Virology 405 (2010) 314–321

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Comparison of selection pressures on the HA gene of pandemic (2009) and seasonal human and swine influenza A H1 subtype viruses

Yuki Furuse, Kozue Shimabukuro, Takashi Odagiri, Rumi Sawayama, Takashi Okada, Irona Khandaker, Akira Suzuki, Hitoshi Oshitani *

Department of Virology, Tohoku University Graduate School of Medicine, 2-1 Seiryou-machi Aoba-ku, Sendai, 980-8575, Japan

ARTICLE INFO

Article history:
Received 8 April 2010
Returned to author for revision 20 May 2010
Accepted 8 June 2010
Available online 3 July 2010

Keywords:
Influenza
Evolution
Pandemic
Phylogeny
Hemagglutinin

ABSTRACT

The 2009 human pandemic influenza (H1N1) virus possesses the HA gene of the H1 subtype. The evolutionary process of the 2009 H1N1 virus remains to be defined. We performed genetic analyses of the HA gene by comparing the 2009 H1N1 virus with seasonal human and swine viruses.

We analyzed sequences of 116 2009 H1N1 viruses, and obtained 1457 seasonal H1N1, 365 swine H1, and 1332 2009 H1N1 viruses from the database. Selection pressure for the 2009 H1N1 virus was higher than that for the swine virus and equivalent to that for the seasonal virus. Positions 206 and 264 were found to be positively selected sites. We also identified sites under different selection pressures from the seasonal or swine virus that may be involved in imparting significant biological characteristics.

The evolutionary characteristics of the H1 gene of the 2009 H1N1 virus differed from those of seasonal and swine viruses.

Introduction

The influenza virus is a common cause of respiratory infections worldwide. Subtypes of the influenza A virus are determined by the antigenicity of two surface glycoproteins: hemagglutinin (HA) and neuraminidase (NA). The H1N1 and H3N2 subtypes have been circulating within the human population since 1977.

A pandemic occurred from 1918 to 1919 due to the Spanish influenza (H1N1) virus, and its progeny viruses were circulating in the human population until 1957 (Webster et al., 1992). The H1N1 virus disappeared in 1957 and was replaced by the H2N2 virus when the Asian influenza (H2N2) pandemic occurred. Thereafter, the Hong Kong influenza virus (H3N2) caused a pandemic in 1968. However, the H1N1 virus reemerged in 1977 as the Russian influenza, which was similar to H1N1 viruses isolated in the 1950s (Webster et al., 1992). Since then, the progeny of the H1N1 virus has been circulating as a seasonal virus along with the H3N2 virus. Furthermore, in April 2009, outbreaks of influenza in the human population caused by a novel H1N1 virus that had originated from the swine influenza virus were reported from Mexico and the US (Garten et al., 2009; Smith et al., 2009). Reassortment between swine influenza viruses in two distinct lineages, “triple reassortant” and “Eurasian avian-like swine,” led to the generation of the 2009 pandemic H1N1 virus. The source of the swine influenza virus lineage “triple reassortant” itself comprised genes derived from avian, human, and swine influenza virus lineages (Garten et al., 2009; Smith et al., 2009). The novel 2009 pandemic H1N1 virus has spread worldwide, and the World Health Organization (WHO) raised its official pandemic alert level to phase 6 on its six-phase scale on June 11, 2009. According to WHO, more than 208 countries and overseas territories or communities have reported laboratory-confirmed cases of pandemic influenza, including at least 12,799 deaths, as of January 2010: http://www.who.int/csr/don/2010_01_08/en/index.html (The World Health Organizaion, 2010).

HA protein is responsible for virus attachment, and the subsequent fusion of viral and cellular membranes (Earp et al., 2005; Skehel and Wiley, 2000; Steinhauer, 1999). Furthermore, HA may also play a structural role in budding and particle formation (Earp et al., 2005; Skehel and Wiley, 2000; Steinhauer, 1999). It is a rod-shaped molecule with its stalk inserted into the viral membrane and projecting like a spike. It is synthesized as a single polypeptide chain (HA0), and the HA0 precursor cleaves into HA1 and HA2 subunits: a long fibrous stem, comprised of an HA2 component, and a globular head, comprised of an HA1 component (Gamblin et al., 2004). The receptor binding site lies within the globular head of the molecule (Skehel and Wiley, 2000). The receptor binding specificity of HA, which is determined by the nature of the amino acids that form the receptor binding pocket, is responsible for the host range restriction of the virus (Rogers and Paulson, 1983).

Antigenic drift is an accumulation of point mutations in HA or NA genes, leading to minor and gradual antigenic changes (Webster et al., 1992). Drift variants emerge due to positive selection of spontaneous mutants that can evade the existing immunity. HA protein is a major

* Corresponding author.
E-mail address: oshitanih@mail.tains.tohoku.ac.jp (H. Oshitani).

0042-6822/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2010.06.018

antigenic component of the virus. For H1 viruses, antigenic sites are designated as Ca1, Ca2, Cb, Sa, and Sb (Gerhard et al., 1981; Stevens et al., 2004). Antigenic sites surround the sialic acid receptor binding site. Single point mutations in an HA antigenic site can be sufficient for antigenic variation. In addition, oligosaccharide side chains to HA are believed to facilitate viral escape from an immune response (Schulze, 1997; Wei et al., 2010). The number of N-glycosylation sequons (Asn-Xaa-Ser/Thr, where Xaa is any amino acid except Pro) in HA sequences has been increasing in the human population since the emergence of Spanish influenza (Igarashi et al., 2008). From 1918 to 1957, and from 1977 to 2009, further substantial antigenic evolution of human H1N1 viruses occurred (Kilbourne et al., 2002; Tumpey et al., 2004). Antigenic evolution of the virus was sufficient to recommend eight different strains for the H1N1 component of the influenza virus vaccine from 1977 to 2009 (Hay et al., 2001; Health Protection Agency, 2009). Based on the theory of natural selection, positively selected sites in HA could be found by calculating the selection pressure (Suzuki, 2006, 2008). This may be useful for identifying the epitopes involved in the elimination of viruses from infected patients.

