INFLUENZA VIRUS VACCINE
COMPOSITION AND METHODS OF USE

Inventors: Catherine J. Luke, Frederick, MD
(US); Adrian Violta, San Diego, CA
(US); Mary K. Wloch, San Diego, CA
(US); Thomas G. Evans, San Diego,
CA (US); Andrew J. Geall, San
Marcos, CA (US); Gretchen S.
Jimenez, San Diego, CA (US)

Correspondence Address:
STERNE, KESSLER, GOLDSMITH & FOX PLLC
1100 NEW YORK AVENUE, N.W.
WASHINGTON, DC 20005 (US)

Publication Classification

Int. Cl.
C12Q 1/70 (2006.01)
C07H 21/04 (2006.01)
C12P 21/06 (2006.01)
A61K 39/145 (2006.01)
C07K 14/11 (2006.01)

U.S. Cl. 435/5; 435/69.1; 435/456;
435/325; 424/209.1; 530/350;
536/23.72

ABSTRACT

The present invention is directed to enhancing the immune response of a human in need of protection against IV infection by administering in vivo, into a tissue of the human, at least one polynucleotide comprising one or more regions of nucleic acid encoding an IV protein or a fragment, a variant, or a derivative thereof. The present invention is further directed to enhancing the immune response of a human in need of protection against IV infection by administering, in vivo, into a tissue of the human, at least one IV protein or a fragment, a variant, or derivative thereof. The IV protein can be, for example, in purified form or can be an inactivated IV, such as those present in inactivated IV vaccines. The polynucleotide is incorporated into the cells of the human in vivo, and an immunologically effective amount of an immunogenic epitope of an IV, or a fragment, variant, or derivative thereof is produced in vivo. The IV protein (in purified form or in the form of an inactivated IV vaccine) is also administered in an immunologically effective amount.

NP ORF
SEQ ID NO: 23

Consensus
ATGGCCCTCTC ArGGGACTcAA AGCATCTTAC GAAAGAATAG AGAAGAGATTG

NP ORF
SEQ ID NO: 23

Consensus
AGAACCGCCAG AAATGACTCG CAAATGACTCG CAGAAGATCG AGAGATGATAG

NP ORF
SEQ ID NO: 23

Consensus
GTGGAATTGG AGACATCTC CTCATGACTG GACACGCACT CAACTGACGT

NP ORF
SEQ ID NO: 23

Consensus
GATTATGAAAG GACGGTGGAT CCAAAACGAC CCAAAATAG AGAGAACATGGT

NP ORF
SEQ ID NO: 23

Consensus
GCTTATCGAA TGGACGCTGAG GACCAGCACA ATACCTTGGG GAGATCTCCTC

NP ORF
SEQ ID NO: 23

Consensus
swGCsGAsAA rGATCCGGAA AAAACTGCGG GACATCTGAc TsGAGAGATG

NP ORF
SEQ ID NO: 23

Consensus
AACACCTAAAGG AGATGAGCAG GACGCTTCTGTTT TATGACAAAG AAGAAAATAG

NP ORF
SEQ ID NO: 23

Consensus
GGCGGATGCTG CCGAAAAGCT TATGACAAAG AAGAAAATCG

NP ORF
SEQ ID NO: 23

Consensus
GGCGGATGCTG CCGAAAAGCT TATGACAAAG AAGAAAATCG

NP ORF
SEQ ID NO: 23

Consensus
GGCGGATGCTG CCGAAAAGCT TATGACAAAG AAGAAAATCG
Consensus

ACmGrGCyCyTmGOTkmGmAC mGGMATGqGT cCmGrATGT GCTCwCTkAT

Consensus

GCArGgTCTm ACyCTcCCwA GGAGGwcyGG rGcmG3CwGGT GCyGCAGTC

Consensus

AAGGrGTkGG AACwATGGTG ATGGArTyTR TsmGwATGT yAArCGyGGs

Consensus

ATmAATGAYC GsAYTTCTG GmGGGKgGAr AAYGGACGAA ArA5AAmGmAT

Consensus

TGcWTATgGA mGMATGqTGCA AyATTCtCAA rGGGAAATTy CARAckGCTG

Consensus

CwC4A4A4GC mATGATGGAY cARGTGAgGR AGwsmmGrAA CCCAGGsAAY
Figure 1C
Figure 1D
0 - 2°C below cloud point

846 µl of 1.277 mM BAK in PBS
10 min

0 - 2°C below cloud point

Poloxamer added via positive Displacement pipette, 27 µl
60 min

Room temperature above cloud point

Add 2.727 ml of 6.4 mg/ml VR4700 to BAK/Poloxamer solution
15 min

Cycle through the cloud point

Stirring solution warmed on bench
15 min

Cycle through the cloud point

one cycle

Cycle through the cloud point

one cycle

Cycle through the cloud point

one cycle

Cycle through the cloud point

one cycle

Cycle through the cloud point

one cycle

Cycle through the cloud point

one cycle

Freeze at -80°C

FIG. 2
34mg/ml → 1.557ml of VR4700 in PBS
50mg/ml → 1.525ml of VR4700 in PBS

10 min

Poloxamer added via positive Displacement pipette,
34mg/ml → 68μl poloxamer
50mg/ml → 100μl poloxamer

30 min

0 - 2°C below cloud point
BAK added slowly
Drop wise, 1.6mM, 375μl
30 min

Room temperature above cloud point
Stirring solution warmed on bench
15 min

Stirring solution cooled in ice bath
15 min
One cycle

Stirring solution warmed on bench
15 min

Cycled through cloud point
2 more cycles

1:2 dilution in PBS

Room temperature above cloud point
Filtration

Product in vials

0 - 2°C below cloud point
Cooled below cloud point and frozen @ -80°C

FIG. 3
BAK solution
0.78 ml @ 0.77 mM

0 - 2 °C below cloud point
Poloxamer added via positive Displacement pipette, 15 μl

DNA solution added slowly Drop wise, 8.3 mg/ml, 1.205ml

0 - 2 °C below cloud point
Filter sterilization

Room temperature above cloud point
2 x 900 μl aliquots in glass vials

0 - 2 °C below cloud point
Cooled below cloud point and frozen @ -80 °C

FIG. 4
Figure 5

Anti-NP antibody, prior to boost

Figure 6

Anti-NP antibody, post-boost
Figure 9A

Two dose regimen immunization study with pDNA encoding Influenza A HA (H3) (Survival)

% of mice challenged that survived

Days Post Infection

Group A: A/PR/86; 500 PFU
Group B: HA; 5,000 PFU
Group C: HA; 500 PFU
Group D: HA; 50 PFU
Group E: VR10551; 5,000 PFU
Group F: VR10551; 500 PFU
Group G: VR10551; 50 PFU

Figure 9B

Two dose regimen immunization study with plasmid DNA encoding Influenza A HA H3

Body mass (g)

Days Post Infection

Group A: A/PR/86; 500 PFU
Group B: HA; 5,000 PFU
Group C: HA; 500 PFU
Group D: HA; 50 PFU
Group E: VR10551; 5,000 PFU
Group F: VR10551; 500 PFU
Group G: VR10551; 50 PFU
**Figure 11A**

Expression of eM2NP

<table>
<thead>
<tr>
<th>Marker kDa</th>
<th>VR4757</th>
<th>VR4758</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>111</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>eM2NP</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 11B**

Expression of NP pDNAs

<table>
<thead>
<tr>
<th>Marker kDa</th>
<th>VR4700</th>
<th>VR4761</th>
<th>VR4762</th>
<th>VR4551</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP</td>
<td>61</td>
<td>48</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
NP consensus vs. 1990-2000 strains

<table>
<thead>
<tr>
<th>NP consensus</th>
<th>1 masqgtkrsyeqmetdgerqaneteirsavgkmidgigfyigmctekilksdyegrilqns</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000trans is</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>2000trans is</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1999trans ay</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1999trans af</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1999trans aj</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1998trans ab</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1998trans AF</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1998trans af</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1997trans AF</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1997trans az</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1996trans AF</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1996trans af</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1995trans u7</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1994trans u7</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1993trans af</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1993trans AB</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1993trans a7</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1992trans af</td>
<td>1 ........................................................................</td>
</tr>
<tr>
<td>1991trans lo</td>
<td>1 ........................................................................</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NP consensus</th>
<th>61 ltiervylarvdermrnyloepshagdpkktqplyrvdkgmrelvlydkeirriw</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000trans is</td>
<td>61 ........................................................................</td>
</tr>
<tr>
<td>2000trans is</td>
<td>61 ........................................................................</td>
</tr>
<tr>
<td>1999trans ay</td>
<td>49 i. ........................................................................</td>
</tr>
<tr>
<td>1999trans af</td>
<td>61 ........................................................................</td>
</tr>
<tr>
<td>1999trans aj</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1998trans ab</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1998trans AF</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1998trans af</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1997trans AF</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1997trans af</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1997trans af</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1997trans AF</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1995trans AB</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1995trans u7</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1994trans u7</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1993trans af</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1993trans a7</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1992trans s5</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1990trans 10</td>
<td>61 i. ........................................................................</td>
</tr>
<tr>
<td>1990trans lo</td>
<td>61 i. ........................................................................</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NP consensus</th>
<th>121 qaanggedataalthmihwhsnldtttyqlrtaavtqamdrpmslmqstlprrsqaag</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000trans is</td>
<td>121 ........................................................................</td>
</tr>
<tr>
<td>2000trans is</td>
<td>121 ........................................................................</td>
</tr>
<tr>
<td>1999trans ay</td>
<td>109 ........................................................................</td>
</tr>
<tr>
<td>1999trans af</td>
<td>121 ........................................................................</td>
</tr>
<tr>
<td>1999trans aj</td>
<td>121 ........................................................................</td>
</tr>
<tr>
<td>1998trans ab</td>
<td>121 ........................................................................</td>
</tr>
<tr>
<td>1998trans AF</td>
<td>121 ........................................................................</td>
</tr>
<tr>
<td>1997trans af</td>
<td>121 ........................................................................</td>
</tr>
<tr>
<td>1997trans AJ</td>
<td>121 ........................................................................</td>
</tr>
</tbody>
</table>

Figure 12A
NP consensus vs. 1990-2000 strains

Figure 12B
NP consensus vs. 1990-2000 strains


Figure 12C
NP consensus vs. 1990-2000 strains

1990trans lo 421 ..............................................k...........................................
NP consensus 480 dmsnegsyffgdnaeydn
2000trans is 480 e.................................
2000trans is 480 w..................................9
1999trans ay 468 ......................................
1999trans af 480......................................
1999trans aj 480......................................
1998trans ab 480......................................
1998trans AF 480......................................d...
1998trans af 480......................................
1997trans AJ 480......................................
1997trans AF 480......................................
1997trans AF 480......................................
1997trans af 480......................................
1997trans af 480......................................
1996trans af 480......................................
1995trans AB 480......................................
1995trans u7 480......................................
1994trans u7 480......................................
1993trans af 480......................................
1991trans l2 480......................................
1991trans z5 480......................................
1990trans 10 480......................................
1990trans lo 480......................................
<table>
<thead>
<tr>
<th>Plasmid #</th>
<th>Gene Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>VR4756</td>
<td></td>
</tr>
<tr>
<td>VR4759</td>
<td></td>
</tr>
<tr>
<td>VR4762</td>
<td></td>
</tr>
<tr>
<td>VR4764</td>
<td></td>
</tr>
<tr>
<td>VR4765</td>
<td></td>
</tr>
<tr>
<td>VR4766</td>
<td></td>
</tr>
<tr>
<td>VR4767</td>
<td></td>
</tr>
<tr>
<td>VR4768</td>
<td></td>
</tr>
<tr>
<td>VR4769</td>
<td></td>
</tr>
<tr>
<td>VR4770</td>
<td></td>
</tr>
<tr>
<td>VR4771</td>
<td></td>
</tr>
<tr>
<td>VR4772</td>
<td></td>
</tr>
<tr>
<td>VR4773</td>
<td></td>
</tr>
<tr>
<td>VR4774</td>
<td></td>
</tr>
<tr>
<td>VR4775</td>
<td></td>
</tr>
<tr>
<td>VR4776</td>
<td></td>
</tr>
<tr>
<td>VR4777</td>
<td></td>
</tr>
<tr>
<td>VR4778</td>
<td></td>
</tr>
<tr>
<td>VR4779</td>
<td></td>
</tr>
<tr>
<td>VR4780</td>
<td></td>
</tr>
</tbody>
</table>

**Gene Elements:**
- **CMV** Human CMV promoter
- **RSV** RSV promoter from Alavedin XbaI site (down) near transcriptional start
- **Flank** Modified RSV promoter (G. Herstellung changed XbaI to asp1I)
- **Intron A** Intron A from human CMV immediate early
- **PolyA** R 120bp from US region of HTLV-I. NcoI hybrid link with HTLV-I donor CMV-IE intron A acceptor
- **N1A** codon optimized, consensus
- **Seg 7** Segment 7 consensus sequence encoding M1 and M2
- **L1A** M2 codon optimized, consensus
- **BGH** Bovine Growth Hormone terminator
- **mRBG** Modified rabbit beta-globin terminator, Proudfoot
Influenza A Western blots: M2 and NP expression in VM92 cells

Figure 14
Influenza A Western Blots:
M1, M2, and NP
eexpression in VM92 cells
INFLUENZA VIRUS VACCINE COMPOSITION AND METHODS OF USE

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims the benefit of the filing date of U.S. Provisional Application No. 60/571,854 filed May 18, 2004, which is incorporated herein by reference in its entirety.

REFERENCE TO A SEQUENCE LISTING SUBMITTED ON A COMPACT DISC

[0002] This application includes a “Sequence Listing,” which is provided as an electronic document on a compact disk (CD-R). This compact disk contains the file “Sequence Listing.txt” (340,000 bytes, created on May 18, 2005), which is hereby incorporated by reference in its entirety.

BACKGROUND OF THE INVENTION

[0003] The present invention relates to influenza virus vaccine compositions and methods of treating or preventing influenza infection and disease in mammals. Influenza is an acute febrile illness caused by infection of the respiratory tract. There are three types of influenza viruses: A, B, and C “IAV,” “IBV” or “IAC,” respectively, or generally “IV.” Type A, which includes several subtypes, causes widespread epidemics and global pandemics such as those that occurred in 1918, 1957 and 1968. Type B causes regional epidemics. Type C causes sporadic cases and minor, local outbreaks. These virus types are distinguished in part on the basis of differences in two structural proteins, the nucleoprotein, found in the center of the virus, and the matrix protein, which forms the viral shell.

[0004] The disease can cause significant systemic symptoms, severe illness requiring hospitalization (such as viral pneumonia), and complications such as secondary bacterial pneumonia. More than 20 million people died during the pandemic flu season of 1918/1919, the largest pandemic of the 20th century. Recent epidemics in the United States are believed to have resulted in greater than 10,000 (up to 40,000) excess deaths per year and 5,000-10,000 deaths per year in non-epidemic years.

[0005] The best strategy for prevention of morbidity and mortality associated with influenza is vaccination. Vaccination is especially recommended for people in high-risk groups, such as residents of nursing or residential homes, as well as for diabetes, chronic renal failure, or chronic respiratory conditions.

[0006] Traditional methods of producing influenza vaccines involve growth of an isolated strain in embryonated hens’ eggs. Initially, the virus is recovered from a throat swab or similar source and isolated in eggs. The initial isolation in egg is difficult, but the virus adapts to its host and subsequent propagation in eggs takes place relatively easily. It is widely recognized, however, that the egg-derived production of IV for vaccine purposes has several disadvantages. One disadvantage is that such production process is rather vulnerable due to the varying (micro)biological quality of the eggs. Another disadvantage is that the process completely lacks flexibility if demand suddenly increases, i.e., in case of a serious epidemic or pandemic, because of the logistical problems due to the non-availability of large quantities of suitable eggs. Also, vaccines thus produced are contra-indicated for persons with a known hypersensitivity to chicken and/or egg proteins.

[0007] The influenza vaccines currently in use are designated whole virus (WV) vaccine or subvirion (SV) (also called “split” or “purified surface antigen”). The WV vaccine contains intact, inactivated virus, whereas the SV vaccine contains purified virus disrupted with detergents that solubilize the lipid-containing viral envelope, followed by chemical inactivation of residual virus. Attenuated viral vaccines against influenza are also in development. A discussion of methods of preparing conventional vaccine may be found in Wright, P. F. & Webster, R. G., FIELDS VIROLOGY, 4th Ed. (Knipe, D. M. et al. Eds.), 1464-65 (2001), for example.

Virus Structures

[0008] An IV is roughly spherical, but it can also be elongated or irregularly shaped. Inside the virus, eight segments of single-stranded RNA contain the genetic instructions for making the virus. The most striking feature of the virus is a layer of spikes projecting outward over its surface. There are two different types of spikes: one is composed of the molecule hemagglutinin (HA), the other of neuraminidase (NA). The HA molecule allows the virus to “stick” to a cell, initiating infection. The NA molecule allows newly formed viruses to exit their host cell without sticking to the cell surface or to each other. The viral capsid is comprised of viral ribonucleic acid and several so-called “internal” proteins (polymerases P1, P2, and PA, matrix protein M1) and nucleoprotein (NP). Because antibodies against HA and NA have traditionally proved the most effective in fighting infection, much research has focused on the structure, function, and genetic variation of these molecules. Researchers are also interested in a two non-structural proteins M2 and NS1, both molecules play important roles in viral infection.

[0009] Type A subtypes are described by a nomenclature system that includes the geographic site of discovery, a lab identification number, the year of discovery, and in parentheses the type of HA and NA it possesses, for example, A/Hong Kong/156/97 (H5N1). If the virus infects non-humans, the host species is included before the geographical site, as in A/Chicken/Hong Kong/G9/77 (H9N2).

[0010] Virions contain 7 segments (influenza C virus) to 8 segments (influenza A and B virus) of linear negative-sense single stranded RNA. Most of the segments of the virus genome code for a single protein. For many influenza viruses, the whole genome is now known. Genetic reassortment of the virus results from intermixing of the parental gene segments in the progeny of the viruses when a cell is co-infected by two different viruses of a given type. This phenomenon is facilitated by the segmental nature of the genome of influenza virus. Genetic reassortment is manifested as sudden changes in the viral surface antigens.

[0011] Antigenic changes in HA and NA allow the influenza virus to have tremendous variability. Antigenic drift is the term used to indicate minor antigenic variations in HA and NA of the influenza virus from the original parental virus, while major changes in HA and NA which make the new virions significantly different, are called Antigenic shift. The difference between the two phenomena is a matter of degree.
[0012] Antigenic drift (minor changes) occurs due to accumulation of point mutations in the gene which results in changes in the amino acids in the proteins. Changes which are extreme, and drastic (too drastic to be explained by mutation alone) result in antigenic shift of the virus. The segmented genomes of the influenza viruses reassort readily in double infected cells. Genetic reassortment between human and non-human influenza virus has been suggested as a mechanism for antigenic shift. Influenza is a zoonotic disease, and an important pathogen in a number of animal species, including swine, horses, and birds, both wild and domestic. Influenza viruses are transferred to humans from other species.

[0013] Because of antigenic shift and antigenic drift, immunity to an IV carrying a particular HA and/or NA protein does not necessarily confer protective immunity against IV strains carrying variant, or different HA and/or NA proteins. Because antibodies against HA and NA have traditionally proved the most effective in fighting IV infections, much research has focused on the structure, function and genetic variation of these molecules.

Recent IV Vaccine Candidates

[0014] During the past few years, there has been substantial interest in testing DNA-based vaccines for a number of infectious diseases where the need for a vaccine, or an improved vaccine, exists. Several well-recognized advantages of DNA-based vaccines include the speed, ease and cost of manufacture, the versatility of developing and testing multivalent vaccines, the finding that DNA vaccines can produce a robust cellular response in a wide variety of animal models as well as in humans, and the proven safety of using plasmid DNA as a delivery vector (Donnelly, J. J., et al., Annu. Rev. Immunol. 15:617-648 (1997); Manickan, E., et al., Crit. Rev. Immunol. 17(2):139-154 (1997); U.S. Pat. No. 6,214,804). DNA vaccines represent the next generation in the development of vaccines (Nossal, G., Nat. Med. 4(5 Suppl.):475-476 (1998)) and numerous DNA vaccines are in clinical trials. The above references are herein incorporated by reference in their entirety.

[0015] Studies have already been performed using DNA-based vaccines in animals. Ulmer, J. B. et al., Science 259:1749-5 (1993) revealed that mice could be protected by an IV nucleoprotein DNA vaccine alone against severe disease and death resulting from either a homologous or a heterologous IV challenge. Further studies have substantiated this model, and comparative studies of live influenza vaccines versus DNA influenza vaccines show them to be relatively equivalent in immune induction and protection in the murine model.

[0016] WO 94/21797, incorporated herein by reference in its entirety, discloses IV vaccine compositions comprising DNA constructs encoding NP, HA, M1, PB1 and NS1. WO 94/21797 also discloses methods of protecting against IV infection comprising immunization with a prophylactically effective amount of these DNA vaccine compositions.


[0018] Heterologous “prime boost” strategies have been effective for enhancing immune responses and protection against numerous pathogens. Schneider et al., Immunol. Rev. 170:29-38 (1999); Robinson, H. L., Nat. Rev. Immunol. 2:239-50 (2002); Gonzalez, R. M. et al., Vaccine 20:1226-31 (2002); Tanghe, A., Infect. Immun. 69:3041-7 (2001). Providing antigen in different forms in the prime and the boost injections appears to maximize the immune response to the antigen. DNA vaccine priming followed by boosting with protein in adjuvant or by viral vector delivery of DNA encoding antigen appears to be the most effective way of improving antigen specific antibody and CD4+ T-cell responses or CD8+ T-cell responses respectively. Shiver J. W. et al., Nature 415:331-5 (2002); Gilbert, S. C. et al., Vaccine 20:1039-45 (2002); Billault-Mulot, O. et al., Vaccine 19:95-102 (2000); Sin, J. I. et al., DNA Cell Biol. 18:771-9 (1999). Recent data from monkey vaccination studies suggests that adding CRL1005 poloxamer (12 kDa, 5% POE), to DNA encoding the HIV gag antigen enhances T-cell responses when monkeys are vaccinated with an HIV gag DNA prime followed by a boost with an adenviral vector expressing HIV gag (Ad5-gag). The cellular immune responses for a DNA/poloxamer prime followed by an Ad5-gag boost were greater than the responses induced with a DNA (without poloxamer) prime followed by Ad5-gag boost or for Ad5-gag only. Shiver, J. W. et al, Nature 415:331-5 (2002). U.S. patent application Publication No. US 2002/0165172 A1 describes simultaneous administration of a vector construct encoding an immunogenic portion of an antigen and a protein comprising the immunogenic portion of an antigen such that an immune response is generated. The document is limited to hepatitis B antigens and HIV antigens. Moreover, U.S. Pat. No. 6,501,432 is directed to methods of enhancing an immune response of nucleic acid vaccination by simultaneous administration of a polynucleotide and polypeptide of interest. According to the patent, simultaneous administration means administration of the polynucleotide and the polypeptide during the same immune response, preferably within 0-10 or 3-7 days.
of each other. The antigens contemplated by the patent include, among others, those of Hepatitis (all forms), HSV, HIV, CMV, EBV, RSV, VZV, HPV, polio, influenza, parasites (e.g., from the genus *Plasmodium*), and pathogenic bacteria (including but not limited to *M. tuberculosis*, *M. leprae*, *Chlamydia*, *Shigella*, *B. burgdorferi*, enterotoxigenic *E. coli*, *S. typhosa*, *H. pylori*, *V. cholerae*, *B. pertussis*, etc.). All of the above references are herein incorporated by reference in their entireties.

SUMMARY OF THE INVENTION

[0019] The present invention is directed to enhancing the immune response of a vertebrate in need of protection against IV infection by administering in vivo, into a tissue of the vertebrate, at least one polynucleotide, wherein the polynucleotide comprises one or more nucleic acid fragments, where the one or more nucleic acid fragments are optionally fragments of codon-optimized coding regions operably encoding one or more IV polypeptides, or fragments, variants, or derivatives thereof. The present invention is further directed to enhancing the immune response of a vertebrate in need of protection against IV infection by administering, in vivo, into a tissue of the vertebrate, a polynucleotide described above plus at least one isolated IV polypeptide or a fragment, a variant, or derivative thereof. The isolated IV polypeptide can be, for example, a purified subunit, a recombinant protein, a viral vector expressing an isolated IV polypeptide, or can be an inactivated or attenuated IV, such as those present in conventional IV vaccines. According to either method, the polynucleotide is incorporated into the cells of the vertebrate in vivo, and an immunologically effective amount of an immunogen epitope of the encoded IV polypeptide, or a fragment, variant, or derivative thereof, is produced in vivo. When utilized, an isolated IV polypeptide or a fragment, variant, or derivative thereof is also administered in an immunologically effective amount.

[0020] According to the present invention, the polynucleotide can be administered either prior to, at the same time (simultaneously), or subsequent to the administration of the isolated IV polypeptide. The IV polypeptide or fragment, variant, or derivative thereof encoded by the polynucleotide comprises at least one immunogenic epitope capable of eliciting an immune response to influenza virus in a vertebrate. In addition, an isolated IV polypeptide or fragment, variant, or derivative thereof, when used, comprises at least one immunogenic epitope capable of eliciting an immune response in a vertebrate. The IV polypeptide or fragment, variant, or derivative thereof encoded by the polynucleotide can, but need not, be the same protein or fragment, variant, or derivative thereof as the isolated IV polypeptide which can be administered according to the method.

[0021] The polynucleotide of the invention can comprise a nucleic acid fragment, where the nucleic acid fragment is a fragment of a codon-optimized coding region operably encoding any IV polypeptide or fragment, variant, or derivative thereof, including, but not limited to, HA, NA, NP, M1, or M2 proteins or fragments (e.g., eM2), variants or derivatives thereof. A polynucleotide of the invention can also encode a derivative fusion protein, wherein two or more nucleic acid fragments, at least one of which encodes an IV polypeptide or fragment, variant, or derivative thereof, arejoined in frame to encode a single polypeptide, e.g., NP fused to eM2. Additionally, a polynucleotide of the invention can further comprise a heterologous nucleic acid or nucleic acid fragment. Such heterologous nucleic acid or nucleic acid fragment may encode a heterologous polypeptide fused in frame with the polynucleotide encoding the IV polypeptide, e.g., a hepatitis B core protein or a secretory signal peptide. Preferably, the polynucleotide encodes an IV polypeptide or fragment, variant, or derivative thereof comprising at least one immunogenic epitope of IV, wherein the epitope elicits a B-cell (antibody) response, a T-cell (e.g., CTL) response, or both.

[0022] Similarly, the isolated IV polypeptide or fragment, variant, or derivative thereof to be delivered (either a recombinant protein, a purified subunit, or viral vector expressing an isolated IV polypeptide, or in the form of an inactivated IV vaccine) can be any isolated IV polypeptide or fragment, variant, or derivative thereof, including but not limited to the HA, NA, NP, M1 or M2 proteins or fragments (e.g., eM2), variants or derivatives thereof. In certain embodiments, a derivative protein can be a fusion protein, e.g., NP-eM2. In other embodiments, the isolated IV polypeptide or fragment, variant, or derivative thereof can be fused to a heterologous protein, e.g., a secretory signal peptide or the hepatitis B virus core protein. Preferably, the isolated IV polypeptide or fragment, variant, or derivative thereof comprises at least one immunogenic epitope of IV, wherein the antigen elicits a B-cell antibody response, a T-cell antibody response, or both.

[0023] Nucleic acids and fragments thereof of the present invention can be altered from their native state in one or more of the following ways. First, a nucleic acid or fragment thereof which encodes an IV polypeptide or fragment, variant, or derivative thereof can be part or all of a codon-optimized coding region, optimized according to codon usage in the animal in which the vaccine is to be delivered. In addition, a nucleic acid or fragment thereof which encodes an IV polypeptide can be a fragment which encodes only a portion of a full-length polypeptide, and/or can be mutated so as to, for example, remove from the encoded polypeptide non-desired protein motifs present in the encoded polypeptide or virulence factors associated with the encoded polypeptide. For example, the nucleic acid sequence could be mutated so as to not encode a membrane anchoring region that would prevent release of the polypeptide from the cell as with, e.g., eM2. Upon delivery, the polynucleotide of the invention is incorporated into the cells of the vertebrate in vivo, and a prophylactically or therapeutically effective amount of an immunologic epitope of an IV is produced in vivo.

[0024] Similarly, the proteins of the invention can be a fragment of a full-length IV polypeptide and/or can be altered so as to, for example, remove from the polypeptide non-desired protein motifs present in the polypeptide or virulence factors associated with the polypeptide. For example, the polypeptide could be altered so as not to encode a membrane anchoring region that would prevent release of the polypeptide from the cell.

[0025] The invention further provides immunogenic compositions comprising at least one polynucleotide, wherein the polynucleotide comprises one or more nucleic acid fragments, where each nucleic acid fragment is a fragment of a codon-optimized coding region encoding an IV
polypeptide or a fragment, a variant, or a derivative thereof; and immunogenic compositions comprising a polynucleotide as described above and at least one isolated IV polypeptide or a fragment, a variant, or derivative thereof. Such compositions can further comprise, for example, carriers, excipients, transfection facilitating agents, and/or adjuvants as described herein.

[0026] The immunogenic compositions comprising a polynucleotide and an isolated IV polypeptide or fragment, variant, or derivative thereof as described above can be provided so that the polynucleotide and protein formulation are administered separately, for example, when the polynucleotide portion of the composition is administered prior (or subsequent) to the isolated IV polypeptide portion of the composition. Alternatively, immunogenic compositions comprising the polynucleotide and the isolated IV polypeptide or fragment, variant, or derivative thereof can be provided as a single formulation, comprising both the polynucleotide and the protein, for example, when the polynucleotide and the protein are administered simultaneously. In another alternative, the polynucleotide portion of the composition and the isolated IV polypeptide portion of the composition can be provided simultaneously, but in separate formulations.

[0027] Compositions comprising at least one polynucleotide comprising one or more nucleic acid fragments, where each nucleic acid fragment is optionally a fragment of a codon-optimized coding region operably encoding an IV polypeptide or fragment, variant, or derivative thereof together with one or more isolated IV polypeptides or fragments, variants or derivatives thereof (as either a recombinant protein, a purified subunit, a viral vector expressing the protein, or in the form of an inactivated or attenuated IV vaccine) will be referred to herein as “combinatorial polynucleotide (e.g., DNA) vaccine compositions” or “single formulation heterologous prime-boost vaccine compositions.”

[0028] The compositions of the invention can be univalent, bivalent, trivalent or multivalent. A univalent composition will comprise only one polynucleotide comprising a nucleic acid fragment, where the nucleic acid fragment is optionally a fragment of a codon-optimized coding region encoding an IV polypeptide or a fragment, variant, or derivative thereof, and optionally the same IV polypeptide or a fragment, variant, or derivative thereof in isolated form. In a single formulation heterologous prime-boost vaccine composition, a univalent composition can include a polynucleotide comprising a nucleic acid fragment, where the nucleic acid fragment is optionally a fragment of a codon-optimized coding region encoding an IV polypeptide or a fragment, variant, or derivative thereof and an isolated polypeptide having the same antigenic region as the polynucleotide. A bivalent composition will comprise, either in polynucleotide or protein form, two different IV polypeptides or fragments, variants, or derivatives thereof, each capable of eliciting an immune response. The polynucleotide(s) of the composition can encode two IV polypeptides or alternatively, the polynucleotide can encode only one IV polypeptide and the second IV polypeptide would be provided by an isolated IV polypeptide of the invention as in, for example, a single formulation heterologous prime-boost vaccine composition. In the case where both IV polypeptides of a bivalent composition are delivered in polynucleotide form, the nucleic acid fragments operably encoding those IV polypeptides need not be on the same polynucleotide, but can be on two different polynucleotides. A trivalent or further multivalent composition will comprise three IV polypeptides or fragments, variants or derivatives thereof, either in isolated form or encoded by one or more polynucleotides of the invention.

[0029] The present invention further provides plasmids and other polynucleotide constructs for delivery of nucleic acid fragments of the invention to a vertebrate, e.g., a human, which provide expression of IV polypeptides, or fragments, variants, or derivatives thereof. The present invention further provides carriers, excipients, transfection facilitating agents, immunogenicity-enhancing agents, e.g., adjuvants, or other agent or agents to enhance the transfection, expression or efficacy of the administered gene and its product.

[0030] In one embodiment, a multivalent composition comprises a single polynucleotide, e.g., plasmid, comprising one or more nucleic acid regions operably encoding IV polypeptides or fragments, variants, or derivatives thereof. Reducing the number of polynucleotides, e.g., plasmids in the compositions of the invention can have significant impacts on the manufacture and release of product, thereby reducing the costs associated with manufacturing the compositions. There are a number of approaches to include more than one expressed antigen coding sequence on a single plasmid. These include, for example, the use of Internal Ribosome Entry Site (IRES) sequences, dual promoters/expression cassettes, and fusion proteins.

[0031] The invention also provides methods for enhancing the immune response of a vertebrate to IV infection by administering to the tissues of a vertebrate one or more polynucleotides each comprising one or more nucleic acid fragments, where each nucleic acid fragment is optionally a fragment of a codon-optimized coding region encoding an IV polypeptide or fragment, variant, or derivative thereof, and optionally administering to the tissues of the vertebrate one or more isolated IV polypeptides, or fragments, variants, or derivatives thereof. The isolated IV polypeptide can be administered prior to, at the same time (simultaneously), or subsequent to administration of the polynucleotides encoding IV polypeptides.

[0032] In addition, the invention provides consensus amino acid sequences for IV polypeptides, or fragments, variants or derivatives thereof, including, but not limited to the HA, NA, NP, M1 or M2 proteins or fragments (e.g. eM2), variants or derivatives thereof. Polynucleotides which encode the consensus polypeptides or fragments, variants or derivatives thereof, are also embodied in this invention. Such polynucleotides can be obtained by known methods, for example by backtranslation of the amino acid sequence and PCR synthesis of the corresponding polynucleotide as described below.

BRIEF DESCRIPTION OF THE FIGURES

[0033] FIG. 1 shows an alignment of nucleotides 46-1542 of SEQ ID NO:1 (native NP coding region) with a coding region fully codon-optimized for human usage (SEQ ID NO:23).
FIG. 2 shows the protocol for the preparation of a formulation comprising 0.3 mM BAK, 7.5 mg/ml CRL 1005 and 5 mg/ml of DNA in a final volume of 3.6 ml, through the use of thermal cycling.

FIG. 3 shows the protocol for the preparation of a formulation comprising 0.3 mM BAK, 34 mg/ml or 50 mg/ml CRL 1005 and 2.5 mg/ml DNA in a final volume of 4.0 ml, through the use of thermal cycling.

FIG. 4 shows the protocol for the simplified preparation (without thermal cycling) of a formulation comprising 0.3 mM BAK, 7.5 mg/ml CRL 1005 and 5 mg/ml DNA.

FIG. 5 shows the anti-NP antibody response three weeks after a single administration of a combinatorial prime-boost vaccine formulation against the influenza virus NP protein.

FIG. 6 shows the anti-NP antibody response twelve days after a second administration of a combinatorial prime-boost vaccine formulation against the influenza virus NP protein.

FIG. 7 shows the CD8+ T Cell response to a combinatorial prime-boost vaccine formulation against the influenza virus NP protein.

FIG. 8 shows the CD4+ T Cell response to a combinatorial prime-boost vaccine formulation against the influenza virus NP protein.

FIGS. 9A and 9B show the results of a two dose mouse immunization regimen study with plasmid DNA encoding IAV HA (H3)

FIGS. 10A and 10B show the in vitro expression of M1 and M2 from segment 7 and an M1M2 fusion.

FIGS. 11A and 11B show the in vitro expression of cM2-NP and codon-optimized influenza virus NP protein.

FIG. 12 shows the influenza A NP protein consensus amino acid sequence aligned with 22 full length NP sequences.

FIG. 13 is a schematic diagram of various vectors encoding influenza proteins described herein.

FIG. 14 are the results of western blot experiments as described in Example 1, Experiment 3. The blots show lysates of VM92 cells transfected with plasmids which express M2 or NP to compare expression of the influenza protein from different expression vectors.

FIG. 15 are the results of western blot experiments as described in Example 3, Experiment 3. The blots show lysates of VM92 cells transfected with plasmids which express M1, M2 or NP to compare expression of the influenza protein from expression vectors.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to compositions and methods for enhancing the immune response of a vertebrate in need of protection against IV infection by administering in vivo, into a tissue of a vertebrate, at least one polynucleotide comprising one or more nucleic acid fragments, where each nucleic acid fragment is optionally a fragment of a codon-optimized coding region operably encoding an IV polypeptide, or a fragment, variant, or derivative thereof in cells of the vertebrate in need of protection. The present invention is also directed to administering in vivo, into a tissue of the vertebrate the above described polynucleotide and at least one isolated IV polypeptide, or a fragment, variant, or derivative thereof. The isolated IV polypeptide or fragment, variant, or derivative thereof can be, for example, a recombinant protein, a purified subunit protein, a protein-encoding RNA or a nucleic acid encoding a viral vector expressing the protein, or can be an inactivated IV, such as those present in conventional, commercially available, inactivated IV vaccines. According to either method, the polynucleotide is incorporated into the cells of the vertebrate in vivo, and an immunologically effective amount of the influenza protein, or fragment or variant encoded by the polynucleotide is produced in vivo. The isolated protein or fragment, variant, or derivative thereof is also administered in an immunologically effective amount. The polynucleotide can be administered to the vertebrate in need of that either prior, at the same time (simultaneously), or subsequent to the administration of the isolated IV polypeptide or fragment, variant, or derivative thereof.
[0050] The amino acid sequence of the NP protein of Influenza A/Puerto Rico/8/34 (H1N1), encoded by nucleotides 46-1494 of SEQ ID NO:1 is as follows, referred to herein as SEQ ID NO: 2:

```
NASQGTKEVQNETQCTGRQFATNHEVNEVIQGQFTQCTEKLS
```

[Dyergi, 1998]

[0053] The amino acid sequence of the M2 protein of Influenza A/Puerto Rico/8/34/Mount Sinai (H1N1), encoded (in spliced form) by nucleotides 26 to 51 and 740 to 1007 of SEQ ID NO:3 is as follows, referred to herein as SEQ ID NO:5:

```
NSLETTETVFLISIPGGPKKQAIQRLEDVPXNQNTLDEELMVXNLRT
```


[0055] eM2 is a derivative of NP and eM2 described herein is encoded by a construct which encodes the first 24 amino acids of M2 and all or a portion of NP. The fusion constructs may be constructed with the eM2 sequences followed by the NP sequences, or with the NP sequences followed by the eM2 sequences. Exemplary fusion constructs using the NP and M2 sequences from Influenza A/Puerto Rico/8/34/Mount Sinai (H1N1) are set out below. A sequence, using the original influenza virus nucleotide sequences, which encodes the first 24 amino acids of M2 fused at its 3' end to a sequence which encodes NP in its entirety eM2-NP is referred to herein as SEQ ID NO:6.
1 AGCGATCTCC TAAAGGAGCG GAAAGCGGCT TACGAAAGCG AATGGCGGCG TAGATGCGAC
2 GCAGAATGAT ATAGGCGCTT TCTAGGAACG GAAAGCGGCT TACGAAAGCG AATGGCGGCG
3 TACGGAAGCT CACGATAAGT GACGTGCTG TACGGAAGCT CACGATAAGT GACGTGCTG
4 ATACCAGGAC GAGGCGGCG TGGAGGAGT AAGAGAGGAC AAGAGAGGAC TGGAGGAGT
5 GCAAGGAGT TATAGCCTG GATAGGAGT AAGAGAGGAC AAGAGAGGAC TGGAGGAGT
6 ATATGTAAC TGGATTGCAT AGTACGTAGT ATATGTAAC TGGATTGCAT AGTACGTAGT
7 TATGTAAC TGGATTGCAT AGTACGTAGT ATATGTAAC TGGATTGCAT AGTACGTAGT
8 ATATGTAAC TGGATTGCAT AGTACGTAGT ATATGTAAC TGGATTGCAT AGTACGTAGT
9 ATATGTAAC TGGATTGCAT AGTACGTAGT ATATGTAAC TGGATTGCAT AGTACGTAGT
10 ATATGTAAC TGGATTGCAT AGTACGTAGT ATATGTAAC TGGATTGCAT AGTACGTAGT
11 ATATGTAAC TGGATTGCAT AGTACGTAGT ATATGTAAC TGGATTGCAT AGTACGTAGT
12 ATATGTAAC TGGATTGCAT AGTACGTAGT ATATGTAAC TGGATTGCAT AGTACGTAGT
13 ATATGTAAC TGGATTGCAT AGTACGTAGT ATATGTAAC TGGATTGCAT AGTACGTAGT
14 ATATGTAAC TGGATTGCAT AGTACGTAGT ATATGTAAC TGGATTGCAT AGTACGTAGT
15 ATATGTAAC TGGATTGCAT AGTACGTAGT ATATGTAAC TGGATTGCAT AGTACGTAGT

[0056] The amino acid sequence of the eM2-NP fusion protein of Influenza A/PR/8/341 (H1N1), encoded by nucleotides 1 to 1566 SEQ ID NO:6 is as follows, referred to herein as SEQ ID NO:7 (eM2 amino acid sequence underlined):

-continued

[0057] A sequence, using the original influenza virus nucleotide sequences, which encodes NP in its entirety fused at its 3' end to the first 24 amino acids of M2 fused to a sequence which encodes NP in its entirety is referred to herein as SEQ ID NO:8:

AGACAGACAAAGCCCAAGAACATGAAAGAGTGACGACGTATG AGACAGACAAAGCCCAAGAACATGAAAGAGTGACGACGTATG

STLRSRHWRAHRGSGGTHQREAQOSIQTPSFVGSNSLHLPFORTTYN
AAGFOGHTGRTDMGFIEIIMSAADDTEFVDFSGHGFSDKDEAAASPIV
PSFDMSNGSGFFGNGAEEYDN

...
[0059] The construction of functional fusion proteins often requires a linker sequence between the two fused fragments, in order to adopt an extended conformation to allow maximal flexibility. We used program LINKER (C. J. Croato and Feng, J. Protein Engineering 13:309-312 (2000)), which can automatically produce a set of linker sequences, which are known to adopt extended conformations as determined by X-ray crystallography and NMR. Examples of suitable linkers to use in various cM2-NP or NP-cM2 fusion proteins are as follows:

1. GHTRFA (SEQ ID NO:10)
2. FQNSFT (SEQ ID NO:11)
3. FSVNLK (SEQ ID NO:12)
4. GTSNGTVT (SEQ ID NO:13)
5. VNEKTIDSD (SEQ ID NO:14)

[0060] The nucleotide sequence of the NP protein of Influenza B/LEE/40 is available as GenBank Accession Number K01395, and has the following sequence, referred to here as SEQ ID NO:15:

1 ATGCTCAAGC TTCAGTAA TGAAGAAG GTTCTTTG GACACATGC AAACAGAAG

61 CGAAGCGCG GGCGACAG CAGCGGTTT TAGAAGCGCG

121 ACCAGCGG GCCCGTGAAT CGCGTGAAGCT

181 GATGCTGG AAGCGACG GGCGACAG CAGCGGTTT

241 AAAGCGTA GTTCTTTG GACACATGC AAACAGAAG
-continued

301 GATGACATCG AAAGAAACTT AATTCGAAAT GCACAAGCTG TGAGAAGAT CCACTGCGT
306 GCAGCTGAGT ACAAGAACAC TGAATACGAA AAGAAGAGAA AGTCCAGAGA TGCTCAAAGA
311 CGGAGAAGAGA AATAGAAGCA GCAACAGACA GGGAGGAGGT TTTATAGAT GGTAGAAGAT
316 GATGAAGCAA TTCTACTTCG GGTATTAAA AATCCTTCTT TAAAGAGGAG GGGAGAAGAT
321 AGTATCGAAC CCAAGCCATGG GATGTAAAGTG TCTAGTGGAC TAAATCATCA TATATTAGCA
326 CATTCGAGGA GCAAGCTGAT CTGTTTCCAAA AAGTAAAGGG GACTAAAAGG GTTGAGAATCT
331 GACCCCTCAT TAATGACAT TTTGCAAGGA AGGACAGTAC CCAGAGAGAC AGCTACACCT
336 GGCCTGTACCA CCAAGAAGAGA AGTCTCAGCT CTGAGGAGAG TGTACAGCTG
341 CTCCGAGGCT TATGGAAAA CCGCAGGCAG AGCTCTAGGT TGACATAGTG
346 ATGCACATGA GGGAGAGGAGA GCTGAGAAGAT GGTGGAGAGT GGAGTAAGAG
351 AGAGAGATGA TGAGAGAAAGG TGAGAGAAGG TGGAGAGAAGG AGAGAGAAGG
356 TAATAGAAATT CGGGCCAGCC GACACAGAAG CTCACTCTAT GCTGCGCTCA
361 1AAATATGCTC AATAGAAGCA GCAACAGACA GGGAGGAGGT TTTATAGAT GGTAGAAGAT
366 GATGAAGCAA TTCTACTTCG GGTATTAAA AATCCTTCTT TAAAGAGGAG GGGAGAAGAT
371 AGTATCGAAC CCAAGCCATGG GATGTAAAGTG TCTAGTGGAC TAAATCATCA TATATTAGCA
376 CATTCGAGGA GCAAGCTGAT CTGTTTCCAAA AAGTAAAGGG GACTAAAAGG GTTGAGAATCT
381 GACCCCTCAT TAATGACAT TTTGCAAGGA AGGACAGTAC CCAGAGAGAC AGCTACACCT
386 GGCCTGTACCA CCAAGAAGAGA AGTCTCAGCT CTGAGGAGAG TGTACAGCTG
391 CTCCGAGGCT TATGGAAAA CCGCAGGCAG AGCTCTAGGT TGACATAGTG
396 ATGCACATGA GGGAGAGGAGA GCTGAGAAGAT GGTGGAGAGT GGAGTAAGAG
401 AGAGAGATGA TGAGAGAAAGG TGAGAGAAGG TGGAGAGAAGG AGAGAGAAGG
406 TAATAGAAATT CGGGCCAGCC GACACAGAAG CTCACTCTAT GCTGCGCTCA

[0061] The amino acid sequence of the NP protein of IBV B/LEE/40, encoded by nucleotides 1-1680 of SEQ ID NO:1 is as follows, referred to herein as SEQ ID NO:16:

MSHMDILEPDTVDKPEHLPFGATFIPKIFLATLAFSNERHEPE
FERTTSGESEDTQIKQEREPEQPEIPTMKWYMVYPEGFMVQVYKAGL
DCIMRSLNQAGVARRLLAATDQKTEYQKEESEQAGYVKEKEEIKHKT
GGFTYKWRQDIYIPSIKIEFLKEEKVMTYKETM5EDGFPSLGHLNIMIG
HSQNSDQVPIQKEKKLVQGQDLIESFTAGSLPERSGQTIVVAKSGGTL
VDE11FRGAMRDRLDIAXTAYKILNKLSCSPQQELIQDVQ
IGRRMFQ1ADIEEUDDLAKGTVAKVRFPSVAVXKKLYSLFISIYAKTIFQLOQENTE
EYSNQYEMALNAYMPVS1RLKQDAAKSKQFLPMHGCAGAIEELSLVL
SALYTPIFPRSLKCGPQVHVPAXEVEGSAALSIKQFWVFMTSUG
NYE52EG6GG0QCSPYFAYERPAFLSSEQVQRVRLMSINVSEDEAVKGNL

[0062] Non limiting examples of nucleotide sequences encoding the IAV hemagglutinin (HA) are as follows. It should be noted that HA sequences vary significantly between IV subtypes. Virtually any nucleotide sequence encoding an IV HA is suitable for the present invention. In fact, HA sequences included in vaccines and therapeutic formulations of the present invention (discussed in more detail below) might change from year to year depending on the prevalent strain or strains of IV.

[0063] The partial nucleotide sequence of the HA protein of IAV A/New_York/1/18(H1N1) is available as GenBank Accession Number AF116576, and has the following sequence, referred to herein as SEQ ID NO: 17:
1 atggagggca gactacttgt cttgtaatgt gcatctgcaag ctacaacagt cagcaacaata 
61 tgtataggt taccagcga ttacactaacc gacacttgtt acacagtaat cgaacgaaat 
121 tgttacgga caacacttct tcacgcttcgccacagcgacagcgacgcacagcggaa aatcctgaat 
181 ttaaagacaa taccccaaaacc gacattggg gaataagtaata taagcagggg cttacatctt 
241 aacccaggaat gacctactactt tccacaccaagg cagccaccaagt acaccaaatc 
301 acacacagga atggacacagt taacccagga aggattacagc acacccgagc 
361 caactcgagca tgtgtaatcgc gcacacacagc ttcagatttt ctaacacgagc 
421 cccacacagcg aacacagca aggcagcagc acacacagcg atcgagcaagc cagcaacaatc 
481 tttttacagca atggatgctc gtgtacagcgac aaggaagcaag ctctacaggc gagcagcagc 
541 tacatgtgcc accaccaacct gagacagct cttggccgtc acaccaaccttg cttgatagct 
601 acacactgtac atacaccaaat ctaatagacagcctcgaagctcgtc ttagggctata 
661 atacacatagc accaactgc gacaagcgca atacagcgagc gaccaacacagc aagagtgcaaa 
721 ggtgagcgaga aacatcatactt atcagacatta ctaacacacagc 
781 gcacacacagc accaactgtac ggtgctcac gtcagactccc tgtgtctgagtctgatgg 
841 gcacccggatc acgcgcgtctc gtgcggatcccctgctggtgagc gtcagactccc 
901 ggtgtactact accagcagc aacagcagcgtcagtctgctggtgagc gttcgtgcgtc 
961 cccacaaggat cccacaaggat cccacaaggat cccacaaggat cccacaaggat 
1021 atcacaaccagc accttggaggct cggcttggagct cggcttggagct cggcttggagct 
1081 atggcagagc atggactgca tgcagcagc accttggagct 
1141 gaccaacacagc accaactgc gacaacagcata accaactgc gacaacagcata 
1201 gagactacag ccacacagc 

[0064] The amino acid sequence of the partial HA protein of IAV A/New York/1/18 (H1N1), encoded by nucleotides 1 to 1218 of SEQ ID NO:17 as follows, referred to herein as SEQ ID NO:18:

MEAKLLVLCACAAAPNADTICGHYANNSTDDYTVLEEXNVTYVSVNLL 
EDSHNRKLCRKLGIAPLQGLRCHNIAWGNNLPPCCOLLITASSWSYIVETS 
NSEDTPGPDLYELREQLLLLVS4PPEKEFIPKLSWPNHITGTGT 
AACEYAGAAGFYPNNLWNLTKGSSYTPKLSVYNNRAGKEVLVLACUVHPT 
GTDQSSLQNADAVSVGSXKHPFRTE1AARFXVBQQAGNmNYPWILL

[0065] The nucleotide sequence of the IAV A/Hong Kong/482.97 hemagglutinin (H5) is available as GenBank Accession Number AF046098, and has the following sequence, referred to herein as SEQ ID NO:19:

1 cctcggagat ggagcgataa gctggtcctc tgcaacagt cagctgctgtt cctgctgatc 
61 aaggttacct atgtattaact gcaacacact gacgagcaga ggctgcgacagcgacagcgacagcgacgcacagcggaa aatcctgaat 
121 aaggttacct atgtattaact gcaacacact gacgagcaga ggctgcgacagcgacagcgacagcgacgcacagcggaa aatcctgaat 
181 aaggttacct atgtattaact gcaacacact gacgagcaga ggctgcgacagcgacagcgacagcgacgcacagcggaa aatcctgaat 
241 aaggttacct atgtattaact gcaacacact gacgagcaga ggctgcgacagcgacagcgacagcgacgcacagcggaa aatcctgaat 
301 aaggttacct atgtattaact gcaacacact gacgagcaga ggctgcgacagcgacagcgacagcgacgcacagcggaa aatcctgaat

[0064] -continued

XPDBTITFEATGKLIAPWYAFALRGGSGEECIPYSSAPVHDCNHFRCFYPPHS 
AANSNLPSQRIFPVTIGECFKYRTFELMATGLHRHPSIQERSLGPAIA 
GFPSSGGWGTNIOWGTYSEHOSEKGGGYAAADQRKQSTQNAIGITNKNVSVIE 
KMMFQ

[0065] -continued

XPDBTITFEATGKLIAPWYAFALRGGSGEECIPYSSAPVHDCNHFRCFYPPHS 
AANSNLPSQRIFPVTIGECFKYRTFELMATGLHRHPSIQERSLGPAIA 
GFPSSGGWGTNIOWGTYSEHOSEKGGGYAAADQRKQSTQNAIGITNKNVSVIE 
KMMFQ
The amino acid sequence of the HA protein of IAV A/Hong Kong/482/97 (H5), encoded by nucleotides 9 to 1715 of SEQ ID NO: 20, is as follows, referred to herein as SEQ ID NO: 21:

```
AISGEISEGWQDNQVYYSYHSNQGQDAERKEKTSQGKIDNVE
IHHNHTQPEAVGREGFKMLERAOELMLRNMEDGFLUVWYHAEELVLH
ERTLOPHJUSVKNYVCIQRLLHARAKLGQUALGEFCYHKDHECNESVK
GVEPQFYQEERAINRKEEISGYKLSMVGQILS1ISTVASSLALAINAV
```

The nucleotide sequence of the IAV A/Hong Kong/1073-99(H9N2) is available as GenBank Accession Number INA04626, and has the following sequence, referred to herein as SEQ ID NO: 21:

```
1 gcaaagcag gggaaataa taatacagaa aatggaacaa atataatcag aactaatcctg
```
[0068] The amino acid sequence of the HA protein of IAV A/Hong Kong/1073/99 (H9N2), encoded by nucleotides 32 to 1711 of SEQ ID NO:21 is as follows, referred to herein as SEQ ID NO:22:

```
[0689] The present invention also provides vaccine compositions and methods for delivery of IV coding sequences to a vertebrate with optimal expression and safety conferred through codon optimization and/or other manipulations. These vaccine compositions are prepared and administered in such a manner that the encoded gene products are optimally expressed in the vertebrate of interest. As a result,
```
these compositions and methods are useful in stimulating an immune response against IV infection. Also included in the invention are expression systems, delivery systems, and codon-optimized IV coding regions.

[0070] In a specific embodiment, the invention provides combinatorial polynucleotide (e.g., DNA) vaccines which combine both a polynucleotide vaccine and polypeptide (e.g., either a recombinant protein, a purified subunit protein, a viral vector expressing an isolated IV polypeptide, or in the form of an inactivated or attenuated IV vaccine) vaccine in a single formulation. The single formulation comprises an IV polypeptide-encoding polynucleotide vaccine as described herein, and optionally, an effective amount of a desired isolated IV polypeptide or fragment, variant, or derivative thereof. The polypeptide may exist in any form, for example, a recombinant protein, a purified subunit protein, a viral vector expressing an isolated IV polypeptide, or in the form of an inactivated or attenuated IV vaccine. The IV polypeptide or fragment, variant, or derivative thereof encoded by the polynucleotide vaccine may be identical to the isolated IV polypeptide or fragment, variant, or derivative thereof. Alternatively, the IV polypeptide or fragment, variant, or derivative thereof encoded by the polynucleotide may be different from the isolated IV polypeptide or fragment, variant, or derivative thereof.

[0071] It is to be noted that the term “a” or “an” entity refers to one or more of that entity; for example, “a polynucleotide” is understood to represent one or more polynucleotides. As such, the terms “a” or “an”, “one or more,” and “at least one” can be used interchangeably herein.

[0072] The term “polynucleotide” is intended to encompass a singular nucleic acid or nucleic acid fragment as well as plural nucleic acids or nucleic acid fragments, and refers to an isolated molecule or construct, e.g., a virus genome (e.g., a non-infectious viral genome), messenger RNA (mRNA), plasmid DNA (pDNA), or derivatives of pDNA (e.g., minicircles as described in Darquet, A-M et al., Gene Therapy 4:1341-1349 (1997)) comprising a polynucleotide. A polynucleotide may comprise a conventional phosphodiester bond or a non-conventional bond (e.g., an amide bond, such as found in peptide nucleic acids (PNA)).

[0073] The term “nucleic acid” or “nucleic acid fragment” refer to any one or more nucleic acid segments, e.g., DNA or RNA fragments, present in a polynucleotide or construct. A nucleic acid or fragment thereof may be provided in linear (e.g., mRNA) or circular (e.g., plasmid) form as well as double-stranded or single-stranded forms. By “isolated” nucleic acid or polynucleotide is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment. For example, a recombinant polynucleotide contained in a vector is considered isolated for the purposes of the present invention. Further examples of an isolated polynucleotide include recombinant polynucleotides maintained in heterologous host cells or purified (partially or substantially) polynucleotides in solution. Isolated RNA molecules include in vivo or in vitro transcripts of the polynucleotides of the present invention. Isolated polynucleotides or nucleic acids according to the present invention further include such molecules produced synthetically.

[0074] As used herein, a “coding region” is a portion of nucleic acid which consists of codons translated into amino acids. Although a “stop codon” (TAG, TGA, or TAA) is not translated into an amino acid, it may be considered to be part of a coding region, but any flanking sequences, for example promoters, ribosome binding sites, transcriptional terminators, and the like, are not part of a coding region. Two or more nucleic acids or nucleic acid fragments of the present invention can be present in a single polynucleotide construct, e.g., on a single plasmid, or in separate polynucleotide constructs, e.g., on separate (different) plasmids. Furthermore, any nucleic acid or nucleic acid fragment may encode a single IV polypeptide or fragment, derivative, or variant thereof, e.g., or may encode more than one polypeptide, e.g., a nucleic acid may encode two or more polypeptides. In addition, a nucleic acid may encode a regulatory element such as a promoter, ribosome binding site, or a transcription terminator, or may encode heterologous coding regions fused to the IV coding region, e.g., specialized elements or motifs, such as a secretory signal peptide or a heterologous functional domain.

