Acta virologica **54**: 113 – 118, 2010 doi:10:4149/av_2010_02_113

Selection pressure on the hemagglutinin gene of Influenza A (H1N1) virus: adaptation to human and swine hosts in Asia

Z. XU, R. ZHOU, M. JIN, H. CHEN*

Division of Animal Infectious Disease, State Key Laboratory of Agricultural Microbiology, College of Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, P.R. China

Received August 17, 2009; accepted March 30, 2010

Summary. – Influenza A (H1N1) virus is an important pathogen that can be transmitted in the enzootic cycle involved in influenza epidemics. In this study, we analyzed natural selection on the 260 sequences of subtype 1 hemagglutinin (HA1) gene of H1N1 viruses isolated from humans and swine in Asia. Nonsynonymous and synonymous substitution rates (d_N, d_S) were determined for each amino acid site in the HA1 coding region. Four and two positively selected sites were identified in human and swine viruses, respectively. These sites might be involved in the change of antigenic epitopes to evade the host immune system. The results indicate that more diversifying and less purifying selection controls HA1 of human viruses compared with the swine viruses. Furthermore, thirteen sites evolving differently in human and swine viruses were identified in the HA1. These sites may play a role in the cross-species transmission. The analysis of adaptive evolution of the HA1 can provide valuable information for accelerating vaccine and drug development for prevention and treatment of influenza A (H1N1) in Asia.

Keywords: Influenza A virus; H1N1; hemagglutinin; positive selection; antigenic epitope

Introduction

Influenza viruses, the etiological agents of influenza (the family *Orthomyxoviridae*), are classified into three types A, B, C, and the type A viruses belong to the most virulent human pathogens (Webster *et al.,* 1982; Suzuki, 2006). Influenza A virus contains RNA genome in 8 segments encoding hemagglutinin (HA), neuraminidase (NA), nucleoprotein (NP), matrix proteins (M1 and M2), polymerase proteins (PB2, PB1, and PA) and nonstructural proteins (NS1 and NS2) (Webster *et al.,* 1992).

The swine influenza pandemic that occurred in April 2009, was a global outbreak caused by a new variant strain of influenza A virus subtype H1N1. Historically, the H1N1 virus caused disastrous Spanish influenza in 1918–1919 leading to approximately 50 million deaths (Taubenberger and Morens, 2006).

The HA protein of influenza A viruses is a surface glycoprotein. It has also become an important antigen as the major target of humoral immunity against Influenza A virus (Caton *et al.,* 1982). The mature HA can be cleaved into two subunits HA1 and HA2 by the host trypsin-like proteases (Webster *et al.,* 1992). HA of human influenza viruses that is under the host immune pressure is evolving much more rapidly than the internal protein genes PB1, PB2, PA, NP and M1 (Webster *et al.,* 1992; Suzuki, 2008). The evidence of positive selection has been reported for the HA1 genes of different subtypes such as H3N2 and H5N1 (Bush *et al.,* 1999; Campitelli *et al.,* 2006; Kumar *et al.,* 2006; Suzuki, 2008).

Estimation of the rates of nonsynonymous (d_N) to synonymous substitutions (d_S) per site has become a standard tool of the selection pressure: $d_N/d_S>1$ and $d_N/d_S<1$ indicate positive and negative selection, respectively (Hughes and Nei, 1988; Nielsen and Yang, 1998). These analyses have been widely

*Corresponding author. E-mail: chenhch@mail.hzau.edu.cn; fax: +8627-87281795.

Abbreviations: d_N = nonsynonymous substitutions per site; d_S = synonymous substitutions per site; FEL = fixed effects likelihood; HA1, 2 = hemagglutinin subunits 1, 2; ML = maximum likelihood; SLAC = single likelihood ancestor counting

Table 1. The list of 260 sequences of HA1 gene of influenza A (H1N1) viruses used in this study