The novel 2009 pandemic H1N1 virus (from here, we refer to the virus as "2009 H1N1 virus") possesses the HA gene of the H1 subtype, which is also possessed by the seasonal H1N1 influenza virus. However, the similarity of HA genes between seasonal and 2009 H1N1 viruses is only ~73%. Therefore, the existing immunity for the seasonal virus would not prevent infection with the 2009 H1N1 virus (Garten et al., 2009; Hancock et al., 2009; Itoh et al., 2009). The virus has spread rapidly worldwide with substantial morbidity and mortality rates. The mechanisms by which the 2009 H1N1 virus has evaded host range restriction are a major topic for study (Furuse et al., 2010; Mehle and Doudna, 2009). Besides, the evolutionary process of the 2009 H1N1 virus remains to be defined since its introduction in the human population. The virus responsible for the Spanish influenza pandemic from 1918 to 1919 may have evolved by becoming much more lethal, thereby leading to a second wave with higher mortality (Reid et al., 2003). In addition, influenza viruses from different hosts have evolved under different selection pressures (Furuse et al., 2009).

Thus, genetic analysis comparing 2009 H1N1 as well as seasonal human and swine viruses should provide a critical insight into the evolutionary course and mechanisms of the influenza virus. It would be intriguing to know whether or not seasonal (progeny of Spanish influenza virus) and 2009 H1N1 viruses are under different selection pressures in the same human population, and whether or not 2009 H1N1 and swine viruses are under different selection pressures, since the 2009 H1N1 virus was derived from the swine virus.

In this study, we performed phylogenetic analyses of the H1 subtype HA gene and showed the difference and similarity in evolutionary characteristics among seasonal H1N1, swine H1, and 2009 H1N1 viruses.

Results and discussion

We calculated the evolutionary rates of and selection pressures on the entire HA gene of seasonal H1N1, swine H1, and 2009 H1N1 viruses. The evolutionary rate is described by the number of substitutions/site/year. A higher selection pressure indicates that the gene or site was under stronger selection, i.e., positive selection for amino acid substitution. Lower selection pressure indicates that the gene or site was under stronger negative selection to retain the same amino acid(s), because changes may lead to incompetence or abortion (Pond et al., 2007; Suzuki and Gojobori, 1999).

The evolutionary rate of seasonal H1N1 (1918–1957), seasonal H1N1 (1977–2009), swine H1 (1930–2009), and 2009 H1N1 (March–December 2009) viruses was 2.9×10^{-3}, 1.7×10^{-3}, 1.9×10^{-3}, and 0.9×10^{-3} substitutions/site/year, respectively (Fig. 1). The evolutionary rate of 2009 H1N1 virus was much lower than that of the other viruses. However, the evolutionary rate of 2009 H1N1 virus was unreliable as the correlation coefficient was very low ($R^2=0.01$). We observed the evolutionary trend of 2009 H1N1 virus for only 10 months. A longer period of observation is needed to establish the evolutionary rate of the virus.

The selection pressure for the swine H1 virus was the lowest among the viruses analyzed (Fig. 2). However, 2009 H1N1 virus, which originated from the swine H1 virus, had similar or even higher selection pressure compared with the seasonal H1N1 virus. Adaptation to a new host, humans, may have increased the selection pressure of 2009 H1N1 virus.

The antigenic evolution of swine viruses occurred more slowly than that of human viruses, although rates of genetic evolution of the HA gene were similar for both (de Jong et al., 2007; Sugita et al., 1991). We also showed that the selection pressure for the swine H1 virus was lower than that for the seasonal H1N1 virus, whereas the evolutionary rates for swine H1 and seasonal H1N1 viruses from 1977 to 2009 were equivalent. This could be the consequence of a lower selection pressure affecting the antigenic properties of HA in pigs. A vast majority of domestic pigs are killed at the age of 6 months, and thus the susceptible pig population is continuously renewed, thereby limiting the build-up of immune pressure. Because of the short average life span of pigs, evolution of the swine influenza virus may be determined only to a limited extent by the immune pressure, a driving force for the antigenic drift of influenza viruses in humans.

Furthermore, comparison of the data for the seasonal H1N1 virus from 1918 to 1957 and from 1977 to 2009 indicated that the evolutionary rate and selection pressure of this virus decreased after its reemergence in 1977. These results suggest that the seasonal H1N1 virus did not have amino acid substitutions as frequently as before, and these results are compatible with those of previous reports (Raymond et al., 1986; Shen et al., 2009). The virus might already have been well adapted to humans when it reemerged in 1977. The selection pressure for 2009 H1N1 virus is as high as that for the seasonal virus from 1918 to 1957. The selection pressure could be high, especially in the early stage of viral pandemics, due to increased adaptation to a new host.

Fig. 3 shows positively selected sites in the HA1 region. Selection profiles are different among viruses. Eight positively selected sites were identified for the seasonal H1N1 virus. Among these, sites at positions 187, 190, 192, and 225, identified by H3 numbering, are located at antigenic sites. Mutations in these sites could be due to positive selection imposed by the host's immune response, leading to antigenic drift. Besides, mutations in positions 190 and 225 are key determinants for effective binding to human-like receptors (Kobasa et al., 2004; Stevens et al., 2006; Tumpey et al., 2007). Mutations in these sites might result in a change in the receptor binding affinity in humans. Due to the overlapping locations of the antigenic sites and receptor binding sites of HA, the biological significance of mutations at these sites is controversial (Shen et al., 2009). Hensley et al. recently showed that neutralizing antibodies selected escape mutants with a mutation in HA that increased viral binding to cell receptors (Hensley et al., 2009).

Four sites were found to be positively selected for the swine H1 virus. Among these, positions 83 and 192 were located at antigenic sites. Swine influenza virus is considered to be under weak selection pressure by the host's immune system (de Jong et al., 2007), and we showed lower selection pressure on swine H1 virus (see above). Although the selection pressure on swine virus is weak, results indicate that selection pressure has influenced the evolution of the virus, leading to amino acid substitutions at antigenic sites. de Jong et al. proved the occurrence of the antigenic evolution of swine viruses (de Jong et al., 2007). Furthermore, no sites involved in receptor binding specificity were under positive selection in the swine H1 virus. The receptors of swine cells can bind to both human and avian influenza viruses, which possess both N-acetylsialic acid attached to galactose with an α2,6 linkage (SAα2,3Gal) and SAα2,6Gal (Ito et al.,

