

Defective HIV-1 proviruses produce novel protein-coding RNA species in HIV-infected patients on combination antiretroviral therapy

Hiromi Imamichi^a, Robin L. Dewar^b, Joseph W. Adelsberger^b, Catherine A. Rehm^a, Una O'Doherty^c, Ellen E. Paxinos^d, Anthony S. Fauci^{a,1}, and H. Clifford Lane^a

^aLaboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892; ^bClinical Services Program, Applied and Development Research Directorate, Leidos Biomedical Research, Inc., Frederick, MD 21072; ^cDepartment of Pathology and Laboratory Medicine, University of Pennsylvania School of Medicine, Philadelphia, PA 19104; and ^dApplications and Collaborations, Pacific Biosciences, Menlo Park, CA 94025

Contributed by Anthony S. Fauci, June 9, 2016 (sent for review April 23, 2016; reviewed by Scott Hammer and Nelson L. Michael)

Despite years of plasma HIV-RNA levels <40 copies per milliliter during combination antiretroviral therapy (cART), the majority of HIV-infected patients exhibit persistent seropositivity to HIV-1 and evidence of immune activation. These patients also show persistence of proviruses of HIV-1 in circulating peripheral blood mononuclear cells. Many of these proviruses have been characterized as defective and thus thought to contribute little to HIV-1 pathogenesis. By combining 5'LTR-to-3'LTR single-genome amplification and direct amplicon sequencing, we have identified the presence of "defective" proviruses capable of transcribing novel unspliced HIV-RNA (usHIV-RNA) species in patients at all stages of HIV-1 infection. Although these novel usHIV-RNA transcripts had exon structures that were different from those of the known spliced HIV-RNA variants, they maintained translationally competent ORFs, involving elements of *gag*, *pol*, *env*, *rev*, and *nef* to encode a series of novel HIV-1 chimeric proteins. These novel usHIV-RNAs were detected in five of five patients, including four of four patients with prolonged viral suppression of HIV-RNA levels <40 copies per milliliter for more than 6 y. Our findings suggest that the persistent defective proviruses of HIV-1 are not "silent," but rather may contribute to HIV-1 pathogenesis by stimulating host-defense pathways that target foreign nucleic acids and proteins.

HIV-1 | provirus | activation | latency

Although once considered a terminal illness, HIV-1 infection has now become a chronic manageable disease. With the use of combination antiretroviral therapy (cART), prolonged viral suppression of plasma HIV-RNA levels (pVL) < 40 copies per milliliter is achievable in the majority of patients with HIV-1 infection (1). One paradoxical observation is that, despite prolonged viral suppression with no evidence of active viral replication, the majority of HIV-infected patients exhibit persistent presence of HIV-1 proviruses (2–4), persistent seropositivity to HIV-1 (5), and evidence of immune activation (6–10). The only setting in which this has not been the case has been following bone marrow transplantation in which seronegativity has been reported in association with a loss of peripheral blood proviral DNA (11, 12). Over 90% of proviruses in the peripheral blood are thought to be "defective" by having lethal genetic alterations that include G-to-A hypermutations (13, 14) and small insertions/deletions (indels) that disrupt ORFs (13), or large internal deletions (13, 15, 16). Because these defective proviruses are unable to encode intact viruses, the peripheral blood pool of proviruses has been thought to largely represent a silent "graveyard" of viral sequences. In a single case, we previously reported the ability of a provirus with a stop codon in the protease to transcribe viral RNA (14). Cells harboring these "defective" proviruses were found in a clonally expanded population of effector memory CD4⁺ T cells that had persisted for 17 y. To better evaluate the characteristics of these persistent proviruses, we developed a system that allows simultaneous isolation of genomic DNA and cytoplasmic RNA from the same populations of CD4⁺ T cells,

amplification of up to 8.9 kb HIV-1 DNA and 8.4 kb cell-associated unspliced HIV-RNA (usHIV-RNA), and sequencing of the resulting near full-length HIV-1 genome fragments. This approach allowed us to better characterize the genetic variability in HIV-1 proviral genomes and to determine which "defective" proviruses were transcribed and capable of encoding viral proteins.

Results

Characterization of HIV-1 Proviruses in Patients with HIV-1 Infection. The 5'LTR-to-3'LTR single genome amplification and direct amplicon sequencing of HIV-1 proviruses were performed for four patients with pVL \geq 40 copies per milliliter (range 26,758–225,658); four patients with pVL < 40 copies per milliliter for >6 y (range 6.1–11.9 y); and one patient before and after suppressive cART. Proviral DNA lengths were estimated based on migration distances of the 5'LTR-to-3'LTR PCR-amplified DNA fragments on agarose gels and confirmed by sequencing (Fig. 1 A–C). For the four patients with pVL \geq 40 copies per milliliter, 42 of 48 (88%) full-length proviral genomes encoded for an intact virus. In contrast, for the four patients with pVL < 40 copies per milliliter, full-length genomes were found in only one patient [patient (Pt) 8]. In that patient, six of the seven sequenced

Significance

The presence of "defective" HIV-1 proviruses in HIV-infected patients has been well documented. The current consensus view of the "defective" proviruses is that these are dead-end products that do not give rise to progeny virus and thus collectively represent a "graveyard" of viruses. We describe the presence of defective HIV-1 proviruses capable of transcribing novel unspliced HIV-RNA species in HIV-infected patients on combination antiretroviral therapy. We propose that the proviruses persistently present in combination antiretroviral therapy-treated patients are not defective in a conventional sense, but rather represent incomplete forms of proviruses encoding translationally competent HIV-RNA transcripts. Strategies directed toward curing HIV-1 infection and eliminating the state of persistent immune activation need to include approaches designed to eliminate cells harboring such proviruses.

