Strain-space model for Sars-CoV-2

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

[1]Sunnybrook Research Institute, Toronto, Canada

[2]University of Toronto, Toronto, Canada

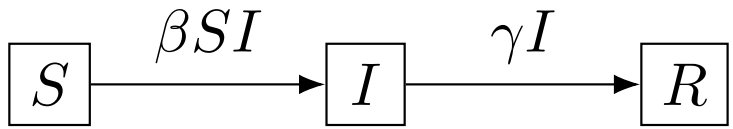
[3]Dalhousie University, Halifax, Canada

[4]Simon Fraser University, Burnaby, Canada

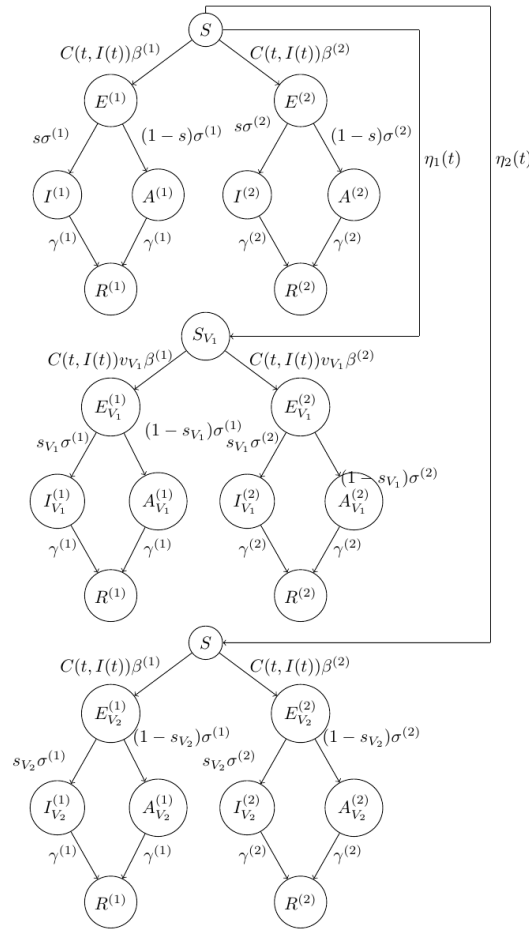
[5]Shared Hospital Laboratory, Toronto, Canada

June 29, 2022

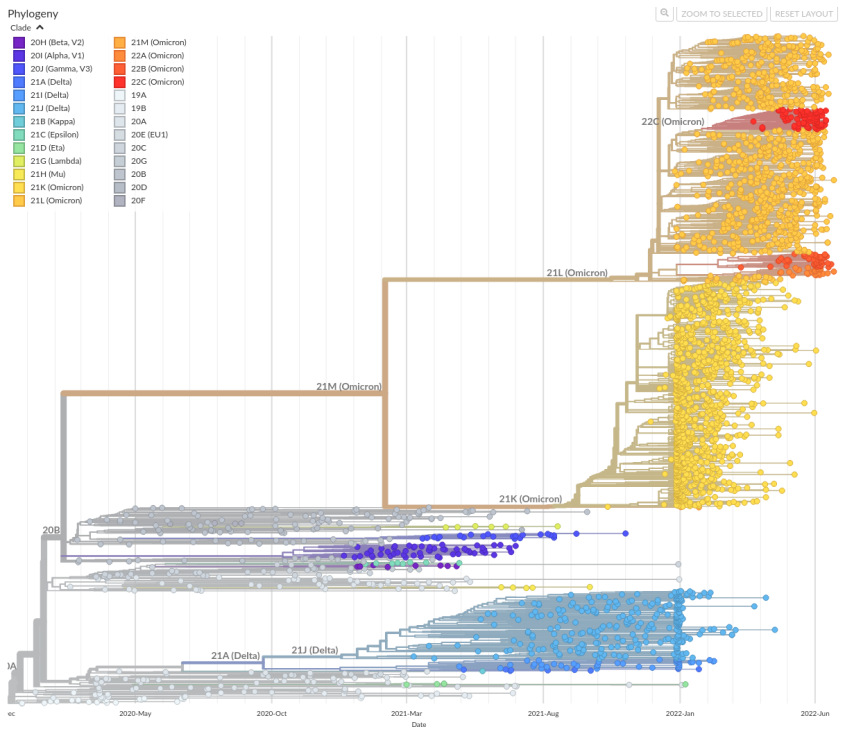
- Infection spread is often modelled using compartmental models
- Represent subsets of a host population and rates of movement between them



