

Creative Commons Attribution 4.0 International (CC BY 4.0)

<https://creativecommons.org/licenses/by/4.0/>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

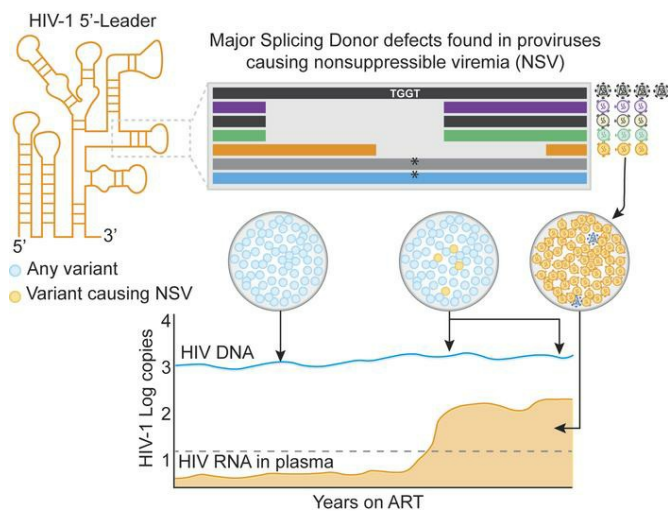
Clonally expanded HIV-1 proviruses with 5'-Leader defects can give rise to nonsuppressible residual viremia

Jennifer A. White, ... , Janet D. Siliciano, Francesco R. Simonetti

J Clin Invest. 2023. <https://doi.org/10.1172/JCI165245>.

Clinical Medicine In-Press Preview AIDS/HIV

Graphical abstract



Find the latest version:

<https://jci.me/165245/pdf>



Clonally expanded HIV-1 proviruses with 5'-Leader defects can give rise to nonsuppressible residual viremia

Authors

Jennifer A. White^{1*}, Fengting Wu^{1*}, Saif Yasin², Milica Moskovljevic¹, Joseph Varriale¹, Filippo Dragoni¹, Angelica Camilo-Contreras¹, Jiayi Duan¹, Mei Y. Zheng², Ndeh F. Tadzong², Heer B. Patel², Jeanelle Mae C. Quiambao², Kyle Rhodehouse¹, Hao Zhang³, Jun Lai¹, Subul A. Beg¹, Michael Delannoy⁴, Christin Kilcrease¹, Christopher J. Hoffmann¹, Sébastien Poulin⁵, Frédéric Chano⁵, Cécile Tremblay^{6,7}, Jerald Cherian¹, Patricia Barditch-Crovo¹, Natasha Chida¹, Richard Moore¹, Michael F. Summers^{2,8}, Robert F. Siliciano^{1,8}, Janet D. Siliciano¹, Francesco R. Simonetti¹.

Affiliations:

¹Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA

²Department of Chemistry and Biochemistry, University of Maryland, Baltimore County, Baltimore, MD, USA

³Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

⁴Institute for Basic Biomedical Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA

⁵Clinique L'Agora, Montréal, Canada.

⁶Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Canada.

⁷Département de Microbiologie, Immunologie et Infectiologie, Université de Montréal, Montréal, Canada.

⁸Howard Hughes Medical Institute, Baltimore, MD, USA

*JAW and FW contributed equally to this work.

Correspondence:

Francesco R. Simonetti, M.D, Ph.D.
Department of Medicine, Johns Hopkins University School of Medicine,
733 N Broadway, 21205, Baltimore, MD, USA.
Phone: 401-955-7757
Email: fsimonetti@jhmi.edu

Conflict-of-interest statement: Aspects of the intact proviral DNA assay (IPDA) are the subject of a patent application ("Compositions and methods Related to characterizing proviral reservoirs," PCT/US16/28822) filed by Johns Hopkins University, and RFS is one of the inventors on this application. RFS is a consultant for Neximmune. The other authors have declared that no conflict of interest exists.

Abstract

Background. Antiretroviral therapy (ART) halts HIV-1 replication, decreasing viremia to below the detection limit of clinical assays. However, some individuals experience persistent nonsuppressible viremia (NSV) originating from CD4⁺ T cell clones carrying infectious proviruses. Defective proviruses represent over 90% of all proviruses persisting during ART and can express viral genes, but whether they can cause NSV and complicate ART management is unknown.

Methods. We carried an in-depth characterization of proviruses causing NSV in 4 study participants with optimal adherence and no drug resistance. We investigated the impact of the observed defects on 5'-Leader RNA properties, virus infectivity, and gene expression. Integration-site specific assays were used to track these proviruses over time and among cell subsets.

Results. Clones carrying proviruses with 5'-Leader defects can cause persistent NSV up to ~10³ copies/mL. These proviruses had small, often identical deletions or point mutations involving the major splicing donor site (MSD) and showed partially reduced RNA dimerization and nucleocapsid binding. Nevertheless, they were inducible and produced non-infectious virions containing viral RNA but lacking Envelope.

Conclusion. These findings show that proviruses with 5'-Leader defects in CD4⁺ T cell clones can give rise to NSV, affecting clinical care. Sequencing of the 5'-Leader can help understanding failure to completely suppress viremia.

Funding. Office of the NIH Director and National Institute of Dental & Craniofacial Research, NIH; Howard Hughes Medical Institute; Johns Hopkins University Center for AIDS Research; National Institute for Allergy and Infectious Diseases, NIH, to the PAVE, BEAT-HIV and DARE Martin Delaney collaboratories.

63 **Introduction**

64 Treatment with antiretroviral therapy (ART) rapidly reduces plasma HIV-1 to below the
65 detection limit of clinical assays, prevents infection of new cells, and dramatically reduces HIV-
66 1-associated morbidity and mortality (1). However, ART is not curative because the virus
67 persists in a stable latent reservoir in resting memory CD4⁺ T cells, necessitating life-long
68 adherence to ART (2-5).

69 A small fraction of reservoir cells become activated daily and produce trace levels of free
70 virus (1-3 copies/mL) detectable with ultrasensitive assays (6, 7). These viruses are non-
71 evolving archival sequences sensitive to the current regimen (8-10), indicating that residual
72 viremia reflects virus release from stable reservoirs rather than ongoing replication cycles.
73 Residual viremia cannot be suppressed by treatment intensification (11, 12) and is often
74 dominated by identical sequences reflecting extensive proliferation of infected T cells (10, 13-
75 15), which contributes to reservoir stability (16-20).

76 Some treated PLWH experience sustained periods of detectable viremia, prompting re-
77 assessment of adherence, drug resistance, and drug concentrations. Regimen optimization or
78 intensification often fail to suppress viremia (11, 12, 21). Recent studies have shown that
79 expanded T cell clones carrying infectious proviruses can cause nonsuppressible viremia (NSV)
80 (14, 15).

81 In treated people living with HIV-1 (PLWH), >90% of proviruses are defective (22,
82 23) with large deletions and/or APOBEC3G/F-induced G-to-A hypermutations (24). Some
83 defective proviruses can express HIV-1 RNA and produce viral proteins (25-30). A recent study
84 suggested that a fraction of viruses found in plasma before ART was defective due to small
85 deletions and frameshifts (31). Some defective proviruses have small deletions in the 5'-Leader
86 (5'-L) upstream of *gag* (23, 32, 33). This region contains regulatory elements orchestrating
87 genomic RNA dimerization and packaging, reverse transcription, proviral gene expression, and
88 RNA splicing and translation (34-39). Proviruses with small 5'-L deletions comprise 5-10% of
89 proviral populations, are found in most PLWH, often in expanded clones (23, 32, 40-42), and
90 may produce viral mRNA and p24 protein (26, 33). Given the potential of 5'-L defects to affect
91 multiple steps in the viral cycle, these proviruses would not be expected to cause NSV or viral
92 rebound. However, here we demonstrate that these proviruses are a common cause of NSV
93 that complicates clinical management.
94

Results

Emergence of NSV despite ART. Participants were referred for persistently detectable plasma viremia despite optimal adherence. Their characteristics and HIV-1 reservoir measurements are summarized in Table S1, Figures 1, S1, and S2). The four participants had been living with HIV-1 for more than >15 years (range 15-32 years) and had been on long-term ART (range 7.8-27 years) to which they responded with peripheral CD4⁺ T cell recovery and undetectable viral loads. After years of viral suppression (<20 copies/mL), they experienced detectable viremia not explained by changes in adherence, drug bioavailability, or resistance. Treatment optimization, intensification, or de-intensification had no effect (Figure 1A, 1B, and Figures S1 and S2). Participant 1 (P1) had persistently detectable viremia for 4.3 years, with a median of 80 HIV-1 RNA copies/mL of plasma (range 37-156 copies/mL). P2 experienced intermittent periods of detectable viremia for more than 10 years, with a median of 75 copies/mL (range <20-300 copies/mL). P3 had persistently detectable viremia for 4.6 years, with a median of 123 HIV-1 RNA copies/mL (range 26-857 copies/mL) (Figure S1). P4 represents an extreme example of NSV, maintaining a median of 2979 copies/mL (range 1145-5138 copies/mL) for almost 2 years (Figure S2). Clinical histories are summarized in Supplemental Results.

Clonal origin of HIV-1 viremia. To investigate the cause of NSV, we recovered longitudinal single genome sequences of plasma RNA (Figure 1 and Figures S1, S2, and S3). In P1, 44 plasma *env* sequences were identical, except for 2 sequences with 1 nucleotide difference each, likely reflecting PCR error (43). The plasma sequences at all four time points spanning 1.5 years were identical, suggesting a predominant plasma clone resulting from cell proliferation and proviral expression (10, 14, 15). Analyses of U5-*gag* and p6-RT sequences confirmed that viremia was due to a single variant with no resistance mutations (Figure S3A, B). *Env* sequences in plasma at the time of ART initiation (2013) were diverse and belonged to two distinct lineages (Figure S3C). The provirus causing NSV on ART fell immediately outside one of these two main clades and showed no significant divergence from pre-ART sequences (Figure S3C). Proviruses from 6.5 and 7.8 years on ART (Figure 1C) were diverse but showed no increase in diversity (4.2% and 3.6%, respectively) or genetic shift (test for Panmixia $p=0.52$) over time. Most importantly, one proviral variant was identical to the plasma sequences and represented ~50% of proviral sequences at each time point (17/35 out of 35 and 21/39 out of 39, respectively).

In P2, multiple variants contributed to viremia (Figure 1D and S3), likely reflecting a larger reservoir as indicated by IPDA and QVOA (Table S1). From the first time point, we obtained 48 p6-RT RNA sequences belonging to 10 distinct variants that represented between 2.1% and 41.7% of virus in plasma. HIV-1 viremia decreased at later time points, resulting in lower sampling depth (11 and 8 sequences, respectively). Across three time points, we observed a variable representation of HIV-1 variants in plasma, and only 5 variants were found more than once, likely due to lower sampling and the variable virus production from infected clones over time (44). Three variants matched sequences recovered from viral outgrowth assays, representing replication-competent proviruses contributing to NSV (15). Interestingly, the most abundant plasma variant in plasma at the first time point (41.7%) did not match any of the 142 proviral sequences or the 16 outgrowth sequences from QVOA (Figure S3). This observation is not uncommon, as infected cells from this clone could be at very low frequency in blood, or are tissue-resident cells (45). For P3, 27/28 plasma p6-RT sequences represented a single variant (Figure S1B) that was rare among infected cells ($2.7\pm 5.2\%$, Figure S1B, S1C). For P4, 20/20 plasma p6-RT RNA sequences from 26.5 years on ART were identical, drug sensitive, and matched four RNA sequences obtained previously with a clinical HIV-1 genotype assay (Figure S2). Proviral sequences ($n=35$) were diverse and none matched the plasma variant, reflecting the low frequency of the provirus causing viremia (less than $2.8\pm 5.4\%$),

(Figure S2B). Of note, we did not detect drug resistance in any of the sequences recovered from the 4 participants.

Together these results show that NSV is often comprised of identical sequences despite enormous proviral diversity.

NSV can arise from proviruses with 5'-Leader defects. To investigate replication-competence of the proviruses causing NSV, we performed quantitative viral outgrowth assays (QVOAs) (46). In P1, despite the input of 11.12 million CD4⁺ T cells, there was no exponential outgrowth after 28 days, (<0.06 IUPM, Figure S4). One well showed a borderline p24 signal (~3 pg/mL) on day 21. Single genome U5-*gag* sequences from cell-associated RNA and supernatant virions (Figure S4) were identical to plasma virus, suggesting in vitro virus production without exponential replication of the provirus causing NSV.

The QVOA from P2 revealed a large reservoir size (15 IUPM) allowing recovery of 6 unique variants from 11/24 p24⁺ wells. The most abundant QVOA variant represented 37% of positive wells and matched one plasma clone found at all time points (Figure 1D). Two additional outgrowth viruses matched other clonal plasma sequences. Other plasma viruses from P2 were not detected in the QVOA. QVOAs were not performed for P3 and P4 due to sample availability.

To explore the replication competence of plasma clones not detected in the QVOA, we performed whole genome amplification of CD4⁺ T cell DNA (Figure 2A). Near full length proviral sequencing revealed subtype-B, R-tropic, drug sensitive proviruses (Figure S2D) with intact open reading frames. However, they all showed defects in the 5'-Leader (Figure 2B). Surprisingly, proviruses causing NSV from P1, P2 and P4, shared the same 22-nucleotide deletion (HXB2 positions 727-748) affecting the dimerization hairpin and the major splicing donor site (MSD, TGGT) (Figure 2B). The appearance of the same deletion in different individuals is likely favored by short repeats (GAG) at the deletion junction (Figure 2B), consistent with microhomology-driven template switching during minus strand synthesis (47). Two more proviruses causing NSV in P2 had 5'-L defects: a 21-nucleotide deletion (HXB2 positions 740-760) and a T-to-A mutation in position 745, both affecting the MSD site. Similarly, in P3 we found that 100% of plasma virus had an MSD mutation, also at position 745 (T-to-C) (Figure S1C). We confirmed the clonal nature of these recurring defective proviruses using integration site analysis (Figure 2C, Table S2, and Supplemental Results). These 5'-L defective proviruses are labeled with the host gene symbol and the specific 5'-L defect (e.g., ADK.d22). These results indicate that otherwise intact, clonally expanded proviruses with 5'-L defects can cause sustained NSV.

