Characterizing evolution pressures in Sars-CoV-2 on a broader scale

Peter C. Jentsch[1,4], Finlay Maguire[3,5], and Samira Mubareka[1,2]

[1] *Sunnybrook Research Institute, Toronto, Canada*
[2] *University of Toronto, Toronto, Canada*
[3] *Dalhousie University, Halifax, Canada*
[4] *Simon Fraser University, Burnaby, Canada*
[5] *Shared Hospital Laboratory, Toronto, Canada*

May 13, 2022

1 Introduction

(Peter: need to rework much of this text...)

The objective of this work is to study the SARS-CoV-2 viral diversity in Ontario and Quebec. Initially, the focus will be comparing differences between the provinces. The prevalence of the Gamma and Alpha variants of concern (VOC) as differed significantly between Ontario and Quebec during the 3rd wave of the pandemic in Canada (February 2021 to June 2021). The key differences between Alpha and Gamma are in immune escape and transmissibility [7]. Alpha has less immune escape (for neutralising sera from vaccinated or convalescent individuals), but binds ACE2 more readily and therefore relatively higher transmission. We will not need to consider the S:E484k mutation, which affects immune escape, since it was in a negligible proportion of Canadian Alpha infections during this time period. On the other hand, the Gamma lineage likely has greater immune escape but is probably relatively less transmissible. In contrast to Ontario, during this time the government of Quebec was actively prioritizing the distribution of mRNA vaccines to groups more vulnerable to hospitalization and communities with VOC outbreaks. Therefore, both the mRNA (Pfizer, Moderna) and AstraZeneca vaccine will be included in our model.

Compartmental differential equation models have been used extensively to model competing viral strains in the past (e.g. [2, 13, 16, 21]), and have been invaluable to informing policy throughout the pandemic [4,9,15]. One approach to address this research question is to develop a parsimonious model that accurately reproduces observed data. Such a model would generate generalizable insights into the effect of NPIs more broadly. To this end, here we suggest a compartmental model to act as a jumping-off point (Section **??**, Figure **??**),

and request data fields to parameterize it (Section **??**). Most disease parameters can be obtained from literature. We will estimate underreporting and underascertainment multipliers by comparing observed case data with serological studies [9].

2 Model

The evolution of Sars-CoV-2 throughout the pandemic is marked by antigenic drift [24], giving rise to new variants that exhibit significant immune escape. The model of Gog and Grenfell [8] constrains strain space to a one or two dimensional lattice, thereby making the analysis of strain evolution tractable. In two dimensional strain space, cross-immunity of a pathogen is given specified by a coefficient σ_{ijkl}, and mutation is implemented as discrete diffusion with some fixed speed. These ideas were generalized to n-dimensional strain space, and applied to modeling drift in influenza A by Kryazhimskiy et al [10]. This application of these models to influenza was partially motivated by the work of Lapedes and Farber [11], and Smith et al. [18], which suggests that the antigenic evolution of Influenza A primarily occurs within two dimensions. The idea for this line of Sars-CoV-2 research is to handle strain space similar to the aforementioned work, but include further compartments and dynamics specific to Sars-CoV-2 to test hypotheses of evolution.

$$S'_{ij}(t) = -\sum_{kl} \beta_{kl}\sigma_{ijkl}S_{ij}I_{kl} + \gamma R_{ij} \tag{1}$$

$$I'_{ij}(t) = \beta_{ij}S_{ij}I_{ij} - \xi I_{ij} + M\left(-4I_{ij} + I_{i-1,j} + I_{i+1,j} + I_{i,j-1} + I_{i,j+1}\right) \tag{2}$$

$$R'_{ij}(t) = \xi I_{ij} - \gamma R_{ij} \tag{3}$$

Symbol	Description
N	Size of variant grid
S_{ij}	Population susceptible to variant $(i,j) \in [0,N]^2$
I_{ij}	Population infected by variant $(i,j) \in [0,N]^2$
R_{ij}	Recovered/Immune to variant $(i,j) \in [0,N]^2$
σ_{ijkl}	Probability that exposure to variant (i,j) causes immunity to variant (k,l)
β_{ij}	Transmission rate of variant (i,j)
ξ	Recovery rate of all strains
γ	Rate of immunity loss of all strains

