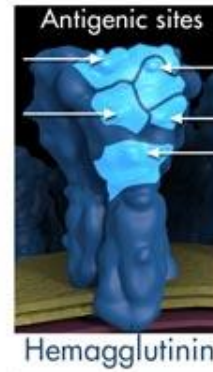
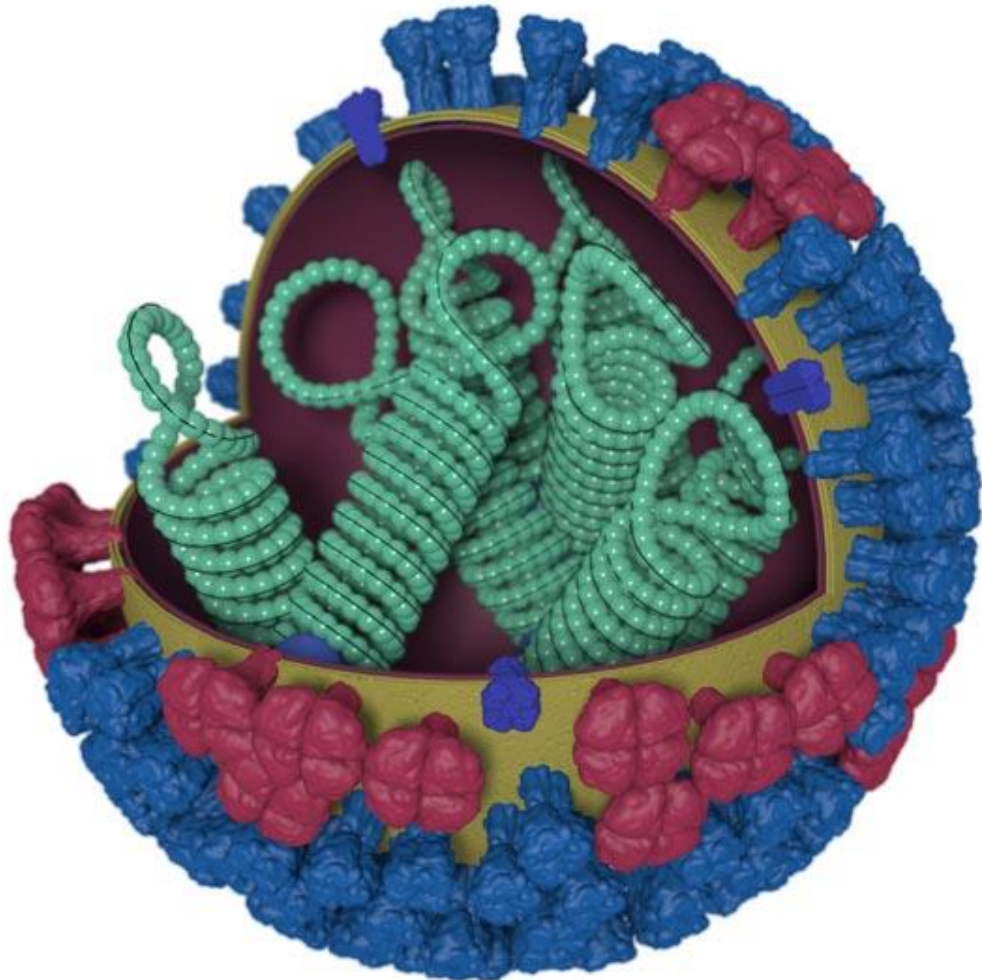


Mapping the Antigenic and Genetic Evolution of Influenza Virus

**Derek J. Smith, Alan S. Lapedes, Jan C. de Jong, Theo M. Bestebroer,
Guus F. Rimmelzwaan, Albert D. M. E. Osterhaus, Ron A. M. Fouchier**
Science 305, 371 (2004)

O. Liota Weinbaum: biology
Brendan Donahue: methods
Esther Rodriguez: results
Audience: discussion

AN INFLUENZA VIRUS



“antibody binding sites”
“antigenic sites”
“antigens”

Molecular regions that the immune system can recognize. One molecule may have many recognizable regions.