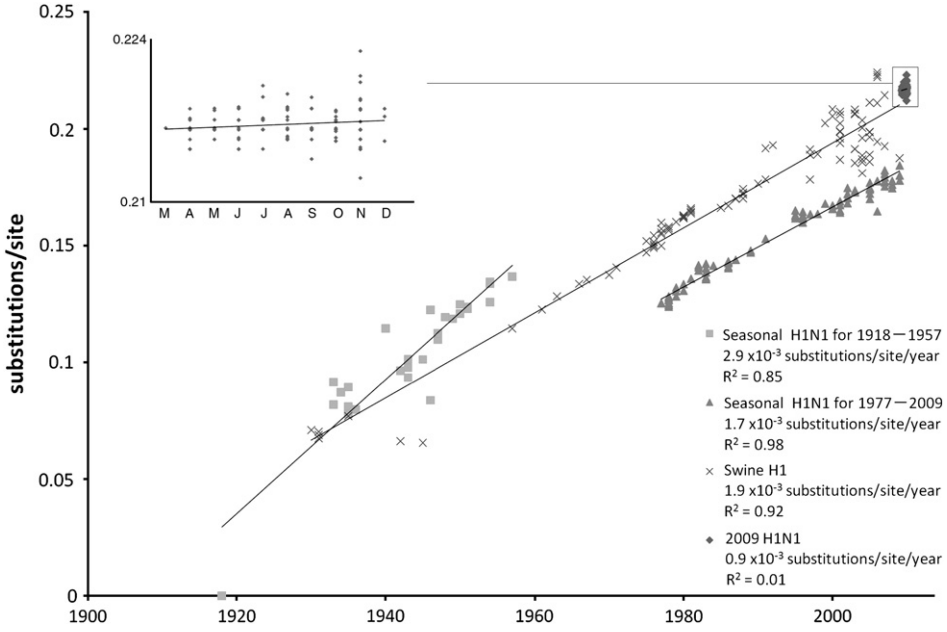


Fig. 1. Evolutionary rate. Numbers of nucleotide substitutions compared with A/South Carolina/1/18 (H1N1) are plotted. Evolutionary rates are calculated from the slope of the tangent of a simple regression line (number of substitutions/site/year) for seasonal H1N1 (1918–1957 and 1977–2009), swine H1, and 2009 H1N1 viruses. The square of the correlation coefficient (R^2) was estimated using Pearson correlation. Expansion for 2009 H1N1 virus is also shown.

1998; Kida et al., 1994). This may be a reason why positive selection was not observed for sites related to receptor binding. We calculated an average for dN/dS at sites involved in receptor binding specificity (Fig. 3). The average for selection pressures was very low for the swine H1 virus.

For 2009 H1N1 virus, positions 206 and 264 were found to be positively selected sites (Fig. 3). It should be noted that no sites under positive selection for 2009 H1N1 virus overlapped with those for the seasonal or swine virus, i.e., sites under positive selection in 2009 H1N1 virus were not positively selected in the seasonal or swine virus.

As observed with the seasonal H1N1 virus, one of the sites under positive selection for 2009 H1N1 virus, position 206, was involved in antigenicity. Mutation in this site could be due to positive selection imposed by the host's immune response. It is surprising that selection pressure had already been imposed on 2009 H1N1 virus, although only a year has passed since its introduction into the human population, which does not possess immunity to 2009 H1N1 virus, except for cross-reactive immunity observed in some elderly people (Garten et al., 2009; Itoh et al., 2009). The virus must be circulating very rapidly and extensively in the human population, evolving under positive selection through the host's immune response. Another possibility is that mutation at the site can be advantageous for the virus, e.g., by increasing receptor binding affinity and/or replication ability.

We also identified a site at position 264 as being positively selected for 2009 H1N1 virus. However, its specific role is unknown. Experimental studies are needed to clarify the significance of this mutation. The site may possess unknown characteristics such as being part of unidentified antigenic sites. Furthermore, no sites involved in receptor binding specificity were under significant positive selection (Fig. 3A). 2009 H1N1 virus is considered to be able to recognize human-type receptors (Childs et al., 2009; Yang et al., 2010). Receptor binding sites may not have to be mutated to increase adaptation in humans. However, we identified many nonsynonymous mutations at sites involved in receptor binding specificity, even if they were not positively selected. Averaged selection pressure on sites involved in receptor binding specificity for 2009 H1N1 virus was considerably higher than that for the swine H1 virus (Fig. 3A). Matrosovich et al. showed that the receptor binding specificity of HA is altered soon after the transmission of an avian virus to mammalians and, therefore, may be a prerequisite for the highly effective replication and spread that characterizes epidemic strains (Matrosovich et al., 2000). We should continue to monitor whether mutation at receptor binding sites can occur, leading to changes in receptor binding specificity.

We also observed a differential in site-by-site selection pressures between viruses. We identified and mapped sites under significantly different selection pressures between 2009 H1N1 and seasonal H1N1 viruses, and between 2009 H1N1 and swine H1 viruses (Fig. 4 and Supplementary Table S1). The number of sites under higher selection pressure in 2009 H1N1 virus compared with the seasonal H1N1 or swine H1 virus (yellow sites in Fig. 4) was higher than the number of sites under lower selection pressure in 2009 H1N1 virus compared with the seasonal H1N1 or swine H1 virus (purple sites in Fig. 4). These results are compatible with those in Fig. 3B, which shows that 2009 H1N1 virus possesses more sites at which dN/dS was >1 than either the seasonal H1N1 or swine H1 virus. These sites may play significant roles in 2009 H1N1 virus. Another possibility is that their selection pressure was overestimated. As intensive surveillance had been conducted for 2009 H1N1 virus; incidental mutations that would not be maintained by purifying selection (also known as "negative selection") have also been found.

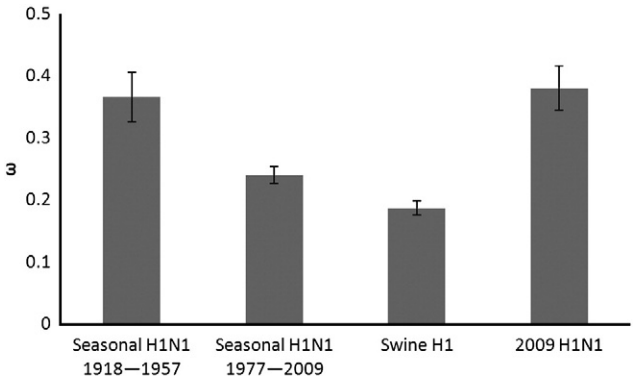


Fig. 2. Selection pressure. Selection pressures for the whole sequence (ω) are calculated for the entire coding region of the HA gene. Error bar shows 95% confidence interval.

Sites under higher selection pressure in 2009 H1N1 virus are located in various parts of HA. They are situated at antigenic sites, receptor binding sites, or sites whose roles are unknown. The HA gene of 2009 H1N1 virus must have evolved with several functions. Furthermore, many sites that were under lower selection pressure in 2009 H1N1 virus than in either the seasonal H1N1 or swine H1 virus are located at antigenic sites (see blue sites in Fig. 4A and purple sites in Figs. 4B and C), and, some of these are N-glycosylation sequons in the seasonal H1N1 virus (see green sites in Fig. 4A). Additional N-glycosylations during future antigenic changes in HA of 2009 H1N1 virus have been predicted (Igarashi et al., 2010). Wei et al. showed that additional N-glycosylation sequons in HA of 2009 H1N1 virus led to resistance to neutralizing antibodies (Wei et al., 2010). However, 2009 H1N1 virus has not acquired any additional N-glycosylation sequons since its emergence (data not shown).

