The evolution of human influenza viruses

Alan J. Hay[*]**, Victoria Gregory, Alan R. Douglas and Yi Pu Lin**

National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

The evolution of influenza viruses results in (i) recurrent annual epidemics of disease that are caused by progressive antigenic drift of influenza A and B viruses due to the mutability of the RNA genome and (ii) infrequent but severe pandemics caused by the emergence of novel influenza A subtypes to which the population has little immunity. The latter characteristic is a consequence of the wide antigenic diversity and peculiar host range of influenza A viruses and the ability of their segmented RNA genomes to undergo frequent genetic reassortment (recombination) during mixed infections. Contrasting features of the evolution of recently circulating influenza AH1N1, AH3N2 and B viruses include the rapid drift of AH3N2 viruses as a single lineage, the slow replacement of successive antigenic variants of AH1N1 viruses and the co-circulation over some 25 years of antigenically and genetically distinct lineages of influenza B viruses. Constant monitoring of changes in the circulating viruses is important for maintaining the efficacy of influenza vaccines in combating disease.

Keywords: influenza virus; evolution; influenza A; influenza B; genetic reassortment; antigenic drift

1. INTRODUCTION

The outstanding feature of influenza viruses is their capacity for evading host immunity and causing recurrent annual epidemics of disease and, at infrequent intervals, major worldwide pandemics due to the introduction of antigenically novel viruses into an immunologically naive human population. Shortly after the isolation of the first human influenza virus in 1933 (Smith *et al.* 1933), recognition of the changeability in antigenic characteristics led to the World Health Organization's (WHO) establishment of a global Influenza Surveillance Network for monitoring changes in the viruses causing outbreaks of influenza throughout the year in different parts of the world. The network was initially centred around the World Influenza Centre in London. Today, it includes four international WHO collaborating centres in Atlanta, London, Melbourne and Tokyo together with *ca.* 110 national influenza centres located in some 80 different countries. In addition, a WHO collaborating centre in Memphis, TN studying animal and avian influenza viruses provides an important interface for the rapid identification of exotic human isolates. The principal objectives are: (i) the early detection and characterization of novel subtypes of human influenza A with the potential for causing pandemics, such as the recent cases of influenza AH5N1 and AH9N2 infection in Hong Kong (Claas *et al.* 1998; Lin *et al.* 2000) and (ii) the identification of new antigenic variants among currently circulating influenza A and B viruses in order to ensure that influenza vaccines contain components that reflect the immunological characteristics of the prevalent viruses. The WHO has published formal recommendations for the compositions of influenza vaccines based on cumulative data since 1973. More recently, since 1998, two recommendations per year

in February and September relate to vaccines for use during the following winters in the northern and southern hemispheres, respectively.

The epidemiology of human influenza reflects the peculiar characteristics of the virus genome (segmented single-stranded RNA) and the diversity and host range of the viruses. Two types, namely influenza A and influenza B viruses, are responsible for recurrent annual epidemics. They are antigenically distinct and do not exhibit cross-immunity, nor do they undergo intertypic genetic reassortment (recombination). These viruses co-circulate and either may predominate in a particular influenza season; an increased incidence of influenza B frequently follows a peak of influenza A activity. In recent years, B viruses have tended to be prominent every 2–3 years. The interplay between these viruses and their host may influence their impact and evolution.

Although influenza B viruses have been responsible for severe epidemics, the impact of influenza A viruses is greater in terms of annual epidemics as well as the infrequent more devastating pandemics. The latter characteristic and the evolution of influenza A viruses is a consequence of their greater genetic diversity and, in particular, their unique host range. Whereas influenza B viruses almost exclusively infect humans, influenza A viruses are essentially avian viruses that periodically transmit to other species, including mammals. Furthermore, influenza A viruses comprise a large variety of antigenically distinct subtypes, with different combinations of 15HA and 9NA subtypes, that replicate asymptomatically in the intestine of aquatic birds, particularly ducks and constitute a large reservoir of potential pandemic viruses (Webster *et al.* 1978, 1992). Only a few subtypes are known to have established in humans, three of them since 1933. The population is therefore vulnerable to severe infection by antigenically novel viruses to which it is immunologically naive, as emphatically demonstrated by the
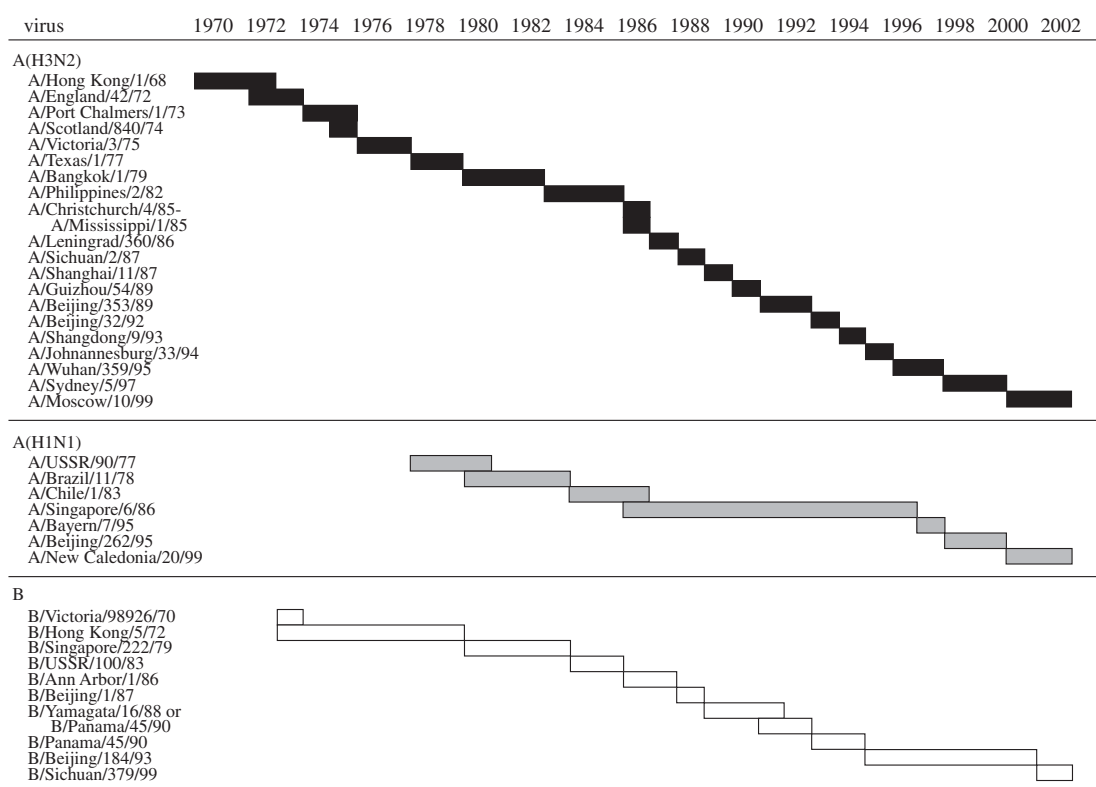
Figure 1. Changes in the influenza vaccine compositions recommended by the WHO, 1973–2001. The viruses listed are the prototypes recommended for inclusion in the bivalent or trivalent (1977 onwards) vaccine. Adapted from World Health Organization (1996, fig. 1).

