Applying bioinformatics to examine the structural and functional relationships between TMPRSS2 SNPs and SARS-CoV-2 and their potential implications for Covid-19 infection.

Annika Dinulos

Loyola Marymount University

Kam Dahlquist, Ph.D.

Annotated Bibliography

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

This research article provides expression information of TMPRSS2 and ACE2 in specific host tissues. Expression data was then sorted and displayed according to gender, and results show that gender disparities in Covid-19 infection are likely not due to ACE2 and TMPRSS2 expression. This research is important for understanding background information about TMPRSS2, where it can be found, and some overall population data.

David, A., Khanna, T., Beykou, M., Hanna, G., & Sternberg, M. J. (2020). Structure, function and variants analysis of the androgen-regulated TMPRSS2, a drug target candidate for COVID-19 infection. *bioRxiv.*

Information in this article includes a structural model of TMPRSS2, known TMPRSS2 variants and population data. The methods section provides a good background for databases and software that are very useful in this research. Both SIFT and PolyPhen2 report “damaging” statistics for TMPRSS2 SNPs and are good resources to sort and prioritize SNPs.

Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases, 20*(5), 533-534.

Dong et. al report data on Covid-19 positive cases, deaths, and recoveries across multiple countries. This information is useful to examine and track the rate of infection, death, and recoveries and is relevant to understand the severity and importance of Covid-19 research.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N. H., Nitsche, A., Müller, M. A., Drosten, C., & 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.e8. https://doi.org/10.1016/j.cell.2020.02.052

This article provides information on the mechanisms of SARS-CoV-2 cell entry. It states that TMPRSS2 is an S protein primer and indicates that TMPRSS2 could provide avenues for future treatment options. It identifies the importance of TMPRSS2 research and also describes cell entry information about both SARS-CoV and SARS-CoV-2.

Hussain, M., Jabeen, N., Amanullah, A., Baig, A. A., Aziz, B., Shabbir, S., ... & Uddin, N. (2020). Molecular docking between human TMPRSS2 and SARS-CoV-2 spike protein: conformation and intermolecular interactions. *AIMS microbiology, 6*(3), 350.

Hussain et. al provide very useful information on TMPRSS2 and SARS-CoV-2 spike protein interactions. They provide docking and conformational models that visualize several possible interactions. Several different polymorphisms are provided in their results section that should be examined.

Klaassen, K., Stankovic, B., Zukic, B., Kotur, N., Gasic, V., Pavlovic, S., & Stojiljkovic, M. (2020). Functional prediction and comparative population analysis of variants in genes for proteases and innate immunity related to SARS-CoV-2 infection. *bioRxiv.*

Information reported in this research compares the function of proteases and genes that are believed to be involved with SARS-CoV-2 infection. The researchers also included population genetics data that point out different variances in these proteases and genes, including TMPRSS11a. They did not find notable differences within TMPRSS2 but point to connections between TMPRSS11 and TMPRSS2.

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

This research provides a good, comprehensive background on the serine protease TMPRSS2 and its properties when interacting with SARS-CoV-2. It also mentions the other known functions and related diseases such as prostate cancer and androgen signalling, which allows for comparison between the two mechanisms. Expression data in smaller populations is also provided and is compared to other published data.

Ravikanth V. et al. (2020) Genetic variants in TMPRSS2 and Structure of SARS-CoV-2 spike glycoprotein and TMPRSS2 complex. *bioRxiv*. doi:10.1101/2020.06.30.179663

This new research provides another good list of functionally relevant SNPs within ACE2 and TMPRSS2. It also shows the frequency of the variants and is sorted for quality and variants in both genes. Structure images are provided of the subunits of the S protein and are a good representation of promoters and other domains.

Russo, R., Andolfo, I., Lasorsa, V. A., Iolascon, A., & Capasso, M. (2020). Genetic analysis of the coronavirus SARS-CoV-2 host protease TMPRSS2 in different populations. *Frontiers in genetics, 11, 872.*

Genetic markers and allele frequencies were examined in this paper, along with tissue expression data of TMPRSS2. The researchers looked at this data across multiple tissues and populations in different countries, and localized two variants that may be associated with increased SARS-CoV-2 expression. These variants could be indicators of differing Covid-19 presentations.

Shen, L.W.; Mao, H.J.;, Wu, Y.L.; Tanaka,Y.; Zhang,W. (2017) TMPRSS2: A potential target for treatment of influenza virus and coronavirus infections, *Biochimie, 142,* 1-10. https://doi.org/10.1016/j.biochi.2017.07.016.

Structure models of TMPRSS2 are provided in this paper, as gene maps and schematics of TMPRSS2 membrane fusion. This paper also provides good images that show the replication cycle of coronaviruses and compare it to influenza. The comparisons of coronavirus and influenza infections in respect to TMPRSS2 are important as they highlight potentials for future therapies and interventions.

Yamamoto, N., Ariumi, Y., Nishida, N., Yamamoto, R., Bauer, G., Gojobori, T., Shimotohno, K., & Mizokami, M. (2020). SARS-CoV-2 infections and COVID-19 mortalities strongly correlate with ACE1 I/D genotype. *Gene, 758*, 144944. https://doi.org/10.1016/j.gene.2020.144944

Yamamoto et al. highlight potential polymorphisms in ACE1, either an insertion or deletion, and hypothesize that the insertion allele increases the risk of Covid-19 death. This research uses population data to determine if the presence of these alleles is influenced by human migration patterns and can be localized to specific regions. This provides good information on the potential for population genetic differences in ACE1 frequencies - providing good background information on how to continue onwards with TMPRSS2 research.

Zang, R., Gomez Castro, M. F., McCune, B. T., Zeng, Q., Rothlauf, P. W., Sonnek, N. M., Liu, Z., Brulois, K. F., Wang, X., Greenberg, H. B., Diamond, M. S., Ciorba, M. A., Whelan, S., & Ding, S. (2020). TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Science Immunology, 5*(47), eabc3582. https://doi.org/10.1126/sciimmunol.abc3582

Information about TMPRSS2 and TMPRSS4 is provided in this article, which inspired the idea to further investigate the relevance and relationship between TMPRSS4 and SARS-CoV-2 infection. TMPRSS4 is more expressed in intestinal enterocytes and is also relevant in influenza A infection. This paper also discusses that infection of human small intestine enterocytes is shed through the colon, discussing that another possible infection route is through fecal-oral spread.

Zhou, L., Niu, Z., Jiang, X., Zhang, Z., Zheng, Y., Wang, Z., ... & Sun, Q. (2020). Systemic analysis of tissue cells potentially vulnerable to SARS-CoV-2 infection by the protein-proofed single-cell RNA profiling of ACE2, TMPRSS2 and Furin proteases. *bioRxiv*.

This research profiles ACE2, TMPRSS2 and furin proteases in different cell tissues derived from different organs. They hypothesize that co-expression of TMPRSS2 and furin proteases with ACE2 in tissues actually increases the susceptibility of viral infection. It is mentioned that these co-expressions are likely highest in lung tissue, providing a possible reason for the severity of the respiratory symptoms of Covid-19.