

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
27 November 2008 (27.11.2008)

PCT

(10) International Publication Number
WO 2008/143717 A2

(51) International Patent Classification:	11/923,559	24 October 2007 (24.10.2007)	US
<i>C07K 14/11</i> (2006.01)	<i>C12Q 1/70</i> (2006.01)	60/982,338	24 October 2007 (24.10.2007) US
<i>A61K 39/145</i> (2006.01)	<i>C07H 21/00</i> (2006.01)	60/991,676	30 November 2007 (30.11.2007) US
<i>C07K 16/08</i> (2006.01)			

(21) International Application Number:
PCT/US2008/000645

(22) International Filing Date: 18 January 2008 (18.01.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/880,966	18 January 2007 (18.01.2007)	US
60/898,097	30 January 2007 (30.01.2007)	US
11/755,597	30 May 2007 (30.05.2007)	US
60/954,743	8 August 2007 (08.08.2007)	US
60/935,499	16 August 2007 (16.08.2007)	US
60/935,816	31 August 2007 (31.08.2007)	US
60/982,333	24 October 2007 (24.10.2007)	US
60/982,336	24 October 2007 (24.10.2007)	US

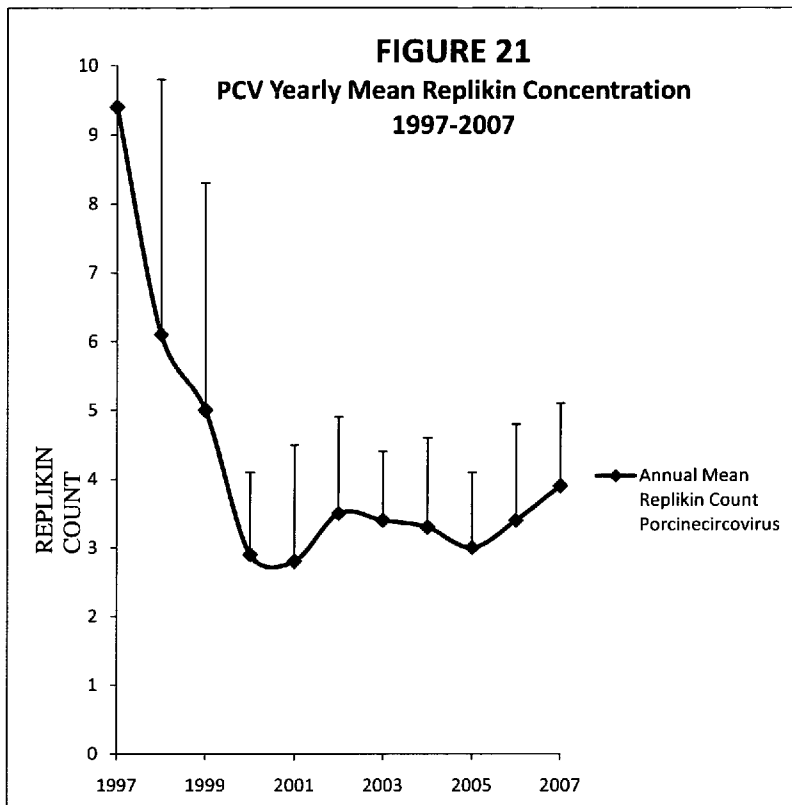
(71) Applicants and
(72) Inventors: **BOGOCH, Samuel** [CA/US]; 46 East 91st Street, New York, NY 10028 (US). **BOGOCH, Elenore, S.** [US/US]; 46 East 91st Street, New York, NY 10028 (US). **BOGOCH, Samuel, Winston** [CA/US]; 429 62nd Street, Oakland, CA 94609 (US). **BORSANYI, Anne, Elenore** [CA/US]; 122 Carlton Street, Brookline, MA 02466 (US).

(74) Agent: **AMOROSO, Stephanie**; Milbank, Tweed, Hadley & McCloy, LLP, International Square Building, 1850 K Street, NW, Suite 1100, Washington, DC 20005 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN,

[Continued on next page]

(54) Title: METHODS OF DETERMINING LETHALITY OF PATHOGENS AND MALIGNANCIES INVOLVING REPLIKIN PEAK GENES



(57) Abstract: The present invention provides methods of identifying lethal, virulent and rapidly replicating viruses, organisms, and malignancies comprising comparing Replikin concentrations among different viruses, organisms, or malignancies. The present invention further provides isolated Replikin Peak Genes associated with increased lethality, virulence and rapid replication, for diagnostic, therapeutic and predictive purposes.

WO 2008/143717 A2



MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),

Published:

— *without international search report and to be republished upon receipt of that report*

METHODS OF DETERMINING LETHALITY OF PATHOGENS AND MALIGNANCIES INVOLVING REPLIKIN PEAK GENES

The instant application contains a "lengthy" Sequence Listing which has been submitted via CD-R in lieu of a printed paper copy, and is hereby incorporated by reference in its entirety. Said CD-R, recorded on January 17, 2008, are labeled CRF, "Copy 1," "Copy 2" and "Copy 3", respectively, and each contains only one identical 1.36 Mb file (27129302.txt).

[0001] This application claims priority to U.S. Provisional Appln. Ser. No. 60/991,676, filed November 30, 2007, U.S. Appln. Ser. No. 11/923,559, filed October 24, 2007, U.S. Provisional Appln. Ser. No. 60/982,336, filed October 24, 2007, U.S. Provisional Appln. Ser. No. 60/982,333, filed October 24, 2007, U.S. Provisional Appln. Ser. No. 60/982,338, filed October 24, 2007, U.S. Provisional Appln. Ser. No. 60/935,816, filed August 31, 2007, U.S. Provisional Appln. Ser. No. 60/935,499 filed August 16, 2007, U.S. Provisional Appln. Ser. No. 60/954,743, filed August 8, 2007, U.S. Appln. Ser. No. 11/755,597, filed May 30, 2007, U.S. Provisional Appln. Ser. No. 60/898,097, filed January 30, 2007, and U.S. Provisional Appln. Ser. No. 60/880,966, filed January 18, 2007, each of which is incorporated herein by reference in its entirety. This application additionally incorporates herein by reference: U.S. Provisional Appln. Ser. No. 60/853,744, filed October 24, 2006, U.S. Appln. Ser. No. 11/355,120, filed February 16, 2006, U.S. Appln. Ser. No. 11/116,203, filed April 28, 2005, U.S. Appln. Ser. No. 10/860,050, filed June 4, 2004, U.S. Appln. Ser. No. 10/189,437, filed July 8, 2002, U.S. Appln. Ser. No. 10/105,232, filed March 26, 2002, now U.S. Patent No. 7,189,800, U.S. Appln. Ser. No. 09/984,057, filed October 26, 2001, and U.S. Appln. Ser. No. 09/984,056, filed October 26, 2001, now U.S. Patent No. 7,176,275, each in its entirety.

TECHNICAL FIELD OF THE INVENTION

[0002] This invention relates generally to identifying virulent and lethal strains of pathogenic viruses, pathogenic organisms and malignancies through identifying concentrations of the class of small peptides known as Replikins, and to diagnosis, prevention and treatment of disease from such virulent and lethal pathogens and malignancies.

BACKGROUND OF THE INVENTION

[0003] Rapid replication is characteristic of virulence in, among other things, certain bacteria, viruses and malignancies. The inventors have described a quantitative chemistry common to rapid replication in different organisms, viruses and malignancies. The chemistry of rapid replication described by the inventors is present in a family of conserved small

protein sequences related to rapid replication called Replikins. A correlation between increased concentrations of Replikin sequences and increased replication and virulence has been observed in a range of viruses and organisms. Replikin sequences, therefore, offer new targets for developing effective methods of predicting and treating viral outbreaks.

Replikin Sequences in Malignancies and Viral and Bacterial Pathogens

[0004] A Replikin sequence is an amino acid sequence of 7 to about 50 amino acids comprising a Replikin motif. A Replikin motif comprises (1) at least one lysine residue located at a first terminus of the motif and at least one lysine residue or at least one histidine residue located at a second terminus of the motif; (2) a first lysine residue located six to ten residues from a second lysine residue; (3) at least one histidine residue; and (4) at least 6% lysine residues. A Replikin sequence may comprise a terminal lysine and may further comprise a terminal lysine or a terminal histidine. A Replikin peptide or Replikin protein is a peptide or protein consisting of a Replikin sequence.

[0005] The inventors have identified Replikin sequences in oncogenic cells and in viral and organismal proteins associated with rapid replication and virulence. Additionally, higher concentrations of Replikin sequences in the genomic code have now been associated with a variety of infectious and pathogenic agents including human cancer, HIV, plant viruses, and a range of pathogenic animal and human viruses. Further, the correlation between the concentration of Replikin sequences in viral or organismal proteins and major outbreaks of disease and the correlation between the concentration of Replikin sequences in malignancies and poor prognoses are both significant.

[0006] Replikin sequences have been observed to be conserved in human cancers generally and in many pathogenic organisms and viruses, including conservation in both intrastrain and interstrain influenza viruses, for as long as 90 years based on data going back to the 1917-18 flu pandemic. Concentration of Replikin sequences in viral genomes has been shown to increase prior to strain-specific outbreaks and increased mortality in SARS, in influenza, in H5N1 bird flu and now in many other viral and non-viral pathogens. An increase in concentration of production of proteins containing Replikin sequences also has been shown in cancer as replication increases.

[0007] Within the last century there have been three influenza pandemics, each strain specific: H1N1 in 1918; H2N2 in 1957; and H3N2 in 1968. The inventors have established that prior to each pandemic there was a strain-specific increase in the concentration of

Replikin sequences within the strain. The strain-specific increase in Replikin concentration was followed by a decrease in Replikin concentration and several years later a rebound increase in Replikin concentration associated with a strain-specific rebound epidemic. The Replikin algorithm provided the first chemistry that correlated with influenza epidemics and pandemics.

[0008] A similar correlation between the outbreaks of H5NI (Bird Flu) between 1997 and 2007 and the concentration of Replikin sequences in the viral proteins has been demonstrated. Likewise, a correlation has been established between the global outbreak of SARS coronavirus in 2003 and an increase in the concentration of Replikin sequences in the proteins of coronavirus. In another study, Replikins in two strains of human HIV-1 virus demonstrated that the Replikin concentration in the rapidly replicating strain was six fold greater than that of the slowly replicating strain. No instances of rapid replication have been observed in all the viruses and organisms examined wherein the Replikin concentration did not significantly increase as compared to the Replikin concentration in the dormant state.

[0009] The Replikin algorithm was initially discovered in Glycoprotein 10B, a membrane glycoprotein isolated from brain glioblastoma multiforme, lymphoma and breast cancer cells (U.S. Pat. No. 6,242,578 B1). A constituent peptide of Aglyco 10B, malignin, was observed to be enriched in cell membranes tenfold during anaerobic replication while cell number was observed to increase only five-fold. This increase in membrane concentration of the malignin protein in rapid replication of glioma cells suggested an integral relationship of the Replikins in malignin to replication of the glioblastoma multiforme.

[00010] Hydrolysis and mass spectrometry of malignin yielded a 16-mer peptide that included the Replikin sequence kagvafhkk (SEQ ID NO:3658). This peptide, which is absent from the normal human genome, was assumed to be acquired. Homologues of the Replikin sequence were found in all tumor viruses (that is viruses that cause cancer), and in replicating proteins of algae, plants, fungi, viruses and bacteria.

[00011] When the glioma Replikin was synthesized in vitro and administered as a synthetic vaccine to rabbits, abundant antimalignin antibody was produced. This production of abundant antimalignin antibody established that the peptide alone is an epitope, that is, it is a sufficient basis for an immune response observed in cancer patients wherein antimalignin antibodies are naturally produced. A 16-mer peptide containing the glioma Replikin produced both IgM and IgG forms of the antibody.

[00012] A study of 8,090 serum specimens from cancer patients and controls demonstrated that the concentration of antimalignin antibody increases with age in healthy individuals, as the incidence of cancer in the population increases, and increases further two to three-fold in early malignancy, regardless of cell type. In vitro this antibody was observed to be cytotoxic to cancer cells at picograms (femtomoles) per cancer cell, and in vivo the concentration of antimalignin antibody related quantitatively to the survival of cancer patients. As shown in glioma cells, the stage in cancer at which cells have only been transformed to the immortal malignant state but remain quiescent or dormant, now can be distinguished from the more active life-threatening replicating state which is characterized by the increased concentration of Replikins.

[00013] Using the sequence of the glioma Replikin peptide (kagvafhkk) (SEQ ID NO:3658) as a template, and constructing a “3-point- recognition” method to visually scan protein sequences of several different organisms, a new class of peptides, the Replikins, was revealed in organisms as diverse as algae, yeast and viruses. Surprisingly, these peptides were found to be concentrated in larger “replicating” and “transforming” proteins.

[00014] An infrequent occurrence of homologues was observed in “virus peptides” as a whole (1.5%), and in other peptides not designated as associated with malignant transformation or replication such as “brain peptides” and “neuropeptides” (together 8.5%). A surprisingly high frequency of occurrence of homologues was identified in tumor viruses, transforming proteins and cancer cell proteins. For example, 100% of identified tumor viruses contain Replikin sequences. 85% of transforming proteins contained Replikin sequences and 97% of cancer proteins contained Replikin sequences.

[00015] Further, Replikins were identified in such proteins as *Saccharomyces cerevisiae* replication binding protein; the replication associated protein A of maize streak virus; the replication-associated protein of *Staphylococcus aureus*; the DNA replication protein of bovine herpes virus 4; and the mealgrid herpes virus 1 replication binding protein. Replikin-containing proteins also are associated frequently with redox functions, and protein synthesis or elongation, as well as with cell replication.

[00016] The highest concentration of Replikin sequences in an organism or virus that had been analyzed and reported was 111 Replikin sequences per 100 amino acids in the extraordinarily-rapidly-replicating parasitic protozoa *Plasmodium falciparum* (reportedly responsible for 90% of malarial deaths in humans) (herein sometimes referred to as malaria).

P. falciparum has been observed to replicate 11,000 times in 48 hours during passage of the parasite from liver to blood in the host.

[00017] A significant feature of Replikin sequences observed in *P. falciparum* was a marked overlapping of Replikin structures throughout malarial proteins. For example, there are nine overlapping Replikins in the 39 amino acid sequence of SEQ ID NO. 3667 (Replikin concentration=23.1/100 amino acids); and 15 overlapping Replikins in the 41 amino acids of SEQ ID NO. 3668 (Replikin concentration=36.6/100 amino acids). Both of these overlapping Replikin structures occur in blood stage trophozoites and schizonts. This mechanism of lysine multiples was also seen in the Replikins of cancer proteins such as in gastric cancer transforming protein, ktkkgnrvsptmkvth (SEQ ID NO: 3669), and in transforming protein P21B (K-RAS 2B) of lung, khkekmskdgkkkkkks (SEQ ID NO: 3670).

Replikin Scaffolds

[00018] In monitoring Replikin sequences in influenza virus, the inventors have additionally identified a sub-family of conserved Replikin sequences known as Replikin Scaffolds or Replikin Scaffold sequences. Replikin Scaffolds were initially identified in conserved structures in particularly virulent influenza viruses. Included among these strains were the viruses causing the pandemics of 1918, 1957, 1968 and virulent strains of the H5N1 "bird flu" strain of influenza virus. Analogues of Replikin Scaffold sequences have since been identified in the virulent and rapidly replicating SARS coronavirus. See U.S. Published Application No. 2007/0026009.

[00019] Scaffolding of Replikin sequences homologous but not identical to the algorithm of the identified Replikin Scaffold has also been identified in *P. falciparum*. Replikin scaffolding in general has been related to an increase in Replikin concentrations in pathogenic genomes where it has been identified. In *P. falciparum*, scaffolding contributes significantly to the very high Replikin concentration noted in the proteins of the protozoa.

Influenza

[00020] Virulent and lethal outbreaks of influenza are a continuing challenge to world health and the medical practitioner is increasingly aware of the continued threat of virulent and lethal influenza pandemics that require new methods of predicting virulence and lethality and will require new methods and compounds for treatment. Influenza is an acute respiratory illness of global importance. Despite international attempts to control influenza virus outbreaks through vaccination, influenza infections remain an important cause of morbidity

and mortality. Worldwide influenza pandemics have occurred at irregular and previously unpredictable intervals throughout history and it is expected that influenza pandemics will continue to occur in the future. The impact of pandemic influenza is substantial in terms of morbidity, mortality and economic cost.

[00021] Influenza vaccines remain the most effective defense against influenza virus, but because of the ability of the virus to mutate, and the availability of non-human host reservoirs, it is expected that influenza will remain an emergent or re-emergent infection. Global influenza surveillance indicates that influenza viruses may vary within a country and between countries and continents during an influenza season. Virologic surveillance is of importance in monitoring antigenic shift and drift. Disease surveillance is also important in assessing the impact of epidemics. Both types of information have provided the basis of vaccine composition and use of antivirals. However, traditionally there has been only annual post hoc hematological classification of the increasing number of emerging influenza virus strains, and no specific chemical structure of the viruses was identified as an indicator of approaching influenza epidemic or pandemic. Until recently, the only basis for annual classification of influenza virus as active, inactive or prevalent in a given year was the activities of the virus hemagglutinin and neuraminidase proteins.

[00022] There is a need in the art for methods of predicting increases in virulence and lethality of influenza prior to outbreaks. There is likewise a need in the art for methods of preventing and treating outbreaks caused by virulent strains of influenza. Because of the annual administration of influenza vaccines and the short period of time when a vaccine can be administered, strategies directed at improving vaccine coverage are of critical importance.

Equine Influenza Virus

[00023] Equine influenza is a common upper respiratory disease of the horse currently caused by the H3N8 strain of equine influenza virus (EIV). Typical symptoms of equine influenza include a dry hacking cough, nasal discharge, and fever. The viral disease is considered enzootic in Europe, the United States and parts of Asia. Significant outbreaks have also been observed in South America, China, and India.

[00024] The first outbreak of equine influenza in Japan since 1972 was recently reported and 2007 saw the first ever report of equine influenza in Australia. So far, no fatalities have been reported. Equine influenza is, however, sometimes fatal in young foals.

[00025] Quarantine has been thought to be the best prevention against the spread of equine influenza. South Africa, Australia and Japan have used quarantine of imported horses to stop

the spread of, among other diseases, equine influenza. The quarantine practice has apparently not been fully successful suggesting possible incidental transfer of the disease through human handlers of the horses.

[00026] The influenza virus is highly mutable and, as a result, development of long-term therapies has been difficult. Vaccines generally have needed to be updated as virulent mutants of the virus have arisen. Annual review of worldwide outbreaks of the virus provides data for recommended production of vaccines against the most relevant strains of virus. Significant time elapses between identification of the most relevant strains and commercialization of vaccines.

[00027] There is a need in the art for methods of identifying emerging equine influenza viruses prior to outbreaks so that preventive measures may be taken against such emerging viruses. There is likewise a need in the art for methods of preventing and treating outbreaks caused by virulent strains of EIV including vaccines.

Foot and Mouth Disease

[00028] Foot and Mouth Disease is a highly contagious and sometimes fatal viral disease of cattle, pigs and other animals including bovids with cloven hooves cause by foot and mouth disease virus (FMDV). FMDV is a single-stranded RNA aphthovirus of the Picornaviridae virus family. There are said to be seven different FMDV serotypes: O, A, C, SAT-1, SAT-2, SAT-3, and Asia-1.

[00029] There is a need in the art for methods of predicting increases in virulence of FMDV prior to outbreaks of Foot and Mouth Disease. There is likewise a need in the art for methods of preventing and treating outbreaks of Food and Mouth Disease caused by virulent strains of FMDV.

West Nile Virus

[00030] West nile virus (WNV), in a small percentage of infected humans, causes encephalitis and other serious neuroinvasive diseases. In about four percent of reported cases of WNV infection, the resulting neuroinvasive disease results in death. WNV is flaviviridae virus that was first observed in North America in 1999 and is now considered endemic in the United States. The virus is spread to humans through mosquito (and related insect) bites. Infection with WNV causes diseases such as encephalitis, meningitis and meningoencephalitis in less than about one percent of infected humans. In about 20 percent of infected humans, less severe illness, characterized by fever, headache, tiredness, aches and

sometimes rashes, may occur. Of the total number of U.S. cases of WNV infection reported, about four percent have resulted in death.

[00031] WNV is a single-stranded sense RNA virus and is a member of the Japanese encephalitis virus antigenic complex, which includes several medically important viruses associated with human encephalitis: Japanese encephalitis, St. Louis encephalitis, Murray Valley encephalitis, and Kunjin encephalitis, an Australian subtype of WNV.

[00032] Since introduction of the disease to the United States in 1999, there have been more than 16,000 reported cases of WNV in humans and more than 650 reported deaths. In addition, more than 21,000 cases have been reported in horses. Currently, the only available approved strategies to combat WNV in humans are nationwide active surveillance in conjunction with mosquito control efforts and individual protection with insect repellents. There is a need in the art, therefore, for methods of predicting increases in virulence of WNV prior to epidemics. There is likewise a need in the art for methods of preventing and treating outbreaks caused by virulent strains of WNV.

Viral Diseases in Pigs

[00033] Two severe viral diseases now endemic in swine in many countries and presently causing great economic losses worldwide are Porcine Reproductive and Respiratory Syndrome (PRRS) and Porcine Circovirus Associated Diseases (PCVAD), caused by porcine reproductive and respiratory syndrome virus (PRRSV) and porcine circovirus (PCV), respectively. Each disease has a significant impact on the hog industry and, in both diseases, current control measures are proving inadequate.

[00034] PRRS is a relatively recently recognized disease in pigs. The infectious virus is classified in the family Arteriviridae and order Nidovirales and did not have a standardized name in the past but is now known as porcine reproductive and respiratory syndrome virus (PRRSV). The disease is characterized by reproductive failure, death in young pigs and mild respiratory disease.

[00035] The pig is the only known host for PRRSV but evidence suggests that another host or hosts may have existed prior to identification of PRRS in the United States in 1987 and Europe in 1990. PRRS is now endemic in the United States and many European countries. Evidence of infection (whether serological or virological or both) has been found in Japan, Korea, the Philippines, Vietnam, South America and the Caribbean.

[00036] The disease has been associated with reproductive failure in sows and respiratory disease in all stages of pig development. Clinical signs of the disease include: fever,

anorexia, depression, reduced conception rates, abortion, weak piglets, respiratory distress and increased rates of other endemic diseases.

[00037] PRRSV is a positive-sense single-stranded small envelope RNA virus with at least nine open reading frames (ORFs) in its genome encoding about 20 putative proteins: ORF 1a and 1b encode replication proteins; ORF 2a and 2b encode unknown structure proteins; ORF 3, 4 and 5 encode envelope proteins; ORF 6 encodes membrane proteins and ORF 7 encodes nucleocapsid proteins.

[00038] Two types of PRRSV have been identified: European (Type I) and North American (Type II). The two types share about 60% sequence identity. PRRSV strains are known to differ markedly in pathogenicity. In 2006, highly pathogenic outbreaks of PRRSV occurred in China and Mexico. The cost of PRRSV infection to the U.S. pork industry has been estimated at between \$560 million and \$761 million annually. PRRSV infection has been associated with a reduction in the number of pigs weaned per litter, a reduction in birthing rate, increased mortality, reduced feed conversion and reduced average daily weight gain.

[00039] Porcine Circovirus Associated Diseases (PCVAD) have also only recently been recognized in pigs (1996). PCVAD is a term used to define the entire range of disease associated with porcine circovirus (PCV) infection. The range of disease in pigs includes: Postweaning Multisystemic Wasting Syndrome (PMWS); respiratory illness; pneumonia; diarrhea; reproductive disorders and high mortality. PCVAD symptoms may include detection of PCV within lesions that form on growing pigs, inflammation in, for example, the spleen, thymus, intestines, lymph nodes, lung, kidney, liver, and tonsils, and depletion of lymphoid cells. PCV infection is thought to pose no apparent risk to human health. PCVAD is presently severely affecting the Canadian swine industry.

[00040] Two antigenically distinct types of PCV have been identified. Porcine Circovirus 1 (PCV1), which may be non-pathogenic, and Porcine Circovirus 2 (PCV2), which appears to be the strain that causes PCVAD. PCV1 and PCV2 share about 65% amino acid identity in open reading frame 2 of the virus genome.

[00041] The incidence of PCV infection associated disease has increased by 4% between 2000 and 2006 in Canada and new outbreaks have been observed in Western Canada. In some studies, more than 80% of Canadian pigs have been found to be infected with PCV2 at slaughter. In infected herds, an increase in mortality rates has also been observed. As incidence of PCV infection has increased, pork production has decreased due to pig death and

decreased productivity. Production in Canada in 2006 is expected to decrease 1.5 percent below 2005 production due to PCV-influenced disease.

[00042] There is a need in the art for methods of predicting increases in virulence of PRRSV and PCV prior to outbreaks. There is likewise a need in the art for methods of preventing and treating outbreaks caused by virulent strains of PRRSV and PCV.

WSSV and TSV Shrimp Pathogens

[00043] White spot syndrome virus (WSSV) (also known as white spot baculoform virus) and taura syndrome virus (TSV) are global lethal pathogens in shrimp.

[00044] Taura syndrome is a viral disease in shrimp that significantly impacts the shrimp farming industry worldwide. Taura Syndrome is caused by the taura syndrome virus (TSV), which is a member of the Discistroviridae family in the genus Cripavirus that has a single positive stranded genome of about 10,000 nucleotides. The genome contains two open reading frames (ORF). ORF1 reportedly contains coding for a helicase, a protease and an RNA-dependent RNA polymerase. ORF2 reportedly contains coding for three capsid proteins.

[00045] Taura syndrome is now considered endemic in the Americas and outbreaks have been observed in Asia. Infected shrimp generally have a red tail, are anorexic and erratic in their behavior, tail muscles may become opaque and the cuticle may become soft. Mortality rates between 5% and 95% have been observed during the acute phase of the disease. Shrimp that survive outbreaks of TSV seem to be refractory to reinfection while remaining infectious.

[00046] White spot syndrome (WSS) is a highly contagious and lethal viral infection of shrimp often destroying entire farm populations within several days of observation of the first symptoms. The first reported epidemic of the disease was in Taiwan in 1992 and the disease is now known to be present in all shrimp growing regions globally except Australia. The virus has a wide host range including most cultured penaeid shrimp including *Fenneropenaeus indicus*, *Penaeus monodon*, *Litopenaeus vannamei*, and *Marsupenaeus japonicas*, other non-penaeid shrimp, crabs, spiny lobsters and others.

[00047] WSSV is a rod-shaped double-stranded DNA virus. The complete DNA sequence of WSSV genome has reportedly been assembled into a circular sequence of 292,967 base pairs. Clinical signs of WSSV infection include white spots on the carapace, often reddish discoloration, and reduction in food consumption and loss of energy. There is a need in the

art for methods of preventing and treating viral infections of shrimp such as TSV and WSSV by manipulating the replicating function of Replikin sequences and for identifying molecular targets related to the replicating function of Replikin sequences for treatment of virulent viral.

SUMMARY OF THE INVENTION

[00048] The present invention provides a method of identifying a first virus, first organism or first malignancy with a higher lethality than at least one second virus of the same species as the first virus, second organism of the same species as the first organism or second malignancy of the same species as the first malignancy which comprises comparing the Replikin Count of the Replikin Peak Gene of the first virus, first organism or first malignancy to the Replikin Count of the Replikin Peak Gene of at least one second virus, second organism, or second malignancy to determine that the virus, organism or malignancy with the higher Replikin Count is the more lethal.

[00049] In one embodiment, the first malignancy is a lung malignancy, a brain malignancy, a breast malignancy, an ovarian malignancy or a lymph malignancy. In a specific embodiment, the first malignancy is a non-small cell lung carcinoma.

[00050] In another embodiment, the first organism is a *Mycobacterium tuberculosis*, *Mycobacterium mucogenicum*, *Staphylococcus aureus*, or *Plasmodium falciparum*.

[00051] In a further embodiment, the virus is influenza virus, foot and mouth disease virus, west Nile virus, porcine respiratory and reproductive syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus, hemorrhagic septicemia virus, or tobacco mosaic virus.

[00052] In a specific embodiment, the first virus is a strain of *Influenza A* virus of H1N1, H2N2, H3N2, H5N1, or H3N8.

[00053] In a further embodiment, said at least one Replikin sequence within the protein or protein fragment of the identified Replikin Peak Gene is isolated from influenza A strain H5N1 and is selected from the group consisting of SEQ ID NOS:1685-1691, SEQ ID NOS:1702-1717. In a further embodiment, said at least one Replikin sequence within the protein or protein fragment of the identified Replikin Peak Gene is isolated from equine influenza virus (H3N8) and is selected from the group consisting of SEQ ID NOS:547-562.

[00054] The present invention further provides a isolated or synthesized Replikin Peak Gene of a virus, organism or malignancy wherein said Replikin Peak Gene is identified as the portion of the genome, protein or protein fragment of a virion of the virus, a cell of the organism or a malignant cell of the malignancy consisting of the highest number of

continuous Replikin sequences per 100 amino acids as compared to other portions of the genome, protein or protein fragment of the virion of the virus, the cell of the organism or the malignant cell of the malignancy.

[00055] In one embodiment, the isolated or synthesized Replikin Peak Gene is the portion of a protein or protein fragment consisting of the highest number of continuous Replikin sequences per 100 amino acids as compared to all other proteins or protein fragments in the virion of the virus, in the cell of the organism or in the malignant cell of the malignancy.

[00056] In a specific embodiment, the isolated or synthesized Replikin Peak Gene is isolated from a lung malignancy, a brain malignancy, a breast malignancy, an ovarian malignancy, or a lymph malignancy. In another specific embodiment, the isolated or synthesized Replikin Peak Gene is isolated from a non-small cell lung carcinoma or glioblastoma multiforme.

[00057] In yet another embodiment, the isolated or synthesized Replikin Peak Gene of is isolated from *Mycobacterium tuberculosis*, *Mycobacterium mucogenicum*, *Staphylococcus aureus*, or *Plasmodium falciparum*.

[00058] According to a further embodiment, the isolated or synthesized Replikin Peak Gene is isolated from influenza virus, foot and mouth disease virus, west nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus, hemorrhagic septicemia virus, or tobacco mosaic virus.

[00059] In one embodiment, the isolated or synthesized Replikin Peak Gene is from influenza virus, particularly an *Influenza A* virus. In a specific embodiment, the *Influenza A* virus is a strain H1N1, H2N2, H3N2, H5N1 or H3N8. In another specific embodiment, the Replikin Peak Gene is isolated from the pB1 gene area of an influenza virus.

[00060] According to another embodiment, the isolated or synthesized Replikin Peak Gene is from foot and mouth disease virus. In a specific embodiment, the isolated or synthesized Replikin Peak Gene is identified within the VP1 gene of a foot and mouth disease virus.

[00061] In yet another embodiment, the isolated or synthesized Replikin Peak Gene is from a west nile virus. In a specific embodiment, the isolated or synthesized Replikin Peak Gene is isolated from the envelope protein of west nile virus.

[00062] In a further embodiment, the isolated or synthesized Replikin Peak Gene is from a porcine respiratory and reproductive syndrome virus. In a specific embodiment, the isolated

or synthesized Replikin Peak Gene is isolated from a nucleocapsid protein of a porcine respiratory and reproductive syndrome virus.

[00063] In yet another embodiment, the isolated or synthesized Replikin Peak Gene is from a porcine circovirus. In a specific embodiment, the isolated or synthesized Replikin Peak Gene is isolated from a replicase protein of a porcine circovirus.

[00064] In still a further embodiment, the isolated or synthesized Replikin Peak Gene is from a white spot syndrome virus. In a specific embodiment, the isolated or synthesized Replikin Peak Gene is isolated from a ribonucleotide reductase protein of a white spot syndrome virus.

[00065] In yet another embodiment, the isolated or synthesized Replikin Peak Gene is from a tobacco mosaic virus.

[00066] In a further embodiment, the isolated or synthesized Replikin Peak Gene is from a hemorrhagic septicemia virus in fish. In a specific embodiment, the isolated or synthesized Replikin Peak Gene is isolated from a glycoprotein in a hemorrhagic septicemia virus.

[00067] In another specific embodiment, the isolated or synthesized Replikin Peak Gene comprises a sequence of SEQ ID NO:1741, SEQ ID NO: 3664, SEQ ID NO:3660, SEQ ID NO:3665, SEQ ID NO:1996, SEQ ID NO:1665, SEQ ID NO:1684, SEQ ID NO:1701, SEQ ID NO:546, SEQ ID NO:124, SEQ ID NO:130, SEQ ID NO:311, SEQ ID NOS:341-344, SEQ ID NO:286, SEQ ID NO:287, SEQ ID NO:288, SEQ ID NO:289, SEQ ID NO:290, SEQ ID NOS:233-238, SEQ ID NO:415, SEQ ID NO:421, SEQ ID NO:438, SEQ ID NO:451, SEQ ID NO:462, SEQ ID NO:498, SEQ ID NO:669, SEQ ID NO:1168, SEQ ID NO: 1531, SEQ ID NO:1548, or SEQ ID NO:1939.

[00068] The present invention further provides an immunogenic composition comprising the isolated or synthesized Replikin Peak Gene. In a specific embodiment, the immunogenic composition comprises a Replikin sequence of SEQ ID NOS:2902-2925, SEQ ID NOS: 2312-2544, SEQ ID NOS:2701-2711, SEQ ID NOS: 2713-2718, SEQ ID NOS:3282-3285, 3287-3291, 3293, 3295, 3297, 3299, 3300, 3302, 3304, 3306, and 3308, SEQ ID NOS:1685-1691, SEQ ID NOS:1702-1717, SEQ ID NO:106, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NOS:125-129, SEQ ID NOS:131-156, SEQ ID NOS:233-244, SEQ ID NOS:286-290, SEQ ID NOS:312-323, SEQ ID NOS:354-366, SEQ ID NOS:368-380, SEQ ID NOS:383-393, SEQ ID NOS:395-401, SEQ ID NOS:403-414, SEQ ID NOS:291-307, SEQ ID NOS:308-310, SEQ ID NOS:324-327, SEQ ID NOS:328-340, SEQ ID NOS:416-419, SEQ ID NOS:422-437, SEQ ID NOS: 440-445, SEQ ID NOS:452-457, SEQ ID NOS:464-476,

SEQ ID NOS:482-484 and SEQ ID NOS:487-492, SEQ ID NOS:547-562, SEQ ID NOS:663-667, SEQ ID NOS:670-1166, SEQ ID NOS:1169-1529, SEQ ID NOS:1532-1542, SEQ ID NO:1548, SEQ ID NOS:3788-3823), or SEQ ID NOS:1637-1663.

[00069] A non-limiting embodiment of the present invention provides computer readable medium having stored thereon instructions which, when executed, cause the processor to perform a method for identifying a Replikin Peak Gene of a virus, organism or malignancy comprising identifying, within amino acid sequences or nucleic acid sequences that encode amino acid sequences of said virus, organism or malignancy, the portion of the genome, or protein or protein fragment of said virus, said organism or said malignancy consisting of the highest number of continuous Replikin sequences per 100 amino acids as compared to other portions of the genome, or protein or protein fragment of the malignancy, organism or virus.

[00070] In one embodiment, the computer readable medium comprises instructions which, when executed, cause the processor to perform a method for predicting an increase in lethality or virulence of said virus, organism or malignancy that comprises said identified Replikin Peak Gene or an outbreak of said virus or organism that comprises said identified Replikin Peak Gene by: (1) determining that the Replikin Count of said Replikin Peak Gene or that the Replikin Count of a protein or gene area comprising said Replikin Peak Gene is higher than another Replikin Peak Gene or a protein or gene area comprising said other Replikin Peak Gene identified within the genome or within a protein or protein fragment of at least one other virus of the same species as said virus, at least one other organism of the same species as said organism or at least one other malignancy of the same type as said malignancy wherein said other virus, said other organism or said other malignancy is isolated at an earlier time point than said virus, said organism or said malignancy, and (2) predicting an increase in lethality or virulence of said virus, organism or malignancy or predicting an outbreak of said virus or organism.

[00071] The invention also provides a method of predicting the strain, the host or the geographic region of an outbreak or increase in lethality or virulence of a virus or organism by (1) identifying a Replikin Peak Gene or a protein or gene area comprising a Replikin Peak Gene within the genome of a first virus or organism of a first strain, from a first host, or isolated from a first geographic region or within a protein or protein fragment of the first virus or organism that has a higher Replikin Count than a Replikin Peak Gene or protein or gene area comprising a Replikin Peak Gene identified within the genome or within a protein or protein fragment of at least one second virus of the same species as the first virus or at

least one second organism of the same species as the first organism wherein said first virus or said first organism is isolated at a later time point than said first virus or said first organism and is the same strain, from the same or another host or isolated from the same or another geographic region as the first virus or first organism, and (2) predicting an outbreak or an increase in lethality or virulence of said first strain, in said first host or within said first geographic region of said first virus or organism.

[00072] In one embodiment, the protein or gene area comprising said Replikin Peak Gene within the genome of a first virus or organism is identified as having a higher Replikin Count than said protein or gene area comprising a Replikin Peak Gene identified within the genome or within a protein or protein fragment of said at least one second virus or organism.

[00073] In another embodiment, the first virus or first organism is isolated at least six months to three years later than the second organism or said second virus. In a specific embodiment, the first organism or first virus is *Mycobacterium tuberculosis*, *Mycobacterium mucogenicum*, *Staphylococcus aureus*, and *Plasmodium falciparum*, influenza virus, foot and mouth disease virus, west nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus in tomato plants, hemorrhagic septicemia virus, or tobacco mosaic virus. In another embodiment, the *Staphylococcus aureus* is methicillin-resistant.

[00074] In a further embodiment, the influenza virus is a strain of *Influenza A* virus. In a specific embodiment, the first virus is an influenza virus of the strain H1N1, H2N2, H3N2, H5N1 or H3N8.

[00075] In a further embodiment of the invention, the protein or gene area comprising the Replikin Peak Gene is the pB1 gene area of the influenza virus.

[00076] In yet another embodiment, the protein or gene area is a nucleocapsid protein of porcine respiratory and reproductive syndrome virus.

[00077] In a further embodiment, the protein or gene area is an envelope protein of west nile virus.

[00078] In a further embodiment, the protein or gene area is a VP1 protein of foot and mouth disease virus.

[00079] In still another embodiment, the protein or gene area is an ATP-ase of *Plasmodium falciparum*.

[00080] In yet a further embodiment, the protein or gene area is a replicase protein of porcine circovirus.

[00081] In another embodiment, the protein or gene area is a ribonucleotidase of said white spot syndrome virus.

[00082] The present invention further provides a method of identifying a first virus, organism or malignancy associated with higher lethality, higher virulence or more rapid replication than a second virus of the same species as the first virus, a second organism of the same species as the first organism or a second malignancy of the same type as the first malignancy comprising identifying a Replikin Peak Gene encoded within the genome of at least one virion of the first virus, or at least one cell of the first organism, or at least one malignant cell of the first malignancy, or within a protein or protein fragment of at least one virion of the first virus, or at least one cell of the first organism, or at least one malignant cell of the first malignancy that has a higher Replikin Count than a Replikin Peak Gene identified encoded within the genome of at least one virion of the second virus, or at least one cell of the second organism, or at least one malignant cell of the second malignancy or within a protein or protein fragment of at least one virion of the second virus, or at least one cell of the second organism, or at least one malignant cell of the second malignancy wherein said first virus, first organism or first malignancy has higher lethality, higher virulence or more rapid replication than said second virus, second organism or second malignancy, and wherein the Replikin Peak Gene is defined as a protein or protein fragment having the highest concentration of continuous Replikin sequences per 100 amino acids as compared to the remaining proteins or protein fragments in the same virion of the virus, the same cell of the organism, or the same malignant cell, or the portion of the genome encoding the protein or protein fragment.

[00083] Further provided is a method of identifying a first virus, first organism or first malignancy with a higher lethality than at least one second virus of the same species as the first virus, second organism of the same species as the first organism or second malignancy of the same species as the first malignancy comprising comparing the Replikin Count of the whole genome of a virus, organism or malignancy to the Replikin Count of the whole genome of at least one second virus, second organism, or second malignancy to determine that the virus, organism or malignancy with the higher Replikin Count is the more lethal.

[00084] According to a specific embodiment, the first virus is a coronavirus, a foot and mouth disease virus, a white spot syndrome virus, a taura syndrome virus, a porcine circovirus, or an influenza virus.

[00085] In one specific embodiment, the first virus is an H5N1 strain of influenza virus.

[00086] In another specific embodiment, the influenza virus is an *Influenza A* virus. In a further specific embodiment, the *Influenza A* virus is H1N1, H2N2, H3N2, H5N1 or H3N8.

[00087] According to another embodiment, the Replikin Peak Gene is isolated from the pB1 gene area of an influenza virus.

[00088] The present invention also provides method for obtaining an isolated or synthesized Replikin Peak Gene of a virus, organism or malignancy for diagnosis, prevention or treatment of an infection of said virus or said organism or for diagnosis, prevention or treatment of said malignancy comprising: (1) obtaining a plurality of isolates of virus of the same species, a plurality of organisms of the same species, or a plurality of malignancies of the same type; (2) analyzing the protein sequences or protein sequence fragments of each individual isolate of the plurality of isolates of virus, a cell of each individual organism of the plurality of organisms, or a malignant cell of each individual malignancy of the plurality of malignancies for the presence and concentration of Replikin sequences; (3) identifying the protein sequence or the protein sequence fragment having the highest concentration of continuous Replikin sequences in the malignant cell of each individual malignancy, the cell of each individual organism or each individual virus isolate; (4) selecting the protein sequence or protein sequence fragment having the highest concentration of continuous Replikin sequences among the plurality of isolates of virus, the plurality of organisms, or the plurality of malignancies; (5) identifying the amino acid sequence of the selected protein sequence or protein sequence fragment as the Replikin Peak Gene of the plurality of virus isolates, organisms or malignancies; and (6) isolating or synthesizing the identified Replikin Peak Gene of at least one of the plurality of virus isolates, organisms or malignancies wherein the isolated or synthesized identified Replikin Peak Gene is useful for diagnosis, prevention or treatment of said infection of said virus or said organism or said malignancy.

[00089] Further provided is an immunogenic composition comprising at least one isolated or synthesized Replikin Peak Gene isolated according to the above method. In a specific embodiment, the immunogenic composition is isolated from an emerging strain of a virus or organism, and optionally further comprises a pharmaceutically acceptable carrier.

[00090] The present invention also provides a vaccine comprising at least one isolated or synthesized Replikin Peak Gene. In a specific embodiment, the vaccine comprises a Replikin Peak Gene isolated from an emerging strain of virus or organism. In another specific embodiment, the vaccine comprises SEQ ID NO:1741, SEQ ID NO: 3664, SEQ ID NO:3660, SEQ ID NO:3665, SEQ ID NO:1996, SEQ ID NO:1665, SEQ ID NO:1684, SEQ ID

NO:1701, SEQ ID NO:546, SEQ ID NO:124, SEQ ID NO:130, SEQ ID NO:311, SEQ ID NOS:341-344, SEQ ID NO:286, SEQ ID NO:287, SEQ ID NO:288, SEQ ID NO:289, SEQ ID NO:290, SEQ ID NOS:233-238, SEQ ID NO:415, SEQ ID NO:421, SEQ ID NO:438, SEQ ID NO:451, SEQ ID NO:462, SEQ ID NO:498, SEQ ID NO:669, SEQ ID NO:1168, SEQ ID NO: 1531, SEQ ID NO:1548, positions 81-204 of SEQ ID NO:3787, or SEQ ID NO:1939.

[00091] In yet a further embodiment, the vaccine comprises a Replikin Peak Gene isolated from a virus.

[00092] In a specific embodiment, the virus is influenza virus, foot and mouth disease virus, west nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus, hemorrhagic septicemia virus, or tobacco mosaic virus.

[00093] In one embodiment, the Replikin Peak Gene in the vaccine is isolated from *Influenza A*, or specifically strains H1N1, H2N2, H3N2, H5N1 or H3N8.

[00094] In another embodiment, the vaccine comprises a Replikin Peak Gene isolated from an organism.

[00095] In a further embodiment, Replikin Peak Gene is isolated from *Mycobacterium tuberculosis*, *Mycobacterium mucogenicum*, *Staphylococcus aureus*, or *Plasmodium falciparum*. In a specific embodiment, the *Staphylococcus aureus* is methicillin-resistant.

[00096] In still another embodiment, the Replikin Peak Gene is isolated from a malignancy.

[00097] In a specific embodiment, the Replikin Peak Gene is isolated from a lung malignancy, a brain malignancy, a breast malignancy or a lymph malignancy. In another embodiment, the Replikin Peak Gene is isolated from a non-small cell lung carcinoma. In a further embodiment, the Replikin Peak Gene is isolated from glioblastoma multiforme.

[00098] The present invention further provides an immunogenic composition comprising a Replikin Peak Gene, optionally in combination with a pharmaceutically acceptable carrier. In one embodiment, the immunogenic composition comprises SEQ ID NO:1741, SEQ ID NO: 3664, SEQ ID NO:3660, SEQ ID NO:3665, SEQ ID NO:1996, SEQ ID NO:1665, SEQ ID NO:1684, SEQ ID NO:1701, SEQ ID NO:546, SEQ ID NO:124, SEQ ID NO:130, SEQ ID NO:311, SEQ ID NOS:341-344, SEQ ID NO:286, SEQ ID NO:287, SEQ ID NO:288, SEQ ID NO:289, SEQ ID NO:290, SEQ ID NOS:233-238, SEQ ID NO:415, SEQ ID NO:421,

SEQ ID NO:438, SEQ ID NO:451, SEQ ID NO:462, SEQ ID NO:498, SEQ ID NO:669, SEQ ID NO:1168, SEQ ID NO: 1531, SEQ ID NO:1548, or SEQ ID NO:1939.

[00099] The present invention further provides an isolated or synthesized Replikin sequence isolated from a protein or protein fragment a Replikin Peak Gene or isolated from a protein comprising a Replikin Peak Gene.

[000100] In one embodiment, the Replikin sequence is from a Replikin Peak Gene isolated from *Mycobacterium tuberculosis*, *Mycobacterium mucogenicum*, *Staphylococcus aureus*, or a *Plasmodium falciparum*. In a specific embodiment, the Replikin sequence is from a Replikin Peak Gene isolated from *Mycobacterium mucogenicum*. In a further embodiment, the the Replikin Peak Gene is SEQ ID NOS:2902-2925. In another specific embodiment, the Replikin sequence is from a Replikin Peak Gene isolated from *Plasmodium falciparum*. In a further embodiment, the Replikin Peak Gene is one of SEQ ID NOS: 2312-2544, SEQ ID NOS:2701-2711, SEQ ID NOS: 2713-2718, SEQ ID NOS: 3282-3285, 3287-3291, 3293, 3295, 3297, 3299, 3300, 3302, 3304, 3306, or SEQ ID NO: 3308.

[000101] In another embodiment, the Replikin sequence is from a Replikin Peak Gene isolated from influenza virus, foot and mouth disease virus, west nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus, hemorrhagic septicemia virus, or tobacco mosaic virus.

[000102] In a specific embodiment, the influenza virus is *Influenza A* virus. In another specific embodiment, the *Influenza A* virus is H1N1, H2N2, H3N2, H5N1 or H3N8. In a further specific embodiment, the *Influenza A* virus is H5N1 and the Replikin sequence is one of SEQ ID NOS:1685-1691, SEQ ID NOS:1702-1716 or SEQ ID NO:1717. In a further specific embodiment, the *Influenza A* virus is H3N8 and the Replikin sequence is one of SEQ ID NOS:547-561 or SEQ ID NO:562.

[000103] In another embodiment, the Replikin sequence is from a Replikin Peak Gene isolated from foot and mouth disease virus. In a specific embodiment, the Replikin sequence from the foot and mouth disease virus is one of SEQ ID NO:106, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NOS:125-129, SEQ ID NOS:131-155 or SEQ ID NO:156.

[000104] In still another embodiment, the Replikin sequence is from a Replikin Peak Gene isolated from west nile virus. In a specific embodiment, the Replikin sequence from the west nile virus is one of SEQ ID NOS:233-243 or SEQ ID NO:244.

[000105] In a further embodiment, the Replikin sequence is from a Replikin Peak Gene isolated from porcine reproductive and respiratory virus. In a specific embodiment, the Replikin sequence from porcine reproductive and respiratory virus is one of SEQ ID NOS:286-290, SEQ ID NOS:312-323, SEQ ID NOS:354-366, SEQ ID NOS:368-380, SEQ ID NOS:383-393, SEQ ID NOS:395-401, SEQ ID NOS:403-413 or SEQ ID NO:414.

[000106] In another embodiment, the Replikin sequence is from a Replikin Peak Gene isolated from porcine circovirus. In a specific embodiment, the Replikin sequence from porcine circovirus is one of SEQ ID NOS:291-307, SEQ ID NOS:308-310, SEQ ID NOS:324-327, SEQ ID NOS:328-340, SEQ ID NOS:416-419, SEQ ID NOS:422-437, SEQ ID NOS: 440-445, SEQ ID NOS:452-457, SEQ ID NOS:464-476, SEQ ID NOS:482-484, SEQ ID NOS:487-491 or SEQ ID NO:492.

[000107] In still a further embodiment, the Replikin sequence is from a Replikin Peak Gene isolated from white spot syndrome virus. In a specific embodiment, the Replikin sequence from white spot syndrome virus is one of SEQ ID NOS:663-667, SEQ ID NOS:670-1166, SEQ ID NOS:1169-1529, SEQ ID NOS:1532-1542 and SEQ ID NO:1548.

[000108] According to the present invention provided is a vaccine for prevention and/or treatment of an viral or organismal infection or a malignancy wherein the vaccine comprises at least one isolated or synthesized Replikin sequence within a protein or protein fragment of a Replikin Peak Gene or a protein comprising a Replikin Peak Gene identified in said virus, organism, or malignancy.

[000109] In a further embodiment, the at least one isolate or synthesized Replikin sequence in the vaccine is one of SEQ ID NOS:2902-2925, SEQ ID NOS: 2312-2544, SEQ ID NOS:2701-2711, 2713-2718, SEQ ID NOS: 3282-3285, 3287-3291, 3293, 3295, 3297, 3299, 3300, 3302, 3304, 3306, 3308, SEQ ID NOS:1685-1691, SEQ ID NOS:1702-1717, SEQ ID NOS:547-562, SEQ ID NO:106, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NOS:125-129, and SEQ ID NOS:131-156, SEQ ID NOS:233-244, SEQ ID NOS:286-290, SEQ ID NOS:312-323, SEQ ID NOS:354-366, SEQ ID NOS:368-380, SEQ ID NOS:383-393, SEQ ID NOS:395-401, SEQ ID NOS:403-414, SEQ ID NOS:291-307, SEQ ID NOS:308-310, SEQ ID NOS:324-327, SEQ ID NOS:328-340, SEQ ID NOS:416-419, SEQ ID NOS:422-437, SEQ ID NOS: 440-445, SEQ ID NOS:452-457, SEQ ID NOS:464-476, SEQ ID NOS:482-484 SEQ ID NOS:487-492, SEQ ID NOS:663-667, SEQ ID NOS:670-1166, SEQ ID NOS:1169-1529, SEQ ID NOS:1532-1542, SEQ ID NO:1548, SEQ ID NOS:1637-1662, or SEQ ID NO:1663.

[000110] In one embodiment, the vaccine is for prevention and/or treatment of a viral infection. In a specific embodiment, the vaccine is for a viral infection is caused by influenza virus, foot and mouth disease virus, west nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus, hemorrhagic septicemia virus, or tobacco mosaic virus.

[000111] In another specific embodiment, the influenza virus is *Influenza A* virus. In a further specific embodiment, the *Influenza A* virus is H1N1, H2N2, H3N2, H5N1 or H3N8 *Influenza A* virus.

[000112] In another specific embodiment, the virus is hemorrhagic septicemia virus.

[000113] In another embodiment, the vaccine is for prevention and/or treatment of an organismal infection.

[000114] In one specific embodiment, the organismal infection is caused by *Mycobacterium mucogenicum*, *Mycobacterium tuberculosis*, *Staphylococcus aureus*, or *Plasmodium falciparum*. In a further specific embodiment, the *Staphylococcus aureus* is methicillin-resistant.

[000115] In another embodiment, the vaccine is for prevention of a malignancy.

[000116] In one specific embodiment, the malignancy is a lung malignancy, a brain malignancy, a breast malignancy, an ovarian malignancy, or a lymph malignancy. In a further specific embodiment, the malignancy is non-small cell lung carcinoma or glioblastoma multiforme.

[000117] The invention also provides an immunogenic compound comprising at least one isolated or synthesized Replikin sequence within the protein or protein fragment of a Replikin Peak Gene or within a protein comprising a Replikin Peak Gene wherein said Replikin Peak Gene is identified in a virus, an organism or a malignancy, optionally further comprising a pharmaceutically acceptable carrier.

[000118] In another aspect, the present invention provides a method of stimulating the immune system, comprising administering in an animal at least one isolated or synthesized Replikin sequence identified within a protein or protein fragment of a Replikin Peak Gene or within a protein or gene area comprising a Replikin Peak Gene identified in a virus, organism, or malignancy. In a specific embodiment, the animal is a human.

[000119] The invention further provides an antibody to at least one isolated or synthesized Replikin sequence within a protein or protein fragment of Replikin Peak Gene or within protein or gene area comprising a Replikin Peak Gene.

[000120] Also provided by the present invention is a method of identifying a lethal strain of malignancy, organism or virus comprising: (1) obtaining a plurality of isolates of said malignancy, organism or virus; (2) identifying the Replikin Peak Gene in each isolate of the plurality of isolates of said malignancy, organism or virus; (3) analyzing the amino acid sequence of a protein or protein fragment of the Replikin Peak Gene of each isolate of the plurality of isolates for the presence and concentration of Replikin sequences; (4) comparing the concentrations of Replikin sequences in each of the proteins or protein fragments of the Replikin Peak Gene of each isolate of the plurality of isolates to the concentration of Replikin sequences in each of the proteins or protein fragments of the Replikin Peak Gene of each of the other isolates of the plurality of isolates; and (5) identifying the isolate having the highest concentration of continuous Replikin sequences in the protein or protein fragment of the Replikin Peak Gene as a virulent or lethal strain of said malignancy, organism or virus.

[000121] Further provided is a method of selecting a peptide from a malignancy, organism or virus for inclusion in a preventive or therapeutic vaccine or immunogenic compound for a malignancy, organism or virus comprising identifying at least one difference in the amino acid sequence of an otherwise conserved Replikin sequence or Replikin Peak Gene between at least two isolates of said malignancy, organism or virus and correlating the identified at least one difference in the amino acid sequence with the highest virulence, morbidity or host mortality among the at least two isolates and selecting an otherwise conserved Replikin sequence, Replikin Peak Gene or Replikin sequence within a Replikin Peak Gene having the identified at least one amino acid sequence difference as the peptide for inclusion in a preventive or therapeutic vaccine or immunogenic compound.

[000122] In one embodiment, the method further comprises predicting the isolate comprising the selected conserved Replikin sequence or Replikin Peak Gene having the at least one difference in the amino acid sequence to be lethal isolate of said malignancy, organism or virus.

[000123] In a specific embodiment, the malignancy, organism or virus is a malignancy.

[000124] In another specific embodiment, the malignancy is a lung malignancy, a brain malignancy, a breast malignancy or a lymph malignancy. In a further specific embodiment, the malignancy is a non-small cell lung carcinoma or a glioblastoma multiforme.

- [000125] In another aspect, the malignancy, organism or virus is an organism.
- [000126] In a first specific embodiment, the organism is *Mycobacterium tuberculosis*, *Mycobacterium mucogenicum*, *Staphylococcus aureus*, or *Plasmodium falciparum*. In another specific embodiment, the *Staphylococcus aureus* is methicillin-resistant.
- [000127] In another aspect, the malignancy, organism or virus is a virus.
- [000128] In one specific embodiment, the virus is influenza virus, foot and mouth disease virus, west nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus, hemorrhagic septicemia virus or tobacco mosaic virus.
- [000129] The invention further provides a method of determining a source of a case of lung malignancy comprising identifying at least one peptide in a Replikin Peak Gene of a lung cancer cell that is also present in a Replikin Peak Gene of an isolate of tobacco mosaic virus, wherein the peptide is involved with the source of the lung malignancy.
- [000130] In one embodiment, a plurality of peptides is identified in the Replikin Peak Gene of the lung cancer cell wherein each one of the plurality of peptides is also identified in the Replikin Peak Gene of an isolate of tobacco mosaic virus.
- [000131] In another embodiment, the the at least one peptide in the Replikin Peak Gene of the lung cancer cell and the at least one peptide in Replikin Peak Gene of the isolate of tobacco mosaic virus is a peptide of about 10 amino acids or less comprising at least two lysines and at least one histidine.
- [000132] In a further embodiment, the at least one peptide in the Replikin Peak Gene of the lung cancer cell and the at least one peptide in Replikin Peak Gene of the isolate of tobacco mosaic virus is a peptide of about 10 amino acids or less comprising at least three lysines and at least one histidine.
- [000133] In yet another embodiment, the the at least one peptide in the Replikin Peak Gene of the lung cancer cell and the at least one peptide in the Replikin Peak Gene of the isolate of tobacco mosaic virus is about 7 amino acids or less comprising at least three lysines and at least one histidine.
- [000134] In a further embodiment, the at least one peptide in the Replikin Peak Gene of the lung cancer cell and the at least one peptide in the Replikin Peak Gene of the isolate of tobacco mosaic virus is about 4 amino acids comprising three lysines and one histidine.

[000135] In a specific embodiment, the the at least one peptide in the Replikin Peak Gene of the lung cancer cell and in the Replikin Peak Gene of the isolate of tobacco mosaic virus is KHKK (SEQ ID NO:1584).

[000136] In another embodiment, the more than one KHKK (SEQ ID NO:1584) peptide is identified in the Replikin Peak Gene of the lung cancer cell and in the Replikin Peak Gene of the isolate of tobacco mosaic virus.

[000137] In one specific embodiment, at least 10 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the lung cancer cell and at least 10 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the isolate of tobacco mosaic virus.

[000138] In another specific embodiment, at least 20 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the lung cancer cell and at least 20 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the isolate of tobacco mosaic virus.

[000139] In a third specific embodiment, at least 30 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the lung cancer cell and at least 30 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the isolate of tobacco mosaic virus.

[000140] In a fourth specific embodiment, at least 50 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the lung cancer cell and at least 50 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the isolate of tobacco mosaic virus.

[000141] The present invention further provides a method of identifying a first case of malignancy of the lung having a higher rate of replication, aggressive growth pattern or lethality as compared with a second case of malignancy of the lung comprising identifying a Replikin Peak Gene in a malignant cell from a first case of malignancy of the lung that has a higher Replikin Count in the Replikin Peak Gene than a Replikin Peak Gene identified in a malignant cell from a second case of malignancy of the lung.

[000142] In one embodiment, first and second cases of malignancy of the lung are non-small cell lung malignancies.

[000143] Further provided is an isolated or synthesized Replikin Peak Gene in a lung malignancy for diagnosis, prevention or treatment of lung cancer by the method comprising: (1) obtaining at least one malignant cell from a lung malignancy; (2) analyzing the protein

sequences or protein sequence fragments of the at least one malignant cell for the presence and concentration of Replikin sequences; (3) identifying the protein sequence or the protein sequence fragment having the highest concentration of continuous Replikin sequences in the at least one malignant cell; (4) selecting the protein sequence or protein sequence fragment having the highest concentration of continuous Replikin sequences; (5) identifying the amino acid sequence of the selected protein sequence or protein sequence fragment as the Replikin Peak Gene; and (6) isolating or synthesizing the identified Replikin Peak Gene of the at least one malignant cell, wherein the isolated or synthesized identified Replikin Peak Gene is useful for diagnosis, prevention or treatment of lung cancer.

In one aspect, the lung malignancy is a non-small cell lung malignancy.

[000144] In another aspect, at least one isolated or synthesized Replikin sequence within the protein or protein fragment of the identified Replikin Peak Gene for diagnosis, prevention or treatment of lung cancer.

[000145] In a specific embodiment, the at least one isolated or synthesized Replikin sequence within the protein or protein fragment of the identified Replikin Peak Gene is one of SEQ ID NOS: 1585-1635 of SEQ ID NO:1636.

[000146] The invention also provides an immunogenic composition for prevention and treatment of lung cancer, wherein the immunogenic composition comprises at least one isolated or synthesized Replikin sequence within the protein or protein fragment of the identified Replikin Peak Gene.

[000147] Also provided is method of stimulating the immune system, comprising administering in an animal the at least one isolated or synthesized Replikin sequence identified within the Replikin Peak Gene of the lung malignancy for prevention, treatment or diagnosis of lung cancer in an animal. In a specific embodiment, the animal is a human.

[000148] In another embodiment, the present invention provides a method of identification of a lethal form of lung cancer comprising: (1) obtaining at least one malignant cell from a plurality of lung tumors; (2) identifying the Replikin Peak Gene in the at least one malignant cell of each of the plurality of lung tumors; (3) analyzing the amino acid sequence of a protein or protein fragment of the Replikin Peak Gene in the at least one malignant cell of each of the plurality of lung tumors for the presence and concentration of Replikin sequences; (4) comparing the concentrations of Replikin sequences in each of the proteins or protein fragments of the Replikin Peak Gene in the at least one malignant cell of each of the plurality of lung tumors; and (5) identifying the lung tumor having the highest concentration

of continuous Replikin sequences in the protein or protein fragment of the Replikin Peak Gene as a lethal form of lung cancer.

[000149] In a further embodiment, the present invention provides a method of identification of a more lethal form of lung cancer among at least two lung cancers, comprising: (1) obtaining at least one malignant cell from each of at least two lung cancers; (2) identifying the Replikin Peak Gene in the at least one malignant cell of each of the at least two lung cancers; (3) analyzing the amino acid sequence of a protein or protein fragment of the Replikin Peak Gene in the at least one malignant cell of each of the at least two lung cancers for the presence and concentration of Replikin sequences; (4) comparing the concentrations of Replikin sequences in each of the proteins or protein fragments of the Replikin Peak Gene in the at least one malignant cell of each of the at least two lung cancers; and (5) identifying the lung cancer having the highest concentration of continuous Replikin sequences in the protein or protein fragment of the Replikin Peak Gene as the more lethal form of lung cancer.

[000150] The invention further provides a method of determining an expected increase in lethality or virulence of a virus or organism which method comprises: (1) obtaining a plurality of isolates of said virus or organism wherein each isolate is isolated within a known time period and wherein at least two of said isolates is isolated about six months to about 5 years later than at least two other of said isolates; (2) identifying a Replikin Peak Gene in each isolate of said plurality of isolates; (3) analyzing the identified Replikin Peak Gene of each isolate of the plurality of isolates to determine the Replikin Count of each Replikin Peak Gene of each isolate of the plurality of isolates, or analyzing a protein, protein fragment, or gene area comprising the identified Replikin Peak Gene of each isolate of the plurality of isolates to determine the Replikin Count of the protein, protein fragment, or gene area of the plurality of isolates; (4) determining a mean Replikin Count within the Replikin Peak Gene or within the protein, protein fragment, or gene area comprising said identified Replikin Peak Gene for each known time period; (5) comparing the mean Replikin Count within the Replikin Peak Gene or within the protein, protein fragment, or gene area for each known time period one to another; (6) identifying an increase in the mean Replikin Count between at least two known time periods; and (7) identifying an expected increase in lethality or virulence of said virus, or organism within about six months to about three years following said identified increase in the mean Replikin Count.

[000151] In one specific embodiment, the known time period is about 1 year. In another specific embodiment, the increase in mean Replikin Count occurs over one year. In a further specific embodiment, the increase in mean Replikin Count occurs over three years. In another embodiment, the increase in mean Replikin Count is significant between at least two known time periods. In a further embodiment, the increase in mean Replikin Count has a significance of $p = < 0.001$.

BRIEF DESCRIPTION OF THE DRAWINGS

[000152] Figure 1 illustrates the localization of the pB1 gene area as the Replikin Peak Gene in the genome of the H5N1 strain of influenza virus. Replikin Peak Genes are the places in the genome where Replikin sequences are continuous and most concentrated. The pB1 gene area comprises a Replikin Peak Gene in the H5N1 genome and the Replikin Count of the pB1 gene area correlates increases in virulence and mortality. Dark gray columns represent mean Replikin Counts for designated gene areas in isolates of H5N1 virus isolated during the given year. Light gray columns represent standard deviation from the mean in the population of isolates in a given year. Standard deviation of the means is shown in light gray columns on top of the means, rather than in the usual 'T' symbols. This style is used to emphasize the diverse expanding virus population with regard to the Replikin Count. Replikin Counts for isolates of H5N1 virus isolated in years 2003 through 2006 with genetic information publicly available at pubmed.com were determined separately by analysis of the number of Replikin sequences observed in each of the eight genome areas of human H5N1 influenza virus for isolates in a given year. The eight genome areas that have been identified are nucleocapsid, matrix, pB2, neuraminidase, pA, NS, hemagglutinin, and pB1 gene areas.

[000153] Figure 2 illustrates an increase in Replikin Count before and accompanying each influenza A pandemic and outbreak since 1918 and low Replikin Counts during quiescent periods of influenza A infection and continually in non-lethal *Influenza B*. The graph provides annual Replikin Counts from 1917-2007 for all Replikin Peak Genes isolated *in silico* in the pB1 gene area of influenza strains having amino acid or nucleic acid sequences publicly available at PubMed. Data is provided (1) for non-lethal human *Influenza B* between 1940 and 2007 (thick dashed medium gray line) and (2) for both the lethal and non-lethal periods of human *Influenza A* viruses between 1917 and 2007. Human *Influenza A* strains are (1) H1N1 (thick medium gray line), (2) H2N2 (thin light gray line), (3) H3N2 (thin medium gray line), and (4) H5N1 (thin black line grey). H5N1 strains isolated from chicken are illustrated by a thick medium gray line. The total number of sequences analyzed for the

data (N) is 14,227. Listed pandemics, epidemics and outbreaks are the 1918 H1N1 pandemic, the 1930's H1N1 epidemic, the 1957 H2N2 pandemic, the 1968 H3N1 pandemic, the 1977-78 H3N2 outbreaks and the H5N1 outbreaks of 1997, 2001-2004 and 2007. Over a ninety year period, pandemics, epidemics and outbreaks are associated with Replikin Counts of four or above in the RPG of influenza strains. Over the same period constant low Replikin Counts of less than four may be observed during quiescent non-lethal periods of influenza A infections and low Replikin Counts of less than four may be observed in non-lethal *Influenza B*.

[000154] Figure 3 illustrates successive “emerging” strains of influenza virus between 1930 and 2007. Mean Replikin Counts per year of isolation of various strains of influenza are provided for the polymerase area (marked with circles), the pB1 area (marked with triangles), and the pB1-F2 area (marked with squares). Data for H1N1 and H3N2 continue through 2007. Gaps represent years where no data was available for these genomic areas on PubMed. Dramatic increases in Replikin Count may be observed just before outbreak in the rebound epidemic of H1N1 beginning in the 1930's, in the pandemics of H2N2 and H2N3, which occurred in 1957 and 1968, respectively, and the outbreaks of H5N1 between 1997 and 2007. The largest increase in Replikin Count may be observed in the pB1-F2 area of the genome, which is contained within the pB1 area of the genome. The next largest increase in Replikin Count may be observed in the pB1 area of the genome, which is contained in the polymerase area of the genome. The smallest increase in Replikin Count may be observed in the polymerase area of the genome. It may be observed, therefore, that the Replikin Count becomes magnified as measured within the pB1 area as compared to the polymerase area and within the pB1-F2 area as compared to the pB1 area.

[000155] Figure 4 illustrates the relationship of Replikin Count of the Replikin Peak Gene pB1 gene area in human H5N1 to percent human mortality between 2003 and 2007 in human cases of H5N1 infection. An increase in Replikin Count in the pB1 gene area of H5N1 is observed to be quantitatively related to higher mortality in the host. In the graph, (1) light gray represents the mean Replikin Count of whole virus isolates at a given year, (2) medium gray represents the mean Replikin Count in the pB1 area of publicly available sequences of isolates of human H5N1 at a given year, (3) the colorless bars represent the standard deviation from the mean of Replikin Count in a given year, and (4) black represents ten times the percent mortality of identified human cases of H5N1 infection in the given year.

[000156] Figure 5 illustrates a 2005 through 2007 upregulation of human H5N1 in humans as compared to H5N1 in goose, duck and chicken. Dark grey represents mean Replikin Count in the Replikin Peak Gene pB1 gene area of H5N1 isolates from goose, duck, chicken and human in isolates from 2001 through 2006 where data was publicly available at www.pubmed.com. Light grey represents standard deviation from the mean.

[000157] Replikin analysis was performed separately for H5N1 Replikin Peak Genes of each host group, namely, goose, duck, chicken and human. Low levels of Replikin count, below 4, were observed in each host group until 2005-2006. In 2005-2006 epidemics began to increase in Asian countries. While duck H5N1 counts decreased in 2006, they continued to increase in chicken H5N1 in 2006. Human RPG activity was upregulated in 2005-2006 and overtook RPG activity in chickens. This transition of Replikin Count increase from duck to chicken to human is in agreement with epidemiological evidence of the order of transfer of the virus between hosts. Changes in Replikin Count in the Replikin Peak Gene of the H5N1 isolates as in Figure 5 allow for identification of those hosts in which the influenza virus strain is more virulent than other hosts.

[000158] Figure 6 illustrates localization of human H5N1 isolates having the highest lethality by measuring mean Replikin Counts in isolates of human H5N1 from different geographic areas isolated in a given year. Figure 6 is a bar graph depicting the number (with standard deviation) of Replikins per 100 amino acids in the pB1 gene area (Replikin Peak Gene) of H5N1 influenza virus strains identified annually in humans in Japan, Russia, Egypt, China, Vietnam, Thailand and Indonesia between 2003 and 2006.

[000159] Replikin analysis was performed separately for human H5N1 RPGs of each country. The results are shown for the Replikin Count for all data available on PubMed each year from 2003-2006. Low levels of Replikin count, below 4, were observed in each host group until 2005-2006, when human H5N1 increased in Asian countries. Human RPG activity was upregulated in 2005-2006 most prominently in Indonesia. The country most likely to first experience the increased human mortality was predicted in 2006 to be Indonesia. This prediction was proven correct in 2007 where incidence of human morbidity and mortality in the Indonesian outbreak were exceptionally high and evidence of possible human to human transmission was observed. Changes in Replikin Count in the Replikin Peak Gene of the H5N1 isolates such as in Figure 6 allow for identification of those geographic areas in which the influenza virus strain is more virulent than other geographic areas.

[000160] Figure 7 illustrates a relationship between Replikin Counts of Replikin Peak Genes identified within the pB1, pB2, and pA genomic areas of equine influenza 1977-2007 and epidemics of equine encephalitis caused by H3N8 equine influenza. Series 1 reflects the mean Replikin Count identified in the Replikin Peak Gene in the pB1 area of the genome. Series 2 reflects the standard deviation from mean Replikin Count in the pB1 gene area. Series 3 reflects the Replikin Count identified in the Replikin Peak Gene in the pA gene area of the genome, which neighbors the pB1 gene area. Series 4 reflects the Replikin Count identified in the Replikin Peak Gene in the pB2 gene area of the genome, which also neighbors the pB1 gene area. Replikin Count increases in the pB1 gene area are observed to occur one to three years before epidemic outbreaks while no increase in Replikin Count is observed in the pB2 and pA gene areas.

[000161] Figure 8 illustrates an increasing Replikin concentration of the whole hemagglutinin protein in the H5N1 strain of influenza virus that preceded three "Bird Flu" Epidemics between 1997 and 2004. In H5N1 influenza, the increasing strain-specific Replikin concentration (Replikin Count, Means \pm SD) 1995 to 1997 preceded the Hong Kong H5N1 epidemic of 1997 (E1); the increase from 1999 to 2001 preceded the epidemic of 2001 (E2); and the increase from 2002 to 2004 preceded the epidemic in 2004 (E3). The decline in 1999 occurred with the massive culling of poultry in response to the E1 epidemic in Hong Kong. Figure 8 demonstrates that although Replikin Count increases in RPGs occur in ranges four to eight fold greater than the increases which can be observed in whole proteins or genomes (see, *e.g.*, Figures 1 and 2), changes in the Replikin Counts of whole proteins or genomes have the advantage of completeness and may be large enough to be detected and statistically significant.

[000162] Figure 9 illustrates an increase in Replikin Count in spike and nucleocapsid coronavirus proteins preceding the SARS coronavirus epidemic of 2003. The x-axis indicates the year and the y-axis indicates the Replikin Count. The appearance of the SARS outbreak and the eight countries involved in the outbreak is shown by the conical shaded area. The solid black symbols represent the mean Replikin concentration for spike coronavirus proteins and the vertical black bars represent the standard deviation of the mean.

[000163] Although SARS was first identified in 2003, Applicants wondered whether the emergence of the SARS strain of coronavirus might have been presaged in the activity of the whole group of coronaviruses. The pre-pandemic increase in both nucleocapsid and spike coronavirus proteins is in accord with, and might have served as a warning of, the finding that

a coronavirus would be responsible for the 2003 first SARS emergence. It may be seen that the Replikin Count rose between 1995 and 2002, consistent with the SARS coronavirus outbreak, which emerged at the end of 2002 and persisted into 2003. The decline in Replikin Count correctly signaled the end of the SARS outbreak and had already begun its return to pre-outbreak levels when the outbreak emerged. A similar decline occurred on termination of *Influenza A* epidemics and pandemics (Figure 2). As also seen in Figure 2, however, this decline has not occurred in the case of H5N1 in 2006 and 2007, so that the ongoing H5N1 outbreak may be assumed not to be over.

