BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 7(2)

A novel strategy for exploring the reassortment origins of newly emerging influenza virus

Deqiao Tian, Yumin Wang, Tao Zheng*

Beijing Institute of Biotechnology, Beijing 100071, China; Tao Zheng - Email: zt19721@hotmail.com; *Corresponding author

Received August 16, 2011; Accepted August 17, 2011; Published September 06, 2011

Abstract:
In early 2009, new swine-origin influenza A (H1N1) virus emerged in Mexico and the United States. The emerging influenza virus had made global influenza pandemic for nearly one year. To every emerging pathogen, exploring the origin sources is vital for viral control and clearance. Influenza virus is different from other virus in that it has 8 segments, making the segment reassortment a main drive in virus evolution. In exploring reassortment evolution origins of a newly emerging influenza virus, integrated comparing of the origin sources of all the segments is necessary. If some segments have high homologous with one parental strain, lower homologous with another parental strain, while other segments are reverse, can we proposed that this emerging influenza virus may re-assort from the two parental strains. Here we try to explore the multilevel reassortment evolution origins of 2009 H1N1 influenza virus using this method. By further validating the fidelity of this strategy, this method might be useful in judging the reassortment origins of newly emerging influenza virus.

Keywords: influenza virus; reassortment; origin sources

Background:
The genome of the influenza virus contains eight single stranded negative RNA segments. When two or more different influenza viruses co-infect the same host cell, newly produced influenza virus might contain the RNA from a combination of segments from all the parental strains, so the gene reassortment is a main drive in influenza evolution. In March and early April of 2009, a new swine-origin influenza A (H1N1) virus emerged in Mexico and United States. In April 2009, the Centers for Disease Control and Prevention in the United States submitted the new influenza A (H1N1) virus genome sequence (A/California/04/2009). Based on each segments homology analysis respectively, some researchers analyzed the origins of the novel influenza A H1N1 virus **[1-7]**, and drew a conclusion that this strain might be generated from a triple reassortment of human, swine and avian viruses. Respective judging each segment's origin is useful, but as a matter of fact, some inaccuracy may occur conditionally. For example, there might be a judgment that influenza X is re-assorted from Y, but in fact Y may re-assort from X, or X and Y are just mutation relationship in evolution. In exploring reassortment evolution origins of emerging influenza virus, integrated comparing of the origin sources of all segments is necessary. If some segments have high homologous with one parental strain, lower homologous with another parental strain, while other segments are reverse, we can propose that this emerging influenza virus may re-assort from the two parental strains. Here we try to explore the reassortment evolution origin of 2009 H1N1 influenza virus by this strategy.

Materials and Methodology:
Sequence data:
The accession numbers of influenza A (H1N1) virus in this study are all obtained from Genbank of NCBI (www.ncbi.nlm.nih.gov) and listed in **Table 1 (see supplementary material).** A/California/04/2009(H1N1) comes from a 10 years old boy in California San Diego of United States. A/Brevig Mission/1/1918 (H1N1) is the first pandemics influenza strain, which was recovered from an Alaskan influenza victim who was buried in 1918 **[8]**. A/Japan/305/1957(H2N2) is a second pandemic strain, which contains three segments from an avian strain (PB1, HA and NA)

[9, 10]. A/Hong Kong/1/1968(H3N2) is a third pandemic strain, which contains two segments from avian viruses (PB1 and HA) **[9, 10].** A/New Jersey/1976(H1N1) was isolated in United States from an army recruit at Fort Dix in 1976, which has been considered as a potential pandemics strain **[11].**

