Infection, Genetics and Evolution 32 (2015) 107–112

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Similarity of currently circulating H1N1 virus with the 2009 pandemic clone: Viability of an imminent pandemic

Rachana Banerjee [a], Ayan Roy [b], Santasabuj Das [c], Surajit Basak [d,e,*]

[a] Department of Bio-Physics, Molecular Biology and Bioinformatics, University of Calcutta, 92, Acharya Prafulla Chandra Road, Kolkata 700009, India
[b] NBU Bioinformatics Facility, Department of Botany, University of North Bengal, Raja Rammohanpur, Siliguri 734013, India
[c] Biomedical Informatics Center, National Institute of Cholera and Enteric Diseases, P-33, C.I.T Road, Scheme-XM, Beliaghata, Kolkata 700010, India
[d] Department of Molecular Biology & Bioinformatics, Tripura University, Suryamaninagar, Tripura 799 022, India
[e] Bioinformatics Centre, Tripura University, Suryamaninagar, Tripura 799 022, India

ARTICLE INFO

Article history:
Received 23 December 2014
Received in revised form 22 February 2015
Accepted 23 February 2015
Available online 28 February 2015

Keywords:
H1N1
Hemagglutinin
Pandemic
Pathogenicity
Evolution

ABSTRACT

The first influenza pandemic in the 21st century commenced in March, 2009 causing nearly 300,000 deaths globally within the first year of the pandemic. In late 2013 and in early 2014, there was gradual increase in the reported case of H1N1 infection and according to World Health Organization (WHO) report, influenza activity increased in several areas of the Southern Hemisphere and was dominated by the H1N1 pandemic strain of 2009. In the present study, a comprehensive comparison of the global amino acid composition and the structural features of all HA gene sequences of H1N1, available in the Flu Database (NCBI), from 1918 to December, 2014 has been performed to trace out the possibility of a further H1N1 pandemic in near future. The results suggest that the increased potential to enhance pathogenicity for the H1N1 samples of 2013 (latter part) and 2014 could lead to a more severe outbreak in the near future.

1. Introduction

The first influenza pandemic of the present century, commonly known as "Mexican flu", commenced in March 2009 with the emergence of a new swine-origin (triple-reassorted) H1N1 influenza A virus that has been remarkably different from the previously circulating H1N1 strains (Dawood et al., 2009). Clinical symptoms associated with the 2009 pandemic H1N1 strain (H1N1 pdm09) include mild respiratory irritations that extend up to severe pneumonia, associated with acute respiratory distress syndrome (ARDS) with the gradual progress of infection (Chowell et al., 2009; Perez-Padilla et al., 2009; Go et al., 2012). The H1N1 pandemic that initiated in 2009 resulted in nearly 300,000 deaths globally within the first year of its outbreak (Dawood et al., 2012). A distinctive feature of influenza virus infection is a successful viral entry into the host system followed by subsequent destruction of the host immune responses and inception of disease (Go et al., 2012; de Jong et al., 2006; Kobasa et al., 2007). Continuous antigenic variations occur in influenza A virus that ensure its competent replication inside the human host (Schmolke and Garcia-Sastre, 2010). H1N1 subtype of influenza A virus has been found to be display variations at the amino acid level which allow them to evade the host immune signals and establish the disease with absolute efficacy. H1N1 pdm09 influenza A virus is a classic example supporting the fact that the antigenic variations correlate strongly with host immune elusion and crucially govern the efficacy of the virus in pronouncing the disease. The origin of human-isolated H1N1 virus has been well explored from the various phylogenetic analyses however, proper know-how of the determinants of cross-species transmission still remain somewhat obscure (Liu et al., 2014; Lin et al., 2009; Landolt and Olsen, 2007; Banerjee et al., 2012).

Hemagglutinin (HA) has been confirmed to be one of the major glycoproteins present on the surface of influenza virus that facilitate proper viral attachment to the host cellular receptors (Guarnaccia et al., 2013; Caton et al., 1982; Gerhard et al., 1981). Successful attachment of HA with host sialic-acid receptors plays a significant role to pave the way for the onset of infection of the respiratory epithelial cells. HA serves to be the primary antigen of the influenza virus and is also efficient in evading the host immune response. Interestingly, mutations causing antigenic drift are typically restricted to the antigenic sites adjacent to the receptor binding site on the globular head of HA (Yewdell and Gerhard, 1981; Both et al., 1983; Skehel et al., 1984). Therefore, it is always

* Corresponding author at: Department of Molecular Biology & Bioinformatics, Tripura University, Suryamaninagar, Tripura 799 022, India. Tel.: +91 9862924152.
E-mail address: basaksurajit@gmail.com (S. Basak).

