Supplementary Materials for:

Long-term persistence of transcriptionally-active "defective" HIV-1 proviruses: Implications for persistent immune activation during antiretroviral therapy

Kanal Singh[a], Ven Natarajan[b], Robin Dewar[b], Adam Rupert[b], Yuden Badralmaa[b], Tracey Zhai[a], Nicole Winchester[a], Francesca Scrimieri[b], Mindy Smith[a], Ivery Davis[b], Perrine Lallemand[b], Aude Giglietti[a], Jack Hensien[a], Thomas Buerkert[a], Bruktawit Goshu[a], Catherine A. Rehm[c], Zonghui Hu[d], H. Clifford Lane[a], and Hiromi Imamichi[a]

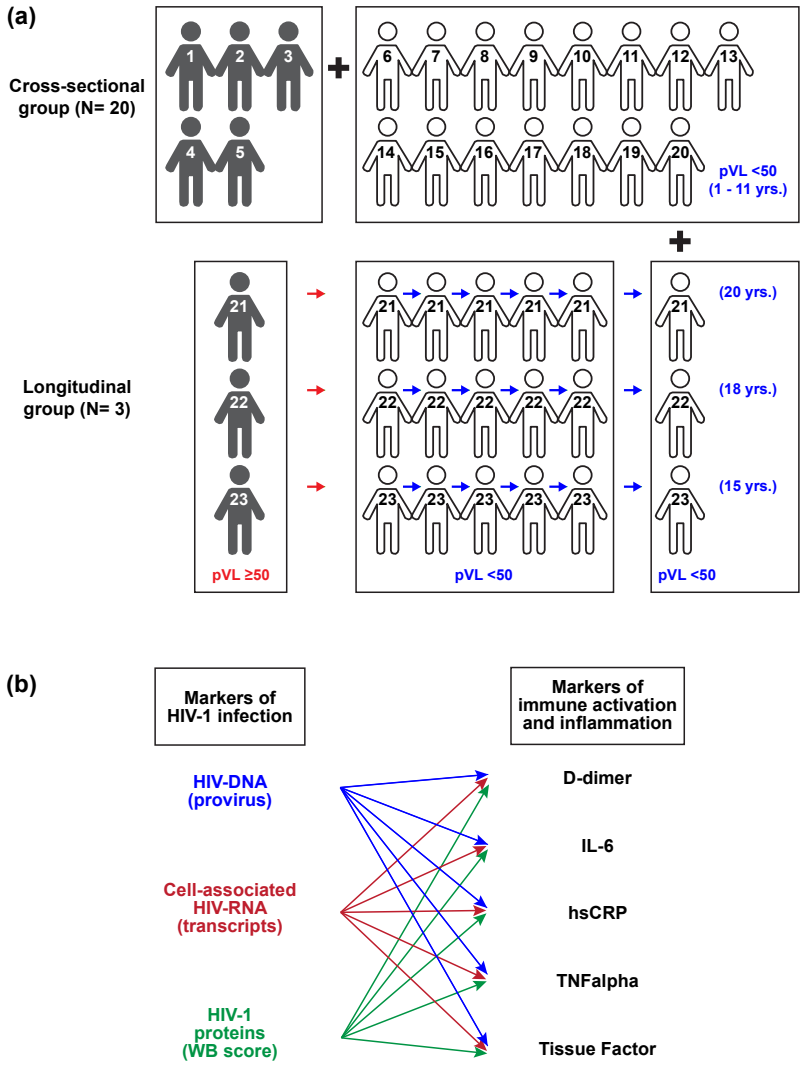
[a]Clinical and Molecular Retrovirology Section, Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, MD 20892, USA.

[b]Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA.

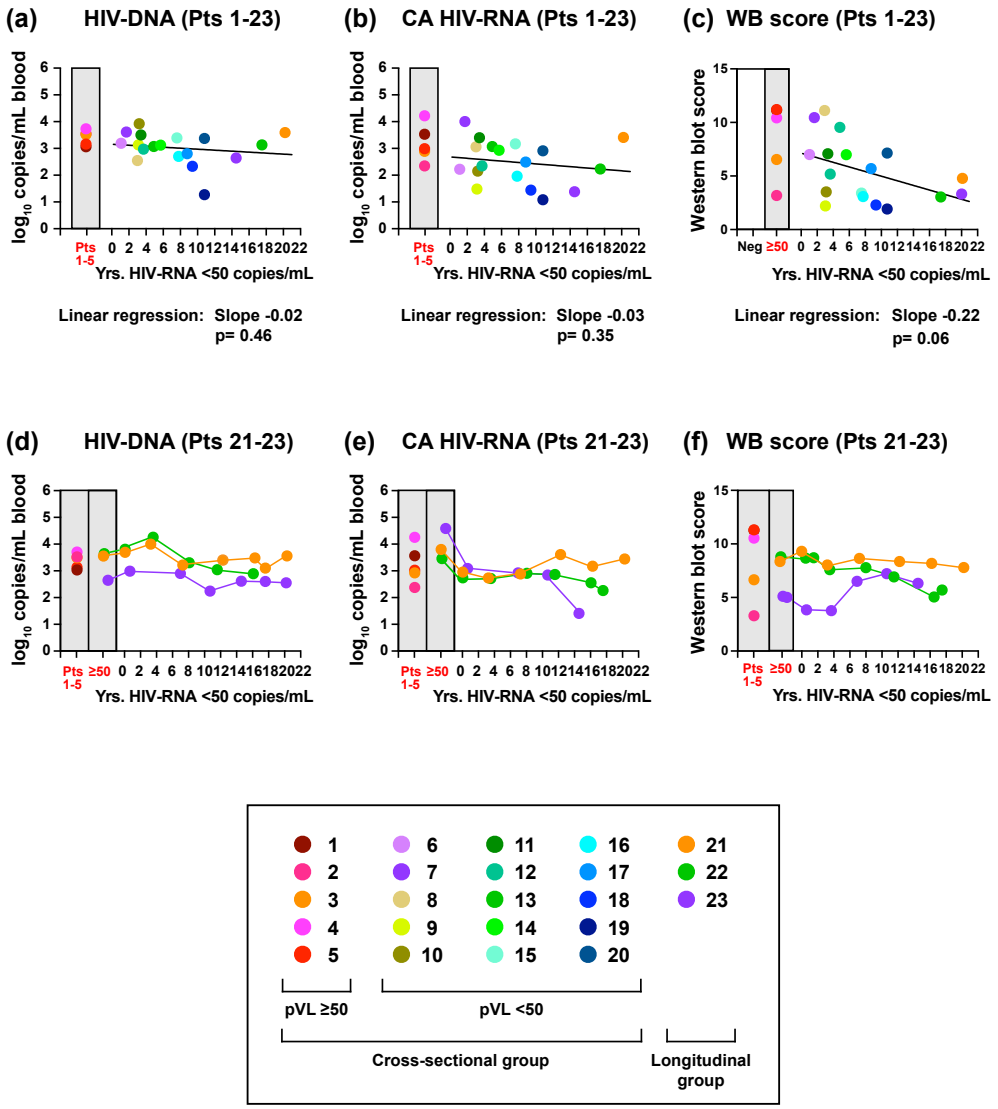
[c]Clinical Research Section, Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, MD 20892, USA.