Defective 5'-Leaders exhibit modest changes in dimerization and nucleocapsid binding.

The 22-nt (5'-L^{d22}) and 21-nt (5'-L^{d21}) deletions remove the MSD and portions of the dimer initiation site (DIS) and packaging (ψ , ψ) hairpins, respectively, and disrupt the tandem three-way junction (Figure 3A) involved in genome packaging (37, 39) and in high-affinity binding to Gag (48). Therefore, we compared dimerization of 5'-L^{d22} and 5'-L^{d21} mutants to that of the HIV-1 NL4-3 (5'-L^{WT}) using native agarose gels. Under physiologic ionic conditions, all three RNAs dimerized (Figure 3B), with the 5'-L^{d22} form showing a modest reduction in dimer stability (~5-fold higher K_d), probably because the deletion occurs within the lower stem of the DIS hairpin. In contrast, the deletions in the 5'-L^{d21} RNA reside outside the DIS hairpin element and do not appear to influence RNA dimerization (Figure 3B).

To determine whether the deletions affect Gag binding, we used isothermal titration calorimetry (ITC). The 5'-L^{WT} RNA bound ~32 nucleocapsid (NC) molecules, whereas the 5'-L^{d21} and 5'-L^{d22} bound 20 and 25 NC proteins, respectively (Figure S5, Table S3). To probe for NC-induced RNA unwinding associated with the initial, highest affinity NC binding sites, we conducted additional ITC titrations to at lower NC-to-RNA ratios (37). ITC titration profiles for 5'-

L^{WT} exhibited a characteristic endothermic contribution to NC binding (37), but profiles for both mutants lack this feature (Figure 3D, E). Prior work with fragments of the native 5'-L indicate that this endothermic term corresponds to approximately four highest-affinity NC binding sites ($K_d \approx 40$ nM) and is found to be essential for efficient RNA packaging (37).

Taken together, these experiments show that 5'-L deletions found in proviruses causing NSV produce only modest reductions in dimerization and NC binding, consistent with retained genomic RNA packaging.

5'-Leader-defective proviruses are inducible and show rescue of gene expression

through alternative splicing. To investigate the inducibility of proviruses responsible for NSV, we stimulated CD4⁺ T cells from P1 and P2 with anti-CD3/anti-CD28 beads, and quantified cell-associated HIV-1 RNA (R-U5) at 0, 24, and 48 hours post stimulation (Figure 4). In addition, to selectively measure transcripts from proviruses of interest (*ADK.d22* and *DNAJB14.d21* from P1 and P2, respectively), we designed assays that would amplify only unspliced RNAs with those deletions (Figure 4A and S6). Finally, to rule out readthrough transcripts (49) we used an amplicon spanning the host gene-U3 junction. Upon stimulation, we observed a significant increase in total R-U5 RNA, and in transcripts originating from the *ADK.d22* and *DNAJB14.d21* proviruses, which at 48 hours represented 20% and 5% of the total viral RNA, respectively. Of note, we detected *ADK.d22* RNA at baseline, consistent with low-level spontaneous HIV-1 expression from this clone in peripheral blood cells. We did not detect chimeric RNA at the host gene-LTR junction. This was expected, given the opposite transcriptional orientation of the proviruses relative to the surrounding host genes, and suggests that provirus-specific transcripts are LTR-driven. We also detected low levels of p24 in culture supernatant and provirus-specific virion-associated RNA at 24 and 48 hours (Figure 4B), supporting the conclusion that genomic RNA from these defective proviruses can be packaged.

Given the absence of the MSD site due to the 22-nucleotide deletion, we analyzed splicing in cells carrying the *ADK.d22* provirus. CD4⁺ T cells from P1 were activated for 48 hours with anti-CD3/CD28 beads, and HIV-1 cDNA was synthesized from cell-associated RNA using primers annealing immediately downstream of two major splicing acceptors A5 and A7. Single genome amplification of singly- and multiply-spliced HIV-1 RNA revealed a novel non-canonical splicing donor site AGA*GT, created by the 22-nucleotide deletion in the *ADK* provirus. This deletion results in the fusion of two alternative splice donor sites, D1b and D1c (Figure 4C). Of 36 mRNA sequences, 35 were derived from the *ADK.d22* provirus, and 33 of these used the novel splice donor. The remaining 2 sequences used a previously described alternative splice donor ATGG*GT at the *gag* gene translation start site (26). We detected mRNAs for Env, Vpu, Tat, and Nef, but not for Rev which is required for the efficient export of intron-containing RNAs, and for Vif, which however represent only ~1% of all 4kb class of mRNA (50-52). Together, these results show that proviruses with an MSD deletion, can be induced, and can express some viral genes using a novel strong donor site.

Small 5'-Leader deletions result in non-infectious viral particles with decreased Envelope incorporation.

To investigate the impact of the 5'-L defects on replicative fitness, we introduced the deletions observed into a reference proviral construct (NL4-3) and generated virions by transfection of 293T cells (Figure 5A). At 72 hours post-transfection, virus production was reduced 10-fold relative to wildtype (Figure 5B). However, after normalization by p24, d22 and d21 had only a modest reduction in HIV-1 RNA packaging, as predicted by the 5'-L analyses described above (Figure 3).

To investigate replication capacity, we spinoculated activated primary CD4⁺ T cells with wildtype or 5'-L mutant viruses (Figure 5C). Only the wildtype virus showed an exponential increase in supernatant p24, while d22 and d21 showed persistently low p24 levels that were not affected by antiretrovirals and were likely the result of p24 carryover from the spinoculation

(Figure 5C). Reductions in RNA dimerization can reduce replicative capacity by affecting reverse transcription efficiency (53-55). Therefore, we measured late products of reverse transcription containing the U5-PBS junction in primary CD4⁺ T cells. The d22 and d21 mutants showed no cDNA at 12 hours post-infection (Figure 5D), suggesting a defect early in or upstream of reverse transcription. To test whether the deletions would prevent the initiation of reverse transcription due to a disruption of primer binding site (PBS) secondary RNA structure, we performed an in vitro tRNA binding assay. Both wildtype and mutant 5'-L RNAs bound tRNA (Figure S7), suggesting that the deletions did not affect PBS structure.

Based on these observations, we hypothesized that the deletions could prevent viral entry. After normalization by input p24, western blots of pelleted d22 and d21 virions showed negligible gp160 and gp41 (Figure 5E), indicating that the loss of infectivity is due to insufficient Envelope incorporation. To determine whether the lack of virion-associated Env is caused by reduced surface Env on infected cells, we studied Env on 293T cells 24 hours after transfection by flow cytometry (Figure 5F). Compared to wildtype, d22 and d21 showed 5.1- and 17.5-fold reductions in Env⁺ cells (WT 54%, d22 10.5%, d21 3.1% Env⁺ cells, respectively) (Figure 5G). Moreover, mean fluorescence intensity was significantly lower, suggesting that Env expression was also reduced at the individual cell level ($p < 0.0001$). We hypothesized that the lower Env expression is caused by alterations in splicing due to the 5'-L deletions. Therefore, we measured cell-associated HIV-1 RNA with assays targeting all polyadenylated transcripts (56), 4kb-class of spliced mRNA including polycistronic *vpu/env* transcripts, and multiply-spliced *tat/rev* mRNA (57). 5'-L deletions led to a significant reduction not only in absolute copies of 4kb-class of spliced mRNA, but also in its relative percentage among all transcripts (wild type 10%, d22 0.9%, d21 2.7%, $p < 0.0001$). Conversely, the relative abundance of *tat/rev* mRNA was increased in the 5'-L mutants (Figure 5I). Finally, we imaged 293T cells by transmission electron microscopy to visualize viral particles at 24 hours post-transfection (Figure S8). While a large fraction of cells transfected with wildtype NL4-3 showed high virus production and underwent cell death, transfection with d22 and d21 led to low-level production of viral particles and little cytopathic effect. Of note, viral particles looked had an immature appearance, with the typical radial distribution of Gag polyproteins (58). Together, these results show that the 5'-L defects observed in the plasma clones from P1, P2, and P4 cause splicing defects that reduce but do not completely abrogate the production of mRNAs encoding Env, resulting in lower virion Env protein levels and non-infectious viral particles.

Tracking cells carrying 5'-Leader defective proviruses over time. The above results indicate that proviruses with 5'-L defects precluding replication can cause NSV. To study the genesis and persistence of the T cell clones causing NSV, we performed digital PCR experiments to quantify the frequency and percentage of all infected cells belonging to *ADK.d22* or *DNAJB14.d21* clones (see Methods). We analyzed longitudinal samples collected from P1 before (3.8 years on ART) and after the onset of NSV (4.9, 6.5, 7, and 7.8 years on ART, respectively; Figure 6C). Although LTR copies remained stable (range 463-595 copies per 10⁶ CD4⁺ T cells, t-test between the first and last time point $p = 0.11$), *ADK.d22* copies significantly increased upon the onset of NSV (Figure 6C), from below the limit of detection (less than 1.13 copies per 10⁶ cells), to a new plateau of about 50 copies per 10⁶ CD4⁺ T cells. *ADK.d22* contributed to 17% of all LTR copies – about 1 out of 3 HIV-infected cells, if we assume most proviruses have both LTRs – and with an estimated total-body clone size of 10⁷ cells (see Methods). For P2, *DNAJB14.d21* had a frequency of 83 and 119 copies/10⁶ cells, corresponding to 2.2 and 3.3% of all LTR copies, at two time points during NSV. This reflects the much larger pool of total infected cells compared to P1. Despite this small proportion, *DNAJB14.d21* reached an estimated total body size of 24 million cells. In P3 and P4, the defective proviruses causing of NSV had a frequency of 20 and 44 copies/10⁶ cells,

representing only 0.5 and 1.8% of all LTR copies, respectively, and total body sizes of 4 and 6.7 million cells. Of note, infected clone sizes poorly correlated with plasma viremia.

Taken together, these results show that infected clones contributing to NSV are stable and have large total-body sizes, suggesting that only a small, constant fraction of these cells produces virus at any given time (15, 25).

Proviruses causing NSV are compartmentalized in Effector Memory T cells. To investigate the distribution of proviruses causing NSV among CD4⁺ T cell subsets, we quantified LTR and integration site copies in subsets identified based on CCR7 and CD45RA expression (Figure 7). P1 showed a significant shift in T cell subset percentages relative to the expected range for individuals of his age (59) (Figure 7A), with a marked increase in the more differentiated effector memory (EM) and effector memory CD45RA⁺ cells (EMRA) which also had a higher level of infection as assessed by LTR copies (494, 587, and 1170 copies per 10⁶ cells in CM, EM, and EMRA, respectively, $p=0.04$, Figure 7C). Interestingly, the *ADK.d22* copies were found almost exclusively in EM cells (79.2 copies per 10⁶ cells), which contributed to 96% of the cells carrying this provirus (Figure 7E, 7G).

P2 showed the expected distribution of subsets (Figure 7B) and a higher frequency of LTR copies in more differentiated cells ($p<0.0001$). The *DNAJB14.d21* provirus was also significantly enriched in EM cells (455 copies per 10⁶ cells, $p<0.0001$), in which this provirus represented 7.5% of all LTR copies, while it was not found in CM and EMRA cells (Figure 7F). The EM subset contributed to more than 96% of cells carrying *DNAJB14.d21* (Figure 7H). Thus, both clones causing NSV are compartmentalized EM cells. Given that these cells are characterized by shorter half-lives (60), their maintenance relies on frequent proliferation and differentiation from CM progenitors. Thus, it was surprising that *ADK.d22* and *DNAJB14.d21* proviruses were not found in CM and EMRA cells, suggesting that the two clones may defy the canonical “differentiation flux” from central to terminally differentiated memory cells (61, 62). To investigate whether this observation is common among all CD4⁺ T cells, and not unique to the two infected clones causing viremia, we analyzed TCR β repertoires from total, CM, and EM cells from P1, and found that about 20% of TCR β sequences were unique to EM cells (see Supplemental Results and Figure S10).

We and others have shown that infected cells can persist through extensive proliferation in response to viral antigens (42, 63-65). We sorted CD4⁺ T cells reactive to CMV and HIV-1 Gag to test whether the cells harboring *ADK.d22* were specific to these antigens. Although P1 showed responses to both CMV and Gag (6.7% and 1.6% of all CD4⁺ T cells, respectively, Figure S11), we did not find *ADK.d22* within these antigen-reactive cells. Thus, the antigen-specificities of the *ADK.d22* clone and those causing viremia from P2, P3, and P4 remain unknown.

Defective proviruses causing NSV evade cellular and humoral immune pressures. We hypothesized that a lack of clearance by the immune system of both virus-producing cells and virions could favor the occurrence of NSV. To investigate CTL escape, we analyzed full-length genomes of proviruses contributing to NSV and identified well-characterized HLA class-I restricted epitopes (Figure 8A). In P1, *ADK.d22* had escape mutations in 8 out of 19 epitopes (58%). This was particularly evident for B*57-restricted epitopes, as 9 out of 11 (82%) had previously documented mutations resulting from CTL escape (Figure 8B), including the A1P and I2L mutations in the Gag epitope AW10, and the T242N mutation in the Gag epitope TW10, the latter known to reduce viral fitness (Figure 8B and Table S3) (66-68). The high frequency of CTL mutations is likely the result of virus evolution during >20 years of untreated infection before P1 started ART. A similar pattern was observed for *CCND3.d22* from P4, in which 80% of B*57-restricted epitopes had nearly identical escape mutations (Figure 8B). Proviruses from

P2 and P3 also showed some escape mutations but retained ~60% of susceptible epitopes (Figure 8A).

We also evaluated susceptibility of the *ADK.d22* and *DNAJB14.d21* envelopes to autologous neutralizing antibodies (see Supplemental Results) and found that both were resistant. These results suggest that persistent low-level viral expression did not elicit neutralizing antibodies over time, and future studies should investigate whether this observation is unique to defective plasma clones possibly due to low Envelope incorporation or also typical of replication-competent proviruses causing NSV (15).

