

Influenza A viruses in feral Canadian ducks: extensive reassortment in nature

Todd F. Hatchette,¹ David Walker,¹ Christie Johnson,¹ Ashley Baker,¹ S. Paul Pryor³ and Robert G. Webster^{1,2}

Correspondence

Robert G. Webster

Robert.Webster@stjude.org

¹Division of Virology, Department of Infectious Diseases, St Jude Children's Research Hospital, 332 North Lauderdale Street, Memphis, TN 38105-2794, USA

²Department of Pathology, University of Tennessee Health Science Center, Memphis, TN, USA

³Canadian Wildlife Service, Environment Canada, Edmonton, Alberta, Canada

The current dogma of influenza accepts that feral aquatic birds are the reservoir for influenza A viruses. Although the genomic information of human influenza A viruses is increasing, little of this type of data is available for viruses circulating in feral waterfowl. This study presents the genetic characterization of 35 viruses isolated from wild Canadian ducks from 1983 to 2000, as the first attempt at a comprehensive genotypic analysis of influenza viruses isolated from feral ducks. This study demonstrates that influenza virus genes circulating in Canadian ducks have achieved evolutionary stasis. The majority of these duck virus genes are clustered in distinct North American clades; however, some H6 and H9 genes are clustered with those from Eurasian viruses. Genes appeared to reassort in a random fashion. None of the genotypes identified remained present throughout all of the years examined and most PA and PB2 genes that crossed over into swine were clustered in one phylogenetic grouping. Additionally, matrix genes were identified that branch very early in the evolutionary tree. These findings demonstrate the diversity of the influenza virus gene pool in Canadian ducks, and suggest that genes which cluster in specific phylogenetic groupings in the PB2 and PA genes can be used for markers of viruses with the potential for crossing the species barrier. A more comprehensive study of this important reservoir is needed to provide further insight into the genomic composition of viruses that crossover the species barrier, which would be a useful component to pandemic planning.

Received 4 December 2003

Accepted 19 April 2004

INTRODUCTION

Aquatic birds are the reservoirs of influenza A viruses (Horimoto & Kawaoka, 2001; Webster *et al.*, 1992). The relationship between the virus and aquatic birds has progressed to the point of 'evolutionary stasis', in which the viral genes have achieved maximal fitness in their natural avian host as compared with other species (Gorman *et al.*, 1990, 1992; Ito *et al.*, 1991).

Genes from this pool have crossed over the species barrier into human and other domestic avian and swine species, either as complete avian viruses or as genetically reassorted viruses following interspecies transmission. Phylogenetic studies have shown that the human pandemic strains of 1957 and 1968 were derived from reassortant viruses with genes of human and avian origin (Kawaoka *et al.*, 1989; Schafer *et al.*, 1993; Scholtissek *et al.*, 1978).

Swine have been proposed as the 'mixing vessel' in which

these pandemic influenza viruses arise (Scholtissek & Naylor, 1988). Recently, H3N2 viruses have established a stable lineage in USA swine. In these animals, two reassortants were identified, a 'double reassortant', between the classical swine H1N1 and human H3N2 viruses and a 'triple reassortant', which included two avian internal protein genes (PA and PB2) in addition to those seen in the double reassortant (Zhou *et al.*, 1999a). The triple reassortant has become the dominant circulating strain in USA swine and continues to evolve. Subsequent reassortant events with the co-circulating classical H1N1 swine viruses have generated novel H1N2 viruses. Two of these H1N2 isolates have a third avian internal protein gene (PB1), which is closely related to an avian H5N2 virus (Choi *et al.*, 2002; Karasin *et al.*, 2002). In Europe, avian-like H1N1 viruses have replaced the classical H1N1 in European swine and multiple reassortant events similar to those seen in the USA have occurred in pigs in Great Britain (Brown *et al.*, 1997, 1998; Scholtissek *et al.*, 1983). These findings suggest that although influenza viruses continue to evolve in swine, the conservation of these avian internal protein genes confer a selective advantage.

The GenBank accession numbers for the sequences used in the phylogenetic analysis described in this study are AY633116–AY633395.

Although avian influenza viruses can infect other mammals, it was thought that direct transmission to humans in nature was unlikely (reviewed by Horimoto & Kawaoka, 2001). However in 1997, a highly pathogenic H5N1 avian influenza virus was directly transmitted to humans from poultry in Hong Kong, suggesting that adaptation in an intermediate mammalian host was not necessary (Claas *et al.*, 1998; De Jong *et al.*, 1997; Subbarao *et al.*, 1998).

Despite the depopulation of the bird markets and institution of preventative measures, the H5N1 subtype re-emerged in 2000, 2001, 2002 and 2003 with multiple genotypes (Brammer *et al.*, 2003; Guan *et al.*, 2002a, b; Webster *et al.*, 2002). Although a number of recent studies have examined phylogenetic data from virus samples collected in the bird markets of Hong Kong and southern China, little data are available regarding the gene pool in feral waterfowl (Guan *et al.*, 2000, 2002a, b; Hoffmann *et al.*, 2000; Lin *et al.*, 1994; Liu *et al.*, 2003; Webster *et al.*, 2002; Zhou *et al.*, 1999b).

In the current study, we provide the first genotypic analysis of this vast gene pool by detailing phylogenetic data from 35 viruses collected from Canadian ducks sampled in the Alberta wilderness during a 17-year period. From this data, we sought to enhance the current database and answer the following questions: do these duck isolates support the current dogma that influenza in waterfowl has achieved 'evolutionary stasis' and that genes can be grouped into separate North American and Eurasian avian clades? Do these gene segments reassort within this pool? Are there

specific haemagglutinin (HA) subtypes that have a greater propensity for reassortment? Does this feral waterfowl gene pool contribute to infection in domestic species and, if so, which genotypes have a propensity to crossover into other species?

METHODS

Virus isolates. The 35 viruses used in this study were chosen from the St Jude Children's Research Hospital repository (Table 1). The viruses reflected the predominate influenza A subtypes that circulate in the Alberta duck gene pool (H3, H4 and H6) and included subtypes which have been found in humans and swine (H1, H2 and H3) and in domestic avian species (H6 and H9). Subtypes rarely isolated in this duck gene pool (H5, H7, H8, H11 and H12) were not included (Sharp *et al.*, 1993). Although the viruses examined spanned a 17-year period, the majority of viruses were isolated between 1991 and 2000. The viruses were isolated from cloacal swabs collected from ducks in Alberta Canada, as part of a longitudinal study to determine the diversity of the influenza gene pool in this duck population and investigate whether the isolates were antigenically related to human strains (Hinshaw *et al.*, 1980). Sampling techniques are described elsewhere (Hinshaw *et al.*, 1979). The viruses were revived from -70°C storage and grown in 10-day-old embryonated chicken eggs. Most of the viruses were passaged twice in eggs prior to being used in this study.

Antigenic analysis. HA subtyping was determined by haemagglutination inhibition (HI) assay performed in 96-well microtitre plates using reference sera that were treated with receptor-destroying enzyme (Palmer *et al.*, 1975). Neuraminidase (NA) subtype was determined by the NA inhibition assay using reference sera, as described elsewhere (Aymard-Henry *et al.*, 1973).

Table 1. Feral Canadian duck influenza A virus isolates used in the present study

Mallard, *Anas platyrhynchos*; Northern pintail, *Anas acuta*; redhead, *Aythya americana*; blue-winged teal, *Anas discors*.

