Genetic Analysis of the 1918 Influenza Virus

Ann H. Reid

Armed Forces Institute of Pathology

Washington, DC

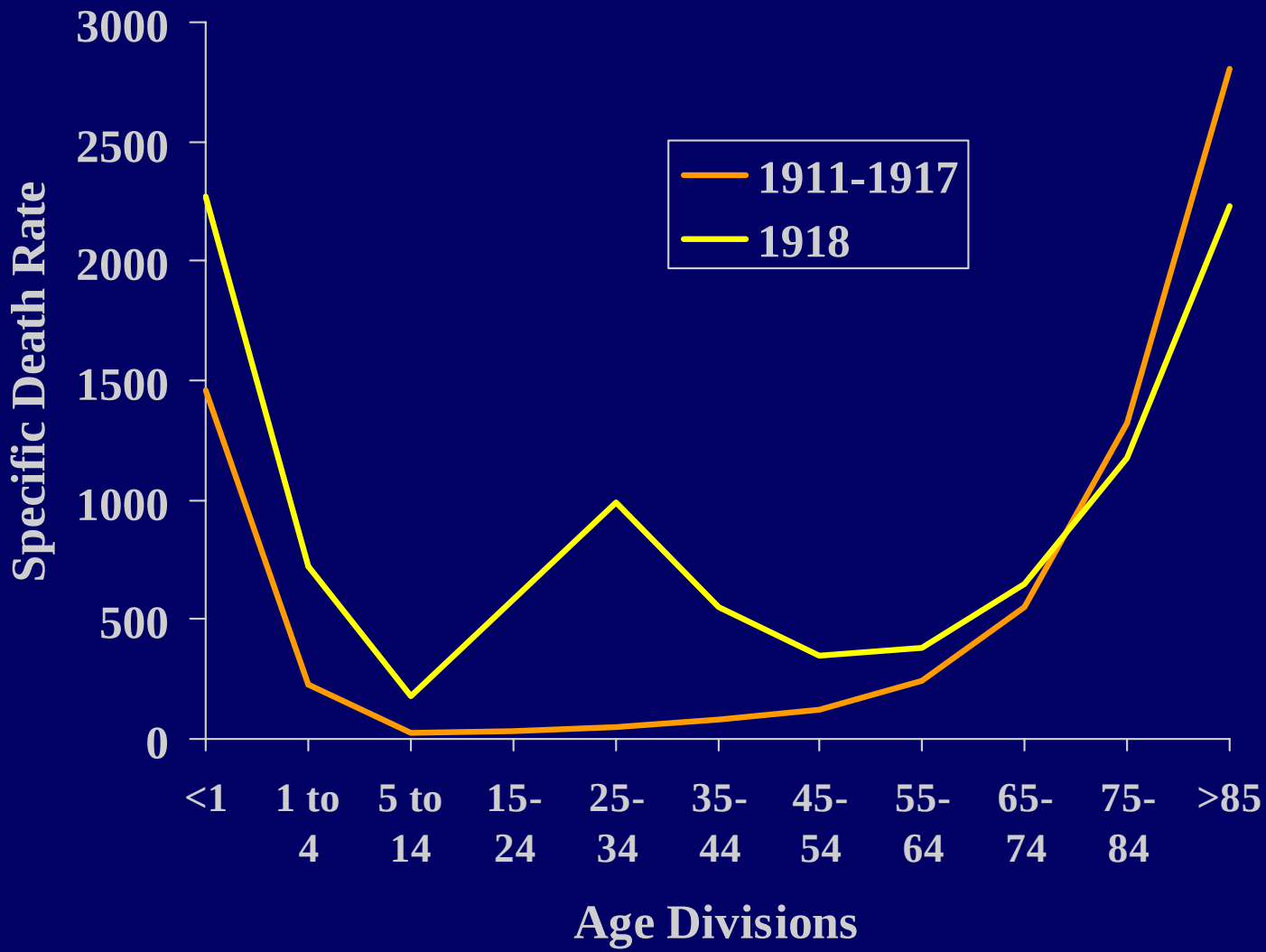
1918 ‘Spanish’ Influenza

- Origin in spring of 1918, possibly in U.S. Spring wave infectious but not very lethal
- Spread through Europe from late spring through summer
- Second wave started at end of August on 3 continents, exploded into worldwide pandemic in September-November, 1918