[d]Biostatistics Research Branch, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, MD 20852, USA.

Correspondence: Hiromi Imamichi, PhD, Clinical and Molecular Retrovirology Section, Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, NIH, 10 Center Drive, Bethesda, MD 20892 (himamichi@nih.gov)

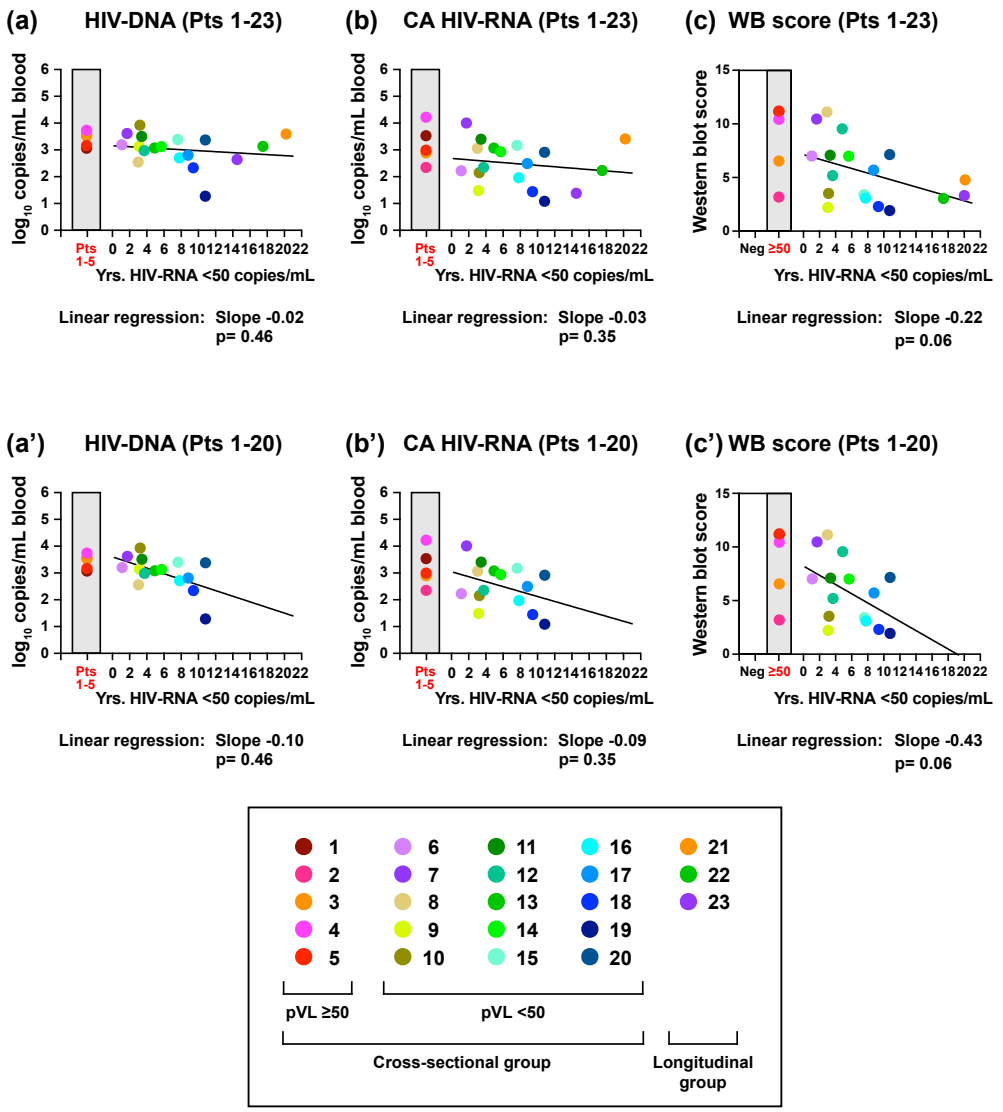


Supplementary Figure S1. Study design. Schematic diagram illustrating the study design and sampling schema is shown. **(a)** At the time of participant assessment, participants 1-5 (Pts 1-5) had detectable viremia with plasma HIV-RNA ≥ 50 copies/mL, whereas participants 6-20 (Pts 6-20) were virologically suppressed on antiretroviral therapy (ART) with plasma HIV-RNA <50 copies/mL. Participants 21-23 (Pts 21-23) were followed longitudinally up to 20 years. Samples were obtained from 5-7 different time points: 1 time point when their plasma HIV-RNA levels were ≥50 copies/mL; and 4-6 time points during suppressive ART with plasma HIV-RNA <50 copies/mL. The furthest time points from Pts 21-23 were included with the cross-sectional group for some analyses. **(b)** Associations between levels of HIV-DNA, cell-associated (CA) HIV-RNA, and western blot (WB) score and biomarkers of immune activation and inflammation were analyzed.



Supplementary Figure S2. Changes in the levels of HIV-DNA, CA HIV-RNA, and WB score as a function of duration of viral suppression.

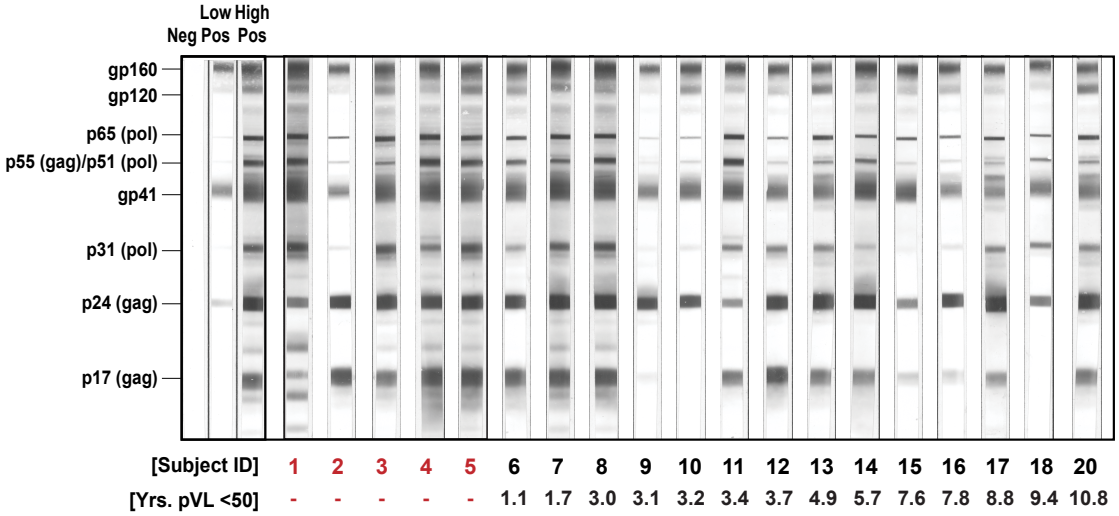
(**a-c**) HIV-DNA levels, CA HIV-RNA levels, and WB score were analyzed as a cross-sectional panel (Pts 1-23). (**d-f**) The same analyses were performed for three participants followed longitudinally (Pts 21-23). The slopes and p-values for linear regression during periods of suppression of plasma HIV-RNA <50 copies/mL are shown for Supplementary Figure S2 a-c. F-test was used to assess whether the slopes are significantly non-zero. The grey shaded areas represent time points with plasma HIV-RNA levels ≥50 copies/mL. Values from the five participants with plasma HIV-RNA ≥50 copies/mL from the cross-sectional panel (Pts 1-5) are plotted together as references in Figure S2 d-f. A color-coding scheme to indicate each participant is provided in the key.



