Tsinghua Science and Technology

Volume 18 | Issue 5 Article 10

2013

Mutations in Hemagglutinin of a Novel Avian-Origin H7N9 Virus That Are Critical for Receptor Binding Specificity

Wei Hu
Department of Computer Science, Houghton College, Houghton, NY 14744, USA

TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 10/11 pp522-529
Volume 18, Number 5, October 2013

Mutations in Hemagglutinin of a Novel Avian-Origin H7N9 Virus That Are Critical for Receptor Binding Specificity

Wei Hu*

Abstract: A novel avian-origin H7N9 influenza virus was discovered in March in China and has caused a total of 131 people infected including 39 deaths in China as of June 9, 2013. Adaptation of avian viruses to efficiently infect humans requires the viral hemagglutinin (HA) binding switches from avian to human type receptors with help of some mutations in HA. As such it is critical for pandemic assessment to discover these mutations as hallmarks of adaptation. To continue our previous study of this novel H7N9 virus, we identified two sets of mutations in HA. The first set of mutations are present in the current circulating strains of 2013 H7N9 in China, and the second set are potential mutations that were found when compared to the HAs of previous human H7 subtype. These two sets of mutations exhibited unique features. The first group of mutations, on average, enhanced the HA binding to human type receptors whereas reduced that to avian types. Further the reduction of avian binding was almost three times of the increase of the human binding. The second group increased the binding to both human and avian types. But the increase in human types was almost three times of that in the avian types. Though different in their way of changing the binding preference, these two sets of mutations both contained more mutations to decrease the avian binding and increase the human binding than those that did the opposite. Our research highlighted the pandemic potential of this novel virus by showing the important mutations that could potentially help it to adapt to human hosts. Our findings offered new insights into the current state of evolution of this virus, which might be helpful for the continued surveillance of the emergence of H7N9 strains having the ability of human-to-human transmission.

Key words: H7N9; influenza; receptor binding specificity; mutation; HA gene

1 Introduction

Influenza A viruses can infect a wide range of host species including birds, horses, pigs, and humans, of which waterfowl are considered the natural reservoir. Occasionally, they can switch hosts and cause outbreaks in new species. The main subtypes of influenza currently circulating in humans are H1N1 and H3N2. Human infections with avian influenza are sporadic, including the H5N1 human infection in Hong Kong in 1997. However, human infections from an animal source might signal the start of an influenza pandemic like the 2009 H1N1 pandemic. Avian influenza usually circulates among birds. In general, they do not replicate efficiently or cause disease in humans. However, a novel avian origin H7N9, capable of poultry to person transmission, was discovered from three patients with severe lower respiratory tract disease in March 2013 in China, raising global public health concerns[1,2]. This virus was found to be better adapted than other avian influenza viruses infecting mammals. The same virus has also been detected in poultry in China, originating from multiple reassortment events since its eight gene segments were found to be closely related to some of the previous

• Wei Hu is with the Department of Computer Science, Houghton College, Houghton, NY 14744, USA. E-mail: wei.hu@houghton.edu.
∗ To whom correspondence should be addressed.
Manuscript received: 2013-08-06; accepted: 2013-08-07

avian influenza viruses in Asia[1]. This was the first time report of human infections with H7N9, although some other H7 viruses such as H7N2, H7N3, and H7N7 have occasionally been found to infect humans. In 2003 there was an H7N7 poultry outbreak in the Netherlands, causing at least 86 poultry workers infected and one of them died. Furthermore, two persons were infected during an outbreak of H7N3 in Canada in 2004[3]. The concern is that the adaptability of this new H7N9 virus appears to be superior to that of the H5N1 influenza viruses[4].

Because the human population does not have pre-existing immunity to H7N9, as of June 9, 2013, 131 people have been infected with this virus and 39 have died, showing a lower fatality rate than that of the H5N1 virus (http://news.xinhuanet.com). Many of the human cases of H7N9 appeared to have contact with infected poultry or contaminated environments. Empirical data suggests that this virus tends to infect middle-aged to elderly males. In contrast, the 2009 H1N1 pandemic virus favored to infect young people because the older adults may posse partial immunity against 2009 H1N1[5]. Avian influenza viruses need undergo polygenic molecular changes to overcome barriers so they can efficiently replicate in and transmit among humans. Unfortunately, the complete mechanism of avian viruses cross species barriers to infect humans is not yet fully elucidated.

