Clustering and reverse transcription of HIV-1 genomes in nuclear niches of macrophages

Elena Rensen[1,2], Florian Mueller[1*], Viviana Scoca[2], Jyotsana Parmar[1], Philippe Souque[2], Christophe Zimmer[1*#], Francesca Di Nunzio[2*#]

[1]Imaging and Modeling Unit, Institut Pasteur, UMR 3691 CNRS, C3BI USR 3756 IP CNRS, Paris, France

[2]Molecular Virology and Vaccinology, Institut Pasteur, Paris, France

*corresponding authors

#co-senior authors

Summary

In order to replicate, the Human Immunodeficiency Virus (HIV-1) reverse transcribes its RNA genome into DNA, which subsequently integrates into host cell chromosomes. These two key events of the viral life cycle are commonly viewed as separate not only in time but also in cellular space, since reverse transcription (RT) is thought to be completed in the cytoplasm before nuclear import and integration. However, the spatiotemporal organization of the early replication cycle in macrophages, natural non-dividing target cells that constitute reservoirs of HIV-1 and an obstacle to curing AIDS, remains unclear. Here, we demonstrate that infected macrophages display large nuclear foci of viral DNA and viral RNA, in which multiple genomes cluster together. These clusters form in the absence of chromosomal integration, sequester the paraspeckle protein CPSF6 and localize to nuclear speckles. Strikingly, we show that viral RNA clusters consist mostly of genomic, incoming RNA, both in cells where RT is pharmacologically suppressed and in untreated cells. We demonstrate that, after temporary inhibition of RT, RT can resume in the nucleus and lead to vDNA accumulation in these clusters. We further show that nuclear RT can result in transcription competent viral DNA. These findings change our understanding of the early HIV-1 replication cycle, and may have implications for understanding HIV-1 persistence.

Productive infection by the Human Immunodeficiency Virus 1 (HIV-1), the causative agent of AIDS, requires reverse transcription (RT) of the viral RNA (vRNA) genome into double-stranded viral DNA (vDNA) and subsequent integration of the vDNA into host cell chromosomes. Studies in immortalized cell lines such as HeLa and activated CD4+ T cells have established a spatiotemporal sequence of events where: (i) vDNA is synthesized by RT in the cytoplasm, with concomitant degradation of the template vRNA, (ii) the vDNA genome translocates into the nucleus, (iii) the integrase enzyme (IN) inserts the vDNA into the genome, (iv) the integrated vDNA genome undergoes transcription that leads to viral progeny (Freed, 2001; Hu & Hughes, 2012; Hulme *et al*, 2011; Dismuke & Aiken, 2006) . How the HIV-1 replication cycle proceeds in other cell types remains comparatively underexplored. Indeed, different cell types can exhibit widely divergent responses to viral attacks, owing to e.g. different restriction factors or immune cell defense mechanisms (Stremlau *et al*, 2004; Goujon *et al*, 2013; Lahaye *et al*, 2018; Rasaiyaah *et al*, 2013). Macrophages are terminally differentiated, non-dividing cells derived from blood monocytes, which play a critical role in the innate and adaptive immune response (Koppensteiner *et al*, 2012; Gordon & Taylor, 2005; Ganor *et al*, 2019). Along with activated CD4+ T cells, macrophages are also natural target cells for HIV-1 and accumulating evidence points to a critical role of these cells in viral persistence, which remains a major roadblock to eradicating HIV (Ganor *et al*, 2019; Honeycutt *et al*, 2017). Despite this importance, the early steps of HIV-1 infection in macrophages remain elusive. Here, we use imaging approaches to visualize and quantify the cellular localizations of vDNA and vRNA in infected macrophages and study their link with RT. Our data reveal that genomic vRNA forms nuclear clusters that associate with nuclear speckle factors, and provide surprising evidence that these structures can harbor a nuclear RT activity.

HIV-1 genomes form large nuclear foci in infected macrophages

To study HIV infection in macrophages, we used ThP1 cells, a human monocytic cell line, and differentiated them into macrophage-like cells by stimulation with phorbol esters (Schwende *et al*, 1996). We infected these cells with VSV-G pseudotyped HIV-1 carrying the HIV-2 accessory protein Vpx (unless stated otherwise), which overcomes the natural resistance of macrophages to viral infection by counteracting the host restriction factor SAMHD1 (Laguette *et al*, 2011). To enable fluorescent labeling of the virus, we used a virus carrying the endogenous viral integrase (IN) gene fused to an HA-tag (Petit *et al*, 2000, 1999). The tagged virus is similarly infectious as the untagged virus (Petit *et al*, 1999; Blanco-Rodriguez *et al*, 2020). We analyzed reverse transcription (RT) with qPCR by measuring viral DNA synthesis and the presence of nuclear viral DNA (vDNA) forms including 2LTRs at different times post infection (p.i.). In parallel, we also measured the number of integrated proviruses (Lelek *et al*, 2015), revealing a low level of integration in these cells at 24 h p.i. (**Fig. S1**). These assays indicated that RT peaks at ~24 h p.i., and that the formation of episomal nuclear forms (2LTRs), peaks at ~30 h p.i., confirming that these early steps of HIV-1 infection in ThP1 macrophage-like cells are delayed relative to HeLa cells (Arfi *et al*, 2008).

To visualize the reverse transcribed viral DNA genome, we infected ThP1 cells in the presence of the nucleotide analog EdU for 24 hours (Peng *et al*, 2014; Stultz *et al*, 2017) with a multiplicity of infection (MOI) of 50 (unless otherwise stated), as measured by qPCR on 293T cells. Fluorescent visualization of EdU was performed at 48 h p.i. by click chemistry, unless stated otherwise. Because the vast majority of ThP1 cells were terminally differentiated, EdU did not incorporate into host cell chromosomes. A

minority of cells (~5 %) exhibited bright EdU signal throughout the nucleus, clearly indicating that these cells failed to differentiate and continued to replicate their DNA (**Fig. S2A**). We excluded these cells from further analysis throughout this study and thus only considered the remaining, terminally differentiated cells. Our images revealed strikingly large and bright EdU foci in the nuclei of infected cells, whereas uninfected control cells displayed only a very weak background signal (**Fig. 1A, Movie S1**, **Fig. S2B-C**). While the nuclear envelope of some cells displayed invaginations, dual color imaging of EdU with immunostained lamin in infected cells confirmed that these foci are located within the nuclear lumen (**Fig. 1C**, **Fig. S3**). Quantifications indicated that ~73% of cells contained one or more foci (**Fig. 1J**) and that foci had a median size (as measured by the full width at half maximum, FWHM, of intensity profiles) of ~600 nm (interquartile range ~175 nm; n=40 foci) (**Fig. 1B,F**). In addition to these large foci, some infected cells also showed discrete nuclear punctae of much lower brightness, with a size (FWHM) of ~370 nm, close to the microscope's theoretical resolution of ~300 nm and hence consistent with particles of similar or smaller size (**Fig. 1A,F**). Triple color imaging of EdU together with immunolabeled capsid (CA) and integrase (IN) exhibited clear colocalization, thereby simultaneously confirming the viral nature of EdU-labelled DNA, and showing that the EdU foci (hereafter referred to as 'vDNA foci') are enriched in these viral proteins (**Fig. 1D, Fig. S4**). Importantly, although our experiments used viral particles incorporating Vpx to increase HIV-1 infection efficiency, we observed that in absence of Vpx, infected ThP1 cells also form nuclear vDNA foci, ruling out the possibility that these foci are an artifact due to the presence of Vpx (**Fig. S5**).

We next asked if the observed vDNA foci might correspond to individual viral genomes by further analyzing their size and dependence on the MOI. First, we considered the

possible size of a single ~10 Kb long piece of DNA, the approximate length of the HIV-1 genome. We used polymer simulations that model the chromatinized vDNA as a linear or circular chain of nucleosomes (Arbona *et al*, 2017; Geis & Goff, 2019) (**Fig. 1E, Fig. S6**). Our simulations predict a distribution of apparent sizes for a single 10 Kb long HIV-1 genome that is close to the microscope's spatial resolution of ~300 nm, hence is much smaller than the observed sizes of nuclear vDNA foci (**Fig. 1F**) ($p<10^{-10}$). Thus, nuclear vDNA foci are larger than expected for single genomes, suggesting that they may comprise multiple genomes. Second, we analyzed the intensity of vDNA foci for MOIs of ~10, ~50 and ~100 (**Fig. 1G,H**). Note that EdU labeling does not involve signal amplification, enabling a quantitative interpretation of the measured intensities. If individual foci correspond to individual genomes, their number is expected to increase proportionally with the MOI, but their intensity should not depend on MOI. In our data, the peak intensity of vDNA foci significantly increased from MOI 10 to MOI 50, and again from MOI 50 to MOI 100 ($p<10^{-4}$) (**Fig. 1H**). Therefore, the MOI dependence of the vDNA intensity also argues for the coexistence of multiple genomes within these foci.

Nuclear vDNA/vRNA foci contain multiple HIV-1 genomes

In order to analyze the potential association of vDNA foci with viral RNA (vRNA), we used RNA-FISH probes directed against the POL gene of HIV-1. We verified the specificity of RNA-FISH using uninfected cells, and ruled out the possibility that RNA-FISH probes bind vDNA using infected cells treated with RNase (**Fig. S7A-B**). The dual color EdU/RNA-FISH images showed bright vRNA foci that displayed high and significant colocalization with the vDNA foci (64% of EdU foci contained vRNA, and

30% of vRNA spots contained vDNA; $p<0.01$), indicating that viral RNA and DNA occupy the same nuclear space (**Fig. 2A,B; Fig. S8**). The vDNA and vRNA intensities in colocalizing spots exhibited a strong positive correlation (Spearman's $r=0.57$, $p<10^{-30}$) (**Fig. 2C**).