http://dx.doi.org/10.1016/j.meegid.2015.02.023
1567-1348/

essential to thoroughly scrutinize the properties of the HA protein in H1N1 virus, as it can provide good evidence for elucidating the complex infective mechanisms of the virus. Mutations that involve amino acid changes in antigenic regions of influenza proteins assist proficient immune escape and are frequently termed as antigenic drifts. Such mutations commonly occurs with the genes encoding the HA surface glycoprotein. Thus, HA proteins appear to be the major targets of neutralizing antibodies that circulate as a result of vaccination (Guarnaccia et al., 2013). Several reports have been published in the past one year, highlighting the unique mutations of the HA gene and their impact on the structure and function of the protein, especially the human receptor binding affinity (Wu et al., 2014; Łepek et al., 2014; Linderman et al., 2014).

The new swine-origin H1N1pdm09 (the pandemic strain of H1N1 of 2009) virus was first isolated in April, 2009, from patients with febrile respiratory illnesses in the United States and Mexico and it spread rapidly across the world by human-to-human transmission. Later, the World Health Organization declared the 2009 H1N1 infection a global pandemic. The rapid spread of this swine influenza virus, mainly among young healthy adults and outside the classical influenza season added to the unpredictability of this virus. Thus, the virus and its molecular evolution raised a number of questions, which are of prime international public health concern. After the first reported case of infection with the H1N1pdm09 virus, the virus was found to be transmitted in different countries. The epidemiology of H1N1pdm09 virus in the United Kingdom during 2009–2011 was characterized by 3 distinct waves: first wave, April–August 2009; second wave, September 2009–April 2010; and third wave, August 2010–April 2011 (Sridhar et al., 2013). In the post pandemic period, the virus remained in circulation and showed unusually increased activity and severity in March and April, for three consecutive years, which is certainly unseasonal (Dakhave et al., 2013). Since November 2013, the Public Health Agency of Canada has received a number of reports of illness caused by the influenza A H1N1 flu virus among young and middle-aged adults (http://www.phac-aspc.gc.ca/influenza/ah1n1-eng.php). The most alarming scenario was reported for the 2013–2014 season when CDC (Centers for Disease Control and Prevention) received reports of severe flu illness among young and middle-aged adults, many of whom were infected with the H1N1pdm09 virus (http://www.cdc.gov/washington/fluBrief/CDCInfluenza_E-Brief.pdf). It has also been detected by CDC's influenza surveillance system that most of the patients affected with flu-infection were aged between 18 and 64 years. Notably, similar records were evident among non-elderly adults during the 2009 H1N1 pandemic (http://www.cdc.gov/flu/pastseasons/1314season.htm). In the present study, a comprehensive comparison of the global amino acid composition of all HA gene sequences of H1N1 available in the Flu Database (NCBI) from 1918 to December, 2014 has been performed to trace out significant features pointing towards a further probable H1N1 pandemic in the near future.

2. Methods and calculations

All the gene sequences of haemagglutinin (HA) of the influenza A (H1N1) viruses isolated until October, 2014 were downloaded from the NCBI Influenza Virus Resource database (http://www.ncbi.nlm.nih.gov/genomes/FLU/SwineFlu.html). Single linkage cluster analysis on amino acid usage of all HA genes was performed using STATISTICA (version 6.0, published by Statsoft Inc., USA). Since amino acid usage is multivariate in nature, therefore, it is necessary to analyze this data with multivariate statistical techniques i.e., Correspondence analysis (Greenacre, 1984; Sabbía et al., 2007; Basak et al., 2004; Banerjee et al., 2004; Basak and Ghosh, 2006). Major trend in amino acid usage variation among the HA genes was predicted by Correspondence analysis available in CodonW 1.4.2 (Peden, 2000; http://www.molbiol.ox.ac.uk/cu/).