No.	Acc. No	Virus	No.	Acc. No	Virus
1	U46783	A/swine/Beijing/47/1991	2	EU004452	A/swine/Henan/01/2006
3	EU502884	A/swine/Shanghai/1/2005	4	EU502885	A/swine/Shanghai/2/2005
5	EU004444	A/swine/Tianjin/01/2004	6	FJ415610	A/swine/Zhejiang/1/2007
7	X57491	A/swine/Hong Kong/1/1974	8	U44482	A/swine/Hong Kong/168/1993
9	U46020	A/swine/Hong Kong/172/1993	10	U45451	A/swine/Hong Kong/176/1993
11	U45452	A/swine/Hong Kong/273/1994	12	GQ229285	A/swine/Hong Kong/8512/2001
13	GQ229357	A/swine/Hong Kong/9656/2001	14	GQ229325	A/swine/Hong Kong/NS1179/2007
15	GQ229293	A/swine/Hong Kong/NS1659/2001	16	GQ229301	A/swine/Hong Kong/NS29/2009
17	GQ229261	A/swine/Hong Kong/NS837/2001	18	D28518	A/sw/Obihiro/5/1992
19	X57494	A/swine/Ehime/1/1980	20	AB434392	A/swine/Hokkaido/2/1981
21	AB434384	A/swine/Kyoto/3/1979	22	AB434400	A/swine/Niigata/1/1977
23	EU798778	A/swine/Korea/CAN01/2004	24	EU798779	A/swine/Korea/CAS08/2005
25	AB434328	A/swine/Chachoengsao/NIAH587/2005	26	EU296605	A/swine/Chonburi/06CB2/2006
27	AB434320	A/swine/Chonburi/NIAH589/2005	28	EU296599	A/swine/Chonburi/NIAH589/2005
29	AB434304	A/swine/Chonburi/NIAH9469/2004	30	AB434312	A/swine/Chonburi/NIAH977/2004
31	AB434296	A/swine/Ratchaburi/NIAH550/2003	32	FJ688266	A/swine/Thailand/HF6/2005
33	FJ375207	A/Cambodia/365/2007	34	FJ375205	A/Cambodia/0371/2007
35	FJ375206	A/Cambodia/0374/2007	36	FJ375208	A/Cambodia/502/2007
37	FJ743471	A/Cambodia/21/2007	38	GQ200287	A/Shandong/1/2009
39	GQ268003	A/GuangzhouSB/01/2009	40	GQ225365	A/Fujian/1/2009
41	GQ225357	A/Shanghai/1/2009	42	GQ293077	A/Zhejiang/2/2009
43	GQ166223	A/Sichuan/1/2009	44	GQ359758	A/Zhejiang/DTID-ZJU01/2009
45	GQ280264	A/Nanjing/1/2009	46	GQ232093	A/Beijing/4/2009
47	GQ221694	A/GuangzhouSB/01/2009	48	CY020573	A/Tientsin/78/1977
49	EU382981	A/Guangzhou/483/2006	50	EU382982	A/Guangzhou/506/2006
51	EU382985	A/Guangzhou/665/2006	52	EU382983	A/Guangzhou/555/2006
53	EU382984	A/Guangzhou/657/2006	54	EU382987	A/Guangzhou/1684/2006
55	EU382986	A/Guangzhou/1561/2006	56	AY289930	A/Shenzhen/227/1995
57	L19017	A/Qingdao/28/1991	58	L19024	A/Sichuan/4/1988
59	AJ457900	A/Beijing/262/1995	60	CY033614	A/Beijing/262/1995
