

H5N1 Clade 2.2 Polymorphism Tracing Identifies Influenza Recombination and Potential Vaccine Targets

Niman HL^{*}, Saad MD[#], Boynton BR[#], Tjaden, J[#], Earhart KC[#], Mansour MM[#],
Elsayed NM⁺, Nayel AI⁺, Abdelghani AS⁺, Essmat HM⁺, Labib EM[#], Ayoub EA[#],
Aly MM[%], Arafa A-SA[%], Monteville, MR[#]

^{*}Recombinomics, Inc., Pittsburgh, Pennsylvania, USA, [#]U.S. Naval Medical Research Unit 3 (NAMRU-3), Cairo, EGY, [%]Central Laboratory for Veterinary Quality Control, Giza, EGY, ⁺Ministry of Health, Arabic Republic of Egypt, Cairo, EGY.

Department of Influenza Recombination, Recombinomics, Inc, 648 Field Club Road, Pittsburgh, PA 15238, USA

Correspondence to Henry L Niman, E-mail: henry_niman@recombinomics.com

617.877.0987, 412.963.1362 FAX

Text word count: 1666

2Figures

Running Title: H5N1 Polymorphism Tracing

Submitted to Proceedings of Options for the Control of Influenza VI

June 17-23, 2007 Toronto, Ontario, Canada

Highly pathogenic Influenza A H5N1 was first identified in Guangdong Province in 1996, followed by human cases in Hong Kong in 1997 ¹. The number of confirmed human cases now exceeds 300 and the associated Case Fatality Rate exceeds 60% ². The genetic diversity of the serotype continues to increase. Four distinct clades or sub-clades have been linked to human cases ^{3,4}. The gradual genetic changes identified in the sub-clades have been attributed to copy errors by viral encoded polymerases that lack an editing function, thereby resulting in antigenic drift ⁵. We traced polymorphism acquisition in Clade 2.2 sequences. We report here the concurrent acquisition of the same polymorphism by multiple, genetically distinct, Clade 2.2 sub-clades in Egypt, Russia and Ghana. These changes are not easily explained by the current theory of “random mutation” through copy error, and are more easily explained by recombination with a common source. This conclusion is supported by additional polymorphisms shared by Clade 2.2 isolates in Egypt, Nigeria and Germany including aggregation of regional polymorphisms from each of these areas into a single Nigerian human hemagglutinin gene.

The study of influenza evolution in nature has been aided by the emergence of a new strain (Clade 2.2) first identified at Qinghai Lake in central China in the spring of 2005. Sequencing of all eight genes ^{6,7} showed that isolates from migratory waterfowl were easily distinguishable from previous isolates linked to poultry and human infections in eastern and southeastern Asia. The new strain was subsequently found in outbreaks in Russia, Kazakhstan and Mongolia ⁸. Prior to these Clade 2.2 outbreaks the highly pathogenic Asian version of H5N1

had never been reported west of China. The H5N1 detection in migratory waterfowl in summer 2005 in Russian and Mongolian migratory bird sanctuaries signaled the start of a major geographical expansion of H5N1. In the following 12 months, almost 50 countries west of China reported H5N1 for the first time. All infections were Clade 2.2. The expanded geographical reach included Europe, the Middle East and Africa.

This expansion offered a unique opportunity to study the evolution of H5N1 as it migrated into new regions, including human cases in Turkey, Iraq, Azerbaijan, Egypt, Djibouti in 2006 and Nigeria in 2007. Sequence analysis indicated all cases were Clade 2.2. The outbreaks were due to multiple introductions and isolates had regional specific polymorphisms.

We isolated H5N1 from patients and poultry in Egypt. The first poultry isolates were collected February 2006. The first human case developed symptoms in March 2006. Analysis of the H5N1 isolates collected between February and May 2006 defined a series of HA and NA regional markers. These markers were also found in the human case from Djibouti and poultry isolates from Israel and Gaza.