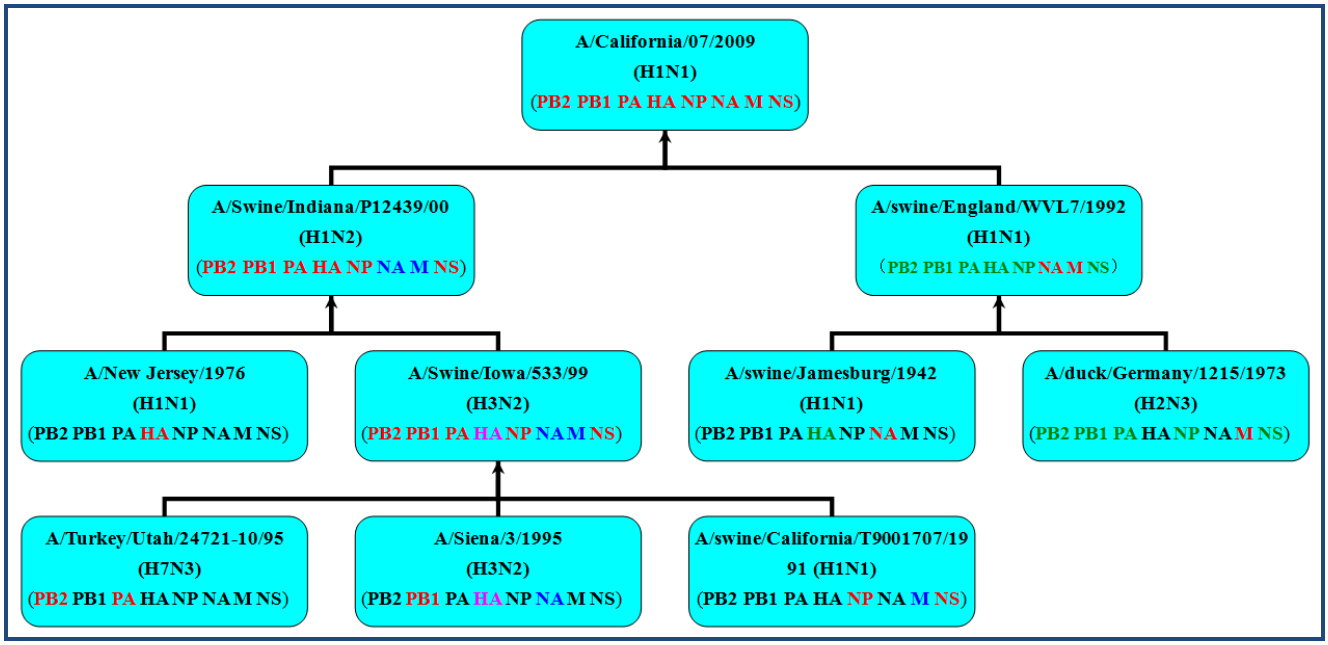

Figure 1: Origins of A/California/07/2009(H1N1) based on segment reassortment

Homologous analysis:

Sequence homologs were searched by NCBI BLAST on 2009.12.22-2009.12.23 from Database: Nucleotide collection (nr/nt), Optimize for: blastn, Max target, Sequences display: 5000. A homology search was performed for each segment of influenza virus A/California/07/2009(H1N1) using BLAST, and the results obtained were combined, compared and analyzed using EXCEL software. The re-assortment event was then judged based on the total scores of each BLAST result. During this process, the virus, which was collected before the isolate time of A/California/07/2009(H1N1), was chosen. The virus has a total of 8 segment sequences for further analysis. Using the same method, three other strains analysis was carried out. In judging the reassortment strategy, priority search of a strain that has seven segments high homology with the strain, and another strain with only one high homology segment were monitored. If an ideal result cannot be obtained, strains which have 6 high homology segments and 2 high homology segments, et al may be explored. If we still cannot get an ideal result, three strains reassortment may be hypothesized, and judging its three parental strains by the same way can be explored. In this study, DNAMAN software Version 5.2 dynamic method was used for two homology calculations.

Results:

Reassortment analysis:

The eight segments of A/California/07/2009 were analyzed using blast program, based on the sequence data in GenBank, and found that it might have come from the reassortment of a virus strain similar to A/Swine/Indiana/P12439/00(H1N2) and A/swine/England/WVL7/1992 (H1N1). Using the same method, A/Swine/Indiana/P12439/00(H1N2) was analyzed, and found that it may have come from the reassortment of virus strain similar to A/Swine/Iowa/533/99 (H3N2) and A/New Jersey/1976(H1N1), A/swine/England/WVL7/1992(H1N1) may have come from the reassortment of strain similar to A/swine/Jamesburg/1942(H1N1) and A/duck/Germany/1215/1973(H2N3). A/Swine/Iowa/533/99 (H3N2) may have come from the reassortment of strain similar to A/Turkey/Utah/24721-10/95 (H7N3), A/Siena/3/1995(H3N2) and A/swine/California/T9001707/1991(H1N1) **(Figure 1)**. The virus strains in **Figure 1** are not the actual reassortment strain. Before and after the reassortment, some sequence mutations or re-assortments might occur, since the virus sequences, which have been submitted to Genbank is only a very small part. The analysis results can only be considered as a clue for exploring the origins of a novel virus.