Author contributions: H.I., C.A.R., U.O., E.E.P., A.S.F., and H.C.L. designed research; H.I., R.L.D., J.W.A., and H.C.L. performed research; H.I., R.L.D., J.W.A., C.A.R., U.O., E.E.P., A.S.F., and H.C.L. analyzed data; and H.I., A.S.F., and H.C.L. wrote the paper.

Reviewers: S.H., Columbia University; and N.L.M., US Military HIV Research Program.

The authors declare no conflict of interest.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. KU677989–KU678196).

¹To whom correspondence should be addressed. Email: afauci@niaid.nih.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1609057113/-DCSupplemental.

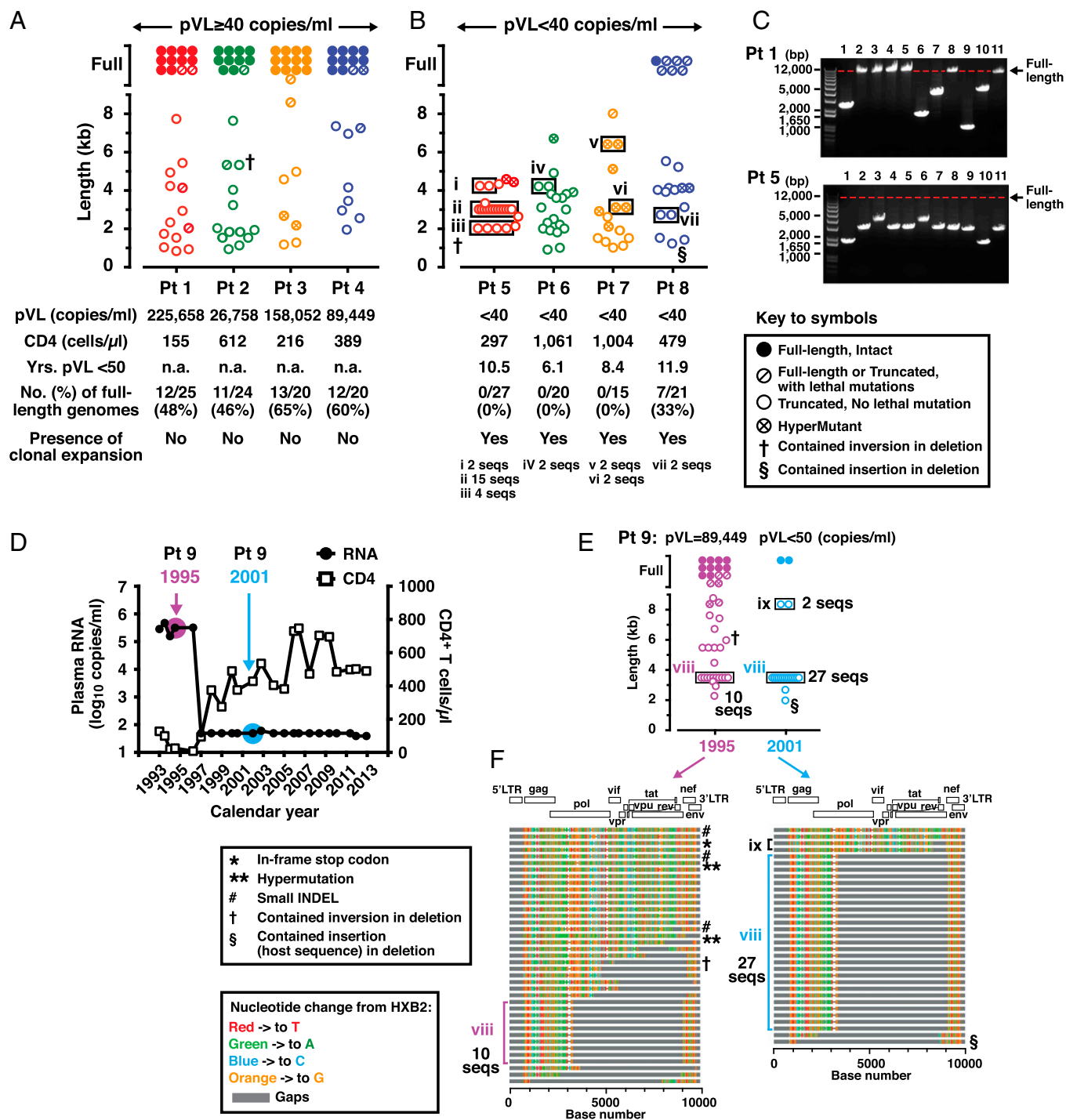


Fig. 1. Characterization of full-length and truncated proviruses in patients with HIV-1 infection. (A) Distribution of proviral DNA lengths for four patients with pVL \geq 40 copies per milliliter and (B) four patients with pVL < 40 copies per milliliter. Descriptions of symbols for all panels are provided in the key. Expanded proviral clones are depicted by the black rectangles. Distinct clonally expanded populations are identified by Roman numerals. (C) Agarose gel pictures depicting sizes of proviral DNA PCR fragments for one patient with pVL \geq 40 copies per milliliter (Pt 1) and one patient with pVL < 40 copies per milliliter (Pt 5) generated using the 5'LTR-to-3'LTR PCR. Only positive PCR reactions generated at limiting dilutions are shown. The red dotted lines indicate the migration point of DNA bands corresponding to full-length HIV-1 proviruses (8.9 kb). (D) Longitudinal changes over 20 y in CD4⁺ T-cell counts and HIV-RNA levels for a single patient with HIV-1 infection (Pt 9, first tested positive for HIV-1 in 1985). (E) Distribution of HIV-1 proviral DNA lengths and clonal expansions for Pt 9 in 1995 and 2001. Descriptions of symbols for the panels are provided in the key in A and B. (F) The Highlighter program (v2.2.3) analyses of the sequence data from 1995 and 2001 for HIV-1 proviral DNA derived from patient 9. Of note, identical expanded proviral clones (viii) were noted in 1995 and 2001.

proviruses were defective with lethal mutations in their genomes (Fig. 1B). The higher proportion of full-length proviruses observed in patients with pVL \geq 40 copies per milliliter likely

reflects the presence of substantial ongoing viral replication that was taking place in these patients at the time of sampling. Of note, such truncated proviruses were not detected when either a

plasmid DNA containing the full-length NL4.3 strain of HIV-1 or genomic DNA from the chronically infected 8E5 cell line were used as PCR templates (Fig. S1), thus minimizing the possibility that these findings were a result of PCR artifacts. However, these truncated proviruses could be detected as early as 24 h following in vitro infection of primary CD4⁺ T cells with the DH12 clone of HIV-1 (Fig. S2).

The presence of clonally expanded defective provirus populations was evident in all four patients with pVL < 40 copies per milliliter (Fig. 1B, clones i–vii). Of particular note, the 3.0-kb clone ii proviruses (Fig. 1B) accounted for 56% (15 of 27) of the total number of proviruses found in patient 5. Clonally expanded proviruses were not detected in any of the four patients with pVL ≥ 40 copies per milliliter. We believe this dichotomy likely reflects the relative proportions of intact and truncated proviruses in these two subpopulations of patients.

