z-plane corresponding to the spot centroid location.

SC35/HIV-1 MLO distances analysis

The 3D confocal images were processed for multi-channel image splitting, vDNA and SC35 segmentation. The automated 3D segmentation of cell nuclei included a preliminary non-local means denoising step to cope with the strong signal heterogeneity and the segmentation of SC35 was readjusted to low contrast conditions. The 3D boundary-to-boundary distance (rather than centroid-to-centroid which was done for vDNA/vRNA-NE analysis) between each vDNA or vRNA spot and its closest SC35 speckle or the closest HIV-1 MLO (SC35+CPSF6) were computed.

Acknowledgements

We wish to thank Guillermo Blanco-Rodriguez, Stella Frabetti, Philippe Souque and Blandine Monel for experimental help, Nicoletta Casartelli for critical reading of the manuscript. We thank Fabrizio Mammano, Edouard Bertrand, Florian Mueller, and Edward Campbell for sharing reagents. We gratefully acknowledge the UtechS Photonic BioImaging (Imagopole), C2RT, Institut Pasteur, supported by the French National Research Agency (France BioImaging; ANR-10–INSB–04; Investments for the Future). We thank the NIH AIDS Reagents program to support us with precious reagents. This work was funded by the ANRS REACTing grant ECTZ88162 with a nominative PhD student fellowship ECTZ88177 for V.S., the Sidaction/FRM grant VIH20170718001, the Pasteur Institute.

Competing financial interests: The authors declare no competing financial interests.

References

Achuthan, V., J.M. Perreira, G.A. Sowd, M. Puray-Chavez, W.M. McDougall, A. Paulucci-Holthauzen, X. Wu, H.J. Fadel, E.M. Poeschla, A.S. Multani, S.H. Hughes, S.G. Sarafianos, A.L. Brass, and A.N.

Engelman. 2018. Capsid-CPSF6 Interaction Licenses Nuclear HIV-1 Trafficking to Sites of Viral DNA Integration. *Cell Host Microbe*. 24:392-404 e398.

Albanese, A., D. Arosio, M. Terreni, and A. Cereseto. 2008. HIV-1 pre-integration complexes selectively target decondensed chromatin in the nuclear periphery. *PLoS One*. 3:e2413.

Ballandras-Colas, A., D.P. Maskell, E. Serrao, J. Locke, P. Swuec, S.R. Jonsson, A. Kotecha, N.J. Cook, V.E. Pye, I.A. Taylor, V. Andresdottir, A.N. Engelman, A. Costa, and P. Cherepanov. 2017. A supramolecular assembly mediates lentiviral DNA integration. *Science*. 355:93-95.

Bejarano, D.A., K. Peng, V. Laketa, K. Borner, K.L. Jost, B. Lucic, B. Glass, M. Lusic, B. Muller, and H.G. Krausslich. 2019. HIV-1 nuclear import in macrophages is regulated by CPSF6-capsid interactions at the nuclear pore complex. *Elife*. 8.

Blanco-Rodriguez, G., A. Gazi, B. Monel, S. Frabetti, V. Scoca, F. Mueller, O. Schwartz, J. Krijnse-Locker, P. Charneau, and F. Di Nunzio. 2020. Remodeling of the core leads HIV-1 pre-integration complex in the nucleus of human lymphocytes. *J Virol*.

Buffone, C., A. Martinez-Lopez, T. Fricke, S. Opp, M. Severgnini, I. Cifola, L. Petiti, S. Frabetti, K. Skorupka, K.K. Zadrozny, B.K. Ganser-Pornillos, O. Pornillos, F. Di Nunzio, and F. Diaz-Griffero. 2018. Nup153 Unlocks the Nuclear Pore Complex for HIV-1 Nuclear Translocation in Nondividing Cells. *J Virol*. 92.

Burdick, R.C., K.A. Delviks-Frankenberry, J. Chen, S.K. Janaka, J. Sastri, W.S. Hu, and V.K. Pathak. 2017. Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. *PLoS Pathog*. 13:e1006570.

Burdick, R.C., C. Li, M. Munshi, J.M.O. Rawson, K. Nagashima, W.S. Hu, and V.K. Pathak. 2020. HIV-1 uncoats in the nucleus near sites of integration. *Proc Natl Acad Sci U S A*. 117:5486-5493.

Bushman, F.D. 2003. Targeting survival: integration site selection by retroviruses and LTR-retrotransposons. *Cell*. 115:135-138.

Butler, S.L., M.S. Hansen, and F.D. Bushman. 2001. A quantitative assay for HIV DNA integration in vivo. *Nat Med*. 7:631-634.

Campbell, E.M., and T.J. Hope. 2015. HIV-1 capsid: the multifaceted key player in HIV-1 infection. *Nat Rev Microbiol*. 13:471-483.

Charneau, P., G. Mirambeau, P. Roux, S. Paulous, H. Buc, and F. Clavel. 1994. HIV-1 reverse transcription. A termination step at the center of the genome. *J Mol Biol*. 241:651-662.

Cherepanov, P., G. Maertens, P. Proost, B. Devreese, J. Van Beeumen, Y. Engelborghs, E. De Clercq, and Z. Debyser. 2003. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J Biol Chem*. 278:372-381.

Chin, C.R., J.M. Perreira, G. Savidis, J.M. Portmann, A.M. Aker, E.M. Feeley, M.C. Smith, and A.L. Brass. 2015. Direct Visualization of HIV-1 Replication Intermediates Shows that Capsid and CPSF6 Modulate HIV-1 Intra-nuclear Invasion and Integration. *Cell Rep*. 13:1717-1731.

Ciuffi, A., M. Llano, E. Poeschla, C. Hoffmann, J. Leipzig, P. Shinn, J.R. Ecker, and F. Bushman. 2005. A role for LEDGF/p75 in targeting HIV DNA integration. *Nat Med*. 11:1287-1289.

Craig, E.A., and J. Marszalek. 2002. A specialized mitochondrial molecular chaperone system: a role in formation of Fe/S centers. *Cell Mol Life Sci*. 59:1658-1665.

de Chaumont, F., S. Dallongeville, N. Chenouard, N. Herve, S. Pop, T. Provoost, V. Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. Le Montagner, T. Lagache, A. Dufour, and J.C. Olivo-Marin. 2012. Icy: an open bioimage informatics platform for extended reproducible research. *Nat Methods*. 9:690-696.

De Wit, F., S.R. Pillalamarri, A. Sebastian-Martin, A. Venkatesham, A. Van Aerschot, and Z. Debyser. 2019. Design of reverse transcriptase-specific nucleosides to visualize early steps of HIV-1 replication by click labeling. *J Biol Chem*. 294:11863-11875.

Dharan, A., S. Talley, A. Tripathi, J.I. Mamede, M. Majetschak, T.J. Hope, and E.M. Campbell. 2016. KIF5B and Nup358 Cooperatively Mediate the Nuclear Import of HIV-1 during Infection. *PLoS Pathog*. 12:e1005700.

Di Nunzio, F. 2013. New insights in the role of nucleoporins: a bridge leading to concerted steps from HIV-1 nuclear entry until integration. *Virus Res*. 178:187-196.

Di Nunzio, F., T. Fricke, A. Miccio, J.C. Valle-Casuso, P. Perez, P. Souque, E. Rizzi, M. Severgnini, F. Mavilio, P. Charneau, and F. Diaz-Griffero. 2013. Nup153 and Nup98 bind the HIV-1 core and contribute to the early steps of HIV-1 replication. *Virology*. 440:8-18.

Di Primio, C., V. Quercioli, A. Allouch, R. Gijsbers, F. Christ, Z. Debyser, D. Arosio, and A. Cereseto. 2013. Single-cell imaging of HIV-1 provirus (SCIP). *Proc Natl Acad Sci U S A*. 110:5636-5641.

Durand, S., X.N. Nguyen, J. Turpin, S. Cordeil, N. Nazaret, S. Croze, R. Mahieux, J. Lachuer, C. Legras-Lachuer, and A. Cimarelli. 2013. Tailored HIV-1 vectors for genetic modification of primary human dendritic cells and monocytes. *J Virol*. 87:234-242.

Emiliani, S., A. Mousnier, K. Busschots, M. Maroun, B. Van Maele, D. Tempe, L. Vandekerckhove, F. Moisant, L. Ben-Slama, M. Witvrouw, F. Christ, J.C. Rain, C. Dargemont, Z. Debyser, and R. Benarous. 2005. Integrase mutants defective for interaction with LEDGF/p75 are impaired in chromosome tethering and HIV-1 replication. *J Biol Chem*. 280:25517-25523.

Farnet, C.M., and W.A. Haseltine. 1991. Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. *J Virol*. 65:1910-1915.

Ferris, A.L., X. Wu, C.M. Hughes, C. Stewart, S.J. Smith, T.A. Milne, G.G. Wang, M.C. Shun, C.D. Allis, A. Engelman, and S.H. Hughes. 2010. Lens epithelium-derived growth factor fusion proteins redirect HIV-1 DNA integration. *Proc Natl Acad Sci U S A*. 107:3135-3140.

Fox, A.H., Y.W. Lam, A.K. Leung, C.E. Lyon, J. Andersen, M. Mann, and A.I. Lamond. 2002. Paraspeckles: a novel nuclear domain. *Curr Biol*. 12:13-25.

Francis, A.C., C. Di Primio, V. Quercioli, P. Valentini, A. Boll, G. Girelli, F. Demichelis, D. Arosio, and A. Cereseto. 2014. Second generation imaging of nuclear/cytoplasmic HIV-1 complexes. *AIDS Res Hum Retroviruses*. 30:717-726.

Francis, A.C., M. Marin, J. Shi, C. Aiken, and G.B. Melikyan. 2016. Time-Resolved Imaging of Single HIV-1 Uncoating In Vitro and in Living Cells. *PLoS Pathog*. 12:e1005709.

Francis, A.C., M. Marin, P.K. Singh, V. Achuthan, M.J. Prellberg, K. Palermino-Rowland, S. Lan, P.R. Tedbury, S.G. Sarafianos, A.N. Engelman, and G.B. Melikyan. 2020. HIV-1 replication complexes accumulate in nuclear speckles and integrate into speckle-associated genomic domains. *Nat Commun*. 11:3505.

Francis, A.C., and G.B. Melikyan. 2018. Single HIV-1 Imaging Reveals Progression of Infection through CA-Dependent Steps of Docking at the Nuclear Pore, Uncoating, and Nuclear Transport. *Cell Host Microbe*. 23:536-548 e536.

Geis, F.K., and S.P. Goff. 2019. Unintegrated HIV-1 DNAs are loaded with core and linker histones and transcriptionally silenced. *Proc Natl Acad Sci U S A*. 116:23735-23742.

Germier, T., S. Kocanova, N. Walther, A. Bancaud, H.A. Shaban, H. Sellou, A.Z. Politi, J. Ellenberg, F. Gallardo, and K. Bystricky. 2017. Real-Time Imaging of a Single Gene Reveals Transcription-Initiated Local Confinement. *Biophys J*. 113:1383-1394.

Graham, T.G., X. Wang, D. Song, C.M. Etson, A.M. van Oijen, D.Z. Rudner, and J.J. Loparo. 2014. ParB spreading requires DNA bridging. *Genes Dev*. 28:1228-1238.

Greig, J.A., T.A. Nguyen, M. Lee, A.S. Holehouse, A.E. Posey, R.V. Pappu, and G. Jedd. 2020. Arginine-Enriched Mixed-Charge Domains Provide Cohesion for Nuclear Speckle Condensation. *Mol Cell*. 77:1237-1250 e1234.

Hare, S., F. Di Nunzio, A. Labeja, J. Wang, A. Engelman, and P. Cherepanov. 2009. Structural basis for functional tetramerization of lentiviral integrase. *PLoS Pathog*. 5:e1000515.

Hrecka, K., C. Hao, M. Gierszewska, S.K. Swanson, M. Kesik-Brodacka, S. Srivastava, L. Florens, M.P. Washburn, and J. Skowronski. 2011. Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature*. 474:658-661.

Hubner, W., G.P. McNerney, P. Chen, B.M. Dale, R.E. Gordon, F.Y. Chuang, X.D. Li, D.M. Asmuth, T. Huser, and B.K. Chen. 2009. Quantitative 3D video microscopy of HIV transfer across T cell virological synapses. *Science*. 323:1743-1747.

Hulme, A.E., Z. Kelley, D. Foley, and T.J. Hope. 2015. Complementary Assays Reveal a Low Level of CA Associated with Viral Complexes in the Nuclei of HIV-1-Infected Cells. *J Virol*. 89:5350-5361.

880 Komatsu, T., C. Quentin-Froignant, I. Carlon-Andres, F. Lagadec, F. Rayne, J. Ragues, R.H. Kehlenbach,
881 W. Zhang, A. Ehrhardt, K. Bystricky, R. Morin, J.M. Lagarde, F. Gallardo, and H. Wodrich. 2018.
882 In Vivo Labelling of Adenovirus DNA Identifies Chromatin Anchoring and Biphasic Genome
883 Replication. *J Virol*. 92.
884 Lagache, T., A. Grassart, S. Dallongeville, O. Faklaris, N. Sauvonnet, A. Dufour, L. Danglot, and J.C. Olivo-
885 Marin. 2018. Mapping molecular assemblies with fluorescence microscopy and object-based
886 spatial statistics. *Nat Commun*. 9:698.
887 Laguette, N., B. Sobhian, N. Casartelli, M. Ringeard, C. Chable-Bessia, E. Segeral, A. Yatim, S. Emiliani,
888 O. Schwartz, and M. Benkirane. 2011. SAMHD1 is the dendritic- and myeloid-cell-specific HIV-
889 1 restriction factor counteracted by Vpx. *Nature*. 474:654-657.
890 Lamond, A.I., and D.L. Spector. 2003. Nuclear speckles: a model for nuclear organelles. *Nat Rev Mol*
891 *Cell Biol*. 4:605-612.
892 Lee, K., Z. Ambrose, T.D. Martin, I. Oztop, A. Mulky, J.G. Julias, N. Vandegraaff, J.G. Baumann, R. Wang,
893 W. Yuen, T. Takemura, K. Shelton, I. Taniuchi, Y. Li, J. Sodroski, D.R. Littman, J.M. Coffin, S.H.
894 Hughes, D. Unutmaz, A. Engelman, and V.N. KewalRamani. 2010. Flexible use of nuclear import
895 pathways by HIV-1. *Cell Host Microbe*. 7:221-233.
896 Lelek, M., N. Casartelli, D. Pellin, E. Rizzi, P. Souque, M. Severgnini, C. Di Serio, T. Fricke, F. Diaz-Griffero,
897 C. Zimmer, P. Charneau, and F. Di Nunzio. 2015. Chromatin organization at the nuclear pore
898 favours HIV replication. *Nat Commun*. 6:6483.
899 Lelek, M., F. Di Nunzio, R. Henriques, P. Charneau, N. Arhel, and C. Zimmer. 2012. Superresolution
900 imaging of HIV in infected cells with FlAsH-PALM. *Proc Natl Acad Sci U S A*. 109:8564-8569.
901 Liu, R., F.R. Simonetti, and Y.C. Ho. 2020. The forces driving clonal expansion of the HIV-1 latent
902 reservoir. *Virol J*. 17:4.
903 Llano, M., D.T. Saenz, A. Meehan, P. Wongthida, M. Peretz, W.H. Walker, W. Teo, and E.M. Poeschla.
904 2006. An essential role for LEDGF/p75 in HIV integration. *Science*. 314:461-464.
905 Maldarelli, F. 2016. The role of HIV integration in viral persistence: no more whistling past the proviral
906 graveyard. *J Clin Invest*. 126:438-447.
907 Mamede, J.I., G.C. Cianci, M.R. Anderson, and T.J. Hope. 2017. Early cytoplasmic uncoating is
908 associated with infectivity of HIV-1. *Proc Natl Acad Sci U S A*. 114:E7169-E7178.
909 Mariame, B., S. Kappler-Gratias, M. Kappler, S. Balor, F. Gallardo, and K. Bystricky. 2018. Real-Time
910 Visualization and Quantification of Human Cytomegalovirus Replication in Living Cells Using
911 the ANCHOR DNA Labeling Technology. *J Virol*. 92.
912 Marini, B., A. Kertesz-Farkas, H. Ali, B. Lucic, K. Lisek, L. Manganaro, S. Pongor, R. Luzzati, A. Recchia, F.
913 Mavilio, M. Giacca, and M. Lusic. 2015. Nuclear architecture dictates HIV-1 integration site
914 selection. *Nature*. 521:227-231.
915 Naganuma, T., and T. Hirose. 2013. Paraspeckle formation during the biogenesis of long non-coding
916 RNAs. *RNA Biol*. 10:456-461.
917 Olson, A., B. Basukala, W.W. Wong, and A.J. Henderson. 2019. Targeting HIV-1 proviral transcription.
918 *Curr Opin Virol*. 38:89-96.
919 Peng, K., W. Muranyi, B. Glass, V. Laketa, S.R. Yant, L. Tsai, T. Cihlar, B. Muller, and H.G. Krausslich.
920 2014. Quantitative microscopy of functional HIV post-entry complexes reveals association of
921 replication with the viral capsid. *Elife*. 3:e04114.
922 Petit, C., O. Schwartz, and F. Mammano. 1999. Oligomerization within virions and subcellular
923 localization of human immunodeficiency virus type 1 integrase. *J Virol*. 73:5079-5088.
924 Petit, C., O. Schwartz, and F. Mammano. 2000. The karyophilic properties of human immunodeficiency
925 virus type 1 integrase are not required for nuclear import of proviral DNA. *J Virol*. 74:7119-
926 7126.
927 Pradeepa, M.M., G.R. Grimes, G.C. Taylor, H.G. Sutherland, and W.A. Bickmore. 2014. Psip1/Ledgf p75
928 restrains Hox gene expression by recruiting both trithorax and polycomb group proteins.
929 *Nucleic Acids Res*. 42:9021-9032.