61	AY289928	A/Beijing/262/1995	62	CY016228	A/Nanchang/8/1996
63	CY013853	A/Nanchang/26/1996	64	CY013813	A/Nanchang/13/1996
65	CY013821	A/Nanchang/15/1996	66	CY017011	A/Nanchang/14/1996
67	CY013829	A/Nanchang/16/1996	68	GQ351314	A/Hong Kong/2369/2009
69	GQ168606	A/Hong Kong/01/2009	70	CY009292	A/Hong Kong/117/1977
71	CY031342	A/Hong Kong/2652/2006	72	EU516329	A/HongKong/CUHK-13003/2002
73	AF386781	A/Hong Kong/1134/1998	74	AJ457885	A/Hong Kong/4847/1998
75	AJ457872	A/Hong Kong/747/2001	76	AJ457895	A/Hong Kong/470/1997
77	AJ457901	A/Hong Kong/59/1994	78	AJ457899	A/Hong Kong/52/1995
79	AF386776	A/Hong Kong/1035/1998	80	AF386775	A/Hong Kong/1131/1998
81	CY020453	A/India/6263/1980	82	CY031293	A/Iraq/AF1193/2007
83	CY031288	A/Iraq/AF1188/2007	84	CY031291	A/Iraq/AF1191/2007
85	CY031296	A/Iraq/AF1196/2007	86	CY031303	A/Iraq/AF1203/2007
87	CY031297	A/Iraq/AF1197/2007	88	CY031289	A/Iraq/AF1189/2007
89	CY041968	A/Israel/276/2009	90	GQ287619	A/Fukuoka-C/1/2009
91	GQ287625	A/Tokushima/1/2009	92	GQ261277	A/Shiga/1/2009
93	GQ334355	A/Utsunomiya/1/2009	94	GQ365446	A/Sapporo/1/2009
95	GQ261275	A/Sakai/2/2009	96	CY043078	A/Japan/1070/2009
97	AB043499	A/Yokohama/24/2000	98	CY034527	A/Japan/AF1884/2008
99	CY031469	A/Japan/AF1274/2008	100	CY031470	A/Japan/AF1275/2008
101	CY031467	A/Japan/AF1272/2008	102	CY031442	A/Japan/AF1278/2007
103	CY031468	A/Japan/AF1273/2008	104	CY031466	A/Japan/AF1271/2008
105	CY034522	A/Japan/AF1963/2007	106	CY031471	A/Japan/AF1276/2008
107	AB043493	A/Aichi/25/1996	108	AB043492	A/Kamata/69/1996
109	AB043491	A/Kamata/381/1995	110	L19014	A/Fukushima/2/1988
111	AB043488	A/Nagano/1669/1989	112	D31949	A/Yamagata/32/1989
113	D13574	A/Suita/1/1989/	114	D13573	A/Suita/1/1989
115	AB043490	A/Aichi/24/1992	116	AB043489	A/Nagano/92/1991
117	D00841	A/Yamagata/120/1986	118	AB043487	A/Kamata/85/1987
119	AB434105	A/Morioka/7/2005	120	AB434100	A/Morioka/3/2005
121	AB434102	A/Morioka/21/2005	122	AB434099	A/Morioka/17/2002
123	AB434103	A/Morioka/27/2002	124	AB043495	A/Yokohama/50/1998
125	DQ397950	A/Hiroshima/37/2001	126	AB043496	A/Aichi/94/1999
127	AB043498	A/Yokohama/12/2000	128	AY029292	A/Misawa/1226/00

