TISSUE RESERVOIRS OF HIV-1: INSIGHTS FROM THE CENTRAL NERVOUS SYSTEM

Lauren A. Rackoff Tompkins

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Science in the Microbiology and Immunology Department in the School of Medicine.

Chapel Hill
2015

Approved by:

Ronald Swanstrom

Mark Heise

Nathaniel Moorman

Glenn Matsushima

ABSTRACT

Lauren A. Rackoff Tompkins: Tissue Reservoirs of HIV-1: Insights from the Central Nervous System
(Under the direction of Ronald Swanstrom)

Cellular and anatomical reservoirs of HIV-1 preclude a cure to infection. Efforts to characterize these reservoirs are an important part of developing a strategy to eradicate all forms of HIV-1. The central nervous system (CNS) is a unique bodily compartment that can support viral replication independent of that in the blood (compartmentalization) and may be an anatomical reservoir of unique viruses. In this 'proof of principle' study, we characterized viral sequences (RNA and provirus) from the blood, cerebrospinal fluid, brain, and liver of two infected donors who died with HIV-1-associated dementia and disparate states of viral replication (compartmentalized versus equilibrated). We show that selective pressures exist within CNS and liver tissue to drive expansion of particular viral species. We found macrophage-tropic viral lineages archived in the brain, implicating macrophages as a potential cellular reservoir. By including more donors and tissues, our study provides insight towards HIV-1 reservoirs and cure research.

ACKNOWLEDGEMENTS

I would like to thank my mentor, Ron, for his valuable guidance, unwavering support, and friendship. He stuck by me through thick-and-thin, was always inspiring, and I am a better scientist for knowing him. The Swanstrom Lab is not only my scientific home base, but also a place where I've met lifelong friends who I cherish greatly. We share laughs and beers, prank each other mercilessly, and always seem to make light of even the toughest parts of being a scientist. To my closest friends in the lab, you know who you are, thank you from the bottom of my heart for being the incredible friends you are. You made work fun, which is such a gift, especially when you're in the throws of grad school. More so, you've been nothing but supportive, encouraging, and wonderful people - I am truly blessed for having known you.

To my committee, I cannot express how grateful I am to have had you with me throughout this journey. Your guidance has been invaluable to me and has shaped my journey as a scientist. I've learned so much from each of you and could always come to you with questions or to ask for advice. You've been my cheerleaders when I needed them most and never lost faith in me. I couldn't have done it without you.

I would also like to thank UNC and all the people who have helped me along the way, especially Dixie Flannery, Bob Bourret, and Bill Goldman of the Microbiology & Immunology Department. Thank you for all of your help in navigating grad school and taking a personal interest in my graduate career as well as those of each and every one of my classmates. I am overcome with gratitude for the immense support I've received from every person I've met at UNC – this is truly a school that cares about its students.

My greatest appreciation is reserved for my family and friends: thank you for being in my life. You are the people who have always been there and will always be there, whose love in unconditional and so essential. My parents are truly great people; their selflessness and generosity inspires me every day. To Matt, my husband and soul mate: you are my rock and my world, I love you more than anything. You make everything in this world shine a little brighter.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I. ASSESSING TISSUE RESERVIORS OF HIV-1 IN DONORS WHO DIED WITH HIV-ASSOCIATED DEMENTIA

INTRODUCTION

1.1 HIV-1 Disease Burden And Molecular Biology

1.1.1 The HIV-1 pandemic: treatable, but not curable

Human immunodeficiency virus type 1 (HIV-1) is the cause of acquired immunodeficiency syndrome (AIDS), where uncontrolled infection in the blood and lymphoid organs depletes CD4+ T cells and thus diminishes immune competence over time. HIV-1 primarily infects CD4+ T cells and productive infection usually results in cell death via toxic effects of viral replication or immune attack. When a sufficient number of CD4+ T cells are lost, HIV-infected people become susceptible to opportunistic infections, cancers, and other co-morbidities. Without treatment, HIV-1 disease progresses to AIDS in about 10-15 years, and most people die within a few years of AIDS diagnosis. Roughly 37 million people are currently living with HIV-1, with an estimated 5,600 new infections occurring per day. The disease burden is greatest in low- and middle-income countries: 70% of people with HIV-1 live in Africa and 65% of AIDS-related deaths also occur in Africa [1].

The introduction of antiretroviral therapy (ART) has drastically changed the landscape of HIV-1 morbidity and mortality. ART controls viral replication and thus disease progression, thereby extending longevity in HIV-infected people. At the end of 2013, 12.9 million people globally were receiving ART, 11.7 million of them in low- and middle-income countries [1, 2]. An additional 1.9 million people were newly enrolled in ART in 2014, which is one of the largest annual increases in treatment initiation to date. Still, only 14.8 of the 37 million HIV-1+ people are receiving ART, which means that the majority of people are untreated and capable of transmitting the virus (60% or 22.2 of 37 million) [1, 2]. Efforts are underway to continually increase the number of people on therapy, and the World Health Organization is now

recommending a “treat-all” approach where all populations and age groups are recommended to receive ART as soon after diagnosis as possible [1].

HIV-1 infection is indeed treatable, but there is currently no cure. Although ART is successful at achieving viral suppression (i.e. very low or undetectable viral load), these drugs must be taken throughout life to control infection. A major focus of current research is to eradicate HIV-1 and cure infected patients. Viral persistence in cellular and anatomical reservoirs precludes a cure, thus efforts to characterize these reservoirs are an important part of developing a strategy for eradicating all forms of HIV-1.

1.1.2 HIV-1 particle structure and genome organization

HIV-1 is a single-stranded RNA (ssRNA), enveloped virus that belongs to the *Lentivirus* genus, *Retroviridae* family. The HIV-1 genome is ~9.7 kilobases in length, has a 5’ cap and 3’ polyA tail, and encodes for nine genes that are flanked by long terminal repeats (LTR): *gag*, *pol*, *env*, *vif*, *vpr*, *tat*, *rev*, *vpu*, and *nef*. Structural proteins are encoded in the *env* gene (viral envelope, “Env protein”) and the *gag* gene (the Gag polyprotein precursor, which is processed into the matrix [MA], capsid [CA], and nucleocapsid [NC] structural proteins). The *pro* and *pol* genes encodes the viral enzymes protease (Pro), integrase (IN), and reverse transcriptase (RT). The remainder of genes encode accessory proteins (Vif, Vpr, Vpu, and Nef) or regulatory proteins (Tat and Rev) [3].

Two identical copies of the ssRNA viral genome are packaged within a conical capsid core. The dimeric ssRNA genome associates with NC protein to form a NC-RNA complex that is surrounded by CA protein to form the conical capsid core [4]. A spherical shell composed of MA protein surrounds the core. MA protein is embedded in the viral envelope, a lipid bilayer derived from the infected host cell membrane. Trimeric Env proteins are also embedded in the viral envelope and are exposed as spikes on the outer surface of the virion. A mature virion contains all of the components required for infectivity: the dimeric ssRNA genome, cellular tRNALys3 molecules to prime cDNA synthesis, the viral structural proteins that form a functional virion and permit attachment to viral receptors expressed on permissive cells, and the three viral enzymes. Vpr, Vif, and Nef are also packaged in the virion, as well as cellular proteins incorporated during budding from an infected cell.

1.1.3 The viral life cycle of HIV-1

The first step of the HIV-1 replication cycle is attachment to the surface of a permissive cell via interactions between HIV-1 Env protein and the cell-surface viral receptor, CD4, and co-receptors, CCR5 or CXCR4 [5]. Viral Env is a heavily glycosylated trimeric protein composed of the gp120 surface glycoprotein, which is exposed on the outer face of the viral particle, and the gp41 transmembrane glycoprotein, which is embedded in the viral membrane (viral envelope). These proteins remain tethered to one another noncovalently in the viral envelope. Both gp120 and gp41 originate from the same precursor protein, gp160, which is proteolytically cleaved by host furin-like proteases in an infected cell.

Cell attachment begins with gp120 binding CD4 on the cell surface. This interaction initiates a cascade of conformational changes in viral Env gp120 and gp41 eventually leading to fusion with the cellular membrane [5]. First, gp120-CD4 association induces conformational changes in gp120 that allow interaction with the viral co-receptor, CCR5 or CXCR4. Co-receptor binding causes exposure of the hydrophobic gp41 fusion peptide, which inserts into the host cell membrane thereby tethering the viral and cell membranes. Each gp41 monomer within the trimer bends at a hinge region to form a stable six-helix bundle, which is the driving force in forming a fusion pore. Contents of the viral particle are then delivered into the cell cytoplasm through the fusion pore.

Upon release of the viral core into the cell cytoplasm, the viral capsid is uncoated and viral genomic RNA and proteins are released [5]. In the cytoplasm, viral RNA is reverse transcribed to DNA by viral RT. Viral RT has an RNA-dependent DNA polymerase for synthesizing a DNA copy of the viral genome and RNase H activity for degrading RNA in the RNA:DNA replication intermediate. The pre-integration complex (PIC), composed of viral genomic DNA and viral and cellular proteins, is then formed and translocates into the cell nucleus through the intact nuclear envelope, allowing HIV-1 to replicate in non-dividing cells [6]. With the help of viral IN, viral DNA is integrated into a host cell chromosome (provirus), which is an essential step of the viral life cycle [7]. Integration can happen anywhere in the host genome, but tends to favor transcriptionally active genes. A recent study has highlighted the importance of the nuclear architecture in integration site selection, as integration occurs near the nuclear pore and favors active transcriptional units as opposed to chromatin regions located deeper within the

nucleus [8]. Viral integration allows the virus to persist in the form of an infected cell that can become a reservoir when dormant (latent infection).

The HIV-1 provirus can then be transcribed using cellular machinery, namely RNA polymerase II, with the help of the viral Tat protein [9]. Viral transcripts, spliced or unspliced, are then transported out of the nucleus, a process mediated in part by the viral Rev protein, for translation of viral proteins or packaging into a nascent virion (viral assembly) in the cell cytoplasm [10]. Viral *gag*, *pro*, *pol*, and *env* are translated into polyprotein precursors, which are then further processed/cleaved by the viral Pro or a host protease during and after viral assembly and budding [11]. Viral assembly occurs at the cell surface, where two copies of an unspliced, full-length viral RNA genome, viral proteins, and components of the cell cytosol are packaged into an immature virion that buds from the cell surface. In a highly ordered process of maturation, viral Pro cleaves the structural viral polyprotein precursors (Gag and Gag-Pro-Pol), ultimately leading to the formation of a mature, infectious viral particle.

1.2 Tissue Reservoirs Of HIV-1

1.2.1 HIV-1 persistence in viral reservoirs

HIV-1 persistence in cellular and anatomical reservoirs precludes a cure to infection. Two essential criteria exist to define a viral reservoir of HIV-1 [12]. First, a reservoir must preserve replication-competent virus in some form (i.e. viral particles or viral genomes) so that the virus can reestablish productive infection in the future. Second, a reservoir must have mechanisms of longevity. For example, reservoirs composed of virions would require escape from biochemical decay, as seen with virion particles that become trapped extracellularly in dendritic cell processes. Cell-associated viral reservoirs, however, require cell survival and escape from immune control including cytotoxic T cell (CTL) activity. Latent infection, a state with no active replication, is the best characterized cellular reservoir of HIV-1, but another type of reservoir could be composed of productively infected cells with slow turnover. The concept of a reservoir is further complicated by the detection of cells that clonally expand in vivo through transactivation of cellular growth-promoting genes by integrated viral DNA [13].

A reservoir will most likely occur in cells that are normally infected with HIV-1. This virus infects cells that express the viral receptor CD4 and co-receptor (CCR5 or CXCR4). Activated CD4+ T cells are the most permissive cell type for HIV-1. Latently infected resting memory CD4+ T cells are the hallmark reservoir of HIV-1 infection and are thought to predominantly arise from productive infection of activated CD4+ T cells as they are transitioning back into a resting state [14]. Latently infected T cells contain replication-competent HIV-1 genomic DNA (provirus) integrated within the human genome in the absence of ongoing viral replication. Although dormant, latently infected resting T cells can be induced to become activated and thereby transcribe integrated DNA to generate new progeny virions capable of productive infection. It is important to note that integrated DNA can be either intact or defective, thus many infected T cells contain HIV-1 DNA that is incapable of producing functional virions [15]. Latently infected cells are established early after a person becomes infected and these cells persist even in the presence of ART-mediated viral suppression. Persistence of the latent reservoir is due in part to the virus escaping immune surveillance, as integrated viral DNA in latently infected cells is likely transcriptionally silent and does not produce antigens that signal immune attack [16].

A second type of HIV-1 reservoir could be a productively infected cell with slow turnover, which is illustrated in HIV-1 infection of the central nervous system (CNS). The CNS is a bodily compartment separated from the periphery by the blood-brain barrier. HIV-1 enters the CNS and can establish productive infection with features that are distinct from those in the blood [17-22]. One such feature involves comparing disparate viral decay kinetics in the blood and cerebrospinal fluid (CSF), the latter of which bathes the CNS and is an indirect measure of CNS infection. After the onset of suppressive ART, which prevents new infections without affecting previously infected cells, the blood viral load decreases rapidly (1-2 weeks) due to rapid turnover of short-lived infected T cells [23-25]. For most people, the same pattern of rapid viral decay is observed in the CSF. However, occasionally the viral load decays much more slowly in the CSF than in the blood, which suggests that cells with a longer half-life than T cells can support HIV-1 replication in the CNS [20, 26].

The concept of anatomical reservoirs is a current topic of interest, and the CNS is a unique anatomical compartment capable of sustaining viral replication and potentially harboring viral reservoirs that differ from those in the blood. Tissues exposed to the peripheral blood may also harbor unique viral

reservoirs, as tissue microenvironments are highly structured and functionally specialized. Importantly, only 1-2% of lymphocytes in the body are present in the blood at any given time [27, 28], whereas the vast majority of lymphocytes are tissue-resident, indicating that a greater frequency of permissive cells are present in the tissues compared to the blood [29]. Consistent with the idea of tissue reservoirs differing from the blood, reservoir establishment in at least some tissues appears to precede that in the blood, as demonstrated by a recent rhesus macaque animal model of HIV-1 infection [30]. In macaques treated with suppressive ART at day three post-infection with simian immunodeficiency virus (SIV), viral DNA was found in lymph node and gut tissue before detectable viremia and in the absence of SIV DNA in PBMCs. Importantly, treatment interruption was associated with rebound virus, which indicates that the reservoir was indeed established in these macaques. Understanding how tissues act as anatomical compartments for viral replication and reservoir establishment will be important for disease treatment methods and advancement towards finding a cure for HIV-1.

