Available online at www.sciencedirect.com
SCIENCE DIRECT
Virus Research 114 (2005) 89–100
Virus Research
www.elsevier.com/locate/virusres

Phylogenetic analyses of type A influenza genes in natural reservoir species in North America reveals genetic variation

Erica Spackman [a,*], David E. Stallknecht [b], Richard D. Slemons [c], Kevin Winker [d], David L. Suarez [a], Melissa Scott [a], David E. Swayne [a]

[a] *Southeast Poultry Research Laboratory, USDA-ARS, 934 College Station Road, Athens, GA 30605, USA*
[b] *Southeastern Cooperative Wildlife Disease Study, College of Veterinary Medicine, University of Georgia, Athens, GA 30602, USA*
[c] *Department of Veterinary Preventative Medicine, The Ohio State University, 1900 Coffey Road, Columbus, OH 43210, USA*
[d] *University of Alaska Museum, 907 Yukon Drive, Fairbanks, AK 99775, USA*

Received 17 March 2005; received in revised form 24 May 2005; accepted 25 May 2005
Available online 21 July 2005

Summary

The matrix (M) (98 isolates), nucleoprotein (NP) (67 isolates), non-structural (NS) (107 isolates), H4 subtype HA1 (21 isolates) and H6 (38 isolates) subtype HA1 region of the hemagglutinin (HA) gene were sequenced from avian influenza viruses isolated from North American wild aquatic birds between 1969 and 2003. Sequences were compared phylogenetically with all available wild aquatic bird isolate sequences and sequences from isolates from gallinaceous poultry and mammalian host species. Among the wild aquatic bird isolates the NS subtype B NS1 and subtype B NS2 proteins were the most conserved with minimum amino acid identities of 98.0 and 99.0%, respectively. The H6 HA1, M and NS subtype A genes were more divergent; both nucleotide and amino acid divergence levels were similar to those of the same genes from gallinaceous poultry and mammalian origin type A influenza isolates. Phylogenetically none of the genes assorted geographically (within North America), chronologically or by species of origin, unlike isolates from non-natural host species, suggesting that selection pressure is low. A lack of clear temporal or spatial grouping shows that multiple lineages of virus co-circulate and precludes the development of an epidemiological map for influenza virus from these species.
Published by Elsevier B.V.

1. Introduction

The natural host species for type A influenza viruses are wild-waterfowl and shorebirds (Slemons et al., 1974) and all 16 hemagglutinin subtypes are found in these species, however the subtypes are not evenly distributed among species or locations. Viruses of the H6 and H4 hemagglutinin (HA) subtypes are two of the most common subtypes isolated from waterfowl in North America (Ito et al., 1995; Stallknecht et al., 1990). Both H4 and H6 viruses isolated from waterfowl have been shown to be able to infect humans by experimental inoculation (Beare and Webster, 1991) and both subtypes have caused disease outbreaks in commercial poultry in the US (Johnson et al., 1977; Woolcock et al., 2003).

Compared to the vast amount of genetic data available for type A influenza in non-natural host species, there is relatively little information from natural reservoir species for type A influenza. Genetic variability is expected to be relatively low based on a lack of selection pressure, therefore the virus has been proposed to be in evolutionary stasis (Gorman et al., 1992). However reports focusing on wild aquatic bird isolates are limited. Previous reports have focused on comparing sequences from wild aquatic bird sequences with isolates from outbreaks in poultry (Chin et al., 2002; Hoffmann et al., 2000; Okazaki et al., 2000; Webby et al., 2002). Other reports based on partial gene sequences have demonstrated that re-assortment occurs among duck isolates (Hatchette

* Corresponding author. Tel.: +1 706 546 3617; fax: +1 706 546 3161.
E-mail addresses: espackman@seprl.usda.gov (E. Spackman), dstall@vet.uga.edu (D.E. Stallknecht), slemons.1@osu.edu (R.D. Slemons), ffksw@uaf.edu (K. Winker), dsuarez@seprl.usda.gov (D.L. Suarez), mscott@seprl.usda.gov (M. Scott), dswayne@seprl.usda.gov (D.E. Swayne).

0168-1702/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.virusres.2005.05.013

et al., 2004) and that the North American duck and shorebird isolates may have sequences consistent with Eurasian gene lineage viruses (Hatchette et al., 2004; Makarova et al., 1999).

Generating influenza sequence data from natural host species from different years and different species is essential to understanding the basic biology, ecology and evolution of the virus in the natural host and will help to elucidate the molecular mechanisms of cross species transmission. Sequence data from wild waterfowl and shorebirds can also be used to develop an epidemiological map for avian influenza that can be used during surveillance and outbreaks. Finally, although shorebirds and waterfowl are considered the natural reservoir for influenza, it is not known whether there are differences in the viruses among duck and shorebird species.

In this report sequence data from avian influenza viruses (AIV) collected from free-flying waterfowl and shorebirds in North America between 1969 and 2003 are reported. The HA1 region of the HA genes of 38 H6 and 21 H4 isolates and the full coding regions of three other gene segments, the matrix (M) genes from 98 isolates, the non-structural (NS) genes from 107 isolates and the nucleoprotein (NP) genes from 67 isolates, were sequenced and analyzed with other wild aquatic bird isolates and representative isolates from poultry and mammalian hosts.

2. Materials and methods

2.1. Viruses

Avian influenza viruses were obtained from the Southeast Poultry Research Laboratory (SEPRL) repository or collected during a multi-institutional cooperative wild bird virus surveillance project between University of Georgia Southeastern Cooperative Wildlife Disease Study, the University of Alaska Museum, the Ohio State University Department of Veterinary Preventive Medicine and USDA-SEPRL. Viruses were obtained from cloacal swab material collected from captured free-flying birds. Avian influenza was isolated by standard virus isolation methods in embryonating chicken eggs (ECE) (Swayne et al., 1998). Chorioallantoic fluid from the infected embryos was used as the virus source. The isolates collected and included in this study are listed in Table 1.

2.2. Sequencing

The H6 and H4 HA genes were chosen because these subtypes appear to be common in North American wild waterfowl. Additionally, numerous H6 outbreaks have been reported in commercial poultry world wide (Chin et al., 2002; Woolcock et al., 2003) which has increased the need for H6 HA sequence data. Also, previously reported sequence data on the H4 HA subtype is limited and needs to be expanded.