In order to directly test for the presence of multiple viral genomes in nuclear foci, we co-infected cells with two HIV strains containing different reporter genes, GFP and luciferase (LUC). We then performed dual-color RNA-FISH with two different sets of probes directed against GFP and LUC (**Fig. 2D**). We confirmed the specificity of our RNA-FISH probes using control cells infected with only one of the two reporter viruses (**Fig. S9**). In co-infected cells, RNA-FISH revealed strong colocalization of LUC and GFP RNA (45% of LUC RNA foci contained GFP RNA, and 35% of GFP RNA foci contained LUC RNA; $p<0.01$) (**Fig. 2E**). In colocalizing foci, intensities of GFP and LUC RNA correlated positively (Spearman's $r=0.5$, $p<10^{-9}$). Taken together with our observation that vRNA foci occupy the same nuclear regions as vDNA foci (**Fig. 2A,B**), these data firmly establish that nuclear foci are clusters consisting of multiple HIV-1 genomes.

Viral genome clusters are associated with nuclear body factors

Inspection of our dual EdU/DAPI images indicates that vDNA clusters are located in nuclear regions with lower densities of host cell DNA (**Fig. S10**). This opens the possibility of an association of HIV genomes with nuclear bodies (Spector, 2001; Matera, 1999), membrane-less compartments located in the interchromosomal space. We therefore examined the possible association of viral clusters with host proteins, starting with the cleavage and polyadenylation-specific factor subunit 6 (CPSF6).

CPSF6 is known to interact with CA and has been implicated in the regulation of different steps of the viral life cycle from HIV-1 nuclear import to integration site distribution (Price *et al*, 2012; Achuthan *et al*, 2018, 6; Buffone *et al*, 2018; Lee *et al*, 2010; Burdick *et al*, 2020). In uninfected cells, CPSF6 displayed a diffuse nucleoplasmic signal; in infected cells, however, CPSF6 accumulated in a small number of nuclear foci, which colocalized with vDNA, in agreement with a recent report (Bejarano *et al*, 2019) (median Pearson's r=0.30; Costes p<0.01 for 16 out of 19 regions of interest) (**Fig. 3A,D; Fig. S11**). CPSF6 is known to associate with paraspeckles, nuclear bodies often found in the vicinity of nucleoli and speckles (Fox *et al*, 2002; Naganuma & Hirose, 2013). To test whether viral clusters associate with paraspeckles, we imaged the non-coding RNA NEAT1 (Nuclear Enriched Abundant Transcript 1/Human nuclear paraspeckle assembly transcript 1), an essential architectural component of paraspeckles (Naganuma & Hirose, 2013; Zhang *et al*, 2013, 1; Yamazaki *et al*, 2018, 1). In uninfected cells, NEAT1 was enriched in a small number of nuclear clusters, as expected for paraspeckles (**Fig. 3B; Fig. S12A**). In infected cells, NEAT1 displayed a similar localization pattern, and NEAT1 foci appeared in close vicinity to the vDNA clusters, but did not overlap with them and appeared to be excluded (median Pearson's r=-0.19; Costes p<0.01 for 19/19 regions of interest) (**Fig. 3B,D; Fig. S12B**). The core paraspeckle protein NONO (p54nrb), recently found to be essential for triggering an immune response to HIV in dendritic cells through interaction with CA (Lahaye *et al*, 2018), was also excluded from vDNA (**Fig. S13**). Thus, viral clusters do not localize to paraspeckles, but upon infection, the paraspeckle factor CPSF6 relocates to viral clusters.

Paraspeckles are often found in proximity to speckles, nuclear bodies enriched in pre-mRNA splicing factors (Spector & Lamond, 2011). Therefore, our data raised the

possibility that viral clusters are associated with speckles. To investigate this, we imaged the non-small nuclear ribonucleoprotein particle factor SC35, a well-studied splicing regulator and bona fide marker of speckles (Spector & Lamond, 2011). Images of SC35 in uninfected cells showed a large number (typically ~10-15) large nuclear bodies (typical size ~1-2 µm), as expected for speckles (Lamond & Spector, 2003) (**Fig. 3C**). In HIV-1 infected cells, SC35 displayed a similar pattern. Interestingly, nuclear vDNA clusters colocalized with a subset of these SC35 positive bodies (median Pearson's r=0.14; Costes p<0.01 in 16/17 regions) (**Fig. 3C,D; Fig. S14**). Thus, our data suggests that HIV-1 genome clusters enriched in the paraspeckle protein CPSF6 associate with the splicing and speckle factor SC35.

Viral integration is not required for viral DNA cluster formation

To determine whether the formation of vDNA clusters requires integration of the viral genome into host chromosomes, we infected cells with an integration-deficient mutant virus (D116A). Interestingly, we again observed the presence of large vDNA clusters which are also associated with vRNA (**Fig. 4A, Fig. S15**). As for the virus containing the functional integrase (**Fig. 2A-C**), the colocalization of vDNA and vRNA was highly significant and intensities in colocalizing clusters correlated significantly (Spearman's r=0.51, $p=1.6x10^{-5}$) (**Fig. 4B,C**). Additionally, when we treated cells at the time of infection with 10 µM of Raltegravir (RAL), an inhibitor of HIV integration, we also observed large and bright vDNA clusters associated with vRNA (**Fig. S16B**). These experiments demonstrate that the formation of vDNA/vRNA clusters does not depend on integration of the viral genome into host chromosomes.

Viral RNA clusters contain mostly genomic RNA

The dual color EdU/RNA-FISH images reported above clearly indicate that vDNA clusters colocalize with vRNA (**Fig. 2A-C**). However, our RNA-FISH probes against the HIV-1 POL gene cannot distinguish between incoming (genomic) vRNA and newly transcribed (messenger) vRNA. We therefore sought to verify the nature of the observed vRNA by treating cells with 10 µM of Nevirapine (NVP), a potent RT inhibitor. The efficacy of the drug was demonstrated by the absence of detectable vDNA in infected nuclei; to our surprise, however, we still observed bright vRNA clusters in the nucleus, whether we exposed the cells to NVP for 24 h, 48 h or 72 h (**Fig. 4D,E; Fig. S1, S17A, S23**). Because HIV-1 transcripts cannot be produced in the absence of vDNA, we conclude that nuclear vRNA clusters at these time points contain mostly incoming, genomic vRNA, in NVP treated cells.

We next asked if genomic vRNA clusters also exist in absence of RT inhibition. Intensity distributions of vRNA clusters in NVP treated cells at 48 h p.i. were similar to those in untreated cells (p=0.97; **Fig. S17B**), strongly suggesting that the vRNA clusters in the untreated cells are also mostly composed of genomic vRNA. In order to verify this directly, we turned to a different labeling technique that, unlike RNA-FISH, highlights the genomic, but not the transcribed, vRNA. Specifically, we used 5-ethynyl uridine (EU), another nucleoside analog that incorporates into nascent RNA during transcription and can be fluorescently detected by click-chemistry, as for EdU (Jao & Salic, 2008; Xu *et al*, 2013). As demonstrated previously (Xu *et al*, 2013), we used EU to label HIV RNA in 293T producer cells, then we infected NVP treated or untreated Thp1 cells with this virus (MOI 50), and performed fixation at 24 h p.i., followed by click chemistry labeling and fluorescence imaging, in combination with RNA-FISH (**Fig. 4F, Fig. S18**). Despite the presence of a significant background signal, we observed EU

spots in the nucleus of both NVP treated and untreated cells that clearly coincided with vRNA clusters (**Fig. 4G, Fig. S18**). These data therefore strongly support the presence of genomic vRNA in the nucleus, regardless of whether cells have been pharmacologically treated.

Nuclear viral RNA clusters can undergo reverse transcription

The unexpected presence of genomic vRNA clusters in the nucleus (**Fig. 4D,E,F,G**) raises the intriguing question as to whether they can serve as templates for RT. To test this possibility, we took advantage of the reversibility of RT inhibition by NVP. We reasoned that if RT occurs locally in vRNA clusters, allowing RT to resume after temporary RT inhibition would lead to the detection of newly synthesized vDNA within these nuclear structures.

We first exposed cells to NVP starting from the time of infection during 24, 48 or 72 hours. Dual color images obtained at all three time points showed an absence of nuclear vDNA signal, as expected, while the vRNA clusters remained clearly visible, as above (**Fig. S17A, Fig. 4D**). Next, we exposed cells to NVP starting from the time of infection for 48 h or 72 h, then removed NVP by wash-out and imaged cells 24 h later. Strikingly, in both experiments, we observed clear vDNA clusters that colocalized with the vRNA clusters, and displayed positive correlation of their intensities ($p<10^{-4}$) (**Fig. 5A,B,E,F; Fig. S19**), as in untreated cells (**Fig. 2A-C**). This appearance of vDNA at vRNA clusters can be explained by a local RT activity within the nuclear vRNA clusters. Alternatively, the observed vDNA clusters might reflect nuclear import of vDNA synthesized exclusively in the cytoplasm, between the time points of NVP washout and fixation. In the first case (nuclear RT), we expect that clusters containing

more vRNA also exhibit higher amounts of vDNA. This is indeed the case, as evidenced by the positive correlations between vRNA and vDNA intensities in colocalizing clusters in all wash-out experiments (**Fig. 5E**).

To further assess both possibilities, we next quantified the amount of vRNA in the cytoplasmic and nuclear compartments using an automated analysis based on FISH-quant and ImJoy (Mueller *et al*, 2013; Ouyang *et al*, 2019). These analyses indicated that the nuclear vRNA pool exceeded the cytoplasmic pool by at least ~7-fold after 72 h of NVP treatment (**Fig. S17C**). Thus, the amount of vRNA available for cytoplasmic RT and subsequent nuclear import during the 24 h after wash-out is much smaller than the amount of vRNA available for nuclear RT. Therefore, the nuclear vDNA detected in our NVP wash-out experiments more likely arises from local RT in nuclear clusters rather than import of vDNA synthesized in the cytoplasm.