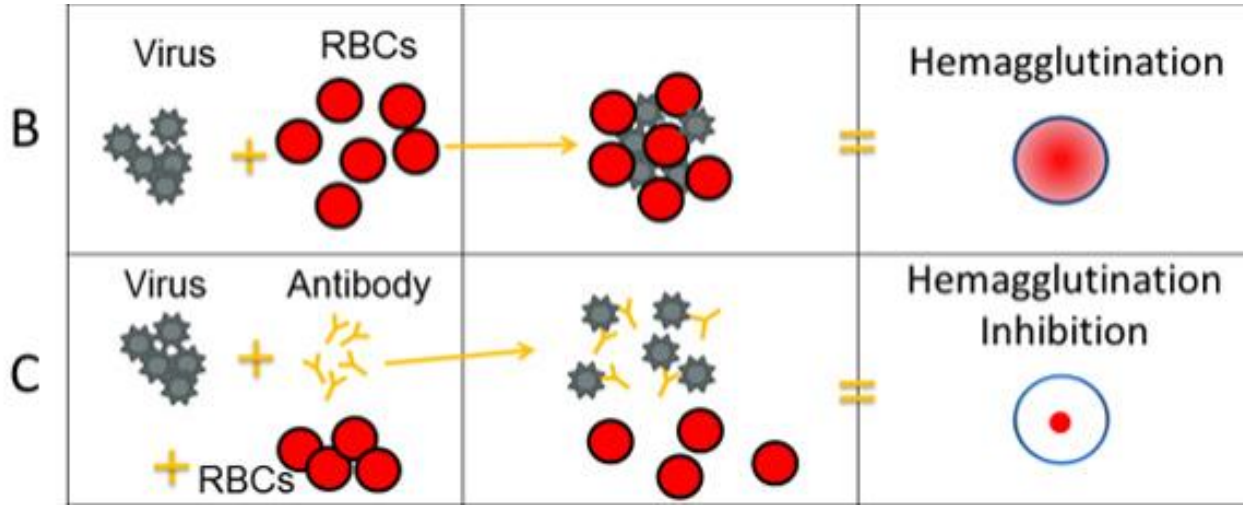
Recognition means creating a complementary molecule that will bind to the antigen. The complimentary molecule is called an antibody.



When many antibodies bind to a pathogen, the pathogen becomes disabled.



HI assay



Row B: when an influenza virus is added to the RBC solution, the virus' **hemagglutinin (HA)** will bind to RBCs, forming a lattice structure that keeps the RBCs suspended in solution instead of sinking to the bottom. This is called **hemagglutination**.

Row C: antibodies that are complementary to a virus being tested will bind to that virus. This prevents the virus and RBCs from binding, and therefore, hemagglutination does not occur: **hemagglutination inhibition** occurs instead (**HI**). If antibodies that resulted from exposure to the comparison vaccine bind to the influenza virus from a sick patient, this indicates that the comparison virus is antigenically similar to the influenza virus obtained from the sick patient. If not, HI won't occur and the result will look more like Row B.

The higher the dilution, the fewer antibodies are needed to block hemagglutination, the more antigenically similar the two viruses being compared are to each other.

HI assay checks for functionality of specific Ab to specific viral Ag measured as the reactivity in serial dilutions

Ag map relates based on this phenotype

ML tree based on **genotype**

HAI domain is section of viral genome that codes for the antigens in question
evolutionary history from most ancestral to most recent
statistics, based on parsimony