[0075] The terms “fragment,” “variant,” “derivative” and “analog” when referring to IV polypeptides of the present invention include any polypeptides which retain at least some of the immunogenicity or antigenicity of the corresponding native polypeptide. Fragments of IV polypeptides of the present invention include proteolytic fragments, deletion fragments and in particular, fragments of IV polypeptides which exhibit increased secretion from the cell or higher immunogenicity or reduced pathogenicity when delivered to an animal. Polypeptide fragments further include any portion of the polypeptide which comprises an antigenic or immunogenic epitope of the native polypeptide, including linear as well as three-dimensional epitopes. Variants of IV polypeptides of the present invention include fragments as described above, and also polypeptides with altered amino acid sequences due to amino acid substitutions, deletions, or insertions. Variants may occur naturally, such as an allelic variant. By an “allelic variant” is intended alternate forms of a gene occurring a given locus on a chromosomal sequence of an organism or virus. Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985), which is incorporated herein by reference. For example, as used herein, variations in a given gene product. When referring to IV NA or HA proteins, each such protein is a “variant” in that native IV strains are distinguished by the type of NA and HA proteins encoded by the virus. However, within a single NA or HA variant type, further naturally or non-naturally occurring variations such as amino acid deletions, insertions or substitutions may occur. Non-naturally occurring variants may be produced using art-known mutagenesis techniques. Variant polypeptides may comprise conservative or non-conservative amino acid substitutions, deletions or additions. Derivatives of IV polypeptides of the present invention, are polypeptides which have been altered so as to exhibit additional features not found on the native polypeptide. Examples include fusion proteins. An analog is another form of an IV polypeptide of the present invention. An example is a proprotein which can be activated by cleavage of the proprotein to produce an active mature polypeptide.

[0076] The terms “infectious polynucleotide” or “infectious nucleic acid” are intended to encompass isolated viral polynucleotides and/or nucleic acids which are solely sufficient to mediate the synthesis of complete infectious virus particles upon uptake by permissive cells. Thus, “infectious nucleic acids” do not require pre-synthesized copies of any
of the polypeptides it encodes, e.g., viral replicases, in order to initiate its replication cycle in a permissive host cell.

[0077] The terms “non-infectious polynucleotide” or “non-infectious nucleic acid” as defined herein are polynucleotides or nucleic acids which cannot, without additional materials, e.g., polypeptides, mediate the synthesis of complete infectious virus particles upon uptake by permissive cells. An infectious polynucleotide or nucleic acid is not made “non-infectious” simply because it is taken up by a non-permissive cell. For example, an infectious viral polynucleotide from a virus with limited host range infectious if it is capable of mediating the synthesis of complete infectious virus particles when taken up by cells derived from a permissive host (i.e., a host permissive for the virus itself). The fact that uptake by cells derived from a non-permissive host does not result in the synthesis of complete infectious virus particles does not make the nucleic acid “non-infectious.” In other words, the term is not qualified by the nature of the host cell, the tissue type, or the species taking up the polynucleotide or nucleic acid fragment.

[0078] In some cases, an isolated infectious polynucleotide or nucleic acid may produce fully-infectious virus particles in a host cell population which lacks receptors for the virus particles, i.e., is non-permissive for virus entry. Thus viruses produced will not infect surrounding cells. However, if the supernatant containing the virus particles is transferred to cells which are permissive for the virus, infection will take place.

[0079] The terms “replicating polynucleotide” or “replicating nucleic acid” are meant to encompass those polynucleotides and/or nucleic acids which, upon being taken up by a permissive host cell, are capable of producing multiple, e.g., one or more copies of the same polynucleotide or nucleic acid. Infectious polynucleotides and nucleic acids are a subset of replicating polynucleotides and nucleic acids; the terms are not synonymous. For example, a defective virus genome lacking the genes for virus coat proteins may replicate, e.g., produce multiple copies of itself, but is NOT infectious because it is incapable of mediating the synthesis of complete infectious virus particles unless the coat proteins, or another nucleic acid encoding the coat proteins, are exogenously provided.

[0080] In certain embodiments, the polynucleotide, nucleic acid, or nucleic acid fragment is DNA. In the case of DNA, a polynucleotide comprising a nucleic acid which encodes a polypeptide normally also comprises a promoter and/or other transcription or translation control elements operably associated with the polypeptide-encoding nucleic acid fragment. An operable association is when a nucleic acid fragment encoding a gene product, e.g., a polypeptide, is associated with one or more regulatory sequences in such a way as to place expression of the gene product under the influence of control of the regulatory sequence(s). Two DNA fragments (such as a polypeptide-encoding nucleic acid fragment and a promoter associated with the 5' end of the nucleic acid fragment) are “operably associated” if induction of promoter function results in the transcription of mRNA encoding the desired gene product and if the nature of the linkage between the two DNA fragments does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the expression regulatory sequences to direct the expression of the gene product, or (3) interfere with the ability of the DNA template to be transcribed. Thus, a promoter region would be operably associated with a nucleic acid fragment encoding a polypeptide if the promoter was capable of effecting transcription of that nucleic acid fragment. The promoter may be a cell-specific promoter that directs substantial transcription of the DNA only in predetermined cells. Other transcription control elements, besides a promoter, for example enhancers, operators, repressors, and transcription termination signals, can be operably associated with the polynucleotide to direct cell-specific transcription. Suitable promoters and other transcription control regions are disclosed herein.

[0081] A variety of transcription control regions are known to those skilled in the art. These include, without limitation, transcription control regions which function in vertebrate cells, such as, but not limited to, promoter and enhancer segments from cytomegaloviruses (the immediate early promoter, in conjunction with intron-A), simian virus 40 (the early promoter), and retroviruses (such as Rous sarcoma virus). Other transcription control regions include those derived from vertebrate genes such as actin, heat shock protein, bovine growth hormone and rabbit IgG, as well as other sequences capable of controlling gene expression in eukaryotic cells. Additional suitable transcription control regions include tissue-specific promoters and enhancers as well as lymphokine-inducible promoters (e.g., promoters inducible by interferons or interleukins).

[0082] Similarly, a variety of translation control elements are known to those of ordinary skill in the art. These include, but are not limited to ribosome binding sites, translation initiation and termination codons, elements from picornaviruses (particularly an internal ribosome entry site, or IRES, also referred to as a CITE sequence).

[0083] A DNA polynucleotide of the present invention may be a circular or linearized plasmid or vector, or other linear DNA which may also be non-infectious and nonintegrating (i.e., does not integrate into the genome of vertebrate cells). A linearized plasmid is a plasmid that was previously circular but has been linearized, for example, by digestion with a restriction endonuclease. Linear DNA may be advantageous in certain situations as discussed, e.g., in Cheng, J. Y., et al., J. Control. Release 60:343-53 (1999), and Chen, Z. Y., et al. Mol. Ther. 3:403-10 (2001), both of which are incorporated herein by reference. As used herein, the terms plasmid and vector can be used interchangeably.

[0084] Alternatively, DNA virus genomes may be used to administer DNA polynucleotides into vertebrate cells. In certain embodiments, a DNA virus genome of the present invention is nonreplicative, noninfectious, and/or nonintegrating. Suitable DNA virus genomes include without limitation, herpesvirus genomes, adenovirus genomes, adeno-associated virus genomes, and poxvirun genomes. References citing methods for the in vivo introduction of non-infectious virus genomes to vertebrate tissues are well known to those of ordinary skill in the art, and are cited supra.

[0085] In other embodiments, a polynucleotide of the present invention is RNA, for example, in the form of messenger RNA (mRNA). Methods for introducing RNA sequences into vertebrate cells are described in U.S. Pat. No. 5,580,859, the disclosure of which is incorporated herein by reference in its entirety.
[0086] Polynucleotides, nucleic acids, and nucleic acid fragments of the present invention may be associated with additional nucleic acids which encode secretory or signal peptides, which direct the secretion of a polypeptide encoded by a nucleic acid fragment or polynucleotide of the present invention. According to the signal hypothesis, proteins secreted by mammalian cells have a signal peptide or secretory leader sequence which is cleaved from the mature protein once export of the growing protein chain across the rough endoplasmic reticulum has been initiated. Those of ordinary skill in the art are aware that polypeptides secreted by vertebrate cells generally have a signal peptide fused to the N-terminus of the polypeptide, which is cleaved from the complete or “full length” polypeptide to produce a secreted or “mature” form of the polypeptide. In certain embodiments, the native leader sequence is used, or a functional derivative of that sequence that retains the ability to direct the secretion of the polypeptide that is operably associated with it. Alternatively, a heterologous mammalian leader sequence, or a functional derivative thereof, may be used. For example, the wild-type leader sequence may be substituted with the leader sequence of human tissue plasminogen activator (TPA) or mouse β-glucuronidase.

[0087] In accordance with one aspect of the present invention, there is provided a polynucleotide construct, for example, a plasmid, comprising a nucleic acid fragment, where the nucleic acid fragment is a fragment of a codon-optimized coding region operably encoding an IV-derived polypeptide, where the coding region is optimized for expression in vertebrate cells, of a desired vertebrate species, e.g., humans, to be delivered to a vertebrate to be treated or immunized. Suitable IV polypeptides, or fragments, variants, or derivatives thereof may be derived from, but are not limited to, the IV HA, NA, NP, M1, or M2 proteins. Additional IV-derived coding sequences, e.g., coding for HA, NA, NP, M1, M2 or EM2, may also be included on the plasmid, or on a separate plasmid, and expressed, either using native IV codons or codons optimized for expression in the vertebrate to be treated or immunized. When such a plasmid encoding one or more optimized influenza sequences is delivered, in vivo to a tissue of the vertebrate to be treated or immunized, one or more of the encoded gene products will be expressed, i.e., transcribed and translated. The level of expression of the gene product(s) will depend to a significant extent on the strength of the associated promoter and the presence and activation of an associated enhancer element, as well as the degree of optimization of the coding region.

[0088] As used herein, the term “plasmid” refers to a construct made up of genetic material (i.e., nucleic acids). Typically a plasmid contains an origin of replication which is functional in bacterial host cells, e.g., Escherichia coli, and selectable markers for detecting bacterial host cells comprising the plasmid. Plasmids of the present invention may include genetic elements as described herein arranged such that an inserted coding sequence can be transcribed and translated in eukaryotic cells. Also, the plasmid may include a sequence from a viral nucleic acid. However, such viral sequences normally are not sufficient to direct or allow the incorporation of the plasmid into a viral particle, and the plasmid is therefore a non-viral vector. In certain embodiments described herein, a plasmid is a closed circular DNA molecule.

[0089] The term “expression” refers to the biological production of a product encoded by a coding sequence. In most cases a DNA sequence, including the coding sequence, is transcribed to form a messenger-RNA (mRNA). The messenger-RNA is then translated to form a polypeptide product which has a relevant biological activity. Also, the process of expression may involve further processing steps to the RNA product of transcription, such as splicing to remove introns, and/or post-translational processing of a polypeptide product.

[0090] As used herein, the term “polypeptide” is intended to encompass a singular “polypeptide” as well as plural “polypeptides,” and comprises any chain or chains of two or more amino acids. Thus, as used herein, terms including, but not limited to “peptide,” “dipeptide,” “tripeptide,” “protein,” “amino acid chain,” or any other term used to refer to a chain or chains of two or more amino acids, are included in the definition of a “polypeptide,” and the term “polypeptide” can be used instead of, or interchangeably with any of these terms. The term further includes polypeptides which have undergone post-translational modifications, for example, glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, or modification by non-naturally occurring amino acids.

[0091] Also included as polypeptides of the present invention are fragments, derivatives, analogs, or variants of the foregoing polypeptides, and any combination thereof. Polypeptides, and fragments, derivatives, analogs or variants thereof of the present invention can be antigenic and immunogenic polypeptides related to IV polypeptides, which are used to prevent or treat, i.e., cure, ameliorate, lessen the severity of, or prevent or reduce contagion of infectious disease caused by the IV.

[0092] As used herein, an “antigenic polypeptide” or an “immunogenic polypeptide” is a polypeptide which, when introduced into a vertebrate, reacts with the vertebrate’s immune system molecules, i.e., is antigenic, and/or induces an immune response in the vertebrate, i.e., is immunogenic. It is quite likely that an immunogenic polypeptide will also be antigenic, but an antigenic polypeptide, because of its size or conformation, may not necessarily be immunogenic. Examples of antigenic and immunogenic polypeptides of the present invention include, but are not limited to, e.g., HA or fragments or variants thereof, e.g., NP, or fragments thereof, e.g., P1, or fragments or variants thereof, e.g., NS1 or fragments or variants thereof, e.g., M1 or fragments or variants thereof, and e.g., M2 or fragments or variants thereof including the extracellular fragment of M2 (eM2), or e.g., any of the foregoing polypeptides or fragments fused to a heterologous polypeptide, for example, a hepatitis B core antigen. Isolated antigenic and immunogenic polypeptides of the present invention in addition to those encoded by polynucleotides of the invention, may be provided as a recombinant protein, a purified subunit, a viral vector expressing the protein, or may be provided in the form of an inactivated IV vaccine, e.g., a live-attenuated virus vaccine, a heat-killed virus vaccine, etc.

[0093] By an “isolated” IV polypeptide or a fragment, variant, or derivative thereof is intended an IV polypeptide or protein that is not in its natural form. No particular level of purification is required. For example, an isolated IV
The term "epitopes," as used herein, refers to portions of a polypeptide having antigenic or immunogenic activity in a vertebrate, for example a human. An "immunogenic epitope," as used herein, is defined as a portion of a protein that elicits an immune response in an animal, as determined by any method known in the art. The term "antigenic epitope," as used herein, is defined as a portion of a protein to which an antibody or T-cell receptor can immunospecifically bind as determined by any method well known in the art. Immunospecific binding excludes non-specific binding but does not exclude cross-reactivity with other antigens. Where all immunogenic epitopes are antigenic, antigenic epitopes need not be immunogenic.

The term "immunogenic carrier" as used herein refers to a first polypeptide or fragment, variant, or derivative thereof which enhances the immunogenicity of a second polypeptide or fragment, variant, or derivative thereof. Typically, an "immunogenic carrier" is fused to or conjugated to the desired polypeptide or fragment thereof. An example of an "immunogenic carrier" is a recombinant hepatitis B core antigen expressing, as a surface epitope, an immunogenic epitope of interest. See, e.g., European Patent No. EP 0385610 B 1, which is incorporated herein by reference in its entirety.

In the present invention, antigenic epitopes preferably contain a sequence of at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, at least 25, or between about 8 to about 30 amino acids contained within the amino acid sequence of an IV polypeptide of the invention, e.g., an NP polypeptide, an M1 polypeptide or an M2 polypeptide. Certain polypeptides comprising immunogenic or antigenic epitopes are at least 5, 10, 15, 20, 25, 30, 35, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 amino acid residues in length. Antigenic as well as immunogenic epitopes may be linear, i.e., be comprised of contiguous amino acids in a polypeptide, or may be three dimensional, i.e., where an epitope is comprised of non-contiguous amino acids which come together due to the secondary or tertiary structure of the polypeptide, thereby forming an epitope.

As to the selection of peptides or polypeptides bearing an antigenic epitope (e.g., that contain a region of a protein molecule to which an antibody or T cell receptor can bind), it is well known that in that art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein. See, e.g., Sutcliffe, J. G., et al., Science 219:660-666 (1983), which is herein incorporated by reference.

Peptides capable of elicits an immunogenic response are frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are confined neither to immunodominant regions of intact proteins nor to the amino or carboxyl terminals. Peptides that are extremely hydrophobic and those of six or fewer residues generally are ineffective in inducing antibodies that bind to the mimicked protein; longer peptides, especially those containing proline residues, usually are effective. Sutcliffe et al., supra, at 661. For instance, 18 of 20 peptides designed according to these guidelines, containing 8-39 residues covering 75% of the sequence of the IV hemagglutinin HAI polypeptide chain, induced antibodies that reacted with the HAI protein or intact virus; and 12/12 peptides from the MuLV polymerase and 18/18 from the rabies glycoprotein induced antibodies that precipitated the respective proteins.

Codon Optimization

"Codon optimization" is defined as modifying a nucleic acid sequence for enhanced expression in the cells of the vertebrate of interest, e.g., human, by replacing at least one, more than one, or a significant number, of codons of the native sequence with codons that are more frequently or most frequently used in the genes of that vertebrate. Various species exhibit particular bias for certain codons of a particular amino acid.

In one aspect, the present invention relates to polynucleotides comprising nucleic acid fragments of codon-optimized coding regions which encode IV polypeptides, or fragments, variants, or derivatives thereof, with the codon usage adapted for optimized expression in the cells of a given vertebrate, e.g., humans. These polynucleotides are prepared by incorporating codons preferred for use in the genes of the vertebrate of interest into the DNA sequence. Also provided are polynucleotide expression constructs, vectors, and host cells comprising nucleic acid fragments of codon-optimized coding regions which encode IV polypeptides, and fragments, variants, or derivatives thereof, and various methods of using the polynucleotide expression constructs, vectors, host cells to treat or prevent influenza disease in a vertebrate.

As used herein the term "codon-optimized coding region" means a nucleic acid coding region that has been adapted for expression in the cells of a given vertebrate by replacing at least one, or more than one, or a significant number, of codons with one or more codons that are more frequently used in the genes of that vertebrate.

Deviations in the nucleotide sequence that comprise the codons encoding the amino acids of any polypeptide chain allow for variations in the sequence coding for the gene. Since each codon consists of three nucleotides, and the nucleotides comprising DNA are restricted to four specific bases, there are 64 possible combinations of nucleotides, 61 of which encode amino acids (the remaining three codons encode signals ending translation). The "genetic code" which shows which codons encode which amino acids is reproduced herein as Table 1. As a result, many amino acids are designated by more than one codon. For example, the
amino acids alanine and proline are coded for by four triplets, serine and arginine by six, whereas tryptophan and methionine are coded by just one triplet. This degeneracy allows for DNA base composition to vary over a wide range without altering the amino acid sequence of the proteins encoded by the DNA.

### TABLE 1

<table>
<thead>
<tr>
<th>Codon</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTT</td>
<td>Phe</td>
</tr>
<tr>
<td>TAT</td>
<td>Tyr</td>
</tr>
<tr>
<td>TAC</td>
<td>Tyr</td>
</tr>
<tr>
<td>TCA</td>
<td>Ser</td>
</tr>
<tr>
<td>TCT</td>
<td>Ser</td>
</tr>
<tr>
<td>TGA</td>
<td>Val</td>
</tr>
<tr>
<td>TGG</td>
<td>Trp</td>
</tr>
<tr>
<td>CTT</td>
<td>Leu</td>
</tr>
<tr>
<td>CCT</td>
<td>Leu</td>
</tr>
<tr>
<td>CTC</td>
<td>Leu</td>
</tr>
<tr>
<td>CTA</td>
<td>Leu</td>
</tr>
<tr>
<td>CGT</td>
<td>Arg</td>
</tr>
<tr>
<td>AGT</td>
<td>Ser</td>
</tr>
<tr>
<td>ATT</td>
<td>Ile</td>
</tr>
<tr>
<td>GGT</td>
<td>Val</td>
</tr>
<tr>
<td>GTC</td>
<td>Val</td>
</tr>
<tr>
<td>GTA</td>
<td>Val</td>
</tr>
<tr>
<td>GAA</td>
<td>Gln</td>
</tr>
<tr>
<td>GAG</td>
<td>Gln</td>
</tr>
</tbody>
</table>

### TABLE 2

**Codon Usage Table for Human Genes (Homo sapiens).**

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon</th>
<th>Number</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe</td>
<td>UUU</td>
<td>320146</td>
<td>0.4525</td>
</tr>
<tr>
<td>Phe</td>
<td>UUC</td>
<td>394680</td>
<td>0.5475</td>
</tr>
<tr>
<td>Leu</td>
<td>UUA</td>
<td>723826</td>
<td>0.0728</td>
</tr>
<tr>
<td>Leu</td>
<td>UUG</td>
<td>242151</td>
<td>0.1266</td>
</tr>
<tr>
<td>Leu</td>
<td>CUU</td>
<td>246206</td>
<td>0.1287</td>
</tr>
<tr>
<td>Leu</td>
<td>CUC</td>
<td>374262</td>
<td>0.1956</td>
</tr>
<tr>
<td>Leu</td>
<td>CUA</td>
<td>133980</td>
<td>0.0700</td>
</tr>
<tr>
<td>Leu</td>
<td>CUG</td>
<td>777017</td>
<td>0.4062</td>
</tr>
<tr>
<td>Ile</td>
<td>AUA</td>
<td>391292</td>
<td>0.3584</td>
</tr>
<tr>
<td>Ile</td>
<td>AUC</td>
<td>307721</td>
<td>0.3584</td>
</tr>
<tr>
<td>Ile</td>
<td>UAU</td>
<td>414843</td>
<td>0.4850</td>
</tr>
<tr>
<td>Ile</td>
<td>UAA</td>
<td>136399</td>
<td>0.1596</td>
</tr>
<tr>
<td>Met</td>
<td>AUG</td>
<td>854603</td>
<td>1.0000</td>
</tr>
<tr>
<td>Val</td>
<td>GUU</td>
<td>430946</td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>GUC</td>
<td>210423</td>
<td>0.1773</td>
</tr>
<tr>
<td>Val</td>
<td>GCA</td>
<td>282445</td>
<td>0.2380</td>
</tr>
<tr>
<td>Val</td>
<td>GCG</td>
<td>134991</td>
<td>0.1137</td>
</tr>
<tr>
<td>Val</td>
<td>GUG</td>
<td>550544</td>
<td>0.4710</td>
</tr>
<tr>
<td>Ser</td>
<td>UCU</td>
<td>118603</td>
<td>0.1840</td>
</tr>
<tr>
<td>Ser</td>
<td>UCC</td>
<td>336349</td>
<td>0.2191</td>
</tr>
<tr>
<td>Ser</td>
<td>UCA</td>
<td>229071</td>
<td>0.1472</td>
</tr>
<tr>
<td>Ser</td>
<td>UCG</td>
<td>86761</td>
<td>0.0565</td>
</tr>
<tr>
<td>Ser</td>
<td>AGU</td>
<td>230874</td>
<td>0.1499</td>
</tr>
<tr>
<td>Ser</td>
<td>AGC</td>
<td>373362</td>
<td>0.2433</td>
</tr>
<tr>
<td>Pro</td>
<td>CCA</td>
<td>1534889</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>CCC</td>
<td>333705</td>
<td>0.2834</td>
</tr>
<tr>
<td>Pro</td>
<td>CCG</td>
<td>386462</td>
<td>0.3281</td>
</tr>
<tr>
<td>Pro</td>
<td>CCA</td>
<td>322220</td>
<td>0.2756</td>
</tr>
<tr>
<td>Pro</td>
<td>CGG</td>
<td>120317</td>
<td>0.1149</td>
</tr>
<tr>
<td>Thr</td>
<td>ACC</td>
<td>1177704</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>ACA</td>
<td>247913</td>
<td>0.2419</td>
</tr>
<tr>
<td>Thr</td>
<td>ACC</td>
<td>371420</td>
<td>0.3624</td>
</tr>
<tr>
<td>Thr</td>
<td>ACG</td>
<td>285655</td>
<td>0.2787</td>
</tr>
<tr>
<td>Thr</td>
<td>ACG</td>
<td>120322</td>
<td>0.1171</td>
</tr>
<tr>
<td>Ala</td>
<td>GCU</td>
<td>102510</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>GCC</td>
<td>361146</td>
<td>0.2637</td>
</tr>
<tr>
<td>Ala</td>
<td>GCA</td>
<td>551452</td>
<td>0.4037</td>
</tr>
<tr>
<td>Ala</td>
<td>GAG</td>
<td>308034</td>
<td>0.2255</td>
</tr>
<tr>
<td>Ala</td>
<td>GCG</td>
<td>146233</td>
<td>0.1071</td>
</tr>
<tr>
<td>Tyr</td>
<td>UAU</td>
<td>3365865</td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>UAC</td>
<td>232240</td>
<td>0.3437</td>
</tr>
<tr>
<td>His</td>
<td>CAU</td>
<td>534218</td>
<td>0.4113</td>
</tr>
<tr>
<td>His</td>
<td>CAC</td>
<td>201389</td>
<td>0.5653</td>
</tr>
<tr>
<td>His</td>
<td>CAC</td>
<td>288200</td>
<td>0.5887</td>
</tr>
<tr>
<td>Glu</td>
<td>CAA</td>
<td>490559</td>
<td>0.2541</td>
</tr>
<tr>
<td>Glu</td>
<td>CAG</td>
<td>227472</td>
<td>0.2541</td>
</tr>
<tr>
<td>Glu</td>
<td>GCG</td>
<td>668391</td>
<td>0.7459</td>
</tr>
<tr>
<td>Asp</td>
<td>AAU</td>
<td>896133</td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>AAC</td>
<td>322271</td>
<td>0.4614</td>
</tr>
<tr>
<td>Asp</td>
<td>AAG</td>
<td>378210</td>
<td>0.5896</td>
</tr>
<tr>
<td>Lys</td>
<td>AAA</td>
<td>698481</td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>AAG</td>
<td>462660</td>
<td>0.4212</td>
</tr>
<tr>
<td>Lys</td>
<td>AAG</td>
<td>635752</td>
<td>0.5788</td>
</tr>
</tbody>
</table>

[0103] Many organisms display a bias for use of particular codons to code for insertion of a particular amino acid in a growing peptide chain. Codon preference or codon bias, differences in codon usage between organisms, is afforded by degeneracy of the genetic code, and is well documented among many organisms. Codon bias often correlates with the efficiency of translation of messenger RNA (mRNA), which is in turn believed to be dependent on, inter alia, the properties of the codons being translated and the availability of particular transfer RNA (tRNA) molecules. The predominance of selected tRNAs in a cell is generally a reflection of the codons used most frequently in peptide synthesis. Accordingly, genes can be tailored for optimal gene expression in a given organism based on codon optimization.

[0104] Given the large number of gene sequences available for a wide variety of animal, plant and microbial species, it is possible to calculate the relative frequencies of codon usage. Codon usage tables are readily available, for example, at the "Codon Usage Database" available at http://www.kazusa.or.jp/codon/ (visited Jul. 9, 2002), and these tables can be adapted in a number of ways. See Nakamura, Y., et al. "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28:292 (2000), which is incorporated by reference. As examples, the codon usage tables for human, mouse, domestic cat, and cow, published from GeneBank Release 128.0 (15 Feb. 2002), are reproduced below as Tables 2-5. These Tables use mRNA nomenclature, and so instead of thymine (T) which is found in DNA, the Tables use uracil (U) which is found in RNA. The Tables have been adapted so that frequencies are calculated for each amino acid, rather than for all 64 codons.
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon</th>
<th>Number</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>GAC</td>
<td>50296</td>
<td>0.5387</td>
</tr>
<tr>
<td>Asp</td>
<td>GAU</td>
<td>430744</td>
<td>0.4613</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>933864</td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>GAA</td>
<td>561277</td>
<td>0.4161</td>
</tr>
<tr>
<td>Glu</td>
<td>GAG</td>
<td>787712</td>
<td>0.5859</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>3348989</td>
<td></td>
</tr>
<tr>
<td>Cys</td>
<td>UGU</td>
<td>199062</td>
<td>0.4468</td>
</tr>
<tr>
<td>Cys</td>
<td>UGC</td>
<td>236400</td>
<td>0.5552</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>427362</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>UGG</td>
<td>240083</td>
<td>1.0000</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>240083</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>CGU</td>
<td>90899</td>
<td>0.0839</td>
</tr>
<tr>
<td>Arg</td>
<td>CCG</td>
<td>210931</td>
<td>0.1927</td>
</tr>
<tr>
<td>Arg</td>
<td>CGA</td>
<td>122555</td>
<td>0.1120</td>
</tr>
<tr>
<td>Arg</td>
<td>CGG</td>
<td>228970</td>
<td>0.2052</td>
</tr>
<tr>
<td>Arg</td>
<td>AGA</td>
<td>221223</td>
<td>0.2021</td>
</tr>
<tr>
<td>Arg</td>
<td>AGG</td>
<td>220119</td>
<td>0.2011</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1094695</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>GGU</td>
<td>209450</td>
<td>0.1632</td>
</tr>
<tr>
<td>Gly</td>
<td>GGC</td>
<td>441320</td>
<td>0.3438</td>
</tr>
<tr>
<td>Gly</td>
<td>GGA</td>
<td>315726</td>
<td>0.2459</td>
</tr>
<tr>
<td>Gly</td>
<td>GGG</td>
<td>337263</td>
<td>0.2471</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1283579</td>
<td></td>
</tr>
<tr>
<td>Stop</td>
<td>UAA</td>
<td>13963</td>
<td></td>
</tr>
<tr>
<td>Stop</td>
<td>UAG</td>
<td>10631</td>
<td></td>
</tr>
<tr>
<td>Stop</td>
<td>UGA</td>
<td>24607</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Codon Usage Table for Mouse Genes (Mus musculus)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino Acid</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>Ser</td>
</tr>
<tr>
<td>Ser</td>
</tr>
<tr>
<td>Ser</td>
</tr>
<tr>
<td>Ser</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Pro</td>
</tr>
<tr>
<td>Pro</td>
</tr>
<tr>
<td>Pro</td>
</tr>
<tr>
<td>Pro</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Thr</td>
</tr>
<tr>
<td>Thr</td>
</tr>
<tr>
<td>Thr</td>
</tr>
<tr>
<td>Thr</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Ala</td>
</tr>
<tr>
<td>Ala</td>
</tr>
<tr>
<td>Ala</td>
</tr>
<tr>
<td>Ala</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Tyr</td>
</tr>
<tr>
<td>Tyr</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>His</td>
</tr>
<tr>
<td>His</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Gln</td>
</tr>
<tr>
<td>Gln</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Asp</td>
</tr>
<tr>
<td>Asp</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Glu</td>
</tr>
<tr>
<td>Glu</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Cys</td>
</tr>
<tr>
<td>Cys</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Thr</td>
</tr>
<tr>
<td>Stop</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Arg</td>
</tr>
<tr>
<td>Arg</td>
</tr>
<tr>
<td>Arg</td>
</tr>
<tr>
<td>Arg</td>
</tr>
<tr>
<td>Arg</td>
</tr>
<tr>
<td>Arg</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Gly</td>
</tr>
<tr>
<td>Gly</td>
</tr>
<tr>
<td>Gly</td>
</tr>
<tr>
<td>Gly</td>
</tr>
</tbody>
</table>

[0105]

TABLE 3

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon</th>
<th>Number</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe</td>
<td>UUU</td>
<td>150467</td>
<td>0.4321</td>
</tr>
<tr>
<td>Phe</td>
<td>UUC</td>
<td>197905</td>
<td>0.5679</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>348262</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>UUA</td>
<td>55635</td>
<td>0.0625</td>
</tr>
<tr>
<td>Leu</td>
<td>UUG</td>
<td>116210</td>
<td>0.1356</td>
</tr>
<tr>
<td>Leu</td>
<td>CUU</td>
<td>114669</td>
<td>0.1289</td>
</tr>
<tr>
<td>Leu</td>
<td>CUC</td>
<td>179246</td>
<td>0.2015</td>
</tr>
<tr>
<td>Leu</td>
<td>CUA</td>
<td>69237</td>
<td>0.0778</td>
</tr>
<tr>
<td>Leu</td>
<td>CUG</td>
<td>354745</td>
<td>0.3987</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>889772</td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>AUU</td>
<td>137513</td>
<td>0.3367</td>
</tr>
<tr>
<td>Ile</td>
<td>AUC</td>
<td>208533</td>
<td>0.5106</td>
</tr>
<tr>
<td>Ile</td>
<td>AUA</td>
<td>62349</td>
<td>0.1527</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>408395</td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>AUG</td>
<td>204546</td>
<td>1.0000</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>204546</td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>GUU</td>
<td>93754</td>
<td>0.1673</td>
</tr>
<tr>
<td>Val</td>
<td>GUC</td>
<td>140762</td>
<td>0.2513</td>
</tr>
<tr>
<td>Val</td>
<td>GUA</td>
<td>64417</td>
<td>0.1150</td>
</tr>
<tr>
<td>Val</td>
<td>GUG</td>
<td>261308</td>
<td>0.4664</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>560241</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>UCU</td>
<td>139576</td>
<td>0.1936</td>
</tr>
<tr>
<td>Ser</td>
<td>UCC</td>
<td>160313</td>
<td>0.2224</td>
</tr>
</tbody>
</table>

[0105]
### TABLE 3-continued

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon</th>
<th>Number</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>592474</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stop</td>
<td>UAA</td>
<td>4497</td>
<td></td>
</tr>
<tr>
<td>Stop</td>
<td>UAG</td>
<td>4661</td>
<td></td>
</tr>
<tr>
<td>Stop</td>
<td>UGA</td>
<td>10356</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 4

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon</th>
<th>Number</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe</td>
<td>UUU</td>
<td>1204.00</td>
<td>0.4039</td>
</tr>
<tr>
<td>Phe</td>
<td>UUC</td>
<td>1777.00</td>
<td>0.5961</td>
</tr>
<tr>
<td>Total</td>
<td>2981</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>UUA</td>
<td>404.00</td>
<td>0.0570</td>
</tr>
<tr>
<td>Leu</td>
<td>UUG</td>
<td>857.00</td>
<td>0.1209</td>
</tr>
<tr>
<td>Leu</td>
<td>CUU</td>
<td>791.00</td>
<td>0.1136</td>
</tr>
<tr>
<td>Leu</td>
<td>CUC</td>
<td>1513.00</td>
<td>0.2135</td>
</tr>
<tr>
<td>Leu</td>
<td>CUA</td>
<td>488.00</td>
<td>0.0688</td>
</tr>
<tr>
<td>Leu</td>
<td>CUG</td>
<td>2035.00</td>
<td>0.2824</td>
</tr>
<tr>
<td>Total</td>
<td>7088</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>AUA</td>
<td>3018.00</td>
<td>0.2584</td>
</tr>
<tr>
<td>Ile</td>
<td>AUC</td>
<td>1835.00</td>
<td>0.5580</td>
</tr>
<tr>
<td>Ile</td>
<td>AUA</td>
<td>558.00</td>
<td>0.1636</td>
</tr>
<tr>
<td>Total</td>
<td>3411</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>AUG</td>
<td>1553.00</td>
<td>0.0036</td>
</tr>
<tr>
<td>Total</td>
<td>1553</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>GUU</td>
<td>696.00</td>
<td>0.1512</td>
</tr>
<tr>
<td>Val</td>
<td>GUC</td>
<td>1279.00</td>
<td>0.2779</td>
</tr>
<tr>
<td>Val</td>
<td>GUA</td>
<td>463.00</td>
<td>0.1006</td>
</tr>
<tr>
<td>Val</td>
<td>GUG</td>
<td>2164.00</td>
<td>0.4702</td>
</tr>
<tr>
<td>Total</td>
<td>4602</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>UCU</td>
<td>940.00</td>
<td>0.1875</td>
</tr>
<tr>
<td>Ser</td>
<td>UCC</td>
<td>1200.00</td>
<td>0.2513</td>
</tr>
<tr>
<td>Ser</td>
<td>CCA</td>
<td>600.00</td>
<td>0.1213</td>
</tr>
<tr>
<td>Ser</td>
<td>UCG</td>
<td>332.00</td>
<td>0.0662</td>
</tr>
<tr>
<td>Ser</td>
<td>AGU</td>
<td>672.00</td>
<td>0.1540</td>
</tr>
<tr>
<td>Ser</td>
<td>AGC</td>
<td>1202.00</td>
<td>0.2597</td>
</tr>
<tr>
<td>Total</td>
<td>5014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>UUC</td>
<td>956.00</td>
<td>0.2626</td>
</tr>
<tr>
<td>Pro</td>
<td>CCC</td>
<td>1375.00</td>
<td>0.3769</td>
</tr>
<tr>
<td>Pro</td>
<td>CCA</td>
<td>850.00</td>
<td>0.2330</td>
</tr>
<tr>
<td>Pro</td>
<td>CGG</td>
<td>465.00</td>
<td>0.1255</td>
</tr>
<tr>
<td>Total</td>
<td>3648</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>ACA</td>
<td>822.00</td>
<td>0.2127</td>
</tr>
<tr>
<td>Thr</td>
<td>ACC</td>
<td>1574.00</td>
<td>0.4072</td>
</tr>
<tr>
<td>Thr</td>
<td>ACA</td>
<td>903.00</td>
<td>0.2336</td>
</tr>
<tr>
<td>Thr</td>
<td>ACC</td>
<td>566.00</td>
<td>0.1464</td>
</tr>
<tr>
<td>Total</td>
<td>3865</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>GCC</td>
<td>1129.00</td>
<td>0.2946</td>
</tr>
<tr>
<td>Ala</td>
<td>GCC</td>
<td>1051.00</td>
<td>0.2431</td>
</tr>
<tr>
<td>Ala</td>
<td>GCA</td>
<td>883.00</td>
<td>0.2152</td>
</tr>
<tr>
<td>Ala</td>
<td>GCG</td>
<td>561.00</td>
<td>0.1340</td>
</tr>
<tr>
<td>Total</td>
<td>4524</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>UAU</td>
<td>837.00</td>
<td>0.2211</td>
</tr>
<tr>
<td>Tyr</td>
<td>UAC</td>
<td>1378.00</td>
<td>0.3779</td>
</tr>
<tr>
<td>Total</td>
<td>2215</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 5

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon</th>
<th>Number</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe</td>
<td>UUUA</td>
<td>13002</td>
<td>0.4112</td>
</tr>
<tr>
<td>Phe</td>
<td>UUC</td>
<td>18614</td>
<td>0.5888</td>
</tr>
<tr>
<td>Total</td>
<td>31616</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>UUA</td>
<td>4467</td>
<td>0.0590</td>
</tr>
<tr>
<td>Leu</td>
<td>UUG</td>
<td>9024</td>
<td>0.1192</td>
</tr>
<tr>
<td>Leu</td>
<td>CUU</td>
<td>9069</td>
<td>0.1298</td>
</tr>
<tr>
<td>Leu</td>
<td>CUC</td>
<td>1603</td>
<td>0.2114</td>
</tr>
<tr>
<td>Leu</td>
<td>CUA</td>
<td>4608</td>
<td>0.0809</td>
</tr>
<tr>
<td>Leu</td>
<td>CUG</td>
<td>22536</td>
<td>0.4298</td>
</tr>
<tr>
<td>Total</td>
<td>75707</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 4-continued

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon</th>
<th>Number</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>His</td>
<td>CAU</td>
<td>594.00</td>
<td>0.3738</td>
</tr>
<tr>
<td>His</td>
<td>CAC</td>
<td>995.00</td>
<td>0.6262</td>
</tr>
<tr>
<td>Total</td>
<td>1589</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>CAA</td>
<td>747.00</td>
<td>0.2783</td>
</tr>
<tr>
<td>Gln</td>
<td>CAG</td>
<td>937.00</td>
<td>0.7217</td>
</tr>
<tr>
<td>Total</td>
<td>2684</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>AAU</td>
<td>1109.00</td>
<td>0.3949</td>
</tr>
<tr>
<td>Asn</td>
<td>AAC</td>
<td>1699.00</td>
<td>0.6051</td>
</tr>
<tr>
<td>Total</td>
<td>2808</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>AAA</td>
<td>1445.00</td>
<td>0.4088</td>
</tr>
<tr>
<td>Lys</td>
<td>AAG</td>
<td>2090.00</td>
<td>0.5912</td>
</tr>
<tr>
<td>Total</td>
<td>3535</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>GAU</td>
<td>925.00</td>
<td>0.4055</td>
</tr>
<tr>
<td>Asp</td>
<td>GAC</td>
<td>1140.00</td>
<td>0.5948</td>
</tr>
<tr>
<td>Total</td>
<td>1865</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>GAA</td>
<td>763.00</td>
<td>0.4164</td>
</tr>
<tr>
<td>Glu</td>
<td>GAG</td>
<td>1284.00</td>
<td>0.5826</td>
</tr>
<tr>
<td>Total</td>
<td>2047</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cys</td>
<td>UGU</td>
<td>719.00</td>
<td>0.4425</td>
</tr>
<tr>
<td>Cys</td>
<td>UGC</td>
<td>906.00</td>
<td>0.5575</td>
</tr>
<tr>
<td>Total</td>
<td>1625</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>UGG</td>
<td>1073.00</td>
<td>1.0000</td>
</tr>
<tr>
<td>Total</td>
<td>1073</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>CGU</td>
<td>236.00</td>
<td>0.0700</td>
</tr>
<tr>
<td>Arg</td>
<td>CGC</td>
<td>629.00</td>
<td>0.1865</td>
</tr>
<tr>
<td>Arg</td>
<td>CCA</td>
<td>354.00</td>
<td>0.1050</td>
</tr>
<tr>
<td>Arg</td>
<td>CGG</td>
<td>662.00</td>
<td>0.1963</td>
</tr>
<tr>
<td>Arg</td>
<td>AGA</td>
<td>712.00</td>
<td>0.2112</td>
</tr>
<tr>
<td>Arg</td>
<td>AGG</td>
<td>779.00</td>
<td>0.2310</td>
</tr>
<tr>
<td>Total</td>
<td>3372</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>GGU</td>
<td>648.00</td>
<td>0.1498</td>
</tr>
<tr>
<td>Gly</td>
<td>GGC</td>
<td>1586.00</td>
<td>0.3551</td>
</tr>
<tr>
<td>Gly</td>
<td>GGA</td>
<td>2065.00</td>
<td>0.4262</td>
</tr>
<tr>
<td>Gly</td>
<td>GGG</td>
<td>1077.00</td>
<td>0.2490</td>
</tr>
<tr>
<td>Total</td>
<td>4326</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stop</td>
<td>UAA</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>Stop</td>
<td>UAG</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>Stop</td>
<td>UGA</td>
<td>110</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 5

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon</th>
<th>Number</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe</td>
<td>UUUA</td>
<td>13002</td>
<td>0.4112</td>
</tr>
<tr>
<td>Phe</td>
<td>UUC</td>
<td>18614</td>
<td>0.5888</td>
</tr>
<tr>
<td>Total</td>
<td>31616</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>UUA</td>
<td>4467</td>
<td>0.0590</td>
</tr>
<tr>
<td>Leu</td>
<td>UUG</td>
<td>9024</td>
<td>0.1192</td>
</tr>
<tr>
<td>Leu</td>
<td>CUU</td>
<td>9069</td>
<td>0.1298</td>
</tr>
<tr>
<td>Leu</td>
<td>CUC</td>
<td>1603</td>
<td>0.2114</td>
</tr>
<tr>
<td>Leu</td>
<td>CUA</td>
<td>4608</td>
<td>0.0809</td>
</tr>
<tr>
<td>Leu</td>
<td>CUG</td>
<td>22536</td>
<td>0.4298</td>
</tr>
<tr>
<td>Total</td>
<td>75707</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 5-continued

**Codon Usage Table for Cow Genes (Bos taurus)**

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon</th>
<th>Number</th>
<th>Frequency of usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ile</td>
<td>AUA</td>
<td>5581</td>
<td>0.1429</td>
</tr>
<tr>
<td></td>
<td>AUU</td>
<td>12474</td>
<td>0.3313</td>
</tr>
<tr>
<td></td>
<td>AUC</td>
<td>19800</td>
<td>0.5258</td>
</tr>
<tr>
<td>Met</td>
<td>AUG</td>
<td>17770</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>AUU</td>
<td>17770</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>AUG</td>
<td>17770</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>GUA</td>
<td>4932</td>
<td>0.0982</td>
</tr>
<tr>
<td></td>
<td>GUG</td>
<td>24222</td>
<td>0.4824</td>
</tr>
<tr>
<td></td>
<td>UUA</td>
<td>50212</td>
<td>0.1804</td>
</tr>
<tr>
<td></td>
<td>UUC</td>
<td>11069</td>
<td>0.2684</td>
</tr>
<tr>
<td></td>
<td>UCA</td>
<td>123258</td>
<td>0.0235</td>
</tr>
<tr>
<td></td>
<td>UCG</td>
<td>3470</td>
<td>0.0609</td>
</tr>
<tr>
<td></td>
<td>AGU</td>
<td>9040</td>
<td>0.1410</td>
</tr>
<tr>
<td></td>
<td>AGC</td>
<td>14279</td>
<td>0.2503</td>
</tr>
<tr>
<td></td>
<td>CCU</td>
<td>11069</td>
<td>0.2684</td>
</tr>
<tr>
<td></td>
<td>CCC</td>
<td>15221</td>
<td>0.2493</td>
</tr>
<tr>
<td></td>
<td>CCA</td>
<td>11039</td>
<td>0.2533</td>
</tr>
<tr>
<td></td>
<td>CGG</td>
<td>50212</td>
<td>0.1290</td>
</tr>
<tr>
<td></td>
<td>CAC</td>
<td>24576</td>
<td>0.2923</td>
</tr>
<tr>
<td></td>
<td>ACC</td>
<td>16574</td>
<td>0.3895</td>
</tr>
<tr>
<td></td>
<td>ACA</td>
<td>10892</td>
<td>0.2560</td>
</tr>
<tr>
<td></td>
<td>AGC</td>
<td>5712</td>
<td>0.1342</td>
</tr>
<tr>
<td></td>
<td>GCU</td>
<td>42550</td>
<td>0.2592</td>
</tr>
<tr>
<td></td>
<td>GCC</td>
<td>23073</td>
<td>0.4295</td>
</tr>
<tr>
<td></td>
<td>GCA</td>
<td>30704</td>
<td>0.1992</td>
</tr>
<tr>
<td></td>
<td>GCG</td>
<td>4025</td>
<td>0.1121</td>
</tr>
<tr>
<td></td>
<td>UAU</td>
<td>53725</td>
<td>0.3882</td>
</tr>
<tr>
<td></td>
<td>UAC</td>
<td>14882</td>
<td>0.6118</td>
</tr>
<tr>
<td></td>
<td>CAU</td>
<td>24932</td>
<td>0.3649</td>
</tr>
<tr>
<td></td>
<td>CAC</td>
<td>11363</td>
<td>0.6351</td>
</tr>
<tr>
<td></td>
<td>CAA</td>
<td>17891</td>
<td>0.2430</td>
</tr>
<tr>
<td></td>
<td>CAG</td>
<td>25108</td>
<td>0.3750</td>
</tr>
<tr>
<td></td>
<td>AAU</td>
<td>33168</td>
<td>0.4088</td>
</tr>
<tr>
<td></td>
<td>AAC</td>
<td>12491</td>
<td>0.5912</td>
</tr>
<tr>
<td></td>
<td>AAA</td>
<td>30554</td>
<td>0.2897</td>
</tr>
<tr>
<td></td>
<td>AAG</td>
<td>27000</td>
<td>0.6103</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>44244</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 5-continued

**Codon Usage Table for Cow Genes (Bos taurus)**

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon</th>
<th>Number</th>
<th>Frequency of usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>GUA</td>
<td>16615</td>
<td>0.4329</td>
</tr>
<tr>
<td>Asp</td>
<td>GAC</td>
<td>22580</td>
<td>0.5761</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>39195</td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>GAA</td>
<td>21102</td>
<td>0.4007</td>
</tr>
<tr>
<td>Glu</td>
<td>GAG</td>
<td>31555</td>
<td>0.5993</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>52657</td>
<td></td>
</tr>
<tr>
<td>Cys</td>
<td>UGU</td>
<td>7566</td>
<td>0.4200</td>
</tr>
<tr>
<td>Cys</td>
<td>UGC</td>
<td>10426</td>
<td>0.5800</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>17992</td>
<td></td>
</tr>
<tr>
<td>Trp</td>
<td>UGG</td>
<td>10706</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>17992</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>CGU</td>
<td>3391</td>
<td>0.0824</td>
</tr>
<tr>
<td>Arg</td>
<td>CGC</td>
<td>7998</td>
<td>0.1943</td>
</tr>
<tr>
<td>Arg</td>
<td>CGG</td>
<td>4588</td>
<td>0.1108</td>
</tr>
<tr>
<td>Arg</td>
<td>CGA</td>
<td>8300</td>
<td>0.2017</td>
</tr>
<tr>
<td>Arg</td>
<td>AGA</td>
<td>8287</td>
<td>0.2001</td>
</tr>
<tr>
<td>Arg</td>
<td>AGG</td>
<td>8671</td>
<td>0.2107</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>41155</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>GGU</td>
<td>8508</td>
<td>0.1616</td>
</tr>
<tr>
<td>Gly</td>
<td>GGC</td>
<td>18517</td>
<td>0.3518</td>
</tr>
<tr>
<td>Gly</td>
<td>GGA</td>
<td>12838</td>
<td>0.2439</td>
</tr>
<tr>
<td>Gly</td>
<td>GGG</td>
<td>12772</td>
<td>0.2427</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>52635</td>
<td></td>
</tr>
<tr>
<td>Stop</td>
<td>UAA</td>
<td>555</td>
<td></td>
</tr>
<tr>
<td>Stop</td>
<td>UAG</td>
<td>394</td>
<td></td>
</tr>
<tr>
<td>Stop</td>
<td>UGA</td>
<td>392</td>
<td></td>
</tr>
</tbody>
</table>

[0108] By utilizing these or similar tables, one of ordinary skill in the art can apply the frequencies to any given polypeptide sequence, and produce a nucleic acid fragment of a codon-optimized coding region which encodes the polypeptide, but which uses codons more optimal for a given species. Codon-optimized coding regions can be designed by various different methods.

[0109] In one method, termed “uniform optimization,” a codon usage table is used to find the single most frequent codon used for any given amino acid, and that codon is used each time that particular amino acid appears in the polypeptide sequence. For example, referring to Table 2 above, for leucine, the most frequent codon in humans is CUG, which is used 41% of the time. Thus all the leucine residues in a given amino acid sequence would be assigned the codons CUG. A coding region for IAV NP (SEQ ID NO:2) optimized by the “uniform optimization” method is presented herein as SEQ ID NO 24:

```
1 ATGCCCCGCG ACCGACGAA CGTGAGCTTC GACGAGCTT GAGCCCCCG CCGGCGGAGC
61 ACCGACCG AGATCCGGCG CGACGTGGCG AGTGATGC CGTCGGCGC CCGGCTCCAG
121 ATCGATGCT GACGATGCC AGATGACG GACGAGGAG CGCGTGCTG AGACGACGCG
181 GTGCCAGCG AGCGAGCGG GCTGAGCGCC TGCGAGCAC CGCAGCAGCG
```
-continued

241 GAGCCACCCCA GCAGCGCCCA GAGCCAGCC AAGCCGCCG GAGCCGGCTA CCCCCCTCA GCCGGGCTG
301 AAGCCAGCAG AGAGGGCGAG GAGAGCTCTG TACGAGCAGG AGAGAGGCGG GAGCCGTCCG
361 CCCCCGGAAC AAGCCGGGCA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
421 AAGCCAGCAG AGAGGGCGAG GAGAGCTCTG TACGAGCAGG AGAGAGGCGG GAGCCGTCCG
481 CCCCCGGAAC AAGCCGGGCA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
541 GCCGGGCTTG AAGCCGGGCA CACGAGGTG AGAGGCTCTG TACGAGCAGG AGAGAGGCGG GAGCCGTCCG
601 ATCCCGACGG GAGAGCTCTG TACGAGCAGG AGAGAGGCGG GAGCCGTCCG ATCCCGGCGA
661 CCCCCGGGAAC AAGCCGGGCA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
721 CGGAGGGCGA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
781 CGGAGGGCGA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
841 CCCCCGGGAAC AAGCCGGGCA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
901 ATCCCGACGG GAGAGCTCTG TACGAGCAGG AGAGAGGCGG GAGCCGTCCG ATCCCGGCGA
961 AAGCCGGGCA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
1021 CCCCCGGGAAC AAGCCGGGCA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
1081 CCCCCGGGAAC AAGCCGGGCA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
1141 CCCCCGGGAAC AAGCCGGGCA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
1201 CCCCCGGGAAC AAGCCGGGCA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
1261 CCCCCGGGAAC AAGCCGGGCA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
1321 CCCCCGGGAAC AAGCCGGGCA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
1381 CCCCCGGGAAC AAGCCGGGCA CAGCGCCGGCC CGCGCGCTTTA CCCAGGGAT GAGCGGACAC
1441 ATCCCGACGG GAGAGCTCTG TACGAGCAGG AGAGAGGCGG GAGCCGTCCG ATCCCGGCGA

[0110] In another method, termed "full-optimization," the actual frequencies of the codons are distributed randomly throughout the coding region. Thus, using this method for optimization, if a hypothetical polypeptide sequence had 100 leucine residues, referring to Table 2 for frequency of usage in humans, about 7, or 7% of the leucine codons would be UUA, about 13, or 13% of the leucine codons would be UUG, about 13, or 13% of the leucine codons would be CUU, and about 41, or 41% of the leucine codons would be CUG. These frequencies would be distributed randomly throughout the leucine codons in the coding region encoding the hypothetical polypeptide. As will be understood by those of ordinary skill in the art, the distribution of codons in the sequence can vary significantly using this method; however, the sequence always encodes the same polypeptide.

[0111] As an example, a nucleotide sequence for NP (SEQ ID NO:2) fully optimized for human codon usage, is shown as SEQ ID NO:23. An alignment of nucleotides 46-1542 of SEQ ID NO:1 (native NP coding region) with the codon-optimized coding region (SEQ ID NO:23) is presented in FIG. 1.

[0112] In using the "full-optimization" method, an entire polypeptide sequence may be codon-optimized as described above. With respect to various desired fragments, variants or derivatives of the complete polypeptide, the fragment variant, or derivative may first be designed, and is then codon-optimized individually. Alternatively, a full-length polypeptide sequence is codon-optimized for a given species resulting in a codon-optimized coding region encoding the entire polypeptide, and then nucleic acid fragments of the codon-optimized coding region, which encode fragments, variants, and derivatives of the polypeptide are made from the original codon-optimized coding region. As would be well understood by those of ordinary skill in the art, if codons have been randomly assigned to the full-length coding region based on their frequency of usage in a given species, nucleic acid fragments encoding fragments, variants, and derivatives would not necessarily be fully codon-optimized for the given species. However, such sequences are still much closer to the codon usage of the desired species than the native codon usage. The advantage of this approach is that synthesizing codon-optimized nucleic acid fragments encoding each fragment, variant, and derivative of a given polypeptide, although routine, would be time consuming and would result in significant expense.

[0113] When using the "full-optimization" method, the term "about" is used precisely to account for fractional percentages of codon frequencies for a given amino acid. As used herein, "about" is defined as one amino acid more or one amino acid less than the value given. The whole number value of amino acids is rounded up if the fractional frequency of usage is 0.5 or greater, and is rounded down if
the fractional frequency of use is 0.49 or less. Using again the example of the frequency of usage of leucine in human genes for a hypothetical polypeptide having 62 leucine residues, the fractional frequency of codon usage would be calculated by multiplying 62 by the frequencies for the various codons. Thus, 7.28 percent of 62 equals 4.51 UUA codons, or "about 5," i.e., 4, 5, or 6 UUA codons, 12.66 percent of 62 equals 7.85 UUG codons or "about 8," i.e., 7, 8, or 9 TUG codons, 12.87 percent of 62 equals 7.98 CUU codons, or "about 8," i.e., 7, 8, or 9 CUU codons, 19.56 percent of 62 equals 12.12 CUC codons or "about 12," i.e., 11, 12, or 13 CUC codons, 7.00 percent of 62 equals 4.34 CUA codons or "about 4," i.e., 3, 4, or 5 CUA codons, and 40.62 percent of 62 equals 25.19 CUG codons, or "about 25," i.e., 24, 25, or 26 CUG codons.

[0114] In a third method termed "minimal optimization," [0115] coding regions are only partially optimized. For example, the invention includes a nucleic acid fragment of a codon- optimized coding region encoding a polypeptide in which at least about 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 100% of the codon positions have been codon-optimized for a given species. That is, they contain a codon that is preferentially used in the genes of a desired species, e.g., a vertebrate species, e.g., humans, in place of a codon that is normally used in the native nucleic acid sequence. Codons that are rarely found in the genes of the vertebrate of interest are changed to codons more commonly utilized in the coding regions of the vertebrate of interest.

[0115] Thus, those codons which are used more frequently in the IV gene of interest than in genes of the vertebrate of interest are substituted with more frequently-used codons. The difference in frequency at which the IV codons are substituted may vary based on a number factors as discussed below. For example, codons used at least twice more per thousand in IV genes as compared to genes of the vertebrate of interest are substituted with the most frequently used codon for that amino acid in the vertebrate of interest. This ratio may be adjusted higher or lower depending on various factors such as those discussed below. Accordingly, a codon in an IV native coding region would be substituted with a codon used more frequently for that amino acid in IV coding regions of the vertebrate of interest if the codon is used 1.1 times, 1.2 times, 1.3 times, 1.4 times, 1.5 times, 1.6 times, 1.7 times, 1.8 times, 1.9 times, 2.0 times, 2.1 times, 2.2 times, 2.3 times, 2.4 times, 2.5 times, 2.6 times, 2.7 times, 2.8 times, 2.9 times, 3.0 times, 3.1 times, 3.2 times, 3.3 times, 3.4 times, 3.5 times, 3.6 times, 3.7 times, 3.8 times, 3.9 times, 4.0 times, 4.1 times, 4.2 times, 4.3 times, 4.4 times, 4.5 times, 4.6 times, 4.7 times, 4.8 times, 4.9 times, 5.0 times, 5.5 times, 6.0 times, 6.5 times, 7.0 times, 7.5 times, 8.0 times, 8.5 times, 9.0 times, 9.5 times, 10.0 times, 10.5 times, 11.0 times, 11.5 times, 12.0 times, 12.5 times, 13.0 times, 13.5 times, 14.0 times, 14.5 times, 15.0 times, 15.5 times, 16.0 times, 16.5 times, 17.0 times, 17.5 times, 18.0 times, 18.5 times, 19.0 times, 19.5 times, 20.0 times, 21 times, 22 times, 23 times, 24 times, 25 times, 26 times, 27 times, greater frequency in IV coding regions than in coding regions of the vertebrate of interest.

[0116] This minimal human codon optimization for highly variant codons has several advantages, which include but are not limited to the following examples. Since fewer changes are made to the nucleotide sequence of the gene of interest, fewer manipulations are required, which leads to reduced risk of introducing unwanted mutations and lower cost, as well as allowing the use of commercially available site-directed mutagenesis kits, and reducing the need for expensive oligonucleotide synthesis. Further, decreasing the number of changes in the nucleotide sequence decreases the potential of altering the secondary structure of the sequence, which can have a significant impact on gene expression in certain host cells. The introduction of undesirable restriction sites is also reduced, facilitating the subcloning of the genes of interest into the plasmid expression vector.

[0117] The present invention also provides isolated polynucleotides comprising coding regions of IV polypeptides, e.g., NP, M1, M2, HA, NA, PB1, PB2, PA, NS1 or NS2, or fragments, variants, or derivatives thereof. The isolated polynucleotides can also be codon-optimized.

[0118] In certain embodiments described herein, a codon-optimized coding region encoding SEQ ID NO:2 is optimized according to codon usage in humans (Homo sapiens). Alternatively, a codon-optimized coding region encoding SEQ ID NO:2 may be optimized according to codon usage in any plant, animal, or microbial species. Codon-optimized coding regions encoding SEQ ID NO:2, optimized according to codon usage in humans are designed as follows. The amino acid composition of SEQ ID NO:2 is shown in Table 6.

