

Epidemiology and full genome sequence analysis of H1N1pdm09 from Northeast China

Xiaoyan Li · Mei Kong · Jinying Chen ·
Xu Su · Ming Zou · Liru Guo · Lin Li ·
Jianqing Ding

Received: 23 February 2013 / Accepted: 31 May 2013 / Published online: 15 June 2013
© Springer Science+Business Media New York 2013

Abstract Pandemic influenza A (H1N1) 2009 virus (H1N1pdm09) was a novel tri-assortment virus that emerged in Mexico and North America in 2009 and caused the first influenza pandemic in the 21st century. This study investigated the prevalence pattern and molecular characteristics of H1N1pdm09 in three continuous years from April 2009 to March 2012 in populations of Tianjin, Northeast China. Totally, 3,068 influenza viruses (25.4 %) were detected from 12,089 respiratory specimens. Among them, 41.4 % (1,269/3,068) were H1N1pdm09 positive. 15.1 % (192/1,269) severe respiratory infection cases were H1N1pdm09 positive. H1N1pdm09 was the predominant prevalence subtype in October 2009–March 2010 (69.1 %, 930/1,346) and October 2010–March 2011 (42.1 %, 220/523). Eight isolated H1N1pdm09 viruses from severe infection/death cases in three different years were selected to sequence the whole genome through splicing the sequences following 46 PCRs. HA sequences of seven H1N1pdm09 isolates from mild infection cases were detected. Phylogenetic analysis showed that HA, NA, M, NP and NS genes of H1N1pdm09 viruses gathered

together with swine influenza A (H1N1), whereas PB2 and PA genes originated from avian influenza virus, and PB1 gene originated from human seasonal influenza virus. Identity analysis indicated that all the genes were highly conserved. Compared with vaccine strain A/California/07/2009(H1N1), the maximal mutation gene was HA (0.7–2.6 %), then NA (0.6–1.7 %), last one was M (mutation rate 0–0.6 %). More site substitutions were observed in 2011 isolates than in 2009 and 2010 isolates of HA ($p = 0.002$), NA ($p = 0.003$) and PA ($p = 0.001$) proteins. The amino acid substitution rates were varied among eight gene segments, ranging from 7.39×10^{-4} for PB2 to 7.40×10^{-3} for NA. The higher d_N/d_S rates were observed in HA, PA and NS segments in H1N1pdm09 in Tianjin. Three HA amino acid site substitutions occurred at the HA receptor-binding sites and antigenic determinant, including S179N and K180T (located at antigenic site Sa) in A/Tianjinhedong/SWL44/2011(H1) and A/Tianjinjinnan/SWL41/2011(H1), and D239N (located at antigenic site Ca) in A/Tianjinninghe/SWL49/2009(H1). Antigenic drift may have occurred in H1N1pdm09 with time. No oseltamivir-resistance site substitution was observed at 275 and 295 sites. Amino acid residue site at 31 in M2 protein was N in all 8 isolates, which suggested that H1N1pdm09 was resistant to amantadine.

Electronic supplementary material The online version of this article (doi:10.1007/s11262-013-0931-1) contains supplementary material, which is available to authorized users.

X. Li · J. Chen (✉)
School of Basic Medicine, Tianjin Medical University,
No.22 Qixiangtai Road, Heping District, Tianjin 300070,
People's Republic of China
e-mail: jinyingchen1@163.com

X. Li
e-mail: xiaoyanli1291@163.com

X. Li · M. Kong · X. Su · M. Zou · L. Guo · L. Li · J. Ding
Institute of Pathogenic Biology, Tianjin Centers for Disease
Control and Prevention, No. 6 Huayue Road, Hedong District,
Tianjin 300011, People's Republic of China

Keywords Pandemic influenza A (H1N1) 2009 ·
Whole genome sequencing · Phylogenetic analysis ·
Clinical severity

Introduction

Influenza virus infection is a great burden on health, causing seasonal outbreaks that affect approximately 10 % of the population each year [1]. Antigen variation often happens to influenza viruses which can lead to influenza

prevalence, even worldwide pandemic. Influenza pandemic occurred four times in the last century: the 1918 Spanish influenza, the 1957 Asian influenza, the 1968 Hong Kong influenza and the 1977 Russian influenza [1, 2]. Among them, the Spanish and Russian influenza pandemic were caused by H1N1 subtype influenza virus.

In April 2009, a novel H1N1 influenza virus emerged in Mexico and North America [3, 4] and rapidly spread to more than 70 countries, causing approximately 28,000 infections and 144 deaths before the World Health Organization announced it as the first pandemic in the 21st century, known as pandemic H1N1 2009 (H1N1pdm09) [5]. This virus was a novel gene re-assortment influenza virus from North American avian and classical swine influenza, human influenza, and Eurasian swine influenza [6].

The first laboratory-confirmed case in China was detected on 10 May 2009 [isolated virus strain was A/Sichuan/1/2009(H1N1)] in Sichuan Province, southern part of China [7]. Since the outbreak of the H1N1pdm09 influenza infection, a large-scale surveillance was carried out in China on molecular level, the evolutionary and spatial dynamics of this virus in human populations. The network laboratories expanded from original 83 to 410. Six new laboratories were developed in Tianjin since then. Influenza in the south part (2 prevalent peaks, winter and summer) and north part (1 peak, winter) of China have different prevalence patterns. Tianjin is located in the northeast part of China, which showed one prevalent peak usually in winter season. To clarify the prevalent pattern of H1N1pdm09 in Tianjin and its molecular characteristics, epidemiology and full genome sequence analysis were performed in April 2009–March 2012.

Materials and methods

Sample collection

11,324 Throat swabs were collected from patients with influenza-like illness (ILI) in 9 sentinel influenza surveillance hospitals and ILI outbreak cases in Tianjin between April 2009 and March 2012 [8]. 765 Other respiratory samples were collected from severe respiratory infection cases in the same period in Tianjin [8].

