

# **Phylogenetics for Predicting Virus Evolution**

---

**Can we anticipate next seasons dominant influenza strain from sequence alignment?**

**Paul Wiesemeyer**

## Problem



HOUSTON,  
we have a — CHOO!

- Influenza evolves rapidly and evades immunity and vaccines
- Estimated 3 – 11% of the population catch symptomatic influenza each season [Tokars et al., 2018]
- RNA recombination events can cause sudden epidemics and even pandemics
- Science needs to respond timely to health issues

# Problem

- About 8 months from submission to publication of a medical paper [AAMC, 2018]
- About 6 months from vaccine strain selection to distribution:

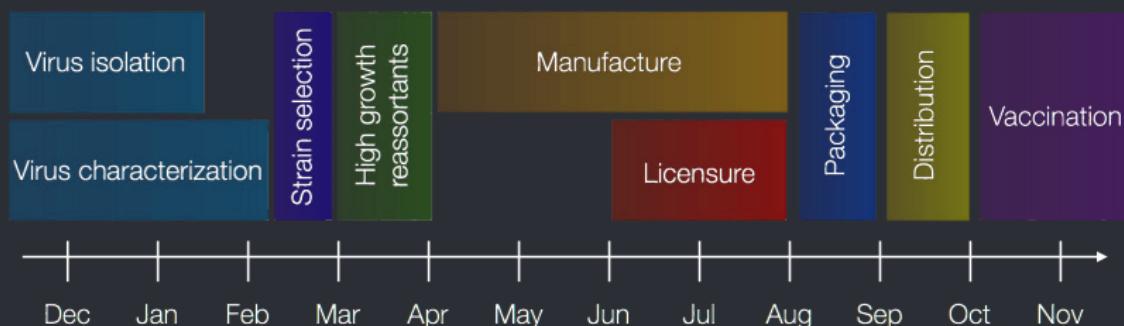


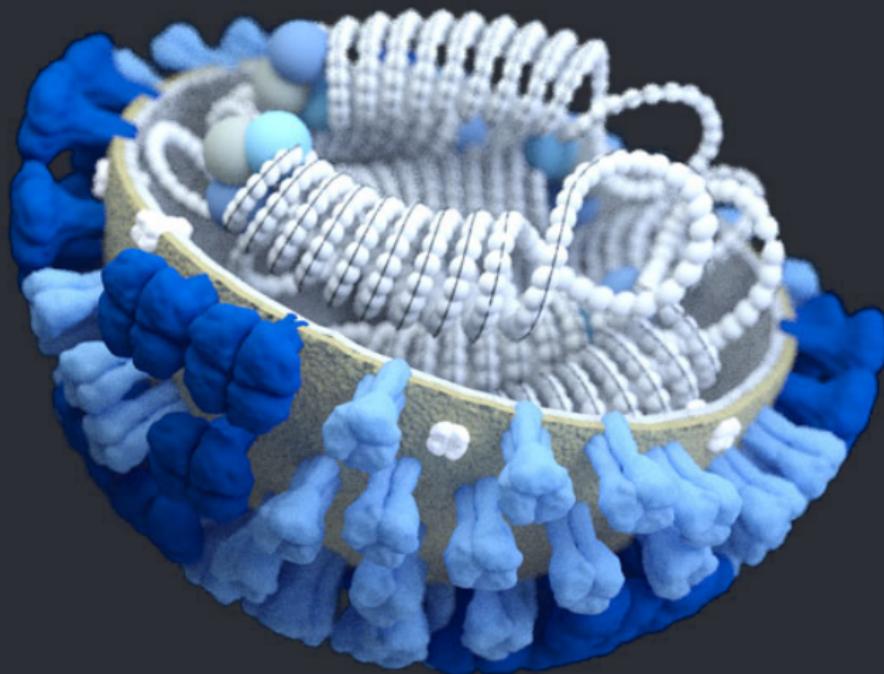
Figure: [Bedford, 2015]

# Outline

1. Influenza and Vaccines
  - 1.1 Basics
  - 1.2 The Hemagglutination Inhibition Assay
2. Predicting the next strain of influenza
  - 2.1 Results
3. Phylogenetics
  - 3.1 The Molecular Clock
  - 3.2 The Molecular Clock
  - 3.3 Sequence Alignment
  - 3.4 Building A Phylogenetic Tree
4. *Nextstrain*
  - 4.1 How to use the Framework
  - 4.2 The Powerful Meta Data
  - 4.3 Confidence Levels and Limitations

# Influenza and Vaccines

*Influenza—an artful disguise master*



## Basics

*Where does influenza come from?*

- *Host switch events*, the jumping from one species to another.
- *Antigenic shift events*, when different virus types infect the same host cell and recombine their 8 RNA parts.
- *Antigenic drift*, driven by point mutations and *indels*

