Madeleine King

19 March 2021

Introduction:

Severe acute respiratory syndrome coronavirus, 2 (SARS-CoV-2) is a newly emerged virus that has caused a tragic pandemic. However, coronaviruses have been studied for over 50 years. Coronaviruses are single stranded, positive sense RNA viruses that cause zoonotic diseases that transfer from animal species to humans. They are highly susceptible to recombination and mutation, making them a highly diverse virus (Aronson 2020). Due to their genetic makeup and zoonotic transmission, recent human epidemics have been caused by coronaviruses such as SARS-CoV in 2002, MERS-CoV in 2012, and presently, SARS-CoV-2 which first emerged in 2019. SARS-CoV was an infectious disease that first emerged in China and spread to various countries, making it the first pandemic of the 21st century. However, the outbreak was able to be controlled within about 8 months. In 2012, MERS-CoV was first discovered in the Arabian Peninsula. The virus is still around today, but it has not become a large-scale health concern like SARS (Walls 2020). Contrasting from other coronaviruses, SARS-CoV-2 is highly infectious and the symptoms it causes can range from asymptomatic to severe. This makes SARS-CoV-2 very difficult to control, and the cause behind the variation of symptoms is not completely clear.

The Spike (S) glycoprotein of SARS-CoV-2 is essential to virus attachment and entry (Walls, 2020). It forms homotrimers on the enveloped protein's surface that not only attach to host receptors but also to get activated by host proteases. The glycoprotein contains two subunits: S1 and S2. The receptor-binding domain in the S1 subunit allows for recognition by the

host cell receptor and contributes to stabilization of the S protein. The action of the S1 subunit binding to a host cell receptor causes conformation changes that will trigger the S2 subunit to eventually fuse the viral and host cell membrane together (Walls, 2020). Like its predecessor, SARS-CoV-2 has been shown to bind to angiotensin-converting enzyme 2 (ACE2) using its S1 subunit to attach onto the surface of a cell (Hoffmann, 2020). In addition, all coronaviruses must be cleaved at the S2' site for exposure of the fusion peptide, leading to host and viral membrane fusion. Unique of its kind, SARS-CoV-2 S protein has a furin-motif cleavage site at the S1/S2 site that is a common characteristic of highly pathogenic bird flu viruses (Walls, 2020). Host proteases, furin and transmembrane protease, serine 2 (TMPRSS2) have been shown to cleave the S protein at the S1/S2 site and S2' site, respectively (Bestle, 2020).

Due to the S transmembrane protein's role in host cell entry, it is the target protein for therapeutics and neutralizing antibodies (Walls, 2020). In particular, inhibitors of ACE2 or TMPRSS2 have been considerable approaches in preventing the S protein from being activated, and therefore, the virus from infecting host cells. However, TMPRSS2 may be a safer approach as a target for COVID-19 therapeutic due to several factors. Firstly, ACE-2 is part of the renal-angiotensin system, which maintains blood pressure, electrolyte balance, and vascular resistance. ACE-2-knockout mice were observed to have severe cardiac defects and therefore ACE2 was discovered to be essential for heart function (Baughn, 2020). In contrast, TMPRSS2's physiological function is not completely understood. It is most known for being linked to prostate cancer, as heightened expression of the protease is observed in prostate cancer tissue (Mollica, 2020). Furthermore, TMPRSS2-knockout mice developed normally and survived to adulthood with no problems or noticeable differences (Baughn, 2020). In a different study, rats with knock-out TMPRSS2 had a less severe immune response when infected with SARS-CoV

and MERS-CoV and had less viral spread in the respiratory system (Iwata-Yoshikawa, 2019). The act of inhibiting TMPRSS2 is also not unheard of in medicine. Camostat mesylate, an inhibitor of TMPRSS2, has been shown in vitro to block SARS-CoV-2 entry and is a promising drug candidate (Hoffmann 2020). Bromhexine is found in over-the-counter mucolytic cough suppressant and it has been shown to inhibit TMPRSS2 (Lucas 2014).