influenza AH5N1 infections in Hong Kong in 1997, one-third (six out of 18) of which were fatal (Bender *et al.* 1999). Although various genetic factors have been shown to determine species specificity in experimental situations, we still have a poor understanding of the combination of factors that are important in restricting or facilitating interspecies transmission. Genetic reassortment clearly has an important role to play in introducing novel antigens. Although the origin of the AH1N1 subtype, that circulated between 1918 and 1957, is not yet known (Taubenberger *et al.* 1997; Reid *et al.* 1999), the subtypes that emerged to cause the 1957 Asian and 1968 Hong Kong pandemics were the products of genetic reassortment between the genomes of the previously circulating human virus and avian viruses (Schafer *et al.* 1993). The mechanism whereby this occurred and possible intermediate hosts are discussed in Webby (2001). The new subtypes that emerged in 1957 and 1968 replaced the preceding subtype. In contrast, the AH1N1 viruses that reappeared in 1977 and were closely related to viruses circulating in the human population in 1950 (Kendal *et al.* 1978) did not replace the preceding H3N2 subtype. Thus, two influenza A subtypes, namely H1N1 and H3N2, have co-circulated with influenza B viruses since 1977 and current trivalent inactivated vaccines contain representatives of each A subtype and B virus.

Despite the severity of pandemics, such as those of 1957 and 1968, the cumulative impact of annual epidemics during interpandemic periods exceeds that of pandemics. The timing and speed as well as the nature of antigenic changes are unpredictable. Changes in the prevalence of H3N2 variants generally occur rapidly, for example A/

Sydney/5/97(H3N2)-like viruses spread to all parts of the world within 6 months of their initial detection and rapidly replaced the previously prevalent A/Wuhan/359/95-like viruses. Co-circulation of antigenically distinguishable variants of influenza AH1N1 and B viruses reflects the slower emergence of novel antigenic strains. The extent of variation in the antigenic properties of the viruses is reflected in the number of changes recommended in vaccine composition (figure 1). The greater antigenic variability of AH3N2 viruses has required 19 changes in the vaccine component over 29 years (since 1972). In contrast, the 10 changes to the influenza B component and six changes to the AH1N1 component made during this period reflect the slower rates of antigenic change of these viruses.

2. INFLUENZA A H3N2 VIRUSES

Antibody to the virus haemagglutinin, that neutralizes virus infectivity by preventing attachment to cell receptors (Fleury *et al.* 1999), is of prime importance in immunity. Antigenic drift results from the accumulation of amino acid changes that abrogate antibody binding and, consequently, reduces immunity to subsequent infection by antigenic variants. Table 1 illustrates the antigenic differences between selected successive H3N2 variants that are distinguished by haemagglutination inhibition using post-infection ferret antisera. Ferrets readily succumb to infection by human influenza viruses with symptoms comparable with those of the human disease and produce antibodies that are particularly useful in distinguishing antigenic variants that are of epidemiological significance

Table 1. Antigenic drift of influenza AH3N2 viruses: haemagglutination inhibition titres.

(Homologous titres are marked in bold. $< = <40$.)

	post-infection ferret sera												
virus	A/HK 1/68	A/Eng 42/72	A/Vic 3/75	A/Tex 1/77	A/Bk 1/79	A/Phil 2/82	A/Miss 1/85	A/Shan 11/87	A/Beij 352/89	A/Beij 32/92	A/Jhb 33/94	A/Wuh 359/95	A/Syd 5/97
A/Hong Kong/1/68	**1280**	320	<	<	<	<	<	<	<	<	<	<	<
A/England/42/72	40	**640**	40	<	<	<	<	<	<	<	<	<	<
A/Victoria/3/75	<	<	**640**	<	<	<	<	<	<	<	<	<	<
A/Texas/1/77	40	40	80	**1280**	320	160	<	40	<	<	<	<	<
A/Bangkok/1/79	<	40	40	320	**1280**	160	<	80	40	<	<	<	<
A/Philippines/2/82	<	<	40	80	80	**640**	80	160	80	<	<	<	<
A/Mississippi/1/85	<	<	<	40	80	80	**1280**	160	80	<	<	<	<
A/Shanghai/11/87	<	40	<	40	80	80	40	**640**	80	<	<	<	<
A/Beijing/352/89	<	<	<	<	<	<	<	80	**2560**	<	<	<	<
A/Beijing/32/92	<	<	<	<	<	<	<	<	80	**640**	80	<	<
A/Johannesburg/33/94	<	<	<	<	<	<	<	<	40	80	**640**	80	<
A/Wuhan/359/95	<	<	<	<	<	<	<	<	<	40	40	**1280**	160
A/Sydney/5/97	<	<	<	<	<	<	<	<	<	<	<	160	**2560**

in the human population. Antigenic difference defined in this way is the principal criterion upon which changes in vaccine composition are based.

