Trends in Bioinformatics

ISSN 1994-7941

science alert

ANSInet
an open access publisher
http://ansinet.com

Trends in Bioinformatics 6 (1): 1-9, 2013
ISSN 1994-7941 / DOI: 10.3923/tb.2013.1.9

Evolutionary Network to Predict the Reassortment of Avian-human A/H5N1 Influenza Virus in India

R. Mathur and N. Adlakha
Department of Applied Mathematics and Humanities, S.V. National Institute of Technology, Surat-395007, Gujarat, India

Corresponding Author: N. Adlakha, Department of Applied Mathematics and Humanities, S.V. National Institute of Technology, Surat-395007, Gujarat, India

ABSTRACT

Highly Pathogenic Avian Influenza (HPAI) A/H5N1 virus has posed a great threat to humans with its high rate of fatality and mortality. These viruses have different segments in its genome. The evolution of these viruses involves mutation and reassortment, where these segments are interchanged among viruses and co-infect the host called mixing vessel for the reassortment of viruses. In this study, the viral strain of the Hemagglutinin (HA) gene of the surface protein of human A/H5N1 virus is analyzed by diversification rate to produce the evolutionary network among the viruses of several countries. The evolutionary network obtained provides the information about the evolution of reassorted viruses. Results suggested that the various viruses of Indian isolates are clustered well with the viruses of other countries indicating the transmission of viruses from these places to India. It is also observed that the reassortment between the pairs of sequences of (human USA-1918, avian West Bengal 2008) and (human USA-1918, human Cambodia-2011) can occur to create new pandemic H5N1 viruses and hence better understanding is required to control the emergence of these new viruses in India.

Key words: Highly pathogenic avian influenza A/H5N1 virus, HA gene, evolutionary network, reassortment, pandemic, India

INTRODUCTION

Reassortment is one of the major mechanisms among the influenza viruses which lead to the new fatal viruses that can cause a threat to the human race (Li *et al*., 2010; Octaviani *et al*., 2010). Influenza viruses are classified into three categories, designated A, B and C. The C type Viruses are common that generally cause no symptoms or mild respiratory illness and are not considered for public health concern. The viruses of type B cause sporadic outbreaks of more severe respiratory disease, particularly among young children. Both B and C type viruses are human viruses; C viruses are stable but A and B viruses are prone to mutation. Initially, Influenza A viruses are of great concern among birds, pigs, mammals, etc but now they are transmitting towards the human populations as well (Khanna *et al*., 2006; Rao, 2008).

The influenza virus is contagious to humans and a number of animals with specific contagiousness towards some other species. The common symptoms of this disease are cold, headache, cough and fever (Tambunan *et al*., 2010). The complete genome of the influenza A virus is composed of eight segments of single stranded RNA namely Hemagglutinin (HA), Neuraminidase (NA), matrix protein (M), Nonstructural (NS) protein, Nucleocapsid Protein (NP), Polymerase Basic (PB1 and PB2) and Polymerase Acidic (PA) but HA and NA are two major glycoproteins responsible

for viral particle located on the virus surface. Also, during the process of reassortment these segments are able to be interchanged or replicated with each other. The mainly used tool for phylogenetic analysis is the phylogenetic tree that does not suits well to represent reassorted events. In typical phylogenetic thinking, changes such as mutations are vertically mapped along the branches between an ancestor and a descendant. In contrast to mutations, reassortment or reticulate events would have to be mapped horizontally across the branches representing divergent lineages of species.

As the evolution of different species generally based on the concepts of vertical and horizontal change, the analysis of reassortment of influenza A/H5N1 virus and other organisms that evolve via Horizontal Gene Transfer (HGT) is best represented by a network. So, considerably there is keen interest in developing alternative methods that can capture HGT in such type of viruses (Bokhari and Janies, 2010; Sinha *et al*., 2009).

