UC San Diego
UC San Diego Electronic Theses and Dissertations

Title
Single Genome Amplification of Full-length pol Shows HIV Drug Resistance Exists Throughout Multiple Anatomical Tissues

Permalink
https://escholarship.org/uc/item/7963s9xs

Author
Chen, Aimee Lu

Publication Date
2020

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Single Genome Amplification of Full-length *pol* Shows HIV Drug Resistance Exists Throughout Multiple Anatomical Tissues

A thesis submitted in partial satisfaction of the requirements
for the degree of Master of Science

in

Biology

by

Aimee Chen

Committee in charge:

Professor David M. Smith, Chair
Professor Matthew Daugherty, Co-chair
Professor Justin Meyer

2020

The Thesis of Aimee Chen is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

____

____
Co-Chair

____
Chair

University of California San Diego

2020

TABLE OF CONTENTS

SIGNATURE PAGE iii
TABLE OF CONTENTS iv
LIST OF TABLES v
LIST OF FIGURES vi
ACKNOWLEDGEMENTS vii
ABSTRACT OF THESIS viii
INTRODUCTION 1
MATERIALS AND METHODS 11
RESULTS 17
DISCUSSION 27
REFERENCES 34

LIST OF TABLES

Table 1. Overview of LG03 and LG05. 11

Table 2. Tissues and samples collected from LG03 and LG05. 13

Table 3. Frequencies of DRMs and HIV sequences as represented across different tissues in LG03 and LG05. 20

Table 4. Overview of DRMs found in different tissues of LG03 and LG05. 21

Table 5. Frequencies of the types of DRMs found in various tissues of LG03 and LG05. 22

Table 6. Comparison of the M184V mutation in LG03 tissues. 26

Table 7. Comparison of the M184V mutation in LG05 tissues. 26

LIST OF FIGURES

Figure 1. General schematics of the HIV-1 genome. The viral genome is composed of three major genes (gag, pol, and env) flanked by LTR regions. Each gene is specific for encoding different functions for the virus with focus on pol, which encodes for the three main enzymes used by HIV-1. 2

Figure 2. Last Gift study design. HIV-infected terminally-ill participants in the LG cohort who are taking ART are enrolled in the study. For these patients, we will examine HIV populations and DRMs in blood while on ART (green) and in anatomical tissues upon death. 12

Figure 3. Pipeline to analyze HIV-1 pol sequences and the two sets of HIV primers used for PCR. (A) After the participants have passed away, tissues and samples are collected through a rapid autopsy protocol. DNA is extracted from those collected samples and subject to double rounds of PCRs. 15

Figure 4. Phylogenetic analysis of LG03 and LG05. IQ-TREE was used to create phylogenies for FL HIV pol sequences obtained from PBMC and multiple anatomical tissues in LG03 and LG05 after rapid autopsy. 18

Figure 5. Clinical histories of LG03 and LG05. The clinical histories of LG03 and LG05 are shown along with ART regimens they have taken post-HIV diagnosis. The timeline shows significant events from HIV diagnosis to autopsy. 23

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Davey Smith for his great support as the chair of my committee. I appreciate the valuable knowledge and skills that I have been able to gather during my time in his lab. This process would not have been possible without Dr. Smith.

I would like to acknowledge all members of the Smith lab. Without them, my research definitely would have taken much longer and been much more difficult. They have all been such good mentors and have taught me plenty during my time in the lab.

I would also like to acknowledge my immediate mentor, Sergio Ita, for always being there to help and guide me. He was willing to sit down, for as long as necessary, to discuss and figure things out with me. Sergio was always patiently answered my endless questions and all of this would have been impossible without him.

The figures in the introduction, methods, and results sections are currently unpublished but are coauthored with Ita, Sergio; Chaillon, Antoine; Kamalu, Maris; Baham, Jonny; Porrachia, Magali; Ignacio, Caroline; Rawlings, Stephen; Wagner, Gabriel; Gianella, Sara; Smith, Davey. The thesis author was a coauthor of these figures.

ABSTRACT OF THESIS

Single Genome Amplification of Full-length *pol* Shows HIV Drug Resistance Exists Throughout Multiple Anatomical Tissues

by

Aimee Chen

Master of Science in Biology

University of California San Diego, 2020

Professor David M. Smith, Chair
Professor Matthew Daugherty, Co-Chair

Research on human immunodeficiency virus type I (HIV-1) has led to the development and use of different antiretroviral therapies (ART) to effectively inhibit viral replication and to decrease the transmission rates. However, poor adherence or interruption of therapy may select for drug resistance which leads to a decrease in the overall effectiveness of ART. Even though there is research performed to understand drug resistance mutations through blood sampling, there is a need to understand HIV drug resistance in tissues outside of blood in people living with HIV. For this study, various tissues of two participants with HIV in the Last Gift cohort were

examined. HIV drug-resistance mutations and viral populations were obtained through single-genome amplification of the full-length HIV *pol* gene. The results obtained showed that there were many sequences found within and across various tissues that had nearly identical sequences, which might be due to viral migration. The lymph nodes and spleen exhibited high levels of DRMs to more than one class of ART. We also observed that drugs taken early after HIV diagnosis possibly selected for DRMs that were archived in HIV reservoirs. Also, PBMCs from whole blood did not have certain DRMs that were present in other tissues of the corresponding participants. Overall, the study suggests blood may not be fully capable of characterizing viral reservoirs in other anatomical locations. These findings, in hopes, can provide aid in improving current ART regimens for people living with HIV worldwide.

INTRODUCTION

The human immunodeficiency virus (HIV) is perhaps one of the most well-known infectious pathogens in the world. This RNA virus belongs to a family of retroviruses, including a subfamily of lentiviruses, that is known to infect humans [1]. HIV infection is also recognized to be the underlying cause of acquired immunodeficiency syndrome (AIDS) which is characterized by a compromised immune system that can lead to various opportunistic diseases and infections [1]. Overall, there are two major types of HIV: HIV-1 and HIV-2. HIV-1 is the most common, consisting of multiple viral groups: M, N, O, and P [1]. The M group is more widely spread and is responsible for the HIV/ AIDS pandemic worldwide [1]. Reasoning behind this is that the M group harbors various recombinant forms that aid in creating a selective advantage that further facilitates the spread and persistence of the virus [1]. HIV-2, on the other hand, is not as common or infectious as HIV-1; it was first discovered in 1986 to be confined to West Africa [1]. HIV-1 emerged in humans through cross-species transmission of a simian immunodeficiency virus (SIV) that naturally infects chimpanzees, while HIV-2 from exposure to sooty mangabeys [1]. The general course of HIV infection begins with a primary infection, which causes flu-like symptoms in the host, then progresses to the asymptomatic stage, and finally AIDS, in which the person becomes highly susceptible to other infections and diseases [1].

Viral Structure

HIV contains multiple envelope glycoproteins consisting of gp120 surface subunits that protrude from the surface of the membrane and gp40 subunits that is embedded in the viral membrane [1]. Inside the lipid membrane, there is an inner protein shell (matrix) and a p24 core

protein or capsid, which encompasses genomic viral RNA and the reverse transcriptase enzyme (RT) that will subsequently be used to convert host DNA into more viral RNA [1].

The HIV-1 Genome

The virus is comprised of genes that aid in the production of crucial proteins for its survival and infectious activity. There are three major genes involved: *gag, pol,* and *env* [1, Fig.1]. *Gag* is responsible for producing proteins that form the core of the virus [1]. *Pol* is a gene that codes for protease, integrase, and reverse transcriptase, which are the three main enzymes that are utilized by the virus [1]. *Env,* on the other hand, encodes for the envelope glycoproteins [1]. HIV contains six other genes, making it much more complex than other retroviruses [1].

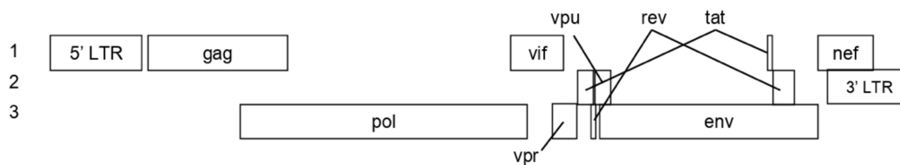


Figure 1. General schematics of the HIV-1 genome. The viral genome is composed of three major genes (gag, pol, and env) flanked by LTR regions. Each gene is specific for encoding different functions for the virus with focus on pol, which encodes for the three main enzymes used by HIV-1: protease (PR), reverse transcriptase (RT), and integrase (IN).

With its complexity, HIV is highly prone to acquiring mutations and alterations in coding sequences. Mutations can arise through base substitutions, insertions, and deletions of nucleotides in addition to the numerous errors produced by the HIV RNA polymerases [1-3]. The massive genetic variability of this virus also occurs through the high frequency of genetic recombination [4, 5]. The fast replication rates, excess mutations, and genetic recombination contribute to the rise of heterogeneity and minor HIV variants, which will be explained later in further detail.

The General HIV-1Replication Cycle

The replication cycle of HIV is fascinating in the way that it is able to rewire the host's cellular activities to produce more viral copies which can continue to infect other cells. The replication process begins with the binding of the external gp120 protein to receptors that are present on the host cell surface [1]. The receptors that gp120 binds to is the CD4 molecule that is present on T lymphocytes, and other immune cells, that help to aid the immune system [1]. Once gp120 binds to CD4 on the surface of host cells, it goes through a conformational change that increases its efficiency of binding to one of the two co-receptors: CCR5 or CXCR4 [1]. Then the gp41 protein is able to break through the host cell's plasma membrane to allow the virus membrane to fuse with the host cell membrane [1]. Upon fusion, the shell of the capsid unravels and enables the viral RNA and RT enzyme to enter the host cell for reverse transcription to convert viral RNA into double-stranded viral DNA [1]. Next, the viral enzyme integrase enables the double-stranded viral DNA to be incorporated into the host cell's DNA [1]. After integration of viral DNA, the combination and interaction of HIV proteins, enzymes, and genomic RNA can form a new viral particle at the plasma membrane of the host cell [1]. It is through the host cell's plasma membrane that the newly formed pre-mature virion is able to bud off [1]. As the last step, the viral protease enzyme cleaves the gag-pol precursor to transform the new virion into a mature virion which can circulate and infect other host cells to repeat this replication process [1]. In the average untreated HIV-positive patient, the viral replication occurring in the body can produce 10^{10} -10^{11} virions per day [6]. However, every step of this replication cycle, including the viral enzymes involved, is important for the production of new infectious viruses and has the potential to be targeted and inhibited by antiviral drugs and combination therapy.

