

Direct tracking of single proviruses reveals HIV-1/LEDGF complexes excluded from virus-induced membraneless organelles

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

▶ To cite this version:

Viviana Scoca, Marion Louveaux, Renaud Morin, Dmitry Ershov, Jean-Yves Tinevez, et al.. Direct tracking of single proviruses reveals HIV-1/LEDGF complexes excluded from virus-induced membraneless organelles. 2020. pasteur-03214329v1

HAL Id: pasteur-03214329 https://pasteur.hal.science/pasteur-03214329v1

Preprint submitted on 4 Mar 2021 (v1), last revised 1 May 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Direct tracking of single proviruses reveals HIV-1/LEDGF complexes excluded from virus-induced membraneless organelles

4

Viviana Scoca^{1, 2}, Marion Louveaux^{*3, 4}, Renaud Morin^{*5}, Dmitry Ershov^{4, 6}, Jean Yves Tinevez⁴, Francesca Di Nunzio¹

- 7
- ¹Department of Virology, Unit of Molecular Virology and Vaccinology, Pasteur Institute,
 Paris, France
- 10 ² BioSPC Doctoral School, Université de Paris, Paris, France
- ³ Department of Cell Biology and Infection, Biological Image Analysis, Pasteur Institute,
 Paris, France
- 13 ⁴ Image Analysis Hub/ C2RT, Pasteur Institute, Paris, France
- ¹⁴ ⁵ Imactiv-3D, 1 Place Pierre Potier, 31106 Toulouse, France
- 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: HIV-1 MLOs retain integrase and release viral DNA
- 20 Abstract

21 The ultimate goal of HIV-1 is integration into the host chromatin to optimize the release 22 of high levels of viral progeny and discretely coexist with the host. To uncover the HIV-23 1 DNA fate in the nuclear landscape we directly tracked the viral DNA (vDNA) and the 24 viral RNA (vRNA) by coupling HIV-1 ANCHOR technology with RNA FISH or MCP-25 MS2 RNA-tagging bacterial system. Our computational imaging analysis revealed that proviral forms are early located in proximity of the nuclear periphery of mitotic and non-26 27 mitotic cells. We also observed that HIV-1 infection prompts clustering formation of the host factor CPSF6 restructuring membraneless organelles enriched in both viral 28 29 proteins and speckle factors. Interestingly, we observed that integrase proteins are retained in CPSF6 clusters, while the late retrotranscribed DNA was excluded from 30 HIV-induced membranelless organelles (HIV-1 MLOs), indicating that those structures 31 are not proviral sites, but orchestrate viral events prior to the integration step. HIV-1 32 MLOs are in the vicinity of pre-existing LEDGF clusters. Importantly, we identified 33 actively transcribing proviruses localize, outside HIV-1 MLOs, in LEDGF-abundant 34 35 regions, known to be active chromatin sites. This study highlights single functional host-proviral complexes in their nuclear landscape, which is markedly restructured by 36 37 HIV-1 to favor viral replication.

38

39 Introduction

40 Immediately after fusion at the plasma membrane, HIV-1 cores are released into the cytoplasm and move towards the nucleus, while the viral RNA genome begins the 41 42 process of reverse transcription into double-stranded DNA (dsDNA) (Campbell and 43 Hope, 2015; Di Nunzio, 2013). Once imported in the nucleus, the interplay between the HIV-1 DNA genome and the host chromatin compartment is crucial for the fate of 44 the virus-host coexistence. HIV-1 can adopt either an episomal or a proviral form 45 46 (Butler et al., 2001). The choice between these two possibilities and their surrounding 47 chromatin landscape dictates the evolution of HIV infection (Maldarelli, 2016; Sharkey 48 et al., 2011; Sharkey et al., 2000). The mechanisms determining the fate of the proviral 49 DNA in the nucleus are still under investigation (Liu et al., 2020; Olson et al., 2019) 50 also because of the limits imposed by available technologies. Real-time imaging 51 approaches provide new insight into unprecedented spatial information of individual 52 HIV-1 infected cells. Several strategies have enabled the visualization of the viral components during nuclear entry: Integrase (IN)-TC/FIAsH, APOBEC 3F (A3F)-YFP, 53 54 IN-YFP, Gag-iGFP, Cyclophilin A (CypA)-DsRed/ Capsid (CA), INsfGFP, and CA-GFP 55 (Burdick et al., 2017; Hubner et al., 2009; Lelek et al., 2012) (Francis et al., 2016; Francis and Melikyan, 2018) (Mamede et al., 2017) (Albanese et al., 2008; Francis et 56 57 al., 2014). Previously, biochemistry studies supported only a limited and secondary role of the viral CA in the early steps of infection (Farnet and Haseltine, 1991). State-58 59 of-the-art imaging technologies developed in recent years have shed light on the central role of the viral CA in nuclear import(Blanco-Rodriguez et al., 2020; Burdick et 60 al., 2020) dictated by the CA interplay with host factors, such as CPSF6 (Buffone et 61 al., 2018; Francis et al., 2020; Price et al., 2014) and Nup153 (Di Nunzio et al., 2013; 62 63 Lelek et al., 2015). Although these techniques improved our knowledge in the HIV field, the labeling of viral proteins only allows the visualization of steps prior to the 64 65 establishment of proviral or episomal forms. Mechanisms underlying viral post-nuclear entry steps and the role of nuclear structures induced by HIV-1 during viral replication 66 could be revealed only by the direct co-labeling of viral DNA (vDNA) forms and host 67 factors. Thus far, the vDNA can be visualized either by supplying 5-Ethynyl-2'-68 deoxyuridine (EdU) during infection (Peng et al., 2014), which is limited to fixed non-69 dividing cells (De Wit et al., 2019) or by DNA-FISH (Marini et al., 2015), which is 70 71 characterized by harsh conditions of fixed sample preparation that limits its coupling

with other techniques, impeding the study of the interplay between the vDNA and the 72 73 host factors. Here, we were able to live track episomal and proviral forms, showing 74 their divergent intranuclear behavior. In parallel, we detect the location of functional 75 proviruses in the nuclear space of the main target cells of the virus, such as CD4⁺T cells and macrophages, by using HIV-1 ANCHOR methodology (Blanco-Rodriguez et 76 77 al., 2020) to label the vDNA coupled to RNA FISH or to the live bacterial approach 78 MCP-MS2 (Tantale et al., 2016) to identify the vRNA foci. Next, we deepened the study of the nuclear landscape that surrounds proviruses. We observed that viral infection 79 80 reprogramed the nuclear location of a viral partner, CPSF6. CPSF6 is a paraspeckle 81 factor (Fox et al., 2002; Naganuma and Hirose, 2013) involved in viral nuclear entry 82 though its interaction with the viral CA (Buffone et al., 2018; Francis et al., 2020; Lee 83 et al., 2010; Price et al., 2014). Here we found that CPSF6 relocates in HIV-1 induced 84 membraneless organelles (HIV-1 MLOs), enriched in speckle factors, SC35 (Spector and Lamond, 2011) and SON (Sharma et al., 2010), near LEDGF clusters. Previous 85 86 studies reported the importance of LEDGF in HIV-1 integration sites distribution (Ciuffi et al., 2005), however a complex between the transcribing proviral DNA and LEDGF 87 88 has never been visualized. In this study, we determined the nuclear position of several 89 viral forms, including actively transcribing proviruses, with respect to HIV-1- MLOs or 90 the nuclear envelope (NE). In addition, we found that the late retro-transcribed DNA 91 co-exists, but remains separated, from viral INs in the host nucleus. We observed that 92 the viral INs are sequestered in HIV-1 MLOs, suggesting the role of these nuclear 93 structures in the Pre-Integration Complex (PIC) maturation. On the other hand, the late reverse transcripts are fully excluded from HIV-1 MLOs, in fact we found that HIV-1 94 95 MLOs are not the sites of viral integration. Finally, we identified functional host-viral complexes, formed by vDNA/vRNA associated to LEDGF, excluded from the inner 96 97 cavity of HIV-1 MLOs (avg distances from HIV-1 MLOs < 0.6 μ m) and located in the 98 vicinity of the NE (avg of distances from NE < $2 \mu m$). In this study we show that HIV-1 99 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 results provide new insights into how HIV 102 reprograms and markedly restructures the nuclear environment to orchestrate viral 103 replication steps.

105 **Results**

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

107 With the aim of studying HIV-1 DNA nuclear fate in different cell types, we benefited of 108 the HIV-1 ANCHOR DNA labeling system (Fig. 1A) (Blanco-Rodriguez et al., 2020). 109 Once fully retrotranscribed, the HIV-1 genome carrying the bacterial sequence ANCH3 110 is specifically recognized by OR-GFP protein, which is expressed in the target cells, genetically modified by the lentiviral vector (LV) OR-GFP (Blanco-Rodriguez et al., 111 112 2020). The accumulation of OR-GFP, which is a modified version of the bacterial ParB 113 protein (Graham et al., 2014; Sanchez et al., 2015), on the ANCH3 sequence 114 generates bright nuclear signals (Fig.1A). Importantly, OR protein binds exclusively double stranded DNA (Saad et al., 2014). We cloned ANCH3 in the nef gene, thus, 115 116 HIV-1 ANCHOR system exclusively detects late reverse transcripts, because only late 117 retrotranscripts contain ANCH3 sequence as double stranded DNA (Fig. 1A). Therefore, HIV-1 ANCHOR system allows only the visualization of potential functional 118 119 vDNA. We optimized and set up our imaging approach in HeLa P4R5 cells (Charneau 120 et al., 1994), a widely used cellular model for HIV-1 infection. Previous studies showed 121 that HeLa cells enable single particle tracking of HIV-1 proteins (Burdick et al., 2020; 122 Francis and Melikyan, 2018). However, since HIV-1 genome dissociates from viral proteins to become a proviral DNA or an extrachromosomal DNA, the track of these 123 124 viral forms is the only direct measurement to investigate the fate of the viral genome 125 into the host nuclear space. Unlike previous studies, we are now able to directly 126 visualize by live imaging the nuclear vDNA as GFP puncta appearing in the host 127 nucleus once retrotranscribed (Fig. 1A). On the contrary, bright individual spots are 128 rarely visualized in the cytoplasm, where the vDNA is only partially accessible to OR-129 GFP, as we previously demonstrated (Blanco-Rodriguez et al., 2020). HIV-1 ANCHOR 130 is able to specifically label the vDNA as shown by the correlation between the nuclear GFP puncta and the MOI (multiplicity of infection) used (Fig. S1A). This specificity was 131 132 corroborated using the reverse transcription inhibitor, Nevirapine (NEV), which completely prevents the detection of vDNA (Fig. 1B, S1B). We then investigated 133 134 whether HIV-1 ANCHOR allows the detection of different intranuclear vDNAs, such as 135 episomal and integrated forms. To generate episomal forms, either we infected cells 136 with HIV-1 ANCH3 in the presence of the integration inhibitor Raltegravir (RAL), or using an integration-deficient virus, which is mutated at the catalytic site of the 137

138 integrase (IN) enzyme (HIV-1 ANCH3 IN_{D116A}). We detected intranuclear GFP puncta 139 under both conditions (Fig. 1B), meaning that HIV-1 ANCHOR permits the detection of 140 nuclear episomal vDNA forms. The experiments were validated by gPCR, which 141 confirmed the extrachromosomal forms generated during infection when the 142 integration step is impeded (Fig. 1B). We next assessed whether HIV-1 ANCHOR is 143 sufficiently sensitive to allow the detection of integrated vDNA at the level of a single 144 provirus. We selected single HeLa P4R5 clones carrying the HIV-1 ANCH3 genome. The imaging data correlates with results obtained by ALU PCR (Fig. 1C), 145 146 demonstrating that this technology is powerful enough to detect a single provirus. 147 Thus, we used this system to live track proviruses and episomal forms to investigate 148 their behaviour in the nucleus (movies S1, S2, and S3). We performed live tracking of 149 vDNAs in cells at six days post infection (p.i.). (movie S1), which contain a negligible 150 amount of episomal forms, likely lost during cell division (Fig. S1C), in the single-cell 151 clone (movie S2) or in cells containing only the episomal vDNA (movie S3). Then, we 152 compared vDNA trajectories in the aforementioned experimental conditions (Fig. 1D). The stable clone and the cells at six days p.i. show vDNA with similar trajectories and 153 154 diffusion coefficient (D (nm^2/s) < 5). Instead, infected cells carrying unintegrated 155 viruses harbored two populations composed of episomal forms with divergent 156 behaviors, one with a low D (D (nm^2/s) < 5) and the other with a high D (5 < D (nm^2/s)) 157 < 20) (Fig. 1D). Our data hint that most, but not all, episomal forms are retained in 158 particular nuclear regions, most likely interacting with host factors or chromatin 159 compartments, consistent with the findings of other studies (Geis and Goff, 2019; Zhu 160 et al., 2018), whereas some episomal forms are free to move in the nuclear space. On 161 the other hand, the integrated forms showed a uniform, low-diffusive behavior.