[000164] Figure 10 illustrates that mortality rates in humans from *Plasmodium falciparum* correlate with Replikin Count in the *P. falciparum* ATP-ase enzyme. High malaria morbidity and mortality rates occurred in the late 1990s and was thought to be due to adaptation of the microorganism and decreased effectiveness of anti-malarials. ATP-ase is a primary target of artemisinin treatment of malaria. With increased use of artemisinin, and improved public health measures, morbidity and mortality rates declined from 1998 to 2006. The Replikin Count of *P. falciparum* ATP-ase increased from 1997 to 1998 along with an increase in mortality per 250 malaria cases. The Replikin Count of *P. falciparum* ATP-ase decreased along with mortality rates from 1998 to 2006. Mortality rates per 250 cases for 1997 to 2006 were as follows: 1997 mortality rates was 7.7; 1998 mortality rate was 6.6; 1999 mortality rate was 9.1; 2000 mortality rate was 10.5; 2001 mortality rate was 8.1; 2002 mortality rate was 9.9; 2003 mortality 2.5; 2004 mortality rate was 4; 2005 mortality rate was 3.9; 2006 mortality rate was 2.6. Mortality rates declared by the World Health Organization, *see* www.who.int.

[000165] Figure 11 illustrates a relationship between Replikin Counts observed in the VP1 protein (Replikin Peak Gene) of isolates of publicly-available foot and mouth disease virus serotype-O between 1969 and 2006 and certain observed outbreaks of Foot and Mouth Disease. Standard deviations are represented by vertical light grey capped lines above mean Replikin Counts. Observed European and UK outbreaks of Foot and Mouth Disease are noted including outbreaks in the UK in 1967, 1981, 2001 and 2007, in Baltic states in 1991 and 1993 through 1996, and Japan, Korea and Greece in 2000. Increases in Replikin Counts from baseline values between 1969 and 1978 preceded repeated increased Replikin counts 1979 forward, which in turn preceded outbreaks of foot and mouth disease 1981 to 2007.

[000166] Figure 12 illustrates a relationship between Replikin Counts observed in the envelope protein of isolates of west nile virus and total human morbidity and mortality. The

data for Figure 12 is contained in Table 10. A correlation between Replikin Count in the envelope protein (the protein containing the RPG of the virus), and Morbidity and Mortality is demonstrated. Figure 12 is a graph comparing (1) the concentration of Replikin (Replikin Count) of publicly available sequences of the envelope protein of isolates of west Nile virus between 1982 and 2007 (with standard deviation bars for each data point), (2) total morbidity reported in the United States on a year by year basis by the Center for Disease Control (total U.S. morbidity is the value denoted on the y-axis times 100) between 1999 and 2007, and (3) total mortality resulting from WNV infection reported in the United States on a year by year basis by the Center for Disease Control between 1999 and 2007.

[000167] Figure 13 illustrates Replikin Counts in the nucleocapsid protein of the porcine respiratory and reproductive syndrome virus (PRRSV) in isolates from 2004 through 2007. Mean Replikin Count is shown in grey columns. Standard deviation from the mean is shown in colorless columns. The Replikin Count of PRRSV nucleocapsid protein is seen to increase between 2004 and 2007. This increase correlates with a major outbreak of PRRSV in China. Standard deviation from the mean in 2005 is considerably larger than other years demonstrating a marked increase in Replikin Count was occurring in 2005 and measured as an increase in mean Replikin Count in 2006. The large standard deviation observed in 2005 indicates that more members of the class had increasing Replikin Counts. Standard deviation in 2005 was an early warning prior to the increase in the mean in 2006 and 2007. A similar phenomenon is observable in Figure 7.

[000168] Figure 14 illustrates a correlation between cumulative survival of *Litopenaeus vannamei* shrimp challenged with four different taura syndrome virus isolates over 15 days (unless 100% mortality occurred prior to 15 days) and the Replikin concentration of Open Reading Frame 1 (ORF1) of each isolate. Translated amino acid sequences of ORF1 of the genome of individual isolates of TSV from Belize, Thailand, Hawaii and Venezuela were analyzed for Replikin Count. Replikin Count was determined to be 3.5 for the Belize isolate, 3.4 for the Thailand isolate, 3.3 for the Hawaii isolate and 3.0 for the Venezuela isolate. Graph A illustrates observed percent survival in three trials of shrimp challenged with the Belize isolate of TSV. In one trial, total mortality was observed on day 6. In the other trials, total mortality was observed on day 11. Graphs B, C and D illustrate observed percent survival of shrimp challenged with the Thailand isolate, the Hawaii isolate and the Venezuela isolate, respectively, each in three trials over 15 days. In the Thailand isolate, a mean of 80% percent mortality was observed on day 15. In the Hawaii isolate, a mean of 78.3% mortality

was observed on day 15. In the Venezuela isolate, a mean of 58.3% mortality was observed on day 15.

[000169] Figure 15A illustrates a direct sequential correlation between Replikin Count in isolates of taura syndrome virus (TSV) collected from Belize, Thailand, Hawaii and Venezuela, respectively, and mean number of days to 50% mortality in *Litopenaeus vannamei* shrimp challenged with the respective TSV isolates beginning on day one through day three. Statistical differences between the Replikin concentration for each isolate are significant at a level of $p < 0.001$.

[000170] Figure 15B illustrates a direct correlation between Replikin Count in isolates of taura syndrome virus (TSV) collected from Belize, Thailand, Hawaii and Venezuela, respectively, and mean cumulative survival of *Litopenaeus vannamei* shrimp at 15 days after challenge with the respective TSV isolate. Statistical differences between the Replikin concentrations for each isolate are significant at a level of $p < 0.001$.

[000171] Figure 16 illustrates a magnification of the effect of increases in Replikin Count on human mortality from H5N1 infections when Replikin concentration is observed in the pB1 gene area (containing a RPG) as compared to the polymerase gene or as compared to the entire genome of the H5N1 virus. In Figure 16, a correlation is established between human mortality and (1) mean concentration of Replikin sequences in the whole genome, (2) mean concentration of Replikin sequences in the polymerase gene, and (3) mean concentration of Replikin sequences in the Replikin Peak Gene (pB1 gene area) of H5N1 influenza strains. Replikin concentration in the Replikin Peak Gene (pB1 gene area) of the H5N1 genome is seen to correlate most significantly with human mortality as compared to Replikin Counts in the whole genome and the polymerase gene.

[000172] Figure 17 illustrates a significant eight-fold increase in Replikin concentration in the pB1 gene area (Replikin Peak Gene) of isolates of H5N1 from 2003 through the first quarter of 2007 (that correlates with an increase in host mortality in humans), while no significant increase is observed in neighboring gene areas of the pB1 gene area, namely, the pA gene area and the pB2 gene area. Figure 17 graphically compares percent human mortality from H5N1 infections in years 2005 through the first quarter of 2007 to mean concentration of Replikin sequences in (1) the pB1 gene area, (2) the pB2 gene area, and (3) the pA gene area, respectively, of H5N1 influenza strains isolated in 2003 through the first quarter of 2007.

[000173] Figure 18 illustrates a correlation between the mean Replikin Count and standard deviation of Replikin sequences observed in publicly available amino acid sequences of white spot syndrome virus (WSSV) isolated between 1995 and 2007 and a significant outbreak of WSSV in 2001. The remarkably high Replikin concentration in 2000 of 97.6 predicts the 2001 outbreak. Furthermore, an even more remarkable Replikin concentration of 103.8 was observed in a ribonucleotide reductase protein sequence from a 2000 isolate of WSSV wherein a Replikin Peak Gene was identified with an even higher Replikin concentration of 110.7.

[000174] Figure 19 illustrates a correlation between increased Replikin Count in the genome of taura syndrome virus and outbreaks of the virus in 2000 and 2007 in shrimp. taura syndrome virus peptide sequences available at www.pubmed.com were analyzed by the inventors for mean Replikin concentration in the publicly available sequences. Figure 19 is a graph comparing mean Replikin concentration for each year in which peptide sequences were publicly available between 2000 and 2005 (with standard deviation) and dates of significant outbreaks of taura syndrome virus. Significant outbreaks of the disease are noted at years 2000 and 2007. It may be observed from the graph that outbreaks of the virus occur following an increase in Replikin concentration. In year 2000, TSV had a Replikin Count of 2.7. Between 2001 and 2004, TSV had a lower mean Replikin Count, as low as 0.7, and an identified Replikin Scaffold disappeared. In 2005 the Replikin Scaffold reappeared, with an increase in lysines and histidines, and a commensurate increase in Replikin concentration to 1.8, followed by an increase in TSV outbreaks in 2006-2007.

[000175] Figure 20 illustrates the total hemagglutinin Replikin Counts in the three influenza pandemics of the last century. Strain-specific high Replikin Counts accompany each of the three pandemics of the last century: 1918, 1957, and 1968. In each case this peak is followed by a decline (likely due to immunity in the hosts), then by a recovery and a “rebound” epidemic. The probability is very low that these correlations are due to chance, since they are specific for each strain, specific for each of the three pandemic years out of the century, specific for each post-pandemic decline, and specific for each rebound epidemic. Example 13 provides an example of analysis of hemagglutinin Replikin Counts in publicly available sequences between 1918 and 2007.

[000176] Figure 21 illustrates an annual mean Replikin Count observed in isolates of porcine circovirus (PCV) having publicly available accession numbers on a year by year basis between 1997 and 2007 (with standard deviation bars for each Replikin Count data

point) and demonstrates a correlation between increases in Replikin Count from 2000 through 2007 and reported increased in morbidity and mortality in Canada between 2000 and 2006 and an outbreak in China in 2007.

DETAILED DESCRIPTION OF THE INVENTION

Definitions

[000177] As used herein, a Replikin Peak Gene (RPG) (or sometimes a Replikin Peak Gene Area-RPGA) is to mean a segment of a genome, protein, segment of protein, or protein fragment in which an expressed gene or gene segment has a highest concentration of continuous, non-interrupted and overlapping Replikin sequences (number of Replikin sequences per 100 amino acids) when compared to other segments or named genes of the genome. Generally, a whole protein or gene or gene segment that contains the amino acid portion having the highest concentration of continuous Replikin sequences is also referred to as the Replikin Peak Gene. More than one RPG may be identified within a gene, gene segment, protein, or protein fragment. An RPG may have a terminal lysine or a terminal histidine, two terminal lysines, or a terminal lysine and a terminal histidine. For diagnostic, therapeutic and preventive purposes, an RPG may have a terminal lysine or a terminal histidine, two terminal lysines, or a terminal lysine and a terminal histidine or may likewise have neither a terminal lysine nor a terminal histidine so long as the terminal portion of the RPG contains a Replikin sequence or Replikin sequences defined by the definition of a Replikin sequence, namely, an amino acid sequence having about 7 to about 50 amino acids comprising:

- (1) at least one lysine residue located six to ten amino acid residues from a second lysine residue;
- (2) at least one histidine residue; and
- (3) at least 6% lysine residues.

Further, for diagnostic, therapeutic, preventive and predictive purposes, an RPG may include the protein or protein fragment that contains an identified RPG. For example, an RPG is herein identified in the pB1 gene area of H5N1. For predictive purposes, a Replikin Count in the RPG may be used to track changes in virulence and lethality. Likewise the RPG may be used as an immunogenic compound or as a vaccine. Additionally, however, as described herein, a Replikin Count in the pB1 gene area of influenza strains (like, for example, H5N1, H1N1 and H3N8), which contains but is not limited to an identified RPG having highest

concentration of continuous, non-interrupted and overlapping Replikin sequences, is particularly useful for predicting changes in lethality and virulence. Other examples of predictive use of Replikin Counts in proteins in which RPGs have been identified are the VP1 protein of foot and mouth disease virus, the envelope protein in the west Nile virus, and the nucleocapsid protein in porcine respiratory and reproductive syndrome virus, among many other viruses and organisms. Whole proteins or protein fragments containing RPGs are likewise useful for diagnostic, therapeutic and preventive purposes, such as, for example, to be included in immunogenic compounds, vaccines and for production of therapeutic or diagnostic antibodies.

[000178] As used herein, a Replikin sequence is an amino acid sequence having about 7 to about 50 amino acids comprising:

- (1) at least one lysine residue located six to ten amino acid residues from a second lysine residue;
- (2) at least one histidine residue; and
- (3) at least 6% lysine residues.

[000179] A Replikin sequence may comprise a terminal lysine and may further comprise a terminal lysine or a terminal histidine. A Replikin peptide or Replikin protein is a peptide or protein consisting of a Replikin sequence. A Replikin sequence may also be described as a Replikin sequence of about 7 to about 50 amino acids comprising or consisting of a Replikin motif wherein the Replikin motif comprises:

- (1) at least one lysine residue located at a first terminus of said isolated influenza virus peptide and at least one lysine residue or at least one histidine residue located at a second terminus of said isolated influenza virus peptide;
- (2) a first lysine residue located six to ten residues from a second lysine residue;
- (3) at least one histidine residue; and
- (4) at least 6% lysine residues.

For the purpose of determining Replikin concentration, a Replikin sequence must have a lysine residue at one terminus and a lysine or a histidine residue at the other terminus.

[000180] The term "Replikin sequence" can also refer to a nucleic acid sequence encoding an amino acid sequence having about 7 to about 50 amino acids comprising:

- (1) at least one lysine residue located six to ten amino acid residues from a second lysine residue;
- (2) at least one histidine residue; and

(3) at least 6% lysine residues, wherein the amino acid sequence may comprise a terminal lysine and may further comprise a terminal lysine or a terminal histidine.

[000181] As used herein, “animal” includes mammals, such as humans.

[000182] As used herein, the term “peptide” or “protein” refers to a compound of two or more amino acids in which the carboxyl group of one amino acid is attached to an amino group of another amino acid via a peptide bond. As used herein, “isolated” or “synthesized” peptide or biologically active portion thereof refers to a peptide that is, after purification, substantially free of cellular material or other contaminating proteins or peptides from the cell or tissue source from which the peptide is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized by any method, or substantially free from contaminating peptides when synthesized by recombinant gene techniques or a protein or peptide that has been isolated *in silico* from nucleic acid or amino acid sequences that are available through public or private databases or sequence collections. An “encoded” or “expressed” protein, protein sequence, protein fragment sequence, or peptide sequence is a sequence encoded by a nucleic acid sequence that encodes the amino acids of the protein or peptide sequence with any codon known to one of ordinary skill in the art now or hereafter. It should be noted that it is well-known in the art that, due to redundancy in the genetic code, individual nucleotides can be readily exchanged in a codon and still result in an identical amino acid sequence. As will be understood by one of skill in the art, a method of identifying a Replikin amino acid sequence also encompasses a method of identifying a nucleic acid sequence that encodes a Replikin amino acid sequence wherein the Replikin amino acid sequence is encoded by the identified nucleic acid sequence.

[000183] As used herein, “reservoir” is any source of Replikin sequences that may be shared with a virus, organism or malignancy including any host of a virus, organism or malignancy, any food source of a host of the virus, organism or malignancy, any vector of virus, organism or malignancy, or any substance wherein the genetic information of a virus, organism or malignancy that may be shared, mingled, mixed, exchanged, or come into the proximity of the Replikin sequences of the reservoir.

[000184] As used herein, “different time periods” or “different time points” is any two time periods or points that may be differentiated one from another. For example, an isolate of virus isolated during the year 2004 is isolated in a different time period than an isolate of the same virus isolated during the year 2005. Likewise, an isolate of virus isolated in May 2004

is isolated in a different time period than an isolate of the same virus isolated in June 2004. When comparing Replikin concentrations of different isolates, it is preferred to use comparable time periods for comparison. For example, an isolate from 2004 is preferably compared to at least one other isolate from some other year such as 2002 or 2005. Likewise, an isolate from May 2004 is preferably compared to at least one isolate from some other month of some year, for example, an isolate from December 2003 or from June 2004. An isolate is any virus isolated from a natural source wherein a natural source includes, but is not limited to, a reservoir of a virus, a vector of a virus or a host of a virus. "Obtaining" an isolate is any action by which an amino acid or nucleic acid sequence within an isolate is obtained including, but not limited to, isolating an isolate and sequencing any portion of the genome or protein sequences of the isolate, obtaining any nucleic acid sequence or amino acid sequence of an isolate from any medium, including from a database such as PubMed, wherein the nucleic acid sequence or amino acid sequence may be analyzed for Replikin concentration, or any other means of obtaining the Replikin concentration of a virus isolated from a natural source at a time point.

[000185] As used herein, an earlier-arising virus or organism or a virus or organism isolated at an earlier time period is a specimen of a virus or organism collected from a natural source of the virus or organism on a date prior to the date on which another specimen of the virus or organism was collected from a natural source. For viruses, a natural source includes, but is not limited to, a reservoir of a virus, a vector of a virus, or a host of the virus. A later-arising virus or organism or a virus or organism isolated at a later time period is a specimen of a virus or organism collected from a natural source of the virus (including, but not limited to, a reservoir, a vector, or a host) or a natural source of the organism on a date subsequent to the date on which another specimen of the virus or organism was collected from a natural source.

[000186] As used herein, "emerging strain" refers to a strain of a virus identified as having an increased or increasing concentration of Replikin sequences in one or more of its protein sequences relative to the concentration of Replikins in other strains of such organism. The increased or increasing concentration of Replikins occurs over a period of preferably at least about six months, at least about one year or at least about three years, but may be a much shorter period of time for highly mutable viruses. An emerging strain of virus indicates an increase in lethality, virulence or replication.

[000187] As used herein, “bird” is any avian species including migratory and domestic birds, wherein said migratory and domestic birds includes, for example, chickens, ducks of all kinds, geese, pigeons, gulls, seabirds *etc.*

[000188] As used herein, “outbreak” is an increase in virulence, morbidity or mortality in a viral disease as compared to a baseline of an earlier occurring epidemiological pattern of infection in the same viral disease. One of ordinary skill in the art will know how to determine an epidemiological baseline. As used herein, “morbidity,” is the number of cases of a disease caused by the virus, either in excess of zero cases in the past or in excess of a baseline of endemic cases in the past. Therefore the baseline of endemic cases, in epidemiological terms, may, for example, relate to whether no or some cases were present in a geographic region in the immediate past. The past, in epidemiological terms, may mean more than one year and can mean several years or more as understood by one of ordinary skill in the art. The past may also mean less than one year as determined by one of ordinary skill in the art. In the case of annually-recurrent common influenza, for example, the baseline reflects an annual recurrence of common influenza.

[000189] As used herein, “mutation” refers to a change in the structure and properties of a virus or organism caused by substitution of amino acids. In contrast, the term “conservation” as used herein, refers to conservation of particular amino acids due to lack of substitution. A “point mutation” may refer to a change in a single amino acid residue or may refer to a change in a small number of amino acid residues.

[000190] As used herein, “segment” or “portion” of a genome, protein or protein fragment refers to any nucleic acid sequence of any size within a genome or any amino acid sequence of any size within a protein or protein fragment wherein the termini of the nucleic acid sequence may be any two nucleic acid residues within the genome and the termini of the amino acid sequence may be any two amino acid residues within the protein or protein fragment.

[000191] As used herein, “Replikin Count” or “Replikin Concentration” refers to the number of Replikins per 100 amino acids in a protein, protein fragment, virus, or organism. A higher Replikin concentration in a first strain of a virus or organism has been found to correlate with more rapid replication of the first virus or organism as compared to a second, earlier-arising or later-arising strain of the virus or organism having a lower Replikin concentration.

[000192] As used in this patent application, the term “continuous Replikin sequences” means a series of two or more Replikin sequences that are overlapped or are directly covalently linked.

[000193] As used herein a “Replikin Scaffold” refers to a series of conserved Replikin peptides wherein each of said Replikin peptide sequences comprises about 16 to about 34 amino acids, and preferably about 27 to about 33 amino acids and further comprises: (1) a terminal lysine and optionally a lysine immediately adjacent to the terminal lysine; (2) a terminal histidine and optionally a histidine immediately adjacent to the terminal histidine; (3) a lysine within 6 to 10 amino acid residues from another lysine; and (4) about 6% lysine. “Replikin Scaffold” also refers to an individual member or a plurality of members of a series of Replikin Scaffolds.

[000194] In an influenza virus, a Replikin Scaffold may refer to a Replikin peptide sequence comprising about 16 to about 34 amino acid residues, and in a preferred embodiment about 28 to about 30 amino acid residues. In a white spot syndrome virus, a Replikin Scaffold may refer to a Replikin peptide sequence comprising about 16 to about 34 amino acid residues, and in a more preferred embodiment about 29 to about 31 amino acid residues. In a taura syndrome virus, a Replikin Scaffold may refer to a Replikin peptide sequence comprising about 16 to about 34 amino acid residues, and in a more preferred embodiment about 29 to about 33 amino acid residues.

I. Replikin Count in Replikin Peak Gene is Predictive of and Related to Virulence and Lethality in Malignancies, Influenza and Other Pathogens and Replikin Peak Genes and Associated Replikin Sequences are Useful for Diagnostic, Therapeutic and Predictive Purposes

[000195] A virus Replikin gene related to lethality and virulence was first identified by Applicants in human H5N1 *Influenza* And was labeled a Replikin Peak Gene. Replikin Peak Genes were subsequently isolated *in silico* in numerous other viruses, bacteria, and protozoa. Replikin Peak Genes have now been associated with lethality in plant, fish, crustacea and vertebrate hosts. Because of their association with lethality, virulence and rapid replication, Replikin Peak Genes are now available as excellent targets for therapeutic and preventive treatments for a wide range of malignancies and pathogens.

[000196] Replikins, a class of peptides related to rapid replication, are 7 to 50 amino acids long, containing at least 2 lysine groups 6 to 10 amino acids apart, at least 1 histidine group, and at least 6% lysine. The phenomenon of the association of Replikins with rapid replication and virulence has been fully described in U.S. Patent No. 7,189,800, U.S. Patent

No. 7,176,275, U.S. Appln. Ser. No. 11/355,120, U.S. Appln. Ser. No. 10/860,050 and U.S. Appln. Ser. No. 10/105,232. Both Replikin concentration (number of Replikins per 100 amino acids) and Replikin composition have been correlated with the functional phenomenon of rapid replication.

[000197] Using an algorithm constructed to identify, count, and track Replikin sequences historically, Replikins were analyzed in 130,488 protein and genome sequences, representing all the accession numbers for common strains of influenza and some other lethal virus isolations published between 1917 and 2007 and reported on PubMed. Genomic areas with the highest concentration of continuous Replikins were isolated and named Replikin Peak Genes (RPGs).

[000198] Analysis of all publicly available protein and genome sequences for lethal *Influenza A* strains, including H5N1, revealed 10,182 RPGs. RPGs were found to be present in isolates from all outbreaks of lethal influenza between 1917 and 2007 and the number of Replikin sequences per 100 amino acids (Replikin Count) in the identified RPGs was consistently observed to be above four and increased to as high as 29. In a significant control in *Influenza B* virus, which is non-lethal in humans, the Replikin Count in all 371 RPGs in *Influenza B* between 1940 and 2006 were found never to exceed four. Replikin Counts below four in the RPG of *Influenza B* virus contrasts with lethal *Influenza A* strains (with Replikin Counts as high as 29). RPG Replikin Counts during quiescent (or non-lethal) periods of *Influenza A*, were regularly observed to be four or below.

[000199] Replikin Counts below four for non-lethal isolates of influenza may be compared to highly lethal or virulent viruses such as *ebola virus*, which has been observed to have a Replikin Count of 32, Porcine Reproductive and Respiratory Virus (PRRSV) in pigs, which has been observed to have a Replikin Count of 43, gemini yellow leaf curl virus in tomato plants observed to have a Replikin Count of 56, hemorrhagic septicemia virus in fish observed to have a Replikin Count of 59, and white spot syndrome virus in shrimp, which has been observed to have a Replikin Count of 106. All of these viruses were observed to return to low counts during quiescent periods. Increased Replikin Counts in RPGs also were found in *Mycobacterium tuberculosis* (28), in methicillin-resistant *Staphylococcus aureus* (81), in *Plasmodium falciparum* (malaria) (153), and in lung cancer (261).

[000200] Analysis of Replikin Counts in genomic and proteomic sequences alone prospectively correctly predicted: 1) the order of lethality in shrimp of four strains of taura syndrome virus (prediction was made blind in a laboratory study); 2) a 2007 increase in

H5N1 percent mortality in humans; and 3) the country in which the increased percent mortality would occur most significantly, namely, Indonesia.

[000201] In addition to high Replikin Counts, analysis of rapidly replicating, virulent and lethal virus has revealed a series of conserved Replikin peptides associated with rapid replication, virulence and lethality known as Replikin Scaffolds. Replikin Scaffolds were observed in influenza virus strains where, for example, a 29-amino acid Replikin Scaffold has been conserved for 90 years in the genome of successive influenza virus strains. The scaffold has been present in each of the lethal influenza pandemics of 1918, 1957 and 1968 and in each lethal H5N1 outbreaks. Repeating signatures such as a KHKK (SEQ ID NO:1584) signature has been observed in Replikin sequences within RPGs of lethal malignancies, viruses and organisms. The KHKK (SEQ ID NO:1584) signature has been observed eleven times within the RPG of the protozoa that causes most malaria, *P. falciparum*. The KHKK (SEQ ID NO:1584) signature has been observed 20 times within the RPG of a tobacco mosaic virus that induced exacerbated cell death in a pepper plant. The KHKK (SEQ ID NO:1584) signature has been observed 57 times in non-small cell lung carcinoma within 52 Replikins observed within the 18 amino acid RPG identified in chromosome 9 of a non-small cell lung carcinoma. The presence of such a high number of KHKK (SEQ ID NO:1584) signatures within the 18 amino acid RPG of the non-small cell lung carcinoma is explained by overlapping of the signatures. Overlapping of Replikin sequences and repeated signatures such as KHKK (SEQ ID NO:1584) has now been associated with lethality, virulence and rapid replication. Together, these data indicate that a Replikin gene is quantitatively associated with lethal functions, and may be a mobile agent of lethality transferring between strains and species.

[000202] Whether Replikins can arise by synthesis *de novo* or are transferred from one organism or virus to another (or both) is yet to be determined. There is some beginning evidence for both. In one experiment, Replikin synthesis and/or transfer was facilitated in the laboratory in glioblastoma multiforme cells growing in tissue culture. The event, which facilitated the synthesis and/or transfer, was induced anoxia. Whether the anoxia stimulates increased rate of Replikin synthesis or membrane impairment facilitates Replikin transfer, or both, is yet to be determined.

[000203] Counting of Replikin sequences within a malignancy, a virus, a protozoon, a plant or an animal is aided by computer review of databases of gene and protein sequences. Bacteria were accepted as real when the light microscope permitted them to be seen as

discrete entities, sufficiently discrete that they could be counted. Similarly, viruses were accepted as real when the electron microscope permitted them to be seen as discrete entities, sufficiently discrete that they could be counted. Likewise, Replikins can now be accepted as real since the “computer microscope” permits them to be seen as discrete entities, sufficiently discrete that they can be counted. Hence, the Replikin Count, or determination of number of Replikin Sequences in 100 amino acids in any given genomic or proteomic sequence, is facilitated on a large scale by computer analysis and comparison of Replikin Counts has provided the necessary evidence to associate increased Replikin Counts (in both whole genomes and Replikin Peak Genes) with lethality.

[000204] Visualization and counting of Replikin sequences in a wide range of genomes has now revealed that Replikin sequences are not scattered throughout the genome of lethal, virulent and rapidly replicating entities but, instead, are concentrated in particular areas of the genome. The concentration of Replikin sequences in a particular area of the genome has now been identified as a Replikin Peak Gene (RPG). Concentration of Replikin sequences in a RPG provides a magnification of the Replikin Count and a magnification of the developmental, growth and disease associations with the presence of Replikin Sequences. *See, e.g.*, Figures 1, 3, 16 and 17. This magnification not only makes identification and counting easier, but facilitates the discovery of both the structural history and the functional associations of Replikins, as seen, for example, in the increase in Replikin Count of the RPG of human H5N1 with the increased percent human mortality between 2003 and 2007. Figures 4, 16 and 17.

[000205] The magnification effect of analyzing the Replikin Count of a Replikin Peak Gene as compared to Replikin Counts from other parts of a genome or the whole genome is demonstrated in Figures 16 and 17. There, mortality in humans from H5N1 infection correlates strongly with an increase in Replikin Count in the pB1 gene area (RPG) of the virus while correlating less strongly with an increase in Replikin Count in the polymerase gene or the whole genome of the virus.

[000206] By means of visual and software inspection, Applicants have analyzed 130,488 protein and genome sequences from common strains of influenza and other lethal viruses, isolated from 1917 to 2007 and accessible in PubMed. Replikin sequences in these 130,488 sequences have been identified, counted and annually tracked. This extensive analysis revealed the Replikin Peak Gene that has not been found to be quantitatively related to lethality in several hosts, including plants, fish, crustacea and vertebrates, such as humans.

II. Prediction of Pathogenic Outbreaks and Lethal Malignancies

[000207] Prediction of epidemics and future outbreaks of viruses such as *Influenza A* (including H1N1, H2N2, H3N2, H3N8 and H5N1), foot and mouth disease virus, west nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, tobacco mosaic virus, coronavirus, and SARS virus, may be made, for example, by reviewing the Replikin concentration of isolates of a virus strain and comparing the Replikin concentration for a particular time period with Replikin concentrations from another time period. Prediction of outbreaks or increases in virulence or lethality of organism may also be made, for example, by reviewing the Replikin concentration of isolates of an organism and comparing the Replikin concentration for a particular time period with Replikin concentrations from another time period. Organisms for which outbreaks or increases in virulence and lethality may be predicted include, for example, *P. falciparum*, *M. mucogenicum* and *S. aureus*.

[000208] The difference in time period may be, for example, one month, six months, one year, three years or more. Preferably, the difference in time period is six months to three years. Also preferably, the difference in time period is one year. A significant increase in Replikin concentration from one year to the next and preferably over one, two, three or five years provides predictive value of an emerging strain of virus or organism that may begin an outbreak. A viral or other pathogenic outbreak may be predicted within about six months to about one to about three-years from the observation of a significant increase in Replikin concentration. The outbreak is preferably predicted within about one to about two years. An outbreak of virus or other pathogen, therefore, may be predicted within 1 to about 2 years as demonstrated in Figures 2, 3, 7, 11 and 19 wherein an epidemic occurred at about 1 to about 2 years following each peak of the measured Replikin Count of the particular viruses and organisms.

[000209] Significant increases may be observed over a time period of more than one year, such as three, four, five or more years. An outbreak may likewise be predicted within about six months to about one year or more from the initial observation of an observable decrease in Replikin concentration following a significant increase.

[000210] The correlation between Replikin concentration and viral outbreaks noted throughout this application provide a method of predicting outbreaks of virus and other pathogens by monitoring increases or decreases in Replikin concentration in the RPG of isolates of the virus or other pathogen. Likewise, the lethality of an organism may be

predicted by comparing the Replikin Count of the identified RPG of a malignancy with the Replikin Count of the identified RPG of another malignancy of the same type.

III. Replikin Peak Gene Sequences in Diagnostics and Therapies

[000211] High Replikin concentrations and RPGs have been shown to be related to rapid replication, viral outbreaks, epidemics, morbidity and host mortality in, for example, influenza virus strains, including H5N1, in SARS coronavirus, in shrimp in taura syndrome virus and in white spot syndrome virus, in foot and mouth disease virus, porcine reproductive and respiratory syndrome virus and in porcine circovirus, and in malignancies such as non-small cell lung carcinoma, among others. Because Replikin sequences in general (and particularly RPGs) are chemically defined, the sequences may be synthesized by organic chemistry rather than biological techniques, and thus are more specific, more reproducible and more reliable than other targets for diagnostics and therapeutics. The chemically defined Replikin sequences identified by the applicants are likewise potentially free of adverse reactions that are characteristic of biologically derived vaccines and antibodies.

[000212] In addition to the diagnostic power of Replikin technology shown in these examples, it is clear that recognition for the first time of this class of virus peptides, and the discovery that they are related to rapid replication, virus outbreaks and high morbidity and mortality, makes the Replikins, and particularly the Replikin Peak Gene structures illustrated herein, new conserved prime targets for treatment and vaccines against these and other viruses.

[000213] Presence of the Replikin Peak Gene correlates with an increase in virulence in various species and an increase in mortality rate in humans in influenza virus, malaria and lung cancer and in pigs in PRRSV and porcine circovirus. Because an increase in virulence and mortality rate can be correlated with the Replikin Peak Gene (RPG), portions or fragments of the RPG are available as preferred targets for treatment with vaccines, antibodies or other blocking agents. Replikins in the gene are further preferred targets for identification of virulent strains of virus and other pathogens and for prediction of outbreaks of virus and other pathogens.

IV. Immunogenic Compounds, Vaccines, Antibodies and Blocking Agents

[000214] The observations of specific Replikins and their concentration in proteins of viral and organismal pathogens and malignancies provides the first specific quantitative early chemical correlates of outbreaks and provides for production and timely administration of vaccines tailored specifically to treat the prevalent emerging or re-emerging strain virus in a particular region of the world. By analyzing the protein sequences of isolates of a virus for

the presence, concentration and/or conservation of Replikins, virus outbreaks and epidemics can be predicted and treatments developed. Furthermore, the severity of such outbreaks can be significantly lessened by administering a peptide immunogenic compound or vaccine based on the Replikin sequences identified herein or using the methods provided herein or Replikin sequences found to be most abundant or shown to be on the rise in virus isolates over a given time period, such as about one to about three years.

[000215] Vaccine products against SARS Replikin sequences and H5N1 influenza virus Replikin Scaffolds have been demonstrated by Applicants. *See, e.g.*, U.S. Appln. Ser. No. 11/355,120, filed February 16, 2006 (Examples 6 and 7), incorporated herein by reference. Replikin sequences added to the feed source of shrimp have likewise imparted measurable resistance to challenges with taura syndrome virus. *See* Example 19. To date, all Replikin sequences tested in rabbit or chicken have induced an immune response and the glioma Replikin sequence (SEQ ID NO:3658) has been identified and synthesized in peptides that induce an immune response and react with natural antibody responses in humans. *See* U.S. Patent No. 6,638,505.

[000216] An immunogenic compound or peptide vaccine of the invention may include a single Replikin peptide sequence or may include a plurality of Replikin sequences observed in particular virus strains. Preferably, the peptide vaccine is a Replikin Peak Gene or a Replikin sequence isolated within a Replikin Peak Gene. Further, the peptide vaccine may be based on Replikin sequence(s) shown to be increasing in concentration over a given time period and conserved for at least that period of time. A vaccine may also include a conserved Replikin peptide(s) in combination with a new Replikin(s) peptide or may be based on new Replikin peptide sequences. The Replikin peptides can be synthesized by any method, including chemical synthesis or recombinant gene technology, and may include non-Replikin sequences, although vaccines based on peptides containing only Replikin sequences, Replikin Peak Genes or Replikin sequences identified within a Replikin Peak Gene are preferred. Preferably, vaccine compositions of the invention also contain a pharmaceutically acceptable carrier and/or adjuvant.

[000217] The immunogenic compounds and vaccines of the present invention can be administered alone or in combination with antiviral drugs, such as gancyclovir; interferon; interleukin; M2 inhibitors, such as, amantadine, rimantadine; neuraminidase inhibitors, such as zanamivir and oseltamivir; and the like, as well as with combinations of antiviral drugs.

[000218] The vaccine of the present invention may be administered to any animal capable of producing antibodies in an immune response. For example, the vaccine of the present invention may be administered to a rabbit, a chicken, a pig or a human. Because of the universal nature of Replikin sequences, a vaccine of the invention may be directed at a range of strains of virus or a particular strain of virus.

V. Increased Replikin Counts in Replikin Peak Gene of pB1 Area of Influenza A Strains Correlates with Pandemics and Lethal Outbreaks

[000219] Applicants have identified Replikin Peak Genes as a segment of a genome, protein, segment of protein, or protein fragment in which an expressed gene or gene segment has the highest concentration of continuous, non-interrupted and overlapping Replikin sequences (number of Replikin sequences per 100 amino acids) as compared to other segments or named genes of a genome. The inventors have likewise identified gene areas or proteins or protein fragments containing the highest concentration of continuous, non-interrupted and overlapping Replikin sequences (number of Replikin sequences per 100 amino acids) as Replikin Peak Genes.

[000220] Increased Replikin Counts in the Replikin Peak Gene identified in the pB1 gene area of *influenza A* strains has now been correlated by Applicants with pandemics and lethal outbreaks of influenza. These findings correspond to the Applicants' discovery that quantitative measurement of the concentration of Replikin peptides in proteins allows for correlation of Replikin peptide concentration per 100 amino acids with virulence, morbidity, mortality, epidemics and pandemics in malignancies, and organismal and viral infections. A correlation between increased Replikin Counts in the RPG of malignancies and pathogens has been established by Applicants in, for example, human pandemic influenza viruses, H5N1 ("Bird Flu") influenza virus, white spot syndrome virus, foot and mouth disease virus, west Nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, equine influenza virus, tobacco mosaic virus, malaria and non-small cell lung malignancies, among others. An increase in Replikin Count in these pathogens and malignancies allows for prediction of increased lethality or virulence and prediction of forthcoming outbreaks of infections.

A. Replikin Peak Gene in H5N1 Associated with Lethal Outbreak

[000221] Applicants initially identified a Replikin Peak Gene in the pB1 gene area of the genome of the H5N1 strain of influenza virus (*e.g.*, SEQ ID NO: 1684) and observed that outbreaks of the H5N1 virus and lethality in infections from the virus correlated with increases in Replikin Count in the identified Replikin Peak Gene. Figure 1 illustrates the

localization of the pB1 gene area as the Replikin Peak Gene in the genome of the H5N1 strain of influenza virus. The data for Figure 1 is contained in Table 1. The eight genome areas identified in the H5N1 genome are the nucleocapsid, matrix, pB2, neuraminidase, pA, NS, hemagglutinin, and pB1 gene areas. The graph in Figure 1 reveals that Replikin sequences were found to be most concentrated in the pB1 gene area of the H5N1 virus genome. When the inventors identified the Replikin Peak Gene (RPG) in the pB1 gene area, they discovered that an increase in Replikin Count in the RPG correlated with an increase in lethality in virus infectious. As such, “upregulation” of the RPG in H5N1 was observed in 2005 and 2006 as a significant increase in mean Replikin Count and standard deviation from the mean Replikin Count were observed and this upregulation correlated with increased lethality and virulence.

[000222] Table 1 provides mean Replikin Count and standard deviation from mean for publicly available sequences at PubMed for each of the eight gene areas in isolates of H5N1 between 2003 and 2006. Where no data is available for a given year, the year is not included in the table.

Table 1-H5N1 Influenza

Year	PubMed Accession Number-Replikin Count	No. of Isolates per year	Mean Replikin Count per year	S.D.	Significance
VI. Human H5N1 pB1 Area					
2003	BAE07199 15	1	2.0	0.0	prev p<.30
2004	ABE97546 15 ABE97545 15 ABE97544 15 ABE97543 15 ABE97542 15 ABE97541 15 ABE97540 15 ABE97539 15 ABE97538 15 ABE97537 15 ABE97536 15 AAV35116 15 AAV32652 15 AAV32644 15 ABE97535 15 ABE97534 15 ABE97533 15 ABE97532 15 ABE97531 15 ABE97530 15 ABE97529 15 ABE97528 15 ABE97527 15 ABE97526 15 ABE97525 15 ABE97524 15 ABE97523 15 ABE97522 15 ABE97521 15 ABE97520 15 ABE97519 15 ABE97518 15 ABE97517 15 ABE97516 15 ABE97515 15 ABE97514 15 ABE97513 15 ABE97512 15 ABE97511 15 ABE97509 15 ABE97508 15 ABE97507 15 ABE97506 15 ABE97505 15 ABE97504 15 ABE97503 15 ABE97502 15 ABE97501 15 ABE97500 15 ABE97499 15 ABE97498 15 ABE97497 15 ABE97496 15 ABE97495 15 ABE97494 15 ABE97493 15 ABE97492 15 ABE97491 15 ABE97490 15 ABE97489 15 ABE97488 15 ABE97487 15 ABE97486 15 ABE97485 15 ABE97484 15 ABE97483 15 ABE97482 15 ABE97481 15 ABE97480 15 ABE97479 15 ABE97478 15 ABE97477 15 ABE97476 15 ABE97475 15 ABE97474 15 ABE97473 15 ABE97472 15 ABE97471 15 ABE97470 15 ABE97469 15 ABE97468 15 ABE97467 15 ABE97466 15 ABE97465 15 ABE97464 15 ABE97463 15 ABE97462 15 ABE97461 15 ABE97460 15 ABE97510 15 AAV73985 3	91	2.0	0.1	low p<.001, prev p<.30

2005	<u>ABI36230 16 ABI36225 16 ABI36220 16 ABI36216 16</u> <u>ABI36214 15 ABI36209 15 ABI36009 15 ABI36000 14</u> <u>ABG78564 15 ABF56656 27 ABG78565 14 ABD16290 15</u> <u>ABC72648 15</u>	13	8.0	7.7	low p<.01, prev p<.005
2006	<u>ABK34974 14 ABL31777 16 ABL31774 21 ABL31763 21</u> <u>ABL31752 21 ABL31741 16 ABI49393 16 ABL07027 21</u> <u>ABL07016 16 ABL07005 16 ABI49404 16 ABI36472 16</u> <u>ABI36461 16 ABI36452 16 ABI36441 16 ABI36430 16</u> <u>ABI36420 16 ABI36408 16 ABI36397 16 ABI36386 16</u> <u>ABI36375 16 ABI36364 16 ABI36353 15 ABI36342 15</u> <u>ABI36331 15 ABI36320 15 ABI36309 16 ABI36303 16</u> <u>ABI36292 16 ABI36283 16 ABI36271 16 ABI36268 16</u> <u>ABI36265 16 ABI36261 16 ABI36257 16 ABI36252 16</u> <u>ABI36249 16 ABI36244 14 ABI36241 14 ABI36236 14</u> <u>ABI36232 14 ABI36195 16 ABI36184 16 ABI36174 16</u> <u>ABI36163 16 ABI36152 16 ABI36141 16 ABI16502 14</u>	48	16.1	5.7	low p<.001, prev p<.001
Human H5n1 Hemagglutinin Area					
2003	<u>BAE07201 22</u>	1	3.9	0.0	
2004	<u>AAS65618 23 AAS65615 22 ABE97634 22 ABE97633 22</u> <u>ABE97632 22 ABE97631 26 ABE97630 22 ABE97629 22</u> <u>ABE97628 22 ABE97627 22 ABE97626 21 ABE97625 22</u> <u>ABE97624 22 AAV34704 22 AAS89004 22 AAV32636 22</u> <u>AAV65826 22 ABE97623 22 ABE97622 22 ABE97621 22</u> <u>ABE97620 22 ABE97619 22 ABE97618 22 ABE97617 22</u> <u>ABE97616 22 ABE97615 22 ABE97614 22 ABE97613 20</u> <u>ABE97612 22 ABE97611 20 ABE97610 19 ABE97609 22</u> <u>ABE97608 22 ABE97607 22 ABE97606 22 ABE97605 22</u> <u>ABE97604 11 ABE97603 22 ABE97602 22 ABE97601 22</u> <u>ABE97600 22 ABE97599 22 ABE97598 22 ABE97597 22</u> <u>ABE97596 22 ABE97595 22 ABE97594 22 ABE97593 22</u> <u>ABE97592 22 ABE97591 18 ABE97590 18 ABE97589 18</u> <u>ABE97588 22 ABE97587 22 ABE97586 22 ABE97585 22</u> <u>ABE97584 22 ABE97583 21 ABE97582 22 ABE97581 22</u> <u>ABE97580 22 ABE97579 22 ABE97578 22 ABE97577 22</u> <u>ABE97576 22 ABE97575 24 ABE97574 22 ABE97573 22</u> <u>ABE97572 22 ABE97571 22 ABE97570 22 ABE97569 22</u> <u>ABE97568 22 ABE97567 22 ABE97566 22 ABE97565 22</u> <u>ABE97564 21 ABE97563 22 ABE97562 22 ABE97561 22</u> <u>ABE97560 22 ABE97559 22 ABE97558 22 ABE97557 19</u> <u>ABE97556 23 ABE97555 22 ABE97554 24 ABE97553 22</u> <u>ABE97552 22 ABE97551 22 ABE97550 22 ABE97549 22</u> <u>ABE97548 23 ABE97547 22 AAW59559 8 AAW59558 7</u> <u>AAW59556 4 AAW59554 6 AAW59552 16 AAW59550 16</u> <u>AAW59548 12 AAV73980 16 AAV73975 8 AAV73972 16</u>	104	3.9	0.3	low p<.001, prev p>.50
2005	<u>ABC59833 17 ABB00582 17 ABG78567 17 ABG78549 17</u> <u>ABG20478 21 ABG20476 14 ABC70167 17 ABB86287 17</u> <u>ABC72655 22 ABF56648 21 ABI36045 23 ABI36044 23</u> <u>ABI36043 27 ABI36042 27 ABI36041 23 ABI36040 21</u> <u>ABI36012 23 ABD16284 22</u>	18	4.1	0.3	low p<.001, prev p<.02
2006	<u>ABD66293 21 ABG23657 24 ABI16504 27 ABG20472 21</u> <u>ABG20468 23 ABJ90343 21 ABG45944 19 ABM54180 21</u> <u>ABM54179 21 ABL31780 23 ABI49415 25 ABI49396 23</u> <u>ABL31766 23 ABL31755 23 ABL31744 23 ABL07030 23</u> <u>ABL07019 23 ABL07008 23 ABK32782 20 ABK32780 21</u> <u>ABK32781 21 ABK32779 21 ABK32778 21 ABK32777 21</u> <u>ABK32776 21 ABK32775 20 ABI49407 23 ABI36480 23</u> <u>ABI36469 23 ABI36450 22 ABI36439 23 ABI36428 23</u> <u>ABI36423 22 ABI36406 22 ABI36395 23 ABI36384 23</u> <u>ABI36373 23 ABI36362 22 ABI36351 23 ABI36340 23</u>	64	4.1	0.3	low p<.001, prev p>.50

	<u>ABI36329 22 ABI36318 22 ABI36307 23 ABI36295 23</u> <u>ABI36286 23 ABI36275 23 ABI36198 23 ABI36187 22</u> <u>ABI36177 22 ABI36166 22 ABI36155 22 ABI36144 22</u> <u>ABI36057 24 ABI36056 23 ABI36055 23 ABI36054 23</u> <u>ABI36053 23 ABI36051 23 ABI36050 23 ABI36049 23</u> <u>ABI36048 23 ABI36047 23 ABI36046 23 ABI23979 21</u>				
Human H5N1 NS					
2004	<u>AAV35114 2 AAV35113 2</u>	2	1.7	0.0	low p<.002
2005	<u>ABF56654 27 ABF56653 27</u>	2	3.8	0.0	low p<.005
2006	<u>ABI16510 2 ABI16509 2</u>	2	1.7	0.0	low p<.002
Human H5N1 Pa					
2003	<u>BAE07200 27</u>	1	3.8	0.0	
2004	<u>ABL67793 27 ABL67792 27 ABL67791 27 ABL67780 31</u> <u>ABL67769 27 ABE97897 29 ABE97896 29 ABE97895 29</u> <u>ABE97894 29 ABE97893 29 ABE97892 29 ABE97891 29</u> <u>ABE97890 29 ABE97889 29 ABE97888 29 AAV35115 24</u> <u>AAV32651 27 AAV32643 27 AAW59560 1 AAW59557 10</u> <u>AAW59551 27 AAW59549 27 ABE97887 29 ABE97886 29</u> <u>ABE97885 29 ABE97884 29 ABE97883 29 ABE97882 29</u> <u>ABE97881 29 ABE97880 29 ABE97879 29 ABE97878 29</u> <u>ABE97877 29 ABE97876 29 ABE97875 29 ABE97874 29</u> <u>ABE97873 29 ABE97872 29 ABE97871 29 ABE97870 29</u> <u>ABE97869 29 ABE97868 29 ABE97867 29 ABE97866 29</u> <u>ABE97865 29 ABE97864 29 ABE97863 29 ABE97862 29</u> <u>ABE97861 29 ABE97860 29 ABE97859 29 ABE97858 29</u> <u>ABE97857 29 ABE97856 29 ABE97855 29 ABE97854 29</u> <u>ABE97853 29 ABE97852 29 ABE97851 29 ABE97850 29</u> <u>ABE97849 29 ABE97848 29 ABE97847 29 ABE97846 29</u> <u>ABE97845 29 ABE97844 29 ABE97843 29 ABE97842 29</u> <u>ABE97841 29 ABE97840 29 ABE97839 29 ABE97838 29</u> <u>ABE97837 29 ABE97836 29 ABE97835 29 ABE97834 29</u> <u>ABE97833 29 ABE97832 29 ABE97831 29 ABE97830 29</u> <u>ABE97829 29 ABE97828 29 ABE97827 29 ABE97826 29</u> <u>ABE97825 29 ABE97824 29 ABE97823 29 ABE97822 29</u> <u>ABE97821 29 ABE97820 29 ABE97819 29 ABE97818 29</u> <u>ABE97817 29 ABE97816 29 ABE97815 29 ABE97814 29</u> <u>ABE97813 29 ABE97812 29 ABE97811 29 AAV73982 7</u> <u>AAV73977 4 AAV73974 4</u>	102	4.0	0.5	low p<.001
2005	<u>ABL67826 27 ABL67815 28 ABL67804 28 ABI36229 33</u> <u>ABI36224 33 ABI36222 26 ABI36219 26 ABI36212 27</u> <u>ABI36011 27 ABI36002 27 ABG78563 22 ABG78562 22</u> <u>ABF56655 27 ABD16289 27 ABC72649 24</u>	15	3.8	0.4	low p>.50, prev p<.05
2006	<u>ABK34973 28 ABL31779 31 ABL31765 27 ABL31754 27</u> <u>ABL31743 27 ABI49414 27 ABI49395 27 ABL07029 27</u> <u>ABL07018 31 ABL07007 31 ABI49406 27 ABI36481 26</u> <u>ABI36470 26 ABI36451 27 ABI36440 27 ABI36429 27</u> <u>ABI36419 27 ABI36407 27 ABI36396 32 ABI36385 32</u> <u>ABI36374 32 ABI36363 27 ABI36352 27 ABI36341 27</u> <u>ABI36330 23 ABI36319 23 ABI36308 27 ABI36302 27</u> <u>ABI36291 27 ABI36282 27 ABI36273 27 ABI36267 27</u> <u>ABI36260 27 ABI36256 27 ABI36255 27 ABI36251 27</u> <u>ABI36247 27 ABI36240 27 ABI36238 27 ABI36235 27</u> <u>ABI36197 27 ABI36186 27 ABI36176 27 ABI36165 27</u> <u>ABI36154 27 ABI36143 27 ABI16503 24</u>	47	3.8	0.3	low p<.20, prev p>.50
human H5N1 Neuraminidase					
2003	<u>BAE07203 8</u>	1	1.7	0.0	
2004	<u>AAS65617 27 AAS65616 24 ABE97722 23 ABE97721 23</u> <u>ABE97720 23 ABE97719 23 ABE97718 23 ABE97717 23</u> <u>ABE97716 23 ABE97715 23 ABE97714 23 ABE97713 23</u>	100	5.0	0.3	low p<.001, prev p<.001

	<u>ABE97712</u> 23 <u>AAS89006</u> 18 <u>AAS89005</u> 16 <u>AAV32637</u> 24 <u>ABL67796</u> 18 <u>ABL67783</u> 18 <u>ABL67772</u> 18 <u>AAV65827</u> 18 <u>ABE97692</u> 23 <u>ABE97711</u> 23 <u>ABE97710</u> 23 <u>ABE97709</u> 23 <u>ABE97708</u> 23 <u>ABE97707</u> 23 <u>ABE97706</u> 23 <u>ABE97705</u> 23 <u>ABE97704</u> 23 <u>ABE97703</u> 23 <u>ABE97702</u> 23 <u>ABE97701</u> 23 <u>ABE97700</u> 23 <u>ABE97699</u> 23 <u>ABE97698</u> 23 <u>ABE97697</u> 23 <u>ABE97696</u> 23 <u>ABE97695</u> 23 <u>ABE97694</u> 23 <u>ABE97693</u> 23 <u>ABE97691</u> 23 <u>ABE97690</u> 23 <u>ABE97689</u> 23 <u>ABE97688</u> 23 <u>ABE97687</u> 23 <u>ABE97686</u> 23 <u>ABE97685</u> 23 <u>ABE97684</u> 23 <u>ABE97683</u> 23 <u>ABE97682</u> 23 <u>ABE97681</u> 23 <u>ABE97680</u> 23 <u>ABE97679</u> 23 <u>ABE97678</u> 23 <u>ABE97677</u> 23 <u>ABE97676</u> 23 <u>ABE97675</u> 23 <u>ABE97674</u> 23 <u>ABE97673</u> 23 <u>ABE97672</u> 23 <u>ABE97671</u> 23 <u>ABE97670</u> 23 <u>ABE97669</u> 23 <u>ABE97668</u> 23 <u>ABE97667</u> 23 <u>ABE97666</u> 23 <u>ABE97665</u> 23 <u>ABE97664</u> 23 <u>ABE97663</u> 23 <u>ABE97662</u> 23 <u>ABE97661</u> 23 <u>ABE97660</u> 23 <u>ABE97659</u> 23 <u>ABE97658</u> 23 <u>ABE97657</u> 23 <u>ABE97656</u> 23 <u>ABE97655</u> 23 <u>ABE97654</u> 23 <u>ABE97653</u> 23 <u>ABE97652</u> 23 <u>ABE97651</u> 23 <u>ABE97650</u> 23 <u>ABE97649</u> 23 <u>ABE97648</u> 23 <u>ABE97647</u> 23 <u>ABE97646</u> 23 <u>ABE97645</u> 23 <u>ABE97644</u> 23 <u>ABE97643</u> 23 <u>ABE97642</u> 23 <u>ABE97641</u> 23 <u>ABE97640</u> 23 <u>ABE97639</u> 23 <u>ABE97638</u> 23 <u>ABE97637</u> 23 <u>ABE97636</u> 23 <u>ABE97635</u> 23 <u>AAV73978</u> 8 <u>AAV73976</u> 8 <u>AAV73973</u> 18				
2005	<u>ABC59832</u> 6 <u>ABB00581</u> 8 <u>ABL67829</u> 18 <u>ABL67818</u> 18 <u>ABL67807</u> 18 <u>ABG78555</u> 8 <u>ABG78554</u> 6 <u>ABC72646</u> 16 <u>ABC70166</u> 8 <u>ABB86286</u> 8 <u>ABF56651</u> 27 <u>ABI36097</u> 13 <u>ABI36095</u> 13 <u>ABI36094</u> 13 <u>ABI36085</u> 13 <u>ABI36084</u> 13 <u>ABI36014</u> 13 <u>ABD16286</u> 20	18	3.0	0.8	low p<.001, prev p<.001
2006	<u>ABG23658</u> 12 <u>ABI16506</u> 13 <u>ABG20474</u> 7 <u>ABG20470</u> 8 <u>ABM68050</u> 8 <u>ABM68049</u> 8 <u>ABM68048</u> 8 <u>ABL31782</u> 15 <u>ABL31768</u> 13 <u>ABL31757</u> 13 <u>ABL31746</u> 13 <u>ABI49417</u> 13 <u>ABI49398</u> 13 <u>ABL07032</u> 12 <u>ABL07021</u> 15 <u>ABL07010</u> 15 <u>ABI49409</u> 13 <u>ABI36476</u> 13 <u>ABI36465</u> 13 <u>ABI36457</u> 13 <u>ABI36446</u> 13 <u>ABI36435</u> 13 <u>ABI36426</u> 13 <u>ABI36413</u> 13 <u>ABI36402</u> 13 <u>ABI36391</u> 13 <u>ABI36380</u> 13 <u>ABI36369</u> 13 <u>ABI36358</u> 13 <u>ABI36347</u> 13 <u>ABI36336</u> 13 <u>ABI36325</u> 13 <u>ABI36314</u> 13 <u>ABI36298</u> 13 <u>ABI36287</u> 13 <u>ABI36278</u> 13 <u>ABI36200</u> 13 <u>ABI36189</u> 13 <u>ABI36179</u> 13 <u>ABI36168</u> 13 <u>ABI36157</u> 13 <u>ABI36146</u> 13 <u>ABI36096</u> 13 <u>ABI36093</u> 13 <u>ABI36092</u> 13 <u>ABI36090</u> 13 <u>ABI36089</u> 11 <u>ABI36088</u> 11 <u>ABI36087</u> 13 <u>ABI36086</u> 13 <u>ABI23981</u> 7	51	2.8	0.4	low p<.001, prev p<.20
2007	<u>ABO38180</u> 8 <u>ABM90546</u> 13 <u>ABM90535</u> 13 <u>ABM90524</u> 13 <u>ABM90513</u> 13 <u>ABM90502</u> 15 <u>ABM90491</u> 15 <u>ABM90480</u> 15 <u>ABM90469</u> 15 <u>ABM90458</u> 15 <u>ABM90447</u> 15 <u>ABM90436</u> 15 <u>ABQ43810</u> 8	13	3.0	0.5	low p<.001, prev p<.05
Human H5N1 Pb2 Area					
2003	<u>BAE07198</u> 18 <u>BAE07197</u> 18	2	2.4	0.0	
2004	<u>ABL67788</u> 18 <u>ABL67777</u> 18 <u>ABL67766</u> 18 <u>ABF01751</u> 21 <u>ABF01750</u> 21 <u>ABF01749</u> 21 <u>ABF01748</u> 21 <u>ABF01747</u> 21 <u>ABF01746</u> 21 <u>ABF01745</u> 21 <u>ABF01744</u> 21 <u>ABF01743</u> 21 <u>AAV35117</u> 18 <u>AAV32653</u> 17 <u>AAV32645</u> 21 <u>ABF01742</u> 21 <u>ABF01741</u> 21 <u>ABF01740</u> 21 <u>ABF01739</u> 21 <u>ABF01738</u> 21 <u>ABF01737</u> 21 <u>ABF01736</u> 21 <u>ABF01735</u> 21 <u>ABF01734</u> 21 <u>ABF01733</u> 21 <u>ABF01732</u> 21 <u>ABF01731</u> 21 <u>ABF01728</u> 21 <u>ABF01726</u> 21 <u>ABF01724</u> 21 <u>ABF01722</u> 21 <u>ABF01709</u> 21 <u>ABF01708</u> 21 <u>ABF01706</u> 21 <u>ABF01705</u> 21 <u>ABF01704</u> 21 <u>ABF01703</u> 21 <u>ABF01702</u> 21 <u>ABF01701</u> 21 <u>ABF01700</u> 21 <u>ABF01699</u> 21 <u>ABF01698</u> 21 <u>ABF01697</u> 21 <u>ABF01696</u> 21 <u>ABF01695</u> 21 <u>ABF01694</u> 21 <u>ABF01693</u> 21 <u>ABF01692</u> 21 <u>ABF01691</u> 21 <u>ABF01690</u> 21 <u>ABF01689</u> 21 <u>ABF01688</u> 21	73	2.8	0.3	low p<.001, prev p<.001

	<u>ABF01687 21 ABF01686 21 ABF01685 21 ABF01684 21</u> <u>ABF01683 21 ABF01682 21 ABF01681 21 ABF01680 21</u> <u>ABF01679 21 ABF01678 21 ABF01677 21 ABF01676 21</u> <u>ABF01675 21 ABF01674 21 ABF01673 21 ABF01672 21</u> <u>ABF01670 21 ABF01669 21 ABF01668 21 ABF01667 21</u> <u>AAV73984 12</u>				
2005	<u>ABL67823 18 ABL67812 18 ABL67801 18 ABI36228 18</u> <u>ABI36227 18 ABI36223 18 ABI36218 18 ABI36213 18</u> <u>ABI36211 19 ABI36008 19 ABI35999 18 ABG78566 13</u> <u>ABF56657 27 ABG78548 13 ABD16291 18 ABC72647 18</u>	16	2.4	0.4	low p>.50, prev p<.001
2006	<u>ABL31776 18 ABL31773 18 ABL31762 18 ABL31751 18</u> <u>ABL31740 18 ABI49392 18 ABL07026 18 ABL07015 18</u> <u>ABL07004 18 ABI49403 18 ABI36482 18 ABI36471 18</u> <u>ABI36454 18 ABI36443 18 ABI36432 18 ABI36422 18</u> <u>ABI36410 18 ABI36399 18 ABI36388 18 ABI36377 18</u> <u>ABI36366 18 ABI36355 17 ABI36344 17 ABI36333 18</u> <u>ABI36322 18 ABI36311 18 ABI36306 18 ABI36305 18</u> <u>ABI36294 18 ABI36285 18 ABI36274 18 ABI36270 18</u> <u>ABI36264 18 ABI36263 18 ABI36259 18 ABI36254 18</u> <u>ABI36248 18 ABI36246 18 ABI36243 18 ABI36239 18</u> <u>ABI36234 18 ABI36194 18 ABI36173 18 ABI36162 18</u> <u>ABI36151 18 ABI36140 18 ABG23659 21 ABI16501 18</u>	48	2.4	0.1	low p>.50, prev p>.50
human H5N1 matrix area					
2003	<u>BAE07204 3</u>	1	1.2	0.0	prev p<.05
2004	<u>AAV35110 3 AAV32646 3 AAV32638 3 AAV35111 2</u> <u>AAV32647 2 AAV32639 2 AAV73981 3</u>	7	1.6	0.4	low p<.04
2005	<u>ABF01926 3 ABF01924 3 ABF01922 3 ABF01920 3</u> <u>ABF01918 3 ABF01916 3 ABF01914 3 ABF01912 3</u> <u>ABF01910 3 ABF01908 3 ABF01906 3 ABF01927 2</u> <u>ABF01925 2 ABF01923 1 ABF01921 2 ABF01919 2</u> <u>ABF01917 2 ABF01915 2 ABF01913 2 ABF01911 2</u> <u>ABF01909 2 ABF01907 2 ABG78552 3 ABG78551 3</u> <u>ABG20477 3 ABF56649 3 ABI36082 5 ABI36080 3 ABI36060</u> <u>5 ABI36058 3 ABI36015 3 ABI36004 3 ABF01904 3</u> <u>ABF01902 3 ABF01900 3 ABF01898 3 ABF01896 3</u> <u>ABF01894 3 ABF01892 3 ABF01890 3 ABF01888 3</u> <u>ABF01886 3 ABF01884 3 ABF01882 3 ABF01880 3</u> <u>ABF01878 3 ABF01876 3 ABF01874 3 ABF01872 3</u> <u>ABF01870 3 ABF01868 3 ABF01866 3 ABF01864 3</u> <u>ABF01862 3 ABF01860 3 ABF01858 3 ABF01856 3</u> <u>ABF01854 3 ABF01852 3 ABF01850 3 ABF01848 3</u> <u>ABF01846 3 ABF01844 3 ABF01842 3 ABF01840 3</u> <u>ABF01838 3 ABF01836 3 ABF01834 3 ABF01832 3</u> <u>ABF01830 3 ABF01828 3 ABF01826 3 ABF01824 3</u> <u>ABF01822 3 ABF01820 3 ABF01818 3 ABF01816 3</u> <u>ABF01814 3 ABF01812 3 ABF01810 3 ABF01808 3</u> <u>ABF01806 3 ABF01804 3 ABF01802 3 ABF01800 3</u> <u>ABF01798 3 ABF01796 3 ABF01794 3 ABF01792 3</u> <u>ABF01790 3 ABF01788 3 ABF01786 3 ABF01784 3</u> <u>ABF01782 3 ABF01780 3 ABF01778 3 ABF01776 3</u> <u>ABF01774 3 ABF01772 3 ABF01770 3 ABF01768 3</u> <u>ABF01766 3 ABF01764 3 ABF01762 3 ABF01760 3</u> <u>ABF01758 3 ABF01756 3 ABF01754 3 ABF01752 3</u> <u>ABD16285 3 ABC72651 3 ABF56650 2 ABI36083 2</u> <u>ABI36081 4 ABI36061 2 ABI36059 4 ABI36016 4 ABI36005 4</u> <u>ABF01905 2 ABF01903 2 ABF01901 1 ABF01899 2</u> <u>ABF01897 2 ABF01895 2 ABF01893 2 ABF01891 2</u> <u>ABF01889 2 ABF01887 2 ABF01885 2 ABF01883 1</u> <u>ABF01881 1 ABF01879 2 ABF01877 2 ABF01875 2</u>	198	1.6	0.7	low p<.001, prev p>.50

	<u>ABF01873</u> 2 <u>ABF01871</u> 2 <u>ABF01869</u> 2 <u>ABF01867</u> 2 <u>ABF01865</u> 2 <u>ABF01863</u> 2 <u>ABF01861</u> 1 <u>ABF01859</u> 2 <u>ABF01857</u> 2 <u>ABF01855</u> 2 <u>ABF01853</u> 2 <u>ABF01851</u> 2 <u>ABF01849</u> 2 <u>ABF01847</u> 2 <u>ABF01845</u> 2 <u>ABF01843</u> 2 <u>ABF01841</u> 2 <u>ABF01839</u> 2 <u>ABF01837</u> 2 <u>ABF01835</u> 2 <u>ABF01833</u> 2 <u>ABF01831</u> 1 <u>ABF01829</u> 2 <u>ABF01827</u> 1 <u>ABF01825</u> 2 <u>ABF01823</u> 2 <u>ABF01821</u> 2 <u>ABF01819</u> 2 <u>ABF01817</u> 2 <u>ABF01815</u> 2 <u>ABF01813</u> 2 <u>ABF01811</u> 2 <u>ABF01809</u> 2 <u>ABF01807</u> 2 <u>ABF01805</u> 2 <u>ABF01803</u> 2 <u>ABF01801</u> 2 <u>ABF01799</u> 2 <u>ABF01797</u> 2 <u>ABF01795</u> 1 <u>ABF01793</u> 1 <u>ABF01791</u> 4 <u>ABF01789</u> 2 <u>ABF01787</u> 2 <u>ABF01785</u> 4 <u>ABF01783</u> 2 <u>ABF01781</u> 1 <u>ABF01779</u> 2 <u>ABF01777</u> 2 <u>ABF01775</u> 1 <u>ABF01773</u> 2 <u>ABF01771</u> 2 <u>ABF01769</u> 4 <u>ABF01767</u> 2 <u>ABF01765</u> 2 <u>ABF01763</u> 1 <u>ABF01761</u> 1 <u>ABF01759</u> 1 <u>ABF01757</u> 1 <u>ABF01755</u> 4 <u>ABF01753</u> 2 <u>ABC72652</u> 2 <u>ABG78553</u> 1 <u>ABG78550</u> 1				
2006	<u>ABG23663</u> 3 <u>ABI16508</u> 3 <u>ABG20473</u> 3 <u>ABG20469</u> 3 <u>ABG23664</u> 2 <u>ABL31783</u> 5 <u>ABL31769</u> 5 <u>ABL31758</u> 5 <u>ABL31747</u> 5 <u>ABI49418</u> 5 <u>ABI49399</u> 5 <u>ABL07033</u> 5 <u>ABL07022</u> <u>5</u> <u>ABL07011</u> 5 <u>ABI36062</u> 5 <u>ABI49410</u> 5 <u>ABI36483</u> 5 <u>ABI36474</u> 5 <u>ABI36463</u> 5 <u>ABI36455</u> 5 <u>ABI36444</u> 5 <u>ABI36433</u> 5 <u>ABI36424</u> 3 <u>ABI36411</u> 3 <u>ABI36400</u> 5 <u>ABI36389</u> 5 <u>ABI36378</u> 5 <u>ABI36367</u> 3 <u>ABI36356</u> 3 <u>ABI36345</u> 3 <u>ABI36334</u> 3 <u>ABI36323</u> 3 <u>ABI36312</u> 5 <u>ABI36296</u> 5 <u>ABI36276</u> 5 <u>ABI36201</u> 5 <u>ABI36190</u> 3 <u>ABI36180</u> 3 <u>ABI36169</u> 3 <u>ABI36158</u> 3 <u>ABI36147</u> 3 <u>ABI36078</u> 3 <u>ABI36076</u> 3 <u>ABI36074</u> 3 <u>ABI36070</u> 5 <u>ABI36068</u> 5 <u>ABI36066</u> 5 <u>ABI36064</u> 5 <u>ABI23980</u> 3 <u>ABL31784</u> 2 <u>ABL31770</u> 2 <u>ABL31759</u> <u>2</u> <u>ABL31748</u> 2 <u>ABI49419</u> 2 <u>ABI49400</u> 2 <u>ABL07034</u> 2 <u>ABL07023</u> 2 <u>ABL07012</u> 2 <u>ABI36063</u> 2 <u>ABI49411</u> 2 <u>ABI36484</u> <u>2</u> <u>ABI36475</u> 2 <u>ABI36464</u> 2 <u>ABI36456</u> 2 <u>ABI36445</u> 2 <u>ABI36434</u> <u>2</u> <u>ABI36425</u> 2 <u>ABI36412</u> 2 <u>ABI36401</u> 2 <u>ABI36390</u> 2 <u>ABI36379</u> <u>2</u> <u>ABI36368</u> 2 <u>ABI36357</u> 4 <u>ABI36346</u> 4 <u>ABI36335</u> 4 <u>ABI36324</u> <u>4</u> <u>ABI36313</u> 2 <u>ABI36297</u> 2 <u>ABI36277</u> 2 <u>ABI36202</u> 2 <u>ABI36191</u> <u>2</u> <u>ABI36181</u> 2 <u>ABI36170</u> 2 <u>ABI36159</u> 2 <u>ABI36148</u> 2 <u>ABI36079</u> <u>4</u> <u>ABI36077</u> 4 <u>ABI36075</u> 4 <u>ABI36071</u> 2 <u>ABI36069</u> 2 <u>ABI36067</u> <u>2</u> <u>ABI36065</u> 2 <u>ABI16507</u> 2	93	2.0	0.7	low p<.001, prev p<.001
Human H5N1 Nucleocapsid Area					
2004	<u>ABL67795</u> 2 <u>ABL67782</u> 2 <u>ABL67771</u> 2 <u>AAV35112</u> 2	4	0.4	0.0	
2005	<u>ABL67828</u> 2 <u>ABL67817</u> 2 <u>ABL67806</u> 2 <u>ABI36110</u> 2 <u>ABI36109</u> <u>2</u> <u>ABI36108</u> 2 <u>ABI36099</u> 2 <u>ABI36098</u> 2 <u>ABI36013</u> 2 <u>ABI36003</u> <u>2</u>	10	0.4	0.0	low p<.001, prev p>.50
2006	<u>ABL31781</u> 2 <u>ABL31767</u> 2 <u>ABL31756</u> 2 <u>ABL31745</u> 2 <u>ABI49416</u> 2 <u>ABI49397</u> 2 <u>ABL07031</u> 2 <u>ABL07020</u> 2 <u>ABL07009</u> <u>2</u> <u>ABI49408</u> 2 <u>ABI36477</u> 2 <u>ABI36466</u> 2 <u>ABI36458</u> 2 <u>ABI36447</u> <u>2</u> <u>ABI36436</u> 2 <u>ABI36427</u> 2 <u>ABI36414</u> 2 <u>ABI36403</u> 2 <u>ABI36392</u> <u>2</u> <u>ABI36381</u> 2 <u>ABI36370</u> 2 <u>ABI36359</u> 2 <u>ABI36348</u> 2 <u>ABI36337</u> <u>2</u> <u>ABI36326</u> 2 <u>ABI36315</u> 2 <u>ABI36299</u> 2 <u>ABI36288</u> 2 <u>ABI36279</u> <u>2</u> <u>ABI36199</u> 2 <u>ABI36188</u> 2 <u>ABI36178</u> 2 <u>ABI36167</u> 2 <u>ABI36156</u> <u>2</u> <u>ABI36145</u> 2 <u>ABI36107</u> 2 <u>ABI36106</u> 2 <u>ABI36105</u> 2 <u>ABI36104</u> <u>2</u> <u>ABI36103</u> 2 <u>ABI36102</u> 2 <u>ABI36101</u> 2 <u>ABI36100</u> 2	43	0.4	0.0	low p<.001, prev p>.50

[000223] Analysis of the number of Replikin sequences present in the areas in the genome adjacent to the pB1 area revealed no more than a two-fold increase in Replikin Count in the seven other areas of the genome as compared to an eight-fold increase (p<0.001) in the Replikin Count in the pB1 area between years 2003 to 2006. The specificity of the localization of the upregulated RPG in the pB1 area is underlined by the fact that other parts

of the polymerase gene area of which pB1 is a part, namely, the pB2 and pA gene areas do not have the same amount of increase in Replikin Count even though the gene areas are immediately adjacent to pB1.

[000224] In the illustration, the standard deviation of the means is shown in light gray columns on top of the means, rather than in the usual 'T' symbols, to emphasize the diverse expanding virus population with regard to the Replikin Count. As Replikin Count increases in a population, a diversity of Replikin Counts may be observed as the lethality and virulence of the virus increases. An increasing standard deviation within a virus population is, therefore, itself an index of viral outbreaks. Here in Figure 1 and in the figures that follow, small standard deviations from mean Replikin Count are seen to accompany quiescent inter-outbreak periods of the virus.