HIV Transmission

As an infectious pathogen, there are many methods in which HIV can be transmitted in humans. The main course of HIV transmission is through sexual contact, whether it is heterosexual or male-male contact [1]. Often, different body fluids are involved in sexual contact and exposure to HIV-containing bodily secretions can heighten the risk of viral transmissions. It was found that the virus has the tendency to concentrate and collect in the seminal fluid [1]. There is a higher risk of viral transmission in homosexual intercourse than heterosexual intercourse [1]. In addition, there is a greater transmission rate from male-to-female than female-to-male [1]. Studies have revealed that the presence of existing sexually transmitted infections, such as chlamydia and gonorrhea, factors into the increased risk of HIV infection [1].

Apart from transmission through sexual contact, one can also be infected by HIV through blood and similar blood products [1]. It is widely seen in patients who have received HIV-contaminated blood, blood products, or transplanted tissue [1]. It is estimated that over 90% of people who came into contact with or received HIV-positive blood, such as through blood transfusion for example, become infected [1]. The probability-per-act of such blood transfusions is 9,250 positive cases out of 10,000 individuals (92.5%) [1]. The greater amount of viable HIV in the plasma, the greater the risk of infection. To prevent and decrease the risk of infection by blood, donated blood is now screened and tested for HIV-1 and HIV-2 while blood donors are cautiously chosen by utilizing health history questionnaires [1].

Another commonly known method of HIV transmission is the sharing of needles and such intravenous equipment without proper sterilization between drug-users [1]. Penetrating infected tissues and coming into contact with infected blood can contaminate needles and other drug-use related equipment. Although the per-act risk of HIV infection for injection drug-use is

much lower (~0.6%) [1], numerous people are highly susceptible to this risk or have already been infected.

A lesser-known means of transmission is maternal transmission in which an infected mother can pass along the virus to her fetus which typically happens during the first or second trimesters of pregnancy [1]. If the mother has not received antiretroviral therapy in an appropriate form, the per-act of HIV-1 transmission to her child is approximately 15%-25% in most advanced countries [1].

Although there are multiple methods of HIV-1 transmission in humans, it has not been confirmed that such transmission can occur through surface-level skin-to-skin contact or exposure to salivary fluid, tears, sweat, or urine [1]. Even if the risk of infection seems significantly lower than other infectious pathogens, HIV-1 has been declared a global pandemic and major efforts are continuously being made to abolish any of risk of infection and transmission.

HIV Pathology

The most devastating result of an established HIV-1 infection is having a compromised immune system due to the lack of sufficient T lymphocytes, or otherwise known as helper T cells [1]. The CD4 receptors that HIV recognizes to begin the initial attack on host cells are the defining characteristics of helper T cells [1]. To put it simply, HIV is a pathogen that targets and attacks CD4+ helper T cells. As observed in people living with HIV-1, the immunodeficiency caused by loss of CD4+ T cells occurs through different ways. The virus can directly infect and annihilate CD4+ T cells even while the immune system destroys some HIV-infected cells [1]. Furthermore, cellular apoptosis can also occur from dysfunctional immune activation and immune fatigue can contribute to low levels of CD4+ T cells [1]. As mentioned earlier, HIV can

lead to the development of AIDS. Those who are considered immunocompromised have CD4+ T cell levels that are below a numerical threshold and therefore, are particularly susceptible to AIDS and life-threatening opportunistic diseases [1].

Upon initial HIV-1 infection, the body has natural mucous barriers that aid in blocking further infection [1]. Viruses that have successfully entered the host body then target resting and activated CD4+ T cells to attempt to replicate and increase proliferation [1]. The typical human body contains fewer activated than resting T cells, but active T cells are capable of producing more of the virus [1]. For an established HIV-1 infection, each individual infected cell is able to infect at least one other cell [1]. Within the first few days of infection, HIV cannot be detected in the plasma, but after a period of days up to weeks, the virus disperses to the lymph nodes and then to other CD4+ T cell-rich locations in the body [1]. The simultaneous circulation and replication of HIV greatly increases the viral count. The higher the viral count, the higher the chance of transmission.

When viremia levels are high and infection is permanent, HIV-1 is able to evade the normal immune response in a host's body [1]. Since HIV-1 targets CD4+ T cells, which regularly play an important role in immunity, the body is unable to produce an effective immune response to eradicate the viruses and prevent viral reproduction [1]. Due to the escape of immune responses by HIV, nearly all infected people have latently infected reservoirs of the virus which contain resting CD4+ T cells that have host-integrated HIV DNA [1]. These resting CD4+ T cells are inactivated and remain in this state until an activation event triggers functional replication [1]. It is due to this latent reservoir that the infection becomes chronic and lasts for the host's lifetime. Sometimes viruses in the latent reservoir do not respond well to treatment and ART, causing the CD4+ T cell count to drop below 200 cells/mm^3 in which at this stage, the

patient has developed AIDS [1]. As the CD4+ T cell count can continue to decrease from this point, the lower the T cell count the greater the likelihood of developing opportunistic diseases and infections.

Antiretroviral Therapy and HIV-1 Drug Resistance

Over the past few decades, ART has been used as a global treatment for patients living with HIV. There are a few classes of ART that have been developed thus far and each type of drug functions differently to suppress further HIV infection and replication. The classes of ART include: nucleoside/ nucleotide reverse transcriptase inhibitors (NRTI), non-nucleoside reverse transcriptase inhibitors (NNRTI), protease inhibitors (PI), integrase inhibitors, fusion inhibitors, and CCR5 antagonists [1]. NRTIs function by inhibiting the RT enzyme from creating viral DNA from host RNA [1]. While NNRTIs also prevent the production of HIV DNA, it works by binding directly to the RT enzyme [1]. Protease inhibitors and integrase inhibitors work to interfere with their respective enzyme targets while fusion inhibitors and CCR5 antagonists interrupt HIV entry into the host cells [1]. Although monotherapy is possible, it is highly recommended that combinations of drugs be given to HIV-infected patients in order to reduce the likelihood of acquired drug-resistant mutations (DRMs) that can cause treatment failures [7]. However, it was observed in some studies that poor adherence to ART can cause a rise in HIV minority variants harboring drug-resistant mutations [8-10]. Nonetheless, ART is a crucial factor for prolonging life for those living with HIV.

As mentioned earlier, HIV goes through fast, erroneous replication that can result in numerous acquired mutations. These mutated strains of virus contribute to its vast genetic diversity. Due to this, a single human can have various HIV strains simultaneously, making them difficult to eradicate [1]. When humans are infected with the wild-type virus, it can acquire many

mutations to diverge into genetically different forms [1]. These HIV variants that reside in reservoirs have the potential to become resistant to certain drugs and pose a massive complication to treatment. Even though variants are typically less infectious, there have been studies that confirmed the ability of drug-resistant minority variants to be transmitted [11, 12]. Minority variants exist in all HIV-infected individuals and can be present in nearly every part of the human body; the type and quantity of variants at each location differ depending on certain selective pressures [13]. It is important to gain a thorough understanding of HIV minority variants to be able to construct broader and more effective treatment options for patients worldwide.

SHIV and Spatial Dispersion of Drug-Resistant HIV

Many early studies were performed to observe HIV minority variants as present in blood. However, not much is known about the behavior of HIV-1 in different parts of the body. As an approach to further understand HIV minority variants, Feder et al. conducted a study on the spread of drug resistance in simian/human immunodeficiency viruses (SHIV) across different tissues in macaques. A chimeric SIV-HIV virus (RT-SHIV) was studied in infected macaques due to its relatability to HIV-1 in humans [14]. The goal of this study was to understand how drug-resistant mutations of the virus develop and how they disperse throughout tissues in a host's body [14].

The experiment involved four RT-SHIV-infected macaques that underwent daily antiretroviral therapy (monotherapy) to induce the production of drug-resistant mutations [14]. vRNA samples from the blood, lymph nodes, gut, and vagina were taken before, during, and after the appearance of drug-resistant variants and the vRNA samples were sequenced using single-genome sequencing [14]. Results have shown that in the 3 macaques displaying treatment

resistance, drug-resistant mutations in the plasma and the other tissues have surfaced around the same time [14]. However, the frequency of drug resistance in the different compartments varied [14]. There was lower drug resistance in the lymph nodes than the PBMC (peripheral blood mononuclear cells; from blood), plasma, gut and vagina [14]. Feder et al. explained that this might be caused by slower replication of the virus in the lymph nodes and the fact that there is less selective pressure due to the weak circulation of drugs throughout the lymph nodes [14-16]. On the other hand, the gut showed a much higher level of vDNA DRMs than all other tissues and blood [14]. The common drug-resistant mutation also seen in HIV-1 in humans, M184V, was observed to be in all the compartments sampled [14]. This suggests that viral migration was possible and able to spread DRMs from one location to another [14]. Even though identical variant genotypes were noted at multiple compartments, the quantities, or frequencies of those genotypes were different between locations [14]. The researchers defined this phenomenon, in which two compartments have different variant frequencies, as compartmentalization [14]. Compartmentalization throughout periods of time is constantly changing and evolving [14]. As concluded, the most significant differences in compartmentalization was seen in lymph nodes and mucosal tissues such as the gut and vagina or between the gut and vagina themselves [14].