- Multiple infections (e.g. competing VoCs) can be represented as more compartments
- Work on multiple infections is usually here due to lack of data, increasing complexity

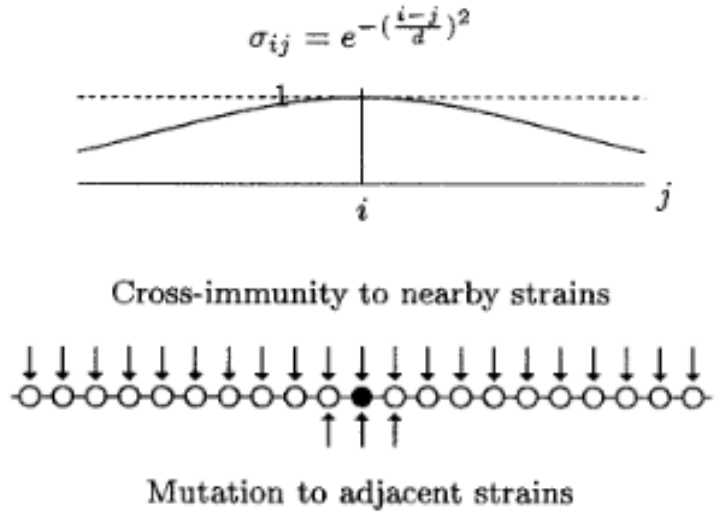


This only represents a tiny amount of the genomic data we have for Sars-CoV-2!