[000225] Examples 1- 3 are provided below as examples of analysis of Replikin Peak Genes in sequences publicly available in accession numbers at PubMed. Examples 2 and 3 illustrate how identification of a Replikin Peak Gene allows for a magnification of the effect of increases in Replikin Count in an isolate where the increase may be correlated with and predict increases in virulence and lethality. For example, Example 2 provides a 2003 isolate of H5N1 from Hong Kong with a whole pB1 gene area (SEQ ID NO: 1683) Replikin Count of 2.0 and an RPG Replikin Count of 14.6. Example 3 provides a 2006 isolate from Indonesia with a whole pB1 gene area Replikin Count of 17.8 and an RPG Replikin Count of 22.5. Indonesia experienced a highly lethal outbreak of H5N1 with evidence of human to human transmission in 2007. The high Replikin Counts in isolates from Indonesia in 2006 allowed the inventors to prospectively predict the lethal Indonesian outbreak.

[000226] The Replikin count of the whole genome from the 2006 Indonesian isolate demonstrates a significant increase as compared to the 2003 Hong Kong isolate. The isolation of the Replikin peak gene (RPG) area, that is the area of the genome which shows the highest concentration (count) of continuous Replikins per 100 amino acids, magnifies the effect. For this reason, whole genome counts are used for first approximations of Replikin count increases, and where more detailed specific gene areas or open reading frame data is not available. However, when available, the RPG is used for more definitive "higher power" examinations. This is illustrated in Figures 1-4.

B. Increasing Replikin Count in RPG of Influenza A Associated with Pandemics and Lethality

[000227] The inventors have now associated an increase in Replikin Count in the RPG of *Influenza A* virus with pandemics, epidemics and lethal outbreaks of influenza. Figure 2 illustrates an increase in Replikin Count before and accompanying each *Influenza A* pandemic and outbreak since 1918 and low Replikin Counts during quiescent periods of *Influenza A* infection and continually in non-lethal *Influenza B*. The graph provides annual Replikin Counts from 1917-2007 for all Replikin Peak Genes isolated *in silico* in the pB1 gene area of influenza strains having amino acid or nucleic acid sequences publicly available at PubMed. The total number of sequences analyzed for the data is 14,227. The Replikin Count of each influenza-*in-silico* isolate was obtained separately and objectively through time for each species by computer software (FluForecast[®], available through Replikins LLC, Boston, MA). The software queried publicly available sequences at www.pubmed.com. The software measures solely the number of Replikins per 100 amino acids in the publicly available sequences and provides a mean Replikin Count with standard deviation from the mean for all isolates available in a given strain of influenza in a given year.

[000228] Over a ninety year period, the graph in Figure 2 demonstrates an increase in Replikin Count before and accompanying each *Influenza A* pandemic and outbreak, namely, the 1918 H1N1 pandemic, the 1930's H1N1 epidemic, the 1957 H2N2 pandemic, the 1968 H3N1 pandemic, the 1977-78 H3N2 outbreaks and the H5N1 outbreaks of 1997, 2001-2004 and 2007. In Figure 2, p values at <0.001 are supportive of the significance of the differences between the pandemic and epidemic groups on the one hand and clinically quiescent periods on the other hand.

[000229] Over the same ninety year period, constant low Replikin Counts of less than four may be observed during quiescent non-lethal periods between epidemic outbreaks in all 3 pandemic strains of *Influenza A* including H1N1, H2N2, H3N2 and low Replikin Counts of less than four may also be observed in human H5N1 and H5N1 in chicken in relatively quiescent periods. Low Replikin Counts may likewise be observed in non-lethal *Influenza B* through the entire period of observation. In particular, the absence of increase in Replikin Count above five in *Influenza B* corresponds to the absence of any observed lethal outbreak. As such, during the observation period, *Influenza B* is always non-lethal. This absence of Replikin Counts of five or more in non-lethal *Influenza B* provides an important control for the study of Replikin Count as a correlate of lethality. In *Influenza A*, an increase in Replikin

Count indicates an increase in lethality and a decrease in Replikin Count indicates a decrease in lethality.

[000230] Analysis of publicly available sequences for isolates of *Influenza B* virus between 1940 and 2007 is provided below in Table 2. Years in which not data was available are not included in the table.

Table 2 – Influenza B

Year	PubMed Accession Number-Replikin Count	No. of Isolates per year	Mean Replikin Count per year	S.D.	Significance
1940	NP_056657 14 ABG85176 14 ABG85165 14 P07832 14 BAA00002 14 AAA43767 14 AAF06886 16 AAF06851 10 NP_056659 16 NP_056658 10	10	1.8	0.3	low p<.001
1966	AAF89738 25 P13872 14 P13871 14 Q9IMP4 25	4	2.6	0.8	low p<.10
1969	ABQ81851 14	1	1.9	0.0	prev p<.10
1972	ABF21251 18 ABF21252 18	2	2.4	0.0	
1979	AAF06873 14 AAF06856 10	2	1.6	0.4	low p<.05, prev p<.10
1984	AAF06870 14 AAF06888 16 AAF06853 10	3	1.8	0.5	low p<.02, prev p>.50
1985	AAF06868 14 AAF06885 16 AAF06850 10	3	1.8	0.5	low p<.02, prev p>.50
1987	ABL77253 14 AAF06874 17 AAF06891 14 AAF06857 10	4	1.8	0.4	low p<.002, prev p>.50
1988	ABN50611 14 ABL77264 14 AAF06875 14 AAF06860 14 AAF06892 17 AAF06877 14 AAF06858 10 AAF06842 10 P12236 6	9	1.8	0.3	low p<.001, prev p>.50
1989	ABL77275 14 P21796 5	2	1.8	0.1	low p<.01, prev p>.50
1990	ABN58670 16 ABN50633 14 ABN50622 14 ABL76703 14 AAB72043 14 O36430 14	6	1.9	0.1	low p<.001, prev p<.20
1991	ABN50644 27 ABN51204 10 ABN51193 10 ABL77286 14	4	2.0	1.1	low p<.05, prev p>.50
1992	ABN50655 14 ABL77308 14 ABL77297 14	3	1.9	0.0	prev p>.50
1993	CAA05486 40 ABN50666 14 ABL77341 14 ABL77330 14 ABL77319 14 AAF06869 14 AAF06865 14 AAF06861 14 CAG96502 14 AAF06887 17 AAF06882 14 AAF06878 17 AAF06852 10 AAF06846 10 AAF06843 10	15	2.1	1.0	low p<.001, prev p<.40
1994	AAU94857 14 AAU94856 14 AAU94855 14 AAU94854 14 AAU94853 14 AAU94852 14 AAU94851 14 AAU94850 14 AAU94849 14 AAU94848 14 AAU94847 14 AAU94846 14 AAU94845 14 AAF89734 14 ABR16004 14 ABN50721 14 ABN50710 14 ABN50699 14 ABN50688 14 ABN50677 14 ABL77363 14 ABL77352 14 ABL77000 14 AAF06866 15 AAF06864 14 AAF06883 17 AAF06881 14 AAF06848 10 AAF06847 10	29	1.8	0.2	low p<.001, prev p<.40
1995	ABR16015 14 ABN50732 14 ABL77385 14 ABL77374 14 ABL76945 14	5	1.9	0.0	prev p>.50
1996	ABL76967 14 ABL76714 14	2	1.9	0.0	
1997	ABN59454 14 ABN50413 14 ABN50402 14 ABN50391 14 ABL76978 14 ABL76285 14 ABL76274 14 ABL76263 14 AAK95906 14 AAF06867 14 AAF06862 14 CAG96500 14 ABI96727 14 ABI96738 13 AAP22114 14 AAP22106 14	20	1.8	0.2	low p<.001, prev p>.50

	<u>AAF06884</u> 17 <u>AAF06879</u> 17 <u>AAF06849</u> 10 <u>AAF06844</u> 10				
1998	<u>ABN50743</u> 14 <u>ABN50512</u> 14 <u>ABN50457</u> 14 <u>ABL77022</u> 14 <u>ABL77011</u> 14 <u>ABL76989</u> 14 <u>ABL76956</u> 14 <u>ABL76780</u> 14 <u>ABL76769</u> 14 <u>ABL76296</u> 14 <u>AAF06876</u> 14 <u>AAF06872</u> 14 <u>AAF06871</u> 14 <u>AAF06863</u> 14 <u>AAU00993</u> 14 <u>AAF06893</u> 17 <u>AAF06890</u> 17 <u>AAF06889</u> 17 <u>AAF06880</u> 17 <u>AAF06859</u> 10 <u>AAF06855</u> 10 <u>AAF06854</u> 10 <u>AAF06845</u> 10	23	1.8	0.3	low p<.001, prev p>.50
1999	<u>ABL77055</u> 14 <u>ABL77044</u> 14 <u>ABL77033</u> 14 <u>ABL76813</u> 14 <u>ABL76802</u> 14 <u>ABL76362</u> 14 <u>ABL76351</u> 14 <u>ABL76340</u> 14 <u>ABL76329</u> 14 <u>ABL76318</u> 14 <u>ABL76307</u> 14 <u>ABQ81840</u> 14 <u>CAG96499</u> 14 <u>ABI94772</u> 20	14	1.9	0.2	low p<.001, prev p<.40
2000	<u>ABL77110</u> 14 <u>ABL77066</u> 14 <u>ABL76901</u> 14 <u>ABL76890</u> 14 <u>ABL76879</u> 14 <u>ABL76868</u> 14 <u>ABL76857</u> 14 <u>ABL76824</u> 14 <u>ABL76791</u> 14 <u>ABL76395</u> 14 <u>ABL76384</u> 14 <u>ABL76373</u> 14 <u>ABL84349</u> 14 <u>CAG96513</u> 14 <u>AAT69423</u> 14	15	1.9	0.0	low p<.001, prev p<.30
2001	<u>ABR15982</u> 14 <u>ABO72385</u> 14 <u>ABN50600</u> 14 <u>ABN50567</u> 14 <u>ABN50534</u> 14 <u>ABN50523</u> 14 <u>ABN50490</u> 14 <u>ABN50479</u> 14 <u>ABN50435</u> 14 <u>ABN50424</u> 14 <u>ABL77187</u> 14 <u>ABL77143</u> 14 <u>ABL77099</u> 14 <u>ABL77088</u> 14 <u>ABL77077</u> 14 <u>ABL76417</u> 14 <u>ABL76406</u> 14 <u>AAT69445</u> 14 <u>CAG96504</u> 14 <u>AAT69434</u> 14 <u>CAG96509</u> 14 <u>ABJ09524</u> 14 <u>ABJ09472</u> 15 <u>ABJ15707</u> 16 <u>ABI96775</u> 20 <u>ABI96695</u> 16	26	2.0	0.3	low p<.001, prev p<.05
2002	<u>ABN50754</u> 14 <u>ABN50578</u> 14 <u>ABN50556</u> 14 <u>ABN50545</u> 14 <u>ABL77396</u> 14 <u>ABL77176</u> 14 <u>ABL77165</u> 14 <u>ABL77154</u> 14 <u>ABL77132</u> 14 <u>ABL76483</u> 14 <u>ABL76472</u> 14 <u>ABL76461</u> 14 <u>ABL76450</u> 14 <u>ABL76439</u> 10 <u>ABL76428</u> 14 <u>CAG96515</u> 14 <u>CAG96514</u> 14 <u>CAG96511</u> 14 <u>CAG96510</u> 14 <u>CAG96503</u> 14 <u>CAG96501</u> 14 <u>ABI97312</u> 16 <u>ABJ09504</u> 17 <u>ABJ09486</u> 15 <u>ABI98925</u> 16 <u>ABI96765</u> 15 <u>ABK00110</u> 15 <u>ABI97340</u> 15 <u>ABI97331</u> 21 <u>ABI97322</u> 21 <u>ABI94786</u> 14 <u>ABI94737</u> 11	32	2.0	0.3	low p<.001, prev p>.50
2003	<u>ABR15993</u> 14 <u>ABN50589</u> 14 <u>ABL77209</u> 14 <u>ABL77198</u> 14 <u>ABL76835</u> 14 <u>ABL76626</u> 14 <u>ABL76615</u> 14 <u>ABL76604</u> 14 <u>ABL76593</u> 14 <u>ABL76582</u> 14 <u>ABL76571</u> 14 <u>ABL76560</u> 14 <u>ABL76549</u> 14 <u>ABL76538</u> 14 <u>ABL76527</u> 14 <u>ABL76516</u> 14 <u>ABL76505</u> 14 <u>ABL76494</u> 14 <u>CAG96520</u> 14 <u>CAG96519</u> 14 <u>CAG96517</u> 14 <u>CAG96512</u> 14 <u>CAG96508</u> 14 <u>CAG96507</u> 14 <u>CAG96506</u> 14 <u>CAG96505</u> 14 <u>ABJ98940</u> 16 <u>ABJ09534</u> 15 <u>ABJ09509</u> 15 <u>ABJ09482</u> 15 <u>ABI98936</u> 14 <u>ABI98912</u> 15 <u>ABI98908</u> 15 <u>ABK00142</u> 16 <u>ABK00088</u> 8 <u>ABJ80591</u> 8 <u>ABJ52572</u> 8 <u>ABJ52553</u> 16 <u>ABI97342</u> 14	39	1.9	0.2	low p<.001, prev p<.10
2004	<u>ABN50468</u> 14 <u>ABL77231</u> 14 <u>ABL77220</u> 14 <u>ABL77121</u> 14 <u>ABL76923</u> 14 <u>ABL76912</u> 14 <u>ABL76846</u> 14 <u>ABL76758</u> 14 <u>ABL76637</u> 14 <u>CAG96518</u> 14 <u>CAG96516</u> 10 <u>AAT70178</u> 14 <u>ABJ16471</u> 15 <u>ABJ09543</u> 14 <u>AAT78590</u> 14 <u>ABK00130</u> 15 <u>ABJ52559</u> 14 <u>ABJ15718</u> 15 <u>ABI97302</u> 13 <u>ABI96707</u> 13	20	1.8	0.1	low p<.001, prev p<.30
2005	<u>ABN50501</u> 14 <u>ABN50446</u> 14 <u>ABL77242</u> 14 <u>ABL76934</u> 14 <u>ABL76692</u> 14 <u>ABL76681</u> 14 <u>ABL76670</u> 14 <u>ABL76659</u> 14 <u>ABL76648</u> 14 <u>ABI96712</u> 14	10	1.9	0.1	low p<.001, prev p<.20
2006	<u>ABL76747</u> 14 <u>ABL76736</u> 14 <u>ABL76725</u> 14 <u>ABR16026</u> 14	4	1.9	0.0	low p<.001, prev p<.30
2007	<u>ZP_01998985</u> 50 <u>EDN71014</u> 50	2	3.1	0.0	prev p<.001

[000231] While the Replikin Count in non-lethal *Influenza B* remains remarkable constant, the Replikin Count in *Influenza A* shows significant variation that correlates with outbreaks, epidemics and pandemics. For example, mean Replikin Count of the RPG in

Figure 2 may be observed to be greater in the 1918 (H1N1) pandemic than in the 1957 (H2N2) and 1968 (H3N2) pandemics in approximate scale to the mortality rates observed in those three pandemics. The 1918 pandemic is thought to have resulted in 675,000 deaths in the U.S. and 50 million deaths globally. A Replikin Count of 19 in 1917 may be observed in Figure 2. The 1957 pandemic is thought to have resulted in 70,000 deaths in the U.S. and 1-2 million deaths globally. A Replikin Count of 4 in 1957 with a standard deviation of 4.9 may be observed in Figure 2. The 1968 pandemic is thought to have resulted in 34,000 deaths in the U.S. with 700,000 deaths globally. A Replikin Count of 7.2 in 1968 with a standard deviation of 8 may be observed in Figure 2.

[000232] The dominance of H5N1 over other *Influenza A* strains between 1990 and 2007 is also evident in Figure 2. This dominance in mean Replikin Count is reflected in high global lethality in birds during that time period and outbreaks resulting in human lethality in 1997, 2001-2004 and continuing lethality and possible human-to-human transmission in 2007. The counts for human H5N1 in 2004 through 2007 are increasing and approaching the mean Replikin Count level of the 1918 H1N1 pandemic. The mean Replikin Count of the RPG in chicken H5N1 is also observed to increase to a lesser degree over this time period and observed to decrease in 2007. The standard deviation of the means (SD) for all strains is shown in light grey columns with caps, on top of the column for the mean Replikin count, and emphasizes the broad distribution of Replikin Counts in the RPG of the expanding virus population. This broad distribution of Replikin Counts illustrates rapid changes in distribution of Replikin Counts during the rapid replication that is associated with virus outbreaks. During quiescent periods, the standard deviation is observed to be approximately 10% or less of the mean. In contrast, when an outbreak develops, the standard deviation is observed to be 50% or greater than the mean (the same phenomenon is observed in Figure 7 for H3N8 equine encephalitis). The data for mean Replikin Count in human H5N1 for 2005, 2006 and 2007 suggest that the current epidemic is not over. For example, in each of the H1N1, H2N2 and H3N2 pandemics, a decline may be observed in Replikin Count prior to the end of the outbreak. This decline prior to end of an outbreak was also seen in the SARS outbreak of 2003. *See* Figure 9. As such, since observed mean Replikin Count has not yet begun to decline in H5N1, the current epidemic is expected to continue.

[000233] Figure 3 illustrates successive “emerging” strains of influenza virus between 1930 and 2007. Mean Replikin Counts per year of isolation of various strains of influenza are provided for the polymerase area (marked with circles), the pB1 area (marked with

triangles), and the pB1-F2 area (marked with squares). Data for H1N1 and H3N2 continue through 2007. Gaps represent years where no data was available on these genomic areas on PubMed.

[000234] Dramatic increases in Replikin Count may be observed just before outbreak in the rebound epidemic of H1N1 beginning in the 1930's, in the pandemics of H2N2 and H2N3, which occurred in 1957 and 1968, respectively, and the outbreaks of H5N1 between 1997 and 2007. The largest increase in Replikin Count may be observed in the pB1-F2 area of the genome, which is contained within the pB1 area of the genome and contains an identified RPG (*e.g.*, SEQ ID NO: 1723). The next largest increase in Replikin Count may be observed in the pB1 area of the genome, which is contained in the polymerase area of the genome. The smallest increase in Replikin Count may be observed in the polymerase area of the genome. It may be observed, therefore, that the Replikin Count becomes magnified as measured within the pB1 area as compared to the polymerase area and within the pB1-F2 area as compared to the pB1 area.

[000235] As in Figure 2, Figure 3 illustrate the constancy of Replikin Counts during quiescent periods of the strain, and a marked increase in Replikin Peak Gene Replikin Counts one year in advance of, or simultaneous with, outbreaks of specific strains. Figures 2 and 3 demonstrate that neither increases in Replikin Count nor outbreaks occur in more than one influenza strain at the same time. The figures further demonstrate a "rise" of H3N2 in 1968 that occurs simultaneous with a "fall" of H2N2.

C. Replikin Count in RPG of H5N1 Directly Correlates with Human Mortality

[000236] The inventors have now demonstrated that increased Replikin Counts in the RPG of H5N1 influenza virus (*e.g.*, SEQ ID NO: 1684) may be directly correlated with human mortality. Figure 4 illustrates the relationship of Replikin Count of the Replikin Peak Gene in human H5N1 to percent human mortality between 2003 and 2007 in human cases of H5N1 infection. An increase in Replikin Count in the Replikin Peak Gene of H5N1 is observed to be quantitatively related to higher mortality in the host. The Replikin Peak Gene in human H5N1 is the pB1 gene area, which has the highest concentration of continuous Replikin sequences in publicly available sequences of the H5N1 genome.

[000237] Magnification of Replikin Count may be observed in Figure 4 when the mean Replikin Count in the whole virus a given year is compared with the mean Replikin Count in the pB1 gene area (identified as the Replikin Peak Gene area of the virus). For example, annual mean Replikin Count in the whole genome increased 33% from 2005 to 2007 while

annual mean Replikin Count in the Replikin Peak Gene (pB1 gene area) increased nine-fold from 2003 to 2007 and 222% from 2005 to 2007 with a statistical p value less than 0.001. Annual percent mortality of human H5N1 cases increased approximately 100% from 2005 to 2007. The 2007 data, while unfortunately sparse because of withholding of data by some countries, does not indicate a decrease in whole genome Replikin count or in RPG Replikin Count. A significant decrease in Replikin Count would typically signal the end of an outbreak or epidemic. *See, e.g.*, SARS coronavirus data in Figure 9. No such decrease has yet to be observed.

[000238] Figures 16 and 17 likewise demonstrate that increased Replikin Counts in the RPG of H5N1 is more strongly correlated with lethality in a given year than increased Replikin Counts in other portions of the H5N1 genome. The data for Figures 16 and 17 are contained in Table 3 below.

Table 3 – H5N1 Replikin Concentration and Human Mortality

	2003	2004	2005	2006	2007
H5N1 Whole Virus Replikin Count in Humans	2.2+/- 1.2	2.4+/- 1.4	2.3+/- 2.6	3.8+/- 4.6	3.7+/- 4.5
H5N1 Polymerase Replikin Count in Humans	2.6+/- 0.8	2.9+/- 0.9	4.8+/- 5.0	7.4+/- 7.0	7.3+/- 6.7
H5N1 pB1 gene area (RPG) Replikin Count in Humans	2.0+/-0	2.0+/- 0.1	8.0+/-7.7	16.1+/- 5.7	15.4+/- 5.9
H5N1 pB2 gene area Replikin Count in Humans	2.4+/-0	2.8+/- 0.3	2.4+/-0.4	2.4+/- 0.1	2.4+/- 0.3
H5N1 pA gene area Replikin Count in Humans	3.8+/-0	4.0+/- 0.6	3.8+/- 0.4	3.8+/- 0.3	4.2+/- 0.3
H5N1 Human Mortality Percent			45	69	85

[000239] In Figure 16, a correlation was established between human mortality and (1) mean concentration of Replikin sequences in the whole genome, (2) mean concentration of Replikin sequences in the polymerase gene, and (3) mean concentration of Replikin sequences in the Replikin Peak Gene (pB1 gene area) of H5N1 influenza strains. As Replikin concentration increased by these three measures, human mortality was observed to increase. However, while all three measures provided a correlation with human mortality, changes in the Replikin Count in the polymerase gene correlated more significantly with human mortality, and changes in the Replikin Count in the Replikin Peak Gene (pB1 gene area) of the H5N1 genome correlated still more significantly with human mortality. Figure 16 suggests, therefore, that identification of Replikin Peak Genes within viral genomes improves identification and prediction of virulence and mechanisms of virulence using Replikin concentration data.

[000240] Figure 17 illustrates a significant eight-fold increase in Replikin concentration in the pB1 gene area (Replikin Peak Gene) of isolates of H5N1 while no significant increase is observed in neighboring gene areas of the pB1 gene area, namely, the pA gene area and the pB2 gene area. Figure 17 illustrates a significant correlation between human mortality and the Replikin Peak Gene (pB1 gene area) of isolates of H5N1 influenza virus. No correlation is observed in neighboring gene areas of the pB1 gene area, namely the pB2 and pA gene areas. In addition to the correlative aspect of the increase in Replikin concentration being related to percent mortality, Figure 17 provides strong confirmation of the power and validity of the methodology of predicting changes in virulence and outbreaks of virus by monitoring changes in Replikin concentration.

[000241] Table 3 provides mortality data for H5N1 infections from 2005 through 2007 and does not include earlier mortality data. Mortality data prior to 2005 has not been included in Table 3 because data prior to 2005 is inconsistent and understood by those of skill in the art to contain errors including errors caused by underreporting. The first generally agreed occasion when there were human deaths caused by proven H5N1 infection was in Hong Kong in 1997-1998. (This is probably incorrect, however, since there probably was mortality between 1959, when H5N1 was first reported, and 1997). The usual figures cited for 1997 are: 30 human cases, 8 deaths with mortality rate of about 27%. The number of cases (morbidity) and the number who died (mortality) that were not reported is unknown, but suspected to be significant. These errors are usually high in geographic areas where the medical care is less structured and scientific and the reporting is incomplete. Press reports between 1998 and 2002 were few, scattered, and not in agreement. Mortality data between 2005 and 2007 appear to be more consistent and have a higher level of reliability. Table 3, therefore, contains data from these years.

D. Replikin Count in RPG Correctly Identifies Host Lethality and Geographic Location of Outbreaks

[000242] The inventors have now demonstrated that Replikin Count in a Replikin Peak Gene provides a method for predicting and identifying outbreaks of pathogens such as H5N1 influenza by host and by geographic area. Figure 5 demonstrates the predictive capacity for identifying outbreaks in particular hosts and Figure 6 demonstrates the predictive capacity for identifying the lethality of an outbreak in a particular geographic area. The data for Figures 5 and 6 are contained in Table 4 below.

Table 4-Host and Geographic Predictions
(Replikin Count and Standard Deviation)

	Goose	Chicken	Duck	Human	Japan	Russia	Egypt	China	Vietnam	Thailand	Indonesia
2003	No data	2±1.1	2.1±1.1	2.2±1.2	2±0	No data	No data	1.9±0.3	No data	1.9±0.2	No data
2003	No data	3±1.5	3.8±1.8	3.3±1.3	2±0	No data	No data	2±0.1	2±0.2	2±0.2	2±0
2005	2.6±1.2	3.2±2.8	4±5	3.7±4.1	2.1±0.1	2±0.1	No data	2.7±0.5	3±0.2	4.1±4.7	3.9±3.6
2006	2.5±1.2	3.2±3.1	2.7±1.6	5±5.7	2±0	2±0	2±0	3±0.2	No data	6.7±6.7	16.7±4.9

[000243] Increased Replikin Counts in pathogens in particular hosts is predictive of an increase in probability of an outbreak of the pathogen. For example, Figure 5 illustrates a 2005 through 2007 upregulation of human H5N1 in humans as compared to H5N1 in goose, duck and chicken. Replikin analysis was performed separately for H5N1 Replikin Peak Genes of each host group, namely, goose, duck, chicken and human. Low levels of Replikin Count, below 4, were observed in each host group until 2005-2006, when epidemics increased in Asian countries. While duck H5N1 counts decreased in 2006, Replikin Counts continued to increase in chicken H5N1 in 2006. Human RPG activity was upregulated in 2005-2006 and overtook RPG activity in chickens. This transition of increased Replikin Count from duck to chicken to human is in agreement with epidemiological evidence of the order of transfer of the virus between hosts. Changes in Replikin Count in the Replikin Peak Gene of the H5N1 isolates in Figure 5 allows for identification of those hosts in which the influenza virus strain is more virulent than other hosts.

[000244] Increased Replikin Counts in pathogens in particular geographic areas is predictive of an increase in lethality of the pathogen in the identified geographic area. For example, Figure 6 illustrates localization of human H5N1 isolates having the highest lethality by measuring mean Replikin Counts in isolates of human H5N1 from different geographic areas isolated in a given year. Replikin analysis was performed separately for human H5N1 RPGs of each country. The results are shown for the Replikin Count for all data available on PubMed each year 2003-2006. Low levels of Replikin count, below 4, were observed in each host group until 2005-2006, when human H5N1 increased in Asian countries. Human RPG activity was upregulated in 2005-2006 most prominently in Indonesia. Using this data, Applicants predicted Indonesia would be the country most likely to first experience increased human mortality. The prediction was proven correct in 2007 when incidence of human

morbidity and mortality in the Indonesian outbreak were exceptionally high and evidence of possible human to human transmission was observed. Changes in Replikin Count in the Replikin Peak Gene of the H5N1 isolates such as in Figure 6 allows for identification of those geographic areas in which the influenza virus strain is more virulent than other geographic areas.

E. Replikin Peak Genes as Predictors of Outbreaks

[000245] Identification of the pB1 Replikin Peak Gene as a more significant gene area for changes in Replikin concentration effecting virulence reflects the same phenomenon in equine influenza as demonstrated in the pB1 gene area of H5N1. *See* Figure 7 and compare to Figures 2 and 3. Figure 7 additionally demonstrates the cyclical nature of changes in Replikin Count over a period of years. These cycles are like those observed previously for H1N1 since 1918. *See* Figures 2 and 3 and U.S. Patent No. 7,189,800 (Tables 3-6). It is noteworthy that increases in Replikin Count in virulent influenza isolates have generally ranged between 2 and 5, that is 2- to 3-fold above other influenza isolates. Replikin Counts in the Replikin Peak Gene of virulent isolates, however, have been observed to range between 2 and 20, that is a 10-fold change in concentration. This magnification makes sense in terms of the concentration of the Replikins in the Replikin Peak Gene, rather than an even distribution throughout other parts of the virus genome.

F. Replikin Concentration in Replikin Peak Gene of pB1 Area Correlates with Equine Influenza Epidemics

[000246] As with other influenza strains, an increase in Replikin concentration in equine influenza virus (EIV) has likewise been shown to be predictive of an increase in virulence of the virus and allows for prediction of forthcoming outbreaks or increases in morbidity and, in extreme cases, mortality. A review of publicly available amino acid sequences of isolates of EIV that demonstrates an increase in Replikin Count in the genome or a genome segment, or in a protein or protein fragment of the virus over time or between isolates is used as a predictor of an increase in outbreaks and morbidity in horses, donkeys, mules and other effected animals. Publicly available sequences for isolates of EIV from PubMed or other public or private sources may be analyzed by hand or using the FluForecast[®] search tool. (REPLIKINS LLC, Boston, MA).

[000247] Applicants have established a correlation between Replikin Count in the pB1 gene area (RPG) in EIV and an increase in virulence of the virus resulting in epidemics. The Applicants have reviewed publicly available amino acid sequences of isolates of EIV having

accession numbers at www.pubmed.com and have identified increases in Replikin concentration in the Replikin Peak Gene of the pB1 gene area of the genome of the virus that relate to and predict an increase in outbreaks.

[000248] Applicants' initial analysis determined the Replikin Peak Genes within publicly available sequences of the pB1, pB2 and pA proteins of the H3N8 strain of influenza virus by analyzing publicly available sequences for the gene areas of the pB1, pB2 and pA proteins and identifying the protein segment having the highest concentration of continuous Replikin sequences within each gene area.

[000249] Applicants then compared the mean Replikin concentration in the identified Replikin Peak Gene for each of the three gene areas for isolates in each year having publicly available sequence information between 1977 and 2007. Applicants further analyzed all publicly available whole genome sequences for H3N8 between 1977 and 2007.

[000250] Figure 7 illustrates a relationship between Replikin Counts of Replikin Peak Genes identified within the pB1, pB2, and pA genomic areas of equine influenza virus 1977-2007 and epidemics of equine encephalitis caused by H3N8 equine influenza. Replikin Count increases in the pB1 gene area are observed to occur one to three years before epidemic outbreaks while no increase in Replikin Count is observed in the pB2 and pA gene areas. Standard deviation of the means is again shown separately (as a clear column) to draw attention to the increase of some individual viruses with higher Replikin counts prior to the maximal Replikin count elevation, followed by viral outbreak.

[000251] Replikin Counts of the RPGs of the pA and pB2 genomic areas, which are immediately adjacent to the pB1 area in the H3N8 genome, remain below 5 and do not increase to the extent of the Replikin Count of the RPG of the pB1 area. These observed increases in the pB1 area and absence of increases in the pB2 and pA areas are in direct agreement with the data on H5N1 influenza reflected in Figure 1.

[000252] The range of Replikin Counts in the RPGs of H3N8 may be observed to be similar to the range of Replikin Counts in other *Influenza A* species. *See, e.g.*, Figure 2. Further, Replikin Counts in H3N8 during quiescent periods are comparable to Replikin Counts in *Influenza B* at all observed times and comparable to other influenza species during quiescent periods, that is between lethal outbreaks. Additionally, Replikin Counts in H3N8 during epidemics are comparable to outbreak levels reached prior to epidemics in *Influenza A*. *See, e.g.*, Figure 2.

[000253] The data for Figure 7 is provide in Table 4A below, which provides the yearly mean Replikin concentrations (with Standard Deviation) of publicly available peptide sequences of the identified Replikin Peak Gene (RPG) of the pB1 gene area, the yearly mean Replikin concentrations of publicly available peptide sequences of the identified Replikin Peak Gene (RPG) of the pA gene area, and the yearly mean Replikin concentrations of publicly available peptide sequences of the identified Replikin Peak Gene (RPG) of the pB2 gene area.

Table 4A – Equine Influenza

Year	Replikin Concentration of pB1 Gene	SD	Replikin Concentration of RPG in pB1	SD	Replikin Concentration of RPG in pA	Replikin Concentration of RPG in pB2
1972	1.8	0				2.4
1977	16.7	0	16.7	0		2.4
1978	11	12.7	11	12.7		2.4
1979	22.2	0	22.2	0		2.4
1980	6.7	8.8	6.7	8.8		2.3
1982	17.8	0	17.8	0		2.4
1985	20.6	2.4	20.6	2.4		2.4
1986	13.3	11.1	13.3	11.1	3.4	2.3
1987	19.1	2.7	19.1	2.7		2.4
1991	9.3	10.4	9.3	10.4	4.6	2.4
1992	15.6	0	15.6	0.2		2.4
1998	2.2	0	2.2	0		3.8
1999	2	0	2	0		0.2
2000						0.6
2001	9.9	11.2	9.9	11.2	2.2	2.4
2002	4.6	6.1	4.6	6.1	2.2	2.4
2003	2	0	2	0.2	3.1	2.4
2004	2.2	0.2	2.2	0.2	2.2	2.4
2005	18.5	2.8	18.5	2.8		2.3
2006					3.5	
2007					2.1	

[000254] In Figure 7, Series 1 reflects the mean Replikin concentration identified in the Replikin Peak Gene in the pB1 area of the genome. Series 2 reflects the standard deviation from mean Replikin concentration in the pB1 gene area. The large standard deviations in the first column of every pair are noteworthy as the Standard Deviation then drops as the mean Replikin concentration increases. This increase in standard deviation in the Replikin Peak Gene pB1 area probably reflects heterogeneity in the virus population once a more virulent strain of virus having a higher Replikin concentration has become present. The higher standard deviation suggests a more diverse population of the virus in which some members are relatively dormant whereas an increasing number are rapidly replicating. As the “build up” increases prior to the outbreak, more members are increasingly, rapidly replicating, thus

raising the mean Replikin concentration. In contrast, as seen in Table 4A above, the stability of the Replikin concentration in neighboring genomic areas such as pA and pB2 demonstrate both the reproducibility of the quantitative measurement of the Replikin concentration, the constancy over many years of the Replikin concentration in dormant areas, and the high degree of specificity of the increases in the pB1 area. The standard deviation then drops as the more virulent strain or strains enters an epidemic stage and less virulent strains (having lower Replikin concentrations) become less competitive and less present as a percentage of isolates in the host population. To the inventors' knowledge no such highly specific changes in virus structure have been observed to correlate with outcomes in the host.

[000255] Specifically, Series 3 in Figure 7 reflects the Replikin concentration identified in the Replikin Peak Gene in the pA gene area of the genome, which neighbors the pB1 gene area. The Replikin concentration of the Replikin Peak Gene in the pA gene area is observed to be remarkably constant over the analyzed years, never going above 5. This constancy stands in marked contrast to the extensive changes in Replikin concentration noted in the pB1 gene area. These control data validate the location of the most significant Replikin Peak Gene for the present isolates of virus in the pB1 gene area. Because the pA gene is right next to the pB1 gene, the differences in magnitude of change in Replikin concentration between these neighboring areas is quite remarkable.

[000256] Specifically, Series 4 in Figure 7 reflects the Replikin concentration identified in the Replikin Peak Gene in the pB2 gene area of the genome, which also neighbors the pB1 gene area. The Replikin concentration of the Replikin Peak Gene in the pB1 gene area is also observed to be remarkably constant over the analyzed years, not going above 4. This constancy again stands in marked contrast to the extensive changes in Replikin concentration noted in the pB1 gene area. Again, the control data validate the location of the most significant Replikin Peak Gene for the present isolates of virus in the pB1 gene area. Because the pB2 gene is right next to the pB1 gene, the differences in change in Replikin concentration between these neighboring areas is also remarkable.

VII. Methods of Predicting and Treating Outbreaks of Foot and Mouth Disease Virus (FMDV) Using RPGs and Related Replikin Sequences

[000257] An increase in Replikin concentration in the VP1 protein (containing an RPG of the virus genome) of foot and mouth disease virus (FMDV) is predictive of an increase in virulence and lethality of the virus and allows for prediction of forthcoming outbreaks or increases in virulence or lethality. Applicants have reviewed all publicly available amino acid sequences of isolates of FMDV having accession numbers at www.pubmed.com

between 1969 and 2006 and have identified increases in Replikin concentration in the VP1 protein of FMDV (*e.g.*, SEQ ID NO: 157) that relate to and predict certain known outbreaks of Foot and Mouth Disease. Figure 11 illustrates the correlation of Replikin Count observed in the VP1 protein of isolates of foot and mouth disease virus on a year by year basis and observed outbreaks.

[000258] Applicants reviewed Accession No. ABM63320 (SEQ ID NO:157), which provides the amino acid sequence of the entire serotype-O FMDV VP1 polyprotein, and identified two RPGs. The first RPG begins at amino acid residue 925 and continues through amino acid residue 1018 and was isolated *in silico* as SEQ ID NO:124. Five Replikin sequences were isolated (SEQ ID NOS:125-129) in the first RPG, which gave the first RPG a Replikin Count of 6.3. The first RPG represents the Replikin Peak Gene of a fragment of the VP1 polyprotein.

[000259] The second Replikin Peak Gene begins at amino acid residue 1300 and continues through amino acid residue 1481 and was isolated *in silico* as SEQ ID NO:130. Twenty-six Replikin were isolated in the second RPG (SEQ ID NOS: 131-156). The second Replikin Peak Gene Area has a Replikin Count of 14.3 and represents the Replikin Peak Gene of the entire reported VP1 polyprotein. Conserved Replikins within the RPG at SEQ ID NO:130 are also contained, for example, in sequence fragments reported at Accession Nos. ABA46641, AAG43385, AAP81678 and ABG77564. Likewise, parts of the RPG of SEQ ID NO:124 are contained in these accession numbers.

[000260] In the amino-terminal of SEQ ID NO:157 (Accession No. ABM63320) SEQ ID NOS:158-160 were isolated as Replikins. In the mid-molecule, SEQ ID NOS:161-194 were isolated as Replikins. In the carboxy-terminal, SEQ ID NOS:195-213 were isolated as Replikins. Each of these Replikin sequences is a preferred sequence for immunogenic compositions and vaccines and for other diagnostic, therapeutic and predictive purposes as described herein.

[000261] Figure 11 illustrates the concentration of Replikin sequence observed in the VP1 protein of isolates of the common serotype-O of foot and mouth disease virus having publicly available accession numbers on a year by year basis between 1969 and 2006. Observed European and UK outbreaks of Foot and Mouth Disease are noted and relate to observed increases in Replikin Count prior to disease outbreak.

[000262] Prediction of the listed epidemics as well as future outbreaks may be made, for example, by reviewing the Replikin Counts of isolates of FMDV and comparing the Replikin

Counts of the VP1 protein or the RPG within the VP1 protein for a particular year with Replikin Counts from other years. A significant increase in Replikin Count from one year to the next and preferably over one, two or three years provides predictive value of an emerging strain of FMDV that may begin an outbreak of Foot and Mouth Disease. A Foot and Mouth Disease outbreak may be predicted within about six months to about one year or more from the observation of a significant increase in Replikin Count.

[000263] More preferably, an outbreak of Foot and Mouth Disease may be predicted within about six months to about one year from the observation of a significant increase in Replikin count over two or three years. An outbreak may likewise be predicted within about six months to about one year from the initial observation of a decrease in Replikin Count following a significant increase. Using this method, Applicants predicted the August 3, 2007 outbreak of FMDV in the United Kingdom months prior to the outbreak.

[000264] The data for Figure 11 is provided in Table 5 below. Note that data is available for 1958 and 1962, but was not included in Figure 11. Note also that no data was available for 1959 through 1961, 1963 through 1968 and 2004.

Table 5 – FMDV Serotype O Replikin Counts

Year	Accession Records for Serotype-O FMDV VP1	Replikin Count	Standard Deviation	Significance (compared to lowest value)	Significance (compared to previous year)
1958	1	3.8	0.0		
1962	172	0.8	0.2	low p<0.001	
1969	1	0.4	0.0		prev p<0.001
1970	2	0.7	0.4	low p<0.05	prev p<0.40
1971	1	3.3	0.0		prev p<0.05
1972	1	0.5	0.0		
1973	1	0.5	0.0		
1974	7	1.1	0.3	low p<0.001	prev p<0.001
1975	6	1.0	0.5	low p<0.001	prev p>0.50
1976	4	1.2	0.1	low p<0.001	prev p<0.40
1977	4	1.0	0.4	low p<0.001	prev p<0.20
1978	8	1.0	0.1	low p<0.001	prev p>0.50
1979	47	2.5	1.4	low p<0.001	prev p<0.001
1980	5	1.5	0.5	low p<0.001	prev p<0.001
1981	2	0.8	0.3	low p<0.04	prev p<0.10
1982	21	0.9	0.1	low p<0.001	prev p>0.50
1983	6	0.9	0.4	low p<0.001	prev p>0.50
1984	1	1.2	0.0		prev p<0.10
1985	3	1.2	0.7	low p<0.02	prev p>0.50
1986	2	0.8	0.5	low p<0.05	prev p>0.50
1987	1	3.6	0.0		prev p<0.05
1988	3	1.2	0.0		
1989	6	1.3	0.8	low p<0.001	prev p>0.50
1990	5	2.0	1.3	low p<0.02	prev p<0.30
1991	7	3.5	2.4	low p>0.50	prev p<0.10

1992	9	2.1	1.0	low p<0.001	prev p<0.10
1993	16	2.2	1.4	low p<0.001	prev p>0.50
1994	18	2.2	1.1	low p<0.001	prev p>0.50
1995	12	2.0	0.8	low p<0.001	prev p>0.50
1996	12	1.6	1.1	low p<0.001	prev p<0.30
1997	48	2.6	1.1	low p<0.001	prev p<0.01
1998	72	1.2	1.1	low p<0.001	prev p<0.001
1999	49	2.3	1.3	low p<0.001	prev p<0.001
2000	61	1.3	0.8	low p<0.001	prev p<0.001
2001	8	1.8	0.9	low p<0.001	prev p<0.10
2002	2	1.1	0.2	low p<0.02	prev p<0.05
2003	8	1.3	0.5	low p<0.001	prev p<0.40
2005	8	1.3	1.0	low p<0.001	prev p>0.50
2006	3	3.8	3.1	low p>0.50	prev p<0.20

A. Prediction Using VP1 Protein of All Serotypes

[000265] In addition to FMDV VP1 proteins of serotype-O, Applicants also analyzed publicly available sequences for isolates of all reported serotypes of FMDV VP1 protein from PubMed. The data is provided in Table 6 below. Note the increase in Replikin Count correlated with two epidemics in the United Kingdom (and other European countries) in 2001 and in the United Kingdom in 2007. Also note the low Replikin Counts during quiescence. Replikin Count increases from 1.6 in 1998, to 2.5 in 1999, to 2.7 in the year of the epidemic, 2001. Then post-epidemic, three lower Replikin Count years are noted, 1.5 in 2002, 1.5 in 2003, and 1.1 in 2005 (there were no publicly available sequences from 2004). The Replikin Count then rose to 2.8 in 2006 just prior to the outbreak in 2007. Note that the p values are less than 0.001 with respect to previous Replikin Counts.

Table 6 – FMDV (all isolates)

Year	Accession Records for FMDV VP1 protein	Replikin Count	Standard Deviation	Significance (compared to lowest value)	Significance (compared to previous year)
1998	92	1.6	±1.2	low p<0.001	prev p<0.001
1999	60	2.5	±1.3	low p<0.001	prev p<0.001
2000	115	1.7	±1.4	low p<0.001	prev p<0.001
2001	32	2.7	±1.0	low p<0.001	prev p<0.001
2002	3	1.5	±0.8	low p<0.05	prev p<0.02
2003	10	1.5	±0.8	low p<0.001	prev p>0.50
2005	43	1.1	±0.6	low p<0.001	prev p<0.10
2006	36	2.8	±0.9	low p<0.002	prev p<0.001

B. Prediction Using VP1 Protein of Serotype C

[000266] Table 7 provides Replikin Count data for isolates of serotype-C FMDV for some years between 1955 and 2006. Note the significant increases over the low value in Replikin Count in 1998 and 1999 (prior to the 2001 epidemic in the UK) and the significant increase over the low value in 2006 (prior to the 2007 outbreak in the UK). Years having no available data are not reflected in the table.

Table 7 - FMDV Serotype C

Year	Accession Records for Serotype C	Replikin Count	Standard Deviation	Significance (compared to lowest value)	Significance (compared to previous year)
1955	1	2.7	0.0		
1957	1	2.8	0.0		
1979	12	1.4	0.5	low p<0.001	prev p<0.001
1982	2	2.1	1.6	low p>0.50	prev p>0.50
1988	1	1.1	0.0		prev p>0.50
1989	1	1.1	0.0		
1991	3	0.5	0.0	low p<0.001	prev p<0.001
1992	2	0.5	0.0	low p<0.001	prev p>0.50
1993	5	0.5	0.0	low p<0.001	prev p<0.40
1997	2	1.4	1.3	low p<0.30	prev p<0.30
1998	2	2.9	0.0		prev p<0.20
1999	4	3.1	0.4	low p<0.10	prev p<0.40
2006	10	3.0	0.2	low p<0.001	prev p>0.50

[000267] The correlation between Replikin concentration and viral outbreaks noted above and illustrated in Figure 11 provide a method of predicting outbreaks of Foot and Mouth Disease by monitoring increases in Replikin concentration in the VP1 protein of all available FMDV isolates. The method may also employ all available serotype-O isolates or serotype-C isolates of the virus.

[000268] The epidemiology and virology FMDV is different from the epidemiology and virology of some other viruses discussed herein such as Influenza virus. Nevertheless, a correlation between increases in Replikin Count in the FMDV VP1 protein and outbreaks of the virus provides compounding data establishing a shared phenomenon of rapid replication and virulence with an overwhelming number of other tested viruses and organisms.

C. Replikins Conserved in Serotype O FMDV RPGs

[000269] In serotype-O of FMDV, two conserved Replikin sequences contained within the Replikin Peak Gene are hkqkivapvk (SEQ ID NO:91) and hpsearchkqkivapvk (SEQ ID NO:92). A point mutant of the hpsearchkqkivapvk sequence to hptearhkqkivapvk (SEQ ID NO:93) (mutation underlined) reportedly occurred in isolates from 1967 and 2007. The Replikin sequence hkqkivapvk (SEQ ID NO:91) has been conserved from 1962 to 2006. The Replikin sequence hpsearchkqkivapvk (SEQ ID NO:92) has been conserved from 1962 to 2006 except for the point mutation hptearhkqkivapvk (SEQ ID NO:93), which is present in isolates reportedly having caused the 1967 outbreak (isolate O₁BFS) and now the 2007 outbreak in the United Kingdom. These isolated conserved Replikin sequence are embodiments of the invention of particular preference for predictive, diagnostic and therapeutic capacity.

[000270] Table 8 provides the accession numbers of isolates between 1962 and 2006 containing the conserved sequence hkqkivapvk (SEQ ID NO:91) and the amino acid position within the VP1 protein sequence where the conserved Replikin sequence begins.

Table 8 – FMDV Conserved SEQ ID NO:91

1962	CAC22210 position 202 , AAP81678 position 153 , AAP81677 position 153 , AAP81676 position 153 , AAP81675 position 153 , AAP81674 position 153 , ABA46701 position 201 , ABA46700 position 201 , ABA46699 position 201 , ABA46698 position 201 , ABA46697 position 201 , ABA46696 position 201 , ABA46695 position 201 , ABA46693 position 201 , ABA46692 position 201 , ABA46691 position 201 , ABA46690 position 201 , ABA46689 position 201 , ABA46688 position 201 , ABA46687 position 201 , ABA46686 position 201 , ABA46685 position 201 , ABA46684 position 201 , ABA46683 position 201 , ABA46682 position 201 , ABA46681 position 201 , ABA46679 position 201 , ABA46678 position 201 , ABA46677 position 201 , ABA46675 position 201 , ABA46674 position 201 , ABA46673 position 201 , ABA46672 position 201 , ABA46671 position 201 , ABA46670 position 201 , ABA46669 position 201 , ABA46668 position 201 , ABA46666 position 201 , ABA46665 position 201 , ABA46664 position 201 , ABA46663 position 201 , ABA46662 position 201 , ABA46661 position 201 , ABA46660 position 201 , ABA46659 position 201 , ABA46658 position 201 , ABA46657 position 201 , ABA46655 position 201 , ABA46654 position 201 , ABA46653 position 201 , ABA46652 position 201 , ABA46651 position 201 , ABA46650 position 201 , ABA46649 position 201 , ABA46648 position 201 , ABA46647 position 201 , ABA46644 position 201 , ABA46643 position 201 , ABA46642 position 201 , ABA46641 position 201 , ABA46640 position 201 , ABA46639 position 201 , ABA46638 position 201 , ABA46637 position 201 , ABA46614 position 201 , ABA46613 position 201 , ABA46612 position 201 , ABA46611 position 201 , ABA46610 position 201 , ABA46609 position 201 , ABA46606 position 201 , ABA46605 position 201 , ABA46604 position 201 , ABA46603 position 201 , ABA46602 position 201 , ABA46601 position 201 , ABA46600 position 201 , ABA46597 position 201 , ABA46596 position 201 , ABA46594 position 201 , ABA46591 position 201 , ABA46590 position 201 , ABA46589 position 201 , ABA46588 position 201 , ABA46586 position 201 , ABA46585 position 201 , ABA46583 position 201 , ABA46582 position 201 , ABA46581 position 201 , ABA46580 position 201 , ABA46579 position 201 , ABA46578 position 201 , ABA46576 position 201 , ABA46574 position 201 , ABA46573 position 201 , ABA46571 position 201 , ABA46570 position 201 , ABA46569 position 201 , ABA46568 position 201 , ABA46566 position 201 , ABA46565 position 201 , ABA46563 position 201 , ABA46561 position 201 , ABA46560 position 201 , ABA46542 position 201 , ABA46541 position 201 , ABA46539 position 201 , ABA46538 position 201 , ABA46537 position 201 , ABA46536 position 201 , ABA46535 position 201 , ABA46534 position 201 , ABA46533 position 201 , ABA46532 position 201 , ABA46531 position 201 , ABA46559 position 201 , ABA46540 position 201
1969	CAB62584 position 724
1972	CAC22304 position 202.
1974	CAC22211 position 202 , AAK69575 position 153 , AAR85362 position 153 , AAR85361 position 153 , AAR22955 position 153 , AAR22953 position 153 .
1975	AAK69576 position 153 , CAC20174 position 201 , AAR85363 position 153 , AAG35653 position 724.
1976	AAR22952 position 153 , AAR22933 position 153 , AAR22932 position 153.
1977	AAR22963 position 153 , AAR22950 position 153 , CAC48179 position 201.
1978	ABA46745 position 201 , ABA46744 position 201 , ABA46743 position 201 , ABA46742 position 201 , ABA46740 position 201 , AAR22930 position 153 .
1979	CAC22173 position 43 , AAQ88330 position 153 , AAQ88328 position 153 , AAQ88327 position 153 , AAQ88325 position 153 , AAQ88324 position 153 , AAQ88323 position 153 , AAQ88322 position 153 , AAQ88321 position 153 , AAQ88320 position 153 , AAQ88319 position 153 , AAQ88318 position 153 , AAQ88317 position 153 , AAQ88316 position 153 , AAQ88315 position 153 , AAQ88314 position 153 , AAQ88313 position 153 , AAQ88312 position 153 , AAG28368 position 43 , AAG28367 position 43 , AAG28366 position 43 , AAG28362 position 43 , AAG28357 position 43 , AAG28356 position 43 , AAG28355 position 43 , AAG28354 position 43 , AAG28353 position 43 , AAG28352 position 43 , AAG28348 position 43 .
1980	AAR22962 position 153 , AAR22959 position 153 , AAR22941 position 153.
1981	AAR22951 position 153
1982	CAC20178 position 201 , AAZ31360 position 201 , AAZ31359 position 201 , AAZ31358 position

	201 , AAZ31357 position 201 , AAZ31356 position 201 , AAZ31355 position 201 , AAZ31354 position 201 , AAZ31353 position 201 , AAZ31352 position 201 , AAZ31351 position 201 , AAZ31350 position 201 , AAZ31349 position 201 , AAZ31348 position 201 , AAZ31347 position 201 , AAZ31346 position 201 , AAZ31345 position 201 , AAZ31344 position 201 , AAZ31343 position 201 , AAZ31342 position 201 .
1983	AAR22960 position 153, AAR22938 position 153, AAR22937 position 153.
1985	CAC22326 position 90.
1986	AAR22954 position 153.
1987	AAK62003 position 43.
1988	AAK69568 position 153, AAK69567 position 153.
1989	CAC22174 position 90, AAR22961 position 153, AAK62024 position 69.
1990	CAC22178 position 43, CAC22327 position 58.
1991	CAC22175 position 43, CAC22328 position 62.
1992	CAC22176 position 43, CAC22240 position 85, CAC48182 position 201.
1993	CAC22179 position 43, CAC40792 position 201, CAC40789 position 201, CAC40796 position 102.
1994	CAC22180 position 76 , CAC22233 position 62 , CAC22227 position 60 , CAC22215 position 47 , CAC22208 position 82 , CAC22201 position 43 , CAC22167 position 43 , AAK62012 position 43 , CAC40794 position 102 , CAC40790 position 201 , CAC40795 position 102 , CAC40797 position 201 .
1995	CAC22231 position 152, CAC22216 position 44, CAC22171 position 103, AAK62022 position 69.
1996	CAC22194 position 127 , CAC51235 position 201 , AAR22945 position 153 , AAR22942 position 153 , AAK62005 position 69 .
1997	CAC51273 position 201 , CAC51268 position 201 , CAC51249 position 201 , CAC51236 position 201 , AAL05257 position 43 , AAL05249 position 43 , AAL05248 position 85 , AAL05247 position 62 , AAL05246 position 76 , AAL05245 position 43 , AAL05243 position 56 , AAL05242 position 43 , AAL05236 position 43 , AAL05235 position 65 , AAL05234 position 43 , AAL05233 position 43 , AAL05232 position 43 , AAL05231 position 43 , AAL05230 position 43 , AAL05229 position 43 , AAL05228 position 43 , AAL05227 position 85 , AAL05226 position 43 , AAL05225 position 76 , AAL05223 position 43 , AAL05222 position 43 , AAL05221 position 43 , AAL05220 position 122 , AAL05219 position 43 , AAL05218 position 52 , AAL05217 position 43 , AAL05216 position 66 , AAL05214 position 43 , AAL05213 position 93 , AAL05211 position 58 , AAL05207 position 43 , AAL05206 position 62 , AAL05205 position 67 , AAL05196 position 64 .
1998	CAC22229 position 201 , ABI16250 position 201 , ABI16249 position 201 , ABI16248 position 201 , ABI16247 position 201 , ABI16246 position 201 , ABI16245 position 201 , ABI16244 position 201 , ABI16242 position 201 , ABI16241 position 201 , ABI16240 position 201 , ABI16239 position 201 , ABI16238 position 201 , ABI16237 position 201 , ABI16236 position 201 , ABI16235 position 201 , ABI16234 position 201 , ABI16233 position 201 , ABI16232 position 201 , ABI16231 position 201 , ABI16230 position 201 , ABI16229 position 201 , ABI16228 position 201 , ABI16227 position 201 , CAC51269 position 201 , CAC51239 position 201 , CAC51238 position 201 , AAR85364 position 153 , AAR22957 position 153 , AAL05256 position 43 , AAL05255 position 43 , AAL05254 position 43 , AAL05253 position 43 , AAL05250 position 43 , AAL05244 position 43 , AAL05241 position 43 , AAL05240 position 43 , AAL05238 position 43 , AAL05237 position 45 , AAL05212 position 43 .
1999	CAC22228 position 100 , CAC22200 position 100 , AAG43385 position 43 , CAC51332 position 143 , CAC51270 position 175 , CAC51255 position 201 , CAC51318 position 201 , CAC51247 position 201 , CAC51246 position 201 , CAC51245 position 201 , CAD62370 position 925 , CAD62369 position 925 , CAD62208 position 925 , CAC20187 position 201 , AAR22956 position 153 , AAR22940 position 153 , AAF06146 position 43 , AAD41912 position 81 , AAD41131 position 81 , AAL05251 position 43 , AAL05215 position 43 , AAL05210 position 43 , AAL05209 position 43 , AAL05208 position 43 , AAL05204 position 43 , AAL05203 position 45 , AAL05202 position 43 , AAL05201 position 43 , AAL05200 position 43 , AAL05199 position 43 , AAL05198 position 43 , AAL05197 position 70 , AAL05195 position 59 , AAL05194 position 58 , AAL05193 position 43 , AAL05192 position 43 , AAL05191 position 43 .
2000	CAC22209 position 201 , AAL09392 position 153 , AAL09391 position 153 , AAK69397 position 153 , ABF18551 position 43 , ABF18550 position 43 , ABF18549 position 43 , ABF18548 position 43 , CAC51275 position 201 , CAC51271 position 201 , CAC51267 position 201 , CAC51264 position 201 , CAC51263 position 201 , CAC51261 position 201 , CAC51258 position 201 , CAC51257 position 201 , BAC06475 position 925 , CAD62372 position 925 , CAD62371 position 925 , AAG27038 position 153 , AAG27037 position 153

2001	CAD62373 position 925 , AAK92375 position 925 , CAC35464 position 201 , CAC35463 position 201 , CAC35462 position 201 , CAC35461 position 201 , CAG23917 position 925 , CAC86575 position 925 .
2002	AAR07959 position 153, AAM62134 position 201.
2003	AAQ93493 position 925 , AAR07963 position 153 , AAR07962 position 153 , AAR07961 position 153 , AAR07960 position 153 , AAR07965 position 153 , AAR07964 position 153 .
2005	ABD14417 position 201 , ABC55721 position 43 , CAJ51080 position 201 , CAJ51079 position 201 , CAJ51078 position 201 , CAJ51077 position 201 , CAJ51076 position 201 , CAJ51075 position 201 .
2006	ABG77563 position 197 , ABG77564 position 30

[000271] Table 9 provides the accession numbers of FMDV isolates between 1962 and 2006 containing the conserved sequence hpsearchkqkivapvk (SEQ ID NO:92) or the point mutation hptearhkqkivapvk (SEQ ID NO:93) and the amino acid position within the VP1 protein sequence where the conserved Replikin sequence begins.

Table 9 - FMDV SEQ ID NO:92 OR SEQ ID NO: 93

1962	AAP81678 position 147 , AAP81677 position 147 , ABA46700 position 195 , ABA46699 position 195 , ABA46698 position 195 , ABA46697 position 195 , ABA46696 position 195 , ABA46695 position 195 , ABA46693 position 195 , ABA46692 position 195 , ABA46691 position 195 , ABA46690 position 195 , ABA46689 position 195 , ABA46688 position 195 , ABA46687 position 195 , ABA46686 position 195 , ABA46685 position 195 , ABA46684 position 195 , ABA46683 position 195 , ABA46682 position 195 , ABA46681 position 195 , ABA46679 position 195 , ABA46678 position 195 , ABA46677 position 195 , ABA46675 position 195 , ABA46673 position 195 , ABA46672 position 195 , ABA46671 position 195 , ABA46670 position 195 , ABA46666 position 195 , ABA46665 position 195 , ABA46664 position 195 , ABA46663 position 195 , ABA46662 position 195 , ABA46661 position 195 , ABA46659 position 195 , ABA46658 position 195 , ABA46657 position 195 , ABA46655 position 195 , ABA46654 position 195 , ABA46649 position 195 , ABA46648 position 195 , ABA46647 position 195 , ABA46644 position 195 , ABA46643 position 195 , ABA46642 position 195 , ABA46640 position 195 , ABA46639 position 195 , ABA46638 position 195 , ABA46637 position 195 , ABA46614 position 195 , ABA46613 position 195 , ABA46612 position 195 , ABA46611 position 195 , ABA46609 position 195 , ABA46606 position 195 , ABA46605 position 195 , ABA46604 position 195 , ABA46603 position 195 , ABA46602 position 195 , ABA46601 position 195 , ABA46600 position 195 , ABA46588 position 195 , ABA46581 position 195 , ABA46574 position 195 , ABA46573 position 195 , ABA46571 position 195 , ABA46570 position 195 , ABA46569 position 195 , ABA46568 position 195 , ABA46566 position 195 , ABA46565 position 195 , ABA46563 position 195 , ABA46561 position 195 , ABA46539 position 195 , ABA46538 position 195 , ABA46537 position 195 , ABA46536 position 195 , ABA46535 position 195 , ABA46531 position 195 , ABA46559 position 195 .
1974	AAR85362 position 147, AAR85361 position 147.
1975	CAC20174 position 195.
1977	AAR22963 position 147, AAR22950 position 147.
1978	ABA46743 position 195, ABA46742 position 195, ABA46740 position 195.
1979	AAQ88330 position 147 , AAQ88328 position 147 , AAQ88325 position 147 , AAQ88324 position 147 , AAQ88323 position 147 , AAQ88321 position 147 , AAQ88319 position 147 , AAQ88317 position 147 , AAQ88316 position 147 , AAQ88315 position 147 , AAQ88314 position 147 , AAQ88313 position 147 , AAQ88312 position 147 , AAG28362 position 37 , AAG28355 position 37 .
1985	CAC22326 position 84.
1987	AAK62003 position 37.
1989	CAC22174 position 84, AAK62024 position 63.
1990	CAC22178 position 37.
1991	CAC22175 position 37, CAC22328 position 56.
1992	CAC22240 position 79.
1994	CAC22233 position 56.

1995	CAC22216 position 38, CAC22171 position 97.
1996	CAC22194 position 121, CAC51235 position 195, AAR22945 position 147, AAR22942 position 147, AAK62005 position 63.
1997	CAC51273 position 195, CAC51268 position 195, CAC51249 position 195, CAC51236 position 195, AAL05249 position 37, AAL05248 position 79, AAL05247 position 56, AAL05246 position 70, AAL05245 position 37, AAL05243 position 50, AAL05242 position 37, AAL05236 position 37, AAL05235 position 59, AAL05234 position 37, AAL05233 position 37, AAL05229 position 37, AAL05228 position 37, AAL05221 position 37, AAL05207 position 37, AAL05196 position 58.
1998	ABI16250 position 195, ABI16249 position 195, ABI16248 position 195, ABI16247 position 195, ABI16246 position 195, ABI16245 position 195, ABI16244 position 195, ABI16242 position 195, ABI16241 position 195, ABI16240 position 195, ABI16239 position 195, ABI16238 position 195, ABI16237 position 195, ABI16236 position 195, ABI16235 position 195, ABI16234 position 195, ABI16232 position 195, ABI16231 position 195, ABI16229 position 195, ABI16227 position 195, CAC51239 position 195, CAC51238 position 195, AAR22957 position 147, AAL05256 position 37, AAL05255 position 37, AAL05254 position 37, AAL05253 position 37, AAL05250 position 37, AAL05244 position 37, AAL05241 position 37, AAL05240 position 37, AAL05238 position 37, AAL05237 position 39.
1999	CAC22228 position 94, AAG43385 position 37, CAC51332 position 137, CAC51255 position 195, CAC51318 position 195, CAC51247 position 195, CAC51246 position 195, CAC51245 position 195, CAD62370 position 919, CAD62208 position 919, CAC20187 position 195, AAR22956 position 147, AAF06146 position 37, AAL05251 position 37, AAL05210 position 37, AAL05209 position 37, AAL05208 position 37, AAL05204 position 37, AAL05203 position 39, AAL05202 position 37, AAL05201 position 37, AAL05200 position 37, AAL05198 position 37, AAL05195 position 53, AAL05194 position 52, AAL05193 position 37.
2000	CAC22209 position 195, AAL09392 position 147, AAL09391 position 147, AAK69397 position 147, ABF18551 position 37, ABF18550 position 37, ABF18549 position 37, ABF18548 position 37, CAC51275 position 195, CAC51271 position 195, CAC51267 position 195, CAC51264 position 195, CAC51263 position 195, CAC51261 position 195, CAC51258 position 195, CAC51257 position 195, BAC06475 position 919, CAD62372 position 919, CAD62371 position 919, AAG27038 position 147, AAG27037 position 147, ABA46733 position 195, ABA46732 position 195, ABA46731 position 195, ABA46730 position 195, ABA46729 position 195, ABA46728 position 195, ABA46727 position 195, ABA46726 position 195, ABA46725 position 195, ABA46724 position 195, ABA46722 position 195, ABA46721 position 195, ABA46720 position 195, ABA46719 position 195, ABA46717 position 194, ABA46716 position 195, ABA46715 position 195, ABA46714 position 195, ABA46713 position 195, ABA46712 position 195, ABA46711 position 195, ABA46709 position 195, ABA46708 position 195, ABA46706 position 195, ABA46705 position 195, ABA46704 position 195, BAB18050 position 195.
2001	CAD62373 position 919, AAK92375 position 919, CAC35464 position 195, CAC35463 position 195, CAC35462 position 195, CAC35461 position 195, CAG23917 position 919.
2002	AAM62134 position 195.
2003	AAQ93493 position 919.
2005	ABD14417 position 195, ABC55721 position 37.
2006	ABG77563 position 191.

[000272] Accession No. AAG43385 (SEQ ID NO:107) reports an FMDV serotype O isolate from 1999 that partly contains the RPG of SEQ ID NO:124 and contains the conserved sequence SEQ ID NO:91. In SEQ ID NO:107, no Replikin sequences were identified in the amino-terminal. Replikin sequence SEQ ID NO:108 was identified in the mid-molecule. Replikin sequence SEQ ID NO:91 was identified in the carboxy-terminus.

[000273] Accession No. AAP81678 (SEQ ID NO:111), reports an FMDV serotype O isolate from 1962 that partly contains the RPG of SEQ ID NO:124 and contains the conserved sequence SEQ ID NO:91. Accession No. ABA46641 (SEQ ID NO:114) likewise

reports an FMDV serotype O isolate from 1962 that partly contains the RPG of SEQ ID NO:124 and contains the conserved sequence of SEQ ID NO:91 and the conserved sequence of SEQ ID NO:92 but for a single unknown residue at position 199 (SEQ ID NO:115). In SEQ ID NO:114, no Replikin sequences were identified in the amino-terminus or mid-molecule portion of the sequence. SEQ ID NOS:115 and 116 were isolated in the carboxy-terminus.

[000274] Accession No. ABG77564 (SEQ ID NO:118) reports an FMDV serotype O isolated from 2006 that partly contains the RPG of SEQ ID NO:124 and contains the conserved sequence SEQ ID NO:91. In SEQ ID NO:118, no Replikins were identified in the amino terminus of the sequence. SEQ ID NOS:119-121 and 91 were identified as Replikins in the mid-molecule. And no Replikins were identified in the carboxy terminus.

[000275] In addition to the diagnostic power of Replikin technology shown in these examples, it is clear that recognition for the first time of this class of virus peptides, and the discovery that they are related to rapid replication, virus outbreaks and high morbidity and mortality, makes the Replikins, and particularly the Replikin Peak Gene structures illustrated here, new conserved prime targets for treatment and vaccines in FMDV and other viruses. For example, the Replikin sequences (SEQ ID NOS:91-93) provide invariant targets for such a vaccine. Likewise, the RPGs of SEQ ID NOS:124 and 130 and the Replikin sequence identified in the accession number sequences (SEQ ID NOS:108, 115-116 and 119-121) are preferred sequences for immunogenic compositions and vaccines. An embodiment of the invention, therefore, is a vaccine comprising at least one of the sequences SEQ ID NOS:91-93 or SEQ ID NOS:108, 115-116 and 119-121 or any combination thereof.

VIII. Methods of Predicting and Treating Outbreaks of West Nile Virus Using RPGs and Related Replikin Sequences

[000276] Applicants have now demonstrated a correlation between an increase in Replikin Count in a Replikin Peak Gene of the west nile virus (WNV) (*e.g.*, SEQ ID NO:245) and outbreaks, morbidity and mortality in the viral disease. *See* Figure 12. Applicants have also demonstrated a correlation between Replikin Count in the whole virus genome and morbidity and mortality. *See* U.S. Prov. Appln. Ser. No. 60/853,744, filed Aug. 16, 2007.

[000277] Review of publicly available sequences of isolates of WNV from 1982-2007 revealed a Replikin Peak Gene in the envelope protein of west nile virus that has now been associated with virulence and lethality. In comparison with morbidity and mortality data in the United States between 1999 and 2006, an association between Replikin Count in the envelope protein of west nile virus and morbidity and mortality data is clear. *See* Figure 12.

Applicants' analysis of a Replikin Peak Gene in an envelope protein sequence of Accession No. ABA54585 (e.g., SEQ ID NO: 245) is provided in Example 7 below.

[000278] Figure 12 illustrates the Replikin Count of Replikins observed in the envelope protein in PubMed accession numbers on a year by year basis between 1982 and 2006. Increases in Replikin Count on a year by year basis are correlatable with both reported morbidity of the virus in the United States and reported mortality from viral infections in the United States.

[000279] The data for Figure 12 is provided in Table 10 below. Years in which no data were available are not included in the table. It may be observed that Replikin Count correlates with changes in both morbidity and mortality in the U.S. population between 1999 and 2006. The data further make clear a relative decrease in Replikin Count in 2004 followed by a time of relative quiescence of the west Nile virus in the United States in 2004 and 2005. Additionally, beginning in 2000, morbidity and mortality increases in relation to the increasing Replikin Count.

