

MOLECULAR VIROLOGY

Was the 1918 pandemic caused by a bird flu?

Arising from: J. K. Taubenberger *et al.* *Nature* **437**, 889–893 (2005)

Taubenberger *et al.*¹ have sequenced the polymerase genes of the pandemic 'Spanish' influenza A virus of 1918, thereby completing the decoding of the genome of this virus^{2–6}. The authors conclude from these sequences that the virus jumped from birds to humans shortly before the start of the pandemic and that it was not derived from earlier viruses by gene shuffling, a process called reassortment. However, we believe that their evidence does not convincingly support these conclusions and that some of their results even indicate that, on the contrary, the virus evolved in mammals before the pandemic began and that it was a reassortant. In light of this alternative interpretation, we suggest that the current intense surveillance of influenza viruses should be broadened to include mammalian sources.

The inferences by Taubenberger *et al.*¹ are based on two lines of evidence. First, they report phylogenetic trees for nine of the eleven genes of the virus. The trees would have supported the conclusions if, in all or at least in most of them, the 1918 virus linked the bird influenza viruses with the later human and swine influenza viruses: that is, if it had been placed on the root of the cluster of mammalian viruses (Fig. 1a). That topology would indicate that the 1918 virus was the ancestor of all post-1918 human influenza viruses, as Taubenberger *et al.*¹ assume.

However, none of the nine phylogenetic trees described by Taubenberger and his co-workers, including those of the polymerases, shows that topology^{1–6}. In five of the published trees, the 1918 gene is placed next to the main cluster of human influenza viruses, and the classical swine influenzas are linked to the tree between the branches of the human and avian viruses (Fig. 1b). The other four trees have the reverse topology; the 1918 gene lies next to the classical swine viruses, and the main cluster of human viruses is between them and the avian viruses (Fig. 1c). The combined mammalian cluster is linked to the bird cluster directly in five of the trees, but in the other four trees it is linked first to other lineages from mammals (that is, those of pigs and horses).

Taubenberger and colleagues^{1–6} report that the 1918 virus was placed with the bird influenza viruses in some unpublished trees and that in all the trees, both published and unpublished, the 1918 virus lies close to the root of the mammalian lineages of influenza viruses. However, the 1918 virus does not lie on the root, and therefore the phylogenetic analyses of Taubenberger *et al.* do not support their conclusions: their results indicate instead

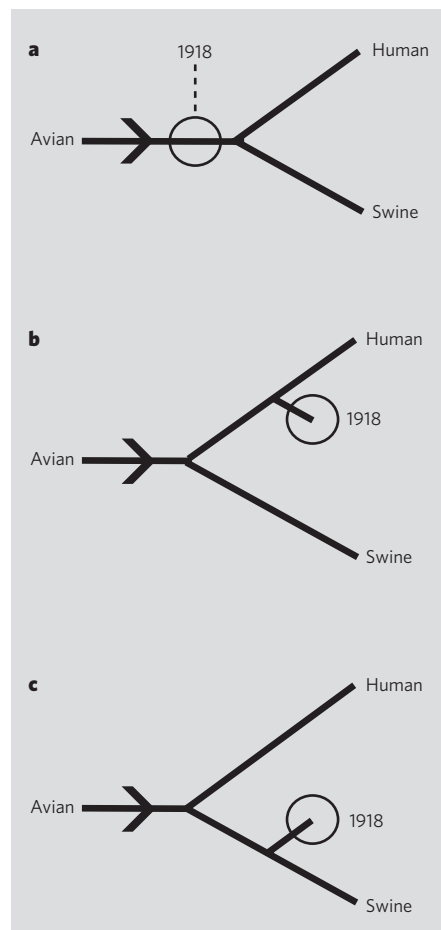


Figure 1 | Hypothetical and observed phylogeny of influenza-virus genes. **a**, Phylogenetic tree that would be expected if the 1918 virus had come directly from birds. The avian influenza-virus source would be linked to the later influenzas isolated from mammals, namely on the branch linking the avian viruses to the mammalian viruses. **b**, Summary of the trees found by Taubenberger and co-workers^{1–6} for five of the influenza-virus genes (*PA*, *PB1*, *HA*, *M1*, *M2*; see ref. 1 for details), which placed the 1918 virus next to the later human influenzas: that is, they are sister groups. **c**, Summary of the remaining four influenza virus gene trees found by Taubenberger and co-workers^{1–6} (*PB2*, *NP*, *NS1*, *NA*; see ref. 1 for details), which placed the 1918 virus next to the later swine viruses. The avian-virus source cluster in most of the gene trees also includes virus isolates from mammals.

that the virus evolved in people or pigs for an unknown period of time before the pandemic started and that the virus may have been a reassortant.