To directly determine the effects of therapy and to evaluate changes in the distribution of HIV-1 proviral DNA lengths over time, samples from a single patient (patient 9) from 1995 and 2001 were examined (Fig. 1D–F). Before initiation of cART and in the setting of a higher viral load (pVL = 89,449 copies per milliliter), the frequency of full-length proviruses was 36% (14 of 39). Six years later and 3 y after achieving pVL < 50 copies per milliliter, this percentage had decreased to 6% (2 of 33). This finding is consistent with the cross-sectional data presented in Fig. 1A and B in that the frequencies of full-length proviruses were higher when there was evidence of active viral replication.

Deletion Junction Analysis of HIV-1 Proviruses. Among the truncated proviruses derived from the nine patients, deletion sizes ranging from 251 bp to 8,033 bp in length were identified (Fig. S3A). Approximately 40% of deletions were associated with directly repeated sequences at the deletion junctions, suggesting that they had occurred during the negative-strand synthesis phase of reverse transcription (Fig. S3) (15, 17, 18). In one instance in patient 5 and two instances in patient 9 (Fig. S3A), these deletions included the 15-nt stretch of purine-rich sequence referred to as the polypurine tract (PPT) that is located in the pol gene (cPPT) and the *nef/U3* region of HIV-1 (3'PPT) (19). The mechanisms responsible for the remaining 60% of deletions are unknown. Of note, in patient 9, we identified a provirus with a “deletion-insertion-deletion” with insertion of host sequence *B4GALT5* (Fig. 1E and F). “Deletion-inversion-deletion” complex genome structures were found in patient 2 (Fig. 1A and Fig. S3A), patient 9 (Fig. 1E and F), and patient 5 (Fig. 1B and Fig. S3A). Sequence analyses of these proviruses indicated that incorporation of the host sequence by readthrough of HIV-1 polyadenylation signal transcription (20–22) or a single polymerase jump during reverse transcription (15, 17, 18) were unlikely mechanisms to explain the unique “insertion or inversion in deletions” genome structures we observed. Taken together, these findings highlight the fact that multiple mechanisms may be responsible for alterations in the HIV-1 proviral genome. As noted below, some of these alterations may lead to unique ORFs.

As seen in Fig. S3A, many of the sequenced truncated proviruses contained, at a minimum, a short stretch of the 5' end of the *gag-pol* intronic region and a part of the *Nef* coding region, although lacking the second intronic region encoding most of the accessory proteins and *Env* (Fig. S4). If transcription occurred from any of these truncated proviruses, the transcripts would resemble unspliced (genomic) HIV-RNA with removal of the second intronic region.

Characterization of Novel usHIV-1 RNA Species. To determine whether or not any of these “defective” proviruses led to the production of novel usHIV-RNA that could lead to the production of novel proteins in the setting of prolonged viral suppression, we performed limiting-dilution single-genome amplification and directly sequenced the resulting single-genome amplicons. Nested PCR reactions were performed with two forward PCR primers located immediately

downstream of the major 5' splice donor site (D1) within the *gag-pol* intronic region (Fig. S4). These custom-designed PCR primers enabled us to selectively amplify unspliced forms of HIV-RNA species of up to 8.4 kb in size from patient 1 (pVL = 225,638 copies per milliliter) and patients 5–8 (pVL < 40 copies per milliliter). The distribution of cell-associated usHIV-RNA lengths are shown in Fig. 2A and plotted alongside the length distributions of the proviral DNA genomes derived from the same populations of CD4⁺ T cells.

