Natural Occurring Polymorphisms in HIV-1 Integrase and RNase H

Regulate Viral Release and Autoprocessing

Running title: Integrase and RNase H regulate GagPol autoprocessing

Tomozumi Imamichi, a # John G. Bernbaum, b Sylvain Laverdure, a Jun Yang, a Qian Chen, a
Helene Highbarger, c Ming Hao, a Hongyan Sui, a Robin Dewar, c Weizhong Chang, a

and

H. Clifford Lane d

a Laboratory of Human Retrovirology and Immunoinformatics, Applied and Developmental Directorate, Frederick National Laboratory, Frederick, Maryland, 21702, USA.
b Integrated Research Facility, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Frederick, Maryland, USA.
c Virus Isolation and Serology Laboratory, Applied and Developmental Directorate, Frederick National Laboratory, Frederick, Maryland, USA.
d Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA.

# Address correspondence to Tomozumi Imamichi, timamichi@mail.nih.gov.

The word count for the abstract: 220
The word count for the text: 6349
Abstract

Recently, a genome-wide association study using plasma HIV RNA from antiretroviral therapy naïve patients reported that 14 naturally occurring non-synonymous single nucleotide polymorphisms (SNPs) in HIV derived from anti-retrovirus drugs naïve patients were associated with virus load (VL). Those SNPs were detected in reverse transcriptase, RNase H, integrase, envelope, and Nef. However, the impact of each mutation on viral fitness was not investigated. Here, we constructed a series of HIV variants encoding each SNP and examined their replicative abilities. An HIV variant containing Met-to-Ile change at codon 50 in integrase (HIV(IN:M50I)) was found as an impaired virus. Despite the mutation being in integrase, the virus release was significantly suppressed (P<0.001). Transmission electron microscopy analysis revealed that abnormal bud accumulation on the plasma membrane and the released virus particles retained immature forms. Western blot analysis demonstrated a defect in autoprocessing of GagPol and Gag polyproteins' autoprocessing in the HIV(IN:M50I) particles, although Förster Resonance Energy Transfer (FRET) assay displayed that GagPol containing IN:M50I forms homodimer with a similar efficiency with GagPol (WT). The impaired maturation and replication were rescued by two other VL-associated SNPs, Ser-to-Asn change at codon 17 of integrase or Asn-to-Ser change at codon 79 of RNase H. These data demonstrate that Gag and GagPol assembly, virus release, and autoprocessing are not only regulated by integrase but also RNase H.

Importance

A nascent HIV-1 is a noninfectious viral particle. Cleaving Gag and GagPol polyproteins in the particle by mature HIV protease (PR), the nascent virus becomes an infectious virus. PR is initially translated as an inactive embedded enzyme in a GagPol polyprotein. The embedded PR
in homodimerized GagPol polyproteins catalyzes a proteolytic reaction to release the mature PR.

This excision step by a self-cleavage is called autoprocessing. Here, during the evaluation of the roles of naturally emerging non-synonymous SNPs in HIV RNA, we found that autoprocessing is inhibited by Met-to-Ile change at codon 50 in integrase GagPol. Co-existing other SNPs, Ser-to-Asn change at codon 17 in integrase or Asn-to-Ser mutation at codon 79 in RNase H, recovered this defect, suggesting that autoprocessing is regulated by not only integrase but also RNase H in GagPol polyprotein.

Introduction

Nascent human immunodeficiency virus type 1 (HIV-1) particles are released from infected host cells as immature and non-infectious viruses (1). These immature viral particles contain Gag and GagPol polyproteins, accessory proteins including Nef precursor, and viral genomic RNAs. The Gag and GagPol polyproteins are composed of viral structural proteins and the viral enzymes: protease (PR), reverse transcriptase (RT), RNase H (RH), and integrase (IN). During proteolytic maturation, the polyproteins in the immature particle are cleaved by the viral PR at 11 sites on the Gag and GagPol polyproteins and one site on the Nef precursor (1-3), and then the immature particle is converted into its infectious mature form. While these catalytic activities of PR are well-described, there are other aspects of PR that may play a role in viral replication. During viral Gag and GagPol assembly and budding at the cell membrane, GagPol polyproteins dimerize, followed by dimerization of the PR domains of the GagPol dimer, leading to the release of a functional dimerized PR in the viral particles. Inhibition of dimerization of the GagPol polyproteins has been considered a unique therapeutic target (4). Autoprocessing, which is the initial activity of releasing the embedded PR in the GagPol polyprotein by its processing, results
in the formation of mature PR (5-7) that can catalyze the other cleavage reactions. Therefore, the homodimerization of the GagPol polyproteins in the immature virion is required for this excision of PR. IN is a component of the GagPol polyprotein, and a biologic role of the integrase protein of HIV-1 has been thought to be limited to inducing integration of proviral DNA to the host cell genome following infection. However, it is demonstrated that a truncated GagPol lacking IN suppress viral release (8), and drug resistance mutations to an experimental integrase inhibitor, KF116, interferes with autoprocessing (9, 10). These reports have suggested a potential role for IN in viral release and autoprocessing.

In a recent sub-study of the Strategic Timing of Antiretroviral Therapy (START) study (11), the investigators identified a series of 14 naturally-occurring non-synonymous single nucleotide polymorphisms (SNPs) in HIV-1 that correlated with different levels of viremia in treatment-naïve patients (12). However, the specific impact of each amino acid (aa) substitution associated with the SNP on viral fitness was not investigated. In the present study, we created a series of recombinant viruses containing each of the SNPs using site-directed mutagenesis and characterized the role of each mutation in viral fitness. Of all the mutants, only the Met-to-Ile substitution at codon 50 of IN (IN:M50I) resulted in the loss of replication capability leading to the release of abnormally shaped virions and interfering with autoprocessing. Of note, this "off-target" effect was restored by compensatory mutations in IN(IN:S17N) or RH (RH:N79S) in the 14 SNPs. These results identify critical roles for RH and IN in the autoprocessing of the HIV-1 GagPol polyprotein.