Table 10 – WNV Envelope Protein

Year	PubMed Accession Number-Replikin Count	Isolates	Mean RC	S.D.	Significance	Morbidity	Mortality
1982	84028435 111	1	3.2	0.0			
1985	P06935 109 NP_776014 109 NP_041724 109 NP_776013 109 NP_776012 109 AAA48498 109	6	3.2	0.0	low p <.001		
1988	P14335 95 BAA00176 95	2	2.8	0.0	prev p<.001		
1995	AAW80621 9	1	4.7	0.0			
1996	P51681 8	1	2.3	0.0			
1998	AAW81711 107 AAD28624 41	2	4.2	1.5	low p >.50, prev p<.30		
1999	AAL10755 6 AAL10754 6 AAL10752 6 AAL10751 6 AAL10750 7 AAL10749 6 AAG49029 3 AAG49028 3 AAG49027 2 AAD31720 32 AAL10753 9 AAF26360 40 AAL10748 6 AAL10747 6 AAL10746 6 AAL10745 6 AAL10744 6 AAL10743 6 AAL10742 6 AAL10741 4 AAL10740 6 AAL10738 6 AAL10737 6 AAL10736 6 AAL10735 6 AAL10734 6 AAL10733 6 AAL10732 6 AAL10731 6 AAL10730 6 AAL10729 6 AAL10728 6 AAL10727 6 AAL10725 10 AAL10724 6 AAL10723 6 id=15919195 6 id=12246899 6 id=12246897 6 id=12246895 6 id=12246893 6 AAG49629 6 AAG49628 6 AAL10739 6 AAL10726 6 AAD28623 40 AAF20092 97 AAG02040 97 AAF18443 97 AAF20205 97 AAF20207 7 AAF20206 7 AAF20204 7	53	3.6	0.9	low p <.005, prev p>.50	62	7
2000	AAK06624 97 AAG02039 98 AAG02038 97	3	2.8	0.0	low p <.001, prev p<.00	21	2

2001	<p><u>AAM70028 28 AAL14222 30 AAL14221 30 AAL14220 30 AAL14219 30 AAL14218 30 AAL14217 30 AAL14216 30 AAL14215 30 AAK58104 30 AAK58103 31 AAK58102 30 AAK58101 30 AAK58100 30 AAK58099 31 AAK58098 30 AAK58097 30 AAK58096 30 AAK52303 30 AAK52302 30 AAK52301 30 AAK52300 30 AAK62766 32 AAK62765 32 AAK62764 32 AAK62763 32 AAK62762 32 AAK62761 32 AAK62760 32 AAK62759 32 AAK62758 32 AAK62757 32 AAK62756 32 AAL07765 6 AAL07764 6 AAL07763 6 AAL07762 6 AAL07761 6 AAK91592 20 AAM81753 97 AAM81752 97 AAM81751 97 AAM81750 97 AAM81749 97 AAM21941 32 AAK67141 7 AAK67140 7 AAK67139 7 AAK67138 7 AAK67137 7 AAK67136 7 AAK67135 7 id=14550088 7 id=14550086 7 id=14550084 7 id=14550082 7 id=14550080 7 id=14550078 7 id=14550076 7 id=14550074 7 id=14550072 7 id=14550070 7 AAK67124 3 AAK67123 7 AAK67122 7 AAK67121 7 AAK67120 7 AAK67119 7 AAK67118 7 AAK67117 7 AAK67116 7 AAK67115 7 AAK67114 7 AAK67113 7 AAK67112 7 AAK67111 7 AAK67110 7 AAK67109 7 AAK67108 7 AAK67107 7 AAK67106 7 AAK67105 7 AAK67104 7 AAK67103 7 AAK67102 7 AAK67101 7 AAK67100 7 AAK67099 7 AAK67098 7 AAK67097 7 AAK67096 7 AAK67095 7 AAK67094 7 AAK67093 7 AAK67092 7 AAK67091 7 AAK67090 7 AAK67089 7 AAK67088 7 AAK67087 7 AAK67086 7 AAK67085 7 AAK67084 7 AAK67083 7 AAK67082 7 AAK67081 7 AAK67080 7 AAK67079 7 AAK67078 7 AAK67077 7 AAK67076 7 AAK67075 7 AAK67074 7 AAK67073 7 AAK67072 7 AAK67071 7 AAK67070 5 AAK67069 7 AAK67068 7 AAK67067 7 AAK67066 7 AAK67065 7 AAK67064 7 AAL87748 19 AAL87747 18 AAL87746 19 AAL87745 18 AAL37596 18 AAM21944 24</u></p>	129	3.6	2.0	low p <.02, prev p<.001	66	9
2002	<p><u>AAO26579 30 AAO26578 30 AAN77484 3 AAM09856 6 AAM09855 6 AAM09854 6 AAN85090 97 AAO73303 36 AAO73302 36 AAO73301 36 AAO73300 36 AAO73299 36 AAO73298 36 AAO73297 36 AAO73296 36 AAO73295 36 AAL87234 96</u></p>	17	4.8	1.4	low p <.001, prev p<.002	4,156	284
2003	<p><u>AAP20887 96 AAR17575 32 AAR17574 32 AAR17573 32 AAR17572 32 AAR17571 32 AAR17570 32 AAR17569 32 AAR17568 32 AAR17567 32 AAR17566 32 AAR17565 32 AAR17564 32 AAR17563 32 AAR17562 32 AAR17561 32 AAR17560 32 AAR17559 32 AAR17558 32 AAR17557 32 AAR17556 32 AAR17555 32 AAR17554 32 AAR17553 32 AAR17552 32 AAR17551 32 AAR17550 32 AAR17549 32 AAR17548 32 AAR17547 32 AAR17546 32 AAR17545 32 AAR17544 32 AAR17543 32 AAR17542 32 AAQ87608 16 AAQ87607 16 AAQ87606 14 AAR10804 6 AAR10803 6 AAR10802 6 AAR10801 6 AAR10800 6 AAR10799 6 AAR10798 6 AAR10797 6 AAR10796 6 AAR10795 6 AAR10794 6 AAR10793 6 AAR10792 6 AAR10791 6 AAR10790 6 AAR10789 6 AAR10788 6 AAR10787 6 AAR10786 6 AAR10785 6 AAR10784 6 AAR10783 6 AAR10782 6 AAR10781 6 AAR10780 6 AAQ88403 10 AAQ88402 10 id=40288320 36 AAQ55854 97 AAR14153 36 id=92919472 97 AAR84614 95 AAR06948 36 AAR06947 36 id=37993725 36 id=37993723 36 id=37993721 36 id=37993719 36 id=37993717 36 AAR06941 36 AAR06940 36 AAR06939 36 AAR06938 36 AAR06937 36 id=37993705 35 id=37993703 36 id=37993701 36 id=37993699 36 id=37993697 36 AAR06931 36 AAQ00999 100 AAQ00998 97 AAP22087 97 AAP22086 97 AAP22089 97 AAP22088 96 AAP85247 6 AAP85246 6 AAP85245 6 AAP85244 6</u></p>	107	5.2	1.5	low p <.001, prev p<.30	9,862	264

	<p><u>AAP85243 6 AAP85242 6 AAP85241 6 AAP85240 6 AAP85239 6</u> <u>AAP85238 6 AAP85237 6 AAP78942 95 AAP78941 95</u></p>					
2004	<p><u>1S6NA 4 AAT11553 32 AAT11552 32 AAT11551 32 AAT11550 32</u> <u>AAT11549 32 AAT11548 32 AAT11547 32 AAT11546 32</u> <u>AAT11545 32 AAT11544 32 AAT11543 32 AAT11542 32</u> <u>AAT11541 32 AAT11540 32 AAT11539 32 AAT11538 32</u> <u>AAT11537 32 AAT11536 32 AAT11535 32 AAT11534 32</u> <u>AAS75296 6 AAS75295 6 AAS75294 6 AAS75293 6 AAS75292 6</u> <u>AAS75291 6 id=51095222 108 AAU00153 96 id=55669122 97</u> <u>BAD34490 97 BAD34489 97 BAD34488 97 AAV68177 97</u> <u>id=73913544 106 id=59876233 97 AAT92099 97 AAT92098 97</u> <u>AAT02759 111 AAV52690 96 AAV52689 97 AAV52688 97</u> <u>AAV52687 97 AAV49728 6 AAV49727 6 AAV49726 6 AAV49725</u> <u>6 AAV49724 6 AAW56064 97 AAW56066 97 AAW56065 97</u> <u>AAW28871 97</u></p>	52	4.3	1.8	low p <.001 , prev p<.00 2	2,53 100
2005	<p><u>ABC18309 8 ABC18308 9 ABC02196 3 1ZTXE 5 AAY67877 9</u> <u>AAY67876 11 AAY67875 11 AAY67874 8 AAY67873 8</u> <u>AAY67872 8 AAY67871 8 AAY67870 8 AAY67869 8 AAY67868 8</u> <u>AAY67867 8 AAY67866 8 AAY57985 8 ABB01532 97 AAZ32750</u> <u>97 AAZ32749 97 AAZ32748 97 AAZ32747 97 AAZ32746 97</u> <u>AAZ32745 97 AAZ32744 97 AAZ32743 97 AAZ32742 97</u> <u>AAZ32741 97 AAZ32739 97 AAZ32737 97 AAZ32736 97</u> <u>AAZ32734 97 AAZ32733 97 id=71483607 97 id=71483605 97</u> <u>id=71483603 97 id=71483601 97 ABC40712 100 ABB01533 101</u> <u>AAV55949 97 ABA62343 97 id=63098704 36 id=63098702 36</u> <u>AAV29684 6 AAY29685 6 AAY29683 6 AAY29682 6 AAY29681 6</u> <u>AAV29680 6 AAY29679 6 AAY29678 6 AAY29677 7 AAY29676 7</u> <u>id=84028433 111 id=76446583 37 ABA43045 37 ABA43044 37</u> <u>ABA43043 37 ABA43042 37 ABA43041 37 ABA43040 37</u> <u>ABA43039 37 ABA43038 37 ABA43037 37 ABA43036 37</u> <u>ABA43035 37 ABA43034 37 ABA43033 37 ABA43032 37</u> <u>ABA43031 37 ABA43030 37 ABA43029 37 ABA43028 37</u> <u>ABA43027 37 ABA43026 37 ABA43025 37 ABA43024 37</u> <u>ABA43023 37 ABA43022 37 ABA43021 37 ABA43020 37</u> <u>ABA43019 37 ABA43017 37 ABA43016 37 id=76446521 37</u> <u>id=76446519 37 id=76446517 37 id=76446515 37 id=76446513 37</u> <u>ABA43010 37 ABA43009 37 ABA43008 37 ABA43007 37</u> <u>ABA43006 37 id=76446501 37 id=76446499 37 id=76446497 37</u> <u>id=76781572 105 id=76781570 105 ABA54593 105 ABA54592 105</u> <u>ABA54591 105 ABA54590 105 ABA54589 105 ABA54588 105</u> <u>ABA54587 105 ABA54586 105 ABA54585 105 ABA54584 105</u> <u>ABA54583 105 ABA54582 105 ABA54581 105 ABA54580 105</u> <u>ABA54579 105 id=76781538 105 id=76781536 105 id=76781534</u> <u>105 id=76781532 105 AAY54162 97</u></p>	119	4.4	1.9	low p <.001 , prev p>.50	3,00 119
2006	<p><u>ABI81406 34 ABI81405 34 ABI81404 34 ABI81403 34 ABI81402</u> <u>34 ABI81401 34 ABI81400 34 ABI81399 34 ABI81398 34</u> <u>ABI81397 34 ABI81396 34 ABI81395 34 ABI81394 34 ABI81393</u> <u>34 ABI81392 34 ABI81391 34 ABI81390 34 ABI81389 34</u> <u>ABI81388 34 ABI81387 34 ABI81386 34 ABI81385 34 ABI81384</u> <u>34 ABI81383 34 ABI81382 34 ABI81381 34 ABI81380 34</u> <u>ABI81379 34 ABI81378 34 ABI81377 34 ABI81376 34 ABI81375</u> <u>34 ABI81374 34 ABI81373 34 ABI81372 34 ABI81371 34</u> <u>ABI81370 34 ABI81369 34 ABI81368 34 ABI81367 34 ABI81366</u> <u>34 ABI81365 34 ABI81364 34 ABI81363 34 ABI81362 34</u> <u>ABI81361 34 ABI81360 34 ABI81359 34 ABI81358 34 ABI81357</u> <u>34 ABI81356 34 ABI81355 34 ABI81354 34 ABI81353 34</u> <u>ABI81352 34 ABI81351 34 ABI81350 34 ABI81349 34 ABI81348</u> <u>34 ABI81347 34 ABI81346 34 ABI81345 34 ABI81344 34</u> <u>ABI81343 34 ABI81342 34 ABI81341 34 ABI81340 34 ABI81339</u></p>	279	6.4	1.4	low p <.001 , prev p<.00 1	4269 177

<p><u>34 ABI81338 34 ABI81337 34 ABI81336 34 ABI81335 34 ABI81334 34 ABI81333 34 ABI81332 34 ABI81331 34 ABI81330 34 ABI81329 34 ABI81328 34 ABI81327 34 ABI81326 34 ABI81325 34 ABI81324 34 ABI81323 34 ABI81322 34 ABI81321 34 ABI81320 34 ABI81319 34 ABI81318 34 ABI81317 34 ABI81316 34 ABI81315 34 ABI81314 34 ABI81313 34 ABI81312 34 ABI81310 34 ABI81308 34 ABI81307 34 ABI81306 34 ABI81305 34 ABI81304 34 ABI81303 34 ABI81302 34 ABI81301 34 ABI81300 34 ABI81299 34 ABI81298 34 ABI81297 34 ABI81296 34 ABI81295 34 ABI81294 34 ABI81293 34 ABI81292 34 ABI81291 34 ABI81290 34 ABI81289 34 ABI81288 34 ABI81287 34 ABI81286 34 ABI81285 34 ABI81284 34 ABI81283 34 ABI81282 34 ABI81281 34 ABI81280 34 ABI81279 34 ABI81278 34 ABI81277 34 ABI81276 34 ABI81275 34 ABI81274 34 ABI81273 34 ABI81272 34 ABI81271 34 ABI81270 34 ABI81269 34 ABI81268 34 ABI81267 34 ABI81266 34 ABI81265 34 ABI81264 34 ABI81263 34 ABI81262 34 ABI81260 34 ABI81259 34 ABI81258 34 ABI81257 34 ABI81256 34 ABI81255 34 ABI81254 34 ABI81253 34 ABI81252 34 ABI81251 34 ABI81250 34 ABI81249 34 ABI81248 34 ABI81247 34 ABI81246 34 ABI81245 34 ABI81244 34 ABI81243 34 ABI81242 34 ABI81241 34 ABI81240 34 ABI81239 34 ABI81238 34 ABI81237 34 ABI81236 34 ABI81235 34 ABI81234 34 ABI81233 34 ABI81232 34 ABI81231 34 ABI81230 34 ABI81229 34 ABI81228 34 ABJ90133 32 ABJ90132 32 ABJ90131 32 ABJ90130 32 ABJ90129 32 ABJ90128 32 ABJ90127 32 ABJ90126 32 ABJ90125 32 ABJ90124 32 ABJ90123 32 ABJ90122 32 ABJ90121 32 ABJ90120 32 ABJ90119 32 ABJ90118 32 ABJ90117 32 ABJ90116 32 ABJ90115 32 ABJ90114 32 ABJ90113 32 ABJ90112 32 ABJ90111 32 ABJ90110 32 ABJ90109 32 ABJ90108 32 ABJ90107 32 ABJ90106 32 ABJ90105 32 ABJ90104 32 ABJ90103 32 ABJ90102 32 ABJ90101 32 ABJ90100 32 ABJ90099 32 ABJ90098 32 ABJ90097 32 ABJ90096 32 ABJ90095 32 ABJ90094 32 ABJ90093 32 ABJ90092 32 ABJ90091 32 ABJ90090 32 ABJ90087 32 ABJ90086 32 ABJ90085 32 ABJ90084 32 ABJ90083 32 ABJ90082 32 ABJ90080 32 ABJ90079 32 ABJ90078 32 ABJ90077 32 ABJ90076 32 ABJ90075 32 ABJ90074 32 ABJ90073 32 ABJ90072 32 ABJ90071 32 ABJ90070 32 ABJ90069 32 ABJ90068 32 ABJ90067 32 ABJ90066 32 ABD85083 99 ABD85082 99 ABD85081 99 ABD85080 99 ABD85079 99 ABD85078 99 ABD85077 99 ABD85076 99 id=90025142 99 id=90025140 99 id=90025138 99 id=90025136 99 id=90025134 99 ABD85070 99 ABD85069 99 ABD85068 99 ABD85067 99 ABD85066 99 ABD85065 99 ABD85064 99 ABG36517 36 ABD19513 97 ABD19512 96 ABD19511 97 ABD19510 97 ABI26622 40 ABI26621 40 id=89340787 97 id=89340785 97 id=89340783 97 ABD67759 97 ABD67758 97 ABD67757 97 ABD67756 97 ABD19642 97 id=87116125 97 id=87116123 97 2I69A 24</u></p>					
<p>2007 <u>ABQ52692 97 ABO69610 36 ABO69609 36 ABO69608 36 ABO69607 36 ABO69606 36 ABO69605 36 ABO69604 36 ABO69603 36 ABO69602 36 ABO69601 36 ABO69600 36 ABO69599 36 ABO69598 36 id=134285072 36 id=134285070 36 id=134285068 36 id=134285066 36 id=134285064 36 ABO69592 36</u></p>	20	5.3	0.6	low p <.001, prev p<.001	

[000280] Upon analysis of Replikin Counts of publicly available sequences from the entire genome of WNV and comparison with WNV morbidity and mortality data from the United States Center for Disease Control, the applicants observed that the mean Replikin

Count of WNV increased significantly between years 2000, 2004, 2005 and 2006, respectively. As seen in Table 11, the mean Replikin Count of 2.8 ± 0 observed in 2000 was found to be significantly different ($p < 0.001$) from the mean Replikin Count of 3.8 ± 1.7 observed in 2004, the mean Replikin Count observed in 3.8 ± 1.7 in 2004 was found to be significantly different ($p < 0.01$) from the mean Replikin Count observed in 4.5 ± 1.8 in 2005, and, finally, the mean Replikin Count observed in 4.5 ± 1.8 in 2005 was found to be significantly different ($p < 0.001$) from the mean Replikin Count observed in 6.0 ± 1.1 in 2006.

Table 11 – WNV Whole Genome

Year	Accession Records for WNV	Replikin Count	Standard Deviation	Significance (compared to previous listed year)	Morbidity	Mortality
2000	2	2.8	± 0.0		CDC 21	CDC 2
2004	68	3.8	± 1.7	prev $p < 0.001$	CDC 2,539	CDC 100
2005	137	4.5	± 1.8	prev $p < 0.01$	CDC 3000	CDC 119
2006	211	6.0	± 1.1	prev $p < 0.001$	CDC 4269	CDC 177

[000281] In the summer of 2007, Applicants reviewed the data for the whole WNV genome in publicly available sequences as provided in Table 11 and expressly predicted that a virulent increase in infection of WNV would likely follow the significant increase observed between each of the analyzed years. Immediately after Applicants' prediction, the California Department of Public Health confirmed Applicants' prediction by reporting that infections of WNV in California through August 2, 2007 had been three times greater than infections seen in the previous year and a health emergency for three California counties was declared.

[000282] The epidemiology and virology of WNV is different from the epidemiology and virology of some other viruses discussed herein such as influenza, FMDV, PRRSV and PCV. Nevertheless, a correlation between increases in Replikin Count in the WNV envelope protein and morbidity and mortality provides compounding data establishing a shared phenomenon of rapid replication and virulence with an overwhelming number of other tested viruses and organisms.

[000283] In WNV and the other viruses and pathogens described herein, prediction of epidemics and future outbreaks may be made, for example, by (1) reviewing the Replikin Counts of isolates of WNV and identifying a RPG, for example, and RPG in the envelope protein (e.g., SEQ ID NO: 245), (2) comparing the Replikin Counts in the RPG, in the protein or gene area containing the RPG, or in the whole virus genome for a particular year with Replikin Counts from other years. A significant increase in Replikin Count from one year to

the next and preferably over one, two or three years provides predictive value of an emerging strain of WNV that may begin an outbreak of more highly virulent WNV. A WNV outbreak may be predicted within about six months to about one year or more from the observation of a significant increase in Replikin Count.

[000284] More preferably, an outbreak of WNV may be predicted within about six months to about one year from the observation of a significant increase in Replikin Count over two or three years or, as in inventors' prediction in 2007, following the observation of strongly significant increases over several years such as wherein Replikin Counts between 2000, 2004 and 2006 had p values of less than at least 0.001 and frequently less than 0.001. As such, significant increases may be observed over a time period of more than one year, such as three, four, five or more years. An outbreak may likewise be predicted within about six months to about one year from the initial observation of an observable decrease in Replikin Count following a significant increase. Using this method, Applicants prospectively predicted the beginnings of a 2007 outbreak of WNV. The method may also employ isolates of individual strains or isolates of all strains of WNV.

[000285] An embodiment of the invention provides a segment of the genome or a protein or segment of a protein of the WNV in which the expressed gene or expressed gene segment has the highest concentration of Replikins, or Replikin Count (number of Replikins per 100 amino acids), when compared to other segments or named genes of the genome, namely the RPG. An RPG (SEQ ID NO:245) in Accession No. ABA54585 is reported in Example 7 below. Twelve Replikin sequences (SEQ ID NOS:246-257) are identified in the RPG diagnostic, preventive, therapeutic and predictive applications. These Replikin sequences are preferred embodiments of immunogenic compositions and vaccines. The invention further provides Replikin sequences within the identified RPG that are conserved in the genome over time and, as such, are available as relatively invariant preferred targets for diagnosis and manipulation of rapid replication and virulence in WNV through immunogenic responses and vaccines.

IX. Methods of Predicting and Treating Outbreaks of Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) Using Replikin Sequences

[000286] An increase in Replikin concentration in PRRSV is predictive of an increase in virulence of the virus and allows for prediction of forthcoming outbreaks or increases in mortality. A review of publicly available amino acid sequences of isolates of PRRSV that demonstrate an increase in Replikin concentration in the genome or a genome segment, or in a protein or protein fragment of the virus over time or between isolates is used as a predictor

of an increase in outbreaks and morbidity and mortality of pigs infected with PRRSV. Publicly available sequences for isolates of PRRSV from PubMed or other public or private sources may be analyzed by hand or using proprietary search tool software (ReplikinForecast™ from REPLIKINS LLC, Boston, MA).

[000287] The inventors have now identified a Replikin Peak Gene in the nucleocapsid protein of the Porcine Respiratory and Reproductive Syndrome Virus (PRRSV) and have demonstrated a correlation between increased Replikin Count in the nucleocapsid protein of PRRSV between 2004 and 2007 and major outbreaks of PRRSV in China. Example 8.

[000288] Figure 13 illustrates Replikin Counts in the nucleocapsid protein of PRRSV (SEQ ID NO: 353). The Replikin Count is seen to increase between 2004 and 2007. This increase correlates with a major outbreak of Porcine Reproductive and Respiratory Syndrome in China. Further, standard deviation from the mean in 2005 is considerably larger than other years demonstrating a marked increase in Replikin Count was occurring in 2005. The large increase that was occurring in 2005 based on increases in standard deviation is confirmed as an increase in mean Replikin Count in 2006. The large standard deviation observed in 2005 indicates that more members of the class had increasing Replikin Counts. Standard deviation in 2005 was an early warning prior to the increase in the mean count in 2006 and 2007. A similar phenomenon is observable in Figure 7. These data provide further confirmation of the predictive value of the RPG Replikin Count in viral outbreaks and provide specific support for RPG Replikin Count as a predictive tool in PRRSV and viruses in pigs generally.

[000289] The invention provides RPGs and Replikin sequences within the identified RPGs for diagnostic, preventive and therapeutic applications. For example, each Replikin sequences identified within an identified RPG in PRRSV and other viruses, organisms and malignancies is available for diagnostic and therapeutic applications including vaccines, immunogenic compositions and antibody therapies. The entire Replikin Peak Gene sequence or fragments thereof are likewise available for diagnostic, preventive, therapeutic and predictive applications. Further, the presence of the Replikin Peak Gene in an isolate of the virus is indicative of rapid replication.

[000290] As discussed herein, applicants have identified RPGs of available PRRSV isolates within the nucleocapsid protein of PRRSV. Identification of these RPGS is different, for example, from the Replikin Peak Gene previously identified by applicants in H5N1 influenza in one polymerase area, namely the RNA-directed RNA polymerase or pB1 protein. Identification of Replikin Peak Genes in different structures of different viruses is made

possible through the strict criteria for a Replikin sequence as defined by the applicants. The proprietary software ReplikinForecast™ (licensable from REPLIKINS LLC, Boston, MA) provides an efficient survey of publicly available Replikin sequences and identification and isolation *in silico* of the Replikin Peak Gene.

[000291] The size of a Replikin Peak Gene, both in terms of the number of amino acids and the Replikin Count, will depend upon the size of the sequence of the entire genome, protein or fragment thereof that has been isolated and reported. The invention further provides Replikin sequences within the identified Replikin Peak Gene or Area that are conserved in the genome over time and, as such, are available as relatively invariant targets for diagnosis and manipulation of rapid replication and virulence in PRRSV.

[000292] Further, the following RPGs have been identified in PRRSV isolates from China reported at Accession Nos. AAM18565, AAP81809 and ABL60920, respectively:

- (1) k⁷ q⁸ q⁹ k¹⁰ k¹¹ k¹² k¹³ g¹⁴ n¹⁵ g¹⁶ q¹⁷ p¹⁸ v¹⁹ n²⁰ q²¹ l²² c²³ q²⁴ m²⁵ l²⁶ g²⁷ k²⁸ i²⁹ i³⁰ a³¹ q³² q³³ n³⁴ q³⁵ s³⁶ r³⁷ g³⁸ k³⁹ g⁴⁰ p⁴¹ g⁴² k⁴³ k⁴⁴ **s⁴⁵ k⁴⁶** k⁴⁷ k⁴⁸ n⁴⁹ p⁵⁰ e⁵¹ k⁵² p⁵³ h⁵⁴ f⁵⁵ p⁵⁶ l⁵⁷ a⁵⁸ t⁵⁹ e⁶⁰ d⁶¹ d⁶² v⁶³ r⁶⁴ h⁶⁵ h⁶⁶ (China 2000) (SEQ ID NO: 341)
- (2) k⁷ q⁸ q⁹ k¹⁰ r¹¹ k¹² k¹³ g¹⁴ d¹⁵ g¹⁶ q¹⁷ p¹⁸ v¹⁹ n²⁰ q²¹ l²² c²³ q²⁴ m²⁵ l²⁶ g²⁷ k²⁸ i²⁹ i³⁰ a³¹ q³² q³³ n³⁴ q³⁵ s³⁶ r³⁷ g³⁸ k³⁹ g⁴⁰ p⁴¹ g⁴² k⁴³ k⁴⁴ **n⁴⁵ k⁴⁶** k⁴⁷ k⁴⁸ n⁴⁹ p⁵⁰ e⁵¹ k⁵² p⁵³ h⁵⁴ f⁵⁵ p⁵⁶ l⁵⁷ a⁵⁸ t⁵⁹ e⁶⁰ d⁶¹ d⁶² v⁶³ r⁶⁴ h⁶⁵ h⁶⁶ (China 2003) (SEQ ID NO:342), and
- (3) k⁷ q⁸ q⁹ k¹⁰ k¹¹ k¹² k¹³ g¹⁴ n¹⁵ g¹⁶ q¹⁷ p¹⁸ v¹⁹ n²⁰ q²¹ l²² c²³ q²⁴ m²⁵ l²⁶ g²⁷ k²⁸ i²⁹ i³⁰ a³¹ q³² q³³ n³⁴ q³⁵ s³⁶ r³⁷ g³⁸ k³⁹ g⁴⁰ p⁴¹ g⁴² k⁴³ k⁴⁴ **n⁴⁵ r⁴⁶** k⁴⁷ k⁴⁸ n⁴⁹ p⁵⁰ e⁵¹ k⁵² p⁵³ h⁵⁴ f⁵⁵ p⁵⁶ l⁵⁷ a⁵⁸ t⁵⁹ e⁶⁰ d⁶¹ d⁶² v⁶³ r⁶⁴ h⁶⁵ h⁶⁶ (China 2006) (SEQ ID NO:343).

[000293] The identified RPG sequences are identical across the 2000, 2003 and 2006 isolates except for point mutations at positions 45 and 46 (underlined in bold). These sequences are, therefore, relatively invariant targets for diagnosis and manipulation of rapid replication and virulence in PRRSV and are available as vaccines against the disease.

[000294] Point mutations, such as in positions 45 and 46 in the above-listed Chinese isolates, provide excellent predictive capacity. In the highly virulent and fatal Chinese variant disclosed in 2006 at ABL60920 (SEQ ID NO: 343), the asparagine and arginine at positions 45 and 46 are the same residues in the same relative positions as asparagine and arginine at residues 21 and 22 in the RPG of the highly virulent PRRSV 2006 Mexican isolate publicly available at Accession No. ABF19568 (comparable mutated residues underlined in bold): k¹⁴ g¹⁵ p¹⁶ g¹⁷ k¹⁸ k¹⁹ k²⁰ **n²¹ r²²** k²³ r²⁴ n²⁵ p²⁶ e²⁷ k²⁸ p²⁹ h³⁰ f³¹ p³² l³³ a³⁴ t³⁵ e³⁶ d³⁷ d³⁸ v³⁹ r⁴⁰ h⁴¹ h⁴² (SEQ ID NO:344).

[000295] These two RPG sequences are, therefore, especially predictive of virulence and are preferred sequences for immunogenic compositions and vaccines. Identification of these residues in other RPG sequences in PRRSV provides a high likelihood of virulence and an excellent target for attack of the virus through antibody therapies, vaccines and other treatments.

X. Methods of Predicting and Treating Outbreaks of Porcine Circovirus (PCV) Using Replikin Sequences

[000296] An increase in Replikin concentration in PCV is predictive of an increase in virulence of the virus and allows for prediction of forthcoming outbreaks or increases in mortality. A review of publicly available amino acid sequences of isolates of PCV that demonstrate an increase in Replikin concentration in the genome or a genome segment, or in a protein or protein fragment of the virus over time or between isolates is used as a predictor of an increase in outbreaks and morbidity and mortality of pigs infected with PCV. Publicly available sequences for isolates of PCV from PubMed or other public or private sources may be analyzed by hand or using software described herein.

[000297] Applicants have now established a correlation between Replikin Counts in PCV and an increase in virulence. Applicants reviewed publicly available amino acid sequences of isolates of PCV having accession numbers at www.pubmed.com and identified increases in Replikin Counts in the genome of the virus that predict an increase in outbreaks and mortality of pigs infected with PCV.

[000298] The data for Figure 21 is provided in Table 25 in Example 15 below. A general increase in Replikin Count from 2000 through 2007 is observable and may be correlated with an increase in incidence of and mortality from the disease between 2000 and 2006 as reported in Canada. Further, the very large Replikin Count number in 1997 followed by a marked decrease in 1998 through 2000 may be correlated with the beginning of increased outbreaks in 2000. In other viruses, outbreaks have been observed about 1 to 3 years after a large increase in Replikin Count that is followed by a notable decrease thereafter. *See, e.g.,* Figures 2, 3 and 9. The graph in Figure 21 demonstrates a cyclical pattern of Replikin Counts that is reminiscent of the correlation of Replikin Count with epidemics shown, for example, in influenza and SARS in Figures 2, 3 and 9.

[000299] In particular, the Replikin Count of PCV is observed at 9.4 (\pm 10.8) in 1997 and decreases rapidly to 2.9 (\pm 1.2) in 2000. Replikin Count then rises to 3.5 (\pm 1.4) in 2002 and rises again to 3.9 (\pm 1.2) through 2007. During this time period, the virulence and mortality observed in swine herds in Canada (with additional reported incidence in Central America)

were increasing. The large standard deviation seen in 1997-1999 evidences a virus population that is undergoing rapid change in the concentration of Replikin sequences in the genome and points to forthcoming changes in virulence, morbidity and mortality.

[000300] Prediction of epidemics and future outbreaks may be made, for example, by reviewing the Replikin Counts of RPGs or other portions of isolates of PCV or PRRSV or other virus or pathogen and comparing the Replikin Counts for a particular year with Replikin Counts from other years. A significant increase in Replikin Count from one year to the next and preferably over one, two or three or more years provides predictive value of an emerging strain of PCV that may begin an outbreak of more highly virulent and/or more highly lethal PCV.

[000301] A PCV outbreak may be predicted within about six months to about one year or more from the observation of a significant increase in Replikin Count. More preferably, an outbreak of PCV may be predicted within about six months to about one year from the observation of a significant increase in Replikin Count over two or three years or following the observation of strongly significant increases over several years such as wherein Replikin Counts of PCV between 2000 and 2002 and between 2005 and 2007 increased with p values each year over lowest mean Replikin Count in the series of less than 0.001.

[000302] Significant increases may be observed over a time period of more than one year, such as three, four or five years or more. An outbreak may likewise be predicted within about six months to about one year or more from the initial observation of an observable decrease in Replikin Count following a notable increase. For example, the marked decrease from 1997 to 2000 in PCV Replikin Counts predicts the increase of incidence and mortality in viral infections beginning in 2000 and continuing through at least 2006 (morbidity and mortality data for 2007 have not been made available at this time). Using this method, Applicants, for example, prospectively predicted the beginnings of a 2007 outbreak of WNV. *See Figure 12.*

[000303] The inventors have identified a Replikin Peak Gene in the replicase protein of the Porcine Circovirus (PCV). Examples of the identification of a Replikin Peak Gene (RPG) in an isolate of PCV in Manitoba, Canada in 1997 and an RPG in an isolated of PCV in China in 2007 are provided in Example 9 (SEQ ID NOS:520 and 525). Example 9 demonstrates comparably high Replikin Counts of the identified RPGs and provides prediction that the isolated strains of the virus have high virulence. Example 9 further provides RPGs and

Replikin sequences within the identified RPGs as targets for production of immunogenic compositions and vaccines.

[000304] The invention provides Replikin sequences within the identified Replikin Peak Gene gene or gene segment for diagnostic, preventive and therapeutic applications. SEQ ID NOS:324-328 are Replikin sequences provided in an RPG from Accession No. AAC59472. See Example 9. SEQ ID NOS:329-340 are provided in an RPG from Accession No. ABP68657. See Example 9. For example, each of the above-listed sequences as Replikin sequences identified within an identified RPG are available for diagnostic and therapeutic applications including vaccines and antibody therapies. The entire Replikin Peak Gene sequence or fragments thereof are likewise available for diagnostic, preventive, therapeutic and predictive applications. Further, the presence of the Replikin Peak Gene in an isolate of the virus is indicative of rapid replication.

[000305] Replikin Peak Genes (RPG) have also been identified in PCV isolates in Accession Nos. AAC98885, AAL01075 and ABP68667 (SEQ ID NOS: 481, 438, and 451). See Example 9. For each identified RPG, continuous, non-interrupted and overlapping Replikin sequences have been identified for predictive and therapeutic applications.

[000306] Applicants have to date identified RPGs of available PCV isolates both within open reading frame 1 in a putative replicase protein and within open reading frame 11 in a predicted 1.8 kD protein. Identification of Replikin Peak Genes in different structures of different viruses is made possible through the strict criteria for a Replikin sequence as defined by the applicants. The size of a Replikin Peak Gene, both in terms of the number of amino acids and the Replikin Count, will depend upon the size of the sequence of the entire genome, protein or fragment thereof that has been isolated and reported. The invention further provides Replikin sequences within the identified Replikin Peak Gene that are conserved in the genome over time and, as such, are available as relatively invariant targets for diagnosis and manipulation of rapid replication and virulence in PCV.

XI. Conservation of Replikin Structure Relates to Virulence and Lethality

[000307] The conservation of any structure is critical to whether that structure provides a stable invariant target to attack and destroy or to stimulate. Replikin sequences have been shown to generally be conserved. When a structure is tied in some way to a basic survival mechanism of the organism, the structures tend to be conserved. A varying structure provides an inconstant target, which is a good strategy for avoiding attackers, such as antibodies that have been generated specifically against the prior structure and thus are

ineffective against the modified form. This strategy is used by influenza virus, for example, so that a previous vaccine may be quite ineffective against the current virulent virus.

[000308] Certain structures too closely related to survival functions, however, apparently cannot change constantly. An essential component of the Replikin structure is histidine (h), which is known for its frequent binding to metal groups in redox enzymes and is a probable source of energy needed for replication. Since the histidine structure remains constant, Replikin sequence structures remain all the more attractive a target for destruction or stimulation.

A. Replikin Conservation in HIV

[000309] Conservation of Replikin sequences has been observed in trans-activator (Tat) proteins in isolates of HIV. Tat (trans-activator) proteins are early RNA binding proteins regulating lentiviral transcription. These proteins are necessary components in the life cycle of all known lentiviruses, such as the human immunodeficiency viruses (HIV). Tat is a transcriptional regulator protein that acts by binding to the trans-activating response sequence (TAR) RNA element and activates transcription Initiation and/or elongation from the LTR promoter. HIV cannot replicate without tat, but the chemical basis of this has been unknown. In the HIV tat protein sequence from 89 to 102 residues, we have found a Replikin that is associated with rapid replication in other organisms. The amino acid sequence of this Replikin is hclvckqkkglgisygrkk (SEQ ID NO:3666) In fact, Applicants found that this Replikin is present in every HIV tat protein. Some tat amino acids are substituted frequently, as shown in Table 12, by alternate amino acids (in small size fonts lined up below the most frequent amino acid, the percentage of conservation for the predominant Replikin (hclvcfqkkglgisygrkk) (SEQ ID NO:3314). These substitutions have appeared for most of the individual amino acids. However, the key lysine and histidine amino acids within the Replikin sequence, which define the Replikin structure, are conserved 100% in the sequence; while substitutions are common elsewhere in other amino acids, both within and outside the Replikin, none occurs on these key histidine amino acids. The sequences listed in Table 12 are SEQ ID NO:3314 and the denoted variations of formula peptide SEQ ID NO:3315.

[000310] The substitutions cannot be considered to be at random because amino acids were substituted except for the lysines and histidines which define the Replikin structure. It is not just that lysine per se is “immune” to substitution, because the lysine not 6 to 10 amino acids from another lysine was freely substituted, while those lysines which do define the Replikin structure were not substituted.

Table 12 – HIV TAT Conservation (SEQ ID NOS: 3314 and 3671)

% Replikin CONSERVATION of each constituent amino acid in the first 117 different isolates of HIV tat protein as reported in PubMed:

38 (100) 57 86 (100) (100) 66 76 (100) 99 57 49 (100) 94 (100) 97 98 85 97 99 (100)(100)(100)%

Neighboring

Amino acids

tat Replikin

k (e) s y [(h) (c) l v (c) f q k (k) g (l) g i s y g (r) (k) (k)]

below are the amino acid substitutions observed for each amino acid above:

h	c	f	q	i	l	b	t	a	a	l	y	h	q
r	w	p	l	l	l	i	h	q	v				
y	s		s		l	m		r	s				
i			s		m	s							
s			r		n								
v													
a													
f													
p													
q													

B. Conservation in Replikin Peak Genes in H5N1 in Humans and Chickens

[000311] A series of conserved Replikin sequences (SEQ ID NOS:1-11 and 14) were isolated *in silico* by Applicants in human and chicken isolates of H5N1 influenza virus. SEQ ID NO:1 was identified in the following accession numbers in the following years at the following amino acid residue positions: (1997) AAK49342 beginning at position 134, AAK49340, 134, AAF74320, 134, AAF74319, 134, AAF74318, 134, AAF74317, 134, AAK49344, 134, AAK49343, 134, AAK49341, 134, AAK49339, 134, AAK49338, 134; (1998) AAK49345, 134; (2003) BAE07200, 134; (2004) AAW59551, 131, AAW59549, 129, ABE97897, 123, ABE97896, 123, ABE97895, 123, ABE97892, 123, ABE97891, 123, AAV32651, 134, AAV32643, 134; (2005) ABG78563, 109, ABG78562, 109, ABF56657, 127, ABF56656, 127, ABF56655, 127; (2006) ABK34973, 134, ABL31779, 134, ABL31765, 134, ABL31754, 134, ABL07029, 134, ABL07018, 119, ABL07007, 134, ABI49406, 134, ABI36481, 134, ABI36470, 134, ABI36451, 134, and ABI36440, 134.

[000312] SEQ ID NO:11 was identified in the following accession numbers in the following years at the following amino acid residue positions: (2003) BAE07200, beginning at position 19; (2004) AAW59551, 16, AAW59549, 14, ABE97897, 8, ABE97896, 8, ABE97895, 8, ABE97894, 8, ABE97893, 8, ABE97892, 8, ABE97891, 8, ABE97890, 8, ABE97889, 8, ABE97888, 8, AAV35115, 19, AAV32651, 19, AAV32643, 19; (2005) ABC72649, 19, ABF56657, 12, ABF56656, 12, ABF56655, 12; (2006) ABK34973, 19, ABL31779, 19, ABL31765, 19, ABL31754, 19, ABL31743, 19, ABI49414, 19, ABL07029, 19, ABL07018, 4, ABL07007, 19, ABI49406, 19, ABI36481, 19, ABI36470, 19, ABI36451, 19, ABI36440, 19, ABI36429, 19.

[000313] SEQ ID NO:14 was identified in the following accession numbers in 2006 at the following amino acid residue positions: ABL31777, beginning at position 41, ABI49393, 41, ABL07016, 41, ABL07005, 41, ABI49404, 41, ABI36472, 41, ABI36461, 41, ABI36452, 41, ABI36441, 41, and ABI36430, 41.

[000314] SEQ ID NO:14 was isolated *in silico* from the pB1 gene area sequence disclosed at Accession No. ABI36441 (SEQ ID NO:15). Replikin sequences (SEQ ID NOS:16-17) were identified in the amino-terminus. Replikin sequences (SEQ ID NOS:18-32) were identified in the mid-molecule. No Replikin sequences were identified in the carboxy-terminus. Sixteen Replikin sequences in 90 amino acid residues gave a Replikin Count of 17.8.

[000315] SEQ ID NO:14 was also isolated *in silico* from Accession No. ABI36430 (SEQ ID NO:33). Replikin sequences (SEQ ID NOS:34-35) were identified in the amino-terminus. Replikin sequences (SEQ ID NOS:36-49) were identified in the mid-molecule. No Replikin sequences were identified in the carboxy-terminus.

[000316] SEQ ID NO:14 was also isolated *in silico* from Accession No. ABL07027 (SEQ ID NO:50). Replikin sequences (SEQ ID NOS:51-52) were identified in the amino-terminus. Replikin sequences (SEQ ID NOS:53-68) were identified in the mid-molecule. Replikin sequences (SEQ ID NO:69-71) were identified in the carboxy-terminus.

[000317] SEQ ID NO:2 was identified in the following accession numbers in the following years at the following amino acid residue positions: (1997) Q9WLS3, 184, O89749, 184, AAK49358, 184, AAF74316, 184, AAK49362, 184, AAK49357, 184, AAK49356, 184, CAB95863, 184; (2003) BAE07199, 184; and (2004) ABL97546, 184, ABE97545, 184, ABE97544, 184, ABE97543, 184, ABE97542, 184, ABE97540, 184, ABE97564, 179, ABC72648, 184, ABK34974, 184.

[000318] SEQ ID NO:3 was identified in the following accession numbers in the following years at the following amino acid residue positions: (1997) Q9WLS3, 184, O89749, 184, AAK49358, 184, AAF74316, 184, AAF74315, 184, AAF74314, 184, AAK49362, 184, AAK493761, 184, AAK49359, 184, AAK49357, 184, AAK49356, 184, CAB95863, 184; (1998) AAK49363, 184; (2003) BAE07199, 184; (2004) ABE97546, 184, ABE97545, 184, AGE97544, 184, ABE97543, 184, ABE97542, 184, ABE97541, 184, ABE97540, 184, ABE97539, 184, ABE97538, 184, ABE97537, 184, ABE97536, 184, AAV35116, 184, AAV32644, 184; (2005) ABG78564, 184, ABC72648, 184; and (2006) ABK34974, 184.

[000319] SEQ ID NO:7 was identified in the following accession numbers in the following years at the following amino acid residue positions: (2003) BAE07200, 128; (2004) AAW59551, 125, AAW59549, 123, ABE97897, 117, ABE97896, 117, ABE97895, 117, ABE97894, 117, ABE97893, 117, ABE97892, 117, ABE97891, 117, ABE97890, 117, ABE97889, 117, ABE97888, 117, AAV32651m 128, AAV32643, 128; (2005) ABG78563, 103, ABG78562, 103, ABF56657, 121, ABF56656, 121, ABF56655, 121; and (2006) ABL31779, 128, AB31765, 128, ABL31754, 128, ABL31743, 128, ABI49414, 128, ABI49395, 128, ABL07029, 128, ABI36470, 128, ABI36451, 128, ABI36440, 128, ABI36429, 128.

[000320] SEQ ID NO:8 was identified in the following accession numbers in the following years at the following amino acid residue positions: (1997) Q9WLS3, 184, O89749, 184, AAK49360, 168, AAK49356, 168, AAF74316, 168, AAK49362, 168, AAK49359, 168, AAK49357, 168, AAK49356, 168, CAB5863, 168; (2003) BAE07199, 168; (2004) ABE97546, 168, ABE97545, 168, ABE97544, 168, ABE97543, 168, ABE97542, 168, ABE97541, 168, ABE97539, 168, ABE97538, 168, ABE97537, 168, ABE97536, 168, AAV35116, 168, AAV32644, 168; (2005) ABG78564, 163, ABC72648, 168; and (2006) ABK34974, 168.

[000321] The series of conserved Replikin sequences discussed above are preferred embodiments of the invention and are particularly useful as immunogenic compounds and vaccines and the presence of these sequences has particular predictive value for timing, geographic position and lethality of H5N1 outbreaks.

C. Conservation in Replikin Scaffolds in Influenza A strains

[000322] Table 13, below, provides support for the role of Replikin Scaffolds as Replikin Peak Genes in lethal outbreaks of influenza in humans and birds. In Table 13, the history of the Goose Replikin and its homologues are tracked from 1917 to the present outbreak of avian H5N1 virus. Table 13 demonstrates conservation of the “scaffold” homology of the Goose Replikin in virulent strains of influenza.

[000323] Table 13 illustrates the history, by year or smaller time period, of the existence in the protein structure of the Goose Replikin and its homologues in other influenza Replikins. Table 13 further illustrates the history of amino acid substitutions in those homologues and the conservation of certain amino acids of the Replikin structure which are essential to the definition of a Replikin and the function of rapid replication supplied by Replikins.

[000324] Table 13 illustrates a Fixed Replikin Peak Gene Scaffold with ordered non-random substitution in the 90 year conservation of influenza virus Replikin peptides, from a 1917 goose flu and 1918 human pandemic to a 2007 H5N1 ‘Bird Flu’ homologue.

[000325] The Goose Replikin is a 29 amino acid peptide RPG in the hemagglutinin protein of influenza virus beginning with kk and ending with hh (SEQ ID NO: 3672). Replikins may contain overlapping Replikins. This Replikin Scaffold appears in the virus genome only when the Replikin count rises above 3, and disappears again when the clinical outbreak is over and the Replikin count declines to less than 3.

Table 13 –Goose Replikin Scaffolds

Residues identical to Goose Replikin 1917 amino acids in *grey*; amino acid substitutions in *light grey* and *black*

1917 Goose Replikin and Homologues. Peptide Length	Year and Virus Strain in which Each Goose Replikin Homologue Was Found	
[<-----29 Amino Acids----->]	Year	Strain; Host -chicken/human unless other
kkgtsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3672) 1917	H1N1 Influenza Goose Replikin
kkgssypklsksyvnnkgkevlvlwgvhh	(SEQ ID NO: 3673) 1918	H1N1 Human Influenza Pandemic
kkensypklsksyvnnkgkevlvlwgvhh	(SEQ ID NO: 3674) 1930	H1N1
kkqdsypkltnsyvnnkgkevlvlwgvhh	(SEQ ID NO: 3675) 1933	H0N1
kkgtsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3676) 1976	H1N1
kkgtsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3677) 1977	H1N1
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3678) 1979	H1N1 swine Arnsberg Germany
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3679) 1980	H1N1 swine Netherlands
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3680) 1980	H1N1 swine Netherlands
kkqtsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3681) 1980	H1N1
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3682) 1981	H1N1 swine Potsdam Germany
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3683) 1981	H1N1 swine Potsdam Germany
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3684) 1981	H1N1 swine Germany
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3685) 1981	H1N1
kkqtsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3686) 1981	H1N1
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3687) 1984	H1N1 swine Germany
kkqtsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3688) 1985	H1N1
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3689) 1991	Unidentified swine Germany
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3690) 1992	H1N1 swine Germany
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3691) 1992	H1N1 swine England
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3692) 1992	H1N1 swine Germany
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3693) 1996	H1N1 swine Netherlands
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3694) 1996	H1N1 swine Netherlands
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3695) 1996	H1N1 swine Netherlands
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3696) 1996	H1N1 swine Netherlands
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3697) 1996	H1N1 swine Netherlands
kkqssypklsksyvnnkgkevlvlwgvhh	(SEQ ID NO: 3698) 1997	H1N1
kkqssypklsksyvnnkgkevlvlwgvhh	(SEQ ID NO: 3699) 1997	H1N1
kkqssypklsksyvnnkgkevlvlwgvhh	(SEQ ID NO: 3700) 1998	H1N1
-knglypnlsksyannkekevlvlwgvhh	(SEQ ID NO: 3701) 1999	H1N1
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3702) 1999	H1N1
-knglypnlsksyannkekevlvlwgvhh	(SEQ ID NO: 3703) 2000	H1N1
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3704) 2000	H1N1
-knglypnlsksyannkekevlvlwgvhh	(SEQ ID NO: 3705) 2001	H1N1
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3706) 2001	H1N1 swine Germany
-knglypnlsksyannkekevlvlwgvhh	(SEQ ID NO: 3707) 2002	H1N1
kkqnsypklsksytnnkgkevlvlwgvhh	(SEQ ID NO: 3708) 2002	H1N1
kkqssypklsksyvnnkgkevlvlwgvhh	(SEQ ID NO: 3709) 2003	H1N1
-knglypnlsksyannkekevlvlwgvhh	(SEQ ID NO: 3710) 2003	H1N1

kkqnsypklsksyttnkqkevlvlgvvh	(SEQ ID NO: 3711)	2003	H1N1swineSpain
-knglypnlksyannkekevlvlgvvh	(SEQ ID NO: 3712)	2003	H1N1 New York
kkqnsypklsksyttnkqkevlvlgvvh	(SEQ ID NO: 3713)	2003	H1N1swineSpain
-knglypnlksyannkekevlvlgvvh	(SEQ ID NO: 3714)	2003	H1N1 New York
-knglypnlksyannkekevlvlgvvh	(SEQ ID NO: 3715)	2004	H1N1
kkqnsypklsksyttnkqkevlvlgvvh	(SEQ ID NO: 3716)	2004	H1N1 Spain
-knglypnlksyannkekevlvlgvvh	(SEQ ID NO: 3717)	2004	H1N1 Ontario
-knglypnlksyannkekevlvlgvvh	(SEQ ID NO: 3718)	2005	H1N1 New York
-knglypnlksyannkekevlvlgvvh	(SEQ ID NO: 3719)	2006	H1N1 Japan
-knglypnlksyannkekevlvlgvvh	(SEQ ID NO: 3720)	2006	H1N1 Arizona
-knglypnlksyannkekevlvlgvvh	(SEQ ID NO: 3721)	2006	H1N1 Kentucky
-knglypnlksyannkekevlvlgvvh	(SEQ ID NO: 3722)	2007	H1N1 Thailand
kkqnsypklsksyinnkekevlvlgvvh	(SEQ ID NO: 3723)	1999	H1N2 Influenza
kkqnsypklsksyinnkkkevlvlgvvh	(SEQ ID NO: 3724)	2000	H1N2 Minnesota
kkqnsypklsksyinnkqkvlvlgvvh	(SEQ ID NO: 3725)	2001	H1N2
kkqtsypklsksyttnkqkevlvlgvvh	(SEQ ID NO: 3726)	2001	H1N2
-knglypnlksyannkekevlvlgvvh	(SEQ ID NO: 3727)	2002	H1N2
-knglypnlksyannkekevlvlgvvh	(SEQ ID NO: 3728)	2002	H1N2
kkqnyypvakrsynttsgeqmlilvlgvvh	(SEQ ID NO: 3729)	1957H2N2	Human Influenza Pandemic
kkqnyypvakrsynttsgeqmlilvlgvvh	(SEQ ID NO: 3730)	1957H2N2	Human Influenza Pandemic
kkensypklrksiiinkkevlvlgvvh	(SEQ ID NO: 3731)	1968H3N2	Human Influenza Pandemic
-----ksykntrkdpalilvlgvvh	(SEQ ID NO: 3732)	1979-2003	H7N7 Influenza
kknnayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3733)	2002	H5N2 Influenza
kknnayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3734)	1959	H5N1 Influenza (Scotland)
kknnayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3735)	1975	H5N1 (Wisconsin)
kknnayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3736)	1981	H5N1 (Minnesota)
kkqnyypvakrsynttsgeqmlilvlgvvh	(SEQ ID NO: 3737)	1983	H5N1 (Pennsylvania)
kknttyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3738)	1988	H5N1 (Scotland)
kknsayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3739)	1996	H5N1 (China)
kknsayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3740)	1997	H5N1 (China)
kknsayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3741)	1998	H5N1 (China)
kknsayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3742)	1999	H5N1 (China)
kknsayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3743)	2000	H5N1 (China)
kknsayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3744)	2001	H5N1 (China)
kknnayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3745)	2001	H5N1 (China)
kknstyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3746)	2001	H5N1 (China)
kknsayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3747)	2002	H5N1 (China)
kknstyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3748)	2002	H5N1 (Thailand)
kknstyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3749)	2002	H5N1 (Vietnam)
kknstyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3750)	2003	H5N1 (Vietnam)
kknstyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3751)	2003	H5N1 (Thailand)
kknstyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3752)	2003	H5N1 (Sindong, China)
kknnayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3753)	2003	H5N1 (China)
kknstyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3754)	2004	H5N1 (Vietnam, highly pathogenic)
kknsayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3755)	2004	H5N1 (Vietnam, "highly patho", gull)
kknstyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3756)	2004	H5N1 Vietnam highly pathogenic
kknstyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3757)	2004	H5N1 (Thailand, highly pathogenic)
kknstyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3758)	2004	H5N1 (Thailand, highly pathogenic)
kknsayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3759)	2004	H5N1 (China, highly pathogenic)
kknsayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3760)	2004	H5N1 (China, "highly pathogenic", goose)
kknsayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3761)	2004	H5N1 Japan
kknnayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3762)	2005	H5N1 Turkey
kknstyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3763)	2005	H5N1 Thailand
kkdnayptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3764)	2005	H5N1 Russia
kknttyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3765)	2006	H5N1 China (Guiyang)
kknttyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3766)	2006	H5N1 China (Anhui) (extra R)
kknstyptikrtyntnvedlilvlgvvh	(SEQ ID NO: 3767)	2006	H5N1 Indonesia Human (extra R)

Qenrtypatikrsynntnqedllvlwgihh	(SEQ ID NO: 3768) 2006	H5N1IndonesiaHuman (extra R)
Qanntypatikrsynntnqedllvlwgihh	(SEQ ID NO: 3769) 2006	H5N1IndonesiaHuman
Qenstypatikesynntnqedllvlwgihh	(SEQ ID NO: 3770) 2006	H5N1IndonesiaHuman
Qdnayypatikrsynntnqedllvlwgihh	(SEQ ID NO: 3771) 2006	H5N1 Nigeria
Qendayypatikrsynntnqedllvlwgihh	(SEQ ID NO: 3772) 2006	H5N1 Italy
Qdnayypatikrsynntnqedllvlwgihh	(SEQ ID NO: 3773) 2006	H5N1 Iraq
Qendtyypatikrsynntnqedllvlwgihh	(SEQ ID NO: 3774) 2006	H5N1 Iran

D. Replikin Scaffold in 2007 Isolate of H1N1

[000326] A Replikin Scaffold hemagglutinin Replikin Peak Gene has now been identified in one human case of H1N1 isolated in 2007 in Thailand. This evidence suggests H1N1 is making a comeback. The H1N1 Replikin Scaffold that has been identified is knglypnlsksyannkekevvlwgvhh (SEQ ID NO:2011), which is associated with a whole hemagglutinin Replikin Count of 8.1, and Replikin Count in the RPG of 28. The Replikin Count in the RPG of the 2007 Thailand isolate is higher than the Replikin Count in the RPG of an H1N1 isolate from the 1918 pandemic, Accession No: IRUZL, which has a Replikin Count in its RPG of 19. Example 5 provides the inventors analysis of the 2007 Thailand isolate.

E. Homologous Replikin Scaffold Sequences in Influenza, WSSV, and TSV

[000327] The inventors have further established a relationship between virulent influenza virus and shrimp viruses WSSV and TSV in the Replikin Scaffold portions of the viruses as may be seen in Table 14 below. Although there is extensive substitution, several short Replikins of the Shrimp white spot syndrome virus demonstrates significant homologies to the influenza virus Replikin sequences, especially with regard to length and key lysine (k) and histidine (h) residues. Similar, but less extensive, homologies are seen in taura syndrome virus. These homologies suggest that the sequences are derived from a shared reservoir and/or that similar mechanisms of Replikin production are used in both virus groups.

Table 14 –Shrimp White Spot and Taura Syndrome Scaffolding

Qcgtstypklsksytnnkqkqevl vlwgvhh	(SEQ ID NO: 3775) 1917	H1N_ Influenza goose peptide
Qcgnstypklsksytnnkqkqevl v wgvhh	(SEQ ID NO: 3776) 2002	H1N1 Swine Influenza
Qenvksakqlphlkvlkldvrgakqlph	(SEQ ID NO: 3777) 2000	Shrimp White Spot SyndromeVirus
-lvhldvlgvqqlhlhkvrdvrgakqlh	(SEQ ID NO: 3778) 2000	Shrimp White Spot SyndromeVirus
Qcensypklsksytnnkqkqevl vlwgihh	(SEQ ID NO: 3779) 1968	H3N2 Human Influenza Pandemic
-----ksykntrkdpali wgihh	(SEQ ID NO: 3780) 1979-2003	H7N7 Influenza
Qcgpnyypvakrsynntsgqemli wgvhh	(SEQ ID NO: 3781) 1957	H2N2 Human Influenza Pandemic
Qcgpnyypvakrsynntsgqemli wgihh	(SEQ ID NO: 3782) 1957	H2N2 Human Influenza Pandemic
Qannayypatikrtynntnvedllvlwgihh	(SEQ ID NO: 3783) 2002	H5N2 Influenza
Qannayypatikrsynntnqedllvlwgihh	(SEQ ID NO: 3784) 1959	H5N1 Influenza
Qlvqanktrvfaasnqqlalalrryylsflhdh	(SEQ ID NO: 3785) 2000	Taura Syndrome Virus
Qacrnagykeaclheldcksfl laqqragah	(SEQ ID NO: 3786) 2005	Taura Syndrome Virus

Residues identical to original 1917 Goose Replikin residues are shown in **medium grey**. Amino acid substitutions in light grey and **bold**.

[000328] TSV is less virulent than WSSV and the structure of the TSV Replikin Scaffold is less closely related to influenza virus than are the structures of WSSV Replikin Scaffolds. In year 2000, TSV had a Replikin concentration of 2.7. Between 2001 and 2004, TSV had a lower mean Replikin concentration, as low as 0.7, and its Replikin Scaffold disappeared. In 2005 the Replikin Scaffold reappeared, with an increase in lysines and histidines, and a commensurate increase in Replikin concentration to 1.8, followed by an increase in TSV outbreaks in 2006-2007. *See* Table 19.

F. Replikin Peak Genes Provide Increased Predictive and Therapeutic Capacity

[000329] Since the identification of the Replikin structure, correlation between increased concentrations of Replikin sequences and increased replication and virulence has been observed in a range of viruses and organisms. These observations are made more accurate by the present isolation *in silico* of Replikin Peak Genes. While increased concentration of Replikin sequences in the genome of a virus offers both advance warning and new targets for developing effective methods of predicting and treating viral outbreaks, identification of an increase in concentration of Replikin sequences in a Replikin Peak Gene of a genome or protein heightens the predictive capacity of the change in Replikin concentration and the efficacy of new targets.

[000330] For example, more precise predictions of increased virulence are now available through identification of a Replikin Peak Gene in, among other viruses, the H5N1 strain of influenza (Figures 1-6), west nile virus (Figure 12) and foot and mouth disease virus (Figure 11). In these and other viruses, increased concentration of Replikin sequences in the whole genome, in a protein of the genome, in a Replikin Peak Gene of the genome, or in a protein containing a RPG, offer both advance warning and new targets for developing effective methods of predicting and treating viral outbreaks.

[000331] By monitoring changes in concentrations of Replikin sequences in viral genomes generally, emerging viral diseases can be identified in virus reservoirs and vectors in advance of their appearance in animal or human hosts. Identification of the emerging viruses and the Replikin sequences within the virus genome allows for appropriate, advance control efforts, including isolation and quarantine, and provide sufficient time for the synthesis and testing of vaccines specific to the sequences of the emerging virus.

[000332] As discussed above, the inventors have identified the pB1 gene area of the H3N8 strains of influenza virus (SEQ ID NO 545) as the region of the genome of the virus

having the highest concentration of Replikin sequences. A Replikin Peak Gene has also been identified in H5N1 influenza virus and has been correlated with epidemics, increased virulence, morbidity and human mortality. (Figures 1-6) Likewise, a Replikin Peak Gene has been identified in the VP1 protein of foot and mouth disease virus and has been correlated with outbreaks of the virus. Figure 11. A second Replikin Peak Gene (or Replikin Peak Gene Area) has additionally been identified in a fragment of the VP1 protein of foot and mouth disease virus and two particular Replikin sequences within the Replikin Peak Gene Area of the virus have been correlated with virulence of foot and mouth disease virus (*e.g.*, SEQ ID NOS: 124 and 130). A Replikin Peak Gene has also been identified in west Nile virus (*e.g.*, SEQ ID NO: 258). Figure 12. Replikin Peak Genes have further been identified in the nucleocapsid protein of the porcine reproductive and respiratory syndrome virus and in Porcine Circovirus (*e.g.*, SEQ ID NOS: 341 and 520, respectively). Figures 13 and 19.

[000333] The invention provides Replikin sequences within the identified Replikin Peak Gene gene or gene segment (gene area) for diagnostic, preventive and therapeutic applications. For example, each Replikin sequence identified within an identified RPG is available for diagnostic and therapeutic applications including vaccines and antibody therapies. The entire Replikin Peak Gene sequence or fragments thereof are likewise available for diagnostic, preventive, therapeutic and predictive applications. Further, the presence of the Replikin Peak Gene in an isolate of the virus is indicative of rapid replication. For each identified RPG, continuous, non-interrupted and overlapping Replikin sequences have been identified for predictive and therapeutic applications. The size of a Replikin Peak Gene or Replikin Peak Gene Area, both in terms of the number of amino acids and the Replikin Count, will depend upon the size of the sequence of the entire genome, protein or fragment thereof that has been isolated and reported.

[000334] The invention further provides Replikin sequences within the identified Replikin Peak Gene or Replikin Peak Gene Area that are conserved in the genome over time and, as such, are available as relatively invariant targets for diagnosis and manipulation of rapid replication and virulence in EIV.

[000335] Point mutations within an RPG provide excellent predictive capacity when the point mutation is correlated with high virulence and provide an excellent target for attack of the virus through antibody therapies, vaccines and other treatments, as well as excellent predictive capacity when such point mutations are identified in emerging strains of the virus.

[000336] A further aspect of the invention provides utilizing software that searches for Replikin Peak Genes and enables the discovery of the point or points in the genome that have the highest concentration of Replikins, the years in which they have occurred, the strain or strains in which they occur, the host or hosts in which they occur, the geographic locations in which they occur, their increase or decrease in the above years, strains, hosts and geographic locations and point or small mutations that are correlatable with virulence.

[000337] The *in silico* detection of the Replikin Peak Gene by software methodology now permits both host and geographic localization of upregulated Replikin gene activity both in viruses and in other organisms. As seen in this study, the annual RPG Replikin analysis, by its correlation with a function such as epidemic outbreak or increase in lethality, can for the first time actually provide evidence for the correlation with the function.

[000338] The Replikin count in the whole genome or RPG make possible the prediction in advance of epidemic outbreaks of high mortality infections, such as those caused by influenza viruses, as seen for H5N1 in Figures 1-6. Such detection and localization permits advance focused public health preparations for better protection of the host, whether animal or human, and give time for the production and testing of new vaccines. The high Replikin Count of the RPG has now been shown to be associated consistently with a high percent lethality in the host, whether the host is a plant, fish, shrimp, or vertebrate, including human cases of H5N1 bird flu. The increase in count was frequently detected one year or more before the outbreak was clinically apparent (Figures 2, 3, 10, 11, 19, *etc.*). Vaccines may now be produced that directly target rapid replication as represented structurally by the Replikins in the whole genome and concentrated in a Replikin Peak Gene, rather than, as now, being targeted at virus epitopes whose function is unknown.

[000339] It may be concluded that Replikins represent a specific class of peptides that are widely distributed, conserved, quantitative markers of lethality. While not wishing to be bound by theory, evidence from the apparent transfer of conserved Replikin structures between strains suggest they may be mobile agents of lethality, transferring horizontally between carrier viruses to reach multicellular hosts, where they may replicate rapidly with lethal consequences. As newly recognized targets for prevention and therapy, Replikins offer a platform from which specifically to control rapid replication and lethality of organisms and cells, without necessarily destroying them.

G. Conserved Replikins in PCV for Diagnostics and Therapies

[000340] In review of the publicly available sequences for Porcine Circovirus, the applicants have identified three Replikin sequences from Accession No. ABQ10608 that are conserved across many isolates from 1997 or 1998 through 2007: kngrsgpqqhk (SEQ ID NO:345); hlqgfanfkkqtfnk (SEQ ID NO:346) and kkqtfnkvwylgarch (SEQ ID NO:347). Because these sequences are conserved, they have predictive value and provide novel and preferred targets for diagnostic and therapeutic applications such as, for example, vaccines. Furthermore, two of these sequences, hlqgfanfkkqtfnk (SEQ ID NO:346) and kkqtfnkvwylgarch (SEQ ID NO:347) are contained within the identified RPG of Accession No. ABQ10608. These sequences, therefore, are of preferred value in predicting virulent strains when such strains contain the sequences. Also, the sequences provide preferred targets for diagnostic and therapeutic applications such as, for example, vaccines.

[000341] Table 15 provides the accession numbers of isolates of PCV between 1997 and 2007 containing the conserved sequence kngrsgpqqhk (SEQ ID NO:345) and the amino acid position within the PCV protein sequence wherein the conserved Replikin sequence begins.

Table 15 – Conserved PCV Sequence

1997	AAC59462 position 5 .
1998	AAC35330 position 5 , AAC35320 position 5 , AAC35309 position 5 , AAC35298 position 5 , CAA11157 position 5 , AAC61860 position 5 , AAC61741 position 5 , AAC61739 position 5 , AAC61737 position 5 , AAD03086 position 5 , AAD03071 position 5 , AAD03061 position 5 , NP_048061 position 5 , AAD11928 position 5 .
1999	BAA88133 position 5 , AAD50432 position 5 , AAD38398 position 5 , AAG41226 position 5 , AAD37776 position 5 , AAD45580 position 5 , AAF35304 position 5 , AAF35302 position 5 , AAF35300 position 5 , AAF35298 position 5 , AAF35296 position 5 , AAF35294 position 5 , AAF35292 position 5 , AAD12308 position 5 .
2000	CAC41085 position 5 , CAC41084 position 5 , CAC41083 position 5 , AAL09364 position 5 , AAL09363 position 5 , AAF87238 position 5 , AAF87236 position 5 , AAF87234 position 5 , AAF87232 position 5 , AAF87230 position 5 , AAF87228 position 5 .
2001	AAK60462 position 5 , AAL58397 position 5 , BAB69441 position 5 , BAB69437 position 5 , BAB69432 position 5 , AAK56300 position 5 , AAK56298 position 5 , AAK56296 position 5 , AAL01075 position 5 .
2002	AAM61272 position 5 , AAM61262 position 5 , AAM61268 position 5 , AAM61266 position 5 , AAM61270 position 5 , AAM61264 position 5 , AAO39760 position 5 , AAM21847 position 5 , AAM21846 position 5 , AAM21845 position 5 , AAM21844 position 5 , AAO24128 position 5 , AAO24124 position 5 , AAO24122 position 5 , AAO23147 position 5 , AAO23145 position 5 , AAN81597 position 5 , AAN06826 position 5 , AAN62769 position 5 , AAN62767 position 5 , AAN62765 position 5 , AAL69968 position 5 , AAM76057 position 5 , Q8BB16 position 5 .
2003	AAP51128 position 5 , AAS65993 position 5 , AAS65991 position 5 , AAS65989 position 5 , AAS65987 position 5 , AAS65985 position 5 , AAS65983 position 5 , AAS65981 position 5 , AAS65979 position 5 , AAS65977 position 5 , AAS65975 position 5 , AAP83635 position 5 , AAP83633 position 5 , AAP83631 position 5 , AAP83629 position 5 , AAP83627 position 5 , AAP83625 position 5 , AAP83623 position 5 , AAP83621 position 5 , AAP83619 position 5 , AAP83617 position 5 , AAP83615 position 5 , AAP83613 position 5 , AAP83611 position 5 , AAP83609 position 5 , AAP83607 position 5 , AAP83605 position 5 , AAP83603 position 5 , AAP83601 position 5 , AAP83599 position 5 , AAP83597 position 5 , AAP83595 position 5 ,

	<p> AAP83593 position 5 , AAP83591 position 5 , AAR03722 position 5 , AAR03720 position 5 , AAR03718 position 5 , AAR03716 position 5 , AAQ94098 position 5 , AAQ94096 position 5 , AAQ94094 position 5 , AAQ94092 position 5 , AAQ94090 position 5 , AAQ94088 position 5 , AAP44188 position 5 , AAP44182 position 5 , AAR97517 position 5 , AAQ96327 position 5 , AAQ23155 position 5 , AAP42468 position 5 , AAP42466 position 5 , AAO61136 position 5 , NP_937956 position 5 , AAR03714 position 5 . </p>
2004	<p> AAW78475 position 5 , AAW78473 position 5 , AAW78471 position 5 , AAW78469 position 5 , AAW78467 position 5 , AAW78465 position 5 , AAW78463 position 5 , AAV34139 position 5 , AAU87519 position 5 , AAU87515 position 5 , AAU87511 position 5 , AAU87509 position 5 , AAU34001 position 5 , AAT97650 position 5 , AAT97648 position 5 , AAT97646 position 5 , AAT36358 position 5 , AAX49397 position 5 , AAU01966 position 5 , AAT72901 position 5 , AAT58234 position 5 , AAS45844 position 5 , AAS45843 position 5 , CAJ31064 position 5 , AAU13780 position 5 , AAX52911 position 5 , AAU87505 position 5 , AAT39479 position 5 , AAT39460 position 5 , AAT37493 position 5 , AAS66198 position 5 , AAS66196 position 5 , AAS66194 position 5 , AAS66192 position 5 , AAS90297 position 5 , AAS89260 position 5 , CAF25171 position 5 . </p>
2005	<p> ABJ98317 position 5 , AAZ20800 position 5 , AAZ20796 position 5 , AAZ20794 position 5 , AAW79865 position 5 , ABC26025 position 5 , ABA40480 position 5 , AAZ78351 position 5 , AAY40292 position 5 , ABB29423 position 5 , ABB29419 position 5 , ABB29417 position 5 , ABB29415 position 5 , ABB29413 position 5 , ABB29411 position 5 , ABB29409 position 5 , ABB29407 position 5 , ABB29405 position 5 , ABB29403 position 5 , ABB29401 position 5 , ABB29399 position 5 , ABA60807 position 5 , ABA60805 position 5 , ABA40399 position 5 , ABA40397 position 5 , AAX10150 position 5 , AAX62053 position 5 , AAX62051 position 5 , AAX62049 position 5 , AAX62047 position 5 , AAX62045 position 5 , AAX62043 position 5 , AAX62041 position 5 , ABC75103 position 5 , ABB20934 position 5 , ABA26910 position 5 , ABA26908 position 5 , AAY34249 position 5 . </p>
2006	<p> ABI29887 position 5 , ABG21279 position 5 , ABG21277 position 5 , ABG21275 position 5 , ABG21273 position 5 , ABG21271 position 5 , ABG21269 position 5 , ABG21267 position 5 , ABJ98319 position 5 , ABI93799 position 5 , ABI93797 position 5 , ABD59347 position 5 , ABD42928 position 5 , ABM88864 position 5 , ABM88862 position 5 , ABM88860 position 5 , ABI17537 position 5 , ABI17535 position 5 , ABI17533 position 5 , ABI17531 position 5 , ABI17529 position 5 , ABI17527 position 5 , ABI17525 position 5 , ABI17523 position 5 , ABG37023 position 5 , ABF71465 position 5 . </p>
2007	<p> ABQ10608 position 5 , ABQ10606 position 5 , ABQ10604 position 5 , ABQ10603 position 5 , ABP68669 position 5 , ABP68665 position 5 , ABP68661 position 5 , ABP68657 position 5 , ABP68655 position 5 , ABP68651 position 5 , ABP68647 position 5 , ABP68645 position 5 , ABP68643 position 5 , ABP68641 position 5 , ABP68639 position 5 , ABP68635 position 5 , ABP68633 position 5 , ABP68631 position 5 , ABP68629 position 5 , ABP68627 position 5 , ABP68625 position 5 , ABP68623 position 5 , ABP68621 position 5 , ABP68619 position 5 , ABP68617 position 5 , ABP68615 position 5 , ABO38130 position 5 , ABM97550 position 5 , ABQ63072 position 5 , ABQ63070 position 5 , ABQ63068 position 5 , ABQ63066 position 5 , ABQ63064 position 5 , ABQ63062 position 5 , ABQ51920 position 5 , ABQ51918 position 5 , ABR14585 position 5 , ABP49176 position 5 , ABP48091 position 5 , ABP48089 position 5 , ABP48087 position 5 , ABP48085 position 5 , ABP48083 position 5 , ABP48081 position 5 , ABO09999 position 5 , ABO09997 position 5 , ABO09995 position 5 , ABO09993 position 5 , ABO09991 position 5 , ABO09989 position 5 , ABO09987 position 5 , ABP23690 position 5 . </p>

[000342] Table 16 provides the accession numbers of PCV isolates between 1997 and 2007 containing the conserved sequence hlqgfanfvkkqtfnk (SEQ ID NO:346) and the amino acid position within the PCV protein sequence wherein the conserved Replikin sequence begins.

