

# The novel swine-origin H1N1 influenza A virus riddle: is it a domestic bird H1N1-derived virus?

Muhammed Babakir-Mina<sup>1</sup>, Salvatore Dimonte<sup>2</sup>, Massimo Ciccozzi<sup>3</sup>,  
Carlo Federico Perno<sup>1,2</sup>, Marco Ciotti<sup>1</sup>

<sup>1</sup>Laboratory of Molecular Virology, Foundation University Hospital Tor Vergata, Rome, Italy;

<sup>2</sup>Virology, Department of Experimental Medicine and Biochemistry, University of Rome Tor Vergata, Italy;

<sup>3</sup>Department of Infectious, Parasitic and Immunomediated Diseases, Istituto Superiore di Sanità, Rome, Italy

## SUMMARY

To understand the role of domestic birds in the 2009 H1N1 influenza A outbreak, a phylogenetic analysis of hemagglutinin, neuraminidase and matrix protein genes from human, avian and swine H1N1 viruses was carried out. Analysis of the H1 sequences revealed that the virus evolved most likely from American swine as well as intermixing between Asian swine and American domestic bird H1N1 viruses. Neuroaminidase and matrix protein analysis showed that the H1N1 2009 viruses were more closely related to the H1N1 isolates from Euro-Asiatic domestic birds and swine than wild birds. Domestic birds could act as intermediate hosts of H1N1 reassortants.

**KEY WORDS:** Swine flu, Avian flu, Phylogenetic analysis, Molecular evolution, Influenza pandemic

Received September 8, 2009

Accepted September 30, 2009

On 24<sup>th</sup> April 2009 the Center for Disease Control and Prevention (CDC) reported human cases of swine influenza A (H1N1) occurring in the USA and Mexico (Centers for Disease Control and Prevention, CDC, 2009). The number of recorded cases increased rapidly: there was a clear human to human transmission and it appeared that the outbreak originated in Mexico in mid-March or earlier (Centers for Disease Control and Prevention, CDC, 2009). Soon after, it was proposed that the current flu outbreak is caused by a new influenza A (H1N1) virus generated from a triple reassortment of human, swine and avian viruses (Butler, 2009; Cohen, 2009). Recent publications, including ours, showed that this new influenza A (H1N1) virus most likely evolved from recent swine viruses (Babakir-Mina *et al.*, 2009; Garten *et al.*, 2009; Dawood *et al.*, 2009; Solovyov *et al.*, 2009). The segmented nature of the genome

of influenza virus and the receptor molecules present in the swine respiratory tract contribute to the evolutionary diversity of influenza viruses (Brown, 2000). Indeed, this genome structure enables influenza viruses to exchange their genes when two or more viruses of different origin infect a host at the same time. Also the existence of 2-3 and 2-6 linked sialic acid in the swine respiratory tract allows type A influenza viruses of human and avian origin to infect pigs (Ito *et al.*, 1998). Our previous phylogenetic study showed that the H1N1 isolates possessed an evolutionarily distinct set of surface antigens and matrix proteins with the H1 gene derived from the classical American swine lineage and the N1 and matrix protein (MP) genes from the European swine lineage (Babakir-Mina *et al.*, 2009). In order to explore further the extent of the genetic diversity, we analysed the whole sequence of hemagglutinin (HA), neuraminidase (NA) and matrix protein (MP) genes of the H1N1 viruses (avian, swine, and human strains) isolated from 1918 to 2009. We attempted to identify the origin and the genetic relationship between the avian (domestic and wild birds), swine and human H1N1 isolates worldwide. Moreover, we considered the

Corresponding author

Dr. Marco Ciotti

Laboratory of Molecular Virology

Foundation University Hospital Tor Vergata

Viale Oxford, 81 - 00133, Rome, Italy

E-mail: marco.ciotti@ptvonline.it

evolutionary effect of H1N1 viruses isolated from domestic and wild birds on the novel swine influenza virus.

In this study, 225 HA, 186 NA, and 153 MP sequences from human, avian, and swine H1N1 viral isolates were downloaded from Genbank (<http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html>) and examined by phylogenetic analysis. The sequences cover the years from 1918 to 2009 and their accession numbers are reported in the supplementary Table 1. Three different data set were built up for HA, NA, and MP gene segments. Multiple sequence alignments of HA, NA, and MP segments were performed using CLUSTAL X (THOMPSON *et al.*, 1994) and manually edited with the Bioedit software (Hall, 1999). Positions containing gaps were removed from the final alignment. For our data set, the best fitting nucleotide substitution model was tested with a hierarchical likelihood ratio test following the strategy described by Swofford and Sullivan 2003 (Swofford and Sullivan, 2003), using a neighbor-joining (NJ) base-tree with LogDet corrected distances. Maximum likelihood (ML) trees were then inferred with the selected model and ML estimated substitution parameters. The heuristic search for the ML tree was performed using an NJ tree as starting tree and the TBR branch-swapping algorithm. NJ trees were also estimated using pairwise distances inferred by ML with the best fitting nucleotide substitution model. Calculations were performed with PAUP\* 4.0b10 according to Swofford and Sullivan, 2003. Statistical support for internal branches in the NJ trees was obtained by bootstrapping (1,000 replicates) and with the ML-based zero-branch-length test for the ML trees (Swofford and Sullivan, 2003). The phylogenetic trees were unrooted.

