UCLA
UCLA Previously Published Works

Title

Analysis of Viral Genetics for Estimating Diffusion of Influenza A H6N1.

Permalink

https://escholarship.org/uc/item/6rc5z59z

Authors

Scotch, Matthew
Suchard, Marc A
Rabinowitz, Peter M

Publication Date

2015

Peer reviewed

Analysis of Viral Genetics for Estimating Diffusion of Influenza A H6N1

Matthew Scotch, PhD, MPH[1], Marc A. Suchard, MD, PhD[2], Peter M. Rabinowitz, MD[3]
[1]Arizona State University, Tempe, AZ, USA; [2]University of California, Los Angeles, Los Angeles, CA; [3]University of Washington, Seattle, WA

Abstract

H6N1 influenza A is an avian virus but in 2013 infected a human in Taiwan. We studied the phylogeography of avian origin H6N1 viruses in the Influenza Research Database and the Global Initiative on Sharing Avian Influenza Data EpiFlu Database in order to characterize their recent evolutionary spread. Our results suggest that the H6N1 virus that infected a human in Taiwan is derived from a diversity of avian strains of H6N1 that have circulated for at least seven years in this region. Understanding how geography impacts the evolution of avian influenza could allow disease control efforts to focus on areas that pose the greatest risk to humans. The serious human infection with a known avian influenza virus underscores the zoonotic potential of diverse avian strains of influenza, and the need for comprehensive influenza surveillance in animals and the value of public sequence databases including GISAID and the IRD.

Introduction

In June, 2013 the Taiwan CDC identified a case of H6N1 influenza A in a 20 year-old female[1]. The index case developed pneumonia, was hospitalized, yet survived. Initial phylogenetic analysis of the viral genome determined that this isolate evolved from chickens in Taiwan[2]. As of September 2014, there has been no documentation of person-to-person transmission of the virus. However, there is still a limited understanding of its phylogeography that might identify vital geographic routes of its genetic lineage. This could enable health agencies to curb future outbreaks of the avian virus and reduce the potential for human-to-human transmission.

Methodology

In order to study the spread of H6N1 avian influenza viruses, we downloaded the entire genome (eight gene segments) of H6N1 avian influenza from the Influenza Research Database (IRD)[3]. We also downloaded sequences of the human H6N1 isolate from the Global Initiative on Sharing Avian Influenza Data (GISAID) EpiFlu database[4]. We created separate FASTA files for each gene and used the strain name to extract geographic and temporal metadata for each stored sequence. For locations outside of China and Taiwan, we altered the definition line of each sequence (indicated by ">") to include the continent rather than the province. We used BEAST v 1.8[5] to perform a Bayesian discrete phylogeography reconstruction[6] of the evolutionary spread of the virus between geographic locations. We then created maximum clade credibility (MCC) trees for each gene from our posterior samples in order to construct single "best" gene trees.

For each gene, we specified a Markov chain Monte Carlo (MCMC) chain length of 30,000,000, sub-sampling every 1,000 steps. We analyzed the effective sample size (ESS) of the parameters using Tracer[7] and, if necessary, re-initiated new chains that we combined with LogCombiner[5]. We used TreeAnnotator[5] to specify an MCC with a 10% burn-in to disregard the initial steps in the MCMC. We used FigTree v. 1.4.2[8] to time-scale the MCC by years and color-code the branches by their most probable geographic state. In addition, we calculated the association index (AI) and parsimony scores (PS) using a program called BaTS to determine if the diffusion of H6N1 is geographically structured[9]. These two statistics test the hypothesis that tips in the tree are no more likely to share the same location (trait) with adjoining taxa than by chance alone[9].

Results

We included the following number of H6N1 sequences in the analysis: 223 PB2 sequences, 221 PB1, 227 PA, 303 HA, 213 NP, 316 NA, 258 M, and 349 NS sequences. In the figure, we show the phylogeographic MCC tree for each influenza gene segment. The time-scaled trees have an x-axis that indicates the years of evolution for the H6N1 virus. The posterior mean estimate of the origin of most of the genes is sometime between 1935-1943. HA (posterior mean: 1925) and PB2 (posterior mean: 1913) are a little earlier than that while NS is a hundred years earlier (posterior mean: 1841). These differences in time could be an indication of reassortment events among the gene segments[6]. We draw an arrow to indicate the human virus. For all genes, the human virus is located within the diversity of avian sequences collected in Taiwan from 1997 - 2010. Seven of the eight genes depict that early in its evolution, H6N1 was most likely to be spreading in North America. For example, other than PB1, at least one of the