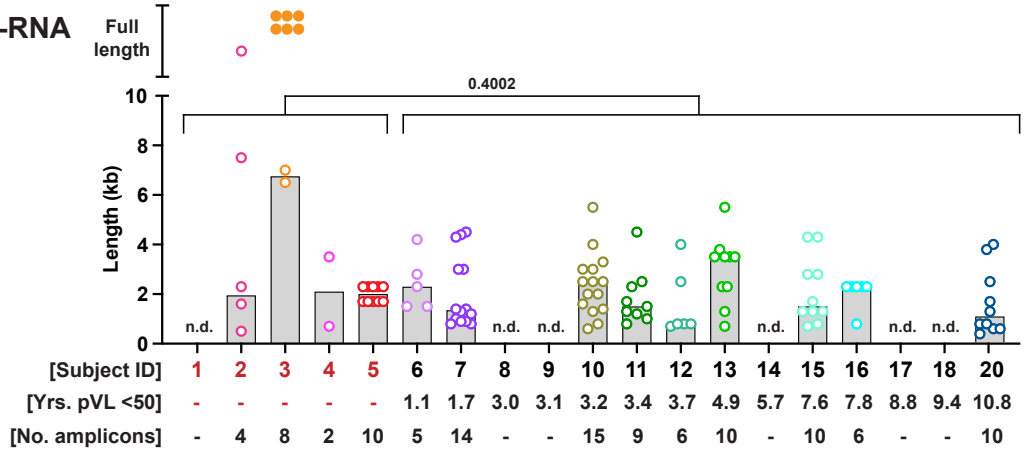
Supplementary Figure S2b. Changes in the levels of HIV-DNA, CA HIV-RNA, and WB score as a function of duration of viral suppression excluding three participants from the longitudinal cohorts (Pts 21-23).

(**a-c**) HIV-DNA levels, CA HIV-RNA levels, and WB score were analyzed as a cross-sectional panel (Pts 1-20). The slopes and p-values for linear regression during periods of suppression of plasma HIV-RNA <50 copies/mL are shown. The grey shaded areas represent time points with plasma HIV-RNA levels ≥50 copies/mL. A color-coding scheme to indicate each participant is provided in the key.

