Deep sequencing of HIV-1 variants from paired plasma and cerebrospinal fluid during primary HIV infection

Arjet Gega¹, Michael J Kozal¹, Jennifer Chiarella¹, Evelyn Lee², Julia Peterson², Frederick M Hecht³, Teri Liegler³, Elizabeth P St John⁴, Birgitte B Simen⁴, Richard W Price², and Serena S Spudich⁵,*

¹Department of Medicine, School of Medicine, Yale University, New Haven, CT, USA
²Department of Neurology, University of California San Francisco, CA, USA
³Department of Medicine, University of California San Francisco, CA, USA
⁴454 Life Sciences (a Roche Company), Branford, CT, USA
⁵Department of Neurology, School of Medicine, Yale University, New Haven, CT, USA

Abstract

Background—Limited data exist comparing viral quasispecies between cerebrospinal fluid (CSF) and plasma compartments during primary HIV infection. Deep sequencing is a new method to examine the HIV plasma and CSF quasispecies.

Methods—in this pilot study, deep sequencing of protease (PR) and reverse transcriptase (RT) was performed in plasma and CSF from participants during primary HIV infection. Estimated mutational load was calculated by mutant variant frequency multiplied by HIV-RNA level.

Results—Paired plasma and CSF samples were studied from five antiretroviral therapy-naïve male participants with median 109 days post estimated transmission, age 32 years, CD4 cell count 580 cells/μL, HIV-RNA 5.18 log₁₀ copies/mL in plasma and 3.67 log₁₀ copies/mL in CSF. Plasma samples averaged 7,124 reads of PR and 2,448 reads of RT, whereas CSF samples averaged 7,082 and 2,792 reads, respectively. A distinct drug-resistance pattern with linked mutations present at significant levels (5–10%) was detected in one participant in CSF. Other low abundance variants (>0.2%) were detected in plasma and CSF of four out of five participants.

Conclusions—Deep sequencing of CSF HIV is technically possible with sufficient HIV-RNA levels. Differences between the quasispecies in the two compartments detected in one participant, which were present with a high mutational load in CSF at an estimated 3.6 months after HIV infection, suggest that early CNS compartmentalisation may be revealed by sensitive deep-
sequencing methods. The presence of distinct low abundance (<1%) resistance variants in plasma and CSF of three other subjects may be significant, but further investigation is needed.

**Keywords**
HIV-1; resistance mutations; cerebrospinal fluid; compartmentalisation; HIV-associated neurocognitive disorder; primary HIV infection

---

**Introduction**

HIV enters the central nervous system (CNS) in primary HIV infection (PHI), most likely via the migration of infected immune cells across the blood–brain barrier [1,2]. In untreated patients, HIV can be detected in cerebrospinal fluid (CSF) throughout the course of infection [3], which in part may reflect establishment of a local reservoir for HIV infection within the CNS [4]. Early establishment of CNS HIV infection may contribute to neurocognitive impairment and neurological complications observed in chronic HIV despite combination antiretroviral therapy (cART) [5].

In HIV-associated dementia (HAD), genetic compartmentalisation frequently exists between viral variants in the plasma and CSF [6,7]. Additionally, the presence of discordantly resistant variants in CSF may be associated with virological ‘escape’ within the CNS in patients on ART, which may be more common than previously recognised [8], and in rare cases associated with overt clinical neurological disease [9,10]. Although biologically important compartmentalisation has been thought to arise late in the course of HIV, recent recognition that differences in the viral quasispecies in the CNS and blood compartments may arise in early infection [11] may have implications for HIV treatment and eradication strategies.

We performed a pilot study employing deep-sequencing (DS) methods to assess HIV from paired CSF and plasma samples obtained during PHI to determine the feasibility of amplifying adequate sequences from CSF, to detect mutations within the low-level CSF viral variants, and to assess for differences in sequences between the CSF and blood compartments during early infection.

**Methods**

**Study design**

Samples were obtained in the context of an observational longitudinal neurological study of PHI in San Francisco, USA [1]. Paired baseline plasma and CSF samples from five cART-naïve participants who enrolled between January 2008 and July 2010 were selected based on specimen HIV-RNA >3,000 copies/mL to optimise success with DS. Confirmation of PHI (<1 year after transmission) and estimated timing of infection were ascertained as previously described [1]. Informed consent was obtained from participants and the study was approved by the involved Institutional Review Boards.
**Specimen sampling and laboratory studies**

CSF, blood and general medical and neurological assessments were obtained as previously described [1]. Background CSF clinical values and blood CD4+ and CD8+ T lymphocyte counts by flow cytometry were measured in the San Francisco General Hospital Clinical Laboratory on fresh samples. CSF was spun at 1,200 \( \times \) g for 10 min then supernatants and blood plasma were aliquoted and stored within 6 hours of collection in \( -70^\circ \)C freezers.