1918 ‘Spanish’ Influenza

- Total deaths in 1918-1919 estimated to be 21-50 million worldwide
- U.S. deaths = 675,000
- Flu deaths in Philadelphia in October 1918 = 10,959. Total flu deaths = 15,785
- U.S. military deaths to flu = 43,000 (out of ~100,000 U.S. troop casualties in W.W.I.)

Influenza and Pneumonia Deaths by Age



1918 influenza, AFIP lung block

1918 Cases:

- Case 1: 21 y.o., PVT, USA, Ft. Jackson, SC, died after 6 day course on 26 Sept. 1918; Dx: Influenza and pneumonia
- Case 2: 30 y.o., PVT, USA, at Camp Upton, NY, died after 3 day course on 26 Sept. 1918; Dx: Massive pulmonary edema
- Case 3: ? y.o. Inuit female from Teller Mission, Alaska, died in <5 days in Nov. 1918; Exhumation and lung biopsy in Aug. 1997; Dx: Pneumonia and hemorrhage

Purpose of gene sequencing:

- Where did the 1918 flu virus come from?
- How did it get into people?
- Why was it lethal?
- What genetic features can be related to its behavior or to the host response?

Why was it lethal?

- Unique mutation in single gene
- Combination of genes working together
- Host-pathogen relationship