# Basics

*Antigenic shifts and zoonoses can cause pandemics*

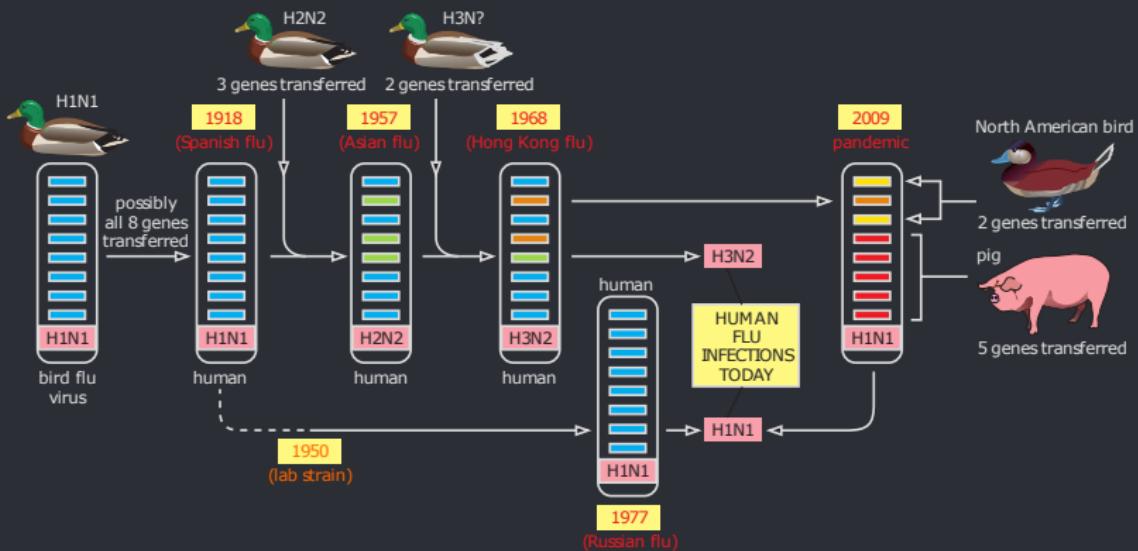


Figure: taken from [Alberts, 2015]

## Basics

*Antigenic drift—a race*

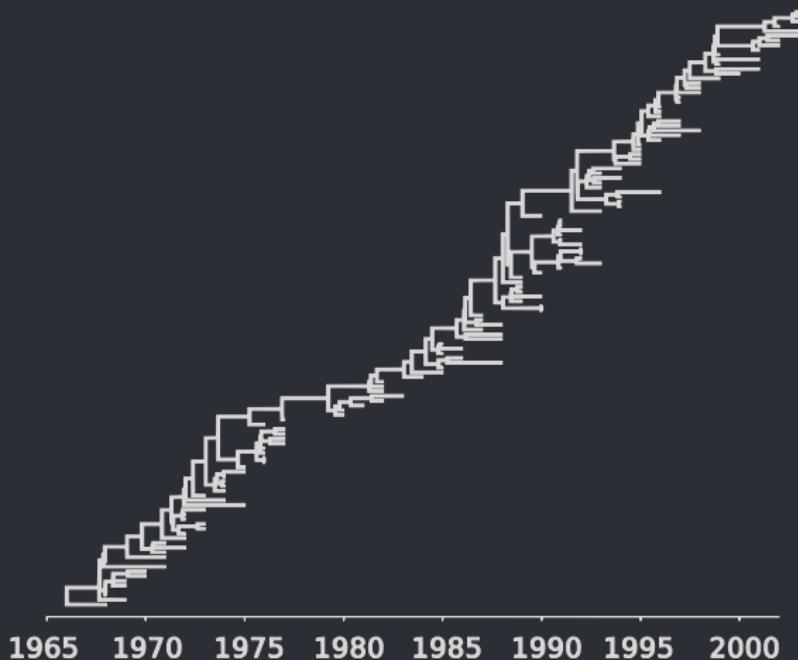


Figure:

# Basics

*Antigenic drift—a race*



Figure: It's like getting more and more stars, having to switch escape vehicle

# The Hemagglutination Inhibition Assay

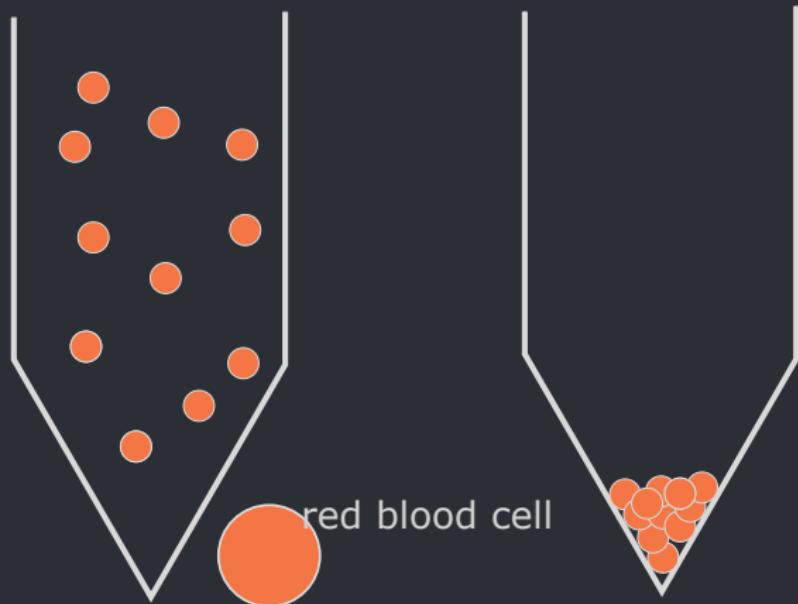


Figure: Red blood cells (RBC) precipitate.

