Supporting Information

Josefsson et al. 10.1073/pnas.1107729108

SI Materials and Methods

Detailed Description of the Single-Cell Sequencing Assay (SCS). ***Isolation of peripheral blood mononuclear cells.*** Blood samples from the nine patients were collected (one time point each for patients 1–5 and 9, and two different time points 6 mo apart for patients 6–8), and peripheral blood mononuclear cells were separated from the plasma using Ficoll.

Monoclonal antibodies. The monoclonal antibodies used for phenotypic isolation of $CD4^+$ T cells were CD3-FITC/CD8-PE/CD45-PerCP/CD4-APC in an antibody mix (Becton Dickinson). Monocytes were isolated using CD14 conjugated to Pacific Blue and CD16 conjugated to PE-Cy7 (Becton Dickinson).

Fluorescence-activated cell sorting. FACS was carried out on a FACSAria (Becton Dickinson). To rule out the possibility of cross-contamination from earlier patient cell sorts, before sorting the cells from each patient, beads were sorted into a PCR plate as a negative control plate. After the negative plate, patient $CD4^+$ T cells were sorted by gating on $CD4^+/CD3^+/CD45^+/CD8^-$ and monocytes based on $CD14^{hi}/CD3^{lo}$ into two separate tubes, one for $CD4^+$ T cells and one for monocytes. During a second sort, the two subsets of monocytes (CD16 and CD14) and $CD4^+$ T cells were placed directly into 96 wells on six separate PCR plates at concentrations of 30, 100, 300, 500, 1,000 and, if enough cells were available, 10,000 cells per well.

Cell lysis. To each well on the PCR plate, 50 μL of lysis buffer containing 10 mM Tris·HCL (Invitrogen), 0.5% Nonidet P-40, 0.5% Tween-20 (Invitrogen), and proteinase K (Ambion) at a concentration of 0.3 mg/mL was added (1). Cells were lysed and DNA released by incubating plates at 55 °C for 1 h and subsequently at 85 °C for 15 min. After cell lysis all plates were stored at −20 °C.

DNA amplification and detection. New PCR plates were prepared for the HIV-1 DNA amplification step. These plates were set up with 5 μL in each of 86 wells containing the following components at the indicated final amounts or concentrations in sterile molecular-grade water: 2XPCR Buffer (Invitrogen), 4 mM $MgSO_4$, 400 nM of specific outer primers, 0.08 U/μL Taq platinum polymerase (Invitrogen), and 400 μM nucleotide mix (Promega). The frozen cell lysate was then thawed for each cell concentration (30–10,000). Gag-pol (p6-RT) fragments from the released intracellular HIV-1 DNA molecules were amplified, and the infection frequency was determined for each cell concentration by the following procedure. Five microliters from the first column of eight wells on the lysate plate containing 50 μL/well were transferred to the first column of eight wells on a second PCR plate prepared for amplification of HIV-1 DNA. This step was repeated to transfer the remaining lysate across nine columns on the PCR plate for a total of 80 wells containing 5 μL of cell lysate and 5 μL of PCR mix each. Six control wells were included on each PCR plate. A plasmid containing DNA from of a laboratory strain (pNL4-3) of HIV-1 was added to two wells as a positive control (1 μL of 10 copies/μL), and molecular-grade water was added to four wells as a negative control. As an additional negative control, the beads sorted before each patient sort were analyzed for HIV-1 DNA by adding 5 μL of the bead mixture to 80 wells on a PCR plate containing PCR reagents. The HIV-1 DNA in the wells from all of the plates was then amplified. The first round of PCR amplified a 1.6-kb region of the viral DNA using the outer primers 1894^+ GATGACAGCATGTCAGGGAG and 3500- CTATTAAGTCTTTTGATGGGTCATAA (2). Forty-two cycles of PCR (94 °C for 30 s, 50 °C for 30 s, 72 °C for 1 min 30 s) were performed after activating the hot start polymerase for 2 min at 94 °C. The first-round PCR products were diluted 1:10 by adding 90 μL 5 mM Tris·HCL, and 5 μL from the first-round PCR were added to a new PCR plate containing 15 μL master mix with second-round primers encompassing a 1.3-kb region of Gag and Pol 1870^+ GAGTTTTGGCTGAGGCAATGAG and 3410- CAGTTAGTGGTATTACTTCTGTTAGTGCTT. The second-round PCR was performed for 41 cycles (94 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min), followed by a 3-min extension at 72 °C. Positive HIV-1 DNA amplicons were identified on a 96-well 1% agarose gel with ethidium bromide (Invitrogen). None of the negative control plates or wells was positive. From the number of HIV-1 DNA molecules amplified by PCR the frequency of infection for each sorted cell concentration was calculated for each patient to determine the concentration level of cells containing far less than one HIV-1–infected cell. On the basis of these calculations, up to 11 additional HIV-1 DNA amplification PCR plates were set up for the particular sorted cell concentrations determined to contain much less than one HIV-1–infected cell. The amplification products identified on the second-round PCR plates were sequenced by direct dideoxyterminator sequencing (Applied Biosystems) in both directions using overlapping internal primers (2).

