Decoding the biogenesis of HIV-induced CPSF6 puncta and their fusion with the nuclear speckle.

Chiara Tomasini[1*], Celine Cuche[1*], Selen Ay[1], Maxence Collard[1], Bin Cui[2], Mohammad Rashid[2], Shaoni Bhattacharjee[1], Julian Buchrieser[3], Charlotte Luchsinger[2], Cinzia Bertelli[2], Vladimir N. Uversky[4], Felipe Diaz-Griffero[2#], Francesca Di Nunzio[1#]

[1] *Institut Pasteur, Advanced Molecular Virology Unit, Department of Virology, Université Paris Cité, 75015 Paris, France*

[2]*Albert Einstein College of Medicine, Department of Immunology and Microbiology, New York 10461, USA*

[3] *Institut Pasteur,* Virus and Immunity Unit*, Department of Virology, Université Paris Cité, 75015 Paris, France.*

[4] *Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA*

 # Address correspondence to Francesca Di Nunzio, dinunzio@pasteur.fr, or Felipe Diaz-Griffero, felipe.diaz-griffero@einsteinmed.edu

** These authors have equally contributed*

Highlights

- The formation of HIV-induced CPSF6 puncta is critical for restoring HIV-1 nuclear reverse transcription (RT).
- CPSF6 protein lacking the FG peptide cannot bind to the viral core, thereby failing to form HIV-induced CPSF6 puncta.
- The FG peptide, rather than low-complexity regions (LCRs) or the mixed charge domains (MCDs) of the CPSF6 protein, drives the formation of HIV-induced CPSF6 puncta.
- HIV-induced CPSF6 puncta form individually and later fuse with nuclear speckles (NS) via the intrinsically disordered region (IDR) of SRRM2.

Summary

Viruses rely on host cellular machinery for replication. After entering the nucleus, the HIV genome accumulates in nuclear niches where it undergoes reverse transcription and integrates into neighboring chromatin, promoting high transcription rates and new virus progeny. Despite anti-retroviral treatment, viral genomes can persist in these nuclear niches and reactivate if treatment is interrupted, likely contributing to the formation of viral reservoirs.

The post-nuclear entry dynamics of HIV remain unclear, and understanding these steps is critical for revealing how viral reservoirs are established.

In this study, we elucidate the formation of HIV-induced CPSF6 puncta and the domains of CPSF6 essential for this process. We also explore the roles of nuclear speckle scaffold factors, SON and SRRM2, in the biogenesis of these puncta. Through genetic manipulation and depletion experiments, we demonstrate the key role of the intrinsically disordered region of SRRM2 in enlarging nuclear speckles in the presence of the HIV capsid.
We identify the FG domain of CPSF6 as essential for both puncta formation and binding to the viral core, which serves as the scaffold for CPSF6 puncta. While the low-complexity regions (LCRs) modulate CPSF6 binding to the viral capsid, they do not contribute to puncta formation, nor do the disordered mixed charge domains (MCDs) of CPSF6. These results demonstrate how HIV evolved to hijack host nuclear factors, enabling its persistence in the host.
Of note, this study provides new insights into the underlying interactions between host factors and viral components, advancing our understanding of HIV nuclear dynamics and offering potential therapeutic targets for preventing viral persistence.

52 **Introduction**

53 Since the discovery of HIV (Barre-Sinoussi et al., 1983), the initial stages of the viral life cycle
54 have been understood to primarily occur within the host cytoplasm. Only the pre-integration
55 complex, carrying the fully reverse-transcribed viral DNA, was believed to enter the nucleus
56 for integration into the host chromatin (Suzuki and Craigie, 2007). Recent studies highlighted
57 that the viral genome is transported in the nucleus via a shuttle that shields it from the hostile
58 cellular environment (Rasaiyaah et al., 2013). This shuttle is constituted by the viral capsid
59 (Blanco-Rodriguez and Di Nunzio, 2021; Blanco-Rodriguez et al., 2020; Chen et al., 2016;
60 Selyutina et al., 2020; Yamashita and Emerman, 2004; Zila et al., 2021), which comprises 250
61 hexamers and 12 pentamers (Pornillos et al., 2009). Within the structure of the capsid,
62 hydrophobic pockets exist between hexamers, which serve as targets for various nucleoporins,
63 especially the ones carrying FG repeats, facilitating the translocation of the capsid through the
64 nuclear pore complex (NPC) (Buffone et al., 2018; Di Nunzio, 2013; Di Nunzio et al., 2012;
65 Lelek et al., 2015; Matreyek et al., 2013; Price et al., 2014). Recent studies suggest that HIV
66 uses multiple FG regions of several nucleoporins to translocate through the NPC, acting as a
67 chaperone by itself (Dickson et al., 2024; Fu et al., 2024). Using a reductionist system of Nup98
68 condensates, it has been demonstrated that FG-mediated phase partitioning identifies specific
69 sites on the capsid that allow it to interact autonomously with these phases. These findings
70 complement the evolving understanding of the early stages of HIV infection, which has been
71 revisited in recent years. They unveil that crucial stages of early viral infection occur within the
72 host nucleus (Burdick et al., 2020; Dharan et al., 2020; Francis et al., 2020; Scoca et al., 2023;
73 Selyutina et al., 2020). Notably, it has been shown that the pre-integration complex forms within
74 the host nucleus (Muller et al., 2021; Scoca et al., 2023), and incoming viral RNA genomes
75 accumulate in nuclear niches containing , such as like cleavage and polyadenylation specificity
76 factor subunit 6 (CPSF6), RNA-binding protein SON and Splicing component, 35 kDa (SC35,
77 also known as serine/arginine-rich splicing factor 2 (SRSF2) (Francis et al., 2020; Rensen et
78 al., 2021; Scoca et al., 2023). However, the mechanistic requirements that govern post-nuclear
79 entry phases, which are crucial for a successful viral infection and the establishment of viral
80 reservoirs, remain enigmatic.

81

82 Specifically, the mechanism behinds the formation of nuclear niches containing viral
83 components and nuclear speckle (NS) factors, such as SC35—traditionally used as a marker
84 for nuclear speckles remains unclear. However, recent studies have revealed that SON and
85 serine/arginine repetitive matrix protein 2 (SRRM2) are essential for NS biogenesis (Fu and
86 Maniatis, 1990). Particularly, intrinsically disordered regions play a critical role in the NS
87 formation (Ilik et al., 2020). These membraneless organelles (MLOs) fulfil various cellular
88 functions besides splicing. Recent evidence indicates a direct role of NSs in cellular

transcription regulation, as their spatial proximity correlates with gene expression amplification, as demonstrated by live-cell imaging of heat-shock responsive genes (Chen et al., 2018; Zhang et al., 2021). HIV, being a virus capable of generating new particles through splicing and integrating into active host genes, finds NSs highly conducive for viral replication. HIV particles, along with their RNA genome, accumulate within nuclear MLOs enriched in NS factors (Rensen et al., 2021; Scoca et al., 2023). Notably, CPSF6, a paraspeckle factor first identified as a viral partner by KewalRamani's laboratory (Lee et al., 2010), has been clearly detected in HIV-induced CPSF6 puncta (Francis et al., 2020; Lee et al., 2010; Luchsinger et al., 2023; Rensen et al., 2021; Scoca et al., 2023). These puncta serve as hubs for nuclear reverse transcription (RT) and the formation of pre-integration complexes, which generate active proviruses detected outside but in close proximity to NSs (Li et al., 2021; Scoca et al., 2023).

In our study, we aim to elucidate how these HIV-induced CPSF6 form and identify the NS factors involved in their formation. CPSF6, along with NS factors, contains intrinsically disordered regions that can guide HIV-1 to the correct nuclear location for successful infection or allow the virus to remain sequestered during drug treatment, forming reservoirs. Importantly, viral reservoirs are the major bottleneck for curing the infection.

In this study, we investigate which disordered domain of CPSF6 is responsible for tracking the viral core and generating HIV-induced CPSF6 puncta in the host nucleus. Simultaneously, we elucidate the key component of NSs that, through its intrinsically disordered regions (IDRs), enable fusion with HIV-induced CPSF6 puncta, likely stabilizing them. Notably, we observed that the virus rebounds when anti-reverse transcription drugs are removed, but only if nuclear niches containing HIV, NS factors, and CPSF6 are present. If these niches are pharmacologically dismantled, viral rebound does not occur.

Overall, studying the biogenesis of HIV-induced nuclear niches is crucial for understanding how the virus navigates and persists in the nucleus of infected cells, and for designing new antiretroviral strategies.

Results

Critical Role of HIV-Induced CPSF6 Puncta in Restoring Nuclear RT After Anti-RT Therapy Discontinuation.

Upon nuclear entry, HIV enhances the formation of CPSF6 clusters, where RT ends. The treatment with the reversible RT inhibitor nevirapine (NEV) can trap the viral RNA genome in

these nuclear niches (Rensen et al., 2021; Scoca et al., 2023). Once NEV is removed, the trapped vRNA can resume RT entirely within the nucleus, a process we term nuclear RT . Here we demonstrate that this phenomenon is dependent on the presence of CPSF6 puncta, as their disruption by high doses of PF74 (25 μM) significantly impairs nuclear RT, shown by the absence of luciferase expression, similar to full NEV treatment (Figure 1A-B).

CPSF6 FG Domain is Required for HIV-Induced Puncta Formation.

Formation of CPSF6 puncta upon HIV-1 infection hinges into two key events: the entry of the HIV-1 core into the nucleus and the binding of CPSF6 to the HIV-1 core (Blanco-Rodriguez and Di Nunzio, 2021; Blanco-Rodriguez et al., 2020; Buffone et al., 2018; Zila et al., 2021). To determine the contribution of CPSF6's disordered domains for the formation of CPSF6 puncta upon HIV-1 infection, we correlated the binding of CPSF6 to the HIV-1 core with the formation of CPSF6 puncta. To this end, we first generated CPSF6 knockout (KO) THP-1 cells (Figure 2A-B) to eliminate the interference from the endogenous protein, which could affect the interpretation of results regarding the role of the analyzed CPSF6 domains. CPSF6 depletion in THP-1 cells was performed by CRISPR Cas9 technology. To completely eliminate the expression of the CPSF6 gene we selected single clones by limiting dilution. We identified a clone that was completely KO for CPSF6, confirmed through western blot and immunofluorescence (Fig. 2A-B) and we infected this clone and the control clone with HIV. CPSF6 puncta were detected only in the control-infected cells and not in the KO clone (Fig. 2B; Suppl. Fig.3A). The viral integrase (IN) was observed within CPSF6 puncta, consistent with previous studies (Francis et al., 2020; Rensen et al., 2021; Scoca et al., 2023), but absent in CPSF6 KO cells where viral IN was predominantly observed in the cytoplasm (Fig. 2B; Suppl. Fig.3A). Thus, we used KO cells for CPSF6 to assess the role of selected CPSF6 domains in HIV-induced condensates. We designed various CPSF6 deletion mutants (Figure 2C) to specifically assess the significance of the main disordered regions of CPSF6 protein (Suppl. Figure 2A) such as, the FG motif, the low complexity regions (LCRs), and the mixed charge domain (MCD), in the ability of CPSF6 to bind to the core and facilitate the formation of CPSF6 puncta. We investigated the role of the FG peptide by generating a mutant that exclusively lacks the FG peptide (ΔFG). Previous *in vitro* studies have shown that the FG peptide binds to the hydrophobic pocket formed between capsid hexamers (Buffone et al., 2018; Price et al., 2014). Here, we want to investigate the role of FG peptide in the context of the protein.

To further explore this, we developed an alternative plasmid by expanding the FG peptide deletion to include surrounding prion-like LCRs (ΔFG ΔLCR). These regions, outside the

158 CPSF6 context, have been identified as crucial for facilitating strong CPSF6 binding to capsid
159 lattices (Wei et al., 2022). In our study, we aim to evaluate their role within a more physiological
160 setting. Additionally, we assessed a CPSF6 variant that carries the 15-mer FG peptide flanked
161 by non-LCR sequences, such as those derived from Beta-adducin (ADD2), kindly provided by
162 Mamuka Kvaratskhelia (ΔLCR+ADD2). These protein segments are known for their high
163 flexibility, akin to the LCR of CPSF6 . Furthermore, to elucidate the contribution of the LCRs of
164 CPSF6 in the formation of CPSF6 puncta, we generated a mutant lacking both LCRs (ΔLCR)
165 (Figure 2C). Analysis of the MCD contribution to both the ability of CPSF6 to bind to the core
166 and formation of CPSF6 puncta was achieved by deleting the MCD and adding 3 nuclear
167 localization signals (3xNLS ΔMCD) since the deletion of the MCD results in a protein that
168 localizes mainly into the cytoplasm (Figure 2C; Suppl. Figure 2B).

169 To correlate the ability of CPSF6 to bind to the HIV-1 core with formation of CPSF6 puncta,
170 we expressed wild type and mutant CPSF6 constructs in THP-1 cells knockout for CPSF6.
171 Subsequently, we infected these cells with HIV-1 and analyzed the presence or absence of
172 CPSF6 clusters at 24 hours post-infection. Importantly, for the imaging experiment we
173 expressed CPSF6 WT and mutants without tags to avoid the formation of aggregates that
174 could interfere with our conclusions. Our data show that HIV-induced CPSF6 clusters can form
175 extremely rarely with the deletion mutant CPSF6 ADD2ΔLCR and with the mutant lacking the
176 FG or both the FG peptide and the LCRs (Fig. 2D-F; Suppl. Fig.3B). However, when we
177 analyzed the role of the MCD domain in CPSF6 puncta formation, which was indicated to be
178 important for condensing CPSF6 in NS (Greig et al., 2020), comparing the number of CPSF6
179 WT puncta induced by HIV infection with CPSF6 mutants revealed that the MCD domain does
180 not play a critical role in HIV-induced CPSF6 puncta formation (Fig. 2D-F). In addition, we
181 observed that the majority of analysed CPSF6 3xNLSΔMCD puncta contain vRNA inside,
182 similar to CPSF6 WT puncta (Figure 2G), thus corroborating the lack of a role for this
183 intrinsically disordered domain in HIV-induced CPSF6 puncta. Since the NLS domain from
184 SV40, which replaces the MCD, is highly basic and could potentially induce condensates, we
185 fused CPSF6 with a non-basic NLS (PY-NLS) or removed the NLS entirely. Even though these
186 two proteins do not efficiently enter the nucleus, the few that do manage to reach the nucleus
187 can host viral particles, as evidenced by the presence of IN. Many viruses are typically blocked
188 in the cytoplasm due to the presence of these mutants that are mainly cytoplasmic. However,
189 because we used a high viral dose, the blockage in the cytoplasm was not complete. As a
190 result, the viruses that successfully entered the nucleus induced the formation of clusters
191 associated with CPSF6-deleted mutants, indicating that the MCD is not critical for the formation
192 of HIV-induced CPSF6 puncta (Fig.2H). Similar to the MCD, when we compared CPSF6
193 truncated for the LCRs with CPSF6 WT, we observed that the LCRs do not contribute to CPSF6

194 puncta formation. Therefore, the FG peptide alone, without the LCRs, is the only CPSF6
195 domain required for their formation (Fig. 2D-F; Suppl. Fig.3B).

196 Next, we tested the ability of the different CPSF6 deletion mutants for their ability to bind the
197 viral core using a previously described capsid binding assay (Selyutina et al., 2018). Wild type
198 and mutant CPSF6 proteins were expressed in human 293T cells at similar levels (INPUT)
199 (Figure 3A). Extracts containing wild type and mutant CPSF6 proteins were incubated with
200 stabilized HIV-1 capsid tubes for 1 h at 25˚ C in the presence of 10 μM of PF74, which is a
201 small molecule that competes with CPSF6 for binding to the hydrophobic pocket formed
202 between hexamers that constitute the viral core (Buffone et al., 2018; Price et al., 2014). HIV-
203 1 capsid stabilized tubes were washed, and the bound proteins were eluted using Laemli buffer
204 (BOUND). For every construct, the percentage of bound protein relative to input in the
205 presence or absence of PF74 is shown (Figure 3B). Our results revealed that the absence of
206 the FG peptide (ΔFG) entirely abolished CPSF6's ability to bind to the viral core. In agreement,
207 simultaneous deletion of the FG motif and LCRs (ΔFG ΔLCR) resulted in a construct unable to
208 bind to the viral core. Similar outcomes were observed when the LCRs were replaced with
209 sequences derived from ADD2, even if the FG was present.

210 LCR-FG is notably more disordered than ADD2-FG, containing a high proportion of prolines
211 (48 out of 98 residues), which makes it mostly non-foldable (Figure 3C-L). Since proline is a
212 structure-disrupting residue, LRC-FG is not expected to adopt any secondary structure. In
213 contrast, ADD2-FG contains fewer prolines (15 out of 98 residues) but has many charged
214 residues. It is predicted to form two short α helices and a ß strand, arranged as: α helix - FG –
215 ß strand - α helix. ADD2-FG may form a flexible collapsed state, as its oppositely charged
216 residues are evenly distributed, potentially allowing polyelectrostatic compaction. This
217 suggests that FG within ADD2-FG may be less accessible for the interaction with the viral
218 core's hydrophobic pocket due to its involvement in this collapsed conformational ensemble
219 (Figure 3C-G, suppl. Fig.4). This aligns with the inability of CPSF6 carrying ADD2 in place of
220 the LCRs to induce CPSF6 puncta (Figure 2F). On the other hand, the deletion of only the two
221 LCRs, while keeping the FG peptide intact, resulted in unexpected findings. The ΔLCR mutant
222 exhibited a stronger binding affinity for the viral core when compared to the wild-type protein
223 (Figure 3B). These results suggest that the LCRs surrounding the FG motif are modulating the
224 affinity of CPSF6 to the viral core, which might be important for function. By contrast, deletion
225 of the MCD (ΔMCD) but retention of other regions, such as the FG peptide and the LCRs,
226 demonstrated a binding affinity to the viral core similar to that of the wild-type protein. These
227 results suggest that the MCD domain is not involved in the binding of CPSF6 to the viral core,
228 which is not surprising since the CPSF6 (1-358), which does not have an MCD, binds to the
229 viral core (Lee et al., 2010).

230 Thus, the viral capsid, through the FG peptide of CPSF6, constitutes the scaffold of HIV-
231 induced CPSF6 puncta.

232 In summary, our results suggest that the FG peptide is the main determinant involved in the
233 binding of CPSF6 to the viral capsid. Interestingly, our work implies that the LCRs may be
234 modulating the affinity of the FG motif for the viral core (Figure 3B). Recognition motifs that
235 mediate protein-protein interactions, such as the FG motif of CPSF6, are usually embedded
236 within longer intrinsically disordered regions that can modulate affinity of the interaction
237 (Karlsson et al., 2022). Taken together, our data show that the FG peptide coordinates both
238 the binding to the viral core and the induction of CPSF6 puncta. This coordination suggests
239 that the FG peptide plays a critical dual role in recognizing the viral capsid and facilitating the
240 cellular clustering of CPSF6, which may be part of the cellular response to viral entry.

241

242 **HIV-induced CPSF6 mutants puncta and Nuclear Speckles.**

243 The NS factor SC35, commonly used as a marker of NS, has been detected in HIV-induced
244 CPSF6 puncta (Figure 4A)**.** In this study, we investigated whether HIV-induced CPSF6 mutant
245 puncta are associated with SC35. We infected cells expressing CPSF6 wild-type (WT), CPSF6
246 3xNLS ∆MCD, CPSF6 ∆LCR, CPSF6 ∆MCD, and CPSF6 PY NLS ∆MCD, and examined
247 whether the nuclear puncta formed by the various CPSF6 proteins associate with SC35.
248 Imaging analysis revealed no significant difference in the association of SC35 with CPSF6 WT
249 or the CPSF6 mutants (Figure 4B), confirming that the disordered domains, LCRs, and MCDs
250 are dispensable for the formation of HIV-induced CPSF6 puncta that localize in the canonical
251 nuclear niche marked by SC35.

252

253 **Biogenesis of HIV-Induced CPSF6 Puncta carrying Nuclear Speckle factors.**

254 Nuclear speckle factors, particularly those involved in their biogenesis, such as SON and
255 SRRM2 (Ilik et al., 2020), have been identified as constituents of HIV-induced CPSF6 puncta
256 . However, the specific role of NSs in these puncta remains unclear. Most of the existing results
257 have been obtained through immunofluorescence experiments conducted several days post-
258 infection. In this study, we conducted a time course experiment to investigate the biogenesis
259 of HIV-induced CPSF6 puncta, which contain NS factors. Our goal was to capture the fusion
260 event between the HIV-induced CPSF6 puncta and NS, providing insights into the dynamics
261 of how HIV manipulates host nuclear structures during infection. We hypothesized that NS
262 factors might either be recruited during the initial formation of HIV-induced CPSF6 puncta,
263 shortly after the virus is released from the nuclear basket of the NPC, or later via the fusion
264 between NSs and HIV-induced CPSF6 puncta. To investigate this, we performed live imaging

265 to track CPSF6-mNeonGreen and NSs in cells expressing endogenous SRRM2 fused with a
266 Halo tag, using CRISPR Paint (courtesy of Roy Parker) (Lester et al., 2021).We fixed and
267 labeled samples at different time points post-infection, ranging from 6 h.p.i. to 30 h.p.i. (Figure
268 5A, Suppl. Fig.1G). At 6 h.p.i., 27% of HIV-induced CPSF6 puncta were still individual,
269 compared to only 9% at 30 h.p.i. (Figure 5B). Concurrently, 61% of HIV-induced CPSF6 puncta
270 were fused with NS at 6 h.p.i., rising to 75% at 30 h.p.i. (Figure 5B). This indicates a
271 progressive increase in the number of HIV-induced CPSF6 puncta fusing with NS over time
272 (Figure 5C). Overall, we detected individual CPSF6 clusters (green) and NSs (red) that quickly
273 fused, confirming that this fusion occurs within the two independent puncta, CPSF6 and
274 SRRM2, rather than during the formation of HIV-induced CPSF6 puncta (Figure 5A-C, movies
275 1A, 1B).
276 Taken together, these results suggest that HIV-induced CPSF6 puncta first form independently
277 of NS and later fuse with NS, causing an enlargement of NS as part of the hijacking process
278 by HIV-1.
279

280 **Role of SON and SRRM2 in the Fusion and Stabilization of HIV-Induced CPSF6 Puncta**
281 **within Nuclear Speckles.**

282 Macrophage-like cells, THP-1, were depleted for SON or SRRM2 using AUM*silence*™ ASO
283 technology (Gao et al., 2024; Marasca et al., 2022; Mazzeo et al., 2024; Zhang et al.). The
284 level of depletion of SON and SRRM2 was evaluated by immunofluorescence using antibodies
285 against SON and SRRM2 (Figure 6A). Both depleted cells were analyzed also for the presence
286 of NS, labelled by SC35 (Suppl. Fig.1A-E). However, recent findings suggest that the primary
287 target of the SC35 mAb is SRRM2 . To confirm this, cells depleted of SRRM2 were labelled
288 with antibodies against both SRRM2 and SC35 (Suppl. Figure 5). We have observed that the
289 reduction of SRRM2 resulted in a slight decrease in the mean intensity of SC35, whereas the
290 depletion of SON (Suppl. Fig.1F) did not have the same effect.

291 Subsequently, we infected THP-1 depleted cells for SRRM2 and SON or control cells with HIV-
292 1 and we fixed them at 48 h post-infection for immunofluorescence. We calculated the
293 percentage of CPSF6 clusters in HIV infected THP-1 control cells (approximately 78%) and in
294 HIV infected THP-1 cells depleted for SRRM2 (about 43%) and SON (around 66%). The results
295 indicate that the partial depletion of SRRM2 affects the formation of HIV-induced puncta, while
296 the depletion of SON slightly reduces their establishment (Figure 6B).

297

298 **The IDR of SRRM2 is a crucial element for the fusion of HIV-Induced CPSF6 puncta to**
299 **the nuclear speckles.**

CPSF6 contains several disordered regions (Di Nunzio et al., 2023) as well as NS factors(Ilik et al., 2020). Previous studies have established the role of IDR of SRRM2 in the biogenesis of NS (Ilik et al., 2020).

Here, we investigate the role of IDRs within NS factors, with a specific emphasis on SRRM2, in the fusion and stabilization of HIV-induced CPSF6 clusters. To address this inquiry, we utilized previously published HEK293 cell lines generated using the CRISPaint system (Lester et al., 2021), comprising two distinct lines: HaloTag SRRM2, SRRM2 full-length (FL) (1-2748 aa) fused with halo tag(insertion at amino acid [aa] 2,708), and a cell line lacking the C-terminal IDR of SRRM2, known as ∆IDR HaloTag SRRM2(1-429 aa, with halo insertion at aa 430) .

As anticipated, the truncated form of SRRM2 displayed a more diffuse distribution within the nucleus, without recruitment to NS (suppl. Fig.1A-E), and consequently lacked nuclear puncta (Figure 6C). On the other hand, the number of SON puncta was highly similar between the two cell lines (Figure 6D).

Subsequently, we quantified the formation of CPSF6 puncta in both cell lines infected with HIV-1. No significant difference in HIV-induced CPSF6 puncta formation was observed between HEK293 and HEK293 carrying the SRRM2 halo tag (~27%). However, a substantial reduction in HIV-induced CPSF6 puncta was evident in the cell line carrying the SRRM2 form that lacks the IDR (~11%) (Figure 6E; Suppl. Fig.1). Collectively, these results underscore the pivotal role of the SRRM2 IDR in the stabilization of HIV-induced CPSF6 puncta through their fusion with NSs (Figure 5A-C).

323 **Discussion**

324 HIV-1 capsid has transformed long-standing assumptions in the field. Previously considered
325 an undruggable viral target due to the belief that it disassembled shortly after HIV entered
326 target cells, whereas recent findings have revealed its essential role in nuclear import (Ay and
327 Di Nunzio, 2023; Blanco-Rodriguez and Di Nunzio, 2021; Blanco-Rodriguez et al., 2020; Chen
328 et al., 2016; Taylor and Fassati, 2024; Yamashita and Emerman, 2004; Zila et al., 2021).
329 Notably, the first-in-class antiretroviral capsid inhibitor, Lenacapavir, has shown remarkable
330 results in patients, demonstrating that the capsid, contrary to previous belief, can indeed be a
331 viable therapeutic target. While Lenacapavir improves patients' quality of life with only two
332 injections per year (Link et al., 2020; Segal-Maurer et al., 2022), no current antiretroviral drugs
333 provide a cure. This may be due to our incomplete understanding of certain aspects of HIV
334 biology. In this study, we shed light on the post-nuclear entry steps, a critical phase for the
335 establishment of viral reservoirs, which represent the main barrier to a cure. Recent findings
336 have shown that not only the HIV viral capsid translocates through the NPC, but that the viral
337 nuclear entry also enhances the formation of CPSF6 puncta. Additionally, it has been revealed
338 that RT is completed within the nucleus (Burdick et al., 2020; Dharan et al., 2020; Rensen et
339 al., 2021; Scoca et al., 2023; Selyutina et al., 2020). Furthermore, incoming viral RNA has been
340 observed to be sequestered in nuclear niches in cells treated with the reversible reverse
341 transcriptase inhibitor, NEV. When macrophage-like cells are infected in the presence of NEV,
342 the incoming viral RNA is held within the nucleus (Rensen et al., 2021; Scoca et al., 2023).
343 This scenario is comparable to what is observed in patients undergoing antiretroviral therapy.
344 Interestingly, we found that if CPSF6 puncta are pharmacologically dismantled, the nuclear
345 reverse transcription cannot be restored, indicating that HIV-induced CPSF6 puncta play a
346 crucial role in the viral life cycle, particularly for their potential role in forming viral reservoirs.
347 Understanding the biogenesis of these puncta could be a significant step towards deepening
348 our knowledge of HIV biology and providing additional tools to combat this pandemic virus.