two direct ancestors to the origin (*i.e.* root of the tree) are most likely from North America. In most of these genes, the virus shows evidence of local North American spread that ultimately formed a clade. Here, we define a clade as viruses that are grouped together in the tree that share the same ancestor[10]. This is distinct from the diversity of avian sequences collected in Taiwan that includes the human virus. The emergence of this diversity was a result of virus spread across different geographical areas including Europe, China, and Hong Kong. Also, we note the most recent geographic location before the diversification of Taiwanese sequences. For seven of the eight genes, the ancestor was most likely split between Hong Kong (PB2, PA, HA, NP) and Europe (PB1, NA, M). One gene, NS, was most likely descended from Korea.
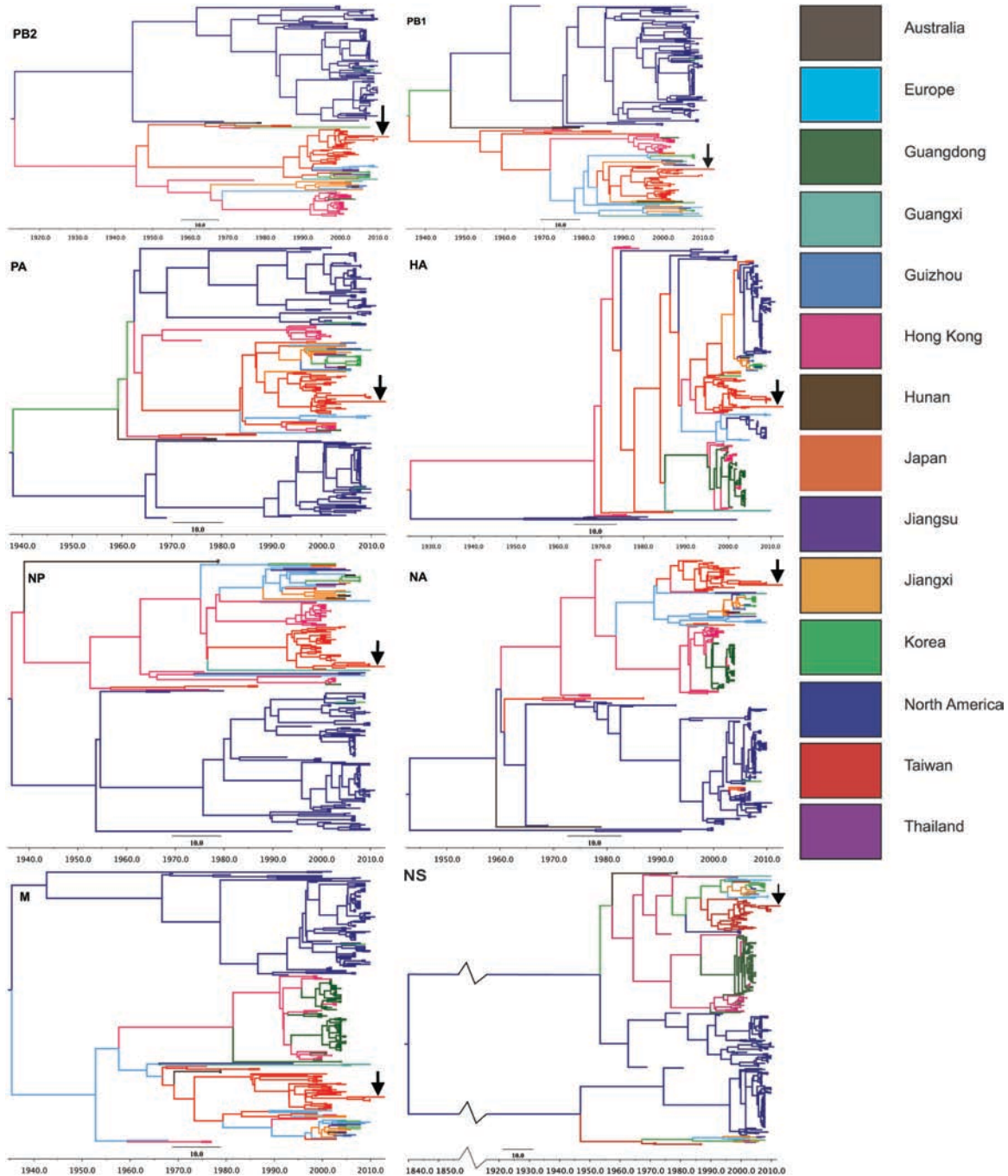

Figure 1. Phylogeographic MCC trees of eight H6N1 influenza A gene segments. Human virus indicated by arrow. Note, we have truncated the long reassortment of the NS segment before the root of the other segments for visualization purposes.

In Table 1, we report the results of the correlation between virus phylogeny and geography by calculating the AI and PS. Here, we present the estimator mean values and the 95% confidence intervals under the observed and null distribution of geographic states. For all genes, the observed values were statistically different from their null estimates, suggesting that the evolution of H6N1 is geographically structured for all gene segments.