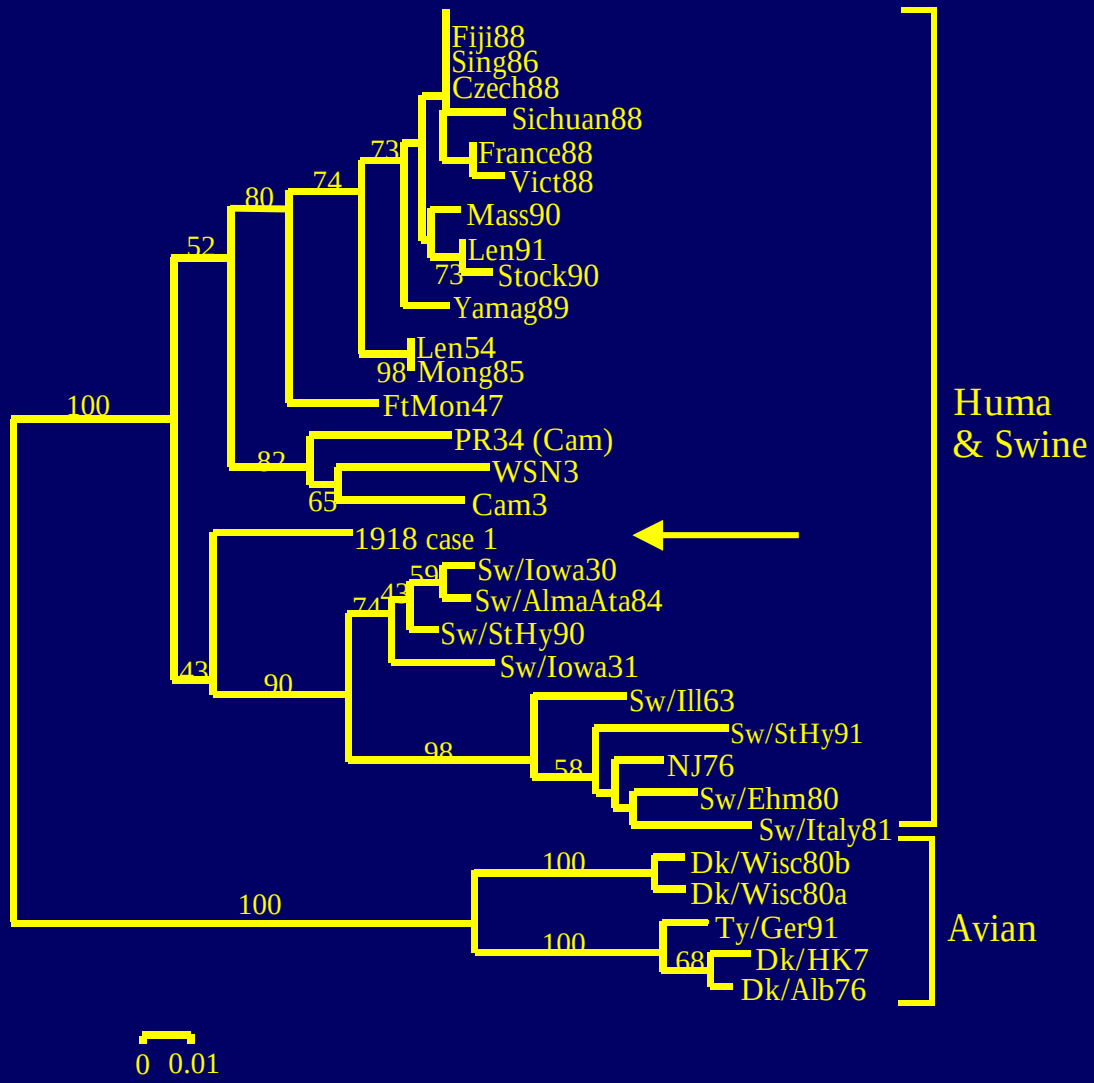
Sequencing results: Mar. 2002

- Surface glycoprotein gene segments sequenced:
  - Hemagglutinin (PNAS 1999, 96:1651)
  - Neuraminidase (PNAS 2000, 97:6785)
- Internal protein gene segments sequenced:
  - Non-structural (PNAS 2001, 98:2746)
  - Matrix (complete sequence, submitted)
  - Nucleoprotein (complete sequence, phylogenetic analysis underway)

Hemagglutinin (HA) summary:

- No cleavage site mutation as seen in virulent poultry influenza strains (e.g. Hong Kong H5N1)
- Receptor binding pattern: avian with minimal mammalian adaptation
- Antigenic sites: 37/41 avian consensus

