A MAJORITY OF CLONALLY-EXPANDED T CELLS CONTAINING

REPLICATION-COMPETENT HIV-1 PROVIRUSES ARE

TRANSCRIPTIONALLY SILENT

by

Andrew Musick

B.A. Biological Chemistry (Hood College) 2015

THESIS

Submitted in partial satisfaction of the requirements

for the degree of

MASTER OF SCIENCE

in

BIOMEDICAL SCIENCE

in the

GRADUATE SCHOOL

of

HOOD COLLEGE

February 2019

Accepted:

Dr. Ann Boyd, Ph.D. Committee Member Dr. Ann Boyd, Ph.D. Director, Biomedical Science Program

Dr. John Coffin, Ph.D. Committee Member

Dr. Mary Kearney, Ph.D. Thesis Adviser Dr. April Boulton, Ph.D. Dean of the Graduate School

STATEMENT OF USE AND COPYRIGHT WAIVER

I do authorize Hood College to lend this thesis, or reproductions of it, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

DEDICATION

I dedicate my Master's Thesis work to my family. To my Mother and Father, Stephanie Musick and David Rampersaud, I am incredibly grateful to have such supportive and loving parents. Both of you have always guided me with words of encouragement, love, and knowledge. The importance placed on knowledge and learning helped guide my path in life. In addition, you both have always said to follow your passion and now I am in a field where my job does not feel like a job, but rather a passion. In this field (my passion) there is always something new to learn. Thank you both for everything you have done to help me.

I also dedicate my Master's Thesis work to my sister, Victoria Musick. Thank you for the support throughout the years. I am very proud of the work you are doing and the field you are going into. Remember there is always something new to learn every day. I know you will do well in your Master's program.

I dedicate this work to my wife, Katrina Musick. You have been supportive of my education and learning from the time we started dating. You have been by my side through the good and bad, but always inspirational and encouraging. I thank you for being a supportive wife and always having a positive outlook.

Thank you all again for all the support, love, and guidance you have given throughout the years. I love you all. Below is one of my favorite quotes:

"Live as if you were to die tomorrow. Learn as if you were to live forever"-Mahatma Gandhi

iii

ACKNOWLEDGEMENTS

I wish to thank my committee members who were more than generous to offer their expertise and time. I would like to thank Dr. Ann Boyd for all the help and advice. You are one of the most passionate professors I have had. There was never a dull moment in class. I have learned a great deal from being in your classes. In addition, a special thank you for all the edits and insights to help make my writing better.

I wish to thank Dr. John Coffin for all of the useful discussions, edits and insights. From presentations to papers, you have been there to help me improve and become a better scientist. I appreciate all the help that you have given me over the years.

I wish to thank the patient volunteer and his family. Without them, this study would not have been possible.

Finally, I would like to thank my mentor Dr. Mary Kearney for all of the help, advice, insight, and discussions. You have always encouraged me to go to meetings in order to gain more knowledge in the HIV field. Thank you for giving me those opportunities, I have learned a great deal. Thank you for always being there to help prepare and practice for presentations. The help and advice you have given has really helped me become a better speaker and scientist. In the lab, if there was ever a problem, I could always come to you for an insightful and helpful discussion. Thank you again for all your help and support!

Thank you, Dr. Ann Boyd, Dr. John Coffin, and Dr. Mary Kearney, for agreeing to be on my thesis committee and all the support!

iv

TABLE OF CONTENTS

	Page
ABSTRACT	vii
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF ABBREVIATIONS	Х
INTRODUCTION	1
MATERIALS AND METHODS	7
Donor and samples in study	7
Viral Load (VL) and Single Genome Sequencing of plasma and	7
Proviral DNA	
Ethics Statement	7
Cell-Associated RNA and DNA-Single Genome Sequencing	7
Sequencing Analysis	9
Sorting CD4+ T cell subsets	10
Statistics	10
RESULTS AND DISCUSSION	11
Participant Description, Phylogenetic Analysis of Plasma Viremia	11
and Viral Outgrowth Assays	
HIV Proviral Sequences Reveal Clones in PBMC	13
Probable Clones in CD4+ T Cell Subsets	16
Fraction and Levels of Cell-Associated HIV-1 RNA in Cell Clones	19
Sources of Drug Resistant Plasma Virus During ART	25

Comparison of High Expressing Cells Within Wildtype and	29
Drug Resistant Populations	

CONCLUSION	32
REFERENCES	35

ABSTRACT

Current antiretroviral therapy (ART) is highly effective at blocking HIV-1 replication but does not cure the infection due to the persistence of latently-infected cells that are able to undergo cellular proliferation (1). The majority of HIV-1 proviruses that persist during ART are defective. Of the minority that are intact and replication competent, it is not known what fraction are transcriptionally active in vivo versus those that are transcriptionally silent (latently infected). To address this question, I determined the fraction of HIV-1 proviruses in populations of expanded cell clones that express unspliced, cell-associated RNA during ART in one individual. In total, 34 different cell clones carrying either intact or defective proviruses in "Patient 1" from Maldarelli, et al. (1) were assessed. We found that a median of 2.3% of cells within clones harboring replicationcompetent proviruses contained unspliced HIV-1 RNA. Highest levels of HIV-1 RNA were found in the effector memory T cell subset, including for the replication-competent AMBI-1 clone, which was the source of persistent viremia on ART. The fraction of cells within a clone that contained HIV-1 RNA was not different in clones with replicationcompetent vs. defective proviruses. However, higher fractions and levels of RNA were found in cells with proviruses containing multiple drug resistance mutations, including those contributing to rebound viremia. These findings suggest that the vast majority of HIV-1 proviruses in persistently-infected cells, including replication-competent proviruses, are transcriptionally silent at any given time. This silence, if maintained over time, may allow infected cells to persist and expand during effective ART.

LIST OF TABLES

Table		Page
1	Fraction of Proviruses Expressing ca-HIV RNA and Levels of Expression in Single Cells	23
2	Gating for CD4+ T Naïve, CTM, and EM Cell Subsets from PBMC	24
3	Mean Fraction of Infected Cells Containing ca-HIV RNA and Levels of ca-HIV RNA in Single Cells: Wildtype vs. Drug Resistant Populations	28
4	Levels of Unspliced ca-HIV RNA in Cells with Proviruses Containing Major Drug Resistance Mutations	30

LIST OF FIGURES

Figure		Page
1	Visual representation of HIV structural proteins and proviral structure	2
2	HIV replication cycle overview: attachment, entry, reverse transcription, integration, assembly, release, and maturation.	3
3	Longitudinal HIV-1 viral load plot of Patient 1 in Maldarelli et. al. (1) and Simonetti et, al. (2) from the time of ART initiation to his death from cancer	11
4	Neighbor Joining distance tree of PBMC viral DNA sequences from the timepoint indicated in red in Figure 3	15
5	NJ distance tree of proviral sequences from CD4+ T cell memory subsets from the first timepoint (red) in Figure 3	17
6	Gating for CD4+ T Naïve, CTM, and EM cell subsets from PBMC	18
7	Plot of the fraction of cells that contained ca-HIV RNA for each of the 34 different clones/probable clones.	22
8	NJ distance tree of drug-resistant variants	27
9	Levels of ca-HIV RNA in single cells that carried the wildtype (blue) and drug resistant (red) proviruses.	31

LIST OF ABBREVIATIONS

ART	Antiretroviral Therapy
РВМС	Peripheral Blood Mononuclear Cells
СТМ	Central Transitional Memory
EM	Effector Memory
CARD-SGS	Cell Associated RNA and DNA Single-Genome Sequencing
CAR	Cell Associated RNA
cDNA	complementary DNA
HIV-1	Human Immunodeficiency Virus
PCR	Polymerase Chain Reaction

INTRODUCTION

Human Immunodeficiency Virus Type 1 (HIV-1) infects the body's immune system, more specifically CD4+ T cells. As HIV-1 infection destroys CD4+ T cells of the immune system, opportunistic infections and/or cancer ensue, leading to eventual death in most individuals without treatment. HIV-1 is a global epidemic with approximately 36.7 million people currently infected and the majority infections are in sub-Saharan Africa. In 2016 alone, approximately 1 million people died from HIV related causes.