930 Pradeepa, M.M., H.G. Sutherland, J. Ule, G.R. Grimes, and W.A. Bickmore. 2012. Psip1/Ledgf p52 binds
931 methylated histone H3K36 and splicing factors and contributes to the regulation of alternative
932 splicing. *PLoS Genet*. 8:e1002717.
933 Price, A.J., D.A. Jacques, W.A. McEwan, A.J. Fletcher, S. Essig, J.W. Chin, U.D. Halambage, C. Aiken, and
934 L.C. James. 2014. Host cofactors and pharmacologic ligands share an essential interface in HIV-
935 1 capsid that is lost upon disassembly. *PLoS Pathog*. 10:e1004459.
936 Rensen, E., F. Mueller, V. Scoca, J.J. Parmar, P. Souque, C. Zimmer, and F. Di Nunzio. 2020. Clustering
937 and reverse transcription of HIV-1 genomes in nuclear niches of macrophages. *EMBO*
938 *J*:e105247.
939 Rensen, E., F. Mueller, V. Scoca, J.J. Parmar, P. Souque, C. Zimmer, and F. Di Nunzio. 2021. Clustering
940 and reverse transcription of HIV-1 genomes in nuclear niches of macrophages. *EMBO J*.
941 40:e105247.
942 Saad, H., F. Gallardo, M. Dalvai, N. Tanguy-le-Gac, D. Lane, and K. Bystricky. 2014. DNA dynamics during
943 early double-strand break processing revealed by non-intrusive imaging of living cells. *PLoS*
944 *Genet*. 10:e1004187.
945 Sanchez, A., D.I. Cattoni, J.C. Walter, J. Rech, A. Parmeggiani, M. Nollmann, and J.Y. Bouet. 2015.
946 Stochastic Self-Assembly of ParB Proteins Builds the Bacterial DNA Segregation Apparatus. *Cell*
947 *Syst*. 1:163-173.
948 Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden,
949 S. Saalfeld, B. Schmid, J.Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A.
950 Cardona. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods*.
951 9:676-682.
952 Schroder, A.R., P. Shinn, H. Chen, C. Berry, J.R. Ecker, and F. Bushman. 2002. HIV-1 integration in the
953 human genome favors active genes and local hotspots. *Cell*. 110:521-529.
954 Scoca, V., and F. Di Nunzio. 2021a. The HIV-1 Capsid: From Structural Component to Key Factor for
955 Host Nuclear Invasion. *Viruses*. 13.
956 Scoca, V., and F. Di Nunzio. 2021b. Membraneless organelles restructured and built by pandemic
957 viruses: HIV-1 and SARS-CoV-2. *J Mol Cell Biol*.
958 Sharkey, M., D.Z. Babic, T. Greenough, R. Gulick, D.R. Kuritzkes, and M. Stevenson. 2011. Episomal viral
959 cDNAs identify a reservoir that fuels viral rebound after treatment interruption and that
960 contributes to treatment failure. *PLoS Pathog*. 7:e1001303.
961 Sharkey, M.E., I. Teo, T. Greenough, N. Sharova, K. Luzuriaga, J.L. Sullivan, R.P. Bucy, L.G. Kostrikis, A.
962 Haase, C. Veryard, R.E. Davaro, S.H. Cheeseman, J.S. Daly, C. Bova, R.T. Ellison, 3rd, B. Mady,
963 K.K. Lai, G. Moyle, M. Nelson, B. Gazzard, S. Shaunak, and M. Stevenson. 2000. Persistence of
964 episomal HIV-1 infection intermediates in patients on highly active anti-retroviral therapy. *Nat*
965 *Med*. 6:76-81.
966 Shun, M.C., Y. Botbol, X. Li, F. Di Nunzio, J.E. Daigle, N. Yan, J. Lieberman, M. Lavigne, and A. Engelman.
967 2008. Identification and characterization of PWWP domain residues critical for LEDGF/p75
968 chromatin binding and human immunodeficiency virus type 1 infectivity. *J Virol*. 82:11555-
969 11567.
970 Shun, M.C., N.K. Raghavendra, N. Vandegraaff, J.E. Daigle, S. Hughes, P. Kellam, P. Cherepanov, and A.
971 Engelman. 2007. LEDGF/p75 functions downstream from preintegration complex formation to
972 effect gene-specific HIV-1 integration. *Genes Dev*. 21:1767-1778.
973 Singh, P.K., M.R. Plumb, A.L. Ferris, J.R. Iben, X. Wu, H.J. Fadel, B.T. Luke, C. Esnault, E.M. Poeschla, S.H.
974 Hughes, M. Kvaratskhelia, and H.L. Levin. 2015. LEDGF/p75 interacts with mRNA splicing
975 factors and targets HIV-1 integration to highly spliced genes. *Genes Dev*. 29:2287-2297.
976 Spector, D.L., and A.I. Lamond. 2011. Nuclear speckles. *Cold Spring Harb Perspect Biol*. 3.
977 Stultz, R.D., J.J. Cenker, and D. McDonald. 2017. Imaging HIV-1 Genomic DNA from Entry through
978 Productive Infection. *J Virol*. 91.
979 Suzuki, Y., and R. Craigie. 2007. The road to chromatin - nuclear entry of retroviruses. *Nat Rev*
980 *Microbiol*. 5:187-196.

Tantale, K., F. Mueller, A. Kozulic-Pirher, A. Lesne, J.M. Victor, M.C. Robert, S. Capozi, R. Chouaib, V. Backer, J. Mateos-Langerak, X. Darzacq, C. Zimmer, E. Basyuk, and E. Bertrand. 2016. A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat Commun*. 7:12248.

Tarantino, N., J.Y. Tinevez, E.F. Crowell, B. Boisson, R. Henriques, M. Mhlanga, F. Agou, A. Israel, and E. Laplantine. 2014. TNF and IL-1 exhibit distinct ubiquitin requirements for inducing NEMO-IKK supramolecular structures. *J Cell Biol*. 204:231-245.

Tinevez, J.Y., N. Perry, J. Schindelin, G.M. Hoopes, G.D. Reynolds, E. Laplantine, S.Y. Bednarek, S.L. Shorte, and K.W. Eliceiri. 2017. TrackMate: An open and extensible platform for single-particle tracking. *Methods*. 115:80-90.

Tsanov, N., A. Samacoits, R. Chouaib, A.M. Traboulsi, T. Gostan, C. Weber, C. Zimmer, K. Zibara, T. Walter, M. Peter, E. Bertrand, and F. Mueller. 2016. smiFISH and FISH-quant - a flexible single RNA detection approach with super-resolution capability. *Nucleic Acids Res*. 44:e165.

Vansant, G., H.C. Chen, E. Zorita, K. Trejbalova, D. Miklik, G. Filion, and Z. Debyser. 2020. The chromatin landscape at the HIV-1 provirus integration site determines viral expression. *Nucleic Acids Res*.

Wang, G.P., B.L. Levine, G.K. Binder, C.C. Berry, N. Malani, G. McGarrity, P. Tebas, C.H. June, and F.D. Bushman. 2009. Analysis of lentiviral vector integration in HIV+ study subjects receiving autologous infusions of gene modified CD4+ T cells. *Mol Ther*. 17:844-850.

Wang, H., M.C. Shun, X. Li, F. Di Nunzio, S. Hare, P. Cherepanov, and A. Engelman. 2014. Efficient Transduction of LEDGF/p75 Mutant Cells by Gain-of-Function HIV-1 Integrase Mutant Viruses. *Mol Ther Methods Clin Dev*. 1.

Zhu, Y., G.Z. Wang, O. Cingoz, and S.P. Goff. 2018. NP220 mediates silencing of unintegrated retroviral DNA. *Nature*. 564:278-282.

FIGURE LEGENDS

Figure 1. Detection and live tracking of nuclear HIV-1 DNA forms. A) Scheme of HIV-1 ANCHOR system created with BioRender.com. ANCH3 sequence has been cloned in the HIV-1 genome in place of Nef gene (HIV-1 ANCH3). The late retrotranscribed DNA of HIV-1 ANCH3 can be visualized in cells stably expressing the OR-GFP protein, thanks to OR accumulation on the double-stranded ANCH3 sequence. Viral DNA detection occurs mostly in the nucleus, where late reverse transcripts are abundant and are more likely to be exposed to the binding of OR-GFP. Starting from early time points of infection, it is possible to follow HIV-1 ANCH3 in target cells. The labelling results in the appearance of a bright fluorescent spot that can be tracked over time post infection. **B)** Confocal images of HeLa P4R5 cells expressing OR-GFP infected with HIV-1 ANCH3 (MOI 30), HIV-1 ANCH3 + RAL (20 µM), HIV-1 ANCH3 IN_{D116A} and HIV-1 ANCH3 + NEV (10 µM), 24 h p.i. The histogram plots of 2-LTR circles and ALU-PCR represent the nuclear import and integration rates, respectively. One-way ANOVA followed by Turkey's multiple comparison test, ns=non-significant, ****=p ≤ 0.0001. **C)** Confocal image of HIV-1 ANCH3 provirus in a single-cell clone. The clone was transduced with OR-GFP LV for the detection of HIV-1 integrated DNA. On the right, histogram plot of ALU-PCR and β-galactosidase expression of the HIV-1 ANCH3 single-cell clone compared to uninfected cells (NI). Unpaired t test, **=p ≤ 0.01, ****=p ≤ 0.0001. **D)** Tukey's box plots of the diffusion coefficient (D (nm^2/s)) of episomal and integrated HIV-1 DNA forms tracked in 3D time lapses (video 1, 2, 3). On the right, an example of the live track for each experimental group is shown. Low D corresponds to D (nm^2/s) < 5 and high D to 5 < D (nm^2/s) < 20. The color range indicates the time up to 10 min. All graphs contain ± SD. Scale bar: 5 µm.

Figure 2. HIV-1 transcriptional activity in dividing and non-dividing cells. Co-labeling of vDNA and vRNA by coupling HIV-1 ANCHOR with RNA FISH. **A)** Confocal images of HIV-1 RNA FISH in HeLa P4R5 cells (MOI 30) with or without RAL (20 µM), 24 h p.i. The scatter plots with bars show the mean fluorescence intensity of the vRNA signal per nucleus and of vDNA spots (vRNA: 13 (-RAL), 16 (+RAL) cells; vDNA: 95 (-RAL) and 141 (+RAL) spots). The analysis was performed in 2D. Unpaired t test, ****=p

≤ 0.0001. **B)** Confocal image of HIV-1 RNA FISH in Jurkat CD4+T cells (MOI 5), 48 h p.i. The scatter plot with bars shows the percentage of vDNA associated to vRNA signal and vice versa, per nucleus (12 cells; 262 vDNAs; 64 vRNA foci). The analysis was performed in 3D. Unpaired t test, **=p ≤ 0.01. **C)** Confocal image of HIV-1 RNA FISH in macrophage-like THP-1 cells (MOI 20) 3 days p.i. The scatter plot with bars shows the percentage of vDNA associated to vRNA signal and vice versa, per nucleus (8 cells; 130 vDNAs; 99 vRNA foci). The analysis was performed in 3D. Unpaired t test, *=p ≤ 0.05. **D)** Confocal image of HIV-1 RNA FISH in primary activated CD4+T cells (MOI 50), 3 days p.i. and primary monocyte-derived macrophages (MDMs), (MOI 40), 4 days p.i.. All graphs contain ± SD. Scale bar: 5 µm.

Figure 3. Nuclear localization of HIV-1 provirus in dividing and non-dividing cells. Co-labeling of vDNA and vRNA by coupling of HIV-1 ANCHOR with RNA FISH. **A)** Confocal images of HIV-1 RNA FISH in HeLa P4R5 cells infected with HIV-1 ANCH3 (MOI 5) at 48 h p.i. compared to 7 days p.i.. The scatter plot with bars indicates the distance from the nuclear envelope (NE) per detected vDNA spot (green) and vRNA focus (red) (n cells = 11 (48 h p.i.), 10 (+RAL), 5 (IN_{D116A}), and 9 (7 days p.i); vDNA: 83 (48 h p.i.), 87 (+RAL), 29 (IN_{D116A}), 16 (7 d p.i.); vRNA: 52 (48h p.i.) and 18 (7d p.i.)). One-way ANOVA followed by Turkey's multiple comparison test, ns=non-significant, *=p ≤ 0.05, **=p ≤ 0.01. **B)** Confocal image of HIV-1 RNA FISH in THP-1 cells infected with HIV-1 ANCH3 (MOI 20), 3 days p.i.. The scatter plot with bars indicates the distance from the nuclear envelope (NE) per detected vDNA (green) or vRNA (red) (n cells = 8 (vDNA, vRNA), 13 (IN_{D116A}); n spots= 139 (vDNA), 76 (IN_{D116A}), 126 (vRNA)). Unpaired t test, *=p ≤ 0.05, ***=p ≤ 0.001. **C)** Scatter plot with bars comparing the distances from the NE per detected vRNA focus in HeLa P4R5 and THP-1 cells (n = 11 (HeLa) and 8 (THP-1) cells; vRNAs: 52 (HeLa) and 126 (THP-1)). Unpaired t test, ns=non-significant. All the analyses were performed in 3D. All graphs contain ± SD. Scale bar: 5 µm, inset: 1 µm.

Figure 4. Sequestration of viral components in CPSF6 clusters. A) Late reverse transcripts separation from integrase (IN) foci: confocal images of infected primary activated $CD4^+$ T cells (MOI 100, 20 h p.i.) and macrophage-like cells, THP-1 cell line, (MOI 10, 30h p.i.), compared to uninfected cells. The scatter plot with bars compares

the percentages of vDNAs associated to IN, per nucleus (CD4+T = 8 cells, 54 vDNAs; THP-1 = 13 cells, 100 vDNAs). Unpaired t test, ns = non-significant. The analysis was performed in 3D. **B)** On the left, quantification of IN foci found associated to vDNA signal for each time-lapse video (at frame 0) (n = 13 cells, 81 IN foci). Six different videos were performed at sequential time p.i. On the right, example of vDNA separation in live, tracked down in the 18-hours p.i. video (video 6). Pearson's correlation per time frame is shown below. **C)** Confocal images of infected primary activated $CD4^+$ T cells (MOI 100, 20 h p.i.) and THP-1 cells (MOI 10, 30h p.i.), compared to uninfected cells. The graph shows the percentages of IN associated to CPSF6, per nucleus ($CD4^+T$ = 12 cells, 33 IN foci; THP-1 = 40 cells, 132 IN foci). Unpaired t test, ns = non-significant. **D)** Confocal images of Immuno-RNA FISH in SupT1R5 CD4+T cells and THP-1 cells infected with HIV-1 ANCH3 (MOI 20; 2 days p.i.). The graph shows the percentages of vRNA associated to CPSF6 per condition. (CD4+T = 19 (-NEV), 16 (+NEV) cells, 79 (-NEV), 30 (+NEV) vRNAs; THP-1 = 56 (-NEV), 34 (+NEV) cells, 151 (-NEV), 61 (+NEV) vRNAs). Unpaired t test, **=p ≤ 0.01, ****=p ≤ 0.0001. All graphs contain ± SD. Scale bar: 5 µm, inset: 1 µm.

