Long-read sequencing assay allows accurate characterization of the HIV-1 reservoir

Laurens Lambrechts[1,2], Rita Verstraeten[1,2], Noah Bonine[1,2], Ytse Noppe[1], Marion Pardons[1], Basiel Cole[1,*], Linos Vandekerckhove[1,*]

[1]HIV Cure Research Center, Ghent University, 9000 Ghent, Belgium
[2]BioBix, Ghent University, 9000 Ghent, Belgium
*These authors contributed equally.

Introduction

The advent of near full-length (NFL) HIV-1 proviral genome sequencing greatly expanded our understanding of the quality of the viral reservoir, revealing that 2-5% of the persistent proviruses can be considered genome-intact in ART-treated individuals. However, current NFL assays are labor-intensive and based on costly principles of repeated PCRs at limiting dilution, limiting their scalability. We developed a long-read sequencing assay to characterize many proviral genomes in parallel from bulk DNA.

Methods

The long-read sequencing assay was performed on 15 chronic ART-suppressed individuals using a fixed input of 500 ng DNA extracted from peripheral blood CD4 T cells (reservoir sizes ranging from 321 to 6581 total HIV-1 DNA cp/Mio CD4 T cells). Individual proviruses were tagged with a different unique molecular identifier (UMI) at each end during a single reaction, followed by NFL PCR amplification and long-read sequencing on an Oxford Nanopore MinION (Figure 1B). UMI-based demultiplexing allowed for the construction of highly accurate consensus genomes, while excluding aberrant chimeric PCR artefacts. In addition, Full-Length Individual Provirus Sequencing (FLIPS) was performed on 2 individuals (Figure 1A). Data from both assays were compared through phylogenetic analyses.

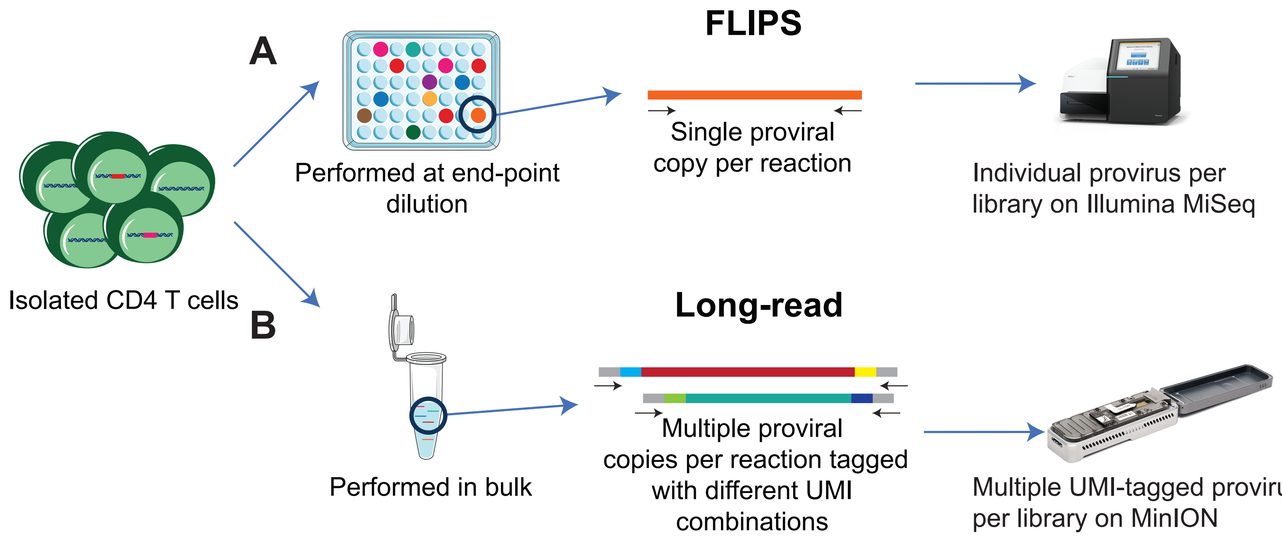


Figure 1: Schematic overview of the used assays. In A) the workflow of the gold standard full-length individual proviral sequencing (FLIPS) assay reyling on limiting dilution is depicted while B) shows the long-read assay relying on unique molecular identifiers (UMI).

