Rebound human immunodeficiency virus type 1 (HIV-1) in cerebrospinal fluid after antiviral therapy interruption is dominated by clonal amplification of R5 T cell-tropic virus

Laura Kincer[1, 2], Sarah Beth Joseph[1,2,3,4], Blake M. Hauser[2,5], Maria M. Gilleece[2,6], Sabrina Sizemore[2], Clara Di Germanioe[7,8], Henrik Zetterberg[9,10,11,12,13], Dietmar Fuchs[14], Steven G. Deeks[15], Serena Spudich[16], Magnus Gisslen[17,18], Richard W. Price[19] and Ronald Swanstrom[2,4,20]

[1]Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

[2]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

[3]UNC HIV Cure Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

[4]UNC Center for AIDS Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA,

[5]Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, USA

[6]Biogen, Research Triangle Park, NC, USA

[7]Vitalant Research Institute, San Francisco, CA, USA

[8]Department of Laboratory Medicine, University of California San Franscisco, San Francisco, CA, USA

[9]Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Mölndal, Sweden

[10]Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden

[11]Department of Neurodegenerative Disease, UCL Institute of Neurology, Queen Square, London, UK

[12]UK Dementia Research Institute at UCL, London, UK

[13]Hong Kong Center for Neurodegenerative Diseases, Clear Water Bay, Hong Kong, China

[14]Division of Biological Chemistry, Biocenter, Innsbruck Medical University, Innsbruck, Austria

[15]Division of HIV, Infectious Diseases, and Global Medicine, Zuckerberg San Francisco General Hospital, University of California San Francisco, San Francisco, USA

[16]Department of Neurology, Yale University School of Medicine, New Haven, Connecticut, USA.

[17]Department of Infectious Diseases, Sahlgrenska Academy at the University of Gothenburg, Sweden,

[18]Region Västra Götaland, Sahlgrenska University Hospital, Department of Infectious Diseases, Gothenburg, Sweden

[19]Department of Neurology, University of California San Francisco, San Francisco, California, USA.

[20]Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

*To whom correspondence should be addressed: Ronald Swanstrom (risunc@med.unc.edu).

Abstract (150 words)

HIV-1 persists as a latent reservoir in people receiving suppressive antiretroviral therapy (ART). When ART is interrupted (treatment interruption/TI), rebound virus re-initiates systemic infection in the lymphoid system. During TI, HIV-1 is also detected in cerebrospinal fluid (CSF), although the source of this rebound virus is unknown. To investigate whether there is a distinct HIV-1 reservoir in the central nervous system (CNS), we compared rebound virus after TI in the blood and CSF of 11 participants. Peak rebound CSF viral loads vary and we show that high viral loads and the appearance of clonally amplified viral lineages in the CSF are correlated with the transient influx of white blood cells. We found no evidence of rebound macrophage-tropic virus in the CSF, even in one individual who had macrophage-tropic HIV-1 in the CSF pre-therapy. We propose a model in which R5 T cell-tropic virus is released from infected T cells that enter the CNS from the blood (or are resident in the CNS during therapy), with clonal amplification of infected T cells and virus replication occurring in the CNS during TI.

Introduction (3904 words)

In the absence of antiretroviral therapy (ART), HIV-1 undergoes continuous replication. The virus predominantly replicates in CD4+ T cells using CCR5 as a co-receptor and requires the high density of CD4 present on these cells for efficient entry (R5 T cell-tropic) (see[1,2] and reviewed by[3]). ART reduces the amount of detectable virus in the blood to low levels. Discontinuation of ART invariably results in the reappearance of virus[4-6] which, in most cases, returns to pre-therapy levels[7]. Thus, HIV-1 is largely maintained in a latent state during years of therapy but rapidly reemerges after treatment interruption (TI).

HIV-1 RNA can typically be detected in the blood one to four weeks after TI, but time to rebound is variable[7,8]. The rapid return of viremia after TI is typically due to viruses in multiple cells reactivating, a process that likely occurs throughout the time on therapy but is evident as increasing viremia when ART is discontinued[9-13]. There are cases where the rebound population was likely initiated by a single (or small number) of virus-producing cells, namely cases where the reservoir is extremely small, such as when ART is initiated shortly after birth in a case of vertical transmission[14], very soon after sexual transmission[5], or after an allogenic stem cell transplant[15]. Sequence analysis also suggests that rebound can be generated by reactivation of virus in cells in different compartments[16] and possibly different cell subsets [9].

Resting CD4+ T cells are thought to be the main source of the persisting long-lived reservoir[17] and rebound virus, but other HIV-infected cells in different states of activation can be detected in people on long-term ART[18,19]. Most of our understanding of long-lived,

latently infected CD4+ T cells comes from analysis of cells in the blood. It is now clear that CD4+ T cells can circulate between the blood compartment and tissues or differentiate into tissue-resident cells with limited migratory ability (reviewed by[20]). Studies of HIV-infected CD4+ T cells in the gut and blood suggest that cells in these compartments differ in the mechanisms enforcing latency[21,22].

The central nervous system (CNS) is isolated by the blood-brain barrier. The CNS contains only a small number of T cells[23-25] and its fluid component, cerebrospinal fluid (CSF), has low antibody concentrations (see[26] and reviewed by[25]). CD4+ T cells that traffic from the periphery into the CNS may release virus that is genetically similar to populations in the blood, giving rise to a viral load in CSF that is typically 1-10% the viral load compared to the blood[27]. Within the first 2 years of infection approximately 20% of untreated people have genetically distinct viral populations in their CSF or indirect evidence for viral replication in that compartment[27], and rates of compartmentalized replication in the CNS are likely much higher in individuals late in disease with severe neurocognitive impairment, including HIV-associated dementia (HAD)[28]. Long-term replication within the CNS compartment can lead to viral adaptation to replication in myeloid cells with low levels of surface CD4, such as macrophages and microglia (macrophage tropism). Decline in viral RNA in the CSF after ART initiation indicates that viral replication in this compartment ceases after ART initiation[29]; decline of viral load in the CSF can be rapid or slow, and this has been attributed to the loss of virus being produced from T cells in the former case and from myeloid cells (after viral adaptation) in the latter case [28]. Between 5-15% of people on ART have detectable levels of HIV-1 in their CSF with undetectable levels in their plasma (*i.e.*, asymptomatic CSF escape[30-35]),

and a subset of these cases represent ongoing viral replication in the CNS during ART thus representing an actively replicating reservoir[32,36]. The ability of T cells to cross the blood-brain barrier and the ability of virus to evolve to infect new cell types within the CNS raise the possibility that there are features of latent infection within the CNS that are distinct from those of the lymphoid system.