These results indicate that selection pressure at antigenic sites in either swine or seasonal virus was not present at the same sites in 2009 H1N1 virus. Until now, immune pressure appears to have been working at different sites for 2009 H1N1 virus and seasonal H1N1 or swine H1 virus. Sites under lower selection pressure in 2009 H1N1 virus than in the seasonal H1N1 or swine H1 virus might also be mutating in 2009 H1N1 virus to escape any future immune response.

There is the possibility of the existence of common epitopes for neutralizing antibodies cross-reactive to both HAs for the pandemic viruses of 1918 and 2009 (Igarashi et al., 2010; Yang et al., 2010). Also, Igarashi et al. showed that early human H1N1 viruses isolated from the 1930s to the 1940s still harbored some of the original epitopes that are also found in 2009 H1N1 virus (Igarashi et al., 2010). We should monitor whether mutations might occur at additional sites, including sites positively selected in the seasonal or swine virus. If this occurs, it would be interesting to discover whether these mutations alter the biological characteristics of the virus.

In the present study, we included sequences both from isolates (using cell lines or eggs) and clinical samples, or unknown origin. Mutations can occur during isolation or passage in cell lines (Fitch et al., 1997; Zhirnov et al., 2009). Egg-adapted influenza viruses also have non-natural host-associated modifications in HA sequences (Robertson et al., 1987; Shen et al., 2009). Therefore, the results of the present study may contain some false positives. However, sites under positive selection in 2009 H1N1 virus (positions 206 and 264) have not been reported previously as egg-adapted mutations (Gambaryan et al., 1999; Robertson et al., 1987; Xu et al., 1993).

Conclusions

The 2009 H1N1 virus was introduced into the human population from the swine population. Within a year, the virus has evolved rapidly and extensively following its emergence. In conclusion, evolutionary characteristics of the H1 gene of 2009 H1N1 virus differ from those of the seasonal H1N1 and swine H1 viruses. We identified sites under positive selection and different selection pressures from the seasonal H1N1 or swine H1 virus that may be involved in imparting significant biological characteristics such as antigenicity and receptor binding specificity.

Materials and methods

Viruses and sequencing

Nasopharyngeal swabs were collected from patients with influenza-like illnesses and who had visited pediatric clinics in Sendai City, Japan, from September to December 2009. Clinical specimens were inoculated into Madin–Darby canine kidney (MDCK) cells with 3.5 μl/mL of trypsin. Viral RNA was extracted from isolates using the RNeasy Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. Viral RNA was reverse-transcribed into complementary DNA using an influenza A generic primer, Uni12, as reported elsewhere (Hoffmann et al., 2001). PCR was then performed to amplify all segments, including the HA gene, as reported elsewhere (Inoue et al., 2010). Templates were labeled by performing a cycle-sequencing reaction using BigDye Terminator ver. 3.1 (Applied Biosystems, Foster City, CA, USA), and the products were analyzed using an automatic sequencer (3730xl Genetic Analyzer, Applied Biosystems) according to the manufacturer's instructions. A total of 116 pandemic 2009 H1N1 viruses were analyzed (supplementary file).

Sequence data

In addition to the sequence data we analyzed for 2009 H1N1 virus, data were also obtained from the influenza sequence database (Influenza Virus Resource: http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html, accessed January 14, 2010) (Bao et al., 2008). Sequences used were from both isolates (using MDCK cells or eggs) and clinical samples, or unknown origin. All sequencing data for viruses with a full-length H1 subtype HA gene of human and swine influenza A viruses were included. Sequences derived from laboratory strains were excluded, and 1457, 365, and 1332 sequences were obtained for seasonal H1N1, swine H1, and 2009 H1N1 viruses, respectively (supplementary file). Sequences containing minor insertions, minor deletions, or untranslatable codons were excluded. A multiple alignment of nucleotide sequences was constructed using ClustalW.

A phylogenetic tree for the swine influenza virus was constructed by the neighbor-joining method using Molecular Evolutionary Genetics Analysis (MEGA) ver. 4 (Kumar et al., 2008). Based on the tree structure, swine viruses of the North American lineage, including triple reassortant viruses from which 2009 H1N1 virus originated (Garten et al., 2009; Smith et al., 2009), were selected for further analyses.

Evolutionary rate

The evolutionary rate of each group was calculated as follows: up to 100 sequences of each group were selected, and the evolutionary rate was analyzed for selected sequences with a linear regression model as the number of substitutions/site/year compared with A/South Carolina/1/18 (GenBank accession no. AF117241). Periods of evolutionary rates calculated were 40 years, 33 years, 80 years, and 10 months for the seasonal H1N1 (1918–1957), seasonal H1N1 (1977–2009), swine H1 (1930–2009), and 2009 H1N1 viruses (March–December 2009), respectively.

All results were based on pairwise analysis, which was performed using the Maximum Composite Likelihood method in MEGA (ver. 4). The significance of correlations was estimated using Pearson correlation.

