Design of genus-specific primers panel for detection and identification of viral DNA in environmental samples using next-generation sequencing

ShT.18-1

K. Khafizov[I,II], A. Ayginin[I,II], E. Pimkina[I], A. Matsvay[I], A. Speranskaya[I], M. Safonova[I], E. Blinova[I], I. Artyushin[III], V. Dedkov[I], G. Shipulin[I]

[I]Central Research Scientific Institute of Epidemiology, Moscow, Russia, [II]Moscow Institute of Physics and Technology (State University), Dolgoprudny, Russia, [III]Moscow State University, Moscow, Russia

The advances in the next generation sequencing (NGS) technologies have significantly increased our ability to detect new viral pathogens and systematically determine the spectrum of those, which persist in various biological samples. This approach has led to the discovery of new viral pathogens and established the associations of viromes with many diseases. However, unlike the metagenomic studies using *16S* rRNA for bacterial detection, it is impossible to create universal oligonucleotides to target all known and novel viruses due to the viral genome diversity and variability, whereas whole-genome sequencing is still expensive and relatively low-sensitive in such purposes. The existing approaches for designing oligonucleotides for targeted enrichment are usually oriented at developing primers for the detection of a particular viral species or genera using PCR, but not the families or higher taxonomic orders. In this study, we developed a computational pipeline for designing the oligonucleotides that will cover a high number of known viruses belonging to the different taxonomic orders and also their novel variants. We subsequently designed a genus-specific oligonucleotide panel for targeted enrichment of viral nucleic acids in different samples and demonstrated the possibility of its application for virus detection in bird samples. Our panel has been tested using a number of collected samples and demonstrated superior efficiency in pathogen detection and identification. Since a reliable bioinformatic analysis pipeline for the rapid classification of the sequences is crucial, in this work an NGS-based data analysis module has been developed as well and its functionality has been demonstrated both for detecting novel viruses and analyzing the virome diversity. This approach resulted in a better viral genome coverage. This work was supported by the RSF grant (17-74-20096).