HIV-1 can be suppressed in infected individuals with antiretroviral therapy (ART), however, ART does not cure the infection. Only two individuals, the first being Timothy Ray Brown, have been cured of HIV-1 thus far (3, 4). Brown was diagnosed with HIV-1 and later acute myeloid leukemia. He underwent whole body irradiation and two bone marrow transplants with cells from an HIV-resistant donor (coreceptor CCR5 deleted). The treatment cured the cancer, as well as the HIV-1 infection. A more recent study demonstrated that it was the loss of the CCR5 receptor that was responsible for curing the infection (5). Although ART is an effective therapy, it is life-long and associated with long term toxicities (6). Consequently, a cure that is more applicable to the general HIV-1 infected population than bone marrow transplantation is needed.

HIV is a retrovirus, utilizing the enzyme reverse transcriptase (RT) to make viral DNA from the packaged viral RNA. This viral DNA, called a provirus, is then integrated into the host genome. HIV-1 has 3 main genes *gag*, *pol*, and *env*, shown in Figure 1. *Gag* codes for structural proteins, *pol* codes for viral enzymes, and *env* codes for envelop associated proteins.



Figure 1. Visual representation of HIV-1 structural proteins and proviral structure. Animated by Janet Iwasa. <u>http://scienceofhiv.org/wp/?page_id=20</u>

The HIV replication cycle has 7 key steps: attachment, entry, reverse transcription, integration, assembly, release, and maturation (Figure 2). HIV envelope attaches to the CD4 receptor and coreceptor CCR5 or CXCR4 on the host cells. Next there is a membrane fusion mediated entry that releases the capsid inside the cell. The RNA genome is then reverse transcribed by the RT enzyme to make viral DNA. Viral DNA enters the nucleus and integrates into host genome via integrase. Transcription and translation are mediated by host machinery. There is assembly at the plasma membrane as a new virus particle is prepared. Egress is via budding and maturation occurs through protease cleavage. In the absence of ART, HIV accumulates high genetic diversity *in vivo* because RT is highly error prone, introducing new mutations with each round of viral replication.

ART blocks various steps in the life cycle to halt virus replication. Some targets include attachment, entry, reverse transcription, and integration. A combination of antiretroviral drugs, typically 3 or more, is given to patients to prevent the virus from becoming resistant to one compound. If the patient has unscheduled treatment

interruptions, the virus can replicate and mutate to become resistant to the drugs with longer half-lives. After resistance, a new regimen must be used to regain viral suppression.



Figure 2. HIV replication cycle overview: attachment, entry, reverse transcription, integration, assembly, release, and maturation.

ART prevents new cells from becoming infected but does not kill cells infected prior to its initiation. Integration is the hallmark of retroviruses and enables the persistence of infected cells. When the viral DNA is integrated, the provirus is treated the same as host genes. Transcription and translation are tied to the host and the provirus will persist as the cell divides and until the cell dies.

When a patient is sustained on ART, the plasma HIV RNA (measured in plasma copies per mL) is reduced from $\sim 10^2$ - 10^6 to 1. The result is low-level viral persistence

during ART. In patients sustained on treatment, the majority of proviruses, >95%, are defective due to hypermutations and large internal deletions (7). There are very few proviruses that are fully intact and replication competent. Latently-infected cells are those in an inactive state, not being transcribed to produce viral RNA or virus particles and, therefore, are able to escape immune recognition and viral cytopathogenesis. Latency is likely the mechanism that allows HIV-1 to persist despite ART. The reservoir for HIV-1 comprises infected cells that persist on long-term ART and carry intact proviruses. The fraction of cells within expanded, infected cell clones that are transcriptionally silent (latently infected) vs. those that are transcriptionally active is not known. The source of rapid rebound viremia is likely the expanded clones with intact proviruses that were already transcribing HIV-1 prior to the ART interruption (8). These cells are the reservoir for HIV-1 that must be targeted if a cure is to be achieved.

Proliferation of cells, most likely infected prior to ART initiation, that carry replication-competent proviruses (those that are fully intact and capable of producing infectious virus particles) is an important mechanism for maintaining the HIV-1 reservoir (1, 2, 9–13). It has been proposed that these cells are latently-infected and consequently transcriptionally silent, allowing them to evade virus-induced cytopathogenicity and immune-mediated clearance (2). By this model, persistent and rebound viremia originates from the occasional activation of a small fraction of the pool of latently-infected cells. It has also been demonstrated that latently-infected cells may be killed upon such activation (14, 15), although some studies have challenged this idea (12, 16).

Little is known about the fraction of HIV-1 infected cells persisting on ART that are latently-infected versus transcriptionally active. Here, I used the cell-associated RNA and DNA single-genome sequencing assay (CARD-SGS) (17) - a method that can be used to quantify the proportion of infected cells that have HIV-1 RNA and estimate the levels of expression in such cells. Preliminary studies suggested that during *in vivo* infection, greater than 80% of HIV-1 infected cells have proviruses that are transcriptionally-silent after long-term ART and that cells harboring transcriptionally-active proviruses contain only low levels of unspliced, cell-associated (ca) HIV-1 RNA (median 1 ca-HIV RNA/cell) (17). However, the fractions of transcriptionally-silent proviruses versus transcriptionallyactive proviruses remained unknown within populations of clonally-expanded infected cells, each of which contains the identical provirus at the identical host site, including those that carry intact proviruses (2). Furthermore, it is also not known which CD4+ T cell subsets expand and support the expression of HIV-1 proviruses that persist on ART.

To address these questions, I investigated peripheral blood mononuclear cells (PBMC) from a patient who presented with low level detectable viremia after prolonged ART. Previous analyses revealed that the on ART viremia in this individual originated from two sources: 1) viral replication of drug-resistant variants and 2) virus expression from a highly expanded T cell clone harboring a replication-competent, wildtype HIV-1 provirus denoted AMBI-1 (1, 2). Cells containing AMBI-1 comprised the largest infected cell clone in this individual (approximately 10⁷ cells) and were the sole source of wildtype persistent viremia during ART (2). I investigated samples from this patient to measure levels of HIV-1 production both from cells infected via possible ongoing replication (drug resistant virus) and from long-lived reservoirs (wild-type virus). The region of the virus I sequenced was from the gene coding for the P6 protein in *gag*, and the *pol* gene from Protease (PR) to RT, approximately a 1.2kb fragment.

I identified a total of 34 different wildtype infected cell clones and probable clones (proviruses that are identical in P6-PR-RT), and used CARD-SGS (17) to determine the fraction of PBMC within each clone, including the AMBI-1 clone, that had detectable amounts of unspliced cell-associated HIV-1 RNA (ca-HIV RNA). I also examined if the nature of the provirus (intact or defective) was associated with the fraction of infected PBMC that contained ca-HIV RNA and quantified the levels of ca-HIV RNA in single infected cells in each of the 34 different infected cell clones and in cells infected with drug resistant variants. I determined that a relatively small proportion of PBMC produce ca-HIV RNA during ART, and within a clone of identical cells, on average, less than 10% are producing ca-HIV RNA at any given time. Similar fractions and levels of expression were observed in clones harboring replication-competent proviruses or defective proviruses. However, higher fractions and levels of expression were observed in cells infected from probable ongoing viral replication of drug resistant variants.