Results

Impact of SNPs on Virus Replication Fitness
A total of 14 non-synonymous SNPs were reported to correlate with plasma viral loads in a cohort of 3,592 anti-retroviral drug-naïve HIV-infected patients (12). To characterize the impact of each aa substitution associated with each SNP (Table 1) on viral fitness, we constructed a series of HIV variants containing each change using site-directed mutagenesis using the cloned HIV laboratory strain HIVNL4.3. Since this strain contains an Asn-to-Ser substitution at codon 79 of RH (RH:N79S), this site was back-mutated from Ser to Asn to yield a clone equivalent to the RH of the clade B consensus sequence (RH:N79). We used this construct, pNL(RH:N79WT), as the backbone wild-type HIV, HIV(WT), as a control in the current study. Each non-synonymous SNP was introduced onto the WT control backbone using site-directed mutagenesis. Plasmids encoding each SNP were transfected into HEK293T cells, and then viral stocks were prepared as described in the Materials and Methods. Pelleted viral particles were resuspended in 1/100 volume of the starting materials and used as viral stocks. To determine the amounts of virus in each variant stock, we measured HIV p24 antigen concentrations in each stock. Almost all variants except HIV(IN:M50I) demonstrated comparable amounts of p24 concentrations with that of HIV(WT) (89.0 ± 12 μg/ml, n=7) (Supplemental Table S1). In striking contrast, the p24 concentrations of HIV(IN:M50I) stocks were 0.30 ± 0.087 μg/ml (n=7) and thus around 0.3 % of that in HIV(WT) (P<0.001). To define whether the suppression of HIV(IN:M50I) release is caused by only HEK293T cells, Hela cells were also transfected with HIV(WT) or HIV(IN:M50I) construct and then released HIV amounts were quantified. HIV(IN:M50I) virus was released by 0.3 % of HIV(WT) (Supplemental Table S1) from the cells, indicating that the inhibition is not caused in HEK293T specific manner. It is known that Hela cells produce Interferons (IFNs) in response to transfected plasmid DNA as an innate immune response (13), while HEK293T cells lack the inducing activity (14). To avoid any impact of the
produced IFNs in downstream experiments for HIV production, we used HEK293T cells in entire studies as an HIV-producing cell.

To investigate the impact of each mutation on viral replication, we infected PHA-stimulated primary CD4+ T cells from three independent healthy donors with each HIV variant, and viral replication was monitored for 14 days. All mutants except HIV(IN:M50I) replicated comparably to HIV(WT), and the HIV(IN:M50I) variant showed little to no replication (Fig. 1a – 1c, Supplemental Fig. S1a – S1d). Compared to HIV(WT), in the presence of IN:M50I mutation, HIV replication was suppressed by 99.9 ± 0.030 % (n=7, P<0.01) on the 7th day after infection, and even when the HIV(IN:M50I)-infected cells were cultured for additional 7 days, significant viral replication was not detected.

To further investigate the biologic properties of the IN:M50I mutant, we examined the morphology of the transfected HEK293T cells and virus particles 24 hours after transfection using transmission electron microscopy (TEM). The cell surface of an HIV(WT)-producing cell demonstrated buds and viral neck formation and virus release, while HIV(IN:M50I)-producing cells showed accumulation of virus buds on the cell surface (Fig. 2a and 2b), indicating that HIV(IN:M50I) release is suppressed. While the WT particles were relatively uniform in size (110 – 130 nm in diameter), the M50I particles were highly variable in size (190 – 300 nm in diameter) (Fig 2c and 2d). TEM of HIV(WT) particles revealed virions with the expected outer envelope (Fig 2e and 2f); 90-95 % having an electron-dense core (Fig. 2g). In contrast, TEM of HIV(IN:M50I) particles revealed a similar outer envelope; with no evidence of a core formation. These findings indicated that HIV(IN:M50I)-producing cells exhibited abnormalities in budding formations on the cell surface and releasing of viral particles from the cells. Unlike HIV(WT) with its dense core, the released mutant viruses had ring, doughnut-shaped, or teardrop...
structures. These results indicated that HIV(IN:M50I) was a typical defective virus, suggesting a lack of PR activity in the HIV(IN:M50I) virus particles. Gag and GagPol polyproteins play an important role in the intra-cytoplasmic trafficking of viral genomic RNAs (vgRNAs) and their eventual incorporation into the immature budding virions (1, 15-17). To determine whether or not the mutant phenotype of IN:M50I virions was due to a defect in the delivery of vgRNAs to the budding viral particles, we measured vgRNA copy numbers using purified HIV(WT) and HIV(IN:M50I) particles. No significant differences were observed in vgRNA copy number per microgram of virion protein between HIV(WT) and HIV(IN:M50I) (P>0.05, n=3) (Fig. 2h), indicating that the IN:M50I mutation had no impact on the incorporation of vgRNAs into virions.

Given the immature appearance of IN:M50I virions, we next sought to determine whether the mutation led to any changes in autoprocessing and proteolytic cleavage of the GagPol polyproteins. Purified particles were analyzed for viral proteins by Western blot (WB) using a series of antibodies specific to p24 capsid protein (CA), PR, RT, IN or Nef. In contrast to the wild-type virions, HIV(IN:M50I) virions contained uncleaved Gag and GagPol polyproteins and immature Nef and did not contain the cleaved fragments of CA, PR, RT, and IN as detected in HIV(WT) (Fig. 3a – 3e). These results indicated that the IN:M50I mutation was somehow interfering with the activity of the HIV-1 PR, likely at the initial GagPol cleavage step, autoprocessing, and suppresses the production of mature CA, PR, RT, IN, and Nef.

**Impact of IN:M50I mutation on GagPol dimerization**

The autoprocessing of GagPol polyprotein is dependent upon the formation of GagPol polyprotein homodimers (5, 18), thus we hypothesized that the IN:M50I mutation might lead to a
structural hindrance in GagPol interfering with homodimerization. It is presumed that GagPol homodimerization initiated at p2 (SP1) region of Gag domain like the dimerization of Gag polyprotein (19). To determine whether the IN:M50I mutation suppresses homodimer formation, we employed a Förster resonance energy transfer (FRET) assay using the plasmids pGag(MA/EGFP/CA), pGagPol(MA/mSB/CA), pGag(MA/mSB/CA), pGagPol(MA/EGFP/CA) (20). Since the pGag(MA/EGFP/CA) and pGagPol(MA/mSB/CA) contained RH:N79S, they were back-mutated to RH:S79N, and IN:M50I was induced in both plasmids.

We assessed the efficiency of homodimerization among Gag-to-Gag (G-G), GagPol(WT)-to-GagPol(WT) [GP(WT)-GP(WT)], and GagPol(IN:M50I)-to-GagPol(IN:M50I) [GP(M50I)-GP(M50I)] by transfection of equal amounts of each plasmid DNA. Consistent with another report (20), FRET signals from G-G (Fig. 4a) were more robust than the signals from other pairs (Fig. 4b and 4c). The efficiency of the homodimerization of GP(WT)-to-GP(WT) and the GP(IN:M50I)-to-GP(IN:M50I) were 34.9 ± 6.4 % (n=3, P<0.01) and 57 ± 4.4 % (n=3, P<0.01) of that of the G-G homodimerization, respectively (Fig. 4d). Interestingly, the efficiency of GagPol(IN:M50I) homodimerization was 1.7 ± 0.45-fold higher than that of GagPol(WT) homodimerization (n=3, P<0.05). These data indicated that the emergence of IN:M50I has no interference in the formation of GagPol homodimerization; however, it suppresses the autoprocessing. It was considered that even the IN:M50I mutation might change the distribution of GagPol proteins in the cells. To compare the distribution of each Gag, GagPol(WT), and GagPol(IN:M50) in cells, we transfected each GFP-construct into HEK293T cells and analyzed protein distribution. Gag proteins were distributed throughout cells (the cytosol and the plasma membrane) (Fig 4e). On the contrary, both GagPol(WT) and GagPol(IN:M50I) were predominantly located in the cytosol with a polarization (Fig 4f and 4g), indicating that IN:M50I
mutation has no impact on the GagPol distribution in the cells. Taken together,
GagPol(IN:M50I) distributes in a similar location with GagPol(WT) and can form a homodimer at the membrane. Therefore, it was speculated that IN:M50I may suppress the autoprocessing after the homodimerization.