Figure 5. HIV- MLOs as nuclear hubs of HIV-1 reverse transcription. A) Confocal images of Immuno-RNA FISH in THP-1 cells infected with HIV-1 ANCH3 (MOI 20) ± NEV (10 µM). The graph shows the average percentage per nucleus of vRNA found inside CPSF6 clusters, ± SEM (2 days p.i. = 56 cells, 151 vRNA (-NEV), 61 vRNA (+NEV); 3 days p.i. = 69 cells, 228 vRNA (-NEV), 129 vRNA (+NEV); 7 days p.i. = 45 cells, 148 vRNA (-NEV); 57 vRNA (+NEV). **B)** Chromatin Immunoprecipitation of endogenous CPSF6 coupled to real-time PCR of different reverse transcription (RT) products in THP-1 cells infected with HIV-1 (MOI 5), 2 days p.i. ± SD. One-way ANOVA followed by Turkey's multiple comparison test, *=p ≤ 0.05, **=p ≤ 0.01, ****=p ≤ 0.0001. ChIP-CPSF6 products compared to the input in infected cells by western blot using an Ab anti-CPSF6. **C)** Confocal images of THP-1 cells infected with HIV-1 ANCH3 (MOI 10) ± NEV (10 µM), 2 days p.i. Scale bars: 5 µm.

Figure 6. HIV-1 proviruses locate outside HIV-1 MLOs. A) Confocal images of THP-1 infected with HIV-1 ANCH3, 3 days p.i. The graph shows the intensity profiles of Hoechst, CPSF6 and SC35 signals along the segment crossing the MLO (yellow). **B)** Confocal images of THP-1 cells expressing OR-GFP infected with HIV-1 ANCH3, 3

days p.i., compared to uninfected cells. On the right, the graphs show: the comparison of Pearson's r coefficients of SON-SC35 colocalization between infected and uninfected cells, per nucleus (n = 25 cells (HIV-1 ANCH3), 19 cells (NI); unpaired t test, ns = non-significant); the percentages of CPSF6 clusters associated to vDNA or to SC35 in 3D, per nucleus (n = CPSF6 clusters > 50); the distance of each vDNA from the closest SC35-speckle border (n= 83 vDNAs, 7 cells). **C)** Confocal images of THP-1 cells expressing MCP-GFP infected with HIV-1 ANCH3 MS2 (MOI 2.5), 3 days p.i., ± NEV (10µM). The graph shows the intensity profiles of Hoechst, MCP-GFP, CPSF6 and SC35 signals along the segment crossing the MLO (yellow). **D)** Confocal images of THP-1 cells expressing MCP-GFP infected with HIV-1 ANCH3 MS2 (MOI 2.5), 3 days p.i., compared to uninfected cells. The graph shows the percentages of CPSF6 clusters associated to vRNA foci or to SC35 in 3D (n = 42 CPSF6 clusters, 8 cells) and the distance of each vRNA from the closest SC35-speckle border (n = 31 vRNAs, 8 cells). **E)** Scatter plot with bars indicating the percentage of vDNA or vRNA associated to CPSF6, per nucleus in 3D (n = 69 vDNAs, 30 vRNAs, 8 cells). Unpaired t test, ns = non-significant. **F)** Scatter plot with bars indicating the distance of vDNA or vRNA from the border of the closest SC35/CPSF6 MLO (n = 75 vDNAs, 7 cells; 34 vRNAs, 8 cells). Unpaired t test, **=p ≤ 0.01. **G)** Confocal image of Immuno-RNA FISH in THP-1 cell expressing OR-GFP infected with HIV-1 ANCH3 (MOI 20), 3 days p.i. Scale bar: 5 µm, inset: 1 µm.

Figure 7. HIV-1 transcribes outside HIV-1 MLOs in primary MDMs. Confocal images of Immuno-RNA FISH in primary MDMs infected with HIV-1 GFP (MOI 20) ± NEV (10 µM), 2-3 days p.i. The scatter plots show the GFP intensity per cell (n = 22 (-NEV) 20 (+NEV) cells) and the percentage of vRNA outside CPSF6 clusters per nucleus (n = 22 (-NEV) 20 (+NEV) cells; vRNAs: 268 (-NEV), 66 (+NEV)); unpaired t test, ****=p ≤ 0.0001. Their correlation in infected cells (Pearson's r coefficient) is displayed in the graph (n = 22 cells). The FACS histogram displays CD14+ expression of the MDMs. The FACS dot plot shows the correlation of GAG signal and GFP expression, with the related percentages per population. Scale bar: 5 µm

Figure 8. HIV-1 proviruses locate in LEDGF-enriched sites. A) Confocal images of THP-1 cells expressing OR-GFP infected with HIV-1 ANCH3, 3 days p.i., compared to uninfected cells. The scatter plots show the percentages of CPSF6 found associated

to LEDGF clusters per nucleus in 3D (n = 16 cells; unpaired t test, ***=p ≤ 0.001) and the LEDGF and CPSF6 count per nucleus in 3D (n = 16 cells); unpaired t test, ns = non-significant, ****=p ≤ 0.0001. Scale bar: 5 µm, inset: 2 µm. **B)** Confocal images of THP-1 cells expressing OR-GFP infected with HIV-1 ANCH3 and expressing MCP-GFP infected with HIV-1 ANCH3 MS2. The scatter plot shows the percentage of vDNA or vRNA associated to LEDGF in 3D (n= 33 (vDNA), 15 (vRNA)) ± SD. Unpaired t test, ns = non-significant. The pie charts show the percentages of the 2 populations of LEDGF clusters associated to vDNA or vRNA: LEDGF alone or LEDGF in proximity of CPSF6. Scale bar: 5 µm, inset: 1 µm.

Figure 9. Remodeling of the nuclear landscape by HIV-1 to replicate. New, refined model of viral intranuclear steps. Upon infection, HIV-1 triggers the formation of CPSF6 clusters which localize in nuclear speckles (NSs) (marked by SC35) forming HIV-1 MLOs, which retain viral genomic RNA and integrase proteins. HIV-1 MLOs are sites of the reverse transcription and of the HIV-1 PIC maturation. Once the mature PIC is formed, it leaves the CPSF6 cluster to proceed with the integration step. HIV-1 integration occurs in active chromatin regions enriched in LEDGF, close to the nuclear envelope and to HIV-1 MLOs. Indeed, HIV-1 replication might be favored for viral post-transcriptional processing by the proximity to NSs that host splicing factors. IN, Integrase; gRNA, viral genomic RNA; PIC, Pre-Integration Complex; NPC, Nuclear Pore Complex; vRNA, viral RNA; LTR, Long Terminal Repeat. Created with BioRender.com.

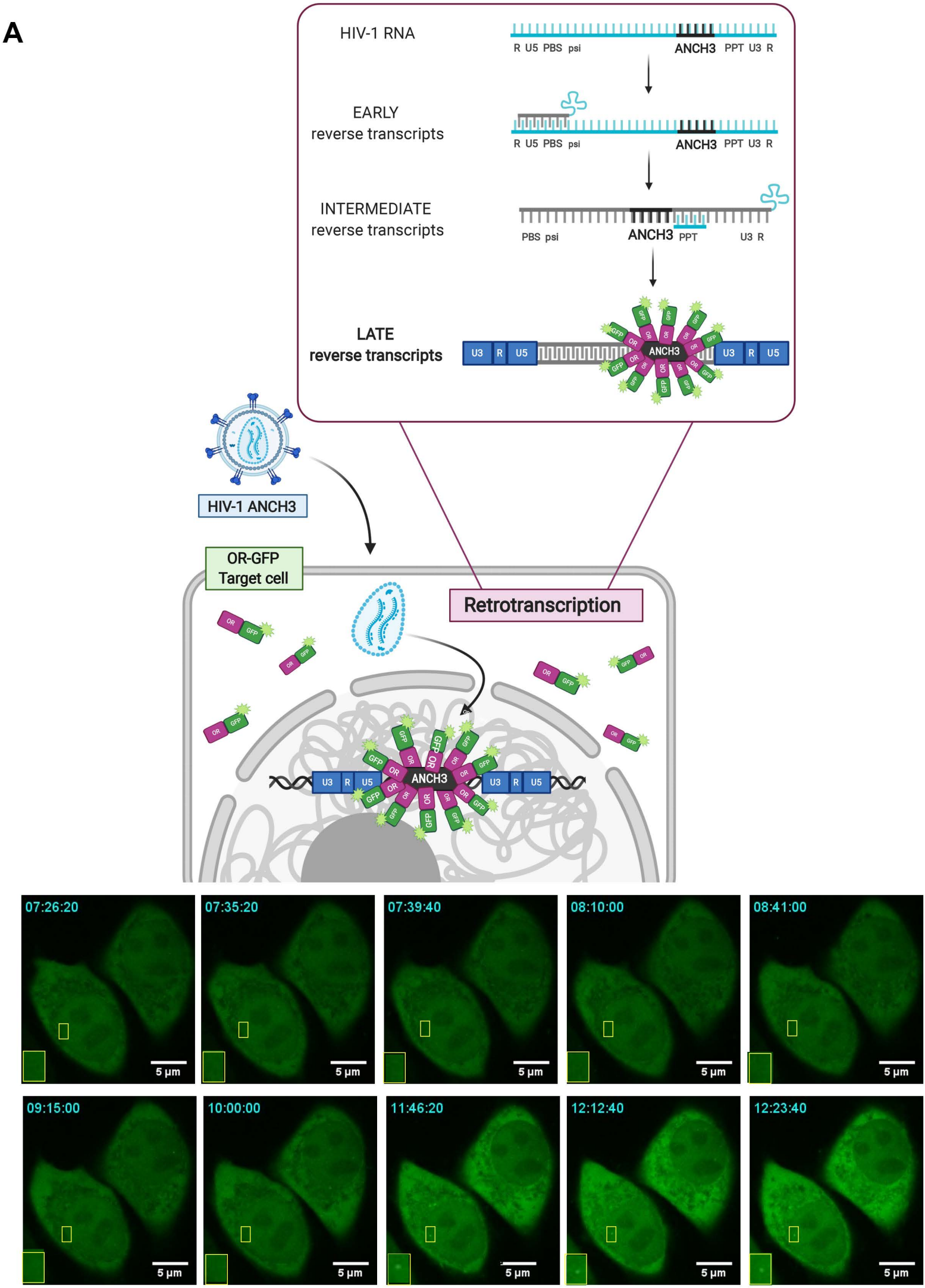

A
HIV-1 RNA
R U5 PBS psi
ANCH3
PPT U3 R
EARLY
reverse transcripts
R U5 PBS psi
ANCH3
PPT U3 R
INTERMEDIATE
reverse transcripts
PBS psi
ANCH3
PPT
U3 R
LATE
reverse transcripts
U3 R U5
ANCH3
U3 R U5
HIV-1 ANCH3
OR-GFP
Target cell
Retrotranscription
U3 R U5
ANCH3
U3 R U5
07:26:20
07:35:20
07:39:40
08:10:00
08:41:00
09:15:00
10:00:00
11:46:20
12:12:40
12:23:40
5 µm

Figure 1

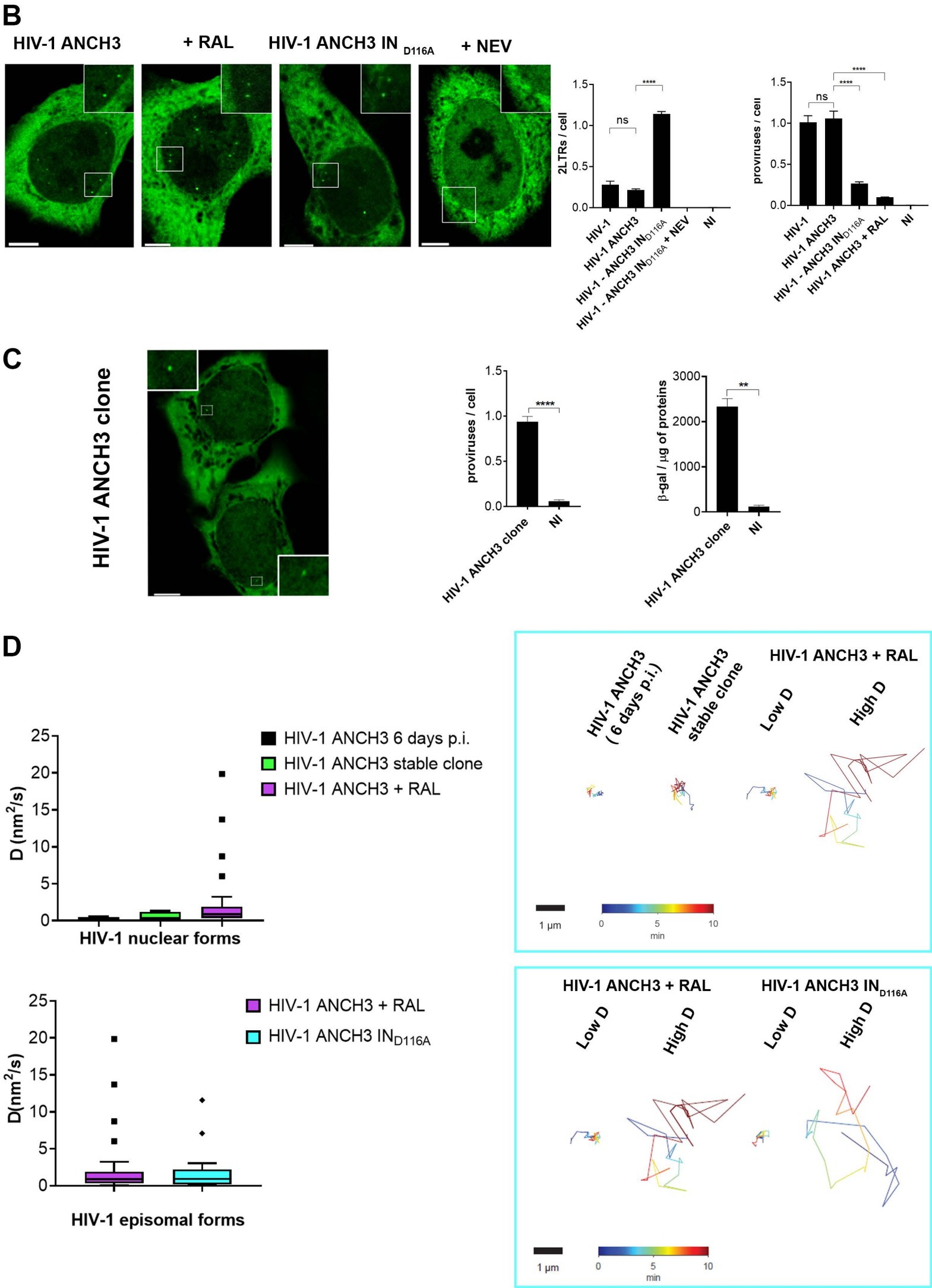

B
HIV-1 ANCH3
+ RAL
HIV-1 ANCH3 IN D116A
+ NEV
2LTRs / cell
ns
****
HIV-1
HIV-1 ANCH3
HIV-1 - ANCH3 IND116A
HIV-1 - ANCH3 IND116A + NEV
NI
proviruses / cell
HIV-1 ANCH3 + RAL
C
HIV-1 ANCH3 clone
β-gal / µg of proteins
**
D
HIV-1 ANCH3 6 days p.i.
HIV-1 ANCH3 stable clone
HIV-1 ANCH3 + RAL
D (nm2/s)
HIV-1 nuclear forms
HIV-1 ANCH3 IND116A
D(nm2/s)
HIV-1 episomal forms
HIV-1 ANCH3 (6 days p.i.)
Low D
High D
1 µm
min

