Virus Research 250 (2018) 43–50

Contents lists available at ScienceDirect

Virus Research

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

H1N1 seasonal influenza virus evolutionary rate changed over time

Ornpreya Suptawiwat[a], Alita Kongchanagul[b,*], Chompunuch Boonarkart[c], Prasert Auewarakul[c]

[a] Faculty of Medicine and Public Health, HRH Princess Chulabhorn College of Medical Science, Chulabhorn Royal Academy, 54 Kamphaeng Phet 6 Talat Bang Khen, Lak Si, Bangkok, 10210, Thailand
[b] The Institute of Molecular Biosciences, Mahidol University, 25/25 Phuttamonthon 4 Road, Salaya, Nakhon Pathom, 73170, Thailand
[c] Department of Microbiology, Faculty of Medicine Siriraj Hospital, Mahidol Universit, 2 Prannok Road. Bangkoknoi, Bangkok, 10700, Thailand

ARTICLE INFO

Keywords:
Influenza virus
H1N1
Antigenic drift
Sequence entropy
Evolution

ABSTRACT

It was previously shown that the seasonal H1N1 influenza virus antigenic drift occurred at a slower rate than the seasonal H3N2 virus during the first decade of the 21th century. It was hypothesized that the slower antigenic evolution led to a decrease in average ages of infection, which in turn resulted in lower level of global viral circulation. It is unclear what caused the difference between the two viruses, but a plausible explanation may be related to the fact that the H1N1 virus had been in human population for much longer than the H3N2 virus. This would suggest that H1N1 antigenic drift in an earlier period may have been different from a more recent period. To test this hypothesis, we analyzed seasonal H1N1 influenza sequences during various time periods. In comparison to more recent H1N1 virus, the older H1N1 virus during the first half of the 20th century showed evidences of higher nonsynnonymous/synonymous ration (dN/dS) in its hemagglutinin (HA) gene. We compared amino acid sequence changes in the HA epitopes for each outbreak season and found that there were less changes in later years. Amino acid sequence diversity in the epitopes as measured by sequence entropy became smaller for each passing decade. These suggest that there might be some limit to the antigenic drift. The longer an influenza virus has drifted in human population, the less flexibility it may become. With less flexibility to adapt and escape the host immunity, the virus may have to rely more on younger naïve population.

1. Introduction

Seasonal influenza viruses are highly variable. Antigenic drift occurs frequently and makes necessary the annual update of vaccine strains. Hemagglutinin (HA) is the major viral envelope protein and contains major antigenic determinants. Influenza HA evolves under a strong positive selective pressure from the host immune response (Fitch et al., 1991). Because of this strong immune pressure, old influenza strains neutralizable by herd immunity become regularly extinct. This strong immune pressure allows antigenically escape strains to outgrow old strains and cause antigenic drift and new outbreaks (Hay et al., 2001). This antigenic drift is a global event, in which a new variant spreads out globally (Rambaut et al., 2008). Although it is generally accepted that seasonal influenza viruses evolve under strong positive selection exerted by host immunity, the degree of the positive selection can be different. It has been previously shown that during the first decade of the 21st century the H1N1 and influenza B viruses had lower rates of antigenic escape than the H3N2 virus with (Bedford et al., 2015). While the H3N2 virus strictly followed the single lineage phylogeny with periodic emergence of new global strains and extinction of old strains, the H1N1 and influenza B viruses had lower level of global viral circulation and higher level of local persistence resulting in a transient multi-branch phylogeny (Bedford et al., 2015). It was hypothesized that the lower rates of immune escape of influenza B and H1N1 may have led to younger average ages of infection as compared to H3N2. It was also hypothesized that the lower average ages of infection may explain the reduced global viral circulation as children travel long-distances less frequently than adults (Bedford et al., 2015). However, it is unclear why influenza B and the H1N1 virus showed lower rates of antigenic escape. It is possible that this is a difference in intrinsic properties of the viruses. However, as influenza B and H1N1 were older than H3N2, it is also possible that this difference is related to the length of time that the viruses have circulated in human population. The H1N1 virus entered human population in 1918 and influenza B had circulated in human population for much longer, whereas the H3N2 virus came into human population in 1968 (Cox and Subbarao, 2000).

Influenza viruses are believed to be in an optimal balance with their natural hosts. In water fowls, influenza A viruses were shown to be in an evolutionary stasis. This is probably a result of a long-term co-evolution of the viruses and hosts (Webster et al., 1992). Accordingly, the

* Corresponding author at: The Institute of Molecular Biosciences, Mahidol University, 25/25 Phuttamonthon 4 Road, Salaya, Nakhon Pathom, 73170, Thailand.
E-mail address: alita.kon@mahidol.ac.th (A. Kongchanagul).

https://doi.org/10.1016/j.virusres.2018.03.020
Received 21 November 2017; Received in revised form 17 March 2018; Accepted 29 March 2018
Available online 30 March 2018
0168-1702/

Table 1
The values of dN/dS and LRT tests for HA1 sequences by CodeML analysis.