Table 1: Table of symbols for Model 2

Equations 4-5 represent the model of [8] with a two-dimensional strain space, with a few changes to better reflect mechanisms of of Sars-CoV-2. I have added

an immunity period, changing the model to an SIRS mechanism, and removed the vital dynamics, as I do not think natural population birth rates are significant in the time scale of the pandemic. A key assumption made by this model is that exposure grants complete immunity to some fraction of individuals, rather than partial immunity (interpreted as reduced transmission rates) to all exposed individuals. Many other methods of dealing with cross-immunity are possible, but this method gives a simpler state space [1].

To incorporate vaccination, consider each vaccine affecting a different region of strain space. That is, for a vaccine v we can associate a matrix $v_{ij} \in [0, 1]$ which determines the relative effect of that vaccine on the immunity of hosts to strain i, j. The function $eta(t)$ represents some base rate of vaccination. This results in the following equations for $S'_{ij}(t)$ and $R'_{ij}(t)$ (the infected equation 2 is unchanged).

$$S'_{ij}(t) = -\sum_{kl} \beta_{kl}\sigma_{ijkl}S_{ij}I_{kl} + \gamma R_{ij} - \eta(t)v_{ij}S_{ij} \tag{4}$$

$$R'_{ij}(t) = \xi I_{ij} - \gamma R_{ij} + \eta(t)v_{ij}S_{ij} \tag{5}$$

This model could be used to test the effect of NPIs on Sars-CoV-2 evolution. Mechanisms for NPIs and additional compartments for heterogeneity are straightforward to add to the model equations. If a given NPI mechanism is more able to fit data with a mechanism that does not act on all strains equally, then this could be evidence that NPIs are affecting Sars-CoV-2 evolution.

2.1 Continuous strain-space

The above model can be viewed as simply a first-order finite difference approximation of a continuous space reaction-diffusion model. Accordingly, we can generalize it to continuous strain-space as

$$S_t(x, y, t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \beta(x', y')\sigma(x, y, x', y')S(x, y, t)I(x', y', t)dx'dy' + \gamma R_{ij} - \eta(t)v(x, y)S(x, y, t) \tag{6}$$

$$I_t(x, y, t) = \beta(x, y)S(x, y, t)I(x, y, t) - \xi I(x, y, t) + M\left(I_x(x, y, t) + I_y(x, y, t)\right) \tag{7}$$

$$R_t(x, y, t) = \xi I(x, y, t)I(x, y, t) - \gamma R(x, y, t) + \eta(t)v(x, y)S(x, y, t) \tag{8}$$

where β, σ, v have been generalized to their continuous counterparts. This formulation is similar to the 1-dimension strain space model described in [3]. Then, given a dispersion kernel $K(x, y) \in L_2 : \mathbb{R}^2 \to \mathbb{R}$ this can be generalised to non-local diffusion as follows

$$I_t(x, y, t) = \beta(x, y)S(x, y, t)I(x, y, t) - \xi I(x, y, t) + M\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K(x - x', y - y')I(x', y', t)dx'dy'\right) \tag{9}$$

2.2 Parameterization

Previous work on these models do not use data to estimate parameters, focusing on broad characterization of dynamics with numerical or analytical approaches. However, genomic data does include a huge amount of admittedly very noisy information. There is some recent work on parameter estimation by comparing simulated phylogeny with observed phylogeny using a suite of summary statistics for tree structures [6,12,17], but these models do not explicitly include genomic structure, which might provide additional inferential power.