In this study, we examined rebound virus in the blood and CSF to determine if there is a distinct viral population that emerges in the CSF after TI that would provide indirect evidence of a persistent CNS-specific viral reservoir. The appearance and level of viral RNA in the CSF trailed that in the blood with lower levels except in those cases (approximately one-half) where there was an influx of white blood cells (WBCs/pleocytosis) into the CNS/CSF coincident with viral RNA levels in the CSF becoming equivalent to those in the blood. Overall, the viral populations in the CSF consisted of clonally amplified sequences that were similar to variants in the blood but were over-represented in the CSF. Furthermore, in several cases where pre-ART samples were available, viral populations present in the CSF before therapy were over-represented in the CSF during TI. We did not detect macrophage-tropic virus in CSF rebound virus, even in one instance where an M-tropic virus population was present in the CSF at the time of therapy initiation. Our data are consistent with virus being released from a limited number of clonally amplified cells within the CNS to initiate further amplification of the virus through viral replication in the inflowing CD4+ T cells. This raises the possibility that latently infected CD4+ T cells within the CNS could form a compartmentalized viral reservoir during therapy that can influence the nature of rebound virus. The influx of white blood cells during TI was associated with an increase in

inflammatory markers in the CSF, especially CXCL10, but this did not predict an increase in the neuron damage marker neurofilament light protein (NfL).

Results and Discussion

Cohort and specimens available to examine rebound virus in blood plasma and CSF. Eight men were identified as failing their ART regimen based on detectable plasma viral loads and were studied during a subsequent 'strategic' treatment interruption [37,38]. In the present study we examined archived samples from these previously described participants[37-39]. They ranged in age from 39 to 56 years old with CD4+ T cell counts between 143 and 834 cells/μl (Table 1). Blood plasma and CSF specimens were collected and viral loads were measured at 3 to 8 longitudinal timepoints near the time of TI. Figure 1 features blood and CSF viral loads, and CSF WBC counts for five participants where the CSF viral load approached or reached that of the blood after TI. Figure 2 includes data for the remaining three participants whose CSF viral load remained 10-fold or more lower than the blood viral load after TI.

Sequencing and phylogenetic analyses were performed at a subset of timepoints (see closed symbols in viral load graphs; Figures 1 and 2) with an effort made to analyze the rebounding viral populations in the CSF and blood at the earliest timepoint with sufficiently high CSF and blood viral loads. As in a previous study[40], the rebound viral load in the CSF trailed that in the blood by approximately 2 weeks. As a consequence, the earliest analyses of viral populations in the CSF were performed when the viral load was increasing (see Figures 1 and 2). Thus, our genetic analyses were able to characterize early populations that emerged in the CSF.

Compartmentalized rebound virus in CSF is influenced by the influx of white blood cells. Viral RNA was extracted from virus particles pelleted from blood plasma and CSF. cDNA synthesis included the use of a unique molecular identifier (UMI/Primer ID) to tag each individual RNA template[41,42]. The cDNA products were then amplified by PCR and sequenced using the Illumina MiSeq 300 base paired-end sequencing protocol to generate sequence information spanning most of the *env* V1-V3 amplicon. Phylogenetic analysis of blood-derived viral sequences revealed that there were multiple viral lineages present in the blood at all timepoints including those just after TI (Figures 1 and 2, center and right panels), an observation consistent with what was previously observed with samples from another cohort where the rebound virus was analyzed at even earlier time points relative to TI and at lower viral loads[12]. Viral sequences were often clustered as identical sequences, indicating rapid expansion of specific variants released from the latent reservoir, and clonal expansion shifted among viral lineages over time.

The nature of the rebound viral populations within the CSF fell into two distinct categories. In the absence of pleocytosis (*i.e.*, low WBC counts) the virus in the CSF largely overlapped the virus in the blood with the CSF viral load being significantly lower than the viral load in the blood. In the presence of elevated WBC counts the viral load dramatically increased in the CSF and the viral population became dominated by a small number of clonally expanded viral lineages.

The proportionately low viral load in the CSF relative to the blood with equilibrated populations mimics what is seen in chronic infection in the absence of CNS-specific replication and in the absence of pleocytosis[27]. We have attributed this low level of viremia

to the background migration of T cells into the CNS, some of which are infected and release virus similar to that in the blood. We believe these are infected T cells for two reasons. First, in people co-infected with HIV-1 and HCV, only HIV-1 is found in the CSF (in preparation). Second, the transporting infected cells are likely T cells rather than monocytes since the virus in the blood requires a high density of CD4 for efficient entry into cells (T cell-tropic) and is inefficient at infecting cells with a low density of CD4, such as monocytes[1,2,43].