Together, these results show that defective proviruses contributing to NSV present variable degrees of CTL escape and were not neutralized by autologous IgGs. Owing to the small sample size of proviruses causing NSV and the lack of data from other intact or defective proviruses not found in plasma, we were not able to determine whether immune escape plays a fundamental role in the selection of proviruses driving NSV.

Discussion

A small fraction of proviruses is reactivated each day, giving rise to residual viremia (6, 69) that can be detected with ultrasensitive assays in most infants and adults on ART (70, 71). In some individuals this process results in years of detectable viremia produced by expanded clones carrying infectious proviruses (14, 15). The true frequency of NSV during ART is unknown. However, based on the study from Halvas and colleagues (19), and the cases referred from our clinic, we estimate that 1 in ~250 individuals on ART experiences persistent NSV. These cases are extreme examples of how the reservoir persists in all individuals on ART (16, 18, 19), but which factors lead a handful of infected clones – out of a myriad – to cause NSV remain unknown.

Some defective proviruses can produce HIV-1 RNA and viral proteins (26, 28, 29, 72), potentially leading to inflammation and/or immune activation, and deflecting CTL responses (26, 29, 30). However, whether defective proviruses can cause NSV was unknown. A previous report described defective viral variants in the context of residual viremia below 50 copies/mL (73). Here, we describe four consecutive cases of PLWH on long-term ART with persistent NSV of up to 10^3 copies/mL due to one or multiple proviral clones with defects in the 5'-L. In participants P1, P3, and P4 plasma virus was entirely caused by a single defective provirus, while participant in P2, who had a large reservoir, a mixture of infectious and 5'-L-defective proviruses contributed to NSV. All 6 defective genomes described here had alterations affecting the MSD site. Surprisingly, proviruses from three distinct participants shared the same d22 deletion, likely favored by the "GAG" repeats driving misplaced recombination events during reverse transcription (47). The remaining two proviruses had single nucleotide mutations at the critical T in the MSD (D1) site: T745A in P2, and T745C in P3, both previously reported (33, 74) and associated with a loss of replication-competence (73). Both mutations affect the second position of the conserved "GU" dinucleotide immediately downstream of the splicing site. The mutated nucleotide is included in the GURAGU motif recognized by the U1 snRNA, which is required to initiate the splicing process (75). Of note, besides these 5'-L mutations, these 6 proviruses described here had genetically intact promoters and open reading frames, including the antisense protein ASP, which has been shown to promote latency (76).

Although nearly all 5'-L defects abrogate replication-competence, their variable size and position complicate prediction of which 5'-L functions are disrupted. Das *et al.*, proposed that MSD-Ψ mutations could prevent proviral expression by activating the 5'-polyadenylation site, resulting in ultra-short non-coding HIV-1 transcripts (77). However, previous transfection experiments (26), assays based on T cell stimulation *ex vivo* (33), and the data presented here demonstrate that proviruses with 5'-L deletions can be induced and can produce virions. Our studies of dimerization and capsid binding show that these functions are only partially affected by the 5'-L mutants causing NSV. Experiments with competition-based packaging assays and incorporation of radio-labeled dNTPs may tease out more granular differences in efficiency relative to intact proviruses (78). However, our results show that these proviruses can contribute to viremia at levels that can be detected by clinical assays despite modest dimerization and packaging defects.

To maintain replicative fitness, HIV-1 must exploit the full spectrum of its alternative splicing possibilities (50). Deviations from the delicate balance between unspliced and spliced mRNAs can influence virus replication (79). The MSD site (or D1) plays a fundamental role in the regulation of HIV-1 splicing and is found in nearly all spliced transcripts (50). However, alternative splicing donors have been described (51), and can take over if D1 is mutated or missing (26). We confirmed this phenomenon by sequencing spliced mRNA forms in CD4⁺ T cells from P1, in which the 5'-L deletion found in *ADK.d22* created a new donor found in >90% of spliced transcripts. However, this alternative donor does not fully rescue generation of all HIV-1 mRNAs. We did not detect spliced mRNAs coding for Rev, which is fundamental for efficient

nuclear export of intron-retaining viral transcripts (50). Most importantly, although we detected spliced RNAs encoding Env in P1 cells stimulated ex vivo, the d22 and d21 deletions caused a striking reduction in Env expression in transfected cells that mirrored the lack of gp160/gp41 incorporation in viral particles.

Longitudinal studies of PLWH on long-term ART have shown that proviral populations are dynamic and under ongoing selective pressures, resulting in the more rapid decay of intact proviruses relative to proviruses with major defects (24, 80-83). The high frequency of 5'-L defective proviruses in individuals on ART, both within the proviral landscape overall and among highly expanded clones, suggests that these nearly intact proviruses have a selective advantage relative to intact proviruses, even if frequently expressed. 5'-L defects like d22 and d21 may allow the production of viral particles upon latency reversal while maintaining lower levels of viral proteins involved in cytopathic effects and immune recognition, especially Envelope. We hypothesize that 5'-L defective proviruses may be a common cause of residual viremia. In the four participants described here – who were not pre-selected – one or more 5'-L defective proviruses contributed to NSV. The development of clinical ultrasensitive assays that sequence the 5'-L as well as the *pol* gene would improve care of individuals with NSV by reducing concerns over cryptic viral replication and risk of HIV-1 transmission, unnecessary laboratory tests, and changes in ART regimens (84).

The genomic context of HIV-1 integration plays a role in the inducibility, proliferative potential, and persistence of infected cells (28, 85). The six integrants causing NSV described here were all located in gene-rich regions, within introns of coding genes in opposite orientation relative to their host gene transcription. Five out of six were within 20kb of known H3K27Ac, and H3K4Me1, H3K4Me3 histone marks, associated with enhancer regions in uninfected cells (86), and in genes expressed at medium-to-high levels in CD4⁺ T cells, involved in housekeeping functions such as nucleotide metabolism, DNA damage response, vesicle-mediated transport, and cell cycle progression. The exception is the *ZFYVE9.745C* provirus from P3, which is integrated into a gene expressed at very low levels in CD4⁺ T cells. Overall, although HIV-1 integration in these genes has not been linked to a selective advantage, these proviruses are likely located in genomic *loci* favorable for HIV-1 gene expression (28).

Clonal proliferation is a major mechanism of persistence of cells carrying both intact and defective proviruses (85). Here, we directly quantified four clonal proviruses by integration-specific digital PCR assays. Total body clone sizes – on the order of 10⁶-10⁷ cells – were comparable to those observed in HIV-1⁺ CMV-responding clonotypes and clones contributing to NSV (15, 25, 42). Interestingly, *ADK.d22* represented about 30% of all infected cells and 50% of all *env* proviral sequences in P1. This finding is in striking contrast with previous studies, in which proviruses matching the predominant plasma clones are rarely found in resting CD4⁺ T cells (10), and represented a small fraction of HIV-1 DNA single genome sequences (0-10.7%) and integration sites (0.03-1.10%) (15). The high proportion of *ADK.d22* proviruses can be explained by the small pool of infected cells in P1, a B*57⁺ viremic controller who remained off of ART for >20 years. This observation is in line with reservoir studies from elite controllers in which few large clones dominate the proviral landscape (87, 88). For the first time, we show that the onset of viremia is concurrent with the rapid waxing of the cells containing the *ADK.d22* provirus, suggesting an immune event that activated its cognate clonotype, leading to proliferation and frequent virus production. In addition, our results on *ADK.d22* and *DNAJB14.d21* indicate that they are stable and compartmentalized in short-lived effector memory cells, suggesting persistence by frequent proliferation, rather than homeostatic survival signals and long half-lives (60). Immune responses to recurrent antigens, such as HIV-1 and cytomegalovirus, can drive the proliferation of infected clones (42). However, the cognate antigens of the clones causing NSV are yet to be determined. NSV may be a reflection of specific T cell responses against antigens that cannot be cleared, such as chronic infections – including HIV-1 itself –, commensal pathogens, or self-antigens. This may explain why only rare

462 proviruses are frequently transcriptionally active, including those that will ultimately lead to viral
463 rebound upon treatment interruption (33, 89, 90).

464 Taken together, our results are consistent with 5'-L defective viruses produced from the
465 same provirus-containing clones; however, given the lack of sampling at multiple time points
466 before and during ART documenting the same integration sites, alternatives cannot be
467 excluded. These include the production of virus with limited diversity from different cells, or rare
468 new infection events with reseeding of new virus-producing cells over time. In addition, future
469 studies should include a systematic comparison between defective and infectious proviruses
470 cause of NSV, in order to tease out unique features, including susceptibility to cell-mediated and
471 humoral immunity, resistance to cell death, and genomic location.

472 In conclusion, we demonstrate that proviruses with 5'-L defects, when in favorable
473 genomic and immunological conditions, are a frequent cause of persistent nonsuppressible
474 viremia despite effective ART. Although incapable of causing viral rebound, these proviruses
475 contribute to variable levels of viremia and complicate ART management (84) leading to
476 additional testing, unnecessary ART changes and intensifications, and by causing anxiety and
477 frustration in patients and clinical care providers. Our work reveals additional complexity in
478 residual viremia and should prompt the development of assays that would allow detection of 5'-L
479 defects.

480

Methods

Study Participants. The study participants were referred to our attention from their HIV-1 care providers at the Bartlett Specialty Clinic, Johns Hopkins University (P1-P3) and at the Clinique I.D. of Saint-Jérôme (Quebec, Canada). Peripheral blood samples (180mL) were collected at one or multiple time points (2019-2022). For P1, historical samples (2013-2018) were obtained through a longitudinal study at the Bartlett Specialty Clinic. Samples from P4 were collected at the Centre Hospitalier de l'Université de Montréal (CHUM), in Montréal, Canada.

Study of HIV-1 sequences in plasma and CD4⁺ T cells. Blood samples were spun at 400g for 10' at 4C and plasma was spun again at 400g for 10' and frozen at -80°C. Upon thawing, plasma was spun at 3500g for 15' at 4C and transferred to tubes for ultracentrifugation and spun at 50000 rpm for 30' at 4C. Viral pellets underwent RNA extraction (91). RNA was used immediately for reverse transcription with Super Script III with primers located in *gag*, *RT*, or *env*. The cDNA was then used for single genome sequencing as previously described (42). For all participants, we initially recovered p6-RT sequences to exclude drug resistance. For P1, due to the small reservoir size and the partial control of HIV replication pre-ART, we focused on *env* SGS to avoid the risk of overestimating clonal sequences.

Intact proviral DNA assay (IPDA). IPDA was performed as previously described (24).

Quantitative viral outgrowth assay (QVOA). QVOAs from total CD4⁺ T cells were performed as previously described (46). Supernatants from p24 positive wells were processed as previously described to sequence virion-associated HIV-1 RNA (42).

Combined analysis of integration site and proviral genome. Endpoint-diluted gDNA was subjected to whole genome amplification, as previously described (92). Wells positive for HIV-1 genomes were detected by *gag*, *env*, or p6RT PCR. Wells with defective proviruses matching the predominant plasma clones were subjected to integration site analysis, and near full-length genome sequencing (92). Additional PCRs were performed to confirm the site of HIV-1 integration and recover the LTR sequence. Primers are provided in Supplemental Table S5.

Analyses of HIV-1 sequences. Sanger sequencing data was processed and analyzed as previously described (42). In brief, neighbor joining trees were performed based on a p-distance and bootstrap analysis with 1000 replicates. For the analysis of pre-ART sequences from P1 and sequences from P2 shown in Figure 1, we used maximum likelihood method based on the HKY+G substitution model with 1000 replicates, identified with Mega v7.0. For the tree analysis on near full genomes in Figure S2D we used a GTR+G+I substitution model. Intactness of proviral full-length genomes was assessed by Proseq-IT (93).

5'-Leader RNA studies. Methods can be found in the Supplemental Methods section.

Testing the impact of 5' Leader deletions. We used site-directed mutagenesis (New England Biolabs) to introduce each 5'-L deletion into an NL4-3 backbone, obtained through the NIH HIV Reagent Program, Division of AIDS, NIAID, NIH: ARP-114, contributed by Dr. M. Martin. Plasmids carrying the wild type (NL4-3wt) and mutants (NL4-3d22 and NL4-3d21) were used to transfect 293T cells (26). After 72 hours, we collected the supernatant which was then filtered and concentrated by ultracentrifugation with a 20% sucrose gradient. Virus recovery was measured by p24 ELISA (Perkin-Elmer) and RT-PCR measuring polyadenylated HIV-1 RNA. We then infected primary CD4⁺ T cells from a healthy donor after 3 days of activation with anti-CD3/anti-CD28 antibody-coated beads (1:1 cell to bead ratio). Each condition was carried in triplicates by spinoculation, with 10ng of p24 per 1 million cells in 100µl of media. After spinoculation, cells were washed 5 times with ice-cold PBS and plated at 1M/mL in RPMI with 10% FBS and 20u/mL of IL-2. To control for background p24 and plasmid carry over, an

identical experimental set up was carried in parallel with media containing tenofovir disoproxil fumarate (10uM), emtricitabine (10uM), and dolutegravir (10nM). Culture supernatants were collected at 0, 12, 24, 48 and 72 hours post-spinoculation and assayed by p24 ELISA. Cells were collected at 0, 6 and 12 hours post-spinoculation to extract genomic DNA and quantify late reverse transcriptase cDNA products by ddPCR probing the U5-PBS junction. Cell-associated RNA at 24 hours post 293T cell transfection underwent cDNA synthesis as described above and quantified by digital PCR (Qiagen) targeting all polyadenylated RNA, and spliced mRNA belonging to the 4kb-class and encoding for tat/rev (Supplemental Table S5). To capture mRNA from both wt and mutant viruses, the 4kb-class assay was designed to capture mRNA regardless of the D1 splicing donor site.

Western blots. Methods can be found in the Supplemental Methods section.

