Acta virologica **55**: 195 – 202, 2011 doi:10.4149/av_2011_03_195

Analysis of antigen epitopes and molecular pathogenic characteristics of the 2009 H1N1 pandemic influenza A virus in China

J.J. ZHOU[1,2#], J. TIAN[1#], D.Y. FANG[1,2], Y. LIANG[1,2], H.J. YAN[1,2], J.M. ZHOU[1,2], H.L. GAO[1], C.Y. FU[1], Y. LIU[1], H.Z. NI[3], C.W. KE[3,*], L.F. JIANG[1,2,*]

[1]Department of Microbiology, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, P.R.China; [2]Key Laboratory of Tropical Disease Control (Ministry of Education), Sun Yat-sen University, Guangzhou, P.R. China; [3]Microbiology Laboratory, Center for Disease Control and Prevention of Guangdong Province, Guangzhou, P.R. China

Received January 27, 2011; accepted June 27, 2011

Summary. – In order to further predict the epidemic trend and develop vaccines for 2009 H1N1 virus, we monitored its epitopes and molecular pathogenic characteristics during the epidemic process. We also analyzed the similarity of antigenic and genetic characteristics among the novel 2009 H1N1, representative seasonal H1N1 strains, and vaccine strains. 2009 H1N1 isolates had high similarity of hemagglutinin (HA) antigenic sites with H1N1 viruses isolated before 1940 and up to 80.0% similarity with 1918 H1N1. The elderly people born before 1940 have relatively low 2009 H1N1 infection rate, which might be responsible for their previous infection with either 1918 H1N1 virus or an early progeny. Compared to seasonal H1N1 vaccine strains from 1999 to 2010, the HA, neuraminidase (NA), and nucleoprotein (NP) proteins of the isolates had highly conserved CTL epitopes (60.5–65.8%, 69.6–82.6%, and 76.7%, respectively). The seriousness and mortality rate of 2009 H1N1 infections were similar to seasonal influenza, which may be related to the molecular characteristics of low toxicity of 2009 H1N1 and cross-T-cell immunity, due to vaccination or exposure to seasonal H1N1 virus. Some strains of 2009 H1N1 acquired mutations at antigenic and glycosylation sites. It is of particular interest that Haishu/SWL110/10 and Beijing/SE2649/09, isolated after November 2009, gained a new glycosylation site at the position 179 of HA protein, near the RBD. Thus, in the future, vaccination with glycosylated 2009 H1N1 virus may prevent the seasonal epidemic caused by strains with glycosylation site mutation near the receptor binding domain (RBD).

Keywords: antigen epitope; glycosylation site; influenza A virus; 2009 H1N1

Introduction

There have been three global influenza pandemics in the 20th century: the "Spanish flu" caused by 1918 H1N1, the "Asian flu" caused by 1957 H2N2, and the "Hong Kong flu" caused by 1968 H3N2. In 1977, the H1N1 subtype reappeared together with the H3N2 influenza virus; both strains were relatively prevalent and, until now, had regularly resurfaced as seasonal epidemics (Palese, 2004; Neumann *et al*., 2009).

In March 2009, a novel influenza A H1N1 virus appeared and spread rapidly to many countries and regions worldwide, developing into a global influenza pandemic (http://www.who.int). Studies showed that the 2009 H1N1 was a novel swine-origin recombinant virus, of which 6 genes (PB2, PB1, PA, HA, NP, and NS) were similar to triple-reassortant swine influenza viruses circulating in pigs in North America, while two genes (NA and M) were similar to the versions seen in influenza strains common to Eurasian swine (Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team, 2009; Smith *et al*., 2009; Trifonov *et al*., 2009).

To date, this novel cross-species transmitted virus – 2009 H1N1 – has been epidemic in the population for more than

*Corresponding authors. E-mail: Jianglf@mail.sysu.edu.cn; kecw1965@yahoo.com.cn; fax: +8620-87332368. #These authors contributed equally to this work.

Abbreviations: CTL = cytotoxic T-lymphocyte; GD06 = Guangdong/06/09; HA = hemagglutinin; NA = neuraminidase; NP = nucleoprotein; RBD = receptor binding domain

a year. Following steps will play an important role in the prediction of its epidemic trend and in vaccine development: 1) analysis of the similarity of antigenic and genetic characteristics among the 2009 H1N1 strains, representative seasonal H1N1 strains from different periods, and vaccine strains; 2) monitoring of its antigenic and molecular pathogenic characteristics in the epidemic process; and 3) improving the understanding of its genetic and variation rules. We therefore isolated, identified and analyzed the genome sequence of the 2009 H1N1 pandemic strains in Guangdong, China. We also analyzed the epitope, genetic and evolutionary characteristics of 2009 H1N1 isolated in China, in other countries and regions, seasonal H1N1 virus strains from different periods, and vaccine strains.