Validation of the SCS. To validate the SCS we used tissue culture 293T cells sequentially infected with NL4-3 and 89.6 based vectors at a low multiplicity of infection. Infected cells were enriched by cell sorting until >95% of the cells contained both viruses (3). To mimic in vivo conditions these infected cells were diluted in a background of uninfected 293T cells to 0.3 infected cells per 100 uninfected cells. One hundred cells from this mixture were lysed in 48 wells on a PCR plate. Six new PCR plates were then prepared for the HIV-1 DNA amplification step. These plates were set up with 5 μL in each of 86 wells containing the first-round PCR reagents, as described above. Five microliters from each of eight wells containing 50 μL/well were transferred to eight rows of 10 wells on the six PCR plates, for a total of 48 rows of 10 wells, each containing 5 μL of cell lysate and 5 μL of PCR mix each. The results of these control experiments validated the assay. For example, the expected total number of positive rows on six PCR plates containing a 0.3% infected to uninfected cell mixture would be 14; the actual number of positive rows produced by the control experiment containing amplified HIV-1 DNA was 12 (Table S1). Each infected cell contained two integrated HIV-1 proviruses; therefore, the expected total number of viral DNA molecules on six PCR plates was 28, and the actual number of wells with amplified HIV-1 DNA was 32. Moreover, the number of rows with more than one infected cell corresponded closely to that predicted by the Poisson distribution. These results indicate that (*i*) the HIV-1 DNA molecules from the lysed infected cells were equally distributed across the 10 wells of each row, and (*ii*) the SCS is sensitive enough to amplify and detect all HIV-1 DNA molecules from cells containing multiple HIV-1 DNA genomes.

Calculations and Statistical Methods. The rate of infected $CD4^+$ T cells was calculated using two methods. In method 1, all multiple DNA molecules were assumed to derive from singly infected cells (no occurrence of multiple infection). In method 2, all multiple DNA molecules were assumed to derive from multiply infected cells (maximal occurrence of multiple infection). For each method, the rate of infection was estimated by the number of infected cells divided by the total number of cells analyzed,

and exact binomial 95% confidence intervals were calculated for the rate of infection. Values are displayed as the inverse of the infection rate, as total number of cells per HIV-1 DNA molecule. For each method, the rate of infection was compared between patients with early infection vs. patients with chronic infection using the Mann–Whitney *U* test. For patients with multiple assessments, the geometric mean of the values was calculated before analysis.

For each patient, the null hypothesis that only single infection was present was tested by assessing the observed number of rows with two or more DNA molecules. The *P* value was calculated using the Poisson distribution with a mean value corresponding to the rate of infection calculated under the assumption that no multiple infection was present, as described above. This hypothesis was also tested for all patients with early infection combined and for all patients with chronic infection combined.

