SURVEY AND SUMMARY

Factors that mold the nuclear landscape of HIV-1 integration

Gregory J. Bedwell[1,2] **and Alan N. Engelman** [1,2,*]

[1]Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, MA 02215, USA and
[2]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA

ABSTRACT

The integration of retroviral reverse transcripts into the chromatin of the cells that they infect is required for virus replication. Retroviral integration has far-reaching consequences, from perpetuating deadly human diseases to molding metazoan evolution. The lentivirus human immunodeficiency virus 1 (HIV-1), which is the causative agent of the AIDS pandemic, efficiently infects interphase cells due to the active nuclear import of its preintegration complex (PIC). To enable integration, the PIC must navigate the densely-packed nuclear environment where the genome is organized into different chromatin states of varying accessibility in accordance with cellular needs. The HIV-1 capsid protein interacts with specific host factors to facilitate PIC nuclear import, while additional interactions of viral integrase, the enzyme responsible for viral DNA integration, with cellular nuclear proteins and nucleobases guide integration to specific chromosomal sites. HIV-1 integration favors transcriptionally active chromatin such as speckle-associated domains and disfavors heterochromatin including lamina-associated domains. In this review, we describe virus-host interactions that facilitate HIV-1 PIC nuclear import and integration site targeting, highlighting commonalities among factors that participate in both of these steps. We moreover discuss how the nuclear landscape influences HIV-1 integration site selection as well as the establishment of active versus latent virus infection.

INTRODUCTION

Retroviruses are enveloped viruses that contain two copies of plus-stranded RNA. The HIV-1 ribonucleoprotein complex, composed of the RNA bound by viral nucleocapsid (NC) protein as well as integrase (IN) and reverse transcriptase (RT) enzymes, is housed within a capsid shell made from ∼200 capsid protein (CA) hexamers and 12 CA pentamers (1,2). Together, these elements form the viral core. HIV-1 infects CD4+ cells including T cells and macrophages by fusing its membrane with the cellular plasma membrane (3). Once membrane fusion is complete, the core is released into the cytoplasm and reverse transcription ensues within the confines of the reverse transcription complex (RTC), a high molecular weight derivative of the viral core (Figure 1) (4). The RTC interacts with components of the cell cytoskeleton to enable its inward movement through the cytoplasm and toward the nucleus [reviewed in (5)]. Reverse transcription yields linear double-stranded viral DNA (vDNA) with several internal discontinuities amid the plus-strand (6,7).

A multimer of IN binds and bridges both ends of vDNA together to form the intasome nucleoprotein complex [reviewed recently in (8)]. Two IN activities, 3′ processing and strand transfer, are required for integration. IN hydrolyzes vDNA ends during 3′ processing to yield recessed CA_{OH}-3′ termini, converting the RTC to the preintegration complex (PIC) (9) (Figure 1). After engaging a suitable target DNA (tDNA) acceptor site, IN uses the vDNA 3′-OHs to cut the major groove in staggered fashion, joining the vDNA ends to the resulting 5′-phosphate groups. The gaps between vDNA 5′ ends and tDNA 3′ ends in the hemi-integrant are repaired by cellular machinery to yield stably integrated provirus flanked by a short duplication (4–6 bp across retroviruses; 5 bp for HIV) of chromosomal sequence that was cut during strand transfer. A detailed overview of retroviral integration can be found in reference (10).

*To whom correspondence should be addressed. Tel: +1 617 632 4361; Fax: +1 617 632 4338; Email: alan_engelman@dfci.harvard.edu
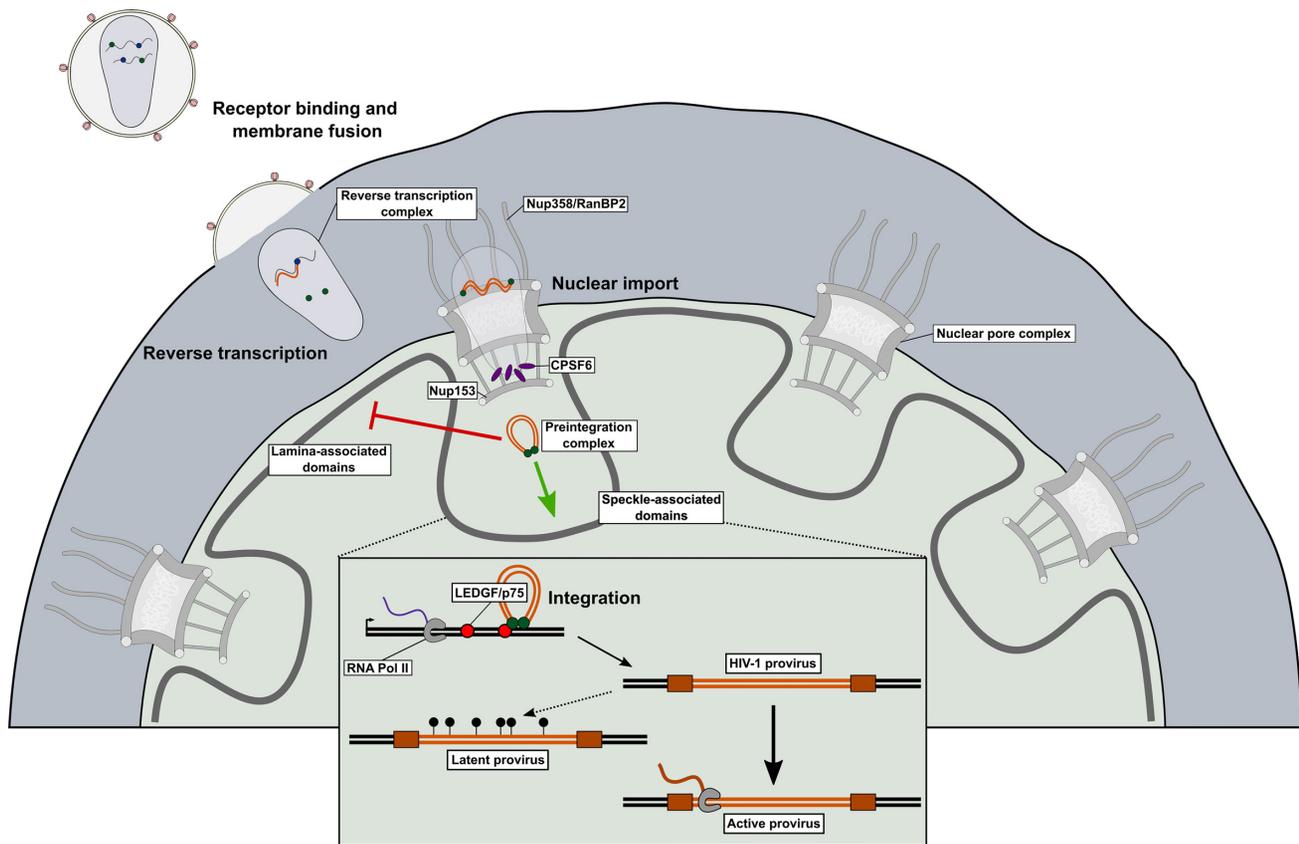
Figure 1. Overview of HIV-1 cellular ingress and principle determinants of integration targeting. Infection is initiated by receptor binding and membrane fusion, which releases the viral core into the cell cytoplasm where reverse transcription begins. During reverse transcription the core is trafficked to the nuclear pore where it is transported into the nucleus via interactions between HIV-1 CA and several nucleoporins, including Nup358 and Nup153. Following translocation, CPSF6 frees the core from the nuclear pore complex and facilitates progression of the PIC beyond the nuclear periphery and into the nuclear interior. Integration is highly biased away from lamina-associated domains and towards speckle-associated domains, which are characterized by active transcription and high gene density. The interaction of PIC-borne IN with LEDGF/p75 directs integration into the interior of gene bodies. HIV-1 proviruses are typically well expressed by cellular RNA polymerase following integration. However, a small population of proviruses (marked by lollipops) are not expressed and become latent. These latent proviruses can be activated upon stimulation years after the initial infection.

HIV-1 efficiently infects non-dividing cells (11,12) due to the active nuclear import of its PIC (13). Once inside the nucleus, integration preferentially occurs in regions of the genome characterized by high gene density and transcriptional activity (14). In particular, HIV-1 integration has been mapped to genomic regions in close proximity to speckle-associated domains (SPADs) and far from heterochromatin markers such as lamina-associated domains (LADs) (Figure 1) (15–18). These integration site selection biases are influenced by a number of factors, including interactions between the PIC and host proteins, the route of nuclear entry, chromatin accessibility, and local nuclear environment (16–25). In this review, we provide an overview of the factors known to influence HIV-1 integration targeting in the human genome. In addition, we discuss how integration site selection relates to the establishment of active versus latent HIV infection.

Access of cell nuclei by HIV-1 PICs

Nucleocytoplasmic transport. Because some of the host factors that help to determine chromosomal sites of HIV-1 integration also play roles in viral nuclear import, a review of integration site targeting necessitates a parallel discussion of PIC nuclear translocation. In this section we briefly review the process of cellular nucleocytoplasmic transport, paying particular attention to aspects that pertain to HIV-1. Readers interested in comprehensive reviews of cellular nuclear import are directed elsewhere (26,27).

Nucleocytoplasmic transport of large macromolecules and macromolecular complexes is regulated by the nuclear pore complex (NPC), a huge ~110 MDa assembly of ~1000 proteins composed of 33 nucleoporins (Nups) arranged in 8-fold rotational symmetry [reviewed in (27)] (Figure 2A–C). The NPC is constructed from Nup subcomplexes referred to as the coat Nup complex or the Y-complex, inner ring Nups, pore membrane proteins (POMs), cytoplasmic filament Nups, and nuclear basket Nups (Figure 2A). About one-fourth of Nup proteins possess FG dipeptide repeats within intrinsically disordered domains that are enriched for polar amino acid residues and depleted of charged residues (28). The central channel of the NPC, ~42 nm in diameter in human cells, is lined with FG repeat Nups that restrict the passive diffusion of proteins greater than ~40 kDa (29).
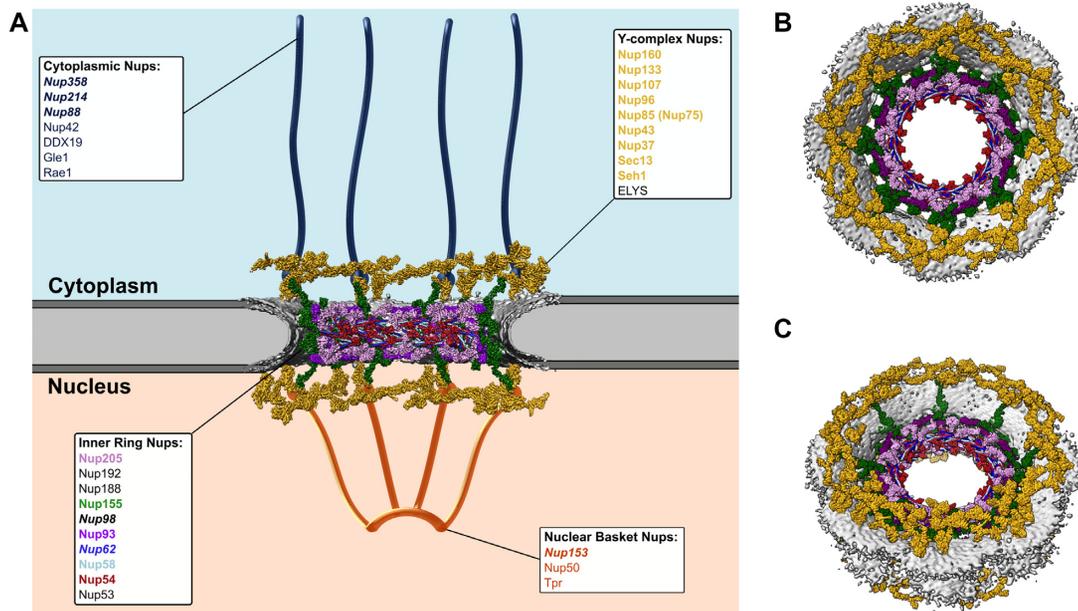
Figure 2. Nuclear pore complex organization and CA-interacting components. (**A**) Cartoon depiction of the nuclear pore complex. Structural elements were derived from entry 3103 in the Electron Microscopy Data Bank (EMDB) and PDB entries 5a9q and 5ijn in the Protein Data Bank (PDB). The diagram depicts the nuclear pore complex as a vertical cross-section through the 8-fold symmetric architecture, revealing Y-complex Nups and several inner ring Nups. The locations of cytoplasmic filament Nups and nuclear basket Nups were approximated manually. POMs are not depicted. The identities of individual nucleoporins that are depicted in the cartoon are labeled in matching colors and bold-face font. Nups previously shown to facilitate PIC nuclear import and/or interact with HIV-1 CA are labeled in italicized bold-face font. (**B, C**) Different perspectives of the intact 8-fold symmetrical nuclear pore. (**B**) Top-view and (**C**) view tilted 40° from the top clearly highlight the overall toroidal architecture of the complex.