For the patient with clear evidence of substantial active virus replication (Pt 1 in Fig. 2A and Fig. S5), 12 of 25 (48%) of HIV-1 proviruses were full-length, as were 8 of 11 (73%) of the cell-associated usHIV-RNA transcripts. In contrast, for the four patients (Pts 5–8) with pVL < 40 copies per milliliter, full-length proviruses were found in only one patient (Pt 8) and full-length usHIV-RNA species were not detected in any of these patients (Fig. 2A and Fig. S5).

The in vivo presence of clonally expanded proviral populations has been reported by ourselves and others (14, 23–26). However, there is minimal evidence for transcription of defective proviruses (14, 27). Here, we were able to identify the presence of clonally expanded defective provirus populations and to demonstrate ongoing transcription of these expanded “defective” proviral clones in two different patients (Pts 5 and 7) with pVL < 40 copies per milliliter (clones ii and vi in Fig. 2A and Fig. S5). In two additional instances (pairs “a” and “b”) in the same two patients (Pts 5 and 7), we were able to identify proviral DNA–usHIV-RNA pairs, providing strong evidence that, at best, some of these “defective” proviruses were responsible for the production of novel usHIV-RNA species in patients with “undetectable” plasma viremia (Fig. 2A and Fig. S5).

Sequence analyses of these novel usHIV-RNA transcripts revealed exon structures that were very different from those of the known spliced HIV-1 mRNA variants and they ranged in size from 0.5 kb to 2.6 kb (Fig. 2B). These novel usHIV-RNA transcripts had appropriate transcription start and stop signals and thus possessed translationally competent ORFs. At a minimum, a part of the *gag* and *nef* genes were retained in all of these truncated novel usHIV-RNA transcripts. Some also included elements of *pol*, *env*, and *rev* to encode a series of novel HIV-1 chimeric proteins (Fig. 2C).

As noted in Fig. 2B, the novel usHIV-RNA transcripts reported in this study did not contain any of the classic splice acceptor/donor sites in their genomes (28), nor did they contain cryptic splice sites that have been reported elsewhere (29, 30). These data support the hypothesis that the “defective” proviruses seen in patients with HIV-1 infection lead to the production of novel HIV-RNA transcripts faithfully transcribed from these proviruses.

All of the usHIV-RNA transcripts found in patients 5–8 were missing a gene segment encoding the Rev response element (RRE) (Fig. 2B). Two usHIV-RNA transcripts contained a partial *rev* gene without the RRE (2.0 kb in Pt 5 and 1.3 kb in Pt 6). In general, HIV-RNAs without any intact introns (i.e., fully spliced 2-kb HIV-RNA transcripts) or *cis*-acting repressive sequences (CRS) elements can be exported to the cytoplasm independent of Rev/RRE via the host cell TAP-dependent mRNA export pathway that mediates general mRNA export (28, 31) and are able to productively engage the host cell's translational machinery. In fact, a partial removal of the CRS elements (the ones within *pol*) has led to Rev/RRE-independent nuclear export of the transcripts and protein production, albeit at low levels in in vitro systems (32, 33). The fact that novel usHIV-RNAs reported in the present study do not contain any intact intronic regions (defined by splice donor/acceptor sites: D1-A1 and D2-A7) (Fig. 2B) or the CRS element within *pol* and thus are of “intron-free” structures might explain the Rev/RRE-independent nuclear export of these variants.

Assessment of the Pattern of HIV-1 Protein Production in Vivo. Taken together, these data support our hypothesis that patients with HIV-1 infection, despite plasma pVL < 40 copies per milliliter,

