

## Review

Correspondence  
Jeffery Taubenberger  
taubenbe@afip.osd.mil

# The origin of the 1918 pandemic influenza virus: a continuing enigma

Ann H. Reid and Jeffery K. Taubenberger

Division of Molecular Pathology, Department of Cellular Pathology and Genetics, Armed Forces Institute of Pathology, 1413 Research Blvd, Building 101, Room 1057, Rockville, MD 20850-3125, USA

Influenza A virus is a major public health threat, killing more than 30 000 per year in the USA alone, sickening millions and inflicting substantial economic costs. Novel influenza virus strains emerge periodically to which humans have little immunity, resulting in devastating pandemics. The 1918 pandemic killed nearly 700 000 Americans and 40 million people worldwide. Pandemics in 1957 and 1968, while much less devastating than 1918, also caused tens of thousands of deaths in the USA. The influenza A virus is capable of enormous genetic variability, both by continuous, gradual mutation and by reassortment of gene segments between viruses. Both the 1957 and 1968 pandemic strains are thought to have originated as reassortants, in which one or both human-adapted viral surface proteins were replaced by proteins from avian influenza virus strains. Analyses of the surface proteins of the 1918 pandemic strain, however, suggest that this strain may have had a different origin. The haemagglutinin gene segment of the virus may have come directly from an avian source different from those currently circulating. Alternatively, the virus, or some of its gene segments, may have evolved in an intermediate host before emerging as a human pathogen. Determining whether pandemic influenza virus strains can emerge via different pathways will affect the scope and focus of surveillance and prevention efforts.

## INTRODUCTION

That there will be epidemics of influenza every year is a virtual certainty. That they will begin in late winter and last a month or two is also very likely (Brammer *et al.*, 2002). However, beyond those general rules, predicting the timing, magnitude and severity of influenza epidemics is a formidable public health challenge. Influenza A viruses circulate widely in humans and spread in several epidemiologically distinct ways: as localized outbreaks, as yearly regional epidemics and, occasionally, as global pandemics. In the USA, influenza leads to the hospitalization of over 100 000 and kills over 30 000 people in an average year (Simonsen *et al.*, 2000; Thompson *et al.*, 2003). Every 2 or 3 years, influenza epidemics boost the yearly number of deaths past the average, causing 10 000–15 000 additional deaths. Occasionally, and unpredictably, influenza sweeps the world, infecting 20 to 40 % of the population in a single year. In these pandemic years, which have occurred every 10 to 50 years for at least several centuries, the number of deaths can be dramatically above average (Beveridge, 1977; Cox & Subbarao, 2000; Wright & Webster, 2001). It is very likely that influenza will return in pandemic form. Recently, it has been estimated for the USA alone that the next

influenza pandemic may result in up to 207 000 deaths, 734 000 hospitalizations, 42 million outpatient visits and 47 million additional illnesses (Meltzer *et al.*, 1999). The estimated economic impact would be 70–170 billion dollars, excluding disruptions to commerce and society. Currently, it is impossible to predict the timing or severity of the next pandemic outbreak, but study of the genetic and epidemiological characteristics of past pandemics may suggest where surveillance and research would be directed best (Layne *et al.*, 2001; Taubenberger & Layne, 2001).

The 1918 influenza pandemic fits the classic pattern of influenza epidemiology in many ways. It occurred 28 years after the previous pandemic of 1890 and emerged globally with explosive suddenness in September 1918 after a limited wave earlier in the year. Most communities experienced morbidity of 25–40 % and the vast majority of cases were self-limiting. Age-specific morbidity was also similar to other pandemics, with children under 15 years of age experiencing the highest rates of infection (Jordan, 1927). Clinically, the 1918 pandemic presented the same symptoms and course as influenza of other years and, pathologically, the disease was similar to other pandemics in that damage was confined largely to the respiratory tract (Wolbach, 1919; Winternitz *et al.*, 1920). However, the 1918 pandemic differed from other pandemics in a few key respects. First,

while the clinical course in the majority of cases was mild, a substantially higher percentage of cases developed severe pneumonic complications. As a result, the case mortality rate in the USA averaged 2.5%, several times higher than the contemporary average. Also, mortality during the 1918 pandemic was concentrated in an unusually young age group (Linder & Grove, 1943; Marks & Beatty, 1976; Rosenau & Last, 1980). People under the age of 65 accounted for more than 99% of excess influenza-related deaths in 1918. In 1957 and 1968, people under 65 accounted for only 36 and 48% of excess deaths due to influenza (Simonsen *et al.*, 1998). The age group affected most severely by the 1918 pandemic was between 20 and 40 years and this group accounted for almost half of influenza deaths during the pandemic.

Until recently, the 1918 pandemic strain was not available for study, since influenza viruses were not isolated and cultured until the 1930s. By then, 15 years of circulation in humans had altered significantly the antigenicity of the circulating H1 haemagglutinin (HA), as assessed serologically (Shope, 1936; Taubenberger *et al.*, 2001), and only indirect analyses of the 1918 strain could be performed. Recently, extraction of RNA from fixed and frozen lung tissues from victims of the 1918 pandemic has allowed the sequencing of the 1918 influenza virus genome (Taubenberger *et al.*, 1997). Four of eight gene segments have been sequenced (Reid *et al.*, 1999, 2000, 2002; Basler *et al.*, 2001). This work has two principal goals: to determine the genetic contribution to the virulence of the 1918 influenza and to determine the origin of the pandemic virus. Understanding the basis of the virulence of the 1918 strain could help in the development of influenza treatment and prevention, while knowing where and how the strain developed could help direct surveillance and prevention efforts.