A variety of mechanisms have been characterized to facilitate the nuclear import of large cytoplasmic cargos. Classically, soluble β- or α-karyopherin nuclear transport factors engage cargo proteins through modular nuclear localization signals (NLSs). In some cases, the NLS-containing cargo first binds an α-karyopherin adapter protein before complexing with a β-karyopherin partner, whereas in other cases the β-karyopherin engages the NLS-cargo protein directly [reviewed in (26)]. The β-karyopherin component of the complex then docks to the NPC to effect nuclear translocation [see (30) for review]. Some proteins, such as transcription factor SPL1, by contrast can gain access to the NPC through direct binding by Nup proteins, in this case via Nup62 and Nup153 (31).

HIV-1 PIC nuclear import. Initial HIV-1 nuclear import studies took reductionist approaches guided by the classical view of nucleocytoplasmic transport to analyze karyophilic properties of individual PIC components. The reasoning was that if proteins in isolation were karyophilic and possessed *bona fide* NLSs, these would function in the PIC to affect HIV-1 nuclear import. Although such approaches revealed that matrix (32), IN (33,34), and viral protein R (35) were indeed karyophilic, the relevance of these findings to PIC nuclear import proved difficult to reproduce in independent studies (36–40). A 99 nt overlap in the mid-region of the vDNA plus-strand, termed the central DNA flap, was also proposed to mediate PIC nuclear import (41). Although follow-up work discounted a major role for the flap (39,42,43), it can modestly influence the kinetics of nuclear vDNA accumulation (40,44,45). The defining ex-

periment in the field of HIV-1 nuclear import came from studying chimeric viruses between HIV-1 and Moloney murine leukemia virus (MLV), a gammaretrovirus that unlike HIV-1 is unable to infect growth-arrested cells. These data clarified that CA is the determinant required for HIV-1 to productively infect non-dividing cells (46). Because HIV-1 gains nuclear access similarly in cycling and growth-arrested cells (47), results derived from growth-arrested cells pertain to interphase cells as well.

CA is composed of two alpha-helical domains, the N-terminal domain (NTD) and C-terminal domain (CTD), which are separated by a flexible linker (5). Intermolecular NTD-NTD and CTD-NTD interactions juxtapose CA molecules into ringlike hexameric and pentameric capsomers (Figure 3A) (48,49). Higher-order CTD-CTD interactions among adjoining capsomeres template the canonical honeycomb pattern formation within the assembled conical capsid shell (Figure 3B) (1,2,50,51).

CA was shown recently to bind the β-karyopherin transportin 1 (TRN-1) (52) and can also interact with NPC proteins Nup62, Nup88, Nup98, Nup153, Nup214, and Nup358 (Figure 2) (20,53–56). Other CA-binding proteins that can affect HIV-1 PIC nuclear import include cyclophilin A (CypA) (57,58) and cleavage and polyadenylation specificity factor 6 (CPSF6) (Figure 4A) (59–62).

Host factors engage CA via two common binding regions. One is the CypA-binding loop, which links alpha helices 4 and 5 within the NTD (Figure 4B) (63). Nup358, which is a cytoplasmic filament Nup, contains a C-terminal cyclophilin-homology domain (CHD) that likewise engages the CypA binding loop (Figure 4B) (19,20,64). Biochem-
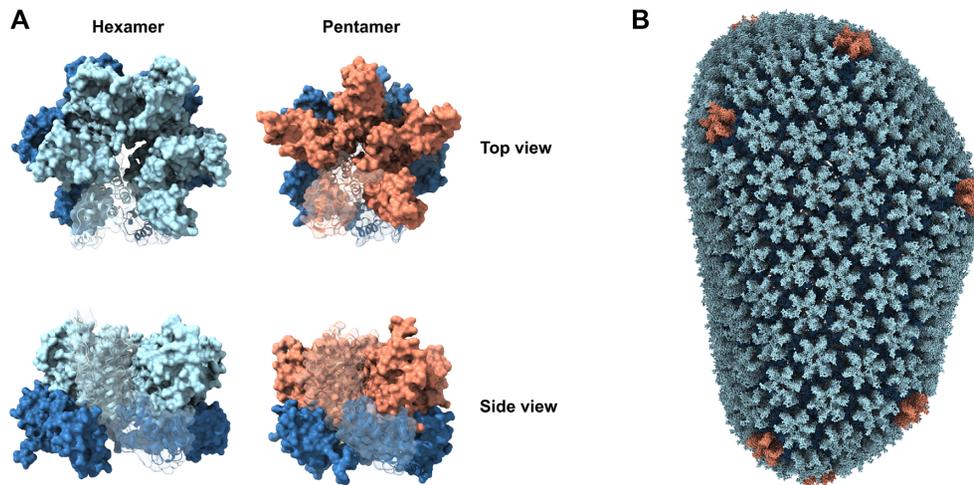
Figure 3. HIV-1 capsomeres and the capsid lattice. (**A**) The capsid lattice is comprised of exactly 12 CA pentamers and ∼200 hexamers, which are shown in top and side views. In both capsomeres the CA CTD is shown in dark blue. While the CA NTD within the hexamer is light blue, it is shown in orange within the pentamer. In each capsomere, a single CA subunit is depicted as a ribbon diagram with matched coloring scheme. (**B**) An all-atom model of the assembled capsid shell derived from PDB entry 3j3y using the color scheme defined in panel A.
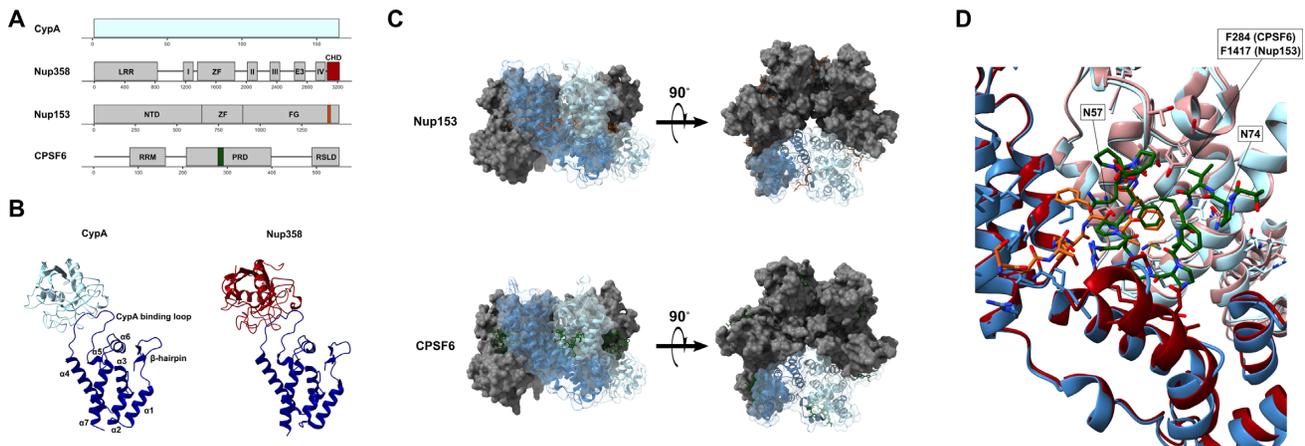
Figure 4. Capsid interactions in HIV-1 integration targeting. (**A**) Organization of principle CA-binding proteins including CypA, Nup358, Nup153 and CPSF6. Locations of CA-binding portions of Nup358, Nup153 and CPSF6 are colored as in subsequent panels. Domain labels are as follows: LRR – leucine-rich region; roman numerals I–IV – Ran binding domains I–IV; ZF – zinc finger; E3 – E3 ligase domain; CHD – cyclophilin homology domain; NTD – N-terminal domain; FG – phenylalanine/glycine repeat domain; RRM – RNA recognition motif; PRD – proline-rich domain; RSLD – arginine/serine-like domain. (**B**) Interactions of Cyp-like protein domains with the CA NTD. CypA and the C-terminus CHD of Nup358 both interact with the conserved CypA-binding loop, one of two principle binding sites within HIV-1 capsomeres. The CypA and Nup358 structures were derived from PDB entries 1ak4 and 4lqw, respectively. Secondary structural elements of CA are noted on the leftward image. (**C**) Structures of hexameric capsomeres with peptides derived from Nup153 (top) and CPSF6 (bottom) from PDB entries 4u0c and 4wym, respectively. These peptides lie in a pocket formed between two individual CA subunits. The binding orientations of the respective peptides are non-identical, highlighting the promiscuity of this binding pocket for mediating CA-host factor interactions. (**D**) Detailed superposition of Nup153 and CPSF6 peptides in complex with CA. Nup153-bound CA molecules are shown in blue while the CPSF6 CA pair is shown in red. For both pairs of CA molecules, individual monomers are differentiated by light and dark coloring. The interaction of each host factor with respective background CAs (light coloring) is anchored by a phenylalanine residue [F284 in CPSF6 (isoform 1 numbering scheme) and F1417 in Nup153]. Docking of this phenylalanine facilitates main chain hydrogen bonding of each host factor with CA residue N57. N74 in CA by contrast preferentially interacts with CPSF6 (green sticks) and not Nup153 (orange sticks). The most pronounced differences in binding modes are derived from interactions with the foreground (darker) CA monomer. CPSF6 adopts a nearly cyclic conformation and interacts primarily with the CTD of the second monomer. In contrast, Nup153 is more linear and interacts with the NTD of the foreground CA subunit.

ical, genetic and molecular modeling experiments suggest that TRN-1 also engages CA via the CypA binding loop (52) though, unlike CypA and the Nup358 CHD, a CA-TRN-1 complex structure has not been solved using wet-bench approaches such as X-ray crystallography. The second region within CA, a pocket that is primarily formed by NTD alpha helices 3–5 with contributions from a neighboring CA within the hexamer, is where Nup153 and CPSF6

bind (Figure 4C) (54,55,65–67). The FG repeat and central proline-rich domains of Nup153 and CPSF6, respectively, confer binding to CA (54,65). Due to the intrinsically disordered nature of these protein domains, host-CA structures to date have been restricted to Nup153- and CPSF6-derived peptides (55,65–67). For both proteins, a phenylalanine residue of an FG dipeptide occupies the binding pocket, with additional interactions made with the ad-

joining CA (Figure 4D). Both peptides accordingly bound hexameric capsomeres ∼10-fold more efficiently than the isolated NTD (66). Although Nup62, Nup88, Nup98 and Nup214 in cell extracts co-pelleted with nanotubes assembled from recombinant CA or CA-NC proteins *in vitro* (53,55,56), direct interactions with CA, as demonstrated through the use of purified Nup proteins or peptides, have not in these cases been confirmed.

The shedding of CA from the core/RTC defines the process of uncoating. Though initially thought to occur soon after virus entry [see (68) for a review], recent evidence suggests that the PIC structurally resembles the intact or nearly intact core during nuclear import (62,69,70). Uncoating (62,69) and the termination of reverse transcription (18,62,69,71,72) are accordingly now thought to occur after nuclear entry. The following scenario can be envisaged for HIV-1 nuclear import. The CA-Nup358 interaction initially docks the PIC to the NPC (20,73,74). One potential role for CypA in nuclear import could be regulation of the CA-Nup358 interaction. Because the CHD has been shown to be dispensable for HIV-1 infection, it seems possible the PIC could also engage one or several of the FG repeats present throughout Nup358 (75). Although cytoplasmic filament Nup214 can co-sediment with CA-NC *in vitro* (55), its role in HIV-1 infection has been mapped to the post-integration step of mRNA nuclear export (73). After docking, the PIC is shuttled through the NPC in a process possibly involving TRN-1 and/or additional FG repeat Nups such as Nup62 and Nup153. Because the diameter of the wide end of the conical core is ∼60 nm (51,76), it is unclear how an intact or nearly intact core could pass through. We envision that structural flexibility, possibly imparted from both within the viral complex [e.g. vDNA plus-strand discontinuities (77) that enable remodeling (78)] and external to the core [e.g. dynamic nature of the NPC (79,80)], work together to directionally 'massage' the oversized cargo. Eventually, the core docks on the nuclear side of the NPC at the nuclear basket via the CA-Nup153 interaction (54,55,67). Recent evidence has suggested that CPSF6 displaces Nup153 from the capsid lattice by competing for the same binding pocket and frees the HIV-1 PIC from the NPC to further its journey into the nucleus (Figure 1) (61).

