Spatio-Temporal Character of Selection and Diversity over the H3 Influenza Hemagglutinin

Keyao Pan[1] and Michael W. Deem[2]

***Short Abstract* — The fast evolution of H3 hemagglutinin occurs in a complicated pattern. To depict the pattern of evolution, we introduce relative entropy and Shannon entropy as two independent state variables of the subtype H3N2 virus evolution. This entropy method discovers 56 positions under selection. Amino acids in Epitope A and B at the top of hemagglutinin undergo the most intense selection. The dominance of the subtype H3N2 influenza virus is a necessary condition of the high selection during the annual flu season. Both the selection and the virus diversity increase and then decrease during the years of antigenic drift.**

***Keywords* — Influenza, Evolution, Selection, Diversity, Entropy**

I. Background

Influenza A virus circulates in human population every year, causing 3-5 million severe illnesses and 250,000-500,000 fatalities all over the world [1]. The evolution of the hemagglutinin causes the mismatch between the virus and the vaccine, and decreases vaccine effectiveness or even induces original antigenic sin in some years [2].

Previous studies have identified the positions critical to the immune escape in the evolutionary history of influenza, by resolving the historical mutation in each position using d*N*/d*S* ratio [3] or amino acid substitutions. More positions have been identified as under immune selection as the number of available sequences increases [4].

The influenza viruses circulating in the current year are the offspring of the viruses from the previous year [5]. The change of distribution between these two set of sequences sampled in two consecutive years indicates the presence of selection, and helps to depict the general picture of the influenza evolution.

II. Data

The 6896 H3 hemagglutinin sequences with amino acids 1-328 were obtained from NCBI Influenza Virus Resource on 8 March 2010, and were aligned. The dominance of H3 hemagglutinin in each year is defined as the fraction of H3N2 virus in all the samples, and is from the historical surveillance data [6].

[1]Department of Bioengineering, Rice University, Houston, TX 77005. E-mail: kpan@rice.edu

[2]Department of Bioengineering and Physics & Astronomy, Rice University, Houston, TX 77005. E-mail: mwdeem@rice.edu

III. Analysis

Diversity of the virus in each position in each year is represented by the Shannon entropy. A measure of selection in each position in each year is defined by the relative entropy between the 20-bin histogram in the current year and that in the previous year, and denoted *S*. We show that Shannon entropy and relative entropy are two independent state variables of influenza evolution.

A. Selection in different regions of hemagglutinin

By calculating the relative entropy, 56 positions under selection are identified. Among five antibody binding sites, Epitope A and B display the highest level of selection and diversity. Unlike other regions, Epitope A and B were dominated by disruptive selection in most years.

B. Selection in different years of circulation

The value of *S* averaged over all positions changes continually and gradually. Local maxima of average *S* mostly emerged when H3N2 virus was dominant.

However, in one single position, the measured selection is both punctual and asynchronous. The positions inducing the transition of antigenic clusters [7] showed high disruptive selection followed by stabilizing selection in the transition years of antigenic clusters. Epitopes A and B were also under intense selection in transition years.

IV. Conclusion

The evolution of influenza is observed in a subset of residues of hemagglutinin and is driven first by immune system pressure and later by non-immune pressures.

References

[1] World Health Organization Media Centre influenza fact sheet 211, 2009. http://www.who.int/mediacentre/factsheets/fs211/en/index.html

[2] Gupta V, Earl DJ, Deem MW (2006) Quantifying influenza vaccine efficacy and antigenic distance. *Vaccine* **24**, 3881-3888.

[3] Bush RM, et al. (1999) Positive selection on the H3 hemagglutinin gene of human influenza virus A. *Mol Biol Evol*. **16**, 1457-1465.

[4] Shih AC, et al. (2007) Simultaneous amino acid substitutions at antigenic sites drive influenza A hemagglutinin evolution. *PNAS* **104**, 6283-6288.

[5] Russell CA, et al. (2008) The global circulation of seasonal influenza A (H3N2) viruses. *Science* **320**, 340-346.

[6] Morbidity and Mortality Weekly Report, Influenza activity, 1981-2008.

[7] Smith DJ, et al. (2004) Mapping the antigenic and genetic evolution of influenza virus. *Science* **305**, 371-376.