Flow-cytometry analysis of cells expressing HIV-1 Envelope. Transfected 293T cells were briefly dissociated using 100uL of TrypLE for 2 minutes followed by one wash in DMEM supplemented with 10% FBS. Cells were resuspended in 10% DMEM at a concentration of 2M/mL to seed 100uL into 96-well V-bottom plates. Cells were incubated with unlabeled primary antibody, 3BNC117 (NIAID, ARP-12474), at a final concentration of 15uM for 1hr at 37C. After two washes, cells were then stained with BV421-labeled secondary antibody against hu-IgG Fc (1:40 diluted, Biolegend, clone M1310G05) and viability dye (1:1000 diluted, eBioscience™ eFluor™ 780) for 30min at 4C. After two washes to remove excess antibodies, cells were analyzed using an Intellicyt iQue cytometer. Non-specific signal was assessed by staining cells with only BV421-labeled secondary antibody, and by staining cells transfected with a delta-*env* NL4-3 expression vector (NIAID, HRP-11100). Finally, background measured by using an anti-Human IgG primary antibody was subtracted from cells stained with 3BNC117.

Transmission electron microscopy. Methods can be found in the Supplemental Methods section.

Analysis of HIV-1 expression upon T-cell activation. Total CD4⁺ T cells were isolated from PBMCs by negative selection and plated in 24-well plates at 2 million/mL in RPMI with 10% FBS, 10nM dolutegravir, and anti-CD3/CD28 antibody-coated magnetic beads (cell to bead ratio 1:1). Cells and culture supernatant was collected at 0, 24, and 48 hours. Cell-associated genomic DNA and RNA were extracted as previously described (94). RNA fractions were subjected to cDNA synthesis as described above using a gag-specific primer. We used three probe-based assays to quantify: i) total HIV RNA (RU5) from any provirus, ii) unspliced RNA from defective proviruses of interest using probes annealing across the 5'-Leader deletion (Figure S6), and iii) read-through RNA from upstream the LTR using probes annealing across provirus-specific integration site. For Participant 1 we generated cDNA from cell-associated RNA in two separate reactions using primers annealing downstream of splicing acceptors A5 and A7 to capture 2kb and 4kb transcripts, respectively (52). Subsequently, cDNA was subjected to limiting-dilution 2-step PCR to amplify- singly and multiply-spliced transcript using forward primers upstream of both canonical and novel splicing donors. See Supplemental Table S5 for additional details.

Duplex quantification of total LTR copies and specific provirus. Methods can be found in the Supplemental Methods section.

T cell subset analysis. Methods can be found in the Supplemental Methods section.

Analysis of HLA-restricted epitopes. Participants were HLA typed at high resolution by the Johns Hopkins University Immunogenetics Laboratory. The CTL epitopes for each participant were identified using the Los Alamos National Lab's "Best-defined CTL/CD8+ Epitope

Summary". When functional annotations of certain mutations were not available, we used the prediction software NetMHC4.0 to classify epitope as strong, weak, or non-binder (95).

Neutralization assay. The complete sequence of the *env* gene was recovered as previously described (96). The starting material was viral RNA isolated from supernatant of autologous CD4⁺ T cells activated ex vivo for 24 hours with anti-CD3/CD28 beads (as described above). The PCR products, matching 100% of the *ADK.d22* and *DNAJB14.d21* proviruses, were cloned into an *env* expression plasmid and used to co-transfect HEK293T cells, together with an *env* deleted NL4-3 vector, to generate pseudotyped virus. Pseudoviruses were titrated and assayed for neutralization in TZM-bl cells as previously described (96). Inhibition of infection was expressed as a fraction of maximum infection. IC₅₀ was calculated as previously described (97).

Isolation of antigen-reactive CD4⁺ T cells. Experiments were conducted as described in Simonetti *et al.*, (42).

Quantification and statistical analyses. Descriptive statistics, tests for normality, 2-tailed Student's t-test and one-way ANOVA tests were used to determine statistical significance using GraphPad Prism v8.0. A P value less than 0.05 was considered significant, unless otherwise stated.

Data availability. HIV-1 sequences are available on GenBank (OQ092462- OQ092467, additional [accession numbers pending from GenBank](#)). Integration site sequencing data can be found on the NCI Retrovirus Integration Database (<https://rid.ncicrf.gov/>). TCRβ sequencing data can be accessed through the ImmuneAccess database (<https://doi.org/10.21417/JAW2023JCI>).

Study approval. The Johns Hopkins Institutional Review Board approved this study. The Study Participants provided written informed consent before enrollment. For P1, historical samples (2013-2018) were obtained through an Institutional Review Board-approved longitudinal study at the Bartlett Clinic. P4 provided written consent to an Institutional Review Board-approved study from the laboratory of Dr. Cécile Tremblay at the Centre Hospitalier de l'Université de Montréal (CHUM), Canada.

Author contributions

FRS, JDS and RFS conceived the study. FRS, JAW, FW, MM, JD, FD, ACC, and KR performed experiments and analyzed data. SAB and JL enrolled the study participants, performed sample processing, and viral outgrowth assays. SY, MYZ, NFT, HBP, JMCQ, and MFS conceived and performed the study of 5'-Leader RNA structure and function. CH, CK, JC, NC, PBC, RM, SP, FC and CT provided clinical care to the study participants and gathered their clinical history and samples. HZ conducted cell sorting experiments. JV performed pseudotyped virus neutralization assays. MD prepared samples and acquired electron microscopy images. FRS conducted analyses and generated figures. FRS wrote the manuscript with input from JDS and RFS, and received feedback and final approval from all authors. JW and FW equally contributed to this work and their order as co-first authors is alphabetical.

Acknowledgments

We deeply thank the study participants and their families for the commitment in volunteering in this study. We thank Annie Chamberland, Emma Neubert and Issac Chaudry for their assistance in sample preparation, and Monica Sullivan for administrative support. We thank Emily Fray, Annie A. Antar, Frank Maldarelli, Wei-Shau Hu, and Vinay Pathak for the useful discussions leading to this work. FRS is supported by the Office of the NIH Director and National Institute of Dental & Craniofacial Research (DP5OD031834), the Johns Hopkins University CFAR (P30AI094189) and the NIAID (PAVE, UM1AI164566). SY is supported by the training grant T32HL007698. This work was also supported by the NIH grant R01AI150498 (MFS), UM1AI164556, UM1AI164570, UM1AI164560 (RFS), and by the Howard Hughes Medical Institute.

730 **References**

- 731 1. Sengupta S, and Siliciano RF. Targeting the Latent Reservoir for HIV-1. *Immunity*.
732 2018;48(5):872-95.
- 733 2. Chun TW, Stuyver L, Mizell SB, Ehler LA, Mican JA, Baseler M, et al. Presence of an
734 inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc Natl Acad*
735 *Sci U S A*. 1997;94(24):13193-7.
- 736 3. Wong JK, Hezareh M, Gunthard HF, Havlir DV, Ignacio CC, Spina CA, et al. Recovery of
737 replication-competent HIV despite prolonged suppression of plasma viremia. *Science*.
738 1997;278(5341):1291-5.
- 739 4. Finzi D, Hermankova M, Pierson T, Carruth LM, Buck C, Chaisson RE, et al.
740 Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy.
741 *Science*. 1997;278(5341):1295-300.
- 742 5. Siliciano JD, Kajdas J, Finzi D, Quinn TC, Chadwick K, Margolick JB, et al. Long-term
743 follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T
744 cells. *Nat Med*. 2003;9(6):727-8.
- 745 6. Hill AL, Rosenbloom DI, Siliciano JD, and Siliciano RF. Insufficient Evidence for Rare
746 Activation of Latent HIV in the Absence of Reservoir-Reducing Interventions. *PLoS*
747 *Pathog*. 2016;12(8):e1005679.
- 748 7. Jacobs JL, Halvas EK, Tosiano MA, and Mellors JW. Persistent HIV-1 Viremia on
749 Antiretroviral Therapy: Measurement and Mechanisms. *Front Microbiol*. 2019;10:2383.
- 750 8. Hermankova M, Ray SC, Ruff C, Powell-Davis M, Ingersoll R, D'Aquila RT, et al. HIV-1
751 drug resistance profiles in children and adults with viral load of <50 copies/ml receiving
752 combination therapy. *JAMA*. 2001;286(2):196-207.
- 753 9. Nettles RE, Kieffer TL, Kwon P, Monie D, Han Y, Parsons T, et al. Intermittent HIV-1
754 viremia (Blips) and drug resistance in patients receiving HAART. *JAMA*.
755 2005;293(7):817-29.
- 756 10. Bailey JR, Sedaghat AR, Kieffer T, Brennan T, Lee PK, Wind-Rotolo M, et al. Residual
757 human immunodeficiency virus type 1 viremia in some patients on antiretroviral therapy
758 is dominated by a small number of invariant clones rarely found in circulating CD4+ T
759 cells. *J Virol*. 2006;80(13):6441-57.
- 760 11. Dinoso JB, Kim SY, Wiegand AM, Palmer SE, Gange SJ, Cranmer L, et al. Treatment
761 intensification does not reduce residual HIV-1 viremia in patients on highly active
762 antiretroviral therapy. *Proc Natl Acad Sci U S A*. 2009;106(23):9403-8.
- 763 12. Gandhi RT, Zheng L, Bosch RJ, Chan ES, Margolis DM, Read S, et al. The effect of
764 raltegravir intensification on low-level residual viremia in HIV-infected patients on
765 antiretroviral therapy: a randomized controlled trial. *PLoS Med*. 2010;7(8).
- 766 13. Tobin NH, Learn GH, Holte SE, Wang Y, Melvin AJ, McKernan JL, et al. Evidence that
767 low-level viremias during effective highly active antiretroviral therapy result from two
768 processes: expression of archival virus and replication of virus. *J Virol*.
769 2005;79(15):9625-34.
- 770 14. Simonetti FR, Sobolewski MD, Fyne E, Shao W, Spindler J, Hattori J, et al. Clonally
771 expanded CD4+ T cells can produce infectious HIV-1 in vivo. *Proc Natl Acad Sci U S A*.
772 2016;113(7):1883-8.
- 773 15. Halvas EK, Joseph KW, Brandt LD, Guo S, Sobolewski MD, Jacobs JL, et al. HIV-1
774 viremia not suppressible by antiretroviral therapy can originate from large T cell clones
775 producing infectious virus. *J Clin Invest*. 2020;130(11):5847-57.
- 776 16. Lorenzi JC, Cohen YZ, Cohn LB, Kreider EF, Barton JP, Learn GH, et al. Paired
777 quantitative and qualitative assessment of the replication-competent HIV-1 reservoir and
778 comparison with integrated proviral DNA. *Proc Natl Acad Sci U S A*.
779 2016;113(49):E7908-E16.

17. Maldarelli F, Wu X, Su L, Simonetti FR, Shao W, Hill S, et al. HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science*. 2014;345(6193):179-83.
18. Bui JK, Sobolewski MD, Keele BF, Spindler J, Musick A, Wiegand A, et al. Proviruses with identical sequences comprise a large fraction of the replication-competent HIV reservoir. *PLOS Pathogens*. 2017;13(3):e1006283.
19. Hosmane NN, Kwon KJ, Bruner KM, Capoferri AA, Beg S, Rosenbloom DI, et al. Proliferation of latently infected CD4(+) T cells carrying replication-competent HIV-1: Potential role in latent reservoir dynamics. *J Exp Med*. 2017;214(4):959-72.
20. Lee GQ, Orlova-Fink N, Einkauf K, Chowdhury FZ, Sun X, Harrington S, et al. Clonal expansion of genome-intact HIV-1 in functionally polarized Th1 CD4+ T cells. *J Clin Invest*. 2017;127(7):2689-96.
21. McMahon D, Jones J, Wiegand A, Gange SJ, Kearney M, Palmer S, et al. Short-course raltegravir intensification does not reduce persistent low-level viremia in patients with HIV-1 suppression during receipt of combination antiretroviral therapy. *Clin Infect Dis*. 2010;50(6):912-9.
22. Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DI, et al. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell*. 2013;155(3):540-51.
23. Bruner KM, Murray AJ, Pollack RA, Soliman MG, Laskey SB, Capoferri AA, et al. Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nature Medicine*. 2016;22(9):1043-9.
24. Bruner KM, Wang Z, Simonetti FR, Bender AM, Kwon KJ, Sengupta S, et al. A quantitative approach for measuring the reservoir of latent HIV-1 proviruses. *Nature*. 2019;566(7742):120-5.
25. Musick A, Spindler J, Boritz E, Perez L, Crespo-Velez D, Patro SC, et al. HIV Infected T Cells Can Proliferate in vivo Without Inducing Expression of the Integrated Provirus. *Front Microbiol*. 2019;10:2204.
26. Pollack RA, Jones RB, Perteza M, Bruner KM, Martin AR, Thomas AS, et al. Defective HIV-1 Proviruses Are Expressed and Can Be Recognized by Cytotoxic T Lymphocytes, which Shape the Proviral Landscape. *Cell Host Microbe*. 2017;21(4):494-506 e4.
27. Sannier G, Dube M, Dufour C, Richard C, Brassard N, Delgado GG, et al. Combined single-cell transcriptional, translational, and genomic profiling reveals HIV-1 reservoir diversity. *Cell Rep*. 2021;36(9):109643.
28. Einkauf KB, Osborn MR, Gao C, Sun W, Sun X, Lian X, et al. Parallel analysis of transcription, integration, and sequence of single HIV-1 proviruses. *Cell*. 2022;185(2):266-82 e15.
29. Imamichi H, Dewar RL, Adelsberger JW, Rehm CA, O'Doherty U, Paxinos EE, et al. Defective HIV-1 proviruses produce novel protein-coding RNA species in HIV-infected patients on combination antiretroviral therapy. *Proc Natl Acad Sci U S A*. 2016;113(31):8783-8.
30. Imamichi H, Smith M, Adelsberger JW, Izumi T, Scrimieri F, Sherman BT, et al. Defective HIV-1 proviruses produce viral proteins. *Proc Natl Acad Sci U S A*. 2020;117(7):3704-10.
31. Fisher K, Wang XQ, Lee A, Morcilla V, de Vries A, Lee E, et al. Plasma-Derived HIV-1 Virions Contain Considerable Levels of Defective Genomes. *J Virol*. 2022;96(6):e0201121.
32. Ho Y-C, Shan L, Hosmane Nina N, Wang J, Laskey Sarah B, Rosenbloom Daniel IS, et al. Replication-Competent Noninduced Proviruses in the Latent Reservoir Increase Barrier to HIV-1 Cure. *Cell*. 2013;155(3):540-51.