HIV PR cleaves a total of 9 cleavage sites in GagPol polyprotein (Fig. 5a) (1). PR cleaves GagPol in an ordered sequence leading to virus maturation (21). Generally, it is considered that the initial cleavage as the autoprocessing occurs between the p2 spacer peptide and the NC protein (at cleavage site 3 in Fig. 5a and 5b ) in HIV(22-28) as an intramolecular reaction (cis-reaction) (5, 18, 29, 30). Since, in the presence of IN:M50I mutation, GagPol suppressed the autoprocessing even it forms homodimers, it was speculated that the mutation may change the conformation of GagPol. Crystal analysis of the full-length of GagPol even an enter length of IN was not completed due to disorder at the C terminus domain of IN and unclear in the resulting crystal structure; instead, many partial structure analyses were conducted (31-35); thus, we attempted to perform in silico Structure analysis of full-length IN using the RosettaCM protocol to predict a structure difference. We have not observed any difference in both predicted IN models (Supplemental Fig. S2), suggesting that the IN:M50I may suppress the initiation of autoprocessing via an uncharacterized manner.

When we performed WB using viral particles of HIV(IN:M50I) and polyclonal anti-PR antibodies, the antibody detected uncleaved GagPol polyproteins, but a mature PR band was not detected (Fig 3b and 5f); however, in addition to the uncleaved GagPol band, unexpected additional bands were detected (Fig 5f, Supplemental Fig. S3). The molecular sizes of the bands were ~ 67 and ~74 kDa. The detection levels of each band were varied by samples and antibody lot-dependent manners; however, a ~67 kDa band was consistently detected among assays,
implicating that, in HIV(IN:M50I), GagPol was partially digested at a lower level. GagPol autoprocessing is initiated at site 3 (5, 22, 30, 36) followed by site 5 (30) or site 1 (36). Since the 67 kDa polypeptide band was consistently detected by anti-PR antibody, we speculated that the 67 kDa band contains PR, and the polypeptide might result from the cleaved product between PR and RT at cleavage site 7 in HIV(IN:M50I) (Fig. 5b) and be composed of MA/CA/p2/NC/p1/p6*/PR (Fig. 5c). To define the component, WB using anti-MA or anti-CA antibodies was conducted. The antibodies detected the 67kDa band (Fig 5g and 5h), indicating that the 67kDa band contains MA, CA and PR, and is most likely a cleaved product at site 7. To address this hypothesis, we created mutant viruses in which cleavage site 7 in HIV(WT) and HIV(IN:M50I) was changed from Phe-Pro to Val-Pro by point mutagenesis, and the resulting viruses were designated HIV(WT_Δ7) and HIV(M50I_Δ7), respectively (Fig 5d and 5e). A virus that lacks cleavage site 7, which makes a PR and RT fusion protein (PR-RT) still possesses a functional PR activity (37). We, therefore, expected that if the IN:M50I mutation had no direct impact on PR activity, HIV (M50I_Δ7) would function as well as HIV (WT_Δ7) at GagPol processing. Those constructs were transfected into HEK293T cells, viral particles were collected, and then WB analysis was performed using those virus lysates with a polyclonal anti-PR antibody. As anticipated, the 67 kDa band was no longer present in HIV (M50I_Δ7) (Fig. 5f); instead, dominant bands at 73 kDa and 90 kDa with other minor bands were detected, which were also detected at a comparable level in HIV(WT_Δ7). WB analysis using anti-MA, -CA and -IN antibodies were also conducted, and comparable levels of mature MA-, CA- and IN-sized bands were detected in both HIV(WT_Δ7) and HIV(IN:M50I_Δ7) (Fig. 5g, 5h, and 5i). These findings indicated that the IN:M50I mutation may alter the order of the autoprocessing rather
than directly inhibiting PR function, and consequently, maturation of the released virions fails due to inhibition of the initial cleavage at cleavage site 3.

**Identification of Compensatory Mutations**

Our viral fitness results demonstrated that the IN:M50I mutation was a lethal mutation when introduced as a single change; however, since it was identified from a study of circulating virions in anti-retroviral drug-treatment-naïve patients, we postulated that the circulating viruses must also contain a compensatory mutation(s) elsewhere in the genome. It has been reported that HIVNL4.3 carrying IN:M50I mutation is replication-competent *in vitro* (38), and HIVNL4.3 exhibits RH:N79S polymorphism (39). This variant was also found among the quasi-species from the patient pool from which IN:M50I was identified (Table1). Thus, we presumed that RH:N79S might be a compensatory mutation for IN:M50I. Also, another mutation in IN, IN:S17N, was also present among the swarm of replicating viruses, suggesting that this might also be a compensatory mutation. To test these hypotheses, we constructed variants carrying a combination of IN:M50I and RH:N79S (HIV(RH:N79S/IN:M50I)), or IN:M50I and IN:S17N (HIV(IN:S17N/IN:M50I)) and analyzed their ability to replicate. Both double-mutant virions restored the ability to replicate (Fig. 6a and 6b, Supplemental Fig. S4a and S4b) and led to the formation of mature virions containing the cleaved PR and CA proteins (Fig. 6c and 6d) in WB and amounts of released virus from the producing cells were comparable to that of the wild-type (Supplementary Table S1). As anticipated, TEM demonstrated that the released viral particles of the double mutants were indistinguishable from that of HIV(WT) containing core proteins, and in the presence of the compensatory mutation, we did not detect the accumulated buds on the cell surface (Fig. 6e and 6f).
To determine the potential clinical relevance of these findings, we analyzed the population of viruses carrying these mutations using the Los Alamos HIV Sequence Database (40) since it has been observed IN:M50I in clinical studies (41, 42). Of 5100 HIV sequences, 401 (7.9 % of the total sequences) carried the IN:M50I mutation (Supplemental Table S2 and S3). Two hundred fifty-two sequences carried either RN:N79S or IN:S17N (the population of IN:M50I with RH:N79S and IN:S17N were 3.8 % and 1.1 %, respectively). Thus, additional compensatory mutations for IN:M50I may be present in the remaining 149 variants.

This population analysis also revealed several changes at IN:M50. In addition to IN:M50I, we observed IN:M50L, IN:M50R, IN:M50T, and IN:M50V mutations. To define the role of these mutations in viral fitness, we constructed variants containing each mutation (HIV(IN:M50V), HIV(IN:M50R), HIV(IN:M50L), and HIV(IN:M50T)) and assessed replication fitness. Unlike M50I, those mutants replicated at the same level as HIV (WT) (Fig. 7), highlighting the uniqueness of the M50I variant.