1.2.2 The CNS is a unique bodily compartment and a potential anatomical reservoir of HIV-1

HIV-1 is detectable in the CSF early in acute infection, indicating that virus enters the CNS and may replicate there [22]. HIV-1 infection of the CNS can cause neurocognitive dysfunction ranging in severity from mild forms of impairment to full-blown dementia (HIV-associated dementia, HAD), diseases collectively termed HIV Associated Neurocognitive Dysfunction (HAND). Incidence of HAD in HIV-infected people has decreased dramatically with the use of antiretroviral therapy (ART), yet milder forms of HAND have increased in prevalence in people undergoing ART [31]. Up to 50% of patients being treated for HIV-1 have some neurocognitive impairment, and the mechanisms of mild HAND pathogenesis in ART-mediated, virally suppressed people are unknown. Studies have highlighted the likely contribution of chronic low-grade CNS inflammation to neurocognitive disease [32-35]. Indeed, people with ART-suppressed plasma viremia can have higher viral load in the CSF, which correlates with increased CNS inflammation and immune activation. Whether persistent CNS viral replication or release of virus from cellular reservoirs is responsible for chronic inflammation is yet to be determined. Understanding the complexity of CNS infection and mechanisms of pathology requires knowledge of CNS anatomy and physiology.

The CNS, consisting of the brain and spinal cord, is an anatomical compartment isolated from the rest of the body by the blood-brain barrier (BBB) and blood-cerebrospinal fluid barrier (BCSFB) [36]. The BBB and BCSFB are semipermeable barriers that limit exchange of substances between the CNS and peripheral blood. These physiologically distinct barriers influence the composition of CNS fluid, including the CSF and CNS interstitial fluid, which differs greatly from the blood. For example, concentrations of white blood cells, albumin, and immunoglobulin G (IgG) in the CSF are all <1% of that in the blood in spite of the fact that the water portion of the CSF is derived from the blood plasma. Key to barrier function of the BBB and BCSFB are intercellular tight junctions, which connect cerebrovascular endothelial cells (BBB) or choroid plexus epithelial cells (BCSFB). Otherwise, the BBB and BCSFB differ in physical composition and function.

The BBB is present along CNS blood vessels throughout the brain. Tight junction-connected BBB endothelial cells are separated from the brain parenchyma (the brain tissue proper) by two basement membranes: the endothelial and parenchymal. At the capillary level, these membranes are fused, but at all other vascular levels, these membranes separate to delineate the CSF-filled perivascular space [37]. The BBB exhibits specialized function based on its position within the overall CNS vasculature: nutrient transport occurs primarily at capillaries that lie in close proximity to neurons, whereas immune modulation occurs at the postcapillary venule (a small vessel that blood flows through after leaving the capillaries) where the perivascular space can accommodate the presence and movement of cells [38].

The BCSFB resides in the choroid plexus located in each of the four brain ventricles. The choroid plexus is composed of fenestrated, permeable blood vessels (vessels containing endothelial pores) surrounded by tight junction-connected epithelial cells that are directly exposed to the CSF [39]. Ependymal cells of the choroid plexus produce CSF from arterial blood. Newly produced CSF fills the brain ventricles, circulates around the exterior surfaces and perivascular spaces of the brain, and ultimately is reabsorbed into venous blood at the meninges. CSF flow is mediated by pulsation of the choroid plexus and action of ependymal cell cilia. Interstitial fluid of the brain parenchyma is also drained into the CSF, a fluid that acts as a surrogate for lymph by mediating immune surveillance throughout the CNS.

The CSF is considered to be an immunologically active fluid as it houses T cells, B cells, and monocytes, cells which have limited access to the CNS [36]. During physiologic conditions, T cells are primarily restricted to the CSF. However, monocytes can exit blood vessels, enter the brain, and differentiate into macrophages that primarily concentrate around CNS vasculature (perivascular macrophages), but can also be found in the meninges (meningeal macrophages) and choroid plexus (choroid plexus macrophages). CNS myeloid cells, as well as microglia (resident macrophages of the brain), have antigen-presenting capability and are considered important for immune surveillance and interaction with circulating central memory T cells.

HIV-1 is found in the CNS very early after transmission, although the associated mechanisms of entry are not well understood [22]. One theory, which is the most favored, proposes trafficking of HIV-infected CD4+ T cells into the CNS as part of routine immune surveillance or in the context of neuroinflammation. Alternatively, HIV-1 virions could cross the BBB/BCSFB, likely in the setting of high blood viral load. Consistent with this hypothesis, some studies suggest that HIV-1 can specifically interfere with the production of proteins involved in the maintenance of tight junctions thereby disrupting the integrity of the BBB [34]. In contrast, the work of Price and colleagues has shown that, in the absence of neurocognitive symptoms, the CSF viral load decreases late in infection while the viral load in the blood is on average increasing. This disparity in viral load occurs as CD4+ T cells are being depleted in the blood and white blood cell (WBC) count drops in the CSF, the latter of which suggests that T cells in the CSF are also reducing in number. These data suggest that the reduction in CSF viral load is due to a loss of CD4+ T cells that come from the blood, and thus virus enters the CNS in the form of trafficking T cells [35].

Brain pathology affects barrier function of the BBB/BCSFB, and as an associated immune response is mounted, this increases the number of CNS-infiltrating immune cells. During inflammation, CNS inflammatory cells secrete leukocyte-attracting chemokines and endothelial cells upregulate expression of adhesion molecule receptors, thereby aiding immune cell recruitment and extravasation into the CNS [38]. Neuroinflammation can be a protective immune response to CNS tissue damage or infection, but immune pathology can also compromise the BBB/BCSFB, which alters CNS homeostasis and the proportions of immune cells in the brain.

HIV-1 infection is associated with increased inflammation and immune activation both systemically and in the CNS. Neuroinflammation and BBB/BCSFB integrity both appear to affect the population of HIV-1 present in an infected CNS in a multifaceted, dynamic manner. Theories behind some of this complexity have been proposed based on studies comparing viral populations in the CSF, a surrogate for examining the CNS in living people, with those in the peripheral blood of the same individuals [22]. First, in people who have very low or undetectable CSF viral load, it is likely that virus is not replicating, at least not appreciably, in the CSF or CNS, but perhaps gains transient entry to the CNS. CSF virus likely comes from infected CD4+ T cells in the blood that cross the BBB/BCSFB and release viral progeny in the CNS. Second, in people who have elevated levels of HIV-1 in the CSF that is genetically similar to virus in the peripheral blood, it is likely that CSF virus also comes from migrating infected T cells, but now in higher levels as part of an immune response associated with increased white blood cell count in the CSF (pleocytosis). Low-level, focal replication in the CSF or CNS could also account for increased viral load with or without pleocytosis, although without establishment of a persistent CNS infection. Third, in people who have detectable CSF virus that is genetically distinct from that in the peripheral blood, CSF virus could come from transient, clonal amplification of certain viral species in the CSF or CNS that may or may not establish persistent infection over time. The term “compartmentalized” viral replication is used to describe independent replication of HIV-1 within a given bodily compartment, and is illustrated by genetic differences in the viral population between compartments, such as the CSF/CNS and blood. “Equilibrated” viral replication defines a state where HIV-1 populations are genetically similar between two compartments due to ongoing or recent intercompartmental movement of viruses.

Roughly 30% of acutely HIV-1 infected people have pleocytosis, which generally correlates with higher viral load in the CSF [22, 40]. Thus, increased viral burden in the CSF could result from an influx of infected cells in response to neuroinflammation. Consistent with this hypothesis, BBB dysfunction, indicated by an increased CSF/blood albumin ratio, often accompanies pleocytosis. A loss of barrier integrity in the context of enhanced immune cell trafficking to the CNS could allow more infected cells and/or cell-free virus to enter the CNS. Viral replication in the CSF/CNS also increases viral load in the CSF, irrespective of pleocytosis. Yet, pleocytosis occurs in a fraction of people with compartmentalized

CNS replication, and in an even greater proportion of people with equilibrated replication. How pleocytosis affects viral replication in the CNS is unclear, and it is plausible that pleocytosis occurs as a consequence of CNS HIV-1 infection. On the other hand, an influx of permissive cell types for infection could also promote viral replication in the CNS (Sturdevant 2015; Spudich 2005).

1.2.3 Potential cellular reservoirs in the CNS

CD4+ T cells

The primary target of HIV-1 infection is the CD4+ T cell, however, there are relatively few T cells in the healthy CNS. The concentration of T cells found in the CSF is less than 1% of that found in blood and even fewer, if any, are seen in the brain parenchyma. Despite the low absolute number of T cells present, the CSF has a relatively large proportion of permissive T cells; the CSF cellular composition includes primarily T cells (90% of total CSF cells), which are mostly of memory phenotype (central and effector) and recently activated (CD69+) [36]. As noted above, pro-inflammatory conditions promote immune cell influx into the CSF/CNS, thus increasing the number of potential target cells for HIV-1 replication.

The traditionally described cellular reservoir of HIV-1 is the latently infected T cell. For such a cell to contribute to a CNS reservoir of HIV-1, the cell must reside over time in the CNS. CD8+ T cells have been shown to persist in the CNS of mice infected with vesicular stomatitis virus (VSV) [41]. These cells are CD103+, which is an integrin found on tissue-resident CD8+ T cells, and expression of CD103 follows antigen recognition in the brain. Furthermore, CD103 appears to be important for retention of CD8+ T cells in the CNS, as knockdown of this molecule resulted in reduced accumulation of CNS T cells. Interestingly, CNS-resident CD8+ T cells in the brain parenchyma were shown to form clusters, some of which contained CD4+ T cells. Although CNS CD4+ T cells were not analyzed thoroughly in this study, another study showed that tissue-resident CD4+ T cells in the skin are antigen-experienced and express CD103 [42]. Taken together, these data suggest that a population of CNS-resident CD4+CD103+ T cells could exist in the CNS and harbor HIV-1.

Some evidence exists in support of HIV-1 replication in CNS T cells. In some cases there is elevated viral load in the CSF sufficient to indicate HIV-1 replication in the CSF/CNS and rapid viral decay

in the CSF upon ART initiation, suggesting that this virus was replicating in a short-lived cell, such as a T cell [26]. Some people with rapid CSF viral decay have compartmentalized replication of HIV-1 in the CSF, and CSF virus is T cell-tropic, meaning that the virus replicates best in T cells compared to other cell types. Such individuals likely have a CNS-derived population of HIV-1 arising from infected T cells, a population that differs from T cell-tropic virus in the blood. Alternatively, people with relatively high CSF viral load, equilibrated viral population, and evidence of pleocytosis may also have HIV-1 replicating in CNS T cells due to an increase in CSF/CNS T cell concentration, but in a manner that does not result in a distinct population of virus in the CSF compared to the blood [22].

Macrophages

Slow decay of CSF virus with ART suggests that, in this case, HIV-1 is being produced from a longer-lived cell type than a T cell. HIV-1 cell tropism depends at least in part on CD4 receptor expression density on the surface of a cell. R5 T cell-tropic virus replicates robustly in cells that express high levels of CD4 (T cells), but poorly in cells that express low levels of CD4, including macrophages, which have a similar number of cell surface CD4 molecules to T cells, but the molecules are less densely packed due to the larger surface area of macrophages [26].

The ability to use low levels of CD4 for cell entry is an evolved feature of the viral Env protein that cannot be attributed to a single mutation [43]. Rather, macrophage tropism likely evolves as an adaptation to the lack of CD4-rich target cells in the CNS, and the evolution of macrophage tropism appears to involve multiple genetic changes that differ between people. Pleocytosis further complicates macrophage-tropic evolution as it may alter the relative proportion of permissive cell types in the CNS thereby supporting viral replication from either T cell- or macrophage-tropic lineages. Indeed, a rhesus macaque animal model of HIV-1 infection showed that infection causes activation of bone marrow-derived monocytes and increased traffic of activated monocytes to the CNS with subsequent differentiation in to CNS macrophages [44].

The CNS is rich in macrophages that could serve as a reservoir for HIV-1. Perivascular macrophages, choroid plexus macrophages, and meningeal macrophages are all bone marrow-derived and are named for anatomical regions in which they reside. These cells could be infected by

macrophage-tropic virus or with much lower efficiency by an R5 T cell-tropic virus. Perivascular macrophages are likely exposed to cell-free or cell-associated virus that crosses the BBB. Such virus could come from either the blood or CSF, depending on barrier physiology at the point of entry. Indeed, immunohistochemical staining of autopsied brain shows the presence of HIV-1 nucleic acid and protein in perivascular macrophages. Similarly, meningeal macrophages, located at the superficial brain meninges, are likely also exposed to blood or CSF virus that crosses the leptomeningeal BBB. Choroid plexus macrophages, on the other hand, are likely exposed to predominantly blood virus, as these macrophages are located in the choroid plexus stroma, which harbors fenestrated capillaries that provide blood for the production of CSF.

A rhesus macaque animal model of HIV-1 infection suggested that the virus could migrate between the CNS meningeal and parenchymal regions or replicate autonomously in each of them [45]. The rapid migration of genetically homogeneous virus throughout the brain was associated with faster disease progression and widespread encephalitis. Furthermore, compartmentalized replication in the meninges versus parenchyma was associated with localized detrimental inflammation within these respective brain regions. One macaque with compartmentalized replication in the meninges versus parenchyma was suggested to have macrophage-tropic virus present in both regions, an observation that exemplifies how regional macrophages may contribute to CNS infection and disease. Collectively, these data suggest that the uncontrolled replication of HIV-1 in different brain regions, and potentially in regional macrophages, may be detrimental for local physiology through induction of pathological inflammation.

Microglia

Microglia are resident macrophages of the CNS and the predominant immune cell type in the brain parenchyma. Unlike macrophages, microglia are not bone marrow-derived, rather they arise during embryonic development and are maintained throughout adulthood via local proliferation [46]. Microglia have immune functions including phagocytic ability, inflammatory cytokine secretion, and weak antigen-presentation. Studies using HIV-infected human brain tissue at autopsy show that microglia can contain HIV-1 nucleic acid and protein. Furthermore, as with monocyte-derived macrophages, HIV-1 can infect microglia in vitro, suggesting that this cell type is permissive to HIV-1 infection. Like macrophages,

microglia have low surface densities of CD4, so virus capable of successfully infecting these cells is likely macrophage-tropic [26]. Microglia are thought to have very long life spans, even longer than CNS bone marrow-derived macrophages, thus persistent infection of these cells could constitute a CNS reservoir of HIV-1. Alternatively, persistent replication in this cell compartment deep in the brain parenchyma could maintain an active reservoir even in the face of poorly penetrating anti-HIV-1 therapy.