It is also suggested from the literature that the high pathogenicity of these viruses like A/H5N1 is responsible for so many pandemic outbreaks in the world. The first pandemic is the H1N1 influenza pandemic of 1918-1919, also called as mother of all pandemics (Flavia and Natarajaseenivasan, 2011). The other one is the Highly Pathogenic Avian Influenza (HPAI) H5N1 virus representing the increasing global concern. The viral strain of the first avian influenza outbreak was observed in 1997 and prior to that it was circulating in the poultry populations of certain regions of Asia. Since the first avian influenza outbreak in 1997, the main concern is that, the highly pathogenic influenza A/H5N1 virus might either mutate or reassort its gene segments with human influenza viruses during the co-infection of a single host called mixing vessel (Fig. 1), resulting in a new virus that would be highly lethal and transmissible from person to person (Bourouiba *et al*., 2011; Khanna *et al*., 2006).

Till now, H5N1 remains an avian epidemic and by the previous literature it is not much prevalent in human population. Worldwide, as of 02 May 2012, WHO has reported 603 laboratory confirmed human cases of infection with avian influenza and 356 deaths since the first case was reported in 2003 and only one case of human influenza is present in India till date (WHO, 2012). But as a matter of precaution, it should be informed that such viruses can acquire the ability to spread efficiently among humans either through adaptation or reassortment or both (Octaviani *et al*., 2010).