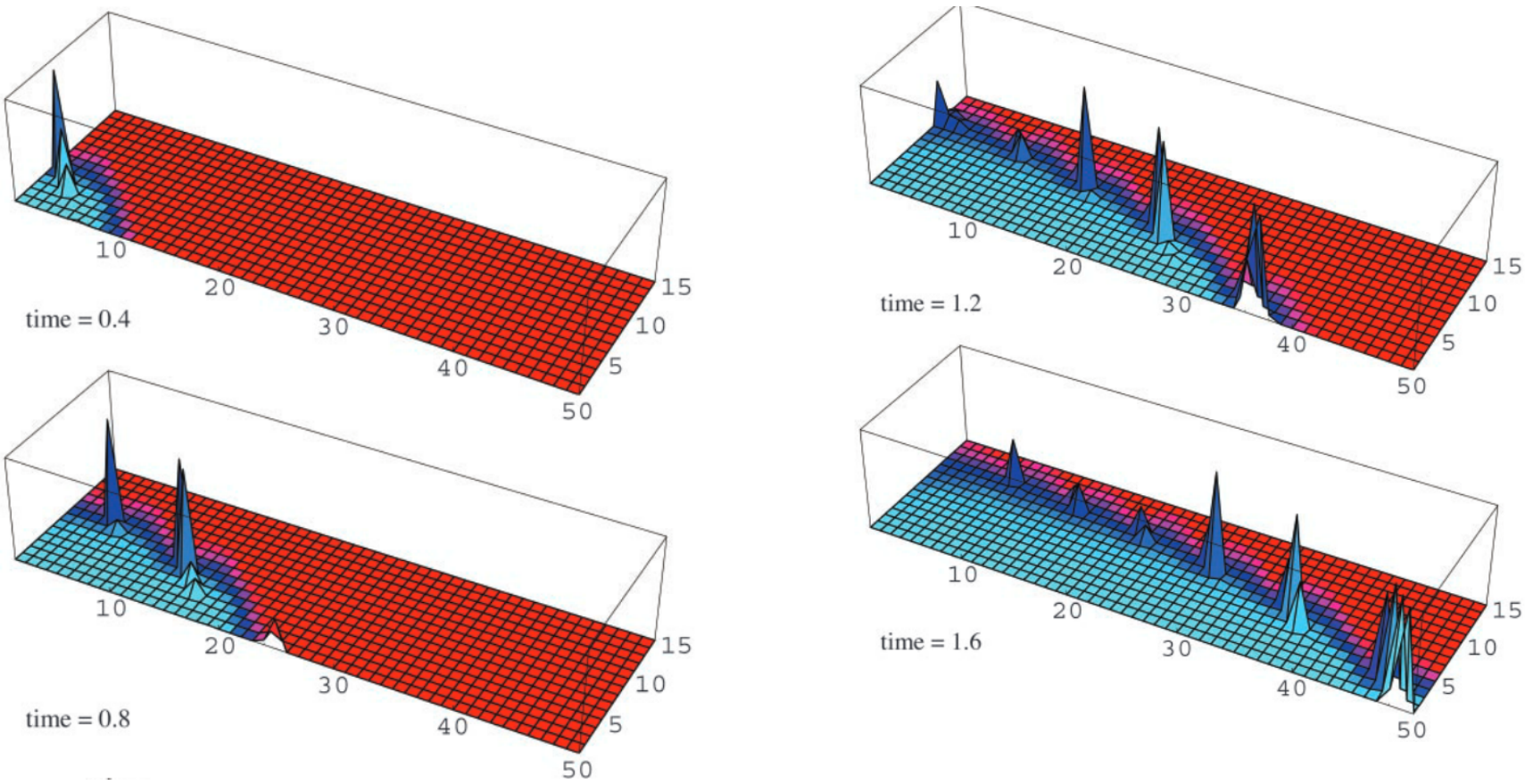


- Can extend these models to a sequence of variants
- Assume each variant is indexed by i
- The dynamics at each variant i are determined by a simple compartmental model

- Variants are related by a function $\sigma(i, j)$ that determines how much an infection by variant i reduces probability of infection to variant j.
- A variant i mutates to neighbouring indices $i + 1, i - 1$ proportional to the population of variant i



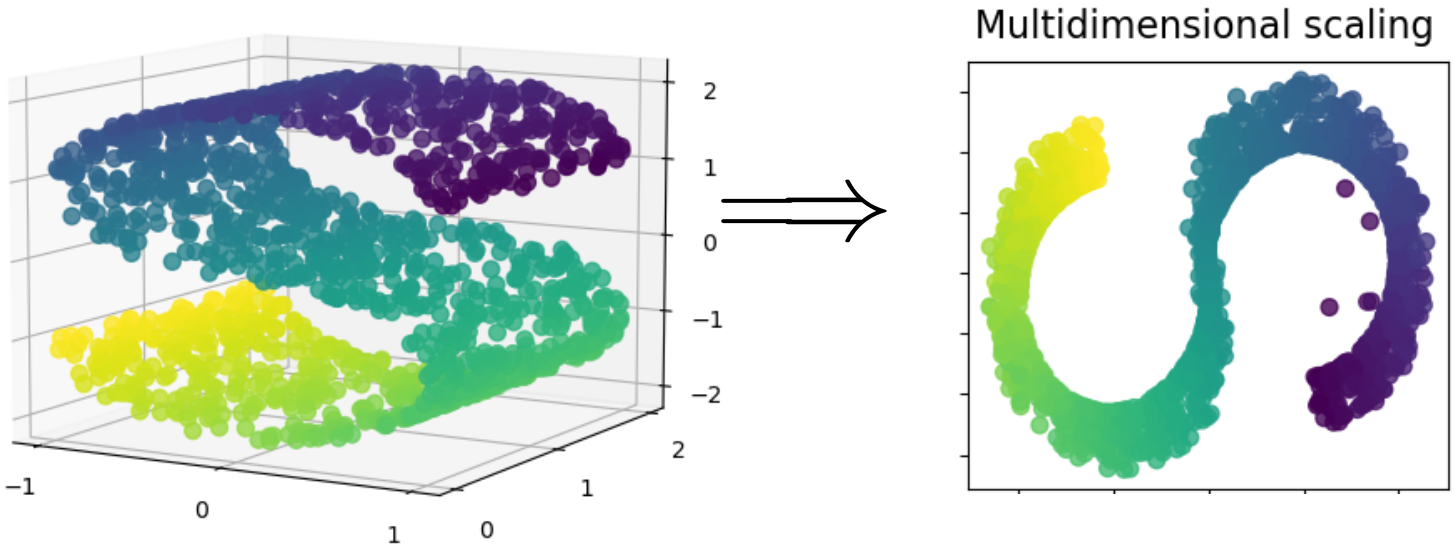
[Gog and Grenfell, 2002]



[Gog and Grenfell, 2002]