Figure 1

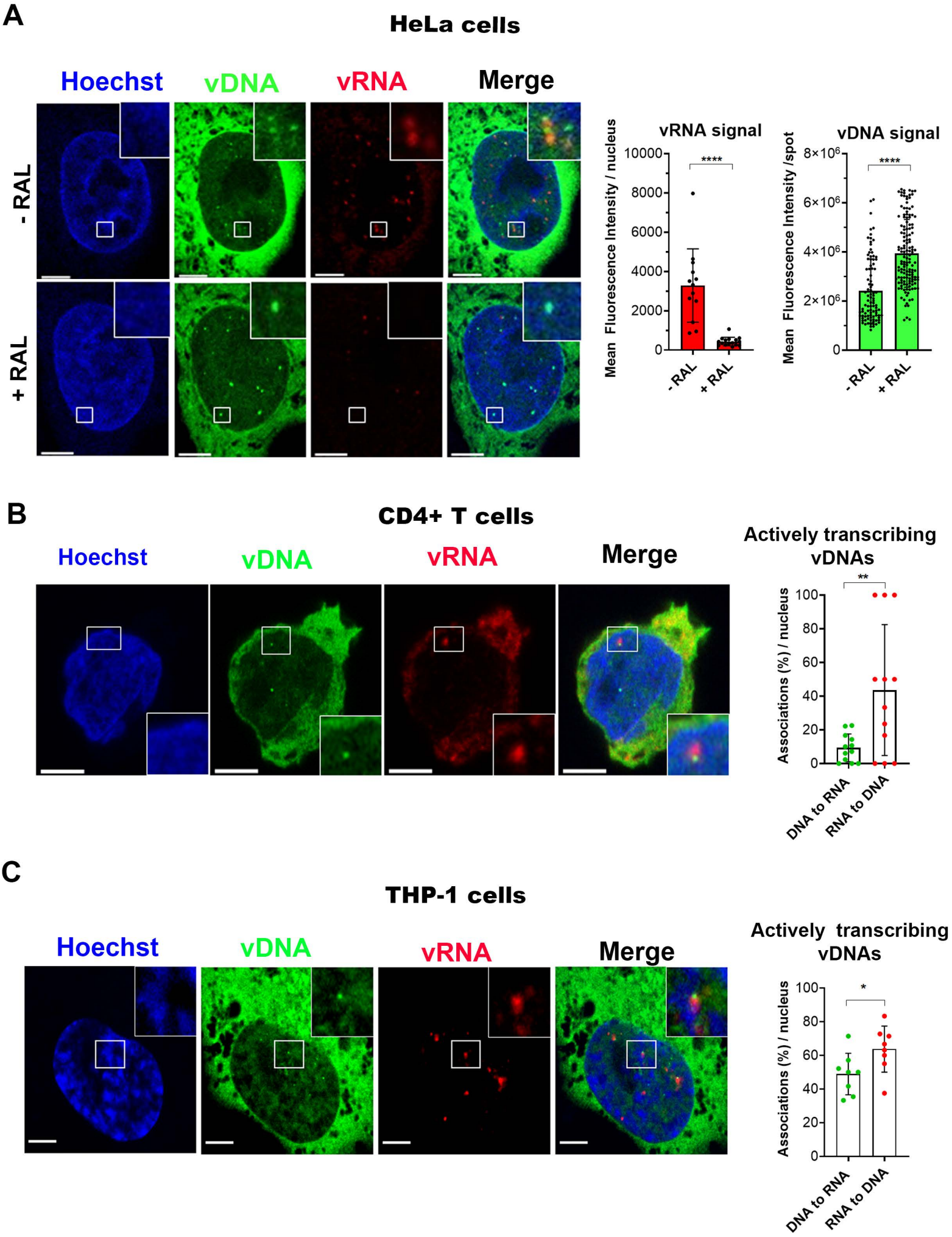
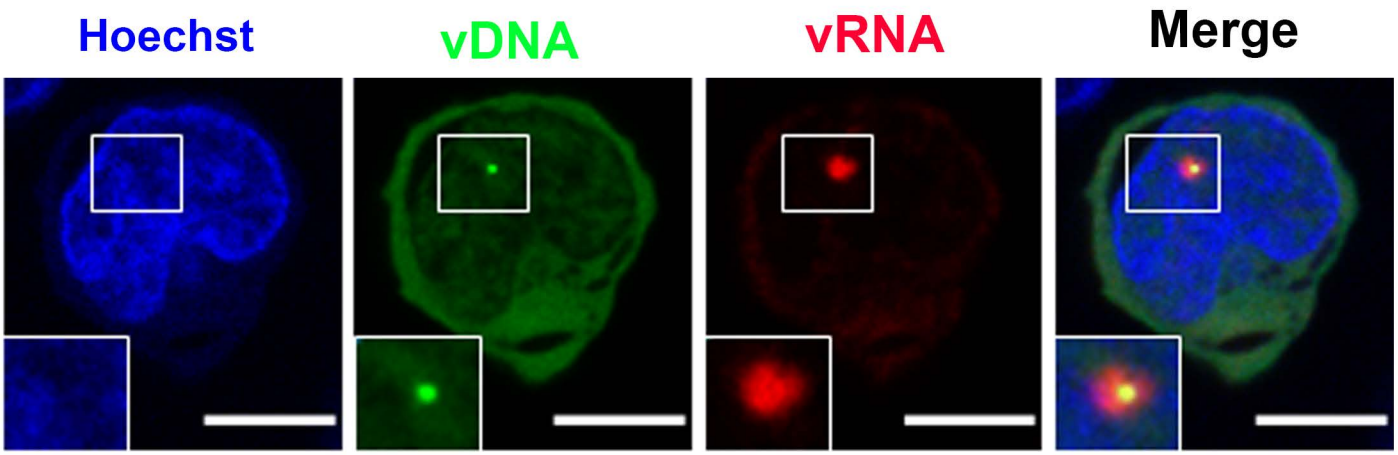

A
HeLa cells
Hoechst
vDNA
vRNA
Merge
- RAL
+ RAL
vRNA signal
Mean Fluorescence Intensity / nucleus
10000
8000
6000
4000
2000
0
****
- RAL
+ RAL
vDNA signal
Mean Fluorescence Intensity /spot
8×10^6
6×10^6
4×10^6
2×10^6
0
****
- RAL
+ RAL
B
CD4+ T cells
Hoechst
vDNA
vRNA
Merge
Actively transcribing vDNAs
Associations (%) / nucleus
100
80
60
40
20
0
**
DNA to RNA
RNA to DNA
C
THP-1 cells
Hoechst
vDNA
vRNA
Merge
Actively transcribing vDNAs
Associations (%) / nucleus
100
80
60
40
20
0
*
DNA to RNA
RNA to DNA

Figure 2

D

Primary CD4+ T cells



Primary macrophages

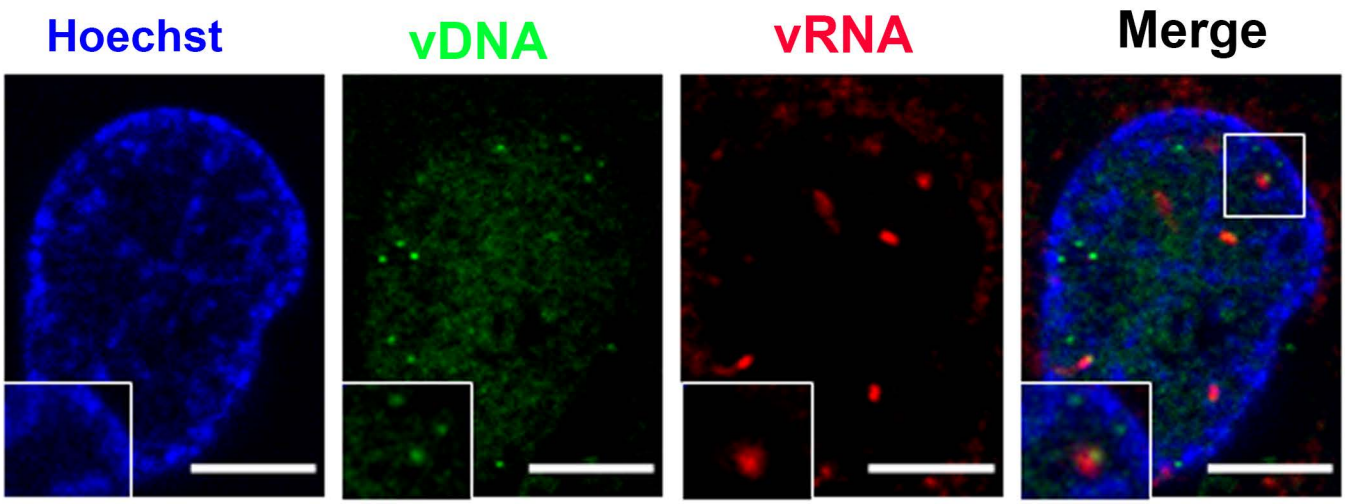
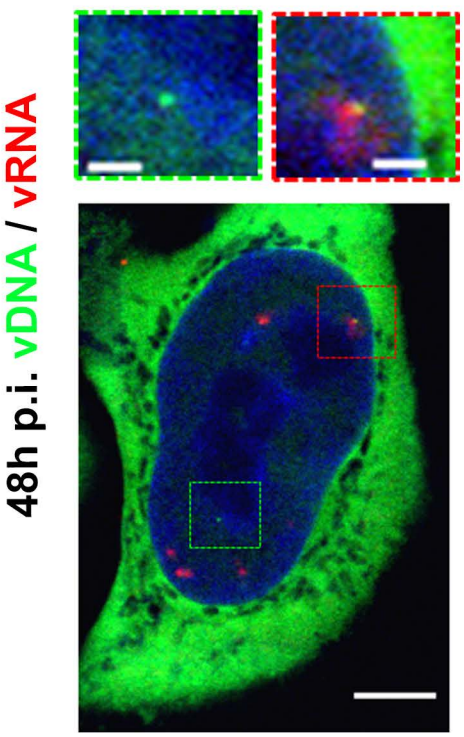
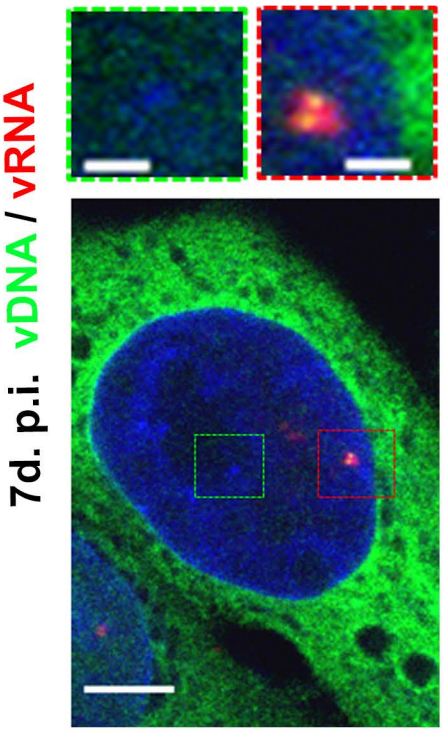


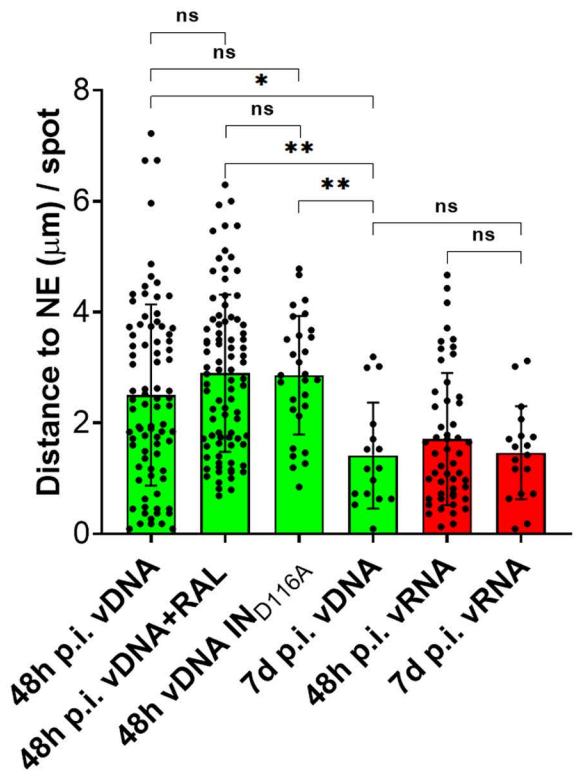
Figure 2

A

Dividing cells: HeLa

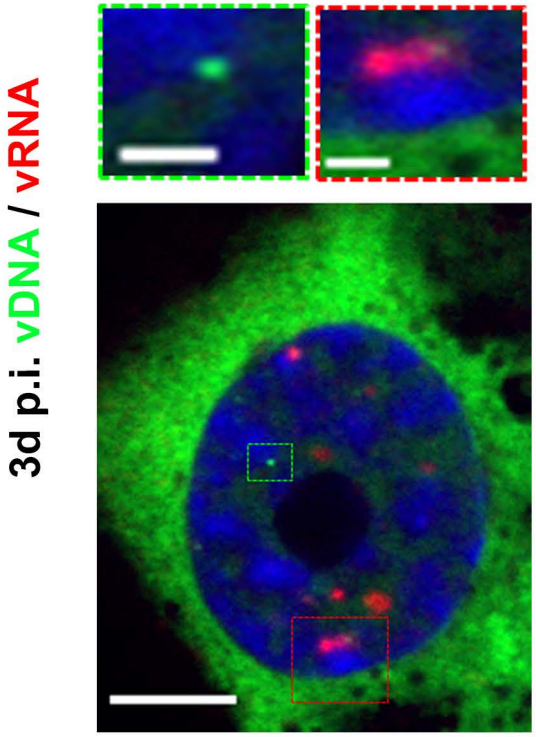


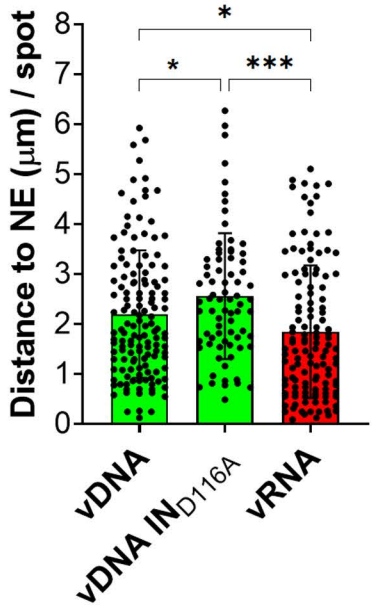




B

Non-dividing cells: THP-1





C

HeLa vs THP-1

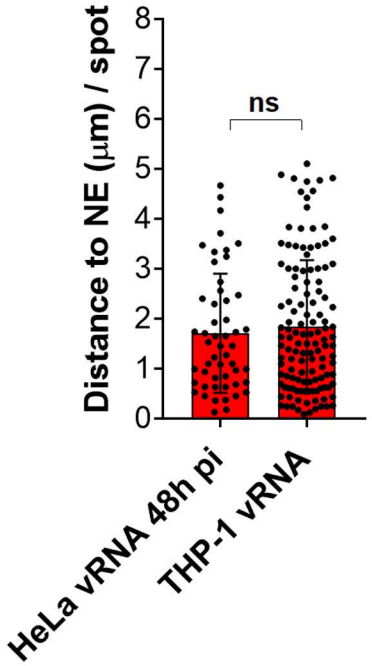
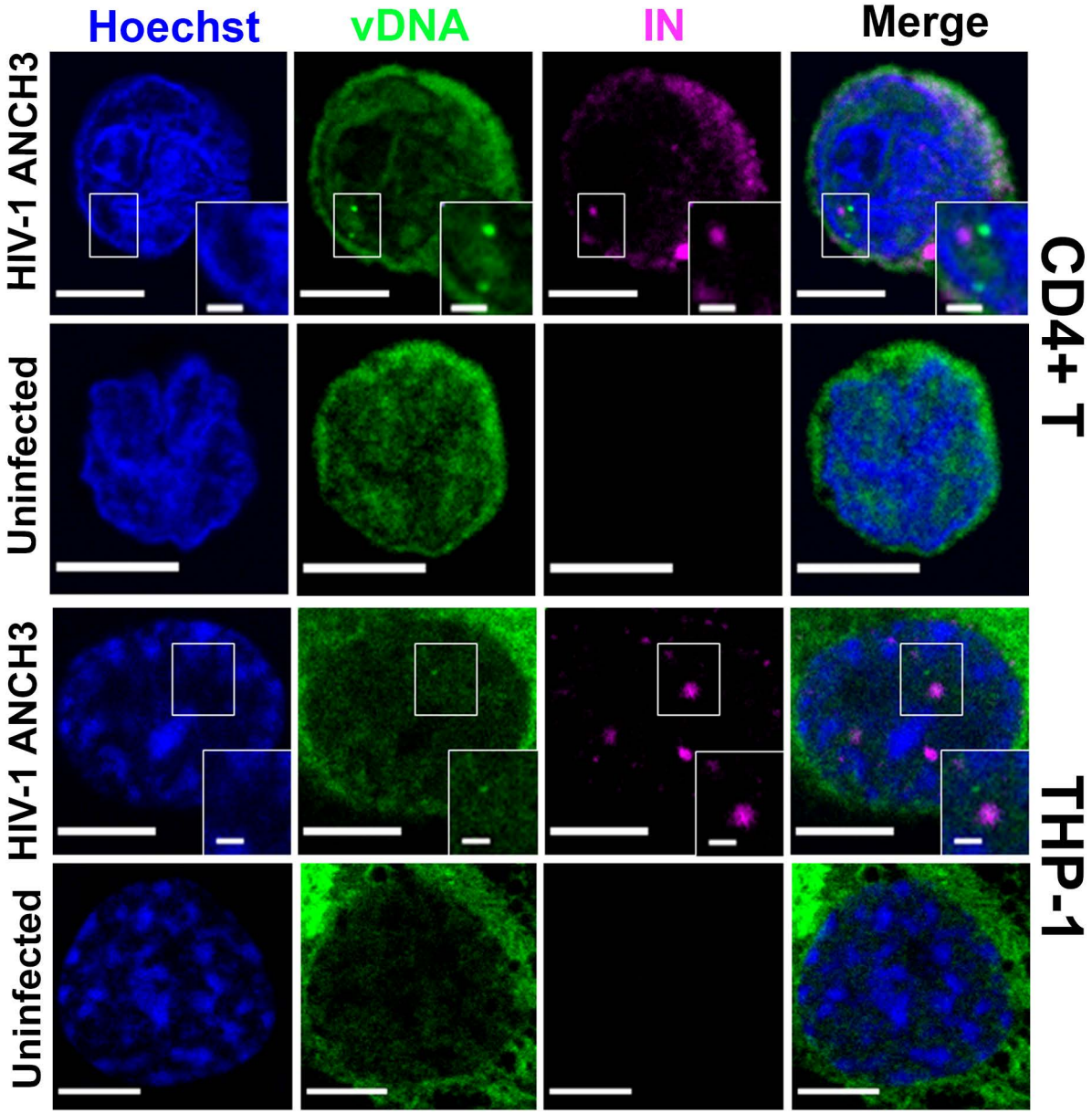
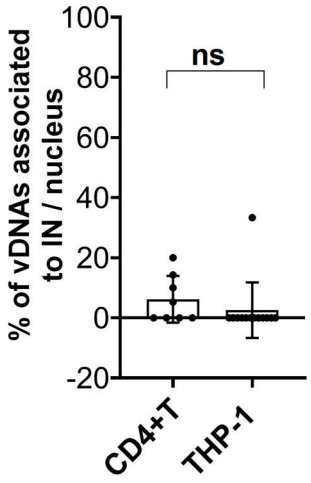