No.	Acc. No	Virus	No.	Acc. No	Virus
129	AB043497	A/Aichi/102/1999	130	AB043500	A/Ibaraki/90/1998
131	AB043494	A/Kamata/159/1997	132	AB043481	A/Tokyo/1/51
133	AB043480	A/TF/15/1951	134	AB043486	A/Saga/2/1957
135	AB043482	A/Kojiya/1/1952	136	AB043484	A/Yamagishi/55
137	AB043485	A/Meguro/1/1956	138	S62154	A/Alma Ata/1417/84
139	CY028091	A/Tennessee/UR06-0523/2007	140	CY031350	A/Macau/189/2007
141	FJ899924	A/Macau/229/2008	142	EF566337	A/Macau/13076/2006
143	EF566338	A/Macau/14585/2006	144	EF566330	A/Macau/10258/2006
145	EF566237	A/Macau/227/2005	146	EF566236	A/Macau/122/2005
147	CY031354	A/Malaysia/330/2007	148	CY031368	A/Malaysia/862/2007
149	EF566325	A/Malaysia/37/2006	150	EF566324	A/Malaysia/36/2006
151	EF566326	A/Malaysia/38/2006	152	EF566327	A/Malaysia/56/2006
153	EF566358	A/Malaysia/88/2004	154	EF566126	A/Malaysia/1513/2004
155	EF566373	A/Malaysia/1513/2004	156	EF566042	A/Malaysia/643/2003
157	EF566041	A/Malaysia/1003/2003	158	CY009340	A/Malaysia/1954
159	CY021053	A/Malaya/302/1954	160	Z54287	A/Mongolia/153/88
161	Z54286	A/Mongolia/231/85	162	GQ243760	A/Philippines/2006/2009
163	GQ243749	A/Philippines/2009/2009	164	GQ243753	A/Philippines/2001/2009
165	FJ743469	A/Philippines/1159/2008	166	CY030867	A/Philippines/905/2006
167	EF566346	A/Philippines/1392/2006	168	CY030863	A/Philippines/1279/2006
169	EF566342	A/Philippines/906/2006	170	EF566347	A/Philippines/624/2006
171	CY031338	A/Philippines/673/2006	172	EF566125	A/Philippines/936/2004
173	EF566124	A/Philippines/987/2004	174	CY034523	A/Qatar/AF1966/2007
175	CY031305	A/Qatar/AF1205/2007	176	GQ392017	A/Singapore/TLL01/2009
177	CY031360	A/Singapore/43/2007	178	EF566374	A/Singapore/14/2004
179	CY030213	A/Singapore/69/2006	180	EF566354	A/Singapore/41/2006
181	CY030215	A/Singapore/36/2006	182	EF566319	A/Singapore/107/2005
183	EF566375	A/Singapore/23/2004	184	L20111	A/Singapore/10/1990
185	L20110	A/Singapore/10/1990	186	L20106	A/Singapore/03/1990
187	L19026	A/Singapore/6/1990	188	L20108	A/Singapore/06/1990
189	L20107	A/Singapore/03/1990	190	L20116	A/Singapore/12/1990
191	L20117	A/Singapore/12/1990	192	L20112	A/Singapore/11/1990
193	L20113	A/Singapore/11/1990	194	CY020477	A/Singapore/6/1986
195	D00406	A/Singapore/6/1986	196	AJ457896	A/Singapore/15/1996
197	EU304355	A/Busan/03/2007	198	CY031352	A/Jeju/2279/2007
199	EU304354	A/Busan/02/2007	200	CY031496	A/Korea/AF1307/2008
201	CY031493	A/Korea/AF1304/2008	202	CY034541	A/Korea/AF1899/2008
203	CY034540	A/Korea/AF1898/2008	204	CY031451	A/Korea/AF1312/2007
205	L33743	A/Seoul/20/1991	206	AY299494	A/Pusan/24/2002
207	AY299501	A/Seoul/15/2002	208	AY299506	A/Chungbuk/50/2002
209	AY297154	A/Chonnam/07/2002	210	AY299499	A/Gwangju/90/2002
211	AY299508	A/Gwangju/57/2002	212	AY299500	A/Seoul/13/2002
213	AY297157	A/Pusan/23/2002	214	AY299496	A/Pusan/45/2002
215	AY299507	A/Gwangju/55/2002	216	AY299509	A/Gwangju/58/2002
217	AY299495	A/Seoul/33/2002	218	CY031356	A/Sri Lanka/22/2007
219	DQ415318	A/TW/4845/1999	220	CY040058	A/Taiwan/71720/2007
221	CY040090	A/Taiwan/70013/2008	222	CY040106	A/Taiwan/70132/2008
223	EF566122	A/Taiwan/1559/2004	224	DQ415316	A/TW/130/1996
225	AF026160	A/Taiwan/118/1996-3	226	AF026159	A/Taiwan/118/1996-2
227	AF026158	A/Taiwan/118/1996-1	228	AF026156	A/Taiwan/117/1996-2
229	AF026155	A/Taiwan/117/1996-1	230	AF026153	A/Taiwan/112/1996-1
231	AF055426	A/Taiwan/2243+C44/1992	232	D00407	A/Taiwan/1/1986
233	X17224	A/Taiwan/1/1986	234	DQ508873	A/Taiwan/01/1986
235	EF566234	A/Taiwan/50/2005	236	DQ249260	A/Taiwan/2985/2002
237	CY040146	A/Taiwan/52/2002	238	CY040154	A/Taiwan/123/2002
239	CY040074	A/Taiwan/567/2002	240	CY040138	A/Taiwan/5072/1999
241	DQ415317	A/TW/3355/1997	242	AB043483	A/Taiwan/13/1954
243	EF101749	A/Thailand/271/2005	244	GQ150342	A/Nonthaburi/102/2009
245	CY031380	A/Thailand/486/2007	246	CY030872	A/Thailand/39/2008
247	EU021256	A/Thailand/CU57/2006	248	CY031382	A/Thailand/545/2007
249	EU021254	A/Thailand/CU51/2006	250	CY031344	A/Thailand/798/2006
251	EU021252	A/Thailand/CU88/2006	252	EF566351	A/Thailand/380/2006
253	EU021250	A/Thailand/CU67/2006	254	EU021246	A/Thailand/CU41/2006
255	EU021258	A/Thailand/CU44/2006	256	EU021262	A/Thailand/CU75/2006
257	EU021264	A/Thailand/CU32/2006	258	EF566235	A/Thailand/28/2005
259	EF566372	A/Bangkok/1406/2004	260	EF568930	A/Thailand/Siriraj-07/2000