Results

Assay can be used on diverse reservoir sizes

The long-read assay could successfully be applied on our cohort and yielded an average of 14 distinct HIV-1 proviruses per participant (range: 3-42). Note that some participants (eg. 1 and 11), despite having a large reservoir, yielded only 3 different proviruses which could be due to the presence of large infected cell clones sharing the same provirus.

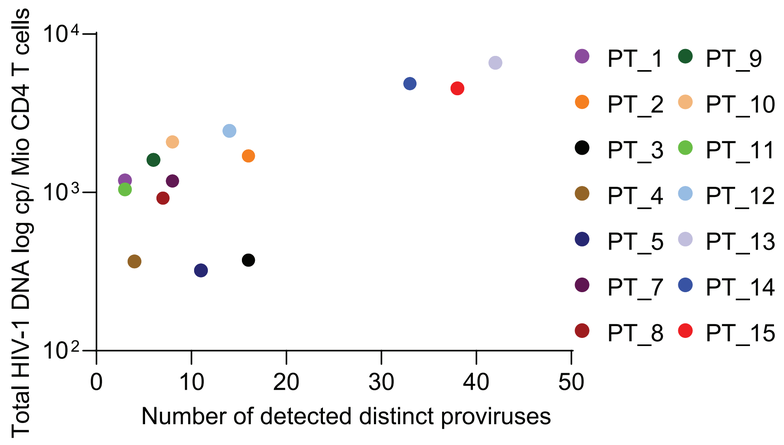


Figure 2: Reservoir size and number of distinct proviruses recoverd using the long-read assay for each participant.

Allows for HIV-1 genome classification

Across all participants, 213 distinct proviruses were retrieved of which 8% were considered as putatively intact.

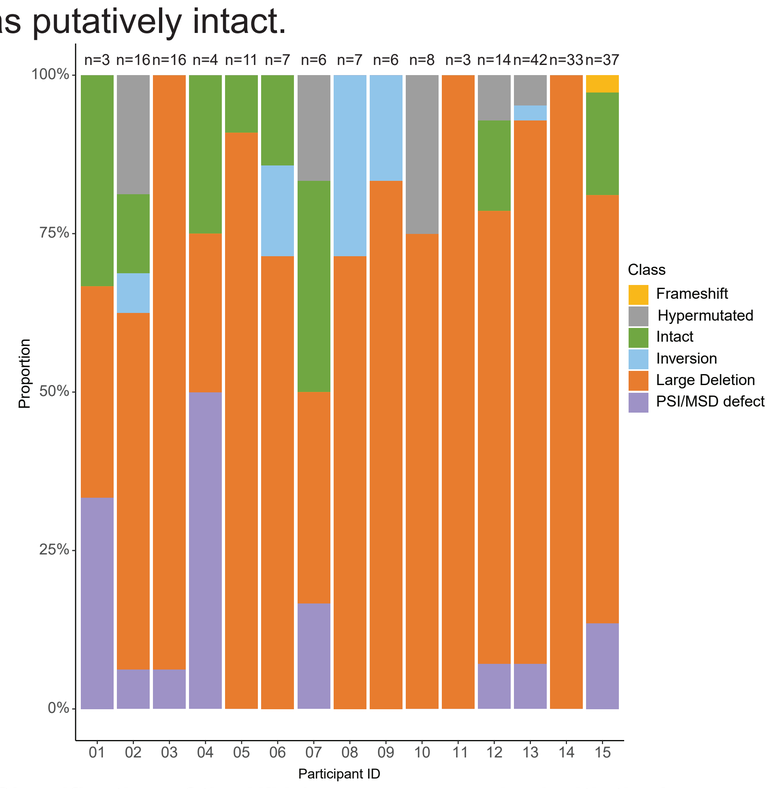


Figure 3: Classification of the HIV-1 genomes acquired with the long-read assay.

Overall agreement on observed proviral reservoir

In Participant 14, who displayed limited clonality (2/35 FLIPS genomes were clonal), only 1 overlapping provirus was found. While an overlap of 3 proviruses was observed in Participant 15 with higher clonality (33 of 36 FLIPS genomes were clonal). Comparing the 4 overlapping consensus proviral genomes to their matching FLIPS counterparts showed an average sequence accuracy of 99,97%.