Table 1
Avian influenza isolates collected from North American free-flying birds and their genes included this report

Isolate	HA subtype	Gene segment			
		M	NS	HA	NP
Black duck/MD/M037/98	H6	X[a]	A[b]	X	X
Blue-winged teal/LA/3B/88	H4	X	A		X
Blue-winged teal/LA/B182/88	H4	X	A	X	X
Blue-winged teal/TX/11/01	H7	X	A		
Blue-winged teal/LA/166B/88	H4	X	A	X	X
Blue-winged teal/LA/240B/88	H4	X	B	X	X
Blue-winged teal/LA/69B/87	H4	X		X	X
Blue-winged teal/LA/B156/88	H4	X		X	
Duck/NJ/5832-13/96	H6	X		X	X
Duck/NY/14933/95	H6		A		
Duck/PA/69	H6	X	A	X	
Duck/NJ/5406-27/94	H4			X	
Guinea fowl/NJ/14190-23/96	H4	X		X	X
Green-winged teal/AK/5937/01	H3	X			
Green-winged teal/TX/7/01	H8	X	A		
Green-winged teal/AK/281/98	H4	X	A		
Green-winged teal/AK/832/00	H6			X	
Green-winged teal/OH/196/99	H6	X	B		X
Green-winged teal/OH/72/99	H6	X	A		X
Laughing gull/NY/2455/00	H7	X	B		
Laughing gull/NY/470/00	H6	X	A	X	X
Mallad/OH/115/89	H6	X	B	X	X
Mallard/AK/3211/02	H3	X	B		X
Mallard/MD/F123/98	H6	X	A	X	X
Mallard/MD/R274/98	H6	X	B	X	X
Mallard/MD/R326/98	H6	X	B	X	X
Mallard/MN/107/98	H6	X	A	X	X
Mallard/MN/133/98	H5	X	A		
Mallard/MN/145/99	H4	X	A	X	X
Mallard/MN/153/98	H9		A		
Mallard/MN/154/98	H6	X	A	X	X
Mallard/MN/154/99	H4	X			X
Mallard/MN/157/98	H6	X	A	X	X
Mallard/MN/17/99	H7	X	A		X
Mallard/MN/175/99	H6	X	A	X	X
Mallard/MN/182/98	H3	X	A		
Mallard/MN/186/99	H9	X	A		
Mallard/MN/187/98	H7	X	A		
Mallard/MN/190/99	H3	X	B		
Mallard/MN/192/99	H4	X	A	X	X
Mallard/MN/195/99	H4	X	A	X	X
Mallard/MN/198/99	H4	X	A	X	X
Mallard/MN/2/98	H4	X	A	X	X
Mallard/MN/212/99	H4	X	B	X	X
Mallard/MN/218/99	H3	X	B		
Mallard/MN/220/98	H4	X	A		
Mallard/MN/225/99	H3	X	B		
Mallard/MN/231/99	H3	X			
Mallard/MN/232/98	H9	X	A		
Mallard/MN/253/99	H6	X	A	X	X
Mallard/MN/26/98	H2		A		
Mallard/MN/263/99	H4	X	B	X	X
Mallard/MN/27/98	H2		A		
Mallard/MN/280/99	H3	X	A		
Mallard/MN/281/98	H9	X	A		
Mallard/MN/281/99	H3		A		
Mallard/MN/283/99	H3	X			
Mallard/MN/284/98	H7	X	A		X
Mallard/MN/290/99	H3		B		
Mallard/MN/304/98	H9	X	A		

Table 1 (*Continued*)

Isolate	HA subtype	Gene segment			
		M	NS	HA	NP
Mallard/MN/309/98	H9		A		
Mallard/MN/31/98	H6	X	A	X	X
Mallard/MN/327/98	H6		A	X	X
Mallard/MN/33/00	H4	X	B	X	X
Mallard/MN/334/99	H6	X	A	X	X
Mallard/MN/346233/00	H6	X	B	X	X
Mallard/MN/347/99	H4	X	B	X	X
Mallard/MN/351/99	H3	X	B		
Mallard/MN/352/99	H4	X	B	X	X
Mallard/MN/354/99	H3	X	B		
Mallard/MN/355803/00	H6	X	A	X	X
Mallard/MN/365/99	H9	X	A		
Mallard/MN/371/98	H4	X	A	X	X
Mallard/MN/403/99	H11	X	B		
Mallard/MN/51/98	H2	X	A		
Mallard/MN/530/00	H4		A	X	
Mallard/MN/68/99	H3	X	B		
Mallard/MN/7/98	H11	X	A		
Mallard/MN/88/98	H7	X	A		
Mallard/MN348/00	H4	X	B	X	X
Mallard/OH/242/98	H6	X	A	X	X
Mallard/OH/249/98	H6	X	A	X	X
Mallard/OH/338/86	H4	X	A	X	X
Mallard/OH/386/88	H6	X	A	X	X
Mallard/OH/66/99	H6	X	A	X	X
Mallard/OH/73/89	H6	X	B	X	X
Mottled duck/LA/32M/87	H6	X	B	X	X
Mottled duck/LA/38M/87	H6	X	A	X	X
Northern pintail/TX/828189/02	H6	X	B	X	X
Northern pintail/TX/828197/02	H6		B	X	X
Northern pintail/AK/3209/02	H3	X	B		X
Pintail/MN/410/99	H6	X	A	X	X
Pintail/MN/423/99	H7	X	B		
Pintail/MN/431/99	H6	X	B	X	X
Pintail/MN/479/99	H3	X	A		
Pintail/OH/25/99	H6	X	A	X	X
Ring-billed gull/DE/421733/01	H6	X	A	X	X
Red knot/NJ/828235/01	H6	X	A	X	X
Ruddy turnstone/DE/241731/01	H6	X	A	X	X
Ruddy turnstone/DE/650679/02	H6	X	A	X	X
Ruddy turnstone/NJ/1407/01	H4	X	A		X
Ruddy turnstone/DE/108/00	H10		A		
Ruddy turnstone/DE/1121/00	H10		A		
Ruddy turnstone/DE/1444/00	H12		A		
Ruddy turnstone/DE/1538/00	H7	X	A		
Ruddy turnstone/DE/1602/00	H12		A		
Ruddy turnstone/DE/1749/00	H12	X	A		
Ruddy turnstone/DE/629/00	H7	X	A		
Ruddy turnstone/DE/748/00	H12		A		
Ruddy turnstone/NJ/1143/00	H10		A		
Ruddy turnstone/NJ/2242/00	H5	X	A		
Ruddy turnstone/NJ/892/02	H6	X	B	X	X
Sanderling/DE/650680/02	H6	X	A	X	X
Semipalmated sandpiper/DE/2109/00	H11	X	A		
Semipalmated sandpiper/DE/2144/00	H11		A		
Wild bird/DE/1185/02	H6	X	B	X	X
Wood duck/MD/M58/98	H9	X	A		X

[a] X: indicates that the specified gene from was included in the study.

[b] A or B indicate the NS gene subtype and that the gene was included in the study.

The remaining genes analyzed; M, NP and NS, are the most commonly used in epidemiological tracing of AIV outbreaks in commercial poultry.

All viruses were sequenced after not more than 1–3 passages in ECE. RNA was extracted with Trizol LS reagent (Invitrogen, Inc., Carlsbad, CA) in accordance with manufacturers instructions. Individual influenza genes were then amplified by RT-PCR as previously described (Suarez et al., 1999). Because HA gene sequence variability made direct sequencing difficult, the full length HA gene or HA1 region of the HA gene was cloned into the pCR2.1 vector with the TopoTA kit (Invitrogen Inc., Carlsbad, CA) in accordance with the kit instructions. Only the HA1 regions were included the analysis as follows: the HA1 from H6 isolates included bases 1–1200 of the coding region and the HA1 of H4 isolates included bases 1–1125 of the coding region. RT-PCR products of the NS, M and NP genes including the full coding region (with the exception of the NP for which the analyzed product was missing the 15 bases from the 3′end) were purified by agarose gel extraction with the Qiaquick gel extraction kit (Qiagen, Inc., Valencia, CA) and were directly sequenced. The BigDye terminator kit (Applied Biosystems, Foster City, CA) was used for cycle sequencing and subsequently run on an ABI 3730 (Applied Biosystems, Foster City, CA).