To directly rule out a contribution of cytoplasmically synthesized vDNA to nuclear vDNA clusters after NVP removal, we aimed to pharmacologically block nuclear import using the molecule PF-3450074 (PF74). At low concentrations (1.25-2.5 µM), PF74 impedes nuclear import of HIV-1 without abolishing RT (Blanco-Rodriguez *et al*, 2020; Bejarano *et al*, 2019; Francis & Melikyan, 2018; Balasubramaniam *et al*, 2019; Blair *et al*, 2010). Indeed, when we exposed cells to 1.5 µM of PF74 for 72 h after infection, qPCR indicated that RT was roughly halved, while 2LTR formation was virtually eliminated (**Fig. 5G**). We observed no discernable nuclear vRNA in PF74 treated cells at 48 h or 72 h p.i. and a higher cytoplasmic vRNA pool as compared to NVP treated cells at the same time point, confirming inhibition of nuclear import (**Fig. 5C**, **Fig. S20A,B; Fig. S21**). Finally, we treated cells with NVP for 24 h p.i., then washed NVP out and immediately exposed the cells to 1.5 µM PF74 for another 24 h. Despite the block of nuclear import, we again observed clear vDNA clusters colocalizing with

vRNA clusters in the nucleus, with positively correlating intensities (r=0.41-0.49; $p<10^{-6}$) (**Fig. 5H; Fig. S20C**). We made the same observation when exposing cells to NVP for 48 h, applying PF74 at 36 h, and washing out NVP (while keeping the cells exposed to 1.5 µM of PF74) at 48 h p.i (**Fig. S22**). We additionally performed qPCR analyses of circular 2LTRs, an exclusively nuclear form of vDNA that can persist in the nucleus for a long time (**Fig. S1**). Interestingly, we were able to amplify 2LTRs in samples after NVP wash-out, whether or not followed by exposure to PF74, albeit with a 3-7 fold reduction compared to untreated cells; this reduction was ~10-fold when applying PF74 at 36 h p.i. and washing out NVP at 48 h p.i. (**Fig. S23**). Nevertheless, the detection of 2LTRs in these experiments corroborates the synthesis of complete vDNA by RT in the nucleus. We also analyzed the presence of proviral DNA using ALU-PCR (Lelek *et al*, 2015; Di Nunzio *et al*, 2013). While we could amplify these integrated vDNA forms in untreated cells or in simple NVP wash-out experiments, we failed to do so in the experiments involving PF74 exposure (**Fig. S23**). This suggests that the majority of the vDNA detected consist of unintegrated genomes, in agreement with our above findings using an integration deficient virus (**Fig. 4A-C**) or when inhibiting integration pharmacologically (**Fig. S15**). Notwithstanding, our data clearly argue against the possibility that the nuclear vDNA clusters arise exclusively from import of cytoplasmically synthesized vDNA and therefore constitute compelling evidence for a nuclear RT activity within vRNA clusters after release of RT inhibition.

Finally, we asked whether the viral DNA synthesized in the nucleus after pharmacological block followed by release of RT is competent for transcription. To address this, we infected cells with the virus carrying the GFP reporter gene used above (**Fig. 2D, S24**). After imaging large fields of view, we quantified the proportion of GFP positive cells under different conditions, using a conservative (high) intensity

threshold. Cells infected with MOI 20 and fixed at 72 h p.i. displayed 1.8% of GFP positive cells (n=73/4071), vs. none for uninfected cells (n=0/1583) (**Fig. S24A,D,F**). When we inhibited RT by NVP treatment for 72 h p.i., allowed RT to resume by washing out NVP and fixed cells for imaging 72 h later, 1.0% (n=43/4347) of cells were found GFP positive (**Fig. S24B,F**). When we repeated this experiment while exposing cells to 1.5 µM of PF74 concomitantly with NVP wash-out to block nuclear import of cytoplasmic vDNA, we still found 0.4% (n=16/4356) of cells displaying a clear GFP signal (**Fig. S24BC,F**). Thus, our data indicate that nuclear RT activity, at least in conditions where RT was initially suppressed, can lead to the production of transcription competent viral DNA.

Discussion

Our study uses imaging of viral DNA and RNA to shed new light on the early replication cycle of HIV-1 in macrophages. We demonstrated that vDNA and vRNA genomes cluster together in nuclear niches associated with speckle factors and that these clusters can form in absence of viral integration into the host genome. We further showed that genomic vRNA clusters can form in absence of vDNA after pharmacological inhibition of RT, but, importantly, also in untreated cells. Our observation of genomic RNA in nuclei agrees with previous studies showing that RT is dispensable for nuclear import (Bejarano *et al*, 2019; Burdick *et al*, 2013, 2017, 2020). However, the potential role of genomic vRNA in the nucleus remained unknown. We note that several recent studies have evoked the possibility of RT in the nucleus (Bejarano *et al*, 2019; Burdick *et al*, 2017, 2020). However, our experiments combining reversible RT inhibition with direct visualization of the synthesized vDNA provide the

first clear demonstration that these genomic RNA clusters can serve as templates for RT in the nucleus. We additionally showed that the viral DNA resulting from this nuclear RT activity can serve as a template for transcription.

We emphasize that our data do not implicate that RT occurs exclusively, or even majoritarily, in the nucleus. Our results allow the possibility that RT is initiated in the cytoplasm and prolonged in the nuclear compartment, or that some viral RNA genomes are entirely reverse transcribed in the cytoplasm. We also acknowledge that our demonstration of nuclear RT was achieved after temporary pharmacological inhibition of RT. Nevertheless, our observation of nuclear genomic vRNA clusters in absence of drug treatment opens the possibility that RT also occurs in the nuclei of untreated macrophage-like cells. At any rate, our clear evidence for nuclear RT in human macrophages-like cells under the aforementioned conditions stands in stark contrast with the classical picture of the HIV replication cycle, according to which RT is entirely restricted to the cytoplasmic compartment. An intriguing speculation is that these vRNA clusters may act as nuclear microreactors that concentrate reverse transcriptase enzymes to enable efficient vDNA synthesis in macrophage-like cells, much as intact capsid cores are believed to concentrate these enzymes in the cytoplasm of HeLa or CD4+ cells (Warrilow *et al*, 2007).

Previous *in vitro* experiments have suggested a potential link between HIV-1 and nuclear speckles (Bell *et al*, 2001; Pendergrast *et al*, 2002) and earlier studies have reported that HIV-1 RNA colocalizes with the speckle factor SC35 (Cardinale *et al*, 2007; Bøe *et al*, 1998). Here, we directly show that HIV-1 RNA/DNA genome clusters localize to nuclear niches enriched in SC35. While it has been proposed that the nuclear pool of unspliced HIV-1 RNAs may be stored in nuclear paraspeckles (Zhang *et al*, 2013, 1), our data instead indicate that viral clusters sequester the host cell factor

CPSF6 away from its canonical localization in paraspeckles (Cardinale *et al*, 2007; Bøe *et al*, 1998; Burdick *et al*, 2020) and relocate this protein in speckles.

Other studies highlighted the ability of other viruses to alter the cytoplasmic or nuclear organization and form microenvironments that locally concentrate viral and host cell factors required for the synthesis of viral progeny and/or to protect the virus from cellular defense mechanisms (Schmid *et al*, 2014; McSwiggen *et al*, 2019; Heinz *et al*, 2018). One intriguing, albeit speculative, possibility is that the crowded microcompartments formed by cellular RNAs in nuclear bodies shroud the viral DNA and protect it from mediators of the innate immune response, such as the cyclic GMP-AMP synthase (cGAS), a viral DNA sensor that associates with HIV DNA in nuclei of infected macrophages (Lahaye *et al*, 2018; Sumner *et al*, 2019).

Our observation of vRNA/vDNA clusters in cells infected by an integration-defective virus or upon inhibition of integration further suggests that most of the vDNA in nuclear clusters may be unintegrated. This dovetails with the common observation of large amounts of unintegrated viral DNA in the nucleus, including in viremic HIV-1 patients after combination antiretroviral therapy (cART) (Bell *et al*, 2001; Hamid *et al*, 2017). These latent extrachromosomal genomes persist in human macrophages for a long time and can act as viral reservoirs, which constitute an important obstacle to successful treatment (Hamid *et al*, 2017). Furthermore, it has been shown that, in presence of integrated vDNA, unintegrated vDNA can boost viral progeny in macrophages (Gelderblom *et al*, 2008). Thus, the unintegrated vDNA observed in our study may be relevant to understanding HIV reactivation in patients.

More research is now in order to further explore the functional role and formation mechanisms of these vRNA/vDNA clusters. Our discovery of RT-competent clusters

of viral RNA and DNA opens new perspectives for understanding and hence combating HIV-1 replication in natural target cells.

ACKNOWLEDGEMENTS

We thank Mickaël Lelek and Andrey Aristov for help with microscopy and Benoît Lelandais for help with image analysis. We thank Marie-Anne Welti, Edouard Bertrand, Xavier Darzacq, Olivier Schwartz, Arnaud Echard and Michaela Mueller for useful discussions and/or comments on the manuscript. We thank Fabrizio Mammano for kindly providing the modified IN-HA HIV-1 virus and Stella Frabetti for technical support. We also thank the UtechS Photonic BioImaging, C2RT, Institut Pasteur, which is supported in part by the French National Research Research Agency (France BioImaging; ANR-10–INSB–04; Investments for the Future). We thank the NIH AIDS Reagents program for reagents. This work was funded by Institut Pasteur, Fondation pour la Recherche Médicale en France (DEQ 20150331762), Institut Carnot Pasteur MS, Sidaction (VIH20170718001) and ANRS (grants N. ECTZ4469-ECTZ74440).