# The Hemagglutination Inhibition Assay

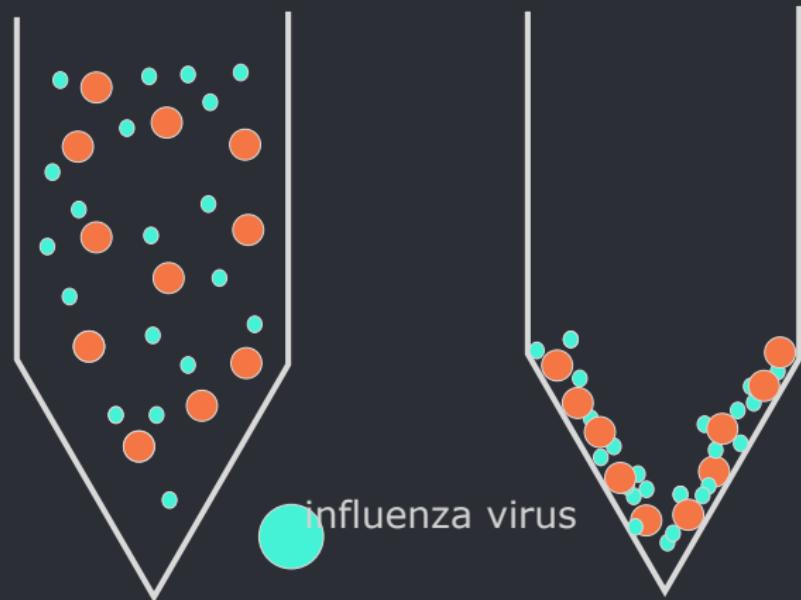


Figure: Influenza Hemagglutinin (HA) coagulates the RBC, forming a mat.

# The Hemagglutination Inhibition Assay

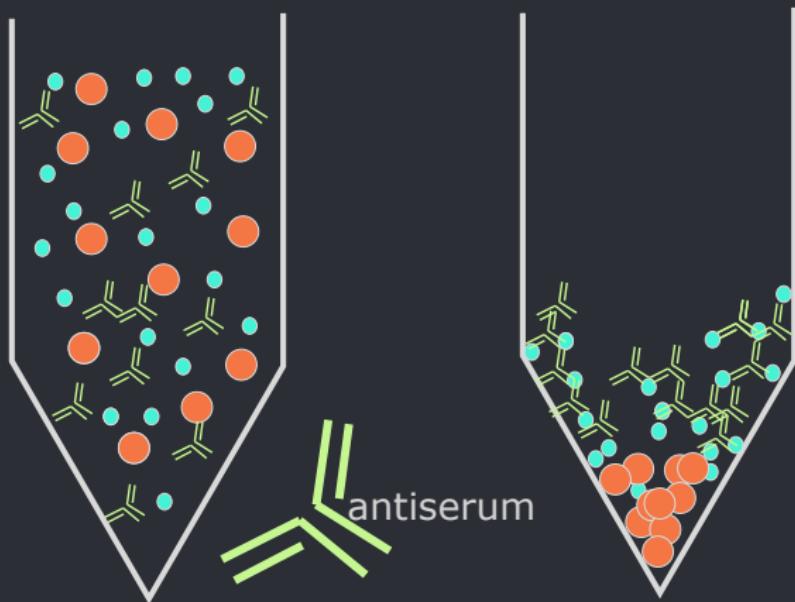


Figure: Antisera of the same serotype clump the HA, letting the RBC sink to the bottom. This is an (antiserum) concentration dependent process.