Astrocytes

Astrocytes provide mechanical and metabolic support for neurons and are the most abundant cell type in the brain [47]. Viral DNA has been detected in astrocytes of HIV-infected people, and astrocytes can be infected at low levels in vitro [48]. However, it is unclear whether astrocytes are productively infected in vivo, as they express no CD4 [49], the viral receptor. Indeed, an analysis of macrophage-tropic HIV-1 *env* genes from individuals with HIV-associated dementia failed to detect CD4-independent infection by their encoded Env proteins [50]. Yet, CD4-independent infection was characterized in another study where rhesus macaques were infected with chimeric human/simian immunodeficiency virus (SHIV) that contains an R5 T cell-tropic HIV-1 *env* in the context of an SIV backbone. Still, only one HIV-1 *env* clone isolated from the CNS of a single infected macaque was able to infect CD4- cells in vitro, so it is difficult to draw definitive conclusions from a single observation [51]. An alternative explanation for the presence of HIV-1 nucleic acid in astrocytes is that these cells have phagocytic ability and could ingest infected T cells [52]. The extent to which HIV-1 can enter cells in the absence of receptor and/or co-receptor is a poorly studied issue that deserves more attention given the number of cells without viral receptors present in a person and the concern that these alternative cell types could contribute to the reservoir.

1.2.4 Logistics and complications in evaluating the CNS reservoir

The defining criterion of a reservoir is that the virus is preserved in some form that allows for reestablishment of productive infection. In the case of the blood reservoir, latently infected CD4+ T cells can be isolated from the blood and used in a viral outgrowth assay to directly assess the prototypical latent reservoir [16]. Such methods are essentially moot for analyzing the CNS reservoir due to the

logistics of collecting viable CNS cells postmortem in a timely manner. An alternative method of determining whether HIV-1 in the CNS can reestablish infection is to characterize virus in the CSF that emerges following ART treatment interruption (rebound virus).

HIV-1 rebound virus appears in the CSF roughly two weeks after the detection of virus in the blood [53]. Phylogenetic analysis of rebound viral populations in the CSF versus blood could be used to determine if populations in these two compartments differ, indicating that “compartmentalized” CSF rebound virus comes from the CNS and thus illustrates the presence of a CNS reservoir of HIV-1. Furthermore, if viral species previously confined to the CSF arise in the blood during rebound, then re-establishment of systemic infection would be influenced by the CNS reservoir. Although it would be difficult to prove what cell type recrudesced CNS virus originates from, the use of in vitro infectivity assays would illustrate cell tropism of rebound virus. The phenotype of CSF rebound virus likely depends on the state of viral replication in the CNS prior to initiation of ART, therefore, studies are required to characterize viral populations throughout the brain of ART-naïve as well as experienced individuals.

Although evidence greatly supports the concept of HIV-1 persistence being attributable to latency, the immune privileged CNS may represent a unique anatomical reservoir, as crosstalk with the periphery is limited due to the BBB, and many ART drugs are relatively poor at penetrating the CSF/CNS. Treatment intensification using ART drugs with optimal CNS penetrance (relative to others) does not reduce levels of HIV-1 RNA in the CSF [54]. These data suggest that low level viral replication does not account for the presence of residual CSF virus. However, treatment intensification studies are limited in informing our understanding of latency in the CNS. Examination of CSF rebound virus could help fill this gap in knowledge. CNS viral persistence is further complicated by the fact that the cellular composition of the CNS is macrophage-rich with limited exposure to T cells, which reside primarily within the CSF. Thus it is possible that mechanisms of HIV-1 persistence differ between the CSF/meninges and the CNS/parenchyma and these mechanisms are affected by the presence of neuroinflammation, which alters the interaction between these bodily compartments. Finally, we do not understand the extent to which viral replication in the parenchyma is "reported" as virus in the CSF.

1.3 Examining HIV-1 Proviruses Archived in Tissues on a Whole Body Scale

We hypothesize that at least some tissues, likely the CNS, are unique environments where HIV-1 replicates and diversifies, thereby seeding anatomical reservoirs that differ from the blood reservoir. To address our hypothesis, we analyzed cell-free virus from bodily fluids (blood and CSF) and cell-associated viral sequences obtained from autopsied tissues of HIV-1-infected donors using phylogenetic methods. Our initial analysis focused on donors who died with HAD, as they are more likely to exhibit compartmentalized replication in the CNS and thus harbor distinct viruses potentially capable of infecting multiple cell types. Since we utilized samples from ART-naïve patients, cells containing HIV-1 DNA represent active infectious events (either productive or nonproductive) as well as persistent, dormant infections constituting cellular reservoirs. Furthermore, we characterized potential cell types supporting recent viral replication using a cell entry assay that illustrates viral infectivity of different permissive cell types based on receptor expression. Future directions include expanding our analysis to additional tissues of HAD donors, as well as including a greater number of HAD donors and a group of ART-treated donors.

METHODS

Study Design

We obtained repository specimens from the National NeuroAIDS Tissue Consortium (NNTC), which stores tissues generously donated for research by people who died with HIV-1. We obtained blood, CSF, and at least one type of tissue from 11 donors who died with HAD (Table 1). Brain tissue was available for five of 11 donors; only liver tissue was available for the remaining six. The majority of donors had a final CD4 count below 400 cell/mm^3 (82%; n = 9/11) and all were viremic at death.

PCR Amplification of HIV-1 Genes and Sequencing

Viral RNA was extracted from bodily fluids (blood and CSF) using the QIAmp Viral RNA Mini Kit (Qiagen). Prior to extraction, virus was concentrated by pelleting via centrifugation at 25,000xg for 1.5

hours at 4°C. Purified viral RNA was reverse transcribed using Superscript III Reverse Transcriptase (Invitrogen) and oligo-d$(T)_{20}$ according to the manufacturer's instructions. cDNA was subjected to single genome amplification (also known as end-point dilution PCR) for analysis of individual viral genomes. cDNA was serially diluted to end-point and the full-length *env* gene was obtained using Platinum Taq High Fidelity polymerase (Invitrogen) and primers that anneal within *vif* (F5010; 5'-TGCCAAGAAAAGCAAAGATCATTAG-3') and 3' LTR U3 (LTRDN1; 5'-GACTCTCGAGAAGCACTCAAGGCAAGCTTTATTGAG-3'), followed by semi-nested PCR using primers LTRDN1 and B5957UP1 (5'-GATCAAGCTTTAGGCATCTCCTATGGCAGGAAGAAG-3'). Full-length *env* (~2.5 kb) was sequenced from ~3.7 kb PCR amplicons at the UNC-CH Genome Analysis Facility. Chromatograms were analyzed for quality in Sequencher and chromatograms with double peaks, indicating amplification from more than one cDNA template, were excluded from analysis.

For analysis of proviruses in tissues, total DNA was extracted using DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's protocol. Briefly, ~2-3mm^3 pieces of tissue were cut, mechanically disrupted, and digested overnight in the presence of proteinase K, prior to DNA extraction using a column-based method. DNA was quantified using UV-Vis spectrophotometry (Eppendorf). DNA was end-point diluted and subjected to multiple rounds of PCR amplification of various HIV-1 genes as described previously [15]. A thorough description of this protocol is included below in the "Method Development" portion of the results section.

Phylogenetic Analysis of Viral Genes

DNA sequence alignments of viral *env* or *gag* obtained from bodily fluids or tissues (i.e. viral RNA or DNA, respectively) were performed using ClustalΩ. Phylogenetic trees were inferred using the neighbor-joining method and 500 bootstrap replicates (MEGA 5.2.2). Compartmentalized viral replication was assessed using the Slatkin & Maddison statistical test of population structure with 10,000 permutations (HyPhy). Highlighter plots were generated using the tool available through the Los Alamos National Laboratory HIV Database.

Cloning HIV-1 *env* for Protein Expression

Viral *env* amplicons were chosen for cloning based on the phylogenetic tree structure, so that at least one amplicon was used to represent each viral lineage. First-round PCR products were used as template in a nested PCR reaction using Phusion Hot Start High Fidelity DNA Polymerase (Finnzymes) and cloning primers. These primers were identical to those used in nested amplification of viral *env* except for the addition of 5'-CACC-3' at the 5' end of the forward primer for the purpose of topoisomerase cloning. PCR amplicons were gel-purified using the QIAquick Gel Extraction Kit (Qiagen) and 50ng of purified amplicons were cloned into the pcDNA3.1D/V5-His-TOPO expression vector (Invitrogen) using the pcDNA 3.1 Directional TOPO Expression Kit (Invitrogen). The entire cloning reaction was transformed into MAX Efficiency Stbl2 competent cells (50ul) per the manufacturer's instructions. Bacterial colonies were screened for unidirectional insertion of viral *env* using colony PCR (Platinum Taq DNA Polymerase, Invitrogen). DNA was extracted from 3-6 colonies using QIAprep Spin Miniprep Kit (Qiagen).

Co-transfection for Generating Env-pseudotyped Virus

For transfection, 293T cells were seeded at a density of 4.8×10^5 cells/well in 6-well tissue culture plates (DMEM, 10% FBS, 100 mg/ml penicillin and streptomycin culture medium). A 1:1 w/w ratio of *env* clone and backbone (pNL4-3.LucR-E-; NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH) was used for co-transfection of 293T cells in serum-free DMEM using Fugene 6 Transfection Reagent (Roche). Transfection medium was replaced five hours later with culture medium and cells were incubated at 37°C, 5% CO_2 for 48 hours. Supernatants containing pseudotyped virus were harvested, passed through a 0.45 uM filter (Millipore), and stored at -80°C.

Single-cycle Infection of Reporter Cells for Determining Cell tropism

Affinofile cells are HEK293 cell derivatives that constitutively express viral co-receptor CXCR4 and can be differentially induced to express variable levels of viral receptor CD4 and viral co-receptor CCR5 using doxycycline and ponasterone A (Invitrogen), respectively. Titration of Env-pseudotyped luciferase reporter viruses was performed in triplicate on Affinofile cells with maximum induction levels for both CD4 (6 ng/ml doxycycline) and CCR5 (5 uM ponasterone A) surface expression. In order to ensure

that subsequent experiments were performed within the linear range of cell infectivity, a volume of virus stock equivalent to 800,000 relative light units (RLUs) of luciferase expression was calculated for use in cell tropism experiments.

For all Affinofile cell experiments, 96-well black bottom tissue culture plates were treated with 10% poly-L lysine and then seeded with 1.85 x 10^4 cells/well (DMEM, 10% 12-14 kD dialyzed FBS, 50 mg/ml blasticidin culture medium). 18 to 24 hours later, expression of CD4 and CCR5 was induced at two conditions in triplicate: $CD4^{hi}/CCR5^{hi}$ (6 ng/ml doxycycline and 5 uM ponasterone A, respectively) and $CD4^{lo}/CCR5^{hi}$ (5 uM ponasterone A only). Induction medium was removed 18 to 24 hours later and replaced with 100 ul/well of fresh, warmed culture medium containing 800,000 RLUs of Env-pseudotyped luciferase reporter virus. Plates were spinoculated at 2,000 rpm for 2 hours at 37°C, and then incubated at 37°C, 5% CO_2 for 48 hours. Infection medium was removed prior to cell lysis for luciferase expression analysis (Firefly Luciferase Assay System, Promega). Lysates were stored at -80°C prior to thawing and analysis using a luminometer. Between any medium change and before cell lysis, cells were washed twice with 1X PBS. The ability to utilize low levels of CD4 expression for cell entry (macrophage tropism) was defined by percent infectivity of $CD4^{lo}/CCR5^{hi}$ relative to $CD4^{hi}/CCR5^{hi}$ cells in terms of RLUs with a 12% cutoff for defining macrophage tropism.

RESULTS

Method Development: Probe Enrichment of Provirus

Our primary goal was to analyze viral DNA sequences from both productively and latently infected cells present in tissues from HIV-infected donors. To this end, we attempted to develop a sensitive polymerase chain reaction (PCR)-based method for amplifying HIV-1 DNA extracted from infected tissues. Total DNA (cellular and viral DNA) was extracted from *in vitro* infected cells (control experiments) or ~3 mm^3 pieces of infected tissue (Qiagen DNeasy). Extracted total DNA contains cellular DNA in excess of viral DNA (perhaps by 1 billion-fold by mass), which increases the likelihood of off-target amplification. In order to improve the specificity of HIV-1 DNA amplification, we attempted to enrich for HIV-1 DNA in total DNA extracted from cells using a positive-selection purification method. In short,

we designed a nucleic acid probe for hybridization to HIV-1 sequence present in total DNA, followed by streptavidin-biotin purification of nucleic acid hybrids on magnetic beads. The DNA fraction enriched for viral DNA was then subjected to single genome amplification (SGA), also known as end-point dilution PCR, where a nucleic acid template is diluted for use in subsequent PCR steps to ensure that a single viral genome is amplified. This technique allows for phylogenetic analysis of viral populations within an individual.

We designed two ~50 base pair (bp) RNA probes complementary to conserved regions in *env* and *nef* of the HIV-1 genome. These probes are composed of "Locked Nucleic Acids" (LNAs), for optimal hybridization efficiency. LNAs have a methylene bridge that locks the ribose ring of each nucleotide in the optimal confirmation for base pairing, thereby increasing binding affinity and stability for specific, low abundance sequences [55]. The 5' end of each LNA probe is conjugated to biotin via a triethyleneglycol (TEG) spacer 15 atoms in length. The TEG spacer is more than twice the length of the standard six-carbon spacer, which increases the distance between the hybridized molecule and the surface of the magnetic sphere for reduced steric hindrance. The 5' biotin-TEG modification allows for purification of nucleic acid hybrids on streptavidin-coated magnetic beads. The 3' end of each LNA probe is modified with hexanediol, a six-carbon glycol spacer that blocks extension by DNA polymerase. This modification prevents the hybridized probe from interfering with downstream PCRs by blocking extension of DNAs at the probe-binding site.