Based on the phylogenetic analysis of the HA gene, the viruses with H1 HA were divided into three main clades, Figure 1. The first clade contained American wild birds, Euro-Asiatic domestic birds, and European swine H1N1 isolates. Within the first clade there are three clusters; the first cluster contains American wild birds, American domestic birds (Duck/NY/13152-13/1994, Turkey/MO/21939/1987), and a Canadian domestic bird (Duck/Alberta/35/1976) (black triangle); the second cluster presents an intermixing between Australian (Duck/Australia/749/1980), Asian (Duck/Hokkaido/55/1996, Duck/Hong

Kong/717/1979, Duck/Miyagi/66/1977) and European (Duck/Italy/69238/2007, Duck/Italy/281904/2006) domestic birds; the third cluster contains all European H1N1 swine isolates. Interestingly, the second and third clusters (domestic birds and swine isolates) were more closely related compared with the first one (wild birds). So, the intermixing between Australian, European and Asiatic domestic birds and swine isolates suggests that the bird H1N1 isolates as well as swine play a role in the influenza virus evolution. Clade two comprises the novel H1N1 viruses (white triangle) and most American classical swine isolates as well as some intermixing with Asiatic swine isolates (Swine/Hokkaido/2/1981, Swine/Chonburi/NIAH9469/2004, Swine/Ratchaburi/NIAH1481/2000, Swine/Ratchaburi/NIAH550/2003, Swine/Korea/CAS08/2005, Swine/Korea/CAN01/2004) from Japan, Thailand, and Korea, respectively. This result also indicates that the Asiatic swine isolates as well as the American swine isolates contribute to the evolution of H1 of the novel H1N1 viruses. Interestingly, the American domestic bird (Turkey/NC/19762/1988) clustered with the novel H1N1 isolates suggesting that the American domestic birds have a role in the evolution of the H1 gene of the novel swine originated influenza H1N1 viruses (Figure 1). The third clade contains all classical H1N1 human isolates (gray triangle) as well as an American wild bird (Quail/Nanchang/12-340/2000), two Asian swine isolates (Swine/Henan/01/2006, Swine/Tianjin/01/2004), and one American swine isolate (Swine/Cambridge/1939). Also the 1918 pandemic virus (South Carolina/1/1918) fell in this group.

The phylogenetic tree generated by the analysis of the N1 NA gene presented three main clusters, Fig.2. Within the first clade, two clusters were observed. The first contains most of the American H1N1 viruses isolated from wild birds (black triangle), while the second one contains Euro-Asiatic swine isolates, Euro-Asiatic domestic birds (Duck/Miyagi/66/77, Duck/Hokkaido/55/1996, Duck/Italy/281904/2006, Duck/Italy/69238/2007, Goose/Italy/296426/2003) and all novel H1N1 influenza viruses (white triangle). This supports also a role for domestic birds in the evolution of novel influenza A (H1N1) viruses beside the swine isolates. The second clade contains all classical H1N1 human isolates (gray triangle)

as well as an American wild bird (Quail/Nanchang/12-340/2000), two Asian swine isolates (Swine/Henan/01/2006, Swine/Tianjin/01/2004). The third clade contains the American influenza H1N1 viruses isolated in swine, two Asian H1N1 isolates from swine (Swine/Korea/CAN01/2004, Swine/Korea/CAS08/2005), and three American isolates from domestic birds (Turkey/NC/19762/1988, Turkey/IA/21089-3/1992, Turkey/MO/21939/1987) (Figure 2). The presence of swine and turkey viral isolates within the same group shows that an NA gene intermixing may have occurred between American domestic birds and swine H1N1 influenza isolates.