The stepwise emergence of the principal antigenic variants is emphasized by the phylogenetic relationships between HA genes (figure 2a), which show that the H3 HA gene has evolved as a single lineage since the introduction of H3N2 viruses into the human population in 1968, with the epidemiologically significant variants close to the main trunk of the tree (Bush *et al.* 1999). The rate of accumulation of mutations is *ca.* 4×10^{-3} substitutions per nucleotide per year, with *ca.* 5×10^{-3} amino acid substitutions per residue per year in HA1. Most of these changes are on the surface of HA1 and are present in antigenic sites that were identified using monoclonal antibodies and are mainly located close to the receptor-binding site (Wiley *et al.* 1981; Skehel & Wiley 2000). In addition to the direct structural consequences of amino acid substitutions, changes in glycosylation are important in modifying antibody binding to HA (Skehel *et al.* 1984). This is reflected in an increase in the number of glycosylation sites on the upper surface of HA1 during evolution of the H3N2 viruses, from three in A/Hong Kong/68 to seven in recent A/Sydney/5/97-like viruses (listed in the legend to figure 3). Changes in conserved amino acids in the receptor-binding site, in particular residue 226 that is an important determinant of receptor specificity (Rogers *et al.* 1983), have also been associated with evolution of H3N2 viruses since 1992. Variation at this position between leucine or glutamine in the HAs of viruses isolated between 1992 and 1995, that correlated with growth capacity in mammalian or avian cells, respectively, was followed by a predominance of isoleucine and, in more recent A/Sydney/5/97-like variants, by valine 226. The significance of these changes is unclear. However, changes in the affinity of receptor binding may be associated with complementary changes in the receptor-destroying activity of the neuraminidase, as well as immune evasion.

Although antibodies to the neuraminidase do not neutralize virus infectivity, they have an important role in reducing virus spread and they also impose selection pressure. The NA genes of AH3N2 viruses exhibit a similar phylogenetic relationship to that of HA genes (figure 2b) and have undergone a similar degree of change during the past 33 years. Many of the amino acid changes associated with antigenic drift cluster on the upper surface of the enzyme encircling the catalytic site (Colman *et al.* 1983).

As illustrated by the 10 changes in vaccine composition recommended between 1986 and 1998 (figure 1), new antigenic variants can arise with a frequency of one per 1–2 years, although antigenic changes have been less marked since the emergence of the A/Sydney/5/97-like viruses. Replacement of the pre-existing variant therefore occurs quite quickly. For example, following identification of the A/Sydney/5/97-like viruses in the southern hemisphere in mid-1997 they spread rapidly to other parts of the world, were the predominant viruses circulating in the northern hemisphere winter of 1997–1998 and had replaced the former A/Wuhan/359/95-like viruses by April of 1998. Sequence analyses, that define the variants in terms of common amino acid differences within HA and NA, can also distinguish the emergence of other variants that are antigenically indistinguishable from the vaccine strain, some of which may represent intermediates in the emergence of subsequent antigenic variants. For example, a variant represented by A/Thessaloniki/1/ 95 was intermediate in amino acid sequence (HA and NA) between those of A/Johannesburg/33/94 and A/Wuhan/ 359/95 and was the predominant A/Johannesburg/33/94-like (antigenically) variant circulating during 1995–1996 prior to the emergence of A/Wuhan/359/95-like viruses. In turn, A/South Africa/1147/96, that was representative of viruses prominent during 1996, was intermediate in the emergence of A/Sydney/5/97 (figure 3). In contrast, another variant, that was represented by A/Bratislava/6/ 97 and was prevalent during 1996–1997, was more distantly related (in both HA and NA genes) to subsequent viruses and represented a more distinct branch on the phylogenetic tree. In particular, the NA gene of these viruses

(a)

A/Prague/5/00
A/Ireland/10586/99
A/Karditsa/186/00
A/Sachsen/434/01
A/Hannover/47/00
A/Israel/83/00
A/Hong Kong/1319/00
A/PANAMA/2007/99
A/Denmark/41/00
A/Bucharest/226/99
A/Perugia/5/99
A/Finland/620/99
A/Hong Kong/1221/99
A/Darwin/7/99
A/MOSCOW/10/99
A/Victoria/318/98
A/Sichuan/346/98
A/Johannesburg/29/98
A/Genoa/8/98
A/SYDNEY/5/97
A/Hong Kong/374/97
A/South Africa/1147/96
A/Zamora/3/96
A/Bratislava/6/97
A/WUHAN/359/95
A/NANCHANG/933/95
A/Thessaloniki/1/95
A/JOHANNESBURG/33/94
A/SHANDONG/9/93
A/BEIJING/32/92
A/BEIJING/353/89
A/GUIZHOU/54/89
A/SICHUAN/2/87
A/SHANGHAI/11/87
A/CHRISTCHURCH/4/85
A/LENINGRAD/360/86
A/MISSISSIPPI/1/85
A/PHILIPPINES/2/82
A/BANGKOK/1/79
A/TEXAS/1/77
A/VICTORIA/3/75
A/PORT CHALMERS/1/73
A/ENGLAND/42/72
A/AICHI/2/68

100 nucleotides

(b)

A/Prague/5/00
A/Ireland/10586/99
A/Latvia/1097/00
A/Lyon/1242/00
A/Hong Kong/1789/00
A/Finland/620/99
A/MOSCOW/10/99
A/Hong Kong/1221/99
A/Denmark/511/98
A/Perugia/5/99
A/PANAMA/2007/99
A/Philippines/26/99
A/Hong Kong/6/99
A/SYDNEY/5/97
A/Genoa/5/98
A/Sichuan/346/98
A/Johannesburg/29/98
A/South Africa/1147/96
A/NANCHANG/933/95
A/Johannesburg/23/97
A/Genoa/9/96
A/Bratislava/6/97
A/Thessalonika/1/95
A/JOHANNESBURG/33/94
A/Lyon/7/94
A/SHANDONG/9/93
A/BEIJING/353/89
A/Washington/15/91
A/Shanghai/24/90
A/BEIJING/32/92
A/Hong Kong/34/90
A/SHANGHAI/11/87
A/England/427/88
A/MISSISSIPPI/1/85
A/BANGKOK/1/79
A/Udon/72
A/VICTORIA/3/75
A/Hong Kong/107/71
A/Hong Kong/6/68
A/Tokyo/3/67
A/Leningrad/17/57