**Virological methods**

HIV-RNA levels were measured in previously frozen CSF and plasma using the ultrasensitive (20 copies/mL lower limit of detection) Amplicor HIV Monitor (version 1.5; Roche Molecular Diagnostic Systems, Branchburg, NJ) or the Abbott RealTime HIV-1 (Abbot Laboratories, Abbot Park, IL, USA) assays. Paired blood and CSF measurements were made in the same PCR run.

**Deep-sequencing methods**

Using SuperScriptII (Invitrogen, Waltham, MA, USA), first-strand cDNA was generated with three gene-specific oligonucleotides and treated with RNAse H (NEB). Subsequently, eight partly overlapping amplicons tagged with sequencing adaptors were generated from the cDNA by 40 cycles of PCR with FastStart HiFi Polymerase (Roche, Penzberg, Germany). The amplicons were purified with AMPure magnetic beads (Beckman Coulter, Brea, CA, USA) and quantified by use of PicoGreen (Invitrogen). After equimolar pooling of all 8 amplicons per sample, clonal amplification on beads (emulsion PCR) was performed using reagents that enabled sequencing in both the forward and reverse directions (kit II and kit III, respectively; Roche–454 Life Sciences, Branford, CT, USA). After bead isolation (breaking), enriched DNA-containing beads were counted on a Multisizer3 (Beckman Coulter). Thirty thousand beads per sample were prepared for ultra-deep sequencing and loaded in one region on a PicoTiter plate fitted with a 16-lane gasket. Sequencing was performed on a Genome Sequencer FLX (Roche–454 Life Sciences). The minimum accepted sequence coverage was 700 clonal reads for each position of interest, with average coverage for the sample set estimated at approximately 1,700 clonal reads per position. GS Amplicon Variant Analyzer (AVA) software was used to analyse the ultra-deep sequencing results. Amplicon nucleotide sequence reads were aligned to a consensus sequence generated from >6,000 sequences obtained from the Los Alamos HIV sequence database (www.hiv.lanl.gov/). The underlying flowgram signals were used in concert with each read’s base-called nucleotides to facilitate alignment accuracy. The direct application of the flowgram signals, rather than quality scores, allowed the consideration of any potentially undercalled nucleotides during the alignment process – bases for which no quality scores would exist. Interread similarity was also exploited for the alignment of any insertion bases relative to the reference sequence. Reads from both orientations and from overlapping amplicons were combined into a single alignment, and primer regions were automatically trimmed to avoid artefacts from the nucleotide content of the synthesised primers. AVA software was used with predefined codons conferring drug resistance based on the Stanford HIVdb Genotypic Resistance (http://hivdb.stanford.edu) mutations.
Analysis

Inclusion threshold was a variant frequency of at least 0.2%. Estimated mutational load (ML) was calculated by multiplying mutant variant frequency by HIV-RNA level, which is a method used frequently in similar studies [14].

Results

Subject characteristics are shown in Table 1. DS generated abundant reads in all samples. Plasma samples averaged 7,124 reads of PR and 2,448 reads of RT, whereas CSF samples averaged 7,082 and 2,792 reads, respectively. In CSF samples of subjects 9024, 9039 and 9044, and plasma of 9024, RT codons 88–148 were not covered.

Mutations associated with transmitted drug resistance were detected in low abundance in CSF and blood in multiple participants (see Table 1).

In subject 9058, sequencing of the RT region from both the CSF and plasma revealed NRTI and NNRTI drug resistance, while the PI region for plasma showed no significant drug resistance mutations. However, in the PI region for the CSF sample, D30N (9.8%, ML 451 copies/mL), M46I (8.1%, ML 374 copies/mL), V82A (0.27%, 12 copies/mL) and G73S (5.2%, ML 238 copies/mL) mutations were detected. In this sample, D30N, M46I and V82A mutations were found to be linked. Results of their interpretation through the Stanford HIVdb Genotypic Resistance Interpretation Algorithm are shown in Table 1. Numerous other non-drug-resistance mutations that deviated from consensus sequence were found (not shown). Figure 1 shows longitudinal plasma and CSF HIV-RNA levels of participants prior to and after cART, indicating viral suppression to the level of assay detection in blood and CSF in all four subjects who initiated cART during follow-up. Rate of HIV-RNA decay after cART could not be compared between study subjects due to variable sampling intervals, but was rapid in CSF even in the participant harbouring compartmentalised PI mutations in the CSF, and, based on the available data, was similar in the blood and CSF compartments in all four subjects.