Homology calculation:

To further confirm the results, the homology of each segment of A/California/07/2009(H1N1) with A/Swine/Indiana/P12439/00 (H1N2) and A/swine/England/WVL7/1992(H1N1) was calculated using DNAMAN software. Using the same method, the homology of each segment of A/Swine/Indiana/P12439/00 (H1N2), A/swine/England/WVL7/1992(H1N1) and A/Swine/Iowa/533/99 (H3N2) was calculated with their parental strains. **(Table 2, see Supplementary material)**. The results further supported our hypothesis. The whole genome sequence (join to one) was also analyzed by two-pair comparison **(Table 3, see Supplementary material)**. From these results, it is shown A/Siena/3/1995(H3N2) is similar to the third pandemic influenza strain A/Hong Kong/1/1968(H3N2) with a homology of 93.63%; A/swine/Jamesburg/1942(H1N1) is similar to the first pandemic influenza strain A/Brevig

Mission/1/1918(H1N1) with a homology of 95.19%. Based on these results using this new strategy, it is concluded that this method provided more accurate information on the source origin of the influenza virus, indicating that this strategy might be useful in judging the reassortment origins of newly emerging influenza virus.

Discussion:

Three worldwide pandemic influenza outbreaks occurred in the 20th century on 1918, 1957, and 1968. They are now known to represent three different antigenic subtypes of influenza A virus: H1N1, H2N2, and H3N2, respectively **[9]**. Occasionally, an Influenza virus from one host can leap into a different host **[12]**. Exploring the rules of influenza gene re-assortment can allow us to understand the evolution of human Influenza virus and thus predict the potential future evolution of the newly emerged viruses. Based on influenza virus sequence analysis, we hypothesized the evolution origins of the novel influenza virus. We also postulated that the first pandemic influenza strain changed to a strain similar to A/swine/Jamesburg/1942(H1N1), and then re-assorted with avain influenza virus strain similar to A/duck/Germany/1215/1973(H2N3). A new H1N1 strain similar to A/swine/England/WVL7/1992 (H1N1) emerged thereafter. The third pandemic influenza strain similar to A/Siena/3/1995 (H3N2) re-assorted with swine H1N1 strain and avian influenza strain, then a H3N2 strain similar to A/Swine/Iowa/533/99 (H3N2) emerged, and it re-assorted with A/New Jersey/1976 (H1N1) similar strain, which is a potential pandemic strain in 1976, and then a H1N2 strain similar to A/Swine/Indiana/P12439/00(H1N2) emerged. A/Swine/Indiana/P12439/00 similar strain may re-assort with A/swine/England/WVL7/1992 similar strain, leading to the emergence of 2009 novel H1N1 influenza. Nowadays, with the increasing contact of human-avian, human-swine, swine-avian and with the bulk production of influenza vaccine, the gene reassortment possibility increases. In recent years, the genomic sequence data for thousands of available influenza A virus strains have been submitted to Genbank. These sequences have been isolated all around the world during the past years. Using the information in these databases, we may better understand the evolution of the viruses. Thus, the Origins of 2009 Influenza A (H1N1) virus were analysed based on total segments homology analysis, indicating that this strategy may be useful in judging the origin of a newly emerging influenza virus.

Acknowledgements:

We thank Hui Zhong, Junjie Yue, Tao He, Li Wang from Beijing Institute of Biotechnology for reviewing the manuscript and for many helpful discussions.

