



- (51) **International Patent Classification:**
C07K 14/005 (2006.01) *A61K 39/12* (2006.01)
- (21) **International Application Number:**
PCT/GB2017/050329
- (22) **International Filing Date:**
9 February 2017 (09.02.2017)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
1602535.5 12 February 2016 (12.02.2016) GB
- (71) **Applicant:** THE UNIVERSITY COURT OF THE UNIVERSITY OF EDINBURGH [GB/GB]; Old College, South Bridge, Edinburgh Lothian EH8 9YL (GB).
- (72) **Inventor:** DIGARD, Paul; Chair of Virology, The Roslin Institute, The University of Edinburgh, Easter Bush, Edinburgh Lothian EH25 9RG (GB).
- (74) **Agent:** HGF LIMITED; Document Handling - (HGF) York, 1 City Walk, Leeds Yorkshire LS11 9DX (GB).
- (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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM,

DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

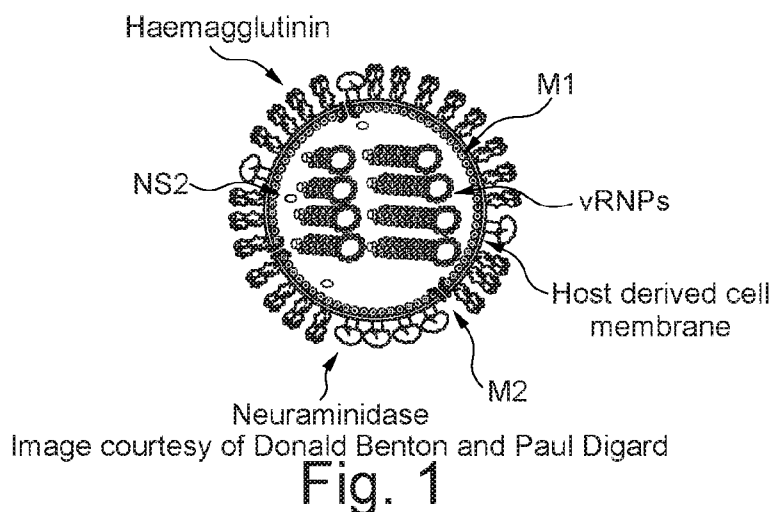
Declarations under Rule 4.17:

— of inventorship (Rule 4.17(iv))

Published:

- with international search report (Art. 21(3))
— with sequence listing part of description (Rule 5.2(a))

(54) **Title:** IMPROVED FLU VACCINE YIELD



(57) **Abstract:** The present invention relates to mutant PR8 segment 3 nucleic acid molecules, expression cassettes, expression systems and host cells, and their uses in the production of reassortant influenza A viruses and associated viral vaccines.

Improved Flu vaccine yield

The present invention relates to mutant PR8 segment 3 nucleic acid molecules, expression cassettes, expression systems and host cells, and their uses in the production of reassortant influenza A viruses and associated viral vaccines.

BACKGROUND

Influenza A viruses and influenza B viruses cause seasonal human influenza but influenza A viruses pose an additional risk of zoonotic infection, with the potential of a host switch and the generation of pandemic influenza. Influenza A and B viruses contain eight genomic segments each of which encode for one or several virus proteins. Six genomic segments of influenza A viruses (segment 1, segment 2, segment 3, segment 5, segment 7 and segment 8) encode the ten "internal" proteins (PB2, PB1, PB1-F2, PA, PA-X, NP, M1, M2, NS1 and NS2); and two genomic segments (segment 4 and segment 6) encode the surface glycoproteins (the antigenic determinants; HA and NA) of the virus. Influenza A viruses are divided into subtypes by antigenicity of the two surface glycoproteins haemagglutinin (HA) and neuraminidase (NA) shown in the cartoon of the virion in Figure 1. The virion contains a greater proportion of HA molecules when compared with NA (Murti & Webster, 1986; illustrated in Figure 1) and so HA is the major antigenic component of the virus.

Genomic segment 3 of influenza A viruses encodes the polymerase acidic (PA) protein. Segment 3 has also recently been shown to produce another protein, PA-X, by ribosomal frameshifting during translation into a +1 open reading frame (ORF) termed the X ORF. PA-X is a 29 kDa protein that contains the N-terminal endonuclease domain of PA, and in most isolates a 61 amino acid C-terminus from the X ORF (Jagger *et al.*, 2012). It has been shown that PA-X from the 1918 H1N1 virus shuts off host cell protein synthesis and at the whole animal level, modulates the immune response. Unexpectedly, mutant forms of the 1918 virus lacking PA-X showed increased pathogenicity in mice (Jagger *et al.*, 2012). PA-X activity in repressing cellular protein synthesis is strain dependent (Naffakh *et al.*, 2001; Desmet *et al.*, 2013).

Influenza epidemics occur almost every year as the virus undergoes antigenic drift but pandemics are rare and have occurred with the H1N1 (1918, 1977 and 2009), H2N2 (1957) and H3N2 (1968) subtype viruses. The 1918 'Spanish flu' pandemic was by far the worst, resulting in 40-100 million deaths worldwide (Johnson & Mueller, 2002), while the 2009 swine flu pandemic caused an estimated 200,000 deaths worldwide (Dawood *et al.*, 2012).

Vaccination is the most important public measure to reduce the impact of influenza epidemics and pandemics. The principle prophylactic approach is to vaccinate individuals using inactivated virus, the “flu shot” composed of quadrivalent or trivalent inactivated virus (TIV); an alternative approach uses live attenuated virus (LAIV) vaccines (administered as nasal spray). Inactivated influenza vaccines are produced using technology that has been around for decades; they are generally safe and effective, but weaknesses in their manufacture and regulation have been apparent in recent years, not least following the experience of the pandemic of 2009 (Albein et al., 2011).

Before the production of inactivated influenza vaccines can commence, and to aid production, high-yielding candidate vaccine viruses (CVVs) need to be prepared based on current circulating influenza viruses or on newly emerging viruses of pandemic concern. The current influenza vaccine manufacturing technique is classical reassortment. Figure 2 is a schematic diagram of this process. In this process, the candidate vaccine virus (strain 1) along with a high yielding virus adapted to growth in eggs (the A/Puerto Rico/8/34 strain (also known herein as “A/PR/8/34” or “PR8”)) are used to co-infect eggs. Vaccine manufacturers then select the viruses which contain the glycoproteins (the antigenic determinants) of the candidate vaccine viruses and test to see which viruses give the best yield in embryonated hen’s eggs. This is a time consuming process, as the manufacturers can’t predict which reassortant will give the best yield. The annual seasonal vaccine production process takes six months - in the case of a pandemic you want to produce vaccine as soon as possible. And in the case of the 2009 swine flu pandemic, which grew poorly in eggs, vaccine manufacturers were unable to produce enough vaccine in the timely manner-promised which was 120 million doses by October; only 12 million doses were ready.

Influenza viruses can be generated in the lab by a technique called reverse genetics, first described by Neumann et al., 1999 and improved by Hoffmann et al., 2000. As shown in Figure 3, this method relies on the transfection of eight plasmids, each encoding one of the eight genomic segments of influenza. Typically, the six viral backbone segments (segments 1, 2, 3, 5, 7 and 8) are derived from the donor strain PR8, whereas the two segments encoding NA and HA are derived from the candidate vaccine virus (strain 1). When taken up by the cells, these plasmids allow for both virus genomic RNA and protein expression in the cells, resulting in the production of virus particles. This virus can then be produced on a large scale in embryonated hen’s eggs. In the case of production of candidate vaccine viruses for highly pathogenic avian influenza, reverse genetics is the only currently viable method.

A limited number of donor strains for influenza virus vaccine manufacture currently exist. The strain that is most commonly utilized in both classical reassortment and reverse genetics is PR8. However, although PR8 is widely accepted as the donor strain for vaccine manufacture, reassortant influenza viruses comprising the PR8 backbone segments do not always grow sufficiently well to ensure efficient vaccine manufacture.

There is a clear need for new reagents and methods for influenza A vaccine production.

BRIEF SUMMARY OF THE DISCLOSURE

The invention is based on the surprising finding that a reduction in PA-X expression during candidate vaccine virus manufacture results in increased virus yield.

The inventors have generated a mutant PR8 segment 3 nucleic acid molecule that exhibits reduced PA-X polypeptide expression compared to the PA-X polypeptide expression of a parent PR8 segment 3 nucleic acid molecule. Using reverse genetics, the mutant PR8 segment 3 was used, together with parent PR8 backbone segments 1, 2, 5, 7 and 8, to generate a candidate vaccine virus (CVV) bearing the two antigenic determinants (glycoproteins NA and HA) of the 2009 pandemic (pdm09) (H1N1) virus. Surprisingly, the mutant CVV (also known herein as "FS mutant" or "6:2FS") gives on average 2-3 fold higher HA antigen yield in embryonated hen's eggs than the wild-type 6:2 counterpart (i.e. a CVV with parent PR8 backbone segments 1, 2, 3, 5, 7 and 8, and NA and HA segments from the 2009 pandemic (pdm09) (H1N1) virus). These data suggest that a reduction in PA-X expression can be used as a novel mechanism to increase HA antigen yield during vaccine manufacture.

A specific frameshift site mutation was introduced into the PR8 segment 3 nucleic acid sequence to demonstrate the invention (see details below). However, the invention is more broadly applicable, and covers any mutation in the PR8 segment 3 nucleic acid sequence that results in the desired phenotype (i.e. a reduction in PA-X expression). The invention also encompasses other mechanisms by which PA-X activity is reduced (e.g. an increase in PA-X protein degradation; or a reduction in PA-X protein stability; or a truncation of the PA-X gene).

The invention was demonstrated using reverse genetics to generate a mutant candidate vaccine virus, wherein the segment 3 nucleic acid sequence exhibited reduced PA-X protein expression. However, the inventive concept may be considered to encompass a recombinant PR8 donor strain with a mutated PR8 segment 3 nucleic acid sequence such

that PA-X protein expression is reduced, and its use in the manufacture of candidate vaccine viruses using classical reassortment techniques.

Advantageously, the invention may be applied to improving the yield of candidate vaccine viruses bearing glycoproteins (i.e. HA and/or NA) of potential pandemic viruses.

The invention provides a mutant PR8 segment 3 nucleic acid molecule that exhibits reduced PA-X polypeptide expression compared to the PA-X polypeptide expression of a parent PR8 segment 3 nucleic acid molecule under the same expression conditions.

Optionally, the parent PR8 segment 3 nucleic acid molecule encodes a PA-X polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO: 13.

Optionally, the parent PR8 segment 3 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity to the nucleic acid sequence of SEQ ID NO: 3.

Optionally, the mutant PR8 segment 3 nucleic acid molecule:

(a) comprises a mutation in the frameshift region of the parent PR8 segment 3 nucleic acid molecule, wherein the mutation reduces the expression of the PA-X polypeptide compared to the PA-X expression of the parent PR8 segment 3 nucleic acid molecule under the same expression conditions;

(b) comprises a premature stop codon within the X-ORF of the parent PR8 segment 3 nucleic acid molecule, wherein the premature stop codon reduces the expression of the PA-X polypeptide compared to the PA-X expression of the parent PR8 segment 3 nucleic acid molecule under the same expression conditions; and/or

(c) does not comprise the X-ORF of the parent PR8 segment 3 nucleic acid molecule.

Optionally, the parent PR8 segment 3 nucleic acid molecule comprises a frameshift region comprising the nucleic acid sequence UCC UUU CGU C (SEQ ID NO: 14), and wherein the mutant PR8 segment 3 nucleic acid molecule comprises a mutation within this frameshift region.

Optionally, the mutant PR8 segment 3 nucleic acid molecule comprises a mutated frameshift region comprising the nucleic acid sequence AGC UUC AGA (SEQ ID NO: 15).

Optionally, the mutant PR8 segment 3 nucleic acid molecule comprises or consists of the nucleic acid sequence of any one of SEQ ID NOs: 7 to 11.

Optionally, the mutant PR8 segment 3 nucleic acid molecule exhibits at least a 5%, 10%,
5 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or 95% reduction in PA-X polypeptide expression compared to the PA-X polypeptide expression of the parent PR8 segment 3 nucleic acid molecule under the same expression conditions.

The invention provides an expression cassette comprising the mutant nucleic acid molecule
10 of the invention, wherein the nucleic acid molecule is functionally connected to a regulated or constitutive promoter.

Optionally, the promoter is an RNA Pol I or an RNA pol II promoter.

15 Optionally, the expression cassette comprises a pHW2000 plasmid.

Optionally, the expression cassette comprises or consists of the sequence of SEQ ID NO:12.

The invention provides an expression system comprising the expression cassette of the
20 invention.

Optionally, the expression system further comprises at least one of:

- (i) a PR8 segment 1 nucleic acid molecule;
 - (ii) a PR8 segment 2 nucleic acid molecule;
 - 25 (iii) a PR8 segment 5 nucleic acid molecule;
 - (iv) a PR8 segment 7 nucleic acid molecule; and,
 - (v) a PR8 segment 8 nucleic acid molecule,
- wherein each of (i) to (v) are functionally connected to a regulated or constitutive promoter.

30 Optionally, the at least one of (i) to (v) is part of the same or part of a different expression cassette than the expression cassette of the invention.

Optionally:

- (i) the PR8 segment 1 nucleic acid molecule encodes a RNA polymerase subunit PB2
35 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:16; and /or

(ii) the PR8 segment 2 nucleic acid molecule encodes a RNA polymerase subunit PB1 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:17; and/or

(iii) the PR8 segment 5 nucleic acid molecule encodes a nucleoprotein NP polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:18; and/or

(iv) the PR8 segment 7 nucleic acid molecule encodes a matrix protein M1 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:19 and a matrix protein M2 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:20; and/or

(v) the PR8 segment 8 nucleic acid molecule encodes a non-structural protein NS1 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:21 and a non-structural protein NEP polypeptide that has at least 95%, at least 99% or at least 100% sequence identity to the amino acid sequence of SEQ ID NO:22.

Optionally:

(i) the PR8 segment 1 nucleic acid molecule encodes a RNA polymerase subunit PB2 polypeptide that has the amino acid sequence of SEQ ID NO:16;

(ii) the PR8 segment 2 nucleic acid molecule encodes a RNA polymerase subunit PB1 polypeptide that has the amino acid sequence of SEQ ID NO:17;

(iii) the PR8 segment 5 nucleic acid molecule encodes a nucleoprotein NP polypeptide that has the amino acid sequence of SEQ ID NO:18;

(iv) the PR8 segment 7 nucleic acid molecule encodes a matrix protein M1 polypeptide that has the amino acid sequence of SEQ ID NO:19 and a matrix protein M2 polypeptide that has the amino acid sequence of SEQ ID NO:20;

(v) the PR8 segment 8 nucleic acid molecule encodes a non-structural protein NS1 polypeptide that has the amino acid sequence of SEQ ID NO:21 and a non-structural protein NEP polypeptide that has the amino acid sequence of SEQ ID NO:22.

Optionally:

(i) the PR8 segment 1 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 1; and/or

(ii) the PR8 segment 2 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 2; and/or

(iii) the PR8 segment 5 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 4; and/or

(iv) the PR8 segment 7 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO:5; and/or

(v) the PR8 segment 8 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 6.

5

Optionally:

(i) the PR8 segment 1 nucleic acid molecule has the sequence of SEQ ID NO: 1;

(ii) the PR8 segment 2 nucleic acid molecule has the sequence of SEQ ID NO: 2;

(iii) the mutant PR8 segment 3 nucleic acid molecule has the sequence of SEQ ID NO:

10 11;

(iv) the PR8 segment 5 nucleic acid molecule has the sequence of SEQ ID NO: 4;

(v) the PR8 segment 7 nucleic acid molecule has the sequence of SEQ ID NO: 5; and

(vi) the PR8 segment 8 nucleic acid molecule has the sequence of SEQ ID NO: 6.

15 Optionally, the expression system further comprises at least one of:

(i) an expression cassette comprising a segment 4 nucleic acid molecule, wherein the segment 4 nucleic acid molecule is from a different influenza A virus strain than PR8; and

(ii) an expression cassette comprising a segment 6 nucleic acid molecule, wherein the segment 6 nucleic acid molecule is from a different influenza A virus strain than PR8;

20 wherein each of (i) to (ii) are functionally connected to a regulated or constitutive promoter.

Optionally, the segment 4 nucleic acid molecule and segment 6 nucleic acid molecule are from the same influenza A virus strain.

25 The invention provides a eukaryotic host comprising the mutant nucleic acid molecule of the invention, the expression cassette of the invention, or the expression system of the invention.

Optionally, the host is an embryonated hen egg or a mammalian cell; optionally the

30 mammalian cell is an MDCK, Vero or PerC6 cell.

The invention provides a recombinant PR8 virus comprising the mutant nucleic acid molecule of the invention.

35 The invention provides use of a mutant nucleic acid molecule according to the invention, an expression cassette according to the invention, an expression system according to the

invention, a host cell according to the invention, or a recombinant PR8 virus according to the invention, for the production of a reassortant influenza A virus (IAV).

Optionally, the reassortant IAV is a candidate vaccine virus.

5

The invention provides a reassortant influenza A virus (IAV) comprising the nucleic acid molecule of the invention.

10 Optionally, the reassortant influenza A virus further comprises a PR8 segment 1 nucleic acid molecule, a PR8 segment 2 nucleic acid molecule, a PR8 segment 5 nucleic acid molecule, a PR8 segment 7 nucleic acid molecule, a PR8 segment 8 nucleic acid molecule, a segment 4 nucleic acid molecule from a different influenza A virus strain than PR8, and a segment 6 nucleic acid molecule from a different influenza A virus strain than PR8.

15 Optionally, the segment 4 nucleic acid molecule and segment 6 nucleic acid molecule are from the same influenza A virus strain.

Optionally:

20 (i) the PR8 segment 1 nucleic acid molecule encodes a RNA polymerase subunit PB2 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:16; and /or

(ii) the PR8 segment 2 nucleic acid molecule encodes a RNA polymerase subunit PB1 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:17; and/or

25 (iii) the PR8 segment 5 nucleic acid molecule encodes a nucleoprotein NP polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:18; and/or

30 (iv) the PR8 segment 7 nucleic acid molecule encodes a matrix protein M1 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:19 and a matrix protein M2 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:20; and/or

35 (v) the PR8 segment 8 nucleic acid molecule encodes a non-structural protein NS1 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:21 and a non-structural protein NEP polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:22.

Optionally:

- (i) the PR8 segment 1 nucleic acid molecule encodes a RNA polymerase subunit PB2 polypeptide that has the amino acid sequence of SEQ ID NO:16;
- (ii) the PR8 segment 2 nucleic acid molecule encodes a RNA polymerase subunit PB1 polypeptide that has the amino acid sequence of SEQ ID NO:17;
- (iii) the PR8 segment 5 nucleic acid molecule encodes a nucleoprotein NP polypeptide that has the amino acid sequence of SEQ ID NO:18;
- (iv) the PR8 segment 7 nucleic acid molecule encodes a matrix protein M1 polypeptide that has the amino acid sequence of SEQ ID NO:19 and a matrix protein M2 polypeptide that has the amino acid sequence of SEQ ID NO:20;
- (v) the PR8 segment 8 nucleic acid molecule encodes a non-structural protein NS1 polypeptide that has the amino acid sequence of SEQ ID NO:21 and a non-structural protein NEP polypeptide that has the amino acid sequence of SEQ ID NO:22.

Optionally:

- (i) the PR8 segment 1 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 1; and/or
- (ii) the PR8 segment 2 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 2; and/or
- (iii) the PR8 segment 5 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 4; and/or
- (iv) the PR8 segment 7 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO:5; and/or
- (v) the PR8 segment 8 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 6.

Optionally:

- (i) the PR8 segment 1 nucleic acid molecule has the sequence of SEQ ID NO: 1;
- (ii) the PR8 segment 2 nucleic acid molecule has the sequence of SEQ ID NO: 2;
- (iii) the mutant PR8 segment 3 nucleic acid molecule has the sequence of SEQ ID NO: 11;
- (iv) the PR8 segment 5 nucleic acid molecule has the sequence of SEQ ID NO: 4;
- (v) the PR8 segment 7 nucleic acid molecule has the sequence of SEQ ID NO: 5; and the PR8 segment 8 nucleic acid molecule has the sequence of SEQ ID NO: 6.

The invention provides a method of preparing a reassortant influenza A virus (IAV) comprising the steps of:

- (a) introducing an expression cassette of the invention, or an expression system of the invention into a host cell; and
- (b) culturing the host cell of step (a) in order to produce the reassortant IAV; and optionally
- (c) purifying the reassortant IAV obtained in step (b).

5

The invention provides a method for producing an influenza A virus (IAV) comprising the steps of:

- (a) infecting a host cell with a reassortant influenza A virus of the invention or a recombinant PR8 virus of the invention;
- 10 (b) culturing the host cell from step (a) to produce the influenza A virus; and optionally
- (c) purifying the influenza A virus produced in step (b).

Optionally, the host cell is an embryonated hen egg.

- 15 Optionally, the culture host is a mammalian cell, optionally wherein the cell is an MDCK, Vero or PerC6 cell.

Optionally, the host cell grows adherently or in suspension.

- 20 The invention provides an influenza A virus obtained by a method of the invention.

The invention provides a method of preparing a vaccine, comprising the steps of (a) preparing a virus by a method of the invention and (b) preparing a vaccine from the virus.

- 25 Optionally, step (b) involves inactivating the virus.

Optionally, the vaccine is a whole virion vaccine, a split virion vaccine, a surface antigen vaccine or a virosomal vaccine.

- 30 The invention provides a pharmaceutical composition comprising a vaccine obtained by the method the invention, and a pharmaceutically acceptable carrier or diluent.

The invention provides a vaccine obtained by the method of the invention and a pharmaceutically acceptable carrier or diluent.

35

The invention provides a vaccine of the invention for the treatment or prophylaxis of influenza.

BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the invention are further described hereinafter with reference to the accompanying drawings:

5

Figure 1 illustrates the schematic organisation of the influenza A virion. A lipid envelope derived from the host cell plasma membrane contains the viral HA, NA and M2 integral membrane proteins, overlies a matrix layer comprised of the M1 polypeptide and surrounds 8 ribonucleoprotein particles containing each of the 8 genome segments encapsidated by the NP, PB1, PB2 and PA polypeptides.

10

Figure 2 illustrates the conventional process for generating CVVs. Embryonated hen's eggs are inoculated with two strains of influenza A virus, typically the PR8 strain and the strain containing the vaccine candidate virus of interest. Simultaneous infection of individual cells with both viruses leads to the generation of reassortant viruses containing mixtures of segments from the two parental viruses. A CVV containing the HA and NA genes of antigenic interest and (typically) 5 or 6 segments from PR8 is selected for its high growth properties and used to prepare vaccine.

15

Figure 3 illustrates the principle of 8-plasmid reverse genetics systems. Highly transfectable 293T cells are transfected with a pool of 8 plasmids each containing cDNA copies of one of the 8 influenza A virus genomic segments. The plasmids are engineered such that bidirectional promoters flanking the viral cDNA produce either mRNA encoding the virus polypeptides or negative sense genomic viral RNA (vRNA). This system allows the self-assembly of infectious virus of the desired sequence that is then amplified by inoculation into embryonated hen's eggs.

20

25

Figure 4 illustrates the differential pathology induced by WT and PA-X mutant influenza A viruses. 10 day old embryonated hen's eggs were inoculated with 1000 PFU of either WT or FS mutant PR8 strains (or mock infected with PBS) and harvested 2 days post infection. (a) Embryos were examined for gross pathology changes. Note the bloody, stunted and fragile nature of WT virus-infected embryos (in contrast to mock-infected specimens) and the intermediate nature of the damage induced by FS mutant. (b) lung tissue sections were taken and (b i) stained with haematoxylin and eosin (H&E) stain or (b ii) by immunocytochemistry for viral NP. Note the loss of normal parabronchiole architecture in WT-virus infected sections but its preservation in FS virus-infected samples, despite the abundant levels of virus antigen.

30

35

Figure 5 provides collated data from independent experiments comparing HA yield of PR8:pdm09 CVV 6:2 and 6:2FS viruses. Average HA titres from allantoic fluid of eggs inoculated at the infection dose which gave maximum yield are shown. Data points represent eight independent experiments using three independently rescued RG virus stocks. In (a) the data are presented as scatter plots, in (b), paired observations are shown. The outcome of statistical (Student's T) tests are indicated.

Figure 6 illustrates actual HA protein yield from virus partially purified from pooled allantoic fluid from eggs infected with a virus dose which gave maximum HA titre. Allantoic fluid was clarified of cell debris by centrifugation and then virus was pelleted through a 30% sucrose cushion by ultracentrifugation. Virus pellets from each experiment were resuspended in equal volumes of PBS and treated with (+) or without (-) the enzyme N-glycosidase F to remove glycosylation. A i) Protein content from equivalent volumes of each virus in an experiment were analysed by SDS-PAGE and staining with Coomassie Blue dye on 12% polyacrylamide gels. A representative image of a gel from one experiment is shown. ii) HA content from partially purified virus preparations was analysed by SDS-PAGE and detection of HA1 in western blot using rabbit polyclonal anti-swine HA antibody. A representative western blot from one experiment is shown.

(b) HA1 yield quantified by densitometry. De-glycosylated HA1 was quantified by western blot (using rabbit polyclonal anti-swine HA antibody) from partially purified virus preparations from pooled allantoic fluid of eggs at the infection dose which gave maximum yield. Data points represent eight independent experiments using three independently rescued RG virus stocks. Data are presented as a scatter plot in (i) and paired observations are shown in (ii). The outcome of statistical (Student's T) tests are indicated.

Figure 7 illustrates relative HAU and HA1 yields of 6:2FS virus compared with 6:2 candidate vaccine viruses (CVVs) containing glycoproteins of several influenza A virus strains, including (in order) two from the 2009 pandemic, two from the 1968 pandemic and a low pathogenicity avian virus isolate. Numerical data are the fold increase in HA titre or HA1 yield (as assessed by SDS-PAGE, western blot and densitometry) resulting from the FS mutation in comparison to a WT segment 3. Also tabulated are the number of independent experiments, the number of independent virus rescues and the numbers of small scale and large scale egg preparations experiments that comprise the overall data. Note that in each case, the average yields were improved by the FS mutation, although the magnitude of this

increase was greatest with the pdm2009 viruses. *An outlier from one experiment was ignored when taking the average.

Figure 8 provides the nucleic acid sequence of SEQ ID NO:1 (parent PR8 segment 1).

5

Figure 9 provides the nucleic acid sequence of SEQ ID NO:2 (parent PR8 segment 2).

Figure 10 provides the nucleic acid sequence of SEQ ID NO:3 (parent PR8 segment 3).

10 Figure 11 provides the nucleic acid sequence of SEQ ID NO:4 (parent PR8 segment 5).

Figure 12 provides the nucleic acid sequence of SEQ ID NO:5 (parent PR8 segment 7).

Figure 13 provides the nucleic acid sequence of SEQ ID NO: 6 (parent PR8 segment 8).

15

Figure 14 provides the nucleic acid sequence of SEQ ID NO: 7 (PR8 segment 3 PTC mutant 1).

Figure 15 provides the nucleic acid sequence of SEQ ID NO:8 (PR8 segment 3 PTC mutant 2).

20

Figure 16 provides the nucleic acid sequence of SEQ ID NO:9 (PR8 segment 3 PTC mutant 3).

25 Figure 17 provides the nucleic acid sequence of SEQ ID NO:10 (PR8 segment 3 PTC mutant 4).

Figure 18 provides the nucleic acid sequence of SEQ ID NO: 11 (PR8 segment 3 frameshift mutant).

30

Figure 19 provides the nucleic acid sequence of SEQ ID NO: 12 (vector plasmid pHW2000_RF483- annotated) In order from start: bold text, CMV pol II (continued at bottom of sequence), grey shading, T7 promoter; underlined, RNA pol 1 termination signal; bold print in parentheses, BsmBI site; capital letters, influenza A PR8 segment 3 FS mutant; bold capital, mutated FS site; strikethrough, RNA pol 1 promoter; italic, PMB1 origin; double underline, AmpR. Aligned with GenBank: FR669675.1 (Reverse genetics of influenza C viruses).