Materials and Methods

Isolation and identification of the virus. Nasopharyngeal swabs of influenza-suspected patients were taken and analyzed using a novel H1N1 real-time PCR kit (Daan, China). Since this study was part of the public health program implemented by the Ministry of Health of the People's Republic of China, there was no need of ethical approval. Positive samples were inoculated into MDCK cells and maintained in Opti-MEM minimum essential medium (Invitrogen, USA) containing 0.5 mg/ml trypsin. The cells were cultured at 35°C in the presence of 5% CO_2. After significant cytopathic effect occurred, hemagglutination and hemagglutination inhibition identification tests were performed (Kendal *et al.,* 1982).

Nucleotide sequencing. The RNeasy mini kit (Qiagen, Germany) was used to extract viral RNA from cell culture supernatants. Reverse transcription of viral cDNA was performed using random primers (N6). Forty-five pairs of primers were designed to amplify the entire viral genome (Hoffmann *et al.*, 2001; Wang and Lee, 2009; Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team, 2009). A gel extraction kit (Omega, USA) was used for the isolation of amplified PCR products. Sequencing was performed using an ABI sequencer (Applied Biosystems, USA).

Analyses of genetic and evolutionary characteristics. Protein sequences of 2009 H1N1 from different countries and regions obtained before August 2010 were downloaded from the influenza virus genome database; protein sequences of representative seasonal H1N1 strains from different periods, vaccine strains and swine influenza viruses were also downloaded. Clustal X and Mega 4.0 software were used to analyze variations in glycosylation sites, virulence-related sites, and antigenic sites of H1N1 viruses. 3D-JIGSAW was used for online prediction of 3D structures of viral HA and NA proteins.

Epitope prediction and analysis of conservation. The sequences of H1N1 viruses were input into the Immune Epitope Database (IEDB) and the artificial neural network (ANN) method was used for online prediction of HLA Class I epitope (CTL epitope). The unit of CTL epitope prediction measurement was IC50 (nM); lower IC50 values represented higher affinity. Typically, IC50 values of less than 50 nM represented high binding capacity, 50–500 nM represented moderate binding capacity, and higher than 500 nM represented low binding capacity (http://tools.immuneepitope.org).

CTL epitopes that were both nine amino acids long and had IC50 values of less than 500 were selected. The epitope conservancy analysis tool was used to analyze cross-conserved CTL epitopes among the 2009 H1N1 viruses, representative seasonal H1N1 strains and vaccine strains. Some T-cell epitopes show allele promiscuity; epitopes with up to two altered amino acids out of any 9-mer frame can still bind with the allele groove of HLA and the original T-cell receptor (Lichterfeld *et al.*, 2006). When evaluating the epitope conservation, we set the sequence similarity to ≥7/9 (77.8%).

Results

Virus isolation and sequencing

135 strains of 2009 H1N1 were isolated during the epidemic period. These isolates were named according to the isolation dates. We selected the representative epidemic strains from the different periods and sequenced the full-length genome of 2009 H1N1. The sequences have been submitted to GenBank, with accession numbers of GQ244322-GQ244325, GQ223434-GQ223447, HM780470-HM780492, and HQ011407-HQ011424.

Analysis of glycosylation sites

The HA protein of 2009 H1N1 virus had seven glycosylation sites at positions 28, 40, 104, 293, 304, 498, and 557. These glycosylation sites were located on the lateral side of HA molecule, with no glycosylation sites near the RBD. However, further analysis indicated that Haishu/SWL110/10 and Beijing/SE2649/09 from China, and Russia/178/09, Russia/180/09 and Salekhard/01/09 from other regions, isolated after November 2009, gained a novel glycosylation site at the position 179 near the RBD. Interestingly, the 1918 H1N1 virus also has no glycosylation sites near the RBD, but the seasonal H1N1 viruses gained in the epidemic process at least two glycosylation sites near the RBD of their HAs, such as at positions 144 and 179 in Marton/43, at positions 142, 172, and 177 in Denver/1/57, and at positions 142 and 177 in Guangzhou/506/06 (Fig. 1).

The NA protein of 2009 H1N1 virus had seven glycosylation sites at positions 50, 58, 63, 68, 88, 146, and 235. The strains isolated in China have no mutation at the glycosylation sites. However, Rome/623/09, Texas/45103737/09, and Texas/45071344/09 isolated in other regions after September 2009 had a new glycosylation site at the position 44. We also