Placing individual samples into the collapsed strain space could be done by using neutralizing antibody studies. The neutralizing response has been characterized for most variants of concern with respect to monoclonal antibodies, convalescent plasma, and plasma from vaccinated persons [20]. To an even higher level of detail. Starr et al. characterized polyclonal antibody binding over all mutations of the Receptor Binding Domain (RBD) [19]. If the neutralizing response is known, we assign strains positions such that the distance between any two strains is equal to that neutralizing response. If the response is not known, we just assume that the position is close to the most recent ancestor for which the response is known (likely a VoC). This approach is especially appropriate for recurrent mutations, since it is unlikely that a mutation that has occurred many times in the viral history but has not evolved further confers a significant advantage. The immune-space position of vaccinate-induced immunity can also be computed this way, although we will still need to make significant assumptions on their geometry. It might be best to assume a simple symmetric shape with a centre given by neutralization results.

Mirroring the work on mapping the strain space for influenza [5,11,18] computing this embedding is a metric multidimensional scaling problem. This method has been applied to the antigenic space of the Sars-CoV-2 as well [14,22,23].

References

[1] *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods, and Theory*, volume 126 of *The IMA Volumes in Mathematics and its Applications*. Springer New York, 2002.

[2] Samuel Alizon and Minus van Baalen. Multiple infections, immune dynamics, and the evolution of virulence. *The American Naturalist*, 172(4):E150–E168, Oct 2008.

[3] Nikolai Bessonov, Gennady Bocharov, Andreas Meyerhans, Vladimir Popov, and Vitaly Volpert. Existence and dynamics of strains in a nonlocal reaction-diffusion model of viral evolution. *SIAM Journal on Applied Mathematics*, 81(1):107–128, Jan 2021.

[4] Kate M Bubar, Kyle Reinholt, Stephen M Kissler, Marc Lipsitch, Sarah Cobey, Yonatan H Grad, and Daniel B Larremore. Model-informed

covid-19 vaccine prioritization strategies by age and serostatus. *Science*, 371(6532):916–921, 2021.

[5] Zhipeng Cai, Tong Zhang, and Xiu-Feng Wan. A computational framework for influenza antigenic cartography. *PLoS computational biology*, 6(10):e1000949, 2010.

[6] Gonché Danesh, Victor Virlogeux, Christophe Ramière, Caroline Charre, Laurent Cotte, and Samuel Alizon. Quantifying transmission dynamics of acute hepatitis c virus infections in a heterogeneous population using sequence data. *PLoS pathogens*, 17(9):e1009916, 2021.

[7] Agency for Clinical Innovation. Living evidence - sars-cov-2 variants. `https://aci.health.nsw.gov.au/covid-19/critical-intelligence-unit/sars-cov-2-variants`, Jan 2022.

[8] J. R. Gog and B. T. Grenfell. Dynamics and selection of many-strain pathogens. *Proceedings of the National Academy of Sciences*, 99(26):17209–17214, December 2002.

[9] Peter C Jentsch, Madhur Anand, and Chris T Bauch. Prioritising covid-19 vaccination in changing social and epidemiological landscapes: a mathematical modelling study. *The Lancet Infectious Diseases*, 2021.

[10] Sergey Kryazhimskiy, Ulf Dieckmann, Simon A. Levin, and Jonathan Dushoff. On State-Space Reduction in Multi-Strain Pathogen Models, with an Application to Antigenic Drift in Influenza A. *PLOS Computational Biology*, 3(8):e159, August 2007.

[11] Alan Lapedes and Robert Farber. The Geometry of Shape Space: Application to Influenza. *Journal of Theoretical Biology*, 212(1):57–69, September 2001.

[12] Gabriel E Leventhal, Roger Kouyos, Tanja Stadler, Viktor Von Wyl, Sabine Yerly, Jürg Böni, Cristina Cellerai, Thomas Klimkait, Huldrych F Günthard, and Sebastian Bonhoeffer. Inferring epidemic contact structure from phylogenetic trees. *PLoS computational biology*, 8(3):e1002413, 2012.