Virus screening and isolation

All samples were processed and detected by real-time PCR for influenza virus typing and subtyping according to the CDC protocol [9]. Next, H1N1pdm09 viruses were isolated from real-time PCR-positive samples. Samples were inoculated onto Madin–Darby canine kidney cells (MDCK) and cultured with serum-free minimum essential medium (MEM; Gibco, USA) in the presence of 2.0 µg/ml of

trypsin (Sigma, USA) [10]. Virus was identified by hemagglutination inhibition (HI) tests with virus type-specific anti-serum which was distributed by National Influenza Center (NIC) and stored at −80 °C.

RNA extraction and sequencing

RNA for real-time PCR was extracted from 500 µl respiratory samples and RNA for sequencing after RT-PCR amplification was extracted from 50 µl viral aliquots using NucliSENS easyMAG (BioMerieux, France) according to the manufacturer's instructions.

The whole genome of eight H1N1pdm09 isolates from patients with severe respiratory infection (reported as unexplained pneumonia) and HA segment of seven H1N1pdm09 isolates from upper respiratory infection (mild symptom infection) were sequenced by performing one-step RT-PCR amplification (Qiagen, Germany), followed by direct sequencing of the products using the WHO recommended protocol with some modification in M13F/M13R-tailed primers [3]. In brief, 46 pairs of M13F/M13R-tailed primers for eight segments were used to perform RT-PCR separately. Product of each primer has 200–300 bp overlapping sequences. All sequencing procedure was conducted by Invitrogen Biotechnology Co. Ltd., Shanghai, China using an ABI 3730 DNA Analyzer (Applied Biosystems, USA). The splicing of each primer products' sequence was performed using Seqman software in DNASTar (version 5.01).

Sequence alignment and phylogenetic analysis

Nucleotide sequence alignments were generated using the ClustalW algorithm of the MEGA4.0.1 (<http://www.mega-software.net/mega.html>) software. The nucleotide and induced amino acid sequences identities were calculated using Megalign software in DNASTar (version 5.01). Vaccine strain A/California/07/2009(H1N1), the early isolate A/California/04/2009(H1N1) and first Chinese isolate A/Sichuan/01/2009(H1N1) were taken as the reference genome and whole open reading frames (ORFs) were adopted for phylogenetic tree construction. Other representative worldwide strains were included from America, Australia, southeastern Asia, Europe and south and north China [10, 11]. All data were downloaded from GenBank. Phylogenetic trees were constructed by MEGA4.0.1 version using the neighbor-joining method. Bootstrap proportions were plotted at the main internal branches of the phylogram to show support values.

Substitution rate estimation

We estimated the substitution rate of the H1N1pdm09 virus sequences using a Bayesian Markov chain Monte Carlo

(MCMC) approach as implemented in the BEAST package v1.7.5 (<http://beast.bio.ed.ac.uk/>). The HKY substitution model was used with a gamma parameter of site heterogeneity model. A strict molecular clock model was used. For each analysis, a chain length of 10,000,000 was used and sampled for every 1,000 states. Convergence was confirmed with Tracer v1.5.

Amino acid mutation rate was calculated as formula:

No. of amino acid mutation/

No. of all amino acid in single protein $\times 100\%$.

Selection pressure measurement

To determine the overall selection pressure faced by each gene segment in H1N1pdm09, we estimated the mean numbers of nonsynonymous substitutions (d_N) and synonymous substitutions (d_S) per site using single-likelihood ancestor counting (SLAC) method, accessed through the Datamonkey interface (<http://www.datamonkey.org>). In all cases, d_N/d_S estimates were based on neighbor-joining trees under the substitution model selected by website.

Statistical analyses

Chi-square test was applied in this study to perform statistical analysis on amino acid mutation rate using SPSS11.5 software.

Results

Prevalence of influenza viruses

Totally, 3,068 influenza viruses (25.4 %) were detected from 12,089 respiratory specimens in April 2009–March 2012 (Table 1; Fig. 1). From April 2009 to September 2009, called summer surveillance season, seasonal H1N1, H3N2, H1N1pdm09 and B type influenza viruses co-circulated, but seasonal H3N2 and H1N1pdm09 were the

predominant subtypes, accounting for 57.7 % (184/319) and 37.0 % (118/319), respectively, in all influenza viruses. Only two seasonal H1N1s were detected. From October 2009 to March 2010, called winter surveillance season, seasonal H3N2, H1N1pdm09 and B type influenza virus co-circulated, and H1N1pdm09 (69.1 %, 930/1,346) was absolutely the predominant subtype. In the next summer season, seasonal H3N2 (79.7 %, 114/143) was the major subtype. In the next winter season, seasonal H3N2 (38.4 %, 201/523) and H1N1pdm09 (42.1 %, 220/523) moved back as the advantage prevalent subtypes. In the summer season in 2011, only type B was detected. And in the following winter season, type B (90.0 %, 633/703) still remained the absolute prevalent type. Since October 2009, no seasonal H1N1 was detected.

Prevalence of H1N1pdm09

Among 3,068 influenza virus positive specimens, 41.4 % (1,269/3,068) was H1N1pdm09 positive. 15.1 % (192/1,269) severe respiratory infection cases were H1N1pdm09 positive. The H1N1pdm09 constituent ratio was shown in Table 1 and Fig. 1. None of the H1N1pdm09 was detected since April 2011. H1N1pdm09 was the predominant prevalence subtype in October 2009–March 2010 and October 2010–March 2011.