<table>
<thead>
<tr>
<th>AMINO ACID</th>
<th>Number in SEQ ID NO: 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Ala</td>
</tr>
<tr>
<td>R</td>
<td>Arg</td>
</tr>
<tr>
<td>C</td>
<td>Cys</td>
</tr>
<tr>
<td>G</td>
<td>Gly</td>
</tr>
<tr>
<td>H</td>
<td>His</td>
</tr>
<tr>
<td>I</td>
<td>Ile</td>
</tr>
<tr>
<td>L</td>
<td>Leu</td>
</tr>
<tr>
<td>K</td>
<td>Lys</td>
</tr>
<tr>
<td>M</td>
<td>Met</td>
</tr>
<tr>
<td>F</td>
<td>Phe</td>
</tr>
<tr>
<td>P</td>
<td>Pro</td>
</tr>
<tr>
<td>S</td>
<td>Ser</td>
</tr>
<tr>
<td>T</td>
<td>Thr</td>
</tr>
<tr>
<td>W</td>
<td>Trp</td>
</tr>
<tr>
<td>Y</td>
<td>Tyr</td>
</tr>
<tr>
<td>V</td>
<td>Val</td>
</tr>
<tr>
<td>N</td>
<td>Asn</td>
</tr>
<tr>
<td>D</td>
<td>Asp</td>
</tr>
<tr>
<td>Q</td>
<td>Gln</td>
</tr>
<tr>
<td>E</td>
<td>Glu</td>
</tr>
</tbody>
</table>

[0119] Using the amino acid composition shown in Table 6, a human codon-optimized coding region which encodes SEQ ID NO:2 can be designed by any of the methods discussed herein. For "uniform" optimization, each amino acid is assigned the most frequent codon used in the human genome for that amino acid. According to this method, codons are assigned to the coding region encoding SEQ ID NO:2 as follows: the 18 phenylalanine codons are TTC; the 33 leucine codons are CTG, the 26 isoleucine codons are ATC, the 25 methionine codons are ATG, the 23 valine codons are GTG, the 40 serine codons are AGC, the 17 prolline codons are CCC, the 26 threonine codons are ACC, the 30 alanine codons are GCC, the 15 tyrosine codons are
TAC, the 6 histidine codons are CAC, the 21 glutamine codons are CAG, the 26 asparagine codons are AAC, the 21 lysine codons are AAG, the 22 aspartic acid codons are GAC, the 36 glutamic acid codons are GAG, the 6 tryptophan codons are TGG, the 49 arginine codons are CGG, AGA, or AGG (the frequencies of usage of these three codons in the human genome are not significantly different), and the 41 glycine codons are GGC.

[0120] Alternatively, a human codon-optimized coding region which encodes SEQ ID NO:2 can be designed by the “full optimization” method, where each amino acid is assigned codons based on the frequency of usage in the human genome. These frequencies are shown in Table 6 above. Using this latter method, codons are assigned to the coding region encoding SEQ ID NO:2 as follows: about 8 of the 18 phenylalanine codons are TTT, and about 10 of the phenylalanine codons are TTC; about 2 of the 33 leucine codons are TTA, about 4 of the leucine codons are TTG, about 4 of the leucine codons are CTT, about 6 of the leucine codons are CTC, about 2 of the leucine codons are CTG, and about 13 of the leucine codons are CGT; about 9 of the 26 isoleucine codons are ATI, about 13 of the isoleucine codons are AIC, and about 4 of the isoleucine codons are ATG; the 25 methionine codons are ATG; about 4 of the 23 valine codons are GTT, about 5 of the valine codons are GGT, about 3 of the valine codons are GAT, and about 11 of the valine codons are GAG; about 7 of the 40 serine codons are CTT, about 9 of the serine codons are TCC, about 6 of the serine codons are TCA, about 2 of the serine codons are TCG, about 6 of the serine codons are AGT, and about 10 of the serine codons are AGC; about 5 of the 17 proline codons are CCT, about 6 of the proline codons are CCC, about 5 of the proline codons are CCA, and about 2 of the proline codons are CGG; about 7 of the 28 threonine codons are ACT, about 10 of the threonine codons are ACC, about 8 of the threonine codons are ACA, and about 3 of the threonine codons are ACC; about 10 of the 39 alanine codons are GCT, about 16 of the alanine codons are GCC, about 9 of the alanine codons are GCA, and about 4 of the alanine codons are GCG; about 7 of the 15 tyrosine codons are TAT and about 8 of the tyrosine codons are TAC; about 2 of the 6 histidine codons are CAT and about 4 of the histidine codons are CAC; about 5 of the 21 glutamine codons are CAA and about 16 of the glutamine codons are CAG; about 12 of the 26 asparagine codons are AAT and about 14 of the asparagine codons are AAC; about 9 of the 21 lysine codons are AAA and about 12 of the lysine codons are AAG; about 10 of the 22 aspartic acid codons are GAT and about 12 of the aspartic acid codons are GAC; about 11 of the 26 glutamic acid codons are GAA and about 15 of the glutamic acid codons are GAG; about 3 of the 6 cysteine codons are TGT and about 3 of the cysteine codons are TGC; the 6 tryptophan codons are TGG; about 4 of the 49 arginine codons are CGT, about 9 of the arginine codons are CGC, about 5 of the arginine codons are CAG, about 10 of the arginine codons are AGA, and about 10 of the arginine codons are AGG; and about 7 of the 41 glycine codons are GGT, about 14 of the glycine codons are GGC, about 10 of the glycine codons are GGA, and about 10 of the glycine codons are GGG.

[0121] As described above, the term “about” means that the number of amino acids encoded by a certain codon may be one more or one less than the number given. It would be understood by those of ordinary skill in the art that the total number of any amino acid in the polypeptide sequence must remain constant; therefore, if there is one “more” of one codon encoding a given amino acid, there would have to be one “less” of another codon encoding that same amino acid.

[0122] A representative “fully optimized” codon-optimized coding region encoding SEQ ID NO:2, optimized according to codon usage in humans is presented herein as SEQ ID NO:23.

[0123] Additionally, a minimally codon-optimized nucleotide sequence encoding SEQ ID NO:2 can be designed by changing only certain codons found more frequently in IV genes than in human genes, as shown in Table 7. For example, if it is desired to substitute more frequently used codons in humans for those codons that occur at least 2 times more frequently in IV genes (designated with an asterisk in Table 7), Arg AGA, which occurs 2.3 times more frequently in IV genes than in human genes, is changed to, e.g., CGG; Asn AAT, which occurs 2.0 times more frequently in IV genes than in human genes, is changed to, e.g., GCC; Ile AIA, which occurs 3.6 times more frequently in IV genes than in human genes, is changed to, e.g., ATC; and Leu CTA, which occurs 2.0 times more frequently in IV genes than is human, is changed to, e.g., CTG.

### Table 7

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Codon</th>
<th>Human</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>A</td>
<td>GCA</td>
<td>16</td>
</tr>
<tr>
<td>Arg</td>
<td>R</td>
<td>AGA</td>
<td>12</td>
</tr>
<tr>
<td>Asn</td>
<td>N</td>
<td>AAC</td>
<td>20</td>
</tr>
<tr>
<td>Asp</td>
<td>D</td>
<td>GAC</td>
<td>26</td>
</tr>
<tr>
<td>Cys</td>
<td>C</td>
<td>TGC</td>
<td>12</td>
</tr>
<tr>
<td>Glu</td>
<td>Q</td>
<td>CAA</td>
<td>12</td>
</tr>
<tr>
<td>Gly</td>
<td>G</td>
<td>GGA</td>
<td>16</td>
</tr>
<tr>
<td>His</td>
<td>H</td>
<td>CAC</td>
<td>15</td>
</tr>
<tr>
<td>Ile</td>
<td>I</td>
<td>AIA</td>
<td>7</td>
</tr>
<tr>
<td>Leu</td>
<td>L</td>
<td>CTG</td>
<td>7</td>
</tr>
<tr>
<td>Lys</td>
<td>K</td>
<td>AAA</td>
<td>24</td>
</tr>
<tr>
<td>Met</td>
<td>M</td>
<td>ATG</td>
<td>22</td>
</tr>
<tr>
<td>Phe</td>
<td>F</td>
<td>TTC</td>
<td>21</td>
</tr>
<tr>
<td>Pro</td>
<td>P</td>
<td>CCA</td>
<td>17</td>
</tr>
</tbody>
</table>

*Note: * denotes the optimized codon.
### TABLE 7-continued

<table>
<thead>
<tr>
<th>Codon Usage Table for Human Genes and IV Genes</th>
<th>Amino Acid</th>
<th>Codon</th>
<th>Human</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGG</td>
<td>Ser</td>
<td>AGC</td>
<td>19</td>
<td>14</td>
</tr>
<tr>
<td>GGT</td>
<td></td>
<td>AGT</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>GGA</td>
<td></td>
<td>TCA</td>
<td>12</td>
<td>23</td>
</tr>
<tr>
<td>GCC</td>
<td></td>
<td>TCG</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>GCA</td>
<td></td>
<td>TCC</td>
<td>18</td>
<td>12</td>
</tr>
<tr>
<td>GCT</td>
<td></td>
<td>TCT</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>ACA</td>
<td>Thr</td>
<td>TAC</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td>ACC</td>
<td></td>
<td>TAT</td>
<td>12</td>
<td>19</td>
</tr>
<tr>
<td>TGG</td>
<td></td>
<td>GTA</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>GTG</td>
<td></td>
<td>GGA</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

### TABLE 7-continued

<table>
<thead>
<tr>
<th>Codon Usage Table for Human Genes and IV Genes</th>
<th>Amino Acid</th>
<th>Codon</th>
<th>Human</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTC</td>
<td></td>
<td>TAA</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>GTC</td>
<td></td>
<td>TAG</td>
<td>0.5</td>
<td>0.4</td>
</tr>
<tr>
<td>GCA</td>
<td></td>
<td>TGA</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

[0124] In another form of minimal optimization, a Codon Usage Table (CUT) for the specific IV sequence in question is generated and compared to CUT for human genomic DNA (see Table 7, supra). Amino acids are identified for which there is a difference of at least 10 percentage points in codon usage between human and IV DNA (either more or less). Then the wild type IV codon is modified to conform to predominant human codon for each such amino acid. Furthermore, the remainder of codons for that amino acid are also modified such that they conform to the predominant human codon for each such amino acid.

[0125] A representative “minimally optimized” codon-optimized coding region encoding SEQ ID NO:2, minimally optimized according to codon usage in humans by this latter method, is presented herein as SEQ ID NO:25:

```
1 ATGCTCTCAC AGGCCACAA CGCGGTTAT GACGAGATGC AGGCCATTG AGGGAGGACAG
61 AGGCCACAG AGACTGAGCC CTGGTTGGGC AGAGATGACG GCGCCATCG CGCGTCCTAT
121 ATGCTATTGT GCACGAGCT GACCTGAGCC GACTGGAGG GCAGATTGAT TGAAGACTCT
181 CTGCCATCATG AGGAGATGCT CCGAGTGGCC TGGCTATGAG GACGGAACCA GATCTCTGGAG
241 GACGGATCCCT CGCAGGGCCA GAGACCCAGG AGAGGAAAGC GCCCAATYAT TACGAGATGTT
301 AGGCCGACGT GAGAGAAGAG GCTCAGATCT TACAGATGAC AGGAGATCGC CAGATATGCG
361 AGGCCGACCA AGAGAGGAGA CATGGCTACC GCGGCGTGT CAGATAGATG CAGATGGCAC
421 AGTACGGCTA AGAGGCGAC CTACGAGAG ACAGGAGGGT TGCTGGCGAC GCGGCGATAT
481 GCCGCGTACT GTACGGCTAT GCGGCGGCT CCGACGCGG GAGGAGCTGG CCGGGGGGC
541 CGCGCGGCAGT ACGCGTGGG CACAGAAGTG ATGGAGCTGG TGGGAGATGC CACAGAGGCC
601 ATAGACGCTC GGGATTTTCG GAGGGCGAG AGACGGACGA AGACGGAGAT AGCCTATAGG
661 CGATCGGCA AGTACTGAGC AGACTGAGCG CCGAGTCTGC CCGAGAGGGC ATGATGAGAT
721 CGAGCGTCGGG AGACGGAAAA CGCGGGCAGG CGCGGTTACG AGAGCTGAGC CAGGCTGGCC
781 AGAGCCTGCCC TGACCTGAG GCGTGCTGTA GCCACAGCT CTCCGGGCGG CGCGGCTCTG
841 TACGGGGCGG CCGGGGGGTC GCGGATCGAGG AGGCTATCG CCGGCTGCTG
901 ATGCGATCCT TTACGAGCTC GCGAGAATCT CAGGTCAGCA CGTGATATGC ACGGACGCG
961 AGGCCGCGAGT ACAAGGCACA GCTGCTGAG ACGCGGCGG ACAGGCGGGC CGCGGCGGAC
1021 CGCGGCGGG TGGTTTATT GAGGGACAGA AGGCTCTGCC CGCGGGGCA CGCTGGCTACT
1081 AGGCCGCGCC AGAAGGCTCT CAAGGAGAAAC AGAGGACGAGC TGGGAGATAG TACGAGATAG
1141 CGCGGCGGCT GCTACGAGCC CAGCGAGATG AGGGAGCTGG GCGGAGACCA CAGAGAGGCC
1201 GCCGCGGGCA GCCGAGACAG GTCCCGGGG ACCTGAGCT CAGGAGAGCA CGCGGCGGCTTG
1261 GTAGAGAAGT CTGTTGAGCG CGCGAGGTCT GCGACATAG ATGAGGAGAG GACGGAGGAGAT
```
[0126] In certain embodiments described herein, a codon-optimized coding region encoding SEQ ID NO:4 is optimized according to codon usage in humans (*Homo sapiens*). Alternatively, a codon-optimized coding region encoding SEQ ID NO:4 may be optimized according to codon usage in any plant, animal, or microbial species. Codon-optimized coding regions encoding SEQ ID NO:4, optimized according to codon usage in humans are designed as follows. The amino acid composition of SEQ ID NO:4 is shown in Table 8.

**TABLE 8**

<table>
<thead>
<tr>
<th>AMINO ACID</th>
<th>Number in SEQ ID NO:4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Ala 25</td>
</tr>
<tr>
<td>R</td>
<td>Arg 17</td>
</tr>
<tr>
<td>C</td>
<td>Cys 3</td>
</tr>
<tr>
<td>G</td>
<td>Gly 16</td>
</tr>
<tr>
<td>H</td>
<td>His 5</td>
</tr>
<tr>
<td>I</td>
<td>Ile 11</td>
</tr>
<tr>
<td>L</td>
<td>Leu 26</td>
</tr>
<tr>
<td>K</td>
<td>Lys 13</td>
</tr>
<tr>
<td>M</td>
<td>Met 14</td>
</tr>
<tr>
<td>F</td>
<td>Phe 7</td>
</tr>
<tr>
<td>P</td>
<td>Pro 8</td>
</tr>
<tr>
<td>S</td>
<td>Ser 18</td>
</tr>
<tr>
<td>T</td>
<td>Thr 18</td>
</tr>
<tr>
<td>W</td>
<td>Trp 1</td>
</tr>
<tr>
<td>Y</td>
<td>Tyr 5</td>
</tr>
<tr>
<td>V</td>
<td>Val 16</td>
</tr>
<tr>
<td>N</td>
<td>Asn 11</td>
</tr>
<tr>
<td>D</td>
<td>Asp 6</td>
</tr>
<tr>
<td>Q</td>
<td>Gln 15</td>
</tr>
<tr>
<td>E</td>
<td>Glu 17</td>
</tr>
</tbody>
</table>

[0127] Using the amino acid composition shown in Table 8, a human codon-optimized coding region which encodes SEQ ID NO:4 can be designed by any of the methods discussed herein. For “uniform” optimization, each amino acid is assigned the most frequent codon used in the human genome for that amino acid. According to this method, codons are assigned to the coding region encoding SEQ ID NO:4 as follows: the 7 phenylalanine codons are TTC, the 26 leucine codons are CTG, the 11 isoleucine codons are ATC, the 14 methionine codons are ATG, the 16 valine codons are GTG, the 18 serine codons are AGC, the 8 proline codons are CCC, the 18 threonine codons are ACC, the 25 alanine codons are GCC, the 5 tyrosine codons are TAC, the 5 histidine codons are CAC, the 15 glutamine codons are CAG, the 11 asparagine codons are AAC, the 13 lysine codons are AAG, the 6 aspartic acid codons are GAC, the 17 glutamic acid codons are GAG, the 3 tryptophan codon is TGG, the 17 arginine codons are CGG, AGA, or AGG (the frequencies of usage of these three codons in the human genome are not significantly different), and the 16 glycine codons are GGC. The codon-optimized coding region designed by this method is presented herein as SEQ ID NO:27.

[0128] Alternatively, a human codon-optimized coding region which encodes SEQ ID NO:4 can be designed by the “full optimization” method, where each amino acid is assigned codons based on the frequency of usage in the human genome. These frequencies are shown in Table 8 above. Using this latter method, codons are assigned to the coding region encoding SEQ ID NO:4 as follows: about 3 of the 7 phenylalanine codons are TTC, and about 3 of the 26 leucine codons are CTG, about 3 of the 11 isoleucine codons are ATC, and about 2 of the 14 methionine codons are ATG; the 16 valine codons are GTG, about 3 of the leucine codons are CTT, about 5 of the leucine codons are TTG, about 3 of the leucine codons are TTA, about 3 of the leucine codons are TAT, about 3 of the leucine codons are TAA, and about 4 of the 11 isoleucine codons are ATG, about 4 of the 14 methionine codons are ACG, about 3 of the 16 valine codons are GTT, about 4 of the 16 valine codons are GTG, about 2 of the valine codons are GAA, about 8 of the valine codons are GGT, about 3 of the 16 valine codons are GTG, about 2 of the valine codons are GGA, and about 8 of the valine codons are GTT, about 3 of the 16 valine codons are GTG, about 2 of the valine codons are GGA, about 8 of the valine codons are GTT, about 3 of the 16 valine codons are GTG, about 2 of the valine codons are GGA, and about 8 of the valine codons are GTT.
ACG; about 7 of the 25 alanine codons are GCT, about 10 of the alanine codons are GCC, about 6 of the alanine codons are GCA, and about 3 of the alanine codons are GCG; about 2 of the 5 tyrosine codons are TAT and about 3 of the tyrosine codons are TAC; about 2 of the 5 histidine codons are CAT and about 3 of the histidine codons are CAC; about 14 of the 15 glutamine codons are CAA and about 11 of the glutamine codons are CAG; about 5 of the 11 asparagine codons are AAT and about 6 of the asparagine codons are AAC; about 5 of the 13 lysine codons are AAA and about 8 of the lysine codons are AAG; about 3 of the 6 aspartic acid codons are GAT and about 3 of the aspartic acid codons are GAC; about 7 of the 17 glutamic acid codons are GAA and about 10 of the glutamic acid codons are GAG; about 1 of the 3 cysteine codons is TGT and about 2 of the cysteine codons is TCG; the 1 tryptophan codons is TGG; about 1 of the 17 arginine codons is CGT, about 3 of the arginine codons are CGC, about 2 of the arginine codons are CGA, about 4 of the arginine codons are CGG, about 3 of the arginine codons are AGA, and about 3 of the arginine codons are AGG; and about 3 of the 16 glycine codons are GGT, about 6 of the glycine codons are GGC, about 4 of the glycine codons are GGA, and about 4 of the glycine codons are GGG.

[0129] As described above, the term “about” means that the number of amino acids encoded by a certain codon may be one more or one less than the number given. It would be understood by those of ordinary skill in the art that the total number of any amino acid in the polypeptide sequence must remain constant, therefore, if there is one “more” of one codon encoding a give amino acid, there would have to be one “less” of another codon encoding that same amino acid.

[0130] A representative “fully optimized” codon-optimized coding region encoding SEQ ID NO:4, optimized according to codon usage in humans is presented herein as SEQ ID NO:26:

```
AGATGCTTTGCTAAGACAGAGTGGATAGCCACTATTTCTCTAG
CGGCCCTTTAAAAGCCGAAATGCTTCCAGCTGAATCTTTTCGG
GCaGACACAAGCGACGGCTGTAATGAGAGTAGGCTTCAGAGG
ATTCGAGCCCGGCTGACTAAAGGAAAATCTGCGCTGGTTTATATTGAG
CGGCTGCTTCAGAGGAGGGTCTCCAAAGAACGGGCCTGCGAGAAGCCT
TAAAACGGAAGGGGCACCCAAATATTTGTAAAGCCGAGAATTATAT
GCGAAATTTAAGGCGGAAATACTCTTCGAGTGAAGCAGGATCTCTCC
GGTTCTACGTCGAGTGCTTCACGTGATGAACTTCTACACTACCC
GATGGGGCGGCCGTACACAGAAAGTGCTCGGTCCCTGCTGTGGCTGGCCACT
TGCGAAGAATGCTTGACGACGCGCGGCCATCCACAAAATGTACAC
CACACAAGACCTGGATGACTACAGAAAGATGCGTGGTTCTACATCTGAG
CTACAGCGAAAGCAATGAAACAAAFGGCGGCGACAGCTGCGAGGCTTCC
GAGGCGGCGGTAGTGGCGCTCCCGAGGCGGAGAAATGACGCGAGGTAGGAG
AATCTGCTGACAGAGCAGCCGCTTCCCTACCTGCGTAGGAGAGTCTTCC
TTGAGAACTCGAGGCGCTACAGAAGCGGATGCGTGTCACAGTAGAGGA
```

[0131] Additionally, a minimally codon-optimized nucleotide sequence encoding SEQ ID NO:4 can be designed by changing only certain codons found more frequently in IV genes than in human genes, as shown in Table 7. For example, if it is desired to substitute more frequently used codons in humans for those codons that occur at least 2 times more frequently in IV genes (designated with an asterisk in Table 7), Arg AGA, which occurs 2.3 times more frequently in IV genes than in human genes, is changed to, e.g., CGG; Asn AAT, which occurs 2.0 times more frequently in IV genes than in human genes, is changed to, e.g., AAC; Ile AUA, which occurs 3.6 times more frequently in IV genes than in human genes, is changed to, e.g., ATC; and Leu CTA, which occurs 2.0 times more frequently in IV genes than is human, is changed to, e.g., CTG.

[0132] In another form of minimal optimization, a Codon Usage Table (CUT) for the specific IV sequence in question is generated and compared to CUT for human genomic DNA (see Table 7, supra). Amino acids are identified for which there is a difference of at least 10 percentage points in codon usage between human and IV DNA (either more or less). Then the wild type IV codon is modified to conform to predominant human codon for each such amino acid. Furthermore, the remainder of codons for that amino acid are also modified such that they conform to the predominant human codon for each such amino acid.

[0133] A representative “minimally optimized” codon-optimized coding region encoding SEQ ID NO:4, minimally optimized according to codon usage in humans by this latter method, is presented herein as SEQ ID NO:28:

```
AGTAGATATCCGACTGACGAGCTGGATGCTACACGATCTCTAG
AGGCCCTTTAAAAGCCGAAATGCTTCCAGCTGAATCTTTTCGG
GCaGACACAAGCGACGGCTGTAATGAGAGTAGGCTTCAGAGG
ATTCGAGCCCGGCTGACTAAAGGAAAATCTGCGCTGGTTTATATTGAG
CGGCTGCTTCAGAGGAGGGTCTCCAAAGAACGGGCCTGCGAGAAGCCT
TAAAACGGAAGGGGCACCCAAATATTTGTAAAGCCGAGAATTATAT
GCGAAATTTAAGGCGGAAATACTCTTCGAGTGAAGCAGGATCTCTCC
GGTTCTACGTCGAGTGCTTCACGTGATGAACTTCTACACTACCC
GATGGGGCGGCCGTACACAGAAAGTGCTCGGTCCCTGCTGTGGCTGGCCACT
TGCGAAGAATGCTTGACGACGCGCGGCCATCCACAAAATGTACAC
CACACAAGACCTGGATGACTACAGAAAGATGCGTGGTTCTACATCTGAG
CTACAGCGAAAGCAATGAAACAAAFGGCGGCGACAGCTGCGAGGCTTCC
GAGGCGGCGGTAGTGGCGCTCCCGAGGCGGAGAAATGACGCGAGGTAGGAG
AATCTGCTGACAGAGCAGCCGCTTCCCTACCTGCGTAGGAGAGTCTTCC
TTGAGAACTCGAGGCGCTACAGAAGCGGATGCGTGTCACAGTAGAGGA
```

[0134] In certain embodiments described herein, a codon-optimized coding region encoding SEQ ID NO:5 is opti-
mized according to codon usage in humans (Homo sapiens).
Alternatively, a codon-optimized coding region encoding
SEQ ID NO:5 may be optimized according to codon usage
in any plant, animal, or microbial species. Codon-optimized
coding regions encoding SEQ ID NO:5, optimized according
to codon usage in humans are designed as follows. The
amino acid composition of SEQ ID NO:5 is shown in Table
9.

<table>
<thead>
<tr>
<th>AMINO ACID</th>
<th>Number in SEQ ID NO:5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td>R</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>Cys</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>G</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>8</td>
</tr>
<tr>
<td>H</td>
<td>His</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>I</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>8</td>
</tr>
<tr>
<td>L</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td>K</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td>M</td>
<td>Met</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>F</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td>P</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td>S</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>7</td>
</tr>
<tr>
<td>T</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td>W</td>
<td>Trp</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Y</td>
<td>Tyr</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>V</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td>N</td>
<td>Asn</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>D</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Q</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>E</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td>9</td>
</tr>
</tbody>
</table>

[0135] Using the amino acid composition shown in Table
9, a human codon-optimized coding region which encodes
SEQ ID NO:5 can be designed by any of the methods
discussed herein. For “uniform” optimization, each amino
acid is assigned the most frequent codon used in the human
genome for that amino acid. According to this method,
codons are assigned to the coding region encoding SEQ ID
NO:5 as follows: the 4 phenylalanine codons are TTC, the
10 leucine codons are CTG, the 8 isoleucine codons are
ATC, the 2 methionine codons are AEG, the 4 valine codons
are GTG, the 7 serine codons are AGC, the 4 proline codons
are CCC, the 4 threonine codons are ACC, the 5 alanine
codons are GCC, the 3 tyrosine codons are TAC, the 2
histidine codons are CAC, the 2 glutamine codons are CAG,
the 3 asparagine codons are AAC, the 5 lysine codons are
AAG, the 5 aspartic acid codons are GAC, the 9 glutamic
codons are GAG, the 2 tryptophan codons are TGG, the
7 arginine codons are CGG, AGA, or AGG (the frequencies
of usage of these three codons in the human genome are
not significantly different), and the 8 glycine codons are
GGC. The codon-optimized PA coding region designed by this
method is presented herein as SEQ ID NO:30:

1ATGAGCCCTGC TGGACGGAGT GAGACCCCC ATCCGGAAAC AGTGGGCTCG CCGTGCAC
61CAGCACAGCC ACCTCCCTGC CATGCCGAC AACAATACG GCATCTTGCA CTCGACCTCG
121TGATGCCCTG ACCTCCCTCG CTCTCAAAGC ATTCAGCGCC GTCAGAAGT CAAGCTCGAG
181GCGCCGCCCA GCACAGCCGC CATGGCAGCG AAGATACGCG AGAGATACG GAGAGAGCAG
241CGAGACGGCG CGGGCCGCAA TCTGAGCAA CTGAGCTGGA GTGA
As described above, the term “about” means that the number of amino acids encoded by a certain codon may be one more or one less than the number given. It would be understood by those of ordinary skill in the art that the total number of any amino acid in the polypeptide sequence must remain constant, therefore, if there is one “more” of one codon encoding a give amino acid, there would have to be one “less” of another codon encoding that same amino acid.

A representative “fully optimized” codon-optimized coding region encoding SEQ ID NO:5, optimized according to codon usage in humans is presented herein as SEQ ID NO:29:

[0139] Additionally, a minimally codon-optimized nucleotide sequence encoding SEQ ID NO:5 can be designed by changing only certain codons found more frequently in IV genes than in human genes, as shown in Table 7. For example, if it is desired to substitute more frequently used codons in humans for those codons that occur at least 2 times more frequently in IV genes (designated with an asterisk in Table 7), Arg AGA, which occurs 2.3 times more frequently in IV genes than in human genes, is changed to, e.g., CGG; Asn AAT, which occurs 2.0 times more frequently in IV genes than in human genes, is changed to, e.g., AAC; Ile ATA, which occurs 3.6 times more frequently in IV genes than in human genes, is changed to, e.g., ATC; and Leu CTA, which occurs 2.0 times more frequently in IV genes than is human, is changed to, e.g., CTG.

[0140] In another form of minimal optimization, a Codon Usage Table (CUT) for the specific IV sequence in question is generated and compared to CUT for human genomic DNA (see Table 7, supra). Amino acids are identified for which there is a difference of at least 10 percentage points in codon usage between human and IV DNA (either more or less). Then the wild type IV codon is modified to conform to predominant human codon for each such amino acid. Furthermore, the remainder of codons for that amino acid are also modified such that they conform to the predominant human codon for each such amino acid.

A representative “minimally optimized” codon-optimized coding region encoding SEQ ID NO:5, minimally optimized according to codon usage in humans by this latter method, is presented herein as SEQ ID NO:31:
[0142] In certain embodiments described herein, a codon-optimized coding region encoding SEQ ID NO:7 is optimized according to codon usage in humans (Homo sapiens). Alternatively, a codon-optimized coding region encoding SEQ ID NO:7 may be optimized according to codon usage in any plant, animal, or microbial species. Codon-optimized coding regions encoding SEQ ID NO:7, optimized according to codon usage in humans are designed as follows. The amino acid composition of SEQ ID NO:7 is shown in Table 10.

<table>
<thead>
<tr>
<th>AMINO ACID</th>
<th>Number in SEQ ID NO:7</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>20</td>
</tr>
<tr>
<td>R</td>
<td>51</td>
</tr>
<tr>
<td>C</td>
<td>8</td>
</tr>
<tr>
<td>G</td>
<td>45</td>
</tr>
<tr>
<td>H</td>
<td>6</td>
</tr>
<tr>
<td>I</td>
<td>27</td>
</tr>
<tr>
<td>L</td>
<td>35</td>
</tr>
<tr>
<td>K</td>
<td>21</td>
</tr>
<tr>
<td>M</td>
<td>26</td>
</tr>
<tr>
<td>F</td>
<td>18</td>
</tr>
<tr>
<td>P</td>
<td>18</td>
</tr>
<tr>
<td>S</td>
<td>43</td>
</tr>
<tr>
<td>T</td>
<td>30</td>
</tr>
<tr>
<td>W</td>
<td>7</td>
</tr>
<tr>
<td>Y</td>
<td>15</td>
</tr>
<tr>
<td>V</td>
<td>24</td>
</tr>
<tr>
<td>N</td>
<td>28</td>
</tr>
<tr>
<td>D</td>
<td>23</td>
</tr>
<tr>
<td>Q</td>
<td>21</td>
</tr>
<tr>
<td>E</td>
<td>39</td>
</tr>
</tbody>
</table>

[0143] Using the amino acid composition shown in Table 10, a human codon-optimized coding region which encodes SEQ ID NO:7 can be designed by any of the methods discussed herein. For “uniform” optimization, each amino acid is assigned the most frequent codon used in the human genome for that amino acid. According to this method, codons are assigned to the coding region encoding SEQ ID NO:7 as follows: the 18 phenylalanine codons are TTC, the 35 leucine codons are CTG, the 27 isoleucine codons are ATC, the 26 methionine codons are ATG, the 24 valine codons are GTG, the 43 serine codons are AGC, the 18 proline codons are CCC, the 30 threonine codons are ACC, the 39 alanine codons are GCC, the 15 tyrosine codons are TAC, the 6 histidine codons are CAC, the 21 glutamine codons are CAG, the 28 asparagine codons are AAC, the 21 lysine codons are AAG, the 23 aspartic acid codons are GAC, the 39 glutamic acid codons are GAG, the 7 tryptophan codons are TGG, the 51 arginine codons are CAG, AGA, or AGG (the frequencies of usage of these three codons in the human genome are not significantly different), and the 43 glycine codons are GGC. The codon-optimized PA coding region designed by this method is presented herein as SEQ ID NO:33:

ATGAGCTCATGACCACCCAGTAGATCAATGAGGAGCCCCATACAGGAGACGAGAGGCTGGGCT
GCAGGTGCAACGGCCAGCAAGCTGAAACAGCAGCACGCCCTACGAGGACAAAGGACAG
TACGAGCAGTGAGACCCGAGCGGAGGAGCAGCAGCCCGAGATTACAG

[0144] Alternatively, a human codon-optimized coding region which encodes SEQ ID NO:7 can be designed by the “full optimization” method, where each amino acid is assigned codons based on the frequency of usage in the human genome. These frequencies are shown in Table 10 above. Using this latter method, codons are assigned to the coding region encoding SEQ ID NO:7 as follows: about 8 of the 18 phenylalanine codons are TTG; about 10 of the 27 leucine codons are CTC; about 3 of the 35 leucine codons are TTA; about 4 of the 24 valine codons are GTT; about 5 of the leucine codons are CTT; about 7 of the leucine codons are TAC, about 2 of the leucine codons are CTA, and about 14 of the leucine codons are CTG; about 10 of the 27 isoleucine codons are ATG; about 13 of the isoleucine codons are ATT; about 13 of the isoleucine codons are ATT; and about 4 of the isoleucine codons are ATT; the 26 methionine codons are ATG; about 4 of the 24
valine codons are GGT, about 6 of the valine codons are GTG, about 3 of the valine codons are GTA, and about 11 of the valine codons are GGA; about 8 of the 43 serine codons are CTC; about 9 of the serine codons are TCC; about 6 of the serine codons are TCA, about 2 of the serine codons are TCG, about 6 of the serine codons are AGT, and about 10 of the serine codons are AGG; about 5 of the 18 proline codons are CCT; about 6 of the proline codons are CCC, about 5 of the proline codons are CCA, and about 2 of the proline codons are CGG; about 7 of the 30 threonine codons are ACT, about 11 of the threonine codons are ACC, about 8 of the threonine codons are ACA, and about 4 of the threonine codons are ACG; about 10 of the 39 alanine codons are GGT, about 16 of the 16 alanine codons are GCA, and about 4 of the alanine codons are GCC; about 7 of the 15 tyrosine codons are TAT and about 8 of the tyrosine codons are TAC; about 2 of the 6 histidine codons are CAT and about 4 of the histidine codons are CAG; about 5 of the 21 glutamine codons are CAU and about 16 of the glutamine codons are CAG; about 13 of the 28 asparagine codons are AAI and about 15 of the asparagine codons are AAC; about 9 of the 21 lysine codons are AAA and about 12 of the lysine codons are AAG; about 11 of the 23 aspartic acid codons are GAT and about 12 of the aspartic acid codons are GAC; about 16 of the 39 glutamic acid codons are GAA and about 23 of the glutamic acid codons are GAG; about 4 of the 8 cysteine codons are TGT and about 4 of the cysteine codons are TGC; the 7 tryptophan codons are TGG; about 4 of the 51 arginine codons are CGT, about 10 of the arginine codons are CGG, about 6 of the arginine codons are CGA, about 11 of the arginine codons are AGA, about 10 of the arginine codons are AGG; and about 7 of the 43 glycine codons are GGT, about 15 of the glycine codons are GGC, about 11 of the glycine codons are GGA, and about 11 of the glycine codons are GGG.

[0145] As described above, the term “about” means that the number of amino acids encoded by a certain codon may be one more or one less than the number given. It would be understood by those of ordinary skill in the art that the total number of any amino acid in the polypeptide sequence must remain constant, therefore, if there is one “more” of one codon encoding a give amino acid, there would have to be one “less” of another codon encoding that same amino acid.

[0146] A representative “fully optimized” codon-optimized coding region encoding SEQ ID NO:7, optimized according to codon usage in humans is presented herein as SEQ ID NO:52:

```
ATGAGCCCTTCTCACAAAGATGCAAACATTATCAGAAGATGAAAGGATG
CGAGTGCAATGCGTCGATGTATGCGCTCTCAGGGCTGAAAAAGAAGCT
ACGAAGCAAAAGGAAGAGTGGAGAGAACGCAAGAACGAAATAGA
GCATCCGGCCGGAGAAGATGAGAACACTGCAATCACTGAGAT
GTGGCAACAGCTAAACATCAGGCTATGAGGGAGGCGATNACAAAGAAAFA
GGCTATGCGTCAGGGAGATGCTGCTGCTGATATTGAGGAGAACGAGA
AAATACCTGGAAGACACGCCCTGGGAAGAGACACACAAAAAAACTTG
AGTGCCGATAACGCGAGAAGATGGTGGTTGGTTGGTTGGTTGGTTGGT
```

---continued---

```
TUTATGATAAGAGAAATAGACCGCCTCTGGCGACGCTAAATATGGA
GACAGCCGCTACTGCGGCTACACGATATGTCATCGTTACCTATT
GAATAGGCAACACGCCAAGAACCGCGACACTGTCGCGAGGGAATG
ACCCCGATATATCTCCTAGGCTACGTTACACCTGAGGGAATG
GAGCAGCTGGACAGCAGTTAAGGCTCTGAGAAGCTGAGATGAGCT
GATGGAATATAGAAAGCGGATGAAATTCGAGCTACCTTGCGAGG
AAACGCGGCAAAACTAGTTGACATAGAGGAGATGTGCAAGATCTCCT
AAACAAAGGAAATCCGCTCAGCCACCAAAAGAAGGCGGATACTAG
CGAAAGTACTGGTGATCCTGAGAAATGCGATAGCTTGACACCTCTG
CGGAGGAGCCCTCCTGCTCAGCGCTAGGCGCTATGCTGCATGGAT
CTGTCCTGCTGTATATCGTCTGCTGCAGCTGAGTCTGCTGCGAGG
AGGGCGCCCTCCGTTGGATAGTGCATATGCTAGCTGAGTCTAGG
CTGTCCTCTATAGAAATGGCTACGCTGAGCGCTAGCTGAGTCTAGG
CTGTCCTCTATAGAAATGGCTACGCTGAGCGCTAGCTGAGTCTAGG
CTGTCCTCTATAGAAATGGCTACGCTGAGCGCTAGCTGAGTCTAGG
CTGTCCTCTATAGAAATGGCTACGCTGAGCGCTAGCTGAGTCTAGG
```

[0147] In certain embodiments described herein, a codon-optimized coding region encoding SEQ ID NO:9 is optimized according to codon usage in humans (Homo sapiens). Alternatively, a codon-optimized coding region encoding SEQ ID NO:9 may be optimized according to codon usage in any plant, animal, or microbial species. Codon-optimized coding regions encoding SEQ ID NO:9, optimized according to codon usage in humans are designed as follows. The amino acid composition of SEQ ID NO:9 is shown in Table 11.

<table>
<thead>
<tr>
<th>AMINO ACID</th>
<th>Number in SEQ ID NO: 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>39</td>
</tr>
<tr>
<td>R</td>
<td>51</td>
</tr>
<tr>
<td>C</td>
<td>8</td>
</tr>
<tr>
<td>G</td>
<td>43</td>
</tr>
<tr>
<td>H</td>
<td>6</td>
</tr>
<tr>
<td>I</td>
<td>27</td>
</tr>
<tr>
<td>L</td>
<td>35</td>
</tr>
<tr>
<td>K</td>
<td>21</td>
</tr>
</tbody>
</table>

Table 11
TABLE 11-continued

<table>
<thead>
<tr>
<th>AMINO ACID</th>
<th>Number in SEQ ID NO: 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>Met</td>
</tr>
<tr>
<td>F</td>
<td>Phe</td>
</tr>
<tr>
<td>P</td>
<td>Pro</td>
</tr>
<tr>
<td>S</td>
<td>Ser</td>
</tr>
<tr>
<td>T</td>
<td>Thr</td>
</tr>
<tr>
<td>W</td>
<td>Tyr</td>
</tr>
<tr>
<td>Y</td>
<td>Tyr</td>
</tr>
<tr>
<td>V</td>
<td>Val</td>
</tr>
<tr>
<td>N</td>
<td>Asn</td>
</tr>
<tr>
<td>D</td>
<td>Asp</td>
</tr>
<tr>
<td>Q</td>
<td>Gln</td>
</tr>
<tr>
<td>E</td>
<td>Gla</td>
</tr>
</tbody>
</table>

[0148] Using the amino acid composition shown in Table 11, a human codon-optimized coding region which encodes SEQ ID NO:9 can be designed by any of the methods discussed herein. For “uniform” optimization, each amino acid is assigned the most frequent codon used in the human genome for that amino acid. According to this method, codons are assigned to the coding region encoding SEQ ID NO:9 as follows: the 18 phenylalanine codons are TTC, the 35 leucine codons are CTA, the 27 isoleucine codons are ATT, the 26 methionine codons are ATG, the 24 valine codons are GTG, the 43 serine codons are AGC, the 18 proline codons are CCC, the 30 threonine codons are ACC, the 39 alanine codons are GCC, the 15 tyrosine codons are TAC, the 6 histidine codons are CAC, the 21 glutamine codons are CAG, the 28 asparagine codons are AAC, the 21 lysine codons are AAG, the 23 aspartic acid codons are GAC, the 39 glutamic acid codons are GAG, the 7 tryptophan codons are TGG, the 51 arginine codons are CGG, AGA, or AGG (the frequencies of usage of these three codons in the human genome are not significantly different), and the 43 glycine codons are GGC. The codon-optimized PA coding region designed by this method is presented herein as SEQ ID NO:35:

ATGCCCACCAAGCGCCACCAAGCGAACTCGAAGCCGATTGAGAGCGCAAGGAG
CGAGGAGACAGAACCACCATCAAGAGCGGACCGCTGCCAGAAGTCATG
GGCCGATCGCAATCTGTTGTTTACAGGAAAGCTGCGGCAAGGCTTAAAGTAC
GATACAGGGCAGGATGTAAGCAAGAAAATCCTAGCAAGGAGATTAGT
GCTGAGCCCTCGAAGAAAGAGGACACATCTGAGGAGGCACACCCCA
GCCGCCGCAAGGCCAAGAGAGCGCGGCGCCATCTACAGGAGGATG
GACCGCGCGGCTGAGAGAGCAAGGGCNCTCTGAGCAAGGAGGAGATCA
GGGAGATCTGAGCGCGGACCGCAACGGCAAGCCGACGCCGCGCCGCGGCTTG
ACCCACACATGATCTGAGCGACGACAACCTGACAGGCGCGCCACCTGAG
GACCGCGCGGCTGAGAGAGCAAGGGCNCTCTGAGCAAGGAGGAGATCA
TGGAGCGGAGCCACCTGCGGAGCGAGGACGCGGCGCCGCGGCGGCGGCTTG
AGGGCGCTGAGCGCGGACCGCAACGGCAAGCCGACGCCGCGCCGCGGCTTG
CATACGCGGAGAGAATCTGAGGAGGCAGCGGCGCCGCGGCGGCGGCGGCTTG
TGGAGCGGAGCCACCTGCGGAGCGAGGACGCGGCGCCGCGGCGGCTTG

[0149] Alternatively, a human codon-optimized coding region which encodes SEQ ID NO:9 can be designed by the “full optimization” method, where each amino acid is assigned codons based on the frequency of usage in the human genome. These frequencies are shown in Table 11 above. Using this latter method, codons are assigned to the coding region encoding SEQ ID NO:9 as follows: about 8 of the 18 phenylalanine codons are TTT, and about 10 of the phenylalanine codons are TTC; about 3 of the 35 leucine codons are TTA, about 4 of the leucine codons are TTG, about 5 of the leucine codons are CTT, about 7 of the leucine codons are CTG, about 2 of the leucine codons are CTA, and about 14 of the leucine codons are CTG; about 10 of the 27 isoleucine codons are ATT, about 13 of the isoleucine codons are ATC, and about 4 of the isoleucine codons are ACT; the 26 methionine codons are ATG; about 4 of the 24 valine codons are GTT, about 6 of the valine codons are GTG, about 3 of the valine codons are GTA, and about 11 of the valine codons are GGA; about 8 of the 43 serine codons are TCT, about 9 of the serine codons are TCC, about 6 of the serine codons are TCA, about 2 of the serine codons are CGT, about 6 of the serine codons are CTT, about 7 of the serine codons are CTC, about 7 of the serine codons are CCT, about 7 of the serine codons are CCA, and about 2 of the serine codons are CCG; about 5 of the 18 proline codons are CCG; about 7 of the 30 threonine codons are ACC, about 11 of the threonine codons are ACC, about 8 of the threonine codons are ACA, and about 4 of the threonine codons are ACC; about 10 of the 39 alanine codons are AGG, about 16 of the alanine codons are GCC, about 9 of the alanine codons are GCA, and about 4 of the alanine codons are GCG; about 7 of the 15 tyrosine codons...
are TAT and about 8 of the tyrosine codons are TAC; about 
2 of the 6 histidine codons are CAT and about 4 of the 
histidine codons are CAC; about 5 of the 21 glutamine 
codons are CAA and about 16 of the glutamine codons 
are CAG; about 13 of the 28 asparagine codons are AAT 
and about 15 of the asparagine codons are AAC; about 9 of 
the 21 lysine codons are AAA and about 12 of the lysine 
codons are AAG; about 11 of the 23 aspartic acid codons 
are GAT and about 12 of the aspartic acid codons are GAC. 
About 16 of the 39 glutamic acid codons are GAA and about 23 of 
the glutamic acid codons are GAG; about 4 of the 8 cysteine 
codons are TGT and about 4 of the cysteine codons are TGC; 
the 7 tryptophan codons are TGG; about 4 of the 51 arginine 
codons are CGT; about 10 of the arginine codons are CGC, 
about 6 of the arginine codons are CGA, about 11 of the 
arginine codons are CGG, about 10 of the arginine codons 
are AGA, and about 10 of the arginine codons are AGG; and 
about 7 of the 43 glycine codons are GGT, about 15 of the 
glycine codons are GGC, about 11 of the glycine codons 
are GGA, and about 11 of the glycine codons are GGG.

[0150] As described above, the term "about" means that 
the number of amino acids encoded by a certain codon may 
be one more or one less than the number given. It would be 
understood by those of ordinary skill in the art that the total 
number of any amino acid in the polypeptide sequence must 
remain constant, therefore, if there is one "more" of one 
codon encoding a give amino acid, there would have to be 
one "less" of another codon encoding that same amino acid.

[0151] A representative "fully optimized" codon-optimized 
coding region encoding SEQ ID NO:9, optimized 
coding according to us humans is presented herein as 
SEQ ID NO:34:

```
ATGCCAGCCAGCCGACAAAAGCAGTCCGAGACAGAGACATGTG
TGAGGGGCAAGGACCAACGAGCACTCCGGCCCTGGCAAGATAGTGG
GTGGGCAACAGGATGATTCAGTTGCTAGCGCATCTTGGGTG
GATTACAGGCGCTTATTAACGAGACTCCGATCTCGACAGGGATGT
CTTGGACCTCAGGAGGCGTTTAACTAGTACGCCAGACAGGTGCT
CTGTCGAGAAGAGCAGACGCTACGGGCACTCTACCGAGGATGT
AACGAGAAGGAAGGAGCAGAGGCTACGAGAGAGAGATCGG
TAGTGCCGAGACAGCTTAAATAGGGATGAGCTCAGCTGCGTG
CCCAATTGGATATGGCATAGCAGAATGCAGCAGGACATTCACGAC
ACTGAGACATCCGAGGACCTGGTACGAGCCACACATCTTCAGAT
GCAAGTCAAGACATGGTCCTGGAAGCCGAGCGCCGGCCCTGG
AGGGGTGCGCAGAATGGTGAAGGACTCGTGCAATCTGCAAAGGAG
AATCAATGACAGAAGATTTTGCGCGAGGAAAAAGCGCGAAGAGCGA
TGCTAGCCAGCGTGCTTAACATTAAAAGGAAATACCTAGCTGAG
CCCAAGAACGATGAGCCCGAACCAAGATGGAGAGAAAGAAAAGCGAAT
GCGAGTTTGAAGCAGCCGCTTTGCTGAGAAGAGCTGCTACGCCG
GCGTCTGCGCGCCACAGAAGTCCGCTCCTCCGGTCCTTACAGGCGCG
CGCTGCGAGFGCTCAGATTTTGCGAGGAGGTTACTCTGCTACG
```

[0152] In certain embodiments described herein, a codon-optimized 
coding region encoding SEQ ID NO:16 is optimized 
coding according to us humans (Homo sapiens). 
Alternatively, a codon-optimized coding region encoding 
SEQ ID NO:16 may be optimized according to codon usage 
in any plant, animal, or microbial species. Codon-optimized 
coding regions encoding SEQ ID NO:16, optimized 
coding according to codon usage in humans are designed as follows. The 
amino acid composition of SEQ ID NO:16 is shown in Table 12.

<table>
<thead>
<tr>
<th>AMINO ACID</th>
<th>Number in</th>
<th>Seq ID NO: 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

[0153] Using the amino acid composition shown in Table 12, 
a human codon-optimized coding region which encodes 
SEQ ID NO: 16 can be designed by any of the methods 
discussed herein. For "uniform" optimization, each amino 
acid is assigned the most frequent codon used in the human 
genome for that amino acid. According to this method,
codons are assigned to the coding region encoding SEQ ID NO:16 as follows: the 21 phenylalanine codons are TTC, the 39 leucine codons are CTA, the 38 isoleucine codons are ATC, the 27 methionine codons are ATG, the 32 valine codons are GTG, the 40 serine codons are AGC, the 26 proline codons are CCC, the 38 threonine codons are ACC, the 41 alanine codons are GCC, the 14 tyrosine codons are TAC, the 4 histidine codons are CAU, the 19 glutamine codons are CAG, the 25 asparagine codons are AAC, the 52 lysine codons are AAG, the 34 aspartic acid codons are GAC, the 30 glutamic acid codons are GAG, the 1 tryptophan codon is TGG, the 30 arginine codons are CGG, the 30 arginine codons are AGA, and the 30 arginine codons are AGG (the frequencies of usage of these three codons in the human genome are not significantly different), and the 44 glycine codons are GGC. The codon-optimized PA coding region designed by this method is presented herein as SEQ ID NO:37:

-continued

[0154] Alternatively, a human codon-optimized coding region which encodes SEQ ID NO:16 can be designed by the “full optimization” method, where each amino acid is assigned codons based on the frequency of usage in the human genome. These frequencies are shown in Table 12 above. Using this latter method, codons are assigned to the coding region encoding SEQ ID NO:16 as follows: about 10 of the 21 phenylalanine codons are TTC, and about 12 of the phenylalanine codons are TTC; about 3 of the 39 leucine codons are TTA, about 5 of the leucine codons are TTG, about 5 of the leucine codons are CTG, about 8 of the leucine codons are CTC, about 5 of the leucine codons are CTA, and about 16 of the leucine codons are CTG; about 14 of the 38 isoleucine codons are ATT, about 18 of the isoleucine codons are ATC, and about 6 of the isoleucine codons are AIA; the 27 methionine codons are ATG, about 6 of the 32 valine codons are GTG, about 4 of the valine codons are GTA, and about 15 of the valine codons are GTG; about 7 of the 40 serine codons are TCT, about 9 of the serine codons are TTC, about 6 of the serine codons are TCA, about 2 of the serine codons are TCG, about 6 of the serine codons are AGT, and about 10 of the serine codons are AGC; about 7 of the 26 proline codons are CCT, about 9 of the proline codons are CCC, about 7 of the proline codons are CCA, and about 3 of the proline codons are CGG; about 9 of the 38 threonine codons are ACT, about 14 of the threonine codons are ACC, about 11 of the threonine codons are ACA, and about 4 of the threonine codons are ACG; about 11 of the 41 alanine codons are GGT, about 17 of the alanine codons are GCC, about 9 of the alanine codons are GCA, and about 4 of the alanine codons are GCG; about 6 of the 14 tyrosine codons are TAT and about 8 of the tyrosine codons are TAC; about 2 of the 4 histidine codons are CAT and about 2 of the histidine codons are CAC; about 5 of the 19 glutamine codons are CAA and about 14 of the glutamine codons are CAG; about 12 of the 25 asparagine codons are AAT and about 13 of the asparagine codons are AAC; about 22 of the 52 lysine codons are AAA and about 30 of the lysine codons are AAG; about 16 of the 34 aspartic acid codons are GAT and about 18 of the aspartic acid codons are GAC; about 12 of the 30 glutamic acid codons are GAA and about 18 of the glutamic acid codons are GAG; about 2 of the 5 cysteine codons are TGT and about 3 of the cysteine codons are TGC; the single tryptophan codon is TGG; about 2 of the 30 arginine codons are CGG, about 3 of the arginine codons are CGA, about 6 of the arginine codons are CAG, about 6 of the arginine codons are AGG, and about 6 of the arginine codons are AGG; and about 7 of the 44 glycine codons are GGT, about 15 of the glycine codons are GGC, about 11 of the glycine codons are GGA, and about 11 of the glycine codons are GGG.
[0155] As described above, the term “about” means that the number of amino acids encoded by a certain codon may be one more or one less than the number given. It would be understood by those of ordinary skill in the art that the total number of any amino acid in the polypeptide sequence must remain constant, therefore, if there is one “more” of one codon encoding a given amino acid, there would have to be one “less” of another codon encoding that same amino acid.

[0156] A representative “fully optimized” codon-optimized coding region encoding SEQ ID NO:16, optimized according to codon usage in humans is presented herein as SEQ ID NO:36:

```
ATGTCGAAACTATGTGACATCGACCACTTACACACGATGACTTTAGTCAAAAC
CCCGGAGACAGACTAACCCCGTGGACCTTACGAGAAGACGGCCCAAT
AACGCAGCCCTCTCGGCCTTCTAGCAATAGAGGAGCCCGAATCTCAAT
CTTGAAGAGGACACTTACTACATGCGGAACTGATGCTCCGGAATTCA
AAAAAACAAAGCCCGCAAGATAGAAGAAGACTTTTCTTAACTGTTGG
TAAAGCTCGGATTTTTATACGAGATGTGCTACCGGCGCTTAAC
GACAGATGCGACGACCACTTATATACGATGACGACGACGACGACGAT
ACTTCTCGGCGCTCTATAAGCAAAAGAAGCTACTTCATCAAAAGA
ATGCGGGAAAGTGAAGAAAGAAAGAAGAATTCAGCACCAAAACT
GGGCGGCGACATCTTAAAGATGTGCGGCAATGACATGCACTTATTGAG
CCCGCATGATATACCCCTGGAAGGGAGGGATGCTTTAAAACAGGG
GGCATGGGGACACATGAGTTTCTCAGGCGAATTTCAAAAGCCTG
CTACGGGACAGTATACGTTTCTATTTGTGGGCTCGGACGATATGCT
```

[0157] Additionally, a minimally codon-optimized nucleotide sequence encoding SEQ ID NO:16 can be designed by changing only certain codons found more frequently in IV genes than in human genes, as shown in Table 7. For example, if it is desired to substitute more frequently used codons in humans for those codons that occur at least 2 times more frequently in IV genes (designated with an asterisk in Table 7), Arg AGA, which occurs 2.3 times more frequently in IV genes than in human genes, is changed to, e.g., CGG; Asn AAT, which occurs 2.0 times more frequently in IV genes than in human genes, is changed to, e.g., AAC; Ile AIF, which occurs 3.6 times more frequently in IV genes than in human genes, is changed to, e.g., ATC; and Leu CTA, which occurs 2.0 times more frequently in IV genes than in human, is changed to, e.g., CTG.

[0158] In another form of minimal optimization, a Codon Usage Table (CUT) for the specific IV sequence in question is generated and compared to CUT for human genomic DNA (see Table 7, supra). Amino acids are identified for which there is a difference of at least 10 percentage points in codon usage between human and IV DNA (either more or less). Then the wild type IV codon is modified to conform to predominant human codon for each such amino acid. Further, the remainder of codons for that amino acid are also modified such that they conform to the predominant human codon for each such amino acid.

[0159] A representative “minimally optimized” codon-optimized coding region encoding SEQ ID NO:16, minimally optimized according to codon usage in humans by this latter method, is presented herein as SEQ ID NO:38:

```
ATGTCGAAACTATGTGACATCGACCACTTACACACGATGACTTTAGTCAAAAC
CCCGGAGACAGACTAACCCCGTGGACCTTACGAGAAGACGGCCCAAT
AACGCAGCCCTCTCGGCCTTCTAGCAATAGAGGAGCCCGAATCTCAAT
CTTGAAGAGGACACTTACTACATGCGGAACTGATGCTCCGGAATTCA
AAAAAACAAAGCCCGCAAGATAGAAGAAGACTTTTCTTAACTGTTGG
TAAAGCTCGGATTTTTATACGAGATGTGCTACCGGCGCTTAAC
GACAGATGCGACGACCACTTATATACGATGACGACGACGACGACGAT
ACTTCTCGGCGCTCTATAAGCAAAAGAAGCTACTTCATCAAAAGA
ATGCGGGAAAGTGAAGAAAGAAAGAAGAATTCAGCACCAAAACT
GGGCGGCGACATCTTAAAGATGTGCGGCAATGACATGCACTTATTGAG
CCCGCATGATATACCCCTGGAAGGGAGGGATGCTTTAAAACAGGG
GGCATGGGGACACATGAGTTTCTCAGGCGAATTTCAAAAGCCTG
CTACGGGACAGTATACGTTTCTATTTGTGGGCTCGGACGATATGCT
```

[0157] Additionally, a minimally codon-optimized nucleotide sequence encoding SEQ ID NO:16 can be designed by changing only certain codons found more frequently in IV genes than in human genes, as shown in Table 7. For example, if it is desired to substitute more frequently used codons in humans for those codons that occur at least 2 times more frequently in IV genes (designated with an asterisk in Table 7), Arg AGA, which occurs 2.3 times more frequently in IV genes than in human genes, is changed to, e.g., CGG; Asn AAT, which occurs 2.0 times more frequently in IV genes than in human genes, is changed to, e.g., AAC; Ile AIF, which occurs 3.6 times more frequently in IV genes than in human genes, is changed to, e.g., ATC; and Leu CTA, which occurs 2.0 times more frequently in IV genes than is human, is changed to, e.g., CTG.

[0158] In another form of minimal optimization, a Codon Usage Table (CUT) for the specific IV sequence in question is generated and compared to CUT for human genomic DNA (see Table 7, supra). Amino acids are identified for which there is a difference of at least 10 percentage points in codon usage between human and IV DNA (either more or less). Then the wild type IV codon is modified to conform to predominant human codon for each such amino acid. Further, the remainder of codons for that amino acid are also modified such that they conform to the predominant human codon for each such amino acid.