We used a control cell line called "8E5" to optimize our method. 8E5 cells are a derivative of A3.01 cells, which is a continuous human CD4+ T cell line that is permissive to HIV-1 infection [56]. The majority of A3.01 cells die from infection, but some CD4-downregulated cells survive and contain provirus. 8E5 cells were selected from a pool of A3.01 survivor cells infected with LAV (1% nt pairwise distance from HXB2) [57]. 8E5 cells contain a single copy of defective provirus. In the region of *pol* that encodes RT, a single nucleotide insertion introduces a frameshift in the open reading frame (ORF) 3, which is corrected by a compensatory mutation in *vpr* to maintain expression of viral *env* and other genes in ORF3. This mutation results in the translation of an enzymatically inactive, truncated RT protein [58]. 8E5 cells support provirus transcription and virus production, but RT-defective progeny virions are not infectious [57, 58]. Therefore, we are able to extra total DNA from a known number of 8E5 cells and

equate cell number with provirus copy number. In method development, we controlled for provirus copy number in experiments using 8E5 provirus as template DNA. We also added controlled amounts of 293T cell DNA relative to 8E5 DNA in order to reflect a putative range of viral-to-cellular DNA ratios in natural infection. Thus, our control system better represents naturally infected cells as opposed to using a plasmid control. Our immediate goal was to amplify full-length *env* sequences from infected tissues in order to characterize viral populations and replication, as well as cell tropism.

Although our protocol was successful at purifying provirus from background DNA, it was unsuccessful at purifying all or even a majority of input proviruses: we could only purify ~1% of input proviruses with this method. We spent a great deal of time attempting to optimize our method, but were not able to improve our efficiency (Table 2). It is likely that the addition of more probes spanning the HIV-1 genome would improve provirus recovery as another lab developed such a method for *Plasmodium falciparum* (malaria) sequencing, and we are currently investigating this possibility [59]. We then moved on to alternative methods of provirus amplification.

Method Development: Full-Length Provirus PCR

Robert Siliciano and colleagues (Johns Hopkins U.) published a promising method for PCR amplification of HIV-1 DNA in latently infected T cells [15]. In short, resting CD4+ T cells were purified from PBMCs of patients on suppressive ART and total DNA was extracted for PCR amplification of proviruses. Proviruses were amplified in limiting dilution PCR and then a series of second-round nested PCRs. First-round limiting dilution PCR amplifies essentially the entire viral genome using primers that anneal within the U5 regions of the 5' and 3' LTRs (9.1 kb; Figure 1). Replicate PCR wells in the first-round reaction were then screened for the presence of provirus with a nested second-round PCR. In similar additional PCRs, overlapping fragments spanning the viral genome are obtained for sequence analysis (Figure 1).

We used the same mixed DNA controls as before (mixtures of 8E5 and 293T cell DNA) in reproducing the method published by Siliciano and colleagues. We first assessed the efficiency of this method using known quantities of proviral 8E5 DNA. Total DNA was extracted from 8E5 cells and quantified with UV spectrophotometry. Serially diluted 8E5 DNA was subjected to first-round PCR in

replicates. In order to screen for the presence of provirus in first-round PCR replicate wells, 1 ul of first-round product was used as template in the second-round nested PCR targeting *gag*. We found that *gag* amplification consistently resulted in a single band of approximately the correct size (~1.5 kb) and almost always was the expected HIV-1 sequence.

All first-round PCR wells that were positive for *gag* were then subjected to *env* amplification in a second nested PCR. Unfortunately, we found that only ~30% of the time we could obtain both *gag* and *env* from the same first-round PCR well. Furthermore, *env* PCR resulted in non-specific amplification of the human genome, resulting in a band of incorrect size or multiple bands that were sequence-verified as human genomic sequence. Since obtaining *env* is important for cell tropism analysis, we sought to determine why the *gag* PCR appeared relatively efficient compared to *env*. We found that the same first-round PCR wells were consistently positive for any subsequent PCRs, whereas those that are *gag*+/*env*- were always negative in subsequent PCRs. We also designed primers to amplify a small ~500 bp fragment of the LTR, and found that additional wells (*gag*-) were positive for LTR, indicating the presence of provirus missed by the *gag* PCR. PCR efficiency was usually best for LTR (~500 bp), then for *gag* (~1500 bp), then for *env* (~3kb), and finally for overlapping fragments spanning the whole genome (~5 kb each) indicating that PCR of smaller fragments is more efficient than larger ones. Our experiments show that first-round PCR of the essentially full viral genome is inefficient, resulting in too few template copies for subsequent PCRs.

After thorough analysis and speaking directly with the Siliciano lab, we found that increasing the primer concentration was important for improving *env* amplification efficiency. In a side-by-side comparison of 1 uM versus 0.4 uM primer in all PCRs using 8E5 template DNA, we found that only *gag*+ reactions obtained with 1 uM primer were also positive for *env*+ in every replicate well. However, increasing primer concentration does not increase the total number of LTR+ wells. These data collectively indicate that the effect of increasing primer concentration in first-round PCR increased the copy number of proviruses rather than our ability to detect them. We think that this improvement in first-round PCR efficiency is due to a greater probability of correctly priming the HIV-1 genome in the initial cycles of first-round PCR. When too little primer is present, the human genome, which is in vast excess to the single provirus present in any given well, is primed in the initial cycles of PCR and therefore amplified instead of

the provirus. Indeed, when HIV-1 sequence was not obtained from a given PCR amplicons, the amplicons almost always corresponded to human genomic sequence. We reformatted our general workflow for the remainder of this study moving forward and an associated protocol is attached (Supplemental 1).

Assessing Compartmentalized Viral Replication

We conducted a preliminary phylogenetic analysis of two individuals with disparate states of viral replication (donors 30005 and 10129). Blood and CSF were available from roughly 11 months prior to the date of death for each donor. We first compared viral *env* sequences from circulating virus (viral RNA) in the blood plasma and CSF. In donor 30005, the population of virus in the blood is distinct from that in the CSF, indicating compartmentalized replication within the CSF/CNS (Figure 2). The Slatkin & Maddison statistical test of population structure was used to confirm compartmentalization ($p = <0.0001$). Regions in *env* with the greatest sequence variability between the blood and CSF viral populations are in V1/V2 and V4/V5 with the majority of differences in the V1 and V4 hypervariable regions.

A similar analysis was completed for donor 10129. In this donor, blood and CSF virus are mixed within the overall population, which indicates that virus is similar between compartments and thus is in an equilibrated state (Figure 2). A lack of compartmentalized replication was confirmed with the Slatkin & Maddison statistical test of population structure ($p = 0.4637$). Due to a relatively low CSF viral load in this donor (328 cp/ml), few *env* sequences were obtained from the entirety of the sample, thereby sacrificing some statistical power. Low CSF viral load in this donor suggests that virus predominantly replicates in the peripheral blood and tissues of this donor, but not appreciably in the CNS.

Eight additional donors will eventually be included in this study, two of which have compartmentalized viral replication in the CNS as determined by viral *env* sequences in the CSF and blood (Table 1, work completed by Laura Kincer and Sarah Joseph). Multiple brain and peripheral tissues are available for both additional compartmentalized donors for whole body analysis of proviruses.

HIV-1 *env* Sequences from Brain Tissue Resemble Those in the CSF and Distinct Viral Lineages Exist within Different Brain Regions

Our analysis of proviral sequences was initially based on viral *env*, since the viral Env protein determines cell tropism based on the efficiency of CD4 receptor utilization. Assessing cell tropism of an archived provirus reveals which cell type(s) the virus was capable of replicating in. To analyze proviruses archived in brain tissue (viral DNA), 42 viral *env* sequences were obtained from total DNA extracted from donor brain tissues (Table 1). 31 of 42 *env* sequences came from donor 30005, 21 of 31 *env* genes from this donor are defective as they contained deletions and 10 *env* genes are putatively functional. Donor 30005 had highly compartmentalized viral replication in the CSF (viral RNA) 11 months prior to death, and five full-length *env* sequences (four functional and one defective) were obtained from provirus in the frontal and occipital lobes for incorporation into the phylogenetic tree of cell-free virus described above. Proviral sequences in the brain most closely resemble virus in the CSF compared to the blood, suggesting compartmentalized replication within the overall CNS at death as well as 11 months prior. Three sequences obtained from the frontal lobe suggest the presence of a brain region-specific viral lineage (100% bootstrap), which appears to be distinct from CSF virus, although the statistical significance is low (<70% bootstrap).

PCR amplification of *env* in tissue DNA is relatively inefficient compared to *gag*, at least with our protocol. Therefore, we had a greater number of *gag* sequences to analyze with phylogenetics. We compared *gag* sequences obtained from virus in the blood plasma and CSF (viral RNA) and virus in tissues (viral DNA; Figure 3) of donor 30005. As predicted, virus obtained form brain tissue more closely resembles CSF virus as opposed to that in the blood. Furthermore, compartmentalized viral replication in the CNS at death was verified using predominantly proviral sequences. The presence of brain region-specific viral lineages was also confirmed with our *gag* data as the frontal and occipital lobes contain lineages that are distinct from one another. Furthermore, multiple lineages exist within each lobe: three in the frontal lobe and two in the occipital lobe.

Clonal amplification of particular viral species occurred in both lobes, which, in the absence of ART, suggests that selective pressure exists within these brain regions. Clonal viral sequences could have arisen from either a local burst of viral replication or from clonal expansion of latently infected cells.

The presence of two CSF viruses, which came from a sample collected 11 months prior to death, within the clonal lineage in the occipital lobe suggests that clonally expanded cells arose from an infectious event, likely from viral clones of CSF virus C6, that occurred at least 11 months prior to death.

A highlighter plot of each clonal viral lineage within the tree for donor 30005 was used to find evidence of viral replication as opposed to clonally expanded cells (Figure 4 and data not shown). We analyzed highly similar occipital lobe sequences using a representative of eight identical *gag* sequences (seven from brain tissue and one from the CSF) as the master sequence with which every other sequence is compared: the CSF virus (C6) was chosen, as it represents actively replicating virus 11 months prior to tissue collection at death. This analysis compares sequence variation relative to the seven clonal sequences within a lineage composed of highly similar sequences. If clonal expansion of cells occurred, then we would expect to see identical nucleotide differences in the remaining sequences when grouped by branch length (i.e. the five sequences with medium branch lengths and the three with longer). However, this lineage appears to have come from a single infectious event followed by local replication in neighboring cells, as nucleotide differences at any given position in the remaining sequences are randomly distributed. The same was observed in a similar analysis of the frontal lobe (data not shown).

11 *env* genes were obtained for donor 10129, and all but one (occipital lobe) came from the liver. 7 of 11 *env* genes are putatively functional (1 sequence needs to be verified) and 4 of 11 are defective (3 with deletions and 1 with stop codons). In donor 10129, the CSF viral load was very low compared with a high blood viral load (328 v. >750,000 cps/ml), suggesting that no appreciable viral replication occurred in the CNS. Similarly to the CSF, few HIV-1 proviral sequences (*env* or *gag*) were obtained from the brain of this donor (occipital and frontal lobes and cerebellum were sampled). Indeed, only one *env* sequence was obtained from the brain, particularly the occipital lobe, and this sequence corresponded to a defective provirus due to a deletion from *pol* to *env* (data not shown). Thus, no functional *env* genes (viral DNA) were obtained from the brain of donor 10129 and only 7 functional *env* genes (viral RNA) were obtained from 1 mL of CSF. Furthermore, there was no evidence of compartmentalized or clonal replication in the CSF/CNS. These data collectively suggest that virus did not replicate in the CNS of this donor, irrespective of HAD diagnosis for this donor.

The CSF we received from donor 10129 was exhausted in the initial analysis of viral *env* due to low CSF viral load, therefore CSF *gag* sequences could not be obtained for this donor. Few proviral *gag* sequences were obtained from the brain, 2 from the frontal and 1 from the occipital lobe, again suggesting a lack of appreciable viral replication in the brain. The equilibrated state of virus throughout the body in this donor was also verified with *gag* sequences (Figure 5).

Individual Viral Lineages Also Exist within the Liver, a Peripheral Tissue Exposed to Blood Virus

In donor 30005, 3 of 5 full-length proviral *env* sequences obtained from the liver (viral DNA) grouped with those in the blood (viral RNA), indicating that virus can travel between these compartments, as predicted. There are two liver viruses distinct from those that grouped with the blood: E4 and D3. Liver virus E4 is more similar to blood virus than CSF virus, but is still a statistically significant distinct entity (93% bootstrap). Liver virus D3 is positioned near a unique viral lineage (blood virus A8 and CSF virus B3, 100% bootstrap) in the phylogenetic tree, but with a poor bootstrap value of 51%. These distinct liver sequences may represent minor viral lineages that are sampled poorly with *env* PCR. Therefore, we analyzed *gag* sequences, as a greater number were available than *env*.

Comparison of HIV-1 proviral *gag* sequences from donor 30005 revealed individual viral lineages within a given tissue, including the liver (Figure 3). At least two viral lineages were found within the liver: one that resembles virus in the blood and another that resembles CSF virus. Since sequences obtained from the frontal lobe as well as the liver resemble the clonal lineage within the occipital lobe, we created a highlighter plot to determine if any sequence features are shared between compartments. Indeed, when designating CSF virus H2 as the master sequence, two additional CSF sequences (A3 and B2), four frontal lobe sequences (G8, H5, F9, and H9), and one liver sequence (B12) all contain the same two mutations observed in the occipital lobe clonal lineage (Figure 4). These data suggest that virus produced in the occipital lobe can travel to the frontal lobe, likely via CSF, and replicate there. It is important to note that CSF flow directionality favors this hypothesis. CSF virus that seeded the brain and locally replicated may have then exited the CNS, entered the periphery, and replicated in liver tissue.

The same first-round PCR products (amplified provirus) were used as template in subsequent, nested *gag* and *env* PCRs so that both genes could be characterized for an individual provirus when

possible. Both *gag* and *env* sequences were obtained for eight viruses in donor 30005. The *env* and *gag* trees for this donor have the same cluster patterns (Figure 3), and since more *gag* sequences are available than *env*, the *gag* tree can inform our understanding of how prevalent viruses with certain *env* genes are. Indeed, unique liver virus D3 represents a major lineage in the *gag* tree, which is more closely related to the CSF/CNS viral population than that in the blood. Furthermore, this lineage appears to have developed from viral replication shortly before death, as no blood or CSF (collected 11 months prior to death) sequences are present in this lineage, but are present in others.