Phylogenetic analysis of the MP gene revealed three main clades, Figure 3. The first clade divided into two clusters. The first contained most of the American H1N1 viruses isolated from wild birds (black triangle), while the second contained all novel H1N1 influenza isolates (white triangle) beside the Euro-Asiatic swine and domestic bird viruses (Duck/Italy/281904/2006, Duck/Italy/69238/2007, Goose/Italy/296426/2003, Duck/Miyagi/66/77, Duck/Hokkaido/55/1996). In addition, an Australian domestic bird isolate (Duck/Australia/749/1980) fell in this second cluster. The second clade contains all classical H1N1 human isolates (gray triangle) as well as an



FIGURE 1 - Phylogenetic analysis of the haemagglutinin (HA) gene (1701 nt) of the H1N1 influenza A viruses was carried out using HKY+ I + C (alpha-parameter = 0.0763) as the best evolutionary model. Branch lengths were estimated with the best fitting nucleotide substitution model according to a hierarchical likelihood ratio test [12], and were drawn in scale with the bar at the bottom indicating 0.08 nucleotide substitutions per site. One \* along a branch represents significant statistical support for the clade subtending that branch ( $P < 0.001$  in the zero-branch-length test) and bootstrap support  $> 95\%$ . The tree was unrooted. Three main clades are presented and indicated as I, II, and III. Clade I is divided into three clusters (I, II, III). All H1N1 isolated from 2000 to 2008 are presented as a gray triangle, all novel 2009 H1N1 as a white triangle, and the wild birds as a black triangle.

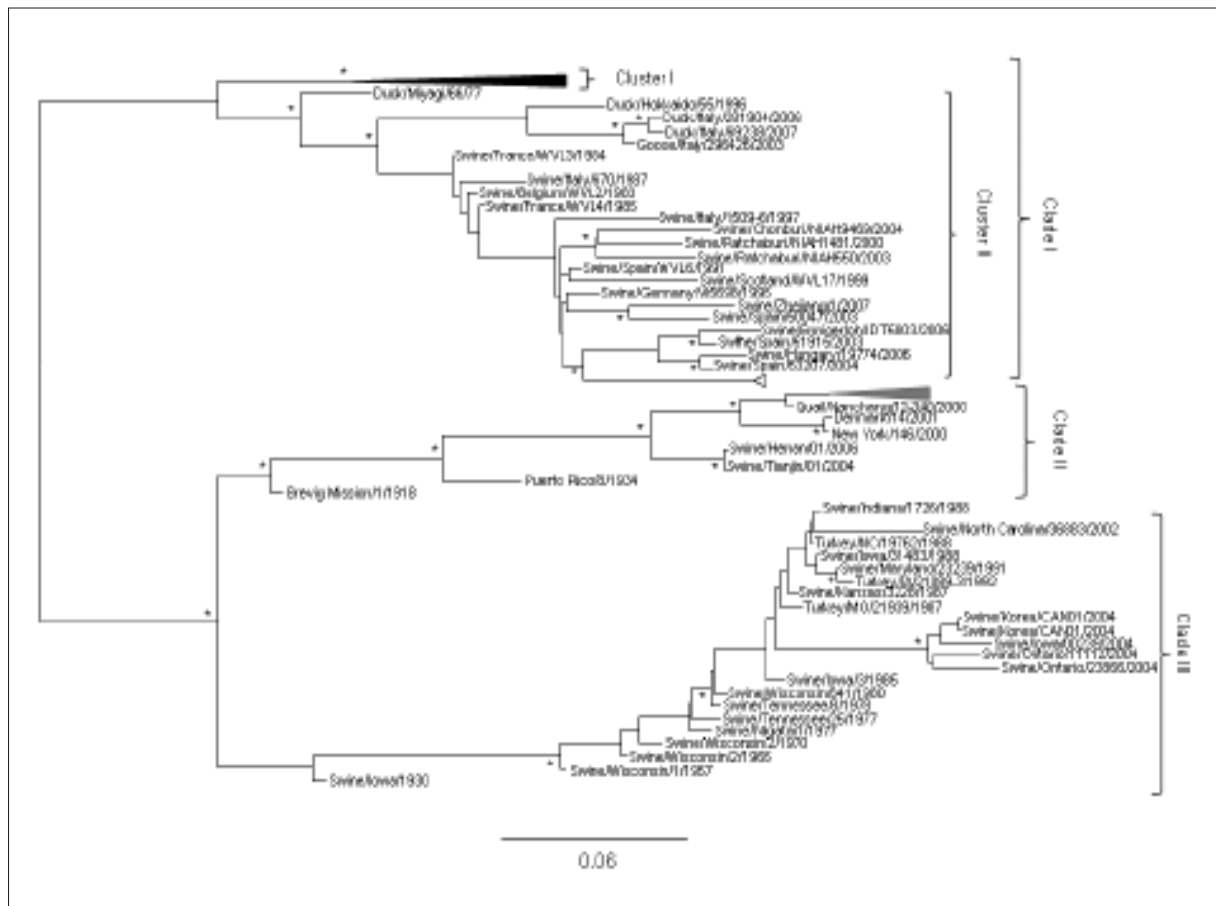


FIGURE 2 - Phylogenetic analysis of the Neuraminidase (NA) gene (1413 nt) of the H1N1 influenza A viruses was carried out using HKY+ I + C (alpha-parameter = 0.0763) as the best evolutionary model. Branch lengths were estimated with the best fitting nucleotide substitution model according to a hierarchical likelihood ratio test [12], and were drawn in scale with the bar at the bottom indicating 0.06 nucleotide substitutions per site. One \* along a branch represents significant statistical support for the clade subtending that branch ( $P < 0.001$  in the zero-branch-length test) and bootstrap support  $> 95\%$ . The tree was unrooted. Three main clades are indicated as I, II, and III. Clade I is divided in two clusters (I, II). All H1N1 isolated from 2000 to 2008 are presented as a gray triangle, the novel 2009 H1N1 as a white triangle, and the wild birds as a black triangle.