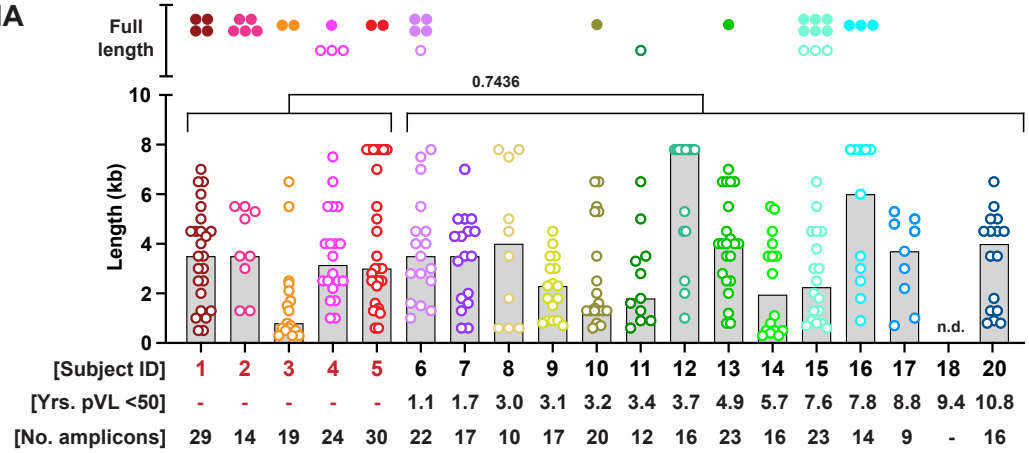
(a) Western blot



(b) CA HIV-RNA

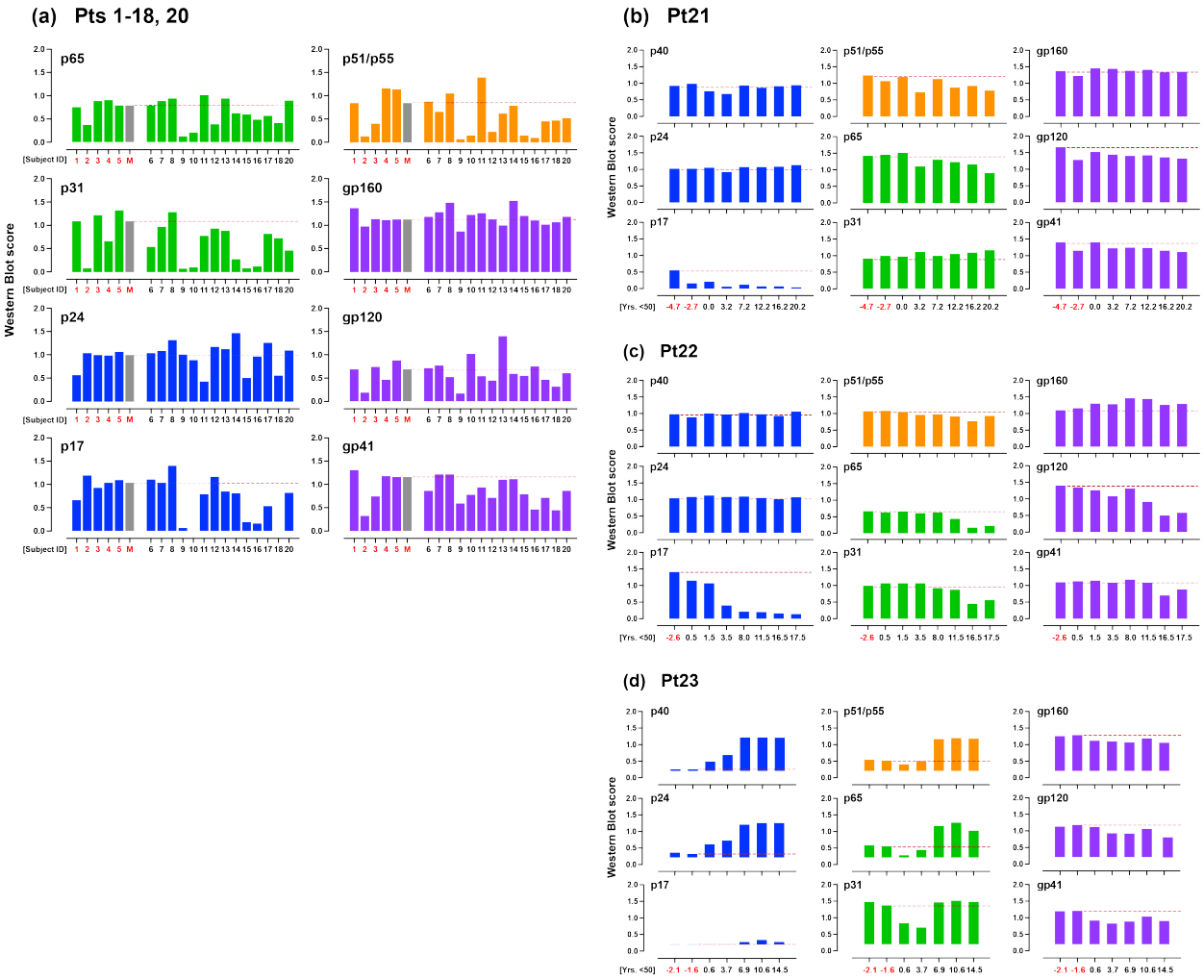


(c) HIV-DNA



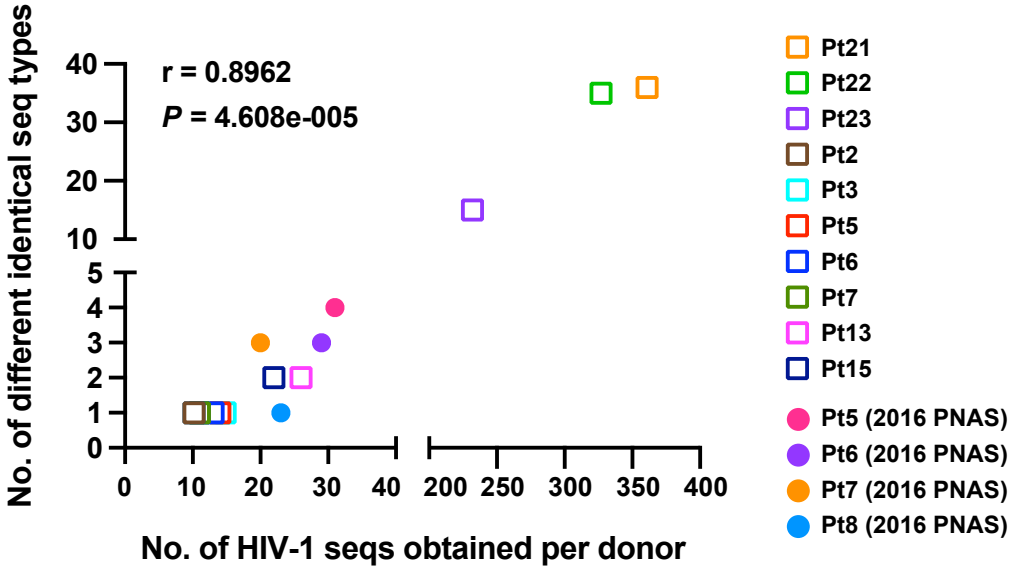
Supplementary Figure S3. HIV-1 western blot analysis and sequence length characterization of HIV-DNA and cell-associated (CA) HIV-RNA before and after suppression of plasma HIV-RNA <50 copies/mL on ART.

(a) Western blot patterns for 19 participants in the cross-sectional group (Pts 1-18 and 20) are shown. The results for Pt 19 were not included, as the western blot assay failed. Negative, low positive and high positive controls were run in parallel and shown on the left side of the panel. Participants sampled at timepoints when plasma HIV-RNA levels ≥50 copies/mL are indicated in red. Sequence length distributions of **(b)** HIV-DNA and **(c)** CA HIV-RNA are shown. Filled circles represent full-length intact species; and open circles represent non-full-length HIV-1 species or full-length containing out-of-frame indels, premature stop codons, or hypermutations. The height of the shaded box represents median length (kb) calculated for non-full-length HIV-1 species. Years of HIV-RNA levels <50 copies/mL and numbers of amplicons obtained for HIV-DNA or CA HIV-RNA analysis for each time point are indicated below the plots. Statistical significance was assessed by Mann-Whitney test for non-full-length HIV-1 species comparing study participants sampled prior to vs. during virologic suppression on ART with plasma HIV-RNA levels <50 copies/mL.



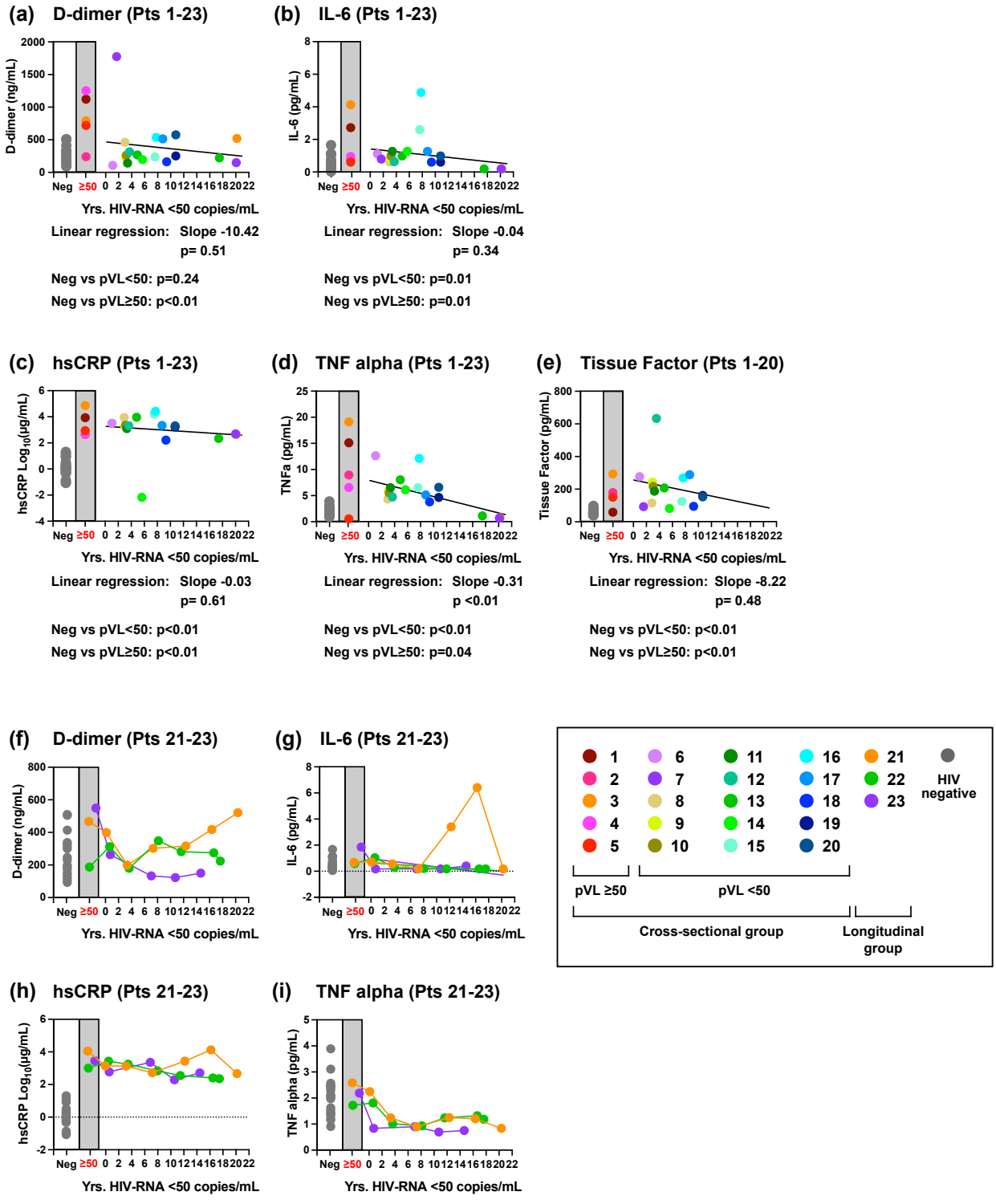
Supplementary Figure S4. HIV-1 Western blot analysis.

