

# **Phylogenetics for Predicting Virus Evolution**

---

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

**Paul Wiesemeyer**

## Problem

- Influenza evading immunity
- Pandemics as with SARS-CoV-2
- "Houston, we have aah—CHOO!"
- 3 – 11% of the population catch symptomatic influenza each season. CITATION
- Influenza mutates quickly at  $\sim$  2 mutations / kilobases / year
- No one can tell next seasons circulating influenza strains.
- Antigenic shift can cause sudden epidemics and even pandemics.
- How can we take informed counter measures on a global level?

# Outline for Section 1

1. Influenza and Vaccines
  - 1.1 Basics
  - 1.2 The Hemagglutination Inhibition Assay
2. Phylogenetics
  - 2.1 The Molecular Clock
  - 2.2 Sequence Alignment
  - 2.3 Building A Phylogenetic Tree
3. *Nextstrain*
  - 3.1 How to use the Framework
  - 3.2 The Powerful Metadata
  - 3.3 Confidence Levels and Limitations

# Influenza and Vaccines

*Influenza—an artful disguise master*

## Basics

Where does influenza come from?

- *Zoonotic 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*

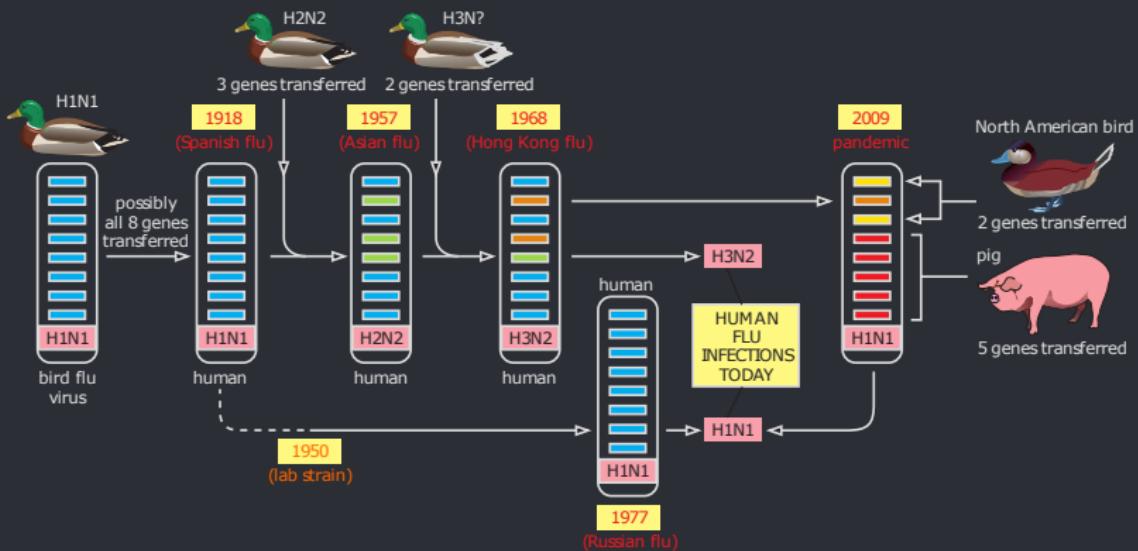


Figure: taken from [Alberts, 2015]

## Basics

## Antigenic drift—a race

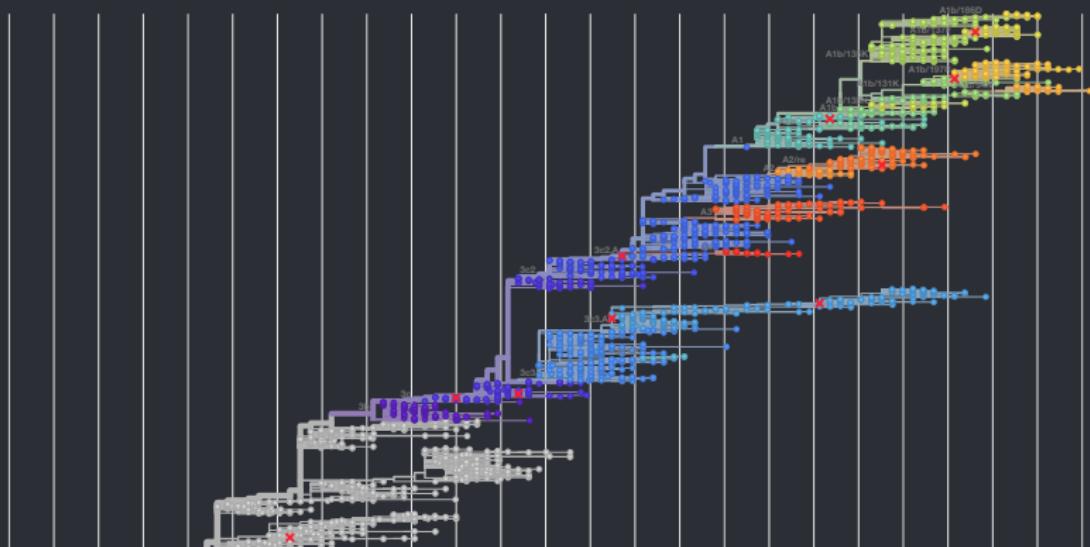


Figure: From NEXTSTRAIN.org Divergence (= subs/HA-genome-length)