The clonal amplification of viral lineages in the CSF linked to pleocytosis is an unexpected observation. There is close linkage between the high viral loads and the high level of expansion of these lineages. Since virus is always present in the CSF it cannot be that infectious virus is limiting for seeing this phenomenon. Also, the lineages in the CSF are disproportionately represented in the CSF versus the blood indicating a distinct phenomenon within the CNS; in fact the small amount of these lineages that appears in the blood could be the result of virus exiting the CNS as CSF fluid is returned to the blood. The influx of T cells into the CNS could be playing one of two roles (that are not mutually exclusive). First, the CD4+ T cells could be supporting the high level of viral replication leading to the high viral loads. Second, antigen-specific CD4+ T cells could define a limited number of infected cells that expand *in situ* to seed the CNS infection with homogeneous virus.

Our results are distinct from those reported by Gianella et al.[44] where they observed compartmentalized rebound virus. Their approach did not use UMIs as part of the sequencing strategy which can contribute to two artifacts in the sequencing. First, collapsing identical sequences into "haplotypes" loses information about clonal

amplification, which we observe as a dominant feature of the rebound virus in the CSF. Second, given that deep sequencing platforms are error-prone, the inability to create a consensus sequence for each template to reduce sequencing errors results in these errors being recorded as diversity in the viral population creating the appearance of compartmentalization. We believe the difference in the inclusion of a UMI in the sequencing approach is sufficient to explain our different observations.

Clonally amplified virus in the CSF/CNS during TI is undergoing active replication and has an entry phenotype that has been selected for replication in CD4+ T cells. Clonally amplified HIV-1 sequences are well known as part of the low level viremia (LLV) that persists in the blood of people on suppressive ART[45-51]. Given that LLV occurs in the face of antiviral drugs blocking viral replication, LLV is most likely produced solely by clonally expanded, HIV-infected cells rather than replication of genetically similar viruses. This is supported by integration site analyses identifying clonally expanding cells with viral DNA[51-54] and studies linking these cells to clonally amplified viral sequences in LLV[47,48,51]. Differences between viral dynamics in LLV and CSF rebound suggest that these clonally amplified populations may be produced by different mechanisms. Specifically, the fact that the viral load of LLV is typically several logs lower than that of virus rebounding in the CSF ($\sim 10^1$-10^2 RNA cp/ml in LLV blood vs. $\sim 10^4$-10^6 RNA cp/ml in rebounding CSF) makes it unlikely that clonally expanded cells alone can generate the high viral loads observed in the CSF during rebound.

In order to explore the mechanisms that generate clonal amplification of virus in the CSF after TI, we analyzed longer sequences from the viral population, specifically the

entire viral *env* gene both to assess whether there was any evidence of ongoing viral replication (in the form of small amounts of sequence diversity) and to determine the entry phenotype as indirect evidence of the cell type that had been harboring the virus. We examined *env* genes from three CSF specimens chosen at times when clonally amplified virus was apparent. The positions of the V1/V3 region of the *env* amplicons that were cloned are shown with stars in the figures as indicated: 4026 at 62 days post TI (Figure 1A); 6008 at 28 days post TI (Figure 1D); 6005 at 78 days post TI (Figure 2A)). A phylogenetic tree of the full length *env* gene sequences (left panel) for CSF samples from these three participants is shown in Figure 3 A-C. Again, the sequences designated with stars were those cloned for testing of the entry phenotype, targeting both sequences within the clonally amplified lineages and those from the more diverse portions of the tree. We first examined the level of sequence diversity within the clonally amplified sequences (middle panel). The synthesis of the cDNA product by reverse transcriptase in vitro is reported to have an error rate of one in 10,000 nucleotides by the manufacturer, and we previously examined this directly and found an error rate of one in 20,000 nucleotides in this PCR end-point dilution protocol[42]. While diversity in the amplified sequences is low, it is greater than this expected error rate (Figure 3, highlighter plots, middle panel). The accumulation of this low level diversity is most easily explained by the clonally amplified virus undergoing a brief period of viral replication to attain the high viral loads. Thus, while proliferation of an infected cell could play a role in the appearance of clonal lineages, the high level of virus in the CSF and the accumulation of some genetic diversity suggests ongoing viral replication.

HIV-1 typically replicates in CD4+ T cells which have a high density of CD4 and the virus targets these cells by requiring a high density for efficient entry, a phenotype termed T cell-tropic[43]. Under circumstances where CD4+ T cells are limiting, the virus adapts to enter cells with a low density of CD4, a phenotype termed macrophage tropism (reviewed by[3]). By comparing viral infectivity at high and low density CD4, an assay conveniently done using Affinofile cells where the level of CD4 can be regulated[55], it is possible to identify cells that have significant infectivity at a low density of CD4 (macrophage tropism) versus those that do not (T cell tropism). We used this assay to assess the entry phenotype encoded by the *env* gene clones from both the clonally amplified sequences and from the more diverse sequences (Figure 3 A-C, right panel). As can be seen, all of the *env* gene clones tested encoded Env proteins that had very poor infectivity at a low density of CD4. Thus we infer that the major populations of rebound virus in the CSF, including the variants in clonally amplified lineages, were previously selected for replication in CD4+ T cells. The simplest explanation of our results is that viruses that had been replicating in CD4+ T cells entered the latent reservoir in resting T cells and these were the source of the rebound virus in both the blood and the CNS/CSF.

Viral lineages over-represented in the CSF at therapy initiation can contribute to rebound after TI. We examined three participants from a second cohort who underwent unsupervised TI with less frequent sampling during the rebound period but with pre-therapy plasma and CSF specimens available. As can be seen in Figure 4, participants 5207 (Figure 4A), 5299 (Figure 4B), and 51126 (Figure 4C) all experienced rebound with

viral loads in the CSF approaching those in the blood, with low to moderate CSF WBC influx (left panels) and with largely clonal rebound virus in the CSF from one or several variants (middle/right panels, phylogenetic trees); these patterns are consistent with what was observed in the first cohort (Figures 1 and 2).

