

Molecular evolution of novel swine-origin A/H1N1 influenza viruses among and before human

Na Ding · Nana Wu · Qinggang Xu ·
Keping Chen · Chiyu Zhang

Received: 17 June 2009 / Accepted: 6 August 2009 / Published online: 20 August 2009
© Springer Science+Business Media, LLC 2009

Abstract We find that the novel A/H1N1 influenza viruses exhibit very low genetic divergence and suffer strong purifying selection among human population and confirm that they originated from the reassortment of previous triple-reassortant swine influenza viruses including genomic segments from both avian and human lineages with North American and Eurasian swine lineages. The longer phylogenetic branch length to their nearest genetic neighbors indicates that the origin of the novel A/H1N1 is unlikely to be a very recent event. Seventy-six new unique mutations are found to be monomorphically fixed in the novel A/H1N1 virus lineages, suggesting a role of selective sweep in the early evolution of this virus.

Keywords Novel A/H1N1 influenza virus · Reassortment · Adaptive evolution · Selective sweep

Introduction

A new influenza pandemic starting in Mexico from April 2009 has been caused by a novel swine-origin H1N1 influenza A virus [1]. It spread rapidly to 69 countries/

regions around the world, and caused 21,940 infection cases including 125 deaths worldwide in the last few months [2]. It had been genetically classified as a novel swine-origin A/H1N1 influenza virus [1, 3].

Unlike most avian and swine A influenza viruses that can result in sporadic human infection via animals to human transmission, but lack the ability of human to human transmission [4, 5], the novel swine-origin H1N1 influenza A virus shows a strong ability to transmit from human to human since its first emergence in Mexico [1]. Recent research showed that six gene segments (polymerase PB2 (PB2), PB1, polymerase PA (PA), hemagglutinin (HA), nucleoprotein (NP), and nonstructural protein (NS)) of the novel A/H1N1 virus have the closest homology to those of the triple-reassortant swine influenza viruses previously circulating in pigs in North America, and neuraminidase (NA) and M protein (M) gene segments have the closest homology to ones in the Eurasian lineage of swine influenza viruses [6, 7], suggesting a origin of the novel H1N1 virus via a reassortment of swine influenza viruses between North American and Eurasian lineages. Pigs have been well demonstrated to be a mixing vessel for various influenza viruses to exchange their genetic material and to evolve into a new hybrid virus lineage capable of causing a human pandemic [5, 8]. Because the North American triple-reassortant swine influenza viruses were originated from the reassortment between avian, human, and swine lineages [9], as a sequences, the novel A/H1N1 was initially assumed to be derived from the reassortment between three influenza lineages from avian, swine, and human.

As an important animal reservoir for influenza virus infection in human populations, although up to now no pig was found to be infected by the novel A/H1N1 influenza virus, the new influenza pandemic was believed to be caused by the cross-species transmission [5]. Some adaptive

N. Ding and N. Wu have contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s11262-009-0393-7) contains supplementary material, which is available to authorized users.

N. Ding · N. Wu · Q. Xu · K. Chen · C. Zhang (✉)
Institute of Life Sciences, Jiangsu University, Zhenjiang,
Jiangsu 212013, China
e-mail: zhangcy1999@hotmail.com

mutations, possibly driven by positive selection or selective sweep, are required for animal viruses overcoming species barrier and accomplishing cross-species transmission [10–14]. The current global outbreak of influenza A/H1N1 virus indicates that the novel virus not only accomplishes the cross-species transmission from its original hosts to human but also gains the ability to spread efficiently among human. Therefore, the evolution of this novel influenza A/H1N1 virus is hypothesized to be driven by positive selection at least at early phase of cross-species transmission. To test the hypothesis, we performed phylogenetic and adaptive evolution analyses to understand the evolutionary history of this novel influenza A/H1N1 virus among and before humans.

Method

Sequences

The genome sequences of the novel influenza A viruses were downloaded from the NCBI Influenza Virus Resource (<http://www.ncbi.nlm.nih.gov/genomes/FLU/SwineFlu.html>) in May 12, 2009. These viruses were isolated from North America, Europe, Africa, Australia, and Asia. Because the novel influenza A viruses were demonstrated to belong to Influenza A virus subtype H1N1 (A/H1N1), other A/H1N1 sequences sampled from human, avian, and swine around the world during the period 1918–2009 were also retrieved from this database. Recent research showed that the HA genes of novel swine-origin A/H1N1 had a closer genetic relationship with some A/H1N2 than A/H1N1 [6, 7]. Therefore, HA gene sequences of some A/H1N2 viruses from human and swine were downloaded. Furthermore, some H5N1 and H3N2 subtype sequences were retrieved to be used in the analyses of NP and PB1 genes.