162

163 Detection of actively transcribing viral DNA at single-cell level in HIV-1 target164 cells.

The identification of functional vDNA in the nucleus of infected cells has been one of our goals. For this purpose, we combined HIV-1 ANCHOR technology with RNA FISH (single molecule inexpensive FISH, smiFISH) against HIV-1 polymerase transcripts, to detect vRNA foci of transcription (Tsanov et al., 2016). As expected, in cells treated with RAL, where the integration step is blocked, vRNA signal is highly depleted. In untreated cells it is possible to observe vDNA-vRNA associations, highlighting the proviruses location (Fig. 2A), in fact proviruses are the only transcribing forms. We

172 strengthen our results on the ability to visualize proviruses by showing the vDNA-vRNA 173 association by live imaging. HIV-1 transcriptional activity was tracked by MCP-GFP 174 system (kind gift from E. Bertrand) to label vRNA (Tantale et al., 2016) coupled with 175 HIV-1 ANCHOR technology (Fig. S2A, B). Next, we compared the intensity of 176 ANCHOR signals of the episomal forms (RAL) with those of the integrated forms 177 (untreated cells) and observed that the signal of unintegrated DNA was two-fold 178 brighter than the proviral DNA spot (Fig. 2A). Interestingly, we also remarked, during live imaging, that the vDNA of the integration-deficient virus (HIV-1 ANCH3 IN_{D116A}) 179 180 aggregated to form "flower-shaped structures" in the nucleus (Fig. S3, movie S4). 181 Aggregation of extrachromosomal forms could prompt a brighter appearance of viral 182 episomal DNA compared to a dimmer signal of the individual proviral DNA (Fig. 2A). 183 Next, we tested the HIV-1 ANCHOR system in the context of the main target cell of 184 HIV-1, primary CD4⁺ T cells, derived from healthy donors. CD4⁺ T cells challenged with 185 various MOIs of HIV-1 ANCH3 showed that the number of nuclear spots increased 186 with the increasing of the MOI (Fig. S4), meaning that HIV-1 ANCHOR specifically detects the vDNA in the nucleus of primary lymphocytes. The correlation between 187 188 different MOIs used and the number of proviruses (r=0.9412), analysed by qPCR, 189 indicates that OR-GFP does not impede the viral integration step (Fig. S4). We tested, 190 as well, the co-detection of viral transcriptional foci and vDNA in T lymphocytes 191 (primary CD4⁺ T and Jurkat cells) through RNA FISH (Fig. 2B, Fig. S5). We observed 192 that not all vDNAs associated to a transcriptional focus at 48 h p.i., most likely because 193 of the presence of silent episomal forms. On the other hand, the majority of cells 194 analyzed contain vRNA foci colocalized with vDNA (Fig. 2B), meaning that we are able 195 to detect single transcribing proviruses. We also performed similar experiments in non-196 dividing cells, THP-1, which are macrophage-like cells. Thus, we imaged and analyzed 197 THP-1 cells at 3 days p.i. and we observed that also in these cells we could distinguish 198 individual transcribing proviruses from silent viral forms (Fig. 2C). Of note, the RNA 199 signal considered in our images was exclusively due to the detection of viral 200 transcription foci, since samples treated with NEV did not show any significant intra-201 nuclear signal (Fig.S6). Taken together, our results suggest that we are able to 202 distinguish episomal forms (mainly silent) from integrated forms that are the only ones 203 able to replicate.

Nuclear distance of actively transcribing proviruses from the nuclear envelope at single-cell level in dividing and non-dividing cells.

206 The location of actively transcribing proviruses in the nuclear space of dividing cells is 207 still a subject of debate (Achuthan et al., 2018; Burdick et al., 2017; Chin et al., 2015; 208 Di Primio et al., 2013; Lelek et al., 2015; Marini et al., 2015), whereas their nuclear 209 location in terminally differentiated cells, as far as we know, has never been studied 210 before. In dividing cells, the 3D analysis of the nuclear location of the vDNA showed 211 the vDNA signals at 48 h p.i. to be more randomly distributed throughout the nucleus 212 compared to the vDNA spots at seven days p.i. (Avg distances from the NE: ~2.5 µm at 48 h p.i. vs ~1.4 µm at 7 days p.i.) (Fig. 3A). The difference in the average distance 213 214 of the vDNA spots from the NE between the early and late time points p.i. may be due 215 to the co-existence of heterogeneous viral forms at 48 h p.i. (episomal and proviral 216 forms), whereas the vDNA population is more homogeneous at seven days p.i. (Fig. 1D, 3A). In fact, at seven days p.i., almost all remaining vDNAs in the nucleus consisted 217 218 of integrated forms, as suggested by the imaging of dividing cells following prolonged 219 treatment with RAL (Fig. S1C). Interestingly, we did not detect differences between the 220 distances of the vRNA foci from the NE (~ 1.5μ m) at both time points (48h p.i. and 7d 221 p.i.) and the vDNA spots at seven days p.i. (~ 1.4 μ m), indicating that in dividing cells 222 the integration of the proviral DNA occurs early during infection near to the NE (Fig. 223 3A). We performed similar 3D analysis to calculate the distances of vDNAs and vRNAs 224 from the NE in non-dividing cells at three days p.i.. The lack of mitosis hinders the loss 225 of the episomal forms in non-dividing cells, so we could not discriminate between 226 integrated and unintegrated vDNA forms. However, when we compared vDNAs 227 (episomal and proviral forms) versus vRNA foci (actively transcribing proviral forms) 228 their analysis of distances from the NE revealed a major vicinity of the proviruses to 229 the NE compared to the total vDNA (~2.2 μ m for the vDNA vs ~1.8 μ m for the vRNA) 230 (Fig.3B). Overall, our data show a similar nuclear spatial distance of actively 231 transcribing proviral forms from the NE at single-cell level in both, dividing and non-232 dividing cells (distance of vRNA foci from the NE: \sim 1.7 µm in HeLa cells and \sim 1.8 µm in THP-1 cells) (Fig.3C). 233

The majority of IN proteins are retained in CPSF6 clusters and separate from the late retrotranscribed DNA

A long-standing question in the early steps of HIV life cycle is whether viral IN proteins 236 237 remain all associated with late reverse transcribed DNA for long time or only for a short 238 time after the nuclear entrance. This association may not be detectable because of the 239 short-time nature of the event or due to the technical limitations of vDNA and IN 240 visualization. Thus, we exploited our vDNA tracking system to investigate about IN 241 proteins localization in primary CD4⁺T cells in comparison to terminally differentiated cells, THP-1. We challenged both cell types with HIV-1 ANCH3 and we analysed them 242 at 20h and 30h p.i., respectively. We found that in both cells the vDNA/IN association 243 244 event was very rare (only one single event in 1 cell for THP-1 and in 3 cells for CD4+T 245 cells) (Fig. 4A). Nuclear IN protein clusters were found in the nucleus, but not at the 246 site of the late reverse transcripts (Fig. 4A). To confirm our data we performed live-247 imaging, asking whether we could pinpoint the dynamic evolution of the association 248 between the vDNA and the IN in the nucleus of infected cells by coupling the HIV-1 249 ANCHOR system with the GIR virus (generous gift from Edward Campbell)(Dharan et 250 al., 2016; Hulme et al., 2015). We produced the GIR virus using a viral genome carrying 251 the IN with a mutation in the catalytic site (HIV-1 ANCH3 IN_{D116A}), as consequence, 252 the only active IN in these viral particles was the IN-Ruby (Fig. S7). Analysis of the 253 dynamics of the vDNA / IN association shows the two signals to be separated (~ 0.6 254 to 0.8 µm) within the nucleus for several hours (Fig. S7, movie S5). Our results suggest 255 that, after nuclear entry, the main localization of IN is not at the exact site of the viral 256 DNA forms, even though we can sometimes spot them in proximity. Next, we 257 investigated whether the viral IN was retained in particular nuclear regions by some 258 nuclear host factors. We first found that the viral IN occupies the same positions as the 259 viral CA and its host nuclear partner, CPSF6 (Fig. S8A). We confirmed that HIV-1 reprograms the nuclear localization of CPSF6 from being randomly dispersed in the 260 261 nucleoplasm to form particular aggregates in both primary lymphocytes and (Rensen 262 macrophages (Francis et al., 2020) et al., BioRxiv 263 doi.org/10.1101/2020.04.12.038067) (Fig. 4B). We also observed that CPSF6-IN-CA 264 clustering is independent of DNA synthesis (Fig. S8B). Then, we quantified the 265 percentage of association per nucleus between the viral IN and CPSF6. We found an average of association of IN to CPSF6 ~ 94.4% in CD4+T cells and ~92.5% in THP-1 266 267 cells (Fig. 4B). Taken together our results show that in dividing and in non-dividing cells there is an intra-nuclear separation of the viral DNA and of the IN proteins. 268 269 Interestingly, the viral IN proteins persist after viral nuclear entry associated to CPSF6 clusters. This observation suggests that the late reverse transcripts are released from
the majority of IN proteins which instead are retained in a separated nuclear location
occupied by CPSF6. Thus, CPSF6 clusters could serve as sites to release a mature
PIC that could contain just few IN proteins necessary for the integration step
(Ballandras-Colas et al., 2017; Hare et al., 2009).