In donor 10129, *env* sequences could only be obtained from liver tissue (none from the brain). The presence of one cluster of *env* sequences from the liver suggests the presence of tissue-specific lineages. Additionally, there was clonal amplification of virus in the liver in the absence of ART or compartmentalized viral replication in the CNS (based on *gag*). Since the proviruses composing the clonal lineage are all hypermutants and therefore non-infectious, these sequences likely came from expansion of infected cells through cell division. Otherwise, analysis of *gag* sequences from this donor confirmed that virus was generally equilibrated throughout the body (Figure 5).

Macrophage-Tropic Virus is Archived as Provirus in Brain Tissue

Viral *env* sequences from the blood and CSF of donors 30005 and 10129 were analyzed for cell tropism using our Affinofile assay. Viral *env* genes were individually cloned into an expressional vector for generating pseudotyped virus. 293T cells were co-transfected with an *env* clone and an *env*-deficient viral backbone genome (pNL4-3) for the production of pseudotyped virus capable of a single infection. The pNL4-3 backbone also contains a luciferase gene in viral *nef* for the reporter. Affinofile cells were induced to express CCR5 and either $CD4^{hi}$ or $CD4^{lo}$ prior to infection with pseudotyped virus. Luciferase expression, indicative of single-round infection, was measured in relative light units (RLU) and normalized to background (backbone-only). Cell tropism was characterized by percent infectivity of $CD4^{lo}$ relative to $CD4^{hi}$, where macrophage-tropic virus was defined as equal or greater efficiency of $CD4^{lo}$ utilization compared to a macrophage-tropic control (pseudotyped BaL). T cell-tropic virus was defined as poorly utilizing $CD4^{lo}$ compared to BaL and a T cell-tropic control virus was used (pseudotyped JR-CSF).

The CSF of donor 30005, with highly compartmentalized viral replication in the CSF/CNS, contained macrophage-tropic virus in each viral lineage, whereas the blood contained T cell-tropic virus (Figure 2, work completed by Sarah Joseph and Laura Kincer). It is important to note that two CSF viruses grouped with virus in the blood and one of the two was confirmed to be T cell-tropic. These viruses either represent a small number of T cell-tropic CSF virus, or, perhaps more likely, minor blood contamination during lumbar puncture in collecting this donor's CSF.

Although cell tropism analyses have not yet been completed for proviral sequences, three proviruses from the frontal lobe and two from the occipital lobe grouped with the overall CSF viral population, and may represent macrophage-tropic viruses. One provirus in particular (occipital lobe virus C10) is most similar to a CSF viral lineage that is macrophage-tropic and is closely related in amino acid composition to the CSF virus that was assessed for tropism (CSF virus E7; pairwise distance = 0.8%). These data suggest that macrophage-tropic virus is archived as provirus in brain tissue. Likewise, three proviruses from the liver grouped with the blood virus population, which contains T cell-tropic virus, and likely have the same tropism.

Liver virus D3 has a unique *env* that is distinct from those in the major blood and CSF viral populations, yet its *gag* resembles macrophage-tropic CNS virus (Figure 3). In contrast, liver virus E4 has a unique *env*, but its *gag* resembles T cell-tropic blood virus. Viruses D3 and E4 may represent macrophage-tropic and T cell-tropic viral lineages in the liver, respectively. Furthermore, since virus D3 is part of a viral lineage in the liver that appears to have arisen shortly before death when the CD4+ T cell count was very low, it is tempting to speculate that the loss of permissive T cells in a macrophage-rich tissue, such as the liver, favored viral replication in tissue macrophages. We are currently cloning all full-length, putatively functional *env* genes obtained from tissues of this donor for cell tropism analysis.

In donor 10129, virus in the blood and CSF are all T cell-tropic (Figure 2). This donor had an equilibrated state of virus throughout the body and no indication of compartmentalized viral replication in the CNS. Coupled with a low CSF viral load, these data indicate that virus was not replicating appreciably in the CNS and therefore had limited selective pressure to replicate in an alternative cell type, such as a macrophage. If such a selective pressure exists within macrophage-rich peripheral tissues, such as the liver, then a greater sample size of donors may be required to address this possibility.

DISCUSSION AND FUTURE DIRECTIONS

HIV-1 infection is incurable due to viral persistence in cellular and anatomical reservoirs despite ART-mediated viral suppression. The CNS is a unique bodily compartment that can support viral replication independent of that in the periphery and may be an anatomical reservoir for HIV-1. The presence of barriers between the CNS and periphery allows compartmentalized viral replication in the CNS. “Compartmentalization” is statistically inferred from a phylogenetic tree where the viral population in the CSF, representative of the CNS, is distinct from that in the blood. This type of replication often correlates with macrophage-tropism of virus sampled from the CSF. Macrophages may serve as a cellular reservoir within the CNS due to their long half-life, ability to support HIV-1 replication, and high frequency relative to T cells in the CNS.

We conducted a preliminary analysis of two donors with disparate states of viral replication (30005 – compartmentalized; 10129 – equilibrated). Our pilot study showed that HIV-1 *env* and *gag* sequences from brain tissue resemble those in the CSF, specifically in donor 30005. The frontal and occipital lobes of this donor contain lineages that are distinct from one another and multiple lineages exist within each lobe. These data suggest some degree of compartmentalized replication in the context of tissue architecture and associated selective pressures. Indeed, we found clonal amplification of virus in the occipital and frontal lobes. Furthermore, we found that virus can intermix between different regions of the brain, thereby diversifying the overall population of virus in the brain, yet not to a degree of diversity seen in the periphery.

A recent SIV study in macaques showed that compartmentalized viral replication can occur in the meninges versus parenchyma of the brain in the context of both rapid and conventional disease progression with encephalitis [45]. CSF virus grouped only with virus in the meninges for one animal, whereas for the rest, CSF virus grouped with both the meningeal and parenchymal viruses. Furthermore, similar viruses were found within two distinct perivascular lesions. These data collectively highlight the importance of CNS structure with respect to viral replication in the brain. Our study confirmed this observation and furthermore did so in the context of full-length *env* as opposed to V1/V2 only, where the former allows characterization of functional versus defective *env* sequences. We found that the majority

of *env* sequences in the brain of donor 30005 were defective, which is consistent with previous studies of latent virus in PBMCs [15]. We also showed that different regions in the brain (i.e. brain lobes) harbor distinct populations of virus that can share genetic information.

We found that individual viral lineages exist within the liver of donor 30005, a peripheral tissue exposed to blood virus. The same appears to be true for donor 10129, although a greater number of sequences are required for confirmation. Clonal amplification of hypermutated provirus also occurred in the liver of donor 10129, in the absence of compartmentalized viral replication in the CNS. This lineage likely arose from expansion of infected cells through cell division, which was not observed in donor 30005, where instead clonal sequences in the brain came from local viral replication. The presence of clonal lineages suggests that selective pressures exist within the peripheral tissues as well as the CNS to drive viral evolution. It is possible that clonal lineages composed of highly similar viral sequences arose by chance through founder effect or a population bottleneck, but the presence of “tissue-specific” lineages composed of more diverse viral sequences within a cluster reiterate the likelihood of selective pressure imposed by the tissue environment.

Clonal amplification of virus in particular tissues has been shown previously, but identical viral clones were also found in other tissues and no tissue-specific clustering patterns were observed [60, 61]. In contrast, our study shows both clonal amplification and tissue-specificity (of more diverse sequences within a cluster) of virus in the liver. It is possible that when more tissues are sampled, viral clustering patterns in the liver will change, intermixing virus in the liver with virus in other tissues. Still, we hypothesize that, due to either selective pressure by the tissue microenvironment or chance (i.e. founder effect or a population bottleneck), certain viral species are more fit than others to replicate in a tissue, which is likely influenced by the local population of permissive immune cells (i.e. relative proportions of T cells and macrophages).

The CSF of donor 30005, with highly compartmentalized viral replication in the CSF/CNS, contained macrophage-tropic virus in each viral lineage, whereas the blood contained T cell-tropic virus. Macrophage-tropic virus is also archived as provirus in brain tissue, as three proviruses from the frontal lobe and two from the occipital lobe grouped with the overall CSF viral population. Importantly, ART-treated patients can have undetectable or very low viral load in the blood and higher viral load in the CSF,

termed “CSF viral escape”, where escape virus could originate from either trafficking cells that release virus in the CNS for local replication or from CNS-resident macrophages releasing virus from a reservoir [62]. The presence of provirus containing functional *gag* and *env* in macrophage-tropic viral lineages in the brain supports the idea of macrophages as a cellular reservoir for HIV in the CNS. When considering the observed similarity between certain proviruses in the brain and liver of donor 30005, it is possible that evolutionary features of CNS virus (i.e. macrophage-tropism) are transferred to the population of virus in the liver. The liver is a tissue rich in resident macrophages and could be an ideal source for macrophage-tropic virus in the periphery.

Studies have found that macrophage tropism occurs only in people who have compartmentalized viral replication in the CSF/CNS. Indeed, donor 30005 died with HAD and a high viral load in the CSF due to CSF/CNS compartmentalization. Since the CSF of this donor contained entirely macrophage-tropic virus, it is likely that macrophages sustained viral replication in the CNS when CD4+ T cell count was falling. Importantly, macrophage-tropic virus persisted in the CNS of this donor for at least 11 months prior to death, suggesting that selective pressures to drive macrophage-tropism exist prior to total CD4+ T cell depletion. The general paucity of CD4+ T cells in the CNS and/or the influence of tissue microenvironments may select for macrophage-tropic virus gradually, as proposed [43].

This pilot study was a ‘proof of principle’ trial in analyzing proviral sequences archived in various tissues throughout the body. Here we characterized the differences in viral populations within brain and liver tissues, as well as in blood and cerebrospinal fluids, of a highly CNS compartmentalized donor versus an equilibrated donor. With the development of a plausible method and workflow (Supplemental 1), additional tissues from the two initial donors, as well as those from new donors (Table 1), are currently being analyzed for incorporation to this study. With a greater sample size, we hope to gain statistical power and further characterize viral populations on a whole body scale for donors in disparate states of viral replication.

Table 1: Study Population Characteristics

Donor[a]	Blood VL (cps/ml)[b]	CSF VL (cps/ml)[b]	CSF/blood VL Ratio	CSF/CNS Compartment.? (p value)	Tissues Available[c]	# Sequences (*gag*, *env*)	
30005	287309	31251	0.11	Yes, p < 0.0001	Blood	8, 30	Initial Analysis
					CSF	10, 23	
					Frontal lobe	14, 6	
					Occipital lobe	22, 18	
					Cerebellum	1, 1	
					Basal Ganglia		
					Liver	35, 8	
					Lymph node		
					Spleen		
					Testis		
10129	>750000	328	<0.0004	No, p = 0.4637	Blood	6, 22	
					CSF	0, 7	
					Frontal lobe	2, 0	
					Occipital lobe	1, 0	
					Cerebellum	0, 0	
					Basal Ganglia		
					Choroid plexus		
					Liver	50, 8	
					Lymph node		
					Spleen		
					Testis		
6800127569	>750,000	>750,000	n/a, ~1	Yes, p < 0.0007	Blood	n/a, 18	Additional Compartmentalized
					CSF	n/a, 18	
					Liver		
					Frontal lobe (gray matter)		
					Frontal lobe (white matter)		
					Occipital lobe		
					Cerebellum		
					Basal Ganglia		
					Leptomeninges		
					Lymph node		
7100616568	489796	>750000	>1.53	Yes, p < 0.0002	Blood	n/a, 11	
					CSF	n/a, 14	
					Liver		
					Frontal lobe (gray matter)		
					Frontal lobe (white matter)		
					Occipital lobe		
					Cerebellum		
					Basal Ganglia		
					Choroid plexus		
					Leptomeninges		
					Spleen		
7200537480	40133	2747	0.07	No, p = 0.2495	Blood	n/a, 18	Additional Not Compartmentalized
					CSF	n/a, 8	
					Liver		
					Frontal lobe (gray matter)		
					Frontal lobe (white matter)		
					Occipital lobe		
					Cerebellum		
					Basal Ganglia		
					Choroid plexus		
					Spleen		
6800207680	15906	76	0.005	n/a	Blood	n/a, 17	
					CSF	n/a, 1	
					Liver		
10119	312240	<50	<0.0002	n/a	*Blood*		
					CSF	n/a, 0	
					Liver		
30020	>750000	<50	n/a, ~0.0001	n/a	*Blood*		
					CSF	n/a, 0	
					Liver		
20024	222840	501	0.002	n/a	*Blood*		
					CSF	n/a, 3	
					Liver		
10105	5722	650	0.11	n/a	*Blood*		
					CSF	n/a, 2	
					Liver		
10051	20103	180	0.009	n/a	*Blood*		
					CSF	n/a, 0	
					Liver		

Footnotes

a: Five additional donors with low CSF viral load (4/5 below detection limit) could also be included in this study.

b: Final blood and CSF viral loads (at or one year prior to death) were provided by NNTC.

c: Tissues were collected at death and were provided by NNTC. Tissues in italics have not yet been analyzed.