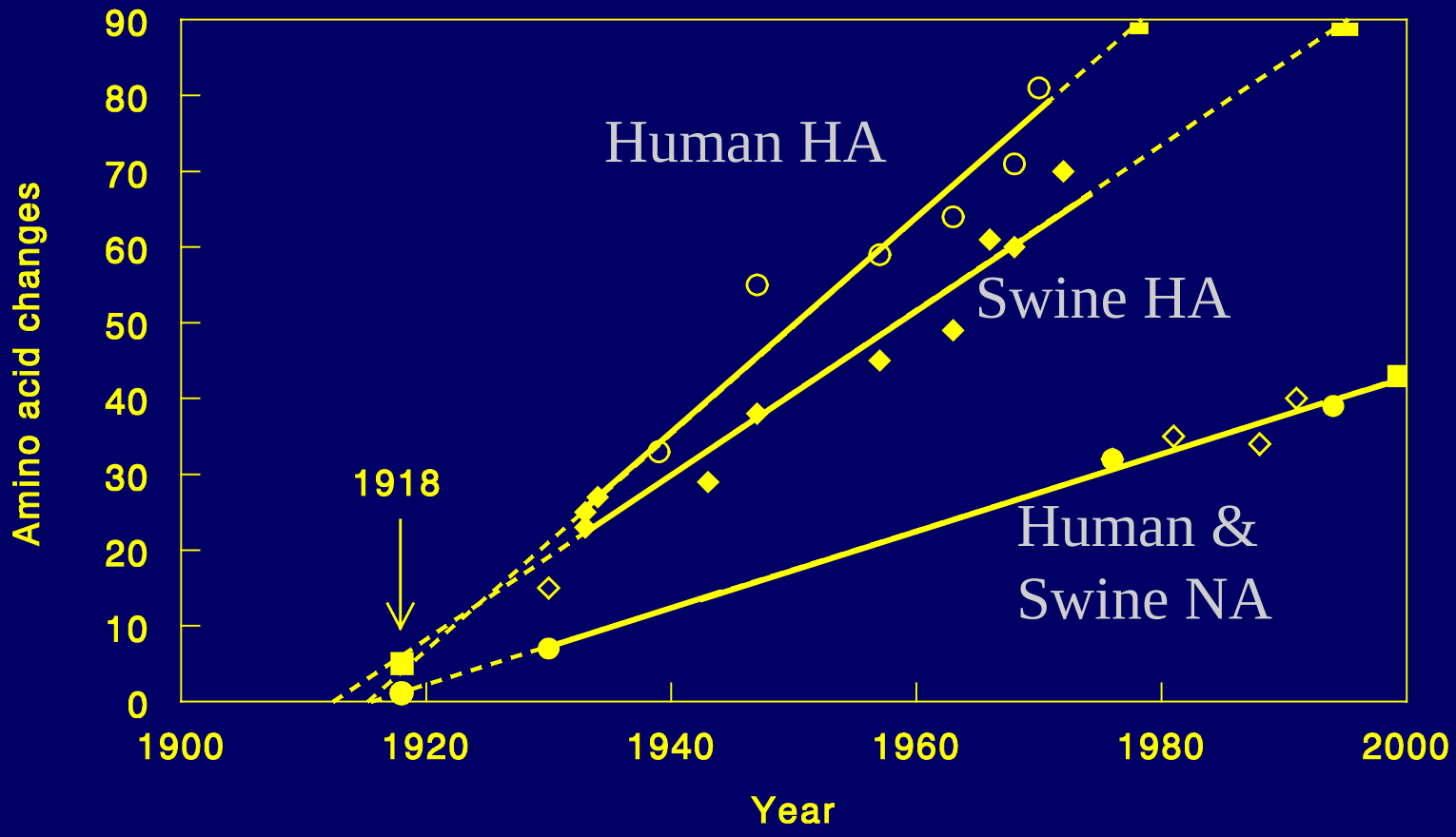
HA Phylogeny



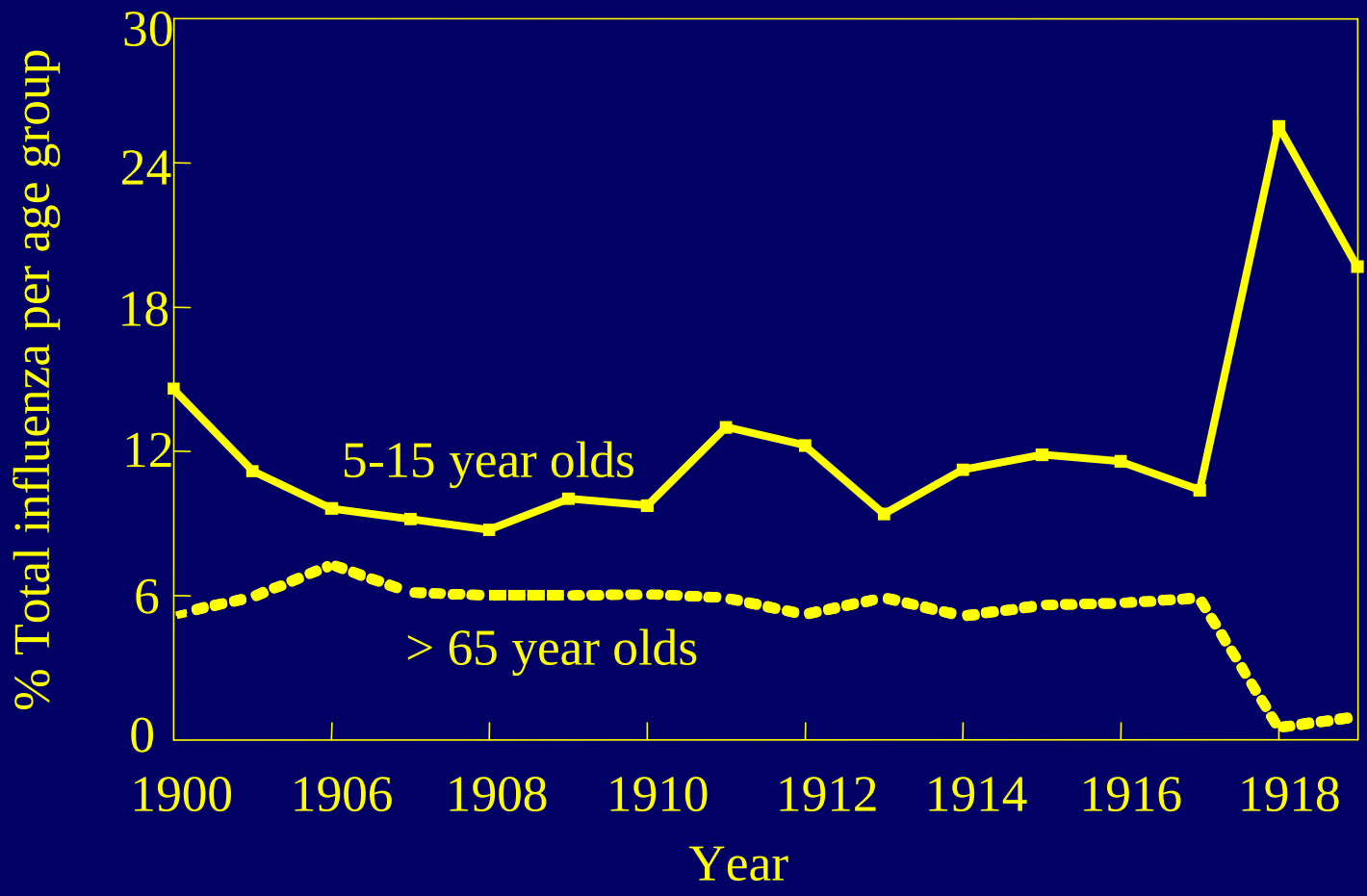
Phylogenetic Results

- For all the sequences analyzed so far, the phylogenetic analyses place the 1918 genes within but very near the root of the mammalian clade
- The 1918 sequences are the most avian-like of all mammalian sequences
- The 1918 sequences more closely resemble early mammalian sequences than any existing avian sequences