(a) Western blot (WB) scores for 19 participants in the cross-sectional group (Pts 1-18 and 20) are plotted. (**b-d**) Changes in western blot scores over time within each of the three persons were calculated and plotted for Pts 21-23 in the longitudinal group. Changes of the band intensity of each HIV-1 viral protein relative to that of the high positive control were calculated and plotted. The red dotted lines indicate the baseline values for each HIV-1 viral protein. The grey bars indicated as “M” represent the median values of the western blot scores for the five individuals with HIV-RNA ≥50 copies/mL (Pts 1-5) and are used as baseline values for Pts 1-18, 20. The results for Pt 19 were not included, as the western blot assay failed.



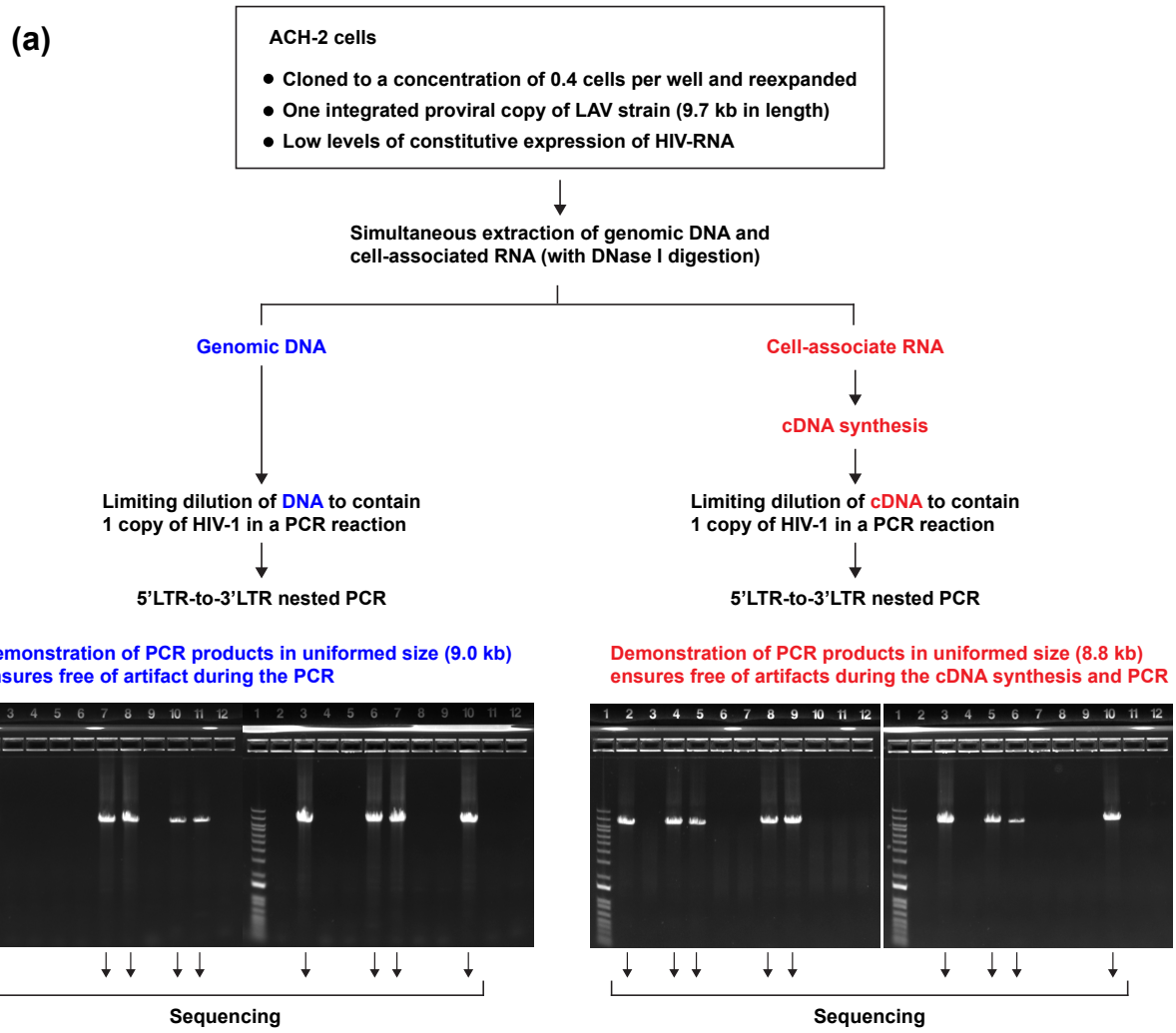
Supplementary Figure S5. Relationship between the number of different clonal HIV-1 sequences and the total number of HIV-1 sequences per individual.