2.3. Phylogenetic and sequence analysis

Sequences were aligned with Clustal V (Lasergene, DNAStar, Madison, WI). Phylogenetic analysis was performed with PAUP* 4.0b10 (Sinauer Associates, Inc., Sunderland, MA) using the maximum parsimony tree building method, with heuristic search with 100 bootstrap replicates. Genes were determined to be in separate phylogenetic groups based on 7.5% or greater divergence among the H4 subtype HA, M, NP and NS genes. Because of greater sequence variability 10% or greater divergence was used to separate H6 subtype HA phylogenetic groups.

Each gene was analyzed both with all available wild waterfowl and shorebird isolate sequences and in comparison with virus genes from gallinaceous poultry (chicken, turkey and quail) and mammalian (including human, swine and equine) origin isolates. The analysis included 232 M, 225 NP and 268 NS subtype A genes. Data from 102 NS subtype B genes from gallinaceous poultry and one mammalian isolate; Equine/Jilin/89 were also included in the analysis. The H6 subtype HA included sequences from 55 additional isolates and the H4 HA included 14 additional isolates.

2.4. GenBank accession numbers

The sequences reported here have the following GenBank numbers – matrix gene: DQ021685-DQ021781; nucleoprotein gene: DQ021782-DQ021847; non-structural gene: DQ021543-DQ021648; H4 hemagglutinin: DQ021848-DQ021868; H6 hemagglutinin: DQ021649-DQ021684.

Table 2
Range of nucleotide and amino acid percent identities among all reported influenza isolates from North American wild aquatic birds

Gene/protein	Percent nucleotide identity	Percent amino acid identity
M	87.4–99.8	
M1		95.7–99.6
M2		87.6–100
NS subtype A	82.1–98.2	
NS1		85.7–99.1
NS2		94.2–100
NS subtype B	96.1–100	
NS1		98.0–99.6
NS2		99.0–100
NP	82.9–99.3	95.7–99.8
H4 HA1	89.8–100	92.0–100
H6 HA1	70.1–100	77.7–100

3. Results

3.1. Hemagglutinin

The HA1 of the HA genes of 38 H6 subtype duck and shorebird origin isolates were sequenced. Phylogenetic analysis was performed with all available H6 HA1 sequences (Chin et al., 2002; Hatchette et al., 2004; Hoffmann et al., 2000; Smirnov et al., 1999; Webby et al., 2002; Woolcock et al., 2003) and included a total of 93 isolates. The H6 HA genes assorted into four groups (Fig. 1). Groups 1, 3 and 4 contained only North American origin isolates. Group 2 contained both North American and Eurasian origin viruses. The most distantly related wild bird origin isolates had 70.1% nucleotide (nt) identity (between isolates in groups 1 and 4) (Table 2). Within each group there was approximately 92.5–100% nt identity and between the groups there was 70.1–90% nt identity. The H6 HA1 protein had between 77.7 and 100% amino acid (aa) identity among the North American origin wild bird isolates. The phylogenetic assortment of the proteins was similar to that of the nt sequences. The isolates did not assort by chronology, geography or species aside from the North American affinities of groups 1, 3 and 4.

Among the 38 H6 isolates there were six different HA cleavage site sequences. No sequences were consistent with what has been observed in H5 and H7 isolates that have been classified as highly pathogenic AI viruses for gallinaceous poultry.

A total of 21 H4 subtype HA1 genes were sequenced from wild aquatic bird isolates. All available H4 HA sequences (Donis et al., 1989; Karasin et al., 2000) were included in the phylogenetic analysis for a total of 35 isolates. The H4 HA1 genes assorted into four groups (Fig. 2), including one group with Eurasian origin isolates and Seal/MA/133/82. Groups 1–3 contained only North American origin viruses. Within each group there was 92.5–100% nt identity and 96.0–100% aa identity. Between the four groups there was 79.5–84% nt identity and 91–94% aa identity including the wild bird, domestic bird and mammalian isolates. The most distantly related wild bird isolates had 89.8% nt identity.

Among the 21 wild bird isolate sequences three different cleavage site sequences were observed. None were consistent with sequences from H5 and H7 isolates that have been classified as highly pathogenic AI viruses for gallinaceous poultry.

3.2. Nucleoprotein gene

The nucleoprotein genes of 67 isolates were sequenced. Among the wild-bird isolates there was between 82.9 and 99.3% nt identity and 95.7–99.8% aa identity (Table 2). There were two phylogenetic groups of NP genes of the North American wild aquatic birds (Fig. 3), with all but three avian isolates were in group 1. There were also two separate groups representing Eurasian avian and mammalian isolates. Isolates in group 1 were closely related with at least 96.7% nt identity among the isolates. However the second group had approximately 83% nt identity to group 1 and contained the isolates previously reported by Gorman et al. (1990). The NP aa tree had similar topology to the nt tree. The isolates did not clearly group by year of isolation, location within North America or species of origin.

Phylogenetic analysis with poultry and mammalian isolates revealed that the wild aquatic bird NP genes assorted separately from the mammalian and Eurasian avian viruses (Fig. 3) and were most closely related to North American poultry isolates and seal/MA/1/80.

3.3. Non-structural gene

The NS genes from 107 isolates were sequenced; 75 (70%) were subtype A and 32 (30%) were subtype B. Phylogenetic analysis revealed three subtype A groups and two subtype B groups among all North American wild aquatic bird isolates (Fig. 4). The NS nt trees had identical topology to the NS1 and NS2 protein trees. Among the wild bird isolates there was between 82.1 and 98.2% nt identity among subtype A genes (Table 2). Within each NS subtype A group there was between 92.7 and 99.9% nt identity and 82.1–92.1% nt identity between groups. Amino acid identity among wild bird origin subtype A genes was between 85.7 and 99.1% in the NS1 and from 94.2 to 100% in the NS2 protein. There was approximately 67% nt identity between subtype A and B. Among all subtype B genes there was at least 96.1% nt identity (Table 2). Within the NS subtype B group 1, the isolates from the 1970s appeared separate from the more recent isolates; the nt differences were 93.9–96% between these sub-groups and may represent a temporal sub-grouping but sampling is inadequate. Within the NS subtype B groups there was between 92.9 and 100% nt identity and 85.7–90.7% nt identity between the two groups. Among subtype B genes amino acid identity was approximately 98.1% in the NS1 and 99.0% in the NS2. No deletions or insertions were observed in the NS1 protein from any isolate of either subtype.