Because changes in CA residues can dramatically alter HIV-1's dependency on specific nuclear import factors, it seems that alternative nuclear import pathways must exist for the PIC (59,72). Although the dominant viral determinant for HIV-1 nuclear import is CA (39,46), IN has continued to garner significant focus (81–86) and IN-host interactions could in theory predominate if alternative import pathways are less reliant on CA interactors. IN can interact with numerous soluble transport receptors including α-karyopherin KPNA2/β-karyopherin KPNB1, α-karyopherin KPNA4, TRN-1 and β-karyopherin transportin 3 (TRN-SR2/TNPO3) [reviewed in (87)], as well as NPC components Nup62 (88) and Nup153 (89). However, the relevance of several of these interactions, including those with KPNA4 (86), TNPO3 (90–92), and Nup153 (53,54), have been brought into question. Although HIV-1 infection is significantly reduced via TNPO3 depletion (19,83,90,93,94), this appears to be an indirect consequence of restriction of virus infection due to enhanced cytoplasmic

CPSF6 accumulation (59,92,95–97). Comparatively weak, ∼0.1 mM, small molecule inhibitors of IN-TNPO3 (98) and IN-KPNA2/KPNB1 (99) interactions have been reported. Inhibitors with minimally 10-fold increases in potency together with the selection of drug resistance that maps to IN should be accomplished to convincingly demonstrate a pharmacological role for IN in HIV-1 PIC nuclear import. For comparison, low μM to sub-nM HIV-1 inhibitors that bind the CA Nup153/CPSF6 binding pocket displace these proteins and, as part of their multimodal mechanisms of action, inhibit PIC nuclear import (54,55,65,100–102).

HIV-1 integration site targeting

Integration site preferences. Several processes impact the selection of retroviral integration sites in animal cell genomes, with viruses that make up the different genera of Retroviridae displaying largely similar preferences for functional elements such as genes and promoter regions [reviewed in (103)]. High-resolution mapping studies demanded genome-wide capabilities, which were enabled in 2001 via the release of the draft human genome (104). The first genome-wide study, conducted by the Bushman laboratory, revealed that HIV-1 integration is highly biased towards gene-dense regions and highly expressed genes (14). HIV-1 integration was subsequently shown to track with histone modifications associated with active chromatin such as H4K16ac, H3K36me3 and H3K4me1, and disfavor repressive heterochromatin markers such as H3K9me3, H3K27me3 and LADs (15,16). Recent results have clarified that HIV-1 integration highly favors SPADs (18), which are genomic DNA regions that physically associate with nuclear speckles (105,106).

Comparisons of genic HIV-1 integration frequencies across studies revealed the presence of recurrent integration genes or RIGs, which by definition were genes targeted for integration in two or more studies (16,25). RIGs can also be tabulated as genes that are experimentally targeted more frequently than expected based on random chance (17). Imaging HIV-1 proviruses and RIGs in activated CD4+ T cells revealed association of both with the nuclear periphery, defining a specific nuclear architecture for HIV-1 integration site targeting (16). These results were consistent with an independent study that highlighted HIV-1 targeting of chromatin in the peripheral region of the nucleus in a manner dependent on the nuclear basket Nup protein Tpr (23). While some prior studies supported the notion of preferential localization of PICs and proviruses at the nuclear periphery (107,108), subsequent work has highlighted a more pan-nuclear distribution of HIV-1 PICs and integrated proviruses (17,60,62,109,110). In our hands, the vast majority of RIGs harbored pan-nuclear distributions in both transformed HEK293T and activated primary CD4+ T cells (17,18).

Activation of CD4+ T cells yields gross rearrangements in nuclear architecture including actin network formation (111,112). While pan-nuclear RIG distribution was corroborated in resting CD4+ T cells, activation resulted in RIG relocation closer to the nuclear periphery (25). The implications of this reorganization on integration site targeting are not entirely clear. While comparatively slow reverse tran-

scription kinetics limits the efficiency of resting T cell infection *in vitro* (113), these cells nevertheless support HIV-1 integration (114,115). Although genic integration targeting frequencies were similar in resting and activated CD4+ T cells infected with HIV-1 *in vitro*, integration in resting cells occurred in modestly less gene-dense regions of chromosomal DNA (116). The limited number of integration sites recovered from such analyses has precluded detailed RIG analyses (25,116). Scaled-up studies should be performed to ascertain whether RIG usage differs in resting versus activated CD4+ T cells. Because SPADs track with gene-rich chromosomal regions (18), such studies would also critically address SPAD integration targeting as a function of T cell activation.

Super-enhancers (SEs) are genomic regions enriched in enhancers, activating epigenetic marks such as H3K4me1 and H3K27ac, as well as binding regions for certain transcription factors [reviewed in (117)]. Recently, SEs were shown to correlate with RIGs, though not with bulk HIV-1 integration sites (18,25). Because SEs are enriched in SPADs (105), the observed correlations between integration sites and SPADs/SEs are likely convoluted. Additional work is required to clarify the contributions of these overlapping genomic markers as predictors of bulk versus RIG-specific HIV-1 integration targeting frequencies.

IN-binding host factors and HIV-1 integration site targeting. Initial observations that viruses from different retroviral genera displayed dramatically different preferences for promoters versus gene bodies (118,119) indicated that genera-specific host factors could play a role in integration site targeting (120). Indeed, the first cell protein shown to play a significant role in retroviral integration site targeting (121), lens epithelium-derived growth factor (LEDGF)/p75, specifically binds the IN proteins of lentiviruses (Figure 5) (122–124). LEDGF/p75 is a transcriptional co-activator that harbors two conserved domains, an N-terminal Pro-Trp-Trp-Pro (PWWP) domain important for chromatin binding (125–127) and a downstream IN-binding domain (IBD) that is necessary and sufficient to bind HIV-1 IN (128) (Figure 5A and B). LEDGF/p75 significantly stimulated lentiviral IN catalytic activities *in vitro* (124,126,128–131) and tethered ectopically-expressed HIV-1 IN to cellular chromatin (132). The LEDGF/p75 PWWP domain can engage the trimethylated H3K36me3 modification on nucleosomes assembled *in vitro* (133–135) and LEDGF/p75 binding sites in cells correlate with H3K36me2 and H3K36me3 marks (136,137). Through its IBD, LEDGF/p75 tethers several different cellular proteins to chromatin to effect transcriptional programming and leukemogenic transformation (138–140).

Efficient knockdown of LEDGF/p75 by RNA interference yielded at best marginal changes in HIV-1 integration site targeting (24,121). By contrast, genic HIV-1 integration targeting was reduced significantly by knocking out *PSIP1*, the gene that encodes for LEDGF/p75 (141–143). Thus, the normal cellular complement of LEDGF/p75 apparently exceeds by several fold that required by HIV-1 for efficient integration site targeting. LEDGF/p75 is a member of the hepatoma-derived growth factor (HDGF) family, of which one other member, HDGF like protein 2 (HDGFL2), contains an IBD that is homologous to the LEDGF/p75 IBD (128,144). Unlike LEDGF/p75, HDGFL2 at steady-state is found in the nucleoplasm as compared to chromatin-associated (144), which may account for why HDGFL2 appeared to play little if any role in HIV-1 integration site targeting in cells that express LEDGF/p75. A subsidiary role for HDGFL2 in integration site targeting was observed in cells that lacked LEDGF/p75 (145,146).

LEDGF/p75's tell-tale signature in HIV-1 integration targeting came from analyzing genic integration site distributions across all targeted genes. This analysis first revealed that while MLV integration preferred promoter regions, HIV-1 favored the interior regions of gene bodies (118). HIV-1's genic integration targeting preference shifted toward gene 5′ end regions in the absence of LEDGF/p75 (24,147). LEDGF/p75 can interact with numerous mRNA splicing factors (136,147) and overcome the transcriptional block imposed by nucleosomes *in vitro* (137). Thus, current models predict that LEDGF/p75's function in HIV-1 integration site targeting is determined through interactions with cellular mRNA splicing and/or transcriptional elongation machineries.

Although >200 cellular proteins have been reported to interact with HIV-1 IN [reviewed in (103)], we are unaware of studies that directly implicate any of these beyond LEDGF/p75 and HDGFL2 in HIV-1 integration site targeting. IN-binding factors that seemingly could play such a role include IN interactor 1 (INI1)/SMARCB1 (148), which is a component of BAF and PBAF chromatin remodeling complexes [reviewed in (149)], as well as the histone acetyltransferase enzyme EP300 (150). Additional work is required to ascertain whether these or other IN-binding factors beyond LEDGF/p75 and HDGFL2 play a role in HIV-1 integration site targeting.

CA interactors and HIV-1 integration site targeting. Initial glimpses of CA-binding partner functionalities in HIV-1 integration site targeting indicated these may fundamentally differ from LEDGF/p75. Cellular depletion of TNPO3 or Nup358, but not LEDGF/p75, yielded significant reductions in the number of genes per Mb (gene density) that surrounded HIV-1 integration sites (19). Because an HIV/MLV chimeric virus carrying the MLV *gag* gene, which among other things encodes for CA, yielded the same result, Ocweija *et al*. concluded that HIV-1 Gag proteins interact with TNPO3 and Nup358 to target integration to gene-rich chromosomal regions (19). Although as discussed above we now believe that the result with TNPO3 was indirect due to dysregulated CPSF6 localization, preferential disruption of integration into gene enriched regions was observed subsequently in cells depleted for Nup153 (151) or CPSF6 (24). CA mutant viruses with single amino acid changes that disrupt binding to CPSF6, including N74D (20,24,55,59,151) and A77V (152), phenocopied CPSF6 depletion and shifted integration from gene-rich chromosomal regions to gene-sparse regions. Contrastingly, CA changes G89V or P90A, which disrupt the CA-CypA interaction (153,154), retargeted HIV-1 integration to regions marginally more enriched in genes than those targeted by the wild type virus (20).
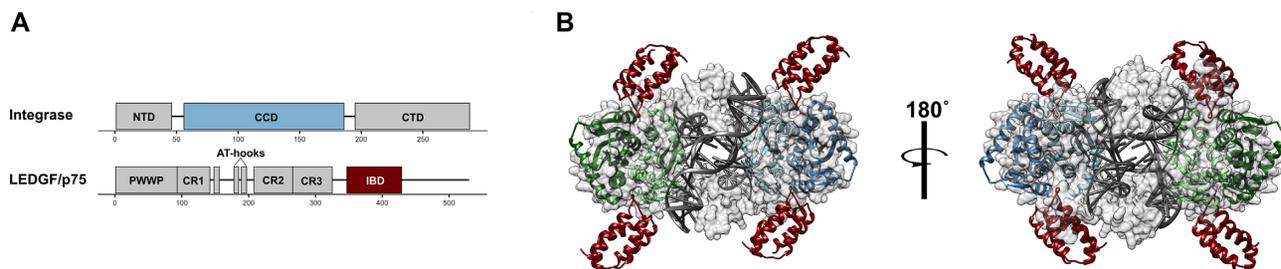
Figure 5. The integrase-LEDGF/p75 interaction. (**A**) Domain organization of HIV-1 IN and LEDGF/p75. Domain annotations are as follows: NTD – N-terminal domain; CCD – catalytic core domain; CTD – C-terminal domain; PWWP – Pro-Trp-Trp-Pro domain; CR – charged region; AT-hooks – adenosine/thymine DNA binding motif; IBD – integrase binding domain. The key interacting domains, the IN CCD and LEDGF/p75 IBD, are colored blue and dark red, respectively. (**B**) Depiction of the core tetramer of the HIV-1 strand transfer complex intasome (PDB 5u1c) bound by the LEDGF/p75 IBD, which was created by superimposing the CCDs of the IBD–HIV-1 IN CCD structure (PDB 2b4j) with the CCDs of the strand transfer complex. Because LEDGF/p75 interacts with integrase at the interface between two CCD dimers, a single intasome contains multiple potential LEDGF/p75 binding sites. Whether all or just a fraction of bound LEDGF/p75 molecules participates in HIV-1/lentiviral integration targeting is not presently known. The IBD color in panel B matches panel A; one of the two CCD dimers in B also matches the panel A coloring.

Results obtained via imaging virus-infected cells have greatly informed the role of CPSF6 in HIV-1 integration site targeting. Depletion of CPSF6 or infection with binding-defective viruses such as N74D and A77V CA mutants resulted in PIC and proviral accumulation in the peripheral region of the nucleus (17,60–62,110,155,156). Concurrent genomic DNA analyses revealed significant upticks in LAD-proximal integration site targeting with parallel reductions in integrations into SPADs (17,18,157). HIV-1 PICs and CPSF6 were accordingly seen to colocalize with nuclear speckles in a variety of acutely-infected cell types including primary CD4+ T cells and macrophages (18).

CPSF6 functions as part of the cleavage factor I mammalian (CFIm) complex, which is one of many complexes that compose the cleavage and polyadenylation complex that processes mRNA 3′ ends for polyadenylation [reviewed in (158)]. CFIm is composed of a heterotetramer of CPSF5 and one of two homologues, CPSF6 or CPSF7 (159). CPSF6 harbors three domains, an N-terminal RNA recognition motif (RRM) that mediates the interaction with CPSF5 (160), a central proline-rich domain that mediates binding to HIV-1 CA (65,161), and a C-terminal arginine/serine-like domain (RSLD) that is enriched in R(D/E) dipeptides and mediates TNPO3 binding (92,97) (Figure 4A). It is somewhat unclear whether CPSF6 function in PIC nuclear import and integration site targeting occurs in the context of CFIm. The vast majority of cellular CPSF6 is sequestered in CFIm (162) and CPSF6- but not CPSF7-containing CFIm colocalized with nuclear PICs (61). However, expression of a CPSF6 RRM deletion mutant defective for CPSF5 binding (163) efficiently restored integration site targeting to *CPSF6* knockout cells (162). These data suggest that CPSF6 need not be complexed with CPSF5 to effect PIC nuclear trafficking to speckles for integration into SPADs. As the CPSF6 RSLD was recently shown to play a role in nuclear speckle condensation (164), we suspect that it largely underlies CPSF6-dependent directional PIC trafficking to nuclear speckles (18).