There are eight gene segments of influenza A viruses that encode at least 11 proteins with their Hemagglutinin (HA) protein serving both as the viral receptor-binding protein and fusion protein. Influenza viruses have HAs with different specificities for recognizing the Sialic Acids (SA) linked to galactose on the surface of host cells and the different structures of the sialic acid-galactose linkages in birds and humans provide a natural barrier to host shift events. The viruses adapted to birds have an HA receptor binding specificity for α2,3 SA, while those adapted to humans have a specificity for α2,6 SA. Despite their differences in receptor specificity, avian viruses could infect humans. To transmit efficiently among humans, HA derived from an avian virus has to mutate to preferentially recognize the human type receptors. The amino acids at two critical positions 226 and 228 (H3 numbering) in the receptor binding sites of HA are important for the receptor binding specificity, and the mutations Q226L and G228S in HA increase the binding affinity of H2, H3, and H7 subtypes to human type receptors[6, 7]. Therefore, the HA gene plays a key role in determining host range. A recent study surprisingly demonstrated that merely five mutations, four in HA and one in PB2, are required for H5N1 to become transmissible between ferrets via the air[8] and another similar study can be seen in Ref. [9].

In the first study of this novel H7N9 virus[1], mutations Q226L and T160A in HA (H3 numbering), a characteristic of human viruses, were identified from some of the initial three H7N9 patients in China. Q226L was one of the few mutations shown to be associated with transmission of H5N1 viruses by respiratory droplets in ferrets[8, 9]. In addition to Q226L, a study in Ref. [10] suggested that double mutations S158N and T160A in the HA protein of H5N1 could enhance the binding to human type receptors without loss of binding to avian type receptors. In Ref. [11] we applied mutations S158N and T160A to HA of A(H7N9)Hangzhou/1/2013, and found that the avian binding was decreased by 0.6% whereas the human binding was increased by 6.4%. Our findings thus indicated that these two mutations had the same HA binding shift effect on Hangzhou/1/2013 as on the H5N1 virus.

The HA gene of this new virus was found to be close to that of A(H7N3)/duck/Zhejiang/12/2011, the NA gene close to that of A(H7N9)/wild bird/Korea/A14/2011, and the six internal genes close to those of A(H9N2)/brambling/Beijing/16/2012. For pandemic risk assessment, 23 substitutions in HA were discovered in a follow-up study when this new virus was compared with previous avian H7N9[11]. Individual and collective effects of these mutations were analyzed using a bioinformatics approach developed in Refs. [12-27]. Collectively, these 23 mutations tended to enhance the HA binding to human type receptors and decreased that to avian types[11], which clearly is a cause for concern. As more protein sequences from this new virus are made available, we wanted to extend the results in Ref. [11] by uncovering new substitutions in HA and to analyze them for their impact on receptor binding preference. This was the purpose of this study, i.e., to continue search for indicators of adaptation for this virus to human hosts.

2 Matertials and Methdos

2.1 Sequence data

The HA protein sequences of 2013 H7N9 in China

used in this study were retrieved from the EpiFlu Database (http://platform.gisaid.org) of GISAID. These sequences were aligned with MAFFT[28] and phylogeny analysis of these sequences was conducted with MEGA[29] (Fig. 1), which indicated that the sequences of human and avian HAs of 2013 H7N9 in China were closely related.

2.2 Informational spectrum method

The Informational Spectrum Method (ISM) is a computational technique that can be employed to analyze protein sequences[30]. The idea of this approach is to translate a protein sequence into a numerical sequence based on Electron-Ion Interaction Potential (EIIP) of each amino acid. Then the Discrete Fourier Transform (DFT) is applied to this numerical sequence, and the resulting DFT coefficients are used to produce the energy density spectrum. The Informational Spectrum (IS) comprises the frequencies and the amplitudes of this energy density spectrum. According to the ISM theory, the peak frequencies of IS of a protein sequence encode its biological or biochemical functions. The Consensus Informational Spectrum (CIS) of a collection of protein sequences is the product of IS of each individual sequence. The ISM was successfully applied to quantify the effects of HA mutations on the receptor binding preference in Refs. [31-33].

3 Results

In this study, we sought to find HA substitutions from the strains of this new virus in two ways. The first was to identify substitutions that have occurred by examining the HA proteins sequences of this virus, and the second was to compare the HA of this virus with those of the previous human H7 to discover some potential substitutions that may occur in the future. As found in Ref. [11], we consider IS frequency $F(0.285)$ as avian and $F(0.326)$ as human binding frequency for 2013 H7N9 in China.