Phylogenetic analysis

The sequences of each genomic segment of the novel influenza A/H1N1 viruses were aligned together with other influenza A viruses using CLUSTAL W program implemented in MEGA 4. The sequence alignments were performed under default condition. The gap open and gap extension penalties in the sequence alignments were 15 and 6.66, respectively. The phylogenetic trees were obtained by using neighbor-joining (NJ) method (MEGA 4) [15], and the reliability of the trees was evaluated by the bootstrap method with 1,000 replications. The genetic diversity within and between the novel A/H1N1 viruses and their closest evolutionary relatives was assessed by MEGA 4.

Adaptive evolution analysis

Positive selection drives viral evolution during cross-species transmission and human infection. To investigate whether the novel influenza A/H1N1 virus is driven by positive selection, the standard McDonald–Kreitman test (<http://mkt.uab.es/mkt/MKT.asp>) was applied to detect the natural selection acting on the novel A/H1N1 virus [16]. The MKT test is based on a comparison of the amount of variation (synonymous and nonsynonymous) within a species and the divergence between species. Under neutrality, the ratio of nonsynonymous to synonymous polymorphisms within species (P_n/P_s) should be equal to the ratio of nonsynonymous to synonymous fixed substitutions between species (D_n/D_s) [16]. The Neutrality Index (NI) is used to indicate the extent to which the levels of amino acid polymorphism depart from the expected in the neutral model. The NI is calculated by the formula $NI = P_n D_s / P_s D_n$. A NI value <1 indicates an excess of fixation of non-neutral replacements due to positive selection, $=1$ means under neutral selection, and >1 reflects negative selection that prevents the fixation of harmful mutations [16]. In phylogenetic trees of all eight genomic segments, the novel influenza A/H1N1 viruses form a well-supported clade, clustering closely with another clade including swine influenza A viruses (referred to as background clade). Both clades were subjected to the standard MKT analyses.

To confirm the result of MKT, further adaptive evolution analyses were performed using CODEML program implemented in PAML 4.0 software package [17]. For influenza virus genes, the protein-coding sequences within two above clades were aligned based on translated protein sequences using Clustal W program implemented in MEGA 4. The site-specific and branch-site models in PAML 4.0 were employed to detect selective pressure among these sequences [18, 19]. In both models, the selective pressure is measured by comparing the rate of nonsynonymous nucleotide substitutions per nonsynonymous site (dN) with that of synonymous substitutions per synonymous site (dS). The dN/dS ratio (ω) is traditionally used as an index to assess positive selection. A $\omega >1$ is taken as evidence of positive (diversifying) selection, $=1$ indicates neutral selection, and <1 reflects strong purifying selection [19]. To detect whether positive selection affects a small number of sites along the novel A/H1N1 influenza virus lineage, this lineage was set to be foreground branch and its nearest-neighbor lineage to be background branch in the branch-site model A (model A). Three ω ratios ($0 < \omega_0 < 1$, $\omega_1 = 1$, $\omega_2 > 1$) and two ω ratios ($0 < \omega_0 < 1$, $\omega_1 = 1$) are assigned for foreground and background branches in model A, respectively. The null model (model A') is same as model A, but with $\omega_2 = 1$ is fixed [18].

Results and discussion

The new A/H1N1 influenza virus originates from the reassortment between lineages from triple-reassortant swine, Eurasia swine, and North American swine

To investigate the evolutionary relationship with other influenza viruses, the phylogenetic analyses of eight genomic segments of novel influenza A/H1N1 virus were performed using MEGA4.0 [15]. In eight phylogenetic trees, all influenza viruses are divided into two large clades, implying a large genetic divergence in the ancient ancestral phase of the evolution of A/H1N1 influenza viruses (data not shown). For HA, PB2, NP, and NS gene segments, one of the two large clades includes not only the influenza virus lineages circulating in avian and swine but also three human-adapted lineages able to spread efficiently among human (i.e., seasonal human A/H1N1, 1918 Spanish A/H1N1, and the new 2009 A/H1N1 influenza viruses), relative to another large clade that only includes influenza viruses prevalent in swine and avian. It implies that some of the influenza viruses in this clade have higher potential to evolve the ability to transmit from human to human after

animals-to-human cross-species transmission. As an example, the tree of HA segment is shown in supplementary Fig. S1. For NA, PB1, PA, and M gene segments, however, three human-adapted influenza virus lineages are dispersed into two large clades. For example, the novel H1N1 lineages in NA tree cluster one large clade, whereas seasonal human A/H1N1 and 1918 Spanish A/H1N1 cluster another large clade (Fig. S2). The different topological position of three human-adapted lineages within the phylogenetic trees of various gene segments suggests that independent evolution and reassortment play a role in gaining human-to-human transmission ability [5, 10].