# The Hemagglutination Inhibition Assay

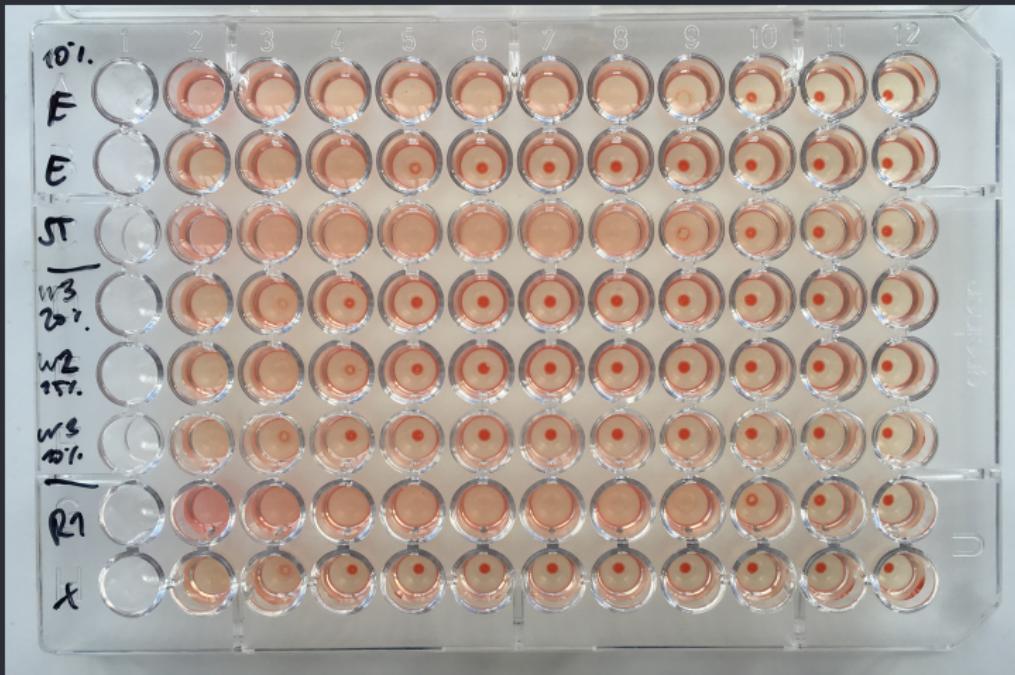
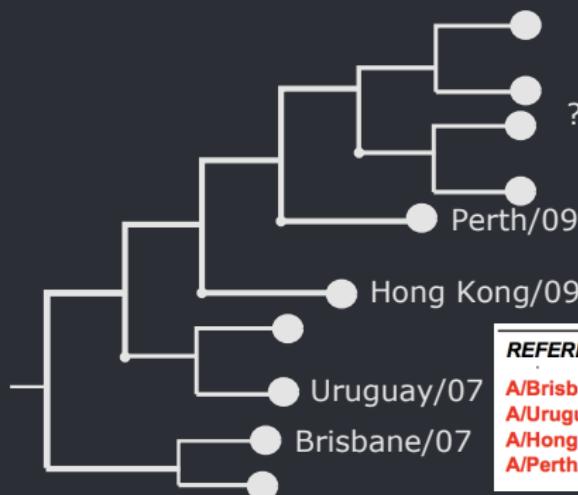


Figure: Here, one antiserum is tested in 12 different dilutions against 8 different virus strains. The highest dilution that prevents agglutination is called the titer.

# The Hemagglutination Inhibition Assay

*How this used to be looked at*

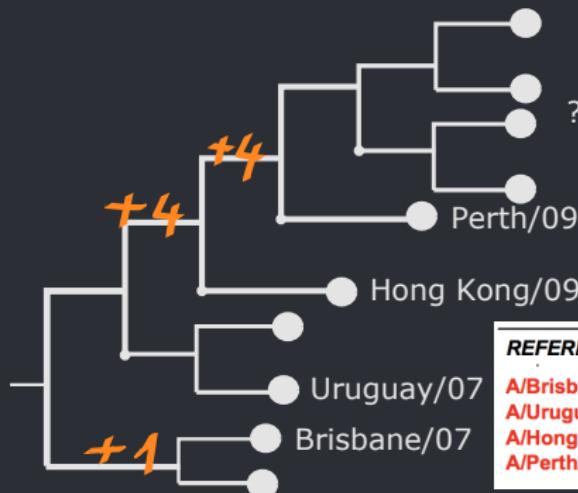
# The Hemagglutination Inhibition Assay



REFERENCE VIRUSES	A/Bris	A/Uru	A/HK	A/Perth
A/Brisbane/10/2007	2560	2560	80	<
A/Uruguay/716/2007	1280	2560	<	<
A/Hong Kong/1985/2009	80	160	1280	640
A/Perth/16/2009	<	40	640	640

Figure: Here, one antiserum is tested in 12 different dilutions against 8 different virus strains. The highest dilution that prevents agglutination is called the titer.

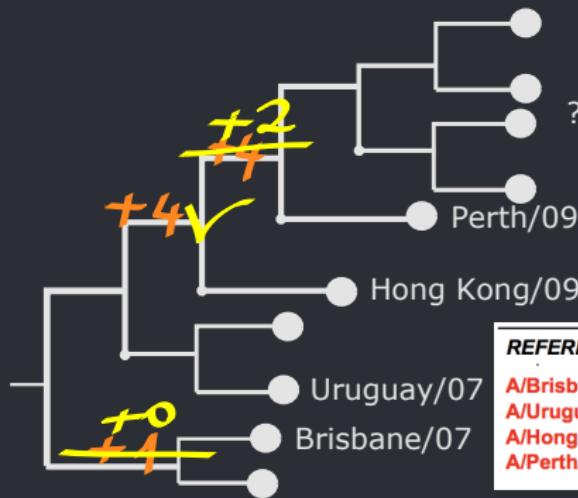
# The Hemagglutination Inhibition Assay



REFERENCE VIRUSES	A/Bris	A/Uru	A/HK	A/Perth
A/Brisbane/10/2007	2560	2560	80	<
A/Uruguay/716/2007	1280	2560	<	<
A/Hong Kong/1985/2009	80	160	1280	640
A/Perth/16/2009	<	40	640	640

Figure: Here, one antiserum is tested in 12 different dilutions against 8 different virus strains. The highest dilution that prevents agglutination is called the titer.

