HIV-induced membraneless organelles orchestrate post-nuclear entry steps

Viviana Scoca, Marion Louveaux, Renaud Morin, Dmitry Ershov, Jean-Yves
Tinevez, Francesca Di Nunzio

1 **HIV-induced membraneless organelles orchestrate post-**
2 **nuclear entry steps.**

3

4

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

7

8 [1]*Department of Virology, Advanced Molecular Virology and Retroviral Dynamics*
9 *Group, Pasteur Institute, Paris, France*
10 [2] *BioSPC Doctoral School, Université de Paris, Paris, France*
11 [3] *Department of Cell Biology and Infection, Biological Image Analysis, Pasteur Institute,*
12 *Paris, France*
13 [4] *Image Analysis Hub/ C2RT, Pasteur Institute, Paris, France*
14 [5] *Imactiv-3D, 1 Place Pierre Potier, 31106 Toulouse, France*
15 [6] *Department of Computational Biology, Hub of Bioinformatics and Biostatistics,*
16 *Pasteur Institute, USR 3756 CNRS, Paris, France.*

17

18 *equal contribution

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

20 **Abstract**

21 HIV integration occurs in chromatin sites that favor the release of high levels of viral

22 progeny, alternatively the virus is also able to coexist discreetly with the host. To

23 uncover the HIV-1 DNA fate in the nuclear landscape, we directly tracked the viral DNA

24 and the viral RNA by coupling HIV-1 ANCHOR technology with RNA FISH or MCP-

25 MS2 RNA-tagging bacterial system. Our computational imaging analysis revealed that

26 proviral forms are early located in proximity of the nuclear periphery of mitotic and non-

27 mitotic cells. HIV-1 triggers the nuclear clustering of the host factor CPSF6, but yet its

28 role is poorly understood. Our data show that CPSF6 clusters are part of HIV-induced

29 membraneless organelles (HIV-1 MLOs). Interestingly, we observed that viral genomic

30 RNAs and the integrase proteins are sequestered in CPSF6 clusters, in which the virus

31 can reverse transcribe, while the late retrotranscribed DNA is excluded from these

32 structures. HIV-1 MLOs are formed in the interchromatin space, indicating that they

33 are not proviral sites, but orchestrate viral events prior to the integration step. We found

34 transcription-competent proviruses localize outside HIV-1 MLOs, in LEDGF-abundant

35 regions, known to be active chromatin sites. This study highlights single functional

36 host-proviral complexes in their nuclear landscape, which is markedly restructured by

37 HIV-1 to favor viral replication.

38

Introduction

40 Immediately after fusion at the plasma membrane, HIV-1 cores are released into the

41 cytoplasm and move towards the nucleus, while the viral RNA (vRNA) genome begins

42 the process of reverse transcription (RT) into double-stranded DNA (dsDNA)

43 (Campbell and Hope, 2015; Di Nunzio, 2013; Scoca and Di Nunzio, 2021a). Once

44 imported in the nucleus, the interplay between the HIV-1 DNA genome and the host

45 chromatin compartment is crucial for the fate of the virus-host coexistence. HIV-1 can

46 adopt either an episomal or a proviral form (Butler et al., 2001). The fate of these viral

47 DNA (vDNA) forms and their surrounding chromatin landscape dictate the evolution of

48 HIV infection (Maldarelli, 2016; Sharkey et al., 2011; Sharkey et al., 2000). The

49 mechanisms determining the fate of the proviral DNA in the nucleus are still under

50 investigation (Liu et al., 2020; Olson et al., 2019) also because of the limits imposed

51 by available technologies. Real-time imaging approaches provide new insight into

52 unprecedented spatial information of individual HIV-1 infected cells. Several strategies

53 have enabled the visualization of the viral components during nuclear entry to indirectly

54 track the HIV-1 pre-integration complex (PIC) : Integrase (IN)-TC/FlAsH, APOBEC 3F

55 (A3F)-YFP, IN-YFP, Gag-iGFP, Cyclophilin A (CypA)-DsRed/ Capsid (CA), INsfGFP,

56 and CA-GFP (Burdick et al., 2017; Hubner et al., 2009; Lelek et al., 2012) (Francis et

57 al., 2016; Francis and Melikyan, 2018) (Mamede et al., 2017) (Albanese et al., 2008;

58 Francis et al., 2014). Recent new discoveries in the field shed light on the central role

59 of the viral capsid in leading the viral genome in the host nucleus (Blanco-Rodriguez

60 et al., 2020; Burdick et al., 2020), contrary to what was widely believed (Farnet and

61 Haseltine, 1991; Suzuki and Craigie, 2007). The viral nuclear translocation is dictated

62 by the capsid interplay with host factors, such as Cleavage and polyadenylation

63 specificity factor subunit 6 (CPSF6)(Buffone et al., 2018; Lee et al., 2010; Price et al.,

64 2014), which is a paraspeckle factor (Fox et al., 2002; Naganuma and Hirose, 2013)

65 and Nup153 (Di Nunzio et al., 2013; Lelek et al., 2015). CPSF6, vDNA, vRNA and

66 transcription factors have been found in nuclear speckles (NSs), suggesting that their

67 co-appearance is consistent with productive integration in these nuclear bodies

68 (Francis et al., 2020; Stultz et al., 2017). Interestingly, CPSF6 contains intrinsically

69 disordered mixed-charge domains which are responsible for the formation of liquid

70 condensates *in vitro* (Greig et al., 2020) and they may account for the remodeling of

71 membraneless organelles (MLOs) in infected cells (Scoca and Di Nunzio, 2021b).

72 However, although NSs contain viral components they are by definition interchromatin
73 granules (Lamond and Spector, 2003), thus they cannot be HIV-1 integration sites.
74 Thus, the specific role of these host/viral nuclear organelles remain unclear. Of note,
75 mechanisms underlying viral post-nuclear entry steps and the role of nuclear structures
76 induced by HIV-1 during viral replication could be revealed only by the direct co-
77 labeling of vDNA forms and host factors. Thus far, all vDNA forms from the early to the
78 late RT products can be indistinguishably visualized in fixed cells, either by supplying
79 5-Ethynyl-2´-deoxyuridine (EdU) during infection (Peng et al., 2014; Rensen et al.,
80 2020), which is limited to non-dividing cells (De Wit et al., 2019) or by DNA-FISH
81 (Marini et al., 2015) or bDNA (Chin et al., 2015). However, the exclusive labeling of
82 late reverse transcribed DNA products in fixed and in live cells is challenging but
83 important, because they are the only vDNA forms included in the mature PIC.

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

104 **Results**

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

106 With the aim of studying HIV-1 DNA nuclear fate in different cell types, we benefited of

107 the HIV-1 ANCHOR DNA labeling system (Fig. 1A) (Blanco-Rodriguez et al., 2020).

108 Once fully retrotranscribed, the HIV-1 genome carrying the bacterial sequence ANCH3

109 is specifically recognized by OR-GFP protein, which is expressed in the target cells,

110 genetically modified by the lentiviral vector (LV) OR-GFP (Blanco-Rodriguez et al.,

111 2020). The accumulation of OR-GFP, which is a modified version of the bacterial ParB

112 protein (Graham et al., 2014; Sanchez et al., 2015), on the ANCH3 sequence

113 generates bright nuclear signals (Fig.1A). Importantly, OR protein binds exclusively the

114 double stranded DNA (Saad et al., 2014). We cloned ANCH3 in the nef gene, thus,

115 HIV-1 ANCHOR system exclusively detects late reverse transcripts, because the

116 ANCH3 sequence is one of the last sequences to be converted into double stranded

117 DNA (Fig. 1A). Therefore, the HIV-1 ANCHOR approach allows only the visualization

118 of potential functional vDNA. Since the HIV-1 genome dissociates from viral proteins

119 to become a proviral DNA or an extrachromosomal DNA, the track of these viral forms

120 is the only direct measurement to investigate the fate of the viral genome into the host

121 nuclear space. Unlike previous studies (Burdick et al., 2017; Burdick et al., 2020; Chin

122 et al., 2015; Francis et al., 2020; Francis and Melikyan, 2018), we are now able to

123 directly visualize by live imaging the nuclear vDNA as GFP puncta appearing in the

124 host nucleus once retrotranscribed (Fig. 1A). HIV-1 ANCHOR is able to specifically

125 label the vDNA as shown by the correlation between the nuclear GFP puncta and the

126 MOI (multiplicity of infection) used (Fig. S1A). This specificity was corroborated using

127 the reverse transcription inhibitor, Nevirapine (NEV), which completely prevents the

128 detection of vDNA (Fig. 1B, Fig.S1B). We then investigated whether HIV-1 ANCHOR

129 allows the detection of different intranuclear vDNAs, such as episomal and integrated

130 forms. To generate episomal forms, either we infected cells with HIV-1 ANCH3 in the

131 presence of the integration inhibitor Raltegravir (RAL), or using an integration-deficient

132 virus, which is mutated at the catalytic site of the integrase (IN) enzyme (HIV-1 ANCH3

133 IN$_{D116A}$). We detected intranuclear GFP puncta under both conditions (Fig. 1B),

134 meaning that HIV-1 ANCHOR permits the detection of nuclear episomal vDNA forms.

135 The experiments were validated by qPCR, which confirmed the extrachromosomal

136 forms generated during infection when the integration step is impeded (Fig. 1B). We

137   next assessed whether HIV-1 ANCHOR is sufficiently sensitive to allow the detection

138   of integrated vDNA at the level of a single provirus. We selected single HeLa P4R5

139   clones carrying the HIV-1 ANCH3 genome. The imaging data correlates with results

140   obtained by ALU PCR (Fig. 1C), demonstrating that this technology is powerful enough

141   to detect a single provirus. Thus, we used this system to live track both proviruses and

142   episomal forms to investigate their behaviour in the nucleus (video 1, 2, and 3). We

143   performed live tracking of vDNAs in cells at six days post infection (p.i.) (video 1), which

144   contain a negligible amount of episomal forms, likely lost during cell division (Fig. S1C),

145   in the single-cell clone (video 2) or in cells containing only the episomal vDNA (video

146   3). Then, we compared vDNA trajectories in the aforementioned experimental

147   conditions (Fig. 1D). The stable clone and the cells at six days p.i. show vDNA with

148   similar trajectories and diffusion coefficient (D (nm^2/s) < 5). Instead, infected cells

149   carrying unintegrated viruses harbored two populations composed of episomal forms

150   with divergent behaviors, one with a low D (D (nm^2/s) < 5) and the other with a high D

151   (5 < D (nm^2/s) < 20) (Fig. 1D). Our data hint that most, but not all, episomal forms are

152   retained in particular nuclear regions, most likely interacting with host factors or

153   chromatin compartments, consistent with the findings of other studies (Geis and Goff,

154   2019; Zhu et al., 2018), whereas some episomal forms are free to move in the nuclear

155   space. On the other hand, the integrated forms showed a uniform, low-diffusive

156   behavior.

157

158   **Detection of actively transcribing proviruses at single-cell level in HIV-1 target**

159   **cells.**

160   The identification of the functional vDNA in the nucleus of infected cells has been one

161   of our goals. For this purpose, we combined HIV-1 ANCHOR technology with RNA

162   FISH (single molecule inexpensive FISH, smiFISH) against HIV-1 polymerase

163   transcripts, to detect vRNA foci of transcription (Tsanov et al., 2016), using ethanol

164   treatment and 5h of probe hybridization (Fig. S4). As expected, in cells treated with

165   RAL, where the integration step is blocked, vRNA signal is highly depleted. In

166   untreated cells it is possible to observe vDNA-vRNA associations, highlighting the

167   proviruses location (Fig. 2A), in fact proviruses are the only transcribing forms. We

168   strengthen our results on the ability to visualize proviruses by showing the vDNA-vRNA

169   association by live imaging. HIV-1 transcriptional activity was tracked by MCP-GFP

170   system (kind gift from E. Bertrand) to label the vRNA (Tantale et al., 2016) coupled

171   with HIV-1 ANCHOR technology (Fig. S2A and B). Next, we compared the intensity of

172   ANCHOR signals of the episomal forms (+ RAL) with those of the integrated forms (-

173   RAL) and observed that the signal of unintegrated DNA was two-fold brighter than the

174   proviral DNA spot (Fig. 2A) in HeLa cells. Interestingly, we also remarked, during live

175   imaging, that the vDNA of the integration-deficient virus (HIV-1 ANCH3 IN$_{D116A}$)

176   clusters in the nucleus (Fig. S2C, video 4). Aggregation of extrachromosomal forms

177   could prompt a brighter appearance of viral episomal DNA compared to a dimmer

178   signal of the individual proviral DNA (Fig. 2A), at least in HeLa cells. Next, we

179   pinpointed transcriptionally active proviruses in the main HIV-1 target cells, CD4$^+$ T

180   cells and macrophages, by co-detecting vRNA and vDNA (Fig. 2 B-D). We observed

181   that not all vDNAs were associated to a transcriptional focus, most likely because of

182   the presence of silent episomal forms. On the other hand, an abundant part of vRNA

183   foci colocalized with vDNA (Fig. 2B and C), meaning that we are able to detect single

184   transcribing proviruses. Taken together, our results suggest that we are able to track

185   down transcribing proviruses in the nuclear space of different cells and distinguish

186   them from silent forms.

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

189   The location of actively transcribing proviruses in the nuclear space of dividing cells is

190   still a subject of debate (Achuthan et al., 2018; Burdick et al., 2017; Chin et al., 2015;

191   Di Primio et al., 2013; Lelek et al., 2015; Marini et al., 2015), whereas their nuclear

192   location in terminally differentiated cells, as far as we know, has never been studied

193   before. In dividing cells, the 3D analysis of the nuclear location of the vDNA showed

194   the vDNA signals at 48 h p.i. to be more randomly distributed throughout the nucleus

195   compared to the vDNA spots at seven days p.i. (avg distances from the nuclear

196   envelope (NE): ~2.5 µm at 48 h p.i. vs ~1.4 µm at 7 days p.i.) (Fig. 3A). The difference

197   in the average distance of the vDNA spots from the NE between the early and late time

198   points p.i. may be due to the co-existence of heterogeneous viral forms at 48 h p.i.

199   (episomal and proviral forms), whereas the vDNA population is more homogeneous at

200   seven days p.i. (Fig. 1D, Fig. 3A). The episomal forms are distributed at a slightly

201   greater distance from the NE than the mixed vDNAs population at 48 hour p.i. when

202   compared to the vDNA at 7 d p.i. (Fig. 3A). Interestingly, we did not detect differences

203   between the distances of the vRNA foci from the NE at both time points (48h p.i. and

204 7d p.i.) and the vDNA spots at seven days p.i. (< 2 μm in avg), indicating that in dividing
205 cells the integration of the proviral DNA occurs early during infection near to the NE
206 (Fig. 3A). We performed similar 3D analysis to calculate the distances of vDNAs and
207 vRNAs from the NE in non-dividing cells at three days p.i.. The lack of mitosis hinders
208 the loss of the episomal forms in non-dividing cells (Fig. 3B). However, when we
209 compared the integration-defective viruses (only episomal forms) with the integration-
210 efficient viruses (episomal and proviral forms), the vDNAs of the latter were found
211 slightly closer to the NE (Fig. 3B). While the vRNAs (viral transcription foci) were closer
212 to the NE with respect to episomal forms (~2.5 μm for the vDNA IN_{D116A} vs ~1.8 μm for
213 the vRNA) (Fig.3B). Finally, our data show a similar nuclear spatial distance of viral
214 transcriptional foci from the NE at single-cell level in both dividing and non-dividing
215 cells (~1.8 μm in HeLa cells and in THP-1 cells) (Fig.3C).

216 **Late reverse transcribed DNAs separate from IN proteins sequestered in CPSF6**
217 **clusters.**
218 A long-standing question in the early steps of HIV life cycle is related to the study of
219 IN-vDNA dynamics after nuclear entry. How long does the association last and can we
220 detect it? There are, indeed, multiple difficulties due to the likely short-time nature of
221 the event or due to the technical limitations of vDNA and IN visualization. Therefore,
222 we exploited our vDNA tracking system to investigate about IN proteins localization in
223 primary CD4[+]T cell and THP-1 cells, challenging both cell types with HIV-1 ANCH3.
224 We found in both cell types that the vDNA/IN association event was very rare (Fig. 4A).
225 Nuclear IN proteins were found in the nucleus separated from the late reverse
226 transcripts (Fig. 4A). To confirm our data we performed live-imaging, asking whether
227 we could pinpoint the dynamic evolution of vDNA/IN association in the nucleus of
228 infected cells by coupling the HIV-1 ANCHOR system with the GIR virus (generous gift
229 from Edward Campbell) (Dharan et al., 2016; Hulme et al., 2015). GIR virus is
230 produced using a viral genome carrying the IN mutated in the catalytic site (HIV-1
231 ANCH3 IN_{D116A}) together with Gag-integrase-Ruby (GIR) plasmid. As consequence,
232 the only active IN in these viral particles is the IN-Ruby (Fig. 4B, Fig. S3). During a time
233 lapse of 4 hours in HeLa cells (Fig. S3), we have observed the signals of the IN and of
234 the vDNA always separated (~ 0.6 to 0.8 μm) (Fig. S3, video 5) corroborating the
235 results in fixed cells (Fig. 4A). Importantly, we did not synchronize the infection to have
236 the opportunity to visualize in a single cell different viral dynamics. We found several

237   IN foci alone and then colocalizing with the vDNA before the separation (Fig. 4B, movie

238   6, 7). We analysed the association of the two signals that rapidly (~10 min) separated

239   (Fig. 4B, video 6,7). These data suggest that the IN proteins that are part of the mature

240   PIC cannot be detected by a conventional light microscopy. Next, we investigated

241   whether the viral IN proteins are retained in particular nuclear regions by some nuclear

242   host factors. We found that the viral IN occupies, in nearly all cases analyzed, the same

243   positions of the host Cleavage and Polyadenylation Specificity Factor subunit 6

244   (CPSF6), in both dividing and non-dividing cells (IN to CPSF6 ~ 94.4% in CD4+T cells

245   and ~92.5% in THP-1 cells) (Figure 4C). Of note, HIV-1 reprograms the nuclear

246   localization of CPSF6 from being randomly dispersed in the nucleoplasm to form

247   particular aggregates in both primary lymphocytes and macrophages (Fig. 4C). Taken

248   together our results indicate that IN proteins after viral nuclear entry are retained in

249   CPSF6 clusters, while the late reverse transcribed DNA separates from these sites

250   (Fig. 4A-C). We next wondered if other viral components are also hosted in CPSF6

251   MLOs together with the IN. Adapting our RNA FISH protocol with a longer probe

252   hybridization and coupling it with the immunostaining of CPSF6 (Immuno-RNA FISH)

253   (Fig. S4), we uncovered the presence of incoming vRNA inside the MLOs (Fig. 4D).

254   Upon NEV treatment, HIV-1 genomes locate in CPSF6 clusters, both in CD4$^+$T and

255   THP-1 cells. In physiological conditions it seems that vRNA/CPSF6 association events

256   are more abundant in non-dividing cells than in dividing cells (~87% vs. ~33%), at 2

257   days post infection (Fig. 4D). Overall, the sequestration of viral IN proteins and viral

258   genomes in CPSF6 clusters hints that CPSF6 MLOs could serve as sites to release a

259   mature PIC that contains only few IN proteins needed for the integration step

260   (Ballandras-Colas et al., 2017; Hare et al., 2009). We called these particular nuclear

261   niches HIV-induced membraneless organelles (HIV-1 MLOs).

262   **HIV-induced membraneless organelles (HIV-1 MLOs) are sites of nuclear reverse**

263   **transcription.**

264   To better characterize the role of HIV-1 MLOs, we performed Immuno-RNA FISH at

265   different time post infection in THP-1 cells. We observed the majority of vRNAs

266   retained in CPSF6 clusters at 2 days p.i., while at 3 days p.i. and even more at 7 days

267   p.i. almost half of the total vRNA foci were excluded from CPSF6 clusters (Figure 5A).