	dN/dS	LRT (M7–M8)	P-value
1918–1957 (72 strains)	0.436	56.21	< 0.00001
1977–2008 (71 strains)	0.271	27.88	< 0.00001
1933–1943 (22 strains)	0.485	14.09	0.00017
1945–1957 (45 strains)	0.387	46.22	< 0.00001
1977–1984 (48 strains)	0.286	21.35	< 0.00001
1985–1997 (83 strains)	0.339	63.66	< 0.00001
1998–2008 (107 strains)	0.256	31.46	< 0.00001
2009–2015 (106 strains)	0.237	8.10	0.00442

longer a virus circulates in a host species, the more it should become closer to this evolutionary stasis. This hypothesis would predict that antigenic drift of a seasonal influenza virus will decrease with time. It is, however, unclear how long it would take to see this effect. The evolutionary stasis of avian influenza A viruses in water fowls may have taken thousands of years. A century is a relatively short time period in comparison. However, if the observed difference in the antigenic drift of H1N1 and H3N2 was a direct result of the difference in time length of the viral circulation, half a century must have been sufficient to cause an observable difference. Comparing old and recent H1N1 evolution may provide some insight into this hypothesis. Since the H1N1 epidemic was interrupted between 1950s to 1977 by H2N2 and H3N2, we decided to compare the evolution of H1N1 before and after 1977.

2. Materials and methods

2.1. Sequences

We focused our analyses on the HA1 gene because it codes for the main antigenic protein of the virus that showed antigenic drift and its sequences are the most abundantly available. Four hundred and eleven full length sequences of H1N1 HA1 were retrieved from NCBI Influenza virus database of the full length sequences were selected and aligned with Bioedit program. After alignment, sequences with 100% similarity were excluded from the analyses. All available sequences from 1918 to 1987 were included, and sequences were randomly selected from the more the recent years (1988–2015) to cover all geographical regions. The sequences were arbitrarily divided into 6 groups by decade (Table 1).

2.2. Phylogenetic analyses

Phylogenetic trees based on HA1 nucleotide sequences were constructed by maximum likelihood method implemented in PAUP version 4.0. The resulted trees were further used as guide trees for estimating selection pressure in CodeML application program in Phylogenetic Analysis by Maximum Likelihood (PAML) package (Yang, 2007). We used models M7 and M8, where M7 contains 10 ω categories to describe ω amongst sites, all constrained to be < 1; M8 differs from M7 only in that it estimates ω for an extra class of sites (p10) at which ω can be > 1 (Yang, 1997). Models were compared using a likelihood ratio test and the Bayes Empirical Bayes (BEB) method was used for a posteriori estimation of individual codons under positive selection (Yang et al., 2005). Phylogenetic tree based on HA1 amino acid sequences were constructed by maximum likelihood method in MEGA program version 6.0.

2.3. Hamming distance and $P_{epitope}$

Hamming distance of HA1 epitopes was calculated by comparing consensus sequence of each year to that of the previous year. $P_{epitope}$, which has been previously shown to correlate with antigenic distance (Deem and Pan, 2009), was also calculated by comparing consensus sequences of the epitope residues of two consecutive years. For each epitope, the P-value is defined as the proportion of different amino

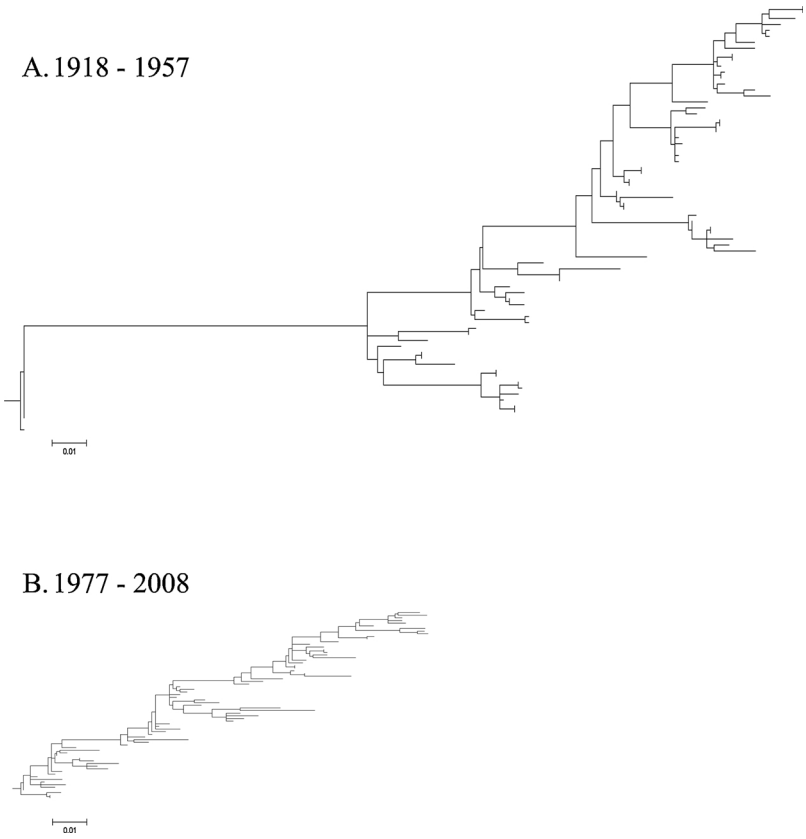

Fig. 1. Phylogenetic analysis of H1N1 HA1 nucleotide sequences. The phylogeny trees were generated by maximum likelihood from HA1 nucleotide sequences of human influenza A viruses isolated from 1918 to 1957 (a) and 1977 to 2008 (b). Branch lengths are drawn to scale. The scale bar represents 0.01 units of nucleotide change per site.