In this cohort we also had the opportunity to sequence the virus present in the blood and CSF at the time point just before the initiation of ART and preceding TI. For participant 5207 (Figure 4A) the clonal rebound virus in the blood and the CSF was from a lineage that pre-therapy had been over-represented in the CSF compared to the blood (lineage branch point circled blue). Similarly, for participant 5299 (Figure 4B) the largely clonal TI virus in the CSF also appears to be from a lineage (branch point circled blue) that was rare pre-therapy, but observed at a higher frequency in the CSF during TI. The simplest explanation for these two patterns is that virus most similar to lineages in the CNS pre-therapy gave rise to at least some of the rebound virus. Also, the predominant virus in the CSF can appear at a low level in the blood, consistent with transport of the predominant virus out of the CNS into the blood.

For participant 51126 (Figure 4C; diagnosed with HAD) samples were available at therapy initiation and rebound as well as at two intermediate time points. The rebound virus in the CSF (46 days post TI) represents several lineages within a branch (branch point circled blue) not detected prior to therapy (-2173 days from TI), while the rebound virus in the blood (46 days post TI) is from both this new lineage as well as from the lineage in the blood present prior to therapy. Viral RNA declined slowly after ART initiation and sequencing of residual viral populations present after 8 weeks of ART (-2074 days from TI) revealed the presence of the rebound lineage that was drug sensitive and

14

genetically distinct from the variants detected in the blood and CSF before ART (-2173 days from TI). By 14 weeks of ART (-2027 days from TI), ~50% of variants in this newly detected lineage were resistant to ralegravir (IN N155H), and this virus contributed to an increasing viral load in the blood. A subsequent change in therapy improved suppression in the blood and CSF. After treatment interruption the rebound virus in the blood represented drug sensitive variants from both the original blood lineage and the lineage that appeared with therapy initiation. Similarly, the rebound virus in the CSF contained only drug sensitive variants but only from the lineage revealed during therapy initiation. Overall, these phylogenetic patterns suggest that viral lineages over-represented in the CSF pre-ART (or revealed during viral load decay post-ART) can contribute to rebound of virus in the CSF and blood. It seems likely that inadequate regimen potency and penetration into the CNS allowed localized viral replication in the CNS when therapy was initiated.

The dynamics of the viral populations in participant 51126 has another feature (Figure 4C). The virus in the CSF pre-therapy represents a highly evolved macrophage-tropic virus population (Figure 4D), while the population revealed during therapy initiation is T cell-tropic, whose appearance is also associated with low level pleocytosis. When we sampled the rebound virus in the CSF to a high level (2200 templates sequenced) we failed to observe any virus from the M-tropic lineage (Supplemental Figure 1). In this one case we failed to detect any contribution of a fulminant CNS infection with a macrophage-tropic virus contributing to virus rebound after TI.

Assessment of CSF biomarkers of inflammation and neuronal damage during TI.

We examined the longitudinal specimens from four participants (6004, 6005, 6011, and 6012) for a set of inflammatory markers (Figure 5 A-D). Two of these participants had strong evidence of clonal amplification of viral lineages with WBCs migrating into the CNS (6004 and 6011, Figure 1) and two did not (6005 and 6012, Figure 2). Some of these data (including CSF NfL and neopterin) have been reported previously[56] but are included here with the other inflammatory biomarkers to allow a more comprehensive comparison. Our results suggest that the initial increase in the CSF viral load and the associated influx of WBCs are associated with several inflammatory markers. As can be seen in Figure 5, rising CXCL10/IP-10 levels were closely linked to the peaking of WBC counts during rebound as seen for participants 6004, 6011, and to a lesser extent for participant 6005 (Figure 5 A-C) where there was only a slight increase in CSF WBCs; this was not the case for participant 6012 (Figure 5D) where there was no influx of cells into the CSF. MCP-1 (CXCL2), TNFα, and MMP9 were more variable in their appearance, sometimes tracking with CXCL 10/IP-10 and in some cases lagging; neopterin also increased in concentration but always lagged CXCL10/IP-10 (Supplemental Fig. 2). By contrast, in the two cases where an increase in NfL was detected after TI (participants 6011 and 6012; Figure 5C and D) this increase occurred well after the initial rise in inflammatory markers. No other markers were consistently linked to the increase in NfL due to both their variable timing of appearance and their presence even in the absence of a rise in NfL (Supplementary Figure 2). Thus the neuronal damage indicated by the increased presence of NfL is not obviously linked to the inflammatory signals recorded in the CSF during TI.

Conclusions

Our results show that the most abundant portions of the viral populations that rebound in the CSF represent clonally amplifed virus with a T cell-tropic entry phenotype (requiring a high density of CD4 for efficient entry). This suggests that the major rebounding populations in both the blood and CNS compartments are likely produced from T cell reservoirs. The primary difference observed between rebounding populations in these two compartments was that rebound virus in the CSF is strongly influenced by transient pleocytosis within the CSF that often accompanies TI. The influx of WBCs into the CNS compartment during TI is closely linked to the clonal amplification of virus, elevated viral loads, and elevated inflammatory markers in the CSF. Our results did not identify any evidence of rebound virus emerging from HIV-infected myeloid cells in the CNS (even when documented to be present pre-therapy) as assessed by virus entry phenotype, which would have been evidence of a possible myeloid cell reservoir. We did observe T cell-tropic viral lineages enriched in the CSF prior to ART that were also observed in the rebounding virus population in the CSF and blood during TI. This raises the possibility that latently infected CD4+ T cells present in the CNS during ART contribute to rebound virus in the CSF and blood. An alternative model is that antigen-specific T cells enter the CNS compartment after TI and expand, with some of them being infected and thus able to release homogeneous virus. Understanding the nature of the putative antigen (locally produced HIV-1?) and the extent of clonality and the antigen specificity of the cells in the CSF during TI pleocytosis will help give further insight into virus rebound in the CNS compartment.