[0159] A representative “minimally optimized” codon-optimized coding region encoding SEQ ID NO:16, minimally optimized according to codon usage in humans by this latter method, is presented herein as SEQ ID NO:38:
[0160] Randomly assigning codons at an optimized frequency to encode a given polypeptide sequence using the “full-optimization” or “minimal optimization” methods, can be done manually by calculating codon frequencies for each amino acid, and then assigning the codons to the polypeptide sequence randomly. Additionally, various algorithms and computer software programs are readily available to those of ordinary skill in the art. For example, the “EditSeq” function in the Lasergene Package, available from DNASTar, Inc., Madison, Wis., the backtranslation function in the VectorNTI Suite, available from InforMax, Inc., Bethesda, Md., and the “backtranslate” function in the GCG—Wisconsin Package, available from Accelrys, Inc., San Diego, Calif. In addition, various resources are publicly available to codon-optimze coding region sequences. For example, the “backtranslation” function found at http://www.entelechon.com/cng/backtranslation.html (visited Jul. 9, 2002), and the backtranseq” function available at http://bioinfo.pbi.nrc.ca:8090/EMBOSS/index.html (visited Oct. 15, 2002). Constructing a rudimentary algorithm to assign codons based on a given frequency can also easily be accomplished with basic mathematical functions by one of ordinary skill in the art.

[0161] A number of options are available for synthesizing codon-optimized coding regions designed by any of the methods described above, using standard and routine molecular biological manipulations well known to those of ordinary skill in the art. In one approach, a series of complementary oligonucleotide pairs of 80-90 nucleotides each in length and spanning the length of the desired sequence are synthesized by standard methods. These oligonucleotide pairs are synthesized such that upon annealing, they form double stranded fragments of 80-90 base pairs, containing cohesive ends, e.g., each oligonucleotide in the pair is synthesized to extend 3, 4, 5, 6, 7, 8, 9, 10, or more bases beyond the region that is complementary to the other oligonucleotide in the pair. The single-stranded ends of each pair of oligonucleotides is designed to anneal with the single-stranded end of another pair of oligonucleotides. The oligonucleotide pairs are allowed to anneal, and approximately five to six of these double-stranded fragments are then allowed to anneal together via the cohesive single stranded ends, and then they ligated together and cloned into a standard bacterial cloning vector, for example, a TOPO vector available from Invitrogen Corporation, Carlsbad, Calif. The construct is then sequenced by standard methods. Several of these constructs consisting of 5 to 6 fragments of 80 to 90 base pair fragments ligated together, i.e., fragments of about 500 base pairs, are prepared, such that the entire desired sequence is represented in a series of plasmid constructs. The inserts of these plasmids are then cut with appropriate restriction enzymes and ligated together to form the final construct. The final construct is then cloned into a standard bacterial cloning vector, and sequenced. Additional methods would be generally apparent to the skilled artisan. In addition, gene expression is readily available commercially.

[0162] The codon-optimized coding regions can be versions encoding any gene products from any strain, derivative, or variant of IV, or fragments, variants, or derivatives of such gene products. For example, nucleic acid fragments of codon-optimized coding regions encoding the NP, M1 and M2 polypeptides, or fragments, variants or derivatives thereof. Codon-optimized coding regions encoding other IV polypeptides or fragments, variants, or derivatives thereof (e.g. HA, NA, PB1, PB2, PA, NS1 or NS2), are included within the present invention. Additional, non-codon-optimized polynucleotides encoding IV polypeptides or other polypeptides are included as well.

Consensus Sequences

[0163] The present invention is further directed to specific consensus sequences of influenza virus proteins, and fragments, derivatives and variants thereof. A “consensus sequence” is, e.g., an idealized sequence that represents the amino acids most often present at each position of two or more sequences which have been compared to each other. A consensus sequence is a theoretical representative amino acid sequence in which each amino acid is the one which occurs most frequently at that site in the different sequences which occur in nature. The term also refers to an actual sequence which approximates the theoretical consensus. A consensus sequence can be derived from sequences which have, e.g., shared functional or structural purposes. It can be defined by aligning as many known examples of a particular structural or functional domain as possible to maximize the homology. A sequence is generally accepted as a consensus when each particular amino acid is reasonably predominant at its position, and most of the sequences which form the basis of the comparison are related to the consensus by rather few substitutions, e.g., from 0 to about 100 substitutions. In general, the wild-type comparison sequences are at least about 50%, 75%, 80%, 90%, 95%, 96%, 97%, 98% or 99% identical to the consensus sequence. Accordingly,
polypeptides of the invention are about 50%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to the consensus sequence. Consensus amino acid sequences can be prepared for any of the influenza antigens. By analyzing amino acid sequences from influenza A strains sequenced since 1990, consensus amino acid sequences were derived for the influenza A NP (SEQ ID NO: 76), M1 (SEQ ID NO: 77) and M2 (SEQ ID NO: 78) proteins (Example 3).

[0164] A "consensus amino acid" is an amino acid chosen to occupy a given position in the consensus protein. A system which is organized to select consensus amino acids can be a computer program, or a combination of one or more computer programs with "by hand" analysis and calculation. When a consensus amino acid is obtained for each position of the aligned amino acid sequences, then these consensus amino acids are "lined up" to obtain the amino acid sequence of the consensus protein.

[0165] Another embodiment of this invention is directed to a process for the preparation of a consensus protein comprising a process to calculate an amino acid residue for nearly all positions of a so-called consensus protein and to synthesize a complete gene from this sequence that could be expressed in a prokaryotic or eukaryotic expression system.

[0166] Polynucleotides which encode the consensus influenza polypeptides, or fragments, variants or derivatives thereof, are also part of this invention. Such polynucleotides can be obtained by known methods, for example by back-translation of the amino acid sequence and PCR synthesis of the corresponding polypeptide.

Compositions and Methods

[0167] In certain embodiments, the present invention is directed to compositions and methods of enhancing the immune response of a vertebrate in need of protection against IV infection by administering in vivo, into a tissue of a vertebrate, one or more polynucleotides comprising at least one codon-optimized coding region encoding an IV polypeptide, or a fragment, variant, or derivative thereof. Polynucleotides may be administered either prior to, at the same time (simultaneously), or subsequent to the administration of the isolated polypeptide.

[0168] The coding regions encoding IV polypeptides or fragments, variants, or derivatives thereof may be codon optimized for a particular vertebrate. Codon optimization is carried out by the methods described herein, for example, in certain embodiments codon-optimized coding regions encoding polypeptides of IV, or nucleic acid fragments of such coding regions encoding fragments, variants, or derivatives thereof are optimized according to the codon usage of the particular vertebrate. The polynucleotides of the invention are incorporated into the cells of the vertebrate in vivo, and an immunologically effective amount of an IV polypeptide or a fragment, variant, or derivative thereof is produced in vivo. The coding regions encoding an IV polypeptide or a fragment, variant, or derivative thereof may be codon optimized for mammals, e.g., humans, apes, monkeys (e.g., owl, squirrel, cebus, thomas, African green, patas, cynomolgus, and cercopithecus), orangutans, baboons, gibbons, and chimpanzees, dogs, wolves, cats, lions, and tigers, horses, donkeys, zebras, cows, pigs, sheep, deer, giraffes, bears, rabbits, mice, ferrets, seals, whales, birds, e.g., ducks, geese, terns, shawwaters, gulls, turkeys, chickens, quail, pheasants, geese, starlings and budgerigars, or other vertebrates.

[0169] In one embodiment, the present invention relates to codon-optimized coding regions encoding polypeptides of IV, or nucleic acid fragments of such coding regions fragments, variants, or derivatives thereof which have been optimized according to human codon usage. For example, human codon-optimized coding regions encoding polypeptides of IV, or fragments, variants, or derivatives thereof are prepared by substituting one or more codons preferred for use in human genes for the codons naturally used in the DNA sequence encoding the IV polypeptide or a fragment, variant, or derivative thereof. Also provided are polynucleotides, vectors, and other expression constructs comprising codon-optimized coding regions encoding polypeptides of IV, or nucleic acid fragments of such coding regions encoding fragments, variants, or derivatives thereof, pharmaceutical compositions comprising polynucleotides, vectors, and other expression constructs comprising codon-optimized coding regions encoding polypeptides of IV, or nucleic acid fragments of such coding regions encoding fragments, variants, or derivatives thereof, and various methods of using such polynucleotides, vectors and other expression constructs. Coding regions encoding IV polypeptides can be uniformly optimized, fully optimized, minimally optimized, codon-optimized by region and/or not codon-optimized, as described herein.

[0170] The present invention is further directed towards polynucleotides comprising codon-optimized coding regions encoding polypeptides of IV antigens, for example, HA, NA, NP, M1 and M2, optionally in conjunction with other antigens. The invention is also directed to polynucleotides comprising codon-optimized nucleic acid fragments encoding fragments, variants and derivatives of these polypeptides, e.g., an eM2 or a fusion of NP and eM2.

[0171] In certain embodiments, the present invention provides an isolated polynucleotide comprising a nucleic acid fragment, where the nucleic acid fragment is a fragment of a codon-optimized coding region encoding a polypeptide at least 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to an IV polypeptide, e.g., HA, NA, NP, M1 or M2, and where the nucleic acid fragment is a variant of a codon-optimized coding region encoding an IV polypeptide, e.g., HA, NA, NP, M1 or M2. The human codon-optimized coding region can be optimized for any vertebrate species and by any of the methods described herein.

Isolated IV Polypeptides

[0172] The present invention is further drawn to compositions which include at least one polynucleotide comprising one or more nucleic acid fragments, where each nucleic acid fragment is optionally a fragment of a codon-optimized coding region operably encoding an IV polypeptide or fragment, variant, or derivative thereof; together with one or more isolated IV component or isolated polypeptide. The IV component may be inactivated virus, attenuated virus, a viral
vector expressing an isolated influenza virus polypeptide, or an influenza virus protein, fragment, variant or derivative thereof.

[0173] The polypeptides or fragments, variants or derivatives thereof, in combination with the codon-optimized nucleic acid compositions may be referred to as “combinatorial polynucleotide vaccine compositions” or “single formulation heterologous prime-boost vaccine compositions.”

[0174] The isolated IV polypeptides of the invention may be in any form, and are generated using techniques well known in the art. Examples include isolated IV proteins produced recombinantly, isolated IV proteins purified from their natural milieu, recombinant (non-IV) virus vectors expressing an isolated IV protein, or proteins delivered in the form of an inactivated IV vaccine, such as conventional vaccines.

[0175] When utilized, an isolated IV polypeptide or fragment, variant or derivative thereof is administered in an immunologically effective amount. Conventional IV vaccines have been standardized to micrograms of viral antigens HA and NA. See Subbarao, K., *Advances in Viral Research* 54:349-373 (1999), incorporated herein by reference in its entirety. The recommended dose for these vaccines is 15 µg of each HA per 0.5 ml Id. The effective amount of conventional IV vaccines is determinable by one of ordinary skill in the art based upon several factors, including the antigen being expressed, the age and weight of the subject, and the precise condition requiring treatment and its severity, and route of administration.

[0176] In the instant invention, the combination of conventional antigen vaccine compositions with the codon-optimized nucleic acid compositions provides for therapeutically beneficial effects at dose sparing concentrations. For example, immunological responses sufficient for a therapeutically beneficial effect in patients predeterined for an approved commercial product, such as for the conventional product described above, can be attained by using less of the approved commercial product when supplemented or enhanced with the appropriate amount of codon-optimized nucleic acid. Thus, dose sparing is contemplated by administration of conventional IV vaccines administered in combination with the codon-optimized nucleic acids of the invention.

[0177] In particular, the dose of conventional vaccine may be reduced by at least 5%, at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60% or at least 70% when administered in combination with the codon-optimized nucleic acid compositions of the invention.

[0178] Similarly, a desirable level of an immunological response afforded by a DNA based pharmaceutical alone may be attained with less DNA by including an aliquot of a conventional vaccine. Further, using a combination of conventional and DNA based pharmaceuticals may allow both materials to be used in lesser amounts while still affording the desired level of immune response arising from administration of either component alone in higher amounts (e.g. one may use less of either immunological product when they are used in combination). This may be manifest not only by using lower amounts of materials being delivered at any time, but also to reducing the number of administrations points in a vaccination regime (e.g. 2 versus 3 or 4 injections), and/or to reducing the kinetics of the immunological response (e.g. desired response levels are attained in 3 weeks in stead of 6 after immunization).

[0179] In particular, the dose of DNA based pharmaceuticals, may be reduced by at least 5%, at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60% or at least 70% when administered in combination with conventional IV vaccines.

[0180] Determining the precise amounts of DNA based pharmaceutical and conventional antigen is based on a number of factors as described above, and is readily determined by one of ordinary skill in the art.

[0181] In addition to dose sparing, the claimed combinatorial compositions provide for a broadening of the immune response and/or enhanced beneficial immune responses. Such broadened or enhanced immune responses are achieved by: adding DNA to enhance cellular responses to a conventional vaccine; adding a conventional vaccine to a DNA pharmaceutical to enhance humoral response; using a combination that induces additional epitopes (both humoral and/or cellular) to be recognized and/or more desirably (epitope broadening); employing a DNA-conventional vaccine combination designed for a particular desired spectrum of immunological responses; obtaining a desirable spectrum by using higher amounts of either component. The broadened immune response is measurable by one of ordinary skill in the art by standard immunological assay specific for the desirable response spectrum.

[0182] Both broadening and dose sparing can be obtained simultaneously.

[0183] The isolated IV polypeptide or fragment, variant, or derivative thereof to be delivered (either a recombinant protein, a purified subunit, or viral vector expressing an isolated IV polypeptide, or in the form of an inactivated IV vaccine) can be any isolated IV polypeptide or fragment, variant, or derivative thereof, including but not limited to the HA, NA, NP, M1, or M2 proteins or fragments, variants or derivatives thereof. Fragments include, but are not limited to, the eM2 protein. In certain embodiments, a derivative protein can be a fusion protein, e.g., NP-eM2. It should be noted that any isolated IV polypeptide or fragment, variant, or derivative thereof described herein can be combined in a composition with any polynucleotide comprising a nucleic acid fragment, where the nucleic acid fragment is optionally a fragment of a codon-optimized coding region operably encoding an IV polypeptide or fragment, variant, or derivative thereof. The proteins can be different, the same, or can be combined in any combination of one or more isolated IV proteins and one or more polynucleotides.

[0184] In certain embodiments, the isolated IV polypeptides, or fragments, derivatives or variants thereof can be fused to or conjugated to a second isolated IV polypeptide, or fragment, derivative or variant thereof, or can be fused to other heterologous proteins, including for example, hepatitis B proteins including, but not limited to the hepatitis B core antigen (HBcAg), or those derived from diphtheria or tetanus. The second isolated IV polypeptide or other heterologous protein can act as a “carrier” that potentiates the immunogenicity of the IV polypeptide or a fragment, variant, or derivative thereof to which it is attached. Hepatitis B virus proteins and fragments and variants thereof useful as
The use of recombinant particles comprising hepatitis B core antigen ("HBcAg") and heterologous protein sequences as potent immunogenic moieties is well documented. For example, addition of heterologous sequences to the amino terminus of a recombinant HBcAg results in the spontaneous assembly of particulate structures which express the heterologous epitope on their surface, and which are highly immunogenic when inoculated into experimental animals. See Clarke et al., Nature 330:381-384 (1987). Heterologous epitopes can also be inserted into HBcAg particles by replacing approximately 40 amino acids of the carboxy terminus of the protein with the heterologous sequences. These recombinant HBcAg proteins also spontaneously form immunogenic particles. See Stahl and Murray, Proc. Natl. Acad. Sci. USA, 86:6283-6287 (1989). Additionally, chimeric HBcAg particles may be constructed where the heterologous epitope is inserted in or replaces all or part of the sequence of amino acid residues in one or more central regions of the HBcAg protein, in an immunodominant loop, thereby allowing the heterologous epitope to be displayed on the surface of the resulting particle. See EP Patent No. 0421835 B1. Shown below are the DNA and amino acid sequences of the human hepatitis B core protein (HBc), subtype ayw (SEQ ID Nos. 39 and 40), as described in Galibert, et al., Nature 281:646-650 (1979); see also U.S. Pat. Nos. 4,818,527, 4,882,145 and 5,143,726. All of the above references are incorporated herein by reference in their entirety. The nucleotide and amino acid sequences are presented herein as SEQ ID NO 39:

```
ATGAGATCGACACCTTATGAAAGTAATGGGATCTGAGTGCCTTCCTGAACTCACAC
GTTCGGTCACACCTGTACGGTAACTGGCGGACTGAGATGAC
CTCGGCCGAGCTGCAATGTATGAC
CTACGCGGCGGCATACGAC
CTGAGCTCGAGCTGAC
CTGAGCTCGAGCTGAC
TTCAGGCGGCTGTACGACCAC
CTCAGGCGGCTGTACGACCAC
ATCTCCTGGTTGCTGGGCTGAGATGAC
GAACTGAGATGAC
TTGAGTTGCTGGGCTGAGATGAC
GAAGCTGGGAATCTGAGATGAC
```

[0188] and SEQ ID NO:42:

```
ATGAGATCGACACCTTATGAAAGTAATGGGATCTGAGTGCCTTCCTGAACTCACAC
GTTCGGTCACACCTGTACGGTAACTGGCGGACTGAGATGAC
CTCGGCCGAGCTGCAATGTATGAC
CTACGCGGCGGCATACGAC
CTGAGCTCGAGCTGAC
CTGAGCTCGAGCTGAC
TTCAGGCGGCTGTACGACCAC
CTCAGGCGGCTGTACGACCAC
ATCTCCTGGTTGCTGGGCTGAGATGAC
GAACTGAGATGAC
TTGAGTTGCTGGGCTGAGATGAC
GAAGCTGGGAATCTGAGATGAC
```

[0189] Chimeric HBcAg particles comprising isolated IV proteins or variants, fragments or derivatives thereof are prepared by recombinant techniques well known to those of ordinary skill in the art. A polynucleotide, e.g., a plasmid, which carries the coding region for the HBcAg operably associated with a promoter is constructed. Convenient restrictions sites are engineered into the coding region encoding the N-terminal, central, and/or C-terminal portions of the HBcAg, such that heterologous sequences may be inserted. A construct which expresses a HBcAg/IV fusion protein is prepared by inserting a DNA sequence encoding an IV protein or variant, fragment or derivative thereof, in frame, into a desired restriction site in the coding region of the HBcAg. The resulting construct can be inserted into a suitable host cell, e.g., E. coli, under conditions where the chimeric HBcAg will be expressed. The chimeric HBcAg self-assembles into particles when expressed, and can then be isolated, e.g., by ultracentrifugation. The particles formed resemble the natural 27 nm HBcAg particles isolated from a hepatitis B virus, except that an isolated IV protein or variant, fragment, variant or derivative thereof is contained in the particle, preferably exposed on the outer particle surface.

[0190] The IV protein or fragment, variant, or derivative thereof expressed in a chimeric HBcAg particle may be of
any size which allows suitable particles of the chimeric HBcAg to self-assemble. As discussed above, even small antigenic epitopes may be immunogenic when expressed in the context of an immunogenic carrier, e.g., a HBcAg. Thus, HBcAg particles of the invention may comprise at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, at least 25, or between about 15 to about 30 amino acids of an IV protein fragment of interest inserted therein. HBcAg particles of the invention may further comprise immunogenic or antigenic epitopes of at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 amino acid residues of an IV protein fragment of interest inserted therein.

[0191] The immunodominant loop region of HBcAg was mapped to about amino acid residues 75 to 83, to about amino acids 75 to 85 or to about amino acids 130 to 140. See Colucci et al., J. Immunol. 141:4376-4380 (1988), and Salfeld et al., J. Virol. 63:798 (1989), which are incorporated by reference. A chimeric HBcAg is still cloned into the immunodominant loop. Thus, for example, amino acids of the IV protein fragment may be inserted into the sequence of HBcAg amino acids at various positions, for example, at the N-terminus, from about amino acid 75 to about amino acid 85, from about amino acid 75 to about amino acid 83, from about amino acid 130 to about amino acid 140, or at the C-terminus. Where amino acids of the IV protein fragment replace all or part of the native core protein sequence, the inserted IV sequence is generally not shorter, but may be longer, than the HBcAg sequence it replaces.

[0192] Alternatively, if particle formation is not desired, full-length IV coding sequences can be fused to the coding region for the HBcAg. The HBcAg sequences can be fused either at the N- or C-terminus of any of the Influenza antigens described herein, including the eM2-26-NT constructs. Fusions could include flexible protein linkers as described for NP-cM2 fusions above. Examples of IV coding sequences for the HBcAg coding sequence of SEQ ID NO:41 include an IAV NP-HBcAg fusion (SEQ ID NO:43),
[0193] an IBV NP-HBcAg fusion (SEQ ID NO:44),

[0194] or an IAV M1-HBcAg fusion (SEQ ID NO:45),
but is not limited to humans; primates such as apes, monkeys
(e.g., owl, squirrel, rhesus, African green, patas,
cynomolgus, and cercopithecus), orangutans, baboons, gib-
bons, and chimpanzees; canids such as dogs and wolves;
felids such as cats, lions, and tigers; equines such as horses,
donkeys, and zebras; food animals such as cows, pigs, and
sheep; ungulates such as deer and giraffes; birds such as
bears; and others such as rabbits, mice, ferrets, seals, whales.
In particular, the mammal can be a human subject, a food
animal or a companion animal.

[0201] The term “bird” is intended to encompass a singular
“bird” and plural “birds,” and includes, but is not limited
to feral water birds such as ducks, geese, terns, shearwaters,
and gulls; as well as domestic avian species such as turkeys,
chickens, quail, pheasants, geese, and ducks. The term
“bird” also encompasses passerine birds such as starlings
and budgerigars.

[0202] The present invention further provides a method for
generating, enhancing or modulating an immune
response to an W comprising administering to a vertebrate
one or more of the compositions described herein. In this
method, the compositions may include one or more isolated
polynucleotides comprising at least one nucleic acid frag-
ment where the nucleic acid fragment is optionally a frag-
ment of a codon-optimized coding region encoding an IV
polypeptide, or a fragment, variant, or derivative thereof.
When the protein is provided as a recombinant protein,
in particular, a fusion protein, a purified subunit, viral vector
expressing the protein, or in the form of an inactivated IV
vaccine. Thus, the latter compositions include both a poly-
ucleotide encoding an IV polypeptide or a fragment, vari-
ant, or derivative thereof and an isolated IV polypeptide
or a fragment, variant, or derivative thereof. The IV polypep-
tide or a fragment, variant, or derivative thereof encoded
by the polynucleotide of the compositions include the same
as the isolated IV polypeptide or a fragment, variant,
and derivative thereof of the compositions. Compositions to be
used according to this method may be univalent, bivalent,
trivalent or multivalent.

[0203] The polynucleotides of the compositions may com-
prise a fragment of a human (or other vertebrate) codon-
optimized coding region encoding a protein of the IV, or a
fragment, variant, or derivative thereof. The polynucleotides
are incorporated into the cells of the vertebrate in vivo, and
an antigenic amount of the IV polypeptide, or fragment,
variant, or derivative thereof, is produced in vivo. Upon
administration of the composition according to this method,
the IV polypeptide or a fragment, variant, or derivative
thereof is expressed in the vertebrate in an amount suffi-
cient to elicit an immune response. Such an immune response
might be used, for example, to generate antibodies to the IV
for use in diagnostic assays or as laboratory reagents, or as
therapeutic or preventative vaccines as described herein.

[0204] The present invention further provides a method for
generating, enhancing, or modulating a protective and/or
therapeutic immune response to IV in a vertebrate, compris-
ing administering to a vertebrate in need of therapeutic
and/or preventative immunity one or more of the composi-
tions described herein. In this method, the compositions
include one or more polynucleotides comprising at least one nucleic acid fragment, where the nucleic acid fragment is optionally a fragment of a codon-optimized coding region encoding an IV polypeptide, or a fragment, variant, or derivative thereof. In a further embodiment, the composition used in this method includes both an isolated polynucleotide comprising at least one nucleic acid fragment, where the nucleic acid fragment is optionally a fragment of a codon-optimized coding region encoding an IV polypeptide, or a fragment, variant, or derivative thereof; and at least one isolated IV polypeptide, or a fragment, variant, or derivative thereof. Thus, the latter composition includes both an isolated polynucleotide encoding an IV polypeptide or a fragment, variant, or derivative thereof and an isolated IV polypeptide or a fragment, variant, or derivative thereof, for example, a recombinant protein, a purified subunit, viral vector expressing the protein, or an inactivated virus vaccine. Upon administration of the composition according to this method, the IV polypeptide or a fragment, variant, or derivative thereof is expressed in the human in a therapeutically or prophylactically effective amount.

[0205] As used herein, an "immune response" refers to the ability of a vertebrate to elicit an immune reaction to a composition delivered to that vertebrate. Examples of immune responses include an antibody response or a cellular, e.g., cytotoxic T-cell, response. One or more compositions of the present invention may be used to prevent influenza infection in vertebrates, e.g., as a prophylactic vaccine, to establish or enhance immunity to IV in a healthy individual prior to exposure to influenza or contraction of influenza disease, thus preventing the disease or reducing the severity of disease symptoms.

[0206] As mentioned above, compositions of the present invention can be used both to prevent IV infection, and also to therapeutically treat IV infection. In individuals already exposed to influenza, or already suffering from influenza disease, the present invention is used to further stimulate the immune system of the vertebrate, thus reducing or eliminating the symptoms associated with that disease or disorder. As defined herein, "treatment" refers to the use of one or more compositions of the present invention to prevent, cure, retard, or reduce the severity of influenza disease in a vertebrate, and/or result in no worsening of influenza disease over a specified period of time in a vertebrate which has already been exposed to IV and is thus in need of therapy. The term "prevention" refers to the use of one or more compositions of the present invention to generate immunity in a vertebrate which has not yet been exposed to a particular strain of IV, thereby preventing or reducing disease symptoms if the vertebrate is later exposed to the particular strain of IV. The methods of the present invention therefore may be referred to as therapeutic vaccination or preventative or prophylactic vaccination. It is not required that any composition of the present invention provide total immunity to influenza or totally cure or eliminate all influenza disease symptoms. As used herein, a "vertebrate in need of therapeutic and/or preventative immunity" refers to an individual for whom it is desirable to treat, i.e., to prevent, cure, retard, or reduce the severity of influenza disease symptoms, and/or result in no worsening of influenza disease over a specified period of time. Vertebrates to treat and/or vaccinate include humans, apes, monkeys (e.g., owl, squirrel, cebus, rhesus, African green, patas, cynomolgus, and cercopithecus), orangutans, baboons, gibbons, and chimpanzees, dogs, wolves, cats, lions, and tigers, horses, donkeys, zebras, cows, pigs, sheep, deer, giraffes, bears, rabbits, mice, ferrets, seals, whales, ducks, geese, terns, shawwaters, gulls, turkeys, chickens, quail, pheasants, geese, starlings and budgerigars.

[0207] One or more compositions of the present invention are utilized in a "prime boost" regimen. An example of a "prime boost" regimen may be found in Yang, Z. et al. J. Virol. 77:799-803 (2002), which is incorporated herein by reference in its entirety. In these embodiments, one or more polynucleotide vaccine compositions of the present invention are delivered to a vertebrate, thereby priming the immune response of the vertebrate to an IV, and then a second immunogenic composition is utilized as a boost vaccination. One or more compositions of the present invention are used to prime immunity, and then a second immunogenic composition, e.g., a recombinant viral vaccine or vaccines, a different polynucleotide vaccine, or one or more purified subunit isolated IV polypeptides or fragments, variants or derivatives thereof is used to boost the anti-IV immune response.

[0208] In one embodiment, a priming composition and a boosting composition are combined in a single composition or single formulation. For example, a single composition may comprise an isolated IV polypeptide or a fragment, variant, or derivative thereof as the priming component and a polynucleotide encoding an influenza protein as the boosting component. In this embodiment, the compositions may be contained in a single vial where the priming component and boosting component are mixed together. In general, because the peak levels of expression of protein from the polynucleotide does not occur until later (e.g. 7-10 days) after administration, the polynucleotide component may provide a boost to the isolated protein component. Compositions comprising both a priming component and a boosting component are referred to herein as "combinatorial vaccine compositions" or "single formulation heterologous prime-boost vaccine compositions." In addition, the priming composition may be administered before the boosting composition, or even after the boosting composition, if the boosting composition is expected to take longer to act.

[0209] In another embodiment, the priming composition may be administered simultaneously with the boosting composition, but in separate formulations where the priming component and the boosting component are separated.

[0210] The terms "priming" or "primary" and "boost" or "boosting" as used herein may refer to the initial and subsequent immunizations, respectively, i.e., in accordance with the definitions these terms normally have in immunology. However, in certain embodiments, e.g., where the priming component and boosting component are in a single formulation, initial and subsequent immunizations may not be necessary as both the "priming" and the "boost" compositions are administered simultaneously.

[0211] In certain embodiments, one or more compositions of the present invention are delivered to a vertebrate by methods described herein, thereby achieving an effective therapeutic and/or an effective preventative immune response. More specifically, the compositions of the present invention may be administered to any tissue of a vertebrate, including, but not limited to, muscle, skin, brain tissue, lung tissue, liver tissue, spleen tissue, bone marrow tissue, thy-
mus tissue, heart tissue, e.g., myocardium, endocardium, and pericardium, lymph tissue, blood tissue, bone tissue, pancreas tissue, kidney tissue, gall bladder tissue, stomach tissue, intestinal tissue, testicular tissue, ovarian tissue, uterine tissue, vaginal tissue, rectal tissue, nervous system tissue, eye tissue, glandular tissue, tongue tissue, and connective tissue, e.g., cartilage.

[0212] Furthermore, the compositions of the present invention may be administered to any internal cavity of a vertebrate, including, but not limited to, the lungs, the mouth, the nasal cavity, the stomach, the peritoneal cavity, the intestine, any heart chamber, veins, arteries, capillaries, lymphatic cavities, the uterine cavity, the vaginal cavity, the rectal cavity, joint cavities, ventricles in brain, spinal canal in spinal cord, the ocular cavities, the lumen of a duct of a salivary gland or a liver. When the compositions of the present invention is administered to the lumen of a duct of a salivary gland or liver, the desired polypeptide is expressed in the salivary gland and the liver such that the polypeptide is delivered into the blood stream of the vertebrate from each of the salivary gland or the liver. Certain modes for administration to secretory organs of a gastrointestinal system using the salivary gland, liver and pancreas to release a desired polypeptide into the bloodstream is disclosed in U.S. Pat. Nos. 5,837,693 and 6,004,944, both of which are incorporated herein by reference in their entirety.

[0213] In certain embodiments, the compositions are administered into embryonated chicken eggs or by intramuscular injection into the defleshed breast area of chicks described in Kodihalli S. et al., Vaccine 18:2592-9 (2000), which is incorporated herein by reference in its entirety.

[0214] In certain embodiments, the compositions are administered to muscle, either skeletal muscle or cardiac muscle, or to lung tissue. Specific, but non-limiting modes for administration to lung tissue are disclosed in Wheeler, C. J., et al., Proc. Natl. Acad. Sci. USA 93:11454-11459 (1996), which is incorporated herein by reference in its entirety.

[0215] According to the disclosed methods, compositions of the present invention can be administered by intramuscular (i.m.), subcutaneous (s.c.), or intrapulmonary routes. Other suitable routes of administration include, but are not limited to intratracheal, transdermal, intracutaneous, intranasal, inhalation, intravenous, intravenous (i.v.), intraocular (e.g., into the pancreas) and intraparenchymal (i.e., into any tissue) administration. Transdermal delivery includes, but not limited to intradermal (e.g., into the dermis or epidermis), transdermal (e.g., percutaneous) and transmucosal administration (i.e., into or through skin or mucosal tissue). Intracavity administration includes, but not limited to administration into oral, vaginal, rectal, nasal, peritoneal, or intestinal cavities as well as, intrabuccal (i.e., into spinal canal), intraventricular (i.e., into the brain ventricles or the heart ventricles), intrathecal (i.e., into the heart atrium) and subarachnoid (i.e., into the subarachnoid spaces of the brain) administration.


[0217] Determining an effective amount of one or more compositions of the present invention depends upon a number of factors including, for example, the antigen being expressed or administered directly, e.g., HA, NA, NP, M1 or M2, or fragments, e.g., eM2, variants, or derivatives thereof, the age and weight of the subject, the precise condition requiring treatment and its severity, and the route of administration. Based on the above factors, determining the precise amount, number of doses, and timing of doses are within the ordinary skill in the art and will be readily determined by the attending physician or veterinarian.

[0218] Compositions of the present invention may include various salts, excipients, delivery vehicles and/or auxiliary agents as are disclosed, e.g., in U.S. patent application Publication No. 2002/0019358, published Feb. 14, 2002, which is incorporated herein by reference in its entirety.

[0219] Furthermore, compositions of the present invention may include one or more transfection facilitating compounds that facilitate delivery of polynucleotides to the interior of a cell, and/or to a desired location within a cell. As used herein, the terms “transfection facilitating compound,” “transfection facilitating agent”, and “transfection facilitating material” are synonymous, and may be used interchangeably. It should be noted that certain transfection facilitating compounds may also be “adjuvants” as described infra, i.e., in addition to facilitating delivery of polynucleotides to the interior of a cell, the compound acts to alter or increase the immune response to the antigen encoded by that polynucleotide. Examples of the transfection facilitating compounds include, but are not limited to inorganic materials such as calcium phosphate, alum (aluminum sulfate), and gold particles (e.g., “powder” type delivery vehicles); peptides that are, for example, cationic, intercell targeting (for selective delivery to certain cell types), intracellular targeting (for nuclear localization or endosomal escape), and
amphipathic (helix forming or pore forming); proteins that are, for example, basic (e.g., positively charged) such as histones, targeting (e.g., asialoprotein), viral (e.g., Sendai virus coat protein), and pore-forming; lipids that are, for example, cationic (e.g., DMRIE, DOSPA, DC-Chol), basic (e.g., steryl amine), neutral (e.g., cholesterol), anionic (e.g., phosphatidyl serine), and zwitterionic (e.g., DOPE, DOPC); and polymers such as dextran, star-polymers, “homogeneous” poly-amino acids (e.g., poly-lysine, poly-arginine), “heterogeneous” poly-amino acids (e.g., mixtures of lysine & glycine), co-polymers, polyvinylpyrrolidone (PVP), poloxamers (e.g. CRL 1005) and polyethylene glycol (PEG).

A transfection facilitating material can be used alone or in combination with one or more other transfection facilitating materials. Two or more transfection facilitating materials can be combined by chemical bonding (e.g., covalent and ionic such as in lipilized polylysine, PE-glylated polylysine) (Toncheva, et al., Biochim. Biophys. Acta 1380(3):354-368 (1988)), mechanical mixing (e.g., free moving materials in liquid or solid phase such as “polylysine+cationic lipids” (Gao and Huang, Biochemistry 35:1027-1036 (1996); Trubetskoy, et al., Biochim. Biophys. Acta 1131:311-313 (1992)), and aggregation (e.g., co-precipitation, gel forming such as in cationic lipids+poly-lactide, and polylysine+gelatin). Each of the references cited in this paragraph is incorporated herein by reference in its entirety.

[0220] One category of transfection facilitating materials is cationic lipids. Examples of cationic lipids are 5-carboxyyspermyl glycerol diocatclylamide (DOGS) and dipalmityl-phosphatidylethanolamine-5-carboxyspermylamide (DPPES). Cationic cholesterol derivatives are also useful, including [3]-[N,N,N'-dimethylamino]ethane-1-carboxyl]-cholesterol (DC-Chol). Dimethyl dioctadecylammonium bromide (DDAB), N-[3-(amino propyl)-N-[3-(bis(2- tetradecyloxyethyl))-N-methyl-ammonium bromide (PA DEMO). N-[3-(amino propyl)-N-[3-(bis(2- dodecylcxyloxyethy)l]-N-methyl-ammonium bromide (PA-DELO). N,N,N-tri-(2-dodecylcxyloxyethyl)-N-(3 amino)propyl-ammonium bromide (PA-TELO), and N1-[3 amino propyl]-(2-dodecylcxyloxyethyl)N2-(2 dodecylcxyloxyethyl)ethyl-N1-piperazinammonium bromide (GA-LOEBP) can also be used in the present invention.

[0221] Non-dicher cationic lipids, such as DL-1,2-dioleoyl-3-dimethylaminopropyl-p-hydroxyethyl ammonium (DORI diester), 1-O-octyl-2-oleoyl-3-dimethylaminopropyl-p-hydroxyethyl ammonium (DORI ester/ether), and their salts promote in vivo gene delivery. In some embodiments, cationic lipids comprise groups attached via a heteroatom attached to the quaternary ammonium moiety in the head group. A glyceryl spacer can connect the linker to the hydroxyl group.

[0222] Specific, but non-limiting cationic lipids for use in certain embodiments of the present invention include DMRIE ((2-N-[2-hydroxyethyl]-N,N-dimethyl-2,3-bis(tetradecyl oxy)l-1-propanium bromide), GAP-DMORIE ((2-N-[3-amino propyl]-N,N-dimethyl-2,3-bis(syn-9-tetradecyn-1-oxo)l-1-propanium bromide), and GAP-DM RIE ((2-N-[3-amino propyl]-N,N-dimethyl-2,3-bis(dodecyl oxy)l-1-propanium bromide).

[0223] Other specific but non-limiting cationic surfactants for use in certain embodiments of the present invention include Bn-DHRIE, DhzRIE, DhzRIE-OAc, DhzRIE-OBz and Pr-DOctRIE-OAc. These lipids are disclosed in copending U.S. patent application Ser. No. 10/725,015. In another aspect of the present invention, the cationic surfactant is Pr-DOctRIE-OAc.

[0224] Other cationic lipids include [(2-N-[2-(sperminecarboxamido)ethyl]-2,3-bis(dodecyl oxy)l-1 propaninium pentahydrochloride (DOSPA), (2-N-[2-amino ethyl]-N,N-dimethyl-2,3-bis(tetradecyl oxy)l-1 propaninium bromide (β-aminoethyl-DMRIE or βAE DMRIE) (Wheeler, et al., Biochim. Biophys. Acta 1280:1-11 (1996), and (2-N-[3-amino propyl]-N,N-dimethyl-2,3-bis(dodecyl oxy)l-1 propaninium bromide (GAP-DLRIE) (Wheeler, et al., Proc. Natl. Acad. Sci. USA 93:11454-11459 (1996)), which have been developed from DMRIE. Both of the references cited in this paragraph are incorporated herein by reference in their entirety.

[0225] Other examples of DMRIE-derived cationic lipids that are useful for the present invention are (2-N-[3-amino propyl]-N,N-dimethyl-2,3-bis(decetyl oxy)l-1-propanumin bromide (GAP-DRRIE), (2-N-[3-amino propyl]-N,N-dimethyl-2,3-bis(tetradecyl oxy)l-1-propaninium bromide (GAP-DMRIE), (2-N-[2-hydroxethyl]-N-N-[2-hydroxyethyl]-N,N-dimethyl-2,3-bis(2-hexadecyl oxy)l-1-propaninium bromide (GMU-DRIE), and (2-N-[2-hydroxyethyl]-N,N-dimethyl-2,3-bis(2-hexadecyl oxy)l-1-propaninium bromide (HP-DORIE).

[0226] In the embodiments where the immunogenic composition comprises a cationic lipid, the cationic lipid may be mixed with one or more co-lipids. For purposes of definition, the term “co-lipid” refers to any hydrophobic material which may be combined with the cationic lipid component and includes amphiphatic lipids, such as phospholipids, and neutral lipids, such as cholesterol. Cationic lipids and co-lipids may be mixed or combined in a number of ways to produce a variety of non-covalently bonded macroscopic structures, including, for example, liposomes, multilamellar vesicles, unilamellar vesicles, micelles, and simple films. One non-limiting class of co-lipids are the zwitterionic phospholipids, which include the phosphatidylethanolamines and the phosphatidylcholines. Examples of phosphati dylethanolamines, include DOPE, DMPE and DPyPE. In certain embodiments, the co-lipid is DPyPE, which comprises two phytanoyl substituents incorporated into the diallylphosphatidyl ethanolamine skeleton. In other embodiments, the co-lipid is DOPE, CAS name 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine.

[0227] When a composition of the present invention comprises a cationic lipid and co-lipid, the cationic lipid:co-lipid molar ratio may be from about 9:1 to about 1:9, from about 4:1 to about 1:4, from about 2:1 to about 1:2, or about 1:1.

[0228] In order to maximize homogeneity, the cationic lipid and co-lipid components may be dissolved in a solvent such as chloroform, followed by evaporation of the cationic lipid:co-lipid solution under vacuum to dryness as a film on the inner surface of a glass vessel (e.g., a Rotovap round-bottomed flask). Upon suspension in an aqueous solvent, the amphipathic lipid component molecules self-assemble into homogenous lipid vesicles. These lipid vesicles may subsequently be processed to have a selected mean diameter of uniform size prior to complexing with, for example, a
codon-optimized polynucleotide of the present invention, according to methods known to those skilled in the art. For example, the sonication of a lipid solution is described in Feltner et al., **Proc. Natl. Acad. Sci. USA** 8, 7413-7417 (1987) and in U.S. Pat. No. 5,264,618, the disclosures of which are incorporated herein by reference.

[0229] In those embodiments where the composition includes a cationic lipid, polynucleotides of the present invention are complexed with lipids by mixing, for example, a plasmid in aqueous solution and a solution of cationic lipid/ion complex as prepared herein are mixed. The concentration of each of the constituent solutions can be adjusted prior to mixing such that the desired final plasmid/cationic lipid:co-lipid ratio and the desired plasmid final concentration will be obtained upon mixing the two solutions. The cationic lipid/co-lipid mixtures are suitably prepared by hydrating a thin film of the mixed lipid materials in an appropriate solvent at elevated temperature to allow for mixing at ambient temperatures for about 1 minute. The thin films are prepared by admixing chloroform solutions of the individual components to afford a desired molar ratio followed by aliquoting the desired volume of the solutions into a suitable container. The solvent is removed by evaporation, first with a stream of dry, inert gas (e.g. argon) followed by high vacuum treatment.

[0230] Other hydrophobic and amphiphilic additives, such as, for example, steroids, fatty acids, gangliosides, glycolipids, lipopeptides, liposaccharides, neobioses, niosomes, prostaglandins and sphingolipids, may also be included in compositions of the present invention. In such compositions, these additives may be included in an amount between about 0.1 mol % and about 99.9 mol % (relative to total lipid), about 1-50 mol %, or about 2-25 mol %.

[0231] Additional embodiments of the present invention are drawn to compositions comprising an auxiliary agent which is administered before, after, or concurrently with the polynucleotide. As used herein, an “auxiliary agent” is a substance included in a composition for its ability to enhance the vaccine efficacy of a combination which is identical except for the inclusion of the auxiliary agent, the entry of polynucleotides into vertebrate cells in vivo, and/or the in vivo expression of polypeptides encoded by such polynucleotides. Certain auxiliary agents may, in addition to enhancing entry of polynucleotides into cells, enhance an immune response to an immunogen encoded by the polynucleotide. Auxiliary agents of the present invention include nonionic, anionic, cationic, or zwitterionic surfactants or detergents, with nonionic surfactants or detergents being preferred, chelators, DNase inhibitors, poloxamers, agents that aggregate or condense nucleic acids, emulsifying or solubilizing agents, wetting agents, gel-forming agents, and buffers.

[0232] Auxiliary agents for use in compositions of the present invention include, but are not limited to non-ionic detergents and surfactants IGEPLA® 6300, NONIDENT NP-40, Nonidet® P-40, Tween-20™, Tween-80™, Pluronic® F-68® (ave. MW: 8400; approx. MW of hydrophobe, 1800; approx. wt. % of hydrophile, 80%), Pluronic F77® (ave. MW: 6600; approx. MW of hydrophobe, 2100; approx. wt. % of hydrophile, 70%), Pluronic F67® (ave. MW: 3400; approx. MW of hydrophobe, 1800; approx. wt. % of hydrophile, 50%), Triton X-100™, and Triton X-114™; the anionic detergent sodium dodecyl sulfate (SDS); the sugar

stachyose; the condensing agent DMSO; and the chelator/ DNAse inhibitor EDTA, CRL 1005 (12 kDa, 5% POE), and BAK (Benzalkonium chloride 50% solution, available from Reger Chemical Co. Inc.). In certain specific embodiments, the auxiliary agent is DMSO, Nonidet P40, Pluronic F68® (ave. MW: 8400; approx. MW of hydrophobe, 1800; approx. wt. % of hydrophile, 80%), Pluronic F77® (ave. MW: 6600; approx. MW of hydrophobe, 2100; approx. wt. % of hydrophile, 70%), Pluronic F67® (ave. MW: 3400; approx. MW of hydrophobe, 1800; approx. wt. % of hydrophile, 50%), Pluronic L64® (ave. MW: 2900; approx. MW of hydrophobe, 1800; approx. wt. % of hydrophile, 40%), and Pluronic F108® (ave. MW: 14600; approx. MW of hydrophobe, 3000; approx. wt. % of hydrophile, 80%). See, e.g., U.S. patent application Publication No. 2002/0019358, published Feb. 14, 2002, which is incorporated herein by reference in its entirety.

[0233] Certain compositions of the present invention can further include one or more adjuvants before, after, or concurrently with the polynucleotide. The term “adjuvant” refers to any material having the ability to (1) alter or increase the immune response to a particular antigen or (2) increase or aid an effect of a pharmacological agent. It should be noted, with respect to polynucleotide vaccines, that an “adjuvant,” can be a transfection facilitating material. Similarly, certain “transfection facilitating materials” described supra, may also be an “adjuvant.” An adjuvant may be used with a composition comprising a polynucleotide of the present invention. In a prime-boost regimen, as described herein, an adjuvant may be used with either the priming immunization, the booster immunization, or both. Suitable adjuvants include, but are not limited to, cytokines and growth factors; bacterial components (e.g., endotoxins, in particular superantigens, exotoxins and cell wall components); aluminum-based salts; calcium-based salts; silica; polynucleotides; toxoids; serum proteins, viruses and virally-derived materials, poisons, venoms, imidazoqui- niline compounds, poloxamers, and cationic lipids.

[0234] A great variety of materials have been shown to have adjuvant activity through a variety of mechanisms. Any compound which may increase the expression, antigenicity or immunogenicity of the polypeptide is a potential adjuvant. The present invention provides an assay to screen for improved immune responses to potential adjuvants. Potential adjuvants which may be screened for their ability to enhance the immune response according to the present invention include, but are not limited to: inert carriers, such as alum, bentonite, latex, and acrylic particles; pluronic block polymers, such as TiterMax® (a block copolymer CRL-8941, squalane (a metabolizable oil) and a microparticulate silica stabilizer); depot formers, such as Freunds adjuvant, surface active materials, such as saponin, lyssolecithin, retinal, Quil A, liposomes, and pluronic polymer formulations; macrophage stimulators, such as bacterial lipopolysaccharide; alternate pathway complement activators, such as insulin, zymosan, endotoxin, and levamisole; and non-ionic surfactants, such as poloxamers, poly(oxyethylene)-poly-(oxypropylene) tri-block copolymers. Also included as adjuvants are transfection-facilitating materials, such as those described above.

[0235] Poloxamers which may be screened for their ability to enhance the immune response according to the present invention include, but are not limited to, commercially
available poloxamers such as Pluronic® surfactants, which are block copolymers of propylene oxide and ethylene oxide in which the propylene oxide block is sandwiched between two ethylene oxide blocks. Examples of Pluronic® surfactants include Pluronic® L121 (ave. MW: 4400; approx. MW of hydrophobe, 3600; approx. wt. % of hydrophobe, 10%), Pluronic® L101 (ave. MW: 3800; approx. MW of hydrophobe, 3000; approx. wt. % of hydrophobe, 10%), Pluronic® L81 (ave. MW: 2750; approx. MW of hydrophobe, 2400; approx. wt. % of hydrophobe, 10%), Pluronic® L61 (ave. MW: 2000; approx. MW of hydrophobe, 1800; approx. wt. % of hydrophobe, 10%), Pluronic® L31 (ave. MW: 1100; approx. MW of hydrophobe, 900; approx. wt. % of hydrophobe, 10%), Pluronic® L122 (ave. MW: 5000; approx. MW of hydrophobe, 3600; approx. wt. % of hydrophobe, 20%), Pluronic® L92 (ave. MW: 3650; approx. MW of hydrophobe, 2700; approx. wt. % of hydrophobe, 20%), Pluronic® L72 (ave. MW: 2750; approx. MW of hydrophobe, 2100; approx. wt. % of hydrophobe, 20%), Pluronic® L62 (ave. MW: 2500; approx. MW of hydrophobe, 1800; approx. wt. % of hydrophobe, 20%), Pluronic® L42 (ave. MW: 1630; approx. MW of hydrophobe, 1200; approx. wt. % of hydrophobe, 20%), Pluronic® L3 (ave. MW: 2650; approx. MW of hydrophobe, 1800; approx. wt. % of hydrophobe, 30%), Pluronic® L4 (ave. MW: 2200; approx. MW of hydrophobe, 1200; approx. wt. % of hydrophobe, 40%), Pluronic® L3S (ave. MW: 2000; approx. MW of hydrophobe, 900; approx. wt. % of hydrophobe, 50%), Pluronic® P123 (ave. MW: 5750; approx. MW of hydrophobe, 3600; approx. wt. % of hydrophobe, 30%), Pluronic® P103 (ave. MW: 4950; approx. MW of hydrophobe, 3000; approx. wt. % of hydrophobe, 30%), Pluronic® L104 (ave. MW: 5900; approx. MW of hydrophobe, 3000; approx. wt. % of hydrophobe, 40%), Pluronic® P84 (ave. MW: 4200; approx. MW of hydrophobe, 2400; approx. wt. % of hydrophobe, 40%), Pluronic® P105 (ave. MW: 6500; approx. MW of hydrophobe, 3000; approx. wt. % of hydrophobe, 50%), Pluronic® P85 (ave. MW: 4600; approx. MW of hydrophobe, 2400; approx. wt. % of hydrophobe, 50%), Pluronic® P75 (ave. MW: 4150; approx. MW of hydrophobe, 2100; approx. wt. % of hydrophobe, 50%), Pluronic® P65 (ave. MW: 3400; approx. MW of hydrophobe, 1800; approx. wt. % of hydrophobe, 50%), Pluronic® P127 (ave. MW: 12600; approx. MW of hydrophobe, 3600; approx. wt. % of hydrophobe, 70%), Pluronic® P98 (ave. MW: 13000; approx. MW of hydrophobe, 2700; approx. wt. % of hydrophobe, 80%), Pluronic® 187 (ave. MW: 7700; approx. MW of hydrophobe, 2400; approx. wt. % of hydrophobe, 70%), Pluronic® P108 (ave. MW: 14600; approx. MW of hydrophobe, 3000; approx. wt. % of hydrophobe, 80%), Pluronic® F88 (ave. MW: 13000; approx. MW of hydrophobe, 2700; approx. wt. % of hydrophobe, 80%), Pluronic® F89 (ave. MW: 11400; approx. MW of hydrophobe, 2400; approx. wt. % of hydrophobe, 80%), Pluronic® F68 (ave. MW: 8400; approx. MW of hydrophobe, 1800; approx. wt. % of hydrophobe, 80%), Pluronic® F38 (ave. MW: 47000; approx. MW of hydrophobe, 900; approx. wt. % of hydrophobe, 80%)

[0237] Other commercially available poloxamers which may be screened for their ability to enhance the immune response according to the present invention include compounds that are block copolymer of polyethylene and polypropylene glycol such as Synercon® L 1214 (ave. MW: 4400), Synercon® L 122 (ave. MW: 5000), Synercon® P 104 (ave. MW: 5850), Synercon® P 105 (ave. MW: 6500), Synercon® P 123 (ave. MW: 5750), Synercon® P 85 (ave. MW: 4600) and Synercon® P 94 (ave. MW: 4600), in which L indicates that the surfactants are liquids, P that they are pastes, the first digit is a measure of the molecular weight of the polypropylene portion of the surfactant and the last digit of the number, multiplied by 10, gives the percent ethylene oxide content of the surfactant; and compounds that are nonylphenyl polyethylene glycol such as Synercon® NP10 (nonylphenol ethoxylated surfactant—10% solution), Synercon® NP30 (condensate of 1 mole of nonylphenol with 30 moles of ethylene oxide) and Synercon® NPS (condensate of 1 mole of nonylphenol with 5.5 moles of naphthalene oxide).

[0238] Other poloxamers which may be screened for their ability to enhance the immune response according to the present invention include: (a) a polyester block copolymer comprising an A-type segment and a B-type segment, wherein the A-type segment comprises a linear polymeric segment of relatively hydrophilic character, the repeating units of which contribute an average Hansch-Leo fragmental constant of about -0.4 or less and have molecular weight contributions between about 30 and about 500, wherein the B-type segment comprises a linear polymeric segment of relatively hydrophobic character, the repeating units of which contribute an average Hansch-Leo fragmental constant of about -0.4 or more and have molecular weight...
contributions between about 30 and about 500, wherein at least about 80% of the linkages joining the repeating units for each of the polymeric segments comprise an ether linkage; (b) a block copolymer having a polyester segment and a polycation segment, wherein the polycation segment comprises at least an A-type block, and the polycation segment comprises a plurality of cationic repeating units; and (c) a polyester-polycation copolymer comprising a polyester, a polycation segment and a polycationic segment comprising a plurality of cationic repeating units of formulas N-H-R, wherein R is a straight chain aliphatic group of 2 to 6 carbon atoms, which may be substituted, wherein said polycarbonate segments comprise at least one of an A-type of B-type segment. See U.S. Pat. No. 5,656,611, by Kabonov, et al., which is incorporated herein by reference in its entirety. Ortho poloxamers of interest include CRL1005 (12 kDa, 5% POE), CRL3800 (11 kDa, 5% POE), CRL2690 (12 kDa, 10% POE), CRL4505 (15 kDa, 5% POE) and CRL1415 (9 kDa, 10% POE).

[0239] Other auxiliary agents which may be screened for their ability to enhance the immune response according to the present invention include, but are not limited to Accacia (gum arabic); the poloxamers ether R-O—(C2H4O)n—H (BR®), e.g., polyethylene glycol dodecyl ether (BR® 35, x=23), polyethylene glycol dodecyl ether (BR® 30, x=4), polyethylene glycol hexa decyl ether (BR® 52 x=2), polyethylene glycol hexadecyl ether (BR® 56, x=10), polyethylene glycol hexadecyl ether (BR® 58, x=20), polyethylene glycol octadecyl ether (BR® 72, x=2), polyethylene glycol octadecyl ether (BR® 76, x=10), polyethylene glycol octadecyl ether (BR® 78, x=20), polyethylene glycol dodecyl ether (BR® 92V, x=2), and polyoxyl 10 oleyl ether (BR® 57, x=10); poly-D-glucosamine (chitosan); chlorbutanol; cholesterol; diethanolamine; diglutaric; dimethylsulfoxide (DMSO), ethylenediamine tetraacetic acid (EDTA); glycerol monostearate; lanolin alcohols; mono- and di-glycerides; monothanolamine; nonylphenol polyethylene oxide ether (NP-40®); octylphenoxypolyethoxythanol (NONIDET NP-40 from Amresco); ethyl phenol poly (ethylene glycol ether), n=11 (Nonidet P-40 from Roche); octyl phenol ethylene oxide condensate with about 9 ethylene oxide units (nonidet P40); IGEPLA. CA 630®, [(cetyl phenoxyl) polyethoxyethanol; structurally same as NONIDET NP-40]; oleic acid; oleyl alcohol; polyethylene glycol 8000; polyoxyl 20 cetosteryl ether; polyoxyl 35 castor oil; polyoxyl 40 hydrogenated castor oil; polyoxyl 40 stearate; polyoxyethylene sorbitan monolaurate (polysorbate 20, or TWEEN®-20); polyoxyethylene sorbitan monolaurate (polysorbate 80, or TWEEN®-80); propylene glycol dicarate; propylene glycol monostearate; protamine sulfate; proteolytic enzymes; sodium dodecyl sulfate (SDS); sodium monolaurate; sodium stearate; sorbitan derivatives (SPAN®), e.g., sorbitan monopalmitate (SPAN® 40), sorbitan monostearate (SPAN® 60), sorbitan tristearate (SPAN® 85), sorbitan monoleate (SPAN® 80), and sorbitan tristearate (SPAN® 85); 2,6,10,15, 19,23-hexamethyl-2,6,10,14,18,22-tetracosahexane (squalene); stachyose; stearic acid; sucrose; surfactin (lipopeptidic antibiotic from Bacillus subtilis); dodecylpoly (ethylene glycol ether), (Thesit®) MW 582,9; octyl phenol ethylene oxide condensate with about 9-10 ethylene oxide units (Triton X-100®); octyl phenol ethylene oxide condensate with about 7-8 ethylene oxide units (Triton X-114®); triis(2-hydroxyethyl)amine (trolamine); and emulsifying wax.

[0240] In certain adjuvant compositions, the adjuvant is a cytokine. A composition of the present invention can comprise one or more cytokines, chemokines, or compounds that induce the production of cytokines and chemokines, or a polynucleotide encoding one or more cytokines, chemokines, or compounds that induce the production of cytokines and chemokines. Examples include, but are not limited to granulocyte macrophage colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF), macrophage colony stimulating factor (M-CSF), colony stimulating factor (CSF), erythropoietin (EPO), interleukin 2 (IL-2), interleukin-3 (IL-3), interleukin 4 (IL-4), interleukin 5 (IL-5), interleukin 6 (IL-6), interleukin 7 (IL-7), interleukin 8 (IL-8), interleukin 10 (IL-10), interleukin 12 (IL-12), interleukin 15 (IL-15), interleukin 18 (IL-18), interferon alpha (IFNα), interferon beta (IFNβ), interferon gamma (IFNγ), interferon omega (IFNo), interferon tau (IFNτ), interferon gamma inducing factor I (IGIF), transforming growth factor beta (TGFβ), RANTES (regulated upon activation, normal T-cell expressed and presumably secreted), macrophage inflammatory proteins (e.g., MIP-1 alpha and MIP-1 beta), Leishmania elongation initiating factor (LEIF), and Flt-3 ligand.

[0241] In certain compositions of the present invention, the polynucleotide construct may be complexed with an adjuvant composition comprising (α)-N-(3-aminopropyl)-N,N-dimethyl-2,3-bis(syn-9-tetradeceneoxy)-1-propanamine bromide (GAP-DMORIE). The composition may also comprise one or more co-lipids, e.g., 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), 1,2-diphtanoylsn-glycero-3-phosphoethanolamine (DPPPE), and/or 1,2-dimyristoyl-glycero-3-phosphoethanolamine (DMPE). An adjuvant composition comprising GAP-DMORIE and DPPPE at a 1:1 molar ratio is referred to herein as Vaxfectin™. See, e.g., PCT Publication No. WO 00/57917, which is incorporated herein by reference in its entirety.