In Table 2, we focus solely on the one branch of the tree that contains the human virus. Here, we report the most recent ancestor and year for the human H6N1 virus isolate across all eight gene segments. We determined the most recent ancestor by examining the leaf node of the human virus in the phylogeographic tree (Figure 1). We considered the location with the highest probability to be the most recent ancestor and calculated the Kullback-Leilber score (KL)[6] using the R statistical package[11] to determine the statistical support for the difference between the posterior and prior probability estimates of location. We considered the age of the leaf node to represent the posterior mean year of divergence. We found agreement across all genes for a Taiwan H6N1 strain to be the most recent ancestor with very strong statistical support as indicated by the large KL scores. The age range of the most recent divergence was 1999-2006 suggesting that the virus that infected the individual has been persisting in Taiwan for at least seven years.

Table 1. Association index and parsimony scores for all eight gene segments when considering location as a trait*#

Gene	Association index		Parsimony score	
	Observed Mean	**Null Mean**	**Observed Mean**	**Null Mean**
PB2	1.63 (1.47, 1.86)	16.13 (14.70,17.70)	21.19 (21.00, 22.00)	94.38 (90.43, 97.95)
PB1	2.22 (1.92,2.51)	14.96 (13.30, 16.50)	23.98 (20.00, 25.00)	89.52 (85.98, 92.50)
PA	2.39 (2.06, 2.71)	15.43 (14.00, 16.92	24.67 (24.00, 25.00)	92.39 (89.74, 95.31)
HA	1.48 (1.15, 1.82)	23.90 (22.50, 25.54)	28.57 (27.00, 29.00)	147.00 (143.14, 151.86)
NP	3.09 (2.85, 3.33)	15.73 (14.22, 17.23)	27.92 (27.00, 28.00)	94.39 (90.81, 97.14)
NA	2.30 (1.91, 2.71)	24.03 (22.47, 25.57)	30.52 (29.00, 32.00)	157.05 (152.64, 161.07)
M	2.22 (1.79, 2.66)	21.80 (20.37, 23.10)	29.43 (28.00, 31.00)	138.27 (132.23, 143.55)
NS	1.96 (1.72, 2.27)	27.31 (25.79, 28.66)	29.50 (29.00, 30.00)	172.51 (166.93, 177.27)

*PB2, polymerase basic 2; PB1, polymerase basic 1; PA, polymerase acidic; HA, hemagglutinin; NP, nucleoprotein; NA, neuraminidase; M, matrix; NS, non-structural.

#The p-values for all observed versus mean values are < 0.05 confirming that the diffusion of H6N1 is geographically structured.

Table 2. Location and year of the most recent ancestor of human H6N1 virus for all eight gene segments.

Gene	Most Recent Ancestor (Highest posterior probability)	Posterior Mean Year	Kullback-Leibler divergence
PB2	Taiwan (0.99)	2002	> 7
PB1	Taiwan (0.99)	1999	> 6
PA	Taiwan (0.99)	1999	> 7
HA	Taiwan (0.99)	2003	> 6
NP	Taiwan (0.99)	2005	> 6
NA	Taiwan (0.99)	2005	> 7
M	Taiwan (0.99)	2005	> 6
NS	Taiwan (0.99)	2006	> 8

Conclusion

Our results suggest that the H6N1 virus that infected a human in Taiwan is derived from a diversity of avian strains of H6N1 that have circulated for at least seven years in this region. The vast majority of viruses included in this

diversity were from chickens. In addition, in the HA and NA trees, this diversity included all of the viruses from a Taiwan chicken clade identified by Lee *et al.*[5, 12]. Our finding of Taiwan as the geographic source is consistent with other works on the origin of this virus including Shi *et al.*[13] In addition, we found that North America contributed to the early diffusion of the virus, likely among migratory North American birds, however this has resulted in the formation of a distinct localized clade. Conversely, the formation of the diversity of H6N1 in Taiwan is a result of geographic mixing in Europe and Asian with Hong Kong serving as an important geographic location in the diffusion process.

Understanding how geography impacts the evolution of avian influenza could allow disease control efforts to focus on areas that pose the greatest risk to humans. Epidemiologists can then study local factors including poultry trade and avian migration in order to identify key points for interventions. In addition, the recent cases of human infection in other avian influenza viruses such as H7N9[14] underscores the zoonotic potential of diverse avian strains of influenza, and the need for comprehensive influenza surveillance in animals and the value of public sequence databases including GISAID and the IRD.

