Prediction of specific virus outbreaks made from the increased concentration of a new class of virus genomic peptides, replikins.

Samuel Bogoch* and Elenore S. Bogoch**

Abstract

Advance warning of pathogen outbreaks has not been possible heretofore. A new class of genomic peptides associated with rapid replication was discovered and named replikins. Software was designed to analyze replikins quantitatively. Replikin concentration changes were measured annually prior to, and “real time” every few days during, the 2009 H1N1 influenza pandemic. Replikins were seen by both linear sequence representation and three-dimensional X-ray diffraction, and found to expand on the virus hemagglutinin surface prior to and during the H1N1 pandemic.

A highly significant increased concentration of virus replikins was found a) retrospectively in three pandemics from 1918 to 1999 (14,227 sequences)(p<0.001), and b) prospectively before the H1N1 2009 pandemic (12,806 sequences) (in the hemagglutinin gene (N=8,046), p values by t-test = 1/10^{130}, by linear regression = 1/10^{24} and 1/10^{29}, by Spearman correlation < 2/10^{16}, by Wilcoxon rank sum<1/10^{16}, by multiple regression adjusting for correlation between consecutive years = 2/10^{22}. Rising replikin concentration in H1N1 from 2006 to 2008, predicted one year in advance the H1N1 outbreak of 2009; and in H5N1, predicted the lethal outbreaks of H5N1 1997-2010.

The possible combination of influenza strains H1N1 (high infectivity) and H5N1 (high lethality) is a matter of global concern (1,2). The risk of a combined H1N1 (high infectivity) - H5N1 (high lethality) outbreak may have increased because first, the Replikin Counts of the two virus strains have risen simultaneously, not seen previously; second, the rise is to the highest levels recorded since 1918 for H1N1, in Mexico (16.7), and since 1957 for H5N1, in Egypt (23.3); and third, clinical outbreaks of each strain are occurring in 2011. These simultaneous conditions may increase the risk that the two virus strains might come into contact with each other more frequently, facilitating transfer of genomic material to form a hybrid.

*Senior Scholar, Boston University School of Medicine
**Director, Foundation for Research on the Nervous System (FRNS), 36 The Fenway, Boston MA 02215

Introduction

No structures of infectious organisms have been described to date which correlate quantitatively and temporally with epidemic outbreaks, course, and lethality, and permit early or advance warning of such outbreaks. Replikins are genomic structures related to rapid replication defined by the authors’ algorithm: peptides 7 to 50 amino acids long, containing two or more lysines, six to ten amino acids apart, at least one histidine, and a lysine concentration of 6% or more (8).
Replikins are the first reported conserved virus structures whose increasing concentration correlates quantitatively with, and predict, strain-specific virus outbreaks.

As observed in the 2009 H1N1 pandemic, advanced Replikins warning was published one year in advance, in April 2008 (12). The outbreak occurred in April 2009. Because of the time taken to produce vaccine, vaccine was not available when the brunt of the pandemic struck in April 2009. Only 20% of the world’s population at risk had vaccine available, and that was eight months after the outbreak, when the virus Replikin Count had already indicated that the lethal aspects of the pandemic would soon be over except for a brief recurrence in December 2010, also predicted (Figure 1a). Fortunately, so far, the H1N1 pandemic has been less severe than the three pandemics in the last century. The 2011 outbreaks have begun for both H1N1 and H5N1. Initial ‘scout’ virus outbreaks of H1N1 have occurred, again in Mexico, with Replikin Counts of the Infectivity Gene up to a record 16.7 (see below) and a human mortality rate of 10.7% (2); and outbreaks of H5N1 in Egypt, better established, have begun with a current cumulative mortality rate of from 34.7% (3,4) to 37.8% (5). It is generally agreed that new approaches are required for the control of acute emergent diseases (6.7). Five year plans with the hope to have a vaccine available in five years may not be relevant to the current threat (24).

The benefits of having more time to prepare for and to respond to acute lethal environmental events has been demonstrated in satellite warnings for hurricanes. The consequences of having little or no advance warning have been recently demonstrated in earthquake-tsunamis and in the 2009 H1N1 influenza pandemic. To date there have been no reliable technologies to predict emergence of specific virus strains. The only global surveillance available is post outbreak and based solely on epidemiological data (1). Acute emergent infectious diseases coupled with current global travel pose challenges to timely implementation of public health measures such as tracking and isolation of cases, and the design, testing and distribution of specific effective vaccines and therapeutics to the world’s population.

Methods

Software based on the authors’ algorithm (9) first identified and then counted the replikin peptides in each genomic sequence (Replikin Count = number of replikins per 100 amino acids). For each group of specimens’ replikins, the mean and standard deviation of the mean (SD) were calculated and compared over the past 93 years. Highly statistically significant increases and decreases were examined, for example by strain, host, country, history, year, month or week; by substitution, morbidity, and lethality. The terms ‘increase’ and ‘decrease’ of Replikin Counts were used only when the p level was less than 0.001. Counts in H1N1, H5N1 and other influenza strains were each monitored separately, retrospectively from 1918 to 1999, and prospectively from 2000-2011 for all countries reporting to Pubmed. During outbreaks, Replikin Counts were compared to Counts for the same strain in non-outbreak (‘resting’) time periods. Statistical analyses of rate of change, trend, pattern, and growth models in the evolution of each virus strain were initiated. Replikin genes were isolated in silico by scanning and identifying those areas of the virus genome which had the highest concentration of replikins. When the eight H5N1 genomic areas were examined year by year, gene areas were found which became upregulated when associated with particular outbreaks. When the upregulation was found to be associated with high infectivity (morbidity over time period), the high Count area was named Replikin Infectivity Gene. When a high Count area in a sequence was found to be associated with high mortality rates, the high Count area was named Replikin Lethality Gene. The Replikin Count of these two genes in the H1N1 virus were determined annually from 2001 to 2008 before the pandemic, then during the pandemic ‘real-time’ every
few days, then weekly, from April 2009 to February 2011. When Replikin Counts exceeded 5 per 100 amino acids, stacking was sought and found. Counts were expressed in two ways: 1) for the entire gene, e.g. for the entire hemagglutinin gene in H1N1 (as employed in Figures 1-4); and 2) for the highest concentration of Replikins within each infectivity or lethality gene, which was designated the Replikin Peak Gene (Figures 4 and 5). Human morbidity and mortality rates data during particular time periods (CDC and WHO)\(^1\)(10,11) were compared with Replikin Counts. Replikin peptides were visualized by two means: a) by linear display of sequences of contiguous numbered amino acids in the primary structure and b) by X-ray diffraction analysis of the 3-dimensional folded structure.