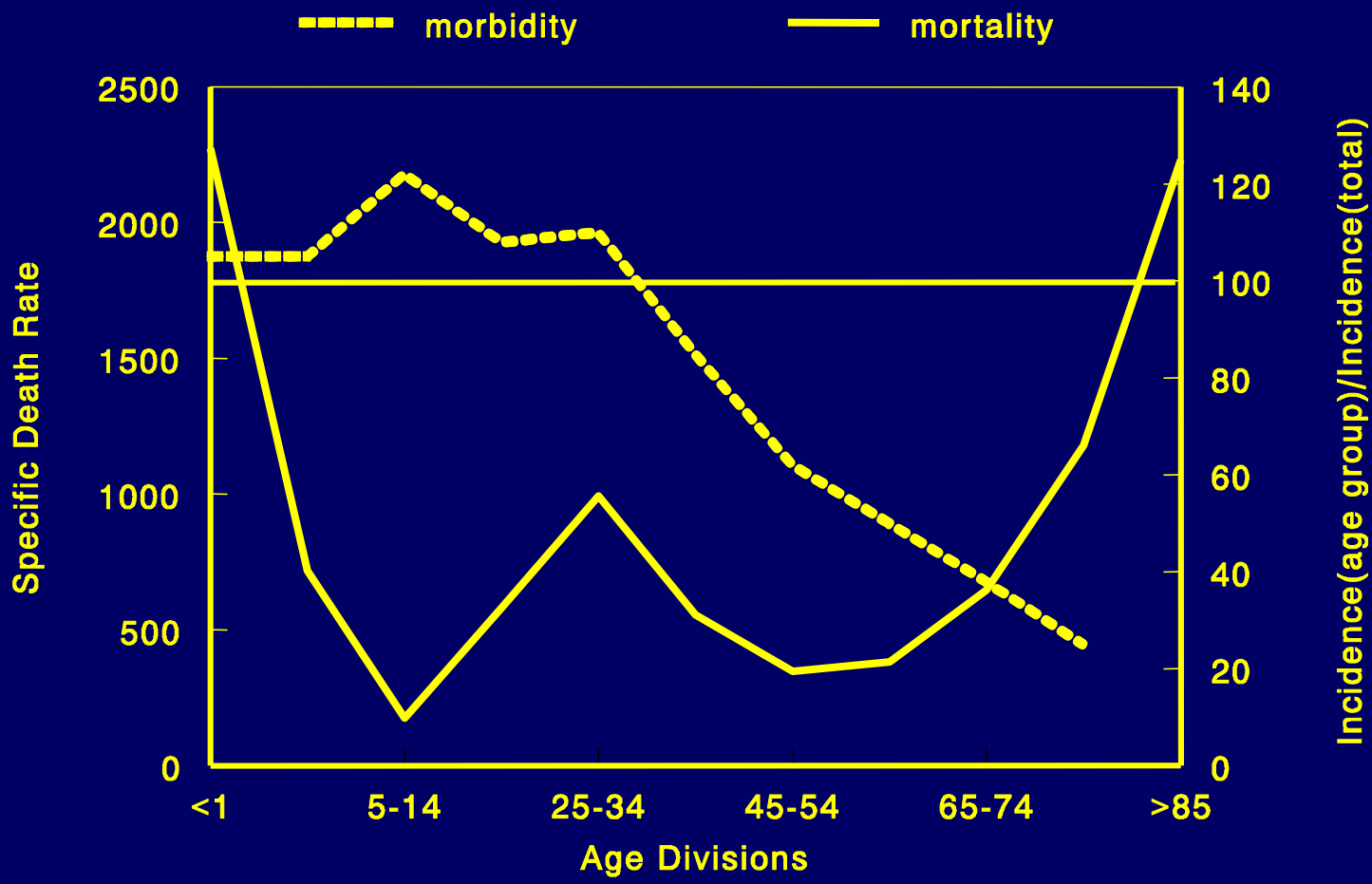
Rate of Change of HA and NA genes: H1N1 human and swine strains



Influenza prevalence by age, 1900-1918, Copenhagen, Denmark



Flu morbidity and mortality by age



Current evidence about origin:

- 1918 HA & NA genes almost certainly different from previously circulating strains
- Likely derived from an avian strain shortly before pandemic, but:
  - Seemingly different from the 1957 or 1968 pandemics
  - Period of adaptation?
  - Drift of avian flu sequences?
- Internal protein genes may or may not have reassorted

Genetic characterization of influenza A viruses recovered from birds in the Smithsonian Collection, 1915-1920

Protocol

- Wild waterfowl collected 1915-1920, catalogued and preserved in ethanol
- Cloacal biopsy, cloacal lavage, large intestinal biopsy performed on juvenile birds
- 25 total birds sampled, representing 16 species
- 6 birds positive for influenza A RNA by RT-PCR
- HA2 domain consensus primers used to screen for H1 subtype