Analyses of cells knocked out for LEDGF/p75 and/or CPSF6 expression have helped clarify the role of each of these factors in HIV-1 integration site targeting. Because the shift in genic integration site distribution to gene 5′ end

regions observed in *PSIP1* knockout cells was retained in cells knocked out for both factors but absent from *CPSF6* knockout cells, we concluded that LEDGF/p75's primary function is positional integration targeting into gene mid-regions (24). Conversely, peripheral nuclear accumulation of PICs and proviruses with integrations mapping nearby LADs was observed in double knockout as well as *CPSF6* knockout cells, indicating that the main CPSF6 role is enabling PIC passage from the periphery into the nuclear lumen to engage nuclear speckles for SPAD-proximal integration (Figure 1) (17,18). This model contrasts prior ones that invoked NPC-proximal integration targeting as a function of Nup153, Tpr and LEDGF/p75 (16,23). Zones of transcriptional activity map to the nuclear periphery in association with NPCs as well as the nuclear interior in close association with nuclear speckles (105). Additional work conducted in primary cells of HIV-1 infection including CD4+ T cells and macrophages should help to clarify the roles of different host factors in integration targeting under physiologically relevant conditions.

Given the role of CPSF6 in PIC nuclear import together with peripheral PIC and proviral accumulation in the absence of CA-CPSF6 binding, our model certainly invokes that CPSF6 acts prior to LEDGF/p75 (Figure 1). While chromatin-binding is essential for LEDGF/p75's role in HIV-1 integration (127,165), it is less clear if CPSF6 directly tethers PICs to chromatin for integration. Recent research that indicates the PIC retains its CA complement post-nuclear import (62,69,70) is consistent with a model whereby CPSF6 remains PIC-associated post nuclear entry to deliver it to nuclear speckles for integration into SPADs (18) (Figure 1).

The roles of other CA-binding proteins such as Nup358, Nup153 and CypA in HIV-1 integration targeting are less clear. The fact that the Nup358 CHD and CypA share the same binding region on CA (Figure 4B) yet invoke opposite effects on integration into gene-dense regions upon disruption of these virus-host interactions sheds little insight. Possibly, loss of CypA enhances integration into gene-dense chromosomal regions via enhancing the CPSF6-CA interaction and/or slowing the rate of nuclear PIC uncoating. Indeed, CypA was recently shown to block HIV restriction

by the antiviral factor TRIM5α by impeding its interaction with capsid (166). It is possible that this same effect is at-play for other host factors involved in HIV-1 biology. CypA was also recently shown to make novel contacts with two additional hexamers in the assembled capsid honeycomb, though the physiological relevance of these findings is not clear (167). Given the breadth of full-length Nup358 with associated FG repeats, it would not be surprising if novel Nup358-CA interactions await discovery.

Many Nups play important cell biology functions outside of their roles as structural components of the NPC [reviewed in (168)]. For example, Y-complex Nup components ELYS and Sec13 (Figure 2) can alter chromatin functionality via interacting with chromatin remodeling complexes (169). Nup153 was recently shown to effect chromatin organization via interacting with architectural proteins CTCF (CCCTC-binding factor) and cohesion (170). Given that Nup153 and CPSF6 share the same binding pocket on CA (Figure 4C and D), additional work is required to discern whether regulating the CA-CPSF6 interaction or perhaps a novel pathway involving CTCF/cohesion underlies Nup153's role in HIV-1 integration site targeting.

Nucleosomes and tDNA flexibility in integration targeting. Early biochemical studies demonstrated that HIV-1 integration *in vitro* preferentially occurred in the exposed major groove of nucleosomal DNA (171,172). Subsequent work revealed that a direct interaction between HIV-1 IN and the tail region of histone H4 stimulated integration into nucleosomal DNA *in vitro* (173,174). Interestingly, DNA minicircles in large part recapitulated the stimulatory effect of tDNA distortion on HIV-1 integration *in vitro* in the absence of bound protein factors (175).

Studies of prototype foamy virus (PFV) intasomes greatly informed the structural basis of tDNA distortion in retroviral integration. To accommodate scissile phosphodiester bonds that are separated by 4 bp into two IN active sites within the intasome, the tDNA major groove had to distort significantly, to 26.3 Å, with concomitant minor groove compression to 9.6 Å (176). Pyrimidine-purine (YR) dinucleotides, which are inherently flexible, are accordingly naturally selected at the center of the tDNA cut made by PFV IN in cells and *in vitro* (176). Site-directed mutagenesis revealed roles for PFV IN residues Ala188 and Arg329 in dictating nucleobase selection at integration sites (176). Intasome models based on the PFV structures implicated similar roles for HIV-1 IN residues Ser119 and Arg231 (21,22). Statistical analysis of tDNA sequence preferences of HIV-1 integration revealed weak but significant bias towards the consensus sequence RYXRY, which, akin to PFV, enforces YR at the two dinucleotides that span the center of the 5 bp sequence (22). Interestingly, in addition to altering tDNA bases at integration sites, certain Ser119 and Arg231 substitutions marginally shifted sites of HIV-1 integration to gene-sparse genomic regions (21). Because not all Ser119 and Arg231 substituents conveyed this phenotype, the mechanistic basis for global integration retargeting in these cases is unclear.

Genomic features of active versus latent infection. The vast majority of cells that become infected with HIV-1 support active transcription and the production of new viral progeny (177,178). The advent of combinatorial antiretroviral therapy (ART) enabled acute measures of viral and infected cell dynamics, which revealed that infected cells persist with a half-life of ~1–2 days due to either virus-induced cell death or immune system eradication (177–181). ART treatment also helped to unveil a latent population of HIV-1 proviruses in patient-derived samples (182–185) that is established early during the course of infection (186,187). Today, it is widely recognized that this population of reactivatable proviruses is the principle barrier to curative HIV-1 strategies (177,178,188,189).

The precise mechanism responsible for the establishment of latent infection is not clear. HIV-1 primarily infects activated CD4+ T cells, but, as previously mentioned, can infect resting CD4+ T cells *in vitro*. The latent reservoir is most probably established by infection of activated CD4+ T cells that then transition to a resting (memory) state (187,190,191). Interestingly, the vast majority (estimated to be ≥98% in some studies) of integrated proviruses in patient cells are defective due to deletions or hypermutation and thus are incapable of supporting virus replication (189,192–194). These defective proviruses accumulate rapidly during the acute phase of infection (194). A significant fraction of persistently infected cells in patients have moreover been shown to clonally expand (192,195,196). Clonal expansion of infected cells closely parallels seeding of the latent reservoir and tends to increase with time (192,197). Because the majority of the proviruses in these cells are defective, it was initially thought that viral recrudescence upon ART cessation was due to non-clonally expanded but quiescent CD4+ T cells that harbor intact provirus (192). However, subsequent studies have shown that sufficient intact proviruses exist in the clonally expanded population to support the resurgence of viral replication (193,198–200).

Like integration site preferences *in vitro*, the majority of integrations in chronically infected patients are observed in genes (192,195,196). Repressive chromatin marks such as H3K9me3, H3K27me3 and CpG methylation have been linked to the establishment of latency (201–206). Interference of HIV-1 transcription caused by active host gene expression and provirus orientation have been proposed to modulate latency (207–209). Components of the mTOR complex and related downstream factors have also been shown to influence latency, possibly through modulating TCR/CD28 signaling and/or NF-κB activation (210). These observations are consistent with the notion that suppression of viral transcription is a key component of HIV-1 latency. The use of barcoded viruses to track individual integration sites has accordingly indicated that latent proviruses are more distal from activating epigenetic marks than are expressed viruses (206,211). Recently, proviruses from elite controllers, which represent a minority of patients that control their infection in the absence of ART, were shown to adopt a state of 'deep latency' due to integration into heterochromatic regions such as centromeric satellite DNA and Krüppel-associated box domain-containing zinc finger genes (212).

Subsets of genic HIV-1 proviruses have been linked to persistence and clonal expansion of infected cells in patients on long-term ART (195,196,213,214). Integration

into specific regions of *MKL2*, *BACH2* and *STAT5B*, for instance, are highly enriched in patient samples (192,195,196). Interestingly, these genes are linked to tumorigenesis, T cell homeostasis, B cell development, and/or immune signaling (215–220). It has been hypothesized that these innate biological functions underlie their overrepresentation in persistent/clonally expanded infected T cell populations (195). Consistent with this hypothesis, it has been shown that many (but not all) integrations into *BACH2* and *STAT5B* result in splicing-induced fusion of viral sequences to the first protein coding exon of these genes (221) – a mechanism conceptually similar to the fusion proteins found in many cancers (222). These fusions can increase the proliferation and survival of T regulatory cells without imparting deleterious effects on their function (221).

CONCLUSIONS

We have witnessed significant progress in understanding the intricacies of HIV-1 integration site targeting over the past two decades. Animal cell genomic sequences have enabled the mapping of individual retroviral integration sites on massive scales. Intasome studies have informed the structural bases of nucleobase selection at sites of vDNA joining. Advances in cell biology have enabled rapid, targeted ablation of specific cell factors by RNA interference and knockout strategies such as CRISPR-Cas9, greatly accelerating the pace of research. Such approaches are crucial to inform the roles of virus-host interactions in HIV-1 integration targeting. Disruption of HIV-host interactions important for integration site targeting has importantly informed novel antiviral inhibitor development. Small molecule inhibitors of CA-Nup153/CPSF6 (102) and IN-LEDGF/p75 [reviewed in reference (87)] interactions engage multiple copies of their respective viral targets, eliciting multipronged allosteric antiviral responses.

Roles for integration sites in establishing and regulating latency are beginning to emerge. Fundamental questions remain, however. From the perspective of a cure, one of the most pressing questions is the relationship between integration site and proviral transcription. Although chromatin landscape around HIV-1 integration sites can influence viral gene expression (206,211), how this relates to latency and cellular persistence/clonal expansion in patients is not explicitly known. Additional research on understanding the reasons why particular integration sites are enriched in clonally expanded and persistent cells *in vivo* is surely warranted. Plausibly, this could inform the development of novel therapeutics to eradicate viral recrudescence that otherwise widely pervades ART cessation. The observation that a small fraction of patients seemingly self-cure via the elimination of cells that otherwise could reseed virus replication indicates that an immunological approach to HIV cure may be plausible (212).

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Ganser,B.K., Li,S., Klishko,V.Y., Finch,J.T. and Sundquist,W.I. (1999) Assembly and analysis of conical models for the HIV-1 core. *Science*, **283**, 80–83.
2. Zhao,G., Perilla,J.R., Yufenyuy,E.L., Meng,X., Chen,B., Ning,J., Ahn,J., Gronenborn,A.M., Schulten,K., Aiken,C. *et al.* (2013) Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature*, **497**, 643–646.
3. Chen,B. (2019) Molecular mechanism of HIV-1 entry. *Trends Microbiol.*, **27**, 878–891.
4. Fassati,A. and Goff,S.P. (2001) Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1. *J. Virol.*, **75**, 3626–3635.
5. Yamashita,M. and Engelman,A.N. (2017) Capsid-dependent host factors in HIV-1 infection. *Trends Microbiol.*, **25**, 741–755.
6. Charneau,P. and Clavel,F. (1991) A single-stranded gap in human immunodeficiency virus unintegrated linear DNA defined by a central copy of the polypurine tract. *J. Virol.*, **65**, 2415–2421.
7. Miller,M.D., Wang,B. and Bushman,F.D. (1995) Human immunodeficiency virus type 1 preintegration complexes containing discontinuous plus strands are competent to integrate in vitro. *J. Virol.*, **69**, 3938–3944.
8. Engelman,A.N. and Cherepanov,P. (2017) Retroviral intasomes arising. *Curr. Opin. Struct. Biol.*, **47**, 23–29.
9. Miller,M.D., Farnet,C.M. and Bushman,F.D. (1997) Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *J. Virol.*, **71**, 5382.
10. Lesbats,P., Engelman,A.N. and Cherepanov,P. (2016) Retroviral DNA integration. *Chem. Rev.*, **116**, 12730–12757.
11. Weinberg,J.B., Matthews,T.J., Cullen,B.R. and Malim,M.H. (1991) Productive human immunodeficiency virus type 1 (HIV-1) infection of nonproliferating human monocytes. *J. Exp. Med.*, **174**, 1477–1482.
12. Lewis,P., Hensel,M. and Emerman,M. (1992) Human immunodeficiency virus infection of cells arrested in the cell cycle. *EMBO J.*, **11**, 3053–3058.
13. Bukrinsky,M.I., Sharova,N., Dempsey,M.P., Stanwick,T.L., Bukrinskaya,A.G., Haggerty,S. and Stevenson,M. (1992) Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 6580–6584.
14. Schröder,A.R.W., Shinn,P., Chen,H., Berry,C., Ecker,J.R. and Bushman,F. (2002) HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*, **110**, 521–529.
15. Wang,G.P., Ciuffi,A., Leipzig,J., Berry,C.C. and Bushman,F.D. (2007) HIV integration site selection: Analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Res.*, **17**, 1186–1194.
16. Marini,B., Kertesz-Farkas,A., Ali,H., Lucic,B., Lisek,K., Manganaro,L., Pongor,S., Luzzati,R., Recchia,A., Mavilio,F. *et al.* (2015) Nuclear architecture dictates HIV-1 integration site selection. *Nature*, **521**, 227–231.
17. Achuthan,V., Perreira,J.M., Sowd,G.A., Puray-Chavez,M., McDougall,W.M., Paulucci-Holthauzen,A., Wu,X., Fadel,H.J., Poeschla,E.M., Multani,A.S. *et al.* (2018) Capsid-CPSF6 interaction licenses nuclear HIV-1 trafficking to sites of viral DNA integration. *Cell Host Microbe*, **24**, 392–404.