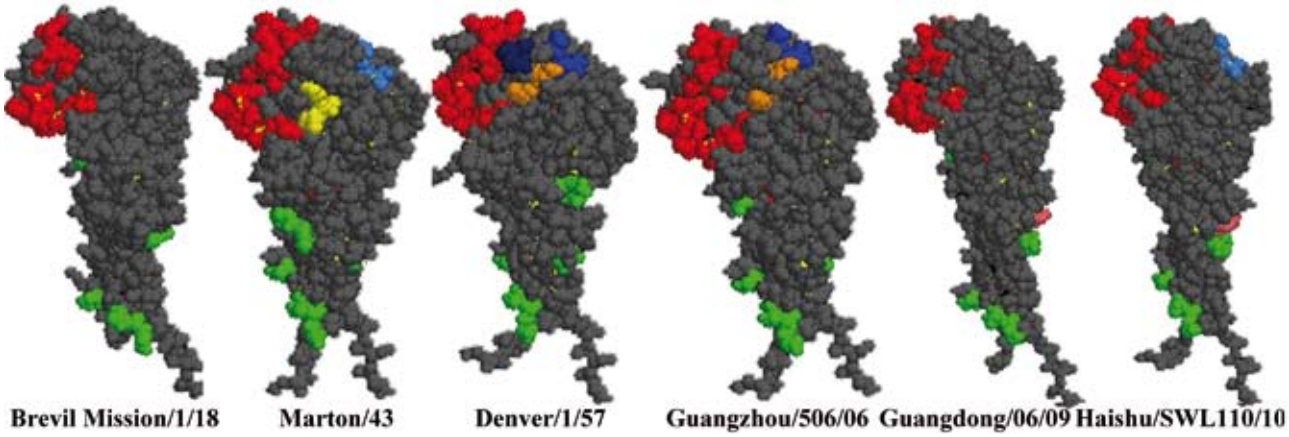

Fig. 1

Analysis of glycosylation sites of H1N1 influenza A virus HA protein

The receptor binding domains (RBD) of the HA protein are shown in red. Glycosylation sites on the side and the stem of HA protein are shown in green. Marton/43 has two new glycosylation sites at the position 144 (yellow), and the position 179 (sky blue) near the RBD of the head of HA protein. Denver/1/57 has three new glycosylation sites, at positions 142 (gold), 172 (dark blue) and 177 (blue) near the RBD. Guangzhou/506/06 had two glycosylation sites, at position 142 (gold), and at position 177 (blue) near the RBD. Apart from a novel glycosylation site at position 293 (pink), the 2009 H1N1 strains, such as Guangdong/06/09, had the same glycosylation sites as Brevil Mission/1/18, and no glycosylation sites were found near the RBD. However, several 2009 H1N1 viruses isolated after November 2009 (Haishu/SWL110/10, etc), had a new glycosylation site at position 179 (skyblue) near the RBD of the head of the HA protein.

found that 1918 H1N1 has no glycosylation site at the position 44, but seasonal H1N1 viruses isolated after 1933 gained a glycosylation site at the position 44 in the epidemic process.

Analysis of antigenic sites

X-ray crystal diffraction and monoclonal antibody analysis showed that the head of the H1N1 HA protein contained five antigenic sites: Cb (residues 78–83), Sa (residues 128–129, 156–160, and 162–167), Sb (residues 187–198), Ca1 (residues 169–173, 206–208, and 238–240), and Ca2 (residues 140–145 and 224–225) (H3 numbering) (Brownlee and Fodor, 2001). Analysis of antigenic sites showed that some 2009 H1N1 isolates had mutations at antigenic sites of the HA protein. These variations were primarily random mutations, but at two sites a variation trend was identified: at the Ca1 site, a Ser206Thr mutation was found in most epidemic strains, and at the Ca2 site, a Asp to Glu/Gly mutation of residue 225 was identified. These variations were minor and mainly restricted to changes of a single amino acid at a single antigenic site. However, there were several strains, which exhibited two or more mutated sites, such as the amino acid variations at Cb, Sa and Ca1 sites in Beijing/SE2649/09, and the amino acid variations at Cb, Ca1 and Ca2 sites in Zhejiang-Yiwu/11/09. Further, in several of these isolated strains, two amino acid mutations were found at the same antigenic site, such as the mutations of residues 78 and 81 at the Cb site in Zhejiang-Yiwu/11/09. The NA protein contains antigenic sites at residues 153, 197–199, 328–336, 339–347, 367–370, 400–403, and 431–434 (N2 numbering) (Colman *et al.*, 1983). The variation of NA antigenic sites of 2009 H1N1 strains was lower than observed in the HA protein. From 95 strains isolated from China, only 5 isolated strains, such as Guangdong/2361/09, showed mutations of one amino acid residue (Table 1).

A/Guangdong/06/09 (abbreviated as GD06), an early isolate from Guangdong, China, was used as a representative strain and compared with four seasonal H1N1 vaccine strains recommended by the WHO (BJ/262/95, NC/20/99, SI/3/06 and BS/59/07) in years 1999 to 2010. The results showed that the similarities between the antigenic sites of the HA proteins of GD06 and the four seasonal vaccine strains listed above were only 44.0%, 50.0%, 50.0%, and 50.0%, respectively. GD06 showed only 44–52% antigen similarity with representative seasonal H1N1 strains from 1940 and later. But it had a high degree of similarity of antigenic sites with Henry/36 (68.0%) and strains from before 1940, especially with antigenic sites of Brevil Mission/1/18 of the 1918 pandemic strain (80.0%). The similarity of NA antigenic sites between GD06, representative seasonal H1N1 strains, and vaccine strains resembled that observed for the HA protein (Table 2).