After determining the fraction and levels of unspliced HIV-1 RNA in the 34 different infected cell clones and in the populations of PBMC infected with drug resistant variants, I sought to determine the specific CD4+ T cell subsets that harbored each of these clonal populations, especially those with intact proviruses, and to identify the cell subsets that supported their expression. I found probable clonal populations to be present in central/transitional memory (CTM) and effector memory (EM) T cell subsets; however, a higher fraction was found in effector memory cells, including in the clones harboring replication-competent proviruses. This result may be an indication that although all clones can be found across the various cell subsets, cells harboring replication competent proviruses can be more readily detected in the EM cell subset.

MATERIALS AND METHODS

Donor and Samples in Study

One hundred million PBMC were obtained from the timepoint indicated in red in Figure 3 from "Patient 1" in Maldarelli et. al. (1) and Simonetti et. al. (2). The donor was enrolled in NIH protocol 00-I-0110 conducted at the NIH Clinical Center in Bethesda, MD and was approved by the NIH internal review board. The sample was collected specifically for this study and the donor provided written informed consent.

Viral Load (VL) and Single Genome Sequencing of plasma and proviral DNA

The viral load (VL) or HIV RNA (copies per mL) in the blood was measured using a commercial viral load assay (Roche) at various timepoint throughout the >11 year ART treatment from "Patient 1" shown in Figure 3 (1, 2). The VL at the timepoint tested in this thesis project was 134 copies per mL, the patient was failing therapy with drug resistant variants present. Single-genome sequencing was completed on plasma RNA (18).

Cell-Associated RNA and DNA-Single Genome Sequencing (CARD-SGS)

CARD-SGS was performed exactly as described in Wiegand et. al. (17) on 72 aliquots containing one million PBMC each. The assay begins by aliquoting cells to about 100 proviruses per vial measured by the single copy assay (19, 20). Cell-associated RNA is extracted and cDNA is synthesized. Single cDNA molecules are PCR amplified and sequenced. Sequences are phylogenetically analyzed using neighbor joining (NJ) trees. The fraction of cells containing HIV RNA is then calculated as described in Wiegand et. al. (17). Cell-associated DNA measured by integrase (iCAD) is an assay used to measure the number of infected cells (19, 20). This assay measures integrase copy numbers with a real time PCR to yield an estimate of HIV infected cells (19) The number of infected cells in each aliquot was determined by iCAD (21) and the frequency of cells in each of the 34 different probable clones was determined by their frequency relative to the total proviral population by SGS.

Probable intact clones are rakes of identical proviral sequences, that when analyzed did not have any major insertions or deletions or stop codons for the region analyzed. Because it has been shown that >95% of proviruses *in vivo* are defective, most of these sequences likely have mutations elsewhere that render them defective. The probable clones also did not grow out in the viral outgrowth assay, indicating that they may not be replication competent. Probable clones were designated as such based on finding identical sequences in the region analyzed. For example, the AMBI-1 clone was found to comprise 11% of the total proviral population by SGS and the number of infected cells in each aliquot by iCAD was about 200 (measured by iCAD). Therefore, each aliquot contained ~20 AMBI-1 infected cells.

The fraction of cells within each clone that contained ca-HIV RNA was determined by assuming that each aliquot contained (on average) less than one "expressing" cell from each clone following the endpoint dilutions. Levels of ca-RNA in single-cells were determined by adding the number of identical ca-HIV-1 RNA molecules in each aliquot. Because SGS on HIV-1 RNA is known to include errors introduced in the RT step (18), variants that differed by a single nucleotide from a group of identical sequences of 5 or more that did not appear more than once were not considered to be a different variant, as previously described (17). Replication-competent proviruses were defined as P6-PR-RT sequences that had identical matches to variants that grew out in the viral outgrowth assay (VOA) (22, 23). Defective proviruses were those that contained stop codons or other obvious defects. Proviruses termed "non-induced" did not contain any obvious defects (i.e. stop codons) in the P6-PR-RT region analyzed but no outgrowth was observed in VOA at our level of sampling. The VOA is an assay used to determine or find intact replication competent provirus. The assay plates infected cells on irradiated feeder cells at varying dilutions, PHA stimulation to bind sugars of TCR triggering pathways to T cell activation. There is a 16-28 day cultivation, followed by P24 detection. This was completed as described in previous studies (2, 12, 24, 25). A total of about 10,450 HIV-1 infected cells were analyzed in the 72 aliquots of PBMC.

Sequencing Analysis

Sequences obtained by CARD-SGS were aligned using an alignment tool called ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/), which aligns sequences based on their pairwise genetic distances. An online software called MEGA6 (www.megasoftware.net) was used to construct neighbor joining phylogenetic trees (NJ). NJ trees are used to visualize sequences, observe the number of nucleotide changes and observe where sequences cluster. In addition, phylogenetic analysis and NJ trees provide a quick easy representation of identical sequences and allowing identifying clones through incorporation of a reference sequence. The diversity can be calculated as well. Inferred trees were used to match ca-HIV RNA sequences to each other and to the proviral DNA variants. Drug resistant variants were identified using the Stanford Database

(hivdb.stanford.edu). Sequences obtained by CARD-SGS are available in GenBank (www.ncbi.nlm.nih.gov/genbank) [MK148701-MK152485].

Sorting CD4+ T cell subsets

PBMC were obtained by leukapheresis. Following lysis of red blood cells, they were preserved in freezing medium at -80C/LN. 3.0×10^8 PBMC (93.1% viability) were thawed and stained with LIVE/DEAD Aqua stain (Molecular Probes, L34957, 5 minutes, 22°C), and with the following antibodies to sort for single CD4+ T cells: CD3-APC-H7 Biosciences, 641406), CD4-BV785 (BioLegend, 317442), CD8-QD0t655 (BD (Invitrogen, Q10055), CD11c-PE (BD Biosciences, 347637), CD14-PE (BD Biosciences, 555398), CD27-PE/Cy5 (Beckman Coulter, 6607107), CD45RO-ECD (Beckman Coulter, IM2712U), CD56-APC (BioLegend, 304610), CD57-BV421 (NIAID Vaccine Research Center (VRC) Ab), CCR7-Ax700 (VRC Ab), and TCR γδ (BD Biosciences, 555718) for 15 minutes at 22°C. Cells were washed, kept on ice, and sorted on a BD FACSAria into CD4+ memory subsets as follows: Naïve (CD3+Aqua-CD8-CD4hiCD56-TCRγδ-CD14-CD11c-CD27+CD45RO-CCR7+CD57-), Central/Transitional Memory (CTM, CD3+Aqua-CD8-CD4hiCD56-TCRγδ-CD14-CD11c-CD27+CD45RO+), and Effector Memory (EM, CD3+Aqua-CD8-CD4hiCD56-TCRγδ-CD14-CD11c-CD27-) (Figure 6).

Statistics

Standard statistical analyses were performed with the use of packages in R. The statistical tests used are indicated in the text and/or the tables.

RESULTS AND DISCUSSION

Participant Description, Phylogenetic Analysis of Plasma Viremia, and Viral Outgrowth Assays

The patient (1, 2) was an African-American male who was diagnosed with advanced HIV-1 infection (16 CD4+ T cells/µl) in May of 2000. He was treated with ART for 13 years; however, unplanned treatment interruptions did occur (Figure 3A) (2). While on ART, the HIV plasma RNA levels declined to below the limit of detection (less than 50 HIV-RNA copies/ml) and CD4+ T cell count partially recovered. However, low level rebound viremia occurred after about 11 years on ART (Figure 3) that was coincident with the development of a squamous cell carcinoma.