The previous study performed by Feder et al. greatly aided in providing sufficient information and a visual imagery of how HIV-1 minority variants behave and exist in different tissues. It is now known that drug-resistant minority variants do not only reside in blood; they are distributed among many tissues, such as lymph nodes and mucosal tissues, and have the ability to further acquire drug-resistant mutations. Even though this experiment provided useful insight as to minority variants in certain tissues of macaque-SHIV models, not much is known about how HIV-1 minority variants are present in human tissues. Although some studies examined

HIV in humans, characterization of HIV reservoirs was mainly done in easier-to-access samples such as the blood and gut [17-20], while there were still limitations in accessing fresh tissues of a broad variety. The Feder et al. study provided as a basis to our project and we ask the question: how and to what extent are HIV variants, with drug resistance mutations integrated into the genome of the host cell, in HIV reservoirs present in different human tissues? In addition, how does drug resistance differ between various tissues and how is it different from blood? Our goal is to utilize single-genome amplification to sequence and analyze the entire HIV *pol* gene from samples taken from human tissues for DRMs. Rather than just studying the lymph nodes, gut, and vaginal tissues, we will also sample tissues from several other organs, such as kidneys and spleen, from deceased HIV-infected individuals. The sequences obtained will ultimately be used to depict DRMs in the anatomical tissues and the results of this study will provide necessary insights on the patterns of HIV DRMs in human tissues. Since HIV-1 is a fast replicating and mutating retrovirus, the production of a vaccine is painstakingly difficult. Therefore, HIV-infected patients currently still rely heavily on ART as a treatment option. Hence, the outcomes of this study will hopefully inform healthcare professionals in improving the effectiveness of HIV treatments and regimens for patients so that they can maintain good, quality lives.

Acknowledgements: The figure in the introduction section is currently unpublished but is coauthored with Ita, Sergio; Chaillon, Antoine; Kamalu, Maris; Baham, Jonny; Porrachia, Magali; Ignacio, Caroline; Rawlings, Stephen; Wagner, Gabriel; Gianella, Sara; Smith, Davey. The thesis author was a coauthor of these figures.

MATERIALS AND METHODS

The study design included 15 terminally-ill HIV-positive patients that were actively participating in the Last Gift (LG) cohort and were receiving ART for a minimum of 48 weeks. The general procedure was to distinguish the HIV populations in their blood while taking ART and in multiple tissues in their bodies after they passed away [Fig. 2]. As a result, the cohort has enrolled 15 participants total in which we collected body tissue samples from 9 who have died. The contents of this thesis will primarily investigate participants LG03 and LG05 who have both continued receiving ART until time of death, making them highly suitable for the characterization of HIV in tissues [Table 1].

Table 1. Overview of LG03 and LG05.

ID	*Age*	*Diagnosis*	*Stopped ART*	*Last ART Regimen*	*Last CD4 count (cells/mm^3)*	*Last HIV RNA level*	*Deceased*
LG03	72	Pancreatic Tumor	No	FTC/TAF+DTG	330 (time of death)	undetectable	May 2018
LG05	57	ALS	No	TAF/FTC/RPV+DTG	347 (time of death)	undetectable	March 2019

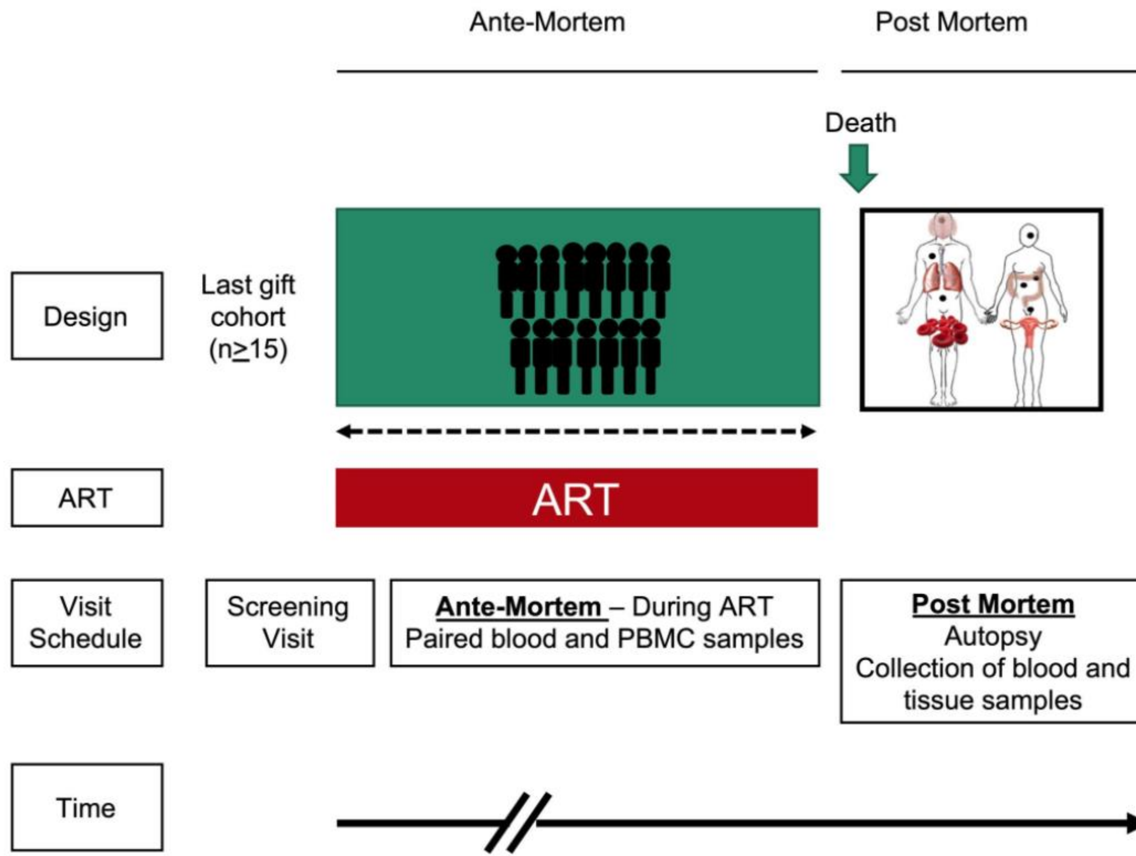


Figure 2. Last Gift study design. HIV-infected terminally-ill participants in the LG cohort who are taking ART are enrolled in the study. For these patients, we will examine HIV populations and DRMs in blood while on ART (green) and in anatomical tissues upon death.

For the participants ante-mortem, blood was collected and viral load was measured weekly using a BioRad ddPCR. The BioRad ddPCR helped to determine the HIV DNA quantities in the blood collected. For each participant, ART-use history was also collected for informational purposes. Post-mortem, the bodies were collected and a rapid autopsy protocol was performed on each participant by the *Clinical and Anatomic Pathology Core*. Various tissues and high volumes of blood were collected less than 12 hours from the time of death. Collecting such samples as early as possible ensured that cells remained viable and that nucleic acids are stable enough for upcoming experiments. Generally, from each participant, we attempted to collected blood plasma, PBMC, cerebral spinal fluid, urine, semen from seminal vesicles, cervical swabs,

stool samples, portions of the brain, eyes, spinal cord, muscle, peripheral nerves, gut, lymph nodes, spleen, liver, bone marrow, heart, kidney, lungs, and adipose tissue. Table 2 notes the specific tissues and samples obtained from LG03 and LG05. The collected body samples were snap frozen and fixed in formalin to be preserved. After tissue and blood samples were collected, DNA was extracted from those samples.

Table 2. Tissues and samples collected from LG03 and LG05.

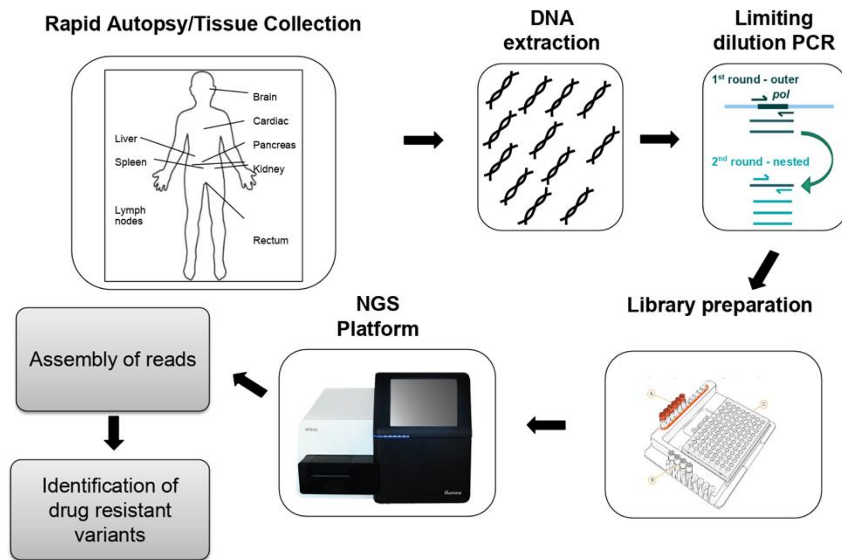
LG03	*LG05*
PBMC	PBMC
Prostate	X
Colon (Right)	X
Spleen	Spleen
Kidney	Kidney
Axillary Lymph Node (LNX)	Axillary Lymph Node (LNX)
Aortic Lymph Node (LNA)	X
Duodenum	Duodenum
Ileum	Ileum
Testes	X
X	Jejunum