In order to assess the extent of divergence between different groups or clusters, we used Mahalanobis distance, which is well known in multivariate statistical analysis. Mahalanobis distance (D^2) between two clusters is calculated as:

$$D^2 = (X - \bar{Y})^T S^{-1} (X - \bar{Y})$$

where X is a vector of amino acid usage values for the data points in the Cluster 2, Y is a mean vector of amino acid usage values calculated from the data points in the Cluster 1, S is the variance–covariance matrix of the amino acid usage calculated from the data points in the Cluster 1 (S^{-1} is the inverse matrix of S), and the superscript T is the transposition operator. To test if the group means significantly differ between the clusters, we used Hotelling's T-square distribution that is applied in multivariate statistics in undertaking tests of differences between the multivariate means of different populations. Let vector d be the difference between sample means, n_1 and n_2 the sample sizes, p the number of variables and S the variance–covariance matrix. Then, the test statistic is

$$T^2 = \frac{(n_1 + n_2 - p - 1) n_1 n_2}{p(n_1 + n_2 - 2)(n_1 + n_2)} d' S^{-1} d$$

A transformation of T^2 yields an exact F distribution, so that:

$$F = \frac{n_1 + n_2 - p - 1}{p(n_1 + n_2 - 2)} T^2 \sim F_{p, n_1 + n_2 - p - 1}$$

This can be evaluated on p and $(N - p - 1)$ degrees of freedom, where p is the number of dependent variables and $N = n_1 + n_2$. Therefore, F can be evaluated in terms of statistical significance by computing the p-values when F and the two degrees of freedom are given.

We have constructed homology based structure of seven HA proteins taking one representative from each of the six groups and one HA protein of the year 2013 from group 6, using the crystal structure of HA protein of 2009 H1N1 influenza virus (PDB ID: 3LZG) as template. We also created the heavy chain of the antibody 2D1 computationally using the 2D1 from crystal structure of Fab 2D1 in complex with the 1918 Influenza Virus Hemagglutinin (PDB ID: 3LZF) as template. The interaction between the HA protein and the heavy chain of the antibody 2D1 and energy minimization of the docked structures were carried out using Accelrys Discovery Studio software (version 2.0).

3. Results and discussion

3.1. Similarity in RAAU of HA genes from 2009 and 2014

All HA gene sequences from 1918 to December, 2014 were retrieved from GenBank that counted a total of 4671 and Relative Amino Acid Usage (RAAU) was calculated. Single linkage clustering on RAAU identified six major clusters (Fig. 1). It is noteworthy that all the HA gene sequences isolated from various seasonal outbreaks from the year 1933 to 2009 have been found to be present under Cluster 1, which has been revealed to be the root of the six clusters as denoted by the single linkage distances. Fig. 1 also indicates the year wise variation of amino acid composition of HA genes starting from 1918 and successfully detects very similar patterns of amino acid composition as reflected from the close association between Cluster 3 and Cluster 6. In other words, very similar patterns of amino acid composition of HA genes exists for 2009 (April to December) and June, 2013–2014. When the actual values of RAAU were compared between HApdm09 (HA gene of pandemic H1N1 strain of 2009) and HA genes collectively of 2013 (June to

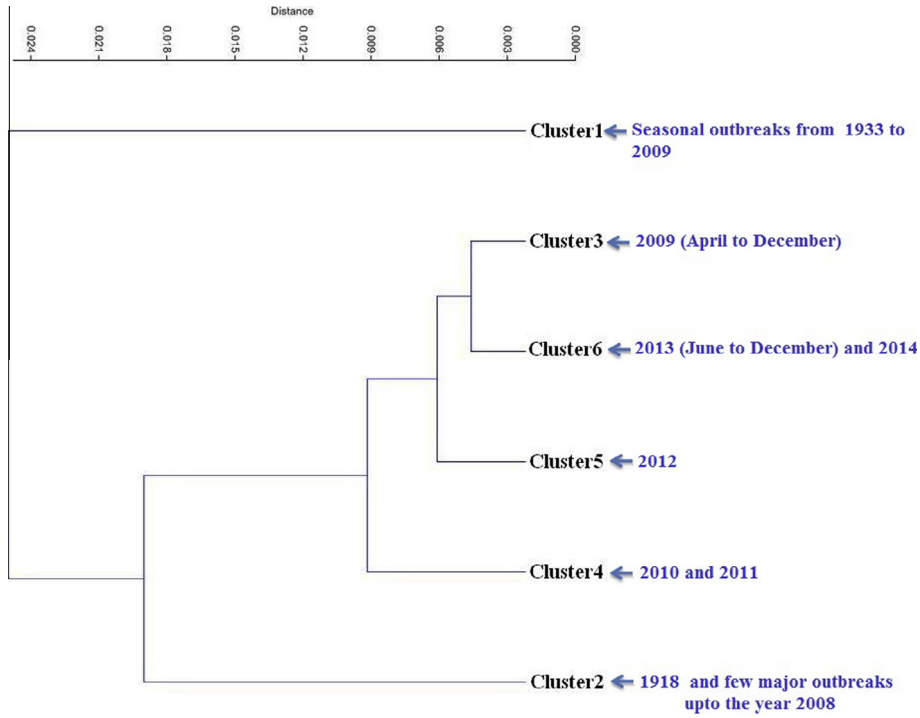

Fig. 1. Single-linkage (Euclidean distances) clustering based on Relative Amino Acid Usage (RAAU) of a total of 4671 HA sequences of H1N1. Six major clusters have been identified. The description illustrating the years of isolation of the sequences representing each of the clusters have been mentioned in the right panel in blue color. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