275 **Proviruses are excluded from HIV-1-induced membraneless organelles.**

276 CPSF6 participates to the viral nuclear translocation and post-nuclear entry steps by 277 interacting with the viral CA (Bejarano et al., 2019; Burdick et al., 2017; Lee et al., 2010; Price et al., 2014) (Francis et al., 2020). Recent studies reported a high level of 278 279 co-localization between the paraspeckle factor, CPSF6, or the speckle factor and 280 splicing regulator, SC35, (Spector and Lamond, 2011) and the vDNA labeled with EdU 281 (Francis et al., 2020) (Bejarano et al., 2019)(Rensen et al., BioRxiv 282 doi.org/10.1101/2020.04.12.038067). Importantly, it has been shown that EdU labels 283 nuclear clusters composed of multiple viral episomal forms (Francis et al., 2020) 284 (Rensen et al., BioRxiv doi.org/10.1101/2020.04.12.038067). Therefore, EdU labeling 285 can detect multiple nascent vDNAs but it is likely not sufficiently sensitive to detect an individual viral genome. Contrary to EdU labeling, the HIV-1 ANCHOR system allows 286 287 the visualization of a single provirus (Fig.1C) and does not interfere with viral 288 (Rensen transcription (Blanco-Rodriguez et al., 2020) al.,BioRxiv et 289 doi.org/10.1101/2020.04.12.038067). Thus, we used HIV-1 ANCHOR technology to 290 visualize the late retrotranscripts in infected macrophage-like cells at 3 days p.i.. We 291 observed that late retrotranscripts do not colocalize with CPSF6 (Fig. 5A). However, 292 upon viral infection, CPSF6 relocates and fully colocalizes with speckle-simile 293 organelles labeled with SC35 (Fig. 5A). SC35 is strongly associated to another speckle 294 factor, SON, independently of viral infection (r pearson ~0.995). More specifically, we 295 found that the nuclear reorganization of CPSF6 is not required for the integration step. 296 as the nuclear organization of CPSF6 was similar in cells treated or not with RAL (Fig. 297 S9). Then, we investigated whether the vDNA colocalizes with SC35. We computed 298 vDNA distances in 3D space from the boundary of SC35 bodies in cells treated with 299 RAL (avg ~ 0.4 μ m (Fig. S9)) or in absence of RAL (avg ~ 0.6 μ m (Fig. 5A)). Therefore, 300 similarly to CPSF6 also SC35 clusters exclude late reverse transcripts (Fig. 5A). In 301 addition, we found that viral transcriptional foci where also located outside 302 CPSF6/SC35 organelles (Fig. 5B). These results are independent of the presence of

vpx (Fig. S10) that we used to increase viral infectivity in THP-1 (Hrecka et al., 2011;
Laguette et al., 2011). Taken together analysis of the nuclear positions of vDNA
(episomal and proviral forms) and vRNA foci (proviruses) showed that these viral
nuclear forms are excluded from HIV-1 MLOs (Fig. 5C), indicating that those
organelles cannot be sites for viral integration or replication.

308

Late reverse transcripts DNA and viral RNA foci form complexes with LEDGF, spatially independent of CPSF6 clusters.

311 We extended our study on THP-1 cells at 3 days p.i. to investigate whether Lens 312 Epithelium-Derived Growth Factor (LEDGF), which is a partner of the viral IN and leads 313 HIV-1 integration sites selection (Cherepanov et al., 2003; Ciuffi et al., 2005; Emiliani 314 et al., 2005; Ferris et al., 2010; Hare et al., 2009; Lelek et al., 2015; Llano et al., 2006; 315 Shun et al., 2008; Shun et al., 2007; Wang et al., 2014), is associated to the proviruses 316 forming a complex visible by imaging. We observed that CPSF6 proteins cluster in 317 proximity of LEDGF sites (~ 43%), without co-localizing (Fig. 6A). There is a similar number of LEDGF clusters in both infected and uninfected cells (Fig. 6A), while CPSF6 318 319 clusters are mainly formed upon infection, suggesting that CSPF6 is prompted by the 320 virus to localize near the pre-existing LEDGF regions (Fig.6A). Of note, the PWWP 321 domain of LEDGF has been shown to interact with H3K36me3 (Pradeepa et al., 2014; 322 Pradeepa et al., 2012), which is a marker of euchromatin. In particular, this post-323 translational modification is enriched in HIV-1 integration sites (Ciuffi et al., 2005; Lelek 324 et al., 2015; Vansant et al., 2020; Wang et al., 2009). LEDGF also interacts with 325 splicing factors (Singh et al., 2015). Therefore, we analysed the location of proviral foci of transcription in relation to LEDGF and CPSF6. Our previous data shows that CPSF6 326 327 co-localizes with SC35, however the distance between the vDNA and SC35 has been 328 estimated to be ~ 0.6 μ m (Fig. 5A), therefore, the late retrotranscribed DNA is too 329 distant from those HIV-1 MLOs to be considered part of the same complex. While, the 330 vDNA of the late retrotranscripts is mostly excluded from CPSF6 clusters, we found 331 that ~64% of the vDNA and ~73% of the vRNA foci associate with LEDGF (Fig. 6B). 332 Next we have characterized in more detail the viral genome associations with LEDGF 333 and CPSF6. We calculated that ~ 87.9% of the vDNA/LEDGF complexes and ~80% 334 of vRNA foci/LEDGF complexes are uncoupled to CPSF6 (Figure 6B). We found only few vDNA (~12.1%) and vRNA (~20%) foci associated to LEDGF clusters coupled to 335 336 CPSF6 (Fig. 6B). Our results suggest that the two cellular factors, CPSF6 and LEDGF.

orchestrate two sequential steps of viral replication, the maturation of the PIC and its
integration. Our data indicate that the functional vDNAs are not located inside
CPSF6/SC35 clusters, but in external active chromatin domains, enriched in LEDGF,
likely to benefit of a favorable environment for viral replication.

341

342 **Discussion**

351 Studies of parasites that permanently integrate into the host genome show that they 352 have evolved to target chromatin in such a way as to optimize their coexistence with 353 the host and favor the release of new viral progeny (Bushman, 2003; Craig and 354 Marszalek, 2002). HIV-1 needs to rapidly release high levels of newly generated 355 viruses after infection, intuitively explaining the preference of active genes as 356 integration targets (Ciuffi et al., 2005; Di Nunzio et al., 2013; Lelek et al., 2015; Maldarelli, 2016; Schroder et al., 2002). However, the nuclear location of individual 357 358 proviruses actively transcribing at the dividing or non-dividing single cell level has 359 never been visualized before. Thus far, the direct detection of HIV-1 genomes after 360 reverse transcription and the distribution of proviral transcriptional foci has been 361 technically challenging. Recent studies have shown the possibility to follow HIV 362 infection in live using surrogate viruses (IN GFP virus (Albanese et al., 2008), a mix of viruses carrying fluorescent and non-fluorescent IN proteins (Hulme et al., 2011), or 363 364 indirectly follow the viral journey by labeling host viral partners, such as CypA (Francis 365 and Melikyan, 2018) or APOBEC 3F (Burdick et al., 2017)). All these approaches allow 366 to follow viral or host proteins but not viral replication complexes. HIV-1 ANCHOR 367 system allows the live tracking of different nuclear late retrotranscribed DNA forms (Fig.1A), which makes possible to probe their behavior via the measurement of the 368 369 mean-squared displacement (MSD). We observed a homogenous population formed 370 by proviral DNA with low diffusion coefficient. Episomal forms showed a 371 heterogeneous behavior, characterized by two different populations, implying that 372 these two viral forms can interact with different factors in the nuclear space (Fig. 1D). 373 Moreover, imaging analysis showed that viral episomal forms identified by the 374 accumulation of OR-GFP on ANCH3 sequences gave rise to brighter signals than 375 proviral forms (Fig. 2A). This phenomenon could account for the differences in 376 behavior between integrated and unintegrated forms. Thus, unintegrated forms can be 377 brighter for the following reasons: a) the absence of transcriptional machinery

378 associated with the vDNA may reduce the steric hindrance that can interfere with the 379 accumulation of OR-GFP on the HIV-1 tag sequence (ANCH3) (Fig. 2A), b) episomal 380 forms can aggregate in a flower shape, generating a brighter signal (Fig. S3, movie 381 S4), c) episomal forms are differently chromatinized than the integrated vDNAs, 382 becoming more accessible to OR-GFP (Geis and Goff, 2019). Next, to track the nuclear location of functional proviruses we co-labeled vDNA and vRNA by combining 383 384 HIV-1 ANCHOR technology with RNA FISH in the main target cells of HIV-1, CD4⁺T cells and macrophages (Fig.2A,B,C). We also corroborated results on the visualization 385 386 of functional proviruses using a live track method, MCP-MS2 system (Tantale et al., 387 2016), coupled to HIV-1 ANCHOR (Fig.S2). Once the performance of our approach for 388 labeling actively transcriptional proviruses was validated, we investigated the spatial 389 location of proviruses in dividing and non-dividing cells. Three-dimensional analysis of 390 the distribution of vDNAs during early (48 h p.i.) infection in mitotic cells displays that 391 they are more randomly distributed than the proviral DNA from the single clone or from 392 cells at 7days p.i. (Fig. 3A). We also observed that the location of viral transcription 393 sites (vRNA) is highly similar between early and late time points of infection, indicating 394 that proviral DNAs, the only transcribing forms (Fig. 2A), immediately localize to near 395 the NE (~1.5 μ m) early after infection (Fig. 3A) in dividing and non-dividing cells 396 (Fig.3B,C). We next investigated the interplay among viral components and their 397 association with host factors during post-nuclear entry steps. We corroborated our 398 published data (Rensen et al.,BioRxiv doi.org/10.1101/2020.04.12.038067) showing 399 that IN proteins strongly co-localize with CPSF6 clusters in macrophage-like cells and 400 in this study we observed a similar association between IN and CPSF6 also in CD4⁺ T 401 lymphocytes at early time of infection (Fig. 4B). Contrary to CPSF6, the late vDNA 402 products of retrotranscription rarely co-localize with IN in both types of cells (Fig. 4A). 403 These results are in line with the dynamic coexistence of the IN and vDNA, which 404 remain separated but in the proximity for several hours, as it has been shown by real-405 time imaging in HeLa cells (Fig. S7, movie S5). This result can appear in contrast with 406 recent data from other groups and from us obtained with another DNA labeling 407 technology, EdU (Francis et al., 2020) (Rensen et al., BioRxiv, 2020), showing the 408 vDNA and the IN strongly colocalized with CPSF6. This apparent discrepancy can be 409 explained by the intrinsic features of vDNA labeling tools. EdU system is capable of detecting nascent episomal forms of vDNA abundant in CPSF6 clusters, sites of 410 411 nuclear reverse transcription (Rensen et al.,BioRxiv