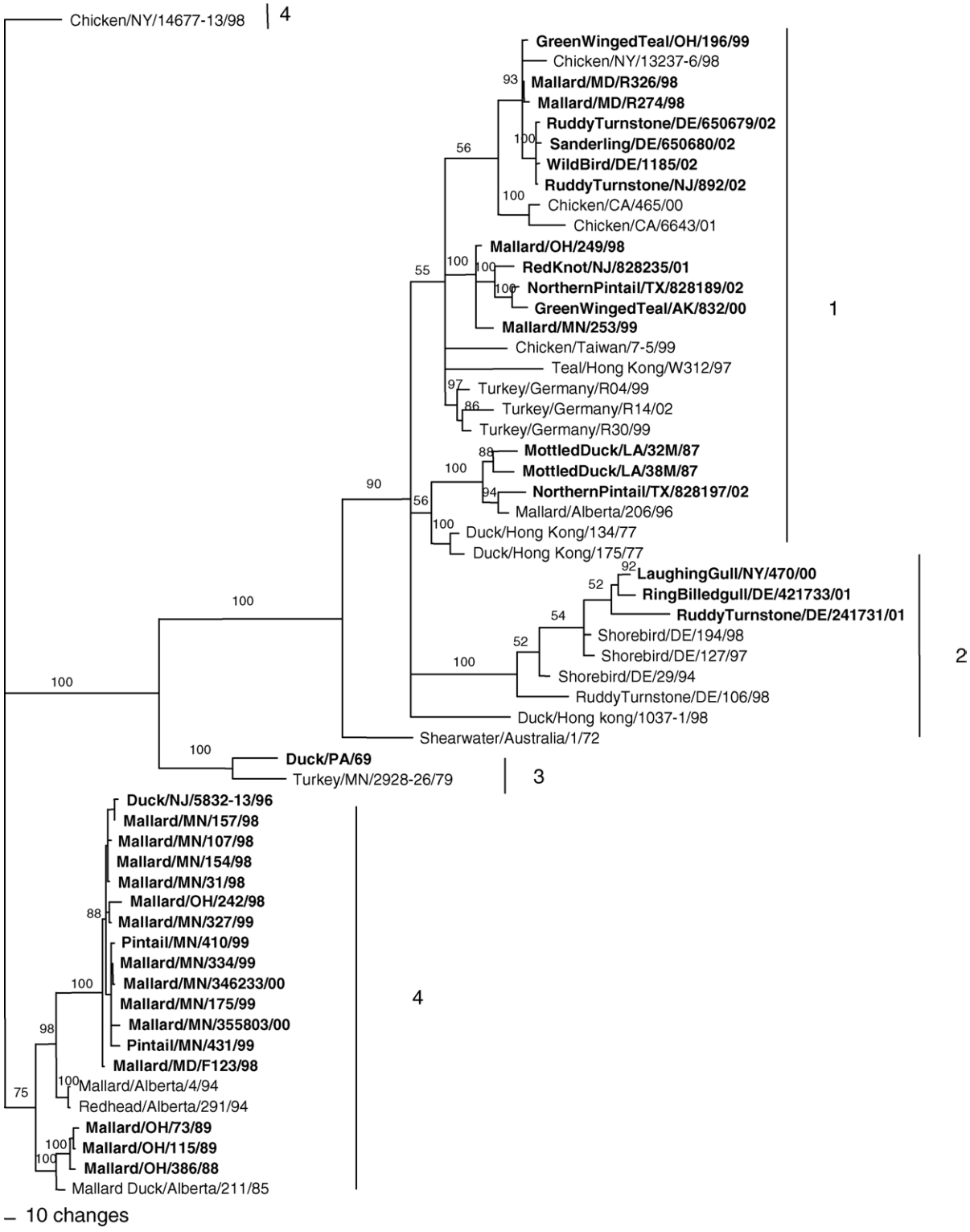

Fig. 1. Unrooted phylogenetic tree of the H6 HA1 including selected North American origin wild aquatic birds isolates (from this report and previously reported) and selected isolates from avian and mammalian species. The tree was constructed with PAUP* 4.0b10 (Sinauer Associates, Sunderland, MA) using maximum parsimony, heuristic search and 100 bootstrap replicates (bootstrap values are shown on tree). Sequence produced for this report is in boldface type. States are abbreviated by their standard two letter postal code.