Figure 3

A


Hoechst
vDNA
IN
Merge
HIV-1 ANCH3
Uninfected
CD4+ T
HIV-1 ANCH3
Uninfected
THP-1


% of vDNAs associated to IN / nucleus
100
80
60
40
20
0
-20
ns
CD4+T
THP-1

B

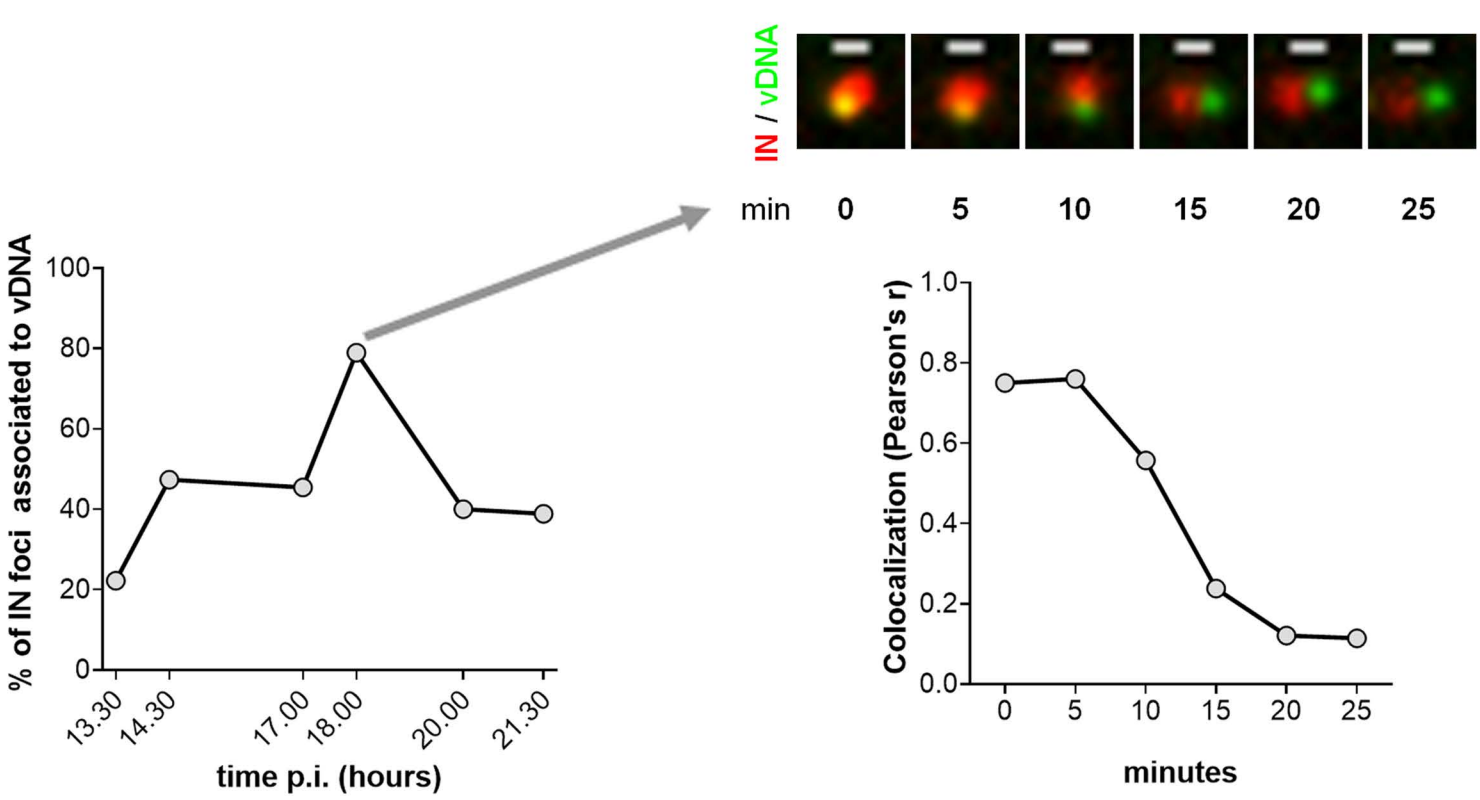
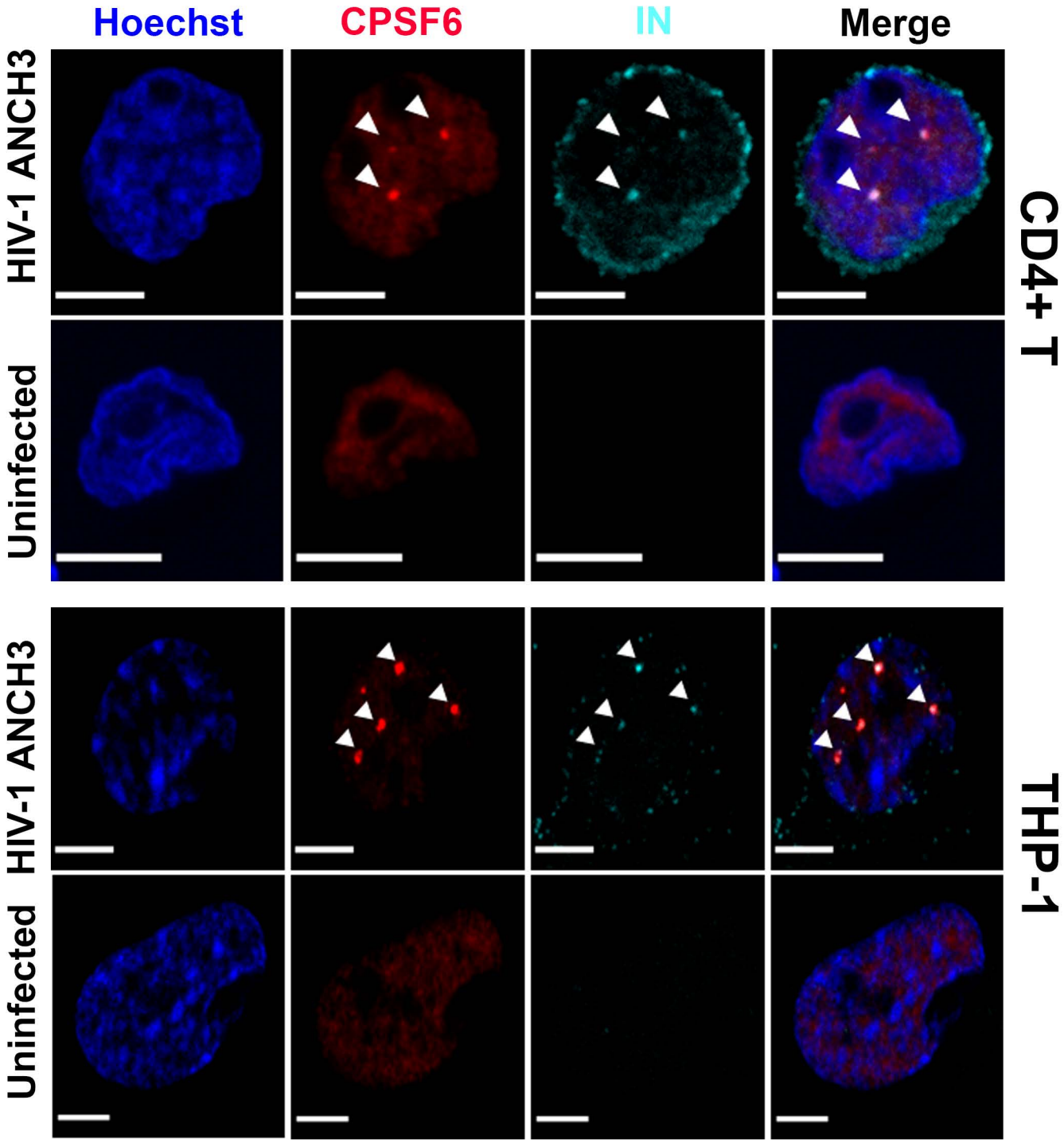

IN / vDNA
min
0
5
10
15
20
25
% of IN foci associated to vDNA
100
80
60
40
20
0
13.30
14.30
17.00
18.00
20.00
21.30
time p.i. (hours)
Colocalization (Pearson's r)
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
minutes

Figure 4


C
Hoechst
CPSF6
IN
Merge
HIV-1 ANCH3
Uninfected
CD4+ T
HIV-1 ANCH3
Uninfected
THP-1

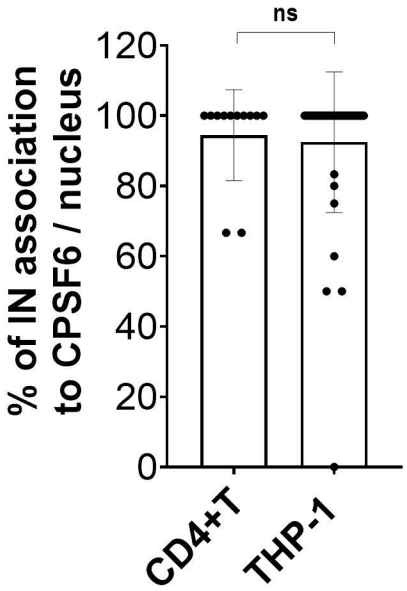
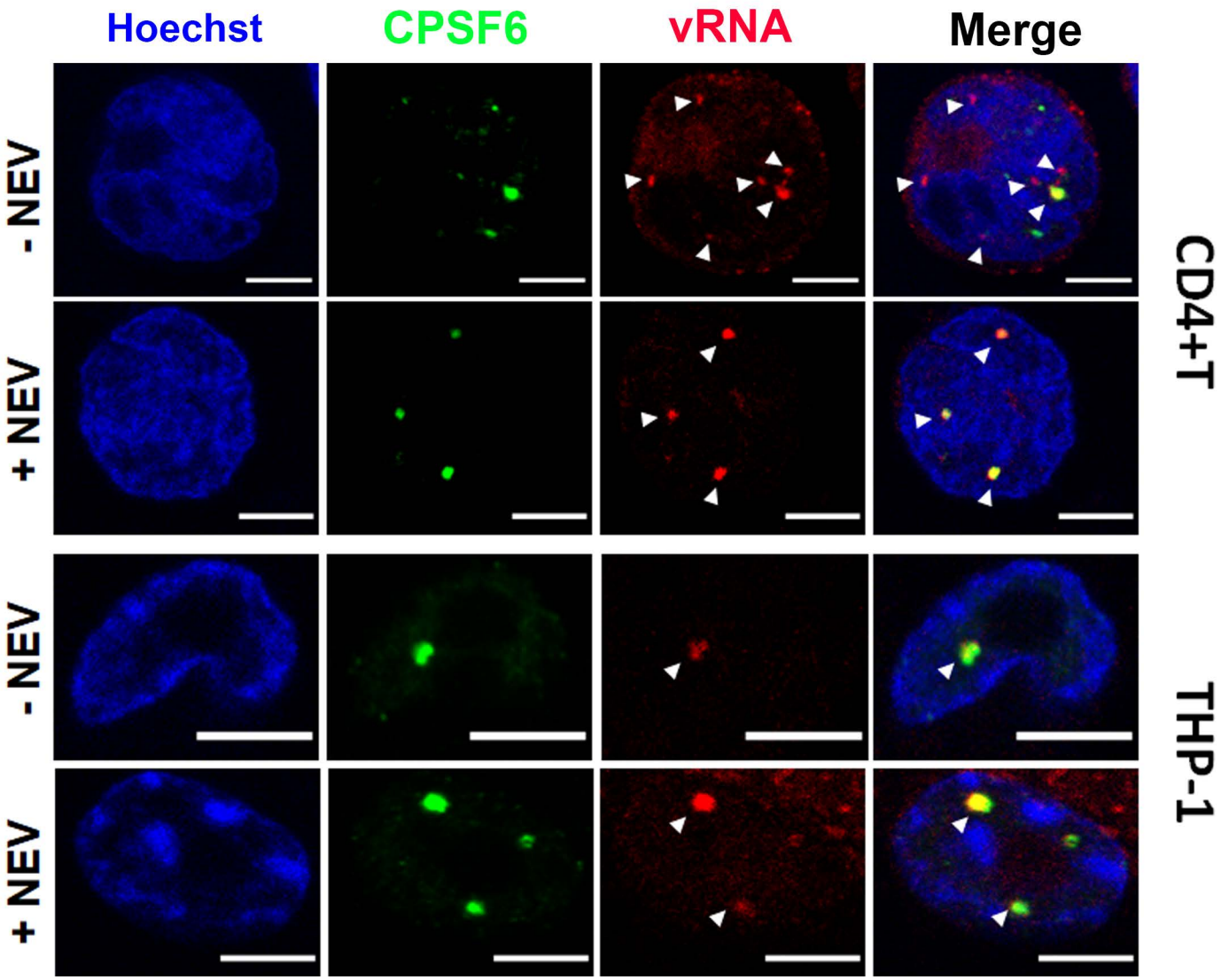