used for the study of adaptive evolution of protein-coding genes of different viruses. In this way, positive selection has been detected in several systems, including Human immunodeficiency virus 1 genes, Hepatitis C virus subtype 1b polyprotein and St. Louis encephalitis virus envelope protein (Nielsen and Yang, 1998; Suzuki and Gojobori, 2001; Baillie *et al.,* 2008).

In this study, we analyzed natural selection on 260 coding sequences of HA1 gene of influenza A (H1N1) viruses isolated from humans and swine in Asia. By estimating the d_N/d_S ratio for each amino acid site we could identify the positively or negatively selected sites in the HA1 and hence the differences in the adaptation of the virus to human and swine host.

Materials and Methods

Sequence data sets. The sequence data of the HA1 gene of influenza A (H1N1) viruses isolated in Asia (updated in July, 2009) were retrieved from the Influenza Virus Resource database (http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html). Two data sets for human and swine viruses were compiled. The identical sequences were excluded from each data set. After removing the sequences derived from laboratory strains, 228 and 32 sequences originated from human and swine viruses, respectively, were used for the following analysis. Acc. Nos. of these sequences are listed in Table 1.

Detection of homologous recombination and positively/negatively selected sites. Multiple sequence alignments were performed using the program MUSCLE 3.6 and then manually examined (Edgar, 2004). Each sequence alignment consisted of 972 nt encoding the HA1 region (324 aa). Briefly, a phylogenetic tree for each data set was built using a maximum likelihood (ML) method implemented in PhyML 3.0 (Guindon and Gascuel, 2003). A substitution model HKY85 with the estimated ratio of transition/transversion rate (κ), gamma distributed rate heterogeneity of 4 categories ($\Gamma 4$) and a proportion of invariable sites was used in the tree reconstruction. The generated ML trees were used for estimating the site-by-site variation rates. Since recombination can lead to spurious results in the estimating of selection pressure, first we tested the recombination in each data set by using a genetic algorithm for recombination detection (GARD) (Kosakovsky *et al.,* 2006a). To deduce amino acid sites under positive/negative selection on HA1, we used two likelihood procedures available in HyPhy and DataMonkey: (i) single likelihood ancestor counting (SLAC) and (ii) the more sensitive fixed effects likelihood (FEL) (Kosakovsky *et al.*, 2005a, b). To take account of nucleotide biases that probably affect the accuracy of evaluating nonsynonymous and synonymous substitution rates, a best fitting model of the nucleotide substitution was estimated for each data set according to the Akaike's Information Criterion. The ratio d_N/d_S was calculated by a codon model obtained by crossing MG94 and the best nucleotide model (Muse and Gaut, 1994). Furthermore, a method for identifying sites that were selected differentially in two populations was implemented in the HyPhy package (Kosakovsky *et al.,* 2006b). The significance threshold for likelihood ratio tests in this study was set at 0.05.