349 Here we identify the disordered FG peptide essential for the binding with the viral capsid as
350 the inducer of HIV-induced CPSF6 puncta. Notably, CPSF6 protein lacking the FG peptide is
351 incapable of forming nuclear puncta. Moreover, we discovered that the two major intrinsically
352 disordered regions of CPSF6, the MCDs and the LCRs, are dispensable for the formation of
353 viral nuclear puncta. The MCDs of CPSF6 have been shown to provide cohesion for NS
354 condensation (Greig et al., 2020). In the context of HIV infection, the viral capsid induces the
355 formation of CPSF6 puncta that are depleted of MCDs carrying the viral RNA genome. These
356 CPSF6 puncta are highly similar to those formed by wild-type CPSF6, suggesting that MCDs
357 do not play a significant role in this process. Additionally, the LCRs of CPSF6 also do not
358 appear to influence puncta formation, as their depletion does not reduce the number of CPSF6

359 puncta. This indicates that neither the MCDs nor the LCRs are involved in CPSF6 puncta
360 formation during HIV infection. Surprisingly, when we assessed the ability of CPSF6 domains
361 to bind to the viral capsid, we observed that the deletion of LCRs increases CPSF6's ability to
362 bind to the viral capsid. We hypothesize that a change in charges may alter the binding
363 mechanism of CPSF6 when LCRs are absent. In scenarios where the FG motif is depleted,
364 we observed a dramatic inhibition of HIV-induced CPSF6 puncta formation and a lack of
365 binding to the viral core *in vitro*. On the flip side, the linkage between CPSF6 entities is
366 facilitated by the FG peptides' interaction with certain hydrophobic CA pockets along adjoining
367 hexamers (Wei et al., 2022). Therefore, it's conceivable that FG peptides, not involved in the
368 capsid's binding, could coalesce similarly to FG-Nups. These undergo phase separation,
369 forming condensates with nuclear pore complex-like permeability barrier features (Hülsmann
370 et al., 2012).

371 We have also identified HIV-induced CPSF6 puncta formation independent of NS at early
372 stages post-infection, with a progressive increase in CPSF6 puncta co-localizing with NS over
373 time. These results were obtained by focusing our studies on two scaffold proteins involved in
374 NS biogenesis: SON and SRRM2. Over the past approximately 0.6–1.2 billion years of
375 metazoan evolution, these two factors have undergone significant lengthening, unlike many
376 other proteins involved in splicing. This extension primarily occurred within their IDRs, which
377 are commonly associated with liquid-liquid phase separation (LLPS) and the formation of
378 biomolecular condensates (Rai et al., 2018). Co-depleting SRRM2 with SON, or depleting SON
379 in a cell line where it is deleted the intrinsically disordered C-terminus of SRRM2, abolished
380 the formation of NSs (Ilik et al., 2020). However, the depletion of only one factor does not
381 abolish NSs. Consistent with this finding, we observed that the depletion of SRRM2 does not
382 affect the presence of SON nuclear clusters. Similar results were obtained with cells genetically
383 modified to express the truncated form of SRRM2 lacking the IDRs. When the ΔIDR SRRM2-
384 halo tag was detected by the halo ligand, there was no recruitment of the truncated SRRM2
385 form in NSs. However, if SC35 is used as a target for antibodies, it can still be detected, albeit
386 with a much lower intensity signal than in cells expressing the full-length SRRM2. Additionally,
387 we observed a significant reduction in the detection of HIV-induced CPSF6 puncta. The few
388 CPSF6 clusters detected in these cells colocalized with the weak SC35 signal. These results
389 suggest that the IDR of SRRM2 plays an important role in HIV-induced CPSF6 puncta
390 stabilization, but rare clusters can appear, likely induced by redundant NS factors. This
391 indicates that further investigation is needed to better understand the hijacking of NS by HIV.

392 Taken together, our results reveal the intricate interplay between individual CPSF6 domains
393 and the viral capsid in dictating the nuclear fate of the virus. Concurrently, the IDR domain of
394 SRRM2 contributes to the enlargement of particular nuclear speckles with the fusion of HIV-

395 induced CPSF6 puncta. Lastly, these HIV-induced CPSF6 puncta necessitate of FG peptide
396 to engage the viral core.

397 Overall, this study could provide insights into the understanding of viral invasion and
398 persistence within the host.

399

400

401 **Materials and methods**

402

403 **Cell lines**

404 THP-1 cells are immortalized monocytic cells, which, once seeded, differentiate into
405 macrophage-like cells under phorbol 12-myristate 13-acetate (PMA) treatment (160 nM). THP-
406 1 cells were also engineered knocking out CPSF6. These cell populations These cells were
407 cultivated in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS) and 1%
408 penicillin–streptomycin solution (100 U/ml). HEK293T cells are human embryonic kidney cells
409 used to produce LVs. For figure 2, we used two engineered HEK293 strains, Halo tagged
410 SRRM2 HEK293 cells and Halo tagged SRRM2 ∆IDR HEK293 cells, kind gifts from Roy
411 Parker's lab (Lester et al., 2021), characterized by the Halo tagged SRRM2 protein. In Halo
412 tagged SRRM2 ∆IDR HEK293 cells Parker's lab deleted also the SRRM2 sequence encoding
413 for aa 430-2748. All the HEK293 strains were cultivated in Dulbecco's modified Eagle medium
414 supplemented with 10% FBS and 1% penicillin–streptomycin (100 U/ml).

415

416 **Bacteria strains**

417 All E. coli bacteria strains were grown in Luria-Bertani (LB) medium at 37°C. DHα competent
418 cells and Stellar Competent Cells were used for molecular cloning, while E.coli One-Shot
419 BL21star (DE3) cells were exploited for protein production.

420

421 **Plasmids**

422 To express the WT CPSF6/WT CPSF6-mNeonGreen and the mutant CPSF6/mutant CPSF6-
423 mNeonGreen clones, the correspondent coding sequences were engineered in pSICO
424 plasmids. The two original plasmids used were pSICO CPSF6-mNeonGreen and pLPCX
425 CPSF6 ADD2, gift from Mamuka's lab. HIV-1ΔEnvINHAΔNef plasmid encodes the ΔEnvHIV-
426 1 LAI (BRU) viral genome where the IN protein is fused to the HA tag while pNL4.3 ∆env ∆Nef
427 IRES GFP plasmid encodes the ΔEnvHIV-1 NL4.3 viral genome and contains also a GFP

428 sequence headed by an IRES. The pNL4.3 ∆env ∆NefLuc has the Luciferase cDNA as reporter
429 gene.

430

431 **CRISPR-Cas9 knockout in THP-1 cells**

432 To target CPSF6, three different crRNAs were used simultaneously (specific sequence: 5'-
433 TCGGGCAAATGGCCAGTCAAAGG-3', 5'-AGGACGGGGCCGTTTTCCAGGGG-3', and 5'-
434 CATGTAATCTCGGTCTTCTGGGG-3', all ordered from Integrated DNA Technologies, IDT).
435 Pre-designed unspecific crRNA was used as control (IDT). crRNA and tracrRNA were
436 resuspended in IDT Duplex Buffer according to the manufacturer's instructions. On the day of
437 the nucleofection, duplexes were formed by mixing equimolar concentration of crRNA and
438 tracrRNA, followed by 5-min annealing at 95°C. RNA duplexes were then mixed (1:2) with
439 TrueCut™ Cas9 Protein v2 for 10 min at RT to generate ribonucleoprotein (RNP) complexes.
440 2×10^5 THP1 cells were resuspended in P3 Primary Cell Nucleofector™ Solution, mixed with
441 RNP and Alt-R® Cas9 Electroporation Enhancer (90 pmol, IDT), and nucleofected in a 4D-
442 Nucleofector™ System using the P3 Primary Cell 4D-Nucleofector™X Kit S (program FI-110).
443 After nucleofection, cells were seeded in complete RPMI medium with 20% FBS. Three days
444 after nucleofection, cells were plated for clonal selection.

445

446 **Clonal selection of KO cell lines**

447 Seventy-two hours post nucleofection, cells were diluted in RPMI medium containing 20% FBS
448 and plated in five 96-well plates at 1 and 5 cells/well condition. After one-month, selected
449 microcolonies (50-100) are placed into 24- wells plate. Once that wells were near confluence,
450 cells were transferred into the well of a 6- well plate. After growing for another one-month, cells
451 were proceeded for western blot.

452

453 **Western blot**

454 Proteins were extracted on ice from THP-1 cells using RIPA buffer (20 mM HEPES, pH 7.6,
455 150 mM NaCl, 1% sodium deoxycholate, 1% Nonidet P-40, 0.1% SDS, 2mM EDTA, complete
456 protease inhibitor), and protein concentration was quantified using a detergent-compatible
457 (DC) protein assay (Pierce™ BCA Protein Assay Kit) with bovine serum albumin (BSA) as a
458 standard. 90 µg of total protein lysate were loaded onto SDS-PAGE 4- 12% Bis-Tris gels
459 (Invitrogen); an Ab rabbit anti-CPSF6 (1:500) and an anti-rabbit HRP-conjugated (1:5000)

460 were used for the detection of CPSF6, whereas the normalization was done by an Ab anti-
461 actin HRP-conjugated (1:3000). Visualization was carried out using an ECL solution .

462 **AUMsilence™ ASO**

463 AUMsilenceTM 352 ASOs were synthesized by AUM BioTech, LLC (Philadelphia, USA). THP-
464 1 negative control, THP-1 KD SRRM2 and THP-1 KD SON cells were differentiated with PMA
465 (160 nM) for 48 h then incubated 72 h with 10 µM of a AUMsilenceTM ASOs complementary
466 to the mRNA of SRRM2 and SON, respectively (scramble control AUMscrambleTM). All cells
467 were kept in incubator at 37°C and 5% CO2. These cells were cultivated in RPMI 1640 medium
468 supplemented with 10% fetal bovine serum (FBS) and 1% penicillin–streptomycin solution (100
469 U/ml).

470 **Cloning**

471 pSICO CPSF6-mNeonGreen was used to generate deletion mutants by exceeding different
472 regions from the original sequence according to Table2. The mutants were produced with and
473 without the mNeonGreen tag, except for pSICO CPSF6-ΔMCD and pSICO CPSF6-ΔMCD PY
474 NLS. All primers are specified in Table1.

475 For pSICO CPSF6-mNeonGreen ΔLCR ADD2, CPSF6 ΔLCR ADD2 sequence was amplified
476 by PCR from pLPCX CPSF6 ADD2 and specific primers were designed to add also BamHI
477 restriction sites at the extremities. Phusion Flash High-Fidelity PCR Master Mix was used and
478 the reaction was performed according to the manufacture datasheet (100 ng template DNA
479 and 56° annealing temperature). PCR products were treated with DpnI for 1h at 37°C and then
480 digested with BamHI restriction enzyme at 37°C for 1h. 2 ul of the backbone pSICO CPSF6-
481 mNeonGreen were digested with BamHI at 37°C for 3h and further treated with CIP. After gel
482 extraction of the DNA and purification, insert and backbone were ligated with T4 Ligase for 2
483 h at 22°C. 5 ul of the product were used to transform 50 ul of DH5α bacteria (30 min at 4°C,
484 45 sec at 42°C, 2 min at 4°C, incubation in SOC medium for 1h at 37°C and plating on LB agar
485 dishes).

486 All the other mutants were obtained using In-Fusion®Snap Assembly protocol, which allowed
487 to amplify the original plasmid deleting the small region of interest. The different primers used
488 are reported in Table1. As reported in the In-Fusion®Snap Assembly protocol
489 (https://www.takarabio.com/documents/User%20Manual/In/InFusion%20Snap%20Assembly
490 %20User%20Manual.pdf), 5 ng of the original plasmid were amplified with PrimeSTARMax
491 DNA polymerase in 35 PCR cycles (10 sec at 98°C, 15 sec at 55°C, 5sec/kb at 72°C). The
492 PCR products were then digested for 1h at 37°C using DpnI to get read of the original plasmid.
493 After plasmid cleaning up, it was circularized through a ligation of 15 min at 50°C. 2.5 ul of

494 products were used in Stellar™ Competent Cells' transformation, following the same
495 procedure used for DH5α transformation, previously explained.

496 **Lentiviral vectors and viral productions**

497 LVs and HIV-1 viruses were produced by transient transfection of HEK293T cells through
498 calcium chloride co-precipitation. Co-transfection was performed as follows: for LVs, 10 µg of
499 transfer vector, 10 µg of packaging plasmid (gag-pol-tat-rev), and 2.5 µg of pHCMV-VSV-G
500 envelope plasmid; for VSV- HIV-1ΔEnvIN$_{HA}$ΔNef-VPX viruses and VSV-pNL4.3 ∆env ∆Nef
501 IRES GFP-VPX, 10 µg HIV-1ΔEnvIN$_{HA}$ΔNef plasmid or pNL4.3 ∆env ∆Nef IRES GFP, 2.5 µg
502 of pHCMV-VSV-G plasmid and 3 ug of SIVMAC Vpx (Durand et al., 2013). After the collection
503 of the supernatant 48 h post-transfection, lentiviral particles were concentrated by
504 ultracentrifugation for 1 h at 22000 rpm at 4°C and stored at −80°C. LVs and viruses were
505 tittered by qPCR in HEK293T cells 3 days post-transduction.

506

507 **Cell transduction and infection**

508 THP-1 ctrl CRISPR clone 2 cells and THP-1 (duplex1-2-3 CRISPR) KO clone 4 cells were
509 differentiated with PMA (160 nM) for 72 h then transduced with different mutants of CPSF6 for
510 72h (MOI=1) and then infected for 30h with HIV-1ΔEnvIN$_{HA}$ΔNef Vpx (MOI=10) in presence of
511 Nevirapine (10uM). The medium was always supplemented with PMA (160 nM) .

512 For Halo tagged SRRM2 HEK 293 cells and Halo tagged SRRM2 ∆IDR HEK 293 cells, 2×10^5
513 cells were seeded on coverslips coated with polylysine in complete growth medium (DMEM,
514 GlutaMAX™-I, 10% FBS, and 1% P/S) and incubated at 37°C (5% CO_2) for 24 h. Cells were
515 then infected with the HIV-1ΔEnvIN$_{HA}$ΔNef (LAI) Bru (MOI 10) in complete growth medium
516 supplemented with Nevirapine (10µM) for 24h.

517 THP-1 control (scramble), THP-1 KD SRRM2 Cells and THP-1 KD SON cells were seeded on
518 coverslips and differentiated with PMA (160 nM) for 72 h. Then incubated 48 h with 10 µM of
519 a FANA ASOs (scramble control FANA (SCR-FANA), SRRM*2*-FANA and *SON*-FANA). ASOs
520 used in this study were designed and synthesized by AUM LifeTech (Philadelphia, PA, USA).
521 Next, cells were infected with the pNL4.3 ∆env ∆Nef IRES GFP-VPX (MOI 25) in complete
522 growth medium and incubated at 37°C in 5% CO_2 for 4 days.

523 In the four timepoints experiment, THP-1 cells were differentiated with PMA (160 nM) for 72 h.
524 Cells were then infected with HIV-1ΔEnvIN$_{HA}$ΔNef Vpx (MOI=10) in presence of Nevirapine
525 (10uM), in complete growth medium supplemented with PMA (160 nM), and incubated at 37°C
526 in 5% CO_2 for 6 h, 9 h, 12 h, 30 h post infection.

527

528 **Immunofluorescense Microscopy**

529 *Immunostaining*

530 On the day of fixation, the cells were washed with PBS and fixed with 4% PFA for 15 min. Cells
531 were treated with glycine 0.15% for 10 min, permeabilized with 0.5% Triton X-100 for 30 min,
532 and blocked with 1% BSA for 30 min. All antibody incubations were carried out at room
533 temperature in a dark humid chamber, for 1 h with primary antibodies and for 45 min with
534 secondary antibodies. Washes between antibody incubations and antibody dilutions were
535 done in 1% BSA.

536 Primary antibodies were diluted as follows: anti-HA 1:500, anti-CPSF6 1:400, anti-SC35 1:200,
537 anti-SON 1:200, anti-SRRM2 1:200, Secondary antibodies used were goat anti-rabbit Alexa-
538 488 1:300 and donkey anti-Rabbit Cy3 1:1000 for CPSF6, SON and SRRM2, goat anti-rat
539 Alexa-647 1:100 for IN-HA, goat anti-mouse Alexa-647 1:300 or goat anti-mouse Alexa-488
540 1:300 for SC35, donkey anti-mouse Alexa-647 1:300 for SRRM2.

541 Finally, cells were stained with Hoechst 33342 1:10000 for 5 min. Coverslips were mounted
542 on glass slides (Star Frost) with Prolong Diamond Antifade Mountant.

543 *HaloTag Labelling*

544 To detect the HaloTag in Halo tagged SRRM2 HEK 293 cells and Halo tagged SRRM2 ΔIDR
545 HEK 293 cells, we used HaloTag®TMR Ligand following the Technical Manual available at
546 www.promega.com/protocols. Specifically, cells were incubated in DMEM supplemented with
547 HaloTag®TMR Ligand (5 μM) for 15 min at 37°C. The ligand-containing medium was then
548 removed and replaced with an equal volume of 1X PBS, repeating the step twice and ending
549 with warm complete medium. Cells were incubated in complete culture medium for 30 min at
550 37°C. The medium was then removed and replaced with an equal volume of fresh warm culture
551 medium.

552 *Immuno-RNA FISH*

553 On the day of fixation, fixed cells were incubated in Permeabilization/Blocking buffer (1% BSA,
554 0.3% Triton X-100, 2mM Vanadyl Ribonucleoside complexes (VRCs) in RNase-free PBS) for
555 1 h before the antibodies' incubations. Antibodies were diluted in Permeabilization/Blocking
556 buffer. After the primary and secondary antibody staining (respectively of 1h and 45 min) with
557 the respective washes, cells were fixed for a second time in PFA 4% (in RNase-free PBS) for
558 10 min at RT with subsequent washes with RNase-free PBS. In the meantime, 40 pmol of
559 primary smiFISH probes (Tsanov et al., 2016) (24 smiFISH probes designed against HIV-1 pol

560 sequence (Rensen et al., 2021) were hybridized with 50 pmol of secondary FLAP probe
561 conjugated to a Cy5 fluorophore (Cy5/AATGCATGTCGACGAGGTCCGAGTGTAA/Cy5Sp/) in
562 1X NEBuffer 3 (diluted in RNase-free H_2O) using a thermocycler. The program setting follows:
563 3 min at 85°C, 3 min at 65°C and 5 min at 25°C. FISH-probe solution was then diluted 1:50 in
564 Hybridization buffer (90% Stellaris® RNA-FISH Hybridization Buffer, 10% Deionized
565 Formamide). After the samples were washed in Wash A buffer (20% Stellaris® RNA-FISH
566 Wash Buffer A, 10% Deionized Formamide, in RNase-free H_2O) at RT for 5 min, they were
567 placed on parafilm, covered with 50 ul of FISH-probe solution in Hybridization buffer and
568 incubated overnight at 37°C. The next day, cells were washed with Wash A buffer in the dark
569 at 37°C for 30 min. Afterwards, the samples were incubated for 10 min in Hoechst 333342
570 diluted 1:10000 in RNase-free H_2O and then washed with Wash B buffer for 5 min at RT in the
571 dark. Finally, the cells were washed in RNase-free H_2O before the coverslips were mounted
572 on microscopy slides using ProLongTM Diamond Antifade mounting medium. The mounting
573 medium was cured overnight at RT under the chemical hood and away from light.

574 *Images acquisition*

575 Images were acquired using a Zeiss LSM700 confocal inverted microscope, with a 63×
576 objective (Plan Apochromat, oil immersion, NA = 1.4), using diode lasers at 405 nm, 488 nm,
577 555 nm and 639 nm for the respective fluorophores. A pixel size of 0.07 µm was used.

578

579 **Live imaging on HIV-1 infected Halo tagged SRRM2 HEK 293 cells**

580 Halo tagged SRRM2 HEK293 cells were seeded and transduced with CPSF6-mNeonGreen
581 lentiviral vector (MOI 0.5) for 24 h. 0.3 x 10^6 Cells were then transferred on poly-L-Lysin coated
582 Ibidi-dishes and infected with HIV-1ΔEnvIN$_{HA}$ΔNef (MOI = 40) in presence of Nevirapine (10
583 uM) for 2 h. After having changed the medium and labeled with HaloTag®TMR Ligand (5 µM)
584 (see previous "HaloTag labelling" section) and Hoechst 333342 diluted 1:80000 in complete
585 DMEM medium, 4D movies were acquired.

586 3D movies were acquired using a Nikon Ti2-E Confocal Inverted Spinning Disk microscope,
587 using a 63x objective (Plan Apochromat, oil immersion, NA = 1.4) and a sCMOS Hamamatsu
588 camera, Orca Flash 4. Pixel size 6.5 µm, 2048 x 2044 pixel, QE 82%. For the Z stuck a Z piezo

589 stage was used, with 0.33 µm interval. For live imaging acquisitions, cells were placed in an
590 environmental chamber with 37°C temperature, 5% CO_2 and 21% O_2.

591

592 **Protein expression and purification.**

593 pET-11a vectors were used to express the HIV-1 capsid protein. Point mutations, A14C and
594 E45C were introduced using the QuikChange II site-directed mutagenesis kit (Stratagene)
595 according to the manufacturer's instructions. All proteins were expressed in *Escherichia coli*
596 one-shoot BL21star™ (DE3) cells (Invitrogen). Briefly, LB medium was inoculated with
597 overnight cultures, which were grown at 30°C until mid log-phase (A_{600}, 0.6-0.8). Protein
598 expression was induced with 1 mM isopropyl-β-d-thiogalactopyranoside (IPTG) overnight at
599 18°C. Cells were harvested by centrifugation at 5,000 x g for 10 min at 4°C, and pellets were
600 stored at −80°C until purification. Purification of capsid was carried out as follows. Pellets from
601 2 L of bacteria were lysed by sonication (Qsonica microtip: 4420; A=45; 2 minutes; 2 seconds
602 on; 2 seconds off for 12 cycles), in 40 ml of lysis buffer (50 mM Tris pH=8, 50mM NaCl, 100mM
603 β-mercaptoethanol and Complete EDTA-free protease inhibitor tablets). Cell debris were
604 removed by centrifugation at 40,000g for 20 min at 4°C. Proteins from the supernatant were
605 precipitated by incubation with 1/3 of volume of saturated ammonium sulfate containing 100
606 mM β-mercaptoethanol for 20 min at 4°C and centrifugation at 8,000g for 20 min at 4°C.
607 Precipitated proteins were resuspended in 30 ml of buffer A (25mM MES pH6.5, 100 mM β-
608 mercaptoethanol) and sonicated 2-3 times (Qsonica microtip: 4420; A=45; 2 minutes; 1 second
609 on; 2 seconds off). Sample was dialyzed 3 times in buffer A (2 h, overnight, 2 h). The sample
610 was sonicated and diluted in 500 ml of buffer A and was chromatographed sequentially on a 5
611 ml HiTrap™ Q HP column and on a 5 ml HiTrap™ SP FF column (GE Healthcare), both pre-
612 equilibrated with buffer A. The capsid protein was eluted from HiTrap™ SP FF column using a
613 linear gradient from 0-2 M of NaCl. Absorbance at 280 nm was checked to take the eluted
614 fraction that had higher protein levels. Pooled fractions were dialyzed 3 times (2 h, overnight,
615 2 h) in storage buffer (25mM MES, 2M NaCl, 20 mM β-mercaptoethanol). Sample was
616 concentrated using centricons to a concentration of 20 mg/ml and stored at -80°C.

617

618 **Assembly of stabilized HIV-1 capsid tubes**

619 1 mL of monomeric capsid (3 mg/mL or 1 mg/mL) was dialyzed in SnakeSkin dialysis tubing
620 10,000 MWCO (Thermo Scientific) against a buffer that is high in salt and contains a reducing
621 agent (buffer 1: 50 mM Tris, pH 8, 1 M NaCl, 100 mM β-mercaptoethanol) at 4°C for 8 hours.
622 Subsequently the protein was dialyzed against the same buffer without the reducing agent β-
623 mercaptoethanol (buffer 2: 50 mM Tris, pH 8, 1 M NaCl) at 4°C for 8 hours. The absence of β-
624 mercaptoethanol in the second dialysis allows formation of disulfide bonds between Cysteine

625 14 and 43 inter-capsid monomers in the hexamer. Finally the protein is dialyzed against buffer
626 3 (20 mM Tris, pH 8,0, 40 mM NaCl) at 4°C for 8 hours. Assembled complexes were kept at
627 4°C up to 1 month.

628

629 **Capsid binding assay protocol**

630 Human HEK293T cells were transfected for 24 h with a plasmid expressing the specified
631 CPSF6 variant tagged with mNeonGreen. Cell media was completely removed and cells were
632 lysed in 300 µL of capsid binding buffer (CBB: 10 mM Tris, pH 8,0, 1,5 mM MgCl2, 10 mM
633 KCl) by scrapping off the plate. Cells were rotated at 4°C for 15 min and then centrifuged to
634 remove cellular debris (21,000 *x g* ,15 minutes, 4 °C). Cell lysates were incubated with
635 stabilized HIV-1 capsid tubes for 1 h at 25 °C. Subsequently, stabilized HIV-1 capsid tubes
636 were washed by pelleting the complexes by centrifugation at 21,000 x g for 2 min. Pellets were
637 washed using by resuspension in CBB or PBS. Pellets were resuspended in Laemmli buffer
638 1X and analyzed by Western blotting using anti-p24 or anti-mNeonGreen antibodies.

639

640 **Bioinformatics analysis of CPSF6**

641 Intrinsic disorder propensity of CPSF6 was evaluated using the Rapid Intrinsic Disorder
642 Analysis Online platform (RIDAO) (https://ridao.app/) designed to predict disordered residues
643 and regions in a query protein based on its amino acid sequence (Dayhoff and Uversky, 2022).
644 RIDAO yields results by combining the outputs of several commonly used per-residue disorder
645 predictors, such as PONDR® VLXT (Romero et al., 2001), PONDR® VL3(Peng et al., 2006),
646 PONDR® VLS2B(Peng et al., 2005), PONDR® FIT (Xue et al., 2010), as well as IUPred2
647 (Short) and IUPred2 (Long)(Dosztanyi et al., 2005a, b). RIDAO also computes a mean disorder
648 score for each residue based on these. In the resulting intrinsic disorder profile, the disorder
649 score of 0.5 is the threshold between order and disorder, where residues/regions above 0.5
650 are disordered, and residues/regions below 0.5 are ordered. The disorder score of 0.15 is the
651 threshold between order and flexibility, where residues/regions with the disorder scores above
652 0.15 but below 0.5 are flexible, and residues/regions below 0.15 are highly ordered.

653 Amino acid compositions of the intrinsically disordered C-terminal domain (residues 261-358)
654 of human CPSF6 and its different variants (CPSF6 ΔFG ΔLCR, CPSF6 ΔLCR, CPSF6 ΔFG,
655 and CPSF6 ADD2 ΔLCR) were analyzed to evaluate the relative abundance of prion-like low
656 complexity region (LCR) defining uncharged, charged, and Pro residues in these protein
657 regions. The corresponding values of the relative abundance of these residue groups were
658 calculated by dividing numbers of prion-like LCR defining uncharged (Ala, Gly, Val, Phe, Tyr,
659 Leu, Ile, Ser, Thr, Pro, Asn, Gln), charged (Asp, Glu, Lys, Arg), and Pro residues by total

660 number of amino acids in the corresponding protein fragments. As references, we used the
661 corresponding data for protein sequences deposited to the UniProtKB/Swiss-Prot database
662 that provides information on the overall distribution of amino acids in nature (Bairoch et al.,
663 2005); PDB Select 25 (Berman et al., 2000), which is a subset of structures from the Protein
664 Data Bank with less than 25% sequence identity, biased towards the composition of proteins
665 amenable to crystallization studies; and DisProt 3.4 that is comprised of a set of consensus
666 sequences of experimentally determined disordered regions (Sickmeier et al., 2007). Per-
667 residue intrinsic disorder propensities of the LCR-FG and ADD2-FG sequences were
668 evaluated by PONDR® VLXT (Romero et al., 2001), which is sensitive to local peculiarities of
669 the amino acid sequences freely available at http://www.pondr.com/ (accessed on August 03,
670 2024). Linear distribution of the net charge per residue (NCPR) within the LCR-FG and ADD2-
671 FG sequences were evaluated by CIDER (Holehouse et al., 2017), which is a webserver for
672 the analysis of a wide range of the physicochemical properties encoded by IDP sequences
673 freely available at http://pappulab.wustl.edu/CIDER (accessed on August 03, 2024).
674 Secondary structure propensities of the LCR-FG and ADD2-FG sequences were evaluated by
675 PSIPRED (McGuffin et al., 2000), which is a highly accurate secondary structure prediction
676 method freely available to non-commercial users at http://globin.bio.warwick.ac.uk/psipred/
677 (accessed on August 03, 2024).