Table 16 – Conserved PCV Sequence

1997	AAC59462 position 57 .
1998	AAC35330 position 57 , AAC35320 position 57 , AAC35309 position 57 , AAC35298 position 57 , CAA11157 position 57 , AAC61860 position 57 , AAC61741 position 57 , AAC61739 position 57 , AAC61737 position 57 , AAD03086 position 57 , AAD03071 position 57 , AAD03061 position 57 , NP_048061 position 57 , AAD11928 position 57 .
1999	BAA88133 position 57 , AAD50432 position 57 , AAD38398 position 57 , AAG41226 position 57 , AAD37776 position 57 , AAD45580 position 57 , AAF35304 position 57 , AAF35302 position 57 , AAF35300 position 57 , AAF35298 position 57 , AAF35296 position 57 , AAF35294 position 57 , AAF35292 position 57 , AAD12308 position 57 .
2000	CAC41085 position 57 , CAC41084 position 57 , AAL09364 position 57 , AAL09363 position 57 , AAF87238 position 57 , AAF87236 position 57 , AAF87234 position 57 , AAF87232 position 57 , AAF87230 position 57 , AAF87228 position 57 .
2001	AAK60462 position 57 , AAL58397 position 57 , BAB69441 position 57 , BAB69437 position 57 , BAB69432 position 57 , AAK56300 position 57 , AAK56298 position 57 , AAK56296 position 57 , AAL01075 position 57 .
2002	AAM61272 position 57 , AAM61262 position 57 , AAM61268 position 57 , AAM61266 position 57 , AAM61270 position 57 , AAM61264 position 57 , AAO39760 position 57 , AAM21845 position 57 , AAM21844 position 57 , AAO24128 position 57 , AAO24126 position 57 , AAO24124 position 57 , AAO24122 position 57 , AAO23147 position 57 , AAO23145 position 57 , AAN81597 position 57 , AAN06826 position 57 , AAN62769 position 57 , AAN62767 position 57 , AAN62765 position 57 , AAN16398 position 57 , AAM83186 position 57 , AAM76057 position 57 , Q8BB16 position 57 , AAO95302 position 57 .
2003	AAP51128 position 57 , AAS65993 position 57 , AAS65991 position 57 , AAS65989 position 57 , AAS65987 position 57 , AAS65985 position 57 , AAS65983 position 57 , AAS65981 position 57 , AAS65979 position 57 , AAS65977 position 57 , AAS65975 position 57 , AAP83635 position 57 , AAP83633 position 57 , AAP83631 position 57 , AAP83629 position 57 , AAP83627 position 57 , AAP83625 position 57 , AAP83623 position 57 , AAP83621 position 57 , AAP83619 position 57 , AAP83617 position 57 , AAP83615 position 57 , AAP83613 position 57 , AAP83611 position 57 , AAP83609 position 57 , AAP83607 position 57 , AAP83605 position 57 , AAP83603 position 57 , AAP83601 position 57 , AAP83599 position 57 , AAP83597 position 57 , AAP83595 position 57 , AAP83593 position 57 , AAP83591 position 57 , AAR03722 position 57 , AAR03720 position 57 , AAR03718 position 57 , AAR03716 position 57 , AAQ94098 position 57 , AAQ94096 position 57 , AAQ94094 position 57 , AAQ94092 position 57 , AAQ94090 position 57 , AAQ94088 position 57 , AAP44188 position 57 , AAP44185 position 57 , AAP44182 position 57 , AAR97517 position 57 , AAQ96327 position 57 , AAQ23155 position 57 , AAP42468 position 57 , AAP42466 position 57 , AAP42464 position 57 , AAO61136 position 57 , NP_937956 position 57 , AAR03714 position 57 .
2004	AAW78475 position 57 , AAW78473 position 57 , AAW78471 position 57 , AAW78469 position 57 , AAW78467 position 57 , AAW78465 position 57 , AAW78463 position 57 , AAV34139 position 57 , AAU87519 position 57 , AAU87517 position 57 , AAU87515 position 57 , AAU87513 position 57 , AAU87511 position 57 , AAU87509 position 57 , AAU87507 position 57 , AAU34001 position 57 , AAU01913 position 57 , AAT97650 position 57 , AAT97648 position 57 , AAT97646 position 57 , AAT97644 position 57 , AAT36358 position 57 , AAX49397 position 57 , AAU01966 position 57 , AAT79579 position 57 , AAT72901 position 57 , AAS45844 position 57 , AAS45843 position 57 , CAJ31064 position 57 , AAU13780 position 57 , AAX52911 position 57 , AAU87505 position 57 , AAT39479 position 57 , AAT39460 position 57 , AAT37493 position 57 , AAS66198 position 57 , AAS66196 position 57 , AAS66194 position 57 , AAS66192 position 57 , AAS66190 position 57 , AAS90297 position 57 , CAF25171 position 57 .

2005	ABJ98317 position 57 , ABA29241 position 57 , AAZ20802 position 57 , AAZ20800 position 57 , AAZ20798 position 57 , AAZ20796 position 57 , AAZ20794 position 57 , AAW79865 position 57 , ABC26025 position 57 , ABA40480 position 57 , AAZ78351 position 57 , AAX21515 position 57 , ABB29423 position 57 , ABB29421 position 57 , ABB29419 position 57 , ABB29417 position 57 , ABB29415 position 57 , ABB29413 position 57 , ABB29411 position 57 , ABB29409 position 57 , ABB29407 position 57 , ABB29405 position 57 , ABB29403 position 57 , ABB29401 position 57 , ABB29399 position 57 , ABA60807 position 57 , ABA60805 position 57 , ABA60803 position 57 , ABA40399 position 57 , ABA40397 position 57 , AAZ66792 position 57 , AAX10150 position 57 , AAX62053 position 57 , AAX62051 position 57 , AAX62049 position 57 , AAX62047 position 57 , AAX62045 position 57 , AAX62043 position 57 , AAX62041 position 57 , ABC75103 position 57 , ABB20934 position 57 , ABA26910 position 57 , ABA26908 position 57 , AAY34249 position 57 .
2006	ABI29887 position 57 , ABG21279 position 57 , ABG21277 position 57 , ABG21275 position 57 , ABG21273 position 57 , ABG21271 position 57 , ABG21269 position 57 , ABG21267 position 57 , ABJ98319 position 57 , ABI93799 position 57 , ABI93797 position 57 , ABD59347 position 57 , ABD42928 position 57 , ABM88864 position 57 , ABM88862 position 57 , ABM88860 position 57 , ABI17537 position 57 , ABI17535 position 57 , ABI17533 position 57 , ABI17531 position 57 , ABI17529 position 57 , ABI17527 position 57 , ABI17525 position 57 , ABI17523 position 57 , ABG37023 position 57 , ABF71465 position 57 .
2007	ABQ10608 position 57 , ABQ10606 position 57 , ABQ10604 position 57 , ABQ10603 position 57 , ABP68669 position 57 , ABP68667 position 57 , ABP68665 position 57 , ABP68663 position 57 , ABP68661 position 57 , ABP68659 position 57 , ABP68657 position 57 , ABP68655 position 57 , ABP68653 position 57 , ABP68651 position 57 , ABP68649 position 57 , ABP68645 position 57 , ABP68643 position 57 , ABP68641 position 57 , ABP68639 position 57 , ABP68637 position 57 , ABP68635 position 57 , ABP68633 position 57 , ABP68629 position 57 , ABP68619 position 57 , ABP68617 position 57 , ABP68615 position 57 , ABO38130 position 57 , ABM97550 position 57 , ABQ63072 position 57 , ABQ63070 position 57 , ABQ63068 position 57 , ABQ63066 position 57 , ABQ63064 position 57 , ABQ63062 position 57 , ABQ51920 position 57 , ABQ51918 position 57 , ABR14585 position 57 , ABP49176 position 57 , ABP48091 position 57 , ABP48089 position 57 , ABP48087 position 57 , ABP48083 position 57 , ABP48081 position 57 , ABO9997 position 57 , ABO9995 position 57 , ABO9993 position 57 , ABO9991 position 57 , ABO9989 position 57 , ABO9987 position 57 , ABP23690 position 57 .

[000343] Table 17 provides the accession numbers of PCV isolates between 1998 and 2007 containing the conserved sequence kkqtfnkvwylgarch (SEQ ID NO:347) and the amino acid position within the PCV protein sequence wherein the conserved Replikin sequence begins.

TABLE 17 – Conserved PCV Sequence

1998	AAC35330 position 66 , AAC35320 position 66 , AAC35309 position 66 , AAC35298 position 66 , CAA11157 position 66 , AAC61860 position 66 , AAC61739 position 66 , AAC61737 position 66 , AAD03086 position 66 , AAD03071 position 66 , AAD03061 position 66 , NP_048061 position 66 , AAD11928 position 66 .
1999	AAG41226 position 66 , AAD37776 position 66 , AAD45580 position 66 , AAF35304 position 66 , AAF35302 position 66 , AAF35300 position 66 , AAF35298 position 66 , AAF35296 position 66 , AAF35294 position 66 , AAF35292 position 66 , AAD12308 position 66 .
2000	CAC41085 position 66 , CAC41084 position 66 , AAF87238 position 66 , AAF87236 position 66 , AAF87234 position 66 , AAF87232 position 66 , AAF87230 position 66 , AAF87228 position 66 .
2001	AAL58397 position 66 , BAB69441 position 66 , BAB69437 position 66 , BAB69432 position 66 , AAK56300 position 66 , AAK56298 position 66 , AAK56296 position 66 ,

	AAL01075 position 66 .
2002	AAM61272 position 66 , AAM61262 position 66 , AAM61268 position 66 , AAM61266 position 66 , AAM61270 position 66 , AAM61264 position 66 , AAO39760 position 66 , AAM21845 position 66 , AAM21844 position 66 , AAO24128 position 66 , AAO24124 position 66 , AAO24122 position 66 , AAN81597 position 66 , AAN06826 position 66 , AAN16398 position 66 , AAM83186 position 66 , AAL69968 position 66 , AAM76057 position 66 , Q8BB16 position 66 .
2003	AAP51128 position 66 , AAS65993 position 66 , AAS65991 position 66 , AAS65989 position 66 , AAS65987 position 66 , AAS65985 position 66 , AAS65983 position 66 , AAS65979 position 66 , AAS65977 position 66 , AAS65975 position 66 , AAP83635 position 66 , AAP83633 position 66 , AAP83631 position 66 , AAP83629 position 66 , AAP83627 position 66 , AAP83625 position 66 , AAP83623 position 66 , AAP83621 position 66 , AAP83619 position 66 , AAP83617 position 66 , AAP83615 position 66 , AAP83613 position 66 , AAP83611 position 66 , AAP83609 position 66 , AAP83607 position 66 , AAP83605 position 66 , AAP83603 position 66 , AAP83601 position 66 , AAP83599 position 66 , AAP83597 position 66 , AAP83595 position 66 , AAP83593 position 66 , AAP83591 position 66 , AAR03722 position 66 , AAR03720 position 66 , AAQ94098 position 66 , AAQ94096 position 66 , AAQ94094 position 66 , AAQ94092 position 66 , AAQ94090 position 66 , AAQ94088 position 66 , AAP44188 position 66 , AAP44182 position 66 , AAQ96327 position 66 , AAQ23155 position 66 , AAP42466 position 66 , AAP42464 position 66 , AAO61136 position 66 .
2004	AAW78479 position 63 , AAW78475 position 66 , AAW78471 position 66 , AAW78469 position 66 , AAW78465 position 66 , AAV34139 position 66 , AAU87519 position 66 , AAU87511 position 66 , AAU87509 position 66 , AAU34001 position 66 , AAU01913 position 66 , AAT97648 position 66 , AAT97644 position 66 , AAT36358 position 66 , AAX49397 position 66 , AAT72901 position 66 , AAS45844 position 66 , AAS45843 position 66 , CAJ31064 position 66 , AAU87505 position 66 , AAT39479 position 66 , AAT39460 position 66 , AAT37493 position 66 , AAS90297 position 66 , AAS89260 position 66 , CAF25171 position 66 .
2005	ABJ98317 position 66 , AAZ20802 position 66 , AAZ20800 position 66 , AAZ20798 position 66 , AAZ20796 position 66 , AAZ20794 position 66 , AAW79865 position 66 , AAY40292 position 66 , ABB29423 position 66 , ABB29421 position 66 , ABB29419 position 66 , ABB29417 position 66 , ABB29415 position 66 , ABB29413 position 66 , ABB29411 position 66 , ABB29409 position 66 , ABB29407 position 66 , ABB29405 position 66 , ABB29403 position 66 , ABB29401 position 66 , ABB29399 position 66 , ABA60807 position 66 , ABA60805 position 66 , ABA40399 position 66 , ABA40397 position 66 , AAZ66792 position 66 , AAX10150 position 66 , AAX62051 position 66 , AAX62049 position 66 , AAX62047 position 66 , AAX62045 position 66 , AAX62041 position 66 , ABB20934 position 66 , ABA26908 position 66 , AAY34249 position 66 .
2006	ABI29887 position 66 , ABG21279 position 66 , ABG21277 position 66 , ABG21275 position 66 , ABG21273 position 66 , ABG21271 position 66 , ABG21269 position 66 , ABG21267 position 66 , ABJ98319 position 66 , ABI93799 position 66 , ABD59347 position 66 , ABD42928 position 66 , ABI17537 position 66 , ABI17535 position 66 , ABI17533 position 66 , ABI17531 position 66 , ABI17529 position 66 , ABI17527 position 66 , ABI17525 position 66 , ABI17523 position 66 , ABG37025 position 63 , ABG37023 position 66 , ABF71465 position 66 .
2007	ABQ10608 position 66 , ABQ10606 position 66 , ABQ10604 position 66 , ABQ10603 position 66 , ABP68669 position 66 , ABP68665 position 66 , ABP68661 position 66 , ABP68653 position 66 , ABP68651 position 66 , ABP68643 position 66 , ABP68641 position 66 , ABP68629 position 66 , ABP68619 position 66 , ABP68617 position 66 , ABP68615 position 66 , ABO38130 position 66 , ABQ63072 position 66 , ABQ63070 position 66 , ABQ63068 position 66 , ABQ63066 position 66 , ABQ63064 position 66 , ABQ63062 position 66 , ABQ51920 position 66 , ABQ51918 position 66 , ABR14585 position 66 , ABP49176 position 66 , ABP48091 position 66 , ABP48089 position 66 , ABP48087 position 66 , ABP48083 position 66 , ABO09999 position 66 , ABO09991 position 66 , ABO09989 position 66 , ABP23690 position 66 .

XII. Relationship between Replikin Peak Gene and Lethality in Tobacco Mosaic Virus and Lung Malignancy

[000344] As established above, the Replikin Peak Gene correlates with activity of viruses such as pandemic influenza, Bird Flu, west Nile virus and Bird Flu H5N1, among many others. It has surprisingly now been discovered that the highest activity to date of the Replikin Peak Gene was found in lung cancer (SEQ ID NO: 1741). Although viruses have been amply confirmed to be associated with the causation of several cancers since the work of Rous in sarcoma at the beginning of the last century, and viruses are the basis of current anti-cancer vaccines, how viruses are related to cancer is still not well understood. The antimalignin antibody in serum (AMAS) test is an FDA-permitted Medicare-approved early detection method for cancer that measures production of antibody against peptides containing a key Replikin sequence, namely, the glioma Replikin peptide, kagvafllhkk (SEQ ID NO: 3658), but how AMAS detects cancer regardless of cell type has not been fully understood. Results from separate studies in the areas of viruses and cancer now have converged with the isolation by the inventors of Replikins in both viruses and cancers that are concentrated in proteins where the concentration of Replikins has been related to rapid replication.

[000345] Higher Replikin Counts in RPGs have now been associated consistently with a higher percent lethality in the host; whether the host is a plant, fish, shrimp, or vertebrate, including human cases of H5N1 bird flu. The increase in count has frequently been detected one year or more before outbreaks have become clinically apparent. In addition to the correlation of high counts with virulence and lethality, structures specific to Replikins have been found by the inventors. For example, a 29-amino acid Replikin scaffold (beginning with SEQ ID NO: 3672) conserved for 90 years, appeared in the genome of successive influenza virus strains and each of the lethal pandemics and lethal H5N1 outbreaks. Additionally, repeating specific Replikin sequence signatures in RPGs of certain pathogens and malignancies have been identified and correlated with lethality. For example, an identical signature (SEQ ID NO:1584) was found to repeat eleven times in the RPG of protozoan *P. falciparum*, 20 times in the RPG of tobacco mosaic virus which included exacerbated cell death in a pepper plant, with exacerbated cell death induced by Tobacco Mosaic virus, and 57 times (by overlapping) within 52 Replikins in the 18 amino acid RPG of non-small cell lung carcinoma.

[000346] While the inventors do not wish to be bound by theory, both of the above two studies support the impression that Replikins are mobile agents of lethality. Pathogenic viruses may just be the carrier of the lethal mobile agents. The highest Replikin Count in a

Replikin Peak Gene that has to date been observed in highly lethal non-small cell lung cancer. The Replikin Count was observed to be 289 Replikin sequences per 100 amino acids. Other cancers, such as breast and ovarian have likewise been observed to have very high Replikin Counts in their Replikin Peak Genes with counts above 40 Replikin sequences per 100 amino acids. An RPG was identified and a Replikin Count of 129 was observed in Accession No. EAW84344 in lymphoblastic leukemia. An RPG was likewise identified and a Replikin Count of 23 was observed in Accession No. EAX09769 in myeloid leukemia.

[000347] Replikins chemically synthesized in the laboratory were found experimentally to be immuno-stimulants, producing strong antibody responses in chickens and rabbits. It appears that the antibodies measured in the AMAS test are against the Replikins' chemistry of rapid replication rather than the histological diagnosis of cancer or the cell type. Thus, for example, histologically proven prostate cancer that is "quiescent" (over 90% of such cancers) has low antibody levels in the AMAS test. But when these cells replicate rapidly, antibody levels measured by the AMAS test increase markedly. AMAS warning frequently precedes detection of the production of Prostate-Specific Antigen (PSA), an antigen that is frequently assayed because of its relationship to prostate cancer. AMAS probably precedes PSA because PSA measures protein fragments, the antigens that must be released by the cancer cells into the blood, while AMAS measures antibody to the peptide changes in the cancer cells, an earlier detectable event.

[000348] Since humans are host to and inhabited by thousands of viruses and bacteria that live symbiotically within the body, unless some event like rapid replication creates disease, no pathogenesis exists. Therefore, it may be important to learn how to control symbiosis between host humans and symbiotic viruses and bacteria without necessarily aiming to destroy the organism outright, especially when destruction proves difficult.

[000349] Peptides isolated from cancer cells grown in tissue culture have been found to contain Replikin sequences. When stimulated by anoxia, the cell number in these tissue cultures increased five-fold per week. Surprisingly, however, Replikin sequence concentration increased ten-fold per week (twice that of cell number), demonstrating a correlation of Replikin count with rapid replication in cancer tissue culture. When the structure of these Replikin-containing peptides was determined, separately synthesized chemically, and administered to rabbits, the peptides produced specific antimalignin antibodies in abundance. The production of antimalignin antibodies in response to the

Replikin-containing peptides provided evidence to close the circle of confirmation that AMAS is measuring Replikins activity in malignancy.

[000350] In addition to early detection by the AMAS test of the activity of the group of Replikins that are unique to cancer, Replikins are widely distributed markers of, and probably agents of, lethality. As newly recognized targets for prevention and therapy, Replikins offer a platform from which to control rapid replication and lethality of organisms and cells, without necessarily destroying them.

XIII. Replikin Count Correlates with Virulence and Lethality in Shrimp Taura Syndrome Virus

[000351] Applicants have likewise demonstrated in a blind study using an independent laboratory testing taura syndrome virus (TSV) in shrimp that virulence and mortality in shrimp correlates with Replikin Count in TSV. The inventors analyzed the genome of the TSV of four main isolates from Hawaii, Belize, Thailand and Venezuela to provide predictions ranking the virulence and mortality rate of each isolate. An independent laboratory tested each isolate in shrimp and provided blind data on mortality. The data demonstrate a quantitative linear correlation between Replikin concentration and mortality. *See Example 18.* Despite differences in epidemiology, virology and host, all of these data lend further support for the value of Replikin concentration in predicting outbreaks of pathogens and lethality of pathogens and malignancies.

XIV. Replikin Concentration in Replikin Peak Gene of Ribonucleotide Reductase Gene Area Correlated with a WSSV Epidemic

[000352] An increase in Replikin concentration in white spot syndrome virus (WSSV) is predictive of an increase in virulence of the virus and allows for prediction of forthcoming outbreaks or increases in morbidity and, in extreme cases, mortality. A review of publicly available amino acid sequences of isolates of WSSV that demonstrate an increase in Replikin Count in the genome or a genome segment, or in a protein or protein fragment of the virus over time or between isolates is used as a predictor of an increase in outbreaks in shrimp. Publicly available sequences for isolates of WSSV from PubMed or other public or private sources may be analyzed by hand or using proprietary search tool software (ReplikinForecast™ available in the United States from REPLIKINS LLC, Boston, MA).

[000353] Applicants have established a correlation between Replikin concentrations in WSSV and an increase in virulence of the virus resulting in epidemics. Applicants reviewed publicly available amino acid sequences of isolates of WSSV having accession numbers at www.pubmed.com and have identified a remarkable increase in Replikin concentration in the

Replikin Peak Gene of the ribonucleotide reductase gene area of the genome of the virus (e.g., SEQ ID NO: 669). The remarkable increase occurred just prior to a significant outbreak of WSSV in shrimp in 2001. Figure 18 illustrates a correlation between increases in Replikin Count in WSSV genome in 2000 and a significant outbreak of WSSV in 2001. In 2000, a remarkably high Replikin concentration of 97.6 is observed in WSSV. In the Replikin Peak Gene identified in ribonucleotide reductase in an isolate from 2000, the Replikin concentration spikes as high as 110.7, providing an unmistakable predictive signal for the significant 2001 outbreak of WSSV that followed. Analysis of the ribonucleotide reductase sequence publicly available at Accession No. AAL89390 (SEQ ID NO: 668) is disclosed in Example 10.

A. Analysis of Annual Replikin Count of WSSV

[000354] Applicants analyzed publicly available sequences for isolates of WSSV from PubMed. The data is contained in Table 18 and graphically described in Figure 18. Mean Replikin concentrations were determined for all amino acid sequences for WSSV with accession numbers publicly available at www.pubmed.com. The mean Replikin Count was then determined for all viruses isolated and reported in a particular year. Table WSSV provides the results of the Replikin Count analysis. Years with no data are not included in the table.

Table 18 - WSSV Replikin Count

Year	PubMed Accession Number-Replikin Count	No. Isolates	Mean Replikin Count	S.D.	Significance
1995	<u>CAA88950</u> <u>18</u> <u>CAA91970</u> <u>59</u>	2	4.4	0.6	low p<.10
1996	<u>CAE17687</u> <u>160</u> <u>CAB03144</u> <u>29</u> <u>CAB03173</u> <u>31</u>	3	6.0	2.6	low p>.50, prev p<.30
1998	<u>ABA54417</u> <u>48</u>	1	6.2	0.0	prev p>.50
2000		529	97.6	0.0	low p<.001
	<u>NP 478030 361 NP 478019 361 NP 478001 361 NP 477774 361 NP 477756 361 NP 477753 361</u> <u>NP 477809 361 NP 477768 361 NP 477523 361 NP 477959 361 NP 478053 361 NP 478052 361</u> <u>NP 478051 361 NP 478050 361 NP 478049 361 NP 478048 361 NP 478047 361 NP 478046 361</u> <u>NP 478045 361 NP 478044 361 NP 478043 361 NP 478042 361 NP 478041 361 NP 478039 361</u> <u>NP 478038 361 NP 478037 361 NP 478036 361 NP 478035 361 NP 478034 361 NP 478033 361</u> <u>NP 478032 361 NP 478031 361 NP 478029 361 NP 478028 361 NP 478027 361 NP 478026 361</u> <u>NP 478025 361 NP 478024 361 NP 478023 361 NP 478022 361 NP 478021 361 NP 478020 361</u> <u>NP 478018 361 NP 478017 361 NP 478016 361 NP 478015 361 NP 478014 361 NP 478013 361</u> <u>NP 478012 361 NP 478011 361 NP 478010 361 NP 478009 361 NP 478008 361 NP 478007 361</u> <u>NP 478006 361 NP 478005 361 NP 478004 361 NP 478003 361 NP 478002 361 NP 478000 361</u> <u>NP 477999 361 NP 477998 361 NP 477997 361 NP 477996 361 NP 477995 361 NP 477994 361</u> <u>NP 477993 361 NP 477992 361 NP 477991 361 NP 477990 361 NP 477989 361 NP 477988 361</u> <u>NP 477987 361 NP 477986 361 NP 477985 361 NP 477984 361 NP 477983 361 NP 477982 361</u>				

	<u>NP 477621 361 NP 477620 361 NP 477619 361 NP 477618 361 NP 477617 361 NP 477616 361</u> <u>NP 477615 361 NP 477614 361 NP 477613 361 NP 477612 361 NP 477611 361 NP 477610 361</u> <u>NP 477609 361 NP 477608 361 NP 477607 361 NP 477606 361 NP 477605 361 NP 477604 361</u> <u>NP 477603 361 NP 477602 361 NP 477601 361 NP 477600 361 NP 477599 361 NP 477598 361</u> <u>NP 477597 361 NP 477596 361 NP 477595 361 NP 477594 361 NP 477593 361 NP 477592 361</u> <u>NP 477591 361 NP 477590 361 NP 477589 361 NP 477588 361 NP 477587 361 NP 477586 361</u> <u>NP 477585 361 NP 477584 361 NP 477583 361 NP 477582 361 NP 477581 361 NP 477580 361</u> <u>NP 477579 361 NP 477578 361 NP 477577 361 NP 477576 361 NP 477575 361 NP 477574 361</u> <u>NP 477573 361 NP 477572 361 NP 477571 361 NP 477570 361 NP 477569 361 NP 477568 361</u> <u>NP 477567 361 NP 477566 361 NP 477565 361 NP 477564 361 NP 477563 361 NP 477562 361</u> <u>NP 477561 361 NP 477560 361 NP 477559 361 NP 477558 361 NP 477557 361 NP 477556 361</u> <u>NP 477555 361 NP 477554 361 NP 477553 361 NP 477552 361 NP 477551 361 NP 477550 361</u> <u>NP 477549 361 NP 477548 361 NP 477547 361 NP 477546 361 NP 477545 361 NP 477544 361</u> <u>NP 477543 361 NP 477542 361 NP 477541 361 NP 477540 361 NP 477539 361 NP 477538 361</u> <u>NP 477537 361 NP 477536 361 NP 477535 361 NP 477534 361 NP 477533 361 NP 477532 361</u> <u>NP 477531 361 NP 477530 361 NP 477529 361 NP 477527 361 NP 477526 361 NP 477525 361</u> <u>NP 477524 361</u>				
200 5	<u>AAZ29239 9 XP 001681561 6</u>	2	2.6	2.4	low p>0.20, prev p<0.001
200 6	<u>ABM92267 14 ABP01348 1</u> <u>ABM64218 6 ABI34434 6 ABI93178 4</u> <u>ABI93177 3 ABI93176 6 ABI93174 12</u> <u>ABQ12866 3 ABD65308 2 ABD65303</u> <u>1 ABD65302 4 ABD65300 3</u> <u>ABD65298 1</u>	14	2.7	2.5	low p<.001, prev p>.50
200 7	<u>2ED6 L 1 2ED6 K 1 2ED6 J 1 2ED6 I</u> <u>1 2ED6 H 1 2ED6 G 1 2ED6 F 1</u> <u>2ED6 E 1 2ED6 D 1 2ED6 C 1</u> <u>2ED6 B 1 2ED6 A 1 ABQ12772 15</u> <u>ABQ12773 3 ABQ12771 6 ABQ12770</u> <u>9 ABO69369 2 ABO69368 2 ABS00974</u> <u>5 ABS00973 1 ABQ44211 3 ABQ44210</u> <u>4 ABP52058 4 ABP52057 1 ABP52054</u> <u>5</u>	25	1.3	1.2	low p<.001, prev p<.05

B. Prediction and Treatment of WSSV Outbreaks

[000355] Prediction of epidemics and future outbreaks may be made, for example, by reviewing the Replikin Counts of isolates of WSSV and comparing the Replikin Count for a particular year with Replikin Counts from other years. A significant increase in Replikin Count from one year to the next and preferably over one, two, three or five years or more provides predictive value of an emerging strain of WSSV that may begin an outbreak of more highly virulent WSSV. A WSSV outbreak may be predicted within about six months to about one year, to about three, to about five years or more from the observation of a significant increase in Replikin concentration. The outbreak is preferably predicted within about one to about three years and more preferably within about one to about two years. An outbreak of WSSV, therefore, may be predicted within 1 to about 2 years as demonstrated in Figure 18 wherein an epidemic occurred at about 1 year following a remarkably significant increase in Replikin concentration and in particular in the identified Replikin Peak Gene.

[000356] Significant increases may be observed over a time period of more than one year, such as three, four, five or more years. An outbreak may likewise be predicted within about six months to about one year or more from the initial observation of an observable decrease in Replikin concentration following a notable increase.

[000357] The correlation between Replikin concentration and viral outbreaks noted above provide a method of predicting outbreaks of WSSV by monitoring increases or decreases in Replikin Count in the RPG of isolates of WSSV. The method may employ isolates of individual strains or isolates of all strains of WSSV.

XV. Replikin Count in TSV Epidemic

[000358] An increase in Replikin concentration in taura syndrome virus (TSV) is predictive of an increase in virulence and lethality of the virus and allows for prediction of forthcoming outbreaks or increases in lethality. Figure 19 illustrates a correlation between increased Replikin Count in the genome of TSV and outbreaks of the virus in 2000 and 2007 in shrimp. The Replikin Count data reflected in the graph is found in Table 19. Significant outbreaks of the disease are noted at years 2000 and 2007. It may be observed from the graph that outbreaks of the virus occur following an increase in Replikin concentration. In year 2000, TSV had a Replikin concentration of 2.7. Between 2001 and 2004, TSV had a lower mean Replikin concentration, as low as 0.7, and an identified Replikin Scaffold disappeared. In 2005 the Replikin Scaffold reappeared, with an increase in lysines and histidines, and a commensurate increase in Replikin concentration to 1.8, followed by an increase in TSV outbreaks in 2006-2007.

Table 19 -TSV Replikin Count

Year	PubMed Accession Number-Replikin Count	No. of Isolates per year	Mean Replikin Concentration per year	S.D.	Significance
2000	NP_149058 70 NP_149057 70 AAK72221 70 AAK72220 70 AAG44834 4	5	2.7	1.3	low p<0.02
2001	AAM73766 7	1	0.7	0.0	prev p<0.02
2002	AAN77089 2 AAN77088 2 AAN77087 2 AAN77086 2 AAW32934 2 AAW32932 2 AAW32930 2 AAW32929 1	8	0.7	0.4	low p>0.50
2003	AAR11292 6 AAR11291 6 AAR11290 6	3	0.6	0.0	prev p<0.20
2004	AAX07125 2 AAX07117 2 AAT81157 75 AAT81158 75 AAX07127 2 AAX07126 2 AAX07124 2 AAX07123 2 AAX07122 2 AAX07121 2 AAX07120 2 AAX07119 2 AAX07118 2 AAX07116 2 AAX07115 2 AAX07114 2 AAX07113 2 AAX07112 2 AAX35819 2 AAX35818 1 AAX35817 2 AAX35816 1 AAX35815 2	23	0.8	0.9	low p<0.40, prev p<0.20
2005	AAY56364 71 AAY56363 71 AAY44822 1 AAY44821 1 AAY44820 1 AAY44819 1 AAY44818 1 AAY44817 1 AAY89097 83 AAY89096 83 ABB17263 63 ABB17264 63	12	1.8	1.7	low p<0.02, prev p<0.05

[000359] The TSV is less virulent than WSSV and the structure of the TSV Replikin Scaffold is less closely related to influenza virus than are the structures of WSSV Replikin Scaffolds.

XVI. Software

[000360] A further aspect of the invention provides utilizing software that searches for Replikin Peak Genes and enables the discovery of the point or points in the genome that have the highest concentration of Replikins, the years in which they have occurred, the strain or strains in which they occur, the host or hosts in which they occur, the geographic locations in which they occur, their increase or decrease in the above years, strains, hosts and geographic location and point or small mutations that are correlatable with virulence.

XVII. SARS Replikin Concentration Correlates with Epidemics

[000361] An increase in Replikin concentration in coronaviruses also correlates with the SARS coronavirus epidemic. In particular, as may be seen in Figure 9, Replikin concentration in Spike and Nucleocapsid Coronavirus Proteins preceded the SARS Coronavirus Epidemic of 2003. In Figure 9, the x-axis indicates the year and the y-axis indicates the Replikin concentration. The appearance of the SARS outbreak is shown by the shaded area in the graph between 2003 and 2004. The peak of the shaded area represents a total number of eight countries in which the SARS outbreak occurred in 2003. The solid black symbols represent the mean Replikin concentration for spike coronavirus proteins and the vertical black bars represent the standard deviation of the mean.

[000362] Figure 9 shows a remarkable constancy of low coronavirus Replikin concentration between 1995 and 2001 in the spike proteins, followed by a dramatic increase in 2002, one year before the SARS epidemic appeared in 2003. Replikin concentration of the spike proteins in SARS then returned to their normal pre-2003 levels, which correlated with the disappearance of SARS.

XVIII. Passive Immunity

[000363] In another aspect of the invention, isolated Replikin peptides may be used to generate antibodies, which may be used, for example to provide passive immunity in an individual. Various procedures known in the art may be used for the production of antibodies to Replikin sequences. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, humanized, single chain, Fab fragments and fragments produced by a Fab expression library. Antibodies that are linked to a cytotoxic agent may also be generated. Antibodies may also be administered in combination with an antiviral agent. Furthermore,

combinations of antibodies to different Replikins may be administered as an antibody cocktail.

[000364] Monoclonal antibodies to Replikins may be prepared by using any technique that provides for the production of antibody molecules. These include but are not limited to the hybridoma technique originally described by Kohler and Milstein, (*Nature*, 1975, 256:495-497), the human B-cell hybridoma technique (Kosbor et al., 1983, *Immunology Today*, 4:72), and the EBV hybridoma technique (Cole et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96). In addition, techniques developed for the production of chimeric antibodies (Morrison et al., 1984, *Proc. Nat. Acad. Sci USA*, 81:6851-6855) or other techniques may be used. Alternatively, techniques described for the production of single chain antibodies (U.S. 4,946,778) can be adapted to produce Replikin-specific single chain antibodies.

[000365] Antibodies to any peptides observed to be present in an emerging or re-emerging strain of virus and combinations of such antibodies are useful in the treatment and/or prevention of viral infection, especially RPG peptides and Replikin sequences isolated within RPG peptides.

[000366] Antibody fragments that contain binding sites for a Replikin may be generated by known techniques. For example, such fragments include but are not limited to F(ab')₂ fragments which can be produced by pepsin digestion of the antibody molecules and the Fab fragments that can be generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries can be generated (Huse et al., 1989, *Science*, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

[000367] In another aspect of the invention, immune serum containing antibodies to one or more Replikins obtained from an individual exposed to one or more Replikins may be used to induce passive immunity in another individual or animal. Immune serum may be administered via i.v. to a subject in need of treatment. Passive immunity also can be achieved by injecting a recipient with preformed antibodies to one or more Replikins. Passive immunization may be used to provide immediate protection to individuals who have been exposed to an infectious organism. Administration of immune serum or preformed antibodies is routine and the skilled practitioner can readily ascertain the amount of serum or antibodies needed to achieve the desired effect. One of the reasons that vaccines directed towards a particular protein antigen of a disease causing agent have not been fully effective in

providing protection against the disease (such as foot and mouth vaccine which has been developed against the VP1 protein or large segments of the VP1 protein) is that the best antibodies have not been produced, that is – it is likely that the antibodies to the Replikins have not been produced.

[000368] For example, either epitopes other than Replikins present in the larger protein fragments may interfere according to the phenomenon of antigenic primacy and/or because the hydrolysis of larger protein sequences into smaller sequences for processing to produce antibodies results in loss of integrity of any Replikin structure that is present, *e.g.*, the Replikin is cut in two and/or the histidine residue is lost in the hydrolytic processing. The present studies suggest that for a more effective vaccine to be produced, the Replikin sequences, and no other epitope, should be used as the vaccine. For example, a vaccine of the invention can be generated using any one of the Replikin peptides identified by the three-point recognition system. A more preferred vaccine comprises at least one Replikin sequence isolated in an RPG. Another preferred vaccine comprises an RPG peptide. Among the preferred Replikin peptides for use in a virus vaccine are those conserved Replikins observed to “re-emerge” after an absence from the amino acid sequence for one or more years.

[000369] The Replikin peptides of the invention, alone or in various combinations are administered to a subject, preferably by *i.v.* or intramuscular injection, in order to stimulate the immune system of the subject to produce antibodies to the peptide. Generally the dosage of peptides is in the range of from about 0.1 μg to about 10 mg. In another embodiment, the dosage of the peptides is in the range from about 10 μg to about 1 mg. In a preferred embodiment, the dosage of the peptides is in the range from about 50 μg to about 500 μg . The skilled practitioner can readily determine the dosage and number of dosages needed to produce an effective immune response.

XIX. A Control Test of Reliability of Method of Predicting Outbreaks with Replikin Count

[000370] Table 3, which contains H5N1 data above, provides Replikin Count data across eight gene areas and an increased correlation is observed between mortality data and the whole virus, the polymerase gene and the pB1 gene area (Replikin Peak Gene). *See also, e.g.*, Figures 4, 16 and 17. In addition to the correlative aspect of the increase in Replikin Count and percent mortality, the data in Table 3, and all of the other data contained herein, above provide strong confirmation of the power and validity of the methodology of predicting changes in virulence and outbreaks of virus with changes in Replikin concentration. These data represent an objective test of the method of independently

selecting and examining several thousand individual accession numbers within approximately 12 million total accession numbers in PubMed wherein each selection is independently submitted to the PubMed database under a separate request using objective software. If there were not a reliable principle and a reliable method underlying each request, the potential for obtaining random results, or no results, or results which do not track each other at $p < 0.001$ would markedly increase. Table 3 provides results wherein p was less than 0.001 between each group as compared one to another.

[000371] In Table 3 the structures that are correlated have, to the knowledge of the Applicants, not been correlated before, that is, the inventors have examined the relationship of one internal virus structure to another internal virus structure or structures (*e.g.*, three-way relationship between whole virus gene area, polymerase and Replikin Peak Gene area) and have examined the external relation of these two or more internal structures to a host result of the virus infection, that is, percent mortality.

[000372] Table 3 represents consistent reproducible data, on repeated trials, which is the essence of the reliability of any method. For example, Table 3 provides independent data on (1) whole virus concentration of Replikins, (2) just polymerase concentration of Replikins, and (3) just the Replikin Peak Gene concentration of Replikins. The data is then correlated with H5N1 mortality three times, namely in 2003, 2004 and 2005. The absence of significant changes in the pA and pB2 gene areas provides a control. In each case, the method measures Replikin concentration three ways, each of which correctly predict mortality, independently, thereby confirming the method, and further illustrating in the process, the magnifying function of the Replikin Peak Gene.

EXAMPLES

Example 1

Calculation of Replikin Count of Replikin Peak Gene of an isolate of H3N2 from the pandemic year of 1968

[000373] The inventors queried Accession No. ABB54523 at www.pubmed.com.

Accession No. ABB54523 discloses the amino acid sequence of SEQ ID NO:1664, deduced from the genomic information of an H3N2 strain of *Influenza A* virus isolated in 1968 in Memphis. Upon analysis of SEQ ID NO:1664, the inventors observed a Replikin Peak Gene having continuous Replikin sequences beginning at residue 15 (histidine) and continuing through residue 85 (lysine) (SEQ ID NO: 1665).

[000374] The inventors isolated the RPG (SEQ ID NO:1665) *in silico*. SEQ ID NO:1665 was identified for diagnostic and therapeutic uses in, for example, an immunogenic

compound and a therapeutic vaccine compound and as a predictive sequence for lethal outbreaks of influenza. Seventeen Replikin sequences (SEQ ID NOS:1667-1682) were identified in the RPG of SEQ ID NO:1665 for diagnostic, therapeutic and predictive uses as described herein. SEQ ID NOS:1667-1674 were identified in the amino-terminal of the sequence disclosed in Accession No. ABB54523 (SEQ ID NO:1664), SEQ ID NOS:1675-1682 were identified in the mid-molecule of the sequence.

[000375] The Replikin Count of the amino acid sequence (SEQ ID NO:1664) disclosed at ABB54523 was seventeen Replikin sequences in 90 total amino acids for a Replikin Count of 18.9. The Replikin Count of the RPG (SEQ ID NO:1665) was seventeen Replikin sequences in 71 total amino acids for a Replikin Count of 23.9.

Example 2

Calculation of Replikin Count of Replikin Peak Gene of an isolate of human H5N1 from 2003

[000376] The inventors queried Accession No. BAE07199 at www.pubmed.com. Accession No. BAE07199 discloses an amino acid sequence deduced from the genomic information of the RNA polymerase gene of an H5N1 strain of *Influenza A* virus isolated in 2003 in Hong Kong. The inventors analyzed the whole pB1 gene area (SEQ ID NO:1683) of the polymerase sequence. Upon analysis of SEQ ID NO:1683, the inventors observed a Replikin Peak Gene having continuous Replikin sequences that begin at residue 168 (lysine) and continue through residue 215 (lysine).

[000377] The inventors isolated the RPG (SEQ ID NO:1684) *in silico*. SEQ ID NO:1684 was identified for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for lethal outbreaks of influenza. Seven Replikin sequences (SEQ ID NOS:1685-1691) were identified in the RPG of SEQ ID NO:1684 for diagnostic, therapeutic and predictive uses as described herein. Replikin sequences SEQ ID NOS:1685-1691 were identified in the amino-terminal of the sequence disclosed in Accession No. BAE07199 (SEQ ID NO:1683), Replikin sequences SEQ ID NOS:1692-1694 were identified in the mid-molecule of the sequence, and Replikin sequence SEQ ID NOS:1695-1699 were identified in the carboxy-terminal of the sequence.

[000378] The Replikin Count of the whole pB1 area sequence (SEQ ID NO:1683) was 15 Replikin sequences in 757 total amino acids for a Replikin Count of 2.0. The Replikin Count of the RPG (SEQ ID NO:1684) was seven Replikin sequences in 48 total amino acids for a Replikin Count 14.6.

Example 3

Calculation of Replikin Count of the pB1 gene area and pB1-F2 sub-gene area of an isolate of human H5N1 from Indonesia in 2006

[000379] The inventors queried Accession No. ABI36257 at www.pubmed.com.

Accession No. ABI36257 discloses an amino acid sequence deduced from the genomic information of the pB1 gene area of an H5N1 strain of *Influenza A* virus isolated in 2006 from Indonesia. The inventors analyzed the pB1-F2 gene area (SEQ ID NO:1700). Upon analysis of SEQ ID NO:1700, the inventors observed a Replikin Peak Gene having continuous Replikin sequences that begin at residue 15 (histidine) and continue through residue 85 (lysine) (SEQ ID NO: 1701).

[000380] The inventors isolated the RPG (SEQ ID NO:1701) *in silico* for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for lethal outbreaks of influenza. Sixteen Replikin sequences (SEQ ID NOS:1702-1717) were identified in the RPG of SEQ ID NO:1701 for diagnostic, therapeutic and predictive uses as described herein. Replikin sequences SEQ ID NOS:1702-1703 were identified in the amino-terminal of the sequence of SEQ ID NO:1701, Replikin sequences SEQ ID NOS:1704-1717 were identified in the mid-molecule of the sequence, and no Replikin sequences were identified in the carboxy-terminal.

[000381] The Replikin Count of the whole pB1-F2 gene area sequence (SEQ ID NO:1700) was 16 Replikin sequences in 90 total amino acids for a Replikin Count of 17.8. The Replikin Count of the RPG pB1-F2 subgene area (SEQ ID NO:1701) was 16 Replikin sequences in 71 total amino acids for a Replikin Count 22.57.

Example 4**Calculation of Replikin Count of the pB1 gene area and pB1-F2 sub-gene area of an isolate of human H5N1 from Indonesia in 2007**

[000382] The inventors queried Accession No. ABM90520 at www.pubmed.com.

Accession No. ABM90520 discloses an amino acid sequence deduced from the genomic information of the pB1 gene area of an H5N1 strain of *Influenza A* virus isolated in 2007 in Indonesia. The inventors analyzed the pB1 gene area (SEQ ID NO:1722). Upon analysis of SEQ ID NO:1722, the inventors observed a Replikin Peak Gene having continuous Replikin sequences that begin at residue 15 (histidine) and continue through residue 85 (lysine) in the pB1-F2 gene area (SEQ ID NO: 1723)

[000383] The inventors isolated the RPG (SEQ ID NO:1723) *in silico*. SEQ ID NO:1723 was identified for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for lethal

outbreaks of influenza. Sixteen Replikin sequences (SEQ ID NOS:1724-1739) were identified in the RPG (or pB1-F2 gene subarea) of SEQ ID NO:1723 for diagnostic, therapeutic and predictive uses as described herein. Replikin sequences SEQ ID NOS:1724-1725 were identified in the amino-terminal of the sequence of SEQ ID NO:1723, Replikin sequences SEQ ID NOS:1726-1739 were identified in the mid-molecule of the sequence, and no Replikin sequences were identified in the carboxy-terminal.

[000384] The Replikin Count of the whole pB1-F2 area sequence (SEQ ID NO:1722) was 16 Replikin sequences in 90 total amino acids for a Replikin Count of 17.8. The Replikin Count of the RPG (SEQ ID NO:1723) was 16 Replikin sequences in 71 total amino acids for a Replikin Count 22.5.

Example 5

Calculation of Replikin Count of the RPG of a 2007 H1N1 Isolate from Thailand Having a Replikin Scaffold

[000385] The inventors queried Accession No. ABS71678 at www.pubmed.com. Accession No. ABS71678 discloses an amino acid sequence deduced from the genomic information of the hemagglutinin gene area of an H1N1 strain of *Influenza A* virus isolated in 2007 in Thailand. The inventors analyzed the amino acid sequence provided at ABS71678 (SEQ ID NO:1995). Upon analysis of SEQ ID NO:1995, the inventors observed a Replikin Peak Gene having continuous Replikin sequences that begin at residue 143 (histidine) and continue through residue 235 (lysine) (SEQ ID NO: 1996). A Replikin Scaffold, knglypnlsksyannkekevlvlgvhh (SEQ ID NO:2011) was observed within the RPG.

[000386] The inventors isolated the RPG (SEQ ID NO:1996) *in silico*. SEQ ID NO:1996 was identified for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for lethal outbreaks of influenza. Twenty-six Replikin sequences (SEQ ID NOS:1999-2024) were identified in the RPG of SEQ ID NO:1996 for diagnostic, therapeutic and predictive uses as described herein. Replikin sequences SEQ ID NOS:1997-2016 were identified in the amino-terminal of the sequence of SEQ ID NO:1995, Replikin sequences SEQ ID NOS:2017-2029 were identified in the mid-molecule of the sequence, and SEQ ID NOS:2030-2042 were identified in the carboxy-terminal. The Replikin sequences were isolated for diagnostic, therapeutic and predictive uses.

[000387] The Replikin Count of the whole hemagglutinin sequence (SEQ ID NO:1995) was 46 Replikin sequences in 564 total amino acids for a Replikin Count of 8.1. The

Replikin Count of the RPG area (SEQ ID NO:1996) was 26 Replikin sequences in 93 total amino acids for a Replikin Count of 28.

Example 6

Replikin Peak Gene Identification in EIV Isolate Reported at Accession No. ABS89395

[000388] Applicants reviewed Replikin sequences publicly available at www.pubmed.com to determine the Replikin Peak Gene Area of available isolates. A Replikin Peak Gene was identified in the pB1-F2 gene area of the virus in Accession No. ABS89395 at www.pubmed.com. The following example provides determination of the Replikin Peak Gene in a 2005 isolate of a Maryland strain of H3N8 serotype *Influenza A* virus.

[000389] The inventors queried Accession No. ABS89395 and analyzed the amino acid sequence provided (SEQ ID NO:545). Upon analysis of the sequence, the inventors observed a Replikin Peak Gene having continuous Replikin sequences that begin at residue 15 (histidine) and continue through residue 85 (lysine) (SEQ ID NO: 546).

[000390] The inventors isolated the RPG (SEQ ID NO:546) *in silico* and identified the sequence for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for lethal outbreaks of tuberculosis. Sixteen Replikin sequences (SEQ ID NOS: 547-562) were identified in the RPG for diagnostic, therapeutic and predictive uses as described herein. Replikin sequences SEQ ID NOS:547-548 were identified in the amino-terminal of the sequence, Replikin sequences SEQ ID NOS:549-562 were identified in the mid-molecule of the sequence, and no Replikins were identified in the carboxy-terminal.

[000391] The Replikin Count of the whole pB1-F2 sequence (SEQ ID NO:545) was 16 Replikin sequences in 90 total amino acids for a Replikin Count of 17.8. The Replikin Count of the RPG area (SEQ ID NO:546) was 16 Replikin sequences in 71 total amino acids for a Replikin Count of 22.5.

Example 7

Replikin Peak Gene Identification in West Nile Virus Isolate Reported in Accession No. ABA54585

[000392] Applicants reviewed Replikin sequences publicly available at www.pubmed.com to determine the Replikin Peak Gene of an available West Nile Virus (WNV) isolate. The entire envelope protein of WNV was reported at Accession No. ABA54585. A Replikin Peak Gene was identified in the 3,433 amino acid polyprotein

sequence of the WNV envelope protein. A Replikin Peak Gene was identified beginning at amino acid residue 2797 extending through amino acid residue 2836 (a total of 40 amino acid residues). The number of Replikin sequences in this section was 12. The Replikin Count (Replikins per 100 amino acids) was 30. The Replikin Peak Gene (RPG) of the envelope protein of WNV is SEQ ID NO:258 and the RPG contains 12 uninterrupted Replikins (SEQ ID NOS: 246-257).

Example 8

Calculation of RPGs in Porcine Respiratory Syndrome Virus

[000393] Applicants reviewed Replikin sequences publicly available at www.pubmed.com to determine the Replikin Peak Gene Area of available PRRSV isolates. A Replikin Peak Gene was identified in Accession No. AAO43261 from mRNA encoding a reported nucleocapsid protein of a PRRSV isolate from Mexico in 2003. The inventors analyzed the amino acid sequence provided in Accession No. AAO43261, which is reported with 123 amino acids within ORF 7 of the virus genome. Upon analysis of the sequence, the inventors observed a Replikin Peak Gene having continuous Replikin sequences that begin at residue 7 (lysine) and continue through residue 66 (histidine).

[000394] The inventors isolated the RPG (SEQ ID NO:394) *in silico* and identified the sequence for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for outbreaks of PRRSV. Seven Replikin sequences (SEQ ID NOS: 395-401) were identified in the RPG for diagnostic, therapeutic and predictive uses as described herein. SEQ ID NO:395 was identified in the amino-terminal portion of the sequence and SEQ ID NOS:396-401 were identified in the mid-molecule portion of the sequence.

[000395] The Replikin Count of the whole nucleocapsid sequence at Accession No. AAO43261 was 7 Replikin sequence in 123 amino acid residues or 5.7. The Replikin Count of the RPG area (SEQ ID NO:394) was 7 Replikin sequences in 60 total amino acids for a Replikin Count of 11.7.

[000396] The asparagine and methionine residues at positions 45 and 46 of the RPG (SEQ ID NO:394) were identified by the inventors as non-conserved positions within the RPG as compared to other reported nucleocapsid sequences such as Accession No. ABF19568 discussed immediately below. Non-conserved positions within an RPG that are correlated with changes in lethality and/or virulence are particularly useful in methods of the invention

to predict outbreaks. The presence of these point mutations in other PRRSV nucleocapsid RPG sequences provides evidence of greater virulence and/or lethality.

[000397] A Replikin Peak Gene was also identified in Accession No. ABF19568 from a nucleic acid sequence of a PRRSV 2006 isolate from Mexico. The reported sequence has 99 amino acid residues. The RPG (SEQ ID NO:402) was isolated *in silico* and identified for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for outbreaks of PRRSV. The total length of the RPG is 29 amino acids. The Replikin Count is 41.4. The Replikin sequences of SEQ ID NOS:403-414 were identified in the RPG for diagnostic, therapeutic and predictive uses as described herein. SEQ ID NOS:403-414 were identified in the amino-terminal of the sequence. No Replikin sequences were identified in the mid-molecule or carboxy-terminus.

[000398] The glycine, proline and glycine residues at positions 14 through 16 and the asparagine, arginine, lysine, arginine and asparagine residues at positions 21 through 25 were identified within the RPG (SEQ ID NO:507) as non-conserved positions as compared to other reported nucleocapsid sequences such as Accession No. AAO43261 above. Further, as compared to the RPG in Accession No. AAO43261 above, the RPG identified in the 2006 Mexico isolate demonstrates a shortening of the RPG and noteworthy condensation of Replikin sequences within the shorter RPG. The result is a remarkable increase in Replikin Count between 2003 and 2006 corresponding to a severe outbreak of PRRSV in Mexico in 2006 with an increase in mortality rate.

[000399] Applicants likewise analyzed Accession Nos. AAM18565, AAP81809, ABL60920 having sequences of isolates from China in 2000, 2003, and 2006, respectively, to determine the Replikin Peak Gene of the isolates (SEQ ID NOS: 341, 342, and 343).

[000400] A Replikin Peak Gene was identified in Accession No. AAM18565 between residue 7 (lysine) and residue 66 (histidine). The inventors isolated the RPG (SEQ ID NO:353) *in silico* and identified the sequence for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for outbreaks of PRRSV. Thirteen Replikin sequences (SEQ ID NOS: 354-366) were identified in the RPG for diagnostic, therapeutic and predictive uses as described herein. SEQ ID NOS:354-357 were identified in the amino-terminal portion of the sequence and SEQ ID NOS:358-366 were identified in the mid-molecule portion of the sequence.

[000401] The Replikin Count of the whole sequence at Accession No. AAM18565 was 13 Replikin sequences within 123 amino acid residues or 10.6. The Replikin Count of the RPG area (SEQ ID NO:353) was 13 Replikin sequences in 60 total amino acids for a Replikin Count of 21.7.

[000402] A Replikin Peak Gene was identified in Accession No. AAP81809 between residue 7 (lysine) and residue 66 (histidine). The inventors isolated the RPG (SEQ ID NO:367) *in silico* and identified the sequence for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for outbreaks of PRRSV. Thirteen Replikin sequences (SEQ ID NOS: 368-380) were identified in the RPG for diagnostic, therapeutic and predictive uses as described herein. SEQ ID NOS:368-371 were identified in the amino-terminal portion of the sequence and SEQ ID NOS:372-380 were identified in the mid-molecule portion of the sequence.

[000403] The Replikin Count of the whole sequence at Accession No. AAP81809 was 13 Replikin sequences within 123 amino acid residues or 10.6. The Replikin Count of the RPG area (SEQ ID NO:367) was 13 Replikin sequences in 60 total amino acids for a Replikin Count of 21.7.

[000404] A Replikin Peak Gene was identified in Accession No. ABL60920 between residue 7 (lysine) and residue 66 (histidine). The inventors isolated the RPG (SEQ ID NO:382) *in silico* and identified the sequence for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for outbreaks of PRRSV. Ten Replikin sequences (SEQ ID NOS: 384-393) were identified in the RPG for diagnostic, therapeutic and predictive uses as described herein. SEQ ID NOS:384-387 were identified in the amino-terminal portion of the sequence and SEQ ID NOS:388-393 were identified in the mid-molecule portion of the sequence.

[000405] The Replikin Count of the whole sequence at Accession No. ABL60920 was 10 Replikin sequences within 123 amino acid residues or 8.1. The Replikin Count of the RPG area (SEQ ID NO:367) was 10 Replikin sequences in 60 total amino acids for a Replikin Count of 16.7.

Example 9

Calculation of RPG in Porcine Circovirus

[000406] Applicants reviewed Replikin sequences publicly available at www.pubmed.com to determine the Replikin Peak Gene of available isolates of PCV. The

inventors identified and compared a Replikin Peak Gene (RPG) of a protein fragment at Accession No. AAC59472 of a strain of PCV isolated from infected pigs in Manitoba, Canada in 1997 and a RPG of a putative truncated replicase protein at Accession No. ABP68657 of a strain of PCV isolated from infected pigs in China in 2007. The AAC59472 fragment was identified from nucleic acid encoding a predicted 1.8 kDa protein in open reading frame 11 of the isolate. The ABP68657 putative truncated replicase protein was identified in open reading frame 1 of the isolate.

[000407] In Accession No. AAC59472, the inventors identified an RPG (SEQ ID NO: 520) for diagnostic, therapeutic and predictive purposes as described herein. The RPG begins at residue 2 (lysine) and continues through residue 12 (lysine). Four Replikin sequences (SEQ ID NOS:521-524) were identified for diagnostic, therapeutic and predictive uses as described in herein. The total length of the RPG is 11 amino acids. The Replikin Count is 36.4. The Replikin Count of the entire fragment is four Replikin sequences in fourteen amino acids or 28.6.

[000408] In Accession No. ABP68657, the inventors identified an RPG (SEQ ID NO:525) for diagnostic, therapeutic and predictive purposes as described herein. Thirteen Replikin sequences (SEQ ID NOS:526-538) were identified for diagnostic, therapeutic and predictive uses as described in herein. The total length of the RPG is 38 amino acids. The Replikin Count is 34.2. The Replikin Count of the entire putative truncated protein is 6.2.

[000409] The reported sequence at Accession No. AAC59472 has only 14 amino acid residues. Nevertheless, the high concentration of continuous, non-interrupted and overlapping Replikin sequences within the RPG (Replikin Count 36.4) is a predictor of virulence and provides sequences available as vaccines. In comparison, the RPG of the truncated replicase protein reported at Accession No. ABP68657 has 306 amino acid residues but the identified RPG has 13 Replikin sequences and a comparable Replikin Count of 34.2, which is likewise a predictor of virulence and provides sequences available as vaccines. Likewise, the high Replikin Count RPGs provide a target for production of immunogenic compounds for treatment and prevention of PCV.

[000410] A Replikin Peak Gene was identified in an isolate of PCV from 1997 publicly available at Accession No. AAC9885 between residues 4 (lysine) and 99 (histidine). The inventors isolated the RPG (SEQ ID NO:421) *in silico* and identified the sequence for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for outbreaks of PCV. Fourteen Replikin

sequences (SEQ ID NOS: 422-435) were identified in the RPG for diagnostic, therapeutic and predictive uses as described herein. SEQ ID NOS:422-435 were identified in the amino-terminal portion of the whole sequence disclosed at the accession number and SEQ ID NOS:436-437 were identified in the mid-molecule portion of the sequence. No Replikin sequences were identified in the carboxy-portion of the sequence.

[000411] The Replikin Count of the whole sequence at Accession No. AAC9885 was 16 Replikin sequences within 312 amino acid residues or 5.1. The Replikin Count of the RPG area (SEQ ID NO:421) was 14 Replikin sequences in 96 total amino acid residues for a Replikin Count of 14.6.

[000412] A Replikin Peak Gene was identified in an isolate of PCV from 2001 publicly available at Accession No. AAL01075 between residues 57 (histidine) and 94 (lysine). The inventors isolated the RPG (SEQ ID NO:438) *in silico* and identified the sequence for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for outbreaks of PCV. Twelve Replikin sequences (SEQ ID NOS: 439-450) were identified in the RPG for diagnostic, therapeutic and predictive uses as described herein. SEQ ID NOS:439-445 were identified in the amino-terminal portion of the whole sequence disclosed at the accession number and SEQ ID NOS:446-450 were identified in the mid-molecule portion of the sequence. No Replikin sequences were identified in the carboxy-portion of the sequence.

[000413] The Replikin Count of the whole sequence at Accession No. AAC9885 was 12 Replikin sequences within 314 amino acid residues or 3.8. The Replikin Count of the RPG area (SEQ ID NO:438) was 12 Replikin sequences in 90 total amino acids for a Replikin Count of 13.3.

[000414] A Replikin Peak Gene was identified in an isolate of PCV from Canada in 2007 that is publicly available at Accession No. ABP68657. The RPG was identified between residues 57 (histidine) and 94 (lysine). The inventors isolated the RPG (SEQ ID NO:462) *in silico* and identified the sequence for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for outbreaks of PCV. Fourteen Replikin sequences (SEQ ID NOS: 462-476) were identified in the RPG for diagnostic, therapeutic and predictive uses as described herein. SEQ ID NOS:462-476 were identified in the amino-terminal portion of the whole sequence disclosed at the accession number and SEQ ID NOS:477-481 were identified in the mid-molecule

portion of the sequence. No Replikin sequences were identified in the carboxy-portion of the sequence.

[000415] The Replikin Count of the whole sequence at Accession No. ABP68657 was 19 Replikin sequences within 306 amino acid residues or 6.2. The Replikin Count of the RPG area (SEQ ID NO:462) was 14 Replikin sequences in 38 total amino acids for a Replikin Count of 36.8.

[000416] In Applicants' review of RPGs in publicly available PCV sequences, the inventors identified an RPG from Accession No. ABQ10608 that contained each of the highly conserved Replikin sequences discussed in Section XI.G. above, namely, kngrsgpqqphk (SEQ ID NO:345); hlqgfanfkkqtfnk (SEQ ID NO:346) and kkqtfnkvwylgarch (SEQ ID NO:347).

[000417] The RPG was identified between residues 57 (histidine) and 94 (lysine). The inventors isolated the RPG (SEQ ID NO:498) *in silico* and identified the sequence for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for outbreaks of PCV. Six Replikin sequences (SEQ ID NOS: 487-492) were identified in the RPG for diagnostic, therapeutic and predictive uses as described herein. SEQ ID NOS:486-492 were identified in the amino-terminal portion of the whole sequence disclosed at the accession number and SEQ ID NOS:493-497 were identified in the mid-molecule portion of the sequence for therapeutic, diagnostic and predictive purposes. No Replikin sequences were identified in the carboxy-portion of the sequence.

[000418] The Replikin Count of the whole sequence at Accession No. ABQ10608 was 12 Replikin sequences within 314 amino acid residues or 3.8. The Replikin Count of the RPG area (SEQ ID NO:498) was six Replikin sequences in 38 total amino acids or 15.8.

Example 10

Calculation of RPG in Tuberculosis Pathogen *Mycobacterium*

[000419] The inventors queried Accession No. AAS59518 at www.pubmed.com. Accession No. AAS59518 discloses an amino acid sequence from *Mycobacterium mucogenicum* strain CIP 105384. The inventors analyzed the amino acid sequence provided at AAS59518 (SEQ ID NO:2901). Upon analysis of SEQ ID NO:2901, the inventors observed a Replikin Peak Gene having continuous Replikin sequences that begin at residue 3 (histidine) and continue through residue 88 (histidine) (SEQ ID NO: 3649).

[000420] The inventors isolated the RPG (SEQ ID NO:3659) *in silico* for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for lethal outbreaks of tuberculosis. Twenty-four Replikin sequences (SEQ ID NOS:2902-2925) were identified in the RPG of SEQ ID NO:3659 for diagnostic, therapeutic and predictive uses as described herein. Replikin sequences SEQ ID NOS:2902-2924 were identified in the amino-terminal of the sequence of SEQ ID NO:2901, Replikin sequences SEQ ID NO: 2925 was identified in the mid-molecule of the sequence, and no Replikins were identified in the carboxy-terminal. All were isolated for diagnostic, therapeutic and predictive purposes.

[000421] The Replikin Count of the whole hemagglutinin sequence (SEQ ID NO:2901) was 24 Replikin sequences in 147 total amino acids for a Replikin Count of 16.3. The Replikin Count of the RPG area (SEQ ID NO:3659) was 24 Replikin sequences in 87 total amino acids for a Replikin Count of 27.6.

Example 10

Determination of Very High Replikin Count in WSSV Ribonucleotide Reductase from Accession No. AAL89390

[000422] Replikin concentration was determined for ribonucleotide reductase of a white spot syndrome virus (WSSV) isolate publicly available at Accession No. AAL89390. The amino acid sequence was translated from the total genome of a 2000 isolate of WSSV made publicly available at Accession No. NC 003225.1. The Replikin concentration in the protein was an unusually high 103.8 and the Replikin concentration of the Replikin Peak Gene of the protein was a yet higher 110.7.

[000423] The amino acid sequence of the protein publicly available at Accession No. AAL89390 is of particular interest because it demonstrates an overlapping of Replikin sequences that results in very high Replikin concentrations comparable to *P. falciparum*. The high concentrations of Replikin sequences provide a reservoir for transfer to influenza viruses.

[000424] In Accession No. AAL89390, SEQ ID NO:668 is disclosed as a ribonucleotide reductase protein of white spot syndrome virus. Within SEQ ID NO:668, the inventors identified a Replikin Peak Gene (SEQ ID NO:669). The Replikin Peak Gene is observed to occupy most of the disclosed protein of SEQ ID NO:668. The expansiveness of the Replikin Peak Gene across most of the amino acid sequence of the protein is highly unusual and creates a remarkably high Replikin concentration.

[000425] Replikin Count of SEQ ID NO:668 was determined by dividing the number of Replikin sequences identified in the amino acid sequence of the protein, 497 Replikin sequences, by the total amino acid length of the protein, 479 amino acids, to arrive at 103.8 Replikin sequences per 100 amino acids. The Replikin Count of the RPG of SEQ ID NO:669 was determined by dividing the number of Replikin sequences identified in the segment of the protein containing the highest concentration of continuous Replikin sequences, 497 Replikin sequences, by the total amino acid length Replikin Peak Gene, 449 amino acids, to arrive at 110.7.

[000426] Within the RPG of SEQ ID NO:669, SEQ ID NOS:670-1166 were identified as Replikin sequences. SEQ ID NOS:669-866 were identified in the amino-terminus of the peptide, SEQ ID NOS:867-1065 were identified in the middle portion, and SEQ ID NOS:1066-1166 were identified in the carboxy-terminus.

[000427] SEQ ID NO:669 was further observed to contain significant Replikin Scaffold sequences. SEQ ID NOS: 663-667 were identified as Replikin Scaffold repeats and were isolated for diagnostic, therapeutic and predictive uses.

[000428] Replikin Count was determined for a functionally undefined protein in the genome of a 2000 isolate WSSV at Accession No. NP 478030 (SEQ ID NO:1167). The Replikin Count in the protein was again an unusually high 97.6 Replikin sequences per 100 amino acids determined by dividing the number of Replikin sequences identified in the amino acid sequence of the protein, 361 Replikin sequences, by the total amino acid length of the protein, 370 amino acids.

[000429] An RPG (SEQ ID NO:1168) was identified within SEQ ID NO:1167 between residues 22 (histidine) and 361 (lysine) and is available for diagnostic, therapeutic and predictive uses as described herein. Total Replikin sequences identified in the RPG were 361 with total amino acid residues of 361, for a Replikin Count in the RPG of 100. SEQ ID NOS:1169-1330 were identified in the amino-terminus of the RPG. SEQ ID NOS:1331-1465 were identified in the mid-molecule of the RPG and SEQ ID NOS:1466-1529 were identified in the carboxy-terminus of the RPG. Each Replikin sequence is available for diagnostic, therapeutic and predictive purposes as described herein.

[000430] The amino acid sequence of Accession No. NP 478030 is of interest because, like the protein in Accession No. AAL89390, it demonstrates an overlapping of Replikin sequences that results in very high Replikin concentrations comparable to the highly-replicating *P. falciparum* of malaria. Overlapping Replikin sequences are exceptional targets

for therapies such as immunogenic agents and vaccines and have excellent predictive capacities.

[000431] In 2006 and 2007 WSSV has been observed to be dormant in shrimp. This continued decline of WSSV into “quiescent” or “dormant” levels in 2006-2007 is demonstrated in mean Replikin Counts for viruses isolated during 2005-2007 that are very low as compared to years wherein the virus demonstrated greater virulence, such as 2001. The continued quiescence in WSSV in 2007 may be contrasted with an observed rising of Replikin concentration in taura syndrome virus Replikin during this period.

[000432] As may be seen from the analysis below, Accession Nos. ABS00973 and AAW88445 have low observed Replikin concentrations. ABS00973 contains a single Replikin sequence (SEQ ID NO:1548) in the entire disclosed amino acid sequence of 240 residues at SEQ ID NO:1547. The Replikin concentration of Accession No. ABS00973 is an inordinately low 0.5. AAW88445 contains a white spot syndrome virus protein of 261 amino acid residues (SEQ ID NO:1530). An RPG of 34-105 was identified (SEQ ID NO:1531). Within the RPG, eleven Replikin sequences were identified (SEQ ID NOS:1532-1542). SEQ ID NOS:1532-1542 were identified in the amino-terminus of SEQ ID NO:1530 and SEQ ID NOS:1543-1546 were identified in the carboxy-terminus of SEQ ID NO:1530.

Example 11

Calculation of Replikin Count in Accession No. AAM73766 and AAY89096 in Taura Syndrome Virus

[000433] The inventors queried Accession No. AAM73766 at www.pubmed.com. Accession No. AAM73766 discloses an amino acid sequence from a 2005 isolate of TSV (SEQ ID NO:3566). Applicants identified SEQ ID NOS:3567-3569 as Replikin sequences in the amino-terminus of the sequence and SEQ ID NOS:3570-3573 as Replikin sequences in the carboxy-terminus of the sequence. Each sequence was isolated *in silico* for diagnostic, therapeutic and predictive purposes as described herein. No Replikin sequence was identified in the mid-molecule. The Replikin Count of SEQ ID NO:3566 was seven Replikin sequences in 1011 amino acid residues or 0.7.

[000434] The inventors queried Accession No. AAY89096 at www.pubmed.com. Accession No. AAY89096 discloses an amino acid sequence from a 2005 isolate of TSV (SEQ ID NO:3574). Applicants identified SEQ ID NOS:3575-3587 in the amino-terminus of the sequence. SEQ ID NOS:3588-3634 were identified as Replikin sequences in the mid-molecule. And SEQ ID NOS: 3635-3657 were identified as Replikin sequences in the carboxy-terminus of the sequence. Each sequence was isolated *in silico* for diagnostic,

therapeutic and predictive purposes as described herein. Replikin Count of SEQ ID NO:3574 was 83 Replikin sequences in 2107 amino acid residues or 3.9.

Example 12

Analysis of pB1-F2 Gene Area for All Publicly Available Influenza A Strains 2002-2007 [000435] The inventors queried www.pubmed.com with the software program

FluForecast[®] available from Replikins LLC of Boston, MA to analyze all amino acid sequences from the pB1-F2 gene area of all isolates of *Influenza A* available between 2002 and 2007. Table 20 provides the results of the query. The data for mean Replikin count for 2005, 2006 and 2007 suggest that the current epidemic is not over. For example, the SARS data in Figure 9 demonstrates that prior to a decline in epidemic infections, a decrease in Replikin Count is expected. Such a decline is not seen in the data in Table 20.

[000436]

Table 20 – PB1-F2 Influenza A 2002-2007

Year	PubMed Accession Number-Replikin Count	No. of Isolates per year	Mean Replikin Count per year	S.D.	Significance
2002	ABD59827 15 ABD59825 15 ABD59823 12	3	1.8	0.2	low p<.05, prev p<.001
2003	ABD59830 12 ABK40004 12 AAZ79547 12 AAZ79504 15	4	1.7	0.2	low p<.005, prev p<.30
2004	ABD59835 12 ABD59833 12	2	1.6	0.0	prev p<.30
2005	ABI36231 16 ABI36226 16 ABI36221 16 ABI36217 16 ABI36215 15 ABI36210 15 ABI36032 14 ABI36021 14 ABI36010 15 ABI36001 14	10	9.6	8.1	low p<.02
2006	BAF37385 16 BAE97585 16 ABL31778 16 ABL31775 21 ABL31764 21 ABL31742 16 ABI49394 16 ABL07028 21 ABL07017 16 ABL07006 16 ABI49405 16 ABI36473 16 ABI36462 16 ABI36453 16 ABI36442 16 ABI36431 16 ABI36421 16 ABI36409 16 ABI36398 16 ABI36387 16 ABI36376 16 ABI36365 16 ABI36354 15 ABI36343 15 ABI36332 15 ABI36321 15 ABI36310 16 ABI36304 16 ABI36293 16 ABI36284 16 ABI36272 16 ABI36269 16 ABI36266 16 ABI36262 16 ABI36258 16 ABI36253 16 ABI36250 16 ABI36245 14 ABI36242 14 ABI36237 14 ABI36233 14 ABI36206 16 ABI36196 16 ABI36185 16 ABI36175 16 ABI36164 16 ABI36153 16 ABI36142 16	48	16.6	4.7	low p<.001, prev p<.01
2007	ABM90520 16 ABM90542 16 ABM90531 16 ABM90509 16 ABM90498 16 ABM90487 16 ABM90476 16 ABM90465 16 ABM90454 16 ABM90443 16 ABM90432 16	11	17.8	0.0	low p<.001

Example 13

H1N1 hemagglutinin (whole hemagglutinin) Replikin Count by Year

[000437] The inventors queried www.pubmed.com with the software program

FluForecast[®] available from Replikins LLC of Boston, MA to analyze all amino acid

sequences of all isolates of H1N1 *Influenza A* available between 1917 and 2007. Table 21 provides the results of the query.