35

Figure 20 provides the amino acid sequence of SEQ ID NO: 13 (PR8 PA-X amino acid sequence).

- 5 Figure 21 provides the amino acid sequence of SEQ ID NO: 16 (PR8 PB2 amino acid sequence).

Figure 22 provides the amino acid sequence of SEQ ID NO: 17 (PR8 PB1 amino acid sequence).

10

Figure 23 provides the amino acid sequence of SEQ ID NO: 18 (PR8 NP amino acid sequence).

- 15 Figure 24 provides the amino acid sequence of SEQ ID NO: 19 (PR8 M1 amino acid sequence).

Figure 25 provides the amino acid sequence of SEQ ID NO: 20 (PR8 M2 amino acid sequence).

- 20 Figure 26 provides the amino acid sequence of SEQ ID NO: 21 (PR8 NS1 amino acid sequence).

Figure 27 provides the amino acid sequence of SEQ ID NO: 22 (PR8 NEP amino acid sequence).

25

Figure 28 provides the amino acid sequence of SEQ ID NO: 23 (PR8 PA amino acid sequence).

- 30 Figure 29 provides the nucleic acid sequence of SEQ ID NO: 24 (segment 4 nucleic acid molecule) and the amino acid sequence of SEQ ID NO: 25 (corresponding HA amino acid sequence).

- 35 Figure 30 provides the nucleic acid sequence of SEQ ID NO: 26 (segment 6 nucleic acid molecule) and the amino acid sequence of SEQ ID NO: 27 (corresponding NA amino acid sequence).

DETAILED DESCRIPTION

The invention is based on the surprising finding that a reduction in PA-X protein expression during candidate vaccine virus manufacture results in increased virus yield.

5 The findings presented herein suggest that a reduction in PA-X expression can be used as a novel mechanism to increase HA antigen yield during vaccine manufacture.

The invention therefore provides a new approach to the manufacture of candidate vaccine viruses, with improvements in viral yield.

10 Mutant PR8 segment 3 nucleic acid molecules

The invention provides a mutant PR8 segment 3 nucleic acid molecule that exhibits reduced PA-X polypeptide expression compared to the PA-X polypeptide expression of a parent PR8 segment 3 nucleic acid molecule under the same expression conditions.

15 As used herein, the term “nucleic acid molecule” includes DNA molecules (e.g., a cDNA or genomic DNA) and RNA molecules (e.g., a vRNA or mRNA) and analogs of the DNA or RNA generated, e.g., by the use of nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded.

20 Unless otherwise specified, any reference to a DNA or RNA molecule is intended to include the reverse complement of that molecule. Except where single-strandedness is required by the text herein, DNA or RNA molecules, though written to depict only a single strand, encompass both strands of the molecule. Thus, a reference to the nucleic acid molecule that encodes a specific protein, or a fragment thereof, encompasses both the sense strand and
25 its reverse complement. Thus, for instance, it is appropriate to generate probes or primers from the reverse complement sequence of the disclosed nucleic acid molecules.

It is noted that Influenza A virus sequences are reported herein in positive sense, according to convention. Note that virus particles contain a reverse complement, negative polarity copy
30 of the sequence, while infected cells contain both polarities.

In some embodiments, reference to a DNA nucleic acid sequence, such as that set forth in SEQ ID NOs: 1 to 11; SEQ ID NO: 24 and SEQ ID NO:26, includes the corresponding RNA sequence. Conversely, reference to an RNA sequence, such as that set forth in SEQ ID
35 NOs: 14 and 15, includes the corresponding DNA sequence.

By way of a specific example, a “mutant PR8 segment 3 nucleic acid molecule” may have the sequence of SEQ ID NO:11. The mutant nucleic acid molecule may be DNA (e.g. if it is to be incorporated into an expression cassette such as pHW2000). Expression of the DNA nucleic acid molecule in a host may result in the production of virus that comprises an RNA sequence corresponding to SEQ ID NO:11. In this example, both the DNA and corresponding RNA sequence are mutant PR8 segment 3 nucleic acid molecules comprising the sequence of SEQ ID NO:11.

As used herein, the term “gene” refers to nucleic acid molecules which include an open reading frame encoding protein, and can further include non-coding regulatory sequences and introns.

As used herein, the term “PR8” refers to the influenza A/Puerto Rico/8/34 virus strain (also known herein as “A/PR/8/34”). Different PR8 lineages (e.g. variants) are known (see for example Johnson et al., PLOS, 2015). All such different lineages are encompassed herein by the term PR8.

The PR8 genome is made up of eight genomic segments (named segments 1 to 8 respectively). As described above, six genomic segments (also known as backbone segments herein) encode the “internal” proteins of the virus. These segments are known as segment 1 (also called the “PB2 backbone segment” or “PB2 segment”), segment 2 (also called the “PB1 backbone segment” or “PB1 segment”), segment 3 (also called the “PA backbone segment” or “PA segment”), segment 5 (also called the “NP backbone segment” or “NP segment”), segment 7 (also called the “M backbone segment” or “M segment”) and segment 8 (also called the “NS backbone segment” or NS segment). The additional two genomic segments are segment 4 (also known as the “HA segment”) and segment 6 (also known as the “NA segment”) which encode the surface glycoproteins of the viral particle (HA and NA respectively).

Accordingly, the phrase “PR8 segment 3 nucleic acid molecule” refers to a nucleic acid molecule that corresponds to genomic segment 3 of the PR8 influenza A virus strain. The PR8 segment 3 nucleic acid molecule may be DNA or RNA.

As used herein, the term “parent” PR8 segment 3 nucleic acid molecule refers to a PR8 segment 3 nucleic acid sequence that has not been mutated. The terms “parent”, “wildtype” and “WT” are used interchangeably herein. These terms encompass wild-type segment 3 nucleic acid sequences derived from different PR8 lineages. By way of example, but not by

way of limitation, the PR8 segment 3 nucleic acid molecules discussed in Johnson et al., PLOS, 2015 are considered to represent parent PR8 segment 3 nucleic acid molecules. The parent PR8 segment 3 nucleic acid molecule used by the inventors is shown in Figure 10 as SEQ ID NO:3.

5

The parent PR8 segment 3 nucleic acid molecule of the invention may have at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the nucleic acid sequence of SEQ ID NO:3 (or the corresponding RNA sequence). Preferably, the parent PR8 segment 3 nucleic acid molecule of the invention has at least 90%, at least 95%, at least 99% or has 100% sequence identity to the nucleic acid sequence of SEQ ID NO:3 (or the corresponding RNA sequence). Most preferably, the parent PR8 segment 3 nucleic acid molecule of the invention comprises or consists of (e.g. has) the nucleic acid sequence of SEQ ID NO:3 (or the corresponding RNA sequence).

10

15

Due to the degeneracy of the genetic code, it is possible to have the same polypeptide(s) encoded by several nucleic acids with different sequences. Thus, a "parent" PR8 segment 3 nucleic acid molecule includes nucleic acid molecules that have a different sequence to that of SEQ ID NO:3 but still encode the same polypeptides (i.e. still encode the same amino acid sequences as the PA and PA-X polypeptides encoded by the nucleic acid sequence of SEQ ID NO:3). The amino acid sequence of PR8 PA and PR8 PA-X are shown in figures 28 (SEQ ID NO: 23) and 20 (SEQ ID NO: 13) respectively.

20

Variations in the DNA and the amino acid sequence may stem from spontaneous mutations which can occur during passaging of the viruses. Such variant influenza strains can also be used as "parent" strains in the invention (i.e. a "parent" PR8 segment 3 nucleic acid molecule may be derived from one of such variant strains, and mutated (e.g. by the introduction of a mutation in the frameshift region, or the introduction of a premature stop codon in the X-ORF, or (partial) deletion of the X-ORF) to generate a mutant PR8 segment 3 nucleic acid molecule in accordance with the invention).

25

30

Variation in PA-X amino acid sequence is also encompassed within the invention. For example, parent PR8 segment 3 nucleic acid molecule may encode a PA-X polypeptide that has at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:13.

35

Calculations of sequence homology or identity (the terms are used interchangeably herein) between sequences are performed as follows.

To determine the percent identity of two amino acid sequences, or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 75%, 80%, 82%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the length of the reference sequence. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman et al. (1970) *J. Mol. Biol.* 48:444-453) algorithm which has been incorporated into the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a BLOSUM 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. A particularly preferred set of parameters (and the one that should be used if the practitioner is uncertain about what parameters should be applied to determine if a molecule is within a sequence identity or homology limitation of the invention) are a BLOSUM 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

Alternatively, the percent identity between two amino acid or nucleotide sequences can be determined using the algorithm of Meyers et al. (1989) *CAB/OS* 4:11-17) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

5

The nucleic acid and protein sequences described herein can be used as a “query sequence” to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (1990) *J. Mol. Biol.* 215:403-410). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to protein molecules of the invention. To obtain gapped alignments for comparison purposes, gapped BLAST can be utilized as described in Altschul et al. (1997, *Nucl. Acids Res.* 25:3389-3402). When using BLAST and gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See <<http://www.ncbi.nlm.nih.gov>>.

15

A parent segment 3 nucleic acid molecule comprises two open reading frames (ORFs); the PA ORF and more recently identified X-ORF. Ribosomal frameshifting during translation of parent segment 3 nucleic acid molecules into a +1 open reading frame of X-ORF results in the expression of PA-X, a protein that has recently been shown to be involved in shutting off host cell protein synthesis and at the whole animal level, modulating the immune response.

20

As used herein, the term “PA-X” refers to a protein that contains the N-terminal endonuclease domain of PA and a 61 amino acid (or 41 amino acid) C-terminus from the X ORF. By way of example, the PA-X produced by PR8 is a 29 kDa protein that contains the N-terminal endonuclease domain of PA and a 61 amino acid C-terminus from the X-ORF. The amino acid sequence of PR8 PA-X is shown in Figure 20. It is noted that the amino acid sequence of PA-X may vary from this (e.g. it may have at least 95%, at least 99% or 100% sequence identity to the sequence of PA-X shown in the figure 20, provided that it retains PA-X activity; which includes variable levels of host cell shut off activity (the ability to repress cellular RNA polymerase II-mediated gene expression) as well as the ability to enhance pathogenicity in chicken embryos).

25

30

35

As used herein, the term “mutant” PR8 segment 3 nucleic acid molecule refers to a PR8 segment 3 nucleic acid molecule that is not “wildtype” i.e. with a different genotype to the

corresponding parent PR8 segment 3 nucleic acid molecule. In other words, a mutant nucleic acid sequence comprises at least one nucleotide change i.e., deletion, insertion, or substitution, with respect to the corresponding region of the parent nucleic acid molecule. Molecular biological techniques for cloning and engineering genes and cDNAs, for mutating DNA, and for expressing polypeptides from polynucleotides in host cells are well known in the art as exemplified "Molecular cloning, a laboratory manual", third edition, Sambrook, J. & Russell, D.W. (eds), Cold Spring Harbor Laboratory press, Cold Spring Harbor, NY, incorporated herein by reference.

Any mutation within the PR8 segment 3 nucleic acid molecule that reduces the expression of PA-X is encompassed by the invention. The invention therefore encompasses any mutation in the PR8 segment 3 nucleic acid sequence that results in the desired phenotype (i.e. a reduction in PA-X expression).

By way of example, this includes mutations in the frameshift region that reduce or abolish ribosomal frameshifting to the +1 open reading frame of X-ORF. One such mutation is shown in Figure 18 (SEQ ID NO:11). Other mutations that reduce or abolish ribosomal frameshifting are well known and/or readily identifiable to a person of skill in the art. The frameshift region of RNA PR8 segment 3 has been well documented and has the consensus sequence of UCC UUU CGU C (+1 frameshift motif). Accordingly, any mutations within this region (or the corresponding DNA region) that reduce or prevent translation of the X-ORF are encompassed within the invention. Examples of appropriate mutations in this region include mutations to the "rare" CGU codon thought to stimulate frameshifting or mutations to the upstream UCC or UUU codons or final C nucleotide that hinder base-pairing of the tRNA molecules within a frameshifted ribosome.

As used herein, the phrase "frameshift region" (or "frameshift site" (FS)) refers to a nucleotide sequence (e.g. within the PR8 segment 3 nucleic acid sequence) which can induce ribosomal frameshifting during translation, such that, e.g. the ribosome continues translation in the +1 ORF (when previously translation occurred in the +0 ORF). This change in reading frame results in altered protein production. In this particular case, for example, translation of the (parent) PR8 segment 3 nucleic acid molecule in the +0 ORF produces the PA protein, whereas ribosomal frameshifting into the +1 ORF of segment 3 within the frameshift region results in a 29 kDa fusion protein (PA-X), comprising the amino terminus endonuclease domain of PA, and in most isolates a 61 amino acid C-terminus from the X ORF.

A reduction (e.g. to undetectable levels) in PA-X expression may also be achieved by introducing a premature stop codon (also known herein as a “premature termination codon” or “PTC”) into the X-ORF of segment 3. Examples of such mutations are shown in Figures 14 to 17 (SEQ ID NO:7 to 10 respectively). Other positions into which a PTC may be introduced into segment 3 to reduce or abolish PA-X expression are well known and/or readily identifiable to a person of skill in the art.

As used herein, “stop codon” means any codon in a DNA or RNA sequence which would indicate the point at which protein translation stops. Thus, in one example, the stop codon is selected from: TAG, TAA, and TGA (equivalent to UAG, UAA, and UGA, respectively, in RNA). As used herein, a “premature termination codon” or “premature stop codon” refers to the occurrence of a stop codon where a codon corresponding to an amino acid should be.

PA-X expression may also be reduced (e.g. abolished) by generating a mutant segment 3 nucleic acid molecule that lacks the full-length X-ORF (i.e. wherein the X-ORF is partially or completely deleted; e.g. the mutant PR8 segment 3 nucleic acid molecule does not comprise the full-length X-ORF of the parent PR8 segment 3 nucleic acid molecule). In other words, a mutant PR8 segment 3 nucleic acid molecule may be generated that only includes the PA ORF, and does not include the X-ORF.

Methods for generating mutant segment 3 nucleic acid molecules (such as those described herein) are well known in the art.

The invention also encompasses other mechanisms by which PA-X activity is reduced (e.g. mechanisms that induce an increase in PA-X protein degradation; or reduce PA-X protein stability). Such methods are also well known and/or readily identifiable by a person of skill in the art.

The mutant segment 3 nucleic acid molecule of the invention exhibits (e.g. has) reduced PA-X polypeptide expression compared to the PA-X polypeptide expression of a parent PR8 segment 3 nucleic acid molecule under the same expression conditions. In other words, when subjected to the same conditions as the parent, a virus that comprises the mutant PR8 segment 3 nucleic acid molecule will express less PA-X than the corresponding virus comprising the parent PR8 segment 3 nucleic acid molecule.

An example of the type of conditions (“expression conditions”) under which the difference in expression would be observed include *in vitro* translation of segment 3 mRNA in rabbit reticulocyte lysate.

Therefore, by way of example, a mutant PR8 segment 3 molecule of the invention may exhibit reduced PA-X polypeptide expression compared to the PA-X polypeptide expression of a parent PR8 segment 3 nucleic acid molecule when both the mutant (mRNA) and parent (mRNA) are subjected to *in vitro* translation in rabbit reticulocyte lysate, for the same period of time.

As used herein, “reduced PA-X polypeptide expression” refers to a reduction (i.e. decrease) in full length PA-X protein expression from the mutant nucleic acid molecule compared to the level (e.g. amount) of PA-X protein expression observed from the parent nucleic acid molecule (when the mutant and parent are subjected to equivalent (e.g. the same) conditions). By way of example, but not by way of limitation, mutant PR8 segment 3 nucleic acid molecules with a mutation in the frameshift region wherein the mutation reduces ribosomal frameshifting (i.e. reduces X-ORF translation) will exhibit reduced PA-X polypeptide expression. By way of an alternative example, but not by way of limitation, mutant PR8 segment 3 nucleic acid molecules that encode a truncated PA-X polypeptide (e.g. due to the presence of a premature stop codon within the X-ORF) also exhibit “reduced PA-X polypeptide expression” in the context of the invention, as a reduction (i.e. decrease) in full length PA-X protein expression will be observed for the mutant nucleic acid molecule compared to the level of full-length PA-X protein expression observed from the parent nucleic acid molecule (when the mutant and parent are subjected to equivalent (e.g. the same) conditions).

A “reduction” may be any observable decrease in expression, for example, at least a 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or 100% reduction in protein expression. The reduction in protein expression may mean that PA-X expression is not detectable when the mutant PR8 segment 3 nucleic acid molecule is expressed. There are several methods for measuring protein expression that are well known to the person of skill in the art. Accordingly, whether or not a particular mutant nucleic acid molecule exhibits reduced PA-X polypeptide expression compared to the parent nucleic acid molecule can easily be identified by a person of skill in the art.

The phrases “mutant PR8 segment 3 nucleic acid molecule” and “mutant nucleic acid molecule” are used interchangeably herein, unless the context specifies otherwise.

Reverse genetics (RG)

The invention is particularly suitable for producing a reassortant influenza virus through reverse genetics techniques. In these techniques, the viruses are produced in culture hosts (also called “host(s)” herein) using an expression cassette, or an expression system, wherein the expression system comprises at least one expression cassette.

Expression cassette

The invention provides an expression cassette comprising the mutant nucleic acid molecule (i.e. the mutant PR8 segment 3 nucleic acid molecule) described herein, wherein the nucleic acid molecule is functionally connected to a regulated or constitutive promoter.

As used herein, the term “expression cassette” refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. The terms “vector” and “expression cassette” are used interchangeably herein. The vector can be capable of autonomous replication or it can integrate into a host DNA. The vector may include restriction enzyme sites for insertion of recombinant DNA and may include one or more selectable markers. The vector can be a nucleic acid in the form of a plasmid, a bacteriophage or a cosmid. As an alternative, an expression cassette may be a linear expression construct. Preferably the vector is capable of propagation in a host cell and is stably transmitted to future generations. Such standard expression cassettes are well known in the art.

In one example, the expression cassette may comprise a pHW2000 plasmid of the type described by de Wit et al., 2004. Such plasmids are well known in the art.

As used herein, the phrase “functionally connected” is intended to mean that the nucleic acid sequence of interest is linked to the regulatory sequence(s) (promoter(s)) in a manner that allows for expression of the nucleic acid sequence (e.g. in an in vitro transcription/translation system or in a host cell when the vector is introduced into the host cell). See Auer H, Nature Biotechnol. 2006; 24: 41-43. The promoter therefore does not need to be directly upstream of the nucleic acid sequence; it may be located elsewhere in the expression cassette, provided that it allows for expression of the nucleic acid sequence.

"Regulatory sequences" as used herein, refers to, DNA or RNA elements that are capable of controlling gene expression. Examples of expression control sequences include promoters, enhancers, silencers, Shine Dalgarno sequences, TATA- boxes, internal ribosomal entry sites (IRES), attachment sites for transcription factors, transcriptional terminators, polyadenylation sites, RNA transporting signals or sequences important for UV-light mediated gene response. Preferably the vector includes one or more regulatory sequences operatively linked ("functionally connected") to the nucleic acid sequence to be expressed. Regulatory sequences include those which direct constitutive expression, as well as tissue-specific regulatory and/or inducible sequences.

"Promoter", as used herein, refers to the nucleotide sequences in DNA or RNA to which RNA polymerase binds to begin transcription. The promoter may be inducible ("regulated") or constitutively expressed. Alternatively, the promoter is under the control of a repressor or stimulatory protein.

As used herein, the term "regulated promoter" refers to a promoter that is activated in the cell in response to specific stimuli. Conversely, the term "constitutive promoter" refers to a promoter that is continuously active (in all circumstances in the cell). Both types of promoter are well known in the art.

"Transcriptional terminator" as used herein, refers to a DNA element, which terminates the function of RNA polymerases responsible for transcribing DNA into RNA. Preferred transcriptional terminators are characterized by a run of T residues preceded by a GC rich dyad symmetrical region.

"Translational control element", as used herein, refers to DNA or RNA elements that control the translation of mRNA. Preferred translational control elements are ribosome binding sites. Preferably, the translational control element is from a homologous system as the promoter, for example a promoter and its associated ribozyme binding site. Preferred ribosome binding sites are T7 or T3 ribosome binding sites.

"Restriction enzyme recognition site" as used herein, refers to a motif on the DNA recognized by a restriction enzyme.

"Selectable marker" as used herein, refers to proteins that, when expressed in a host cell, confer a phenotype onto the cell which allows a selection of the cell expressing said selectable marker gene. Generally this may be a protein that confers resistance to an

antibiotic such as ampicillin, kanamycin, chloramphenicol, tetracyclin, hygromycin, neomycin or methotrexate. Further examples of antibiotics are Penicillins; Ampicillin HCl, Ampicillin Na, Amoxycillin Na, Carbenicillin sodium, Penicillin G, Cephalosporins, Cefotaxim Na, Cefalexin HCl, Vancomycin, Cycloserine. Other examples include Bacteriostatic Inhibitors such as:

- 5 Chloramphenicol, Erythromycin, Lincomycin, Tetracyclin, Spectinomycin sulfate, Clindamycin HCl, Chlortetracycline HCl.

An expression cassette may be uni-directional or bi-directional. Use of bi-directional expression cassettes may reduce the total number of expression cassettes required by the host cell in order to produce a reassortant influenza A virus. Thus, the method of the invention may utilise at least one bi-directional expression construct wherein a nucleic acid molecule (e.g. gene or cDNA) is located between an upstream pol II promoter and a downstream non-endogenous pol I promoter (or vice versa). Transcription of the gene or cDNA from the pol II promoter produces capped positive-sense viral mRNA which can be translated into a protein, while transcription from the non-endogenous pol I promoter produces negative-sense vRNA.

Bi-directional expression cassettes contain at least two promoters which drive expression in different directions (i.e. both 5' to 3' and 3' to 5') from the same expression cassette. The two promoters can be functionally connected to different strands of the same double stranded DNA. Preferably, one of the promoters is a pol I promoter and at least one of the other promoters is a pol II promoter. This is useful as the pol I promoter can be used to express uncapped vRNAs while the pol II promoter can be used to transcribe mRNAs which can subsequently be translated into proteins, thus allowing simultaneous expression of RNA and protein from the same construct. Where more than one expression cassette is used within an expression system, the promoters may be a mixture of endogenous and non-endogenous promoters.

Preferably, RNA pol I and/or RNA pol II promoters are used.

30

The pol I and pol II promoters used in the expression cassette may be endogenous to an organism from the same taxonomic order from which the host cell is derived. Alternatively, the promoters can be derived from an organism in a different taxonomic order than the host cell. The term "order" refers to conventional taxonomic ranking, and examples of orders are primates, rodentia, carnivora, marsupialia, cetacean, etc. For example, the human pol I promoter can be used to express viral segments in canine cells (e.g. MDCK cells). Another

35

example of a promoter that may be used (e.g. together with the human pol I promoter) is the CMV pol II promoter.

5 It is possible to include more than one, for example two, three four, five or six viral segments on the same expression cassette.

A person of skill in the art will be aware of the molecular techniques available for the preparation of expression cassettes.

10 The design of the expression cassette depends on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression cassettes of the invention can be introduced into host cells to thereby produce proteins or polypeptides as described herein.

15 The expression cassettes of the invention can be introduced into host cells using any technique known to those of skill in the art (for example by electroporation, DEAE-dextran, calcium phosphate precipitation, liposomes, microinjection, or microparticle-bombardment, although these examples are not limiting).

20 Expression system

The invention provides an expression system comprising the expression cassette of the invention.

25 An expression system may contain at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, at least eleven, at least twelve or more expression cassettes.

30 By way of example, in addition to comprising an expression cassette with the mutant PR8 segment 3 nucleic acid molecule of the invention, an expression system may further include at least one of:

- (vi) a PR8 segment 1 nucleic acid molecule;
 - (vii) a PR8 segment 2 nucleic acid molecule;
 - (viii) a PR8 segment 5 nucleic acid molecule;
 - (ix) a PR8 segment 7 nucleic acid molecule; and,
 - 35 (x) a PR8 segment 8 nucleic acid molecule,
- wherein each of (i) to (v) are functionally connected to a regulated or constitutive promoter.

Each of (i) to (v) may be part of the same or part of a different expression cassette than the expression cassette that contains the mutant PR8 segment 3 nucleic acid molecule (in other words, an expression cassette that includes the mutant PR8 segment 3 nucleic acid molecule may also include one, two, three, four or five of (i) to (v)).

5

The phrases “PR8 segment 1 nucleic acid molecule” and “parent PR8 segment 1 nucleic acid molecule” are used herein interchangeably. The PR8 segment 1 nucleic acid molecule may be DNA or RNA.

10 An example of a PR8 segment 1 nucleic acid molecule is shown in Figure 8 (SEQ ID NO:1). However, other PR8 segment 1 nucleic acid sequences are also encompassed; for example PR8 segment 1 nucleic acid sequences that are derived from different PR8 lineages. By way of example, but not by way of limitation, the PR8 segment 1 nucleic acid molecules discussed in Johnson et al., PLOS, 2015 are considered to represent PR8 segment 1
15 nucleic acid molecules.

The PR8 segment 1 nucleic acid molecule of the invention may have at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the nucleic acid sequence of SEQ ID NO:1 (or the corresponding RNA sequence).

20 Preferably, the parent PR8 segment 1 nucleic acid molecule of the invention has at least 90%, at least 95%, at least 99% or has 100% sequence identity to the nucleic acid sequence of SEQ ID NO:1 (or the corresponding RNA sequence). Most preferably, the parent PR8 segment 1 nucleic acid molecule of the invention comprises or consists of (e.g. has) the nucleic acid sequence of SEQ ID NO:1 (or the corresponding RNA sequence).

25

Due to the degeneracy of the genetic code, it is possible to have the same polypeptide(s) encoded by several nucleic acids with different sequences. Thus, a PR8 segment 1 nucleic acid molecule includes nucleic acid molecules that have a different sequence to that of SEQ ID NO:1 but still encode the same polypeptide (i.e. still encode the same amino acid
30 sequence of RNA polymerase subunit PB2 as the RNA polymerase subunit PB2 encoded by the nucleic acid sequence of SEQ ID NO:1). The amino acid sequence of PR8 RNA polymerase subunit PB2 is shown in figure 21 (SEQ ID NO: 16).

Variation in RNA polymerase subunit PB2 amino acid sequence is also encompassed within
35 the invention. For example, the PR8 segment 1 nucleic acid molecule may encode a PR8 RNA polymerase subunit PB2 polypeptide that has at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or has

100% sequence identity to the amino acid sequence of SEQ ID NO:16 (e.g. provided that the variant retains RNA polymerase subunit PB2 activity; RNA polymerase subunit PB2 activity being well documented in the art).

- 5 The phrases “PR8 segment 2 nucleic acid molecule” and “parent PR8 segment 2 nucleic acid molecule” are used herein interchangeably. The PR8 segment 2 nucleic acid molecule may be DNA or RNA.

An example of a PR8 segment 2 nucleic acid molecule is shown in Figure 9 (SEQ ID NO:2).
 10 However, other PR8 segment 2 nucleic acid sequences are also encompassed; for example PR8 segment 2 nucleic acid sequences that are derived from different PR8 lineages. By way of example, but not by way of limitation, the PR8 segment 2 nucleic acid molecules discussed in Johnson et al., PLOS, 2015 are considered to represent PR8 segment 2 nucleic acid molecules.

15

The PR8 segment 2 nucleic acid molecule of the invention may have at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the nucleic acid sequence of SEQ ID NO:2 (or the corresponding RNA sequence). Preferably, the parent PR8 segment 2 nucleic acid molecule of the invention has at least
 20 90%, at least 95%, at least 99% or has 100% sequence identity to the nucleic acid sequence of SEQ ID NO:2 (or the corresponding RNA sequence). Most preferably, the parent PR8 segment 2 nucleic acid molecule of the invention comprises or consists of (e.g. has) the nucleic acid sequence of SEQ ID NO:2 (or the corresponding RNA sequence).

