

ABSTRACT

The continuing COVID-19 pandemic has brought to light how crucial it is to comprehend the SARS-CoV-2 virus and its proteins in order to develop efficient treatment and diagnostic approaches. However, there are difficulties with storage, processing, and analysis because of the enormous number of SARS-CoV-2 protein sequence data. In order to analyze SARS-CoV-2 data, this study examines the use of UNIQmin, a protein sequence reduction program. Also, to improve the performance of UNIQmin. This study gives light on the program's ability to improve the interpretation of SARS-CoV-2 proteomics data by examining how well UNIQmin reduces SARS-CoV-2 protein sequences while maintaining important information and its efficiency in producing the result. Using the gathered SARS-CoV-2 sequences from December 2022, The result shows a 98.6% percentage of reduced sequences from the initial nr dataset. The efficiency of the program also has shown improvement with multithreading implementation.

Keywords: COVID-19, Peptidome analysis, SARS-CoV-2, UNIQmin