Because a row with two HIV-1 DNA molecules arising from a multiply infected cell cannot be differentiated from a row with two HIV-1 DNA molecules arising from two singly infected cells, an upper confidence interval was calculated using a conservative assumption for the rate of multiple infection. Specifically, each row with multiple HIV-1 DNA molecules was assumed to represent a cell with multiple infection. An exact one-sided binomial 95% confidence interval was calculated for each patient using the number of rows with two or more HIV-1 DNA molecules as the numerator and the number of rows with at least one provirus as the denominator. Thus, the upper bound of the confidence interval is a conservative estimate because some rows with multiple HIV-1 DNA molecules may arise from multiple singly infected cells. Confidence intervals were calculated in a similar manner for all patients with early infection combined and all patients with chronic infection combined.

For the eight patients for whom rows with multiple HIV-1 DNA molecules were detected, simulations were conducted to estimate the rate of multiple infection most consistent with the number of such rows. For a given rate of multiple infection, each simulation determined the expected number of rows with multiple HIV-1 DNA molecules, according to the Poisson distribution with a mean value corresponding to the estimated rate of infection for that patient. The rate of multiple infection that gave the expected number of rows with multiple HIV-1 DNA molecules that most closely matched the observed number for the patient was determined. Rates of multiple infection from 0 to 40% were tested, with 10,000 simulations for each rate.

Phylogenetic Analyses. Analyses of intracellular and extracellular HIV-1 populations were performed using an in-house computer program written in Perl scripting language (available upon request). Using this program we could identify all sequencing reactions with more than one amplified genome and discard these from the analyses. In total this quality control revealed that when performing SCS no more than 3.7% of the positive wells had more than one amplified genome, further demonstrating that the cellular lysate was equally distributed over the 10 wells on each PCR plate. For phylogenetic analysis of the HIV-1 DNA and RNA populations, maximum likelihood trees were constructed using the PhyML 3.0 program (4). The model of evolution was estimated for each patient data using the find model tool on the Los Alamos HIV Sequence Database (http://www.hiv.lanl.gov/content/index). Statistical support of the tree structures was obtained by 100 bootstrap replicates. All sequences obtained from the nine patients were compared by phylogenetic analysis to each other and standard laboratory viruses to ensure that no contamination between patient samples had occurred. Measurements of diversity (average pairwise distance) within cell and plasma populations were calculated using MEGA4.0 (http://www.megasoftware.net/). Diversities are reported here as percent differences.

SI Discussion

Although previous studies demonstrated that monocytes can be infected with HIV-1 in vitro, studies of in vivo infection of monocytes have yielded conflicting results (5–8). Two main subsets of monocytes have been described: the major $CD14^{hi}CD16^{-}$ subset, referred to as "classic monocytes," and the minor $CD14^{lo}CD16^{hi}$ "non-classic" subset, which resembles mature tissue macrophages (9). Both of these populations have been shown to be permissive to HIV-1 in vivo, with $CD14^{lo}CD16^{hi}$ being infected to a higher degree (0.01% for $CD14^{+}$ and 0.05% for $CD16^{+}$) in treatment-experienced patients (10). We analyzed monocytes from both subsets ($CD14^{+}$ in five of nine patients and $CD16^{hi}CD14^{lo}$ in four of nine patients) and did not find a single infected cell (Table S2). Given the low frequency of infection for $CD14^{+}$ monocytes (0.01%), our findings may result from analyzing too few cells (2,000–28,000). However, we analyzed up to 200,000 $CD16^{+}$ monocytes from one patient (patient 6, time point 1), which at an infection rate of 0.05% as shown by others (10) should have yielded ≈100 infected cells. The patients in our study were treatment-naïve. It is unclear, however, whether treatment correlates with monocyte infection, with some studies demonstrating infected monocytes in patients receiving highly active retroviral treatment and others not (10–12). Our results suggest that monocytes have a much lower frequency of infection than $CD4^{+}$ T cells in treatment-naïve patients (less than one HIV-1 DNA molecule/2,320–200,000 monocytes).