25

Due to the degeneracy of the genetic code, it is possible to have the same polypeptide(s) encoded by several nucleic acids with different sequences. Thus, a PR8 segment 2 nucleic acid molecule includes nucleic acid molecules that have a different sequence to that of SEQ ID NO:2 but still encode the same polypeptide (i.e. still encode the same amino acid sequence of RNA polymerase subunit PB1 as the RNA polymerase subunit PB1 encoded by
 30 the nucleic acid sequence of SEQ ID NO:2). The amino acid sequence of PR8 RNA polymerase subunit PB1 is shown in figure 22 (SEQ ID NO: 17).

Variation in RNA polymerase subunit PB1 amino acid sequence is also encompassed within the invention. For example, the PR8 segment 2 nucleic acid molecule may encode a PR8
 35 RNA polymerase subunit PB1 polypeptide that has at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:17 (e.g. provided the

variant retains RNA polymerase subunit PB1 activity; RNA polymerase subunit PB1 activity being well documented in the art).

The phrases "PR8 segment 5 nucleic acid molecule" and "parent PR8 segment 5 nucleic acid molecule" are used herein interchangeably. The PR8 segment 5 nucleic acid molecule may be DNA or RNA.

An example of a PR8 segment 5 nucleic acid molecule is shown in Figure 11 (SEQ ID NO:4). However, other PR8 segment 5 nucleic acid sequences are also encompassed; for example PR8 segment 5 nucleic acid sequences that are derived from different PR8 lineages. By way of example, but not by way of limitation, the PR8 segment 5 nucleic acid molecules discussed in Johnson et al., PLOS, 2015 are considered to represent PR8 segment 5 nucleic acid molecules.

The PR8 segment 5 nucleic acid molecule of the invention may have at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the nucleic acid sequence of SEQ ID NO:4 (or the corresponding RNA sequence). Preferably, the parent PR8 segment 5 nucleic acid molecule of the invention has at least 90%, at least 95%, at least 99% or has 100% sequence identity to the nucleic acid sequence of SEQ ID NO:4 (or the corresponding RNA sequence). Most preferably, the parent PR8 segment 5 nucleic acid molecule of the invention comprises or consists of (e.g. has) the nucleic acid sequence of SEQ ID NO:4 (or the corresponding RNA sequence).

Due to the degeneracy of the genetic code, it is possible to have the same polypeptide(s) encoded by several nucleic acids with different sequences. Thus, a PR8 segment 5 nucleic acid molecule includes nucleic acid molecules that have a different sequence to that of SEQ ID NO:4 but still encode the same polypeptide (i.e. still encode the same amino acid sequence of nucleoprotein NP as the nucleoprotein NP encoded by the nucleic acid sequence of SEQ ID NO:4). The amino acid sequence of PR8 nucleoprotein NP is shown in figure 23 (SEQ ID NO: 18).

Variation in nucleoprotein NP amino acid sequence is also encompassed within the invention. For example, the PR8 segment 5 nucleic acid molecule may encode a nucleoprotein NP polypeptide that has at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:18 (e.g. provided that the

variant retains nucleoprotein NP activity; nucleoprotein NP activity being well documented in the art).

The phrases "PR8 segment 7 nucleic acid molecule" and "parent PR8 segment 7 nucleic acid molecule" are used herein interchangeably. The PR8 segment 7 nucleic acid molecule may be DNA or RNA.

An example of a PR8 segment 7 nucleic acid molecule is shown in Figure 12 (SEQ ID NO:5). However, other PR8 segment 7 nucleic acid sequences are also encompassed; for example PR8 segment 7 nucleic acid sequences that are derived from different PR8 lineages. By way of example, but not by way of limitation, the PR8 segment 7 nucleic acid molecules discussed in Johnson et al., PLOS, 2015 are considered to represent PR8 segment 7 nucleic acid molecules.

The PR8 segment 7 nucleic acid molecule of the invention may have at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the nucleic acid sequence of SEQ ID NO:5 (or the corresponding RNA sequence). Preferably, the parent PR8 segment 7 nucleic acid molecule of the invention has at least 90%, at least 95%, at least 99% or has 100% sequence identity to the nucleic acid sequence of SEQ ID NO:5 (or the corresponding RNA sequence). Most preferably, the parent PR8 segment 7 nucleic acid molecule of the invention comprises or consists of (e.g. has) the nucleic acid sequence of SEQ ID NO:5 (or the corresponding RNA sequence).

Due to the degeneracy of the genetic code, it is possible to have the same polypeptide(s) encoded by several nucleic acids with different sequences. Thus, a PR8 segment 7 nucleic acid molecule includes nucleic acid molecules that have a different sequence to that of SEQ ID NO:5 but still encode the same polypeptide (i.e. still encode the same amino acid sequences of matrix proteins M1 and M2 as the matrix proteins M1 and M2 encoded by the nucleic acid sequence of SEQ ID NO:5). The amino acid sequence of PR8 matrix proteins M1 and M2 are shown in figures 24 and 25 (SEQ ID NO: 19 and SEQ ID NO:20).

Variation in matrix proteins M1 and M2 amino acid sequence is also encompassed within the invention. For example, the PR8 segment 7 nucleic acid molecule may encode a M1 polypeptide that has at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:19 (e.g. provided that the variant retains matrix protein M1 activity; matrix protein M1 activity being well documented in the art); and it may encode a

M2 polypeptide that has at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:20 (e.g. provided that the variant retains matrix protein M2 activity; matrix protein M2 activity being well documented in the art).

5

The phrases “PR8 segment 8 nucleic acid molecule” and “parent PR8 segment 8 nucleic acid molecule” are used herein interchangeably. The PR8 segment 8 nucleic acid molecule may be DNA or RNA.

10 An example of a PR8 segment 8 nucleic acid molecule is shown in Figure 13 (SEQ ID NO:6). However, other PR8 segment 8 nucleic acid sequences are also encompassed; for example PR8 segment 8 nucleic acid sequences that are derived from different PR8 lineages. By way of example, but not by way of limitation, the PR8 segment 8 nucleic acid molecules discussed in Johnson et al., PLOS, 2015 are considered to represent PR8
15 segment 8 nucleic acid molecules.

The PR8 segment 8 nucleic acid molecule of the invention may have at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the nucleic acid sequence of SEQ ID NO:6 (or the corresponding RNA sequence).

20 Preferably, the parent PR8 segment 8 nucleic acid molecule of the invention has at least 90%, at least 95%, at least 99% or has 100% sequence identity to the nucleic acid sequence of SEQ ID NO:6 (or the corresponding RNA sequence). Most preferably, the parent PR8 segment 8 nucleic acid molecule of the invention comprises or consists of (e.g. has) the nucleic acid sequence of SEQ ID NO:6 (or the corresponding RNA sequence).

25

Due to the degeneracy of the genetic code, it is possible to have the same polypeptide(s) encoded by several nucleic acids with different sequences. Thus, a PR8 segment 8 nucleic acid molecule includes nucleic acid molecules that have a different sequence to that of SEQ ID NO:6 but still encode the same polypeptide (i.e. still encode the same amino acid
30 sequences of non-structural proteins NS1 and NEP as the non-structural proteins NS1 and NEP encoded by the nucleic acid sequence of SEQ ID NO:6). The amino acid sequence of PR8 non-structural proteins NS1 and NEP are shown in figures 26 and 27 (SEQ ID NO: 21 and SEQ ID NO:22).

35 Variation in non-structural proteins NS1 and NEP amino acid sequence is also encompassed within the invention. For example, the PR8 segment 8 nucleic acid molecule may encode a NS1 polypeptide that has at least 70%, at least 75%, at least 80%, at least

90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:21 (e.g. provided that the variant retains NS1 activity; NS1 activity being well documented in the art); and it may encode a NEP polypeptide that has at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:22 (e.g. provided that the variant retains NEP activity; NEP activity being well documented in the art).

The expression system of the invention may further comprise a segment 4 nucleic acid molecule and/or a segment 6 nucleic acid molecule, wherein each nucleic acid molecule is functionally connected to a regulated or constitutive promoter. Each nucleic acid molecule may be part of an expression cassette. The segment 4 nucleic acid molecule may be part of the same (or part of a different) expression cassette as the segment 6 nucleic acid molecule. Furthermore, the segment 4 nucleic acid molecule (and/or the segment 6 nucleic acid molecule) may be part of the same (or part of a different) expression cassette as the mutant PR8 segment 3 nucleic acid molecule of the invention.

Usually, one (or most typically both) of the segment 4 and segment 6 nucleic acid molecules are from a different influenza A virus strain than PR8 (e.g. they are usually (both) derived from the virus strain against which a vaccine is being generated; i.e. the virus of interest).

The phrase "segment 4 nucleic acid molecule" refers to any nucleic acid molecule that encodes for influenza A virus hemagglutinin (HA), the major antigenic determinant of the virus. Examples of a segment 4 nucleic acid molecule and the encoded HA are shown in Figure 29 (SEQ ID NO:24 and SEQ ID NO:25 respectively). However, it is noted that the segment 4 nucleic acid molecule will vary for each virus of interest and that each of these different segment 4 nucleic acid molecules can be used in the context of the invention.

The phrase "segment 6 nucleic acid molecule" refers to any nucleic acid molecule that encodes for influenza A virus neuraminidase (NA), the second antigenic determinant of the virus. Examples of a segment 6 nucleic acid molecule and the encoded HA are shown in Figure 30 (SEQ ID NO: 26 and SEQ ID NO:27 respectively). However, it is noted that the segment 6 nucleic acid molecule will vary for each virus of interest and that each of these different segment 6 nucleic acid molecules can be used in the context of the invention.

As used herein, the phrase "virus of interest" refers to the viral strain against which a vaccine is being generated (in other words, the current circulating influenza virus or a newly

emerging virus of pandemic concern). The phrases “virus of interest” and “vaccine strain” are used interchangeably herein.

The polypeptides described herein can have amino acid sequences sufficiently or substantially identical to the amino acid sequence of SEQ ID NO:13, or any one of SEQ ID NO:16 to 23. The terms “sufficiently identical” or “substantially identical” are used herein to refer to a first amino acid or nucleotide sequence that contains a sufficient or minimum number of identical or equivalent (e.g. with a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences have a common structural domain or common functional activity. For example, amino acid or nucleotide sequences that contain a common structural domain having at least about 60%, or 65% identity, likely 75% identity, more likely 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity are defined herein as sufficiently or substantially identical. The invention therefore encompasses naturally occurring functional allelic variants of the polypeptides comprising the amino acid sequence shown in SEQ ID NO:13, or any one of SEQ ID NO:16 to 23.

Functional allelic variants will typically contain only conservative substitutions of one or more amino acids of SEQ ID NO:13, or any one of SEQ ID NO:16 to 23, or substitution, deletion or insertion of non-critical residues in non-critical regions of the protein. Nucleic acid molecules corresponding to functional natural allelic variants and homologues of the nucleic acid molecules of the invention can be isolated based on their homology to the nucleic acid molecules of the invention using the nucleotide sequences described in SEQ ID NO:1 to 6 or a portion thereof, as a hybridization probe under stringent hybridization conditions.

As used herein, the term “hybridizes under stringent conditions” describes conditions for hybridization and washing. Stringent conditions are known to those skilled in the art and can be found in available references (e.g., Current Protocols in Molecular Biology, John Wiley & Sons, N.Y., 1989, 6.3.1-6.3.6). Aqueous and non-aqueous methods are described in that reference and either can be used. A preferred example of stringent hybridization conditions are hybridization in 6× sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2× SSC, 0.1% (w/v) SDS at 50°C. Another example of stringent hybridization conditions are hybridization in 6× SSC at about 45°C, followed by one or more washes in 0.2× SSC, 0.1% (w/v) SDS at 55°C. A further example of stringent hybridization conditions are hybridization in 6× SSC at about 45°C, followed by one or more washes in 0.2× SSC, 0.1% (w/v) SDS at 60°C. Preferably, stringent hybridization conditions are

hybridization in 6× SSC at about 45°C, followed by one or more washes in 0.2× SSC, 0.1% (w/v) SDS at 65°C. Particularly preferred stringency conditions (and the conditions that should be used if the practitioner is uncertain about what conditions should be applied to determine if a molecule is within a hybridization limitation of the invention) are 0.5 molar sodium phosphate, 7% (w/v) SDS at 65°C, followed by one or more washes at 0.2× SSC, 1% (w/v) SDS at 65°C.

As used herein, a “naturally-occurring” nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

As used herein, the term “recombinant” refers to a biomolecule, for example a gene or a protein that (1) has been removed from its naturally occurring environment, (2) is not associated with all or a portion of a nucleic acid molecule as it is found in nature, (3) is operatively linked to a polynucleotide which it is not linked to in nature, or (4) does not occur in nature.

A “non-essential” amino acid residue is a residue that can be altered from the wild-type sequence of (e.g., the sequence of SEQ ID NO:13, or any one of SEQ ID NO:16 to 23) without abolishing or, more preferably, without substantially altering a biological activity, whereas an “essential” amino acid residue results in such a change. For example, amino acid residues that are conserved among the polypeptides of the present invention are predicted to be particularly non-amenable to alteration, except that amino acid residues in transmembrane domains can generally be replaced by other residues having approximately equivalent hydrophobicity without significantly altering activity.

A “conservative amino acid substitution” is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), non-polar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a nonessential amino acid residue in protein is preferably replaced with another amino acid residue from the same side chain family.

As used herein, a “biologically active portion” of protein or a protein portion with “biological activity” includes a fragment of protein that participates in an interaction between molecules and non-molecules. Biologically active portions of protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequences of the protein, e.g., the amino acid sequences shown in SEQ ID NO: 13, or any one of SEQ ID NO: 16 to 23, which include fewer amino acids than the full length protein, and exhibit at least one activity of the encoded protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the protein.

Any of the nucleic acid molecules described previously may comprise specific changes in the nucleotide sequence so as to optimize codons and mRNA secondary structure for translation in the host cell. Preferably, the codon usage of the nucleic acid is adapted for expression in the host cell, for example codon optimisation can be achieved using CalcGene, Hale, RS and Thomas G. *Protein Exper. Purif.* 12, 185-188 (1998), UpGene, Gao, W *et al. Biotechnol. Prog.* 20, 443-448 (2004), or Codon Optimizer, Fuglsang, A. *Protein Exper. Purif.* 31, 247-249 (2003). Amending the nucleic acid according to the preferred codon optimization can be achieved by a number of different experimental protocols, including, modification of a small number of codons, Vervoort *et al. Nucleic Acids Res.* 25: 2069-2074 (2000), or rewriting a large section of the nucleic acid sequence, for example, up to 1000 bp of DNA, Hale, RS and Thomas G. *Protein Exper. Purif.* 12, 185-188 (1998). Rewriting of the nucleic acid sequence can be achieved by recursive PCR, where the desired sequence is produced by the extension of overlapping oligonucleotide primers, Prodromou and Pearl, *Protein Eng.* 5: 827-829 (1992). Rewriting of larger stretches of DNA may require up to three consecutive rounds of recursive PCR, Hale, RS and Thomas G. *Protein Exper. Purif.* 12, 185-188 (1998), Te'o *et al, FEMS Microbiol. Lett.* 190: 13-19, (2000).

Alternatively, the level of cognate tRNA can be elevated in the host cell. This elevation can be achieved by increasing the copy number of the respective tRNA gene, for example by inserting into the host cell the relevant tRNA gene on a compatible multiple copy plasmid, or alternatively inserting the tRNA gene into the expression vector itself.

Any of the nucleic acid molecules described previously may comprise specific changes in the nucleotide sequence so as to optimize expression, activity or functional life of the encoded polypeptide(s). Preferably, the nucleic acids described previously are subjected to genetic manipulation and disruption techniques. Various genetic manipulation and disruption techniques are known in the art including, but not limited to, DNA Shuffling (US

6,132,970, Punnonen J et al, *Science & Medicine*, 7(2): 38-47, (2000), US 6,132,970), serial mutagenesis and screening. One example of mutagenesis is error-prone PCR, whereby mutations are deliberately introduced during PCR through the use of error-prone DNA polymerases and reaction conditions as described in US 2003152944, using for example commercially available kits such as The GeneMorph® II kit (Stratagene®, US). Randomized DNA sequences are cloned into expression vectors and the resulting mutant libraries screened for altered or improved protein activity.

Host cell

The invention also provides a eukaryotic host comprising the mutant nucleic acid molecule, expression cassette or the expression system of the invention.

The terms “host”, “host cell”, “cell” and “culture host” are used interchangeably herein. These terms are intended to refer to a cell into which a mutant nucleic acid molecule, expression cassette, expression system or virus of the invention has been (or will be) introduced. It should be understood that such terms refer not only to the particular subject cell but to the progeny, or potential progeny, of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but can be still included within the scope of the term as used herein.

The host for use in the present invention can be any eukaryotic cell that can produce the virus of interest.

The host cell may be a cell line, however, primary cells may be used as an alternative. Typically, the host cell is mammalian or avian. The host may be an embryonated hen egg or a mammalian cell. Suitable mammalian cells include, but are not limited to, hamster, cattle, primate (including humans and monkeys) and dog cells. Various cell types may be used, such as kidney cells, fibroblasts, retinal cells, lung cells, etc. Suitable monkey cells are e.g. African green monkey cells, such as kidney cells as in the Vero cell line. Suitable dog cells are e.g. kidney cells, such as the CLDK and MDCK cell lines.

Further suitable cells include, but are not limited to: CHO; 293T; BHK; MRC 5; PER.C6; FRhL2; WI-38; etc. Suitable cells are widely available e.g. from the American Type Cell Culture (ATCC) collection, from the Coriell Cell Repositories, or from the European Collection of Cell Cultures (ECACC). Preferably, the host cell is a mammalian cell such as MDCK, Vero or PER6.

Preferred cells for use in the invention are MDCK cells, which are derived from Madin Darby canine kidney. The original MDCK cells are available from the ATCC as CCL 34, however, derivatives of MDCK cells can be used. By way of example, derivatives that have been adapted for growth in suspension culture are well known in the art. An example of a suitable MDCK cell line is MDCK 33016 (DSM ACC2219).

More than one cell type may be used within the context of the invention. Alternatively, cells from a single cell line may be used. Furthermore, the same cell line may be used for reassorting the virus and for any subsequent propagation of the virus.

Preferably, the cells are cultured in the absence of serum, to avoid a common source of contaminants. Various serum-free media for eukaryotic cell culture are known to the person skilled in the art (e.g. Iscove's medium, ultra CHO medium (BioWhittaker), EX-CELL (JRH Biosciences)). Furthermore, protein-free media may be used (e.g. PF-CHO (JRH Biosciences)). Otherwise, the cells for replication can also be cultured in the customary serum-containing media (e.g. MEM or DMEM medium with 0.5% to 10% of fetal calf serum).

The cells may be in adherent culture (i.e. the cells grow adherently) or in suspension.

As used herein, "adherent cells", or "cells grown adherently" refer to cells which adhere to the culture containers in a cell culture. Examples of adherent cells include monolayer cells, which are cells that form a single layer of cells on the surface of a culture container.

"Suspension cells", cells grown in a suspension" or "suspended cells" refer to cells which do not adhere to culture containers in a cell culture. Suspension cells can be grown in a "spin culture", which is a culture in which the culture medium is stirred continuously during the culture process.

Recombinant PR8 virus

The invention also provides a recombinant PR8 virus comprising the mutant PR8 segment 3 nucleic acid molecule of the invention.

As used herein, the term "recombinant PR8 virus" refers to a PR8 virus particle that comprises at least one recombinant biomolecule, in this case at least the mutant PR8 segment 3 nucleic acid molecule of the invention.

Advantageously, the recombinant PR8 virus may be used as a donor strain in classical reassortment methods to generate a reassortant influenza A virus (e.g. a candidate vaccine virus (CVV)). The recombinant PR8 virus may comprise at least one, at least two, at least three, at least four or at least five other viral backbone segment nucleic acid molecules in addition to the mutant PR8 segment 3 nucleic acid molecule. The additional viral backbone segment nucleic acid molecules may also be derived from PR8 (e.g. may be at least one, at least two, at least three, at least four or at least five of: PR8 segment 1 nucleic acid molecule, PR8 segment 2 nucleic acid molecule, PR8 segment 5 nucleic acid molecule, PR8 segment 7 nucleic acid molecule and PR8 segment 8 nucleic acid molecule; all of which have been described in detail herein).

Classical reassortment methods are described briefly herein and are well known to a person of skill in the art (Robertson et al., 2011).

Reassortant influenza A viruses and methods of production

The mutant nucleic acid molecule, expression cassette, expression system, host cell, or recombinant PR8 virus described herein may be used for the production of a reassortant influenza A virus.

Advantageously, reassortant Influenza A virus strains of the invention (that comprise the mutant PR8 segment 3 nucleic acid molecule of the invention) grow to higher viral titres (e.g. in MDCK cells and/or in eggs) in the same time period and under the same growth conditions compared with reassortant influenza A virus strains that comprise a parent PR8 segment 3 nucleic acid molecule.

As used herein, the phrases “same growth conditions” and “same expression conditions” are used interchangeably. Examples of such conditions are provided elsewhere herein.

As used herein, the terms “influenza A virus”, “IAV”, “influenza virus” and “virus” are used interchangeably, unless the context specifies otherwise.

As used herein, the phrase “candidate vaccine virus” (or “CVV”) refers to a virus (e.g. a reassortant virus of the invention) that is a candidate for vaccine production.

As used herein, the term “reassortant” refers to a mix of genomic segments from different viruses. Accordingly, a reassortant virus comprises at least one gene segment that is from a different virus (i.e. does not naturally occur within the virus per se).

In general, a reassortant influenza A virus will contain only one of each (genomic) backbone segment (i.e. a total of six backbone segments; one of each of segment 1, 2, 3 (e.g. mutant PR8 segment 3 as described herein), 5, 7 and 8). In addition, the reassortant influenza virus will typically contain one of each segment that encodes the antigenic determinants of the virus of interest (i.e. one segment 4 nucleic acid molecule and one segment 6 nucleic acid molecule).

When the reassortant viruses are reassortants comprising the backbone segments from a single donor strain (e.g. PR8), the reassortant viruses will generally include segments from the donor strain and the vaccine strain (the "virus of interest") in a ratio of 1:7, 2:6, 3:5, 4:4, 5:3, 6:2 or 7:1. Having a majority of segments from the donor strain, in particular a ratio of 6:2, is typical. The mutant PR8 segment 3 nucleic acid molecule of the invention is derived from PR8, therefore is considered as a backbone segment from PR8. Accordingly, a reassortant virus comprising parent PR8 segments 1, 2, 5, 7 and 8 together with mutant PR8 segment 3; and segment 4 and segment 6 from a different viral strain (a vaccine strain) would be considered to have a ratio of 6:2 (donor strain segments:vaccine strain segments). Reassortant viruses with a ratio of 6:2 are also referred to as 6:2 genotype reassortant (6:2R) viruses.

The invention also encompasses reassortants which comprise viral segments from more than one vaccine strain provided that the reassortant comprises a backbone according to the present invention (i.e. comprises at least the mutant PR8 segment 3 nucleic acid molecule of the invention). For example, the reassortant influenza viruses may comprise the HA segment (segment 4) from one vaccine strain and the NA segment (segment 6) from a different vaccine strain.

The reassortant viruses of the invention can grow to higher viral titres than the wild-type vaccine strain from which some of the viral segment(s) of the reassortant virus are derived in the same time (for example 12 hours, 24 hours, 48 hours or 72 hours) and under the same growth conditions. Furthermore, the reassortant viruses of the invention (i.e. those comprising the mutant PR8 nucleic acid molecule of the invention) can grow to higher viral titres than the equivalent reassortant virus wherein the mutant PR8 segment 3 nucleic acid molecule is replaced with the parent PR8 segment 3 nucleic acid molecule.

Viral titre can be determined by standard methods known to those of skill in the art (e.g. by measuring HA yield).

The reassortant viruses of the invention (i.e. those comprising the mutant PR8 nucleic acid molecule of the invention) can achieve a viral titre which is at least 10% higher, at least 20% higher, at least 50% higher, at least 100% higher, at least 200% higher, at least 500% higher, or at least 1000% higher than the viral titre of the equivalent reassortant virus wherein the mutant PR8 segment 3 nucleic acid molecule is replaced with the parent PR8 segment 3 nucleic acid molecule.

The invention is suitable for reassorting pandemic as well as inter-pandemic (seasonal) influenza vaccine strains. The reassortant influenza strains may contain the influenza A virus HA subtypes H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15 or H16. They may contain the influenza A virus NA subtypes N1, N2, N3, N4, N5, N6, N7, N8 or N9. The reassortants influenza strains may also contain the HA segment of an influenza B strain. Suitable vaccine strains will be readily identifiable by a person of skill in the art.

As used herein, the term "isolated" refers to in vitro preparation and/or isolation of a nucleic acid molecule, e.g., vector or plasmid, peptide or polypeptide (protein), or virus of the invention, so that it is not associated with in vivo substances, or is substantially purified from in vitro substances. An isolated virus is generally obtained by in vitro culture and propagation, and/or via passage in eggs, and is substantially free from other infectious agents.

As used herein, "substantially free" means below the level of detection for a particular infectious agent using standard detection methods for that agent.

The invention provides a method of preparing the reassortant viruses of the invention. The method comprises the steps of:

- (i) introducing an expression cassette (or expression system) of the invention into a host cell; and
- (ii) culturing the host cell of step (a) in order to produce the reassortant IAV; and optionally
- (iii) purifying the reassortant IAV obtained in step (b).

The methods may further comprise steps of: (iv) infecting a culture host with the virus obtained in step (ii) or step (iii); (v) culturing the culture host from step (iv) to produce further virus; and optionally (vi) purifying the virus obtained in step (v).

Suitable methods for introducing an expression cassette or expression system into a host cell; culturing the host cell; purifying the reassortant IAV; and infecting a culture host with a virus are well known to a person of skill in the art (for example see the standard methods described in Hoffmann et al., 2000; Robertson et al., 2011).

5

Vaccines and methods of production

The invention also provides a method of preparing a vaccine, comprising steps of (a) preparing a virus by the methods of any one of the embodiments described above and (b) preparing a vaccine from the virus.

10

Suitable methods for preparing a vaccine from a virus are well known to a person of skill in the art (for example see the standard methods described in Robertson et al., 2011; IFPMA, 2016).

15

The term "vaccine", as used herein, refers to an immunogenic composition for in vivo administration to a subject, which may be a primate, especially a human, to confer protection against a disease, particularly a viral disease.

20

Various forms of influenza virus vaccine are currently available, and vaccines are generally based either on live virus or on inactivated virus. Inactivated vaccines may be based on whole virions, split virions, or on purified surface antigens. Influenza antigens can also be presented in the form of virosomes. The invention can be used with any of these types of vaccine, but will typically be used with inactivated vaccines.

25

The step of preparing a vaccine from a virus may therefore involve inactivating the virus (i.e. generating an inactivated virus). As used herein, the term "inactivated virus" refers to previously virulent virus which has undergone treatment to inactivate, kill or otherwise modify the virus to substantially eliminate its virulent properties while retaining its characteristic property of immunogenicity. Methods for inactivating a virus are well known to a person of skill in the art.

30

A vaccine of the invention may be part of a composition (e.g. a pharmaceutical composition). A pharmaceutical composition may comprise the vaccine and a pharmaceutically acceptable excipient, adjuvant, diluent and/or carrier.

35

Pharmaceutical compositions may routinely contain pharmaceutically acceptable concentrations of salt, buffering agents, preservatives, compatible carriers, supplementary immune potentiating agents such as adjuvants and cytokines and optionally other therapeutic agents or compounds.

5

As used herein, "pharmaceutically acceptable" refers to a material that is not biologically or otherwise undesirable, i.e., the material may be administered to an individual along with the selected compound without causing any undesirable biological effects or interacting in a deleterious manner with any of the other components of the pharmaceutical composition in which it is contained.