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

19. Ocwieja,K.E., Brady,T.L., Ronen,K., Huegel,A., Roth,S.L., Schaller,T., James,L.C., Towers,G.J., Young,J.A.T., Chanda,S.K. *et al.* (2011) HIV integration targeting: A pathway involving transportin-3 and the nuclear pore protein RanBP2. *PLoS Pathog.*, **7**, e1001313.

20. Schaller,T., Ocwieja,K.E., Rasaiyaah,J., Price,A.J., Brady,T.L., Roth,S.L., Hue,S., Fletcher,A.J., Lee,K., KewalRamani,V.N. *et al.* (2011) HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. *PLoS Pathog.*, **7**, e1002439.

21. Demeulemeester,J., Vets,S., Schrijvers,R., Madlala,P., De Maeyer,M., De Rijck,J., Ndung'u,T., Debyser,Z. and Gijsbers,R. (2014) HIV-1 integrase variants retarget viral integration and are associated with disease progression in a chronic infection cohort. *Cell Host Microbe*, **16**, 651–662.

22. Serrao,E., Krishnan,L., Shun,M.-C., Li,X., Cherepanov,P., Engelman,A. and Maertens,G.N. (2014) Integrase residues that determine nucleotide preferences at sites of HIV-1 integration: implications for the mechanism of target DNA binding. *Nucleic Acids Res.*, **42**, 5164–5176.

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

24. Sowd,G.A., Serrao,E., Wang,H., Wang,W., Fadel,H.J., Poeschla,E.M. and Engelman,A.N. (2016) A critical role for alternative polyadenylation factor CPSF6 in targeting HIV-1 integration to transcriptionally active chromatin. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, E1054–E1063.

25. Lucic,B., Chen,H.-C., Kuzman,M., Zorita,E., Wegner,J., Minneker,V., Wang,W., Fronza,R., Laufs,S., Schmidt,M. *et al.* (2019) Spatially clustered loci with multiple enhancers are frequent targets of HIV-1 integration. *Nat. Commun.*, **10**, 4059.

26. Stewart,M. (2007) Molecular mechanism of the nuclear protein import cycle. *Nat. Rev. Mol. Cell. Biol.*, **8**, 195–208.

27. Lin,D.H. and Hoelz,A. (2019) The structure of the nuclear pore complex (An update). *Annu. Rev. Biochem.*, **88**, 725–783.

28. Rout,M.P. and Wente,S.R. (1994) Pores for thought: nuclear pore complex proteins. *Trends Cell. Biol.*, **4**, 357–365.

29. Timney,B.L., Raveh,B., Mironska,R., Trivedi,J.M., Kim,S.J., Russel,D., Wente,S.R., Sali,A. and Rout,M.P. (2016) Simple rules for passive diffusion through the nuclear pore complex. *J. Cell Biol.*, **215**, 57–76.

30. Lott,K. and Cingolani,G. (2011) The importin β binding domain as a master regulator of nucleocytoplasmic transport. *Biochim. Biophys. Acta*, **1813**, 1578–1592.

31. Zhong,H., Takeda,A., Nazari,R., Shio,H., Blobel,G. and Yaseen,N.R. (2005) Carrier-independent nuclear import of the transcription factor PU.1 via RanGTP-stimulated binding to Nup153. *J. Biol. Chem.*, **280**, 10675–10682.

32. Bukrinsky,M.I., Haggerty,S., Dempsey,M.P., Sharova,N., Adzhubel,A., Spitz,L., Lewis,P., Goldfarb,D., Emerman,M. and Steveson,M. (1993) A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells. *Nature*, **365**, 666–669.

33. Depienne,C., Mousnier,A., Leh,H., Le Rouzic,E., Dormont,D., Benichou,S. and Dargemont,C. (2001) Characterization of the nuclear import pathway for HIV-1 integrase. *J. Biol. Chem.*, **276**, 18102–18107.

34. Gallay,P., Hope,T., Chin,D. and Trono,D. (1997) HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. *Proc. Natl. Acad. Sci. USA*, **94**, 9825–9830.

35. Lu,Y.L., Spearman,P. and Ratner,L. (1993) Human immunodeficiency virus type 1 viral protein R localization in infected cells and virions. *J. Virol.*, **67**, 6542–6550.

36. Freed,E.O., Englund,G. and Martin,M.A. (1995) Role of the basic domain of human immunodeficiency virus type 1 matrix in macrophage infection. *J. Virol.*, **69**, 3949–3954.

37. Reil,H., Bukovsky,A.A., Gelderblom,H.R. and Göttlinger,H.G. (1998) Efficient HIV-1 replication can occur in the absence of the viral matrix protein. *EMBO J.*, **17**, 2699–2708.

38. Lu,R., Limón,A., Devroe,E., Silver,P.A., Cherepanov,P. and Engelman,A. (2004) Class II integrase mutants with changes in putative nuclear localization signals are primarily blocked at a postnuclear entry step of human immunodeficiency virus type 1 replication. *J. Virol.*, **78**, 12735–12746.

39. Yamashita,M. and Emerman,M. (2005) The cell cycle independence of HIV infections is not determined by known karyophilic viral elements. *PLoS Pathog.*, **1**, e18.

40. Rivière,L., Darlix,J.L. and Cimarelli,A. (2010) Analysis of the viral elements required in the nuclear import of HIV-1 DNA. *J. Virol.*, **84**, 729–739.

41. Zennou,V., Petit,C., Guetard,D., Nerhbass,U., Montagnier,L. and Charneau,P. (2000) HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell*, **101**, 173–185.

42. Dvorin,J.D., Bell,P., Maul,G.G., Yamashita,M., Emerman,M. and Malim,M.H. (2002) Reassessment of the roles of integrase and the central DNA flap in human immunodeficiency virus type 1 nuclear import. *J. Virol.*, **76**, 12087–12096.

43. Limón,A., Nakajima,N., Lu,R., Ghory,H.Z. and Engelman,A. (2002) Wild-type levels of nuclear localization and human immunodeficiency virus type 1 replication in the absence of the central DNA flap. *J. Virol.*, **76**, 12078–12086.

44. Marsden,M.D. and Zack,J.A. (2007) Human immunodeficiency virus bearing a disrupted central DNA flap is pathogenic in vivo. *J. Virol.*, **81**, 6146–6150.

45. Iglesias,C., Ringeard,M., Di Nunzio,F., Fernandez,J., Gaudin,R., Souque,P., Charneau,P. and Arhel,N. (2011) Residual HIV-1 DNA Flap-independent nuclear import of cPPT/CTS double mutant viruses does not support spreading infection. *Retrovirology*, **8**, 92.

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

47. Katz,R.A., Greger,J.G., Boimel,P. and Skalka,A.M. (2003) Human immunodeficiency virus type 1 DNA nuclear import and integration are mitosis independent in cycling cells. *J. Virol.*, **77**, 13412–13417.

48. Pornillos,O., Ganser-Pornillos,B.K., Kelly,B.N., Hua,Y., Whitby,F.G., Stout,C.D., Sundquist,W.I., Hill,C.P. and Yeager,M. (2009) X-ray structures of the hexameric building block of the HIV capsid. *Cell*, **137**, 1282–1292.

49. Pornillos,O., Ganser-Pornillos,B.K. and Yeager,M. (2011) Atomic-level modelling of the HIV capsid. *Nature*, **469**, 424–427.

50. Byeon,I.J., Meng,X., Jung,J., Zhao,G., Yang,R., Ahn,J., Shi,J., Concel,J., Aiken,C., Zhang,P. *et al.* (2009) Structural convergence between Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell*, **139**, 780–790.

51. Mattei,S., Glass,B., Hagen,W.J., Kräusslich,H.G. and Briggs,J.A. (2016) The structure and flexibility of conical HIV-1 capsids determined within intact virions. *Science*, **354**, 1434–1437.

52. Fernandez,J., Machado,A.K., Lyonnais,S., Chamontin,C., Gärtner,K., Léger,T., Henriquet,C., Garcia,C., Portilho,D.M., Pugnière,M. *et al.* (2019) Transportin-1 binds to the HIV-1 capsid via a nuclear localization signal and triggers uncoating. *Nat. Microbiol.*, **4**, 1840–1850.

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

54. Matreyek,K.A., Yucel,S.S., Li,X. and Engelman,A. (2013) Nucleoporin NUP153 phenylalanine-glycine motifs engage a common binding pocket within the HIV-1 capsid protein to mediate lentiviral infectivity. *PLoS Pathog.*, **9**, e1003693.

55. 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**, e00648-18.

56. Kane,M., Rebensburg,S.V., Takata,M.A., Zang,T.M., Yamashita,M., Kvaratskhelia,M. and Bieniasz,P.D. (2018) Nuclear pore heterogeneity influences HIV-1 infection and the antiviral activity of MX2. *Elife*, **7**, e35738.

57. Luban,J., Bossolt,K.L., Franke,E.K., Kalpana,G.V. and Goff,S.P. (1993) Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B. *Cell*, **73**, 1067–1078.

58. De Iaco,A. and Luban,J. (2014) Cyclophilin A promotes HIV-1 reverse transcription but its effect on transduction correlates best with its effect on nuclear entry of viral cDNA. *Retrovirology*, **11**, 11.

59. Lee,K., Ambrose,Z., Martin,T.D., Oztop,I., Mulky,A., Julias,J.G., Vandegraaff,N., Baumann,J.G., Wang,R., Yuen,W. *et al.* (2010) Flexible use of nuclear import pathways by HIV-1. *Cell Host Microbe*, **7**, 221–233.

60. Chin,C.R., Perreira,J.M., Savidis,G., Portmann,J.M., Aker,A.M., Feeley,E.M., Smith,M.C. and Brass,A.L. (2015) Direct visualization of HIV-1 replication intermediates shows that capsid and CPSF6 modulate HIV-1 intra-nuclear invasion and integration. *Cell Rep.*, **13**, 1717–1731.

61. Bejarano,D.A., Peng,K., Laketa,V., Börner,K., Jost,K.L., Lucic,B., Glass,B., Lusic,M., Müller,B. and Kräusslich,H.-G. (2019) HIV-1 nuclear import in macrophages is regulated by CPSF6-capsid interactions at the nuclear pore complex. *Elife*, **8**, e41800.

62. 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.

63. Gamble,T.R., Vajdos,F.F., Yoo,S., Worthylake,D.K., Houseweart,M., Sundquist,W.I. and Hill,C.P. (1996) Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell*, **87**, 1285–1294.

64. Lin,D.H., Zimmermann,S., Stuwe,T., Stuwe,E. and Hoelz,A. (2013) Structural and functional analysis of the C-terminal domain of Nup358/RanBP2. *J. Mol. Biol.*, **425**, 1318–1329.

65. Price,A.J., Fletcher,A.J., Schaller,T., Elliott,T., Lee,K., KewalRamani,V.N., Chin,J.W., Towers,G.J. and James,L.C. (2012) CPSF6 defines a conserved capsid interface that modulates HIV-1 replication. *PLoS Pathog.*, **8**, e1002896.

66. Bhattacharya,A., Alam,S.L., Fricke,T., Zadrozny,K., Sedzicki,J., Taylor,A.B., Demeler,B., Pornillos,O., Ganser-Pornillos,B.K., Diaz-Griffero,F. *et al.* (2014) Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 18625–18630.

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

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

69. Selyutina,A., Persaud,M., Lee,K., KewalRamani,V. and Diaz-Griffero,F. (2020) Nuclear import of the HIV-1 core precedes reverse transcription and uncoating. *Cell Rep.*, **32**, 108201.