Whole genome sequencing of H1N1pdm09

Eight H1N1pdm09 isolates from severe respiratory infection or death cases (some severe respiratory infection cases were finally died) were selected to sequence the whole genome (Table 2). Among them, four were isolated from 2009, two were from 2010 and the other two were isolated from 2011. As the HA protein was the major antigenic protein and most variable protein, seven HA genes were sequenced from throat swabs which were collected from upper respiratory tract infection cases to analyze the relationship between amino acid changes in HA and clinical

Table 1 Prevalence of influenza virus from April 2009 to March 2012

Time	No. of tested	No. of positive specimens (%)	Type/subtype and constituent ratio (%)			
			H1N1pdm09 ^a (%)	Seasonal H1 (%)	Seasonal H3 (%)	B (%)
2009.4–2009.9	1,619	319 (19.7)	118 (37.0)	2 (0.6)	184 (57.7)	15 (4.7)
2009.10–2010.3	3,938	1,346 (34.2)	930 (69.1)	0 (0)	166 (12.3)	250 (18.6)
2010.4–2010.9	988	143 (14.5)	1 (0.7)	0 (0)	114 (79.7)	28 (19.6)
2010.10–2011.3	2,172	523 (24.1)	220 (42.1)	0 (0)	201 (38.4)	102 (19.5)
2011.4–2011.9	912	34 (3.7)	0	0	0	34 (100)
2011.10–2012.3	2,460	703 (28.6)	0	0	70 (10.0)	633 (90.0)
Total	12,089	3,068 (25.4)	1,269 (41.4)	2 (0.06)	735 (24.0)	1,062 (34.6)

^a H1N1pdm09: pandemic influenza A (H1N1) in 2009

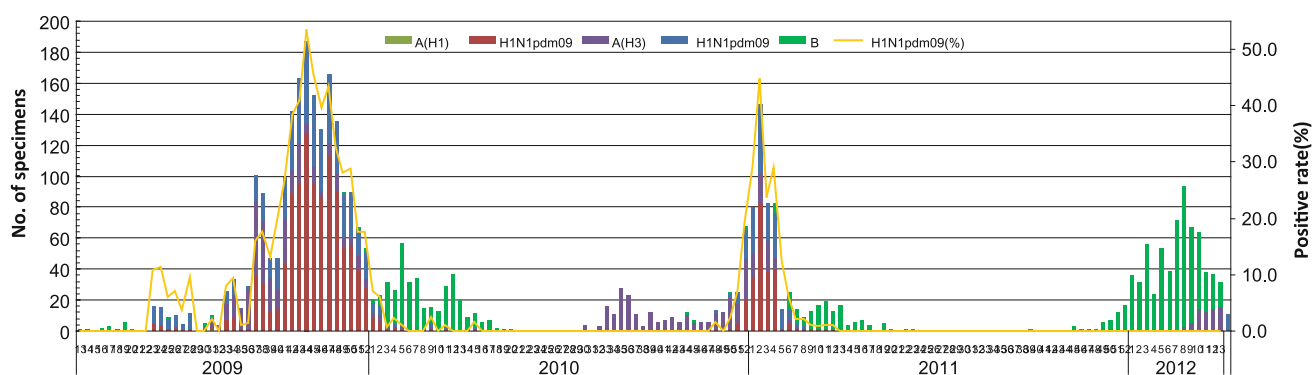


Fig. 1 Prevalence of influenza virus from April 2009 to March 2012. Seasonal H3 seasonal influenza virus H3 subtype. Seasonal H1 seasonal influenza virus H1 subtype. H1N1pdm09 pandemic influenza

A (H1N1) 2009 virus. H1N1pdm09(%) the constituent ratio of pandemic influenza A (H1N1) 2009 virus

Table 2 Information of the selected patients and virus strains sequenced

No.	Age (years)	Gender	Clinical diagnosis	Lab no. of specimen	Date of specimen collected	Name of virus strains
1	41	M	S and D	YQ2009-270	2009.11.12	A/Tianjinheping/SWL313/2009(H1N1)
2	31	M	S and D	2009TF-ZL-0306	2009.11.19	A/Tianjintanggu/SWL432/2009(H1N1)
3	21	F	S and D	2009TF-ZL-0439	2009.12.05	A/Tianjinjinghe/SWL49/2009(H1N1)
4	31	M	S	2009TF-ZL-0625A	2009.12.30	A/Tianjinjinnan/SWL414/2009(H1N1)
5	28	F	S	2010TF-ZL-0017A	2010.01.02	A/Tianjinnankai/SWL49/2010(H1N1)
6	23	F	S	2010TF-ZL-0017E	2010.01.05	A/Tianjinnankai/SWL413/2010(H1N1)
7	45	M	S	2011TF-FY-0004	2011.01.05	A/Tianjinhedong/SWL44/2011(H1N1)
8	71	M	S	2011TF-FY-0007	2011.01.28	A/Tianjinjinnan/SWL41/2011(H1N1)
9	8	F	URI	2009TF-ZL-0164	2009.08.06	A/Tianjinjinnan/SWL164/2009(H1N1)
10	17	F	URI	2009TF-ZL-0172	2009.08.17	A/Tianjinhongqiao/SWL172/2009(H1N1)
11	11	F	URI	2009TF-ZL-0177	2009.08.22	A/Tianjinnankai/SWL177/2009(H1N1)
12	11	M	URI	2009TF-ZL-0178	2009.08.23	A/Tianjinnankai/SWL178/2009(H1N1)
13	11	F	URI	2009TF-ZL-0182	2009.08.24	A/Tianjinhedong/SWL182/2009(H1N1)
14	40	F	URI	2009TF-ZL-0198	2009.09.01	A/Tianjinnankai/SWL198/2009(H1N1)
15	30	M	URI	2009TF-ZL-0299	2009.11.18	A/Tianjinheping/SWL299/2009(H1N1)

M male, F female, S severe respiratory infection, D death, URI upper respiratory infection

severity. Sequences were submitted to GenBank, whose accession numbers are shown in Table 4. The patients and virus strains information are shown in Table 2.