REFERENCES

Achuthan V, Perreira JM, Sowd GA, Puray-Chavez M, McDougall WM, Paulucci-Holthauzen A, Wu X, Fadel HJ, Poeschla EM, Multani AS, Hughes SH, Sarafianos SG, Brass AL & Engelman AN (2018) Capsid-CPSF6 Interaction Licenses Nuclear HIV-1 Trafficking to Sites of Viral DNA Integration. *Cell Host Microbe* **24:** 392-404.e8

Arbona J-M, Herbert S, Fabre E & Zimmer C (2017) Inferring the physical properties of yeast chromatin through Bayesian analysis of whole nucleus simulations. *Genome Biol.* **18:** 81

Arfi V, Rivière L, Jarrosson-Wuillème L, Goujon C, Rigal D, Darlix J-L & Cimarelli A (2008) Characterization of the Early Steps of Infection of Primary Blood Monocytes by Human Immunodeficiency Virus Type 1. *J. Virol.* **82:** 6557–6565

Balasubramaniam M, Zhou J, Addai A, Martinez P, Pandhare J, Aiken C & Dash C (2019) PF74 Inhibits HIV-1 Integration by Altering the Composition of the Preintegration Complex. *J. Virol.* **93:** Available at: https://jvi.asm.org/content/93/6/e01741-18 [Accessed November 10, 2019]

Bejarano DA, Peng K, Laketa V, Börner K, Jost KL, Lucic B, Glass B, Lusic M, Müller B & Kräusslich H-G (2019) HIV-1 nuclear import in macrophages is regulated by CPSF6-capsid interactions at the nuclear pore complex. *eLife* **8:** e41800

Bell P, Montaner LJ & Maul GG (2001) Accumulation and Intranuclear Distribution of Unintegrated Human Immunodeficiency Virus Type 1 DNA. *J. Virol.* **75:** 7683–7691

Blair WS, Pickford C, Irving SL, Brown DG, Anderson M, Bazin R, Cao J, Ciaramella G, Isaacson J, Jackson L, Hunt R, Kjerrstrom A, Nieman JA, Patick AK, Perros M, Scott AD, Whitby K, Wu H & Butler SL (2010) HIV Capsid is a Tractable Target for Small Molecule Therapeutic Intervention. *PLOS Pathog.* **6:** e1001220

Blanco-Rodriguez G, Gazi A, Monel B, Frabetti S, Scoca V, Mueller F, Schwartz O, Krijnse-Locker J, Charneau P & Nunzio FD (2020) Remodeling of the core leads HIV-1 pre-integration complex in the nucleus of human lymphocytes. *J. Virol.* Available at: https://jvi.asm.org/content/early/2020/03/27/JVI.00135-20 [Accessed April 7, 2020]

Bøe S-O, Bjørndal B, Røsok B, Szilvay AM & Kalland K-H (1998) Subcellular Localization of Human Immunodeficiency Virus Type 1 RNAs, Rev, and the Splicing Factor SC-35. *Virology* **244:** 473–482

Buffone C, Martinez-Lopez A, Fricke T, Opp S, Severgnini M, Cifola I, Petiti L, Frabetti S, Skorupka K, Zadrozny KK, Ganser-Pornillos BK, Pornillos O, Nunzio FD & Diaz-Griffero F (2018) Nup153 Unlocks the Nuclear Pore Complex for HIV-1 Nuclear Translocation in Nondividing Cells. *J. Virol.* **92:** e00648-18

Burdick RC, Delviks-Frankenberry KA, Chen J, Janaka SK, Sastri J, Hu W-S & Pathak VK (2017) Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. *PLOS Pathog.* **13:** e1006570

Burdick RC, Hu W-S & Pathak VK (2013) Nuclear import of APOBEC3F-labeled HIV-1 preintegration complexes. *Proc. Natl. Acad. Sci. U. S. A.***:** 1315996110-

Burdick RC, Li C, Munshi M, Rawson JMO, Nagashima K, Hu W-S & Pathak VK (2020) HIV-1 uncoats in the nucleus near sites of integration. *Proc. Natl. Acad. Sci.* **117:** 5486–5493

Cardinale S, Cisterna B, Bonetti P, Aringhieri C, Biggiogera M & Barabino SML (2007) Subnuclear Localization and Dynamics of the Pre-mRNA 3′ End Processing Factor Mammalian Cleavage Factor I 68-kDa Subunit. *Mol. Biol. Cell* **18:** 1282

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

Dismuke DJ & Aiken C (2006) Evidence for a functional link between uncoating of the human immunodeficiency virus type 1 core and nuclear import of the viral preintegration complex. *J Virol* **80:** 3712–3720

Fox AH, Lam YW, Leung AKL, Lyon CE, Andersen J, Mann M & Lamond AI (2002) Paraspeckles: A Novel Nuclear Domain. *Curr. Biol.* **12:** 13–25

Francis AC & Melikyan GB (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.e6

Freed EO (2001) HIV-1 Replication. *Somat. Cell Mol. Genet.* **26:** 13–33

Ganor Y, Real F, Sennepin A, Dutertre C-A, Prevedel L, Xu L, Tudor D, Charmeteau B, Couedel-Courteille A, Marion S, Zenak A-R, Jourdain J-P, Zhou Z, Schmitt A, Capron C, Eugenin EA, Cheynier R, Revol M, Cristofari S, Hosmalin A, et al (2019) HIV-1 reservoirs in urethral macrophages of patients under suppressive antiretroviral therapy. *Nat. Microbiol.* **4:** 633

Geis FK & Goff SP (2019) Unintegrated HIV-1 DNAs are loaded with core and linker histones and transcriptionally silenced. *Proc. Natl. Acad. Sci.* Available at: https://www.pnas.org/content/early/2019/10/30/1912638116 [Accessed November 15, 2019]

Gelderblom HC, Vatakis DN, Burke SA, Lawrie SD, Bristol GC & Levy DN (2008) Viral complementation allows HIV-1 replication without integration. *Retrovirology* **5:** 60

Gordon S & Taylor PR (2005) Monocyte and macrophage heterogeneity. *Nat. Rev. Immunol.* **5:** 953

Goujon C, Moncorgé O, Bauby H, Doyle T, Ward CC, Schaller T, Hué S, Barclay WS, Schulz R & Malim MH (2013) Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection. *Nature* **502:** 559–562

Hamid FB, Kim J & Shin C-G (2017) Distribution and fate of HIV-1 unintegrated DNA species: a comprehensive update. *AIDS Res. Ther.* **14:** 9

Heinz S, Texari L, Hayes MGB, Urbanowski M, Chang MW, Givarkes N, Rialdi A, White KM, Albrecht RA, Pache L, Marazzi I, García-Sastre A, Shaw ML & Benner C (2018) Transcription Elongation Can Affect Genome 3D Structure. *Cell* **174:** 1522-1536.e22

Honeycutt JB, Thayer WO, Baker CE, Ribeiro RM, Lada SM, Cao Y, Cleary RA, Hudgens MG, Richman DD & Garcia JV (2017) HIV persistence in tissue macrophages of humanized myeloid-only mice during antiretroviral therapy. *Nat. Med.* **23:** 638–643

Hu W-S & Hughes SH (2012) HIV-1 Reverse Transcription. *Cold Spring Harb. Perspect. Med.* **2:** a006882

Hulme AE, Perez O & Hope TJ (2011) Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. *Proc Natl Acad Sci U A* **108:** 9975–9980

Jao CY & Salic A (2008) Exploring RNA transcription and turnover in vivo by using click chemistry. *Proc. Natl. Acad. Sci. U. S. A.* **105:** 15779–15784

Koppensteiner H, Brack-Werner R & Schindler M (2012) Macrophages and their relevance in Human Immunodeficiency Virus Type I infection. *Retrovirology* **9:** 82

Laguette N, Sobhian B, Casartelli N, Ringeard M, Chable-Bessia C, Ségéral E, Yatim A, Emiliani S, Schwartz O & Benkirane M (2011) SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* **474:** 654–657

Lahaye X, Gentili M, Silvin A, Conrad C, Picard L, Jouve M, Zueva E, Maurin M, Nadalin F, Knott GJ, Zhao B, Du F, Rio M, Amiel J, Fox AH, Li P, Etienne L, Bond CS, Colleaux L & Manel N (2018) NONO Detects the Nuclear HIV Capsid to Promote cGAS-Mediated Innate Immune Activation. *Cell* **175:** 488-501.e22

Lamond AI & Spector DL (2003) Nuclear speckles: a model for nuclear organelles. *Nat. Rev. Mol. Cell Biol.* **4:** 605

Lee K, Ambrose Z, Martin TD, Oztop I, Mulky A, Julias JG, Vandegraaff N, Baumann JG, Wang R, Yuen W, Takemura T, Shelton K, Taniuchi I, Li Y, Sodroski J, Littman DR, Coffin JM, Hughes SH, Unutmaz D, Engelman A, et al (2010) Flexible use of nuclear import pathways by HIV-1. *Cell Host Microbe* **7:** 221–233

Lelek M, Casartelli N, Pellin D, Rizzi E, Souque P, Severgnini M, Serio CD, Fricke T, Griffero FD, Zimmer C, Charneau P & Nunzio FD (2015) Chromatin organization at the nuclear pore favours HIV replication. *Nat. Commun.*

Matera AG (1999) Nuclear bodies: multifaceted subdomains of the interchromatin space.