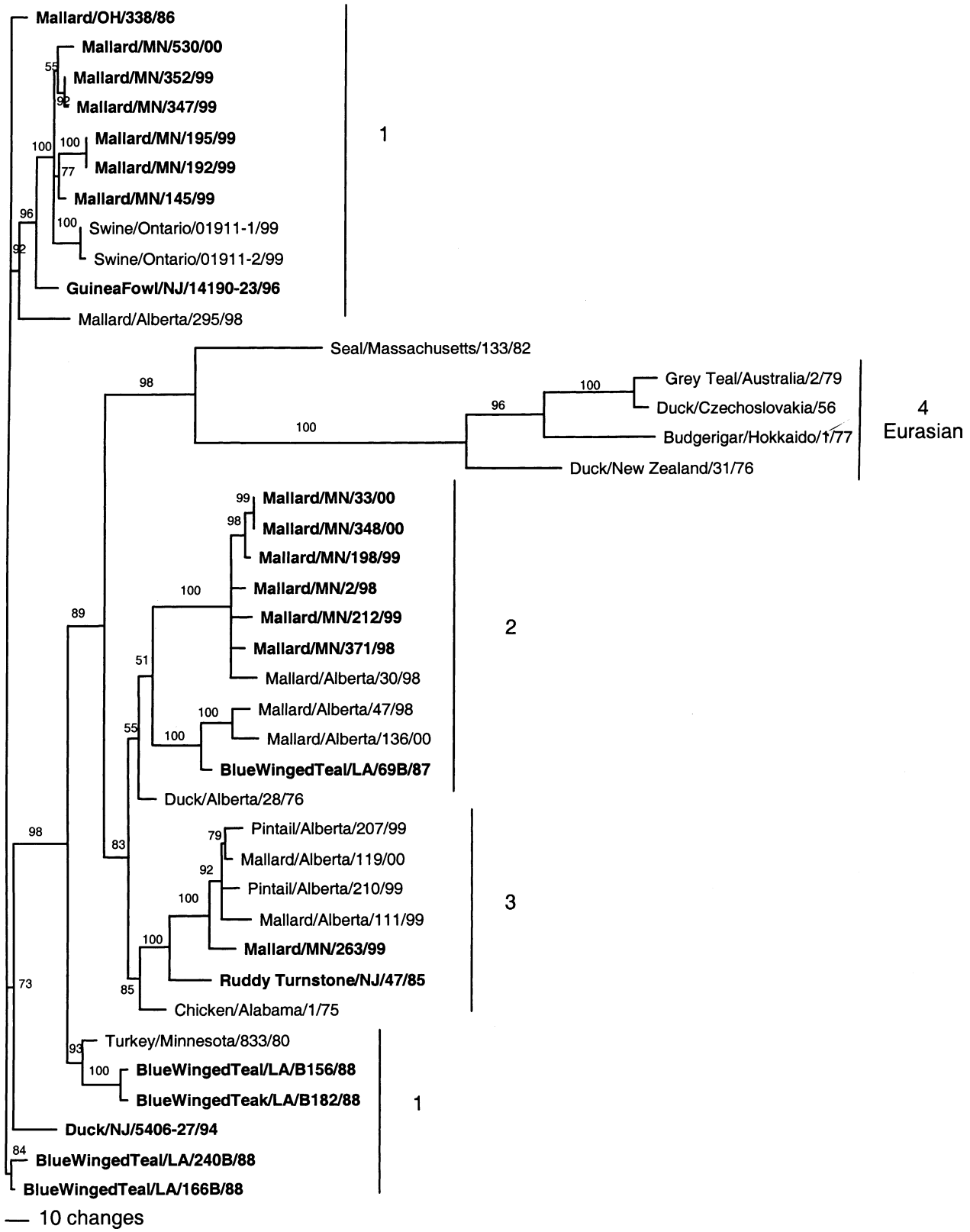

Fig. 2. Unrooted phylogenetic tree of all available H4 HA1 sequence. The tree was constructed with PAUP* 4.0b10 (Sinauer Associates, Sunderland, MA) using maximum parsimony, heuristic search and 100 bootstrap replicates (bootstrap values are shown on tree). Sequence produced for this report is in boldface type. States are abbreviated by their standard two letter postal code.

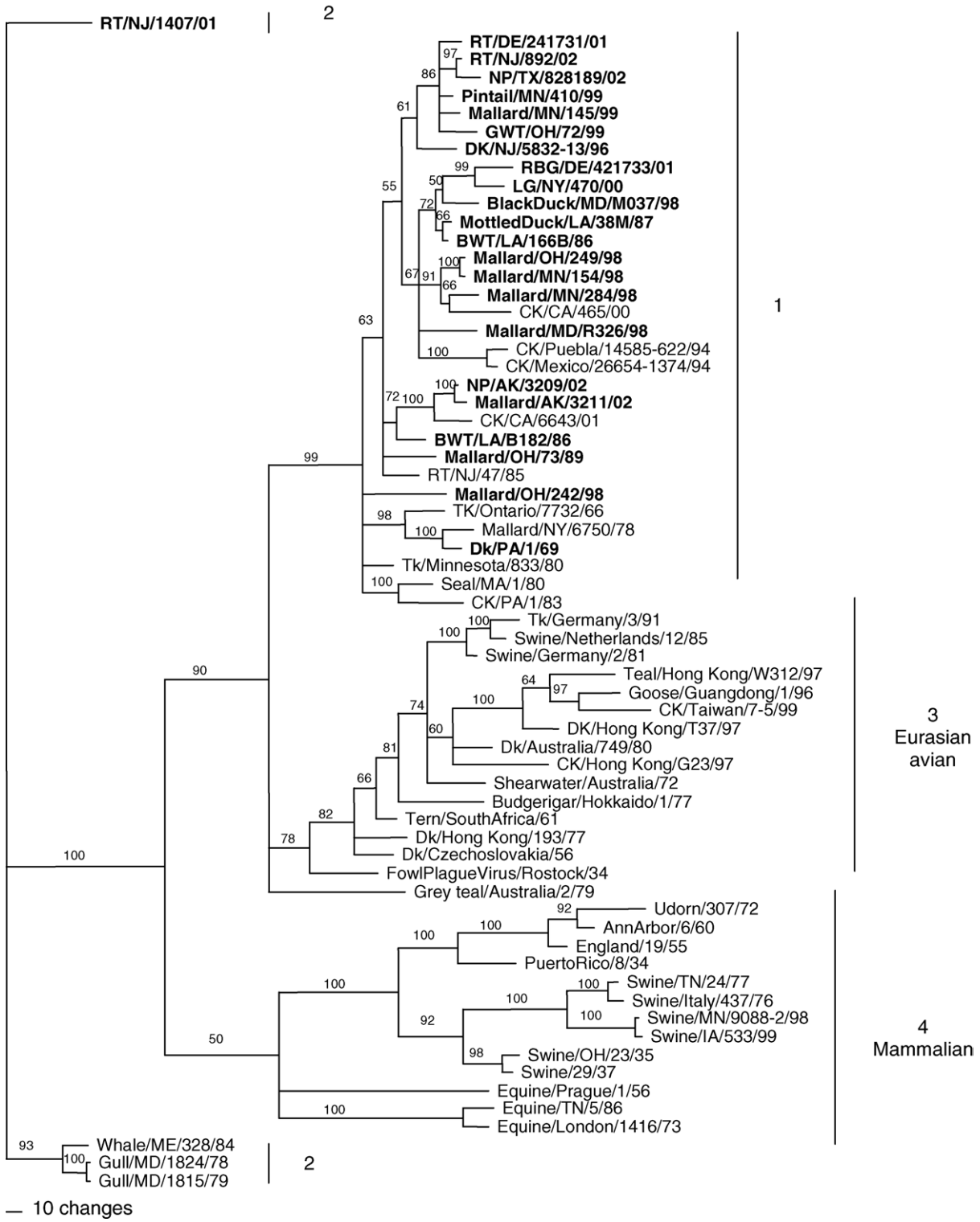

Fig. 3. Unrooted phylogenetic tree of the NP gene including selected North American origin wild aquatic birds isolates (from this report and previously reported) and selected isolates from avian and mammalian species. The tree was constructed with PAUP* 4.0b10 (Sinauer Associates, Sunderland, MA) using maximum parsimony, heuristic search and 100 bootstrap replicates (bootstrap values are shown on tree). Sequence produced for this report is in boldface type. Abbreviations: BWT = blue-winged teal, CK = chicken, DK = duck, GWT = green-winged teal, LG = laughing gull, NP = northern pintail, RBG = ring-billed gull, RT = ruddy turnstone, TK = turkey. States are abbreviated by their standard two letter postal code.