Population analysis

To further define whether the variants RH:N79S, IN:S17N, and IN:M50I emerge in specific HIV-1 subtypes or unique subtypes, we performed a population analysis using the Los Alamos Database. Although IN:M50I population in subtype D (1.5%) was relatively lower than other subtypes, overall, the mutation was detected in all analyzed subtypes, including CRFs (Fig 8a, Supplemental table S4); thus IN:M50I emerges without any subtype specificity. Of interest, the compensatory mutation RH:N79S and IN:S17N were not detected in Subtypes H, K, and L (Fig 8b and Supplemental Table S5) and subtype J (Fig 8c and Supplemental Table S6), respectively, suggesting that those subtypes contain other compensatory mutations. To define the frequency of
RH:N79S/IN:M50I or IN:S17N/IN:M50I in all subtypes, further population analysis was conducted. Even though the combination of RH:N79S/IN:M50I or IN:S17N/IN:M50 was detected in most subtypes, the frequency was not 100% (Fig 8d, 8e, Supplemental table S7, and S8), suggesting that the presence of other compensatory mutations against IN:M50I in all subtypes.

Discussion

Recently, a total of 14 non-synonymous SNPs correlating with different levels of viremia in treatment-naïve HIV-infected patients were reported (12). However, the specific impact of each aa substitution on viral fitness was not investigated. In the present study, we generated HIV-1 variants carrying each of the 14 SNPs to examine the impact of these mutations on viral fitness. We found that the IN:M50I mutation is associated with the loss of replicative capacity. This defect appears to be due to the inhibition of PR autoprocessing by the embedded PR in GagPol polyprotein. Furthermore, we demonstrated that compensatory mutations in IN and RH rescue the loss of replicative capacity and maturation.

It has been reported that GagPol processing is sensitive to the dynamics of viral assembly and that PR activation is required for correct polyprotein processing (43). Interestingly, our TEM images of HIV(IN:M50I) demonstrated the abnormal assembly of virions at the cell membrane. WB studies illustrated that HIV(IN:M50I) mutant lacked mature PR, implicating inhibition of autoprocessing. A series of FRET analyses illustrated that the GagPol(IN:M50I) could form homodimers, and the dimerization efficiency was higher than GagPol(WT). Therefore, IN:M50I suppresses autoprocessing after homodimerization, most likely at the initial cleavage step is regulated by the mutation. We presumed this suppression
might be caused via a structural hindrance; however, in silico predicted modeling did not indicate any significant difference in the structure between Wt and IN:M50I, therefore, IN:M50I mutation may alter the structure in GagPol polyprotein or increase the stiffness in the protein dynamics followed by decreasing the flexibility of the polyprotein or enhancing the efficiency of oligomerization GagPol polymerization. This change may subsequently suppress the initial autoprocessing. Further study define the mechanism of the inhibition. Crystal analysis of the entire length of GagPol polyprotein, even a full-length of IN, was not completed due to unclear structure at the C terminus domain (31, 33). Amino acid substitution of IN in the crystal structure analysis had greatly improved solubility (44); thus, the M50I mutation may contribute to the structure analysis of IN.

The molecular events leading to viral assembly and release have been extensively investigated using Gag polyproteins (45-49). Using sequential deleted IN mutated virus, it is reported that IN plays a role in assembly (8, 31, 50-53). A role for IN in proteolytic processing of GagPol and Gag polyproteins has been reported using several mutants: aa substituted or domain-truncated IN (8, 50, 51). Selected mutations associated with resistance to an integrase inhibitor KF116 regulated proteolytic processing of Gag and GagPol (9, 54), and quinoline-based allosteric integrase inhibitors (ALLINIs) not only suppress integration but also block the formation of mature virions by impairing their maturation (55, 56). These results indicate that the GagPol structure and the activity of autoprocessing by the embedded PR are intimately linked with the structure of IN; however, the mechanism of the regulation is poorly investigated. In the present study, we demonstrated that in addition to IN domain, RH domain in the GagPol polyprotein is also involved in the regulation of autoprocessing. To define the molecular mechanism of the role of RH:N79S mutation in GagPol, further studies are needed to elucidate
the mechanism of the RH-mediated regulation of autoprocessing. This study may reveal a novel target to regulate HIV assembly and budding.

When we performed the Western blotting using different lot of HIV(IN:M50I) lysates with different lot numbers of polyclonal antibodies, we consistently observed the ~67 KDa band but never detected a mature PR band (the cleaved product at sites 6 and 7 in Fig 5b). Therefore, we assumed that PR is still functional in the presence of IN:M50I, but the mutations may interrupt the conventional initial autoprocessing by PR and induce a disorder in the processing. The order of autoprocessing by PR is well-investigated. It is generally agreed that initial autocleavage is occurred at site 3 (Fig 5a) (5, 23, 36), followed by cleavage at site 5 (5) or site 1 (36) as an intramolecular process (5). The enzymatic digestion at site 7 in HIV(WT) is considered a late step in the processing (36). Using HIV(IN:M50I_Δ7), we demonstrated that the profile of GagPol-cleaved products in the mutant was comparable to that in HIV(IN:WT_Δ7)–in essence, restoring the defect created by the mutation in IN and identifying this as a critical site for the interaction of IN with the GagPol polyproteins and autoprocessing. Since the suppression was restored in HIV(IN:M50I_Δ7), it may be caused by an indirect effect at site 3, but it is regulated by site 7. We need further study to delineate the mechanism of the inhibition of the initiation and regulation by sites 7. This study may further define the regulation mechanism in the initiation of autoprocessing.

In summary, the current study has characterized naturally occurring SNPs in IN and RH regulate assembly, leading to abnormal virion formations and suppressing autoprocessing in the released virions. These findings illustrate the multi-functional aspects of several HIV-1 proteins and indicate that the overall structure-function relationships of the intact virion proteins may involve more than Gag. Further study of functional and structural analyses of RH and IN in
GagPol polyproteins and the GagPol interaction mechanism may provide new insights into the regulation of viral assembly, budding, and viral maturation in HIV. The study may disclose new targets to develop novel anti-HIV drugs that may be effective against HIV variants, especially multi-class drug-resistant viruses.

Materials and Methods

Approval for these studies including all sample materials was granted by the National Institute of Allergy and Infectious Diseases Institutional Review Board and participants were informed written consent prior to blood being drawn. All experimental procedures in these studies were approved by the National Cancer Institute at Frederick and Frederick National Laboratory for Cancer Research and performed in accordance with the relevant guidelines and regulations.

Cells: Peripheral blood mononuclear cells (PBMCs) were isolated from healthy donors’ apheresis packs using lymphocyte separation medium (ICN Biomedical, Aurora, OH, USA) (57). CD4(+) T cells were purified from PBMCs using CD4 MicroBeads (Miltenyi Biotec, Auburn, CA, USA) according to the manufacturer’s instructions. The purity of the cell types was at least 90%, based on flow cytometric analysis. Cell viability was determined using trypan blue (Thermo Fisher, Waltham, MA, USA) exclusion method. HEK293T cells and HeLa cells were obtained from ATCC (ATCC, Manassas, VA, USA) and maintained as previously described (57).