268   Genomic vRNAs localize 100% with CPSF6 upon NEV treatment at all times p.i. (Fig.

269   5A), this suggests that the vRNA transcription foci locate outside CPSF6 clusters. What

270   is the functional meaning of the incoming vRNA sequestration in CPSF6 clusters? To
271   answer to this question we performed Chromatin Immunoprecipitation (ChIP) of the
272   endogenous CPSF6 protein in infected and uninfected THP-1 cells at 2 days p.i.,
273   followed by real-time PCR (Fig. 5B). Early, intermediate and late RT products were
274   amplified using appropriate primers. Of note, early transcripts were more abundant
275   than others, revealing an "ongoing" reverse transcription inside CPSF6 clusters. (Fig.
276   5B). Our results explain previous data showing EdU-labeled viral episomal forms
277   clustering in nuclear niches of macrophages (Francis et al., 2020; Rensen et al., 2020).
278   Importantly, the presence of the IN inside HIV-1 MLO is independent of the presence
279   of the reverse transcription, highlighting that the retrotranscribed DNA is not the driving
280   force of the HIV-1 MLOs formation (Fig 5C). Since CPSF6 participates to the viral
281   nuclear translocation and post-nuclear entry steps by interacting with the viral CA
282   (Bejarano et al., 2019; Burdick et al., 2017; Lee et al., 2010; Price et al., 2014) (Francis
283   et al., 2020), it might be that the viruses inside HIV-1 MLOs have not yet completed
284   the uncoating process.

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

287   CPSF6 phase separation does not randomly occur in the nuclear space. CPSF6
288   clusters fully colocalize with the marker of nuclear speckles (NSs), SC35 (Fig. 6 A-C).
289   Importantly, we remarked that CPSF6 and SC35 signals mainly occupy the
290   interchromatin space (Fig. 6A). NSs are, indeed, interchomatin granules (Spector and
291   Lamond, 2011). SC35 strongly associates to another speckle factor, SON,
292   independently of viral infection (Pearson's r ~0.995), indicating that CPSF6 clusters in
293   the host NS (Fig. 6B). Contrary to EdU labeling (Rensen et al., 2020), the HIV-1
294   ANCHOR system allows the visualization of a single provirus (Fig.1C) and does not
295   interfere with viral transcription (Blanco-Rodriguez et al., 2020). Thus, we asked where
296   in the nuclear space the viral expression takes place. We first observed that 3 days p.i.
297   almost all late reverse transcripts are excluded from CPSF6 clusters (Fig.6B), as
298   expected from the vDNA/IN separation data (Fig. 4A and B). We computed, as well,
299   the distances of the vDNAs from the boundary of the closest NS (avg ~ 0.6 μm) (Fig.
300   6B). Similar data were obtained upon integration inhibition (+RAL), where the majority
301   of episomal DNA forms locate outside CSPF6 clusters and in average ~ 0.4 μm from
302   SC35 (Fig. S5A). Taking into account that we found the late reverse transcribed DNA

303 located outside HIV-1 MLO, to understand the exact location of transcriptionally active

304 proviruses we used MCP-MS2 RNA labelling system. We observed that MCP-GFP is

305 unable to label the viral incoming RNA (Fig.6C). Under NEV treatment, MCP-GFP is

306 filtered out of HIV-1 MLOs as shown by the intensity profiles, whether in physiological

307 conditions it labels the viral transcription foci (Fig. 6C). Active proviruses are excluded

308 from CPSF6 clusters (Fig. 6D) and located in average at a distance from the closest

309 CS35 NS ~ 0.6 µm). These results are independent of the presence of vpx (Fig. S5B)

310 that we used to increase viral infectivity in THP-1 cells (Hrecka et al., 2011; Laguette

311 et al., 2011). Taken together, we display that both vDNA (episomal and proviral forms)

312 and vRNA foci (active proviruses) are excluded from HIV-1 MLOs (Fig. 6E), clearly

313 indicating that these structures cannot be sites for viral integration or replication. More

314 in detail, we found foci of viral transcription located in avg at a distance ~ 0.9 µm from

315 the closest HIV-1 MLO (CPSF6+SC35), while total vDNAs were found ~1.8 µm from

316 the nearest HIV-1 MLO (Fig. 6F). These data were corroborated by another vRNA

317 labelling approach, immuno-RNA FISH, coupled to HIV-1 ANCHOR to detect the

318 transcribing proviruses, indicating that viral integration and transcription occur outside

319 CPSF6 clusters (Fig. 6G). Of note, we confirmed the results obtained in THP-1 cells

320 (Fig. 5A) in primary monocyte-derived macrophages (MDMs), using an HIV-1 GFP

321 reporter virus (Fig. 7). Also in the MDMs we could appreciate the blockade of the

322 incoming vRNA in CPSF6, while the transcription foci locate outside these organelles

323 (Fig. 7). Finally, we were able to correlate the percentage of vRNA foci that locate

324 outside the MLOs per nucleus and the viral GFP expression (Pearson's r ~0.63),

325 showing that those are integration and transcription sites of HIV-1 (Fig.7).

326

327 **HIV-1 integration sites are marked by LEDGF.**

328 We extended our study on THP-1 cells at 3 days p.i. to investigate whether Lens

329 Epithelium-Derived Growth Factor (LEDGF), which is a partner of the viral IN and a

330 HIV-1 integration sites selection factor (Cherepanov et al., 2003; Ciuffi et al., 2005;

331 Emiliani et al., 2005; Ferris et al., 2010; Hare et al., 2009; Lelek et al., 2015; Llano et

332 al., 2006; Shun et al., 2008; Shun et al., 2007; Wang et al., 2014), is associated to the

333 proviruses forming a complex visible by imaging. We observed that CPSF6 proteins

334 cluster in proximity of LEDGF sites (~ 43%), without co-localizing (Fig. 8A). There is a

335 similar number of LEDGF clusters in both infected and uninfected cells (Fig. 8A), while

336 CPSF6 clusters are mainly formed upon infection, suggesting that CSPF6 is prompted

337 by the virus to localize near the pre-existing LEDGF regions (Fig.8A). Of note, the

338 PWWP domain of LEDGF has been shown to interact with H3K36me3 (Pradeepa et

339 al., 2014; Pradeepa et al., 2012), which is a marker of euchromatin. In particular, this

340 post-translational modification is enriched in HIV-1 integration sites (Ciuffi et al., 2005;

341 Lelek et al., 2015; Vansant et al., 2020; Wang et al., 2009). Therefore, we analysed

342 the location of proviral foci of transcription in relation to LEDGF. In average ~64%

343 vDNA and ~73% vRNA foci per nucleus were found in complex with LEDGF (Fig. 8B).

344 Importantly, LEDGF is also known to interact with splicing factors (Singh et al., 2015).

345 This is in agreement with our data showing vDNA and vRNA localized in avg ~ 0.6 μm

346 from the closest NS (Fig. 6B and D). Next we have characterized in more detail the

347 viral genome associations with LEDGF with respect to CPSF6. We calculated that ~

348 87.9% of the vDNA/LEDGF complexes and ~80% of vRNA foci/LEDGF complexes are

349 uncoupled to CPSF6 (Figure 8B). We found only few vDNA (~12.1%) and vRNA

350 (~20%) foci associated to LEDGF clusters near to CPSF6 (Fig. 8B). Our results

351 suggest that the two cellular factors, CPSF6 and LEDGF, orchestrate two sequential

352 steps of viral replication, the maturation of the PIC and its integration, respectively (Fig.

353 9). Functional proviruses are not located inside HIV-1 MLOs, but in external active

354 chromatin domains, enriched in LEDGF (Fig. 8B) and in proximity of SC35 NS (Fig.

355 6D), likely to benefit of a favorable environment for viral replication.

356

357 **Discussion**

366 Studies of parasites that permanently integrate into the host genome show that they

367 have evolved to target chromatin in such a way as to optimize their coexistence with

368 the host and favor the release of new viral progeny (Bushman, 2003; Craig and

369 Marszalek, 2002). HIV-1 needs to rapidly release high levels of newly generated

370 viruses after infection, intuitively explaining the preference of active genes as

371 integration targets (Ciuffi et al., 2005; Di Nunzio et al., 2013; Lelek et al., 2015;

372 Maldarelli, 2016; Schroder et al., 2002). However, the nuclear location of actively

373 transcribing individual proviruses in dividing or non-dividing cells has never been

374 visualized before. Thus far, the direct detection of HIV-1 genomes after reverse

375 transcription and the distribution of proviral transcriptional foci has been technically

376 challenging. HIV-1 ANCHOR system allows the live tracking of different nuclear late

377 retrotranscribed DNA forms (Fig.1A-C). We observed a homogenous population

378   formed by proviral DNA with low diffusion coefficient. Episomal forms showed a

379   heterogeneous behavior, characterized by two different populations, implying that

380   these two viral forms can interact with different factors in the nuclear space (Fig. 1D).

381   Moreover, imaging analysis showed that viral episomal forms identified by the

382   accumulation of OR-GFP on ANCH3 sequences gave rise to brighter signals than

383   proviral forms in HeLa cells (Fig. 2A). This phenomenon could explain the differences

384   in behavior between integrated and non-integrated forms (Fig.1D). Intriguingly, we

385   found that episomal vDNAs form clusters, which could result in a brighter signal than

386   integrated forms (Fig. S2C, video 4). Next, to track the nuclear location of functional

387   proviruses, we co-labeled vDNA and vRNA by combining HIV-1 ANCHOR technology

388   with RNA FISH or with MCP-MS2 system (Tantale et al., 2016) in the main target cells

389   of HIV-1 (Fig.2A-D, Fig.S2A and B). Therefore, we investigated the spatial location of

390   proviruses in dividing and non-dividing cells. Three-dimensional analysis of the

391   distribution of vDNAs during early (48 h p.i.) infection in mitotic cells displays that they

392   are more randomly distributed than the proviral DNA from cells at 7days p.i. (Fig. 3A).

393   We also observed that the location of viral transcription sites (vRNA) is highly similar

394   between early and late time points of infection, indicating that in HeLa cells the proviral

395   DNAs, the only transcribing forms (Fig. 2A), immediately localize to near the NE early

396   after infection (Fig. 3A). This vicinity to the NE was detected as well in non-dividing

397   cells (Fig. 3B), highlighting that in both mitotic and non-mitotic cells the proviruses

398   locate in the nuclear periphery (Fig.3C). We next investigated the interplay among viral

399   components and their association with host factors during post-nuclear entry steps. IN

400   proteins strongly co-localize with CPSF6 clusters in macrophage-like cells and in CD4[+]

401   T lymphocytes at early time of infection (Fig. 4C), in agreement with another study

402   (Francis et al., 2020). Contrary to CPSF6, the late vDNA products of retrotranscription

403   rarely co-localize with IN in both types of cells (Fig. 4A). However, in live imaging we

404   pinpointed the budding of newly synthesized late reverse transcribed DNA from the IN

405   clusters, which occurs within minutes (Fig.4B, video 6,7).These results are in line with

406   the dynamic coexistence of the IN and vDNA, which remain separated but in the

407   proximity for several hours, as it has been shown by real-time imaging in HeLa cells

408   (Fig. S3, Table S1, Table S2, video 5). This result can appear in contrast with recent

409   data from other groups and from us obtained with another DNA labeling technology,

410   EdU (Francis et al., 2020; Rensen et al., 2020), showing the vDNA and the IN strongly

411   colocalized with CPSF6. This apparent discrepancy can be explained by the intrinsic

412   features of vDNA labeling tools. While EdU labeling cannot provide direct information
413   on viral replication, as EdU reduces viral transcription (Rensen et al., 2020), HIV-1
414   ANCHOR system allows the detection of late reverse transcripts (Fig.1A) that can
415   generate functional proviral DNA, as we observed in HeLa cells, human lymphocytes,
416   and macrophages (Fig. 2A-D). Thus, we asked whether CPSF6 clusters are
417   physiological sites of nuclear RT and of viral replication. First, we performed ChIP of
418   the endogenous CPSF6 followed by real time PCR in THP-1 cells infected or
419   uninfected (48 h p.i.). The PCR results show an abundance of early and intermediate
420   RT products in comparison to late RT products, indicating an ongoing RT process in
421   CPSF6 sites. These results are in agreement with a detection of the nascent vDNA
422   (Fig. 4B) and with the labeling of the vDNA in these sites (Rensen et al., 2020). Second,
423   we were able to correlate the progression of the viral infection with the nuclear location
424   of the vRNA foci (Fig. 5A). At early time p.i. almost all vRNA are sequestered in CPSF6
425   clusters to be reverse transcribed (Fig.5A). Concurrently with the advancement of the
426   viral infection we observed an increase in vRNA foci outside CPSF6 clusters, these
427   foci are never present in cells treated with NEV where the viral replication is completely
428   blocked (Fig. 5A). Importantly, our results show a positive correlation between the
429   number of vRNA foci excluded from CPSF6 clusters and the GFP expression (HIV-1
430   reporter gene) (Fig. 7). Thus, HIV-1 induces the formation of novel membraneless
431   organelles (MLOs) retaining CPSF6 proteins and viral components, such as IN and
432   viral RNA genomes to host viral reverse transcription (Fig. 4C and D, 5A and B). HIV-
433   1 MLOs form in the interchromatin space, remodeling the host nuclear speckles,
434   characterized by the presence of SC35 and SON (Fig.6A and B). Likely, CPSF6
435   relocates to speckles prompted by the virus to usurp functions linked to these MLOs.
436   We show that the formation of CPSF6 cluster is independent of viral DNA integration
437   or viral DNA synthesis (Fig. S5A, Fig. 5C) and that late reverse transcribed products
438   are mainly excluded from them (Fig. 6B and E, Fig.7). These results are in agreement
439   with a rapid exclusion of late retrotranscribed DNAs from the IN proteins that are
440   retained in CPSF6 clusters (Fig. 4A-C). Next, we calculated the distribution of the
441   actively transcribing proviruses in the nuclear landscape. Interestingly, we observed
442   that the MCP-MS2 system detects only the vRNA associated with viral replication sites,
443   as MCP-GFP is unable to penetrate HIV-1 MLOs that shield the incoming vRNA
444   genome, acting as selective filters (Fig. 6C). When we compared the distances from
445   the boundary of the closest HIV-1 MLOs (CPSF6 + SC35), we observed that the vRNA

446   transcription foci were closer than the vDNAs, indicating that the HIV-1 MLOs-
447   neighbouring regions are the hub of viral replication (Fig. 6 F-G). We also observed
448   that CPSF6 relocates to the vicinity of LEDGF during HIV infection (Fig. 8A). LEDGF
449   associates to highly-spliced transcriptional units through its direct interaction with
450   several splicing factors (Singh et al., 2015) and determines HIV-1 integration sites
451   distribution in active chromatin sites (Ciuffi et al., 2005; Ferris et al., 2010; Shun et al.,
452   2007). However, the visualization of a complex formed by the transcribing provirus and
453   LEDGF remained utopian for a long time. Here, we were able to visualize by
454   fluorescence microscopy a complex formed by the vDNA or vRNA and LEDGF. Almost
455   80% of those complexes are formed by LEDGF uncoupled to CPSF6 (Fig. 8B).
456   Nonetheless, we observed that less than 20% of them can include CPSF6 associated
457   to LEDGF. Whether there is a functional difference between these two complexes it
458   can be further investigated in future studies. Overall, our study highlights the proviral
459   location in the nuclear landscape remodelled by the viral infection (Fig. 9). We found
460   that HIV-1-MLOs sequester the viral IN (Fig. 4C, Fig.5C), as well viral RNA genomes
461   (Fig. 5A) and rapidly release the late retrotranscribed DNA, acting as HIV-1 reverse
462   transcription and PIC maturation sites (Fig.4B, Fig. 5A and B). Thus, HIV-1 MLOs play
463   a role prior to the integration step. Taken together our results show that functional
464   proviruses establish near to the NE (Fig. 3C), excluded from CPSF6 clusters, but
465   locate in a favorable transcriptional environment (Fig. 6D, Fig.8B) at less than 1 μm
466   from HIV-1 MLOs (Fig.6F). Our study supports how new single-cell level approaches
467   are pivotal in the study of functional viral nuclear dynamics (Fig.9), to discriminate the
468   cells susceptible to fuel viremia or that concur to viral persistence.

469

Materials and Methods

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

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

478 immortalized monocyte cells derived from acute monocytic leukemia, which, once

479 seeded, were differentiated into macrophage-like cells under Phorbol 12-myristate 13-

480 acetate (PMA) treatment (167nM). SupT1R5 cells are T lymphoblastic lymphoma cells

481 which were transduced to stably express CCR5 (kind gift of N. Casartelli). HEK293T

482 cells and HeLa cells were cultivated in DMEM medium supplemented with 10% Fetal

483 Bovine Serum (FBS) and 1% penicillin-streptomycin. Jurkat, SupT1R5 and THP-1 cells

484 were cultivated in RPMI 1640 medium supplemented with 10% Fetal Bovine Serum

485 (FBS) and 1% penicillin-streptomycin solution. Primary $CD4^+T$ and MDM cells were

486 purified from healthy donors' blood obtained via EFS (Etablissement Français du Sang,

487 Paris), through density gradient centrifugation with Ficoll 400 and isolation of PBMCs

488 buffy coat. $CD4^+T$ cells positive selection was performed with human CD4 Microbeads

489 (Miltenyi Biotec #130-045-101). The day after $CD4^+T$ cells were activated with T cell

490 Activation/Expansion kit (Miltenyi Biotec #130-091-441) according to the datasheet,

491 complete RPMI 1640 medium was supplemented with interleuykin-2 (IL-2) (400

492 IU/mL). For MDMs the PBMCs were incubated at 37°C for 2 hours, then, non-adherent

493 cells were washed away and complete RPMI 1640 medium was supplemented with

494 M-CSF (10 ng/mL). After 3 days, the medium was changed with RPMI without M-CSF

495 and cells were left to differentiate for other 4 days before proceeding with infection.

496 HIV-1 ANCH3 single-cell clone was generated starting from HeLa P4R5 cells which

497 were infected with HIV-1 ANCH3 virus at MOI 1. Cells were diluted to 1 cell per well,

498 in 96-well plates. Cell-clone colonies were tested for β-galactosidase expression to

499 check viral expression (kit Roche #11758241001). Positive clones were transduced

500 with lentiviral vector (LV) CMV-OR-GFP for imaging analyses of HIV-1 ANCH3

501 provirus. All cells were kept in incubator at 37°C and 5% CO_2.

502

503 **Plasmids, lentiviral vectors and viral productions**

504 HIV-1ΔEnvIN$_{HA}$ΔNef plasmid encodes for the ΔEnvHIV-1 LAI (BRU) viral genome

505 where the Integrase (IN) protein is fused to the hemagglutinin (HA) tag (kindly provided

506 by F. Mammano)(Petit et al., 1999; Petit et al., 2000). HIV-1ΔEnvIN$_{HA}$ΔNef ANCH3

507 (HIV-1 ANCH3) was obtained through ANCH3 insertion: ANCH3 sequence was cloned

508 by PCR using the template plasmid pANCH3 as we previously described in Blanco et

509 al. (Blanco-Rodriguez et al., 2020). The ANCHOR™ technology and sequences are

510 exclusive property of NeoVirTech (Germier et al., 2017; Mariame et al., 2018). HIV-

511 1ΔEnv IN$_{HA}$(D116A) ΔNef ANCH3 was obtained by insertional mutagenesis using the

512 QuikChange II XL Site-Directed Mutagenesis kit (Agilent #200522), for integrase (IN)

513 mutation at D116. The lentiviral vector plasmids CMV OR-GFP and CMV OR-

514 SANTAKA were obtained by cloning OR-GFP/OR-SANTAKA (plamids from

515 NeoVirTech) in a pTripCMV (ΔU3) plasmid through PCR and cloned using restriction

516 site, AgeI and SgrDI. HIV-1ΔEnvIN$_{HA}$ΔNef ANCH3 MS2 plasmid was obtained by

517 inserting the MS2x64 sequence amplified by PCR from pMK123-MS2x64 (Tantale et

518 al., 2016) in HIV-1ΔEnvIN$_{HA}$ΔNef ANCH3. Lentiviral vectors and HIV-1 viruses were

519 produced by transient transfection of HEK293T cells through calcium chloride

520 coprecipitation. Co-transfection was performed as following, LV carrying OR-GFP/OR-

521 SANTAKA: 10 µg of OR-GFP/OR-SANTAKA LV, 10 µg of NDK plasmid (gag-pol-tat-

522 rev) and 2.5 µg of pHCMV-VSV-G envelope expression plasmid; LV carrying MCP-

523 GFP (kindly provided by E. Bertand): 10 µg of MCP-GFP, 10 µg of NDK plasmid (gag-

524 pol-tat-rev) and 2.5 µg of pHCMV-VSV-G envelope expression plasmid; HIV-1ΔEnv

525 pseudotyped with VSV-G: 10 µg HIV-1ΔEnvIN$_{HA}$ΔNef (IN WT or IN(D116A)) plasmid

526 and 2.5 ug of pHCMV-VSV-G plasmid. VSV-G/HIV-1ΔEnv IN$_{HA}$(D116A) ΔNef ANCH3

527 has been also produced in combination with GIR (Gag-IN-Ruby plasmid, kindly

528 provided by E. Campbell) (Dharan et al., 2016; Hulme et al., 2015). For THP-1

529 infection, HIV-1 ANCH3 or HIV-1 ANCH3 MS2 viruses and LV OR-GFP or LV MCP-

530 GFP were all produced adding 3 µg of SIV$_{MAC}$ Vpx (Durand et al., 2013) at the moment

531 of HEK293T transfection. MDMs were infected with a virus HIV-1 ΔEnv Δnef GFP

532 complemented with Vpx (Rensen et al., 2021). After the collection of the supernatant

533 48h post-transfection, lentiviral particles were concentrated by ultracentrifugation for 1

534 h at 22000 rpm at 4°C and stored at −80°C. Lentiviral vectors and viruses were tittered

535 through late reverse transcripts (LTR) qPCR in HEK293T cells.