Virus	Subtype	Virus	Subtype
A/mallard/Alberta/211/98	H1N1	A/pintail/Alberta/207/99	H4N8
		A/pintail/Alberta/210/99	H4N6
A/mallard/Alberta/202/96	H2N5	A/mallard/Alberta/119/2000	H4N6
A/teal/Alberta/16/97	H2N9	A/mallard/Alberta/136/2000	H4N6
A/pintail/Alberta/22/97	H2N9		
A/mallard/Alberta/205/98	H2N3	A/mallard/Alberta/98/85	H6N2
A/mallard/Alberta/226/98	H2N3	A/pintail/Alberta/113/85	H6N2
		A/mallard/Alberta/203/92	H6N5
A/mallard/Alberta/117/97	H3N8	A/pintail/Alberta/179/93	H6N1
A/pintail/Alberta/156/97	H3N8	A/mallard/Alberta/76/94	H6N8
A/mallard/Alberta/242/98	H3N8	A/pintail/Alberta/155/94	H6N8
A/mallard/Alberta/279/98	H3N8	A/mallard/Alberta/232/94	H6N8
A/pintail/Alberta/37/99	H3N8	A/redhead/Alberta/291/94	H6N8
A/mallard/Alberta/199/99	H3N6	A/mallard/Alberta/206/96	H6N8
A/mallard/Alberta/127/2000	H3N8	A/mallard/Alberta/215/99	H6N8
A/mallard/Alberta/47/98	H4N1	A/mallard/Alberta/743/83	H9N1
A/mallard/Alberta/30/98	H4N6	A/mallard/Alberta/321/88	H9N2
A/mallard/Alberta/295/98	H4N6	A/mallard/Alberta/17/91	H9N2
A/mallard/Alberta/111/99	H4N6	A/mallard/Alberta/11/91	H9N2

RNA isolation, RT-PCR and sequencing. Viral RNA was isolated from allantoic fluid using the Qiagen RNeasy kit (Qiagen) according to the manufacturer's protocol. Reverse transcription of viral RNA and subsequent PCR was performed using primers specific for each viral gene, as described previously (Hoffmann *et al.*, 2001). PCR products were gel purified and cleaned using the QIAquick PCR purification kit (Qiagen) according to the manufacturer's protocol. PCR amplicons were sequenced by the Hartwell Center for Bioinformatics and Biotechnology at St Jude Children's Research Hospital using rhodamine or dRhodamine dye-terminator cycle sequence-ready reaction kits with AmpliTaq DNA polymerase FS (Perkin-Elmer Applied Biosystems; PE/ABI) and synthetic oligonucleotides. Electrophoresis and analysis of the samples were done using the PE/ABI model 373 or model 377 DNA sequencers.

Phylogenetic studies. DNA sequences were aligned by the Lasergene sequence analysis software package (DNASTAR). Multiple sequences were aligned using CLUSTAL X (v 1.81; Thompson *et al.*, 1997), and phylogenetic trees were generated using the neighbour-joining and parsimony algorithms and bootstrap analysis in PAUP (v 4.0b10; D. Swofford, Sinauer Associates). Phylogenetic trees were generated based on the alignment of the complete or partial nucleotide sequences of each viral gene. The regions of each internal gene used in analysis included: PB2, nt position 65–603; PB1, 354–565; PA, 47–525; NP, 41–631; M, 40–668 and NS, 101–621. The regions used in the analysis of the HA and NA genes and trees not illustrated in this paper can be viewed at <http://www.stjuderesearch.org/data/feralduck>.

Generation of phylogenetic groupings. Related viruses were assigned phylogenetic groupings on the basis of their position in the trees, which were generated by neighbour-joining and maximum-parsimony methods and the following criteria. (i) Phylogenetic distinctions generated by the trees must be supported by high bootstrap scores (>95%) at the nodal branches (Nei & Kumar, 2000). (ii) When the bootstrap scores at the nodal branches are less than this arbitrary cut-off value, the sequence similarities of the gene segments for each virus in the proposed genotype group must have a 95% or greater sequence similarity with other members of the group and less than 94.9% identity with the gene segments of viruses belonging to other phylogenetic groups. (iii) In cases where the trees predict groupings in which some viral gene segments within a group have less than 95% similarities (e.g. 94.5%), those segments may be included in that group, as long as the sequence identity is higher with members of the proposed grouping than it is with members from other groupings. (iv) If the position of the virus isolate is in the same phylogenetic grouping in trees generated by two different methods (neighbour-joining and maximum-parsimony), but there is discrepancy between groupings based on sequence similarity, the genotype group of the virus isolate will be dictated by where it is placed in the calculated phylogenetic trees. (v) If the virus isolate is not in the same position in both trees generated by the two different methods and there is discrepancy between sequence identities, then a definitive grouping cannot be assigned.

RESULTS

Phylogenetic data

All of the influenza virus genes isolated from Canadian feral ducks, except the H9 and some H6, are clustered within the North American avian clade and could be subgrouped into two to six different clusters depending on the gene. As an example of how the subgroupings in the trees were established, the phylogenetic tree for the PA gene is

presented in Fig. 1 and will be outlined in detail. Because of space constraints, most of the trees will not be presented here, but are available at <http://www.stjuderesearch.org/data/feralduck>. However, notable features of each tree and the number of subgroupings used in the development of genotypes are outlined in Table 2.

Genes in Canadian feral duck viruses are in evolutionary stasis

Evolutionary stasis is evident by the small degree of change seen in internal viral genes isolated from different viruses many years apart. For example, the NP genes of A/mallard/Alberta/743/83 (H9N1) and A/mallard/Alberta/111/99 (H4N6) had a nucleotide sequence identity of 95.1%, with the predicted proteins having a 99.6% identity at the amino acid level. Similarly, two isolates with the greatest difference in nucleotide sequence similarity of their NP genes (92.2%) only had 0.5% difference in amino acids of the predicted protein fragment. The high degree of amino acid identity in the M1 and M2, NP and NS1 proteins of the feral duck viruses also supports this dogma. Similar observations are evident in the surface genes. Two H6 viruses spanning a 7-year period (A/mallard/Alberta/203/92 and A/mallard/Alberta/215/99) share sequence and amino acid similarity of 97.6 and 98.1%, respectively. Most impressive is the high degree of sequence similarity between A/mallard/Alberta/211/98 (H1N1) and A/Brant goose/Alaska/1/1917 (H1N?). Despite the fact that these two isolates circulated in feral waterfowl 81 years apart they share a 95.3% sequence identity over the gene fragment available for comparison.

Identification of matrix genes that form a distinct early branch in the phylogenetic tree

Although matrix genes are clustered into three main groups within the North American avian clade, two genetically identical viruses isolated from two different species of ducks during the same year, A/mallard/Alberta/98/85 (H6N2) and A/pintail/Alberta/113/85 (H6N2) form a distinct branch early in the phylogenetic tree. To establish better their phylogenetic relationship, these two 'outlier' matrix genes were completely sequenced and compared with the complete sequences from 219 matrix genes compiled from viruses sequenced in this study and from the influenza database (Macken *et al.*, 2001). The resulting tree suggests that these isolates form their own distinct branch with sequence similarities less than 92.6%, compared with that of the other 219 sequences. Bootstrap analysis using the neighbour-joining method rooted to A/equine/Prague/56 (H7N7), places these two outliers at the base of the North American avian clade (Fig. 2); bootstrap analysis using the maximum-parsimony method places the outliers just off the root of the tree (see website for tree). However, the numeric value was low (i.e. 34 of 100 trees) with either method.