Site-directed Mutagenesis: Mutations of interest were induced on pNL4.3, a plasmid encoding a full-length of HIVNL4.3 (39) (the plasmid was obtained from Dr. M. Martin through the AIDS
Research and Reference Reagent Program, National Institute of Allergy and Infectious Diseases, National Institutes of Health) by the QuickChange Lightning kit (Agilent Technologies, Santa Clara, CA, USA) according to the manufacture protocol. Briefly, an ApaI (New England Biolab (NEB), Ipswich, MA, USA) and EcoRI (NEB) fragment of pNL4.3 was cloned into the pCR2.1 vector (Thermo Fisher). This clone was used as a shuttle vector and served as the backbone for mutagenesis studies. Primers used in the mutagenesis assays are listed in Supplementary Table S9. All mutagenesis was confirmed by Sanger DNA sequencing using the BigDye terminator v3 (Thermo Fisher) and SeqStudio Genetic Analyzer (Thermo Fisher). After confirming the sequence, the intended clones were digested with ApaI and EcoRI and then the fragments from ApaI and EcoRI digestion were used to replace with the corresponding fragment of pNL4.3. DNA sequencing was used to ascertain that each clone possessed the intended mutations.

Recombinant HIV-1 Viruses: Recombinant HIV-1s were prepared by transfection of pNL4.3 following a method previously reported (58). Briefly, 4x10^6 HEK293T cells or 2x10^6 HeLa cells seeded in a 100 mm dish were transfected with 10 μg of each plasmid purified using the endotoxin-free plasmid isolation kit (Qiagen, Germantown, MD, USA) with TransIT-293 (Mirus, Houston, TX, USA) for HEK293T cells or TransIT-HeLa MONSTER (Mirus) for Hela cells. Culture supernatants were collected at 48 hours after transfection. After centrifugation at 500 x g for 5 min, the supernatants were filtered through a 0.45 μm pore size membrane filter (MiliporeSigma, Louis, MO, USA). Viral particles in the filtrate (8 mL) were ultra-centrifuged on 20 % sucrose in HBS [10 mM HEPES (Quality Biochemical Inc, QBI, Gaithersburg, MD, USA)-150 mM NaCl] cushion for 2 hours at 4 °C and then was resuspended in 80 μL of RP10 or PBS (QBI) and stored at -80 °C until use. Viruses were only used after a single thaw.
Concentration of HIV in each stock was determined by a p24 antigen capture kit (PerkinElmer, Waltham, MA, USA). As HIV(IN:M50I) variant demonstrated a lower level of virus production from HEK293T cells (Supplemental Table S1), for Western blots and a transmission electron microscopic analysis, the variant was produced using three to 5 of 100 mm dishes at a time and combined the produced particles from the multiple dishes for downstream assays.

**HIV Replication Assay:** The levels of HIV-1 replication were determined using primary CD4(+) T cells as follows. CD4(+) T cells were stimulated with 5 μg/mL phytohemagglutinin (PHA; MiliporeSigma) in complete RPMI-1640 (Thermo Fisher) supplemented with 10 mM HEPES, 10% (vol/vol) fetal bovine serum (FBS; Thermo Fisher), and 50 μg/mL gentamicin (Thermo Fisher) (RP10). The PHA-stimulated CD4(+) T cells (10x10^6 cells) were infected with 10 ng of p24 of HIV in 1 mL for two hours at 37°C and then cultured at 1x10^6 cells/mL in the completed RPMI-1640 supplemented with 20 units/mL of recombinant IL-2 (MiliporeSigma) (RP10) for 14 days at 37 °C in T25 flasks. Half of the culture supernatants were exchanged with fresh RP10 every 3 or 4 days of incubation. HIV-1 replication activity was determined by measuring p24 antigen levels in the culture supernatants using the p24 antigen capture assay as described above.

**Western Blot:** Virus lysates for Western Blot analysis were prepared using radioimmuno-precipitation assay buffer (RIPA) lysis buffer (Boston biology, Boston, MA, USA) with a proteinase inhibitor cocktail (MiliporeSigma). Total protein concentration in the samples were quantified using a BCA protein assay kit (Thermo Fisher) and one or five μg of total viral proteins were subjected. SDS-PAGE was run using 4-12% NuPAGE Bis-Tris Gels (Thermo Fisher) in MOPS buffer (Thermo Fisher) for most of cases (58); however, to detect PR, the
SDS-PAGE was run in MES buffer (Thermo Fisher), and Western blot analyses were conducted using the ECL Prime Western Blot Detection system (Thermo Fisher) as previously described (57). Mouse monoclonal anti-HIV-1 CA (p24) antibody (Cat#: ab9071), Rabbit polyclonal anti-HIV PR antibody (Cat#:ab211627), Rabbit polyclonal anti-HIV1 RT antibody (Cat#: ab63911), and mouse monoclonal anti-HIV IN antibody (Cat#: ab66645) were obtained from Abcam (Cambridge, MA, USA). Goat polyclonal anti-MA(p17) antibody and Rabbit polyclonal anti-Nef antibody were kindly provided by Dr. R. Gorelick in Frederick National Laboratory for Cancer Research and Dr. R. Swanstrom (through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH (Catalog #: 2949)) (59), respectively. HRP-conjugated anti-Rabbit IgG and anti-mouse IgG antibodies (Cat#:NA931V and NA934V) were obtained from Thermo Fisher.

**Transmission Electronic Microscopic Analysis:** Plasmid-transfected HEK293T cells were cultured for 24 hours. Cell-free transfection supernatants containing recombinant viruses, or the transfected cells were fixed with 2.5 % Glutaraldehyde (E.M. Sciences, Warrington, PA) in Millonig’s Sodium Phosphate Buffer (Tousimis Research, Rockville, MD) (G-MPB) for 1 min at room temperature. Virus were pelleted as described above and the fixed cells were harvested using a scraper and centrifuged at 500 xg for 15min. The virus and the cell pellets were stored in a fresh G-MPB at 4 °C for overnight. Virus particles were pelleted as described above and then pelleted particle were fixed at 4 °C using G-MPB without disturbing the pellets. All fixed samples were washed repeatedly in Millonig’s Buffer, and then incubated for 2 hours in 1.0 % Osmium Tetroxide (E.M. Sciences), in Millonig’s buffer. Following rinsing steps in Ultrapure Water and en bloc staining with 2.0 % Uranyl Acetate (E.M. Sciences), the samples were dehydrated in a series of graded Ethanol and infiltrated and embedded in Spurr’s plastic resin.
Embedded blocks were sectioned using a Leica UC7 Ultramicrotome. 70-80 nm sections were collected on 150 mesh copper grids, and post-stained with Reynold’s Lead Citrate. Samples were examined in a FEI Tecnai Spirit Twin transmission electron microscope, operating at 80 kV.