The second line of evidence used by Taubenberger *et al.*¹ to support their conclusions comes from comparing similarities in

polymerase amino-acid sequences. Phylogenetic analysis is usually preferred to this approach for studying origins, because similarities can sometimes occur by coincidence through parallel evolution. Indeed, Taubenberger *et al.*¹ discovered examples of identical amino-acid residues in bird and human influenza virus polymerases that resulted from parallel evolution, and we contend that those discoveries undermine their conclusions.

Although the 1918 virus polymerase proteins are placed closest to the typical sequences of some bird influenza viruses, the ranking depends on very few residues; parallel evolution could therefore have affected the ranking. In contrast, when nucleotide similarity was measured, rather than amino-acid similarity, the authors found the 1918 virus to be significantly closer to influenza viruses from mammals. The evidence of sequence similarity uncovered by Taubenberger *et al.*¹ is important for understanding how influenza viruses adapt to humans, but it does not prove their conclusions.

The events that led to the emergence of the 1918 virus are unclear and will probably remain so, at least until the immediate ancestors of the virus have been characterized and the mutation rate of the virus lineage is known. Data from other influenzas show that the mutation rate varies from close to 0% to almost 1% per year⁷, introducing uncertainty about timing inferred from influenza-virus sequences. The haemagglutinin gene from an influenza virus from a goose captured in 1917 has already been sequenced⁸: perhaps unluckily, the 1917 virus was not an ancestor of the 1918 pandemic virus, but it was similar to contemporary H1N1 avian influenzas.

Mark J. Gibbs*, **Adrian J. Gibbs†**

*School of Botany and Zoology, The Australian National University, Canberra, ACT 0200, Australia

e-mail: mark.gibbs@anu.edu.au

†7 Hutt Street, Yarralumla, ACT 2600, Australia

1. Taubenberger, J. K. *et al.* *Nature* **437**, 889–893 (2005).
2. Basler, C. F. *et al.* *Proc. Natl Acad. Sci. USA* **98**, 2746–2751 (2001).
3. Reid, A. H., Fanning, T. G., Hultin, J. V. & Taubenberger, J. K. *Proc. Natl Acad. Sci. USA* **96**, 1651–1656 (1999).
4. Reid, A. H., Fanning, T. G., Janczewski, T. A. & Taubenberger, J. K. *Proc. Natl Acad. Sci. USA* **97**, 6785–6790 (2000).
5. Reid, A. H., Fanning, T. G., Janczewski, T. A., McCall, S. & Taubenberger, J. K. *J. Virol.* **76**, 10717–10723 (2002).
6. Reid, A. H., Fanning, T. G., Janczewski, T. A., Lourens, R. M. & Taubenberger, J. K. *J. Virol.* **78**, 12462–12470 (2004).
7. Webster, R. G., Bean, W. G., Gorman, O. T., Chambers, T. M. & Kawakita, Y. *Microbiol. Rev.* **56**, 152–179 (1992).
8. Fanning, T. G. *et al.* *J. Virol.* **76**, 7860–7862 (2002).

doi:10.1038/nature04823

MOLECULAR VIROLOGY

Was the 1918 flu avian in origin?

Arising from: J. K. Taubenberger *et al.* *Nature* **437**, 889–893 (2005)

Taubenberger *et al.*¹ claim that the 1918 influenza virus was derived from an avian source and adapted to humans shortly before the pandemic. However, we do not believe that this conclusion, which has been widely disseminated in the popular press and in scientific journals^{2–4}, is supported by their phylogenetic evidence.

The authors' conclusion is based on DNA sequences from three RNA-polymerase genes, each of which resides on a separate RNA segment of the viral genome. The nucleotide phylogenies that they present¹ show, with strong statistical support, that the 1918 influenza virus is found within clades containing human

and classical swine influenza viruses and is not basal to those clades. Moreover, the relationship of the 1918 strain to avian strains (rather than to equine or other mammalian strains) is unresolved because the trees are unrooted.