**Results**

**Prospective Prediction of Outbreak Solely from Replikins Analysis**

An increase in H1N1 virus Replikin Counts was found before the outbreak of the 2009 H1N1 pandemic.

A rising mean Replikin Count of the hemagglutinin area of H1N1, between 2002 and 2009, globally from 3.5 to 9.7, was suggested by data from just one country, Mexico; confirmed in Peru-Argentina, Austria, Japan-Vietnam and globally (Table I and II, Figures 1a,1b, 2, 3).

**Table 1 – Early warning of coming outbreak by virus Replikin Count increases in human cases over the seven years before the H1N1 2009 pandemic**

| Year | Annual Replikin Counts, Mean +/-SD | Clinical Chronology |
|------|------------------------------------|---------------------|
|      | Mexico                             | Peru-Argentina      | Austria | Japan-Vietnam | Global          |
| 2002 | 3.6+/-.0.6 (N=2)                   | 4.8+/-.1.7(N=8)     |         | 3.2+/-.1.7(N=11) | 3.5+/-.1.9(N=65) |
| 2003 | 4.7+/-.0(N=18)                     | 5.0+/-.3.1(N=27)    |         | 4.8+/-.1.3(N=92) |
| 2004 | 5.2+/-.4.5(N=26)                   | 5.0+/-.2.8(N=87)    |         | 5.0+/-.3.1(N=27) |
| 2005 | 5.2+/-.1.6(N=16)                   | 4.7+/-.2.6(N=145)   |         | 5.0+/-.3.1(N=27) |
| 2006 | 6.3+/-.1.1(N=2)                    | 7.4+/-.2.6(N=42)    |         | 6.1+/-.1.6(N=583) |
| 2007 | 7.6+/-.1.6(N=47)                   | 8.0+/-.1.5(N=67)    |         | 6.8+/-.1.8(N=444) |
| 2008 | 12.2+/-.1.3(N=57)                  | 9.3+/-.1.1(N=378)   |         | 9.7+/-.1.9(N=5,899) |
| 2009 | 10.7+/-.1.5 (N=5)                  | 9.2+/-.0.2(N=13)    |         | 9.5+/-.2.8(N=453) |
| 2010 | 10.4+/-.0.5(N=3)                   | 9.2+/-.0.2(N=13)    |         | 9.5+/-.2.8(N=453) |

N = Number of cases with genomic sequence analysis, thus permitting replikins analysis. All sequences published in Pubmed were analyzed. Authors’ warning April 2008 (12) when Replikin Counts reached the range of the level previously found in the 1918 H1N1 Pandemic and the 1933-1935 severe H1N1 outbreaks (Figure 4), Beginning and end of pandemic per WHO announcements (13)
Figure 1a: Increasing Replikin Counts (p<0.001) in Human H1N1 Influenza Virus Infectivity Gene 2002 to April 2008 predicted the outbreak of April 2009. Persistence of the elevated Infectivity Gene Count in February 2010 predicted that a clinical recurrence was likely and a clinical recurrence occurred in December 2010.

Replikin Count

---Dates in 2009 and 2010 are publication dates of sequences in Pubmed—1-4 months after the specimens were collected

---2001------------2008-// 2009-----------------------------end2009------2010--------1.18.11

Cumulative number of specimens:
N HA= ----------144---------491----//-118---------------231----------------345--------------------------422--------------------------548------------------1,042//--------8,046-----
N polymerase=122---------157----// 147----------------794-----------------1,165------------------1,274------------------1,705------------------2,886//--------4,760-----
Total N= ----------266--------648-----// 265-------------1,025----------------1,510----------------1,696----------------2,253----------------3,928//----------------12,806--------

Legend for Figure 1a: Replikin Counts Annually, 2001-2008, followed by ‘real-time’ Counts 2009-2010, every 2 to 3 days, then weekly.
Replikin Infectivity Count of entire hemagglutinin gene (red, mean=heavy bar, and standard deviation of mean=light bar);
Replikin Lethality Count of entire polymerase gene (including pA, p B1-F2, p B1 and p B2 frames) (black, mean=heavy bar, and standard deviation of mean=light bar)

y axis: Replikin Count: (Number of replikins/100 amino acids)
x axis: dates genome sequences were published in Pubmed
Figure 1b continues the analysis into 2011. The Mean Replikin Count (red) of the Infectivity Gene of the virus genome shows two major increases: 1) from 2001 to 2009, a 250% increase, before the pandemic became clinically apparent; and 2) from December 2009 to January 2011. The Standard Deviation of the Mean (SD)(yellow) turned to its ‘resting’ level near zero in January 2010, then increased to greater than its 2009 pandemic level. Warning therefore was given in April of 2010 that H1N1 would reoccur, which it did in December 2010. As of January 18, 2011, the Global Mean Replikin Count for the Infectivity Gene had reached 14, and the SD had reached 4.0, both the highest respective values recorded for H1N1 since the major outbreak of H1N1 of 1930-1933. In Mexico, the Count had reached 16.7.
Figure 1b – Infectivity Gene Activation 2001-2009, and 2010-2011

Legend for Figure 1b:
H1N1 Infectivity Gene: Replikin Count, Mean (red); Standard Deviation of the Mean (SD, yellow).