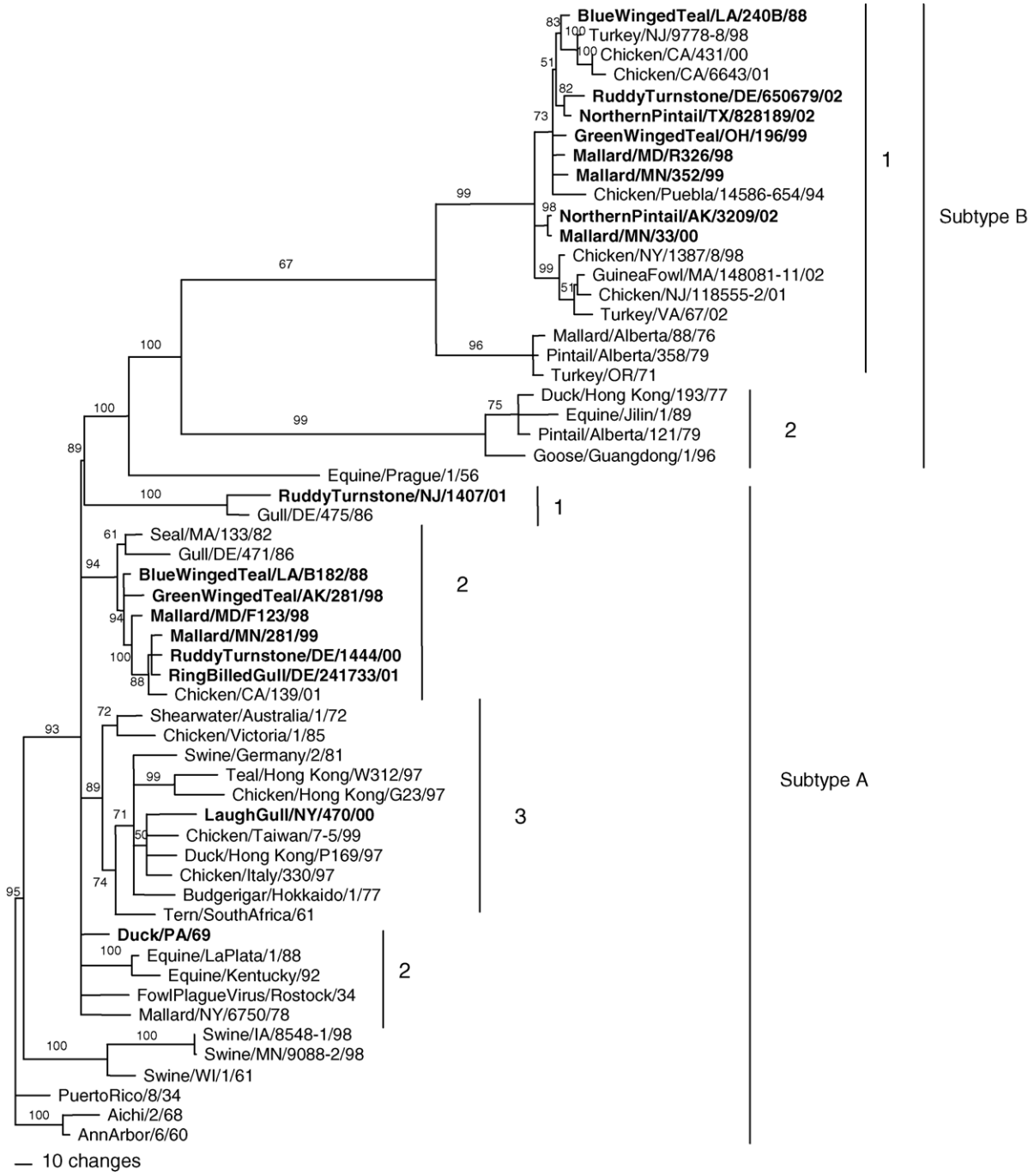

Fig. 4. Unrooted phylogenetic tree of the NS gene including selected North American origin wild aquatic birds isolates (from this report and previously reported) and selected isolates from avian and mammalian species. The tree was constructed with PAUP* 4.0b10 (Sinauer Associates, Sunderland, MA) using maximum parsimony, heuristic search and 100 bootstrap replicates (bootstrap values are shown on tree). Sequence produced for this report is in boldface type. States are abbreviated by their standard two letter postal code.

There was no clear geographical (within North America), chronological or species assortment of the wild aquatic bird origin isolates. When analyzed with poultry, swine, equine and human isolates, the mammalian and poultry origin isolates grouped separately by Eurasian and North American lineages, several subtype A wild aquatic bird isolates from North America assorted with Eurasian origin poultry and mammalian viruses. The subtype B NS also had one North American isolate that grouped with Eurasian lineage viruses; Pintail/Alberta/121/79 (Treanor et al., 1989).

The chicken origin subtype B genes assorted closely with the wild aquatic bird subtype B NS genes and had 3–5.5% nt divergence, however there were 26 unique nt, changes in the

poultry isolates that corresponded to 11 unique NS1 and 2 unique NS2 amino acid changes.

3.4. Matrix gene

Matrix genes from 98 isolates were sequenced in this study. Among these isolates the nt identity was between 87.4 and 99.8% (Table 2). Among all M1 amino acid sequences there was at least 95.7% identity among the isolates. Amino acid identity among the M2 genes was between 87.6 and 100%.

The M genes of North American wild aquatic birds assorted into three phylogenetic groups and three groups representing Eurasian avian and mammalian isolates, equine isolates and human and swine isolates (Fig. 5). Within group three numerous isolates from various times, locations and species, attached to the trunk and had at least 98% nt identity with other isolates attached to the trunk. Among the North American wild aquatic bird isolates there was no clear species, location (within North America) or chronological assortment of the genes; for example isolates collected at the same time from the same species and location (i.e. mallard/MN/7/98 and mallard/MN/154/98), could have as little as 90.4% nt identity. Conversely, duck/PA/69 had 98.5% nt identity with blue-winged teal/LA/166B/88, which are from different locations (and migratory fly-ways) and were collected 19 years apart.

The topology of the M2 protein tree was the same as the nt tree, however the M1 protein tree consisted of two groups one corresponded to groups two and three of the nt tree and the M1 protein from the nt tree group one was also a separate group on the protein tree.

When the M genes of the wild aquatic birds were compared to those of poultry and mammalian isolates, the most closely related isolates were from outbreaks in poultry.

4. Discussion

Historically, it has been assumed that avian influenza viruses are genetically stable and at evolutionary stasis in their natural host species, aquatic birds (Gorman et al., 1992). However influenza viruses, as characteristic of RNA viruses, have a relatively high substitution rate (Parvin et al., 1986), therefore some level of genetic drift is expected. The levels of variation in nt sequence identities observed among isolates from wild aquatic birds in the NS subtype A, M and H6 subtype HA1 genes were similar to that observed between the most distantly related type A influenza isolates from non-natural host species (82 and 87% minimum identity for both NS subtype A and the M gene segment and around 70% for the H6 HA1). Similar results of 82.7% minimum identity have been reported for the M gene previously (Widjaja et al., 2004). Importantly, the amino acid sequence identities reflect the level of variation observed in nt sequences indicating that these are often non-synonymous changes. Too little information for the NS subtype B and H4 HA1 is available from other host species to draw conclusions, although the NS subtype B appears to be very highly conserved.

Importantly, although, the levels of nucleotide divergence among H6 HA1, NS subtype A and M genes among wild aquatic birds vary to a similar extent to the levels observed in other species, the phylogenetic assortment patterns are different from what is seen in isolates from non-natural host species. Unlike patterns observed among isolates from poultry, swine and humans (reviewed by Gorman et al., 1992; Suarez, 2000), wild aquatic bird origin AIV isolate genetic relatedness was not clearly associated with geographical location within North America or date of isolation or species of origin. Also, there did not seem to be species based differences between duck isolates and shorebird isolates, although there is a sample bias toward more duck isolates. The observed variance is attributed to differences in selection in natural hosts (i.e., wild aquatic birds) and host adaptation in non-natural hosts (i.e. poultry and mammals). Additionally, viruses collected at a single location and at the same sample time could vary substantially, as numerous isolates had less identity with other isolates collected at the same time and location from the same species, than with isolates with no geographical or temporal relationship, suggesting that multiple lineages circulate through populations simultaneously. Mixing of migratory birds at a single location also may be responsible for the high level of variation. Similarly, equine influenza has also been shown to persist as multiple co-circulating lineages based on HA sequences (Daly et al., 1996; Manuguerra et al., 2000).