10

Excipients are natural or synthetic substances formulated alongside an active ingredient (e.g. a compound of the invention), included for the purpose of bulking-up the formulation or to confer a therapeutic enhancement on the active ingredient in the final dosage form, such as facilitating drug absorption or solubility. Excipients can also be useful in the manufacturing process, to aid in the handling of the active substance concerned such as by facilitating powder flowability or non-stick properties, in addition to aiding *in vitro* stability such as prevention of denaturation over the expected shelf life. Pharmaceutically acceptable excipients are well known in the art. A suitable excipient is therefore easily identifiable by one of ordinary skill in the art. By way of example, suitable pharmaceutically acceptable excipients include water, saline, aqueous dextrose, glycerol, ethanol, and the like.

15

20

Adjuvants are pharmacological and/or immunological agents that modify the effect of other agents in a formulation. Pharmaceutically acceptable adjuvants are well known in the art. A suitable adjuvant is therefore easily identifiable by one of ordinary skill in the art.

25

Diluents are diluting agents. Pharmaceutically acceptable diluents are well known in the art. A suitable diluent is therefore easily identifiable by one of ordinary skill in the art.

Carriers are non-toxic to recipients at the dosages and concentrations employed and are compatible with other ingredients of the formulation. The term "carrier" denotes an organic or inorganic ingredient, natural or synthetic, with which the active ingredient is combined to facilitate the application. Pharmaceutically acceptable carriers are well known in the art. A suitable carrier is therefore easily identifiable by one of ordinary skill in the art.

30

35

A vaccine of the invention may be used in the treatment or prophylaxis of influenza in a subject.

As used herein, the terms “treat”, “treating” and “treatment” are taken to include an intervention performed with the intention of preventing the development or altering the pathology of a disorder or symptom. Accordingly, “treatment” refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathological disorder or symptom.

As used here in the term “subject” refers to an individual, e.g., a human, pig, horse, mouse, cow, rat etc having or at risk of having influenza. The subject may be a patient i.e. a subject in need of treatment in accordance with the invention. The subject may have already received treatment for influenza previously. Alternatively, the subject has not been treated prior to treatment in accordance with the present invention.

As used herein, the terms “disease” and “disorder” are used interchangeably.

The compounds, combinations and/or compositions described herein can be administered to the subject by any conventional route, including injection or by gradual infusion over time. The administration may, for example, be topical, oral, parenteral, intravenous, intraperitoneal, intramuscular, intravascular, intracavity, intranasal, intracerebral, intratracheal, intralesional, intraperitoneal, intratumoural, rectal, subcutaneous, transdermal, epidural, percutaneous, or by infusion.

The compositions described herein may be in a form suitable for the above modes of administration. For example, suitable forms for oral administration include a tablet or capsule; suitable forms for nasal administration or administration by inhalation include a powder or solution; suitable forms for parenteral injection (including intravenous, subcutaneous, intramuscular, intravascular or infusion) include a sterile solution, suspension or emulsion; suitable forms for topical administration include an ointment or cream; and suitable forms for rectal administration include a suppository.

The compositions described herein are for administration in an effective amount. An “effective amount” (or “therapeutically effective amount”) is an amount that alone, or together with further doses, produces the desired (therapeutic) response. The (therapeutically) effective amount to be used will depend, for example, upon the therapeutic objectives, the route of administration, and the condition of the patient. For example, the dosage of the vaccine for a given patient will be determined by the attending physician, taking into consideration various factors known to modify the action of drugs including severity and type

of disease, body weight, sex, diet, time and route of administration, other medications and other relevant clinical factors. The dosages and schedules may be varied according to the particular disease state and the overall condition of the patient. Therapeutically effective dosages may be determined by either *in vitro* or *in vivo* methods.

5

The compositions of the present invention are advantageously presented in unit dosage form.

Each one of the features described above in the context of any one of the mutant nucleic acid molecule, expression cassette, expression system, host cell, recombinant PR8 virus, reassortant influenza A virus, uses, methods of production/preparation, pharmaceutical compositions and vaccines applies equally to each aspect of the invention (where the context allows). Accordingly, features, integers, characteristics, compounds, chemical moieties or groups described in conjunction with a particular aspect, embodiment or example of the invention are to be understood to be applicable to any other aspect, embodiment or example described herein unless incompatible therewith.

10

15

20

Throughout the description and claims of this specification, the words “comprise” and “contain” and variations of them mean “including but not limited to”, and they are not intended to (and do not) exclude other moieties, additives, components, integers or steps.

25

Throughout the description and claims of this specification, the singular encompasses the plural unless the context otherwise requires. In particular, where the indefinite article is used, the specification is to be understood as contemplating plurality as well as singularity, unless the context requires otherwise.

Aspects of the invention are demonstrated by the following non-limiting examples.

Examples

30

Materials and methods

Reverse genetics rescue of viruses

35

293T cells grown in 293T growth medium were seeded the day before transfection into 6 cm dishes (Nunc) to reach approximately 70-80% confluency of the next day. On the day of transfection, cell culture medium was changed to Opti-MEM (Gibco) 2 hours before transfection. 293T cells were transfected with eight pHW2000 plasmids each encoding one of the influenza segments in Opti-MEM. The mixture was left at room temperature for 5 mins Lipofectamine™ 2000 (Invitrogen) was added to Opti-MEM. The plasmids and transfection

reagent were mixed gently and left at room temperature for 20 mins. The transfection mixture was added to 293T cells. Cells were incubated at 37°C, 5% CO₂ before medium was replaced with 293T serum-free virus growth medium (6 ml). Cells were incubated at 37°C, 5% CO₂. At 2 days post-transfection, trypsin (0.5 µg/ml) was added to cells. Cell culture supernatants were harvested at 3 days post-transfection. 293T cell culture supernatants were clarified and used to infect 11 day old embryonated hens eggs. At 3 days post-infection, eggs were chilled overnight and virus stocks were sequenced to confirm identity. Constructs with a mutation of the frameshift site in influenza segment 3 (FS) are as described by Jagger et al., 2012. Sequences are provided of mutant forms FS and C-terminally truncated forms PTC1-4.

Site-directed mutagenesis

The QuikChange® Lightning site-directed mutagenesis kit (Stratagene) was used for mutagenesis according to the manufacturer's instructions. Primers used for site-directed mutagenesis of the segment 3 gene were designed using the primer design tool from Agilent technologies. A PCR mix was prepared containing a final concentration 1x kit reaction buffer, 100 ng plasmid DNA, 125 ng forward primer, 125 ng reverse primer, 0.75 µl Quiksoln reagent, 0.5 µl Quikchange lightning enzyme in a total volume of 25 µl. An initial denaturing step of 95°C for 2 min was followed by 18 cycles of 95°C for 20 secs, 60°C for 10 secs, 68°C for 2.5 mins (30 secs per Kb of plasmid length), then final extension 68°C for 5 mins. The PCR product was then digested with kit DpnI to remove all of the original plasmid. 2 µl of kit DpnI was added to each reaction and incubated for 30 mins at 37°C. 5 µl of sample was separated by gel electrophoresis on 1% agarose gel to confirm the absence of plasmid DNA and 7-10 µl of reaction was used for transformation.

Virus titration

Plaque assay

Plaque assays quantify infectious virus in a sample by ability of virus to form plaque forming units (PFU). In this assay, confluent monolayers of MDCK or MDCK-SIAT cells were infected with ten-fold serial dilutions of virus for 1 hour at room temperature, shaking every 10 mins to avoid cells drying. After removal of the inoculum, cells were overlaid with avicel overlay containing 1 µg/ml TPCK trypsin. The cytopathic effect (CPE) as a result of this localised infection can be visualised by formation of plaques, each corresponding to a single infectious unit. Plaque assays were fixed and stained at various times p.i. with toluidine blue or by immunostaining for influenza nucleoprotein (NP), as indicated in the results later on. Plaque numbers are scored.

Toluidine blue staining:

After removal of the plaque assay overlay, cells were fixed and stained in PBS solution containing both 4.1% formaldehyde and 0.2% toluidine blue overnight at room temperature. The solution was washed from the plates to visualise plaques.

5

Immunostaining for NP:

After removal of the plaque assay overlay or virus growth medium, cells were fixed with 4.1% formaldehyde/PBS for 20 mins at room temperature. Cells were washed three times with PBS. Cells were permeabilised with 0.2% Triton X 100/PBS for 20 mins at room temperature. Cells were washed three times with PBS. Cells were then immunostained with mouse monoclonal influenza A nucleoprotein (NP) specific 1° Ab in 5% BSA/PBS for 1 hour at room temperature. Cells were washed three times with PBS. Cells were immunostained with anti-mouse horse radish peroxidase (HRP) conjugated 2° Ab diluted in 5% BSA/PBS for 1 hour at room temperature. Cells were again washed three times with PBS. Plaques were visualised by action of HRP on a tetra methyl benzidine based (TMB) substrate, TrueBlue™ (KPL # 50-78-02) for 10 mins at room temperature. Cells were washed twice with dH2O and plates were dried.

10

15

Haemagglutinin assay

In a v-bottom 96-well plate, 50 µl of virus is added to 50 µl PBS and serially diluted two-fold across the plate. 50 µl of 1% chicken red blood cells were added and the plate 50 µl incubated for 30 mins at room temperature. HA titre is the reciprocal of the last virus dilution where agglutination can be observed.

20

Virus purification

Allantoic fluid from infected eggs were clarified of cells by centrifugation twice at 6,500 x g for 10 mins. Virus was partially purified from allantoic by ultracentrifugation at 128,000 x g rpm for 1.5 hours at 4°C through a 30% sucrose cushion.

25

For further purification, virus pellets were resuspended in PBS and loaded onto 15-60% sucrose density gradients and centrifuged at 210,000 x g for 40 mins (no brake) at 4°C. Virus bands were extracted from gradients and virus was pellet by centrifugation at 128,000 x g rpm for 1.5 hours at 4°C. Pellets were resuspended in equivalent volumes of PBS and either treated or untreated with N-glycosidase F (New England Biolabs). Following treatment, virus pellets were lysed in SDS gel loading buffer. Samples were separated by SDS-PAGE on a 10% or 12% polyacrylamide gel under reducing conditions and protein bands were visualised by coomassie blue staining or detected by immunostaining in western blot.

30

35

Coomassie staining

Protein from samples in SDS gel loading buffer were separated by SDS-PAGE on 10% or 12% polyacrylamide gels under reducing conditions using the Mini-Protean® II electrophoresis cell system (BIORAD). Gels were run at 120-150V for 1-2 hours. Gels were fixed three times for 5 mins using 10% acetic, 30% methanol (in dH₂O) and stained for 1 hour at room temperature with 0.2% coomassie blue (fix solution). After removal of the stain, gels were washed by shaking overnight in 5% acetic acid, 15% methanol (in dH₂O). Gels were scanned using an Epson Scanner and bands were quantified using ImageJ.

Western blot

Proteins from samples in SDS gel loading buffer were separated by SDS-PAGE on 10% or 12 % polyacrylamide gels (mini gels) under reducing conditions. Gels were run at 120V-150V for 1-2 hours. Following electrophoresis, gels were equilibrated in transfer buffer (Biorad) for 15 mins and transferred onto nitrocellulose membranes using the turbo blot semi dry apparatus (Biorad) Mini-gels were transferred for 7 min at 100V. Membranes were blocked with 5% milk/PBS-Tween for 20 mins at room temperature. Membranes were washed three times for 5 mins with PBS. Membranes were incubated with 1° Ab in PBST for overnight at 4°C. Membranes were washed again three times for 5 mins with PBS. Membranes were incubated with 2° Ab in blocking solution for 45 mins at room temperature. Membranes were washed again three times for 5 mins with TBS. Membranes were either scanned on Odyssey v1.2 when a Dylight800 or Alexa fluor 680 -conjugated 2° Ab was used and bands were quantified using ImageStudio Lite software (Odyssey).

RNA extraction, RT-PCR and sequence analysis.

All viral RNA extractions were performed using the QIAamp viral RNA mini kit (QIAGEN) using on-column DNase digestion (QIAGEN). Reverse transcription was performed with the Uni12 primer using theVerso cDNA kit (Thermo Scientific). PCR reactions were performed using Pfu Ultra II fusion 145 HS polymerase (Stratagene #600674-51) according to the manufacturer's protocol.or Taq Polymerase (Invitrogen)) according to the manufacturer's protocol. PCR products were purified for sequencing by Illustra GFX PCR DNA and Gel Band Purification kit (GE Healthcare). Primers and purified DNA were sent to GATC biotech (Lightrun method) for sequencing. Sequences were analysed using the DNASTar software.

Data

Abolishing PA-X expression reduces pathogenicity in chicken embryos

Chicken embryos infected with FS mutant virus (Δ PA-X) show less pathology compared with WT viruses on examination of gross pathology and tissue sections (Figure 4). Lungs show necrosis as well as secondary autolysis when infected with WT virus but appear relatively normal when infected with FS virus. Immunohistochemistry for virus nucleoprotein, NP (shown in brown) shows that infection is systemic.

Improved reverse genetics PR8 backbone (lacking PA-X gene) for candidate vaccine viruses

Using reverse genetics, the inventors have developed a candidate vaccine virus (CVV) bearing the two antigenic determinants (glycoproteins) of the 2009 pandemic (pdm09) (H1N1) virus in the high yielding donor background of PR8 virus strain (containing six internal genes) which lacks PR8 PA-X (FS mutant), termed 6:2FS. This CVV virus gives on average 2-3 fold higher HA antigen yield in embryonated hen's eggs (depending on the assay used to quantify antigen yield) when compared with its counterpart wild type CVV 6:2 virus which expresses PA-X. Antigen yield of candidate vaccine viruses is assessed by two methods:

1. Haemagglutinin assay (HA assay) which quantifies the ability of virus antigen HA to agglutinate red blood cells;

2. Quantification of HA protein expression from partially purified virus preparations by separating and thus visualizing individual virus proteins by Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and densitometry.

HA yield in eggs of CVV PR8:pdm09 6:2FS viruses was compared with wild-type 6:2 viruses in several experiments using three independently made reverse genetics virus stocks, assessed by HA assay, is collated and shown in Figure 5. For each reverse genetics virus stock set, several experiments were performed where eggs were inoculated at different inoculation doses to determine which dose gave the highest titres. In each experiment, HA titres of allantoic fluid from individual eggs were determined. Each data point represents average HA titres for each virus at the inoculation dose which gave the maximum titre from each experiment. There was an average 3-fold increase in HA titres from the 6:2FS virus compared with its 6:2 counterpart, a significant difference using paired t-test.

HA yield in eggs of CVV PR8:pdm09 6:2FS viruses was compared with wild-type 6:2 viruses in several experiments using three independently made reverse genetics virus stocks, was also assessed by SDS-PAGE and densitometry of partially purified virus preparations. A

representative image of an SDS-PAGE gel with viral protein components from PR8:pdm09 (A/California/07/2009, Cal7 strain) 6:2 and 6:2FS viruses distinguished, using as blue bands using a protein stain called coomassie blue, is shown in Figure 6a i). Equivalent volumes of partially purified virus pellets of each virus were run on gels, but there was a consistently higher amount of virus proteins from the 6:2FS virus compared with the 6:2 virus, such as NP and M1, annotated on the gel. However, the major antigenic component of the virus, HA1, could not be clearly distinguished on such gels as it migrated at a similar size to the NP protein. Therefore, HA1 protein was detected by a specific antibody in western blot (Figure 6a ii)). Furthermore, quantification of glycoproteins from gels has been found to be more accurate when the carbohydrates are removed. Carbohydrates from glycoproteins are cleaved by addition of an enzyme, N-glycosidase F (PNGase F). From western blot, a clear increase in the amount of HA1 was observed from the 6:2FS virus preparations compared with the 6:2 virus. Figure 6b shows collated data from several experiments with three independent reverse genetics experiments. Each data point represents relative fluorescence intensity values quantified by western blot, for each virus at the inoculation dose which gave the maximum titre from each experiment. In one of our large scale experiments, HA1 yield of 6:2FS was ~ 20-fold higher compared to its 6:2 counterpart. However, in all other experiments, the 6:2FS virus gave an ~ 1.5-3-fold increase in HA1 yield when compared with the 6:2 virus. When the outlier was ignored, average HA1 yield from the other 7 experiments showed the CVV 6:2FS gave 2-fold higher yield than the 6:2 virus, a significant difference using paired t-test.

Conclusions

Using reverse genetics, the inventors have developed a candidate vaccine virus (CVV) bearing the two antigenic determinants (glycoproteins) of the 2009 pandemic (pdm09) (H1N1) virus in the high yielding donor background of PR8 virus strain (containing six internal genes) which lacks PR8 PA-X (FS mutant), termed 6:2FS. This CVV virus gives on average 2-3 fold higher HA antigen yield in embryonated hen's eggs than the wild-type 6:2 counterpart. This improved reverse genetics PR8 backbone may be applied to make candidate vaccines viruses by reverse genetics bearing glycoproteins of potential pandemic viruses with may give improved yields compared with viruses containing the PR8 PA-X gene.

References

Abelin, A., Colegate, T., Gardner, S., Hehme, N., Palache, A., (2011) *Vaccine* 29(6), 1135-8.

- Dawood, F. S., Iuliano, A. D., Reed, C., Meltzer, M. I., Shay, D. K., Cheng, P. Y., Bandaranayake, D., Breiman, R. F., Brooks, W. A., Buchy, P. Feikin, D. R., Fowler, K. B., Gordon, A., Hien, N. T., Horby, P., Huang, Q. S., Katz, M. A., Krishnan, A., Lal, R., Montgomery, J. M., Mølbak, K., Pebody, R., Presanis, A. M., Razuri, H., Steens, A., Tinoco, Y. O., Wallinga, J., Yu, H., Vong, S., Bresee, J. and Widdowson, M. A. (2012), Estimated Global Mortality Associated with the First 12 Months of 2009 Pandemic Influenza A H1N1 Virus Circulation: a Modelling Study. *The Lancet Infectious Diseases*, 12(9): 687-95.
- Desmet, Emily A, Bussey, Kendra A, Stone, Raychel and Takimoto, Toru (2013), Identification of the N-Terminal Domain of the Influenza Virus PA Responsible for the Suppression of Host Protein Synthesis. *Journal of virology*, 87(6): 3108–3118.
- de Wit E, Spronken MI, Bestebroer TM, Rimmelzwaan GF, Osterhaus AD, Fouchier RA. (2004). Efficient generation and growth of influenza virus A/PR/8/34 from eight cDNA fragments. *Virus Res.* 103:155-61.
- Gao, Huijie, Sun, Yipeng, Hu, Jiao, Qi, Lu, Wang, Jinliang, Xiong, Xin, Wang, Yu, He, Qiming, Lin, Yang, Kong, Weili and others (2015), The Contribution of PA-X to the Virulence of Pandemic 2009 H1N1 and Highly Pathogenic H5N1 Avian Influenza Viruses. *Scientific reports*, 5:8262.
- Gao, Huijie, Sun, Honglei, Hu, Jiao, Qi, Lu, Wang, Jinliang, Xiong, Xin, Wang, Yu, He, Qiming, Lin, Yang, Kong, Weili and others (2015), The 20 Amino Acids at the C-Terminus of PA-X Are Associated with Increased Influenza A Virus Replication and Pathogenicity. *Journal of General Virology*: vir-0.[Epub ahead of print]
- Hoffmann E, Neumann G, Kawaoka Y, Hobom G, Webster RG. A DNA transfection system for generation of influenza A virus from eight plasmids. (2000). *Proc Natl Acad Sci U S A.* 97:6108-13.
- Hayashi, Tsuyoshi, MacDonald, Leslie A and Takimoto, Toru (2015), Influenza A Virus Protein PA-X Contributes to Viral Growth and Suppression of the Host Antiviral and Immune Responses. *Journal of virology*: JVI-00319. [Epub ahead of print]
- Hu, Jiao, Mo, Yiqun, Wang, Xiaoquan, Gu, Min, Hu, Zenglei, Zhong, Lei, Wu, Qiwen, Hao, Xiaoli, Hu, Shunlin, Liu, Wenbo and others (2015), PA-X Decreases the Pathogenicity of

Highly Pathogenic H5N1 Influenza A Virus in Avian Species by Inhibiting Virus Replication and Host Response. *Journal of virology*, 89(8): 4126–4142.

International Federation of Pharmaceutical Manufacturers and Associations (IFPMA),

- 5 Vaccine manufacture:Virus processing (2016) accessed 10 February 2016
< <http://www.ifpma.org/resources/influenza-vaccines/influenza-vaccines/vaccine-manufacture-virus-processing.html>>

- Jagger, BW, Wise, HM, Kash, JC, Walters, K-A, Wills, NM, Xiao, Y-L, Dunfee, RL,
10 Schwartzman, LM, Ozinsky, A, Bell, GL and others (2012), An Overlapping Protein-Coding Region in Influenza A Virus Segment 3 Modulates the Host Response. *Science*, 337(6091): 199–204.

- Johnson, N. and Mueller, J. (2002), Updating the Accounts: Global Mortality of the 1918-
15 1920“ Spanish” Influenza Pandemic. *Bulletin of the History of Medicine*, 76(1): 105-115.

Khapersky, Denys A, Emara, Mohamed M, Johnston, Benjamin P, Anderson, Paul, Hatchette, Todd F and McCormick, Craig (2014), Influenza A Virus Host Shutoff Disables Antiviral Stress-Induced Translation Arrest. *PLoS pathogens*, 10(7): e1004217.

- 20 Khapersky, Denys A and McCormick, Craig (2015), Timing Is Everything: Coordinated Control of Host Shutoff by Influenza A Virus NS1 and PA-X Proteins. *Journal of virology*: JVI-00386. [Epub ahead of print]

- 25 Murti, K. G. and Webster, R. G. (1986), Distribution of Hemagglutinin and Neuraminidase on Influenza Virions as Revealed by Immunoelectron Microscopy. *Virology*, 149(1): 36-43.

- Naffakh, Nadia, Massin, Pascale and van der Werf, Sylvie (2001), The Transcription/replication Activity of the Polymerase of Influenza A Viruses Is Not Correlated
30 with the Level of Proteolysis Induced by the PA Subunit. *Virology*, 285(2): 244–252.

Neumann G, Watanabe T, Ito H, Watanabe S, Goto H, Gao P, Hughes M, Perez DR, Donis R, Hoffmann E, Hobom G, Kawaoka Y. (1999). Generation of influenza A viruses entirely from cloned cDNAs. *Proc Natl Acad Sci U S A*. 96:9345-50.

- 35 Robertson, James S, Nicolson, Carolyn, Harvey, Ruth, Johnson, Rachel, Major, Diane, Guilfoyle, Kate, Roseby, Sarah, Newman, Robert, Collin, Rebecca, Wallis, Chantal and

others. (2011). The Development of Vaccine Viruses against Pandemic A (H1N1) Influenza. *Vaccine*, 29(9): 1836–1843.

Claims

1. A mutant PR8 segment 3 nucleic acid molecule that exhibits reduced PA-X polypeptide expression compared to the PA-X polypeptide expression of a parent PR8 segment 3 nucleic acid molecule under the same expression conditions.

5

2. The nucleic acid molecule of claim 1, wherein the parent PR8 segment 3 nucleic acid molecule encodes a PA-X polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO: 13.

10

3. The nucleic acid molecule of claim 1 or 2, wherein the parent PR8 segment 3 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity to the nucleic acid sequence of SEQ ID NO: 3.

15

4. The nucleic acid molecule of any preceding claim, wherein the mutant PR8 segment 3 nucleic acid molecule:

(a) comprises a mutation in the frameshift region of the parent PR8 segment 3 nucleic acid molecule, wherein the mutation reduces the expression of the PA-X polypeptide compared to the PA-X expression of the parent PR8 segment 3 nucleic acid molecule under the same expression conditions;

20

(b) comprises a premature stop codon within the X-ORF of the parent PR8 segment 3 nucleic acid molecule, wherein the premature stop codon reduces the expression of the PA-X polypeptide compared to the PA-X expression of the parent PR8 segment 3 nucleic acid molecule under the same expression conditions; and/or

(c) does not comprise the X-ORF of the parent PR8 segment 3 nucleic acid molecule.

25

5. The nucleic acid molecule of any preceding claim, wherein the parent PR8 segment 3 nucleic acid molecule comprises a frameshift region comprising the nucleic acid sequence UCC UUU CGU C (SEQ ID NO: 14), and wherein the mutant PR8 segment 3 nucleic acid molecule comprises a mutation within this frameshift region.

30

6. The nucleic acid molecule of claim 5, wherein the mutant PR8 segment 3 nucleic acid molecule comprises a mutated frameshift region comprising the nucleic acid sequence AGC UUC AGA (SEQ ID NO: 15).

35

7. The nucleic acid molecule of any preceding claim, wherein the mutant PR8 segment 3 nucleic acid molecule comprises or consists of the nucleic acid sequence of any one of SEQ ID NOs: 7 to 11.

8. The nucleic acid molecule of any preceding claim, wherein the mutant PR8 segment 3 nucleic acid molecule exhibits at least a 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or 95% reduction in PA-X polypeptide expression compared to the PA-X polypeptide expression of the parent PR8 segment 3 nucleic acid molecule under the same expression conditions.

9. An expression cassette comprising the mutant nucleic acid molecule of any one of claims 1 to 8, wherein the nucleic acid molecule is functionally connected to a regulated or constitutive promoter.

10. The expression cassette of claim 9, wherein the promoter is an RNA Pol I or an RNA pol II promoter.

11. The expression cassette of any one of claims 9 to 10, wherein the expression cassette comprises a pHW2000 plasmid.

12. The expression cassette of any one of claims 9 to 11, wherein the expression cassette comprises or consists of the sequence of SEQ ID NO:12

13. An expression system comprising the expression cassette of any one of claims 9 to 12.

14. The expression system of claim 13, further comprising at least one of:

- (i) a PR8 segment 1 nucleic acid molecule;
- (ii) a PR8 segment 2 nucleic acid molecule;
- (iii) a PR8 segment 5 nucleic acid molecule;
- (iv) a PR8 segment 7 nucleic acid molecule; and,
- (v) a PR8 segment 8 nucleic acid molecule,

wherein each of (i) to (v) are functionally connected to a regulated or constitutive promoter.

15. The expression system of claim 14, wherein the at least one of (i) to (v) is part of the same or part of a different expression cassette than the expression cassette of any one of claims 9 to 12.

16. The expression system of claim 14 or 15, wherein:

- (i) the PR8 segment 1 nucleic acid molecule encodes a RNA polymerase subunit PB2 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:16; and /or
- (ii) the PR8 segment 2 nucleic acid molecule encodes a RNA polymerase subunit PB1 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:17; and/or
- (iii) the PR8 segment 5 nucleic acid molecule encodes a nucleoprotein NP polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:18; and/or
- (iv) the PR8 segment 7 nucleic acid molecule encodes a matrix protein M1 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:19 and a matrix protein M2 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:20; and/or
- (v) the PR8 segment 8 nucleic acid molecule encodes a non-structural protein NS1 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:21 and a non-structural protein NEP polypeptide that has at least 95%, at least 99% or at least 100% sequence identity to the amino acid sequence of SEQ ID NO:22.
17. The expression system of claim 16, wherein:
- (i) the PR8 segment 1 nucleic acid molecule encodes a RNA polymerase subunit PB2 polypeptide that has the amino acid sequence of SEQ ID NO:16;
- (ii) the PR8 segment 2 nucleic acid molecule encodes a RNA polymerase subunit PB1 polypeptide that has the amino acid sequence of SEQ ID NO:17;
- (iii) the PR8 segment 5 nucleic acid molecule encodes a nucleoprotein NP polypeptide that has the amino acid sequence of SEQ ID NO:18;
- (iv) the PR8 segment 7 nucleic acid molecule encodes a matrix protein M1 polypeptide that has the amino acid sequence of SEQ ID NO:19 and a matrix protein M2 polypeptide that has the amino acid sequence of SEQ ID NO:20;
- (v) the PR8 segment 8 nucleic acid molecule encodes a non-structural protein NS1 polypeptide that has the amino acid sequence of SEQ ID NO:21 and a non-structural protein NEP polypeptide that has the amino acid sequence of SEQ ID NO:22.
18. The expression system of any one of claims 14 to 17, wherein:
- (i) the PR8 segment 1 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 1; and/or

- (ii) the PR8 segment 2 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 2; and/or
 - (iii) the PR8 segment 5 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 4; and/or
 - 5 (iv) the PR8 segment 7 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 5; and/or
 - (v) the PR8 segment 8 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 6.
- 10 19. The expression system of claim 18, wherein:
- (i) the PR8 segment 1 nucleic acid molecule has the sequence of SEQ ID NO: 1;
 - (ii) the PR8 segment 2 nucleic acid molecule has the sequence of SEQ ID NO: 2;
 - (iii) the mutant PR8 segment 3 nucleic acid molecule has the sequence of SEQ ID NO: 11;
 - 15 (iv) the PR8 segment 5 nucleic acid molecule has the sequence of SEQ ID NO: 4;
 - (v) the PR8 segment 7 nucleic acid molecule has the sequence of SEQ ID NO: 5; and
 - (vi) the PR8 segment 8 nucleic acid molecule has the sequence of SEQ ID NO: 6.
20. The expression system of any one of claims 13 to 19, further comprising at least one of:
- 20 (i) an expression cassette comprising a segment 4 nucleic acid molecule, wherein the segment 4 nucleic acid molecule is from a different influenza A virus strain than PR8; and
 - (ii) an expression cassette comprising a segment 6 nucleic acid molecule, wherein the segment 6 nucleic acid molecule is from a different influenza A virus strain than PR8;
- wherein each of (i) to (ii) are functionally connected to a regulated or constitutive promoter.
- 25
21. The expression system of claim 20, wherein the segment 4 nucleic acid molecule and segment 6 nucleic acid molecule are from the same influenza A virus strain.
22. A eukaryotic host comprising the mutant nucleic acid molecule of any one of claims 1 to 8, the expression cassette of any one of claims 9 to 12, or the expression system of any one of claims 13 to 21.
- 30
23. The host of claim 22, wherein the host is an embryonated hen egg or a mammalian cell, optionally wherein the mammalian cell is an MDCK, Vero or PerC6 cell.
- 35
24. A recombinant PR8 virus comprising the mutant nucleic acid molecule of any one of claims 1 to 8.