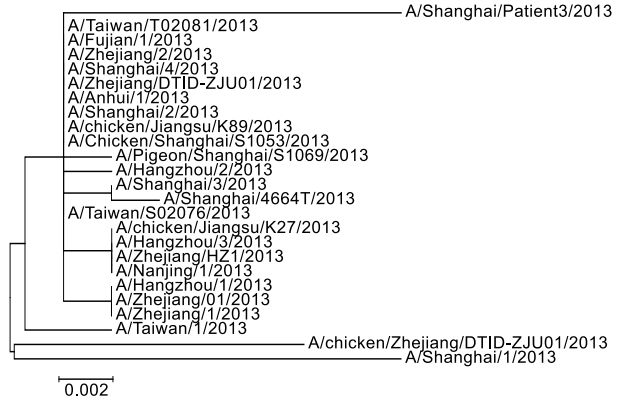


Fig. 1 A polygenic tree of HA sequences of human and avian 2013 H7N9 in China.

3.1 Substitutions in HAs from the current strains of 2013 H7N9 in China

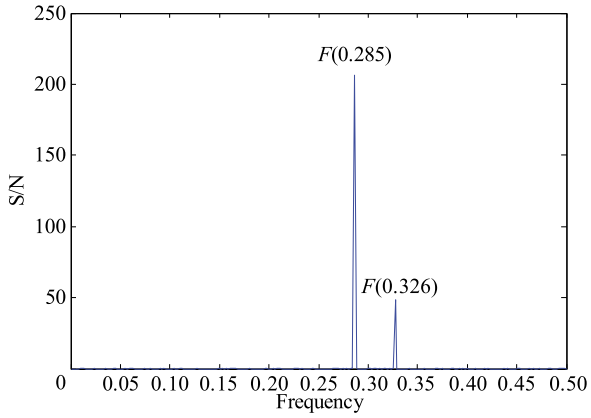
After examining the HA sequences of 2013 H7N9 in China, we found 17 new substitutions in HA (Table 1), in addition to those found in Ref. [11]. The amplitudes of these substitutions were evaluated on the consensus HA sequence of 2013 H7N9 in China at the primary and secondary IS frequencies. Two mutations, A241T and I280M, stand out in Table 1 for their power to switch the primary and secondary binding preferences. On average, the mutations in Table 1 reduced the binding to avian receptors while increased that to human types, manifesting the same trend as the mutations found in Ref. [11]. Two recent reports in Refs. [1, 34] discussed

Table 1 Substitutions in HAs from current strains of 2013 H7N9 in China, where an up arrow indicates an increase and a down arrow suggests a decrease of the IS amplitude compared to the baseline.

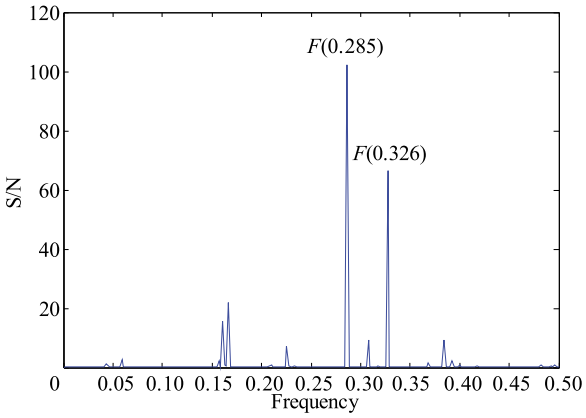
Substitution	IS amplitude at $F(0.285)$	IS amplitude at $F(0.326)$
Baseline (consensus HA of 2013 H7N9 in China)	6.9940	6.4703
V24G	6.9969 ↑	6.4718 ↑
R47M	6.9792 ↓	6.4466 ↓
R47K	6.9000 ↓	6.3406 ↓
C54G	6.8115 ↓	6.0450 ↓
L56P	7.0263 ↑	6.4904 ↑
L57R	7.6597 ↑	7.0615 ↑
I78M	7.5939 ↑	7.0069 ↑
A151S	7.1796 ↑	6.4350 ↓
M155I	6.4072 ↓	5.9294 ↓
S165N	6.6775 ↓	6.6731 ↑
L217P	6.9471 ↓	6.4722 ↑
A241T	6.5615 ↓	6.7706 ↑ (reversed)
N267D	7.7571 ↑	5.6829 ↓
Y273H	7.1140 ↑	6.6548 ↑
I280M	6.6348 ↓	6.7200 ↑ (reversed)
S281N	6.4944 ↓	6.1365 ↓
R303G	6.8164 ↓	6.7644 ↑
Count of ↑	↑ 7 (41%)	↑ 10 (59%)
Count of ↓	↓ 10 (59%)	↓ 7 (41%)
Average	6.973 947 ↓	6.476 571 ↑
Increase	−0.020 05 (−0.29%)	0.006 271 (0.10%)

four mutations in HA, A138S, T160A, G186V, and Q226L, which were all inspected in Ref. [11] for their impact on the HA receptor binding affinity of 2013 H7N9.