Phylogenetic and genetic diversity analysis

The eight whole genome sequences sampled from Tianjin were subjected to phylogenetic analysis. Two genes were selected as representatives: HA (Fig. 2a) and NA (Fig. 2b). Phylogenetic trees of other genes are shown in Figure S1. Phylogenetic analysis showed that HA, NA, M, NP and NS genes of H1N1pdm09 viruses gathered together with swine influenza A(H1N1), whereas PB2 and PA genes originated from avian influenza virus, and PB1 gene originated from human seasonal influenza virus. The strains isolated in

2011 were clustered together in a separate branch, a little far from other strains isolated in 2009 and 2010.

Compared with vaccine strain A/California/07/2009(H1N1), HA amino acid sequence identities were between 97.4 and 99.5 %. Identities of HA sequences isolated from severe infection cases and mild infection cases were 97.4–99.3 and 98.9–99.3 %, respectively. Mutation rate of HA sequences isolated in 2011 were 2.3–2.6 %, higher than that in 2009 (0.7–1.1 %) and 2010 (1.1–2.0 %). Mutation rate of NA sequences isolated in 2011 was 1.7 %, higher than that in 2009 (0.6–0.9 %) and 2010 (0.6–1.1 %). Totally, among eight gene segments, the maximal mutation gene was HA, then NA, last one was M (mutation rate 0–0.6 %), as shown in Tables 4 and 5. The amino acid substitution rates were varied among eight

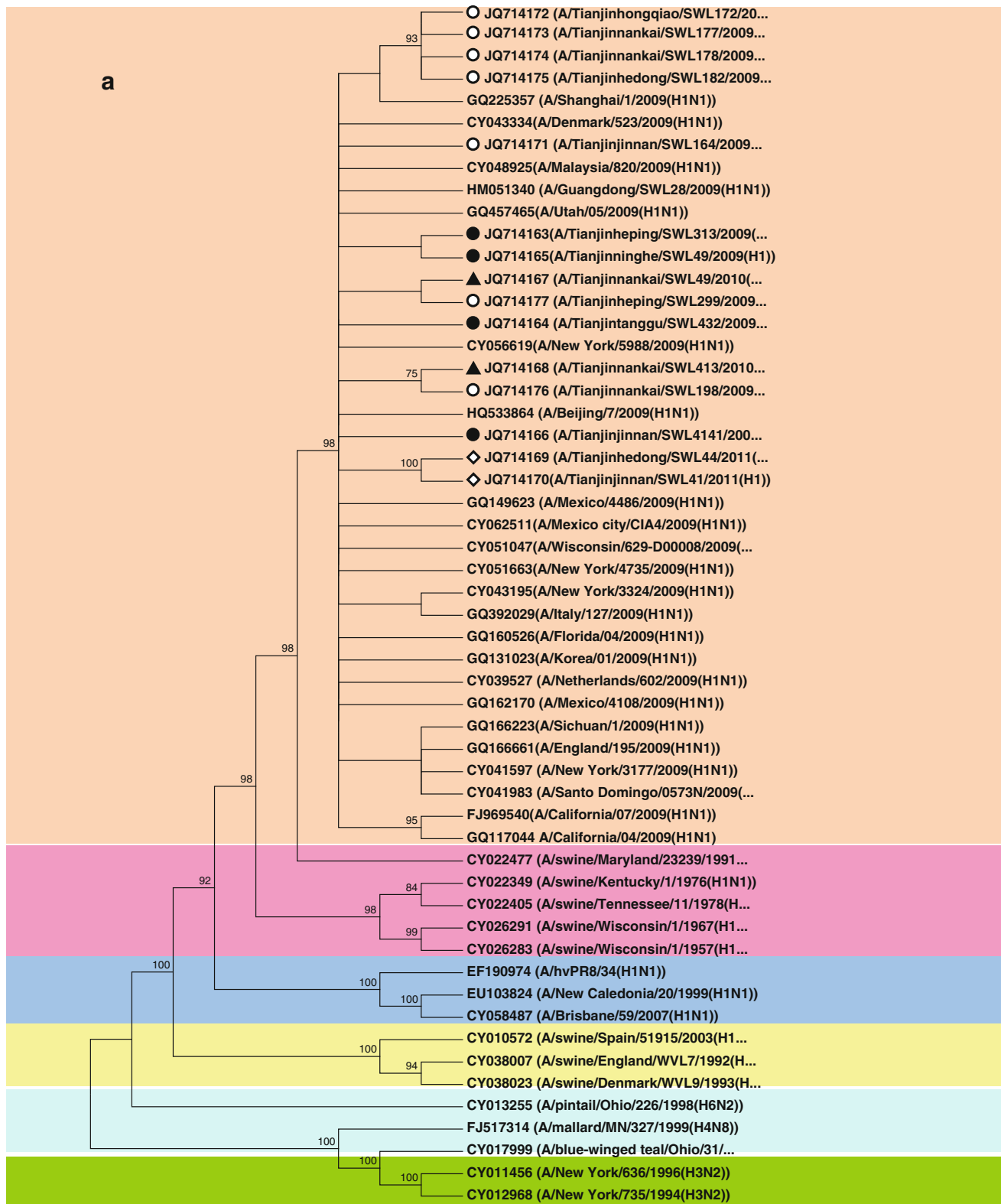


Fig. 2 Phylogenetic tree of HA (a) and NA (b) genes of H1N1pdm isolates collected in Tianjin, China from April 2009 through March 2012. MEGA 4.0 was used to draw phylogenetic tree by neighbor-joining method. The values at the branches denoted the percentage of 1,000 times of boot strap re-sampling and the bootstrap values >70 % are shown.