Trends Cell Biol. **9:** 302–309
McSwiggen DT, Hansen AS, Teves SS, Marie-Nelly H, Hao Y, Heckert AB, Umemoto KK, Dugast-Darzacq C, Tjian R & Darzacq X (2019) Evidence for DNA-mediated nuclear compartmentalization distinct from phase separation. *eLife* **8:** e47098
Mueller F, Senecal A, Tantale K, Marie-Nelly H, Ly N, Collin O, Basyuk E, Bertrand E, Darzacq X & Zimmer C (2013) FISH-quant: Automatic counting of transcripts in 3D FISH images. *Nat. Methods* **10:**
Naganuma T & Hirose T (2013) Paraspeckle formation during the biogenesis of long non-coding RNAs. *RNA Biol.* **10:** 456–461
Ouyang W, Mueller F, Hjelmare M, Lundberg E & Zimmer C (2019) ImJoy: an open-source computational platform for the deep learning era. *Nat. Methods* **16:** 1199–1200
Pendergrast PS, Wang C, Hernandez N & Huang S (2002) FBI-1 Can Stimulate HIV-1 Tat Activity and Is Targeted to a Novel Subnuclear Domain that Includes the Tat-P-TEFb—containing Nuclear Speckles. *Mol. Biol. Cell* **13:** 915–929
Peng K, Muranyi W, Glass B, Laketa V, Yant SR, Tsai L, Cihlar T, Müller B & Kräusslich H-G (2014) Quantitative microscopy of functional HIV post-entry complexes reveals association of replication with the viral capsid. *eLife* **3:** e04114
Petit C, Schwartz O & Mammano F (1999) Oligomerization within Virions and Subcellular Localization of Human Immunodeficiency Virus Type 1 Integrase. *J. Virol.* **73:** 5079–5088
Petit C, Schwartz O & Mammano F (2000) The Karyophilic Properties of Human Immunodeficiency Virus Type 1 Integrase Are Not Required for Nuclear Import of Proviral DNA. *J. Virol.* **74:** 7119–7126
Price AJ, Fletcher AJ, Schaller T, Elliott T, Lee K, KewalRamani VN, Chin JW, Towers GJ & James LC (2012) CPSF6 Defines a Conserved Capsid Interface that Modulates HIV-1 Replication. *PLOS Pathog.* **8:** e1002896
Rasaiyaah J, Tan CP, Fletcher AJ, Price AJ, Blondeau C, Hilditch L, Jacques DA, Selwood DL, James LC, Noursadeghi M & Towers GJ (2013) HIV-1 evades innate immune recognition through specific cofactor recruitment. *Nature* **503:** 402–405
Schmid M, Speiseder T, Dobner T & Gonzalez RA (2014) DNA Virus Replication Compartments. *J. Virol.* **88:** 1404–1420
Schwende H, Fitzke E, Ambs P & Dieter P (1996) Differences in the state of differentiation of THP- 1 cells induced by phorbol ester and 1,25- dihydroxyvitamin D3. *J. Leukoc. Biol.* **59:** 555–561
Spector DL (2001) Nuclear domains. *J Cell Sci* **114:** 2891–2893
Spector DL & Lamond AI (2011) Nuclear Speckles. *Cold Spring Harb. Perspect. Biol.* **3:** a000646
Stremlau M, Owens CM, Perron MJ, Kiessling M, Autissier P & Sodroski J (2004) The cytoplasmic body component TRIM5α restricts HIV-1 infection in Old World monkeys. *Nature* **427:** 848
Stultz RD, Cenker JJ & McDonald D (2017) Imaging HIV-1 Genomic DNA from Entry through Productive Infection. *J. Virol.* **91:** e00034-17
Sumner RP, Harrison L, Touizer E, Peacock TP, Spencer M, Zuliani-Alvarez L & Towers GJ (2019) Disrupting HIV-1 capsid formation causes cGAS sensing of viral DNA. *bioRxiv***:** 838011
Warrilow D, Stenzel D & Harrich D (2007) Isolated HIV-1 core is active for reverse transcription. *Retrovirology* **4:** 77
Xu H, Franks T, Gibson G, Huber K, Rahm N, De Castillia CS, Luban J, Aiken C, Watkins S, Sluis-Cremer N & Ambrose Z (2013) Evidence for biphasic uncoating during HIV-1 infection from a novel imaging assay. *Retrovirology* **10:** 70
Yamazaki T, Souquere S, Chujo T, Kobelke S, Chong YS, Fox AH, Bond CS, Nakagawa S, Pierron G & Hirose T (2018) Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation. *Mol. Cell* **70:** 1038-1053.e7
Zhang Q, Chen C-Y, Yedavalli VSRK & Jeang K-T (2013) NEAT1 Long Noncoding RNA and

Paraspeckle Bodies Modulate HIV-1 Posttranscriptional Expression. *mBio* **4:** e00596-12

FIGURE LEGENDS

Figure 1: HIV genomes form large nuclear foci in macrophages

(**A-C**) EdU-labelled viral DNA forms large foci in ThP1 cells infected by HIV-1. EdU images are shown in red, DAPI images in blue. (**A**) Left: nucleus of an infected cell at 48 h p.i. displays three large and bright EdU foci (colored arrows), as well as small and dim punctae (grey arrows). Center: Magnified views have colored borders matching the colors of the arrows. 3D images show that EdU foci are located within the nucleoplasm (**Movie S1**, **Fig. S2**). Right: nucleus of an uninfected cell displays no EdU signal. (**B**) Graph shows EdU intensity profiles along lines shown in **A**, with colors matching the corresponding arrows. (**C**) Dual-color image of an infected cell with EdU (red) and immunolabeled lamin (grey). Curves to the right plot the intensity profile of EdU and lamin along the dashed blue line. The EdU intensity peaks do not coincide with lamin enrichment. See also **Fig. S3**. (**D**) Multi-color images of an infected cell showing EdU (red) with CA (yellow) and integrase (green). The colocalization of nuclear EdU foci with CA and IN confirms that EdU specifically labels viral DNA and shows the presence of these proteins in vDNA foci. See also **Fig. S4**. (**E,F**) vDNA foci are much larger than the predicted size of single viral genomes. (**E**) Left image shows an observed, EdU-labelled vDNA focus. Right: simulations of a single linear or circular 10 Kb long chromatinized DNA polymer chain and corresponding predicted images in diffraction-limited (~300 nm resolution) microscopy. (**F**) Boxplots show the distribution of sizes (FWHM) of n=40 measured vDNA foci (for MOI 100) compared to the sizes predicted for linear and circular chains based on n=100 simulated configurations each. All differences are highly significant ($p<10^{-15}$). (**G**) Images of vDNA foci in ThP1 cells infected with multiplicities of infection (MOI) 10, 50 and 100. Images for MOI 10 are

shown at two different contrast levels to reveal dimmer spots. (**H**) Boxplots compare the peak intensities inside individual vDNA foci for MOIs 10, 50 and 100. Intensities increase significantly with MOI (MOI 10 vs. MOI 50: $p=0.006$; MOI 50 vs. MOI 100: $p<10^{-10}$; MOI 10 vs. MOI 100: $p<10^{-6}$). (**J**) Histograms show the number of foci per nucleus for MOIs of 10 (blue), 50 (green) and 100 (red).

Figure 2: Nuclear foci contain multiple viral RNAs and DNA genomes

(**A-C**) vDNA foci are also foci of vRNA. (**A**) Dual-color image of an infected ThP1 cell showing the viral DNA (red) and the RNA visualized by RNA-FISH (green). See also **Fig. S8**. (**B**) Venn diagram shows the number of vDNA foci and vRNA foci and the number of vDNA foci colocalizing with vRNA foci. (**C**) Scatter plot shows the intensities of vDNA and vRNA in colocalizing foci. (**D,E,F**) In ThP1 cells co-infected with a HIV1-GFP virus and a HIV1-LUC virus, nuclear foci contain mixtures of both viruses. (**D**) Dual-color image shows RNA-FISH against GFP (red) and RNA-FISH against LUC (green). See **Fig. S9**. (**E**) Venn diagram shows the number of GFP RNA foci, the number of LUC RNA foci, and the number of GFP RNA foci colocalizing with LUC RNA foci. (**F**) Scatter plot shows intensities of GFP RNA and LUC RNA in colocalizing foci.

Figure 3: Viral clusters contain specific nuclear body factors

(**A-C**) Images of infected (right) or uninfected (left) ThP1 cells showing the vDNA (EdU) in red, the nucleus (DAPI) in blue, and selected nuclear body factors in green. (**A**) Green image shows immunolabeling of CPSF6. (**B**) Green image shows RNA-FISH against NEAT1. (**C**) Green image shows immunolabeling of SC35. See also

Figs. S11-14. (**D**) Boxplots show Pearson correlations between vDNA (EdU) and CPSF6, NEAT1 or SC35 in 17-19 regions of interest (ROIs). Significance of positive or negative correlations was assessed using the Costes method of random ROI shifts. Highly significant ($p<0.01$) positive correlations between vDNA and CPSF6 intensities are found in 16 out of 19 ROIs; highly significant negative correlations between vDNA and NEAT1 are found in 19/19 ROIs, and highly significant positive correlations between vDNA and SC35 are found for 16/17 ROIs. (**E**) Intensity profiles of EdU and DAPI along the dotted lines in the vDNA images above.

Figure 4: Nuclear clusters can form in absence of integration and contain incoming vRNA

(**A**) Image of a ThP1 cell infected with an integration-deficient HIV-1 carrying the mutation D116A in the catalytic site of IN. The EdU-labeled vDNA is shown in red and the vRNA detected by RNA-FISH in green. The nucleus (DAPI) is shown in blue. See **Fig. S15** for a larger image region. (**B**) Venn diagram shows the number of vRNA clusters, the number of vDNA clusters, and the number of vRNA clusters colocalizing with vDNA clusters. (**C**) Scatter plot shows vDNA and vRNA intensities in colocalizing vDNA and vRNA clusters. (**D**) ThP1 cells were infected with HIV-1 at an MOI of 20, in absence (left) or presence (right) of nevirapine (NVP, 10 µM). Blue: DAPI staining. Green: vRNA (RNA-FISH). Red: vDNA (EdU). (**E**) Scatter plots show intensities of vDNA and vRNA in detected vRNA clusters. (**F**) Image of a ThP1 cell infected by HIV-1 with EU labeled RNA in absence (left) or presence (right) of nevirapine (NVP, 10 µM). Blue: DAPI staining. Green: vRNA (RNA-FISH). Grey: EU-labeled vRNA. See

Fig. S22 for a larger image region. (**G**) Intensity profiles of EU and RNA-FISH along the dotted lines in panel **F** (top).

Figure 5: Reverse transcription in nuclear clusters of infected macrophages

The left panel shows the timeline of drug exposure experiments. (**A,B**) Cells were exposed to NVP for 48 h (**A**) or 72 h after infection (**B**), then NVP was washed out, and cells were fixed for click chemistry 24 h later. (**C**) Cells were exposed to PF74 for 72 h after infection, then fixed for click chemistry. (**D**) Cells were exposed to NVP for 24 h after infection, then NVP was washed out, and cells were exposed to PF74, before being fixed for click chemistry 24 h later. (**A-D**) Images show vRNA and vDNA in an infected ThP1 cell (MOI 20) for each of the four experimental conditions. See **Fig. S18-S20** for images of larger regions. (**E,F,H**) Scatter plots show vDNA and vRNA intensities in colocalizing clusters. Replicates of the four experiments yielded similar results. (**G**) qPCR measures DNA synthesis (left) and nuclear import (right) at 72 h p.i. in absence of drug treatment, or after exposure to NVP or PF74.