doi.org/10.1101/2020.04.12.038067). Of note, EdU labeling cannot provide direct 412 413 information on viral replication, as EdU reduces viral transcription (Rensen et al., 414 BioRxiv doi.org/10.1101/2020.04.12.038067). Contrariwise, HIV-1 ANCHOR system 415 allows the detection of late reverse transcripts (Fig.1A) that can generate functional 416 proviral DNA, as we observed in HeLa cells, human lymphocytes, and macrophage-417 like cells (Fig. 2A,B,C, Fig.S2A,B, Fig. S5). Thus, we combined HIV-1 DNA and HIV-1 RNA labeling to detect the nuclear landscape where functional viral forms are located. 418 Interestingly, we observed that HIV-1 induces the remodeling of speckles 419 420 characterized by the presence of SC35 and SON, in novel MLOs enriched in viral 421 proteins, such as IN and CA, as well in cellular factors, like the paraspeckle member, 422 CPSF6, (Fig. 4B, 5A,B, S8). Of note, we observed that CPSF6 is fully included inside 423 organelles marked by SC35 during viral infection, otherwise CPSF6 is usually 424 randomly dispersed in the host nucleus of uninfected cells (Fig.4B, Fig.5A). Likely, 425 CPSF6 relocates to speckles prompted by the virus to usurp functions linked to these 426 MLOs. We show that the formation of CPSF6 cluster is independent of viral DNA 427 integration or viral DNA synthesis (Fig. S8B, S9) and that late reverse transcribed 428 products are excluded from them (Fig. 5A). It is possible that once the vRNA is fully 429 retrotranscribed in vDNA (Rensen et al., BioRxiv doi.org/10.1101/2020.04.12.038067), 430 it will quickly leave the HIV-1 MLOs. In fact, vDNAs were located on average at a 431 distance of ~ 600 nm from the boundary of SC35 (Fig. 5A). MCP-labelled vRNA foci 432 shows that actively transcribing proviruses are also located outside HIV-1 MLOs, 433 indicating that those are not sites of viral integration and replication (Fig.5B,C). We 434 also observed that CPSF6 relocates to the vicinity of LEDGF during HIV infection (Fig. 6A). LEDGF is required for the targeting of highly-spliced transcriptional units through 435 436 its direct interaction with several splicing factors (Singh et al., 2015). LEDGF is also 437 known to determine HIV-1 integration sites distribution in active chromatin sites (Ciuffi 438 et al., 2005; Ferris et al., 2010; Shun et al., 2007), however the visualization of a 439 complex formed by the transcribing provirus and LEDGF remained utopian for a long 440 time. Here, we were able to visualize by fluorescence microscopy a complex formed by the vDNA or vRNA and LEDGF. More than 80% of those complexes are formed by 441 442 LEDGF uncoupled to CPSF6 (Fig.6B). However, we observed that less than 20% of 443 them can include CPSF6 associated to LEDGF, suggesting that few proviruses need to be near to HIV-1 MLOs without co-localizing, against the majority of them that can 444 445 replicate farther (Fig. 6B). Whether there is a functional difference between these two

complexes it can be further investigated in future studies. Overall, our study highlights 446 447 the proviral location in the nuclear landscape remodelled by the viral infection (Fig.7). 448 We found that HIV-1-MLOs retain the viral IN but exclude the late retrotranscribed 449 DNA, indicating a role for HIV-1 MLOs prior to the integration step. In particular, we 450 found foci of viral transcription outside HIV-1 MLOs, located in the nuclear periphery, 451 in complex with LEDGF, indicating that HIV-1 MLOs cannot be sites of viral integration 452 and viral transcription. Furthermore, the visualization of the formation of an actively transcribing provirus in complex of a pre-existing cluster of LEDGF, which is known to 453 454 lay on active chromatin sites (Ciuffi et al., 2005; Singh et al., 2015) (Fig.7), indicates 455 that LEDGF is not part of the viral complex during nuclear entry (Fig.6A). Taken 456 together our results show that functional proviruses establish near to the NE, excluded 457 from CPSF6 clusters, but locate in a favorable transcriptional environment ~ 600nm 458 from HIV-1 MLOs (Fig.7). Our study supports how new single-cell level approaches 459 are pivotal in the study of functional viral nuclear dynamics, to discriminate the cells 460 susceptible to fuel viremia or that concur to viral persistence.

461

462 Materials and Methods

463 Cell lines, primary cells and generation of genetically modified single-cell464 clones.

465 HEK 293T cells (ATCC) are human embryonic kidney cells used to produce lentiviral 466 vectors and HIV-1 viruses. HeLa P4R5 reporter cells are HeLa CD4⁺ CCR5⁺ CXCR4⁺, 467 carrying the LacZ gene under the control of the HIV-1 LTR promoter (Charneau et al., 1994). HeLa MCP-GFP is a HeLa clone stably expressing MCP-GFP bacterial fusion 468 469 protein (kind gift from E. Bertrand) (Tantale et al., 2016). Jurkat cells (ATCC) are 470 human immortalized T lymphocytes derived from acute T cells leukemia. THP-1 cells 471 (ATCC) are human immortalized monocyte cells derived from acute monocytic 472 leukemia, which, once seeded, were differentiated into macrophage-like cells under 473 Phorbol 12-myristate 13-acetate (PMA) treatment (167nM). HEK293T cells and HeLa 474 cells were cultivated in DMEM medium supplemented with 10% Fetal Bovine Serum 475 (FBS) and 1% penicillin-streptomycin. Jurkat and THP-1 cells were cultivated in RPMI 1640 medium supplemented with 10 % Fetal Bovine Serum (FBS) and 1% penicillin-476 477 streptomycin. Primary CD4⁺T cells were purified from healthy donors' blood obtained

via EFS (Etablissement Français du Sang, Paris), through density gradient 478 479 centrifugation with Ficoll 400 and isolation of PBMCs buffy coat. Subsequent positive 480 selection was performed with human CD4⁺ Microbeads (Miltenyi Biotec #130-045-481 101). The day after CD4⁺ T cells were activated with T cell Activation/Expansion kit 482 (Miltenyi Biotec #130-091-441) according to protocol, complete RPMI 1640 medium 483 was supplemented with interleuvkin-2 (IL-2) (400 IU/mL). HIV-1 ANCH3 single-cell 484 clone was generated starting from HeLa P4R5 cells which were infected with HIV-1 485 ANCH3 virus at MOI 1. Cells were diluted to 1 cell per well, in 96-well plates. Cell-clone colonies were tested for β-galactosidase expression to check viral expression (kit 486 487 Roche #11758241001). Positive clones were transduced with lentiviral vector (LV) 488 CMV-OR-GFP for imaging analyses of HIV-1 ANCH3 provirus. All cells were kept in 489 incubator at 37°C and 5% CO₂.

490

491 Plasmids and lentiviral vector productions

492 HIV-1 Δ EnvIN_{HA} Δ Nef plasmid encodes for the Δ EnvHIV-1 LAI (BRU) viral genome 493 where the Integrase (IN) protein is fused to the hemagglutinin (HA) tag (kindly provided 494 by F. Mammano)(Petit et al., 1999; Petit et al., 2000). HIV-1ΔEnvIN_{HA}ΔNef ANCH3 495 (HIV-1 ANCH3) was obtained through ANCH3 insertion: ANCH3 sequence was cloned 496 by PCR using the template plasmid pANCH3 as we previously described in Blanco et 497 al. (Blanco-Rodriguez et al., 2020). The ANCHOR[™] technology and sequences are 498 exclusive property of NeoVirTech (Germier et al., 2017; Mariame et al., 2018). HIV-499 $1\Delta Env IN_{HA}(D116A) \Delta Nef ANCH3 was obtained by insertional mutagenesis using the$ 500 QuikChange II XL Site-Directed Mutagenesis kit (Agilent #200522), for integrase (IN) 501 mutation at D116. The lentiviral vector plasmids CMV OR-GFP and CMV OR-502 SANTAKA were obtained by cloning OR-GFP/OR-SANTAKA (plamids from 503 NeoVirTech) in a pTripCMV (Δ U3) plasmid through PCR and cloned using restriction 504 site, Agel and SgrDI. HIV-1ΔEnvIN_{HA}ΔNef ANCH3 MS2 plasmid was obtained by inserting the MS2x64 sequence amplified by PCR from pMK123-MS2x64 (Tantale et 505 506 al., 2016) in HIV-1ΔEnvIN_{HA}ΔNef ANCH3. Lentiviral vectors and HIV-1 viruses were produced by transient transfection of HEK293T cells through calcium chloride 507 508 coprecipitation. Co-transfection was performed as following, LV carrying OR-GFP/OR-509 SANTAKA: 10 µg of OR-GFP/OR-SANTAKA LV, 10 µg of NDK plasmid (gag-pol-tat-510 rev) and 2.5 µg of pHCMV-VSV-G envelope expression plasmid; LV carrying MCP-GFP (kindly provided by E. Bertand): 10 µg of MCP-GFP, 10 µg of NDK plasmid (gag-511

512 pol-tat-rev) and 2.5 μg of pHCMV-VSV-G envelope expression plasmid; HIV-1ΔEnv pseudotyped with VSV-G: 10 μ g HIV-1 Δ EnvIN_{HA} Δ Nef (IN WT or IN(D116A)) plasmid 513 514 and 2.5 ug of pHCMV-VSV-G plasmid. VSV-G/HIV-1ΔEnv IN_{HA}(D116A) ΔNef ANCH3 515 has been also produced in combination with GIR (Gag-IN-Ruby plasmid, kindly 516 provided by E. Campbell) (Dharan et al., 2016; Hulme et al., 2015). For THP-1 infection, HIV-1 ANCH3 or HIV-1 ANCH3 MS2 viruses and LV OR-GFP or LV MCP-517 GFP were all produced adding 3 µg of SIV_{MAC} Vpx (Durand et al., 2013) at the moment 518 519 of HEK293T transfection. After the collection of the supernatant 48h post-transfection, 520 lentiviral particles were concentrated by ultracentrifugation for 1 h at 22000 rpm at 4°C 521 and stored at -80°C. Lentiviral vectors and viruses were tittered through late reverse 522 transcripts (LTR) qPCR in HEK293T cells or HeLa P4R5 cells.

523

524 **Quantitative PCR and primers**

525 Total DNA of infected cells was extracted at 6 hours post infection for Late Reverse Transcripts (LRTs) qPCR and 24 hours post infection for 2LTR-circles and ALU-PCR. 526 through QIAmp DNA micro kit (QIAGEN #56304). Real-Time PCR of LRTs was 527 528 performed to assess DNA synthesis and used to normalize qPCR data on viral input. 529 The reactions were carried on in 20 µL, in iTaqUniversal SYBR Green Supermix (Bio-530 #1725124) U3 U3 Rad using primers for sequence: FX: 5'-531 TTCCGCTGGGGACTTTCCAGGG-3', U3 RX: 5'-AGGCTCAGATCTGGTCTAACC-3'. 532 Real-Time PCR of 2LTR-circles was used to asses nuclear import efficiency. 533 Reactions were performed in 20 µL, in Maxima Probe/ROX qPCR Mastermix 534 (ThermoFisher #K0232) using primers for 2LTR-circle junction: 2LTR FX: 5'-AACTAGGGAACCCACTGCTTAAG-3', 535 2LTR RX: 5'-536 TCCACAGATCAAGGATATCTTGTC-3'. 2-LTR probe: 5'-(FAM)-537 ACACTACTTGAAGCACTCAAG-GCAAGCTTT-(TAMRA)-3'. **Real-Time** PCR of 538 proviral integrations consisting in a first non-kinetic PCR in 50 µL in Platinum SuperFi 539 DNA Polymerase (ThermoFisher #12351250) and in a second step of qPCR reaction 540 in 20 µL in Maxima Probe/ROX qPCR Mastermix (ThermoFisher #K0232). Primer Alu 5'-TCCCAGCTACTCGGGAGGCTGAGG-3', 2: 5'-541 166: Alu GCCTCCCAAAGTGCTGGGATTACAG-3', 5'-542 LambdaU3: ATGCCACGTAAGCGAAACTTTCCGCTGGGGACTTTCCAGGG-3' for ALU PCR; 543 Lambda: 5'-ATGCCACGTAAGCGAAACT-3', U5: 5'-CTGACTAAAAGGGTCTGAGG-544 545 3', Probe: 5'-(FAM)- TTAAGCCTCAATAAAGCTTGCCTTGAGTGC-(TAMRA) for ALU

qPCR(Di Nunzio et al., 2013). In all experiments β-actin detection was used for
normalization. β-actin FX: 5'-AACACCCCAGCCATGTACGT-3', β-actin RX: 5CGGTGAGGATCTTCATGAGGTAGT-3', β-actin probe: (FAM)CCAGCCAGGTCCAGACGCAGGA-(BHQ1).