SGA of Full Length HIV pol

Within the HIV DNA, there is the *pol* gene which includes three main enzymes required for the virus's survival and replication: integrase, protease, and reverse transcriptase. In order to continue the study of drug resistance in HIV, it was necessary to create a protocol that allowed for the amplification of the *pol* gene, which is approximately 3.5 kilobases long [Fig. 3A]. From

the blood plasma and tissue samples obtained, HIV copy number for each specimen was generated through droplet digital PCR (ddPCR-Bio-Rad QX200 Droplet Reader) using HIV DNA. The copy number was then utilized to calculate the amount of HIV DNA template needed to yield one single template per PCR reaction using full length (FL) *pol* specific primers [Fig. 3B]. A first-round limiting dilution PCR (Advantage 2 PCR kit, Takara) was performed with HIV-1 *pol*-specific primers. Another round of PCR using an internal set of *pol*-specific primers was used to further amplify 1 μ*l* of the first round of PCR products [Fig. 3B]. The second-round PCR products were screened through gel electrophoresis to determine which wells in the PCR plate yielded FL *pol* amplicons. This was generally characterized by a single ~3.5 kb pair products around the 3 kilobase (kb) mark when compared to the NEB 1 kb pair standard. Single genome amplification is obtained when approximately 30% of the total number of wells produce correct-sized products (based on a Poisson distribution). If ~30% of total PCR reactions yield product for a tested tissue sample, those products were moved forward to the sequencing step. The amplicons from the positive wells were then PCR purified and run through the TapeStation to confirm that the amplicons were, in fact, the correct length. The second-round PCR purified products were then tested using the Qubit dsDNA HS Assay Kit to determine the concentrations of DNA. The library underwent Illumina sequencing to create short reads (150 bp or 350 bp) that were then analyzed with an in-house bioinformatics pipeline.

A



B

Outer primer set						
Primer ID	Primer Name	Sequence	Length	Positions	Tm	GC content
pol_outerF	Pan-HIV-1_2F	GGGAAGTGA**Y**ATAGC**W**GGAAC	21	1031-1051	52.4	48
pol_outerR	R3580	GCAGGTGATGATTGTGTGGC	20	5055-5074	53.8	55
Inner primer set						
Primer ID	Primer Name	Sequence	Length	Positions	Tm	GC content
pol_innerF	Pol-F2/F1736	GGATGACAGAAACCTTGTTGG	21	1736-1756	53.9	48
pol_innerR	Pan-HIV-1_2R	CTGCCATCTGTTTTCCATA**R**TC	22	4604-4583	5.1	41

Figure 3. Pipeline to analyze HIV-1 pol sequences and the two sets of HIV primers used for PCR. (A) After the participants have passed away, tissues and samples are collected through a rapid autopsy protocol. DNA is extracted from those collected samples and subject to double rounds of PCRs. Final round PCR products from different samples are all purified and barcoded to obtain single amplicon library. Samples are sequenced using Illumina MiSeq and a bioinformatics pipeline will compile high-quality pol reads. Lastly, variations in HIV-1 populations with DRMs can be identified. (B) The first round of PCR utilizes an outer primer set that flanks the pol region while the second round of PCR requires a nested set of primers. Sequence letters in bold indicate sites where there can be varying nucleotides.

Bioinformatics approach

A computational analysis pipeline was used to analyze HIV pol read data. The analysis screened for reads that met the minimum quality requirement. Meaning, the sequences obtained by Illumina sequencing were assembled to the reference HIV viral sequence HXB2 to determine sequence variations. If a segment of DNA was sufficiently matched with the HXB2 reference

sequence, it has met the minimum quality requirement. All reads from samples were further matched to the reference sequence to produce a consensus sequence which has the most common base pair at each position. Then, reads from the samples were then aligned to the consensus sequence to determine variability in the HIV *pol* sequences, which was characteristic of HIV variants. An algorithm that is able to analyze sites of drug mutations was used to identify drug resistance mutations within reverse transcriptase, protease, and integrase genes using a curated database of HIV drug resistance mutations (Stanford HIV Drug Resistance Database). To compare the frequencies of specific mutations between tissues in the two participants, the Prism version 8 software was utilized to perform Fisher's exact tests to determine significance.

Acknowledgements: The figures in the methods section are currently unpublished but are coauthored with Ita, Sergio; Chaillon, Antoine; Kamalu, Maris; Baham, Jonny; Porrachia, Magali; Ignacio, Caroline; Rawlings, Stephen; Wagner, Gabriel; Gianella, Sara; Smith, Davey. The thesis author was a coauthor of these figures.

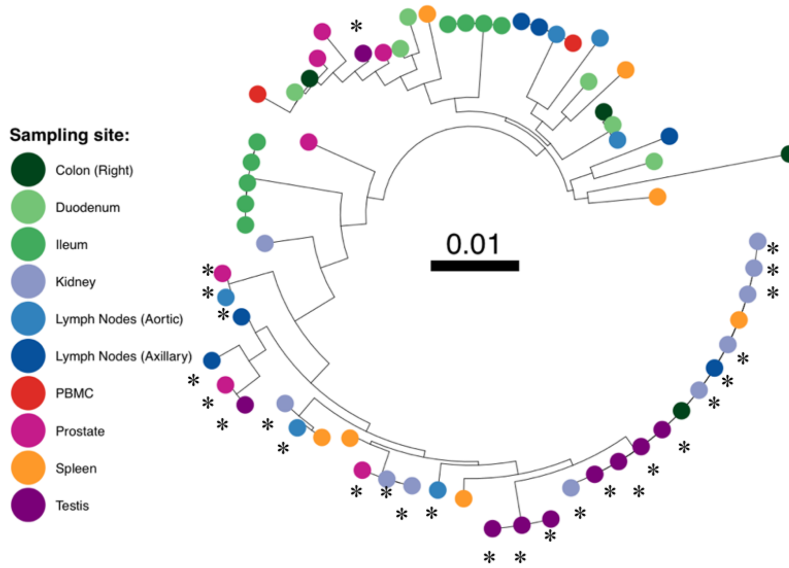
RESULTS

From the 7 deceased HIV-infected patients, we focused on working with LG03 and LG05, who both continued receiving ART until death. Since both patients had taken ART long-term, their CD4 cell counts around times of death were 330 cells/mm^3 and 347 cells/mm^3 respectively, indicating that infection did not progress into AIDS (<200 cells/mm^3) [Table 1]. In addition, both of these patients have received ART for at least 48 weeks. LG03 and LG05 did not have detectable HIV RNA levels in blood plasma at the time of death.

Phylogenetic trees were generated using the FL *pol* sequences obtained from PBMCs and various tissues in LG03 and LG05 [Fig. 4]. Each individual colored dot is representative of a single-genome sequences obtained from a specific sample. The distance scale of 0.01 represents the number of nucleotide substitutions per site. Therefore, the greater the distance between two variants, the more dissimilar the genotypes and sequences.

For participant LG03, we produced a total of 66 FL *pol* sequences. Out of these total sequences, there were the fewest number of sequences produced in colon right (4 sequences) and PBMC (2 sequences) [Fig. 4A]. However, the highest number of sequences was produced in the kidney (10), ileum (9), and the testis (9) [Fig. 4A]. One sequence from colon right was highly genotypically different compared to other variants found in other compartments due to a phylogenetic distance of approximately 0.02 [Fig.4A]. From a subset of the sequences, we found evidence of clonal expansion. The ileum was found to have a set of 5 sequences of nearly identical FL *pol* sequences (99% similarity) and a different set of 4 nearly-alike sequences. The testis indicated a set of 3 sequences and a separate group of 4 nearly identical sequences. In

addition, 2 sequences found in the LNX had almost identical FL *pol* sequences while the kidney had 2 sequences that were alike and 3 other similar sequences. 13 nearly identical sequences



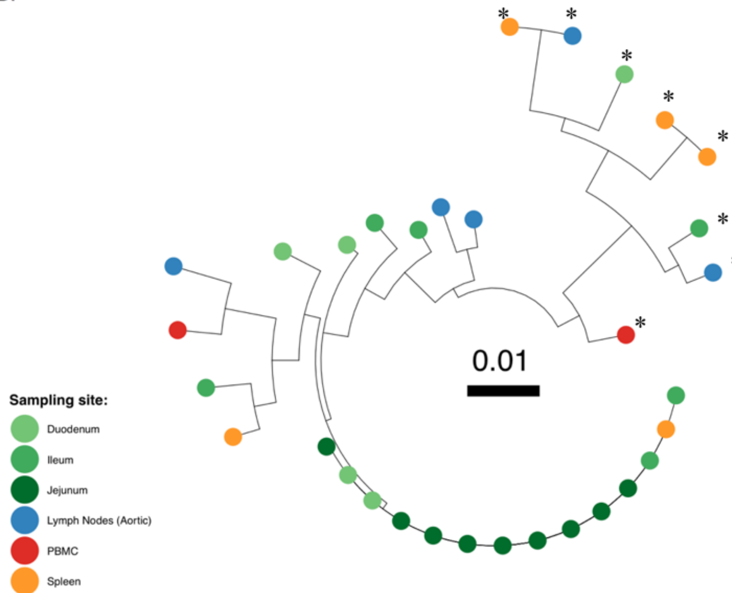


Figure 4. Phylogenetic analysis of LG03 and LG05. IQ-TREE was used to create phylogenies for FL HIV pol sequences obtained from PBMC and multiple anatomical tissues in LG03 and LG05 after rapid autopsy. Different compartments along with the corresponding colored dot are labeled under "sampling sites". Distances between dots symbolize the degrees of genotypic variance. (A) Phylogenetic tree for FL pol sequences found in LG03. (B) Phylogenetic tree for LG05. Colored dots displayed in series indicate nearly identical FL pol populations (99% identical). Asterisks (*) indicate variants with DRMs.

were discovered across 5 different compartments: kidney, testis, colon (R), LNA, and spleen [Fig. 4A]. The presence of matching sequences with similar DRM patterns in multiple compartments suggests that HIV-1 variants underwent clonal expansion and spread to other anatomical tissues and locations in the host's body.