Table 21 – H1N1 Influenza A 1918-2007

Year	PubMed Accession Number-Replikin Count	No. of Isolates per year	Mean Replikin Count per year	S.D.	Significance
1918	AAO65768 13	1	7.0	0.0	
1919	AAO65769 13	1	7.0	0.0	
1930	O9WCD9 27 AAD25303 27 AAB52905 16 CAA40729 16 P05779 3	5	4.4	0.7	low p<.001, prev p<.001
1931	AAA19935 16 ABD79255 27	2	4.7	0.1	low p<.01, prev p<.30
1933	ABD77796 31 ABF47955 31 P03454 31 AAA43209 31 P05780 2 P03470 11	6	4.4	1.7	low p<.01, prev p>.50
1934	AAA58799 21 ABP64731 38 ABP64721 35 AAA43661 7 ABO21709 38 ABD77675 38 P03452 38 AAA43194 21 AAM75158 38 YP_163736 38 YP_163735 38 NP_040980 38 P03468 11 P06821 6	14	6.1	1.3	low p<.02, prev p<.04
1935	ABD62781 38 ABW71481 22 ABO38384 41 ABN59412 38	4	6.1	1.5	low p<.30, prev p>.50
1936	ABO38351 39	1	6.9	0.0	prev p<.30
1937	AAA67181 14	1	4.3	0.0	
1938	AAA67182 18	1	5.5	0.0	
1939	AAA67183 4 BAA00718 19	2	2.3	1.5	low p<.10, prev p<.20
1940	ABI20826 38	1	6.7	0.0	prev p<.10
1941					
1942	ABD62843 28 ABW38010 28	2	4.9	0.0	
1943	AAM76687 14 ABD79101 40 AAM76691 9 AAM76688 9 AAM76686 9 ABO38373 28 ABO38054 28	7	4.1	1.7	low p<.002, prev p<.20
1945	ABP49327 28	1	4.9	0.0	prev p<.20
1946	ABD79112 38	1	6.7	0.0	
1947	AAM76690 10 ABD77807 36 AAM76689 12 AAA67338 36 AAC53844 36 BAA96109 14 AAA67339 8 AAA67340 2 AAA67341 2 AAA67336 3 AAA67337 3 AAB39916 3 AAB39915 3 CAA67497 18 CAA67496 18 CAA67499 11 CAA67498 11 CAA67500 26 P26070 8 Q8JSD9 10 Q82571 11 Q82573 18 P03506 2 Q82570 26 CAB50889 2 CAB50888 2 CAB50887 8 CAB50886 26	28	2.7	1.5	low p<.001, prev p<.001
1948	BAA96110 14 ABN59401 37	2	5.3	1.7	low p<.40, prev p<.05
1949	ABN59434 37	1	6.5	0.0	prev p<.40
1950	ABD61735 29 ABP49316 41 P10921 3	3	5.2	2.1	low p<.20, prev p<.30
1951	BAA96112 16 BAA96111 10 ABR15808 36 ABQ44471 36 ABQ01311 38 ABP49481 36	6	5.6	1.5	low p<.05, prev p>.50
1952	BAA96113 13	1	3.8	0.0	prev p<.02
1954	ABD60966 34 BAA96114 4 ABO52280 34	3	4.4	2.8	low p<.20, prev p>.50
1955	BAA96115 6	1	1.7	0.0	prev p<.20
1956	BAA96116 9 AAF99713 18 AAF99712 16	3	2.9	0.3	low p<.001, prev p<.02
1957	ABD15258 24 AAG22555 3 BAA96117 8 ABV82573 36	4	4.3	1.6	low p<.04, prev p<.10

1961	<u>Q9WCD8 36 AAD25302 36</u>	2	6.4	0.0	prev p<.05
1963	<u>CAA40730 18</u>	1	5.0	0.0	
1966	<u>ABV82595 34</u>	1	6.0	0.0	
1967	<u>ABV82584 32</u>	1	5.7	0.0	
1970	<u>ABR28724 36</u>	1	6.4	0.0	
1971	<u>ABR28702 36</u>	1	6.4	0.0	
1972	<u>ABF21276 48 ABF21278 48 ABF21277 48 ABF21274 48</u> <u>ABF21272 48</u>	5	8.2	0.0	
1974	<u>CAA40728 18</u>	1	5.0	0.0	
1975	<u>ABU80188 41 ABR28680 39 ABR28603 39</u>	3	7.0	0.2	low p>.50, prev p<.002
1976	<u>ABS18465 32 AAF99717 24 AAF99716 30 AAF99715 27</u> <u>AAF99714 30 ABQ45533 41 ABW36366 42 ABW36322 41</u> <u>ABV45838 43 ABR28625 41 ABR28614 43 ABR15819 39</u> <u>ABQ45458 42 ABQ45447 41 ABQ45436 41 ABQ45425 42</u> <u>ABQ45414 42 ABQ44394 40 P26562 31 P03455 40 AAB52910</u> <u>16 AAD25304 31 AAB50962 18 AAB50961 16 AAB39851 40</u> <u>BAA01280 31 Q76WJ1 6 Q9IGQ0 4</u>	28	6.0	1.8	low p<.005, prev p<.005
1977	<u>ABD95350 37 ABD60944 37 ABD60933 37 ABW71492 38</u> <u>ABU80410 41 ABW36410 38 ABW36399 38 ABW36388 38</u> <u>ABW36377 38 ABV29524 41 ABU80287 38 ABU80254 41</u> <u>ABU80243 41 ABU80232 41 ABU80221 41 ABU80210 38</u> <u>ABU80199 41 ABD95712 37 ABS49921 41 ABR28647 41</u> <u>ABR28581 41 ABR28570 41 ABR28559 41 ABR28548 41</u> <u>ABR28537 43 ABR15874 41 ABR15863 41 ABR15852 41</u> <u>ABR15841 41 ABR15830 41 ABO44134 37 ABB19667 31</u> <u>ABB19529 31 ABB19518 31 P03453 37 AAD25308 30</u> <u>AAA43240 16 AAA43206 27 AAB52908 15 BAF03627 29</u> <u>P35938 3 P03469 8</u>	42	6.4	1.2	low p<.005, prev p<.20
1978	<u>ABW86585 41 ABW86574 41 ABW71503 41 ABU80265 39</u> <u>ABR28691 39 ABP49448 37 ABP49338 37 ABO38065 37</u> <u>ABO32992 37 ABO32981 39 ABN59423 37 ABK79948 37</u> <u>ABG26813 37 ABF47737 37 ABF47726 37 ABF47715 37</u> <u>ABF47704 37 ABF47693 37 AAA74287 16 AAA65552 16</u> <u>AAA65548 16</u>	21	6.4	0.8	low p<.005, prev p>.50
1979	<u>ABS18464 32 AAA43172 37 ABW36311 37 ABR28636 45</u> <u>ABQ01322 38 ABN50756 37 ABB19551 30 ABB19540 31</u> <u>P18875 37 CAA86563 15 P31348 8</u>	11	5.8	1.6	low p<.02, prev p<.20
1980	<u>AAA16879 31 AAA16880 29 ABS18466 41 AAB50965 16</u> <u>ABU80276 44 ABS49954 44 ABR28757 44 ABR28746 44</u> <u>ABR28735 44 ABR28713 39 ABO38362 38 ABO33006 37</u> <u>ABI84478 32 ABF47748 38 Q9WCE3 32 AAB52909 17</u> <u>AAD25309 35 AAD25307 32 CAA40731 20</u>	19	6.5	1.0	low p<.05, prev p<.20
1981	<u>AAB50964 16 BAA02766 5 AAZ15840 15 AAZ15839 15</u> <u>ABW36355 41 ABW36344 41 ABW36333 41 ABS49932 41</u> <u>ABR28669 41 ABR28658 41 ABO52258 38 ABI84617 31</u> <u>ABB21772 31 BAA02767 5 BAA02765 7 Q9WCE1 31</u> <u>AAB52906 22 AAD25301 44 AAD25305 31 CAA86562 14</u> <u>CAA82950 38 AAK51352 20 AAK51351 20 AAK51350 20</u> <u>AAK51349 20 AAK51348 20 AAK51347 20 AAK51346 20</u> <u>AAK51345 20 AAK51344 20 AAK51343 20 AAK51342 20</u> <u>AAK51341 20</u>	33	5.9	1.0	low p<.001, prev p<.05
1982	<u>P26142 3 ABD95339 38 ABD77818 38 AAA16905 3</u> <u>ABO52797 42 P10757 44 AAA65553 17 CAC86623 11</u>	8	5.2	2.2	low p<.05, prev p<.30
1983	<u>ABG66977 17 ABG66976 17 ABG66975 20 ABG66974 20</u> <u>ABG66973 20 ABO38340 41 ABO37988 41 ABO33025 42</u> <u>ABN50917 41 ABN50900 38 ABM66886 41 ABM66908 41</u> <u>ABM66897 41 ABM22235 41 ABM22224 41 ABM22213 41</u> <u>ABM22202 41 ABM22191 41 ABM22180 41 ABM22169 41</u>	56	6.9	0.9	low p>.50, prev p<.02

	<u>ABM22158 41 ABL67264 41 ABL67253 41 ABK80047 41</u> <u>ABK80036 41 ABK80025 41 ABK40601 41 ABK40590 41</u> <u>ABK40579 41 ABK40568 41 ABK40557 41 ABK40546 41</u> <u>ABK40534 41 ABK40510 41 ABI92302 38 ABI30378 41</u> <u>ABI20859 41 ABG88344 38 ABG88333 38 ABF47825 41</u> <u>ABF47770 41 ABG79952 41 ABF47847 41 ABF47836 41</u> <u>ABF47759 41 ABF47792 41 ABG26835 41 ABG26824 41</u> <u>ABF47814 41 ABF47803 41 ABF47781 41 BAF63173 20</u> <u>ABW91185 41 AAD25311 34 CAA35094 39 P11485 8</u>				
1984	<u>AAA43171 36 AAZ15838 15 ABP49349 39 ABO38406 42</u> <u>P18876 36 AAB27052 21 AAA65557 21 AAA65556 20</u> <u>AAA65555 20</u>	9	6.0	1.0	low p<.02, prev p<.01
1985	<u>CAA91080 27 AAB50966 13 ABW86596 41 ABR29615 41</u> <u>ABR29605 41 Q9WCE8 32 AAB52907 15 AAD25306 30</u> <u>AAD25312 32 P31349 8</u>	10	5.3	1.7	low p<.01, prev p<.30
1986	<u>AAA43236 17 BAA00309 18 BAA00308 17 ABP49360 34</u> <u>ABO44123 37 ABO38395 37 ABM22246 34 AAC57166 36</u> <u>BAF63172 16 P12590 17 AAA65547 17 BAA00722 17</u> <u>CAA35097 17</u>	13	5.5	0.7	low p<.001, prev p>.50
1987	<u>BAA96118 17 AAZ15842 16 ABU80420 41 ABS50111 41</u> <u>ABR29575 41 ABV29590 33 ABS49943 33 ABQ44416 34</u> <u>ABG88212 29 ABN50940 34 ABN50928 34 AAD25310 33</u> <u>AAA43680 17 AAA65550 17 AAA65549 17 CAA35095 44</u> <u>P05778 4</u>	17	5.9	1.0	low p<.001, prev p<.20
1988	<u>CAA91081 38 AAA43238 17 AAA43233 17 AAA43231 17</u> <u>AAA43170 16 AAA43169 17 AAA43166 17 AAA43161 18</u> <u>AAA43157 14 ABU80400 40 ABS50121 40 ABR29595 42</u> <u>ABR29585 42 ABB19607 30 P26140 39 AAB52904 18</u> <u>CAA42444 14 ABF71860 39 AAA74300 14 AAA74299 17</u> <u>AAA74298 14 AAA74285 14 AAA65551 17</u>	23	5.4	1.2	low p<.001, prev p<.20
1989	<u>AAA43168 18 AAA43158 14 AAA58800 21 AAA58801 21</u> <u>BAA06719 14 BAA96119 14 AAZ15841 20 BAA02768 31</u> <u>BAA02769 33 BAF63171 13 AAA74286 14 ABG57284 2</u> <u>ABG57283 2 ABG57282 2 ABG57281 2</u>	15	4.0	2.3	low p<.001, prev p<.02
1990	<u>AAA43235 14 AAA43234 14 AAA43232 14 AAA43190 14</u> <u>AAA43173 14 AAA43153 14 AAA91616 27 AAZ15844 16</u> <u>AAZ15843 8 ABG88201 34 AAB57740 16 AAA16778 14</u> <u>AAA16779 14 AAA16815 14 AAA16814 14 AAA16813 14</u> <u>AAA16812 13 AAA16811 14 AAA16810 14 AAA16809 14</u> <u>AAA16808 15 ABG66980 2 ABG66979 2</u>	23	3.9	1.2	low p<.001, prev p>.50
1991	<u>AAA43225 14 CAA91082 38 AAP34322 36 AAA43167 14</u> <u>AAA43283 42 BAA96120 14 ABW71521 42 ABR29565 42</u> <u>ABQ10099 16 ABD60955 33 AAA19934 37 S69887 26 S69888</u> <u>36 S69889 42 AAB50963 42 CAA91083 34 CAA86560 39</u> <u>CAA86567 17 AAA43142 14 AAA74297 14 AAA74296 14</u> <u>AAA74295 14 AAA74294 14 AAA74293 14 AAA74292 17</u> <u>AAA74291 14 AAA74290 14 AAA74289 14 AAA74288 14</u> <u>AAA65546 14 AAA65545 14 AAA65544 14 CAD29945 19</u> <u>ABG66981 2 ABG66978 2 ABE73717 2</u>	36	4.8	1.8	low p<.001, prev p<.02
1992	<u>AAB29091 41 BAA05874 14 BAA96121 14 ABB19618 30</u> <u>AAU09400 12 AAC57167 38 HMIV17 39 CAA86561 39</u> <u>AAC14275 14 AAA51481 5 AAA72339 41</u>	11	5.2	2.0	low p<.01, prev p>.50
1993	<u>ABM21960 34 ABI92181 30 AAC57169 35 AAC57168 36</u> <u>ABO52170 39 AAB50960 31 AAB50958 17 AAB50957 17</u>	8	5.9	1.0	low p<.01, prev p<.30
1994	<u>ABS70427 30 AAB03292 42 AAB03291 42 AAB50959 31</u> <u>CAD29938 14</u>	5	6.3	1.5	low p<.30, prev p>.50
1995	<u>AAK70450 37 AAK70449 37 AAP34325 38 AAP34323 20</u> <u>BAC82887 19 BAC82881 19 BAA96122 19 AAZ17358 16</u> <u>ABQ10100 19 ABP51995 1 ABG88322 36 ABG26791 37</u>	47	6.1	1.7	low p<.001, prev p>.50

	<u>ABF47638 36 ABJ53438 37 ABI92313 40 ABI30367 36</u> <u>ABI20870 31 ABI20837 36 ABG88311 36 ABG88300 36</u> <u>ABF47627 37 ABG47840 36 ABG26780 36 ABF47605 37</u> <u>ABE26991 36 ABE12032 37 ABE11942 36 ABE11922 36</u> <u>ABE11900 36 ABE11889 36 ABE11878 36 ABE11867 37</u> <u>AAL60449 36 AAL60444 36 AAL60443 36 AAK67336 17</u> <u>AAK67335 17 AAK67332 17 AAK67331 17 AAK67330 17</u> <u>ABS70438 30 CAC86625 18 CAC86619 14 CAC86617 18</u> <u>CAD29944 19 CAD29937 1 CAD29936 14</u>				
1996	<u>BAC82896 19 BAC82893 19 BAC82884 19 BAA96124 19</u> <u>BAA96123 19 AAK73345 16 AAK73344 16 AAK73343 16</u> <u>AAK73342 16 AAK73341 16 AAK73340 16 AAK73339 16</u> <u>AAK73338 16 AAK73337 16 AAK73336 16 AAK73335 16</u> <u>AAK73334 16 AAK73333 16 AAK73332 16 AAK73331 16</u> <u>AAK73330 16 AAK73329 16 AAK73328 16 AAK73327 16</u> <u>AAK73326 12 AAK73325 16 AAK73324 16 AAK73323 16</u> <u>AAK73322 16 AAK73321 16 AAK73320 16 ABO52225 37</u> <u>ABO38010 37 ABN51066 39 ABN50973 37 ABN50962 37</u> <u>ABN50951 36 ABF47649 36 ABM22290 36 ABM22279 37</u> <u>ABM22268 36 ABM22257 36 ABB19571 30 ABJ53504 31</u> <u>ABJ53493 31 ABI95283 31 ABI95272 31 ABI95261 31</u> <u>ABI95250 31 ABI93028 31 ABI21574 31 ABI21563 31</u> <u>ABI21552 31 ABI21541 31 ABI21530 29 ABI21519 31</u> <u>ABI20848 39 ABG47829 36 ABF47660 37 AAK67328 17</u> <u>AAK67327 17 AAK67326 17 AAK67325 17 AAK67324 17</u> <u>AAK67323 22 AAK67322 17 AAP60039 37 AAP60038 39</u> <u>AAP60037 42 AAP60036 37 AAF06947 16 AAF06946 13</u> <u>AAF06945 13 CAC86611 17 AAB81463 23 AAB81462 19</u> <u>AAB81461 19 AAB81460 19 AAB81459 23 AAB81458 19</u> <u>AAB81457 19 AAB81456 19 CAC86616 38 BAF03629 30</u> <u>ABD59847 36 CAD29943 18 CAD29933 14</u>	87	7.1	1.9	low p<.40, prev p<.002
1997	<u>AAD17229 29 BAA96125 14 ABG26246 19 ABG26245 19</u> <u>ABG26244 19 ABG26243 19 ABG26242 19 AAQ10369 19</u> <u>AAQ10368 16 AAQ10367 19 AAK67337 12 AAF87281 48</u> <u>AAF87280 48 AAF87279 42 AAF87278 42 AAF87277 42</u> <u>AAF87276 42 AAF87275 39 AAF87274 40 AAP79975 1</u> <u>AAP79973 1 CAC86608 42 CAC86606 33 ABD59848 34</u> <u>CAD29934 14 CAD29932 14 CAD29928 14 CAC86615 17</u>	28	5.7	2.0	low p<.001, prev p<.001
1998	<u>AAK70464 22 AAK70459 7 AAK70458 18 AAK70457 31</u> <u>AAK70456 31 AAK70455 14 AAK70454 29 AAK70453 29</u> <u>AAK70452 29 AAK70451 31 AAD17218 21 BAC82898 14</u> <u>BAC82877 14 BAC82871 17 BAA96131 14 BAA96126 14</u> <u>AAD17219 21 ABQ10144 13 ABQ10143 18 ABQ10087 16</u> <u>ABB19574 31 AAK67319 12 AAF87284 45 AAF87283 41</u> <u>AAF87282 42 AAT00438 23 CAB42465 30 AAT65329 30</u> <u>AAO88265 20 CAC86609 6 CAC86624 28 CAD29935 17</u> <u>CAD29931 14 CAD29929 13 CAD29927 14 CAD29922 14</u> <u>CAC86335 17 CAC86620 32</u>	38	5.1	1.2	low p<.001, prev p<.10
1999	<u>AAP34324 31 ABV25643 49 ABV25640 49 ABV25638 49</u> <u>ABV25637 49 ABV25636 49 ABV25635 49 ABV25634 49</u> <u>ABQ10137 15 ABL67055 30 ABL67066 30 ABJ53427 30</u> <u>ABG88256 30 BAC82894 17 BAC82892 14 BAC82885 9</u> <u>BAC82883 14 BAC82876 17 BAC82875 14 BAC82873 20</u> <u>BAC82872 14 BAA96128 14 BAA96127 14 ABO21723 19</u> <u>ABK40006 31 ABJ16609 31 AAQ10385 19 AAQ10380 19</u> <u>AAQ10373 19 AAQ10372 19 AAK67343 12 AAK67342 12</u> <u>AAK67341 12 AAK67340 12 AAK67339 12 AAK67334 17</u> <u>AAK67333 12 AAK67329 12 BAF63169 13 BAF63168 13</u> <u>BAF63167 13 BAF63166 13 AAF80098 14 AAF80099 14</u>	65	5.3	2.0	low p<.001, prev p<.40

	<p><u>CAC86337</u> <u>51 CAC86336</u> <u>19 CAC86610</u> <u>51 CAC86622</u> <u>31 CAC86605</u> <u>19 ABD59849</u> <u>28 CAD29942</u> <u>19 CAD29921</u> <u>14 CAD29917</u> <u>14 CAD29916</u> <u>14 ABV25653</u> <u>9 ABV25650</u> <u>9 ABV25648</u> <u>9 ABV25647</u> <u>9 ABV25646</u> <u>9 ABV25645</u> <u>9 ABV25644</u> <u>9 BAF63165</u> <u>13 CAC86334</u> <u>17 CAC86626</u> <u>14 CAC86621</u> <u>23</u></p>				
2000	<p><u>ABC66246</u> <u>1 ABC66232</u> <u>13 AAY42122</u> <u>15 AAY42121</u> <u>15 AAY42120</u> <u>15 AAY42119</u> <u>15 AAY42118</u> <u>15 AAY42117</u> <u>15 AAY42116</u> <u>15 AAY42115</u> <u>15 AAY42114</u> <u>15 BAC82897</u> <u>19 BAC82895</u> <u>14 BAC82891</u> <u>13 BAC82890</u> <u>14 BAC82889</u> <u>14 BAC82888</u> <u>13 BAC82886</u> <u>14 BAC82882</u> <u>13 BAC82880</u> <u>14 BAC82879</u> <u>14 BAC82878</u> <u>14 BAC82874</u> <u>19 BAC82870</u> <u>14 BAC82865</u> <u>14 AAN83988</u> <u>20 BAA96130</u> <u>19 BAA96129</u> <u>14 AAK40315</u> <u>14 AAK40318</u> <u>14 AAK40316</u> <u>14 AAK40317</u> <u>14 AAK40314</u> <u>14 AAK40313</u> <u>14 ABV45849</u> <u>31 ABU80309</u> <u>31 ABU80298</u> <u>32 ABS49987</u> <u>31 ABS49976</u> <u>31 ABR28801</u> <u>31 ABR28779</u> <u>31 ABR28768</u> <u>31 ABR14657</u> <u>1 ABR14641</u> <u>1 ABR14640</u> <u>14 ABR14639</u> <u>14 ABR15918</u> <u>31 ABR15907</u> <u>31 ABR15896</u> <u>18 ABR15885</u> <u>31 ABQ10097</u> <u>20 ABQ10095</u> <u>13 ABP49382</u> <u>37 ABP49305</u> <u>31 ABP49217</u> <u>31 ABO44046</u> <u>31 ABO21725</u> <u>19 ABO21724</u> <u>21 ABO21716</u> <u>19 ABM22026</u> <u>39 ABL67209</u> <u>37 ABL67187</u> <u>37 ABK79970</u> <u>37 ABK40050</u> <u>37 ABK40039</u> <u>37 ABK40028</u> <u>37 AAX56530</u> <u>41 ABJ53515</u> <u>37 ABJ53449</u> <u>31 ABJ16730</u> <u>37 ABJ16719</u> <u>37 ABJ16642</u> <u>37 ABJ09327</u> <u>18 ABI95294</u> <u>18 ABI95217</u> <u>31 ABG88553</u> <u>31 ABG88542</u> <u>31 ABG80183</u> <u>31 ABG80172</u> <u>31 ABG67477</u> <u>31 ABG48049</u> <u>31 ABG37362</u> <u>31 ABF47891</u> <u>31 ABF47880</u> <u>31 ABF47869</u> <u>31 ABG72870</u> <u>1 ABG72869</u> <u>1 ABG47818</u> <u>31 ABG47807</u> <u>31 ABE11668</u> <u>18 ABE11657</u> <u>31 ABD95031</u> <u>18 ABD95020</u> <u>31 ABD95009</u> <u>31 ABD94998</u> <u>31 ABD94987</u> <u>31 ABD94976</u> <u>18 ABD94965</u> <u>31 ABD94756</u> <u>31 ABD78038</u> <u>31 ABD78027</u> <u>31 ABD78016</u> <u>31 ABD78005</u> <u>31 ABD77994</u> <u>31 ABD77983</u> <u>31 ABD77972</u> <u>31 ABD77961</u> <u>31 ABD77950</u> <u>31 ABD77939</u> <u>31 ABD77928</u> <u>33 ABD77917</u> <u>31 ABD77730</u> <u>31 ABD77719</u> <u>31 ABD77708</u> <u>31 ABD63063</u> <u>31 ABD61540</u> <u>31 ABD61518</u> <u>31 ABD60900</u> <u>31 ABD60889</u> <u>31 ABD60878</u> <u>31 ABD60867</u> <u>31 ABD60856</u> <u>31 ABA08497</u> <u>18 ABA08486</u> <u>18 AAQ10391</u> <u>2 AAK67344</u> <u>12 AAK67338</u> <u>12 AAK67321</u> <u>12 AAK67320</u> <u>12 AAL15459</u> <u>27 CAC86333</u> <u>28 CAC86607</u> <u>28 CAC86612</u> <u>28 CAD29941</u> <u>19 CAD29940</u> <u>19 CAD29939</u> <u>24 CAD29930</u> <u>1 CAD29926</u> <u>1 CAD29924</u> <u>1 CAD29920</u> <u>16 CAD29919</u> <u>14 CAD29899</u> <u>14 CAD57622</u> <u>6 CAC86618</u> <u>31 CAC86614</u> <u>17 CAC86613</u> <u>17 CAC18525</u> <u>12</u></p>	147	5.0	1.6	low p<.001, prev p<.10
2001	<p><u>AAP79964</u> <u>12 ABC66233</u> <u>14 BAC82869</u> <u>1 BAC82868</u> <u>1 BAC82867</u> <u>1 BAC82866</u> <u>1 BAC82864</u> <u>1 BAC82863</u> <u>1 BAC82862</u> <u>14 BAC82861</u> <u>1 BAC82860</u> <u>13 BAC82859</u> <u>14 BAC82858</u> <u>14 BAC82857</u> <u>14 BAC82846</u> <u>1 BAC82843</u> <u>1 ABI55088</u> <u>1 AAZ17359</u> <u>25 AAY56898</u> <u>32 ABR28845</u> <u>18 ABR28834</u> <u>18 ABR14668</u> <u>14 ABR14667</u> <u>14 ABR14666</u> <u>14 ABR14665</u> <u>14 ABR14664</u> <u>14 ABR14663</u> <u>13 ABR14662</u> <u>14 ABR14661</u> <u>14 ABR14660</u> <u>14 ABR14659</u> <u>14 ABR14658</u> <u>14 ABR14656</u> <u>1 ABR14655</u> <u>1 ABR14654</u> <u>1 ABR14653</u> <u>1 ABR14652</u> <u>1 ABR14651</u> <u>1 ABR14650</u> <u>1 ABR14649</u> <u>14 ABR14648</u> <u>14 ABR14647</u> <u>14 ABR14646</u> <u>14 ABR14645</u> <u>14 ABR14644</u> <u>14 ABR14643</u> <u>14 ABR14642</u> <u>14 BAF63035</u> <u>13 BAF63032</u> <u>13 ABQ10092</u> <u>1 ABQ10091</u> <u>1 ABQ10090</u> <u>1 ABQ10089</u> <u>14 ABO38329</u> <u>18 ABO38318</u> <u>18 ABO38043</u> <u>18 ABO38032</u> <u>31 ABO38021</u> <u>31 ABO32959</u> <u>18 ABO32948</u> <u>18 ABN51143</u> <u>18 ABN51077</u> <u>18 ABM66864</u> <u>18 ABJ09151</u> <u>19</u></p>	193	3.2	1.6	low p<.001, prev p<.001

	<p><u>ABG67491</u> 18 <u>ABG37395</u> 18 <u>ABG37384</u> 18 <u>ABG26945</u> 18 <u>ABF82940</u> 18 <u>ABF82929</u> 18 <u>ABF82918</u> 18 <u>ABF82907</u> 18 <u>ABF82896</u> 18 <u>ABF82885</u> 18 <u>ABF82874</u> 18 <u>ABF82863</u> 18 <u>ABF82852</u> 18 <u>ABF82841</u> 18 <u>ABF82830</u> 18 <u>ABF82819</u> 31 <u>ABG72867</u> 1 <u>ABG72866</u> 1 <u>ABF47671</u> 18 <u>ABF47572</u> 18 <u>ABF47561</u> 29 <u>ABG37120</u> 18 <u>ABF82684</u> 18 <u>ABF82673</u> 18 <u>ABF82662</u> 18 <u>ABF47583</u> 18 <u>ABE12248</u> 18 <u>ABE11856</u> 18 <u>ABE11845</u> 18 <u>ABE11834</u> 18 <u>ABE11823</u> 18 <u>ABE11812</u> 18 <u>ABE11734</u> 18 <u>ABE11723</u> 18 <u>ABE11712</u> 18 <u>ABE11701</u> 18 <u>ABE11690</u> 18 <u>ABE11679</u> 26 <u>ABD95328</u> 18 <u>ABD95317</u> 18 <u>ABD95306</u> 18 <u>ABD95295</u> 18 <u>ABD95284</u> 18 <u>ABD95273</u> 18 <u>ABD95262</u> 31 <u>ABD95251</u> 18 <u>ABD95240</u> 18 <u>ABD95229</u> 18 <u>ABD95218</u> 18 <u>ABD95207</u> 31 <u>ABD95196</u> 18 <u>ABD95185</u> 18 <u>ABD95174</u> 18 <u>ABD95163</u> 18 <u>ABD95152</u> 18 <u>ABD95141</u> 18 <u>ABD95130</u> 18 <u>ABD95119</u> 18 <u>ABD95108</u> 18 <u>ABD95097</u> 18 <u>ABD95086</u> 18 <u>ABD95075</u> 18 <u>ABD95064</u> 18 <u>ABD95053</u> 18 <u>ABD95042</u> 18 <u>ABD94811</u> 18 <u>ABD94800</u> 18 <u>ABD94789</u> 18 <u>ABD94778</u> 18 <u>ABD78093</u> 18 <u>ABD78082</u> 18 <u>ABD78071</u> 18 <u>ABD78060</u> 18 <u>ABD62061</u> 31 <u>ABD60911</u> 31 <u>ABC86237</u> 36 <u>ABC40533</u> 18 <u>ABB02814</u> 18 <u>ABA87231</u> 18 <u>ABA87091</u> 18 <u>ABC02277</u> 18 <u>ABB82194</u> 18 <u>ABB80045</u> 18 <u>ABB79990</u> 18 <u>ABB79979</u> 18 <u>ABB53707</u> 18 <u>ABB02936</u> 18 <u>ABB02924</u> 31 <u>ABB02913</u> 18 <u>ABB02825</u> 18 <u>ABA87045</u> 18 <u>ABA43189</u> 18 <u>ABA42575</u> 18 <u>ABA42324</u> 31 <u>ABA42258</u> 18 <u>ABA42236</u> 18 <u>ABA18037</u> 18 <u>ABA12715</u> 18 <u>ABA08519</u> 31 <u>ABA08464</u> 18 <u>AAZ85126</u> 18 <u>AAZ83299</u> 18 <u>AAZ79604</u> 18 <u>AAZ38627</u> 18 <u>AAT85679</u> 18 <u>AAU25871</u> 12 <u>AAU25861</u> 12 <u>AAL47667</u> 14 <u>AAL47668</u> 19 <u>CAD29925</u> 1 <u>CAD29923</u> 1 <u>CAD29918</u> 14 <u>CAD29911</u> 1 <u>CAD29910</u> 1 <u>CAD29909</u> 14 <u>CAD29908</u> 14 <u>CAD29907</u> 1 <u>CAD29906</u> 1 <u>CAD29905</u> 19 <u>CAD29904</u> 1 <u>CAD29903</u> 14 <u>CAD29902</u> 19 <u>CAD29901</u> 14 <u>CAD29900</u> 1 <u>CAD57623</u> 33 <u>CAD57621</u> 6 <u>CAD57620</u> 27 <u>CAD57619</u> 28 <u>CAD57617</u> 1</p>				
2002	<p><u>AAU25851</u> 29 <u>BAC82856</u> 1 <u>BAC82855</u> 1 <u>BAC82854</u> 1 <u>BAC82853</u> 1 <u>BAC82852</u> 1 <u>BAC82851</u> 1 <u>BAC82850</u> 1 <u>BAC82849</u> 1 <u>BAC82848</u> 1 <u>BAC82847</u> 1 <u>BAC82845</u> 1 <u>BAC82844</u> 1 <u>BAC82842</u> 19 <u>AAP69688</u> 1 <u>AAP69687</u> 1 <u>AAP69686</u> 1 <u>AAP69685</u> 1 <u>AAP69684</u> 1 <u>AAP69683</u> 1 <u>AAP69682</u> 1 <u>AAP69681</u> 1 <u>AAP69680</u> 12 <u>AAP69679</u> 1 <u>AAP69678</u> 3 <u>AAP69677</u> 1 <u>AAP69676</u> 1 <u>AAP69675</u> 1 <u>AAP69674</u> 1 <u>AAP69673</u> 1 <u>AAP69692</u> 1 <u>AAP69691</u> 1 <u>AAP69690</u> 1 <u>AAP69689</u> 1 <u>ABS76427</u> 18 <u>BAF63057</u> 12 <u>BAF63054</u> 12 <u>BAF63050</u> 12 <u>BAF63046</u> 17 <u>BAF63043</u> 13 <u>BAF63039</u> 13 <u>ABB19628</u> 31 <u>ABG72868</u> 1 <u>ABA87080</u> 31 <u>ABB82216</u> 31 <u>ABB51962</u> 17 <u>AAZ83253</u> 28 <u>AAT12706</u> 29 <u>ABS70411</u> 30 <u>ABS70400</u> 30 <u>ABS70389</u> 30 <u>ABS70378</u> 30 <u>ABS70367</u> 30 <u>ABS70337</u> 30 <u>ABS70326</u> 30 <u>ABS70316</u> 30 <u>ABS70305</u> 30 <u>CAD29958</u> 14 <u>CAD29914</u> 14 <u>CAD57618</u> 28 <u>CAD57616</u> 14 <u>CAD35678</u> 19</p>	62	2.6	2.5	low p<.001, prev p<.05
2003	<p><u>ABB86907</u> 43 <u>ABB86887</u> 37 <u>ABB86877</u> 37 <u>BAF63102</u> 13 <u>BAF63101</u> 13 <u>BAF63097</u> 13 <u>BAF63096</u> 13 <u>BAF63095</u> 13 <u>BAF63091</u> 13 <u>BAF63090</u> 13 <u>BAF63089</u> 13 <u>BAF63085</u> 13 <u>BAF63084</u> 13 <u>BAF63083</u> 13 <u>BAF63082</u> 13 <u>BAF63078</u> 13 <u>BAF63075</u> 13 <u>BAF63072</u> 13 <u>BAF63068</u> 13 <u>BAF63064</u> 13 <u>BAF63061</u> 13 <u>ABQ10105</u> 16 <u>ABQ10104</u> 14 <u>ABQ10103</u> 14 <u>ABQ10102</u> 14 <u>ABO37999</u> 31 <u>ABN51088</u> 31 <u>ABM67051</u> 31 <u>ABI96088</u> 14 <u>ABI96127</u> 14 <u>ABI96126</u> 14 <u>ABI96125</u> 14 <u>ABI96124</u> 11 <u>ABI96123</u> 14 <u>ABI96122</u> 14 <u>ABI96121</u> 14 <u>ABI96120</u> 14 <u>ABI96119</u> 16 <u>ABI96118</u> 17 <u>ABI96117</u> 14</p>	94	4.9	1.1	low p<.001, prev p<.001

	<u>ABI96116 14 ABI96115 14 ABI96114 14 ABI96113 14</u> <u>ABI96112 11 ABI96111 14 ABI96110 14 ABI96109 31</u> <u>ABI96108 30 ABI96107 31 ABI96105 14 ABI96103 31</u> <u>ABI96100 14 ABI96098 14 ABI96097 14 ABI96096 14</u> <u>ABI96095 14 ABI96094 14 ABI96093 14 ABI96090 14</u> <u>ABE12634 57 ABD78104 51 ABD60779 28 ABD15515 28</u> <u>ABC41714 31 ABB03123 31 ABA87057 31 AAZ83977 31</u> <u>ABB82205 36 ABB80103 31 ABB53740 31 ABB03145 31</u> <u>ABB02803 31 ABB02792 28 ABA42247 36 ABA18145 31</u> <u>ABA12729 28 ABA12707 31 ABA12696 31 ABA08475 36</u> <u>ABK39995 31 ABB86917 31 IRUZ M 19 IRUZ L 19</u> <u>IRUZ K 19 IRUZ J 19 IRUZ I 19 IRUZ H 19 IRUY M 15</u> <u>IRUY L 15 IRUY K 15 IRUY J 15 IRUY I 15 IRUY H 15</u>				
2004	<u>ABB86946 40 ABB86937 44 ABS00326 34 ABQ10005 26</u> <u>ABQ10004 20 ABQ10003 14 ABQ10002 14 ABQ10001 26</u> <u>ABQ09988 14 ABQ09838 22 ABQ09837 28 ABQ09784 14</u> <u>ABQ09783 29 ABQ09782 29 ABQ09780 27 ABQ09779 29</u> <u>ABI96132 14 ABI96130 14 ABI96129 14 ABI54447 13</u> <u>ABI54446 13 ABI54445 13 ABI54444 13 ABI54443 13</u> <u>ABI54442 13 ABI54441 13 ABI54440 13 ABI54439 13</u> <u>ABI54438 13 ABI54437 13 ABE27153 46 ABC42750 46</u> <u>ABB86929 31 ABB86899 31 AAV68006 15 AAV67984 13</u>	36	5.3	1.4	low p<.001, prev p<.10
2005	<u>ABJ51893 13 ABI51313 13 ABO52104 30 ABC66245 14</u> <u>ABC66244 14 ABC66243 14 ABC66242 14 ABC66241 14</u> <u>ABC66240 14 ABC66239 14 ABC66238 14 ABC66237 14</u> <u>ABC66236 14 ABC66235 14 ABC66234 14 ABW75642 18</u> <u>ABI19015 41 ABK57093 35 ABR28900 31 ABJ51895 13</u> <u>ABJ51894 13 ABJ51892 13 ABJ51891 13 ABJ51890 13</u> <u>BAF63115 13 BAF63111 13 BAF63107 13 BAF63103 13</u> <u>ABQ09953 14 ABQ09950 27 ABQ09949 14 ABQ09948 13</u> <u>ABQ09947 14 ABQ09915 14 ABQ09914 14 ABQ09913 14</u> <u>ABQ09904 14 ABQ09874 14 ABQ09873 14 ABQ09872 14</u> <u>ABQ09871 16 ABP51970 14 ABP49393 31 ABO32970 31</u> <u>ABO32678 37 ABO21731 14 ABO21730 14 ABK40689 31</u> <u>ABJ16686 31 ABJ16675 31 ABJ16664 31 ABJ16653 31</u> <u>ABJ09184 31 ABI96148 14 ABI96147 32 ABI96146 14</u> <u>ABI96145 27 ABI96144 13 ABI96143 29 ABI96142 14</u> <u>ABI96141 14 ABI96140 14 ABI96139 14 ABI96138 14</u> <u>ABI96137 14 ABI96135 14 ABI96134 14 ABI96128 16</u> <u>ABI92379 31 ABI30565 31 ABI22109 31 ABI21233 31</u> <u>ABI21222 31 ABI21211 31 ABI21200 31 ABI21189 31</u> <u>ABG72865 1 ABG72864 1 ABG72863 1 BAE53730 14</u> <u>BAE53729 14 ABB84190 14</u>	82	4.9	1.5	low p<.001, prev p<.10
2006	<u>BAE96542 25 BAE96541 14 BAE96540 14 BAE96539 25</u> <u>BAE96538 14 BAE96537 14 BAE96536 25 BAE96535 14</u> <u>BAE96534 14 BAE96533 14 ABW71294 31 ABW23335 40</u> <u>ABW23328 31 ABW23327 14 ABW23324 27 ABW23321 14</u> <u>ABW23320 31 ABW23318 14 ABW23317 27 ABW23314 30</u> <u>ABW23311 14 ABW23310 14 ABW23308 13 ABW23307 14</u> <u>ABW23306 14 ABW23303 14 ABW23300 14 ABW23296 31</u> <u>ABW23295 31 ABW23294 29 ABW23291 14 ABW23290 31</u> <u>ABW23289 17 ABW23287 14 ABW23286 31 ABW23285 13</u> <u>ABW23284 14 ABW23283 1 ABW23280 14 ABW23279 14</u> <u>ABW23275 14 ABV45654 31 ABV29557 31 ABV29546 31</u> <u>ABV29535 18 ABU99109 44 ABU99069 25 ABU99067 27</u> <u>ABU50589 44 ABU50588 44 ABU50587 44 ABU50586 42</u> <u>ABU50574 18 ABU50573 31 ABU50571 31 ABU50570 31</u> <u>ABU50569 31 ABU50568 31 ABU50567 31 ABU50566 18</u> <u>ABU50565 31 ABU50556 11 ABU50555 26 ABU50554 14</u>	135	5.2	1.7	low p<.001, prev p<.20

	<p><u>ABU50553</u> 14 <u>ABU50552</u> 26 <u>ABU50551</u> 23 <u>ABU50550</u> 23 <u>ABU50549</u> 14 <u>ABU50546</u> 17 <u>ABU50545</u> 28 <u>ABU50544</u> 13 <u>ABU50543</u> 14 <u>ABU50542</u> 1 <u>ABU50541</u> 13 <u>ABU50540</u> 1 <u>ABU50539</u> 14 <u>ABU50538</u> 23 <u>ABU50537</u> 14 <u>ABU50536</u> 1 <u>ABU50535</u> 14 <u>ABU50534</u> 31 <u>ABU50518</u> 27 <u>ABU50517</u> 23 <u>ABS00315</u> 31 <u>BAF63135</u> 14 <u>BAF63131</u> 14 <u>BAF63127</u> 25 <u>BAF63123</u> 18 <u>BAF63119</u> 14 <u>ABQ09984</u> 25 <u>ABQ09981</u> 25 <u>ABQ09977</u> 31 <u>ABQ09976</u> 25 <u>ABQ09975</u> 25 <u>ABQ09974</u> 27 <u>ABQ09973</u> 25 <u>ABQ09972</u> 25 <u>ABQ09969</u> 27 <u>ABQ09968</u> 25 <u>ABQ09961</u> 29 <u>ABQ09960</u> 29 <u>ABQ09958</u> 29 <u>ABQ09957</u> 26 <u>ABQ09956</u> 25 <u>ABQ09955</u> 19 <u>ABK79959</u> 31 <u>ABI96174</u> 27 <u>ABI96173</u> 14 <u>ABI96172</u> 14 <u>ABI96171</u> 14 <u>ABI96170</u> 14 <u>ABI96169</u> 14 <u>ABI96168</u> 14 <u>ABI96167</u> 14 <u>ABI96166</u> 25 <u>ABI96165</u> 14 <u>ABI96164</u> 14 <u>ABI96163</u> 14 <u>ABI96162</u> 19 <u>ABI96161</u> 14 <u>ABI96160</u> 14 <u>ABI96159</u> 14 <u>ABI96158</u> 14 <u>ABI96157</u> 14 <u>ABI96156</u> 14 <u>ABI96155</u> 14 <u>ABI96154</u> 14 <u>ABI96153</u> 27 <u>ABI96152</u> 27 <u>ABI96151</u> 14 <u>ABI96150</u> 14 <u>ABI96149</u> 27 <u>ABH07371</u> 29 <u>ABH07372</u> 29</p>				
<p>2007</p>	<p><u>ABS71673</u> 46 <u>ABS71672</u> 46 <u>ABS71671</u> 46 <u>ABS71670</u> 46 <u>ABS71669</u> 46 <u>ABS71668</u> 46 <u>ABS71667</u> 46 <u>ABS71666</u> 46 <u>ABS71665</u> 46 <u>ABS71664</u> 46 <u>ABQ52695</u> 35 <u>ABW34451</u> 23 <u>ABW86606</u> 31 <u>ABW86552</u> 31 <u>ABW86541</u> 31 <u>ABW86530</u> 31 <u>ABW86519</u> 31 <u>ABW86508</u> 31 <u>ABW86497</u> 18 <u>ABW86486</u> 31 <u>ABW86475</u> 31 <u>ABW86464</u> 31 <u>ABW86453</u> 31 <u>ABW86442</u> 31 <u>ABW86431</u> 31 <u>ABW86420</u> 31 <u>ABW86409</u> 31 <u>ABW86398</u> 18 <u>ABW86387</u> 31 <u>ABW86376</u> 31 <u>ABW86365</u> 31 <u>ABW86354</u> 31 <u>ABW86343</u> 31 <u>ABW86332</u> 30 <u>ABW86321</u> 18 <u>ABW71470</u> 31 <u>ABW71459</u> 31 <u>ABW71448</u> 31 <u>ABW71437</u> 31 <u>ABW71426</u> 31 <u>ABW71415</u> 31 <u>ABW71404</u> 31 <u>ABW71393</u> 31 <u>ABW71382</u> 31 <u>ABW71371</u> 31 <u>ABW71360</u> 18 <u>ABW71338</u> 31 <u>ABW71327</u> 31 <u>ABW71316</u> 31 <u>ABW71305</u> 18 <u>ABW40675</u> 31 <u>ABW40664</u> 18 <u>ABW40642</u> 31 <u>ABW40620</u> 31 <u>ABW40609</u> 31 <u>ABW40598</u> 18 <u>ABW40576</u> 31 <u>ABW40565</u> 31 <u>ABW40554</u> 30 <u>ABW40543</u> 31 <u>ABW40532</u> 31 <u>ABW40521</u> 31 <u>ABW40510</u> 18 <u>ABW40499</u> 31 <u>ABW40488</u> 31 <u>ABW40477</u> 31 <u>ABW40466</u> 31 <u>ABW40455</u> 31 <u>ABW40444</u> 31 <u>ABW40433</u> 18 <u>ABW40422</u> 31 <u>ABW40411</u> 31 <u>ABW40400</u> 31 <u>ABW40389</u> 31 <u>ABW40367</u> 31 <u>ABW40356</u> 31 <u>ABW40345</u> 18 <u>ABW40334</u> 18 <u>ABW40312</u> 30 <u>ABW40301</u> 30 <u>ABW40290</u> 31 <u>ABW40279</u> 31 <u>ABW40257</u> 31 <u>ABW40235</u> 31 <u>ABW40224</u> 31 <u>ABW40213</u> 31 <u>ABW40202</u> 31 <u>ABW40180</u> 31 <u>ABW40158</u> 31 <u>ABW40147</u> 31 <u>ABW40125</u> 31 <u>ABW40114</u> 31 <u>ABW40103</u> 31 <u>ABW40092</u> 33 <u>ABW40070</u> 43 <u>ABW40059</u> 31 <u>ABW40048</u> 33 <u>ABW40037</u> 31 <u>ABW40015</u> 18 <u>ABW40004</u> 31 <u>ABW39993</u> 31 <u>ABW39982</u> 31 <u>ABW39971</u> 42 <u>ABW39960</u> 31 <u>ABW39949</u> 31 <u>ABW39927</u> 31 <u>ABW39916</u> 40 <u>ABW39905</u> 31 <u>ABW39894</u> 31 <u>ABW39883</u> 33 <u>ABW39861</u> 31 <u>ABW39850</u> 40 <u>ABW39839</u> 27 <u>ABW39828</u> 46 <u>ABW39817</u> 30 <u>ABW39806</u> 31 <u>ABW39777</u> 31 <u>ABW36300</u> 30 <u>ABW36289</u> 31 <u>ABW36278</u> 31 <u>ABW36267</u> 31 <u>ABW36256</u> 43 <u>ABW36245</u> 31 <u>ABW36234</u> 31 <u>ABW36223</u> 31 <u>ABW36212</u> 31 <u>ABW36201</u> 31 <u>ABW36190</u> 31 <u>ABW36179</u> 31 <u>ABW23343</u> 33 <u>ABW23342</u> 44 <u>ABW23341</u> 36 <u>ABW23340</u> 31 <u>ABW23339</u> 30 <u>ABW23338</u> 31 <u>ABW23337</u> 31 <u>ABW23336</u> 36 <u>ABW23329</u> 14 <u>ABW23326</u> 14 <u>ABW23325</u> 31 <u>ABW23323</u> 31 <u>ABW23322</u> 27 <u>ABW23319</u> 31 <u>ABW23316</u> 14 <u>ABW23315</u> 40 <u>ABW23313</u> 31 <u>ABW23312</u> 46 <u>ABW23309</u> 14 <u>ABW23305</u> 31 <u>ABW23304</u> 31 <u>ABW23302</u> 31 <u>ABW23301</u> 14 <u>ABW23299</u> 46 <u>ABW23298</u> 13 <u>ABW23297</u> 1 <u>ABW23293</u> 31 <u>ABW23292</u> 18 <u>ABW23288</u> 31 <u>ABW23282</u> 31 <u>ABW23281</u> 31 <u>ABW23278</u> 14 <u>ABW23277</u> 31 <u>ABW23276</u> 38 <u>ABW23274</u> 1</p>	<p>285</p>	<p>5.5</p>	<p>1.2</p>	<p>low p<.001, prev p<.02</p>

<u>ABW23273</u> 31 <u>ABV82551</u> 30 <u>ABV45959</u> 31 <u>ABV45948</u> 31				
<u>ABV45937</u> 31 <u>ABV45926</u> 30 <u>ABV45915</u> 31 <u>ABV45893</u> 30				
<u>ABV45882</u> 31 <u>ABV45871</u> 31 <u>ABV30624</u> 31 <u>ABV30613</u> 32				
<u>ABV30602</u> 31 <u>ABV30591</u> 31 <u>ABV30580</u> 31 <u>ABV30569</u> 31				
<u>ABV30558</u> 31 <u>ABV30547</u> 30 <u>ABV30536</u> 31 <u>ABV30525</u> 31				
<u>ABV30503</u> 31 <u>ABV30492</u> 31 <u>ABV30459</u> 31 <u>ABV30371</u> 31				
<u>ABV30360</u> 31 <u>ABV30349</u> 31 <u>ABV30338</u> 31 <u>ABV30327</u> 18				
<u>ABV30316</u> 31 <u>ABV30305</u> 31 <u>ABV30294</u> 18 <u>ABV30283</u> 31				
<u>ABV30195</u> 31 <u>ABV30184</u> 31 <u>ABV30173</u> 31 <u>ABV30162</u> 31				
<u>ABV30151</u> 31 <u>ABV30140</u> 21 <u>ABV30129</u> 31 <u>ABV30107</u> 31				
<u>ABV30096</u> 31 <u>ABV30085</u> 31 <u>ABV30052</u> 31 <u>ABV30041</u> 31				
<u>ABV30030</u> 31 <u>ABV30019</u> 31 <u>ABV30008</u> 31 <u>ABV29997</u> 31				
<u>ABV29986</u> 31 <u>ABV29975</u> 31 <u>ABV29964</u> 31 <u>ABV29953</u> 31				
<u>ABV29942</u> 30 <u>ABV29920</u> 31 <u>ABV29887</u> 31 <u>ABV29876</u> 31				
<u>ABV29865</u> 31 <u>ABV29854</u> 31 <u>ABV29843</u> 31 <u>ABV29832</u> 31				
<u>ABV29799</u> 31 <u>ABV29788</u> 31 <u>ABV29777</u> 30 <u>ABV29766</u> 31				
<u>ABV29755</u> 31 <u>ABV29744</u> 31 <u>ABV29733</u> 30 <u>ABV29700</u> 31				
<u>ABV29689</u> 31 <u>ABV29678</u> 31 <u>ABV29667</u> 31 <u>ABV29656</u> 31				
<u>ABV29645</u> 31 <u>ABV29634</u> 31 <u>ABV29612</u> 31 <u>ABV29601</u> 30				
<u>ABV29579</u> 35 <u>ABV29568</u> 31 <u>ABU50572</u> 31 <u>ABS71683</u> 46				
<u>ABS71682</u> 46 <u>ABS71681</u> 46 <u>ABS71680</u> 46 <u>ABS71679</u> 46				
<u>ABS71678</u> 46 <u>ABS71677</u> 46 <u>ABS71676</u> 46 <u>ABS71675</u> 46				
<u>ABS71674</u> 46 <u>ABW91636</u> 31 <u>ABW91625</u> 31 <u>ABW91614</u> 31				
<u>ABW91603</u> 18 <u>ABW91592</u> 31 <u>ABW91581</u> 31 <u>ABW91570</u> 30				
<u>ABW91559</u> 31 <u>ABW91537</u> 31 <u>ABW91526</u> 30 <u>ABW91515</u> 31				
<u>ABW91504</u> 31 <u>ABW91493</u> 35 <u>ABW91482</u> 18 <u>ABW91471</u> 31				
<u>ABW91460</u> 18 <u>ABW91449</u> 18 <u>ABW91427</u> 31 <u>ABW91416</u> 31				
<u>ABW91405</u> 18 <u>ABW91383</u> 31 <u>ABW91372</u> 31 <u>ABW91361</u> 30				
<u>ABW91350</u> 31 <u>ABW91339</u> 31 <u>ABW91328</u> 28 <u>ABW91317</u> 31				
<u>ABW91306</u> 31 <u>ABW91295</u> 31 <u>ABW91284</u> 42 <u>ABW91273</u> 31				
<u>ABW91218</u> 31				

Example 14

Replikin Count Analysis in Equine Influenza Virus H3N8

[000438] Applicants analyzed publicly available sequences for isolates of EIV from PubMed using proprietary search tool software (ReplikinForecast™ available in the United States from REPLIKINS LLC, Boston, MA). The data is contained in Table 22, below, and Table 4, above, and graphically described in Figure 7.

[000439] Table 22 provides the data for Replikin concentration for publicly available sequences of the pB1 gene area of the H3N8 strain of influenza virus from 1963 to 2005. Sequences were publicly available under accession numbers at www.pubmed.com. Standard deviation and significance as compared to the mean Replikin Count of the previous year and of the lowest mean Replikin Count within the data set are also provided along with the mean Replikin Count for each year. Where data was not available in a given year the year is not presented in the table.

Table 22 –H3N8 pB1

Year	PubMed Accession Number-Replikin Count	No. of Isolates per year	Mean Replikin Count per year	S.D.	Significance
1963	ABB88376 14	1	1.8	0.0	
1972	ABI84585 14	1	1.8	0.0	
1977	ABB19675 15	1	16.7	0.0	
1978	ABB20350 15 ABB88317 18	2	11.0	12.7	low p<0.40, prev p>0.50
1979	ABB87407 20 ABB86793 20	2	22.2	0.0	prev p<0.30
1980	ABB87797 20 ABB20412 14 BAF32965 14 Q08II5 14 Q08II4 5	5	6.7	8.8	low p<0.20, prev p<0.01
1982	ABI84945 16	1	17.8	0.0	prev p<0.04
1985	ABB19731 20 ABB19720 17	2	20.6	2.4	low p<0.05, prev p<0.30
1986	ABO52134 17 ABJ09104 25 ABP49539 14 P16505 17 AAA43638 17	5	13.3	11.1	low p<0.05, prev p<0.20
1987	ABL67228 14 ABL67849 15 ABM66861 20 ABI95482 15	4	19.1	2.7	low p<0.001, prev p<0.30
1991	ABM21946 15 ABI84419 15	2	9.3	10.4	low p<0.40, prev p<0.20
1992	ABB88191 14	1	15.6	0.0	prev p>0.50
1998	AAT65279 6 AAT65275 6 AAT65264 6 AAT65263 6 AAT65251 6 AAT65249 6	6	2.2	0.0	low p<0.001, prev p<0.001
1999	AAZ23576 15	1	2.0	0.0	prev p<0.001
2001	ABB19766 16 AAN15147 5	2	9.9	11.2	low p<0.40, prev p>0.50
2002	AAZ23577 15 ABI47978 14 ABO51892 14 ABO51870 14 AAX23573 15	5	4.6	6.1	low p<0.30, prev p>0.50
2003	AAZ23578 15 AAZ23575 15 AAZ23574 15 AAZ23573 15 AAZ23572 15 ABB17181 15	6	2.0	0.0	low p<0.001, prev p<0.30
2004	AAZ23571 15 ABJ53167 5 ABD27777 6 ABD27776 6 ABD27775 6	5	2.2	0.2	low p<0.01, prev p<0.02
2005	ABS89395 16 ABR37470 16 ABO52651 16 ABO76921 16 ABO52684 16 ABO52640 16 ABL67140 16 ABL67118 16 ABK79945 16 ABJ09126 16 ABI92277 16 ABI92266 16 ABO52101 25	13	18.5	2.8	low p<0.001, prev p<0.001

[000440] As may be seen in Table 22 over the 42 year period for which sequence information is publicly available for H3N8 isolates, the cyclic nature of changes in Replikin concentration becomes evident. Where Replikin concentrations reach a high within the cycle, an epidemic occurs within about 1 to about 2 years. For example, a high of 22.2 Replikin sequences per 100 amino acids in 1979 falls to 2 Replikin sequences per 100 amino acids in 1998 and 1999 with no epidemics reported between 1995 and 2001. The Replikin concentration then appears to be on its way up again in 2001 with epidemics following in the United Kingdom and in Germany in 2002 and 2003, respectively, and then falls back to around 2 in 2003 and 2004 with a marked increase in 2005 to 18.5 that approaches the highs of 1979. Epidemics follow the 2005 increase in Australia, Italy and Japan in 2007.

[000441] Table 23 provides the data for Replikin concentration for publicly available sequences of the pB2 gene of the H3N8 strain of influenza virus from 1963 to 2005. Sequences were available under accession numbers at www.pubmed.com. Standard

deviation and significance as compared to the mean Replikin Count of the previous year and of the lowest mean Replikin Count within the data set are also provided along with the mean Replikin Count for each year

Table 23 –H3N8 pB2

Year	PubMed Accession Number-Replikin Count	No. of Isolates per year	Mean Replikin Count per year	S.D.	Significance
1963	ABB88378 18	1	2.4	0.0	
1972	ABI84587 18	1	2.4	0.0	
1977	ABB19677 18	1	2.4	0.0	
1978	ABB20352 18 ABB88319 18	2	2.4	0.0	
1979	ABB87409 18 ABB86795 18	2	2.4	0.0	
1980	ABI84488 14 ABB87799 19 BAF32964 18 Q08II6 18	4	2.3	0.3	low p>0.50, prev p>0.50
1981					
1982	ABI84947 18	1	2.4	0.0	prev p>0.50
1985	ABB19733 18 ABB19722 18	2	2.4	0.0	
1986	ABO52136 18 ABJ09106 18 ABP49541 17 AAA43133 18 P26105 18	5	2.3	0.1	low p<0.30, prev p<0.30
1987	ABL67230 18 ABL67851 18 ABM66863 18 ABI95484 18	4	2.4	0.0	prev p<0.30
1991	ABM21948 18 ABI84421 18	2	2.4	0.0	
1992	ABB88193 18	1	2.4	0.0	
1998	AAT65244 18 AAT65240 18 AAT65229 18 AAT65228 18 AAT65216 18 AAT65214 18	6	3.8	0.0	
1999	AAZ23564 1	1	0.2	0.0	
2000	AAG10729 1	1	0.6	0.0	
2001	ABB19768 18	1	2.4	0.0	
2002	AAZ23563 18 ABI47980 18 ABO51894 18 ABO51872 18 AAX23572 18	5	2.4	0.1	low p<0.30, prev p<0.30
2003	AAZ23567 18 AAZ23566 18 AAZ23565 18 AAZ23562 18 AAZ23561 18 ABB17182 18	6	2.4	0.0	low p<0.30, prev p>0.50
2004	AAZ23560 18	1	2.4	0.0	prev p<0.30
2005	ABS89397 18 ABR37472 18 ABO52653 18 ABO76923 16 ABO52686 18 ABO52642 18 ABL67142 18 ABL67120 18 ABK79947 18 ABJ09128 18 ABI92279 18 ABI92268 16 ABO52103 18	13	2.3	0.1	low p<0.10, prev p<0.10

[000442] A review of the Replikin concentrations of available sequences for the pB2 gene area of the H3N8 strain of influenza virus reveals much less variability in the Replikin concentration through the years. The pB2 Replikin concentration can be considered control data that validate the location of the most significant Replikin Peak Gene for the present isolates of virus in the pB1 gene area. Because the pB2 gene is right next to the pB1 gene, the difference in variability in Replikin Count between these neighboring areas is remarkable.

[000443] Data from a review of the Replikin Count of available sequences for the pA gene area of the H3N8 strain of influenza virus may be seen in Figure 7. The data also reveal much less variability in the Replikin concentration through the years as compared to the pB1 gene area. As with the pB2 Replikin concentration, the pA Replikin concentration can be considered control data that validate the location and uniqueness of the most significant Replikin Peak Gene for the present isolates of virus in the pB1 gene area. The significance of these observations is further increased when it is realized that these quantitative annual measures for each of three areas of the EIV genome are an objective determination by software scanning and counting of the virus proteins of each of the viruses isolated and reported annually at www.pubmed.com.

[000444] Applicants analyzed publicly available sequences for isolates of EIV from years 1942 to 2007 and determined the mean whole genome Replikin Count for all isolates having genomic sequences in each year for which they were available.

[000445] A list of the accession numbers analyzed by FluForecast® (REPLIKINS LLC, Boston, MA) for the presence and concentration of Replikin sequences is provided in Table 24 below. The mean Replikin concentration for each year is provided following the list of accession numbers from isolates in each corresponding year. Standard deviation and significance as compared to the mean Replikin concentration of the previous year and of the lowest mean Replikin concentration within the data set are also provided along with the mean Replikin concentration for each year.

Table 24 – Equine Influenza Whole Genome

Year	PubMed Accession Number-Replikin Count	No. of Isolates per year	Mean Replikin Count per year	S.D.	Significance
1942	<u>BAF49412</u> 9	1	1.9	0.0	
1956	<u>ABB20499</u> 4 <u>AAA43411</u> 1 <u>ABB20504</u> 2 <u>ABB20503</u> 2 <u>ABB20502</u> 8 <u>ABB20500</u> 4 <u>ABB20501</u> 16 <u>AAC57418</u> 16 <u>AAA43290</u> 4 <u>AAA43289</u> 4 <u>AAA43140</u> 18 <u>CAA44429</u> 30 <u>AAA43108</u> 3 <u>AAC35566</u> 1 <u>AAA52233</u> 8 <u>AAB51006</u> 2 <u>AAB51005</u> 2 <u>P88838</u> 16 <u>P26101</u> 30 <u>P26107</u> 18 <u>P16980</u> 8	21	2.2	1.3	low p<0.30
1957	<u>AAA64363</u> 43 <u>AAA64366</u> 43 <u>AAA64365</u> 43 <u>AAA64364</u> 43 <u>AAA64362</u> 36	5	7.4	0.6	low p<0.001, prev p<0.001
1963	<u>AAA43114</u> 7 <u>AAA43105</u> 7 <u>CAA44430</u> 30 <u>AAA43164</u> 7 <u>AAA43106</u> 3 <u>AAA43409</u> 1 <u>AAC31272</u> 3 <u>AAC31273</u> 3 <u>Q07579</u> 1 <u>P15658</u> 7 <u>P17002</u> 7 <u>P26094</u> 30 <u>P16979</u> 3	13	1.6	1.7	low p<0.40, prev p<0.001
1964	<u>CAA43815</u> 32 <u>CAA44432</u> 32 <u>AAC35580</u> 2 <u>P26097</u> 32 <u>P26096</u> 32	5	4.7	2.1	low p<0.04, prev p<0.01
1966	<u>CAA44433</u> 29 <u>P26098</u> 29	2	5.1	0.0	prev p>0.50
1969	<u>AAA43429</u> 3 <u>Q07581</u> 3	2	0.6	0.0	
1971	<u>AAA43111</u> 12 <u>P17000</u> 12	2	2.1	0.0	
1972	<u>AAA43100</u> 12 <u>CAA44434</u> 34 <u>AAA43355</u> 1 <u>Q07576</u> 1 <u>P26103</u> 34	6	2.8	2.6	low p<0.40,

	<u>P16994 12</u>				prev p>0.50
1973	<u>AAA43174 32 AAA43141 18 CAA44437 32 CAA44435 32</u> <u>AAA43104 27 AAA43637 17 AAA43457 7 AAB51000 2</u> <u>AAB50999 2 P26099 32 P26095 32 P26106 18 P16504 17 P13168</u> <u>27 P15673 7</u>	15	3.4	1.7	low p<0.005, prev p>0.50
1974	<u>AAA43093 16 P08327 16</u>	2	3.4	0.0	prev p>0.50
1976	<u>AAA43107 6 AAA43101 6 CAA44436 32 ABF60576 6 P16995 6</u> <u>P26102 32 P16997 6</u>	7	2.4	2.2	low p>0.50, prev p<0.20
1977	<u>AAC31296 3 CAA44431 32 AAQ90292 32 AAQ90288 7</u> <u>AAC31297 3 AAC31251 2 AAC31250 2 AAQ90293 16 P26100 32</u>	9	3.0	2.0	low p<0.10, prev p>0.50
1978					
1979	<u>AAA43427 3 AAC31274 3 AAC31275 3 AAC31249 2 AAC31248 2</u>	5	1.3	0.4	low p<0.02, prev p<0.02
1980	<u>AAA43109 6 P16998 6</u>	2	1.1	0.0	prev p<0.30
1981	<u>AAB02560 6 AAC31276 3 AAQ55062 7 AAC31277 3 AAA43245 3</u> <u>P08326 3 Q82559 6</u>	7	1.0	0.3	low p<0.001, prev p>0.50
1985	<u>AAQ90291 6 AAQ90289 15 AAQ90290 3 AAA43112 6 AAA43110</u> <u>4 Q6TXB9 3 P16999 4 Q6TXC0 15</u>	8	1.4	1.0	low p<0.10, prev p<.40
1986	<u>AAA43292 3 AAA43288 3 AAA43291 3 AAA43287 3 AAA43133</u> <u>18 AAA43102 6 AAA43113 32 AAA43638 17 AAA43479 15</u> <u>AAA43458 15 AAB51002 2 AAB51001 2 AAA43430 3 NMIVEA</u> <u>16 NMIVEK 3 Q07582 3 P17001 6 P19699 6 P26105 18 P16505 17</u> <u>P13169 32 P67915 15 P67914 15</u>	23	2.0	1.2	low p>0.50, prev p<0.10
1987	<u>CAC84083 3 AAA43103 6 CAD23745 2 P16996 6</u>	4	1.3	0.7	low p<0.10, prev p<0.05
1988	<u>AAC31278 3 CAA74382 3 AAC31279 3 AAC31255 2 AAC31254 2</u> <u>AAC31253 2 AAC31252 2</u>	7	1.4	0.3	low p<0.005, prev p>0.50
1989	<u>AAC31280 3 AAA43254 4 AAC31281 3 AAC31257 2 AAC31256 2</u> <u>AAA52247 3 AAA43253 4 AAA43151 12 AAA43374 3 CAA48482</u> <u>6 HMIVEE 6 HMIVET 6 HMIVE9 6 HMIVE8 4 HMIVE7 6</u> <u>HMIVE6 6 HMIVE5 6 HMIVE4 12 HMIVE3 12 HMIVE2 7</u> <u>HMIVE1 7 Q07578 3 Q08011 6 Q03909 12 P26068 3</u>	25	1.3	0.5	low p<0.001, prev p<0.30
1990	<u>AAB36977 9</u>	1	1.6	0.0	prev p<0.002
1991	<u>ABM21939 3 AAB36980 6 ABM21948 18 ABM21946 15</u> <u>ABM21945 33 ABM21944 2 ABM21943 2 ABM21942 15</u> <u>ABM21940 3 ABM21938 6 AAC31284 3 AAC31282 3 ABM21947</u> <u>15 ABM21941 3 CAA64893 6 CAA64894 6 AAC31285 3</u> <u>AAC31283 3 AAC31261 2 AAC31260 2 AAC31259 2 AAC31258 2</u> <u>AAA43354 3 AAC31286 3 AAC31287 3 AAC31263 2 AAC31262 2</u> <u>Q07575 3</u>	28	1.5	0.8	low p<0.01, prev p>0.50
1992	<u>AAB36979 6 AAC31292 3 AAC31293 3 AAC31269 2 AAC31268 2</u> <u>AAA62470 6 AAC31290 3 AAC31288 3 A45591 30 AAC31291 3</u> <u>AAC31289 3 AAC31267 2 AAC31266 2 AAC31265 2 AAC31264 2</u>	15	1.6	1.0	low p<0.20, prev p>0.50
1993	<u>AAB27733 6 AAB36978 6 AAB36975 6 AAC31294 4 S33703 6</u> <u>AAC31295 4 AAC31271 2 AAC31270 2</u>	8	1.3	0.3	low p<0.001, prev p<0.30
1994	<u>AAB36976 6</u>	1	1.1	0.0	prev p<0.02
1995	<u>NP_034974 16</u>	1	0.6	0.0	
1996	<u>AAC23906 23</u>	1	3.5	0.0	
1998	<u>AAF22345 3</u>	1	0.9	0.0	
1999	<u>ABA39843 6 AAZ23576 15 AAZ23564 1 ABA39854 3</u>	4	1.0	0.8	low p<0.05, prev p>0.50
2000	<u>AAQ18435 3 NP_898880 2 NP_057292 2</u>	3	0.9	0.0	low p<0.001, prev p>0.50
2001	<u>CAC69619 13 CAC69618 14 CAC69617 13 CAC69616 13</u> <u>CAC69615 14 CAC69614 13 CAC69613 14 CAC69612 33</u> <u>CAC69611 33 CAC69610 29 CAC69609 29 CAC69608 3</u> <u>CAC69607 4 CAC69606 4 CAC69605 14 CAC69604 15 CAC69603</u> <u>10 CAC69602 11 CAC69601 11 CAC69600 11 CAC69599 4</u>	37	3.1	2.0	low p<0.001, prev p<0.001

	<u>CAC69598 3 CAC69597 4 CAC69593 20 CAC69592 18 CAC69589 20 CAC69588 18 ABF69262 3 ABF69265 18 ABF69264 11 ABF69263 27 ABF69261 3 ABF69260 3 ABF69259 2 ABF69258 2 ABF69256 13 ABF69257 6</u>				
2002	<u>AAX23578 3 AAX23579 2 AAX23576 15 AAX23575 6 AAX23574 32 AAX23573 15 AAX23572 18 ABA42430 2 ABA42429 2 ABA39845 6 AAZ23582 15 AAZ23577 15 AAZ23563 18 AAX23577 3 ABA39853 3 AAM88392 4 AAM88391 4 AAN99161 14 AAN99160 15 AAN99159 10 AAN99158 11 AAN99157 11 AAN99156 11 AAN99155 4 AAN99154 4 AAN99153 4 AAN99152 3 AAN99151 4 AAN99147 20 AAN99146 18 AAN99143 20 AAN99142 18 AAN99141 6 AAN99140 6 AAN99139 3 AAN99138 3</u>	36	2.0	1.1	low p>0.50, prev p<0.002
2003	<u>AAZ23570 32 ABA42439 3 ABA42437 3 ABA42440 3 ABA42438 3 ABA42432 2 ABA42431 2 ABA42428 2 ABA42427 2 ABA39849 6 ABA39848 6 ABA39847 6 ABA39846 6 AAZ23583 15 AAZ23581 2 AAZ23578 15 AAZ23575 15 AAZ23574 15 AAZ23573 15 AAZ23567 18 AAZ23566 18 AAZ23565 18 AAZ23562 18 ABB17177 3 ABA39858 3 ABA39857 2 ABA39856 3 ABA39855 3 ABB17182 18 ABB17181 15 ABB17180 33 ABB17179 4 ABB17178 4 ABB17176 3 ABB17175 15 ABB17173 6 ABA42435 3 ABB17174 3 ABA42436 3 ABA42426 2 ABA42425 2 ABA39842 6 AAZ23580 15 AAZ23572 15 AAZ23569 32 AAZ23561 18 ABA39852 3 AAQ41575 20 AAQ41574 18 AAQ41573 6 AAQ41572 6 AAQ41571 3 AAQ41570 3</u>	53	1.8	1.1	low p>0.50, prev p>0.50
2004	<u>ABA42441 3 ABA42433 3 ABA42442 3 ABA42434 3 ABA42424 2 ABA42423 2 ABA39850 6 ABA39844 6 AAZ23579 15 AAZ23571 15 AAZ23568 32 AAZ23560 18 ABA39851 5 AAW17902 20 AAW17901 18 AAW17900 6 AAW17899 6 AAW17898 3 AAW17897 3 AAS32565 14 AAS32564 15 AAS32563 10 AAS32562 11 AAS32561 11 AAS32560 11 AAS32559 4 AAS32558 4 AAS32557 4 AAS32556 3 AAS32555 4 AAS32551 20 AAS32550 18 AAS32547 20 AAS32546 18 AAS32545 6 AAS32544 6 AAS32543 3 AAS32542 3</u>	38	2.0	1.2	low p>0.50, prev p<0.40
2005	<u>ABM47075 6</u>	1	1.1	0.0	prev p<0.001
2006	<u>ABI00612 2 ABI00611 2 ABI00610 13 ABI00609 14 ABI00608 13 ABI00607 14 ABI00606 13 ABI00605 14 ABI00604 13 ABI00603 14 ABI00602 33 ABI00601 33 ABI00600 29 ABI00599 29 ABI00598 3 ABI00597 4 ABI00596 4 ABI00595 14 ABI00594 15 ABI00593 10 ABI00592 11 ABI00591 11 ABI00590 11 ABI00589 4 ABI00588 3 ABI00587 4 ABI00583 20 ABI00582 18 ABI00579 20 ABI00578 18</u>	30	3.5	2.1	low p<0.001, prev p<0.001
2007	<u>BAF49411 30 ABS92756 20 ABS92755 18 ABS92754 6 ABS92753 6 ABS92752 3 ABS92751 3 ABN35533 14 ABN35532 15 ABN35531 10 ABN35530 11 ABN35529 11 ABN35528 11 ABN35527 4 ABN35526 4 ABN35525 4 ABN35524 3 ABN35523 4 ABN35519 20 ABN35518 18 ABN35515 20 ABN35514 18 ABN35513 6 ABN35512 6 ABN35511 3 ABN35510 3</u>	26	2.2	1.4	low p<0.20, prev p<0.005

Example 15

Analysis of Replikin Count in PCV to Predict Increased Morbidity and Mortality in PCV

[000446] Applicants analyzed publicly available sequences for isolates of PCV from www.pubmed.com using proprietary search tool software (ReplikinForecast™ available in the United States from REPLIKINS LLC, Boston, MA) from years 1997 to 2007 and

determined the mean Replikin Count for all isolates in each of years 1997 through 2007. Applicants then compared the mean Replikin Count for each year with qualitative changes in infection rates and mortality in pigs in Canada.

[000447] A list of the accession numbers analyzed for the presence and concentration of Replikin sequences is provided in Table 25 below. The mean Replikin Count for each year is provided following the list of accession numbers from isolates in each corresponding year. Standard deviation and significance as compared to the mean Replikin Count of the previous year and of the lowest mean Replikin Count within the data set are also provided along with the mean Replikin Count for each year.

