proceedings

MDPI

Abstract

Metabarcoding-Like Approach for High Throughput Detection and Identification of Viral Nucleic Acids †

Alina Matsvay [1,2], Daniel Kiselev [1,3], Andrey Ayginin [1], Ivan Abramov [1], Vladimir Dedkov [4,5], German Shipulin [1] and Kamil Khafizov [1,*]

1 FSBI "Center of Strategical Planning" of the Federal Medical Biological Agency, 119435 Moscow, Russia; arity767@gmail.com (A.M.); neurolynx13@gmail.com (D.K.); ayginin75@gmail.com (A.A.); abriv2013@gmail.com (I.A.); shipgerman@gmail.com (G.S.)
2 Moscow Institute of Physics and Technology, National Research University, Phystech School of Biological and Medical Physics, 117303 Moscow, Russia
3 I.M. Sechenov First Moscow State Medical University, 119146 Moscow, Russia
4 Pasteur Institute, Federal Service on Consumers' Rights Protection and Human Well-Being Surveillance, 197101 Saint-Petersburg, Russia; vgdedkov@gmail.com
5 Martsinovsky Institute of Medical Parasitology, Tropical and Vector Borne Diseases, Sechenov First Moscow State Medical University, 119146 Moscow, Russia
* Correspondence: kkhafizov@gmail.com
† Presented at Viruses 2020—Novel Concepts in Virology, 5–7 February 2020.

Published: 22 July 2020

Abstract: Next generation sequencing (NGS) technologies have greatly enhanced our ability to identify new viral pathogens in various types of biological samples. This approach has led to the discovery of new viruses and has revealed hidden associations of viromes with many diseases. However, unlike the *16S* rRNA, which allows for bacterial detection by metabarcoding, the diversity and variability of viral genomes render the creation of universal oligonucleotides for targeting all known and novel viruses impossible. While whole-genome sequencing solves this problem, its efficiency is inadequate due to the high cost per sample and relatively low sensitivity. Furthermore, the existing approaches to designing oligonucleotides for targeted PCR enrichment are usually incomprehensive, being oriented at detecting a particular viral species or a genus based on the presumption of its presence in the sample. In this study, we developed a computational pipeline for designing genus-specific oligonucleotides that would simultaneously cover a multitude of known viruses from different taxonomic groups. This new tool was used to design an oligonucleotide panel for targeted enrichment of viral nucleic acids in different types of samples, and its applicability for the detection of multiple viral genera at once was demonstrated. Next, we created a custom protocol for NGS library preparation adapted to the new primer panel, which was tested together on a number of samples and proved highly efficient in pathogen detection and identification. Since a reliable algorithm for bioinformatic analysis is crucial for rapid classification of the sequences, in this work, we developed an NGS-based data analysis module and demonstrated its functionality both for detecting novel viruses and analyzing virome diversity. This work was supported by an RSF (Russian Science Foundation) grant (No. 17-74-20096).

Keywords: NGS; viruses; infections