678

679

680 **IMAGING AND STATISTICAL ANALYSIS**

681 All images were analysed using Fiji Software. More in detail, for the count of the nuclear
682 speckles a macro was computed to segment cellular nuclei and to select and count the nuclear
683 speckles. For the nuclei segmentation a Gaussian Blur with sigma=2 was used and for the
684 channel related to the nuclear speckles the set threshold was at 7000-46012 (min and max).
685 For the counting, a “size =10 – Infinity summarize” was used for the nuclei, whereas for the
686 nuclear speckles the size was reduced to “size=0-Infinity summarize”.

687 For live imaging analysis, Arivis software was used to reconstruct the 3D movies.

688 All data were statistically analyzed with GraphPad Prism 9 (GraphPad Software, La Jolla
689 California USA, www.graphpad.com). Calculations were performed and figures were drawn
690 using Excel 365 or GraphPad Prism 8.0. Statistical analysis was performed, with Wilcoxon
691 matched paired t-tests or Mann–Whitney unpaired t-tests. Spearman correlation coefficients
692 (r) were calculated using GraphPad Prism.

693

694

695 **Figure legends**

696 **Figure 1. Role of HIV-induced CPSF6 puncta in the nuclear RT upon removal of NEV. A)**
697 **A)** THP-1 cells, infected with VSV-G ∆env HIV-1 (NL4.3) ∆R LUC (MOI 10) in presence or not
698 of Nevirapine (10 µM) for 5 days, or in presence of Nevirapine (10 µM) for 2 days and then the
699 remaining 3 days without drug or in presence of Nevirapine (10 µM) for 2 days then in presence
700 of PF74 (25 µM). Confocal microscopy images, to verify the presence of CPSF6 clusters, the
701 cells are stained with anti-CPSF6 antibody (green). Nuclei are stained with Hoechst (blue).
702 Scale bar 10 µm. **B)** Luciferase Assay, to verify luciferase expression in the aforementioned
703 samples. Luciferase values were normalized by total proteins revealed with the Bradford kit.
704 One-way ANOVA statistical test with multiple comparison was performed (****=p<0.0001;
705 *=p<0.05; ns=p>0.05). Data are representative of two independent experiments.
706

707 **Figure 2. Role of CPSF6 Domains in HIV-Induced CPSF6 Puncta. A)** Western blots
708 demonstrate CPSF6 depletion using a specific antibody against CPSF6 in THP-1 cells
709 subjected to CRISPR Cas9 methods: CRISPR Cas9 bulk (left), and CRISPR Cas9 clones
710 selected by limiting dilution (right). Each condition is normalized for actin labeling. The ratio
711 between the intensity signal of CPSF6 and actin was analyzed via ImageJ and is plotted below
712 each western blot. **B)** Confocal microscopy images of THP-1 ctrl CRISPR clone 2 cells (Ctrl 2)
713 and THP-1 duplex1-2-3 CRISPR clone 4 cells (CPSF6 KO 4) infected with HIV-
714 1ΔEnvINHAΔNef Vpx (LAI) Bru (MOI 10) in the presence of Nevirapine (10 µM). The cells are
715 stained 30 h p.i. with anti-CPSF6 antibody and anti-HA antibodies to detect HA tagged
716 integrase (IN). **C)** Schema of CPSF6 isoform 588aa deletion mutants. **D)** Confocal microscopy
717 images of THP-1 CPSF6 KO cells, transduced with different mutants of CPSF6, infected with
718 HIV-1ΔEnvINHAΔNef Vpx (LAI) Bru (MOI 10) in presence of Nevirapine (10 µM). The cells are
719 stained with CPSF6 and HA antibody 30 h p.i.. Scale bar 5µm. **E)** Analysis of the number of
720 CPSF6 clusters in THP-1 CPSF6 KO cells transduced with different mutants of CPSF6, not
721 infected or infected in presence of Nevirapine (10 µM) (the number of analyzed cells is shown
722 under the x-axis). Statistical test: ordinary one-way ANOVA (****=p<0.0001; ***=p<0.001;
723 *=p<0.05; ns=p>0.05). **F)** The plot compares the number of CPSF6 clusters per cell in THP-1
724 CPSF6 KO cells transduced with different mutants of CPSF6, infected with HIV-1 in the
725 presence of Nevirapine (10 µM). Statistical test: ordinary one-way ANOVA (****=p<0.0001;
726 ns=p>0.05). **G)** Confocal microscopy images of THP-1 CPSF6 KO clone 4, non-transduced
727 and non-infected or transduced with WT CPSF6 and CPSF6 3xNLSΔMCD and infected with
728 HIV-1ΔEnvINHAΔNef Vpx (LAI) Bru (MOI 10) in presence of Nevirapine (10 µM). Immuno-
729 RNA FISH: the cells are stained with CPSF6 (green) antibody and with 24 probes against HIV-
730 1 Pol sequence (gray)(RNA-FISH) 25 h p.i. Nuclei are stained with Hoechst (blue). Scale bar

731 10 µm. Violin plot presenting the percentage of CPSF6 clusters colocalizing with the viral RNA
732 in THP-1 CPSF6 KO clone 4 cells transduced with LVs expressing CPSF6 WT or CPSF6
733 3xNLSΔMCD (respectively n=73 and n=103) and infected with HIV-1ΔEnvIN$_{HA}$ΔNef Vpx (LAI)
734 Bru (MOI 10) in presence of Nevirapine (10 µM). A total of 198 CPSF6 WT clusters and 264
735 CPSF6 3xNLSΔMCD clusters were counted. Statistical test: unpaired t-test, ns=p>0.05. **H)**
736 Confocal microscopy images of THP-1 KO CPSF6 cells transduced with WT CPSF6 and
737 CPSF6 ΔMCD without NLS, with 3xNLS or with PY NLS, respectively. Cells were differentiated
738 for 3 days, transduced with CPSF6 lentiviral vectors (MOI 1) for 3 days and infected for 24 h
739 with HIV-1ΔEnvINHAΔNef Vpx (LAI) Bru (MOI 10) in presence of Nevirapine (10 uM) (left
740 panels). The panels on the right show transduced and uninfected cells. CPSF6 and the IN
741 tagged with the HA are labeled with anti-CPSF6 (green) and anti-HA (white) antibodies,
742 respectively. Nuclei are stained with Hoechst (blue). The arrows show CPSF6 puncta in
743 colocalization with IN-HA. Scale bar 10µm.

744

745 **Figure 3. Evaluation of CPSF6 Deletion Mutants' Binding Capacity to the Viral Core. A)**
746 Ability of wild type and mutant CPSF6 proteins to bind to the HIV-1 core. Cellular extracts
747 derived from human 293T cells expressing similar levels of the indicated CPSF6 proteins
748 (INPUT) were incubated with HIV-1 capsid stabilized tubes for 1 hour at room temperatures in
749 the presence and absence of 10 µM PF74, as described in materials and methods. As a carrier
750 control, we utilized DMSO. Subsequently, HIV-1 capsid stabilized tubes were washed, and the
751 bound proteins eluted 1X Laemmli buffer 1X. The BOUND fractions were analysed by Western
752 blotting using antibodies against neon-GFP and the HIV-1 capsid. **B)** Experiments were
753 repeated at least three times and the average BOUND fraction relative to the INPUT fraction
754 normalized to wild type binding are shown for the different CPSF6 mutants. *** indicates a p-
755 value < 0.001; **** indicates a p-value < 0.0001; and ns indicates no significant difference as
756 determined by unpaired t-tests. **C)** Physicochemical characteristics of the LCR-FG and ADD2-
757 FG sequences. Intrinsic disorder predispositions evaluated by PONDR® VLXT. Position of the
758 FR segment within the LCR-FG and ADD2-FG sequences is shown as gray shaded area. **D)**
759 Linear distribution of the net charge per residue (NCPR) within the LCR-FG sequence
760 evaluated by CIDER. **E)** Linear distribution of the net charge per residue (NCPR) within the
761 ADD2-FG sequence evaluated by CIDER. **F)** Secondary structure propensity of the LCR-FG
762 sequence evaluated by PSIPRED. **G)** Secondary structure propensity of the ADD2-FG
763 sequence evaluated by PSIPRED. **H)** Analysis of the peculiarities of the amino acid
764 compositions of the intrinsically disordered C-terminal domain (residues 261-358) of human
765 CPSF6 and its different mutants. Relative abundance of prion-like LCR defining uncharged
766 residues in analyzed protein segments. **I)** Relative abundance of proline residues in analyzed

767 protein segments. **L)** Relative abundance of charged residues in analyzed protein segments.
768 The values were calculated by dividing numbers of prion-like LCR defining uncharged (Ala,
769 Gly, Val, Phe, Tyr, Leu, Ile, Ser, Thr, Pro, Asn, Gln), Pro, and charged (Asp, Glu, Lys, Arg)
770 residues by the total number of amino acids in the respective protein fragments. Corresponding
771 values for all protein sequences deposited in the UniProtKB/Swiss-Prot database, PDB
772 Select25, and DisProt are shown for comparison.

773

774 **Figure 4. A)** Epifluorescence microscopy images of both infected and non-infected 293T cells
775 showing the presence of CPSF6 clusters only in the infected condition. CPSF6 and SC35 are
776 labeled with anti-CPSF6 (red) and anti-SC35 (gray) antibodies, respectively. Nuclei are stained
777 with Hoechst (blue). Scale bar 5µm. **B)** Confocal microscopy images of THP-1 KO CPSF6
778 cells, differentiated for 3 days, transduced with CPSF6 lentiviral vector (MOI 1) (specifically
779 WT CPSF6, CPSF6 ∆LCRs, CPSF6 ∆MCD with 3xNLS, without NLS, or with PY NLS) for 3
780 days and infected for 24 h with HIV-1ΔEnvINHAΔNef Vpx (LAI) Bru (MOI 10) in presence of
781 Nevirapine (10 uM). CPSF6 and nuclear speckles were labeled with anti-CPSF6 (green) and
782 anti-SC35 (red) antibodies, respectively. Nuclei are stained with Hoechst (blue). Scale bar
783 10µm. The percentage of CPSF6 puncta associated with SC35 per field of view is shown in
784 the graph. N cells were counted in each condition and a one-way ANOVA statistical test with
785 multiple comparison was performed; ns= p value > 0.05.
786

787

788 **Figure 5. Dynamics of the HIV-induced CPSF6 puncta formation and their fusion with**
789 **NSs. A)** Time course of infection of THP-1, 6h.p.i., 9h.p.i., 12h.p.i., 30h.p.i. or non infected.
790 Cells were stained with antibodies against CPSF6 (green) and SC35 (red). **B)** The graph
791 shows the percentage of CPSF6 puncta associated with NS or adjacent to NS or isolated from
792 NS at different time post-infection. **C)** The graph shows the progression of CPSF6 puncta
793 associated to NS during the time post-infection. N indicates the number of cells analyzed. One-
794 way ANOVA statistical test with multiple comparison was performed; ns= p value > 0.05; ****
795 indicates a p-value < 0.0001.

796

797 **Figure 6. Role of SRRM2 and SON in the Formation of HIV-Induced CPSF6 puncta. A)**
798 Depletion of SON and SRRM2 in THP-1 cells using AUM*silence*™ ASO technology. The
799 degree of depletion is quantified by mean intensity through immunofluorescence using
800 antibodies against SON and SRRM2, respectively. Scale bar 5µm. **B)** The percentage of

801 CPSF6 puncta formation is quantified by IF in THP-1 cells knocked down for SON, SRRM2,
802 and control (Ctrl) infected with HIV-1 (MOI 25) for 48 h. CPSF6 is stained with an antibody
803 against CPSF6 (green), and nuclei are stained with Hoechst (blue). The graph on the right
804 reports the percentage of CPSF6 clusters calculated from more than 100 cells. Scale bar
805 10µm. Experiments were performed at least twice. **C)** Confocal microscopy images of ∆IDR
806 HaloTag SRRM2 HEK293 and HaloTag SRRM2 HEK293 cells stained with the halo tag ligand
807 (red), and nuclei (blue). Scale bar 10µm. **D)** Confocal microscopy images of HaloTag SRRM2
808 HEK293 and ∆IDR HaloTag SRRM2 HEK293 cells, both labeled with anti-SRRM2 (red) and
809 anti-SON (gray) antibodies. Nuclei are stained with Hoechst (blue). Scale bar 10µm. Statistical
810 studies are summarized in the violin plot which displays the distribution of the number of SON
811 puncta per cell in the two conditions. N cells were counted and Kolmogorov Smirnov test was
812 performed, ns= $p>0.05$. **E)** Confocal microscopy images of HaloTag SRRM2 HEK293 and
813 ∆IDR HaloTag SRRM2 HEK293 cells, either non-infected or infected for 24 h with ΔEnvHIV-1
814 LAI (BRU) (MOI 10) in the presence of Nevirapine (10 µM). CPSF6 and SC35 are labeled with
815 anti-CPSF6 (red) and anti-SC35 (gray) antibodies, respectively. Nuclei are stained with
816 Hoechst (blue). The plot shows the mean ± SD of the percentage of cells with CPSF6 clusters
817 calculated in n fields of view (n=24, 29, 32); N is the number of cells analyzed for each of the
818 three different cell lines; an unpaired t-test was performed, ****=$p<0.0001$, ns=$p>0.05$. Scale
819 bar 10µm. Experiments were performed at list twice.

820

821 **Movies 1 A and B. CPSF6 MLOs can form independently from SRRM2 MLOs in HEK293**
822 **SRRM2 HaloTag cells.** Spinning Disk confocal images of HEK293 SRRM2 HaloTag cells
823 acquired 9 hours after ΔEnvHIV-1 LAI (BRU) Vpx (MOI 10) infection in presence of Nevirapine
824 (10 µM). Cells were previously transduced for 24h with CPSF6-mNeonGreen lentiviral vector
825 (MOI 0.5) to detect CPSF6 and SRRM2 was visualized incubating cells with the TMR-halo tag
826 ligand. The nuclei were stained with Hoechst. Acquisitions were performed continuously for 27
827 minutes.

828

829 **Supplementary Figures**

830 **Suppl. Figure 1.** Multiple examples (**A-E**) of confocal microscopy images of HaloTag SRRM2
831 HEK293 and ∆IDR HaloTag SRRM2 HEK293 cells, either non-infected or infected for 24 h with
832 ΔEnvHIV-1 LAI (BRU) (MOI 10) in the presence of Nevirapine (10 µM). CPSF6 and SC35 are
833 labeled with anti-CPSF6 (red) and anti-SC35 (gray) antibodies, respectively. Nuclei are stained
834 with Hoechst (blue). Scale bar 10µm. **F)** HEK 293 HaloTag SRRM2 or ∆IDR were labelled with
835 Halo ligand (red) and an anti-SON antibody (gray), nuclei are stained by Hoechst (blue). Scale

836 bar 10µm. **G)** THP-1 infected with ΔEnvHIV-1 LAI (BRU) Vpx were labelled with specific
837 antibodies against CPSF6 (green) and SC35 (red), nuclei are stained with Hoechst (blue).
838 Scale bar 10µm.

839

840 **Suppl. Figure 2. A)** Per-residue intrinsic disorder propensity of the CPSF6 isoform 588 aa
841 evaluated by the Rapid Intrinsic Disorder Analysis Online platform (RIDAO) (Dayhoff and
842 Uversky, 2022) that yields results for IU-Pred_short (yellow line), IUPred_long (blue line),
843 PONDR® VL3 (green line), PONDR® VLXT (black line), PONDR® VSL2 (red line), and
844 PONDR® FIT (pink line) and computes a mean disorder score for each residue based on these
845 predictors (MDP, thick, dark pink, dashed line). Light pink shadow represents MDP error
846 distribution. The thin black line at the disorder score of 0.5 is the threshold between order and
847 disorder, where residues/regions with disorder scores above 0.5 are disordered, and
848 residues/regions below 0.5 are ordered. The dashed line at the disorder score of 0.15 is the
849 threshold between order and flexibility, where residues/regions above 0.15 are flexible, and
850 residues/regions below 0.15 are highly ordered (upper). Schema of the deletion mutants of
851 CPSF6 (bottom). **B)** Lentiviral Vector Transduction of PMA-Differentiated THP-1 Cells
852 expressing CPSF6 ΔMCD fused to mNeonGreen (left), CPSF6 NLSΔMCD fused to
853 mNeonGreen (center), CPSF6 3xNLSΔMCD fused to mNeonGreen (right). CPSF6 is
854 represented in green, and nuclei are stained in blue. Scale bar 5µm. **C)** Lentiviral Vector
855 Transduction of PMA-Differentiated THP-1 cells expressing CPSF6 deletion mutants fused to
856 mNeonGreen. CPSF6 is represented in green, and nuclei are stained in blue. Scale bar 5µm.

857 **Suppl. Figure 3. A)** Multiple examples of confocal microscopy images of THP-1 ctrl CRISPR
858 clone 2 cells and THP-1 CPSF6 KO 4 cells infected with HIV-1ΔEnvINHAΔNef Vpx (LAI) Bru
859 (MOI 10) in the presence of Nevirapine (10 µM). The cells are stained 30 h p.i. with CPSF6
860 (green) and HA (red) antibodies to detect integrase (IN). **B)** Multiple examples of THP-1 CPSF6
861 KO clone 4 cells transduced with different LVs carrying CPSF6 WT or mutants and stained
862 with CPSF6 and HA antibody 30 h p.i.. Scale bar 5µm.

863

864 **Suppl. Figure 4.** Sequences of FG and LCRs or substituted amino acids sequences analyzed
865 in figure 5.

866

867 **Suppl. Figure 5.** SRRM2 depletion is confirmed using antibodies against both SRRM2 and
868 SC35 by IF. Statistical analysis: One-Way ANOVA (****=p<0.0001; *=p<0.05; ns=p>0.05).

ACKNOWLEDGMENTS

F.D.N. is supported by the Institut Pasteur and ANRS grants (ECTZ192036, ECTZ137593), ANR-PRCI grant, Sidaction grant. C.T. is supported by fellowships Sidaction, ANR-PRCI. S.A. is supported by ANRS fellowship ECTZ204694. We gratefully acknowledge the UtechS Photonic BioImaging platform (Imagopole), C2RT at Institut Pasteur. We thank the NIH AIDS Reagents program to support us with precious reagents and Addgene.

Bibliography

Ay, S., and Di Nunzio, F. (2023). HIV-Induced CPSF6 Condensates. J Mol Biol *435*, 168094.

Bairoch, A., Apweiler, R., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M.*, et al.* (2005). The Universal Protein Resource (UniProt). Nucleic Acids Res *33*, D154-159.

Barre-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C.*, et al.* (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science *220*, 868-871.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res *28*, 235-242.

Blanco-Rodriguez, G., and Di Nunzio, F. (2021). The Viral Capsid: A Master Key to Access the Host Nucleus. Viruses *13*.

Blanco-Rodriguez, G., Gazi, A., Monel, B., Frabetti, S., Scoca, V., Mueller, F., Schwartz, O., Krijnse-Locker, J., Charneau, P., and Di Nunzio, F. (2020). Remodeling of the Core Leads HIV-1 Preintegration Complex into the Nucleus of Human Lymphocytes. J Virol *94*.

Buffone, C., Martinez-Lopez, A., Fricke, T., Opp, S., Severgnini, M., Cifola, I., Petiti, L., Frabetti, S., Skorupka, K., Zadrozny, K.K.*, et al.* (2018). Nup153 Unlocks the Nuclear Pore Complex for HIV-1 Nuclear Translocation in Nondividing Cells. J Virol *92*.

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

Chen, N.Y., Zhou, L., Gane, P.J., Opp, S., Ball, N.J., Nicastro, G., Zufferey, M., Buffone, C., Luban, J., Selwood, D.*, et al.* (2016). HIV-1 capsid is involved in post-nuclear entry steps. Retrovirology *13*, 28.

Chen, Y., Zhang, Y., Wang, Y., Zhang, L., Brinkman, E.K., Adam, S.A., Goldman, R., van Steensel, B., Ma, J., and Belmont, A.S. (2018). Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. J Cell Biol *217*, 4025-4048.

Dayhoff, G.W., 2nd, and Uversky, V.N. (2022). Rapid prediction and analysis of protein intrinsic disorder. Protein Sci *31*, e4496.

Dharan, A., Bachmann, N., Talley, S., Zwikelmaier, V., and Campbell, E.M. (2020). Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. Nature Microbiology *5*, 1088-1095.

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., Danckaert, A., Fricke, T., Perez, P., Fernandez, J., Perret, E., Roux, P., Shorte, S., Charneau, P., Diaz-Griffero, F.*, et al.* (2012). Human nucleoporins promote HIV-1 docking at the nuclear pore, nuclear import and integration. PLoS One *7*, e46037.

Di Nunzio, F., Uversky, V.N., and Mouland, A.J. (2023). Biomolecular condensates: insights into early and late steps of the HIV-1 replication cycle. Retrovirology *20*, 4.

915 Dickson, C.F., Hertel, S., Tuckwell, A.J., Li, N., Ruan, J., Al-Izzi, S.C., Ariotti, N., Sierecki, E., Gambin, Y.,
916 Morris, R.G.*, et al.* (2024). The HIV capsid mimics karyopherin engagement of FG-nucleoporins.
917 Nature.
918 Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005a). IUPred: web server for the prediction of
919 intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics *21*,
920 3433-3434.
921 Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005b). The pairwise energy content estimated
922 from amino acid composition discriminates between folded and intrinsically unstructured proteins. J
923 Mol Biol *347*, 827-839.
924 Durand, S., Nguyen, X.N., Turpin, J., Cordeil, S., Nazaret, N., Croze, S., Mahieux, R., Lachuer, J., Legras-
925 Lachuer, C., and Cimarelli, A. (2013). Tailored HIV-1 vectors for genetic modification of primary human
926 dendritic cells and monocytes. J Virol *87*, 234-242.
927 Francis, A.C., Marin, M., Singh, P.K., Achuthan, V., Prellberg, M.J., Palermino-Rowland, K., Lan, S.,
928 Tedbury, P.R., Sarafianos, S.G., Engelman, A.N.*, et al.* (2020). Publisher Correction: HIV-1 replication
929 complexes accumulate in nuclear speckles and integrate into speckle-associated genomic domains.
930 Nat Commun *11*, 6165.
931 Fu, L., Weiskopf, E.N., Akkermans, O., Swanson, N.A., Cheng, S., Schwartz, T.U., and Görlich, D. (2024).
932 HIV-1 capsids enter the FG phase of nuclear pores like a transport receptor. Nature *626*, 843-851.
933 Fu, X.D., and Maniatis, T. (1990). Factor required for mammalian spliceosome assembly is localized to
934 discrete regions in the nucleus. Nature *343*, 437-441.
935 Gao, L., Zhang, W., Zhang, L., Gromova, B., Chen, G., Csizmadia, E., Cagle, C., Nastasio, S., Ma, Y.,
936 Bonder, A.*, et al.* (2024). Silencing of aryl hydrocarbon receptor repressor restrains Th17 cell
937 immunity in autoimmune hepatitis. Journal of Autoimmunity *143*, 103162.
938 Greig, J.A., Nguyen, T.A., Lee, M., Holehouse, A.S., Posey, A.E., Pappu, R.V., and Jedd, G. (2020).
939 Arginine-Enriched Mixed-Charge Domains Provide Cohesion for Nuclear Speckle Condensation. Mol
940 Cell *77*, 1237-1250.e1234.
941 Holehouse, A.S., Das, R.K., Ahad, J.N., Richardson, M.O., and Pappu, R.V. (2017). CIDER: Resources to
942 Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. Biophys J *112*, 16-21.
943 Hülsmann, B.B., Labokha, A.A., and Görlich, D. (2012). The permeability of reconstituted nuclear
944 pores provides direct evidence for the selective phase model. Cell *150*, 738-751.
945 Ilik, I.A., Malszycki, M., Lubke, A.K., Schade, C., Meierhofer, D., and Aktas, T. (2020). SON and SRRM2
946 are essential for nuclear speckle formation. Elife *9*.
947 Karlsson, E., Schnatwinkel, J., Paissoni, C., Andersson, E., Herrmann, C., Camilloni, C., and Jemth, P.
948 (2022). Disordered Regions Flanking the Binding Interface Modulate Affinity between CBP and NCOA.
949 J Mol Biol *434*, 167643.
950 Lee, K., Ambrose, Z., Martin, T.D., Oztop, I., Mulky, A., Julias, J.G., Vandegraaff, N., Baumann, J.G.,
951 Wang, R., Yuen, W.*, et al.* (2010). Flexible use of nuclear import pathways by HIV-1. Cell Host Microbe
952 *7*, 221-233.
953 Lelek, M., Casartelli, N., Pellin, D., Rizzi, E., Souque, P., Severgnini, M., Di Serio, C., Fricke, T., Diaz-
954 Griffero, F., Zimmer, C.*, et al.* (2015). Chromatin organization at the nuclear pore favours HIV
955 replication. Nat Commun *6*, 6483.
956 Lester, E., Ooi, F.K., Bakkar, N., Ayers, J., Woerman, A.L., Wheeler, J., Bowser, R., Carlson, G.A.,
957 Prusiner, S.B., and Parker, R. (2021). Tau aggregates are RNA-protein assemblies that mislocalize
958 multiple nuclear speckle components. Neuron *109*, 1675-1691 e1679.
959 Li, C., Burdick, R.C., Nagashima, K., Hu, W.S., and Pathak, V.K. (2021). HIV-1 cores retain their integrity
960 until minutes before uncoating in the nucleus. Proc Natl Acad Sci U S A *118*.
961 Link, J.O., Rhee, M.S., Tse, W.C., Zheng, J., Somoza, J.R., Rowe, W., Begley, R., Chiu, A., Mulato, A.,
962 Hansen, D.*, et al.* (2020). Clinical targeting of HIV capsid protein with a long-acting small molecule.
963 Nature *584*, 614-618.
964 Luchsinger, C., Lee, K., Mardones, G.A., KewalRamani, V.N., and Diaz-Griffero, F. (2023). Formation of
965 nuclear CPSF6/CPSF5 biomolecular condensates upon HIV-1 entry into the nucleus is important for
966 productive infection. Sci Rep *13*, 10974.