Table 25 –PCV

Year	Analyzed Accession Numbers with total number of identified Replikin sequences; Replikin Count and Statistical Analysis
1997	AAC98885 16 AAC59472 4 AAC59466 3 AAC59464 5 AAC59462 9 Total Number of Accession Numbers: 5 Mean Replikin Count: 9.4 Standard Deviation: +/-10.8 Significance as compared to lowest mean Replikin Count: p<.40
1998	AAC35336 2 AAC35332 5 AAC35330 12 AAC35326 2 AAC35322 5 AAC35320 12 AAC35316 2 AAC35313 3 AAC35311 5 AAC35309 12 AAC35305 2 AAC35302 3 AAC35300 5 AAC35298 12 NP_065678 16 CAA11157 12 AAC61865 5 AAC61863 3 AAC61861 5 AAC61860 18 AAC61741 18 AAC61739 18 AAC61737 18 AAC34819 16 AAF97593 14 AAD03091 3 AAD03090 2 AAD03087 5 AAD03086 12 AAD03075 4 AAD03080 2 AAD03073 5 AAD03071 12 AAD03069 2 AAD03065 2 AAD03063 5 AAD03061 12 O56124 5 AAC69862 4 NP_047277 4 AAC69861 21 AAC34818 4 NP_047275 21 NP_048062 5 NP_048061 12 AAD11930 5 AAD11928 12 AAC69863 1 AAC34816 21 Total Number of Accession Numbers: 49 Mean Replikin Count: 6.1 Standard Deviation: +/-3.7 Significance as compared to lowest mean Replikin Count: p<.05 Significance as compared to previous mean Replikin Count p<.40
1999	BAA88133 12 AAD50432 12 AAD38398 12 AAG41230 2 AAG41228 5 AAG41226 12 AAD37776 12 AAD45580 16 AAF35304 12 AAF35302 12 AAF35300 12 AAF35298 12 AAF35296 12 AAF35294 12 AAF35292 12 AAD12314 2 AAD12313 2 AAD12309 6 AAD12308 12 AAE81207 4 Total Number of Accession Numbers: 20 Replikin Count: 5.0 Standard Deviation: +/-3.3 Significance as compared to lowest mean Replikin Count: p<0.50 Significance as compared to previous mean Replikin Count p>0.20
2000	CAC41085 12 CAC41084 12 CAC41083 11 AAL09364 12 AAL09363 12 AAF87238 12 AAF87236 12 AAF87234 12 AAF87232 12 AAF87230 12 AAF87228 12 NP_059530 2 NP_573443 5 NP_150370 4 CAD23544 5 CAC50263 4 AAG30569 3 AAG30566 8 AAG30563 8 AAG30560 7 AAG30557 4 AAG30554 4 AAG30551 7 AAG30548 7 CAC24649 35 AAF74197 2 Total Number of Accession Numbers: 26 Replikin Count: 2.9 Standard Deviation: +/-1.2 Significance as compared to lowest mean Replikin Count: p<0.001 Significance as compared to previous mean Replikin Count p>0.005
2001	AAK60464 5 AAK60462 12 AAL58397 12 BAB69442 5 BAB69441 12 BAB69439 2 BAB69438 5 BAB69437 12 BAB69435 3 BAB69433 5 BAB69432 12 AAK51544 14 AAK56300 12 AAK56298 12 AAK56296 12 AAL01081 4 AAL01080 1 AAL01077 5

	<p>AAL01075 12 NP_998971 5 NP_613078 4 AAM00235 4 AAN37998 4 AAN37994 3 AAN37990 3 AAN37986 3 AAN37982 4 AAN37978 3 AAN37974 4 AAN37970 3 AAN37966 4 AAN37962 4 AAN37958 3 AAL13485 3 CAC50253 2 CAC50247 2 Total Number of Accession Numbers: 36 Replikin Count: 2.8 Standard Deviation: +/-1.7 Significance as compared to lowest mean Replikin Count: p<0.001 Significance as compared to previous mean Replikin Count p>0.50</p>
2002	<p>AAM61272 12 AAM61262 12 AAM61274 16 AAM61268 12 AAM61266 12 AAM61270 12 AAM61264 12 AAO24127 1 AAO39760 12 AAM21847 1 AAM21846 1 AAM21845 7 AAM21844 12 AAO24128 12 AAO24126 11 AAO24124 10 AAO24122 12 AAO23147 12 AAO23145 12 AAN81597 12 AAN06826 12 AAN62769 12 AAN62767 12 AAN62765 12 AAN16398 14 AAM83186 11 AAL69968 12 AAN77863 4 AAN77862 2 AAN77861 2 AAN77860 14 AAN77859 16 AAO39666 16 AAM76057 12 Q8BB16 12 YP_164519 4 ABA54889 4 ABA54887 4 ABA54885 4 AAR28043 4 AAO95302 12 AAO95299 Total Number of Accession Numbers: 42 Replikin Count: 3.5 Standard Deviation: +/-1.4 Significance as compared to lowest mean Replikin Count: p<0.001 Significance as compared to previous mean Replikin Count p>0.05</p>
2003	<p>AAP51128 12 AAS65982 1 AAS65993 12 AAS65991 12 AAS65989 12 AAS65987 12 AAS65985 12 AAS65983 12 AAS65981 9 AAS65979 12 AAS65977 12 AAS65975 12 AAP83635 12 AAP83633 12 AAP83631 12 AAP83629 12 AAP83627 12 AAP83625 12 AAP83623 12 AAP83621 12 AAP83619 12 AAP83617 12 AAP83615 12 AAP83613 12 AAP83611 12 AAP83609 12 AAP83607 12 AAP83605 12 AAP83603 12 AAP83601 12 AAP83599 12 AAP83597 12 AAP83595 12 AAP83593 12 AAP83591 12 AAR97518 1 AAR03722 12 AAR03720 12 AAR03718 12 AAR03716 12 AAQ94098 12 AAQ94096 12 AAQ94094 12 AAQ94092 12 AAQ94090 12 AAQ94088 12 AAP44190 5 AAP44188 12 AAP44187 5 AAP44185 12 AAP44184 5 AAP44182 12 AAO61773 12 AAR97517 12 AAQ96327 12 AAQ23156 5 AAQ23155 12 AAP42468 9 AAP42466 10 AAP42464 12 AAO61136 12 NP_937956 12 AAR03714 12 Q805H4 16 YP_209622 4 AAQ93492 6 AAP69227 4 AAR27947 1 AAT00481 5 AAT00473 4 AAT00471 4 AAT00469 4 AAS16932 6 AAS16931 7 AAS16930 6 AAS16929 6 AAS16928 6 AAS16927 6 AAS16926 6 AAS16925 6 AAS16924 6 Total Number of Accession Numbers: 81 Replikin Count: 3.4 Standard Deviation: +/-1.0 Significance as compared to lowest mean Replikin Count: p<0.001 Significance as compared to previous mean Replikin Count p>0.50</p>
2004	<p>AAU87520 1 AAT97651 2 AAT97647 1 YP_077191 16 AAW78483 16 AAW78481 20 AAW78479 20 AAW78477 20 AAW78475 12 AAW78473 12 AAW78471 12 AAW78469 12 AAW78467 12 AAW78465 12 AAW78463 12 AAV34141 20 AAV34139 12 AAU87519 12 AAU87517 12 AAU87515 12 AAU87513 12 AAU87511 12 AAU87509 12 AAU87507 12 AAU34001 10 AAU01913 14 AAT97650 12 AAT97648 12 AAT97646 12 AAT97644 12 AAT77546 1 AAT72755 16 AAT36358 12 AAU13781 1 AAX49397 12 AAU01966 12 AAT79579 12 AAT72901 12 AAT58234 11 AAS66199 1 AAS66197 1 AAS45844 12 AAS45843 12 CAJ31064 12 AAU13780 12 AAX52911 12 AAU87505 12 AAT39479 12 AAT39460 9 AAT37493 9 AAS66198 12 AAS66196 12 AAS66194 12 AAS66192 12 AAS66190 12 AAS90297 12 AAS89260 12 NP_999004 27 YP_271921 7 CAF25171 12 AAT51967 3 BAD90990 3 BAD90989 4 BAD90988 4 BAD90987 4 BAD90986 3 BAD90985 3 AAS86324 4 AAS93283 6 AAS93279 6 AAS93276 5 AAS93272 9 AAS93268 6 AAS89814 4 AAS89813 6 AAS89812 8 AAS89811 6 AAS89810 6 AAS89809 6 AAS89808 8 AAS89807 9 AAS89806 8 AAS89805 5 AAS89804 4 AAS89803 4 AAS89802 5 AAS89801 7 AAS89800 9 AAS89799 6 AAS89798 11 AAS89797 6 AAS89796 6 AAS89795 5 AAS89794 5 AAS89793 6 AAS89792 11 AAS89791 9 AAS89790 9 AAS89789 9 AAS89788 6 AAS89787 6 AAS89786 6 AAS89785 6 AAS89784 6 AAS89783 6 AAS89782 6 AAS89781 6 Total Number of Accession Numbers: 107 Replikin Count: 3.3 Standard Deviation: +/-1.3</p>

	Significance as compared to lowest mean Replikin Count: $p < 0.001$ Significance as compared to previous mean Replikin Count $p > 0.30$
2005	ABJ98317 12 ABA29241 12 AAZ20802 11 AAZ20800 10 AAZ20798 11 AAZ20796 12 AAZ20794 12 AAW79865 9 ABC26025 12 ABA40480 12 AAZ78351 12 AAY40292 12 AAX21515 12 ABB29423 10 ABB29421 12 ABB29419 12 ABB29417 10 ABB29415 10 ABB29413 10 ABB29411 10 ABB29409 10 ABB29407 10 ABB29405 10 ABB29403 8 ABB29401 10 ABB29399 10 ABB36791 1 ABA60807 12 ABA60805 12 ABA60803 11 ABA40399 12 ABA40397 12 AAZ66792 11 AAX10150 12 AAX62053 16 AAX62051 12 AAX62049 12 AAX62047 12 AAX62045 12 AAX62043 16 AAX62041 12 ABC75103 12 ABB20934 12 ABA26910 12 ABA26908 10 AAY34249 12 YP_610962 8 AAZ07884 7 ABB59615 8 ABA39170 3 ABA39166 3 AAX35672 1 ABA39162 4 ABA39158 3 ABA39154 4 ABA39150 3 ABA39146 3 ABA39142 3 ABA39138 3 AAZ68049 2 AAZ68045 2 Total Number of Accession Numbers: 61 Replikin Count: 3.0 Standard Deviation: +/-1.1 Significance as compared to lowest mean Replikin Count: $p < 0.001$ Significance as compared to previous mean Replikin Count $p > 0.20$
2006	ABG21191 3 ABI29887 12 ABG21279 10 ABG21277 10 ABG21275 10 ABG21273 10 ABG21271 10 ABG21269 12 ABG21267 12 ABJ98319 12 ABI93799 16 ABI93797 12 ABD59347 12 ABG48510 14 ABD42928 12 ABM67071 16 ABM88864 12 ABM88862 12 ABM88860 12 ABI17537 12 ABI17535 12 ABI17533 12 ABI17531 12 ABI17529 12 ABI17527 12 ABI17525 12 ABI17523 12 ABG37025 16 ABG37023 12 ABD52438 16 ABG24031 16 ABG24029 16 ABF71465 12 ABF19812 16 ABF19810 14 ABE96824 16 ABE96822 16 ABE96820 16 ABE96818 13 ABE96816 16 YP_803548 1 YP_803551 5 ABI54258 5 ABI54255 1 ABK79791 2 ABK79788 2 ABK79785 2 ABK79782 3 ABK79779 7 ABK79776 9 ABK79773 3 ABI97391 3 ABE03771 4 ABE03767 6 Total Number of Accession Numbers: 54 Replikin Count: 3.4 Standard Deviation: +/-1.4 Significance as compared to lowest mean Replikin Count: $p < 0.001$ Significance as compared to previous mean Replikin Count $p > 0.10$

Example 16

Repeat KHKK (SEQ ID NO: 1584) Signatures in Lung Cancer and Tobacco Mosaic Virus

[000448] Publicly available amino acid sequences at Accession Nos: Q9NS56 and 117607067, for non-small lung cancer and tobacco mosaic virus, respectively, were analyzed for Replikin Peak Genes. The inventors queried Accession No. Q9NS56 at www.pubmed.com. Accession No. Q9NS56 discloses the amino acid sequence of E3 ubiquitin-protein ligase Topors from human chromosome 9 of non-small cell lung cancer (SEQ ID NO:1740). Upon analysis of SEQ ID NO:1740, the inventors observed a Replikin Peak Gene having continuous Replikin sequences that begin at residue 880 (lysine) and continue through residue 897 (histidine).

[000449] The inventors isolated the RPG (SEQ ID NO:1741) *in silico*. SEQ ID NO:1741 was identified for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for lethality. Fifty-two Replikin sequences (SEQ ID NOS:1886-1937) were identified in the RPG of SEQ

ID NO:1741 for diagnostic, therapeutic and predictive uses as described herein. SEQ ID NOS: 1742-1747 were identified in the amino-terminal of the sequence disclosed in Accession No. Q9NS56 (SEQ ID NO:1741), SEQ ID NOS:1748-1780 were identified in the mid-molecule of the sequence, and SEQ ID NOS:1781-1885 were identified in the carboxy-terminal of the sequence.

[000450] The Replikin Count of the amino acid sequence (SEQ ID NO:1740) disclosed at Q9NS56 was 144 Replikin sequences in 1045 total amino acids for a Replikin Count of 13.8. The Replikin Count of the RPG (SEQ ID NO:1741) was 52 Replikin sequences in 18 total amino acids for a Replikin Count of 289, the highest count yet observed.

[000451] Within the Replikin sequences identified in the RPG (SEQ ID NO:1741), the KHKK signature was observed 57 times within 52 Replikin sequences. This high concentration of lethal signatures corresponds to the high lethality of non-small cell lung malignancies.

[000452] The inventors queried Accession No. 117607067 at www.pubmed.com. Accession No. 117607067 discloses the amino acid sequence of a hot pepper 26S proteasome subunit RPN7 induced by tobacco mosaic virus (SEQ ID NO:1938). Upon analysis of SEQ ID NO:1938, the inventors observed a Replikin Peak Gene having continuous Replikin sequences that begin at residue 91 (histidine) and continue through residue 175 (lysine).

[000453] The inventors isolated the RPG (SEQ ID NO:1939) *in silico*. SEQ ID NO:1939 was identified for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for lethal outbreaks of tobacco mosaic virus. Fifty-four Replikin sequences (SEQ ID NOS:1941-1994) were identified in the RPG of SEQ ID NO:1939 for diagnostic, therapeutic and predictive uses as described herein.

[000454] SEQ ID NO:1941 was identified in the amino-terminal of the sequence disclosed in Accession No. 117607067 (SEQ ID NO:1938), SEQ ID NOS:1942-1986 were identified in the mid-molecule of the sequence, and SEQ ID NOS:1987-1994 were identified in the carboxy-terminal of the sequence. Each Replikin sequence was isolated *in silico* for diagnostic, therapeutic and predictive purposes as described herein including for immunogenic compositions and vaccines.

[000455] The Replikin Count of the amino acid sequence (SEQ ID NO:1938) disclosed at Accession No. 117607067 was 55 Replikin sequences in 179 total amino acids for a Replikin

Count of 30.7. The Replikin Count of the RPG (SEQ ID NO:1939) was 54 Replikin sequences in 89 total amino acids for a Replikin Count of 61.

[000456] Within the Replikin sequences identified in 117607067 (SEQ ID NO:1938), the KHKK (SEQ ID NO: 1584) signature was observed twenty times within 61 Replikin sequences. This high concentration of lethal signatures corresponds to the high lethality of tobacco mosaic virus and connects tobacco mosaic virus through KHKK (SEQ ID NO:1584) signatures to lethal lung cancer.

[000457] As discussed above, repeating signatures such as a “KHKK” (SEQ ID NO:1584)signature have been observed in Replikin sequences within RPGs of lethal malignancies, viruses and organisms. The KHKK (SEQ ID NO:1584)signature has been observed eleven times within the RPG of the protozoa that causes most malaria, *P. falciparum*, 20 times within the RPG of tobacco mosaic virus, which caused exacerbated cell death induced by tobacco mosaic virus, and 57 times in non-small cell lung carcinoma within 52 Replikins observed within the 18 amino acid RPG identified in chromosome 9 of a non-small cell lung carcinoma. The presence of such a high number of KHKK (SEQ ID NO:1584) signatures within the 18 amino acid RPG of the non-small cell lung carcinoma is explained by overlapping of the signatures. Overlapping of Replikin sequences and repeated signatures such as KHKK (SEQ ID NO:1584) has now been associated with lethality, virulence and rapid replication. Together, these data indicate that a Replikin gene is quantitatively associated with lethal functions, and may be a mobile agent of lethality transferring between strains and species.

Example 17

Repeat KHKK (SEQ ID NO:1584) Signatures in Malaria

[000458] The inventors queried Accession No. P13817 at www.pubmed.com. Accession No. P13817 discloses an amino acid sequence from *Plasmodium falciparum*. The inventors analyzed the amino acid sequence provided at P13817 (SEQ ID NO:2043). Upon analysis of SEQ ID NO:2043, the inventors observed a Replikin Peak Gene having continuous Replikin sequences that begin at residue 323 (histidine) and continue through residue 473 (lysine) (SEQ ID NO: 3659).

[000459] The inventors isolated the RPG (SEQ ID NO:3659) *in silico*. SEQ ID NO:3659 was identified for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for lethal outbreaks of malaria. Two hundred and thirty-one Replikin sequences (SEQ ID NOS:2312-

2315 and 2317-2544) were identified in the RPG of SEQ ID NO:3659 for diagnostic, therapeutic and predictive uses as described herein. Replikin sequences SEQ ID NOS:2044-2077 were identified in the amino-terminal of the sequence of SEQ ID NO:2043, Replikin sequences SEQ ID NOS: 2079-2080 were identified in the mid-molecule of the sequence, and Replikin sequence SEQ ID NOS:2081-2315 were identified in the carboxy-terminal.

[000460] The Replikin Count of the whole sequence (SEQ ID NO:2043) was 268 Replikin sequences in 473 total amino acids for a Replikin Count of 56.7. The Replikin Count of the RPG area (SEQ ID NO:3659) was 231 Replikin sequences in 151 total amino acids for a Replikin Count of 153.

[000461] The inventors queried Accession No. A44396 at www.pubmed.com. Accession No. A44396 discloses an amino acid sequence from an ATP-ase-like molecule of *P. falciparum* isolated in 1993. The inventors analyzed the amino acid sequence provided at A44396 (SEQ ID NO:2926). Upon analysis of SEQ ID NO: 2926, the inventors observed a Replikin Peak Gene having continuous Replikin sequences that begin at residue 1297 (histidine) and continue through residue 1333 (histidine).

[000462] The inventors isolated the RPG (SEQ ID NO:3661) *in silico*. SEQ ID NO:3661 was identified for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for lethal outbreaks of malaria. Seventeen Replikin sequences (SEQ ID NOS:3282-3285, 3287-3291, 3293, 3295, 3299-3300, 3302, 3304, 3306, 3308, 3310-3313 and 3663) were identified in the RPG of SEQ ID NO:3661 for diagnostic, therapeutic and predictive uses as described herein. Replikin sequences SEQ ID NOS:2546-2632 were identified in the amino-terminal of the sequence of SEQ ID NO:2926, Replikin sequences SEQ ID NO: 2633-2720 were identified in the mid-molecule of the sequence, and SEQ ID NOS: 2721-2900 were identified in the carboxy-terminus.

[000463] The Replikin Count of the whole ATP-ase sequence (SEQ ID NO:2926) was 355 Replikin sequences in 1984 total amino acids for a Replikin Count of 17.9. The Replikin Count of the RPG area (SEQ ID NO:3661) was 15 Replikin sequences in 37 total amino acids for a Replikin Count of 41.

[000464] Eleven signature repeat KHKK (SEQ ID NO:1584) sequences were noted in the 15 Replikin sequences of the RPG. The eleven signature repeats, namely, SEQ ID NOS:3286, 3292, 3294, 3296, 3298, 3662, 3301, 3303, 3305, 3307, and 3309 are respectfully

found within noted Replikin sequences SEQ ID NOS:3286, 3291, 3293, 3295, 3297, 3299, 3300, 3302, 3204, and 3206.

[000465] The presence of such a high number of KHKK (SEQ ID NO:1584) signatures within the fifteen Replikin sequences in the 37 amino acid RPG of the *P. falciparum* is explained by overlapping of the signatures. Overlapping of Replikin sequences and repeated signatures such as KHKK (SEQ ID NO:1584) has now been associated with lethality, virulence and rapid replication as in malaria, which has an exceptionally high rate of replication within its lifecycle.

Example 18

Laboratory Demonstration of Relationship of Replikin Count to Percent Mortality in Taura Syndrome Virus Infection in Shrimp

[000466] To test further the relationship of Replikins to virulence, the relationship of Replikin count of shrimp viruses to mortality in shrimp was examined in a controlled situation. Based on the hypothesis that the Replikin count of a virus is related to virulence of the virus and the percent mortality of the host, as developed from the evidence on H5N1 virus infections in humans, Applicants tested whether it would be possible to predict solely from the Replikin Count of the amino acid sequence of the whole genome what the order of virulence would be of four strains of the virus. Taura syndrome shrimp virus (TSV), which kills most host shrimp within a few days of infection, was chosen to be studied. The amino acid sequences of four strains of taura syndrome virus (Belize, Thailand, Hawaii, and Venezuela) were analyzed with the FluForecast® software of REPLIKINS LLC, Boston, MA and the results held in confidence until the laboratory challenge experiments with the virus were completed, then compared with the percent mortality produced by each strain.

[000467] In the laboratory, there was a significant linear correlation between the mortality rates of the host shrimp challenged with each of the four virus strains and the mortality rates predicted earlier by only the Replikin counts of each strain. These data support the conclusion that virus Replikin peptide concentration, in addition to predicting virus outbreaks, relates quantitatively to host mortality rate and to the increase in virulence over time observed.

A. Replikin Analysis

[000468] Visual Replikin analysis was performed on the sequence information for the taura syndrome virus isolates from Belize, Thailand, Hawaii, and Venezuela by applying the algorithm defining Replikins with computer access to protein and genomic sequences freely

available on PubMed or other public databases. The specific defining algorithm follows: a Replikin is a peptide sequence in a protein or genome, 7 to 50 amino acids long having a terminal lysine and a terminal lysine or histidine, containing at least 2 lysine groups 6 to 10 amino acids apart, at least 1 histidine group, and at least 6% lysine. Overlapping Replikins are common and are counted separately. The quantitative correlations with rapid replication and epidemics and lethality require all components of the algorithm to be in place for each Replikin. Thus for example, if the length and lysine requirements are present but there is no histidine present, the peptide is not a Replikin. Automated Replikin analysis was performed with the FluForecast® software service of Replikins Ltd., Boston, MA.

B. Identification of the Replikin Peak Gene

[000469] The Replikin count was used to identify that area of the genome which had the highest concentration of Replikins, and this area called the Replikin Peak Gene (RPG) area. The further two to eight-fold increase in the Replikin count of the RPG which occurred with outbreaks was further used to confirm the identity of this gene. The function of the gene was therefore used to identify it or isolate it “in silico”.

C. Shrimp Virus Laboratory Methods

[000470] At the Aquaculture Pathology Laboratory, Department of Veterinary Science and Microbiology, University of Arizona, Tucson AZ, small juveniles of specific-pathogen-free *Litopenaeus vannamei* shrimp per tank, mean weight: 1.8 g, were fed minced TSV-infected tissues (infected separately with each of the 4 isolates originating from Belize, Thailand, Venezuela and Hawaii) for 3 days at 5% of their body weight. These shrimp were maintained with pelleted ration (Rangen 35%) for the following 12 days. Each challenge bioassay of a specific isolate was done in triplicate. During the bioassay period, all tanks were checked daily for dead or moribund shrimp. All mortalities were removed from the tank and frozen. One to three moribund shrimp from each isolate were preserved in Davidson's AFA fixative and processed for routine histology to confirm viral infection. For each isolate, six moribund shrimp were collected during the acute phase infection and total RNA was extracted from their gill tissues with a High Pure RNA tissue kit (Roche). The extracted RNA was analyzed for the presence of TSV by real-time RT-PCR. All tanks were outfitted with an acclimated biological filter and aeration, and were covered with plastic to contain aerosols. The average salinity of the water was 23 ppt and the water temperature was 28°C. The challenge study was terminated after 15 days with live animals counted as survivors.

D. Comparison of Virulence

[000471] First mortality was seen on day 2 after exposure to TSV in all 4 isolates. For Belize isolate, most (83%) of shrimp died by day 4 and had a 0% survival at day 11 (Fig 14A, Table 26). For Thailand isolate, 63% mortalities occurred by day 4 and had 20% survivals at the end of 15-day bioassay (Fig 14B, Table 26). For Hawaii isolate, mortalities increased starting at day 2 and reached to a peak at day 5; the cumulative survival is 22% at the end (Fig 14C, Table 26). For Venezuela isolate, mortalities occurred slowly at days 2 and 3 with 22% of shrimp showed mortalities on day 4 and then mortalities were slowing down; there were 42% of shrimp survived in the end (Fig 14D, Table 26). The time period for reaching 50% mortality caused by TSV infection for the isolate of Belize, Thailand, Hawaii and Venezuela were 2.8, 3.5, 4.5 and 7 days, respectively (Table 26).

[000472] Figure 14 provides data for the cumulative survival of *Litopenaeus vannamei* challenged with TSV isolates per os with taura syndrome virus isolates from a: Belize; b: Thailand; c: Hawaii; d: Venezuela. The data from Figure 14 is contained in Table 26 below.

Table 26 – TSV Challenge

TSV isolate	GenBank no. (ORF1)	Survival (%) (Mean)	Day of 50% mortality
Belize	AAT81157	0	2.8
Thailand	AAY56363	20	3.5
US-Hawaii	AAK72220	22	4.5
Venezuela	ABB17263	42	7.0*

*High variation was observed in Venezuela's triplicate tanks, thus the Day of 50% mortality was determined by Kaplan-Meier survival analysis with the Statistix 8 program.

[000473] The correlation of the virulence observed for each of the TSV isolates with the predicted virulence by Replikin Count alone are shown in Figure 15. Figure 15A provides data comparing Replikin Counts of the four isolates with the mean day of 50% mortality as gathered in blind studies. Figure 15B provides data comparing Replikin Counts of the four isolates with mean cumulative mortality as gathered in blind studies. The linear quantitative relationship between the predicted and experimental values is evident.

[000474] Table 27 below provides the histological data that was gathered for the moribund shrimp to demonstrate TSV infection.

Table 27 – Histology

UAZ ID#	TSV Isolate	Days after exposure	TSV lesions ¹	LOS ²
O6-407J/1	Belize	3	G4	G4
O6-407F/1	Thailand	3	G4	G2
O6-407D/1	Thailand	4	G4	G3
O6-407E/1	Thailand	4	G3	G2

06-407A/1	Hawaii	4	G2	G3
06-407C/1	Hawaii	4	G2	G4
06-407H/1	Venezuela	4	G4	G2

Severity grade: G1: sign of infection; G2: moderate signs of infection; G3: moderate to high signs of infection; G4: severe infection.

1. TSV lesions = Presence of TSV pathognomonic lesions in the gills, mouth, stomach, intecumental cuticular epithelium, and appendages.

2. LOS = presence of lymphoid organ spheroids within the lymphoid organ.

[000475] Belize TSV: Acute lesions of diagnostic TSV infection were found in one representative shrimp sample (06-407J/1) at a severity grade of G4. Nuclear pyknosis and karyorrhexis were observed in the cuticular epithelium of the general body surface, appendages, gills, stomach and esophagus. Lymphoid organ spheroids were also found at severity grade G4. Thailand TSV: Severe (G4) TSV infection was detected in 2 out 3 shrimp (06-407D/1, F/1), another shrimp (06-407E/1) showed a moderate to high grade (G3) of infection. Lymphoid organ spheroids were found at severities of G2 and G3. Hawaii TSV: Moderate level (G2) of TSV infection was detected in 2 shrimp (06-407A/1, C/1) collected at day 4. Lymphoid organ spheroids were found at severities of G3 and G4. Venezuela TSV: Severe (G4) TSV infection was detected in one representative shrimp (06-407H/1) sampled at day 4. Lymphoid organ spheroids were found at severity of G2.

[000476] The real-time TSV RT-PCR assay was designed specifically for Hawaii TSV and thus a high level (10^7 copies / μ l RNA) of TSV was detected in the Hawaii-TSV challenged shrimp (Table 28). The target sequence in 3 other isolates has 2 mis-matched nucleotides with the primers/TaqMan probe. Thus, there is 10 times less quantity of TSV (10^6 copies/ μ l RNA) detected in Belize and Thailand samples. The Venezuela samples were detected with 100-100,000 times less: 10^2 - 10^5 copies/ μ l RNA; this may be due to both the effect of mismatches and a lower level of infection in the samples analyzed. Nevertheless, all 24 samples (6 from each isolates) were all positive for TSV infection. This confirms that the mortalities observed from bioassays are from TSV infection. The real-time TSV RT-PCR assay data is found below in the Table 28.

Table 28 – PCR

TSV isolate	Mean (Range) TSV copies:/l RNA
Belize	2.7×10^6 (4.8×10^5 - 4.4×10^6)
Thailand	2.7×10^6 (4.3×10^5 - 7.5×10^6)
Hawaii	5.2×10^7 (2.3×10^7 - 7.5×10^7)
Venezuela	6.5×10^5 (6.5×10^2 - 2.0×10^5)

E. Laboratory Mortality Results Correlated With Replikin Counts

[000477] Virulence of 4 TSV isolates (Hawaii, Belize, Thailand and Venezuela) was compared through a per os laboratory infection in juvenile *Litopenaeus vannamei* (Kona stock, Oceanic Institute, Hawaii). The results showed that the Belize isolate is the most virulent, Thailand is the second, followed by the Hawaii isolate, and the Venezuela isolate is the least virulent. This is based on the analyses of cumulative survivals at the end of the bioassay ($p < 0.047$) and the time when 50% mortality was occurred ($p < 0.001$). That the mortality of the shrimp was caused by TSV infection was confirmed by positive reactions in RT-PCR detection and by the appearance of characteristic lesions observed in histological analysis

F. Laboratory Mortality Results Correlated With Replikin Counts

[000478] Experimentally, Replikin Counts alone prospectively correctly predicted: (1) blind in controlled experiments in the laboratory, the order of lethality in shrimp of four strains of taura syndrome virus (Figures 15A and B); (2) an increasing H5N1 percent mortality in humans (Figure 4); and (3) the host (Figure 5); and (4) the country in which the latter would occur, Indonesia (Figure 6). For both H5N1 influenza in human hosts, and taura syndrome virus infection in shrimp hosts, evidence in this study demonstrates the quantitative relationship of the virus Replikin Count to the mortality rate in the host. The ability to predict blind is of course one of the more definitive proofs of a relationship; the demonstration of a quantitative linear relationship is even more definitive. Thus, the concentration of a class of specific virus peptides, Replikins, has here been quantitatively correlated with the percent mortality these viruses produce in their respective hosts, namely invertebrate crustacean (shrimp) and vertebrates (humans). To our knowledge, no quantitative correlation of virus structure and host lethality has been reported previously.

Example 19**Increased Host Resistance to Taura Syndrome Virus by Administration of Synthetic Replikins**

[000479] Shrimp cultured using the Challenge Methods described in Example 18 above were exposed in a first experiment for two weeks to synthetic Replikins per os mixed in their feed. The Replikins were peptides specific to Replikin sequences present in the TSV Hawaii strain isolate with which the shrimp were challenged.

[000480] In the experiment, mortality was reduced by 50% compared to a control group. The control group was given feed not containing synthetic Replikin sequences. A second control group was fed Replikin sequences synthesized with the covalent binding of additional

amino acids to the same synthetic Replikins fed to the shrimp. The covalently “blocked” Replikins did not increase shrimp resistance to the virus in the same experiment demonstrating that the increase in host resistance was specific to the Replikin peptide structure.

[000481] Because little is known about the details of the immune system of the shrimp (shrimp appear not to produce antibodies), the phenomenon of “resistance” to infection appears to be based in a “primitive immune system” perhaps similar to the “toll receptor” and related systems. Thus the term “increased resistance” is used for the observed phenomenon and Replikin feed is used rather than “vaccine” for the administered substance which increases resistance.

[000482] The surviving shrimp of the first challenge were then set up in a fresh culture, fed for an additional two weeks with feed containing Replikin sequences, then again challenged with the Hawaii strain of taura syndrome virus. The Replikin sequence supplemented feed was maintained while the survivors were again challenged repeatedly by the same virus, in repeated cycles, until 100% of the shrimp survived the TSV challenge.

Example 20

Calculation of RPG in Viral Hemorrhagic Disease in Fish

[000483] The inventors queried Accession No. ABQ42711 at www.pubmed.com. Accession No. ABQ42711 discloses an amino acid sequence from a glycoprotein in hemorrhagic septicemia virus. Hemorrhagic septicemia virus is a cause of hemorrhagic disease in fish. The inventors analyzed the amino acid sequence provided at ABQ42711 (SEQ ID NO:3787). Upon analysis, the inventors observed a Replikin Peak Gene having continuous Replikin sequences that begin at residue 81 (histidine) and continue through residue 204 (histidine).

[000484] The inventors isolated the RPG *in silico* for diagnostic and therapeutic uses in, for example, an immunogenic compound and a therapeutic vaccine compound and as a predictive sequence for lethal outbreaks of hemorrhagic disease in fish. Thirty-six Replikin sequences (SEQ ID NOS:3788-3823) were identified in SEQ ID NO:3787 for diagnostic, therapeutic and predictive uses as described herein. Replikin sequences SEQ ID NOS:3788-3795 were identified in the amino-terminal, Replikin sequences SEQ ID NOS: 3796-3815 were identified in the mid-molecule of the sequence, and Replikin sequences 3816-3823 were identified in the carboxy-terminal. All were isolated for diagnostic, therapeutic and predictive purposes.

[000485] The Replikin Count of the whole sequence (SEQ ID NO:3787) was 36 Replikin sequences in 222 total amino acids for a Replikin Count of 16. The highest Replikin Count of an identified RPG area in hemorrhagic septicemia virus was 73 Replikin sequences in 123 total amino acids for a Replikin Count of 59.

[000486] The inventors queried publicly available sequences from isolates of hemorrhagic viral disease syndrome in fish from 1990 through 2007. The following table provides the accession numbers queried. The highest Replikin Count of an identified RPG area in hemorrhagic septicemia virus was 73 Replikin sequences in 123 total amino acids for a Replikin Count of 59.

[000487] The inventors queried all sequences for hemorrhagic viral disease in fish publicly available at www.pubmed.com between 1990 and 2007 Using FluForecast® (Replikins LLC, Boston, MA), the inventors determined the mean Replikin Count in each year from 1990 -2007. The data is provided in Table 29. The table does not included years in which no data was available.

Table 29 – Hemorrhagic septicemia Mean Replikin Counts

Year	PubMed Accession Number-Replikin Count	No. of Isolates per year	Mean Replikin Count per year	S.D.	Significance
1990	BAA00591 3 CAA00881 29 CAA52082 33 CAA52081 33 CAA52080 33 CAA52079 33 CAA52078 33 CAA52077 33 P24378 3	9	5.0	2.4	low p<.001
1991	CAA41859 29 CAA41858 29 CAA01751 4 CAA41930 4 P27663 2 P27662 29	6	3.5	2.5	low p<.001, prev p<.20
1992	CAA46926 33 P27371 4	2	3.7	3.9	low p<.10, prev p>.50
1993	AAB26115 33 AAT01207 48 AAT01206 48 AAT01205 48 AAT01204 48 AAT01203 48 AAT01202 48 AAT01201 48 AAT01200 48 AAT01199 48 AAT01198 48 AAT01197 48 AAT01196 48 AAT01195 48 AAT01194 48 AAT01193 48 AAT01192 48 AAT01191 48 AAT01190 48 AAT01189 48 AAT01188 48 AAT01187 48 AAT01186 48 AAT01185 48 AAT01184 48 AAT01183 48 AAT01182 48 AAT01181 48 AAT01180 48 AAT01179 48 AAT01178 48 AAT01177 48 AAT01176 48 AAT01175 48 AAT01174 48 AAT01173 48 AAT01172 48 AAT01171 48 AAT01170 48 AAT01169 48 AAT01168 48 AAT01167 48 AAT01166 48 AAT01165 48 AAT01164 48 AAT01163 48 AAT01162 48 AAT01161 48 AAT01160 48 AAT01159 48 AAT01158 48 AAT01157 48 AAT01156 48 AAT01155 48 AAT01154 48 AAT01153	64	9.4	0.4	low p<.001, prev p<.04

	<u>48 AAT01152 48 AAT01151 48 AAT01150 48 AAT01149 48 AAT01148 48 AAT01147 48 AAT01146 48 AAT01145 48</u>				
1994					
1995	<u>AAB88231 33 AAB88230 38 AAB88229 41 AAB88232 30 Q96460 6</u>	5	5.9	2.7	low p<.001, prev p<.002
1997	<u>AAC24962 3 CAB59222 53 CAB59221 53 CAB07737 53 CAB59220 53 CAB59219 53 CAB59218 53 CAB59217 52 CAB59216 52 CAB59215 52 CAB59214 52 CAB59213 52 CAB59212 52 AAB88228 52 AAB88227 30 AAB88226 29 AAB88225 30 AAB88224 30 AAB88223 23 AAB88222 30 CAB07754 15 CAB07753 17 CAB07752 14 CAB07751 7 CAB07750 14 CAB07749 14 CAB07748 14 CAB07747 11 CAB07746 14 CAB07745 14 CAB07744 14 CAB07743 15 CAB07742 15 CAB07741 7 CAB07740 15 CAB07739 14 CAB07738 14 CAB07736 11 CAB07734 20 CAB07733 11 CAB07732 14 CAB07731 15 CAB07730 3 CAB07729 14 CAB07728 14 CAB07727 15</u>	46	8.3	4.4	low p<.001, prev p<.05
1998	<u>NP_049550 55 NP_049549 55 NP_049548 55 NP_049547 55 NP_049546 55 NP_049545 55 CAB57984 1 CAB44726 55 CAB44725 55 CAB44724 55 CAB44723 55 CAB44722 55 CAB44721 55 CAB40833 66 CAB40832 66 CAB40831 66 CAB40830 66 CAB40829 66 CAB40828 66 CAA08837 58</u>	20	2.9	0.6	low p<.001, prev p<.001
1999	<u>AAF04486 57 AAF04485 57 AAF04484 57 AAF04483 57 AAF04482 57 AAF04481 57 AAF04480 53 AAF04479 53 AAF04478 53 AAF04477 53 AAF04476 53 AAF04475 53 BAC29401 14</u>	13	2.9	0.3	low p<.001, prev p>.50
2000					
2001	<u>BAB70674 15 BAB70673 16 BAB70672 16 BAB70671 16 AAL83805 32 AAL83804 35 AAL83803 40</u>	7	7.4	0.7	low p<.001, prev p<.001
2002	<u>AAN85721 30 CAD31945 13 CAD31944 7 CAD31943 7 CAD31941 7 CAD31924 7 CAD31923 7</u>	7	8.1	2.5	low p<.001, prev p>.50
2003	<u>ABF17852 20 ABF17851 20 ABF17850 20 ABF17849 20 ABF17848 20 ABF17847 20 ABF17846 20 ABF17845 20 NP_997523 30 NP_997978 41 NP_997977 21 NP_001013287 33 NP_891987 29</u>	13	8.1	3.6	low p<.001, prev p>.50
2004	<u>AAU12246 26 AAU12245 26 AAU12244 26 AAU12243 26 AAU12242 26 AAU12241 26 AAU12240 26 AAU12239 26 AAU12238 26 AAU12237 26 AAU12236 26 AAU12235 26 AAU12234 26 AAU12233 26 AAU12232 26 AAU12231 26 AAU12230 26 AAU12229 26 AAU12228 26 AAU12227 26 AAU12226</u>	34	10.6	3.2	low p<.001, prev p<.02

	<u>26 AAU12225 26 AAU12224 26 AAU12223 26 AAU12222 26 AAU12221 26 AAU12220 26 AAU12219</u> <u>26 BAD72126 42 BAD72124 30 BAD72123 8 BAD72122 16 BAD72121 4 NP_998029 8</u>				
2005	<u>CAJ31050 28 CAJ31049 28 CAJ31048 28 CAJ31047 28 CAJ31046 28 CAJ31045 28 CAJ31044 28 CAJ31043</u> <u>28 CAJ31042 28 CAJ31041 28 CAJ31040 28 CAJ31039 28 CAJ31038 28 CAJ31037 28 CAJ31036 28 CAJ31035</u> <u>28 CAJ31034 28 CAJ31033 28 CAJ31032 28 CAJ31031 28 CAJ31030 28 CAJ31029 28 CAJ31028 28 CAJ31027</u> <u>28 CAJ31026 28 CAJ31025 28 CAJ31024 28 CAJ31023 28 CAJ31022 28 CAJ31021 28 CAJ31020 28 CAJ31019</u> <u>28 CAJ31018 21 CAJ31017 21 CAJ31016 28 CAJ31015 28 CAJ31014 28 CAJ31013 28 CAJ31012 28 CAJ31011</u> <u>28 CAJ31010 28 CAJ31009 28 CAJ31008 28 CAJ31007 28 CAJ31006 28 CAJ31005 28 CAJ31004 28 CAJ31003</u> <u>28 CAJ31002 28 CAJ31001 28 CAJ31000 28 CAJ30999 28 CAJ30998 28 CAJ30997 28 CAJ30996 28 CAJ30995</u> <u>28 CAJ30994 28 CAJ30993 28 CAJ30992 28 CAJ30991 28 BAE78962 17 BAE78961 16</u>	62	5.6	0.7	low p<.001, prev p<.001
2006	<u>ABN13930 16 ABN13929 16 ABN13928 16 ABN13927 16 ABN13926 2 ABD96102 4 ABD64588 47 ABD64587</u> <u>47 ABD64586 47 ABD64585 47 ABD64584 47 ABD64583 47 ABD64582 47 ABD64581 47 ABD64580</u> <u>47 ABD64579 47</u>	16	7.8	2.8	low p<.001, prev p<.002
2007	<u>ABQ42711 36</u>	1	16.2	0.0	prev p<.001

What is claimed is:

Claim 1. A method of identifying a first virus, first organism or first malignancy with a higher lethality than at least one second virus of the same species as the first virus, second organism of the same species as the first organism, or second malignancy of the same species as the first malignancy, which method comprises comparing the Replikin Count of the Replikin Peak Gene of the first virus, first organism or first malignancy to the Replikin Count of the Replikin Peak Gene of at least one second virus, second organism, or second malignancy to determine that the virus, organism or malignancy with the higher Replikin Count is the more lethal.

Claim 2. The method of claim 1 wherein said first malignancy is a lung malignancy, a brain malignancy, a breast malignancy, an ovarian malignancy, or a lymph malignancy.

Claim 3. The method of claim 2 wherein said first malignancy is a non-small cell lung carcinoma.

Claim 4. The method of claim 1 wherein said first organism is a *Mycobacterium tuberculosis*, *Mycobacterium mucogenicum*, *Staphylococcus aureus*, or *Plasmodium falciparum*.

Claim 5. The method of claim 1 wherein said first virus is influenza virus, foot and mouth disease virus, west nile virus, porcine respiratory and reproductive syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus, hemorrhagic septicemia virus, or tobacco mosaic virus.

Claim 6. The method of claim 5 wherein said first virus is *Influenza A* virus strain H1N1, H2N2, H3N2, H5N1, or H3N8.

Claim 7. An isolated or synthesized Replikin Peak Gene of a virus, organism or malignancy wherein said Replikin Peak Gene is identified as the portion of the genome, protein or protein fragment of a virion of the virus, a cell of the organism or a malignant cell of the malignancy consisting of the highest number of continuous Replikin sequences per 100 amino acids as compared to other portions of the genome, protein or protein fragment of the virion of the virus, the cell of the organism or the malignant cell of the malignancy.

Claim 8. The isolated or synthesized Replikin Peak Gene of claim 7 wherein the Replikin Peak Gene is the portion of a protein or protein fragment consisting of the highest number of continuous Replikin sequences per 100 amino acids as compared to all other proteins or protein fragments in the virion of the virus, in the cell of the organism or in the malignant cell of the malignancy.

Claim 9. The isolated or synthesized Replikin Peak Gene of claim 7 wherein said Replikin Peak Gene is isolated from a lung malignancy, a brain malignancy, a breast malignancy, an ovarian malignancy, or a lymph malignancy.

Claim 10. The isolated or synthesized Replikin Peak Gene of claim 9 wherein said Replikin Peak Gene is isolated from a non-small cell lung carcinoma or glioblastoma multiforme.

Claim 11. The isolated or synthesized Replikin Peak Gene of claim 7 wherein said Replikin Peak Gene is isolated from *Mycobacterium tuberculosis*, *Mycobacterium mucogenicum*, *Staphylococcus aureus*, or *Plasmodium falciparum*.

Claim 12. The isolated or synthesized Replikin Peak Gene of claim 7 wherein said Replikin Peak Gene is isolated from influenza virus, foot and mouth disease virus, west nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus, hemorrhagic septicemia virus, or tobacco mosaic virus.

Claim 13. The isolated or synthesized Replikin Peak Gene of claim 12 wherein the influenza virus is an *Influenza A* virus.

Claim 14. The isolated or synthesized Replikin Peak Gene of claim 13 wherein said *Influenza A* virus is H1N1, H2N2, H3N2, H5N1 or H3N8.

Claim 15. The isolated or synthesized Replikin Peak Gene of claim 13 wherein said Replikin Peak Gene is isolated from the pB1 gene area of an influenza virus.

Claim 16. The isolated or synthesized Replikin Peak Gene of claim 12 wherein said virus is a foot and mouth disease virus.

Claim 17. The isolated or synthesized Replikin Peak Gene of claim 16 wherein said Replikin Peak Gene is identified within the VP1 gene of said foot and mouth disease virus.

Claim 18. The isolated or synthesized Replikin Peak Gene of claim 12 wherein said virus is west Nile virus.

Claim 19. The isolated or synthesized Replikin Peak Gene of claim 18 wherein said Replikin Peak Gene is isolated from the envelope protein of said west Nile virus.

Claim 20. The isolated or synthesized Replikin Peak Gene of claim 12 wherein said virus is porcine respiratory and reproductive syndrome virus.

Claim 21. The isolated or synthesized Replikin Peak Gene of claim 20 wherein said Replikin Peak Gene is isolated from a nucleocapsid protein of porcine respiratory and reproductive syndrome virus.

Claim 22. The isolated or synthesized Replikin Peak Gene of claim 12 wherein said virus is porcine circovirus.

Claim 23. The isolated or synthesized Replikin Peak Gene of claim 22 wherein said Replikin Peak Gene is isolated from a replicase protein of porcine circovirus.

Claim 24. The isolated or synthesized Replikin Peak Gene of claim 12 wherein said virus is white spot syndrome virus.

Claim 25. The isolated or synthesized Replikin Peak Gene of claim 24 wherein said Replikin Peak Gene is isolated from a ribonucleotide reductase protein of white spot syndrome virus.

Claim 26. The isolated or synthesized Replikin Peak Gene of claim 12 wherein said virus is tobacco mosaic virus.

Claim 27. The isolated or synthesized Replikin Peak Gene of claim 12 wherein said virus is hemorrhagic septicemia virus.

Claim 28. The isolated or synthesized Replikin Peak Gene of claim 27 wherein said Replikin Peak Gene is isolated from a glycoprotein in hemorrhagic septicemia virus.

Claim 29. The isolated or synthesized Replikin Peak Gene of claim 7 wherein said Replikin Peak Gene comprises a sequence of SEQ ID NO:1741, SEQ ID NO: 3664, SEQ ID NO:3660, SEQ ID NO:3665, SEQ ID NO:1996, SEQ ID NO:1665, SEQ ID NO:1684, SEQ ID NO:1701, SEQ ID NO:546, SEQ ID NO:124, SEQ ID NO:130, SEQ ID NO:311, SEQ ID NOS:341-344, SEQ ID NO:286, SEQ ID NO:287, SEQ ID NO:288, SEQ ID NO:289, SEQ ID NO:290, SEQ ID NOS:233-238, SEQ ID NO:415, SEQ ID NO:421, SEQ ID NO:438, SEQ ID NO:451, SEQ ID NO:462, SEQ ID NO:498, SEQ ID NO:669, SEQ ID NO:1168, SEQ ID NO: 1531, SEQ ID NO:1548, positions 81-204 of SEQ ID NO:3787, or SEQ ID NO:1939.

Claim 30. An immunogenic composition comprising the isolated or synthesized Replikin Peak Gene of claim 7.

Claim 31. The immunogenic composition comprising a Replikin sequence isolated from a Replikin Peak Gene of claim 7 wherein said Replikin sequence is SEQ ID NOS:2902-2925, SEQ ID NOS: 2312-2544, SEQ ID NOS:2701-2711, SEQ ID NOS: 2713-2718, SEQ ID NOS:3282-3285, 3287-3291, 3293, 3295, 3297, 3299, 3300, 3302, 3304, 3306, and 3308, SEQ ID NOS:1685-1691, SEQ ID NOS:1702-1717, SEQ ID NO:106, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NOS:125-129, SEQ ID NOS:131-156, SEQ ID NOS:233-244, SEQ ID NOS:286-290, SEQ ID NOS:312-323, SEQ ID NOS:354-366, SEQ ID NOS:368-380, SEQ ID NOS:383-393, SEQ ID NOS:395-401, SEQ ID NOS:403-414, SEQ ID NOS:291-307, SEQ ID NOS:308-310, SEQ ID NOS:324-327, SEQ ID NOS:328-340, SEQ ID NOS:416-419, SEQ ID NOS:422-437, SEQ ID NOS: 440-445, SEQ ID NOS:452-457, SEQ ID NOS:464-476, SEQ ID NOS:482-484 and SEQ ID NOS:487-492, SEQ ID NOS:547-562, SEQ ID NOS:663-667, SEQ ID NOS:670-1166, SEQ ID NOS:1169-1529, SEQ ID NOS:1532-1542, SEQ ID NO:1548, positions 81-204 of SEQ ID NO:3787, or SEQ ID NOS1637-1663.

Claim 32. A computer readable medium having stored thereon instructions which, when executed, cause the processor to perform a method for identifying a Replikin Peak Gene of a virus, organism or malignancy comprising identifying, within amino acid sequences or

nucleic acid sequences that encode amino acid sequences of said virus, organism or malignancy, the portion of the genome, or protein or protein fragment of said virus, said organism or said malignancy consisting of the highest number of continuous Replikin sequences per 100 amino acids as compared to other portions of the genome, or protein or protein fragment of the virus, organism or malignancy.

Claim 33. The computer readable medium of claim 32 further comprising instructions which, when executed, cause the processor to perform a method for predicting an increase in lethality or virulence of said virus, organism or malignancy that comprises said identified Replikin Peak Gene or an outbreak of said virus or organism that comprises said identified Replikin Peak Gene comprising:

- (1) determining that the Replikin Count of said Replikin Peak Gene of claim 32 or that the Replikin Count of a protein or gene area comprising said Replikin Peak Gene of claim 32 is higher than another Replikin Peak Gene or a protein or gene area comprising said other Replikin Peak Gene identified within the genome or within a protein or protein fragment of at least one other virus of the same species as said virus, at least one other organism of the same species as said organism or at least one other malignancy of the same type as said malignancy wherein said other virus, said other organism or said other malignancy is isolated at an earlier time point than said virus, said organism or said malignancy, and
- (2) predicting an increase in lethality or virulence of said virus, organism or malignancy or predicting an outbreak of said virus or organism.

Claim 34. A method of predicting the strain, the host or the geographic region of an outbreak or increase in lethality or virulence of a virus or organism comprising

- (1) identifying a Replikin Peak Gene or a protein or gene area comprising a Replikin Peak Gene within the genome of a first virus or organism of a first strain, from a first host, or isolated from a first geographic region or within a protein or protein fragment of the first virus or organism that has a higher Replikin Count than a Replikin Peak Gene or protein or gene area comprising a Replikin Peak Gene identified within the genome or within a protein or protein fragment of at least one second virus of the same species as the first virus or at least one second organism of the same species as

the first organism wherein said first virus or said first organism is isolated at a later time point than said first virus or said first organism and is the same strain, from the same or another host or isolated from the same or another geographic region as the first virus or first organism, and

- (2) predicting an outbreak or an increase in lethality or virulence of said first strain, in said first host, or within said first geographic region of said first virus or organism.

Claim 35. The method of claim 34 wherein said protein or gene area comprising said Replikin Peak Gene within the genome of a first virus or organism is identified as having a higher Replikin Count than said protein or gene area comprising a Replikin Peak Gene identified within the genome or within a protein or protein fragment of said at least one second virus or organism.

Claim 36. The method of claim 34 wherein said first virus or said first organism is isolated at least six months to three years later than said second organism or said second virus.

Claim 37. The method of claim 34 wherein said first organism or said first virus is *Mycobacterium tuberculosis*, *Mycobacterium mucogenicum*, *Staphylococcus aureus*, and *Plasmodium falciparum*, influenza virus, foot and mouth disease virus, west nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus in tomato plants, hemorrhagic septicemia virus in fish, or tobacco mosaic virus.

Claim 38. The method of claim 37 wherein said *Staphylococcus aureus* is methicillin-resistant.

Claim 39. The method of claim 37 wherein said influenza virus is a strain of *Influenza A* virus.

Claim 40. The method of claim 39 wherein said *Influenza A* virus is H1N1, H2N2, H3N2, H5N1 or H3N8.

Claim 41. The method of claim 39 wherein said protein or gene area comprising said Replikin Peak Gene is the pB1 gene area of the influenza virus.

Claim 42. The method of claim 37 wherein said protein or gene area is a nucleocapsid protein of said porcine respiratory and reproductive syndrome virus.

Claim 43. The method of claim 37 wherein said protein or gene area is a VP1 protein of said foot and mouth disease virus.

Claim 44. The method of claim 37 wherein said protein or gene area is an envelope protein of said west nile virus.

Claim 45. The method of claim 37 wherein said protein or gene area is a nucleocapsid protein of said porcine reproductive and respiratory syndrome virus.

Claim 46. The method of claim 37 wherein said protein or gene area is an ATP-ase of said *Plasmodium falciparum*.

Claim 47. The method of claim 37 wherein said protein or gene area is a replicase protein of said porcine circovirus.

Claim 48. The method of claim 37 wherein said protein or gene area is a ribonucleotidase of said white spot syndrome virus.

Claim 49. A method of identifying a first virus, organism or malignancy associated with higher lethality, higher virulence or more rapid replication than a second virus of the same species as the first virus, a second organism of the same species as the first organism or a second malignancy of the same type as the first malignancy comprising identifying a Replikin Peak Gene encoded within the genome of at least one virion of the first virus, or at least one cell of the first organism, or at least one malignant cell of the first malignancy, or within a protein or protein fragment of at least one virion of the first virus, or at least one cell of the first organism, or at least one malignant cell of the first malignancy that has a higher Replikin Count than a Replikin Peak Gene identified encoded within the genome of at least one virion of the second virus, or at least one cell of the second organism, or at least one malignant cell of the second malignancy or within a protein or protein fragment of at least

one virion of the second virus, or at least one cell of the second organism, or at least one malignant cell of the second malignancy wherein said first virus, first organism or first malignancy has higher lethality, higher virulence or more rapid replication than said second virus, second organism or second malignancy, and wherein the Replikin Peak Gene is defined as a protein or protein fragment having the highest concentration of continuous Replikin sequences per 100 amino acids as compared to the remaining proteins or protein fragments in the same virion of the virus, the same cell of the organism, or the same malignant cell, or the portion of the genome encoding the protein or protein fragment.

Claim 50. A method of identifying a first virus, first organism or first malignancy with a higher lethality than at least one second virus of the same species as the first virus, second organism of the same species as the first organism or second malignancy of the same species as the first malignancy comprising comparing the Replikin Count of the whole genome of said virus, organism or malignancy to the Replikin Count of the whole genome of said at least one second virus, second organism, or second malignancy to determine that the virus, organism or malignancy with the higher Replikin Count is the more lethal.

Claim 51. The method of claim 50 wherein said first virus is a coronavirus, a foot and mouth disease virus, a white spot syndrome virus, a taura syndrome virus, a porcine circovirus, or an influenza virus.

Claim 52. The method of claim 50 wherein said first virus is an H5N1 strain of influenza virus.

Claim 53. A method for obtaining an isolated or synthesized Replikin Peak Gene of a virus, organism or malignancy for diagnosis, prevention or treatment of an infection of said virus or said organism or for diagnosis, prevention or treatment of said malignancy comprising:

- (1) obtaining a plurality of isolates of virus of the same species, a plurality of organisms of the same species, or a plurality of malignancies of the same type;
- (2) analyzing the protein sequences or protein sequence fragments of each individual isolate of the plurality of isolates of virus, a cell of each individual organism of the

- plurality of organisms, or a malignant cell of each individual malignancy of the plurality of malignancies for the presence and concentration of Replikin sequences;
- (3) identifying the protein sequence or the protein sequence fragment having the highest concentration of continuous Replikin sequences in the malignant cell of each individual malignancy, the cell of each individual organism or each individual virus isolate;
 - (4) selecting the protein sequence or protein sequence fragment having the highest concentration of continuous Replikin sequences among the plurality of isolates of virus, the plurality of organisms, or the plurality of malignancies;
 - (5) identifying the amino acid sequence of the selected protein sequence or protein sequence fragment as the Replikin Peak Gene of the plurality of virus isolates, organisms or malignancies; and
 - (6) isolating or synthesizing the identified Replikin Peak Gene of at least one of the plurality of virus isolates, organisms or malignancies wherein the isolated or synthesized identified Replikin Peak Gene is useful for diagnosis, prevention or treatment of said infection of said virus or said organism or said malignancy.

Claim 54. An immunogenic composition comprising at least one isolated or synthesized Replikin Peak Gene of claim 53.

Claim 55. The immunogenic composition of claim 54 wherein said isolated or synthesized Replikin Peak Gene is isolated from an emerging strain of a virus or organism.

Claim 56. The immunogenic composition of claim 54 further comprising a pharmaceutically acceptable carrier.

Claim 57. A vaccine comprising at least one isolated or synthesized Replikin Peak Gene of claim 7.

Claim 58. The vaccine of claim 57 wherein said at least one isolated or synthesized Replikin Peak Gene is isolated from an emerging strain of virus or organism.

Claim 59. A vaccine of claim 57 comprising SEQ ID NO:1741, SEQ ID NO: 3664, SEQ ID NO:3660, SEQ ID NO:3665, SEQ ID NO:1996, SEQ ID NO:1665, SEQ ID NO:1684, SEQ ID NO:1701, SEQ ID NO:546, SEQ ID NO:124, SEQ ID NO:130, SEQ ID NO:311, SEQ ID NOS:341-344, SEQ ID NO:286, SEQ ID NO:287, SEQ ID NO:288, SEQ ID NO:289, SEQ ID NO:290, SEQ ID NOS:233-238, SEQ ID NO:415, SEQ ID NO:421, SEQ ID NO:438, SEQ ID NO:451, SEQ ID NO:462, SEQ ID NO:498, SEQ ID NO:669, SEQ ID NO:1168, SEQ ID NO: 1531, SEQ ID NO:1548, positions 81-204 of SEQ ID NO:3787, or SEQ ID NO:1939.

Claim 60. The vaccine of claim 57 wherein said isolated or synthesized Replikin Peak Gene is isolated from a virus.

Claim 61. The vaccine of claim 60 wherein said isolated or synthesized Replikin Peak Gene is isolated from influenza virus, foot and mouth disease virus, west nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus, hemorrhagic septicemia virus, or tobacco mosaic virus.

Claim 62. The vaccine of claim 61 wherein said isolated or synthesized Replikin Peak Gene is isolated from *Influenza A*.

Claim 63. The vaccine of claim 62 wherein said *Influenza A* is H1N1, H2N2, H3N2, H5N1 or H3N8.

Claim 64. The vaccine of claim 57 wherein said isolated or synthesized Replikin Peak Gene is isolated from an organism.

Claim 65. The vaccine of claim 64 wherein said isolated or synthesized Replikin Peak Gene is isolated from *Mycobacterium tuberculosis*, *Mycobacterium mucogenicum*, *Staphylococcus aureus*, or *Plasmodium falciparum*.

Claim 66. The vaccine of claim 65 wherein said *Staphylococcus aureus* is methicillin-resistant.

Claim 67. The vaccine of claim 57 wherein said isolated or synthesized Replikin Peak Gene is isolated from a malignancy.

Claim 68. The vaccine of claim 67 wherein said isolated or synthesized Replikin Peak Gene isolated is isolated from a lung malignancy, a brain malignancy, a breast malignancy or a lymph malignancy.

Claim 69. The vaccine of claim 68 wherein said isolated or synthesized Replikin Peak Gene is isolated from a non-small cell lung carcinoma.

Claim 70. The vaccine of claim 68 wherein said isolated or synthesized Replikin Peak Gene is isolated from glioblastoma multiforme.

Claim 71. An immunogenic composition comprising a Replikin Peak Gene of claim 29.

Claim 72. The immunogenic composition of claim 71 further comprising a pharmaceutically acceptable carrier.

Claim 73. An isolated or synthesized Replikin sequence isolated from a protein or protein fragment of the Replikin Peak Gene of claim 7 or isolated from a protein comprising the Replikin Peak Gene of claim 7.

Claim 74. The Replikin sequence of claim 73 wherein said Replikin Peak Gene is isolated from *Mycobacterium tuberculosis*, *Mycobacterium mucogenicum*, *Staphylococcus aureus*, or a *Plasmodium falciparum*.

Claim 75. The Replikin sequence of claim 74 wherein said Replikin Peak Gene is isolated from *Mycobacterium mucogenicum*.

Claim 76. The Replikin sequence of claim 75 wherein said Replikin sequence is SEQ ID NOS:2902-2925.

Claim 77. The Replikin sequence of claim 74 wherein said Replikin Peak Gene is isolated from *Plasmodium falciparum*.

Claim 78. The Replikin sequence of claim 77 wherein said Replikin Peak Gene is one of SEQ ID NOS: 2312-2544, SEQ ID NOS:2701-2711, SEQ ID NOS: 2713-2718, SEQ ID NOS: 3282-3285, 3287-3291, 3293, 3295, 3297, 3299, 3300, 3302, 3304, 3306, or SEQ ID NO: 3308.

Claim 79. The Replikin sequence of claim 73 wherein said Replikin Peak Gene is isolated from influenza virus, foot and mouth disease virus, west nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus, hemorrhagic septicemia virus, or tobacco mosaic virus.

Claim 80. The Replikin sequence of claim 79 wherein said influenza virus is *Influenza A* virus.

Claim 81. The Replikin sequence of claim 80 wherein said *Influenza A* virus is H1N1, H2N2, H3N2, H5N1 or H3N8.

Claim 82. The Replikin sequence of claim 81 wherein said *Influenza A* virus is H5N1 and the Replikin sequence is one of SEQ ID NOS:1685-1691, SEQ ID NOS:1702-1716 or SEQ ID NO:1717.

Claim 83. The Replikin sequence of claim 81 wherein said *Influenza A* virus is H3N8 and the Replikin sequence is one of SEQ ID NOS:547-561 or SEQ ID NO:562.

Claim 84. The Replikin sequence of claim 79 wherein said Replikin Peak Gene is isolated from foot and mouth disease virus.

Claim 85. The Replikin sequence of claim 84 wherein said Replikin sequence is one of SEQ ID NO:106, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NOS:125-129, SEQ ID NOS:131-155 or SEQ ID NO:156.

Claim 86. The Replikin sequence of claim 79 wherein said Replikin Peak Gene is isolated from west nile virus.

Claim 87. The Replikin sequence of claim 86 wherein said Replikin sequence is one of SEQ ID NOS:233-243 or SEQ ID NO:244.

Claim 88. The Replikin sequence of claim 79 wherein said Replikin Peak Gene is isolated from PRRSV.

Claim 89. The Replikin sequence of claim 88 wherein said Replikin sequence is one of SEQ ID NOS:286-290, SEQ ID NOS:312-323, SEQ ID NOS:354-366, SEQ ID NOS:368-380, SEQ ID NOS:383-393, SEQ ID NOS:395-401, SEQ ID NOS:403-413 or SEQ ID NO:414.

Claim 90. The Replikin sequence of claim 79 wherein said Replikin Peak Gene is isolated from PCV.

Claim 91. The Replikin sequence of claim 90 wherein said Replikin sequence is one of SEQ ID NOS:291-307, SEQ ID NOS:308-310, SEQ ID NOS:324-327, SEQ ID NOS:328-340, SEQ ID NOS:416-419, SEQ ID NOS:422-437, SEQ ID NOS: 440-445, SEQ ID NOS:452-457, SEQ ID NOS:464-476, SEQ ID NOS:482-484, SEQ ID NOS:487-491 or SEQ ID NO:492.

Claim 92. The Replikin sequence of claim 79 wherein said Replikin Peak Gene is isolated from white spot syndrome virus.

Claim 93. The Replikin sequence of claim 92 wherein said Replikin sequence is one of SEQ ID NOS:663-667, SEQ ID NOS:670-1166, SEQ ID NOS:1169-1529, SEQ ID NOS:1532-1542 and SEQ ID NO:1548.

Claim 94. A vaccine for prevention and/or treatment of a viral or organismal infection or a malignancy wherein the vaccine comprises at least one isolated or synthesized Replikin sequence within a protein or protein fragment of a Replikin Peak Gene or a protein comprising a Replikin Peak Gene identified in said virus, organism, or malignancy.

Claim 95. The vaccine of claim 94 wherein said at least one isolated or synthesized Replikin sequence is one of SEQ ID NOS:2902-2925, SEQ ID NOS: 2312-2544, SEQ ID NOS:2701-2711, 2713-2718, SEQ ID NOS: 3282-3285, 3287-3291, 3293, 3295, 3297, 3299, 3300, 3302, 3304, 3306, 3308, SEQ ID NOS:1685-1691, SEQ ID NOS:1702-1717, SEQ ID NOS:547-562, SEQ ID NO:106, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NOS:125-129, and SEQ ID NOS:131-156, SEQ ID NOS:233-244, SEQ ID NOS:286-290, SEQ ID

NOS:312-323, SEQ ID NOS:354-366, SEQ ID NOS:368-380, SEQ ID NOS:383-393, SEQ ID NOS:395-401, SEQ ID NOS:403-414, SEQ ID NOS:291-307, SEQ ID NOS:308-310, SEQ ID NOS:324-327, SEQ ID NOS:328-340, SEQ ID NOS:416-419, SEQ ID NOS:422-437, SEQ ID NOS: 440-445, SEQ ID NOS:452-457, SEQ ID NOS:464-476, SEQ ID NOS:482-484 SEQ ID NOS:487-492, SEQ ID NOS:663-667, SEQ ID NOS:670-1166, SEQ ID NOS:1169-1529, SEQ ID NOS:1532-1542, SEQ ID NO:1548, SEQ ID NOS:1637-1662, or SEQ ID NO:1663.

Claim 96. The vaccine of claim 94 for prevention and/or treatment of a viral infection.

Claim 97. The vaccine of claim 96 wherein said viral infection is caused by influenza virus, foot and mouth disease virus, west nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus, hemorrhagic septicemia virus, or tobacco mosaic virus.

Claim 98. The vaccine of claim 97 wherein the influenza virus is *Influenza A* virus.

Claim 99. The vaccine of claim 98 wherein the *Influenza A* virus is a H1N1, H2N2, H3N2, H5N1 or H3N8.

Claim 100. The vaccine of claim 97 wherein the virus is hemorrhagic septicemia virus

Claim 101. The vaccine of claim 94 for prevention and/or treatment of an organismal infection.

Claim 102. The vaccine of claim 101 wherein the organismal infection is caused by *Mycobacterium tuberculosis*, *Mycobacterium mucogenicum*, *Staphylococcus aureus*, or *Plasmodium falciparum*.

Claim 103. The vaccine of claim 102 wherein said *Staphylococcus aureus* is methicillin-resistant.

Claim 104. The vaccine of claim 94 wherein the malignancy is a lung malignancy, a brain malignancy, a breast malignancy, an ovarian malignancy, or a lymph malignancy.

Claim 105. The vaccine of claim 104 wherein the malignancy is non-small cell lung carcinoma or glioblastoma multiforme.

Claim 106. An immunogenic compound comprising at least one isolated or synthesized Replikin sequence within the protein or protein fragment of a Replikin Peak Gene or within a protein comprising a Replikin Peak Gene wherein said Replikin Peak Gene is identified in a virus, an organism or a malignancy.

Claim 107. The immunogenic compound of claim 106 further comprising a pharmaceutically acceptable carrier.

Claim 108. A method of stimulating the immune system, comprising administering in an animal at least one isolated or synthesized Replikin sequence identified within a protein or protein fragment of a Replikin Peak Gene or within a protein or gene area comprising a Replikin Peak Gene identified in a virus, organism, or malignancy.

Claim 109. The method of claim 108 wherein said animal is a human.

Claim 110. An antibody to at least one isolated or synthesized Replikin sequence within a protein or protein fragment of Replikin Peak Gene or within a protein or gene area comprising a Replikin Peak Gene.

Claim 111. A method of identifying a lethal strain of malignancy, organism or virus comprising:

- (1) obtaining a plurality of isolates of said malignancy, organism or virus;
- (2) identifying the Replikin Peak Gene in each isolate of the plurality of isolates of said malignancy, organism or virus;
- (3) analyzing the amino acid sequence of a protein or protein fragment of the Replikin Peak Gene of each isolate of the plurality of isolates for the presence and concentration of Replikin sequences;
- (4) comparing the concentrations of Replikin sequences in each of the proteins or protein fragments of the Replikin Peak Gene of each isolate of the plurality of isolates to the

- concentration of Replikin sequences in each of the proteins or protein fragments of the Replikin Peak Gene of each of the other isolates of the plurality of isolates; and
- (5) identifying the isolate having the highest concentration of continuous Replikin sequences in the protein or protein fragment of the Replikin Peak Gene as a virulent or lethal strain of said malignancy, organism or virus.

Claim 112. A method of selecting a peptide from a malignancy, organism or virus for inclusion in a preventive or therapeutic vaccine or immunogenic compound for a malignancy, organism or virus comprising identifying at least one difference in the amino acid sequence of an otherwise conserved Replikin sequence or Replikin Peak Gene between at least two isolates of said malignancy, organism or virus and correlating the identified at least one difference in the amino acid sequence with the highest virulence, morbidity or host mortality among the at least two isolates and selecting an otherwise conserved Replikin sequence, Replikin Peak Gene or Replikin sequence within a Replikin Peak Gene having the identified at least one amino acid sequence difference as the peptide for inclusion in a preventive or therapeutic vaccine or immunogenic compound.

Claim 113. The method of claim 112 further predicting the isolate comprising the selected conserved Replikin sequence or Replikin Peak Gene having the at least one difference in the amino acid sequence to be a lethal isolate of said malignancy, organism or virus.

Claim 114. The method of claim 113 wherein said malignancy, organism or virus is a malignancy.

Claim 115. The method of claim 114 wherein said malignancy is a lung malignancy, a brain malignancy, a breast malignancy or a lymph malignancy.

Claim 116. The method of claim 115 wherein said malignancy is a non-small cell lung carcinoma or a glioblastoma multiforme.

Claim 117. The method of claim 112 wherein said malignancy, organism or virus is an organism.

- Claim 118. The method of claim 117 wherein said organism is *Mycobacterium tuberculosis*, *Mycobacterium mucogenicum*, *Staphylococcus aureus*, or *Plasmodium falciparum*.
- Claim 119. The method of claim 118 wherein said *Staphylococcus aureus* is methicillin-resistant.
- Claim 120. The method of claim 112 wherein said malignancy, organism or virus is a virus.
- Claim 121. The method of claim 120 wherein said virus is influenza virus, foot and mouth disease virus, west nile virus, porcine reproductive and respiratory syndrome virus, porcine circovirus, white spot syndrome virus, taura syndrome virus, coronavirus, ebola virus, gemini leaf curl virus, hemorrhagic septicemia virus or tobacco mosaic virus.
- Claim 122. A method of determining a source of a case of lung malignancy comprising identifying at least one peptide in a Replikin Peak Gene of a lung cancer cell that is also present in a Replikin Peak Gene of an isolate of tobacco mosaic virus, wherein the peptide is involved with the source of the lung malignancy.
- Claim 123. The method of claim 122 wherein a plurality of peptides is identified in the Replikin Peak Gene of the lung cancer cell wherein each one of the plurality of peptides is also identified in the Replikin Peak Gene of an isolate of tobacco mosaic virus.
- Claim 124. The method of claim 122 wherein the at least one peptide in the Replikin Peak Gene of the lung cancer cell and the at least one peptide in Replikin Peak Gene of the isolate of tobacco mosaic virus is a peptide of about 10 amino acids or less comprising at least two lysines and at least one histidine.
- Claim 125. The method of claim 124 wherein the at least one peptide in the Replikin Peak Gene of the lung cancer cell and the at least one peptide in Replikin Peak Gene of the isolate of tobacco mosaic virus is a peptide of about 10 amino acids or less comprising at least three lysines and at least one histidine.

Claim 126. The method of claim 125 wherein the at least one peptide in the Replikin Peak Gene of the lung cancer cell and the at least one peptide in the Replikin Peak Gene of the isolate of tobacco mosaic virus is about 7 amino acids or less comprising at least three lysines and at least one histidine.