ns
% of IN association to CPSF6 / nucleus
120
100
80
60
40
20
0
CD4+T
THP-1

Figure 4

D


Hoechst
CPSF6
vRNA
Merge
- NEV
+ NEV
- NEV
+ NEV
CD4+T
THP-1

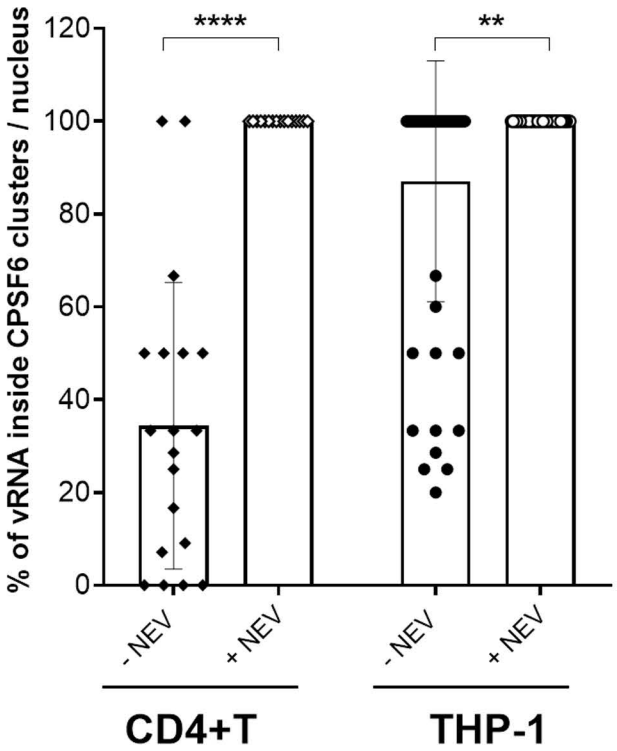
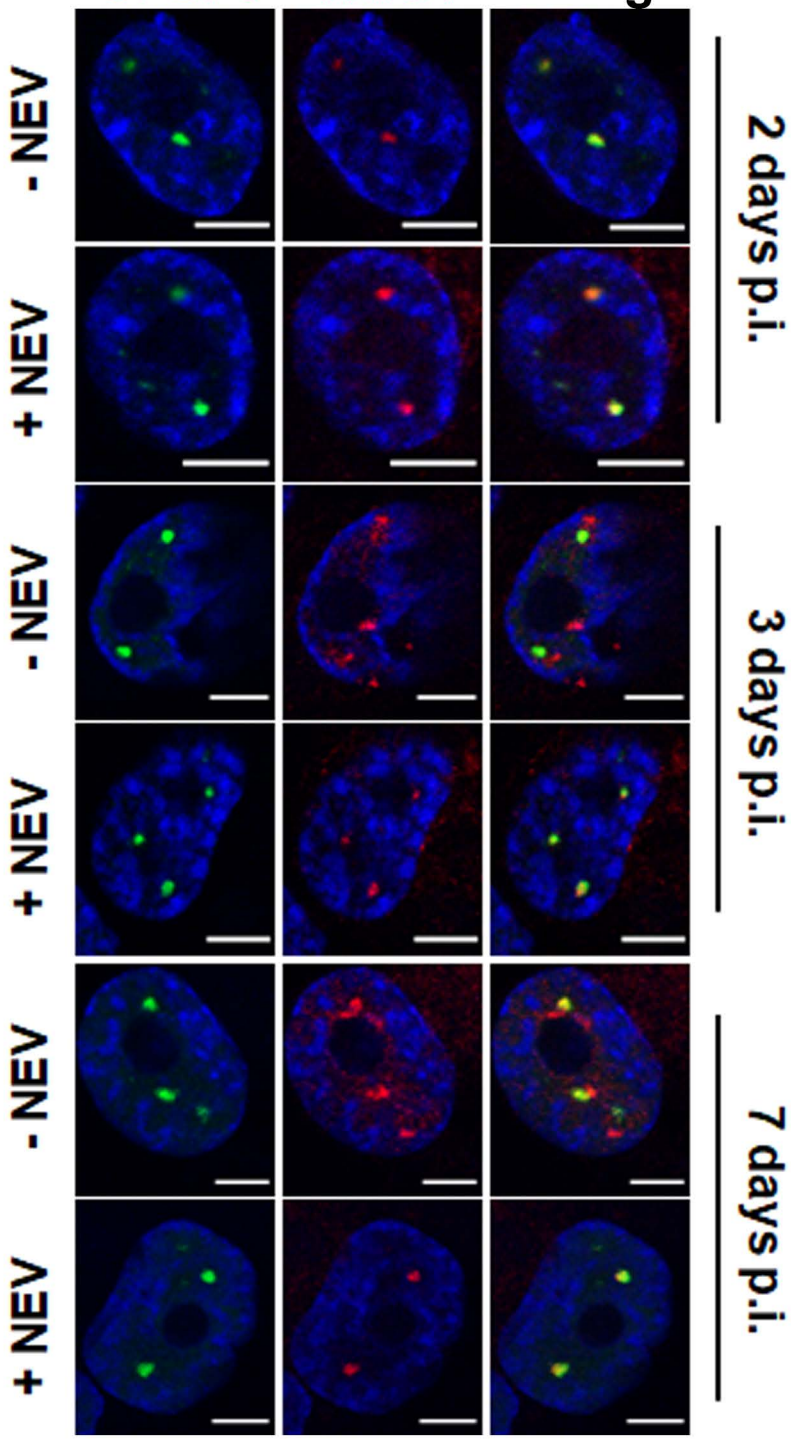
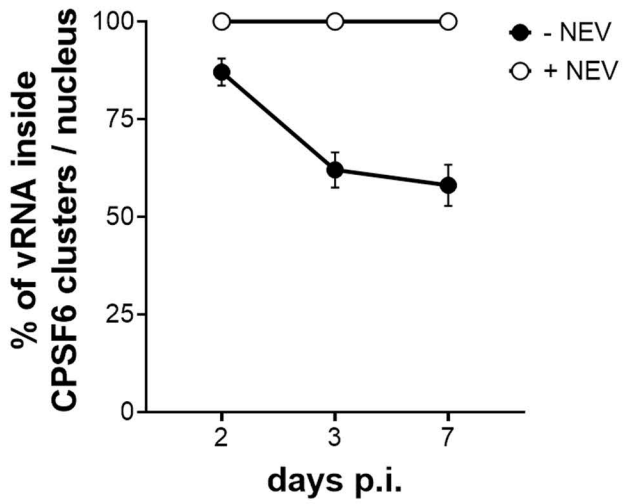

% of vRNA inside CPSF6 clusters / nucleus
120
100
80
60
40
20
0
****
**
- NEV
+ NEV
- NEV
+ NEV
CD4+T
THP-1

Figure 4

A


CPSF6
vRNA
Merge
- NEV
+ NEV
- NEV
+ NEV
- NEV
+ NEV
2 days p.i.
3 days p.i.
7 days p.i.


% of vRNA inside CPSF6 clusters / nucleus
100
75
50
25
0
2
3
7
days p.i.
- NEV
+ NEV

B

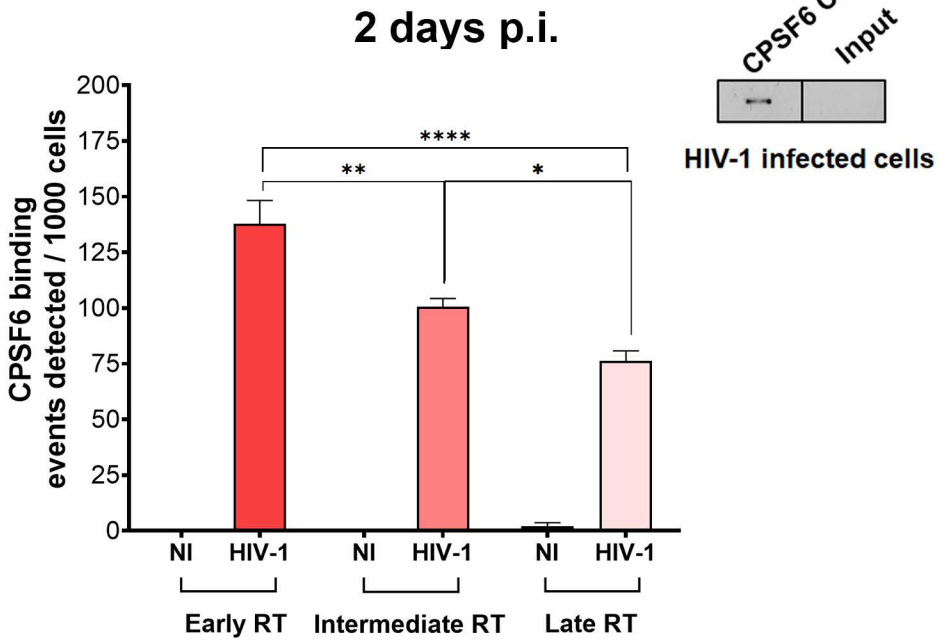

2 days p.i.
CPSF6 binding events detected / 1000 cells
200
175
150
125
100
75
50
25
0
****
**
*
NI
HIV-1
NI
HIV-1
NI
HIV-1
Early RT
Intermediate RT
Late RT
CPSF6 ChIP
Input
HIV-1 infected cells

Figure 5

C

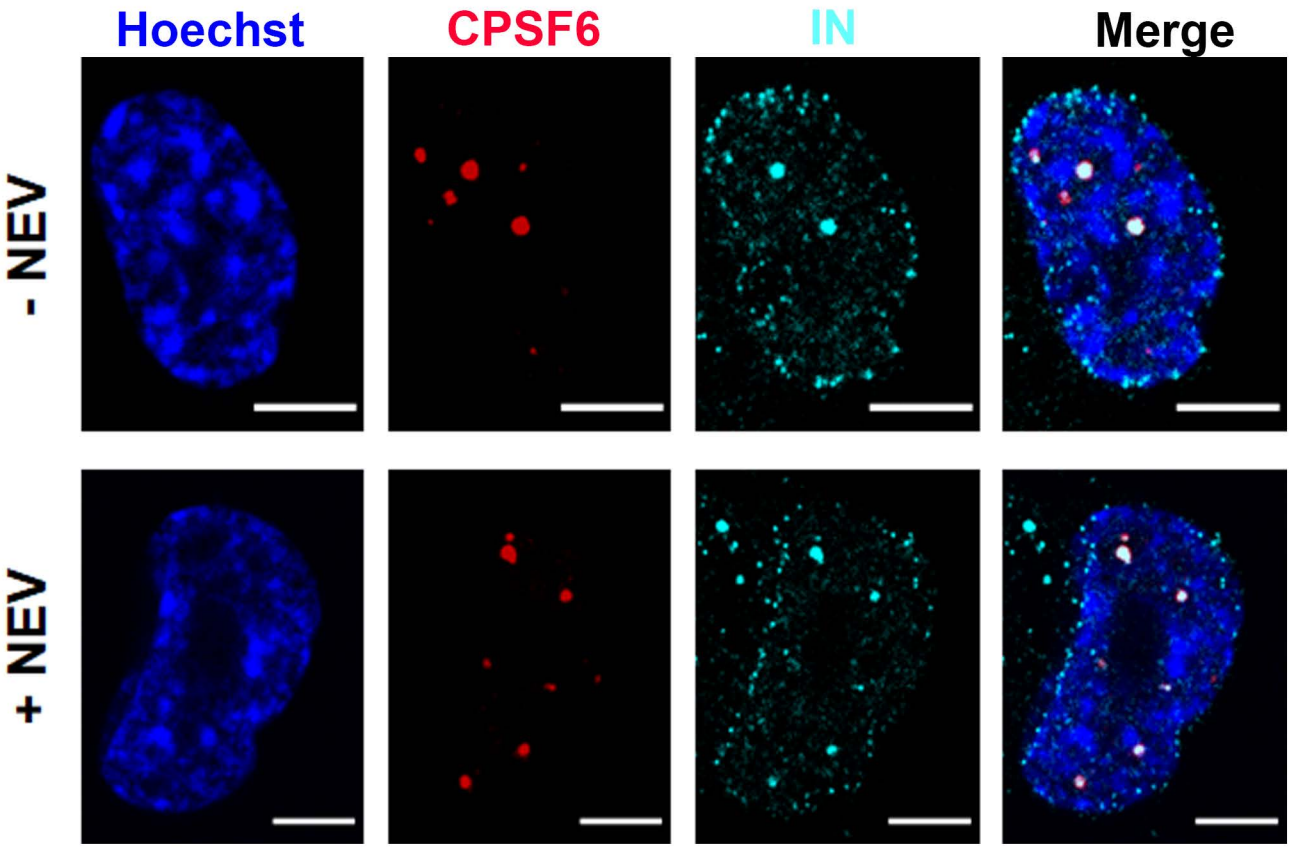
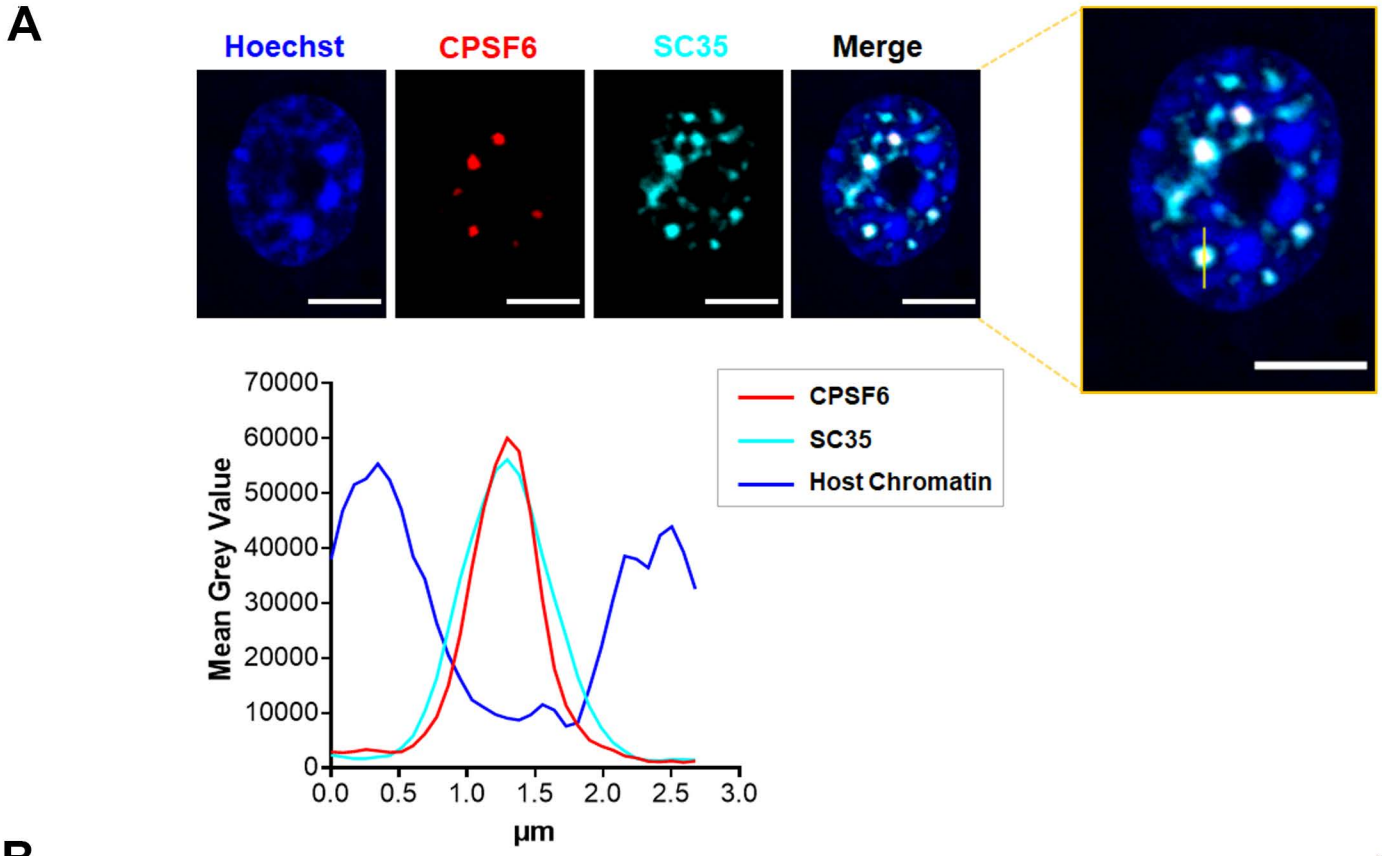


Figure 5


A
Hoechst
CPSF6
SC35
Merge
Mean Grey Value
70000
60000
50000
40000
30000
20000
10000
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
µm
CPSF6
SC35
Host Chromatin