December)-2014, 5 amino acids were significantly different between two datasets (Supplementary Table 1). Single linkage clustering has also been applied to the RAAU of HA genes (available in GenBank) of H3N2 (Supplementary Fig. 1) to compare the pattern of changes of HA genes between H1N1 and H3N2. In Fig. 1, HApdm2009 (Cluster 3) and HA protein of 2014 (Cluster 6) lie on the same clade and is likely to result from the H1N1 2014 variants. The year wise distribution pattern of HA genes of H3N2 (Supplementary Fig. 1) represents the pattern of seasonal circulation and are very much different from HA genes of H1N1. This observation clearly suggests that genetic drift of H1N1pdm09 may result in the onset of a second wave in 2014. Correspondence Analysis was performed to compare RAAU of different HA gene clusters exemplified through single linkage clustering in Fig. 1. Correspondence analysis provides the distribution of 6 different clusters constituted by the HA genes in a two-dimensional plane of coordinate axis. Subsequently Mahalanobis distance was calculated of each of the clusters from that of Cluster 1 using the position of each of the 4671 HA gene in a two-dimensional plane of coordinate axis as revealed by the Correspondence analysis. The results are given in Table 1. Cluster 6 and Cluster 3, were noted to contain much elevated Mahalanobis distance from Cluster 1 (Table 1), where the data points of Cluster 3 correspond to the H1N1 pandemic in 2009 (H1N1pdm09) and Cluster 6 correspond to HA gene sequences during June, 2013–2014.

However, the Mahalanobis distance decreases for the data points correspond to the year 2010 and 2011 (distance between Cluster 4 and Cluster 1) which predicts no major H1N1 outbreak and indeed it was true. Again the Mahalanobis distance increases for the data points correspond to 2012 and 2013 (distance between Cluster 5 and Cluster 1). From the above results it is clear that amino acid usage pattern of HA genes gives valuable insight about the severity in disease outbreak. We hypothesize that as the Mahalanobis distance increases from non-pandemic cluster (Cluster 1), the chance of occurrence of another pandemic gradually increases. The prevailing trend clearly indicates a greater chance of another large-scale H1N1 outbreak in near future. Here it is worthy to mention that the Mahalanobis distance for the data points corresponding to 2013–2014 not only increased but also showing similar value with that of the data points corresponding to the H1N1 pandemic occurred during April–December, 2009 (H1N1pdm09).

Table 1
Mahalanobis distance of Cluster 2 to Cluster 6 from Cluster 1 using the position of each of the HA sequences, considered in the present study, in a two-dimensional plane of coordinate axis as revealed by the Correspondence analysis.

Clusters	Distance	Representing year
Cluster2	11.95	1918
Cluster3	16.46	2009 (April to December)
Cluster4	11.84	2010 and 2011
Cluster5	13.2	2012
Cluster6	16.98	June, 2013–June, 2014

3.2. Identification of unique mutation in HA gene in 2014 and their role in "antigenic drift"

To further validate the above prediction we have performed molecular docking study. From each of the six different clusters showed in Fig. 1 we have taken one representative HA amino acid sequence and constructed the 3D protein structures through homology modeling, using the crystal structure of HA protein of 2009 (HApdm09) H1N1 influenza virus (PDB ID: 3LZG) as template. Information about the above six sequences has been provided in Table 2.

The infection is started by the binding of the HA proteins to the sialic acid receptor of the target cells, the viral genome enters and infects the target cells after the binding. So, inhibiting this binding by antibodies is an important way against flu. The heavy chain of

Table 2
Information about six HA amino acid sequences (chosen as representative of six different clusters from Fig. 1) that have been used for the construction of 3D protein structures through homology modeling.

Cluster	Year of isolation	Country	Accession number
Cluster 1	1934	Puerto Rico	ABO21709
Cluster 2	1918	South Carolina	AAD17229
Cluster 3	2009	California	ACS45035
Cluster 4	2011	Mexico	AEA74031
Cluster 5	2013	Virginia	AGL09824
Cluster 6	2014	California	AHV83788

Table 3
Binding energy for six docked complexes, representing the heavy chain of the antibody 2D1 with each of the six HA proteins as mentioned in Table 2.