ML tree based on amino acid (AA) substitutions

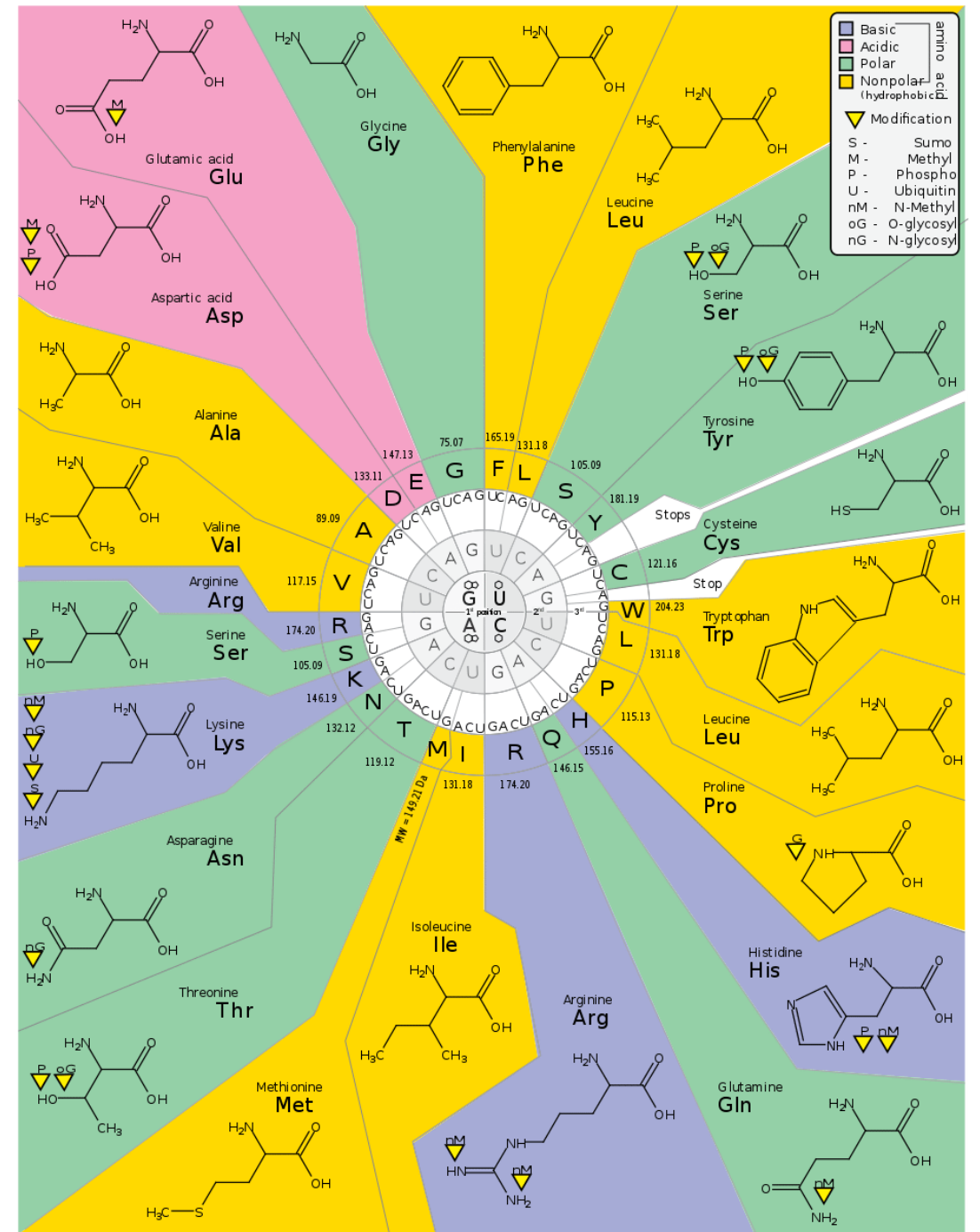
a **bridge between phenotype and genotype**

ignores changes in genotype that don't change AAs
includes AA changes that don't change HI **phenotype**

Contradictory results of the three characterizations reflect **networks** and **neighborhoods (Wagner 2012)**:

- changes in ncDNA
- redundancy of amino acid language
- changes in amino acids that don't change molecular properties

Wikicommons



Quantitative Analysis of Antigenic Data

- Antigenic distance largely believed to be the driving factor behind the evolution of the influenza virus.
- As such, antigenicity is a primary criterion for strain selection for the current influenza season.
- However, while methods to measure antigenic distance have provided insight into the virus's evolution, quantitative analysis through techniques such as, e.g., numerical taxonomy, has proven difficult due to, e.g., misinterpretation of results below sensitivity threshold of assay data, approximation of antigenic distances in an indirect way.

Generation of the Antigenic Map

- Initially proposed by Lapedes and Farber (2001).
- Represent individual antisera and viral strains as two sets of points in a two-dimensional grid, or “shape space”, whose positions on this grid are initially unknown.
- The distance between a single strain and a single antiserum on this grid corresponds to the logarithm of the HI measurement from the assay.

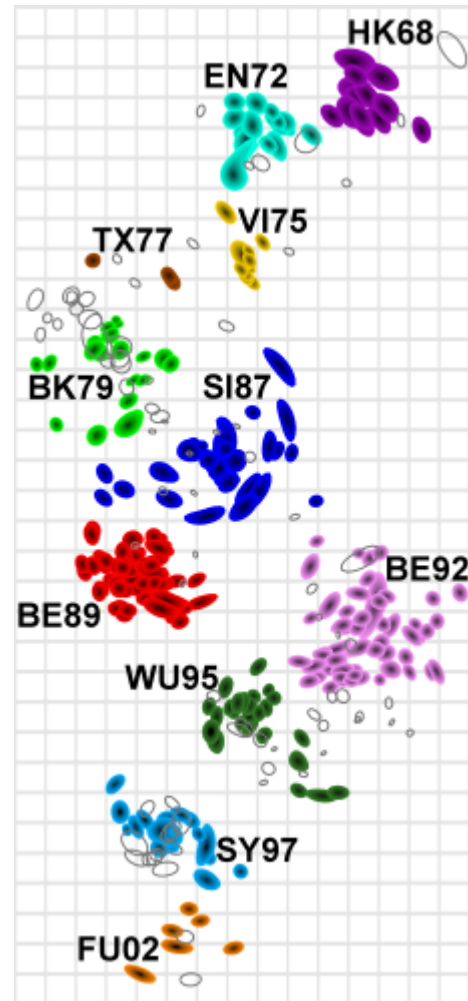
Generation of the Antigenic Map (contd.)