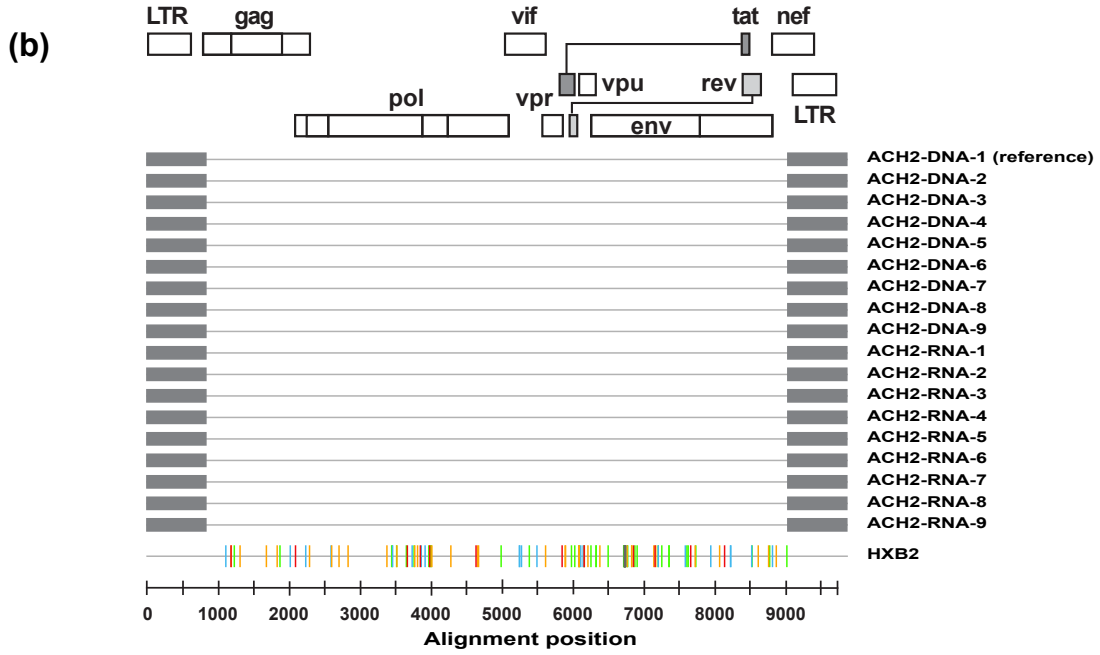
Data from a subset of the study participants in whom ≥10 HIV-1 sequences were obtained (Pts 2, 3, 5, 6, 7, 13, 15, 21-23) were used. Data from Pts 5-8 used in the 2016 PNAS publication [18] were also included in the analysis.


(a) D-dimer (Pts 1-23)
D-dimer (ng/mL)
Neg ≥50 0 2 4 6 8 10 12 14 16 18 20 22
Yrs. HIV-RNA <50 copies/mL
Linear regression: Slope -10.42
p= 0.51
Neg vs pVL<50: p=0.24
Neg vs pVL≥50: p<0.01
(b) IL-6 (Pts 1-23)
IL-6 (pg/mL)
Neg ≥50 0 2 4 6 8 10 12 14 16 18 20 22
Yrs. HIV-RNA <50 copies/mL
Linear regression: Slope -0.04
p= 0.34
Neg vs pVL<50: p=0.01
Neg vs pVL≥50: p=0.01
(c) hsCRP (Pts 1-23)
hsCRP Log10(µg/mL)
Neg ≥50 0 2 4 6 8 10 12 14 16 18 20 22
Yrs. HIV-RNA <50 copies/mL
Linear regression: Slope -0.03
p= 0.61
Neg vs pVL<50: p<0.01
Neg vs pVL≥50: p<0.01
(d) TNF alpha (Pts 1-23)
TNFa (pg/mL)
Neg ≥50 0 2 4 6 8 10 12 14 16 18 20 22
Yrs. HIV-RNA <50 copies/mL
Linear regression: Slope -0.31
p <0.01
Neg vs pVL<50: p<0.01
Neg vs pVL≥50: p=0.04
(e) Tissue Factor (Pts 1-20)
Tissue Factor (pg/mL)
Neg ≥50 0 2 4 6 8 10 12 14 16 18 20 22
Yrs. HIV-RNA <50 copies/mL
Linear regression: Slope -8.22
p= 0.48
Neg vs pVL<50: p<0.01
Neg vs pVL≥50: p<0.01
(f) D-dimer (Pts 21-23)
D-dimer (ng/mL)
Neg ≥50 0 2 4 6 8 10 12 14 16 18 20 22
Yrs. HIV-RNA <50 copies/mL
(g) IL-6 (Pts 21-23)
IL-6 (pg/mL)
Neg ≥50 0 2 4 6 8 10 12 14 16 18 20 22
Yrs. HIV-RNA <50 copies/mL
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
HIV negative
pVL ≥50
pVL <50
Cross-sectional group
Longitudinal group
(h) hsCRP (Pts 21-23)
hsCRP Log10(µg/mL)
Neg ≥50 0 2 4 6 8 10 12 14 16 18 20 22
Yrs. HIV-RNA <50 copies/mL
(i) TNF alpha (Pts 21-23)
TNF alpha (pg/mL)
Neg ≥50 0 2 4 6 8 10 12 14 16 18 20 22
Yrs. HIV-RNA <50 copies/mL

Supplementary Figure S6. Changes in the levels of biomarkers of immune activation and inflammation as a function of duration of viral suppression.

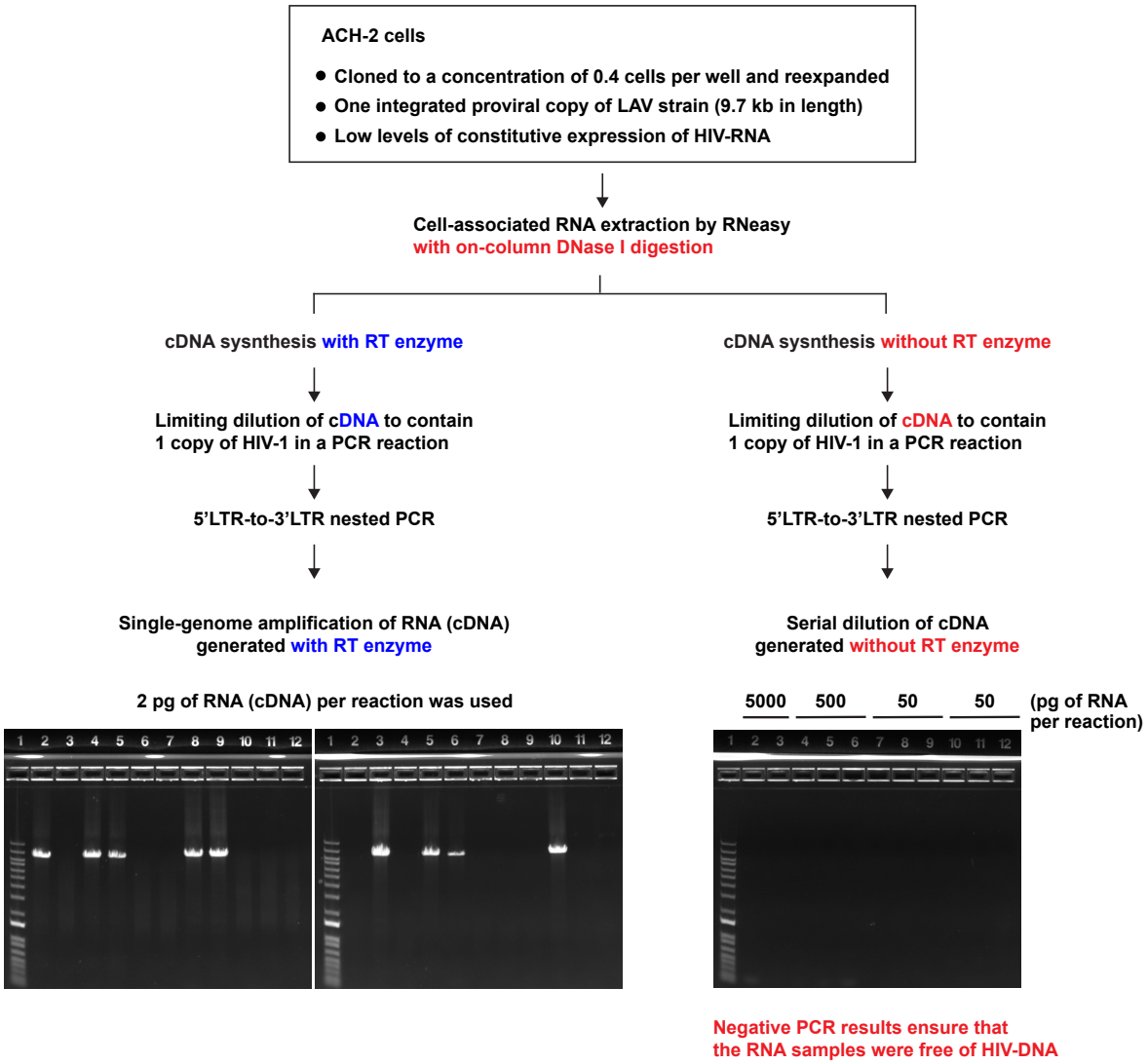
(**a-e**) D-dimer, Interleukin-6 (IL-6), high-sensitivity C-reaction protein (hsCRP), Tumor Necrosis Factor alpha (TNFα), and Tissue Factor were analyzed as a cross-sectional panel (Pts 1-23). Tissue Factor data were not available from Pts 21-23. (**f-i**) The same analyses were performed for the three participants in the longitudinal group (Pts 21-23). The slopes and p-values for linear regression during periods of suppression of plasma HIV-RNA <50 copies/mL are shown for Supplementary Figure S6 a-e. F-test was used to assess whether the slopes are significantly non-zero. The p-values are provided under each plot The grey dots represent values from HIV-negative controls. The grey shaded areas represent time points with HIV-RNA ≥50 copies/mL. A color-coding scheme to indicate each participant is provided in the key.