536

537 **Quantitative PCR and primers**

538 Total DNA of 10^6 infected cells was extracted at 6 hours post infection for Late Reverse

539 Transcripts (LRTs) qPCR and 24 hours post infection for 2LTR-circles and ALU-PCR,

540 through QIAmp DNA micro kit (QIAGEN #56304). Real-Time PCR of LRTs was

541 performed to assess DNA synthesis and used to normalize qPCR data on viral input,

542 the reactions were carried on in 20 µL, in iTaqUniversal SYBR Green Supermix (Bio-

543 Rad #1725124) using primers for U3 sequence: U3 FX: 5'-

544 TTCCGCTGGGGACTTTCCAGGG-3', U3 RX: 5'-AGGCTCAGATCTGGTCTAACC-3'.

545 Real-Time PCR of 2LTR-circles was used to asses nuclear import efficiency.
546 Reactions were performed in 20 µL, in Maxima Probe/ROX qPCR Mastermix
547 (ThermoFisher #K0232) using primers for 2LTR-circle junction: 2LTR FX: 5'-
548 AACTAGGGAACCCACTGCTTAAG-3', 2LTR RX: 5′-
549 TCCACAGATCAAGGATATCTTGTC-3′, 2-LTR probe: 5'-(FAM)-
550 ACACTACTTGAAGCACTCAAG-GCAAGCTTT-(TAMRA)-3'. Real-Time PCR of
551 proviral integrations consisting in a first non-kinetic PCR in 50 µL in Platinum SuperFi
552 DNA Polymerase (ThermoFisher #12351250) and in a second step of qPCR reaction
553 in 20 µL in Maxima Probe/ROX qPCR Mastermix (ThermoFisher #K0232). Primer Alu
554 166: 5'-TCCCAGCTACTCGGGAGGCTGAGG-3', Alu 2: 5'-
555 GCCTCCCAAAGTGCTGGGATTACAG-3', LambdaU3: 5′-
556 ATGCCACGTAAGCGAAACTTTCCGCTGGGGACTTTCCAGGG-3′ for ALU PCR;
557 Lambda: 5'-ATGCCACGTAAGCGAAACT-3', U5: 5'-CTGACTAAAAGGGTCTGAGG-
558 3', Probe: 5'-(FAM)- TTAAGCCTCAATAAAGCTTGCCTTGAGTGC-(TAMRA) for ALU
559 qPCR (Di Nunzio et al., 2013). In all experiments β-actin detection was used for
560 normalization. β-actin FX: 5'-AACACCCCAGCCATGTACGT-3', β-actin RX: 5-
561 CGGTGAGGATCTTCATGAGGTAGT-3', β-actin probe: (FAM)-
562 CCAGCCAGGTCCAGACGCAGGA-(BHQ1).

563

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

578   seeded on poly-L-lysine (Sigma #P4707) coated coverslips and shortly centrifuged to

579   allow attachment. The day of fixation all cells were washed with PBS and fixed with

580   4% PFA for 15 minutes. For protein staining, cells were treated with glycine 0.15% for

581   10 min, permeabilized with 0.5% Triton X-100 for 30 min and blocked with 1% bovine

582   serum albumin (BSA) for 30 min. All incubations were carried out at room temperature,

583   in the dark, in humid chamber, 1h with primary antibodies and 45 min with secondary

584   antibodies. Washes between antibody incubations and antibodies dilution were done

585   in 1% BSA. Primary antibodies were diluted as follows: anti-HA 1:500 (Roche

586   #11867423001), anti-HIV-1 p24 (CA) 1:400 (NIH #3537), anti-CPSF6 1:400 (Novus

587   Biologicals #NBP1-85676), anti-SC35 1:200 (Abcam #ab11826), anti-LEDGF 1:200

588   (BD Bioscience #611715). Secondary antibodies used were the following: Goat anti-

589   Mouse Alexa-647 1:300 (1:100 for CA) (Invitrogen #A21235); Donkey anti-Rat Alexa-

590   488 1:100 (Invitrogen# A21208) for IN-HA or Goat anti-Rat Alexa-647 1:300 (Invitrogen

591   #A21247), 1:100 for IN-HA; Donkey anti-Rabbit Cy3 1:500 (Jackson Lab #711-165-

592   152). Finally, cells were stained with Hoechst 33342 1:5000 (Invitrogen #H3570) for 5

593   minutes. Coverslips were mounted on glass slides (Star Frost) with Prolong Diamond

594   Antifade Mountant (Life Technologies #P36970). Confocal microscopy was carried out

595   with a Zeiss LSM700 inverted microscope, with a 63X objective (Plan Apochromat, oil

596   immersion, NA=1.4).

597

598   **RNA FISH**

599   For RNA FISH studies in adherent cell-lines, the cells were seeded on coverslips. HeLa

600   P4R5 cells expressing OR-GFP were infected with HIV-1 ANCH3 (MOI 30) with or

601   without RAL (20 μM) or infected with HIV-1 ANCH3 (MOI 5). Jurkat cells expressing

602   OR-GFP were infected with HIV-1 ANCH3 (MOI 5). Primary activated CD4$^+$T cells, 3

603   days after activation, were together transduced with OR-GFP LV (MOI 10) and infected

604   with HIV-1 ANCH3 (MOI 50). THP-1 cells were differentiated with PMA for 48h and,

605   after 48h transduction with LV OR-GFP Vpx (MOI 0.5), the cells were infected with

606   HIV-1 ANCH3 Vpx (MOI 20). Primary MDMs were seeded before differentiation on

607   coverslips, and after 7 days, transduced with LV OR-GFP Vpx (MOI 10) and infected

608   with HIV-1 ANCH3 Vpx (MOI 40). The day of fixation suspension cells were seeded on

609   poly-L-lysine (Sigma #P4707) coated coverslips and centrifuged to allow attachment.

610   Then, cells were washed with PBS and fixed with 4% PFA for 15 minutes and

611   incubated in 70% ethanol at -20°C at least for one night. Primary smiFISH probes have
612   been designed against HIV-1 *pol* transcript and a shared readout sequence for
613   secondary probe alignment. Twenty-four smiFISH probes (Table S3) against HIV pol
614   were designed with Oligostan (Tsanov et al., 2016) and purchased from Integrated
615   DNA Technologies (IDT). Primary probes were pre-hybridized with a secondary FLAP
616   probe conjugated to Cy5 fluorophore through pairing with the readout sequence
617   (Eurofins). Washes and hybridization were performed with Stellaris Buffers (WASH
618   buffer A, WASH buffer B, Hybridization Buffer; LGC Biosearch Technologies),
619   following the manufacturer protocol. Hybridization with the probe was carried out at
620   37°C in a dark humid chamber for 5 hours. Finally, cells were stained with Hoechst
621   33342 1:5000 (Invitrogen #H3570) for 5 minutes. Coverslips were mounted on glass
622   slides (Star Frost) with Prolong Diamond Antifade Mountant (Life Technologies
623   #P36970). Confocal microscopy was carried out with a Zeiss inverted LSM700
624   microscope, with a 63X objective (Plan Apochromat, oil immersion, NA=1.4).

625   **IMMUNO-RNA FISH**

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

642   **Time-lapse microscopy**

643   For all live imaging studies, 10^5 cells were plated on a polymer-coverslip bottom μ-Dish
644   35 mm (ibidi #81156) at least 24 hours before 2D or 3D video acquisition with an
645   UltraView VOX inverted Microscope (Perkin-Elmer), based on a CSU-X spinning-disk
646   (Yokogawa), and using a 63X objective (Plan Apochromat, oil immersion, NA=1.4). For
647   integrated provirus imaging, HIV-1 ANCH3 single-cell clone was transduced with OR-
648   GFP LV (MOI 0.2) for 3 days and then plated on the polymer-bottom dish the day
649   before acquisition. On the other hand, HeLa P4R5 cells were infected with HIV-1
650   ANCH3 (MOI 10) and five days post infection cells were plated on the polymer-
651   coverslip bottom dish and transduced with LV OR-GFP for 24h. For the acquisitions of
652   the episomal viral DNA, HeLa P4R5 were firstly transduced with OR-GFP LV (MOI 0.2)
653   and the day after cells were plated on the polymer-bottom dish and infected with HIV-
654   1 ANCH3 (MOI 10) with RAL 20 μM or with the HIV-1 ANCH3 IN_{D116A}. Cells were
655   imaged in 3D in continue for 10 minutes. Experiments of co-live tracking of IN-Ruby
656   (GIR) and vDNA (OR-GFP) were performed in HeLa P4R5 cells transduced with LV
657   OR-GFP (MOI 0.2). The following day, cells were infected with HIV-1 ANCH3 IN_{D116A}
658   complemented with the GIR plasmid using an MOI 50. Frames were acquired every
659   30 seconds in 2D video (Fig. S3, video 5) and every minute for 3D videos (Fig. 4B,
660   video 6,7), at different times post infection. For co-live-imaging of vDNA (OR-
661   SANTAKA) and vRNA foci (MCP-GFP), HeLa MCP-GFP cells were transduced with
662   OR-SANTAKA LV (MOI 1) and 24 h later infected with HIV-1 ANCH3 MS2 (MOI 50).
663

Chromatin Immunoprecipitation with qPCR

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

677 washed, eluted from the beads with SDS buffer, and subjected to RNase and

678 proteinase K treatment. Crosslinks were reversed by incubation overnight at 65°C, and

679 ChIP DNA was purified by phenol-chloroform extraction and ethanol precipitation.

680 Quantitative PCR (qPCR) reactions were carried out in triplicate on specific genomic

681 regions using SYBR Green SuperMix (BioRad). The resulting signals were normalized

682 for primer efficiency by carrying out qPCR for each primer pair using Input DNA.

683 Different RT products were amplified using the following primers, early RT products:

684 5′-GCCTCAATAAAGCTTGCCTTGA-3′ and 5′-TGACTAAAAGGGTCTGAGGGATCT-

685 3′; intermediate RT products: 5′ -CTAGAACGATTCGCAGTTAATCCT-3′ and 5′ -CTAT

686 CCTTTGAT GCACACAATAGAG-3′; late RT products: 5′-

687 TGTGTGCCCGTCTGTTGTGT-3′ and 5′-GAGTCCTGCGTCGAGAGAGC-3′. De-

688 crosslinked ChIP products were checked by western blot against CPSF6, loading 37.5

689 ng of DNA per sample.

690

691 **Imaging Analysis and Statistics**

692 All data were analyzed in GraphPad Prism 8 (GraphPad Software, La Jolla California

693 USA, www.graphpad.com). All images and videos were analyzed in Fiji (Schindelin et

694 al., 2012) and Icy version 2.0.3.0 (de Chaumont et al., 2012) software. Here below the

695 custom-made protocols:

696 *Particle tracking and MSD measurement*

697 3D videos were analyzed in Fiji (Schindelin et al., 2012) and viral nuclear particles

698 were tracked along time in 3D using TrackMate plugin (Tinevez et al., 2017). The

699 resulting tracks were exported in MATLAB (The MathWorks, Natick, USA) and then

700 used for mean-squared displacement (MSD) analysis with @msdanalyzer(Tarantino

701 et al., 2014). MSD curves were fitted by a straight line from which we derived the value

702 of the diffusion coefficient: $D = a / (2*nDim)$, where « a » is the slope of the MSD curve

703 and « nDim » the dimensionality of the image. MSD curves for which the R^2 value was

704 lower than 0.8 were discarded from analysis. Only the initial 25% of the MSD curves

705 were used for the linear fit.

706 *vDNA spot intensity analysis*

707 The 2D confocal images were analysed with Icy software version 2.0.3.0 (de

708 Chaumont et al., 2012) using a custom made protocol

709 (http://icy.bioimageanalysis.org/protocol/batch-spots-detections-and-statistics and

710 doi: 10.5281/zenodo.3925262, example of HeLa confocal image available at doi:

711   10.5281/zenodo.3925213). vDNA spots were detected extracting the green channel

712   and applying first a median filter of half size = 1 and then 3, the Spot Detector with a

713   scale of 1 (sensitivity to 55) on the nuclear ROI only. Data from the spot detection (e.g.

714   mean fluorescence intensity) were exported for analysis. vRNA mean fluorescence

715   intensity signal was computed from the nuclei ROI.

716   *vDNA-vRNA association analysis*

717   For vDNA - vRNA association studies in Jurkat cells, 3D confocal images were

718   analyzed in Icy version 2.0.3.0 using two successive custom made protocols (first

719   protocol:http://icy.bioimageanalysis.org/protocol/nuclei-segmentation-from-maximum-

720   intensity-projections/   with doi: 10.5281/zenodo.3925181) and second protocol:

721   http://icy.bioimageanalysis.org/protocol/spots-detection-and-colocalisation-analysis-

722   with-soda/ with doi: 10.5281/zenodo.3925191, example of Jurkat confocal image

723   available at doi: 10.5281/zenodo.3925213). Nuclei were segmented on a maximum

724   intensity projection of the Hoechst channel using the Icy block Gaussian filter with a

725   radius of sigma x=5, y=5 and z=5 pixels, the Thresholder block with a manual threshold

726   of 5000 and the Label Extractor block, which uses connected components to separate

727   the nuclei. Cy5 spots were detected with the HK Means block (Dufour et al.,

728   doi.org/10.1109/ICPR.2008.4761748).The logical operation block was used to keep

729   only the spots detected in the nucleus. These spots were added as input parameter

730   "List detection 1" of the SODA block. EGFP spots were detected using the log3D block

731   to enhance the spots, inverting the log transformed image using the Math Operation

732   Expression block with the following expression: A*B with A scalar=-1 and B the log

733   transformed image, detecting the spots with the Local Extrema block and segmenting

734   the spots with a dilation and a connected components operations. The logical operation

735   block was used to keep only the spots detected in the nucleus. These spots were

736   added as input parameter "List detection 2" of the SODA block. The nucleus ROI was

737   kept as input parameter for "ROI of Analysis (Cell's shape)". The Cy5 and GFP spots

738   ROIs were then imported in the SODA block (Lagache et al., 2018) to analyse

739   colocalization of spots within the nucleus only with a maximum radius of 10, a step of

740   1 and no fixed search distance.

741   *vDNA/vRNA nuclear envelope distance analysis*

742   The nuclei were segmented on the Hoechst channel using background subtraction

743   followed by a multilevel Otsu thresholding. The resulting nuclei masks were then

744   directly applied on the vDNA and vRNA channels to select the relevant 3D information

745 inside each labelled region. Inside these regions, vRNA spots were identified using 3D

746 non-local means denoising (Buades et al., "Non-local means denoising." *Image*

747 *Processing On Line* 1 (2011): 208-212) and automated segmentation, while vDNA

748 spots were extracted with the image analysis pipeline previously described (Komatsu

749 et al., 2018). For each vDNA and vRNA spot in was computed the distance to the

750 nuclear envelope in the 2D z-plane corresponding to the spot centroid location.

751 *SC35/HIV-1 MLO distances analysis*

752 The 3D confocal images were processed for multi-channel image splitting, vDNA and

753 SC35 segmentation. The automated 3D segmentation of cell nuclei included a

754 preliminary non-local means denoising step to cope with the strong signal

755 heterogeneity and the segmentation of SC35 was readjusted to low contrast

756 conditions. The 3D boundary-to-boundary distance (rather than centroid-to-centroid

757 which was done for vDNA/vRNA-NE analysis) between each vDNA or vRNA spot and

758 its closest SC35 speckle or the closest HIV-1 MLO (SC35+CPSF6) were computed.

759

771

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

773

774 **References**

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

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

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

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

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

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

790 Buffone, C., A. Martinez-Lopez, T. Fricke, S. Opp, M. Severgnini, I. Cifola, L. Petiti, S. Frabetti, K.
791   Skorupka, K.K. Zadrozny, B.K. Ganser-Pornillos, O. Pornillos, F. Di Nunzio, and F. Diaz-Griffero.
792   2018. Nup153 Unlocks the Nuclear Pore Complex for HIV-1 Nuclear Translocation in
793   Nondividing Cells. *J Virol*. 92.

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

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

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

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

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

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

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

810 Chin, C.R., J.M. Perreira, G. Savidis, J.M. Portmann, A.M. Aker, E.M. Feeley, M.C. Smith, and A.L. Brass.
811   2015. Direct Visualization of HIV-1 Replication Intermediates Shows that Capsid and CPSF6
812   Modulate HIV-1 Intra-nuclear Invasion and Integration. *Cell Rep*. 13:1717-1731.

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

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

817 de Chaumont, F., S. Dallongeville, N. Chenouard, N. Herve, S. Pop, T. Provoost, V. Meas-Yedid, P.
818   Pankajakshan, T. Lecomte, Y. Le Montagner, T. Lagache, A. Dufour, and J.C. Olivo-Marin. 2012.
819   Icy: an open bioimage informatics platform for extended reproducible research. *Nat Methods*.
820   9:690-696.

821 De Wit, F., S.R. Pillalamarri, A. Sebastian-Martin, A. Venkatesham, A. Van Aerschot, and Z. Debyser.
822   2019. Design of reverse transcriptase-specific nucleosides to visualize early steps of HIV-1
823   replication by click labeling. *J Biol Chem*. 294:11863-11875.