[0242] In other embodiments, the polynucleotide itself may function as an adjuvant as is the case when the polynucleotides of the invention are derived, in whole or in part, from bacterial DNA. Bacterial DNA containing motifs of unmethylated CpG-dinucleotides (CpG-DNA) triggers innate immune cells in vertebrates through a pattern recognition receptor (including toll receptors such as TLR 9) and thus possesses potent immunostimulatory effects on macrophages, dendritic cells and B-lymphocytes. See, e.g., Wagner, H., Curr. Opin. Microbiol. 5:62-69 (2002); Jung, J. et al., J. Immunol. 169: 2368-73 (2002); see also Klimman, D. M. et al., Proc. NatlAcad. Sci. U.S.A. 93:2879-83 (1996). Methods of using unmethylated CpG-dinucleotides as adjuvants are described in, for example, U.S. Pat. Nos. 6,207, 646, 6,406,705 and 6,429,199, the disclosures of which are herein incorporated by reference.

[0243] The ability of an adjuvant to increase the immune response to an antigen is typically manifested by a significant increase in immune-mediated protection. For example, an increase in humoral immunity is typically manifested by a significant increase in the titer of antibodies raised to the antigen, and an increase in T-cell activity is typically manifested in increased cell proliferation, or cellular cytotoxicity,
or cytokine secretion. An adjuvant may also alter an immune response, for example, by changing a primarily humoral or Th2 response into a primarily cellular, or Th1 response.

[0244] Nucleic acid molecules and/or polynucleotides of the present invention, e.g., plasmid DNA, mRNA, linear DNA or oligonucleotides, may be solubilized in any of various buffers. Suitable buffers include, for example, phosphate buffered saline (PBS), normal saline, Tris buffer, and sodium phosphate (e.g., 150 mM sodium phosphate). Insoluble polynucleotides may be solubilized in a weak acid or weak base, and then diluted to the desired volume with a buffer. The pH of the buffer may be adjusted as appropriate.

In addition, a pharmaceutically acceptable additive can be used to provide an appropriate osmolality. Such additives are within the purview of one skilled in the art. For aqueous compositions used in vivo, sterile pyrogen-free water can be used. Such formulations will contain an effective amount of a polynucleotide together with a suitable amount of an aqueous solution in order to prepare pharmaceutically acceptable compositions suitable for administration to a human.

[0245] Compositions of the present invention can be formulated according to known methods. Suitable preparation methods are described, for example, in Remington’s Pharmaceutical Sciences, 16th Edition, A. Osol, ed., Mack Publishing Co., Easton, Pa. (1980), and Remington’s Pharmaceutical Sciences, 19th Edition, A. R. Gennaro, ed., Mack Publishing Co., Easton, Pa. (1995), both of which are incorporated herein by reference in their entirety. Although the composition may be administered as an aqueous solution, it can also be formulated as an emulsion, gel, solution, suspension, lyophilized form, or any other form of solution known in the art. In addition, the composition may contain pharmaceutically acceptable additives including, for example, diluents, binders, stabilizers, and preservatives.

[0246] The following examples are included for purposes of illustration only and are not intended to limit the scope of the present invention, which is defined by the appended claims. All references cited in the Examples are incorporated herein by reference in their entirety.

EXAMPLES

Materials and Methods

[0247] The following materials and methods apply generally to all the examples disclosed herein. Specific materials and methods are disclosed in each example, as necessary.


Gene Construction

[0249] Constructs of the present invention are constructed based on the sequence information provided herein or in the art utilizing standard molecular biology techniques, including, but not limited to the following. First, a series complementary oligonucleotide pairs of 80-90 nucleotides each in length and spanning the length of the construct are synthesized by standard methods. These oligonucleotide pairs are synthesized such that upon annealing, they form double stranded fragments of 80-90 base pairs, containing cohesive ends. The single-stranded ends of each pair of oligonucleotides are designed to anneal with a single-stranded end of an adjacent oligonucleotide duplex. Several adjacent oligonucleotide pairs prepared in this manner are allowed to anneal, and approximately five to six adjacent oligonucleotide duplex fragments are then allowed to anneal together via the cohesive single stranded ends. This series of annealed oligonucleotide duplex fragments is then ligated together and cloned into a suitable plasmid, such as the TOPO® vector available from Invitrogen Corporation, Carlsbad, Calif. The construct is then sequenced by standard methods. Constructs prepared in this manner, comprising 5 to 6 adjacent 80 to 90 base pair fragments ligated together, i.e., fragments of about 500 base pairs, are prepared, such that the entire desired sequence of the construct is represented in a series of plasmid constructs. The inserts of these plasmids are then cut with appropriate restriction enzymes and ligated together to form the final construct. The final construct is then cloned into a standard bacterial cloning vector, and sequenced. The oligonucleotides and primers referred to herein can easily be designed by a person of skill in the art based on the sequence information provided herein and in the art, and such can be synthesized by any of a number of commercial nucleotide providers, for example Retrogen, San Diego, Calif., and GENEART, Regensburg, Germany.

Plasmid Vectors

[0250] Constructs of the present invention can be inserted, for example, into eukaryotic expression vectors VR1012 or VR10551. These vectors are built on a modified pUC18 background (see Yanisch-Perron, C., et al. Gene 33:103-119 (1985)), and contain a kanamycin resistance gene, the human cytomegalovirus immediate early promoter/enhancer and intron A, and the bovine growth hormone transcription termination signal, and a polylinker for inserting foreign genes. See Hartikka, J., et al., Hum. Gene Ther. 7:1205-1217 (1996). However, other standard commercially available eukaryotic expression vectors may be used in the present invention, including, but not limited to: plasmids pcDNA3, pHCMV/Zeoc, pCRII/His, pNlDG/S, pKcIHC MV2, pSV40/Zeoc, pTRACER-HCMV, pUB6/V5-His, pVAX1,
and pzGeoSV2 (available from Invitrogen, San Diego, Calif.), and plasmid pCI (available from Promega, Madison, Wis.).

[0251] An optimized backbone plasmid, termed VR10551, has minor changes from the VR1012 backbone described above. The VR10551 vector is derived from and similar to VR1012 in that it uses the human cytomegalovirus immediate early (hCMV-IE) gene enhancer/promoter and 5' untranslated region (UTR), including the hCMV-IE Intron A. The changes from the VR1012 to the VR10551 include some modifications to the multiple cloning site, and a modified rabbit β globin 3' untranslated region/polyadenylation signal sequence/transcriptional terminator has been substituted for the same functional domain derived from the bovine growth hormone gene.

[0252] Additionally, constructs of the present invention can be inserted into other eukaryotic expression vector backbones such as VR10682 or VR10686. The VR10682 expression vector backbone (SEQ ID NO:94) contains a modified rous sarcoma virus (RSV) promoter from expression plasmid VSL1.005, the bovine growth hormone (BGH) poly-adenylation site and a polylinker for inserting foreign genes and a kanamycin resistance gene. The RSV promoter in VCL1.005 and VR10682 contains a XbaI endonuclease restriction site near the transcription start site in the sequence TAC TCT AGA CG (SEQ ID NO:82). The modified RSV promoter contained in VR10682. Expression plasmid VCL1.005 is described in U.S. Pat. No. 5,561,064 and is incorporated herein by reference.

[0253] The VR10686 expression vector backbone (SEQ ID NO:112) was created by replacing the West Nile Virus (WNV) antigen insert in VR6430 (SEQ ID NO:89) with the multiple cloning site from the VR1012 vector. The VR10686 and VR6430 expression vector backbones contain the RSV promoter, derived from VCL1.005, which has been modified back to the wild-type RSV sequence (TAC AAT AAA CG (SEQ ID NO:83). The wild-type RSV promoter is fused to the "R" region plus the first 39 nucleotides of the U5 region from Human T-Cell Leukemia Virus I (HTLV-I), hereinafter referred to as the R5 element. The R and U5 elements are regions of the long terminal repeat region (LTR) of HTLV-I which control expression of the HTLV-I transcript and is duplicated at either end of the integrated viral genome as a result of the retroviral integration mechanism. The LTR of HTLV-I and most retroviruses are divided into three regions, U3, R and U5. Transcription from the integrated viral genome commences at the U3-R boundary of the 5' LTR and the transcript is polyadenylated at the R-U5 boundary of the 3' LTR. (See Coff, S. P. Retroviridae, Field's Virology 4th ed. 2:1871-1939 (2001). This R5 HTLV-I element has been shown to be a potent stimulator of translation when fused to the SV40 early gene promoter. See Takebe et al., Mol. Cell. Biol. 8:466-472 (1988). It has been proposed that the stimulation of translation by the HTLV-I R5 element is due to its function, in part, as a translational enhancing internal ribosome entry site (IRES). See Attal et al. FEBS Letters 392:220-224 (1996). Additionally the HTLV-I R5 element provides the 5'-splice donor site. Immediately downstream of the R5 element is the 3'-end of the HCMV intron A sequence containing the splice acceptor sequence. The VR10686 and VR6430 expression vectors contain a hybrid intron composed of the 5'-HTLV I intron sequence fused to the 3'-end of the HCMV intron A, a bovine growth hormone poly-adenylation site, a polylinker for insertion of foreign genes and a kanamycin resistance gene. The VR6430 vector expresses the prM and E West Nile Virus antigens (Genebank Accession No. AF202541).

[0254] The vector backbones described above may be used to create expression vectors which express multiple influenza proteins, fragments, variants or derivatives thereof. An expression vector as described herein may contain an additional promoter. For example, construct VR4774 (described in Example 13), contains a CMV promoter and an RSV promoter. Thus, the vector backbones described herein may contain multiple expression cassettes which comprise a promoter and an influenza coding sequence including, inter alia, polynucleotides as described herein. The expression cassettes may encode the same or different influenza polyepitopes. Additionally, the expression cassettes may be in the same or opposite orientation relative to each other. As such transcription from each cassette may be in the same or opposite direction (i.e. 5' to 3' in both expression cassettes or, alternatively, 5' to 3' in one expression cassette and 3' to 5' in the other expression cassette).

Plasmid DNA Purification

[0255] Plasmid DNA may be transformed into competent cells of an appropriate Escherichia coli strain (including but not limited to the DH5α strain) and highly purified covalently closed circular plasmid DNA was isolated by a modified lysis procedure (Horn, N. A., et al., Hum. Gene Ther. 6:565-573 (1995)) followed by standard double CsCl-ethidium bromide gradient ultracentrifugation (Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y. (1989)). Alternatively, plasmid DNAs are purified using Giga columns from Qiagen (Valencia, Calif.) according to the kit instructions. All plasmid preparations were free of detectable chromosomal DNA, RNA and protein impurities based on gel analysis and the bicinechonic acid protein assay (Pierce Chem. Co., Rockford Ill.). Endotoxin levels were measured using Limulus Amoeboocyte Lysate assay (LAL, Associates of Cape Cod, Falmouth, Mass.) and were less than 0.6 Endotoxin Units/mg of plasmid DNA. The spectrophotometric A260/A280 ratios of the DNA solutions were typically above 1.8. Plasmids were ethanol precipitated and resuspended in an appropriate solution, e.g., 150 mM sodium phosphate (for other appropriate excipients and auxiliary agents, see U.S. patent application Publication 2002/0019358, published Feb. 14, 2002). DNA was stored at -20EC until use. DNA was diluted by mixing it with 300 mM salt solutions and by adding appropriate amount of USP water to obtain 1 mg/ml plasmid DNA in the desired salt at the desired molar concentration.

Plasmid Expression in Mammalian Cell Lines

[0256] The expression plasmids were analyzed in vitro by transfecting the plasmids into a well characterized mouse melanoma cell line (VM-92, also known as UM-449). See, e.g., Wheeler, C. J., Sukhu, L., Yang, G., Tsai, Y., Bustamente, C., Felgner, P. Norman, J & Manthorpe, M. "Converting an Alcohol to an Amine in a Cationic Lipid Dramatically Alters the Co-lipid Requirement, Cellular Transfection Activity and the Ultrastructure of DNA-Cytofectin Complexes," Biochim. Biophys. Acta. 1280:1-11 (1996). Other well-characterized human cell lines can also be used, e.g., MRC-5 cells, ATCC Accession No. CCL-171 or human rhadomyosarcoma cell line RD (ATCC CCL-
The transfection was performed using cationic lipid-based transfection procedures well known to those of skill in the art. Other transfection procedures are well known in the art and may be used, for example electroporation and calcium chloride-mediated transfection (Graham F. L. and A. J. van der Eb, *Virology* 52:456-67 (1973)). Following transfection, cell lysates and culture supernatants of transfected cells were evaluated to compare relative levels of expression of IV antigen proteins. The samples were assayed by western blots and ELISAs, using commercially available polyclonal and/or monoclonal antibodies (available, e.g., from Research Diagnostics Inc., Flanders N.J.), so as to compare both the quality and the quantity of expressed antigen.

**Injections of Plasmid DNA**

136. The quadriceps muscles of restrained awake mice (e.g., female 6-12 week old BALB/c mice from Harlan Sprague Dawley, Indianapolis, Ind.) are injected bilaterally with 1-50 μg of DNA in 50 μl solution (100 μg in 100 μl total per mouse) using a disposable plastic insulin syringe and 28G ½ needle (Becton-Dickinson, Franklin Lakes, N.J., Cat. No. 329430) fitted with a plastic collar cut from a micropipette tip, as previously described (Hartikka, J., et al., *Hum. Gene Ther* 7:1205-1217 (1996)).

**Animal care throughout the study was in compliance with the “Guide for the Use and Care of Laboratory Animals”, Institute of Laboratory Animal Resources, Commission on Life Sciences, National Research Council, National Academy Press, Washington, D.C., 1996 as well as with Wal’s Institutional Animal Care and Use Committee.**

**Example 1**

**Construction of Expression Vectors**

**Plasmid constructs comprising the native coding regions encoding NP, M1, M2, HA, and cM2, IV proteins or fragments, variants or derivatives are constructed as follows.**

The NP, M1, and M2 genes from IV (A/PR/8/34) are isolated from viral RNA by RT PCR, or prepared by direct synthesis if the wildtype sequence is known, by standard methods and are inserted into the vector VR10551 via standard restriction sites, by standard methods.

**Plasmid constructs comprising human codon-optimized coding regions encoding NP, M1, M2, HA, cM2, and/or an cM2-NP fusion; or other codon-optimized coding regions encoding other IV proteins or fragments, variants or derivatives either alone or as fusions with a carrier protein, e.g., HBCAg, are prepared as follows. The codon-optimized coding regions are generated using the full, minimal, or uniform codon optimization methods described herein. The codon optimized coding regions are constructed using standard PCR methods described herein, or are ordered commercially. Oligonucleotides representing about the first 23-24 aa extracellular region of M2 are constructed, and are used in an overlap PCR reaction with the NP coding regions described above, to create a coding region for an cM2-NP fusion protein, for example as shown in SEQ ID NOs 6 and 7. The codon-optimized coding regions are inserted into the vector VR10551 via standard restriction sites, by standard methods.

**Plasmids constructed as above are propagated in** *Escherichia coli* **and purified by the alkaline lysis method (Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., ed. 2 (1989)). CsCl-banded DNA is ethanol precipitated and resuspended in 0.9% saline or PBS to a final concentration of 2 mg/ml for injection. Alternately, plasmids are purified using any of a variety of commercial kits, or by other known procedures involving differential precipitation and/or chromatographic purification.**

**Expression is tested by formulating each of the plasmids in DMRIE/DOPE and transfecting VM2 cells. The supernatants are collected and the protein production tested by Western blot or ELISA. The relative expression of the wild type and codon optimized constructs are compared.**

**Examples of constructs made according to the above methods are listed in Table 13. The experimental procedure for generating the listed constructs is as described above, with particular parameters and materials employed as described herein.**

**TABLE 13**

<table>
<thead>
<tr>
<th>Plasmid #</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VR4700</td>
<td>TPA lender - NP (A/PR/34) in VR 1255</td>
</tr>
<tr>
<td>VR4707</td>
<td>TPA lender-M2 with transmembrane deletion, glycine linker inserted</td>
</tr>
<tr>
<td>VR4710</td>
<td>TPA lender - 1st 24 amino acids of M2 from VR4707 fused to NP from VR4700</td>
</tr>
<tr>
<td>VR4750</td>
<td>full length HA from mouse adapted virus (H5, H9, Hong Kong 68)</td>
</tr>
<tr>
<td>VR4752</td>
<td>full length HA from mouse adapted virus (H5, Puerto Rico 34)</td>
</tr>
<tr>
<td>VR4755</td>
<td>algorithm to codon optimize consensus amino acid sequence, direct fusion M2 to ATG of M1</td>
</tr>
<tr>
<td>VR4756</td>
<td>native sequence from A/Nilgai/137/76 influenza strain (matches amino acid consensus sequence)</td>
</tr>
<tr>
<td>VR4757</td>
<td>Contracted codon optimized - 1st 24 amino acids of M2 from consensus fused to full-length NP consensus</td>
</tr>
<tr>
<td>VR4758</td>
<td>Applicants' codon optimized - 1st 24 amino acids of M2 from consensus fused to full-length NP consensus</td>
</tr>
<tr>
<td>VR4759</td>
<td>Full-length M2 derived from VR4755</td>
</tr>
<tr>
<td>VR4760</td>
<td>Full-length M1 derived from VR4755</td>
</tr>
<tr>
<td>VR4761</td>
<td>Full-length NP derived from VR4757</td>
</tr>
<tr>
<td>VR4762</td>
<td>Full-length NP derived from VR4758</td>
</tr>
<tr>
<td>VR4763</td>
<td>Selectively codon-optimized regions of segment 7</td>
</tr>
</tbody>
</table>

**The pDNA expression vector VR4700 which encodes the influenza NP protein has been described in the art. See, e.g. Sankar, V., Baccaglini, L., Sawdey, M., Wheeler, C. J., Pillemer, S. R., Baum, B. J. and Atkinson, J. C., “Salivary Gland Delivery of pDNA-Cationic Lipopolypexis Elicits Systemic Immune Responses,” *Oral Diseases* 8:275-281 (2002). The following is the open reading frame for TPA-NP (from VR4700), referred to herein as SEQ ID NO:46:**
1 attgatgcaa tgaagagaggg gctctgcggc gttctggccc tgtctggtgg agcttctgtt
61 tegcccaagc ctgagagagct ggaacagggc tccaagggg ccaacggtgc ttcaagacag
121 attggcaggt attggcaggt ccagagagtc actgtaactc gatctgcggt cgganacatt
181 attggcaggt attggcaggt cctgctccag actgctgcgag acctcaact gccttgattat
241 gaggagcgtt tgaggcaaca cagcattaca atagagagaa tggagctcct tcgctttggc
301 gaaagagagaa ataaactacct ggaagaaaccc cccagtgaggg tggagacact taagaaact
361 ggaagcgttat ctagagagag aagaaacaggg ataagactga gagaactcat cttttatgc
421 aagaaagagaa taagagagaa ctgagccagc gtaaaattat gctagagagtg gcggagcgtgt
481 ctgagctca cttgctgctg gcagctcact ttgaagagat caagagctcaca gaggacagga
541 gotctgctcc ttgcggcagag ggcctcgcgg tgtcggcctt tgtgctgaggg ttcagcctgc
601 cctagagagt ctgagcggcgc aaggagcagc ggcaagacag tggagacact ggtgaggaag
661 ttggcagagc tgtcaaacag tggagcact gctgagagat tcttgagggg tggagacagta
721 cggagagacca gaaagctgta tggagagagat tggagctcact tctcgagtgga atctcaact
781 gcgctcggag aagagagagat gcggagcagc gcggccagag ggtgaggaag
841 ttcagagagt ctgagctcact tggtacaaggg acctcagctc ctgctggctg gcggagcagc
901 attgctgagc tgtgctgctg atctcgagag cctcctcagct gcggagcagc gcggagcagc
961 agagagcagc atctgagcct ggaagagagat tggagacact ttcctcagct gcggagcagc
1021 tacagcgcagc tggagacact gcggagcagc gcggagcagc gcggagcagc
1081 tggagagagat gcggagcagc gcggagcagc gcggagcagc gcggagcagc
1141 cttcagcagc ggaagcctcc cattagagcc cctagagac gcagcttcagc cttcagcagc
1201 aatgagagagat gcgctcgcag tggagacact gcggagcagc gcggagcagc
1261 gggagagagat cgaacagagc ggggctcgcag cggagagagat gcgctcgcag cggagagagat
1321 tcagagcgcagc gcagcttcagc tttgcagagc gcagcttcagc cggagagagat gcgctcgcag
1381 acaagagagagat cttgctagagc gcagcttcagc cggagagagat gcgctcgcag cggagagagat
1441 ccagagagagat cttgctcggag cggagagagat cttgctagagc gcagcttcagc acaagagagat
1501 cggagagagat cttgctcggag cggagagagat cttgctagagc gcagcttcagc ccagagagagat
1561 cggagagagat cttgctcggag cggagagagat cttgctagagc gcagcttcagc cccagagagagat

[0265] Purified VR4700 DNA was used to transfet the murine cell line VM92 to determine expression of the NP protein. Expression of NP was confirmed with a Western Blot assay. Western blot analysis showed very low level expression of VR4700 in vitro as detected with mouse polyclonal anti-NP antibody. In vivo antibody response was detected by ELISA with an average titer of 62,578.

[0266] Plasmid VR4707 expresses a secreted form of M2, i.e., TPA-M2. The sequence was assembled using synthetic oligonucleotides in which the oligos were annealed amongst themselves, and then ligated and gel purified. The purified product was then ligated (cloned) into Eco RI/Sal I of VR10551. The M2 sequence lacks the transmembrane domain; the cloned sequence contains amino acids [TPA(1-23)]ARGSG(M2(1-25))GGG(M2(44-97)). Amino acid residues between TPA and M2 and between M2 domains were added as flexible linkers. The following mutations were introduced to generate appropriate T-cell epitopes: 74S→G and 78S→N. The following is the open reading frame for TPA-M2ATM (from VR4707), referred to herein as SEQ ID NO:47:

1 attgatgcaa tgaagagaggg gctctgcggc gttctggccc tgtctggtgg agcttctgtt
61 tegcccaagc ctgagagagct ggaacagggc tccaagggg ccaacggtgc ttcaagacag
[0267] Purified VR4707 DNA was used to transfect the murine cell line VM92 to determine expression of the M2 protein. Expression of M2 was confirmed with a Western Blot assay. Expression was visualized with a commercially available anti-M2 monoclonal antibody. In vivo M2 antibody response to VR4707, as assayed by ELISA, resulted in an average titer of 110, which is lower than the average titer of 9,240 for VR4756, encoding full-length M2 from segment 7. An IFNγ ELISPOT assay for M2-specific T cells resulted in an average of 61 SUF/10⁶ cells versus an average of 121 SUF/10⁶ cells for the segment 7 construct.

[0268] VR4710 was created by fusing the TPA leader and the first 24 amino acids of M2 from VR4707 to the full-length NP gene from VR4700. Primers 5'-GCCGAATTC-GGATGCAATGGAAG-3' (SEQ ID NO:48) and 5'-GGTCGCTTGGGAGGCGCATATACTGAAATATGCCGA-3' (SEQ ID NO:49) were used to amplify the TPA-M2 fragment from VR4707. Primers 5'-TGCAAGAATCAATATGGATGGCGCTGACCATCAACTGAAATATGCCGA-3' (SEQ ID NO:50) and 5'-GCCGACTTATCAATGGCATGCGCTGACCATCAACTGAAATATGCCGA-3' (SEQ ID NO:51) were used to amplify the NP gene from VR4700. Then the N-terminal and C-terminal primers were used to assemble the fusion, and the eM2NB fusion was cloned into VR10551 as an EcoRI-Sall fragment. The following is the open reading frame for TPA-M2-NP (from VR4710), referred to herein as SEQ ID NO:52:

```
1 atggtagcaca tgaagagagg ggttgctgtg gtctgctgc ctgctgggct atgtctgtg
61 tgcgcgacag tctcgagtcag cggattcagt cttctcgacg cggattcag cccattcag
121 aagaaagtgg ggtgctagcag acgagttcag cggccgctg cccatcagc
181 tctttgacag cggctgctg ggtcgcgag cggcgcgag cccatcagc
241 tcgctggagag acctggaaaccttgggagct acctggaaaccttgggagct
301 cttctttg acgatcggag ttttttcgctg gcttctggag ttttttcgctg gcttctggag
361 tctttttg acgatcggag ttttttcgctg gcttttttgcctt cttcttttgcctt cttcttttgcctt
421 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
481 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
541 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
601 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
661 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
721 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
781 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
841 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
901 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
961 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
1021 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
1081 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
1141 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
1201 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
1261 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
1321 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
1381 cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt cttcttttgcctt
```
[0269] Purified VR4710 DNA was used to transfect the murine cell line VM92 to determine expression of the cM2-NP fusion protein. Expression of EM2-NP was confirmed with a Western Blot assay. Expression was visualized with a commercially available monoclonal antibody to M2 and with mouse polyclonal antibody to NP. ELISA assay results following 2 injections of pDNA into mice revealed little antibody response to M2, but an average titer of 66,560 for anti-NP antibody.

[0270] VR4750 was created by first reverse transcribing RNA from the mouse-adapted A/Hong Kong/1/68 virus stock using random hexamer to create a cDNA library. Then primers 5’ GGGCTAGCGCGCCACCATGAAGACCATCAATTGTCT 3’ (SEQ ID NO:53) and 5’ CCGTCGACTCAATGAAAAATTGTTGCA 3’ (SEQ ID NO:54) were employed to PCR the HA gene. The gene was inserted into the Invitrogen TOPO-TA vector first, and then sub-cloned into VR10551 using restriction enzymes NheI and SalI. The following is the open reading frame for HA (H3N2) from mouse-adapted A/Hong Kong/68 (from VR4750), referred to herein as SEQ ID NO:55:

1 atggagaca ttaattcctt gacgtacttt tttgctgctt cttctgycoc aagctttcga
61 ggaaagcaac ccacacacag aagctgtgc cttgagcact aagctgtgc cttgagcact cttgagcact
121 ctgtgaaaact acacacgcat gcaacgca gcaacgcat gcaacgca gcaacgca gcaacgca
181 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
241 acacacgcat gcaacgca gcaacgca gcaacgca gcaacgca gcaacgca gcaacgca
301 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
361 tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt
421 tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt
481 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
541 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
601 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
661 tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt
721 tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt
781 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
841 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
901 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
961 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
1021 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
1081 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
1141 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
1201 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
1261 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
1321 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
1381 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
1441 aagcttttca tttgacactt cttctgacac aagcttttca tttgacactt cttctgacac aagcttttca
GCAAACTACTG 3’ (SEQ ID NO:56) and 5’ CCGTC-GACTCAGATGCTATTCTGCA 3’ (SEQ ID NO:57) were employed to PCR the HA gene. The gene was then cloned into the TOPO-TA vector first, and then sub-cloned into VR10551 using restriction enzymes NheI and Sall. The following is the open reading frame for HA (H1N1) cloned from mouse-adapted A/Puerto Rico/34 (from VR752), referred to herein as SEQ ID NO:58:

1 atggacgca acatagtctg cgtgctatgt gcatactgag ctgcaagagc agacaaaata
15 61 tgcgacgtct acctggagc cagaatgcc agcactgctg ccggaagact gcgaagagac
155 121 tgcgacgtct acctggagc cagaatgcc agcactgctg ccggaagact gcgaagagac
205 181 tgaagagct acctggagc cagaatgcc agcactgctg ccggaagact gcgaagagac
259 241 tgcgacgtct acctggagc cagaatgcc agcactgctg ccggaagact gcgaagagac
313 301 tgcgacgtct acctggagc cagaatgcc agcactgctg ccggaagact gcgaagagac
365 361 tgcgacgtct acctggagc cagaatgcc agcactgctg ccggaagact gcgaagagac
419 421 tgcgacgtct acctggagc cagaatgcc agcactgctg ccggaagact gcgaagagac
473 481 ttcgagagac atggatctgt gcattgagaa cctggagcag cctgctgtct cacgctgtct
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
gagattgctt cagctggatg tggactgctt ctggatgctt ccgctgttct cacgctgttt
[0273] Purified VR4752 DNA was used to transfect the murine cell line VM92 to determine expression of the HA protein. Expression of HA was confirmed with a Western Blot assay. Expression was visualized with a commercially available goat anti-influenza A (H1N1) antibody.

[0274] A direct fusion of the M2 gene to the M1 gene was synthesized based on a codon-optimized sequence derived from methods described in Example 4 using the “universal” optimization strategy. The synthesized gene was received in the pUC119 vector and then sub-cloned into VR10551 as an EcoRI-SalI fragment. The following is the open reading frame for the M2M1 fusion (from VR4755), referred to herein as SEQ ID NO:59:

```plaintext
1  atgagctgag cgcgggaggt ggcagaccccg atcagaaacg agtggggcttg cgagatcagcc
  61 gcacagcaac aacccctgtg cttgagcgctg aagcatactg gcacacgatt ctgtgacagt
g 121 tgatctcctg aggagacttg gttcaagagt tttcagagtc tttcatagtgg ctgagagccta
  181 agcggccca agcggccggg cttgagcctg cttgagcctc aagcactag ccagacagagc
g 241 cgagagcccg aacccctgtg cttgagcctc tttcagagtc ctggagcttg gatgtcctctg
  301 ctgagagcccg agcggccca agcggccggg cttgagcctc tttcagagtc ctggagcttg gatgtcctctg
  361 atcagaaacg agcggccca agcggccggg cttgagcctg cttgagcctc aagcactag ccagacagagc
g 421 gcacagcaac aacccctgtg cttgagcgctg aagcatactg gcacacgatt ctgtgacagt
g 481 ttgagatcctg ccagagacag ggacagacscc cttgagacatg gcagagacac ccagagacagc
g 541 ctgagagcccg aacccctgtg cttgagcctc tttcagagtc ctggagcttg gatgtcctctg
  601 ccagagacag ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac
  661 ctgagagcccg aacccctgtg cttgagcctc tttcagagtc ctggagcttg gatgtcctctg
  721 ccagagacag ggacagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac
  781 ccagagacag ggacagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac
  841 ccagagacag ggacagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac
  901 ccagagacag ggacagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac
  961 ccagagacag ggacagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac ccagagacac
 1021 atgagctgag aacccctgtg cttgagagtc ctggagcttg gatgtcctctg
```

[0275] Purified VR4755 DNA was used to transfect the murine cell line VM92 to determine expression of the M2M1 fusion protein. Expression of M2M1 was confirmed with a Western Blot assay. Expression of the M2M1 fusion was visualized with commercially available anti-M1 and anti-M2 monoclonal antibodies.

[0276] The segment 7 RNA of influenza A encodes both the M1 and M2 genes. A consensus amino acid sequence for M1 and M2 was derived according to methods described herein. The consensus sequences for both proteins, however, are identical to the M1 and M2 amino acid sequences derived from the IV strain A/Niigata/137/96, represented herein as SEQ ID NO:77 and SEQ ID NO:78, respectively. Accordingly, the native sequence for segment 7, A/Niigata/137/96, was synthesized and received as an insert in pUC119. The segment 7 insert was sub-cloned into VR10551 as an EcoRI-SalI fragment. The following is the open reading frame for segment 7 (from VR4755), referred to herein as SEQ ID NO:60:
1atgagccttc taacggagt cgaacggtat gttctctota tcgctcatac aggcccccctc
g1aaagcggaa tcgcggcgac actctgagat gttcttgcgt ggaacaaac agatctttag
21gcttctatgg actgttgaa caacagacca atctctgtaco cctgtggtctggggatat
181ggggttggtc tgcagctcaco gcgtccacagt gcagacagct gcagcgcttac gcgtggtag
c241aaacctgctc taatggggga aattgaatgc acacagactg taacatataat
301agaaagacta aagggggat aacctcctcac ggggcataag aataaggtcg cagtttacctc
361gtcgttccac ttcgtcagcgt caggtggctt atactccatac atacataacca gatgggggcc tggctgccacct
421ggttagcttc tcggctcgac gtctgacgaa tctgcggccttc ctgctgccag gacgcagcag
481ctctgacgac atacgggcc aaaacaaact caactaatg tggcttctgaa gcaagctgtt
541tgcgcagaca ctacgggaa ggtcattggag ccaagggaggc gatccagggcg
gcgcagag
601gaggcctctg aacctctgat tgcgccaacgg cagagggctg agccattgag
ggcagctgctgcagag
661actctctgca ggtcggcagc ctgcctacact gcagtgtcagc atgcggtcag
g611caagcggca tggggggaga gatgcagcag cccctgctgcag
gtgcgccata
g701ctcctgcttctt cttggttttg ggttcttttc cttctctctc
g841ctgctgctctcagggcttcatcagggagccctgttgcagtc
901atcaggga gatctagcgc aagggccagc gactcttgag gcagctgcttgcaggtgctg
961tgctagcatg gtctggagtgtc aa

[0277] SEQ ID NO:77 ("consensus" (A.Niihata/137/96) M1):
MSSLEVEYTVLSPQEGDLKAETQARLEDFPGANMDFEANLMKTRP
ILSFLSRLIGLGFPTLTVPERSKLGRRRFQCHLNGKDFPMQRAVLY
RLKSEITRHPGEAKLGGALASCMILYIRSGAVTTEAVGLVCACT
CEIQADQGRSRHMQWATTNLQTHNHYVLAUSTKAMEQMSGASEQAA
EAMEIA6QARQVMQAMRAIGTHFP666ALGKQDLLENLQTQ7KR6GVQM
QRFPK

[0278] SEQ ID NO:78 ("consensus" (A.Niihata/137/96) M2):
MSSLEVEYTVLSPQEGDLKAETQARLEDFPGANMDFEANLMKTRP
ILSFLSRLIGLGFPTLTVPERSKLGRRRFQCHLNGKDFPMQRAVLY
RLKSEITRHPGEAKLGGALASCMILYIRSGAVTTEAVGLVCACT
CEIQADQGRSRHMQWATTNLQTHNHYVLAUSTKAMEQMSGASEQAA
EAMEIA6QARQVMQAMRAIGTHFP666ALGKQDLLENLQTQ7KR6GVQM
QRFPK

[0279] Purified VR4756 DNA was used to transfect the murine cell line VM92 to determine expression of the mutated proteins encoded by segment 7. Expression of both M1 and M2 was confirmed with a Western blot assay using commercially available anti-M1 and anti-M2 monoclonal antibodies. ELISA assay results following 2 injections of pDNA into mice revealed an average M2 antibody titer of 9,240 versus a 110 average titer for VR4707. An IFNγ ELISPOT assay for M2-specific T cells resulted in an average of 121 SFU/106 cells for VR4756 injected mice versus an average of 61 SFU/106 cells for the VR4707 construct.

[0280] An additional segment 7 sequence is created, VR763, which contains selectively codon-optimized regions of segment 7. Optimization of the coding regions in segment 7 is selective, because segment 7 contains two overlapping coding regions (i.e., encoding M1 and M2), and these coding regions are partially in different reading frames. From the AUG encoded by nucleotides 1 to 3 of segment 7, M1 is encoded by bp 1 through 759 of the segment 7 RNA, while M2 is encoded by a spliced messenger RNA which includes nucleotides 1 to 26 of segment 7 spliced to nucleotides 715 to 982 of segment 7. Optimization of the region from 715 to 759 is avoided because the M1 and M2 coding sequences (in different reading frames) overlap in that region. Due to the splicing that occurs to join bp 26 to an alternate frame at bp 715 of the segment 7 sequence, optimization in these splicing regions is also avoided; adjacent regions that arguably could also participate in splicing are likewise avoided. Optimization is done in a manner to ensure that no new splicing sites are inadvertently introduced. The areas that are optimized are done so using "universal" strategy, e.g. inserting the most frequently used codon for each amino acid. The following is the nucleotide sequence for codon-optimized segment 7 (from VR763), referred to herein as SEQ ID NO:61:

1atgagccttc taacggagt cgaacggtat gttctctota tcgctcatac aggcccccctc
g1aaagcggaa tcgcggcgac actctgagat gttcttgcgt ggaacaaac agatctttag
21gcttctatgg actgttgaa caacagacca atctctgtaco cctgtggtctggggatat
181ggggttggtc tgcagctcaco gcgtccacagt gcagacagct gcagcgcttac gcgtggtag
c241aaacctgctc taatggggga aattgaatgc acacagactg taacatataat
301agaaagacta aagggggat aacctcctcac ggggcataag aataaggtcg cagtttacctc
361gtcgttccac ttcgtcagcgt caggtggctt atactccatac atacataacca gatgggggcc tggctgccacct
421ggttagcttc tcggctcgac gtctgacgaa tctgcggccttc ctgctgccag gacgcagcag
481ctctgacgac atacgggcc aaaacaaact caactaatg tggcttctgaa gcaagctgtt
541tgcgcagaca ctacgggaa ggtcattggag ccaagggaggc gatccagggcg
gcgcagag
601gaggcctctg aacctctgat tgcgccaacgg cagagggctg agccattgag
ggcagctgctgcagag
661actctctgca ggtcggcagc ctgcctacact gcagtgtcagc atgcggtcag
g611caagcggca tggggggaga gatgcagcag cccctgctgcag
gtgcgccata
g701ctcctgcttctt cttggttttg ggttcttttc cttctctctc
g841ctgctgctctcagggcttcatcagggagccctgttgcagtc
901atcaggga gatctagcgc aagggccagc gactcttgag gcagctgcttgcaggtgctg
961tgctagcatg gtctggagtgtc aa

1atgagccttc taacggagt cgaacggtat gttctctota tcgctcatac aggcccccctc
g1aaagcggaa tcgcggcgac actctgagat gttcttgcgt ggaacaaac agatctttag
21gcttctatgg actgttgaa caacagacca atctctgtaco cctgtggtctggggatat
181ggggttggtc tgcagctcaco gcgtccacagt gcagacagct gcagcgcttac gcgtggtag
c241aaacctgctc taatggggga aattgaatgc acacagactg taacatataat
301agaaagacta aagggggat aacctcctcac ggggcataag aataaggtcg cagtttacctc
361gtcgttccac ttcgtcagcgt caggtggctt atactccatac atacataacca gatgggggcc tggctgccacct
421ggttagcttc tcggctcgac gtctgacgaa tctgcggccttc ctgctgccag gacgcagcag
481ctctgacgac atacgggcc aaaacaaact caactaatg tggcttctgaa gcaagctgtt
541tgcgcagaca ctacgggaa ggtcattggag ccaagggaggc gatccagggcg
gcgcagag
601gaggcctctg aacctctgat tgcgccaacgg cagagggctg agccattgag
ggcagctgctgcagag
661actctctgca ggtcggcagc ctgcctacact gcagtgtcagc atgcggtcag
g611caagcggca tggggggaga gatgcagcag cccctgctgcag
gtgcgccata
g701ctcctgcttctt cttggttttg ggttcttttc cttctctctc
g841ctgctgctctcagggcttcatcagggagccctgttgcagtc
901atcaggga gatctagcgc aagggccagc gactcttgag gcagctgcttgcaggtgctg
961tgctagcatg gtctggagtgtc aa
The codon optimized coding region for M1 extends from nucleotide 1 to nucleotide 759 of SEQ ID NO:61 including the stop codon, and is represented herein as SEQ ID NO:79. The codon-optimized coding region for M2 extends from nucleotide 1 to nucleotide 26 of SEQ ID NO:61 spliced to nucleotide 715 through nucleotide 959 of SEQ ID NO:61, including the stop codon, and is represented herein as SEQ ID NO:80.

**Optimized M1 Coding Region (SEQ ID NO:79):**

```
ATGAAGCTGCTGACACGAGGCTGACAACCACTGATGCTCTATGCTCAAGCCAGCAGGAACTGAGAATGCTGAGG
CCGGCCCTGAAAGCCGAGATGCCCGAGCCACGAGATGCTCAGGGAGAGCGATGATGCTG
```

**Optimized M2 Coding Region (SEQ ID NO:80):**

```
ATGAAGCTGCTGACACGAGGCTGACAACCACTGATGCTCTATGCTCAAGCCAGCAGGAACTGAGAATGCTGAGG
CCGGCCCTGAAAGCCGAGATGCCCGAGCCACGAGATGCTCAGGGAGAGCGATGATGCTGAGG
```

**Optimized eCM2-NP:** The eCM2-NP fusion was codon-optimized, inserted in pUC119 and sub-cloned into VR10551 as an EcoRI-SalI fragment. The following is the open reading frame for eCM2-NP: codon-optimized by Contract (from VR4757), referred to herein as SEQ ID NO:62:

```
1 atgagttgcg taaactggaag gagacacgca atcgaaaccg aatggaagatc tagatcgac
61 gatactgac actagctgctc caagggacag aaaaaaagat catgacagat ggaactgac
121 ggagagacag aagacgctac agagatcagc gcatgtgatg gagatgatg aagtgtcagc
```
The eM2-NP fusion gene in VR4758 was codon-optimized and synthesized. The gene was inserted into pUC119 and sub-cloned into VR10551 as an EcoRI-SalI fragment. The following is the open reading frame for eM2-NP: codon-optimized by Applicants (from VR4758), referred to herein as SEQ ID NO:63:

[0285] Purified VR4757 DNA was used to transfect the murine cell line VM92 to determine expression of the eM2-NP fusion protein. Expression of eM2-NP was confirmed with a Western Blot assay. Expression was visualized with a commercially available monoclonal antibody to M2 and with mouse polyclonal antibody to NP. In vivo antibody response to NP was detected by ELISA with an average titer of 51,200.
[0287] Purified VR4758 DNA was used to transfect the murine cell line VM92 to determine expression of the eM2-NP protein. Expression of eM2-NP was confirmed with a Western Blot assay. Expression was visualized with a commercially available monoclonal antibody to M2 and with mouse polyclonal antibody to NP. In vivo antibody response to NP was detected by ELISA with an average titer of 48,640.

[0288] The M2 gene was PCR-amplified from VR4755 using the primers 5’-GCCGAAATCGCCACGATCAGCTGAC-3’ (SEQ ID NO:64) and 5’-GCCGACAGCT- GATCCACGAGCTGAC-3’ (SEQ ID NO:65) and sub-cloned into VR10551 as an EcoRI-Sall fragment. The following is the open reading frame for M2 (from VR4759), referred to herein as SEQ ID NO:66:
[0289] Purified VR4759 DNA was used to transfect the murine cell line VM92 to determine expression of the M2 protein. Expression of M2 was confirmed with a Western Blot assay. Expression was visualized with a commercially available anti-M2 monoclonal antibody.

[0290] The M1 gene was PCR-amplified from VR4755 using the primers 5'-GCCGAATTCGCAACATGCTCTGCACTG-3' (SEQ ID NO:67) and 5'-GCTGCAAGCTATTCTCTGCATG-3' (SEQ ID NO:68) and sub-cloned into VR10551 as an EcoRI-Sall fragment. The following is the open reading frame for M1 (from VR4760), referred to herein as SEQ ID NO:69:

[0291] Purified VR4760 DNA was used to transfect the murine cell line VM92 to determine expression of the M1 protein. Expression of M1 was confirmed with a Western Blot assay. Expression was visualized with a commercially available anti-M1 monoclonal antibody.

[0292] The NP gene was PCR-amplified from VR4757 using primers 5'-GCCGAATTCGCAACATGCTCTGCATG-3' (SEQ ID NO:70) and 5'-GCTGCAAGCTATTCTCTGCATG-3' (SEQ ID NO:71) and sub-cloned into VR10551 as an EcoRI-Sall fragment. The following is the open reading frame for NP: codon-optimized by Contract (from VR4761), referred to herein as SEQ ID NO:72:

1 atgctccgctg acacaggtg gggcctagc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaagctgctaacgc gcagcttggc ttagcacaag}
[0293] Purified VR4761 DNA was used to transfect the murine cell line VM92 to determine expression of the NP protein. Expression of NP was confirmed with a Western Blot assay. Expression was visualized with a mouse polyclonal anti-NP antibody. In vitro expression of VR4761 was significantly higher than VR4700 and comparable to VR4762.

[0294] The NP gene was PCR-amplified from VR4758 using primers 5'-GCGGAATTCCACCAAGGCAAGCCACGAC-3' (SEQ ID NO:73) and 5'-GCGGATCCTGGAGTCTGAGTGATCGTGCTACGTTCC-3' (SEQ ID NO:74) and sub-cloned into VR10551 as an EcoRI-Sall fragment. The following is the open reading frame for NP: codon-optimized by Applicants (from VR4762), referred to herein as SEQ ID NO:75:

```
1  atg goc acc cag ggc acc aag aag cag tac gag cag atg gac acc acc gac
  49  ggc gag aga cag aac cag acc gag atc aga cgc agc gtc ggg aag atg
  97  atc gac ggc atc ggc aga ccc tac acc atg tgc acc acc gag ctc cag
 145  ctc agc gag tac gag ggc aga ctc atc cag acc agc atg acc atc gag
 193  aag atg gtc atc agc ccc cgc cgg gag aga aag aag gag tac atg gag
 241  cag cac ccc aag ggc acc aag gag cac ccc aag aag acc ggc ggc ccc atc
 289  tac aga aga gtt gac ggc gga tgg atg aag aag gtt atg ctt tac gac
 337  aag gag gac atc aga aga atc tgg aga cag ggc aac aac ggc gag gac
 385  ggc acc gcc ggc tgt acc cac atg atc atc tgg cac agc aag agc atg acc
 433  gac acc acc tac cag aag acc aag gcc ctt gtt ctc ggg acc acc aag aag
 481  aag aag atg tgc atg ctc cag ggc agc acc ctc ccc aag aag aag
 529  ggc goc gcc ggc ggc ggc gtt aag ggc atc gcc acc atg tgg atg gac
```
-continued

577 ctg atc aga atg atc aag aga ggc atc aac gac aga aac ccc tgg aga
625 ggc gag aac ggc aga aag aag ggc gac tac gag aga atg tgc aac
673 act ctc aag ggc aag ttc cag acc ggc ccc gac cag aga ggc atg atg gac
721 cac gtc cgg gag aga aac aac ccc ggc aac ggc gag atc gag gac ctg
769 act ttc ctc ggc aga agc ggc tgg atc ctg ctc aga ggc aag ggg tgg ggc cac
817 aag agc tgg ctg ccc ggc tgg ctg cac gcc ccc gcc ggc atg agc aag ggc
865 cac gcc ccc gag aag gag ggg cac acc atg atg ggc aac gcc ccc ccc
913 aag ctc ctc cag aac agc cag tgg tac aag ctc aag atc aag ccc aac gag
961 cac ccc gcc ccc cag aac agc cag ctc ctc ggg ggc aag ggg aac gcc ccc agc
1009 gcc ttc ggct gac ctg aag atg gcc ttc aag gcc aac cag cag cag ggc
1057 ccc ccc aag ggc cag aag ctc aag gcc tgg cag aag atc gcc aag aac
1105 gag aac atg gag aac atg ggc aag agc aac acc ctg atg gtc aag aga aag
1153 tat tgg gcc aac gcc aag aac cag aag cgg ggc aac aac ccc cag cag ggc
1201 gcc agc gcc ggc cag atc aag atg ccc acc ttc agc gtg cag aag
1249 aac ctc ccc ttc gag aag agc cag tgg atg gcc ttc aac gcc aac
1297 aac ggg aag ccc ggc aga atg aag ggc gag atc aag atg atg agc
1345 gag ggc ggc aag ccc gag gag tgg ccc ttc aga aag ggc aag ggc gtt ttc
1393 gag ctg agc ggc gag aag gcc ccc aac acc atc gct atc agc ttc gac
1441 atg agc aag gag ggc agc ttc ttc ggc gac aac gcc gag gag ttc
1489 gcc aac tga

[0295] Purified VR4762 DNA was used to transfect the murine cell line VM92 to determine expression of the NP protein. Expression of NP was confirmed with a Western Blot assay. Expression was visualized with a mouse polyclonal anti-NP antibody. In vitro expression of VR4762 was significantly higher than VR4700 and comparable to VR4761.

[0296] In addition to plasmids encoding single IV proteins, single plasmids which contain two or more IV coding regions are constructed according to standard methods. For example, a polycistronic construct, where two or more IV coding regions are transcribed as a single transcript in eukaryotic cells may be constructed by separating the various coding regions with IRES sequences. Alternatively, two or more coding regions may be inserted into a single plasmid, each with their own promoter sequence.

Example 2

Preparation of Recombinant NP DNA and Protein

[0297] Recombinant NP DNA and protein may be prepared using the following procedure. Eukaryotic cells may be used to express the NP protein from a transfected expression plasmid. Alternatively, a baculovirus system can be used wherein insect cells such as, but not limited to, SF9, SF21, or D.Mel-2 cells are infected with a recombinant baculovirus which can express the NP protein. Cells which have been infected with recombinant baculoviruses, or contain expression plasmids, encoding recombinant NP are collected by knocking and scraping cells off the bottom of the flask in which they are grown. Cells infected for 24 or 48 hours are less easy to detach from flask and may lyse, thus care must be taken with their removal. The flask containing the cells is then rinsed with PBS and the cells are transferred to 250 ml conical tubes. The tubes are spun at 1000 rpm in a J-6 centrifuge (3000g) for about 5-10 minutes. The cell pellets are washed two times with PBS and then resuspended in about 10-20 ml of PBS in order to count. The cells are finally resuspended at a concentration of about 2x10^6 cells/ml in RSB (10 mM Tris pH=7.5, 1.5 mM MgCl2, 10 mM KCl).

[0298] Approximately 10^5 cells are used per lane of a standard SDS-PAGE mini-protein gel which is equivalent to the whole cell fraction for gel analysis purposes. 10% NP40 is added to the cells for a final concentration of 0.5%. The cell-NP40 mixture is vortexed and placed on ice for 10 minutes, vortexing occasionally. After ice incubation, the cells are spun at 1500 rpm in a J-6 centrifuge (600xg) for 10 minutes. The supernatant is removed which is the cytoplasmic fraction. The remaining pellet, containing the nuclei, is washed twice with buffer C (20 mM HEPES pH=7.9, 1.5 mM MgCl2, 0.2 mM EDTA, 0.5 mM PMSE, 0.5 mM DTI) to remove cytoplasmic proteins. The nuclei are resuspended in buffer C to 5x10^7 nuclei/ml. The nuclei are vortexed vigorously to break up particles and an aliquot is removed for the mini-protein gel which is the nuclei fraction.
To the remaining nuclei a quarter of the volume of 5M NaCl is added and the mixture is sonicated for 5 minutes at a maximum output in a bath-type sonicator at 4°C, in 1-2 minute bursts, resting 30 seconds between bursts. The sonicated mixture is stirred at 4°C, then spun at 12000g for 10 minutes. A sample is removed for the protein mini-gel equivalent to approximately 10⁵ nuclei. The sample for the gel is centrifuged and the supernatant is the nuclear extract and the pellet is the nuclear pellet for gel analysis.

For gel analysis, a small amount (about 10⁵ nuclear equivalents) of the nuclear pellet is resuspended directly in gel sample buffer and run with equivalent amounts of whole cells, cytoplasm, nuclei, nuclear extract and nuclear pellet. The above method gives relatively crude NP. To recover NP of a higher purity, 2.1 M NaCl can be added to the nuclear pellet instead of 5M NaCl. This will bring the salt content to 0.42M NaCl. The supernatant will then contain about 60-70% of the total NP plus nuclear proteins. The resulting pellet is then extracted with 1M NaCl and centrifuged as above. The supernatant will contain NP at more than 95% purity.

Example 3

Consensus Amino Acid Sequences of NP, M1 and M2

By analyzing amino acid sequences from influenza strains sequenced since 1990, consensus amino acid sequences were derived for influenza NP, M1 and M2 antigens.

NP Consensus Amino Acid Sequence

The method by which amino acid sequences for influenza NP (strain A) was chosen is as follows. The http://www.flu.lanl.gov database containing influenza sequences for each segment was searched for influenza A strains, human, NP, amino acids. Results gave about 400 sequences, the majority of which were of only partial sequences. The sequences were subsequently narrowed down to 85 approximately full length sequences. If different passages of the same strain were found, the earliest passage was chosen. The sequences were further narrowed down to 28 full length NP sequences isolated from 1990 to 2000 (no full-length sequences from 2001-2003). Five additional sequences were eliminated which were identical to another sequence isolated from the same year based on the assumption that sequences with the same year and identical amino acid sequences were likely to be the same virus strain (in order to avoid double weighting). If there were sequences from the same year with different amino acid sequences, both sequences were kept.

Sequences were aligned to the A/PR/8/34 strain in decending order by most recent, and the consensus sequence was determined by utilizing the amino acid with the majority (FIG. 12). There are 32 amino acid changes between the A/PR/8/34 and the consensus sequence, and all amino acid changes are also present in the two year 2000 NP sequences. For one additional amino acid (aa 275) 15/23 have changed from E (in A/PR/34) to G/D or V (7G, 7D, 1V). Since the two 2000 strains both contain a G at this position, G was chosen. The changes total 33 amino acids, which is about a 7% difference from the A/PR/8/34 strain.

The dominant Ball/c epitope TYQRTRALV is still maintained in the new consensus; changes to other theoretical human epitopes have not been determined as yet.

The A strains used in the last 8 years of flu vaccines (USA) are as follows (no full length sequences are available on any of the these strains’ NP genes):

b. 2001-2002 A/Moscow/10/99, A/New Caledonia/20/99
d. 1999-2000 A/Sydney/05/97, A/Beijing/262/95
e. 1998-1999 A/Sydney/05/97, A/Beijing/262/95
g. 1996-1997 A/Nanchang/933/95, A/Texas/36/91
h. 1995-1996 A/Johannesburg/33/94, A/Texas/36/91

The final NP consensus amino acid sequence derived using this method is referred to herein as SEQ ID NO:7:6:}

lmaaqtqkrly eqnetdgerq natesravgy kmldigiqrfq igmoteiklna dyegrliqna
6ilternres fderrrrynle ehpsaqkdgp ktqgpiyrrv dgkwmervl ydkeesiriw
121rqannegedt eglthamiwv snldtttqyr tralvrtgmd pmoclnmgg tlprrseaqeg
180asvkgiltuq melirmlkrq indnfrvege nkgrtktasye ronnlkikgt qtaasqammd
241lyresrnpq snsediioa reallrlygv abkkeklpacv ypgpavsgyd fekeqysylw
301ldpfilgns qyelipinx meleeqiywv mphkqgivw mxeasafed lrlslfrf kgvrqykilkst
361rygaysann mndmgsatle lrryrsxirt regngntqgr asagqsvqg tfsqvrnlsf
421eketvmasf qntegrrsdn raelirrneq akepeesfry gvyfeledex atnpivpyfod
481mmegyfgy dnsceydn
M1 and M2 Consensus Amino Acid Sequences

Consensus sequences for M1 and M2 were determined in a similar fashion, as follows. The search parameters on the http://www.flu.lanl.gov/website were: influenza A strains, human, segment 7, nucleotide (both M1 and M2 are derived from segment 7). Full-length sequences from 1990-1999 (no 2000+ sequences were available) were chosen. For sequences with the same year and city, only the earliest passage was used. For entries for the same year, sequences were eliminated that were identical to another sequence isolated from the same year (even if different city). Twenty one sequences, full-length for both M1 and M2 from 1993-1999, were compared. At each position, the amino acid with the simple majority was used.

The M1 amino acid consensus sequence is referred to herein as SEQ ID NO:77:

[Example 4]
Codon Optimization Algorithm

The following is an outline of the algorithm used to derive human codon-optimized sequences of influenza antigens.

Back Translation

Starting with the amino acid sequence, one can either (a) manually backtranslate using the human codon usage table from http://www.kazusa.or.jp/codon/

Homo sapiens [gbp1r]: 55194 CDS’s (24298072 codons)

Fields: [triplet][frequency: per thousand] [number]

The M2 amino acid consensus sequence is referred to herein as SEQ ID NO:78:

[Example 5]

Consensus sequences for M1 and M2 were determined in a similar fashion, as follows. The search parameters on the http://www.flu.lanl.gov/website were: influenza A strains, human, segment 7, nucleotide (both M1 and M2 are derived from segment 7). Full-length sequences from 1990-1999 (no 2000+ sequences were available) were chosen. For sequences with the same year and city, only the earliest passage was used. For entries for the same year, sequences were eliminated that were identical to another sequence isolated from the same year (even if different city). Twenty one sequences, full-length for both M1 and M2 from 1993-1999, were compared. At each position, the amino acid with the simple majority was used.

The M1 amino acid consensus sequence is referred to herein as SEQ ID NO:77:

[Example 4]
Codon Optimization Algorithm

The following is an outline of the algorithm used to derive human codon-optimized sequences of influenza antigens.

Back Translation

Starting with the amino acid sequence, one can either (a) manually backtranslate using the human codon usage table from http://www.kazusa.or.jp/codon/

Homo sapiens [gbp1r]: 55194 CDS’s (24298072 codons)

Fields: [triplet][frequency: per thousand] [number]

The M2 amino acid consensus sequence is referred to herein as SEQ ID NO:78:

[Example 5]
* Coding GC 52.45% 1st letter GC 56.04% 2nd letter GC 42.37% 3rd letter GC 58.93% (Table as of Nov. 6, 2003)

[0322] Or (b) log on to www.synthetiegenes.com and use the backtranslation tool, as follows:

[0323] (1) Under Protein tab, paste amino acid sequence;
[0324] (2) Under download codon usage tab, highlight homo sapiens and then download CUT.

<table>
<thead>
<tr>
<th>Codon</th>
<th>Frequency</th>
<th>Codon</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUU</td>
<td>17.1 (415589)</td>
<td>UCU</td>
<td>20.6 (509964)</td>
</tr>
<tr>
<td>UCC</td>
<td>17.6 (427664)</td>
<td>UCA</td>
<td>15.4 (377811)</td>
</tr>
<tr>
<td>UGA</td>
<td>7.5 (1822466)</td>
<td>UAA</td>
<td>12.0 (291788)</td>
</tr>
<tr>
<td>UUG</td>
<td>12.6 (306793)</td>
<td>CGU</td>
<td>4.4 (107809)</td>
</tr>
<tr>
<td>CUC</td>
<td>13.0 (315804)</td>
<td>CCA</td>
<td>17.3 (419521)</td>
</tr>
<tr>
<td>CCG</td>
<td>19.8 (480790)</td>
<td>CAC</td>
<td>20.1 (489224)</td>
</tr>
<tr>
<td>CGA</td>
<td>7.8 (189383)</td>
<td>CAA</td>
<td>16.7 (405320)</td>
</tr>
<tr>
<td>CGG</td>
<td>39.8 (967277)</td>
<td>CGG</td>
<td>6.9 (168542)</td>
</tr>
<tr>
<td>AUU</td>
<td>16.1 (390571)</td>
<td>ACA</td>
<td>13.0 (315736)</td>
</tr>
<tr>
<td>AUC</td>
<td>21.6 (525248)</td>
<td>ACC</td>
<td>19.4 (471273)</td>
</tr>
<tr>
<td>AAG</td>
<td>7.7 (186138)</td>
<td>ACA</td>
<td>15.1 (366753)</td>
</tr>
<tr>
<td>AGG</td>
<td>22.2 (538917)</td>
<td>ACG</td>
<td>6.1 (148277)</td>
</tr>
<tr>
<td>GUU</td>
<td>11.0 (266493)</td>
<td>GCC</td>
<td>18.6 (451517)</td>
</tr>
<tr>
<td>GUC</td>
<td>14.6 (354537)</td>
<td>GCG</td>
<td>20.4 (690382)</td>
</tr>
<tr>
<td>GUG</td>
<td>7.2 (174572)</td>
<td>GCA</td>
<td>16.1 (390964)</td>
</tr>
<tr>
<td>CUG</td>
<td>28.4 (690428)</td>
<td>GGA</td>
<td>7.5 (101803)</td>
</tr>
</tbody>
</table>

(Table as of Nov. 6, 2003)

[0325] (3) Hit Apply button.
[0326] (4) Under Optimize TAB, open General TAB.
[0327] (5) Check use only most frequent codon box.
[0328] (6) Hit Apply button.
[0329] (7) Under Optimize TAB, open Motif TAB.
[0330] (8) Load desired cloning restriction sites into bad motifs; load any undesirable sequences, such as Pribnow Box sequences (TATAAA), Chi sequences (GCTGCGGG), and restriction sites into bad motifs.
[0331] (9) Under Output TAB, click on Start box. Output will include sequence, motif search results (under Report TAB), and codon usage report.