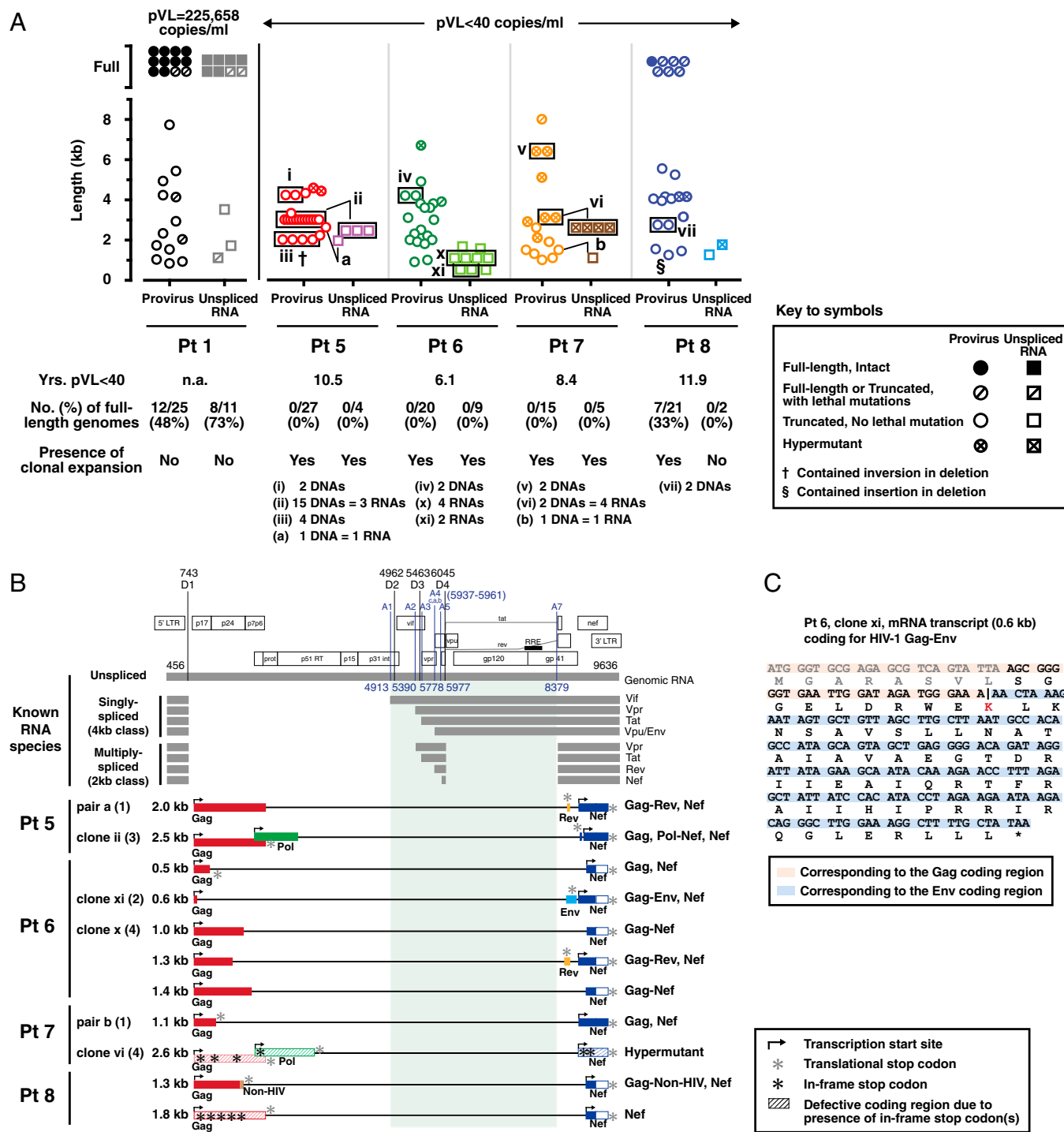


Fig. 2. Characterization of novel unspliced HIV-1 RNA species in patients with HIV-1 infection. (A) Distribution of usHIV-RNA lengths for one patient with pVL ≥ 40 copies per milliliter and four patients with pVL < 40 copies per milliliter. Plots for proviral DNA lengths are the same as those shown in Fig. 1 A and B and replotted here for comparison purposes. Descriptions of symbols for all panels are provided in the key. Clonally expanded populations are depicted by the black rectangles. Distinct expanded proviral clones are identified by Roman numerals. RNAs that had corresponding DNA sequences are identified by lowercase alphabets. (B) Schematic diagrams of novel usHIV-RNA transcripts for four patients with pVL < 40 copies per milliliter. Known unspliced and spliced HIV-RNA species are shown (*Upper*). Locations of the splice donors (D1–D4) and acceptors (A1–A7) for HIV-1 HXB2 are indicated in the HIV-1 genome structure at the top. Coordinates of the splice donors and acceptors are provided based on the sequence coordinates of HIV-1 HXB2. Descriptions of symbols used in the diagram are provided in the key. Only unique HIV-RNA sequences are shown. (C) Amino acid sequence of the ORF of the clone xi mRNA transcript that encodes a part of Gag (17 aa) and a part of Env (50 aa) derived from patient 6. The first 8 aa in the gag were deduced from the proviral DNA sequence from the same patient. The vertical black line indicates a boundary between Gag and Env coding regions.

may still be capable of producing viral proteins. This finding may help explain the persistence of antibodies to HIV-1 proteins greater than 10 y following initiation of suppressive cART (Fig. 3).

Of note, the stochastic element to the fading of bands suggest an asynchronous decay in the viral proteins and is consistent with the production and decay of novel chimeric proteins, which can be

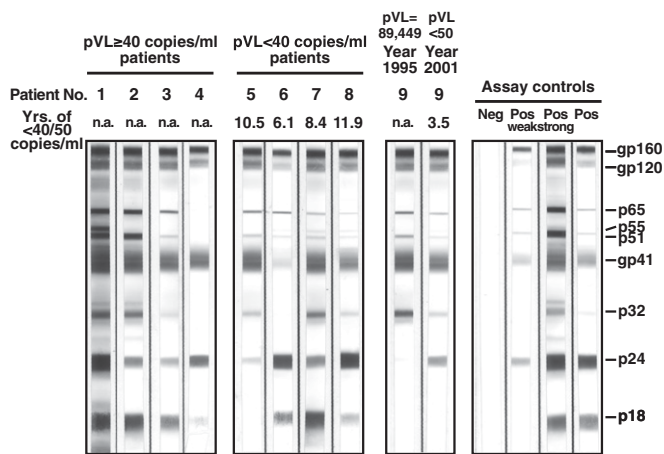


Fig. 3. HIV-1 Western blot analysis of four patients with pVL \geq 40 copies per milliliter (Pts 1–4), four patients with pVL < 40 copies per milliliter (Pts 5–8), and one patient (Pt 9) who exhibited clonal expansion of truncated defective provirus and persistence of the clone. HIV-1 Western blots were performed with plasma samples (diluted in 1:100). Negative and positive controls were run in parallel and are shown on the right panel. In patient 4, faint bands for p32, p51/55, and p65 proteins could be detected on the original strip but are not clearly shown in the scanned image.

seen by comparing the persistence of antibodies to Env in patient 5 and the relative persistence of antibodies to Gag in patient 6 with the uniform presence of antibodies to all viral proteins in patients 1–4.