550

551 Epifluorescence and Immunofluorescence

552 For imaging studies in adherent cell-lines, the cells were plated on coverslips (12mm 553 diameter in 24 well plates or 18 mm diameter in 12 well plate, #1, ThermoFisher) at 554 least 24h before fixation. For HIV-1 ANCH3 imaging, HeLa P4R5 cells were transduced with OR-GFP LV (MOI 0.2) and then infected within the following 3 days 555 556 with HIV-1 ANCH3 or HIV-1 IND116A ANCH3 at different MOIs. NEVIRAPINE (NEV) 10 557 µM or RALTEGRAVIR (RAL) 20 µM were used to block respectively DNA synthesis 558 and integration. For HIV-1 MS2 RNA imaging HeLa-MCP cells were infected with HIV 559 ANCH3 MS2 and fixed 24h post infection. For co-visualization of vDNA and vRNA, 560 HeLa MCP-GFP cells were transduced with OR-SANTAKA LV (MOI 1) and 24h later infected with HIV-1 ANCH3 MS2. All THP-1 cells studies were carried on after 48h of 561 562 differentiation with PMA (167 nM), then, they were transduced with LV OR-GFP Vpx 563 (MOI 5) or LV MCP-GFP Vpx (MOI 20) and 48 h later they were infected with HIV-1 564 ANCH3 Vpx (30 h or 3 days p.i) or HIV-1 ANCH3 MS2 ± Vpx (3 days p.i.). The medium 565 was always supplemented with PMA. For HIV-1 ANCH3 imaging in primary 566 lymphocytes, CD4⁺T cells, 3 days after activation, were transduced with OR-GFP LV (MOI 0.5) for 72 h and then infected with HIV-1 ANCH3 for 20 h. The day of fixation 567 568 cells in suspension were seeded on poly-L-lysine (Sigma #P4707) coated coverslips 569 and centrifuged to allow attachment. The day of fixation all cells were washed with 570 PBS and fixed with 4% PFA for 15 minutes. For protein staining, cells were treated with glycine 0.15% for 10 min, permeabilized with Triton X-100 0.5% for 30 min and 571 572 blocked with 1% bovine serum albumin (BSA) for 30 min. All incubations were carried 573 out at room temperature, in the dark, in humid chamber, 1h with primary antibodies 574 and 45 min with secondary antibodies. Washes between antibody incubations and 575 antibodies dilution were done in 1% BSA. Primary antibodies were diluted as follows: 576 anti-HA 1:500 (Roche #11867423001), anti-HIV-1 p24 (CA) 1:400 (NIH #3537), anti-577 CPSF6 1:400 (Novus Biologicals #NBP1-85676), anti-SC35 1:200 (Abcam #ab11826), 578 anti-LEDGF 1:200 (BD Bioscience #611715). Secondary antibodies used were the

following: Goat anti-Mouse Alexa-647 1:300 (1:100 for CA) (Invitrogen #A21235); 579 Donkey anti-Rat Alexa-488 1:100 (Invitrogen A21208) for IN-HA) or Goat anti-Rat 580 581 Alexa-647 1:300 (Invitrogen #A21247) (1:100 for IN-HA); Donkey anti-Rabbit Cy3 582 1:500 (Jackson Lab #711-165-152). Finally, cells were stained with Hoechst 33342 583 1:5000 (Invitrogen #H3570) for 5 minutes. Coverslips were mounted on glass slides 584 (Star Frost) with Prolong Diamond Antifade Mountant (Life Technologies #P36970). 585 Confocal microscopy was carried out with a Zeiss LSM700 inverted microscope, with a 63X objective (Plan Apochromat, oil immersion, NA=1.4). 586

587

588 RNA FISH

589 For RNA FISH studies in adherent cell-lines, the cells were seeded on coverslips. HeLa 590 P4R5 cells expressing OR-GFP were infected with HIV-1 ANCH3 (MOI 30) with or without RAL 20 µM and fixed 24h post infection or infected with HIV-1 ANCH3 (MOI 5) 591 592 for 48h or 7 days. Jurkat T cells expressing OR-GFP were infected with HIV-1 ANCH3 593 (MOI 5) and fixed 48h post infection. Primary activated CD4⁺T cells, 3 days after 594 activation, were together transduced with OR-GFP LV (MOI 10) and infected with HIV-595 1 ANCH3 (MOI 5) for 3 days, THP-1 cells were differentiated with PMA for 48h and, 596 after 48h transduction with LV OR-GFP Vpx (MOI 5), the cells were infected with HIV-597 1 ANCH3 Vpx (MOI 20) for 3 days. The day of fixation Jurkat and CD4⁺T cells were 598 seeded on poly-L-lysine (Sigma #P4707) coated coverslips and centrifuged to allow 599 attachment. The day of fixation all cells were washed with PBS and fixed with 4% PFA 600 for 15 minutes and incubated in 70% ethanol at -20°C at least for one night. Primary 601 smiFISH probes have a targeting sequence against HIV-1 pol transcript and a shared 602 readout sequence for secondary probe alignment. Twenty-four smiFISH probes (table 603 I) against HIV pol were designed with Oligostan (Tsanov et al., 2016) and purchased 604 from Integrated DNA Technologies (IDT). Primary probes were pre-hybridized with a 605 secondary FLAP probe conjugated to Cy5 fluorophore through pairing with the readout 606 sequence. Washes and hybridization were performed with Stellaris Buffers (WASH buffer A, WASH buffer B, Hybridization Buffer; LGC Biosearch Technologies), 607 608 following the manufacturer protocol. Hybridization with the probe was carried out at 609 37°C in a dark humid chamber for 5 hours. Finally, cells were stained with Hoechst 610 33342 1:5000 (Invitrogen #H3570) for 5 minutes. Coverslips were mounted on glass 611 slides (Star Frost) with Prolong Diamond Antifade Mountant (Life Technologies

#P36970). Confocal microscopy was carried out with a Zeiss inverted LSM700
microscope, with a 63X objective (Plan Apochromat, oil immersion, NA=1.4).

614 **Time-lapse microscopy**

615 For all live imaging studies, the cells were plated on a polymer-coverslip bottom µ-Dish 35 mm (ibidi #81156) or in µ-Dish 35 mm Quad (10,000 cells per chamber, ibidi 616 617 #80416) and 2D or 3D videos were acquired with an UltraView VOX Spinning Disk 618 Microscope (Perkin-Elmer), based on a CSU-X spinning-disk (Yokogawa), and using 619 a 63X objective (Plan Apochromat, oil immersion, NA=1.4). For integrated provirus imaging, HIV-1 ANCH3 single-cell clone was transduced with OR-GFP LV (MOI 0.2) 620 for 3 days and then plated on the polymer-bottom dish (10⁵ cells) the day before 621 622 acquisition. On the other hand, HeLa P4R5 cells were infected with HIV-1 ANCH3 (MOI 623 10) and five days post infection cells were plated on the polymer-coverslip bottom dish 624 and transduced with LV OR-GFP for 24h. For non-integrated viral DNA, HeLa P4R5 were firstly transduced with OR-GFP LV (MOI 0.2) and the day after cells were plated 625 626 on polymer-bottom dish and infected with HIV-1 ANCH3 (MOI 10) with RAL 20 µM or 627 with the mutated HIV-1 IN_{HA} (D116A) ANCH3. Cells were imaged in continue for 10 minutes. Experiments of co-live tracking of IN-Ruby (GIR) and vDNA labeled by OR-628 GFP were performed in HeLa P4R5 cells transduced with LV OR-GFP (MOI 0.2). The 629 630 following day, cells were infected with HIV-1 IN_{HA}(D116A) ANCH3 complemented with 631 the GIR plasmid using an MOI 50. Fluorescence images were taken every 30 seconds 632 for up to 9h post infection. For co-live-imaging of vDNA and vRNA foci, HeLa MCP-633 GFP cells were transduced with OR-SANTAKA LV (MOI 1) and 24 h later infected with 634 HIV-1 ANCH3 MS2 (MOI 50).

635

636 Imaging Analysis and Statistics

All data were analyzed in GraphPad Prism 8 (GraphPad Software, La Jolla California USA, <u>www.graphpad.com</u>), computing Student's t test for 2 comparisons, One-way ANOVA for more than two comparisons and Pearson's correlation. P values symbols: ns = non significant, $* = p \le 0.05$, $** = p \le 0.01$, $*** = p \le 0.001$, $**** = p \le 0.0001$. All images and videos were analyzed in Fiji (Schindelin et al., 2012) and Icy version 2.0.3.0 (de Chaumont et al., 2012) software. All images and video were processed and analyzed in Fiji (Schindelin et al., 2012) or
in Icy software version 2.0.3.0 (de Chaumont et al., 2012). Here below the custommade protocols:

646 Particle tracking and MSD measurement

647 3D videos were analyzed in Fiji (Schindelin et al., 2012) and viral nuclear particles were tracked along time in 3D using TrackMate plugin (Tinevez et al., 2017). The 648 649 resulting tracks were exported in MATLAB (The MathWorks, Natick, USA) and then used for mean-squared displacement (MSD) analysis with @msdanalyzer(Tarantino 650 651 et al., 2014). MSD curves were fitted by a straight line from which we derived the value 652 of the diffusion coefficient: $D = a / (2^*nDim)$, where « a » is the slope of the MSD curve and « nDim » the dimensionality of the image. MSD curves for which the R² value was 653 654 lower than 0.8 were discarded from analysis. Only the initial 25% of the MSD curves 655 were used for the linear fit.

656 vDNA-vRNA association analysis

657 For the analysis of RNA FISH images ± RAL in HeLa P4R5 cells, 2D confocal images 658 were analysed with Icy software version 2.0.3.0 (de Chaumont et al., 2012) using a 659 custom made (http://icy.bioimageanalysis.org/protocol/batch-spotsprotocol 660 detections-and-statistics and doi: 10.5281/zenodo.3925262, example of HeLa 661 confocal image available at doi: 10.5281/zenodo.3925213). vDNA spots were detected extracting the green channel and applying first a median filter of half size = 1 and then 662 3, the Spot Detector with a scale of 1 (sensitivity to 55) on the nuclear ROI only. Data 663 from the spot detection (e.g. mean fluorescence intensity) were exported for analysis. 664 vRNA mean fluorescence intensity signal was computed from the nuclei ROI. For 665 vDNA - vRNA association studies in Jurkat cells, 3D confocal images were analyzed 666 667 in Icy version 2.0.3.0 using two successive custom made protocols (first protocol: 668 http://icy.bioimageanalysis.org/protocol/nuclei-segmentation-from-maximum-

669 <u>intensity-projections/</u> with doi: 10.5281/zenodo.3925181) and second protocol:
 670 <u>http://icy.bioimageanalysis.org/protocol/spots-detection-and-colocalisation-analysis-</u>

677 doi.org/10.1109/ICPR.2008.4761748). The logical operation block was used to keep 678 only the spots detected in the nucleus. These spots were added as input parameter 679 "List detection 1" of the SODA block. EGFP spots were detected using the log3D block 680 to enhance the spots, inverting the log transformed image using the Math Operation 681 Expression block with the following expression: A*B with A scalar=-1 and B the log 682 transformed image, detecting the spots with the Local Extrema block and segmenting 683 the spots with a dilation and a connected components operations. The logical operation 684 block was used to keep only the spots detected in the nucleus. These spots were 685 added as input parameter "List detection 2" of the SODA block. The nucleus ROI was 686 kept as input parameter for "ROI of Analysis (Cell's shape)". The Cy5 and GFP spots 687 ROIs were then imported in the SODA block (Lagache et al., 2018) to analyse 688 colocalization of spots within the nucleus only with a maximum radius of 10, a step of 689 1 and no fixed search distance.