33. Cole B, Lambrechts L, Gantner P, Noppe Y, Bonine N, Witkowski W, et al. In-depth single-cell analysis of translation-competent HIV-1 reservoirs identifies cellular sources of plasma viremia. *Nat Commun.* 2021;12(1):3727.
34. Gremminger T, Song Z, Ji J, Foster A, Weng K, and Heng X. Extended Interactions between HIV-1 Viral RNA and tRNA(Lys3) Are Important to Maintain Viral RNA Integrity. *Int J Mol Sci.* 2020;22(1).
35. Weeks KM, Ampe C, Schultz SC, Steitz TA, and Crothers DM. Fragments of the HIV-1 Tat protein specifically bind TAR RNA. *Science.* 1990;249(4974):1281-5.
36. Stoltzfus CM, and Madsen JM. Role of viral splicing elements and cellular RNA binding proteins in regulation of HIV-1 alternative RNA splicing. *Curr HIV Res.* 2006;4(1):43-55.
37. Ding P, Kharytonchik S, Waller A, Mbaekwe U, Basappa S, Kuo N, et al. Identification of the initial nucleocapsid recognition element in the HIV-1 RNA packaging signal. *Proc Natl Acad Sci U S A.* 2020;117(30):17737-46.
38. Skripkin E, Paillart JC, Marquet R, Ehresmann B, and Ehresmann C. Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *Proc Natl Acad Sci U S A.* 1994;91(11):4945-9.
39. Ding P, Kharytonchik S, Kuo N, Cannistraci E, Flores H, Chaudhary R, et al. 5'-Cap sequestration is an essential determinant of HIV-1 genome packaging. *Proc Natl Acad Sci U S A.* 2021;118(37).
40. Hiener B, Horsburgh BA, Eden J-S, Barton K, Schlub TE, Lee E, et al. Identification of Genetically Intact HIV-1 Proviruses in Specific CD4+ T Cells from Effectively Treated Participants. *Cell Reports.* 2017;21(3):813-22.
41. Pinzone MR, VanBelzen DJ, Weissman S, Bertuccio MP, Cannon L, Venanzi-Rullo E, et al. Longitudinal HIV sequencing reveals reservoir expression leading to decay which is obscured by clonal expansion. *Nature Communications.* 2019;10(1):1-12.
42. Simonetti FR, Zhang H, Soroosh GP, Duan J, Rhodehouse K, Hill AL, et al. Antigen-driven clonal selection shapes the persistence of HIV-1-infected CD4+ T cells in vivo. *J Clin Invest.* 2021;131(3).
43. Jordan MR, Kearney M, Palmer S, Shao W, Maldarelli F, Coakley EP, et al. Comparison of standard PCR/cloning to single genome sequencing for analysis of HIV-1 populations. *J Virol Methods.* 2010;168(1-2):114-20.
44. Wang Z, Gurule EE, Brennan TP, Gerold JM, Kwon KJ, Hosmane NN, et al. Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane. *Proceedings of the National Academy of Sciences.* 2018;115(11):E2575-E84.
45. Bozzi G, Simonetti FR, Watters SA, Anderson EM, Gouzoulis M, Kearney MF, et al. No evidence of ongoing HIV replication or compartmentalization in tissues during combination antiretroviral therapy: Implications for HIV eradication. *Sci Adv.* 2019;5(9):eaav2045.
46. Laird GM, Eisele EE, Rabi SA, Lai J, Chioma S, Blankson JN, et al. Rapid Quantification of the Latent Reservoir for HIV-1 Using a Viral Outgrowth Assay. *PLOS Pathogens.* 2013;9(5):e1003398.
47. Onafuwa-Nuga A, and Telesnitsky A. The remarkable frequency of human immunodeficiency virus type 1 genetic recombination. *Microbiol Mol Biol Rev.* 2009;73(3):451-80, Table of Contents.
48. Keane SC, Heng X, Lu K, Kharytonchik S, Ramakrishnan V, Carter G, et al. RNA structure. Structure of the HIV-1 RNA packaging signal. *Science.* 2015;348(6237):917-21.
49. Pasternak AO, DeMaster LK, Kootstra NA, Reiss P, O'Doherty U, and Berkhout B. Minor Contribution of Chimeric Host-HIV Readthrough Transcripts to the Level of HIV Cell-Associated gag RNA. *J Virol.* 2016;90(2):1148-51.

- 880 50. Sertznig H, Hillebrand F, Erkelenz S, Schaal H, and Widera M. Behind the scenes of
881 HIV-1 replication: Alternative splicing as the dependency factor on the quiet. *Virology*.
882 2018;516:176-88.
- 883 51. Ocwieja KE, Sherrill-Mix S, Mukherjee R, Custers-Allen R, David P, Brown M, et al.
884 Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment
885 and long-read sequencing. *Nucleic Acids Res*. 2012;40(20):10345-55.
- 886 52. Emery A, Zhou S, Pollom E, and Swanstrom R. Characterizing HIV-1 Splicing by Using
887 Next-Generation Sequencing. *J Virol*. 2017;91(6).
- 888 53. Shen N, Jette L, Liang C, Wainberg MA, and Laughrea M. Impact of human
889 immunodeficiency virus type 1 RNA dimerization on viral infectivity and of stem-loop B
890 on RNA dimerization and reverse transcription and dissociation of dimerization from
891 packaging. *J Virol*. 2000;74(12):5729-35.
- 892 54. Shen N, Jette L, Wainberg MA, and Laughrea M. Role of stem B, loop B, and
893 nucleotides next to the primer binding site and the kissing-loop domain in human
894 immunodeficiency virus type 1 replication and genomic-RNA dimerization. *J Virol*.
895 2001;75(21):10543-9.
- 896 55. van Bel N, Das AT, Cornelissen M, Abbink TE, and Berkhout B. A short sequence motif
897 in the 5' leader of the HIV-1 genome modulates extended RNA dimer formation and virus
898 replication. *J Biol Chem*. 2014;289(51):35061-74.
- 899 56. Shan L, Rabi SA, Laird GM, Eisele EE, Zhang H, Margolick JB, et al. A novel PCR assay
900 for quantification of HIV-1 RNA. *J Virol*. 2013;87(11):6521-5.
- 901 57. Yukl SA, Kaiser P, Kim P, Telwatte S, Joshi SK, Vu M, et al. HIV latency in isolated
902 patient CD4(+) T cells may be due to blocks in HIV transcriptional elongation,
903 completion, and splicing. *Sci Transl Med*. 2018;10(430).
- 904 58. Schur FK, Hagen WJ, Rumlova M, Ruml T, Muller B, Krausslich HG, et al. Structure of
905 the immature HIV-1 capsid in intact virus particles at 8.8 Å resolution. *Nature*.
906 2015;517(7535):505-8.
- 907 59. Thyagarajan B, Faul J, Vivek S, Kim JK, Nikolich-Zugich J, Weir D, et al. Age-Related
908 Differences in T-Cell Subsets in a Nationally Representative Sample of People Older
909 Than Age 55: Findings From the Health and Retirement Study. *J Gerontol A Biol Sci*
910 *Med Sci*. 2022;77(5):927-33.
- 911 60. Bacchus-Souffan C, Fitch M, Symons J, Abdel-Mohsen M, Reeves DB, Hoh R, et al.
912 Relationship between CD4 T cell turnover, cellular differentiation and HIV persistence
913 during ART. *PLoS Pathog*. 2021;17(1):e1009214.
- 914 61. Chomont N, El-Far M, Ancuta P, Trautmann L, Procopio FA, Yassine-Diab B, et al. HIV
915 reservoir size and persistence are driven by T cell survival and homeostatic proliferation.
916 *Nature Medicine*. 2009;15(8):893-900.
- 917 62. Grossman Z, Singh NJ, Simonetti FR, Lederman MM, Douek DC, Deeks SG, et al.
918 'Rinse and Replace': Boosting T Cell Turnover To Reduce HIV-1 Reservoirs. *Trends*
919 *Immunol*. 2020;41(6):466-80.
- 920 63. Mendoza P, Jackson JR, Oliveira TY, Gaebler C, Ramos V, Caskey M, et al. Antigen-
921 responsive CD4+ T cell clones contribute to the HIV-1 latent reservoir. *J Exp Med*.
922 2020;217(7).
- 923 64. Gantner P, Pagliuzza A, Pardons M, Ramgopal M, Routy JP, Fromentin R, et al. Single-
924 cell TCR sequencing reveals phenotypically diverse clonally expanded cells harboring
925 inducible HIV proviruses during ART. *Nat Commun*. 2020;11(1):4089.
- 926 65. Collora JA, Liu R, Pinto-Santini D, Ravindra N, Ganoza C, Lama JR, et al. Single-cell
927 multiomics reveals persistence of HIV-1 in expanded cytotoxic T cell clones. *Immunity*.
928 2022;55(6):1013-31 e7.

929 66. Bailey JR, Zhang H, Wegweiser BW, Yang HC, Herrera L, Ahonkhai A, et al. Evolution
930 of HIV-1 in an HLA-B*57-positive patient during virologic escape. *J Infect Dis.*
931 2007;196(1):50-5.

932 67. Bailey JR, O'Connell K, Yang HC, Han Y, Xu J, Jilek B, et al. Transmission of human
933 immunodeficiency virus type 1 from a patient who developed AIDS to an elite
934 suppressor. *J Virol.* 2008;82(15):7395-410.

935 68. Bailey JR, Brennan TP, O'Connell KA, Siliciano RF, and Blankson JN. Evidence of
936 CD8+ T-cell-mediated selective pressure on human immunodeficiency virus type 1 nef in
937 HLA-B*57+ elite suppressors. *J Virol.* 2009;83(1):88-97.

938 69. Cillo AR, Vagratian D, Bedison MA, Anderson EM, Kearney MF, Fyne E, et al. Improved
939 single-copy assays for quantification of persistent HIV-1 viremia in patients on
940 suppressive antiretroviral therapy. *J Clin Microbiol.* 2014;52(11):3944-51.

941 70. Persaud D, Siberry GK, Ahonkhai A, Kajdas J, Monie D, Hutton N, et al. Continued
942 production of drug-sensitive human immunodeficiency virus type 1 in children on
943 combination antiretroviral therapy who have undetectable viral loads. *J Virol.*
944 2004;78(2):968-79.

945 71. Palmer S, Wiegand AP, Maldarelli F, Bazmi H, Mican JM, Polis M, et al. New real-time
946 reverse transcriptase-initiated PCR assay with single-copy sensitivity for human
947 immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol.* 2003;41(10):4531-6.

948 72. Duette G, Hiener B, Morgan H, Mazur FG, Mathivanan V, Horsburgh BA, et al. The HIV-
949 1 proviral landscape reveals that Nef contributes to HIV-1 persistence in effector
950 memory CD4+ T cells. *J Clin Invest.* 2022;132(7).

951 73. Rassler S, Ramirez R, Khoury N, Skowron G, and Sahu GK. Prolonged persistence of a
952 novel replication-defective HIV-1 variant in plasma of a patient on suppressive therapy.
953 *Virol J.* 2016;13(1):157.

954 74. Vibholm LK, Lorenzi JCC, Pai JA, Cohen YZ, Oliveira TY, Barton JP, et al.
955 Characterization of Intact Proviruses in Blood and Lymph Node from HIV-Infected
956 Individuals Undergoing Analytical Treatment Interruption. *J Virol.* 2019;93(8).

957 75. Abramowicz A, and Gos M. Correction to: Splicing mutations in human genetic
958 disorders: examples, detection, and confirmation. *J Appl Genet.* 2019;60(2):231.

959 76. Gholizadeh Z, Iqbal MS, Li R, and Romerio F. The HIV-1 Antisense Gene ASP: The
960 New Kid on the Block. *Vaccines (Basel).* 2021;9(5).

961 77. Das AT, Pasternak AO, and Berkhout B. On the generation of the MSD- class of
962 defective HIV proviruses. *Retrovirology.* 2019;16(1):19.

963 78. Heng X, Kharytonchyk S, Garcia EL, Lu K, Divakaruni SS, LaCotti C, et al. Identification
964 of a minimal region of the HIV-1 5'-leader required for RNA dimerization, NC binding,
965 and packaging. *J Mol Biol.* 2012;417(3):224-39.

966 79. Emery A, and Swanstrom R. HIV-1: To Splice or Not to Splice, That Is the Question.
967 *Viruses.* 2021;13(2).

968 80. Antar AA, Jenike KM, Jang S, Rigau DN, Reeves DB, Hoh R, et al. Longitudinal study
969 reveals HIV-1-infected CD4+ T cell dynamics during long-term antiretroviral therapy. *J*
970 *Clin Invest.* 2020;130(7):3543-59.

971 81. Peluso MJ, Bacchetti P, Ritter KD, Beg S, Lai J, Martin JN, et al. Differential decay of
972 intact and defective proviral DNA in HIV-1-infected individuals on suppressive
973 antiretroviral therapy. *JCI Insight.* 2020;5(4).

974 82. Cho A, Gaebler C, Oliveira T, Ramos V, Saad M, Lorenzi JCC, et al. Longitudinal clonal
975 dynamics of HIV-1 latent reservoirs measured by combination quadruplex polymerase
976 chain reaction and sequencing. *Proc Natl Acad Sci U S A.* 2022;119(4).

977 83. White JA, Simonetti FR, Beg S, McMyn NF, Dai W, Bachmann N, et al. Complex decay
978 dynamics of HIV virions, intact and defective proviruses, and 2LTR circles following
979 initiation of antiretroviral therapy. *Proc Natl Acad Sci U S A.* 2022;119(6).