- Perform the HI Assay on a sufficient number of antiserum-antigen pairs, and record the logarithm of their HI measurements.
- Apply multidimensional scaling (MDS) which (per Wikipedia) given information about pairwise distances (i.e., the logarithm of the HI measurements) between various points, will attempt to map these points to a grid, such that their distances most closely correspond to the input distances.
- A form of “dimensionality reduction” - projecting data from a higher dimensional space to a lower dimensional (and more easily interpreted) space.

Generation of the Antigenic Map (contd.)

- Smith et al. applied this method to 79 antisera and 273 influenza viral isolates taken from 1968 to 2003, and constructed a training dataset of 4215 (out of 21567 possible) individual HI measurements between various antiserum-antigen pairs.
- The map was constructed from this dataset, and the resolution (or accuracy) of the map was determined by sampling 481 pairs not represented in the original dataset, computing their HI value using the assay, and comparing it with that calculated from the map.
- k-clustering is also performed to identify clusters in the map and mark them with the date of the first strain in each cluster.

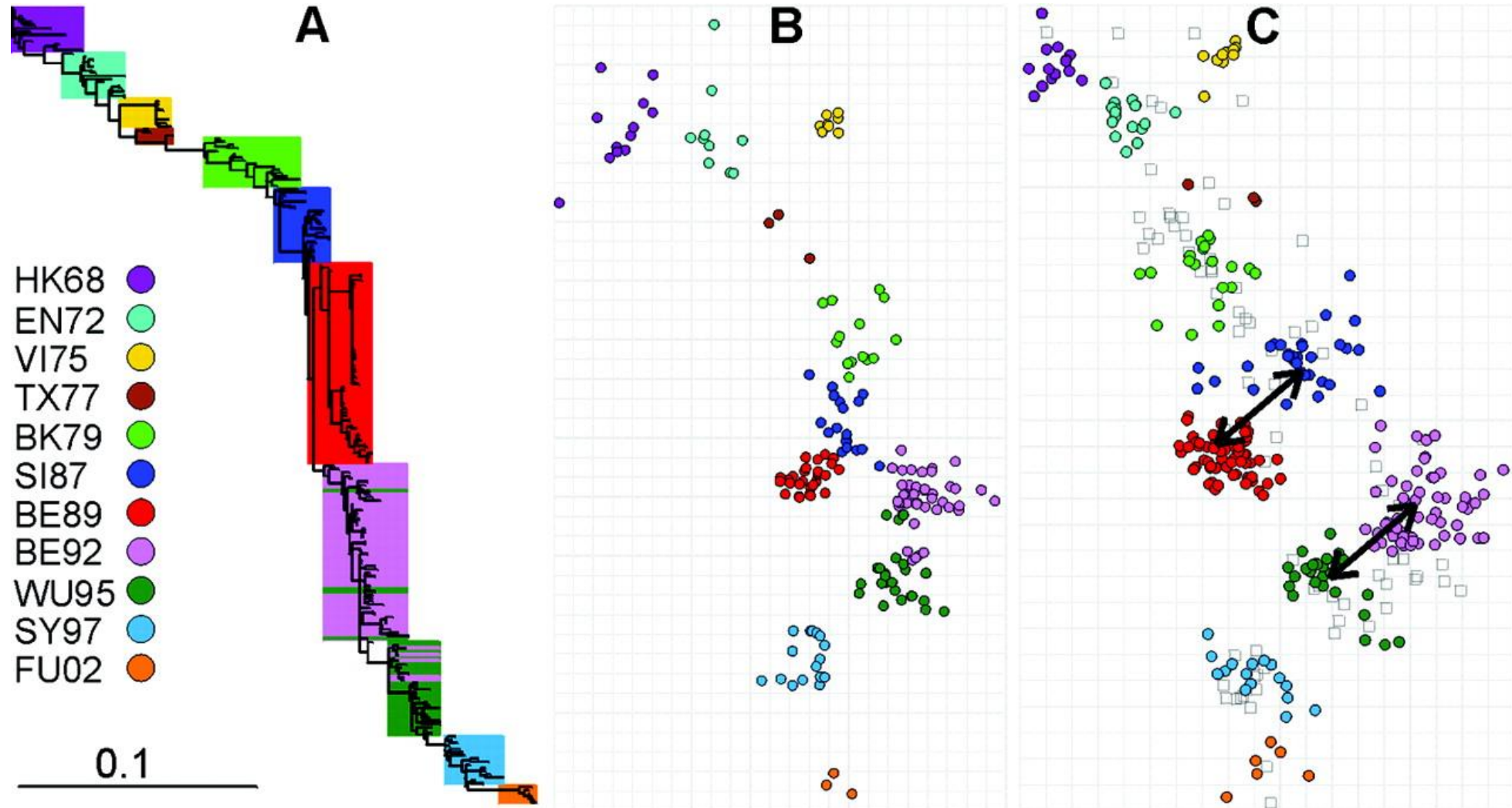
Generation of the Antigenic Map (contd.)