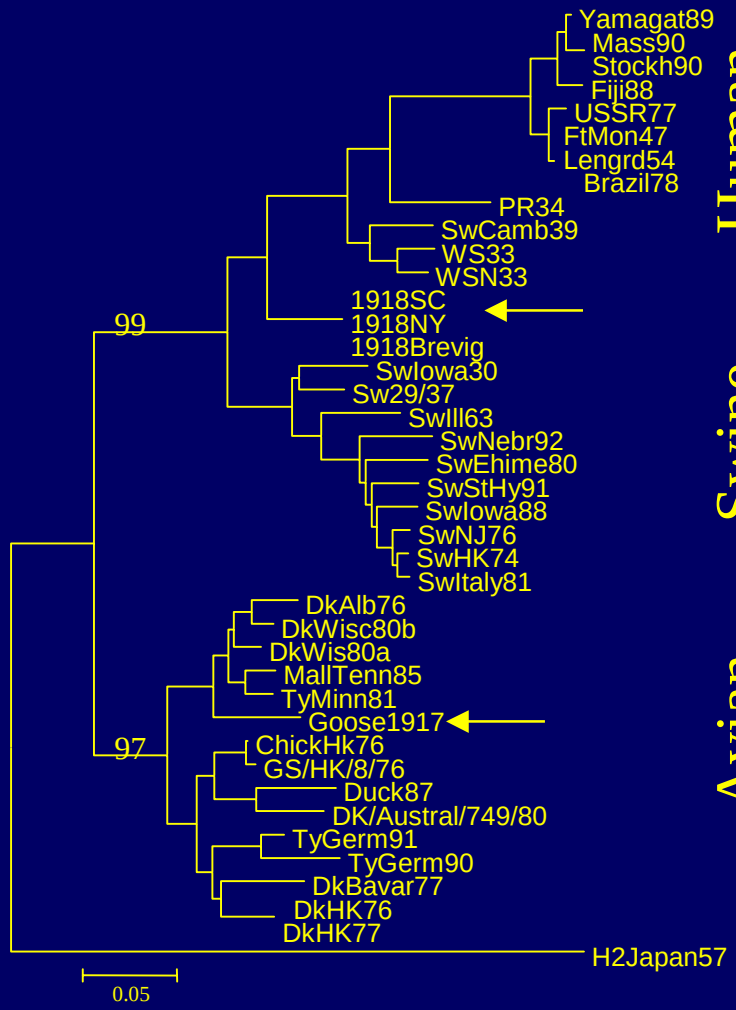
Influenza RNA-positive birds

- **Bear River, Utah, collected by A. Wetmore:**
  - **Bufflehead duck** (*Bucephala albeola*), 15 October 1916
  - **Greenwing teal** (*Anas crecca*), 24 September, 1916
  - **Cinnamon teal** (*Anas cyanoptera*), 11 August, 1916
  - **Ruddy duck** (*Oxyura jamaicensis*), 11 August, 1916
- **Bluewing teal** (*Anas discors*), 25 February, 1919, Dry Tortugas, Florida, collected by C. Johnson
- **Brant goose** (*Branta bernicla*), 9 September, 1917, St. Paul Island, Pribilof Islands, Alaska, collected by G.D. Hanna

A/Brant Goose/1/1917 (H1N?)

- ~500 nucleotides of HA1 domain including antigenic sites and receptor binding sites
  - Antigenic sites: avian consensus, 41/41 residues
  - Receptor binding sites: avian consensus
  - Phylogenetics: near root of North American avian clade
- 250 nucleotides of nucleoprotein
  - Phylogenetics: near root of North American avian clade
- 100 nucleotides of matrix

HA1 phylogeny, 1917 H1 goose



Conclusions

- No significant drift in wild avian influenza A viruses in 80 years
- 1918 pandemic strain phylogenetically distinct from avian clade (equidistant from No. Am. and Eurasian branches)
- Unlike 1957 and 1968, the 1918 strain apparently did not acquire its HA directly from a bird strain but evolved for some time in a mammalian host

Implications for origin of 1918 pandemic

- 1918 HA did not come as directly from avian sources as did 1957 and 1968 HAs
- 1918 NA also has more changes from avian sequences than 1957 NA
- Internal protein genes may or may not have reassorted; difficult to determine without pre-1918 human flu cases. Sequence analysis of MA, NS and NP consistent with retention from previously circulating human strain

Acknowledgements

- Jeff Taubenberger, Tom Fanning, Tom Janczewski, AFIP
- Johan Hultin, San Francisco and the people of Brevig Mission, Alaska
- Dick Slemons, Ohio State University, James Dean, Smithsonian Institution
- This work has been funded by grants from the Department of Veterans Affairs, the American Registry of Pathology, and by intramural funds of the Armed Forces Institute of Pathology