### Influenza A virus biology and ecology

Influenza A viruses are negative-stranded RNA viruses of the family *Orthomyxoviridae* (Lamb & Krug, 2001; Wright & Webster, 2001). Their segmented genome consists of eight RNA segments encoding at least ten proteins. Two glycosylated proteins on the surface of the virus, HA and NA (neuraminidase), are involved in virus attachment and release from host cells. They are also the primary target of the immune system in humans and swine. Two nonstructural proteins, NS1 and NS2 (also called NEP), are involved in regulating numerous aspects of the virus life cycle. Three proteins, PA, PB1 and PB2, are responsible for virus replication, while the nucleoprotein, NP, is the nucleocapsid structural protein. Finally, two membrane proteins, M1 and M2, are involved in nuclear export and pH maintenance, respectively, among other activities (Lamb & Krug, 2001; Wright & Webster, 2001). Influenza A virus is capable of considerable genetic variability. Its polymerases' lack of proofreading capability results in a high mutation rate and the organization of the genome into segments allows reassortment as an important mechanism for generating

diverse strains. Co-infection of one host with two strains can result in novel, reassortant strains, because progeny viruses can be formed with some gene segments from one strain and some from the other.

Pandemic influenza results when an influenza virus strain emerges with an HA protein to which few people have prior immunity (Kilbourne, 1977). It is thought that the source of HA genes new to humans is the extensive pool of influenza viruses that infect wild birds (Wright & Webster, 2001). Periodically, genetic material from avian strains is transferred to strains infectious to humans by reassortment. Of the 15 HA subtypes found in birds (H1–H15), only three (H1, H2 and H3) are known to have caused pandemics in man (Kilbourne, 1997). Recently, wholly avian H5N1 and H9N2 viruses (without reassortment) caused illness in a limited number of people in China (Lin *et al.*, 2000; Hatta & Kawaoka, 2002). Avian and human HAs differ in their ability to bind to different forms of sialic acids and avian HAs bind poorly to the sialic acid receptors prevalent in the human respiratory tract. These different receptor affinities act as a barrier to cross-species infection. Before a virus with an avian HA can replicate and spread efficiently in humans, some adaptation of the HA binding affinity is necessary. It is not known currently whether the HA subtypes that have become established in human strains were able to adapt more easily than other subtypes or whether all 15 avian subtypes pose a similar risk of reassortment.

Since pigs can be infected with both avian and human strains, and various reassortants have been isolated from pigs, they have been proposed as an intermediary in the generation of reassortant pandemic strains (Ludwig *et al.*, 1995). In 1979, an avian influenza A virus began infecting swine in Northern Europe, thereby establishing a stable virus lineage (Ludwig *et al.*, 1995). Since that time, there has been evidence of reassortment between the new swine lineage and human strains circulating currently. Viruses have been detected in swine in which the avian-derived H1 and N1 have been replaced by reassortment with the H3 and N2 HA and NA segments circulating concurrently in humans (Castrucci *et al.*, 1993; Claas *et al.*, 1994; Marozin *et al.*, 2002). However, reassortant strains with the avian-derived H1 and N1 along with human-adapted core protein segments have not been found. Such reassortant strains would be antigenically novel and probably capable of effective replication in humans and, therefore, would have substantial pandemic potential. Similarly, a number of triple reassortant strains, which include gene segments of swine, human and avian origin, have been isolated recently from pigs in the USA. Several reassortant viruses bearing human HA and NA segments have been isolated from swine but, as yet, no viruses with swine or avian surface proteins and human internal protein segments have been detected (Zhou *et al.*, 2000; Marozin *et al.*, 2002; Olsen, 2002).

Until recently, it was thought that reassortment between avian and human strains would be unlikely to take place in humans because there was no evidence that humans could

be infected by a wholly avian influenza virus. However, in 1997, 18 people were infected with avian H5N1 influenza viruses in Hong Kong and six died of complications after infection (Claas *et al.*, 1998; Subbarao *et al.*, 1998). Although these viruses were very poorly transmissible, if at all (Katz *et al.*, 1999), their detection indicates that humans can be infected with wholly avian influenza virus strains. Therefore, it may not be necessary to invoke swine as the intermediary in the formation of a pandemic strain (Scholtissek, 1995), since reassortment could take place directly in humans (Young & Palese, 1979; Palese & Young, 1982).

While reassortment appears to be a critical event for the production of a pandemic virus, a significant amount of data exists to suggest that influenza viruses must also acquire specific adaptations to spread and replicate efficiently in a new host. In addition to the adaptation of the HA protein to host cell receptors, other viral proteins must be able to interact with each other and various host cell proteins. Unfortunately, little is known about which specific genetic features of influenza viruses contribute to the emergence of a virulent pandemic strain. Virulence is complex and involves a number of features, including host adaptation, transmissibility, tissue tropism and virus replication efficiency. The genetic basis for each of these features is not characterized fully yet but is most likely polygenic in nature (Kilbourne, 1977).

### The 1957 and 1968 influenza virus pandemic strains

Prior to recent work on the 1918 virus, only two pandemic influenza virus strains were available for molecular genetic analysis, the H2N2 strain from 1957 and the H3N2 strain from 1968. The 1957 pandemic resulted from the emergence of a reassortant influenza virus in which both the HA and NA segments had been replaced by gene segments related closely to avian strains (Scholtissek *et al.*, 1978a; Schafer *et al.*, 1993; Webster *et al.*, 1995). The 1968 pandemic followed with the emergence of a strain in which the H2 subtype HA gene segment was replaced with an avian-derived H3 HA gene segment (Scholtissek *et al.*, 1978a; Webster *et al.*, 1995), while retaining the N2 gene segment derived in 1957. More recently, it was shown that the PB1 gene segment was replaced in both the 1957 and 1968 pandemics, also with a likely avian derivation in both cases (Kawaoka *et al.*, 1989). The remaining five gene segments: PA, PB2, NP, M and NS, were all preserved from the H1N1 strains circulating prior to 1957. These segments were likely the direct descendants of the gene segments present in the 1918 virus.

