JAMSA
Journal of Asian Medical Students' Association

PAPER COMPETITIONS

Development of immunoinformatics program for amino acid sequence alignment and linear B cell epitope prediction based on spike proteins of SARS-CoV-2

*Preenapun Saokhieo[1], Thanida Auttapracha[1], Ornpailin Wanichthanaolan[1], Anupap Sangjun[1], Piyapong Khumrin[1], Kanokporn Polyiam[2], Yaowaluck Maprang Roshorm[2]

[1] Faculty of Medicine Chaing Mai University

[2] Division of Biotechnology, School of Bioresources and Technology, King Mongkut's University

10.52629/jamsa.vi

Address for Correspondence:

Faculty of Medicine Chaing Mai University 110 Faculty of Medicine, CMU, Inthawarorot Rd., Sri Phum, Muang, Chiang Mai 50200

gampreenapun@gmail.com

Introduction & Objective

Corona Virus Disease 2019 (COVID-19) is caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). With an advance in immunoinformatics, B-cell epitope prediction combining multiple tools has been previously shown to be an effective and accurate method to identify potential B-cell epitopes from the SARS-CoV-2 S protein sequence which could contribute to a better design of COVID-19 vaccine. However, extracting the potential peptides from each prediction tool as well as finalizing the putative B-cell epitopes are still time-consuming and labor-intensive. In this work, we thus aim to develop a program for a rapid and accurate prediction of B-cell epitopes by using SARS-CoV-2 as a test sequence.

Methods

To fill the gaps in B-cell epitope prediction combining multiple prediction tools, a new program can assemble the prediction results obtained from distinct tools and then generate potential B-cell epitopes. To run the program, a full-length sequence of the target protein and the prediction results obtained from B-cell epitope and coil structure prediction using BepiPred tool, and from predictions for accessibility, hydrophilicity, and antigenicity using the methods of Emini, Parker, and Kolaskar & Tongaonkar, respectively. The program runs in 2 phases. In phase I, peptides are extracted from each tool based on the input thresholds. In phase II, the program creates the final putative B-cell epitopes depending on the users' criteria.

Result

The program accuracy was analyzed by comparing the epitope identification outcomes of the wild-type SARS-CoV-2 spike protein which was previously identified. The program accuracy is 64%. The discrepancy may be due to personal experience considered when selecting the B-cell epitopes so as to obtain the most possible epitopes for further applications.

Conclusion

We provide a program with a solution addressing problems associated with B-cell epitope prediction using multiple tools and criteria for epitope identification, thus enabling a more rapid and accurate B-cell epitope prediction.

Keywords

B cell epitope prediction, epitope prediction, spike proteins, SARS-CoV-2, immunoinformatics