Chinese Science Bulletin

Molecular characterization of H1N1 influenza A viruses from human cases in North America

WU Bin[1,3*], WANG ChengMin[1,2*], DONG GuoYing[1], LUO Jing[1,2], ZHAO BaoHua[3] & HE HongXuan[1†]

[1] National Research Center For Wildlife Born Diseases, Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China;
[2] Graduate University of Chinese Academy of Sciences, Beijing 100049, China;
[3] College of Life Science, Hebei Normal University, Shijiazhuang 050016, China

Subtypes of H1N1 influenza virus can be found in humans in North America, while they are also associated with the infection of swine. Characterization of the genotypes of viral strains in human populations is important to understand the source and distribution of viral strains. Genomic and protein sequences of 10 isolates of the 2009 outbreak of influenza A (H1N1) virus in North America were obtained from GenBank database. To characterize the genotypes of these viruses, phylogenetic trees of genes *PB2*, *PB1*, *PA*, *HA*, *NP*, *NA*, NS and *M* were constructed by Phylip3.67 program and *N*-Linked glycosylation sites of HA, NA, PB2, NS1 and M2 proteins were analyzed online by NetNGlyc1.0 program. Phylogenetic analysis indicated that these isolates are virtually identical but may be recombinant viruses because their genomic fragments come from different viruses. The isolates also contain a characteristic lowly pathogenic amino acid motif at their HA cleavage sites (IPSIQSR↓GL), and an E residue at position 627 of the PB2 protein which shows its high affinity to humans. The homologous model of M proteins showed that the viruses had obtained the ability of anti-amantadine due to the mutation at the drug-sensitive site, while sequence analysis of NA proteins indicated that the viruses are still susceptible to the neuraminidase inhibitor drug (i.e. oseltamivir and zanamivir) because no mutations have been observed. Our results strongly suggested that the viruses responsible for the 2009 outbreaks of influenza A (H1N1) virus have the ability to cross species barriers to infect human and mammalian animals based on molecular analysis. These findings may further facilitate the therapy and prevention of possible transmission from North America to other countries.

influenza A viruses, H1N1 subtype, molecular characterization, North America

Influenza A viruses are clinically and/or economically important pathogens in a variety of hosts, including humans, swine, marine mammals, and avian[1]. They are classified into different subtypes on the basis of antigenic properties in the two surface glycoproteins, haemagglutinin (HA) and neuraminidase (NA). There are limited numbers of hosts specific influenza A viruses, i.e., humans (H1N1, H3N2, H2N2, and H1N2) and swine (H1N1, H3N2 and H1N2)[2]. Depending on their ability to cause disease in chickens, influenza A viruses are characterized as lowly pathogenic (LP) or highly pathogenic (HP). Subtypes H1, H2 and H3 influenza A viruses are usually lowly pathogenic to animals and humans but may become highly pathogenic to animals including humans through mutations of multiple basic amino acids in the connecting peptide between HA1 and HA2 domains of the HA0 precursor protein.

Influenza A virus H1N1 subtype, also known as A

Received May 7, 2009; accepted May 22, 2009
doi: 10.1007/s11434-009-0421-y
* Contributed equally to this work
† Corresponding author (email: hehx@ioz.ac.cn)
Supported by the National Key Basic Research and Development Program of China (Grant No. 2007BC109103), Knowledge Innovation Project of the Chinese Academy of Sciences (Grant No. KSCX2-YW-N-063) and National Natural Science Foundation of China (Grant No. 30671576)

Citation: Wu B, Wang C M, Dong G Y, et al. Molecular characterization of H1N1 influenza A viruses from human cases in North America. Chinese Sci Bull, 2009, 54: 2179-2192, doi: 10.1007/s11434-009-0421-y

(H1N1), is the most common cause of influenza (flu) in humans and cause several endemics in humans. The 1918 flu pandemic, which killed 50－100 million people worldwide, is thought to be one of the most deadly pandemics in human history. It was caused by the H1N1 type of influenza virus[1]. The more recent Russian flu, 1977－1978 flu epidemics, was caused by strain A/USSR/90/77 (H1N1). It infected mostly children and young adults under 23 because a similar strain prevalent in 1947－1957 caused most adults to have substantial immunity. Though some called it a flu pandemic, because it only affected the young, it is not considered a true pandemic[2–5]. Minor outbreaks of swine influenza occurred in humans in 1976 and 1988 and in swine in 1998 and 2007. Less virulent H1N1 strains still exist in the wild today, worldwide, causing a small fraction of all influenza-like illness and a large fraction of all seasonal influenza. H1N1 strains caused roughly half of all flu infections in 2006[6]. Other strains of H1N1 are endemic in pigs and birds.

In the 2009 outbreak of influenza A (H1N1) virus, the virus isolated from patients in the United States was found to be made up of genetic elements from 4 different flu viruses－North American Mexican influenza, North American avian influenza, human influenza, and swine influenza virus typically found in Asia and Europe－"an unusually mongrelized mix of genetic sequences"[7]. This new strain appears to be a result of reasserting 4 different H1N1 strains of human influenza and swine influenza viruses. World Organization for Animal Health (OIE) suggests it be called "North-American influenza"[8], and later began referring to it as "Influenza A (H1N1)[9]". Preliminary genetic characterization of the Influenza A (H1N1) isolates showed that the hemagglutinin (*HA*) gene was similar to swine flu viruses present in U.S. pigs since 1999, but neuraminidase (*NA*) and matrix protein (*M*) genes resembled versions present in European swine flu isolates. The 6 genes from American swine flu are themselves mixture of swine flu, bird flu and human flu viruses[10,11]. Nevertheless complete molecular characterizations of viruses remain unknown.

In this study, we collected ten complete genome sequences for U.S. flu cases in 2009 through the Global Initiative on Sharing Avian Influenza Data, and analyzed their phylogenetic relationships based on genes *PB2*, *PB1*, *PA*, *HA*, *NP*, *NA*, *NS* and *M*. Pathogenic and drug esistance were predicted by comparison and analysis of potential *N*-Linked glycosylation sites of HA, NA, PB2, NS1 and M2 proteins. It will be of important significance for human to prevent and treat H1N1 influenza A virus.

1 Materials and methods

1.1 Sequence data

We collected the full genome of 10 H1N1 influenza virus isolates in North America and the reference sequences from GenBank (Table 1). These sequences were analyzed together with all publicly available sequence data, including genes from H1, H2 and H3 viruses. In addition, sequences of HA, NA, NS1, PB2 and M2 proteins were also obtained from GenBank (Table 2).

1.2 Construction of phylogenetic trees

Nucleotide BLASTn analysis (http://www.ncbi.nlm.nih.

1) Anonymous. 1918 Influenza A (H1N1) fact sheet. The Federation of American Scientists. http://www.fas.org/programs/ssp/bio/fact- sheets/H1N1factsheet.html.

2) Anonymous. Interactive health timeline box 1977: Russian flu scare, CNN.

3) Anonymous. Invasion from the Steppes. Time, 1978-2-20. http://www.time.com/time/magazine/article/0,9171,948035,00.html

4) Anonymous. Pandemic Influenza subsection Recent Pandemic Flu Scares. Global Security, http://www.globalsecurity.org/security/ops/ hsc-scen-3_pandemic-influenza.htm

5) Anonymous. Russian flu confirmed in Alaska. Epidemiology, 1978-04-21. http://www.epi.hss.state.ak.us/bulletins/docs/b1978_09. htm

6) CDC, Flu view. A weekly influenza surveillance report prepared by the influenza division. 2008-2009 Influenza Season Week 18 ending May 9, 2009. http://www.cdc.gov/flu/weekly/.

7) MacKenzie D. Deadly new flu virus in US and Mexico may go pandemic. New Scientist. 2009-04-28. http://www.newscientist.com/ article/dn17025-deadly-new-flu-virus-in-us-and-mexico-may-gopandemic.html.