Selection pressure

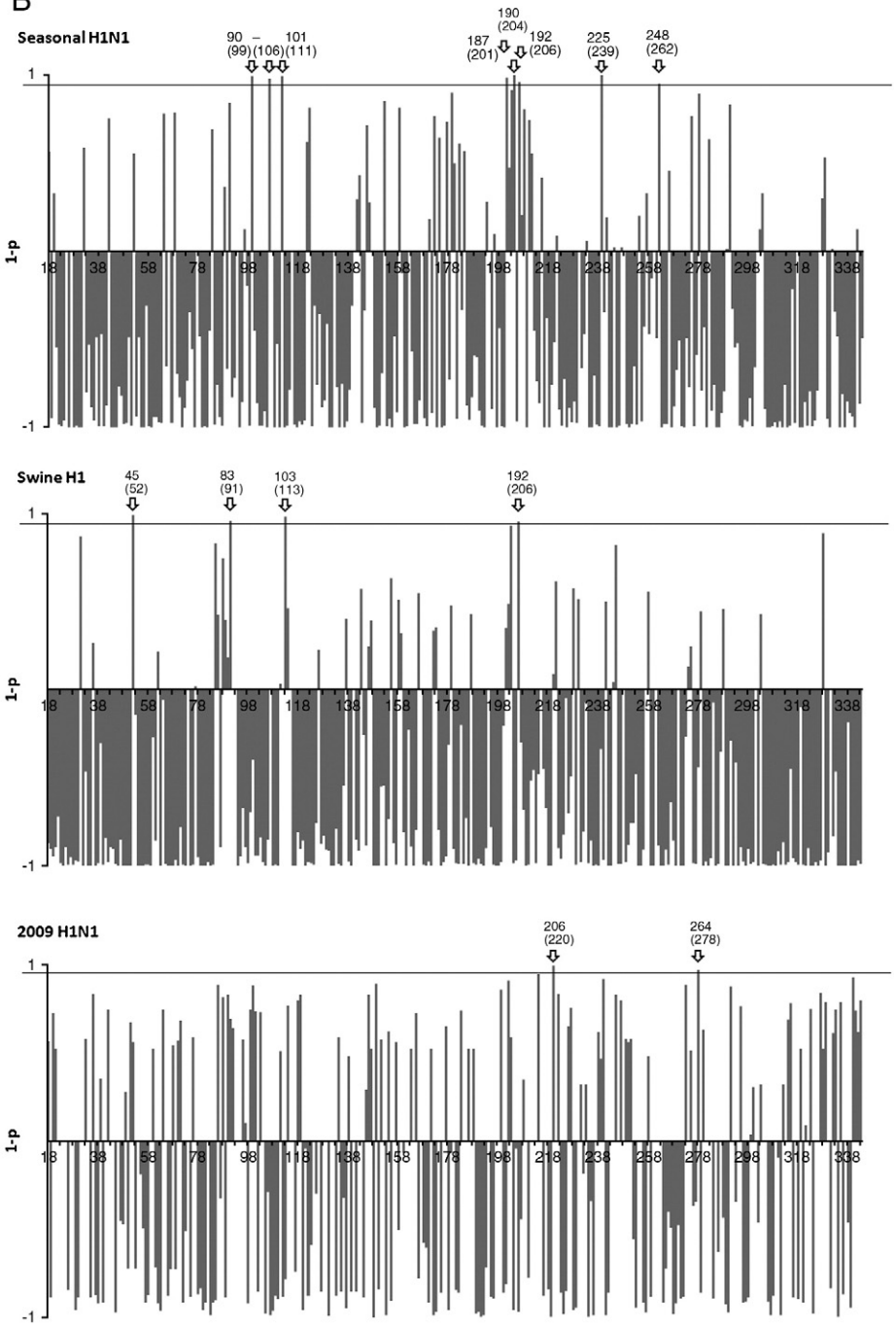
Phylogenetic trees for each dataset of the host were constructed using the maximum likelihood method implemented in PhyML-aLRT (Anisimova and Gascuel, 2006) with the GTR model, which included four rate categories, all parameters of which were estimated from the data.

Positive selection sites were detected using the fixed effects likelihood method, which is based on maximum likelihood estimates. Relative rates of nonsynonymous and synonymous substitutions (dN/dS) in each codon were compared. Sites where $dN/dS>1$ and $dN/dS<1$ were inferred as positively and negatively selected, respectively. Furthermore, global estimates, ω, of dN and dS, averaged over the entire alignment, were compared to calculate the overall strength of selection (Pond et al., 2007). Details of the method are described elsewhere (Campo et al., 2008; Kosakovsky Pond and Frost, 2005; Pond et al., 2007). The differential of evolutionary pressure was also analyzed by HyPhy, which tests the hypothesis as to whether dN/dS at a given site differs between two

A

Virus	Position (H3 [H1] numbering)	dN/dS	*P* value	Description
Seasonal H1N1	90 [99]	inf	0.0063	
	—* [106]	inf	0.022	
	101 [111]	inf	0.0051	
	187 [201]	inf	0.013	antigenicity (Sb)
	190 [204]	9.5	<0.0001	antigenicity (Sb), receptor binding specificity
	192 [206]	2.8	0.037	antigenicity (Sb)
	225 [239]	inf	0.0012	antigenicity (Ca2), receptor binding specificity
	248 [262]	inf	0.048	
	sites involved in receptor binding specificity	1.8**		
Swine H1	45 [52]	Inf	0.0082	
	83 [91]	Inf	0.042	antigenicity (Cb)
	103 [113]	Inf	0.017	
	192 [206]	Inf	0.047	antigenicity (Sb)
	sites involved in receptor binding specificity	0.2**		
2009 H1N1	206 [220]	Inf	0.00046	antigenicity (Ca)
	264 [278]	Inf	0.028	
	sites involved in receptor binding specificity	0.9**		