For LG05, we generated 34 total FL *pol* sequences. From the total, the least number of sequences was produced in PBMC samples (2) while the highest number of sequences was generated in the jejunum (9) [Fig. 4B]. There exists a portion of viral FL *pol* that contained nearly identical sequences in the duodenum, jejunum, ileum, and spleen, again, signifying the possibility of variants spreading throughout the body [Fig. 4B]. The duodenum, jejunum, and ileum are all parts of the small intestine and collectively make up a portion of the gut. These particular tissues with identical sequences seem justified since they are in relatively close proximity to each other. In other words, if minority variants with DRMs have the ability to spread to other areas, the nearby tissues would likely to be affected. The lymph nodes and spleen were noted to also have sufficient numbers of viral sequences. The lymph nodes and spleen are parts of the lymphatic system and there were nearly identical FL *pol* sequences between the two sites.

Both LG03 and LG05 had sequences in multiple tissues that were not quite nearly identical but were very similar due to their short distances on the phylogenetic trees. This signifies that even though certain tissues had almost identical sequences, there is the possibility of dissimilarity, in terms of mutations, within the same compartment. Largely, the phylogenies of LG03 and LG05 demonstrate that some tissues had nearly identical sequences, with clustering of DRMs, which could have arisen due to clonal expansion and viral migration. Also, many FL *pol*

sequences in different tissues were not similar to each other, suggesting the notion that some HIV-1 DRMs and sequences can be specific to certain sites.

Table 3. Frequencies of DRMs and HIV sequences as represented across different tissues in LG03 and LG05.

Participants	Tissue	# of SGA sequences	PI	NRTI	NNRTI
			% of sequences with DRM	**% of sequences with DRM**	**% of sequences with DRM**
LG03	PBMC	2	0	0	0
LG03	Prostate	7	0	42.85	0
LG03	Colon right	4	0	25	50
LG03	Spleen	7	0	71.43	28.57
LG03	Kidney	10	0	90	60
LG03	Axillary LN	6	0	50	16.67
LG03	Aortic LN	6	0	50	16.67
LG03	Duodenum	6	0	0	0
LG03	Ileum	9	0	0	0
LG03	Testes	9	0	11.11	77.78
LG05	PBMC	2	50	50	0
LG05	Axillary LN	5	40	40	0
LG05	Spleen	5	60	60	0
LG05	Duodenum	5	20	20	0
LG05	Jejunum	9	0	0	0
LG05	Ileum	6	16.7	16.67	0
LG05	Kidney	2	0	50	50

Table 4. Overview of DRMs found in different tissues of LG03 and LG05.

Participant	LG03			LG05		
Tissue Type	PI	NRTI	NNRTI	PI	NRTI	NNRTI
PBMC	None	None	None	None	D67N,K70R,M184V	None
Prostate	None	M41L,D67N,K70R,M184V,T215F,K219Q/E	None			
ColonR	None	M41L,D67N,K70R,M184V,T215F,K219E	V108I,Y181C, M230I			
Spleen	None	M41L,D67N,K70R,M184V,T215F,K219Q/E	Y181C,G190A, V108I,Y181C	M46I,I54V,V82T,I84V,L90M	M41L,D67N,K70R,L74I,M184V,T215F,K219Q	A98G
Kidney	None	M41L,D67N,K70R,M184V,T215F,K219Q/E	V108I,Y181C	None	M41L,D67N,K70R,M184V,T215F,K219E	V108I,Y181C
LNX	None	M41L,D67N,K70R,M184V,T215F,K219Q/E	V108I,Y181C			
LNA	None	M41L,D67N,K70R,M184V,T215F,K219Q/E	Y181C,G190A	M46I,I54V,V82T,I84V,L90M	M41L,D67N,K70R,L74I,M184V,T215F,K219Q	A98G
Duodenum	None	None	None	M46I,I54V,V82T,I84V,L90M	M41L,D67N,K70R,L74I,M184V,T215F,K219Q	None
Ileum	None	None	None	M46I,I54V,V82T,I84V,L90M	M41L,D67N,K70R,M184V,T215F,K219Q	None
Testes	None	M41L,D67N,K70R,M184V,T215F,K219Q/E	V108I,Y181C			
Jejunum				None	None	None

The numbers of SGA sequences with DRMs found in different tissues in LG03 and LG05 were quantified in Table 3. The trends showed that LG03 displayed no sequences with DRMs against PI drugs in any of the tissues. However, we did find that LG03 had high percentages of sequences with DRMs against NRTI drugs; the highest was found in the kidney, in which 90% of the 10 SGA sequences showed DRMs against NRTIs. LG03 spleen was also observed to have a relatively high percentage of sequences with DRMs against NRTIs (71.43% of 7). As seen for NNRTIs, the highest numbers of sequences in the testes exhibited DRMs (77.78% of 9). The kidney appeared to have the highest number of sequences with DRMs against HIV drugs overall, while samples tested from PBMC, duodenum, and ileum have no sequences observed to be drug-resistant against any of the three classes of drugs. It is interesting that there were no DRMs against any of the drug classes observed in the duodenum and ileum, considering that sequences from the colon (R) did contain DRMs and that they all are parts of the gut tissues. In contrast, for LG05, majority of the sequences with DRMs were found for PIs and NRTIs with one in the kidney for NNRTI [Table 3]. There were equal frequencies of sequences with DRMs for both PIs and NRTIs. It can be seen that within one sequence, there were also PI and NRTI DRMs present at the same time on the same background [Table 4], which can suggest a linkage of these two

types of DRMs. There may also be a linkage of NNRTI and NRTI DRMs in LG03 [Table 4]. Generally, there were sequences with DRMs in more than one type of ART in various tissues [Table 4]. This proposes that a single variant can be resistant to more than one type of drug, which can be due to HIV acquiring diverse mutations that correspond to decreased susceptibility in multiple drug classes. Although LG05 had DRMs in the duodenum and ileum, none of the HIV sequences produced from the jejunum contained any DRMs to PIs, NRTIs, and NNRTIs. For both LG03 and LG05, no sequences in PBMCs had DRMs for NNRTI drugs.

After we quantified the HIV sequences with DRMs relative to different tissues and samples, we wanted to investigate the presence of specific DRMs that have appeared in those sequences and tissues. The results and data taken from single-genome amplification and sequencing were compiled and analyzed to provide frequencies of specific HIV DRMs that can be further examined using the Stanford HIV Drug Resistance Database [Table 5].

Table 5. Frequencies of the types of DRMs found in various tissues of LG03 and LG05.

Inhibitor and DRM	LG03		LG05	
	#	%	#	%
	NRTI			
M41L	32	50	8	23
M184V	31	48	9	26
D67N	31	48	9	26
K70R	31	48	9	26
T215F	30	47	8	23
K219E	23	36	1	2
K219Q	6	9	7	20
L74I	0	0	5	14
	NNRTI			
Y181C	18	28	1	2
V108I	13	20	1	2
G190A	2	3	0	0
M230I	1	2	0	0
A98G	0	0	2	5
	PI			
M46I	0	0	7	20
I54V	0	0	7	20
V82T	0	0	7	20
I84V	0	0	7	20
L90M	0	0	7	20
L33F	0	0	5	14
G73T	0	0	5	14
G73S	0	0	2	5

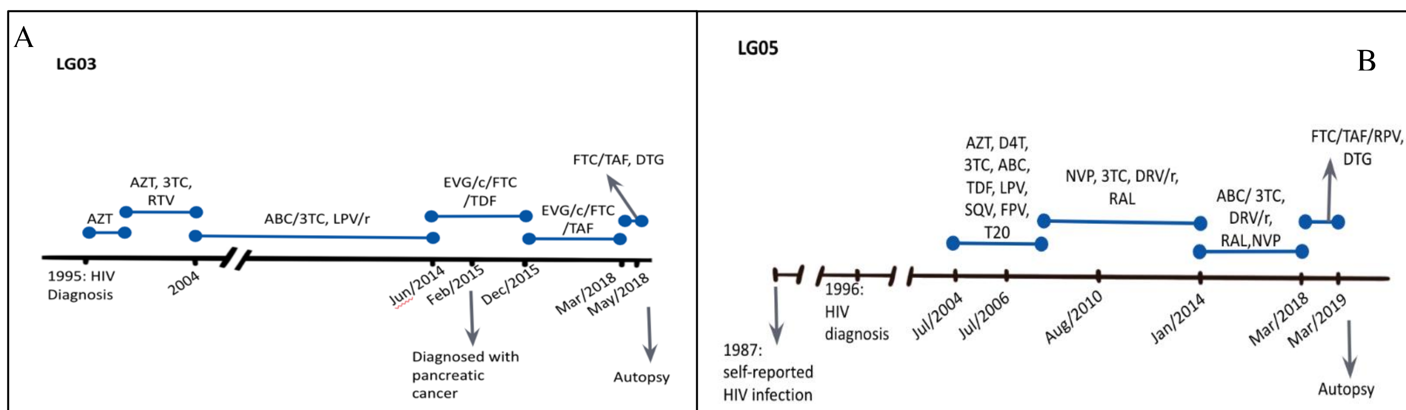


Figure 5. Clinical histories of LG03 and LG05. The clinical histories of LG03 and LG05 are shown along with ART regimens they have taken post-HIV diagnosis. (A) LG03 has taken numerous HIV drugs, many of which have been taken more than once. Drugs that were taken included NRTIs, PIs, and integrase inhibitors. (B) LG05 has taken some of the same drugs as LG03, while some differed. Drugs taken included NRTIs, PIs, NNRTIs, integrase inhibitors, and fusion inhibitors. Figure redrawn from Chaillon, A. Gianella, S., et al. HIV persists throughout deep tissues with repopulation from multiple anatomical sources. *J. Clin. Invest.* 130, 1699-1712 (2020).