100 nucleotides

Figure 2. Phylogenetic relationships of the (a) HA and (b) NA genes of influenza A H3N2 viruses. The sequences of nucleotides 1–1083 encoding the HA1 domain and nucleotides 1–1409 encoding the neuraminidase were compared using *Phylogenetic Analysis Using Parsimony*, version 4.0, (Swofford & Olsen 1990). The lengths of the horizontal lines are proportional to the number of nucleotide differences, as indicated by the bar. Sequences not determined in this study were obtained from Genbank. The vaccine strains are in capitals.

was more closely related to the earlier A/Thessaloniki/1/95 variant (figure 2b). Although distinguishable variants do co-circulate, they generally persist for only short periods and are quickly replaced by subsequent variants. Thus, despite the variation among H3N2 viruses, there is usually a high degree of similarity among the predominant viruses causing outbreaks of disease in different parts of the world, although differences in the geographical distribution of less prevalent variants are apparent.

The interplay between the antigenic drift of the HA and NA during the evolution of these viruses is not well understood. However, it is apparent that they may vary independently and that significant changes in the antigenicity of NA do not necessarily coincide with those in HA (Xu *et al.* 1996). For example, following the emergence

of the A/Sydney/5/97-like viruses, greater change was observed in NA with the emergence of an antigenically distinguishable NA that was represented by that of A/Moscow/10/99 (although antigenic relationships between NAs, as determined by neuraminidase inhibition, are less well defined than those between HAs). Furthermore, co-circulation of several H3N2 variants with different combinations of three distinguishable variants of HA and NA (represented by the NA genes of A/Johannesburg/29/98, A/Panama/2007/99 and A/Moscow/10/99) (figure 2b) during 1998–1999 indicated the importance of reassortment of HA and NA genes in the emergence of the epidemic viruses. Reassortment among the other six genes is also involved in the evolution of these viruses (Lindstrom *et al.* 1998). However, little is known about
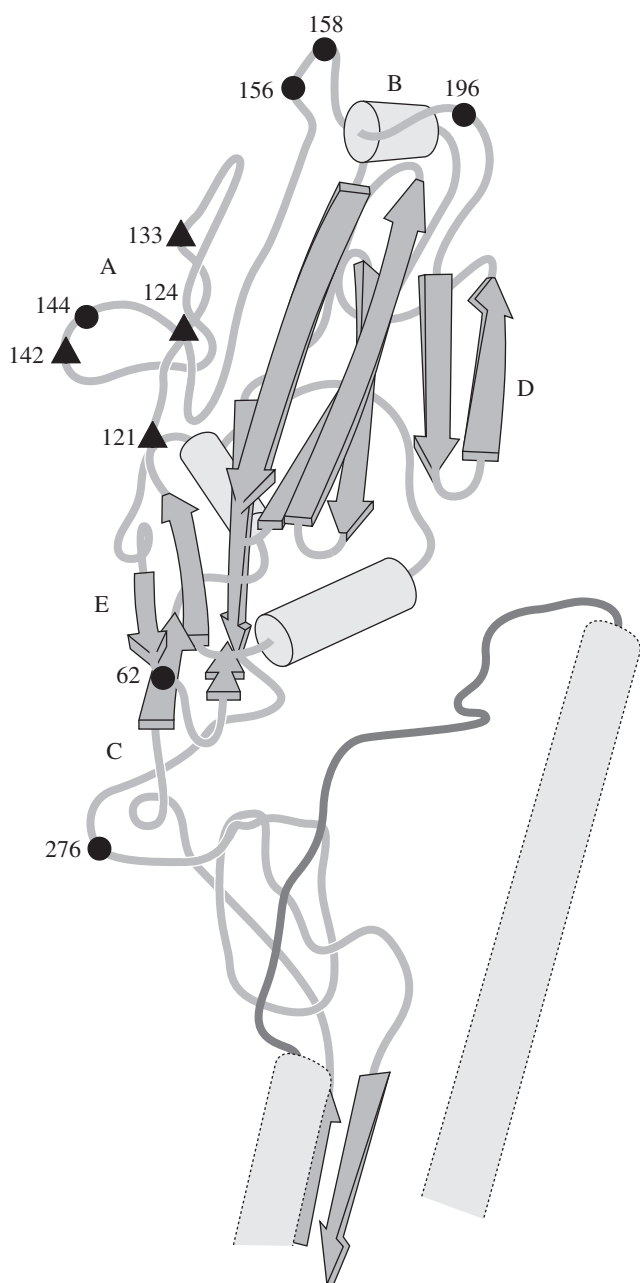
Figure 3. Schematic diagram showing the locations of the principal amino acid differences on the HA1 domain between the HAs of A/Wuhan/359/95 and A/South Africa/1147/96 viruses (solid triangles) and between the HAs of A/South Africa/1147/96 and A/Sydney/5/97 (solid circles) viruses, which together account for the principal differences between A/Wuhan/359/95-like and A/Sydney/5/97-like viruses. The locations of antigenic sites A–E are indicated. Changes to serine 124 and asparagine 133 introduced glysosylation sites at residues 122 and 133 in addition to those at asparagines 8, 22, 38, 63, 126, 165, 246 and 285.

the influence of the variation in the products of these genes on the impact of disease.