A/California/07/2009(H1N1) and A/California/04/2009(H1N1) were taken as references and the ranges adopted for phylogenetic analysis of three genes were the ORFs. The viruses indicated by filled circle (2009 severe cases) open circle (2009 mild infection cases) filled triangle (2010) open diamond (2011) were isolated in Tianjin



Fig. 2 continued

Table 3 Genetic evolutionary of H1N1pdm09 sampled in Tianjin

Viral segment	Protein	NT substitutions per gene site per year	AA substitutions per site per year	d_N/d_S
1	PB2	4.80×10^{-4}	7.39×10^{-4}	0.18
2	PB1	1.39×10^{-3}	1.69×10^{-3}	0.07
3	PA	1.74×10^{-3}	2.36×10^{-3}	0.21^a
4	HA	4.11×10^{-3}	5.96×10^{-3}	0.37
5	NP	3.82×10^{-4}	1.10×10^{-3}	0.02
6	NA	2.10×10^{-3}	7.40×10^{-3}	0.28
7	M1	3.80×10^{-4}	1.57×10^{-3}	5×10^{-9}
7	M2	2.20×10^{-3}	–	– ^b
8	NS1	1.44×10^{-3}	5.66×10^{-3}	0.35
8	NEP (NS2)	2.00×10^{-3}	–	– ^b

NT nucleotide, AA amino acid

^a PA 637 codon had a negatively selected sites substitution, $p = 0.099$

^b The d_N/d_S ratio estimate could not be performed for M2 and NEP genes of H1N1pdm09 viruses because only few variable amino acid sites were available; viral segments exhibiting highest genetic diversity were highlighted in bold font

gene segments, ranging from 7.39×10^{-4} for PB2 to 7.40×10^{-3} for NA. NA, HA, NS1 and PA proteins exhibited highest genetic diversity. The higher d_N/d_S rates were observed in HA, PA and NS segments in H1N1pdm09 in Tianjin (Table 3).

Genetic mutation and antiviral drug resistance

The H1N1pdm09 HA amino acid changes (compared with vaccine strain A/California/07/2009(H1N1)) are shown in Table 4. Among them, three HA amino acid substitution occurred in the HA receptor-binding sites and at antigenic determinant, including S179N and K180T (located at antigenic site Sa) in A/Tianjinhedong/SWL44/2011(H1) and A/Tianjinjinnan/SWL41/2011(H1), and D239N (located at antigenic site Ca) in A/Tianjinninghe/SWL49/2009(H1). In addition, P100S, S200T and I338V happened in all Tianjin isolates (8 from severe infection cases and 7 from mild infection cases), and mild infection strains had the unique mutation V428I. E391G and E391K substitution were observed in severe infection isolates and mild infection isolates, respectively. 391 and 428 were not in receptor-binding site or antigenic sites domain. No significant difference was found in HA mutation rate between isolates from severe infection cases and that from mild infection cases if it was analyzed in isolates from the same year (2009) ($\chi^2 = 0.05$, $p = 0.083$).

Other amino acid substitutions of the H1N1pdm09 are shown in Table 5. More site substitutions were observed in

2011 isolates than in 2009 and 2010 isolates in HA ($\chi^2 = 9.64$, $p = 0.002$), NA ($\chi^2 = 12.95$, $p = 0.003$) and PA ($\chi^2 = 14.8$, $p = 0.001$) proteins. In NA, all isolates had the substitutions at V106I and N248D. No antiviral drug-resistant site substitution was observed at 275 and 295 sites. These indicated that the H1N1pdm09 in Tianjin was still susceptible to oseltamivir.

In M2 protein, 31-amino acid site, which was the most important site related with amantadine resistance, was N in all eight isolates, which suggested that H1N1pdm09 was resistant to amantadine.

P224S in PA, I123V in NS and V100I in NP occurred in all eight Tianjin H1N1pdm09.

Discussion

H1N1pdm09 caused the first influenza pandemic in the 21st century. This virus distinctly differed from the seasonal H1N1 influenza virus that was circulating in humans, enabling it to spread in the thus-far naïve population.

Surveillance of influenza from April 2009 to March 2012 showed that from its first detection in Tianjin in 2009, H1N1pdm09 became the predominant prevalence subtype in that winter surveillance season (October 2009–March 2010), it was up to peak in November (91.1 % influenza viruses were H1N1pdm09 in this month). In the whole surveillance season, 69.1 % influenza viruses were H1N1pdm09. With the increasing of H1N1pdm09-infected individuals and the expanding of vaccinated population, immune barrier developed gradually, only one H1N1pdm09 strain was detected in April 2010–September 2010, which was called summer surveillance season. In the next winter surveillance season (October 2010–March 2011), H1N1pdm09, together with seasonal H3N2, became the predominant prevalence subtype once again. Now WHO has announced that H1N1pdm09 will be prevalent alternatively with seasonal H1N1, H3N2 and B in human beings. Three years' surveillance (April 2009–March 2012) in Tianjin confirmed this viewpoint further.

Influenza A virus has eight gene segments. More attention was often paid to HA gene. There are at least four antigenic determinant domains in the HA gene: Sa, Sb, Ca and Cb [12]. D239G/N (located at Ca antigenic determinant domain) mutation in the HA protein was associated with severe and fatal H1N1pdm09 case [13–15]. In this study, D239N mutation occurred in A/Tianjinninghe/SWL49/2009(H1), and the patient infected with this virus subjected to severe pneumonia and died finally. S179N and K180T substitutions were observed in two 2011 isolates, A/Tianjinhedong/SWL44/2011(H1) and A/Tianjinjinnan/SWL41/2011(H1). These two sites together with 239 were all in antigenic determinant domain of the HA protein and this could improve the adaption of H1N1pdm09 to human

Table 4 Amino acid changes and identity of HA from severe and mild respiratory infection cases compared with A/California/07/2009(H1N1)