1. Brinchmann JE, Albert J, Vartdal F (1991) Few infected CD4+ T cells but a high proportion of replication-competent provirus copies in asymptomatic human immunodeficiency virus type 1 infection. *J Virol* 65:2019–2023.
2. Palmer S, et al. (2005) Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. *J Clin Microbiol* 43:406–413.
3. Rhodes TD, Nikolaitchik O, Chen J, Powell D, Hu WS (2005) Genetic recombination of human immunodeficiency virus type 1 in one round of viral replication: Effects of genetic distance, target cells, accessory genes, and lack of high negative interference in crossover events. *J Virol* 79:1666–1677.
4. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696–704.
5. Gartner S, et al. (1986) The role of mononuclear phagocytes in HTLV-III/LAV infection. *Science* 233:215–219.
6. Ho DD, Rota TR, Hirsch MS (1986) Infection of monocyte/macrophages by human T lymphotropic virus type III. *J Clin Invest* 77:1712–1715.
7. McElrath MJ, Pruett JE, Cohn ZA (1989) Mononuclear phagocytes of blood and bone marrow: Comparative roles as viral reservoirs in human immunodeficiency virus type 1 infections. *Proc Natl Acad Sci USA* 86:675–679.
8. Shen Y, et al. (1994) Blood monocytes from most human immunodeficiency virus type 1-infected patients do not carry proviral DNA. *Clin Diagn Lab Immunol* 1:531–537.
9. Ziegler-Heitbrock HW, et al. (1993) The novel subset of CD14+/CD16+ blood monocytes exhibits features of tissue macrophages. *Eur J Immunol* 23:2053–2058.
10. Ellery PJ, et al. (2007) The CD16+ monocyte subset is more permissive to infection and preferentially harbors HIV-1 in vivo. *J Immunol* 178:6581–6589.
11. Almodóvar S, et al. (2007) HIV-1 infection of monocytes is directly related to the success of HAART. *Virology* 369:35–46.
12. Fulcher JA, et al. (2004) Compartmentalization of human immunodeficiency virus type 1 between blood monocytes and CD4+ T cells during infection. *J Virol* 78: 7883–7893.

Table S1. Validation of the SCS

No. of uninfected cells/row	No. of rows	Expected results	Actual results (top row: number of rows with positive wells; bottom row: total) positive wells						
			Negative	One	Two	Three	Four	Five	Total
100	48	14 positive rows	36	3	3	2	3	1	12
100	48	30 positive wells	360	3	6	6	12	5	32

To validate the SCS, tissue culture 293T cells doubly infected with NL4-3 and 89.6 based vectors at an infection rate of 90% were used. To mimic in vivo conditions these infected cells were diluted in a background of uninfected 293T cells to a ratio of 0.3 infected cells per 100 uninfected cells. This cell mixture was placed in 48 wells at a concentration of 100 cells per well. Cells were lysed and transferred to rows of 10 wells on six PCR plates, for a total of 48 rows of 10 wells each, equating to 0.3 infected cell per row. The viral genome was amplified and detected. The actual number of positive rows and wells from the control experiment was compared with the expected number.

Table S2. Frequency of infection: Monocytes

Patient no.	Frequency of infection (monocytes)*	
	$CD14^{+}$	$CD16^{hi}CD14^{lo}$
Early infection		
1		
2		
3	<1/2,320	
4	<1/25,600	
5	<1/7,200	
Chronic Infection		
6.1		<1/200,000
6.2	<1/28,158	
7.1	<1/28,000	<1/73,210
7.2		<1/38,174
8.1		
8.2		<1/42,753
9.1		<1/5,304

*HIV DNA molecule/no. of cells.