3. INFLUENZA A H1N1 VIRUSES

Viruses of the H1N1 subtype were responsible for the devastating pandemic of 1918–1919 that caused an estimated 20–50 millions deaths worldwide. The genetic

characteristics of these viruses and their relationships to H1N1 viruses subsequently circulating in human, avian and swine populations are described in Taubenberger (2001). Of note is the reappearance of AH1N1 viruses in 1977 following their absence from the human population for 20 years and their similarity to viruses circulating in 1950. As a consequence of this relationship, people aged 27 years or more who had prior exposure to these viruses had substantial immunity to infection and illness occurred mainly in the younger population.

Although the rates of accumulation of mutations in HA and NA genes were comparable with those of H3N2 viruses and the antigenic structure of H1 is similar to that of H3 (Raymond *et al.* 1986), significant antigenic changes occurred less frequently (table 2), which is apparent in the fewer changes recommended for the H1N1 vaccine component since 1977 (figure 1). In particular, there was little change in the antigenic properties of the H1N1 viruses circulating between 1986 and 1995 (table 2) and recent isolates of A/Bayern/7/95-like viruses are still (after 15 years) closely related to A/Singapore/6/86. Furthermore, in contrast to antigenic drift among H3N2 viruses, the antigenic variants of H1N1 viruses that emerged in 1986, as represented by A/Singapore/6/86 and 1995, as represented by A/Beijing/262/95, did not evolve directly from the preceding circulating viruses. For example, the HA and NA genes of the contemporary variants A/Beijing/262/95 and A/Bayern/7/95 had diverged (in the early 1990s) by *ca.* 4% and represent distinct co-circulating lineages (figure 4*a,b*). Deletion of residue 130 of HA1 of A/Beijing/262/95-like viruses, one of 10 characteristic amino acid differences between the viruses of the two lineages, appeared to contribute significantly to the antigenic differences. Sixteen amino acids distinguished the NAs of these viruses. Subsequent (after 1995) divergence of the NAs of viruses of the two lineages has been less marked than for the HAs. The slow increase in prevalence, which is over 5 years, of the A/Beijing/262/95-like viruses to become the predominant viruses causing influenza in 2000–2001 contrasts with the more rapid changes in the predominance of successive AH3N2 variants.

Co-circulation of antigenically distinct variants of the same subtype for extended periods of time complicates recommendations for vaccine composition. However, clinical studies have shown that vaccines containing A/Beijing/262/95 antigens induce antibodies that cross-react with both recent antigenic variants (World Health Organization 1998). Co-circulation of H1N1 and H3N2 subtypes provides the opportunity for the emergence of subtypes with different combinations of HA and NA. Shortly after the introduction of the H1N1 viruses in 1977, H1N1 reassortant viruses possessing M and NS genes of the H1N1 virus and PB1, PB2, PA and NP genes of the contemporary H3N2 viruses were isolated in several countries (Young & Palese 1979). However, the only novel subtype shown to circulate was a H1N2 virus, that derived seven genes from an H3N2 virus and circulated in China during 1988 and 1989 (Guo *et al.* 1992).

4. INFLUENZA B VIRUSES

The evolution of influenza B viruses is characterized by co-circulation of antigenically and genetically distinct

Table 2. *Antigenic drift of influenza AH1N1 viruses: haemagglutination inhibition titres.*

(Homologous titres are marked in bold. $<$ = <40.)

virus	post-infection ferret sera								
	A/USSR 90/77	A/Brazil 11/78	A/Chile 1/83	A/Sing 6/86	A/Taiw 1/86	A/Tex 36/91	A/Bay 7/95	A/Beij 262/96	A/NC 20/99
A/USSR/90/77	**1280**	640	40	$<$	$<$	$<$	$<$	$<$	$<$
A/Brazil/11/78	320	**1280**	80	$<$	$<$	$<$	$<$	$<$	$<$
A/Chile/1/83	80	80	**320**	$<$	$<$	$<$	$<$	$<$	$<$
A/Singapore/6/86	$<$	$<$	$<$	**1280**	320	1280	1280	$<$	$<$
A/Taiwan/1/86	$<$	$<$	$<$	640	**640**	1280	1280	40	$<$
A/Texas/36/91	$<$	$<$	$<$	1280	640	**2560**	1280	40	40
A/Bayern/7/95	$<$	$<$	$<$	640	640	2560	**2560**	80	40
A/Beijing/262/96	$<$	$<$	$<$	$<$	$<$	40	40	**640**	320
A/New Caledonia/20/99	$<$	$<$	$<$	$<$	$<$	$<$	$<$	160	**640**