DRMs in NRTI drugs

NRTIs are commonly prescribed drugs for HIV-infected patients due to the large variety of options and convenient access [1]. However, because NRTIs are regularly used for ART for HIV-infected individuals, there can be a high probability of generating DRMs and HIV minority variants if ART is interrupted. As seen in Table 3, the presence of DRMs to NRTIs was found throughout multiple tissues for both LG03 and LG05. Within these NRTI DRMs, the M41L mutation was at the highest frequency for LG03 (50%) [Table 5]. This signifies that approximately half of all sequences found in tissues of LG03 contained the M41L mutation. Although we produced less FL *pol* sequences for LG05 and this participant had generally lower NRTI mutation frequencies than LG03 (which could be due to fewer tissues and sequences obtained), the percentage of variants with NRTI mutations was still relatively high (26% at the highest). Moreover, both LG03 and LG05 had fractions of sequences (48% and 26% respectively) with the M184V mutation, which is one of the most common HIV mutations to NRTIs that is mainly selected by the drugs Lamivudine (3TC) and Emtricitabine (FTC) [21].

Referencing the clinical histories of LG03 and LG05, both patients have received these two drugs, which can give rise to the M184V mutations [Fig. 5A, B, 29]. M41L, on the other hand, is selected by several NRTI drugs such as Abacavir (ABC) and Tenofovir disoproxil (TDF) [21]; both participants have also taken these two drugs [Fig. 5A, B, 29]. D67N is another NRTI mutation that was more represented in sequences produced from LG03 (48%) and LG05 (26%) than other mutations. D67N, K219Q, and K219E (K219Q/E) are primarily selected by Stavudine (D4T) and Zidovudine (ZDV), with D67N creating a significant reduction in susceptibility to those drugs [21]. It should be noted that ZDV is another name for the drug AZT. Since LG03 took AZT early after HIV diagnosis, any D67N and K219Q/E mutations found in tissues could only have surfaced during this time. Similarly, LG05 has also only taken D4T and AZT early on, which is why we observed the D67N mutation along with K219Q/E [Fig. 5B, 29]. We did detect the presence of D67N with the absence of K219Q/E in some FL *pol* sequences. Interestingly, from the mutations that were present in LG03 and LG05, M41L is a type I thymidine analog mutation (TAM), while most were type II TAMs (D67N, K70R, K219Q/E, and T215F) [Table 4, 22]. Furthermore, M184V was observed to be always present with type II TAMs when referring to the individual HIV sequences in each tissue. Also, the presence of K219Q/E mutations reduce drug susceptibility to D4T and AZT when present with other TAMs [21].

DRMs in NNRTI drugs

Based on the oral history of ART regimens, LG03 did not take any NNRTI drugs throughout the span of HIV-infection while LG05 had taken two NNRTIs: Nevirapine (NVP) and Rilpivirine (RPV) [Fig. 5A, B, 29]. We did find one DRM for Y181C and one for V108I in LG05 likely due to NVP [21]. There were also two instances of the A98G minor accessory mutation in LG05 [21]. In contrast, LG03 had no history of NNRTI use, but has shown many

mutations that correlate with NNRTI drug resistance: Y181C, V108I, G190A, and M230I; there were no mutations found for A98G [Table 4, Fig. 5A, 29]. NVP selects for all the previously named mutations with Y181C, G190A, and M230I causing the highest reduced drug susceptibility while V108I and A98G lead to low reduced susceptibility [21]. It should be noted that M230I is a particularly rare mutation, even if it is selected by RPV and induces highly reduced treatment susceptibility [23]. LG05 has only received RPV treatment once and for a relatively short period of time (~1 year) [29], hence, it cannot be ruled out that the circumstances were not adequate for acquiring M230I mutations.

DRMs in PI drugs

As for the PI class of drugs used to treat HIV, LG03 was recorded to have none of the DRM mutations that were present in LG05 [Table 4]. LG05, conversely, had variants with evenly distributed frequencies for several PI DRMs: M46I, I54V, V82T, I84V, and L90M. LG03's clinical history showed that two PI drugs were taken whereas LG05 has taken 4, with Darunavir/ ritonavir (DRV/r) taken over a span of ~10 years and then interrupted [Fig. 5A, B, 29]. Thus, it is expected that LG05 was at risk for DRMs that affect susceptibility to PIs. LG03 had taken Ritonavir (RTV) and Lopinavir/ritonavir (LPV/r) [29], but the risk for DRMs in HIV for this participant may be lower since LG05 had more PI interruptions. From 2004 to 2006, it was indicated that LG05 took multiple drugs, some included significant drug resistance to 3 of the 4 total PI drugs taken: LPV, Saquinavir (SQV), and Fosamprenavir (FPV) [29]. Since the participant consumed multiple PIs early in the infection, this can give rise to PI DRMs that might make treatment less effective.

Largely, the clinical histories of LG03 and LG05 corresponded well with the presence of certain DRMs. Although the data is not shown, there were no DRMs found for integrase

inhibitors for both participants even though both have taken this class of ART. Certain drugs that selected for specific mutations were generally exhibited in the participants that have consumed them. This signifies that the drugs do, in fact, select for mutations that can cause varying degrees of reduced ART susceptibility and that these mutations do not only show up in the blood, but also in human anatomical sites.

Table 6. Comparison of the M184V mutation in LG03 tissues.

Tissue	*PBMC*	*Kidney*	*Duodenum*	*LN*
PBMC	x	0.0455	>0.9999	0.4633
Kidney	**0.0455**	x	**0.0009**	0.1181
Duodenum	>0.9999	**0.0009**	x	**0.01818**
LN	0.4633	0.1181	**0.01818**	x

Table 7. Comparison of the M184V mutation in LG05 tissues.

Tissue	*PBMC*	*Spleen*	*Duodenum*	*Jejunum*	*LNX*
PBMC	x	>0.9999	0.463	0.1818	>0.9999
Spleen	>0.9999	x	0.2424	**0.0275**	>0.9999
Duodenum	0.463	0.2424	x	0.4	0.5455
Jejunum	0.1818	0.0275	0.4	x	0.1099
LNX	>0.9999	>0.9999	0.5455	0.1099	x

Comparison of DRMs in tissues

For the two participants, we wanted to compare the presence of specific mutations across different anatomical tissues. The Fisher's exact test was used to determine the significance of differences of a mutation between tissues. The Fisher's exact test compared the frequency of a

sequence with DRMs to the frequency of a sequence with no DRMs within a given tissue to determine differences. As displayed in Table 6 for LG03, there were significant differences in the frequencies of the M184V mutation in the PBMC and kidney, in the kidney and the duodenum, and in the duodenum and lymph nodes. Contrastingly, LG05 had only a significant difference between the spleen and the jejunum [Table 7]. Since M184V is one of the most common DRMs observed in people living with HIV, this is a suitable start for identifying tissues that may have significantly different frequencies of certain mutations for each participant.

Acknowledgements: The figures in the results sections are currently unpublished but are coauthored with Ita, Sergio; Chaillon, Antoine; Kamalu, Maris; Baham, Jonny; Porrachia, Magali; Ignacio, Caroline; Rawlings, Stephen; Wagner, Gabriel; Gianella, Sara; Smith, Davey. The thesis author was a coauthor of these figures.

DISCUSSION

For our study, we utilized tissue samples from multiple anatomical parts of the body from HIV- infected participants enrolled in the Last Gift cohort. As discussed, we investigated two participants, LG03 and LG05, who did not stopped ART from the time they have started taking therapy until their passing. The Last Gift study began ante-mortem, in which we collected blood samples, to post-mortem, where their entire bodies were provided for tissue collection. Since prior to this study, many other investigations for HIV were limited to sampling blood only [24-26]. Although some studies examined HIV population in tissues, it was often limited to comparison of blood with one type of tissue that was usually more easily accessible [27, 28]. Therefore, the access to numerous fresh human tissues from HIV- infected individuals who donate their bodies after death can be used to determine and characterize HIV reservoirs in numerous anatomical body tissues. It was valuable that, for the Last Gift study, we had the opportunity to acquire a variety of deep tissues to examine the dynamics of HIV spread. We were able to employ advanced molecular biology techniques and single genome amplification to generate and analyze HIV data. The main findings of our research include: 1) nearly identical sequences were found within tissues and across multiple different tissues; 2) LG03 and LG05 had high frequencies of FL *pol* sequences with DRMs in the lymphoid tissues of the spleen and lymph nodes.

LG03 was shown to contain many HIV-1 FL *pol* sequences within tissues such as the kidney, testis, prostate, spleen, and lymph nodes. The phylogeny depicted that the kidney, testis, right colon, LNA, and spleen had 13 nearly identical (99% identical) sequences [Fig. 4A]. This suggests that there may be migration of HIV variants from one location in the body to another or clonal expansion. In LG05, there were multiple nearly identical sequences found in gut tissues

such as the duodenum, ileum, and jejunum [Fig. 4B]. There was also a sequence from the spleen that was observed to be nearly identical to the gut tissues, also suggesting that the spread of minority variants can occur. In a similar study from our group, which investigated the full-length HIV *env* gene, there was evidence of migration of variants in virally suppressed participants from the Last Gift cohort (the same participants were used, including LG03 and LG05) [29]. It was discovered that most of the dispersion that occurred originated from the lymph nodes and gut tissues [29]. In addition, the researchers found that the testis and prostate can also be possible centers of viral distribution throughout the body, suggesting that genital tract tissues can be key reservoirs of HIV along with lymph nodes and gut tissues as seen in other studies [20, 29-34]. Corresponding to this, we did observe higher quantities of sequences in the testes, including some that were nearly identical to other tissues. So, if the spread of drug-resistant sequences in LG03 and LG05 was observed using the FL HIV *env* gene, there is a possibility that this can also be true using the FL HIV *pol* gene taken from various tissues.