25. Use of a mutant nucleic acid molecule according to any one of claims 1 to 8, an expression cassette according to any one of claims 9 to 12, an expression system according to any one of claims 13 to 21, a host cell according to claim 22 or 23, or a recombinant PR8 virus according to claim 24, for the production of a reassortant influenza A virus (IAV).

5

26. Use according to claim 25, wherein the reassortant IAV is a candidate vaccine virus.

27. A reassortant influenza A virus (IAV) comprising the nucleic acid molecule according to any one of claims 1 to 8.

10

28. The reassortant influenza A virus of claim 27, further comprising a PR8 segment 1 nucleic acid molecule, a PR8 segment 2 nucleic acid molecule, a PR8 segment 5 nucleic acid molecule, a PR8 segment 7 nucleic acid molecule, a PR8 segment 8 nucleic acid molecule, a segment 4 nucleic acid molecule from a different influenza A virus strain than PR8, and a segment 6 nucleic acid molecule from a different influenza A virus strain than PR8.

15

29. The reassortant influenza virus of claim 28, wherein the segment 4 nucleic acid molecule and segment 6 nucleic acid molecule are from the same influenza A virus strain.

20

30. The reassortant influenza virus of claim 28 or 29, wherein:

(i) the PR8 segment 1 nucleic acid molecule encodes a RNA polymerase subunit PB2 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:16; and /or

25

(ii) the PR8 segment 2 nucleic acid molecule encodes a RNA polymerase subunit PB1 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:17; and/or

(iii) the PR8 segment 5 nucleic acid molecule encodes a nucleoprotein NP polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:18; and/or

30

(iv) the PR8 segment 7 nucleic acid molecule encodes a matrix protein M1 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:19 and a matrix protein M2 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:20; and/or

35

(v) the PR8 segment 8 nucleic acid molecule encodes a non-structural protein NS1 polypeptide that has at least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:21 and a non-structural protein NEP polypeptide that has at

least 95%, at least 99% or has 100% sequence identity to the amino acid sequence of SEQ ID NO:22.

31. The reassortant influenza virus of claim 30, wherein:

- 5 (i) the PR8 segment 1 nucleic acid molecule encodes a RNA polymerase subunit PB2 polypeptide that has the amino acid sequence of SEQ ID NO:16;
- (ii) the PR8 segment 2 nucleic acid molecule encodes a RNA polymerase subunit PB1 polypeptide that has the amino acid sequence of SEQ ID NO:17;
- (iii) the PR8 segment 5 nucleic acid molecule encodes a nucleoprotein NP polypeptide that
- 10 has the amino acid sequence of SEQ ID NO:18;
- (iv) the PR8 segment 7 nucleic acid molecule encodes a matrix protein M1 polypeptide that has the amino acid sequence of SEQ ID NO:19 and a matrix protein M2 polypeptide that has the amino acid sequence of SEQ ID NO:20;
- (v) the PR8 segment 8 nucleic acid molecule encodes a non-structural protein NS1
- 15 polypeptide that has the amino acid sequence of SEQ ID NO:21 and a non-structural protein NEP polypeptide that has the amino acid sequence of SEQ ID NO:22.

32. The reassortant influenza virus of any one of claims 28 to 31, wherein:

- (i) the PR8 segment 1 nucleic acid molecule has at least 95%, at least 99%, or has
- 20 100% sequence identity with the sequence of SEQ ID NO: 1; and/or
- (ii) the PR8 segment 2 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 2; and/or
- (iii) the PR8 segment 5 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 4; and/or
- 25 (iv) the PR8 segment 7 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO:5; and/or
- (v) the PR8 segment 8 nucleic acid molecule has at least 95%, at least 99%, or has 100% sequence identity with the sequence of SEQ ID NO: 6.

30 33. The reassortant influenza virus of claim 32, wherein:

- (i) the PR8 segment 1 nucleic acid molecule has the sequence of SEQ ID NO: 1;
- (ii) the PR8 segment 2 nucleic acid molecule has the sequence of SEQ ID NO: 2;
- (iii) the mutant PR8 segment 3 nucleic acid molecule has the sequence of SEQ ID NO: 11;
- 35 (iv) the PR8 segment 5 nucleic acid molecule has the sequence of SEQ ID NO: 4;
- (v) the PR8 segment 7 nucleic acid molecule has the sequence of SEQ ID NO: 5; and the PR8 segment 8 nucleic acid molecule has the sequence of SEQ ID NO: 6.

34. A method of preparing a reassortant influenza A virus (IAV) comprising the steps of:

(a) introducing an expression cassette according to any one of claims 9 to 12, or an expression system according to any one of claims 13 to 21 into a host cell; and

- 5 (b) culturing the host cell of step (a) in order to produce the reassortant IAV; and optionally
(c) purifying the reassortant IAV obtained in step (b).

35. A method for producing an influenza A virus (IAV) comprising the steps of:

- 10 (a) infecting a host cell with a reassortant influenza A virus according to any one of claims 27
to 33 or a recombinant PR8 virus according to claim 24;

(b) culturing the host cell from step (a) to produce the influenza A virus; and optionally

(c) purifying the influenza A virus produced in step (b).

36. The method of claim 34 or 35, wherein the host cell is an embryonated hen egg.

15

37. The method of claim 34 or 35, wherein the culture host is a mammalian cell, optionally wherein the cell is an MDCK, Vero or PerC6 cell.

38. The method of claim 37, wherein the host cell grows adherently or in suspension.

20

39. An influenza A virus obtained by the method of any one of claims 34 to 38.

40. A method of preparing a vaccine, comprising the steps of (a) preparing a virus by the method of any one of claims 34 to 38 and (b) preparing a vaccine from the virus.

25

41. The method of claim 40, wherein step (b) involves inactivating the virus.

42. The method of claim 40 or 41, wherein the vaccine is a whole virion vaccine, a split virion vaccine, a surface antigen vaccine or a virosomal vaccine.

30

43. A pharmaceutical composition comprising a vaccine obtained by the method of any one of claims 40 to 42, and a pharmaceutically acceptable carrier or diluent.

35 44. A vaccine obtained by the method of any one of claims 40 to 42 and a pharmaceutically acceptable carrier or diluent.

45. A vaccine according to claim 44 for the treatment or prophylaxis of influenza.

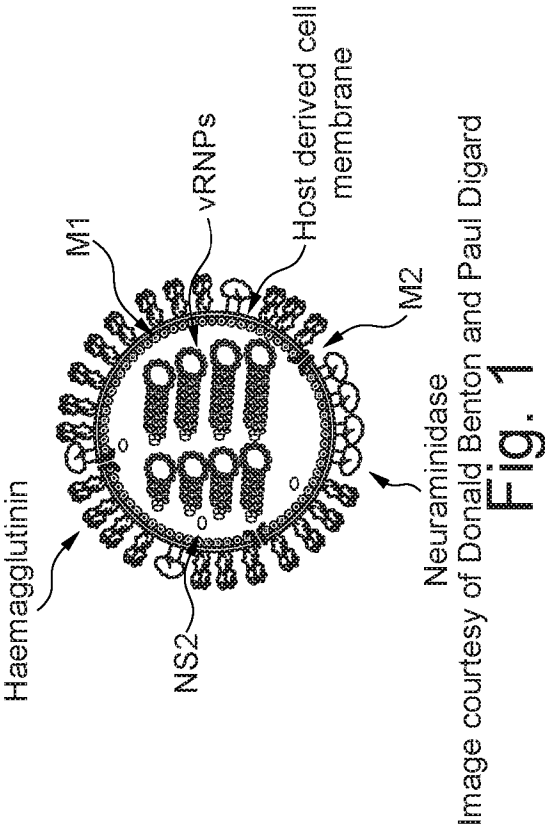


Image courtesy of Donald Benton and Paul Digard

Fig. 1

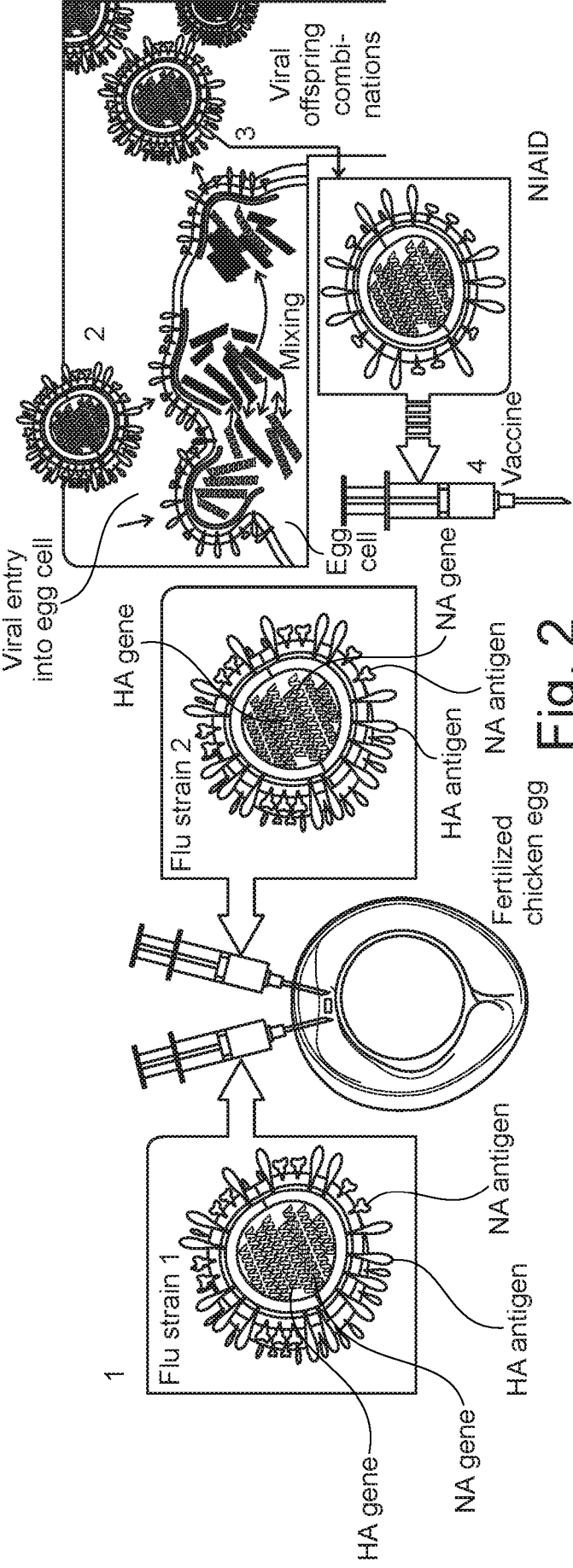


Fig. 2

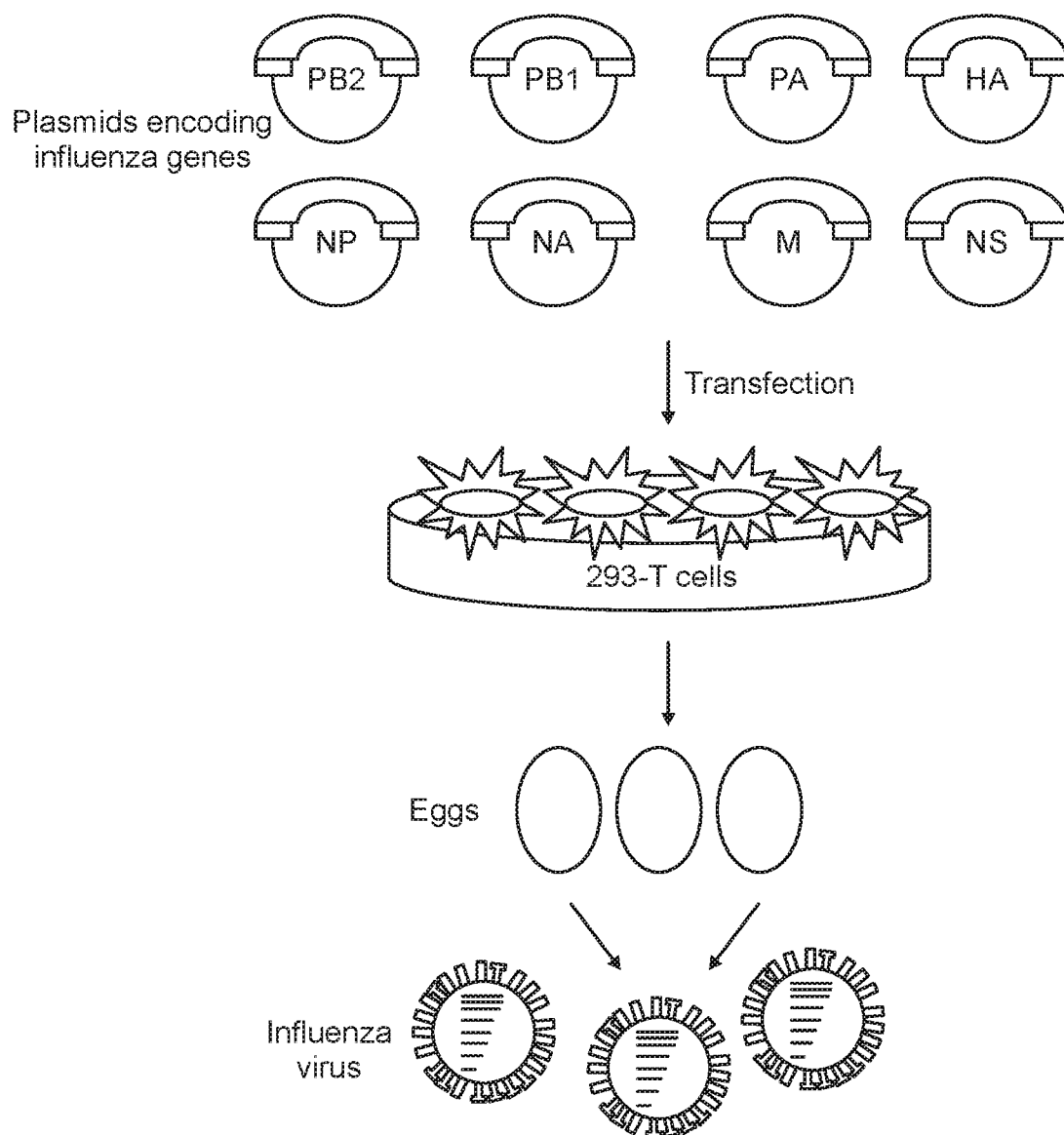


Fig. 3

4a

PR8 PA-X mutant viruses

48 hours p.i.

Mock

WT

FS



4b i)

Lung sections

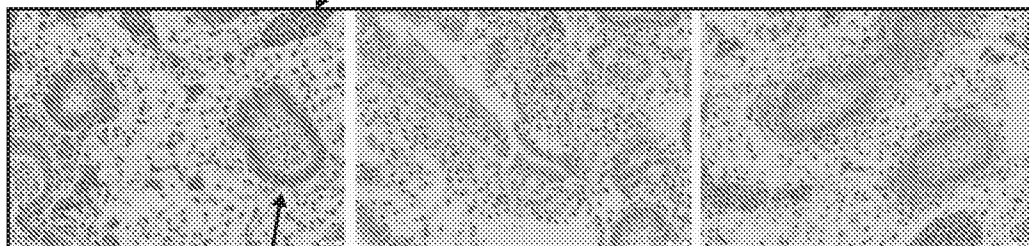
H & E staining:

Mock

Blood vessel

WT

FS



Bronchus

4b ii)

Immunostaining for influenza NP:

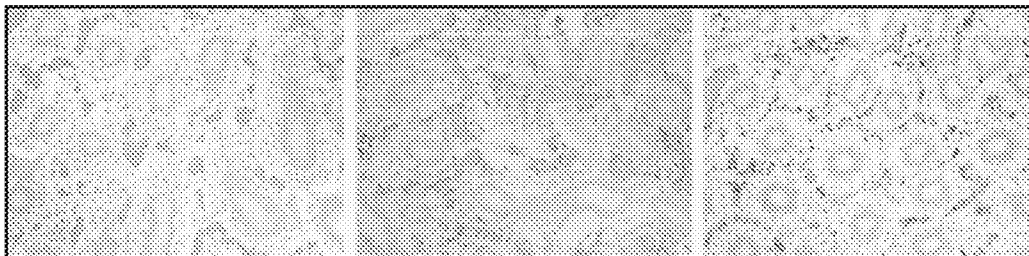
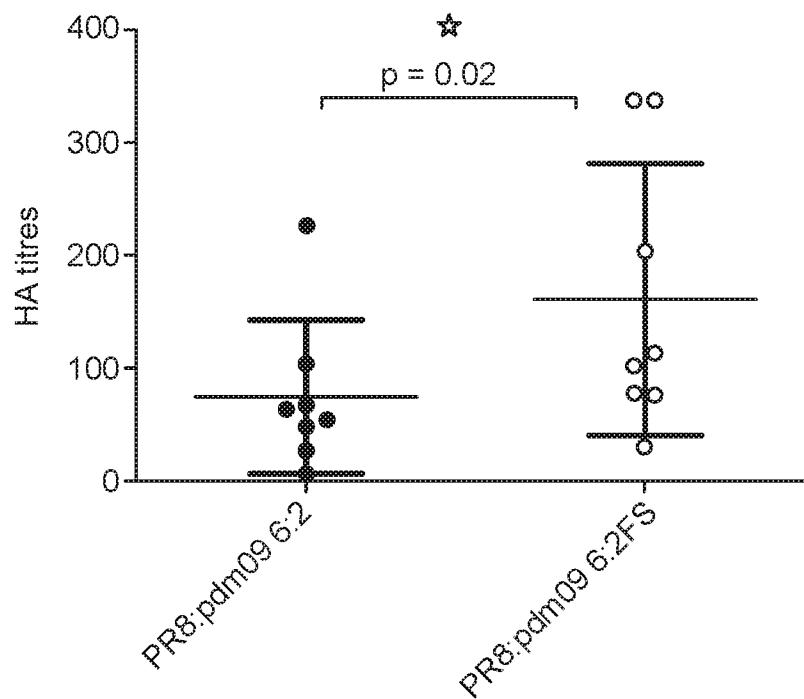


Fig. 4

5a



5b

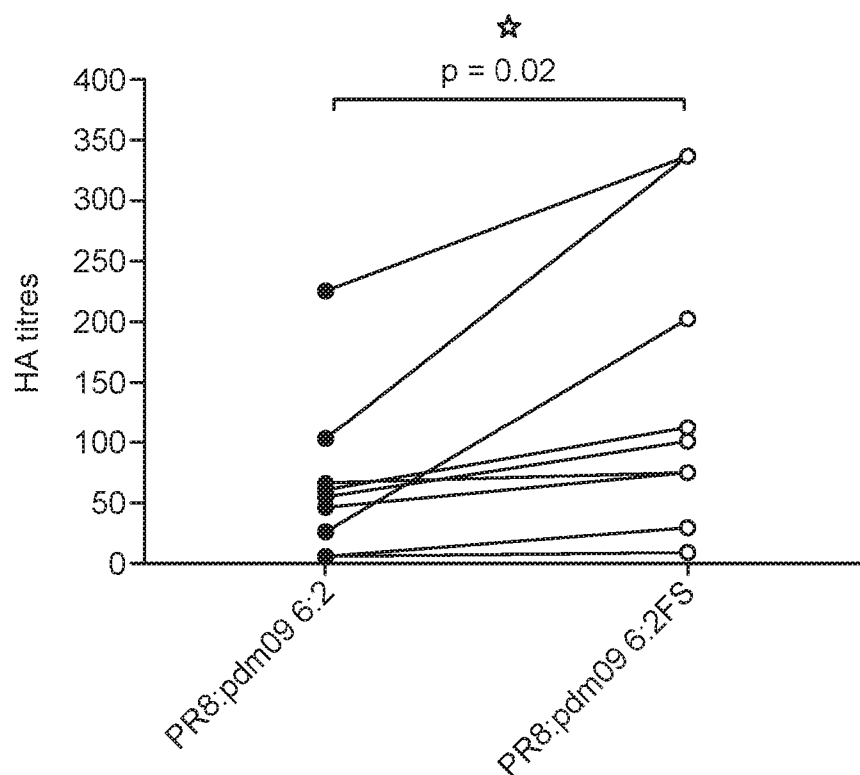
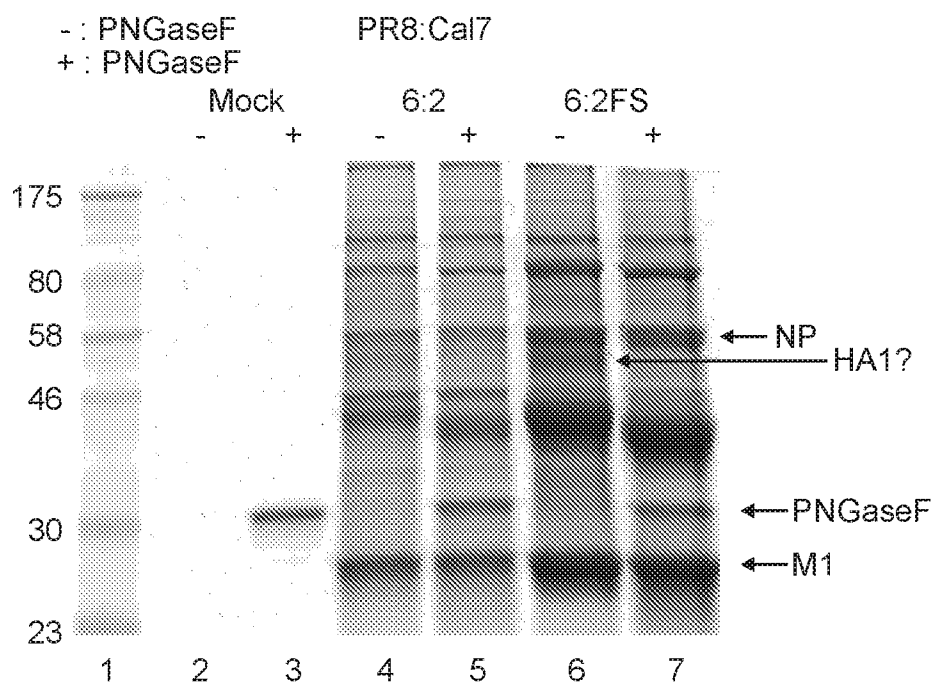


Fig. 5

6a i)



6a ii)

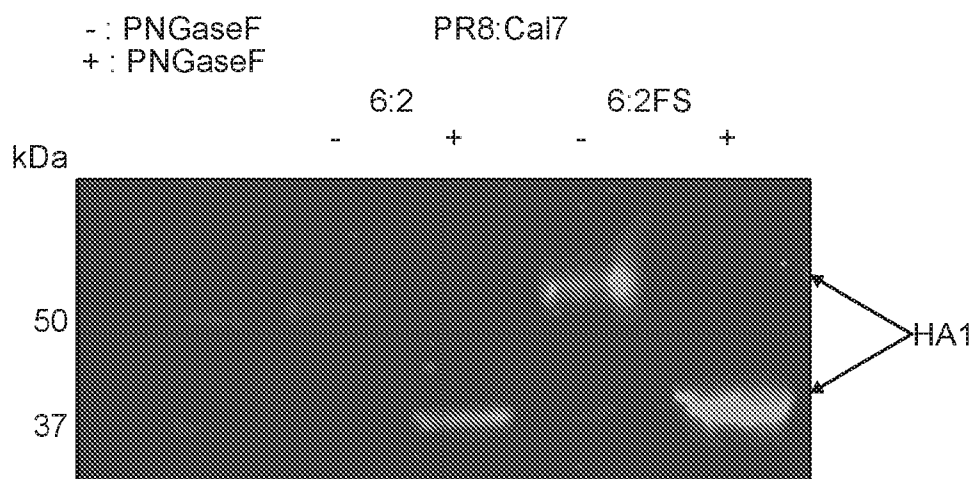
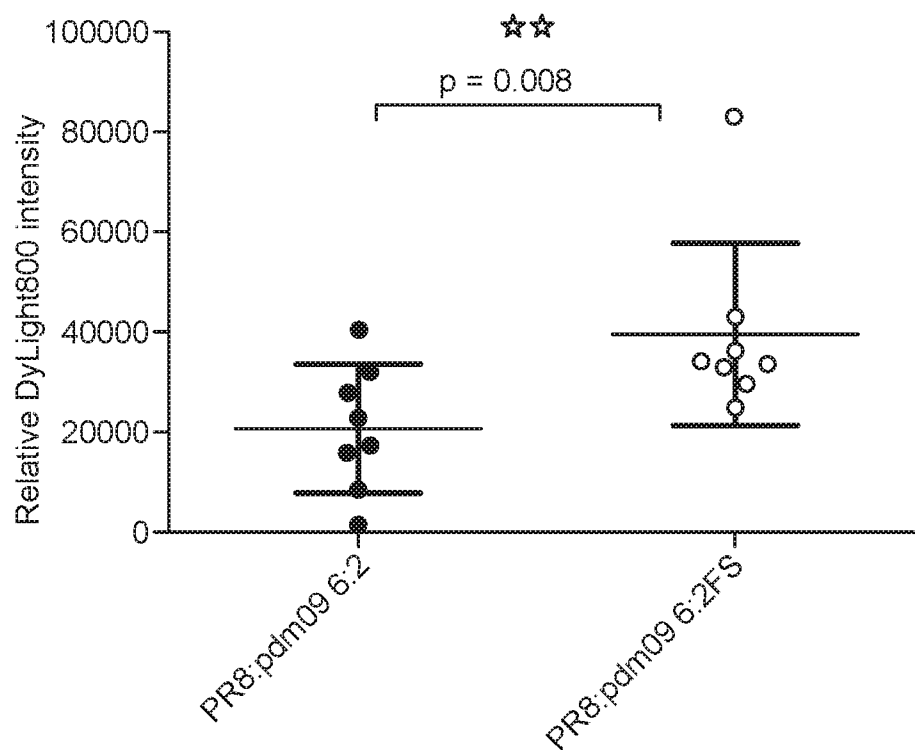


Fig. 6

6b i)



6b ii)

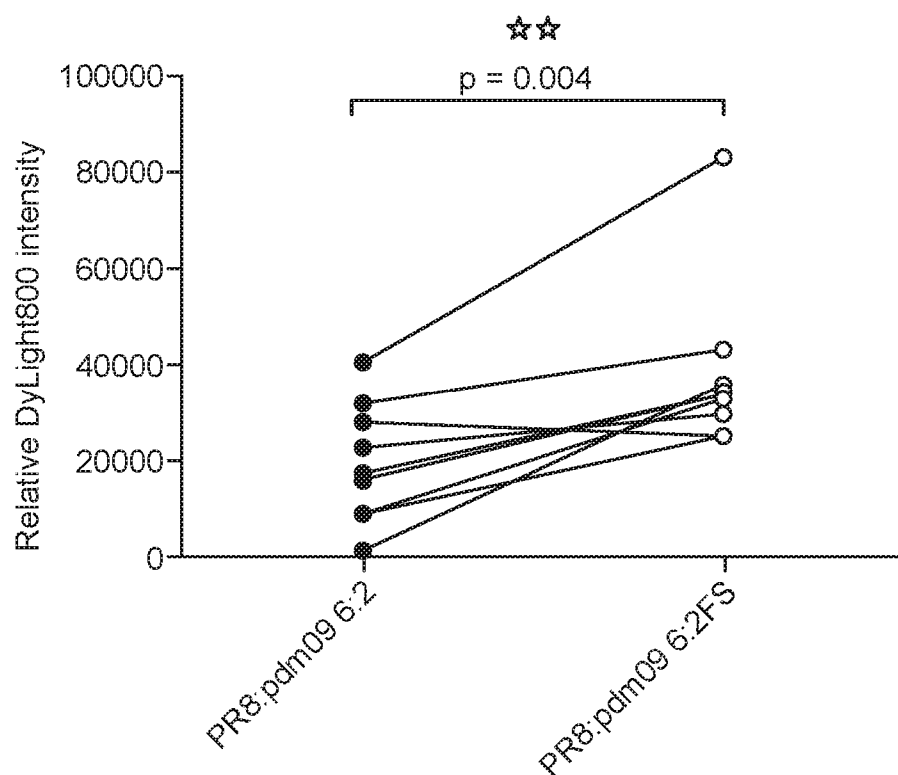


Fig. 6 cont.