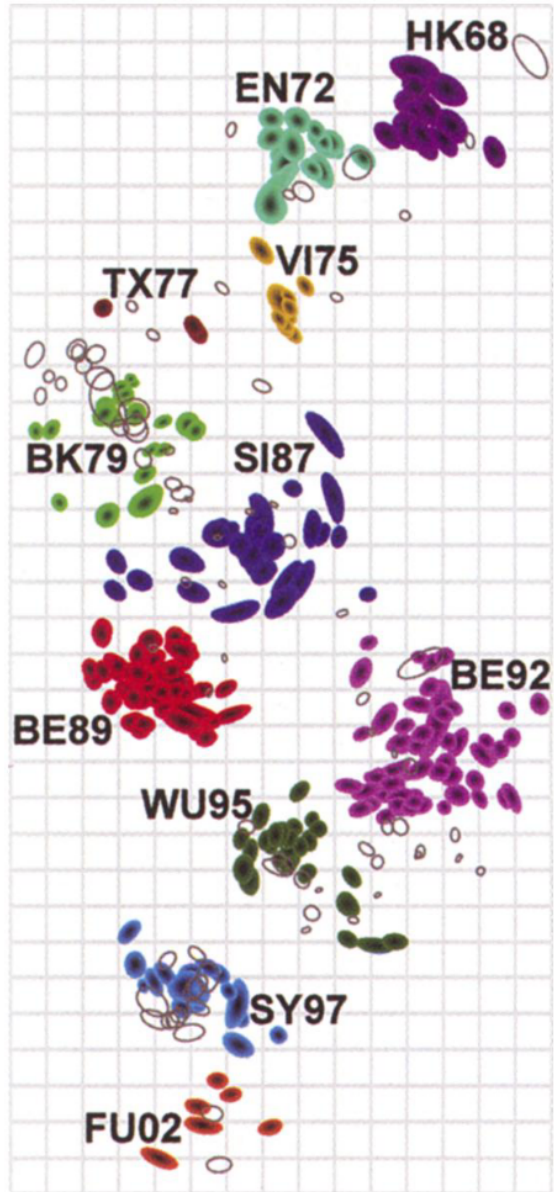
Antigenic cartography

- Practice of mapping out immune responses to related pathogens
- Distance between serums and pathogen is quantified, these points are visualized using multidimensional scaling (MDS)