American wild bird (Quail/Nanchang/12-340/2000), two Asian swine (Swine/Henan/01/2006, Swine/Tianjin/01/2004). The third clade contains swine American H1N1 isolates, two Asian H1N1 swine isolates (Swine/Korea/CAN01/2004, Swine/Korea/CAS08/2005) and two American domestic bird isolates (Turkey/IA/21089-3/1992, Turkey/NC/19762/1988) (Figure 3). Therefore, an intermixing at the MP gene level occurred between the American domestic birds and swine H1N1 influenza isolates.

To our knowledge, this is the first report characterizing the evolution of the H1N1 genome in domestic and wild birds, humans and swine world-

wide from 1918 to 2009. A better knowledge of the circulating strains may provide useful information for predicting the following season strains and to make a more efficient vaccine.

The present analysis examined whether domestic birds could act as intermediate hosts for the transmission of influenza viruses. Multiple subtypes of avian influenza viruses were distributed among domestic birds in Asia (Qiu *et al.*, 2008) and swine-domestic bird interspecies transmission has been reported in the United States (Wright *et al.*, 1992). Our analysis suggests that domestic birds populations may act as special mixing vessels for influenza A (H1N1) viruses.

Because the 1918 pandemic virus was of avian origin (Garten *et al.*, 2009) and caused millions of deaths worldwide, it would be important to strengthen the surveillance and prevention of avian influenza viruses in domestic birds. Moreover, it may be important to study the evolution and the phylogeography of influenza virus types over time (Babakir-Mina *et al.*, 2009).

## REFERENCES

- BABAKIR-MINA M., DIMONTE S., PERNO C.F., CIOTTI M. (2009). Origin of the 2009 Mexico influenza virus: a comparative phylogenetic analysis of the principle external antigens and matrix protein. *Arch. Virol.* **154** (8), 1349-1352.
- BABAKIR-MINA, M., CICOZZI, M., CIOTTI, M., MARCUCCILLI, F., BALESTRA, E., DIMONTE, S., PERNO, C.F., AQUARO, S. (2009). Epidemiological features and phylogenetic analysis of the surface proteins of influenza A (H5N1) viruses isolated in asian and african populations. *New Microbiologica*. (In press).
- BROWN IH. (2000). The epidemiology and evolution of influenza viruses in pigs. *Vet. Microbiol.* **74** 1-2), 29-46.
- BUTLER D. (2009). Swine flu goes global. *Nature.* **458** (7242), 1082-1083.
- CENTERS FOR DISEASE CONTROL AND PREVENTION (CDC). (2009). Swine influenza A (H1N1) infection in two children - southern California, March-April 2009. *MMWR Morb. Mortal. Wkly. Rep.* **58** (15), 400-402.
- CENTERS FOR DISEASE CONTROL AND PREVENTION (CDC). (2009). Update: swine influenza A (H1N1) infections - California and Texas, April 2009. *MMWR Morb. Mortal. Wkly. Rep.* **58** (16), 1-3.
- CENTERS FOR DISEASE CONTROL AND PREVENTION (CDC). (2009). Outbreak of swine-origin influenza A (H1N1) virus infection. Mexico, March - April 2009. *MMWR Morb. Mortal. Wkly. Rep.* **58** (17), 467-470.
- COHEN J. (2009). Swine flu outbreak. Out of Mexico? Scientists ponder swine flu's origins. *Science.* **324** (5928), 700-702.
- GARTEN R.J., DAVIS C.T., RUSSELL C.A., ET AL. (2009). Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science*. [Epub ahead of print].
- HALL T.A. (1999). Bioedit: a user-friendly biological sequence alignment, editor and analysis program for Windows 95/98 NT. *Nucl. Acids Symp. Ser.* **41**, 95-98.
- ITO T., COUCEIRO J.N., KELM S., ET AL. (1998) Molecular basis for the generation in pigs of influenza A viruses with pandemic potential. *J. Virol.* **72**, 7367-7373.
- DAWOOD F.S. JAIN S., FINELLI L., ET AL. (2009). Novel swine-origin influenza A (H1N1) virus investigation team, emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N. Engl. J. Med.* **360** (25), 2605-2615.
- POSADA D., CRANDALL K.A. (1998). MODELTEST: Testing the model of DNA substitution. *Bioinformatics.* **14**, 817-818.
- QIU B., LIU W., PENG D., HU S., LIU X. (2008). Distribution of avian influenza virus subtypes among domestic ducks in eastern China. *Wei Sheng Wu Xue Bao.* **48** (10), 1290-1294.
- SOLOVYOV A., PALACIOS G., BRIESE T., LIPKIN W.I., RABADAN R. (2009). Cluster analysis of the origins of the new influenza A(H1N1) virus. *Euro Surveill.* **14** (21). pii: 19224.
- SWOFFORD D., SULLIVAN J. (2003). Phylogeny inference based on parsimony and other methods with PAUP\*. In: Salemi M, Vandamme AM, eds. The phylogenetic handbook - A practical approach to DNA and protein phylogeny. New York: Cambridge University Press. 160-206.
- THOMPSON J.D., GIBSON T.J., PLEWNIAC F., JEANMOUGIN F., HIGGINS D.G. (1994). CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876-4882.
- WRIGHT S.M., KAWAOKA Y., SHARP G.B., SENNE D.A., WEBSTER R.G. (1992). Interspecies transmission and reassortment of influenza A viruses in pigs and turkeys in the United States. *Am. J. Epidemiol.* **136** (4), 488-497.