Figure 3. A) Longitudinal VL plot of Patient 1 from the time of ART initiation to his death from cancer. VLs below limit of detection (<50 copies per mL) are indicated with open black circles. The red arrow in the right panel indicates the timepoint that was analyzed by CARD-SGS. B) NJ tree showing plasma RNA sequences at three marked timepoints on the VL plot in A. Solid circles indicate sequences with no drug resistance mutations and open circles indicate drug-resistant variants. The ART regimens employed are shown in the boxes at the top.

To investigate HIV population genetics in this low-level viremia, a neighbor joining (NJ) phylogenetic distance tree was constructed using data obtained by P6-PR-RT SGS from plasma samples taken at three timepoints immediately before and after the ART regimen change (Figure 3B). The presence of drug-resistant (DR) virus prompted a change in ART regimen that re-suppressed drug-resistant viremia, while the identical wildtype (WT) sequences persisted (Figure 3). These data indicated that, as expected, the ART-resistant virus was the product of ongoing replication (because the addition of new drugs suppressed the DR virus and rebounding viremia), while the WT virus population was produced from a reservoir of clonally-expanded and/or long-lived cells (because it remained despite the change in ART regimen). The reservoir of long-lived cells is able to generate virus particles, but such particles do not result in ongoing cycles of viral replication because the reverse transcriptase inhibitors and integrase inhibitors block the productive infection of new cells.

Samples were collected for the investigation of HIV-1 production from clonallyinfected, long-lived cells (WT virus) as well as from ongoing replication (DR virus). The DR population was highly diverse indicating recent viral replication, whereas the WT population consisted almost entirely of identical sequences that were previously shown to arise from a highly-expanded clone of CD4+ T cells, called AMBI-1 (1, 2). After the ART regimen change, when viremia was re-suppressed (Figure 3 green timepoint), the DR virus was not detected, but the WT AMBI-1 virus persisted, suggesting that the AMBI-1 virus was not replicating, rather it was being released from clonally expanded cells.

An endpoint-dilution, single-stimulation, viral outgrowth assay (VOA) performed at the timepoint marked in red (Figure 3A) revealed four different replication-competent proviruses: WT AMBI-1, WT Outgrowth-1, WT Outgrowth-2, and DR Outgrowth-1 (2). Sequences matching two of these outgrowth viruses, AMBI-1 and DR Outgrowth, were also previously detected in plasma (Figure 3B). These data suggest that both WT and DR virus are capable of *ex vivo* outgrowth and, likely, *in vivo* rebound if ART is interrupted.

HIV Proviral Sequences Reveal Clones in PBMC

To characterize HIV populations in cells, CARD-SGS was performed on PBMC collected from the first timepoint (Figure 3A, red) shortly after breakthrough viremia on ART, at which time the plasma contained a mixture of about 50% DR and 50% WT virus (mostly AMBI-1) (Figure 3B). We obtained 183 independent P6-PR-RT proviral sequences from this timepoint and constructed a NJ distance tree (Figure 4). Because the total HIV-1 diversity in this individual was very high (2.9% average pairwise distance), as previously reported (1), the probability of obtaining two identical P6-PR-RT sequences by chance in our dataset was $<10^{-6}$ (calculated using an in-house program that considers both the average pairwise distance and the distribution of the sequences). We defined identical WT proviral sequences (or matching RNA sequences in different aliquots of cells) as "probable clones" of HIV-1 infected cells. Probable clones were divided into three categories: those carrying replication-competent proviruses determined by VOA (intact clones), those containing obviously defective proviruses (defective clones), and those having proviruses with no obvious defects in the region sequenced but were not recovered from the VOA (noninduced probable clones) (2).

We found ca-HIV RNA matching 34 of the clonal populations (Figure 4, numbered). For all clones, in all categories, a fraction of cells contained unspliced HIV-1 transcripts *in vivo*, including the WT replication-competent clones AMBI-1, Outgrowth-1,

and Outgrowth-2 (red triangles). The AMBI-1 clone was the largest detected by CARD-SGS. Five other clones with matching ca-HIV RNA carried defective proviruses with stop codons (purple squares), indicating that defective proviruses can express ca-HIV RNA in vivo, as previously reported (8, 17, 26, 27). Identical sequences were detected from 26 other distinct proviruses with matching ca-HIV RNA (blue triangles). These 26 probable clones did not have obvious defects (i.e. stop codons or large deletions) in the region sequenced but failed to produce detectable levels of infectious virus in a VOA despite the detection of ca-HIV RNA in vivo, suggesting defects elsewhere in the provirus or presence at a frequency below the sampling capacity of a single round of VOA. Seventy-seven proviruses for which no matching ca-HIV RNA was detected are shown with gray triangles for WT and squares for hypermutants. The fraction of RNA containing cells in these possible clones may have been below our level of detection or the proviral DNA may be integrated into a region that is less permissive to reactivation. As will be described in more detail below, we detected ca-HIV RNA matching each of the 3 intact proviruses, indicating that the reservoir for replication-competent HIV-1 contains cells with transcriptionallyactive proviruses *in vivo* as well as with silent proviruses.

As expected, DR proviruses were also detected by this analysis. DR sequences constituted only 16% of the proviral population (Figure 4) but represented nearly half of the rebounding plasma virus (Figure 3B), indicating that only a small fraction of the total pool of infected cells contributed to half of viremia. The eight DR variants with matching ca-HIV RNA are indicated with closed orange diamonds. Positive VOA was observed for one of these proviruses, labeled "DR Outgrowth-1" (Figure 4). DR proviruses for which no ca-HIV RNA was detected are indicated with open diamonds (known DR mutations) or

green diamonds (other mutations at DR sites). As in WT proviruses, we observed DR proviruses that were transcriptionally silent, transcriptionally active but not induced, and transcriptionally active and capable of *ex vivo* outgrowth and *in vivo* rebound.



Figure 4. NJ distance tree of PBMC viral DNA sequences from the timepoint indicated in red in Figure 1. The tree contains probable clones with matching ca-HIV RNA (indicated with numbers) including the replication-competent clone AMBI-1, and the probable clones Outgrowth-1 and Outgrowth-2 in red, defective clones in purple, non-induced probable clones in blue, and DR proviruses with matching ca-HIV RNA in orange. Proviruses where no ca-HIV RNA was detected are shown in gray. A total of 183 proviral sequences were analyzed.

Probable Clones in CD4+ T Cell Subsets

To determine the prevalence of the 34 probable clones in each subset, PBMC from the same timepoint were sorted by markers defining them as naïve, central/transitional memory (CTM), and effector memory (EM) CD4+ T cells (Figure 5). Twenty-seven of the 34 clones were detected in 90,000 to 150,000 cells assayed for each subset. The 27 clones identified were about equally represented in CTM (21/27) and EM (16/27) (p=0.3, Fisher exact), but were less prevalent in naïve cells (4/27) (pNaïve,CTM=0.0001, pNaïve,EM=0.004, Fisher's exact test). Although the fraction of CTM and EM cells harboring unique DR variants was not different (p=0.5, Fisher's exact test), a higher fraction of the infected naïve cells harbored DR variants than the CTM or EM subsets (naïve and CTM p=1.2*10-4; naïve and EM p = $7.8*10^{-5}$, Fisher's exact test). HIV uses two coreceptors, CXCR4 and CCR5, to enable entry into cells. The virus can use either of these two or both depending on the mutated HIV-1 envelope. Because naïve cells have higher surface expression of CXCR4 (18, 19), it is possible that the DR mutations arose on a CXCR4-using backbone, explaining their higher prevalence in naïve cells. Env sequencing followed by analysis for tropism using a software called geno2pheno (coreceptor.geno2pheno.org) confirmed that CXCR4-tropic envelopes were prevalent in the proviral population (sequences available in GenBank [MK148701- MK152485]).

