Comparing antigenic distance metrics for influenza

EPID 8200 project

W. Zane Billings

2024-05-05

Abstract

Put the 50 words abstract here

Influenza A virus causes seasonal epidemics worldwide, primarily driven by continual evolution of the virus under selective pressure by host immunity [1]. Development of a universal influenza vaccine which can protect against novel strains of influenza has many challenges, including surveillence of new genomic variants and predicting which will be successful [2]. Statistical modeling and phylodynamic approaches are crucial tools in the development of a more broadly-protective influenza vaccine, but these methods rely on understanding how different each genomic variant of influenza actually is from its predecessors [3,**viboud2020?**].

Many different metrics for assessing the antigenic difference between two influenza strains currently exist, including phylogenetic methods [4], methods based on sequence differences [5,**anderson2018?**], and methods like antigenic cartrography which are based on observed immune responses [6]. To understand the agreement in phylogenetic and sequence-based approaches for determining the distance between influenza strains, we obtained data from an ongoing prospective, open cohort study that has been previously described [7,**abreau2020?**]. Using laboratory data from this study, we were able to compare antigenic cartography based on a panel of hemagglutination inhibition (HAI) titers to phylogenetic methods.

# The Study

Briefly, our study data [7,**abreau2020?**] consisted of volunteers enrolled at three different study sites from 2013 – 2019 who received a FluZone (Sanofi Pasteur) vaccine, and gave pre-vaccination and post-vaccination (21 or 28 day) serum samples. The serum samples were used for HAI assays against a panel of historical viruses. We computed the Hamming distance [**anderson2018?**], -Epitope distance [5], and the absolute difference in the year of isolation of strains [8] from the sequences of all influenza viruses used for HAI assays, and used Racmacs to compute antigenic cartography distances from the HAI data [9]. All of our analyses were conducted separately for H1N1 and H3N2 strains.

In order to compare with phylogenetic methods, we first computed a multiple sequence alignment (MSA) using the MUSCLE algorithm [10]. The Hamming and -Epitope distances were computed based on this MSA. In total, we had 18 H1N1 strains and 21 H3N2 strains which were aligned separately. We then used both alignments to construct maximum likelihood unrooted phylogenetic trees using the FLU amino acid substitution model. We extracted the conphenetic distances between taxa from the ML trees, and compared these distances to each of our other four distance metrics (temporal, Hamming, -Epitope, and cartography) using Pearson’s correlation.

For each of the four distance metrics, we also built distance-based trees using neighbor joining. To compare the methods, we calculated the likelihood of each of the distance-based trees under the same ML framework, then estimated the Shimodaira-Hasegawa test statistic (with one million bootstrap resamples) to compare each of the distance trees to the ML tree. Finally, we computed the Robinson-Foulds distance between each set of trees. Our analyses were implented with R version 4.3.3 (2024-02-29 ucrt) [11] using the packages phangorn [12] and msa [13].

Figure 1: correlation plot between distance metrics, show pearson and spearman correlations

Figure 2: main trees Panel A: ML tree, panel B: 4 subpanels with the distance trees, if it fits. Otherwise we have space for one more figure.

Table 1: table of tree similarity and LRT values

# Conclusions

# Acknowledgement

# References

1. Kim H, Webster RG, Webby RJ. Influenza Virus: Dealing with a Drifting and Shifting Pathogen. Viral Immunology. 2018;31: 174–183. doi:[10.1089/vim.2017.0141](https://doi.org/10.1089/vim.2017.0141)

2. Erbelding EJ, Post DJ, Stemmy EJ, Roberts PC, Augustine AD, Ferguson S, et al. A Universal Influenza Vaccine: The Strategic Plan for the National Institute of Allergy and Infectious Diseases. The Journal of Infectious Diseases. 2018;218: 347–354. doi:[10.1093/infdis/jiy103](https://doi.org/10.1093/infdis/jiy103)

3. Krammer F, Smith GJD, Fouchier RAM, Peiris M, Kedzierska K, Doherty PC, et al. Influenza. Nature Reviews Disease Primers. 2018;4: 1–21. doi:[10.1038/s41572-018-0002-y](https://doi.org/10.1038/s41572-018-0002-y)

4. Durviaux S, Treanor J, Beran J, Duval X, Esen M, Feldman G, et al. Genetic and Antigenic Typing of Seasonal Influenza Virus Breakthrough Cases from a 2008-2009 Vaccine Efficacy Trial. Clinical and Vaccine Immunology. 2014;21: 271–279. doi:[10.1128/CVI.00544-13](https://doi.org/10.1128/CVI.00544-13)

5. Gupta V, Earl DJ, Deem MW. Quantifying influenza vaccine efficacy and antigenic distance. Vaccine. 2006;24: 3881–3888. doi:[10.1016/j.vaccine.2006.01.010](https://doi.org/10.1016/j.vaccine.2006.01.010)

6. Fonville JM, Fraaij PLA, de Mutsert G, Wilks SH, van Beek R, Fouchier RAM, et al. Antigenic Maps of Influenza A(H3N2) Produced With Human Antisera Obtained After Primary Infection. The Journal of Infectious Diseases. 2016;213: 31–38. doi:[10.1093/infdis/jiv367](https://doi.org/10.1093/infdis/jiv367)

7. Nuñez IA, Carlock MA, Allen JD, Owino SO, Moehling KK, Nowalk P, et al. Impact of age and pre-existing influenza immune responses in humans receiving split inactivated influenza vaccine on the induction of the breadth of antibodies to influenza A strains. PloS One. 2017;12: e0185666. doi:[10.1371/journal.pone.0185666](https://doi.org/10.1371/journal.pone.0185666)

8. Auladell M, Phuong HVM, Mai LTQ, Tseng Y-Y, Carolan L, Wilks S, et al. Influenza virus infection history shapes antibody responses to influenza vaccination. Nature Medicine. 2022;28: 363–372. doi:[10.1038/s41591-022-01690-w](https://doi.org/10.1038/s41591-022-01690-w)

9. Wilks S. Racmacs: Antigenic cartography macros. 2023.

10. Edgar RC. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 2004;5: 113. doi:[10.1186/1471-2105-5-113](https://doi.org/10.1186/1471-2105-5-113)

11. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2024.

12. Schliep KP. Phangorn: Phylogenetic analysis in R. Bioinformatics (Oxford, England). 2011;27: 592–593. doi:[10.1093/bioinformatics/btq706](https://doi.org/10.1093/bioinformatics/btq706)

13. Bodenhofer U, Bonatesta E, Horejš-Kainrath C, Hochreiter S. Msa: An R package for multiple sequence alignment. Bioinformatics. 2015;31: 3997–3999. doi:[10.1093/bioinformatics/btv494](https://doi.org/10.1093/bioinformatics/btv494)