70. Zila,V., Margiotta,E., Turonova,B., Müller,T.G., Zimmerli,C.E., Mattei,S., Allegretti,M., Börner,K., Rada,J., Müller,B. *et al.* (2020) Cone-shaped HIV-1 capsids are transported through intact nuclear pores. bioRxiv doi: https://doi.org/10.1101/2020.07.30.193524, 30 July 2020, preprint: not peer reviewed.

71. Galvis,A.E., Fisher,H.E., Nitta,T., Fan,H. and Camerini,D. (2014) Impairment of HIV-1 cDNA synthesis by DBR1 knockdown. *J. Virol.*, **88**, 7054–7069.

72. 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. *Nat. Microbiol.*, **5**, 1088–1095.

73. DiNunzio,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.

74. Dharan,A., Talley,S., Tripathi,A., Mamede,J.I., Majetschak,M., Hope,T.J. and Campbell,E.M. (2016) KIF5B and Nup358 cooperatively mediate the nuclear import of HIV-1 during infection. *PLoS Pathog.*, **12**, e1005700.

75. Meehan,A.M., Saenz,D.T., Guevera,R., Morrison,J.H., Peretz,M., Fadel,H.J., Hamada,M., van Deursen,J. and Poeschla,E.M. (2014) A cyclophilin homology domain-independent role for Nup358 in HIV-1 infection. *PLoS Pathog.*, **10**, e1003969.

76. Briggs,J.A., Wilk,T., Welker,R., Kräusslich,H.G. and Fuller,S.D. (2003) Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J.*, **22**, 1707–1715.

77. Arhel,N.J., Souquere-Besse,S., Munier,S., Souque,P., Guadagnini,S., Rutherford,S., Prévost,M.C., Allen,T.D. and Charneau,P. (2007) HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *EMBO J.*, **26**, 3025–3037.

78. 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**, e00135-20.

79. Knockenhauer,K.E. and Schwartz,T.U. (2016) The nuclear pore complex as a flexible and dynamic gate. *Cell*, **164**, 1162–1171.

80. Stanley,G.J., Fassati,A. and Hoogenboom,B.W. (2018) Atomic force microscopy reveals structural variability amongst nuclear pore complexes. *Life Sci. Alliance*, **1**, e201800142.

81. Ao,Z., Fowke,K.R., Cohen,E.A. and Yao,X. (2005) Contribution of the C-terminal tri-lysine regions of human immunodeficiency virus type 1 integrase for efficient reverse transcription and viral DNA nuclear import. *Retrovirology*, **2**, 62.

82. Hearps,A.C. and Jans,D.A. (2006) HIV-1 integrase is capable of targeting DNA to the nucleus via an importin alpha/beta-dependent mechanism. *Biochem. J.*, **398**, 475–484.

83. Christ,F., Thys,W., De Rijck,J., Gijsbers,R., Albanese,A., Arosio,D., Emiliani,S., Rain,J.C., Benarous,R., Cereseto,A. *et al.* (2008) Transportin-SR2 imports HIV into the nucleus. *Curr. Biol.*, **18**, 1192–1202.

84. Ao,Z., Danappa Jayappa,K., Wang,B., Zheng,Y., Kung,S., Rassart,E., Depping,R., Kohler,M., Cohen,E.A. and Yao,X. (2010) Importin alpha3 interacts with HIV-1 integrase and contributes to HIV-1 nuclear import and replication. *J. Virol.*, **84**, 8650–8663.

85. Levin,A., Hayouka,Z., Friedler,A. and Loyter,A. (2010) Transportin 3 and importin α are required for effective nuclear import of HIV-1 integrase in virus-infected cells. *Nucleus*, **1**, 422–431.

86. De Houwer,S., Demeulemeester,J., Thys,W., Rocha,S., Dirix,L., Gijsbers,R., Christ,F. and Debyser,Z. (2014) The HIV-1 integrase mutant R263A/K264A is 2-fold defective for TRN-SR2 binding and viral nuclear import. *J. Biol. Chem.*, **289**, 25351–25361.

87. Engelman,A.N. (2019) Multifaceted HIV integrase functionalities and therapeutic strategies for their inhibition. *J. Biol. Chem.*, **294**, 15137–15157.

88. Ao,Z., Jayappa,K.D., Wang,B., Zheng,Y., Wang,X., Peng,J. and Yao,X. (2012) Contribution of host nucleoporin 62 in HIV-1 integrase chromatin association and viral DNA integration. *J. Biol. Chem.*, **287**, 10544–10555.

89. Woodward,C.L., Prakobwanakit,S., Mosessian,S. and Chow,S.A. (2009) Integrase interacts with nucleoporin NUP153 to mediate the nuclear import of human immunodeficiency virus type 1. *J. Virol.*, **83**, 6522–6533.

90. Krishnan,L., Matreyek,K.A., Oztop,I., Lee,K., Tipper,C.H., Li,X., Dar,M.J., Kewalramani,V.N. and Engelman,A. (2010) The requirement for cellular transportin 3 (TNPO3 or TRN-SR2) during infection maps to human immunodeficiency virus type 1 capsid and not integrase. *J. Virol.*, **84**, 397–406.

91. Cribier,A., Ségéral,E., Delelis,O., Parissi,V., Simon,A., Ruff,M., Benarous,R. and Emiliani,S. (2011) Mutations affecting interaction of integrase with TNPO3 do not prevent HIV-1 cDNA nuclear import. *Retrovirology*, **8**, 104.

92. Maertens,G.N., Cook,N.J., Wang,W., Hare,S., Gupta,S.S., Öztop,I., Lee,K., Pye,V.E., Cosnefroy,O., Snijders,A.P. *et al.* (2014) Structural basis for nuclear import of splicing factors by human Transportin 3. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 2728–2733.

93. Brass,A.L., Dykxhoorn,D.M., Benita,Y., Yan,N., Engelman,A., Xavier,R.J., Lieberman,J. and Elledge,S.J. (2008) Identification of host proteins required for HIV infection through a functional genomic screen. *Science*, **319**, 921–926.

94. Zhou,L., Sokolskaja,E., Jolly,C., James,W., Cowley,S.A. and Fassati,A. (2011) Transportin 3 promotes a nuclear maturation step required for efficient HIV-1 integration. *PLoS Pathog.*, **7**, e1002194.

95. De Iaco,A., Santoni,F., Vannier,A., Guipponi,M., Antonarakis,S. and Luban,J. (2013) TNPO3 protects HIV-1 replication from

CPSF6-mediated capsid stabilization in the host cell cytoplasm. *Retrovirology*, **10**, 20.

96. Fricke,T., Valle-Casuso,J.C., White,T.E., Brandariz-Nuñez,A., Bosche,W.J., Reszka,N., Gorelick,R. and Diaz-Griffero,F. (2013) The ability of TNPO3-depleted cells to inhibit HIV-1 infection requires CPSF6. *Retrovirology*, **10**, 46.

97. Jang,S., Cook,N.J., Pye,V.E., Bedwell,G.J., Dudek,A.M., Singh,P.K., Cherepanov,P. and Engelman,A.N. (2019) Differential role for phosphorylation in alternative polyadenylation function versus nuclear import of SR-like protein CPSF6. *Nucleic Acids Res.*, **47**, 4663–4683.

98. Demeulemeester,J., Blokken,J., De Houwer,S., Dirix,L., Klaassen,H., Marchand,A., Chaltin,P., Christ,F. and Debyser,Z. (2018) Inhibitors of the integrase-transportin-SR2 interaction block HIV nuclear import. *Retrovirology*, **15**, 5.

99. Wagstaff,K.M., Headey,S., Telwatte,S., Tyssen,D., Hearps,A.C., Thomas,D.R., Tachedjian,G. and Jans,D.A. (2019) Molecular dissection of an inhibitor targeting the HIV integrase dependent preintegration complex nuclear import. *Cell. Microbiol.*, **21**, e12953.

100. Fricke,T., Buffone,C., Opp,S., Valle-Casuso,J. and Diaz-Griffero,F. (2014) BI-2 destabilizes HIV-1 cores during infection and Prevents Binding of CPSF6 to the HIV-1 Capsid. *Retrovirology*, **11**, 120.

101. Peng,K., Muranyi,W., Glass,B., Laketa,V., Yant,S.R., Tsai,L., Cihlar,T., Müller,B. and Kräusslich,H.G. (2014) Quantitative microscopy of functional HIV post-entry complexes reveals association of replication with the viral capsid. *Elife*, **3**, e04114.

102. Bester,S., Wei,G., Zhao,H., Adu-Ampratwum,D., Iqbal,N., Courouble,V.V., Francis,A.C., Annamalai,A.S., Singh,P.K., Shkriabai,N. *et al.* (2020) Structural and mechanistic bases for a potent HIV-1 capsid inhibitor. *Science*, **370**, 360–364.

103. Engelman,A.N. and Maertens,G.N. (2018) Virus-host interactions in retrovirus integration. In: Parent,L.J. (ed). *Retrovirus-Cell Interactions*. Academic Press, San Diego, CA, pp. 163–198.

104. Lander,E.S., Linton,L.M., Birren,B., Nusbaum,C., Zody,M.C., Baldwin,J., Devon,K., Dewar,K., Doyle,M., FitzHugh,W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.

105. 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.

106. Chen,Y. and Belmont,A.S. (2019) Genome organization around nuclear speckles. *Curr. Opin. Genet. Dev.*, **55**, 91–99.

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

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

109. Stultz,R.D., Cenker,J.J. and McDonald,D. (2017) Imaging HIV-1 genomic DNA from entry through productive infection. *J. Virol.*, **91**, e00034.

110. Achuthan,V., Perreira,J.M., Ahn,J.J., Brass,A.L. and Engelman,A.N. (2019) Capsid-CPSF6 interaction: Master regulator of nuclear HIV-1 positioning and integration. *J. Life Sci. (Westlake Village)*, **1**, 39–45.

111. Tsopoulidis,N., Kaw,S., Laketa,V., Kutscheidt,S., Baarlink,C., Stolp,B., Grosse,R. and Fackler,O.T. (2019) T cell receptor-triggered nuclear actin network formation drives CD4(+) T cell effector functions. *Sci. Immunol.*, **4**, eaav1987.

112. Bediaga,N.G., Coughlan,H.D., Johanson,T.M., Garnham,A.L., Naselli,G., Schröder,J., Fearnley,L.G., Bandala-Sanchez,E., Allan,R.S., Smyth,G.K. *et al.* (2020) Activation-induced re-organization of chromatin in human T cells. bioRxiv doi: http://dx.doi.org/10.1101/2020.06.08.135020, 09 June 2020, preprint: not peer reviewed.

113. Zhou,Y., Zhang,H., Siliciano,J.D. and Siliciano,R.F. (2005) Kinetics of human immunodeficiency virus type 1 decay following entry into resting CD4+ T cells. *J. Virol.*, **79**, 2199–2210.

114. Agosto,L.M., Yu,J.J., Dai,J., Kaletsky,R., Monie,D. and O'Doherty,U. (2007) HIV-1 integrates into resting CD4+ T cells even at low inoculums as demonstrated with an improved assay for HIV-1 integration. *Virology*, **368**, 60–72.

115. Dai,J., Agosto,L.M., Baytop,C., Yu,J.J., Pace,M.J., Liszewski,M.K. and O'Doherty,U. (2009) Human immunodeficiency virus integrates directly into naïve resting CD4+ T cells but enters naïve cells less efficiently than memory cells. *J. Virol.*, **83**, 4528–4537.

116. Brady,T., Agosto,L.M., Malani,N., Berry,C.C., O'Doherty,U. and Bushman,F. (2009) HIV integration site distributions in resting and activated CD4+ T cells infected in culture. *AIDS*, **23**, 1461–1471.

117. Peng,Y. and Zhang,Y. (2018) Enhancer and super-enhancer: positive regulators in gene transcription. *Anim. Model Exp. Med.*, **1**, 169–179.

118. Wu,X., Li,Y., Crise,B. and Burgess,S.M. (2003) Transcription start regions in the human genome are favored targets for MLV integration. *Science*, **300**, 1749–1751.

119. Mitchell,R.S., Beitzel,B.F., Schroder,A.R., Shinn,P., Chen,H., Berry,C.C., Ecker,J.R. and Bushman,F.D. (2004) Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol.*, **2**, E234.

120. Engelman,A. (2005) The ups and downs of gene expression and retroviral DNA integration. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 1275–1276.

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

122. Llano,M., Vanegas,M., Fregoso,O., Saenz,D., Chung,S., Peretz,M. and Poeschla,E.M. (2004) LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes. *J. Virol.*, **78**, 9524–9537.

123. Busschots,K., Vercammen,J., Emiliani,S., Benarous,R., Engelborghs,Y., Christ,F. and Debyser,Z. (2005) The interaction of LEDGF/p75 with integrase is lentivirus-specific and promotes DNA binding. *J. Biol. Chem.*, **280**, 17841–17847.