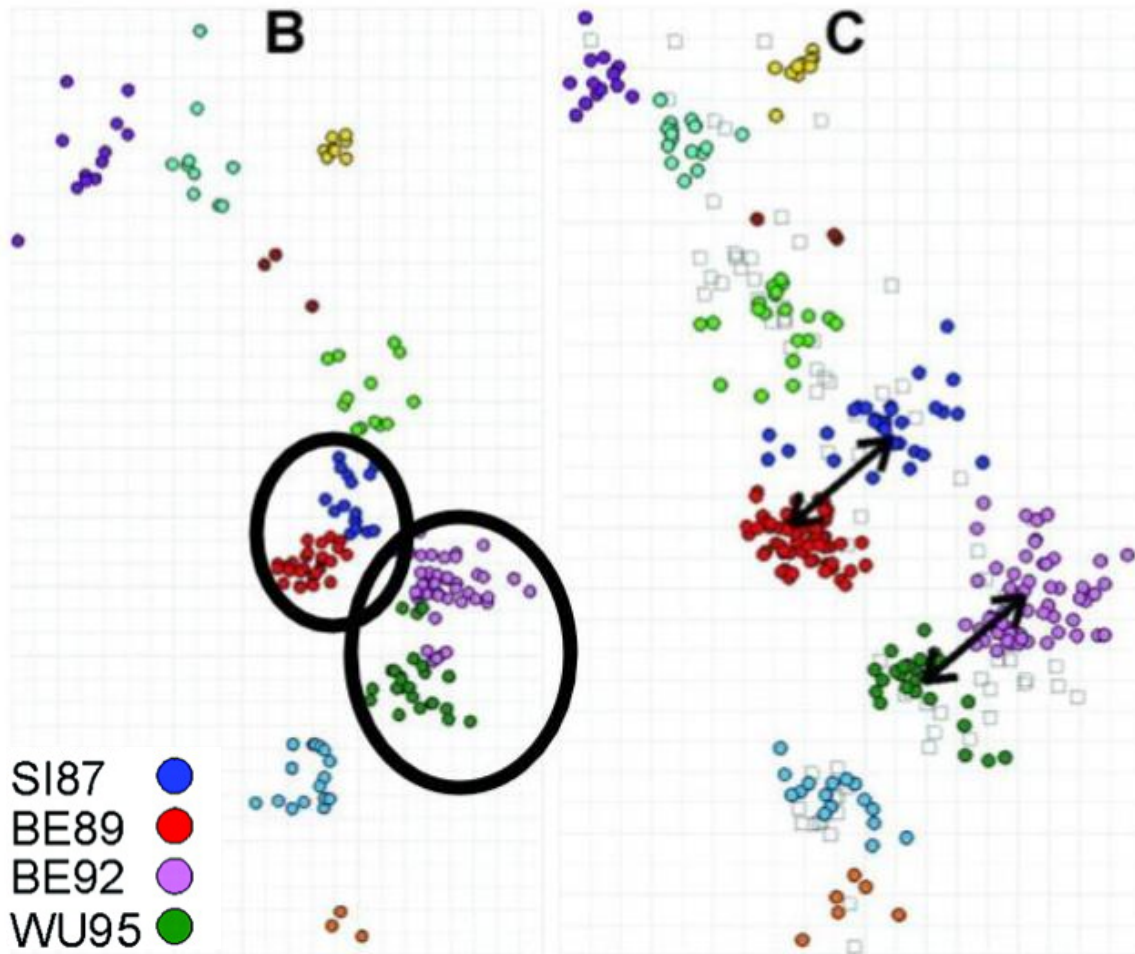
Analysis of the Antigenic Map

- Once the antigenic clusters were identified, it became possible to identify the individual amino acid substitutions that significantly contributed to antigenic differences between clusters.
- The location and types of each substitution between cluster centroids, as well as the antigenic distance (in antigenic map units between clusters) and the genetic distance (in the number of amino acid substitutions between strains) are recorded.
- Genetic distances are then computed for several different pairs of strains (not just cluster centroids) in the dataset.
- A maximum-likelihood phylogenetic tree and genetic map are generated from this data to record the evolution of the virus genome and compare genetic and antigenic evolution.