Single nucleotide polymorphisms (SNPs) are nucleotide changes to an individual's DNA. They occur very readily, as there is a SNP approximately 1 in every 1,000 nucleotides, which means there are 4-5 million SNPs in an individual's DNA. SNPs can also vary by ethnic groups, which has led to genetic studies between populations. There are two types of SNPS in the coding region: nonsynonymous and synonymous variants. A nonsynonymous variant is a nucleotide change that leads to a change in amino acid or premature stop codon. Contrarily, synonymous variants have no change to the amino acid sequence that is translated. A change in amino acid can be detrimental to the structure of the protein it encodes, leading to the function. It also can vary in lethality depending on the conservation of specific regions of the protein. Supporting this claim, SNPs have been seen to possibly increase/decrease risk in infectivity. Baughn (2020) states that there are 2 common missense variants (rs75603675; p.Gly8Val and rs12329760; p.Val197Met) whose frequencies vary by ancestry and geography. Therefore, targeting TMPRSS2 requires more research to determine if these SNPS can play a part on TMPRSS2 catalytic activity. Cheng (2015) studied if genetic predisposition was possible for types of influenza. His findings revealed that individuals with SNPS rs383510 and rs2070788 of the TMPRSS2 gene are more susceptible to get a severe form of influenza A and acute respiratory distress syndrome. Another variant in TMPRSS2, rs2276205, is associated with increased survival to breast cancer, potentially due to sensitization to estrogen modulators. Lastly,

rs12329760 (C>T) is a common variant that has been studied in men with a genetic predisposition to prostate cancer, doubling the chance of cancer diagnosis (Strope, 2020).

Overall, TMPRSS2 has not been well-studied yet. Mostly recent articles and preprints shine light on the relationship between nonsynonymous variants of TMPRSS2 and pathogenicity of SARS-CoV-2. No crystallized structure of TMPRSS2 is available, and only one TMPRSS2-SARS-CoV-2 biomolecular complex has been published up to date (Hussain, 2020).

In this present study, we asked whether nonsynonymous variants in TMPRSS2 in humans lead to differences in infection rates or severity of disease symptoms of SARS-CoV-2. We did a bioinformatics search in order to find possible pathogenic variants that are common the population, modelled TMPRSS2 using various protein prediction software, and docked the structure against SARS-CoV-2 to determine important binding sites. We predicted that SNPs that are found in these interaction regions are more likely to cause a change in TMPRSS2 structure, leading to improper activation of SARS-CoV-2. Therefore, the variation of common TMPRSS2 SNPs in the population may explain the variation of COVID-19 symptoms.

References

Aronson, J. K. (2020, March 25). Coronaviruses - a general introduction. https://www.cebm.net/covid-19/coronaviruses-a-general-introduction/

Baughn, L. B., Sharma, N., Elhaik, E., Sekulic, A., Bryce, A. H., & Fonseca, R. (2020). Targeting TMPRSS2 in SARS-CoV-2 Infection. *Mayo Clinic proceedings*, 95(9), 1989–1999. https://doi.org/10.1016/j.mayocp.2020.06.018

Cheng, Z., Zhou, J., To, K. K. W., Chu, H., Li, C., Wang, D., ... & Yuen, K. Y. (2015). Identification of TMPRSS2 as a susceptibility gene for severe 2009 pandemic A (H1N1) influenza and A (H7N9) influenza. *The Journal of infectious diseases*, *212*(8), 1214-1221. https://doi.org/10.1093/infdis/jiv246

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., ... & Pöhlmann, S. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*, *181*(2), 271-280. https://doi.org/10.1016/j.cell.2020.02.052

Iwata-Yoshikawa, N., Okamura, T., Shimizu, Y., Hasegawa, H., Takeda, M., & Nagata, N. (2019). TMPRSS2 contributes to virus spread and immunopathology in the airways of murine models after coronavirus infection. *Journal of virology*, *93*(6). https://doi.org/10.1128/JVI.01815-18

Lucas, J. M., Heinlein, C., Kim, T., Hernandez, S. A., Malik, M. S., True, L. D., ... & Nelson, P. S. (2014). The androgen-regulated protease TMPRSS2 activates a proteolytic cascade involving components of the tumor microenvironment and promotes prostate cancer metastasis. *Cancer discovery*, *4*(11), 1310-1325. https://doi.org/10.1158/2159-8290

Mollica, V., Rizzo, A., & Massari, F. (2020). The pivotal role of TMPRSS2 in coronavirus

disease 2019 and prostate cancer. *Future Medicine*, 16(27), 2029–2033. https://doi.org/10.2217/fon-2020-0571

Strope, J. D., & Chau, C. H. (2020). TMPRSS2: Potential Biomarker for COVID-19 Outcomes. *Journal of clinical pharmacology*. https://doi.org/10.1002/jcph.1641

Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., & Veesler, D. (2020). Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*, *181*(2), 281-292. https://doi.org/10.1016/j.cell.2020.02.058