Mean increased from Replikin Count of 3.5 in 2002, to 7 in December 2008, to 10.1 before the clinical outbreak in January 2009, to 13 after the clinical recurrence in December 2010. SD increased from Replikin Count 1 to 3 in 2004; then the mean increased before the clinical outbreak in 2009. Before the recurrence in 2010, the SD of the Replikin Count increased from 0 in January 2010 to almost 4 in June, 2010, and the mean again followed, from 8 to 10 for the H1N1 clinical recurrence in December 2010.

---Dates in 2009 through 2011 are publication dates of sequences in Pubmed—1-4 months after the specimens were collected.
Figure 1c shows that the Replikin Count of the replikins Lethality Gene (black), in the polymerase area, increased before the onset of the pandemic in April 2009, first in the SD in 2006 and 2007, then in both the mean and the SD in specimens of December 2008 and January 2009. Rapid reduction in Lethality Gene Replikin Counts, Mean (black) and SD (green) occurred in virus specimens from January to April 2009, and continued until the pre-pandemic Count levels of July 2008 were reached in October 2009. The mortality rate reported from Mexico was initially 27%, then following the Replikin Count, rapidly decreased in Mexico and elsewhere, to less than 1% (14,15).

---

Dates in 2009 through 2011 are publication dates of sequences – 1-4 months after specimens were collected.
The rate of change in the course of a different virus outbreak, that of the coronavirus SARS in 2003, was compared to the rate of change in the H1N1 2009 pandemic. In contrast, in SARS, after the predictive increase in the Replikin Count of the virus spike protein in 2002, before the clinical outbreak in 2003, the decrease of the Replikin Count occurred almost immediately, at the beginning of 2003 (compared to failure of H1N1 infectivity to decrease after 2 years). The total duration of the SARS outbreak also was shorter - approximately 9 months.

Independent Statistical Analysis by Todd MacKenzie, PhD, Associate Professor of Biostatistics, Dartmouth Medical School

Methods

The variable of primary interest is Replikin count. The unit of analysis is a specimen, designated by Accession Number in PubMed. The Replikin count was calculated as the number of Replikins per 100 amino acids. The goal of these analyses was to assess changes in the distribution of Replikin count by year over the calendar period 2001-2010, overall and within specific time windows; 2001-2008 (up to the year before the H1N1 pandemic of 2009), 2006-2008 (the three years before the pandemic), 2009 vs 2008, and during the pandemic, 2010 vs 2009. Analyses used year of specimen collection because month and day were available for less than 10% of specimens during the years 2001-2009. Replikin count is analyzed as a continuous variable, although we do report frequencies of some specific values because the distribution exhibits a discreteness suggesting it is a mixture distribution. Statistical analyses consisted of linear regression of Replikin count on year, Spearman correlations, as well two-sample t-tests and Wilcoxon rank sum tests to compare means between two different years. These analyses assume that Replikin counts are independent between specimens. It is possible that there is a temporal correlation (e.g. correlation of specimens taken close together in time), but assessment of temporal correlation is difficult without day, or at least month of specimens. Existence of a temporal correlation, would lower the level of significance. For the analysis of a slope between years 2001 to 2008 multiple regression was used to test for a slope over time while adjusting for the mean level of Replikin counts in the year before in order to remove any temporal correlation between successive years.

Results

Table II - Descriptive Statistics for Replikin H1N1 HA by Year

| Year | N  | Minimum | 25% | Median | Mean | 75% | Max | St.Dev | SEM |
|------|----|---------|-----|--------|------|-----|-----|-------|-----|
| 2001 | 91 | 3.2     | 3.20| 3.20   | 3.4758| 3.200| 6.4 | 0.74391| 0.077983|
| 2002 | 14 | 3.2     | 3.85| 4.25   | 4.4571| 5.500| 5.5 | 0.93703| 0.250430|
| 2003 | 89 | 0.3     | 4.10| 4.70   | 4.8573| 5.500| 9.8 | 1.29910| 0.137710|
| 2004 | 19 | 3.5     | 6.40| 6.40   | 6.5789| 8.100| 8.1 | 1.59150| 0.365110|
| 2005 | 60 | 0.3     | 4.10| 5.50   | 5.7583| 6.275| 10.1| 2.61550| 0.337660|
| 2006 | 127| 0.3     | 4.00| 4.50   | 4.5378| 6.400| 10.2| 2.68380| 0.238150|
| 2007 | 546| 0.3     | 5.50| 5.50   | 6.0015| 7.075| 13.2| 1.52370| 0.065209|
| 2008 | 414| 0.3     | 6.50| 6.60   | 6.7222| 7.300| 13.2| 1.80100| 0.088515|
| 2009 | 5141| 0.3   | 10.00| 10.10 | 9.8686| 10.200| 18.6| 1.49490| 0.020849|
| 2010 | 399| 0.9     | 9.20| 10.10 | 9.5128| 10.600| 27.5| 2.87850| 0.144100|

* Percentiles
^ Standard Error of the Mean
Figure 2: Distribution of Replikin Counts in H1N1 HA with mean and se from 2001 to 2010

Legend for Figure 2 - The discreteness of the Replikin count distribution is evident in Figure 2. The value of 10.1 occurs in 47% of the 5141 specimens in 2009. The value of 5.5 occurs in 43% of 546 specimens in 2007. The value of 0.3 occurs in 23% of the 127 specimens in 2006.