[0332] The program did not always use the most frequent codon for amino acids such as cysteine proline, and arginine. To change this, go back to the Edit CUT TAB and manually drag the rainbow colored bar to 100% for the desired codon. Then re-do start under the Output TAB.

[0333] The use of CGG for arginine can lead to very high GC content, so AGA can be used for arginine as an alternative. The difference in codon usage is 11.6 per thousand for CGG vs. 11.5 per thousand for AGA.

Splice Donor and Acceptor Site Search

[0335] (2) Check boxes for Human or other and both splice sites.
[0336] (3) Select minimum scores for 5' and 3' splice sites between 0 and 1.
[0337] Used the default setting at 0.4 where:

<table>
<thead>
<tr>
<th></th>
<th>% splice sites recognized</th>
<th>% false positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human 5' Splice sites</td>
<td>93.2%</td>
<td>5.2%</td>
</tr>
<tr>
<td>Human 3' Splice sites</td>
<td>83.8%</td>
<td>3.1%</td>
</tr>
</tbody>
</table>

[0338] Default minimum score is 0.4, where:

[0339] (4) Paste in sequence.
[0340] (5) Submit.
[0341] (6) Based on predicted donors or acceptors, change the individual codons until the sites are no longer predicted.
Add in 5' and 3' Sequences.

[0342] On the 5' end of the gene sequence, the restriction enzyme site and Kozak sequence (gccacc) was added before ATG. On 3' end of the sequence, Ica was added following the stop codon (tga on opposite strand) and then a restriction enzyme site. The GC content and Open Reading Frames were then checked in SEC Central.

Example 5

Preparation of Vaccine Formulations

[0343] Plasmid constructs comprising codon-optimized and non-codon-optimized coding regions encoding NP, M1, M2, HA, EM, and/or an eM2-EP fusion; or alternatively coding regions (either codon-optimized or non-codon optimized) encoding various IV proteins or fragments, variants or derivatives either alone or as fusions with a carrier protein, e.g., HBcAg, as well as various controls, e.g., empty vector, are formulated with the poloxamer CRL 1005 and BAK (Benzalkonium chloride 50% solution, available from Reger Chemical Co. Inc.) by the following methods. Specific final concentrations of each component of the formulae are described in the following methods, but for any of these methods, the concentrations of each component may be varied by basic stoichiometric calculations known by those of ordinary skill in the art to make a final solution having the desired concentrations.

[0344] For example, the concentration of CRL 1005 is adjusted depending on, for example, transfection efficiency, expression efficiency, or immunogenicity, to achieve a final concentration of between about 1 mg/ml to about 75 mg/ml, for example, about 1 mg/ml, about 2 mg/ml, about 3 mg/ml, about 4 mg/ml, about 5 mg/ml, about 6.5 mg/ml, about 7 mg/ml, about 7.5 mg/ml, about 8 mg/ml, about 9 mg/ml, about 10 mg/ml, about 15 mg/ml, about 20 mg/ml, about 25 mg/ml, about 30 mg/ml, about 35 mg/ml, about 40 mg/ml, about 45 mg/ml, about 50 mg/ml, about 55 mg/ml, about 60 mg/ml, about 65 mg/ml, about 70 mg/ml, or about 75 mg/ml of CRL 1005.

[0345] Similarly the concentration of DNA is adjusted depending on many factors, including the amount of a formulation to be delivered, the age and weight of the subject, the delivery method and route and the immunogenicity of the antigen being delivered. In general, formulations of the present invention are adjusted to have a final concentration from about 1 ng/ml to about 30 mg/ml of plasmid (or other polynucleotide). For example, a formulation of the present invention may have a final concentration of about 1 ng/ml, about 5 ng/ml, about 10 ng/ml, about 50 ng/ml, about 100 ng/ml, about 500 ng/ml, about 1 μg/ml, about 5 μg/ml, about 10 μg/ml, about 50 μg/ml, about 200 μg/ml, about 400 μg/ml, about 600 μg/ml, about 800 μg/ml, about 1 mg/ml, about 2 mg/ml, about 2.5, about 3 μg/ml, about 3.5, about 4 mg/ml, about 4.5, about 5 μg/ml, about 5.5 mg/ml, about 6 mg/ml, about 7 mg/ml, about 8 mg/ml, about 9 mg/ml, about 10 mg/ml, about 20 mg/ml, or about 30 mg/ml of a plasmid.

[0346] Certain formulations of the present invention include a cocktail of plasmids (see, e.g., Example 2 supra) of the present invention, e.g., comprising coding regions encoding IV proteins NP, M1 and/or M2 and optionally, plasmids encoding immunity enhancing proteins, e.g., cytokines. Various plasmids desired in a cocktail are combined together in PBS or other diluent prior to the addition to the other ingredients. Furthermore, plasmids may be present in a cocktail at equal proportions, or the ratios may be adjusted based on, for example, relative expression levels of the antigens or the relative immunogenicity of the encoded antigens. Thus, various plasmids in the cocktail may be present in equal proportion, or up to twice or three times as much of one plasmid may be included relative to other plasmids in the cocktail.

[0347] Additionally, the concentration of BAK may be adjusted depending on, for example, a desired particle size and improved stability. Indeed, in certain embodiments, formulations of the present invention include CRL 1005 and DNA, but are free of BAK. In general BAK-containing formulations of the present invention are adjusted to have a final concentration of BAK from about 0.05 mM to about 0.5 mM. For example, the present invention may have a final BAK concentration of about 0.05 mM, 0.1 mM, 0.2 mM, 0.3 mM, 0.4 mM or 0.5 mM.

[0348] The total volume of the formulations produced by the methods below may be scaled up or down, by choosing apparatus of proportional size. Finally, in carrying out any of the methods described below, the three components of the formulation, BAK, CRL 1005, and plasmid DNA, may be added in any order. In each of these methods described below the term “cloud point” refers to the point in a temperature shift, or other titration, at which a clear solution becomes cloudy, i.e., when a component dissolved in a solution begins to precipitate out of solution.

Thermal Cycling of a Pre-Mixed Formulation

[0349] This example describes the preparation of a formulation comprising 0.3 mM BAK, 7.5 mg/ml CRL 1005, and 5 mg/ml of DNA in a total volume of 3.6 ml. The ingredients are combined together at a temperature below the cloud point and then the formulation is thermally cycled to room temperature (above the cloud point) several times, according to the protocol outlined in FIG. 2.

[0350] A 1.28 mM solution of BAK is prepared in PBS, 846 μl of the solution is placed into a 15 ml round bottom flask fitted with a magnetic stirring bar, and the solution is stirred with moderate speed, in an ice bath on top of a stirrer/hotplate (hotplate off) for 10 minutes. CRL 1005 (27 μl) is then added using a 100 μl positive displacement pipette and the solution is stirred for a further 60 minutes on ice. Plasmids comprising codon-optimized coding regions encoding, for example, NP, M1, and M2 as described herein, and optionally, additional plasmids comprising codon-optimized or non-codon-optimized coding regions encoding, e.g., additional IV proteins, and or other proteins, e.g., cytokines, are mixed together at desired proportions in PBS to achieve 6.4 mg/ml total DNA. This plasmid cocktail is added drop wise, slowly, to the stirring solution over 1 min using a 5 ml pipette. The solution at this point (on ice) is clear since it is below the cloud point of the poloxamer and is further stirred on ice for 15 min. The ice bath is then removed, and the solution is stirred at ambient temperature for 15 minutes to produce a cloudy solution as the poloxamer passes through the cloud point.

[0351] The flask is then placed back into the ice bath and stirred for a further 15 minutes to produce a clear solution as
the mixture is cooled below the poloxamer cloud point. The ice bath is again removed and the solution stirred at ambient temperature for a further 15 minutes. Stirring for 15 minutes above and below the cloud point (total of 30 minutes), is defined as one thermal cycle. The mixture is cycled six more times. The resulting formulation may be used immediately, or may be placed in a glass vial, cooled below the cloud point, and frozen at -80 °C for use at a later time.

Thermal Cycling, Dilution and Filtration of a Pre-mixed Formulation, Using Increased Concentrations of CRL 1005

[0352] This example describes the preparation of a formulation comprising 0.3 mM BAK, 34 mg/ml or 50 mg/ml CRL 1005, and 5.0 mg/ml of DNA in a final volume of 4.0 ml. The ingredients are combined together at a temperature below the cloud point, then the formulation is thermally cycled to room temperature (above the cloud point) several times, diluted, and filtered according to the protocol outlined in FIG. 3.

[0353] Plasmids comprising codon-optimized coding regions encoding, for example, NP, M1, and M2 as described herein, and optionally, additional plasmids comprising codon-optimized or non-codon-optimized coding regions encoding, e.g., additional IV proteins, and or other proteins, e.g., cytokines, are mixed together at desired proportions in PBS to achieve 6.4 mg/ml total DNA. This plasmid cocktail is placed into the 15 ml round bottom flask fitted with a magnetic stirring bar, and for the formulation containing 50 mg/ml CRL 1005, 3.13 ml of a solution containing about 3.2 mg/ml of NP encoding plasmid and about 3.2 mg/ml M2 encoding plasmid (about 6.4 mg/ml total DNA) is placed into the 15 ml round bottom flask fitted with a magnetic stirring bar, and the solutions are stirred with moderate speed, in an ice bath on top of a stirrer/hotplate (hotplate off) for 10 minutes. CRL 1005 (136 μl for 34 mg/ml final concentration, and 200 μl for 50 mg/ml final concentration) is then added using a 200 μl positive displacement pipette and the solution is stirred for a further 30 minutes on ice. Solutions of 1.6 mM and 1.8 mM BAK are prepared in PBS, and 734 μl of 1.6 mM and 670 μl of 1.8 mM are then added drop wise, slowly, to the stirring poloxamer solutions with concentrations of 34 mg/ml or 50 mg/ml mixtures, respectively, over 1 min using a 1 ml pipette. The solutions at this point are clear since they are below the cloud point of the poloxamer and are stirred on ice for 30 min. The ice baths are then removed; the solutions stirred at ambient temperature for 15 minutes to produce cloudy solutions as the poloxamer passes through the cloud point.

[0354] The flasks are then placed back into the ice baths and stirred for a further 15 minutes to produce clear solutions as the mixtures cooled below the poloxamer cloud point. The ice baths are again removed and the solutions stirred for a further 15 minutes. Stirring for 15 minutes above and below the cloud point (total of 30 minutes), is defined as one thermal cycle. The mixtures are cycled two more times.

[0355] In the meantime, two Sterilip® 50 ml disposable vacuum filtration devices, each with a 0.22 μm Millipore Express® membrane (available from Millipore, cat # SCGP00525) are placed in an ice bucket, with a vacuum line attached and left for 1 hour to allow the device to equilibrate to the temperature of the ice. The poloxamer formulations are then diluted to 2.5 mg/ml DNA with PBS and filtered under vacuum.

[0356] The resulting formulations may be used immediately, or may be transferred to glass vials, cooled below the cloud point, and frozen at -80 °C for use at a later time.

A Simplified Method Without Thermal Cycling

[0357] This example describes a simplified preparation of a formulation comprising 0.3 mM BAK, 7.5 mg/ml CRL 1005, and 5 mg/ml of DNA in a total volume of 2.0 ml. The ingredients are combined together at a temperature below the cloud point and then the formulation is simply filtered and then used or stored, according to the protocol outlined in FIG. 4.

[0358] A 0.77 mM solution of BAK is prepared in PBS, and 780 μl of the solution is placed into a 15 ml round bottom flask fitted with a magnetic stirring bar, and the solution is stirred with moderate speed, in an ice bath on top of a stirrer/hotplate (hotplate off) for 15 minutes. CRL 1005 (15 μl) is then added using a 100 μl positive displacement pipette and the solution is stirred for a further 60 minutes on ice. Plasmids comprising codon-optimized coding regions encoding, for example, NP, M1, and M2 as described herein, and optionally, additional plasmids comprising codon-optimized or non-codon-optimized coding regions encoding, e.g., additional IV proteins, and or other proteins, e.g., cytokines, are mixed together at desired proportions in PBS to achieve a final concentration of about 8.3 mg/ml total DNA. This plasmid cocktail is added drop wise, slowly, to the stirring solution over 1 min using a 5 ml pipette. The solution at this point (on ice) is clear since it is below the cloud point of the poloxamer and is further stirred on ice for 15 min.

[0359] In the meantime, one Sterilip® 50 ml disposable vacuum filtration devices, with a 0.22 μm Millipore Express® membrane (available from Millipore, cat # SCGP00525) is placed in an ice bucket, with a vacuum line attached and left for 1 hour to allow the device to equilibrate to the temperature of the ice. The poloxamer formulation is then filtered under vacuum, below the cloud point and then allowed to warm above the cloud point. The resulting formulations may be used immediately, or may be transferred to glass vials, cooled below the cloud point and then frozen at -80 °C for use at a later time.

Example 6

Animal Immunizations

[0360] The immunogenicity of the various IV expression products encoded by the codon-optimized polynucleotides described herein are initially evaluated based on each plasmid’s ability to mount an immune response in vivo. Plasmids are tested individually and in combinations by injecting single constructs as well as multiple constructs. Immunizations are initially carried out in animals, such as mice, rabbits, goats, sheep, non-human primates, or other suitable animal, by intramuscular (IM) injections. Serum is collected from immunized animals, and the antigen specific antibody response is quantified by ELISA assay using purified immobilized antigen proteins in a protein—immunized subject antibody—anti-species antibody type assay, accord-
ing to standard protocols. The tests of immunogenicity further include measuring antibody titer, neutralizing antibody titer, T-cell proliferation, T-cell secretion of cytokines, cytolytic T cell responses, and by direct enumeration of antigen specific CD4+ and CD8+ T-cells. Correlation to protective levels of the immune responses in humans are made according to methods well known by those of ordinary skill in the art. See above.

A. DNA Formulations

[0361] Plasmid DNA is formulated with a poloxamer by any of the methods described in Example 3. Alternatively, plasmid DNA is prepared as described above and dissolved at a concentration of about 0.1 mg/ml to about 10 mg/ml, preferably about 1 mg/ml, in PBS with or without transfection-facilitating cationic lipids, e.g., DMRIE/DOPE at a 4:1 DNA:lipid mass ratio. Alternative DNA formulations include 150 mM sodium phosphate instead of PBS, adjuvants, e.g., Vaxfectin™ at a 4:1 DNA:Vaxfectin™ mass ratio, mono-phosphoryl lipid A (detoxified endotoxin) from S. minnesota (MPL) and tricholeaidorynomycolateAF (TDM), in 2% oil (squalene)-Tween 80-water (MPL+TDM, available from Sigma/Aldrich, St. Louis, Mo., catalog # M6536), a solubilized mono-phosphoryl lipid A formulation (AF, available from Corixa), or (a)-N-(3-Acetoxyspropyl)-N,N-dimethyl-2,3-bis(Octyloxyl)-1-propanaminium chloride (compound # VC1240) (see Shriver, J. W. et al., Nature 415:331-335 (2002), and PCT. Publication No. WO 02/00844 A2, each of which is incorporated herein by reference in its entirety).

B. Animal Immunizations

[0362] Plasmid constructs comprising codon-optimized and non-codon-optimized coding regions encoding NP, M1, M2, eM2, and/or an eM2-NP fusion; or alternatively coding regions (either codon-optimized or non-codon optimized) encoding various IV proteins or fragments, variants or derivatives either alone or as fusions with a carrier protein, e.g., HBCAg, as well as various controls, e.g., empty vector, are mixed together at desired proportions in PBS to achieve a final concentration of 1.0 mg/ml. The plasmid cocktail, as well as the controls, are formulated with Vaxfectin™. Groups of 5 BALB/c female mice are injected bilaterally in the rectus femoris muscle with 50 µl of DNA solution (100 µl total/mouse), on days 1 and 21 and 49 with each formulation. Mice are bled for serum on days 0 (prebleed), 20 (bleed 1), and 41 (bleed 2), and 62 (bleed 3), and up to 40 weeks post-injection. Antibody titers to the various IV proteins encoded by the plasmid DNAs are measured by ELISA as described elsewhere.

[0363] The immunization schedule is as follows:

<table>
<thead>
<tr>
<th>Day</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Pre-bleed</td>
</tr>
<tr>
<td>3</td>
<td>Plasmid injections, intramuscular, bilateral in rectus femoris, 5-50 µg/leg</td>
</tr>
<tr>
<td>21</td>
<td>Plasmid injections, intramuscular, bilateral in rectus femoris, 5-50 µg/leg</td>
</tr>
<tr>
<td>49</td>
<td>Plasmid injections, intramuscular, bilateral in rectus femoris, 5-50 µg/leg</td>
</tr>
<tr>
<td>59</td>
<td>Serum collection</td>
</tr>
</tbody>
</table>

[0364] Serum antibody titers are determined by ELISA with recombinant proteins, peptides or transfection supernatants and lysates from transfected VM-92 cells live, inactivated, or lysed virus.

C. Immunization of Mice with Vaccine Formulations Using a Vaxfectin™ Adjuvant

[0365] Vaxfectin™ (a 1:1 molar ratio of the cationic lipid VC1052 and the neutral co-lipid DpyPE) is a synthetic cationic lipid formulation which has shown promise for its ability to enhance antibody titers against when administered with DNA intramuscularly to mice.

[0366] In mice, intramuscular injection of Vaxfectin™ formulated with NP DNA increased antibody titers up to 20-fold to levels that could not be reached with DNA alone. In rabbits, complexing DNA with Vaxfectin™ enhanced antibody titers up to 50-fold. Thus, Vaxfectin™ shows promise as a delivery system and as an adjuvant in a DNA vaccine.

[0367] Vaxfectin™ mixtures are prepared by mixing chloroform solutions of VC1052 cationic lipid with chloroform solutions of DpyPE neutral co-lipid. Dried films are prepared in 2 ml sterile glass vials by evaporating the chloroform under a stream of nitrogen, and placing the vials under vacuum overnight to remove solvent traces. Each vial contains 1.5 µmole each of VC1052 and DpyPE. Liposomes are prepared by adding sterile water followed by vortexing. The resulting liposome solution is mixed with DNA at a phospho:mole:cationic lipid mole ratio of 4:1.

[0368] Plasmid constructs comprising codon-optimized and non-codon-optimized coding regions encoding NP, M1, M2, eM2, and/or an eM2-NP fusion; or alternatively coding regions (either codon-optimized or non-codon optimized) encoding various IV proteins or fragments, variants or derivatives either alone or as fusions with a carrier protein, e.g., HBCAg, as well as various controls, e.g., empty vector, are mixed together at desired proportions in PBS to achieve a final concentration of 1.0 mg/ml. The plasmid cocktail, as well as the controls, are formulated with Vaxfectin™. Groups of 5 BALB/c female mice are injected bilaterally in the rectus femoris muscle with 50 µl of DNA solution (100 µl total/mouse), on days 1 and 21 and 49 with each formulation. Mice are bled for serum on days 0 (prebleed), 20 (bleed 1), and 41 (bleed 2), and 62 (bleed 3), and up to 40 weeks post-injection. Antibody titers to the various IV proteins encoded by the plasmid DNAs are measured by ELISA as described elsewhere.

[0369] Cytolytic T-cell responses are measured as described in Hartikka et al. “Vaxfectin Enhances the Humoral Response to Plasmid DNA-encoded Antigens,” Vaccine 19:1911-1923 (2001) and is incorporated herein in its entirety by reference. Standard ELISPOT technology is used for the CD4+ and CD8+ T-cell assays as described in Example 6, part A.

D. Production of NP, M1 or M2 Antisera in Animals

[0370] Plasmid constructs comprising codon-optimized and non-codon-optimized coding regions encoding NP, M1, M2, eM2, and/or an eM2-NP fusion; or alternatively coding regions (either codon-optimized or non-codon optimized) encoding various IV proteins or fragments, variants or derivatives either alone or as fusions with a carrier protein, e.g., HBCAg, as well as various controls, e.g., empty vector, are prepared according to the immunization scheme described above and injected into a suitable animal for generating polyclonal antibodies. Serum is collected and the antibody titered as above.
Monoclonal antibodies are also produced using hybridoma technology (Kohler et al., *Nature* 256:495 (1975); Kohler et al., *Eur. J. Immunol.* 6:511 (1976); Kohler et al., *Eur. J. Immunol.* 6:292 (1976); Hammerling et al., in *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., (1981), pp. 563-681, one of which is incorporated herein by reference in its entirety). In general, such procedures involve immunizing an animal (preferably a mouse) as described above. The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP20), available from the American Type Culture Collection, Rockville, Md. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al., *Gastroenterology* 80:225-232 (1981), incorporated herein by reference in its entirety. The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the various IV proteins.

Alternatively, additional antibodies capable of binding to IV proteins described herein may be produced in a two-step procedure through the use of anti-idiotypic antibodies. Such a method makes use of the fact that antibodies to themselves are themselves antigens, and that, therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, various IV-specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the IV protein-specific antibody can be blocked by the cognate IV protein. Such antibodies comprise anti-idiotypic antibodies to the IV protein-specific antibody and can be used to immunize an animal to induce formation of further IV-specific antibodies.

It is appreciated that Fab and F(ab')₂ and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')₂ fragments). Alternatively, NP, M₁, M₂, HA, eM₂ and eM₂ binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.


These antibodies are used, for example, in diagnostic assays, as a research reagent, or to further immunize animals to generate IV-specific anti-idiotypic antibodies. Non-limiting examples of uses for anti-IV antibodies include use in Western blots, ELISA (competitive, sandwich, and direct), immunofluorescence, immunoelectron microscopy, radioimmunoassay, immunoprecipitation, agglutination assays, immunodiffusion, immunoelectrophoresis, and epitope mapping (Weir, D. Ed. *Handbook of Experimental Immunology*, 4th ed. Vols. I and II, Blackwell Scientific Publications (1986)).

Example 7

**Mucosal Vaccination and Electrically Assisted Plasmid Delivery**

A. Mucosal DNA Vaccination

Plasmid constructs comprising codon-optimized and non-codon-optimized coding regions encoding NP, M₁, M₂, HA, eM₂, and/or an eM₂-NP fusion; or alternatively coding regions (either codon-optimized or non-codon-optimized) encoding various IV proteins or fragments, variants or derivatives either alone or as fusions with a carrier protein, e.g., HBcAg, as well as various controls, e.g., empty vector, (100 μg/50 μL total DNA) are delivered to BALB/c mice at 0, 2 and 4 weeks via i.m., intranasal (i.n.), intravenous (i.v.), intravaginal (i.vag.), intrarectal (i.r.) or oral routes. The DNA is delivered unformulated or formulated with the cationic lipids DMRIE/DOPE (DD) or GAP-DLRIE/DOPE (GD). As endpoints, serum IgG titers against the various IV antigens are measured by ELISA and splenic T-cell responses are measured by antigen-specific production of IFN-gamma and IL-4 in ELISPOT assays. Standard chromatin release assays are used to measure specific cytotoxic T lymphocyte (CTL) activity against the various IV antigens. Tetramer assays are used to detect and quantify antigen specific T-cells, with quantification being confirmed and phenotypic characterization accomplished by intracellular cytokine staining. In addition, IgG and IgA responses against the various IV antigens are analyzed by ELISA of vaginal washes.

B. Electrically-Assisted Plasmid Delivery


The electroporation procedure can be performed with various electroporation devices. These devices include external plate type electrodes or invasive needle/rod electrodes and can possess two electrodes or multiple electrodes placed in an array. Distances between the plate or needle electrodes can vary depending upon the number of electrodes, size of target area and treatment subject.

The Trigrid needle array, used in examples described herein, is a three electrode array comprising three elongate electrodes in the approximate shape of a geometric
triangle. Needle arrays may include single, double, three, four, five, six or more needles arranged in various array formations. The electrodes are connected through conductive cables to a high voltage switching device that is connected to a power supply.

[0380] The electrode array is placed into the muscle tissue, around the site of nuclear acid injection, to a depth of approximately 3 mm to 3 cm. The depth of insertion varies depending upon the target tissue and size of patient receiving electroporation. After injection of foreign nuclear acid, such as plasmid DNA, and a period of time sufficient for distribution of the nuclear acid, square wave electrical pulses are applied to the tissue. The amplitude of each pulse ranges from about 100 volts to about 1500 volts, e.g., about 100 volts, about 200 volts, about 300 volts, about 400 volts, about 500 volts, about 600 volts, about 700 volts, about 800 volts, about 900 volts, about 1000 volts, about 1100 volts, about 1200 volts, about 1300 volts, about 1400 volts, or about 1500 volts or about 1-1.5 kV/cm, based on the spacing between electrodes. Each pulse has a duration of about 1 μs to 1000 μs, e.g., about 1 μs, about 10 μs, about 50 μs, about 100 μs, about 200 μs, about 300 μs, about 400 μs, about 500 μs, about 600 μs, about 700 μs, about 800 μs, about 900 μs, or about 1000 μs, and a pulse frequency on the order of about 1-10 Hz. The polarity of the pulses may be reversed during the electroporation procedure by switching the connectors to the pulse generator. Pulses are repeated multiple times. The electroporation parameters (e.g., voltage amplitude, duration of pulse, number of pulses, depth of electrode insertion and frequency) will vary based on target tissue type, number of needles used and distance of electrode spacing, as would be understood by one of ordinary skill in the art.

[0381] Immediately after completion of the pulse regimen, subjects receiving electroporation can be optionally treated with membrane stabilizing agents to prolong cell membrane permeability as a result of the electroporation. Examples of membrane stabilizing agents include, but are not limited to, steroids (e.g., dexamethasone, methylprednisone and progesterone), angiotensin II and vitamin E. A single dose of dexamethasone, approximately 0.1 mg per kilogram of body weight, should be sufficient to achieve a beneficial affect.

[0382] EAPD techniques such as electroporation can also be used for plasmids contained in liposome formulations. The liposome—plasmid suspension is administered to the animal or patient and the site of injection is treated with a safe but effective electrical field generated, for example, by a TriGrid needle array. The electroporation may aid in plasmid delivery to the cell by destabilizing the liposome bilayer so that membrane fusion between the liposome and the target cellular structure occurs. Electroporation may also aid in plasmid delivery to the cell by triggering the release of the plasmid, in high concentrations, from the liposome at the surface of the target cell so that the plasmid is driven across the cell membrane by a concentration gradient via the pores created in the cell membrane as a result of the electroporation.

[0383] Female BALB/c mice aged 8-10 weeks are anesthetized with inhalant isoflurane and maintained under anesthesia for the duration of the electroporation procedure. The legs are shaved prior to treatment. Plasmid constructs comprising codon-optimized and non-codon-optimized coding regions encoding NP, M1, M2, HA, eM2, and/or an eM2-NP fusion; or alternatively coding regions (either codon-optimized or non-codon optimized) encoding various IV proteins or fragments, variants or derivatives either alone or as fusions with a carrier protein, e.g., HBCAg, as well as various controls, e.g., empty vector, are administered to BALB/c mice (n=10) via unilateral injection in the quadriceps with 25 μg total of a plasmid-DNA per mouse, region 0.3 cc insulin syringe and a 26 gauge, ½ length needle fitted with a plastic collar to regulate injection depth. Approximately one minute after injection, electrodes are applied. Modified caliper electrodes are used to apply the electrical pulse. See Hartikka J. et al. Mol Ther 188:407-415 (2001). The caliper electrode plates are coated with conductivity gel and applied to the sides of the injected muscle before closing to a gap of 3 mm for administration of pulses. EAPD is applied using a square pulse type at 1-10 Hz with a field strength of 100-500 V/cm, 1-10 pulses, of 10-100 ms each.

[0384] Mice are vaccinated sEAPD at 0, 2 and 4 weeks. As endpoints, serum IgG titers against the various IV antigens are measured by ELISA and splenic T-cell responses are measured by antigen-specific production of IFN-gamma and IL-2 in ELISPOT assays. Standard chromium release assays are used to measure specific cytotoxic T lymphocyte (CTL) activity against the various IV antigens.

[0385] Rabbits (n=3) are given bilateral injections in the quadriceps muscle with plasmid constructs comprising codon-optimized and non-codon-optimized coding regions encoding NP, HA, M1, M2, eM2, and/or an eM2-NP fusion; or alternatively coding regions (either codon-optimized or non-codon optimized) encoding various IV proteins or fragments, variants or derivatives either alone or as fusions with a carrier protein, e.g., HBCAg, as well as various controls, e.g., empty vector. The implantation area is shaved and the TriGrid electrode array is implanted into the target region of the muscle. 3.0 mg of plasmid DNA is administered per dose through the injection port of the electrode array. An injection collet is used to control the depth of injection. Electroporation begins approximately one minute after injection of the plasmid DNA is complete. Electroporation is administered with a TriGrid needle array, with electrodes evenly spaced 7 mm apart, using an Ichor TGP-2 pulse generator. The array is inserted into the target muscle to a depth of about 1 to 2 cm. 4-8 pulses are administered. Each pulse has a duration of about 50-100 μs, an amplitude of about 1-1.2 kV/cm and a pulse frequency of 1 Hz. The injection and electroporation may be repeated.

[0386] Sera are collected from vaccinated rabbits at various time point. As endpoints, serum IgG titers against the various IV antigens are measured by ELISA and PBMC T-cell proliferative responses.

[0387] To test the effect of electroporation on therapeutic protein expression in non-human primates, male or female rhesus monkeys are given either 2 or 6 i.m. injections of plasmid constructs comprising codon-optimized and non-codon-optimized coding regions encoding NP, M1, M2, eM2, and/or an eM2-NP fusion; or alternatively coding regions (either codon-optimized or non-codon optimized) encoding various IV proteins or fragments, variants or derivatives either alone or as fusions with a carrier protein, e.g., HBCAg, as well as various controls, e.g., empty vector,
(0.1 to 10 mg DNA total per animal). Target muscle groups include, but are not limited to, bilateral rectus femoris, cranial tibialis, biceps, gastrocnemius or deltoid muscles. The target area is shaved and a needle array, comprising between 4 and 10 electrodes, spaced between 0.5 to 1.5 cm apart, is implanted into the target muscle. Once injections are complete, a sequence of brief electrical pulses are applied to the electrodes implanted in the target muscle using an Ichor TGP-2 pulse generator. The pulses have an amplitude of approximately 120 - 200V. The pulse sequence is completed within one second. During this time, the target muscle may make brief contractions or twitches. The injection and electroporation may be repeated.

Scra are collected from vaccinated monkeys at various time points. As endpoints, serum IgG titers against the various IV antigens are measured by ELISA and PBMC T-cell proliferative responses are measured by antigen-specific production of IFN-gamma and IL-4 in ELISPOT assays or by tetramer assays to detect and quantify antigen-specific T-cells, with quantification being confirmed and phenotypic characterization accomplished by intracellular cytokine staining. Standard chromatography release assays are used to measure specific cytoxic T lymphocyte (CTL) activity against the various IV antigens.

Example 8
Combintorial DNA Vaccine Using Heterologous Prime-Boost Vaccination

This Example describes vaccination with a combinatorial formulation including one or more polynucleotides comprising one codon-optimized coding regions encoding an IV protein or fragment, variant, or derivative thereof prepared with an adjuvant and/or transduction facilitating agent; and also an isolated IV protein or fragment, variant, or derivative thereof. Thus, antigen is provided in two forms. The exogenous isolated polynucleotides stimulate antigen specific antibody and CD4+ T-cell responses, while the polynucleotide-encoded protein, produced as a result of cellular uptake and expression of the coding region, stimulates a CD8+ T-cell response. Unlike conventional "prime-boost" vaccination strategies, this approach provides different forms of antigen in the same formulation. Because antigen expression from the DNA vaccine doesn’t peak until 7-10 days after injection, the DNA vaccine provides a boost for the protein component. Furthermore, the formulation takes advantage of the immunostimulatory properties of the bacterial plasmid DNA.

A. Non-Codon Optimized NP Gene

This example demonstrates the efficacy of this procedure using a non-codon-optimized polynucleotide encoding NP, however, the methods described herein are applicable to any IV polynucleotide vaccine formulation. Because only a small amount of protein is needed in this method, it is conceivable that the approach could be used to reduce the dose of conventional vaccines, thus increasing the availability of scarce or expensive vaccines. This feature would be particularly important for vaccines against pandemic influenza or biological warfare agents.

An injection dose of 10 μg influenza A/PR/8/34 nucleoprotein (NP) DNA per mouse, prepared essentially as described in Ulmer, J. B., et al., Science 259:1745-49 (1993) and Ulmer, J. B. et al., J. Virol. 72:5648-53 (1998) was pre-determined in dose response studies to induce T cell and antibody responses in the linear range of the dose response and results in a response rate of greater than 95% of mice injected. Each formulation, NP DNA alone, or NP DNAxNP protein formulated with Ribi I or the cationic lipids, DMRIE/DOPE or Vaxfectin™, was prepared in the recommended buffer for that vaccine modulation. For injections with NP DNA formulated with cationic lipid, the DNA was diluted in 2x PBS to 0.2 mg/ml purified recombinant NP protein (produced in baculovirus as described in Example 2) at 0.08 mg/ml. Each cationic lipid was reconstituted from a dried film by adding 1 ml of sterile water for injection (SWFI) to each vial and vortexing continuously for 2 min - then diluted with SWFI to a final concentration of 0.15 mM. Equal volumes of NP DNA (xNP protein) and cationic lipid were mixed to obtain a DNA to cationic lipid molar ratio of 4:1. For injections with DNA containing Ribi I adjuvant (Sigma), Ribi I was reconstituted with saline to twice the final concentration. Ribi I (2x) was mixed with an equal volume of NP DNA at 0.2 mg/ml in saline xNP protein at 0.08 mg/ml. For immunizations without cationic lipid or Ribi, NP DNA was prepared in 150 mM sodium phosphate buffer, pH 7.2. For each experiment, groups of 9 BALB/c female mice at 7-9 weeks of age were injected with 50 μl of NP DNAxNP protein, cationic lipid or Ribi I. Injections were given bilaterally in each rectus femoris on day 0 and day 21. The mice were bled by OAS on day 20 and day 33 and serum titers of individual mice were measured.

NP specific serum antibody titers were determined by indirect binding ELISA using 96 well ELISA plates coated overnight at 4°C with purified recombinant NP protein at 0.5 μg per well in PBS buffer pH 8.3. NP coated wells were blocked with 1% bovine serum albumin in PBS for 1 hour at room temperature. Two-fold serial dilutions of sera were incubated for 2 hours at room temperature and detected by incubating with alkaline phosphatase conjugated (AP) goat anti-mouse IgG-Fc (Jackson Immunoresearch, West Grove, Pa.) at 1:5000 for 2 hours at room temperature. The antibody was developed with 1 mg/ml paranitrophenyl phosphate (Calbiochem, La Jolla, Calif.) in 50 mM sodium bicarbonate buffer, pH 9.8 and 1 mM MgCl2, and the absorbance read at 405 nm. The titer is the reciprocal of the last dilution exhibiting an absorbance value 2 times that of pre-bled samples.

Standard ELISPOT technology, used to identify the number of interferon gamma (IFN-γ) secreting cells after stimulation with specific antigen (spot forming cells per million splenocytes, expressed as SFU/million), was used for the CD4+ and CD8+ T-cell assays. For the screening assays, 3 mice from each group were sacrificed on day 34, 35, and 36. At the time of collection, spleens from each group were pooled, and single cell suspensions made in cell culture media using a dounce homogenizer. Red blood cells were lysed, and cells washed and counted. For the CD4+ and CD8+ assays, cells were serially diluted 3-fold, starting at 10 cells per well and transferred to 96 well ELISPOT plates pre-coated with anti-murine IFN-γ monoclonal antibody. Spleen cells were stimulated with the H-2Kb binding peptide, TYQRTALV (SEQ ID NO:81), at 1 μg/ml and recombinant murine IL-2 at 1 U/ml for the CD8+ assay and with purified recombinant NP protein at 20 μg/ml for the CD4+ assay. Cells were stimulated for 24-24 hours at 37°C in 5% CO2, then the cells were washed out and biotin labeled.
anti-IFN-γ monoclonal antibody added for a 2 hour incubation at room temperature. Plates were washed and horseradish peroxidase-labeled avidin was added. After a 1-hour incubation at room temperature, AEC substrate was added and "spots" developed for 15 min. Spots were counted using the Immunospot automated spot counter (C.T.I., Inc., Cleveland, Ohio). Thus, CD4+ and CD8+ responses were measured in three separate assays, using spleens collected on each of three consecutive days.

Three weeks after a single injection, antibody responses in mice receiving vaccine formulations containing purified protein were 6 to 8-fold higher than for mice receiving NP DNA only (Fig. 5, Table 15). The titers for mice receiving DNA and protein formulated with a cationic lipid were similar to those for mice receiving protein in Ribi adjuvant or DNA and protein in Ribi adjuvant. These data indicate that the levels of antibody seen when protein is injected with an adjuvant can be obtained with DNA vaccines containing DNA and protein formulated with a cationic lipid, without the addition of conventional adjuvant.

Twelve days after a second injection, antibody responses in mice receiving vaccine formulations containing purified protein were 9 to 129-fold higher than for mice receiving NP DNA only (Fig. 6, Table 15). With a mean anti-NP antibody titer of 750,933 at day 33, the titers for mice receiving DNA and protein formulated with Vaxfectin™ were 25-fold higher than for mice receiving DNA alone (mean titer=30,578), and nearly as high as those for mice injected with protein in Ribi adjuvant (mean titer=1,748,133).

<table>
<thead>
<tr>
<th>Formulation</th>
<th>20 days after one injection</th>
<th>12 days after second injection</th>
</tr>
</thead>
<tbody>
<tr>
<td>protein + Ribi</td>
<td>7X (p = 0.0002)</td>
<td>57X (p = 0.002)</td>
</tr>
<tr>
<td>DNA + protein +</td>
<td>6X (p = 0.00005)</td>
<td>9X (p = 0.0002)</td>
</tr>
<tr>
<td>DMRIE-DOPA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA + protein +</td>
<td>8X (p = 0.00003)</td>
<td>25X (p = 0.0004)</td>
</tr>
<tr>
<td>Vaxfectin™</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA + protein + Ribi</td>
<td>7X (p = 0.01)</td>
<td>129X (p = 0.003)</td>
</tr>
</tbody>
</table>

*proteins = purified recombinant NP protein

As expected, an NP specific CD8+ T-cell IFN-γ response was not detected in spleens of mice injected with NP protein in Ribi (Fig. 7). All of the other groups had detectable NP specific CD8+ T-cell responses. The CD8+ T-cell responses for all groups receiving vaccine formulations containing NP DNA were not statistically different from each other.

Mice from all of the groups had detectable NP specific CD4+ T-cell responses (Fig. 8). The CD4+ T-cell responses of splenocytes from groups receiving vaccine formulations containing NP DNA and NP protein formulated with cationic lipid were 2-6 fold higher than the group injected with DNA alone.

B. Codon-Optimized IV Constructs

Plasmid constructs comprising codon-optimized and non-codon-optimized coding regions encoding NP, M1, M2, eM2, and/or an eM2-NP fusion, or alternatively coding regions (either codon-optimized or non-codon optimized) encoding various IV proteins or fragments, variants or derivatives either alone or as fusions with a carrier protein, e.g., HBCag, as well as various controls, e.g., empty vector, are used in the prime-boost compositions described herein. For the prime-boost modalities, the same protein may be used for the boost, e.g., DNA encoding NP with NP protein, or a heterologous boost may be used, e.g., DNA encoding NP with an M1 protein boost. Each formulation, the plasmid comprising a coding region for the IV protein alone, or the plasmid comprising a coding region for the IV protein plus the isolated protein are formulated with Ribi I or the cationic lipids, DMRIE-DOPA or Vaxfectin™. The formulations are prepared in the recommended buffer for that vaccine modality. Exemplary formulations, using NP as an example, are described herein. Other plasmid/protein formulations, including multivalent formulations, can be easily prepared by one of ordinary skill in the art by following this example. For injections with DNA formulated with cationic lipid, the DNA is diluted in 2x PBS to 0.2 mg/ml/purified recombinant NP protein at 0.08 mg/ml. Each cationic lipid is reconstituted from a dried film by adding 1 ml of sterile water for injection (SWFI) to each vial and vortexing continuously for 2 min., then diluted with SWFI to a final concentration of 0.15 mM. Equal volumes of NP DNA (±NP protein) and cationic lipid are mixed to obtain a DNA to cationic lipid molar ratio of 4:1. For injections with DNA containing Ribi I adjuvant (Sigma), Ribi 1 is reconstituted with saline to twice the final concentration. Ribi I (2x) is mixed with an equal volume of NP DNA at 0.2 mg/ml in saline±NP protein at 0.08 mg/ml. For immunizations without cationic lipid or Ribi, NP DNA is prepared in 150 mM sodium phosphate buffer, pH 7.2. For each experiment, groups of 9 BALB/c female mice at 7-9 weeks of age are injected with 50 µ of NP DNA±NP protein, cationic lipid or Ribi I. The formulations are administered to BALB/c mice (n=10) via bilateral injection in each rectus femoris at day 0 and day 21.

The mice are bled on day 20 and day 33 and serum titers of individual mice to the various IV antigens are measured. Serum antibody titers specific for the various IV antigens are determined by ELISA. Standard ELISPOT technology, used to identify the number of interferon gamma (IFN-γ) secreting cells after stimulation with specific antigen (spot forming cells per million splenocytes, expressed as SFU/million), is used for the CD4+ and CD8+ T-cell assays using 3 mice from each group vaccinated above, sacrificed on day 34, 35 and 36, post vaccination.

Example 9

Murine Challenge Model of Influenza

General Experimental Procedure

A murine challenge model with influenza A virus is used to test the efficacy of the immunotherapies. The model is based on that described in Ulmer, J. B., et al., Science 259:1745-49 (1993) and Ulmer, J. B. et al., J Virol. 72:5648-53 (1998), both of which are incorporated herein by reference in their entirety. This model utilizes a mouse-adapted strain of influenza A/HK8/68 which replicates in mouse lungs and is titrated in tissue culture in Madin Darby Canine Kidney cells. The LD₅₀ of this mouse-adapted influenza virus is determined in female BALB/c mice age 13-15
weeks. In this model, two types of challenge study can be conducted: lethal challenge, where the virus is administered intranasally to heavily sedated mice under ketamine anesthesia; and a sub-lethal challenge, where mice are not anesthetized when the viral inoculum is administered (also intranasally). The endpoint for lethal challenge is survival, but loss in body mass and body temperature can also be monitored. The read-outs for the sublethal challenge include lung virus titers and loss in body mass and body temperature.

[0401] In the studies described here, mice are subjected to lethal challenge. Mice that are previously vaccinated with DNA encoding IV antigens are anesthetized and challenged intranasally with 0.02 mL of mouse-adapted influenza A/HK/68 (mouse passage #6), diluted 1 to 10,000 (500 PFU) in PBS containing 0.2% w/v BSA.

[0402] These challenge studies utilize groups of 10 mice. The route of administration is intramuscular in rectus femoris (quadriiceps), using 0.1 μg up to 1 mg total plasmid DNA. Plasmid constructs comprising codon-optimized and non-codon-optimized coding regions encoding NP, M1, M2, eM2, and/or an eM2-NP fusion; or alternatively, coding regions (either codon-optimized or non-codon optimized) encoding various IV proteins or fragments, variants or derivatives either alone or as fusions with a carrier protein, e.g., HBCAg, as well as various controls, e.g., empty vector, are tested singly and in multivalent cocktails for the ability to protect against challenge. The plasmids are formulated with an adjuvant and/or a transfection facilitating agent, e.g., Vaxfectin™ by methods described elsewhere herein. Mice are vaccinated on days 0 and 21 using amounts of plasmids as described in Example 6. Subsequent injections can be administered. Nasal challenge of mice takes place 3 weeks after the final immunization, and animals are monitored daily for body mass, hypothermia, general appearance and then death.

[0403] For each group of mice that are studied, blood is taken at 2 weeks following the second injection, and/or any subsequent injection, and the animals are terminally bled two weeks following the last injection. Antibody titers are determined for M2, M1, and NP using ELISAs as previously described.

Plasmids

[0404] As described above, constructs of the previous invention were inserted into the expression vector VR10551. VR10551 is an expression vector without any transgene insert.

[0405] VR4750 contains the coding sequence for hemagglutinin (HA) (H3N2) from mouse adapted A/Hong Kong/68. The DNA was prepared using Qiagen plasmid purification kits.

Experimental Procedure

[0406] The experimental procedure for the following example is as described above, with particular parameters and materials employed as described herein. In order to provide a pDNA control for protection in the mouse influenza challenge model, the hemagglutinin (HA) gene was cloned from the influenza A/HK/68 challenge virus stock, which was passaged 6 times in mice.

[0407] Mice were vaccinated twice at 3 week intervals with either 100 μg pDNA VR4750 encoding the HA gene cloned directly from the mouse-adapted influenza A/HK/68 strain, or with 100 μg blank vector pDNA (VR10551). An additional control group was immunized intranasally with live A/HK/68/68 virus (500 PFU). Three weeks after the last injection, mice were challenged intranasally with mouse-adapted influenza A/HK/8168 with one of 3 doses (50, 500 and 5,000 PFU). Following viral challenge, mice were monitored daily for symptoms of disease, loss in body mass and body survival.

[0408] FIG. 9 shows that homologous HA-pDNA vaccinated mice are completely protected over a range of viral challenge doses (FIG. 9A) and did not suffer significant weight loss (FIG. 9B) during the 3 week period following challenge.

[0409] Based on these results, future mouse flu challenge studies can include VR4750 (HA) pDNA as a positive control for protection and utilize 500 PFU, which is the LD90 for this mouse-adapted virus, as the challenge dose.

Example 10
Challenge in Non-Human Primates

[0410] The purpose of these studies is to evaluate three or more of the optimal plasmid DNA vaccine formulations for immunogenicity in non-human primates. Rhesus or cynomolgous monkeys (6/group) are vaccinated with plasmid constructs comprising codon-optimized and non-codon-optimized coding regions encoding NP, HA, M1, M2, eM2, and/or an eM2-NP fusion; or alternatively, coding regions (either codon-optimized or non-codon optimized) encoding various IV proteins or fragments, variants or derivatives either alone or as fusions with a carrier protein, e.g., HBCAg, as well as various controls, e.g., empty vector, intramuscularly 0.1 to 2 mg DNA combined with cationic lipid, and/or poloxamer and/or aluminum phosphate based or other adjuvants at 0, 1 and 4 months.

[0411] Blood is drawn twice at baseline and then again at the time of and two weeks following each vaccination, and then again 4 months following the last vaccination. At 2 weeks post-vaccination, plasma is analyzed for humoral response and PBMCs are monitored for cellular responses, by standard methods described herein. Animals are monitored for 4 months following the final vaccination to determine the durability of the immune response.

[0412] Animals are challenged within 2-4 weeks following the final vaccination. Animals are challenged intratracheally with the suitable dose of virus based on preliminary challenge studies. Nasal swabs, pharyngeal swabs and lung lavages are collected at days 0, 2, 4, 6, 8 and 11 post-challenge and will be assayed for cell-free virus titers on monkey kidney cells. After challenge, animals are monitored for clinical symptoms, e.g., rectal temperature, body weight, leukocyte counts, and in addition, hemocrit and respiratory rate. Oropharyngeal swab samples are taken to allow determination of the length of viral shedding. Illness is scored using the system developed by Berndt & Hall (Infect Immun 16:476-479 (1977)), and will be analyzed by analysis of variance and the method of least significant difference.

Example 11
Challenge in Birds

[0413] In this example, various vaccine formulations of the present invention are tested in the chicken influenza
model. For these studies an IV H5N1 virus, known to infect birds, is used. Plasmid constructs comprising codon-optimized and non-codon-optimized coding regions encoding NP, M1, M2, eM2, and/or an eM2-NP fusion; or alternatively coding regions (either codon-optimized or non-codon optimized) encoding various IV proteins or fragments, variants or derivatives either alone or as fusions with a carrier protein, e.g., HBcAg, as well as various controls, e.g., empty vector, are formulated with cationic lipid, and/or poloxamer and/or aluminum phosphate based or other adjuvants. The vaccine formulations are delivered at a dose of about 1-10 μg, delivered IM into the defatted breast area, at 0 and 1 month. The animals are bled for antibody results 3 weeks following the second vaccine. Antibody titers against the various IV antigens are determined using techniques described in the literature. See, e.g., Kodihalli S. et al., Vaccine 18:2592-9 (2000). The birds are challenged intranasally with 0.1 mL containing 100 LD50, 3 weeks post second vaccination. The birds are monitored daily for 10 days for disease symptoms, which include loss of appetite, diarrhea, swollen faces, cyanosis, paralysis and death. Tracheal and cloacal swabs are taken 4 days following challenge for virus titration.

Example 12

Formulation Selection Studies

The potency of different vaccine formulations was evaluated in different experimental studies using the NP protein of Influenza A/PR/8134.

Vaccination Regimen

Groups of nine, six- to eight-week-old BALB/c mice (Harlan-Sprague-Dawley) received bilateral (50 μL/leg) intramuscular (rectus femoris) injections of plasmid DNA. Control mice received DNA in PBS alone. Mice received injections on days 0, 20 and 49. Mice were bled by OSP on day 62, and NP-specific antibodies analyzed by ELISA. Splenocytes were harvested from 3 mice/group/day for three sequential days beginning day 63, and NP-specific specific T cells were analyzed by IFNy ELISPOT using overlapping peptide stimulation.

Cell Culture Media

Splenocyte cultures were grown in RPMI-1640 medium containing 25 mM HEPES buffer and L-glutamine and supplemented with 10% (v/v) FBS, 50 μg/mL β-mercaptoethanol, 100 μM of penicillin G sodium salt, and 100 μg/mL of streptomycin sulfate.

Standard Influenza NP Indirect Binding Assay

NP specific serum antibody titers were determined by indirect binding ELISA using 96 well ELISA plates coated overnight at 4°C with purified recombinant NP protein at 0.5 μg per well in PBS buffer, pH 8.3. NP coated wells were blocked with 1% bovine serum albumin in PBS for 1 hour at room temperature. Two-fold serial dilutions of sera in blocking buffer were incubated for 2 hours at room temperature and detected by incubating with alkaline phosphatase conjugated (AP) goat anti-mouse IgG-Fc (Jackson Immunoresearch, West Grove, Pa.) at 1:5000 for 2 hours at room temperature. Color was developed with 1 mg/ml para-nitrophenyl phosphate (Calbiochem, La Jolla, Calif.) in 50 mM sodium bicarbonate buffer, pH 9.8 and 1 mM MgCl2, and the absorbance read at 405 nm. The titer is the reciprocal of the last dilution exhibiting an absorbance value 2 times that of pre-bleed samples.

Standard NP CD8+ and CD4+ T-Cell ELISPOT Assay

[0418] Standard ELISPOT technology, used to identify the number of interferon gamma (IFN-γ) secreting cells after stimulation with specific antigen (spot forming cells per million splenocytes, expressed as SFU/million), was used for the CD4+ and CD8+ T-cell assays. Three mice from each group were sacrificed on each of three consecutive days. At the time of collection, spleens from each group were pooled, and single cell suspensions were made in cell culture media using a dounce homogenizer. Red blood cells were lysed, and cells were washed and counted. For the CD4+ and CD8+ assays, cells were serially diluted 3- fold, starting at 106 cells per well and transferred to 96 well ELISPOT plates pre-coated with anti-murine IFN-γ monoclonal antibody. Spleen cells were stimulated with the H-2Kb binding peptide, IYQTRALVY at 1 μg/ml and recombinant murine IL-2 at 1 nM/ml for the CD8+ assay and with purified recombinant NP protein at 20 μg/ml for the CD4+ assay. Cells were stimulated for 20-24 hours at 37°C in 5% CO2, and then the cells were washed and biotin labeled anti-IFN-γ monoclonal antibody added for a 2 hour incubation at room temperature. Plates were washed and horseradish peroxidase-labeled avidin was added. After a 1 hour incubation at room temperature, AEC substrate was added and “spots” developed for 15 minutes. Spots were counted using the Immunospot automated spot counter (C.T.L. Inc., Cleveland Ohio).

Experiment 1

[0419] The purpose of this experiment was to determine a dose response to naked DNA (VR4700) and for pDNA formulated with VF-P1205-02A. VR4700 is a plasmid encoding influenza A/PR/8/34 nucleoprotein (NP) in a VR1055 backbone. VR1055 is an expression vector without any transgene insert. VF-P1205-02A is a formulation containing a poloxamer with a POP molecular weight of 12 KDa and POE of 5% (CRL1005) at a DNA:poloxamer:BAK ratio of 5 mg/ml:7.5 mg/ml:0.3 mM. The results of this experiment are shown in the following Table:

<table>
<thead>
<tr>
<th>TABLE 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA dose (μg)</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>100</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>100</td>
</tr>
</tbody>
</table>

[0420] The results of this experiment indicate that increasing the dose of DNA increases both the humoral and cell mediated immune responses. When the DNA is formulated with poloxamer and BAK, increasing the dose also increases both the humoral and cell mediated immune responses.

Experiment 2

[0421] The purpose of this experiment was to determine a dose response to CRL1005, with a fixed pDNA (VR4700)
dose and no BAK. The results of this experiment are shown in the following Table:

### TABLE 17

<table>
<thead>
<tr>
<th>DNA dose (μg)</th>
<th>CRL1005 dose (μg)</th>
<th>Serum Ab titters (total IgG, n = 9)</th>
<th>CD8⁺T cells (SFU/10⁶)</th>
<th>CD4⁺T cells (SFU/10⁶)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>15</td>
<td>27,733</td>
<td>45</td>
<td>46</td>
</tr>
<tr>
<td>10</td>
<td>50</td>
<td>38,460</td>
<td>69</td>
<td>86</td>
</tr>
<tr>
<td>10</td>
<td>150</td>
<td>46,533</td>
<td>68</td>
<td>77</td>
</tr>
<tr>
<td>10</td>
<td>450</td>
<td>54,044</td>
<td>90</td>
<td>97</td>
</tr>
<tr>
<td>10</td>
<td>750</td>
<td>76,800</td>
<td>90</td>
<td>92</td>
</tr>
</tbody>
</table>

*ND - could not be counted due to high background

[0422] The results of this experiment indicate that increasing the dose of CRL1005 increases both the humoral and cell mediated immune responses.

### Experiment 3

[0423] The purpose of this experiment was to compare immune responses of DMRIE:DOPE (1:1, mol:mol) and Vaxfectin™ cationic lipid formulations at different pDNA/cationic lipid molar ratios. The results of this experiment are shown in the following Table:

### TABLE 18

<table>
<thead>
<tr>
<th>DNA dose (μg)</th>
<th>DMRIE:DOPE pDNA/cationic lipid molar ratio</th>
<th>Vaxfectin™ pDNA/cationic lipid molar ratio</th>
<th>Serum Ab titters (total IgG, n = 9)</th>
<th>CD8⁺T cells (SFU/10⁶)</th>
<th>CD4⁺T cells (SFU/10⁶)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>4:1</td>
<td>17,778</td>
<td>57</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2:1</td>
<td>48,356</td>
<td>47</td>
<td>112</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>4:1</td>
<td>49,778</td>
<td>44</td>
<td>133</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2:1</td>
<td>88,178</td>
<td>68</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2:1</td>
<td>190,756</td>
<td>46</td>
<td>363</td>
<td></td>
</tr>
</tbody>
</table>

[0424] The results of this experiment indicate that formulating the plasmid with DMRIE:DOPE or Vaxfectin™ increases both the humoral and cell mediated immune responses.

### Experiment 4

[0425] The purpose of this experiment was first to compare immune responses of DMRIE:DOPE (1:1, mol:mol) at pDNA:cationic lipid molar ratios of 4:1 as an MLV (multi lamellar vesicle formulation—multi-vial) or SUV (small unilamellar vesicles—single-vial) formulation. Second, it was to compare sucrose (lyophilized and frozen) and PBS based formulations. The results of this experiment are shown in the following Table:

### TABLE 19

<table>
<thead>
<tr>
<th>DNA dose (μg)</th>
<th>Formulation</th>
<th>Serum Ab titters (total IgG, n = 9)</th>
<th>CD8⁺T cells (SFU/10⁶)</th>
<th>CD4⁺T cells (SFU/10⁶)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>PBS, pH 7.2</td>
<td>21,333</td>
<td>107</td>
<td>138</td>
</tr>
<tr>
<td>10</td>
<td>PBS, pH 7.2</td>
<td>15,644</td>
<td>144</td>
<td>189</td>
</tr>
<tr>
<td>10</td>
<td>PBS, pH 7.2</td>
<td>13,513</td>
<td>114</td>
<td>173</td>
</tr>
</tbody>
</table>

[0428] The results of this experiment indicate that formulating the plasmid with DMRIE:DOPE stimulates both the humoral and cell mediated immune responses. Changing the co-lipid from DOPE to cholesterol also stimulates both the humoral and cell mediated immune responses.