Cluster	Year of isolation	Binding energy (kcal/mol)
Cluster 1	1934	−2719.87
Cluster 2	1918	−2749.18
Cluster 3	2009	−2769.18
Cluster 4	2011	−2662.7
Cluster 5	2013	−2793.37
Cluster 6	2014	−2850.74

the antibody 2D1 (isolated from an elderly survivor of the 1918 pandemic) (Xu et al., 2010) has been generated computationally using the 2D1 template (PDB id: 3LZF). A docking study was performed of heavy chain of the antibody 2D1 with each of the six HA proteins and the binding energy was calculated for each of the six docked complexes (Table 3). The energy values indicate higher binding energy for the HA protein of 2014 followed by 2013, compared to the binding energy of HA protein of the pandemic year 2009 (HApdm09). It reflects some obvious changes at the HA amino acid sequences of 2014 and 2013 compared to 2009. Multiple sequence alignment of the HA protein sequences from the year 2014, 2013 and 2009 reveals two mutations (S142T and S160T) for HA sequence in 2013 and 2014, with respect to the HApdm09 (Fig. 2). The mutation, S142T has been found to be present within epitopic region Sb and the other (S160T) is present within the epitopic region Ca. In addition, the 2013 HA sequence contains another mutation within the epitopic region Ca i.e., A98T. Moreover, the 2014 HA sequence contains another mutation with respect to the HApdm09, within the epitopic region Sa i.e., K120Q. To determine the role of the above mutations in binding of the HA sequences with the antibody 2D1, we substituted these residues of HA protein from the specified year with the residues present at the corresponding sites in the year 2009 and measured the binding energy. The results are presented in Table 4. The binding energy suggests that the wild form (HA-2D1 complex of 2013 and 2014) are more stable than the mutant form (i.e., HA-2D1 complex of 2009). It provides support to the fact that all of the three mutations have played substantial role in increasing the binding affinity of the HA sequence of the year 2013 and 2014 with the antibody 2D1 in comparison to that of the HApdm09. According to Xu et al. (2010), the Sa epitopic site comprises the centerpiece of the antibody-binding surface for 2D1. Here we should focus to an important finding that the HA sequence from the year 2014 (belonging to Cluster 6), contains a unique mutation within the epitopic region Sa i.e., K120Q. Notably, the binding energy of the HA sequence from the year 2014, is much higher and substitution of the residue Q with K in the position 120 in 2014 HA sequences reduces its binding efficacy remarkably. This result points towards the significance of the mutation K120Q, at the epitopic region Sa, for the 2014 HA sequence, in enhancing its antigenic feature.

The present study provides an evolutionary trend of hemagglutinin with respect to its pandemic potential. Influenza A virus remains an important human pathogen due largely to its ability to evade antibodies specific for HA. This “antigenic drift” is due to accumulation of amino acid substitutions in HA that can alter viral infectivity. Comparison with HA genes of H3N2 clearly suggests that genetic drift of H1N1pdm09 may result in the onset of a second wave pandemic from this strain in 2014–2015. The specific mutations present in the HA gene sequences of 2013 and 2014