8) Anonymous. A/H1N1 influenza like human illness in Mexico and the USA: OIE statement. 2009-4. http://www.oie.int/eng/press/ en_090427.htm

9) Renaming of strain to Influenza A (H1N1). WHO. http://www. who.int/csr/disease/swineflu/en/index.html

10) Watts S. Experts concerned about potential flu pandemic. BBC, 2009-04-25. http://www.bbc.co.uk/blogs/newsnight/susanwatts/ 2009/04/experts concerned_about_potent.html.

11) Niman H. at FluTrackers has described the homologies of the genes as PB2 Avian North America, PB1 Human circa 1993, PA Swine Eurasia and/or North America, HA Swine North America, NP Swine Eurasia and/or North America, NA Swine Eurasia, MP Swine Eurasia, NS Swine Eurasia and/or North America.

Table 1 Accession numbers of the nucleotide sequences used in this study

	PB2	PB1	PA	HA	NP	NA	MP	NS
A/Swine/Indiana/9K035/99 (H1N2)	AF250131	AF250130	AF250129	AF250124	AF250127	AF250126	AF250125	AF250128
A/Swine/Iowa/569/99 (H3N2)	AF251426	AF251421	AF251425	AF251419	AF251423	AF251420	AF251422	AF251424
A/Swine/Minnesota/593/99 (H3N2)	AF251434	AF251429	AF251433	AF251427	AF251431	AF251428	AF251430	AF251432
A/swine/Hong Kong/5212/99(H3N2)	AY363467	AY363485	AY363503	AY363521	AY363539	AY363557	AY363575	AY363593
A/swine/Korea/JNS06/2004(H3N2)	EU301177	EU301400	EU301368	EU301209	EU301304	EU301273	EU301241	EU301336
A/Swine/Indiana/P12439/00 (H1N2)	AF455736	AF455728	AF455720	AF455680	AF455704	AF455696	AF455688	AF455712
A/Swine/Minnesota/55551/00 (H1N2)	AF455734	AF455726	AF455718	AF455678	AF455702	AF455694	AF455686	AF455710
A/Swine/Illinois/100084/01 (H1N2)	AF455738	AF455730	AF455722	AF455682	AF455706	AF455698	AF455690	AF455714
A/duck/NC/91347/01(H1N2)	AY233387	AY233388	AY233389	AY233393	AY233394	AY233391	AY233392	AY233390
A/Swine/Minnesota/9088-2/98 (H3N2)	AF153242	AF153246	AF153250	AF153234	AF153254	AF153238	AF153258	AF153262
A/Swine/Texas/4199-2/98 (H3N2)	AF153241	AF153245	AF153249	AF153233	AF153253	AF153237	AF153257	AF153261
A/Wisconsin/10/98 (H1N1)	AF342824	AF342823	AF342822	AF342821	AF342819	AF342820	AF342818	AF342817
A/Hong Kong/1774/99(H3N2)	AJ293920	AJ293921	AJ293922	AJ293926	AJ293924	AJ293923	AJ293925	AJ293941
A/Turkey/MO/24093/99(H1N2)	AY038017	AY038016	AY038018	AY038014	AY038019	AY038015	AY038020	AY038021
A/swine/Korea/CY05/2007(H3N2)	EU798934	EU798914	EU798894	EU798794	EU798854	EU798834	EU798814	EU798874
A/swine/Denmark/WVL9/1993(H1N1)	CY038020	CY038021	CY038022	CY038023	CY038024	CY038025	CY038026	CY038027
A/swine/England/WVL7/1992(H1N1)	CY038004	CY038005	CY038006	CY038007	CY038008	CY038009	CY038010	CY038011
A/swine/Spain/WVL6/1991(H1N1)	CY037996	CY037997	CY037998	CY037999	CY038000	CY038001	CY038002	CY038003
A/swine/Korea/CAS09/2006(H3N2)	EU798932	EU798912	EU798892	EU798792	EU798852	EU798832	EU798812	EU798872
A/Goose/Guangdong/1/96(H5N1)	AF144300	AF144301	AF144302	AF144305	AF144303	AF144304	AF144306	AF144307
A/Brevig_Mission/1/18(H1N1)	DQ208309	DQ208310	DQ208311	AF116575	AY744935	AF250356	AY130766	AF333238
A/Guiyang/1/1957(H2N2)	CY032276	CY032275	CY032274	CY032269	CY032272	CY032271	CY032270	CY032273
A/Hong Kong/1-1/1968(H3N2)	CY033008	CY033007	CY033006	CY033001	CY033004	CY033003	CY033002	CY033005
A/USSR/90/1977(H1N1)	CY010379	CY010378	CY010377	CY010372	CY010375	CY010374	CY010373	CY010376
A/California/06/2009(H1N1)	FJ966963	FJ966965	FJ966964	FJ966960	FJ966961	FJ971075	FJ966962	FJ971074

Table 2 Accession numbers of the amino acid sequences used in this study

	NS1	M2	PB2	HA	NA
A/California/04/2009(H1N1)	ACP41110/ACP44154	ACP41109/ACP44153	ACP41102/ACP44157	ACP41105	ACP41107
A/California/05/2009(H1N1)		ACP41929	ACP41930	ACP41926	
A/California/08/2009(H1N1	ACP44179/ACP44161	ACP44178/ACP44160		ACP52565	
A/California/07/2009(H1N1)	ACP44186/ACP44172	ACP44185/ACP44171/ACP41955	ACP44175/ACP41956	ACP41953	ACQ63272
A/California/09/2009(H1N1)		ACP41951			
A/California/10/2009(H1N1)		ACP44149		ACP44150	
A/Texas/04/2009(H1N1)	ACQ55367	ACP41961/ACP41965/ACQ55364	ACP44168	ACQ55361	ACP41962
A/Texas/05/2009(H1N1)	ACQ55356/ACP41942	ACQ55353		ACP41934	
A/California/06/2009(H1N1)	ACP52562	ACP41938	ACP41939	ACP41935	ACP52564
A/Ohio/07/2009(H1N1)					ACP44181
A/New York/15/2009(H1N1)					ACQ63262
A/New York/19/2009(H1N1)				ACQ63282	
A/Auckland/2/2009(H1N1)		ACQ42239			
A/Auckland/3/2009(H1N1)		ACQ42237			
A/Denmark/184/2005(H1N1					ABW81483
A/Pusan/46/2002(H1N1)					AAP59861
A/South Australia/26/2000(H1N1)					ABK79973
A/Memphis/15/2000(H1N1)					ABP49308
A/Swine/Wisconsin/458/98(H1N1)				AAF87283	
A/swine/Cotes d'Armor/1488/1999(H1N1))				CAC86337	
A/swine/Guangdong/2/2001(H1N1)				AAY56898	
A/swine/Wisconsin/641/1980(H1N1)			ABR28745		
A/swine/France/WVL4/1985(H1N1)			ACO25096		
A/Hong Kong/470/1997(H1N1)			AAF99674		
A/mallard/Alberta/130/2003(H1N1)			ABB19429		
A/chicken/Hubei/wl/1997(H5N1)	ABI96744				
A/chicken/Hubei/wm/1997(H5N1)	ABI96753				
A/chicken/Hong Kong/915/97(H5N1)	AAF02349				
A/Duck/Hong Kong/p46/97(H5N1)	AAF02351				
A/Hong Kong/117/1977(H1N1)		ABD60946			
A/swine/Tennessee/10/1978(H1N1)		ABU80267			
A/mallard/Maryland/358/2002(H1N1)		ABS70359			

gov/BLAST) was used to identify related genes of the viruses, and the reference sequences were obtained from GenBank. Pair-wise sequence alignments were also performed with the MEGA4.0 program[3] (http://www.megasoftware.net/) to determine nucleotide sequence similarities. Alignments of each influenza virus sequence were generated using program ClustalW[4] (http://clustalw.genome.ad.jp/). To understand the evolutionary characterization of H1N1 viruses isolates in North America, phylogenetic analyses of the aligned sequences for 8 gene segments were performed by the neighbor-joining method with 1000 bootstraps and Maximum-Likelihood with 100 bootstraps by using PHYLIP version 3.67 (http://evolution.gs.washington.edu/phylip.html).