(a)
ACH-2 cells
Cloned to a concentration of 0.4 cells per well and reexpanded
One integrated proviral copy of LAV strain (9.7 kb in length)
Low levels of constitutive expression of HIV-RNA
Simultaneous extraction of genomic DNA and cell-associated RNA (with DNase I digestion)
Genomic DNA
Cell-associate RNA
cDNA synthesis
Limiting dilution of DNA to contain 1 copy of HIV-1 in a PCR reaction
Limiting dilution of cDNA to contain 1 copy of HIV-1 in a PCR reaction
5'LTR-to-3'LTR nested PCR
5'LTR-to-3'LTR nested PCR
Demonstration of PCR products in uniformed size (9.0 kb) ensures free of artifact during the PCR
Demonstration of PCR products in uniformed size (8.8 kb) ensures free of artifacts during the cDNA synthesis and PCR
Sequencing
Sequencing


(b)
LTR
gag
vif
tat
nef
pol
vpr
vpu
rev
env
LTR
ACH2-DNA-1 (reference)
ACH2-DNA-2
ACH2-DNA-3
ACH2-DNA-4
ACH2-DNA-5
ACH2-DNA-6
ACH2-DNA-7
ACH2-DNA-8
ACH2-DNA-9
ACH2-RNA-1
ACH2-RNA-2
ACH2-RNA-3
ACH2-RNA-4
ACH2-RNA-5
ACH2-RNA-6
ACH2-RNA-7
ACH2-RNA-8
ACH2-RNA-9
HXB2
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
Alignment position

Supplementary Figure S7. PCR amplicons generated during the 5'LTR-to-3'LTR PCR were free of artifacts.

(**a**) To rule out the possibility of the PCR amplicons obtained in the present study were not artifacts arisen during the PCR assay, genomic DNA and cell-associated RNA were extracted from ACH-2 cell line. The ACH-2 cell line was cloned to a concentration of 0.4 cells per well and re-expanded to ensure the clonality of the cells. Agarose gel pictures depicting sizes of HIV-DNA (9.0 kb) and CA HIV-RNA (8.8 kb) fragments generated during the 5'LTR-to-3'LTR PCR are shown. Positive PCR reactions generated at limiting dilutions are shown. (**b**) The Highlighter program analyses of the sequence data for the HIV-DNA and CA HIV-RNA derived from the ACH-2 cell line. ACH-2 HIV-RNA sequences corresponded precisely to the HIV-DNA sequences.



Supplementary Figure S8. Cell-associated HIV-RNA samples used in the present study were free of residual DNA contamination.

The ACH-2 cell line was cloned to a concentration of 0.4 cells per well and re-expanded to ensure the clonality of the cells. Cellular RNA was extracted by the RNeasy Mini kit (Qiagen) with on-column DNase I digestion per manufacturer's instruction. Conversion of RNA to cDNA was performed with and without reverse transcriptase (RT) enzyme. In PCR amplicons derived from cDNA product generated with RT enzyme, CA HIV-RNA of expected size (8.8 kb) were detected. PCR reactions derived from cDNA product generated without RT enzyme did not produce any amplicon. The use of excess amount of cDNA (from 25 to 2500 times excess amount required for single-genome amplification) for the PCR reaction did not change the outcome.

