A Universal Vaccine for COVID-19

Course Project Proposal

02-718 Computational Medicine, CMU Fall 2020

Alex Guo (xg2), Ryan Nelson (ryann), Yizhou Wang (yizhouwa)

Introduction

COVID-19 is a disease caused by a SARS-CoV-2 coronavirus. Discovered in 2019 [1], it is responsible for the largest pandemic in over a century. Infection causes a wide range of signs and symptoms, from no symptoms to severe (i.e. multi-organ failure, sepsis, acute respiratory distress syndrome) [2]. One of the most challenging aspects of the virus is its high infection rate (R0 between 2.2 and 5.7) [3]. Despite government lockdowns of businesses and the encouraged use of personal protective equipment, the virus continues to move through the population. A vaccine is needed in order to stop the spread of infection and protect individuals from getting sick.

The goal of vaccination is to cause the body's immune system to recognize and protect against a natural pathogen [4]. MHC, or major histocompatibility complex, is a surface protein present in human cells [5]. Consisting of two main classes, I and II, MHC proteins display antigens on the surface of cells. MHC proteins vary among individuals but are consistent in specific populations [6]. Epitopes are the parts of an antigen which are recognized by antibodies, T cells, and B cells in the immune system [7]. Each MHC variant has a certain affinity with which it binds to various epitopes. Different viruses have different epitopes [8], and vaccines can lead to the production of antibodies which recognize epitopes of a given virus. Peptide vaccines do this by presenting the body with protein sequences similar to naturally occurring proteins from a pathogen [9].

Viruses tend to evolve, and changes to their epitopes can limit the long-term effectiveness of a vaccine. A universal vaccine is meant to be effective against all strains of the virus and to confer long-term immunity against virus mutation [10]. Additionally, vaccines are designed to include multiple epitopes to ensure a wide coverage across different populations.

Our goal is to identify peptide sequences which could be used to develop a universal peptide vaccine for COVID-19. The SARS-CoV-2 coronavirus has spike proteins present on its surface, and these proteins are the mechanism by which the virus infects host cells [11]. The production of antibodies against epitopes on these proteins may prevent the virus from replicating inside its host, and antibodies against highly conserved regions means that this protection can remain effective even as the virus mutates [12].

Methods

The first step is to find the conserved regions across the COVID-19 spike protein sequences. An alignment across the spike protein sequences needs to be created using BLAST or MUSCLE, ClustalW etc., then a sequence variability analysis can be performed to find the most conserved regions. We would then perform variability analysis and select the consensus sequences ourselves for better compatibility with downstream analysis.

The second step is to identify potential epitopes in the conserved regions that bind well with some MHCs. To do this, we take our conserved regions and subdivide them into peptides of lengths 8-17 to match the potential MHCs (length 8-11 for MHC I, length 13-17 for MHC II). We would be able to find MHC sequences that might bind well using the provided IEDB MHC binding prediction tool.

Finally, we choose one or more potential epitopes from the conserved regions such that the MHCs they bind to combined have a population coverage that meets or exceeds a chosen threshold. (A particular epitope might bind well to MHCs of one or few populations. By including multiple epitopes in the vaccine, we can cover many or all populations.) The coverage of MHC could be determined on the IEDB population coverage inquiry tool.

Evaluation

We will evaluate the vaccine in two ways. First, we will see what percentage of the population are covered by the epitopes that we include in our vaccine, based on the MHC allele frequencies of different populations or the IEDB database. Our goal is for our population coverage to achieve 85%.

And then, we compare the epitopes we find for the vaccine with the known epitopes of COVID-19. Those known epitopes can be found in the IEDB database. We also found a published study of vaccine development for COVID-19 by Yarmarkovich et. al.[13], and we will compare the epitopes we found with what they found to see how much our results agree.

Datasets

- Spike Protein Sequences: https://sites.google.com/view/sarswars/home (SARS Wars course)
- Binding Predictions: http://tools.iedb.org/main/datasets/
- Epitope Population Coverage: http://tools.iedb.org/population/
- IEDB database for known COVID-19 epitopes: https://www.iedb.org/home_v3.php

Potential Pitfalls & Strategies to Overcome Them

One possible pitfall is that we may fail to identify enough conserved regions, or not enough conserved regions that can bind well to MHCs, to achieve a high enough coverage in all populations. If that happens, we may need to widen our search and find more spike protein sequences than those in the SARS Wars course datasets. Or we can try to find a vaccine that has high population coverage for certain regions, instead of a universal vaccine.

References

[1] Coronavirus disease 2019. (2020, October 14). Retrieved October 17, 2020, from https://en.wikipedia.org/wiki/Coronavirus_disease_2019

[2] Similarities and Differences between Flu and COVID-19. (2020, October 06). Retrieved October 17, 2020, from https://www.cdc.gov/flu/symptoms/flu-vs-covid19.htm

[3] Sanche, S., Lin, Y., Xu, C., Romero-Severson, E., Hengartner, N., & Ke, R. (2020). High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2. Emerging Infectious Diseases, 26(7), 1470-1477. https://dx.doi.org/10.3201/eid2607.200282

[4] Vaccine. (2020, October 15). Retrieved October 17, 2020, from https://en.wikipedia.org/wiki/Vaccine

[5] Major histocompatibility complex. (2020, October 10). Retrieved October 17, 2020, from https://en.wikipedia.org/wiki/Major_histocompatibility_complex

[6] Janeway CA Jr, Travers P, Walport M, et al. Immunobiology: The Immune System in Health and Disease. 5th edition. New York: Garland Science; 2001. The major histocompatibility complex and its functions. Available from: https://www.ncbi.nlm.nih.gov/books/NBK27156/

[7] Epitope. (2020, October 11). Retrieved October 17, 2020, from https://en.wikipedia.org/wiki/Epitope

[8] Langmead, C. (2020). “Vaccines”. SARS Wars, Episode III: Vaccines & Antivirals. Video retrieved October 17, 2020, from https://sites.google.com/view/sarswars/home/episode-iii

[9] Peptide vaccine. (2020, October 05). Retrieved October 17, 2020, from https://en.wikipedia.org/wiki/Peptide_vaccine

[10] Universal flu vaccine. (2020, September 12). Retrieved October 17, 2020, from https://en.wikipedia.org/wiki/Universal_flu_vaccine

[11] Huang, Y., Yang, C., Xu, X. et al. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. Acta Pharmacol Sin 41, 1141–1149 (2020). https://doi.org/10.1038/s41401-020-0485-4

[12] Padron-Regalado E. (2020). Vaccines for SARS-CoV-2: Lessons from Other Coronavirus Strains. Infectious diseases and therapy, 9(2), 1–20. Advance online publication. https://doi.org/10.1007/s40121-020-00300-x

[13] Yarmarkovich, M., Warrington, J. M., Farrel, A., & Maris, J. M. (2020). Identification of SARS-CoV-2 Vaccine Epitopes Predicted to Induce Long-Term Population-Scale Immunity. *Cell Reports Medicine, 1*(3), 100036. doi:10.1016/j.xcrm.2020.100036