TABLE 1 - Accession numbers of HA, NA and MP nucleotide sequences used in this study.

Strain	Accession Number (HA)	Accession Number (NA)	Accession Number (MP)
A/Amagasaki/1/2009	GQ219574	GQ220730	GQ222023
A/Arizona/01/2009	GQ117067	GQ117064	GQ117066
A/Arizona/02/2009	GQ117079	GQ117077	GQ117078
A/Arizona/13/2007	FJ179358	EU567009	EU716527
A/Arizona/13/2008	FJ686949	FJ686950	N
A/Arkansas/01/2007	EU199356	EU199387	N
A/Auckland/1/2009	FJ973557	FJ973555	FJ973556
A/Auckland/2/2009	N	N	FJ973554
A/Auckland/3/2009	N	FJ973552	FJ973553
A/Auckland/584/2000	CY025026	CY025028	CY025027
A/Auckland/591/2000	CY023010	CY023012	CY023011
A/Auckland/597/2000	CY025034	CY025036	CY025035
A/Beijing/4/2009	GQ232093	GQ232095	GQ232096
A/blue winged teal/LA/B228/1986	EU743306	EU743308	EU743307
A/blue winged teal/TX/43/2002	FJ357061	FJ357063	FJ357063
A/blue-winged teal/Alberta/141/1992	CY004539	CY014568	CY004540
A/Brevig_Mission/1/18	N	AF250356	AY130766
A/California/04/2009	GQ117044	FJ969517	FJ969513
A/California/05/2009	FJ966952	FJ966956	FJ966954
A/California/06/2009	FJ966960	FJ971075	FJ966962
A/California/07/2008	FJ532076	N	N
A/California/07/2009	FJ969540	FJ984386	FJ969527
A/California/08/2009	FJ971076	N	FJ969518
A/California/09/2006	EU199348	EU516124	N
A/California/09/2009	FJ966971	FJ966973	FJ966972
A/California/10/2009	FJ969511	N	FJ969510
A/California/14/2009	GQ117040	GQ117036	GQ117039
A/California/26/2007	EU716572	EU199388	N
A/California/UR06-0232/2007	CY026355	CY026357	CY026356
A/California/UR06-0393/2007	CY026539	CY026541	CY026540
A/Canada-AB/RV1532/2009	GQ132146	GQ132156	GQ132151
A/Canada-NS/RV1536/2009	GQ132147	GQ132158	GQ132152
A/Canada-NS/RV1538/2009	GQ132142	GQ132154	GQ132148
A/Canada-ON/RV1526/2009	GQ132143	GQ132159	GQ132153
A/Canada-ON/RV1529/2009	GQ132144	GQ132157	GQ132149
A/Castilla-La Mancha/GP13/2009	FJ985753	FJ985754	FJ985750
A/Castilla-La Mancha/GP9/2009	FJ985768	FJ985769	FJ985765
A/Catalonia/P148/2009	GQ122099	GQ122100	N
A/Catalonia/P154/2009	GQ122102	GQ122101	N
A/Catalonia/P48/2009	N	GQ132186	N
A/Colorado/01/2008	EU716526	EU716568	N
A/Colorado/03/2009	GQ117119	GQ117118	N
A/Colorado/21/2007	EU516018	EU516273	EU716528
A/Colorado/UR06-0053/2007	CY026523	CY026525	CY026524
A/Denmark/05/2008	FJ264941	FJ264976	FJ265003
A/Denmark/122/2008	FJ264950	FJ264981	FJ265014
A/Denmark/14/2001	EU097937	EU097709	N
A/Denmark/16/2001	EU097938	EU097710	N
A/Denmark/226/2005	EU097947	EU097721	N
A/Denmark/3/2001	EU097934	EU097713	N
A/Denmark/33/2005	EU097949	EU097728	N
A/Denmark/40/2001	EU097933	EU097716	EU097878
A/Denmark/47/2006	EU097956	EU097731	EU097889
A/Denmark/48/2006	EU097957	EU097735	N
A/Denmark/513/2009	FJ982430	FJ982431	FJ982432
A/Denmark/52/2008	FJ264949	FJ264980	FJ265013

*follow*

TABLE 1 - Accession numbers of HA, NA and MP nucleotide sequences used in this study.