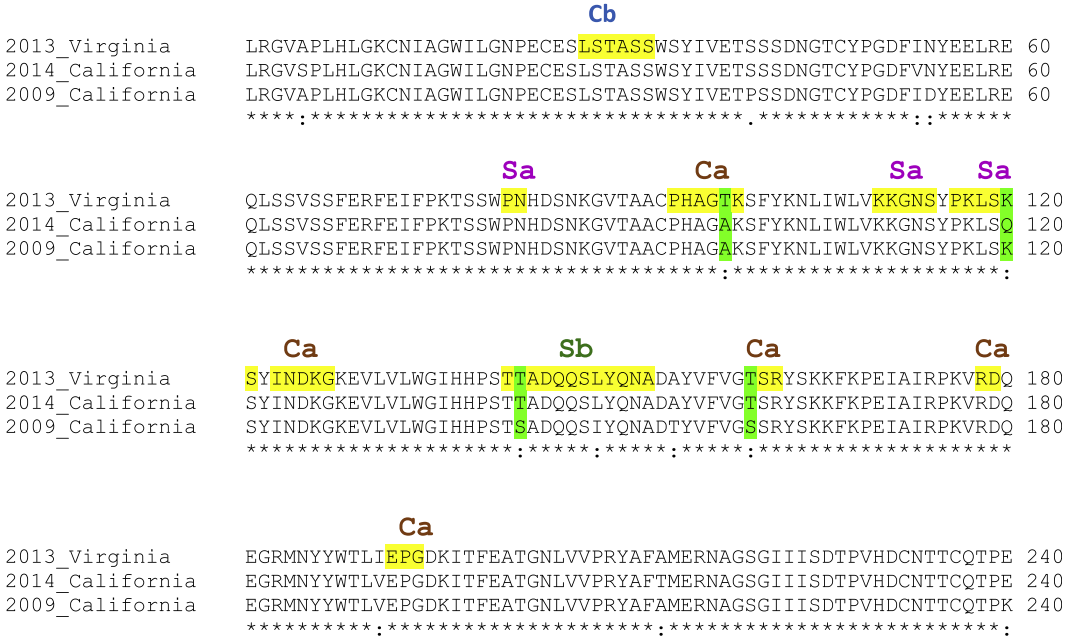

Fig. 2. Multiple sequence alignment of HA sequences from 2009, 2013 and 2014. Epitopic regions have been marked in yellow and unique mutations within the HA sequences of the year 2013 and 2014 have been marked in green. Name of the epitopic regions have been mentioned above each of the regions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Variation of binding energies of wild and mutant types complexes of the heavy chain of the antibody 2D1 with the HA proteins.

Mutant description	Binding energy (kcal/mol)
2013 (wild form)	−2793.37
2013 (T98A)	−2789.01
2014 (wild form)	−2850.74
2014 (Q120K)	−2731.03

are perhaps the result of "antigenic drift" that actually changes the binding energy of the docked complex of HA-2D1 for those years.

4. Conclusions

The evolutionary trend suggests the previously unappreciated potential for antibody escape for the very recent H1N1 samples. The emergence of escape variants should make us rethink pathogenesis, as antibody escape may inadvertently increase pathogenesis. It was observed for the first global H1N1 pandemic (1918) that the second wave became much more lethal (Taubenberger and Morens, 2006). It would not be surprising that H1N1 influenza virus of 2013 and 2014 could also lead to a potentially more dangerous second wave in the near future as suggested by Mahalanobis distance.

Indeed, in the late 2013 and in early 2014, there is gradual increase in the reported case of H1N1 infection. According to World Health Organization (WHO) report, Influenza activity increased in several areas of the Southern Hemisphere and was dominated by the 2009 pandemic H1N1 strain over the 2-week period from Jul 7 through Jul 20, 2013. Data were reported from 73 countries and The Global Influenza Surveillance and Response System (GISRS) tested 24,825 specimens. Of the 1884 specimens positive for influenza, 1590 (84.4%) were influenza A and 294 (15.6%) were influenza B. Of the 1310 type A specimens that were sub-typed, 719 (54.9%) were 2009 H1N1. These findings indicate the need to identify and protect groups at highest risk for adverse outcomes.

The H1N1 virus which emerged in 2009 to trigger a pandemic that year is active again this year. According to recent report (January, 2014) of CDC, the estimated proportion of flu-related deaths is above the epidemic threshold, this year. More importantly, in 2014, the number of pediatric deaths due to flu, as reported to CDC, is increasing rapidly (http://www.cdc.gov/flu/news/flu-severity-rises.htm). CDC reports high rates of hospitalization and deaths among people 18–64 years old. A similar age distribution for hospitalizations for flu was seen during the 2009 H1N1 pandemic. More deaths than usual (nearly 60% of flu deaths) have occurred in the 25–64 age group; a pattern similar to the 2009 pandemic. Fatalities by H1N1 in other parts of the world like Australia, New Zealand and Pakistan have been also reported in 2014 (http://www.un-influenza.org/?q=content/unsic-news-pouch-avian-and-pandemic-influenza-10-mar-2014). According to a very recent report of February 14, 2015, in U.S., the proportion of deaths attributed to influenza like diseases in this season was above the epidemic threshold (http://www.cdc.gov/flu/weekly/#S2). India has seen a massive spurt in cases of swine flu this year with 6298 cases already being reported and the death toll from swine flu in 2015 alone has risen to 585 across the country since February 12 (http://timesofindia.indiatimes.com/india/Swine-flu-death-nears-600-in-India/articleshow/46271578.cms). While H1N1 viruses have continued to circulate since 2009, this is the first season since the pandemic that H1N1 has been circulating so widely (http://www.flu.gov/about_the_flu/current_flu/index.html). Hence, public awareness is very much important to fight against H1N1 infection in the imminent season.

Acknowledgments

Authors acknowledge the facility at DBT-Bioinformatics Centre, Tripura University for carrying out the computational genomics work. Authors also acknowledge the facility at Biomedical Informatics Centre of Indian Council of Medical Research at National Institute of Cholera and Enteric Diseases for carrying out the molecular modeling and docking work.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.meegid.2015.02.023.