# The Hemagglutination Inhibition Assay

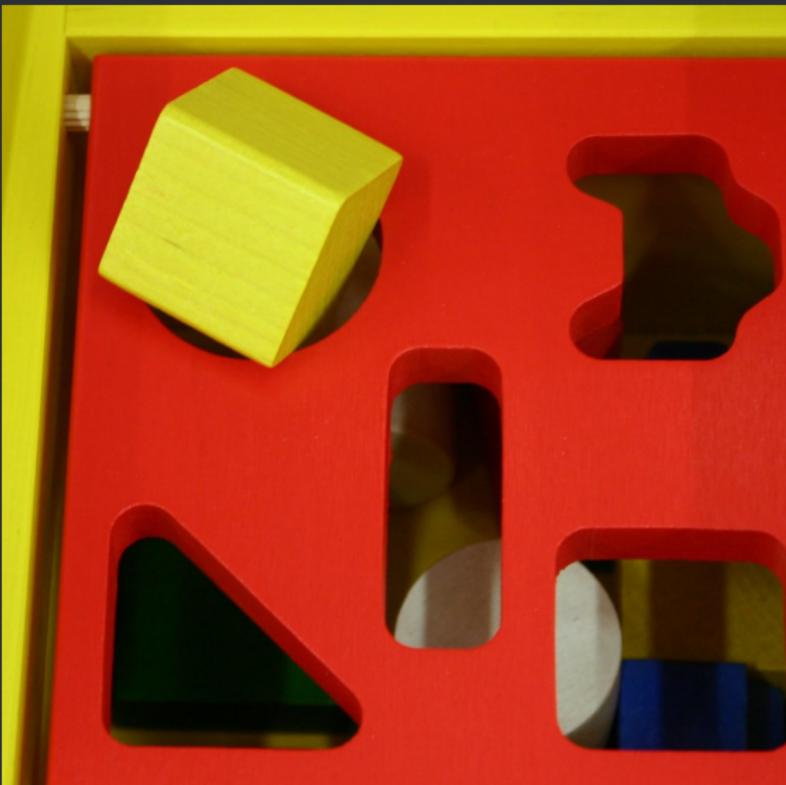


REFERENCE VIRUSES	A/Bris	A/Uru	A/HK	A/Perth
A/Brisbane/10/2007	<b>2560</b>	2560	80	<
A/Uruguay/716/2007	1280	<b>2560</b>	<	<
A/Hong Kong/1985/2009	80	160	<b>1280</b>	640
A/Perth/16/2009	<	40	640	<b>640</b>

Figure: Here, one antiserum is tested in 12 different dilutions against 8 different virus strains. The highest dilution that prevents agglutination is called the titer.

# The Hemagglutination Inhibition Assay

*The asymmetry makes sense, think of it like this but with more dimensions*



## Multi Scale Evolution

If a single event mutation occurs, say **D 186 G** in the HA genome, it is subject to multi scale evolutionary selection:

- this RNA instance vs. the other RNA strands in the same cell
- this cell's mutated viruses vs. other viruses inside the host
- this host's viruses vs. viruses in rest of the population
- this population vs. other populations

These scales are difficult to separate. At the population level *epidemiological* processes may dominate.