Table 2: Attempts to Optimize Efficiency of Probe Method (Unsuccessful)

Parameter Evaluated
Sheared v. unsheared total extracted DNA
Shearing by boiling v. g-Tube
nef probe v. env probe
Removal of excess probe prior to bead purification
More probe
Longer hybridization time (10X)
Slow cool hybridization
More probe + longer hybridization time
Rehybridization of supernatant from bead purification
2nd round bead purification on supernatant from 1st bead purification
More beads
More beads + smaller volume of binding solution
Base removal of purified proviruses from beads

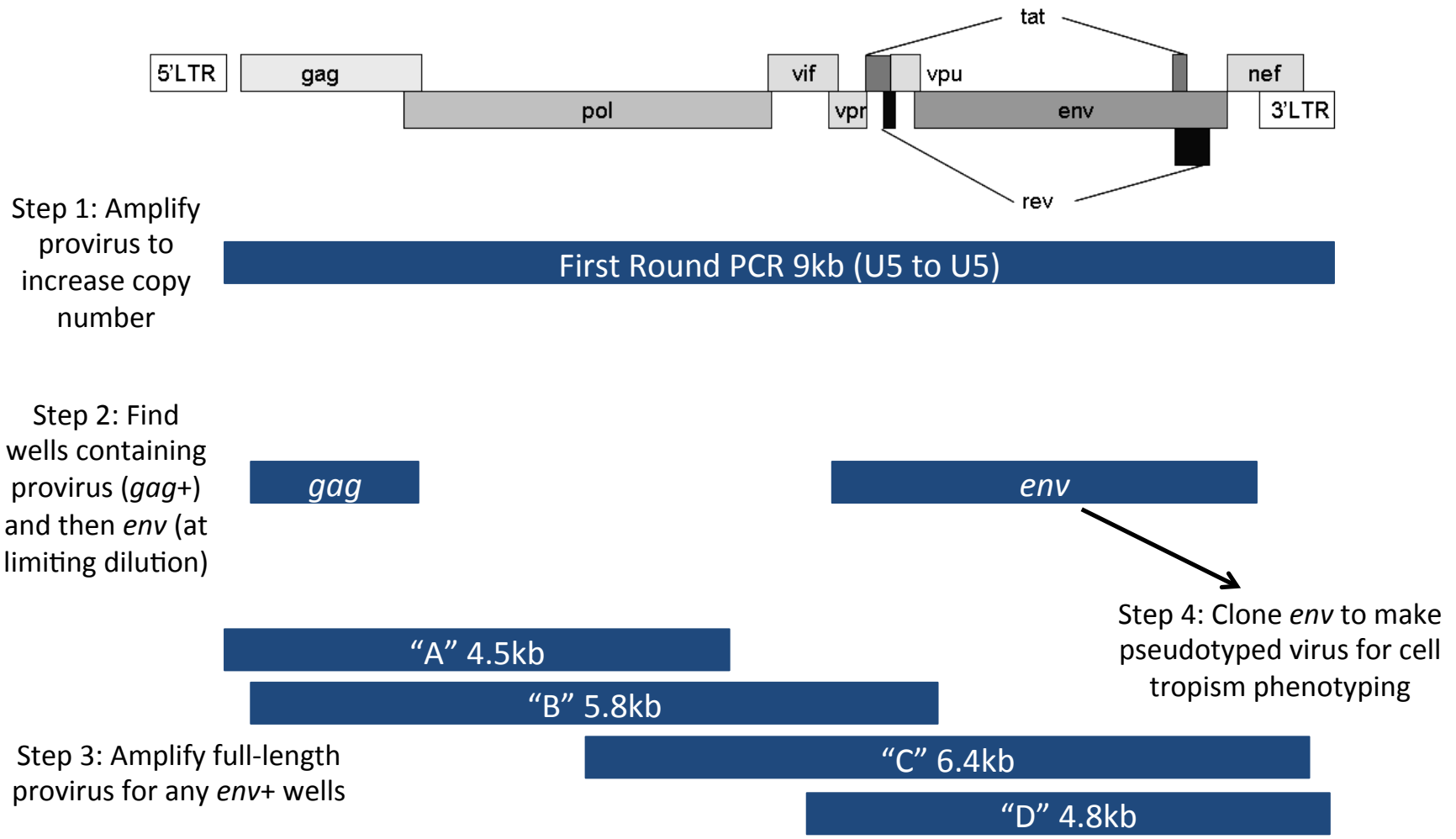


Figure 1. Method workflow for characterizing proviruses present in extracted total DNA from cells or tissue. First, essentially the full-length HIV-1 genome is amplified from 5' to 3' LTRs (U5 to U5) to enrich for provirus copy number. Subsequent PCRs are then performed using the first-round PCR product as template. Sequences are obtained for each amplicon in subsequent PCRs. Note that this method was developed by Robert Siliciano ant colleagues at Johns Hopkins University.

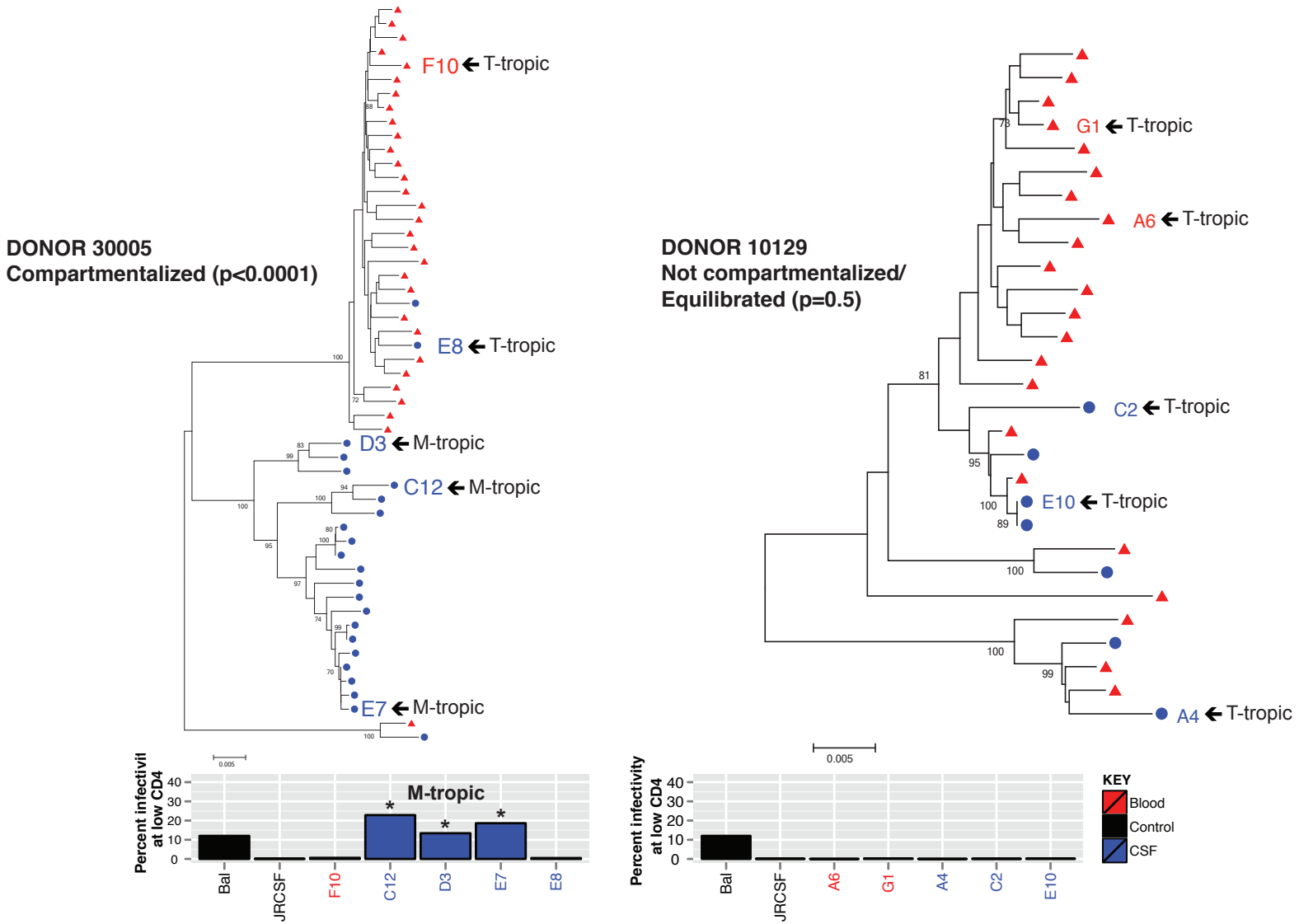


Figure 2. Compartmentalized viral replication in the CSF/CNS is associated with macrophage-tropic virus in the CSF. Viral *env* sequences were analyzed in a Neighbor-joining phylogenetic tree for donors 30005 and 10129. Donor 30005 had statistically significant (Slatkin & Maddison statistical test) compartmentalized viral replication in the CSF/CNS, whereas donor 10129 did not (equilibrated). Viral *env* was cloned for making pseudotyped virus for cell tropism analysis using Affinofile cells. Viral *env* clones were defined as macrophage-tropic if percent infectivity at low CD4 was at least that of the positive control, BaL (12%). Analyses were completed by Sarah Joseph and Laura Kincer.

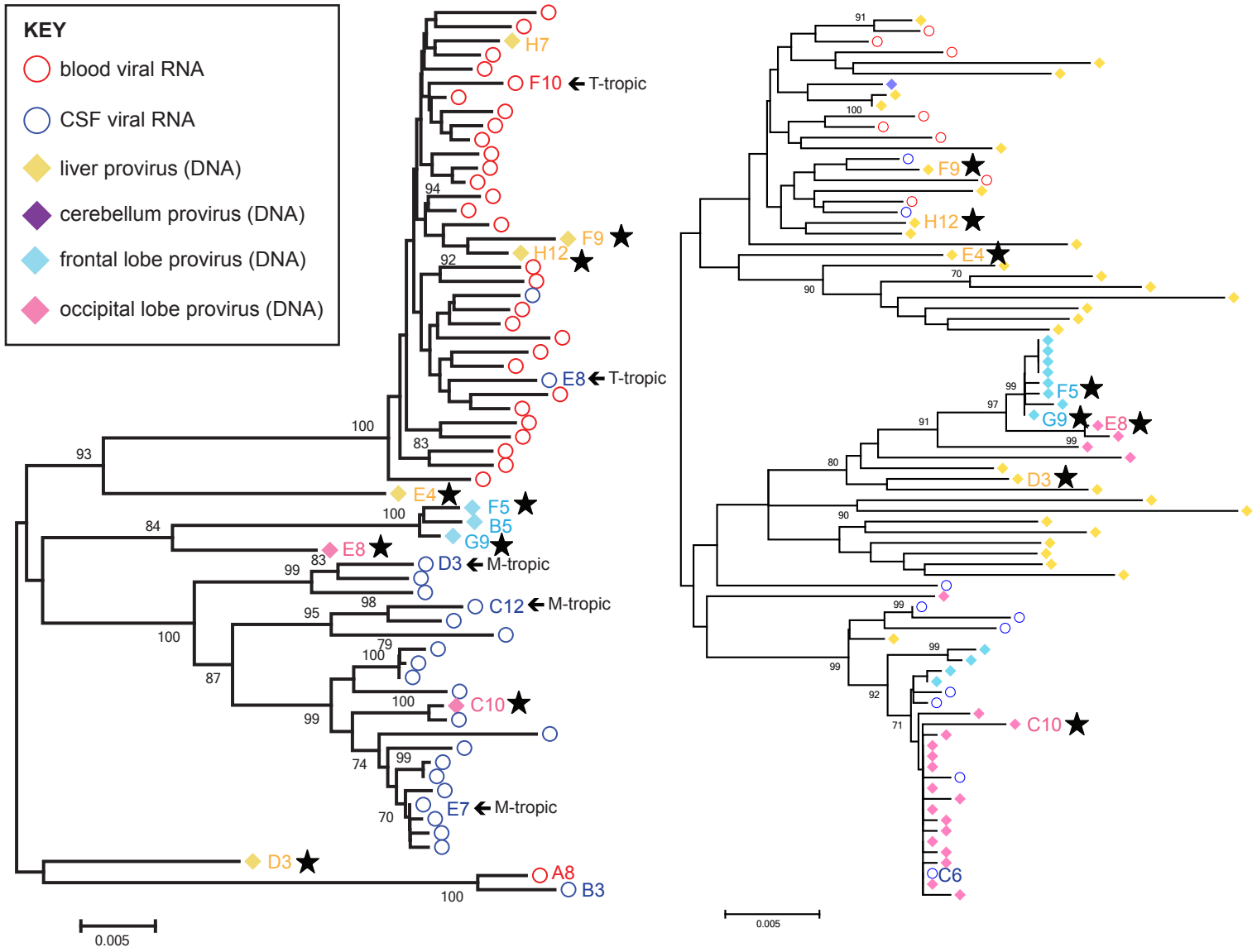


Figure 3. Tissue-specific viral lineages exist throughout the brain and within the liver in the presence of compartmentalized viral replication in the CNS. Viral *env* (left) and *gag* (right) sequences were analyzed in Neighbor-joining phylogenetic trees for donor 30005 (left). Stars (★) designate *env* and *gag* sequences that came from the same provirus (same first-round PCR well); arrows (🡐) designate macrophage-tropic (M-tropic) or T cell-tropic (T-tropic) virus analyzed in Affinofile cells.

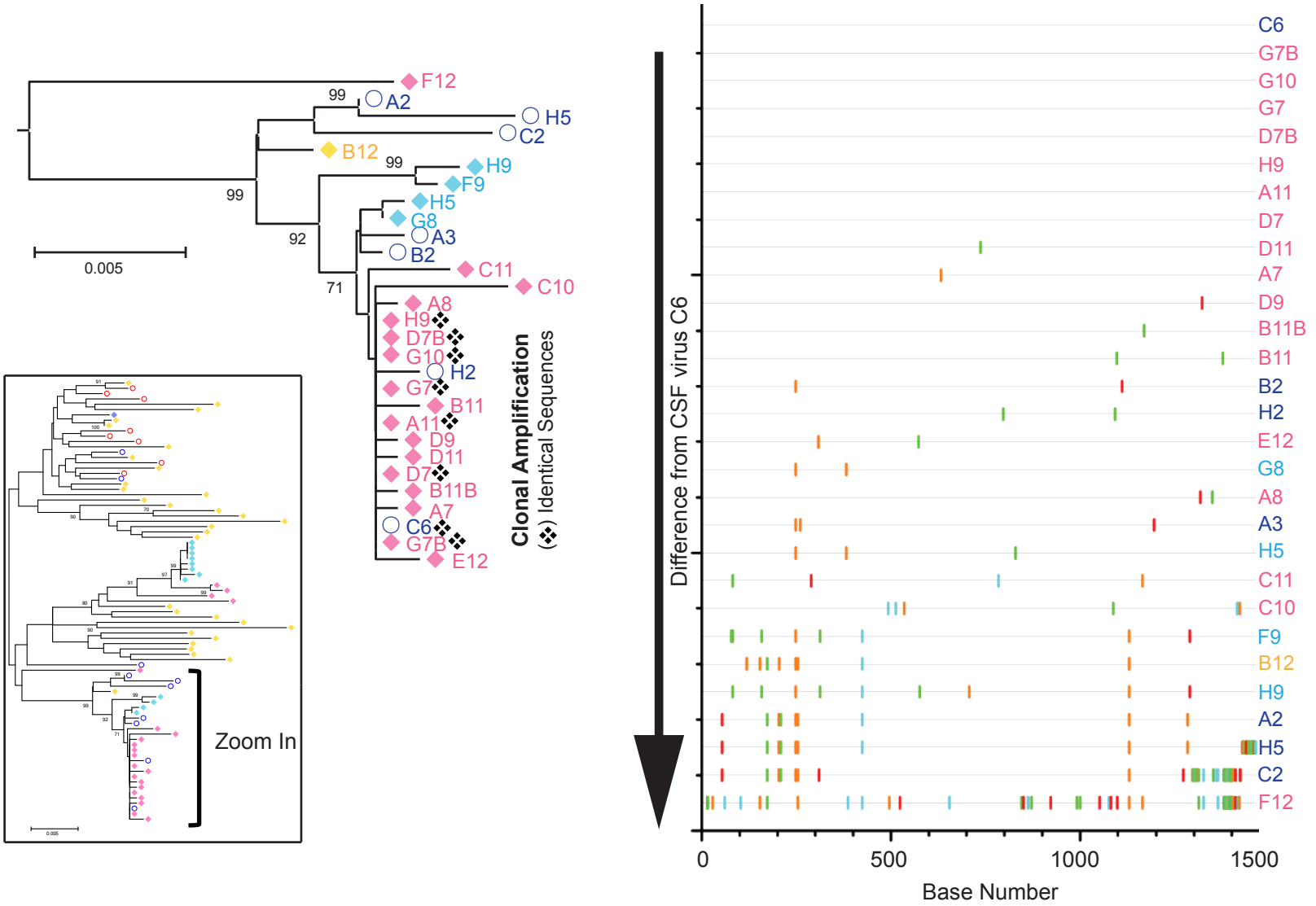


Figure 4. Clonal amplification of viral species through local replication diversifies the population of virus throughout the CNS and in the peripheral liver tissue. Viral *gag* sequences belonging to a group of viral clones were analyzed through phylogenetics (left). Nucleotide differences relative to CSF virus C6 are illustrated in the Highlighter Plot (right). (❖) designate identical gag sequences within the occipital lobe.