Phylogenetic analysis of the remaining seven genes did not

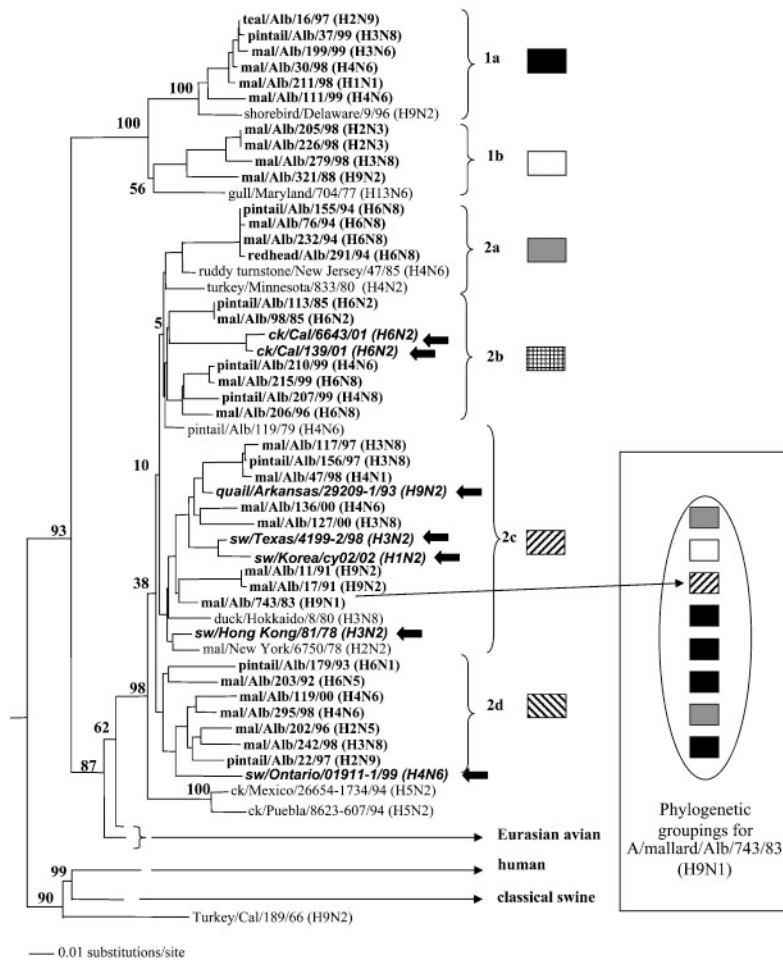


Fig. 1. Phylogenetic tree of the PA gene (nucleotide bases 47–525) analysed with PAUP using bootstrap and the neighbour-joining methods as distance measures. The tree is rooted to A/equine/London/1416/73 (H7N7). The distance bar is shown under the tree and bootstrap values (100 replicates) are given for selected nodes. The Canadian feral duck isolates sequenced in this study cluster in the North American avian clade and are in bold print. Black arrows identify swine, quail and chicken viruses that clustered with Canadian duck viruses; each of the six phylogenetic groupings (1a, b; 2a–d) is assigned a different pattern. To generate the genotype for a particular virus, the patterns representing the phylogenetic groupings for each gene can then be compiled to give a visual representation of the genotype. The genotype of A/mallard/Alberta/743/83 (H9N1) is presented in a box to the right of the tree as an example of this process. The genes illustrated from top to bottom are: PB2, PB1, PA, HA, NP, NA, M and NS. Abbreviations: Alb, Alberta; Cal, California; ck, chicken; mal, mallard; sw, swine.

reveal any unique lineages. All of the genes fell into the North American avian clade closely associated with other feral North American duck viruses.

Although most of the predicted proteins of the Canadian feral duck matrix genes clustered closely together in the North American avian clade, with >99% amino acid identity, the phylogenetic analysis of the M1 and M2 sequences and their predicted proteins of A/mallard/Alberta/98/85 (H6N2) and A/pintail/Alberta/113/85 (H6N2) confirm the unique placement of these outlier genes. The similarity between the predicted M1 protein and an avian consensus sequence generated using sequence data from 73 North American duck and shorebird viruses obtained from the database (Macken *et al.*, 2001), and sequences provided by L. Widjaja (personal communication), was 97.6%; that of the predicted M2 protein was 93.5%. The M1 protein of these outliers had 5 amino acid changes throughout the length of the protein (I15V, S116A, S142V, N224S and N232D). Similarly, the predicted sequence of the M2 protein of the outliers differed from the consensus sequence by 6 amino acids distributed throughout the cytoplasmic, transmembrane and surface portions of the protein (S13N, I27V, F28I, Y50C, K70E and Q77R).

Clustering of PB2 and PA genes from swine and quail viruses within the same phylogenetic grouping

One of the most interesting results from the phylogenetic analysis is the clustering of the PB2 and PA genes from quail and swine isolates with feral Canadian duck viruses in the same phylogenetic grouping (Fig. 1). To illustrate this, the phylogenetic analysis of the PA gene is presented below.

The PA genes clustered into two main groups, with isolates having ≤88.5% nucleotide similarities between the two groups. The two groups were further subdivided as follows: the first group, which included 10 isolates related to influenza A viruses from shorebirds and gulls, divided into two distinct phylogenetic clusters (1a and b). The second group comprised the remaining 25 viruses, which clustered within the North American avian clade. The second group was divided into four phylogenetic sub-groupings (2a–d). Key features of this tree include: (i) the clustering of one shorebird virus [A/ruddy turnstone/New Jersey/47/85 (H4N6)] and two Eurasian viruses from the influenza database [i.e. A/duck/Hokkaido/8/80 (H3N8) and A/swine/Hong Kong/81/78 (H3N2)] with the feral duck genes in the North American avian clade; (ii) the placement

Table 2. Summary of the phylogenetic analysis of the internal protein genes

Gene segment	Phylogenetic subgroups	Notable features of the phylogenetic tree
PB2	5	Two North American swine isolates, A/Swine/Texas/4199-2/98 (H3N2) and A/Swine/Ontario/01911-1/99 (H4N6), and one Eurasian isolate, A/swine/korea/cy02/02 (H1N2) with PB2 genes derived from avian sources all clustered in the same subgroup (subgroup 4).
PB1	2	Two swine viruses with avian PB1 genes, A/Swine/Ontario/01911-1/99 (H4N6) and A/swine/Kansas/13481-T/00 (H1N2) were placed into different phylogenetic groupings (group 2 and 1, respectively).
PA	6	Many 'intermediate host'* isolates cluster in group 2c and share close similarity with A/mallard/Alberta/743/83 (H9N1) (Fig. 1).
NP	2	Two isolates with the greatest difference in sequence similarity (91.7%), A/mallard/Alberta/30/98 (H4N6) and A/mallard/Alberta/210/99 (H4N6), share an amino acid similarity of 99.5%. Two viruses collected 16-years apart, A/mallard/Alberta/743/83 (H9N1) and A/mallard/Alberta/111/99 (H4N6), shared a nucleotide sequence and predicted protein identity of 95.1 and 99.6%, respectively.
M	4	Two isolates from feral Canadian duck viruses were placed on a unique, distinct branch near base of the tree (Fig. 2).
NS	4	Both alleles present which have $\leq 70.6\%$ similarity (2 phylogenetic groupings per allele). The predicted NS1 protein identity of the A allele was 95.2–99.6%, and that of the B allele was 97–99.6%. Similarity between the two alleles was as much as 70.1% at the amino acid level. No deletions or substitutions (E92D) associated with NS genes in virulent H5N1 viruses were noted.