References:

[1] Garten RJ *et al*. *Science* 2009 **325**: 197 [PMID: 19465683]
[2] Fatimah S *et al*. *Engl J Med*. 2009 **360**: 2605 [PMID: 19423869]
[3] Kingsford C *et al*. *Plos one*. 2009 **4**: e6402 [PMID: 19636415]
[4] Nava GM *et al*. *Euro Surveill*. 2009 **14**: 19228 [PMID: 19497253]
[5] Smith GJ *et al*. *Nature* 2009 **459**: 1122 [PMID: 19516283]
[6] Trifonov V *et al*. *N Engl J Med*. 2009 **361**: 115 [PMID: 19474418]
[7] Gibbs AJ *et al*. *Virology J*. 2009 **6**: 207 [PMID: 19930669]
[8] Tumpey TM *et al*. *Science* 2005 **310**: 77 [PMID: 16210530]
[9] Kilbourne ED. *Emerg Infect Dis*. 2006 **12**: 9 [PMID: 16494710]
[10] Suzuki T *et al*. *J Virol*. 2005 **79**: 11705 [PMID: 16140748]
[11] Gaydos JC *et al*. *Emerg Infect Dis*. 2006 **12**: 23 [PMID: 16494712]
[12] Rabadan R & Robins H. *Evol Bioinform Online*. 2007 **3**: 299 [PMID: 19430605]
[13] Karasin AI *et al*. *J Clin Microbiol*. 2002 **40**: 1073 [PMID: 11880444]
[14] Karasin AI *et al*. *Virus Res*. 2000 **68**: 71 [PMID: 10930664]
[15] Dunham EJ *et al*. *J Virol*. 2009 **83**: 5485 [PMID: 19297491]
[16] Obenauer JC *et al*. *Science* 2006 **311**: 1576 [PMID: 16439620]

Edited by P Kangueane
Citation: Tian *et al*. Bioinformation 7(2): 64-68 (2011)

Supplementary material:

Table 1: GenBank number of some influence virus

Influence virus	GenBank Number							
	PB2	**PB1**	**PA**	**HA**	**NP**	**NA**	**M**	**NS**
A/California/04/2009(H1N1) **[1,2]**	FJ966079.1	FJ966080.1	FJ966081.1	FJ966082.1	FJ966083.1	FJ966084.1	FJ966085.1	FJ966086.1
A/Swine/Indiana/P12439/00 (H1N2)**[13]**	AF455736.1	AF455728.1	AF455720.1	AF455680.1	AF455704.1	AF455696.1	AF455688.1	AF455712.1
A/Swine/Iowa/533/99 (H3N2)**[14]**	AF251418.1	AF251413.1	AF251417.1	AF251411.1	AF251415.2	AF251412.3	AF251414.1	AF251416.1
A/Siena/3/1995(H3N2)	CY038510.1	CY038509.1	CY038508.1	CY038503.1	CY038506.1	CY038505.1	CY038504.1	CY038507.1
A/Turkey/Utah/24721-10/95 (H7N3)	EU980473.1	EU980472.1	EU980471.1	EF470585.1	EU980470.1	EU980469.1	AF073201.1	AF074284.1
A/swine/England/WVL7/1992(H1N1)**[15]**	CY038004.1	CY038005.1	CY038006.1	CY038007.1	CY038008.1	CY038009.1	CY038010.1	CY038011.1
A/swine/California/T9001707/1991 (H1N1)	CY028787.1	CY028786.1	CY028785.1	CY028780.1	CY028783.1	CY028782.1	CY028781.1	CY028784.1
A/New Jersey/1976(H1N1)**[11]**	CY021964.1	CY021963.1	CY021962.1	CY021957.1	CY021960.1	CY021959.1	CY021958.1	CY021961.1
A/duck/Germany/1215/1973(H2N3)**[16]**	CY014716.1	CY014715.1	CY014714.1	CY014710.1	CY014712.1	AY207522.1	CY014711.1	CY014713.1
A/swine/Jamesburg/1942(H1N1)	CY026434.1	CY026433.1	CY026432.1	CY026427.1	CY026430.1	CY026429.1	CY026428.1	CY026431.1
A/Brevig Mission/1/1918(H1N1) **[8]**	DQ208309.1	DQ208310.1	DQ208311.1	AF116575.1	AY744935.1	AF250356.2	AY130766.1	AF333238.1
A/Japan/305/1957(H2N2) **[10]**	CY045811.1	CY045810.1	CY045809.1	CY045804.1	CY045807.1	CY045806.1	CY045805.1	CY045808.1
A/Hong Kong/1/1968(H3N2) **[10]**	CY044268.1	CY044267.1	CY044266.1	CY044261.1	CY044264.1	CY044263.1	CY044262.1	CY044265.1