Importantly, the replication-competent provirus, AMBI-1, was found in both CTM and EM T cells. Twelve of 100 AMBI-1 proviruses sequenced in the cell subsets were in CTM and the other 88 in EM (Figure 6). Geno2pheno results on AMBI-1 *env* suggested CCR5-usage, possibly explaining its lack of detection in the naïve cell population. Ten Outgrowth-1 proviruses were identified in the subsets and all were in EM cells. These results suggest that, although clones may be equally prevalent in CTM and EM cells, the replication-competent proviruses may be more readily detected in the EM, consistent with previous findings by Hiener, et al. (28).



Figure 5. Gating for CD4 T Naïve, CTM, and EM Cell Subsets from PBMC (A) not part of multi-cell conjugates (B) that were viable and stained with the T cell markers CD3 (C) and CD4 (D) but not myeloid cell markers CD14 and CD11c, or lineage markers CD56 and TCR $\gamma\delta$ (E) were divided by CD27 and CD45RO staining and collected as T_{CTM} (F, top-right gate) or T_{EM} (F, bottom gate) subsets. Cells that were single-positive (T₂₇₊; F, top-left gate) for CD27 were further divided by CD57 and CCR7 staining to collect T_N (G) subset. Numbers on plots represent percentages of plotted cells falling within the gates shown. Naïve CD4 defined as CD45RO-, CD27+, CCR7+, CD57-. CTM CD4 defined as CD45RO+, CD27+. EM CD4 defined as CD45RO+/-, CD27-.



Figure 6. NJ distance tree of proviral sequences from CD4+ T cell memory subsets from the first timepoint (red) in Figure 1. PBMC were sorted into naïve (pink), central-transitional memory (black), and effector memory (orange) CD4+ T-cell subsets. Probable clones from which ca-HIV RNA was recovered are numbered, including VOA outgrowth viruses. DR mutations are indicated.

Fraction and Levels of Cell-Associated HIV-1 RNA in Cell Clones

We hypothesized that a large fraction of cells within the clonal populations would have little or no viral RNA, possibly allowing them to survive and proliferate despite infection. To understand the persistence of the clonal populations of infected cells on ART and, in the case of AMBI-1, their contribution to low-level viremia during ART, we performed CARD-SGS and used the data to estimate the fraction of PBMC with detectable levels of viral RNA in the three different categories of clones/probable clones. We also used CARD–SGS to determine the levels of ca-HIV RNA in single PBMC and in single naïve, CD27+ memory (central or transitional memory [CTM]), and CD27- memory (effector memory [EM]) cells for each of the clonal populations.

For each clonal population, the fraction of infected PBMC with ca-HIV RNA and the levels of ca-HIV RNA were calculated as described in Wiegand, et al. (17) and are shown in Figure 7, Table 1, and Table 2. CARD-SGS was performed exactly as described in Wiegand et. al. (17) on 72 aliquots containing one million PBMC each. The assay begins by aliquoting cells to about 100-200 proviruses per vial measured by single copy assay. Viral RNA is extracted, and cDNA synthesized. Single cDNA molecules are PCR amplified and sequenced. Sequences were phylogenetically analyzed using neighbor joining (NJ) trees. The fraction with HIV RNA was then calculated using the original number of proviruses aliquoted. Intact clones had a median fraction of 2.3% (range 1-9%) of infected PBMC with detectable levels of ca-HIV RNA and a mean level of 2 (range 1-10) ca-HIV RNA molecules per RNA-containing cell (Figure 7, red bars). Defective clones had a median fraction of 3.5% (1-7%) with ca-HIV RNA and a mean level of 1 (range 1-4) ca-HIV RNA molecule per RNA-containing cell (Figure 7, purple bars). The clones that were not induced in VOA (non-induced probable clones) had a median fraction of 6.6% (1-65%) and a mean level of 2 (range 1-16) ca-HIV RNA molecules per RNA-containing cell (Figure 7, blue bars). Thus, no significant difference in the fractions or levels of P6-PR-RT ca-HIV RNA was observed between intact, defective, and the non-induced clones (p=0.20 to 0.99 Mann Whitney). This result is consistent with a study by Pollack, et al. (15) demonstrating that cells containing defective proviruses can express HIV transcripts, resulting in the recognition and clearance by CD8+ T cells ex vivo. Ca-HIV RNA matching 13 of the 34 cell clones was detected in 250,000 – 350,000 sorted CTM and EM cells (indicated in Figure 7). All ca-HIV RNA detected in naïve cells harbored drug resistance mutations and, hence, did not match any of the WT cell clones. In total, the fraction of naïve, CTM, and EM cells with ca-HIV RNA was not different from the total PBMC (4% of infected naïve cells with ca-HIV RNA, 2% of CTM, and 9% of EM). However, infected EM cells had a significantly higher fraction with ca-HIV RNA than infected CTM cells (7.8x10⁻¹⁰; Chi-sq). Consistent with the findings of others (29), higher levels of ca-HIV RNA were also observed in some single EM cells but the difference did not achieve significance with our level of sampling (p=0.1 Kruskal-Wallis).

Despite the AMBI-1 clone's being the source of WT persistent viremia in this individual (2), only a small fraction of cells (2.3%) within the AMBI-1 clone contained ca-HIV RNA, strongly suggesting that clonal populations persist because the vast majority of the daughter cells at each division produce little or no HIV RNA. These data demonstrate that all the cells within a clone are not uniformly producing HIV RNA and suggest that epigenetic factors may affect HIV transcription. These results also suggest that the production of infectious virus by a small fraction of the cells in the clonal populations that carry intact proviruses is likely to be the source of rapid rebound viremia when ART is interrupted. Likewise, only a small fraction of the clonal cell siblings appear to be the source of viral outgrowth in the VOA: about 3.6% of the AMBI-1 infected CD4+ T cells produced infectious virus in the quantitative VOA (2). Since it is improbable that the small numbers of unspliced viral RNA copies detected per cell are sufficient to support infectious virus production, it is likely that a much smaller fraction of cells with intact proviruses are actually producing virions at any one time *in vivo*. Taken together, these data support that only a small fraction of probable proviral clones are actively being transcribed, independent of intactness or outgrowth potential. Conversely, about 90% or more of the infected cells within clonally expanded populations contain little or no detectable unspliced viral RNA at a given point in time, allowing these cells to proliferate and survive for years, even though they may contain fully intact, infectious proviruses.



Figure 7. Plot of the fraction of cells that contained ca-HIV RNA for each of the 34 different clones/probable clones. Replication-competent clones/probable clones are shown in red. Defective clones are shown in purple. Non-induced probable clones are in blue. About 50 to 1200 cells within each clone/probable clone were assayed for ca-HIV RNA.