The phylogenies described by Taubenberger *et al.*¹ contradict their main conclusions and are presented without discussion of the evolutionary relationships they imply. Instead, evolutionary conclusions are improperly drawn from a similarity between the 1918 and avian influenza viruses in the patterns of base-pair substitution (that is, the synonymous/non-synonymous and transition/transversion ratios, and variation in fourfold-degenerate

sites) and without consideration of the relative similarity of the 1918 flu in these traits to other mammalian strains. These are inappropriate characters from which to infer similarity by descent, and the data from the DNA sequences should not be discounted as they are superior indicators of phylogenetic relatedness.

Taubenberger *et al.*¹ also claim that the amino-acid sequences encoded by these RNA-polymerase genes support the avian nature of the 1918 virus. We reconstructed phylogenies using their amino-acid sequences⁵ and found that the 1918 virus falls within, and not basal to, clades containing strains from other mammalian hosts (Fig. 1). Sequences from the nucleoprotein gene of the influenza virus also fail to provide evidence that the 1918 strain is derived directly from an avian source⁵.

In support of their conclusions, Taubenberger *et al.*¹ cite prior phylogenetic studies⁶ indicating that the 1918 flu may have been derived from an avian source. However, those results simply show that the 1918 flu is phylogenetically unresolved and genetically equidistant from the North American and Asian clades of avian flu, which does not indicate emergence from an avian source.

By stating that the high pathogenicity of the 1918 virus is related to its emergence as a human-adapted avian influenza virus, the authors raise the possibility that an emerging avian strain could resemble the 1918 flu. This alarming implication, which is based on misinterpretation of the phylogenetic data, is completely unjustified and could seriously distort the public perception of disease risk, with grave economic and social consequences.

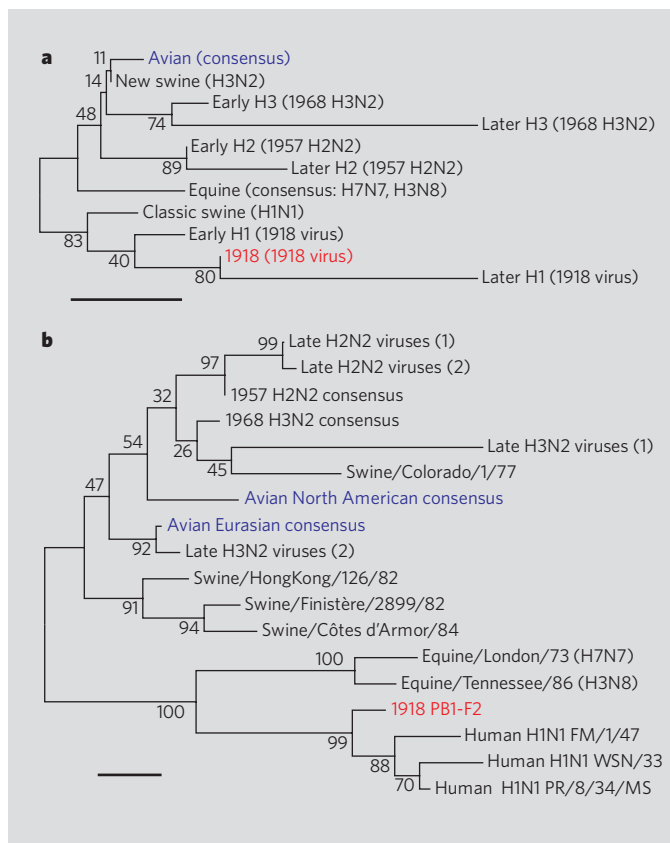
Janis Antonovics, Michael E. Hood, Christi Howell Baker

Department of Biology, University of Virginia, Charlottesville, Virginia 22904, USA
e-mail: ja8n@virginia.edu

1. Taubenberger, J. K. *et al.* *Nature* **437**, 889–893 (2005).
2. Belshe, R. B. *N. Engl. J. Med.* **353**, 2209–2211 (2005).
3. Russell, C. J. & Webster, R. G. *Cell* **123**, 368–371 (2005).
4. Sharp, P. A. *Science* **310**, 17 (2005).
5. Reid, A. H., Fanning, T. G., Janczewski, T. A., Lourens, R. M. & Taubenberger, J. K. *J. Virol.* **78**, 12462–12470 (2004).
6. Reid, A. H., Taubenberger, J. K. & Fanning, T. G. *Nature Rev. Microbiol.* **2**, 909–914 (2004).
7. Kumar, S., Tamura, K. & Nei, M. *Brief. Bioinform.* **5**, 150–163 (2004).

doi:10.1038/nature04824

Figure 1 | Neighbour-joining analysis of the PB1 subunit of the influenza virus A polymerase heterotrimer and of the peptide PB1-F2. **a, b,** Amino-acid alignments for **a**, PB1, and **b**, PB1-F2 (incomplete sequences omitted), taken from Supplementary Figure 2a, b of Taubenberger *et al.*¹. Data were analysed using MEGA 3.1 software⁷; bootstrap values represent 100 replications. Sample names are given in ref. 1. The 1918 and avian viruses are shown in red and blue text, respectively.