690 vDNA/vRNA nuclear envelope distance analysis

691 For the analysis of HeLa P4R5 cells 48 h p.i. and 7 days p.i. and in THP-1 3 days p.i. 3D RNA FISH images, the nuclei were segmented on the Hoechst channel using 692 693 background subtraction followed by a multilevel Otsu thresholding. The resulting nuclei 694 masks were then directly applied on the vDNA and vRNA channels to select the 695 relevant 3D information inside each labelled region. Inside these regions, vRNA spots 696 were identified using 3D non-local means denoising (Buades et al., "Non-local means 697 denoising." Image Processing On Line 1 (2011): 208-212) and automated 698 segmentation, while vDNA spots were extracted with the image analysis pipeline 699 previously described (Komatsu et al., 2018). For each vDNA and vRNA spot in was 700 computed the distance to the nuclear envelope in the 2D z-plane corresponding to the 701 spot centroid location.

702 SC35-vDNA distance analysis

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

711

712

713 Acknowledgements

714 We wish to thank Guillermo Blanco-Rodriguez, Stella Frabetti, Philippe Sougue and 715 Blandine Monel for experimental help. We thank Nicoletta Casartelli for critical reading 716 of the manuscript. We thank Fabrizio Mammano, Edouard Bertrand, Florian Mueller, 717 and Edward Campbell for sharing reagents. We gratefully acknowledge the UtechS 718 Photonic Biolmaging (Imagopole), C2RT, Institut Pasteur, supported by the French National Research Agency (France BioImaging; ANR-10-INSB-04: Investments for 719 720 the Future). We thank the NIH AIDS Reagents program to support us with precious 721 reagents. This work was funded by the ANRS (Agence Nationale de Recherche sur le 722 SIDA) grant ECTZ88162 with a nominative PhD student fellowship ECTZ88177 for 723 V.S., the Sidaction/FRM grant VIH20170718001, the Pasteur Institute.

724

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

726

727 References

- Achuthan, V., J.M. Perreira, G.A. Sowd, M. Puray-Chavez, W.M. McDougall, A. Paulucci-Holthauzen, X.
 Wu, H.J. Fadel, E.M. Poeschla, A.S. Multani, S.H. Hughes, S.G. Sarafianos, A.L. Brass, and A.N.
 Engelman. 2018. Capsid-CPSF6 Interaction Licenses Nuclear HIV-1 Trafficking to Sites of Viral
 DNA Integration. *Cell Host Microbe*. 24:392-404 e398.
- Albanese, A., D. Arosio, M. Terreni, and A. Cereseto. 2008. HIV-1 pre-integration complexes selectively
 target decondensed chromatin in the nuclear periphery. *PLoS One*. 3:e2413.
- Ballandras-Colas, A., D.P. Maskell, E. Serrao, J. Locke, P. Swuec, S.R. Jonsson, A. Kotecha, N.J. Cook, V.E.
 Pye, I.A. Taylor, V. Andresdottir, A.N. Engelman, A. Costa, and P. Cherepanov. 2017. A
 supramolecular assembly mediates lentiviral DNA integration. *Science*. 355:93-95.
- Bejarano, D.A., K. Peng, V. Laketa, K. Borner, K.L. Jost, B. Lucic, B. Glass, M. Lusic, B. Muller, and H.G.
 Krausslich. 2019. HIV-1 nuclear import in macrophages is regulated by CPSF6-capsid
 interactions at the nuclear pore complex. *Elife*. 8.
- Blanco-Rodriguez, G., A. Gazi, B. Monel, S. Frabetti, V. Scoca, F. Mueller, O. Schwartz, J. Krijnse-Locker,
 P. Charneau, and F. Di Nunzio. 2020. Remodeling of the core leads HIV-1 pre-integration
 complex in the nucleus of human lymphocytes. *J Virol*.
- Buffone, C., A. Martinez-Lopez, T. Fricke, S. Opp, M. Severgnini, I. Cifola, L. Petiti, S. Frabetti, K.
 Skorupka, K.K. Zadrozny, B.K. Ganser-Pornillos, O. Pornillos, F. Di Nunzio, and F. Diaz-Griffero.
 2018. Nup153 Unlocks the Nuclear Pore Complex for HIV-1 Nuclear Translocation in
 Nondividing Cells. *J Virol.* 92.

- Burdick, R.C., K.A. Delviks-Frankenberry, J. Chen, S.K. Janaka, J. Sastri, W.S. Hu, and V.K. Pathak. 2017.
 Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. *PLoS Pathog.* 13:e1006570.
- Burdick, R.C., C. Li, M. Munshi, J.M.O. Rawson, K. Nagashima, W.S. Hu, and V.K. Pathak. 2020. HIV-1
 uncoats in the nucleus near sites of integration. *Proc Natl Acad Sci U S A*. 117:5486-5493.
- 754Bushman, F.D. 2003. Targeting survival: integration site selection by retroviruses and LTR-755retrotransposons. Cell. 115:135-138.
- Butler, S.L., M.S. Hansen, and F.D. Bushman. 2001. A quantitative assay for HIV DNA integration in vivo.
 Nat Med. 7:631-634.
- Campbell, E.M., and T.J. Hope. 2015. HIV-1 capsid: the multifaceted key player in HIV-1 infection. *Nat Rev Microbiol.* 13:471-483.
- 760Charneau, P., G. Mirambeau, P. Roux, S. Paulous, H. Buc, and F. Clavel. 1994. HIV-1 reverse761transcription. A termination step at the center of the genome. J Mol Biol. 241:651-662.
- Cherepanov, P., G. Maertens, P. Proost, B. Devreese, J. Van Beeumen, Y. Engelborghs, E. De Clercq,
 and Z. Debyser. 2003. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75
 protein in human cells. *J Biol Chem.* 278:372-381.
- Chin, C.R., J.M. Perreira, G. Savidis, J.M. Portmann, A.M. Aker, E.M. Feeley, M.C. Smith, and A.L. Brass.
 2015. Direct Visualization of HIV-1 Replication Intermediates Shows that Capsid and CPSF6
 Modulate HIV-1 Intra-nuclear Invasion and Integration. *Cell Rep.* 13:1717-1731.
- Ciuffi, A., M. Llano, E. Poeschla, C. Hoffmann, J. Leipzig, P. Shinn, J.R. Ecker, and F. Bushman. 2005. A
 role for LEDGF/p75 in targeting HIV DNA integration. *Nat Med*. 11:1287-1289.
- Craig, E.A., and J. Marszalek. 2002. A specialized mitochondrial molecular chaperone system: a role in
 formation of Fe/S centers. *Cell Mol Life Sci*. 59:1658-1665.
- de Chaumont, F., S. Dallongeville, N. Chenouard, N. Herve, S. Pop, T. Provoost, V. Meas-Yedid, P.
 Pankajakshan, T. Lecomte, Y. Le Montagner, T. Lagache, A. Dufour, and J.C. Olivo-Marin. 2012.
 Icy: an open bioimage informatics platform for extended reproducible research. *Nat Methods*.
 9:690-696.
- De Wit, F., S.R. Pillalamarri, A. Sebastian-Martin, A. Venkatesham, A. Van Aerschot, and Z. Debyser.
 2019. Design of reverse transcriptase-specific nucleosides to visualize early steps of HIV-1
 replication by click labeling. *J Biol Chem.* 294:11863-11875.
- Dharan, A., S. Talley, A. Tripathi, J.I. Mamede, M. Majetschak, T.J. Hope, and E.M. Campbell. 2016.
 KIF5B and Nup358 Cooperatively Mediate the Nuclear Import of HIV-1 during Infection. *PLoS Pathog.* 12:e1005700.
- Di Nunzio, F. 2013. New insights in the role of nucleoporins: a bridge leading to concerted steps from
 HIV-1 nuclear entry until integration. *Virus Res.* 178:187-196.
- Di Nunzio, F., T. Fricke, A. Miccio, J.C. Valle-Casuso, P. Perez, P. Souque, E. Rizzi, M. Severgnini, F.
 Mavilio, P. Charneau, and F. Diaz-Griffero. 2013. Nup153 and Nup98 bind the HIV-1 core and
 contribute to the early steps of HIV-1 replication. *Virology*. 440:8-18.
- Di Primio, C., V. Quercioli, A. Allouch, R. Gijsbers, F. Christ, Z. Debyser, D. Arosio, and A. Cereseto. 2013.
 Single-cell imaging of HIV-1 provirus (SCIP). *Proc Natl Acad Sci U S A*. 110:5636-5641.
- Durand, S., X.N. Nguyen, J. Turpin, S. Cordeil, N. Nazaret, S. Croze, R. Mahieux, J. Lachuer, C. Legras Lachuer, and A. Cimarelli. 2013. Tailored HIV-1 vectors for genetic modification of primary
 human dendritic cells and monocytes. *J Virol*. 87:234-242.
- Finiliani, S., A. Mousnier, K. Busschots, M. Maroun, B. Van Maele, D. Tempe, L. Vandekerckhove, F.
 Moisant, L. Ben-Slama, M. Witvrouw, F. Christ, J.C. Rain, C. Dargemont, Z. Debyser, and R.
 Benarous. 2005. Integrase mutants defective for interaction with LEDGF/p75 are impaired in
 chromosome tethering and HIV-1 replication. *J Biol Chem.* 280:25517-25523.
- Farnet, C.M., and W.A. Haseltine. 1991. Determination of viral proteins present in the human
 immunodeficiency virus type 1 preintegration complex. *J Virol*. 65:1910-1915.
- Ferris, A.L., X. Wu, C.M. Hughes, C. Stewart, S.J. Smith, T.A. Milne, G.G. Wang, M.C. Shun, C.D. Allis, A.
 Engelman, and S.H. Hughes. 2010. Lens epithelium-derived growth factor fusion proteins
 redirect HIV-1 DNA integration. *Proc Natl Acad Sci U S A*. 107:3135-3140.