A pandemic virus faces the twin challenges of being antigenically 'new' to its host, while being supremely well adapted to it. This challenge was met in 1957 and 1968 by reassortment: combining surface proteins novel to humans with human-adapted internal proteins (with the intriguing exception of PB1). The 1968 viral HA appears to have had an avian origin (Fang *et al.*, 1981; Bean *et al.*, 1992). Sequencing of an avian H3 HA gene (A/duck/Ukraine/1/63) isolated in

1963 demonstrated its close molecular similarity to the HA gene of A/Aichi/2/68, the latter being an example of the 1968 pandemic virus. For example, 1605 of 1765 nucleotides (90.9%) are identical between the two viruses, while 542 of 566 amino acids (95.8%) are identical (Fang *et al.*, 1981). In addition, of the approximately 40 amino acid residues involved in antigenic recognition, only four residues differ between A/duck/Ukraine/1/63 and A/Aichi/2/68. For comparison, there are 14 amino acid differences in antigenic residues between A/duck/Ukraine/1/63 and A/Victoria/3/75, a human virus isolated only 7 years after the A/Aichi/2/68 virus (Fang *et al.*, 1981). Also, phylogenetic analyses support strongly the avian origin of the 1968 pandemic HA gene (Bean *et al.*, 1992).

Like the 1968 H3 pandemic strain, the HA of the 1957 pandemic is closely related also to avian H2 sequences. When the 1957 H2 sequences are compared as a group to avian HA sequences, only four amino acids differ consistently between the human and avian groups [N<sup>92</sup>→D, T<sup>114</sup>→K, K<sup>156</sup>→E and I<sup>214</sup>→T; H3 subtype sequence numbering (Winter *et al.*, 1981)]. Residue 226, which is Q in all avian sequences and L in most human sequences, is likely also to reflect a consistent difference between human and avian strains, since it is critical to improving the H2 binding affinity for receptors on human cells. Approximately 40 amino acids have been identified as being involved in antibody binding in the H3 molecule (Wiley *et al.*, 1981) and studies indicate that the H2 subtype has a similar antigenic structure (Tsuchiya *et al.*, 2001). Only one of the five amino acids differing between avian and early human isolates is in an antigenic site, suggesting that there had been little or no antigenic drift pressure on the H2 molecule before it emerged in a pandemic strain. Phylogenetic analyses (Schafer *et al.*, 1993) indicate that the gene was acquired from an avian source shortly before 1957. It appears that the avian source was Eurasian, since the pandemic viral sequences resembled Eurasian avian sequences much more closely than they resembled North American avian sequences.

The 1957 pandemic strain also acquired a novel N2-subtype NA, replacing the N1 of the previous strain. The sequence of the new NA was related very closely to avian N2 sequences, with only six amino acids differing consistently from avian sequences. Thirty-four amino acids have been identified as potentially antigenic residues on the N2 protein (Martinez *et al.*, 1983); none of the six differences are in these antigenic sites, suggesting that the protein had not been under selective antigenic pressure in humans before the pandemic. The avian sequence related most closely to the 1957 sequences is A/chicken/Korea/MS96/96, which differs at over 20 amino acids. There are no full-length N2 sequences from wild birds in the published databases but it seems likely that the recent avian origin of the 1957 N2 will be confirmed by further sequencing of wild avian strains.

The hypothesis that reassortment between avian and human strains is the likely mechanism for the generation of new

pandemic strains has become well accepted. During a recent outbreak of highly pathogenic avian influenza in The Netherlands, efforts were made to minimize the possibility of simultaneous infection with human and avian influenza, especially when the virus began by causing conjunctivitis in humans. Those involved in the culling process were vaccinated against circulating human strains and discouraged from contact with sick birds when suffering flu-like symptoms in an effort to minimize the possibility of an individual being infected simultaneously with human and avian strains (ProMED Mail, 2003). Given that the 1957 and 1968 pandemic strains may well have originated in just such a dual infection, these precautions seem appropriate. It is possible, though, that reassortment is not the only route to a pandemic. In 1947, drift in the prevailing H1 strain resulted in vaccine failure and outbreaks of influenza on a pandemic scale (Kilbourne, 1997). In 1977, an H1N1 virus re-emerged, having been absent since 1957, but failed either to cause a pandemic or to replace the prevailing H3N2 subtype (Nakajima *et al.*, 1978).

### Origin of the 1918 HA gene

The genetic sequences encoding the HA1 domains of five 1918 influenza virus strains have been determined. Two of the strains came from American soldiers who died on 26 September 1918: one in Camp Upton (New York, USA) and one in Fort Jackson (South Carolina, USA). The third came from an Inuit woman who died in mid-November 1918 in a remote village on the Seward Peninsula of Alaska (USA). Two nucleotide differences were found among these three strains, one of which resulted in an amino acid substitution in the receptor-binding site (Taubenberger *et al.*, 1997; Reid *et al.*, 1999). The fourth and fifth strains came from influenza victims treated at the Royal London Hospital in London (UK) and who died of pneumonia on 13 November 1918 and 15 February 1919, respectively (Reid *et al.*, 2003). The five sequences differ from each other by only one to three nucleotides. Our results show that strains separated by over 7500 miles (Brevig Mission, Alaska, USA to London, UK) and by several months (26 September 1918 to 15 February 1919) share a sequence identity of 99 %.