967 Marasca, F., Sinha, S., Vadalà, R., Polimeni, B., Ranzani, V., Paraboschi, E.M., Burattin, F.V., Ghilotti, M.,
968 Crosti, M., Negri, M.L.*, et al.* (2022). LINE1 are spliced in non-canonical transcript variants to regulate
969 T cell quiescence and exhaustion. Nature Genetics *54*, 180-193.
970 Matreyek, K.A., Yucel, S.S., Li, X., and Engelman, A. (2013). Nucleoporin NUP153 phenylalanine-
971 glycine motifs engage a common binding pocket within the HIV-1 capsid protein to mediate lentiviral
972 infectivity. PLoS Pathog *9*, e1003693.
973 Mazzeo, L., Ghosh, S., Di Cicco, E., Isma, J., Tavernari, D., Samarkina, A., Ostano, P., Youssef, M.K.,
974 Simon, C., and Dotto, G.P. (2024). ANKRD1 is a mesenchymal-specific driver of cancer-associated
975 fibroblast activation bridging androgen receptor loss to AP-1 activation. Nature Communications *15*,
976 1038.
977 McGuffin, L.J., Bryson, K., and Jones, D.T. (2000). The PSIPRED protein structure prediction server.
978 Bioinformatics *16*, 404-405.
979 Muller, T.G., Zila, V., Peters, K., Schifferdecker, S., Stanic, M., Lucic, B., Laketa, V., Lusic, M., Muller, B.,
980 and Krausslich, H.G. (2021). HIV-1 uncoating by release of viral cDNA from capsid-like structures in the
981 nucleus of infected cells. Elife *10*.
982 Peng, K., Radivojac, P., Vucetic, S., Dunker, A.K., and Obradovic, Z. (2006). Length-dependent
983 prediction of protein intrinsic disorder. BMC Bioinformatics *7*, 208.
984 Peng, K., Vucetic, S., Radivojac, P., Brown, C.J., Dunker, A.K., and Obradovic, Z. (2005). Optimizing long
985 intrinsic disorder predictors with protein evolutionary information. J Bioinform Comput Biol *3*, 35-60.
986 Pornillos, O., Ganser-Pornillos, B.K., Kelly, B.N., Hua, Y., Whitby, F.G., Stout, C.D., Sundquist, W.I., Hill,
987 C.P., and Yeager, M. (2009). X-ray structures of the hexameric building block of the HIV capsid. Cell
988 *137*, 1282-1292.
989 Price, A.J., Jacques, D.A., McEwan, W.A., Fletcher, A.J., Essig, S., Chin, J.W., Halambage, U.D., Aiken, C.,
990 and James, L.C. (2014). Host cofactors and pharmacologic ligands share an essential interface in HIV-1
991 capsid that is lost upon disassembly. PLoS Pathog *10*, e1004459.
992 Rai, A.K., Chen, J.X., Selbach, M., and Pelkmans, L. (2018). Kinase-controlled phase transition of
993 membraneless organelles in mitosis. Nature *559*, 211-216.
994 Rasaiyaah, J., Tan, C.P., Fletcher, A.J., Price, A.J., Blondeau, C., Hilditch, L., Jacques, D.A., Selwood, D.L.,
995 James, L.C., Noursadeghi, M.*, et al.* (2013). HIV-1 evades innate immune recognition through specific
996 cofactor recruitment. Nature *503*, 402-405.
997 Rensen, E., Mueller, F., Scoca, V., Parmar, J.J., Souque, P., Zimmer, C., and Di Nunzio, F. (2021).
998 Clustering and reverse transcription of HIV-1 genomes in nuclear niches of macrophages. EMBO J *40*,
999 e105247.
1000 Romero, P., Obradovic, Z., Li, X., Garner, E.C., Brown, C.J., and Dunker, A.K. (2001). Sequence
1001 complexity of disordered protein. Proteins *42*, 38-48.
1002 Scoca, V., Morin, R., Collard, M., Tinevez, J.Y., and Di Nunzio, F. (2023). HIV-induced membraneless
1003 organelles orchestrate post-nuclear entry steps. J Mol Cell Biol *14*.
1004 Segal-Maurer, S., DeJesus, E., Stellbrink, H.J., Castagna, A., Richmond, G.J., Sinclair, G.I., Siripassorn,
1005 K., Ruane, P.J., Berhe, M., Wang, H.*, et al.* (2022). Capsid Inhibition with Lenacapavir in Multidrug-
1006 Resistant HIV-1 Infection. N Engl J Med *386*, 1793-1803.
1007 Selyutina, A., Bulnes-Ramos, A., and Diaz-Griffero, F. (2018). Binding of host factors to stabilized HIV-1
1008 capsid tubes. Virology *523*, 1-5.
1009 Selyutina, A., Persaud, M., Lee, K., KewalRamani, V., and Diaz-Griffero, F. (2020). Nuclear Import of the
1010 HIV-1 Core Precedes Reverse Transcription and Uncoating. Cell Rep *32*, 108201.
1011 Sickmeier, M., Hamilton, J.A., LeGall, T., Vacic, V., Cortese, M.S., Tantos, A., Szabo, B., Tompa, P., Chen,
1012 J., Uversky, V.N.*, et al.* (2007). DisProt: the Database of Disordered Proteins. Nucleic Acids Res *35*,
1013 D786-793.
1014 Suzuki, Y., and Craigie, R. (2007). The road to chromatin — nuclear entry of retroviruses. Nature
1015 Reviews Microbiology *5*, 187-196.
1016 Taylor, I.A., and Fassati, A. (2024). The capsid revolution. J Mol Cell Biol *15*.

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

Wei, G., Iqbal, N., Courouble, V.V., Francis, A.C., Singh, P.K., Hudait, A., Annamalai, A.S., Bester, S., Huang, S.W., Shkriabai, N.*, et al.* (2022). Prion-like low complexity regions enable avid virus-host interactions during HIV-1 infection. Nat Commun *13*, 5879.

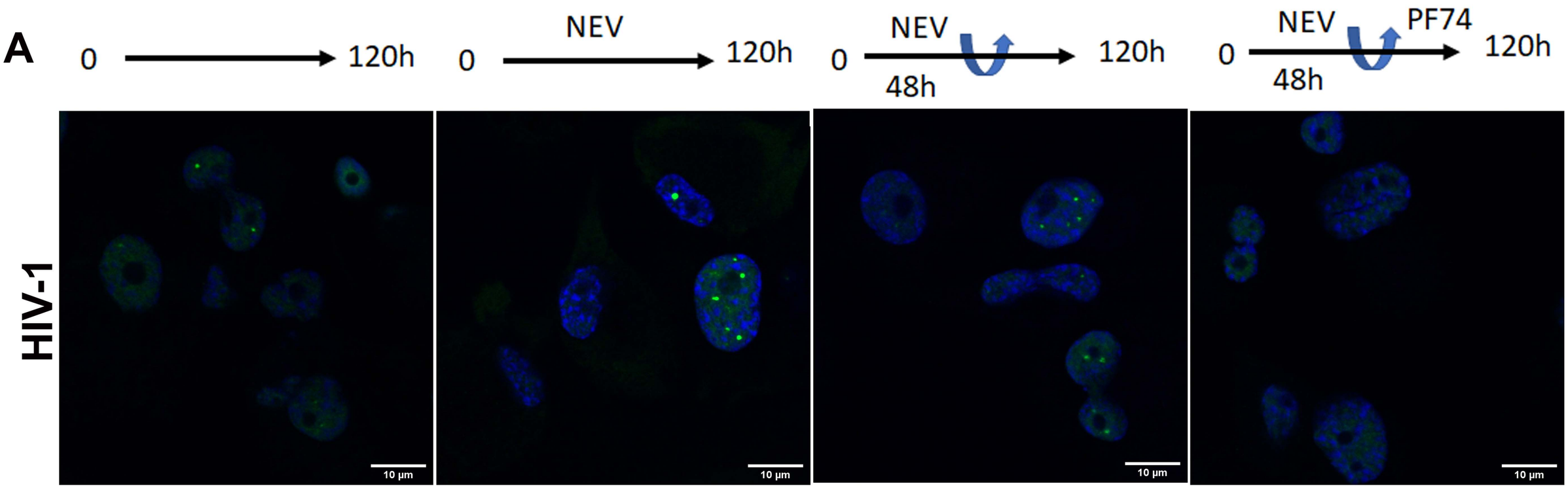
Xue, B., Dunbrack, R.L., Williams, R.W., Dunker, A.K., and Uversky, V.N. (2010). PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. Biochim Biophys Acta *1804*, 996-1010.

Yamashita, M., and Emerman, M. (2004). Capsid is a dominant determinant of retrovirus infectivity in nondividing cells. J Virol *78*, 5670-5678.

Zhang, L., Zhang, Y., Chen, Y., Gholamalamdari, O., Wang, Y., Ma, J., and Belmont, A.S. (2021). TSA-seq reveals a largely conserved genome organization relative to nuclear speckles with small position changes tightly correlated with gene expression changes. Genome Res *31*, 251-264.

Zhang, Z.-Y., Harischandra, D.S., Wang, R., Ghaisas, S., Zhao, J.Y., McMonagle, T.P., Zhu, G., Lacuarta, K.D., Song, J., Trojanowski, J.Q.*, et al.* TRIM11 protects against tauopathies and is down-regulated in Alzheimer's disease. Science *381*, eadd6696.

Zila, V., Margiotta, E., Turonova, B., Muller, T.G., Zimmerli, C.E., Mattei, S., Allegretti, M., Borner, K., Rada, J., Muller, B.*, et al.* (2021). Cone-shaped HIV-1 capsids are transported through intact nuclear pores. Cell *184*, 1032-1046 e1018.


A
0
120h
0
NEV
120h
0
NEV
48h
120h
0
NEV
PF74
48h
120h
HIV-1
10 µm
10 µm
10 µm
10 µm

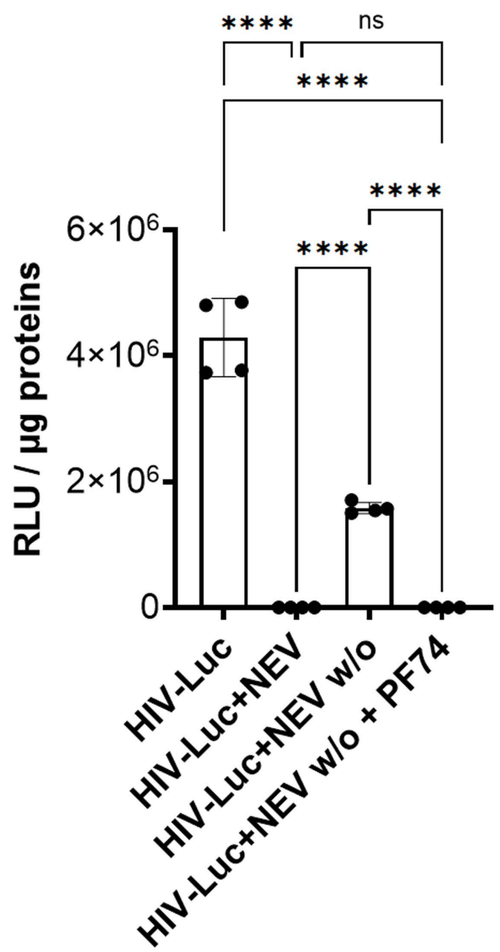

B
****
ns
****
****
6×10^6
4×10^6
2×10^6
0
RLU / µg proteins
HIV-Luc
HIV-Luc+NEV
HIV-Luc+NEV w/o
HIV-Luc+NEV w/o + PF74

Figure 1

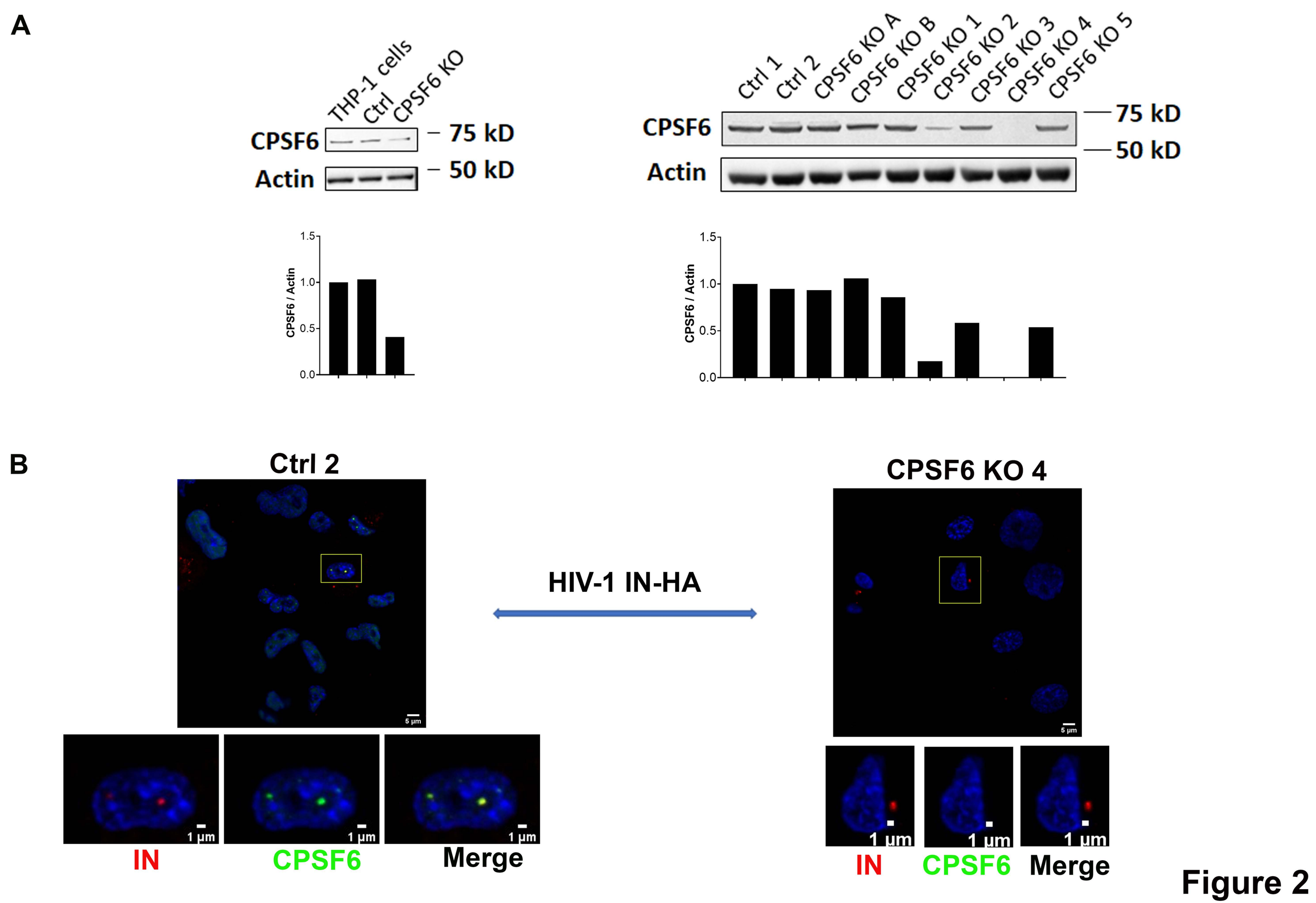

A
THP-1 cells
Ctrl
CPSF6 KO
CPSF6
Actin
75 kD
50 kD
CPSF6 / Actin
1.5
1.0
0.5
0.0
Ctrl 1
Ctrl 2
CPSF6 KO A
CPSF6 KO B
CPSF6 KO 1
CPSF6 KO 2
CPSF6 KO 3
CPSF6 KO 4
CPSF6 KO 5
B
Ctrl 2
CPSF6 KO 4
HIV-1 IN-HA
5 µm
1 µm
IN
CPSF6
Merge

Figure 2

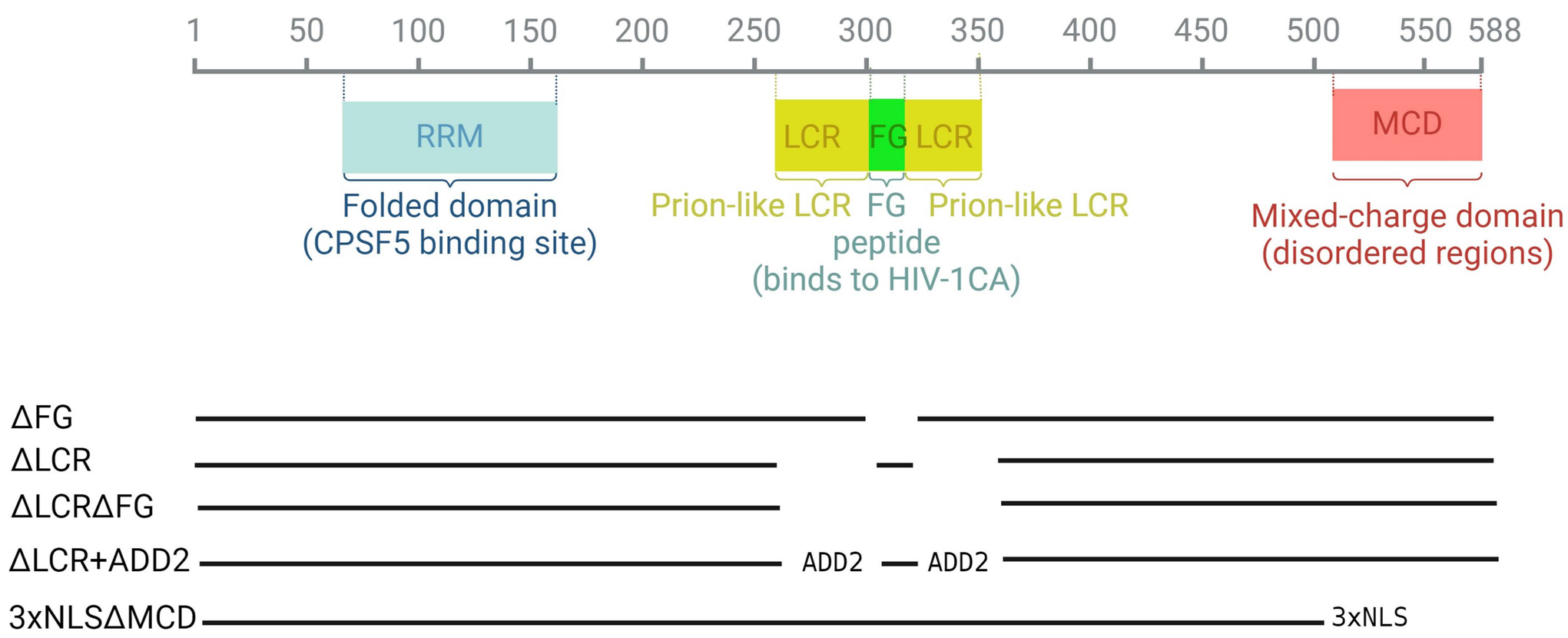

C
1
50
100
150
200
250
300
350
400
450
500
550
588
RRM
LCR
FG
LCR
MCD
Folded domain
(CPSF5 binding site)
Prion-like LCR
FG
peptide
(binds to HIV-1CA)
Prion-like LCR
Mixed-charge domain
(disordered regions)
ΔFG
ΔLCR
ΔLCRΔFG
ΔLCR+ADD2
ADD2
ADD2
3xNLSΔMCD
3xNLS

Figure 2

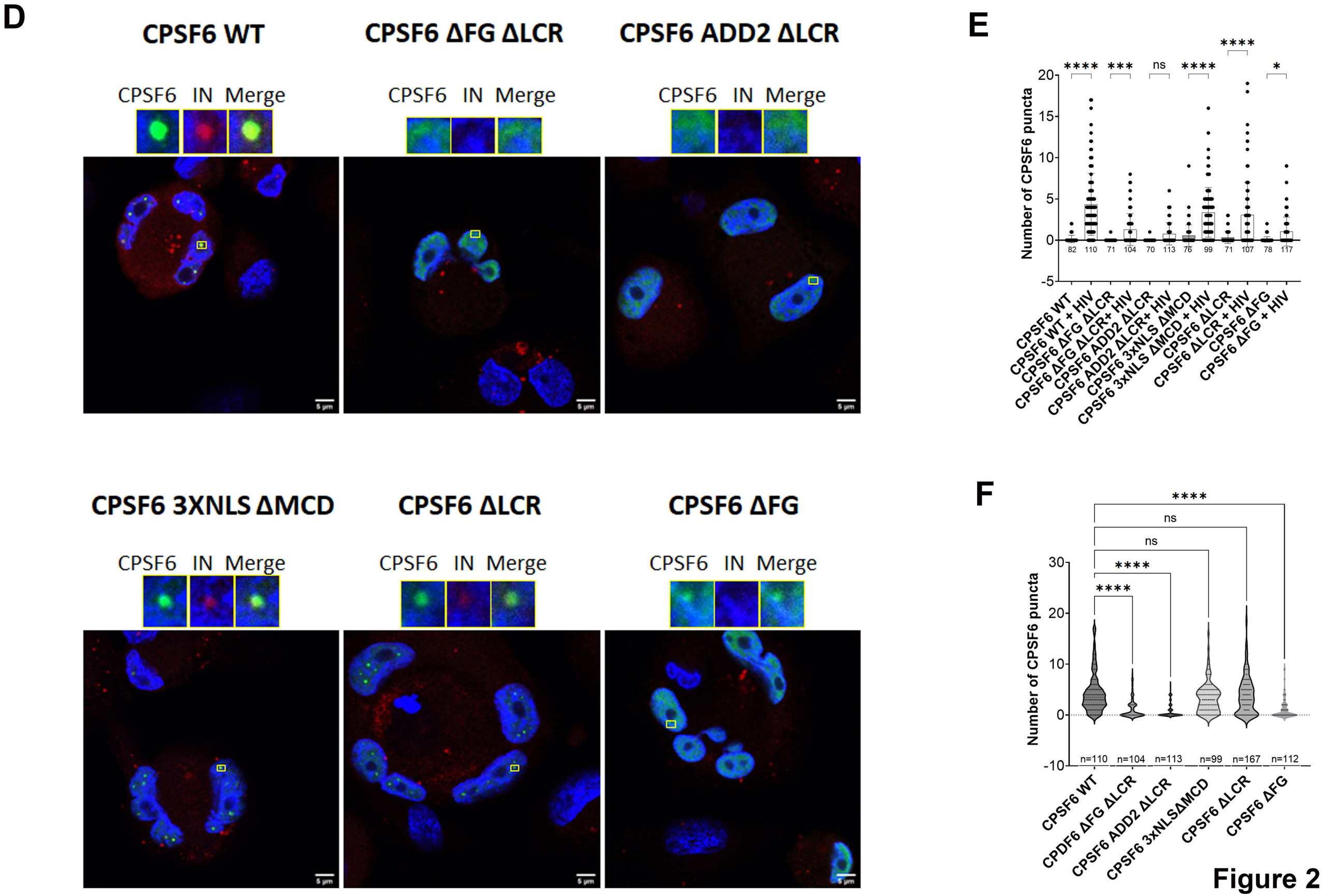

D
CPSF6 WT
CPSF6 ΔFG ΔLCR
CPSF6 ADD2 ΔLCR
CPSF6 IN Merge
CPSF6 3XNLS ΔMCD
CPSF6 ΔLCR
CPSF6 ΔFG
5 µm
E
Number of CPSF6 puncta
82 110 71 104 70 113 76 99 71 107 78 117
CPSF6 WT
CPSF6 WT + HIV
CPSF6 ΔFG ΔLCR
CPSF6 ΔFG ΔLCR+ HIV
CPSF6 ADD2 ΔLCR
CPSF6 ADD2 ΔLCR+ HIV
CPSF6 3xNLS ΔMCD
CPSF6 3xNLS ΔMCD + HIV
CPSF6 ΔLCR
CPSF6 ΔLCR + HIV
CPSF6 ΔFG
CPSF6 ΔFG + HIV
F
Number of CPSF6 puncta
n=110 n=104 n=113 n=99 n=167 n=112
CPSF6 WT
CPDF6 ΔFG ΔLCR
CPSF6 ADD2 ΔLCR
CPSF6 3xNLSΔMCD
CPSF6 ΔLCR
CPSF6 ΔFG

Figure 2

G

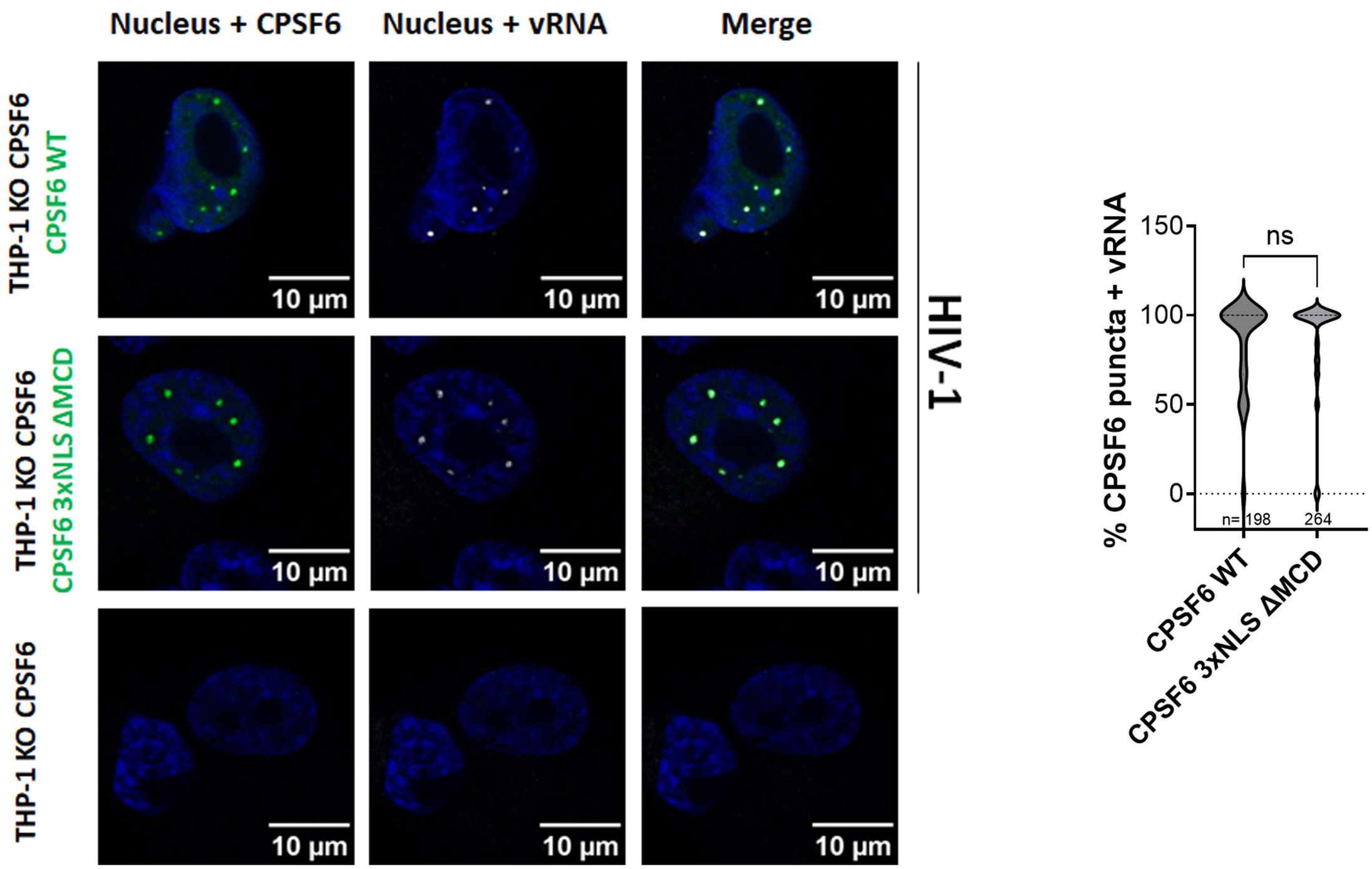

Nucleus + CPSF6
Nucleus + vRNA
Merge
THP-1 KO CPSF6
CPSF6 WT
THP-1 KO CPSF6
CPSF6 3xNLS ΔMCD
THP-1 KO CPSF6
10 µm
HIV-1
% CPSF6 puncta + vRNA
150
100
50
0
ns
n= 198
264
CPSF6 WT
CPSF6 3xNLS ΔMCD