In all phylogenetic trees of eight gene segments, the new influenza A/H1N1 viruses form a well-supported branch (with bootstrap values of ≥ 99), which classically clusters with some swine-origin influenza virus lineages, and is distinct from human-adapted lineages, such as seasonal human influenza A/H1N1 and 1918 Spanish influenza virus (Figs. 1, S1, S2). In the trees of gene segments PB2, PB1, PA, NP, and NS, the clade of novel influenza A/H1N1 viruses closely clusters with triple-reassortant swine influenza A/H1N1 virus lineage (Fig. 1), which is derived from the reassortment between classic North American swine,

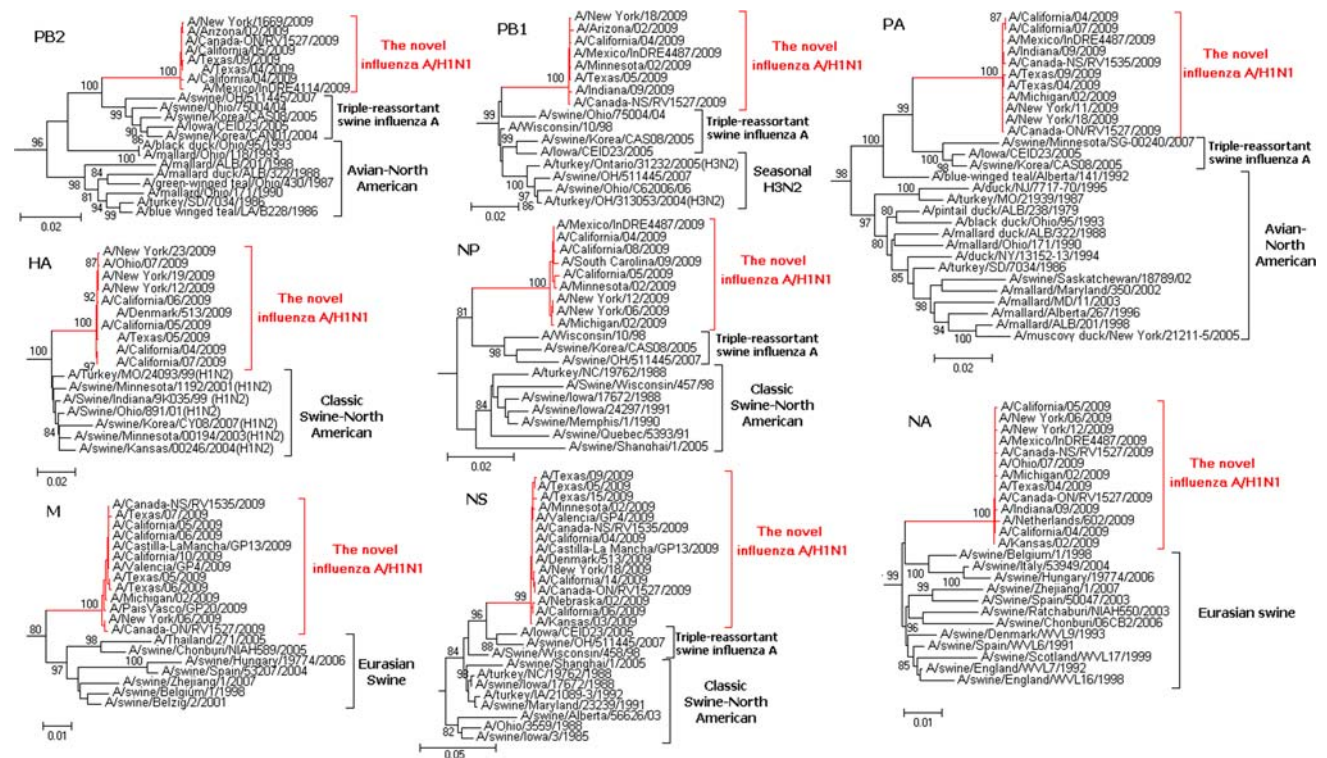


Fig. 1 Phylogenetic subtrees of all eight segments of the novel influenza A/H1N1 viruses. The phylogenetic trees were constructed with MEGA 4.0 by using the NJ method, and the reliability of the tree was evaluated by the bootstrap method with 1,000 replications. Only bootstrap values of >80 are shown at the corresponding nodes. The

triple-reassortant swine influenza A/H1N1 virus lineages that originated from the reassortment between avian, human, and swine lineages are circulating in swine, even occasionally in human, in North America. The GenBank accession numbers of the sequences used in this figure are provided in Table S1

North American avian, and seasonal human H3N2 lineages [9]. The triple-reassortant swine influenza H1 lineages have circulated in swine around the world and occasionally resulted in human infections over the past several years [9, 20, 21]. The analyses of other gene segments show that the NA and M of the novel A/H1N1 viruses are phylogenetically close to Eurasia swine A/H1N1 lineages, whereas HA segment shows the closest phylogenetic relationship with classic North American swine A/H1N1 lineage (Fig. 1). These results distinctly indicate that the new A/H1N1 influenza virus originate from the genetic reassortment between A/H1N1 virus lineages from triple-reassortant swine, Eurasia swine, and North American swine, by which NA and M segments of Eurasia swine lineage and HA of North American swine lineage were imported into the backbone of triple-reassortant swine lineage. This result is well consistent with recent observations [6, 7].