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

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

829 Di Nunzio, F., T. Fricke, A. Miccio, J.C. Valle-Casuso, P. Perez, P. Souque, E. Rizzi, M. Severgnini, F.
830   Mavilio, P. Charneau, and F. Diaz-Griffero. 2013. Nup153 and Nup98 bind the HIV-1 core and
831   contribute to the early steps of HIV-1 replication. *Virology*. 440:8-18.
832 Di Primio, C., V. Quercioli, A. Allouch, R. Gijsbers, F. Christ, Z. Debyser, D. Arosio, and A. Cereseto. 2013.
833   Single-cell imaging of HIV-1 provirus (SCIP). *Proc Natl Acad Sci U S A*. 110:5636-5641.
834 Durand, S., X.N. Nguyen, J. Turpin, S. Cordeil, N. Nazaret, S. Croze, R. Mahieux, J. Lachuer, C. Legras-
835   Lachuer, and A. Cimarelli. 2013. Tailored HIV-1 vectors for genetic modification of primary
836   human dendritic cells and monocytes. *J Virol*. 87:234-242.
837 Emiliani, S., A. Mousnier, K. Busschots, M. Maroun, B. Van Maele, D. Tempe, L. Vandekerckhove, F.
838   Moisant, L. Ben-Slama, M. Witvrouw, F. Christ, J.C. Rain, C. Dargemont, Z. Debyser, and R.
839   Benarous. 2005. Integrase mutants defective for interaction with LEDGF/p75 are impaired in
840   chromosome tethering and HIV-1 replication. *J Biol Chem*. 280:25517-25523.
841 Farnet, C.M., and W.A. Haseltine. 1991. Determination of viral proteins present in the human
842   immunodeficiency virus type 1 preintegration complex. *J Virol*. 65:1910-1915.
843 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.
844   Engelman, and S.H. Hughes. 2010. Lens epithelium-derived growth factor fusion proteins
845   redirect HIV-1 DNA integration. *Proc Natl Acad Sci U S A*. 107:3135-3140.
846 Fox, A.H., Y.W. Lam, A.K. Leung, C.E. Lyon, J. Andersen, M. Mann, and A.I. Lamond. 2002. Paraspeckles:
847   a novel nuclear domain. *Curr Biol*. 12:13-25.
848 Francis, A.C., C. Di Primio, V. Quercioli, P. Valentini, A. Boll, G. Girelli, F. Demichelis, D. Arosio, and A.
849   Cereseto. 2014. Second generation imaging of nuclear/cytoplasmic HIV-1 complexes. *AIDS Res
850   Hum Retroviruses*. 30:717-726.
851 Francis, A.C., M. Marin, J. Shi, C. Aiken, and G.B. Melikyan. 2016. Time-Resolved Imaging of Single HIV-
852   1 Uncoating In Vitro and in Living Cells. *PLoS Pathog*. 12:e1005709.
853 Francis, A.C., M. Marin, P.K. Singh, V. Achuthan, M.J. Prellberg, K. Palermino-Rowland, S. Lan, P.R.
854   Tedbury, S.G. Sarafianos, A.N. Engelman, and G.B. Melikyan. 2020. HIV-1 replication
855   complexes accumulate in nuclear speckles and integrate into speckle-associated genomic
856   domains. *Nat Commun*. 11:3505.
857 Francis, A.C., and G.B. Melikyan. 2018. Single HIV-1 Imaging Reveals Progression of Infection through
858   CA-Dependent Steps of Docking at the Nuclear Pore, Uncoating, and Nuclear Transport. *Cell
859   Host Microbe*. 23:536-548 e536.
860 Geis, F.K., and S.P. Goff. 2019. Unintegrated HIV-1 DNAs are loaded with core and linker histones and
861   transcriptionally silenced. *Proc Natl Acad Sci U S A*. 116:23735-23742.
862 Germier, T., S. Kocanova, N. Walther, A. Bancaud, H.A. Shaban, H. Sellou, A.Z. Politi, J. Ellenberg, F.
863   Gallardo, and K. Bystricky. 2017. Real-Time Imaging of a Single Gene Reveals Transcription-
864   Initiated Local Confinement. *Biophys J*. 113:1383-1394.
865 Graham, T.G., X. Wang, D. Song, C.M. Etson, A.M. van Oijen, D.Z. Rudner, and J.J. Loparo. 2014. ParB
866   spreading requires DNA bridging. *Genes Dev*. 28:1228-1238.
867 Greig, J.A., T.A. Nguyen, M. Lee, A.S. Holehouse, A.E. Posey, R.V. Pappu, and G. Jedd. 2020. Arginine-
868   Enriched Mixed-Charge Domains Provide Cohesion for Nuclear Speckle Condensation. *Mol
869   Cell*. 77:1237-1250 e1234.
870 Hare, S., F. Di Nunzio, A. Labeja, J. Wang, A. Engelman, and P. Cherepanov. 2009. Structural basis for
871   functional tetramerization of lentiviral integrase. *PLoS Pathog*. 5:e1000515.
872 Hrecka, K., C. Hao, M. Gierszewska, S.K. Swanson, M. Kesik-Brodacka, S. Srivastava, L. Florens, M.P.
873   Washburn, and J. Skowronski. 2011. Vpx relieves inhibition of HIV-1 infection of macrophages
874   mediated by the SAMHD1 protein. *Nature*. 474:658-661.
875 Hubner, W., G.P. McNerney, P. Chen, B.M. Dale, R.E. Gordon, F.Y. Chuang, X.D. Li, D.M. Asmuth, T.
876   Huser, and B.K. Chen. 2009. Quantitative 3D video microscopy of HIV transfer across T cell
877   virological synapses. *Science*. 323:1743-1747.
878 Hulme, A.E., Z. Kelley, D. Foley, and T.J. Hope. 2015. Complementary Assays Reveal a Low Level of CA
879   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.

Pradeepa, M.M., H.G. Sutherland, J. Ule, G.R. Grimes, and W.A. Bickmore. 2012. Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLoS Genet*. 8:e1002717.

Price, A.J., D.A. Jacques, W.A. McEwan, A.J. Fletcher, S. Essig, J.W. Chin, U.D. Halambage, C. Aiken, and L.C. James. 2014. Host cofactors and pharmacologic ligands share an essential interface in HIV-1 capsid that is lost upon disassembly. *PLoS Pathog*. 10:e1004459.

Rensen, E., F. Mueller, V. Scoca, J.J. Parmar, P. Souque, C. Zimmer, and F. Di Nunzio. 2020. Clustering and reverse transcription of HIV-1 genomes in nuclear niches of macrophages. *EMBO J*:e105247.

Rensen, E., F. Mueller, V. Scoca, J.J. Parmar, P. Souque, C. Zimmer, and F. Di Nunzio. 2021. Clustering and reverse transcription of HIV-1 genomes in nuclear niches of macrophages. *EMBO J*. 40:e105247.

Saad, H., F. Gallardo, M. Dalvai, N. Tanguy-le-Gac, D. Lane, and K. Bystricky. 2014. DNA dynamics during early double-strand break processing revealed by non-intrusive imaging of living cells. *PLoS Genet*. 10:e1004187.

Sanchez, A., D.I. Cattoni, J.C. Walter, J. Rech, A. Parmeggiani, M. Nollmann, and J.Y. Bouet. 2015. Stochastic Self-Assembly of ParB Proteins Builds the Bacterial DNA Segregation Apparatus. *Cell Syst*. 1:163-173.

Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 9:676-682.

Schroder, A.R., P. Shinn, H. Chen, C. Berry, J.R. Ecker, and F. Bushman. 2002. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*. 110:521-529.

Scoca, V., and F. Di Nunzio. 2021a. The HIV-1 Capsid: From Structural Component to Key Factor for Host Nuclear Invasion. *Viruses*. 13.

Scoca, V., and F. Di Nunzio. 2021b. Membraneless organelles restructured and built by pandemic viruses: HIV-1 and SARS-CoV-2. *J Mol Cell Biol*.

Sharkey, M., D.Z. Babic, T. Greenough, R. Gulick, D.R. Kuritzkes, and M. Stevenson. 2011. Episomal viral cDNAs identify a reservoir that fuels viral rebound after treatment interruption and that contributes to treatment failure. *PLoS Pathog*. 7:e1001303.

Sharkey, M.E., I. Teo, T. Greenough, N. Sharova, K. Luzuriaga, J.L. Sullivan, R.P. Bucy, L.G. Kostrikis, A. Haase, C. Veryard, R.E. Davaro, S.H. Cheeseman, J.S. Daly, C. Bova, R.T. Ellison, 3rd, B. Mady, K.K. Lai, G. Moyle, M. Nelson, B. Gazzard, S. Shaunak, and M. Stevenson. 2000. Persistence of episomal HIV-1 infection intermediates in patients on highly active anti-retroviral therapy. *Nat Med*. 6:76-81.

Shun, M.C., Y. Botbol, X. Li, F. Di Nunzio, J.E. Daigle, N. Yan, J. Lieberman, M. Lavigne, and A. Engelman. 2008. Identification and characterization of PWWP domain residues critical for LEDGF/p75 chromatin binding and human immunodeficiency virus type 1 infectivity. *J Virol*. 82:11555-11567.

Shun, M.C., N.K. Raghavendra, N. Vandegraaff, J.E. Daigle, S. Hughes, P. Kellam, P. Cherepanov, and A. Engelman. 2007. LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev*. 21:1767-1778.

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. Hughes, M. Kvaratskhelia, and H.L. Levin. 2015. LEDGF/p75 interacts with mRNA splicing factors and targets HIV-1 integration to highly spliced genes. *Genes Dev*. 29:2287-2297.

Spector, D.L., and A.I. Lamond. 2011. Nuclear speckles. *Cold Spring Harb Perspect Biol*. 3.

Stultz, R.D., J.J. Cenker, and D. McDonald. 2017. Imaging HIV-1 Genomic DNA from Entry through Productive Infection. *J Virol*. 91.

Suzuki, Y., and R. Craigie. 2007. The road to chromatin - nuclear entry of retroviruses. *Nat Rev Microbiol*. 5:187-196.

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

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

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

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

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

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

999   Wang, H., M.C. Shun, X. Li, F. Di Nunzio, S. Hare, P. Cherepanov, and A. Engelman. 2014. Efficient
1000   Transduction of LEDGF/p75 Mutant Cells by Gain-of-Function HIV-1 Integrase Mutant Viruses.
1001   *Mol Ther Methods Clin Dev*. 1.

1002   Zhu, Y., G.Z. Wang, O. Cingoz, and S.P. Goff. 2018. NP220 mediates silencing of unintegrated retroviral
1003   DNA. *Nature*. 564:278-282.

1004

1005

1006

1007

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, ****=$p \leq 0.0001$. **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, **=$p \leq 0.01$, ****=$p \leq 0.0001$. **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

1041 ≤ 0.0001. **B)** Confocal image of HIV-1 RNA FISH in Jurkat CD4+T cells (MOI 5), 48 h

1042 p.i. The scatter plot with bars shows the percentage of vDNA associated to vRNA

1043 signal and vice versa, per nucleus (12 cells; 262 vDNAs; 64 vRNA foci). The analysis

1044 was performed in 3D. Unpaired t test, **=p ≤ 0.01. **C)** Confocal image of HIV-1 RNA

1045 FISH in macrophage-like THP-1 cells (MOI 20) 3 days p.i. The scatter plot with bars

1046 shows the percentage of vDNA associated to vRNA signal and vice versa, per nucleus

1047 (8 cells; 130 vDNAs; 99 vRNA foci). The analysis was performed in 3D. Unpaired t test,

1048 *=p ≤ 0.05. **D)** Confocal image of HIV-1 RNA FISH in primary activated CD4+T cells

1049 (MOI 50), 3 days p.i. and primary monocyte-derived macrophages (MDMs), (MOI 40),

1050 4 days p.i.. All graphs contain ± SD. Scale bar: 5 μm.

1051

1052 **Figure 3. Nuclear localization of HIV-1 provirus in dividing and non-dividing**

1053 **cells.** Co-labeling of vDNA and vRNA by coupling of HIV-1 ANCHOR with RNA FISH.

1054 **A)** Confocal images of HIV-1 RNA FISH in HeLa P4R5 cells infected with HIV-1

1055 ANCH3 (MOI 5) at 48 h p.i. compared to 7 days p.i.. The scatter plot with bars indicates

1056 the distance from the nuclear envelope (NE) per detected vDNA spot (green) and

1057 vRNA focus (red) (n cells = 11 (48 h p.i.), 10 (+RAL), 5 (IN$_{D116A}$), and 9 (7 days p.i);

1058 vDNA: 83 (48 h p.i.), 87 (+RAL), 29 (IN$_{D116A}$), 16 (7 d p.i.); vRNA: 52 (48h p.i.) and 18

1059 (7d p.i.)). One-way ANOVA followed by Turkey's multiple comparison test, ns=non-

1060 significant, *=p ≤ 0.05, **=p ≤ 0.01. **B)** Confocal image of HIV-1 RNA FISH in THP-1

1061 cells infected with HIV-1 ANCH3 (MOI 20), 3 days p.i.. The scatter plot with bars

1062 indicates the distance from the nuclear envelope (NE) per detected vDNA (green) or

1063 vRNA (red) (n cells = 8 (vDNA, vRNA), 13 (IN$_{D116A}$); n spots= 139 (vDNA), 76 (IN$_{D116A}$),

1064 126 (vRNA)). Unpaired t test, *=p ≤ 0.05, ***=p ≤ 0.001. **C)** Scatter plot with bars

1065 comparing the distances from the NE per detected vRNA focus in HeLa P4R5 and

1066 THP-1 cells (n = 11 (HeLa) and 8 (THP-1) cells; vRNAs: 52 (HeLa) and 126 (THP-1)).

1067 Unpaired t test, ns=non-significant. All the analyses were performed in 3D. All graphs

1068 contain ± SD. Scale bar: 5 μm, inset: 1 μm.

1069

1070 **Figure 4. Sequestration of viral components in CPSF6 clusters. A)** Late reverse

1071 transcripts separation from integrase (IN) foci: confocal images of infected primary

1072 activated CD4[+] T cells (MOI 100, 20 h p.i.) and macrophage-like cells, THP-1 cell line,

1073 (MOI 10, 30h p.i.), compared to uninfected cells. The scatter plot with bars compares

1074   the percentages of vDNAs associated to IN, per nucleus (CD4+T = 8 cells, 54 vDNAs;

1075   THP-1 = 13 cells, 100 vDNAs). Unpaired t test, ns = non-significant. The analysis was

1076   performed in 3D. **B)** On the left, quantification of IN foci found associated to vDNA

1077   signal for each time-lapse video (at frame 0) (n = 13 cells, 81 IN foci). Six different

1078   videos were performed at sequential time p.i. On the right, example of vDNA

1079   separation in live, tracked down in the 18-hours p.i. video (video 6). Pearson's

1080   correlation per time frame is shown below. **C)** Confocal images of infected primary

1081   activated CD4$^+$ T cells (MOI 100, 20 h p.i.) and THP-1 cells (MOI 10, 30h p.i.),

1082   compared to uninfected cells. The graph shows the percentages of IN associated to

1083   CPSF6, per nucleus (CD4$^+$T = 12 cells, 33 IN foci; THP-1 = 40 cells, 132 IN foci).

1084   Unpaired t test, ns = non-significant. **D)** Confocal images of Immuno-RNA FISH in

1085   SupT1R5 CD4+T cells and THP-1 cells infected with HIV-1 ANCH3 (MOI 20; 2 days

1086   p.i.). The graph shows the percentages of vRNA associated to CPSF6 per condition.

1087   (CD4+T = 19 (-NEV), 16 (+NEV) cells, 79 (-NEV), 30 (+NEV) vRNAs; THP-1 = 56 (-

1088   NEV), 34 (+NEV) cells, 151 (-NEV), 61 (+NEV) vRNAs). Unpaired t test, **=$p \leq 0.01$,

1089   ****=$p \leq 0.0001$. All graphs contain ± SD. Scale bar: 5 µm, inset: 1 µm.

1090

1091   **Figure 5. HIV- MLOs as nuclear hubs of HIV-1 reverse transcription. A)** Confocal

1092   images of Immuno-RNA FISH in THP-1 cells infected with HIV-1 ANCH3 (MOI 20) ±

1093   NEV (10 µM). The graph shows the average percentage per nucleus of vRNA found

1094   inside CPSF6 clusters, ± SEM (2 days p.i. = 56 cells, 151 vRNA (-NEV), 61 vRNA

1095   (+NEV); 3 days p.i. = 69 cells, 228 vRNA (-NEV), 129 vRNA (+NEV); 7 days p.i. = 45

1096   cells, 148 vRNA (-NEV); 57 vRNA (+NEV). **B)** Chromatin Immunoprecipitation of

1097   endogenous CPSF6 coupled to real-time PCR of different reverse transcription (RT)

1098   products in THP-1 cells infected with HIV-1 (MOI 5), 2 days p.i. ± SD. One-way ANOVA

1099   followed by Turkey's multiple comparison test, *=$p \leq 0.05$, **=$p \leq 0.01$, ****=$p \leq 0.0001$.

1100   ChIP-CPSF6 products compared to the input in infected cells by western blot using an

1101   Ab anti-CPSF6. **C)** Confocal images of THP-1 cells infected with HIV-1 ANCH3 (MOI

1102   10) ± NEV (10 µM), 2 days p.i. Scale bars: 5 µm.

1103

1104   **Figure 6. HIV-1 proviruses locate outside HIV-1 MLOs. A)** Confocal images of THP-

1105   1 infected with HIV-1 ANCH3, 3 days p.i. The graph shows the intensity profiles of

1106   Hoechst, CPSF6 and SC35 signals along the segment crossing the MLO (yellow). **B)**

1107   Confocal images of THP-1 cells expressing OR-GFP infected with HIV-1 ANCH3, 3

1108 days p.i., compared to uninfected cells. On the right, the graphs show: the comparison

1109 of Pearson's r coefficients of SON-SC35 colocalization between infected and

1110 uninfected cells, per nucleus (n = 25 cells (HIV-1 ANCH3), 19 cells (NI); unpaired t

1111 test, ns = non-significant); the percentages of CPSF6 clusters associated to vDNA or

1112 to SC35 in 3D, per nucleus (n = CPSF6 clusters > 50); the distance of each vDNA from

1113 the closest SC35-speckle border (n= 83 vDNAs, 7 cells). **C)** Confocal images of THP-

1114 1 cells expressing MCP-GFP infected with HIV-1 ANCH3 MS2 (MOI 2.5), 3 days p.i.,

1115 ± NEV (10μM). The graph shows the intensity profiles of Hoechst, MCP-GFP, CPSF6

1116 and SC35 signals along the segment crossing the MLO (yellow). **D)** Confocal images

1117 of THP-1 cells expressing MCP-GFP infected with HIV-1 ANCH3 MS2 (MOI 2.5), 3

1118 days p.i., compared to uninfected cells. The graph shows the percentages of CPSF6

1119 clusters associated to vRNA foci or to SC35 in 3D (n = 42 CPSF6 clusters, 8 cells) and

1120 the distance of each vRNA from the closest SC35-speckle border (n = 31 vRNAs, 8

1121 cells). **E)** Scatter plot with bars indicating the percentage of vDNA or vRNA associated

1122 to CPSF6, per nucleus in 3D (n = 69 vDNAs, 30 vRNAs, 8 cells). Unpaired t test, ns =

1123 non-significant. **F)** Scatter plot with bars indicating the distance of vDNA or vRNA from

1124 the border of the closest SC35/CPSF6 MLO (n = 75 vDNAs, 7 cells; 34 vRNAs, 8 cells).

1125 Unpaired t test, **=p ≤ 0.01. **G)** Confocal image of Immuno-RNA FISH in THP-1 cell

1126 expressing OR-GFP infected with HIV-1 ANCH3 (MOI 20), 3 days p.i. Scale bar: 5 μm,

1127 inset: 1 μm.