### Experiment 5

[0429] The purpose of this experiment was to obtain a dose response to pDNA formulated with DMRIE:DOPE (1:1, mol:mol) at a 4:1 pDNA:cationic lipid molar ratio. The results of this experiment are shown in the following Table:

### TABLE 20

<table>
<thead>
<tr>
<th>DNA dose (μg)</th>
<th>Formulation</th>
<th>Serum Ab titters (total IgG, n = 9)</th>
<th>CD8⁺T cells (SFU/10⁶)</th>
<th>CD4⁺T cells (SFU/10⁶)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>MLV, DMEDP</td>
<td>10,342</td>
<td>65</td>
<td>98</td>
</tr>
<tr>
<td>10</td>
<td>MLV, DMEDP</td>
<td>38,684</td>
<td>70</td>
<td>126</td>
</tr>
<tr>
<td>10</td>
<td>MLV, DMEDP</td>
<td>75,093</td>
<td>82</td>
<td>182</td>
</tr>
<tr>
<td>10</td>
<td>MLV, DMEDP</td>
<td>53,476</td>
<td>88</td>
<td>186</td>
</tr>
<tr>
<td>10</td>
<td>MLV, DMEDP</td>
<td>36,409</td>
<td>93</td>
<td>106</td>
</tr>
<tr>
<td>10</td>
<td>MLV, DMChol</td>
<td>52,338</td>
<td>68</td>
<td>154</td>
</tr>
</tbody>
</table>

[0427] The purpose of this experiment was first to determine what effect changing the ratio of DMRIE to DOPE has on immune response at pDNA/cationic lipid molar ratios of 4:1 as an MLV (multi-vial, in PBS) or SUV (single-vial in PBS) formulation. Second, it was to compare the effect of changing the co-lipid from DOPE to cholesterol. The results of this experiment are shown in the following Table:

### TABLE 21

<table>
<thead>
<tr>
<th>DNA dose (μg)</th>
<th>Formulation</th>
<th>Serum Ab titters (total IgG, n = 9)</th>
<th>CD8⁺T cells (SFU/10⁶)</th>
<th>CD4⁺T cells (SFU/10⁶)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>MLV</td>
<td>22,044</td>
<td>139</td>
<td>160</td>
</tr>
<tr>
<td>10</td>
<td>MLV</td>
<td>22,756</td>
<td>46</td>
<td>97</td>
</tr>
<tr>
<td>10</td>
<td>MLV</td>
<td>45,511</td>
<td>159</td>
<td>250</td>
</tr>
</tbody>
</table>
TABLE 21-continued

<table>
<thead>
<tr>
<th>DNA dose (µg)</th>
<th>Serum Ab-titer (total IgG, n = 9)</th>
<th>CD8+ T cells (SUF/10^6)</th>
<th>CD4+ T cells (SUF/10^6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>MLV</td>
<td>60,444</td>
<td>274</td>
</tr>
<tr>
<td>100</td>
<td>MLV</td>
<td>91,022</td>
<td>277</td>
</tr>
</tbody>
</table>

[0340] The results of this experiment indicate that when the plasmid is formulated with DMRIE:DOPE, increasing the dose also increases both the humoral and cell mediated immune responses.

Example 13

In vitro Expression of Influenza Antigens

Plasmid Vector

[0341] Polyornucleotides of the present invention were inserted into eukaryotic expression vector backbones VR10551, VR10682 and VR6430 all of which are described previously. The VR10551 vector is built on a modified pUC18 background (see Yanisch-Perron, C., et al. Gene 33:103-119 (1985)), and contains a kanamycin resistance gene, the human cytomegalovirus immediate early 1 promoter/enhancer and intron A, and the bovine growth hormone transcription termination signal, and a polylinker for inserting foreign genes. See Haurika, J., et al., Hum. Gene Ther. 7:1205-1217 (1996). However, other standard commercially available eukaryotic expression vectors may be used in the present invention, including, but not limited to: plasmids pcDNA3, plHCMV/Zeo, pCRII/His, pNBD/ GS, pRe/HCMV2, pSV40/Zeo2, pTRACER/HCMV, pUB6/ V5-His, pVAX1, and pZeoS/V2 (available from Invitrogen, San Diego, Calif.), and plasmid pCT (available from Promega, Madison, Wis.).

[0342] Various plasmids were generated by cloning the nucleotide sequence for the following influenza A antigens: segment 7 (encodes both M1 and M2 proteins via differential splicing), M2 and NP into expression constructions as described below and pictured in FIG. 13.

[0343] Plasmids VR4756 (SEQ ID NO:91), VR4759 (SEQ ID NO:92) and VR4762 (SEQ ID NO:93) were created by cloning the nucleotide sequence encoding the consensus sequence for the following influenza A antigens respectively: segment 7 (encoding both M1 and M2 proteins by differential splicing), M2 and NP into the VR10551 backbone. The VR4756, VR4759 and VR4762 plasmids are also described in Table 13.

[0344] The VR4764 (SEQ ID NO:95) and VR4765 (SEQ ID NO:96) plasmids were constructed by bridging the segment 7 and NP coding regions from VR4756 and VR4762 respectively into the VR10682 vector. Specifically, the VR4750 vector was digested with EcoRV and Sall restriction endonucleases and the blunt fragment was ligated into the VR10682 backbone, which had been digested with the EcoRV restriction enzyme. The VR4765 vector was constructed by digesting the VR4762 vector with EcoRV and Nol and ligating the NP coding region into the VR10682 backbone digested with the same restriction endonucleases.

[0345] VR4766 (SEQ ID NO:97) and VR4767 (SEQ ID NO:98) contain a CMV promoter/intron A-NP expression cassette and a RSV promoter (from VCL1005) segment 7 expression cassette in the same orientation (VR4766) or opposite orientation (VR4767). These plasmids were generated by digesting VR4762 with the DraIII restriction endonuclease and cutting the RSV-segment 7-mRBG cassette from VR4764 with EcoRV and BamHI restriction endonucleases. After exonuclease digestion with the Klenow fragment of DNA polymerase I, the EcoRV/BamHI fragment was cloned into the DraIII digested VR4762 vector. Both insert orientations were obtained by this blunt end cloning method.

[0346] VR4768 (SEQ ID NO:99) and VR4769 (SEQ ID NO:100), containing a CMV promoter/intron A-segment 7 expression cassette and a RSV promoter-NP expression cassette, were similarly derived. VR4756 was digested with the DraIII restriction endonuclease and blunted by treatment with the Klenow fragment of DNA Polymerase I. The cassette containing the RSV promoter, NP coding region and mRBG terminator was removed from VR4765 by digesting with KpnI and Ndel restriction endonucleases. The fragment was also blunt-ended with the Klenow fragment of DNA polymerase I and ligated into the DraIII-digested VR4762 vector in both gene orientations.

[0347] VR4770 (SEQ ID NO:101), VR4771 (SEQ ID NO:102) and VR4772 (SEQ ID NO:103) were constructed by cloning the coding regions from VR4756, VR4762 and VR4759 respectively into the VR6430 vector backbone. Specifically, the segment 7 gene from VR4756 was removed using SacI and EcoRV restriction endonucleases and blunt-ended with the Klenow fragment of DNA polymerase I. The VR6430 plasmid was digested with EcoRV and BamHI and the vector backbone fragment was blunt-ended with the Klenow fragment of DNA polymerase I. The segment 7 gene fragment was then ligated into the VR6430 vector backbone. VR4771 was derived by removing the NP insert from VR4762 following EcoRV and BglII restriction endonuclease digestion and the fragment was ligated into the VR6430 vector backbone which had been digested the same restriction endonucleases. VR4772 was derived by subcloning the M2 coding region from VR4759 as a blunt-ended SacI-EcoRV fragment and ligating into the VR6430 vector backbone from a blunt-ended EcoRV-BamHI digest.

[0348] VR4773 (SEQ ID NO:104) and VR4774 (SEQ ID NO:105) contain a CMV promoter/intron A-segment 7 expression cassette and a RSV/V-NP expression cassette with the genes in the same or opposite orientation. These plasmids were generated by digesting VR4756 with the DraIII restriction endonuclease, blunting, and ligating to the RSV/V-NP-BGH fragment from VR4771 (VR4771 digested with Ndel and SfiI and then blunt-ended).

[0349] VR4775 (SEQ ID NO:106) and VR4776 (SEQ ID NO:107) contain a CMV promoter/intron A-NP expression cassette and a RSV/V-segment 7 expression cassette with the genes in the same or opposite orientations. These plasmids were generated by digesting VR4762 with the DraIII restriction enzyme and blunting with the Klenow fragment of DNA polymerase I. The RSV/V-segment 7-BGH fragment was generated by digesting VR4770 with Ndel and SfiI restriction endonucleases and ligating the blunt fragment with the DraIII restriction endonuclease digested VR4762.
VR4777 (SEQ ID NO:108) and VR4778 (SEQ ID NO:109) contain a CMV promoter/intron A-NP expression cassette and a RSV/R-M2 expression cassette in the same or opposite orientation. These plasmids were generated by digesting VR4762 with the MscI restriction endonuclease, digesting VR4772 with NdeI and SfiI restriction endonucleases and treating the RSV/R-M2-BGH with the Klenow fragment of DNA polymerase, followed by ligation of these two gel purified fragments.

VR4779 and VR4780 contain a CMV promoter/intron A-M2 expression cassette and a RSV/R-NP expression cassette in the same or opposite orientation. These plasmids were generated by digesting VR4759 with the MscI restriction endonuclease, digesting VR4771 with NdeI and SfiI restriction endonucleases and treating the RSV/R-NP-BGH segment with the Klenow fragment of DNA polymerase, followed by ligation of these two gel purified fragments.

Plasmid DNA Purification

Plasmid DNA was transformed into Escherichia coli DH5α competent cells, and highly purified covalently closed circular plasmid DNA was isolated by a modified lysis procedure (Horn, N. A., et al., Hum. Gene Ther. 6:565-573 (1995)) followed by standard double CsCl-ethidium bromide gradient ultracentrifugation (Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y. (1989)). All plasmid preparations were free of detectable chromosomal DNA, RNA and protein impurities based on gel analysis and the biocinchonic acid protein assay (Pierce Chem. Co., Rockford III.). Endotoxin levels were measured using Limulus Amebocyte Lysate assay (LAI, Associates of Cape Cod, Falmouth, Mass.) and were less than 0.6 Endotoxin Unit/mg of plasmid DNA. The spectrophotometric A260/A280 ratios of the DNA solutions were typically above 1.8. Plasmids were ethanol precipitated and resuspended in an appropriate solution, e.g., 150 mM sodium phosphate (for other appropriate excipients and auxiliary agents, see U.S. patent application Publication 2002/0019358, published Feb. 14, 2002). DNA was stored at -20°C until use. DNA was diluted by mixing it with 300 mM salt solutions and by adding appropriate amount of USP water to obtain 1 mg/ml plasmid DNA in the desired salt at the desired molar concentration.

Plasmid Expression in Mammalian Cell Lines

The expression plasmids were analyzed in vitro by transfecting the plasmids into a well characterized mouse melanoma cell line (VM-92, also known as UM-449) and the human rhabdomyosarcoma cell line RD (ATCC CCL-136) both available from the American Type Culture Collection, Manassas, Va. Other well-characterized human cell lines may also be used, e.g., MRC-5 cells, ATCC Accession No. CCL-171. The transfection was performed using cationic lipid-based transfection procedures well known to those of skill in the art. Other transfection procedures are well known in the art and may be used, for example electroporation and calcium chloride-mediated transfection (Graham F. L. and A. J. van der Eb Virology 52:456-67 (1973)). Following transfection, cell lysates and culture supernatants of transfected cells were evaluated to compare relative levels of expression of 1V antigen proteins. The samples were assayed by Western blots and ELISAs, using commercially available monoclonal antibodies (available, e.g., from Research Diagnostics Inc., Flanders, N.J.), so as to compare both the quality and the quantity of expressed antigen.

Genes encoding the consensus amino acid sequences (described above) derived for NP, M1 and M2 antigens were cloned in several configurations into several plasmid vector backbones. The pDNAs were tested for in vitro expression and are being assessed in vivo for immunogenicity, as well as for the ability to protect mice from influenza challenge.

Experiment 1

Following the derivation of an amino acid consensus for M1 and M2, a native segment 7 isolate was found to encode this consensus, and this nucleotide sequence was synthesized according to methods described above. An M2-M1 fusion gene was also created and the nucleotide sequence was human codon-optimized using the above described codon optimization algorithm of Example 4. The individual full-length M2 and M1 genes were also cloned via PCR from this fusion.

In vitro expression of influenza antigens in cell lysates was assessed 48 hours after transfection into a mouse melanoma cell line. M2 expression was detected following transfection of VR4756 (segment 7), VR4755 (M2-M1 fusion) and VR4759 (full-length M2) using the anti-M2 monoclonal antibody (14C2) from Affinity BioReagents. The data are shown in FIG. 10 for VR4756 and VR4755. Expression of M1 was detected from transfected VR4756, VR4755 and VR760 (full-length M1) pDNAs, as detected by anti-M1 monoclonal (Serotec) in FIG. 10 for VR4756 and VR7555, or by anti-M1 goat polyclonal (Virostat, data not shown). VR1055 is the empty cloning vector.

Experiment 2

In order to compare alternative human codon-optimization methods, two versions of a fusion of the first 24 amino acids of M2 to full-length NP ("eM2-NP") were constructed. One nucleotide sequence was derived from the above codon optimization algorithm, while the other was done by an outside vendor. Comparison of expression levels from the two eM2-NP pDNAs was measured in vitro, and comparison of immunogenicity in vivo is on-going. Additionally, the full-length NP genes for both codon-optimized versions were sub-cloned from the eM2-NP pDNAs and analyzed for expression in vitro.

In vitro expression was tested to compare eM2-NP and NP pDNAs derived from the above described codon-optimization algorithm and an outside vendor algorithm. The data are shown in FIG. 11. Expression levels were approximately the same for VR4757 (eM2-NP vendor optimization) vs. VR4758 (eM2-NP Applicant optimization), as detected by anti-M2 monoclonal (FIG. 11A) or anti-NP mouse polyclonal (data not shown). Similarly, NP expression was approximately equal for VR7601 (vendor optimization) vs. VR7602 (Applicant optimization), detected by anti-NP mouse polyclonal generated by Applicants (FIG. 11B). NP consensus protein expression in vitro was also detected using a goat polyclonal antibody (Fitzgerald) generated against whole H1N1 or H3N2 virus (data not shown). Expression levels of both of these NP constructs were much higher than a pDNA containing A/PR/34-NP (VR4700).
Experiment 3

Influenza antigen-encoding plasmids were transfected into VM92 cells using methods described above. Cell lysates and media were collected 48 hours after transfection. Cells were lysed in 200 μl of Laemmli buffer, cell debris removed by microcentrifuge spin, and 20 μl was heated and loaded on a 4-12% Bis-Tris gel. To determine expression of those vectors encoding secreted NP protein, 15 μl of media was mixed with 5 μl of loading buffer, heated, and loaded on a gel. Western blots were processed as described above. Primary antibodies were as follows: monoclonal antibody MA1-062 (ABR) to detect M2 protein, monoclonal antibody MCA401 (Seropec) to detect M1 protein, and a polyclonal antibody against VR4762-infected rabbits generated in-house. All primary antibodies were used at a 1:500 dilution.

Fig. 14 shows Western blot results wherein M2 protein expression from segment 7-encoded plasmids are higher in CMV promoter/intron A-segment 7 (VR4756) and RSV/R-segment 7 (VR4770) than VR4764 (RSV promoter). NP expression appeared highest from the RSV/R-NP plasmid (VR4771), followed by CMV/intron A-NP (VR4762) and then RSV-NP (VR4765). Similar results were seen in Western blots from human RD-transfected cells.

For dual promoter plasmids, containing RSV-segment 7 and CMV/intron A-NP (VR4766 and VR4767), M2 expression from segment 7 is very low, independent of orientation. The CMV/intron A-NP expression in these dual promoter plasmids does not differ significantly compared to VR4762. RSV-NP expression in dual promoter plasmids (VR4768 and VR4769), where segment 7 is expressed from CMV/intron A, NP expression decreases somewhat, but not as drastically as M2 expression in the dual promoter VR4766 and VR4767.

Fig. 15 shows expression of the M1 and M2 proteins from segment 7, as well as NP, from CMV promoter/intron A, RSV promoter, and RSV/R-containing plasmids. For these Western blots, dual promoter plasmids contain the CMV promoter/intron A and RSV/R driving either NP or segment 7. Similar results were seen in Western blots from human RD-transfected cells.

Western blot results confirm that the M1 and M2 protein expression from both CMV promoter/intron A-segment 7 (VR4756) and RSV/R-segment 7 (VR4770) is superior to RSV-segment 7 (VR4764). M1 and M2 expression decrease slightly when RSV/R-segment 7 or CMV/intron A-segment 7 is combined with CMV/intron A-NP or RSV/R-NP in a dual promoter plasmid (VR4773, VR4774, VR4775, and VR4776). Results were similar in Western blots from human RD-transfected cells. Human RD cells transfected with M2 antigen encoding plasmids, RSV/R-M2 (VR4772) and CMV/intron A-M2 (VR4759), showed a similar level of M2 expression, which was decreased in dual promoter plasmids (VR4777, VR4778, VR4779, and VR4780). Human RD cells transfected with NP antigen-encoding plasmids, VR4771, VR4777, VR4778, VR4779, and VR4780, all showed similar NP expression levels.

Example 14

Murine Influenza a Challenge Model

A model influenza A challenge model has been established using a mouse-adapted A/BK/8/68 strain. Positive and negative control Hemaglutinin (HA)-containing plasmids were generated by PCR of the HA genes directly from mouse-adapted A/Hong Kong/68 (H3N2) and A/Puerto Rico/34 (H1N1) viruses, respectively.

For all experiments, plasmid DNA vaccinations are given as bilateral, rectus femoris injections at 0 and 3 weeks, followed by orbital sinus puncture (OSP) bleed at 5 weeks and intranasal viral challenge at 6 weeks with 500 pfu (1 LD50) of virus. Mice are monitored for morbidity and weight loss for about 3 weeks following viral challenge. Endpoint antibody titers for NP and M2 were determined by ELISA. For study JS80, 5 additional mice per test group were vaccinated and interferon-γ ELISPOT assays were performed at week number 5.

Study CL88:

A mouse influenza challenge study was initiated to test the M1, M2, Segment 7, and NP-encoding plasmids alone, or in combination. In addition to HA pDNAs, sub-lethal infection and naïve mice serve as additional positive and negative controls, respectively. Mice received 100 μg of each plasmid formulated in poloxamer CRL1005, 02A formulation. The test groups and 21 day post-challenge survival are shown in Table 21:

<table>
<thead>
<tr>
<th>Group</th>
<th>Construct(s)</th>
<th>Total pDNA per vaccination</th>
<th># mice/group</th>
<th>21 day Survival (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>VR4762 (NP)</td>
<td>100 μg</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>B</td>
<td>VR4759 (M2)</td>
<td>100 μg</td>
<td>12</td>
<td>25</td>
</tr>
<tr>
<td>C</td>
<td>VR4760 (M1)</td>
<td>100 μg</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>VR4756 (S7)</td>
<td>100 μg</td>
<td>12</td>
<td>50</td>
</tr>
<tr>
<td>E</td>
<td>VR4762 (NP)</td>
<td>200 μg</td>
<td>12</td>
<td>100</td>
</tr>
<tr>
<td>F</td>
<td>VR4759 (M2)</td>
<td>100 μg</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>G</td>
<td>VR4762 (NP)</td>
<td>100 μg</td>
<td>12</td>
<td>75</td>
</tr>
<tr>
<td>H</td>
<td>VR4750 (HA, HIN2, + control)</td>
<td>100 μg</td>
<td>12</td>
<td>100</td>
</tr>
<tr>
<td>I</td>
<td>VR4752 (HA, H1N1, - control)</td>
<td>100 μg</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>J</td>
<td>Naïve mice (- control)</td>
<td>NA</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>K</td>
<td>Sub-lethal (+ control)</td>
<td>NA</td>
<td>12</td>
<td>100</td>
</tr>
</tbody>
</table>

CL88 Results:

The performance criteria for this study was survival of >90% for the positive controls, ≤10% for the negative controls, and >75% for the experimental groups. Table 21 shows that all of the control groups, as well as two experimental groups met the performance criteria. The M2+NP and S7+NP plasmid DNA combinations resulted in 100% and 75% survival, respectively. There was no statistically significant difference (p<0.05) between the two lead plasmid combinations, but there was statistical significance in the S7, S7+NP, and M2+NP groups vs. the negative controls.

Weight loss data showed that the positive control groups did not exhibit any weight loss following viral challenge, as opposed to the weight loss seen in all of the experimental groups. Mice that survived the viral challenge recovered to their starting weight by the end of the study.

Tables 22 and 23 show endpoint antibody titers for test
groups containing M2, Segment 7, and NP antigens. Shaded boxes represent mice that died following viral challenge.

### TABLE 22

<table>
<thead>
<tr>
<th>mouse</th>
<th>Group D (seg 7)</th>
<th>Group G (NP + seg?)</th>
<th>Group B (M2)</th>
<th>Group E (NP + M2)</th>
<th>Group F (NP + M2)</th>
<th>Group G (NP + seg?)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>800</td>
<td>1600</td>
<td>2500</td>
<td>1600</td>
<td>6400</td>
<td>3200</td>
</tr>
<tr>
<td>2</td>
<td>3200</td>
<td>6400</td>
<td>800</td>
<td>12800</td>
<td>12800</td>
<td>12800</td>
</tr>
<tr>
<td>3</td>
<td>6400</td>
<td>12800</td>
<td>0</td>
<td>3200</td>
<td>3200</td>
<td>3200</td>
</tr>
<tr>
<td>4</td>
<td>12800</td>
<td>0</td>
<td>800</td>
<td>12800</td>
<td>12800</td>
<td>12800</td>
</tr>
<tr>
<td>5</td>
<td>800</td>
<td>12800</td>
<td>12800</td>
<td>3200</td>
<td>3200</td>
<td>3200</td>
</tr>
<tr>
<td>6</td>
<td>12800</td>
<td>0</td>
<td>3200</td>
<td>6400</td>
<td>6400</td>
<td>6400</td>
</tr>
<tr>
<td>7</td>
<td>800</td>
<td>3200</td>
<td>400</td>
<td>800</td>
<td>800</td>
<td>800</td>
</tr>
<tr>
<td>8</td>
<td>12800</td>
<td>3200</td>
<td>6400</td>
<td>3200</td>
<td>3200</td>
<td>3200</td>
</tr>
<tr>
<td>9</td>
<td>12800</td>
<td>1600</td>
<td>3200</td>
<td>400</td>
<td>400</td>
<td>400</td>
</tr>
</tbody>
</table>

**An M2 antibody titer of 0 represents a titer of <100.**

### TABLE 23

<table>
<thead>
<tr>
<th>mouse</th>
<th>Group A (NP)</th>
<th>Group E (NP + M2)</th>
<th>Group F (NP + M1)</th>
<th>Group G (NP + seg?)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>204800</td>
<td>51200</td>
<td>102400</td>
<td>25600</td>
</tr>
<tr>
<td>2</td>
<td>204800</td>
<td>51200</td>
<td>204800</td>
<td>51200</td>
</tr>
</tbody>
</table>

Study GSJ05:

[0460] In order to attempt to distinguish between the two antigen combinations, S7+NP and M2+NP, a dose ranging challenge experiment was undertaken with these two plasmid combinations. Mice were injected with 100 μg, 30 μg, or 10 μg per plasmid in the 02A poloxamer formulation at 0 and 3 weeks, followed by bleed at 5 weeks and viral challenge at 6 weeks. Sixteen mice per group were vaccinated for test groups A-H, while 12 mice per group were vaccinated for the controls. Poloxamer 02A-formulated HA plasmids, VR4750 (HA H3) and VR4752 (HA H1), were included as positive and negative controls, respectively. The test groups and 21 day survival post-challenge are shown in Table 24:

### TABLE 24

<table>
<thead>
<tr>
<th>Group</th>
<th>Construct(s)</th>
<th>Total pDNA per vaccination</th>
<th># mice/group</th>
<th>21 day survival (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>VR4756 (Seg 7) + VR4762 (NP)</td>
<td>200 μg</td>
<td>16</td>
<td>73</td>
</tr>
<tr>
<td>B</td>
<td>VR4756 (Seg 7) + VR4762 (NP)</td>
<td>60 μg</td>
<td>16</td>
<td>81</td>
</tr>
<tr>
<td>C</td>
<td>VR4756 (Seg 7) + VR4762 (NP)</td>
<td>20 μg</td>
<td>16</td>
<td>69</td>
</tr>
<tr>
<td>D</td>
<td>VR4759 (M2) + VR4762 (NP)</td>
<td>200 μg</td>
<td>16</td>
<td>94</td>
</tr>
<tr>
<td>E</td>
<td>VR4759 (M2) + VR4762 (NP)</td>
<td>60 μg</td>
<td>16</td>
<td>81</td>
</tr>
<tr>
<td>F</td>
<td>VR4759 (M2) + VR4762 (NP)</td>
<td>20 μg</td>
<td>16</td>
<td>75</td>
</tr>
<tr>
<td>G</td>
<td>VR4750 (Positive DNA control)</td>
<td>100 μg</td>
<td>12</td>
<td>100</td>
</tr>
<tr>
<td>H</td>
<td>VR4752 (Negative DNA control)</td>
<td>100 μg</td>
<td>12</td>
<td>8</td>
</tr>
</tbody>
</table>
Results

The performance criteria of >90% survival with the HA positive control and ≤10% for the HA negative control plasmid again were met. The performance criteria for the experimental groups, >75% survival at the 30 μg per plasmid dose, was met by both M2+NP and S7+NP (Table 24). In fact, at a dose of 10 μg per plasmid, S7+NP and M2+NP resulted in 69% and 75% survival, respectively. There was no statistical significance (p<0.05) between the three doses of M2+NP or between the 3 doses of S7+NP, nor was the statistical significance when comparing M2+NP to S7+NP at 200 μg, 60 μg, or 20 μg doses. However, there was a statistical difference for the HA positive control vs. S7+NP at 200 μg and 20 μg. Body mass data shows weight loss and recovery by all surviving experimental plasmid DNA-vaccinated groups, while the HA positive control mice did not experience weight loss. Antibody data for M2 and NP are shown in Tables 25 and 26.

<table>
<thead>
<tr>
<th>GRS05 M2 Antibody Titer</th>
<th>mouse #</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
<th>Group F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2600</td>
<td>400</td>
<td>200</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>0</td>
<td>0</td>
<td>2500</td>
<td>1600</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>3200</td>
<td>3200</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>3200</td>
<td>3200</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3200</td>
<td>3200</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>400</td>
<td>200</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>2500</td>
<td>400</td>
<td>2500</td>
<td>200</td>
<td>400</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3200</td>
<td>3200</td>
<td>12800</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3200</td>
<td>3200</td>
<td>12800</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3200</td>
<td>3200</td>
<td>12800</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3200</td>
<td>3200</td>
<td>12800</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3200</td>
<td>3200</td>
<td>12800</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3200</td>
<td>3200</td>
<td>12800</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3200</td>
<td>3200</td>
<td>12800</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3200</td>
<td>3200</td>
<td>12800</td>
<td>0</td>
</tr>
</tbody>
</table>

TABLE 26-continued

<table>
<thead>
<tr>
<th>GRS05 NP Antibody Titer</th>
<th>mouse #</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
<th>Group F</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>2500</td>
<td>12800</td>
<td>2500</td>
<td>2500</td>
<td>12800</td>
<td>12800</td>
<td>12800</td>
</tr>
<tr>
<td>5</td>
<td>51200</td>
<td>102400</td>
<td>6400</td>
<td>2500</td>
<td>12800</td>
<td>12800</td>
<td>12800</td>
</tr>
<tr>
<td>6</td>
<td>2500</td>
<td>51200</td>
<td>2500</td>
<td>2500</td>
<td>12800</td>
<td>6400</td>
<td>400</td>
</tr>
<tr>
<td>7</td>
<td>102400</td>
<td>51200</td>
<td>6400</td>
<td>2500</td>
<td>12800</td>
<td>6400</td>
<td>400</td>
</tr>
<tr>
<td>8</td>
<td>51200</td>
<td>2500</td>
<td>12800</td>
<td>51200</td>
<td>12800</td>
<td>6400</td>
<td>400</td>
</tr>
<tr>
<td>9</td>
<td>12800</td>
<td>51200</td>
<td>2500</td>
<td>102400</td>
<td>12800</td>
<td>12800</td>
<td>12800</td>
</tr>
<tr>
<td>10</td>
<td>2500</td>
<td>2500</td>
<td>2500</td>
<td>2500</td>
<td>12800</td>
<td>12800</td>
<td>12800</td>
</tr>
<tr>
<td>11</td>
<td>51200</td>
<td>2500</td>
<td>51200</td>
<td>2500</td>
<td>2500</td>
<td>3200</td>
<td>3200</td>
</tr>
<tr>
<td>12</td>
<td>51200</td>
<td>51200</td>
<td>3200</td>
<td>2500</td>
<td>2500</td>
<td>3200</td>
<td>3200</td>
</tr>
<tr>
<td>13</td>
<td>51200</td>
<td>2500</td>
<td>51200</td>
<td>2500</td>
<td>2500</td>
<td>12800</td>
<td>12800</td>
</tr>
<tr>
<td>14</td>
<td>51200</td>
<td>2500</td>
<td>51200</td>
<td>2500</td>
<td>2500</td>
<td>6400</td>
<td>12800</td>
</tr>
<tr>
<td>15</td>
<td>2500</td>
<td>6400</td>
<td>2500</td>
<td>2500</td>
<td>2500</td>
<td>12800</td>
<td>12800</td>
</tr>
<tr>
<td>16</td>
<td>51200</td>
<td>51200</td>
<td>2500</td>
<td>2500</td>
<td>2500</td>
<td>12800</td>
<td>6400</td>
</tr>
</tbody>
</table>

Gray shading represents mice that died post-challenge. Group A, mouse 9 (spotted box) died during the OSE bleed procedure.

Study GJS06

The plasmid combination VR4759 (M2) and VR4762 (NP) was utilized in further mouse influenza challenge studies to examine additional formulations.

Using the experimental protocol described above, 12 mice per group were vaccinated with equal weight VR4759 (M2) and VR4762 (NP) in the following formulations:

- Poloxamer 02A used in the previous two challenge experiments.
- DMRIE+Cholesterol (DM:Chol) at a 4:1 molar ratio of DNA to DMRIE; the molar ratio of DM:Chol is 3:1.
- Vaxfectin™ (VC 1052+DpyPE) at a 4:1 molar ratio of DNA: VC1052, the molar ratio of VC1052: DpyPE is 1:1.

GSJ06 study design and 21 day survival post-challenge is found in Table 27.

<table>
<thead>
<tr>
<th>Group</th>
<th>pDNA (ug)</th>
<th>Total pDNA</th>
<th>21 Day Survival (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Poloxamer 02A</td>
<td>20</td>
<td>92</td>
</tr>
<tr>
<td>B</td>
<td>Poloxamer 02A</td>
<td>20</td>
<td>58</td>
</tr>
<tr>
<td>C</td>
<td>DMRIE: Cholesterol</td>
<td>20</td>
<td>58</td>
</tr>
<tr>
<td>D</td>
<td>DMRIE: Cholesterol</td>
<td>20</td>
<td>58</td>
</tr>
<tr>
<td>E</td>
<td>Vaxfectin</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>F</td>
<td>Vaxfectin</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>G</td>
<td>VR4750 (HA, positive)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>H</td>
<td>VR4752 (HA, negative)</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>
Results

Poloxamer 02A and Vaxfectin™-formulated plasmid DNA led to 92% and 100% survival at the 20 μg pDNA dose, and 58% and 75% at the 2 μg dose, respectively (Table 27).

Average weights were tracked for each group of mice starting at the day of challenge. As shown in Table 28, it was noted in this experiment that the weight recovery for group E (Vaxfectin™-formulated pDNA, 20 μg total) began after day 4, as opposed to the other groups’ recovery beginning at day 7. Antibody titers, Tables 29 and 30, were determined for M2 and NP and shaded boxes represent mice that died following viral challenge.

### TABLE 28

<table>
<thead>
<tr>
<th>Group</th>
<th>pDNA</th>
<th>Total pDNA</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>7</th>
<th>9</th>
<th>11</th>
<th>14</th>
<th>16</th>
<th>18</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Poloxamer 02A</td>
<td>20 μg</td>
<td>20.73</td>
<td>19.98</td>
<td>17.98</td>
<td>16.14</td>
<td>17.36</td>
<td>18.74</td>
<td>19.94</td>
<td>20.45</td>
<td>20.60</td>
<td>21.08</td>
</tr>
<tr>
<td>B</td>
<td>Poloxamer 02A</td>
<td>2 μg</td>
<td>21.08</td>
<td>19.91</td>
<td>17.86</td>
<td>15.17</td>
<td>15.45</td>
<td>16.03</td>
<td>16.77</td>
<td>17.41</td>
<td>18.10</td>
<td>19.52</td>
</tr>
<tr>
<td>E</td>
<td>Vaxfectin</td>
<td>20 μg</td>
<td>21.41</td>
<td>19.97</td>
<td>17.83</td>
<td>15.17</td>
<td>15.45</td>
<td>16.03</td>
<td>16.77</td>
<td>17.41</td>
<td>18.10</td>
<td>19.52</td>
</tr>
<tr>
<td>F</td>
<td>Vaxfectin</td>
<td>2 μg</td>
<td>20.47</td>
<td>18.97</td>
<td>16.86</td>
<td>15.17</td>
<td>16.22</td>
<td>16.84</td>
<td>17.87</td>
<td>18.60</td>
<td>19.08</td>
<td>20.22</td>
</tr>
<tr>
<td>H</td>
<td>VR4752 (H2, negative)</td>
<td>100 μg</td>
<td>20.89</td>
<td>20.25</td>
<td>17.57</td>
<td>14.67</td>
<td>15.17</td>
<td>16.22</td>
<td>16.84</td>
<td>17.87</td>
<td>18.60</td>
<td>19.08</td>
</tr>
</tbody>
</table>

Shading represents the lowest group average post-challenge for each test group. Group H (negative control) weight averages are not recorded since the percentage survival has dropped below 50%.

### TABLE 29

<table>
<thead>
<tr>
<th>mouse #</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
<th>Group F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>800</td>
<td>400</td>
<td>200</td>
<td>0</td>
<td>1600</td>
<td>6400</td>
</tr>
<tr>
<td>2</td>
<td>6400</td>
<td>400</td>
<td>200</td>
<td>0</td>
<td>1600</td>
<td>6400</td>
</tr>
<tr>
<td>3</td>
<td>6400</td>
<td>400</td>
<td>200</td>
<td>0</td>
<td>1600</td>
<td>6400</td>
</tr>
<tr>
<td>4</td>
<td>1600</td>
<td>0</td>
<td>400</td>
<td>0</td>
<td>25600</td>
<td>1600</td>
</tr>
<tr>
<td>5</td>
<td>6400</td>
<td>3200</td>
<td>1600</td>
<td>0</td>
<td>400</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>3200</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>12800</td>
<td>1600</td>
</tr>
<tr>
<td>7</td>
<td>800</td>
<td>1600</td>
<td>1600</td>
<td>0</td>
<td>800</td>
<td>3200</td>
</tr>
<tr>
<td>8</td>
<td>400</td>
<td>100</td>
<td>3200</td>
<td>200</td>
<td>6400</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td>1600</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>6400</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>400</td>
<td>1600</td>
<td>100</td>
<td>3200</td>
<td>400</td>
</tr>
<tr>
<td>11</td>
<td>3200</td>
<td>0</td>
<td>800</td>
<td>0</td>
<td>1600</td>
<td>1600</td>
</tr>
<tr>
<td>12</td>
<td>6400</td>
<td>0</td>
<td>400</td>
<td>0</td>
<td>6400</td>
<td>1600</td>
</tr>
</tbody>
</table>

### TABLE 30

<table>
<thead>
<tr>
<th>mouse #</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
<th>Group F</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>25600</td>
<td>1600</td>
<td>6400</td>
<td>3200</td>
<td>20480</td>
<td>102400</td>
</tr>
<tr>
<td>5</td>
<td>25600</td>
<td>6400</td>
<td>25600</td>
<td>3200</td>
<td>51200</td>
<td>51200</td>
</tr>
<tr>
<td>6</td>
<td>51200</td>
<td>12800</td>
<td>25600</td>
<td>12800</td>
<td>102400</td>
<td>51200</td>
</tr>
<tr>
<td>7</td>
<td>25600</td>
<td>25600</td>
<td>12800</td>
<td>12800</td>
<td>6400</td>
<td>25600</td>
</tr>
<tr>
<td>8</td>
<td>25600</td>
<td>3200</td>
<td>12800</td>
<td>12800</td>
<td>6400</td>
<td>25600</td>
</tr>
<tr>
<td>9</td>
<td>25600</td>
<td>6400</td>
<td>51200</td>
<td>400</td>
<td>51200</td>
<td>25600</td>
</tr>
<tr>
<td>10</td>
<td>51200</td>
<td>6400</td>
<td>12800</td>
<td>12800</td>
<td>51200</td>
<td>51200</td>
</tr>
<tr>
<td>11</td>
<td>25600</td>
<td>12800</td>
<td>25600</td>
<td>12800</td>
<td>400</td>
<td>51200</td>
</tr>
<tr>
<td>12</td>
<td>51200</td>
<td>6400</td>
<td>12800</td>
<td>12800</td>
<td>400</td>
<td>51200</td>
</tr>
</tbody>
</table>

Study GSJ08

Further formulation comparisons were done with utilizing VR4759 (M2) and VR4762 (NP). Seventeen mice per test group (A-G) were vaccinated with equal weight VR4759 (M2) and VR4762 (NP) vectors in the following formulations:
Poloxamer 02A

Vaxfectin™ (preparations A and B represent different purifications)

DMRIE:DOPE at a 4:1 molar ratio of DNA to DMRIE

DMRIE:DOPE at a 2.5:1 molar ratio of DNA to DMRIE

PBS (unformulated pDNA)

Twelve mice per test group were challenged with influenza virus at week number 6. Five mice per test group were sacrificed at days 36-38 for T cell assays (IFN-γ ELISPOT). The test groups and 21 day survival post-challenge are shown in Table 31. Groups A-D, and F-G were vaccinated with 20 μg total plasmid DNA per injection to further explore the weight loss/recovery phenomena seen in study GSJ06 with the Vaxfectin™-formulated pDNA.

<table>
<thead>
<tr>
<th>Group</th>
<th>Construct(s)</th>
<th>Total pDNA per vaccination</th>
<th>21 Day Survival (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Poloxamer 02A</td>
<td>20 μg</td>
<td>50</td>
</tr>
<tr>
<td>B</td>
<td>DMRIE:DOPE 4:1</td>
<td>20 μg</td>
<td>92</td>
</tr>
<tr>
<td>C</td>
<td>DMRIE:DOPE 2.5:1</td>
<td>20 μg</td>
<td>92</td>
</tr>
<tr>
<td>D</td>
<td>Vaxfectin - prep A</td>
<td>20 μg</td>
<td>92</td>
</tr>
<tr>
<td>E</td>
<td>Vaxfectin - prep A</td>
<td>2 μg</td>
<td>75</td>
</tr>
<tr>
<td>F</td>
<td>Vaxfectin - prep B</td>
<td>20 μg</td>
<td>100</td>
</tr>
<tr>
<td>G</td>
<td>PBS</td>
<td>20 μg</td>
<td>42</td>
</tr>
<tr>
<td>H</td>
<td>VR4750 (HA, H1N1, +control)</td>
<td>100 μg</td>
<td>300</td>
</tr>
<tr>
<td>I</td>
<td>VR4752 (HA, H1N1, −control)</td>
<td>100 μg</td>
<td>17</td>
</tr>
</tbody>
</table>

Results

The DMRIE:DOPE and Vaxfectin™ formulated groups resulted in 92-100% survival at a 20 μg pDNA dose. Group A (Poloxamer 02A) and Group G (PBS) survival results were not statistically different than the negative control (as measured by Fisher exact p, one-tailed), while the Vaxfectin™ and DMRIE:DOPE Groups (Groups B-F) were shown to be statistically superior (p<0.05) as compared to the negative control. Therefore, the plasmid DNA formulated with lipids appear to provide superior protection in the mouse influenza model challenge.

A repeated measures ANOVA mixed model analysis of weight data for groups B, C, and D of the weight loss and recovery data showed that Group B and Group D were not statistically different, while Group C and Group D were statistically different.

T cell responses, as measured by IFN-γ ELISPOT assay, were conducted on the last 5 mice per group using an M2 peptide encompassing the first 24 amino acids of M2 (TABLE 33), an NP protein expressed in baculovirus (TABLE 34), and an NP CDS+ Bulb/c immunodominant peptide (TABLE 35).

Antibody titers, Tables 36 and 37, were determined for M2 and NP proteins. The first 12 mice listed for each group were challenge at day 42 and the last 5 mice per group were sacrificed for IFN-γ ELISPOT. The shaded boxes represent mice that died following viral challenge.

<table>
<thead>
<tr>
<th>Group</th>
<th>Construct(s)</th>
<th>Total pDNA per vaccination</th>
<th>Avg Body Weights (g)-Days post-challenge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>DMRIE:DOPE 2.5:1</td>
<td>20 μg</td>
<td>15.95</td>
</tr>
<tr>
<td>F</td>
<td>Vaxfectin - prep B</td>
<td>20 μg</td>
<td>21.33</td>
</tr>
<tr>
<td>G</td>
<td>PBS</td>
<td>20 μg</td>
<td>20.84</td>
</tr>
<tr>
<td>I</td>
<td>VR4752 (HA, H1N1, −control)</td>
<td>100 μg</td>
<td>21.67</td>
</tr>
</tbody>
</table>

Shading represents the lowest group average post-challenge for each test group. Group G and I weight averages are not recorded once the percentage survival has dropped below 50%.
### TABLE 33

<table>
<thead>
<tr>
<th>Mouse</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>66</td>
<td>88</td>
<td>145</td>
<td>189</td>
<td>283</td>
<td>253</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>31</td>
<td>115</td>
<td>150</td>
<td>269</td>
<td>62</td>
<td>282</td>
<td>47</td>
</tr>
<tr>
<td>3</td>
<td>115</td>
<td>247</td>
<td>190</td>
<td>233</td>
<td>99</td>
<td>283</td>
<td>112</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>6</td>
<td>51</td>
<td>67</td>
<td>73</td>
<td>93</td>
<td>45</td>
</tr>
<tr>
<td>5</td>
<td>93</td>
<td>277</td>
<td>367</td>
<td>248</td>
<td>202</td>
<td>399</td>
<td>93</td>
</tr>
<tr>
<td>AVG</td>
<td>61</td>
<td>147</td>
<td>187</td>
<td>201</td>
<td>144</td>
<td>262</td>
<td>66</td>
</tr>
</tbody>
</table>

### TABLE 34-continued

<table>
<thead>
<tr>
<th>Mouse</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVG</td>
<td>17</td>
<td>74</td>
<td>119</td>
<td>125</td>
<td>40</td>
<td>84</td>
<td>31</td>
</tr>
</tbody>
</table>

### TABLE 34

<table>
<thead>
<tr>
<th>Mouse</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>32</td>
<td>3</td>
<td>52</td>
<td>72</td>
<td>108</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>83</td>
<td>34</td>
<td>125</td>
<td>8</td>
<td>34</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>22</td>
<td>91</td>
<td>106</td>
<td>293</td>
<td>26</td>
<td>51</td>
<td>73</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>15</td>
<td>80</td>
<td>39</td>
<td>53</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>37</td>
<td>150</td>
<td>374</td>
<td>117</td>
<td>40</td>
<td>217</td>
<td>43</td>
</tr>
</tbody>
</table>

### TABLE 35

<table>
<thead>
<tr>
<th>Mouse</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVG</td>
<td>18</td>
<td>19</td>
<td>12</td>
<td>16</td>
<td>12</td>
<td>34</td>
<td>13</td>
</tr>
</tbody>
</table>

### TABLE 36

<table>
<thead>
<tr>
<th>mouse #</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
<th>Group F</th>
<th>Group G</th>
<th>Group H</th>
<th>ELISPOT #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1600</td>
<td>3200</td>
<td>3200</td>
<td>6400</td>
<td>400</td>
<td>12800</td>
<td>800</td>
<td>6400</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>12800</td>
<td>6400</td>
<td>6400</td>
<td>1600</td>
<td>3200</td>
<td>800</td>
<td>1600</td>
<td>800</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>3200</td>
<td>6400</td>
<td>25600</td>
<td>800</td>
<td>3200</td>
<td>1600</td>
<td>800</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>800</td>
<td>0</td>
<td>6400</td>
<td>1600</td>
<td>400</td>
<td>800</td>
<td>1600</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1600</td>
<td>0</td>
<td>800</td>
<td>12800</td>
<td>1600</td>
<td>800</td>
<td>800</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>8400</td>
<td>3200</td>
<td>1600</td>
<td>6400</td>
<td>200</td>
<td>12800</td>
<td>400</td>
<td>800</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>12800</td>
<td>3200</td>
<td>12800</td>
<td>800</td>
<td>1600</td>
<td>3200</td>
<td>1600</td>
<td>6400</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1600</td>
<td>6400</td>
<td>3200</td>
<td>12800</td>
<td>12800</td>
<td>12800</td>
<td>12800</td>
<td>400</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1600</td>
<td>1600</td>
<td>0</td>
<td>12800</td>
<td>6400</td>
<td>12800</td>
<td>1600</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>3200</td>
<td>1600</td>
<td>12800</td>
<td>12800</td>
<td>1600</td>
<td>800</td>
<td>1600</td>
<td>12800</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1600</td>
<td>6400</td>
<td>3200</td>
<td>3200</td>
<td>0</td>
<td>6400</td>
<td>800</td>
<td>400</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>200</td>
<td>800</td>
<td>6400</td>
<td>25600</td>
<td>1600</td>
<td>800</td>
<td>6400</td>
<td>6400</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1600</td>
<td>800</td>
<td>6400</td>
<td>12800</td>
<td>3200</td>
<td>6400</td>
<td>6400</td>
<td>6400</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>3200</td>
<td>6400</td>
<td>1600</td>
<td>1600</td>
<td>800</td>
<td>12800</td>
<td>3200</td>
<td>12800</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>1600</td>
<td>3200</td>
<td>3200</td>
<td>12800</td>
<td>12800</td>
<td>6400</td>
<td>12800</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>3200</td>
<td>3200</td>
<td>1600</td>
<td>12800</td>
<td>0</td>
<td>12800</td>
<td>200</td>
<td>6400</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>3200</td>
<td>200</td>
<td>400</td>
<td>6400</td>
<td>800</td>
<td>400</td>
<td>1600</td>
<td>3200</td>
<td></td>
</tr>
</tbody>
</table>

**GSF08 M2 Antibody Titers**
TABLE 37

<table>
<thead>
<tr>
<th>mouse</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
<th>Group F</th>
<th>Group G</th>
<th>Group H</th>
<th>ELISPOT #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>51200</td>
<td>25600</td>
<td>6400</td>
<td>51200</td>
<td>12800</td>
<td>51200</td>
<td>51200</td>
<td>25600</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>51200</td>
<td>25600</td>
<td>6400</td>
<td>51200</td>
<td>12800</td>
<td>51200</td>
<td>51200</td>
<td>25600</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>26000</td>
<td>12800</td>
<td>12800</td>
<td>51200</td>
<td>102400</td>
<td>12800</td>
<td>51200</td>
<td>12800</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12800</td>
<td>25600</td>
<td>51200</td>
<td>102400</td>
<td>12800</td>
<td>25600</td>
<td>25600</td>
<td>25600</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>102400</td>
<td>12800</td>
<td>51200</td>
<td>102400</td>
<td>25600</td>
<td>51200</td>
<td>12800</td>
<td>51200</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>51200</td>
<td>25600</td>
<td>51200</td>
<td>12800</td>
<td>51200</td>
<td>51200</td>
<td>25600</td>
<td>25600</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>25600</td>
<td>12800</td>
<td>51200</td>
<td>102400</td>
<td>51200</td>
<td>12800</td>
<td>51200</td>
<td>51200</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>12800</td>
<td>25600</td>
<td>51200</td>
<td>102400</td>
<td>51200</td>
<td>12800</td>
<td>51200</td>
<td>51200</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>6400</td>
<td>12800</td>
<td>51200</td>
<td>25600</td>
<td>12800</td>
<td>51200</td>
<td>25600</td>
<td>25600</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>12800</td>
<td>51200</td>
<td>25600</td>
<td>204800</td>
<td>12800</td>
<td>102400</td>
<td>51200</td>
<td>51200</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>25600</td>
<td>12800</td>
<td>51200</td>
<td>102400</td>
<td>25600</td>
<td>204800</td>
<td>12800</td>
<td>51200</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>25600</td>
<td>12800</td>
<td>51200</td>
<td>102400</td>
<td>25600</td>
<td>204800</td>
<td>12800</td>
<td>51200</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>25600</td>
<td>25600</td>
<td>12800</td>
<td>51200</td>
<td>102400</td>
<td>25600</td>
<td>25600</td>
<td>25600</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>25600</td>
<td>102400</td>
<td>12800</td>
<td>51200</td>
<td>102400</td>
<td>25600</td>
<td>51200</td>
<td>12800</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>25600</td>
<td>12800</td>
<td>102400</td>
<td>51200</td>
<td>25600</td>
<td>102400</td>
<td>51200</td>
<td>12800</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>25600</td>
<td>6400</td>
<td>25600</td>
<td>51200</td>
<td>12800</td>
<td>102400</td>
<td>25600</td>
<td>51200</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>25600</td>
<td>25600</td>
<td>51200</td>
<td>12800</td>
<td>51200</td>
<td>25600</td>
<td>51200</td>
<td>25600</td>
<td></td>
</tr>
</tbody>
</table>

[0489] The present invention is not to be limited in scope by the specific embodiments described which are intended as single illustrations of individual aspects of the invention, and any compositions or methods which are functionally equivalent are within the scope of this invention. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Such modifications are intended to fall within the scope of the appended claims.

[0490] All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

SEQUENCE LISTING

\[<160> NUMBER OF SEQ ID NOS: 112\]
\[<210> SEQ ID NO 1\]
\[<211> LENGTH: 1565\]
\[<212> TYPE: DNA\]
\[<213> ORGANISM: Influenza A virus\]
\[<400> SEQUENCE: 1\]

agcacaagc aggtgatgaa tcactcactg agtgaatcct aatacagcc qgtoaagggc
accacaagct ttacaacaa gatggaacct gatggcaagc gcaagatggt ctcctgaac
agagcactcg tcgaaacat gattggttga atggctgacat ttatacctgaa attgtaacct
aacgcaaaac tcagcttctg gtaggcaagc ttgctccaa aaacgtttaac aattaggaaga
taggtgtct ctgcttttga cgaagagga aataaatgga ttgacgacaa tcgccaggtgc
agcagacact ctcagagac ctgagacct cttatacagga gtaggcaagc aaatggttag
agagactca ttcttttga caaagagaa ataagcgga ccctggtcga agctataaat

---continued

ggtgagcagt caacagctgg ttcgctcaca atgtgatct ggcattcaca ttggaaatgat 480
gcaacctct acgacgaac atcctttgct ctgacgcaac tggctgacac gatggtcgtct 540
cctgctcaag gttcactctt cctgagaggg ttcgagctgc cagctgctgc agtcaaggg 600
gttgcaaca tgcgatgagga attgctcagc atgatcaaac cttggacacac tggagcggcgc 660
tctgagggg tcggataattg ccgaaaaacc aagataattg atgaatattg tgtgacacatt 720
ctgcaaggg gattccatac gttgccaca acagtaatgg tcgatctagt gcagagagagc 780
gcgacaccgc ggaatgtcgtg tcggagatgtc cttaattttc tcgacggttc tcgacactcata 840
ttcgaggtgg cctgctgca acacgctctg ctcgctgctc gtctgtcattg acactgctca 900
gcaggttcgtt acagcttttgca aagagaggggt taccctctctag tcggataatag gcctttttcaga 960
cgctcccaac acacgcaagt gcacagctta atcagagcaca atgagacact acagacacag 1020
agtcaactcg ttgcagagttgt cgcctattct gcctgctattc agatctcaag agttactacag 1080
ttcctcaag gcggcgaggt gcctcacaag ggagccgccc ctctagcggt agcttcacatt 1140
gtccctcaag aacatataag gcctatgg栽 aacactatac ttcggctcag aacagctc 1200
tggctcaata gcacagctaa aggagacaga acacgctaatc acagacagttg cgggacg 1260
atacgtacata aacagctcag tgcggatcct cagcagccatc agaactcag caccccttt 1320
atgctgctc ttcgcagggc taccagattg gcacattgct atcgagggc cgaatcata 1380
aggttagcatt aacagcagctt acagcttttgc gggaggggtg agtctcctag 1440
cctgctgacac gcagccttct gcggccatgct ttcgctgctt gctgagctat gcacagagagc 1500
ttcggagctt cagcagctttgc tgcctgctct gcggcagctt gctgctgctt gctgctgct 1560
catatctt tcggagctttc tgcgtgccag tgcctgcttt gctgctgtcctt acctctgcttt 1620
ctact 1656

<210> SEQ ID NO 2
<211> LENGTH: 498
<212> TYPE: PRT
<213> ORGANISM: Influenza A virus

<400> SEQUENCE: 2

Met Ala Ser Gln Gly Thr Lys Arg Ser Thr Glu Gln Met Glu Thr Asp
1        15
Gly Glu Arg Gln Asn Ala Thr Glu Ile Arg Ala Ser Val Gly Lys Met
20       30
Ile Gly Gly Ile Gly Arg Phe Tyr Ile Glu Met Cys Thr Glu Leu Lys
35       45
Leu Ser Asp Tyr Glu Arg Leu Ile Glu Asn Ser Leu Thr Ile Glu
50       60
Arg Met Val Leu Ser Ala Phe Asp Glu Arg Asp Ala Thr Lys Thr Glu Glu
65       80
His His Pro Ser Ala Lys Asp Pro Lys Lys Thr Gly Gly Pro Ile
85       95
Tyr Arg Arg Val Asn Gly Lys Trp Met Arg Glu Leu Ile Leu Tyr Asp
100      110
Lys Glu Glu Ile Arg Arg Ile Trp Arg Glu Ala Asn Gln Gln Asp Asp
115      125
Ala Thr Ala Gly Leu Thr His Met Met Ile Trp His Ser Asn Leu Asn
130      140
Asp Ala Thr Tyr Gln Arg Thr Ala Leu Val Arg Thr Gly Met Asp
---continued---

145  150  155  160
Pro Arg Met Cys Ser Met Met Gln Gly Ser Thr Leu Pro Arg Arg Ser
165  170  175
Gly Ala Ala Gly Ala Ala Val Lys Gly Val Gly Thr Met Val Met Glu
180  185  190
Leu Val Arg Met Ile Lys Arg Gly Ile Asn Arg Arg Asp Asp Phe Trp Arg
195  200  205
Gly Glu Asn Gly Arg Lys Thr Arg Ile Ala Tyr Glu Arg Met Cys Asn
210  215  220
Ile Leu Lys Gly Lys Phe Gln Thr Ala Ala Gln Lys Ala Met Met Asp
225  230  235  240
Gln Val Arg Glu Ser Arg Asn Pro Gly Asn Ala Glu Phe Glu Asp Leu
245  250  255
Thr Phe Leu Ala Arg Ser Ala Leu Ile Leu Arg Gly Ser Val Ala His
260  265  270
Lys Ser Cys Leu Pro Ala Cys Val Tyr Gly Pro Ala Val Ala Ser Gly
275  280  285
Tyr Asp Phe Glu Arg Glu Gly Tyr Ser Leu Val Gly Ile Asp Pro Phe
290  295  300
Arg Leu Leu Gln Asn Ser Glu Val Tyr Ser Leu Ile Arg Pro Asn Glu
305  310  315  320
Asn Pro Ala His Lys Ser Ser Glu Leu Val Trp Met Ala Cys His Ser Ala
325  330  335
Ala Phe Glu Asp Leu Arg Val Leu Ser Phe Ile Lys Gly Thr Lys Val
340  345  350
Leu Pro Arg Gly Lys Leu Ser Thr Arg Gly Val Glu Ile Ala Ser Asn
355  360  365
Glu Asn Met Glu Thr Met Glu Ser Ser Thr Leu Met Leu Arg Ser Arg
370  375  380
Tyr Thr Ala Ile Arg Arg Ser Gly Gly Arg Thr Asn Gln Gln Arg
385  390  395  400
Ala Ser Ala Gly Glu Ile Ser Ile Gin Pro Thr Phe Ser Val Gin Arg
405  410  415
Asn Leu Pro Phe Asp Arg Thr Val Met Ala Ala Phe Ser Gly Asn
420  425  430
Thr Glu Gly Arg Thr Ser Asp Met Arg Thr Glu Ile Arg Arg Met Met
435  440  445
Glu Ser Ala Arg Pro Glu Val Ser Phe Glu Gly Arg Gly Val Phe
450  455  460
Glu Leu Ser Asp Glu Lys Ala Ala Ser Pro Ile Val Pro Ser Phe Asp
465  470  475  480
Met Ser Asn Gly Ser Tyr Phe Phe Gly Asp Asn Ala Glu Glu Tyr
485  490  495
Asp Asn

<210> SEQ ID NO 3
<211> LENGTH: 1027
<212> TYPE: DNA
<213> ORGANISM: Influenza A virus
<400> SEQUENCE: 3

agcggagaga ggtagatatt gaaagatgg agttctaccc gcagctcagaa ccgagcatcact  60
<table>
<thead>
<tr>
<th>Position</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-15</td>
<td>Met Ser Leu Leu Thr Glu Val Glu Thr Val Leu Ser Ile Ile Pro</td>
</tr>
<tr>
<td>16-20</td>
<td>Gly Pro Leu Lys Ala Glu Ile Ala Gln Arg Leu Glu Asp Val Phe</td>
</tr>
<tr>
<td>21-25</td>
<td>Ala Gly Lys Asn Thr Asp Leu Glu Val Leu Met Glu Trp Leu Lys Thr</td>
</tr>
<tr>
<td>26-30</td>
<td>Arg Pro Ile Leu Ser Pro Leu Thr Lys Gly Ile Leu Gly Phe Val Phe</td>
</tr>
<tr>
<td>31-35</td>
<td>Thr Leu Thr Val Pro Ser Glu Arg Gly Leu Gln Arg Arg Arg Phe Val</td>
</tr>
<tr>
<td>36-40</td>
<td>Gln Asn Ala Leu Asn Gly Asn Gly Asp Pro Asn Asn Met Asp Lys Ala</td>
</tr>
<tr>
<td>41-45</td>
<td>Val Lys Leu Tyr Arg Lys Leu Lys Arg Glu Ile Thr Phe His Gly Ala</td>
</tr>
<tr>
<td>46-50</td>
<td>Lys Glu Ile Ser Leu Ser Tyr Ser Ala Gly Ala Leu Ala Ser Cys Met</td>
</tr>
<tr>
<td>51-55</td>
<td>Gly Leu Ile Tyr Asn Arg Met Gly Ala Val Thr Glu Val Ala Phe</td>
</tr>
<tr>
<td>56-60</td>
<td>Gly Leu Val Cys Ala Thr Cys Glu Gin Ile Ala Asp Ser Gin His Arg</td>
</tr>
<tr>
<td>61-65</td>
<td>Ser His Arg Gin Met Val Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr</td>
</tr>
<tr>
<td>66-70</td>
<td>Ser His Arg Gin Met Val Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr</td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 4
<211> LENGTH: 252
<212> TYPE: PRT
<213> ORGANISM: Influenza A virus