Influenza virus has high potential to reassort [5]. Apart from current novel influenza pandemic, other examples for influenza viruses to gain the ability of cross-species and human-to-human transmission through reassortment were the 1918 Spanish H1N1, 1957 Asian H2N2, and 1968 Hong Kong H3N2 influenza A viruses, in which some avian influenza genomic segments were reassorted with human viruses [5]. Furthermore, influenza virus is able to further facilitate its adaptation to human host through secondary reassortments. Therefore, there is a potential likelihood of the novel A/H1N1 influenza virus to further reassort with other highly pathogenic influenza virus (e.g., H5N1), implying an urgent concern [22].

Long phylogenetic branch indicates that the generation of new A/H1N1 is not a recent event

The genome of influenza virus contains eight RNA segments. Given that the reassortment occurred very recently,

Table 1 Comparison of the mean genetic distance (%) between the clades of the novel influenza A/H1N1 viruses and their closest genetic neighbors

	Mean genetic distance (%)							
	PB2	PB1	PA	HA	NP	NA	M	NS
Within the novel A/H1N1 clade								
Nucleotide	0.1	0.1	0.1	0.4	0.3	0.2	0.2	0.3
Amino acid	0.0	0.2	0.1	0.2	0.4	0.1	0.1	0.8
Within the background clade								
Nucleotide	1.5	2.5	1.6	7.8	1.8	5.9	2.6	10
Amino acid	1.5	1.1	1.6	5.9	1.8	5.9	2.6	10
Between clades								
Nucleotide	4.7	4.7	5.4	5.8	4.7	3.1	5.3	5.6
Amino acid	1.8	2.2	2.5	5.8	2.0	8.3	3.0	8.0

eight genomic segments of the novel A/H1N1 influenza virus should appear longer phylogenetic distance to their nearest genetic neighbors (referred to as the nearest background clade). From the phylogenetic trees of eight genomic segments, we found that the clade of new A/H1N1 had a distinct longer branch length to the nearest background clade (Fig. 1). Further comparison of genetic distances within and between new A/H1N1 and its nearest background clades shows that the mean of genetic distances (0.1–0.8%) interior of the novel A/H1N1 virus clade is obviously smaller than its nearest background clade (1.5–10%) (Table 1), and the mean of genetic distances between both clades is relatively large at the nucleic acid (3.1–5.8%) and amino acid (1.8–8.3%) levels. These results indicate that the novel influenza A/H1N1 virus has a short evolutionary history among human and a relative long evolutionary history before introduction into human, implying that this virus might have been circulating undetected among its animal reservoirs somewhere in the world for a relatively long period of time. More compelling evidence is from the further phylogenetic analyses using BEAST software, which showed that the novel influenza A/H1N1 originated over a period of 10 years (data not shown) [23]. Furthermore, it also suggests that the introduction of the novel A/H1N1 virus into human from animal reservoirs might be a single occasional event or multiple events of genetically homologous lineages [7]. Therefore, it is assumed that there is an animal reservoir where the genetic reassortment of triple-reassortant, Eurasia, and North American swine influenza A viruses occurs. Pigs have been well demonstrated to play a role as a mixing vessel in the co-infection and the reassortment of various influenza viruses [5, 8]. Up to now, although no new A/H1N1 influenza virus is isolated from pigs, the closest genetic relationship of the novel A/H1N1 viruses with swine influenza virus lineages in all eight genomic segments strongly suggests that pigs are the most possible animal reservoir, and as a consequence, the new influenza outbreak was caused by the cross-species transmission of A/H1N1 from pigs to human after reassortment between three swine-origin lineages [6]. However, since pigs do not always appear ill after infection by swine influenza viruses, it increases the difficulty to search for the direct intermediate host of the novel A/H1N1 influenza virus [5].

Purifying selection drives the evolution of new A/H1N1 viruses among human

Some pathogens can obtain the ability of cross-species transmission from animal reservoir species to human, and even human-to-human transmission via adaptive mutation and reassortment of various lineages [10]. Two compelling examples are avian influenza viruses and SARS-CoV. The

former, e.g., the highly pathogenic H5N1 virus, became capable of overcoming the species barrier to establish human infection and caused sporadic human infection cases around the world [5]. However, it is unable to be maintained among human population because of limited ability of human-to-human transmission [22]. The later not only accomplished the cross-species transmission from animals to human [14] but also gained capability to transmit among human population, resulting in the global SARS outbreak in 2003 [13]. Similar to SARS-CoV, but distinct from the H5N1 virus, the novel A/H1N1 influenza virus has caused the current global influenza outbreak, implying the acquirement of the ability of cross-species and human-to-human transmissions. Previous studies had demonstrated that animal viruses suffered positive Darwinian selection during the process of the cross-species transmission and initial stages of human outbreak, followed by purifying (negative) selection when viruses adapted to the new human host in late epidemic [11–13].