124. Cherepanov,P. (2007) LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. *Nucleic Acids Res.*, **35**, 113–124.

125. Llano,M., Vanegas,M., Hutchins,N., Thompson,D., Delgado,S. and Poeschla,E.M. (2006) Identification and characterization of the chromatin-binding domains of the HIV-1 integrase interactor LEDGF/p75. *J. Mol. Biol.*, **360**, 760–773.

126. Turlure,F., Maertens,G., Rahman,S., Cherepanov,P. and Engelman,A. (2006) A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo. *Nucleic Acids Res.*, **34**, 1653–1665.

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

128. Cherepanov,P., Devroe,E., Silver,P.A. and Engelman,A. (2004) Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. *J. Biol. Chem.*, **279**, 48883–48892.

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

130. Pandey,K.K., Sinha,S. and Grandgenett,D.P. (2007) Transcriptional coactivator LEDGF/p75 modulates human immunodeficiency virus type 1 integrase-mediated concerted integration. *J. Virol.*, **81**, 3969–3979.

131. Hare,S., Shun,M.C., Gupta,S.S., Valkov,E., Engelman,A. and Cherepanov,P. (2009) A novel co-crystal structure affords the design of gain-of-function lentiviral integrase mutants in the presence of modified PSIP1/LEDGF/p75. *PLoS Pathog.*, **5**, e1000259.

132. Maertens,G., Cherepanov,P., Pluymers,W., Busschots,K., De Clercq,E., Debyser,Z. and Engelborghs,Y. (2003) LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *J. Biol. Chem.*, **278**, 33528–33539.

133. Eidahl,J.O., Crowe,B.L., North,J.A., McKee,C.J., Shkriabai,N., Feng,L., Plumb,M., Graham,R.L., Gorelick,R.J., Hess,S. *et al.* (2013) Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. *Nucleic Acids Res.*, **41**, 3924–3936.

134. van Nuland,R., van Schaik,F.M., Simonis,M., van Heesch,S., Cuppen,E., Boelens,R., Timmers,H.M. and van Ingen,H. (2013) Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain. *Epigenet. Chromatin*, **6**, 12.

135. Wang,H., Farnung,L., Dienemann,C. and Cramer,P. (2020) Structure of H3K36-methylated nucleosome-PWWP complex reveals multivalent cross-gyre binding. *Nat. Struct. Mol. Biol.*, **27**, 8–13.

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

137. LeRoy,G., Oksuz,O., Descostes,N., Aoi,Y., Ganai,R.A., Kara,H.O., Yu,J.R., Lee,C.H., Stafford,J., Shilatifard,A. *et al.* (2019) LEDGF and HDGF2 relieve the nucleosome-induced barrier to transcription in differentiated cells. *Sci. Adv.*, **5**, eaay3068.

138. Yokoyama,A. and Cleary,M.L. (2008) Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell*, **14**, 36–46.

139. Tesina,P., Čermáková,K., Hořejší,M., Procházková,K., Fábry,M., Sharma,S., Christ,F., Demeulemeester,J., Debyser,Z., Rijck,J.D. *et al.* (2015) Multiple cellular proteins interact with LEDGF/p75 through a conserved unstructured consensus motif. *Nat. Commun.*, **6**, 7968.

140. El Ashkar,S., Schwaller,J., Pieters,T., Goossens,S., Demeulemeester,J., Christ,F., Van Belle,S., Juge,S., Boeckx,N., Engelman,A. *et al.* (2018) LEDGF/p75 is dispensable for hematopoiesis but essential for MLL-rearranged leukemogenesis. *Blood*, **131**, 95–107.

141. Marshall,H.M., Ronen,K., Berry,C., Llano,M., Sutherland,H., Saenz,D., Bickmore,W., Poeschla,E. and Bushman,F.D. (2007) Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. *PLoS One*, **2**, e1340.

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

143. Schrijvers,R., De Rijck,J., Demeulemeester,J., Adachi,N., Vets,S., Ronen,K., Christ,F., Bushman,F.D., Debyser,Z. and Gijsbers,R. (2012) LEDGF/p75-independent HIV-1 replication demonstrates a role for HRP-2 and remains sensitive to inhibition by LEDGINs. *PLoS Pathog.*, **8**, e1002558.

144. Vanegas,M., Llano,M., Delgado,S., Thompson,D., Peretz,M. and Poeschla,E. (2005) Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering. *J. Cell Sci.*, **118**, 1733–1743.

145. Schrijvers,R., Vets,S., De Rijck,J., Malani,N., Bushman,F.D., Debyser,Z. and Gijsbers,R. (2012) HRP-2 determines HIV-1 integration site selection in LEDGF/p75 depleted cells. *Retrovirology*, **9**, 84.

146. Wang,H., Jurado,K.A., Wu,X., Shun,M.C., Li,X., Ferris,A.L., Smith,S.J., Patel,P.A., Fuchs,J.R., Cherepanov,P. *et al.* (2012) HRP2 determines the efficiency and specificity of HIV-1 integration in LEDGF/p75 knockout cells but does not contribute to the antiviral activity of a potent LEDGF/p75-binding site integrase inhibitor. *Nucleic Acids Res.*, **40**, 11518–11530.

147. Singh,P.K., Plumb,M.R., Ferris,A.L., Iben,J.R., Wu,X., Fadel,H.J., Luke,B.T., Esnault,C., Poeschla,E.M., Hughes,S.H. *et al.* (2015) LEDGF/p75 interacts with mRNA splicing factors and targets HIV-1 integration to highly spliced genes. *Genes Dev.*, **29**, 2287–2297.

148. Kalpana,G., Marmon,S., Wang,W., Crabtree,G. and Goff,S. (1994) Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. *Science*, **266**, 2002–2006.

149. Hodges,C., Kirkland,J.G. and Crabtree,G.R. (2016) The many roles of BAF (mSWI/SNF) and PBAF complexes in cancer. *Cold Spring Harb. Perspect. Med.*, **6**, a026930.

150. Cereseto,A., Manganaro,L., Gutierrez,M.I., Terreni,M., Fittipaldi,A., Lusic,M., Marcello,A. and Giacca,M. (2005) Acetylation of HIV-1 integrase by p300 regulates viral integration. *EMBO J.*, **24**, 3070–3081.

151. Koh,Y., Wu,X., Ferris,A.L., Matreyek,K.A., Smith,S.J., Lee,K., KewalRamani,V.N., Hughes,S.H. and Engelman,A. (2013) Differential effects of human immunodeficiency virus type 1 capsid and cellular factors Nucleoporin 153 and LEDGF/p75 on the efficiency and specificity of viral DNA integration. *J. Virol.*, **87**, 648–658.

152. Saito,A., Henning,M.S., Serrao,E., Dubose,B.N., Teng,S., Huang,J., Li,X., Saito,N., Roy,S.P., Siddiqui,M.A. *et al.* (2016) Capsid-CPSF6 interaction is dispensable for HIV-1 replication in primary cells but is selected during virus passage in vivo. *J. Virol.*, **90**, 6918–6935.

153. Franke,E.K., Yuan,H.E. and Luban,J. (1994) Specific incorporation of cyclophilin A into HIV-1 virions. *Nature*, **372**, 359–362.

154. Wiegers,K., Rutter,G., Schubert,U., Grättinger,M. and Kräusslich,H.G. (1999) Cyclophilin A incorporation is not required for human immunodeficiency virus type 1 particle maturation and does not destabilize the mature capsid. *Virology*, **257**, 261–274.

155. Francis,A.C. and Melikyan,G.B. (2018) Single HIV-1 imaging reveals progression of infection through CA-dependent steps of docking at the nuclear pore, uncoating, and nuclear transport. *Cell Host Microbe*, **23**, 536–548.

156. Zila,V., Müller,T.G., Laketa,V., Müller,B. and Kräusslich,H.G. (2019) Analysis of CA content and CPSF6 dependence of early HIV-1 replication complexes in SupT1-R5 cells. *mBio*, **10**, e02501-19.

157. Li,W., Singh,P.K., Sowd,G.A., Bedwell,G.J., Jang,S., Achuthan,V., Oleru,A.V., Wong,D., Fadel,H.J., Lee,K. *et al.* (2020) CPSF6-dependent targeting of speckle-associated domains distinguishes primate from non-primate lentiviral integration. *mBio*, **11**, e02254-20.

158. Tian,B. and Manley,J.L. (2017) Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol. Cell. Biol.*, **18**, 18–30.

159. Rüegsegger,U., Beyer,K. and Keller,W. (1996) Purification and characterization of human cleavage factor Im involved in the 3′ end processing of messenger RNA precursors. *J. Biol. Chem.*, **271**, 6107–6113.

160. Dettwiler,S., Aringhieri,C., Cardinale,S., Keller,W. and Barabino,S.M. (2004) Distinct sequence motifs within the 68-kDa subunit of cleavage factor Im mediate RNA binding, protein-protein interactions, and subcellular localization. *J. Biol. Chem.*, **279**, 35788–35797.

161. Lee,K., Mulky,A., Yuen,W., Martin,T.D., Meyerson,N.R., Choi,L., Yu,H., Sawyer,S.L. and KewalRamani,V.N. (2012) HIV-1 capsid-targeting domain of cleavage and polyadenylation specificity factor 6. *J. Virol.*, **86**, 3851–3860.

162. Rasheedi,S., Shun,M.C., Serrao,E., Sowd,G.A., Qian,J., Hao,C., Dasgupta,T., Engelman,A.N. and Skowronski,J. (2016) The cleavage and polyadenylation specificity factor 6 (CPSF6) subunit of the capsid-recruited pre-messenger RNA cleavage factor I (CFIm) complex mediates HIV-1 integration into genes. *J. Biol. Chem.*, **291**, 11809–11819.

163. Yang,Q., Coseno,M., Gilmartin,G.M. and Doublié,S. (2011) Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping. *Structure*, **19**, 368–377.

164. Greig,J.A., Nguyen,T.A., Lee,M., Holehouse,A.S., Posey,A.E., Pappu,R.V. and Jedd,G. (2020) Arginine-enriched mixed-charge domains provide cohesion for nuclear speckle condensation. *Mol. Cell*, **77**, 1237–1250.

165. Llano,M., Saenz,D.T., Meehan,A., Wongthida,P., Peretz,M., Walker,W.H., Teo,W. and Poeschla,E.M. (2006) An essential role for LEDGF/p75 in HIV integration. *Science*, **314**, 461–464.

166. Selyutina,A., Persaud,M., Simons,L.M., Bulnes-Ramos,A., Buffone,C., Martinez-Lopez,A., Scoca,V., Di Nunzio,F., Hiatt,J., Marson,A. *et al.* (2020) Cyclophilin A prevents HIV-1 restriction in lymphocytes by blocking human TRIM5alpha binding to the viral core. *Cell Rep.*, **30**, 3766–3777.

167. Ni,T., Gerard,S., Zhao,G., Dent,K., Ning,J., Zhou,J., Shi,J., Anderson-Daniels,J., Li,W., Jang,S. *et al.* (2020) Intrinsic curvature of the HIV-1 CA hexamer underlies capsid topology and interaction with cyclophilin A. *Nat. Struct. Mol. Biol.*, **27**, 855–862.

168. Ptak,C. and Wozniak,R.W. (2016) Nucleoporins and chromatin metabolism. *Curr. Opin. Cell Biol.*, **40**, 153–160.

169. Kuhn,T.M., Pascual-Garcia,P., Gozalo,A., Little,S.C. and Capelson,M. (2019) Chromatin targeting of nuclear pore proteins induces chromatin decondensation. *J. Cell. Biol.*, **218**, 2945–2961.

170. Kadota,S., Ou,J., Shi,Y., Lee,J.T., Sun,J. and Yildirim,E. (2020) Nucleoporin 153 links nuclear pore complex to chromatin architecture by mediating CTCF and cohesin binding. *Nat. Commun.*, **11**, 2606.

171. Pruss,D., Bushman,F.D. and Wolffe,A.P. (1994) Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 5913–5917.

172. Pruss,D., Reeves,R., Bushman,F.D. and Wolffe,A.P. (1994) The influence of DNA and nucleosome structure on integration events directed by HIV integrase. *J. Biol. Chem.*, **269**, 25031–25041.

173. Benleulmi,M.S., Matysiak,J., Robert,X., Miskey,C., Mauro,E., Lapaillerie,D., Lesbats,P., Chaignepain,S., Henriquez,D.R., Calmels,C. *et al.* (2017) Modulation of the functional association between the HIV-1 intasome and the nucleosome by histone amino-terminal tails. *Retrovirology*, **14**, 54.

174. Mauro,E., Lesbats,P., Lapaillerie,D., Chaignepain,S., Maillot,B., Oladosu,O., Robert,X., Fiorini,F., Kieffer,B., Bouaziz,S. *et al.* (2019) Human H4 tail stimulates HIV-1 integration through binding to the carboxy-terminal domain of integrase. *Nucleic Acids Res.*, **47**, 3607–3618.