Discussion

In conclusion, cells harboring defective proviral DNA capable of transcribing novel usHIV-RNAs with translationally competent ORFs can be seen in HIV-infected patients even despite a period of prolonged viral suppression. In the present study, we were able to detect matching “defective” proviral DNAs and RNA transcripts in four instances in two separate patients. The presence of transcription-competent “defective” proviruses is a consistent finding in patients at all stages of HIV-1 infection. We have referred to these proviruses as “zombie” proviruses (34), as they are not really “alive” given that they lack the ability to produce intact viruses but can inflict harm by producing foreign nucleic acids and proteins. It is not clear what fraction of cells harboring defective proviruses produce RNA transcripts. It is possible that differences in the activation status of the host-infected cells or epigenetic state of the HIV-1 proviruses might play a role. Persistence of these proviruses may explain the persistent seropositivity to HIV-1 and persistent immune activation seen in patients with “undetectable” virus. In addition to eliciting persistent antibody responses, HIV-1 proteins derived from “defective” proviruses could also be associated with CD4 and CD8 T-cell responses. Future studies will hopefully further delineate the effects of the transcription or translation of defective proviruses on immune responses and immune activation in cART-treated HIV-infected individuals with prolonged viral suppression. Strategies directed toward curing HIV-1 infection and eliminating the state of persistent immune activation associated with HIV-1 infection need to include approaches designed to eliminate cells harboring such proviruses.

Materials and Methods

Study Participants. All participants were enrolled in National Institute of Allergy and Infectious Diseases Institutional Review Board-approved HIV-1 clinical research protocols and provided written informed consent before study participation. All patients (Pts 1–9) were chronically infected with HIV-1 and had been known to be positive for HIV-1 for >10 y (range: 10–25 y) at

the time of sampling. Samples from patients 1–4 and patient 9 (1995 sample) were obtained from time points when these patients had HIV-RNA levels \geq 40 copies per milliliter. All patients (Pts 1–9) currently receive combination antiretroviral therapy.

Purification of CD4⁺ T Cells. Peripheral blood mononuclear cells (PBMCs) were obtained either from leukapheresis packs (Pts 1, 2, 4, 5, 7, 8, and Pt 9’s 2001 sample) or whole blood (Pts 3, 6, and patient 9’s 1995 sample). CD4⁺ T cells were isolated from PBMCs by negative selection using the CD4⁺ T-cell isolation kit II (Miltenyi Biotec).

Simultaneous Isolation of HIV-1 DNA and HIV-1 RNA. Total HIV-1 DNA and cell-associated HIV-1 RNA were simultaneously isolated from the same populations of CD4⁺ T cells. As depicted in Fig. S6, \sim 4–10 \times 10⁶ CD4⁺ T cells were lysed on ice for 10 min in CER I lysis buffer (NE-PER Nuclear and Cytoplasmic Extraction Reagents, Pierce) after addition of RNaseOUT recombinant ribonuclease inhibitor (final 1 unit per microliter; Thermo Fisher Scientific). Lysates were further incubated on ice for 1 min after addition of CER II lysis buffer. Lysates were centrifuged for 5 min at 16,000 \times g at 4 °C. The supernatants and the nuclei pellets were served as sources for cytoplasmic RNA and genomic DNA, respectively. A schematic diagram summarizing the major steps involved in these procedures is shown in Fig. S6. Cytoplasmic RNA was extracted from the supernatants using the RNeasy Mini kit (Qiagen) following the protocol (including an on-column DNase digestion step) provided by the manufacturer. Completeness of the fractionation of cytoplasmic RNAs was confirmed as shown in Fig. S7. Genomic DNA was extracted from the pelleted nuclei using the PicoPure kit (Thermo Fisher Scientific) according to the manufacturer’s protocol.

Amplification of Near Full-Length HIV-1 DNA. Genomic DNA was subjected to limiting dilution before amplification with a near full-length outer PCR followed by nested amplification with an inner PCR (Fig. S4 and Table S1). The PCR was performed with KAPA HiFi HotStart (KAPA Biosystems) with the initial denaturation at 95 °C for 2 min, followed by 30 cycles of denaturation at 98 °C for 20 s, annealing at 65 °C for 15 s, and extension at 72 °C for 6 min, with the final extension at 72 °C for 6 min. PCR reactions were performed in volumes of 25 or 50 μ L in 0.2 mL PCR microtubes. The final concentrations of both primers were 400 nM. The PCR products were purified with either 0.6 \times (a Bead:DNA ratio) AMPure XP beads (Beckman Coulter) or the BluePippin (Sage Science) using the 0.75%DF Marker 51 high-pass 4–10 kb vs.2 cassette definition with BP-Start: 6,700 and BP-end: 10,000, followed by purification with 1.0 \times AMPure XP beads (Beckman Coulter).