METHODS

Cell culture

Human ThP-1 cells (ATCC TIB-202) were grown in RPMI 1640 medium supplemented with 10 % (vol/vol) fetal bovine serum and 1% (vol/vol) penicillin-streptomycin. For infections, 12-well plates containing coverslips were seeded with $1x10^6$ THP-1 cells and treated with PMA (259 µg/ml final concentration) at 37 °C and 5% CO_2. Twenty four hours post-stimulation, non-adherent cells were removed, adherent cells were washed, and further cultured. PMA was present during the entire experiment.

Plasmids and Viral production

The plasmid HIV-1ΔEnv IN_{HA} (D116A)ΔNef was obtained by insertional mutagenesis using the QuikChange II XL Site-Directed Mutagenesis kit (Agilent). HIV-1 viruses were produced by cotransfection with calcium phosphate with 10 µg HIV-1 LAI (BRU) (or NL4.3) ΔEnv Virus (NIH) or with the modified versions HIV-1ΔEnvIN_{HA} (Petit *et al*, 1999) or HIV-1ΔEnv IN_{HA} (D116A)ΔNef in combination with 1 µg of VSV-G envelope expression plasmid pHCMV-G (VSV-G) and 3 µg of SIV_{MAC} Vpx (Durand *et al*, 2013). The viruses collected from 293T cells 48 h post-transfection were ultracentrifuged at 4 ºC for 1h at 22,000 rpm. Virus normalizations were performed by p24 ELISA according to the manufacturer's instructions (Perkin Elmer)

or by qPCR. HIV-1 ΔEnv Δnef LUC and HIV-1 ΔEnv Δnef GFP viruses were produced and titered similarly.

Quantitative PCR

DNA synthesis, nuclear import and integration during HIV-1 infection in ThP1 cells were quantified by qPCR. We analyzed for late reverse transcription (LRT) products representing HIV-1 DNA synthesis in the cell and 2LTRs to measure nuclear import by qPCR. Integration of proviruses into the human genome were measured by ALU PCR. Viruses were treated for 30 min at 37°C with 1,000 U of DNase I (Roche). As a control, 10 μM nevirapine was used in infected cells. Total cellular DNA was isolated using the QIAamp DNA micro kit (Qiagen) at 7 and 24 h p.i. Viral DNA synthesis products and 2LTRs were measured at different time points by real-time PCR. LRT, 2LTRs and Alu PCR were performed as described (Lelek *et al*, 2015; Di Nunzio *et al*, 2013). LRT were amplified using the primers MH531 and MH532, with the standard curve prepared using the plasmid coding for the viral genome. 2LTRs were amplified by primers MH535/536 and probe MH603, using as standard curve the pUC2LTR plasmid which contains the HIV-1 2LTR junction. Integration was assessed by Alu-PCR, using primers designed in the U3 region of LTR (Lelek *et al*, 2015; Di Nunzio *et al*, 2013). The standard curve was prepared as follows: DNA generated from infected cells was end point diluted in DNA prepared from uninfected cells and serial dilutions were made. The control of the first round PCR was the amplification without Alu primers but only U3 primers (Lelek *et al*, 2015; Di Nunzio *et al*, 2013). Dilutions of the first round were processed by real time PCR (Lelek *et al*, 2015; Di

Nunzio *et al*, 2013). All experiments were carried out using internal controls such as infection in presence of RAL (10μM) and/or NVP (10μM). LRT, 2-LTR and Alu-PCR reactions were normalized by amplification of the Actin gene (Lelek *et al*, 2015; Di Nunzio *et al*, 2013).

Virus infection

Cells were infected at least 24 h after initial PMA stimulation with MOIs 10, 50 or 100, based on the viral titer calculated on ThP1 cells ($8.93x10^8$ TU/mL) by qPCR. For experiments with antiretroviral drugs, cells were infected with HIV-1 in the presence of 10 μM RAL or 10 μM NVP or 1.5 μM PF74.

Drug washout experiment

Cells were grown and differentiated as described above. Cells were infected with an MOI of 50 in the presence of EdU and NVP or PF74. For experiments where NVP was washed out, the drug-containing medium was removed, cells were washed twice with 1 mL of medium for 15 min, fresh medium was added and cells were incubated for the indicated amount of time. The medium was changed every 24 h and reconstituted with or without the PF74 drug, as indicated.

Labeling of viral DNA by click chemistry and immunolabeling

For imaging of vDNA, CA and IN, ThP1 cells were seeded in 12-well plates on cover glass in the presence of PMA and infected on the following day with HIV-1 at different MOI in medium containing 5 μM EdU. After 24 hours, the medium was removed and replaced by fresh medium containing 5 μM EdU and incubation was continued at 37 °C. To stop the infection, cells were washed with warm PBS and fixed with 4% paraformaldehyde in PBS for 20 min at room temperature.

For vDNA labeling, cells were washed twice with PBS supplemented with 3% bovine serum albumin (BSA) and permeabilized with 0.5% (vol/vol) Triton X-100 for 30 min. After washing with 3% BSA in PBS twice, click-labeling was performed for 30 min at room temperature using the Click-iT EdU- Alexa Fluor 647 Imaging Kit (Thermo Fisher Scientific) following the manufacturer's instructions.

For immunolabeling, cells were blocked for 30 min with 3% BSA in PBS and permeabilized with 0.5% (vol/vol) Triton X-100 for 30 min. After 2 washes with 3% BSA in PBS, cells were incubated with the primary antibody in 1% BSA in PBS for 1 h at room temperature. After washing with 1% BSA in PBS, a secondary antibody was used for 1 h at room temperature in 1% BSA in PBS. The primary and secondary antibodies used in this study are listed in **Note S1**.

In experiments that combine EdU and protein labeling, the click chemistry reaction was performed prior to immunolabeling.

Labeling of viral RNA with single molecule RNA-FISH

To visualize individual vRNA molecules, we used the smiFISH approach (Tsanov *et al*, 2016). Unlabelled primary probes are designed to target the RNA of interest, and can be pre-hybridized with fluorescently labeled secondary detector oligonucleotides for visualization. All probes, except against Neat1, were designed using either OLIGOSTAN (Tsanov *et al*, 2016) or Stellaris Probe designer, and purchased from Integrated DNA Technologies (IDT). All probe sequences are available in **Tables S1-3**. We used 24 primary probes, each 18-20 nt long, against the HIV-1 POL gene. Probe sets against GFP and LUC comprised 18 and 22 probes, respectively. Secondary probes are conjugated to either Cy3 or Cy5. To detect Neat1, we used Stellaris FISH probes against human NEAT1 5' segment with QUASAR 570 DYE SMF-2036-1.

Cells were fixed as described above, washed twice with PBS and stored in nuclease-free 70% ethanol at -20 °C until labeling. On the day of the labeling, the samples were brought to room temperature, washed twice with wash buffer A (2x SSC in nuclease-free water) for 5 min, followed by two washing steps with washing buffer B (2X SSC and 10% formamide in nuclease-free water) for 5 min.

The target-specific primary probes were pre-hybridized with the fluorescently labeled secondary probes via a complementary binding readout sequence. The stock concentration of the probes were: Luciferase LUC: 148.6 ng/µL, POL: 1497 ng/µL and GFP: 163.3 ng/µL.

The reaction mixture contained primary probes at a final concentration of 40 pm, and secondary probes at a final concentration of 50 pm in 1x NEB3 buffer. Pre-hybridization was performed in a PCR machine with the following cycles: 85 °C

for 3 min, followed by heating to 65 °C for 3 min, and a further 5 min heating at 25 °C. 2 µL of this FISH-probe stock solution was added to 100 µL of hybridization buffer (10% (w/v) dextran, 10% formamide, 2X SSC in nuclease-free water).

Samples were placed on Parafilm in a humidified chamber on 100 µL of hybridization solution, sealed with Parafilm, and incubated overnight at 37°C. The next day, cells were washed in the dark at 37°C without shaking for >30min twice with pre-warmed washing buffer B. Sample were washed once with PBS for 5 min, stained with DAPI in PBS (1:10000) for 5 min, and washed again in PBS for 5 min. Samples were mounted in ProLong Gold antifade mounting medium.

For simultaneous detection of viral DNA with click chemistry and vRNA with RNA-FISH, the click reaction was performed prior to RNA-FISH and as described above, except no BSA was used in order to minimize RNA degradation by RNases. The ratio copper-protectant/copper solution was 1.5:1 to avoid degradation of RNA, as indicated by the manufacturer.

Polymer simulations

To estimate the expected size and images of single vDNA particles, we performed polymer simulations in which the vDNA molecule is represented by a semiflexible chain of beads undergoing random motions (Langevin dynamics) (Arbona *et al*, 2017). We assumed chromatinized (i.e. nucleosome-containing) DNA and performed two types of simulations: one with a linear chain (as expected e.g. for integrated vDNA) and the other with a circular chain (e.g. for unintegrated, episomal vDNA).

Each bead represents a single nucleosome consisting of 175 bp of DNA (nucleosomal + linker DNA) confined in a 11 nm spherical bead. We assumed the bending stiffness of the polymer to be 11 nm (single bead) and connected consecutive beads via spring potentials. We started our simulations from a linear or circular configuration and allowed the polymer to reach equilibrium (as judged by the temporal evolution of the gyration radius) before sampling one configuration for each simulation. Simulations were run using the LAMMPS simulation library (https://lammps.sandia.gov/). Simulated images were obtained by convolving the simulated configuration with a 3D Gaussian kernel of standard deviation 100 nm approximating the microscope point spread function. We used 100 independent simulations to compute the distributions of sizes (FWHM) for each model.

Imaging and analysis

Three-dimensional image stacks with a z-spacing of 200 nm were captured on a wide-field microscope (Nikon TiE Eclipse) equipped with a 60X 1.4 NA objective and an sCMOS camera (Hamamatsu Orcaflash 4), with a lumencor SOLA LED light source with adequate Semrock filters, and controlled with MicroManager 1.4.