Results and Discussion

The detection of recombination was performed for each data set of influenza A (H1N1) viruses, which were found to be free of homologous recombination events. After detection of recombination, the global ratios of d_N/d_S were estimated to be 0.172 and 0.324 in the HA1 coding region for swine and human viruses, respectively. The result indicated that HA1 genes may experience stronger positive selection in the process of adaptation to the human than to the swine. Particularly, the ratio d_N/d_S for HA1 of human H3N2 viruses was estimated to be approximately 0.3 (Suzuki, 2006). It indicated that the HA1 genes from H1N1 and H3N2 viruses are subject to the similar selection pressure in human host environment. Together with other authors (Raymond *et al.,* 1986; Sugita *et al.,* 1991), we can suppose that the HA1 region from different subtypes of Influenza A virus evolves more rapidly in the humans than in swine.

The results of positively selected sites in the HA1 region of influenza A (H1N1) viruses were summarized in Table 2 and the selection profiles inferred by FEL were illustrated in Fig. 1. For the viruses isolated from swine, d_S exceeded d_N in the majority of sites (263 sites; 81.2%), and the strong evidence for negative selection was detected in approximately one-third (127 sites; 39.2%) of sites. On the other hand, d_N exceeded d_S in only 35 sites (10.8%), and two sites 113 and 203 were found to be under significant positive selection (Fig. 1). By comparison, 238 sites (73.5%) where d_S exceeded d_N were identified in the HA1 region of human influenza viruses, among which 113 sites (34.9%) were negatively selected. Moreover, there were totally 57 sites (17.6%), where d_N exceeded d_S for human influenza viruses and four of them were found to be under positive selection (Table 2). The selection results of site-by-site together with the estimation of global ratios d_N/d_S suggested less purifying and more diversifying selection acting on HA1 of human viruses compared with swine viruses.

The positively selected sites are useful for the identification of immunodominant epitopes. The mutations at these sites may provide selective advantage to the mutants by allowing them to escape from the host immune system (Suzuki and Gojobori, 2001; Suzuki, 2006). Although the analysis of positive selection on the HA1 genes have been performed for human or swine influenza A (H1N1) viruses, no positively selected site was found in the HA1 region of swine H1N1 viruses, what is mainly caused by the limited sequence data (Sugita *et*

Table 2. Positively selected amino acid sites in the HA1 of influenza A (H1N1) viruses isolated from humans and swine in Asia

Virus	Number of sequences	Sites	
		SLAC	FEL
Human	228	239, 278	11, 204, 239, 278
Swine	32	113, 203	113, 203

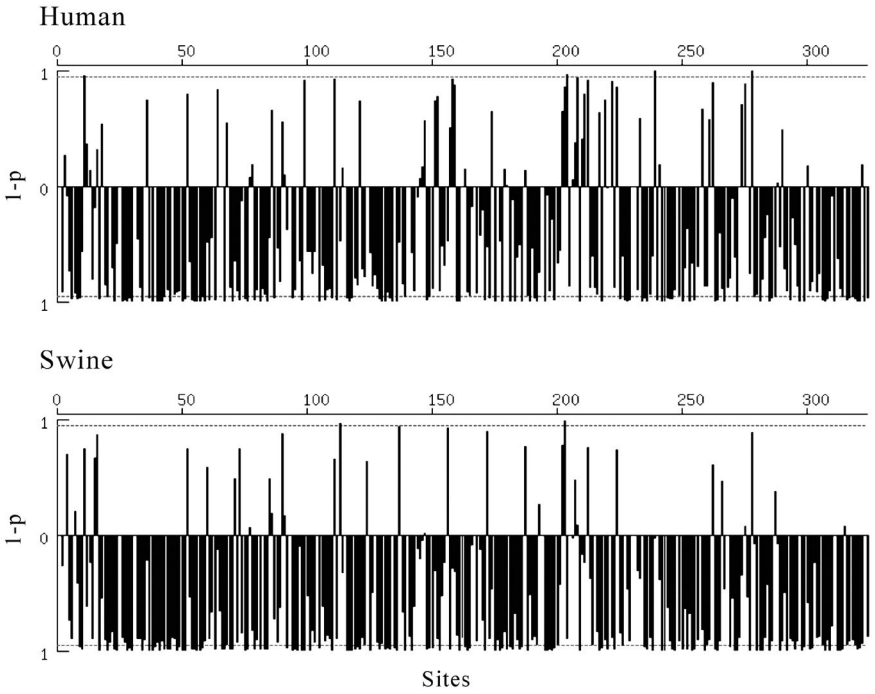

Fig. 1

Selection profiles of the coding region of HA1 gene of influenza A (H1N1) viruses isolated from humans and swine in Asia