There is reason to question whether the 1918 pandemic strain originated in a simple reassortment immediately before the pandemic. Extensive phylogenetic analyses of the HA gene segment, in particular, are difficult to reconcile with the hypothesis of direct avian origin (Reid *et al.*, 1999). The sequence of the 1918 HA, although it is related more closely to avian strains than subsequent mammalian H1 sequences, has many more differences from avian sequences than the 1957 and 1968 HA sequences. If it should prove true that the 1918 pandemic strain acquired a novel HA via a different mechanism than subsequent pandemics, this could have important public health implications. An alternate origin could even have contributed to the exceptional virulence of the 1918 pandemic strain. Despite the current lack of influenza virus samples from before 1918, several indirect experimental approaches have been explored to test

the hypothesis of an alternative origin of the 1918 influenza virus strain.

The sequence of the 1918 HA is related most closely to A/sw/Iowa/30, the first influenza virus isolated from swine (Reid *et al.*, 1999). The similarity suggests that the human pandemic influenza virus became established in swine, in which it changed very slowly over the next 12 years. Unlike the 1957 and 1968 pandemic HAs, phylogenetic analyses do not place the 1918 sequence in the avian clade. However, the 1918 pandemic sequence is related more closely to avian H1s than to any other mammalian H1s and has many avian features. Of the 41 amino acids that have been shown to be targets of the immune system and subject to antigenic drift pressure in humans, 37 match the avian consensus sequence, suggesting that there was little immunologic pressure on the HA protein before the autumn of 1918 (Reid *et al.*, 1999; Brownlee & Fodor, 2001). Another mechanism by which influenza viruses evade the human immune system is the acquisition of glycosylation sites to mask antigenic epitopes. Modern human H1N1s have up to five glycosylation sites in addition to the four found in all avian strains. The 1918 virus has only the four conserved avian sites.

The H1 receptor-binding site apparently required little change from the avian-adapted receptor-binding site configuration [with a preference for  $\alpha(2,3)$  sialic acids] to that of swine H1s [which can bind both  $\alpha(2,3)$  and  $\alpha(2,6)$  sialic acids] (Matrosovich *et al.*, 1997). The receptor-binding residues of the 1918 HA differs by as little as one amino acid (E<sup>190</sup>→D) from the avian consensus.

In spite of the many ways in which the 1918 HA resembles avian viruses, phylogenetic analyses always place the 1918 HA with the mammalian viruses and not with the avian viruses (Reid *et al.*, 1999). Both the 1957 and 1968 pandemic strains appear to have resulted from reassortments of a human-adapted influenza virus strain with HA genes from a Eurasian avian lineage strain (Scholtissek *et al.*, 1978b; Bean *et al.*, 1992; Schafer *et al.*, 1993). In contrast, the 1918 HA is much less avian-like and, while probably novel to humans in 1918, does not appear to have been derived directly from an avian strain (Taubenberger *et al.*, 2001). Table 1 presents the number of amino acid differences between pandemic viruses and the consensus HA sequences of both North American and Eurasian birds. The numbers demonstrate that the HA genes of the pandemic viruses of 1968 and 1957 are more

**Table 1.** Number of amino acid differences between pandemic HA and the avian subtype consensus sequences

Pandemic strain	No. of differences from North American avian consensus sequences	No. of differences from Eurasian avian consensus sequences
1918	24	24
1957	19	5
1968	13	7

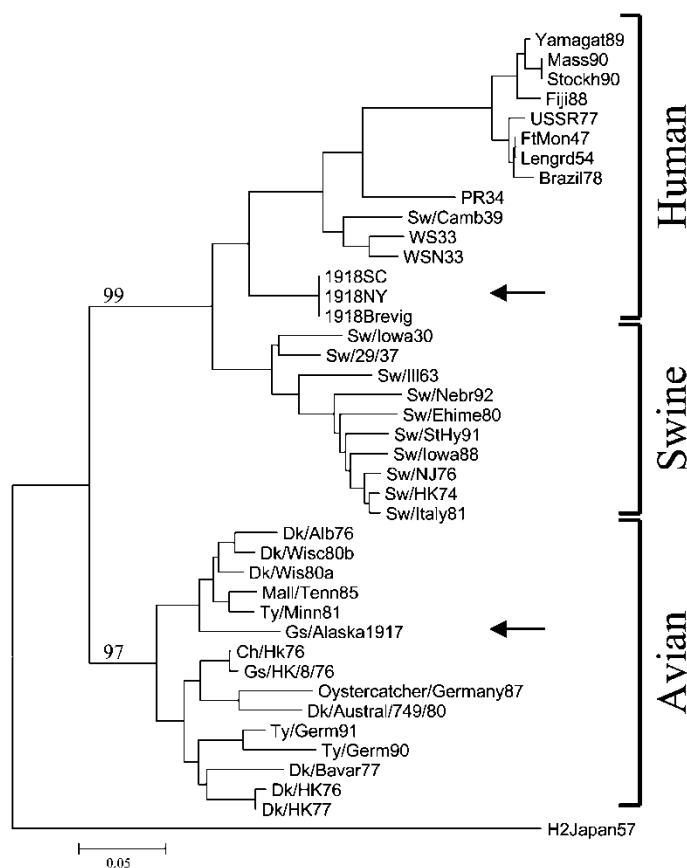
Eurasian avian-like (seven and five differences) than North American avian-like (13 and 19 differences). In contrast, the 1918 pandemic virus HA gene appears much less avian than either the 1968 or 1957 viruses and has no clear affinity with either North American or Eurasian avian viruses (Table 1). A similar situation can be demonstrated with the NA genes of the 1918 and 1957 strains.