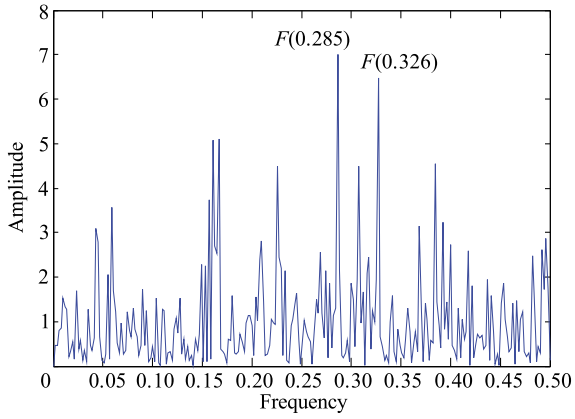
To visualize the IS frequencies and the effect of mutation A241T, we plotted the CIS and IS of the HAs of 2013 H7N9 in China in Fig. 2, which suggested that the primary CIS binding frequency of 2013 H7N9 in China was $F(0.285)$ (avian) and the secondary was $F(0.326)$ (human). The CISs of human and avian H7N9 exhibited very similar patterns as implied by their sequence similarity as seen in Fig. 1. Also we showed that mutation A241T reversed the primary and secondary binding frequencies for the consensus HA of 2013 H7N9. It was intriguing to notice that the consensus of 20 HA sequences of 2013 human H7N9 was identical to that of five 2013 avian H7N9. One recent study demonstrated



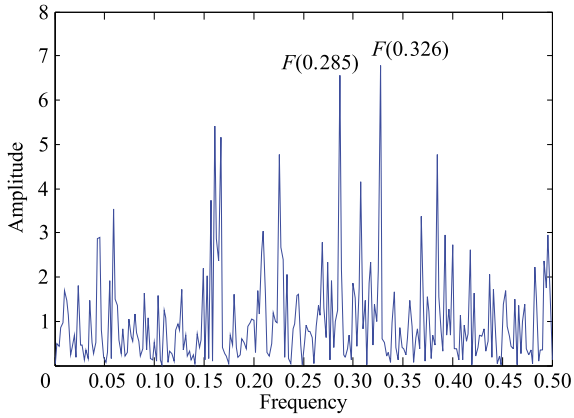
(a) CIS of 20 human H7N9 HAs in China 2013



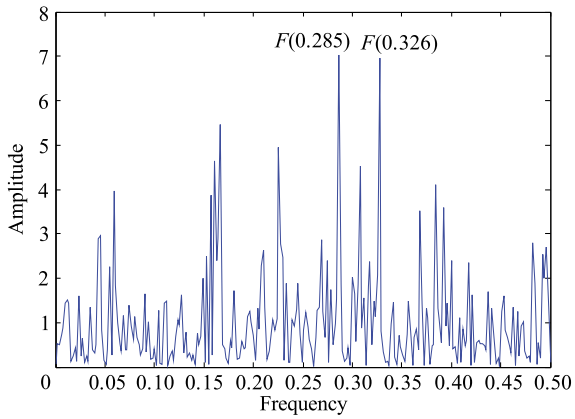
(b) CIS of five avian H7N9 HAs in China 2013



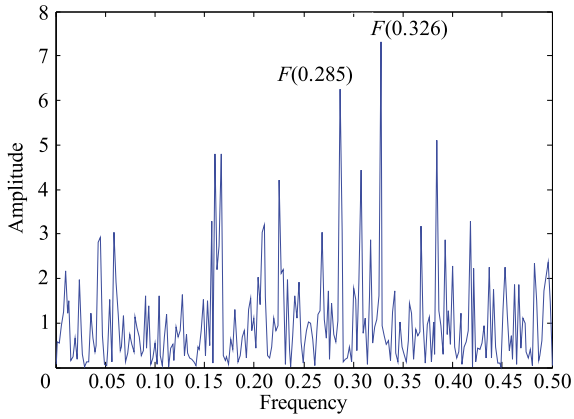
(c) IS of consensus HA seq of 2013 H7N9 in China



(d) IS of consensus HA seq of 2013 H7N9 in China with mutation A241T



(e) IS of consensus HA seq of 2013 H7N9 in China with mutation G228S



(f) IS of HA of A/Shanghai/Patient3/2013 (H7N9)

Fig. 2 CIS or IS of HAs of 2013 H7N9 in China.

that the mutation G228S enhanced significantly the human binding of 2013 H7N9[35], which was in line with our result of this mutation on the consensus of 2013 H7N9 since the consensus already had mutation Q226L (Fig. 2). Finally, we plotted the IS of A(H7N9)/Shanghai/patient3/2013 that contained mutations A241T and I280M to show the changing power of these two mutations on the primary and secondary binding frequencies.