Vertical bars above and below horizontal line represent $d_N/d_S > 1$ and $d_N/d_S < 1$, respectively. The dashed lines represent the significant threshold.

al., 1991; Ina and Gojobori, 1994). In our study, the adaptive evolution for HA1 of human and swine H1N1 viruses isolated in Asia was re-analyzed using a large data set presently available. The positively selected sites in the HA1 region of H1N1 viruses were detected in human as well as in swine viruses. Among these sites found under positive selection pressure, the site 239 is involved in the composition of Ca antigenic region (Caton *et al.,* 1982). according to the structure properties of HA protein, two sites 203 and 204 within the helix region are important for the interactions with receptor (Gamblin *et al.,* 2004). Except for the site 11 in the signal peptide region, the function of other positively selected sites 113 and 278 needs to be further investigated. We deduced that the frequent variations of these sites in the HA1 regions were probably responsible for the modification of antigenic epitopes and generation of the escape mutants.

To further explore the adaptive evolutionary differences between human and swine H1N1 viruses, we compared the profiles of divergent selection acting on the HA1 region. Thirteen out of 324 sites (4.0%) in HA1 were found to be under differential selection at sites 24, 73, 99, 113, 121, 142, 157, 180, 198, 204, 219, 222, and 239 ($p < 0.05$) (Table 3). In order to adapt to the human host environment, the majority of these sites in human H1N1 viruses were subjected to the strong positive selection pressure. Among the sites in HA1 regions that were selected differentially, there were one site (204) involved in the receptor binding and four sites (142, 157, 180 and 239) involved in the immunodominant Sa and Ca antigenic sites (Caton *et al.,* 1982; Gamblin *et al.,* 2004).

Table 3. Amino acid sites in the HA1 evolving differently in human and swine influenza A (H1N1) viruses in Asia

Site[a]	Human		Swine	
	d_S	d_N	d_S	d_N
24	1.065	0.405	12.782	0
73	3.622	0.276	0	0.785
99	0	1.612	8.589	0
113	3.735	2.502	0.848	5.663
121	0	0.956	7.623	0.621
142	3.510	1.906	10.497	0.379
157	0.605	1.242	2.490	0
180	3.213	3.241	3.343	0
198	1.022	0.646	5.141	0
204	3.608	12.278	4.657	0.794
219	0.548	1.686	2.414	0.379
222	0	1.702	7.041	1.366
239	0	6.66	2.292	2.225

[a]Amino acid site numbering is based on A/Shandong/1/2009 virus.

The remaining sites might be associated with the novel host-specific antigenic epitopes and require more experimental work to understand their function in virus infection. These sites under differential selection in the two host populations may play a potential role in the interspecies transmission of influenza A (H1N1) viruses.

In summary, four and two positively selected sites were identified in the HA1 region of human and swine influenza A (H1N1) viruses isolated in Asia, respectively. These positively selected sites may involve the transition of antigenic epitopes and in this way they facilitate influenza viruses to escape the host immune system. The estimated rates of nonsynonymous and synonymous substitutions showed more diversifying selection and less purifying selection acting on HA1 of human viruses compared with swine viruses. Furthermore, we identified 13 sites in the HA1 region that were selected differentially in human or swine, demonstrating the specific adaptation of H1N1 viruses to the particular host. These results provide valuable information for the development of vaccines and antiviral drugs for future prevention and treatment of the influenza A (H1N1) infections in Asia.

Acknowledgements. This work was supported by the grant No. 30530590 from National Natural Science Foundation of China and No. 2006BAD06A01 from National Scientific & Technical Supporting Program of China.

References

Baillie GJ, Kolokotronis SO, Waltari E, Maffei JG, Kramer LD, Perkins SL (2008): Phylogenetic and evolutionary analyses of St. Louis encephalitis virus genomes. Mol. Phylogenet. Evol. 47, 717–728. doi:10.1016/j.ympev.2008.02.015

Bush RM, Fitch WM, Bender CA, Cox NJ (1999): Positive selection on the H3 hemagglutinin gene of human influenza virus A. Mol. Biol. Evol. 16, 1457–1465.