The 1918 HA sequence suggests two contradictory explanations of its origin: either the HA spent some length of time in an intermediate host where it accumulated many changes from the original avian sequence or the 1918 HA came directly from an avian virus with a sequence markedly different from the H1 sequences available currently (Reid *et al.*, 1999). One possibility is that avian sequences might have drifted substantially in the 80 years since the 1918 pandemic. To assess this possibility, our laboratory, in collaboration with J. Dean (Museum of Natural History, Smithsonian Institution, Washington, USA) and R. D. Slemons (Ohio State University, Ohio, USA), was able to identify an H1 avian influenza A virus strain from a Brant Goose collected in 1917 and stored in the collections of the Smithsonian Institution. The H1 sequence obtained is related closely to modern avian North American H1 strains (Fig. 1), suggesting that there has been little drift in avian sequences over the past 80 years (Fanning *et al.*, 2002). Thus, the possibility that the 1918 virus came directly from the ancestor of an avian H1 strain known currently is remote.

However, there remains the possibility that H1-subtype HAs may exist in an avian host that has not been identified yet or that more sequence variation would be found within the H1 subtype upon more extensive sampling. Currently, the influenza virus sequence database contains only 14 avian H1 sequences, of which only nine are full-length sequences. Five of the 14 sequences are from turkeys or chickens and eight are from ducks. Comparison of these sequences shows that the two sequences related most closely (A/goose/Hong Kong/8/76 and A/chicken/Hong Kong/14/76) are 99.6% identical at the nucleotide level. The most distantly related (A/duck/Alberta/35/76 and A/turkey/Germany/2482/90) are 85.2% identical. Within the Eurasian group of avian viruses, sequence similarity ranges from 88 to 99.6%. The North American viruses are more homogeneous, with similarities ranging from 95.7 to 98.3%. The 1918 pandemic strain is equally distant – between 74.5 and 77.6% identity – from all the avian H1s. Extensive sequencing of wild bird H1 strains may identify a strain more similar to the 1918 HA.

#### Involvement of an intermediate host in 1918?

It may be that no avian H1 will be found resembling the 1918 strain because, in fact, the HA did not reassort directly from a bird strain. In this case, an intermediate host must be identified. A possibility that has been widely suggested is the pig (Webster *et al.*, 1992). Known to be susceptible to a wide



**Fig. 1.** Phylogenetic tree showing H1 sequences, including the 1918 pandemic strains and the 1917 Brant goose (arrows) (Fanning *et al.*, 2002).

range of both human and avian viruses and long-lived enough to exert some positive immune selection on the HA gene, swine could well have served as an intermediate host between birds and humans in 1918 (Alexander & Brown, 2000).

Indeed, during the 1918 pandemic, simultaneous outbreaks of influenza were seen in humans and swine. Interestingly, swine influenza was first recognized as a clinical entity in that species in the autumn of 1918 (Koen, 1919) concurrently with the spread of the second wave of the pandemic in humans (Dorset *et al.*, 1922). Investigators were impressed by the clinical and pathological similarities between human and swine influenza in 1918 (Koen, 1919; Murray & Biester, 1930). An extensive review by the veterinarian W. W. Dimoch of the diseases of swine published in August 1918 makes no mention of any swine disease resembling influenza (Dimoch, 1918). Thus, contemporary investigators were convinced that influenza virus had not circulated as an epizootic disease in swine before 1918 and that the virus spread from humans to pigs because of the appearance of illness in pigs after the first wave of the 1918 influenza in humans (Shope, 1936).

Thereafter, the disease became widespread among swine herds in the Midwest USA. The epizootic of 1919–1920 was as extensive as that in 1918–1919. The disease then appeared among swine in the Midwest every year, leading to R. E. Shope's isolation of the first influenza virus in 1930, A/swine/Iowa/30 (Shope & Lewis, 1931), 3 years before the isolation of the first human influenza virus, A/WS/33, by W. Smith, C. Andrewes and P. Laidlaw (Smith *et al.*, 1933). Classical swine viruses have continued to circulate not only in North American pigs but also in swine populations in Europe and Asia (Nerome *et al.*, 1982; Kupradinun *et al.*, 1991; Brown *et al.*, 1997).

During the fall and winter of 1918–1919, severe influenza-like disease outbreaks were noted not only in swine in the USA but also in Europe and China (Koen, 1919; Chun, 1919; Beveridge, 1977). The classical swine H1N1 lineage became endemic in swine herds in the USA and there are good data to support the global circulation of the 1918 influenza virus in pigs concurrently with its circulation in humans. Since 1918, there have been many examples of both H1N1 and H3N2 human influenza A virus strains becoming established in swine (Castrucci *et al.*, 1993; Brown *et al.*, 1998; Zhou *et al.*, 2000). Unfortunately for the argument that swine might have served as the intermediate between avian and humans in 1918, swine influenza virus strains have been isolated only sporadically from humans (Gaydos *et al.*, 1977; Woods *et al.*, 1981; Rimmelzwaan *et al.*, 2001). It seems probable that at least during the height of the 1918 pandemic, the direction of transmission was from humans to pigs. However, is it possible that before the pandemic, the originally avian HA was gradually adapting into a swine influenza virus strain?

Interestingly, an avian H1N1 lineage has become established

in European swine in the last 20 years, providing a model for the evolution of avian viruses in pigs. As noted earlier, the 1918 HA1 sequence had many more amino acid differences from avian sequences than did the 1957 and 1968 pandemic strains but very few of these change were in antigenic sites, suggesting that the 1918 HA had not been subjected to significant selective pressure before emerging as a pandemic. In phylogenetic analyses, the 1918 HA is always placed in the mammalian clade. It would be interesting to note whether, at some point in the evolution of an avian H1N1 lineage in European pigs, a similar degree of divergence from the avian clade would be found. The earliest avian-like H1N1 strains were isolated from swine in Northern Europe in 1979 and 1980. A/swine/Arnsberg/6554/79 has 12 amino acid differences from the avian consensus sequence and A/Swine/Netherlands/3/80 has seven differences. In both cases, three of the differences are in antigenic sites. In contrast, the 1918 HA has 28 amino acid differences from the avian consensus sequence, of which four are in antigenic sites. The latest avian-like H1N1 isolated from swine in Europe from which sequence is available, A/swine/Belgium/117/96, has 17 differences from the avian consensus sequence, of which five are in antigenic sites. Furthermore, phylogenetic analyses place even A/swine/Belgium/117/96 in the avian clade. Thus, it appears that even 20 years of evolution in swine has not resulted in the number of changes from the avian consensus sequence exhibited by the 1918 pandemic strain.