- Fox, A.H., Y.W. Lam, A.K. Leung, C.E. Lyon, J. Andersen, M. Mann, and A.I. Lamond. 2002. Paraspeckles:
 a novel nuclear domain. *Curr Biol*. 12:13-25.
- Francis, A.C., C. Di Primio, V. Quercioli, P. Valentini, A. Boll, G. Girelli, F. Demichelis, D. Arosio, and A.
 Cereseto. 2014. Second generation imaging of nuclear/cytoplasmic HIV-1 complexes. *AIDS Res Hum Retroviruses*. 30:717-726.
- Francis, A.C., M. Marin, J. Shi, C. Aiken, and G.B. Melikyan. 2016. Time-Resolved Imaging of Single HIV 1 Uncoating In Vitro and in Living Cells. *PLoS Pathog*. 12:e1005709.
- Francis, A.C., M. Marin, P.K. Singh, V. Achuthan, M.J. Prellberg, K. Palermino-Rowland, S. Lan, P.R.
 Tedbury, S.G. Sarafianos, A.N. Engelman, and G.B. Melikyan. 2020. HIV-1 replication
 complexes accumulate in nuclear speckles and integrate into speckle-associated genomic
 domains. *Nat Commun.* 11:3505.
- Francis, A.C., and G.B. Melikyan. 2018. Single HIV-1 Imaging Reveals Progression of Infection through
 CA-Dependent Steps of Docking at the Nuclear Pore, Uncoating, and Nuclear Transport. *Cell host & microbe*. 23:536-548 e536.
- Geis, F.K., and S.P. Goff. 2019. Unintegrated HIV-1 DNAs are loaded with core and linker histones and
 transcriptionally silenced. *Proc Natl Acad Sci U S A*. 116:23735-23742.
- 6ermier, T., S. Kocanova, N. Walther, A. Bancaud, H.A. Shaban, H. Sellou, A.Z. Politi, J. Ellenberg, F.
 6allardo, and K. Bystricky. 2017. Real-Time Imaging of a Single Gene Reveals TranscriptionInitiated Local Confinement. *Biophys J.* 113:1383-1394.
- Graham, T.G., X. Wang, D. Song, C.M. Etson, A.M. van Oijen, D.Z. Rudner, and J.J. Loparo. 2014. ParB
 spreading requires DNA bridging. *Genes Dev.* 28:1228-1238.
- Hare, S., F. Di Nunzio, A. Labeja, J. Wang, A. Engelman, and P. Cherepanov. 2009. Structural basis for
 functional tetramerization of lentiviral integrase. *PLoS Pathog.* 5:e1000515.
- Hrecka, K., C. Hao, M. Gierszewska, S.K. Swanson, M. Kesik-Brodacka, S. Srivastava, L. Florens, M.P.
 Washburn, and J. Skowronski. 2011. Vpx relieves inhibition of HIV-1 infection of macrophages
 mediated by the SAMHD1 protein. *Nature*. 474:658-661.
- Hubner, W., G.P. McNerney, P. Chen, B.M. Dale, R.E. Gordon, F.Y. Chuang, X.D. Li, D.M. Asmuth, T.
 Huser, and B.K. Chen. 2009. Quantitative 3D video microscopy of HIV transfer across T cell
 virological synapses. *Science*. 323:1743-1747.
- Hulme, A.E., Z. Kelley, D. Foley, and T.J. Hope. 2015. Complementary Assays Reveal a Low Level of CA
 Associated with Viral Complexes in the Nuclei of HIV-1-Infected Cells. *J Virol.* 89:5350-5361.
- Hulme, A.E., O. Perez, and T.J. Hope. 2011. Complementary assays reveal a relationship between HIV 1 uncoating and reverse transcription. *Proceedings of the National Academy of Sciences of the United States of America*. 108:9975-9980.
- Komatsu, T., C. Quentin-Froignant, I. Carlon-Andres, F. Lagadec, F. Rayne, J. Ragues, R.H. Kehlenbach,
 W. Zhang, A. Ehrhardt, K. Bystricky, R. Morin, J.M. Lagarde, F. Gallardo, and H. Wodrich. 2018.
 In Vivo Labelling of Adenovirus DNA Identifies Chromatin Anchoring and Biphasic Genome
 Replication. J Virol. 92.
- Lagache, T., A. Grassart, S. Dallongeville, O. Faklaris, N. Sauvonnet, A. Dufour, L. Danglot, and J.C. Olivo Marin. 2018. Mapping molecular assemblies with fluorescence microscopy and object-based
 spatial statistics. *Nat Commun*. 9:698.
- Laguette, N., B. Sobhian, N. Casartelli, M. Ringeard, C. Chable-Bessia, E. Segeral, A. Yatim, S. Emiliani,
 O. Schwartz, and M. Benkirane. 2011. SAMHD1 is the dendritic- and myeloid-cell-specific HIV 1 restriction factor counteracted by Vpx. *Nature*. 474:654-657.
- Lee, K., Z. Ambrose, T.D. Martin, I. Oztop, A. Mulky, J.G. Julias, N. Vandegraaff, J.G. Baumann, R. Wang,
 W. Yuen, T. Takemura, K. Shelton, I. Taniuchi, Y. Li, J. Sodroski, D.R. Littman, J.M. Coffin, S.H.
 Hughes, D. Unutmaz, A. Engelman, and V.N. KewalRamani. 2010. Flexible use of nuclear import
 pathways by HIV-1. *Cell Host Microbe*. 7:221-233.
- Lelek, M., N. Casartelli, D. Pellin, E. Rizzi, P. Souque, M. Severgnini, C. Di Serio, T. Fricke, F. Diaz-Griffero,
 C. Zimmer, P. Charneau, and F. Di Nunzio. 2015. Chromatin organization at the nuclear pore
 favours HIV replication. *Nat Commun.* 6:6483.

- Lelek, M., F. Di Nunzio, R. Henriques, P. Charneau, N. Arhel, and C. Zimmer. 2012. Superresolution imaging of HIV in infected cells with FIAsH-PALM. *Proc Natl Acad Sci U S A*. 109:8564-8569.
- Liu, R., F.R. Simonetti, and Y.C. Ho. 2020. The forces driving clonal expansion of the HIV-1 latent reservoir. *Virol J.* 17:4.
- Llano, M., D.T. Saenz, A. Meehan, P. Wongthida, M. Peretz, W.H. Walker, W. Teo, and E.M. Poeschla.
 2006. An essential role for LEDGF/p75 in HIV integration. *Science*. 314:461-464.
- Maldarelli, F. 2016. The role of HIV integration in viral persistence: no more whistling past the proviral
 graveyard. *J Clin Invest*. 126:438-447.
- 860 Mamede, J.I., G.C. Cianci, M.R. Anderson, and T.J. Hope. 2017. Early cytoplasmic uncoating is 861 associated with infectivity of HIV-1. *Proc Natl Acad Sci U S A*. 114:E7169-E7178.
- Mariame, B., S. Kappler-Gratias, M. Kappler, S. Balor, F. Gallardo, and K. Bystricky. 2018. Real-Time
 Visualization and Quantification of Human Cytomegalovirus Replication in Living Cells Using
 the ANCHOR DNA Labeling Technology. J Virol. 92.
- Marini, B., A. Kertesz-Farkas, H. Ali, B. Lucic, K. Lisek, L. Manganaro, S. Pongor, R. Luzzati, A. Recchia, F.
 Mavilio, M. Giacca, and M. Lusic. 2015. Nuclear architecture dictates HIV-1 integration site
 selection. *Nature*. 521:227-231.
- Naganuma, T., and T. Hirose. 2013. Paraspeckle formation during the biogenesis of long non-coding
 RNAs. *RNA Biol.* 10:456-461.
- Olson, A., B. Basukala, W.W. Wong, and A.J. Henderson. 2019. Targeting HIV-1 proviral transcription.
 Curr Opin Virol. 38:89-96.
- Peng, K., W. Muranyi, B. Glass, V. Laketa, S.R. Yant, L. Tsai, T. Cihlar, B. Muller, and H.G. Krausslich.
 2014. Quantitative microscopy of functional HIV post-entry complexes reveals association of
 replication with the viral capsid. *Elife*. 3:e04114.
- Petit, C., O. Schwartz, and F. Mammano. 1999. Oligomerization within virions and subcellular
 localization of human immunodeficiency virus type 1 integrase. *J Virol*. 73:5079-5088.
- Petit, C., O. Schwartz, and F. Mammano. 2000. The karyophilic properties of human immunodeficiency
 virus type 1 integrase are not required for nuclear import of proviral DNA. *J Virol*. 74:7119 7126.
- Pradeepa, M.M., G.R. Grimes, G.C. Taylor, H.G. Sutherland, and W.A. Bickmore. 2014. Psip1/Ledgf p75
 restrains Hox gene expression by recruiting both trithorax and polycomb group proteins.
 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.
- 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.
- 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.
- Sharma, A., H. Takata, K. Shibahara, A. Bubulya, and P.A. Bubulya. 2010. Son is essential for nuclear
 speckle organization and cell cycle progression. *Mol Biol Cell*. 21:650-663.
- 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*. 82:1155511567.
- 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.
- 918Singh, 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.919Hughes, M. Kvaratskhelia, and H.L. Levin. 2015. LEDGF/p75 interacts with mRNA splicing920factors and targets HIV-1 integration to highly spliced genes. Genes Dev. 29:2287-2297.
- 921 Spector, D.L., and A.I. Lamond. 2011. Nuclear speckles. *Cold Spring Harb Perspect Biol.* 3.
- Tantale, K., F. Mueller, A. Kozulic-Pirher, A. Lesne, J.M. Victor, M.C. Robert, S. Capozi, R. Chouaib, V.
 Backer, J. Mateos-Langerak, X. Darzacq, C. Zimmer, E. Basyuk, and E. Bertrand. 2016. A single molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting.
 Nat Commun. 7:12248.
- Tarantino, N., J.Y. Tinevez, E.F. Crowell, B. Boisson, R. Henriques, M. Mhlanga, F. Agou, A. Israel, and E.
 Laplantine. 2014. TNF and IL-1 exhibit distinct ubiquitin requirements for inducing NEMO-IKK
 supramolecular structures. *J Cell Biol*. 204:231-245.
- Tinevez, J.Y., N. Perry, J. Schindelin, G.M. Hoopes, G.D. Reynolds, E. Laplantine, S.Y. Bednarek, S.L.
 Shorte, and K.W. Eliceiri. 2017. TrackMate: An open and extensible platform for single-particle
 tracking. *Methods*. 115:80-90.
- Tsanov, N., A. Samacoits, R. Chouaib, A.M. Traboulsi, T. Gostan, C. Weber, C. Zimmer, K. Zibara, T.
 Walter, M. Peter, E. Bertrand, and F. Mueller. 2016. smiFISH and FISH-quant a flexible single
 RNA detection approach with super-resolution capability. *Nucleic Acids Res.* 44:e165.
- Vansant, G., H.C. Chen, E. Zorita, K. Trejbalova, D. Miklik, G. Filion, and Z. Debyser. 2020. The chromatin
 landscape at the HIV-1 provirus integration site determines viral expression. *Nucleic Acids Res.*
- Wang, G.P., B.L. Levine, G.K. Binder, C.C. Berry, N. Malani, G. McGarrity, P. Tebas, C.H. June, and F.D.
 Bushman. 2009. Analysis of lentiviral vector integration in HIV+ study subjects receiving
 autologous infusions of gene modified CD4+ T cells. *Mol Ther*. 17:844-850.
- Wang, H., M.C. Shun, X. Li, F. Di Nunzio, S. Hare, P. Cherepanov, and A. Engelman. 2014. Efficient
 Transduction of LEDGF/p75 Mutant Cells by Gain-of-Function HIV-1 Integrase Mutant Viruses.
 Mol Ther Methods Clin Dev. 1.
- Shu, Y., G.Z. Wang, O. Cingoz, and S.P. Goff. 2018. NP220 mediates silencing of unintegrated retroviral
 DNA. *Nature*. 564:278-282.
- 945
- 946
- 947
- 948