Table 2: Homology calculation of all the segment sequences of some influence virus

Influence virus(strain)	homology of all the segment sequences								Influence virus(strain)
	PB2	PB1	PA	HA	NP	NA	M	NS	
A/California/07/2009(H1N1)	96.49%	96.17%	95.82%	95.30%	96.66%	51.98%	87.98%	95.11%	A/Swine/Indiana/P12439/00 (H1N2)
	83.99%	86.24%	86.75%	74.38%	84.00%	94.16%	919%	81.83%	A/swine/England/WVL7/1992(H1N1)
A/Swine/Indiana/P12439/00 (H1N2)	98.77%	98.77%	98.33%	52.44%	99.20%	98.79%	97.76%	97.85%	A/Swine/Iowa/533/99 (H3N2)
	86.75%	81.05%	84.33%	91.30%	86.84%	51.11%	90.91%	85.20%	A/New Jersey/1976(H1N1)
A/swine/England/WVL7/1992(H1N1)	84.82%	83.11%	84.15%	77.50%	86.55%	85.04%	90.94%	87.48%	A/swine/Jamesburg/1942(H1N1)
	94.56%	93.54%	89.03%	65.80%	94.44%	52.78%	94.62%	93.74%	A/duck/Germany/1215/1973(H2N3)
A/Swine/Iowa/533/99 (H3N2)	84.36%	98.81%	83.08%	97.65%	83.80%	98.01%	89.58%	82.94%	A/Siena/3/1995(H3N2)
	95.81%	88.35%	95.91%	54.58%	83.57%	54.94%	91.58%	86.52%	A/Turkey/Utah/24721-10/95 (H7N3)
	83.90%	80.18%	81.45%	53.40%	97.78%	52.34%	98.88%	97.28%	A/swine/California/T9001707/1991(H1N1)

Table 3: Whole genome sequence compare

Strain		1	2	3	4	5	6	7	8	9	10	11	12	13
1	A/California/04/2009(H1N1)	100%												
2	A/Swine/Indiana/P12439/00 (H1N2)	90.82%	100%											
3	A/Swine/Iowa/533/99 (H3N2)	85.29%	92.58%	100%										
4	A/Siena/3/1995 (H3N2)	78.32%	83.86%	90.02%	100%									
5	A/Turkey/Utah/24721-10/95 (H7N3)	80.94%	81.57%	82.59%	78.64%	100%								
6	A/swine/England/WVL7/1992 (H1N1)	85.42%	80.69%	78.05%	76.86%	79.95%	100%							
7	A/swine/California/T9001707/1991(H1N1)	85.86%	84.21%	79.15%	76.29%	76.37%	81.79%	100%						
8	A/New Jersey/1976 (H1N1)	84.57%	82.27%	77.41%	79.53%	78.36%	83.56%	90.19%	100%					
9	A/duck/Germany/1215/1973 (H2N3)	80.84%	81.23%	80.29%	78.23%	85.38%	85.36%	77.95%	79.77%	100%				
10	A/Hong Kong/1/1968 (H3N2)	77.93%	82.65%	87.68%	93.63%	79.78%	78.12%	77.00%	81.37%	79.09%	100%			
11	A/Japan/305/1957(H2N2)	79.38%	83.90%	82.64%	87.61%	79.90%	80.05%	79.02%	83.99%	84.78%	91.00%	100%		
12	A/swine/Jamesburg/1942(H1N1)	85.09%	82.43%	78.50%	78.50%	79.04%	84.31%	89.04%	91.31%	80.38%	79.72%	82.14%	100%	
13	A/Brevig Mission/1/1918 (H1N1)	85.14%	82.16%	79.65%	80.98%	80.98%	85.59%	87.66%	92.59%	81.52%	82.20%	84.35%	95.19%	100%