Also, the lymphoid tissues, such the lymph nodes and spleen, exhibited higher frequencies and quantities of DRMs to multiple classes of ARV drugs [Table 3]. These tissues have shown relatively high numbers of sequences with DRMs against mainly NRTI and NNRTI drugs in LG03 and to all three drug classes in LG05. The results we obtained from the lymph nodes seem to differ from the previous Feder et al. study where they found that the lymph nodes had less DRMs [14]. We speculate that this observation is still likely due to the varying degrees of drug penetrance in tissues. It was discovered that the lymph nodes have relatively low drug penetrance [15, 16]. This low selective pressure may have led to high replication rates or a build-up of HIV in the lymph nodes, where there is also a greater immune pressure due to the high

levels of CD4+ T cells [35]. Therefore, the high quantities of DRMs found in the lymphoid tissues may be due to the weakened ability of ART to effectively eliminate HIV.

Our research showed that the PBMC samples we collected in both participants did not contain any sequences with NNRTI DRMs. This is interesting because evidences of sequences with NNRTI DRMs were seen in several other deep tissues. This may be explained by the higher efficacy of ART in blood due to high drug penetrance, while NNRTIs are unable to work as well in deeper tissues [36]. Hence, it is possible that ART has reduced and controlled HIV levels in the blood, making any DRMs, if any, undetectable. It is also possible that we did not detect DRMs in sequences from PBMCs due to the limited number of sequences generated that reflected low viral levels in the blood.

Looking at the clinical history of LG03, this participant took Zidovudine (AZT), a NRTI drug early on after HIV diagnosis [Fig. 5A, 29]. Since AZT selects for K219Q/E and D67N mutations, we did find presence of these in various tissues of LG03. However, it is interesting that even though LG03 stopped taking AZT around 2004 and did not taken this drug again, we still observed these mutations in tissues taken from autopsy. Likewise, LG05 only took D4T (also selects for the D67N and K219Q/E mutations) and AZT from 2004 to 2006 towards the beginning of the infection, but had detectable D67N and K219Q/E mutations in various tissues after death [Fig. 5B, 29]. These findings suggest that D67N and K219Q/E can possibly be archived in HIV reservoirs of an individual for months or even years. This raises the possibility that other mutations selected by certain ART regimens also have the ability to remain in the body over extended periods of time. Hence, in theory, if a patient has similar treatment in the future, archived DRMs could potentially lead to treatment failures. On another note, LG03 was observed to have multiple NNRTI mutations although no NNRTI drugs were recorded in the

clinical history [Table 4, 29]. Since all participants of the Last Gift cohort were questioned about their clinical histories and ART regimens [29], the oral history of LG03 might not be accurate. DRMs are usually selected for by drugs in that same class and since the number of variants with NNRTI DRMs were sufficiently present in LG03, it would be expected that these mutations originated from a history of NNRTI drug use.

There were some significant differences in the M184V mutation frequencies for different tissues [Table 6, 7]. This indicates that a specific mutation can vary in frequency across diverse anatomical locations of the body. Although we cannot yet say that those particular tissues exhibit compartmentalization, a related study found evidence of compartmentalization in the same participants, LG03 and LG05, in which the FL *env* gene was used to examine HIV reservoirs [29]. Therefore, if compartmentalization was observed in the two participants using a related viral gene, this highlights the possibility of compartmentalization in LG03 and LG05 using the FL *pol* gene.

There were a few limitations that were identified throughout the study. Firstly, as seen in the SHIV study by Feder et al., the researchers discovered vast differences in viral DNA and viral RNA [14]. However, for this Last Gift study, only viral DNA was investigated in terms of HIV reservoirs and DRMs. Hence, there may be certain characteristics and HIV dynamics of drug resistance in the sequences that we have missed in viral RNA. Additionally, this means that we were unable to compare the viral DNA to the viral RNA in tissues. Nonetheless, the investigation of HIV DNA has provided adequate information on HIV DRMs in human tissues. Secondly, since SGA, and not bulk deep sequencing, was used to deep sequence single templates of HIV DNA, some results may not be representative of the whole population. SGA requires limiting dilution of extracted DNA in order to obtain amplification of a single genome. Bulk

deep sequencing, on the other hand, does not require dilution; therefore, the sequencing information obtained is representative of the population. So, the frequencies of DRMs that were reported were only representative of what we sequenced. It is possible that the quantities of mutations were actually lower in the population, such as in the individual or in the tissues. However, SGA of the FL *pol* gene was able to provide sufficient information on DRMs seen in each sequence (each single genome) because there is a defined genetic background. Therefore, we still acquired useful information in determining whether different classes of mutations appeared on the same background or sequence. Third, we were not able to obtain the same tissues or the same amount of tissues for each of the participants. Overall, we obtained a greater variety of tissues in LG03 than LG05. There was some overlap in the types of tissues sampled from each participant, but it would be advantageous to acquire additional tissues in LG05 so that data from same tissues in each participant could be properly compared and analyzed. Specifically, for PBMCs, the sampling was not as robust as it was for some of the other tissues. Lastly, as mentioned earlier, the clinical histories of ART regimens used to characterize DRMs found in each of the LG participants relied primarily on self-reporting. Although we concluded that most of their ART histories correlated with the DRMs that were present in their tissues, any discrepancies found may be attributed to an error in their oral histories. ART regimen history was important for predicting and explaining many of the DRMs that arose, but we were still able to well characterize mutations that existed in HIV sequences in different tissues.

The future directions that the path of this study lead us to have great potential in uncovering even more of what is already known about HIV reservoirs. Since there is some preliminary evidence of significant differences in specific mutations across tissues, then further analysis and research can be performed to determine the presence or absence of

compartmentalization in anatomical locations of the human body. Several other studies have reported compartmentalization of viral genotypes [14, 29], hence, this may be potentially true for tissues in LG03 and LG05. If future evidence of compartmentalization was reported, this could exemplify differences in viral dynamics even within human tissues of people living with HIV. Also, more research could be done to determine whether certain tissues act as centers of distribution for certain HIV genotypes to confirm that viral variants can migrate and disperse. Now that we know HIV sequences can harbor multiple classes of DRMs on the same background, other interesting issues that could be further examined would be how and what order those mutations came into place. Last but not least, we could aim for bulk deep sequencing to examine viral dynamics and DRMs as representative of populations.

In conclusion, the Last Gift study provided the remarkable opportunity to investigate HIV characteristics and DRMs using fresh tissues obtained from autopsy from people who lived with HIV. Prior to this study, much of the research performed on HIV did not have access to a variety of deep tissues. Therefore, not much was known about the viral reservoirs and characteristics in human anatomical sites. Yet, we were able to have a glimpse into the spatial dynamics of HIV, including associated DRMs, that were present in different tissues of each participant. ART has been effective in suppressing viral replication and transmission, but efforts are still needed to completely eradicate HIV reservoirs in not just the blood, but also in deep tissues. ART could be theoretically even more efficient at eliminating HIV if the rise of DRMs could be controlled and reduced greatly. With further characterization and understanding of HIV in anatomical parts of the body, we are one step closer to improving the lives of those living with HIV and finding a cure, all in hopes to alleviate this pandemic worldwide.

REFERENCES

1. Fauci, Anthony S., Kasper, D., Hauser, S., Longo, D., Jameson, L. J., Loscalzo, J. *Harrison's Principles of Internal Medicine, 19e* Eds. (McGraw-Hill, Ltd, 2014).

2. Drake, J. W. & Holland, J. J. Mutation rates among RNA viruses. *Proc. Natl. Acad. Sci.* 96, 13910 LP – 13913 (1999).

3. Duffy, S., Shackelton, L. A. & Holmes, E. C. Rates of evolutionary change in viruses: patterns and determinants. *Nat. Rev. Genet.* **9**, 267–276 (2008).

4. Onafuwa-Nuga, A. & Telesnitsky, A. The Remarkable Frequency of Human Immunodeficiency Virus Type 1 Genetic Recombination. *Microbiol. Mol. Biol. Rev.* **73**, 451–480 (2009).

5. Ramirez, B. C., Simon-Loriere, E., Galetto, R. & Negroni, M. Implications of recombination for HIV diversity. *Virus Res.* **134**, 64–73 (2008).

6. Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M. & Ho, D. D. HIV-1 Dynamics in Vivo: Virion Clearance Rate, Infected Cell Life-Span, and Viral Generation Time. *Science (80-.).* **271**, 1582 LP – 1586 (1996).

7. Kumari, G. & K. Singh, R. Anti-HIV Drug Development: Structural Features and Limitations of Present Day Drugs and Future Challenges in the Successful HIV/AIDS Treatment. *Curr. Pharm. Des.* **19**, 1767–1783 (2013).

8. Metzner, K. J., Bonhoeffer, S., Fischer, M., Karanicolas, R., Allers, K., Joos, B., Weber, R., Hirschel, B., G. Kostrikis, L. & Studya, the S. H. I. V. C. Emergence of Minor Populations of Human Immunodeficiency Virus Type 1 Carrying the M184V and L90M Mutations in Subjects Undergoing Structured Treatment Interruptions. *J. Infect. Dis.* **188**, 1433–1443 (2003).

9. Hare, C. B., Mellors, J., Krambrink, A., Su, Z., Skiest, D., Margolis, D. M., Patel, S. S., Barnas, D., Frenkel, L., Coombs, R. W., Aweeka, F., Morse, G. D., Haas, D. W., Boltz, V., Palmer, S., Coffin, J. & Havlir, D. V. Detection of Nonnucleoside Reverse-Transcriptase Inhibitor–Resistant HIV-1 after Discontinuation of Virologically Suppressive Antiretroviral Therapy. *Clin. Infect. Dis.* **47**, 421–424 (2008).