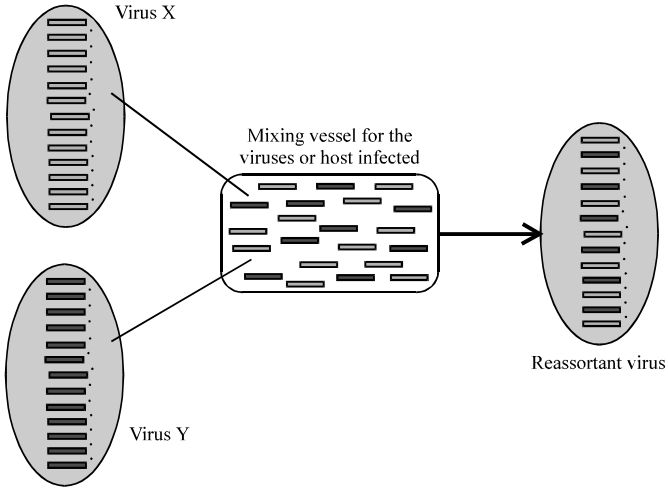

Fig. 1: Reassortant virus comes out by mixing of two different viruses

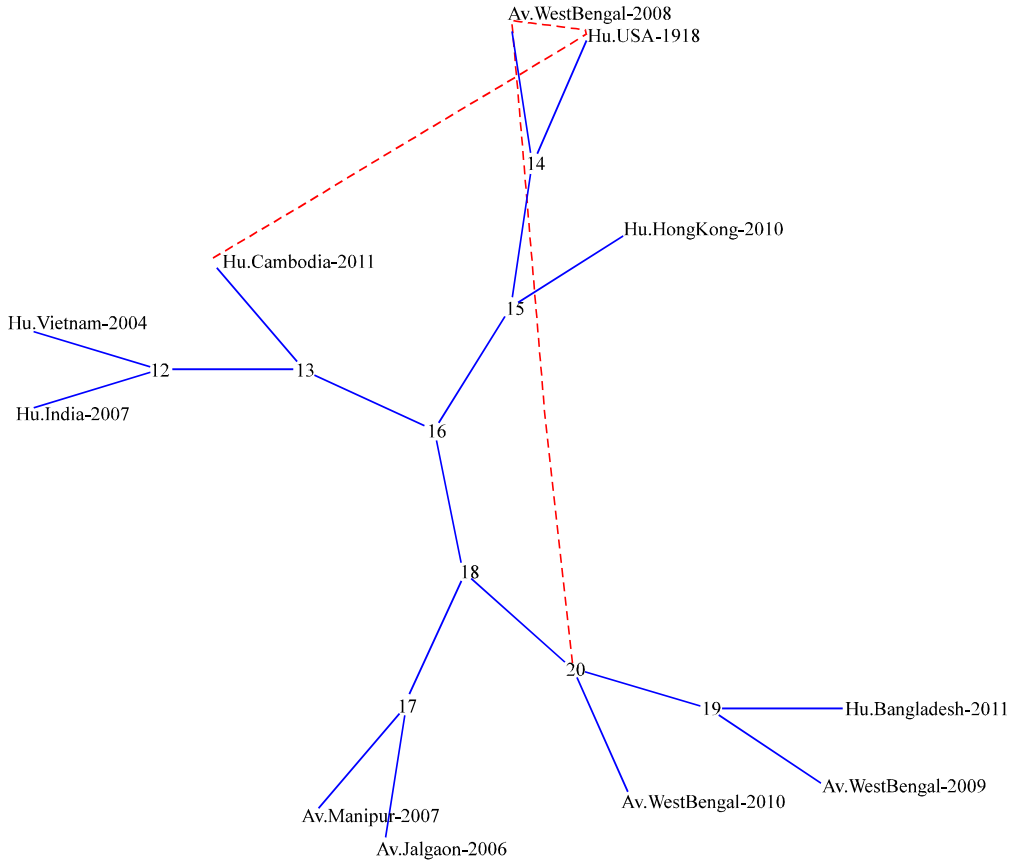

Fig. 2: Evolutionary network phylogeny among A/H5N1 influenza viruses obtained by least square criterion (LSC) and showing new branches added to the tree

This study has focused on the viral strain of HA gene observed in the human influenza H5N1 virus of 2007 in India. The future interactions or reassortment of this virus with the H1N1 virus and the other avian influenza viruses in India are predicted and are shown in Fig. 2. The preventive measures can be made before the reassortment of the virus or the new human influenza pandemic that may be very dangerous to the human race in the coming period of time.

MATERIALS AND METHODS

Materials: The DNA sequences of the Hemagglutinin (HA) gene of the various influenza A/H5N1 viruses are retrieved from the National Centre for Biotechnology Information (NCBI) Influenza Virus Resource (http://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi). The detailed information of the sequences used in this work is given in Table 1.

Analysis of data: The DNA sequence of Hemagglutinin gene of the human case of Influenza A/H5N1 virus found in India is analyzed by using the Blast tool of the NCBI site and corresponding to it some similar sequences of avian influenza A H5N1 virus, human H5N1 and human H1N1 are retrieved for the further prediction of reassortment of viruses that may cause severe influenza epidemic among the human kind (http://www.ncbi.nlm.nih.gov/nucleotide/). The sequences were

Table 1: List of hemagglutinin (HA) gene isolates of human and avian influenza A/H5N1 virus analyzed in this study

Virus	GenBank accession number	Date of isolation	Location	Base pairs
A/India/m777/2007 (H5N1)	EU311220	2007	India (U.P.)	1689
A/Bangladesh/3233/2011 (H5N1)	CY088769	9/3/2011	Bangladesh (Dhaka)	1707
A/HongKong/6841/2010 (H5N1)	HQ636461	16/11/2010	China (Hong Kong)	1586
A/Cambodia/V0203306/2011 (H5N1)	JN588805	3/2/2011	Cambodia (Phnom Penh)	1767
A/Vietnam/1203/2004 (H5N1)	EU122406	2004	Vietnam	1695
A/Brevig Mission/1/18 (H1N1)	AF116575	1918	USA	1220
A/chicken/Jalgaon/8824/2006 (H5N1)	DQ887062	2006	India (Jalgaon, Maharashtra)	1723
A/chicken/Manipur/NIV9743/2007 (H5N1)	FJ719834	19/7/2007	India	1704
A/chicken/WestBengal/82544/2008 (H5N1)	EU871810	1/1/2008	India (WestBengal)	699
A/chicken/WestBengal/155505/2009 (H5N1)	GU271998	1/1/2009	India (WestBengal)	1779
A/chicken/WestBengal/239022/2010 (H5N1)	CY061302	12/1/2010	India (Murshidabad)	1707

aligned using clustalW software of Phylip Package (Felsenstein, 1993; Thompson *et al.*, 1994) and then the distances among the sequences were obtained using K2P method of dnadist tool of Mobyle Pesteur Package (Mobyle@Pasteur v1.0). The distance matrix of these sequences is then used to generate the phylogenetic tree using Neighbor Joining method that shows the evolutionary pathways among the viruses (Saitou and Nei, 1987).