# Basics

*Antigenic drift—a race*



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

# The Hemagglutination Inhibition Assay

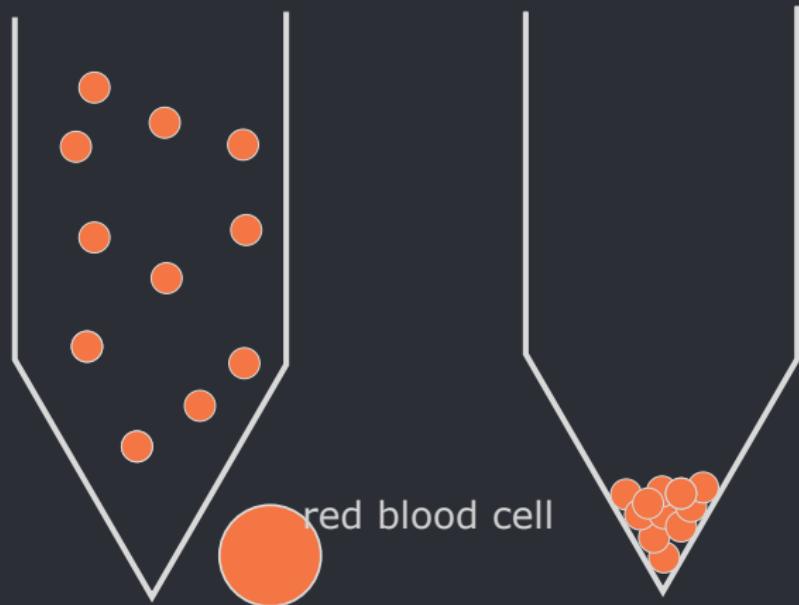


Figure: Red blood cells (RBC) precipitate.

# The Hemagglutination Inhibition Assay

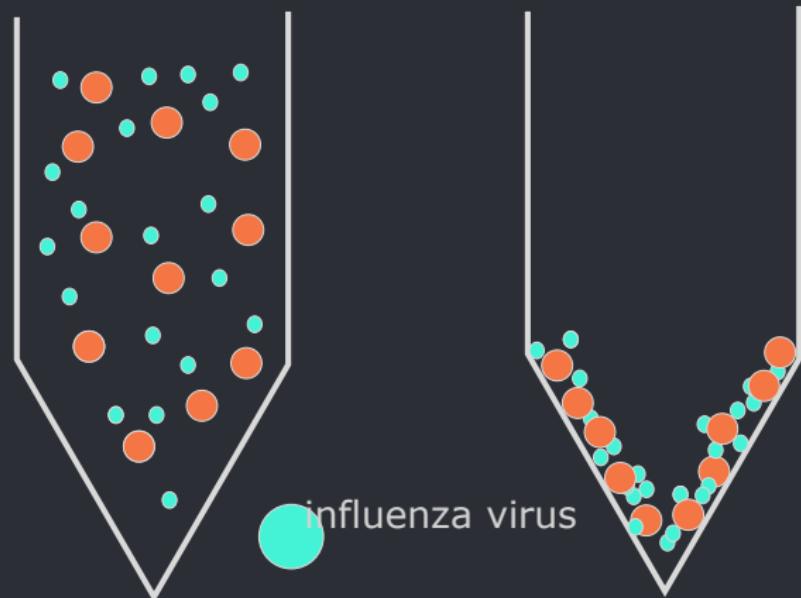


Figure: Influenza Hemagglutinin (HA) coagulates the RBC, forming a mat.