MOLECULAR VIROLOGY

Taubenberger *et al.* reply

Replying to: M. J. Gibbs & A. J. Gibbs *Nature* **440**, doi:10.1038/nature04823; J. Antonovics, M. E. Hood & C. H. Baker *Nature* **440**, doi:10.1038/nature04824 (2006)

We have proposed that the virus responsible for the 1918 human influenza pandemic was avian-like, which to us is a reasonable interpretation of all the available data from phylogenetic, sequence, serological and historical

analysis, combined with what is known of influenza A virus biology^{1–4}. However, Gibbs and Gibbs⁵ and Antonovics *et al.*⁶ question our inferred origin of the 1918 influenza virus.

Gibbs and Gibbs contest that the 1918 virus

evolved in mammals before the pandemic and was a reassortment⁵. They present three simplified trees showing the 1918 virus on the root of the mammalian clade, or in either the human or classical swine subclades, claiming that the 1918 virus was not placed on the root of the tree as expected. Likewise, Antonovics *et al.* state that the 1918 virus sequences are placed in the human and classical swine influenza clades and are not basal to them⁶. That is indeed true for the polymerase-gene trees shown, so here we assess this apparent discrepancy.

When branch lengths are examined, it is evident that the sequence of the 1918 virus is the closest to the common ancestor of the mammalian clade in the genes for each of the three polymerases (PA, PB1, PB2). Similarly, the haemagglutinin and neuraminidase genes of the 1918 virus are clearly placed very near the root of the mammalian clade^{7,8}. In contrast to results shown in Fig. 1 of Antonovics *et al.*⁶, which is derived from a subset of our data¹, a more extensive analysis using a larger data set of amino-acid sequences¹ places the 1918 virus closest to the root of the mammalian clade for each polymerase gene (for example, see Fig. 1).

In an earlier phylogenetic study of swine influenza viruses⁹, it was concluded that the internal genes of classical swine viruses are 'avian-like' and that these genes evolved concomitantly; furthermore, the results were in agreement with the historical account that the 1918 human influenza virus entered the pig population in 1918. We agree with the generally held conclusion that the human and classical swine H1N1 influenza viral lineages descended from the 1918 virus. It is important to note, however, that not all swine influenza viruses are from this lineage. The trees shown by Antonovics *et al.* show sequences from several swine influenza viruses that are not descended from the classical swine lineage, but rather from the European 'avian-like' swine influenza H1N1 lineages of the late 1970s¹⁰ or from an Asian H3N2 swine virus containing an avian-derived PB1 gene¹¹. These lineages share no evolutionary relationship with the 1918 virus.

We have never maintained that the virus entered the human population in 1918: rather, as described earlier^{2,12}, our claim that it entered the human population "shortly" before the pandemic should be interpreted as 'at least several years before the pandemic', as stated in our discussion¹. The path that the precursors of the 1918 pandemic virus took before emerging in humans in 1918 remains unknown. Phylogenetic analysis on its own cannot definitively resolve the issue. As in previous analyses⁹, we analysed the sequences of these genes for clues about their origins and found that the proteins encoded by the 1918 polymerase genes were avian-like in all cases.

Other scenarios can be derived from the available data to explain the origin of the 1918 human influenza virus. For example, a virus from an unknown avian-like source could have infected an unknown mammal, where it may have evolved for several years before