B



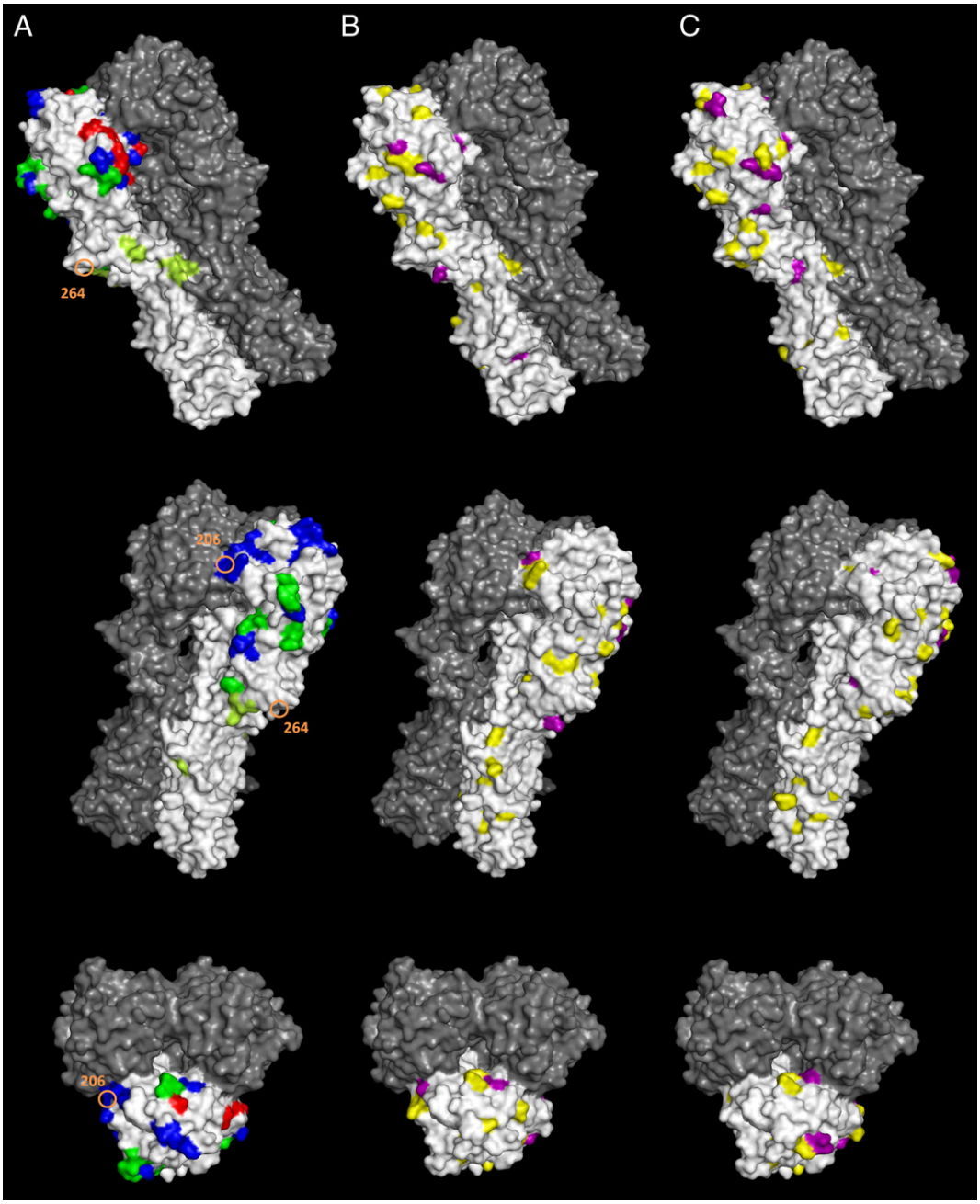


Fig. 4. Sites under different selection pressures. A) Antigenic structure of HA. Antigenic sites are shaded in blue, sites involved in receptor binding specificity are in red, N-glycosylation sequons of Spanish influenza and 2009 H1N1 viruses are in light green, and additional N-glycosylation sequons of the seasonal H1N1 virus are in green. Sites selected positively in 2009 H1N1 virus are indicated by circles. B) Sites under different selection pressure between 2009 H1N1 and seasonal H1N1 viruses. Sites shaded yellow were under higher selection pressure in 2009 H1N1 virus, and those shaded purple were under lower selection pressure. C) Sites under different selection pressure between 2009 H1N1 and swine H1 viruses. Sites shaded yellow were under higher selection pressure in 2009 H1N1 virus and those shaded purple were under lower selection pressure.

datasets along the phylogenetic tree. Details are described elsewhere (Pond et al., 2006, 2007).

3D structures are shown as a solvent-accessible surface representation prepared from A/swine/St-Hyacinthe/148/1990 (H1N1) (PDB ID 1RUY) by PyMOL (Gamblin et al., 2004). Sites were numbered based on H1 and H3 numbering. H1 numbering was performed by counting from the initiation codon to the termination codon based on A/South Carolina/1/18 (H1N1). H3 numbering was performed by separately counting HA1 and HA2 regions (Nobusawa et al., 1991; Stevens et al., 2004). Antigenic sites for H1 and positions of amino

Fig. 3. Selection profile. A) List of sites under positive selection. The significance of fixed effect likelihood (FEL) results for positive selection levels is given as a *P* value. inf, infinity as denominator = 0. *The site does not possess H3 numbering. **Average of selection pressures on sites involved in receptor binding specificity. B) Selection profile for each codon of HA1. The abscissa indicates the codon position. The ordinate indicates the $1-P$ value for each position, and is above or below the horizontal line, respectively, when $dN/dS>1$ or $dN/dS<1$. Upper horizontal lines represent 0.95 so that positions where bars cross the lines above indicate positively selected sites. Positively selected sites are indicated by arrows with their positions as per H3 and H1 numbering.

acids in sites are reported elsewhere (Gerhard et al., 1981; Stevens et al., 2004).

Acknowledgments

This work was supported by the Health and Labor Science Research Grants from the Ministry of Health, Labor, and Welfare, Japan. We are indebted to the medical practitioners Dr. Makoto Shoji, Dr. Kazuhisa Kawamura, Dr. Jun Kayaba, Dr. Syunzo Hayamizu, Dr. Masataka Itano, and Dr. Kikuya Metoki in Sendai, Japan, for collecting clinical specimens. YF is a recipient of a scholarship through the Honjo International Scholarship Foundation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.virol.2010.06.018.