949

FIGURE LEGENDS

950

951 Figure 1. Detection and live tracking of nuclear HIV-1 DNA forms. A) Scheme of 952 vDNA labeled by the HIV-1 ANCHOR system. ANCH3 is the target sequence of the 953 OR protein. ANCH3 has been cloned in the HIV-1 genome in place of Nef gene (HIV-954 1 ANCH3). The late retrotranscribed DNA of HIV-1 ANCH3 can be visualized in cells 955 stably expressing the OR-GFP protein, thanks to the accumulation of OR protein to the 956 double-stranded ANCH3 sequence. Viral DNA detection occurs mostly in the nucleus, 957 where late reverse transcripts are abundant and are more likely to be exposed to the 958 binding of OR-GFP. Starting from early time points of infection, it is possible to follow 959 HIV-1 ANCH3 in target cells. The labelling results in the appearance of a bright 960 fluorescent spot that can be tracked over time (time post-infection is indicated in the 961 upper-left corner). The time-lapse images are representative of three independent 962 experiments. The cartoon model was created with BioRender.com. B) Confocal 963 images of HeLa P4R5 cells expressing OR-GFP infected at MOI 30 with HIV-1 ANCH3, 964 HIV-1 ANCH3 + RAL (20 µM), HIV-1 ANCH3 IN_{D116A} and HIV-1 ANCH3 + NEV 965 (10 µM), 24 h p.i. On the right, histogram plots of 2-LTR circles and ALU-PCR, 966 comparing the different infection conditions, represent the nuclear import and 967 integration rates, respectively. All graphs contain ± SD. C) Confocal image of HIV-1 968 ANCH3 provirus in a single-cell clone. The clone was transduced with OR-GFP LV for 969 the detection of HIV-1 integrated DNA. On the right, histogram plot of ALU-PCR and 970 β-galactosidase expression of the HIV-1 ANCH3 single-cell clone compared to non-971 infected cells (NI). All graphs contain \pm SD. C) On the left, Tukey's box plots of the 972 diffusion coefficient (D (nm²/s)) of episomal and integrated HIV-1 DNA forms tracked 973 in 3D time lapses (movies S1, S2, S3). On the right, an example of live track for each 974 experimental group is shown. Low D corresponds to D $(nm^2/s) < 5$ and high D to 5 < D 975 $(nm^2/s) < 20$. The color range indicates the time up to 10 min. Statistics: Student's t 976 test for two comparisons, one-way ANOVA for more than two comparisons. ns=non-977 significant, *=p ≤ 0.05 , **=p ≤ 0.01 , ***=p ≤ 0.001 , ****=p ≤ 0.0001 . Scale bars: 5 µm. 978

Figure 2. HIV-1 transcriptional activity in dividing and non-dividing cells. A)
Labeling of vDNA and vRNA in HeLa cells expressing OR-GFP obtained by coupling
of HIV-1 ANCHOR (vDNA) with RNA FISH (vRNA). On the left, confocal images of

HIV-1 RNA FISH with or without RAL (20 µM), 24 h p.i. On the right, scatter plot with 982 983 bars showing the mean fluorescence intensity of the vRNA signal (red) per nucleus 984 and of vDNA spots detection (green) (for vRNA: n = 13 cells (-RAL), 16 cells (+RAL); 985 for vDNA: 95 (-RAL) and 141 (+RAL) spots, n = 13 cells (-RAL), 16 cells (+RAL)). The 986 analysis was performed in 2D. B) Labeling of vDNA and vRNA in CD4⁺T cells 987 expressing OR-GFP. On the left, confocal image of HIV-1 RNA FISH in primary 988 activated CD4⁺ T cells, 3 days p.i.. On the right, scatter plot with bars showing the 989 percentage of vDNA associated to vRNA signal and vice versa, per nucleus (n= 262 990 (vDNA spots); 64 (vRNA foci) in infected Jurkat cells, 48 h p.i. (some examples of cells 991 are shown in Fig. S5). The analysis was performed in 3D. C) The same techniques 992 have been used to label vDNA and vRNA in macrophage-like cells, THP-1, expressing 993 OR-GFP. On the left, confocal image of infected THP-1, 3 days p.i. On the right, scatter 994 plot with bars showing the percentage of vDNA associated to vRNA signal and vice 995 versa, per nucleus (n=130 (vDNA spots), 99 (vRNA foci). The analysis was performed in 3D. All graphs contain ± SD. Statistics: Student's t test. ns=non-significant, *=p ≤ 996 997 0.05, **=p ≤ 0.01 , ***=p ≤ 0.001 , ****=p ≤ 0.0001 . Scale bars: 5 µm.

998

999 Figure 3. Distance of HIV-1 provirus from the nuclear envelope in dividing and 1000 **non-dividing cells.** A) On the left, confocal images of HIV-1 DNA labeling by HIV-1 1001 ANCHOR system and HIV-1 RNA labeling by RNA FISH in HeLa cells expressing OR-GFP infected with HIV-1 ANCH3 at 48 h p.i. compared to 7 days p.i.. On the right, 1002 1003 scatter plot with bars indicating the distance from the nuclear envelope (NE) per 1004 detected vDNA spot (green) and vRNA focus (red) (n=vDNA spots: 83 (48 h p.i.) and 1005 16 (7 d p.i.); n=vRNA foci: 52 (48h p.i.) and 18 (7d p.i.)). B) On the left, confocal images 1006 to detect vDNA and vRNA in THP-1 cells expressing OR-GFP infected with HIV-1 1007 ANCH3, 3 days p.i.. On the right, scatter plot with bars indicating the distance from the 1008 nuclear envelope (NE) per detected vDNA spot (green) and vRNA focus (red) (n=139 vDNA spots; n= 126 vRNA foci). C) Scatter plot with bars comparing the distances 1009 1010 from the nuclear envelope (NE) per detected vRNA focus in HeLa P4R5, 48h p.i. and THP-1, 3 days p.i. (vRNA foci: 52 (HeLa), 126 (THP-1)). All the analyses were 1011 1012 performed in 3D. All graphs contain ± SD. Statistics: Student's t test for two 1013 comparisons, one-way ANOVA for more than two comparisons. ns=non-significant, $*=p \le 0.05$, $**=p \le 0.01$, $***=p \le 0.001$, $****=p \le 0.0001$. Scale bars: 5 µm, inset 1 µm. 1014 1015

1016 Figure 4. HIV-1 integrase location in dividing and non-dividing cells. A) Confocal images of primary activated CD4⁺ T and THP-1 cells expressing OR-GFP infected with 1017 1018 HIV-1 ANCH3, 20h and 30h p.i, respectively, compared to uninfected cells. Covisualization of vDNA (green) and viral IN (magenta). On the bottom scatter plot with 1019 1020 bars comparing the percentages of vDNAs associated to IN per nucleus (n=vDNA: 54 1021 (CD4+T), 100 (THP-1)). The analysis was performed in 3D. B) Confocal images of 1022 primary activated CD4⁺ T and THP-1 cells infected with HIV-1 ANCH3, 20h and 30h p.i, respectively, compared to uninfected cells. Co-visualization of CPSF6 (red) and 1023 1024 viral IN (cyan). The graphs show the percentages of IN-CPSF6 association events and vice versa, per nucleus (n= for CD4+T: 46 (CPSF6), 33 (IN); for THP-1: 206 1025 1026 (CPSF6), 132 (IN)). All graphs contain ± SD. Statistics: Student's t test. ns=nonsignificant, $*=p \le 0.05$, $**=p \le 0.01$, $***=p \le 0.001$, $****=p \le 0.0001$. Scale bars: 5 µm. 1027 1028

1029 Figure 5. HIV-1 proviruses are excluded from CPSF6/SC35 MLOs. A) Confocal 1030 images of THP-1 cells expressing OR-GFP infected with HIV-1 ANCH3, 3 days p.i., to co-visualize vDNA (green), CPSF6 (red), SC35 (cyan) or to co-visualize vDNA (green), 1031 1032 SON (red), SC35 (cyan), compared to uninfected cells. On the right, the graphs show: the percentages of CPSF6 clusters associated to vDNA or to SC35, per nucleus (n= 1033 1034 CPSF6 clusters > 50); the comparison of Pearson's r coefficients of SON-SC35 colocalization between infected and uninfected cells, per nucleus (n = 25 cells (HIV-1 1035 ANCH3), 19 cells (NI)); the average distance of each vDNA from the closest SC35-1036 speckle border (n= vDNAs: 83). B) Confocal images of THP-1 cells expressing MCP-1037 GFP infected with HIV-1 ANCH3 MS2, 3 days p.i., to co-visualize vRNA (green), 1038 1039 CPSF6 (red), SC35 (cyan), compared to uninfected cells. The graph shows the 1040 percentages of CPSF6 clusters associated to vRNA foci or to SC35 (n= CPSF6 clusters: 42). C) Scatter plot with bars indicating the percentage of vDNA or vRNA 1041 1042 associated to CPSF6, per nucleus (n= 69 (vDNA), 30 (vRNA)). All the analyses were 1043 performed in 3D. All graphs contain ± SD. Statistics: Student's t test. ns=nonsignificant, $*=p \le 0.05$, $**=p \le 0.01$, $***=p \le 0.001$, $****=p \le 0.0001$. Scale bars: 5 µm, 1044 1045 inset 1 µm.

1046

Figure 6. HIV-1 functional proviruses are in complex with LEDGF. A) Confocal images of THP-1 cells expressing OR-GFP infected with HIV-1 ANCH3, 3 days p.i., to co-visualize CPSF6 (red) and LEDGF (white), compared to uninfected cells. On the

1050 left scatter plot of the percentages of CPSF6 found associated to LEDGF clusters per nucleus; on the right, scatter plot of the LEDGF and CPSF6 count per nucleus (n = 16 1051 cells). Scale bar: 5 µm, inset 2 µm. B) Confocal images of THP-1 cells expressing OR-1052 GFP infected with HIV-1 ANCH3 and expressing MCP-GFP infected with HIV-1 1053 1054 ANCH3 MS2, to co-visualize vDNA or vRNA (green), CPSF6 (red) and LEDGF (white). The scatter plot shows the percentage of vDNA or vRNA associated to LEDGF (n= 33 1055 (vDNA), 15 (vRNA)). The pie charts show the percentages of the 2 populations of 1056 LEDGF clusters associated to vDNA or vRNA: LEDGF alone or LEDGF in proximity of 1057 1058 CPSF6. Scale bar: 5 µm, inset 1 µm. All the analyses were performed in 3D. Statistics: Student's t test. ns=non-significant, *= $p \le 0.05$, **= $p \le 0.01$, ***= $p \le 0.001$, ****= $p \le 0.001$, *** 1059 1060 0.0001.

1061

1062 Figure 7. HIV-1 restructures the nuclear landscape for efficient replication. Once retrotranscribed and imported in the nucleus, HIV-1 genome integrates in open 1063 1064 chromatin regions enriched in LEDGF proteins, in less than 2 micrometers form the nuclear envelope. In addition, upon infection, CPSF6 clusters in SC35 nuclear 1065 1066 speckles with the viral IN restructuring the nuclear landscape and inducing HIV-1 1067 MLOs formation. Late retrotranscribed DNA and proviruses are excluded from CPSF6 clusters. HIV-1 transcription is favored by the nuclear environment surrounding proviral 1068 1069 sites that may profit from the proximity to SC35 and other splicing factors for post-1070 transcriptional processes. Created with BioRender.com.

Α





Figure 1



Figure 1

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.17.385567; this version posted November 17, 2020. The copyright holder for this preprint (which was not certified by peer review) in the utpr/credit st rights reserved. No reuse allowed without permission.



С

Α

Macrophages-like cells



Figure 2

20

0

DNA^{to RNA}

RNADONA

bioRxiv preprint doi: https://doi.org/1.DiviceDin 1988 2001 https://do



Non-dividing cells: THP-1



Α

В

С











Figure 4





100

80

60

40 20

IN to CPSFO

0

•

CP5FotoIN



40-20-0

:

bioRxiv preprint doi: https://doi.org/10.1101/2029 14 17-385567; this version posted November 17, 2020. The coveright holder for this preprint Ore Coveright for the author/funder. All rights reserved. No reuse allowed where the author of th

THP-1

CD4+ T



Uninfected

В













С



Figure 5



В

Α

Figure 6