# Results

## *cumulative antigenic change*

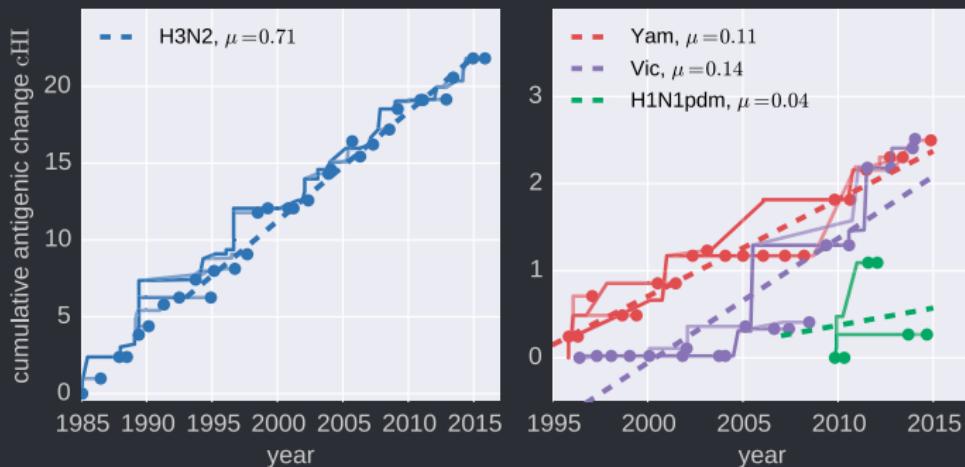


Figure: a

# Results

## *tree model*

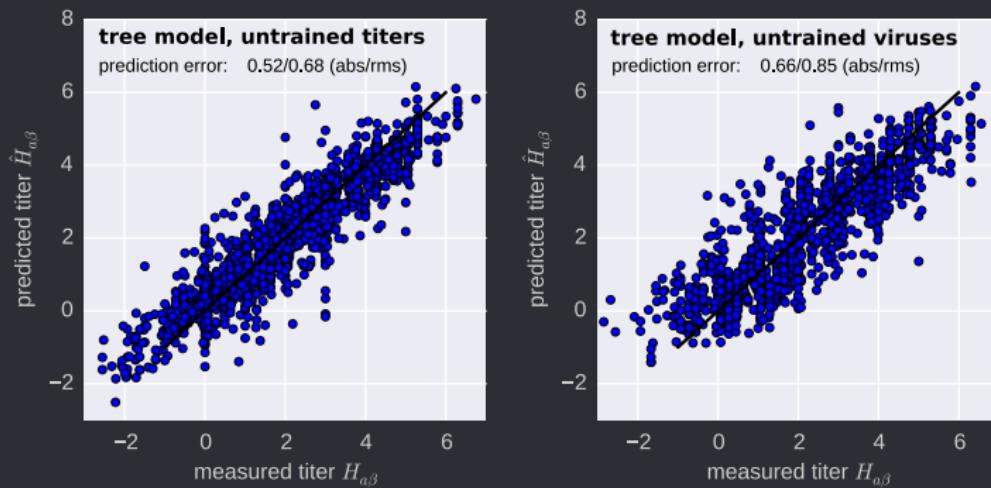


Figure: a

# Results

## *tree model hist*

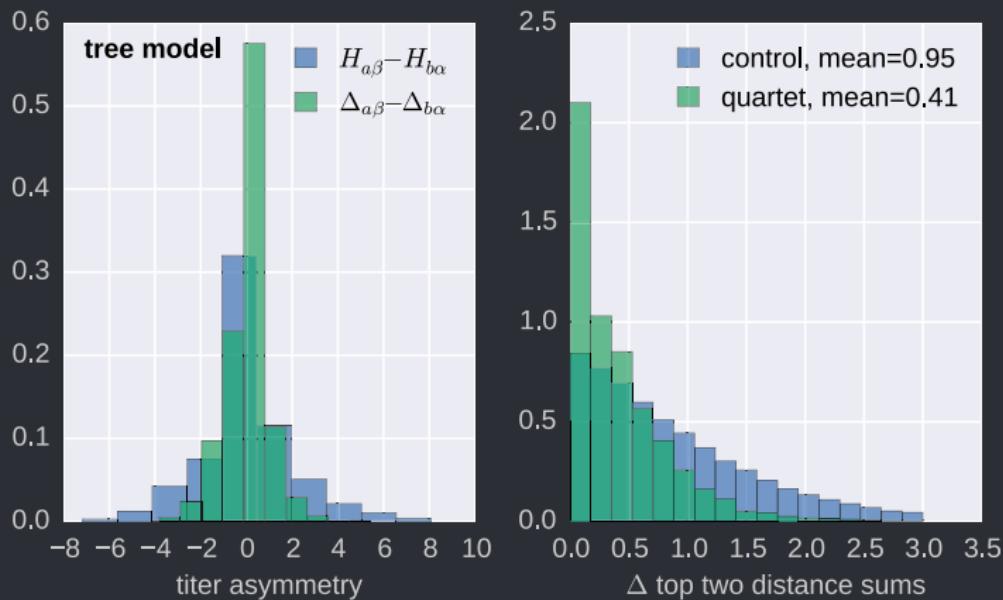


Figure: a

# Results

## *recent antigenic evolution*

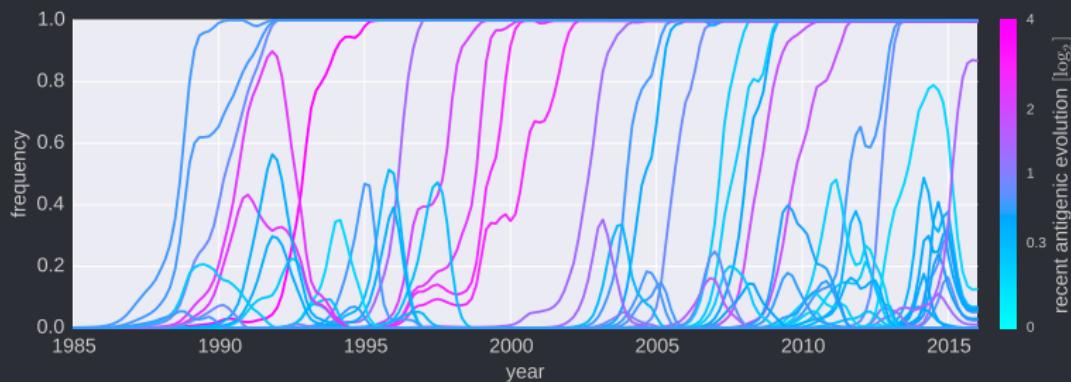


Figure: a