10. Martinez-Picado, J., Morales-Lopetegi, K., Wrin, T., Prado, J. G., Frost, S. D. W., Petropoulos, C. J., Clotet, B. & Ruiz, L. Selection of drug-resistant HIV-1 mutants in response to repeated structured treatment interruptions. *AIDS* **16**, (2002).

11. Little, S. J., Holte, S., Routy, J.-P., Daar, E. S., Markowitz, M., Collier, A. C., Koup, R. A., Mellors, J. W., Connick, E., Conway, B., Kilby, M., Wang, L., Whitcomb, J. M., Hellmann, N. S. & Richman, D. D. Antiretroviral-Drug Resistance among Patients Recently Infected with HIV. *N. Engl. J. Med.* **347**, 385–394 (2002).

12. Wensing, A. M. J., van de Vijver, David A., Angarano, G., Åsjö, B., Balotta, C., Boeri, E., Camacho, R., Chaix, M., Costagliola, D., De Luca, A., Derdelinckx, I., Grossman, Z., Hamouda, O., Hatzakis, A., Hemmer, R., Hoepelman, A., Horban, A., Korn, K., Kücherer, C., Leitner, T., Loveday, C., MacRae, E., Maljkovic, I., de Mendoza, C., Meyer, L., Nielsen, C., Op de Coul, E., Ormaasen, V., Paraskevis, D., Perrin, L. Prevalence of Drug-Resistant HIV-1 Variants in Untreated Individuals in Europe: Implications for Clinical Management. *J. Infect. Dis.* **192**, 958–966 (2005).

13. Paydary, K., Khaghani, P., Emamzadeh-Fard, S., Alinaghi, S. A. S. & Baesi, K. The emergence of drug resistant HIV variants and novel anti-retroviral therapy. *Asian Pac. J. Trop. Biomed.* **3**, 515–522 (2013).

14. Feder, A. F., Kline, C., Polacino, P., Cottrell, M., Kashuba, A. D. M., Keele, B. F., Hu, S. L., Petrov, D. A., Pennings, P. S. & Ambrose, Z. A spatio-temporal assessment of simian/human immunodeficiency virus (SHIV) evolution reveals a highly dynamic process within the host. *PLoS Pathog.* **13**, e1006358 (2017).

15. Solas, C., Lafeuillade, A., Halfon, P., Chadapaud, S., Hittinger, G. & Lacarelle, B. Discrepancies between protease inhibitor concentrations and viral load in reservoirs and sanctuary sites in human immunodeficiency virus-infected patients. *Antimicrob. Agents Chemother.* **47**, 238–243 (2003).

16. Fletcher, C. V., Staskus, K., Wietgrefe, S. W., Rothenberger, M., Reilly, C., Chipman, J. G., Beilman, G. J., Khoruts, A., Thorkelson, A., Schmidt, T. E., Anderson, J., Perkey, K., Stevenson, M., Perelson, A. S., Douek, D. C., Haase, A. T. & Schacker, T. W. Persistent HIV-1 replication is associated with lower antiretroviral drug concentrations in lymphatic tissues. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 2307–2312 (2014).

17. Wong, J. K. & Yukl, S. A. Tissue reservoirs of HIV. *Current Opinion in HIV and AIDS* vol. 11 362–370 (2016).

18. Yukl, S. A., Shergill, A. K., Ho, T., Killian, M., Girling, V., Epling, L., Li, P., Wong, L. K., Crouch, P., Deeks, S. G., Havlir, D. V, Mcquaid, K., Sinclair, E. & Wong, J. K. The Distribution of HIV DNA and RNA in Cell Subsets Differs in Gut and Blood of HIV-Positive Patients on ART: Implications for Viral Persistence. doi:10.1093/infdis/jit308.

19. van Marle, G., Gill, M. J., Kolodka, D., McManus, L., Grant, T. & Church, D. L. Compartmentalization of the gut viral reservoir in HIV-1 infected patients. *Retrovirology* **4**, 87 (2007).

20. Yukl, S. & Wong, J. K. Blood and Guts and HIV: Preferential HIV Persistence in GI Mucosa. *J. Infect. Dis.* **197**, 640–642 (2008).

21. Rhee, S. Y., Gonzales, M. J., Kantor, R., Betts, B. J., Ravela, J. & Shafer, R. W. Human immunodeficiency virus reverse transcriptase and protease sequence database. *Nucleic Acids Research* vol. 31 298–303 (2003).

22. De Luca, A., Di Giambenedetto, S., Romano, L., Gonnelli, A., Corsi, P., Baldari, M., Di Pietro, M., Menzo, S., Francisci, P., Almi, P. & Zazzi, M. Frequency and Treatment-Related Predictors of Thymidine-Analogue Mutation Patterns in HIV-1 Isolates after Unsuccessful Antiretroviral Therapy. *J. Infect. Dis.* **193**, 1219–1222 (2006).

23. Azijn, H., Tirry, I., Vingerhoets, J., De Béthune, M. P., Kraus, G., Boven, K., Jochmans, D., Van Craenenbroeck, E., Picchio, G. & Rimsky, L. T. TMC278, a next-generation nonnucleoside reverse transcriptase inhibitor (NNRTI), active against wild-type and NNRTI-resistant HIV-1. *Antimicrob. Agents Chemother.* **54**, 718–727 (2010).

24. Fulcher, J. A., Hwangbo, Y., Zioni, R., Nickle, D., Lin, X., Heath, L., Mullins, J. I., Corey, L. & Zhu, T. Compartmentalization of Human Immunodeficiency Virus Type 1 between Blood Monocytes and CD4+ T Cells during Infection. *J. Virol.* **78**, 7883–7893 (2004).

25. Potter, S. J., Lemey, P., Dyer, W. B., Sullivan, J. S., Chew, C. B., Vandamme, A. M., Dwyer, D. E. & Saksena, N. K. Genetic analyses reveal structured HIV-1 populations in serially sampled T lymphocytes of patients receiving HAART. *Virology* **348**, 35–46 (2006).

26. Delobel, P., Sandres-Sauné, K., Cazabat, M., L'Faqihi, F. E., Aquilina, C., Obadia, M., Pasquier, C., Marchou, B., Massip, P. & Izopet, J. Persistence of distinct HIV-1 populations in blood monocytes and naive and memory CD4 T cells during prolonged suppressive HAART. *AIDS* **19**, 1739–1750 (2005).

27. Imamichi, H., Degray, G., Dewar, R. L., Mannon, P., Yao, M., Chairez, C., Sereti, I. & Kovacs, J. A. Lack of Compartmentalization of HIV-1 Quasispecies Between the Gut and Peripheral Blood Compartments. doi:10.1093/infdis/jir259.

28. Kearney, M. F., Anderson, E. M., Coomer, C., Smith, L., Shao, W., Johnson, N., Kline, C., Spindler, J., Mellors, J. W., Coffin, J. M. & Ambrose, Z. 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**, (2015).

29. Chaillon, A., Gianella, S., Dellicour, S., Rawlings, S. A., Schlub, T. E., De Oliveira, M. F., Ignacio, C., Porrachia, M., Vrancken, B. & Smith, D. M. HIV persists throughout deep tissues with repopulation from multiple anatomical sources. *J. Clin. Invest.* **130**, 1699–1712 (2020).

30. Telwatte, S., Lee, S., Somsouk, M., Hatano, H., Baker, C., Kaiser, P., Kim, P., Chen, T. H., Milush, J., Hunt, P. W., Deeks, S. G., Wong, J. K. & Yukl, S. A. Gut and blood differ

in constitutive blocks to HIV transcription, suggesting tissue-specific differences in the mechanisms that govern HIV latency. *PLoS Pathog.* **14**, e1007357 (2018).

31. Embretson, J., Zupancic, M., Ribas, J. L., Burke, A., Racz, P., Tenner-Racz, K. & Haase, A. T. Massive covert infection of helper T lymphocytes and macrophages by HIV during the incubation period of AIDS. *Nature* **362**, 359–362 (1993).

32. Pantaleo, G., Graziosi, C., Demarest, J. F., Butini, L., Montroni, M., Fox, C. H., Orenstein, J. M., Kotler, D. P. & Fauci, A. S. HIV infection is active and progressive in lymphoid tissue during the clinically latent stage of disease. *Nature* **362**, 355–358 (1993).

33. Wong, J. K., Günthard, H. F., Havlir, D. V., Zhang, Z. Q., Haase, A. T., Ignacio, C. C., Kwok, S., Emini, E. & Richman, D. D. Reduction of HIV-1 in blood and lymph nodes following potent antiretroviral therapy and the virologic correlates of treatment failure. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12574–12579 (1997).

34. Günthard, H. F., Havlir, Diane V., Fiscus, S., Zhang, Z., Eron, J., Mellors, J., Gulick, R., Frost, S. D. W., Brown, A. J. L., Schleif, W., Valentine, F., Jonas, L., Meibohm, A., Ignacio, C. C., Isaacs, R., Gamagami, R., Emini, E., Haase, A., Richman, D. D., Wong, J. K. Residual Human Immunodeficiency Virus (HIV) Type 1 RNA and DNA in Lymph Nodes and HIV RNA in Genital Secretions and in Cerebrospinal Fluid after Suppression of Viremia for 2 Years. *J. Infect. Dis.* **183**, 1318–1327 (2001).

35. Vanderford, T. H., Bleckwehl, C., Engram, J. C., Dunham, R. M., Klatt, N. R., Feinberg, M. B., Garber, D. A., Betts, M. R. & Silvestri, G. Viral CTL escape mutants are generated in lymph nodes and subsequently become fixed in plasma and rectal mucosa during acute SIV infection of macaques. *PLoS Pathog.* **7**, e1002048 (2011).

36. Cohen, M. S., Smith, M. K., Muessig, K. E., Hallett, T. B., Powers, K. A. & Kashuba, A. D. Antiretroviral treatment of HIV-1 prevents transmission of HIV-1: Where do we go from here? *The Lancet* vol. 382 1515–1524 (2013).