Phylogenetic Tree (A) and Genetic Map (B)



Comparison of Genetic and Antigenic Evolution

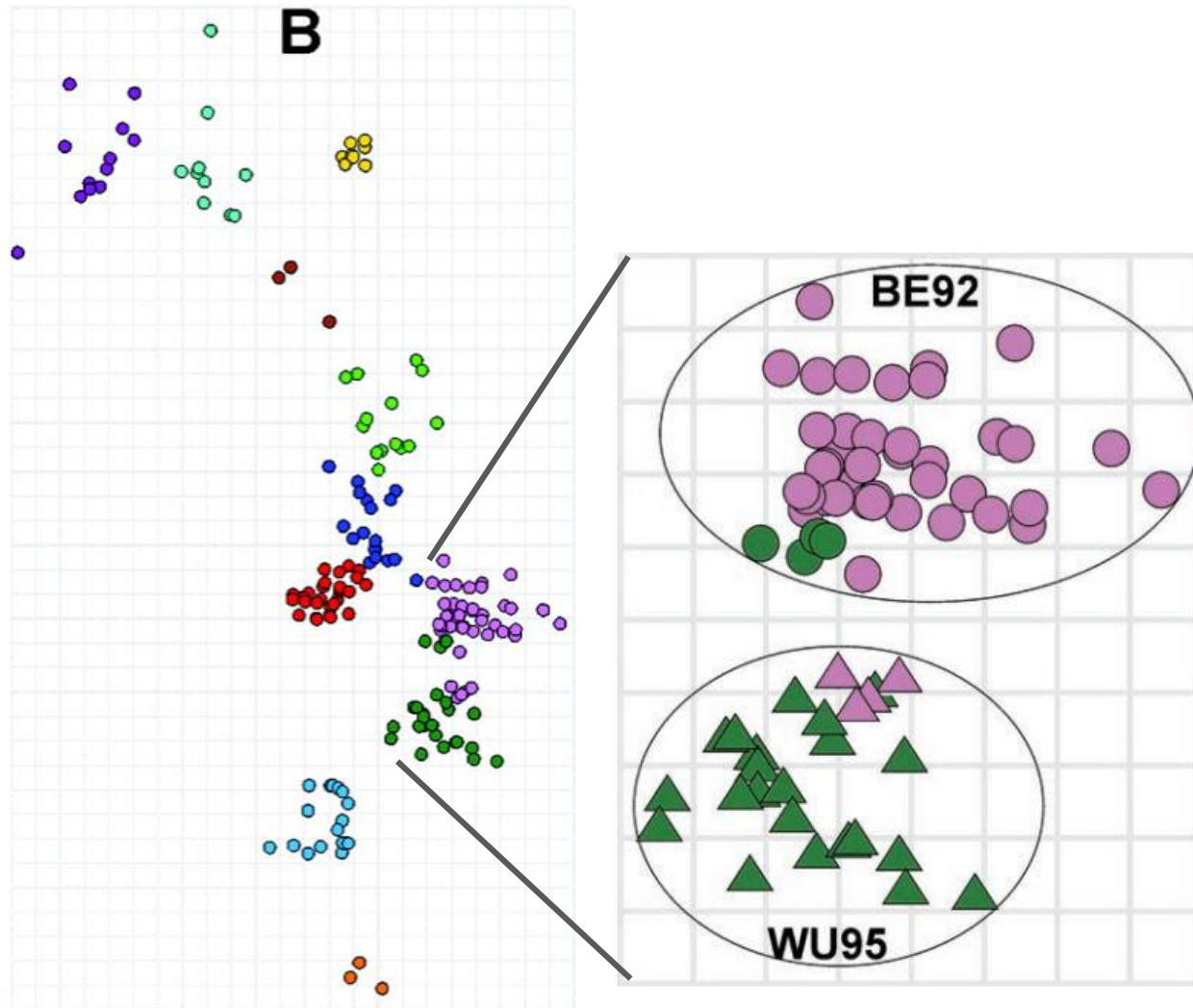


B) Genetic Map

C) Antigenic Map

- Sichuan 1987 (blue) and Beijing 1989 (red) are genetically closely related but antigenically distinct.
- A single amino acid substitution, N145K, is the only cluster-difference substitution between the SI87 (blue) and BE89 (red) clusters.
- On average, a single amino acid substitution causes only 0.37 units of antigenic change.
- Cluster transitions are characterized by multiple cluster-difference substitutions.