1.3 Comparison and analysis of amino acid sequences in HA, NS1, NA, M2 and PB2

We extracted the amino acid sequences of North America H1N1 influenza viruses HA, NS1, NA, M2 and PB2 proteins from the public domain database Entrez Protein, and compared each of them with all known North America H1N1 Influenza viruses proteins (n = 10 as of May 8, 2009) using the software ClusterW[4]. Analysis of *N*-glycosylation sites based on the amino acid sequences was conducted using the NetNGlyc1.0 program (http://www.cbs.dtu.dk/services/NetNGlyc/).

1.4 Molecular models and 3D structural similarity of M2 protein

Based upon comparison and analysis of the amino acid sequences in M2 proteins, a three-dimensional model of the homologous M2 proteins was constructed using the software SWISS-MODEL[5] (available through Web interface at http://swissmodel.expasy.org/SWISS-MODEL.html), and visualized by VMD program (Visual molecular dynamics) (version 1.8.6, which is freely available at http://www.ks.uiuc.edu/Research/vmd/)[6].

2 Results

2.1 Phylogenetic analysis of the North America isolates

To investigate the genetic origin of the North America isolates, the nucleotide sequences of 10 North America isolates and reference strains were obtained from GenBank. The percentage of nucleotide sequence homology among the eight gene segments of the ten North America isolates ranged between 99.9% and 100%, and the predicted amino acid sequence homology is 100%. Thus, their genetic characterizations were virtually identical, suggesting that the 10 isolates originated from the same progenitor virus.

Therefore, we selected the eight gene segments of A/California/06/2009(H1N1) (as a representative strain of 10 isolates) to further conduct genetic distance analysis. The genetic origin of this isolate (A/California/06/2009(H1N1)) was initially inferred from BLAST analysis and pair wise comparisons of each gene segment to the corresponding sequences of reference viruses. In our present studies, the results indicated that the eight gene segments from the isolate share the different sequence homology with other influenza viruses circulating in different regions (Table 3). The isolate shares the highest homology (98%) of the *PB2* gene segment with reference viruses such as A/swine/Korea/JNS06/2004 (H3N2), A/Swine/Minnesota/55551/00 (H1N2) and A/mallard duck/SouthDakota/Sg-00128/2007. For *M* gene, this isolate shares the higher homology (91%) with the human isolate (A/swine/Hong Kong/5212/99 (H3N2)) and the swine isolate (A/Hong Kong/1774/99 (H3N2)) in Hong Kong and the swine isolate (A/swine/ Denmark/WVL9/1993 (H1N1)) in Denmark. Compared to the swine isolates in Europe such as A/Swine/England/195852/92 (H1N1), A/swine/Spain/WVL6/1991 (H1N1) and A/swine/England/WVL7/1992 (H1N1), its homology is only 94%. For North America, this isolate shares the homology (96%) of the *NP* gene segment with A/Swine/Iowa/533/99(H3N2) and A/Swine/Indiana/P12439/00 (H1N2), the homology (96%) of the *PA* gene segment with A/Swine/Illinois/100084/01 (H1N2), A/Swine/Minnesota/593/99 (H3N2) and A/Swine/Iowa/569/99 (H3N2), the homology (96%) of the *PB1* gene segment with A/Wisconsin/10/98 (H1N1) and A/Swine/Indiana/9K035/99 (H1N2), and the homology (96%) of the *NS* gene segment with A/Swine/Minnesota/9088-2/98 (H3N2) and A/Swine/Texas/4199-2/98 (H3N2). According to host species, this isolate shares the homology (96%) of the PB1 gene segment with mallard duck (A/mallard duck/South Dakota/Sg-00125/2007 (H3N2)) and the homology (95%) of the HA gene segment (A/Turkey/MO/24093/99(H1N2).

To further demonstrate the genetic origin of the gene segments of the North America virus isolates, we constructed 8 phylogenetic trees (Figure 1). Reference viruses consisted of H1N1, H1N2, H3N2, H5N1 and H2N2 viruses isolated in poultry, humans and swine

Table 3 Genetic similarity between A/California/06/2009(H1N1) and related reference viruses available in GenBank

	PB2	PB1	PA	HA	NP	NA	MP	NS
A/Swine/Indiana/9K035/99 (H1N2)	0.02193	0.04604	0.06864	0.08377	0.044	0.8326	0.13024	0.04158
A/Swine/Iowa/569/99(H3N2)	0.02401	0.04324	0.04951	2.01479	0.03851	0.8307	0.12681	0.05678
A/Swine/Minnesota/593/99(H3N2)	0.02401	0.04324	0.04951	2.01479	0.04068	0.8307	0.13105	0.04701
A/swine/Hong Kong/5212/99(H3N2)				2.04854	0.20753	0.83529	0.0202	0.23268
A/swine/Korea/JNS06/2004(H3N2)	0.02809	0.052	0.05887	1.99389	0.03969	0.83991		0.0523
A/Swine/Indiana/P12439/00(H1N2)	0.02811	0.04311	0.0497	0.07703	0.03635	0.83166		0.05381
A/Swine/Minnesota/55551/00(H1N2)	0.02805	0.04019	0.04951	0.1674	0.04178	0.83972		0.04156
A/Swine/Illinois/100084/01(H1N2)	0.03215	0.04913	0.05438	0.10128	0.04067	0.83646		0.04428
A/duck/NC/91347/01(H1N2)	0.03226	0.04025	0.0591	0.16235	0.04517	0.82921	0.12329	0.04824
A/Swine/Minnesota/9088-2/98(H3N2)					0.03628	0.83535	0.11997	
A/Swine/Texas/4199-2/98(H3N2))					0.03525	0.83591	0.12338	
A/Wisconsin/10/98(H1N1)		0.04019			0.03628	0.23261	0.12984	0.04835
A/Hong Kong/1774/99(H3N2)	0.21234	0.16816	0.1413	2.05678	0.2061	0.83339	0.02313	0.23268
A/Turkey/MO/24093/99(H1N2)				0.08707	0.03302	0.8326		
A/swine/Korea/CY05/2007(H3N2)	0.03433	0.04611	0.05439	1.98696	0.03411	0.83725	0.12993	0.09701
A/swine/Denmark/WVL9/1993(H1N1)	0.19777	0.16547	0.15758	0.62368	0.20841	0.07015	0.02908	0.223
A/swine/England/WVL7/1992(H1N1)	0.20282	0.15091	0.16934	0.58668	0.20228	0.06133	0.03503	0.22096
A/swine/Spain/WVL6/1991(H1N1)	0.21102	0.15843	0.17496	0.58128	0.21216	0.06211	0.03207	0.22658
A/swine/Korea/CAS09/2006(H3N2)	0.03636	0.05819	0.05417	2.03412	0.03738	0.8395	0.13347	0.04157
A/Goose/Guangdong/1/96(H5N1)	0.17542	0.15617	0.09917	1.15196	0.21039	0.14643	0.08247	0.46003
A/Brevig Mission/1/1918(H1N1)	0.14259	0.21503	0.18062		0.13576	0.19094	0.0955	0.16918
A/Guiyang/1/1957(H2N2)	0.16387	0.1158	0.17658	1.20144	0.1725	0.84947	0.11563	0.19657
A/Hong Kong/1-1/1968(H3N2)	0.17739	0.08414	0.18883	2.02991	0.17306	0.84592	0.11194	0.21397
A/USSR/90/1977(H1N1)	0.17194	0.21133	0.18147	0.53046	0.17302	0.2228	0.10561	0.18994