**RNA Copy Assay:** To quantify copy numbers of HIV genomic RNA in virus stocks, the RealTime HIV-I assay kit (Abbot Laboratories, Abbott Park, IL, USA) was used. Briefly, 100 μL of HIV stocks were 10-fold serially diluted and then combined with an internal control RNA from the kit. All assays were conducted on the automated m2000 System (Abbot) with Abbott mSample Preparation System reagents (Abbot). The detection range of the assay system was 40 to 10x10^6 copies/mL. BCA protein assay was used to determine total protein concentrations in each sample and the results from the copy assay were normalized by the amounts of protein.

**Förster Resonance Energy Transfer (FRET) Assay:** Interactions of Gag-Gag, Gag-GagPol and GagPol-GagPol were measured by FRET assays using a series of the modified-expression plasmids encoding HIV_NL4.3 Gag or GagPol (lacking frame shifting signal and containing inactive PR with D25N mutation) fused with fluorescent protein genes (EGFP or mSB), generating pGag(MA/EGFP/CA), pGag(MA/mSB/CA), pGagPol(MA/EGFP/CA) and pGagPol(MA/mSB/CA) (20). Since pGagPol constructs lacking the frameshifting signal and _gag_ and _pol_ were placed in-frame and Gag-Pol polyproteins were translated without the frameshift (20). To induce an inactive PR in pNL(WT) and pNL(IN:M50I), a D25N mutation was induced in the PR of each plasmid using the point mutagenesis with D25N primers.
digested with *ApaI* and *EcoRI* and the *ApaI*-*EcoRI* region was replaced with a corresponding *ApaI* and *EcoRI* fragment of PR-inactive pNL(WT) or pNL(IN:M50I). The subcloned plasmids were termed pGP(WT)GFP, pGP(WT)mSB, pGP(IN:M50I)GFP, and pGP(IN:M50I)mSB.

**HEK293T cells** (50x10^3 cells) were seeded for 24 hours on μ-Slide 8-Well Glass Bottom chambers (ibidi, GmbH, Planegg, Germany) and then transfected with a total of 0.26 μg of DNA (GFP vector: mSB vector=1:1) using the TransiT-293. 24 hours after transfection, cell images were taken. All images were acquired on a Zeiss Axio Observer.Z1 equipped with the LSM800 confocal module, using a Plan-Apochromat 63x/1.40 objective (Carl Zeiss Microscopy, White Plains, NY, USA); cells were maintained at 37 °C and 5 % CO₂ during the experiments. For FRET analysis, a combination of three images was taken for each field of view, with respective excitation/emission wavelengths of 488/491-509 nm (Donor (EGFP) channel), 561/587-603 nm (Acceptor (mSB) channel) and 488/587-603 nm (FRET channel). The pinhole was set at 35 μm throughout all three FRET channels, i.e., 0.82 A.U. for donor channel and 0.70 A.U. for acceptor and FRET channels. Scan zoom was kept at 1.0x resulting in a 0.099 μm pixel size. At least 24 images from three independent experiments per experimental condition were subjected to FRET analysis using the “FRET and Colocalization Analyzer” plugin of the Fiji app, as follows. For each field of view, a bleed-through-corrected FRET index image was generated using the previously described formula (60).

FRET index image = I_{\text{FRET}} − (α_D I_{\text{Donor}}) − (α_A I_{\text{Acceptor}})

where α_D and α_A are signal bleed through coefficients from donor-only and acceptor-only transfection conditions, generated for each independent experiment, and I_{\text{FRET}}, I_{\text{Donor}} and I_{\text{Acceptor}} are signal intensities from FRET, Donor and Acceptor channels, respectively; a 1-pixel radius

(Supplementary Table S9). The pGagPol(MA/EGFP/CA) and pGagPol(MA/mSB/CA) were digested with *ApaI* and *EcoRI* and the *ApaI*-*EcoRI* region was replaced with a corresponding
median filter was subsequently applied to resulting FRET index images to smooth signal dispersion. Resulting background levels on filtered images were determined for each independent experiment and filtered-out, thereby creating masks mapping every above-background FRET-positive pixel. These masks were then applied to both unfiltered FRET index, Donor and Acceptor images, hence allowing for unbiased analysis of FRET efficiency for each field of view as a whole, and not singled-out cells. Apparent FRET efficiencies were calculated as previously described (20, 61), as a function of acceptor ($E_A$).

$$E_A = \frac{I_{FRET} - (\alpha_D I_{Donor} - \alpha_A I_{Acceptor})}{I_{Acceptor}}$$

**Tubulin Immunostaining**: HEK 293T cells (3x10⁵ cells) were plated on the sterilized round 15mm #1 glass coverslips (VWR) in 12-well plates and seeded for 24 hours prior to transfection. Transfection was conducted using 1.5 µg of pG-GFP, pGP(WT)GFP or pGP(IN:M50I)GFP as described above and then 24 hours after transfection, culture medium was removed. The cells were fixed with MeOH-free 4% formaldehyde (E.M. Sciences) for 15 minutes at room temperature; fixative was subsequently removed, and cells were washed four times in PBS (Quality Biological). Cells were then permeabilized using 0.1% Triton (Calbiochem, MiliporeSigma) in PBS for 5 minutes at room temperature, then washed four times in PBS, after cells were blocked using the BlockAid™ blocking solution (Thermo Fisher Scientific) for one hour at room temperature, coverslips were then recovered from plates and placed upside-down on a 100mL drop of antibody-containing BlockAid™ (ab204686, Recombinant Alexa Fluor 647 Anti-beta Tubulin antibody at 1:100; ab206627, Recombinant Alexa Fluor 555 Anti-beta Tubulin antibody at 1:100) and incubated in a humidity chamber overnight at 4°C in the dark. Coverslips were then washed 4-times in PBS and mounted on glass slides in a drop of ProLong™ Diamond.
Antifade Mountant with DAPI (Thermo Fisher Scientific). Imaging was carried out on a Zeiss Axio Observer.Z1 motorized microscope using a Plan-Apochromat 63x/1.40 objective and the LSM800 confocal module.