1128

1129 **Figure 7. HIV-1 transcribes outside HIV-1 MLOs in primary MDMs.** Confocal

1130 images of Immuno-RNA FISH in primary MDMs infected with HIV-1 GFP (MOI 20) ±

1131 NEV (10 μM), 2-3 days p.i. The scatter plots show the GFP intensity per cell (n = 22 (-

1132 NEV) 20 (+NEV) cells) and the percentage of vRNA outside CPSF6 clusters per

1133 nucleus (n = 22 (-NEV) 20 (+NEV) cells; vRNAs: 268 (-NEV), 66 (+NEV)); unpaired t

1134 test, ****=p ≤ 0.0001. Their correlation in infected cells (Pearson's r coefficient) is

1135 displayed in the graph (n = 22 cells). The FACS histogram displays CD14+ expression

1136 of the MDMs. The FACS dot plot shows the correlation of GAG signal and GFP

1137 expression, with the related percentages per population. Scale bar: 5 μm

1138

1139 **Figure 8. HIV-1 proviruses locate in LEDGF-enriched sites. A)** Confocal images of

1140 THP-1 cells expressing OR-GFP infected with HIV-1 ANCH3, 3 days p.i., compared to

1141 uninfected cells. The scatter plots show the percentages of CPSF6 found associated

1142 to LEDGF clusters per nucleus in 3D (n = 16 cells; unpaired t test, ***=p ≤ 0.001) and

1143 the LEDGF and CPSF6 count per nucleus in 3D (n = 16 cells); unpaired t test, ns =

1144 non-significant, ****=p ≤ 0.0001. Scale bar: 5 µm, inset: 2 µm. **B)** Confocal images of

1145 THP-1 cells expressing OR-GFP infected with HIV-1 ANCH3 and expressing MCP-

1146 GFP infected with HIV-1 ANCH3 MS2. The scatter plot shows the percentage of vDNA

1147 or vRNA associated to LEDGF in 3D (n= 33 (vDNA), 15 (vRNA)) ± SD. Unpaired t test,

1148 ns = non-significant. The pie charts show the percentages of the 2 populations of

1149 LEDGF clusters associated to vDNA or vRNA: LEDGF alone or LEDGF in proximity of

1150 CPSF6. Scale bar: 5 µm, inset: 1 µm.

1151

1152 **Figure 9. Remodeling of the nuclear landscape by HIV-1 to replicate.** New, refined

1153 model of viral intranuclear steps. Upon infection, HIV-1 triggers the formation of CPSF6

1154 clusters which localize in nuclear speckles (NSs) (marked by SC35) forming HIV-1

1155 MLOs, which retain viral genomic RNA and integrase proteins. HIV-1 MLOs are sites

1156 of the reverse transcription and of the HIV-1 PIC maturation. Once the mature PIC is

1157 formed, it leaves the CPSF6 cluster to proceed with the integration step. HIV-1

1158 integration occurs in active chromatin regions enriched in LEDGF, close to the nuclear

1159 envelope and to HIV-1 MLOs. Indeed, HIV-1 replication might be favored for viral post-

1160 transcriptional processing by the proximity to NSs that host splicing factors. IN,

1161 Integrase; gRNA, viral genomic RNA; PIC, Pre-Integration Complex; NPC, Nuclear Pore

1162 Complex; vRNA, viral RNA; LTR, Long Terminal Repeat. Created with BioRender.com.

1163
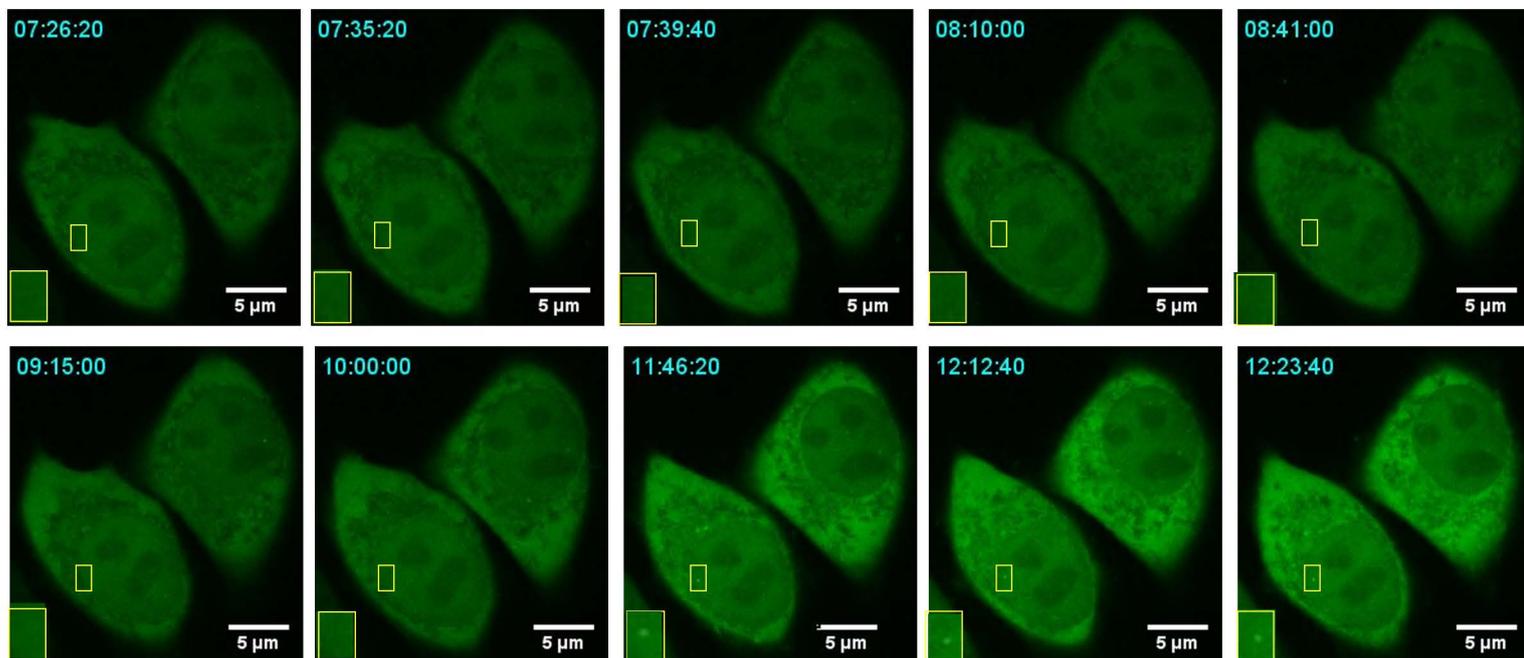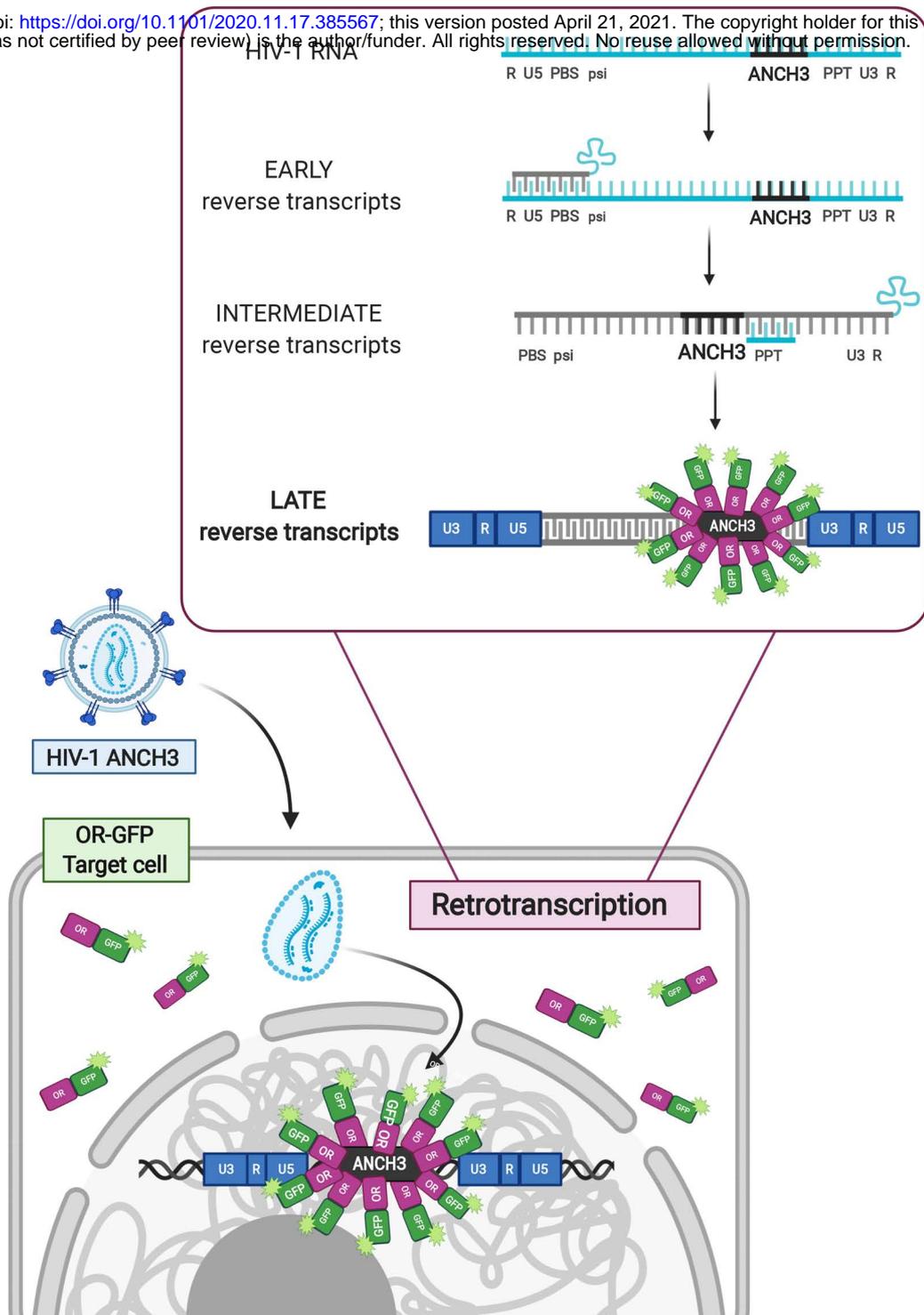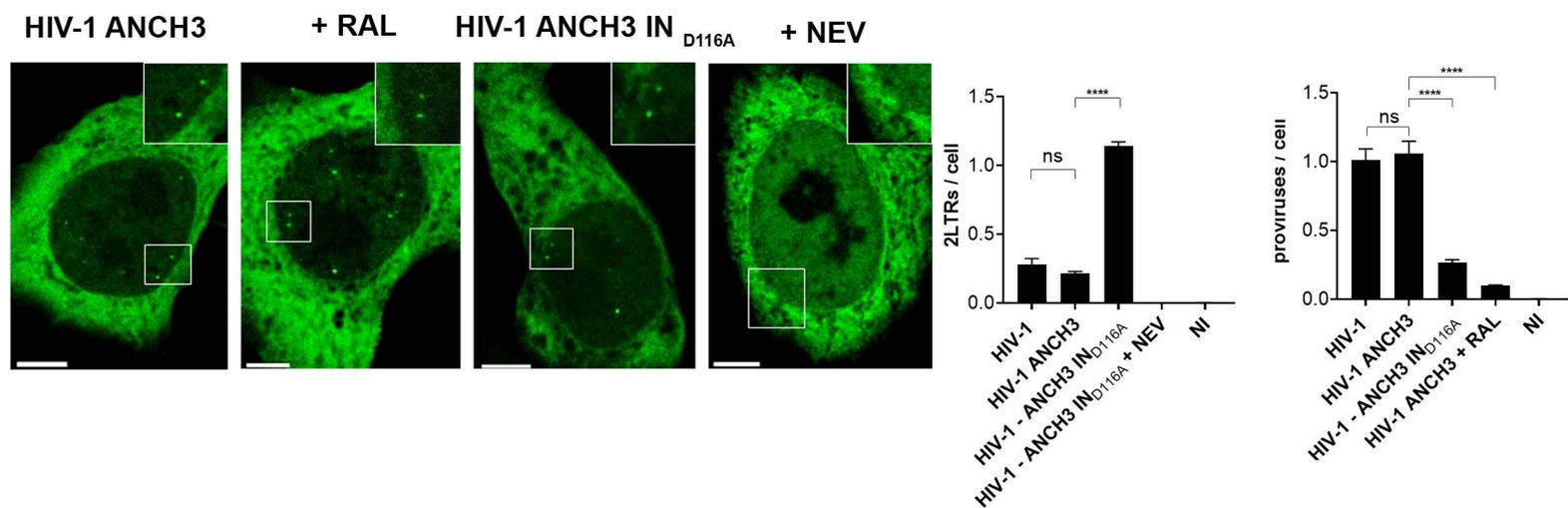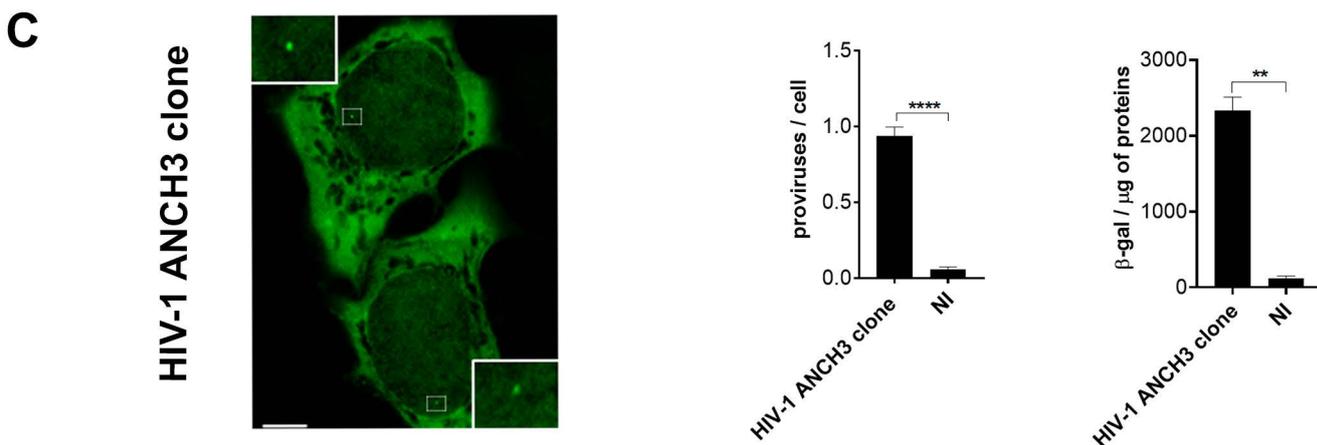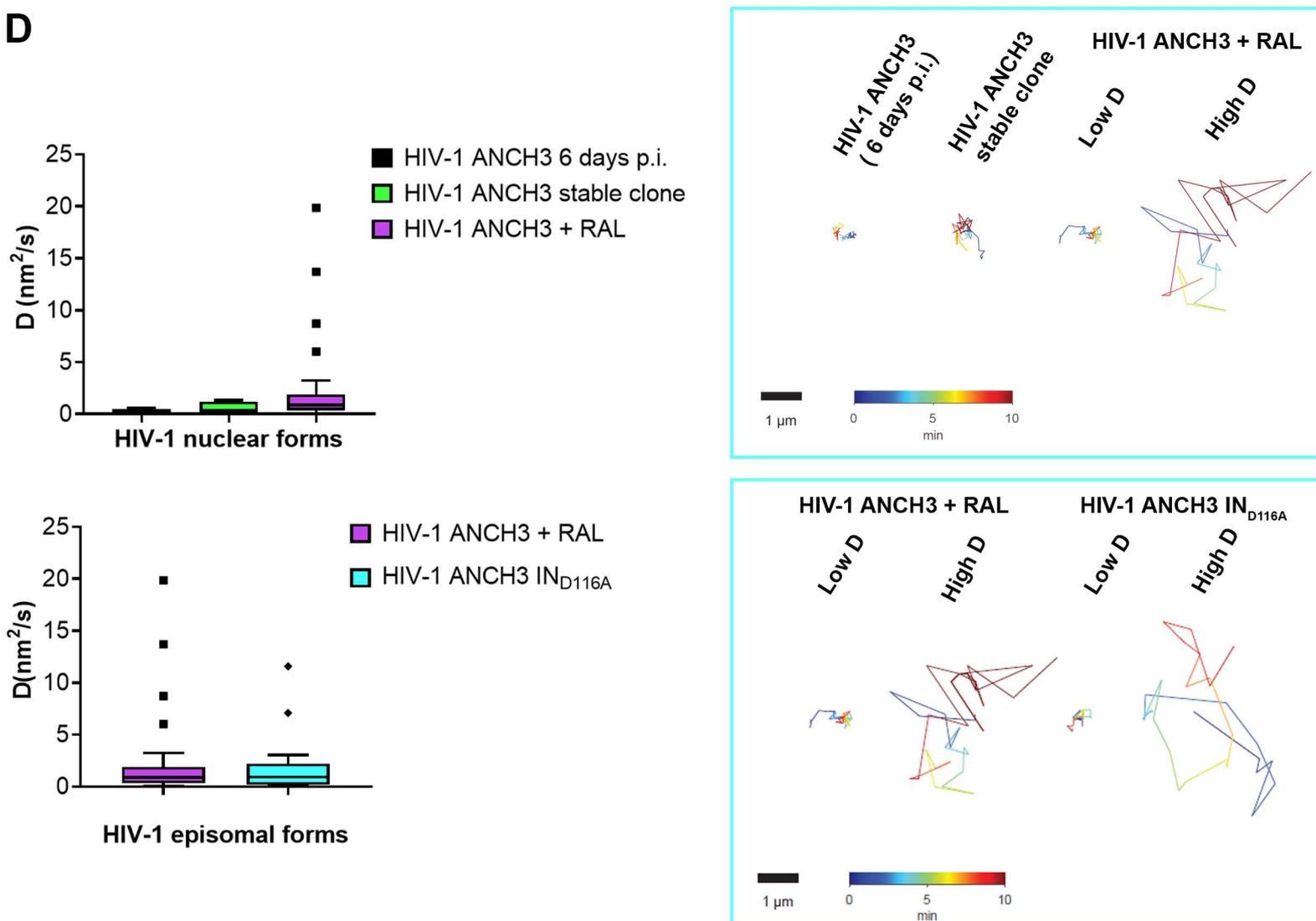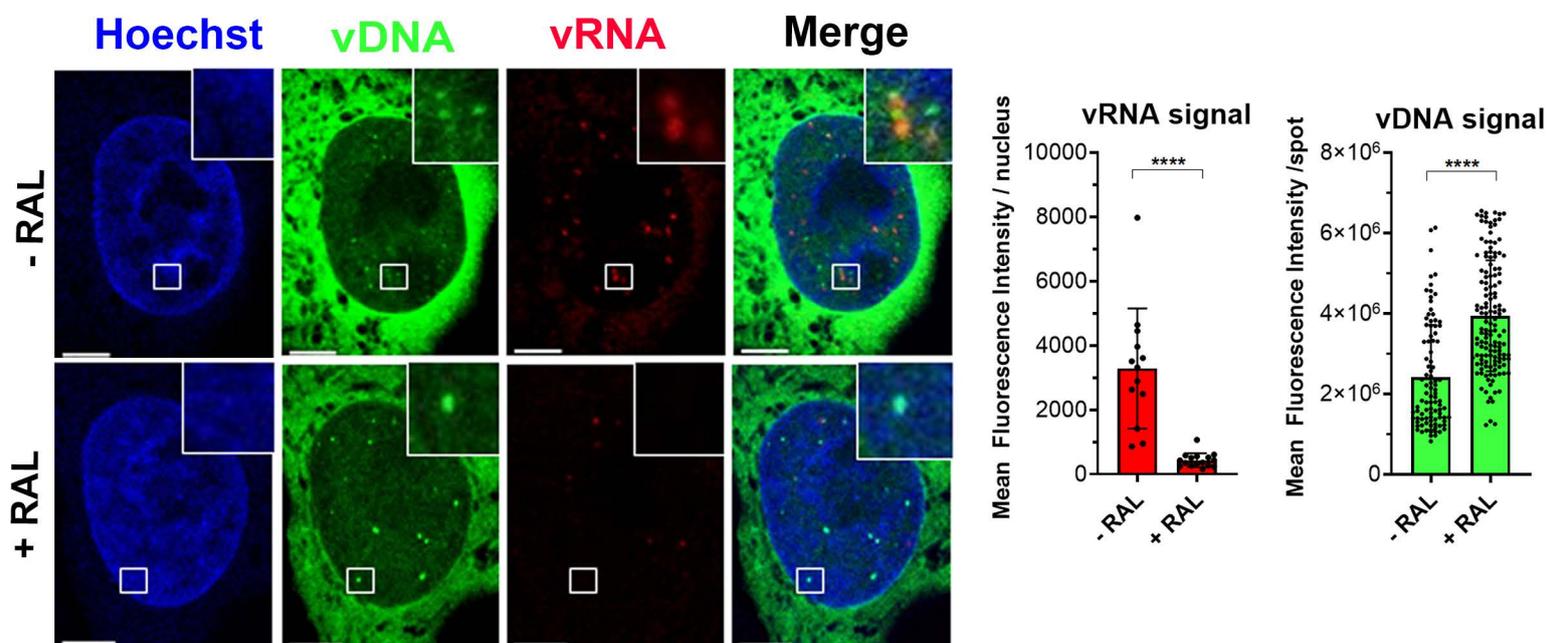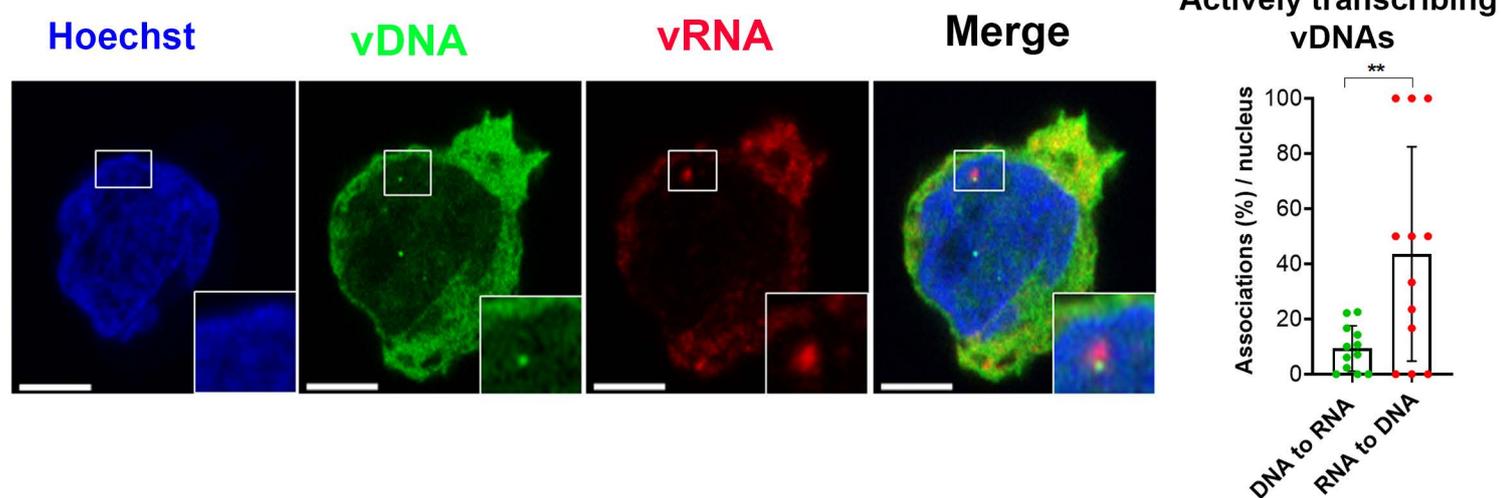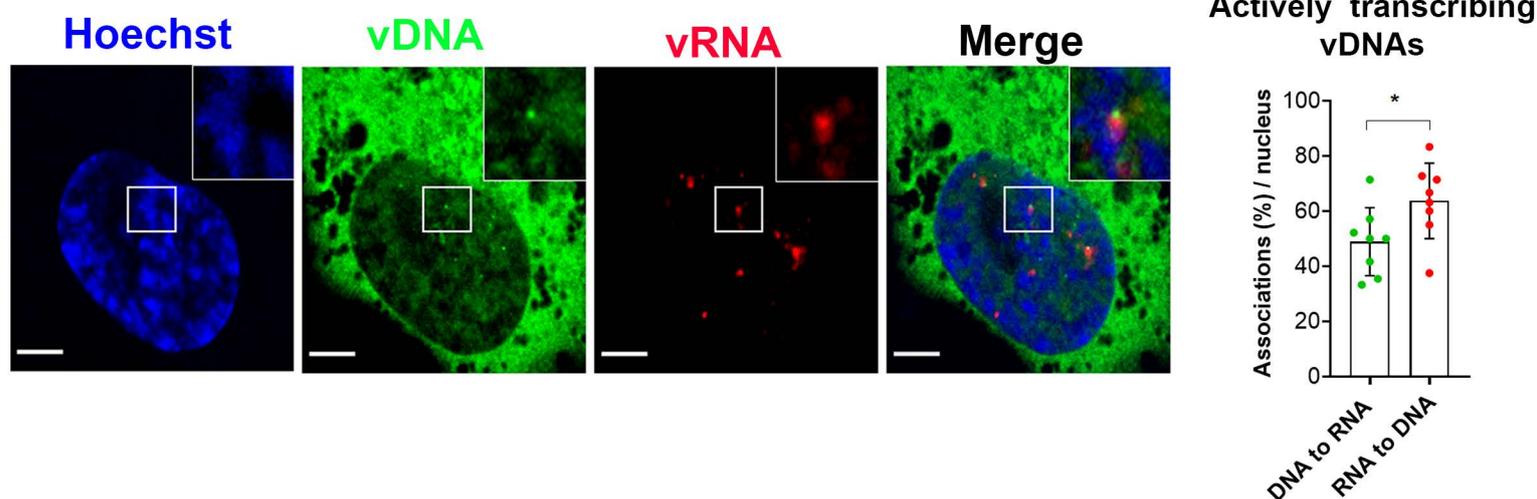
Figure 1