Between years 2006 and 2008 the count increased a mean of 1.0 (95% Conf. Int. 0.8 to 1.1) per year which was statistically significant: p-value (linear regression) = 1/10^{29}, p-value (Spearman correlation) < 2/10^{16}.

Between years 2001 and 2008 the count increased a mean of 0.4 (95% Conf. Int. 0.3 to 0.5) per year which was statistically significant: p-value (linear regression) = 4/10^{23}, p-value (Spearman correlation) < 2/10^{16}, p-value (multiple regression adjusting for correlation between consecutive years) = 2/10^{22}.

The increase in mean counts between 2008 and 2009 was 3.1 (95% Conf. Int. 2.9 to 3.3) which was a statistically significant increase: p-value (t-test)=1/10^{130}, p-value(Wilcoxon rank sum)<1/10^{16}.

From 2009 to 2010 there was a small decrease of 0.4 (95% Conf. Int. 0.1 to 0.6, p-value from t-test=0.0151, p-value (Wilcoxon rank sum)=0.010.

Conclusion Independent Statistical Analysis
Replikin counts increased 2 fold between the years 2001 and 2008, with much of the increase between 2006 and 2008, and they increased significantly between 2008 (the year before the pandemic), and 2009 (the year of the pandemic). They decreased slightly between 2009 and 2010. The distribution of Replikin counts appears to be dominated by a handful of modes, or 10 to 20 classes. This suggests that there are a handful of variations of the virus, such that for specimens within a class the Replikin count is identical. The increase in the proportion of classes with high Replikin counts accounts for the increase in mean Replikin counts leading up to the pandemic.

Virus Replikins Visualized Before and During the H1N1 Pandemic of 2009

The presence and position of the replikins in the H1N1 hemagglutinin gene area prior to and during the course of the H1N1 pandemic were visualized as well as counted. Table II shows that as the Replikin Count increased from 3.2 in 2002 to 11.7 in 2010, a substantial structural change occurred, most evident in the HA1 area, from amino acid 1 to 276, in which more Replikins appeared, increasing from 0 to 34 in number per 100 amino acids. At the same time, the presence of ‘free’ K’s and H’s (not within replikin structures) in the same HA1 area, decreased from 20 to 0. As the number of replikins increased, they were found to increase on the surface the hemagglutinin gene.

Table III: Replikins Counted in Primary Linear Hemagglutinin Gene Structure: Increase and Reorganization Before and During the 2009 Pandemic.

| Year   | 2002 | 2008 | 2009 | 2010 |
|--------|------|------|------|------|
| a) Replikin Count Increase in Entire Hemagglutinin | 3.2  | 5.5  | 10.1 | 11.7 |
|       | (No. replikins/100 amino acids) in entire Hemagglutinin unit |      |      |      |      |
| b) Appearance of New Replikin Structures in HA1 and HA2 Regions | 20   | 16   | 4    | 0    |
| In HA1 Region: | 0    | 13   | 28   | 34   |
| No. of replikins in HA1 (amino acid nos. 1-276) | No. of non-replikin (‘free’) K’s and H’s in HA1 (1-276) | 20   | 16   | 4    | 0    |
| In HA2 Region: | 18   | 18   | 30   | 32   |
| No. of replikins in HA2 (amino acid nos. 277-565) | No. of non-replikin (‘free’) K’s and H’s in HA2 (277-565) | 1    | 1    | 0    | 0    |

Figure 3a. Three-Dimensional Visualization of Virus H1N1 hemagglutinin shows the increasing appearance of replikin structures on the surface of hemagglutinin, and the encirclement of the sialic acid contact area of the virus, as the Replikin Counts increased from 3.2 to 10.1 during the pandemic’s development.
Replikins are in cyan. Contact points for the sialic acid host receptor are shown as orange spheres. Note at 3.2 there are no visible replikins around the sialic acid. At 5.5 the partial, and at 10.1 the almost complete encirclement of the sialic acid site by replikins. The maximal surface coverage is achieved at Replikin Count of 10.1; the surface at Replikin Count of 11.7 (not shown) is the same as that at 10.1.
Figure 3b - Anti-H1N1 Human Antibody Binding Site (red) is part of the replikins structure (cyan)

Legend for Figure 3b. Surface representation of influenza hemagglutinin of the 1918 H1N1 pandemic virus showing replikin distribution (cyan) and CR6261 antibody (16,17) binding site. The antibody binds on the outer surface of the hemagglutinin with contact on and binding to amino acid residues (red) of replikins (cyan). All 18/18 amino acid contact and binding sites in HA2 for the antibody are parts of the additional replikins covering the surface in this area at the Replikin Count of 10.1 and higher, but not present at Replikin Counts of 3.2 and 5.5, earlier in the development of the pandemic (Figure 3a).