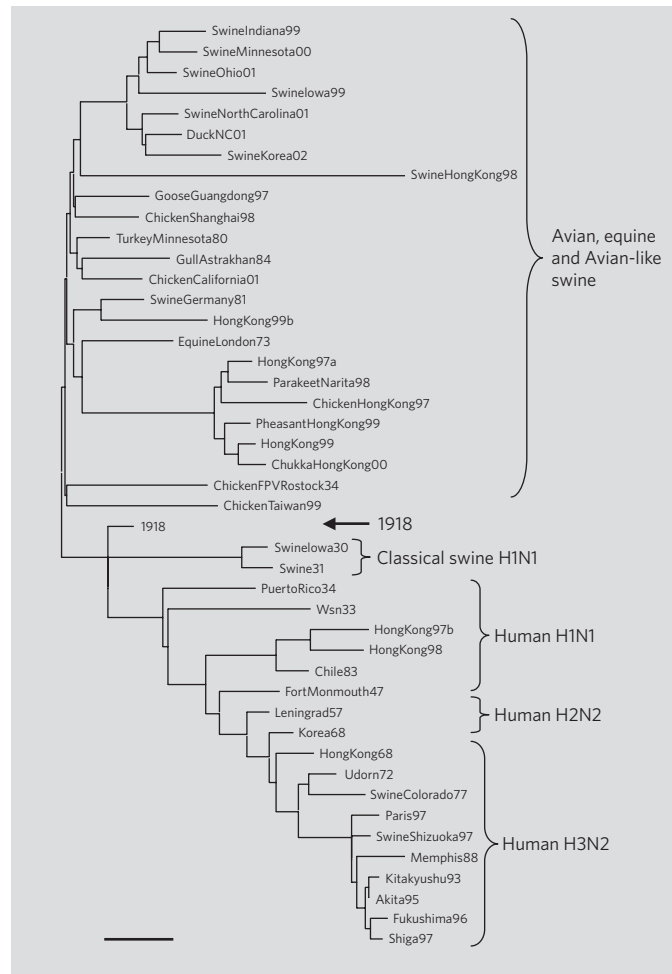


Figure 1 | Neighbour-joining analysis using MEGA 2.1 software¹⁴ of influenza A viral amino-acid sequences of the PB2 polymerase subunit. The arrow designates the 1918 sequence. The scale is proportion of differences.

causing the 1918 pandemic, or it could have infected humans and adapted to them directly several years before the 1918 pandemic. But as no direct precursors of the 1918 virus have been identified, we cannot at present separate these two possibilities. We find the second more plausible, given what we now know about recent human infection with the avian H5N1 influenza virus¹³.

Jeffery K. Taubenberger*, **Ann H. Reid*†**, **Raina M. Lourens*‡**, **Ruixue Wang***, **Guozhong Jin***, **Thomas G. Fanning***

*Armed Forces Institute of Pathology, Department of Molecular Pathology, Rockville, Maryland 20850, USA

e-mail: taubenberger@afip.osd.mil

†Board on Life Sciences, The National Academies, Washington DC 20001, USA

‡Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, Iowa 52242, USA

1. Taubenberger, J. K. *et al.* *Nature* **437**, 889–893 (2005).
2. Reid, A. H., Taubenberger, J. K. & Fanning, T. G. *Nature Rev. Microbiol.* **2**, 909–914 (2004).
3. Taubenberger, J. K., Reid, A. H., Janczewski, T. A. & Fanning, T. G. *Phil. Trans. R. Soc. Lond. B* **356**, 1829–1839 (2001).
4. Gorman, O. *et al.* *J. Virol.* **65**, 3704–3714 (1991).
5. Gibbs, M. J. & Gibbs, A. J. *Nature* **440**, doi:10.1038/nature04823 (2006).
6. Antonovics, J., Hood, M. E. & Baker, C. H. *Nature* **440**, doi:10.1038/nature04824 (2006).
7. Reid, A. H., Fanning, T. G., Hultin, J. V. & Taubenberger, J. K. *Proc. Natl Acad. Sci. USA* **96**, 1651–1656 (1999).
8. Reid, A. H., Fanning, T. G., Janczewski, T. A. & Taubenberger, J. K. *Proc. Natl Acad. Sci. USA* **97**, 6785–6790 (2000).
9. Schultz, U., Fitch, W., Ludwig, S., Mandler, J. & Scholtissek, C. *Virology* **183**, 61–73 (1991).
10. Brown, I. H. *et al.* *J. Gen. Virol.* **78**, 553–562 (1997).
11. Kawaoka, Y., Krauss, S. & Webster, R. G. *J. Virol.* **63**, 4603–4608 (1989).
12. Reid, A. H., Fanning, T. G., Janczewski, T. A., Lourens, R. M. & Taubenberger, J. K. *J. Virol.* **78**, 12462–12470 (2004).
13. Peiris, J. S. *et al.* *Lancet* **363**, 617–619 (2004).
14. Kumar, S., Tamura, K., Jakobsen, I. B. & Nei, M. MEGA2: Molecular Evolutionary Genetics Analysis software, version 2.1 (Arizona State Univ., Tempe, Arizona, 2001).

doi:10.1038/nature04825