Figure 1

Figure 2

D

Primary CD4+ T cells
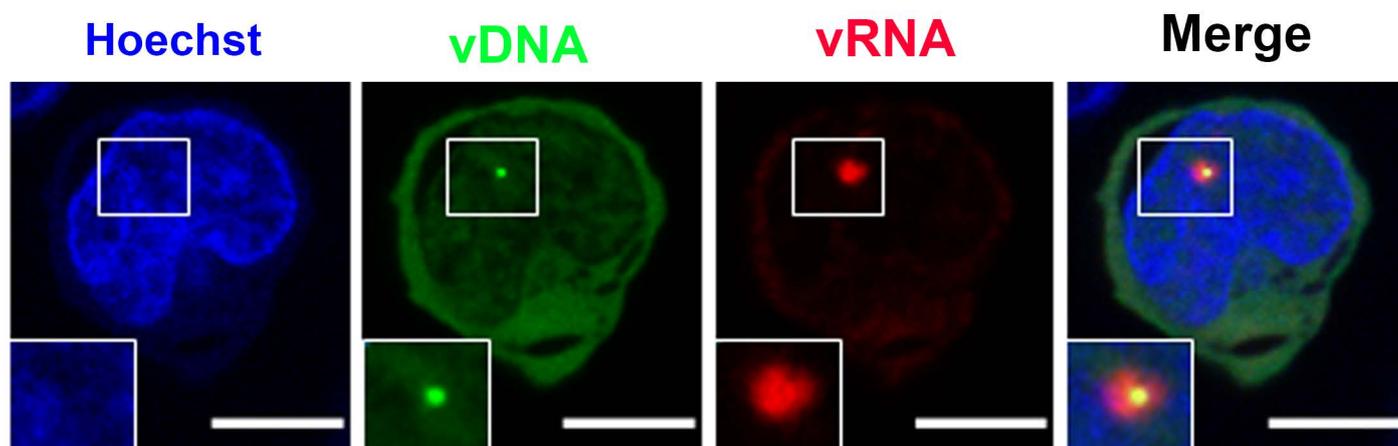
Primary macrophages
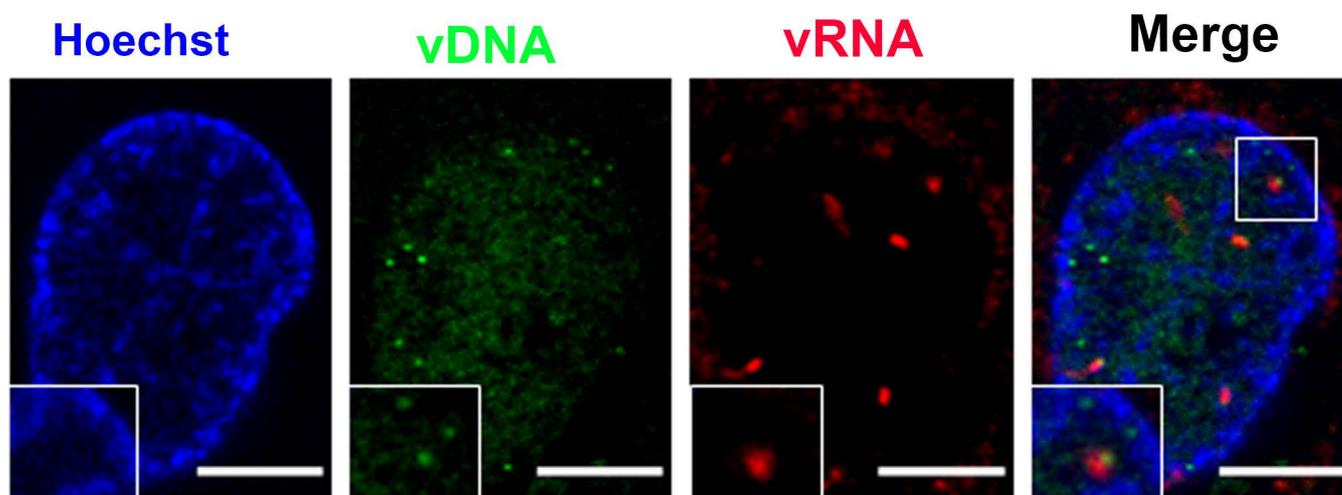
Figure 2

A Dividing cells: HeLa
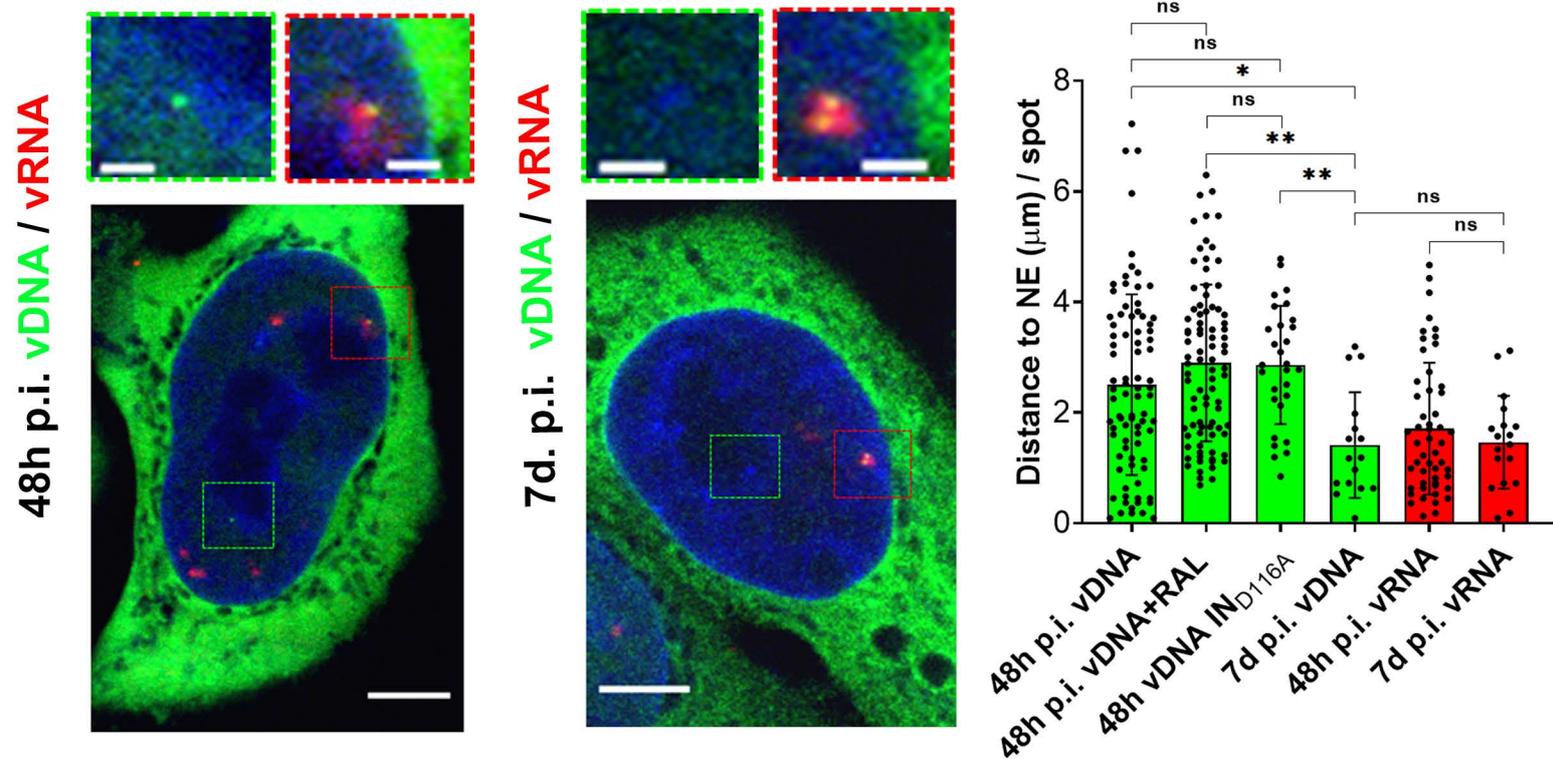
B Non-dividing cells: THP-1
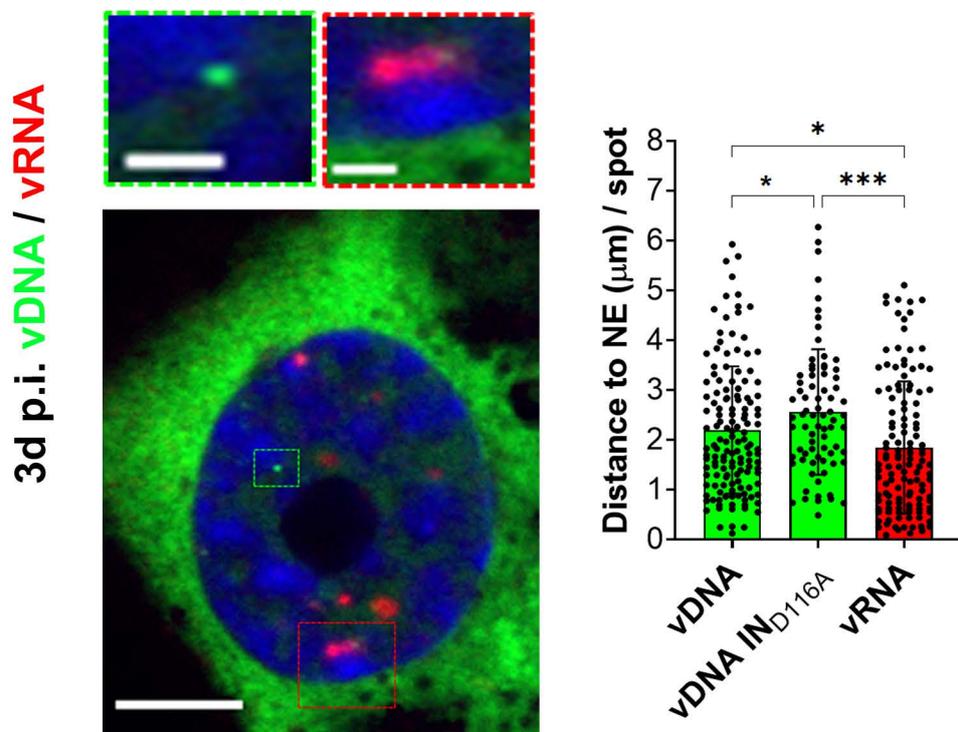
C HeLa vs THP-1
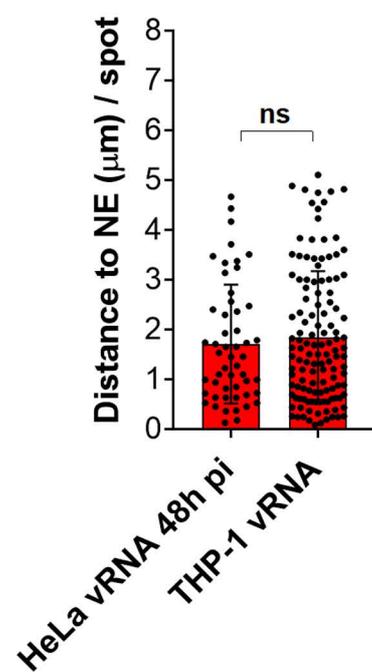
Figure 3