Increased Strain-Specific Replikin Counts Correlate with Pandemics and Outbreaks 1918-2007: Retrospective Analysis 1917-2000; Prospective Analysis 2001-2011

Increases in Replikin Count were observed specifically for each strain, but not for other strains, in those years in which each strain was shown to be responsible for the outbreak (p<0.001) (Figure 4). Replikin Counts of Replikin Peak Genes were examined in all Pubmed records for sequences of H1N1, H2N2, H3N2 and H5N1 and Influenza B. An important natural control is seen in the rarely lethal influenza B over 69 years (1940-2009) for which the mean Replikin Count rarely exceeded 4. When one or both of the infectivity and lethality replikin genes are upregulated, in preparation for, or in the midst of an outbreak, the Replikin Count increases above 4 (p<0.001)(Figures 1-3, 5; Table 1) and replikins increase their presence on the surface of the hemagglutinin gene (Figure 3).
Figure 4: Retrospective: Replikin Peak Gene (RPG) Lethality Gene Replikin Counts 1918-2007 correlate with strain-specific influenza virus outbreaks and pandemics. N=14,227; p<0.001

Legend for Figure 4 - Influenza B (pale blue) through 2007 is rarely associated with lethality compared to the A viruses, and does not have a Replikin Count above 4. Mean H1N1: (black), mean H2N2: (dark blue), mean H3N2: (green), and mean H5N1: (red). Standard deviation of the mean (SD): (grey bars). Note that a virus ‘at rest’ has an absent or small SD. Increased Replikin Counts, above 4, are specific both temporally and with regard to the particular strain for each pandemic or severe outbreak. Spaces indicate years where sequence data was not available on Pubmed.
Evolution, trans-flu strain sharing, and conservation of Replikins

The substitutions which have occurred in the Goose Replikin from 1917 and the 1918 H1N1 pandemic to 2011 shown in Table IV appear to be selective and retained (conserved) rather than random. The conservation data from 1917-2004 (1) is updated to 2010 in Table IV.

Table IV: Evolution, trans-flu strain sharing, and conservation 1917 to 2010 of one 29 amino acid replikin: *The Goose Replikin is a 29 amino acid replikin peptide in the hemagglutinin protein of influenza virus beginning with kk and ending with hh. All ks and hs are shown in bold type. Individual replikin sequences of individual strains of influenza virus containing residues identical to The Goose Replikin amino acids in blue, amino acid substitutions in yellow and red.

| Peptide Length | Year | Strain |
|----------------|------|--------|
| [<--------29 Amino Acids--------->] | 1917 | H1N_Influenza Goose Replikin |
| kkgtsypklskstynkgkvevlwgvhhk | 1918 | H1N1 HumanInfluenzaPandemic |
| kkg-sypklskstynkgkvevlwgvhhk | 1930 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1933 | H0N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1976 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1977 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1979 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1980 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1981 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1981 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1991 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1992 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1996 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1996 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1997 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1998 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1999 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 2000 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 2001 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 2002 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 2009 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 2010 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 2011 | H1N1 |
| kkg-sypklskstynkgkvevlwgvhhk | 1999 | H1N2 Influenza |
| kkg-sypklskstynkgkvevlwgvhhk | 2000 | H1N2 |
| kkg-sypklskstynkgkvevlwgvhhk | 2001 | H1N2 |
| kkg-sypklskstynkgkvevlwgvhhk | 2001 | H1N2 |
| kkg-sypklskstynkgkvevlwgvhhk | 2002 | H1N2 |
| kkg-sypklskstynkgkvevlwgvhhk | 2002 | H1N2 |
| kkg-sypklskstynkgkvevlwgvhhk | 1968 | H3N2 Human Influenza Pandemic |
| kkg-sypklskstynkgkvevlwgvhhk | 1979-2003 | H7N7 Influenza |
| kkg-sypklskstynkgkvevlwgvhhk | 1957 | H2N2 Human Influenza Pandemic |
| kkg-sypklskstynkgkvevlwgvhhk | 1957 | H2N2 Human Influenza Pandemic |
| kkg-sypklskstynkgkvevlwgvhhk | 1959 | H5N1 Influenza (Scotland) |
H5N1 Outbreaks and Lethality

H5N1 is very different from H1N1 in terms of its infectivity and lethality. H5N1’s infectivity is lower and lethality higher than H1N1. During the worst recorded period of H1N1, the 1918 pandemic, the mortality rate was estimated to be only 2.5%, whereas the human mortality rate for H5N1 has been reported by WHO to be as high as 82% (11). From specimens of 1957, when H5N1 sequences were first reported, until 1996, the mean Replikin Count did not exceed 4 (Figure 5). A slight but statistically significant increase in H5N1 Lethality Gene Replikin Count occurred in 1996, in advance of the outbreak in Hong Kong in 1997. Where H5N1 outbreaks had previously been in poultry, in 2007 a human outbreak occurred. In 1996, an increase in Replikin Count precursed the 2007 Hong Kong outbreak in which 30 human cases occurred with a mortality rate of 27%. The Count increased in 2005-2008 (Figure 6) with subsequent animal and human outbreaks in Asia and Egypt in 2008-2010 (WHO)(5).
Figure 5. Replikin Counts in H5N1 Replikin Peak Genes in Infectivity and Lethality Genes in birds, and for H1N1 in humans, in 2004-2008.

Prediction of Geographic Location of H5N1 (Avian flu) Outbreak in Indonesia
Instead of comparing neighboring genes, neighboring countries were compared for the Replikin Counts of H5N1 scout infections in humans over several years. In the replikin prediction of 2005-2006, Indonesia was predicted to be the country that would be worst affected in terms of increased human mortality (18). Following the replikin prediction, 277 human H5N1 cases were reported and the human mortality rate increased in Indonesia from 40% to 82% (WHO)(11).