*Both swine and quail are thought to act as intermediate hosts that facilitate infection across species barriers (Perez *et al.*, 2003; Scholtissek & Naylor, 1988). Isolates clustering in group 2c included A/quail/Arkansas/29209-1/93 (H9N2), A/swine/HongKong/81/78 (H3N2), A/swine/Texas/4199-2/98 (H3N2) and A/swine/Korea/cy02/02 (H1N2).

of a number of swine and quail viruses [i.e. A/swine/Texas/4199-2/98 (H3N2) (sw/TX/98), A/swine/HongKong/81/78 (H3N2) (sw/HK/78) and A/quail/Arkansas/29209-1/93 (H9N2) (qu/AR/93)] within the same phylogenetic subgrouping (2c), which share high nucleotide sequence similarity with the PA gene from one particular duck virus A/mallard/Alberta/743/83 (H9N1) (96.7, 98.3 and 98.5% sequence similarity, respectively). The other North American swine virus with an avian PA gene, A/swine/Ontario/01911-1/99 (H4N6) (sw/ONT/99) belongs to phylogenetic group 2d and is most closely related to A/mallard/Alberta/202/96 (H2N5) and A/pintail/Alberta/22/97 (H2N9) (95.4%) but also shares a 95% sequence identity with A/mallard/Alberta/743/83.

Placement of three H6 and four H9 genes in the Eurasian lineage

The phylogenetic analyses of the genes encoding the surface glycoproteins circulating in the feral Canadian ducks have similar features, such as high degree of nucleotide and amino acid identities between genes from viruses isolated many years apart. Interestingly, three of eight H6 and all four of the H9 genes in these North American duck viruses clustered with Eurasian isolates, whereas the remaining NA and HA genes were all placed in one or two groupings within the North American avian clade. The phylogenetic analysis of the H6 gene is presented (Fig. 3) in greater detail to illustrate this unique feature. The phylogenetic tree for H9 and other HA and NA genes can be viewed at www.stjuderesearch.org/data/feralduck.

In this study, three of eight H6 genes fell within the Eurasian clade (Fig. 3). One isolate, A/pintail/Alberta/179/93 (H6N1), shared a 95.5% sequence similarity and 95.3% amino acid similarity with A/chicken/California/139/01 (H6N2) (ck/CA/01), which represents the current H6N2 viruses circulating in California. The other two isolates, A/mallard/Alberta/203/92 (H6N5) and A/mallard/Alberta/215/99 (H6N8) formed a separate cluster with A/mallard/Alberta/206/96 (H6N8) in the Eurasian clade. The other five feral duck H6 isolates form two distinct branches within the North American avian clade whose nucleotide sequences differ by less than 3.3%. A/pintail/Alberta/113/85 (H6N2) and A/mallard/Alberta/232/94 (H6N8), and whose circulation spanned a 9-year period, have a 96.9% nucleotide and 98.1% amino acid similarity.

A high degree of similarity of genes from influenza A isolated from domestic species and viruses isolated from feral ducks

Feral aquatic birds in North America have been the reservoir of some or all of the genes in influenza viruses that have crossed over into domestic species in the USA, including sw/TX/98 (H3N2), sw/ONT/99 (H4N6) and ck/CA/01 (H6N2) (Karasin *et al.*, 2000; Webby *et al.*, 2002; Zhou *et al.*, 1999a). The nucleotide sequences of the internal genes of these viruses have a higher degree of similarity with genes in feral Canadian duck viruses than they have with previously published sequences (Table 3). Although A/swine/Korea/cy02/02 (H1N2) also has PB2 and PA genes that clustered with North American duck isolates, all of the

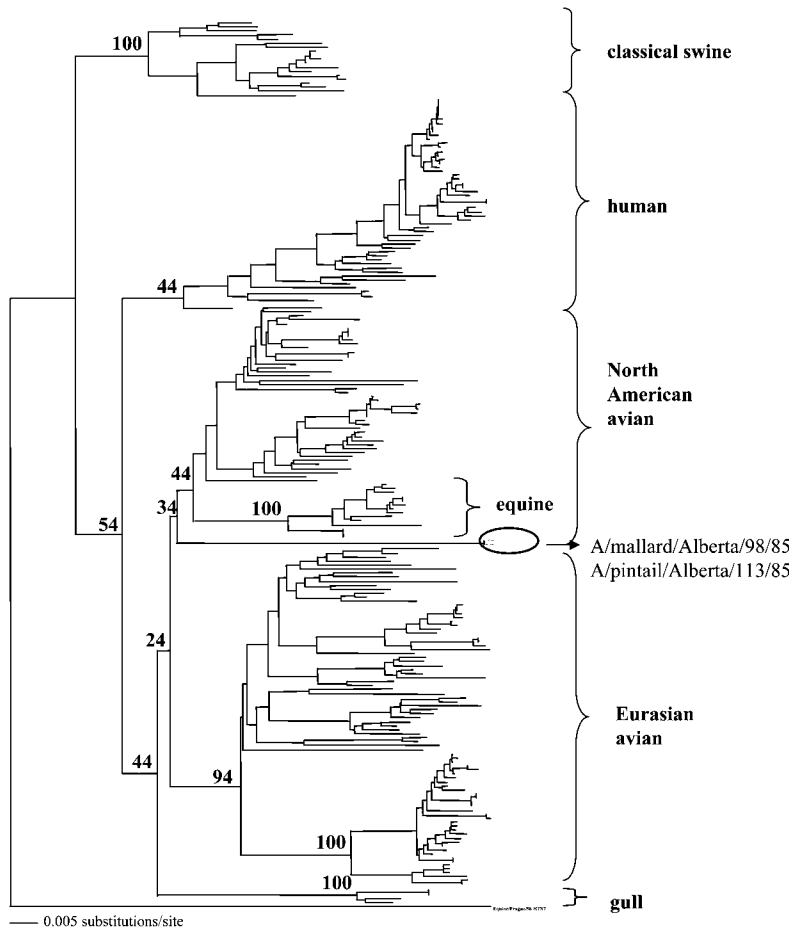


Fig. 2. Phylogenetic tree of the matrix gene obtained by using the complete sequences of 221 M genes (nucleotides bases 26–1004) analysed with PAUP using the bootstrap and the neighbour-joining methods as distance measures. The tree is rooted to A/equine/Prague/56 (H7N7). The distance bar is shown under the tree and bootstrap values (100 replicates) are given for selected nodes. Major phylogenetic divisions are indicated to the right of the tree and the two Canadian feral duck outliers (A/pintail/Alberta/113/85 and A/mallard/Alberta/98/85) are identified at the base of the North American avian clade.