To investigate whether positive selection drives the adaptation to human, the McDonald–Kreitman test was applied to analyze the sequences of the novel A/H1N1 viruses and their nearest genetic neighbors (referred to as background clade) (Fig. 1). The result shows that except PA all influenza genomic segments have the NI value of more than 1 (Table 2), indicating a negative selection. In particular, the posterior probabilities (*p*) of NI values in PB2 and NP sequences achieve the statistical significant level (0.016 and 0.014, respectively), implying strong

negative selection acting on both genes. In contrast, a NI value of 0.846 is observed in PA gene segment (Table 2), suggesting weak positive selection acting on this gene.

To confirm the results of MKT analyses, the above sequence data were further analyzed using CODEML program implemented in PAML4.0 [17]. In the analyses, two robust models—site-specific and branch-site models—were used [18, 19]. In the analyses of branch-site model, the clade of the novel A/H1N1 viruses was labeled as foreground branch and its nearest genetic clade was referred to as background branch (Fig. 1). Table 3 shows the results from the branch-site model. First, the alternative model has significantly lower log-likelihood value than the null model for each genomic segment of influenza viruses, strongly suggesting that the alternative model that allows ω (dN/dS: the rate of nonsynonymous nucleotide substitutions per nonsynonymous site with that of synonymous substitutions per synonymous site) >1 in foreground is unfit for the data. Second, all ω_2 values (from 0.934 to 0.950) for all eight influenza virus genes appear to be near neutral, suggesting no positive selection acting on all of eight genes. An exception was observed in PA gene where site 277H was detected under positive selection. These results indicate that except PA gene under weak positive selection, all influenza genomic segments suffer strong purifying selection during the outbreak, well consistent with that of MKT analyses.

Viral evolution plays an important role in cross-species transmission by overcoming the species barrier [10].

Table 2 McDonald–Kreitman test for the novel influenza A/H1N1 virus lineage

	PB2		PB1		PA		HA	
	Polym.	Div.	Polym.	Div.	Polym.	Div.	Polym.	Div.
Neutral	91	67.54	133	49.06	119	52.83	87	41.87
Non-neutral	27	7.01	32	9.03	23	12.05	57	17.14
Neutrality Index (NI)	2.855		1.307		0.846		1.599	
Proportion of adaptive substitutions (α)	−1.855		−0.307		0.153		−0.599	
χ^2 value	5.698		0.424		0.179		2.000	
<i>p</i> -value	0.016		0.514		0.671		0.157	
	NP		NA		M		NS	
	Polym.	Div.	Polym.	Div.	Polym.	Div.	Polym.	Div.
Neutral	47	39.95	53	60.22	77	19.03	25	14.72
Non-neutral	21	5.01	31	26.42	27	5.02	33	11.12
Neutrality Index (NI)	3.560		1.333		1.328		1.747	
Proportion of adaptive substitutions (α)	−2.560		−0.333		−0.328		−0.747	
χ^2 value	5.945		0.784		0.269		1.377	
<i>p</i> -value	0.014		0.375		0.603		0.240	

Polym. polymorphism, *Div.* divergence

Table 3 Maximum likelihood (ML) estimates for the novel A/H1N1 influenza viruses with the branch-site model

Gene	Parameters under null model (MA') (fix $\omega_2 = 1$)	Parameters under alternative model (MA)	lnL ₀ (lnL ₁)	Positively selected sites
PB2	P0 = 0.317, $\omega_0 = 0.000$, P1 = 0.383 (P2a + P2b = 0.300)	P0 = 0.325, $\omega_0 = 0.947$, P1 = 0.379, $\omega_2 = \mathbf{0.947}$ (P2a + P2b = 0.296)	-6739.85 (-6900.14)	NA
PB1	P0 = 0.330, $\omega_0 = 0.000$, P1 = 0.369 (P2a + P2b = 0.301)	P0 = 0.313, $\omega_0 = 0.950$, P1 = 0.384, $\omega_2 = \mathbf{0.950}$ (P2a + P2b = 0.303)	-5161.44 (-5006.37)	NA
PA	P0 = 0.316, $\omega_0 = 0.000$, P1 = 0.384 (P2a + P2b = 0.298)	P0 = 0.315, $\omega_0 = 0.934$, P1 = 0.390, $\omega_2 = \mathbf{0.934}$ (P2a + P2b = 0.294)	-8471.93 (-9402.76)	277H
HA	P0 = 0.331, $\omega_0 = 0.000$, P1 = 0.370 (P2a + P2b = 0.299)	P0 = 0.310, $\omega_0 = 0.950$, P1 = 0.394, $\omega_2 = \mathbf{0.950}$ (P2a + P2b = 0.295)	-4021.31 (-4099.22)	NA
NP	P0 = 0.316, $\omega_0 = 0.000$, P1 = 0.386 (P2a + P2b = 0.298)	P0 = 0.310, $\omega_0 = 0.950$, P1 = 0.396, $\omega_2 = \mathbf{0.950}$ (P2a + P2b = 0.294)	-4380.30 (-4903.45)	NA
NA	P0 = 0.318, $\omega_0 = 0.000$, P1 = 0.380 (P2a + P2b = 0.302)	P0 = 0.310, $\omega_0 = 0.945$, P1 = 0.390, $\omega_2 = \mathbf{0.945}$ (P2a + P2b = 0.300)	-7218.600 (-7797.14)	NA
M	P0 = 0.319, $\omega_0 = 0.000$, P1 = 0.389 (P2a + P2b = 0.291)	P0 = 0.327, $\omega_0 = 0.948$, P1 = 0.376, $\omega_2 = \mathbf{0.948}$ (P2a + P2b = 0.297)	-2438.53 (-2473.29)	NA
NS	P0 = 0.313, $\omega_0 = 0.000$, P1 = 0.383 (P2a + P2b = 0.304)	P0 = 0.309, $\omega_0 = 0.949$, P1 = 0.383, $\omega_2 = \mathbf{0.949}$ (P2a + P2b = 0.307)	-3466.55 (-3713.25)	NA