After a lull in reported infections over the summer, H5N1 re-emerged in Egypt in September 2006. The more recent isolates had the same regional markers seen the previous season. However, the more recent poultry and human isolates had acquired a series of new polymorphisms. Non-synonymous polymorphisms were identified in samples collected from a cluster of three family members from the

Gharbiya governorate (Nile Delta). HA gene polymorphisms were identified in or near the receptor binding domain, including V223I and M230I, as well as the oseltamivir resistance polymorphism, N294S (NA gene). The patients first developed symptoms in December 2006 and all three infections were fatal (a detailed report on patients and polymorphism tracing will be described elsewhere).

Additional cases in early 2007 included HA sequences with a 3 BP deletion of the nucleotides encoding Ser at position 133 (H3 Numbering), and a case with a novel HA cleavage site, REGRRRKR. The changes were found in multiple patients in central and southern Egypt. The above non-synonymous changes were associated with additional synonymous and non-synonymous changes in the HA and NA sequences. However, these isolates from the 2006/2007 season maintained the regional markers seen in early 2006 in isolates from Egypt, Djibouti, Israel and Gaza.

Chicken isolates from Gharbiya samples collected on February 15, 2007 included one sequence that was closely related to the sequences from the human Gharbiya cluster. The sequence, A/chicken/1892N3-HK49/2007, had the regional markers previously seen in the 2006 and 2007 isolates, as well as HA non-synonymous changes, V223I and M230I. Additionally, an NA synonymous polymorphism, G743A, was appended onto the genetic background of the human Gharbiya cluster, as seen in an NA cladogram (data not shown). This polymorphism was found in two additional chicken isolates, A/chicken/1890N3-

HK45/2007 and A/chicken/1891N3-CLEVB/2007, collected the same day in the Gharbiya governorate, but these two isolates fell onto a separate branch of the tree.

The G743A was subsequently found in human isolates from patients who developed symptoms in April 2007. Included were siblings with HA sequences that had the 3 BP deletion seen in earlier patients from central Egypt. Like the chicken sequences above, the G743A polymorphism was appended onto sequences identified earlier in Egypt. Similarly, distinct sequences from another patient, A/Egypt/2630-NAMRU3/2007, that acquired G743A, also fell onto a separate branch

The distinct branches displayed in the NA cladogram were also seen in the HA cladogram (data not shown). The isolates with G743A are also located at the tips of the branches, supporting a recent polymorphism acquisition.

In February 2007, an H5N1 Clade 2.2 outbreak occurred near Moscow, Russia. Isolates from infected chickens were most closely related to 2006 sequences from Azerbaijan. Figure 1 is an expanded NA cladogram with isolates from Europe, the Middle East and sub-Saharan Africa.



Figure 1 NA phylogram of Clade 2.2 isolates.

NA phylogram of positions 43-1337. Isolates with G743A marked either with red arrows or red bars.

Like the acquisitions in Egypt, the isolates with G743A mapped onto the tip of a branch composed of earlier isolates that did not have the acquisition.

Similarly, in April 2007, an H5N1 clade 2.2 outbreak occurred near Tema, Ghana. Sequences from three chickens were most closely related to turkey

isolates collected in December 2006 in the Ivory Coast. Like the Egypt and Moscow isolates above, the recent isolates with G743A mapped to the tip of a branch containing earlier isolates that did not have the polymorphism.

Additional HA polymorphisms are noted in the HA phylogram (data not shown). Isolates that had the NA polymorphism, G743A, also had a synonymous HA polymorphism, C689T. This polymorphism was also in human and bird isolates from the Nile Delta. Another polymorphism, G754A, that encodes M230I, is in one of the German isolates⁹, A/eagle owl/Germany/R166/2006, and maps to another branch with Egyptian human and poultry isolates from the Nile Delta. A third polymorphism, C1614T, that encodes T517I, is in another German isolate, A/mute swan/Germany/R797/2006, and another branch with human isolates from southern Egypt. The isolates also have the novel HA cleavage site initially found in whooper swan isolates in Mongolia (2005). The polymorphisms found in German isolates in 2006 were in Russian Clade 2.2 isolates in 2005.