Acknowledgement

Methods

Study design. Participants were retrospectively identified at the University of California at San Francisco (UCSF) and at the University of Gothenburg (Gothenburg, Sweden) who had undergone serial lumbar punctures before and after ART interruption with accompanying blood sampling. The decision to interrupt treatment was made by each participant and their primary caregivers. Before and after ART TI, participants underwent site-specific protocols to evaluate responses in the CSF to ART modifications. Site-specific study protocols were approved by the IRB at UCSF and the University of Gothenburg, respectively. All participants provided informed consent.

Specimen collection and viral load analyses. Whole blood was collected in EDTA. Blood plasma was separated from whole blood by centrifugation, and the plasma was stored at - 80°C. CSF was collected in uncoated tubes and cells were removed by centrifugation. CSF supernatant was then stored at - 80°C. HIV-1 RNA levels in blood plasma and CSF were quantified with the Amplicor HIV-1 Monitor assay, version 1.5 (Roche Diagnostic Systems, Hoffman-La Roche, Basel, Switzerland) with a dynamic range down to 50 copies/mL and a detection limit of approximately 20 copies/mL. Each center used their standard clinical laboratory methods to measure CSF WBC counts and peripheral blood CD4+ T lymphocyte counts.

Biomarker analyses. Neopterin was measured by a commercially available enzyme-linked immunosorbent assay (ELISA) (BRAHMS GmbH, Hennigsdorf, Germany) with an upper normal reference value of ≤ 5.8 nmol/L in CSF[56]}. CSF neurofilament light protein

(NfL) concentration was measured by means of a sensitive ELISA (NF-light ELISA kit; UmanDiagnostics AB, Umeå, Sweden) as previously described[57,58]. Upper reference values are age dependent[59].

[Other biomarkers]

Sequence Analyses. Most sequences were generated by Illumina MiSeq deep sequencing with Primer ID[42] however, a subset of sequences was generated by single genome amplification (SGA) in order to facilitate analyses of viral entry. Deep sequencing was performed as follows: viral RNA was extracted from blood plasma or CSF using the QIAmp Viral RNA Mini Kit (Qiagen), and cDNA was generated using a single primer for the V1-V3 region of *env* or a pool of 4 primers for the V1-V3 region and 3 regions in *pro/pol* (partial RT, IN, and PR) for synthesis using reverse transcriptase. Each cDNA primer included a random 11 base Primer ID. cDNAs were amplified by PCR and sequenced using the Illumina MiSeq 300 base paired-end multiplex library preparation protocol. A template consensus sequence (TCS) was generated using the multiple reads for each Primer ID. TCSs were aligned using multiple sequence comparison by log-expression (MUSCLE)[60], and neighbor-joining phylogenetic trees were generated and visualized using FigTree v1.4.4.

Single genome amplification (SGA) was performed as previously described[1]. Briefly, viral RNA was extracted from blood plasma or CSF using the QIAmp Viral RNA Mini Kit (Qiagen) and cDNA was generated using an Oligo(dT)$_{20}$ primer (Invitrogen) using reverse transcriptase. End-point dilution PCR was used to generate full-length *env* genes that were bidirectionally sequenced by Sanger sequencing to yield 2X or greater

coverage. The *env* genes were aligned using MUSCLE and neighbor-joining phylogenetic trees were generated and visualized using FigTree (v1.4.4). Highlighter plots were generated using the Highlighter tool in the Los Alamos National Laboratory's HIV sequence database (hiv.lanl.gov).

The percent clonality of the viral population in the CSF was calculated for each participant as follows. First, the total number of CSF-derived *env* sequences (as TCSs) was counted. Next, all of the identical sequences were collapsed into a single representative sequence and the total number of unique CSF sequences was counted. The percent difference between the total number of CSF sequences and the total number of unique CSF sequences (i.e. after 'collapsing' the sequences) was used as the CSF percent population clonality.

HIV-1 entry phenotype assessment. Full-length *env* genes were cloned into the pcDNA3.1D/V5-His-TOPO expression vector (Invitrogen) using the pcDNA3.1 directional TOPO cloning kit (Invitrogen). We used our established protocol previously described[61] to determine the ability of the *env* genes to facilitate entry of pseudotyped reporter viruses into Affinofile cells expressing a low density of surface CD4, a marker of macrophage tropism.

Statistical Analyses. All statistical analyses were performed using R statistical software (version 3.6.3).

References

1. Ping, L.H.*, et al.* Comparison of viral Env proteins from acute and chronic infections with subtype C human immunodeficiency virus type 1 identifies differences in glycosylation and CCR5 utilization and suggests a new strategy for immunogen design. *J. Virol.* **87**, 7218-7233 (2013).

2. Parrish, N.F.*, et al.* Transmitted/founder and chronic subtype C HIV-1 use CD4 and CCR5 receptors with equal efficiency and are not inhibited by blocking the integrin α4β7. *PLoS Path.* **8**, e1002686 (2012).

3. Joseph, S.B. & Swanstrom, R. The evolution of HIV-1 entry phenotypes as a guide to changing target cells. *J. Leukoc. Biol.* **103**, 421-431 (2018).

4. Colby, D.J.*, et al.* Rapid HIV RNA rebound after antiretroviral treatment interruption in persons durably suppressed in Fiebig I acute HIV infection. *Nat. Med.* **24**, 923-926 (2018).

5. Henrich, T.J.*, et al.* HIV-1 persistence following extremely early initiation of antiretroviral therapy (ART) during acute HIV-1 infection: An observational study. *PLoS. Med.* **14**, e1002417 (2017).

6. Whitney, J.B.*, et al.* Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys. *Nature* **512**, 74-77 (2014).

7. Treasure, G.C.*, et al.* Relationship among viral load outcomes in HIV treatment interruption trials. *J.A.I.D.S.* **72**, 310-313 (2016).

8. Li, J.Z.*, et al.* The size of the expressed HIV reservoir predicts timing of viral rebound after treatment interruption. *AIDS* **30**, 343-353 (2016).