Positive selected sites were identified with posterior probability $p \geq 99\%$, in boldface

Before the adaptation of animal viruses to human, there should be some adaptive mutations to be accumulated. For example, SARS-CoV Spike (S) protein was previously detected under positive selection during both phases of cross-species transmission and early epidemic, and under strong purifying selection in late epidemic [11, 12]. Although the information on the epidemiology and sequences before the influenza outbreak (April 2009) is lacking, it is believed that the novel influenza A/H1N1 virus experiences a similar evolutionary process to SARS-CoV. Therefore, it is hypothesized that positive selection drives the evolution of the novel influenza A/H1N1 virus during and before jumping into human.

Selective sweep may drive the evolution of novel A/H1N1 virus before cross-species transmission to human

By comparing the amino acid sequences between the new A/H1N1 viruses and their nearest genetic neighbors, we found that the novel A/H1N1 viruses have less divergence in amino acid level than their background lineages (Fig. 2). Relative to their neighboring lineages, the majority of mutation sites are monomorphic in the novel A/H1N1 virus lineage (Fig. 2), implying an exclusive fixation. Furthermore, there are some new unique amino acid residues only occurring and being exclusively fixed in the novel A/H1N1 lineage (Fig. 2), suggesting a key role in adaptation to human host, but also providing a potential evidence of

selective sweep in early evolution of the novel A/H1N1 influenza virus. Nine, 6, 8, 17, 5, 16, 5, and 10 new unique mutations were detected in PB1, PB2, PA, HA, NP, NA, M, and NS genes of new A/H1N1 contain, respectively (Fig. 2, Table 4). Despite current lack of direct experimental evidence, some of them are believed to contribute to the adaptation of novel A/H1N1 influenza virus to human host, whereas the other might be fixed through genomic hitchhiking.

Conclusion

In this study, we confirmed that the novel A/H1N1 influenza virus originated from the reassortment between influenza virus lineages from triple-reassortant swine, Eurasia swine, and North American swine, suggesting an urgent concern on the potential reassortment of the new A/H1N1 viruses with the highly pathogenic H5N1 virus. The very low genetic divergence and the longer phylogenetic branch length to their nearest genetic neighbors indicate that the origin of the novel A/H1N1 virus is unlikely to be a very recent event. Strong purifying selection was found to force

Fig. 2 Amino acid variations between the novel influenza A/H1N1 lineage and their background clade. In the multiple alignment, gap open and gap extension penalties are 15 and 6.66, respectively. Only the variant sites were shown. The background clade represents the closest evolutionary neighbors of the novel A/H1N1 lineage in phylogenetic tree (see Fig. 1). Unique mutation sites exclusively fixed in the novel influenza A/H1N1 lineages are shaded