N145K has a Large Antigenic Effect

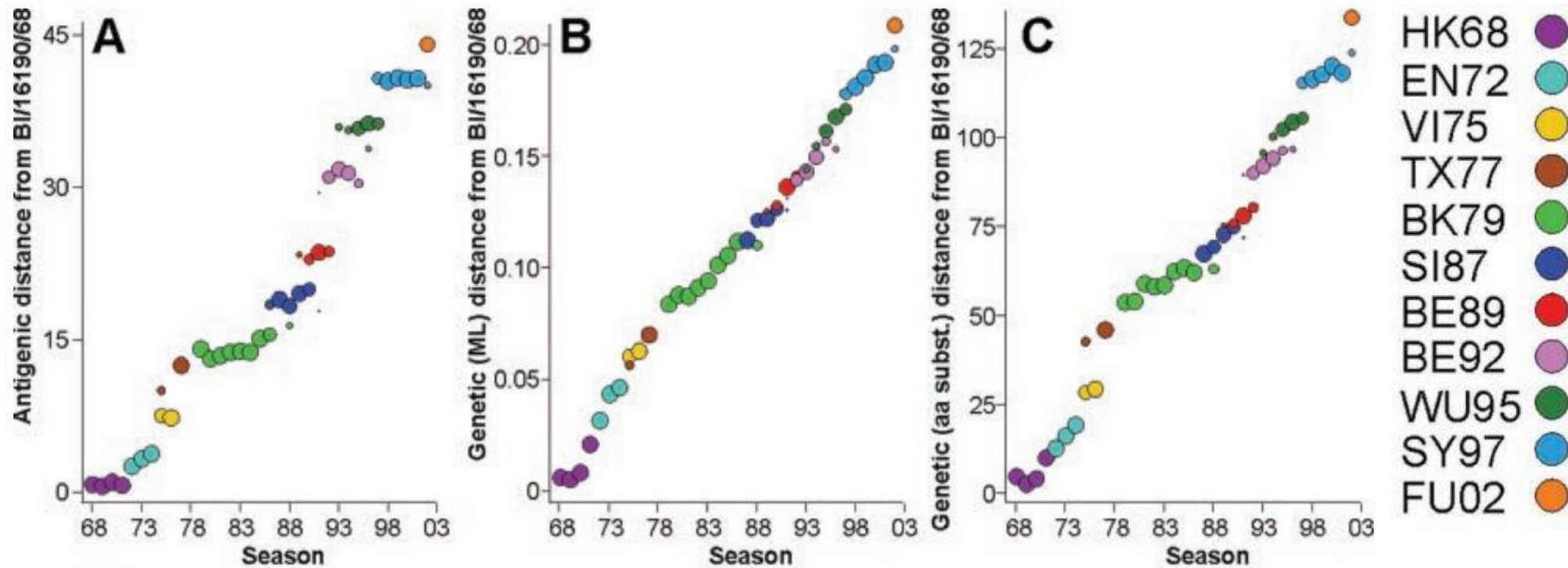


B) Genetic map

Green symbols have **lysine (K)** at position 145 and pink symbols have **asparagine (N)** at 145.

