The Influenza-A mystery: Insight from Bioinformatics resources and analysis

Asad U Khan

Interdisciplinary Biotechnology Unit, Aligarh Muslim University, Aligarh 2020002, India; Asad U Khan – Email: asad.k@rediffmail.com; Phone: +91-9837021912

Sir, Bioinformatics is a tool which can excavate any information, one wants to obtain. I found its implication in handling an emerging and highly pathogenic strain of Influenza virus, H5N1. The first influenza pandemic was happened in 1918 "Spanish flu" followed by Asian flu in 1957 caused by H2N2. Moreover, one million deaths were reported in Hong Kong due to H3N2 outbreak during 1968. The antigenic shift and drift of influenza A virus led to a new strain H1N1 in 1977 "Russian flu". After two decades a new strain, H5N1 was evolved, called avian influenza virus. The outbreaks of this strain were continued till 2008 with the yearly epidemic in the world. In 2009, WHO has declared Influenza pandemic due to the outbreak of H1N1. The disease has infected over 35928 people in 80 countries, with 163 deaths. But it has been most severe in Mexico, which has reported the highest number of fatalities.

The influenza virus is mysterious pathogen known for severe respiratory illness. It causes yearly epidemics in tropical and subtropical countries with epidemic and even pandemic flu, a major cause of morbidity and mortality. Recurrent appearance of influenza in different parts of the world is due to frequent mutations which may lead to either antigenic drift or shift. The structure of influenza A viral genome composed of eight segments of negative-sense single stranded RNA. Surface glycoprotein of influenza A virus, Hemagglutinin (HA) and Neuraminidase (NA) are the main components of viral genome to undergo reassortment and may cause to develop a new subtype or strain of Influenza virus. There are 16 different H antigens (H1 to H16) and nine different N antigens (N1 to N9). The latest H type (H16) was isolated from black-headed gulls found in Sweden and Netherlands during 1999 and was published in 2005. The history on influenza epidemic revealed circulation of these strains (H1N1, H2N2, H3N2, H5N1, H7N7, H1N2, H9N2, H7N2, H7N3, H5N2 and H10N7) among human, pig and bird population of different regions on the globe **[1].**

The strains appeared during outbreaks in different parts of the world and with the gap of different time periods could be analyzed for their phylogeny to understand their origin or source of spread. Phylogenetic trees are built by Neighbour-joining program with Kimura 2 parameter available at phylip3.65 package **[2]**. It is drawn by using TreeView1.6 program **[3].** We can further analyze genome using GOR IV method **[4].** Different domains/motifs were searched by "Composite" available at Expasy server **[5].** Moreover, Amino acids sequences are also compared by MAP MUTATION program **[6]** to analyze the mutations among these isolates.

Homology modeling of template structure related to proteins was searched using Blast algorithms against PDB. Amino acids sequence of PB1-F2 protein of A/HK/156/97 was obtain from Swiss-Prot [ID: P0C0U0]. Homology modeling of PB1-F2 from A/HK/156/97 (H5N1) based on crystal structure of PB1-F2 protein [PDB id: 2HN8] from H1N1 subtypes with >69% identity, was built by modeler **[7].** Rigid body docking of PB1-F2 protein with VDAC1 and ANT3 was performed by HEX5.1 **[8, 9].** Each docking solution was minimized by energy minimization technique. Details of interaction derived from refined structures were analyzed using the LIGPLOT program **[10].** It is trouble-free as well as time saving to understand the insight of microbial genome to answer several hidden mysteries and therefore, the origin of viral strains during an outbreak, by exploring these tools can easily be identified. The viral genome databases, available on different sources, NCBI **[11]** and IVDB **[12]** could be used to retrieve the sequences for analysis.

References:
[1] http://www.who.int/csr/don/2009_06_10a/en/index.html
[2] http://evolution.genetics.washington.edu/phylip.html
[3] http://taxonomy.zoology.gla.ac.uk/rod/treeview.html
[4] http://npsa-pbilibcp.fr
[5] http://expasy.org/tools/scanprosite
[6] T Anwar, AU Khan, *Bioinformation* **1:**92 (2006)
[7] A Sali *et al., Proteins* **23:**318 (1995) [PMID: 8710825]
[8] http://www.csd.abdn.ac.uk/hex
[9] M Danishuddin *et al., J Mol Model.* (2009) [PMID: 19669810]
[10] AC Wallace *et al., Protein Eng.* **8:**127 (1995) [PMID: 7630882]
[11] http://www.ncbi.nlm.nih.gov
[12] http://influenza.psych.ac.cn/statics/Statics.jsp