Prediction of CTL epitopes and analysis of conservation

CTL epitopes of H1N1 were restricted by HLA class I alleles. In this study, we used six HLA class I supertype alleles (A0101, A0201, A0301, A2402, B0702, and B4403) that are

Table 1. Analysis of the antigenic variation of HA and NA proteins of 2009 H1N1 pandemic influenza A virus in China

2009 H1N1 isolates	HA protein												NA protein				
	Cb		Sa		Sb			Ca1			Ca2		197-199	328-336	367-370	400-403	
	78	81	158	165	188	191	198	172	206	208	238	225	198	329	370	400	401
California/07/09	L	A	G	S	S	Q	A	K	S	R	E	D	D	N	S	I	N
Guangdong/06/09	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-
Beijing/SE2649/09	F	S	-	N	-	-	-	-	T	-	-	-	-	-	-	-	-
Zhejiang-Yiwu/11/09	F	S	-	-	-	-	-	-	T	-	-	G	-	-	-	-	-
Jiangyin/68/09	-	T	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-
Haishu/SWL110/10	-	-	-	N	-	-	-	-	T	-	-	-	-	-	-	-	-
Nanjing/1/10	-	-	E	-	-	-	-	-	T	-	-	G	-	-	-	-	-
Chengdu/18/09	-	-	-	-	-	R	-	-	T	-	-	-	-	-	-	-	-
Guangdong/1252/09	-	-	-	-	-	-	V	-	T	-	-	-	-	-	-	-	-
Nanjing/1/09	-	-	-	-	-	-	-	-	-	K	-	-	-	-	-	-	-
Guangdong/2221/09	-	-	-	-	-	-	-	-	T	-	-	E	-	-	-	-	-
Lishui/02w/09	F	S	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-
Hangzhou/1/09	F	S	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-
Changsha/78/09	-	-	-	-	-	-	-	-	T	-	-	E	-	-	-	-	-
Hunan-Kaifu/SWL4142/09	-	-	-	-	N	-	-	-	T	-	-	-	-	-	-	-	-
Dongcheng/SWL45/09	-	-	-	-	-	-	-	R	-	-	-	-	-	-	-	-	-
Chengdu/04/09	-	-	-	-	-	-	-	-	-	-	K	-	-	-	-	-	-
Guangdong/2361/09	-	-	-	-	-	-	-	-	T	-		-	-	-	L	-	-
Jiangsu/1/09	-	-	-	-	-	-	-	-	-	-			-	I	-	-	-
Guangdong/1105/09	-	-	-	-	-	-	-	-	-	-		-	-	-	-	-	H
Guangdong/1202/09	-	-	-	-	-	-	-	-	-			-	-		-	T	-
Nanjing/3/09	-	-	-	-	-	-	-	-	-	-		-	G	-	-	-	-

The antigenic sites of 2009 H1N1 strains isolated in China were analyzed and compared to the strain California/07/09 (2009 pandemic vaccine strain). Identical amino acid residues in the antigenic sites are indicated by "-".

Table 2. The amino acid homology and antigenic similarity of HA and NA proteins between GD06 and H1N1 representative strains from different periods and vaccine strains

H1N1 isolates	GD06 HA amino acid homology	GD06 HA antigenic similarity						GD06 NA amino acid homology	GD06 NA Antigenic similarity							
		Cb	Sa	Sb	Ca1	Ca2	Total (%)		153	197 | 199	328 | 336	339 | 347	367 | 370	400 | 403	431 | 434	Total (%)
California/07/09	99.1	6	13	12	10	8	49/50(98.0)	99.8	1	3	9	9	4	4	4	34/34 (100)
Brevig Mission/1/18	83.7	5	12	10	9	5	40/50(80.0)	87.2	1	3	8	9	3	2	4	30/34(88.2)
PR/8/34	81.1	2	8	7	7	4	28/50(56.0)	83.5	1	3	7	6	3	1	3	24/34(70.6)
Henry/36	82.3	3	10	8	8	5	34/50(68.0)	84	1	3	7	6	3	1	4	25/34(73.5)
Weiss/43	80.6	2	6	5	8	5	26/50(52.0)	82.5	0	2	6	6	3	1	3	21/34(61.8)
FM/1/47	80.2	2	8	4	7	4	25/50(50.0)	84.0	1	2	6	6	3	1	3	22/34(64.7)
Denver/1/57	79.1	2	7	3	6	5	23/50(46.0)	83.4	1	2	6	4	3	1	2	19/34(55.9)
New Jersey/8/76	90.6	4	12	11	9	5	41/50(82.0)	82.3	1	3	5	8	4	1	3	25/34(73.5)
USSR/90/77	79.3	2	8	3	7	3	23/50(46.0)	83.4	1	2	6	4	3	1	2	19/34(55.9)
Taiwan/01/86	79.5	2	9	4	7	3	25/50(50.0)	83.4	1	2	6	5	3	1	3	21/34(61.8)
BJ/262/95	79.5	2	7	3	7	3	22/50(44.0)	81.4	1	3	6	4	1	1	3	19/34(55.9)
NC/20/99	79.4	2	8	4	7	4	25/50(50.0)	81.4	1	3	6	4	1	2	3	20/34(58.8)
SI/3/06	79.3	2	8	4	7	4	25/50(50.0)	80.7	1	3	6	4	1	2	3	20/34(58.8)
BS/59/07	78.9	2	8	4	7	4	25/50(50.0)	80.9	1	3	6	5	1	2	3	21/34(61.8)
swine/Iowa/15/30	85.3	4	12	11	9	5	41/50(82.0)	87.6	1	3	7	9	4	2	4	30/34(88.2)
swine/Indiana/P12439/00	94.9	5	12	12	10	7	46/50(92.0)	/	/	/	/	/	/	/	/	/	
swine/Belgium/1/83	81.3	4	11	10	7	1	33/50(66.0)	92.5	1	3	8	9	3	2	4	30/34(88.2)

