An Example of Genetically Distinct HIV Type 1 Variants in Cerebrospinal Fluid and Plasma DuringSuppressive Therapy

Viktor Dahl,1 Magnus Gisslen,2 Lars Hagberg,2 Julia Peterson,2 Wei Shao,4 Serena Spudich,5 Richard W. Price,3 andSarah Palmer1,6,7
1Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet and Swedish Institute for Communicable Disease Control, Stockholm, and 2Department of Infectious Diseases, University of Gothenburg, Gothenburg, Sweden; 3Department of Neurology, University of California–San Francisco; 4Advanced Biomedical Computing Center, Science Application International Corporation–Frederick, Frederick National Laboratory for Cancer Research, Maryland; 5Department of Neurology, Yale University, New Haven, Connecticut and 6Centre for Virus Research, Westmead Millennium Institute, Westmead, and 7University of Sydney, Australia

We sequenced the genome of human immunodeficiency virus type 1 (HIV-1) recovered from 70 cerebrospinal fluid (CSF) specimens and 29 plasma samples and corresponding samples obtained before treatment initiation from 17 subjects receiving suppressive therapy. More CSF sequences than plasma sequences were hypermutants. We determined CSF sequences and plasma sequences in specimens obtained from 2 subjects after treatment initiation. In one subject, we found genetically distinct CSF and plasma sequences, indicating that they came from HIV-1 from 2 different compartments, one potentially the central nervous system, during suppressive therapy. In addition, there was little evidence of viral evolution in the CSF during therapy, suggesting that continuous virus replication is not the major cause of viral persistence in the central nervous system.

Keywords. HIV; cerebrospinal fluid; CSF; central nervous system; CNS; suppressive therapy; compartmentalization; hypermutants; reservoir.

Human immunodeficiency virus type 1 (HIV-1) infection of the central nervous system (CNS) occurs during primary infection, and HIV-1 virions are detectable almost immediately in the cerebrospinal fluid (CSF) [1, 2]. Over time, compartmentalization between CSF and plasma HIV-1 populations appears, likely related to independent replication within the CNS [3, 4]. Effective combination antiretroviral therapy (cART) can reduce HIV-1 RNA concentrations below the detection threshold for standard assays in both CSF and plasma [5].

By using a highly sensitive assay that detects <1 copy/mL, HIV-1 RNA can be found in both CSF and plasma specimens from some subjects even after years of therapy, and this persistent HIV-1 RNA is unaffected by treatment intensification [6–8]. It is unclear whether the HIV-1 RNA found in the CSF during suppressive therapy stems from virions produced within or outside the CNS. The origin of the virions in the CSF during suppressive therapy is of importance when developing HIV-1 eradication strategies. We investigated whether viral populations in CSF and plasma were compartmentalized during suppressive therapy.

METHODS AND MATERIALS

Clinical Specimens
Subjects were selected on the basis of the number and volume of available archived CSF and plasma sample pairs obtained at the same time point. The specimens had been collected in the context of studies of treatment response that were conducted at the University of California–San Francisco and at the Sahlgrenska Academy, University of Gothenburg, Sweden. Subjects had <40 copies of HIV-1 RNA/mL plasma for at least 1 year and longitudinal samples with <40 copies in CSF. The CNS penetration effectiveness (CPE) of the treatment regimens was calculated as proposed by Letendre et al [9]. We included samples from 17 subjects, of whom 5 started treatment during primary infection (median estimated time after infection, 61 days [range, 49–110 days]; Feibig stage V–VI [1]). All subjects were male and had had a plasma HIV-1 RNA load of <40 copies/mL for a median of 4 years (range, 1–12 years) and, because of less frequent sampling, a CSF HIV-1 RNA load of <40 copies/mL for 3 years (range, 1–10 years). All subjects provided informed consent. The ethics review boards at the University of California–San Francisco, Yale University, and in Stockholm and Gothenburg approved these studies.

Single-Genome Sequencing
We used single-genome sequencing to obtain HIV-1 RNA sequences from CSF and plasma samples [10, 11]. The
complementary DNA generated from CSF and plasma was serially diluted to a single copy. We amplified, sequenced, and analyzed the HIV-1 gene region encompassing p6, protease, and the first 900 nucleotides of reverse transcriptase. To increase the sensitivity of this assay, we pelleted virus from 5–10 mL of CSF and incorporated improved methods to extract viral RNA from samples with very low HIV-1 RNA concentrations [12].

