

Comparative Study of the Emerging COVID-19 Variants Using the msa R Package and ClustalW

Adrian VOICU* and Corina VERNIC

Department of Medical Informatics and Biostatistics, Victor Babeş University of Medicine and Pharmacy, E. Murgu Str., No. 2, 300041 Timișoara, Romania
E-mails: adrian.voicu@umft.ro; cvernic@umft.ro

* Author to whom correspondence should be addressed

Abstract

COVID-19 has become a major problem in the last 2 years. The occurrence of a large number of viral variants raises a legitimate question: are the vaccines against Corona Virus still efficient? To answer that question, we used the R package msa and clustalX to visualize and compare the alignment of the spike, envelope, and membrane proteins for the most important emerging SARS-Cov2 variants with the first Wuhan-1 variant, used for developing all the vaccines in use. As emerging variants, we considered the Delta (B.1.617.2) and B.1.1.7 (or the Alpha variant). Both are considered as variants of concern. The viral proteins were downloaded from the GISAID site in FASTA file format and then imported and analyzed in R, using the msa (multiple sequence alignment) R packages. The results of the alignment and the consensus sequences were visualized using ClustalX, as an extension for msa package. The results revealed a good alignment and consensus for all the proteins in the study. Based on these facts, we can conclude that the vaccines in use are still efficient against the circulating variants of Sars-Cov2. But other new variants, that suffered more mutations at the level of spike protein, such as Lambda (C.37), can be a reason of concern and must be kept under continuous surveillance.

Keywords: SARS-Cov2; Emerging variants; msa R package; ClustalX