The HA protein has five antigenic sites, including 50 amino acid residues. The NA protein has seven antigenic sites, including 34 amino acid residues. Numbers before "/" show the total number of identical amino acid residues between GD06 and H1N1 representative strains.

Table 3. The analysis of the conservation of HLA Class I restricted epitopes of HA, NA, and NP proteins between GD06 and H1N1 representative strains from different periods and vaccine strains

H1N1 isolates	HA								NA								NP							
	A 01 01	A 02 01	A 03 01	A 24 02	B 07 02	B 44 03	Total	Conservation (%)	A 01 01	A 02 01	A 03 01	A 24 02	B 07 02	B 44 03	Total	Conservation (%)	A 01 01	A 02 01	A 03 01	A 24 02	B 07 02	B 44 03	Total	Conservation (%)
GD06	4	16	9	2	4	3	**38**		2	8	4	2	6	1	**23**		2	10	9	1	6	2	**30**	
California/07/09	4	16	9	2	4	3	**38**(38)	100	2	8	4	2	6	1	**23**(23)	100	2	10	9	1	6	2	**30**(30)	100
Brevig Mission/1/18	1	19	5	1	5	2	**33**(31)	81.6	2	11	5	0	5	0	**23**(18)	78.3	2	10	6	1	7	2	**28**(27)	90.0
PR/8/34	2	15	6	1	7	2	**33**(24)	63.2	1	7	6	0	5	1	**20**(15)	65.2	2	10	7	3	7	2	**31**(27)	90.0
Weiss/43	2	17	5	1	5	2	**32**(25)	65.8	3	9	4	0	3	2	**21**(13)	56.5	2	10	8	3	7	2	**32**(30)	100
FM/1/47	1	16	6	1	4	3	**31**(26)	68.4	3	8	2	0	4	1	**18**(11)	47.8	2	9	7	3	6	2	**29**(26)	86.7
Denver/1/57	3	15	5	1	7	3	**34**(23)	60.5	3	8	5	1	5	1	**23**(18)	78.3	3	11	8	4	6	2	**34**(31)	100
New Jersey/8/76	2	16	6	1	5	2	**32**(32)	84.2	2	14	6	1	4	1	**28**(19)	82.6	2	10	9	2	6	2	**31**(31)	100
USSR/90/77	1	17	4	1	4	3	**30**(26)	68.4	4	10	5	0	5	1	**25**(18)	78.3	3	10	6	4	6	2	**31**(27)	90.0
Taiwan/01/86	2	17	3	1	4	4	**31**(26)	68.4	2	9	5	0	5	1	**22**(16)	69.6	2	9	5	4	6	2	**28**(24)	80.0
BJ/262/95	3	17	3	1	4	4	**32**(23)	60.5	3	8	6	0	4	1	**22**(19)	82.6	2	9	5	3	6	2	**27**(23)	76.7
NC/20/99	2	17	4	1	4	4	**32**(24)	63.2	3	10	6	0	4	0	**23**(18)	78.3	2	8	5	3	6	2	**26**(23)	76.7
SI/3/06	3	17	4	1	5	5	**35**(25)	65.8	3	10	6	0	3	0	**22**(16)	69.6	/	/	/	/	/	/	/	
BS/59/07	2	17	4	1	4	5	**33**(24)	63.2	3	11	8	0	4	0	**26**(18)	78.3	/	/	/	/	/	/	/	
Taiwan/71720/07	2	17	4	1	4	5	**33**(23)	60.5	3	11	8	0	4	1	**27**(19)	82.6	2	8	5	3	6	2	**26**(23)	76.7

Numbers in bold represent the total of CTL epitopes restricted by the selected "supertype" HLA of H1N1 isolates. Numbers in ( ) are CTL epitopes restricted by the HLA supertype of H1N1 isolates also conserved in GD06. Conservation (%) represents the percentage of conserved CTL epitopes from the total number of CTL epitopes in GD06.

estimated to cover 99% of the population (de Groot and Berzofsky, 2004). The results showed that the number of CTL epitopes restricted by these six supertype alleles in HA, NA, and NP proteins of GD06 were 38, 23, and 30, respectively, of which the A0201 allele restricted the most. When compared to the 2009 H1N1 vaccine strain California/07/09, the GD06 epitopes showed 100% conservation.