**Homology Modeling of HIV-1 Integrase Wild-type and M50I Mutation Structures.**

The wild-type HIV-1 integrase (WT) protein sequence was based on the NL4.3 genome sequence. The full-length structure was predicted using the Baker Lab hosted Robetta server, which follows the RosettaCM protocol: recombines fragments from the aligned template structures and de novo fragments building the unaligned regions, followed by structural optimization and refinement (62). To construct the homology model of IN based on the DNA-free templates, the sequences of three domains of IN, i.e., N-terminal domain (NTD), catalytic core domain (CCD), and C-terminal domain (CTD), were separately used as query sequences. To identify the corresponding templates using BLAST searching the PDB database (63) where three structures (PDB IDs: 1WJB (64), 1K6Y (34) and 1WJA (64)) were identified as the templates of CCD with the identity of 99.4%; and two final structures (PDB IDs: 2ITG (65), 1B9D (66) and 1ITG (44)) were identified as the templates of CCD with identity of 99.4%; and finally two structures (PDB IDs: 5HOT (67) and 1EX4 (32)) were chosen as the CTD templates with the identity of 100% and 94.7%, respectively. These eight templates were then used as the input of Robetta to perform the multiple-template modeling protocol. Specifically, the WT sequence with a length of 288 amino acids was submitted into the Robetta server. The default options were used for the model-building procedure, and the eight templates were uploaded. Then, the comparative model protocol was selected with 1000 models sampled. Finally, the best WT model was determined by carefully comparing the structures of templates. After that, the mutated sequences
of M50I and the WT sequence were submitted to Robetta with the same protocol as described above, using the predicted WT structure as the single template for modeling corresponding structures. The two final predicted models (IN:WT and IN:M50I) and the alignment between them are shown in Figures 1a - 1c. It appears that the structures are aligned well, and no significant conformation change can be detected by Robetta between the wild-type and M50I mutated sequences.

Population analysis

All subtypes of HIV-1 Pol sequences of group M (subtype A, B, C, D, F, G, H, J, K, L, U, and Circular recombinant form (CRF) ) were obtained by downloading from Los Alamos HIV sequence database (https://www.hiv.lanl.gov/components/sequence/HIV/search/search.html) for genomic region covering RH and IN. One sequence per patient was downloaded. There was a total of 15453 sequences downloaded. In summary, there were 1194 sequences of subtype A, 6357 for subtype B, 2880 for C, 260 for D, 144 for F, 144 for G, 14 for H, 10 for J, 3 for K, 3 for L, 40 for U, and 4404 for CRF. Sequences that do not provide RH Codon79 (RH:79, IN codon 17 (IN:17), or IN codon50 (IN:50) were excluded. The final number of sequences used for the analysis was followed: RH:N79 (748 for subtype A, 5239 for B, 2609 for C, 202 for D, 90 for F, 89 for G, 10 for H, 7 for J, 3 for K, 3 for L and 26 for U, 3483 for CRF), IN:S17 (1194 for subtype A, 6352 for B, 2877 for C, 260 for D, 144 for F, 144 for G, 14 for H, 10 for J, 3 for K, 3 for L, 39 for U, and 4404 for CRF) and IN:M50 (1194 for subtype A, 6351 for B, 2875 for C, 260 for D, 144 for F, 144 for G, 14 for H, 10 for J, 3 for K, 3 for L, 40 for U, and 4403 for CRF).
**Statistical Analysis:** Intergroup comparisons were performed using two-tailed unpaired t-tests using Prism 8 software (GraphPad, San Diego, CA, USA). *P* values <0.05 were considered statistically significant.

**References**


12. Migle Gabrielaite AGZ, Marc Bennedbaek, Christina Ekenberg, Virginia L. Kan, Giota Touloumi, Linos Vandekerckhove, Dan Turner, James Neaton, H. Clifford Lane, Jens D. Lundgren, Marvig. RL. 2020. ASSOCIATION OF HIV AND HOST GENETIC VARIANTS IN ANTIRETROVIRAL THERAPY-NAIVE PERSONS, abstr The Conference on Retroviruses and Opportunistic Infections, Boston, Massachusetts, March 8-11, 2020


Acknowledgments

Authors thank Drs. M. Bennedbæk and J.D. Lundgren for providing information of the 14 SNPs to initiate this study, Dr. Y. Morikawa for providing plasmid constructs for FRET assay, Dr. R. Gorellick for providing anti-MA antibody, Dr. J. Kovacs for discussion, Drs. L. Huzella, X. Jiao, and F. Scrimieri for critical reading of the manuscript. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of
Health, under Contract No. HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This research was supported [in part] by the National Institute of Allergy and Infectious Disease. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

**Author contributions:** T.I. designed all studies, performed assays, analyzed data, oversaw assays, supervised and wrote the manuscript. Q.C. performed assays. J.G.B., S.L., J.Y., H.H., M.H. and H.S. contributed assays and wrote a draft of manuscript. R.D. and W.C. supervised and reviewed the draft manuscript. H.C.L. conceptualized this project and wrote the manuscript. All authors reviewed the final manuscript.

**Competing interests:** Authors declare that they have no competing interests.

**Data and materials availability:** All data are available in the main text or the supplementary materials.
Table 1: Natural polymorphic mutations associated with viral loads in antiviral drug naïve HIV-infected patients.

<table>
<thead>
<tr>
<th>#</th>
<th>HIV Protein *</th>
<th>Codon #</th>
<th>AA Change</th>
<th>Designation †</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RT</td>
<td>322</td>
<td>Ser&gt;Thr</td>
<td>RT:S322T</td>
</tr>
<tr>
<td>2</td>
<td>RH</td>
<td>79</td>
<td>Asn&gt;Ser</td>
<td>RH:N79S</td>
</tr>
<tr>
<td>3</td>
<td>IN</td>
<td>17</td>
<td>Ser&gt;Asn</td>
<td>IN:S17N</td>
</tr>
<tr>
<td>4</td>
<td>IN</td>
<td>50</td>
<td>Met&gt;Ile</td>
<td>IN:M50I</td>
</tr>
<tr>
<td>5</td>
<td>Gp120</td>
<td>252</td>
<td>Arg&gt;Lys</td>
<td>G120:R252K</td>
</tr>
<tr>
<td>6</td>
<td>Gp120</td>
<td>448</td>
<td>Asn&gt;Arg</td>
<td>G120:N448R</td>
</tr>
<tr>
<td>7</td>
<td>Gp41</td>
<td>24</td>
<td>Met&gt;Val</td>
<td>G41:M24V</td>
</tr>
<tr>
<td>8</td>
<td>Gp41</td>
<td>24</td>
<td>Met&gt;Leu</td>
<td>G41:M24L</td>
</tr>
<tr>
<td>9</td>
<td>Gp41</td>
<td>254</td>
<td>Leu&gt;Arg</td>
<td>G41:L254R</td>
</tr>
<tr>
<td>10</td>
<td>Gp41</td>
<td>340</td>
<td>Leu&gt;Ala</td>
<td>G41:L340A</td>
</tr>
<tr>
<td>11</td>
<td>Nef</td>
<td>92</td>
<td>Lys&gt;Arg</td>
<td>NF:K92R</td>
</tr>
<tr>
<td>12</td>
<td>Nef</td>
<td>98</td>
<td>Glu&gt;Asp</td>
<td>NF:E98D</td>
</tr>
<tr>
<td>13</td>
<td>Nef</td>
<td>133</td>
<td>Val&gt;Ile</td>
<td>NF:V133I</td>
</tr>
<tr>
<td>14</td>
<td>Nef</td>
<td>143</td>
<td>Tyr&gt;Phe</td>
<td>NF:Y143F</td>
</tr>
</tbody>
</table>

* : RT: reverse transcriptase, RH: RNase H, IN: integrase, G120: gp120 envelope, G41: gp41 envelope, NF: nef. †: Name of designated mutation in this study.
Figure Legends

Figure 1. Characterization of 14 SNPs on HIV replication.