84. Ryscavage P, Kelly S, Li JZ, Harrigan PR, and Taiwo B. Significance and clinical management of persistent low-level viremia and very-low-level viremia in HIV-1-infected patients. *Antimicrob Agents Chemother.* 2014;58(7):3585-98.
85. Coffin JM, and Hughes SH. Clonal Expansion of Infected CD4+ T Cells in People Living with HIV. *Viruses.* 2021;13(10).
86. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature.* 2009;459(7243):108-12.
87. Jiang C, Lian X, Gao C, Sun X, Einkauf KB, Chevalier JM, et al. Distinct viral reservoirs in individuals with spontaneous control of HIV-1. *Nature.* 2020;585(7824):261-7.
88. Lian X, Gao C, Sun X, Jiang C, Einkauf KB, Seiger KW, et al. Signatures of immune selection in intact and defective proviruses distinguish HIV-1 elite controllers. *Sci Transl Med.* 2021;13(624):eabl4097.
89. Cole B, Lambrechts L, Boyer Z, Noppe Y, De Scheerder MA, Eden JS, et al. Extensive characterization of HIV-1 reservoirs reveals links to plasma viremia before and during analytical treatment interruption. *Cell Rep.* 2022;39(4):110739.
90. Kearney MF, Wiegand A, Shao W, Coffin JM, Mellors JW, Lederman M, et al. Origin of Rebound Plasma HIV Includes Cells with Identical Proviruses That Are Transcriptionally Active before Stopping of Antiretroviral Therapy. *J Virol.* 2016;90(3):1369-76.
91. Tosiano MA, Jacobs JL, Shutt KA, Cyktor JC, and Mellors JW. A Simpler and More Sensitive Single-Copy HIV-1 RNA Assay for Quantification of Persistent HIV-1 Viremia in Individuals on Suppressive Antiretroviral Therapy. *J Clin Microbiol.* 2019;57(3).
92. Einkauf KB, Lee GQ, Gao C, Sharaf R, Sun X, Hua S, et al. Intact HIV-1 proviruses accumulate at distinct chromosomal positions during prolonged antiretroviral therapy. *J Clin Invest.* 2019;129(3):988-98.
93. Shao W, Shan J, Hu WS, Halvas EK, Mellors JW, Coffin JM, et al. HIV Proviral Sequence Database: A New Public Database for Near Full-Length HIV Proviral Sequences and Their Meta-Analyses. *AIDS Res Hum Retroviruses.* 2020;36(1):1-3.
94. Wiegand A, Spindler J, Hong FF, Shao W, Cyktor JC, Cillo AR, et al. Single-cell analysis of HIV-1 transcriptional activity reveals expression of proviruses in expanded clones during ART. *Proc Natl Acad Sci U S A.* 2017;114(18):E3659-E68.
95. Nielsen M, Lundegaard C, Worning P, Lauemoller SL, Lamberth K, Buus S, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci.* 2003;12(5):1007-17.
96. Bertagnolli LN, Variale J, Sweet S, Brockhurst J, Simonetti FR, White J, et al. Autologous IgG antibodies block outgrowth of a substantial but variable fraction of viruses in the latent reservoir for HIV-1. *Proc Natl Acad Sci U S A.* 2020;117(50):32066-77.
97. Shen L, Peterson S, Sedaghat AR, McMahon MA, Callender M, Zhang H, et al. Dose-response curve slope sets class-specific limits on inhibitory potential of anti-HIV drugs. *Nat Med.* 2008;14(7):762-6.

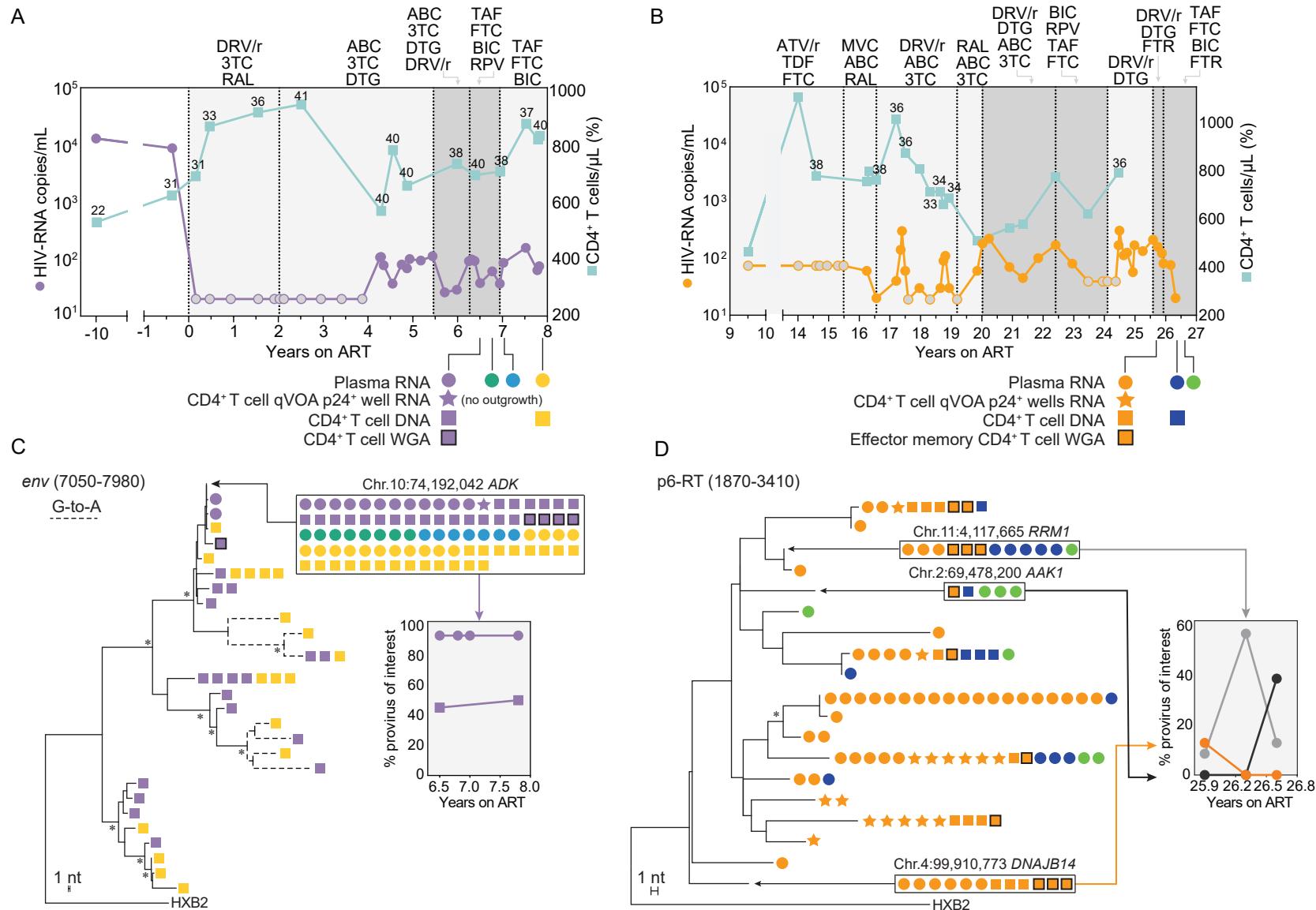


Figure 1. Clinical history of two study participants with NSV, and analysis of HIV-1 populations in plasma and CD4+ T cells. (A and B) Plasma HIV-1 RNA and CD4+ T cell counts over time for P1 and P2; grey circles indicate values below the limit of quantification; numbers above squares indicate CD4+ T cell percentage; light grey areas indicate standard ART; dark grey areas indicate ART intensification; **(C)** Maximum likelihood tree analysis of *env* single genome sequences from P1; dashed branches indicate sequences with hypermutation; tree nodes with bootstrap values above 80 are marked by star symbols; identical sequences matching proviruses with integration and full genome data are highlighted in boxes; chromosomal location is indicated above boxed area; frequencies of variants of interests over time are shown in the graph insert. **(D)** Maximum likelihood tree analysis of P6-RT single genome sequences from P2; only plasma and viral outgrowth RNA sequences are shown, together with matching proviral DNA sequences (the complete tree is shown in Figure S3). 3TC lamivudine, ABC abacavir, FTC emtricitabine, TDF tenofovir disoproxil fumarate, TAF tenofovir alafenamide, DRV/r darunavir-ritonavir, ATV/r atazanavir-ritonavir, RAL raltegravir, DTG dolutegravir, BIC bictegravir, MVC maraviroc, RPV rilpivirine, FTR fostemsavir.

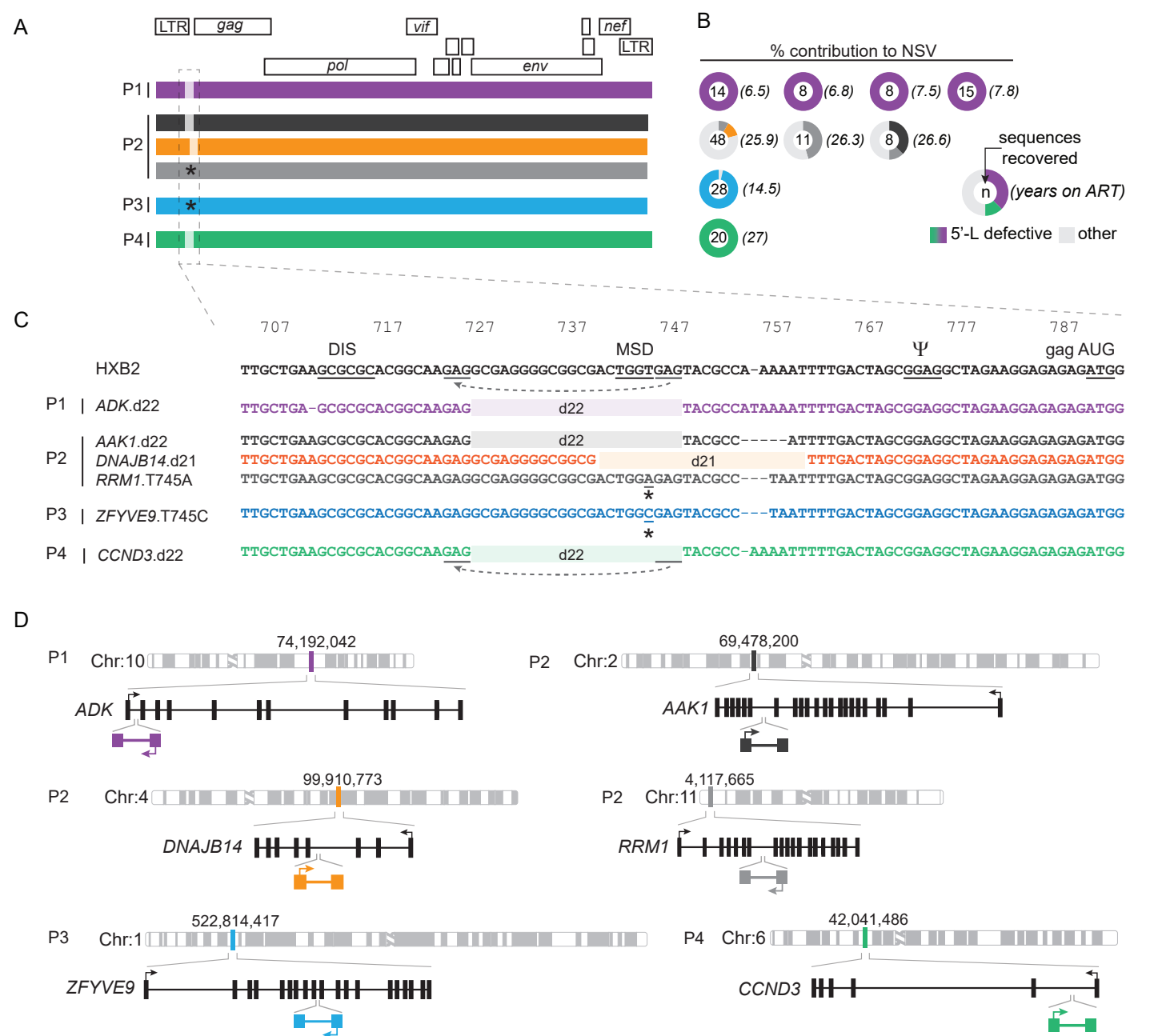


Figure 2. Paired full genome sequencing and integration site analysis of 5'-Leader-defective proviruses causing NSV. (A) Mapped sequences of proviruses contributing to plasma HIV-1 RNA in each participant; proviruses of interest are color-coded throughout the figure; light-shaded areas indicate small 5'-L deletions; star symbols indicate point mutations affecting the major splicing donor site (MSD). (B) Percentage contribution to NSV of defective proviruses of interest (C) 5'-Leader defects aligned to the HXB2 reference; DIS dimerization initiation signal, MSD major splicing donor, PSI packaging signal, AUG Gag start codon; dashes indicate length polymorphisms; grey lines highlight GAG repeats at deletion junctions causing misplaced jumping of reverse transcriptase (dashed grey arrows). (D) Chromosomal and gene locations of proviruses causing viremia; sets of arrows indicate direction of proviral and host gene transcription; schematic gene tracks are shown in black, with vertical bars representing exons.