Population type	Median fraction of cells expressing HIV RNA (range)	Mean number of ca-HIV RNA copies in single cells (range)	Fraction of high-expressing cells (# of cells with >20 ca-HIV RNA molecules/# cells with any # of ca-HIV RNA)
Clones with intact proviruses (N=3) ^a	2.3% (1.2%-8.8%)	2.3 (1-10)	0/35
Clones with defective proviruses (N=5)	3.5% (0.9%-7.0%)	1.3 (1-4)	0/22
Non-induced clones (N=26) Cells with ca-HIV	6.6% (1.3%-64.9%)	2.0 (1-16)	0/277
RNA detected but no DNA detected b^{b}	>2.2%	1.5 (1-65)	9/784 [°]
Cells with DNA detected but no ca- HIV RNA detected	<1.7%	0	0/72

Table 1. Fraction of proviruses expressing ca-HIV RNA and levels of expression in single cells

^aExcluding DR-OG, which may reflect spreading infection ^bIncluding both wildtype and drug resistant populations ^cAll 9 high expressing cells were found in the drug resistant population

Clone Type	Clone ID	Number of infected cells with HIV RNA/total cells analyzed for each clone (%)	Average number of RNA molecules per expressing cell	Number of cells with the # of RNA molecules indicated (# of RNA molecules)
	1 (AMBI-1)	28/1199 (2.3%)	2.5	12(1), 5(2), 6(3), 2(4), 1(5), 1(6), 1(10)
Intact Clones (2 (WT Dutgrowth-1)	2/171 (1.2%)	1	2(1)
(3 (WT Dutgrowth-2)	5/57 (8.8%)	1.6	3(1), 1(2), 1(3)
	4	1/114 (0.9%)	1	1(1)
Defection	5	4/288 (1.8%)	1.3	3(1), 1(2)
Defective	6	2/57 (3.5%)	1	2(1)
Clones	7	11/171 (6.4%)	1.5	8(1), 1(2), 1(3), 1(4)
	8	4/57 (7.0%)	1.3	3(1), 1(2)
	9	3/228 (1.3%)	1	3(1)
	10	1/57 (1.8%)	1	1(1)
	11	1/57 (1.8%)	1	1(1)
	12	2/114 (1.8%)	6	1(1), 1(11)
	13	1/57 (1.8%)	1	1(1)
	14	1/57 (1.8%)	1	1(1)
	15	1/57 (1.8%)	1	1(1)
	16	10/343 (2.9%)	1	10(1)
	17	5/171 (2.9%)	1	5(1)
	18	2/57 (3.5%)	1	2(1)
	19	2/57 (3.5%)	1	2(1)
	20	2/57 (3.5%)	1	2(1)
	21	7/114 (6.1%)	1.2	5(1), 2(2)
Non-	22	4/57 (7.0%)	1	4(1)
Induced	23	17/228 (7.5%)	1.3	14(1), 1(2), 2(3)
Clones	24	23/286 (8.0%)	1.6	14(1), 5(2), 3(3), 1(4)
	25	13/114 (11.4%)	1.5	8(1), 3(2), 2(3)
	26	14/114 (12.3%)	1.6	10(1), 1(2), 2(3), 1(4)
	27	19/114 (16.7%)	1.1	18(1), 1(2)
	28	38/228 (16.7%)	2.8	15(1), 11(2), 4(3), 3(4), 1(5),
				1(6), 1(7), 1(9), 1(10)
	29	10/57 (17.5%)	1.3	8(1), 1(2), 1(3)
	30	11/57 (19.3%)	1.5	7(1), 3(2), 1(3)
	31	13/57 (22.8%)	1.2	12(1), 1(3)
	32	19/57 (33.3%)	2.2	12(1), 3(2), 2(3), 1(8), 1(9)
	33	20/57 (35.1%)	1.9	12(1), 4(2), 1(3), 2(4), 1(6)
	24		2.0	11(1), 3(2), 4(3), 7(4), 6(5),
	34	37/57 (64.9%)	3.9	1(6), 1(7), 1(8), 1(9), 1(11), 1(16)

Table 2. Fraction of cells in clones and probable clones that contain unspliced ca-HIV RNA

Sources of Drug Resistant Plasma Virus During ART

To compare the DR variants in the proviral population to those in the plasma, we constructed a NJ tree (Figure 4) of all DR proviral sequences (diamonds) and DR plasma variants (circles) and identified the mutations in each variant (defined by the Stanford University HIV Drug Resistance Database https://hivdb.stanford.edu/). The DR ca-HIV RNA sequences with matching HIV DNA are also shown in Figure 8 (squares) (discussed later). As in Figure 4, all the DR viral DNA sequences formed a distinct node relative to the WT sequences, implying a common evolutionary origin, which likely arose during one of the treatment interruptions (Figure 3A). Almost all the DR plasma sequences contained four linked mutations in RT (L74V, K103N, V108I, M184V) which, together, conferred resistance to each drug in the contemporaneous ART regimen. However, of 30 proviral DR variants, only three had all four of these linked DR mutations (Figure 8). This observation suggests that, perhaps, as many as 90% of the DR proviral sequences are not making a significant contribution to the plasma viral RNA pool. By contrast to the viral DNA, 13 of the 15 plasma DR viral RNA sequences belonged to the node with all four linked resistance mutations. Of the two sequences that did not have all four mutations, one had K103N alone and the other had K103N linked to V108I and M184V. Overall, the presence of the four linked DR mutations in only three out of a total of 183 proviruses implies that, at the time the sample was taken, less than 2% of the infected cells were the likely source of the replicating DR plasma virus.

The fraction of cells with detectable ca-HIV RNA in the WT versus DR populations is compared in Table 2. At the timepoint analyzed, an average of 4.7% of cells infected with WT proviruses had ca-HIV RNA while 17.2% of cells with DR proviruses had ca-

HIV RNA (p=3x10⁻¹⁴ Mann-Whitney test). Of the cells that carried proviruses that had the four DR mutations, 11.0% had ca-HIV RNA (detected by either DNA with matching ca-RNA as shown in Figure 8 or detected as only ca-HIV RNA in single cells (not shown in Figure 8 but included in sequences submitted to GenBank [MK148701- MK152485]). These results are consistent with our hypothesis that, despite ongoing viral replication of at least some of the DR variants, the majority of cells that carried the DR proviruses contained little or no viral RNA and, again, that only a small fraction of infected cells contribute to plasma viremia at a given time. Figure 8 shows the DR ca-HIV RNA variants for which we detected the matching proviruses among the 30 that were sequenced. Eleven percent of the DR proviruses with ca-HIV RNA had all four of the DR mutations discussed previously, suggesting that some of these proviruses may be in newly infected cells and may be contributing to the DR plasma virus.



Figure 8. NJ distance tree of drug-resistant variants. Plasma SGS are shown as red, open circles, PBMC proviruses with matching ca-HIV RNA are shown as orange diamonds, ca-HIV RNA in single cells are shown in different colored squares. The number of single cells that contained ca-HIV RNA is indicated by the number of different colored squares. The number of ca-HIV RNA molecules in each single cell is indicated by the number of squares of each color. Only RNA sequences identical to DR proviral DNA sequences are shown.

Population Type (# of cells assayed)	Mean fraction of expressing proviruses	р value	Mean number of ca-HIV RNA copies in single cells (range)	р value	High-expressing cells (# of cells with >20 ca-HIV RNA molecules/# cells with any # of ca- HIV RNA)	p value [°]
Cells with WT proviruses (N=820)	4.7%	14	1.6 (1-16)	4	0/820	5
Cells with DR proviruses (N=324)	17.2%	3x10 ⁻¹⁴	2.2 (1-65)	2x10 ⁻⁴	9/324	4x10 ⁻³
Cells with all 4 DR mutations (N=116) ^a	11.0%		2.6 (1-65)		5/116	

Table 3. Mean fraction of infected cells containing ca-HIV RNA and levels of ca-HIV RNA in single cells: Wildtype vs. Drug Resistant populations

^a Subset of all DR proviruses that have linked L74V, K103N, V108I, M184V

ı,

^b Mann-Whitney test ^c 2-tail Fisher Exact test

Comparison of High Expressing Cells Within Wildtype and Drug Resistant Populations

We also compared the distribution of ca-HIV RNA levels in single infected cells for the WT (N=820) and DR (N=324) variants (Table 3, Table 4, and Figure 9). The cells with WT variants had levels of ca-HIV RNA ranging from 1 to 16 copies/single PBMC with ca-HIV RNA (mean=1.6), whereas the DR population had levels ranging from 1 to 65 copies/single PBMC with ca-HIV RNA (mean=2.2) (p= $2x10^{-4}$) (Table 3). Cells with greater than 20 ca-HIV RNA copies, were observed more frequently in the DR variants (0 cells in WT, 9 cells in DR, Table 1) (Table 3, Figure 9). One example of a cell with high levels of ca-RNA is shown in the NJ tree in Figure 8 (pink squares, indicated with a black arrow). Five of the 9 cells with high levels of viral RNA contained proviruses with all four linked DR mutations. It is likely that some or all of these cells are producing virus and may, therefore, be short-lived and, hence, rarely detected. The remaining 4 high-expressing proviruses contained the following DR mutations: 1) K103N, V108I, M184V, H221Y, 2) K103N, V108I, M184V, V90I, 3) K103N, V108I, M184V, V90I, M46I, and 4) K103N, M184V, D67N. This hypothesis is consistent with a higher frequency of high expressing cells detected in the multi-DR population vs. WT.