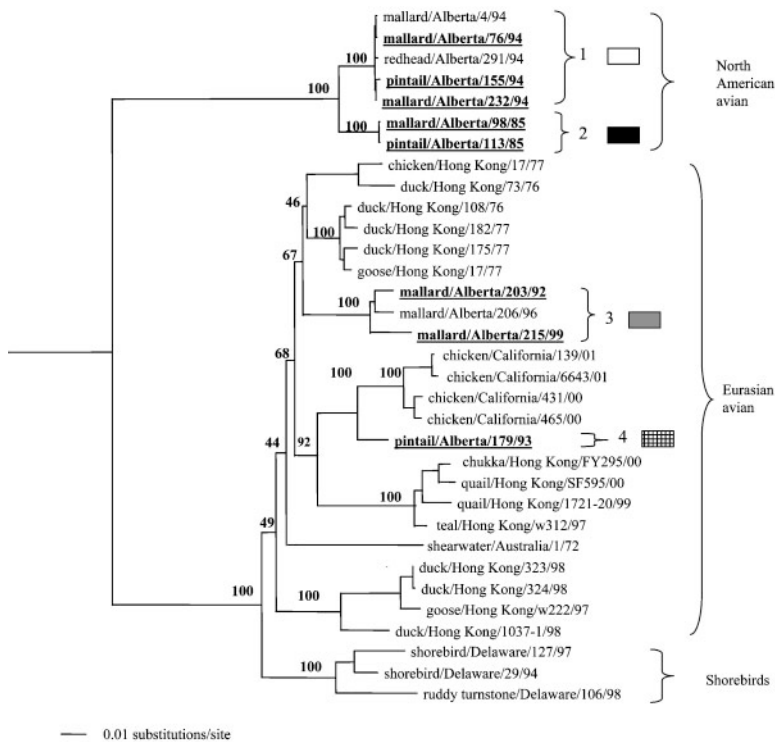


Fig. 3. Phylogenetic tree of the H6 subtype of the HA gene (nucleotides bases 67–1701) analysed with PAUP by using the bootstrap and the neighbour-joining methods as distance measures. The tree is rooted to A/mallard/Pennsylvania/10218/84 (H5N2). The distance bar is shown under the tree and bootstrap values (100 replicates) are given for selected nodes. Major phylogenetic divisions are indicated to the right of the tree. The feral Canadian duck viruses sequenced in this study are identified in bold type and underlined. Each of the four distinct phylogenetic groupings is assigned a different pattern that is used in generating genotypes, which are illustrated to the right of each of the viruses.

Table 3. Nucleotide sequence similarity of influenza viruses from domestic species with viruses isolated from feral ducks in Alberta

Gene	A/swine/Texas/4199-2/98 (H3N2) (sw/tx/98)	A/swine/Ontario/01911-1/99 (H4N6) (sw/ont/99)	A/chicken/California/139/01 (H6N2) (ck/cal/01)
PB2	97 % A/mallard/Alberta/321/88 (H9N2) [Phylogenetic subgroup 4]	98 % A/mallard/Alberta/211/98 (H3N8) [Phylogenetic subgroup 4]	96·7 % A/mallard/Alberta/11/91 (H9N2) [Phylogenetic subgroup 1]
PB1	NA	98 % A/pintail/Alberta/156/97 (H3N8) [Phylogenetic subgroup 2]	98·5 % A/mallard/Alberta/30/98 (H4N6) [Phylogenetic subgroup 1]
PA	96·7 % A/mallard/Alberta/743/83 (H9N1) [Phylogenetic subgroup 2c]	95·4 % A/pintail/Alberta/22/97 (H9N1) A/mallard/Alberta/202/96 (H2N5) [Phylogenetic subgroup 2d]	94·6 % A/mallard/Alberta/743/83 (H9N1) A/mallard/Alberta/98/85 (H6N2) A/pintail/Alberta/113/85 (H6N2) [Phylogenetic subgroup 2b]
HA	NA	95·7 % A/mallard/Alberta/295/98 (H4N6)	95·5 % A/pintail/Alberta/179/93 (H6N4)
NP	NA	96·4 % A/mallard/Alberta/76/94 (H6N8) A/mallard/Alberta/232/92 (H6N8) [Phylogenetic subgroup 1]	96·3 % A/mallard/Alberta/17/91 (H9N2) [Phylogenetic subgroup 1]
NA	NA	96·2 % A/mallard/Alberta/119/00 (H4N6)	91·6 % A/mallard/Alberta/11/91 (H9N2)
M	NA	98·6 % A/mallard/Alberta/206/96 (H6N8) [Phylogenetic subgroup 2]	97·8 % A/pintail/Alberta/202/96 (H2N5) [Phylogenetic subgroup 1]
NS	NA	98·7 % A/mallard/Alberta/203/92 (H6N5) [Phylogenetic subgroup 1]	98·3 % A/mallard/Alberta/205/98 (H2N3) [Phylogenetic subgroup 1]

NA, Not applicable.

genes in this virus share close similarity with currently circulating H1N2 swine viruses in North America and is the result of a swine virus having been imported to Korea from North America (Y. K. Choi, personal communication).

Genotypic data

Virus genotypes were assigned by compiling the phylogenetic groupings of each gene segment together for each virus isolate. A visual representation of the process was created by allocating distinct phylogenetic groupings a different pattern representing a different grouping for each gene (Figs 1 and 4). A dominant genotype did not persist within the majority of HA subtypes, but there were examples of a single predominate genotype in viruses of the H2 and H6 subtypes in 1998 and 1994, respectively (Fig. 4). However, there was no clear evolution of these 'dominant' genotypes from year to year. For example, although the four H6 viruses isolated in 1994 had identical genotypes (suggesting a dominant genotype), the isolate from the previous year only had the NP, M and NS genes in common with the dominant genotype.

Although the majority of influenza A genes appeared to reassort in a random manner, there are examples of some

HA subtypes conserving a phylogenetic subgrouping for a particular gene. Eight PB2 genes of the H4 viruses (isolated 1998–2000) and four PB1 and NP genes of the H9 viruses (1983–1991) belonged to the same phylogenetic cluster. Similarly, the NS gene of the H2 viruses (1996–1998) and matrix genes of the H3 viruses (1997–2000) showed similar associations with HA subtype. Despite these associations, the remaining genes were diverse (Fig. 4). The only example of limited diversity was seen in the four H9 viruses. Although A/mallard/Alberta/321/88 had a unique genotype compared with the others, the remaining three viruses maintained a high degree of similarity over an 8-year period, suggesting limited diversity within this subtype. With the exception of PB2, which clusters in a separate group, all of the internal genes of the viruses A/mallard/Alberta/743/83 (H9N1), A/mallard/Alberta/17/91 (H9N2) and A/mallard/Alberta/11/91 (H9N2) grouped in the same phylogenetic clusters with nucleotide sequence similarities ranging from 96·2 to 97·5 % (Fig. 4).

DISCUSSION

It is the accepted dogma that feral waterfowl and shorebirds are the reservoir for all subtypes of influenza A viruses