Figure 4. Phylogenetic relationships of the (*a*) HA and (*b*) NA genes of influenza A H1N1 viruses. The sequences of nucleotides 6–978 encoding the HA1 domain and nucleotides 1–1384 encoding the NA were analysed as described in figure 2.
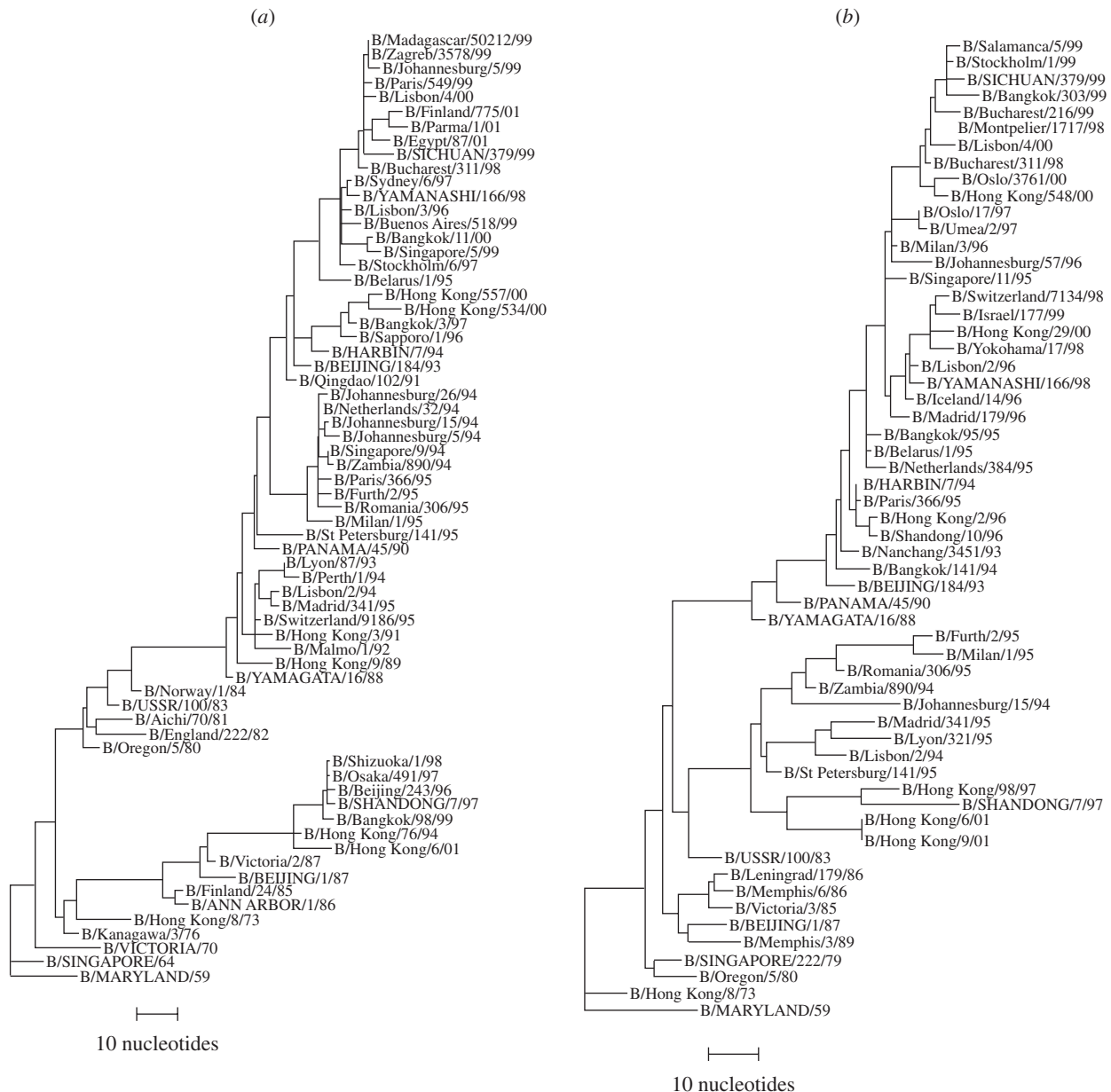
(a)

(b)

Figure 5. Phylogenetic relationships of the (a) HAs and (b) NAs of influenza B viruses. The sequences of nucleotides 1–987 encoding the HA1 domain and nucleotides 1–1439 encoding the NA were analysed as described in figure 2. The Yamagata lineage is drawn above the Victoria lineage.

lineages for extended periods of time (Yamashita *et al.* 1988). Two lineages that are defined by phylogenetic relationships of HA genes diverged in the early to mid-1970s, one represented by B/Victoria/2/87, the 'B/Victoria lineage' and the other represented by the antigenic variant B/Yamagata/16/88, which emerged in 1988, the 'B/Yamagata lineage' (figure 5 and table 3) (Kanegae *et al.* 1990; Rota *et al.* 1990). Viruses of these two lineages have predominanted at different times, as indicated by recommendations for inclusion in influenza vaccines (figures 1 and 5) and have differed in their geographical distribution. Since 1991, viruses of the B/Victoria lineage have been isolated infrequently in Africa, America and Europe, but have continued to circulate in Asia and in some years have been the predominant B virus in certain countries. Co-circulation of the antigenically distinct

lineages of B viruses, their fluctuating prevalence and differences in their geographical distribution have complicated global recommendations for vaccine composition. For example, in order to reflect the predominance of B/Shandong/7/97-like viruses (B/Victoria lineage) in Japan and Thailand in 1998–1999, the WHO recommended use in these areas of vaccine containing a B/Shandong/7/97-like virus rather than a B/Beijing/184/93-like virus (B/Yamagata lineage) that was representative of the viruses prevalent in other parts of the world (figure 1 and table 3) (World Health Organization 1999). More recently, the emergence of an antigenic variant of the B/Victoria lineage, as represented by B/Hong Kong/6/01 and its co-circulation with B/Sichuan/379/99-like (B/Yamagata lineage) viruses has emphasized the persistence of viruses belonging to the two lineages over some 25 years.

Table 3. *Antigenic drift of influenza B viruses: haemagglutination inhibition titres.*

(Homologous titres are marked in bold. $<$ = < 40.)

	post-infection ferret sera								
virus	B/HK 8/73	B/Sing 222/79	B/USSR 100/83	B/Vic 2/87	B/Sha 7/97	B/Yam 16/88	B/Pan 45/90	B/Beij 184/93	B/Sich 379/99
B/Hong Kong/8/73	**640**	160	80	80	$<$	80	40	40	$<$
B/Singapore/222/79	160	**320**	160	40	$<$	80	40	$<$	$<$
B/USSR/100/83	80	80	**320**	80	40	40	40	$<$	$<$
B/Victoria/2/87	40	$<$	40	**320**	160	40	40	$<$	$<$
B/Shangdong/7/97	$<$	$<$	$<$	160	**640**	40	$<$	$<$	$<$
B/Yamagata/16/88	40	40	40	40	$<$	**1280**	80	40	$<$
B/Panama/45/90	80	40	40	$<$	$<$	160	**160**	160	$<$
B/Beijing/184/93	$<$	$<$	$<$	$<$	$<$	$<$	40	**320**	320
B/Sichuan/379/99	$<$	$<$	$<$	$<$	$<$	$<$	$<$	40	**640**

Estimates of the rate of evolution have generally indicated that B viruses have evolved at a rate two to three times lower than those of A viruses. Analyses of the data illustrated in figure 5 gave values of *ca.* 2×10^{-3} nucleotide substitutions per site per year for the accumulation of mutations in HA and NA genes of viruses of the Yamagata lineage, that compare with values of *ca.* 4×10^{-3} for the HA and NA genes of A H3N2 viruses. The reason for these differences, that may be related to the periodic introduction of novel A viruses (antigenic shift) and the greater diversity of influenza A viruses, is not understood.