Strain	Accession Number (HA)	Accession Number (NA)	Accession Number (MP)
A/duck/Alberta/35/76	D10477	N	N
A/duck/Australia/749/1980	CY014627	N	CY014628
A/duck/Australia/749/80	AF091312	N	N
A/duck/Bavaria/1/77	AF091313	N	N
A/duck/Hokkaido/55/96	AB271115	AB271116	N
A/duck/Hong Kong/717/1979	EF679199	N	N
A/duck/Italy/281904/2006	FJ432770	FJ432772	FJ432771
A/duck/Italy/69238/2007	FJ432754	FJ432756	FJ432755
A/duck/Miyagi/66/77	AB271113	AB271114	N
A/duck/NJ/7717-70/1995	EU026110	N	N
A/duck/NY/13152-13/1994	EU026102	N	N
A/duck/WI/1938/80	L25071	N	N
A/Florida/18/2008	FJ686987	FJ686986	N
A/Florida/19/2007	EU716621	EU716535	N
A/Florida/3/2006	EU516297	EU100633	N
A/goose/Italy/296426/2003	FJ432778	FJ432780	FJ432779
A/Indiana/04/2008	EU779625	EU779649	N
A/Indiana/09/2009	GQ117097	GQ117094	GQ117096
A/Israel/644/2009	GQ166760	GQ166759	GQ166758
A/Italy/06/2009	GQ149765	N	N
A/Kansas/02/2009	GQ117059	GQ117058	N
A/Kansas/03/2009	FJ969523	N	N
A/Korea/01/2009	GQ131023	GQ132185	GQ131025
A/Lisboa/26/2009	GQ166752	GQ166750	GQ166751
A/mallard duck/ALB/322/1988	CY004531	CY004533	CY004532
A/mallard duck/ALB/46/1977	CY004466	CY004467	CY004468
A/mallard duck/ALB/46/1977	N	CY004468	N
A/mallard/ALB/201/1998	CY004507	CY004509	CY004508
A/mallard/ALB/267/1996	CY004504	N	N
A/mallard/Alberta/35/1976	AF091309	N	N
A/mallard/Maryland/350/2002	CY032205	CY032207	CY032206
A/mallard/MD/02-184/2002	EU743576	EU743578	EU743577
A/mallard/MD/199/2003	EU743576	EU743578	EU743577
A/mallard/MD/369/2002	EU026058	N	N
A/mallard/MD/42/2003	FJ357128	FJ357130	FJ357129
A/mallard/MD/53/2003	FJ357136	FJ357138	FJ357137
A/mallard/Ohio/118/1993	CY018885	CY018887	CY018886
A/mallard/Ohio/171/1990	CY012800	CY012802	CY012801
A/mallard/Ohio/56/1999	CY012824	CY012826	CY012825
A/mallard/Tennessee/11464/85	AF091311	N	N
A/Massachusetts/06/2009	GQ117043	GQ117042	N
A/Massachusetts/07/2009	GQ117103	GQ117102	N
A/Mexico/4108/2009	GQ149651	GQ149656	GQ162179
A/Mexico/4115/2009	GQ149692	GQ149691	GQ149625
A/Mexico/4482/2009	GQ149671	GQ149672	GQ149669
A/Mexico/4486/2009	GQ149623	GQ149644	GQ149642
A/Mexico/4603/2009	GQ149630	GQ149640	GQ149629
A/Mexico/4604/2009	GQ149634	GQ149631	GQ149638
A/Mexico/InDRE4114/2009	GQ132145	GQ132155	GQ132150
A/Michigan/02/2009	N	GQ117108	GQ117111
A/Michigan/5/2003	CY008524	CY008526	CY008525
A/Minnesota/02/2009	N	GQ117071	N
A/muscovy duck/New York/21211-5/2005	CY029841	CY029843	CY029842
A/Nebraska/02/2009	N	GQ117105	N
A/Netherlands/602/2009	CY039527	CY039528	N
A/New York/01/2008	EU716563	N	N
A/New York/03/2008	EU779614	N	N
A/New York/05/2008	EU851987	N	N
A/New York/06/2009		FJ984340	FJ984338
A/New York/08/2007	EU199355	EU199384	

*follow*

TABLE 1 - Accession numbers of HA, NA and MP nucleotide sequences used in this study.