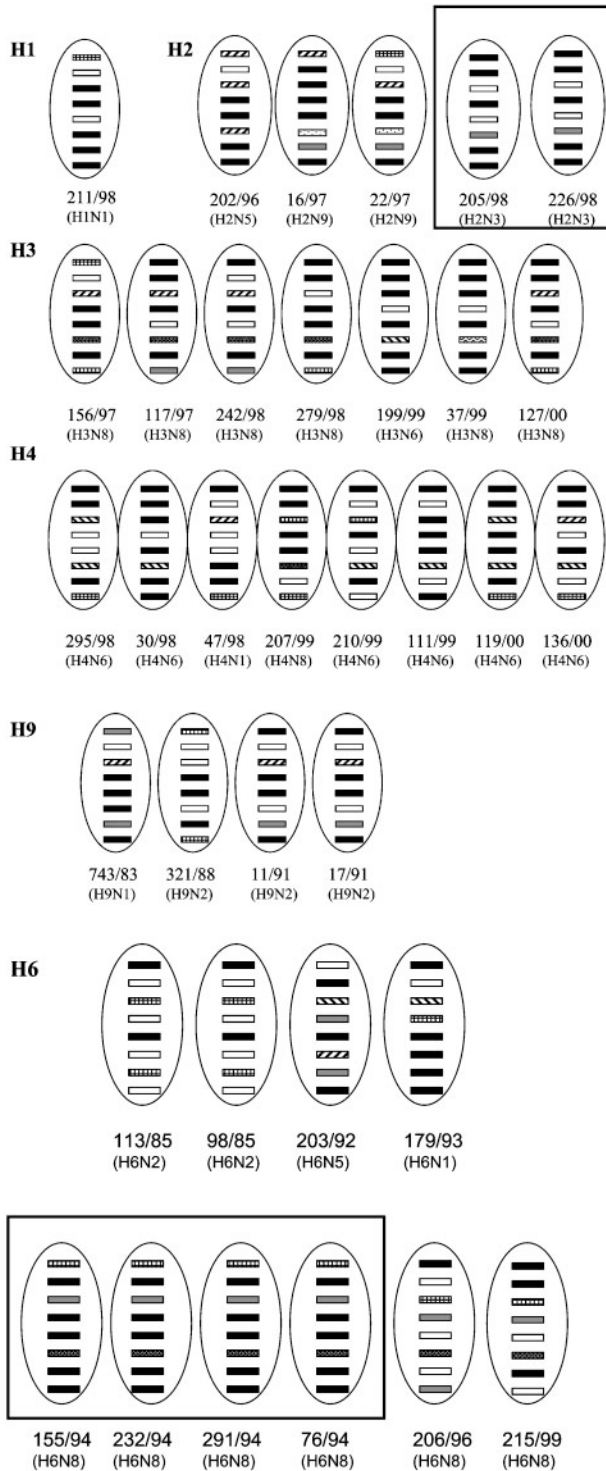


Fig. 4. Graphic representation of the genotypes identified by using the phylogenetic groupings derived from the phylogenetic trees of each gene. Different phylogenetic groupings are assigned different patterns. Viruses are identified by their isolation number and the year that they were collected. The genes in each virus are (from top to bottom) PB2, PB1, PA, HA, NP, NA, M and NS. Dominant genotypes within an HA subtype in a given year are enclosed in boxes.

(Webster *et al.*, 1992). Little phylogenetic information is available regarding the gene pool in North American or Eurasian wild aquatic birds. In this study, we provide the first attempt to systematically examine the avian gene pool circulating in feral North American ducks.

Consistent with accepted dogma, genes from influenza viruses isolated from Canadian feral ducks are in evolutionary stasis and the majority of the genes clustered with other avian species in the North American clade. However in contrast to this, the H9 and some of the H6 HA genes in this study clustered with viruses in the Eurasian avian clade.

Although phylogenetic studies have demonstrated that avian influenza viruses have evolved into two separate lineages (i.e. the North American and Eurasian clades) (Ito *et al.*, 1991, 1995; Kawaoka *et al.*, 1988; Lin *et al.*, 1994; Okazaki *et al.*, 2000; Suarez & Perdue, 1998; Webster *et al.*, 1992), ‘geographical intermixing’ has been described in the H2 subtype in shorebirds in North America. This may be because of the intermixing of birds from different continents in the arctic and subarctic regions and rare trans-Atlantic migration of some of these birds, particularly ruddy turnstones, whose migratory patterns are not fully elucidated (Gorman *et al.*, 1990; Del Hoyo *et al.*, 1992; Makarova *et al.*, 1999). This hypothesis is further supported by the results of this study, which place the PB2 gene of *A/ruddy turnstone/New Jersey/47/85* (H4N6) with Eurasian isolates. Despite evidence that there may be a distinct gene pool in shorebirds (Kawaoka *et al.*, 1988), the clustering of duck PA and PB2 genes with those of shorebirds and gulls in this study suggests an intermixing of genes between the two pools. Alternatively, the possible existence of two major phylogenetic subgroups that do not correlate with the geographical origin of the virus isolates, and therefore do not fit the current classification system, could explain these data. Although convergent evolution has also been suggested as an alternative hypothesis (Webby *et al.*, 2002), this is less likely given that the virus has achieved evolutionary stasis in these avian hosts.

Genes within the feral Canadian duck gene pool reassort independently. Although dominant genotypes may be present in any given year, there was no clear evolution of these dominant genotypes from year to year and the majority of the HA subtypes examined were equally diverse. However, some subtypes had conserved a phylogenetic grouping for a particular gene such as the NS gene in H2, the M gene in H3 and the PB2 gene in H4 viruses. The significance of this finding is not clear and may be related to the limited number of virus isolates examined over a short-time frame. However, the remaining genes in these viruses had multiple phylogenetic groupings, a finding that suggests multiple genotypes that reflect the diverse nature of the gene pool. This contrasts previous findings that suggest although co-infection with two or more influenza A viruses can occur in ducks, it does not occur randomly; thus, the number and types of possible reassortants are limited (Sharp *et al.*, 1997).

The H9 subtype was the only example of restricted diversity in a particular HA subtype in this study; the genotypic pattern of three out of four H9 viruses maintained a high degree of similarity from 1983 to 1991. These data differ from previous studies, which showed that the H9N2 subtypes circulating in domestic species within the avian markets of southeast China had a greater diversity and showed more evidence of reassortment than did other virus subtypes in the gene pool, suggesting that the H9 subtype has a greater propensity for reassortment (Liu *et al.*, 2003). Alternatively, because there is greater host variation in these domestic species, the increased selection pressures may be driving the increase in reassortment seen within the domestic population.

Why the H9 subtypes in the current study did not reassort is unclear but may be related to the infrequency with which this subtype is isolated in the North American duck gene pool. Sharp *et al.* (1993) found that the predominant subtypes that were isolated consistently in feral Alberta ducks included H3, H4 and H6; H9 subtypes were isolated in small, infrequent clusters and would 'disappear' from the gene pool for periods as long as 6 years. In contrast, the H9 subtype is more common in shorebirds and gulls, and this gene pool may 'spillover' into the duck gene pool as these birds migrate north in the spring (Kawaoka *et al.*, 1988). However, if there is limited circulation of this subtype, the lack of diversity over an 8-year period suggests that the source of periodic reintroduction of H9 is either reassorting at a slow rate or in a state of 'frozen evolution', similar to what has been suggested for equine strains isolated from South America in 1987 and from India in 1988 (Endo *et al.*, 1992; Lindstrom *et al.*, 1998). These speculations are based on a limited sample size and need to be interpreted with caution. More genomic data of H9 viruses circulating in Canadian ducks is needed to explore these observations further.

Viruses sequenced in this study were clearly related to viruses that have been isolated from domestic species. By examining each of the eight genes in the feral duck viruses, we can determine better whether there are common genotypes or phylogenetic groupings that crossover the species barrier more frequently. We have shown that the PA genes of swine and quail species and the PB2 genes of swine species that have North American avian related genes clustered together in the same phylogenetic grouping.

Both quail and swine are intermediate hosts, and swine, in particular, are thought to be the mixing vessel from which pandemic influenza arises (Perez *et al.*, 2003; Scholtissek & Naylor, 1988). Currently, the dominant influenza virus strains circulating in Eurasian and North American pigs have various genes encoding internal proteins of avian origin (Brown *et al.*, 1997; Zhou *et al.*, 1999a). This predominance suggests that the genes of the avian viruses impart a selective advantage. Lin *et al.* (1994) suggested that influenza genes in the Eurasian avian gene pool may 'possess unique sequences' that permit transmission

and infection of other hosts. The present study indicates that perhaps this constellation of polymerase genes is optimal for a polymerase complex that allows enough replicative error to optimize species adaptation. Although chickens have been implicated in the transmission of avian influenza viruses to humans, their role as an intermediate host has not been as defined as it has been for quail and swine species. Clustering of viruses isolated from chickens within these phylogenetic groupings would further support this, however, this was not the case (Fig. 1).