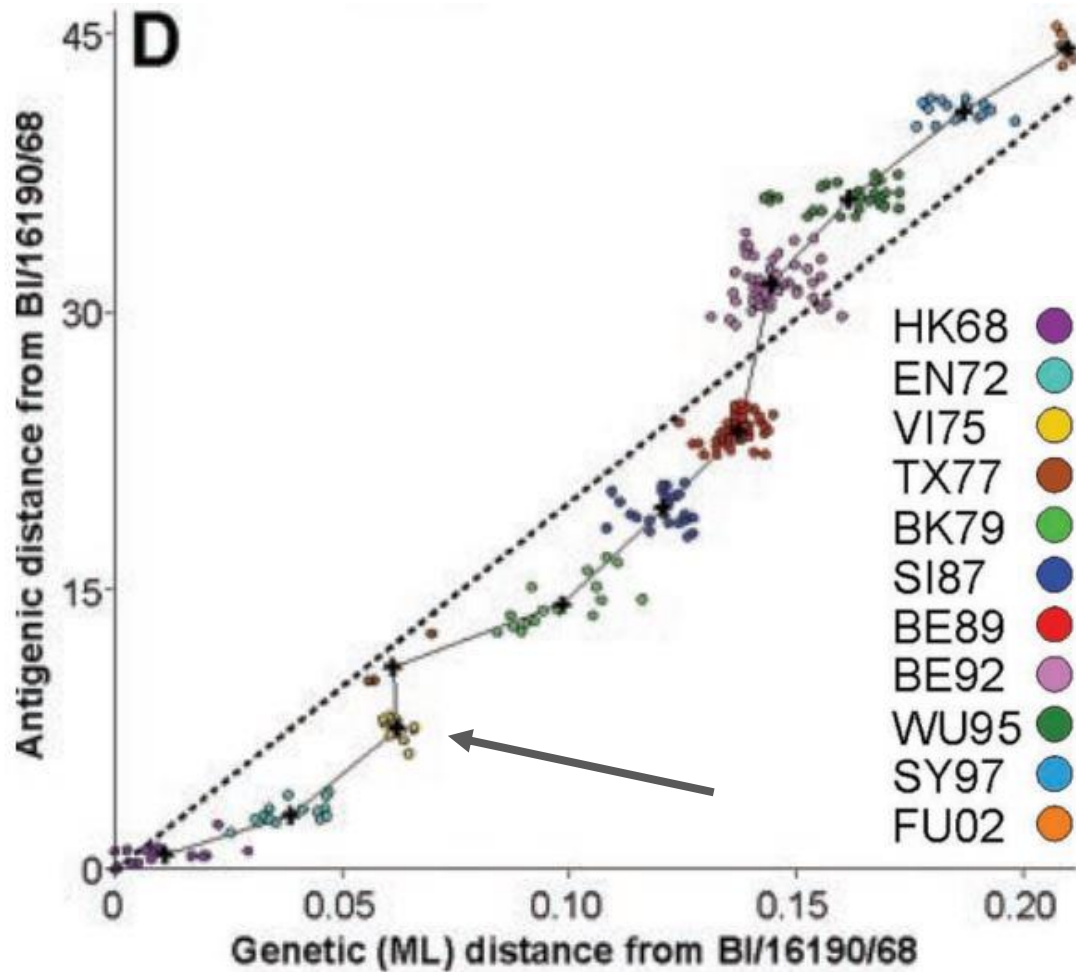
- N145K average antigenic distance is four times the average antigenic distance for other amino acid substitutions at the same position (I145S, N145S) and the same substitution at a different position (N92K).
- There were nine strains in the genetic map for which the genetic cluster did not correspond with the antigenic cluster and N145K was responsible for the difference.

Season-by-season Analysis



- We observe punctuated antigenic evolution with more continuous genetic evolution.
- This suggests that some amino acid substitutions have little antigenic effect or an effect spreading the cluster sideways.

Comparison of Antigenic and Genetic Evolution



- Sometimes the rate of antigenic evolution was faster than genetic evolution.
- Clusters that move away linearly from previous clusters, most effectively escape existing population-level immunity.

Results

- Antigenic evolution is clustered.
- There is a higher rate of antigenic evolution between clusters than within clusters.
- Antigenic evolution is punctuated and genetic evolution is more continuous.
- Remarkable correspondence between antigenic and genetic evolution, with important exceptions of epidemiological significance.
- This quantification of antigenic data provides opportunity for analyses which integrate selection at the phenotypic level with genetic change at the level of individual amino acids.

Results

- These methods increase the value of surveillance data and facilitate vaccine strain selection.
- If immune-escape is a dominant aspect of total fitness, antigenic maps could predict which strains would be more likely to seed a new epidemic.
- Antigenic maps may help increase the efficacy of repeated vaccination by accounting quantitatively for the antigenic distances among vaccine and circulating strains.
- The same methods were applied to the characterization of human H1N1, swine H3N2, and equine H3N8 influenza A viruses as well as human influenza B virus.
- These methods are expected to be useful for a variety of antigenically variable pathogens including HIV and hepatitis C virus.

Discussion questions

“Sometimes the rate of antigenic evolution was faster than genetic evolution...”

Usually not, but, how could this ever occur?

Why is it “remarkable” for there to be high correlation between genetic change and phenotypic change?

References

D. J. Smith et al., "Mapping the antigenic and genetic evolution of influenza virus," *Science*, vol. 305, no. 5682, pp. 371–376, 2004.

2017. Centers for Disease Control and Prevention, National Center for Immunization and Respiratory Diseases. Seasonal Influenza (Flu): About Flu: Understanding Influenza Viruses: Antigenic Characterization. <https://www.cdc.gov/flu/professionals/laboratory/antigenic.htm>

A. Wagner, "The role of robustness in phenotypic adaptation and innovation," *Proc. R. Soc. B Biol. Sci.*, vol. 279, no. January, pp. 1249–1258, 2012.

2019. Wikimedia Foundation. Wikipedia: Multidimensional Scaling. http://en.wikipedia.org/wiki/Multidimensional_scaling

A. Lapedes, R. Farber, "The geometry of shape space: application to influenza," *J. Theor. Biol.*, 212, 57 (2001).