The one grouping observed with these wild aquatic birds sequences which is similar to that of poultry and some previous reports of free-flying bird origin AIV is the broad geographic assortment of numerous AIV genes (M, NS, NP, H5 subtype HA and H7 subtype HA) into North American and Eurasian lineages (Donis et al., 1989; Gorman et al., 1990; Schafer et al., 1993). The H4 HA also appears to assort into North American and Eurasian lineages, however the sequence from Eurasian wild aquatic bird isolates is very limited. Alternatively, several North American origin wild aquatic bird isolates assorted into major clades with genes from Eurasian isolates as shown here with the H6 HA and NS subtype A genes indicating that there is some intercontinental exchange of genes. One NS subtype B gene from Pintail/Alberta/121/79 (Treanor et al., 1989) was closely related to Goose/Gaundong/1/96, the donor virus of the HA of the H5N1 highly pathogenic avian influenza that is currently circulating in Asia. Similar results have been shown with the H2 subtype HA from wild shorebirds (Makarova et al., 1999). This demonstrates that intercontinental, transoceanic dissemination of AIV, although probably rare, may occur with migratory water birds.

Differences in topologies among the wild aquatic bird isolate phylogenetic trees of different genes indicate that avian influenza virus undergoes reassortment in wild aquatic birds, which supports previous reports that reassortment is a normal

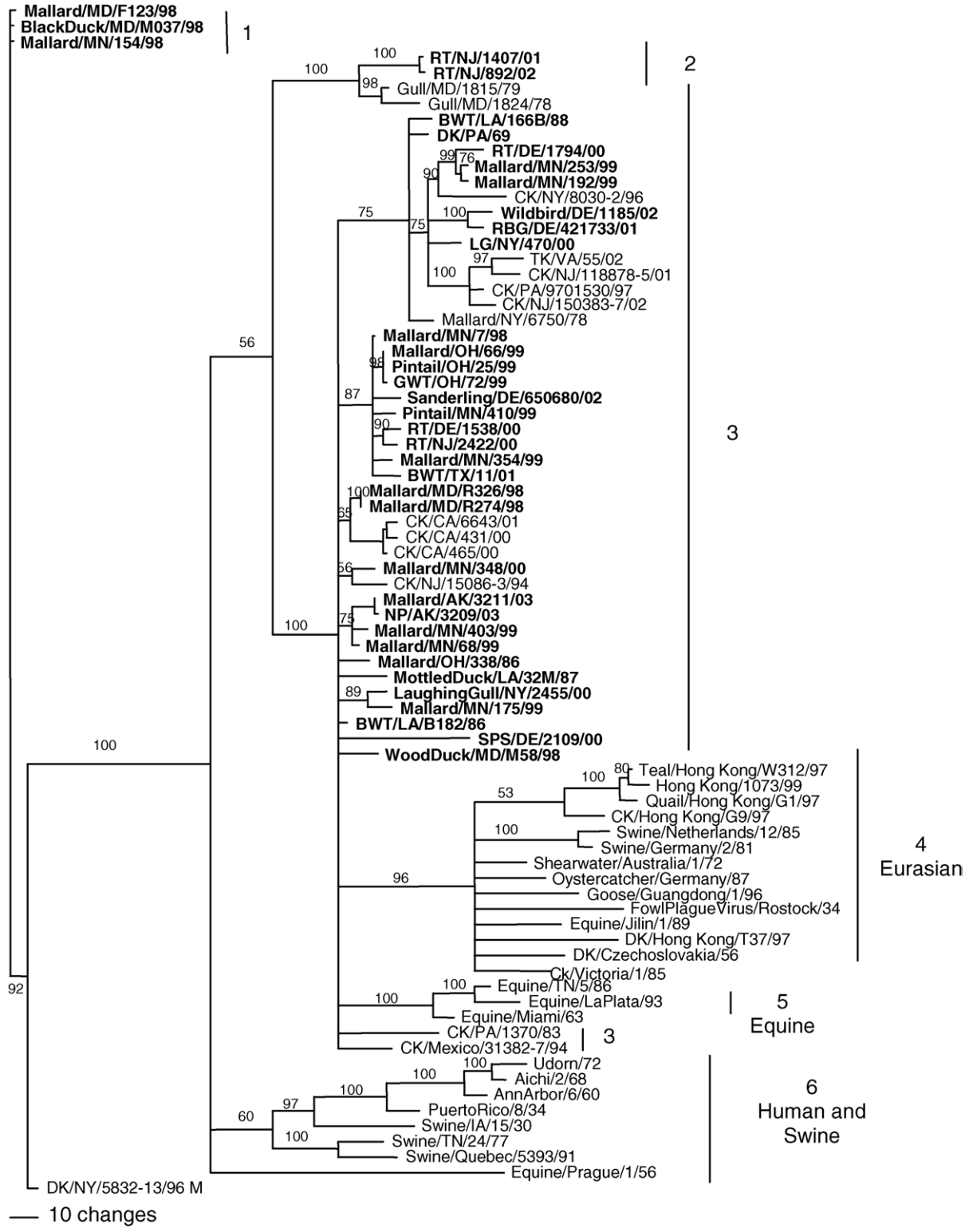

Fig. 5. Unrooted phylogenetic tree of the M gene with selected North American origin wild aquatic birds isolates (from this report and previously reported) and selected isolates from avian and mammalian species. The tree was constructed with PAUP* 4.0b10 (Sinauer Associates, Sunderland, MA) using maximum parsimony, heuristic search and 100 bootstrap replicates (bootstrap values are shown on tree). Sequence produced for this report is in boldface type. Abbreviations: BWT = blue-winged teal, CK = chicken, DK = duck, GWT = green-winged teal, LG = laughing gull, NP = northern pintail, RBG = ring-billed gull, RT = ruddy turnstone, SPS = semipalmated sandpiper, TK = turkey. States are abbreviated by their standard two letter postal code.

part of the ecology of influenza in aquatic birds (Hatchette et al., 2004; Sharp et al., 1993). In order for reassortment of the influenza gene segments to occur, one host must be simultaneously infected with two (or more) virus strains, this suggests therefore that the environmental and biological reservoirs of influenza maintain multiple strains concurrently.

Analysis of individual AIV genes from wild aquatic bird isolates should help in understanding of gene function that can be applied to studying host restriction and pathogenesis as the "natural" sequences of AIV are determined. For example the NS subtype B gene has only been found in avian isolates, with the exception of Euine/Jilin/89 (Guo et al., 1992) and although it is not uncommon in chicken and turkey origin isolates (Ludwig et al., 1991; Suarez et al., 1999; Suarez and Perdue, 1998; Treanor et al., 1989; Webby et al., 2002) it appears to be less common in natural host species than subtype A, comprising only about 30% of the isolates sequenced. The functional importance of NS subtype B has not been defined, although re-assortment studies that have placed NS subtype B genes into mammalian origin viruses have attenuated these viruses in mice (Treanor et al., 1989). Currently it is unclear whether the apparently lower prevalence of NS subtype B is due to decreased fitness in water birds, sample bias or chance. Additionally, the NS1 protein (subtype A) has been shown to vary in length in some influenza isolates from mammalian hosts and poultry, but was a consistent length of 230 aa in the limited number of wild aquatic bird isolates included in a report by Suarez and Perdue (1998). In this report the NS1 gene length was 230 aa in all isolates of either subtype, indicating that this may be the natural length for the protein and adaptation in non-natural hosts may select for different protein lengths.