Lineage	Subtype	Strain	Fold increase in HA titre ($\pm$ S.D.)	Fold increase in HA1 yield ($\pm$ S.D.)	No. of expts	No. of RG rescues	Small scale	Large scale
Human pdm2009	H1N1	A/California/07/2009	2.65 ± 2.16	$1.9 \pm 1.07^*$	7	2	4	3
Human pdm2009	H1N1	A/England/195/2009	3.79 ± 2.21	2.4 ± 0.37	2	1	1	1
Human pdm1968	H3N2	A/Udorn/307/72	1.26 ± 0.47	1.35 ± 0.36	5	2	3	2
Human pdm1968	H3N2	A/Hong Kong/1/68	1.05 ± 0.06	1.22 ± 0.39	3	1	2	1
Avian	H1N1	A/Mallard/Netherlands/10/99	1.30 ± 0.44	1.25 ± 0.19	4	2	3	1

Fig. 7

Figure 8 – SEQ ID NO:1 (parent PR8 segment 1)

AGCGAAAGCA GGTCAATTAT ATTCAATATG GAAAGAATAA AAGAACTAAG AAATCTAATG
TCGCAGTCTC GCACCCGCGA GATACTCACA AAAACCACCG TGGACCATAT GGCCATAATC
AAGAAGTACA CATCAGGAAG ACAGGAGAAG AACCCAGCAC TTAGGATGAA ATGGATGATG
GCAATGAAAT ATCCAATTAC AGCAGACAAG AGGATAACGG AAATGATTCC TGAGAGAAAT
GAGCAAGGAC AAACCTTTATG GAGTAAAATG AATGATGCCG GATCAGACCG AGTGATGGTA
TCACCACTGG CTGTGACATG GTGGAATAGG AATGGACCAA TAACAAATAC AGTTCATTAT
CCAAAAATCT ACAAACCTTA TTTTGAAAGA GTCGAAAGGC TAAAGCATGG AACCTTTGGC
CCTGTCCATT TTAGAAACCA AGTCAAAATA CGTCGGAGAG TTGACATAAA TCCTGGTCAT
GCAGATCTCA GTGCCAAGGA GGCACAGGAT GTAATCATGG AAGTTGTTTT CCCTAACGAA
GTGGGAGCCA GGATACTAAC ATCGGAATCG CAACTAACGA TAACCAAAGA GAAGAAAGAA
GAACTCCAGG ATTGCAAAAT TTCTCCTTTG ATGGTTGCAT ACATGTTGGA GAGAGAACTG
GTCCGCAAAA CGAGATTCCT CCCAGTGGCT GGTGGAACAA GCAGTGTGTA CATTGAAGTG
TTGCATTTGA CTCAAGGAAC ATGCTGGGAA CAGATGTATA CTCCAGGAGG GGAAGTGAGG
AATGATGATG TTGATCAAAG CTTGATTATT GCTGCTAGGA ACATAGTGAG AAGAGCTGCA
GTATCAGCAG ATCCACTAGC ATCTTTATTG GAGATGTGCC ACAGCACACA GATTGGTGGA
ATTAGGATGG TAGACATCCT TAGGCAGAAC CCAACAGAAG AGCAAGCCGT GGATATATGC
AAGGCTGCAA TGGGACTGAG AATTAGCTCA TCCTTCAGTT TTGGTGGATT CACATTTAAG
AGAACAAGCG GATCATCAGT CAAGAGAGAG GAAGAGGTGC TTACGGGCAA TCTTCAAACA
TTGAAGATAA GAGTGCATGA GGGATATGAA GAGTTCACAA TGGTTGGGAG AAGAGCAACA
GCCATACTCA GAAAAGCAAC CAGGAGATTG ATTCAGCTGA TAGTGAGTGG GAGAGACGAA
CAGTCGATTG CCGAAGCAAT AATTGTGGCC ATGGTATTTT CACAAGAGGA TTGTATGATA
AAAGCAGTCA GAGGTGATCT GAATTCGTC AATAGGGCGA ATCAGCGATT GAATCCTATG
CATCAACTTT TAAGACATTT TCAGAAGGAT GCGAAAGTGC TTTTCAAAA TTGGGGAGTT
GAACCTATCG ACAATGTGAT GGAATGATT GGGATATTGC CCGACATGAC TCCAAGCATC
GAGATGTCAA TGAGAGGAGT GAGAATCAGC AAAATGGGTG TAGATGAGTA CTCCAGCACG
GAGAGGGTAG TGGTGAGCAT TGACCGTTTT TTGAGAATCC GGGACCAACG AGGAAATGTA
CTACTGTCTC CCGAGGAGGT CAGTGAAACA CAGGGAACAG AGAAACTGAC AATAACTTAC

TCATCGTCAA TGATGTGGGA GATTAATGGT CCTGAATCAG TGTTGGTCAA TACCTATCAA
TGGATCATCA GAAACTGGGA AACTGTAAA ATTCAGTGGT CCCAGAACCC TACAATGCTA
TACAATAAAA TGGAATTTGA ACCATTTAG TCTTTAGTAC CTAAGGCCAT TAGAGGCCAA
TACAGTGGGT TTGTAAGAAC TCTGTTCCAA CAAATGAGGG ATGTGCTTGG GACATTTGAT
ACCGCACAGA TAATAAACT TCTTCCCTTC GCAGCCGCTC CACCAAAGCA AAGTAGAATG
CAGTTCTCCT CATTTACTGT GAATGTGAGG GGATCAGGAA TGAGAATACT TGTAAGGGGC
AATTCTCCTG TATTCAACTA TAACAAGGCC ACGAAGAGAC TCACAGTTCT CGGAAAGGAT
GCTGGCACTT TAACTGAAGA CCCAGATGAA GGCACAGCTG GAGTGGAGTC CGCTGTTCTG
AGGGGATTCC TCATTCTGGG CAAAGAGAAC AAGAGATATG GGCCAGCACT AAGCATCAAT
GAACTGAGCA ACCTTGCGAA AGGAGAGAAG GCTAATGTGC TAATTGGGCA AGGAGACGTG
GTGTTGGTAA TGAAACGGAA ACGGGACTCT AGCATACTTA CTGACAGCCA GACAGCGACC
AAAAGAATTC GGATGGCCAT CAATTAGTGT CGAATAGTTT AAAAACGACC TTGTTTCTAC

T

Figure 9 –SEQ ID NO:2 (parent PR8 segment 2)

AGCGAAAGCA GGCAAACCAT TTGAATGGAT GTCAATCCGA CCTACTTTT CTAAAAAGTG
CCAGCACAAA ATGCTATAAG CACAACTTTC CTTATACTG GAGACCCTCC TTACAGCCAT
GGGACAGGAA CAGGATACAC CATGGATACT GTCAACAGGA CACATCAGTA CTCAGAAAAG
GGAAGATGGA CAACAAACAC CGAACTGGA GCACCGCAAC TCAACCCGAT TGATGGGCCA
CTGCCAGAAG ACAATGAACC AAGTGGTTAT GCCCAAACAG ATTGTGTATT GGAGGCGATG
GCTTTCCTTG AGGAATCCCA TCCTGGTATT TTTGAAAAC CGTGTATTGA AACGATGGAG
GTTGTTTCAGC AAACACGAGT AGACAAGCTG ACACAAGGCC GACAGACCTA TGA CTGGACT
CTAAATAGAA ACCAACCTGC TGCAACAGCA TTGGCCAACA CAATAGAAGT GTTCAGATCA
AATGGCCTCA CGGCCAATGA GTCTGGAAGG CTCATAGACT TCCTTAAGGA TGTAATGGAG
TCAATGAACA AAGAAGAAAT GGGGATCACA ACTCATTTTC AGAGAAAGAG ACGGGTGAGA
GACAATATGA CTAAGAAAAT GATAACACAG AGAACAATGG GTAAAAAGAA GCAGAGATTG
AACAAAAGGA GTTATCTAAT TAGAGCATTG ACCCTGAACA CAATGACCAA AGATGCTGAG
AGAGGGAAGC TAAAACGGAG AGCAATTGCA ACCCCAGGGA TGCAAATAAG GGGGTTTGTA
TACTTTGTTG AGACACTGGC AAGGAGTATA TGTGAGAAAC TTGAACAATC AGGGTTGCCA
GTTGGAGGCA ATGAGAAGAA AGCAAAGTTG GCAAATGTTG TAAGGAAGAT GATGACCAAT
TCTCAGGACA CCGAACTTTC TTTCACCATC ACTGGAGATA ACACCAAATG GAACGAAAAT
CAGAATCCTC GGATGTTTTT GGCCATGATC ACATATATGA CCAGAAATCA GCCCGAATGG
TTCAGAAATG TTCTAAGTAT TGCTCCAATA ATGTTCTCAA ACAAATGGC GAGACTGGGA
AAAGGGTATA TGTTTGAGAG CAAGAGTATG AAACCTAGAA CTCAAATACC TGCAGAAATG
CTAGCAAGCA TCGATTTGAA ATATTTCAAT GATTCAACAA GAAAGAAGAT TGAAAAATC
CGATCGCTCT TAATAGAGGG GACTGCATCA TTGAGCCCTG GAATGATGAT GGGCATGTTC
AATATGTAA GCACTGTATT AGGCGTCTCC ATCCTGAATC TTGGACAAAA GAGATACACC
AAGACTACTT ACTGGTGGGA TGGTCTTCAA TCCTCTGACG ATTTTGCTCT GATTGTGAAT
GCACCCAATC ATGAAGGGAT TCAAGCCGGA GTCGACAGGT TTTATCGAAC CTGTAAGCTA
CTTGGAATCA ATATGAGCAA GAAAAAGTCT TACATAAACA GAACAGGTAC ATTTGAATTC
ACAAGTTTTT TCTATCGTTA TGGGTTTGTT GCCAATTTCA GCATGGAGCT TCCCAGTTTT
GGGGTGTCTG GGATCAACGA GTCAGCGGAC ATGAGTATTG GAGTTACTGT CATCAAAAAC

AATATGATAA ACAATGATCT TGGTCCAGCA ACAGCTCAAA TGGCCCTTCA GTTGTTTCATC
AAAGATTACA GGTACACGTA CCGATGCCAT AGAGGTGACA CACAAATACA AACCCGAAGA
TCATTTGAAA TAAAGAAACT GTGGGAGCAA ACCCGTTCCA AAGCTGGACT GCTGGTCTCC
GACGGAGGCC CAAATTTATA CAACATTAGA AATCTCCACA TTCCTGAAGT CTGCCTAAAA
TGGAATTGA TGGATGAGGA TTACCAGGGG CGTTTATGCA ACCCACTGAA CCCATTTGTC
AGCCATAAAG AAATTGAATC AATGAACAAT GCAGTGATGA TGCCAGCACA TGGTCCAGCC
AAAAACATGG AGTATGATGC TGTTGCAACA ACACACTCCT GGATCCCCAA AAGAAATCGA
TCCATCTTGA ATACAAGTCA AAGAGGAGTA CTTGAGGATG AACAAATGTA CCAAAGGTGC
TGCAATTTAT TTGAAAAATT CTTCCCAGC AGTTCATACA GAAGACCAGT CGGGATATCC
AGTATGGTGG AGGCTATGGT TTCCAGAGCC CGAATTGATG CACGGATTGA TTTCGAATCT
GGAAGGATAA AGAAAGAAGA GTTCACTGAG ATCATGAAGA TCTGTTCCAC CATTGAAGAG
CTCAGACGGC AAAAATAGTG AATTTAGCTT GTCCTTCATG AAAAAATGCC TTGTTTCTAC

T

Figure 10 – SEQ ID NO:3 (parent PR8 segment 3)

AGCGAAAGCA GGTACTGATC CAAAATGGAA GATTTTGTGC GACAATGCTT CAATCCGATG
ATTGTCGAGC TTGCGGAAAA AACAATGAAA GAGTATGGGG AGGACCTGAA AATCGAAACA
AACAAATTTG CAGCAATATG CACTCACTTG GAAGTATGCT TCATGTATTC AGATTTTCAC
TTCATCAATG AGCAAGGCGA GTCAATAATC GTAGAACTTG GTGATCCAAA TGCACTTTTG
AAGCACAGAT TTGAAATAAT CGAGGGAAGA GATCGCACAA TGGCCTGGAC AGTAGTAAAC
AGTATTTGCA AACTACAGG GGCTGAGAAA CCAAAGTTTC TACCAGATTT GTATGATTAC
AAGGAGAATA GATTCATCGA AATTGGAGTA ACAAGGAGAG AAGTTCACAT ATACTATCTG
GAAAAGGCCA ATAAATTAAT ATCTGAGAAA ACACACATCC ACATTTTCTC GTTCACTGGG
GAAGAAATGG CCACAAAGGC AGACTACACT CTCGATGAAG AAAGCAGGGC TAGGATCAAA
ACCAGACTAT TCACCATAAG ACAAGAAATG GCCAGCAGAG GCCTCTGGGA TTCCTTTCGT
CAGTCCGAGA GAGGAGAAGA GACAATTGAA GAAAGGTTTG AAATCACAGG AACAATGCGC
AAGCTTGCCG ACCAAAGTCT CCCGCCGAAC TTCTCCAGCC TTGAAAATTT TAGAGCCTAT
GTGGATGGAT TCGAACCGAA CGGCTACATT GAGGGCAAGC TGTCTCAAAT GTCCAAAGAA
GTAAATGCTA GAATTGAACC TTTTTTGAAA ACAACACCAC GACCACTTAG ACTTCCGAAT
GGGCTCCCT GTTCTCAGCG GTCCAAATTC CTACTGATGG ATGCCTTAAA ATTAAGCATT
GAGGACCCAA GTCATGAAGG AGAGGGAATA CCGCTATATG ATGCAATCAA ATGCATGAGA
ACATTCTTTG GATGGAAGGA ACCCAATGTT GTTAAACCAC ACGAAAAGGG AATAAATCCA
AATTATCTTC TGTCATGGAA GCAAGTACTG GCAGAACTGC AGGACATTGA GAATGAGGAG
AAAATTCCAA AGACTAAAAA TATGAAGAAA ACAAGTCAGC TAAAGTGGGC ACTTGGTGAG
AACATGGCAC CAGAAAAGGT AGACTTTGAC GACTGTAAAG ATGTAGGTGA TTTGAAGCAA
TATGATAGTG ATGAACCAGA ATTGAGGTCG CTGCAAGTT GGATTCAGAA TGAGTTTAAC
AAGGCATGCG AACTGACAGA TTCAAGCTGG ATAGAGCTCG ATGAGATTGG AGAAGATGTG
GCTCCAATTG AACACATTGC AAGCATGAGA AGGAATTATT TCACATCAGA GGTGTCTCAC
TGCAGAGCCA CAGAATACAT AATGAAGGGG GTGTACATCA ATACTGCCTT GCTTAATGCA
TCTTGTGCAG CAATGGATGA TTTCCAATTA ATTCCAATGA TAAGCAAGTG TAGAACTAAG
GAGGGAAGGC GAAAGACCAA CTTGTATGGT TTCATCATAA AAGGAAGATC CCACTTAAGG
AATGACACCG ACGTGGTAAA CTTTGTGAGC ATGGAGTTTT CTCTCACTGA CCCAAGACTT

GAACCACATA AATGGGAGAA GTACTGTGTT CTTGAGATAG GAGATATGCT TATAAGAAGT
GCCATAGGCC AGGTTTCAAG GCCCATGTTC TTGTATGTGA GAACAAATGG AACCTCAAAA
ATTAAAATGA AATGGGGAAT GGAGATGAGG CGTTGCCTCC TCCAGTCACT TCAACAAATT
GAGAGTATGA TTGAAGCTGA GTCCTCTGTC AAAGAGAAAG ACATGACCAA AGAGTTCTTT
GAGAACAAAT CAGAAACATG GCCCATTGGA GAGTCCCCCA AAGGAGTGGA GGAAAGTTCC
ATTGGGAAGG TCTGCAGGAC TTTATTAGCA AAGTCGGTAT TCAACAGCTT GTATGCATCT
CCACAAC TAG AAGGATTTTC AGCTGAATCA AGAAAAC TGC TTCTTATCGT TCAGGCTCTT
AGGGACAACC TTGAACCTGG GACCTTTGAT CTTGGGGGGC TATATGAAGC AATTGAGGAG
TGCCTGATTA ATGATCCCTG GGTTTTGCTT AATGCTTCTT GGTTC AACTC ATTCCTTACA
CATGCATTGA GTTAGTTGTG GCAGTGCTAC TATTTGCTAT CCATACTGTC CAAAAAAGTA
CCTTGTTTCT ACT

Figure 11 – SEQ ID NO:4 (parent PR8 segment 5)

AGCAAAAGCA GGGTAGATAA TCACTCACTG AGTGACATCA AAATCATGGC GTCTCAAGGC
ACCAAACGAT CTTACGAACA GATGGAGACT GATGGAGAAC GCCAGAATGC CACTGAAATC
AGAGCATCCG TCGGAAAAAT GATTGGTGGA ATTGGACGAT TCTACATCCA AATGTGCACC
GAACTCAAAC TCAGTGATTA TGAGGGACGG TTGATCCAAA ACAGCTTAAC AATAGAGAGA
ATGGTGCTCT CTGCTTTTGA CGAAAGGAGA AATAAATACC TTGAAGAACA TCCCAGTGCG
GGGAAAGATC CTAAGAAAAC TGGAGGACCT ATATACAGGA GAGTAAACGG AAAGTGGATG
AGAGAACTCA TCCTTTATGA CAAAGAAGAA ATAAGGCGAA TCTGGCGCCA AGCTAATAAT
GGTGACGATG CAACGGCTGG TCTGACTCAC ATGATGATCT GGCATTCCAA TTTGAATGAT
GCAACTTATC AGAGGACAAG AGCTCTTGTT CGCACCGGAA TGGATCCCAG GATGTGCTCT
CTGATGCAAG GTTCAACTCT CCCTAGGAGG TCTGGAGCCG CAGGTGCTGC AGTCAAAGGA
GTTGGAACAA TGGTGATGGA ATTGGTCAGA ATGATCAAAC GTGGGATCAA TGATCGGAAC
TTCTGGAGGG GTGAGAATGG ACGAAAAACA AGAATTGCTT ATGAAAGAAT GTGCAACATT
CTCAAAGGGA AATTTCAAAC TGCTGCACAA AAAGCAATGA TGGATCAAGT GAGAGAGAGC
CGGAACCCAG GGAATGCTGA GTTCGAAGAT CTCACTTTTC TAGCACGGTC TGCACTCATA
TTGAGAGGGT CGGTTGCTCA CAAGTCCTGC CTGCCTGCCT GTGTGTATGG ACCTGCCGTA
GCCAGTGGGT ACGACTTTGA AAGGGAGGGA TACTCTCTAG TCGGAATAGA CCCTTTCAGA
CTGCTTCAAA ACAGCCAAGT GTACAGCCTA ATCAGACCAA ATGAGAATCC AGCACACAAG
AGTCAACTGG TGTGGATGGC ATGCCATTCT GCCGCATTG AAGATCTAAG AGTATTAAGC
TTCATCAAAG GGACGAAGGT GCTCCCAAGA GGAAGCTTT CCACTAGAGG AGTTCAAATT
GCTTCCAATG AAAATATGGA GACTATGGAA TCAAGTACAC TTGAACTGAG AAGCAGGTAC
TGGGCCATAA GGACCAGAAG TGGAGGAAAC ACCAATCAAC AGAGGGCATC TCGGGGCCAA
ATCAGCATAC AACCTACGTT CTCAGTACAG AGAAATCTCC CTTTGTGACAG AACAACCATT
ATGGCAGCAT TCAATGGGAA TACAGAGGGG AGAACATCTG ACATGAGGAC CGAAATCATA
AGGATGATGG AAAGTGCAAG ACCAGAAGAT GTGTCTTTCC AGGGGCGGGG AGTCTTCGAG
CTCTCGGACG AAAAGGCAGC GAGCCCGATC GTGCCTTCCT TTGACATGAG TAATGAAGGA
TCTTATTCT TCGGAGACAA TGCAGAGGAG TACGACAATT AAAGAAAAAT ACCCTTGTTT
CTACT

Figure 12 – SEQ ID NO:5 (parent PR8 segment 7)

AGCGAAAGCA GGTAGATATT GAAAGATGAG TCTTCTAACC GAGGTCGAAA CGTACGTACT
CTCTATCATC CCGTCAGGCC CCCTCAAAGC CGAGATCGCA CAGAGACTTG AAGATGTCTT
TGCAGGGAAG AACACCGATC TTGAGGTTCT CATGGAATGG CTAAAGACAA GACCAATCCT
GTCACCTCTG ACTAAGGGGA TTTTAGGATT TGTGTTACG CTCACCGTGC CCAGTGAGCG
AGGACTGCAG CGTAGACGCT TTGTCCAAAA TGCCCTTAAT GGGAACGGGG ATCCAAATAA
CATGGACAAA GCAGTTAAAC TGTATAGGAA GCTCAAGAGG GAGATAACAT TCCATGGGGC
CAAAGAAATC TCACTCAGTT ATTCTGCTGG TGCACTTGCC AGTTGTATGG GCCTCATATA
CAACAGGATG GGGGCTGTGA CCACTGAAGT GGCATTTGGC CTGGTATGTG CAACCTGTGA
ACAGATTGCT GACTCCCAGC ATCGGTCTCA TAGGCAAATG GTGACAACAA CCAATCCACT
AATCAGACAT GAGAACAGAA TGGTTTTAGC CAGCACTACA GCTAAGGCTA TGGAGCAAAT
GGCTGGATCG AGTGAGCAAG CAGCAGAGGC CATGGAGGTT GCTAGTCAGG CTAGACAAAT
GGTGCAAGCG ATGAGAACCA TTGGAATCA TCCTAGCTCC AGTGCTGGTC TGAAAAATGA
TCTTCTTGAA AATTTGCAGG CCTATCAGAA ACGAATGGGG GTGCAGATGC AACGGTTCAA
GTGATCCTCT CACTATTGCC GCAAATATCA TTGGGATCTT GCACTTGACA TTGTGGATTG
TTGATCGTCT TTTTTTCAA TGCATTTACC GTCGCTTTAA ATACGGACTG AAAGGAGGGC
CTTCTACGGA AGGAGTGCCA AAGTCTATGA GGGAAGAATA TCGAAAGGAA CAGCAGAGTG
CTGTGGATGC TGACGATGGT CATTTTGTCA GCATAGAGCT GGAGTAAAAA ACTACCTTGT
TTCTACT

Figure 13 – SEQ ID NO: 6 (parent PR8 segment 8)

AGCAAAAGCA GGGTGACAAA AACATAATGG ATCCAAACAC TGTGTCAAGC TTTCAGGTAG
ATTGCTTTCT TTGGCATGTC CGCAAACGAG TTGCAGACCA AGAACTAGGT GATGCCCCAT
TCCTTGATCG GCTTCGCCGA GATCAGAAAT CCCTAAGAGG AAGGGGCAGT ACTCTCGGTC
TGGACATCAA GACAGCCACA CGTGCTGGAA AGCAGATAGT GGAGCGGATT CTGAAAGAAG
AATCCGATGA GGCACCTAAA ATGACCATGG CCTCTGTACC TCGTCTGCGT TACCTAACTG
ACATGACTCT TGAGGAAATG TCAAGGGACT GGTCCATGCT CATACCCAAG CAGAAAGTGG
CAGGCCCTCT TTGTATCAGA ATGGACCAGG CGATCATGGA TAAGAACATC ATACTGAAAG
CGAACTTCAG TGTGATTTTT GACCGGCTGG AGACTCTAAT ATTGCTAAGG GCTTTCACCG
AAGAGGGAGC AATTGTTGGC GAAATTTAC CATTGCCTTC TCTCCAGGA CATACTGCTG
AGGATGTCAA AAATGCAGTT GGAGTCCTCA TCGGAGGACT TGAATGGAAT GATAACACAG
TTCGAGTCTC TGAAACTCTA CAGAGATTCG CTTGGAGAAG CAGTAATGAG AATGGGAGAC
CTCCACTCAC TCCAAAACAG AAACGAGAAA TGGCGGGAAC AATTAGGTCA GAAGTTTGAA
GAAATAAGAT GGTGATTGA AGAAGTGAGA CACAACTGA AGATAACAGA GAATAGTTTT
GAGCAAATAA CATTTATGCA AGCCTTACAT CTATTGCTTG AAGTGGAGCA AGAGATAAGA
ACTTTCTCGT TTCAGCTTAT TTAGTACTAA AAAACACCCT TGTTTCTACT

Figure 14 – SEQ ID NO: 7 (PR8 segment 3 PTC mutant 1)