3.2 Potential substitutions in HA of 2013 H7N9 in China when compared to previous human H7

In a preliminary investigation, a set of mutations in HA of 2013 H7N9 were discovered in Ref. [11] when it was compared to previous avian H7N9. Here our goal was to find extra potential mutations in HA of 2013 H7N9 when compared to previous human H7. We applied Random Forests[36,37] to find the top 20 sites in HA that could differentiate the HA sequences of 2013 H7N9 and previous human H7 subtype (Fig. 3) and calculated the amino acid distribution at each such site (Table 2). With the exception of two positions 104 and 109, all the positions in Table 2 display two different dominant amino acids, one from 2013 H7N9 and one from previous human H7. From the amino acid information in Table 2, we found 15 mutations (Table 3). Two mutations, L217P and L217Q (226 in H3 numbering), located at the receptor binding site both decreased the binding to avian type receptors and L217P increased that to human types while L217Q decreased that to human types[38]. Mutations, A128S, A151S, V177G, L217P, and L217Q, were reported in Ref. [34], the first of which was also analyzed in Ref. [11] and the remaining four were included in this report.

This set of mutations increased the binding to both avian and human types with larger amount than those in Table 1 and the increase in human binding was almost three times of that in avian binding. However, they both contained more mutations to decrease the avian binding and increase the human binding than those that did the opposite.

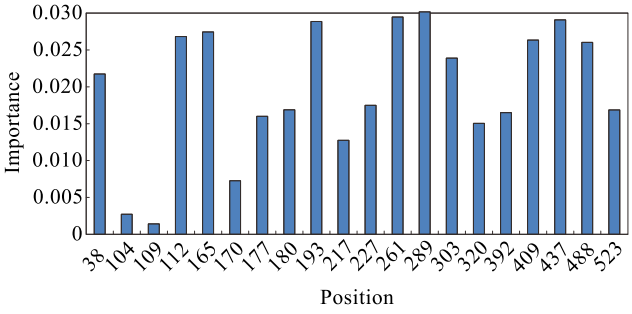


Fig. 3 Top sites in HA identified by Random Forests that could differentiate 2013 H7N9 in China and previous human H7.

Table 2 Amino acid distribution at the important sites in HA in Fig. 3, where the number in parenthesis is the percent of that amino acid.

Position	HAs of 2013 H7N9	HAs of previous human H7	Potential mutation
38	I(100.0)	V(94.1), I(3.9), T(2.0)	I38V
104	E(100.0)	E(64.7), G(29.4), K(3.9), R(2.0)	E104G
109	D(100.0)	D(70.6), Y(25.5), G(2.0), N(2.0)	D108Y
112	A(100.0)	T(94.1), S(5.9)	A112T
165	S(96.0) , N(4.0)	D(94.1), K(5.9)	S165D
170	V(100.0)	I(74.5), V(25.5)	V170I
177	V(96.0), G(4.0)	G(98.0), E(2.0)	V177G
180	A(100.0)	T(94.1), A(3.9), S(2.0)	A180T
193	V(100.0)	I(100.0)	V193I
217	L(76.0), I(12.0), Q(8.0), P(4.0)	Q(100.0)	L217Q
227	M(100.0)	I(90.2), L(5.9), M(3.9)	M227I
261	G(100.0)	E(92.2), D(5.9), K(2.0)	G261E
289	D(100.0)	N(100.0)	D289N
303	R(96.0) , G(4.0)	E(94.1), T(3.9), K(2.0)	R303E
320	G(100.0)	R(90.2), T(3.9), G(3.9), P(2.0)	G320R
392	N(96.0) , T(4.0)	T(94.1), S(3.9), N(2.0)	N392T
409	I(100.0)	M(96.1), L(3.9)	I409M
437	D(100.0)	N(98.0), S(2.0)	D437N
488	M(100.0)	I(94.1), L(5.9)	M488I
523	V(96.0) , A(4.0)	A(100.0)	V523A

4 Discussion and Conclusions

In March 2013, a novel H7N9 virus was isolated from humans in Shanghai and Anhui Province, China. Severe disease in humans caused by a novel influenza virus is a major health concern. Consequently, the advent of this H7N9 has attracted much attention. The H7N9 human infection in China is primarily through contact with clinically normal but infected poultry, in contrast to H5N1 infection that usually results in rapid death in infected poultry. As a result, this novel H7N9 is not easy to detect when it spreads among the poultry. To control human infection, Chinese authorities have slaughtered thousands of birds and closed several live poultry markets.