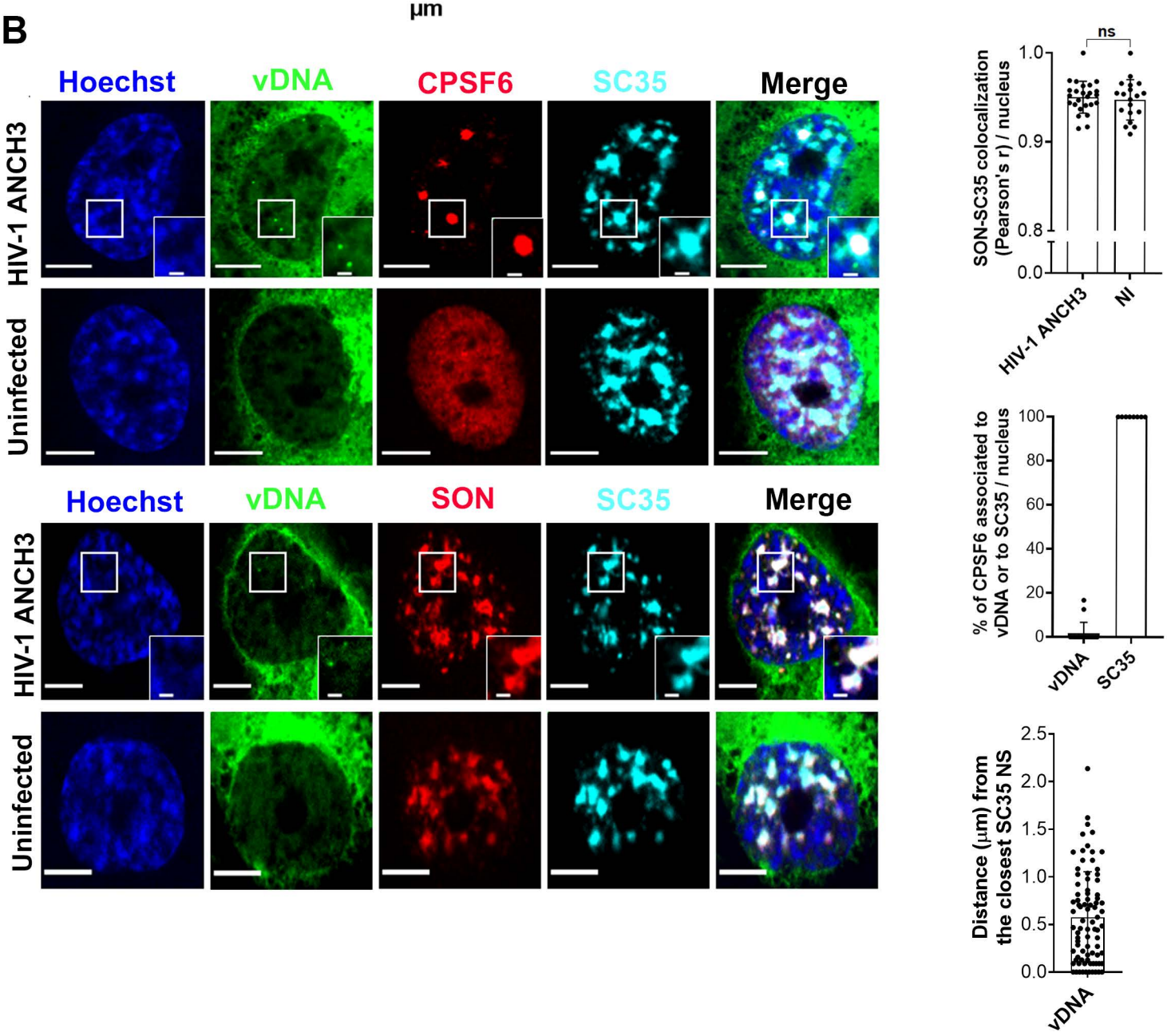
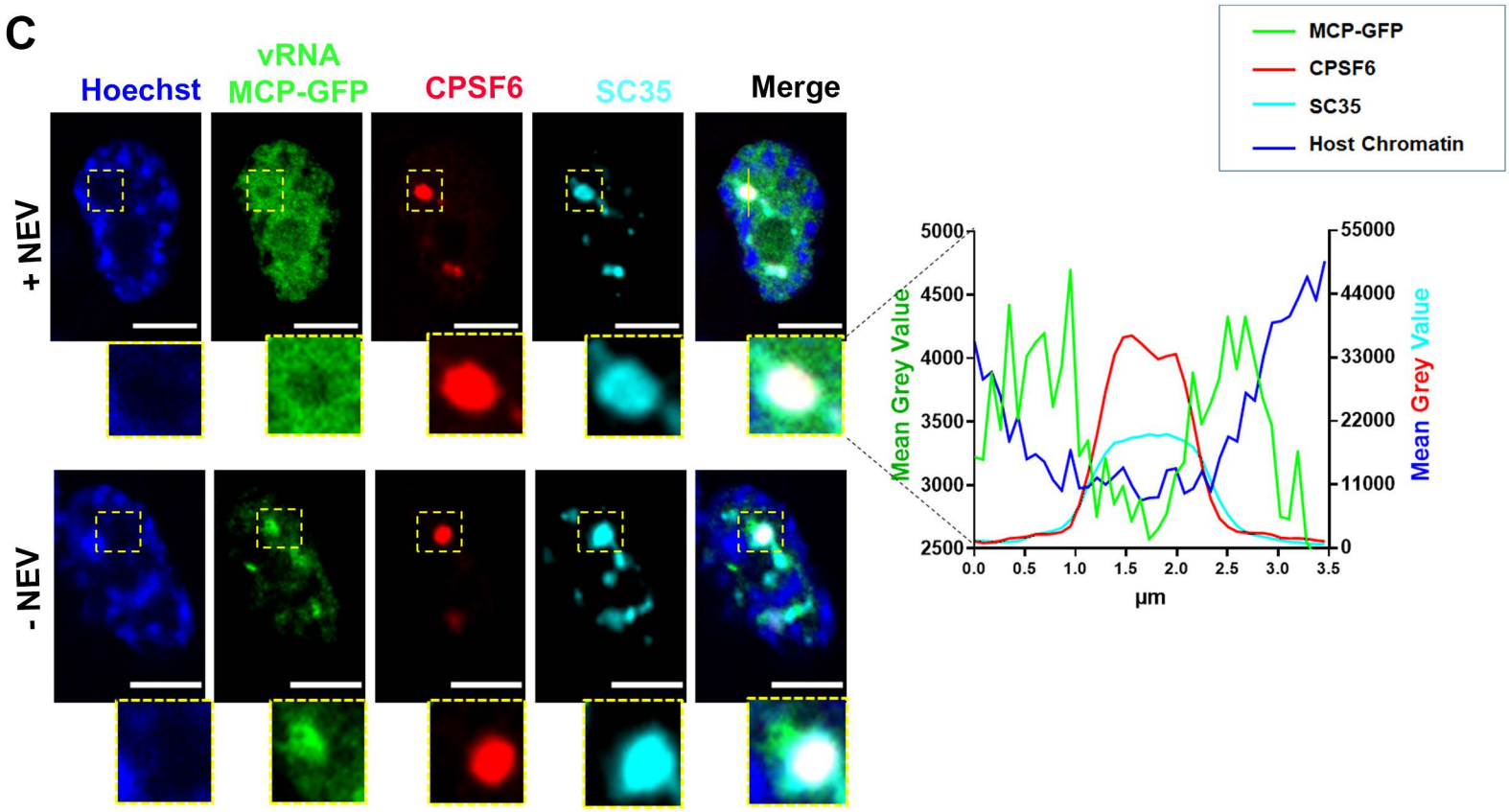
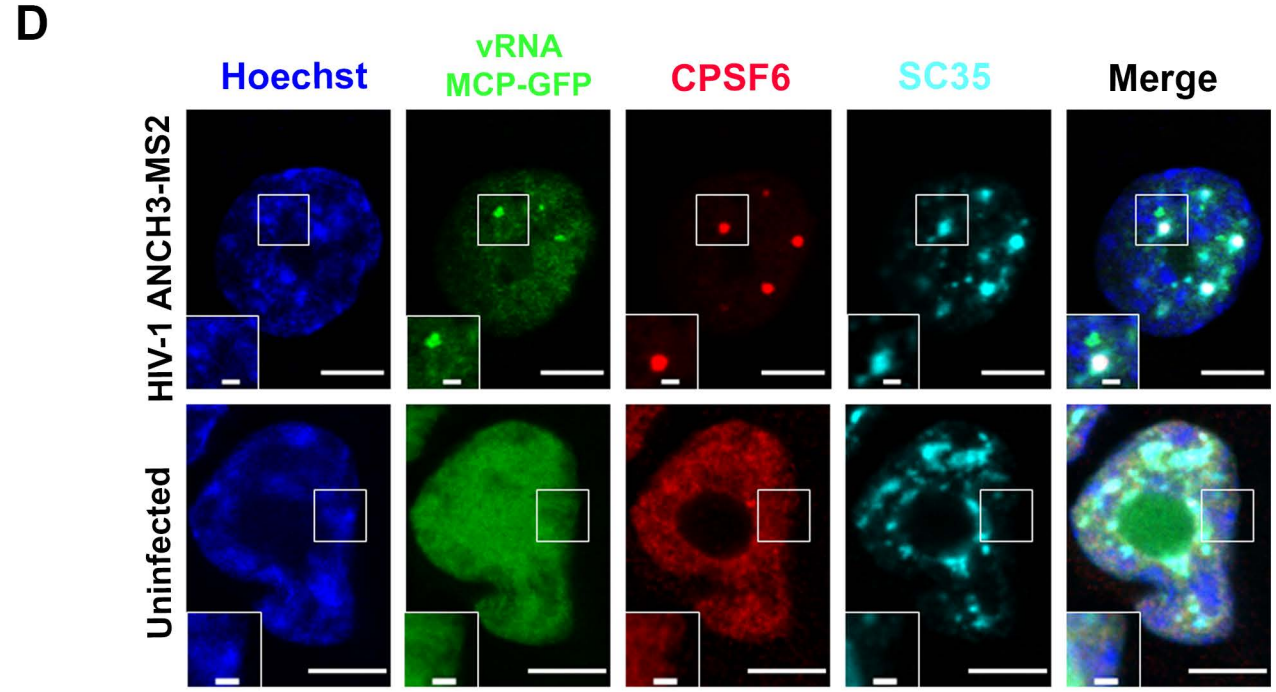

B
Hoechst
vDNA
CPSF6
SC35
Merge
HIV-1 ANCH3
Uninfected
Hoechst
vDNA
SON
SC35
Merge
HIV-1 ANCH3
Uninfected
ns
SON-SC35 colocalization (Pearson's r) / nucleus
1.0
0.9
0.8
0.0
HIV-1 ANCH3
NI
% of CPSF6 associated to vDNA or to SC35 / nucleus
100
80
60
40
20
0
vDNA
SC35
Distance (µm) from the closest SC35 NS
2.5
2.0
1.5
1.0
0.5
0.0
vDNA

Figure 6


C
Hoechst
vRNA
MCP-GFP
CPSF6
SC35
Merge
+ NEV
- NEV
MCP-GFP
CPSF6
SC35
Host Chromatin
Mean Grey Value
Mean Grey Value
µm


D
Hoechst
vRNA
MCP-GFP
CPSF6
SC35
Merge
HIV-1 ANCH3-MS2
Uninfected

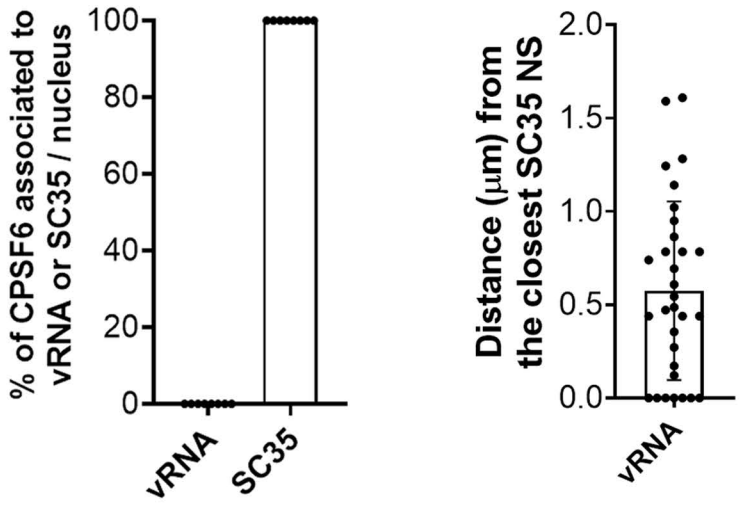
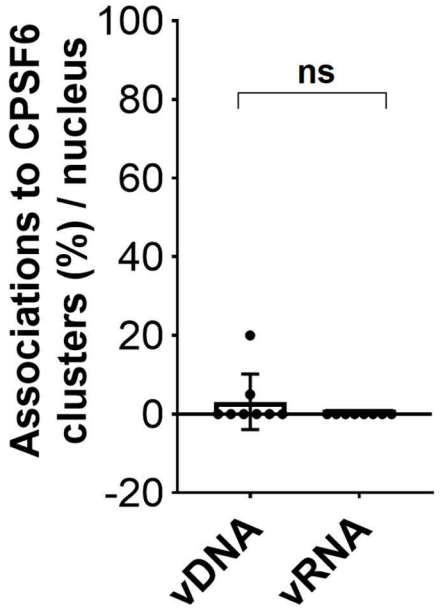
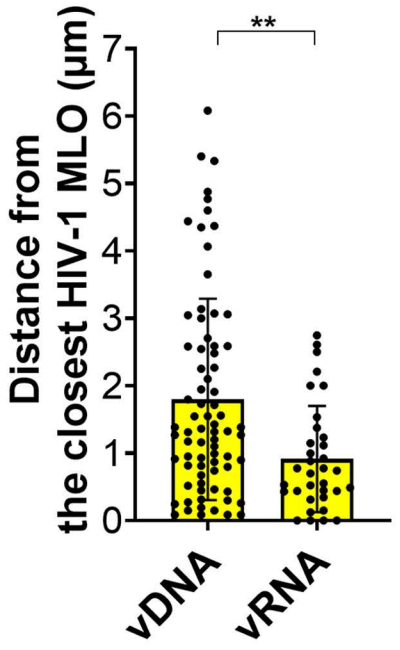

% of CPSF6 associated to vRNA or SC35 / nucleus
vRNA
SC35
Distance (µm) from the closest SC35 NS
vRNA

Figure 6

E


Associations to CPSF6 clusters (%) / nucleus
100
80
60
40
20
0
-20
ns
vDNA
vRNA

F


Distance from the closest HIV-1 MLO (µm)
7
6
5
4
3
2
1
0
**
vDNA
vRNA

G

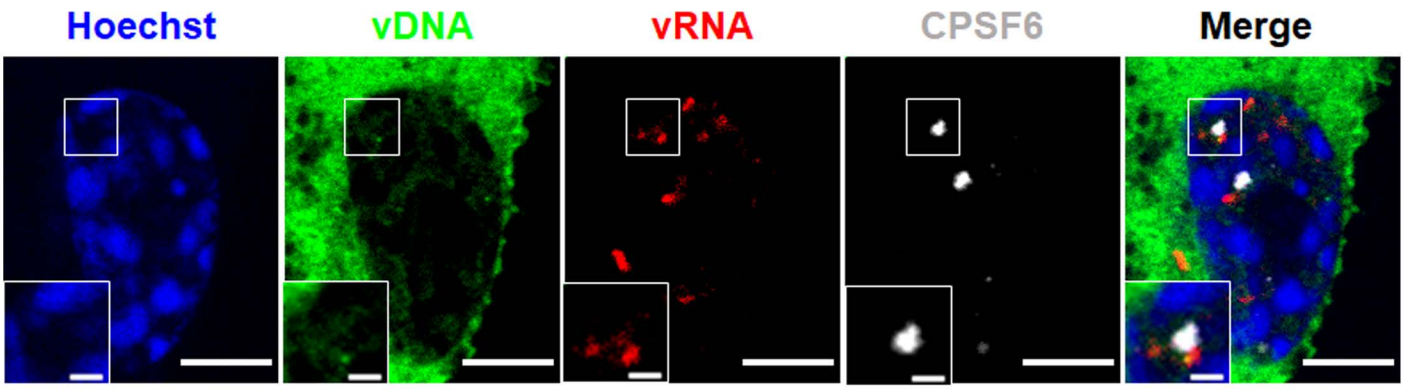
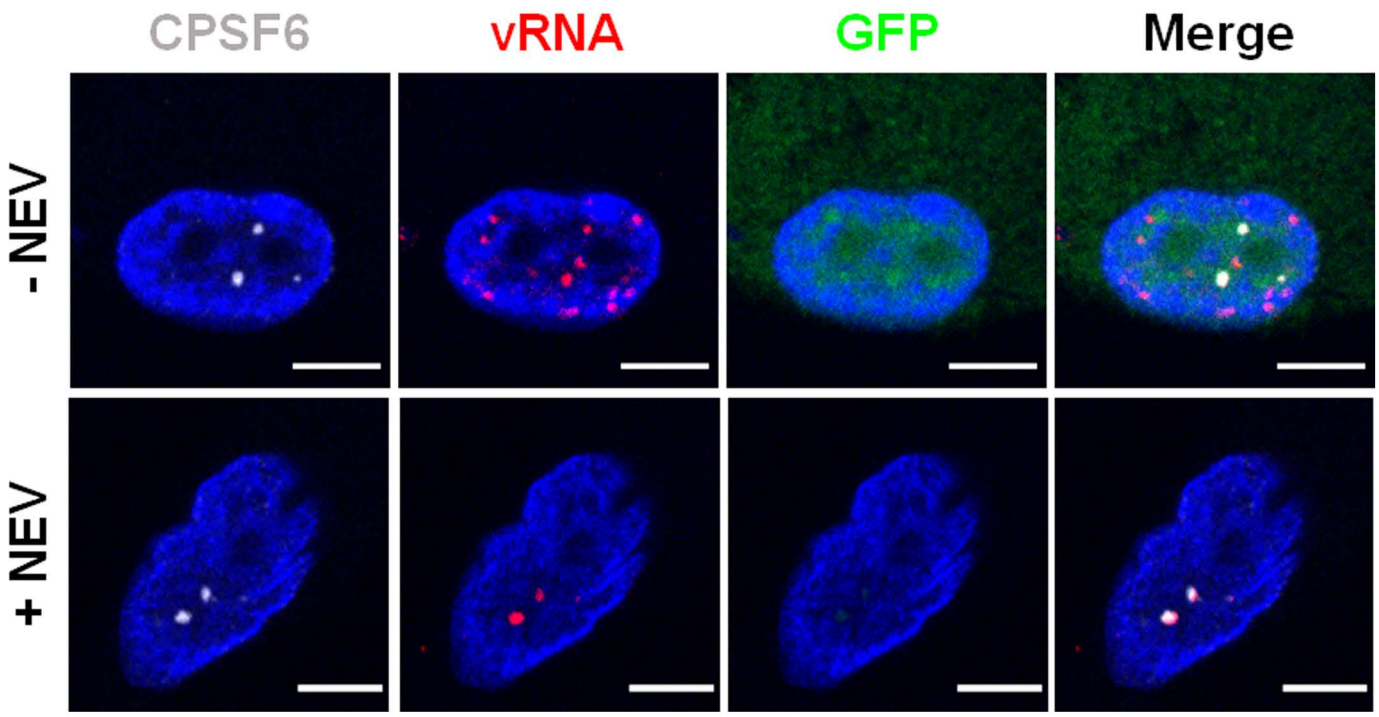
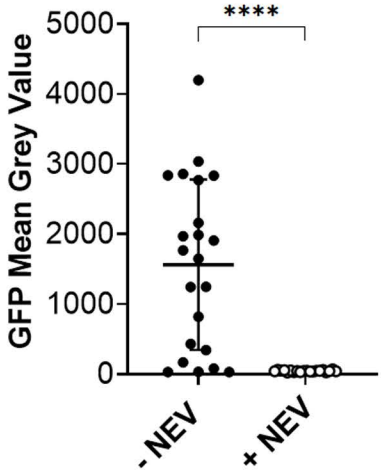
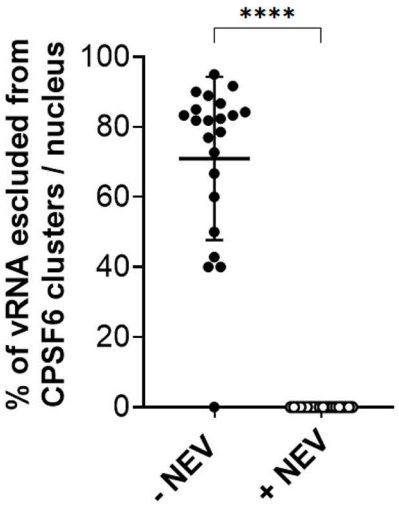