Sequence Analysis
We assembled sequences from the raw sequencing data, using an in-house script written in Perl (available on request). Sequences containing >1 single-nucleotide polymorphism were removed because this indicated that >1 HIV-1 RNA molecule was amplified. Alignments were constructed using MAFFT, version 6.0 (available at: http://mafft.cbrc.jp/alignment/software/). Hypermutants were identified using “Hypermut” (available at: http://www.hiv.lanl.gov/content/sequence/HYPERMUT/hypermut.html) and were removed from the alignments. Alignments were tested for recombinants by using Splitstree4 (available at: http://www.splitstree.org/), and they were removed. For phylogenetic analysis of the HIV-1 populations found in the CSF and plasma, maximum likelihood trees were constructed using PhyML 3.0 (available at: http://www.atgc-montpellier.fr/phyml/). Generalized time reversible plus gamma was selected as substitution model after analyzing sequencing data from each subject with the find model tool on Los Alamos HIV Sequence Database (available at: http://www.hiv.lanl.gov/content/index). All sequences obtained were compared to each other and to standard laboratory viruses to ensure no contamination had occurred. Statistical support of the tree structures was obtained by 200 bootstrap replicates. Bootstrap values of >90% were considered significant. We calculated measurements of the HIV-1 genetic diversity (average pairwise distance) of HIV-1 populations located in the plasma and were removed from the alignments. Alignments were tested for compartmentalization by using the Slatkin-Maddison algorithm as implemented in HyPhy, version 2.1.2 (available at: http://hyphy.org/). To determine the degree of compartmentalization between the plasma- and CSF-derived sequences, 10,000 permutations were performed, with P values of <.05 considered significant.

Statistical Analyses
Average pairwise distances were compared using the Mann–Whitney U test. The proportion of hypermutants was compared using the Fisher exact test. All P values were 2 sided, with values of <.05 considered significant. All statistical analyses were performed using GraphPad Prism 5 software (GraphPad).

RESULTS
Samples Sequenced
Seventy CSF samples with viral RNA levels of <40 copies/mL were selected for analysis (median volume, 8 mL; Figure S1). From 15 of these 70 CSF samples we were able to generate a total of 21 amplicons. These 15 samples were collected from 8 of the 17 subjects in this study. Next we analyzed 29 plasma samples obtained during therapy (median volume, 7 mL) from these 8 subjects. From these 29 plasma samples we generated 14 amplicons. We also analyzed CSF and plasma samples collected before or immediately after therapy initiation (hereafter, “presuppression samples”) from these 8 subjects.

Population Structure and Diversity in Presuppression Samples
Presuppression samples were available for 7 of the 8 subjects. In presuppression samples from subjects 13, 7046, and 7047, we noticed ≥3 identical sequences in CSF samples, indicating clonal expansion. The presuppression CSF and plasma sequences were tested for compartmentalization (subject 13 was excluded from this analysis since the presuppression CSF and plasma samples analyzed for this subject were collected 1 year apart), with significant genetic compartmentalization detected between the CSF and plasma populations in 5 of 6 subjects (Table 1). The average HIV-1 genetic diversity of the presuppression CSF versus plasma samples for these 6 subjects, as measured by percentage average pairwise distance, was not significantly different (1.1% and 1.0% for plasma and CSF, respectively; \( P = 1 \)).

In subject 9037, the genetic diversity was 8-fold lower in CSF as compared to plasma, even though both samples were collected at the same time point. The lower genetic diversity in the CSF suggests that a longer time transpired for viral variants to become established and replicate within the CNS as compared to the peripheral blood in this subject. Alternatively, this subject may have been infected by several different viral variants, with only 1 establishing infection in the CNS. The overall genetic diversity of HIV-1 was lower in the 2 subjects (subjects 9037 and 9058) with presuppression CSF and plasma samples obtained during primary infection.

Population Structure and Diversity During Suppressive Therapy
Phylogenetic analyses and a test for hypermutants revealed that 8 of 21 sequences obtained from the CSF during suppressive therapy were hypermutants, compared with 0 of 13 sequences from plasma (\( P = .013 \)). However, 6 hypermutant CSF-derived sequences came from the same subject.

Phylogenetic trees were constructed for the 5 subjects (subjects 13, 7027, 7046, 7047, and 9058) from whom we obtained valid (ie, nonhypermutant and nonrecombinant) CSF sequences from at least 1 time point during suppressive therapy (Figure 1A–E). In subjects 7027 and 9058, from whom we generated sequences from both CSF and plasma specimens during suppressive therapy, the sequences from CSF and plasma specimens obtained during therapy did not cluster together. For subject 7027, CSF and plasma sequences from the same time point were separated by branches with significant bootstrap
support, suggesting that they came from cells infected by distinct viral populations. For the subjects for whom we had CSF sequences obtained during therapy at several time points, the sequences did not cluster together, and in some cases they were separated by branches with a significant bootstrap support. In general, the CSF sequences obtained during therapy were not found on longer branches, indicating that there was no or very little evolution due to continuous replication during suppressive therapy in this compartment.

Subject 13 had clonal expansion of plasma sequences obtained during therapy that contained drug resistance mutations (DRMs) against nucleoside reverse transcriptase inhibitors in CSF and plasma (M41L, M184V, L210W, and T215Y), as well as a DRM for nonnucleoside reverse transcriptase inhibitors (K103N) in plasma (Table 1). The presence of DRMs likely made the sequences obtained during therapy genetically distinct from the presuppression sequences (Figure 1 and Figure S2). In sequences from other subject samples, we rarely found DRMs (Table 1). The DRMs found were often related to current or previous drug exposure (Table S1). In some cases (subjects 13, 9037, and 9058), we observed DRMs in plasma but not in CSF (Table 1). We never detected DRMs in CSF that were not seen in plasma.