AGCGAAAGCA GGTACTGATC CAAAATGGAA GATTTTGTGC GACAATGCTT CAATCCGATG
ATTGTCGAGC TTGCGGAAAA AACAATGAAA GAGTATGGGG AGGACCTGAA AATCGAAACA
AACAAATTTG CAGCAATATG CACTCACTTG GAAGTATGCT TCATGTATTC AGATTTTCAC
TTCATCAATG AGCAAGGCGA GTCAATAATC GTAGAACTTG GTGATCCAAA TGCACTTTTG
AAGCACAGAT TTGAAATAAT CGAGGGAAGA GATCGCACAA TGGCCTGGAC AGTAGTAAAC
AGTATTTGCA AACTACAGG GGCTGAGAAA CCAAAGTTTC TACCAGATTT GTATGATTAC
AAGGAGAATA GATTCATCGA AATTGGAGTA ACAAGGAGAG AAGTTCACAT ATACTATCTG
GAAAAGGCCA ATAAAATTAA ATCTGAGAAA ACACACATCC ACATTTTCTC GTTCACTGGG
GAAGAAATGG CCACAAAGGC AGACTACACT CTCGATGAAG AAAGCAGGGC TAGGATCAAA
ACCAGACTAT TCACCATAAG ACAAGAAATG GCCAGCAGAG GCCTCTGGGA TTCCTTTCGT
CAGTCCGAGA GAGGAGAAGA GACAATTGAA GAAAGGTTTG AAATAACAGG AACAATGCGC
AAGCTTGCCG ACCAAAGTCT CCCGCCGAAC TTCTCCAGCC TTGAAAATTT TAGAGCCTAT
GTGGATGGAT TCGAACCGAA CGGCTACATT GAGGGCAAGC TGTCTCAAAT GTCCAAAGAA
GTAAATGCTA GAATTGAACC TTTTTTGAAA ACAACACCAC GACCACTTAG ACTTCCGAAT
GGGCCTCCCT GTTCTCAGCG GTCCAAATTC CTA CTGATGG ATGCCTTAAA ATTAAGCATT
GAGGACCCAA GTCATGAAGG AGAGGGAATA CCGCTATATG ATGCAATCAA ATGCATGAGA
ACATTCTTTG GATGGAAGGA ACCCAATGTT GTTAAACCAC ACGAAAAGGG AATAAATCCA
AATTATCTTC TGTCATGGAA GCAAGTACTG GCAGAACTGC AGGACATTGA GAATGAGGAG
AAAATTCCAAGACTAAAAA TATGAAGAAA ACAAGTCAGC TAAAGTGGGC ACTTGGTGAG
AACATGGCAC CAGAAAAGGT AGACTTTGAC GACTGTAAAG ATGTAGGTGA TTTGAAGCAA
TATGATAGTG ATGAACCAGA ATTGAGGTCG CTTGCAAGTT GGATTCAGAA TGAGTTTAAC
AAGGCATGCG AACTGACAGA TTCAAGCTGG ATAGAGCTCG ATGAGATTGG AGAAGATGTG
GCTCCAATTG AACACATTGC AAGCATGAGA AGGAATTATT TCACATCAGA GGTGTCTCAC
TGCAGAGCCA CAGAATACAT AATGAAGGGG GTGTACATCA ATACTGCCTT GCTTAATGCA
TCTTGTGCAG CAATGGATGA TTTCCAATTA ATTCCAATGA TAAGCAAGTG TAGAACTAAG
GAGGGAAGGC GAAAGACCAA CTTGTATGGT TTCATCATAA AAGGAAGATC CCACTTAAGG

AATGACACCG ACGTGGTAAA CTTTGTGAGC ATGGAGTTTT CTCTCACTGA CCCAAGACTT
GAACCACATA AATGGGAGAA GTACTGTGTT CTTGAGATAG GAGATATGCT TATAAGAAGT
GCCATAGGCC AGGTTTCAAG GCCCATGTTC TTGTATGTGA GAACAAATGG AACCTCAAAA
ATTAAAATGA AATGGGGAAT GGAGATGAGG CGTTGCCTCC TCCAGTCACT TCAACAAATT
GAGAGTATGA TTGAAGCTGA GTCCTCTGTC AAAGAGAAAG ACATGACCAA AGAGTTCTTT
GAGAACAAAT CAGAAACATG GCCCATTGGA GAGTCCCCCA AAGGAGTGGA GGAAAGTTCC
ATTGGGAAGG TCTGCAGGAC TTTATTAGCA AAGTCGGTAT TCAACAGCTT GTATGCATCT
CCACAAC TAG AAGGATTTTC AGCTGAATCA AGAAAAC TGC TTCTTATCGT TCAGGCTCTT
AGGGACAACC TTGAACCTGG GACCTTTGAT CTTGGGGGGC TATATGAAGC AATTGAGGAG
TGCCTGATTA ATGATCCCTG GGTTTTGCTT AATGCTTCTT GGTTC AACTC ATTCCTTACA
CATGCATTGA GTTAGTTGTG GCAGTGCTAC TATTTGCTAT CCATACTGTC CAAAAAAGTA
CCTTGTTTCT ACT

Figure 15 – SEQ ID NO:8 (PR8 segment 3 PTC mutant 2)

AGCGAAAGCA GGTACTGATC CAAAATGGAA GATTTTGTGC GACAATGCTT CAATCCGATG
ATTGTCGAGC TTGCGGAAAA AACAATGAAA GAGTATGGGG AGGACCTGAA AATCGAAACA
AACAAATTTG CAGCAATATG CACTCACTTG GAAGTATGCT TCATGTATTC AGATTTTCAC
TTCATCAATG AGCAAGGCGA GTCAATAATC GTAGAACTTG GTGATCCAAA TGCACTTTTG
AAGCACAGAT TTGAAATAAT CGAGGGAAGA GATCGCACAA TGGCCTGGAC AGTAGTAAAC
AGTATTTGCA AACTACAGG GGCTGAGAAA CCAAAGTTTC TACCAGATTT GTATGATTAC
AAGGAGAATA GATTCATCGA AATTGGAGTA ACAAGGAGAG AAGTTCACAT ATACTATCTG
GAAAAGGCCA ATAAAATTAA ATCTGAGAAA ACACACATCC ACATTTTCTC GTTCACTGGG
GAAGAAATGG CCACAAAGGC AGACTACACT CTCGATGAAG AAAGCAGGGC TAGGATCAAA
ACCAGACTAT TCACCATAAG ACAAGAAATG GCCAGCAGAG GCCTCTGGGA TTCCTTTCGT
CAGTCCGAGA GAGGAGAAGA GACAATTGAA GAAAGGTTTG AAATCACAGG AACAATGAGG
AAGCTAGCCG ACCAAAGTCT CCCGCCGAAC TTCTCCAGCC TTGAAAATTT TAGAGCCTAT
GTGGATGGAT TCGAACCGAA CGGCTACATT GAGGGCAAGC TGTCTCAAAT GTCCAAAGAA
GTAAATGCTA GAATTGAACC TTTTTTGAAA ACAACACCAC GACCACTTAG ACTTCCGAAT
GGGCCTCCCT GTTCTCAGCG GTCCAAATTC CTA CTGATGG ATGCCTTAAA ATTAAGCATT
GAGGACCCAA GTCATGAAGG AGAGGGAATA CCGCTATATG ATGCAATCAA ATGCATGAGA
ACATTCTTTG GATGGAAGGA ACCCAATGTT GTTAAACCAC ACGAAAAGGG AATAAATCCA
AATTATCTTC TGTCATGGAA GCAAGTACTG GCAGAACTGC AGGACATTGA GAATGAGGAG
AAAATTCCAA AGACTAAAAA TATGAAGAAA ACAAGTCAGC TAAAGTGGGC ACTTGGTGAG
AACATGGCAC CAGAAAAGGT AGACTTTGAC GACTGTAAAG ATGTAGGTGA TTTGAAGCAA
TATGATAGTG ATGAACCAGA ATTGAGGTCG CTTGCAAGTT GGATTCAGAA TGAGTTTAAC
AAGGCATGCG AACTGACAGA TTCAAGCTGG ATAGAGCTCG ATGAGATTGG AGAAGATGTG
GCTCCAATTG AACACATTGC AAGCATGAGA AGGAATTATT TCACATCAGA GGTGTCTCAC
TGCAGAGCCA CAGAATACAT AATGAAGGGG GTGTACATCA ATACTGCCTT GCTTAATGCA
TCTTGTGCAG CAATGGATGA TTTCCAATTA ATTCCAATGA TAAGCAAGTG TAGAACTAAG
GAGGGAAGGC GAAAGACCAA CTTGTATGGT TTCATCATAA AAGGAAGATC CCACTTAAGG

AATGACACCG ACGTGGTAAA CTTTGTGAGC ATGGAGTTTT CTCTCACTGA CCCAAGACTT
GAACCACATA AATGGGAGAA GTACTGTGTT CTTGAGATAG GAGATATGCT TATAAGAAGT
GCCATAGGCC AGGTTTCAAG GCCCATGTTC TTGTATGTGA GAACAAATGG AACCTCAAAA
ATTAAAATGA AATGGGGAAT GGAGATGAGG CGTTGCCTCC TCCAGTCACT TCAACAAATT
GAGAGTATGA TTGAAGCTGA GTCCTCTGTC AAAGAGAAAG ACATGACCAA AGAGTTCTTT
GAGAACAAAT CAGAAACATG GCCCATTGGA GAGTCCCCCA AAGGAGTGGA GGAAAGTTCC
ATTGGGAAGG TCTGCAGGAC TTTATTAGCA AAGTCGGTAT TCAACAGCTT GTATGCATCT
CCACAAC TAG AAGGATTTTC AGCTGAATCA AGAAAAC TGC TTCTTATCGT TCAGGCTCTT
AGGGACAACC TTGAACCTGG GACCTTTGAT CTTGGGGGGC TATATGAAGC AATTGAGGAG
TGCCTGATTA ATGATCCCTG GGTTTTGCTT AATGCTTCTT GGTTC AACTC ATTCCTTACA
CATGCATTGA GTTAGTTGTG GCAGTGCTAC TATTTGCTAT CCATACTGTC CAAAAAAGTA
CCTTGTTTCT ACT

Figure 16 – SEQ ID NO:9 (PR8 segment 3 PTC mutant 3)

AGCGAAAGCA GGTACTGATC CAAAATGGAA GATTTTGTGC GACAATGCTT CAATCCGATG
ATTGTCGAGC TTGCGGAAAA AACAATGAAA GAGTATGGGG AGGACCTGAA AATCGAAACA
AACAAATTTG CAGCAATATG CACTCACTTG GAAGTATGCT TCATGTATTC AGATTTTCAC
TTCATCAATG AGCAAGGCGA GTCAATAATC GTAGAACTTG GTGATCCAAA TGCACTTTTG
AAGCACAGAT TTGAAATAAT CGAGGGAAGA GATCGCACAA TGGCCTGGAC AGTAGTAAAC
AGTATTTGCA AACTACAGG GGCTGAGAAA CCAAAGTTTC TACCAGATTT GTATGATTAC
AAGGAGAATA GATTCATCGA AATTGGAGTA ACAAGGAGAG AAGTTCACAT ATACTATCTG
GAAAAGGCCA ATAAAATTAA ATCTGAGAAA ACACACATCC ACATTTTCTC GTTCACTGGG
GAAGAAATGG CCACAAAGGC AGACTACACT CTCGATGAAG AAAGCAGGGC TAGGATCAAA
ACCAGACTAT TCACCATAAG ACAAGAAATG GCCAGCAGAG GCCTCTGGGA TTCCTTTCGT
CAGTCCGAGA GAGGAGAAGA GACAATTGAA GAAAGGTTTG AAATCACAGG AACAATGCGC
AAGCTTGCCG ACCAAAGTCT CCCGCCGAAC TTCTCCAGCC TAGAAAATTT TAGAGCCTAT
GTGGATGGAT TCGAACCGAA CGGCTACATT GAGGGCAAGC TGTCTCAAAT GTCCAAAGAA
GTAAATGCTA GAATTGAACC TTTTTTGAAA ACAACACCAC GACCACTTAG ACTTCCGAAT
GGGCCTCCCT GTTCTCAGCG GTCCAAATTC CTA CTGATGG ATGCCTTAAA ATTAAGCATT
GAGGACCCAA GTCATGAAGG AGAGGGAATA CCGCTATATG ATGCAATCAA ATGCATGAGA
ACATTCTTTG GATGGAAGGA ACCCAATGTT GTTAAACCAC ACGAAAAGGG AATAAATCCA
AATTATCTTC TGTCATGGAA GCAAGTACTG GCAGAACTGC AGGACATTGA GAATGAGGAG
AAAATTCCAA AGACTAAAAA TATGAAGAAA ACAAGTCAGC TAAAGTGGGC ACTTGGTGAG
AACATGGCAC CAGAAAAGGT AGACTTTGAC GACTGTAAAG ATGTAGGTGA TTTGAAGCAA
TATGATAGTG ATGAACCAGA ATTGAGGTCG CTTGCAAGTT GGATTCAGAA TGAGTTTAAC
AAGGCATGCG AACTGACAGA TTCAAGCTGG ATAGAGCTCG ATGAGATTGG AGAAGATGTG
GCTCCAATTG AACACATTGC AAGCATGAGA AGGAATTATT TCACATCAGA GGTGTCTCAC
TGCAGAGCCA CAGAATACAT AATGAAGGGG GTGTACATCA ATACTGCCTT GCTTAATGCA
TCTTGTGCAG CAATGGATGA TTTCCAATTA ATTCCAATGA TAAGCAAGTG TAGAACTAAG
GAGGGAAGGC GAAAGACCAA CTTGTATGGT TTCATCATAA AAGGAAGATC CCACTTAAGG

AATGACACCG ACGTGGTAAA CTTTGTGAGC ATGGAGTTTT CTCTCACTGA CCCAAGACTT
GAACCACATA AATGGGAGAA GTACTGTGTT CTTGAGATAG GAGATATGCT TATAAGAAGT
GCCATAGGCC AGGTTTCAAG GCCCATGTTC TTGTATGTGA GAACAAATGG AACCTCAAAA
ATTAAAATGA AATGGGGAAT GGAGATGAGG CGTTGCCTCC TCCAGTCACT TCAACAAATT
GAGAGTATGA TTGAAGCTGA GTCCTCTGTC AAAGAGAAAG ACATGACCAA AGAGTTCTTT
GAGAACAAAT CAGAAACATG GCCCATTGGA GAGTCCCCCA AAGGAGTGGA GGAAAGTTCC
ATTGGGAAGG TCTGCAGGAC TTTATTAGCA AAGTCGGTAT TCAACAGCTT GTATGCATCT
CCACAAC TAG AAGGATTTTC AGCTGAATCA AGAAAAC TGC TTCTTATCGT TCAGGCTCTT
AGGGACAACC TTGAACCTGG GACCTTTGAT CTTGGGGGGC TATATGAAGC AATTGAGGAG
TGCCTGATTA ATGATCCCTG GGTTTTGCTT AATGCTTCTT GGTTC AACTC ATTCCTTACA
CATGCATTGA GTTAGTTGTG GCAGTGCTAC TATTTGCTAT CCATACTGTC CAAAAAAGTA
CCTTGTTTCT ACT

Figure 17 - SEQ ID NO:10 (PR8 segment 3 PTC mutant 4)

AGCGAAAGCA GGTACTGATC CAAAATGGAA GATTTTGTGC GACAATGCTT CAATCCGATG
ATTGTCGAGC TTGCGGAAAA AACAATGAAA GAGTATGGGG AGGACCTGAA AATCGAAACA
AACAAATTTG CAGCAATATG CACTCACTTG GAAGTATGCT TCATGTATTC AGATTTTCAC
TTCATCAATG AGCAAGGCGA GTCAATAATC GTAGAACTTG GTGATCCAAA TGCACTTTTG
AAGCACAGAT TTGAAATAAT CGAGGGAAGA GATCGCACAA TGGCCTGGAC AGTAGTAAAC
AGTATTTGCA AACTACAGG GGCTGAGAAA CCAAAGTTTC TACCAGATTT GTATGATTAC
AAGGAGAATA GATTCATCGA AATTGGAGTA ACAAGGAGAG AAGTTCACAT ATACTATCTG
GAAAAGGCCA ATAAAATTAA ATCTGAGAAA ACACACATCC ACATTTTCTC GTTCACTGGG
GAAGAAATGG CCACAAAGGC AGACTACACT CTCGATGAAG AAAGCAGGGC TAGGATCAAA
ACCAGACTAT TCACCATAAG ACAAGAAATG GCCAGCAGAG GCCTCTGGGA TTCCTTTCGT
CAGTCCGAGA GAGGAGAAGA GACAATTGAA GAAAGGTTTG AAATCACAGG AACAATGCGC
AAGCTTGCCG ACCAAAGTCT CCCGCCGAAC TTCTCCAGCC TTGAAAATTT TAGAGCCTAT
GTAGATGGAT TCGAACCGAA CGGCTACATT GAGGGCAAGC TGTCTCAAAT GTCCAAAGAA
GTAAATGCTA GAATTGAACC TTTTTTGAAA ACAACACCAC GACCACTTAG ACTTCCGAAT
GGGCCTCCCT GTTCTCAGCG GTCCAAATTC CTA CTGATGG ATGCCTTAAA ATTAAGCATT
GAGGACCCAA GTCATGAAGG AGAGGGAATA CCGCTATATG ATGCAATCAA ATGCATGAGA
ACATTCTTTG GATGGAAGGA ACCCAATGTT GTTAAACCAC ACGAAAAGGG AATAAATCCA
AATTATCTTC TGTCATGGAA GCAAGTACTG GCAGAACTGC AGGACATTGA GAATGAGGAG
AAAATTCCAAGACTAAAA TATGAAGAAA ACAAGTCAGC TAAAGTGGGC ACTTGGTGAG
AACATGGCAC CAGAAAAGGT AGACTTTGAC GACTGTAAAG ATGTAGGTGA TTTGAAGCAA
TATGATAGTG ATGAACCAGA ATTGAGGTCG CTTGCAAGTT GGATTCAGAA TGAGTTTAAC
AAGGCATGCG AACTGACAGA TTCAAGCTGG ATAGAGCTCG ATGAGATTGG AGAAGATGTG
GCTCCAATTG AACACATTGC AAGCATGAGA AGGAATTATT TCACATCAGA GGTGTCTCAC
TGCAGAGCCA CAGAATACAT AATGAAGGGG GTGTACATCA ATACTGCCTT GCTTAATGCA
TCTTGTGCAG CAATGGATGA TTTCCAATTA ATTCCAATGA TAAGCAAGTG TAGAACTAAG
GAGGGAAGGC GAAAGACCAA CTTGTATGGT TTCATCATAA AAGGAAGATC CCACTTAAGG

AATGACACCG ACGTGGTAAA CTTTGTGAGC ATGGAGTTTT CTCTCACTGA CCCAAGACTT
GAACCACATA AATGGGAGAA GTACTGTGTT CTTGAGATAG GAGATATGCT TATAAGAAGT
GCCATAGGCC AGGTTTCAAG GCCCATGTTC TTGTATGTGA GAACAAATGG AACCTCAAAA
ATTAAAATGA AATGGGGAAT GGAGATGAGG CGTTGCCTCC TCCAGTCACT TCAACAAATT
GAGAGTATGA TTGAAGCTGA GTCCTCTGTC AAAGAGAAAG ACATGACCAA AGAGTTCTTT
GAGAACAAAT CAGAAACATG GCCCATTGGA GAGTCCCCCA AAGGAGTGGA GGAAAGTTCC
ATTGGGAAGG TCTGCAGGAC TTTATTAGCA AAGTCGGTAT TCAACAGCTT GTATGCATCT
CCACAAC TAG AAGGATTTTC AGCTGAATCA AGAAAAC TGC TTCTTATCGT TCAGGCTCTT
AGGGACAACC TTGAACCTGG GACCTTTGAT CTTGGGGGGC TATATGAAGC AATTGAGGAG
TGCCTGATTA ATGATCCCTG GGTTTTGCTT AATGCTTCTT GGTTC AACTC ATTCCTTACA
CATGCATTGA GTTAGTTGTG GCAGTGCTAC TATTTGCTAT CCATACTGTC CAAAAAAGTA
CCTTGTTTCT ACT

Figure 18 - SEQ ID NO: 11 (PR8 segment 3 frameshift mutant)

AGCGAAAGCA GGTACTGATC CAAAATGGAA GATTTTGTGC GACAATGCTT CAATCCGATG
ATTGTCGAGC TTGCGGAAAA AACAATGAAA GAGTATGGGG AGGACCTGAA AATCGAAACA
AACAAATTTG CAGCAATATG CACTCACTTG GAAGTATGCT TCATGTATTC AGATTTTCAC
TTCATCAATG AGCAAGGCGA GTCAATAATC GTAGAACTTG GTGATCCAAA TGCACTTTTG
AAGCACAGAT TTGAAATAAT CGAGGGAAGA GATCGCACAA TGGCCTGGAC AGTAGTAAAC
AGTATTTGCA AACTACAGG GGCTGAGAAA CCAAAGTTTC TACCAGATTT GTATGATTAC
AAGGAGAATA GATTCATCGA AATTGGAGTA ACAAGGAGAG AAGTTCACAT ATACTATCTG
GAAAAGGCCA ATAAAATTAA ATCTGAGAAA ACACACATCC ACATTTTCTC GTTCACTGGG
GAAGAAATGG CCACAAAGGC AGACTACACT CTCGATGAAG AAAGCAGGGC TAGGATCAAA
ACCAGACTAT TCACCATAAG ACAAGAAATG GCCAGCAGAG GCCTCTGGGA TAGCTTCAGA
CAGTCCGAGA GAGGAGAAGA GACAATTGAA GAAAGGTTTG AAATCACAGG AACAATGCGC
AAGCTTGCCG ACCAAAGTCT CCCGCCGAAC TTCTCCAGCC TTGAAAATTT TAGAGCCTAT
GTGGATGGAT TCGAACCGAA CGGCTACATT GAGGGCAAGC TGTCTCAAAT GTCCAAAGAA
GTAAATGCTA GAATTGAACC TTTTTTGAAA ACAACACCAC GACCACTTAG ACTTCCGAAT
GGGCCTCCCT GTTCTCAGCG GTCCAAATTC CTA CTGATGG ATGCCTTAAA ATTAAGCATT
GAGGACCCAA GTCATGAAGG AGAGGGAATA CCGCTATATG ATGCAATCAA ATGCATGAGA
ACATTCTTTG GATGGAAGGA ACCCAATGTT GTTAAACCAC ACGAAAAGGG AATAAATCCA
AATTATCTTC TGTCATGGAA GCAAGTACTG GCAGAACTGC AGGACATTGA GAATGAGGAG
AAAATTCCAA AGACTAAAAA TATGAAGAAA ACAAGTCAGC TAAAGTGGGC ACTTGGTGAG
AACATGGCAC CAGAAAAGGT AGACTTTGAC GACTGTAAAG ATGTAGGTGA TTTGAAGCAA
TATGATAGTG ATGAACCAGA ATTGAGGTCG CTTGCAAGTT GGATTCAGAA TGAGTTTAAC
AAGGCATGCG AACTGACAGA TTCAAGCTGG ATAGAGCTCG ATGAGATTGG AGAAGATGTG
GCTCCAATTG AACACATTGC AAGCATGAGA AGGAATTATT TCACATCAGA GGTGTCTCAC
TGCAGAGCCA CAGAATACAT AATGAAGGGG GTGTACATCA ATACTGCCTT GCTTAATGCA
TCTTGTGCAG CAATGGATGA TTTCCAATTA ATTCCAATGA TAAGCAAGTG TAGAACTAAG
GAGGGAAGGC GAAAGACCAA CTTGTATGGT TTCATCATAA AAGGAAGATC CCACTTAAGG

AATGACACCG ACGTGGTAAA CTTTGTGAGC ATGGAGTTTT CTCTCACTGA CCCAAGACTT
GAACCACATA AATGGGAGAA GTACTGTGTT CTTGAGATAG GAGATATGCT TATAAGAAGT
GCCATAGGCC AGGTTTCAAG GCCCATGTTC TTGTATGTGA GAACAAATGG AACCTCAAAA
ATTAAAATGA AATGGGGAAT GGAGATGAGG CGTTGCCTCC TCCAGTCACT TCAACAAATT
GAGAGTATGA TTGAAGCTGA GTCCTCTGTC AAAGAGAAAG ACATGACCAA AGAGTTCTTT
GAGAACAAAT CAGAAACATG GCCCATTGGA GAGTCCCCCA AAGGAGTGGA GGAAAGTTCC
ATTGGGAAGG TCTGCAGGAC TTTATTAGCA AAGTCGGTAT TCAACAGCTT GTATGCATCT
CCACAAC TAG AAGGATTTTC AGCTGAATCA AGAAAAC TGC TTCTTATCGT TCAGGCTCTT
AGGGACAACC TTGAACCTGG GACCTTTGAT CTTGGGGGGC TATATGAAGC AATTGAGGAG
TGCCTGATTA ATGATCCCTG GGTTTTGCTT AATGCTTCTT GGTTC AACTC ATTCCTTACA
CATGCATTGA GTTAGTTGTG GCAGTGCTAC TATTTGCTAT CCATACTGTC CAAAAAAGTA
CCTTGTTTCT ACT

Figure 19 SEQ ID NO:12

atgtcgtacaactccgccccattgacgcaaattggcggttaggcgtgtacgggtgggaggtctatataagcagag
 ctctctggctaactagagaaccactgcttactggcttatcgaaatttaatacgaactcactatagggagacc caa
 gctgttaacgctagcagtttaaccggagtagtggctcgacctccgaagttggggggg(agagacgggtaccgtctcc
) AGCGAAAGCA GGTACTGATC CAAAATGGAA GATTTTGTGC GACAATGCTTCAATCCGATGATTGTGCGAGC
 TTGCGGAAAA AACAATGAAA GAGTATGGGG AGGACCTGAA AATCGAAACA AACAAATTTGCAGCAATATG
 CACTCACTTG GAAGTATGCT TCATGTATTC AGATTTTCAC TTCATCAATG AGCAAGGCGA GTCAATAATC
 GTAGAACTTG GTGATCCAAA TGCACTTTTG AAGCACAGAT TTGAAATAAT CGAGGGAAGA GATCGCACAA
 TGGCCTGGAC AGTAGTAAAC AGTATTTGCA AACTACAGG GGCTGAGAAA CCAAAGTTTC TACCAGATTT
 GTATGATTAC AAGGAGAATA GATTCATCGA AATTGGAGTA ACAAGGAGAG AAGTTCACAT ATACTATCTG
 GAAAAGGCCA ATAAATTAAT ATCTGAGAAA ACACACATCC ACATTTTCTC GTTCACTGGG GAAGAAATGG
 CCACAAAGGC AGACTACACT CTCGATGAAG AAAGCAGGGC TAGGATCAAA ACCAGACTAT TCACCATAAG
 ACAAGAAATG GCCAGCAGAG GCCTCTGGGA **TAGCTTCAGAC**AGTCCGAGAGAGGAGAAGAGACAATTGAA
 GAAAGGTTTG AAATCACAGG AACAATGCGC AAGCTTGCCG ACCAAAGTCT CCCGCCGAAC TTCTCCAGCC
 TTGAAAATTT TAGAGCCTAT GTGGATGGAT TCGAACCGAA CGGCTACATT GAGGGCAAGC TGTCTCAAT
 GTCCAAAGAA GTAAATGCTA GAATTGAACC TTTTTGAAA ACAACACCAC GACCACTTAG ACTTCCGAAT
 GGGCCTCCCT GTTCTCAGCG GTCCAAATTC CACTGATGG ATGCCTTAAA ATTAAGCATT GAGGACCCAA
 GTCATGAAGG AGAGGGAATA CCGCTATATG ATGCAATCAA ATGCATGAGA ACATTCTTGGATGGAAGGA
 ACCCAATGTT GTTAAACCAC ACGAAAAGGG AATAAATCCA AATTATCTTC TGTCATGGAA GCAAGTACTG
 GCAGAACTGC AGGACATTGA GAATGAGGAG AAAATTCCAA AGACTAAAAATATGAAGAAAACAAGTCAGC
 TAAAGTGGGC ACTTGGTGAG AACATGGCAC CAGAAAAGGT AGACTTTGAC GACTGTAAAGATGTAGGTGA
 TTTGAAGCAA TATGATAGTG ATGAACCAGA ATTGAGGTCG CTGCAAGTT GGATTCAAGAA TGAGTTTAAC
 AAGGCATGCG AACTGACAGA TTCAAGCTGG ATAGAGCTCG ATGAGATTGG AGAAGATGTGGCTCCAATTG
 AACACATTGC AAGCATGAGA AGGAATTATT TCACATCAGA GGTGTCTCAC TGCAGAGCCA CAGAATACAT
 AATGAAGGGG GTGTACATCA ATACTGCCTT GCTTAATGCA TCTGTGCAG CAATGGATGA TTCCAATTA
 ATTCCAATGA TAAGCAAGTG TAGAACTAAG GAGGGAAGGC GAAAGACCAA CTTGTATGGT TTCATCATAA
 AAGGAAGATC CCACTTAAGG AATGACACCG ACGTGGTAAA CTTTGTGAGC ATGGAGTTTT CTCTACTGA
 CCCAAGACTT GAACCACATA AATGGGAGAA GTACTGTGTT CTGAGATAG GAGATATGCT TATAAGAAGT
 GCCATAGGCC AGGTTTCAAG GCCCATGTTC TTGTATGTGA GAACAAATGG AACCTCAAAA ATTAATGA
 AATGGGGAAT GGAGATGAGG CGTTGCCTCC TCCAGTCACT TCAACAAATT GAGAGTATGA TTGAAGCTGA
 GTCCTCTGTC AAAGAGAAAG ACATGACCAA AGAGTTCTTT GAGAACAAAT CAGAAACATG GCCATTGGA
 GAGTCCCCCA AAGGAGTGA GGAAGTTCC ATTGGGAAGG TCTGCAGGAC TTTATTAGCA AAGTCGGTAT
 TCAACAGCTT GTATGCATCT CCACAACTAG AAGGATTTTC AGCTGAATCA AGAAAACCTGC TTCTTATCGT
 TCAGGCTCTT AGGGACAACC TTGAACCTGG GACCTTTGAT CTTGGGGGGC TATATGAAGC AATTGAGGAG
 TGCCTGATTA ATGATCCCTG GGTTTTGCTT AATGCTTCTT GGTCAACTC ATTCCTTACA CATGCATTGA
 GTTAGTTGTG GCAGTGCTAC TATTTGCTAT CCATACTGTC CAAAAAAGTA CTTGTTTCTACT
 aataacecggcgcccccaaatgcccgaactcggagcgaagatataacctccccggggcggggaggtcgctcacc
 gaccacgcgcggccccaggcgacgcgcgcgacacgggacacctgtcccccaaaacgcaccatcgccagccacacac
 ggagcgccccggggccctctgggtcaacccccaggacacacgcggggagcagcgccggggcgccctccccggc
 ggtcacctaaatgctagagctcgtgatcagcctcgactgtgccttctagttgccagccatctgttgtttgccc
 ctccccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcat
 cgcattgtctgagtaggtgtcattctattctggggggtgggggtggggcaggacagcaagggggaggattgggaa
 gacaatagcaggcatgctggggatgcgggtgggctctatggcttctgaggcggaagaaccagctgcattaatga
 atcgggccaacgcgcggggagaggcggtttgctattgggcgctcttccgcttctcgtcactgactcgtcgcg
 ctcggtcgttcggtcgggcgagcggtatcagctcactcaaaggcggttaatacggttatccacagaatcagggg
 ataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcggttgctggcg
 tttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgac
 aggactataaagataaccaggcgtttccccctggaagctccctcgtgcgctctcctgtttccgacctgcccgtta
 ccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagt
 tcggtgtaggtcggttcgctccaagctgggctgtgtgcacgaacccccgttcagccccgaccgtgcgccttatc