Supplementary Table 1. Characteristics of study participants

Subject ID	Time point	Age	Gender	Race, Ethnicity[1]	HIV-RNA (cps/mL)	$CD4^+$ T cell count (cells/µL)	$CD4^+$ T cell nadir (cells/µL)	Yrs. since HIV-1 diagnosis	Yrs. with HIV-RNA <50 cps/mL[2]		ART Regimen[3]
1	-	45	F	B, N	26100	295	295	3.5	n.a.		Not on ART
2	-	31	M	U, HL	21755	433	433	0.2	n.a.		Not on ART
3	-	46	M	W, N	3159	462	462	0.1	n.a.		Not on ART
4	-	37	F	B, N	2948	555	555	3.0	n.a.		Not on ART
5	-	34	F	B, N	2044	548	548	3.9	n.a.		Not on ART
6	-	26	M	U, HL	<50	520	287	3.6	1.1		ATV, RTV, FTC, TDF
7	-	57	M	B, N	<50	641	420	23.0	1.7		EFV, FTC, TDF
8	-	55	F	B, N	<50	652	305	10.3	3.0		EFV, FTC, TDF
9	-	44	M	W, N	<50	275	192	8.6	3.1		AZT, 3TC IDV, RTV, ddI
10	-	52	M	W, N	<50	597	105	23.4	3.2		DRV, RTV, FTC, TDF, RGV
11	-	49	M	W, N	<50	1036	299	14.2	3.4		ATV, RTV, 3TC, TDF
12	-	58	M	W, N	<50	436	171	3.2	3.7		DRV, RTV AZT, 3TC, RGV
13	-	57	M	W, N	<50	543	154	20.3	4.9		ATV, RTV, TDF
14	-	68	M	W, N	<50	577	356	15.5	5.7		ATV, RTV, ABC, 3TC
15	-	47	M	W, N	<50	1026	225	19.6	7.6		ABC, 3TC, TDF, EFV
16	-	56	M	W, N	<50	593	355	18.8	7.8		ATV, ABC, 3TC
17	-	59	F	U, HL	<50	806	724	15.5	8.8		AZT, 3TC, NVP
18	-	44	M	W, N	<50	343	209	18.7	9.4		FTC, TDF, NVP
19	-	49	M	W, N	<50	541	119	20.6	10.8		FTC, TDF, NVP
20	-	65	M	W, N	<50	916	446	23.3	10.8		fAPV, RTV, FTC, TDF
21	-	33	M	W, N	9151	724		4.6	-4.7	(*)	Not on ART
21	1	35	M	W, N	27000	583	252	6.6	-2.7		Not on ART
21	2	39	M	W, N	<500	448		9.3	0.0		RTV, d4T, 3TC
21	3	42	M	W, N	<50	796		12.5	3.2		RTV, d4T, 3TC
21	4	45	M	W, N	<50	553		16.5	7.2		NVP, ABC+3TC
21	5	50	M	W, N	<50	411		21.5	12.2		NVP, ABC+3TC
21	6	54	M	W, N	<40	499		25.5	16.2		NVP, ABC+3TC
21	7	58	M	W, N	<40	617		29.5	20.2		NVP, ABC+3TC
22	1	46	M	W, N	15000	704	423	9.2	-2.6		AZT
22	2	49	M	W, N	<500	554		12.2	0.5		AZT, 3TC, IND
22	-	49	M	W, N	<500	662		13.2	1.5	(*)	NFV, AZT, 3TC
22	3	52	M	W, N	<50	675		15.3	3.5		NFV, AZT, 3TC
22	4	56	M	W, N	<50	708		19.7	8.0		EFV+FTC+TDF
22	5	60	M	W, N	<50	495		23.3	11.5		EFV+FTC+TDF
22	6	64	M	W, N	<40	521		28.2	16.5		EFV+FTC+TDF
22	7	66	M	W, N	<40	551		29.2	17.5		EFV+FTC+TDF
23	-	38	M	W, HL	89980	27		10.7	-2.1	(*)	Not on ART
23	1	38	M	W, HL	88460	19	6	11.2	-1.6		Not on ART
23	2	41	M	W, HL	<50	197		13.4	0.6		d4T, 3TC, NVP
23	-	44	M	W, HL	<50	879		16.5	3.7	(*)	d4T, 3TC, RTV
23	3	47	M	W, HL	<50	555		19.7	6.9		d4T, 3TC, RTV
23	4	51	M	W, HL	<50	473		23.4	10.6		FTC+TDF, RTV, ATV
23	5	55	M	W, HL	<50	499		27.3	14.5		FTC+TDF, RTV, ATV

n.a.: Not applicable

[1]U: Unknown, W: White, B: Black, HL: Hispanic or Latino; N: Not Hispanic or Latino

[2]Years of stable suppression with plasma HIV-RNA below limit of detection of clinical assays

[3]Antiretroviral drug abbreviations: ABC (abacavir); ATV (atazanavir); AZT (zidovudine); ddI (didanosine); DRV (darunavir); EFV (efavirenz); fAPV (fosamprenavir); FTC (emtricitabine); IDV (indinavir); NFV (nelfinavir); NVP (nevirapine); TDF (tenofovir disoproxil fumarate); RGV (raltegravir); RTV (ritonavir); 3TC (lamivudine)

*: Additional time points examined for Western blot score

Supplementary Table 2

Associations among levels of HIV-1, Western blot (WB) score, biomarkers, and T cells.

	Western blot score				CA HIV-RNA				HIV-DNA			
Data from pVL≥50 copies/mL	Included		Excluded		Included		Excluded		Included		Excluded	
	r	*P* value	*r*	*P* value	*r*	*P* value	*r*	*P* value	*r*	*P* value	*r*	*P* value
HIV-DNA	0.16	0.49	0.20	0.45	0.50	0.01	0.59	0.01	-	-	-	-
HIV-RNA	0.73	**<0.01**	0.69	**<0.01**	-	-	-	-	-	-	-	-
WB score	-	-	-	-	-	-	-	-	-	-	-	-
D-dimer	0.52	0.01	0.31	0.23	0.53	0.01	0.36	0.14	0.26	0.23	0.13	0.61
IL-6	0.17	0.46	0.17	0.99	0.23	0.29	0.22	0.39	0.15	0.50	0.21	0.41
hsCRP	0.23	0.32	0.28	0.29	0.09	0.68	0.14	0.59	-0.05	0.81	0.04	0.88
TNFa	0.19	0.40	0.35	0.18	0.23	0.30	0.21	0.42	0.30	0.18	0.30	0.24
Tissue Factor	-0.41	0.08	-0.30	0.30	-0.45	0.05	-0.35	0.20	0.04	0.87	-0.07	0.81
$CD4^+$ T cells	0.27	0.23	0.51	0.04	0.44	0.04	0.69	**<0.01**	0.28	0.20	0.46	0.05
$CD8^+$ T cells	0.40	0.06	0.50	0.04	0.52	0.01	0.59	0.01	0.65	**<0.01**	0.71	**<0.01**

The Spearman correlation coefficient (r) and p-value are shown.

P-values <0.01 are in bold.

Supplementary Table 3

Associations among levels of HIV-1, Western blot (WB) score, biomarkers, and T cells excluding three participants from the longitudinal cohort (Pts 21-23).

	Western blot score				CA HIV-RNA				HIV-DNA			
Data from pVL≥50 copies/mL	Included		Excluded		Included		Excluded		Included		Excluded	
	r	*P* value	*r*	*P* value	*r*	*P* value	*r*	*P* value	*r*	*P* value	*r*	*P* value
HIV-DNA	0.11	0.66	0.17	0.55	0.42	0.07	0.53	0.04	-	-	-	-
HIV-RNA	0.77	**<0.01**	0.74	**<0.01**	-	-	-	-	-	-	-	-
WB score	-	-	-	-	-	-	-	-	-	-	-	-
D-dimer	0.53	0.02	0.31	0.27	0.46	0.04	0.21	0.46	0.17	0.47	-0.04	0.90
IL-6	-0.10	0.69	-0.15	0.60	0.24	0.31	0.33	0.23	0.20	0.40	0.37	0.17
hsCRP	0.04	0.88	0.10	0.74	0.09	0.70	0.22	0.45	-0.06	0.82	0.07	0.82
TNFa	0.05	0.84	0.19	0.53	0.25	0.31	0.32	0.26	0.41	0.08	0.52	0.06
Tissue Factor	-0.41	0.08	-0.30	0.30	-0.45	0.05	-0.35	0.20	0.04	0.87	-0.07	0.81
$CD4^+$ T cells	0.22	0.36	0.43	0.12	0.41	0.07	0.69	**<0.01**	0.22	0.36	0.43	0.11
$CD8^+$ T cells	0.28	0.25	0.42	0.14	0.46	0.04	0.58	0.03	0.67	**<0.01**	0.76	**<0.01**

The Spearman correlation coefficient (r) and p-value are shown.

P-values <0.01 are in bold.