Strain	Accession Number (HA)	Accession Number (NA)	Accession Number (MP)
A/New York/10/2009	FJ984375	FJ984371	FJ984369
A/New York/11/2009	FJ984347	FJ984344	FJ984342
A/New York/12/2009	FJ984337	FJ984335	N
A/New York/13/2009	GQ117116	GQ117114	N
A/New York/146/2000	CY000449	CY000451	CY000450
A/New York/15/2009	FJ984379	FJ984378	FJ984376
A/New York/1669/2009	N	CY039895	CY039894
A/New York/1682/2009	N	CY039903	CY039902
A/New York/18/2009	FJ984355	FJ984350	FJ984348
A/New York/19/2009	FJ969509	FJ984390	FJ984388
A/New York/20/2009	FJ969542	FJ969541	GQ117085
A/New York/22/2009	GQ117024	GQ117022	GQ117023
A/New York/220/2002	CY002528	CY002530	CY002529
A/New York/23/2009	FJ984364	FJ984362	N
A/New York/26/2006	EU199333	N	N
A/New York/302/2001	CY003312	CY003314	CY003313
A/New York/31/2009	FJ984360	FJ984357	
A/New York/444/2001	CY003288	CY003290	CY003289
A/New York/483/2003	CY008524	CY008526	CY008525
A/New York/493/2003	CY006667	CY006669	CY006668
A/New York/UR06-0199/2007	CY026627	CY026629	CY026628
A/New York/UR06-0386/2007	CY025285	CY025287	CY025286
A/Nonthaburi/102/2009	N	GQ150343	GQ150341
A/northern pintail/Interior Alaska/1/2007	CY039747	N	N
A/Ohio/04/2006	EU516024	EU516133	EU516161
A/Ohio/07/2009	FJ969521	FJ969520	FJ984398
A/Ohio/UR06-0100/2007	CY026003	CY026005	CY026004
A/Ohio/UR06-0296/2007	CY030061	CY030063	CY030062
A/Osaka/1/2009	GQ219578	GQ220734	GQ222028
A/Pais Vasco/GP20/2009	FJ985763	FJ985764	FJ985760
A/pintail duck/ALB/219/1977	CY004474	CY004476	CY004475
A/pintail duck/Alberta/210/2002	CY004546	CY004548	CY004547
A/Puerto Rico/8/34	CY033577	CY033579	CY033578
A/Quail/Nanchang/12-340/2000	AY180460	CY005429	CY005428
A/redhead duck/Minnesota/Sg-00123/2007	CY035250	N	N
A/Regensburg/Germany/01/2009	FJ974021	FJ984953	FJ970928
A/Shandong/1/2009	GQ200287	GQ200288	GQ200293
A/South Carolina/09/2009	GQ117056	GQ117053	GQ117055
A/South Carolina/1/18	AF117241		
A/Stockholm/28/2009	GQ122105	GQ122104	GQ122103
A/swine/Belgium/WVL2/1983	CY037967	CY037969	CY037970
A/swine/Cambridge/1939	U04859	N	N
A/swine/Chonburi/NIAH9469/2004	AB434304	AB434306	AB434307
A/swine/Ennigerloh/IDT5803/2006	GQ161139	GQ161140	
A/swine/France/WVL3/1984	CY037975	CY037977	CY037978
A/swine/France/WVL4/1985	CY037983	CY037985	CY037986
A/swine/Germany/Vi5698/95	AM920728	AM920729	AM920731
A/swine/Haseluenne/IDT2617/03	EU163946	N	N
A/swine/Henan/01/06	EU004452	EU004450	EU004449
A/swine/Hokkaido/2/81	AF091306	N	N
A/swine/Hungary/19774/2006	FJ798777	FJ798780	FJ798779
A/swine/IDT/Re230/92hp	EU163947	N	N
A/swine/Indiana/1726/1988	CY039925	CY039927	CY039926
A/swine/Iowa/00239/2004	EU139832	EU139842	
A/swine/Iowa/15/30	EU139823	EU139833	M33045
A/swine/Iowa/3/1985	CY022325	CY022327	CY022326
A/swine/Iowa/31483/1988	CY022970	CY022972	CY022971
A/swine/Italy/1509-6/97	AJ344017	AJ410880	
A/swine/Italy/1511/98	AJ344016	N	N
A/swine/Italy/670/1987	CY025253	CY025255	CY025254 follow



TABLE 1 - Accession numbers of HA, NA and MP nucleotide sequences used in this study.