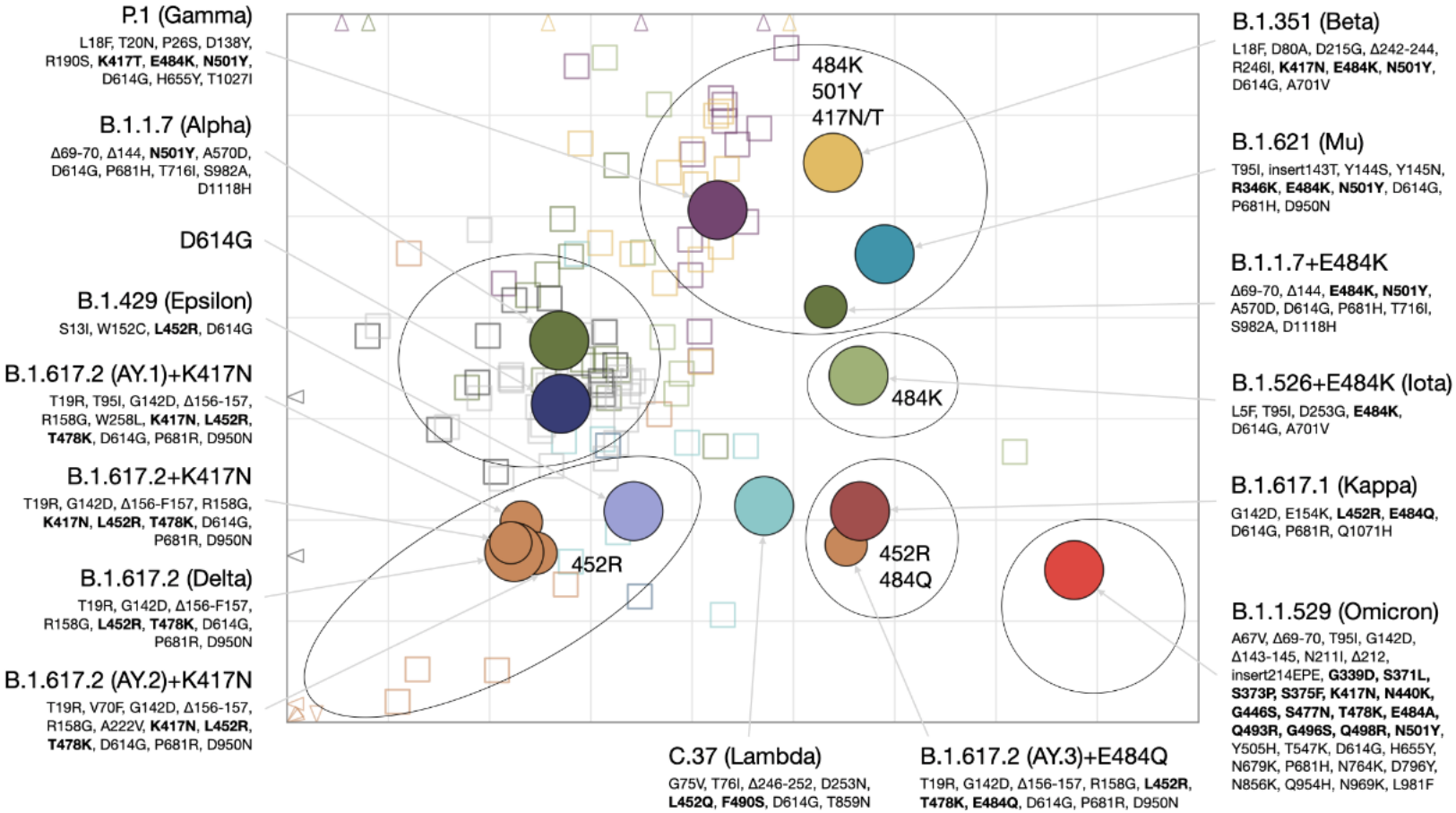


[Pedregosa et al., 2011]

This technique was developed to visualize the antigenic drift of influenza A (H3N2)



[Lapedes and Farber, 2001, Smith et al., 2004]