As might be expected, genetic reassortment between viruses of the co-circulating lineages is also a characteristic feature of the evolution of B viruses. For example, viruses prevalent in various parts of the world, including Africa, America and Europe between 1992 and 1996, such as B/Johannesburg/15/94, possessed an HA gene of the B/Yamagata lineage and an NA gene that is closely related to those of viruses of the B/Victoria lineage (figure 5) (McCullers *et al.* 1999). Frequent reassortment among other genes, e.g. NP, M and NS, encoding internal components of the viruses contributes to variability and emphasizes the potential important role that genetic reassortment plays in the evolution of co-circulating lineages (Lindstrom *et al.* 1999).

5. THE EMERGENCE OF NOVEL SUBTYPES

The major role of genetic reassortment in the evolution of human influenza viruses is in facilitating the emergence of novel A subtypes that cause pandemics. The 1957 AH2N2 viruses resulted from the acquisition of three genes from an avian source encoding the HA, NA and the polymerase component PB1 by a preceding AH1N1 virus and, in turn, the subsequent acquisition of novel avian H3 HA and PB1 components gave rise to the Hong Kong AH3N2 pandemic strain (Kawaoka *et al.* 1989). The identity of the host species in which the genetic reassortment occurred is not known. Pigs have been considered a potential intermediate host in view of their susceptibility to infection by both avian and human viruses, the emergence in pig populations of human–avian reassortants and the frequency of sporadic swine to

human transmissions (Scholtissek *et al.* 1985). This can be explained by the broader specificity of receptors for influenza viruses on cells lining the pig trachea, which include oligosaccharides with sialic acid linked $\alpha 2,3$ or $\alpha 2,6$ to the penultimate galactose (Ito *et al.* 1998), that accommodates the different specificities of avian viruses for the former and human viruses for the latter receptors, that is an important determinant of host range restriction (Connor *et al.* 1994). In addition to facilitating avian–human virus reassortment, the pig trachea also provides an environment for the adaptation of avian HAs to acquire human $\alpha 2,6$ specificity, as observed following the introduction of an avian H1N1 virus into pigs in Europe in the late 1970s (Ito *et al.* 1998).

The importance of pigs as a potential intermediate host in the emergence of novel human viruses and for genetic reassortment between avian and human viruses has been reinforced by the genetic compositions of recently emerging swine viruses (Brown *et al.* 1998; Zhou *et al.* 1999). The human–avian reassortant H3N2 viruses that have circulated in European pigs since 1984 resulted from reassortment between two swine viruses, that derived the HA and NA from a H3N2 virus related to viruses circulating in the human population in the early 1970s and the six internal genes from an 'avian-like' H1N1 virus introduced into pigs in 1979 (Castrucci *et al.* 1993). Cases of human infection by an 'avian-like' H1N1 swine virus in The Netherlands in 1986 and by H3N2 reassortant viruses in The Netherlands in 1993 and in Hong Kong in 1999 indicate the potential for swine to human transmission (Claas *et al.* 1994; Gregory *et al.* 2001). The resistance of these swine viruses to amantadine also emphasizes a potential role in the emergence of drug-resistant human influenza viruses (Gregory *et al.* 2001). The ability of avian internal genes to support efficient replication in human cells was also demonstrated by the severe infections by avian AH5N1 subtype viruses in Hong Kong in 1997; however, efficient person to person transmission did not ensue. The similarity between the six internal genes of the chicken H5N1 viruses and the quail H9N2 viruses, that caused human infections in 1999, (related by genetic reassortment) suggests that certain features of this genetic complement facilitate avian to human transmission (Lin *et al.* 2000). Thus,

domestic poultry also have a potential role in facilitating the emergence of a pandemic strain. A major concern during the latter events was the potential for co-infection by circulating human viruses directly giving rise to a human reassortant pandemic strain. However, the nature of the next pandemic strain, as for the timing and route of its emergence, remains an enigma.

REFERENCES

Bender, C., Hall, H., Huang, J., Klimov, A., Cox, N., Hay, A., Gregory, V., Cameron, K., Lim, W. & Subbarao, K. 1999 Characterization of the surface proteins of influenza A (H5N1) viruses isolated from humans in 1997–1998. *Virology* **254**, 115–123.

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.

Bush, R. M., Bender, C. A., Subbarao, K., Cox, N. J. & Fitch, W. M. 1999 Predicting the evolution of human influenza A. *Science* **286**, 1921–1925.

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.

Claas, E. C. J., 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. J., Osterhaus, A. D. M. E., Van Beek, R., De Jong, J. C., Rimmelzwaan, G. F., Senne, D. A., Krauss, S., Shortridge, K. F. & Webster, R. G. 1998 Human influenza AH5N1 virus related to a highly pathogenic avian influenza virus. *Lancet* **351**, 472–477.

Colman, P. M., Varghese, J. N. & Laver, W. G. 1983 Structure of the catalytic and antigenic sites in influenza virus neuraminidase. *Nature* **303**, 41–44.

Connor, R. J., Kawaoka, Y., Webster, R. G. & Paulson, J. C. 1994 Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates. *Virology* **205**, 17–23.

Fleury, D., Barrere, B., Bizebard, T., Daniels, R. S., Skehel, J. J. & Knossow, M. 1999 A complex of influenza hemagglutinin with a neutralizing antibody that binds outside the virus receptor binding site. *Nat. Struct. Biol.* **6**, 530–534.

Gregory, V., Lim, W., Cameron, K., Bennett, M., Marozin, S., Klimov, A., Hall, H., Cox, N., Hay, A. & Lin, Y. P. 2001 Infection of a child in Hong Kong by an influenza A H3N2 virus closely related to viruses circulating in European pigs. *J. Gen. Virol.* **82**, 1397–1406.