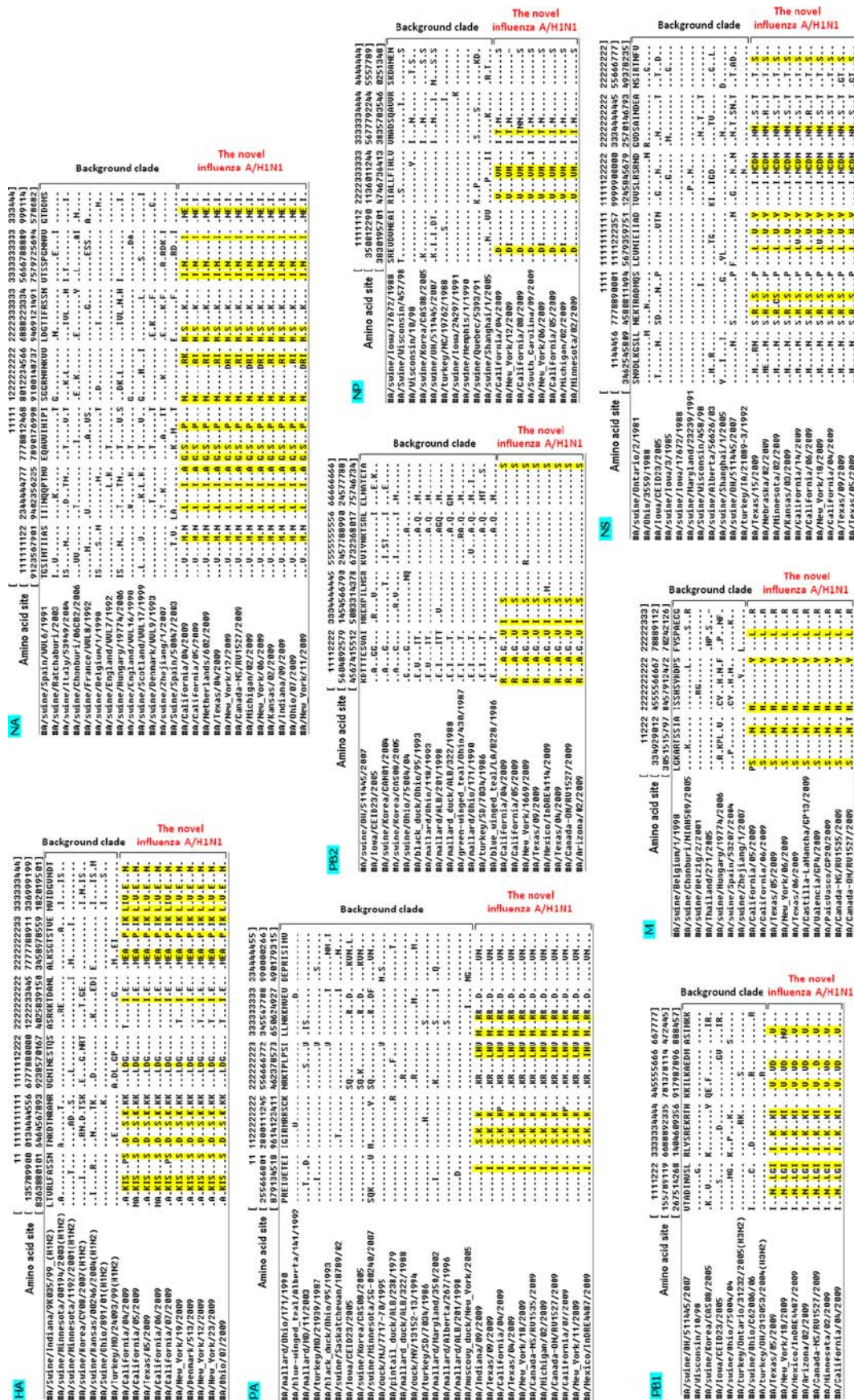


Table 4 Unique mutation sites exclusively fixed in the novel influenza A/H1N1 lineage

Segment	Unique amino acid mutations
PB2	K54R, T184A, I292V, M315I, P453S, A684S
PB1	D175N, V/M121L, L298I, L364I, R386K, T435I, I517V, E618D, I728V
PA	T85I, G186S, R204K, R213K, P275L, S277H, L336M, K356R
HA	R53K, L78I, F188S, N101S, N114D, T/A145S, R147K, N185D, T233I, K275E, S278A, S288P, E319K, M331L, I362V, G391E, D490N
NP	E53D, I217V, F314V, I317M, A374T
NA	I/V19M, I/T40L, L/P46I, V75A, E77G, A/T/K79S, H126P, G/S189N, K257R, V263KI, L/E269M, G286S, T365I, S/R369N, V/A389I, D398E
M	G30S, S207N, I248M, S/F272Y, P284L
NS	K78R, A91S, M119L, I129V, D/N171Y, R206C, N207D, S240D, S240N, N/D272S

the evolution of the novel A/H1N1 virus among human. In addition, 76 new unique mutations are found to be exclusively fixed in whole genome of the novel A/H1N1 influenza virus lineage, not only suggesting a role of selective sweep in the early evolution of this virus but also providing a valuable reference for clinic diagnosis, treatment, and vaccine research of influenza.

Acknowledgments The authors thank Mrs. Neng Yao and Wanqiang Dong for their preparation in the sequence data. The study was supported by the grants from the “top-notch personnel” project of Jiangsu University, and in part from the National Natural Science Foundation of China (No. 30600352).