	1	2	3	4	5	6	7	8
Severe respiratory infection								
Amino acid changes	P100S, S145P, S220T, I338V	P100S, S220T, I338V	P100S, S145P, S220T, D239N, I338V	P100S, V216A, S220T, I338V, E391K	P100S, S220T, R269K, E391K	P100S, D114G, E129K, S220T, I338V, E391K, V540I, S551I, N557H, S559E	P100S, S145P, S179N, K180T, V216A, S220T, P288S, H290R, I312V, I338V, I389F, E391K, D489N	P100S, S145P, S179N, K180T, V216A, S220T, P288S, H290R, I312V, I338V, I389F, E391K, D489N
Percentage of identity (%)	99.4 (99.1)	99.5 (99.3)	99.4 (98.9)	99.4 (98.9)	99.3 (98.9)	98.8 (98.0)	98.5 (97.70)	98.4 (97.4)
GenBank accession number	JQ714163	JQ714164	JQ714165	JQ714166	JQ714167	JQ714168	JQ714169	JQ714170
Mild infection								
Amino acid changes	P100S, S220T, I338VZ	P100S, S220T, I338V, E391G, V428I	P100S, S220T, I338V, E391G, V428I	P100S, S220T, I338V, E391G, V428I	P100S, S220T, I338V, E391G, V428I	P100S, S220T, I338V, E391G, V428I	P100S, S220T, I338V	P100S, S220T, I338V
Percentage of identity (%)	99.5 (99.1)	99.4 (98.9)	99.4 (98.9)	99.4 (98.9)	99.4 (98.9)	99.4 (98.9)	99.5 (99.3)	99.4 (99.3)
GenBank accession	JQ714163	JQ714164	JQ714165	JQ714166	JQ714167	JQ714168	JQ714169	JQ714170

HA amino acid sites numbering starts at the N-terminal of H1 HA1

1: A/Tianjinheping/SWL313/2009(H1N1), 2: A/Tianjintangu/SWL432/2009(H1N1), 3: A/Tianjinninghe/SWL49/2009(H1N1), 4: A/Tianjinjinan/SWL4141/2009(H1N1), 5: A/Tianjinan-kai/SWL49/2010(H1N1), 6: A/Tianjinmankai/SWL413/2010(H1N1), 7: A/Tianjinhedong/SWL44/2011(H1N1), 8: A/Tianjinjinan/SWL41/2011(H1N1), 9: A/Tianjinjinan/SWL164/2009(H1N1), 10: A/Tianjinhongqiao/SWL172/2009(H1N1), 11: A/Tianjinmankai/SWL177/2009(H1N1), 12: A/Tianjinmankai/SWL178/2009(H1N1), 13: A/Tianjinhedong/SWL182/2009(H1N1), 14: A/Tianjinmankai/SWL198/2009(H1N1), 15: A/Tianjinheping/SWL299/2009(H1N1)

Table 5 Amino acid changes and identity of the whole genome (except for HA) compared with A/California/07/2009(H1N1)

Protein	1	2	3	4	5	6	7	8
PB2								
Amino acid changes	T588I	K340N	V344M, T588I	K340N, V511I, K660R	L94P, I176T, V511I, K660R	S107N	V255I, T588I	V255I, T588I
Percentage of identity (%)	99.8 (99.7)	99.7 (99.7)	99.6 (99.6)	99.6 (99.5)	99.4 (99.3)	99.6 (99.7)	99.5 (99.6)	99.5 (99.6)
PB1								
Amino acid changes				T257A, V645I		V200I	V587I, E739D	V587I
Percentage of identity (%)	99.9 (99.9)	99.9 (99.9)	99.8 (99.9)	99.8 (99.6)	99.9 (99.9)	99.7 (99.7)	99.6 (99.6)	99.4 (99.6)
PA								
Amino acid changes	A70V, P224S, D547E	P224S, L226F	A70V, P224S, D547E	P224S, L226F	P224S	V14I, P224S, T570N	A70V, V90M, P224S, R269K, V323A, K328R, D547E, R626K	A70V, I85T, V90M, P224S, R269K, V323A, K328R, D547E, R626K
Percentage of identity (%)	99.6 (99.3)	99.7 (99.6)	99.6 (99.4)	99.6 (99.6)	99.7 (99.7)	99.6 (99.4)	99.2 (98.7)	99.1 (98.6)
NP								
Amino acid changes	V100I, V476L	V100I	V100I	V100I	V100I	V100I	V100I	V100I
Percentage of identity (%)	99.5 (99.4)	99.7 (99.6)	99.7 (99.6)	99.6 (99.6)	99.7 (99.6)	99.7 (99.6)	99.4 (99.6)	99.5 (99.6)
NA								
Amino acid changes	V106I, N248D	V106I, N248D	V106I, N248D	V106I, N248D	V106I, N248D	V106I, N248D, N248D, D284S, N386D	V106I, G147R, N248D, R257K, I365T, N386S, I389M, I436V	V106I, N248D, R257K, I288V, I365T, N386S, I389M, I436V
Percentage of identity (%)	99.3 (99.1)	99.6 (99.4)	99.6 (99.4)	99.5 (99.4)	99.6 (99.4)	99.4 (98.9)	98.8 (98.3)	98.9 (98.3)
MI								
Amino acid changes				S275P		H256Y, F288S		
Percentage of identity (%)	99.8 (100)	99.7 (100)	99.7 (100)	99.7 (99.7)	99.7 (100)	99.6 (99.4)	99.6 (100)	99.6 (100)
NS								
Amino acid changes	I123V, R148K	I123V, N127D	I123V	E72K, I123V	M93I, I123V	I123V	T76A, R108K, I123V	T76A, R108K, I123V
Percentage of identity (%)	99.5 (98.4)	99.6 (99.2)	99.9 (99.6)	99.3 (98.8)	99.7 (99.6)	99.7 (99.6)	99.2 (98.8)	99.0 (98.8)

1: A/Tianjinheping/SWL313/2009(H1N1), 2: A/Tianjintangu/SWL432/2009(H1N1), 3: A/Tianjininghe/SWL49/2009(H1N1), 4: A/Tianjinjinan/SWL4141/2009(H1N1), 5: A/Tianjinnankai/SWL49/2010(H1N1), 6: A/Tianjinnankai/SWL413/2010(H1N1), 7: A/Tianjinhedong/SWL44/2011(H1N1), 8: A/Tianjinhedong/SWL44/2011(H1N1), 9: A/Tianjinhedong/SWL44/2011(H1N1)

hosts [15]. The transmission ability of the influenza virus is mainly determined by the receptor-binding sites (RBS), which are defined by three structural elements, a 130-loop (135–138) and a 190-helix site (190–198) and a 220-loop (221–228) (numbering on the basis of H3 HA sequence) [16, 17]. Compared with A/California/07/2009, our virus isolates were conserved at the 130-loop, 190-helix sites: VTAA (135–138) at loop 130 and DQSLYQNA (190–198) at helix 190. Only one substitution occurred at loop 220 (PKVRDQEG PKVRNQEG) in one HA sequence (A/Tianjinninghe/SWL49/2009(H1), which also had D239N substitution).