Compared with seasonal H1N1 isolates from different periods, the conservation of GD06 NP CTL epitope was the highest (76.7–100%); the conservation of the HA CTL epitope was 60.5–84.2%; and the conservation of NA CTL epitope was comparatively low (47.8–82.6%). NP, HA, and NA CTL epitopes of GD06 shared the highest similarity with the New Jersey/08/76 of the classical swine H1N1 isolated from human, with similarities of 100%, 84.2%, and 82.6%, respectively. NP, HA, and NA CTL epitopes of GD06 were also highly similar to Brevil Mission/1/18, with similarities of 90.0%, 81.6%, and 78.3%, respectively (Table 3).

When compared to the vaccine strains (BJ/262/95, NC/20/99, SI/3/06, and BS/59/07), the conservation of CTL epitopes of the HA protein in GD06 was 60.5–65.8%, including 23 commonly conserved epitopes. Additionally, GD06 had two conserved epitopes in the HA protein: YPKLSKSYI, which was not conserved in BJ/262/95 but was conserved in the other three strains analyzed; and FYKNLIWLV, which was only conserved in SI/3/06. The conservation of the NA CTL epitopes was 69.6–82.6%, including 16 commonly conserved epitopes. In addition, GD06 has three conserved epitopes in the NA protein: NSDTVGWSW and WPDGAELPF were not conservative in SI/3/06, but conservative in the other three strains analyzed; and AELPFTIDK, only conservative in BJ/262/95. Because the NP protein sequence of SI/3/06 and BS/59/07 was not available, we selected Taiwan/71720/07 from the same period for comparison. GD06 had 23 common epitopes with two vaccine strains and Taiwan/71720/07 with the conservation of 76.7% (Table 4).

Discussion

T-cell epitopes (HLA class I and II epitopes) play an important role in the immunological response to the in-

Table 4. The HLA Class I epitope peptides of GD06 also conserved in the seasonal H1N1 vaccine strains (BJ/262/95, NC/20/99, SI/3/06, and BS/59/07)

HLA Class Allele	HA conserved epitopes	NA conserved epitopes	NP conserved epitopes
A0101	WTGMVDGWY ***	VSFNQNLEY **	CTELKLSDY **
		NSDTVGWSW *	HSNLNDATY **
A0201	TVLEKNVTV ***	LQIGNIISI ***	KLSDYDGRL *
	VLWGIHHPS *	ITYENNTWV **	RLIQNSITI **
	RMNYYWTLV *	NTWVNQTYV *	GMDPRMCSL ***
	FQNIHPITI *	AIYSKDNSV *	FLARSALIL ***
	AIDEITNKV *	LMSCPIGEV *	CLPACVYGL **
	KMNTQFTAV ***	CVNGSCFTV **	IMDSNTLEL *
	FLDIWTYNA ***	FSFKYGNGV **	FQGRGVFEL ***
	WTYNAELLV ***		KLLQNSQVV *
	TLDYHDSNV **		
	YQILAIYST ***		
	QILAIYSTV ***		
	ILAIYSTVA ***		
	TVASSLVLV **		
	SLGAISFWM ***		
A0301	KLRLATGLR *	VSGWAIYSK *	LMQGSTLPR ***
	RIYQILAIY *	KIFRIEKGK **	AMELIRMIK *
			RMCNILKGK ***
			ILRGSVAHK ***
			SLVGIDPFK **
A2402	FYKNLIWLV *	IWISHSIQL *	IFLARSALI **
	IYSTVASSL***	VWIGRTKSI**	
B0702	YPKLSKSYI *	EPFISCSPL **	DPKKTGGPI ***
	LPFQNIHPI *	SPLECRTFF **	LPFERATVM *
	VPSIQSRGL **	VPSPYNSRF *	LPRRSGAAG ***
		GPDNGAVAV ***	NPAHKSQLV ***
		WPDGAELPF ***	NPIVPSFDM ***
B4403	KEVLVLWGI **	AELPFTIDK ***	EEIRRVWRQ **
	QEGRMNYYW ***		AEIEDLIFL **
	LENERTLDY **		

YPKLSKSYI was not conservative in the HA of BJ/262/95 but conservative in the HAs of other three strains; FYKNLIWLV was only conservative in the HA of SI/3/06. NSDTVGWSW and WPDGAELPF were not conservative in the NA of SI/3/06 but conservative in the NAs of other three strains; AELPFTIDK was only conservative in the NA of BJ/262/95. *** showed 100% (9/9) conservation, ** 88.9% (8/9) conservation, * 77.8% (7/9) conservation.