**DISCUSSION**

Our analysis of presuppression CSF and plasma samples revealed compartmentalization between CSF and plasma in all but 1 subject. This is similar to what other studies have shown [3, 4, 12, 13]. A possible explanation for this is independent HIV-1 replication within the CNS. We found clonal expansions among the presuppression CSF samples, which may be linked to R5 T-cell–tropic virus and associated with pleocytosis [13]. The 2 subjects (subjects 9037 and 9058) who initiated therapy during early infection had a lower genetic diversity in their presuppression samples, which has been described earlier and may be caused by a genetic bottleneck during HIV-1 transmission [14].

The analysis of HIV-1 sequences from the CSF and plasma specimens obtained during suppressive therapy revealed a larger percentage of hypermutants in the CSF (38%), compared with the plasma (0%; \( P = .013 \)). Hypermutants are induced by cellular cytidine deaminase, referred to as the APOBEC family, and result in replication-incompetent viruses [15]. The high level of HIV-1 hypermutants in the CSF may be related to a higher amount of APOBEC in the cells in the CNS. This finding should be interpreted with some caution because of the small number of subjects and because many of the hypermutants found in CSF came from one of the subjects. Furthermore, fewer plasma-derived sequences than CSF-derived sequences were obtained during therapy (14 vs 23).

In subjects 7027 and 9058, we generated sequences from both CSF and plasma specimens obtained during therapy that did not cluster together. In subject 7027, these were separated
by branches with significant bootstrap support. This is consistent with the hypothesis that these virions were produced by one population of cells within the CNS and another in the periphery. It could also be explained by a population of cells in the periphery being infected with a diverse viral population and producing virions of which some were transported into the CSF. In other subjects, sequences derived from specimens obtained at several time points during therapy did not cluster together and were heterogeneous, indicating that they had not been produced by a single long-lived cell. Sequences collected during therapy did not appear on longer branches, suggesting that they had not diversified further from the presuppression sequences through ongoing replication during suppressive therapy. This is consistent with the hypothesis that the virions found in CSF or plasma during therapy are produced by cells that had been infected before the initiation of cART and are reactivated. In subject 13, 2 clusters of clonal sequences were found in plasma obtained during suppressive therapy, probably because this subject had experienced many treatment failures, as sequence analysis showed several drug resistance mutations.

This study had limitations. Notably, despite the analysis of a large number of CSF samples obtained during suppressive therapy, we were unable to generate many sequences for phylogenetic analysis. This was attributable to the low concentration of HIV-1 RNA in CSF during suppressive therapy and to the unexpected high amount of hypermutants. This, in itself an interesting finding, prevented in-depth phylogenetic analysis.

Despite these limitations, we reach several conclusions concerning HIV-1 populations located in the CSF and plasma during suppressive therapy. A large fraction of the virus found in the CSF consists of replication-incompetent hypermutants. The structure of the phylogenetic trees was consistent with a model in which virions found in the CSF during suppressive therapy are produced by a separate population of cells located in the CNS. These cells were most likely infected before the initiation of therapy, because there was little evidence of ongoing

Figure 1. Phylogenetic trees. Maximum likelihood trees made in PhyML 3.0 under the general time reversible plus gamma nucleotide substitution model. Sequences from plasma are represented by squares, and sequences from cerebrospinal fluid (CSF) are represented by circles. The trees were rooted with a presuppression plasma sequence. Presuppression sequences are represented by black symbols for plasma and by grey symbols for CSF. Longitudinal sequences obtained during full suppression are represented by colored symbols, as follows: time point 1, red; time point 2, green; time point 3, blue; and time point 4, pink. Clonal amplification found in the presuppression samples is circled. Branch support was calculated by 200 bootstrap replicates. Branches with a bootstrap support of ≥90% are marked with a star. Clusters with >3 identical sequences (ie, clonal expansion in presuppression CSF samples) are circled. Data are for subjects 13 (A), 7027 (B), 7046 (C), 7047 (D), and 9058 (E).
replication taking place in the HIV-1 populations from the CNS during suppressive therapy.

**Supplementary Data**

**Supplementary materials** are available at The Journal of Infectious Diseases online (http://jid.oxfordjournals.org/). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copyedited. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

**Notes**

**Financial support.** This work was supported by the Foundation for AIDS Research (grant 108021-49-RFRL); the National Institutes of Health (grants R21 MH096619, R01 MH081772, UL1 RR024131/TR000004, and P30 AI027763); the Swedish Research Council (grant 2007-7092); and Sahlgrenska Academy, University of Gothenburg, Sweden (grant ALFGBG-11067).

**Potential conflicts of interest.** R. W. P. has been a consultant to Merck and received honoraria and travel reimbursement from AbbVie for meeting presentations. S. S. has received an honorarium and support for travel to a meeting from AbbVie. M. G. has received honoraria as speaker and/or advisor from Abbvie, Boehringer-Ingelheim, Bristol-Myers Squibb, Gilead Sciences, GlaxoSmithKline, Janssen-Cilag, Merck, and Pfizer. All other authors report no potential conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

**References**