Guo, Y., Xu, X. & Cox, N. J. 1992 Human influenza A (H1N2) viruses isolated from China. *J. Gen. Virol.* **73**, 383–388.

Ito, T. (and 11 others) 1998 Molecular basis for the generation in pigs of influenza A viruses with pandemic potential. *J. Virol.* **72**, 7367–7373.

Kanegae, Y., Sugita, S., Endo, A., Ishida, M., Senya, S., Osaka, K., Nerome, K. & Oya. A. 1990 Evolutionary pattern of the hemagglutinin gene of influenza B viruses isolated in Japan: cocirculating lineages in the same epidemic season. *J. Virol.* **64**, 2860–2865.

Kawaoka, Y., Krauss, S. & Webster, R. G. 1989 Avian-to-human transmission of the PB1 gene of influenza A viruses in 1957 and 1968 pandemics. *J. Virol.* **63**, 4603–4608.

Kendal, A. P., Noble, G. R., Skehel, J. J. & Dowdle, W. R. 1978 Antigenic similarity of influenza A (H1N1) viruses from epidemics in 1977–1978 to 'Scandinavian' strains isolated in epidemics of 1950–1951. *Virology* **89**, 632–636.

Lin, Y. P. (and 12 others) 2000 Avian-to-human transmission of H9N2 subtype influenza A viruses: relationship between H9N2 and H5N1 human isolates. *Proc. Natl Acad. Sci. USA* **97**, 9654–9658.

Lindstrom, S. E., Hiromoto, Y., Nerome, R., Omoe, K., Sugita, S., Yamazaki, Y., Takahashi, T. & Nerome, K. 1998 Phylogenetic analysis of the entire genome of influenza A (H3N2) viruses from Japan: evidence for genetic reassortment of the six internal genes. *J. Virol.* **72**, 8021–8031.

Lindstrom, S. E., Hiromoto, Y., Nishimura, H., Saito, T., Nerome, R. & Nerome, K. 1999 Comparative analysis of evolutionary mechanisms of the hemagglutinin and three internal protein genes of influenza B virus: multiple cocirculating lineages and frequent reassortment of the NP, M, and NS genes. *J. Virol.* **73**, 4413–4426.

McCullers, J. A., Wang, G. C., He, S. & Webster, R. G. 1999 Reassortment and insertion–deletion are strategies for the evolution of influenza B viruses in nature. *J. Virol.* **73**, 7343–7348.

Raymond, F. L., Caton, A. J., Cox, N. J., Kendal, A. P. & Brownlee, G. G. 1986 The antigenicity and evolution of influenza H1 haemagglutinin, from 1950–1957 and 1977–1983: two pathways from one gene. *Virology* **148**, 275–287.

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. USA* **96**, 1651–1656.

Rogers, G. N., Paulson, J. C., Daniels, R. S., Skehel, J. J., Wilson, I. A. & Wiley, D. C. 1983 Single amino acid substitutions in influenza haemagglutinin change receptor binding specificity. *Nature* **304**, 76–78.

Rota, P. A., Wallis, T. R., Harmon, M. W., Rota, J. S., Kendal, A. P. & Nerome, K. 1990 Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983. *Virology* **175**, 59–68.

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., Burger, H., Kistner, O. & Shortridge, K. F. 1985 The nucleoprotein as a possible major factor in determining host specificity of influenza H3N2 viruses. *Virology* **147**, 287–294.

Skehel, J. J. & Wiley, D. C. 2000 Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *A. Rev. Biochem.* **69**, 531–569.

Skehel, J. J., Stevens, D. J., Daniels, R. S., Douglas, A. R., Knossow, M., Wilson, I. A. & Wiley, D. C. 1984 A carbohydrate side chain on hemagglutinins of Hong Kong influenza viruses inhibits recognition by a monoclonal antibody. *Proc. Natl Acad. Sci. USA* **81**, 1779–1783.

Smith, W., Andrews, C. H. & Laidlaw, P. P. 1933 A virus obtained from influenza patients. *Lancet* ii, 66–88.

Swofford, D. L. & Olsen, G. J. 1990 Phylogeny reconstruction. In *Molecular systematics* (ed D. M. Hilis & C. Moritz), pp. 411–501. Sunderland, MA: Sinauer Associates.

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.

Webby, R. J. & Webster, R. G. 2001 Emergence of influenza A viruses. *Phil. Trans. R. Soc. Lond.* B **356**, 1815–1826.

Webster, R. G., Yakhno, M., Hinshaw, V. S., Bean, W. J. & Murti, K. G. 1978 Intestinal influenza: replication and characterization of influenza viruses in ducks. *Virology* **84**, 268–278.

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.

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.

World Health Organization. 1996 WHO influenza surveillance. *WHO Wkly Epidemiol. Rec.* **71**, 353–356.

World Health Organization. 1998 Recommended composition of influenza virus vaccines for use in the 1998–1999 season. *WHO Wkly Epidemiol. Rec.* **73**, 56–60.

World Health Organization 1999 Recommended composition of influenza virus vaccines for use in the 1999–2000 season. *WHO Wkly Epidemiol. Rec.* **74**, 57–61.

Xu, X., Cox, N. J., Bender, C. A., Regnery, H. L. & Shaw, M. W. 1996 Genetic variation in neuraminidase genes of influenza A (H3N2) viruses. *Virology* **224**, 175–183.

Yamashita, M., Krystal, M., Fitch, W. M. & Palese, P. 1988 Influenza B virus evolution: co-circulating lineages and comparison of evolutionary pattern with those of influenza A and C viruses. *Virology* **163**, 112–122.

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. USA* **76**, 6547–6551.

Zhou, N. N., Senne, D. A., Landgraf, J. S., Swenson, S. L., Erickson, G., Rossow, K., Liu, L., Yoon, K. J., Krauss, S. & Webster, R. G. 1999 Genetic reassortment of avian, swine, and human influenza A viruses in American pigs. *J. Virol.* **73**, 8851–8856.