## CONCLUSION

Understanding the origin of the 1918 pandemic influenza virus strain is not a question of idle historical curiosity. Given the speed with which a new pandemic could spread in the modern world, the emergence of a strain as virulent as that of 1918 would be devastating. Current surveillance efforts focused on rapid identification of novel strains in humans as well as efforts to minimize the possibility of cross-infection between species are aimed at detecting and preventing a new pandemic. However, it is important to recognize that the mechanisms by which pandemic strains originate have not been explained yet. It seems likely that the 1957 and 1968 pandemic strains originated in the reassortment of avian and human strains. However, the actual circumstances and time-course of these reassortment events were not detected at the time and, therefore, it is unclear how long it would take such a reassorted strain to develop into a pandemic. The 1918 pandemic strain is even more puzzling because its HA sequence is neither consistent with direct reassortment with a known avian H1 strain nor with adaptation in swine. Sequencing of more avian H1 strains and research into alternative intermediate hosts than swine, such as poultry or horses, may shed further light on the origins of the 1918 pandemic. Until its origins are understood better, detection and prevention efforts may overlook the beginning of the next pandemic.

## ACKNOWLEDGEMENTS

This work was supported in part by grants from the NIH and by the intramural funds of the Armed Forces Institute of Pathology. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or Department of Defense. This is a US government work; there are no restrictions on its use.