9.   Andrade, V.M.*, et al.* A minor population of macrophage-tropic HIV-1 variants is identified in recrudescing viremia following analytic treatment interruption. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 9981-9990 (2020).

10.   Lu, C.L.*, et al.* Relationship between intact HIV-1 proviruses in circulating CD4(+) T cells and rebound viruses emerging during treatment interruption. *Proc Natl Acad Sci U. S. A.* **115**, E11341-E11348 (2018).

11.   Bar, K.J.*, et al.* Effect of HIV antibody VRC01 on viral rebound after treatment interruption. *N. Engl. J. Med.* **375**, 2037-2050 (2016).

12.   Bednar, M.M.*, et al.* Diversity and tropism of HIV-1 rebound virus populations in plasma level after treatment discontinuation. *J. Infect. Dis.* **214**, 403-407 (2016).

13.   Salantes, D.B.*, et al.* HIV-1 latent reservoir size and diversity are stable following brief treatment interruption. *J. Clin. Invest.* **128**, 3102-3115 (2018).

14.   Luzuriaga, K.*, et al.* Viremic relapse after HIV-1 remission in a perinatally infected child. *N. Engl. J. Med.* **372**, 786-788 (2015).

15.   Henrich, T.J.*, et al.* Antiretroviral-free HIV-1 remission and viral rebound after allogeneic stem cell transplantation report of 2 cases. *Ann. Intern. Med.* **161**, 319-327 (2014).

16.   De Scheerder, M.A.*, et al.* HIV rebound Is predominantly fueled by genetically identical viral expansions from diverse reservoirs. *Cell Host Microbe* **26**, 347-358 (2019).

17.   Chun, T.W.*, et al.* In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency. *Nat. Med.* **1**, 1284-1290 (1995).

18. Hiener, B*., et al.* Identification of genetically intact HIV-1 proviruses in specific CD4(+) T cells from effectively treated participants. *Cell Reports* **21**, 813-822 (2017).

19. Chomont, N*., et al.* HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat. Med.* **15**, 893-900 (2009).

20. Shacklett, B.L., Ferre, A.L. & Kiniry, B.E. Defining T cell tissue residency in humans: Implications for HIV pathogenesis and vaccine design. *Current HHIV/AIDS Reports* **17**, 109-117 (2020).

21. Telwatte, S*., et al.* Mechanistic differences underlying HIV latency in the gut and blood contribute to differential responses to latency-reversing agents. *AIDS* **34**, 2013-2024 (2020).

22. Telwatte, S*., et al.* Gut and blood differ in constitutive blocks to HIV transcription, suggesting tissue-specific differences in the mechanisms that govern HIV latency. *PLoS Path.* **14**, e1007357 (2018).

23. Louveau, A., Harris, T.H. & Kipnis, J. Revisiting the mechanisms of CNS immune privilege. *Trends Immunol.* **36**, 569-577 (2015).

24. Engelhardt, B., Vajkoczy, P. & Weller, R.O. The movers and shapers in immune privilege of the CNS. *Nat. Immunol.* **18**, 123-131 (2017).

25. Iwasaki, A. Immune regulation of antibody access to neuronal tissues. *Trends Mol. Med.* **23**, 227-245 (2017).

26. Burbelo, P.D*., et al.* Anti-Human Immunodeficiency Virus antibodies in the cerebrospinal fluid: Evidence of early treatment impact on central nervous system reservoir? *J. Infect. Dis.* **217**, 1024-1032 (2018).

27. 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 Path.* **11**, e1004720 (2015).

28. Schnell, G., Joseph, S., Spudich, S., Price, R.W. & Swanstrom, R. HIV-1 replication in the central nervous system occurs in two distinct cell types. *PLoS Path.* **7**, e1002286 (2011).

29. Ellis, R.J.*, et al.* Cerebrospinal fluid HIV RNA originates from both local CNS and systemic sources. *Neurology* **54**, 927-936 (2000).

30. Perez-Valero, I.*, et al.* Cerebrospinal fluid viral escape in aviremic HIV-infected patients receiving antiretroviral therapy: prevalence, risk factors and neurocognitive effects. *AIDS* **33**, 475-481 (2019).

31. Trunfio, M.*, et al.* Cerebrospinal fluid HIV-1 escape according to different thresholds and underlying comorbidities: is it time to assess the definitions? *AIDS* **33**, 759-762 (2019).

32. Joseph, S.B.*, et al.* Human Immunodeficiency Virus Type 1 RNA detected in the central nervous system (CNS) after years of suppressive antiretroviral therapy can originate from a replicating CNS reservoir or clonally expanded cells. *Clin. Infect. Dis.* **69**, 1345-1352 (2019).

33. Anderson, A.M.*, et al.* Prevalence and correlates of persistent HIV-1 RNA in cerebrospinal fluid during antiretroviral therapy. *J. Infect. Dis.* **215**, 105-113 (2017).

34. Eden, A.*, et al.* HIV-1 viral escape in cerebrospinal fluid of subjects on suppressive antiretroviral treatment. *J. Infect. Dis.* **202**, 1819-1825 (2010).

35. Eden, A.*, et al.* Asymptomatic cerebrospinal fluid HIV-1 viral blips and viral escape during antiretroviral therapy: A longitudinal study. *J. Infect. Dis.* **214**, 1822-1825 (2016).

36. Joseph, S.B., Trunfio, M., Kincer, L.P., Calcagno, A. & Price, R.W. What can characterization of cerebrospinal fluid escape populations teach us about viral reservoirs in the central nervous system? *AIDS* **33**, S171-S179 (2019).

37. Price, R.W.*, et al.* Cerebrospinal fluid response to structured treatment interruption after virological failure. *AIDS* **15**, 1251-1259 (2001).