Nuclei were automatically detected by a trained neural network implemented within the computational platform ImJoy (Ouyang *et al*, 2019) (Plugin DPNUnet). Spots in Edu and FISH images were detected with a standard spot detection approach: images were first filtered with a Laplacian of Gaussian filter (LoG), then spots were identified with a local maximum filter with a user defined minimum intensity.

Colocalization analysis between two channels was implemented in a custom ImJoy plugin. Only spots located within the detected nuclei were considered. Colocalization was implemented as a linear assignment problem (Python function linear_sum_assignment from SciPy), where each spot from the first channel is assigned to one spot in the second channel. Each spot can be only assigned once, and only assignments below a user defined distance threshold are permitted. Spot intensities for co-localized spots were measured as the maximum intensity in a +/- 1 pixel window. To calculate p-values for the reported co-localization, the colocalization analysis was repeated 100 times after randomly moving ("jitter") detected spots in one channel by 1000 nm in XYZ. The reported p-values are the proportion of jitter iterations where a higher colocalization was obtained than in absence of jitter.

EdU enrichment for Nevirapine experiments was calculated after nuclei segmentation and spot detection as described above. Only RNA foci within the detected nuclei and with intensities above a user defined threshold were considered. FISH and EdU intensities at these locations were measured as the maximum intensity in a window of +/- 300 nm.

To analyze colocalization of immunolabeled proteins and Neat1 with vDNA, we used the Costes method (Costes *et al*, 2004) as implemented in the coloc2 plugin of Fiji (Schindelin *et al*, 2012). First, nuclei were manually selected to exclude background that could lead to inappropriate thresholding during correlation analysis. For each selected region, the plugin automatically thresholds each channel such that the Pearson correlation coefficient of channel intensities below the threshold is 0. The

Pearson correlation of intensities above the thresholds is then reported. Similarly to above, statistical significance was determined by comparing this correlation to that of 100 randomly jittered images, where jittering consisted in moving one channel by 3 pixels. If the Pearson correlation of the unjittered channels is positive (respectively, negative), the reported p-value reflects the proportion of jitter iterations in which the Pearson coefficient is higher (respectively, lower).

To estimate the size of nuclear vDNA foci, we used Fiji (Schindelin *et al*, 2012) to manually draw line profiles across EdU spots. Each intensity profile was then fitted with a Gaussian function and the corresponding full width at half maximum (FWHM) was used as a measure of the vDNA cluster size.

METHODS REFERENCES:

Arbona J-M, Herbert S, Fabre E & Zimmer C (2017) Inferring the physical properties of yeast chromatin through Bayesian analysis of whole nucleus simulations. *Genome Biol.* **18:** 81

Costes SV, Daelemans D, Cho EH, Dobbin Z, Pavlakis G & Lockett S (2004) Automatic and Quantitative Measurement of Protein-Protein Colocalization in Live Cells. *Biophys. J.* **86:** 3993–4003

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

Durand S, Nguyen X-N, Turpin J, Cordeil S, Nazaret N, Croze S, Mahieux R, Lachuer J, Legras-Lachuer C & Cimarelli A (2013) Tailored HIV-1 Vectors for Genetic Modification of Primary Human Dendritic Cells and Monocytes. *J. Virol.* **87:** 234–242

Lelek M, Casartelli N, Pellin D, Rizzi E, Souque P, Severgnini M, Di Serio C, Fricke T, Diaz-Griffero F, Zimmer C, Charneau P & Di Nunzio F (2015) Chromatin organization at the nuclear pore favours HIV replication. *Nat. Commun.* **6:** 6483

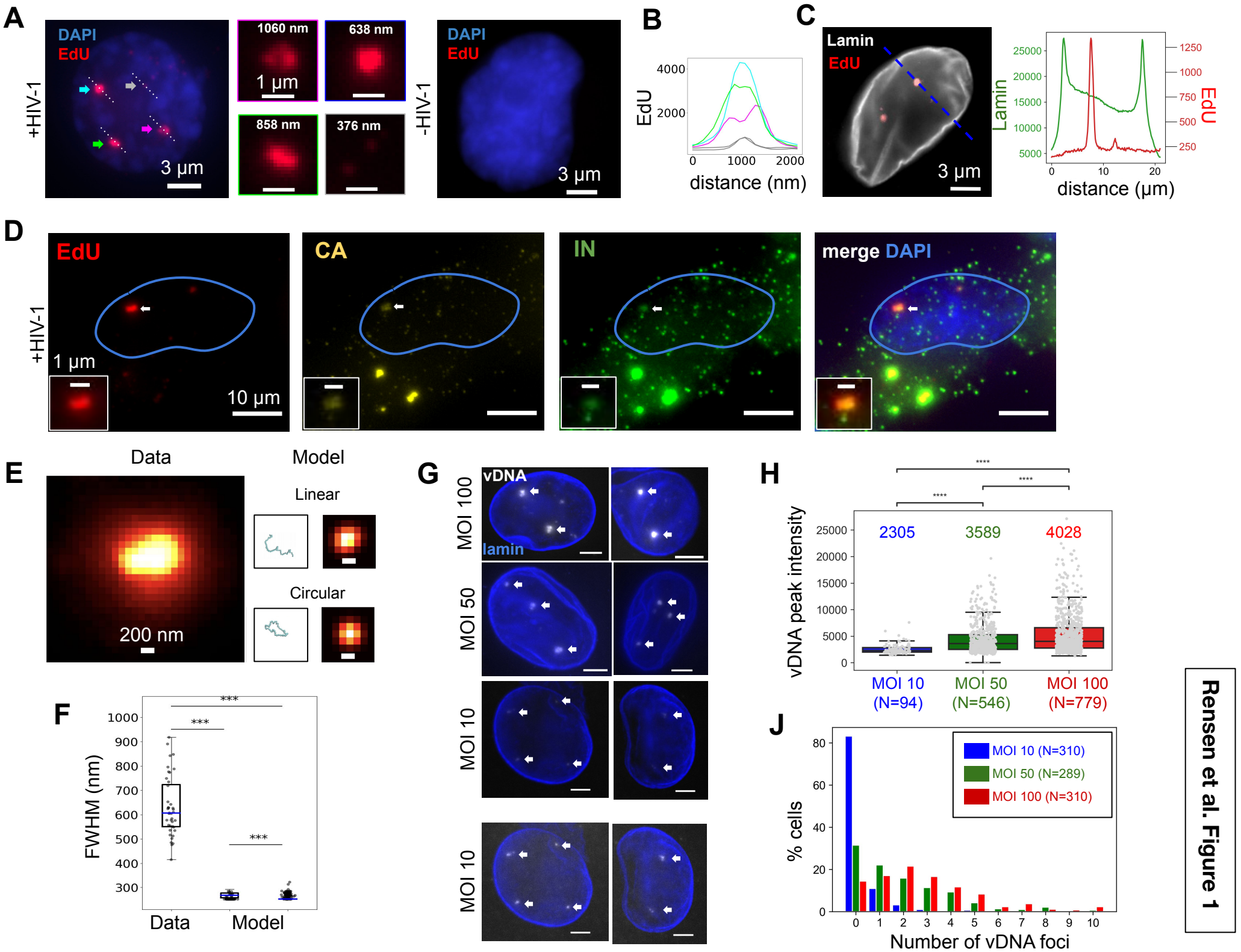
Ouyang W, Mueller F, Hjelmare M, Lundberg E & Zimmer C (2019) ImJoy: an open-source computational platform for the deep learning era. *Nat. Methods* **16:** 1199–1200

Petit C, Schwartz O & Mammano F (1999) Oligomerization within Virions and Subcellular Localization of Human Immunodeficiency Virus Type 1 Integrase. *J. Virol.* **73:** 5079–5088

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K,

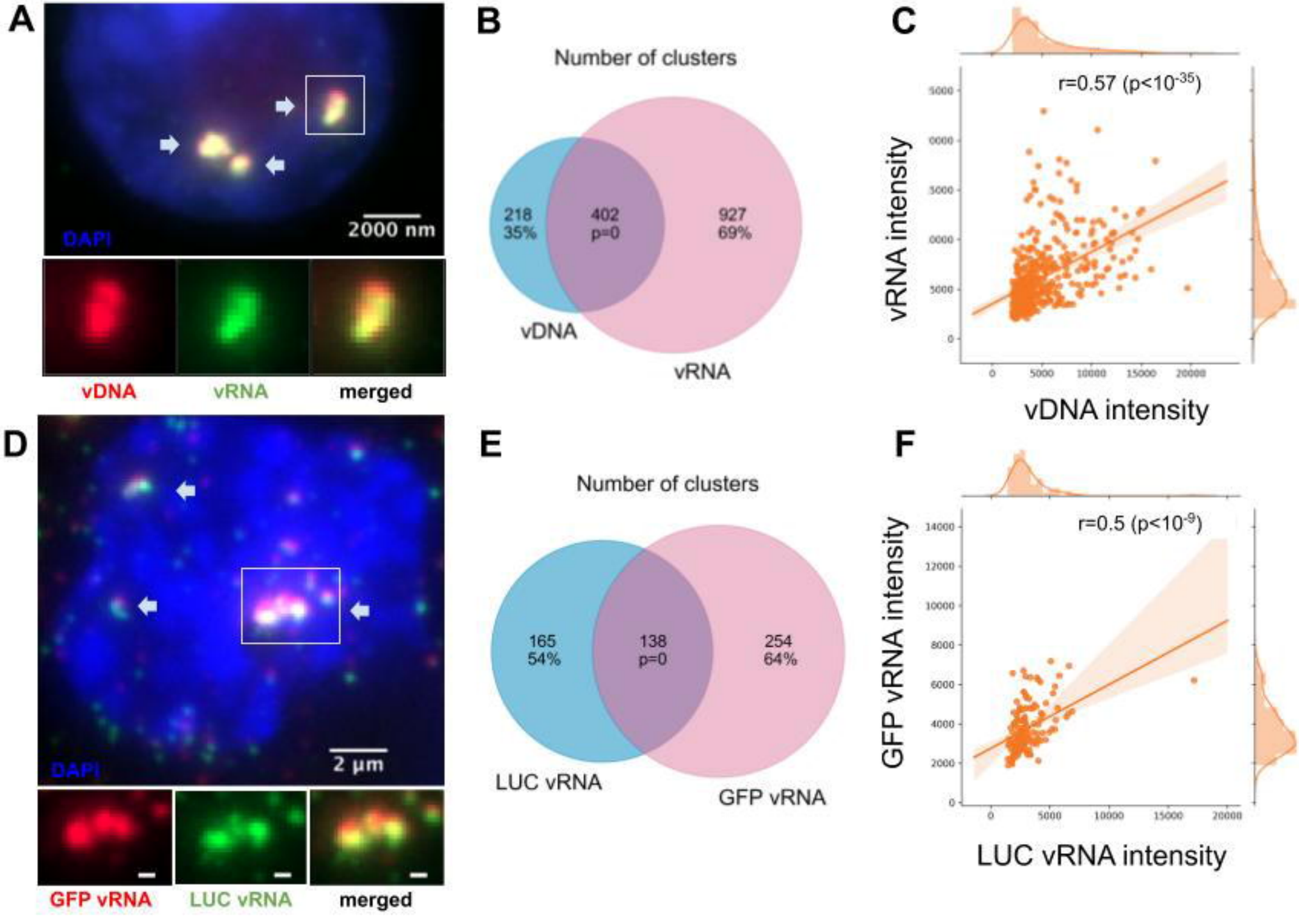
Tomancak P & Cardona A (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9:** 676–682