Table 3 Substitutions in HA of 2013 H7N9 in China when compared to previous human H7, where an up arrow indicates an increase and a down arrow suggests a decrease of the IS amplitude compared to the baseline.

Substitution	IS amplitude at $F(0.285)$	IS amplitude at $F(0.326)$
Baseline (consensus HA of 2013 H7N9 in China)	6.9940	6.4703
I38V	6.9721 ↓	6.5060 ↑
E104G	6.9981 ↑	6.4737 ↑
D108Y	6.7125 ↓	6.4184 ↓
A112T	6.9273 ↓	6.8378 ↑
S165D	7.1069 ↑	6.3067 ↓
V170I	6.9937 ↓	6.4273 ↓
V177G	6.9939 ↓	6.4685 ↓
A180T	7.2600 ↑	6.5719 ↑
V193I	6.9674 ↓	6.5021 ↑
L217Q	6.7662 ↓	6.4342 ↓
M227I	7.5884 ↑	7.0424 ↑
G261E	6.9956 ↑	6.4673 ↓
D289N	6.8491 ↓	6.5577 ↑
R303E	6.8185 ↓	6.7628 ↑
G320R	7.6256 ↑	6.8832 ↑
Count of ↑	↑ 6 (40%)	↑ 9 (60%)
Count of ↓	↓ 9 (60%)	↓ 6 (40%)
Average	7.0383 ↑	6.5773 ↑
Increase from base line	0.0443 (0.63%)	0.1070 (1.65%)

For pandemic risk assessment it is urgent to monitor the gradual process of its adaptation to human hosts, which requires several changes in the virus, most notably in the HA protein. Influenza infection is initiated by viral attachment mediated by its HA protein binding to the receptors on the surface of a host cell. Therefore, HA receptor binding specificity is a major molecular determinant of host range.

A switch in the binding specificity of HA from avian to human types is believed to be a key step towards efficient human to human transmission, although the switching itself alone is not sufficient for this event to happen. Due to the high rates of mutation and replication of influenza, viral adaptation to new hosts primarily manifest as amino acid substitutions in HA to increase the efficiency of viral entry into the new host cells. Therefore, identification of such mutations in HA is of great importance.

As a continuation of our previous study in Ref. [11], the present work aimed to find new mutations in HA of 2013 H7N9 in China as there are more sequences of this virus available now. Using two different approaches, two sets of mutations in the HA of strains of 2013 H7N9 were located. The first set of mutations, collectively, increased the HA binding to human type receptors whereas reduced that to avian types, while the second increased the binding to both types. Furthermore, the first demonstrated the reduction of avian binding was almost three times of the increase of the human binding, but the second increased the binding to human types almost three times of that to avian types.

Our analysis identified mutations in HA that could alter the receptor binding specificity of 2013 H7N9 in China, thereby rendering the current state of the evolution of the HA protein of this new virus, which might be useful to our understanding of this virus.

Acknowledgements

We thank Houghton College for its financial support.

References

[1] R. Gao, B. Cao, Y. Hu, Z. Feng, D. Wang, W. Hu, J. Chen, Z. Jie, H. Qiu, K. Xu, et al., Human infection with a novel avian-origin influenza A (H7N9) virus, *The New England Journal of Medicine*, 2013, doi: 10.1056/NEJMoa1304459.

[2] Y. Chen, W. Liang, S. Yang, N. Wu, H. Gao, J. Sheng, H. Yao, J. Wo, Q. Fang, D. Cui, et al., Human infections with the emerging avian influenza A H7N9 virus from wet market poultry: Clinical analysis and characterisation of viral genome, *The Lancet*, vol. 381, no. 9881, pp. 1916-1925, 2013. doi: 10.1016/S0140-6736(13)60903-4.

[3] J. K. Taubenberger and J. C. Kash, Influenza virus evolution, host adaptation and pandemic formation, *Cell Host Microbe*, vol. 7, no. 6, pp. 440-451, 2010. doi: 10.1016/j.chom.2010.05.009.

[4] J. Z. Shi, G. H. Deng, P. Liu, J. Zhou, L. Guan, W. Li, X. Li, J. Guo, G. Wang, J. Fan, et al., Isolation and characterization of H7N9 viruses from live poultry markets—Implication of the source of current H7N9 infection in humans, *Chin. Sci. Bull.*, 2013. doi: 10.1007/s11434-013-5873-4.