Figure 2

H

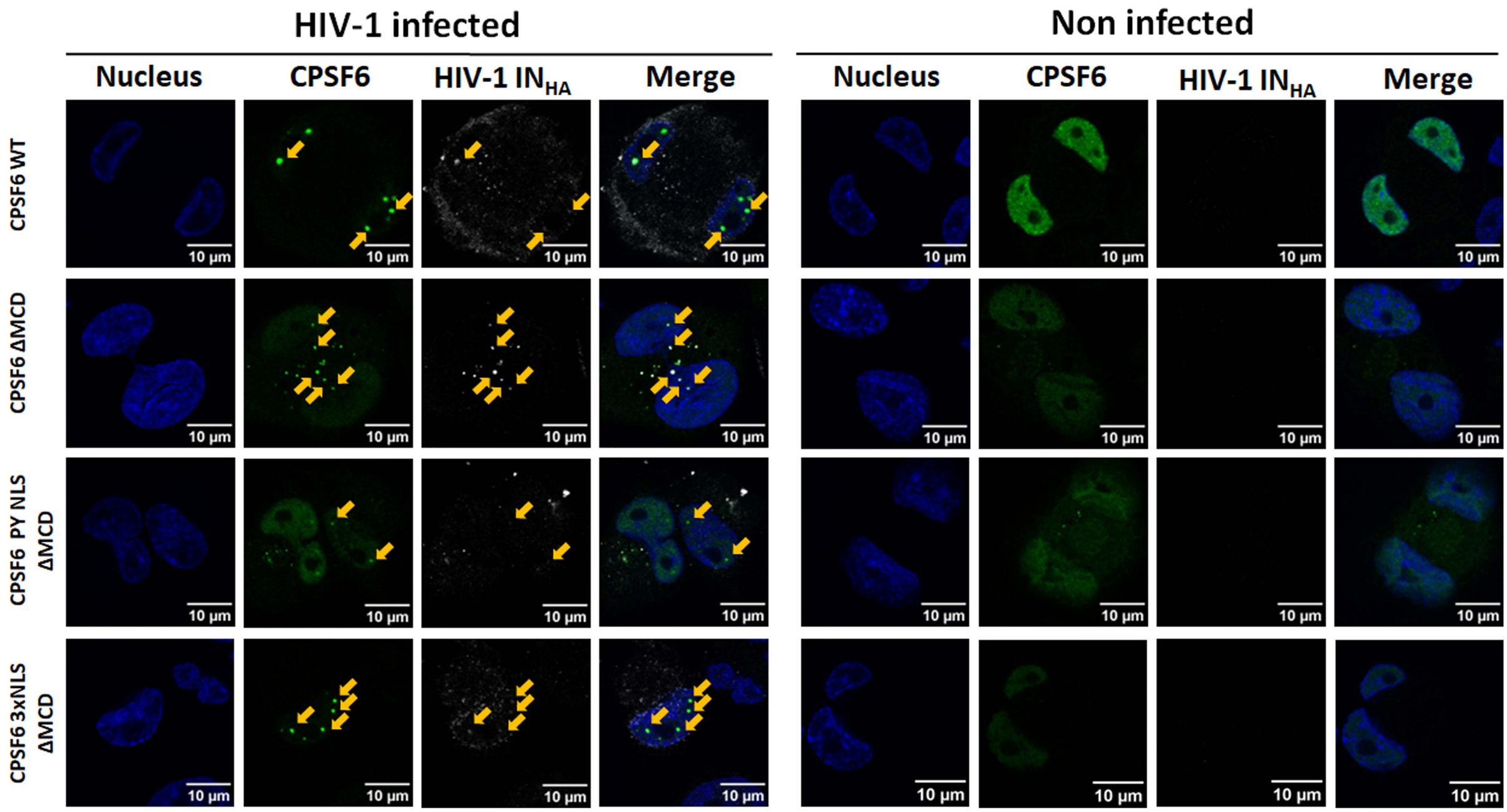

HIV-1 infected
Non infected
Nucleus
CPSF6
HIV-1 IN$_{HA}$
Merge
CPSF6 WT
CPSF6 ΔMCD
CPSF6 PY NLS ΔMCD
CPSF6 3xNLS ΔMCD
10 µm

Figure 2

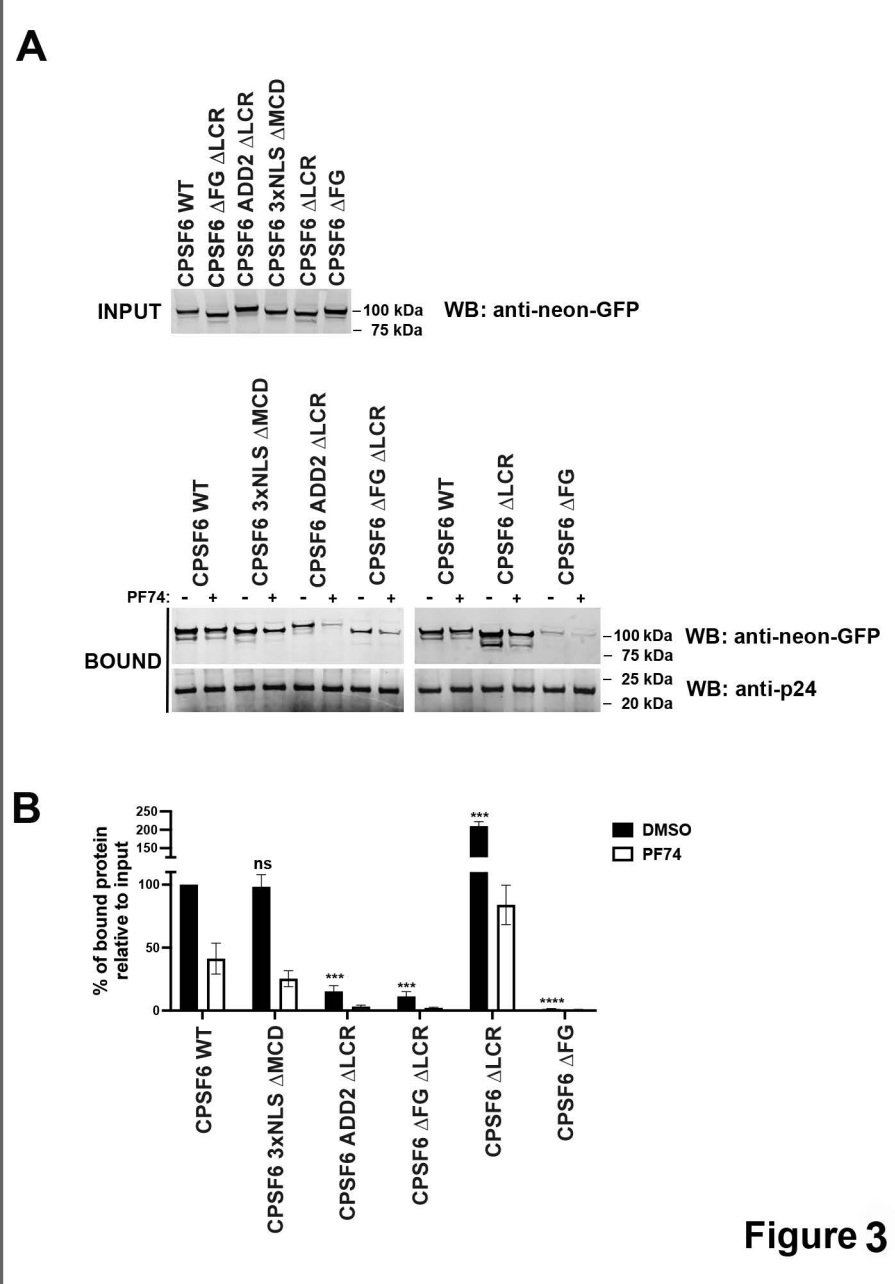

A
CPSF6 WT
CPSF6 ΔFG ΔLCR
CPSF6 ADD2 ΔLCR
CPSF6 3xNLS ΔMCD
CPSF6 ΔLCR
CPSF6 ΔFG
INPUT
100 kDa
75 kDa
WB: anti-neon-GFP
CPSF6 WT
CPSF6 3xNLS ΔMCD
CPSF6 ADD2 ΔLCR
CPSF6 ΔFG ΔLCR
CPSF6 WT
CPSF6 ΔLCR
CPSF6 ΔFG
PF74:
BOUND
100 kDa
75 kDa
WB: anti-neon-GFP
25 kDa
20 kDa
WB: anti-p24
B
% of bound protein relative to input
250
200
150
100
50
0
ns
***
***
***
****
DMSO
PF74
CPSF6 WT
CPSF6 3xNLS ΔMCD
CPSF6 ADD2 ΔLCR
CPSF6 ΔFG ΔLCR
CPSF6 ΔLCR
CPSF6 ΔFG

Figure 3

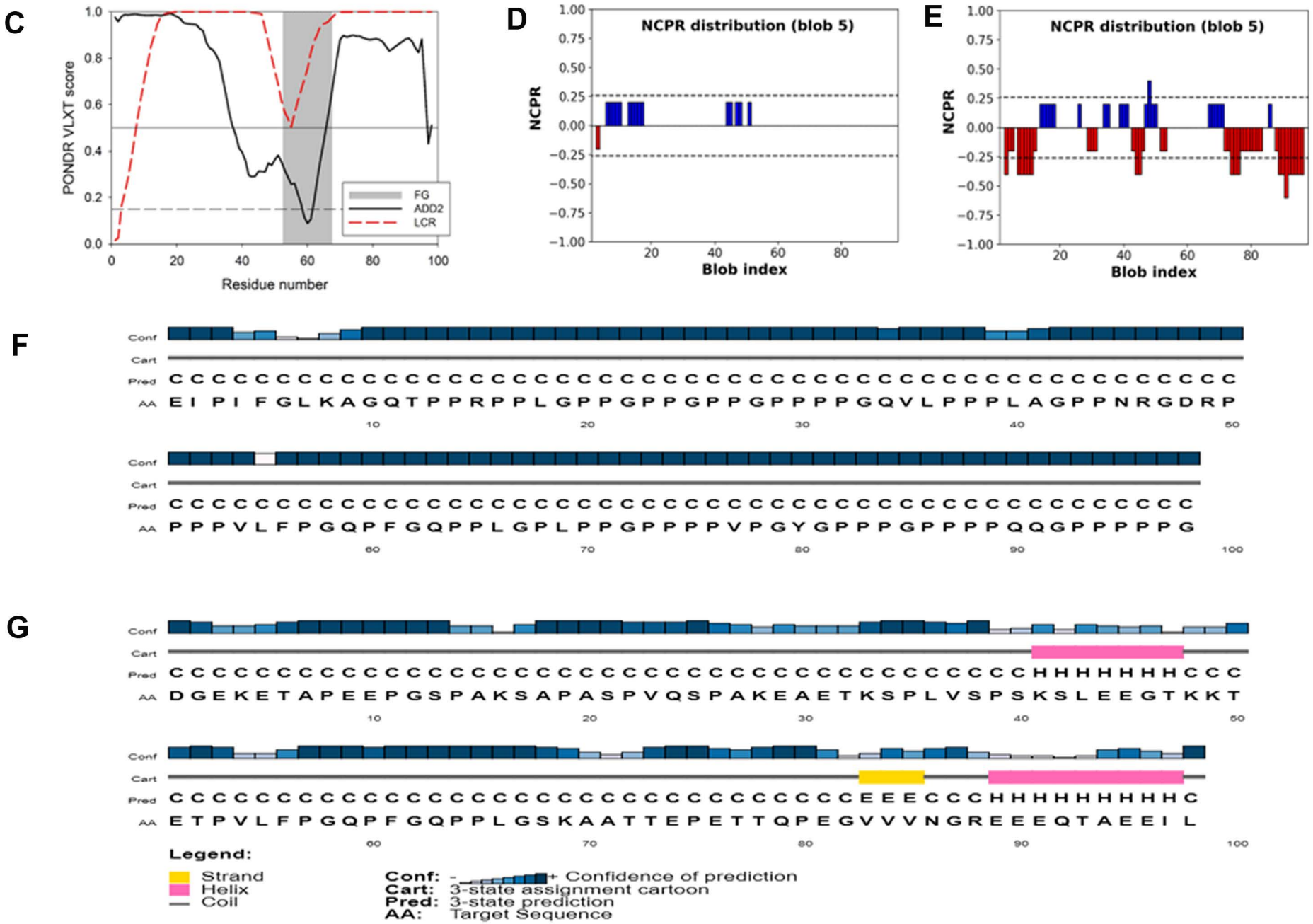
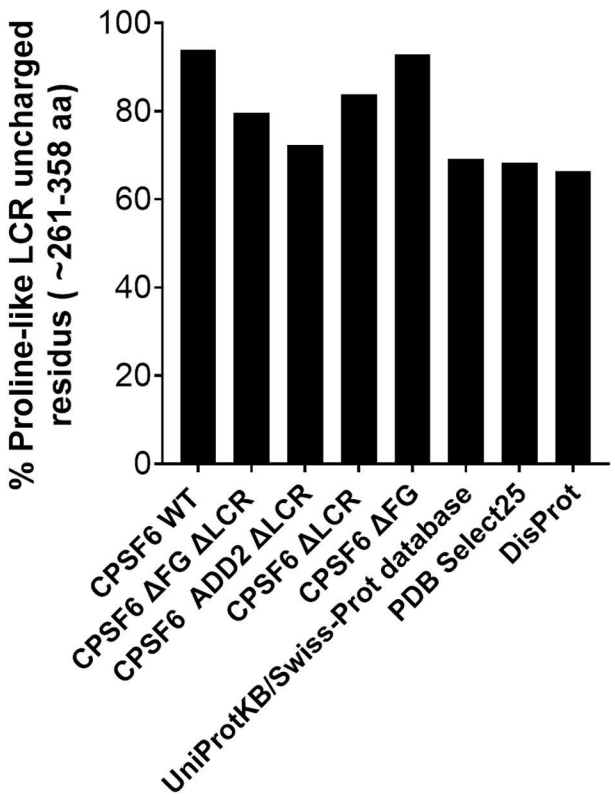
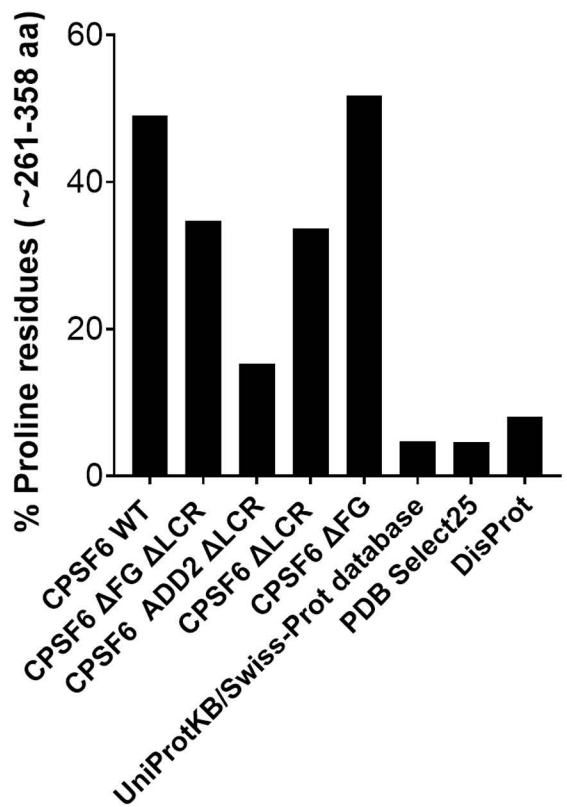

C
PONDR VLXT score
Residue number
FG
ADD2
LCR
D
NCPR distribution (blob 5)
NCPR
Blob index
E
NCPR distribution (blob 5)
NCPR
Blob index
F
Conf
Cart
Pred
AA
G
Conf
Cart
Pred
AA
Legend:
Strand
Helix
Coil
Conf: Confidence of prediction
Cart: 3-state assignment cartoon
Pred: 3-state prediction
AA: Target Sequence

Figure 3

H


% Proline-like LCR uncharged residus (~261-358 aa)
100
80
60
40
20
0
CPSF6 WT
CPSF6 ΔFG ΔLCR
CPSF6 ADD2 ΔLCR
CPSF6 ΔLCR
CPSF6 ΔFG
UniProtKB/Swiss-Prot database
PDB Select25
DisProt

I


% Proline residues (~261-358 aa)
60
40
20
0
CPSF6 WT
CPSF6 ΔFG ΔLCR
CPSF6 ADD2 ΔLCR
CPSF6 ΔLCR
CPSF6 ΔFG
UniProtKB/Swiss-Prot database
PDB Select25
DisProt

L

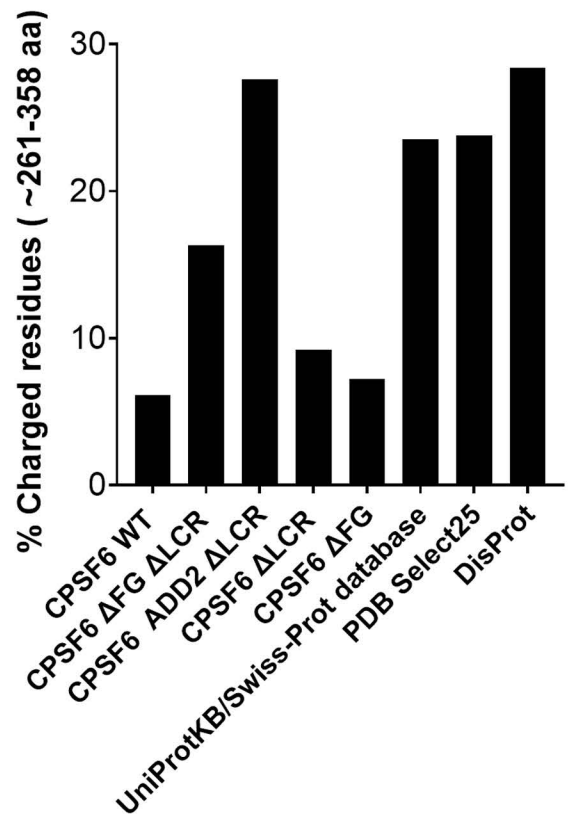
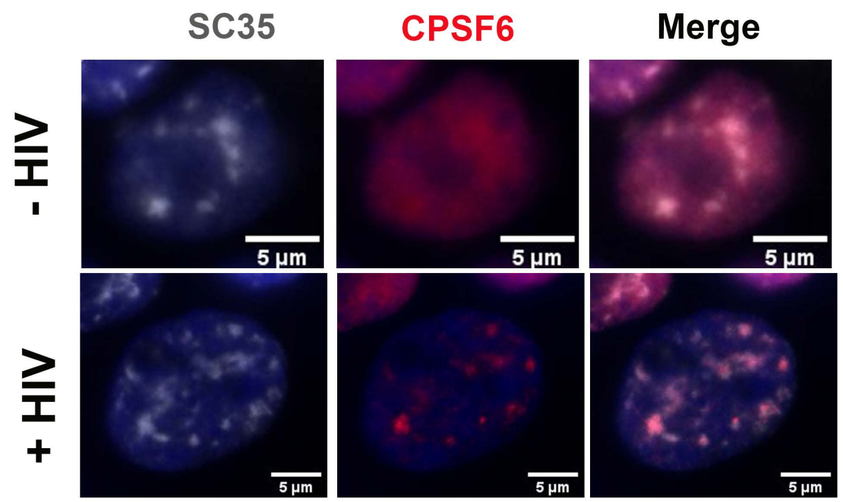
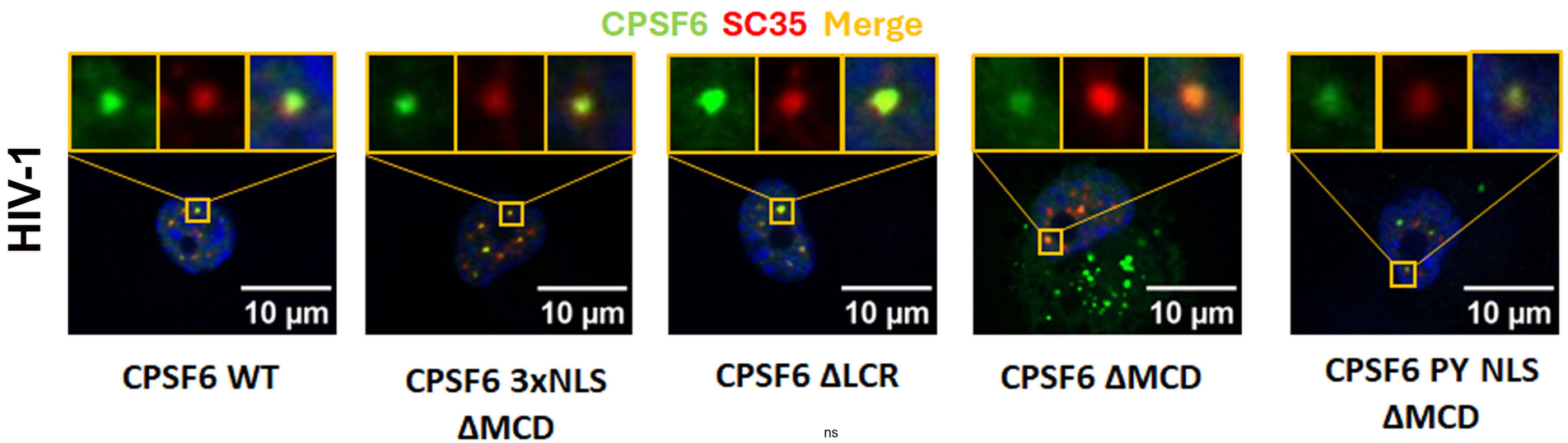

% Charged residues (~261-358 aa)
30
20
10
0
CPSF6 WT
CPSF6 ΔFG ΔLCR
CPSF6 ADD2 ΔLCR
CPSF6 ΔLCR
CPSF6 ΔFG
UniProtKB/Swiss-Prot database
PDB Select25
DisProt

Figure 3

A


SC35
CPSF6
Merge
- HIV
+ HIV
5 µm

B


CPSF6 SC35 Merge
HIV-1
10 µm
CPSF6 WT
CPSF6 3xNLS ΔMCD
CPSF6 ΔLCR
CPSF6 ΔMCD
CPSF6 PY NLS ΔMCD

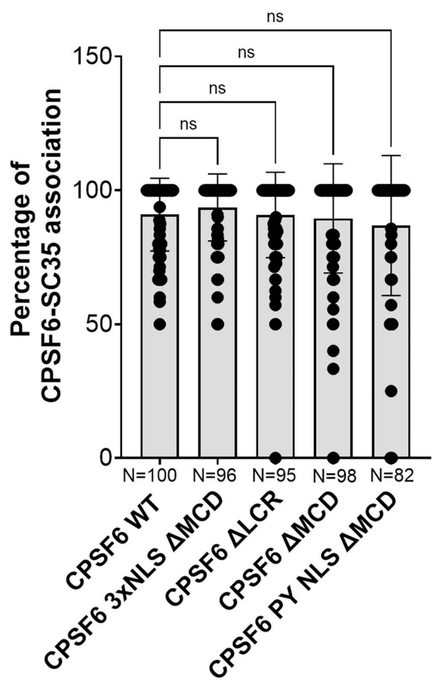

Percentage of CPSF6-SC35 association
150
100
50
0
ns
N=100
N=96
N=95
N=98
N=82
CPSF6 WT
CPSF6 3xNLS ΔMCD
CPSF6 ΔLCR
CPSF6 ΔMCD
CPSF6 PY NLS ΔMCD

Figure 4

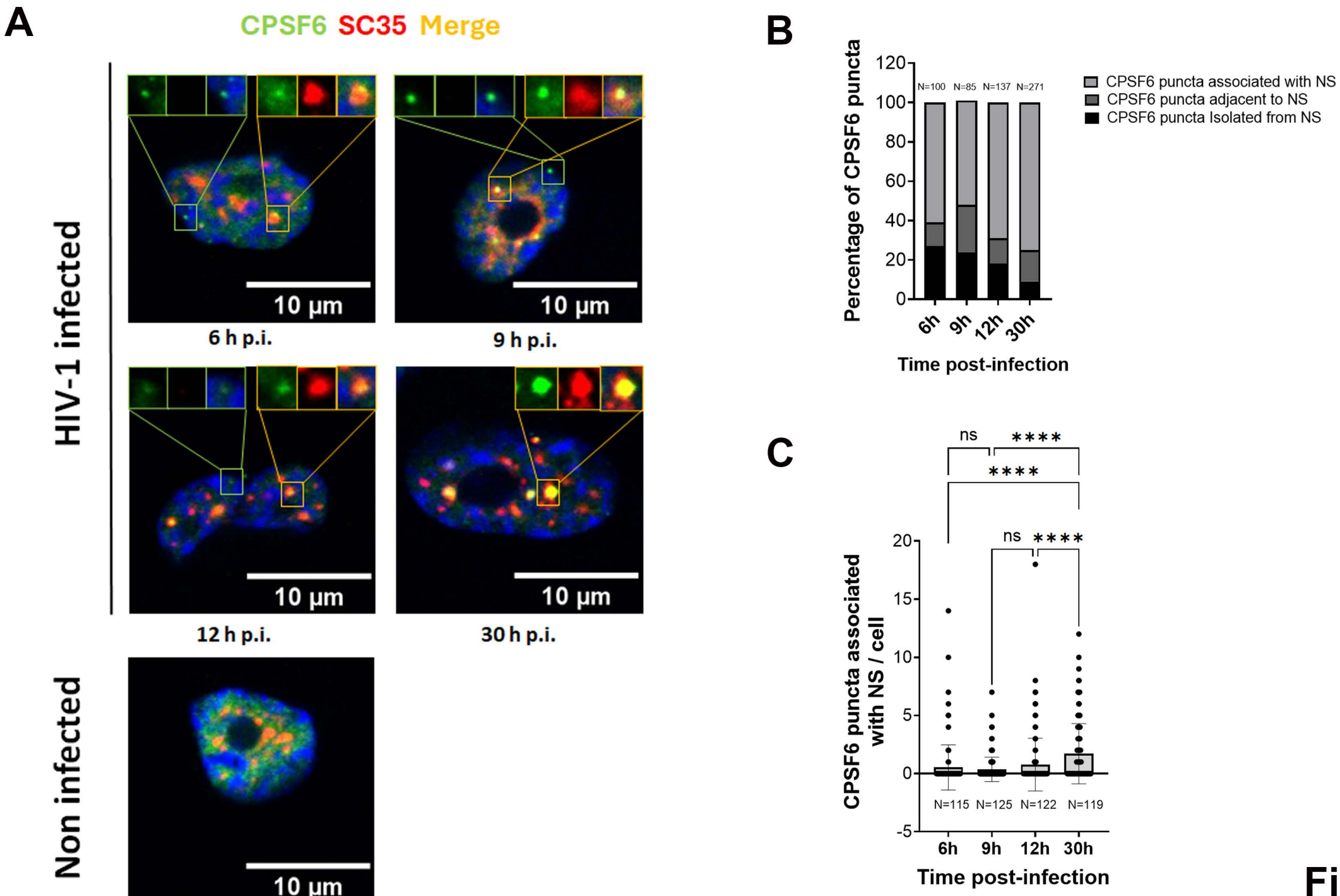

A
CPSF6 SC35 Merge
HIV-1 infected
Non infected
10 µm
6 h p.i.
9 h p.i.
12 h p.i.
30 h p.i.
B
Percentage of CPSF6 puncta
N=100 N=85 N=137 N=271
CPSF6 puncta associated with NS
CPSF6 puncta adjacent to NS
CPSF6 puncta Isolated from NS
6h 9h 12h 30h
Time post-infection
C
ns
****
CPSF6 puncta associated with NS / cell
N=115 N=125 N=122 N=119
6h 9h 12h 30h
Time post-infection

