HIV Surveillance and Parallel Virome Analysis Using Unbiased Next-Generation Sequencing


Infectious Diseases Research, Abbott Diagnostics; University of California San Francisco, Virus Discovery & Diagnostics Center; Université de Yaoundé, Cameroon

Abstract

Background: Next-generation sequencing (NGS) has the potential to revolutionize strategies for HIV surveillance, genome characterization, and minor variant analysis. Moreover, application of this technology in an unbiased manner offers the promise of simultaneous detection and characterization of viral co-infections.

Methodology: Thirty-five plasma specimens from HIV-1-infected Cameroonian blood donors were selected for analysis based on paired genome characterization (gap guk, pol and env IDR) using traditional Sanger sequencing. NGS (DNA libraries prepared using nucleic acid pre-treatment followed by RNA extraction and random-primer reverse transcription were multiplexed (4-5 libraries/run) and run on an Illumina HiSeq3000 platform. Sequence data was analyzed utilizing a custom virus discovery pipeline (SURPI; Naccache, et al., submitted) to identify and sort all viral sequences. Phylogenetic analysis was performed with PHYLIP and HIV-1 recombinants were characterized using SimPlot.

Results: HIV-1 related sequence ranged from 0.63% to 4.51% of total sequences yielding 99% coverage for 95% of samples. Complete genomes were assembled for 25 strains including 7 CRF11_cpx, 2 CRF13_cpx, 1 CRF18_cpx, CRF37_cpx, and 1 CRF37_cpx, circulating recombinant forms with relatively limited data available as well as 9 unique recombinant forms. Average depth of genome coverage ranged from 121- to 8972-fold. Bioinformatics analysis revealed the presence of numerous other viruses. Notably, GBV-C (HPgV) co-infection was detected in 35% of samples. Seven samples tested positive for GBV-C genomes with read numbers ranging from 0.21% to 5.57%. These GBV-C genomes formed a unique cluster in genotype 1 with Ghanaian sequences distinct from S. African and Ugandan reference strains. Additional detected viruses included hepatitis B virus, adeno-associated virus, hepatitis delta virus, human parvovirus B, human rhinovirus C (not expected in blood), and viral sequences with distant homology (<70% aa identity) to small circular DNA viruses (circoviruses, cycloviruses, and geminiviruses).

Conclusions: These data demonstrate the utility of unbiased NGS for HIV-1 surveillance, rapid viral-genome characterization, and for simultaneous screening of known and potential novel viruses. Moreover, the extent of HIV-1 genome coverage substantially increases confidence in phylogenetic classification of strains. The technology coupled with powerful bioinformatics represents an efficient and rapid method for the virome and deeper interrelationships between HIV-1 and other viromes.