[5] Q. Liu, L. Lu, Z. Sun, G. W. Chen, Y. Wen, and S. Jiang, Genomic signature and protein sequence analysis of a novel influenza A (H7N9) virus that causes an outbreak in humans in China, *Microbes and Infection*, 2013. http://dx.doi.org/10.1016/j.micinf.2013.04.004.

[6] K. Srinivasan, R. Raman, A. Jayaraman, K. Viswanathan, and R. Sasisekharan, Quantitative description of glycan-receptor binding of influenza A virus H7 hemagglutinin, *PLoS ONE*, vol. 8, no.2, p. e49597, 2013. doi:10.1371/journal.pone.0049597.

[7] K. Viswanathan, X. Koh, A. Chandrasekaran, C. Pappas, R. Raman, A. Srinivasan, Z. Shriver, T. M. Tumpey, and R. Sasisekharan, Determinants of glycan receptor specificity of H2N2 influenza A virus hemagglutinin, *PLoS One*, vol. 5, p. e13768, 2010. doi: 10.1371/journal.pone.0013768.

[8] S. Herfst, E. J. Schrauwen, M. Linster, S. Chutinimitkul, E. de Wit, V. J. Munster, E. M. Sorrell, T. M. Bestebroer, D. F. Burke, D. J. Smith, et al., Airborne transmission of influenza A/H5N1 virus between ferrets, *Science*, vol. 336, no. 6088, pp. 1534-1541, 2012. doi:10.1126/science.1213362.

[9] M. Imai and Y. Kawaoka, The role of receptor binding specificity in interspecies transmission of influenza viruses, *Current Opinion in Virology*, vol. 2, no. 2, pp. 160-167, 2012. doi:10.1016/j.coviro.2012.03.003.

[10] W. Wang, B. Lu, H. Zhou, A. L. Jr Suguitan, X. Cheng, K. Subbarao, G. Kemble, and H. Jin, Glycosylation at 158N of the hemagglutinin protein and receptor binding specificity synergistically affect the antigenicity and immunogenicity of a live attenuated H5N1 A/Vietnam/1203/2004 vaccine virus in ferrets, *J. Virol.*, vol. 84, no. 13, pp. 6570-6577, 2010.

[11] W. Hu, Receptor binding specificity and sequence comparison of a novel avian-origin H7N9 virus in China, *Journal of Biomedical Science and Engineering*, vol. 6, no. 5, pp. 533-542, 2013.

[12] W. Hu, Analysis of correlated mutations, stalk motifs, and phylogenetic relationship of the 2009 influenza A virus neuraminidase sequences, *Journal of Biomedical Science and Engineering*, vol. 2, no. 7, pp. 550-558, 2009.

[13] W. Hu, The interaction between the 2009 H1N1 Influenza A hemagglutinin and neuraminidase: Mutations, co-mutations, and the NA stalk motifs, *Journal of Biomedical Science and Engineering*, vol. 3, pp. 1-12, 2010.

[14] W. Hu, Novel host markers in the 2009 pandemic H1N1 influenza A virus, *Journal of Biomedical Science and Engineering*, vol. 3, no. 6, pp. 584-601, 2010.

[15] W. Hu, Nucleotide host markers in the influenza A viruses, *Journal of Biomedical Science and Engineering*, vol. 3, no. 7, pp. 684-699, 2010.

[16] W. Hu, Identification of highly conserved domains in hemagglutinin associated with the receptor binding specificity of influenza viruses: 2009 H1N1, avian H5N1, and swine H1N2, *Journal of Biomedical Science and Engineering*, vol. 3, pp. 114-123, 2010.

[17] W. Hu, Subtle differences in receptor binding specificity and gene sequences of the 2009 pandemic H1N1 influenza virus, *Advances in Bioscience and Biotechnology*, vol. 1, pp. 305-314, 2010.

[18] W. Hu, Correlated mutations in the four influenza proteins essential for viral RNA synthesis, host adaptation, and virulence: NP, PA, PB1, and PB2, *Natural Science*, vol. 2, no. 10, pp. 1138-1147, 2010.

[19] D. King, Z. Miller, W. Jones, and W. Hu, Characteristic sites in the internal proteins of avian and human Influenza viruses, *Journal of Biomedical Science and Engineering*, vol. 3, no. 10, pp. 943-955, 2010.

[20] W. Hu, Highly conserved domains in hemagglutinin of influenza viruses characterizing dual receptor binding, *Natural Science*, vol. 2, no. 9, pp. 1005-1014, 2010.