References

Banerjee, T., Basak, S., Gupta, S.K., Ghosh, T.C., 2004. Evolutionary forces in shaping the codon and amino acid usages in *Blochmannia floridanus*. J. Biomol. Struct. Dyn. 22, 13–23.
Banerjee, R., Roy, A., Ahmad, F., Das, S., Basak, S., 2012. Evolutionary patterning of hemagglutinin gene sequence of 2009 H1N1 pandemic. J. Biomol. Struct. Dyn. 29, 733–742.
Basak, S., Ghosh, T.C., 2006. Temperature adaptation of synonymous codon usage in different functional categories of genes: a comparative study between homologous genes of *Methanococcus jannaschii* and *Methanococcus maripaludis*. FEBS Lett. 580, 3895–3899.
Basak, S., Banerjee, T., Gupta, S.K., Ghosh, T.C., 2004. Investigation on the causes of codon and amino acid usages variation between thermophilic *Aquifex aeolicus* and mesophilic *Bacillus subtilis*. J. Biomol. Struct. Dyn. 22, 205–214.
Both, G.W., Sleigh, M.J., Cox, N.J., Kendal, A.P., 1983. Antigenic drift in influenza virus H3 hemagglutinin from 1968 to 1980: multiple evolutionary pathways and sequential amino acid changes at key antigenic sites. J. Virol. 48, 52–60.
Caton, A.J., Brownlee, G.G., Yewdell, J.W., Gerhard, W., 1982. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell 31, 417–427.
Chowell, G., Bertozzi, S.M., Colchero, M.A., Lopez-Gatell, H., Alpuche-Aranda, C., Hernandez, M., Miller, M.A., 2009. Severe respiratory disease concurrent with the circulation of H1N1 influenza. N. Engl. J. Med. 361, 674–679.
Dakhave, M., Khirwale, A., Patil, K., Kadam, A., Potdar, V., 2013. Whole-genome sequence analysis of postpandemic influenza A (H1N1)pdm09 virus isolates from India. Genome Announc. 1, e00727-13.
Dawood, F.S., Jain, S., Finelli, L., Shaw, M.W., Lindstrom, S., Garten, R.J., Gubareva, L.V., Xu, X., Bridges, C.B., Uyeki, T.M., 2009. Emergence of a novel swine-origin influenza A (H1N1) virus in humans. N. Engl. J. Med. 360, 2605–2615.
Dawood, F.S., Iuliano, A.D., Reed, C., Meltzer, M.I., et al., 2012. Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study. Lancet Infect. Dis. 12, 687–695.
de Jong, M.D., Simmons, C.P., Thanh, T.T., Hien, V.M., Smith, G.J., Chau, T.N., Hoang, D.M., Chau, N.V., Khanh, T.H., Dong, V.C., Qui, P.T., Cam, B.V., Ha do, Q., Guan, Y., Peiris, J.S., Chinh, N.T., Hien, T.T., Farrar, J., 2006. Fatal outcome of human influenza A (H5N1) is associated with high viral load and hypercytokinemia. Nat. Med. 12, 1203–1207.
Gerhard, W., Yewdell, J., Frankel, M.E., Webster, R., 1981. Antigenic structure of influenza virus haemagglutinin defined by hybridoma antibodies. Nature 290, 713–717.
Go, J.T., Belisle, S.E., Tchitchek, N., Tumpey, T.M., Ma, W., Richt, J.A., Safronetz, D., Feldmann, H., Katze, M.G., 2012. Pandemic H1N1 influenza virus elicits similar clinical course but differential host transcriptional response in mouse, macaque, and swine infection models. BMC Genomics 13, 627.
Greenacre, M.J., 1984. Theory and Applications of Correspondence Analysis. Academic Press, London, 34.
Guarnaccia, T., Carolan, L.A., Maurer-Stroh, S., Lee, R.T., Job, E., Reading, P.C., Petrie, S., McCaw, J.M., McVernon, J., Hurt, A.C., Kelso, A., Mosse, J., Barr, I.G., Laurie, K.L., 2013. Antigenic drift of the pandemic 2009 A (H1N1) influenza virus in A ferret model. PLoS Pathog. 9, e1003354.
Kobasa, D., Jones, S.M., Shinya, K., Kash, J.C., Copps, J., Ebihara, H., Hatta, Y., Kim, J.H., Halfmann, P., Hatta, M., Feldmann, F., Alimonti, J.B., Fernando, L., Li, Y., Katze, M.G., Feldmann, H., Kawaoka, Y., 2007. Aberrant innate immune response in lethal infection of macaques with the 1918 influenza virus. Nature 445, 319–323.
Landolt, G.A., Olsen, C.W., 2007. Up to new tricks-A review of cross-species transmission of influenza A viruses. Anim. Health Res. Rev. 8, 1–21.
Łepek, K., Pająk, B., Siedlecki, P., Niemcewicz, M., Kocik, J., Wu, H.S., Yang, J.R., Kucharczyk, K., Szewczyk, B., 2014. Genetic diversity of hemagglutinin gene of A (H1N1) pdm09 influenza strains isolated in Taiwan and its potential impact on HA-neutralizing epitope interaction. Hum. Vaccine Immunother. 10, 577–585.
Lin, T., Wang, G., Li, A., Zhang, Q., Wu, C., Zhang, R., Cai, Q., Song, W., Yuen, K.Y., 2009. The hemagglutinin structure of an avian H1N1 influenza A virus. Virology 392, 73–81.
Linderman, S.L., Chambers, B.S., Zost, S.J., Parkhouse, K., Li, Y., Herrmann, C., Ellebedy, A.H., Carter, D.M., Andrews, S.F., Zheng, N.Y., Huang, M., Huang, Y.,