Figure 5

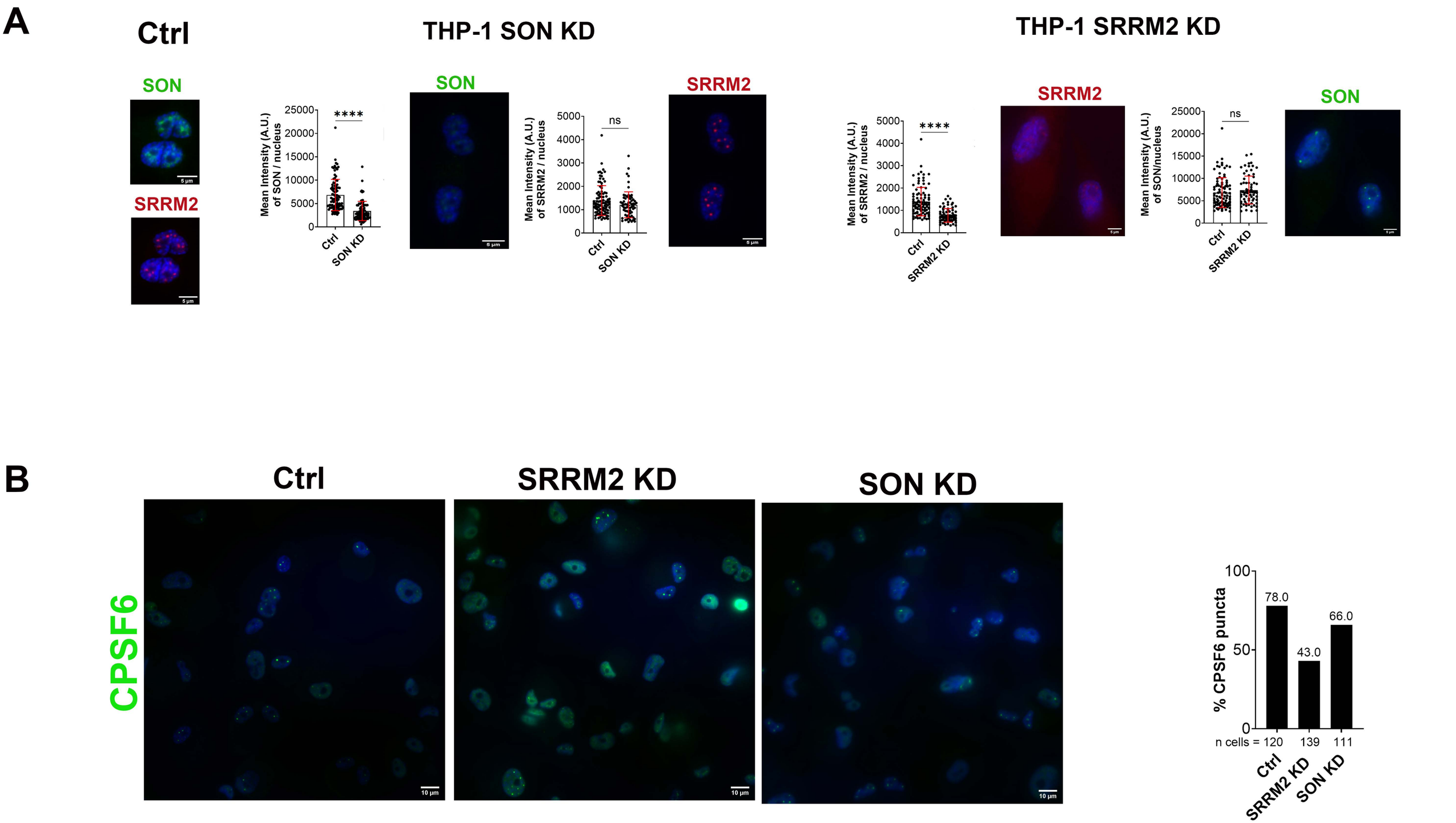
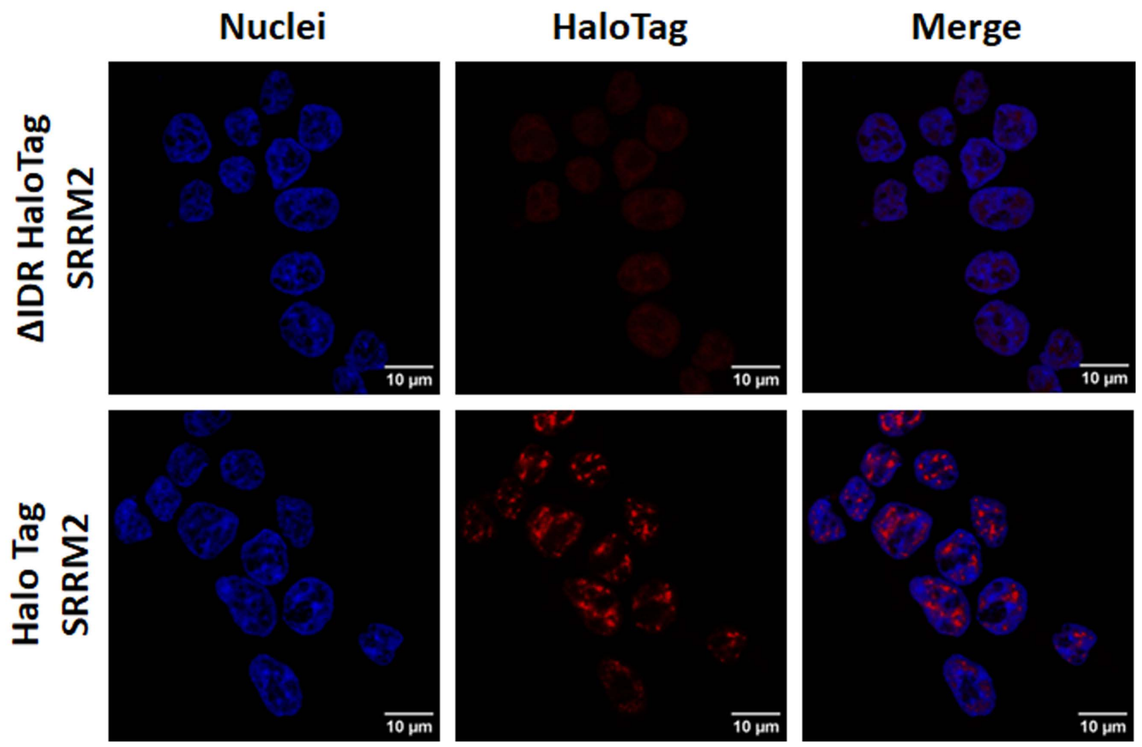
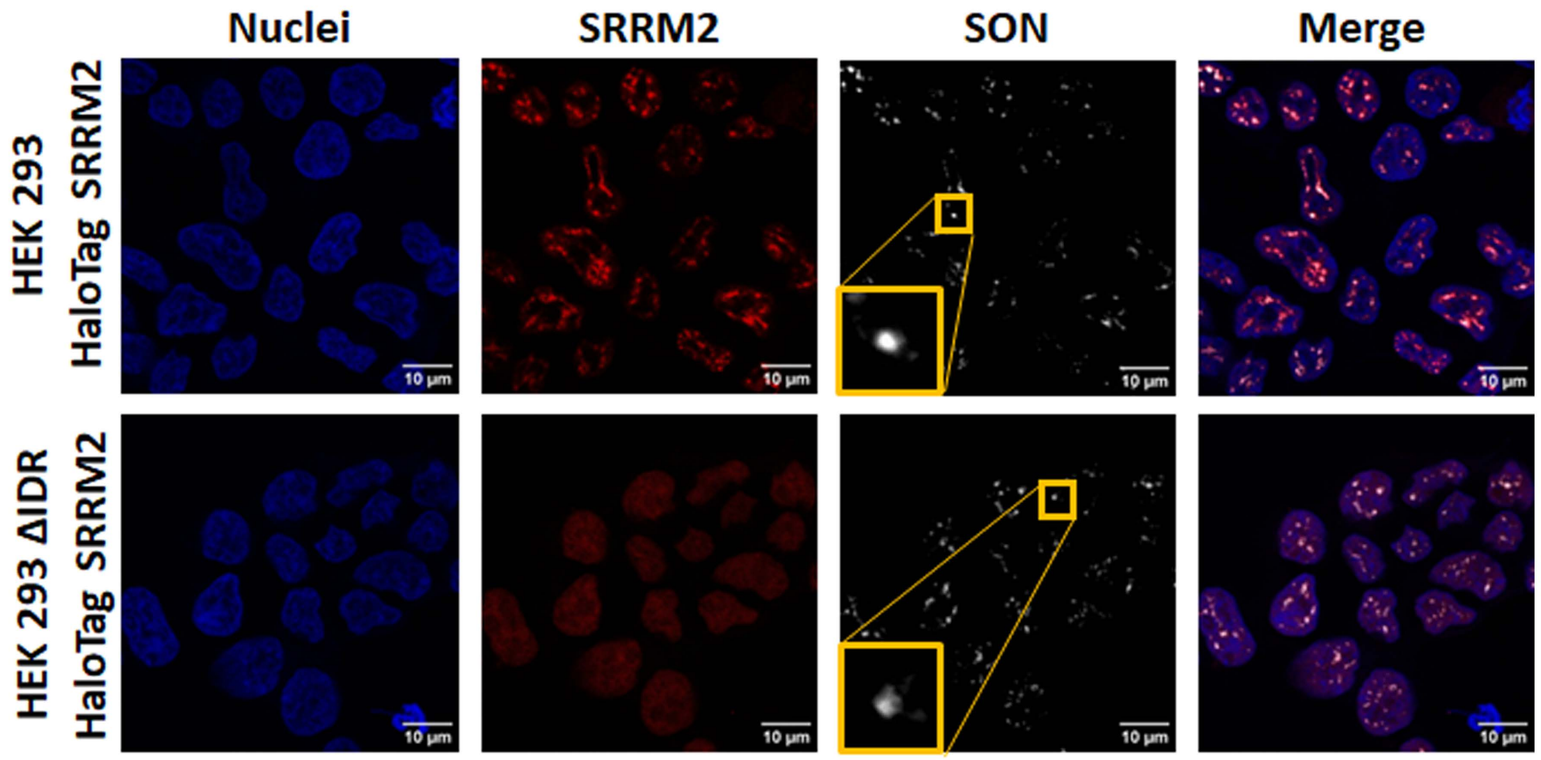

A
Ctrl
THP-1 SON KD
THP-1 SRRM2 KD
SON
SRRM2
Mean Intensity (A.U.) of SON / nucleus
****
Ctrl
SON KD
Mean Intensity (A.U.) of SRRM2 / nucleus
ns
Mean Intensity (A.U.) of SRRM2 / nucleus
****
Ctrl
SRRM2 KD
Mean Intensity (A.U.) of SON/nucleus
ns
B
Ctrl
SRRM2 KD
SON KD
CPSF6
% CPSF6 puncta
78.0
43.0
66.0
n cells = 120
139
111
Ctrl
SRRM2 KD
SON KD

Figure 6

C


Nuclei
HaloTag
Merge
ΔIDR HaloTag SRRM2
Halo Tag SRRM2
10 µm

D


Nuclei
SRRM2
SON
Merge
HEK 293 HaloTag SRRM2
HEK 293 ΔIDR HaloTag SRRM2
10 µm

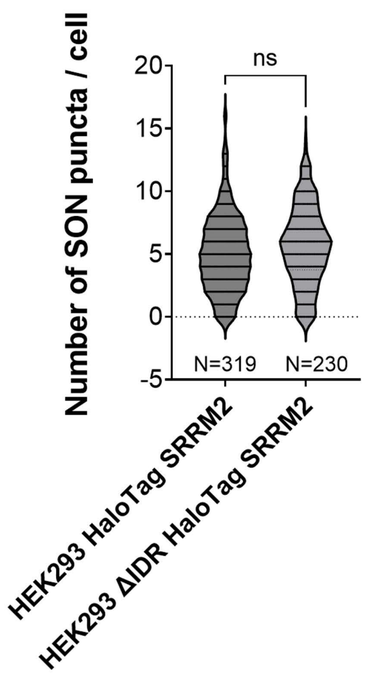
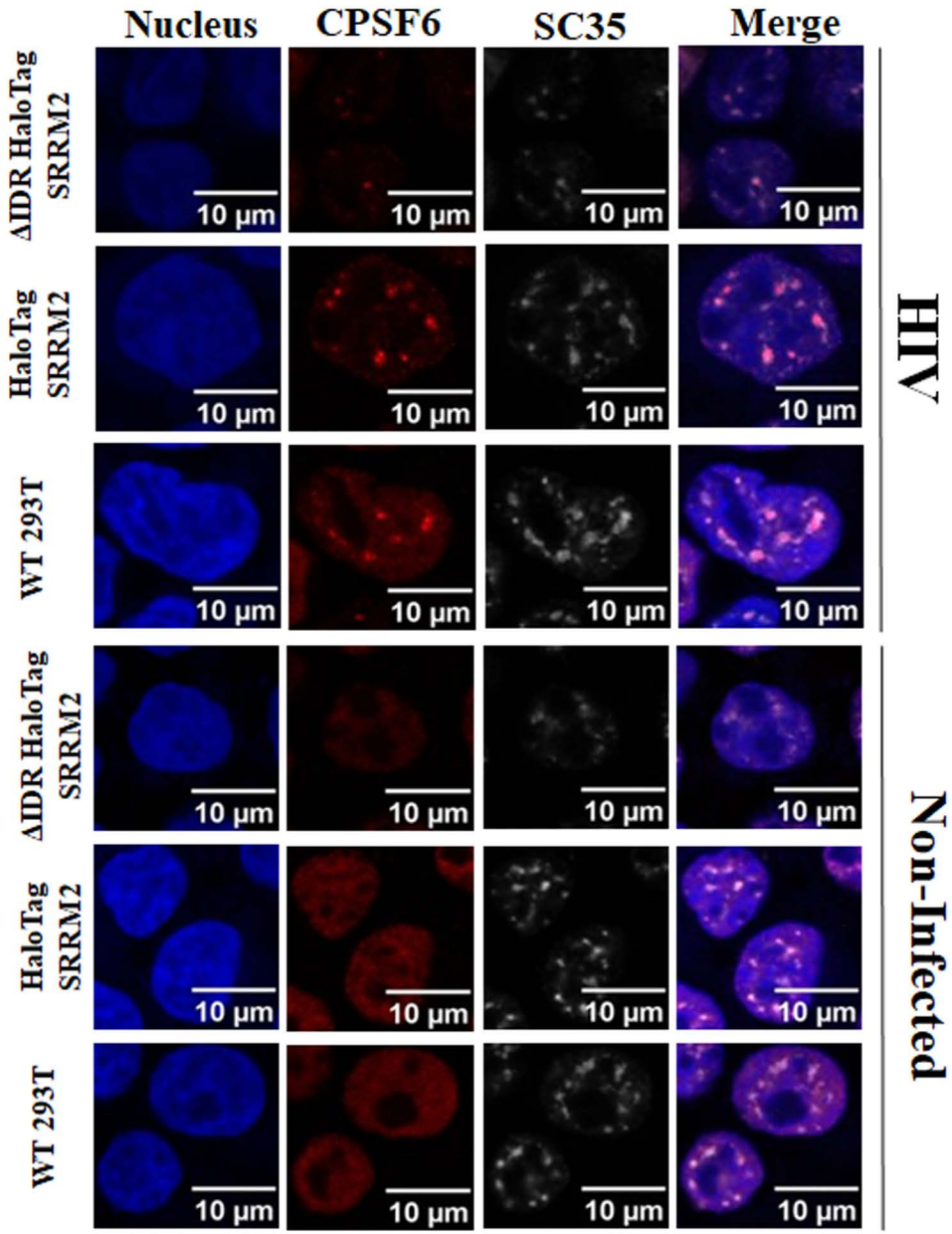

Number of SON puncta / cell
20
15
10
5
0
-5
ns
N=319
N=230
HEK293 HaloTag SRRM2
HEK293 ΔIDR HaloTag SRRM2

Figure 6

E


Nucleus
CPSF6
SC35
Merge
ΔIDR HaloTag SRRM2
HaloTag SRRM2
WT 293T
ΔIDR HaloTag SRRM2
HaloTag SRRM2
WT 293T
10 µm
HIV
Non-Infected

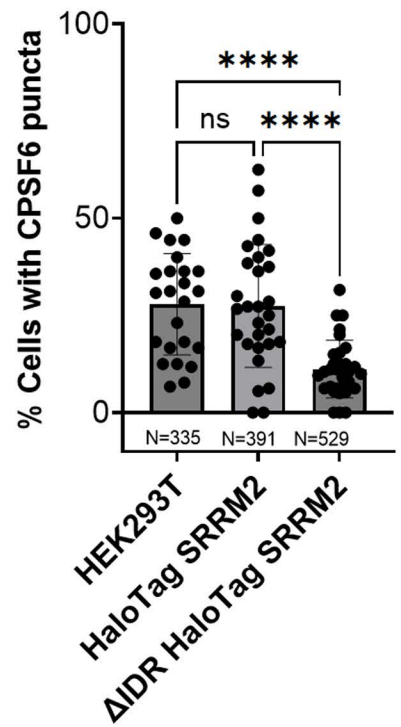
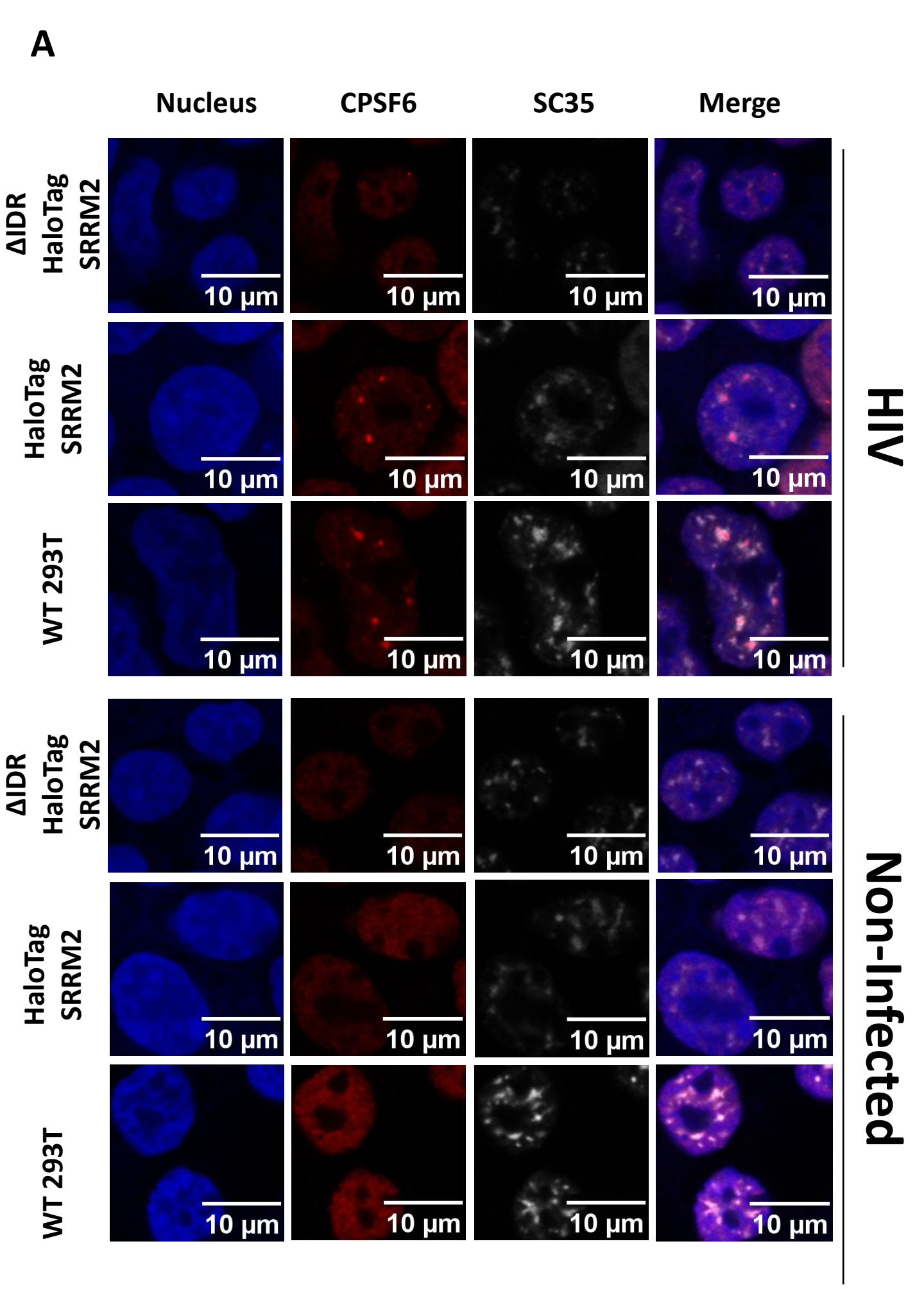
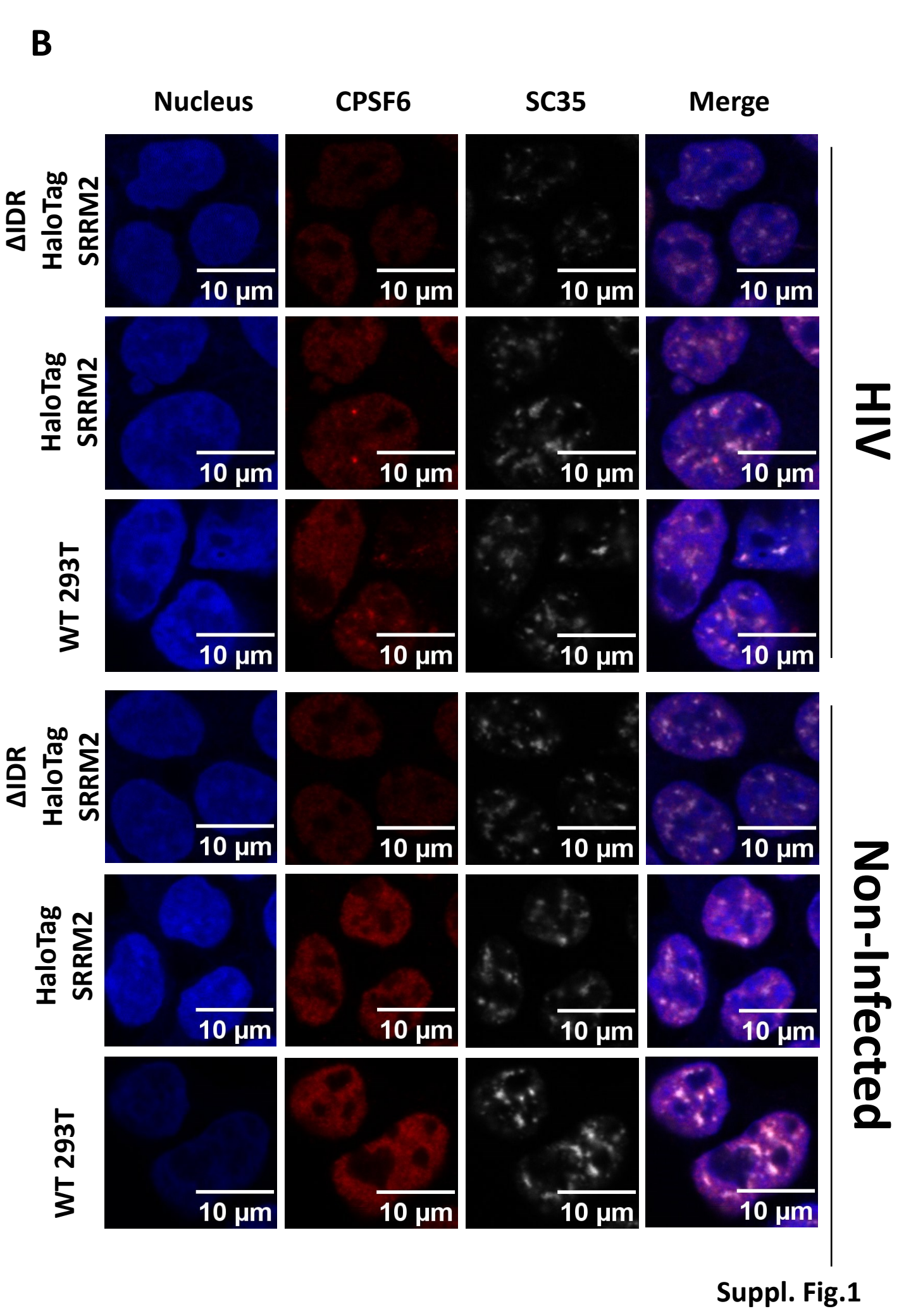

% Cells with CPSF6 puncta
100
50
0
****
ns
****
N=335
N=391
N=529
HEK293T
HaloTag SRRM2
ΔIDR HaloTag SRRM2

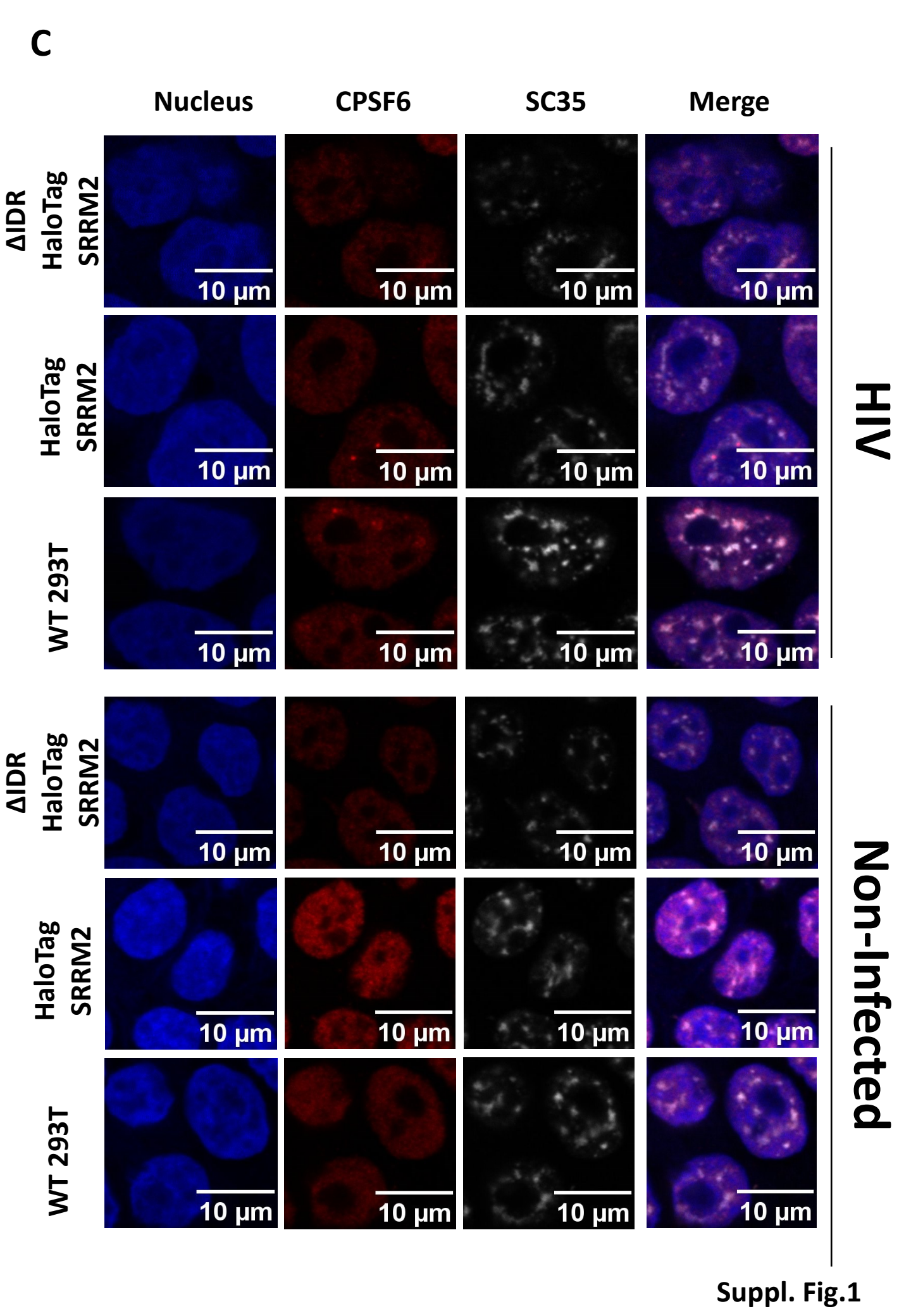
Figure 6


A
Nucleus
CPSF6
SC35
Merge
ΔIDR HaloTag SRRM2
HaloTag SRRM2
WT 293T
10 µm
HIV
Non-Infected