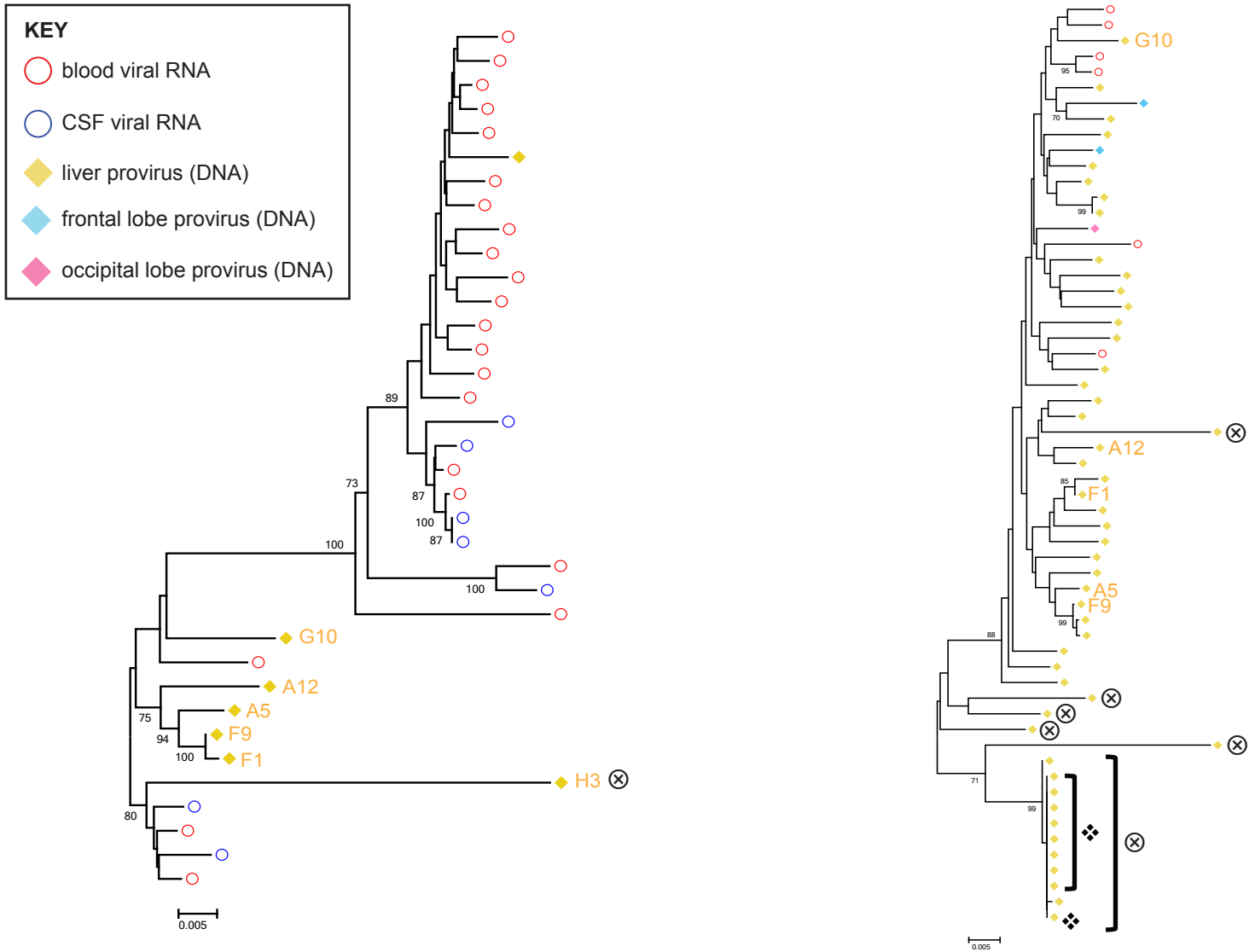
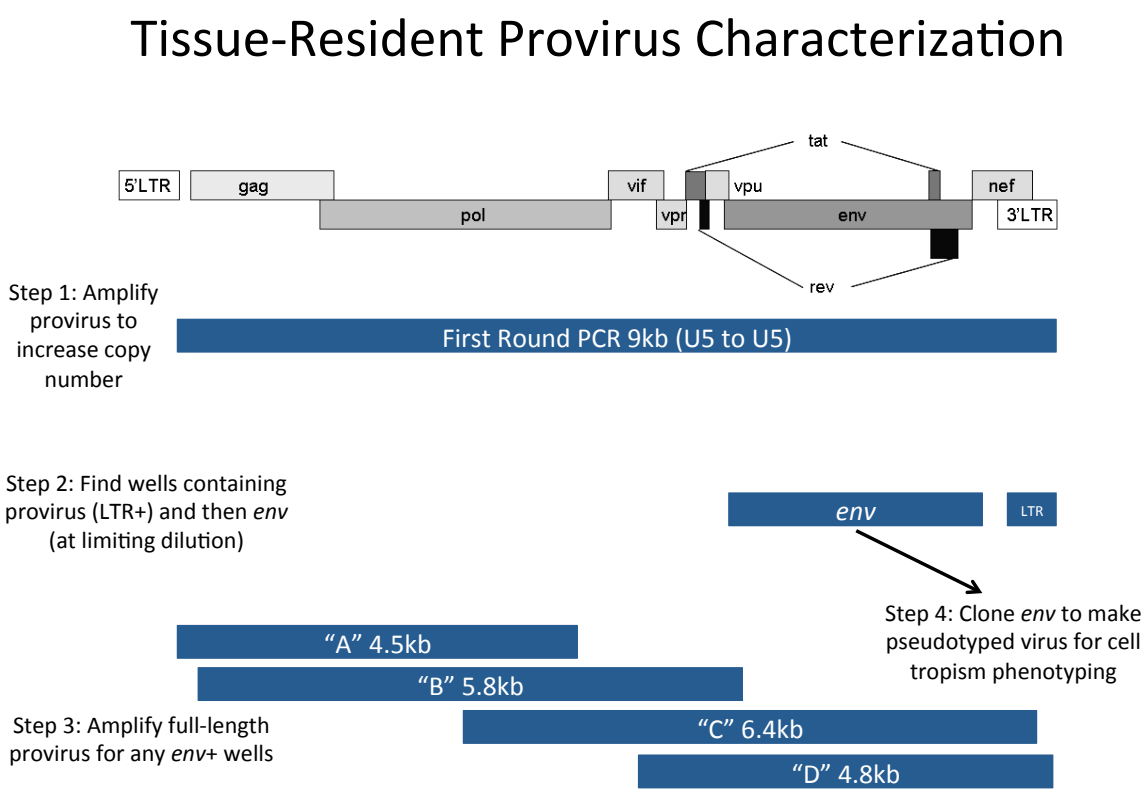


Figure 5. Clonal amplification occurs in the liver irrespective of an otherwise equilibrated viral population throughout the body. Viral *env* (left) and *gag* (right) sequences were analyzed in Neighbor-joining phylogenetic trees for donor 10129. Viral *gag* and *env* sequences corresponding to the same provirus (same first-round PCR well) are designated alphanumerically. A clonal lineage composed of 9 identical, defective *gag* sequences is identified with (❖) and hyptermutants are designated with (⊗). Viral *env* and *gag* corresponding to the same provirus are labeled.

APPENDIX: MODIFIED PROTOCOL FOR PROVIRUS CHARACTERIZATION IN TISSUES

Prepared by Lauren A. R. Tompkins 9/1/15
Adapted from: Katie Bruner (adapted from Ya-Chi Ho
Siliciano lab, Johns Hopkins University School of Medicine)

General Overview



STEP 1: Limiting dilution first round PCR of provirus ("SGA1")

- Obtain DNA extracted from tissue using DNeasy Kit (QIAgen) and record concentration (ng/ul)
- PCR must be done at limiting dilution to amplify from a single provirus template, so first find the best amount of DNA input to achieve ~1/3 LTR+ wells
- Use a maximum of 100ng DNA per well or less and record amount of DNA added per 24 wells (an example is shown below)
- See SGA1 master mix preparation below (copied from Excel spreadsheet)

Set up SGA1: (add to mastermix for 24 wells)

8uL of extracted DNA spread over 24 wells
4uL of extracted DNA spread over 24 wells
2uL of extracted DNA spread over 24 wells
1uL of extracted DNA spread over 24wells

SGA1: Provirus PCR

How many reactions? --> 24

Ingredient	Vol for 1 rxn, uL	Vol for MM, uL
H2O	31.75	762
10X buffer	5	120
MgSO4	2	48
dNTP	1	24
10 uM BLOuterF	5	120
10 uM BLOuterR	5	120
Platinum Taq HiFi	0.25	6
Total	50	1200

Make 4 identical MM's, add DNA directly to it, then distribute over 24 PCR wells each.

Thermalcycling

	94°C	2 mins.
3 cycles of:	94°C	30 secs.
	64°C	30 secs.
	68°C	10 mins.
3 cycles of:	94°C	30 secs.
	61°C	30 secs.
	68°C	10 mins.
3 cycles of:	94°C	30 secs.
	58°C	30 secs.
	68°C	10 mins.
21 cycles of:	94°C	30 secs.
	55°C	30 secs.
	68°C	10 mins.
	68°C	10 mins.
	4°C	∞

STEP 2: LTR PCR (to check for wells containing provirus at limiting dilution/clonality) and Env PCR (to find wells containing env)

- LTR PCR to find a dilution of tissue DNA where ~1/3 wells are positive for provirus (limiting dilution/clonality)
- Env PCR to find proportion of LTR+ wells containing functional env AND to obtain env genes for cloning and Affinofile cell tropism phenotyping
- See LTR and Env master mix preparation below (copied from Excel spreadsheet)

LTR		
How many reactions? -->	100	
Ingredient	Vol for 1 rxn, uL	Vol for MM, uL
H2O	31.75	3175
10X buffer	5	500
MgSO4	2	200
dNTP	1	100
10 uM F9058	5	500
10 uM LTR9556R	5	500
Platinum Taq HiFi	0.25	25
Total	50	5000
Add 1 uL of SGA1		
Thermalcycling		
	94°C	2 mins.
3 cycles of:	94°C	30 secs.
	64°C	30 secs.
	68°C	1 min.
3 cycles of:	94°C	30 secs.
	61°C	30 secs.
	68°C	1 min.
3 cycles of:	94°C	30 secs.
	58°C	30 secs.
	68°C	1 min.
31 cycles of:	94°C	30 secs.
	55°C	30 secs.
	68°C	1 min.
	68°C	10 mins.
	4°C	∞

Env		
How many reactions? -->	100	
Ingredient	Vol for 1 rxn, uL	Vol for MM, uL
H2O	31.75	3175
10X buffer	5	500
MgSO4	2	200
dNTP	1	100
10 uM 5EnvIn	5	500
10 uM 3EnvIn	5	500
Platinum Taq HiFi	0.25	25
Total	50	5000
Add 1 uL of SGA1		
Thermalcycling		
	94°C	2 mins.
3 cycles of:	94°C	30 secs.
	64°C	30 secs.
	68°C	3 min. 30 secs.
3 cycles of:	94°C	30 secs.
	61°C	30 secs.
	68°C	3 min. 30 secs.
3 cycles of:	94°C	30 secs.
	58°C	30 secs.
	68°C	3 min. 30 secs.
31 cycles of:	94°C	30 secs.
	55°C	30 secs.
	68°C	3 min. 30 secs.
	68°C	10 mins.
	4°C	∞

STEP 3: Sequencing verify LTR and Env amplicons

- Goal is to sequence-verify LTR and Env amplicons
  - LTR: sequence with amplification primers (F9058 and LTR9556R); try direct sequencing first, and go back and gel-purify if needed
  - Env: sequence with primers 5EnvIn, R6458, B/CV1, F15, R15, 7320DN, F7510, F7913, R8056, F8231, R8398, F8645; try direct sequencing first, and go back and gel-purify if needed
- Wells containing functional, full-length *env* will be used for characterizing the full-length genome described below

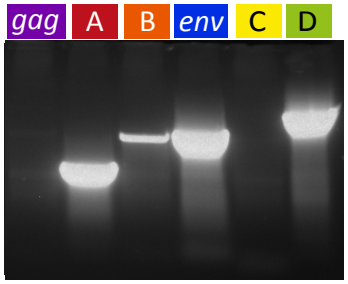
STEP 4: Overlapping PCRs to obtain full-length genome

- Perform 4 PCR reactions to obtain overlapping genomic amplicons for any wells containing functional *env*
- Set up reactions "A", "B", "C", and "D" the same as above for LTR/Env, except using different primers and extension times as indicated:

Reaction Primers and Extension Times
Rxn A: 275F + 3INOut, 5 mins.
Rxn B: 263F + 3AccOut, 6 mins.
Rxn C: 5INOut + BLInnerR, 6 mins. 30 secs.
Rxn D: 5AccOut + 280R, 5 mins.

