High-throughput Analysis of Full-Length Proviral HIV-1 Genomes from PBMCs

Hiromi Imamichi1, Yan Guo2, Colleen Ludka2, Ellen E. Paxinos2, H. Clifford Lane1, Da Wei Huang3, Richard A. Lempicki3, and Michael P.S. Brown2

1NIAID, NIH, DHHS, Bethesda, MD; 2Pacific Biosciences, Menlo Park, CA; 3Leidos Biomedical Research, Inc., Frederick, MD

Background
HIV-1 proviruses in peripheral blood mononuclear cells (PBMCs) are felt to be an important reservoir of HIV-1 infection. However, accurate study of this pool is burdened by difficulties encountered in sequencing a full-length proviral genome, typically accomplished by assembling overlapping pieces and imputing the full genome.

SMRT Sequencing Methods

- Cryopreserved PBMCs collected from a total of nine HIV+ patients from 1995-2001 were used.
- High molecular weight DNA was extracted using Puregene DNA Isolation Kit (Qiagen).
- Genomic DNA was subjected to limiting dilution prior to amplification with a near full-length outer PCR followed by nested amplification with an inner PCR (Fig. 1).
- The PCR was performed with KAPA HiFi HotStart (KAPA Biosystems) with the initial denaturation at 95°C for 2 min, followed by 30 cycles of denaturation at 98°C for 20s, annealing at 67°C for 20s, and extension at 72°C for 2 min 20s, with the final extension at 72°C for 5 min.
- The PCR products were purified with the BluePippin system (Sage Science) using the 0.75%DF Marker S1 high-pass 4-10kb vs2 cassette definition with BP-Start: 6700 and BP-end: 10000, followed by purification with 1.0X AMPure XP beads (Beckman Coulter).
- Single molecules were sequenced as near-full-length amplicons directly from PCR products without shearing on a PacBio® RS II instrument using commercially available chemistries and protocols for SMRT® Sequencing (P4/C2, 180 min movies).

Validation of SMRT Sequencing Results

Figure 2. Samples were sequenced on the PacBio RS II with 180 min movies P4/C2 chemistry.

Sequencing of 8.9 kb HIV-1 Proviruses by PacBio

- It is possible to amplify near-full-length HIV-1 proviruses using the technique described in this study.
- Approximately 5-30% of amplifiable proviruses using this technique were found to be nearly full length.
- Of 24 nearly full-length proviruses obtained in this study, 11 were predicted to be replication-incompetent.

Conclusions

Acknowledgements. This work was funded in part through the intramural research programs of the National Institute of Allergy and Infectious Disease and the National Cancer Institute, National Institutes of Health, Bethesda, Maryland. This work was also supported in part through the intramural research program of the National Institute of Allergy and Infectious Diseases of the National Institutes of Health. Additional support was provided by the United States National Cancer Institute (Bethesda, Maryland) and the Leidos Biomedical Research, Inc., and in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. HHSN261200800001E.

Sequences of near-full-length HIV-1 Proviruses Derived from 9 HIV+ Patients

(1) Intact proviruses

(2) Defective Proviruses

(3) Intact and Defective Proviruses

Sequences (1.6-kb gag-pol region) of plasma-derived virus from the corresponding time points were compared. Nearly full-length provirus sequences were obtained by the PacBio SMRT Sequencing method. Plasma virus sequences were obtained by the Sanger method. Amino acid positions associated with drug resistance (in the gag cleavage site, protease and a part of reverse transcriptase) are shown in the figure.