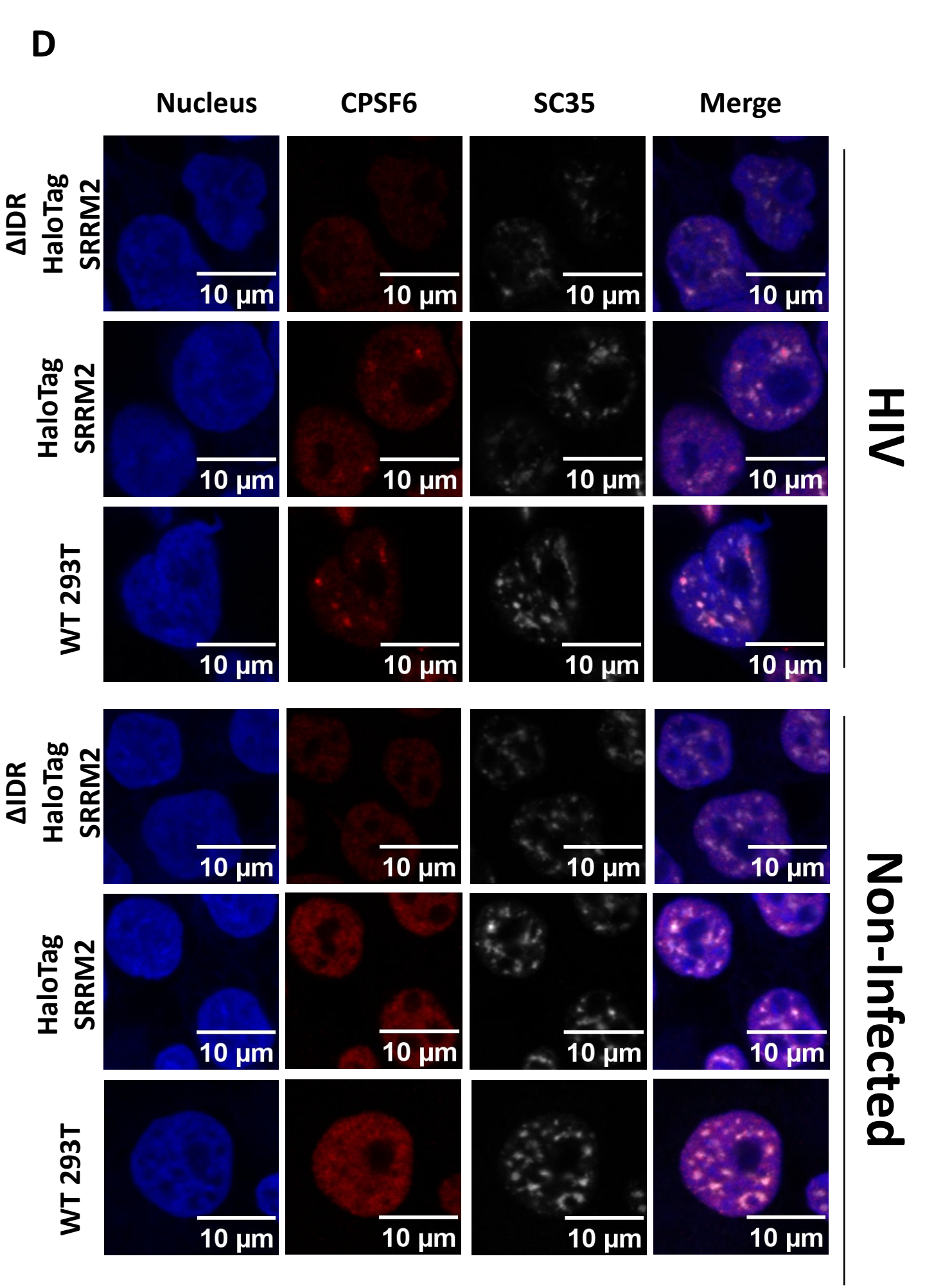
Suppl. Fig.1


B
Nucleus
CPSF6
SC35
Merge
ΔIDR HaloTag SRRM2
HaloTag SRRM2
WT 293T
10 µm
HIV
Non-Infected

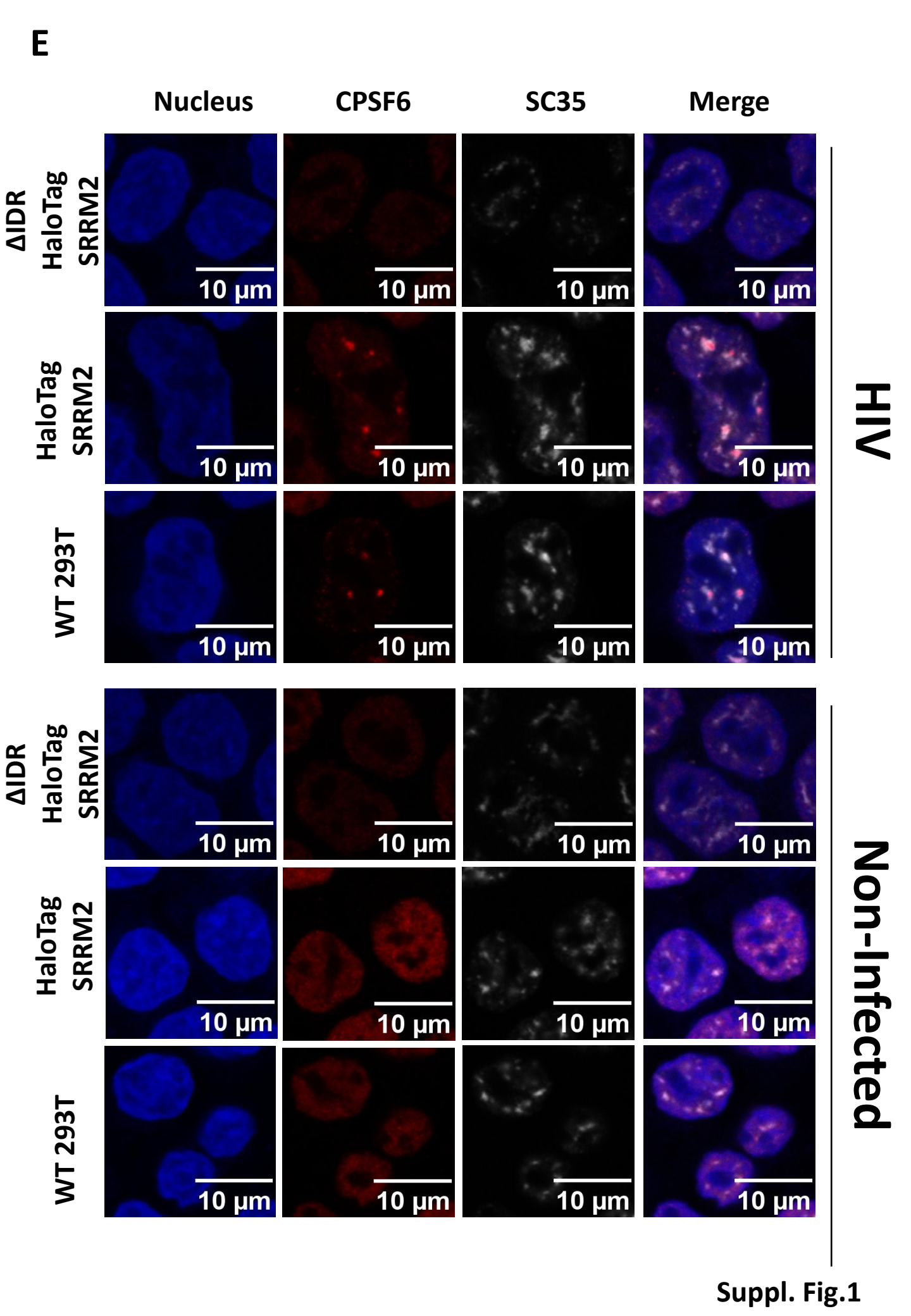
Suppl. Fig.1


C
Nucleus
CPSF6
SC35
Merge
ΔIDR HaloTag SRRM2
HaloTag SRRM2
WT 293T
10 µm
HIV
Non-Infected

Suppl. Fig.1

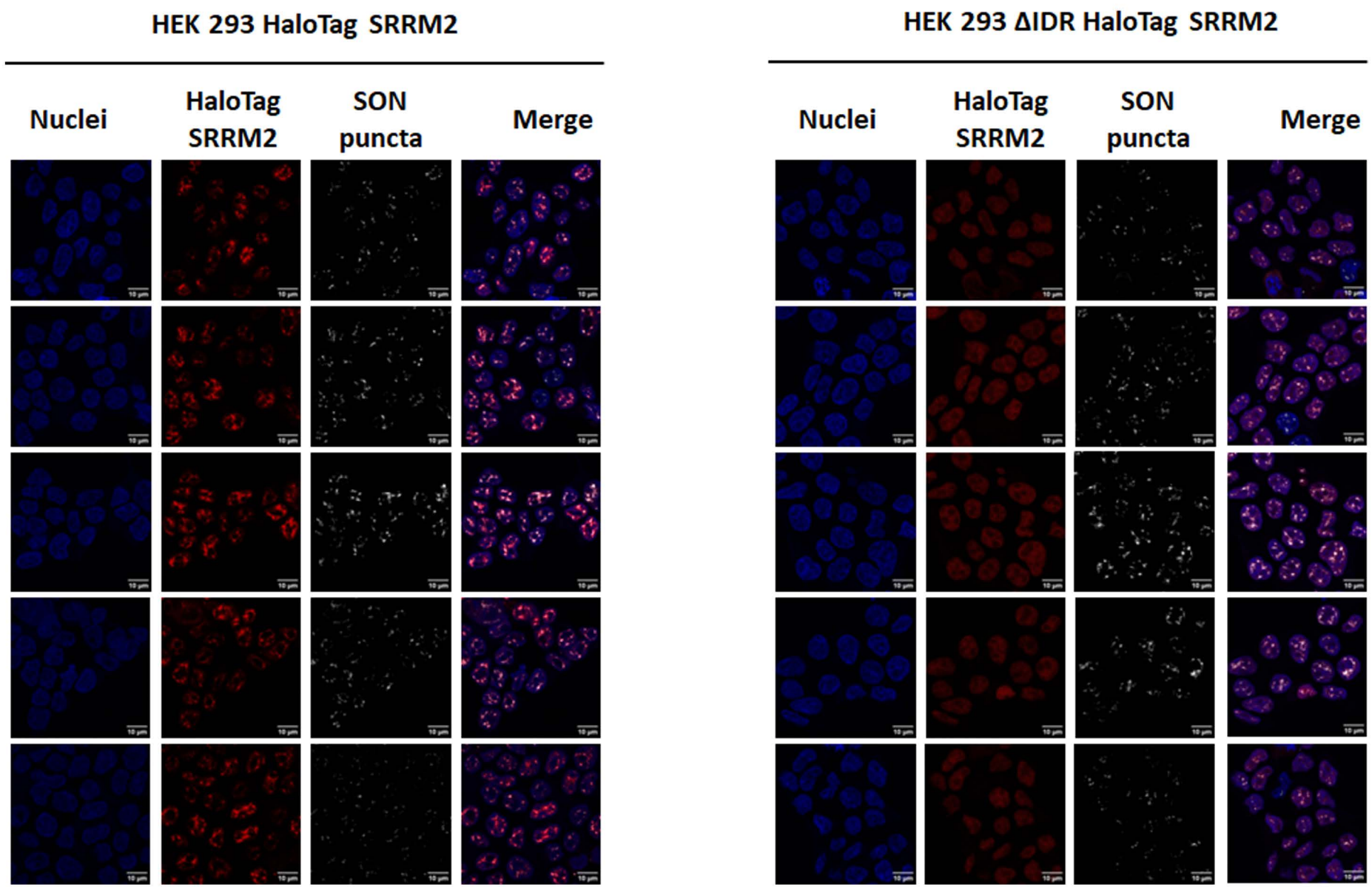

D
Nucleus
CPSF6
SC35
Merge
ΔIDR HaloTag SRRM2
HaloTag SRRM2
WT 293T
10 µm
HIV
Non-Infected

Suppl. Fig.1


E
Nucleus
CPSF6
SC35
Merge
ΔIDR HaloTag SRRM2
HaloTag SRRM2
WT 293T
10 µm
HIV
Non-Infected