(a-c) PHA-stimulated primary CD4(+) T cells from healthy donors were infected with 10 ng p24 amounts of HIV(WT) or variants containing mutation in GagPol (a), envelope (b), or Nef (c) as described in the Materials and Methods. The infected cells were cultured for 14 days with media changed every 3-4 days. HIV replication was monitored using a p24 antigen capture kit. Representative data from three independent assays are presented as mean ± SD (n=3).

Figure 2. Characterization of viral and cell morphologies.

HEK293T cells were transfected with plasmid encoding HIV(WT) (a) or HIV(IN:M50I) (b) gene and cultured for 48 hours; each cell was fixed and then subjected for TEM as described in the Materials and Methods. Scale bars indicate 2 µm. Boxes indicate the magnified cell surface images. (c and d) HEK293T cells were transfected with plasmids encoding HIV(WT) (c) or HIV(IN:M50I) (d), and then cultured for 24 hours. Cells were fixed, and then TEM images were taken as described in the Materials and Methods. (e and f) TEM image analyses of sucrose-purified HIV particles of HIV(WT) (e) and HIV(IN:M50I) (f). Red and yellow arrowheads indicate typical electron-dense particles with a conical core (red) and electron-dense particles without a core (yellow). Scale bars show 100 nm in (c) and (d) and 200 nm in (e) and (f). (g) The relative population of viral morphology was calculated from 1000 particle images by three independent assays. Data indicates mean ± SE (n=3). (h) Viral genomic RNA copy numbers in viral stocks were compared between HIV(WT) and HIV(IN:M50I). Copy numbers were determined using the Real-time HIV-I assay kit from three independent virus stocks, and data were normalized to total viral protein. Data shows mean ± SE (n=3).
Figure 3. Evaluation of Gag and GagPol processing.
A total of 1 μg of viral proteins from HIV(WT) and HIV(IN:M50I) particles were subjected for WB using monoclonal anti-CA or polyclonal anti-PR, RT, IN, or Nef antibodies. Protein bands were detected by using the ECL assay as described in the Materials and Methods. The protein gel for detecting PR was run in MES running buffer (Thermo Fisher), other gels were run in MOPS running buffer (Thermo Fisher). Data are representative of two independent experiments.

Figure 4. Evaluation of Gag and GagPol dimerization using FRET assay.
HEK293T cells were co-transfected with pGag-EGFP and pGag-mSB (a), pGagPol(WT)-EGFP and pGagPol(WT)-mSB (b), or pGagPol(IN:M50I)-EGFP and pGagPol(IN:M50I)-mSB (c), and then FRET assays were conducted as described in the Materials and Methods. In the figures, Gag-EGFP and Gag-mSB are displayed G-EGFP and G-mSB. GagPol(WT)-EGFP, GagPol(WT)-mSB, GagPol(IN:M50I)-EGFP and GagPol(IN:M50I)-mSB are depicted GP(WT)-EGFP, GP(WT)-mSB, GP(M50I)-EGFP and GP(M50I)-mSB, respectively. (d) FRET efficiencies were calculated as described in the Materials and Methods. Gag and GagPol homodimerization are described as G-G and GP-GP, respectively. The efficiencies of GP-GP of WT and M50I were compared with that of the G-G pair. Data indicate mean ± SE from five independent assays. * P<0.05, *** P<0.001. (e-g) HEK293T cells were transfected with pGag-GFP (a), pGagPol(WT)-GFP (b), or pGagPol(IN:M50I)-GFP (c) construct and the protein distribution was observed. After transfection, cells were fixed with a fixation buffer (Abcam), and Tubulins were stained with Alexa Fluor® 555 Anti-beta Tubulin antibody [EPR16774] (ab206627). Scale bars indicate 20 μm.
Figure 5. Effect of the deletion of the cleavage site between PR and RT in GagPol on autoprocessing.

(a-e) Diagrams indicate the structure of GagPol(WT) and variants. Numbers above the diagrams indicate cleavage sites #1 – #9 by mature PR. Molecular size of each domain: MA=17 kDa, CA=24 kDa, p2=2 kDa, NC=7 kDa, TF=1 kDa, P (PR)=10 kDa, RT=51 kDa, RH=10 kDa, IN=35 kDa. The cleavage site 7 between PR and RT in HIV(WT) and HIV(IN:M50I) was deleted by the point mutagenesis as described in the Materials and Methods, and the resulted clones contain a fusion gene of PR and RT (PR-RT). (f-i) Virus particles are isolated using ultracentrifugation as described in the Materials and Methods, and five µg of viral lysates were subjected to WB. PR, CA, and IN were detected by (f) anti-PR, (g) anti-MA, (h) anti-CA, and (i) anti-IN antibodies.

Figure 6. Evaluation of RH:N79S or IN:S17N mutation on HIV(IN:M50I) virus replication and GagPol processing.

(a and b) PHA-stimulated primary CD4(+) T cells were infected with HIV(WT) or HIV(IN:M50I) with HIV(RH:N79S) or HIV(RH:N79S/IN:M50I) (a) or with HIV(IN:S17N) or HIV(IN:S17N/IN:M50I) (b). The infected cells were cultured for 14 days with media changed every 3-4 days. HIV replication was monitored using a p24 antigen capture kit. Representative data from two independent assays are presented. Data are presented as mean ± SD (n=3). (c and d) Viral lysates of HIV(WT), HIV(RH:N79S), HIV(IN:S17N), HIV(IN:M50I), HIV(RH:N79S/IN:M50I), and HIV(IN:S17N/IN:M50I) underwent WB analyses using anti-PR (c) and anti-CA antibodies (d). (e and f) Comparison of the morphology of viral particles using...
TEM. HEK293T cells were transfected with plasmid DNA encoding HIV(RH:N79S/IN:M50I) (e) or HIV(IN:S17N/M50I) (f) and then cultured for 24 hours. Cells were fixed for TEM analysis.

Figure 7. The impact of mutations at codon 50 in IN on HIV replication

PHA-stimulated primary CD4(+) T cells were infected with HIV(WT) or variants containing a different mutation at codon 50 as described in the Materials and Methods. The infected cells were cultured for 14 days with changing media every 3-4 days. HIV replication was monitored using a p24 antigen capture kit. Representative data from two independent assays are presented as mean ± SD (n=3).

Figure 8. Population analysis in HIV subtypes.

A total of 15,453 HIV sequences was obtained from Los Alamos HIV database, and amino acids (aa) sequences were compared to that in the corresponding HXB2 codon sequence. Population (%) of IN:M50I (a), RH:N79S (b), IN:S17N(c), and the frequency of the combination of RH:N79S with IN:M50I (d) and IN:S17N with IN:M50I (e) in each subtype is depicted using pie charts.