Amplification of Near Full-Length usHIV-1 RNA. cDNA was synthesized from RNA that was diluted to an endpoint with the SuperScript IV First-Strand cDNA Synthesis kit (Thermo Fisher Scientific) using an anchored oligo(dT)₂₀ primer that consists of a string of 20 deoxythymidylic acid residues followed by five nucleotides (TGAAG) complementary to the R region of the HIV-1 3’ LTR (Primer 5T25) (35). Single molecules of an \sim 8.4-kb fragment, encompassing near full-length unspliced HIV-1 RNA, obtained through limiting dilution, were PCR-amplified using the PCR primers listed in Fig. S4 and Table S1. For second-round PCR reactions, a primer pair, RNA F2 (796–816)/RNA R2 (9145–9171), was used in most cases. An alternate primer pair, RNA F2 (796–816)/RNA R2’ (9438–9458) was used in cases where the RNA R2 primer failed to yield an adequate band. All spliced (singly spliced and multiply spliced) HIV-1 mRNAs are spliced at the major 5’ splice donor site (D1, 743 in HXB2) and lack intron 1 (Fig. S4). Thus, selective amplification of usHIV-RNA was achieved through using forward PCR primers (RNA F1 and RNA F2) located immediately downstream of the 5’ major splice donor site (D1, 743 in HXB2). PCR conditions were the same as the ones used for amplification of near full-length HIV-1 DNA.

Sequencing. Single molecules were sequenced as near-full-length amplicons directly from PCR products using the 3500xL Genetic Analyzer (Applied Biosystems by Thermo Fisher Scientific) with a BigDye Terminator v3.1 Cycle Sequencing Kit. Single-molecule direct sequencing was also performed for a subset of patient 9 samples using a PacBio RS II instrument without shearing using commercially available chemistries and protocols (P4/C2, 180-min movies) or the MiSeq System (Illumina) with a Nextera XT DNA Library Preparation Kit. During postsequencing analyses of data, samples that contained more than one template (evidenced by having heterogeneous peaks on sequencing chromatograms or having degenerate sequences) were discarded and excluded from the further analyses.

ACKNOWLEDGMENTS. We thank R. Davey, C. Hadigan, J. Kovacs, F. Maldarelli, A. Pau, M. Sneller, and M. Wright for providing patient specimens; A. Ober, G. Roby, and C. Seamon for coordinating patient visits; H. Highbarger, A. Shah,

and I. Davis for running the Abbott HIV-1 assay; D.-W. Huang, R. Lempicki, and D. Liang for running the MiSeq assay; M. Brown, Y. Guo, and C. Ludka for performing the PacBio sequencing; and S. Hughes, T. Imamichi, and L. Dodd for helpful discussion. This work was funded through the intramural

research program of the National Institute of Allergy and Infectious Diseases of the National Institutes of Health, and in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract HHSN261200800001E.