## REFERENCES

- Alexander, D. J. & Brown, I. H. (2000). Recent zoonoses caused by influenza A viruses. *Rev Sci Tech* **19**, 197–225.
- Basler, C. F., Reid, A. H., Dybing, J. K. & 9 other authors (2001). Sequence of the 1918 pandemic influenza virus nonstructural gene (NS) segment and characterization of recombinant viruses bearing the 1918 NS genes. *Proc Natl Acad Sci U S A* **98**, 2746–2751.
- Bean, W. J., Schell, M., Katz, J., Kawaoka, Y., Naeve, C., Gorman, O. & Webster, R. G. (1992). Evolution of the H3 influenza virus hemagglutinin from human and nonhuman hosts. *J Virol* **66**, 1129–1138.
- Beveridge, W. (1977). *Influenza: the Last Great Plague, an Unfinished Story of Discovery*. New York: Prodist.
- Brammer, T. L., Murray, E. L., Fukuda, K., Hall, H. E., Klimov, A. & Cox, N. J. (2002). Surveillance for influenza: United States, 1997–98, 1998–99, and 1999–00 seasons. *MMWR Surveill Summ* **51**, 1–10.
- Brown, I. H., Ludwig, S., Olsen, C. W. & 7 other authors (1997). Antigenic and genetic analyses of H1N1 influenza A viruses from European pigs. *J Gen Virol* **78**, 553–562.
- Brown, I. H., Harris, P. A., McCauley, J. W. & Alexander, D. J. (1998). Multiple genetic reassortment of avian and human influenza A viruses in European pigs, resulting in the emergence of an H1N2 virus of novel genotype. *J Gen Virol* **79**, 2947–2955.
- Brownlee, G. G. & Fodor, E. (2001). The predicted antigenicity of the haemagglutinin of the 1918 Spanish influenza pandemic suggests an avian origin. *Philos Trans R Soc Lond B Biol Sci* **356**, 1871–1876.
- Castrucci, M. R., Donatelli, I., Sidoli, L., Barigazzi, G., Kawaoka, Y. & Webster, R. G. (1993). Genetic reassortment between avian and human influenza A viruses in Italian pigs. *Virology* **193**, 503–506.
- Chun, J. (1919). Influenza including its infection among pigs. *Natl Med J China (Peking)* **5**, 34–44.
- Claas, E. C., Kawaoka, Y., de Jong, J. C., Masurel, N. & Webster, R. G. (1994). Infection of children with avian–human reassortant influenza virus from pigs in Europe. *Virology* **204**, 453–457.
- Claas, E. C., Osterhaus, A. D., van Beek, R., De Jong, J. C., Rimmelzwaan, G. F., Senne, D. A., Krauss, S., Shortridge, K. F. & Webster, R. G. (1998). Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. *Lancet* **351**, 472–477.
- Cox, N. J. & Subbarao, K. (2000). Global epidemiology of influenza: past and present. *Annu Rev Med* **51**, 407–421.
- Dimoch, W. W. (1918). Diseases of swine. *J Am Vet Med Assoc* **54**, 321–340.
- Dorset, M., McBryde, C. N. & Niles, W. B. (1922). Remarks on ‘hog’ flu. *J Am Vet Med Assoc* **62**, 162–171.
- Fang, R., Min Jou, W., Huylebroeck, D., Devos, R. & Fiers, W. (1981). Complete structure of A/duck/Ukraine/63 influenza hemagglutinin gene: animal virus as progenitor of human H3 Hong Kong 1968 influenza hemagglutinin. *Cell* **25**, 315–323.
- Fanning, T. G., Slemons, R. D., Reid, A. H., Janczewski, T. A., Dean, J. & Taubenberger, J. K. (2002). 1917 avian influenza virus sequences suggest that the 1918 pandemic virus did not acquire its hemagglutinin directly from birds. *J Virol* **76**, 7860–7862.
- Gaydos, J., Hodder, R., Top, F. J., Soden, V., Allen, R., Bartley, J., Zabkar, J., Nowosiwsky, T. & Russell, P. (1977). Swine influenza A at Fort Dix, New Jersey (January–February 1976). I. Case finding and clinical study of cases. *J Infect Dis* **136**, 356–362.
- Hatta, M. & Kawaoka, Y. (2002). The continued pandemic threat posed by avian influenza viruses in Hong Kong. *Trends Microbiol* **10**, 340–344.
- Jordan, E. (1927). *Epidemic Influenza: a Survey*, pp. 355. Chicago: American Medical Association.
- Katz, J. M., Lim, W., Bridges, C. B. & 13 other authors (1999). Antibody response in individuals infected with avian influenza A (H5N1) viruses and detection of anti-H5 antibody among household and social contacts. *J Infect Dis* **180**, 1763–1770.
- Kawaoka, Y., Krauss, S. & Webster, R. G. (1989). Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics. *J Virol* **63**, 4603–4608.
- Kilbourne, E. D. (1977). Influenza pandemics in perspective. *JAMA* **237**, 1225–1228.
- Kilbourne, E. D. (1997). Perspectives on pandemics: a research agenda. *J Infect Dis* **176** (Suppl. 1), S29–S31.
- Koen, J. S. (1919). A practical method for field diagnoses of swine diseases. *Am J Vet Med* **14**, 468–470.
- Kupradinun, S., Peanpajit, P., Bhodhikosoom, C., Yoshioka, Y., Endo, A. & Nerome, K. (1991). The first isolation of swine H1N1 influenza viruses from pigs in Thailand. *Arch Virol* **118**, 289–297.
- Lamb, R. & Krug, R. (2001). *Orthomyxoviridae: The viruses and their replication*. In *Fields Virology*, 4th edn, pp. 1487–1531. Edited by D. Knipe & P. Howley. Philadelphia: Lippincott Williams & Wilkins.
- Layne, S. P., Beugelsdijk, T. J., Patel, C. K., Taubenberger, J. K., Cox, N. J., Gust, I. D., Hay, A. J., Tashiro, M. & Lavanchy, D. (2001). A global lab against influenza. *Science* **293**, 1729.
- Lin, Y. P., Shaw, M., Gregory, V. & 10 other authors (2000). Avian-to-human transmission of H9N2 subtype influenza A viruses: relationship between H9N2 and H5N1 human isolates. *Proc Natl Acad Sci U S A* **97**, 9654–9658.
- Linder, F. E. & Grove, R. D. (1943). *Vital Statistics Rates in the United States: 1900–1940*. Washington, DC: Government Printing Office.
- Ludwig, S., Stitz, L., Planz, O., Van, H., Fitch, W. & Scholtissek, C. (1995). European swine virus as a possible source for the next influenza pandemic? *Virology* **212**, 555–561.
- Marks, G. & Beatty, W. K. (1976). *Epidemics*. New York: Scribner.
- Marozin, S., Gregory, V., Cameron, K. & 7 other authors (2002). Antigenic and genetic diversity among swine influenza A H1N1 and H1N2 viruses in Europe. *J Gen Virol* **83**, 735–745.
- Martinez, C., del Rio, L., Portela, A., Domingo, E. & Ortin, J. (1983). Evolution of the influenza virus neuraminidase gene during drift of the N2 subtype. *Virology* **130**, 539–545.
- Matrosovich, M. N., Gambaryan, A. S., Teneberg, S., Piskarev, V. E., Yamnikova, S. S., Lvov, D. K., Robertson, J. S. & Karlsson, K. A. (1997). Avian influenza A viruses differ from human viruses by recognition of sialyloligosaccharides and gangliosides and by a higher conservation of the HA receptor-binding site. *Virology* **233**, 224–234.
- Meltzer, M. I., Cox, N. J. & Fukuda, K. (1999). The economic impact of pandemic influenza in the United States: priorities for intervention. *Emerg Infect Dis* **5**, 659–671.