While generating the tree, the distances between the viruses are diversified from their original distances of evolution. The rate of diversification is measured by using the following mathematical equation:

$$\text{Rate of diversification} = \{\tau(i, j)\text{-}d(i, j): \tau(i, j) > d(i, j)\}$$

The error comes in existence during the construction of the tree occurs due to the diversification of viruses from their original distances and is measured and minimized by the Least Squares Criterion (LSC) denoted by Q (Legendre and Makarenkov, 2002; Makarenkov, 2001; Mathur and Adlakha, 2011):

$$Q = \sum_{i<j} \sum_{i,j \in n} (\tau(i, j) - d(i, j))^2 \rightarrow \min$$

where, n is the total No. of viruses considered in the study, $\tau(i, j)$ is the tree distance between viruses (i, j) and d(i, j) is the original distance between viruses (i, j).

Finally, tree distances obtained are utilized to construct the network for the evolutionary pathways among the viruses by using the T-rex Package given by Makarenkov (2001). The diversification occurred during the construction of the evolutionary network from the tree is minimized by the above formula so that we can reach closer to the original evolutionary distances among the viruses. Henceforth, the events like hybridization, homoplasy, recombination and speciation can also be predicted to find out the viruses that can be reassorted and can produce the new viruses which may be much harmful to the public health of living communities.

RESULTS AND DISCUSSION

The names of viruses used in Fig. 2 of the network suggest the type, name and year of isolation of the particular virus.

Prediction of reassorted viruses by evolutionary network: The A/H5N1 virus is known for its high pathogenicity among the birds, especially in poultry. So, this study is performed to predict the possibility of transmission of virus from birds to human beings. During the generation of phylogenetic tree among the viruses considered in the study, three main clusters come into existence that shows relationships with each other. Figure 2 suggests that the first cluster contains all the three viruses of the human cases of influenza virus A/ H5N1. The second cluster is composed of three viruses i.e., human virus, avian cases of influenza A/H5N1 virus and the swine origin A/H1N1 virus of 1918. The third cluster is composed of the remaining five viruses among which one is from human origin and other four are from avian origin of influenza A/H5N1 virus of India. These clusters signify the evolutionary closeness with each other in the form of distances. The clusters are represented by the phylogenetic tree shown in Fig. 2 notwithstanding the dotted red lines (Sinha *et al*., 2009).

After the construction of the phylogenetic tree, all the 55 combinations among the pairs of the viruses are analyzed for their genetic lengths. Out of 55 combinations of pairs of viruses, only 29 pairs of viruses increased their lengths from their original distances representing that they are diversifying or going away from their common ancestry. The 25 pairs have decreased their evolutionary distances from the original ones representing that they are coming closer to their common ancestry and one pair maintains the same evolutionary distance.

Table 2 shows the diversification values among top ten more diversifying pairs of viruses added to the tree to construct the network with their least squares function of phylogenetic tree or network. The value at Sr. No. 1 in Table 2 represents the LSC value of the phylogenetic tree when no new diversifying branch is considered and the values of LSC decreases continuously with the addition of new branches. The pairs of viruses numbered 3 to 6 in Table 2 are not added to the tree to construct the network as all of them belongs to the same species of the same cluster i.e., avian viruses. The viruses of the same cluster can be hybridized but generally not able to produce new

Table 2: Diversification and LSC values among some most diversified pairs of viruses

Sr. No.	Pairs of viruses	Diversification	LSC (Tree)	LSC (Network)
1	-	-	0.020964007	-
2	USA 1918-WestBengal 2008	0.123405	-	0.005735213
3	WestBengal 2008-WestBengal 2009	0.030000	-	0.004835213
4	WestBengal 2008-WestBengal 2010	0.028638	-	0.004015078
5	Jalgaon 2006-WestBengal 2008	0.023258	-	0.003474143
6	Manipur 2007-WestBengal 2008	0.015845	-	0.003223079
7	Cambodia 2011-USA 1918	0.015186	-	0.002992465
8	USA 1918-HongKong 2010	0.004270	-	0.002974232
9	Vietnam 2004-Manipur 2007	0.004150	-	0.002957009
10	India 2007-Manipur 2007	0.004120	-	0.002940035
11	India 2007-HongKong 2010	0.003672	-	0.002926551