Hoechst
vDNA
vRNA
CPSF6
Merge

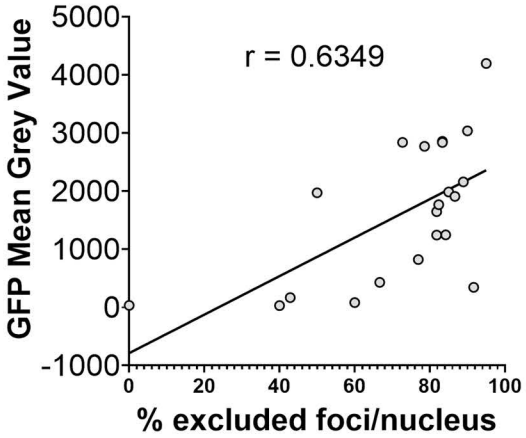
Figure 6

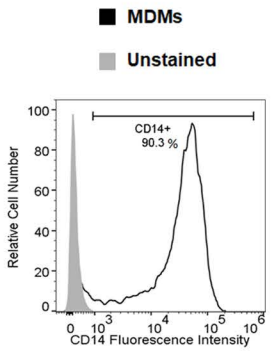
MDM 3 days p.i.











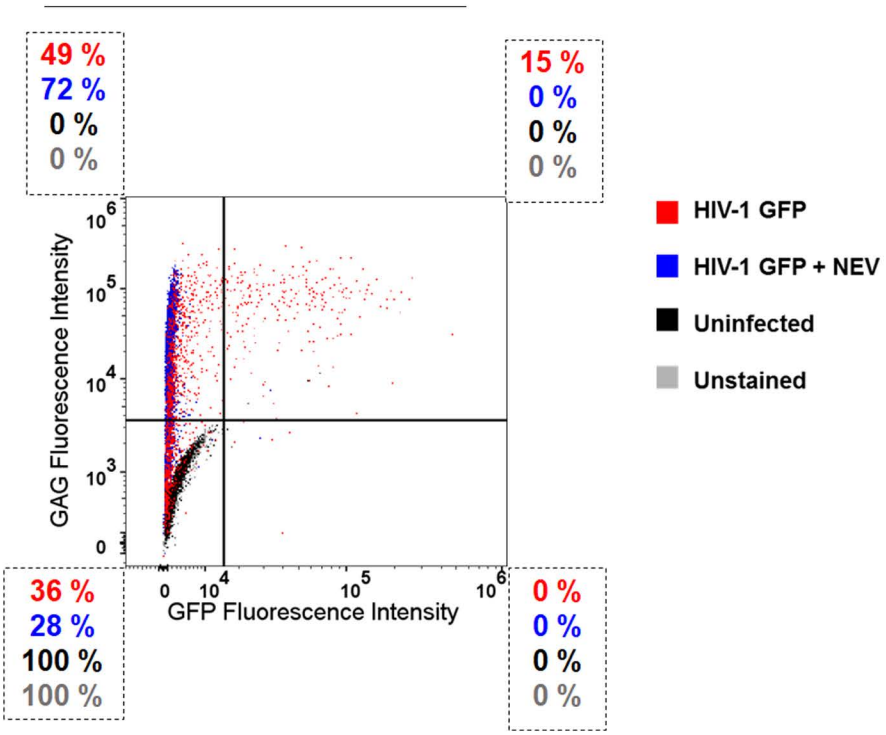
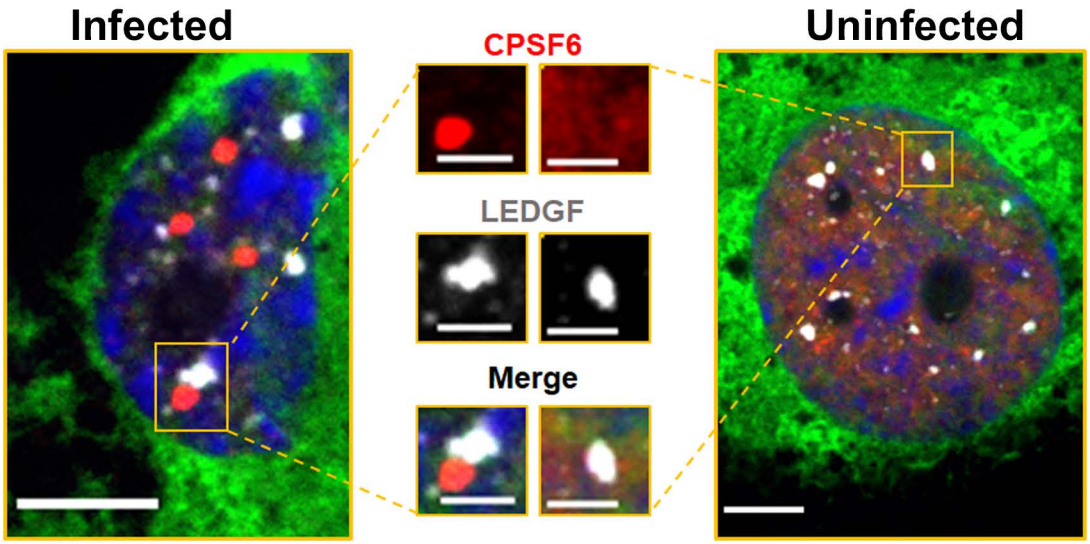
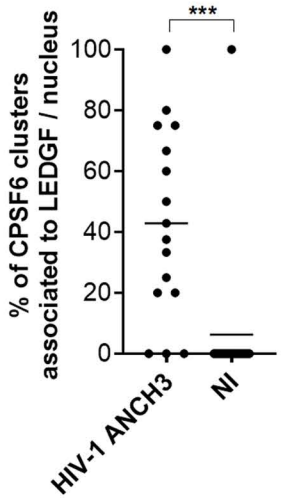
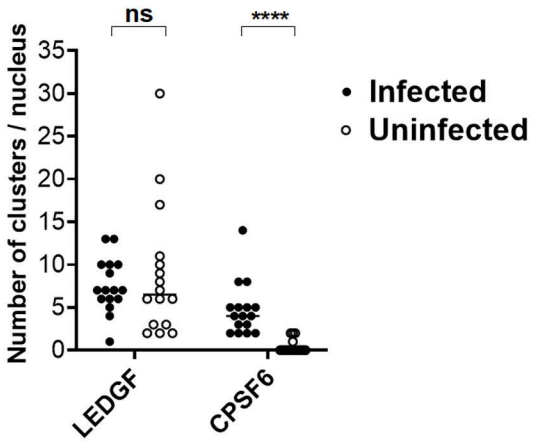


Figure 7

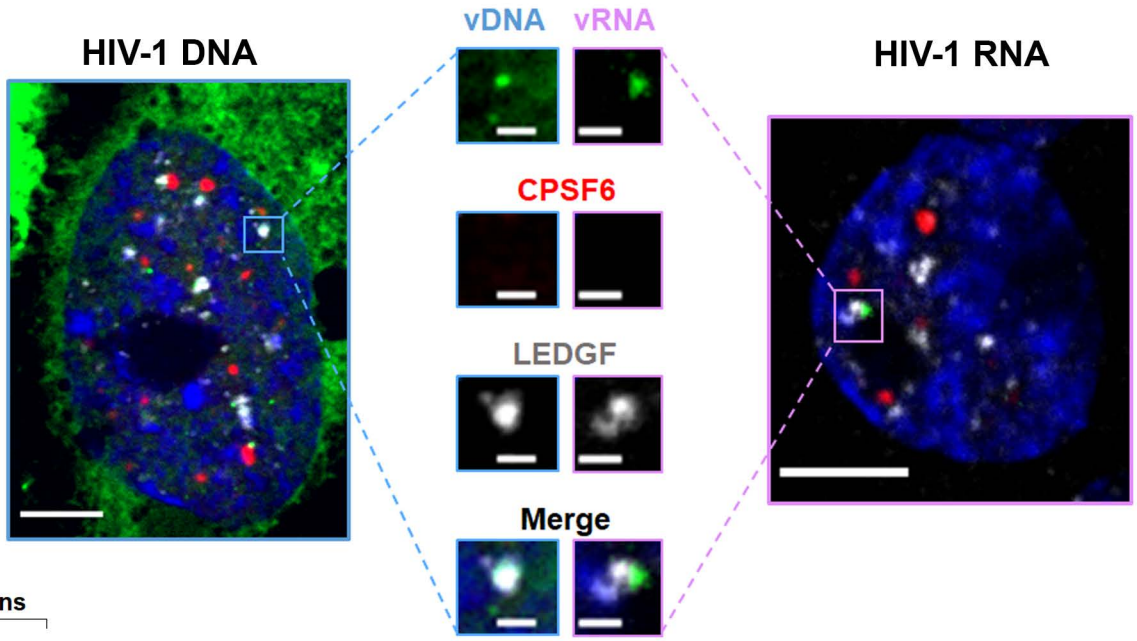
A

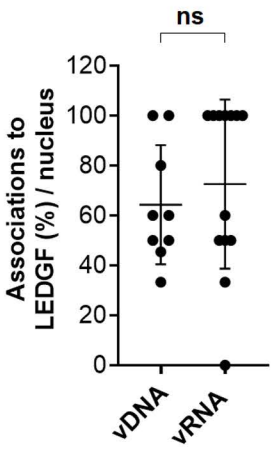

Infected
CPSF6
LEDGF
Merge
Uninfected

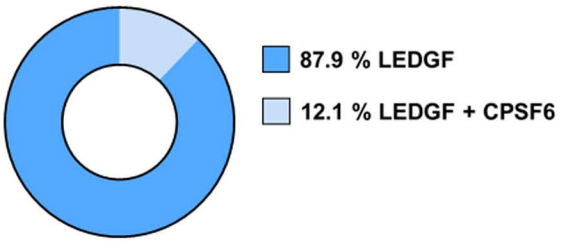

***
% of CPSF6 clusters associated to LEDGF / nucleus
100
80
60
40
20
0
HIV-1 ANCH3
NI


ns
****
Number of clusters / nucleus
35
30
25
20
15
10
5
0
LEDGF
CPSF6
Infected
Uninfected

B


HIV-1 DNA
vDNA
vRNA
CPSF6
LEDGF
Merge
HIV-1 RNA


ns
Associations to LEDGF (%) / nucleus
120
100
80
60
40
20
0
vDNA
vRNA


vDNA
87.9 % LEDGF
12.1 % LEDGF + CPSF6

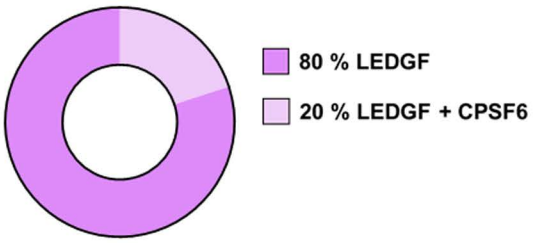

vRNA
80 % LEDGF
20 % LEDGF + CPSF6

Figure 8

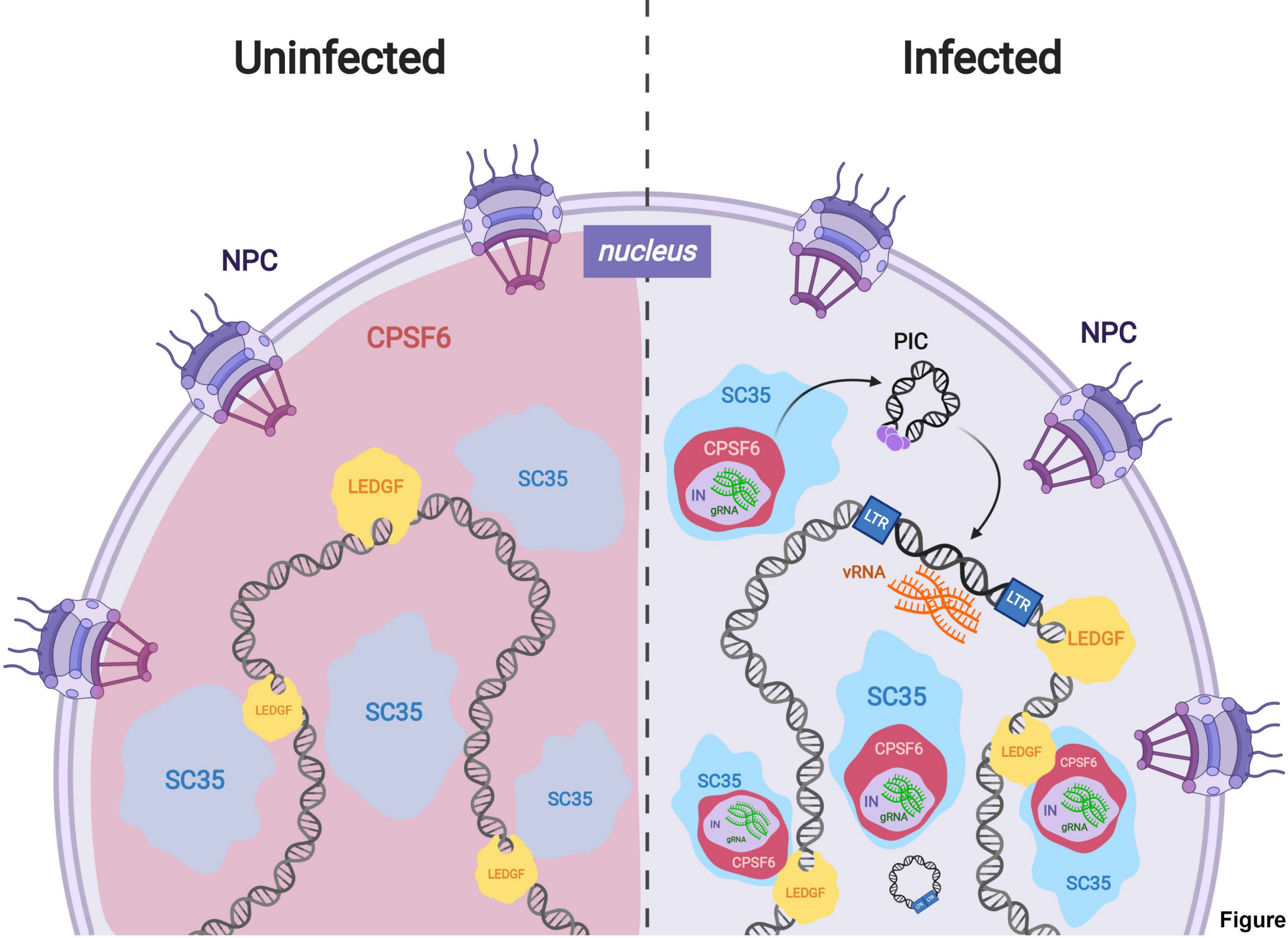

Uninfected
Infected
nucleus
NPC
CPSF6
LEDGF
SC35
LEDGF
SC35
SC35
SC35
LEDGF
PIC
NPC
SC35
CPSF6
IN
gRNA
LTR
vRNA
LTR
LEDGF
SC35
CPSF6
IN
gRNA
SC35
IN
gRNA
CPSF6
LEDGF
LEDGF
CPSF6
IN
gRNA
SC35

Figure 9