Drug Resistance Variant ID	Average number of RNA molecules per cell	Number of cells (# of RNA molecules)
1 (Drug Resistant	3.2	25(1), 7(2), 2(3), 2(4), 1(12),
Outgrowth)	5.2	1(56) ^a
2	1.7	22(1), 6(2), 3(3), 4(4)
3	5.3	3(1), 1(24)
4	1	5(1)
5	1	4(1)
6	1	4(1)
7	1	2(1)
8	1	1(1)

Table 4. Levels of unspliced ca-HIV RNA in cells with proviruses containing major drug resistance mutations

^a **Bold** indicates a high-expressing cell (>20 ca-HIV RNA molecules per cell).



Figure 9. Levels of ca-HIV RNA in single cells that carried the wild-type (blue) and drug resistant (red) proviruses. The number of ca-HIV RNA copies per cell in the wild-type population ranged from 1 to 16. The number of ca-HIV RNA copies per cell in the drug resistant population ranged from 1 to 65. Data were analyzed from 820 cells with wild-type proviruses and 324 cells with drug resistant proviruses. The remaining infected cells did not contain any detectable ca-HIV RNA. The high-expressing cells are indicated with the black arrow.

CONCLUSIONS

ART can suppress HIV but cannot cure it in patients. The current therapy allows patients to live longer and halt virus replication. There are negative aspects of ART, including therapy is life long, access is limited (only about half of infected individuals have access), and is associated with long term toxicities. Development of a cure has great importance and will help millions of individuals. Curative strategies must focus and target the reservoir for HIV. The reservoir is composed of HIV infected expanded T cell clones harboring intact replication-competent proviruses. Most of these cells are latently infected they are not active and are not expressing HIV RNA (17). Because of this latency, cells can escape the immune system and persist despite ART. When the small fraction of cells within expanded clones do express HIV RNA, they are likely recognized and killed by the immune system. Latently infected cells may become activated to replace these cells, in addition to clonal expansion (cellular proliferation). The combination of this allows the fraction of expressing and non-expressing to remain constant. In addition, the latently infected cells can persist, as ART does not kill infected cells (only prevents new cells from becoming infected).

To develop effective HIV-1 cure strategies, it is critically important to define the mechanisms that allow HIV-1 infected cells to persist and proliferate in patients on ART, especially those cells that contain intact proviruses and can be the source of persistent low-level viremia during ART and cause viral rebound when ART is interrupted. Here, I present detailed evidence in one very well-characterized individual that most (>90%) of individual cells within infected cell clones carrying replication-competent proviruses are transcriptionally silent, independent of the T-cell subset, likely allowing them to escape

targeting by the immune system and direct cytopathic effects caused by the virus. We propose that similar mechanisms of HIV-1 persistence exist in other individuals on ART but additional in-depth analyses similar to those reported here will be required to generalize our findings. The life-span and clearance rate of the small fraction of cells with intact proviruses that express ca-HIV RNA is not defined but we hypothesize that they are eliminated and may be replaced by reactivation of siblings in the same clone. Testing of this hypothesis will require new approaches to assess single cell turnover in vivo of specific cell clones that constitute a very small fraction of all clonal CD4+ T cell populations.

The DR virus, which comprised about half of the virus present in the low-level viremia in the donor at the timepoint we sampled, consisted of cells with relatively high levels of ca-HIV RNA, defined here as having greater than 20 individual unspliced ca-HIV RNA molecules per cell. No such cells were found in the infected cells that carried WT proviruses, consistent with a previous study that suggested that cells with high levels of viral RNA may be more common in untreated individuals than in those on ART (14), as expected. We found one match between plasma RNA sequences and a cell with high levels of ca-HIV RNA, suggesting that high-expressing cells may be a source of plasma virus. This finding supports the idea that cells with high levels of viral RNA may be associated with viral replication, which could provide another tool to investigate the possibility of HIV-1 replication in tissues during ART.

Overall, the results of this study provide direct evidence that HIV-1 infected cell clones are able to persist despite ART because a high fraction of the cells in these clones have transcriptionally-silent proviruses. The results also show that latently-infected cells within clones harboring intact proviruses exist in multiple CD4+ memory subsets and are an important component of the HIV-1 reservoir; however, only a small fraction of each clone is transcriptionally-active at any one time, and potentially capable of producing infectious virus that can seed rebound if ART is stopped. In depth characterization of the HIV-1 reservoir and latently-infected cell clones, as performed here, may help define the appropriate goals and therapeutic approaches to achieve an HIV-1 cure in the future.

REFERENCES

- Maldarelli F, Wu X, Su L, Simonetti FR, Shao W, Hill S, Spindler J, Ferris AL, Mellors JW, Kearney MF, Coffin JM, Hughes SH. 2014. HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. Science 345:179–183.
- Simonetti FR, Sobolewski MD, Fyne E, Shao W, Spindler J, Hattori J, Anderson EM, Watters SA, Hill S, Wu X, Wells D, Su L, Luke BT, Halvas EK, Besson G, Penrose KJ, Yang Z, Kwan RW, Van Waes C, Uldrick T, Citrin DE, Kovacs J, Polis MA, Rehm CA, Gorelick R, Piatak M, Keele BF, Kearney MF, Coffin JM, Hughes SH, Mellors JW, Maldarelli F. 2016. Clonally expanded CD4+ T cells can produce infectious HIV-1 in vivo. Proc Natl Acad Sci U S A 113:1883–1888.
- Brown TR. 2015. I Am the Berlin Patient: A Personal Reflection. AIDS Res Hum Retroviruses 31:2–3.
- Pernet O, Yadav SS, An DS. 2016. Stem Cell--Based Therapies for HIV/AIDS. Adv Drug Deliv Rev 103:187–201.
- Gupta RK, Abdul-Jawad S, McCoy LE, Mok HP, Peppa D, Salgado M, Martinez-Picado J, Nijhuis M, Wensing AMJ, Lee H, Grant P, Nastouli E, Lambert J, Pace M, Salasc F, Monit C, Innes AJ, Muir L, Waters L, Frater J, Lever AML, Edwards SG, Gabriel IH, Olavarria E. 2019. HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation. Nature.

- Chawla A, Wang C, Patton C, Murray M, Punekar Y, de Ruiter A, Steinhart C.
 2018. A Review of Long-Term Toxicity of Antiretroviral Treatment Regimens and Implications for an Aging Population. Infect Dis Ther 7:183–195.
- Ho Y-C, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DIS, Lai J, Blankson JN, Siliciano JD, Siliciano RF. 2013. Replication-competent non-induced proviruses in the latent reservoir increase barrier to HIV-1 cure. Cell 155:540–551.
- Kearney MF, Wiegand A, Shao W, Coffin JM, Mellors JW, Lederman M, Gandhi RT, Keele BF, Li JZ. 2016. Origin of Rebound Plasma HIV Includes Cells with Identical Proviruses That Are Transcriptionally Active before Stopping of Antiretroviral Therapy. J Virol 90:1369–1376.
- 9. Josefsson L, von Stockenstrom S, Faria NR, Sinclair E, Bacchetti P, Killian M, Epling L, Tan A, Ho T, Lemey P, Shao W, Hunt PW, Somsouk M, Wylie W, Douek DC, Loeb L, Custer J, Hoh R, Poole L, Deeks SG, Hecht F, Palmer S. 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 U S A 110:E4987-4996.
- Wagner TA, McLaughlin S, Garg K, Cheung CYK, Larsen BB, Styrchak S, Huang HC, Edlefsen PT, Mullins JI, Frenkel LM. 2014. HIV latency. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. Science 345:570–573.