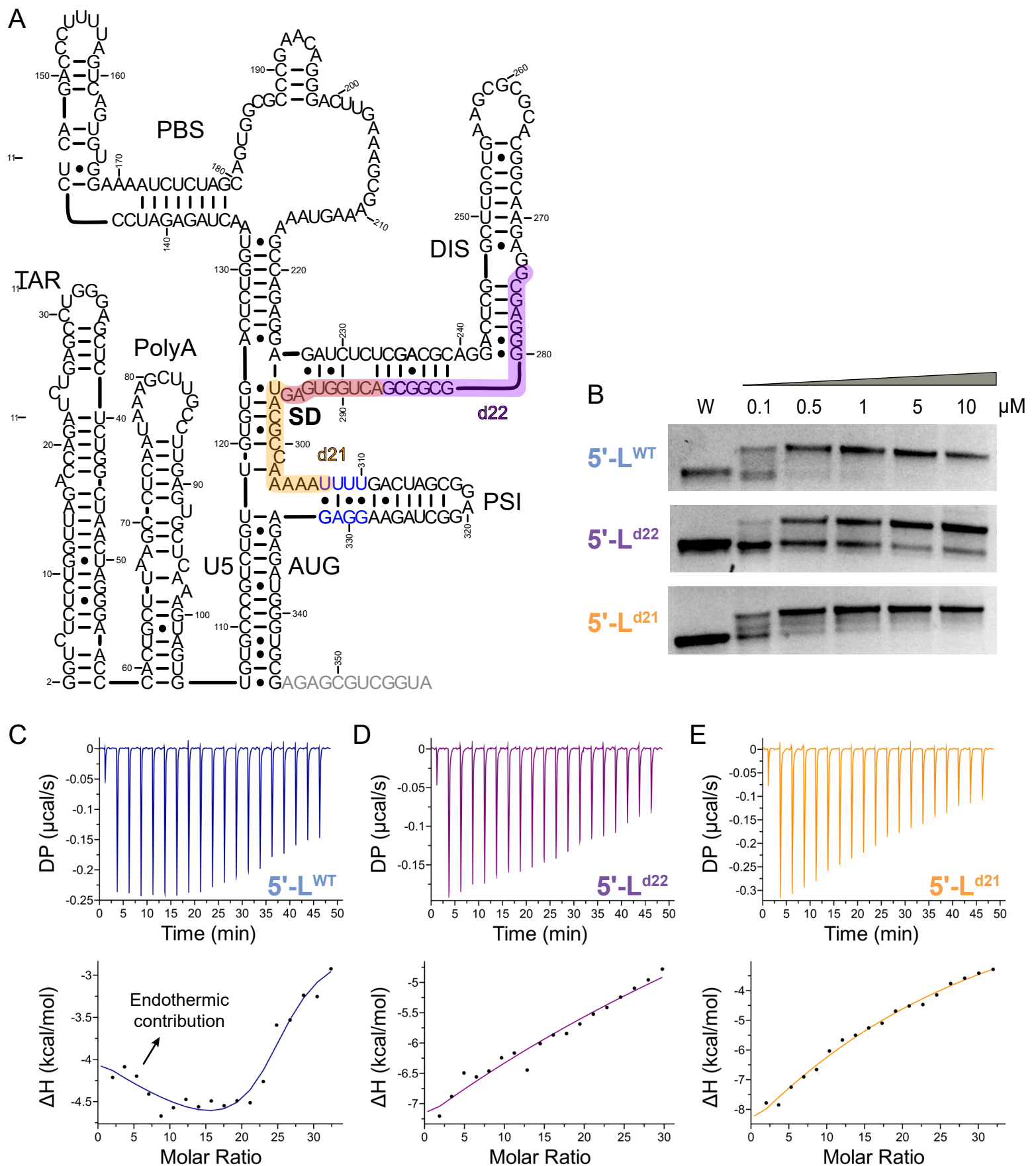


Figure 3. 5'-Leader deletions alter dimerization propensities and nucleocapsid binding properties. (A) Secondary structure of the HIV-1 NL4-3 gRNA 5'-Leader with patient deletions indicated in purple (d22) and orange (d21), with a region of overlap within the splice donor (SD) shown in red. High affinity binding sites with endothermic contribution on ITC binding within the psi hairpin are indicated with blue text. Grey text indicates the portion of AUG truncated to better study the dimer and its initial binding sites by ITC. **(B)** Concentration dependent dimerization assays of the full leader show that the wildtype and d21 constructs maintain similar dimerization propensities, while the d22 variant exhibit reduced dimerization. **(C, D, E)** ITC isotherms for the truncated dimeric 5'-Leader titrated with low protein-to-RNA ratios. WT exhibits previously described initial binding with an endothermic contribution **(C)** that is not seen for the d22 **(D)** or d21 **(E)** 5'-Leader constructs.

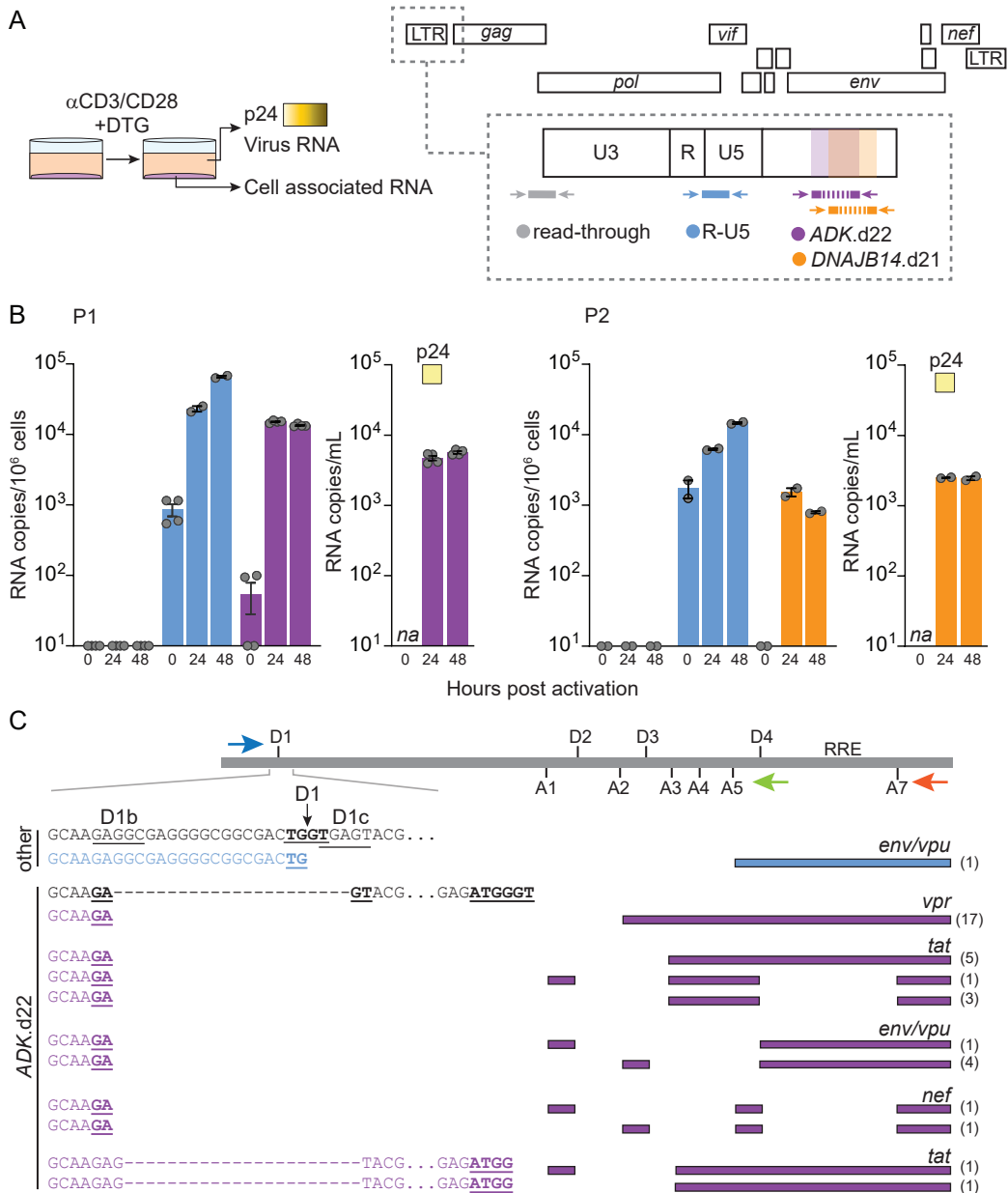


Figure 4. 5'-Leader-defective proviruses are inducible, their genomic RNA is packaged, and can use alternative splicing donors. (A) CD4⁺ T cells from P1 and P2 were cultured for 48 hours in the presence on dolutegravir (DTG) and anti-CD3/CD28 beads; cells and supernatants were collected at 0, 24 and 48 hours; the right panel shows the location of primers and probes used to quantify viral RNA; shaded areas indicate 5'-Leader deletion used to measure provirus-specific transcription. (B) Mean levels of read-through, total and provirus-specific RNA detected in cells and supernatant upon T-cell activation; grey dots represent digital PCR replicate reactions; error bars indicate standard error of the mean; (C) Singly and multiply-spliced transcripts amplified at limiting dilution from cells at 48 hours; arrows indicate primer locations; ticks indicate the locations of splicing donors and acceptors in the HIV-1 genome; major and alternative splicing donor sites are labeled in black; mapped splicing junctions are underlined and in bold; 22-nucleotide deletion is represented by dashed lines; numbers in parentheses indicate the number of sequences recovered for each type of spliced variant.

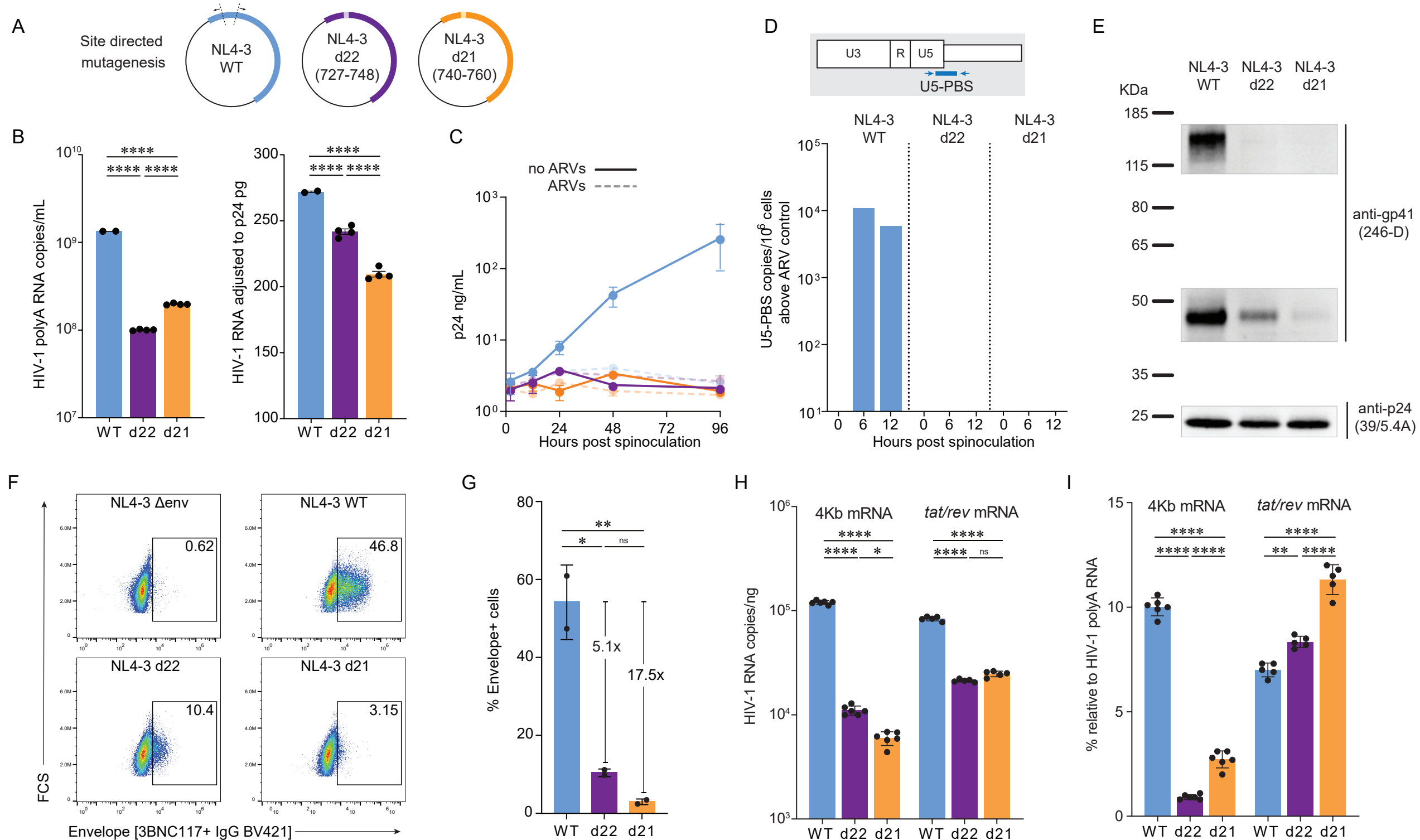
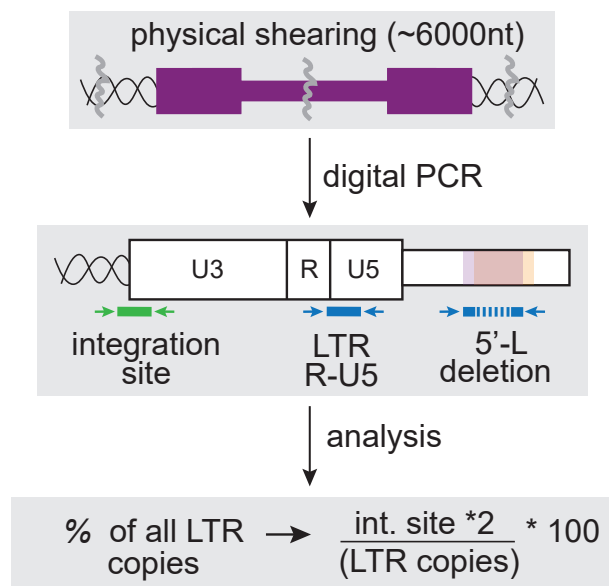
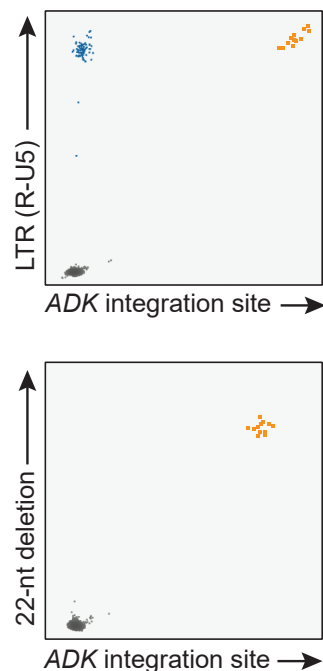


Figure 5. 5'-Leader deletions lead to non-infectious particles lacking Envelope incorporation. (A) Deletions found in proviruses causing viremia were introduced in an NL4-3 expression plasmid by site-directed mutagenesis; deletion start and end positions relative to HXB2 are indicated in parentheses. (B) Copies of HIV-1 RNA recovered at 72 hours post transfection of 293T cells, expressed as copies per mL (left) or normalized by p24 pg/mL (right). (C) Spinoculation of primary CD4⁺ T cells shows exponential increase in p24 levels only with wildtype NL4-3, in the absence of antiretrovirals (ARVs: TDF, FTC, DTG). (D) Reverse transcription was assessed by measuring late cDNA products by ddPCR targeting the U5-PBS junction; primary CD4⁺ T cells were collected at 0, 6, and 12 hours post spinoculation with and without ARVs; U5-PBS copies detected in the presence of ARVs, result of incomplete DNaseI digestion of plasmid carry over from transfection, were subtracted from copies detected in conditions without ARVs. (E) Virus produced upon 293T transfection was pelleted by ultracentrifugation, lysed, normalized by p24 and used for Western blots with primary antibodies specific to p24 and gp41. (F) Surface staining of HIV-1 Envelope on 293T cells 24 hours post transfection. (G) Frequency of Envelope positive cells transfected with wild type versus 5'-L deletions; fold reduction relative to WT is indicated above each mutant; results from two transfection experiments are shown, each dot represents the average of 2 technical replicates. (H) Quantification of cell-associated spliced HIV-1 transcripts belonging to the 4kb class or tat/rev mRNA normalized to RNA ng. (I) Percentage of spliced transcripts relative to total HIV-1 polyA RNA. Error bars indicate standard error of the mean (panels B, H and I), or standard deviation (panels C and F). Statistical significance between conditions was tested by one-way ANOVA.