Figure 4

Figure 4

D

Figure 4

Figure 5

C

Hoechst	CPSF6	IN	Merge

- NEV

+ NEV
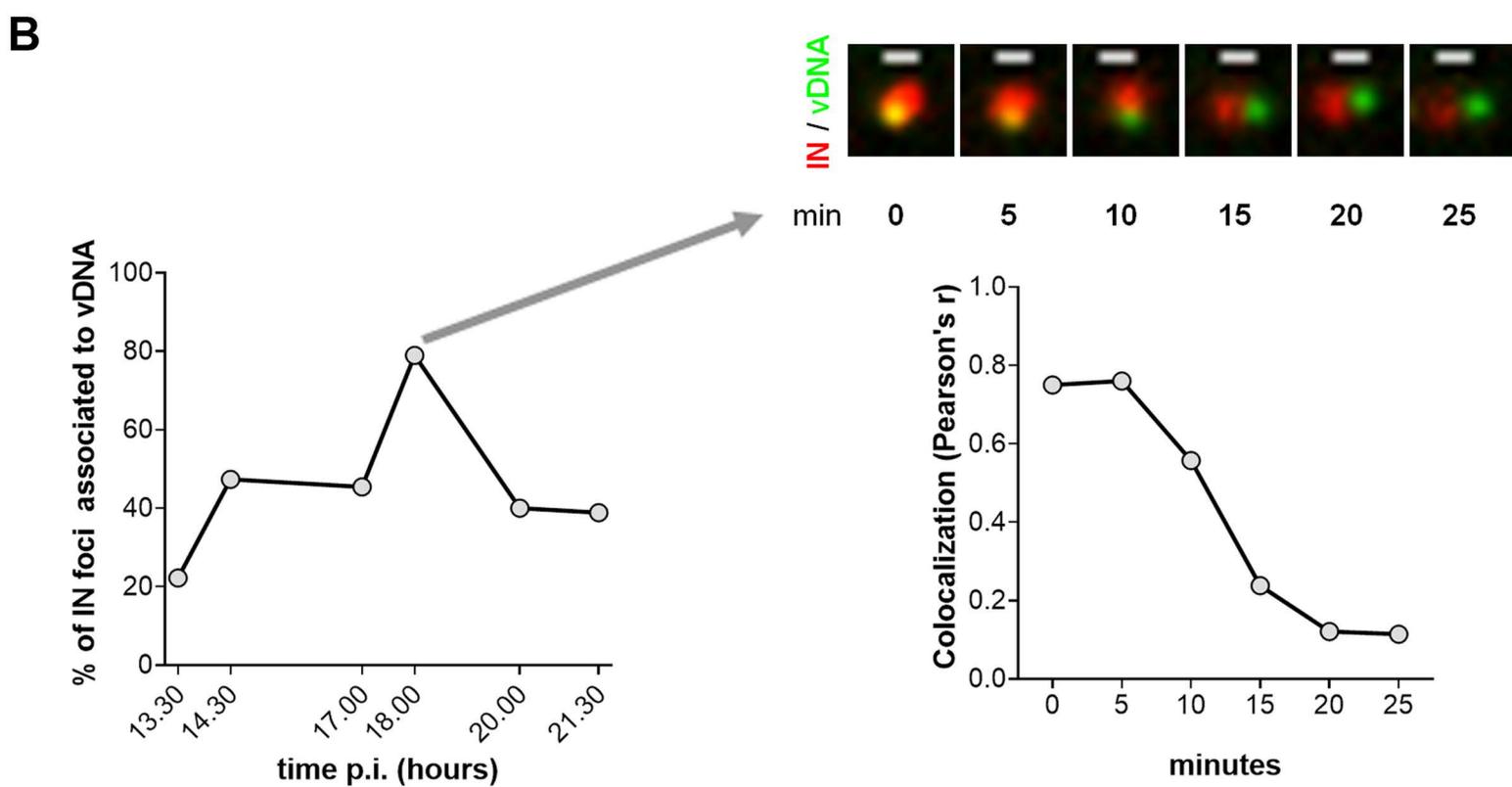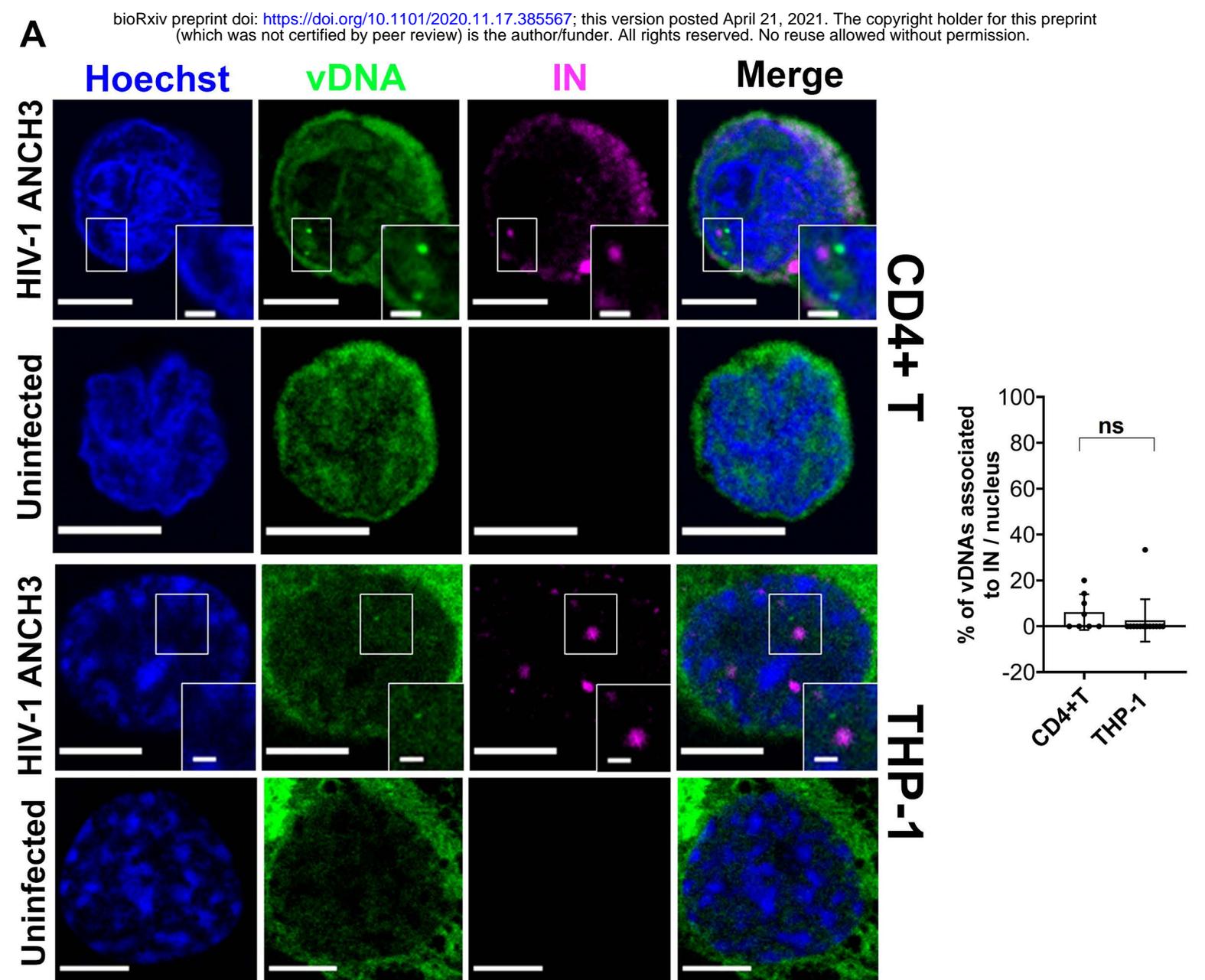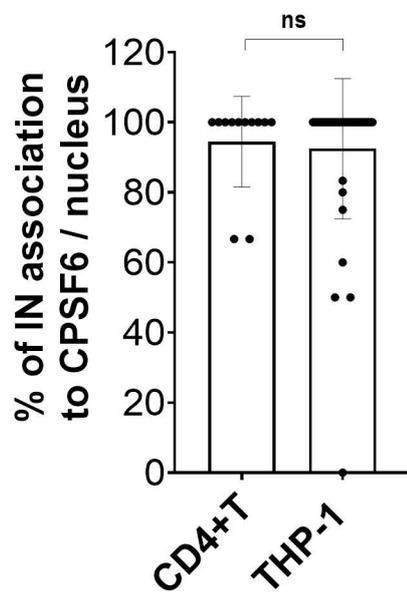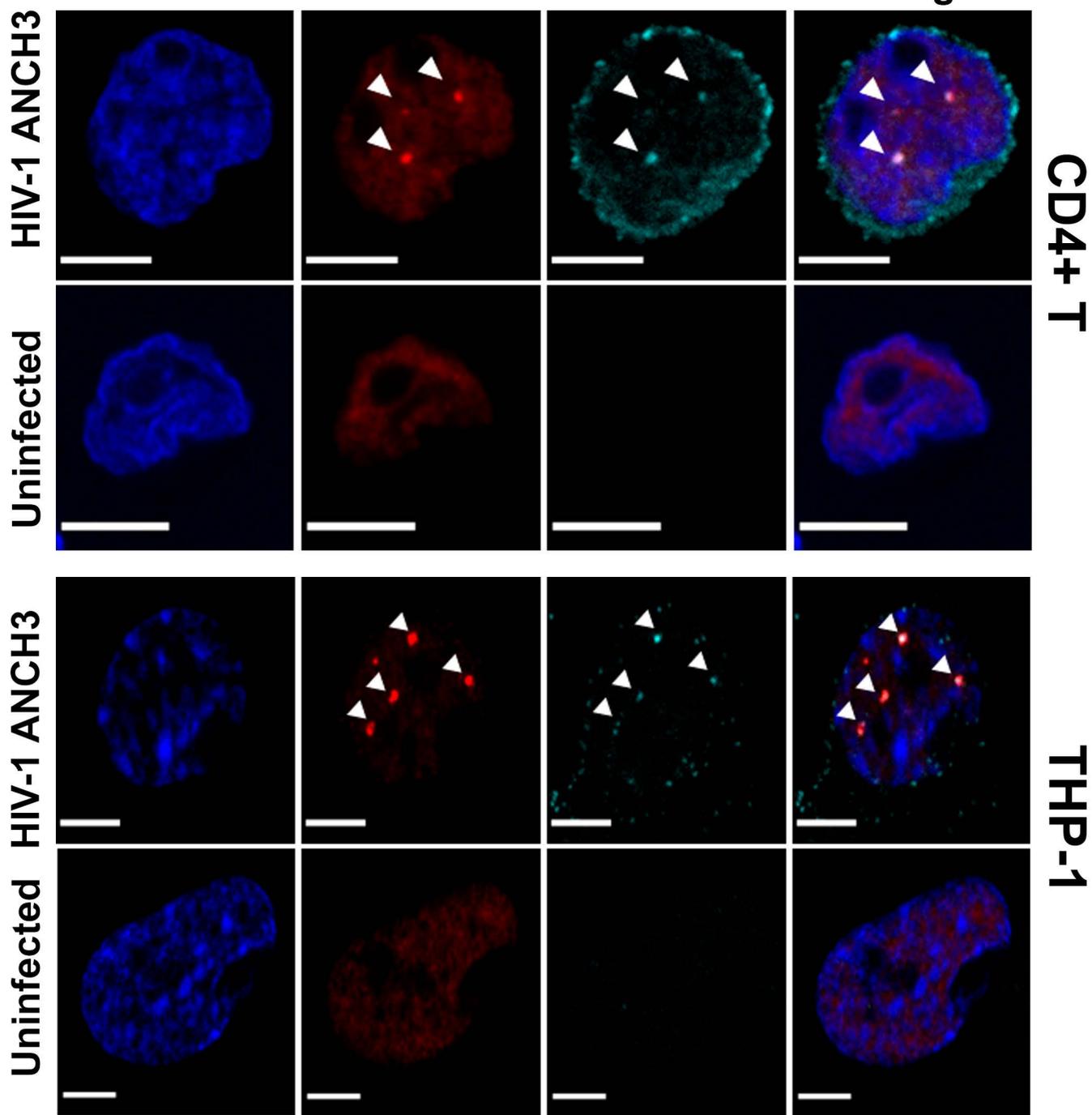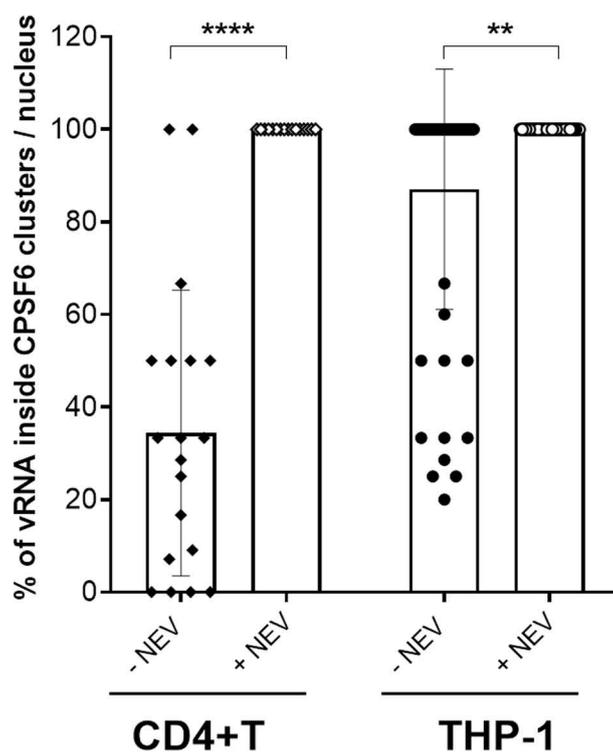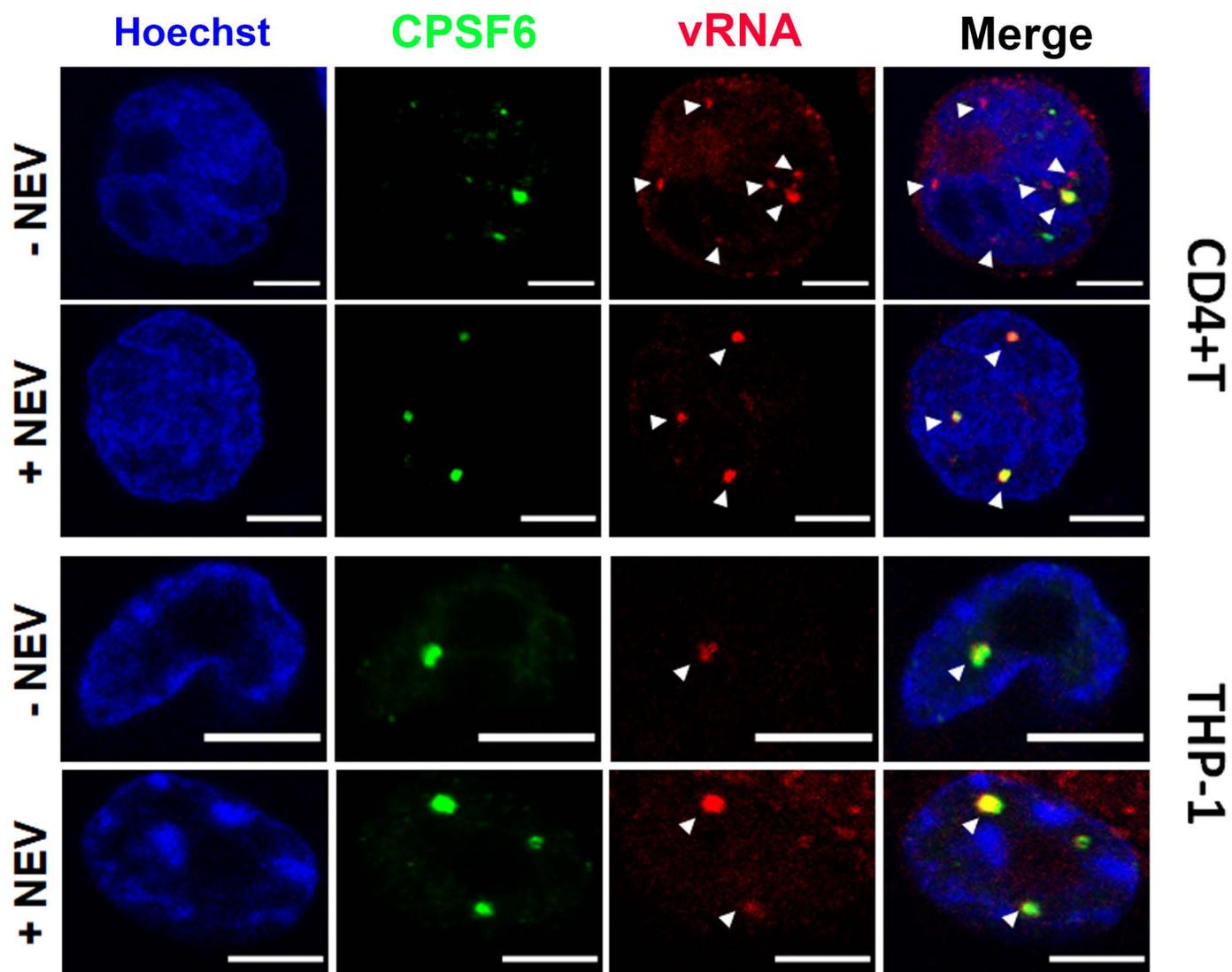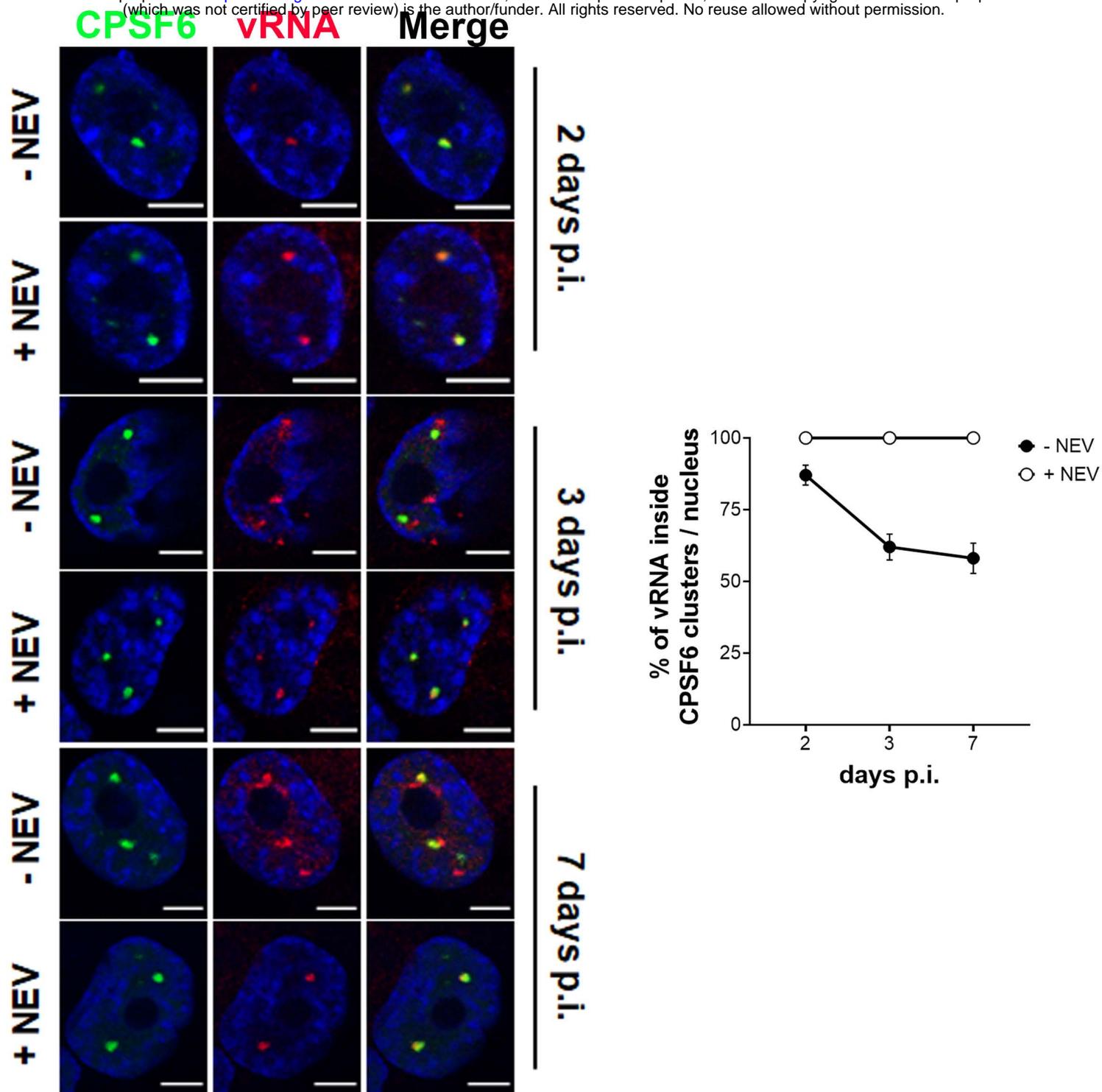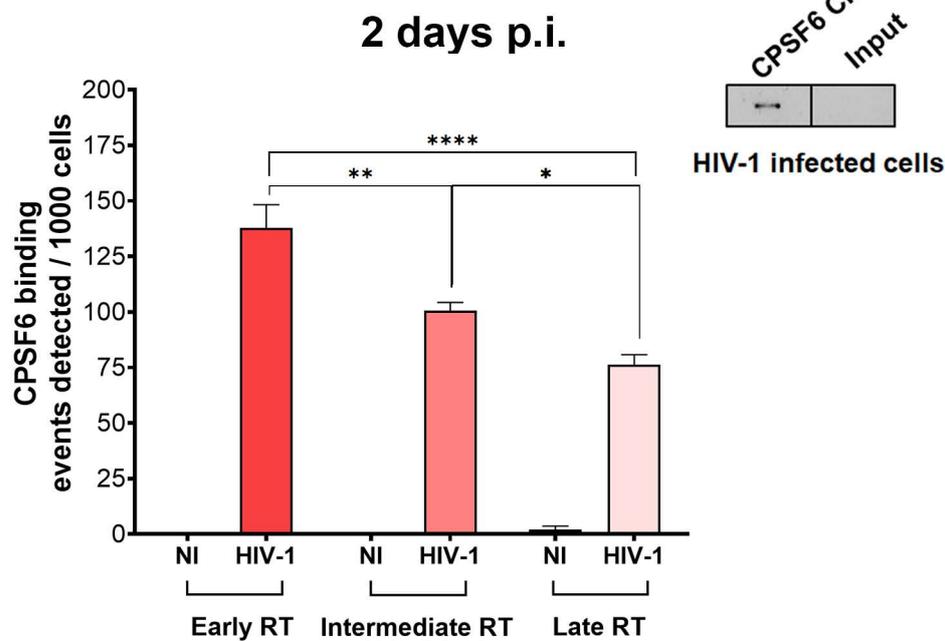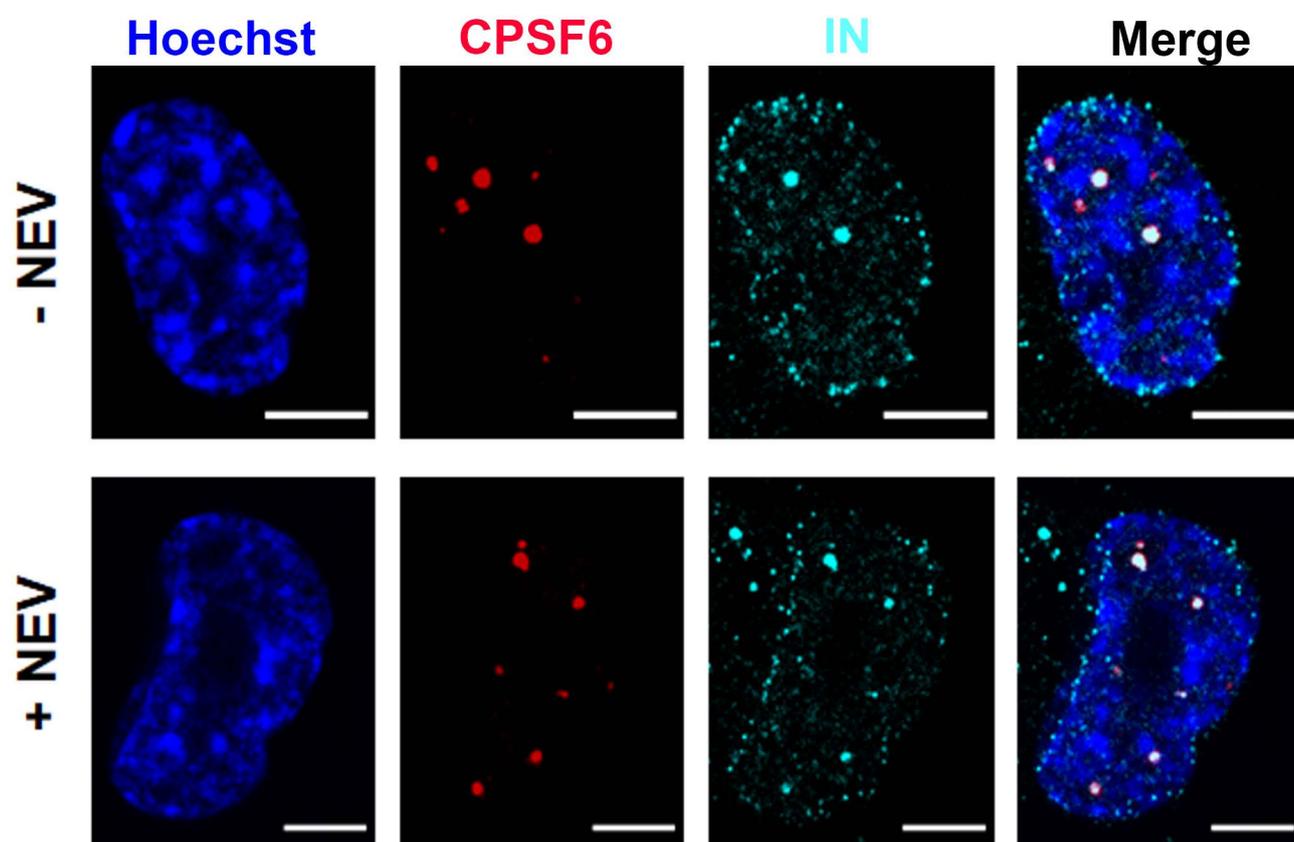