Finally, the seemingly almost random assortment pattern of some gene segments from wild aquatic bird origin influenza isolates precludes the development of an epidemiological map for avian influenza from these hosts. Although virus sequences from outbreaks in other species may be compared to determine how closely related they are to a wild aquatic bird virus sequence, the species or place of origin could not be discerned based on sequence data alone. Also, it may be possible to estimate how long an isolate has been circulating in a new host based on how much identity the sequence has with sequences from characterized wild aquatic bird isolate sequences. Finally, the unexpectedly high degree of variation among isolates from the same species and location at the same collection time re-enforces that surveillance and analysis requires a large number of samples from diverse locations and species for accurate and precise data collection.

Acknowledgments

The authors would like to gratefully acknowledge Joan Beck, Scott Lee, Dawn Williams-Coplin, Joyce Bennett, Michael Wege, M.P. Lutrell, B.A. Hanson and Suzanne DeBlois for technical assistance with this work. This work was funded by USDA CRIS project numbers 6612-32000-039-00D, 6612-32000-041-01S, 6612-32000-041-04S, 6612-32000-041-06S and 6612-32000-039-03S. Mention of trade names or commercial products in this manuscript is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

References

Beare, A.S., Webster, R.G., 1991. Replication of avian influenza viruses in humans. Arch. Virol. 119, 37–42.

Chin, P.S., Hoffmann, E., Webby, R., Webster, R.G., Guan, Y., Peiris, M., Shortridge, K.F., 2002. Molecular evolution of H6 influenza viruses from poultry in southeastern China: prevalence of H6N1 influenza viruses possessing seven A/Hong Kong/156/97 (H5N1)-like genes in poultry. J. Virol. 76, 507–516.

Daly, J.M., Lai, A.C., Binns, M.M., Chambers, T.M., Barrandeguy, M., Mumford, J.A., 1996. Antigenic and genetic evolution of equine H3N8 influenza A viruses. J. Gen. Virol. 77 (Pt. 4), 661–671.

Donis, R.O., Bean, W.J., Kawaoka, Y., Webster, R.G., 1989. Distinct lineages of influenza virus H4 hemagglutinin genes in different regions of the world. Virology 169, 408–417.

Gorman, O.T., Bean, W.J., Kawaoka, Y., Webster, R.G., 1990. Evolution of the nucleoprotein gene of influenza A virus. J. Virol. 64, 1487–1497.

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.

Guo, Y., Wang, M., Kawaoka, Y., Gorman, O., Ito, T., Saito, T., Webster, R.G., 1992. Characterization of a new avian-like influenza A virus from horses in China. Virology 188, 245–255.

Hatchette, T.F., Walker, D., Johnson, C., Baker, A., Pryor, S.P., Webster, R.G., 2004. Influenza A viruses in feral Canadian ducks: extensive reassortment in nature. J. Gen. Virol. 85, 2327–2337.

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 southeastern China: Was H6N1 a derivative or a precursor of H5N1? J. Virol. 74, 6309–6315.

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.

Johnson, D.C., Maxfield, B.G., Moulthrop, J.I., 1977. Epidemiologic studies of the 1975 avian influenza outbreak in chickens in Alabama. Avian. Dis. 21, 167–177.

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.

Ludwig, S., Schultz, U., Mandler, J., Fitch, W.M., Scholtissek, C., 1991. Phylogenetic relationship of the nonstructural (NS) genes of influenza A viruses. Virology 183, 566–577.

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.

Manuguerra, J.C., Zientara, S., Sailleau, C., Rousseaux, C., Gicquel, B., Rijks, I., van der Werf, S., 2000. Evidence for evolutionary stasis and genetic drift by genetic analysis of two equine influenza H3 viruses isolated in France. Vet. Microbiol. 74, 59–70.

Okazaki, K., Takada, A., Ito, T., Imai, M., Takakuwa, H., Hatta, M., Ozaki, H., Tanizaki, T., Nagano, T., Ninomiya, A., Demenev, V.A., Tyaptirganov, M.M., Karatayeva, T.D., Yamnikova, S.S., Lvov, D.K., Kida, H., 2000. Precursor genes of future pandemic influenza viruses are perpetuated in ducks nesting in Siberia. Arch. Virol. 145, 885–893.

Parvin, J.D., Moscona, A., Pan, W.T., Leider, J.M., Palese, P., 1986. Measurement of the mutation rates of animal viruses: influenza A virus and poliovirus type 1. J. Virol. 59, 377–383.

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.

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.

Slemons, R.D., Johnson, D.C., Oborn, J.S., Hayes, F., 1974. Type-A influenza viruses isolated from wild free-flying ducks in California. Avian. Dis. 18, 119–124.

Smirnov, Y.A., Lipatov, A.S., Gitelman, A.K., Okuno, Y., Van Beek, R., Osterhaus, A.D., Claas, E.C., 1999. An epitope shared by the hemagglutinins of H1, H2, H5, and H6 subtypes of influenza A virus. Acta Virol. 43, 237–244.

Stallknecht, D.E., Shane, S.M., Zwank, P.J., Senne, D.A., Kearney, M.T., 1990. Avian influenza viruses from migratory and resident ducks of coastal Louisiana. Avian. Dis. 34, 398–405.

Suarez, D.L., 2000. Evolution of avian influenza viruses. Vet. Microbiol. 74, 15–27.

Suarez, D.L., Garcia, M., Latimer, J., Senne, D., Perdue, M., 1999. Phylogenetic analysis of H7 avian influenza viruses isolated from the live bird markets of the northeast United States. J. Virol. 73, 3567–3573.

Suarez, D.L., Perdue, M.L., 1998. Multiple alignment comparison of the non-structural genes of influenza A viruses. Virus Res. 54, 59–69.

Swayne, D.E., Senne, D.A., Beard, C.W., 1998. Avian influenza. In: Swayne, D.E. (Ed.), A Laboratory Manual for the Isolation and Identification of Avian Pathogens, 4th ed. The American Association of Avian Pathologists, Tallahassee, FL, pp. 150–155.

Treanor, J.J., Snyder, M.H., London, W.T., Murphy, B.R., 1989. The B allele of the NS gene of avian influenza viruses, but not the A allele, attenuates a human influenza A virus for squirrel monkeys. Virology 171, 1–9.

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.

Widjaja, L., Krauss, S.L., Webby, R.J., Xie, T., Webster, R.G., 2004. Matrix gene of influenza A viruses isolated from wild aquatic birds: ecology and emergence of influenza A viruses. J. Virol. 78, 8771–8779.

Woolcock, P.R., Suarez, D.L., Kuney, D., 2003. Low-pathogenicity avian influenza virus (H6N2) in chickens in California, 2000–2002. Avian. Dis. 47, 872–881.