AIDS Alters the Commensal Plasma Virome

Linlin Li,a,b Xutao Deng,a,b Piyada Linsuwanon,c David Bangsberg,d Mwebesa Bosco Bwana,e Peter Hunt,f Jeffrey N. Martin,g Steven G. Deeks,f Eric Delwart,a,b Blood Systems Research Institute, San Francisco, California, USAa, Department of Laboratory Medicine, University of California, San Francisco, San Francisco, California, USAa, Chulalongkorn University, Bangkok, Thailandb; Harvard University, Boston, Massachusetts, USAc, Mbarara University, Mbarara, Uganda5; Positive Health Program, San Francisco General Hospital, San Francisco, California, USAf; Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California, USAa

We compared the plasma viromes of HIV-infected subjects with low versus high CD4+ T cell counts from the United States and Uganda by using deep sequencing and detected HIV, hepatitis C virus, hepatitis B virus, GB virus C, anellovirus, and human endogenous retrovirus (HERV) reads. An increase in the proportion of reads for anelloviruses, a family of highly prevalent and genetically diverse human viruses, was seen in subjects with AIDS from both countries. The proportion of endogenous human retrovirus reads was increased in AIDS subjects from Uganda but not the United States. Progression to AIDS is therefore associated with changes in the plasma concentration of commensal viruses.

A characteristic feature of HIV and pathogenic simian immunodeficiency virus (SIV) infections is the loss of CD4+ T lymphocytes, which coordinate the adaptive T and B cell responses against infections (1). The malfunction or failure of the immune system can result in an increased diversity of plasma and enteric viral and microbial communities (2, 3). A recent study showed that pathogenic SIV infection was associated with a significant expansion of the enteric virome, based on 454 pyrosequencing of the RNA and DNA viruses in feces. Pathogenic SIV infection resulted in more than a 10-fold increase in the number of virus reads in feces, including those of multiple enteric adenoviruses, caliciviruses, parvoviruses, picornaviruses, and polyomaviruses (2).

Another study that used Illumina sequencing to compare the plasma microbiome of HIV/AIDS subjects and healthy adults reported a different bacterial profile and a higher abundance of bacterial DNA in plasma of people with HIV/AIDS (3). Here we analyzed the DNA and RNA virome in the plasma of HIV-positive subjects with low (<20 cells/μl) versus high (>700 cells/μl) CD4+ T cell counts from the United States and Uganda. Stored frozen plasma samples from 35 subjects were selected based on high or low CD4+ cell counts. Viral nucleic acids were enriched by filtration and nuclease treatment and extracted from plasma as previously described (4), and libraries were constructed by using the ScriptSeq library preparation kit (Epigenic, Madison, WI). Nucleic acids from each plasma sample were labeled with a different primer bar code. The DNA library was sequenced using MiSeq (Illumina, San Diego, CA), which generated ~33 million paired-end 250-bp reads. The sequence reads were de-bar-coded, clonal reads were removed, and low-sequencing-quality tails and adaptors were trimmed, leaving ~6 million unique usable sequence reads, which were then de novo assembled separately for each sample by using SOAPdenovo2. The resulting contigs and singlets (>50 bp) were then translated and analyzed via a protein similarity search (using BLASTx), with an E value cutoff of 1 × 10−10. In this analysis, we focused on known human viruses, including human endogenous retrovirus (HERV). Other viral reads observed included sequences encoding the reverse transcriptase used to generate the cDNA libraries as well as sequences related to small circular and linear single-stranded DNA (ssDNA) genomes and iridoviruses that have been reported to contaminate the nucleic acid extraction columns that we used to purify nucleic acids (5). Fourteen sequence reads were also derived from a recently described rhabdovirus (6) that was being resequenced on the same MiSeq machine we used.

The subjects’ information, including gender, age, virus load, and standard antiretroviral therapy (ART) at time of collection, is shown in Table 1. The 12 U.S. subjects were 45 years old on average and 100% male, and 92% were on ART, while the 23 Uganda subjects were 32 years old on average and 35% male, and 43% were on ART.

The plasma viromes of HIV subjects contained viral sequences from HIV, GB virus C (GBV-C), hepatitis B virus (HBV), hepatitis C virus (HCV), anellovirus, and HERV (Table 1). HIV sequences were identified in 13 samples from all but one subject with HIV RNA viral loads of greater than 1 × 105 copies/ml (another subject’s viral load was not available), reflecting the approximate limit of detection for ssRNA viruses with the sample and nucleic acid processing method and the depth of Illumina sequencing used here. HIV subtyping was performed by using the automatic subtyping tool RIP (http://www.hiv.lanl.gov/content/sequence/RIP/RIP.html). The three U.S. positives were all subtype B, while five Ugandans carried subtype A1, two were subtype D, and three had too few reads to achieve subtyping. The subtyping results were consistent with the reported geographical distributions of HIV subtypes.

GBV-C belongs to the Flaviviridae family and has been classified into 7 genotypes with distinct geographical distributions (7). Coinfection with GBV-C is common among HIV-infected people and has been associated with slower disease progression, a lower mortality rate, and longer survival (8, 9). GBV-C was detected in 5/35 samples. Both GBV-C isolates from the U.S. subjects were genotype 2, which is mostly reported in Europe and the United States.

Received 30 May 2013 Accepted 25 July 2013
Published ahead of print 31 July 2013
Address correspondence to Eric Delwart, delwarte@medicine.ucsf.edu.
Copyright © 2013, American Society for Microbiology. All Rights Reserved. doi:10.1128/JVI.01839-13
States (10); the three positives among Ugandan samples had too few reads to be genotyped. One HCV-positive plasma sample (too few reads to be genotyped) and one HBV-positive (genotype A) plasma sample were detected in one U.S. and one Ugandan subject, respectively.