Strauss, Shaz, B.H., Hodinka, R.L., Reyes-Terán, G., Ross, T.M., Wilson, P.C., Ahmed, R., Bloom, J.D., Hensley, S.E., . Potential antigenic explanation for atypical H1N1 infections among middle-aged adults during the 2013–2014 influenza season. Proc. Natl. Acad. Sci. USA 111, 15798–15803.

Liu, Y., Zhang, L., Zhou, Y., 2014. Comparative analysis of H1N1 avian influenza virus by multiple sequence alignment and support vector machine. J. Gene Ther. 2, 4.

Peden, J., 2000. Analysis of Codon Usage (Ph.D. thesis). University of Nottingham.

Perez-Padilla, R., De la Rosa-Zamboni, D., Ponce De Leon, S., Hernandez, M., Quinones-Falconi, F., Bautista, E., Ramirez-Venegas, A., Rojas-Serrano, J., Ormsby, C.E., Corrales, A., Higuera, A., Mondragon, E., Cordova-Villalobos, J.A., 2009. Pneumonia and respiratory failure from swine-origin influenza A (H1N1) in Mexico. N. Engl. J. Med. 361, 680–689.

Sabbía, V., Piovani, R., Naya, H., Rodríguez-Maseda, H., Romero, H., Musto, H., 2007. Trends of amino acid usage in the proteins from the human genome. J. Biomol. Struct. Dyn. 25, 55–59.

Schmolke, M., Garcia-Sastre, A., 2010. Evasion of innate and adaptive immune responses by influenza A virus. Cell. Microbiol. 12, 873–880.

Skehel, J.J., Stevens, D.J., Daniels, R.S., Douglas, A.R., Knossow, M., Wilson, I.A., Wiley, D.C., 1984. A carbohydrate side chain on hemagglutinins of Hong Kong influenza viruses inhibits recognition by a monoclonal antibody. Proc. Natl. Acad. Sci. USA 81, 1779–1873.

Sridhar, S., Begom, S., Bermingham, A., Hoschler, K., Adamson, W., Carman, W., Van Kerkhove, M.D., Lalvani, A., 2013. Incidence of influenza A(H1N1)pdm09 infection, United Kingdom. 2009–2011. Emerg. Infect. Dis. 19, 1866–1869.

Taubenberger, J.K., Morens, D.M., 2006. 1918 Influenza: the mother of all pandemics. Emerg. Infect. Dis. 12, 15–22.

Wu, N.C., Young, A.P., Al-Mawsawi, L.Q., Olson, C.A., Feng, J., Qi, H., Chen, S.H., Lu, I.H., Lin, C.Y., Chin, R.G., Luan, H.H., Nguyen, N., Nelson, S.F., Li, X., Wu, T.T., Sun, R., 2014. High-throughput profiling of influenza A virus hemagglutinin gene at single-nucleotide resolution. Sci. Rep. 4, 4942.

Xu, R., Ekiert, D.C., Krause, J.C., Hai, R., Crowe Jr., J.E., Wilson, I.A., 2010. Structural basis of pre-existing immunity to the 2009 H1N1 pandemic influenza virus. Science 328, 357–360.

Yewdell, J.W., Gerhard, W., 1981. Antigenic characterization of viruses by monoclonal antibodies. Annu. Rev. Microbiol. 35, 185–206.