# The Hemagglutination Inhibition Assay

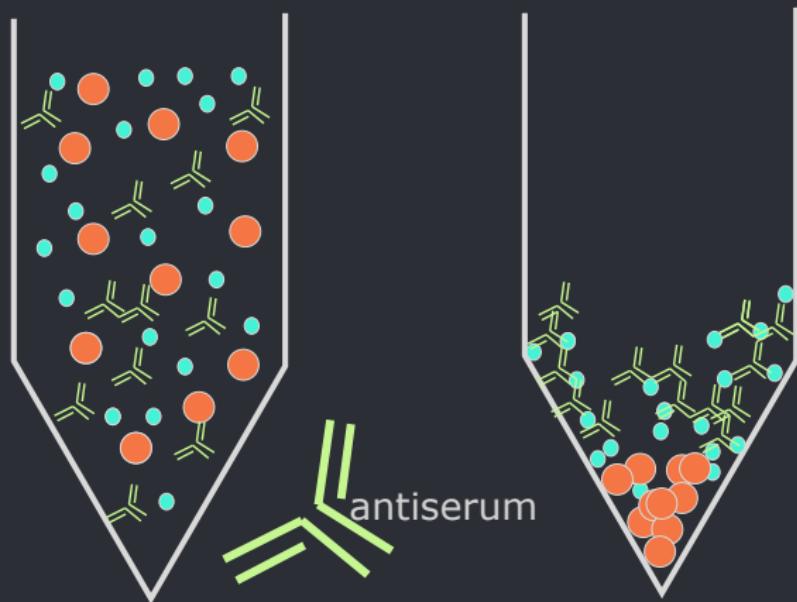


Figure: Antisera of the same serotype clump the HA, letting the RBC sink to the bottom. This is an (antiserum) concentration dependent process.

# The Hemagglutination Inhibition Assay

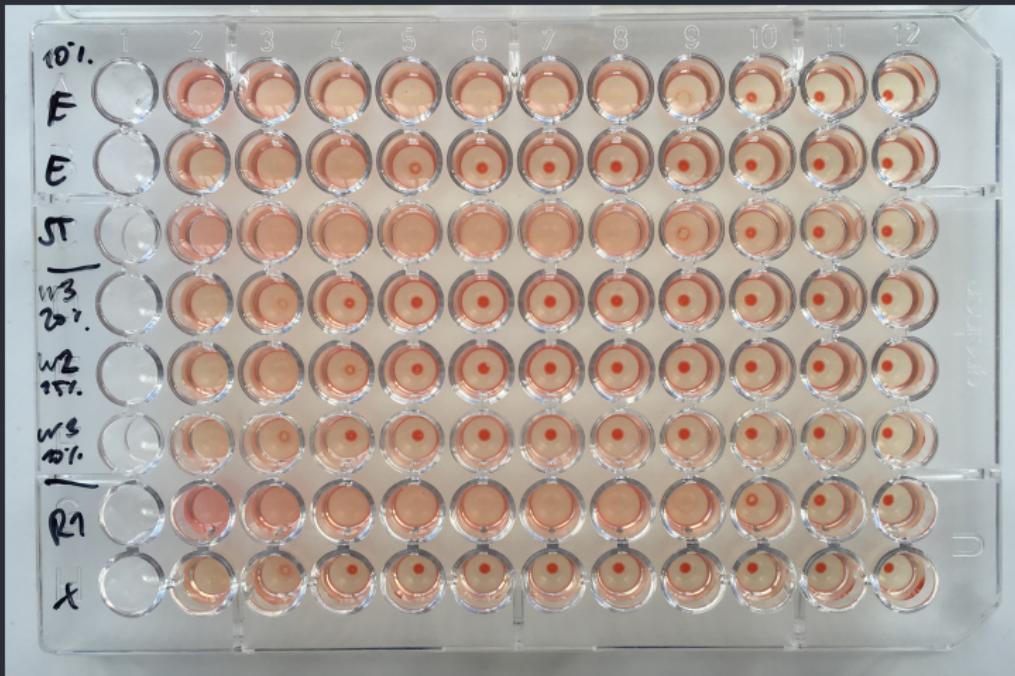
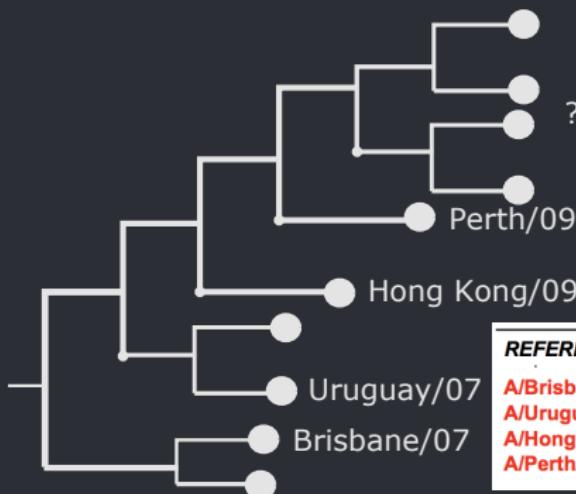


Figure: Here, one antiserum is tested in 12 different dilutions against 8 different virus strains. The highest dilution that prevents agglutination is called the titer.