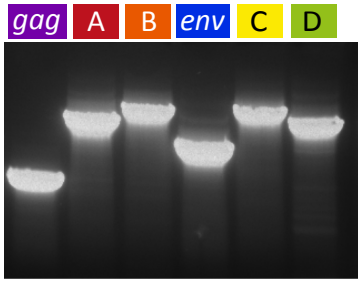
- Run gel to see whether deletions are present in genome:

Example of a 5' deletion



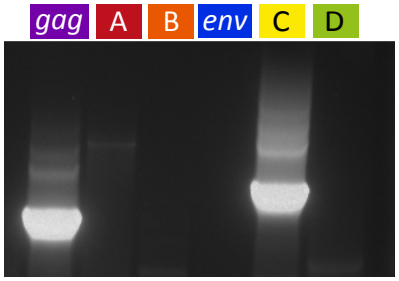
KMB6_E38P13C21-25

Example of an intact provirus



KMB10_E39P19D20-25

Example of a 3' deletion



KMB9_E2425P4-9B20-25

- IF NO DELETIONS ARE PRESENT: sequence the entire genome using WGS primers
- IF DELETIONS ARE PRESENT: sequence only the deletion junction

PCR primers

PCR	Length	Name	HXB2 position	Exact sequence	Extension time
SGA1 – Provirus PCR					
SGA1	9,064	BLOuterF	623 – 649	AAATCTCTAGCAGTGGCGCCCGAACAG	10 m
		BLOuterR	9,662 – 9,686	TGAGGGATCTCTAGTTACCAGAGTC	
Inner PCRs					
A	4,449	275F	646 – 666	ACAGGGACCTGAAAGCGAAAG	5 m
		3INOut	5,072 – 5,094	AATCCTCATCCTGTCTACTTGCC	
B	5,793	263F	651 – 672	GACCTGAAAGCGAAAGGGAAAC	6 m
		3AccOut	6,421 – 6,443	GGCATGTGTGGCCCARAYATTAT	
C	6,385	5INOut	3,248 – 3,270	ACTCCATCCTGATAAATGGACAG	6 m 30 s
		BLInnerR	9,604 – 9,632	GCACTCAAGGCAAGCTTTATTGAGGCTTA	
D	4,778	5AccOut	4,899 – 4,922	CGGGTTTATTACAGGGACARCARA	5 m
		280R	9,650 – 9,676	CTAGTTACCAGAGTCACACAACAGACG	
LTR		F9058	9,058 – 9,083	AGCCACTTTTTAAAAGAAAAGGGGGG	1 m
		LTR9556R	9,556 – 9,577	GAGCTCCCAGGCTCAGATCTGG	
env	2,841	5EnvIn	6201-6231	GAG AAA GAG CAG AAG ACA GTG GCA ATG AGA G	3 m 30 s
		3EnvIn	9007-9042	CTT GTA AGT CAT TGG TCT TAA AGG TAC CTG AGG TCT G	

REFERENCES

1. *World Health Organization HIV/AIDS*. Global Health Observatory (GHO) Data 2015; Available from: http://www.who.int/gho/en/.

2. *UNAIDS*. AIDSinfo 2015; Available from: http://aidsinfo.unaids.org.

3. Peterlin, B.M. and D. Trono, *Hide, shield and strike back: how HIV-infected cells avoid immune eradication.* Nat Rev Immunol, 2003. **3**(2): p. 97-107.

4. Sundquist, W.I. and H.G. Krausslich, *HIV-1 assembly, budding, and maturation.* Cold Spring Harb Perspect Med, 2012. **2**(7): p. a006924.

5. Wilen, C.B., J.C. Tilton, and R.W. Doms, *HIV: cell binding and entry.* Cold Spring Harb Perspect Med, 2012. **2**(8).

6. Craigie, R. and F.D. Bushman, *HIV DNA integration.* Cold Spring Harb Perspect Med, 2012. **2**(7): p. a006890.

7. Coffin, J.M., S.H. Hughes, and H.E. Varmus, *The Interactions of Retroviruses and their Hosts*, in *Retroviruses*, J.M. Coffin, S.H. Hughes, and H.E. Varmus, Editors. 1997: Cold Spring Harbor (NY).

8. Marini, B., et al., *Nuclear architecture dictates HIV-1 integration site selection.* Nature, 2015. **521**(7551): p. 227-31.

9. Liang, C. and M.A. Wainberg, *The role of Tat in HIV-1 replication: an activator and/or a suppressor?* AIDS Rev, 2002. **4**(1): p. 41-9.

10. Cullen, B.R., *Retroviruses as model systems for the study of nuclear RNA export pathways.* Virology, 1998. **249**(2): p. 203-10.

11. Freed, E.O., *HIV-1 assembly, release and maturation.* Nat Rev Microbiol, 2015. **13**(8): p. 484-96.

12. Blankson, J.N., D. Persaud, and R.F. Siliciano, *The challenge of viral reservoirs in HIV-1 infection.* Annu Rev Med, 2002. **53**: p. 557-93.

13. Maldarelli, F., et al., *HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells.* Science, 2014. **345**(6193): p. 179-83.

14. Siliciano, R.F. and W.C. Greene, *HIV latency.* Cold Spring Harb Perspect Med, 2011. **1**(1): p. a007096.

15. Ho, Y.C., et al., *Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure.* Cell, 2013. **155**(3): p. 540-51.

16. Archin, N.M., et al., *Eradicating HIV-1 infection: seeking to clear a persistent pathogen.* Nat Rev Microbiol, 2014. **12**(11): p. 750-64.

17. Harrington, P.R., et al., *Cross-sectional characterization of HIV-1 env compartmentalization in cerebrospinal fluid over the full disease course.* AIDS, 2009. **23**(8): p. 907-15.

18. Schnell, G., et al., *Compartmentalized human immunodeficiency virus type 1 originates from long-lived cells in some subjects with HIV-1-associated dementia.* PLoS Pathog, 2009. **5**(4): p. e1000395.

19. Schnell, G., et al., *Compartmentalization and clonal amplification of HIV-1 variants in the cerebrospinal fluid during primary infection.* J Virol, 2010. **84**(5): p. 2395-407.

20. Schnell, G., et al., *HIV-1 replication in the central nervous system occurs in two distinct cell types.* PLoS Pathog, 2011. **7**(10): p. e1002286.

21. Sturdevant, C.B., et al., *Central nervous system compartmentalization of HIV-1 subtype C variants early and late in infection in young children.* PLoS Pathog, 2012. **8**(12): p. e1003094.

22. Sturdevant, C.B., et al., *Compartmentalized replication of R5 T cell-tropic HIV-1 in the central nervous system early in the course of infection.* PLoS Pathog, 2015. **11**(3): p. e1004720.

23. Ho, D.D., et al., *Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection.* Nature, 1995. **373**(6510): p. 123-6.

24. Perelson, A.S., et al., *HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time.* Science, 1996. **271**(5255): p. 1582-6.

25. Simon, V. and D.D. Ho, *HIV-1 dynamics in vivo: implications for therapy.* Nat Rev Microbiol, 2003. **1**(3): p. 181-90.

26. Joseph, S.B., et al., *HIV-1 target cells in the CNS.* J Neurovirol, 2014.

27. Trepel, F., *Number and distribution of lymphocytes in man. A critical analysis.* Klin Wochenschr, 1974. **52**(11): p. 511-5.

28. Ganusov, V.V. and R.J. De Boer, *Do most lymphocytes in humans really reside in the gut?* Trends Immunol, 2007. **28**(12): p. 514-8.

29. Sathaliyawala, T., et al., *Distribution and compartmentalization of human circulating and tissue-resident memory T cell subsets.* Immunity, 2013. **38**(1): p. 187-97.

30. Whitney, J.B., et al., *Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys.* Nature, 2014. **512**(7512): p. 74-7.

31. Spudich, S., *HIV and neurocognitive dysfunction.* Curr HIV/AIDS Rep, 2013. **10**(3): p. 235-43.

32. Eden, A., et al., *HIV-1 viral escape in cerebrospinal fluid of subjects on suppressive antiretroviral treatment.* J Infect Dis, 2010. **202**(12): p. 1819-25.

33. Yilmaz, A., et al., *Persistent intrathecal immune activation in HIV-1-infected individuals on antiretroviral therapy.* J Acquir Immune Defic Syndr, 2008. **47**(2): p. 168-73.

34. Zayyad, Z. and S. Spudich, *Neuropathogenesis of HIV: from initial neuroinvasion to HIV-associated neurocognitive disorder (HAND).* Curr HIV/AIDS Rep, 2015. **12**(1): p. 16-24.

35. Price, R.W., et al., *Evolving character of chronic central nervous system HIV infection.* Semin Neurol, 2014. **34**(1): p. 7-13.

36. Ransohoff, R.M. and B. Engelhardt, *The anatomical and cellular basis of immune surveillance in the central nervous system.* Nat Rev Immunol, 2012. **12**(9): p. 623-35.

37. Engelhardt, B. and R.M. Ransohoff, *Capture, crawl, cross: the T cell code to breach the blood-brain barriers.* Trends Immunol, 2012. **33**(12): p. 579-89.

38. Obermeier, B., R. Daneman, and R.M. Ransohoff, *Development, maintenance and disruption of the blood-brain barrier.* Nat Med, 2013. **19**(12): p. 1584-96.

39. Spector, R., et al., *A balanced view of choroid plexus structure and function: Focus on adult humans.* Exp Neurol, 2015. **267**: p. 78-86.

40. Spudich, S.S., et al., *Cerebrospinal fluid HIV infection and pleocytosis: relation to systemic infection and antiretroviral treatment.* BMC Infect Dis, 2005. **5**: p. 98.

41. Wakim, L.M., A. Woodward-Davis, and M.J. Bevan, *Memory T cells persisting within the brain after local infection show functional adaptations to their tissue of residence.* Proc Natl Acad Sci U S A, 2010. **107**(42): p. 17872-9.

42. Watanabe, R., et al., *Human skin is protected by four functionally and phenotypically discrete populations of resident and recirculating memory T cells.* Sci Transl Med, 2015. **7**(279): p. 279ra39.

43. Arrildt, K.T., et al., *Phenotypic Correlates of HIV-1 Macrophage Tropism.* J Virol, 2015. **89**(22): p. 11294-311.

44. Burdo, T.H., et al., *Increased monocyte turnover from bone marrow correlates with severity of SIV encephalitis and CD163 levels in plasma.* PLoS Pathog, 2010. **6**(4): p. e1000842.

45. Matsuda, K., et al., *Laser capture microdissection assessment of virus compartmentalization in the central nervous systems of macaques infected with neurovirulent simian immunodeficiency virus.* J Virol, 2013. **87**(16): p. 8896-908.

46. Waisman, A., et al., *Homeostasis of Microglia in the Adult Brain: Review of Novel Microglia Depletion Systems.* Trends Immunol, 2015. **36**(10): p. 625-36.

47. Nedergaard, M., B. Ransom, and S.A. Goldman, *New roles for astrocytes: redefining the functional architecture of the brain.* Trends Neurosci, 2003. **26**(10): p. 523-30.

48. Churchill, M.J., et al., *Use of laser capture microdissection to detect integrated HIV-1 DNA in macrophages and astrocytes from autopsy brain tissues.* J Neurovirol, 2006. **12**(2): p. 146-52.

49. Liu, Y., et al., *CD4-independent infection of astrocytes by human immunodeficiency virus type 1: requirement for the human mannose receptor.* J Virol, 2004. **78**(8): p. 4120-33.

50. Joseph, S.B., et al., *Quantification of entry phenotypes of macrophage-tropic HIV-1 across a wide range of CD4 densities.* J Virol, 2014. **88**(4): p. 1858-69.

51. Zhuang, K., et al., *Emergence of CD4 independence envelopes and astrocyte infection in R5 simian-human immunodeficiency virus model of encephalitis.* J Virol, 2014. **88**(15): p. 8407-20.

52. Chung, W.S., et al., *Astrocytes mediate synapse elimination through MEGF10 and MERTK pathways.* Nature, 2013. **504**(7480): p. 394-400.

53. Monteiro de Almeida, S., et al., *Dynamics of monocyte chemoattractant protein type one (MCP-1) and HIV viral load in human cerebrospinal fluid and plasma.* J Neuroimmunol, 2005. **169**(1-2): p. 144-52.

54. Yilmaz, A., et al., *Treatment intensification has no effect on the HIV-1 central nervous system infection in patients on suppressive antiretroviral therapy.* J Acquir Immune Defic Syndr, 2010. **55**(5): p. 590-6.

55. Vester, B. and J. Wengel, *LNA (locked nucleic acid): high-affinity targeting of complementary RNA and DNA.* Biochemistry, 2004. **43**(42): p. 13233-41.

56. Folks, T., et al., *Characterization of a continuous T-cell line susceptible to the cytopathic effects of the acquired immunodeficiency syndrome (AIDS)-associated retrovirus.* Proc Natl Acad Sci U S A, 1985. **82**(13): p. 4539-43.

57. Folks, T.M., et al., *Biological and biochemical characterization of a cloned Leu-3- cell surviving infection with the acquired immune deficiency syndrome retrovirus.* J Exp Med, 1986. **164**(1): p. 280-90.

58. Gendelman, H.E., et al., *Molecular characterization of a polymerase mutant human immunodeficiency virus.* Virology, 1987. **160**(2): p. 323-9.

59. Melnikov, A., et al., *Hybrid selection for sequencing pathogen genomes from clinical samples.* Genome Biol, 2011. **12**(8): p. R73.

60. Kearney, M.F., et al., *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, 2015. **12**: p. 93.

61. Penton, P.K. and J.T. Blackard, *Analysis of HIV quasispecies suggests compartmentalization in the liver.* AIDS Res Hum Retroviruses, 2014. **30**(4): p. 394-402.

62. Ferretti, F., et al., *Cerebrospinal Fluid HIV Escape from Antiretroviral Therapy.* Curr HIV/AIDS Rep, 2015. **12**(2): p. 280-8.