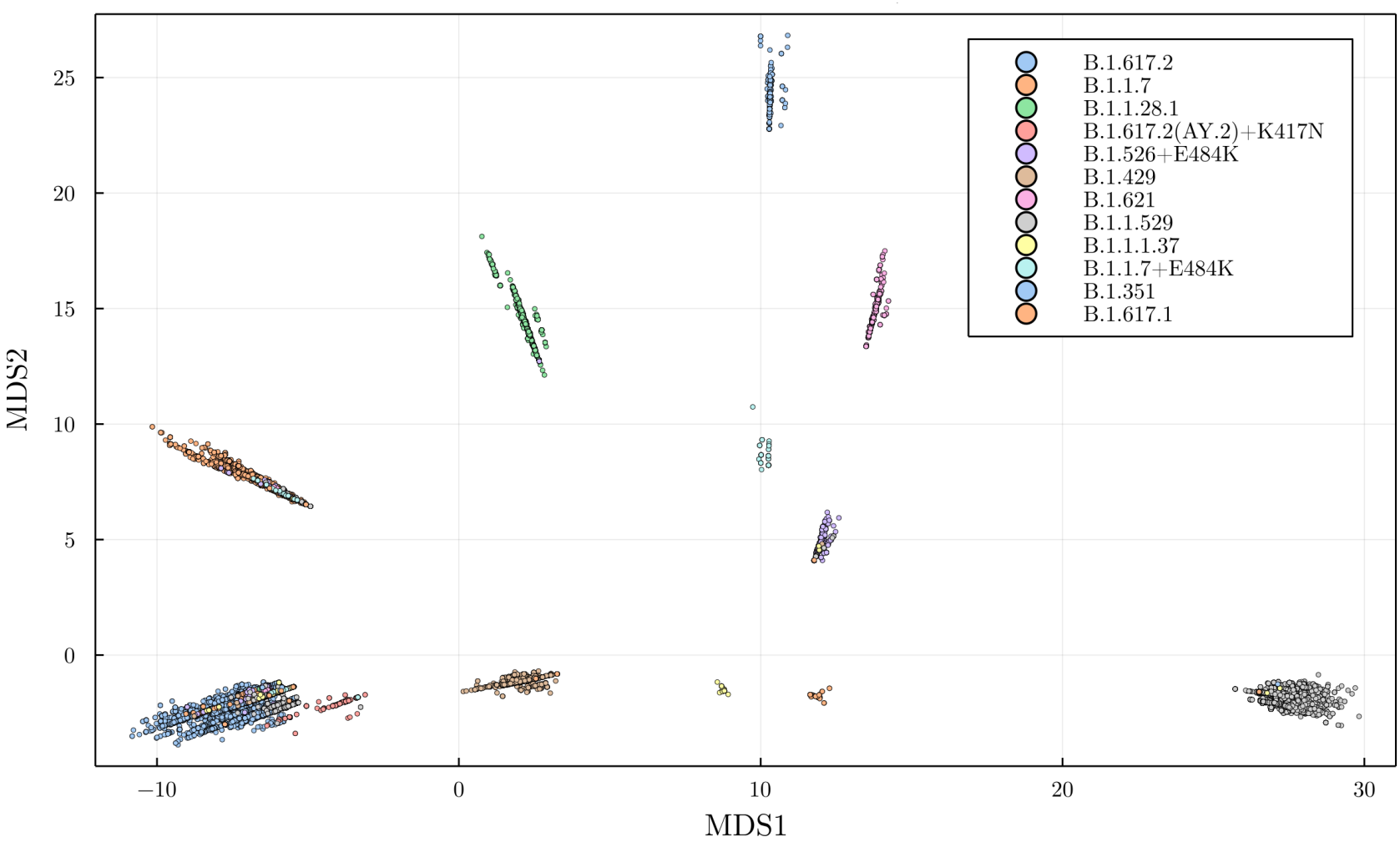


[Wilks et al., 2022]

These results suggest that 2 dimensions might be an adequate approximation to the full space!