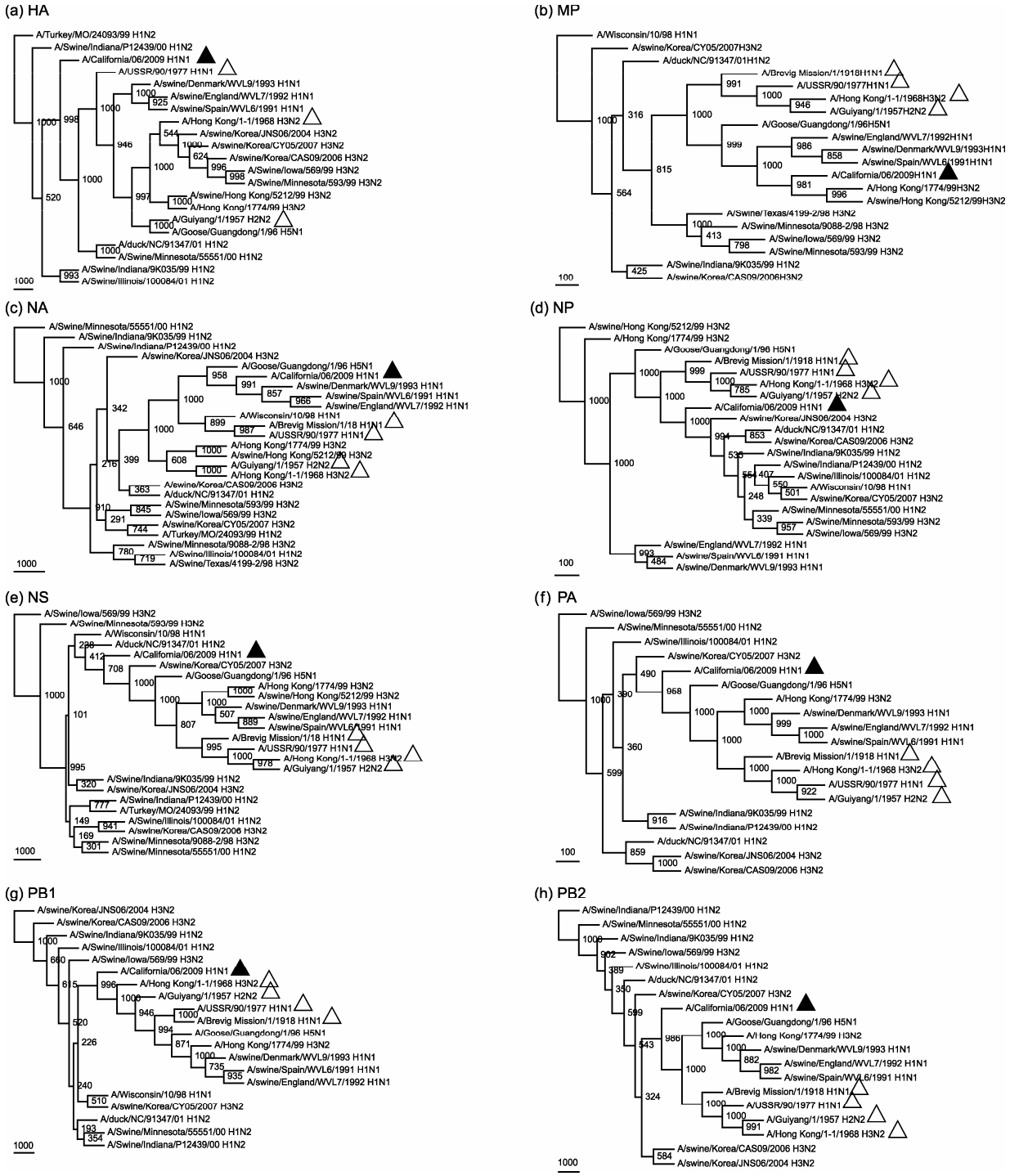

Figure 1 Phylogenetic trees of the nucleotide sequences for the *HA*, *NA*, *NP*, *M*, *NS*, *PB2*, *PB1* and *PA* genes of the North America isolates (A/California/06/2009(H1N1)) and related reference viruses. The evolutionary relationships among these viruses were estimated by the neighbor-joining method with 100 bootstraps by using PHYLIP version 3.67. Alignments of each influenza virus sequence were generated using program Clustal W. The sequence regions compared were as follows: nt 16—1624 (1609 bp) of *HA*, 7—1377 (1371 bp) of *NA*; 17—2242 (2226 bp) of *PB2*, 1—2274 (2274 bp) of *PB1*, 22—2161 (2140 bp) of *PA*, 1—1458 (1458 bp) of *NP*, 9—939 (931 bp) of *M*, 5—802 (798 bp) of *NS*. The tree of *HA* is rooted to A/Turkey/MO/24093/99 (H1N2), the tree of *M* is rooted to A/Wisconsin/10/98 (H1N1), the tree of *NA* is rooted to A/Swine/Minnesota/55551/00 (H1N2), the tree of *NP* is rooted to A/swine/Hong Kong/5212/99 (H3N2), the trees of *NS* and *PA* are rooted to A/Swine/Iowa/569/99 (H3N2), the tree of *PB1* is rooted to A/swine/Korea/JNS06/2004 (H3N2), and the tree of *PB2* is rooted to A/Swine/ Indiana/P12439/00 (H1N2). Black triangles indicate A/California/06/2009 (H1N1) (as a representative of North America isolates), including A/Brevig Mission/1/1918 (H1N1), A/Guiyang/1/1957 (H2N2), A/Hong Kong/1-1/1968 (H3N2), and A/USSR/90/1977 (H1N1).

from 1918 to 2009 (Figure 1). Phylogenetic analysis also indicated the same genetic diversity. As a result, the genetic origin of the North America isolates is complex in several aspects such as geographical distribution (America, Europe and Asia), host species (humans, avian and swine) and genotype (H1N1, H1N2 and H3N2). A larger genetic distance was observed compared to HPAI H5N1 and major epidemic strains that caused 4 pandemics in the 20th century.

2.2 Molecular characterization of the North America isolates

Based on the deduced amino acid sequences, 8 potential N-linked glycosylation sites were found in the HA protein, of which 6 are located in HA1 (positions 27, 28, 40, 104, 293 and 304) and 2 in HA2 (positions 498 and 557) (Figure 2). Compared to the H1N1 virus (A/Brevig_ Mission/1/18(H1N1)) in 1918, the A/California/06/2009 (H1N1) virus isolate exhibited a K295T mutation, which results in the emergence of a potential glycosylation site. This mutation was also found in other isolates from North America in 2009. The A/California/ 06/2009 (H1N1) isolate contained the amino acid motif IPSIQSR/GL at their HA cleavage sites, which is not a characteristic of highly pathogenic influenza viruses.

Several amino acid mutations of the polymerase (PB2) protein may have an important effect on the virulence and adaptation of H1N1 virus in the hosts[17]. Analysis of amino acid sequence in PB2 protein showed clearly that all 10 isolates from North America exhibited the exactly identical amino acid residue (Glu, E) at position 627 (Figure 3). The A/California/06/2009 (H1N1) isolate exhibited exactly the same model of glycosylation sites with the H1N1 virus strain (A/swine/Guangdong/2/ 2001 (H1N1)) in China. All the other 9 virus isolates from North America exhibited the same amino acid residues (Gln, Q) at position 240, but A/California/07/ 2009 (H1N1) isolate exhibited an undetermined amino acid residues (indicates "X")(Figure 3). These results indicated that the amino acid residues (Gln, Q) at position 240 were considered to enhance the binding of H1N1 virus to the sialic acid (SA) α2,6-NeuAcGal linkages of the human cell receptors. Therefore, although influenza virus isolates from North America have obtained the preferential ability to infect humans, yet they are lowly pathogenic for human hosts. This result was consistent with the facts (74 deaths/8829 cases, data from WHO, up to date May 18, 2009).

