Evaluation of MinION nanopore platform for HIV whole coding regions sequencing

A.S. Speranskaya*, A.E. Lopatukhin, K. Khafizov, A.A. Ayginin, E.V. Korneenko, D.E. Kireev, G.A. Shipulin
Central Research Institute of Epidemiology, Moscow, Russia
* e-mail: hanna.s.939@gmail.com

Key words: sequencing technology, genomics, biomedicine, HIV, genome assembly

Motivation and Aim: Human immunodeficiency virus has a significant impact on economies worldwide due to high morbidity rates. The search for the new effective approaches for both rapid and non-expensive identification of drug resistance mutations is needed in order to reduce the HIV-induced mortality. Complete sequencing of the HIV genomes now is possible by various high-throughput second generation sequencing (HTS) methods. Both Illumina and Ion Torrent semiconductor platforms are widely used for the analysis of sequence diversity, recombinants determination and estimation of genetic distance between the HIV-1 quasi-species. Second generation HTS platforms has efficiency, accuracy and ability to detect the minor populations at the level as low as 1 %. However, the usage of these approaches is labour-consuming and therefore the analysis of clinical samples might be significantly delayed. Recently, the long read HTS sequencing of a third generation has become available from Oxford Nanopore (nanopore sequencing technology). The MinION is a low-cost portable device, which requires short time for the sample preparation and potentially can generate up to 10–20 Gb of DNA sequencing data per run. Oxford Nanopore sequencing technology is often described as highly error-prone (~8–40 % error rate). Our study aimed at estimating of the accuracy of the whole HIV coding region identification of HIV viral standards using the MinION sequencer.
Materials and Methods: For experiments we used the pNL4-3 (NIH AIDS Reagent, aidsreagent.org, USA), which are vectors containing the genome-wide replicatively and infectiously active HIV DNA. The four amplified fragments of the HIV genome (of 2,500–3,500 bp length each) were used for library preparation. MinION non-barcoded libraries were prepared using the Oxford Nanopore Genomic DNA Sequencing protocol provided by the manufacturer. The Rapid 1D Sequencing Kit was used to determine if the using of 1D reads would be adequate for HIV single nucleotide analysis. The canu software was used to assemble the reads into contigs. BWA program was then employed to map the contigs to the reference genome.
Results: The MinION produces data in real time so the analysis can be done during the experiment step by step. We estimated what number of reads should be produced for the best results. Finally, we made 700,000 more reads to achieve extra-large covering of pNL4-3 reference sequence (which is 8,883 bp long). The complete genome-wide HIV consensus sequences were assembled using MinION sequencing data. The platforms accuracies were estimated by comparing of the consensuses to the pNL4-3 reference sequence. We found that 98.95 % of the nucleotides analyzed positions were correctly identified when all obtained reads were used for assembling. Another 1.05 % were single nucleotide insertions (0.21 %), deletions (0.54 %) or ambiguous positions (0.29 %).
Conclusion: An efficient approach to obtain HIV genome sequencing data using the portable sequencer MinION was demonstrated.