The HA gene of the H1N1pdm09 viruses was reportedly derived from “classical swine H1N1 virus”, which likely shares a common ancestor with the human H1N1 virus that caused the influenza pandemic in 1918, and whose descendant viruses continued to circulate in the human population with highly altered antigenicity of HA [18]. Some research about 1918 influenza virus reports that three amino acid mutations in 190, 225 and 226 (numbering on the basis of H3 HA sequence) sites can cause a switch in receptor-binding preference from the avian α -2,3 to the human α -2,6 sialic acid which results in the virus host range change [19, 20]. 190D, 239D and 240Q (recognized as 225D and 226Q if numbering on the basis of H3 HA sequence) residues in Tianjin H1N1pdm09 HA protein indicated that the virus bound preferentially to human host α 2-6-linked sialyl receptor and can spread in human population [19, 20].

No significant association between the increased amino acid substitution and clinical severity was observed in this study, but only limited sequence data were obtained, so further sequencing and case epidemiology are essential to make statements about this issue. But we found that three-residue mutations from severe infection viruses located at the antigenic domain [D239N (isolated in 2009), which has reportedly been associated with severe cases [13], S179N and K180T (isolated in 2011)] occurred, while none was found from mild symptom infection viruses. We also noticed that substitution sites increased gradually with time (from 2009 to 2012), the substitution rates of eight proteins (except M2 and NEP) were higher (from 7.39×10^{-4} to 7.40×10^{-3} per site per year), “antigenic drift” may have occurred in H1N1pdm09 with its ongoing evolving. Relatively high d_N/d_S ratios per site (especially PA, HA, NA and NS1) were observed, most likely reflecting different selection pressure at eight gene segments. The high evolutionary rate in H1N1pdm09 might be due to adaption to new host species, indicating strong selective advantages [21].

Single-nucleotide polymorphisms at codons encoding L26F, V27A, A30T S31N and G34E in the M2 gene fragment confer resistance to amantadine, and the most common substitution occurred at S31N [22]. In this study, 26, 27, 30 and 34 were conserved in all 8 H1N1pdm09,

only residue 31 was N, which indicated the virus was resistant to amantadine.

Neuraminidase inhibitors (such as oseltamivir and zanamivir) were the first medicinal choice against H1N1pdm09 influenza, but oseltamivir-resistant H1N1pdm09 strains were reported continually in different countries [23]. H275Y and N295S (N1 numbering) substitutions in the NA gene were confirmed that can lead to oseltamivir resistance. None of the H1N1pdm09 sequenced in this study had these two amino acid changes, but they all had N248D, which was also found in southern China [24]. Pan et al. [25] indicated that 248 site was adjacent to 275, and maybe associated with oseltamivir resistance. But another different opinion emerged from a Japanese researcher that N248D had little effect on E227 through crystal structure analysis; it may not alter the effect of oseltamivir to the virus [26].

One of the main concerns derived from our analysis would be a possibility that identified and forthcoming novel HA mutations may cause an “antigenic drift” which is sufficient to diminish the protective effect of vaccination against H1N1pdm09 in a significant proportion of vaccinated populations. This concern seems relevant because the viral strain utilized for vaccine development (the influenza A/California/7/2009 strain) does not carry the mutant form of the HA protein [27, 28].

Acknowledgments This work was supported by Tianjin Municipal Science and Technology Commission (grant No. 07SYSF05100), and Tianjin Centers for Disease Control and Prevention.

References

1. K.G. Nicholson, J.M. Wood, M. Zambon, *Influenza*. *Lancet* **362**, 1733 (2003)
2. S.U. Schnitzler, P. Schnitzler, *Virus Genes* **39**, 279 (2009)
3. NovelSwine-OriginInfluenzaA (H1N1) Virus investigation team, F.S. Dawood, S. Jain, L. Finelli, M.W. Shaw, S. Lindstrom, R.J. Garten, L.V. Gubareva, X. Xu, C. B. Bridges, T.M. Uyeki, *N. Engl. J. Med.* **360**, 2605 (2009)
4. G. Neumann, T. Noda, Y. Kawaoka, *Nature* **459**, 931 (2009)
5. H. Zaraket, H. Kondo, C. Tabet, R. Hanna-Wakim, Y. Suzuki, G.S. Dbaibo, R. Saito, H. Suzuki, *J. Clin. Virol.* **51**, 170 (2011)
6. G.J. Smith, D. Vijaykrishna, J. Bahl, S.J. Lycett, M. Worobey, O.G. Pybus, S.K. Ma, C.L. Cheung, J. Raghvani, S. Bhatt, J.S. Peiris, Y. Guan, A. Rambaut, *Nature* **25**, 1122 (2009)
7. C. Bin, L. Xingwang, S. Yuelong, J. Nan, C. Shijun, X. Xiayuan, W. Chen, *Emerg. Infect. Dis.* **15**, 1418 (2009)
8. A.J. Blaschke, M.A. Allison, L. Meyers, M. Rogatcheva, C. Heyrend, B. Mallin, M. Carter, B. Lafleur, T. Barney, M.A. Poritz, J.A. Daly, C.L. Byington, *J. Clin. Virol.* **3**, 210 (2011)
9. CDC, WHO, CDC Protocol for real-time RTPCR for influenza A(H1N1)—revision 1 30 April 2009 (2009)
10. M. Graham, B. Liang, G. Van Domselaar, N. Bastien, C. Beaudoin, S. Tyler, B. Kaplen, E. Landry, NationalInfluenzaA/H1N1pdm genomics study team (NIGST). *PLoS ONE* **6**, e16087 (2011)