The NA G743A polymorphism can be traced through public H5N1 sequences¹⁰. The polymorphism was identified in the first reported sequences linked to the spread of H5N1 in Asia in 2003/2004 in South Korea and Japan. The polymorphism was subsequently identified in Clade 1 isolates in southeast Asia, as well as Clade 2.1 isolates in Indonesia and Clade 2.3 isolates in China. The first reported Clade 2.2 isolates were in wild birds in Germany collected in February 2006. The isolates in Germany formed distinct HA and NA branches due to a series of regional markers.

The first human isolate from Nigeria had several markers that were regionally distributed in the 2006 Clade 2.2 isolates. These shared HA polymorphisms are labeled in the HA cladogram (Figure 2).

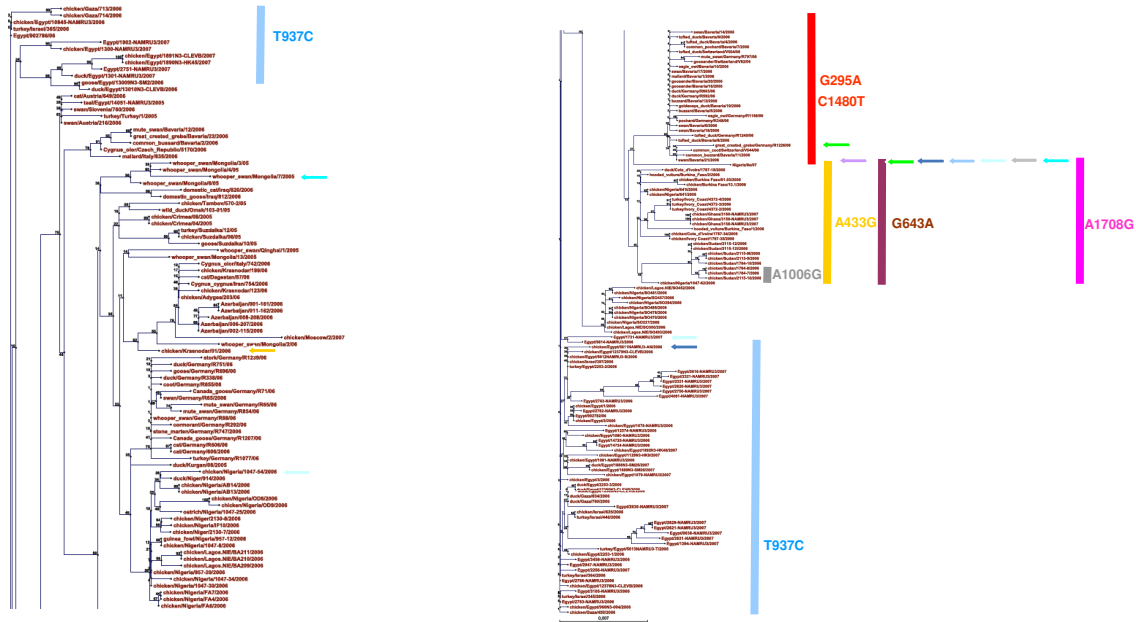


Figure 2 HA Phylogram of H5N1 isolates showing polymorphisms shared with Nigerian isolate.

HA phylogram of positions 93-1688. A/Nigeria/6e/2007 shared polymorphisms annotated with colored arrows and bars.