Concurrent H5N1 and H1N1 Build-Up in 2011
Because of the increase in Replikin Count in birds from 2002 to 2008 (Figure 5), and the increased Counts in its precursor H9N2 in chickens (19,20), the authors issued a warning in January of 2009 that H5N1 outbreaks would surge (19,20). By January of 2010, these H5N1 outbreaks occurred; in addition to outbreaks in birds and chickens in 63 countries (21,22), human cases appeared monthly, most prominently in Egypt, where there were 106 WHO confirmed human cases and 32 deaths (mortality rate 30%)(25) as of March 16, 2010. Also in March, 2011, the case-fatality rate (CFR) was reported to be 34%, versus 60% for other countries with human cases (4). Now, as of June and July, 2011 respectively, the cumulative mortality rate is reported by WHO to have risen since 2006 to 34.7% (3); and the current mortality rate reported by the Egyptian government is 38.7% (36). Globally, as of December 18, 2010, the H5N1 hemagglutinin Infectivity Gene Replikin Count had reached its highest level in humans since 1998.

The H1N1 Infectivity Gene Replikin Count had peaked globally in January 2011 at 13.5+/-4.2; in Mexico it declined by August, 2011 only to a mean of 11.9+/-2.5 but 3/21 specimens had levels of 16.7 (Figure 6a). As stated in a footnote to Figure 1c above, this failure of the Replikin Infectivity gene to return to pre-outbreak levels is in marked contrast to the SARS Replikin Count which promptly returned to pre-outbreak Replikin levels to signal the end of the SARS outbreak. The data strongly suggest that the H1N1 pandemic of 2009 is continuing to develop.

While there were insufficient H1N1 polymerase sequences submitted to PubMed from Mexico as of August, 2011, the Global H1N1 Lethality Gene Replikin Count level in humans, which had decreased at the end of 2010 to its pre-pandemic level of 2.0+/-0.2, as of August 2011 was again increased to 5.6+/-4.9, in the range that it was before the 2009 H1N1 pandemic (Figure 6b).
The persistent elevation of the concentration of both the H1N1 Infectivity and Lethality Genes, from previous examples, is invariably followed by clinical outbreaks. The current combined activation of both the Infectivity and Lethality Genes of H5N1 in 2011 is of concern.

**Figure 6a**
Replikins Infectivity Gene, H1N1, Annually, in Mexico Only  
Mean(Blue), SD(Red)

**Figure 6b**
Replikins Lethality Gene, H1N1, Annually, Globally  
Mean(Blue), SD(Red)

**Legend for Figures 6a and 6b** – In Figure 6a, the persistent increase in the Replikin Count after the 2009 Pandemic had been declared over by WHO (13), suggests that the pandemic is not over. In Figure 6b, as seen in Figure 1a above, the Lethality Gene peak increase occurred 2 years before the pandemic outbreak, then declined promptly to its pre-outbreak level. However, this has been followed by the recurrence in 2011 of a marked increase in Replikin Count again to pre-pandemic levels.

**Discussion**

The Risk of a Combined H1N1-H5N1 Pandemic. The risk of a combined H1N1 (high infectivity) - H5N1 (high lethality) pandemic may have increased because of the simultaneous rise in the Replikin Counts of each of these two virus strains to their highest levels recorded since 1918 and the appearance of ‘scout’ virus outbreaks of H1N1 again in Mexico and of H5N1 in Egypt. The simultaneous emergence of two virus strains with record high Replikin Counts has not been observed previously. All increases observed in the past 93 years have been in only one strain (see Figure 4). Simultaneous pandemic outbreaks could bring the two virus strains more frequently in contact with each other, facilitating transfer of genomic material to form a hybrid with pathogenic capability of each strain. H1N1- H5N1 combination is now only a risk; it has not yet occurred, and it is not certain that it will occur. However, with simultaneous increases in the Replikin Counts of each strain, outbreaks of H1N1 and H5N1 are now in progress in Mexico and
Egypt respectively (2,41,44), and in preparation to meet this threatened combination, the authors have prepared the first completely synthetic TransFlu™ Replikins Vaccine against these two and the other common influenza strains. TransFlu™ was successful in blocking H5N1 (23) in the first of its continuing independent trials in the U.S. and elsewhere.

**Pan-Influenza Vaccines.** A new technology is presented here of Replikins, genomic peptides which act as epitopes, conserved and shared inter-strain. These facts of conservation and sharing of the newly defined Replikins epitopes were not known, and this unawareness and the believed absence of conservation and sharing was the basis of the reason given for decades of not being able to make pan-influenza vaccines which could be used from year to year. However, having reviewed Replikins’ technology since 2003 from data provided by the authors, the National Institutes of Health (NIH) first confirmed the influenza epitopes earlier defined by specific Replikins by landing anti-flu antibodies on them (16). This NIH data was simultaneously independently confirmed by studies at Scripps (17). NIH has now announced that they hope to have a pan-influenza vaccine, possibly in 5 years, based on these epitopes (24). As noted above, the Replikins TransFlu™ pan-influenza vaccine, based on the Replikins epitopes confirmed by the NIH, is available now, and is being tested (23).

**Improved vaccine production.** The need for new vaccines in influenza is widely recognized (6,9). The conservation of replikins over time and the sharing of influenza trans-strain replikins (ref. 1 and Table V) has been used to design and produce effective solid phase synthetic vaccines in as little as seven days, permitting a more rapid response to newly appearing strains (23). The first synthetic replikins pan-influenza vaccine was successful against H5N1; it also blocked virus excretion, providing for the first time the potential to block the development of H5N1 reservoirs (23).