Segment ID	Segment	Country	Collection date	Isolate name	Originating Lab	Submitting Lab	Authors
EPI459852	PB2	Taiwan	2013-May-07	A/Taiwan/2/2013	National Influenza Center, Centers for Disease Control	Taiwan CDC	Ji-Rong, Yang; Ming-Tsan, Liu; Ho-Sheng, Wu; Feng-Yee, Chang
EPI459853	PB1	Taiwan	2013-May-07	A/Taiwan/2/2013	National Influenza Center, Centers for Disease Control	Taiwan CDC	Ji-Rong, Yang; Ming-Tsan, Liu; Ho-Sheng, Wu; Feng-Yee, Chang
EPI459854	PA	Taiwan	2013-May-07	A/Taiwan/2/2013	National Influenza Center, Centers for Disease Control	Taiwan CDC	Ji-Rong, Yang; Ming-Tsan, Liu; Ho-Sheng, Wu; Feng-Yee, Chang
EPI459855	HA	Taiwan	2013-May-07	A/Taiwan/2/2013	National Influenza Center, Centers for Disease Control	Taiwan CDC	Ji-Rong, Yang; Ming-Tsan, Liu; Ho-Sheng, Wu; Feng-Yee, Chang
EPI459856	NP	Taiwan	2013-May-07	A/Taiwan/2/2013	National Influenza Center, Centers for	Taiwan CDC	Ji-Rong, Yang; Ming-Tsan, Liu; Ho-Sheng, Wu;

					Disease Control		Feng-Yee, Chang
EPI459857	NA	Taiwan	2013-May-07	A/Taiwan/2/2013	National Influenza Center, Centers for Disease Control	Taiwan CDC	Ji-Rong, Yang; Ming-Tsan, Liu; Ho-Sheng, Wu; Feng-Yee, Chang
EPI459858	M	Taiwan	2013-May-07	A/Taiwan/2/2013	National Influenza Center, Centers for Disease Control	Taiwan CDC	Ji-Rong, Yang; Ming-Tsan, Liu; Ho-Sheng, Wu; Feng-Yee, Chang
EPI459859	NS	Taiwan	2013-May-07	A/Taiwan/2/2013	National Influenza Center, Centers for Disease Control	Taiwan CDC	Ji-Rong, Yang; Ming-Tsan, Liu; Ho-Sheng, Wu; Feng-Yee, Chang

References

1. CDC Taiwan. Laboratory-confirmed case of human infection with avian influenza A(H6N1) virus in Taiwan recovered. 2013 [cited; Available from: http://www.cdc.gov.tw/english/info.aspx?treeid=bc2d4e89b154059b&nowtreeid=ee0a2987cfba3222&tid=E36A5E9AB3D3A216

2. Wei SH, Yang JR, Wu HS, et al. Human infection with avian influenza A H6N1 virus: an epidemiological analysis. The Lancet Respiratory medicine. 2013 Dec;**1**(10):771-8.

3. Squires RB, Noronha J, Hunt V, et al. Influenza research database: an integrated bioinformatics resource for influenza research and surveillance. Influenza Other Respi Viruses. 2012 Nov;**6**(6):404-16.

4. GISAID. EpiFlu. 2013 [cited 2013 Apr 6]; Available from: http://platform.gisaid.org/

5. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular biology and evolution. 2012 Aug;**29**(8):1969-73.

6. Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. PLoS computational biology. 2009 Sep;**5**(9):e1000520.

7. Rambaut A, Suchard MA, Drummond AJ. Tracer. 2013 [cited 2014 May 13]; Available from: http://tree.bio.ed.ac.uk/software/tracer/

8. Rambaut A. FigTree. 2013 [cited 2014 May 13]; Available from: http://tree.bio.ed.ac.uk/software/figtree/

9. Parker J, Rambaut A, Pybus OG. Correlating viral phenotypes with phylogeny: accounting for phylogenetic uncertainty. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2008 May;**8**(3):239-46.

10. Higgs PG, Attwood T. Bioinformatics and Molecular Evolution. Hoboken, NJ, USA: Wiley-Blackwell; 2009.

11. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.

12. Lee MS, Chang PC, Shien JH, Cheng MC, Chen CL, Shieh HK. Genetic and pathogenic characterization of H6N1 avian influenza viruses isolated in Taiwan between 1972 and 2005. Avian diseases. 2006 Dec;**50**(4):561-71.

13. Shi W, Shi Y, Wu Y, Liu D, Gao GF. Origin and molecular characterization of the human-infecting H6N1 influenza virus in Taiwan. Protein & cell. 2013 Nov;**4**(11):846-53.

14. Hu J, Zhu Y, Zhao B, et al. Limited human-to-human transmission of avian influenza A(H7N9) virus, Shanghai, China, March to April 2013. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2014;**19**(25).