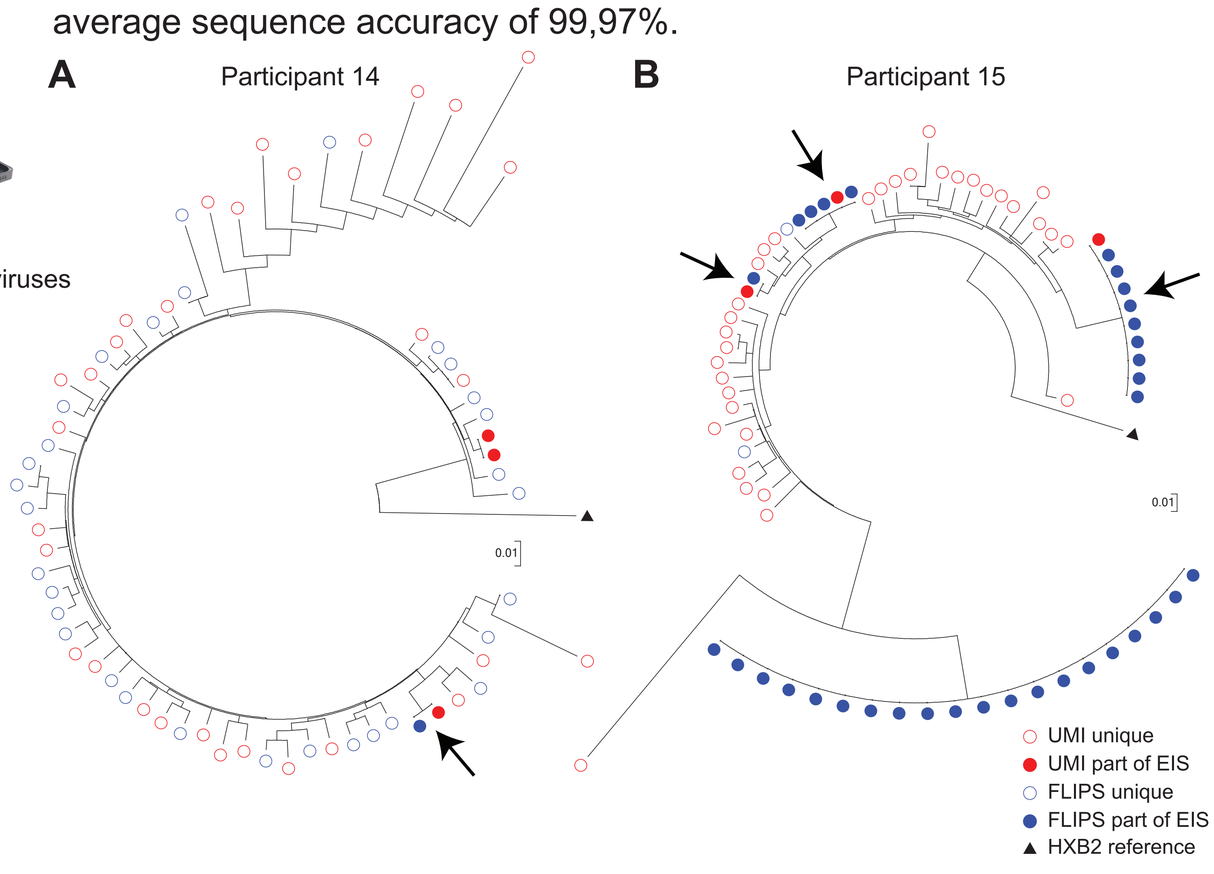
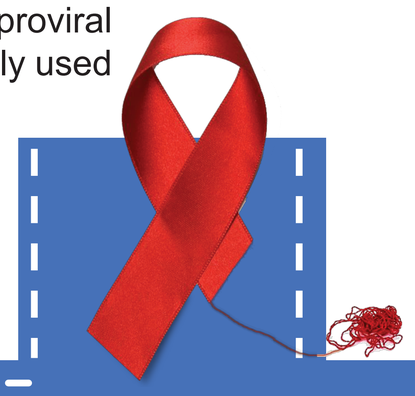


Figure 4: Phylogenetic analysis by Maximum Likelihood Trees. Trees included for Participant 14 and Participant 15. Genomes part of an expansion of identical sequences (EIS) are represented by filled circles. Arrows indicate matches of identical proviruses between the FLIPS and long-read assays.

Conclusions

The long-read assay offers a new NFL sequencing method which enables an accurate characterization of the proviral landscape in a more efficient way than current widely used NFL techniques.



Breach Symposium
30 November 2021

Fonds Wetenschappelijk Onderzoek Vlaanderen
Opening new horizons

HIV Cure Research Center
laurens.lambrechts@ugent.be