**Specific Visible Changes in the Transformation of Virus Structure.** Replikins are shown to be present on the surface of the hemagglutinin unit and the coverage increased as the Replikin Count increased toward the onset and during the course of the H1N1 2009 pandemic (Figure 3a). The early visible reorganization of virus replikin structures from 2002 to 2008 (Table II and Figure 3a) is here shown to occur in advance of and during the pandemic. This change is concomitant with changes from neutral to acidic conformation assisting fusion of the virus with the host membrane during virus entry (27-30). The actual visualization of Replikins, with the ability to count them, supports the reality and practical clinical relevance of Replikin structures. Consequences of the surface increase may be: 1) the replikins themselves may contribute to the increased infectivity and lethality of the virus by encircling and supporting the essential active site of the virus for entry via sialic acid (Figure 3a). The central role of the sialic acid receptor for influenza virus entry into host cells was demonstrated a) in 1959 by the authors by blocking the virus entry into brain cells by sialic acid conjugates from brain gangliosides (25), and b) by the release of similar sialic acid conjugate decoys (sialoresponsins) by the chorioallantoic membrane of the chick egg under influenza attack (26). The molecular definition of the reaction between the hemagglutinin unit and the sialic acid receptor pocket of the host membrane was recently demonstrated (27-30). 2) The increased surface coverage by replikins may represent an increased ’shield’ or ’armour’ against the immune system of the host (Figures 3a and 3b).

**Replikins in other lethal zoonotic organisms and other pathogenic states.** The 2010-2011 Foot and Mouth Disease virus outbreaks were predicted by the Replikin Count in 2009, one year previously (31). Lethal pathogens other than viruses also contain replikins; for example, in bacteria, such as tuberculosis, and in trypanosomes (malaria)(8). Replikins also play a key role in human cancer (8,32,33,34). The authors have postulated that viruses, bacteria, trypanosomes, cancer cells and other biological organisms may be carriers or vectors for mobile pathogenic
replikin sequences associated with rapid replication (communication in preparation). Some replikins are associated with rapid replication in healthy growth throughout biology, as in algae and food plants (8).

Acknowledgements

PubMed and the NCBI were used extensively in these studies. The U.S. Centers for Disease Control (CDC) kindly provided epidemiological data on the strains of influenza associated with influenza outbreaks each year from 1918 to 2005. Dr. S. Winston Bogoch, with the authors, designed the software used in this study, Anne Bogoch Borsanyi was responsible for the preparation of research protocols and contributed to the writing of this manuscript. We thank Dr. Todd MacKenzie, Associate Professor of Biostatistics, Dartmouth Medical School, New Hampshire, for the independent statistical analysis of the data, as seen in Table II and Figure 2. We are grateful to Professor Mark Jackwood, University of Georgia, for the independent testing of synthetic replikins TransFlu™ vaccine against H5N1 in chickens (23). We are grateful to the United Kingdom Department of Trade and Investment for sponsoring academic seminars to introduce the technology of the Replikins Bioradar Global Surveillance System™(35).

References

1. Global Public Health Intelligence Network (GPHIN), Public Health Agency of Canada. [http://www.phac-aspc.gc.ca/media/nr-rp/2004/2004_gphin-rmispbk-eng.php].
2. Cordova, JA. Minister of Health, Mexico. [English.xinhauet.com] April 8, 2011.
3. WHO, [FluTrackers.com] July 7, 2011.
4. Roos, R. CIDRAP News. March 22, 2011.
5. Noamany, M. Minister, Egypt. [FluTrackers.com] July 7, 2011.
6. Fauci, A.S, Lambert LC. Influenza vaccines for the future. N Engl J Med;363:2036-44, 2010.
7. Fineberg, H.V. Draft report of the Institute of Medicine. WHO website, March 9, 2011.
8. Bogoch S, Bogoch ES. Replikins: the chemistry of rapid replication. Begell House, N.Y. 2005. Library of Congress ISBN 1-56700-200-5.
9. ‘Bogoch Software’: FluForecast™ and Bioradar™ from Replikins, Ltd., London Bioscience Innovation Centre, 2 Royal College Street, London NW1 0NH, United Kingdom.
10. CDC (Centers for Disease Control). Estimates of deaths associated with seasonal influenza- United States, 2008-2010.MMWR Morb Mort Wkly Rep 2010;59:1057-62.
11. WHO (World Health Organization). Cumulative Number of Confirmed Human Cases of Avian Influenza A (H5N1) Reported to WHO. 16 March 2010, and 2011.
12. Replikins Press. H1N1 Virus with Highest Replikin Count Since the 1918 Pandemic Identified in the U.S. and Austria. [replikins.com] Report Number 18, April 7, 2008.
13. WHO (World Health Organization). Press releases: Swine flu pandemic declared to have begun on June 11, 2009, and declared over on August 11, 2010.
14. Editorial. The Lancet, Volume 373, Issue 9674, Page 1495, 2 May 2009 >doi:10.1016/S0140-6736(09)60826-6. 29 April 2009.
15. Echevarría-Zuno, S, Mejía-Aranguré, JM, Mar-Obeso, AJ, Grajales-Muñiz, C, Robles-Pérez, E, González-León, M, Ortega-Alvarez, MC, Gonzalez-Bonilla, C, Rascón-Pacheco, RA, Borja-Aburto, VH. Infection and death from influenza A H1N1 virus in
Mexico: a retrospective analysis The Lancet.com Published online November 12, 2009 DOI:10.1016/S0140-6736(09)61638-X.

16. Sui, J, Hwang, WC, Perez, S, Wei, Ge, Aird, D, Chen, Li-mei, Santelli, E, Stec, B, Cadwell, G, Ali, M, Wan, H, Murakami, A, Yammanuru, A, Han, T, Cox, N, Bankston, L, A, Donis, R, Liddington, R, C, and Marasco, W, A. Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. Nature Structural and Molecular Biology, published online: 22 February, 2009, doi:10.1038-nsmb.1566. Nature Structural and Molecular Biology 16:265-273, 2009.

17. Ekeirt, D. C., Bhabha, G, Eislinger, M-A, Friesen, R, H, Jongeneelen, M, Throsby, M, Goudsmit, J, and Wilson, I, A., Antibody Recognition of a Highly Conserved Influenza Epitope, Science DOI: 10.1126-science.1171491, Science Published online, Feb. 23 2009; Science 324:246-251, 2009.