One of the Nigerian human polymorphisms, T937C is an Egyptian regional marker. A/Nigeria/6e/2007 also has T610C matching a small subset of the Egyptian isolates. G881A is in another small Egyptian subset, and C980T is in

isolates from Egypt and Nigeria. Similarly, the Nigerian isolate has two regional markers from Germany, G295A and C1480T, and another marker, A778G, found in one of the German bird isolates, A/great crested grebe/Germany/R1226/2006. The pattern continues in this same Nigerian isolate with the inclusion of sub-Saharan African regional markers, A433G and G643A, as well as another marker from a subset of the Sudan isolates, A1006G. Finally, G1658A from A/whooper swan/Mongolia/7/2006 appears in A/Nigeria/6e/2007. Hence the demonstration shows that the Nigerian isolate has recently aggregated H5N1 Clade 2.2 sequence information, including twelve Clade 2.2 SNPs, from a minimum of three geographically distinct areas. The cobbling together of this series of 2006 regional markers from multiple, geographically distinct locations into a single sequence is most easily explained by recombination.

Similarly, the concurrent acquisition of the same polymorphism by multiple Clade 2.2 sub-clades challenges the current theory of influenza evolution that invokes random mutations as a mechanism for the generation of antigenic drift. The isolates with the newly acquired polymorphisms map to the tips of the branches of the phylogenetic trees, indicating recent acquisition. All referenced isolates on the tips of the branches were collected over a short time frame between February and April 2007 and in three geographically distinct regions. These data do not support a common progenitor sequence because the most closely related sequences to each of the respective recent isolates do not have this change. Similarly, concurrent mutation / selection by eleven isolates that map to

six branches in three countries and were collected over a short time frame is also unlikely.

An alternative explanation for this concurrent acquisition of G743A in multiple Clade 2.2 sub-clades and for the Nigerian isolate aggregating previously dispersed regional markers is homologous recombination between closely related sequences. Polymorphism tracing demonstrates that most of the newly acquired polymorphisms can be traced to the same serotype identified recently at locations that are linked together by migratory bird flyways, raising the possibility that the distribution and acquisition of the polymorphism is linked to recombination between H5N1 sequences transmitted and transported by migratory birds.

The individual polymorphisms recombine to generate sequences that create antigenic drift. Mapping of these pathways and associations can be used to develop novel vaccine targets representing rapidly evolving genomes.

References

1. Subbarao, K. *et al.* Characterization of an Avian Influenza A (H5N1) Virus Isolated from a Child with a Fatal Respiratory Illness *Science* **279**, 393-396 (1998).
2. Cumulative number of confirmed cases of human influenza A/(H5N1)

reported to WHO. Available from

http://www.who.int/csr/disease/avian_influenza/country/cases_table_2007_05_16/en/index.html

3. Peiris, J.S. *et al.* Re-emergence of fatal human influenza A subtype H5N1 disease. *Lancet* **363**, 617-619 (2004).
4. Fouchier, R., Kuiken, T., Rimmelzwaan, G., & Osterhaus, A. Global task force for influenza. *Nature* **435**, 419-420 (2005).
5. Webster, R.G., Bean, W.J., Gorman, O.T., Chambers, T.M. & Kawaoka, Y. Evolution and ecology of influenza A viruses *Microbiol. Rev.* **56**, 152 - 179 (1992).
6. Chen, H. *et al.* Avian flu: H5N1 virus outbreak in migratory waterfowl. *Nature* **436**, 191-192 (2005).
7. Liu, J. *et al.* Highly Pathogenic H5N1 Influenza Virus Infection in Migratory Birds *Science*, **309**, 1266-1267 (2005).
8. Aleksandr, S. *et al.* Influenza (H5N1) Viruses in Poultry, Russian Federation, 2005–2006. *Emerg. Infect. Dis.* **13**, April (2007). Available from <http://www.cdc.gov/eid/content/13/4/539.htm>

9. Weber, S. *et al.* Molecular analysis of highly pathogenic avian influenza virus of subtype H5N1 isolated from wild birds and mammals in northern Germany *J. Gen. Virol.* **88**, 554-558 (2007).

10. Niman, H.L. & and Boynton, B.R. Concurrent Acquisition of a Single Nucleotide Polymorphism in Diverse Influenza H5N1 Clade 2.2 Sub-clades. Available from Nature Precedings
<http://precedings.nature.com/documents/459/version/2> (2007).