-: When reticulation branches are added to the evolutionary tree, it becomes the evolutionary network and gets the LSC values of the network, so the values for tree remain blank, LSC (Network): Least squares criterion for evolutionary network, LSC (Tree): Least squares criterion for evolutionary tree

viruses that are evolved by the combination of the two different parents. Meanwhile, the viruses numbered 2 and 7 in Table 2 are involved in the network since they are descended from the different ancestors and able to produce some new reassorted viruses that may be harmful to the human race in the coming future.

The viruses which are diversifying can be controlled by creating the shortcut pathways of the original evolutionary distances and hence the phylogenetic network can be generated by using the T-rex Package (Makarenkov, 2001). Figure 2 represents the evolutionary network among all the viruses considered and Table 3 represents the corresponding evolutionary network distances among all the pairs of viruses. The significant results of the Table 3 are made bold and having the values 0.566001, 0.613790, 0.040057, 0.046329 and 0.039979. These values represents that the pairs of viruses corresponding to them are more diversifying than other pairs in the evolutionary tree and hence made contribution to construct the evolutionary network. This representation of the network shows that there is the possibility of mixing of viruses or reassortment among the viruses of different clusters shown by the dotted red lines in the Fig. 2.

Figure 2 showed that the first reassortment or reticulation occurs between the pair of Human A/H1N1 1918 sequence of USA and Avian A/H5N1 2008 sequence of West Bengal with the branch length of 0.566001 shown in Table 3. The second reticulation expected to occur between Human A/H1N1 USA 1918 and Human A/H5N1 Cambodia 2011 sequences with the branch length of 0.613790. Third reassortment occurs among Avian A/H5N1 2008 sequence of West Bengal and the common ancestor of the cluster of A/H5N1 sequences (Human Bangladesh 2011, Avian West Bengal 2009 and 2010) with the new reticulation branch length of 0.028651. Finally, there may be the chances of reassortment between the Human influenza virus of India and Avian influenza virus of Manipur, India. The pairs of Human influenza viruses of India 2007 and Hong Kong 2010 can also be reassorted.

Also, it is obvious to say that infection in humans with A/H5N1 virus remains rare but human cases are continued to be reported. As well, A/H5N1 is now considered endemic among poultry in parts of Asia, providing opportunities to this virus to disseminate or spread widely over the world and to mutate and adapt to humans and other mammalian species. It is also essential to collaborate between human and animal health sectors for surveillance, case investigation, virus sharing and risk assessment to understand and reduce the risk of virus transmission at the interface between domestic poultry and humans. So, it is necessary to quickly trace out the changes that may occur

Table 3: Evolutionary network distance matrix among all the pairs of influenza viruses analyzed in this study

	India-2007	Bangladesh-2011	HongKong-2010	Cambodia-2011	Vietnam-2004	USA-1918	Jalgaon-2006	Manipur-2007	WestBengal-2008	WestBengal-2009	WestBengal-2010
India-2007	0.000000										
Bangladesh-2011	0.060293	0.000000									
HongKong-2010	0.071987	0.082163	0.000000								
Cambodia-2011	0.049415	0.083789	0.095483	0.000000							
Vietnam-2004	0.002000	0.060293	0.071987	0.049415	0.000000						
USA-1918	0.605317	0.612331	0.617611	**0.613790**	0.605317	0.000000					
Jalgaon-2006	0.051583	0.038047	0.073453	0.075078	0.051583	0.606783	0.000000				
Manipur-2007	0.052950	0.039414	0.074820	0.076446	0.052950	0.608150	0.023821	0.000000			
WestBengal-2008	0.039316	**0.046329**	0.051609	0.062812	0.039316	**0.566001**	0.040782	0.042149	0.000000		
WestBengal-2009	0.053943	0.015988	0.075813	0.077439	0.053943	0.605981	0.031697	0.033064	**0.039979**	0.000000	
WestBengal-2010	0.054021	0.029084	0.075891	0.077516	0.054021	0.606058	0.031775	0.033142	**0.040057**	0.022734	0.000000