- 11. Lorenzi JCC, Cohen YZ, Cohn LB, Kreider EF, Barton JP, Learn GH, Oliveira T, Lavine CL, Horwitz JA, Settler A, Jankovic M, Seaman MS, Chakraborty AK, Hahn BH, Caskey M, Nussenzweig MC. 2016. Paired quantitative and qualitative assessment of the replication-competent HIV-1 reservoir and comparison with integrated proviral DNA. Proc Natl Acad Sci U S A 113:E7908–E7916.
- 12. Bui JK, Halvas EK, Fyne E, Sobolewski MD, Koontz D, Shao W, Luke B, Hong FF, Kearney MF, Mellors JW. 2017. Ex vivo activation of CD4+ T-cells from donors on suppressive ART can lead to sustained production of infectious HIV-1 from a subset of infected cells. PLoS Pathog 13:e1006230.
- Einkauf KB, Lee GQ, Gao C, Sharaf R, Sun X, Hua S, Chen SMY, Jiang C, Lian X, Chowdhury FZ, Rosenberg ES, Chun T-W, Li JZ, Yu XG, Lichterfeld M. 2019. Intact HIV-1 proviruses accumulate at distinct chromosomal positions during prolonged antiretroviral therapy. J Clin Invest 129:988–998.
- 14. Spina CA, Anderson J, Archin NM, Bosque A, Chan J, Famiglietti M, Greene WC, Kashuba A, Lewin SR, Margolis DM, Mau M, Ruelas D, Saleh S, Shirakawa K, Siliciano RF, Singhania A, Soto PC, Terry VH, Verdin E, Woelk C, Wooden S, Xing S, Planelles V. 2013. An in-depth comparison of latent HIV-1 reactivation in multiple cell model systems and resting CD4+ T cells from aviremic patients. PLoS Pathog 9:e1003834.
- Pollack RA, Jones RB, Pertea M, Bruner KM, Martin AR, Thomas AS, Capoferri AA, Beg SA, Huang S-H, Karandish S, Hao H, Halper-Stromberg E, Yong PC, Kovacs C, Benko E, Siliciano RF, Ho Y-C. 2017. Defective HIV-1 Proviruses Are

Expressed and Can Be Recognized by Cytotoxic T Lymphocytes, which Shape the Proviral Landscape. Cell Host Microbe 21:494-506.e4.

- 16. Deng K, Pertea M, Rongvaux A, Wang L, Durand CM, Ghiaur G, Lai J, McHugh HL, Hao H, Zhang H, Margolick JB, Gurer C, Murphy AJ, Valenzuela DM, Yancopoulos GD, Deeks SG, Strowig T, Kumar P, Siliciano JD, Salzberg SL, Flavell RA, Shan L, Siliciano RF. 2015. Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations. Nature 517:381–385.
- 17. Wiegand A, Spindler J, Hong FF, Shao W, Cyktor JC, Cillo AR, Halvas EK, Coffin JM, Mellors JW, Kearney MF. 2017. Single-cell analysis of HIV-1 transcriptional activity reveals expression of proviruses in expanded clones during ART. Proc Natl Acad Sci U S A 114:E3659–E3668.
- Palmer S, Kearney M, Maldarelli F, Halvas EK, Bixby CJ, Bazmi H, Rock D, Falloon J, Davey RT, Dewar RL, Metcalf JA, Hammer S, Mellors JW, Coffin JM. 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.
- Cillo AR, Vagratian D, Bedison MA, Anderson EM, Kearney MF, Fyne E, Koontz D, Coffin JM, Piatak M, Mellors JW. 2014. Improved Single-Copy Assays for Quantification of Persistent HIV-1 Viremia in Patients on Suppressive Antiretroviral Therapy. J Clin Microbiol 52:3944–3951.

- Tosiano MA, Jacobs JL, Shutt KA, Cyktor JC, Mellors JW. 2019. 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 57.
- Hilldorfer BB, Cillo AR, Besson GJ, Bedison MA, Mellors JW. 2012. New tools for quantifying HIV-1 reservoirs: plasma RNA single copy assays and beyond. Curr HIV/AIDS Rep 9:91–100.
- Laird GM, Rosenbloom DIS, Lai J, Siliciano RF, Siliciano JD. 2016. Measuring the Frequency of Latent HIV-1 in Resting CD4⁺ T Cells Using a Limiting Dilution Coculture Assay. Methods Mol Biol Clifton NJ 1354:239–253.
- 23. Salantes DB, Zheng Y, Mampe F, Srivastava T, Beg S, Lai J, Li JZ, Tressler RL, Koup RA, Hoxie J, Abdel-Mohsen M, Sherrill-Mix S, McCormick K, Overton ET, Bushman FD, Learn GH, Siliciano RF, Siliciano JM, Tebas P, Bar KJ. 2018. HIV-1 latent reservoir size and diversity are stable following brief treatment interruption. J Clin Invest 128:3102–3115.
- Kim M, Hosmane NN, Bullen CK, Capoferri A, Yang H-C, Siliciano JD, Siliciano RF. 2014. A primary cell model of HIV-1 latency that uses activation through the T cell receptor and return to quiescence to establish latent infection. Nat Protoc 9:2755–2770.
- Wang Z, Gurule EE, Brennan TP, Gerold JM, Kwon KJ, Hosmane NN, Kumar MR, Beg SA, Capoferri AA, Ray SC, Ho Y-C, Hill AL, Siliciano JD, Siliciano RF. 2018.

Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane. Proc Natl Acad Sci U S A 115:E2575–E2584.

- 26. Kearney MF, Anderson EM, Coomer C, Smith L, Shao W, Johnson N, Kline C, Spindler J, Mellors JW, Coffin JM, Ambrose Z. 2015. Well-mixed plasma and tissue viral populations in RT-SHIV-infected macaques implies a lack of viral replication in the tissues during antiretroviral therapy. Retrovirology 12:93.
- 27. Imamichi H, Dewar RL, Adelsberger JW, Rehm CA, O'Doherty U, Paxinos EE, Fauci AS, Lane HC. 2016. Defective HIV-1 proviruses produce novel proteincoding RNA species in HIV-infected patients on combination antiretroviral therapy. Proc Natl Acad Sci U S A 113:8783–8788.
- 28. Hiener B, Horsburgh BA, Eden J-S, Barton K, Schlub TE, Lee E, von Stockenstrom S, Odevall L, Milush JM, Liegler T, Sinclair E, Hoh R, Boritz EA, Douek D, Fromentin R, Chomont N, Deeks SG, Hecht FM, Palmer S. 2017. Identification of Genetically Intact HIV-1 Proviruses in Specific CD4+ T Cells from Effectively Treated Participants. Cell Rep 21:813–822.
- Grau-Expósito J, Serra-Peinado C, Miguel L, Navarro J, Curran A, Burgos J, Ocaña I, Ribera E, Torrella A, Planas B, Badía R, Castellví J, Falcó V, Crespo M, Buzon MJ. 2017. A Novel Single-Cell FISH-Flow Assay Identifies Effector Memory CD4+ T cells as a Major Niche for HIV-1 Transcription in HIV-Infected Patients. mBio 8.