Discussion

In this pilot study, we demonstrate the feasibility of investigating compartmentalisation between CSF- and blood-derived HIV through detection of low-abundance HIV drug-resistant viral variants using DS. To our knowledge there has only been one other report of the use of DS to examine the CSF HIV quasispecies [15]. In their report Tong et al. employed the same technique to examine the CSF HIV quasispecies of a limited number of subjects, without other confirmation.

In our study, this highly sensitive method for mutation detection revealed linked drug-resistance mutations present at significant levels (up to 9.77%) in CSF but which were not detected in plasma during PHI in one out of five subjects selected for our pilot study. DS is a recent addition to the armamentarium for detection of viral variants in HIV populations. Low-level variants can be detected through a high-throughput automated assay, as opposed to detection of predominant population sequences through Sanger sequencing. The aim of
this pilot study was to determine the feasibility of applying HIV DS methods to CSF. The high number of reads in the RT and PR genes confirms the success of application of DS to previously frozen CSF samples. Our findings may be influenced by the fact that, as samples were selected based on HIV-RNA levels, our participants had median CSF HIV-RNA levels higher than those in the overall observational study cohort ($3.67 \log_{10}$ copies/mL vs. $2.62 \log_{10}$ copies/mL). However, these data suggest that DS may be employed to examine CSF HIV in other relevant subject groups, provided CSF HIV-RNA levels are adequate for amplification.

The significance of early neuroinvasion for the course of CNS HIV infection remains unclear. One key question is when CNS compartmentalisation, measured through distinctions between CSF and plasma-derived HIV, is established during the course of disease. Compartmentalisation may result from independent replication of virus within the CNS, or a bottleneck or selection of HIV upon entry into the CNS. Previous studies have documented phenotypic or genotypic compartmentalisation between CNS-derived and systemic HIV populations in chronic infection [16, 17]. Specifically, CNS compartmentalisation of env correlates strongly with presence of dementia [7]. Several studies have specifically detected discordant RT and PR resistance genotypes in CSF and plasma using conventional genotyping methods, also noting increased genetic distance between compartments in the setting of HAD [6, 18]. Furthermore, drug-resistance mutations using conventional genotyping methods have been associated with virological failure in CSF in chronic infection [19]. The prospective association between low-level drug resistance and emergence of CSF HIV ‘escape’ has not been investigated. However, recent studies have demonstrated a high prevalence of asymptomatic CSF ‘escape’ [8], or more rarely, CSF escape with progressive CNS disease in patients on systemically suppressive cART [9, 10], which has been associated in some cases with CSF HIV resistance detected by standard genotyping.

As a pilot study to determine the feasibility of performing HIV DS on CSF specimens, the small sample size was one limitation of our study. However, the main objective of this study was to prove the concept and determine the feasibility of performing HIV DS on CSF specimens and provide a rationale for further investigation with a larger number of participants. In the next phase of the study we plan to quantify the cDNA template input for all our samples. Additionally, some of the distinct variants in Table 1 had very low MLs and frequencies near the limit of detection for DS [12, 13, 20]. Thus, some variant difference may have been due to the error rate in HIV, or enzymatic errors due to RT-PCR or DS. Furthermore, although DS can be performed on samples with a low copy number, the levels of mutations identified reflect mutations present in the RT-PCR amplicons and may or may not reflect the variant distribution in the sample [20]. The significance of these mutations at a level below 1% is unknown. Construction of phylogenetic trees based on the numerous variants was beyond the scope of this analysis. Despite these caveats, one of the participants (9058) had clear discordances in the variant distribution between the two compartments at a level between 5–10%. In this participant, three of these PR mutations in the CSF were linked.
Our DS results are consistent with prior findings in this PHI cohort indicating limited compartmentalisation of CSF HIV detected by standard methods [11]. The detection of frequent mutant variants by DS is consistent with a moderate prevalence of plasma-transmitted drug resistance by standard genotyping in the community from which the participants were enrolled [21]. Low-level plasma drug-resistance mutations may impact long-term systemic treatment outcomes, since cART-naïve participants with NNRTI-resistance mutations detected by DS more frequently experience virological failure after initiating NNRTI and NRTI regimens [22]. It is unclear what mutational load is required for these low-abundance drug-resistance mutations to have an impact on clinical outcome, and further investigation is necessary. The treatment outcome of subjects harbouring low-level drug-resistant variants may be successful if a boosted PI-based regimen with a higher resistance barrier is used [20,22].