Bold values are significant

in the virus or in the epidemiology of its spread to humans that signals adaptation to humans. Current exposure data remain too general to specify the prediction of future cases of H5N1 infection in human populations. However, the results of the available studies, including those reporting cases having no contact with poultry, suggest that exposure through the environment may also account for many human cases. Rapid, systematic and standardized collection of detailed information on poultry contact and human case contacts for all suspected and confirmed human cases of H5N1 would improve our understanding of risks of A/H5N1 virus and help to inform development and implementation of appropriate public health risk reduction measures (Bourouiba *et al.*, 2011; Kerkhove *et al.*, 2011).

The reassorted network (Fig. 2) is determined among HA genes of A/H5N1 viruses. The value of LSC of the evolutionary network in this study is decreasing continuously by adding every new branch to the evolutionary tree which shows less diversity than previous one. Hence, the viruses examined in this study have undergone the intra-cluster and inter-cluster reassortment and thus represents the genetic parents of the new reassorted viruses. It is depicted that the avian-human interface could create any mixed virus similar to the recently emerged pandemic A/H5N1 viruses and hence an emphasis should be put on the prevention and control of AI outbreaks in India done by the reassortment. In addition, A/H5N1 viruses should be routinely analyzed for monitoring the genetic diversity of the virus and identifying new reassortment between avian and human H5N1 viruses (Amonsin *et al.*, 2010).

The predictions presented in this study will be helpful to the scientific community for the timely planning and early control of the reassorted virus epidemics in India.

Preventive measures to tackle the next influenza pandemic: Since 1997, highly pathogenic influenza viruses of the H5N1 subtype have been causing human infections with a high mortality rate (Octaviani *et al.*, 2010). And it is also observed by the studies that the Human influenza outbreak has the potential of triggering a pandemic either by adaptation or reassortment or both (Khanna *et al.*, 2006).

Although in past, millions of birds died of avian influenza and caused a growing loss to human lives. As a result, the precautions become must in this kind of disease which harms animals and humans at large scale. The World Health Organization (WHO) has warned of a risk of a pandemic of avian influenza in the future and hence, it is necessary and important to investigate the transmission process of avian influenza and to take effective measures to control the spread of avian influenza (Kim *et al.*, 2010).

Therefore, appropriate surveillance and preventive measures to minimize the risks of reassortment between the human and avian H5N1 viruses interface and to tackle the next influenza pandemic in India are as (Octaviani *et al.*, 2010):

- The most important measure to face the pandemic influenza virus is to use the influenza virions, vaccines and antiviral agents. Four important drugs currently available for the treatment of influenza virus infection are amantadine, rimantadine, zanamivir (Relenza) and oseltamivir (Tamiflu)
- The stockpiling of the drugs and the vaccines should be maintained to combat with any emergency situation
- To stop bird to human, human to human transmission and the spread of the virus to other areas, all the birds within 10 km radius of the affected area should be banned to migrate from one place to the other

- As the H5N1 viruses are extremely sensitive to heat, peoples should eat eggs and chicken cooked at 70°C or above for approximate 30 min
- Exposure to H5N1 contaminated environments like water sources, poultry workers, etc.

CONCLUSION

The reassorted network of A/H5N1 viruses based on the HA gene has been determined in this study. The value of LSC of the evolutionary network is decreasing continuously by adding every new branch to the evolutionary tree which is of less diversity than previous one. Hence, the viruses examined in this work have undergone the intra-cluster and inter-cluster reassortment and thus represents the genetic parents of the new reassorted viruses. Three reassorted branches (Fig. 2) among the viruses has been predicted. It is also depicted that the avian-human interface could create any mixed virus similar to the recently emerged pandemic A/H5N1 viruses and hence, an emphasis should be put on the prevention and control of AI outbreaks in India done by the reassortment. In addition, A/H5N1 viruses should be routinely analyzed for monitoring the genetic diversity of the virus and identifying new reassortment between avian and human A/H5N1 viruses. The authors hope that the predictions presented in this study will be helpful to the scientific community for the timely planning and early control of the reassorted virus epidemics in India.