In addition, NS1 protein plays an important role in the pathogenicity of H1N1 virus in different hosts[8]. Large-scale sequence analysis of avian influenza viruses indicated that the four C-terminal residues of the NS1 protein form a potential PDZ ligand binding motif of the ESEV type[9]. The PDZ ligand binding motifs with the sequence of ESEV or EPEV were found in the NS1 from the highly pathogenic H5N1 viruses isolated in 1997 and 2003 as well as the 1918 pandemic virus (all of avian origin). They are able to bind cellular PDZ-containing proteins involved in host cellular signaling pathways. By contrast, the NS1 proteins in most lowly pathogenic human influenza viruses contain different motifs (RSKV or RSEV), which cannot bind PDZ-containing proteins. A recent study showed that the PDZ-binding motif of NS1 is a new virulence factor of influenza A viruses[10]. But at the C-terminus of NS1 protein of A/California/06/ 2009 (H1N1) virus isolate (Figure 4), we did not find the amino acid motif ESEV or EPEV present in avian highly pathogenic influenza viruses, which was considered to bind to many the avian intracellular receptors[9]. Compared with the nine virus isolates from North America, no mutations were found in the NS1 of A/California/06/ 2009 (H1N1) isolate. Thus, these isolates may show low pathogenicity. This result was also consistent with the facts (74 deaths/8829 cases, data from WHO, up to date May 18, 2009) (http://www.who.int/csr/don/2009_05_ 18/en/index.html).

2.3 Drug resistant characterization of the North America isolates

Antiviral activity is essential for the treatment and prevention of influenza infections. The M2 protein of influenza A virus forms a homotetramer ion channel in the lipid membrane[11]. The protonation of His37 and concomitant cation-pi interaction with Trp41 is a key step in the activation of the M2 ion channel. Trp41 may play a certain role in stabilizing the channel open state through the cation-pi interaction with His37 (Figure 5). A molecular model for the activation of M2 ion channel is proposed on the basis of the gating mechanism[11]. We have studied the structure of the transmembrane domain of the M2 ion channel by using software SWISS-MODEL, with special attention to the side chains of Histidine (His37) and Tryptophan (Trp41) residues. A three-dimensional model of the homologous M proteins was visualized VMD program (visual molecular dynamics). By amino acid sequence analysis of the M2 proteins,

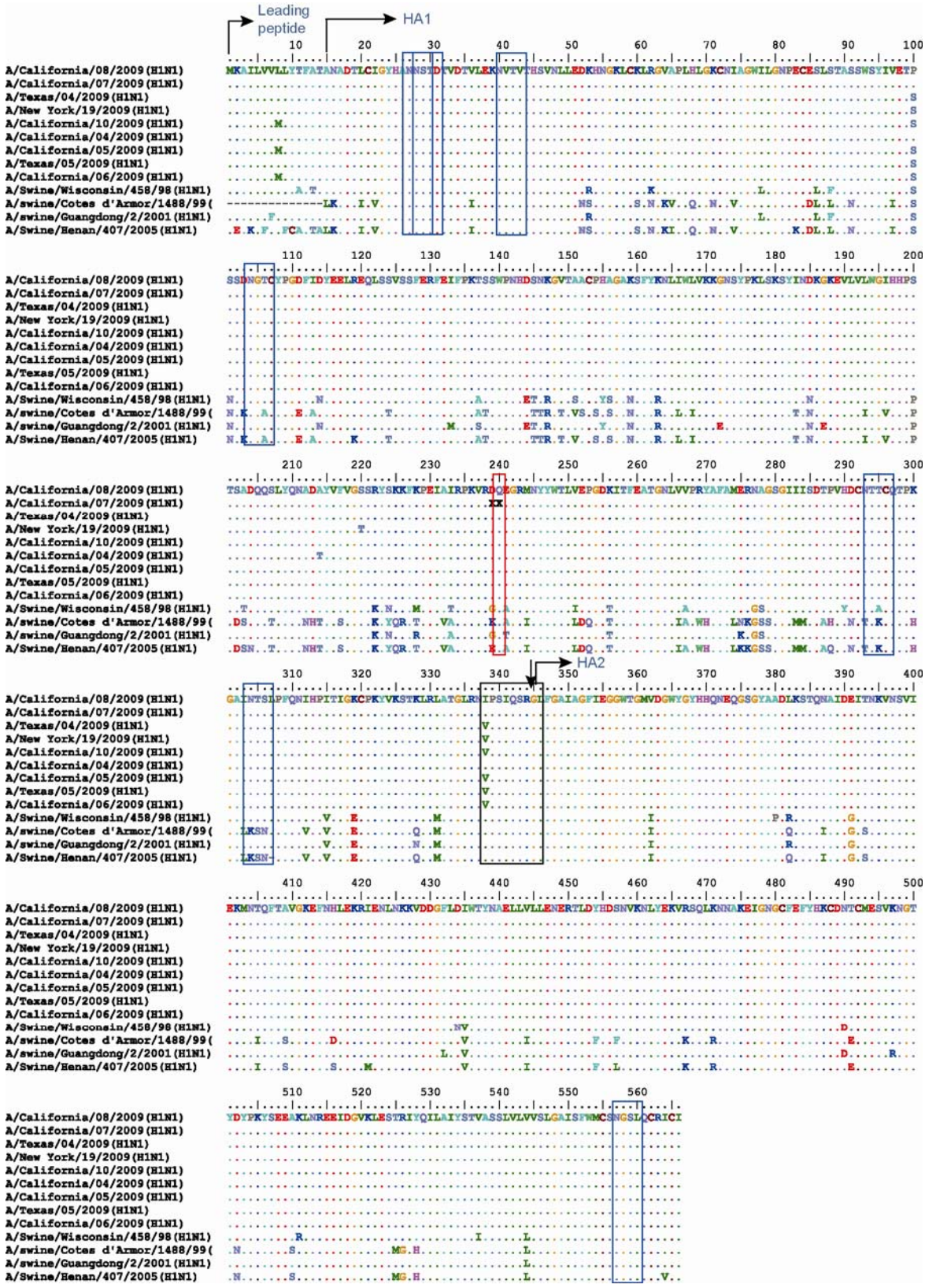

Figure 2 Amino acid sequence comparison of the HA proteins between the isolates from North America and related reference viruses available in GenBank. Only the amino acids different from those in the consensus sequence are indicated. The blue boxed residues indicate the potential glycosylation sites, the black boxed residues indicate the HA cleavage sites, and the red boxed residues indicate the undetermined amino acid sites.

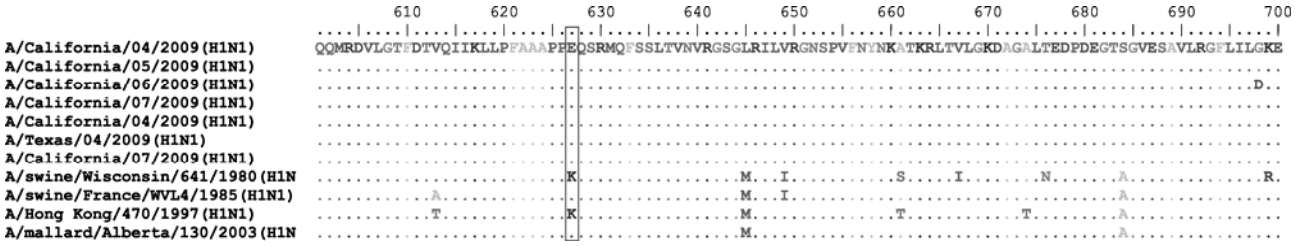

Figure 3 Amino acid sequence comparison of the PB2 proteins between the isolates from North America and related reference viruses available in GenBank. Only the amino acids different from those in the consensus sequence are indicated. The boxed residues indicate the potential glycosylation sites.