Interestingly, although the 1957 and 1968 pandemics were associated with human avian reassortants that contained an avian PB1 gene (Kawaoka *et al.*, 1989; Schafer *et al.*, 1993; Scholtissek *et al.*, 1978), genes from the two swine viruses that contain avian PB1 protein genes [swine/Kansas/13481-T/00 (H1N2) and sw/ONT/99 (H4N6)] did not cluster together in this analysis. However, qu/AR/93 (H9N2), ch/CA/01 (H6N2) and swine/Kansas/13481-T/00 (H1N2) did cluster in the same phylogenetic grouping.

Although there is no direct evidence and the dataset is small, one could speculate that North American avian viruses that possess polymerase protein genes which fall into these genotypic groupings might have a greater propensity to crossover the species barrier and adapt in an intermediate host (to view PB1 and PB2 trees see supplemental figures at www.stjuderesearch.org/data/feralduck).

Another unique finding in this study is the identification of a distinct lineage of matrix genes in feral ducks. These outlier genes are different from other matrix genes circulating in the Canadian duck gene pool, and although definitive placement in the phylogenetic tree is not possible, our analyses suggest that the viruses that contain these outlier genes diverge from other avian isolates at the base of the North American avian clade or earlier. The consequence of amino acid changes in the predicted proteins of the M1 and M2 genes of the outliers (compared with the avian consensus) is not clear. The five differing residues in the M1 protein do not lie within the nuclear targeting domain or within sites of the protein that influence the morphology of the virus (Bourmakina & Garcia-Sastre, 2003; Ye *et al.*, 1995). Although the M2 gene in the outlying duck viruses have an I27V substitution, this amino acid change is not known to confer amantadine resistance (Grambas *et al.*, 1992; Hay *et al.*, 1985).

It is not clear, why this unique lineage of matrix genes in other feral Canadian ducks was not identified. The virus isolate belonged to the H6 subtype that was frequently isolated and had multiple genotypes in different years, a finding that suggests it reassorted freely. Because this study included only 1% of the viruses in our repository, other related viruses probably have yet to be identified. Alternatively, the matrix genes of the outlying viruses may have conferred a survival disadvantage and were subsequently subjected to negative selection. However, this hypothesis is unlikely to be correct; because of the

comparatively low sequence identity between the outliers and other avian matrix genes suggests that the outliers' genes diverged from the other North American avian viral genes long ago and have been maintained.

This study is the first attempt at a systematic review of the influenza genes circulating in its natural host. The main criticism of genotype analysis as a method of analysing the patterns of gene reassortment is that it depends on the strictness of the criteria for grouping viruses. In this study, criteria were established for generation of phylogenetic groupings in an attempt to accurately reflect the relationship of one virus to another. Although arbitrary, this method and the time period chosen provide some indication of how diverse the gene pool is and how genes can reassort over a 17-year period. However, only a few of the genes were completely sequenced and the relatively small number of viruses chosen during this time period make it impossible to assess the year-to-year changes in the gene pool. In addition, sampling bias because of a small sample size cannot be excluded. A more comprehensive and complete genotypic analysis is needed to examine some of the concepts arising from this study, to gain a broader understanding of the depth of the influenza gene pool and to determine how influenza A viruses reassort in aquatic birds. Expansion of the database and alignment of the complete gene sequences or implementation of a set region of the gene segment for analysis would facilitate further 'genomic mining' to help provide a more accurate reflection of the diversity of this gene pool.

ACKNOWLEDGEMENTS

This work was supported by Public Health Service grant AI-95357 from the National Institute of Allergy and Infectious Diseases, by Cancer Center Support CA-21765 from the National Cancer Institute, by the Lebanese Syrian Associated Charities (ALSAC) and funding from the University Internal Medicine Research Foundation, Department of Medicine, Dalhousie University and the Department of Pathology and Laboratory Medicine, Queen Elizabeth II Health Science Center, Halifax, Nova Scotia. We gratefully acknowledge Richard J. Webby, Scott Krauss and Diane Hulse for helpful discussions, advice and technical assistance.

REFERENCES

- Aymard-Henry, M., Coleman, M. T., Dowdle, W. R., Laver, W. G., Schild, G. C. & Webster, R. G. (1973). Influenza virus neuraminidase and neuraminidase-inhibition test procedures. *Bull WHO* **48**, 199–202.
- Bourmakina, S. V. & Garcia-Sastre, A. (2003). Reverse genetics studies on the filamentous morphology of influenza A virus. *J Gen Virol* **84**, 517–527.
- Brammer, L., Postrma, A., Harper, S., Klimov, A. & Cox, N. (2003). Update: influenza activity—United States, 2002–2003 season. *Morb Mortal Wkly Rep* **52**, 224–225.
- 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.
- Choi, Y. K., Goyal, S. M., Farnham, M. W. & Joo, H. S. (2002). Phylogenetic analysis of H1N2 isolates of influenza A virus from pigs in the United States. *Virus Res* **87**, 173–179.
- 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.
- De Jong, J. C., Claas, E. C., Osterhaus, A. D., Webster, R. G. & Lim, W. L. (1997). A pandemic warning? *Nature* **389**, 554.
- Del Hoyo, J., Elliot, A. & Sargatel, J. (1992). *Handbook of the Birds of the World*, vol. 1 & 3. Barcelona: Lynx Edicions.
- Endo, A., Pecoraro, R., Sugita, S. & Nerome, K. (1992). Evolutionary pattern of the H3 haemagglutinin of equine influenza viruses: multiple evolutionary lineages and frozen replication. *Arch Virol* **123**, 73–87.
- Gorman, O. T., Donis, R. O., Kawaoka, Y. & Webster, R. G. (1990). Evolution of influenza A virus PB2 genes: implications for evolution of the ribonucleoprotein complex and origin of human influenza A virus. *J Virol* **64**, 4893–4902.
- Gorman, O. T., Bean, W. J. & Webster, R. G. (1992). Evolutionary processes in influenza viruses: divergence, rapid evolution, and stasis. *Curr Top Microbiol Immunol* **176**, 75–97.
- Grambas, S., Bennett, M. S. & Hay, A. J. (1992). Influence of amantadine resistance mutations on the pH regulatory function of the M2 protein of influenza A viruses. *Virology* **191**, 541–549.
- Guan, Y., Shortridge, K. F., Krauss, S., Chin, P. S., Dyrting, K. C., Ellis, T. M., Webster, R. G. & Peiris, M. (2000). H9N2 influenza viruses possessing H5N1-like internal genomes continue to circulate in poultry in southeastern China. *J Virol* **74**, 9372–9380.
- Guan, Y., Peiris, J. S., Lipatov, A. S., Ellis, T. M., Dyrting, K. C., Krauss, S., Zhang, L. J., Webster, R. G. & Shortridge, K. F. (2002a). Emergence of multiple genotypes of H5N1 avian influenza viruses in Hong Kong SAR. *Proc Natl Acad Sci U S A* **99**, 8950–8955.
- Guan, Y., Peiris, M., Kong, K. F., Dyrting, K. C., Ellis, T. M., Sit, T., Zhang, L. J. & Shortridge, K. F. (2002b). H5N1 influenza viruses isolated from geese in Southeastern China: evidence for genetic reassortment and interspecies transmission to ducks. *Virology* **292**, 16–23.
- Hay, A. J., Wolstenholme, A. J., Skehel, J. J. & Smith, M. H. (1985). The molecular basis of the specific anti-influenza action of amantadine. *EMBO J* **4**, 3021–3024.
- Hinshaw, V. S., Webster, R. G. & Turner, B. (1979). Water-borne transmission of influenza A viruses? *Intervirology* **11**, 66–68.
- Hinshaw, V. S., Webster, R. G. & Turner, B. (1980). The perpetuation of orthomyxoviruses and paramyxoviruses in Canadian waterfowl. *Can J Microbiol* **26**, 622–629.
- Hoffmann, E., Stech, J., Leneva, I., Krauss, S., Scholtissek, C., Chin, P. S., Peiris, M., Shortridge, K. F. & Webster, R. G. (2000). Characterization of the influenza A virus gene pool in avian species in southern China: was H6N1 a derivative or a precursor of H5N1? *J Virol* **74**, 6309–6315.
- Hoffmann, E., Stech, J., Guan, Y., Webster, R. G. & Perez, D. R. (2001). Universal primer set for the full-length amplification of all influenza A viruses. *Arch Virol* **146**, 2275–2289.
- Horimoto, T. & Kawaoka, Y. (2001). Pandemic threat posed by avian influenza A viruses. *Clin Microbiol Rev* **14**, 129–149.
- Ito, T., Gorman, O. T., Kawaoka, Y., Bean, W. J. & Webster, R. G. (1991). Evolutionary analysis of the influenza A virus M gene with comparison of the M1 and M2 proteins. *J Virol* **65**, 5491–5498.