Campitelli L, Ciccozzi M, Salemi M, Taglia F, Boros S, Donatelli I, Rezza G (2006): H5N1 influenza virus evolution: a comparison of different epidemics in birds and humans (1997–2004). J. Gen. Virol. 87, 955–960. doi:10.1099/vir.0.81397-0

Caton AJ, Brownlee GG, Yewdell JW, Gerhard W (1982): The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell 31, 417–427. doi:10.1016/0092-8674(82)90135-0

Edgar RC (2004): MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. doi:10.1093/nar/gkh340

Gamblin SJ, Haire LF, Russell RJ, Stevens DJ, Xiao B, Ha Y, Vasisht N, Steinhauer DA, Daniels RS, Elliot A, Wiley DC, Skehel JJ (2004): The structure and receptor binding properties of the 1918 influenza hemagglutinin. Science 303, 1838–1842. doi:10.1126/science.1093155

Guindon S, Gascuel O (2003): A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704. doi:10.1080/10635150390235520

Hughes AL, Nei M (1988): Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 335, 167–170. doi:10.1038/335167a0

Ina Y, Gojobori T (1994): Statistical analysis of nucleotide sequences of the hemagglutinin gene of human influenza A viruses. Proc. Natl. Acad. Sci. USA 91, 8388–8392. doi:10.1073/pnas.91.18.8388

Kosakovsky Pond SL, Frost SD (2005a): Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21, 2531–2533. doi:10.1093/bioinformatics/bti320

Kosakovsky Pond SL, Frost SD, Muse SV (2005b): HyPhy: hypothesis testing using phylogenies. Bioinformatics 21, 676–679. doi:10.1093/bioinformatics/bti079

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD (2006a): GARD: a genetic algorithm for recombination detection. Bioinformatics 22, 3096–3098. doi:10.1093/bioinformatics/btl474

Kosakovsky Pond SL, Frost SD, Grossman Z, Gravenor MB, Richman DD, Brown AJ (2006b): Adaptation to different human populations by HIV-1 revealed by codon-based analyses. PLoS Comput. Biol. 2, e62. doi:10.1371/journal.pcbi.0020062

Kumar R, Halder P, Poddar R (2006): Adaptive molecular evolution of virulence genes of avian influenza - A virus subtype H5N1: An analysis of host radiation. Bioinformation 1, 321–326.

Muse SV, Gaut BS (1994): A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. Mol. Biol. Evol. 11, 715–724.

Nielsen R, Yang Z (1998): Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148, 929–936.

Raymond FL, Caton AJ, Cox NJ, Kendal AP, Brownlee GG (1986): The antigenicity and evolution of influenza H1 haemagglutinin, from 1950–1957 and 1977–1983: two pathways from one gene. Virology 148, 275–287. doi:10.1016/0042-6822(86)90325-9

Sugita S, Yoshioka Y, Itamura S, Kanegae Y, Oguchi K, Gojobori T, Nerome K, Oya A (1991): Molecular evolution of hemagglutinin genes of H1N1 swine and human influenza A viruses. J. Mol. Evol. 32, 16–23. doi:10.1007/BF02099924

Suzuki Y, Gojobori T (2001): Positively selected amino acid sites in the entire coding region of hepatitis C virus subtype 1b. Gene 276, 83–87. doi:10.1016/S0378-1119(01)00640-0

Suzuki Y (2006): Natural selection on the influenza virus genome. Mol. Biol. Evol. 23, 1902–1911. doi:10.1093/molbev/msl050

Suzuki Y (2008): Positive selection operates continuously on hemagglutinin during evolution of H3N2 human influenza A virus. Gene 427, 111-116. doi:10.1016/j.gene.2008.09.012

Taubenberger JK, Morens DM (2006): 1918 Influenza: the mother of all pandemics. Emerg. Infect. Dis. 12, 15–22.

Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y (1992): Evolution and ecology of influenza A viruses. Microbiol. Rev. 56, 152–179.

Webster RG, Laver WG, Air GM, Schild GC (1982): Molecular mechanisms of variation in influenza viruses. Nature 296, 115–121. doi:10.1038/296115a0