Strain	Accession Number (HA)	Accession Number (NA)	Accession Number (MP)
A/swine/Kansas/3228/1987	CY022469	CY022471	CY022470
A/swine/Korea/CAN01/2004	EU798778	EU798818	EU798798
A/swine/Korea/CAS08/2005	EU798779	EU798819	EU798799
A/swine/Maryland/23239/1991	CY022477	CY022479	CY022478
A/Swine/Nebraska/1/92	L09063	N	N
A/swine/Netherlands/12/85	AF091317	N	M63525
A/swine/Netherlands/3/80	AF091314	N	N
A/swine/Niigata/1/1977	AB434400	AB434402	AB434403
A/swine/North Carolina/36883/2002	EU139829	EU139839	N
A/swine/OH/511445/2007	EU604689	N	N
A/swine/Ontario/11112/04	DQ280250	DQ280251	DQ280252
A/swine/Ontario/23866/04	DQ280243	DQ280242	DQ280244
A/swine/Ratchaburi/NIAH1481/2000	AB434288	AB434290	AB434291
A/swine/Ratchaburi/NIAH550/2003	AB434296	AB434298	AB434299
A/swine/Scotland/WVL17/1999	CY037960	CY037962	CY037963
A/Swine/Spain/50047/2003	CY009892	CY009894	CY009893
A/swine/Spain/51915/2003	CY010572	CY010574	CY010573
A/swine/Spain/53207/2004	CY010580	CY010582	CY010581
A/swine/Spain/WVL6/1991	CY037999	CY038001	CY038002
A/swine/Tennessee/25/1977	CY009916	CY009918	CY009917
A/swine/Tennessee/8/1978	CY027523	CY027525	CY027524
A/swine/Tianjin/01/04	EU004444	EU004442	EU004440
A/swine/Wisconsin/1/1957	CY026283	CY026285	CY026284
A/swine/Wisconsin/2/1966	CY026299	CY026301	CY026300
A/swine/Wisconsin/2/1970	CY022437	CY022439	CY022438
A/swine/Wisconsin/641/1980	CY022445	CY022447	CY022446
A/Swine/Wisconsin/125/97	AF222026	N	N
A/Swine/Wisconsin/136/97	AF222027	N	N
A/Swine/Wisconsin/166/97	AF222030	N	N
A/Swine/Wisconsin/458/98	AF222035	N	N
A/Swine/Wisconsin/464/98	AF222036	N	N
A/swine/Zhejiang/1/2007	FJ415610	FJ415611	FJ415612
A/Texas/04/2009	FJ981612	FJ966981	J966980
A/Texas/05/2007	EU199328	EU516025	
A/Texas/05/2009	FJ966959	FJ966969	FJ981608
A/Texas/06/2007	EU199338	N	N
A/Texas/06/2009	FJ984385	FJ984383	FJ984381
A/Texas/07/2009	GQ117091	N	GQ117090
A/Texas/08/2009	GQ117051	GQ117048	GQ117050
A/Texas/09/2008	FJ532088	N	N
A/Texas/09/2009	GQ117032	GQ117028	GQ117031
A/Texas/11/2008	FJ532089	N	N
A/Texas/15/2008	FJ549051	FJ549050	N
A/Texas/29/2008	FJ686988	FJ686989	N
A/Texas/UR06-0012/2006	CY025213	CY025215	CY025214
A/Texas/UR06-0133/2007	CY027659	CY027661	CY027660
A/Toronto/3141/2009	FJ974026	N	N
A/Toronto/3145/2009	FJ974027	N	N
A/Toronto/3146/2009	FJ974028	N	N
A/Toronto/3178/2009	FJ974025	N	N
A/Toronto/3181/2009	FJ974022	N	N
A/Toronto/3184/2009	FJ974024	N	N
A/turkey/IA/21089-3/1992	EU743159	EU743161	EU743160
A/turkey/Minnesota/1661/81	AF091310	N	N
A/turkey/MO/21939/1987	EU743143	EU743145	EU743144
A/turkey/NC/19762/1988	EU735786	EU735788	EU735787
A/Valencia/GP4/2009	FJ985758	FJ985759	FJ985755
A/Waikato/11/2005	CY016459	CY016461	CY016460
A/Wellington/13/2005	CY013589	CY013590	CY013591
A/Wellington/14/2000	CY016196	CY016198	CY016197

N\*: Sequence not available.



FIGURE 3 - Phylogenetic analysis of the matrix protein (MP) gene (942 nt) of the H1N1 influenza A viruses was carried out using HKY+ I + C (alpha-parameter = 0.0763) as the best evolutionary model. Branch lengths were estimated with the best fitting nucleotide substitution model according to a hierarchical likelihood ratio test [12], and were drawn in scale with the bar at the bottom indicating 0.03 nucleotide substitutions per site. One \* along a branch represents significant statistical support for the clade subtending that branch ( $P < 0.001$  in the zero-branch-length test) and bootstrap support  $> 95\%$ . The tree was unrooted. Three main clades are indicated as I, II, and III. Clade I is divided into two clusters (I, II). All H1N1 isolated from 2000 to 2008 are presented as a gray triangle, all novel 2009 H1N1 as a white triangle, and the wild birds as a black triangle.