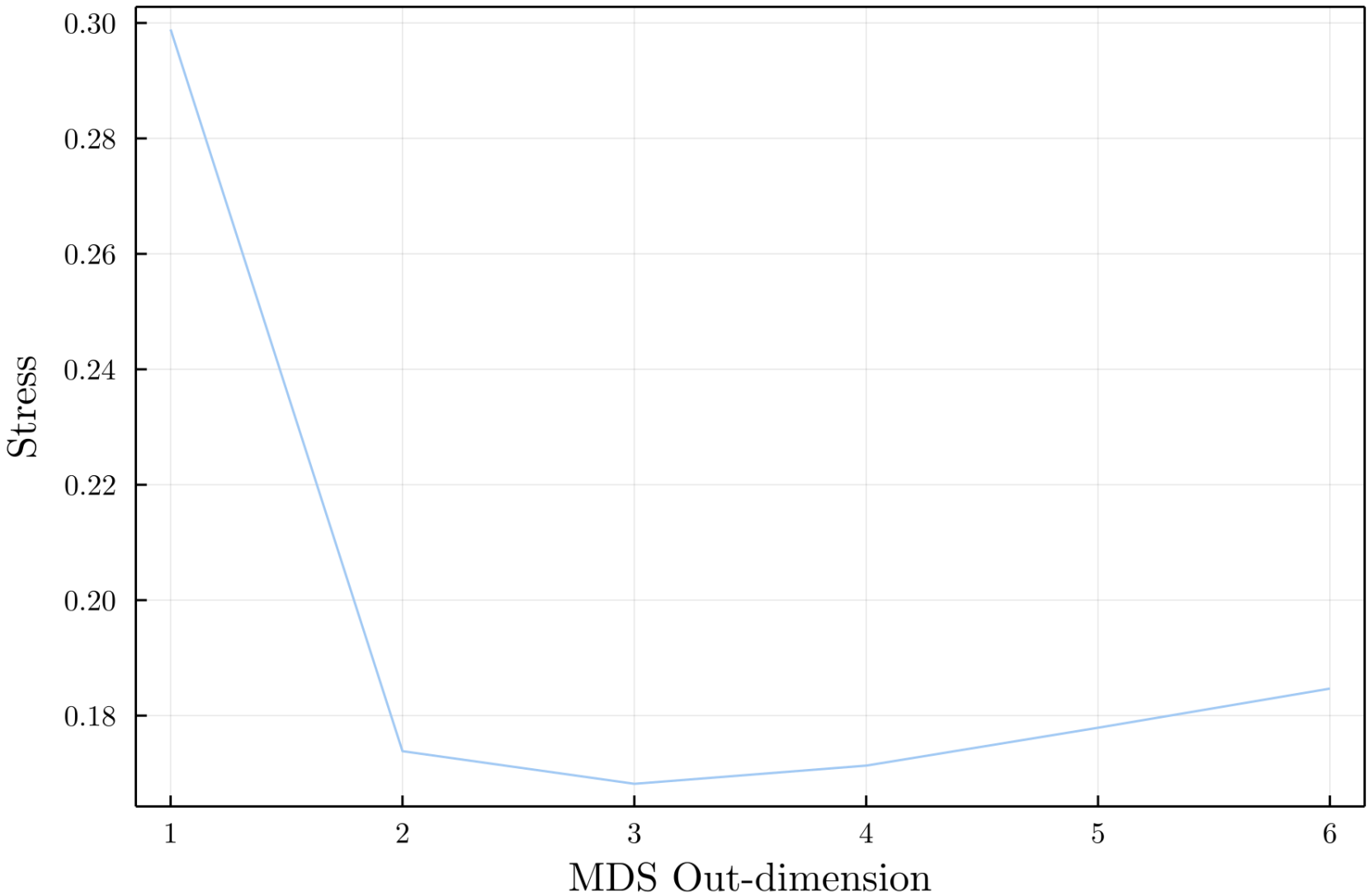
Incorporating more data

- We have millions of samples and hundreds of thousands of unique genomes from random sampling alone
- Assign each sample to its closest lineage in the map from [Wilks et al., 2022]
- Distance between any two samples is determined by the distance between the respective closest lineage in Wilks, plus the SNP distance, where each SNP is weighted by how often it recurs in the global tree



Map of all unique genomes collected in random USA sampling, considering the top 150 most recurrent SNPs

Evaluating the MDS approximation



(animation of kernel approximation)

Model Equations

$$\frac{S_{ij}}{dt} = -\sum_{kl} s(t)\beta_{kl}\sigma_{ijkl}S_{ij}I_{kl} + \gamma R_{ij} - V(t)S \tag{1}$$

$$\frac{I_{ij}(t)}{dt} = s(t)\beta_{ij}S_{ij}I_{ij} - \xi I_{ij} + M\left(-4I_{ij} + I_{i-1,j} + I_{i+1,j} + I_{i,j-1} + I_{i,j+1}\right) \tag{2}$$

$$\frac{R_{ij}(t)}{dt} = \xi I_{ij} - \gamma R_{ij} + V(t)S \tag{3}$$

Boundary conditions: $I_{0,j} = 0, I_{j,0} = 0, I_{N,j} = 0, I_{j,N} = 0$
Initial conditions computed from genomic data in GISAID

Model parameters/variables

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)
$v(t)$	vaccination rate at time t
$s(t)$	stringency at t
ξ	Recovery rate of all strains
γ	Rate of immunity loss of all strains

Table of symbols for Model 2

I am working on fitting to data at the moment


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


Lapedes, A. and Farber, R. (2001).
The Geometry of Shape Space: Application to Influenza.
Journal of Theoretical Biology, 212(1):57–69.


Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011).
Scikit-learn: Machine learning in Python.
Journal of Machine Learning Research, 12:2825–2830.

Smith, D. J., Lapedes, A. S., de Jong, J. C., Bestebroer, T. M., Rimmelzwaan, G. F., Osterhaus, A. D. M. E., and Fouchier, R. A. M. (2004).
Mapping the Antigenic and Genetic Evolution of Influenza Virus.
Science, 305(5682):371–376.

Wilks, S. H., Mühlemann, B., Shen, X., Türeli, S., LeGresley, E. B., Netzl, A., Caniza, M. A., Chacaltana-Huarcaya, J. N., Daniell, X., Datto, M. B., Denny, T. N., Drosten, C., Fouchier, R. A. M., Garcia, P. J., Halfmann, P. J., Jassem, A., Jones, T. C., Kawaoka, Y., Krammer, F., McDanal, C., Pajon, R., Simon, V., Stockwell, M., Tang, H., van Bakel, H., Webby, R., Montefiori, D. C., and Smith, D. J. (2022).
Mapping SARS-CoV-2 antigenic relationships and serological responses.
Preprint, Immunology.