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

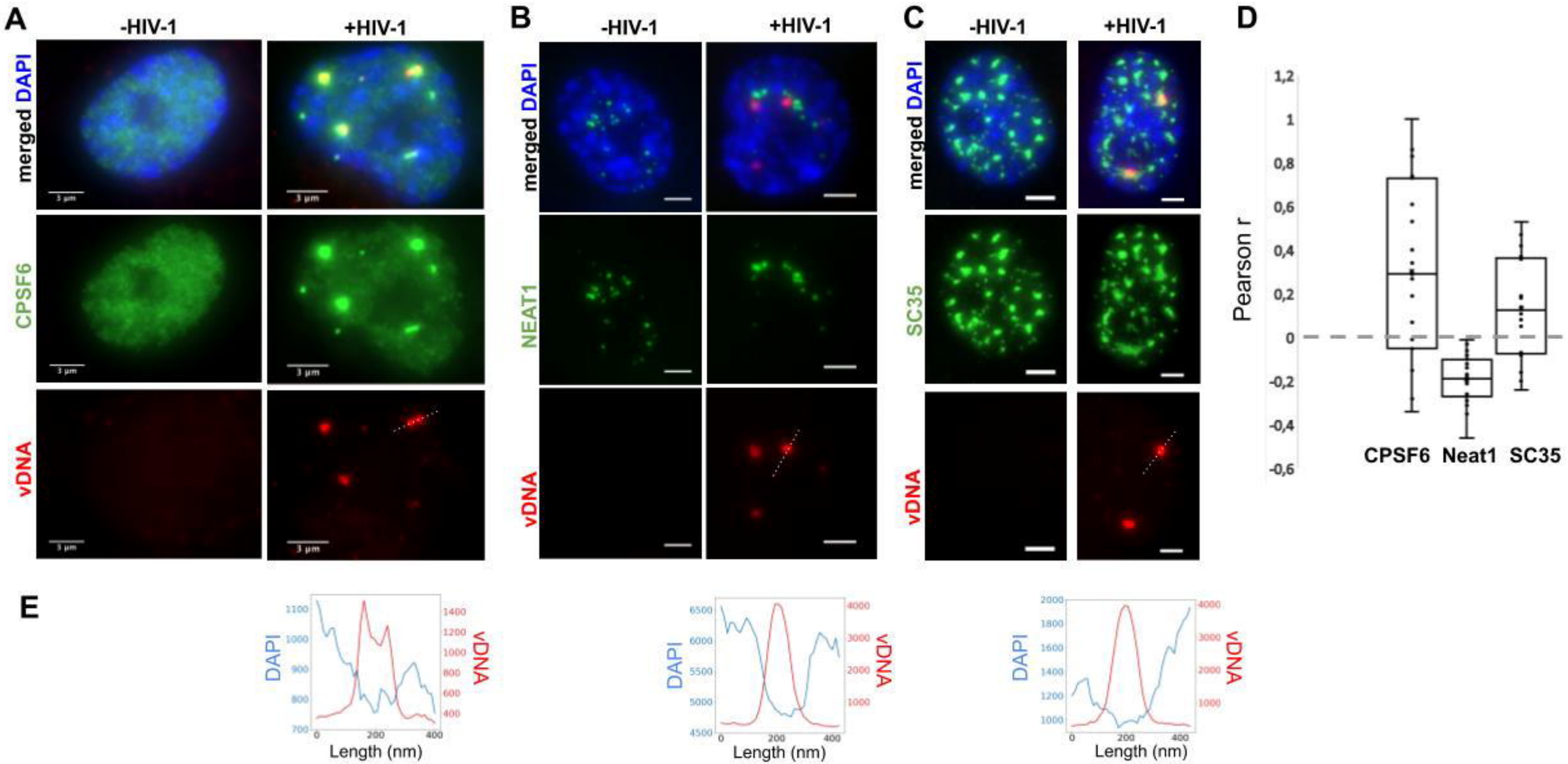


Rensen et al. Figure 1

Rensen et al. Figure 2



Rensen et al. Figure 3



Rensen et al. Figure 4

A

IN D116A

3000 nm

DAPI

vDNA vRNA merged

B

Number of clusters

185
74%

65
p=0

859
92%

vDNA

vRNA

C

r=0.51 ($p<10^{-4}$)

vRNA intensity

vDNA intensity

D

+HIV-1
no Nevirapine

+HIV-1
10 µM Nevirapine

merged

3 µm

vRNA

vDNA

E

vDNA intensity

vRNA intensity
(N=666)

vDNA intensity

vRNA intensity
(N=629)

F

+HIV-1
no Nevirapine

+HIV-1
10 µM Nevirapine

merged

3 µm

RNA-FISH vRNA

EU vRNA

G

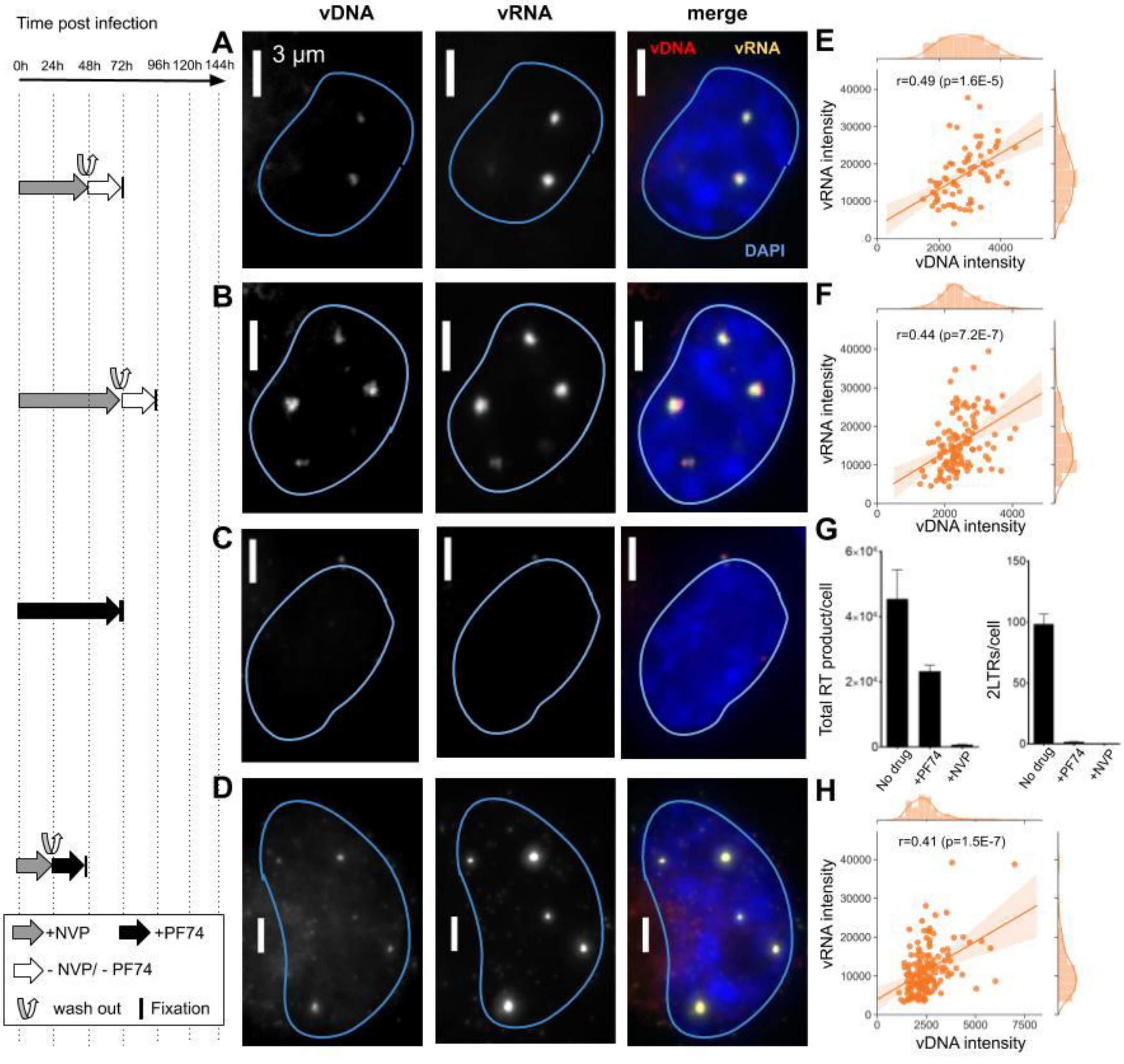
Intensity RNA-FISH

Intensity EU

length (µm)

Intensity RNA-FISH

Intensity EU

length (µm)

Rensen et al. Figure 5