References

Anisimova, M., Gascuel, O., 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Syst. Biol. 55 (4), 539–552.
Bao, Y., Bolotov, P., Dernovoy, D., Kiryutin, B., Zaslavsky, L., Tatusova, T., Ostell, J., Lipman, D., 2008. The influenza virus resource at the National Center for Biotechnology Information. J. Virol. 82 (2), 596–601.
Campo, D.S., Dimitrova, Z., Mitchell, R.J., Lara, J., Khudyakov, Y., 2008. Coordinated evolution of the hepatitis C virus. Proc. Natl. Acad. Sci. USA 105 (28), 9685–9690.
Childs, R.A., Palma, A.S., Wharton, S., Matrosovich, T., Liu, Y., Chai, W., Campanero-Rhodes, M.A., Zhang, Y., Eickmann, M., Kiso, M., Hay, A., Matrosovich, M., Feizi, T., 2009. Receptor-binding specificity of pandemic influenza A (H1N1) 2009 virus determined by carbohydrate microarray. Nat. Biotechnol. 27 (9), 797–799.
de Jong, J.C., Smith, D.J., Lapedes, A.S., Donatelli, I., Campitelli, L., Barigazzi, G., Van Reeth, K., Jones, T.C., Rimmelzwaan, G.F., Osterhaus, A.D., Fouchier, R.A., 2007. Antigenic and genetic evolution of swine influenza A (H3N2) viruses in Europe. J. Virol. 81 (8), 4315–4322.
Earp, L.J., Delos, S.E., Park, H.E., White, J.M., 2005. The many mechanisms of viral membrane fusion proteins. Curr. Top. Microbiol. Immunol. 285, 25–66.
Fitch, W.M., Bush, R.M., Bender, C.A., Cox, N.J., 1997. Long term trends in the evolution of H(3) HA1 human influenza type A. Proc. Natl. Acad. Sci. USA 94 (15), 7712–7718.
Furuse, Y., Suzuki, A., Kamigaki, T., Oshitani, H., 2009. Evolution of the M gene of the influenza A virus in different host species: Large-scale sequence analysis. Virol. J. 6 (1), 67.
Furuse, Y., Suzuki, A., Oshitani, H., 2010. Reassortment between swine influenza A viruses increased their adaptation to humans in pandemic H1N1/09. Infect. Genet. Evol.
Gambaryan, A.S., Robertson, J.S., Matrosovich, M.N., 1999. Effects of egg-adaptation on the receptor-binding properties of human influenza A and B viruses. Virology 258 (2), 232–239.
Gamblin, S.J., Haire, L.F., Russell, R.J., Stevens, D.J., Xiao, B., Ha, Y., Vasisht, N., Steinhauer, D.A., Daniels, R.S., Elliot, A., Wiley, D.C., Skehel, J.J., 2004. The structure and receptor binding properties of the 1918 influenza hemagglutinin. Science 303 (5665), 1838–1842.
Garten, R.J., Davis, C.T., Russell, C.A., Shu, B., Lindstrom, S., Balish, A., Sessions, W.M., Xu, X., Skepner, E., Deyde, V., Okomo-Adhiambo, M., Gubareva, L., Barnes, J., Smith, C.B., Emery, S.L., Hillman, M.J., Rivailler, P., Smagala, J., de Graaf, M., Burke, D.F., Fouchier, R.A., Pappas, C., Alpuche-Aranda, C.M., Lopez-Gatell, H., Olivera, H., Lopez, I., Myers, C.A., Faix, D., Blair, P.J., Yu, C., Keene, K.M., Dotson Jr., P.D., Boxrud, D., Sambol, A.R., Abid, S.H., St George, K., Bannerman, T., Moore, A.L., Stringer, D.J., Blevins, P., Demmler-Harrison, G.J., Ginsberg, M., Kriner, P., Waterman, S., Smole, S., Guevara, H.F., Belongia, E.A., Clark, P.A., Beatrice, S.T., Donis, R., Katz, J., Finelli, L., Bridges, C.B., Shaw, M., Jernigan, D.B., Uyeki, T.M., Smith, D.J., Klimov, A.I., Cox, N.J., 2009. Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. Science 325 (5937), 197–201.
Gerhard, W., Yewdell, J., Frankel, M.E., Webster, R., 1981. Antigenic structure of influenza virus haemagglutinin defined by hybridoma antibodies. Nature 290 (5808), 713–717.
Hancock, K., Veguilla, V., Lu, X., Zhong, W., Butler, E.N., Sun, H., Liu, F., Dong, L., DeVos, J. R., Gargiullo, P.M., Brammer, T.L., Cox, N.J., Tumpey, T.M., Katz, J.M., 2009. Cross-reactive antibody responses to the 2009 pandemic H1N1 influenza virus. N. Engl. J. Med. 361 (20), 1945–1952.
Hay, A.J., Gregory, V., Douglas, A.R., Lin, Y.P., 2001. The evolution of human influenza viruses. Philos. Trans. R. Soc. Lond. B Biol. Sci. 356 (1416), 1861–1870.
Health Protection Agency, 2009. Composition of the Influenza Vaccine.
Hensley, S.E., Das, S.R., Bailey, A.L., Schmidt, L.M., Hickman, H.D., Jayaraman, A., Viswanathan, K., Raman, R., Sasisekharan, R., Bennink, J.R., Yewdell, J.W., 2009. Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. Science 326 (5953), 734–736.
Hoffmann, E., Stech, J., Guan, Y., Webster, R.G., Perez, D.R., 2001. Universal primer set for the full-length amplification of all influenza A viruses. Arch. Virol. 146 (12), 2275–2289.
Igarashi, M., Ito, K., Kida, H., Takada, A., 2008. Genetically destined potentials for N-linked glycosylation of influenza virus hemagglutinin. Virology 376 (2), 323–329.
Igarashi, M., Ito, K., Yoshida, R., Tomabechi, D., Kida, H., Takada, A., 2010. Predicting the antigenic structure of the pandemic (H1N1) 2009 influenza virus hemagglutinin. PLoS ONE 5 (1), e8553.
Inoue, E., Wang, X., Osawa, Y., Okazaki, K., 2010. Full genomic amplification and subtyping of influenza A virus using a single set of universal primers. Microbiol. Immunol. 54 (3), 129–134.
Ito, T., Couceiro, J.N., Kelm, S., Baum, L.G., Krauss, S., Castrucci, M.R., Donatelli, I., Kida, H., Paulson, J.C., Webster, R.G., Kawaoka, Y., 1998. Molecular basis for the generation in pigs of influenza A viruses with pandemic potential. J. Virol. 72 (9), 7367–7373.
Itoh, Y., Shinya, K., Kiso, M., Watanabe, T., Sakoda, Y., Hatta, M., Muramoto, Y., Tamura, D., Sakai-Tagawa, Y., Noda, T., Sakabe, S., Imai, M., Hatta, Y., Watanabe, S., Li, C., Yamada, S., Fujii, K., Murakami, S., Imai, H., Kakugawa, S., Ito, M., Takano, R., Iwatsuki-Horimoto, K., Shimojima, M., Horimoto, T., Goto, H., Takahashi, K., Makino, A., Ishigaki, H., Nakayama, M., Okamatsu, M., Warshauer, D., Shult, P.A., Saito, R., Suzuki, H., Furuta, Y., Yamashita, M., Mitamura, K., Nakano, K., Nakamura, M., Brockman-Schneider, R., Mitamura, H., Yamazaki, M., Sugaya, N., Suresh, M., Ozawa, M., Neumann, G., Gern, J., Kida, H., Ogasawara, K., Kawaoka, Y., 2009. In vitro and in vivo characterization of new swine-origin H1N1 influenza viruses. Nature 460 (7258), 1021–1025.
Kida, H., Ito, T., Yasuda, J., Shimizu, Y., Itakura, C., Shortridge, K.F., Kawaoka, Y., Webster, R.G., 1994. Potential for transmission of avian influenza viruses to pigs. J. Gen. Virol. 75 (Pt 9), 2183–2188.
Kilbourne, E.D., Smith, C., Brett, I., Pokorny, B.A., Johansson, B., Cox, N., 2002. The total influenza vaccine failure of 1947 revisited: major intrasubtypic antigenic change can explain failure of vaccine in a post-World War II epidemic. Proc. Natl. Acad. Sci. U S A 99 (16), 10748–10752.
Kobasa, D., Takada, A., Shinya, K., Hatta, M., Halfmann, P., Theriault, S., Suzuki, H., Nishimura, H., Mitamura, K., Sugaya, N., Usui, T., Murata, T., Maeda, Y., Watanabe, S., Suresh, M., Suzuki, T., Suzuki, Y., Feldmann, H., Kawaoka, Y., 2004. Enhanced virulence of influenza A viruses with the haemagglutinin of the 1918 pandemic virus. Nature 431 (7009), 703–707.
Kosakovsky Pond, S.L., Frost, S.D.W., 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol. Biol. Evol. 22 (5), 1208–1222.
Kumar, S., Nei, M., Dudley, J., Tamura, K., 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief. Bioinform. 9 (4), 299–306.
Matrosovich, M., Tuzikov, A., Bovin, N., Gambaryan, A., Klimov, A., Castrucci, M.R., Donatelli, I., Kawaoka, Y., 2000. Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. J. Virol. 74 (18), 8502–8512.
Mehle, A., Doudna, J.A., 2009. Adaptive strategies of the influenza virus polymerase for replication in humans. Proc. Natl. Acad. Sci. U S A 106 (50), 21312–21316.
Nobusawa, E., Aoyama, T., Kato, H., Suzuki, Y., Tateno, Y., Nakajima, K., 1991. Comparison of complete amino acid sequences and receptor-binding properties among 13 serotypes of hemagglutinins of influenza A viruses. Virology 182 (2), 475–485.
Pond, S.L.K., Frost, S.D.W., Grossman, Z., Gravenor, M.B., Richman, D.D., Brown, A.J.L., 2006. Adaptation to different human populations by HIV-1 revealed by codon-based analyses. PLoS Comput. Biol. 2 (6), e62.
Pond, S. L. K., Poon, A. F. Y., and Frost, S. D. W. (2007). Estimating selection pressures on alignments of coding sequences. Analyses using HyPhy, Vol. 2008.
Raymond, F.L., Caton, A.J., Cox, N.J., Kendal, A.P., Brownlee, G.G., 1986. The antigenicity and evolution of influenza H1 haemagglutinin, from 1950–1957 and 1977–1983: two pathways from one gene. Virology 148 (2), 275–287.
Reid, A.H., Janczewski, T.A., Lourens, R.M., Elliot, A.J., Daniels, R.S., Berry, C.L., Oxford, J.S., Taubenberger, J.K., 2003. 1918 influenza pandemic caused by highly conserved viruses with two receptor-binding variants. Emerg. Infect. Dis. 9 (10), 1249–1253.
Robertson, J.S., Bootman, J.S., Newman, R., Oxford, J.S., Daniels, R.S., Webster, R.G., Schild, G.C., 1987. Structural changes in the haemagglutinin which accompany egg adaptation of an influenza A(H1N1) virus. Virology 160 (1), 31–37.
Rogers, G.N., Paulson, J.C., 1983. Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin. Virology 127 (2), 361–373.
Schulze, I.T., 1997. Effects of glycosylation on the properties and functions of influenza virus hemagglutinin. J. Infect. Dis. 176 (Suppl 1), S24–S28.
Shen, J., Ma, J., Wang, Q., 2009. Evolutionary trends of A(H1N1) influenza virus hemagglutinin since 1918. PLoS ONE 4 (11), e7789.
Skehel, J.J., Wiley, D.C., 2000. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu. Rev. Biochem. 69, 531–569.
Smith, G.J.D., Vijaykrishna, D., Bahl, J., Lycett, S.J., Worobey, M., Pybus, O.G., Ma, S.K., Cheung, C.L., Raghwani, J., Bhatt, S., Peiris, J.S.M., Guan, Y., Rambaut, A., 2009. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature 459 (7250), 1122–1125.
Steinhauer, D.A., 1999. Role of hemagglutinin cleavage for the pathogenicity of influenza virus. Virology 258 (1), 1–20.
Stevens, J., Corper, A.L., Basler, C.F., Taubenberger, J.K., Palese, P., Wilson, I.A., 2004. Structure of the uncleaved human H1 hemagglutinin from the extinct 1918 influenza virus. Science 303 (5665), 1866–1870.
Stevens, J., Blixt, O., Glaser, L., Taubenberger, J.K., Palese, P., Paulson, J.C., Wilson, I.A., 2006. Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. J. Mol. Biol. 355 (5), 1143–1155.