fluenza virus. In particular, the HLA class I epitope (CTL epitope) can activate $CD8^+$ T-cells and is involved in both virus clearance and disease recovery (Boon *et al.*, 2004; Kreijtz *et al.*, 2007). When compared to seasonal H1N1 vaccine strains from 1999 to 2010, the GD06 antigenic sites of HA and NA proteins showed 44.0–50.0% and 55.9–61.8% similarity, respectively. However, CTL epitopes of HA and NA proteins were highly conserved (60.5–65.8% and 69.6–82.6%, respectively) when compared to vaccine strains. The CTL epitopes of the NP protein were 76.7% conserved. Thus the cross-T-cell immunity, acquired due to the vaccination or exposure to seasonal H1N1 virus, may alleviate the disease progress of 2009 H1N1. The severity and mortality of 2009 H1N1 infections were similar to seasonal influenza (Uyeki, 2009; Cowling *et al.*, 2010). This epidemic feature was related to the low-pathogenic molecular characteristics of the 2009 H1N1 virus, and might also be associated with conserved T-cell epitope-induced cross-cell immunity after the infection with seasonal H1N1 viruses or vaccination.

Antibodies against the HA protein of the influenza virus can neutralize infectivity of the virus and, therefore, play an important role in the protection against viral infections (Skehel and Wiley, 2000). The antigenic sites of the HA protein of 2009 H1N1 viruses were highly similar to H1N1 viruses isolated before 1940, in particular to the 1918 H1N1 virus (up to 80.0% similarity). Apart from the HA antigenic similarity, the glycosylation sites of the 2009 pandemic H1N1 virus were

very similar to the 1918 H1N1 virus, with no glycosylation sites near the RBD of the HA protein. Antibodies against the 1918 H1N1 virus showed good cross-neutralization of the 2009 H1N1 (Krause *et al.*, 2010; Manicassamy *et al.*, 2010). The high titer of cross-protective antibodies in people born before 1940, who are likely to have been previously infected with either 1918 H1N1 virus or an early progeny, might be responsible for their relatively low 2009 H1N1 infection rate.

Pandemic strain of influenza virus can evade the host immune response through the mutation of antigenic sites, enabling them to develop into seasonal epidemics. However, viral protein glycosylation, particularly the glycosylation near the RBD, can lead to an antigenic drift and escape from neutralizing antibodies against the virus by shielding antigenic sites. Thus, evasion of antibodies by glycosylation represents another mechanism for the development of seasonal flu from pandemic infections (Wang *et al.*, 2009; Cohen, 2010; Xu *et al.*, 2010). During the epidemic process of 2009 H1N1, some strains acquired mutations at antigenic sites; however, the variation of these sites was small, typically caused by the mutation of single amino acid at a single antigenic site. But it should be noted that several 2009 H1N1 strains had obvious variation at the glycosylation sites. Rome/623/09, Texas/45071344/09, and Texas/45103737/09, isolated after September 2009, gained new glycosylation site at the position 44 of the NA protein, the same as seasonal H1N1 viruses. It is of particular interest that Haishu/SWL110/10, Beijing/SE2649/09, Russia/178/09, Russia/180/09, and Salekhard/01/09, isolated after November 2009, gained a new glycosylation site at the position 179 of the HA protein near the RBD. Critically, the anti-serum against 2009 H1N1 California/04/09 cannot neutralize the infectivity of mutated strains with glycosylation at the position 179 of HA (Settembre *et al.*, 2010; Wei *et al.*, 2010). This trend of resistance through glycosylation should be monitored closely to avoid a seasonal epidemic of 2009 H1N1. Studies have shown that antibody prepared by immunization of mice using HA proteins glycosylated near the RBD of 2009 H1N1 can bind to the HA protein with or without glycosylation (Wei *et al.*, 2010). Thus, in the future strategy of vaccine development, vaccination with glycosylated 2009 H1N1 virus may prevent the seasonal epidemic caused by mutated strains with glycosylation near the RBD.

Acknowledgments. This study was supported by the Science and Technology Project of Guangdong province (Grant No. 2009B020600001), the Natural Science Foundation of Guangdong province (Grant No. 8451008901000445), the National Natural Science Foundation of China (Grant No. 81071367), and the special fund of Sun Yat-sen University (Grant No. 09ykpy82).

References

Boon ACM, de Mutsert G, van Baarle D, Smith DJ, Lapedes AS, Fouchier RAM, Sintnicolaas K, Osterhaus ADME, Rimmelzwaan GF (2004): Recognition of Homo-and Heterosubtypic Variants of Influenza A Viruses by Human CD8+ T Lymphocytes. J. Immunol. 172, 2453–2460.