cggtactatcgtcttgagtccaacccggtgaagacacgacttatcgccactggcagcagccactggtaacagga
 ttagcagagcgagggtatgtaggcgtgctacagagttcttgaagtgggtggcctaactacggctacactagaagg
 acagtatcttggtatctgctgctctgctgaagccagttaccttcgaaaaagagttggtagctcttgatccggcaa
 acaaacaccgctggtagcgggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaggatctcaag
 aagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggatg
 agattatcaaaaaggatcttcacctagatccttttaaaattaaaaatgaagttttaaatcaatctaaagtatata
 tgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgtt
catccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgt
gcaatgataccgagacccacgctcaccgggtccagatttatcagcaataaaccagccagccggaagggccga
gcgagaaagtggctcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagta
gttcgccagtttaatatgtttgcgcaacgttggttgccattgctacaggcatcgtgggtgtcacgctcgtcgtttggt
atggcttcattcagctccggttcccaacgatcaaggcgagttacatgatccccatgtttgtgcaaaaaagcggt
tagctccttcggtcctccgatcgttgtcagaagtaagttggccgagtggttatcactcatggttatggcagcac
tgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactgggtgagtactcaaccaagtcattc
tgagaatagtgatgctggcgaccgagttgctcttgcccggtcaatacgggataataaccgcgccacatagcag
aactttaaaagtgtcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagat
ccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtga
gcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactctt
 cctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttaga
 aaaataaacaataggggttccgcgcacatttccccgaaaagtgccacctgacgtcgatatgccaaagtacgcc
 cctattgacgtcaatgacggtaaatggccgcct**ggcattatgccagtacatgaccttatgggactttcct**
acttggcagtacatctacgtatttagtcacgtattaccatgggtgatgcggttttggcagtacatcaatgggcg
tggatagcggtttgactcacggggatttccaagtctccacccattgacgtcaatgggagtttgttttggcacc
aaaatcaacgggactttccaaa

Figure 20.

SEQ ID NO: 13 (PR8 PA-X amino acid sequence)

MEDFVRQCFNPMIVELAETMKEYGEDLKIETNKFAAICTHLEVCFMYSDFHFINEQGESIIVELGDPNALLKHFEEI
 EGRDRTMAWTVVNSICNTTGAETPKFLPDLYDYKENRFIEIGVTRREVHIYYLEKANKIKSEKTHIHIFSFTGEEMATK
 ADYTLDEESRARIKTRLTIRQEMASRGLWDSFVSPREEKRQLKKGLKSQEQCASLPTKVSRRTPALKILEPMWMD
 SNRTATLRASCLCKPKK

Figure 21.

SEQ ID NO: 16 (PR8 PB2 amino acid sequence)

MERIKELRNLSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMAMKYPITADKRITEMIPERNEQG
 QTLWSKMNDAGSDRVMVSPLAVTWVNRNGPITNTVHYPKIYKTYFERVERLKHGTFGPVHFRNQVKIRRRVDIN
 PGHADLSAKEAQDVIMEVFPNEVGARILTSQTLITKEKKEELQDCKISPLMVAYMLERELVRKTRFLPVAGGTSS
 VYIEVLHLTQGTCEWQMYTPGGEVRNDDVDQSLIAARNIVRRAAVSADPLASLLEMCHSTQIGGIRMVDILRQNP
 TEEQAVDICKAAMGLRISSFSFGGFTFKRTSGSSVKREEEVLGNLQTLKIRVHEGYEFTMVGRRATAILRKATRRL
 IQLIVSGRDEQSIAEAIIVAMVFSQEDCMIAVRGDLNFVNRRANQRLNPMHQLLRHFQKDAKVLQNWGVEPIDN
 VMGMIGILPDMTPSIEMSMRGVRISKMGVDEYSSTERVVVSIDRFLRIRDQRGNVLLSPEEVSETQGTEKLITYSSS
 MMWEINGPESVLVNTYQWIIRNWETVKIQWSQNPTMLYNKMEFEPFQSLVPKAIRGQYSGFVRTLFQQMRDVL

GTFDTAQIIKLLPFAAAPPKQSRMQFSSFTVNVRGSGMRILVRGNSPVFNYNKATKRLTVLGKDAGTLTEDPDEGT
AGVESAVLRGFLILGKENKRYGPALSINELSNLAKGEKANVLIGQGDDVVLVLMKRKRDSILTDSQTATKRIRMAIN

Figure 22.

SEQ ID NO: 17 (PR8 PB1 amino acid sequence)

MDVNPTLLFLKVPAQNAISTTFPYTGDPPYSHGTGTGYTMDTVNRTHQYSEKGRWTTNTETGAPQLNPIDGPLPE
DNEPSGYAQTDCVLEAMAFLEESHPIGIFENSCIETMEVVQQTRVDKLTQGRQTYDWTLNRNQPAATALANTIEVF
RSNGLTANESGRLIDFLKDVME SMNKEEMGITTHFQKRKRRVRDNMTKKMITQRTMGKKKQRLNKR SYLIRALT LN
TMTKDAERGKLRRAIATPGMQIRGFVYFVETLARSICEKLEQSGLPVGGNEKKAKLANVVRKMMTNSQDTELSFT
ITGDNTKWNENQNPRMFLAMITYMTRNQPEWFRNVLSIAPIMFSNKMARLGKGYMFESKSMKLRTQIPAEMLA
SIDLKYFNDSTRKKIEKIRSLIEGTASLSPGMMMGMFNMLSTVLGV SILNLGQKRYTKTTYWWDGLQSSDDFALIV
NAPNHEGIQAGVDRFYRTCKLLGINMSKKKSYINRTGTFFTSFFYRYGFVANFSMELPSFGVSGINESADMSIGVT
VIKNNMINNDLGPATAQMALQLFIKDYRYTYRCHRGDTQIQTRRSFEIKKLWEQTRSKAGLLVSDGGPNLYNIRNL
HIPEVCLKWELMDEDYQGRLCNPLNPFVSHKEIESMNNNAVMMMPAHGPAKNMEYDAVATTHSWIPKRNRSILNT
SQRGVLEDEQMYQRCCNLFKFFPSSSYRRPVGISSMVEAMVSRARIDARIDFESGRIKKEEFTEIMKICSTIEELRRQ
K

Figure 23.

SEQ ID NO: 18 (PR8 NP amino acid sequence)

MASQGTKRSYEQMETDGERQNATEIRASVGKMIGGIGRFYIQMCTELKLSDYEGRLIQNSLTIERMVLSAFDERRN
KYLEEHPSAGKDPKKTGGPIYRRVNGKWMRELILYDKEIRRIWRQANNGDDATAGLTHMMIWHNSLNNDATYQ
RTRALVRTGMDPRMCSLMQGSTLPRRSGAAGAAVKGVGTMMVMELVRLMIKRGINDRNFWRGENGRKTRIAYER
MCNILKGFQTAQAQKAMMDQVRESRNPNGNAEFEDLTFLARSALILRGSVAHKSCLPACVYGPASVGYDFEREY
SLVGIDPFRLLQNSQVYSLIRPNENPAHKSQLVWMACHSAAFEDLRVLSFIKGTKVLPRGKLSTRGVQIASNENMET
MESSTLELRSRYWAIRTRSGGNTNQQRASAGQISIQPTFSVQRNLPDRTTIMAAFNGNTEGRTSDMRTEIIRMM
ESARPEDVSFQGRGVFELSDEKAASPIVPSFDMSNEGSYFFGDNAEEYDN

Figure 24.

SEQ ID NO: 19 (PR8 M1 amino acid sequence)

MSLLTEVETYVLSIIPSGPLKAEIAQRLEDVFAGKNTDLEVLMEWLKTRPILSPLTKGILGFVFTLTVPSEGLQRRRFV
QNALNGNDPNNMDKAVKLYRKLKREITFHGAKEISLSYSAGALASCMGLIYNRMGAVTTEVAFGLVCATCEQIA
DSQHRSHRQMVTTTNPLIRHENRMVLASTTAKAMEQMAGSSEQAAEAMEVASQARQMVQAMRTIGTHPSSS
AGLKNLLENLQAYQKRMGVQMQRFK

Figure 25.

SEQ ID NO: 20 (PR8 M2 amino acid sequence)

MSLLTEVETPIRNEWGCRCNGSSDPLTIAANIIGILHLTLWILDRLLFFKCIYRRFKYGLKGGPSTEGVPKSMREEYRKE
QQSAVDADDGHFVSIELE

Figure 26. SEQ ID NO: 21 (PR8 NS1 amino acid sequence)

MDPNTVSSFQVDCFLWHVRKRVADQELGDAPFLDRLRRDQKSLRGRGSTLGDIKTATRAGKQIVERILKEESDEA
LKMTMASVPASRYLTDMTLEEMSRDWSMLIPKQKVAGPLCIRMDQAIMDKNIILKANFSVIFDRLETILLRAFTEE
GAIVGEISPLPSLPGHTAEDVKNAVGVLLIGGLEWNDNTVRVSETLQRFARSSNENGRPPLTPKQKREMAGTIRSE
V

Figure 27. SEQ ID NO: 22 (PR8 NEP amino acid sequence)

MDPNTVSSFQDILLRMSKMQLESSSEDNGMITQFESLKLYRDSLGEAVMRMGDLHSLQNRNEKWREQLGQKFE
EIRWLIEEVRHKLKITENSFEQITFMQALHLLLEVEQEIRTFSFQLI

Figure 28. SEQ ID NO: 23 (PR8 PA amino acid sequence)

MEDFVRQCFNPMIVELAEKTMKEYGEDLKIETNKFAAICTHLEVCFMYSDFHFINEQGESIIVELGDPNALLKHFEEI
EGRDRTMAWTVVNSICNTTGAEKPKFLPDLYDYKENRFIEIGVTRREVHIYYLEKANKIKSEKTHIHIFSFTGEEMATK
ADYTLDEESRARIKTRLFTIRQEMASRGLWDSFRQSERGEETIEERFEITGTMRKLDQSLPPNFSSLENFRAYVDGF
EPNGYIEGKLSQMSKEVNARIEPFLKTTTPRLPLPNGPPCSQRSKFLMDALKLSIEDPSHEGEGIPLYDAIKCMRTFF
GWKEPNVVKPHEKGINPNYLLSWKQVLAELQDIENEEKIPKTKNMKKTSQKQWALGENMAPEKVDFFDDCKDVG
DLKQYDSDEPELRSLSWQNEFNKACELTDSSWIELDEIGEDVAPIEHASMRNRYFTSEVSHCRATEYIMKGVYIN
TALLNASCAAMDDFQLIPMISKRTKEGRRKTNLYGFIKGRSHLRNDTDVVNFVSMESLTDPRLEPHKWEKYCVL
EIGDMLIRSAIGQVSRPMFLYVRTNGTSKIKMKWGMEMRRCLLQSLQQIESMIEAESSVKEKDMTKEFFENKSET
WPIGESPKGVEESSIGKVCRTLLAKSVFNSLYASPQLEGFSAESRKLIVQALRDNLPGTFDLGGLYEAIEECLINDP
WVLLNASWFNSFLTHALS

Figure 29. SEQ ID NO: 24 (segment 4 nucleic acid molecule)

AGCAAAAGCAGGGGAAAACAAAAGCAACAAAATGAAGGCAATACTAGTAGTTCTGCTATATACATTTGTAA
CCGCAAATGCAGACACATTATGTATAGTTATCATGCGAACAATTCAACAGACACTGTAGACACAGTACTAGA
AAAGAATGTAACAGTAACACACTCTGTAACTTCTAGAAGACAAGCATAACGGGAAACTATGCAAATAAGA
GGGGTAGCCCCATTGCATTTGGGTAAATGTAACATTGCTGGCTGGATCCTGGGAAATCCAGAGTGTGAATCAC
TCTCCACAGCAAGCTCATGGTCCTACATTGTGGAAACACCTAGTTCAGACAATGGAACGTGTTACCCAGGAGA
TTTCATCGATTATGAGGAGCTAAGAGAGCAATTGAGCTCAGTGTCATCATTTGAAAGGTTTGAGATATTCCCC
AAGACAAGTTCATGGCCCAATCATGACTCGAACAAGGTGTAACGGCAGCATGTCCTCATGCTGGAGCAAAA
AGCTTCTACAAAAATTTAATATGGCTAGTTAAAnAnaGGAAATTCATACcCAAAGCTCAGCAAATCCTACATTAA

TGATAAAGGGAAAGAAGTCCTTGTGCTATGGGGCATTACCATCCATCTACTAGTGCTGACCAACAAAGTCTC
TATCAGAATGCAGATGCATATGTTTTGTGGGGTCATCAAGATACAGCAAGAAGTTCAAGCCGGAAATAGCA
ATAAGACCCAAAGTGAGGGATCGAGAAGGGAGAATGAACTATTACTGGACACTAGTAGAGCCGGGAGACAn
AATAACATTcgAAGCAACTGGAAATCTAGTGGTaccGaaaTATGCATTcgCAATGGAAagaAATGCTGGATCTGG
TATTATCATTTTCAGATACacCAGTCCACGATTGCAATACAACCTTGTCAAACACCCAAGGGTGCTATAAACACCA
GCCTCCCATTTTCAGAATATACATCCGATCACAATTGGAAAATGTCCAAaATATGTAAAAAGCACAAAATTGAGA
CTGGCCACAGGATTGAGGAATATCCCGTCTATTCAATCTAGAGGCCTATTTGGGGCCATTGCCGGTTTCATTG
AAGGGGGGTGGACAGGGATGGTAGATGGATGGTACGGTTATCACCATCAAAATGAGCAGGGGTGTCAGGATAT
GCAGCCGACCTGAAGAGCACACAGAATGCCATTGACGAGATTACTAACAAAGTAAATTCTGTTATTGAAAAGA
TGAATACACAGTTCACAGCAGTAGGTAAAGAGTTCAACCACCTAGAAAAAAGAATAGAGAATTTAAATAAAA
AAGTTGATGATGGTTTCCTGGACATTTGGACTTACAATGCCGAAGTGTGGTTCTATTGGAAAATGAAAGAAC
TTTGGACTACCACGATTCAAATGTGAAGAACTTATATGAAAAGGTAAGAAGCCAGCTAAAAACAATGCCAA
GGAAATTGGAAACGGCTGCTTTGAATTTTACCACAAATGCGATAACACGTGCATGGAAAGTGTCAAAAATGG
GACTTATGACTACCCAAAATACTCAGAGGAAGCAAAATTAACAGAGAAGAAATAGATGGGGTAAAGCTGGA
ATCAACAAGGATTACCAGATTTTGGCGATCTATTCAACTGTCGCCAGTTCATTGGTACTGGTAGTCTCCCTGG
GGGCAATCAGTTTCTGGATGTGCTCTAATGGGTCTCTACAGTGTAGAATATGTATTTAACATTAGGATTTTCAGA
AGCATGAGAAAAACAC

SEQ ID NO: 25 (corresponding HA amino acid sequence)

MKAILVLLYTFVTANADTLCIGYHANNSTDVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRGVAPLHLGKCNIAG
WILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAAC
HAGAKSFYKNLIWLXXGNSYPKLSKSYINDKGKEVLVLWGIHPSTSDQQLYQNAAYVFGSSRYSKKFKPEI
AIRPKVRDREGRMNYYWTLVEPGDXITFEATGNLVVPKYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINSLP
FQNIHPITIGKCPKYVKSTKLRLATGLRNIPSIQSRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGSYAADLKS
TQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLLNERTLDYHDSNVKN
LYEKVRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLESTRIYQILAIYSTVASS
LVLVSLGAISFWMCSNGSLQCRICI

Figure 30. SEQ ID NO: 26 (segment 6 nucleic acid molecule)

AGCAAAAGCAGGAGTTTAAATGAATCCAAACCAAAGATAATAACCATTGGTTCGGTCTGTATGACAATTGG
AATGGCTAACTTAATATTACAAATTGGAAACATAATCTCAATATGGATTAGCCACTCAATTCAACTTGGGAATC
AAAATCAGATTGAAACATGCAATCAAAGCGTCATTACTTATGAAAACAACACTTGGGTAAATCAGACATATGT
TAACATCAGCAACACCAACTTTGCTGCTGGACAGTCAGTGGTTTCCGTGAAATTAGCGGGCAATTCCTCTCTCT
GCCCTGTTAGTGGATGGGCTATATACAGTAAAGACAACAGTGTAAGAATCGGTTCCAAGGGGGATGTGTTTG
TCATAAGGGAACCATTCATATCATGCTCCCCCTTGAATGCAGAACCTTCTTCTTGACTCAAGGGGCCTTGCTA

AATGACAAACATTCCAATGGAACCATTAAGACAGGAGCCCATATCGAACCCCTAATGAGCTGTCCTATTGGTG
AAGTTCCCTCTCCATACAACCTCAAGATTTGAGTCAGTCGCTTGGTCAGCAAGTGCTTGTGCATGATGGCATCAAT
TGGCTAACAAATTGGAATTTCTGGCCCAGACAATGGGGCAGTGGCTGTGTTAAAGTACAACGGCATAATAACA
GACACTATCAAGAGTTGGAGAAACAATATATTGAGAACACAAGAGTCTGAATGTGCATGTGTAAATGGTTCTT
GCTTTACTGTAATGACCGATGGACCAAGTAATGGACAGGCCTCATACAAGATCTTCAGAATAGAAAAGGGAA
AGATAGTCAAATCAGTCGAAATGAATGCCCCTAATTATCACTATGAGGAATGCTCCTGTTATCCTGATTCTAGT
GAAATCACATGTGTGTGCAGGGATAACTGGCATGGCTCGAATCGACCGTGGGTGTCTTCAACCAGAATCTG
GAATATCAGATAGGATACATATGCAGTGGGATTTTCGGAGACAATCCACGCCCTAATGATAAGACAGGCAGT
TGTGGTCCAGTATCGTCTAATGGAGCAAATGGAGTAAAAGGGTTTTCATACAAATATGGCAATGGTGTGGGA
TAGGGAGAACTAAAAGCATTAGTTCAAGAAACGGTTTTGAGATGATTTGGGATCCGAACGGATGGACTGGGA
CAGACAATAACTTCTCAATAAAGCAAGATATCGTAGGAATAAATGAGTGGTCAGGATATAGCGGGAGTTTTG
TTCAGCATCCAGAACTAACAGGGCTGGATTGTATAAGACCTTGCTTCTGGGTGAACTAATCAGAGGGCGACC
CAAAGAGAACACAATCTGGACTAGCGGGAGCAGCATATCCTTTTGTGGTGTAACAGTGACACTGTGGGTTG
GTCTTGCCAGACGGTGCTGAGTTGCCATTTACCATTGACAAGTAATTGTTCAAAAAAC

SEQ ID NO: 27 (corresponding NA amino acid sequence)

MNPNQKIITIGSVCMTIGMANLILQIGNIISIWISHSIQLGNQNIETCNQSVITYENNTWVNQTYVNISNTNFAAG
QSVVSVKLAGNSSLCPVSGWAIYSKDNSVRIGSKGDVFIREFISCSPLECRFFLTQGALLNDKHSNGTIKDRSPYR
TLMSCPIGEVPSPYNSRFESVAWSASACHDGINWLTIGISGPDNGAVAVLKYNGIITDTIKSWRNNILRTQESECAC
VNGSCFTVMTDGPSNGQASYKIFRIEKGKIVKSVEMNAPNYHYEECSCYPDSSEITCVCRDNWHGSNRPWVSFNQ
NLEYQIGYICSGIFGDNPRPNDKTGSCGPVSSNGANGVKGFYKYGNVWIGRTKSISSRNGFEMIWDPNGWTGT
DNNFSIKQDIVGINEWSGYSGSFVQHPELTGLDCIRPCFWVELIRGRPKENTIWTSGSSISFCGVNSDTVGWSWPD
GAELPFTIDK

INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2017/050329

A. CLASSIFICATION OF SUBJECT MATTER

INV. C07K14/005 A61K39/12
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K A61K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	HUIJIE GAO ET AL: "The contribution of PA-X to the virulence of pandemic 2009 H1N1 and highly pathogenic H5N1 avian influenza viruses", SCIENTIFIC REPORTS, vol. 5, 5 February 2015 (2015-02-05), page 8262, XP055361841, DOI: 10.1038/srep08262 abstract page 2, column 1, paragraph 2 - column 2, paragraph 2 page 10, column 1, paragraph 1 - paragraph 4; figure 1 ----- -/-	1-45



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

6 April 2017

Date of mailing of the international search report

02/05/2017

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Mossier, Birgit

INTERNATIONAL SEARCH REPORT

International application No.

PCT/GB2017/050329

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed:
 - ☒ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
 - b. ☐ furnished together with the international application under PCT Rule 13~~ter~~.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
 - ☐ in the form of an Annex C/ST.25 text file (Rule 13~~ter~~.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13~~ter~~.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2017/050329

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	TSUYOSHI HAYASHI ET AL: "Influenza A Virus Protein PA-X Contributes to Viral Growth and Suppression of the Host Antiviral and Immune Responses", JOURNAL OF VIROLOGY., vol. 89, no. 12, 15 June 2015 (2015-06-15), pages 6442-6452, XP055362007, US ISSN: 0022-538X, DOI: 10.1128/JVI.00319-15 abstract page 6443, column 2, paragraph 2 - paragraph 4 page 6445, column 2, paragraph 6 - page 6446, column 2, paragraph 1; figure 1 -----	1-45
Y	JIAO HU ET AL: "PA-X Decreases the Pathogenicity of Highly Pathogenic H5N1 Influenza A Virus in Avian Species by Inhibiting Virus Replication and Host ResponseABSTRACT", JOURNAL OF VIROLOGY., vol. 89, no. 8, 15 April 2015 (2015-04-15), pages 4126-4142, XP055361684, US ISSN: 0022-538X, DOI: 10.1128/JVI.02132-14 abstract; figure 1 -----	1-45
Y	B. W. JAGGER ET AL: "An Overlapping Protein-Coding Region in Influenza A Virus Segment 3 Modulates the Host Response", SCIENCE, vol. 337, no. 6091, 13 July 2012 (2012-07-13), pages 199-204, XP055361970, ISSN: 0036-8075, DOI: 10.1126/science.1222213 cited in the application abstract page 201, column 3, paragraph 2 - page 202, column 1, paragraph 2 page 204, column 1 - column 2; figure 1 -----	1-45
Y	US 2014/227310 A1 (LI ZEJUN [CN] ET AL) 14 August 2014 (2014-08-14) paragraph [0006] - paragraph [0008] paragraph [0015] - paragraph [0022]; claims 1,9; examples 1-3; tables 2,3 -----	1-45
Y	WO 2014/086732 A2 (NOVARTIS AG [CH]; SYNTHETIC GENOMICS VACCINES INC [US]) 12 June 2014 (2014-06-12) page 1, line 30 - page 2, line 12 page 3, line 23 - line 36 page 30, line 20 - page 32, line 11 ----- -/--	1-45

INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2017/050329

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2010/070098 A1 (AVIR GREEN HILLS BIOTECHNOLOGY [AT]; MUSTER THOMAS [AT]; WOLSCHEK MARK) 24 June 2010 (2010-06-24) abstract; example 1 -----	1-45
Y	BAEZ M ET AL: "Gene composition of high-yielding influenza vaccine strains obtained by recombination", JOURNAL OF INFECTIOUS DISEASES. JID, UNIVERSITY OF CHICAGO PRESS, CHICAGO, IL, vol. 141, no. 3, 1 March 1980 (1980-03-01), pages 362-365, XP008126560, ISSN: 0022-1899 abstract; table 2 -----	1-45
A	WO 2007/048089 A2 (MEDIMMUNE VACCINES INC [US]; LU BIN [US]; HONG JIN [US]) 26 April 2007 (2007-04-26) abstract paragraph [0204] - paragraph [0216]; figure 16 -----	1-45
A	DENYS A. KHAPERSKYY ET AL: "Timing Is Everything: Coordinated Control of Host Shutoff by Influenza A Virus NS1 and PA-X Proteins", JOURNAL OF VIROLOGY., vol. 89, no. 13, 15 April 2015 (2015-04-15), pages 6528-6531, XP055362010, US ISSN: 0022-538X, DOI: 10.1128/JVI.00386-15 abstract page 6530, column 1, paragraph 2 - column 2 -----	1-45

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2017/050329

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2014227310 A1	14-08-2014	AU 2012306866 A1	10-04-2014
		CA 2848117 A1	14-03-2013
		CN 102304495 A	04-01-2012
		US 2014227310 A1	14-08-2014
		WO 2013034069 A1	14-03-2013

WO 2014086732 A2	12-06-2014	AU 2013354219 A1	02-07-2015
		CA 2893429 A1	12-06-2014
		CN 105120893 A	02-12-2015
		EP 2925356 A2	07-10-2015
		HK 1214959 A1	12-08-2016
		JP 2016501020 A	18-01-2016
		KR 20150110494 A	02-10-2015
		US 2015315548 A1	05-11-2015
		WO 2014086732 A2	12-06-2014

WO 2010070098 A1	24-06-2010	EP 2233152 A1	29-09-2010
		EP 2373334 A1	12-10-2011
		US 2011262481 A1	27-10-2011
		WO 2010070098 A1	24-06-2010

WO 2007048089 A2	26-04-2007	AU 2006304898 A1	26-04-2007
		CA 2626266 A1	26-04-2007
		CA 2891994 A1	26-04-2007
		EP 1945778 A2	23-07-2008
		EP 2614836 A2	17-07-2013
		ES 2402061 T3	26-04-2013
		JP 5349049 B2	20-11-2013
		JP 2009511073 A	19-03-2009
		JP 2013236638 A	28-11-2013
		US 2009297554 A1	03-12-2009
		US 2013102053 A1	25-04-2013
		WO 2007048089 A2	26-04-2007