- Ito, T., Okazaki, K., Kawaoka, Y., Takada, A., Webster, R. G. & Kida, H. (1995). Perpetuation of influenza A viruses in Alaskan waterfowl reservoirs. *Arch Virol* **140**, 1163–1172.
- Karasin, A. I., Brown, I. H., Carman, S. & Olsen, C. W. (2000). Isolation and characterization of H4N6 avian influenza viruses from pigs with pneumonia in Canada. *J Virol* **74**, 9322–9327.
- Karasin, A. I., Landgraf, J., Swenson, S., Erickson, G., Goyal, S., Woodruff, M., Scherba, G., Anderson, G. & Olsen, C. W. (2002). Genetic characterization of H1N2 influenza A viruses isolated from pigs throughout the United States. *J Clin Microbiol* **40**, 1073–1079.
- Kawaoka, Y., Chambers, T. M., Sladen, W. L. & Webster, R. G. (1988). Is the gene pool of influenza viruses in shorebirds and gulls different from that in wild ducks? *Virology* **163**, 247–250.
- 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.
- Lin, Y. P., Shu, L. L., Wright, S., Bean, W. J., Sharp, G. B., Shortridge, K. F. & Webster, R. G. (1994). Analysis of the influenza virus gene pool of avian species from southern China. *Virology* **198**, 557–566.
- Lindstrom, S., Endo, A., Sugita, S., Pecoraro, M., Hiromoto, Y., Kamada, M., Takahashi, T. & Nerome, K. (1998). Phylogenetic analyses of the matrix and non-structural genes of equine influenza viruses. *Arch Virol* **143**, 1585–1598.
- Liu, M., He, S., Walker, D. & 7 other authors (2003). The influenza virus gene pool in a poultry market in South Central China. *Virology* **305**, 267–275.
- Macken, C., Lu, H., Goodman, J. & Boykin, L. (2001). The value of a database in surveillance and vaccine selection. In *Options for the Control of Influenza IV*, pp. 103–106. Edited by A. D. M. E. Osterhaus, N. Cox & A. W. Hampson. Amsterdam: Elsevier Science.
- Makarova, N. V., Kaverin, N. V., Krauss, S., Senne, D. & Webster, R. G. (1999). Transmission of Eurasian avian H2 influenza virus to shorebirds in North America. *J Gen Virol* **80**, 3167–3171.
- Nei, M. & Kumar, S. (2000). Accuracies and statistical tests of phylogenetic trees. In *Molecular Evolution and Phylogenetics*, pp. 165–186. Edited by M. Nei & S. Kumar. New York: Oxford University Press.
- Okazaki, K., Takada, A., Ito, T. & 13 other authors (2000). Precursor genes of future pandemic influenza viruses are perpetuated in ducks nesting in Siberia. *Arch Virol* **145**, 885–893.
- Palmer, D. F., Coleman, W. R., Dowdle, W. R. & Schild, G. C. (1975). *Advanced Laboratory Techniques for Influenza Diagnosis*. Washington, DC, USA: Department of Health, Education and Welfare.
- Perez, D. R., Lim, W., Seiler, J. P., Yi, G., Peiris, M., Shortridge, K. F. & Webster, R. G. (2003). Role of quail in the interspecies transmission of H9 influenza A viruses: molecular changes on HA that correspond to adaptation from ducks to chickens. *J Virol* **77**, 3148–3156.
- 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. & Naylor, E. (1988). Fish farming and influenza pandemics. *Nature* **331**, 215.
- Scholtissek, C., Rohde, W., Von Hoyningen, V. & Rott, R. (1978). On the origin of the human influenza virus subtypes H2N2 and H3N2. *Virology* **87**, 13–20.
- Scholtissek, C., Burger, H., Bachmann, P. A. & Hannoun, C. (1983). Genetic relatedness of hemagglutinins of the H1 subtype of influenza A viruses isolated from swine and birds. *Virology* **129**, 521–523.
- Sharp, G. B., Kawaoka, Y., Wright, S. M., Turner, B., Hinshaw, V. & Webster, R. G. (1993). Wild ducks are the reservoir for only a limited number of influenza A subtypes. *Epidemiol Infect* **110**, 161–176.
- Sharp, G. B., Kawaoka, Y., Jones, D. J., Bean, W. J., Pryor, S. P., Hinshaw, V. & Webster, R. G. (1997). Coinfection of wild ducks by influenza A viruses: distribution patterns and biological significance. *J Virol* **71**, 6128–6135.
- Suarez, D. L. & Perdue, M. L. (1998). Multiple alignment comparison of the non-structural genes of influenza A viruses. *Virus Res* **54**, 59–69.
- 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.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876–4882.
- Webby, R. J., Woolcock, P. R., Krauss, S. L. & Webster, R. G. (2002). Reassortment and interspecies transmission of North American H6N2 influenza viruses. *Virology* **295**, 44–53.
- 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., Guan, Y., Peiris, M. & 9 other authors (2002). Characterization of H5N1 influenza viruses that continue to circulate in geese in southeastern China. *J Virol* **76**, 118–126.
- Ye, Z., Robinson, D. & Wagner, R. R. (1995). Nucleus-targeting domain of the matrix protein (M1) of influenza virus. *J Virol* **69**, 1964–1970.
- Zhou, N. N., Senne, D. A., Landgraf, J. S. & 7 other authors (1999a). Genetic reassortment of avian, swine, and human influenza A viruses in American pigs. *J Virol* **73**, 8851–8856.
- Zhou, N. N., Shortridge, K. F., Claas, E. C., Krauss, S. L. & Webster, R. G. (1999b). Rapid evolution of H5N1 influenza viruses in chickens in Hong Kong. *J Virol* **73**, 3366–3374.