- DHHS Panel on Antiretroviral Guidelines for Adults and Adolescents (2015) Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents. Department of Health and Human Services. Available at <https://aidsinfo.nih.gov/contentfiles/lvguidelines/adultandadolescentgl.pdf>. Accessed July 1, 2016.
- Murray JM, et al.; PINT Study Team (2014) HIV DNA subspecies persist in both activated and resting memory CD4+ T cells during antiretroviral therapy. *J Virol* 88(6):3516–3526.
- Josefsson L, et al. (2013) The HIV-1 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic changes over time. *Proc Natl Acad Sci USA* 110(51):E4987–E4996.
- Buzon MJ, et al. (2014) Long-term antiretroviral treatment initiated at primary HIV-1 infection affects the size, composition, and decay kinetics of the reservoir of HIV-1-infected CD4 T cells. *J Virol* 88(17):10056–10065.
- Mendoza D, et al. (2012) Comprehensive analysis of unique cases with extraordinary control over HIV replication. *Blood* 119(20):4645–4655.
- Kuller LH, et al.; INSIGHT SMART Study Group (2008) Inflammatory and coagulation biomarkers and mortality in patients with HIV infection. *PLoS Med* 5(10):e203.
- Sandler NG, et al.; INSIGHT SMART Study Group (2011) Plasma levels of soluble CD14 independently predict mortality in HIV infection. *J Infect Dis* 203(6):780–790.
- Wilson EM, et al.; Study to Understand the Natural History of HIV/AIDS in the Era of Effective Therapy (SUN Study) Investigators (2014) Monocyte-activation phenotypes are associated with biomarkers of inflammation and coagulation in chronic HIV infection. *J Infect Dis* 210(9):1396–1406.
- Hey-Cunningham WJ, et al.; PINT study team (2015) Early antiretroviral therapy with raltegravir generates sustained reductions in HIV reservoirs but not lower T-cell activation levels. *AIDS* 29(8):911–919.
- Catalfamo M, et al. (2008) HIV infection-associated immune activation occurs by two distinct pathways that differentially affect CD4 and CD8 T cells. *Proc Natl Acad Sci USA* 105(50):19851–19856.
- Hütter G, et al. (2009) Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med* 360(7):692–698.
- Henrich TJ, et al. (2013) Long-term reduction in peripheral blood HIV type 1 reservoirs following reduced-intensity conditioning allogeneic stem cell transplantation. *J Infect Dis* 207(11):1694–1702.
- Ho YC, et al. (2013) Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell* 155(3):540–551.
- Imamichi H, et al. (2014) Lifespan of effector memory CD4+ T cells determined by replication-incompetent integrated HIV-1 provirus. *AIDS* 28(8):1091–1099.
- Sanchez G, Xu X, Chermann JC, Hirsch I (1997) Accumulation of defective viral genomes in peripheral blood mononuclear cells of human immunodeficiency virus type 1-infected individuals. *J Virol* 71(3):2233–2240.
- Li Y, et al. (1991) Molecular characterization of human immunodeficiency virus type 1 cloned directly from uncultured human brain tissue: Identification of replication-competent and -defective viral genomes. *J Virol* 65(8):3973–3985.
- Temin HM (1993) Retrovirus variation and reverse transcription: Abnormal strand transfers result in retrovirus genetic variation. *Proc Natl Acad Sci USA* 90(15):6900–6903.
- Hu WS, Hughes SH (2012) HIV-1 reverse transcription. *Cold Spring Harb Perspect Med* 2(10):a006882.
- Rausch JW, Le Grice SF (2004) 'Binding, bending and bonding': Polypurine tract-primed initiation of plus-strand DNA synthesis in human immunodeficiency virus. *Int J Biochem Cell Biol* 36(9):1752–1766.
- Takebe Y, Telesnitsky A (2006) Evidence for the acquisition of multi-drug resistance in an HIV-1 clinical isolate via human sequence transduction. *Virology* 351(1):1–6.
- Pasternak AO, et al. (2015) Minor contribution of chimeric host-HIV readthrough transcripts to the level of HIV cell-associated gag RNA. *J Virol* 90(2):1148–1151.
- Han Y, et al. (2008) Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. *Cell Host Microbe* 4(2):134–146.
- Wagner TA, et al. (2014) HIV latency. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science* 345(6196):570–573.
- Maldarelli F, et al. (2014) HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science* 345(6193):179–183.
- Cohn LB, et al. (2015) HIV-1 integration landscape during latent and active infection. *Cell* 160(3):420–432.
- von Stockenström S, et al. (2015) Longitudinal genetic characterization reveals that cell proliferation maintains a persistent HIV type 1 DNA pool during effective HIV therapy. *J Infect Dis* 212(4):596–607.
- Kearney MF, et al. (2015) Origin of rebound plasma HIV includes cells with identical proviruses that are transcriptionally active before stopping of antiretroviral therapy. *J Virol* 90(3):1369–1376.
- Saliou JM, et al. (2009) Role of RNA structure and protein factors in the control of HIV-1 splicing. *Front Biosci (Landmark Ed)* 14:2714–2729.
- Ocvieja KE, et al. (2012) Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing. *Nucleic Acids Res* 40(20):10345–10355.
- Carrera C, Pinilla M, Pérez-Alvarez L, Thomson MM (2010) Identification of unusual and novel HIV type 1 spliced transcripts generated in vivo. *AIDS Res Hum Retroviruses* 26(7):815–820.
- Cullen BR (2005) Human immunodeficiency virus: Nuclear RNA export unwound. *Nature* 433(7021):26–27.
- Maldarelli F, Martin MA, Strebel K (1991) Identification of posttranscriptionally active inhibitory sequences in human immunodeficiency virus type 1 RNA: Novel level of gene regulation. *J Virol* 65(11):5732–5743.
- Cochrane AW, et al. (1991) Identification and characterization of intragenic sequences which repress human immunodeficiency virus structural gene expression. *J Virol* 65(10):5305–5313.
- Imamichi H, et al. (2014) Cryptic transcription of HIV-RNA species from "defective" proviruses: A novel pathway for persistent immune activation in patients with HIV-1 infection and mechanism for persistent seropositivity despite "undetectable" levels of virus. *20th International AIDS Conference, Melbourne, Australia, July 20–25, 2014*. Oral abstract WEA0206LB. Available at pag.aids2014.org/Abstracts.aspx?AID=11347. Accessed July 1, 2016.
- Shan L, et al. (2013) A novel PCR assay for quantification of HIV-1 RNA. *J Virol* 87(11):6521–6525.
- Lassen KG, Ramyar KX, Bailey JR, Zhou Y, Siliciano RF (2006) Nuclear retention of multiply spliced HIV-1 RNA in resting CD4+ T cells. *PLoS Pathog* 2(7):e68.
- Gaschen B, Kuiken C, Korber B, Foley B (2001) Retrieval and on-the-fly alignment of sequence fragments from the HIV database. *Bioinformatics* 17(5):415–418.
- Larsson A (2014) AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30(22):3276–3278.
- Salminen MO, et al. (1995) Recovery of virtually full-length HIV-1 provirus of diverse subtypes from primary virus cultures using the polymerase chain reaction. *Virology* 213(1):80–86.
- Carr JK, et al. (2001) The AG recombinant IbNG and novel strains of group M HIV-1 are common in Cameroon. *Virology* 286(1):168–181.
- Laukkanen T, et al. (2000) Virtually full-length subtype F and F/D recombinant HIV-1 from Africa and South America. *Virology* 269(1):95–104.
- Falcone EL, et al. (2013) Cerebrospinal fluid HIV-1 compartmentalization in a patient with AIDS and acute varicella-zoster virus meningomyelomalacia. *Clin Infect Dis* 57(5):e135–e142.