Claim 127. The method of claim 126 wherein the at least one peptide in the Replikin Peak Gene of the lung cancer cell and the at least one peptide in the Replikin Peak Gene of the isolate of tobacco mosaic virus is about 4 amino acids comprising three lysines and one histidine.

Claim 128. The method of claim 127 wherein the at least one peptide in the Replikin Peak Gene of the lung cancer cell and in the Replikin Peak Gene of the isolate of tobacco mosaic virus is KHKK (SEQ ID NO:1584).

Claim 129. The method of claim 128 wherein more than one KHKK (SEQ ID NO:1584) peptide is identified in the Replikin Peak Gene of the lung cancer cell and in the Replikin Peak Gene of the isolate of tobacco mosaic virus.

Claim 130. The method of claim 129 wherein at least 10 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the lung cancer cell and at least 10 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the isolate of tobacco mosaic virus.

Claim 131. The method of claim 130 wherein at least 20 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the lung cancer cell and at least 20 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the isolate of tobacco mosaic virus.

Claim 132. The method of claim 131 wherein at least 30 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the lung cancer cell and at least 30 KHKK peptides are identified in the Replikin Peak Gene of the isolate of tobacco mosaic virus.

Claim 133. The method of claim 132 wherein at least 50 KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the lung cancer cell and at least 50

KHKK (SEQ ID NO:1584) peptides are identified in the Replikin Peak Gene of the isolate of tobacco mosaic virus.

Claim 134. A method of identifying a first case of malignancy of the lung having a higher rate of replication, aggressive growth pattern or lethality as compared with a second case of malignancy of the lung comprising identifying a Replikin Peak Gene in a malignant cell from a first case of malignancy of the lung that has a higher Replikin Count in the Replikin Peak Gene than a Replikin Peak Gene identified in a malignant cell from a second case of malignancy of the lung.

Claim 135. The method of claim 134 wherein the first and second cases of malignancy of the lung are non-small cell lung malignancies.

Claim 136. An isolated or synthesized Replikin Peak Gene in a lung malignancy for diagnosis, prevention or treatment of lung cancer by the method comprising:

- (1) obtaining at least one malignant cell from a lung malignancy;
- (2) analyzing the protein sequences or protein sequence fragments of the at least one malignant cell for the presence and concentration of Replikin sequences;
- (3) identifying the protein sequence or the protein sequence fragment having the highest concentration of continuous Replikin sequences in the at least one malignant cell;
- (4) selecting the protein sequence or protein sequence fragment having the highest concentration of continuous Replikin sequences;
- (5) identifying the amino acid sequence of the selected protein sequence or protein sequence fragment as the Replikin Peak Gene; and
- (6) isolating or synthesizing the identified Replikin Peak Gene of the at least one malignant cell, wherein the isolated or synthesized identified Replikin Peak Gene is useful for diagnosis, prevention or treatment of lung cancer.

Claim 137. The method of claim 136 wherein the lung malignancy is a non-small cell lung malignancy.

Claim 138. The method of claim 137 wherein at least one Replikin sequence within the protein or protein fragment of the identified Replikin Peak Gene is isolated or synthesized for diagnosis, prevention or treatment of lung cancer.

Claim 139. The method of claim 138 wherein the at least one Replikin sequence within the protein or protein fragment of the identified Replikin Peak Gene is one of SEQ ID NOS: 1585-1635 of SEQ ID NO:1636.

Claim 140. An immunogenic composition for prevention and treatment of lung cancer, wherein the immunogenic composition comprises at least one isolated or synthesized Replikin sequence within the protein or protein fragment of an identified Replikin Peak Gene.

Claim 141. A method of stimulating the immune system, comprising administering in an animal the at least one isolated or synthesized Replikin sequence identified within the Replikin Peak Gene of the lung malignancy for prevention, treatment or diagnosis of lung cancer in an animal.

Claim 142. The method of claim 141, wherein the animal is a human.

Claim 143. An antibody to at least one isolated or synthesized Replikin sequence within the Replikin Peak Gene sequence of a lung malignancy.

Claim 144. A method of identification of a lethal form of lung cancer comprising:

- (1) obtaining at least one malignant cell from a plurality of lung tumors;
- (2) identifying the Replikin Peak Gene in the at least one malignant cell of each of the plurality of lung tumors;
- (3) analyzing the amino acid sequence of a protein or protein fragment of the Replikin Peak Gene in the at least one malignant cell of each of the plurality of lung tumors for the presence and concentration of Replikin sequences;
- (4) comparing the concentrations of Replikin sequences in each of the proteins or protein fragments of the Replikin Peak Gene in the at least one malignant cell of each of the plurality of lung tumors; and

- (5) identifying the lung tumor having the highest concentration of continuous Replikin sequences in the protein or protein fragment of the Replikin Peak Gene as a lethal form of lung cancer.

Claim 145. A method of identification of a more lethal form of lung cancer among at least two lung cancers, comprising:

- (1) obtaining at least one malignant cell from each of at least two lung cancers;
- (2) identifying the Replikin Peak Gene in the at least one malignant cell of each of the at least two lung cancers;
- (3) analyzing the amino acid sequence of a protein or protein fragment of the Replikin Peak Gene in the at least one malignant cell of each of the at least two lung cancers for the presence and concentration of Replikin sequences;
- (4) comparing the concentrations of Replikin sequences in each of the proteins or protein fragments of the Replikin Peak Gene in the at least one malignant cell of each of the at least two lung cancers; and
- (5) identifying the lung cancer having the highest concentration of continuous Replikin sequences in the protein or protein fragment of the Replikin Peak Gene as the more lethal form of lung cancer.

Claim 146. A method of determining an expected increase in lethality or virulence of a virus or organism which method comprises:

- (1) obtaining a plurality of isolates of said virus or organism wherein each isolate is isolated within a known time period and wherein at least two of said isolates is isolated about six months to about 5 years later than at least two other of said isolates;
- (2) identifying a Replikin Peak Gene in each isolate of said plurality of isolates;
- (3) analyzing the identified Replikin Peak Gene of each isolate of the plurality of isolates to determine the Replikin Count of each Replikin Peak Gene of each isolate of the plurality of isolates, or analyzing a protein, protein fragment, or gene area comprising

the identified Replikin Peak Gene of each isolate of the plurality of isolates to determine the Replikin Count of the protein, protein fragment, or gene area of the plurality of isolates;

- (4) determining a mean Replikin Count within the Replikin Peak Gene or within the protein, protein fragment, or gene area comprising said identified Replikin Peak Gene for each known time period;
- (5) comparing the mean Replikin Count within the Replikin Peak Gene or within the protein, protein fragment, or gene area for each known time period one to another;
- (6) identifying an increase in the mean Replikin Count between at least two known time periods; and
- (7) identifying an expected increase in lethality or virulence of said virus, or organism within about six months to about three years following said identified increase in the mean Replikin Count.

Claim 147: The method of claim 146, wherein the known time period is 1 year.

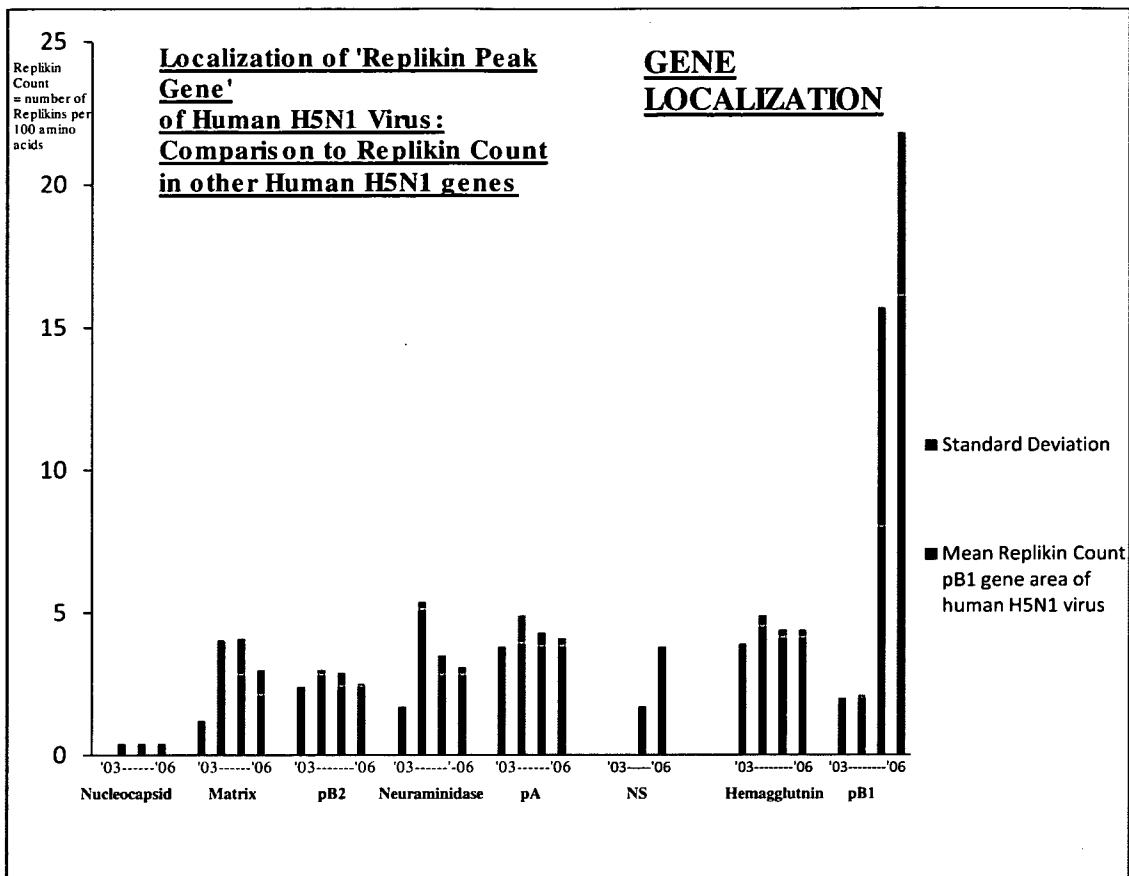
Claim 148: The method of claim 147, wherein the increase in mean Replikin Count occurs over one year.

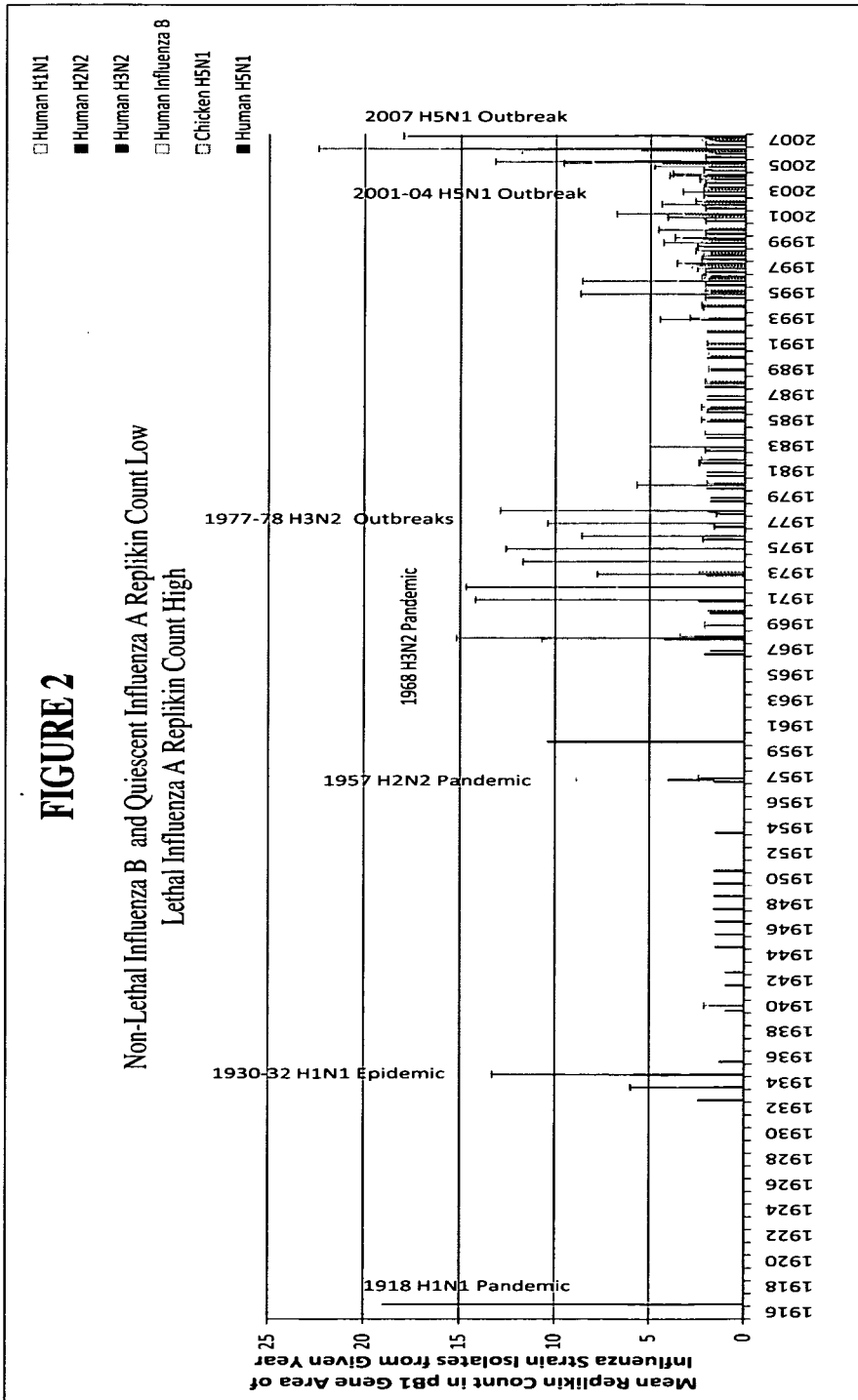
Claim 149. The method of claim 147, wherein the increase in mean Replikin Count occurs over three years.

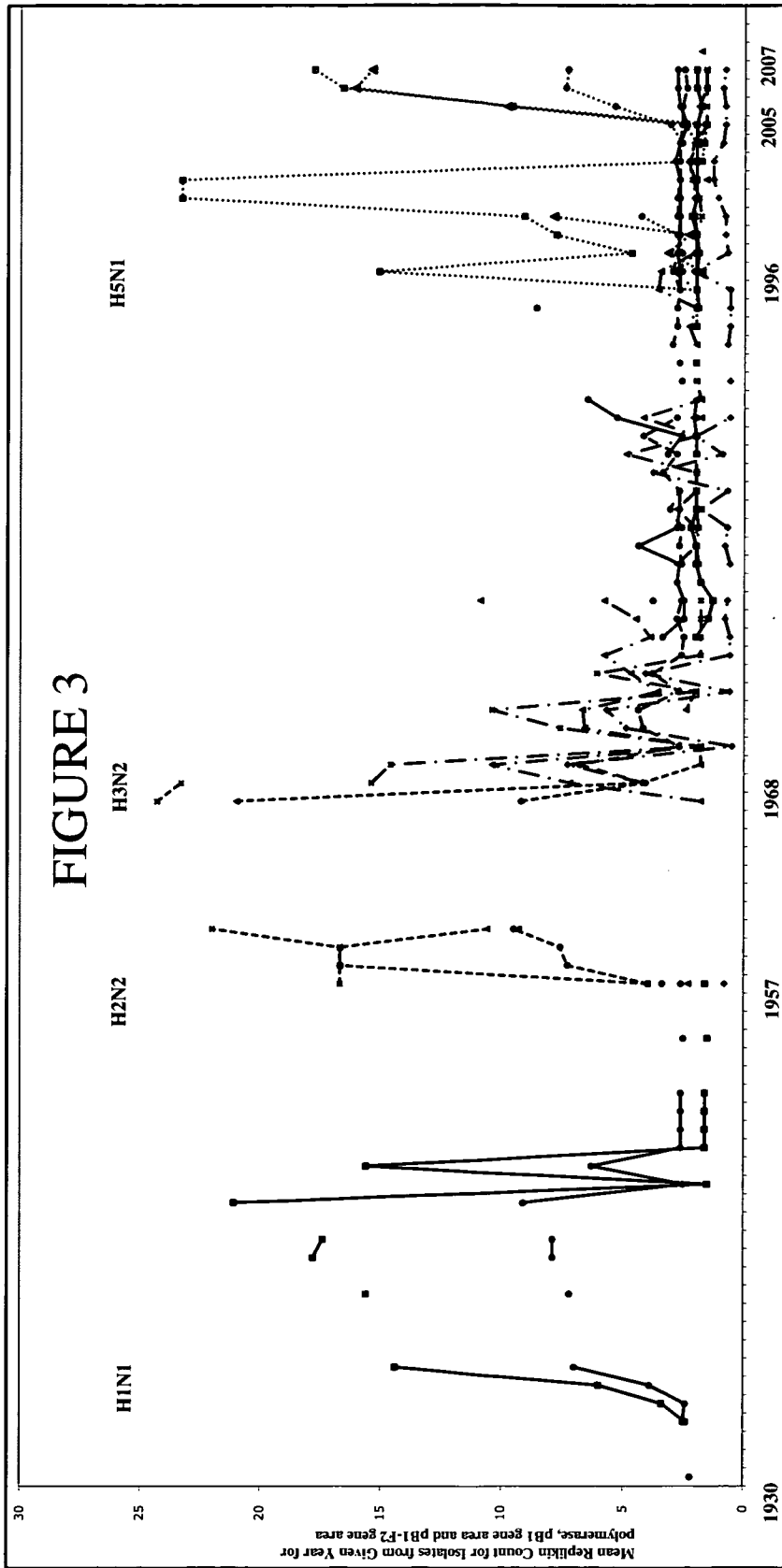
Claim 150: The method of claim 146, wherein said increase in mean Replikin Count is significant between at least two known time periods.

Claim 151: The method of claim 149, wherein said increase in mean Replikin Count has a significance of $p = < 0.001$.

FIGURE 1







Polymerase area (circles), pB1 area (triangles), and pB1-F2 area (squares) of the genome mark successive 'emerging' strains of influenza virus: 1) outbreaks of H1N1(beginning 1930)—solid line; (2) pandemic of H2N2(1957)—short-dashed line, 3) pandemic of H2N3(1968)—long-dashed line, and (4) outbreaks of H5N1(1997 to 2007)—dotted line.

Figure 4 - Relation of Replikin Count of Replikin Peak Gene in pB1 area in Human H5N1 virus to % Human Mortality

Mean Replikin Count of pB1 area in Human H5N1 per year (light gray)

Standard Deviation from Mean Replikin Count (dark gray)

Percent Human Mortality (black)

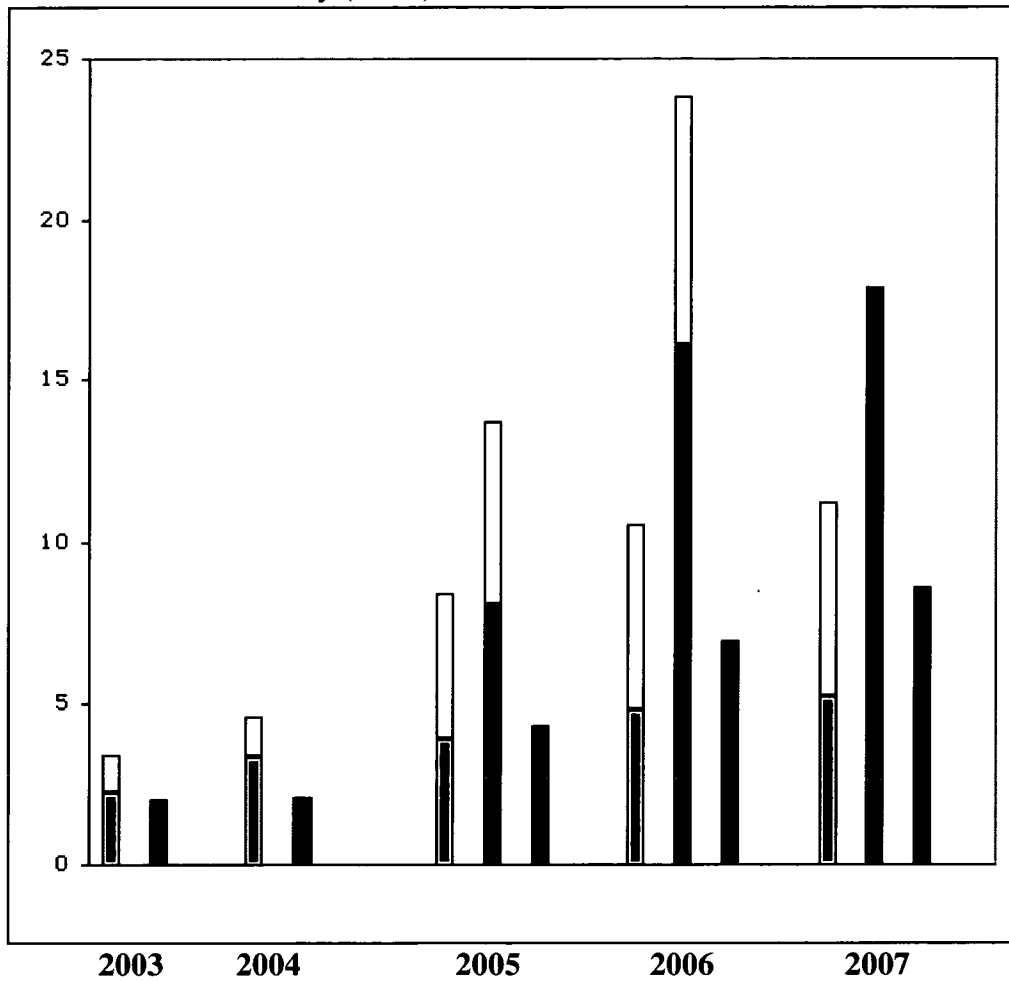


FIGURE 5

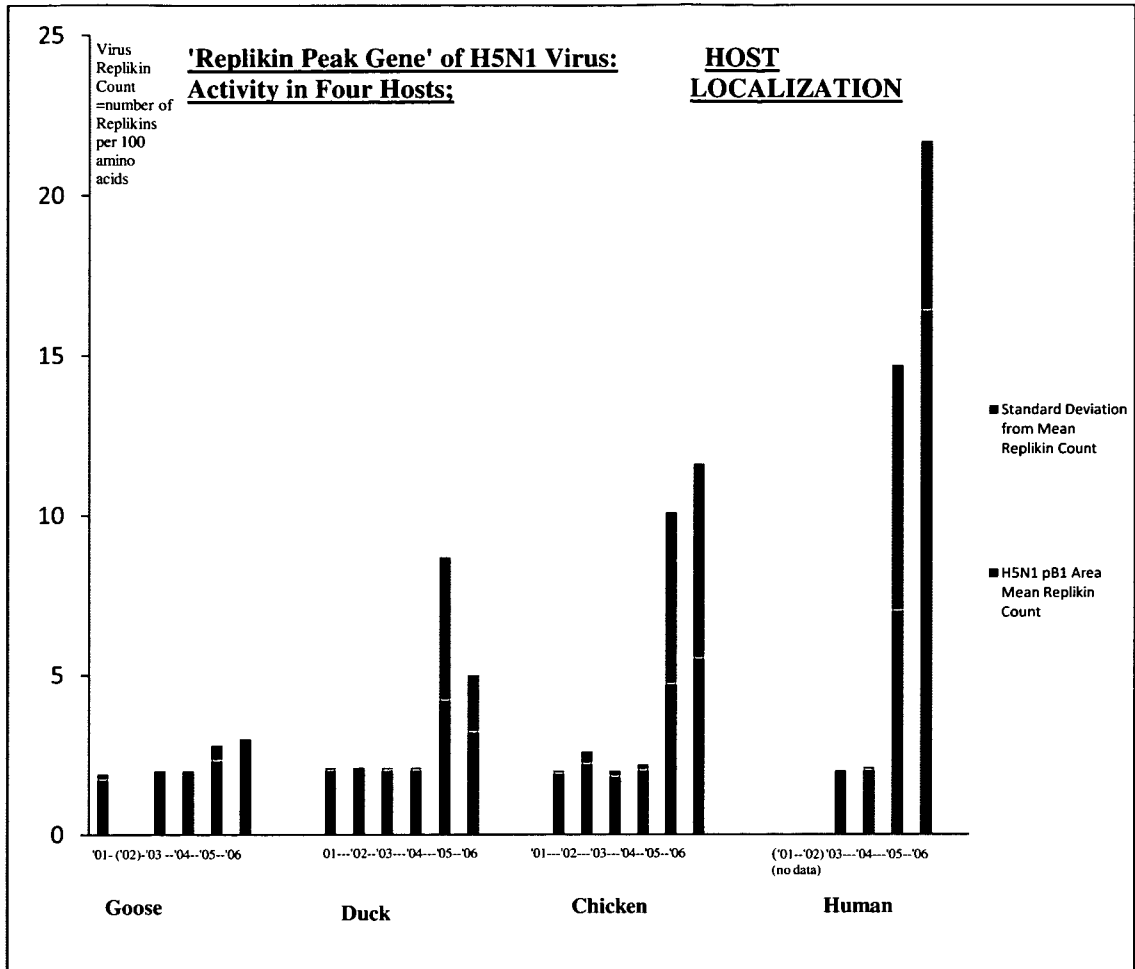


FIGURE 6

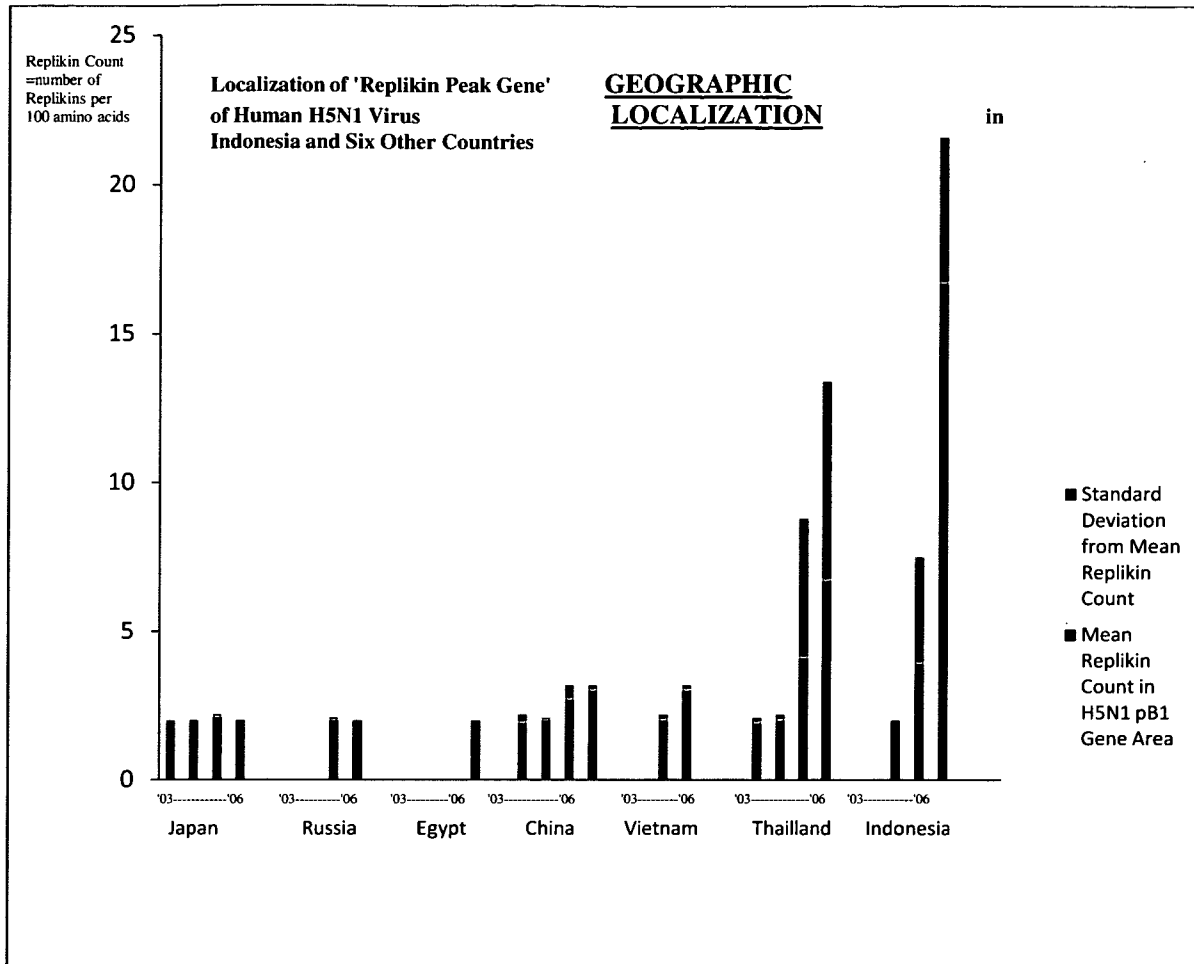


FIGURE 7

Increase in Replikin Count of Replikin Peak Gene (RPG) in pB1 Genomic Area Only, Not in pA or pB2, One to Two Years in Advance of Equine Influenza Epidemics (E) 1977-2007

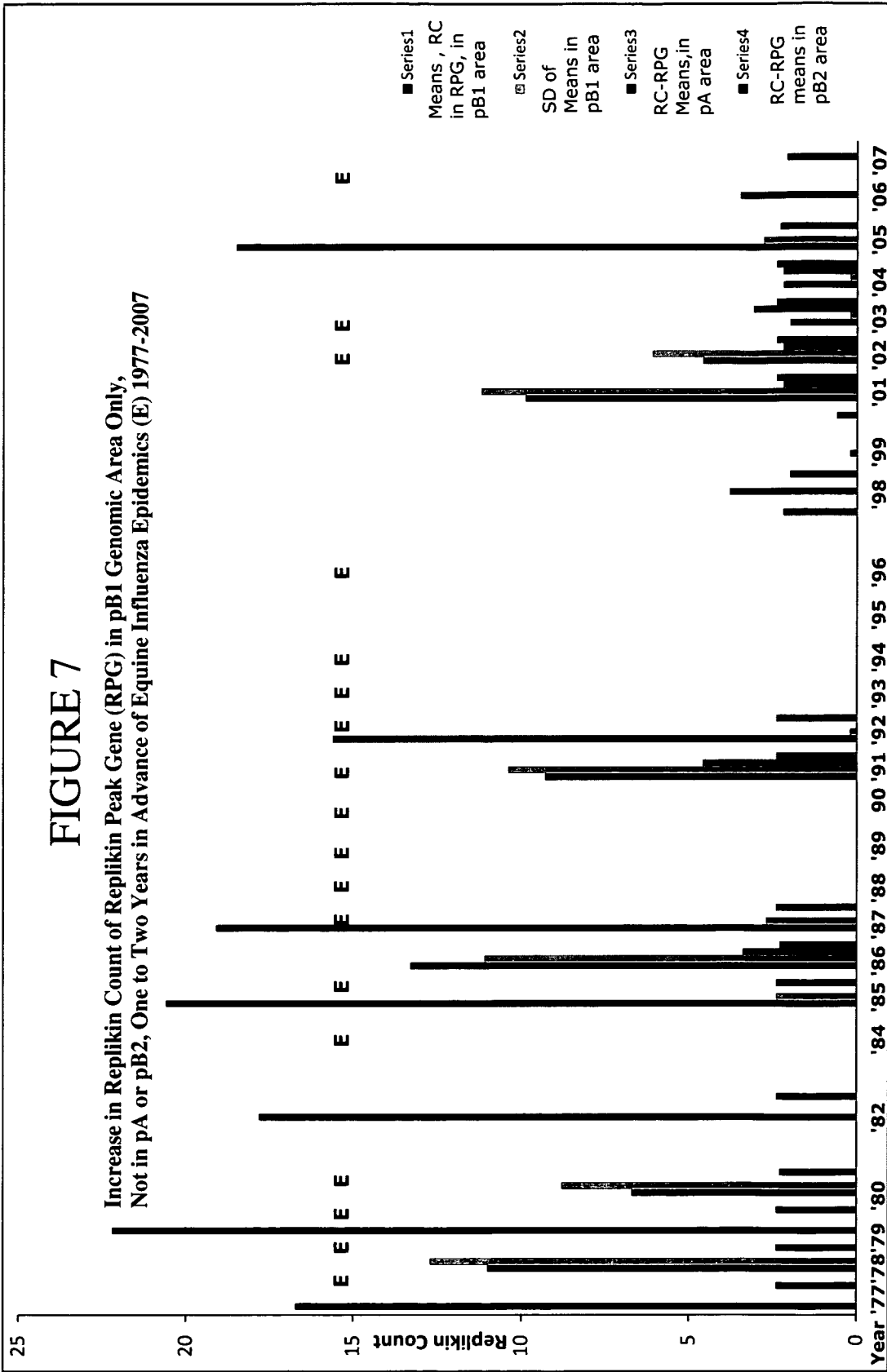
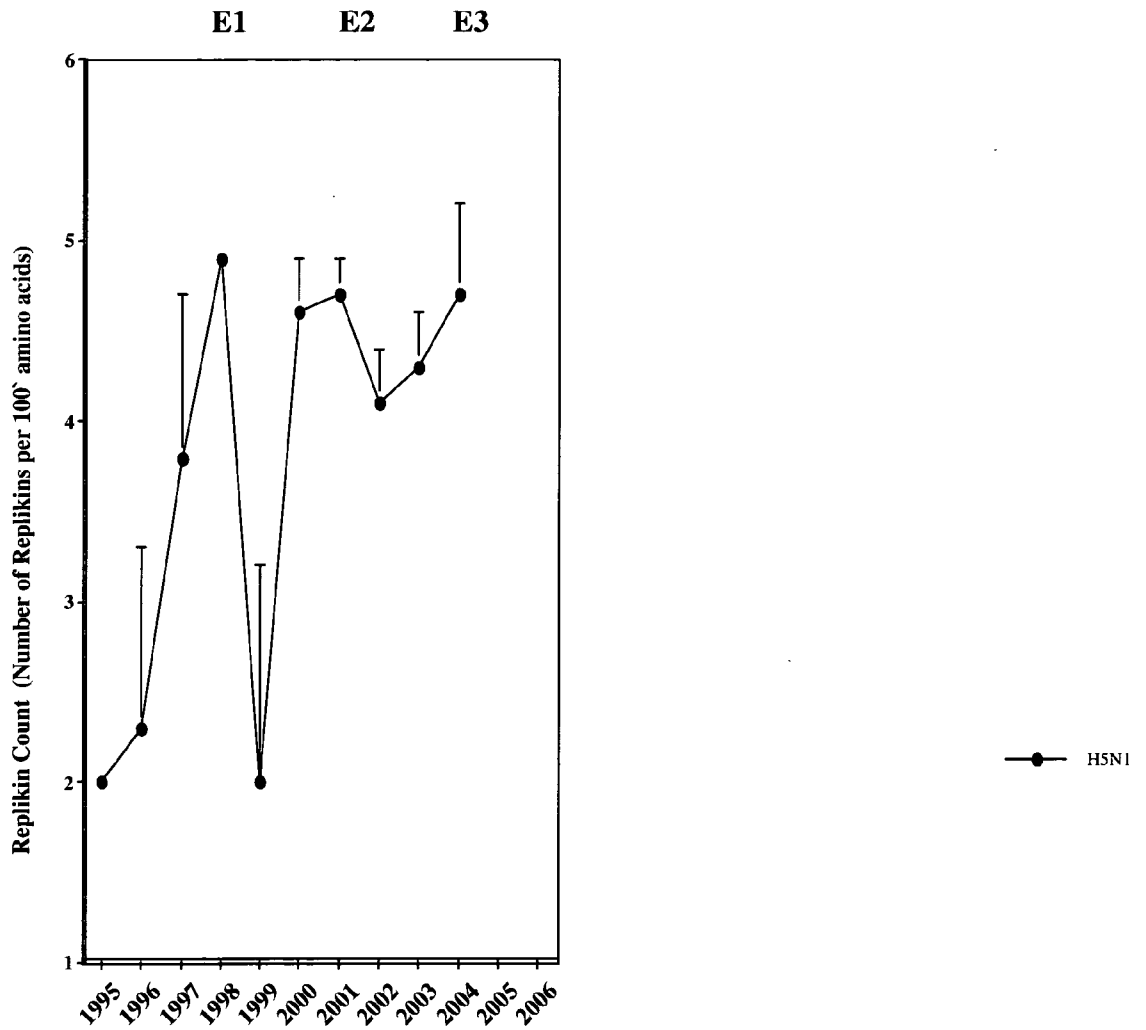
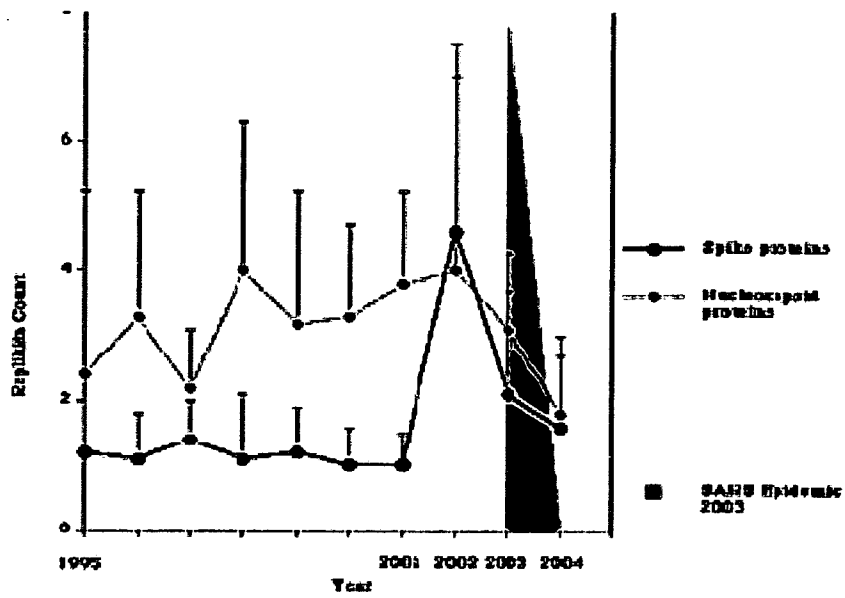


FIGURE 8



Increasing Replikin Count of the whole hemagglutinin protein of H5N1 precedes ‘Bird Flu’ outbreaks and epidemics E1, E2, and E3, 1997-2006.

FIGURE 9



Coronavirus SARS Outbreak

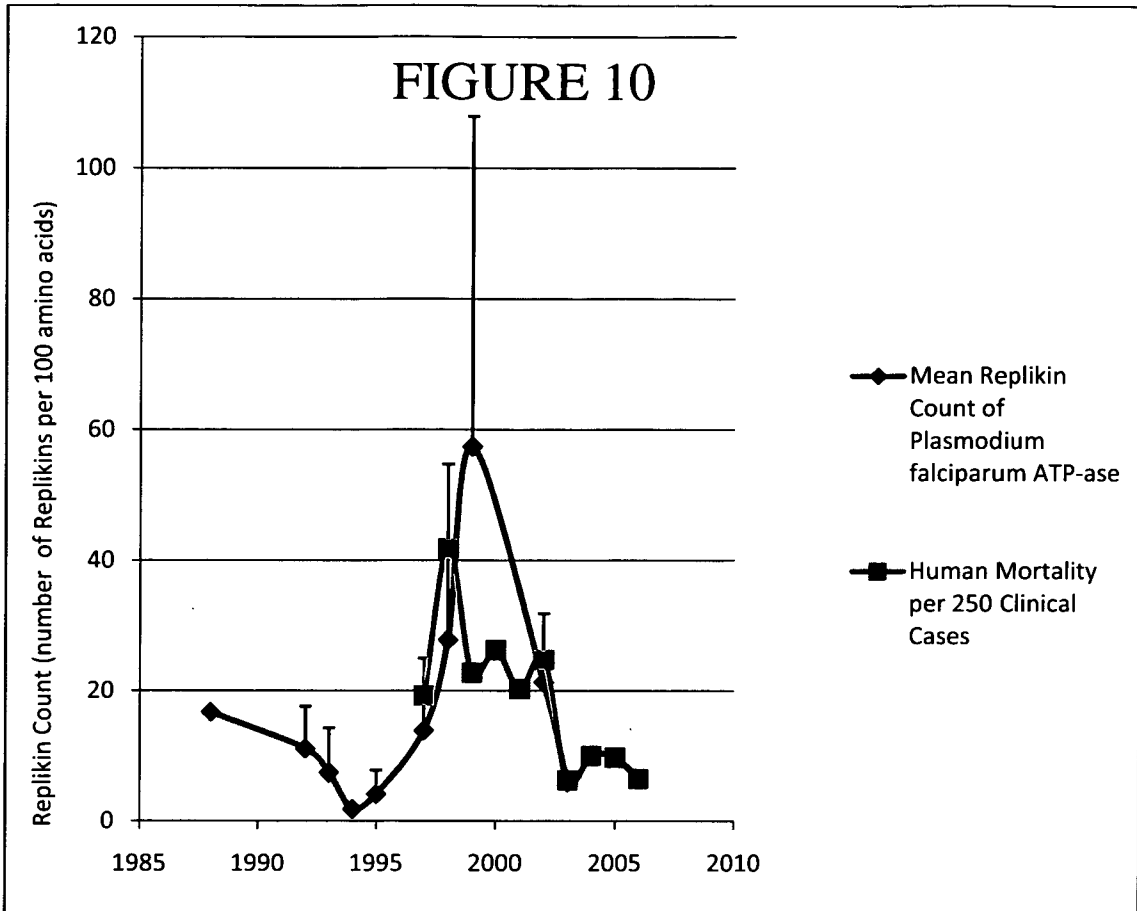
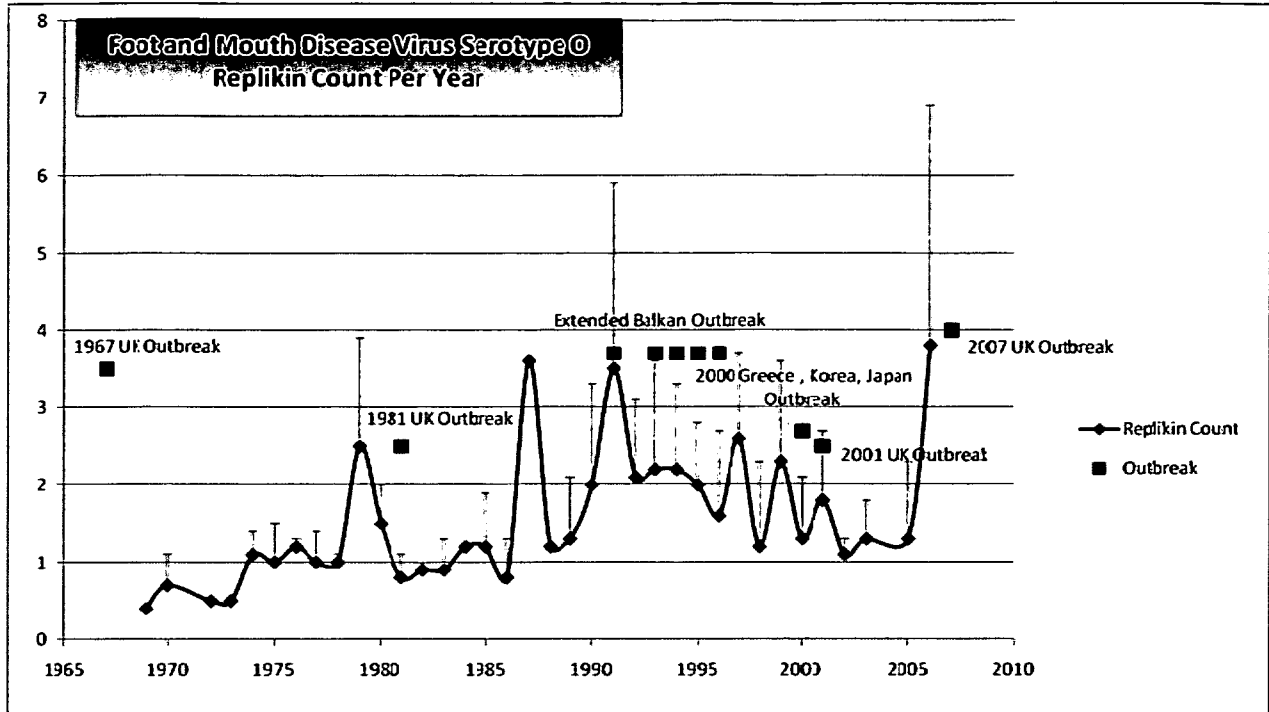


FIGURE 11

Replikin
Count



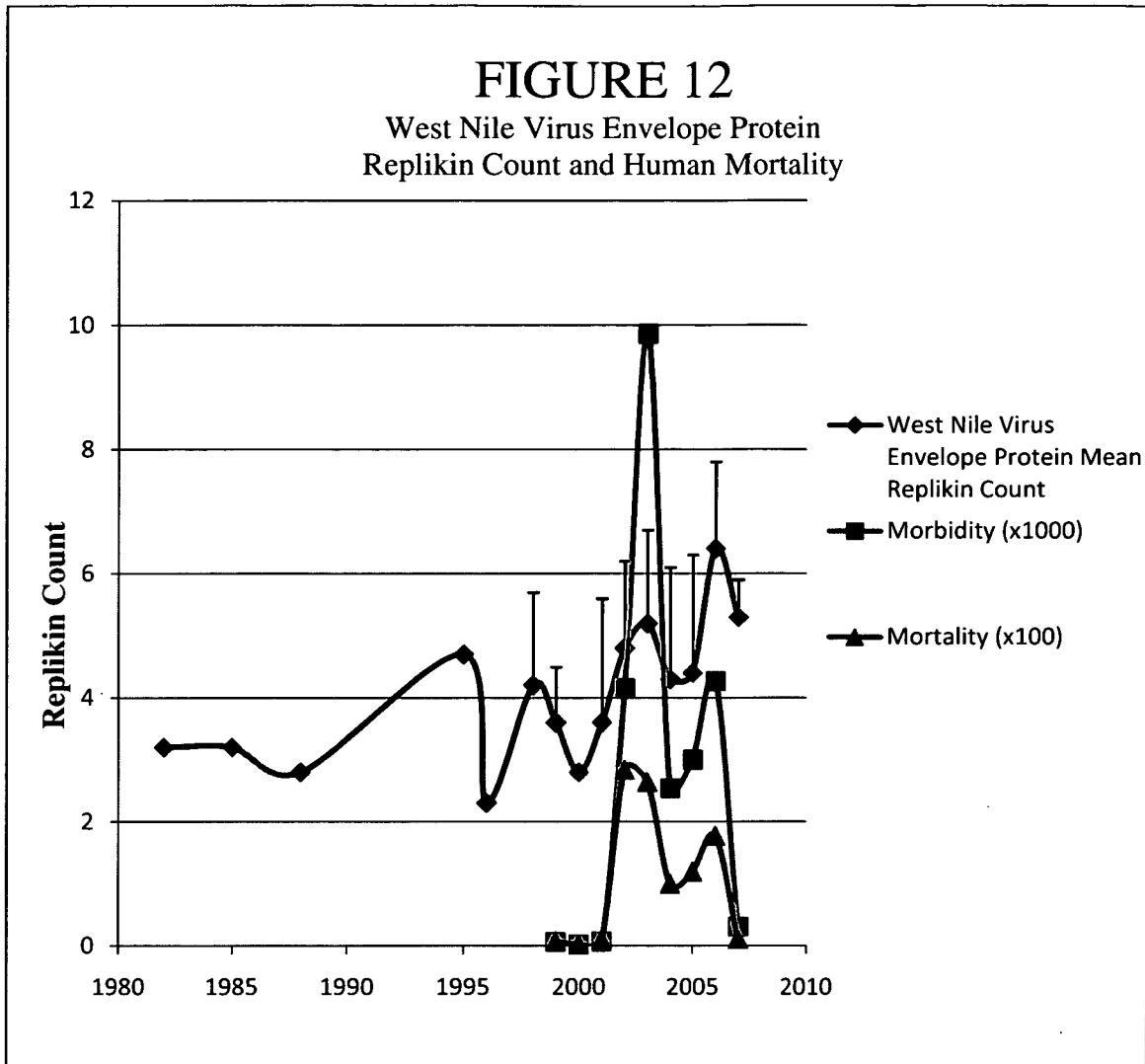
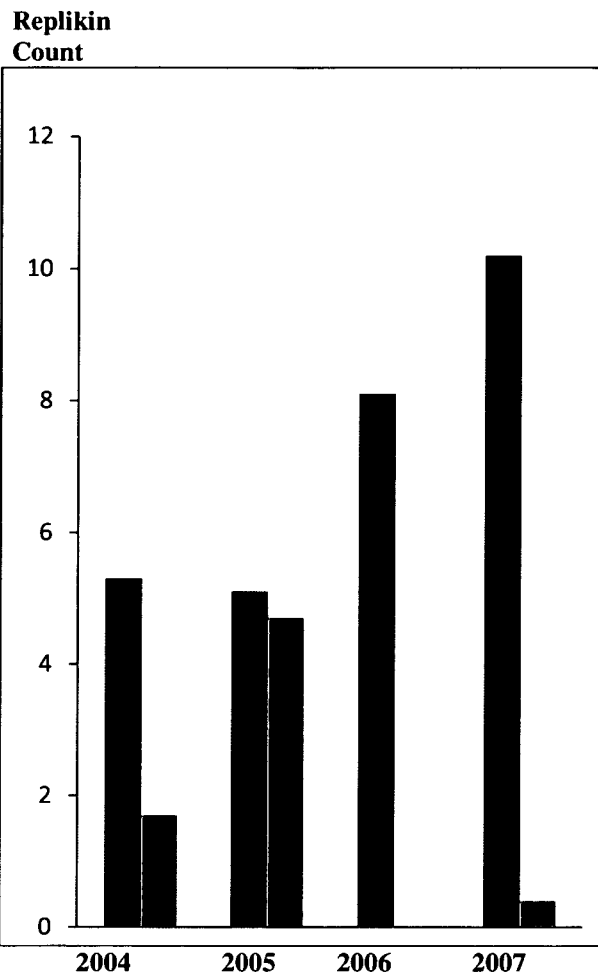


FIGURE 13

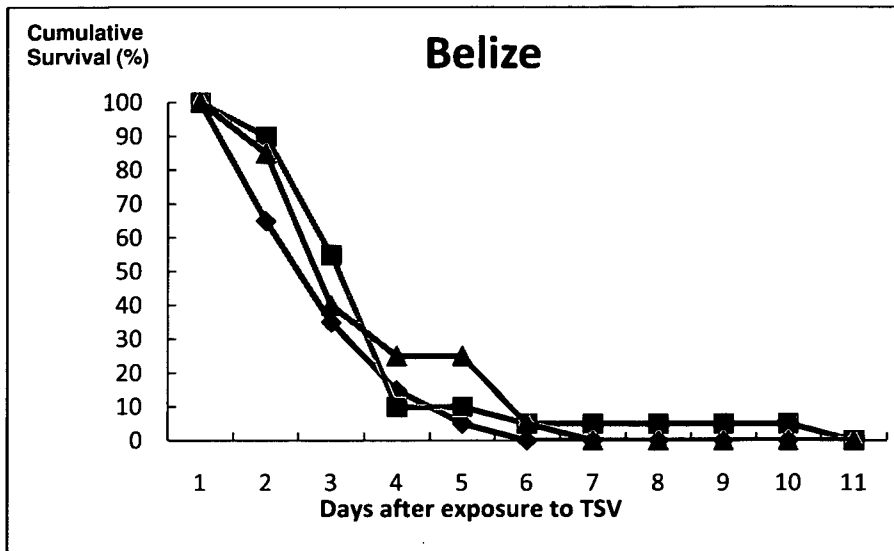


PRRS virus nucleocapsid protein replikin counts 2004 to 2007

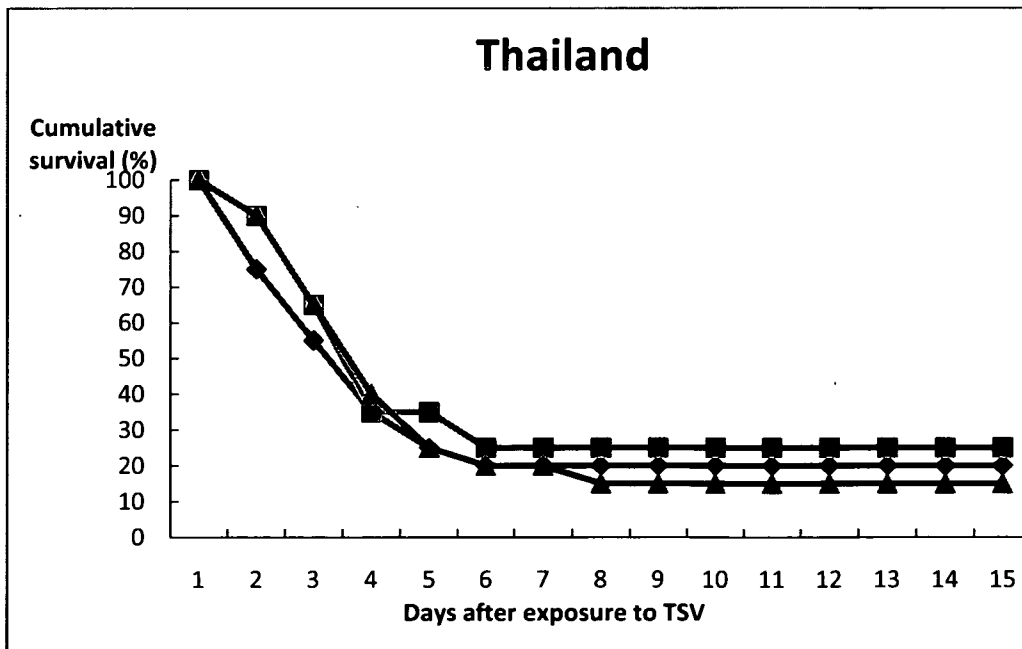
FIGURE 14

Cumulative survival of *Litopenaeus vannamei* per os challenged with TSV isolates: a: Belize; b: Thailand; c: Hawaii; d: Venezuela.

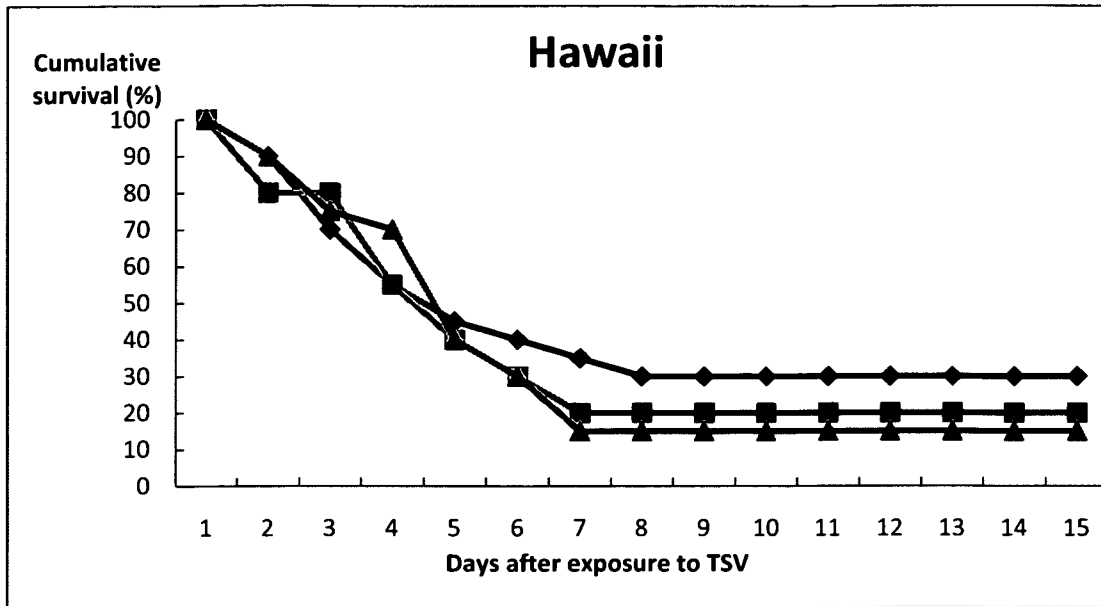
14A



14B.



14C



14D

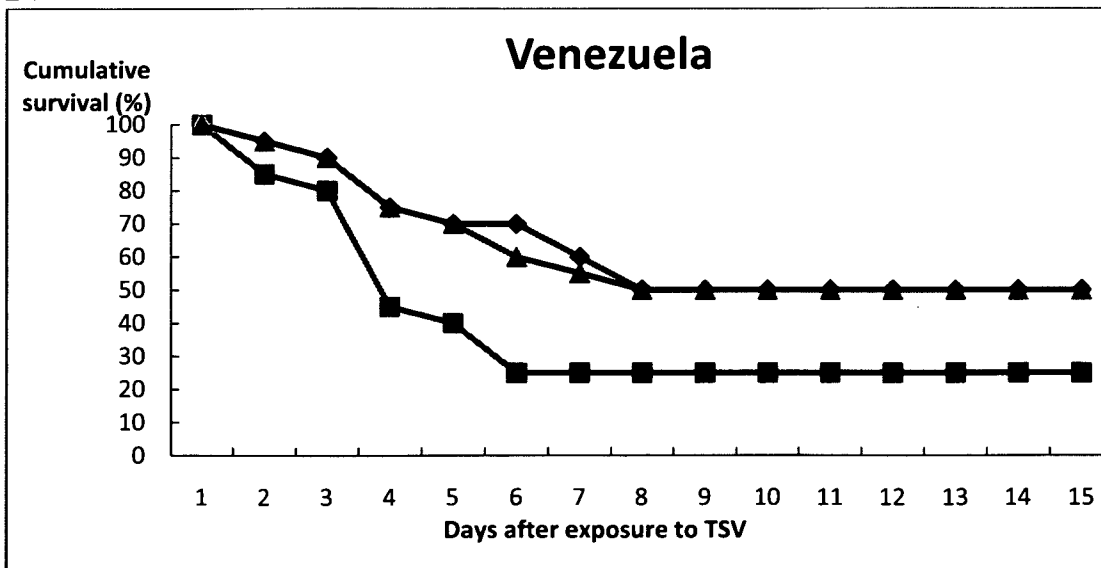
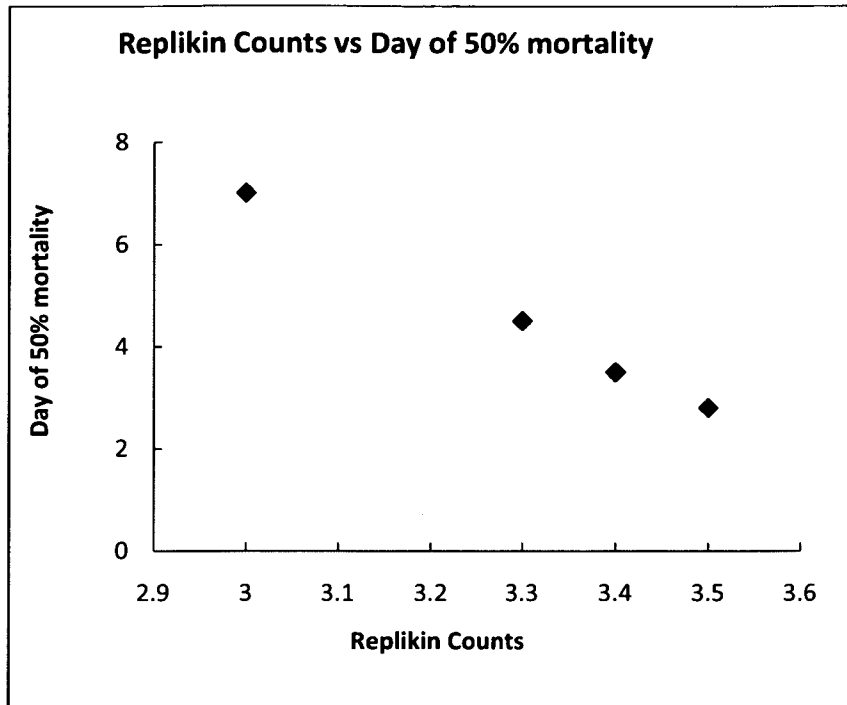
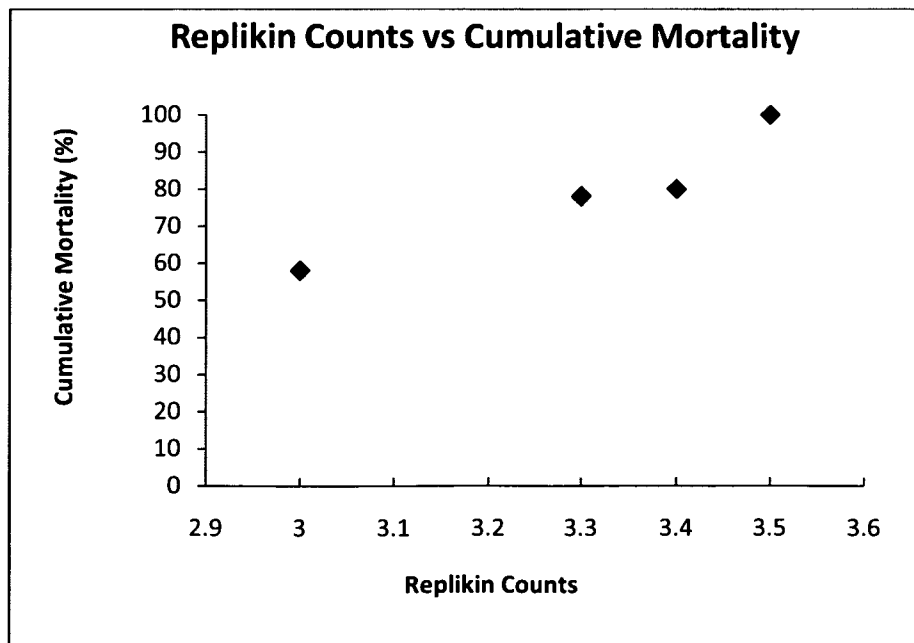


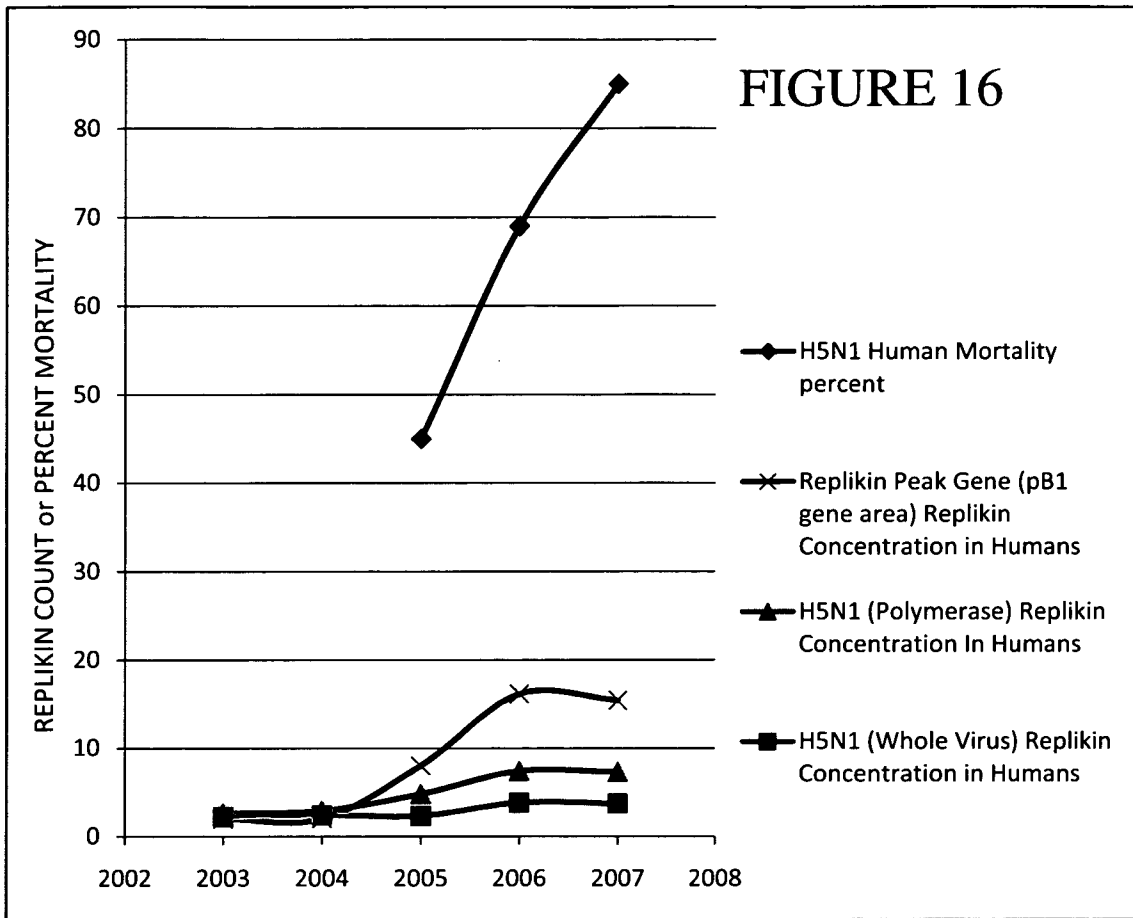
FIGURE 15

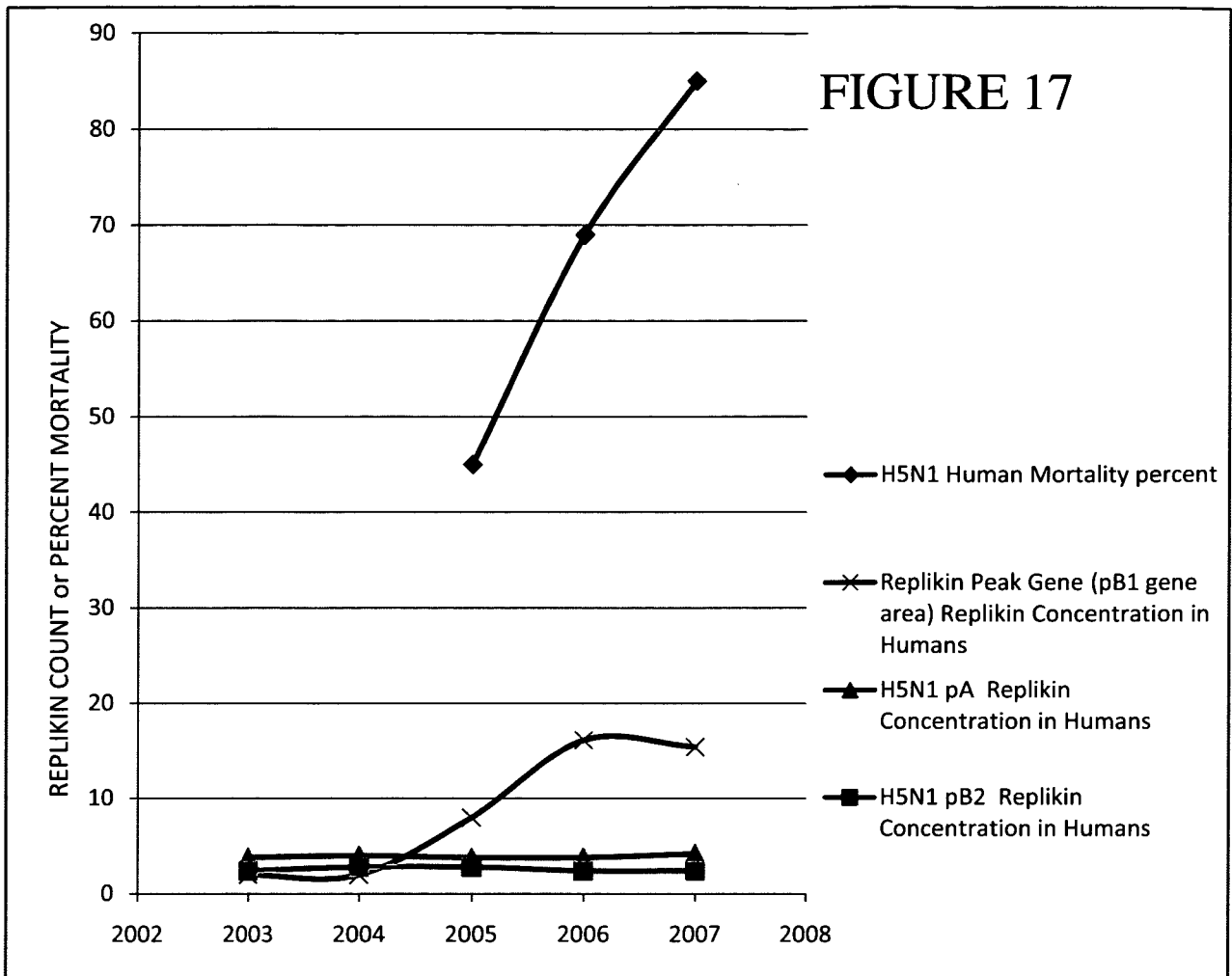
15A

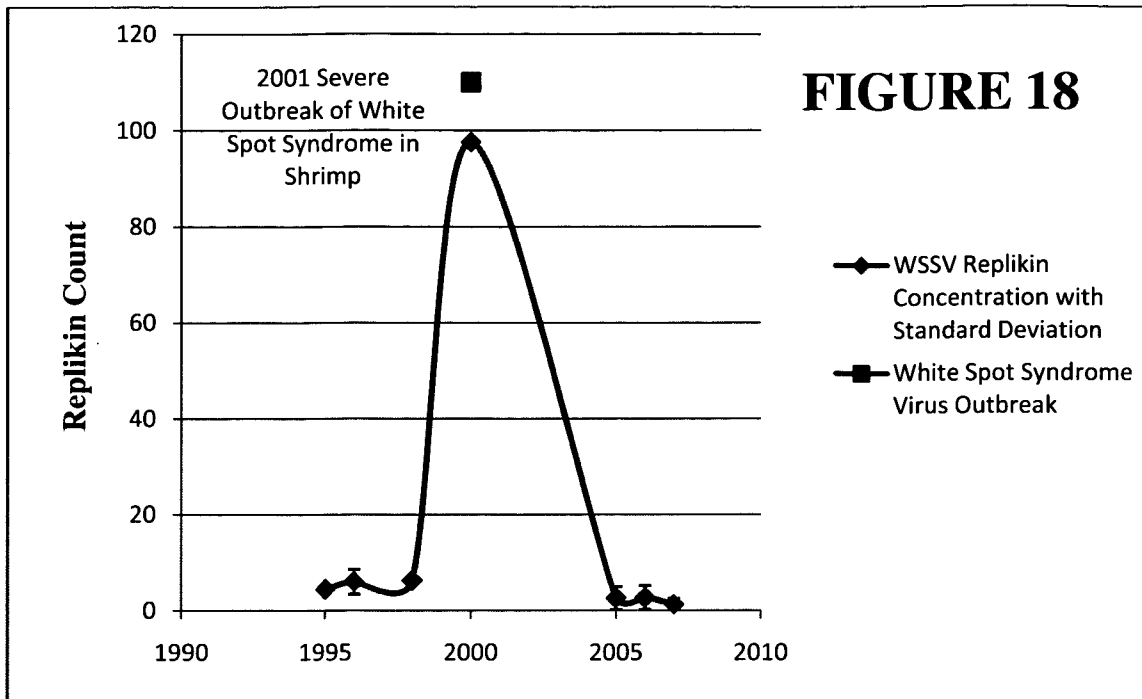


15B









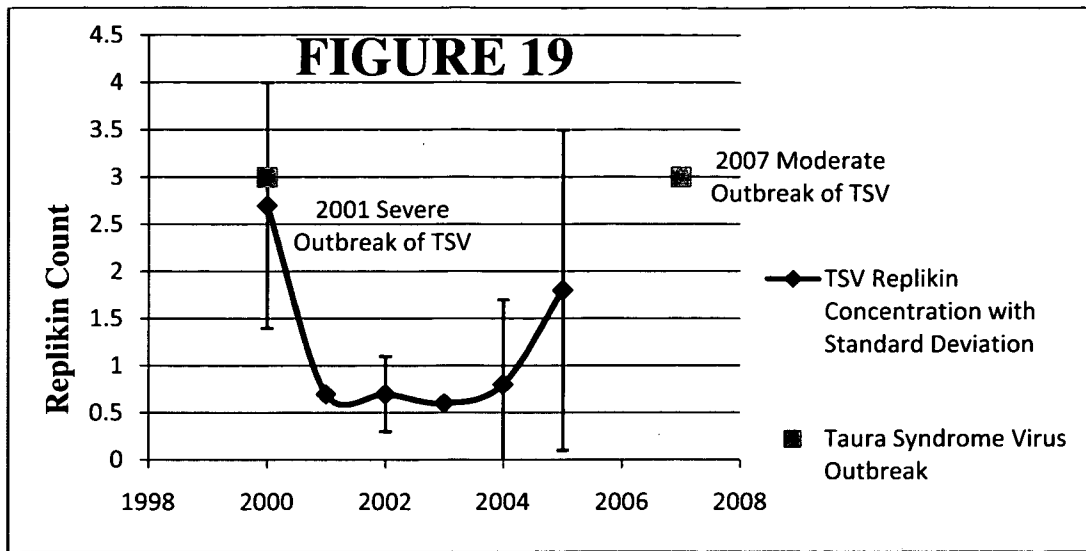


FIGURE 20