The clinical significance of CSF low-prevalence drug-resistant variants in PHI is unknown. In very limited clinical follow-up, initiation of PI/r or integrase-based cART in four of our study subjects led to successful viral suppression in both compartments (Figure 1). Possibly, such variants have no impact on the eventual course of HIV infection provided treatment is initiated and suppresses high-level viral replication. All four subjects in our study who started cART appeared to have similar viral decay rates in both compartments, although sampling frequencies were not adequate to assess initial decay phases in both CSF and blood. Alternatively, low-abundance drug-resistant viral variants may be significant in the CSF if treatment is delayed allowing for further replication of resistant virus, or if regimens are selected with poor CNS penetration or targeting distinct resistance genotypes detected in plasma. Interestingly, the participant in this study (subject 9058) who had low-level resistant variants in CSF prior to cART was found in another study to have unique envelope sequences in CSF as compared to blood in the context of suppressive cART using single-genome sequencing, implying the possibility of persistence of compartmentalised HIV in the CNS on cART [23]. It is possible that low-level drug resistance detected in this analysis may have had an impact on this persistence. Future studies dedicated to investigating low-abundance HIV variants in CSF in different stages of infection, assessing changes over time in dual compartments prior to therapy, as well as observing the response to cART initiation will be likely to yield a new understanding of the importance of early detection and characterisation of low-level CSF variants in individuals infected with HIV.

Acknowledgments

We sincerely thank the participants who volunteered for these studies. We also thank the San Francisco General Hospital (SFGH)/University of California–San Francisco (UCSF) Clinical Research Center, staff at the ARI-UCSF Laboratory of Clinical Virology, and staff at the UCSF Options Project and Magnet for their invaluable help. This work was supported by the National Institutes of Health (grants R01 MH081772, K23 MH074466, and M01 RR000836), and was previously presented in part as a late-breaker platform presentation at the International AIDS Society Conference, Rome, Italy, July 2011.

References


Figure 1.
Longitudinal plasma and CSF HIV-RNA levels prior to and after cART. Longitudinal follow-up of four of five subjects after the baseline samples investigated for this study indicate successful suppression of HIV-RNA in response to ART in both compartments. HIV-RNA levels in (a) plasma and (b) CSF with respect to duration of infection (days post HIV transmission) for each subject enrolled in the study are shown. Initiation of cART is indicated with an arrow on the bottom panel for each of subjects 9055, 9044, 9025 and 9058, and cART regimens are indicated for each subject. Subject 9039 had only a single paired CSF and plasma study sample available. DRV: darunavir; EFV: efavirenz; FTC: emtricitabine; r: ritonavir; RAL, raltegravir; TDF: tenofovir
Table 1

Background clinical characteristics and reverse transcriptase (RT) and protease (PR) mutations detected in paired CSF and plasma specimens in PHI participants

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Duration of HIV infection (days)</th>
<th>CD4 cell count (cells/μL)</th>
<th>HIV-RNA plasma (log_{10} copies/mL)</th>
<th>HIV-RNA CSF (log_{10} copies/mL)</th>
<th>CSF WBC (cells/μL)</th>
<th>Plasma RT %, Mutant variants (copies/mL)</th>
<th>CSF RT %, Mutant variants (copies/mL)</th>
<th>Plasma PR %, Mutant variants (copies/mL)</th>
<th>CSF PR %, Mutant variants (copies/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9024</td>
<td>33</td>
<td>222</td>
<td>974</td>
<td>4.75</td>
<td>4.29</td>
<td>12</td>
<td>Wildtype</td>
<td>0.74 G67C (416)</td>
<td>Wildtype</td>
</tr>
<tr>
<td>9039</td>
<td>33</td>
<td>37</td>
<td>539</td>
<td>5.57</td>
<td>4.30</td>
<td>53</td>
<td>0.33 M184V (1226)</td>
<td>0.42 T227L (1569)</td>
<td>Wildtype</td>
</tr>
<tr>
<td>9044</td>
<td>25</td>
<td>128</td>
<td>533</td>
<td>4.82</td>
<td>3.49</td>
<td>20</td>
<td>0.32 V75A (211)</td>
<td>Wildtype</td>
<td>0.46 M46 (394)</td>
</tr>
<tr>
<td>9055</td>
<td>32</td>
<td>106</td>
<td>619</td>
<td>5.38</td>
<td>3.51</td>
<td>12</td>
<td>0.33 D67C (791)</td>
<td>Wildtype</td>
<td>0.39 V82A (172)</td>
</tr>
<tr>
<td>9058</td>
<td>29</td>
<td>110</td>
<td>237</td>
<td>5.18</td>
<td>3.67</td>
<td>4</td>
<td>0.34 G190E (573)</td>
<td>Wildtype</td>
<td>Wildtype</td>
</tr>
</tbody>
</table>

*J Virus Erad.* Author manuscript; available in PMC 2016 February 05.