References

- Centers for Disease Control and Prevention (CDC), MMWR Morb. Mortal. Wkly Rep. **58**, 467–470 (2009)
- WHO, *Influenza A (H1N1) - update 44* (World Health Organization, Geneva, 2009). http://www.who.int/csr/don/2009_06_05/en/index.html. Accessed 5 June 2009
- Centers for Disease Control and Prevention (CDC), MMWR Morb. Mortal. Wkly Rep. **58**, 400–402 (2009)
- T. Kuiken, E.C. Holmes, J. McCauley, G.F. Rimmelzwaan, C.S. Williams, B.T. Grenfell, *Science* **312**, 394–397 (2006)
- R.J. Webby, R.G. Webster, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **356**, 1817–1828 (2001)
- 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**, 2605–2615 (2009)
- 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. Rivaviller, J. Smagala, M. de Graaf, D.F. Burke, R.A. Fouchier, C. Pappas, C.M. Alpuche-Aranda, H. Lopez-Gatell, H. Olivera, I. Lopez, 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**, 197–201 (2009)
- K. Van Reeth, A. Nicoll, *Euro Surveill.* **14**, pii: 19124 (2009)
- V. Shinde, C.B. Bridges, T.M. Uyeki, B. Shu, A. Balish, X. Xu, S. Lindstrom, L.V. Gubareva, V. Deyde, R.J. Garten, M. Harris, S. Gerber, S. Vagasky, F. Smith, N. Pascoe, K. Martin, D. Dufficy, K. Ritger, C. Conover, P. Quinlisk, A. Klimov, J.S. Bresee, L. Finelli, N. *Engl. J. Med.* **360**, 2616–2625 (2009)
- C.R. Parrish, E.C. Holmes, D.M. Morens, E.C. Park, D.S. Burke, C.H. Calisher, C.A. Laughlin, L.J. Saif, P. Daszak, *Microbiol. Mol. Biol. Rev.* **72**, 457–470 (2008)
- C.Y. Zhang, J.F. Wei, S.H. He, *BMC Microbiol.* **6**, 88 (2006)
- H.D. Song, C.C. Tu, G.W. Zhang, S.Y. Wang, K. Zheng, L.C. Lei, Q.X. Chen, Y.W. Gao, H.Q. Zhou, H. Xiang, H.J. Zheng, S.W. Chern, F. Cheng, C.M. Pan, H. Xuan, S.J. Chen, H.M. Luo, D.H. Zhou, Y.F. Liu, J.F. He, P.Z. Qin, L.H. Li, Y.Q. Ren, W.J. Liang, Y.D. Yu, L. Anderson, M. Wang, R.H. Xu, X.W. Wu, H.Y. Zheng, J.D. Chen, G. Liang, Y. Gao, M. Liao, L. Fang, L.Y. Jiang, H. Li, F. Chen, B. Di, L.J. He, J.Y. Lin, S. Tong, X. Kong, L. Du, P. Hao, H. Tang, A. Bernini, X.J. Yu, O. Spiga, Z.M. Guo, H.Y. Pan, W.Z. He, J.C. Manuguerra, A. Fontanet, A. Danchin, N. Niccolai, Y.X. Li, C.I. Wu, G.P. Zhao, *Proc. Natl. Acad. Sci. USA* **102**, 2430–2435 (2005)
- S.M.E.C. Chinese, *Science* **303**, 1666–1669 (2004)
- W. Li, C. Zhang, J. Sui, J.H. Kuhn, M.J. Moore, S. Luo, S.K. Wong, I.C. Huang, K. Xu, N. Vasilieva, A. Murakami, Y. He, W.A. Marasco, Y. Guan, H. Choe, M. Farzan, *EMBO J.* **24**, 1634–1643 (2005)
- K. Tamura, J. Dudley, M. Nei, S. Kumar, *Mol. Biol. Evol.* **24**, 1596–1599 (2007)
- R. Egea, S. Casillas, A. Barbadilla, *Nucleic Acids Res.* **36**, W157–W162 (2008)
- Z. Yang, *Mol. Biol. Evol.* **24**, 1586–1591 (2007)
- J. Zhang, R. Nielsen, Z. Yang, *Mol. Biol. Evol.* **22**, 2472–2479 (2005)
- Z. Yang, R. Nielsen, N. Goldman, A.M. Pedersen, *Genetics* **155**, 431–449 (2000)
- A.P. Newman, E. Reisdorf, J. Beinemann, T.M. Uyeki, A. Balish, B. Shu, S. Lindstrom, J. Achenbach, C. Smith, J.P. Davis, *Emerg. Infect. Dis.* **14**, 1470–1472 (2008)
- H.M. Yassine, M. Khatri, Y.J. Zhang, C.W. Lee, B.A. Byrum, J. O’Quin, K.A. Smith, Y.M. Saif, *Vet. Microbiol.* doi:10.1016/j.vetmic.2009.1004.1028, 2009
- A.N. Abdel-Ghafar, T. Chotpitayasunondh, Z. Gao, F.G. Hayden, D.H. Nguyen, M.D. de Jong, A. Naghdaliyev, J.S. Peiris, N. Shindo, S. Soeroro, T.M. Uyeki, *N. Engl. J. Med.* **358**, 261–273 (2008)
- 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* **459**, 1122–1125 (2009)