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 (1), 16–23.
Suzuki, Y., 2006. Natural selection on the influenza virus genome. Mol. Biol. Evol. 23 (10), 1902–1911.
Suzuki, Y., 2008. Positive selection operates continuously on hemagglutinin during evolution of H3N2 human influenza A virus. Gene 427 (1–2), 111–116.
Suzuki, Y., Gojobori, T., 1999. A method for detecting positive selection at single amino acid sites. Mol. Biol. Evol. 16 (10), 1315–1328.
The World Health Organizaion, 2010. Pandemic (H1N1) 2009 — update 82, Vol. 2010.
Tumpey, T.M., Garcia-Sastre, A., Taubenberger, J.K., Palese, P., Swayne, D.E., Basler, C.F., 2004. Pathogenicity and immunogenicity of influenza viruses with genes from the 1918 pandemic virus. Proc. Natl. Acad. Sci. U S A 101 (9), 3166–3171.
Tumpey, T.M., Maines, T.R., Van Hoeven, N., Glaser, L., Solorzano, A., Pappas, C., Cox, N.J., Swayne, D.E., Palese, P., Katz, J.M., Garcia-Sastre, A., 2007. A two-amino acid change in the hemagglutinin of the 1918 influenza virus abolishes transmission. Science 315 (5812), 655–659.
Webster, R.G., Bean, W.J., Gorman, O.T., Chambers, T.M., Kawaoka, Y., 1992. Evolution and ecology of influenza A viruses. Microbiol. Rev. 56 (1), 152–179.
Wei, C.J., Boyington, J.C., Dai, K., Houser, K.V., Pearce, M.B., Kong, W.P., Yang, Z.Y., Tumpey, T.M., Nabel, G.J., 2010. Cross-neutralization of 1918 and 2009 influenza viruses: role of glycans in viral evolution and vaccine design. Sci. Transl. Med. 2 (24), 24ra21.
Xu, X., Rocha, E.P., Regenery, H.L., Kendal, A.P., Cox, N.J., 1993. Genetic and antigenic analyses of influenza A (H1N1) viruses, 1986–1991. Virus Res. 28 (1), 37–55.
Yang, H., Carney, P., Stevens, J., 2010. Structure and receptor binding properties of a pandemic H1N1 virus hemagglutinin. PLoS Curr. Influenza RRN1152.
Zhirnov, O.P., Vorobjeva, I.V., Saphonova, O.A., Poyarkov, S.V., Ovcharenko, A.V., Anhlan, D., Malyshev, N.A., 2009. Structural and evolutionary characteristics of HA, NA, NS and M genes of clinical influenza A/H3N2 viruses passaged in human and canine cells. J. Clin. Virol. 45 (4), 322–333.