# Abstract

This work is part of a project focused on the study of the variability of human cytomegalovirus (HCMV) among clinical isolates with the aim to map the geographical distribution of HCMV genotypes, reveal the relationships between genotypes and the severity of HCMV-associated diseases, and identify regions in the HCMV genome with a potential for use as diagnostic and therapeutic targets. Attention was paid to the development of the methodology for the preparation of the material for next-generation sequencing (NGS) from HCMV clinical isolates and evaluation of the obtained sequencing data.

Blood and urine samples collected from hematopoietic stem cell transplantat recipients and congenitally infected children were analyzed. Samples suitable for NGS were sequenced by the Illumina platform and sequences were created by *de novo* assembly followed by mapping assembly.

Urine samples in comparison to blood samples had higher yield of material for NGS. Of the samples positive for HCMV DNA (7 of 50) after amplification in the cell cultures, only one sample had high purity of the viral DNA (98%) while six samples had purity of less than 7%. The sample containing 98% of the viral DNA was fully sequenced and the sequence was compared to the sequences of other clinical isolates from Belgium in 11 polymorphic regions. Only two isolates were identical in all 11 studied regions. The analysis of the whole genome sequences has shown high variability between HCMV clinical isolates. Futhermore, HCMV genes were found to cluster independently. To find out about the patterns of clinical isolates, a lot of samples have to be sequenced. This is possible thanks to the NGS methodology.



## Keywords

Human cytomegalovirus, whole genome, variability, next-generation sequencing, congenital infection, immunosuppresed patients, isolation, clinical isolates, genotypes