[21] W. Hu, Host markers and correlated mutations in the overlapping genes of influenza viruses: M1, M2; NS1, NS2; and PB1, PB1-F2, *Natural Science*, vol. 2, no. 11, pp. 1225-1246, 2010.

[22] W. Hu, Receptor binding specificity and origin of 2009 H1N1 pandemic influenza virus, *Natural Science*, vol. 3, no. 3, pp. 234-248, 2011.

[23] W. Hu, New mutational trends in the HA protein of 2009 H1N1 pandemic influenza virus from May 2010 to February 2011, *Natural Science,* vol. 3, no. 5, pp. 379-387, 2011.

[24] W. Hu, Characterization of Asian and North American avian H5N1, *American Journal of Molecular Biology*, vol. 1, no. 2, pp. 52-61, 2011.

[25] W. Hu, Molecular features of highly pathogenic avian and human H5N1 influenza A viruses in Asia, *Computational Molecular Bioscience*, vol. 2, no. 2, pp. 45-59, 2012.

[26] W. Hu, Molecular determinants for receptor binding in hemagglutinin protein of 2009 pandemic H1N1, in *Proceedings of the 6th International Conference Bioinformatics and Biomedical Engineering (ICBBE2012)*, Shanghai, China, 2012, pp. 629-632.

[27] W. Hu, Mutations in hemagglutinin of H5N1 influenza that switch receptor specificity from avian to human types, *Computational Molecular Bioscience*, vol. 3, pp. 32-37, 2013.

[28] K. Katoh, K. Kuma, H. Toh, and T. Miyata, MAFFT version 5: Improvement in accuracy of multiple sequence alignment, *Nucleic Acids Res.*, vol. 33, pp. 511-518, 2005.

[29] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar, MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Molecular Biology and Evolution*, vol. 28, pp. 2731-2739, 2011.

[30] I. Cosic, *The Resonant Recognition Model of Macromolecular Bioreactivity, Theory and Application*. Berlin, Germany: Birkhauser Verlag, 1997.

[31] V. Veljkovic, H. L. Niman, S. Glisic, N. Veljkovic, V. Perovic, and C. P. Muller, Identification of hemagglutinin structural domain and polymorphisms which may modulate swine H1N1 interactions with human receptor, *BMC Structural Biology*, vol. 9, p. 62, 2009.

[32] V. Veljkovic, N. Veljkovic, C. P. Muller, S. Müller, S. Glisic, V. Perovic, and H. Köhler, Characterization of conserved properties of hemagglutinin of H5N1 and human influenza viruses: Possible consequences for therapy and infection control, *BMC Struct. Biol.*, vol. 7, pp. 9-21, 2009.

[33] W. Hu, Quantifying the effects of mutations on receptor binding specificity of influenza viruses, *Journal of Biomedical Science and Engineering*, vol. 3, pp. 227-240, 2010.

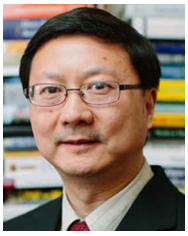
[34] T. Kageyama, S. Fujisaki, E. Takashita, H. Xu, S. Yamada, Y. Uchida, G. Neumann, T. Saito, Y. Kawaoka, and M. Tashiro, Genetic analysis of novel avian A(H7N9) influenza viruses isolated from patients in China, February to April 2013, *Euro Surveill*, vol. 18, no. 15, p. 20453, 2013. http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20453.

[35] K. Tharakaraman, A. Jayaraman, R. Raman, K. Viswanathan, N. W. Stebbins, D. Johnson, Z. Shriver, V. Sasisekharan, and R. Sasisekharan, Glycan receptor binding of the influenza A virus H7N9 hemagglutinin, *Cell*, vol. 153, no. 7, pp. 1486-1493, 2013.

[36] L. Breiman, Random forests, *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.

[37] R. Díaz-Uriarte and S. Alvarez de Andrés, Gene selection and classification of microarray data using random forest, *BMC Bioinformatics,* vol. 7, pp. 3, 2006.

[38] D. Liu, W. Shi, Y. Shi, D. Wang, H. Xiao, W. Li, Y. Bi, Y. Wu, X. Li, J. Yan, et al., Origin and diversity of novel avian influenza A H7N9 viruses causing human infection: Phylogenetic, structural, and coalescent analyses, *The Lancet*, vol. 381, no. 9881, pp. 1926-1932, 2013.



Wei Hu is a professor of math and computer science at Houghton College. He graduated from the University of Kentucky in 1997 with a PhD degree in math and an MS degree in computer science. His research interests are data mining and machine learning.