A



B



C

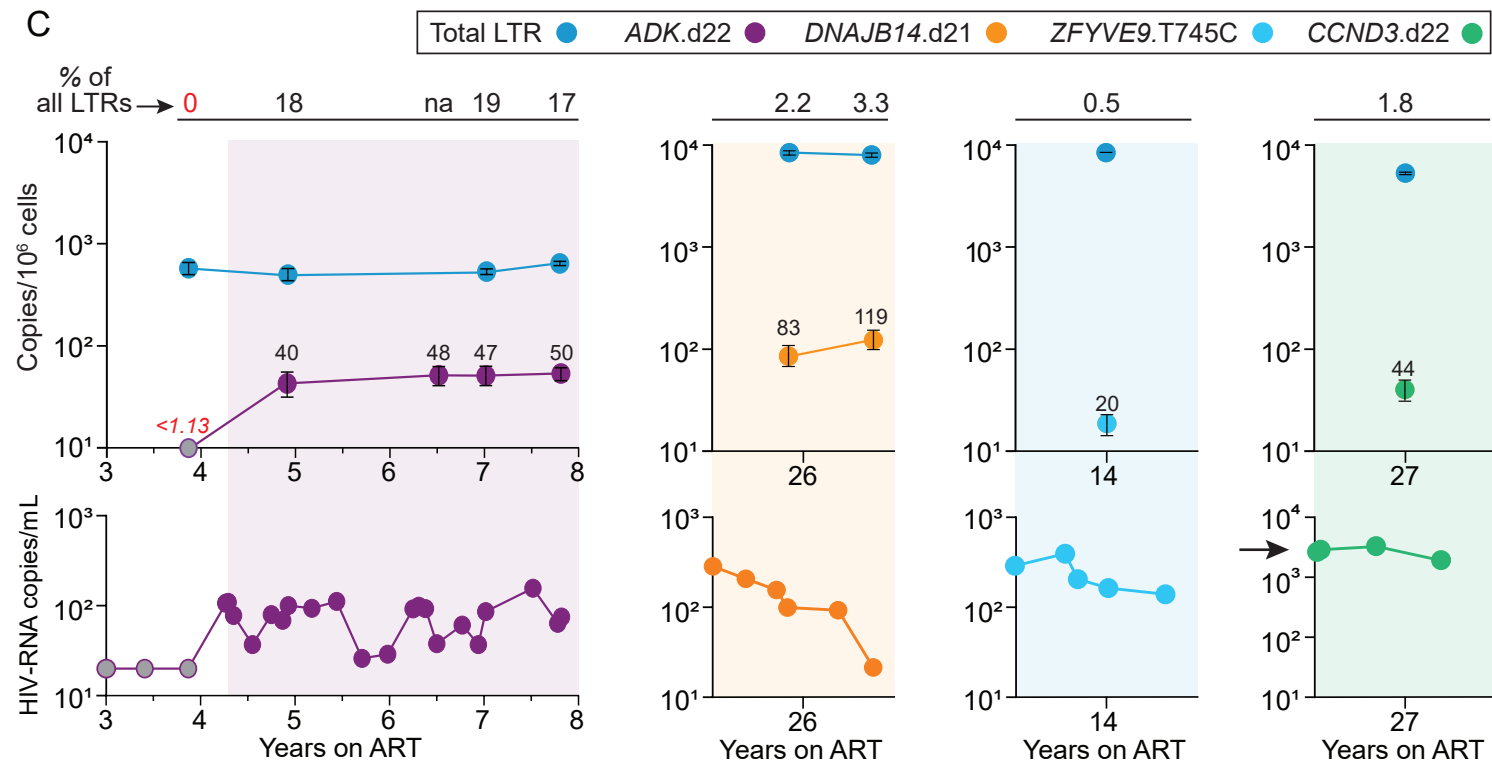


Figure 6. Proviruses source of plasma virus are stable over time and show clonal expansion concurrent with onset of NSV. (A) Experimental approach: genomic DNA is sheared to obtain fragments shorter than 6000nt; total LTR copies are quantified with primers and probe targeting both R-U5 junctions, while proviruses of interest are quantified targeting the site of HIV-1 integration or the 5'-Leader deletion. (B) Representative ddPCR 2D plots of *ADK.d22*-specific assays; due to the proximity of the integration site and 5' the R-U5 junction, most proviruses of interest are double positive. (C) Longitudinal quantification of total LTR copies and proviruses contributing to viremia; numbers above symbols indicate mean copies in each sample; numbers in red indicate limit of detection (grey symbols); error bars indicate the standard error of the mean; shaded areas represent time with plasma HIV-1 RNA above the limit of quantification; black arrow highlights higher viremia in P4; na, not available.

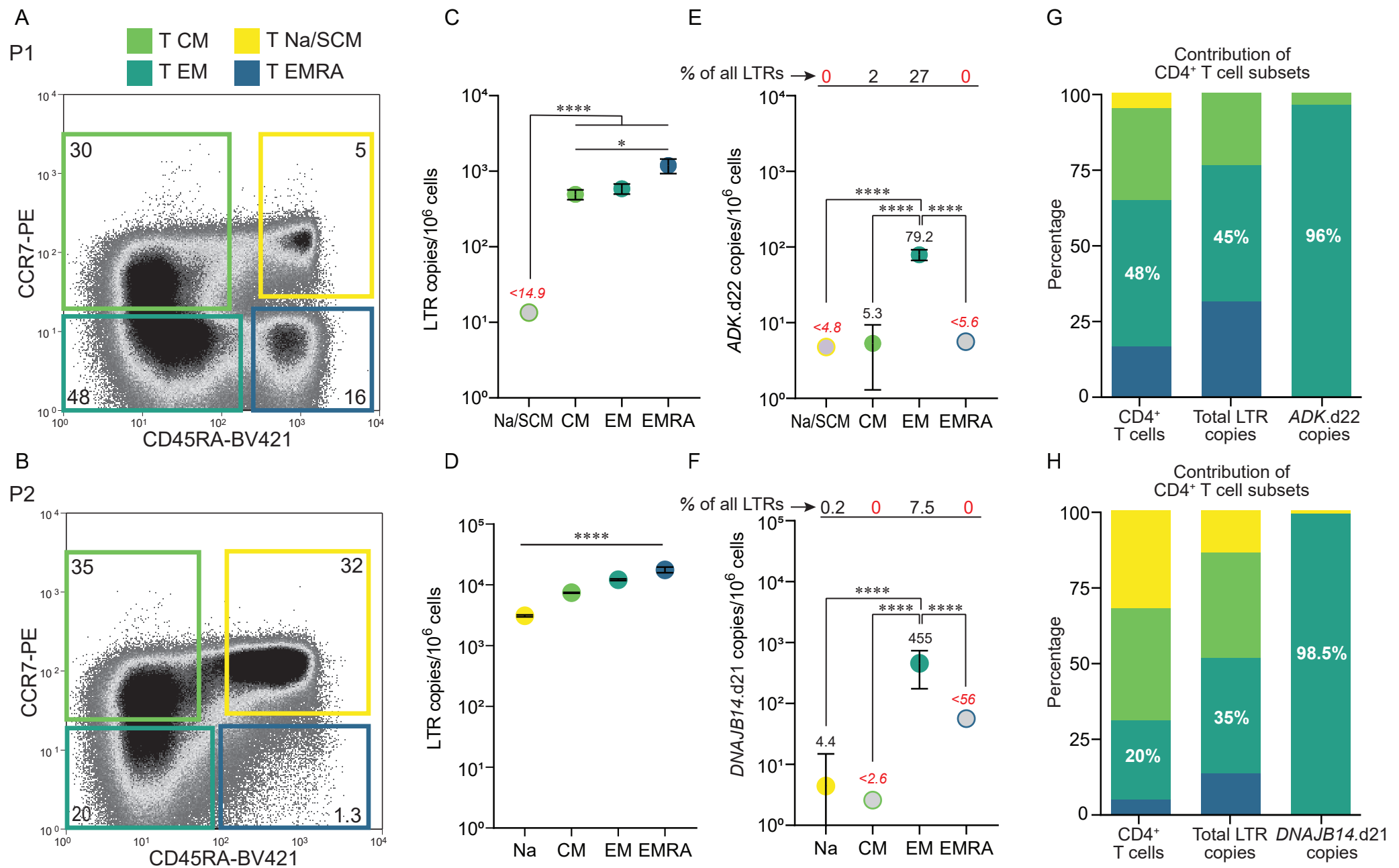


Figure 7. Cells carrying the *ADK.d22* and *DNAJB14.21* proviruses are compartmentalized in Effector Memory cells. (A-B) Distribution of CD3⁺CD4⁺ live cells based on CD45RA and CCR7 surface markers; colored boxes represent sorting gates and numbers indicate the percentage of events in each gate. **(C-D)** Frequency of total LTR copies among sorted CD4 subsets. **(E-F)** Frequency of *ADK.d22* and *DNAJB14.d21* among sorted CD4 subsets; numbers above symbols indicate mean values; percentage of all LTRs represented by the two proviruses are displayed above the graph. **(G-H)** Contribution of subsets to CD4⁺ T cells, total LTR copies and clones carrying the provirus of interest. Grey symbols indicate values below the limit of detection; error bars indicate standard error of the mean; statistical significance of differences among sorted populations was tested by one-way ANOVA. Samples were collected at 7.8 and 26.6 years on ART, respectively.

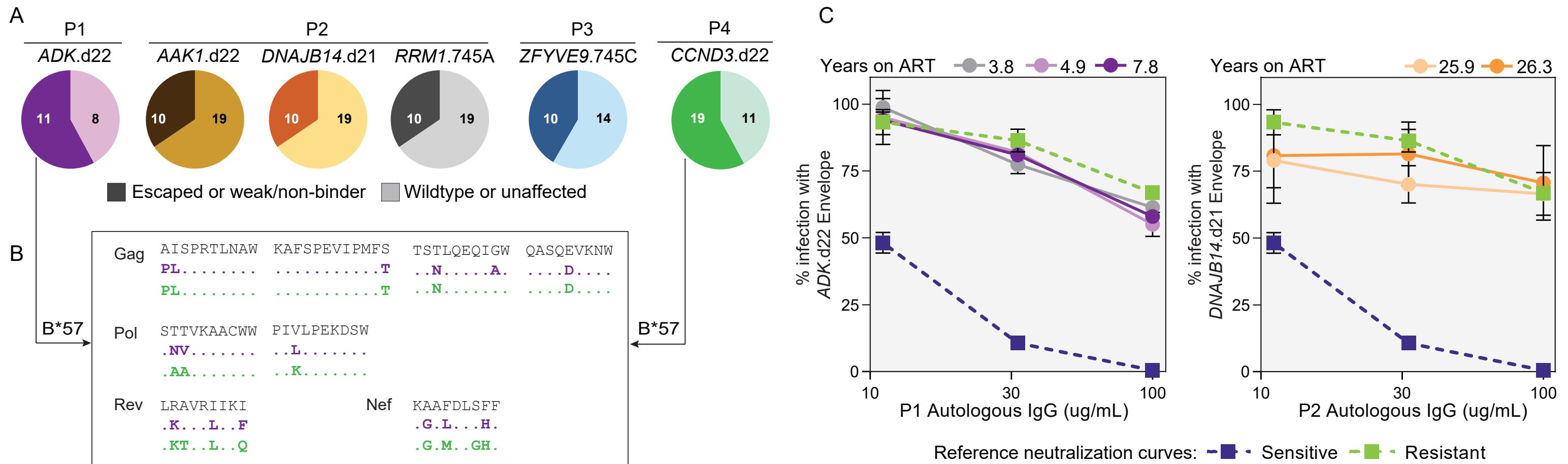


Figure 8. Proviruses contributing to viremia show variable proportion of CTL escape mutations and are resistant to autologous neutralization. (A) Distribution of T cell epitopes from full proviral genomes; documented escape mutations and predicted weak or non-binders are indicated in darker shades (left section of the pie charts); numbers within pie charts represent the number of epitopes analyzed in each category. (B) Representative epitopes restricted by HLA-B*57 showing mutations with documented impact on HLA binding. (C) Neutralization experiments with autologous IgGs and viruses pseudotyped with Envelopes from proviruses of interest; error bars indicate standard deviations of three technical replicates; neutralization curves of stereotypical sensitive (blue) and resistant (green) Envelopes are displayed for reference.