38. Price, R.W. & Deeks, S.G. Antiretroviral drug treatment interruption in human immunodeficiency virus-infected adults: Clinical and pathogenetic implications for the central nervous system. *J. Neurovirol.* **10 Suppl 1**, 44-51 (2004).

39. Gisslen, M., Rosengren, L., Hagberg, L., Deeks, S.G. & Price, R.W. Cerebrospinal fluid signs of neuronal damage after antiretroviral treatment interruption in HIV-1 infection. *AIDS Res. Ther.* **2**, 6 (2005).

40. de Almeida, S.N.*, et al.* Dynamics of monocyte chemoattractant protein type one (MCP-1) and HIV viral load in human cerebrospinal fluid and plasma. *J. Neuroimmun.* **169**, 144-152 (2005).

41. Zhou, S., Bednar, M.M., Sturdevant, C.B., Hauser, B.M. & Swanstrom, R. Deep sequencing of the HIV-1 *env* gene reveals discrete X4 lineages and linkage disequilibrium between X4 and R5 viruses in the V1/V2 and V3 variable regions. *J. Virol.* **90**, 7142-7158 (2016).

42. Zhou, S., Jones, C., Mieczkowski, P. & Swanstrom, R. Primer ID validates template sampling depth and greatly reduces the error rate of next-generation sequencing of HIV-1 denomic RNA populations. *J. Virol.* **89**, 8540-8555 (2015).

43. Joseph, S.B.*, et al.* Quantification of entry phenotypes of macrophage-tropic HIV-1 across a wide range of CD4 densities. *J. Virol.* **88**, 1858-1869 (2014).

44. Gianella, S.*, et al.* Compartmentalized HIV rebound in the central nervous system after interruption of antiretroviral therapy. *Virus Evol.* **2**, vew020 (2016).

45. Aamer, H.A.*, et al.* Cells producing residual viremia during antiretroviral treatment appear to contribute to rebound viremia following interruption of treatment. *PLoS Path.* **16**, e1008791 (2020).

46. Bailey, J.R.*, et al.* Residual human immunodeficiency virus type 1 viremia in some patients on antiretroviral therapy is dominated by a small number of invariant clones rarely found in circulating CD4(+) T cells. *J. Virol.* **80**, 6441-6457 (2006).

47. Cole, B.*, et al.* In-depth single-cell analysis of translation-competent HIV-1 reservoirs identifies cellular sources of plasma viremia. *Nat. Commun.* **12**, 3727 (2021).

48. Halvas, E.K.*, et al.* HIV-1 viremia not suppressible by antiretroviral therapy can originate from large T cell clones producing infectious virus. *J. Clin. Invest.* **130**, 5847-5857 (2020).

49. Rassler, S., Ramirez, R., Khoury, N., Skowron, G. & Sahu, G.K. Prolonged persistence of a novel replication-defective HIV-1 variant in plasma of a patient on suppressive therapy. *Virol. J.* **13**, 157 (2016).

50. Sahu, G.K., Sarria, J.C. & Cloyd, M.W. Recovery of replication-competent residual HIV-1 from plasma of a patient receiving prolonged, suppressive highly active antiretroviral therapy. *J. Virol.* **84**, 8348-8352 (2010).

51. Simonetti, F.R*., et al.* Clonally expanded CD4+ T cells can produce infectious HIV-1 in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 1883-1888 (2016).

52. Cohn, L.B*., et al.* HIV-1 integration landscape during latent and active infection. *Cell* **160**, 420-432 (2015).

53. Maldarelli, F*., et al.* Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science* **345**, 179-183 (2014).

54. Wang, Z*., et al.* Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E2575-E2584 (2018).

55. Johnston, S.H*., et al.* A quantitative affinity-profiling system that reveals distinct CD4/CCR5 usage patterns among human immunodeficiency virus type 1 and simian immunodeficiency virus strains. *J. Virol.* **83**, 11016-11026 (2009).

56. Hagberg, L*., et al.* Cerebrospinal fluid neopterin: an informative biomarker of central nervous system immune activation in HIV-1 infection. *AIDS Res. Ther.* **7**, 12 (2010).

57. Jessen Krut, J*., et al.* Biomarker evidence of axonal injury in neuroasymptomatic HIV-1 patients. *PLoS One* **9**, e88591 (2014).

58. Norgren, N., Rosengren, L. & Stigbrand, T. Elevated neurofilament levels in neurological diseases. *Brain Res.* **987**, 25-31 (2003).

59.   Yilmaz, A*., et al.* Neurofilament light chain protein as a marker of neuronal injury: review of its use in HIV-1 infection and reference values for HIV-negative controls. *Expert Rev. Mol. Diagn.* **17**, 761-770 (2017).

60.   Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792-1797 (2004).

61.   Joseph, S.B., Lee, B. & Swanstrom, R. Affinofile assay for identifying macrophage-tropic HIV-1. *Bio. Protoc.* **4**, e1184 (2014).

Table 1 Participant Characteristics

Study Site[a]	PID[b]	Year of TI	Age (years)	Sex	ARV Before TI	Years From HIV Diagnosis	CD4 Nadir (Cells/μl)	Suppressed (S) or Viremic (V) Before TI[c]
UCSF	4026	1998	51	M	NFV/d4T/ddi		118	V
	6004	1999	51	M	3TC/d4T/IDV	16	116	V
	6005	1999	53	M	3TC/d4T/IDV	12	3	V
	6006	1999	39	M	EFV/NFV/ ABC/Adefovir	17	69	V
	6008	2000	*35*	M	SQV/RTV	6	360	S
	6011	2001	43	M	EFV/ZDV/3TC	20	191	S
	6012	2001	51	M	EFV/d4T/3TC	23	199	S
	6013	2001	42	M	EFV/d4T/3TC	12	275	V
GOT	5207	2001	56	F	d4T/3TC/IDV/r	1	110	S
	5299	2006	55	M	TDF/3TC/AZV/r	3	120	V
	51126	2016	53	M	ABC/3TC/DRV/r	5	60	V
	38002	2014	35	F	ABC/3TC/LPV/r	10	110	S