REFERENCES

Amonsin, A., J. Lapkuntod, K. Suwannakarn, P. Kitikoon and S. Suradhat *et al.*, 2010. Genetic characterization of 2008 reassortant influenza A virus (H5N1) Thailand. Virol. J., Vol 7.

Bokhari, S.H. and D. Janies, 2010. Reassortment networks for investing the evolution of segmented viruses. IEEE/ACM Trans. Comput. Biol. Bioinf., 7: 288-298.

Bourouiba, L., S.A. Gourley, R. Liu and J. Wu, 2011. The interaction of migratory birds and dometic poultry and its role in sustaining avian influenza. SIAM. J. Appl. Math., 71: 487-516.

Felsenstein, J., 1993. PHYLIP (Phylogeny Inference Package) Version 3.67c. Department of Genetics, University of Washington, Seattle.

Flavia, G.B. and K. Natarajaseenivasan, 2011. Phylogenetic analysis of H1N1 sequences from pandemic infections during 2009 in India. Bioinformation, 5: 416-421.

Kerkhove, M.D.V., E. Mumford, A.W. Mounts, J. Breee, S. Ly, C.B. Bridges and J. Otte, 2011. Highly pathogenic avian influenza (H5N1): Pathways of exposure at the animal-human interface, a systematic review. PLoS ONE, 10.1371/journal.pone.0014582

Khanna, M., N. Akther, V. Srivastava, P. Kumar and V.K. Vijayan, 2006. Biological and epidemiological aspects of influenza virus H5N1 in context of India. Indian J. Exp. Biol., 44: 265-278.

Kim, K.I., Z. Lin and L. Zhang, 2010. Avian-human influenza epidemic model with diffusion. Nonlinear Anal. Real World Appl., 11: 313-322.

Legendre, P. and V. Makarenkov, 2002. Reconstruction of biogeographic and evolutionary networks using reticulograms. Syst. Biol., 51: 199-216.

Li, C., M. Hatta, C.A. Nidom, Y. Muramoto, S. Watanabe, G. Neumann and Y. Kawaoka, 2010. Reassortment between avian H5N1 and human H3N2 influenza virus creates hybrid with substantial virulence. PNAS, 107: 4687-4692.

Makarenkov, V., 2001. T-Rex: reconstructing and visualizing phylogenetic trees and reticulation networks. Bioinformatics, 17: 664-668.

Mathur, R. and N. Adlakha, 2011. A least squares method to determine reticulation in eight grass Plastomes. OJB, 12: 230-242.

Octaviani, C.P., M. Ozawa, S. Yamada, H. Goto and Y. Kawaoka, 2010. High level of genetic compatibility between swing-origin H1N1 and highly pathogenic avian H5N1 influenza viruses. J. Virol., 84: 10918-10922.

Rao, A.S., 2008. Modeling the rapid spread of avian influenza (H5N1) in India. Math. Biosci. Eng., 5: 523-537.

Saitou, N. and M. Nei, 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol., 4: 406-425.

Sinha, N.K., A. Roy, B. Das, S. Das and S. Basak, 2009. Evolutionary complexities of swine flu H1N1 gene sequences of 2009. Biochem. Biophys. Res. Commun., 390: 349-351.

Tambunan, U.S.F., A. Limanto and A.A. Parikesit, 2010. *In silico* analysis of hemagglutinin, neuraminidase and matrix2 of H5N1 virus indonesian strain related to its high pathogenicity. IIOAB J., 1: 17-24.

Thompson, J.D., D.G. Higgins and T.J. Gibson, 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res., 22: 4673-4680.

WHO, 2012. Influenza at the human-animal interface. http://www.who.int/influenza/human_animal_interface/H5N1_cumulative_table_archives/en/index.html