# The Hemagglutination Inhibition Assay

*How this used to be looked at*

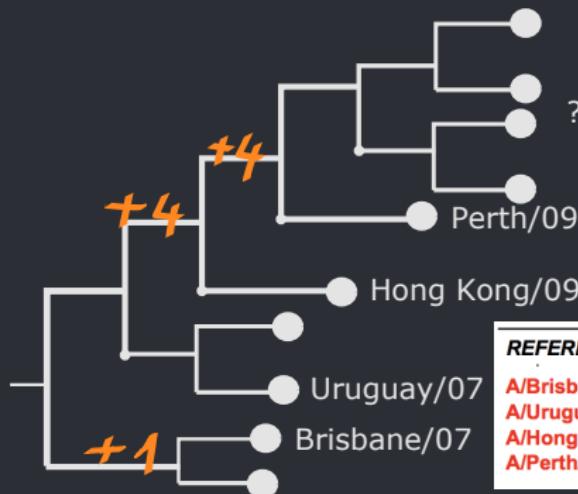
# The Hemagglutination Inhibition Assay



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.

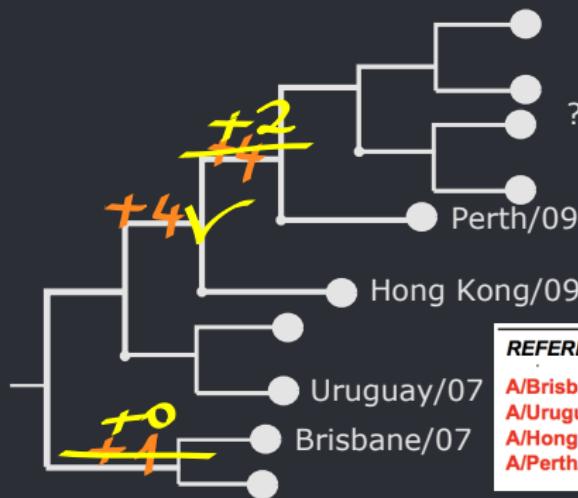
# The Hemagglutination Inhibition Assay



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.

## Multi Scale Evolution

If a single event mutation occurs, say **D 186 G** in the HA genome, it is subject to multiscale 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.

# Outline for Section 2

1. Influenza and Vaccines
  - 1.1 Basics
  - 1.2 The Hemagglutination Inhibition Assay
2. Phylogenetics
  - 2.1 The Molecular Clock
  - 2.2 Sequence Alignment
  - 2.3 Building A Phylogenetic Tree
3. *Nextstrain*
  - 3.1 How to use the Framework
  - 3.2 The Powerful Metadata
  - 3.3 Confidence Levels and Limitations

# Phylogenetics

*An old idea*

# Phylogenetics: The Molecular Clock

# Phylogenetics: Sequence Alignment

# Phylogenetics: Building A Phylogenetic Tree

*—from sequenced data*

# Outline for Section 3

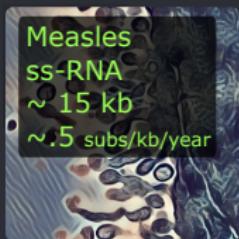
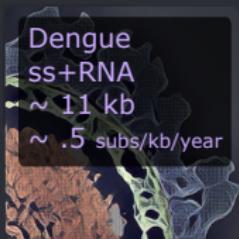
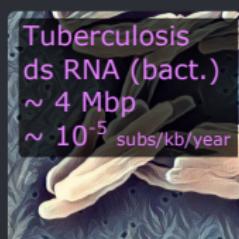
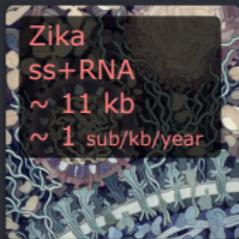
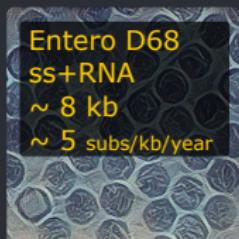
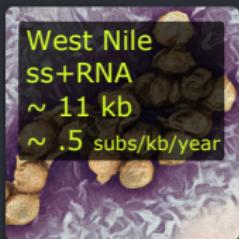
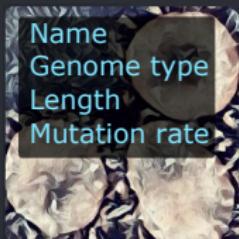
1. Influenza and Vaccines
  - 1.1 Basics
  - 1.2 The Hemagglutination Inhibition Assay
2. Phylogenetics
  - 2.1 The Molecular Clock
  - 2.2 Sequence Alignment
  - 2.3 Building A Phylogenetic Tree
3. *Nextstrain*
  - 3.1 How to use the Framework
  - 3.2 The Powerful Metadata
  - 3.3 Confidence Levels and Limitations

*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 Metadata**

*(this is mirrored in the narrative)*

## *Nextstrain: Confidence Levels and Limitations*

*(this is mirrored in the narrative)*

# Conclusion

*...and why Nextstrain is awesome*

## References

[Alberts, 2015] Alberts, B. (2015).  
*Molecular biology of the cell.*