- Murray, C. & Biester, H. E. (1930).** Swine influenza. *J Am Vet Med Assoc* **76**, 349–355.
- Nakajima, K., Desselberger, U. & Palese, P. (1978).** Recent human influenza A (H1N1) viruses are closely related genetically to strains isolated in 1950. *Nature* **274**, 334–339.
- Nerome, K., Ishida, M., Oya, A. & Oda, K. (1982).** The possible origin H1N1 (Hsw1N1) virus in the swine population of Japan and antigenic analysis of the isolates. *J Gen Virol* **62**, 171–175.
- Olsen, C. W. (2002).** The emergence of novel swine influenza viruses in North America. *Virus Res* **85**, 199–210.
- Palese, P. & Young, J. F. (1982).** Variation of influenza A, B, and C viruses. *Science* **215**, 1468–1474.
- ProMED Mail (2003).** Avian influenza: Netherlands. *ProMED Mail*. Archive no. 20030306.0552.
- Reid, A. H., Fanning, T. G., Hultin, J. V. & Taubenberger, J. K. (1999).** Origin and evolution of the 1918 ‘Spanish’ influenza virus hemagglutinin gene. *Proc Natl Acad Sci U S A* **96**, 1651–1656.
- Reid, A. H., Fanning, T. G., Janczewski, T. A. & Taubenberger, J. K. (2000).** Characterization of the 1918 ‘Spanish’ influenza virus neuraminidase gene. *Proc Natl Acad Sci U S A* **97**, 6785–6790.
- Reid, A. H., Fanning, T. G., Janczewski, T. A., McCall, S. & Taubenberger, J. K. (2002).** Characterization of the 1918 ‘Spanish’ influenza virus matrix gene segment. *J Virol* **76**, 10717–10723.
- Reid, A. H., Janczewski, T. A., Lourens, R. A., Elliot, A. J., Daniels, R. S., Berry, C. L., Oxford, J. S. & Taubenberger, J. K. (2003).** 1918 influenza pandemic caused by highly conserved viruses with two receptor-binding variants. *Emerg Infect Dis* (in press).
- Rimmelzwaan, G. F., de Jong, J. C., Bestebroer, T. M., van Loon, A. M., Claas, E. C., Fouchier, R. A. & Osterhaus, A. D. (2001).** Antigenic and genetic characterization of swine influenza A (H1N1) viruses isolated from pneumonia patients in The Netherlands. *Virology* **282**, 301–306.
- Rosenau, M. J. & Last, J. M. (1980).** *Maxcy-Rosenau Preventative Medicine and Public Health*. New York: Appleton-Century-Crofts.
- Schafer, J. R., Kawaoka, Y., Bean, W. J., Suss, J., Senne, D. & Webster, R. G. (1993).** Origin of the pandemic 1957 H2 influenza A virus and the persistence of its possible progenitors in the avian reservoir. *Virology* **194**, 781–788.
- Scholtissek, C. (1995).** Molecular evolution of influenza viruses. *Virus Genes* **11**, 209–215.
- Scholtissek, C., Koennecke, I. & Rott, R. (1978a).** Host range recombinants of fowl plague (influenza A) virus. *Virology* **91**, 79–85.
- Scholtissek, C., Rohde, W., Von Hoyningen, V. & Rott, R. (1978b).** On the origin of the human influenza virus subtypes H2N2 and H3N2. *Virology* **87**, 13–20.
- Shope, R. E. (1936).** The incidence of neutralizing antibodies for swine influenza virus in the sera of human beings of different ages. *J Exp Med* **63**, 669–684.
- Shope, R. E. & Lewis, P. A. (1931).** Swine influenza. *J Exp Med* **54**, 349–359.
- Simonsen, L., Clarke, M. J., Schonberger, L. B., Arden, N. H., Cox, N. J. & Fukuda, K. (1998).** Pandemic versus epidemic influenza mortality: a pattern of changing age distribution. *J Infect Dis* **178**, 53–60.
- Simonsen, L., Fukuda, K., Schonberger, L. B. & Cox, N. J. (2000).** The impact of influenza epidemics on hospitalizations. *J Infect Dis* **181**, 831–837.
- Smith, W., Andrewes, C. & Laidlaw, P. (1933).** A virus obtained from influenza patients. *Lancet* **225**, 66–68.
- Subbarao, K., Klimov, A., Katz, J. & 13 other authors (1998).** Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. *Science* **279**, 393–396.
- Taubenberger, J. K. & Layne, S. P. (2001).** Diagnosis of influenza virus: coming to grips with the molecular era. *Mol Diagn* **6**, 291–305.
- Taubenberger, J. K., Reid, A. H., Krafft, A. E., Bijwaard, K. E. & Fanning, T. G. (1997).** Initial genetic characterization of the 1918 ‘Spanish’ influenza virus. *Science* **275**, 1793–1796.
- Taubenberger, J. K., Reid, A. H., Janczewski, T. A. & Fanning, T. G. (2001).** Integrating historical, clinical and molecular genetic data in order to explain the origin and virulence of the 1918 Spanish influenza virus. *Philos Trans R Soc Lond B Biol Sci* **356**, 1829–1839.
- Thompson, W. W., Shay, D. K., Weintraub, E., Brammer, L., Cox, N., Anderson, L. J. & Fukuda, K. (2003).** Mortality associated with influenza and respiratory syncytial virus in the United States. *JAMA* **289**, 179–186.
- Tsuchiya, E., Sugawara, K., Hongo, S., Matsuzaki, Y., Muraki, Y., Li, Z. N. & Nakamura, K. (2001).** Antigenic structure of the haemagglutinin of human influenza A/H2N2 virus. *J Gen Virol* **82**, 2475–2484.
- Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M. & Kawaoka, Y. (1992).** Evolution and ecology of influenza A viruses. *Microbiol Rev* **56**, 152–179.
- Webster, R. G., Sharp, G. B. & Claas, E. C. (1995).** Interspecies transmission of influenza viruses. *Am J Respir Crit Care Med* **152**, 25–30.
- Wiley, D. C., Wilson, I. A. & Skehel, J. J. (1981).** Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature* **289**, 373–378.
- Winter, G., Fields, S. & Brownlee, G. G. (1981).** Nucleotide sequence of the haemagglutinin gene of a human influenza virus H1 subtype. *Nature* **292**, 72–75.
- Winternitz, M. C., Wason, I. M. & McNamara, F. P. (1920).** *The Pathology of Influenza*. New Haven: Yale University Press.
- Wolbach, S. B. (1919).** Comments on the pathology and bacteriology of fatal influenza cases, as observed at Camp Devens, Mass. *Johns Hopkins Hospital Bull* **30**, 104–109.
- Woods, G. T., Schnurrenberger, P. R., Martin, R. J. & Tompkins, W. A. (1981).** Swine influenza virus in swine and man in Illinois. *J Occup Med* **23**, 263–267.
- Wright, P. & Webster, R. (2001).** Orthomyxoviruses. In *Fields Virology*, 4th edn, pp. 1533–1579. Edited by D. Knipe & P. Howley. Philadelphia: Lippincott Williams & Wilkins.
- Young, J. F. & Palese, P. (1979).** Evolution of human influenza A viruses in nature: recombination contributes to genetic variation of H1N1 strains. *Proc Natl Acad Sci U S A* **76**, 6547–6551.
- Zhou, N. N., Senne, D. A., Landgraf, J. S. & 7 other authors (2000).** Emergence of H3N2 reassortant influenza A viruses in North American pigs. *Vet Microbiol* **74**, 47–58.