[a]Study sites were the University of California San Francisco (UCSF) and the University of Gothenburg (GOT)

[b]Participant identifiers used in the text and figures

[c]Designation for detectable viral load in the blood

Figure Legends

Fig. 1. HIV-1 rebound populations in participants with high CSF viral load during rebound. Viral load timelines and phylogenetic trees are shown for 5 participants: 4026 (**a**), 6004 (**b**), 6006 (**c**), 6008 (**d**), and 6011 (**e**). The first panel depicts viral loads in the plasma (red squares) and the CSF (blue circles) as well as the CSF WBC count (black triangles) at indicated intervals (in days) after TI. Time points for which sequencing data are available are represented by filled symbols, and time points for which no sequencing data are available are represented by open symbols. Neighbor-joining phylogenetic trees were constructed using all the V1-V3 sequences (about 500 nucleotides in length) generated by Primer ID MiSeq deep sequencing. The trees are shown to the right of the viral load graphs with the time point indicated above each tree. Sequences from the plasma (in red) and CSF (in blue) from each specific time point are colored while the sequences from other time points are in gray. Black stars correspond to the closest positions in the tree of the full length HIV-1 *env* gene sequences that were used in pseudotyping assays for which entry phenotypes were assessed (Fig. 3).

Fig. 2. HIV-1 rebound populations in participants with low CSF viral loads during rebound. Timelines and phylogenetic trees are shown for 3 participants: 6005 (**a**), 6012 (**b**), and 6013 (**c**). All details of this figure are the same as described in the legend of Figure 1.

Fig. 3. Clonally amplified HIV-1 in CSF is replicating and is adapted to growth in T cells. Full length *env* gene amplicons were generated by end-point dilution PCR of cDNA synthesized using viral RNA in the CSF as template. The entry phenotype was assessed for three representative participants with clonal CSF-specific lineages in the viral rebound populations following TI at the indicated time: 4026 (**a**), 6005 (**b**), and 6008 (**c**). Panel 1 shows neighbor-joining phylogenetic trees of full-length *env* gene sequences. The *env* genes that were cloned and used to pseudotype a reporter virus are marked with a black star. Panel 2 to the right is a highlighter plot of the full-length *env* gene sequences depicting the low but detectable sequence diversity within these lineages in each participant. The *env* genes that were cloned and used to pseudotype a reporter virus are again marked with a black star. The third panel (far right) graphs the relative infectivity at low levels of CD4 for macrophage-tropic *env* gene controls used to pseudotype a reporter virus, T cell-tropic *env* gene controls, as well as the participant-derived *env* genes shown in the first panel. All *env* genes cloned from the CSF virus of the three participants were T cell-tropic when used to pseudotype the reporter virus and tested in the $CD4^{low}CCR5^{high}/CD4^{high}CCR5^{high}$ Affinofile cell entry assay.

Fig. 4. HIV-1 rebound populations in CSF and blood compared to viral populations in pre-therapy plasma and CSF. Timelines and phylogenetic trees are shown for 3 participants: 5207 (**a**), 5299 (**b**), and 51126 (**c**). The first panel depicts viral loads in the plasma (red circles) and the CSF (blue squares) as well as the CSF WBC count (black triangles) before ART, during ART, and after TI. Time points prior to TI are shown as negative days. Time points for which there is sequencing data available are represented

by filled symbols, and time points for which there is no sequencing data available are represented by open symbols. Neighbor-joining phylogenetic trees were constructed using up to 50 V1-V3 sequences (or all that were available if less) generated by Primer ID MiSeq deep sequencing at the indicated for each participant. Sequences from the plasma (in red) and CSF (in blue) from each specific time point are colored while the sequences the from other time points are gray.

Fig. 5. Inflammatory biomarkers measured during TI. Biomarkers are shown for four participants in a group of four panels per participant tracking the biomarkers over the course of the TI: 6004 (**a**), 6011 (**b**), 6005 (**c**), 6012 (**d**). The upper left panel in each group reproduces the viral load from blood (red squares) and CSF (blue circles), and CSF WBC counts (black triangles) shown in Figs. 1 and 2. The upper right panel in each group shows blood (red square) and CSF (blue circles) levels of neopterin, and the blood CD4+ T cell count (red circles). The lower left panel of each group shows the levels of IP10 (green circles) and MCP1 (purple squares). The lower right panel of each group shows the levels of IL-6 (light gray circles), TNFα (black circles), and MMP9 (orange squares).

Supplemental Figure 1. Deep sequencing of HIV-1 populations from participant 51126. Neighbor-joining hylogenetic tree containing a large sampling of template consensus sequences (TCS) from MiSeq/Primer ID sequencing from the pretherapy and TI timepoints. Sequences from the blood plasma pretherapy are shown in pink (1,734 TCS). Sequences from the CSF pretherapy are shown in light blue (1,734 TCS). Sequences from the blood plasma post TI are shown in red (2,233 TCS). Sequences from

the CSF post TI are shown in dark blue (2,233 TCS). Also included in gray are 50 sequences from the blood and CSF from each of the two intermediate decay timepoints. On the right the tree is expanded to show the portion where the macrophage-tropic virus lineage was found in the CSF pretherapy.

Supplemental Figure 2. Marker analysis during TI. Further analysis of markers was done for the participants shown in Figure 5. Two additional graphs are presented in vertical columns for each of the participants. In the top graph is shown the level of NfL as a function of time post TI/enrollment (purple circles). Also included is the QNPZ4 score test of neurocognition (tan diamonds). The vertical dashed line shows the peak viral load associated with pleocytosis. In the lower graph of the pair, the values of sCD163 are shown (light blue circles), and sCD14 (orange squares).