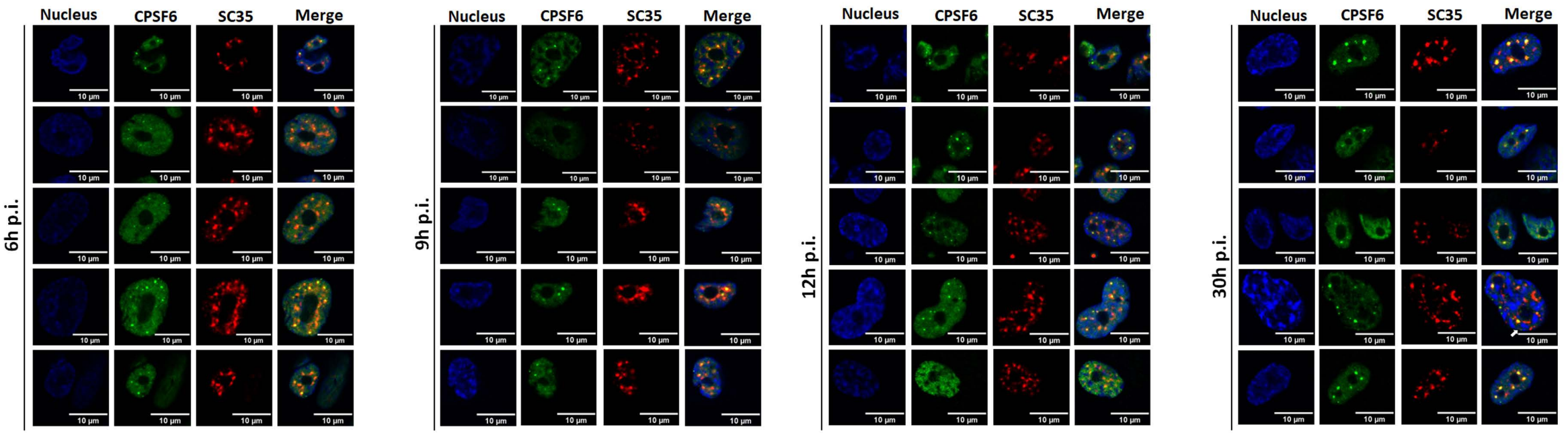
Suppl. Fig.1

F



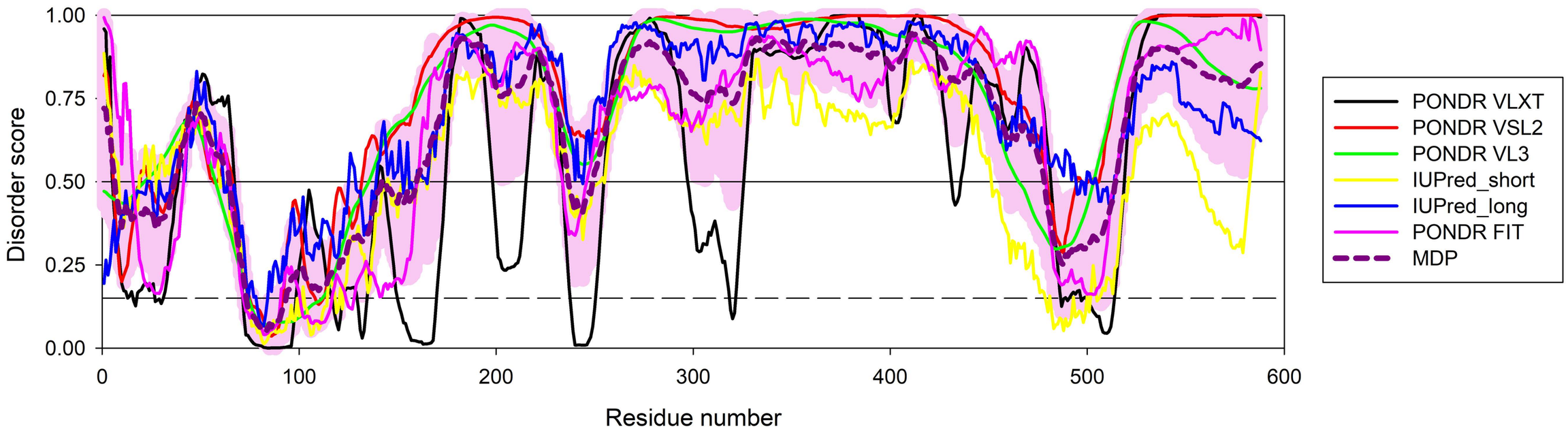
Suppl. Figure 1

G



Suppl. Figure 1

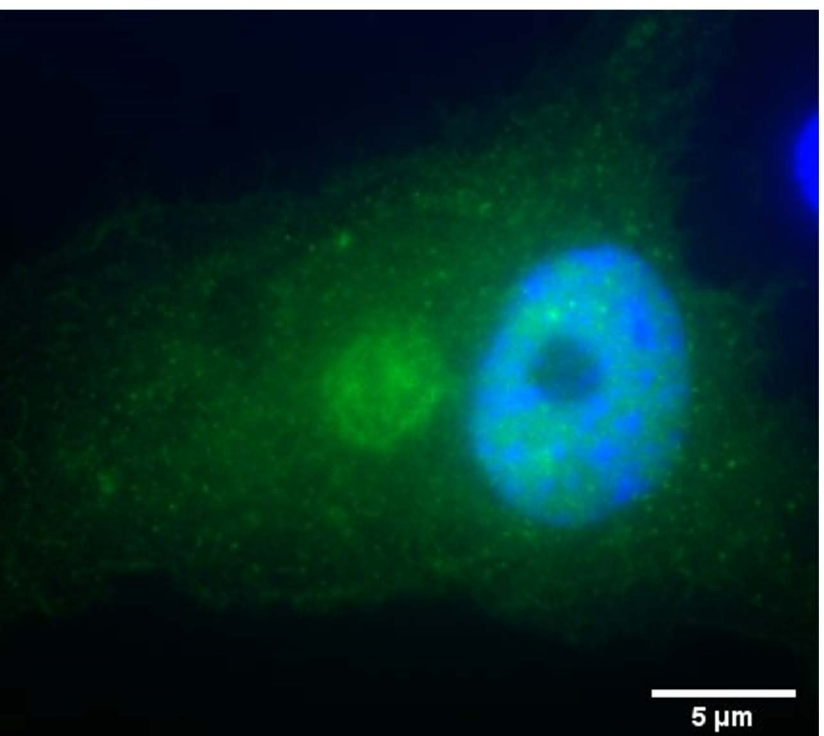
A



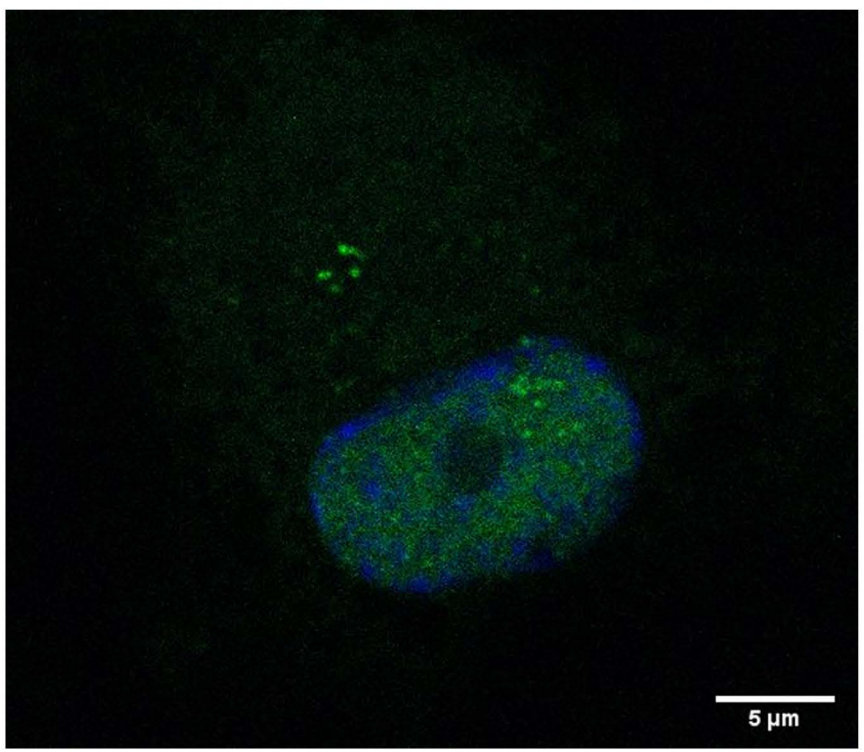
Suppl. Figure 2

B

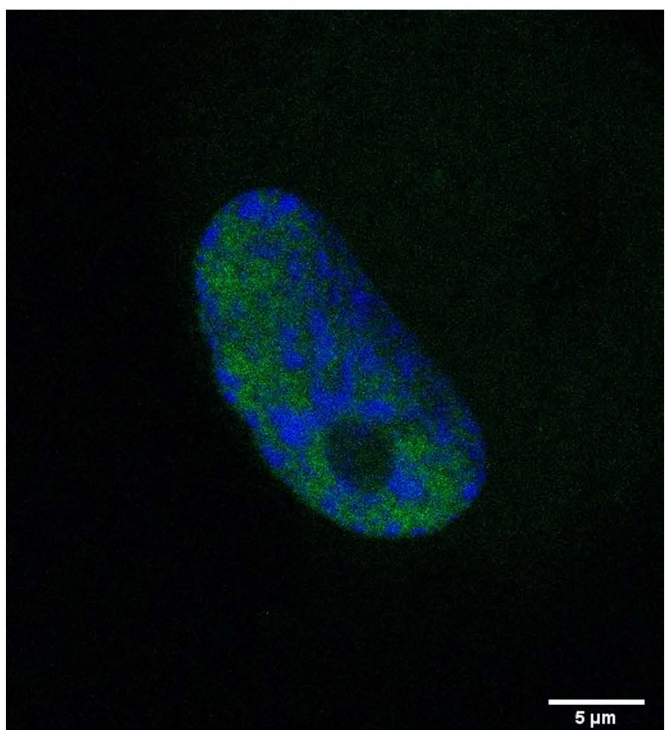
CPSF6 ΔMCDmNeonGreen



CPSF6 NLSΔMCDmNeonGreen



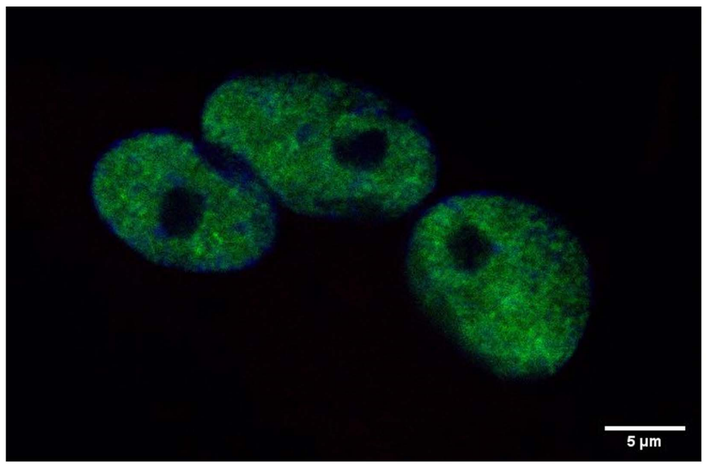
CPSF6 3xNLSΔMCDmNeonGreen



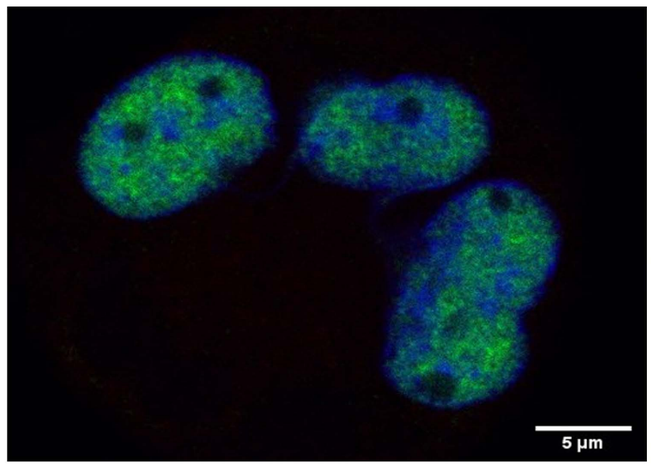
Suppl. Figure 2

C

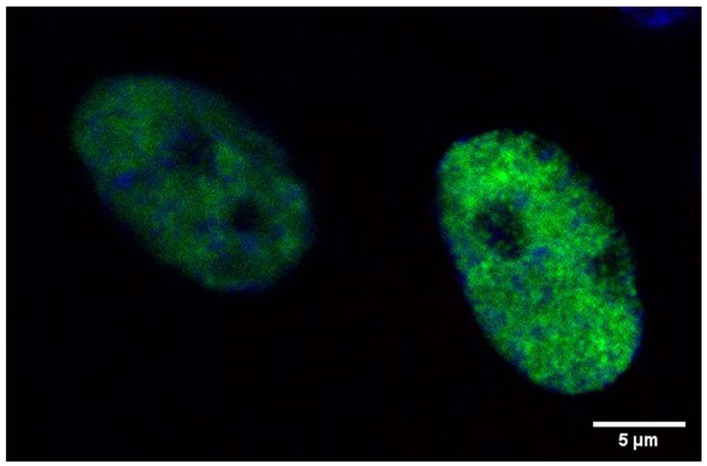
CPSF6



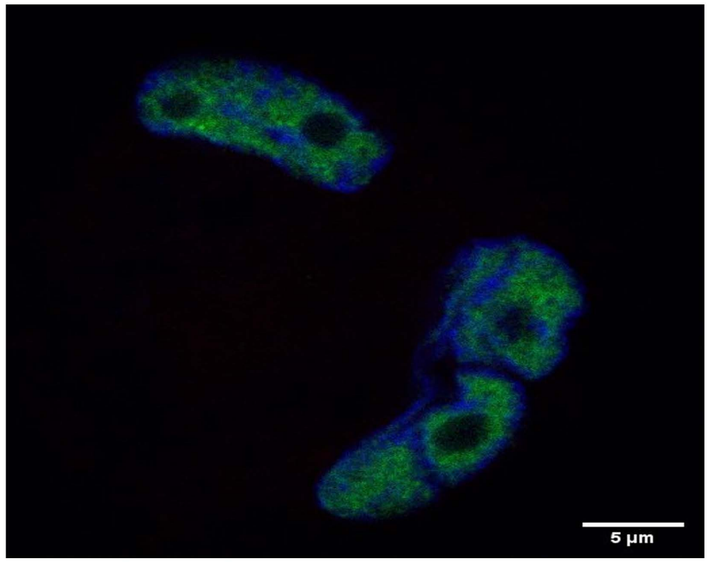
CPSF6 ΔLCRΔFG



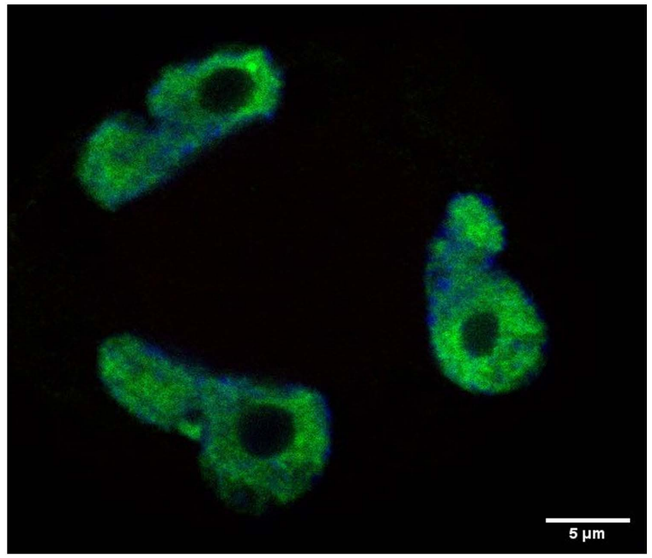
CPSF6 ΔFG



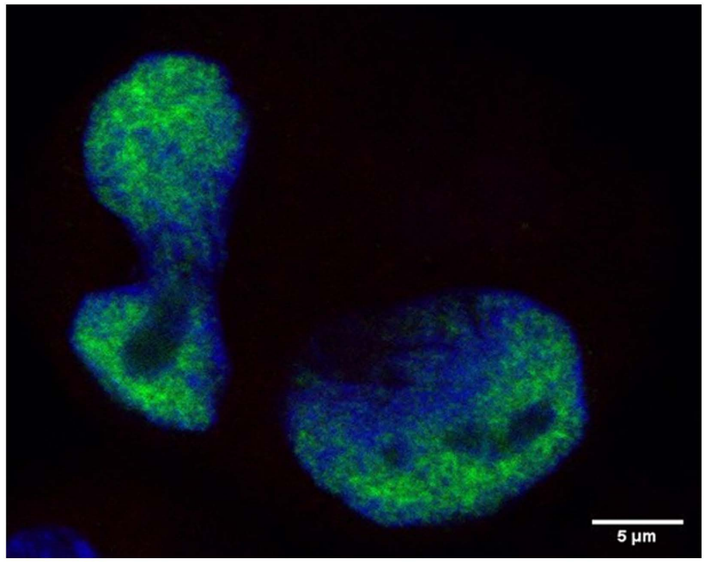
CPSF6 3xNLSΔMCD



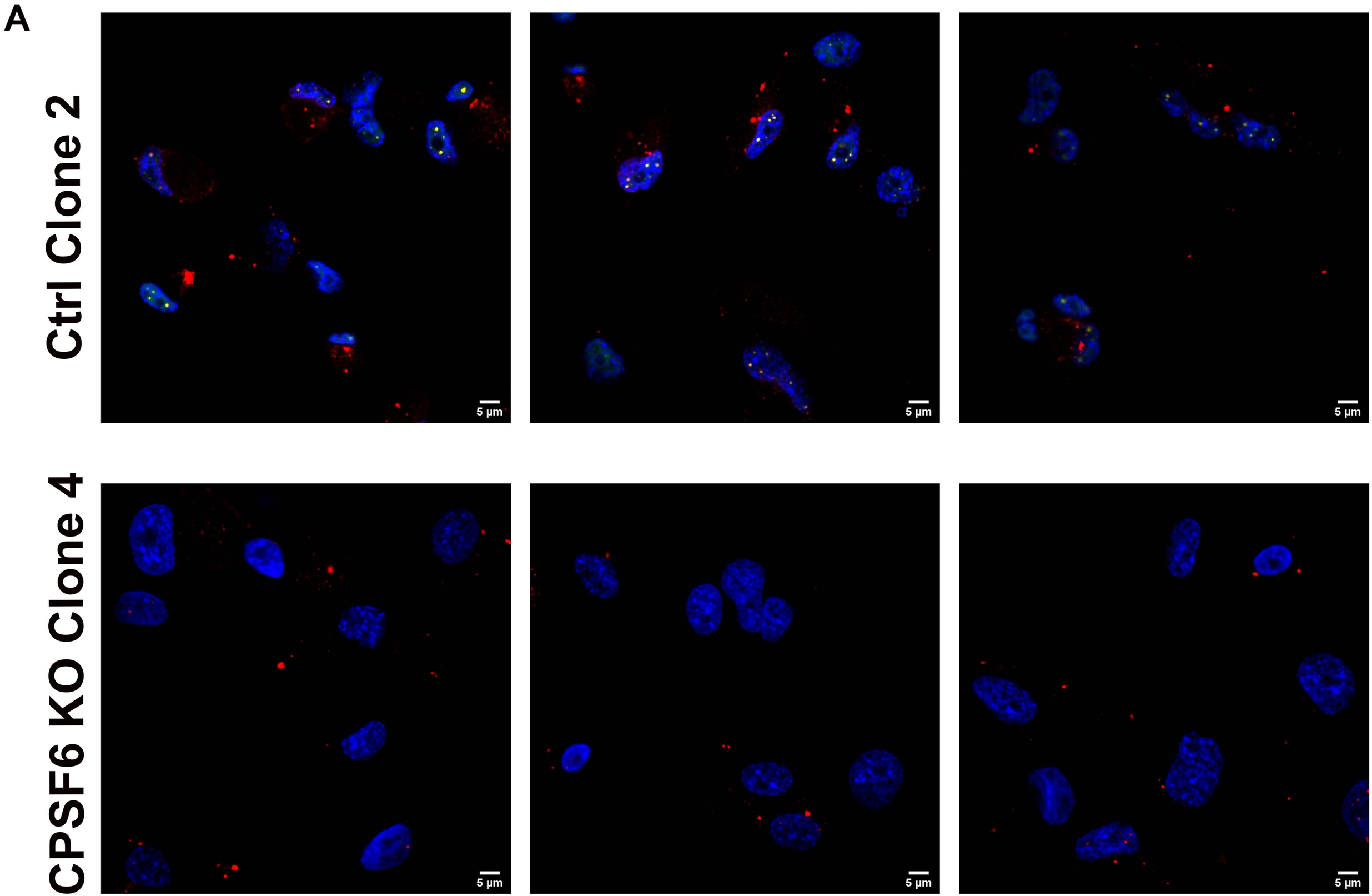
CPSF6 ΔLCR



CPSF6 ADD2 ΔLCR

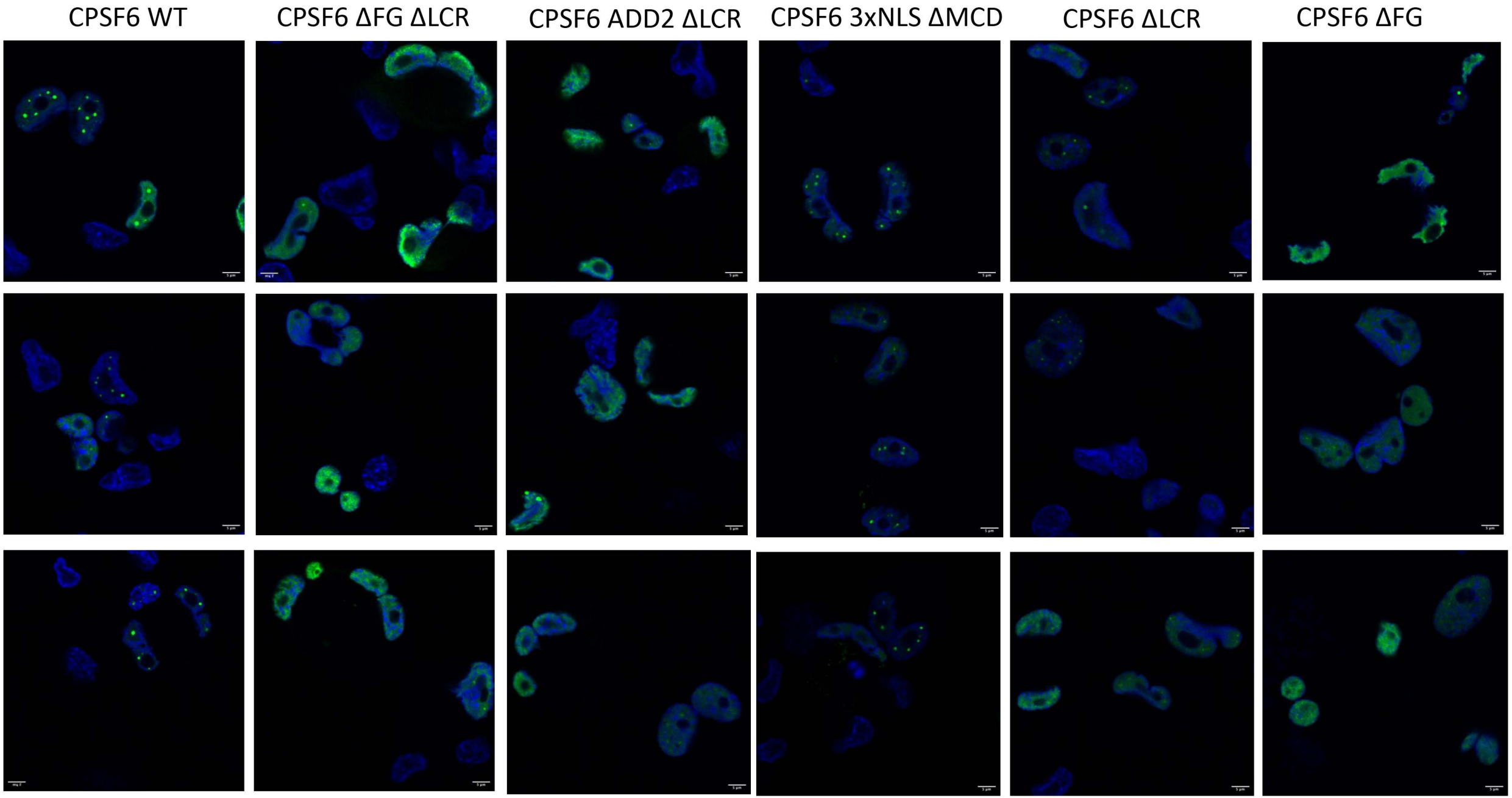


Suppl. Figure 2



Suppl. Figure 3

B


CPSF6 WT
CPSF6 ΔFG ΔLCR
CPSF6 ΔD2 ΔLCR
CPSF6 3xNLS ΔMCD
CPSF6 ΔLCR
CPSF6 ΔFG

Suppl. Fig. 3

Bioinformatic comparison of amino acids sequence ~261-358aa:

CPSF6 domain ΔLRC + FG peptide:
gavpggdrfpgpagpggpppfpgnlikhlvkgtrplfletripwhmghsiePVLFPGQPFGQPPLGpfpprppgplgppltlappphlp
gpppgapp

CPSF6 **Δ**LCRs + ADD2 + FG peptide:
DGEKETAPEEPGSPAKSAPASPVQSPAKEAETKSPLVSPSKSLEEGTKKTETPVLFPGQPFGQPPLGSKAATTEPETTQ
PEGVVVNGREEEQTAEEIL

CPSF6 WT (domain LCRs + FG peptide):

Eipifglkagqtpprpplgppgppgppgpppgqvlppplagppnrgdrpppvlfpgqpfgqpplgplppgpppvpgygpppgpppp
qqgpppppg

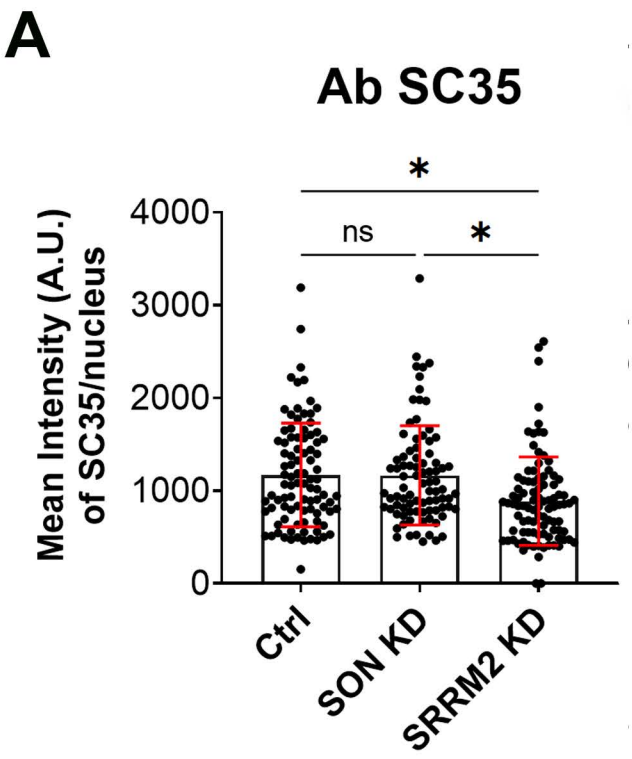
CPSF6 (domain LCRs **Δ** FG peptide):

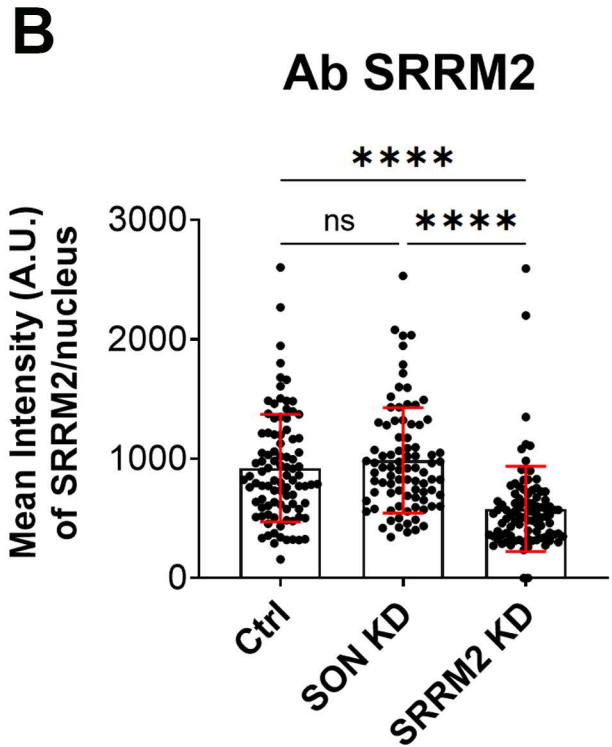
Eipifglkagqtpprpplgppgppgppgpppgqvlppplagppnrgdrpppplppgpppvpgygpppgpppqqgpppppg

CPSF6 (**Δ** domain LCRs **Δ** FG peptide):

Pfpprppgplgppltlappphlpgpppgapppaphvnpaffppptnsgmptsdsrgppptdpygrpppydrgdygppgremdtartpls
eaefeeimn

Suppl. Figure 4


A
Ab SC35
*
ns
*
4000
3000
2000
1000
0
Mean Intensity (A.U.)
of SC35/nucleus
Ctrl
SON KD
SRRM2 KD


B
Ab SRRM2
****
ns
****
3000
2000
1000
0
Mean Intensity (A.U.)
of SRRM2/nucleus
Ctrl
SON KD
SRRM2 KD

Suppl. Fig.5

Appendix 1—key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rat monoclonal anti-HA	Roche	Cat#11867423001
Mouse monoclonal anti-p24	NIH reagent	Cat#NIH183-H12-5C
Rabbit monoclonal anti-CPSF6	Novus Biologicals	Cat#NBP1- 85676
Mouse monoclonal anti-SC35	Abcam	Cat#ab11826
Rabbit monoclonal anti-SON (human)	Sigma-Aldrich	Cat#HPA031755
Rabbit monoclonal anti-SRRM2	Sigma-Aldrich	Cat#HPA066181
Donkey polyclonal anti-mouse Cy3	Jackson Lab	Cat#715-165-150
Goat polyclonal anti-rat Alexa Fluor™ 647	Invitrogen	Cat#A21247
Goat polyclonal anti-rat Alexa Fluor™ 555	Invitrogen	Cat#A21434
Goat polyclonal anti-rabbit Alexa Fluor™ 488	Invitrogen	Cat#A32731
Goat polyclonal anti-mouse Alexa Fluor™ 568	Invitrogen	Cat#A11031
Goat polyclonal anti-mouse Alexa Fluor™ 647	Invitrogen	Cat#A21235
Mouse monoclonal anti-β actin HRP-conjugated	Abcam	Cat#8226
Mouse polyclonal anti-rabbit HRP-conjugated	Santa Cruz	Cat#sc2357
Bacterial and virus strains		
VSV-G ΔEnv HIV-1 (BRU) IN(HA) ΔNef	Petit et al., 1999	
VSV-G ΔEnv HIV-1 (NL4.3) ΔNef IRES GFP	Rensen et al., 2021	
VSV-G Δenv HIV-1 (NL4.3) ΔR LUC	Di Nunzio et al., 2013	
DH5α *E.coli* competent cells		This study
One Shot™ BL21 Star™ (DE3) Chemically Competent *E. coli*	Invitrogen	Cat#C601003
Stellar™ Competent Cells	Takara	Cat#636763
Chemicals, peptides, and recombinant proteins		
BamHI	New England BioLabs	Cat#R3136S
Bovine Serum Albumin (BSA)	Sigma-Aldrich	Cat#A9647
Bradford	Biorad	Cat#500-0006
Calcium Chloride solution	Sigma Aldrich	Cat #21115
CIP	New England BioLabs	Cat#M0525S
cOmplete, EDTA free (tablet)	Sigma Aldrich	Cat#11873580001
Deionized Formamide	Bio Basic	Cat#FB0211
DMEM (1X) + GlutaMAX™-I	Gibco	Cat#31966
DPBS	Gibco	Cat#14200
ECL solution	Cytiva	Cat#RPN2232
Ethanol Absolute	Fisher BioReagents	Cat#BP2818-500
Ethylenediamine tetra-acetic acid disodium salt solution (EDTA)	Sigma Aldrich	Cat#E7889

Foetal Bovine Serum (FBS)	Serana®	Cat#S-FBS-SA-015
Glycine	Sigma	Cat#G8898
HaloTag®TMR Ligand	Promega	Cat#G8252
HEPES-buffered saline, pH 7.0 (2X for transfection)	Fisher Scientific	Cat#15450257
HEPES solution (for cell culture)	Gibco	Cat#15630
Hoechst 33342	Invitrogen	Cat#H3570
In-Fusion® Snap Assembly Master Mix	Takara	Cat#638947
Luciferase Assay System	Promega	Cat#E4030
NEBuffer 3	New England BioLabs	Cat#B7003S
Nevirapine	Sigma Aldrich	Cat#SML0097
NucleoSpin Plasmid	Macherey-Nagel	Cat#740588.50
NuPAGE™ Bis-Tris Mini Protein Gels, 4–12%	Invitrogen	Cat#NP0322BOX
P3 Primary Cell 4D-Nucleofector™ X Kit S	Lonza	Cat# V4XP-3032 (?)
Paraformaldehyde 32%	Electron Microscopy Sciences	Cat#15714
PBS (RNase-free)	Invitrogen	Cat#AM9625
PCR clean-up and gel extraction	Macherey-Nagel	Cat#740609.50
Penicillin Streptomycin (P/S)	Gibco	Cat#15140
PF74	sigma	Cat#SML0835
Phorbol 12-myristate 13-acetate (PMA)	Sigma	Cat#P8139
Pierce™ BCA Protein Assay Kits	Thermo Scientific	Cat#23225
Poly-L-lysine solution	Sigma	Cat#P4707
Precision Plus Protein Western C Standard solution	BioRad	Cat#1610376
ProLong™ Diamond Antifade Mountant	Thermo Fisher Scientific	Cat#P36970
QuikChange II site-directed mutagenesis kit	Stratagene	
Reducing Agent 20X	BioRad	Cat#1610792
RIPA buffer	Sigma-Aldrich	Cat#R0278
RNaseZAP™	Sigma-Aldrich	Cat#R2020
RPMI Medium 1640 (1X) + GlutaMAX™-I	Gibco	Cat#61870
Running buffer 20X	Invitrogen	Cat#NP0001
Sample Buffer 4X	BioRad	Cat#1610791
Stellaris® RNA FISH Hyb. Buffer	Biosearch™ Technologies	Cat#SMF-HB1
Stellaris® RNA FISH Wash Buffer A	Biosearch™ Technologies	Cat#SMF-WA1
Stellaris® RNA FISH Wash Buffer B	Biosearch™ Technologies	Cat#SMF-WB1
T4 Ligase	Thermo Fisher Scientific	Cat#EL0011
Transfer buffer 20X	Invitrogen	Cat#NP0006-1
Triton™ X-100	Sigma	Cat#T8787
TrueCut™ Cas9 Protein v2	Thermo Fisher	Cat# A36496
Trypsin 0.05%-EDTA (1X)	Gibco	Cat#253000
TWEEN® 20	Sigma-Aldrich	Cat#P1379
Phusion Flash High-Fidelity PCR Master Mix	Thermo Scientific	Cat#F-548S

UltraPure™ Distilled Water (H_2O)	Invitrogen	Cat#10977
Vanadyl Ribonucleoside Complexes (VRC)	Sigma	Cat#94742
Experimental models: Cell lines		
HEK293T	American Type Culture Collection (ATCC)	CRL-3216™
Halo tagged SRRM2 HEK293 cells	Parker et al., 2021	
Halo tagged SRRM2 ΔIDR HEK293 cells	Parker et al., 2021	
THP-1	American Type Culture Collection (ATCC)	TIB-202™
THP-1 KO CPSF6	This paper	
Oligonucleotides		
Primary smiFISH probes against HIV-1 pol (× 24)	Integrated DNA Technologies	
Cy5 FLAP (RNA FISH secondary probe)	Eurofins Genomics	
Table1 primers	This paper	
Recombinant DNA		
HIV-1 (BRU) ΔEnv IN(HA) plasmid	Petit et al., 1999 ; Petit et al., 2000	
pNL4.3 Δenv ΔNef IRES GFP	Rensen et al., 2021	
pNL4.3 Δenv ΔNef Luc	Di Nunzio et al., 2013	
pCMV-VSV-G	Addgene	Cat#8454
pSD-GP-NDK	PlasmidFactory.com	
SIVMAC Vpx	Berger G. et al., 2008	
pSICO CPSF6-mNeonGreen	Addgene	Cat#167587
pLPCX CPSF6 ADD2	Gift from Mamuka's lab	
Table2 pSICO-CPSF6 mutants	This paper	
Software and algorithms		
ZEISS arivis Scientific Image Analysis	ZEISS	https://www.arivis.com/
Fiji	Schindelin et al., 2012	https://imagej.net/software/fiji/
Icy	De Chaumont et al., 2012	https://icy.bioimageanalysis.org/download/
Prism	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
Other		
Centrifuge	ThermoFisher Scientific	Sorvall ST4 Plus
BD SYRINGE 60ml (no needle)	Dutscher	309653
Confocal microscope	Zeiss	LSM700 inverted
Confocal microscope 63´ objective	Zeiss	Objective Plan-Apochromat 63x/1.4 Oil DIC M27
Corning® black 96 Well Solid Polystyrene Microplate	Merck	Cat#CLS3916
Enspire 2300	Perkin Elmer	
Falcon® Cell culture 24-well plate	Dutscher	Cat#353047
Glasstic™ Slide 10 with Grids	KOVA™	Cat#87144E
HiTrap™ Q HP column	Cytiva	Cat# 17115401
HiTrapTM SP FF column	Cytiva	Cat#17515701

Ibidi microdishes 35 mm high	Ibidi	Cat#81158
LS Columns	Miltenyi Biotec	Cat#130-042-401
Minisart™ NML Syringe Filter 0.45um	Sartorius	Cat#16555
MS Columns	Miltenyi Biotec	Cat#130-042-201
Open-Top Thinwall Polypropylene Konical Tube, 31.5 mL, 25 ´ 89mm	Beckman Coulter, Inc.	Cat#358126
OPTILUX Petri dish - 100 x 20 mm	Dutscher	Cat#353003
Precision cover glasses 12mm Æ thickness No. 1.5H	Marienfeld	Cat#0117520
Refrigerated benchtop centrifuge	Eppendorf	Centrifuge 5415 R
Rotor for LE-80K	Beckman Coulter, Inc.	SW32Ti
SnakeSkin™ Dialysis Tubing 10,000 MWCO	Thermo Scientific	Cat#68100
Syngene GeneGenius Bio Imaging System Gel Documentation UV Transilluminator	Syngene	
Star-Frost slides 76 x 26 mm	Dutscher	Cat#100204
Stericup Quick Release-HV Sterile Vacuum Filtration System 0.45µm	Millipore	Cat#S2HVU05RE
Thermocycler	Eppendorf	Mastercycler® nexus
Thermocycler (qPCR)	Eppendorf	Mastercycler® realplex^2
Ultracentrifuge	Beckman Coulter, Inc.	Optima™ LE-80K

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Francesca Di Nunzio (francesca.di-nunzio@pasteur.fr).

Materials availability

All materials are available on request.

Table 1: Oligonucleotides

Name	Sequence
For CA insertion	TGGACAACCTTTTGGGCAGCCTCCATTGGGTCCCTTTCCACCTCGTCCACC
Rev CA insertion	CCAAAAGGTTGTCCAGGAAAAAGAACTGGCTCTATGCTGTGCCCCATATGC
For del CA binding	CTCCACCACCACTTCCTCCTGGCCCT
Rev del CA binding	GAAGTGGTGGTGGAGGGCGATCTCCT
For BamHI (ΔLCR+ADD2)	GAACCGTCGGATCCCAGTGTGGTGGTACGGGAATTCAGGAAGATGGCGGACGGCGTGGACCACATAGACATTTACG
Rev BamHI (ΔLCR+ADD2)	GACCGGTGGATCCCGGGCCCGCGGTACCTCACGATGACGATATTCGCGCTCTCGGTCACG
For del MCD + 2NLS	CAAAGTGGGCGGCAGCCCGAAAAAAAAACGCAAAGTGCCAAAAAAGAAAAGAAAAGTTGGCC
Rev del MCD + 2NLS	CTGCCGCCCACTTTGCGTTTTTTTTTCGGTTCCAGAATAAGAACTTTGCAACGATCATC
For del MCD 520-588	AGTCCAAGCGTGAGGTACCGCGGGCC
Rev del MCD 520-588	CCTCACGCTTGGACTCAATTCCATGAAGGC
For CPSF6 260-359	GCATAGAGCCCTTTCCACCTCGTCCACC
Rev CPSF6 260-359	GAAAGGGCTCTATGCTGTGCCCCATATGC
ForCPSF6 w/o mNG	GTCATCGTTAAGCGGCCGCATAACTTCG
RevCPSF6 w/o mNG	CCGCTTAACGATGACGATATTCGCGCTCTC
ForCPSF6 del MCD w/o mNG	AAGTTGGCTAAGCGGCCGCATAACTTCG
RevCPSF6 del MCD w/o mNG	CCGCTTAGCCAACTTTTCTTTTCTTTTTTGGC
Rev seq CPSF6 mut	CATGATCTCGTGAAGCGAGC
For seq CPSF6ADD2	AAATGTCGTAACAACTCCGC
Rev seq CPSF6ADD2	CTTTGCTACCCAATGGAGG
For BamHI CPSF6	GACCGGTGGATCCCGGGCCCGCGGTACCTCACGATGACGATATTCGCGCTCTCGGTCACGC
Rev BamHI CPSF6	GAACCGTCGGATCCCAGTGTGGTGGTACGGGAATTCAGGAAGATGGCGGACGGCGTGGACCACATAGACATTTACG
SRRM2-ASO	ATCTACTACGAGACCTGCGTC
SON-ASO	ACATCTGAGCATCCATGGTGC
SCR-ASO	TGACCCTAtgctgtTCCTATA

Table 2: Composition and compounds used to obtain the engineered CPSF6.

Construct	Modification (aa)	Plasmid/s used	Primers used
pSICO CPSF6 ΔFG	Deletion, aa 313-327	pSICO CPSF6-mNeonGreen	C,D (deletion) M,N (removal of mNG)
pSICO CPSF6 ΔLCR	Deletion, aa 260-312 and 328-359	pSICO CPSF6-mNeonGreen	K,L (deletion) A,B (insert of FG) M,N (removal of mNG)
pSICO CPSF6 ΔLCR-ΔFG	Deletion, aa 260-359	pSICO CPSF6-mNeonGreen	K,L (deletion) M,N (removal of mNG)
pSICO CPSF6-ΔLCR ADD2	Replacement, CPSF6 -ΔLCR ADD2 from pLPCX CPSF6 ADD2 at the place of CPSF6 in pSICO CPSF6-mNeonGreen	pSICO CPSF6-mNeonGreen (backbone) pLPCX CPSF6 ADD2 (insert)	E,F M,N (removal of mNG)
pSICO CPSF6-ΔMCD-3xNLS	Deletion, aa 520-588; Insertion of 3xNLS at C term	pSICO CPSF6-mNeonGreen	I,J (deletion and insertion of NLS) G,H (insertion of 2 NLS) O,P (removal of mNG)