175. Pasi,M., Mornico,D., Volant,S., Juchet,A., Batisse,J., Bouchier,C., Parissi,V., Ruff,M., Lavery,R. and Lavigne,M. (2016) DNA minicircles clarify the specific role of DNA structure on retroviral integration. *Nucleic Acids Res.*, **44**, 7830–7847.

176. Maertens,G.N., Hare,S. and Cherepanov,P. (2010) The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature*, **468**, 326–329.

177. Eisele,E. and Siliciano,R.F. (2012) Redefining the viral reservoirs that prevent HIV-1 eradication. *Immunity*, **37**, 377–388.

178. Coffin,J. and Swanstrom,R. (2013) HIV pathogenesis: Dynamics and genetics of viral populations and infected cells. *Cold Spring Harb. Perspect. Med.*, **3**, a012526.

179. Wei,X., Ghosh,S.K., Taylor,M.E., Johnson,V.A., Emini,E.A., Deutsch,P., Lifson,J.D., Bonhoeffer,S., Nowak,M.A., Hahn,B.H. *et al.* (1995) Viral dynamics in human immunodeficiency virus type 1 infection. *Nature*, **373**, 117–122.

180. Ho,D.D., Neumann,A.U., Perelson,A.S., Chen,W., Leonard,J.M. and Markowitz,M. (1995) Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*, **373**, 123–126.

181. Coffin,J. (1995) HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science*, **267**, 483–489.

182. Chun,T.W., Stuyver,L., Mizell,S.B., Ehler,L.A., Mican,J.A.M., Baseler,M., Lloyd,A.L., Nowak,M.A. and Fauci,A.S. (1997) Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 13193–13197.

183. Chun,T.-W., Carruth,L., Finzi,D., Shen,X., DiGiuseppe,J.A., Taylor,H., Hermankova,M., Chadwick,K., Margolick,J., Quinn,T.C. *et al.* (1997) Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature*, **387**, 183–188.

184. Finzi,D., M.,H., Pierson,T., Carruth,L.M., Buck,C., Chaisson,R.E., Quinn,T.C., Chadwick,K., Margolick,J., Brookmeyer,R. *et al.* (1997) Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science*, **278**, 1295–1300.

185. Wong,J.K., Hezareh,M., Günthard,H.F., Havlir,D.V., Ignacio,C.C., Spina,C.A. and Richman,D.D. (1997) Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science*, **278**, 1291–1295.

186. Chun,T.W., Engel,D., Berrey,M.M., Shea,T., Corey,L. and Fauci,A.S. (1998) Early establishment of a pool of latently infected, resting CD4+ T cells during primary HIV-1 infection. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 8869–8873.

187. Chavez,L., Calvanese,V. and Verdin,E. (2015) HIV latency is established directly and early in both resting and activated primary CD4 T cells. *PLoS Pathog.*, **11**, e1004955.

188. Finzi,D., Blankson,J., Siliciano,J.D., Margolick,J.B., Chadwick,K., Pierson,T., Smith,K., Lisziewicz,J., Lori,F., Flexner,C. *et al.* (1999) Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat. Med.*, **5**, 512–517.

189. Ho,Y.-C., Shan,L., Hosmane,N.N., Wang,J., Laskey,S.B., Rosenbloom,D.I.S., Lai,J., Blankson,J.N., Siliciano,J.D. and Siliciano,R.F. (2013) Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell*, **155**, 540–551.

190. Pace,M.J., Graf,E.H., Agosto,L.M., Mexas,A.M., Male,F., Brady,T., Bushman,F.D. and O'Doherty,U. (2012) Directly infected resting CD4+T cells can produce HIV Gag without spreading infection in a model of HIV latency. *PLoS Pathog.*, **8**, e1002818.

191. Shan,L., Deng,K., Gao,H., Xing,S., Capoferri,A.A., Durand,C.M., Rabi,S.A., Laird,G.M., Kim,M., Hosmane,N.N. *et al.* (2017) Transcriptional reprogramming during effector-to-memory transition renders CD4+ T cells permissive for latent HIV-1 infection. *Immunity*, **47**, 766–775.

192. Cohn,L.B., Silva,I.T., Oliveira,T.Y., Rosales,R.A., Parrish,E.H., Learn,G.H., Hahn,B.H., Czartoski,J.L., McElrath,M.J., Lehmann,C. *et al.* (2015) HIV-1 integration landscape during latent and active infection. *Cell*, **160**, 420–432.

193. Simonetti,F.R., Sobolewski,M.D., Fyne,E., Shao,W., Spindler,J., Hattori,J., Anderson,E.M., Watters,S.A., Hill,S., Wu,X. *et al.* (2016) Clonally expanded CD4+ T cells can produce infectious HIV-1 in vivo. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 1883–1888.

194. Bruner,K.M., Murray,A.J., Pollack,R.A., Soliman,M.G., Laskey,S.B., Capoferri,A.A., Lai,J., Strain,M.C., Lada,S.M., Hoh,R. *et al.* (2016) Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nat. Med.*, **22**, 1043–1049.

195. Wagner,T.A., McLaughlin,S., Garg,K., Cheung,C.Y.K., Larsen,B.B., Styrchak,S., Huang,H.C., Edlefsen,P.T., Mullins,J.I. and Frenkel,L.M. (2014) Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science*, **345**, 570.

196. Maldarelli,F., Wu,X., Su,L., Simonetti,F.R., Shao,W., Hill,S., Spindler,J., Ferris,A.L., Mellors,J.W., Kearney,M.F. *et al.* (2014) Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science*, **345**, 179.

197. Coffin,J.M., Wells,D.W., Zerbato,J.M., Kuruc,J.D., Guo,S., Luke,B.T., Eron,J.J., Bale,M., Spindler,J., Simonetti,F.R. *et al.* (2019) Clones of infected cells arise early in HIV-infected individuals. *JCI Insight*, **4**, e128432.

198. Lorenzi,J.C.C., Cohen,Y.Z., Cohn,L.B., Kreider,E.F., Barton,J.P., Learn,G.H., Oliveira,T., Lavine,C.L., Horwitz,J.A., Settler,A. *et al.* (2016) Paired quantitative and qualitative assessment of the replication-competent HIV-1 reservoir and comparison with integrated proviral DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, E7908–E7916.

199. Bui,J.K., Sobolewski,M.D., Keele,B.F., Spindler,J., Musick,A., Wiegand,A., Luke,B.T., Shao,W., Hughes,S.H., Coffin,J.M. *et al.* (2017) Proviruses with identical sequences comprise a large fraction of the replication-competent HIV reservoir. *PLoS Pathog.*, **13**, e1006283.

200. Hosmane,N.N., Kwon,K.J., Bruner,K.M., Capoferri,A.A., Beg,S., Rosenbloom,D.I.S., Keele,B.F., Ho,Y.-C., Siliciano,J.D. and Siliciano,R.F. (2017) Proliferation of latently infected CD4+ T cells carrying replication-competent HIV-1: Potential role in latent reservoir dynamics. *J. Exp. Med.*, **214**, 959–972.

201. Chéné,I.d., Basyuk,E., Lin,Y.-L., Triboulet,R., Knezevich,A., Chable-Bessia,C., Mettling,C., Baillat,V., Reynes,J., Corbeau,P. *et al.* (2007) Suv39H1 and HP1γ are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. *EMBO J.*, **26**, 424–435.

202. Kauder,S.E., Bosque,A., Lindqvist,A., Planelles,V. and Verdin,E. (2009) Epigenetic regulation of HIV-1 latency by cytosine methylation. *PLoS Pathog.*, **5**, e1000495.

203. Blazkova,J., Trejbalova,K., Gondois-Rey,F., Halfon,P., Philibert,P., Guiguen,A., Verdin,E., Olive,D., Van Lint,C., Hejnar,J. *et al.* (2009) CpG methylation controls reactivation of HIV from latency. *PLoS Pathog.*, **5**, e1000554.

204. Imai,K., Togami,H. and Okamoto,T. (2010) Involvement of histone H3 lysine 9 (H3K9) methyltransferase G9a in the maintenance of HIV-1 latency and its reactivation by BIX01294. *J. Biol. Chem.*, **285**, 16538–16545.

205. Friedman,J., Cho,W.K., Chu,C.K., Keedy,K.S., Archin,N.M., Margolis,D.M. and Karn,J. (2011) Epigenetic silencing of HIV-1 by

the histone H3 lysine 27 methyltransferase enhancer of Zeste 2. *J. Virol.*, **85**, 9078–9089.

206. Vansant,G., Chen,H.-C., Zorita,E., Trejbalová,K., Miklík,D., Filion,G. and Debyser,Z. (2020) The chromatin landscape at the HIV-1 provirus integration site determines viral expression. *Nucleic Acids Res.*, **48**, 7801–7817.

207. Lenasi,T., Contreras,X. and Peterlin,B.M. (2008) Transcriptional interference antagonizes proviral gene expression to promote HIV latency. *Cell Host Microbe*, **4**, 123–133.

208. Han,Y., Lin,Y.B., An,W., Xu,J., Yang,H.-C., O'Connell,K., Dordai,D., Boeke,J.D., Siliciano,J.D. and Siliciano,R.F. (2008) Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. *Cell Host Microbe*, **4**, 134–146.

209. Shan,L., Yang,H.C., Rabi,S.A., Bravo,H.C., Shroff,N.S., Irizarry,R.A., Zhang,H., Margolick,J.B., Siliciano,J.D. and Siliciano,R.F. (2011) Influence of host gene transcription level and orientation on HIV-1 latency in a primary-cell model. *J. Virol.*, **85**, 5384–5393.

210. Besnard,E., Hakre,S., Kampmann,M., Lim,H.W., Hosmane,N.N., Martin,A., Bassik,M.C., Verschueren,E., Battivelli,E., Chan,J. *et al.* (2016) The mTOR complex controls HIV latency. *Cell Host Microbe*, **20**, 785–797.

211. Chen,H.C., Martinez,J.P., Zorita,E., Meyerhans,A. and Filion,G.J. (2017) Position effects influence HIV latency reversal. *Nat. Struct. Mol. Biol.*, **24**, 47–54.

212. Jiang,C., Lian,X., Gao,C., Sun,X., Einkauf,K.B., Chevalier,J.M., Chen,S.M.Y., Hua,S., Rhee,B., Chang,K. *et al.* (2020) Distinct viral reservoirs in individuals with spontaneous control of HIV-1. *Nature*, **585**, 261–267.

213. Hughes,S.H. and Coffin,J.M. (2016) What integration sites tell us about HIV persistence. *Cell Host Microbe*, **19**, 588–598.

214. Maldarelli,F. (2016) The role of HIV integration in viral persistence: no more whistling past the proviral graveyard. *J. Clin. Invest.*, **126**, 438–447.

215. Selvaraj,A. and Prywes,R. (2003) Megakaryoblastic leukemia-1/2, a transcriptional co-activator of serum response factor, is required for skeletal myogenic differentiation. *J. Biol. Chem.*, **278**, 41977–41987.

216. Huang,D., Sumegi,J., Dal Cin,P., Reith,J.D., Yasuda,T., Nelson,M., Muirhead,D. and Bridge,J.A. (2010) C11orf95-MKL2 is the resulting fusion oncogene of t(11;16)(q13;p13) in chondroid lipoma. *Genes Chromosomes Cancer*, **49**, 810–818.

217. Muehlich,S., Hampl,V., Khalid,S., Singer,S., Frank,N., Breuhahn,K., Gudermann,T. and Prywes,R. (2012) The transcriptional coactivators megakaryoblastic leukemia 1/2 mediate the effects of loss of the tumor suppressor deleted in liver cancer 1. *Oncogene*, **31**, 3913–3923.

218. Rani,A. and Murphy,J.J. (2016) STAT5 in cancer and immunity. *J. Interferon Cytokine Res.*, **36**, 226–237.

219. Yang,L., Chen,S., Zhao,Q., Sun,Y. and Nie,H. (2019) The critical role of Bach2 in shaping the balance between CD4+ T cell subsets in immune-mediated diseases. *Mediators Inflamm.*, **2019**, 2609737.

220. Sasaki,S., Ito,E., Toki,T., Maekawa,T., Kanezaki,R., Umenai,T., Muto,A., Nagai,H., Kinoshita,T., Yamamoto,M. *et al.* (2000) Cloning and expression of human B cell-specific transcription factor BACH2 mapped to chromosome 6q15. *Oncogene*, **19**, 3739–3749.

221. Cesana,D., Santoni de Sio,F.R., Rudilosso,L., Gallina,P., Calabria,A., Beretta,S., Merelli,I., Bruzzesi,E., Passerini,L., Nozza,S. *et al.* (2017) HIV-1-mediated insertional activation of STAT5B and BACH2 trigger viral reservoir in T regulatory cells. *Nat. Commun.*, **8**, 498.

222. Yoshihara,K., Wang,Q., Torres-Garcia,W., Zheng,S., Vegesna,R., Kim,H. and Verhaak,R.G.W. (2015) The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene*, **34**, 4845–4854.