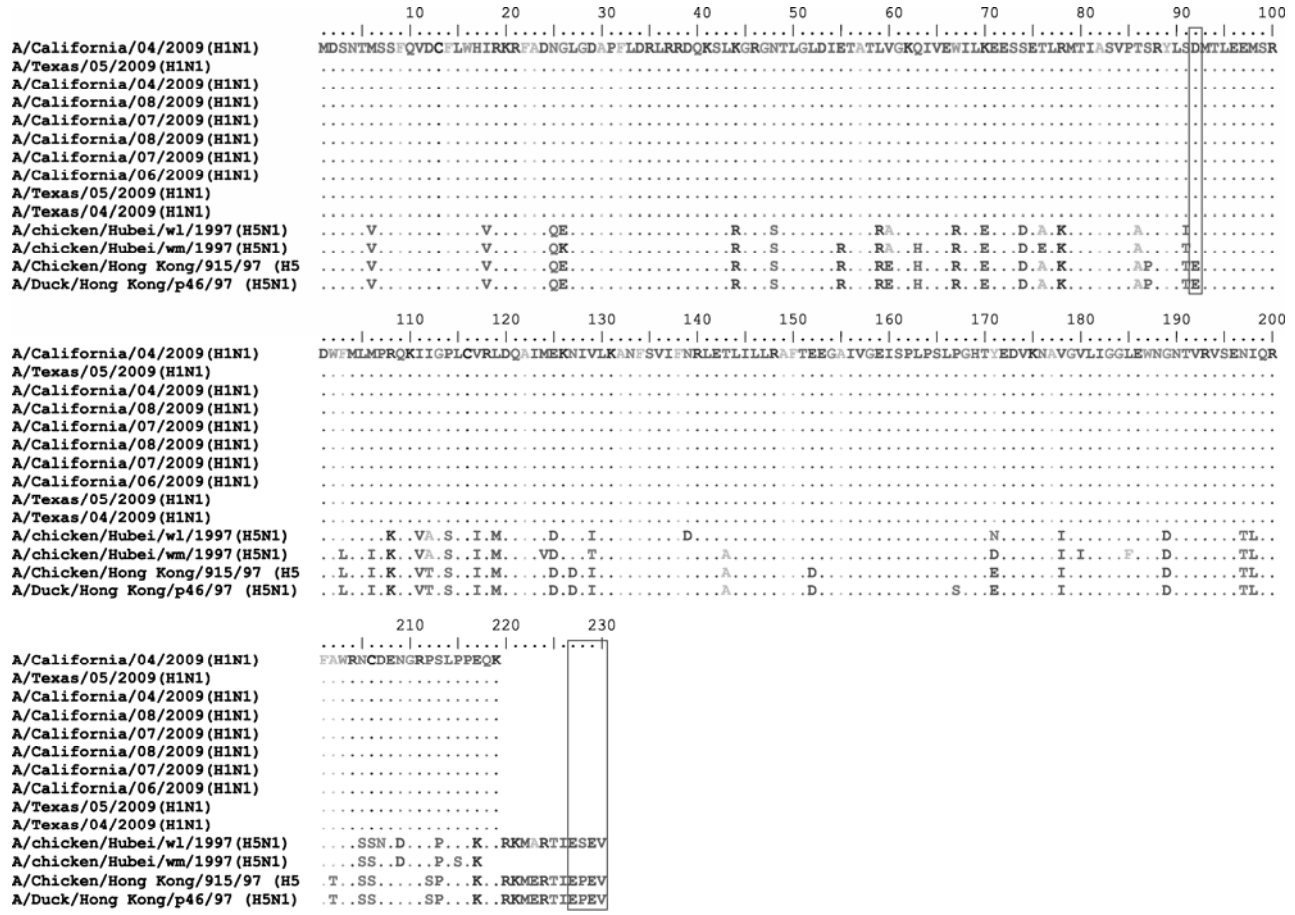

Figure 4 Amino acid sequence comparison of the NS1 proteins between the isolates from North America and related reference viruses available in GenBank. Only the amino acids different from those in the consensus sequence are indicated.

all the isolates for North America exhibited an S31N mutation, which results in the loss of a drug-sensitive site[12]. Amantadine and Rimantadine target the M2 protein, and single mutations in the trans-membrane domain of M2 (e.g. positions 26 (Leu, L), 27 (Val, V), 30 (Ala, A), and 31 (Asn, N)) can confer resistance to these drugs[12]. Thus, it was suggested that the North America isolates are resistant to this class of antiviral drugs. In order to show clearly the spatial relationship of each residue, an almost complete model of M2 protein (Figure 5, a and d), and a partial model of M2 (Figure 5, b—f) were built by SWISS-MODEL. It has been reported that several amino acid mutations, including E119, R152, H275, R293, and N295, could confer viral resistance to NA inhibitors (e.g. oseltamivir and zanamivir)[13]. No similar amino acid substitutions were observed at the conserved residues in the NA of the North America isolated viruses (data not shown). These results suggest that the North

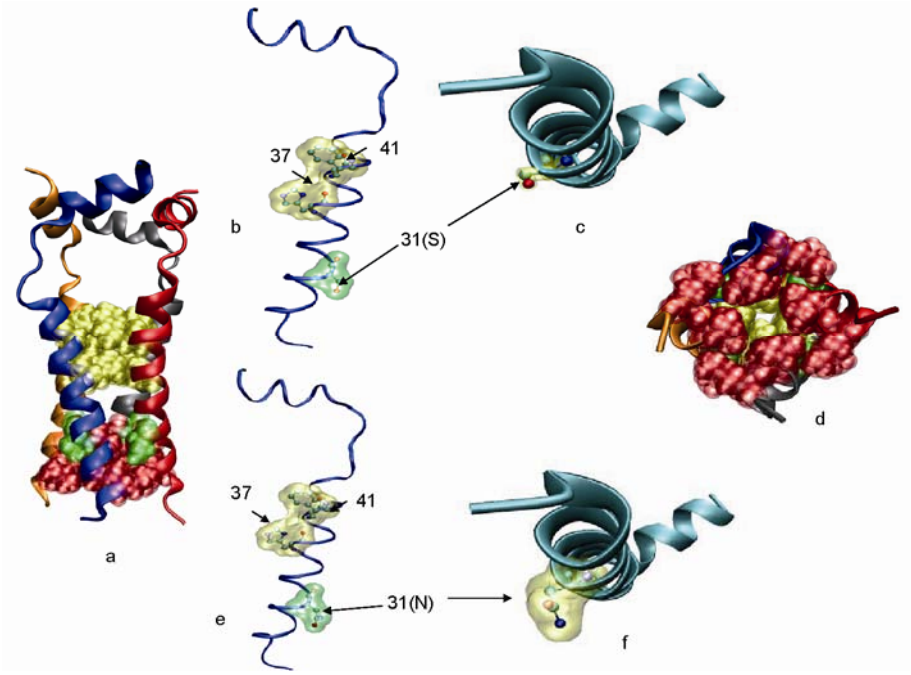

Figure 5 The three-dimensional structures of M2 protein. a and d indicate the frontal view and the bottom view of the M2 protein tetramer in the drug-sensitive virus, respectively; b and c indicate the frontal view and the bottom view of the M2 protein monomer related to drug-sensitive sites, respectively; e and f indicate the frontal view and the bottom view of the M2 protein monomer related to drug-resistant sites, respectively. Yellow indicates the molecular surface of 37 and 41 amino acid residues, which consists of the core of the H^+ ion channel (a and e); red indicates position 26, 27 and 30 amino acid residues (a and d), which may have an important effect on the binding of Amantadine and Rimantadine to A influenza virus. Green indicates position 31 amino acid residues, which exhibit a mutation of S31N in North America isolates.