# Results

*fraction frequency threshold color bar*

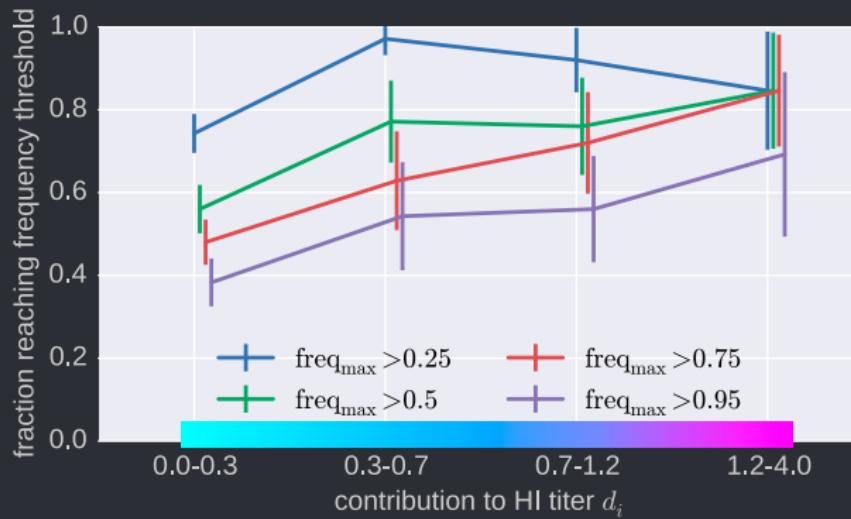


Figure: a

## Results

*distance to season year*

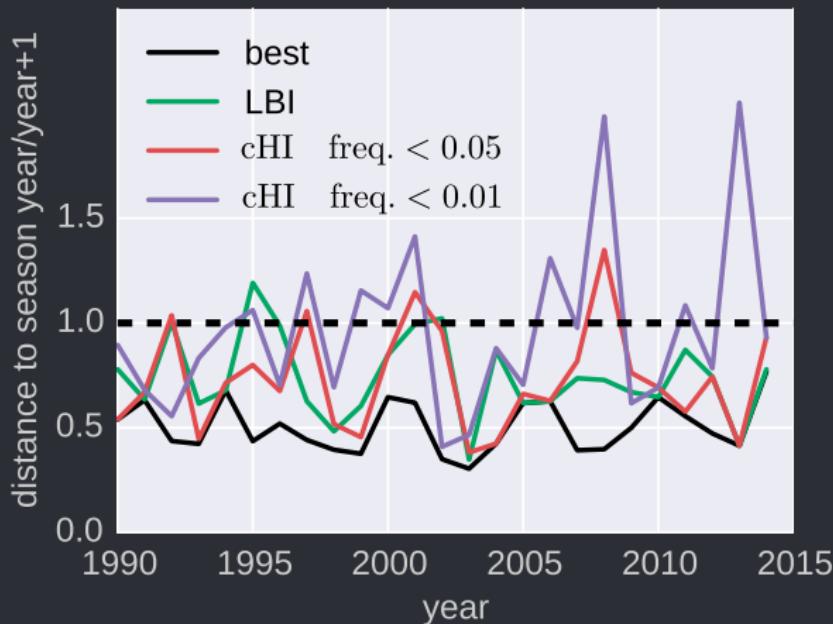


Figure: a

## Results

*years antigenic advance*

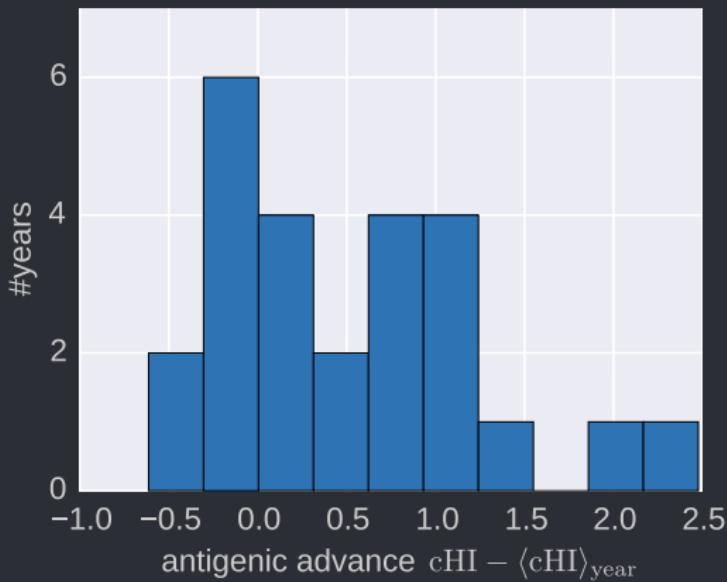


Figure: a

# Phylogenetics

*An old idea*

# Phylogenetics: The Molecular Clock

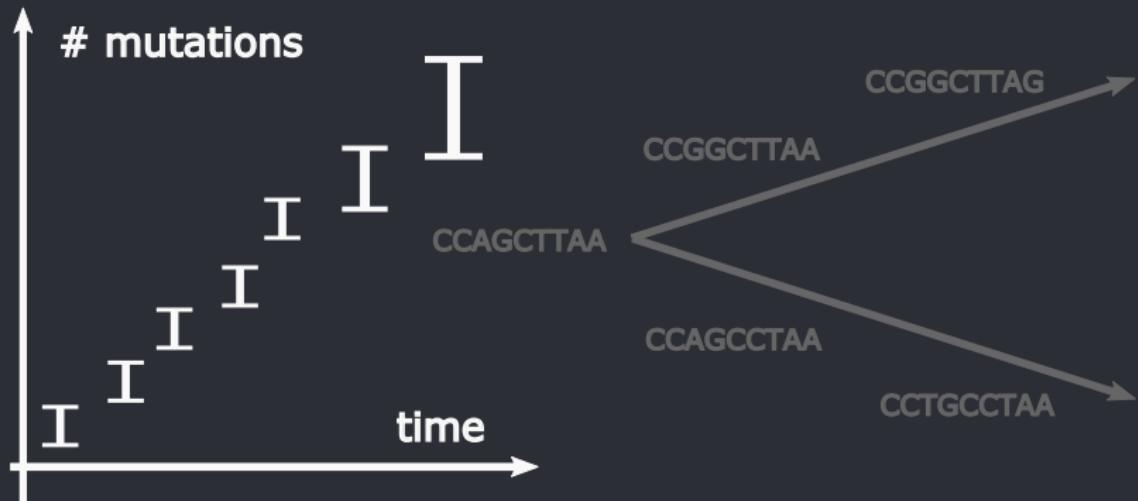


Figure:

# Phylogenetics: The Molecular Clock

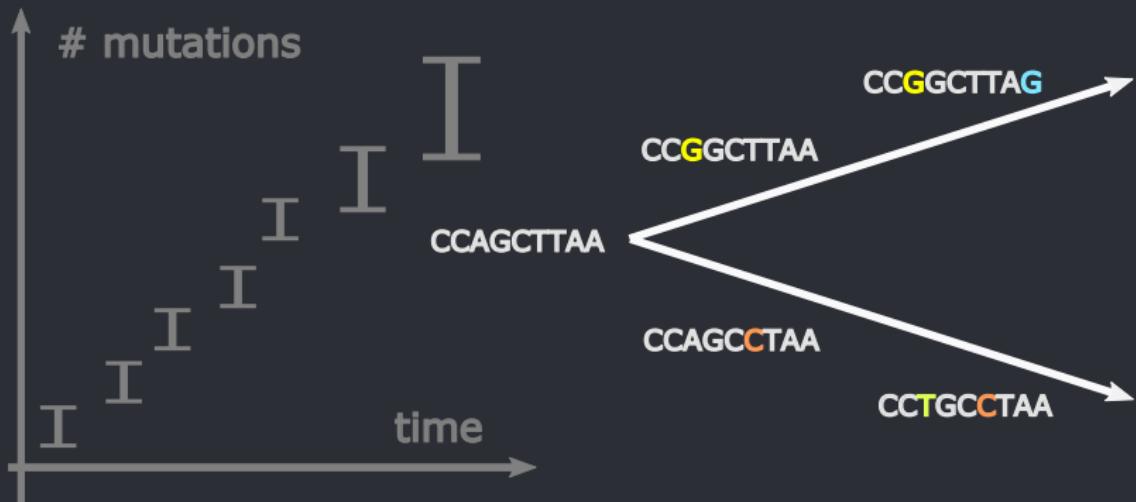


Figure: Linear time-mutation relationship

Corroborated by genetic equidistance.  
Limited by complete turnover times.

# Phylogenetics: Sequence Alignment

# Phylogenetics: Building A Phylogenetic Tree

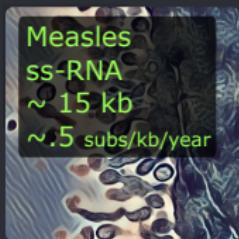
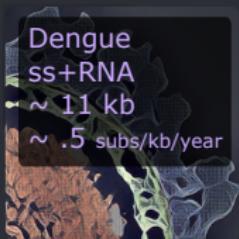
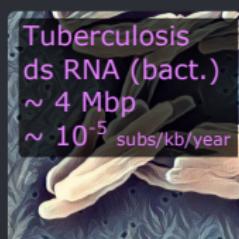
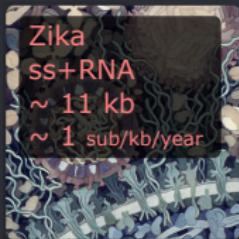
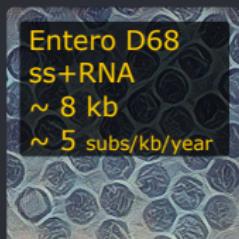
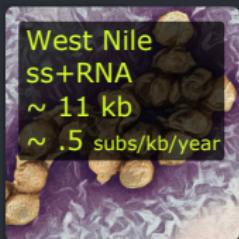
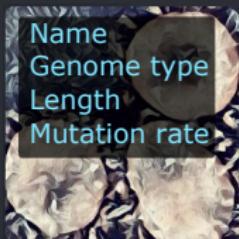
*—from sequenced data*

*Nextstrain*

*Mending pieces together*

Please visit [nextstrain.org/narratives/](https://nextstrain.org/narratives/)

# Nextstrain: How to use the Framework (this is mirrored in the narrative)



## **Nextstrain: The Powerful Meta Data**

*(this is mirrored in the narrative)*

## **Nextstrain: Confidence Levels and Limitations**

*(this is mirrored in the narrative)*

# Conclusion

*...and why Nextstrain is awesome*

# References I

- [AAMC, 2018] AAMC (2018).  
Journal of the Association of American Medical Colleges Frequently Asked Questions.
- [Alberts, 2015] Alberts, B. (2015).  
*Molecular biology of the cell*.
- [Bedford, 2015] Bedford, T. (2015).  
Real-time evolutionary forecasting for influenza vaccine strain selection.
- [rosipaw, 2010] rosipaw (2010).  
146/365 square peg into a round hole.
- [Tokars et al., 2018] Tokars, J. I., Olsen, S. J., and Reed, C. (2018).  
Seasonal Incidence of Symptomatic Influenza in the United States.  
*Clinical Infectious Diseases*, 66(10):1511–1518.  
Publisher: Oxford Academic.