[13] Marc Lipsitch, Caroline Colijn, Ted Cohen, William P. Hanage, and Christophe Fraser. No coexistence for free: Neutral null models for multistrain pathogens. *Epidemics*, 1(1):2–13, Mar 2009.

[14] Nathaniel L. Miller, Thomas Clark, Rahul Raman, and Ram Sasisekharan. An Antigenic Space Framework for Understanding Antibody Escape of SARS-CoV-2 Variants. *Viruses*, 13(10):2009, October 2021.

[15] Swapnil Mishra, Tresnia Berah, Thomas A. Mellan, H. Juliette T. Unwin, Michaela A. Vollmer, Kris V. Parag, Axel Gandy, Seth Flaxman, and Samir Bhatt. On the derivation of the renewal equation from an age-dependent branching process: an epidemic modelling perspective. *arXiv:2006.16487 [q-bio, stat]*, Jun 2020. arXiv: 2006.16487.

[16] Emily J. Nicoli, Diepreye Ayabina, Caroline L. Trotter, Katherine M.E. Turner, and Caroline Colijn. Competition, coinfection and strain replacement in models of bordetella pertussis. *Theoretical Population Biology*, 103:84–92, Aug 2015.

[17] Emma Saulnier, Olivier Gascuel, and Samuel Alizon. Inferring epidemiological parameters from phylogenies using regression-abc: A comparative study. *PLoS computational biology*, 13(3):e1005416, 2017.

[18] Derek J. Smith, Alan S. Lapedes, Jan C. de Jong, Theo M. Bestebroer, Guus F. Rimmelzwaan, Albert D. M. E. Osterhaus, and Ron A. M. Fouchier. Mapping the Antigenic and Genetic Evolution of Influenza Virus. *Science*, 305(5682):371–376, July 2004.

[19] Tyler N Starr, Allison J Greaney, Sarah K Hilton, Daniel Ellis, Katharine HD Crawford, Adam S Dingens, Mary Jane Navarro, John E Bowen, M Alejandra Tortorici, Alexandra C Walls, et al. Deep mutational scanning of sars-cov-2 receptor binding domain reveals constraints on folding and ace2 binding. *Cell*, 182(5):1295–1310, 2020.

[20] Philip Tzou, Kaiming Tao, Janin Nouhin, Soo-Yon Rhee, Benjamin Hu, Shruti Pai, Neil Parkin, and Robert Shafer. Coronavirus antiviral research database (cov-rdb): An online database designed to facilitate comparisons between candidate anti-coronavirus compounds. *Viruses*, 12(9):1006, Sep 2020.

[21] Minus van Baalen and Maurice W. Sabelis. The dynamics of multiple infection and the evolution of virulence. *The American Naturalist*, 146(6):881–910, Dec 1995.

[22] Karlijn van der Straten, Denise Guerra, Marit van Gils, Ilja Bontjer, Tom G Caniels, Hugo D van Willigen, Elke Wynberg, Meliawati Poniman, Judith A Burger, Joey H Bouhuijs, et al. Mapping the antigenic diversification of sars-cov-2. *medRxiv*, 2022.

[23] Samuel H. Wilks, Barbara Mühlemann, Xiaoying Shen, Sina Türeli, Eric B. LeGresley, Antonia Netzl, Miguela A. Caniza, Jesus N. Chacaltana-Huarcaya, Xiaoju Daniell, Michael B. Datto, Thomas N. Denny, Christian Drosten, Ron A. M. Fouchier, Patricia J. Garcia, Peter J. Halfmann, Agatha Jassem, Terry C. Jones, Yoshihiro Kawaoka, Florian Krammer, Charlene McDanal, Rolando Pajon, Viviana Simon, Melissa Stockwell, Haili Tang, Harm van Bakel, Richard Webby, David C. Montefiori, and Derek J. Smith. Mapping SARS-CoV-2 antigenic relationships and serological responses. Preprint, Immunology, January 2022.

[24] Jonathan W. Yewdell. Antigenic drift: Understanding COVID-19. *Immunity*, 54(12):2681–2687, December 2021.