America isolated viruses are sensitive to NA inhibitors.

3 Discussion

In our present studies, all the ten 2009 North America isolates from human cases were exactly identical by comparison of 8 gene segments. Therefore, we selected the A/California/06/2009(H1N1) isolate as a representative of all the North America isolates to continue further analysis. The nucleotide genetic distance of *HA* genes was analyzed between the (A/California/06/2009(H1N1)) isolate and 24 classical reference viruses from North America, Europe and Asia (Table 3). Of the 24 reference strains, the U.S. H1N2 isolate showed the minimal genetic distance with (A/Swine/Indiana/P12439/00 (H1N2)) for *HA* gene. According to comparison of the other genes, the minimal genetic distances with the isolate (A/California/06/2009 (H1N1)) were 0.06133 (A/swine/England/WVL7/1992 (H1N1)) for NA, 0.04951 (A/Swine/Minnesota/55551/00 (H1N2), A/Swine/Minnesota/593/99 (H3N2) and A/Swine/Iowa/ 569/99 (H3N2)) for *PA*, 0.02401 (A/Swine/Iowa/569/99(H3N2) and A/Swine/Minnesota/593/99 (H3N2)) for *PB2*, 0.04025 (A/duck/NC/91347/01 (H1N2)) for *PB1*, 0.03411 (A/swine/Korea/CY05/2007 (H3N2)) for *NP*, 0.0202 (A/swine/Hong Kong/5212/99 (H3N2)) for *NP*, and 0.04156 (A/Swine/Minnesota/55551/00 (H1N2)) for *NS*, respectively. Phylogenetic trees of the nucleotide sequences of all the 8 genes are shown in Figure 1. Phylogenetic analysis clearly showed that all *HA*, *M*, *NP*, *PB2*, *PA*, *NA* and *NS* genes were in different influenza virus lineages. Although most of the genes showed close phylogenetic relationships with the swine influenza virus lineage, all the 8 genes were placed in three different influenza virus lineages including the swine, avian and human. Therefore, these results demonstrated that all the ten North America H1N1 virus isolates are reassortants between classical H1N1, H1N2 and H3N2 swine, human and avian influenza viruses. Our current results may also provide potent evidence for the results reported by Debora MacKenzie and Novel Swine-Orgin Influenza A (H1N1) virus Investigation Team[14], 1).

1) See footnote 7) on page 2180.

tor-binding properties and to its being targeted by neutralizing antibodies since it represents the membrane fusion glycoprotein of influenza virus. The residues within the receptor-binding site are relatively conserved but the residue mainly responsible for NeuAc α 2,6 Gal linkage specific for the H3 subtype was Ile226 instead of Leu226 as previously reported[15,16]. In our present studies, we did not observe significant changes in amino acids residues of the HA protein at position 131, 158, 160, 183, 187, 222 and 227. These amino acids residues were located at either the HA receptor-binding site (positions 158, 160, 183), at the edge of the HA receptor-binding pocket (position 131, 222 and 227), or next to this pocket (position 187)[17]. In another study, a single amino acid residue change, L226Q, at the right edge of the receptor-binding pocket was shown to be responsible for altering the specificity of the HA from an SAα2,3 Gal to SAα2,6Gal linkage[18], as well as altering the host range of the virus. In our present studies, all the North America isolates but A/California/07/2009(H1N1) were amino acid residue Gln (Q) at position 226, which was responsible for SAα2,6Gal linkage. Thus, North America isolates have obtained the ability of transmitting from human to human or swine. The facts (As of 06:00 GMT, 18 May 2009, 40 countries have officially reported 8829 cases of influenza A(H1N1) infection) (http://www.who.int/csr/don/2009_05_18/en/index.html) have proved our results.

Currently, two specific classes of antiviral drugs are available to manage influenza virus infection: the inhibitors of M2 protein, amantadine and its congener rimantadine, and the neuraminidase (NA) inhibitors, zanamivir and oseltamivir[19]. These two classes of drugs target different viral proteins and have different mechanisms of action on the replication cycle of the virus. Amantadine and rimantadine inhibit virus replication during the early stage of infection by blocking the ion channel formed by the transmembrane domain of the M2 protein[20,21], and NA inhibitors interrupt the established replication cycle by preventing virus release and allowing viruses to clump on the cell surface[17,22–24]. M2 blockers have an inhibitory effect on different subtypes of influenza A viruses. However, a major concern about using M2 blockers is their association with rapid emergence of drug resistance and ineffectiveness against influenza A virus infection. In our present study, we observed that M2 protein exhibited an S31N mutation, which results in the loss of a drug-sensitive site. Resistant variants possessed amino acid substitutions at positions 26, 27, 30, 31 and 34 in the target M2 protein[20,21]. In addition, amantadine-resistant variants showed no obvious loss of virulence or transmissibility in animal models or humans, i.e., M2 mutations conferring resistance to amantadine did not compromise the replicative capacity or virulence of the influenza A viruses[25]. Other studies showed that naturally occurring amantadine-resistant influenza A viruses are quiet rare[26]. However, it was reported that the level of amantadine-resistant variants increased among viruses of the H3[27] and H5 HA subtypes[28].

Zanamivir and oseltamivir are potent and specific against all 9 NA subtypes of influenza A and B viruses, and they have minimal adverse effects, as reported in clinical trials[19,29–31]. Unlike M2 blockers, NA inhibitors appear to be associated with a lower frequency of emergence of drug-resistant variants caused by mutations in either the NA or hemagglutinin (HA). The NA substitutions in drug-resistant viruses include amino acid residues at position 119, 152, 274, 292 and 294 of the enzyme's active centre[31–34]. In this study, we did not observe significant amino acid sequence differences at the conserved sites of the NA.

The contribution of NS to the genetic variability of influenza virus and the discovery of NS families with unique proteotype combinations prompted us to examine this gene in detail. The NS gene encodes the proteins NS1 and NS2. NS1 is found only in infected cells and regulates many cell functions during infection[35]. Substantial evidence suggests that NS1 plays a certain role in virulence by abrogating the expression of antiviral genes in host cells, including interferon (IFN), nuclear factor kappa B (NF-κB), and RNA activated protein kinase (PKR) pathways[36–40]. PDZ domains are modular protein interaction domains that bind in sequence-specific fashion to short C-terminal peptides (ESEV-COOH). They are known to function as scaffold proteins that coordinate the assembly of supermolecular complexes that perform localized signaling at particular subcellular locations[41]. Proteins that contain PDZ domains play important roles in many key signaling pathways, including regulating the activity and trafficking of membrane proteins, maintaining cell polarity and morphology, and organizing the postsynaptic density in neuronal cells. The human viruses with the EPEV motif are highly virulent isolates from the 1997—1999 out-

breaks in Hong Kong. The human isolates with the PL motif ESEV were from the 2003－2004 outbreaks in Hong Kong of China, Vietnam, and Thailand. The 1918 virus NS1 protein has a unique PL motif, KSEV. Thus, the recent high-mortality H5N1 outbreaks are all characterized by NS1 proteins with ''avian'' PL motifs, whereas previous low-mortality human pandemics (in 1957 and 1968) are characterized by NS1 proteins with ''human'' PL motif[9]. We examined our collection of the North America isolates NS data at the C terminus of NS1 (ESEV); however, we did not observe the structure "ESEV" at the C terminus of NS1. Therefore, the North America isolates were the lowly pathogenic human virus. This finding reveals a fact that the North America isolates could cause high morbidity but low mortality (40 countries have officially reported 8829 cases of influenza A(H1N1) infection, including 74 deaths.) (http://www.who.int/csr/don/2009 _05_18/en/index.html).