11. M. Nelson, D. Spiro, D. Wentworth, E. Beck, J. Fan, E. Ghedin, R. Halpin, J. Bera, E. Hine, K. Proudfoot, T. Stockwell, X. Lin, S. Griesemer, S. Kumar, M. Bose, C. Viboud, E. Holmes, K. Henrickson, PLoS Curr. RRN1126 (2009)
12. S.M. Luoh, M.W. McGregor, V.S. Hinshaw, J. Virol. **66**, 1066 (1992)
13. A. Kilander, R. Rykkvin, S.G. Dudman, O. Hungnes, Euro. Surveill. **15**, 19498 (2010)
14. H. Chen, X. Wen, K.K. To, P. Wang, H. Tse, J.F. Chan, H.W. Tsoi, K.S. Fung, C.W. Tse, R.A. Lee, K.H. Chan, K.Y. Yuen, J. Infect. Dis. **201**, 1517 (2010)
15. Y. Liu, R.A. Childs, T. Matrosovich, S. Wharton, A.S. Palma, W. Chai, R. Daniels, V. Gregory, J. Uhlenhorff, M. Kiso, H.D. Klenk, A. Hay, T. Feizi, M. Matrosovich, J. Virol. **84**, 12069 (2010)
16. J. Cinatl Jr, M. Michaelis, H.W. Doerr, Med. Microbiol. Immunol. **196**, 181 (2007)
17. S.J. Gamblin, L.F. Haire, R.J. Russell, D.J. Stevens, B. Xiao, Y. Ha, N. Vasisht, D.A. Steinhauer, R.S. Daniels, A. Elliot, D.C. Wiley, J.J. Skehel, Science **303**, 1838 (2001)
18. R.J. Garten, C.T. Davis, C.A. Russell, B. Shu, S. Lindstrom, A. Balish, W.M. Sessions, X. Xu, E. Skepner, V. Deyde, M. Okomo-Adhiambo, L. Gubareva, J. Barnes, C.B. Smith, S.L. Emery, M.J. Hillman, P. Rivaller, J. Smagala, M. de Graaf, D.F. Burke, R.A. Fouchier, C. Pappas, C.M. Alpuche-Aranda, H. López-Gatell, H. Olivera, I. López, C.A. Myers, D. Faix, P.J. Blair, C. Yu, K.M. Keene, P.D. Dotson Jr, D. Boxrud, A.R. Sambol, S.H. Abid, K. St George, T. Bannerman, A.L. Moore, D.J. Stringer, P. Blevins, G.J. Demmler-Harrison, M. Ginsberg, P. Kriner, S. Waterman, S. Smole, H.F. Guevara, E.A. Belongia, P.A. Clark, S.T. Beatrice, R. Donis, J. Katz, L. Finelli, C.B. Bridges, M. Shaw, D.B. Jernigan, T.M. Uyeki, D.J. Smith, A.I. Klimov, N.J. Cox, Science. **325**, 197 (2009)
19. G.N. Rogers, J.C. Paulson, R.S. Daniels, J.J. Skehel, I.A. Wilson, D.C. Wiley, Nature **304**, 76 (1983)
20. T.M. Tumpey, T.R. Maines, N. Van Hoeven, L. Glaser, A. Sol-órzano, C. Pappas, N.J. Cox, D.E. Swayne, P. Palese, J.M. Katz, A. García-Sastre, Science **315**, 655–659 (2007)
21. Y. Qu, R. Zhang, P. Cui, G. Song, Z. Duan, F. Lei, Virol J. **8**, 250 (2011)
22. R.A. Bright, D.K. Shay, B. Shu, N.J. Cox, A.I. Klimov, JAMA **295**, 891 (2006)
23. Z. Zonis, D. Engelhard, M. Hindiyeh, D. Ram, M. Mandelboim, E. Mendelson, D. Glikman, Emerg. Infect. Dis. **16**, 1045 (2010)
24. A. Farooqui, Y. Lei, P. Wang, J. Huang, J. Lin, G. Li, A.J. Leon, Z. Zhao, D.J. Kelvin, J Infect Dev Ctries. **5**, 700 (2011)
25. C. Pan, B. Cheung, S. Tan, C. Li, L. Li, S. Liu, S. Jiang, PLoS ONE **5**, e9549 (2010)
26. J.É. Morlighem, S. Aoki, M. Kishima, M. Hanami, C. Ogawa, A. Jalloh, Y. Takahashi, Y. Kawai, S. Saga, E. Hayashi, T. Ban, S. Izumi, A. Wada, M. Mano, M. Fukunaga, Y. Kijima, M. Shiomi, K. Inoue, T. Hata, Y. Koretsune, K. Kudo, Y. Himeno, A. Hirai, K. Takahashi, Y. Sakai-Tagawa, K. Iwatsuki-Horimoto, Y. Kawaoka, Y. Hayashizaki, T. Ishikawa, PLoS ONE **6**, e18956 (2011)
27. M.E. Greenberg, M.H. Lai, G.F. Hartel, C.H. Wichems, C. Gittleson, J. Bennet, G. Dawson, W. Hu, C. Leggio, D. Washington, R.L. Bassar, N. Engl. J. Med. **361**, 2405 (2009)
28. T.W. Clark, M. Pareek, K. Hoschler, H. Dillon, K.G. Nicholson, N. Groth, I. Stephenson, N. Engl. J. Med. **361**, 2424 (2009)