Brownlee GG, Fodor E (2001): The predicted antigenicity of the haemagglutinin of the 1918 Spanish influenza pandemic suggests an avian origin. Phil. Trans. R. Soc. Lond. B. 356, 1871–1876. doiorg/10.1098/rstb.2001.1001

Cohen J (2010): Swine flu pandemic. What´s old is new: 1918 virus matches 2009 H1N1 strain. Science 327, 1563–1564. doiorg/10.1126/science.327.5973.1563

Colman PM, Varghese JN, Laver WG (1983): Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature 303, 41–44. doi.org/10.1038/303041a0

Cowling BJ, Chan KH, Fang VJ, Lau LLH, So HC, Fung ROP, Ma ESK, Kwong ASK, Chan CW, Tsui WWS (2010): Comparative epidemiology of pandemic and seasonal influenza A in households. N. Engl. J. Med. 362, 2175–2184. doiorg/10.1056/NEJMoa0911530

de Groot AS, Berzofsky JA (2004): From genome to vaccine – new immunoinformatics tools for vaccine design. Methods 34, 425–428. doiorg/10.1016/j.ymeth.2004.06.004

Hoffmann E, Stech J, Guan Y, Webster RG, Perez DR (2001): Universal primer set for the full-length amplification of all influenza A viruses. Arch. Virol. 146, 2275–2289. doiorg/10.1007/s007050170002

Kendal AP, Pereira MS, Skehel J (1982): Concepts and procedures for laboratory-based influenza surveillance. Geneva: World Health Organization.

Krause JC, Tumpey TM, Huffman CJ, McGraw PA, Pearce MB, Tsibane T, Hai R, Basler CF, Crowe JE, Jr. (2010): Naturally occurring human monoclonal antibodies neutralize both 1918 and 2009 pandemic influenza A (H1N1) viruses. J. Virol. 84, 3127–3130. doiorg/10.1128/JVI.02184–09

Kreijtz JH, Bodewes R, van Amerongen G, Kuiken T, Fouchier RA, Osterhaus AD, Rimmelzwaan GF (2007): Primary influenza A virus infection induces cross-protective immunity against a lethal infection with a heterosubtypic virus strain in mice. Vaccine 25, 612–620. doiorg/10.1016/j.vaccine.2006.08.036

Lichterfeld M, Williams KL, Mui SK, Shah SS, Mothe BR, Sette A, Kim A, Johnston MN, Burgett N, Frahm N (2006): T cell receptor cross-recognition of an HIV-1 CD8+ T cell epitope presented by closely related alleles from the HLA-A3 superfamily. Int. Immunol. 18, 1179–1188. doiorg/10.1093/intimm/dxl052

Manicassamy B, Medina RA, Hai R, Tsibane T, Stertz S, Nistal-Villán E, Palese P, Basler CF, García-Sastre A (2010): Protection of mice against lethal challenge with 2009 H1N1 influenza A virus by 1918-like and classical swine H1N1 based vaccines. PloS. Pathog. 6, 1–14. doiorg/10.1371/journal.ppat.1000745

Neumann G, Noda T, Kawaoka Y (2009): Emergence and pandemic potential of swine-origin H1N1 influenza virus. Nature 459, 931–939. doi.org/10.1038/nature08157

Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team (2009): Emergence of a novel swine-origin influenza A (H1N1) virus in humans. N. Engl. J. Med. 360, 2605–2615. doi.org/10.1056/NEJMoa0903810

Palese P (2004): Influenza: old and new threats. Nat. Med. 10 (Suppl. 1), S82–S87. doi.org/10.1038/nm1141

Settembre EC, Dormitzer PR, Rappuoli R (2010): H1N1: can a pandemic cycle be broken? Sci. Transl. Med. 2, 24ps14.

Skehel JJ, Wiley DC (2000): Receptor binding and membrane fusion in virus entry: The influenza hemagglutinin. Annu. Rev. Biochem. 69, 531–569. doi.org/10.1146/annurev.biochem.69.1.531

Smith GJD, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, Ma SK, Cheung CL, Raghwani J, Bhatt S (2009): Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature 459, 1122–1125. doi.org/10.1038/nature08182

Trifonov V, Khiabanian H, Rabadan R (2009): Geographic dependence, surveillance, and origins of the 2009 influenza A (H1N1) virus. N. Engl. J. Med. 361, 115–119. doi.org/10.1056/NEJMp0904572

Uyeki TM (2010): 2009 H1N1 virus transmission and outbreaks. N. Engl. J. Med. 362, 2221–2223. doi.org/10.1056/NEJMe1004468

Wang CC, Chen JR, Tseng YC, Hsu CH, Hung YF, Chen SW, Chen CM, Khoo KH, Cheng TJ, Cheng YSE (2009): Glycans on influenza hemagglutinin affect receptor binding and immune response. Proc. Natl. Acad. Sci. 106, 18137–18142. doi.org/10.1073/pnas.0909696106

Wang L, Lee CW (2009): Sequencing and mutational analysis of the non-coding regions of influenza A virus. Vet. Microbiol. 135, 239–247. doi.org/10.1016/j.vetmic.2008.09.067

Wei CJ, Boyington JC, Dai K, Houser KV, Pearce MB, Kong WP, Yang ZY, Tumpey TM, Nabel GJ (2010): Cross–neutralization of 1918 and 2009 influenza viruses: Role of glycans in viral evolution and vaccine design. Sci. Transl. Med. 2, 24ra21.

Xu R, Ekiert DC, Krause JC, Hai R, Crowe JE, Jr., Wilson IA (2010): Structural basis of preexisting im munity to the 2009 H1N1 pandemic influenza virus. Science 328, 357–360. doi.org/10.1126/science.1186430