In summary, influenza is a globally important contagion that continues to cause an unacceptable number of deaths and substantial economic losses worldwide. The present studies clearly indicate that the North America H1N1 influenza virus isolates are of high affinity to humans but lowly pathogenic for humans. In addition, we have obtained baseline data for the sensitivities of the North America H1N1 influenza virus isolates to the NA inhibitors zanamivir and oseltamivir, and resistance to amantadine. Taken together, these findings may help to eliminate anxiety and panics because influenza can be treated effectively.

1 Webster R G, Bean W J, Gorman O T, et al. Evolution and ecology of influenza A viruses. Microbiol Rev, 1992, 56: 152－179
2 Peiris J S, de J, Guan Y. Avian influenza virus (H5N1): a threat to human health. Clin Microbiol Rev, 2007, 20: 243－267
3 Kumar S, Tamura K, Jakobsen I B, et al. MEGA2: Molecular evolutionary genetics analysis software. Bioinformatics, 2001, 17: 1244－1245
4 Thompson J D, Higgins D G, Gibson T J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res, 1994, 22: 4673－4680
5 Schwede T, Kopp J, Guex N, et al. SWISS MODEL: An automated protein homology-modeling server. Nucleic Acids Res, 2003, 31: 3381－3385
6 William H, Andrew D, Klaus S. VMD: Visual molecular dynamics. J Mol Grap, 1996, 14: 33－38
7 Gabriel G, Herwig A, Klenk H D. Interaction of polymerase subunit PB2 and NP with importin alpha1 is a determinant of host range of influenza A virus. PLoS Pathog, 2008, 4: e11
8 Qi X, Li X, Rider P, et al. Molecular characterization of highly pathogenic H5N1 avian influenza A viruses isolated from raccoon dogs in China. PLoS ONE, 2009, 4(3): e4682, doi:10.1371/journal.pone.0004682
9 Obenauer J C, Denson J, Mehta P K, et al. Large-scale sequence analysis of avian influenza isolates. Science, 2006, 311: 1576－1580
10 Jackson D, Hossain M J, Hickman D, et al. A new influenza virus virulence determinant: The NS1 protein four C-terminal residues modulate pathogenicity. Proc Natl Acad Sci USA, 2008, 105: 4381－4386
11 Atsushi O, Takashi M, Hideo T. Protonation of histidine and histidine-tryptophan interaction in the activation of the M2 ion channel from influenza A virus. Biochemistry, 2001, 40: 6053－6060
12 Suzuki H, Saito R, Masuda H, et al. Emergence of amantadine-resistant influenza A viruses: Epidemiological study. J Infect Chemother, 2003, 9: 195－200
13 Scholtissek C, Quack G, Klenk H D, et al. How to overcome resistance of influenza A viruses against adamantane derivatives. Antiviral Res, 1998, 37: 83－95
14 Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team. Emergence of a Norel Swine-Origin Influenza A (H1N1) Virus in Hamans. N Engl J Med, 2009, doi: 10.1056/NEJMoa0903810
15 Parrish C R, Kawaoka Y. The origins of new pandemic viruses: The acquisition of new host ranges by canine parvovirus and influenza A viruses. Annu Rev Microbiol, 2005, 59: 553－586
16 Skehel J J, Wiley D C. Receptor binding and membrane fusion in virus entry: The influenza hemagglutinin. Annu Rev Biochem, 2000, 69: 531－569
17 Weis W, Brown J H, Cusack S, et al. Structure of the influenza virus hemagglutinin complexed with its receptor, sialic acid. Nature, 1988, 333: 426－431
18 Rogers G N, Paulson J C, Daniels R S, et al. Single amino acid substitutions in influenza hemagglutinin change receptor binding specificity. Nature, 1983, 304: 76－78
19 Monto A S. The role of antivirals in the control of influenza. Vaccine, 2003, 21: 1796－1800
20 Hay A J, Wolstenholm A J, Skehel J J, et al. The molecular basis of the specific anti-influenza action of amantadine. EMBO J, 1985, 4: 3021－3024
21 Pinto L H, Holsinger L J, Lamb L A. Influenza virus M2 protein has ion channel activity. Cell, 1992, 69: 517－528
22 Madren L K, Shipman C, Hayden F G. *In vitro* inhibitory effects of combinations of anti-influenza agents. Antivir Chem Chemother, 1995, 6: 109－113
23 Varghese J H, McKimm-Breschkin J L, Caldwell J B, et al. The structure of the complex between influenza virus neuraminidase and sialic acid, the viral receptor. Proteins, 1992, 14: 327－332
24 Von Itzstein M, Wu W Y, Kok G K, et al. Rational design of potent sialidase based inhibitors of influenza virus protection. Nature, 1993, 363: 418－423
25 Hayden FG. Amantadine and rimantadine—clinical aspects. In:

Richman D D, ed. Antiviral Drug Resistance. New York: Wiley, 1996. 59－77
26 Ziegler T, Hemphill M L, Ziegler M L, et al. Low incidence of rimantadine resistance in field isolates of influenza A viruses. J Infect Dis, 1999, 180: 935－939
27 Bright R A, Medina M, Xu X, et al. Incidence of adamantine resistance among influenza A (H3N2) viruses isolated worldwide from 1994 to 2005: a cause for concern. Lancet, 2005, 366: 1175－1181
28 Ilyushina N A, Govorkova E A, Webster R G. Detection of amantadine-resistant variants among avian influenza viruses isolated in North America and Asia. Virology, 2005, 341: 102－106
29 Monto A S, Fleming D M, Henry D, et al. Efficacy and safety of the neuraminidase inhibitor zanamivir in the treatment of influenza A and B virus infections. J Infect Dis, 1999, 180: 254－261
30 Lew W, Chen X, Kim C U. Discovery and development of GS4104 (oseltamivir): An orally active influenza neuraminidase inhibitor. Curr Med Chem, 2000, 7: 663－672
31 McKimm-Breschkin J L. Resistance of influenza viruses to neraminidase inhibitors —a review. Antivir Res, 2000, 47: 1－17
32 McKimm-Breschkin J L. Management of influenza virus infections with neuraminidase inhibitors: detection, incidence, and implications of drug resistance. Treat Respir Med, 2005, 4: 107－116
33 Kiso M, Mitamura K, Sakai-Tagawa Y, et al. Resistant influenza A viruses in children treated with oseltamivir: Descriptive study. Lancet, 2004, 364: 759－765
34 Gubareva L V, Matrosovich M N, Brenner M K, et al. Evidence for zanamivir resistance in an immunocompromised child infected with influenza B virus. J Infect Dis, 1998, 178: 1257－1262
35 Robert M K, Wei M Y, Diana L N, et al. Intracellular warfare between human influenza viruses and human cells: The roles of the viral NS1 protein. Virology, 2003, 309: 181－189
36 Stacey S C, Naomi D S, Gabriele N, et al. Influenza virus NS1 protein induces apoptosis in cultured cells. J Virol, 2001, 75: 7875
37 Adolfo G S, Andrej E, Demetrius M, et al. influenza a virus lacking the ns1 gene replicates in interferon-deficient systems. Virology, 1998, 252: 324－330
38 Sang H S, Erich H, Robert G W.The *NS1* gene of H5N1 influenza viruses circumvents the host anti-viral cytokine responses. Virus Res, 2004, 103: 107－113
39 Gary K G, Mirella S, Terrence M T, et al. Cellular transcriptional profiling in influenza A virus-infected lung epithelial cells: The role of the nonstructural NS1 protein in the evasion of the host innate defense and its potential contribution to pandemic influenza. Proc Natl Acad Sci USA, 2002, 99: 10736－10741
40 Aleksandr S L, Samita A, Richard J W, et al. Pathogenesis of Hong Kong H5N1 influenza virus NS gene reassortants in mice: The role of cytokines and B- and T-cell responses. J Gen Virol, 2005, 86: 1121－1130
41 Morgan S, Carlo S. PDZ domains and the organization of supramolecular complexes. Annu Rev Neurosci, 2001, 24: 1－29