18. Replikins Press. A new way to predict outbreaks: replikin peptide concentration in H5N1 influenza virus genome as a marker for lethal outbreaks. <replikins.com> Report Number 22, November 12, 2008.

19. Replikins Press. Both H5N1 Bird Flu and H1N1 Swine Flu Replikin Counts™ Increasing. <replikins.com> Report Number 35. April 9, 2010.

20. Replikins Press. New H1N1 Genomic Data Shows Decrease in Replikin Count* of Lethality Gene; Replikin Count of Infectivity Gene Remains High. <replikins.com> Report Number 31, September 30, 2009.

21. Replikins Press. Rising H9N2 Influenza Replikin Count Has Doubled That of H5N1. <replikins.com> Report Number 24. January 15, 2009.

22. CIDRAP (The Center For Infectious Disease Research And Policy) A year's experience with the 2009 H1N1 influenza virus has underscored the endless unpredictability of flu. Apr 23, 2010. CIDRAP (The Center For Infectious Disease Research And Policy) . FAO says H5N1 outbreaks are up this year of FAO AIDE News, Sep 2 issue, 2010.

23. Jackwood MW, Bogoch, S., Bogoch ES, Hilt, D and Williams SW. Efficacy of a Replikin peptide Vaccine Against Low Pathogenicity Avian Influenza H5 Virus. Avian Diseases 53(4): 613-617, 2009.

24. Collins, F. Director, National Institutes of Health. USA Today. Front page. July 27, 2011.

25. Bullough PA, Hughson FM, Skehel JJ, Wiley DC (1994) Structure of influenza hemagglutinin at the pH of membrane fusion. Nature 371:37–43.

26. Russell, RJ, Kerry PS, Stevens, JS, Steinhauser, DA, Martin, SR, Gamblin, SJ, and Skehel, JJ. Structure of influenza hemagglutinin in complex with an inhibitor of membrane fusion. PNAS 105: No. 46: 17736-17741, 2008.

27. Weis, W, Brown JH, Cusack, S., Paulson, JC, Skehel, JJ and Wiley, DC. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. Nature 333, 426-431, 1988.

28. Yamada, S, Suzuki, Y, Le, MQ, Nidom, CA, Sakai-Tagawa, Y, Mueamoto, Y Ito, M, Kiso, M, Horimoto, T, Shinya, K, Sawada, T, Kiso, M, Usui, T, Murata, T, Lin, Y, Hay A, Haire, LF, Stevens, DJ, Russell, RJ, Gamblin, SJ, Skehel, JJ, Kawaoka, Y. Hemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. Nature 444:378-382, 2006.
29. Bogoch, S. Lynch, P. and Levine, A.S. Influence of Brain Ganglioside upon the Neurotoxic Effect of Influenza Virus in Mouse Brain. *Virology* 7:161, 1959
30. Bogoch, S., Gilfillan, R.F. and Evans, P. Sialoresponsin: Accumulation of a Neuraminic Acid-containing Substance(s) in Chorioallantoic Fluid During the First Minutes of Viral Infection. *Nature* 196:649, 1962.
31. Replikins Press. Current Severe Outbreaks of Foot and Mouth Disease Predicted by Replikins' BioRadar™ One Year in Advance. <replikins.com> Report number 36, June 2009.
32. Bogoch, S. Production of a Synthetic General Cancer Vaccine Which Augments the Concentration of Anti-Malignin Antibody in Vivo. Cancer Detection and Prevention 22(1): S-159, 1998.
33. Bogoch, S. Bogoch, E., Tsung, Y. Monoclonal Anti-Malignin Antibodies. The Lancet 2:141-142, 1981.
34. Bogoch, S., Bogoch, E. Malignin Antibody and Early Malignancy. The Lancet 337:977, 1991
35. UK Department of Trade and Investment. Press Release. Replikins Bioradar Global Surveillance System. Proceedings of Seminar sponsored by the Department of Trade and Investment, UK. London, April 30, 2010.
36. Shin Young-Soo, “Official WHO Press Release Warns Of Potential “Reassortment” Between H1N1 and H5N1” WHO Press Release, November 25-29, 2009
37. Mckay, B. “More Deadly Swine Flu? CDC Mixes H1N1, H5N1 Viruses in Tests”. Wall Street Journal. September 17, 2009
38. Bogoch, S and Bogoch, E. “Killer Replikins in Pandemic Flu and Cancer”. Volume II of Ref. 8, 2011, in press.
39. Replikins Press. “H1N1 Influenza Virus with Highest Replikin Count™ Since the 1918 Pandemic Identified in the U.S. and Austria.” April 7, 2008 http://www.replikins.com/press_18.html
40. Replikins Press. “Both H5N1 Bird Flu and H1N1 Swine Flu Replikin Counts™ Increasing”, April 9, 2010 http://www.replikins.com/press_35.html.
41. Pan American Health Organization Alert. Archive Number 20110425.1292. Subject PRO/EDR> Influenza (32): PAHO alert. 25-APR-2011
42. Replikins Press. “Rising Human H5N1 ‘Bird Flu’ High-Virulence Sequences Found By Replikins, Ltd. “ November 6, 2006. http://www.replikins.com/press_7.html
43. Replikins Press. “Indonesia Reports Experiencing Human H5N1 Mortality Increase, as Predicted Last Year by Replikins' FluForecast® Quantitative Virus Analysis “. June 8, 2007. http://www.replikins.com/press_12.html
44. Schnirring, L. CIDRAP News. H5N1 Jun 1, 2011