The Anelloviridae family consists of nonenveloped, circular, single-stranded DNA viruses that are genetically extremely diverse in humans, including multiple genera that each consist of multiple species and genotypes (11). Multiple mammalian species can also be chronically infected. The reported Torque teno virus prevalence in HIV subjects varies, possibly due to the use of different PCR primers for detection, but has been reported to be between 50% and 100% (12–16). Transmission is thought to initially occur from the United States group (P = 0.118), but the difference did not reach statistical significance. The results indicated that AIDS was associated with reduced control of anellovirus replication and higher viral loads (20).

HERVs are remnants of germ line retroviral integration. Approximately 3,900 full-length HERVs with two long terminal repeats have been estimated to reside in the human genome (21). HERVs are generally considered functionally defective, and in cases of autoimmune disorders, malignancies, and HIV infection, higher levels of HERV transcription and RNA levels in plasma have been observed (22, 23). HERV-K represents a more recently integrated and active group and the association of HERV-K with increased of HERV-K RNA but not HERV-H RNA has been reported in the plasma of HIV-infected subjects (26,27).

Here, HERV sequences were detected in 30/35 samples. The majority of HERV reads were matched to the class I ERV group, which includes HERV-H and HERV-W (28), while HERV-K (class II ERV group) only represented the class II ERV group (28), while HERV-K represents a more recently integrated and active group and the association of HERV-K with HERV.H RNA has been reported in the plasma of HIV-infected subjects (26,27).
bers were significantly higher in the Uganda low-CD4\(^+\) group than the Ugandan high-CD4\(^+\) group (HERV, \(P = 0.015\); HERV-K, \(P = 0.015\)). Unexpectedly the same phenomenon of higher HERV read numbers in low-CD4\(^+\) versus high-CD4\(^+\) subjects was not observed in the U.S. groups (HERV, \(P = 0.96\); HERV-K, \(P = 0.98\)). The gender distribution and ART prevalence differences between the U.S. and Ugandan subjects and their possible association with HERV expression were also analyzed. Gender was not associated with a difference in HERV (\(P = 0.63\)) or HERV-K levels (\(P = 0.63\)). Subjects on ART showed lower, although not statistically supported, levels of HERV (\(P = 0.13\)) and HERV-K (\(P = 0.13\)), possibly related to their higher CD4\(^+\) T cell counts (ART-treated average CD4\(^+\) count of 528 cells/\(\mu\)l, while those not ART treated had an average CD4\(^+\) count of 8 cells/\(\mu\)l).

In conclusion, our study provides an initial survey of viral sequences in the plasma of HIV-infected subjects. HIV, HCV, GBV-C, HBV, anelloviruses, and HERV were detected. It is possible that other viruses were undetected, but at least for ssRNA genomes these would likely be at lower viral loads than those ssRNA viruses detected here or have very low protein sequence identity to any already-sequenced viral proteins and thus be undetectable by BLASTx analysis. A higher level of anelloviruses was found in subjects with lower CD4\(^+\) counts from the U.S. group (\(P = 0.001\)) and the Ugandan group (\(P = 0.118\)). Increased anellovirus levels with lower CD4\(^+\) counts may be due to reduced immunological controls on replication of these viruses. The very high genetic diversity of anelloviruses (11) complicates studies of their possible pathogenicity (29), and no generally accepted link to disease has yet been identified for these very highly prevalent human infections (11). Besides possibly further stimulating the immune system (30), the health consequences of increased anellovirus replication during AIDS are unknown.

HERV sequences were detected at higher levels in Ugandan AIDS subjects than HIV-infected Ugandans with high CD4\(^+\) counts. The majority of HERV reads found in our study matched HERV-H and HERV-W reads. Upregulation of both HERV-H and HERV-W levels has been reported in subjects with rheumatoid arthritis or multiple sclerosis, but not in HIV-1-infected subjects (30). The majority of HERV reads found in our study matched HERV-H and HERV-W levels (\(P = 0.001\)) and the Ugandan group (\(P = 0.118\)). Increased anellovirus levels with lower CD4\(^+\) counts may be due to reduced immunological controls on replication of these viruses. The very high genetic diversity of anelloviruses (11) complicates studies of their possible pathogenicity (29), and no generally accepted link to disease has yet been identified for these very highly prevalent human infections (11). Besides possibly further stimulating the immune system (30), the health consequences of increased anellovirus replication during AIDS are unknown.

Recent studies in immunodeficient mice have shown the generation of replication-competent endogenous retrovirus through recombination, which results in higher rates of lymphoma (32, 33). Surprisingly, such ERV activation requires the presence of a full intestinal microbiota (32). It is interesting to speculate on a similar phenomenon occurring in AIDS patients and the possible influence of different microbiota in the U.S. and Ugandan populations.

**Nucleotide sequence accession number.** The sequence data from each plasma sample were deposited in the DNA Data Bank of Japan under accession number SRA091349.

**ACKNOWLEDGMENTS**

This work was supported by the Blood Systems Research Institute and NIH R01 HL105770 to E. Delwart and P30 AI27763, U01 CA066529, R24 AI067039, and R01 MH054097 to J. N. Martin.

We thank Zhanna Kaidarova for assistance with the statistical analysis.

**REFERENCES**


**Journal of Virology**