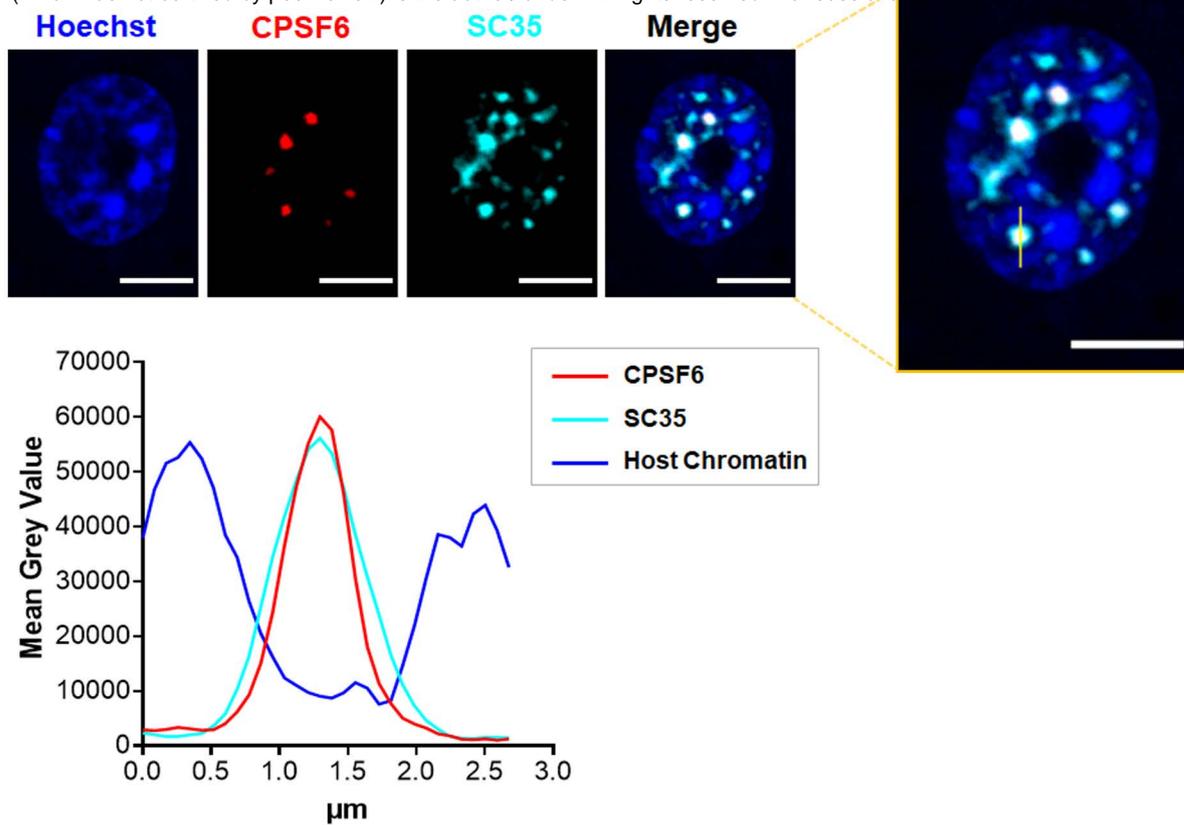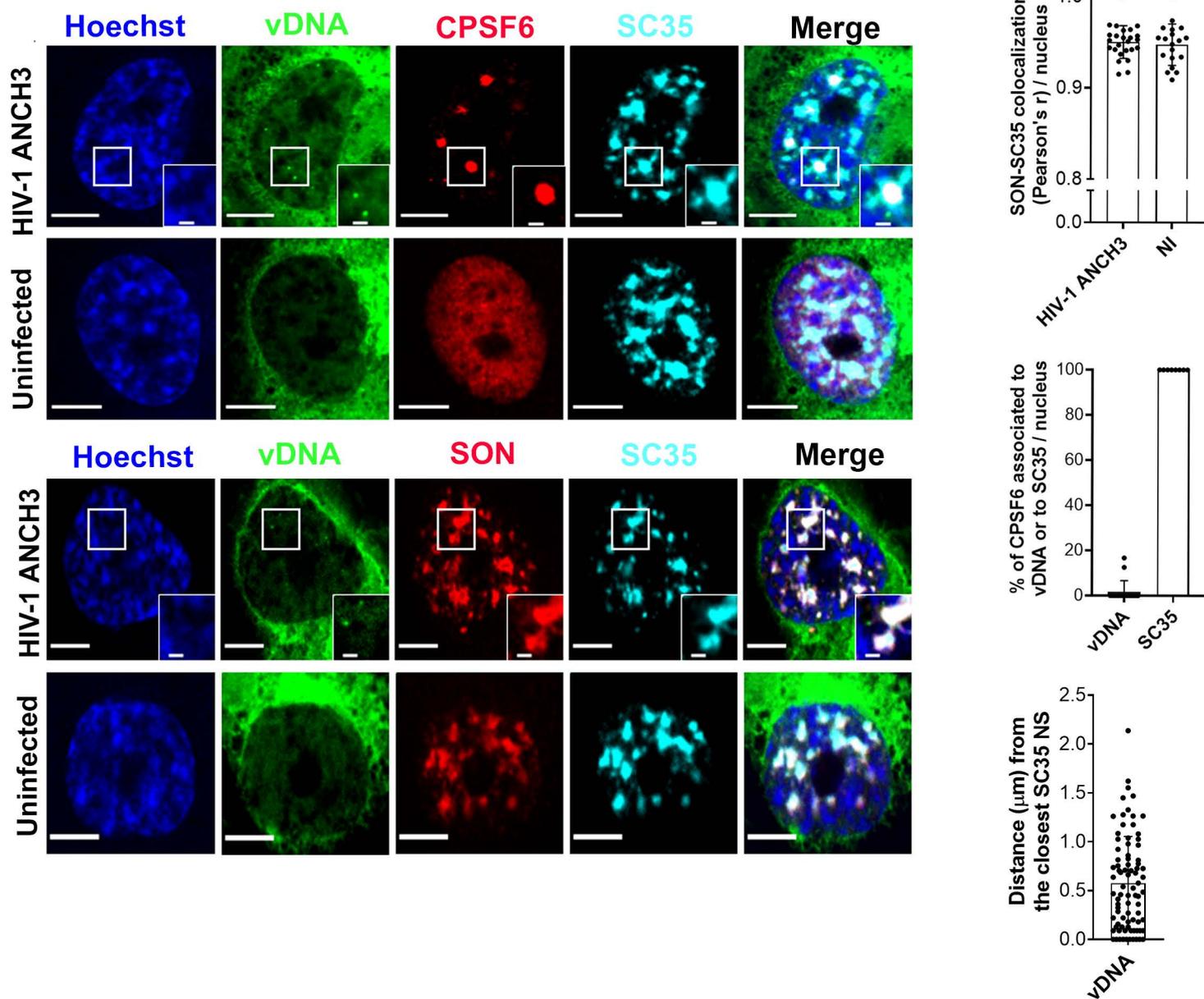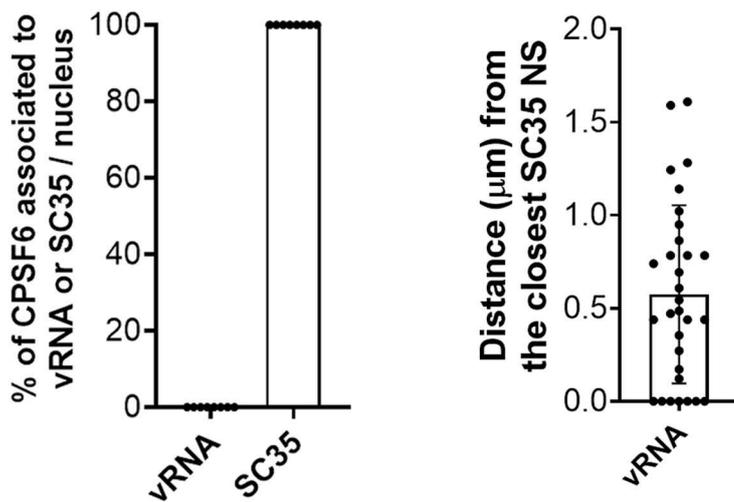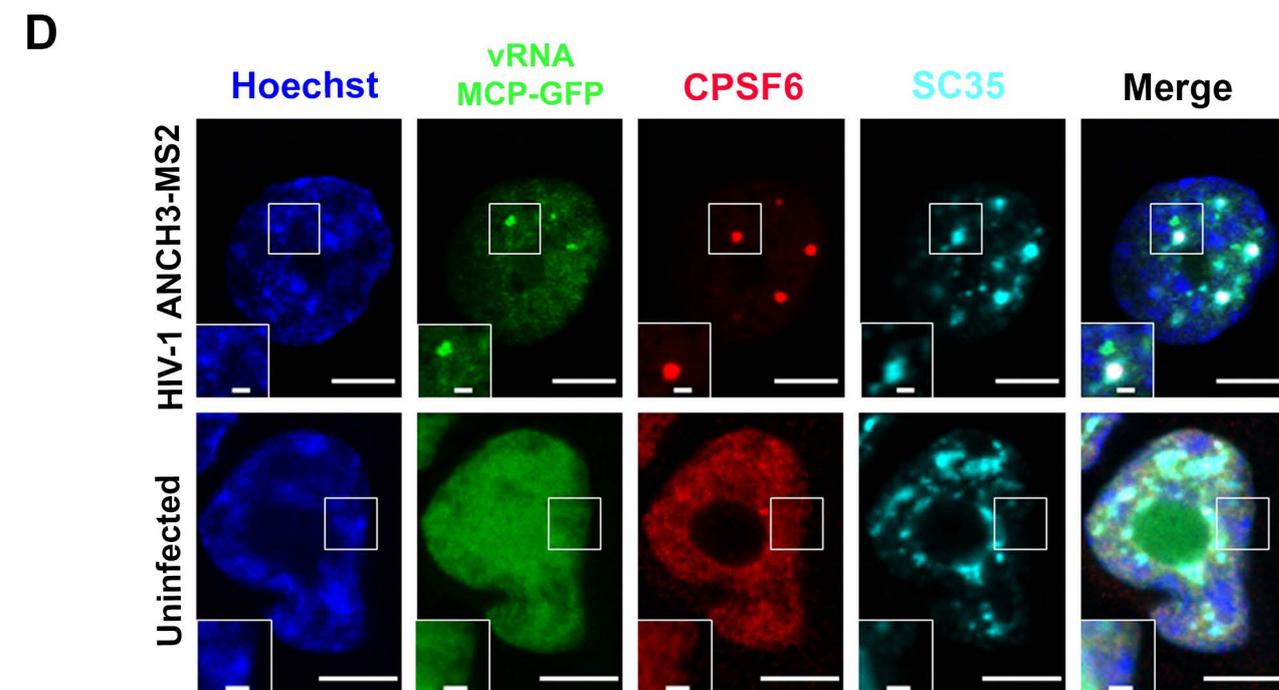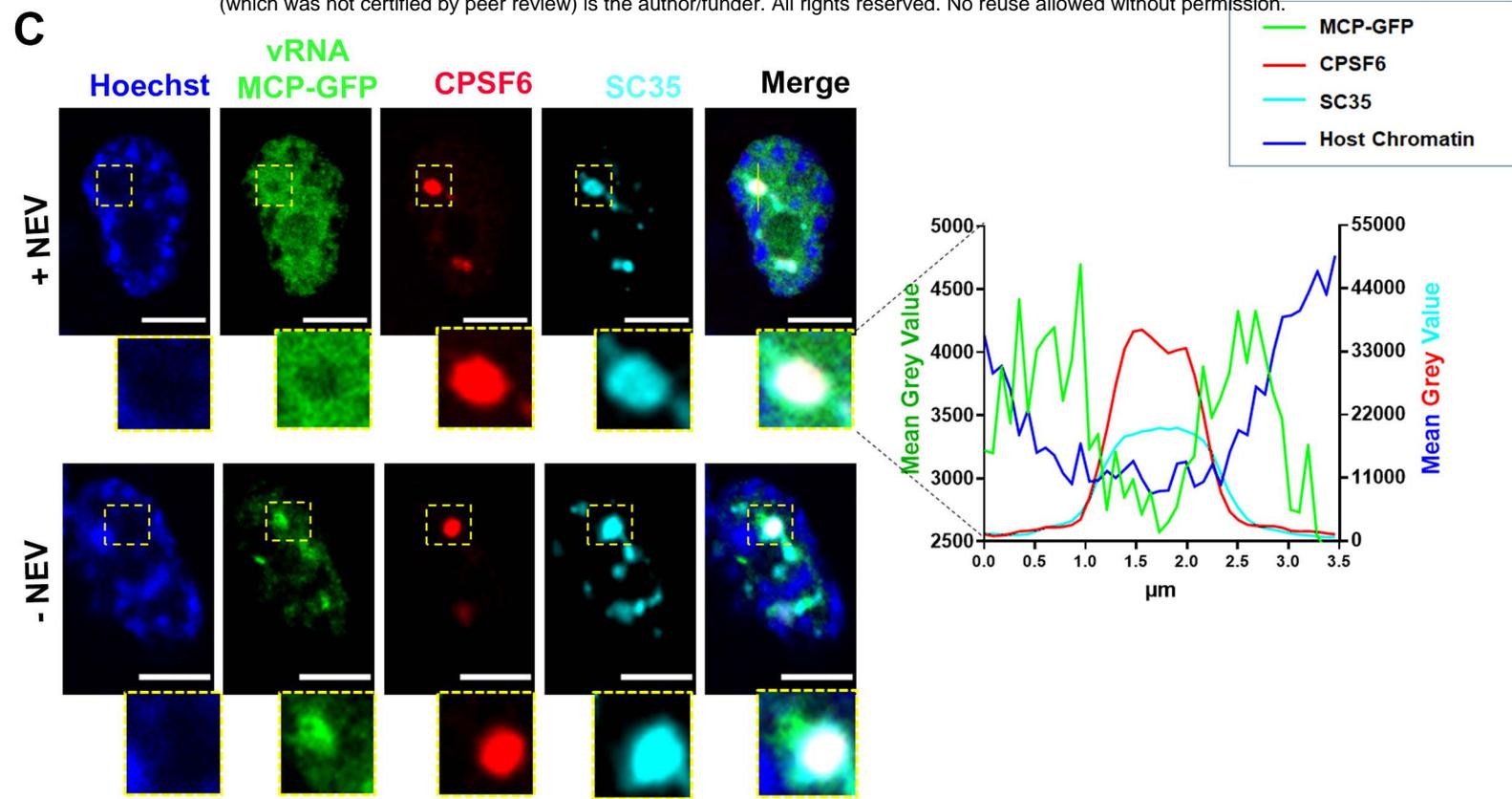
Figure 5
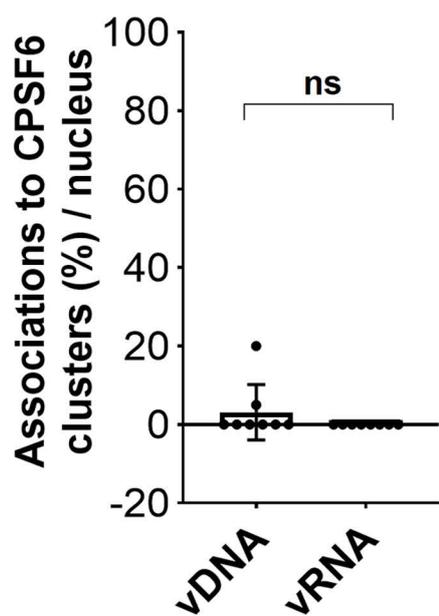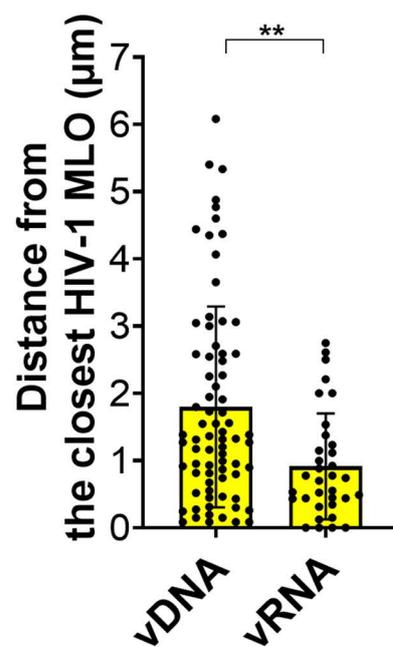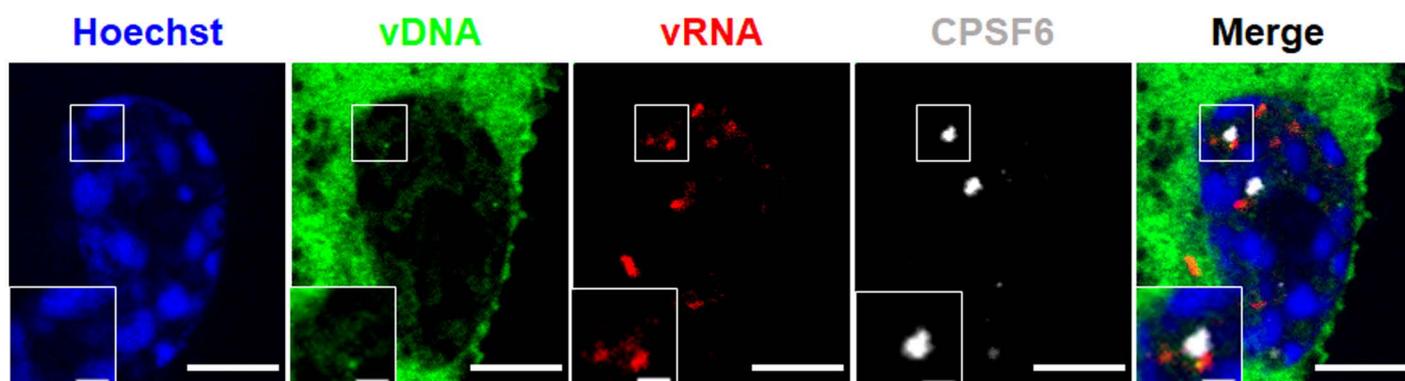
Figure 6

Figure 6

Figure 6

MDM 3 days p.i.

Figure 7

Figure 8

Figure 9