<400> SEQUENCE: 4

Met Ser Leu Leu Thr Glu Val Glu Thr Val Leu Ser Ile Ile Pro
1    5    10    15
Gly Pro Leu Lys Ala Glu Ile Ala Gln Arg Leu Glu Asp Val Phe
20   25    30
Ala Gly Lys Asn Thr Asp Leu Glu Val Leu Met Glu Trp Leu Lys Thr
35   40    45
Arg Pro Ile Leu Ser Pro Leu Thr Lys Gly Ile Leu Gly Phe Val Phe
50   55    60
Thr Leu Thr Val Pro Ser Glu Arg Gly Leu Gln Arg Arg Arg Phe Val
65   70    75    80
Gln Asn Ala Leu Asn Gly Asn Gly Asp Pro Asn Asn Met Asp Lys Ala
85   90    95
Val Lys Leu Tyr Arg Lys Leu Lys Arg Glu Ile Thr Phe His Gly Ala
100  105   110
Lys Glu Ile Ser Leu Ser Tyr Ser Ala Gly Ala Leu Ala Ser Cys Met
115  120   125
Gly Leu Ile Tyr Asn Arg Met Gly Ala Val Thr Glu Val Ala Phe
130  135   140
Gly Leu Val Cys Ala Thr Cys Glu Gin Ile Ala Asp Ser Gin His Arg
145  150   155   160
Ser His Arg Gin Met Val Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr
-continued

Ann Arg Met Val Leu Ala Ser Thr Thr Ala Lys Ala Met Glu Gin Met
180 185 190
Ala Gly Ser Ser Glu Gin Ala Ala Ala Glu Met Glu Val Ala Ser Gin
195 200 205
Ala Arg Gin Met Val Gin Ala Met Arg Thr Ile Gly Thr His Pro Ser
210 215 220
Ser Ser Ala Gly Leu Lys Aaa Aap Leu Leu Glu Aaa Leu Gin Ala Tyr
225 230 235 240
Gln Lys Arg Met Gly Val Gin Met Gin Arg Phe Lys
245 250

<210> SEQ ID NO 5
<211> LENGTH: 97
<212> TYPE: PRT
<213> ORGANISM: Influenza A virus

<400> SEQUENCE: 5

Met Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Aan Gin Glu Trp Gly
1  5   10   15
Cys Arg Cys Aan Gly Ser Ser Aap Pro Leu Ala Ile Ala Ala Aan Ile
20  25  30
Ile Gly Ile Leu His Leu Thr Leu Trp Ile Leu Aap Arg Leu Phe Phe
35  40  45
Lys Cys Ile Tyr Arg Arg Phe Lys Tyr Gly Leu Lys Gly Gly Pro Ser
50  55  60
Thr Glu Gly Val Pro Lys Ser Met Arg Glu Glu Tyr Arg Lys Glu Gin
65  70  75  80
Gln Ser Ala Val Aap Ala Aap Gly His Phe Val Ser Ile Leu
85  90  95
Glu

<210> SEQ ID NO 6
<211> LENGTH: 1566
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE: OTHER INFORMATION: eH2NP fusion

<400> SEQUENCE: 6

atggttttc taaacaggg tgtgaagccgt atcagaaacg aatgggggtg aagatgcaac 60
ggacaaggt atatgggtct tcagggcacc aacaagcttt acaagacgat gagactgat
120
gcagacggg aagatgccac tgaattcag aagactgctcg gaaagtggagt tggggaatt
180
ggacgatct acatccacaa aatgcacgga ctcacacatca gtgattatga ggcgctgttg
240
tagcagaacc cttttaagca gcagcgattt gacttggggc cactactcg ttttttcaag
300
agacttgg aatgaatttc cagcgccg ggtcgccaga aagactacttct aagactacttg aagacctta
360
taagccaggg taaacaggg tgtgaagccgt atcagaaacg aatgggggtg aagatgcaac 420
ggacaaggt atatgggtct tcagggcacc aacaagcttt acaagacgat gagactgat
480
tagcagaacc cttttaagca gcagcgattt gacttggggc cactactcg ttttttcaag
540
agacttgg aatgaatttc cagcgccg ggtcgccaga aagactacttct aagactacttg aagacctta
600
taagccaggg taaacaggg tgtgaagccgt atcagaaacg aatgggggtg aagatgcaac 660
agacttgg aatgaatttc cagcgccg ggtcgccaga aagactacttct aagactacttg aagacctta
720
atgcttattg aagaaatgtg caacatttct aasagggaaat ttcacaactgct tgcacaaaaa 780
gcagtattg atcaagttag agagacgagg aaccaccgaa aaatgtgaaat cagagatcctctc 840
acctttttag acacgtcttg agaggttggg ttgcttccaa atcttgtgcag 900
cctgctgtg tgtatgcacc tcgcgtcggc agtggctagc atcttgaaag gcagggatcc 960
tatcaagtg gatagaggcc ttcacaagct ttcacaaaaa gcacaagtga cagcctaac 1020
agacaaatg agaataccgg aacaagggct aacagtcggt gtagacctg ccaattgctgc 1080
gcatttgag atctaaaagt attaagcctc acaaaaagga gcaaggtgct ccaagaggg 1140
agccttccca cttgagaaag tcaaaattgct tccatgtgaa atcaagagac tagggataca 1200
agtcacagc acctgtgaaac cagcctaagc gcatctaacaa ccaagaaacc 1260
atcacaaga gggcatctgc gcgcacacct gcacatacctc ctgcttcccct gcgtcagaga 1320
aatcctcct tgtcagacac aaccgtaatg gcagctcatt gttgagaaac agagggagag 1380
acatctgaca tgcaagcccata cacttacagg atgatgaaa ggtcaagcacc agaagatgct 1440
tttctccgc gcgcagggatat cttcgctgct ttggcagaaag gcaccagcag ccgctgtgct 1500
cctctgcctg atctaggtas tgaaggtat tatttttctg gagaacatag agaggaatac 1560
gataat 1566

<210> SEQ ID NO 7
<211> LENGTH: 522
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: eM2NP fusion
<400> SEQUENCE: 7

Met Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Arg Ala Glu Trp Gly 1 5 10 15
Cys Arg Cys Asn Gly Ser Ser Asp Met Ala Ser Glu Gly Thr Lys Arg 20 25 30
Ser Tyr Glu Gln Met Glu Thr Asp Gly Glu Arg Glu Ala Asn Thr Glu 35 40 45
Ile Arg Ala Ser Val Gly Lys Met Ile Gly Gly Ile Gly Arg Phe Tyr 50 55 60
Ile Gln Met Cys Thr Glu Leu Lys Leu Ser Asp Tyr Glu Gly Arg Leu 65 70 75 80
Ile Gln Asn Ser Leu Thr Ile Glu Arg Met Val Leu Ser Ala Phe Asp 95 95
Glu Arg Arg Asn Lys Tyr Leu Glu Glu His Pro Ser Ala Gly Lys Asp 100 105 110
Pro Lys Lys Thr Gly Gly Pro Ile Tyr Arg Arg Val Asn Gly Lys Trp 115 120 125
Met Arg Glu Leu Ile Leu Tyr Asp Lys Glu Ile Arg Arg Ile Trp 130 135 140
Arg Gln Ala Asn Asn Gly Asp Asp Ala Thr Ala Gly Leu Thr His Met 145 150 155 160
Met Ile Trp His Ser Asn Leu Asn Asp Ala Thr Tyr Glu Arg Thr Arg 165 170 175
Ala Leu Val Arg Thr Gly Met Asp Pro Arg Met Cys Ser Leu Met Glu 180 185 190
---continued---

Gly Ser Thr Leu Pro Arg Arg Ser Gly Ala Ala Gly Ala Ala Val Lys
     195  200  205
Gly Val Gly Thr Met Val Met Glu Leu Val Arg Met Ile Lys Arg Gly
     210  215  220
Ile Asp Arg Asn Phe Trp Arg Gly Asp Gly Arg Lys Thr Arg
     225  230  235  240
Ile Ala Tyr Glu Arg Met Cys Asn Ile Leu Lys Gly Lys Phe Gln Thr
     245  250  255
Ala Ala Glu Lys Ala Met Met Asp Glu Val Arg Glu Ser Arg Asn Pro
     260  265  270
Gly Asn Ala Glu Phe Glu Asp Leu Thr Phe Leu Ala Arg Ser Ala Leu
     275  280  285
Ile Leu Arg Gly Ser Val Ala His Lys Ser Cys Leu Pro Ala Cys Val
     290  295  300
Tyr Gly Pro Ala Val Ala Ser Gly Tyr Asp Phe Glu Arg Gly Tyr
     305  310  315  320
Ser Leu Val Gly Ile Asp Pro Phe Arg Leu Leu Glu Asn Ser Gln Val
     325  330  335
Tyr Ser Leu Ile Arg Pro Asn Glu Asn Pro Ala His Lys Ser Glu Leu
     340  345  350
Val Trp Met Ala Cys His Ser Ala Ala Phe Glu Asp Leu Arg Val Leu
     355  360  365
Ser Phe Ile Lys Gly Thr Lys Val Leu Pro Arg Gly Lys Leu Ser Thr
     370  375  380
Arg Gly Val Gln Ile Ala Ser Asn Glu Asn Met Glu Thr Met Glu Ser
     385  390  395  400
Ser Thr Leu Glu Leu Arg Ser Arg Tyr Trp Ala Ile Arg Thr Arg Ser
     405  410  415
Gly Gly Asn Thr Asn Gln Glu Arg Ala Ser Ala Gly Glu Ile Ser Ile
     420  425  430
Gln Pro Thr Phe Ser Val Gln Arg Asn Leu Pro Phe Asp Arg Thr Thr
     435  440  445
Val Met Ala Ala Phe Ser Gly Asn Thr Glu Arg Thr Ser Asp Met
     450  455  460
Arg Thr Glu Ile Arg Met Met Gly Ser Ala Arg Pro Glu Asp Val
     465  470  475  480
Ser Phe Gln Gly Arg Gly Val Phe Leu Ser Asp Glu Lys Ala Ala
     485  490  495
Ser Pro Ile Val Pro Ser Phe Asp Ser Met Ser Asn Glu Gly Ser Tyr Phe
     500  505  510
Phe Gly Asp Asn Ala Glu Glu Tyr Asp Asn
     515  520

<210> SEQ ID NO 8
<211> LENGTH: 1566
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<222> OTHER INFORMATION: NPeM2 Fusion Construct
<400> SEQUENCE: 8
atggcgctctc aaggaaacaa acgatctca acaacagstg agactgatg agaagcgaag

---continued---

actgccactg aatctcagc atccgtcaga aatagatgg gtggagattg acgattctac 120
atcataaagt gacagcaact caatactgat gatctacagc gacctctaga ccaaaccagc 180
ttacataag agagatgttg gctctctgct ttgagaga ggagaataaa atacagtaa 240
gaacataccta gtgcgaggag aagctctcaag aacactggag aaccttttaa cagagagaga 300
acaggaag ggtatcagga aactcatttc taatcaaaag aagaaattag gccaatcttg 360
cgccnaagta acaatggtagc ccaggtgctc gtgcctttga ctcaatctgt gatctggcat 420
tacaaattaga atgtgcaacc ttatcagag cccagagcttc ttgctgcaac ccgaggyggt 480
ccagaggtg ccgtctctag gaaagtttca acctctcctta gagggttctgg agccgaggg 540
gtctgcctca aaggagtttg acaagtygttg atggaatttg tccagactgt caaacgtgag 600
atcataagc gaaacctggc gagggyttag aatgacaggaa aacaaagatg tygctttaga 660
aanstgctca aaccttttca aggagacctc caaactctg ccaaaaccgg aggattctgag 720
casgtygagc agagccgaa cccagagaaag gatgtgcttg aagctctcag lccttcagcag 780
cgtcctggac ttaatctggc aggytctggt gcttcaactt cttcacgtgcc tggcttggtg 840
taattggagc cggataagcct tgggcagact ttttgaagggg aggataacagct tagctggg 900
atagacccct taagacctct ctaaaccaagc caaatgtaa gcaatcagag acaatttgag 960
acttcatgcc aacaagacatg acagtctggg agggcatcgcc atctgcggcc aattggagat 1020
tcaagagttc gaaagctggc agaggtctcoc ccaggagggc gttctggcaact 1080
agaagygtacc aatctggtc aatagaaaaat atggaggctc taggacacag taaacctggta 1140
otcggagacgc caaggtggag cataagggc agaaggtctgg gcacacccca ctaacaggg 1200
agtctggg gcaaatagc cataatactg atctctcgag taacagacact tctcctctttt 1260
agaagagcag ccagagctcc agctctccag aagatcagc aagggagac aagctggacat 1320
agaagacca ttcagtcatg gattggagg gcagagcccc aagagtcgctc ttccctgggg 1380
cgggtgctc tagagcccg agagagccag ccagagctgg ctcctttgac 1440
atggagctcgttaaagttctccttcagcgaagattacaagccttattagctaaagt 1500
aggctgtag gttgatcagag ccagagctgctt acgggtagg cagaggttga cagacggtca 1560
agatgt 1566

<210> SEQ ID NO: 9
<211> LENGTH: 522
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE: 
<222> OTHER INFORMATION: NPl92 Fusion Construct

<400> SEQUENCE: 9

Met Ala Ser Gln Gly Thr Lys Arg Ser Tyr Glu Gln Met Glu Thr Asp
1 5 10 15
Gly Glu Arg Gln Asn Ala Thr Glu Ile Arg Ala Ser Val Gly Lys Met
20 25 30
Ile Gly Gly Ile Gly Arg Phe Tyr Ile Glu Met Cys Thr Glu Leu Lys
35 40 45
Leu Ser Asp Tyr Glu Gly Arg Leu Ile Glu Asn Ser Leu Thr Ile Glu
50 55 60
Arg Met Val Leu Ser Ala Phe Asp Glu Arg Arg Asn Lys Tyr Leu Glu
65 70 75 80
Glu His Pro Ser Ala Gly Lys Aas Pro Lys Lys Thr Gly Gly Pro Ile
85  90  95
Tyr Arg Arg Val Asn Gly Lys Trp Met Arg Glu Leu Ile Leu Tyr Asp
100 105 110
Lys Glu Glu Ile Arg Ile Trp Arg Gln Ala Asn Asn Gly Aas Asp
115 120 125
Ala Thr Ala Gly Leu Thr His Met Met Ile Trp His Ser Asn Leu Asn
130 135 140
Asp Ala Thr Tyr Glu Arg Thr Arg Ala Leu Val Arg Thr Gly Met Arg
145 150 155 160
Pro Arg Met Cys Ser Leu Met Glu Gly Ser Thr Leu Pro Arg Arg Ser
165 170 175
Gly Ala Ala Gly Ala Ala Val Lys Gly Val Gly Thr Met Val Met Glu
180 185 190
Leu Val Arg Met Ile Lys Arg Gly Ala Asn Asp Arg Asn Phe Trp Arg
195 200 205
Gly Glu Asn Gly Arg Lys Thr Arg Ile Ala Tyr Glu Arg Met Cys Asn
210 215 220
Ile Leu Lys Gly Lys Phe Gin Thr Ala Ala Gin Lys Ala Met Met Asp
225 230 235 240
Gln Val Arg Glu Ser Arg Asn Pro Gly Asn Ala Glu Phe Glu Asp Leu
245 250 255
Thr Phe Leu Ala Arg Ser Ala Leu Ile Leu Arg Gly Ser Val Ala His
260 265 270
Lys Ser Cys Leu Pro Ala Cys Val Tyr Gly Pro Ala Val Ala Ser Gly
275 280 285
Tyr Asp Phe Glu Arg Glu Gly Tyr Ser Leu Val Gly Ile Asp Pro Phe
290 295 300
Arg Leu Leu Gln Asn Ser Gin Val Tyr Ser Leu Ile Arg Pro Asn Glu
305 310 315 320
Asn Pro Ala His Lys Ser Gin Leu Val Trp Met Ala Cys His Ser Ala
325 330 335
Ala Phe Glu Asp Leu Arg Val Leu Ser Phe Ile Lys Gly Thr Lys Val
340 345 350
Leu Pro Arg Gly Lys Leu Ser Thr Arg Gly Val Gln Ile Ala Ser Asn
355 360 365
Glu Asn Met Glu Thr Met Glu Ser Thr Leu Leu Arg Ser Arg
370 375 380
Tyr Thr Ala Ile Arg Thr Arg Ser Gly Asn Thr Asn Gln Glu Arg
385 390 395 400
Ala Ser Ala Gly Gin Ile Ser Ile Gin Pro Thr Phe Ser Val Gin Arg
405 410 415
Asn Leu Pro Phe Asp Arg Thr Val Met Ala Ala Phe Ser Gly Asn
420 425 430
Thr Glu Gly Thr Ser Asp Met Arg Thr Glu Ile Arg Met Met
435 440 445
Glu Ser Ala Arg Pro Glu Asp Val Ser Phe Gln Gly Arg Gly Val Phe
450 455 460
Glu Leu Ser Asp Glu Lys Ala Ala Ser Pro Ile Val Pro Ser Phe Asp
465 470 475 480
-continued

Met Ser Asn Glu Gly Ser Tyr Phe Phe Gly Asp Asn Ala Glu Glu Tyr 485 490 495

Asp Asn Met Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu 500 505 510

Trp Gly Cys Arg Cys Asn Gly Ser Ser Asp 515 520

<210> SEQ ID NO 10
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Linker Peptide

<400> SEQUENCE: 10
Gly Tyr Ala Thr Arg Ala 1 5

<210> SEQ ID NO 11
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Linker Peptide

<400> SEQUENCE: 11
Phe Gln Met Gly Glu Thr 1 5

<210> SEQ ID NO 12
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Linker Peptide

<400> SEQUENCE: 12
Phe Asp Arg Val Lys His Leu Lys 1 5

<210> SEQ ID NO 13
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Linker Peptide

<400> SEQUENCE: 13
Gly Arg Asn Thr Asn Gly Val Ile Thr 1 5

<210> SEQ ID NO 14
<211> LENGTH: 10
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Linker Peptide

<400> SEQUENCE: 14
Val Asn Glu Lys Thr Ile Pro Asp His Asp 1 5 10

<210> SEQ ID NO 15
atgcccaca tggattgtg cagtaaat accggacaa tgcattaaac accagaaga 60
tctgcttcgg gcaaccagga ccatcataa gcgcaaccag gctgtcctcg 120
cosaagcaca acgcacccgg aaccacacta tccggagc agcccccggc ggtggcacc 180
gatgctgag gaaagctgc gtttctgag cgggatgc tctatagcct 240
aatagttgag taatttggtgg tcatctcatc aaccagagga ggggattgct 300
atgcgttcag aagaaactggt tcctatcactgc tctacagcct ctctatagcct 360
gatgctgag acacggaggg tacacgcaat ccagcagggc ggtggcacc 420
ggaaaggg aacaagcc cacccaagca ggcggcgcttt ttctatttg gccctaagat 480
gataaaacc tcatctcag cccatatata attactttt tattatacagg gtgaggac 540
atgcgaag cccacatcgg ctcctctgt ctcttctgtc ttcagctgtg tttggagct 600
caccagc cttgcgggg aagacagaga caaggaggg aacgttcagc 660
gtcgtggcc ttcagctggt cctctctccc ggcgggcttt gcgcaccctt 720
atgctgctg gaggtcctgg ttggagcagc tttcgtttg ggcgggcttt gcgcaccctt 780
atgcgttcag cccacatcgg ctcctctgt ctcttctgtc ttcagctgtg tttggagct 840
gtaaaacc tcatctcag cccatatata attactttt tattatacagg gtgaggac 900
atgcgaag cccacatcgg ctcctctgt ctcttctgtc ttcagctgtg tttggagct 960
atgcgttcag cccacatcgg ctcctctgt ctcttctgtc ttcagctgtg tttggagct 1020
aatggtgtg aagaaactggt tcctatcactgc tctacagcct ctctatagcct 1080
aatggtgtg aagaaactggt tcctatcactgc tctacagcct ctctatagcct 1140
aatggtgtg aagaaactggt tcctatcactgc tctacagcct ctctatagcct 1200
aatggtgtg aagaaactggt tcctatcactgc tctacagcct ctctatagcct 1260
aatggtgtg aagaaactggt tcctatcactgc tctacagcct ctctatagcct 1320
aatggtgtg aagaaactggt tcctatcactgc tctacagcct ctctatagcct 1380
aatggtgtg aagaaactggt tcctatcactgc tctacagcct ctctatagcct 1440
aatggtgtg aagaaactggt tcctatcactgc tctacagcct ctctatagcct 1500
aatggtgtg aagaaactggt tcctatcactgc tctacagcct ctctatagcct 1560
aatggtgtg aagaaactggt tcctatcactgc tctacagcct ctctatagcct 1620
aatggtgtg aagaaactggt tcctatcactgc tctacagcct ctctatagcct 1680

Met Ser Aan Met Aep Ile Aep Ser Ile Aan Thr Gly Thr Ile Aep Lys
1 5 10 15

<210> SEQ ID NO 16
<211> LENGTH: 560
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Influenza B Virus
<400> SEQUENCE: 16

The sequence ends here, indicated by the `---continued` line.
| Thr Pro Glu Glu Leu Thr Pro Gly Thr Ser Gly Ala Thr Arg Pro Ile |
|------------------|------------------|------------------|
| 20               | 25               | 30               |
| Ile Lys Pro Ala Thr Ala Pro Pro Ser Asn Lys Arg Thr Arg Asn |
| 35               | 40               | 45               |
| Pro Ser Pro Glu Arg Thr Thr Ser Ser Glu Thr Asp Ile Gly Arg |
| 50               | 55               | 60               |
| Lys Ile Gln Lys Lys Gln Thr Pro Thr Glu Ile Lys Lys Ser Val Tyr |
| 65               | 70               | 75               | 80               |
| Lys Met Val Val Lys Leu Gly Glu Phe Tyr Asn Gln Met Met Val Lys |
| 85               | 90               | 95               |
| Ala Gly Leu Asn Asp Asp Met Glu Arg Asn Leu Ile Gln Asn Ala Gln |
| 100              | 105              | 110              |
| Ala Val Glu Arg Ile Leu Leu Ala Ala Thr Asp Asp Lys Thr Glu |
| 115              | 120              | 125              |
| Tyr Gln Lys Arg Asn Ala Arg Asp Val Lys Glu Gly Lys Glu Glu |
| 130              | 135              | 140              |
| Ile Asp His Asn Lys Thr Gly Gly Thr Phe Tyr Lys Met Val Arg Asp |
| 145              | 150              | 155              | 160              |
| Asp Lys Thr Ile Tyr Phe Ser Pro Ile Lys Ile Thr Phe Leu Lys Glu |
| 165              | 170              | 175              |
| Glu Val Lys Thr Met Tyr Lys Thr Thr Met Gly Ser Arg Gly Phe Ser |
| 180              | 185              | 190              |
| Gly Leu Asn His Ile Met Ile Gly His Ser Gln Met Asn Asp Val Cys |
| 195              | 200              | 205              |
| Phe Glu Arg Ser Lys Gly Leu Lys Val Gly Leu Asp Pro Ser Leu |
| 210              | 215              | 220              |
| Ile Ser Thr Phe Ala Gly Ser Thr Leu Pro Arg Arg Ser Gly Thr Thr |
| 225              | 230              | 235              | 240              |
| Gly Val Ala Ile Lys Gly Gly Gly Thr Leu Val Asp Glu Ala Ile Arg |
| 245              | 250              | 255              |
| Phe Ile Gly Arg Ala Met Ala Asp Arg Gly Leu Leu Arg Asp Ile Lys |
| 260              | 265              | 270              |
| Ala Lys Thr Ala Tyr Glu Lys Ile Leu Leu Asn Leu Lys Asn Lys Cys |
| 275              | 280              | 285              |
| Ser Ala Pro Gln Gln Lys Ala Leu Val Asp Gln Val Ile Gly Ser Arg |
| 290              | 295              | 300              |
| Asn Pro Gly Ile Ala Asp Ile Glu Asp Leu Thr Leu Ala Arg Ser |
| 305              | 310              | 315              | 320              |
| Met Val Val Arg Pro Ser Val Ala Ser Lys Val Leu Pro Ile |
| 325              | 330              | 335              |
| Ser Ile Tyr Ala Lys Ile Pro Gln Leu Gly Phe Asn Thr Glu Gly Tyr |
| 340              | 345              | 350              |
| Ser Met Val Gly Tyr Glu Ala Met Ala Leu Tyr Asn Met Ala Thr Pro |
| 350              | 360              | 365              |
| Val Ser Ile Leu Arg Met Gly Asp Ala Lys Asp Lys Ser Gln Leu |
| 370              | 375              | 380              |
| Phe Phe Met Ser Cys Phe Gly Ala Ala Tyr Glu Asp Leu Arg Val Leu |
| 380              | 395              | 400              |
| Ser Ala Leu Thr Gly Thr Glu Phe Lys Pro Arg Ser Ala Leu Lys Cys |
| 395              | 410              | 415              |
-continued

Lys Gly Phe Ala Val Pro Ala Lys Glu Glu Val Gly Met Gly Ala
420 425 430
Ala Leu Met Ser Ile Lys Leu Glu Phe Trp Ala Pro Met Thr Arg Ser
435 440 445
Gly Gly Asn Glu Val Ser Gly Glu Gly Ser Gly Glu Ile Ser Cys
450 455 460
Ser Pro Val Phe Ala Val Glu Arg Pro Ile Ala Leu Ser Lys Gln Ala
465 470 475 480
Val Arg Arg Met Leu Ser Met Asn Val Glu Gly Arg Asp Ala Asp Val
485 490 495
Lys Gly Asn Leu Leu Lys Met Met Asp Ser Met Ala Lys Asp Thr
500 505 510
Ser Gly Asn Ala Phe Ile Gly Lys Met Phe Glu Ile Ser Asp Lys
515 520 525
Aan Lys Val Aan Pro Ile Glu Ile Pro Ile Lys Gln Thr Ile Pro Aan
530 535 540
Phe Phe Gly Arg Asp Thr Ala Glu Asp Tyr Asp Leu Asp Tyr
545 550 555 560

<210> SEQUENCE
<211> LENGTH: 1220
<212> ORGANISM: Influenza A virus
<400> SEQUENCE: 17

atgagcagca gactacgtgt ctttgtatgt gcctttgcag ctaacaatgc agacacoata 60
tgtataagg acaacagcga taacctacacg gacacttggtg acaacactct cgaaagagat 120
gtgtcggcgc cccatcagctc tcacccgtgc gacacccgcc acaacggaga actatatccga 180
atccagggag caccagcagc gttccagtcag gacatcagc gcacatctctg atggtggtt 240
aacacaagcg gttcattga gacaccgctggc gacatttcgcc acaacagcccg acaacgagag 300
aacagagcgc gttcgcagtc atccacccgct gccagctggtg accaggtggtg gacagtccgc 360
ctccacgctgc acctgacgtg caccgggttc ctcctgagag gaccaggtgc acacagaggc 420
cgccccggtt ccctctcccag cccgcacgtg ctcctgggat gacagtttggtg ttttacggag 480
ttcaccggtccttgtgcgt gttgacacag gggagtctcgt ccatccaccag cttgagcccgg 540
ttttacggag cttgtgacag gtaaatggctg acatcctgctg acatccctgctg ttaaagccct 600
acacccctgctg cttaacagctg taaaatgctg gatctagctg cccgccctgctg ttttaccgag 660
tccacccctgctg cttaagctgctg gcatctagctg cccgcctgctg ttttacctgctg 720
gttccggcgt ctttacctgctg gttaccgtctg gtttacctgcc gtttacctgcc gtttacctgcc 780
gcaacggagcg gactctagctg cgctctgccg gcttctgccg gtttacctgctg 840
ggtactgctg ctttacctgcc gcctagctg cttaagctgctg ggttacctgcc gtttacctgctg 900
ggtactgctg ctttacctgcc gcctagctg cttaagctgctg ggttacctgcc gtttacctgctg 960
ccttaagctg ctgtgctgctg gtttacctgcc gtttacctgcc gtttacctgctg 1020
atcctagctg ctttacctgcc gcctagctg cttaagctgctg ggttacctgcc gtttacctgctg 1080
atcctagctg ctttacctgcc gcctagctg cttaagctgctg ggttacctgcc gtttacctgctg 1140
gtttacctgcc gtttacctgcc gcctagctg cttaagctgctg ggttacctgcc gtttacctgctg 1200

1220
<210> SEQ ID NO 18
<211> LENGTH: 406
<212> TYPE: PRT
<213> ORGANISM: Influenza A virus

<400> SEQUENCE: 18

Met Glu Ala Arg Leu Leu Val Leu Leu Cys Ala Phe Ala Ala Thr Asn
1  5  10
Ala Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Ser Thr Asp Thr
20  25  30
Val Asp Thr Val Leu Glu Asn Val Thr Val Thr His Ser Val Asn
35  40  45
Leu Leu Glu Asp Ser His Asn Gly Lys Leu Cys Lys Leu Lys Gly Ile
50  55  60
Ala Pro Leu Gin Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu Gly
65  70  75  80
Asn Pro Glu Cys Asp Leu Leu Thr Ala Ser Ser Thr Ser Tyr Ile
85  90  95
Val Glu Thr Ser Asn Ser Glu Gly Thr Cys Tyr Pro Gly Asp Phe
100 105 110
Ile Asp Tyr Glu Glu Leu Arg Glu Gin Leu Ser Ser Val Ser Ser Phe
115 120 125
Glu Lys Phe Glu Ile Phe Pro Lys Thr Ser Ser Thr Pro Asn His Glu
130 135 140
Thr Thr Lys Gly Val Thr Ala Ala Cys Ser Tyr Ala Gly Ala Ser Ser
145 150 155 160
Phe Tyr Arg Asn Leu Leu Trp Leu Thr Lys Gly Ser Ser Tyr Pro
165 170 175
Lys Leu Ser Ser Tyr Val Asn Lys Gly Lys Glu Val Leu Val
180 185 190
Leu Trp Gly Val His Pro Pro Thr Gly Thr Asp Gin Gin Ser Leu
195 200 205
Tyr Gin Asn Ala Asp Ala Tyr Val Ser Val Ser Gly Ser Ser Tyr Asn
210 215 220
Arg Arg Phe Thr Pro Glu Ile Ala Ala Arg Pro Lys Val Arg Gly Gin
225 230 235 240
Ala Gly Arg Met Asn Tyr Tyr Thr Thr Leu Leu Leu Pro Gly Asp Thr
245 250 255
Ile Thr Phe Glu Ala Thr Gly Asn Thr Leu Ile Ala Pro Thr Tyr Ala Phe
260 265 270
Ala Leu Asn Arg Gly Ser Gly Ser Gly Ile Thr Ser Ser Asp Ala Pro
275 280 285
Val His Asp Cys Asn Thr Lys Cys Gin Thr Pro His Gly Ala Ile Asn
290 295 300
Ser Ser Leu Pro Phe Gin Asn Ile His Pro Val Thr Ile Gly Glu Cys
305 310 315 320
Pro Lys Tyr Val Arg Ser Thr Lys Leu Arg Met Ala Thr Gly Leu Arg
325 330 335
Aan Ile Pro Ser Ile Gin Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
340 345 350
Phe Ile Glu Gly Gly Thr Thr Gly Met Ile Asp Gly Gly Thr Tyr Gly Tyr
<table>
<thead>
<tr>
<th>355</th>
<th>360</th>
<th>365</th>
</tr>
</thead>
<tbody>
<tr>
<td>His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Aep Gln Lys Ser</td>
<td></td>
<td></td>
</tr>
<tr>
<td>370</td>
<td>375</td>
<td>380</td>
</tr>
<tr>
<td>Thr Gln Asn Ala Ile Aep Gly Ile Thr Asn Lys Val Asn Ser Val Ile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>385</td>
<td>390</td>
<td>395</td>
</tr>
<tr>
<td>Glu Lys Met Asn Thr Gln</td>
<td></td>
<td>405</td>
</tr>
</tbody>
</table>

<210> SEQ ID NO: 19
<211> LENGTH: 1741
<212> ORGANISM: Influenza A virus

<400> SEQUENCE: 19

cyttgcaaat ggggacata gttgtccctt tggcaacagc caagtctgggta asagctgact 60
gatatggtc tggctaacct cgaacaact cgcacagga gttgacaca atagdynaa 120
agaaatgcct tacgtcagc gcccagaca tacagaaag gacacacaac gggagctct 180
goyagctaa cagagttcag aacccacatt tggagattc agagagctt 240
tggagaccc tattgactcg gtaaccacag atgtctgcca atgctgttac atagdgag 300
gggagcctgagc ccgataactc cagaggtttc cagacgacta gagaactggag 360
arcactatt cacagacta aaccactttg agaasactta gctactacccc aacaggtctt 420
gtcatactcg ctagggtggg gcctagcact cctactctgg ctgagagctt 480
cacacattc cagacgtaag cagctttcag acacaggtgg gacactttc 540
ggagctcctg ctaacactag ccacagactc tttggtact gttggttcatt cccactctga 600
atgtgtggcg agggactgga aagctcttgat cccctcttcc cccctctgaa 660
cactacatg ccccagacta acccctctag cagaggttaag gttagaactg 720
agagactgtc cagactgac cactactat ctcctgcttg gcctgagctg 780
aagagactgct cagaggttgc acacactgt cccactactg ccactactg 840
ccacactattt cagactgac cactactat ctcctgcttg gcctgagctg 900
tgtggttactg agactatgta aacaccctcc cccctctcct cccctctgag 960
gcccactata tggagtattc cactactac gttggactac gactacactc 1020
agagactgct cagaggttgc acacactgt cccactactg ccactactg 1080
ccacactattt cagactgac cactactat ctcctgcttg gcctgagctg 1140
gatggtgactc cactactac gttggactac gactacactc 1200
tgtggttactg agactatgta aacaccctcc cccctctcct cccctctgag 1260
actacaag ctgactactc aactaattc caccactttg gacagactgt 1320
gtgcctactg cactactac gttggactac gactacactc 1380
actacaag ctgactactc aactaattc caccactttg gacagactgt 1440
acacggtactc cactactac gttggactac gactacactc 1500
tagaagagc cagaggttgc acacactgt cccactactg ccactactg 1560
ccacactattt cagactgac cactactat ctcctgcttg gcctgagctg 1620
tgtggttactg agactatgta aacaccctcc cccctctcct cccctctgag 1680
cactactac gttggactac gactacactc 1740
a

<213> ORGANISM: Influenza A virus

<400> SEQUENCE: 20

 Met Glu Lys Ile Val Leu Leu Leu Ala Thr Val Ser Leu Val Lys Ser
  1    5    10   15
 Asp Gin Ile Cys Ile Tyr His Ala Asn Ser Thr Glu Gin Val
  20   25    30
 Asp Thr Ile Met Glu Asn Val Thr Val Thr His Ala Gin Asp Ile
  35   40    45
 Leu Glu Arg Thr His Asn Gly Lys Leu Cys Asp Leu Asn Gly Val Lys
  50   55    60
 Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
  65    70    75    80
 Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
  85   90    95
 Glu Lys Ala Ser Pro Ala Asn Asp Leu Cys Tyr Pro Gly Asn Phe Asn
 100   105   110
 Asp Tyr Glu Glu Leu His Leu Leu Ser Arg Ile Asn His Phe Glu
 115   120   125
 Lys Ile Gin Ile Ile Pro Lys Ser Ser Ser Thr Val Gly Arg Ser Ser Phe
 130   135   140
 Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Arg Ser Ser Phe Phe
 145   150   155   160
 Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Ala Tyr Pro Thr Ile
 165   170   175
 Lys Arg Ser Tyr Asn Asn Thr Asn Gin Glu Asp Leu Leu Val Leu Trp
 180   185   190
 Gly Ile His His Pro Asn Asp Ala Glu Gin Thr Lys Leu Tyr Gin
 195   200   205
 Aasn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gin Arg
 210   215   220
 Leu Val Pro Glu Ile Ala Thr Arg Pro Lys Val Asn Gly Gin Ser Gly
 225   230   235   240
 Arg Met Glu Phe Phe Thr Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
 245   250   255
 Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
 260   265   270
 Val Lys Lys Gly Ser Phe Thr Ile Met Lys Ser Glu Leu Glu Tyr Gly
 275   280   285
 Asn Cys Asn Thr Lys Cys Gin Thr Pro Met Gly Ala Ile Asn Ser Ser
 290   295    300
 Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Gin Thr Cys Pro Lys
 305   310   315   320
 Tyr Val Lys Ser Aasn Arg Leu Val Ile Thr Gly Leu Arg Asn Thr
 325   330   335
 Pro Gin Arg Glu Arg Arg Arg Lys Arg Gly Leu Phe Gly Ala Ile
 340   345   350
 Ala Gly Phe Ile Glu Gly Gly Trp Gin Gly Met Val Asp Gly Trp Tyr
355 360 365
 Gly Tyr His His Ser Aen Glu Gin Gly Ser Gly Tyr Ala Ala Asp Lys
370 375 380
 Glu Ser Thr Gin Lys Ala Ile Asp Gin Val Thr Asn Val Asn Ser
385 390 395 400
 Ile Ile Asn Lys Met Asn Thr Gin Phe Glu Ala Val Gly Arg Glu Phe
405 410 415
 Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Met Glu Asp
420 425 430
 Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Val Leu Met
435 440 445
 Glu Asn Glu Arg Thr Leu Aen Phe His Asp Ser Aen Val Lys Asn Leu
450 455 460
 Tyr Asp Lys Val Arg Leu Gin Leu Arg Asp Asn Ala Lys Glu Leu Gly
465 470 475 480
 Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu
485 490 495
 Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Gin Tyr Ser Glu Glu Ala
500 505 510
 Arg Leu Asn Arg Glu Ile Ser Gly Val Lys Leu Glu Ser Met Gly
515 520 525
 Thr Tyr Gin Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
530 535 540
 Leu Ala Ile Met Val Ala Gly Leu Ser Leu Trp Met Cys Ser Ser Gly
545 550 555 560
 Ser Leu Gin Cys Arg Ile Cys Ile
565
<210> SEQ ID NO: 21
<211> LENGTH: 1714
<212> TYPE: DNA
<213> ORGANISM: Influenza A virus
<400> SEQUENCE: 21
gcaaaagcac ggacacagtct taactgcgaa atagcagcata atctcaacta taacctacta 60
actaggtaa cgcaacagcta atcgagtaa atctgacta ggcacacagc caacaaccto 120
cacgaacct tgtgacacgc taaccaagac caactttctgt tgaacacagt ccaagaatt 180
gctccacac gacgctagtg gatgcgctgtg tgcacacagc ctgggacact cccctactct 240
agacacctc acttgtgaag gcctatctca tgcacacctt ctctggtgccc tctggtgttgg 300
aggaagagct tctgctcagc tctgcagaa atcgacagc gtagttgga cgttgttaccc 360
tgcttttctg gaaaacctcag aagcaacagcg gacactttttt gttccoycta gttctacaca 420
agacactac aacactccag acctggatag gatactgctg tacagcctgc gacgacagagc 480
tagttcatctatatc gatgctgaag atgctgcact cacaagaccg gtttttaccc 540
tgctcaac acgcaactac cacaactcag gcacagacag atcttctctgtg tggggtgcat 600
atgctcaca cccctactgt ccagcactc ccatgacact asacttgtct atcagcactc acacactac 660
aacagcata aggcaaatg tgaagagagc cttcaacac gttgatgccc caagccctct 720
tgtcagctgt ctcagcagac gattgtgct tatttggtcg gctactaaac cagccacac 780
--continued

attcgagta cgctccagtg ggaatctcaat tcgcctaga gtagcagtctg ttcttgctgg 840
agggcgccc ggaggacccc ggagagctga tttttaaggt ggttaatttggt tagcagtctga 900
tcgcctgaa aaggygttagct taacatactct atggccattc ccaaatctct cgaatata 960
atggagttg tcgcctcct atggattgaat taattgccttc aacagttgagag tgcgggtgag 1020
gaagtgcgc ttcagagctaa gtaggaccc attgagccgct atagcctcct attcagaagg 1080
aggttgccca ggaagtagctc cttgctgttgct ttcattttca cattccaaatt atcaagggct 1140
tggttagctc gcaaggtggcg atccacccctt aaggccattt gataaaatata cattcaggt 1200
gtaattatata gcggcagatc gtagcagttct aacataattag atttgctagcgtg 1260
gttgcctct aaagccatt cgcctgatct gcctgacttatt agagcctatttattg 1320
ggcctctcat cagcttatgt ggtaagttcct ccaaaactgcctt atgcgattaaa 1380
tcgctgctttgtc atcagctcgct aatgctgctctg aagcactagctaa 1440
agatgagaaat gcctggtcttgg atcgttcctat gatactttcct cgcctggttcgtg 1500
tgtagcagggg acaatattgag gtagcagttctaatagcggcctgtaacaggtt 1560
aagctgaggg attgaggggt tagccacacttg ccctctttga cctctctctgct 1620
tcgctgctttc acaagctgctttc tgcctgctttgct ggtggtgtgtc ttcattttca cattccaaatt 1680
cagttgctgct ctcagcattc caagttgagag tgcgggtgag 1714

<210> SEQ ID NO 22
<211> LENGTH: 560
<212> TYPE: PRT
<213> ORGANISM: Influenza A virus
<400> SEQUENCE: 22

Met Glu Thr Ile Ser Leu Thr Ile Leu Val Val Thr Ala Ser
1  5  10  15
Asn Ala Asp Lys Ile Cys Ile Gly His Glu Ser Thr Asn Ser Thr Glu
20 25 30
Thr Val Asp Thr Leu Thr Glu Thr Asn Val Pro Val Thr His Ala Lys
35 40 45
Glu Leu Leu His Thr Glu His Asn Gly Met Leu Cys Ala Thr Ser Leu
50 55 60
Gly His Pro Leu Ile Leu Asp Thr Cys Thr Ile Glu Gly Leu Val Tyr
65 70 75 80
Gly Asn Ser Ser Cys Asp Leu Leu Gly Gly Arg Glu Trp Ser Tyr
85 90 95
Ile Val Glu Arg Ser Ser Ala Val Asn Gly Thr Cys Tyr Pro Gly Asn
100 105 110
Val Glu Asn Leu Glu Glu Leu Arg Thr Leu Phe Ser Ser Ala Ser Ser
115 120 125
Tyr Gin Arg Ile Glu Ile Phe Pro Asp Thr Thr Thr Asn Val Thr Tyr
130 135 140
Thr Gly Thr Ser Arg Ala Cys Ser Gly Ser Phe Tyr Arg Ser Met Arg
145 150 155 160
Trp Leu Thr Gin Lys Ser Gly Phe Tyr Pro Val Glu Asp Ala Gin Tyr
165 170 175
Thr Asn Asn Arg Gly Lys Ser Ile Leu Phe Val Thr Gly Ile His His
180 185 190
Pro Thr Tyr Thr Glu Gln Thr Asn Leu Tyr Ile Arg Asp Thr
195  200  205
Thr Thr Ser Val Thr Glu Asp Leu Asn Arg Thr Phe Lys Pro Val
210  215  220
Ile Gly Pro Arg Pro Leu Val Asn Gly Leu Gln Gly Arg Ile Asp Tyr
225  230  235  240
Tyr Trp Ser Val Leu Lys Pro Gly Gln Thr Leu Arg Val Arg Ser Asn
245  250  255
Gly Asn Leu Ile Ala Pro Trp Tyr Gly His Val Leu Ser Gly Gly Ser
260  265  270
His Gly Arg Ile Leu Lys Thr Asp Leu Lys Gly Gly Asn Cys Val Val
275  280  285
Gln Cys Gln Thr Glu Gly Leu Asn Ser Thr Leu Pro Phe His
290  295  300
Asn Ile Ser Lys Tyr Ala Phe Gly Thr Cys Pro Lys Tyr Val Arg Val
300  310  315  320
Asn Ser Leu Lys Leu Asn Val Gly Leu Arg Asn Val Pro Ala Arg Ser
325  330  335
Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp
340  345  350
Pro Gly Leu Val Ala Gly Trp Tyr Gly Phe Glu His Ser Asn Asp Gln
355  360  365
Gly Val Gly Met Ala Ala Asp Arg Asp Ser Thr Gln Lys Ala Ile Asp
370  375  380
Lys Ile Thr Ser Lys Val Asn Asn Ile Val Asp Lys Met Asn Lys Gln
385  390  395  400
Tyr Glu Ile Ile Asp His Glu Phe Ser Glu Val Glu Thr Arg Leu Asn
400  405  410  415
Met Ile Asn Asn Lys Ile Asp Asp Gln Ile Gln Asp Val Trp Ala Tyr
420  425  430
Asn Ala Glu Leu Leu Val Leu Glu Asn Gln Lys Thr Leu Asp Glu
435  440  445
His Asp Ala Asn Val Asn Leu Tyr Asn Lys Val Arg Ala Leu
450  455  460
Gly Ser Asn Ala Met Glu Asp Gly Lys Cys Phe Glu Leu Tyr His
465  470  475  480
Lys Cys Asp Asp Gln Cys Met Glu Thr Ile Arg Asn Gly Thr Tyr Asn
485  490  495
Arg Arg Lys Tyr Arg Glu Glu Ser Arg Leu Glu Arg Gln Lys Ile Glu
500  505  510
Glu Val Lys Leu Glu Ser Glu Gly Thr Tyr Lys Ile Leu Thr Ile Tyr
515  520  525
Ser Thr Val Ala Ser Ser Leu Val Leu Ala Met Gly Phe Ala Ala Phe
530  535  540
Leu Phe Trp Ala Met Ser Asn Gly Ser Cys Arg Cys Asn Ile Cys Ile
545  550  555  560

<210> SEQ ID NO 23
<211> LENGTH: 1494
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
---continued---

**OTHER INFORMATION:** Human Codon Optimized Influenza A Virus H1N1 Nucleoprotein

**SEQUENCE:** 23

```
atgccccctc aagggcgcaaa gggctcctac gcagcagtag gcacgcaggg aagacggcag 60
aacggagc agatccgaco ctgggagga aagcagtag gcagcaggg tagttctac 120
attcagattg gcagcaggtc taagcttgaag gtttagaag gcagcagtag acagaatctca 180
ctcaccctctc aagacgcagtc gccagcggtc atagcagcaggg ggaagcaggg 240
gagcagcagtc cacccggtaag aacagcagtc gcccagcaggg agcagcaggg 300
acccggagtc acagcaggtc gctagctgc acgggggtc aagcagcaggg gggggtgcag 360
agggggtctc aacagcagtg atagcagcaggg gggggtgcag aagcagcaggg 420
tctacattc aagcagcagtc cacccggtaag aacagcagtc gcccagcaggg 480
acccggagtc acagcaggtc gctagctgc acgggggtc aagcagcaggg gggggtgcag 540
gccagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 600
atcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 660
cagcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 720
cagcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 780
atcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 840
tctacattc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 900
atcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 960
acccggagtc acagcaggtc gctagctgc acgggggtc aagcagcaggg gggggtgcag 1020
cagcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1080
cagcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1140
tctacattc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1200
gccagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1260
ggccagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1320
gccagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1380
gccagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1440
gccagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1494
```

**SEQ ID NO:** 24
**LENGTH:** 1497
**TYPE:** DNA
**FEATURE:**

**OTHER INFORMATION:** Human Codon Optimized Influenza A Virus H1N1 Nucleoprotein

**SEQUENCE:** 24

```
atgccccctc aagggcgcaaa gggctcctac gcagcagtag gcacgcaggg aagacggcag 60
aacggagc agatccgaco ctgggagga aagcagtag gcagcaggg tagttctac 120
attcagattg gcagcaggtc taagcttgaag gtttagaag gcagcagtag acagaatctca 180
ctcaccctctc aagacgcagtc gccagcggtc atagcagcaggg ggaagcaggg 240
gagcagcagtc cacccggtaag aacagcagtc gcccagcaggg agcagcaggg 300
acccggagtc acagcaggtc gctagctgc acgggggtc aagcagcaggg gggggtgcag 360
agggggtctc aacagcagtg atagcagcaggg gggggtgcag aagcagcaggg 420
tctacattc aagcagcagtc cacccggtaag aacagcagtc gcccagcaggg 480
acccggagtc acagcaggtc gctagctgc acgggggtc aagcagcaggg gggggtgcag 540
gccagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 600
atcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 660
cagcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 720
cagcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 780
atcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 840
tctacattc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 900
atcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 960
acccggagtc acagcaggtc gctagctgc acgggggtc aagcagcaggg gggggtgcag 1020
cagcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1080
cagcagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1140
tctacattc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1200
gccagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1260
ggccagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1320
gccagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1380
gccagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1440
gccagcagtc aagcagcaggg aacatagcag atagcagcaggg gggggtgcag 1494
```
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>acgcagaagc gatagacgg gtctaatcgg ttcgcaaaag aggagatcag gcgcccttg</td>
<td>360</td>
</tr>
<tr>
<td>cgccgagcc aacagcccg caacccgaca gcgcggagc caaagatcgc gatctgcagc</td>
<td>420</td>
</tr>
<tr>
<td>agaagactg caagagccat ctcaagcgg gacccggccc tgggtggagc ccgggatgcc</td>
<td>480</td>
</tr>
<tr>
<td>cccccgatg gcagctctgt gcagggccgc aacctgcaccc gccggagcgg cgccgctggc</td>
<td>540</td>
</tr>
<tr>
<td>gcgcggcga aggagcctgg cacccatgcag aagagccttc ccagccagtg gcagcgggcc</td>
<td>600</td>
</tr>
<tr>
<td>acgccacgc ggacgttgc gcggggygg aacggcgag agagccctgc gctctgtgag</td>
<td>660</td>
</tr>
<tr>
<td>cggtgtgca acatcctgcag ggccagtttc caagccggcc cccagagccc cctgttggagc</td>
<td>720</td>
</tr>
<tr>
<td>cgctgtcgg aggcgcggg cccggcggcc gcggtgcttc aggagccagc gtaagctgcac</td>
<td>780</td>
</tr>
<tr>
<td>cgagccgccc tggctgtgc ggagcggct gcggcacaag gcctgttcgc gcgtctgtgg</td>
<td>840</td>
</tr>
<tr>
<td>tcagccgccc cctggtggcg cctgttgagg ttcgacgggg aggagctacg cctgtggtgg</td>
<td>900</td>
</tr>
<tr>
<td>atgcacaccc ttctggtgct gcagacagca caggtgtcag gcgtctgtgg gcggcaacag</td>
<td>960</td>
</tr>
<tr>
<td>aacccggcc aacagagcag gcggctctgc atgccctggcac cggagccggcc ctgggtgaggc</td>
<td>1020</td>
</tr>
<tr>
<td>ctggcgggtg tgcgttcttc caagccggcc aagttgcttc cccggcggcc gcctgacccc</td>
<td>1080</td>
</tr>
<tr>
<td>cgggcggcag gatagccagc caaaagagac aagagacagc gccaagagag ccaagtggag</td>
<td>1140</td>
</tr>
<tr>
<td>ctggcgggag gcgtcctggc ccttcgagag gcgcgggag cccaggactg cggagcttgg</td>
<td>1200</td>
</tr>
<tr>
<td>gcgcgcgggg gcgcgtcttc gcctggccgg acatcctgag gcagccgggg ggtgcttcggc</td>
<td>1260</td>
</tr>
<tr>
<td>gcgcgccagc ccgtctgggc ccctctgag gcgcggcagc ccggagcgtc gcctggagcg</td>
<td>1320</td>
</tr>
<tr>
<td>cgccggccag tcagctccgt gcagctcagc gcggcggccag cggagcgtc gcctggagcg</td>
<td>1380</td>
</tr>
<tr>
<td>gcggcgggcc tcctgtcgtc gcagcaggag gcgcgcggcg ccctctgag gcctggagcg</td>
<td>1440</td>
</tr>
<tr>
<td>atgcacaccc aggagagcag tcctctgggc gcgaagcgcc gcagatcag ccaagtga</td>
<td>1497</td>
</tr>
</tbody>
</table>

SEQ ID NO: 25
LENGTH: 1497
TYPE: DNA
ORGANISM: Artificial sequence
FEATURES:
OTHER INFORMATION: Human Codon Optimized Influenza A Virus H1N1 Nucleoprotein
SEQUENCE: 25

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>atggcttcac aggagccgg gcggagttaa gcggagatgg gcagcagatgg gcgcggacag</td>
<td>60</td>
</tr>
<tr>
<td>aacgccgagc gatagacgg gcctagctgg aggagatcag gcgcccttg</td>
<td>120</td>
</tr>
<tr>
<td>atgcacaccc aggagatcag gcagctctgg gcgcgcgggg gcctggttctc ccagctggagc</td>
<td>180</td>
</tr>
<tr>
<td>ctgcgggtgc gcctcgtgg ccagggccgc ggccttcgcg ccctgcagtc gcgcggccag</td>
<td>240</td>
</tr>
<tr>
<td>gcgcgcggtg gcgcggccag gcggctgcag gcgcggagcc gcccataa tgcagagtt</td>
<td>300</td>
</tr>
<tr>
<td>agggtgcag gcggagagc ttgcctctgc gcagctccttc ccagctggagc ccagtgtgag</td>
<td>360</td>
</tr>
<tr>
<td>agggcgcc aacagcccg caacccgaca gcgcggagc caaagatcgc gatctgcagc</td>
<td>420</td>
</tr>
<tr>
<td>aggagactg caagagcctgc ctcaagcggg aacagagcag cctggtggagc ccgggatgcc</td>
<td>480</td>
</tr>
<tr>
<td>ccctaaaag gatagacgg gcctagctgg aggagatcag gcgcccttg</td>
<td>540</td>
</tr>
<tr>
<td>gcgcgcgggg gcctcgtgg gcgcgggggc gcagctccttc ccagctggagc ccagtgtgag</td>
<td>600</td>
</tr>
<tr>
<td>gcgcgcgcg gcctcgtgg gcgcgggggc gcagctccttc ccagctggagc ccagtgtgag</td>
<td>660</td>
</tr>
<tr>
<td>gcgcgcgcg gcctcgtgg gcgcgggggc gcagctccttc ccagctggagc ccagtgtgag</td>
<td>720</td>
</tr>
<tr>
<td>gcgcgcgggg gcctcgtgg gcgcgggggc gcagctccttc ccagctggagc ccagtgtgag</td>
<td>780</td>
</tr>
</tbody>
</table>
agatctggcc tgatatacgg ggctctgta gcocacaaagc ctgtgctgcc gctctgtgct 840
tacgcgccc ccgctgccg ccgctgac cgctgagcg gggttaactc ccttgctgac 900
tatgaccttt ctattcttg ctacagactc ggtcgatcag aatgctgctg 960
aatgcggcggc tataagctgc gcggctggcg atggcctgcgg atgcgctgcgg 1020
tatgctgcttg tctttttat aagggcagac aagggcagac cccagggcaca gctctgactc 1080
agggctgcttg aatagcccgc cccagagacg ccgggagccg gctacgctgctcgtcag 1140
tatgctgcttg ggcgggctgc cccagagacg gcggcgcagc cccagagacg 1200
gccagcgcgg gcgacgacgc cccagagacg cccagagacg cccagagacg 1260
gattacactc atcagactgc cgctggtctc gcgacacgcc gcggcgcagc gcggcgcagc 1320
cagagacgc ctacagactc ggcgggctgc gcggcgcagc agggctgcttg ggcgggctgc 1380
agggctgcttg cagagacgc cccagacgac cccagacgac cccagagacg 1440
atgctgcttg ggcgggctgc cccagagacg gcggcgcagc gcggcgcagc 1497

<210> SEQ ID NO 26
<211> LENGTH: 756
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Human Codon Optimized Influenza A Virus M1 Protein
<400> SEQUENCE: 26

atgagctggc taaacgaagtt gcgasagta ttcacatcgc ttaagctgcgc 60
gacccggaa gacagtggagt acagagatg agagatgtg gacggatgtg 120
gcagcagtg gcgcagcagtg cgaagcagtg cgaagcagtg cgaagcagtg 180
gcagcagtg gcgcagcagtg cgaagcagtg cgaagcagtg cgaagcagtg 240
gcagcagtg gcgcagcagtg cgaagcagtg cgaagcagtg cgaagcagtg 300
gcagcagtg gcgcagcagtg cgaagcagtg cgaagcagtg cgaagcagtg 360
gcagcagtg gcgcagcagtg cgaagcagtg cgaagcagtg cgaagcagtg 420
gcagcagtg gcgcagcagtg cgaagcagtg cgaagcagtg cgaagcagtg 480
gcagcagtg gcgcagcagtg cgaagcagtg cgaagcagtg cgaagcagtg 540
gcagcagtg gcgcagcagtg cgaagcagtg cgaagcagtg cgaagcagtg 600
gcagcagtg gcgcagcagtg cgaagcagtg cgaagcagtg cgaagcagtg 660
gcagcagtg gcgcagcagtg cgaagcagtg cgaagcagtg cgaagcagtg 720
gcagcagtg gcgcagcagtg cgaagcagtg cgaagcagtg cgaagcagtg 756
gtctgtgg agggttaga caccaagcgg acctagacg gaccaacag 180
ggatcaggt tccaactcac gatgaagcgg gatagggcag ggttctgctg 240
cgacacaag cagccgaacg gcaagaccag tggctctctg cagacaggg 300
aggaagcttg aagttgagat ccccttacac gggcgcggaa agatccaggt gattacagc 360
gcgcgcgcgg tggcacacgt cagagggctct atacacacaa ggacgggcag cggacacac 420
gggtgctg cggctctctg tggctcgacgc tggcgcggac cggccacagg 480
agccacacc gcgcggttgg caccacacac gccctgtctg ggccagcaga cagatgtgg 540
cgctgcgac ccccaacggcg ggcctcagag cagatggcct gcagccgggc gcagccgccc 600
gggacagaag ggtgctctct gggccggcag cagatgtgg cagatgtgg ggcctgccg 660
accccaacac ggcggagacag cggcctgacg cggggagtgg ggcctgccc ggcctgccc 720
cggcgcgggt ggtgctctct cgaacacag ggctgctctct tccaatg 756

<210> SEQ ID NO 28
<211> LENGTH: 756
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Human Codon Optimized Influenza A Virus M1 Protein
<400> SEQUENCE: 28
atgatcttg cggccggtg atggcagtag ctgccctcgt cgtggctttc gaccaacag 60
agacacacc gcgcggttgc gtcggtggag gttgctctct cgaacacag ggctgctct 120
agtctgttg agggttaga caccaagcgg acctagacg gaccaacag 180
ggatcaggt tccaactcac gatgaagcgg gatagggcag ggttctgctg 240
gggtgctg cggctctctg tggctcgacgc tggcgcggac cggccacagg 300
aggaagcttg aagttgagat ccccttacac gggcgcggaa agatccaggt gattacagc 360
gcgcgcgcgg tggcacacgt cagagggctct atacacacaa ggacgggcag cggacacac 420
gggtgctg cggctctctg tggctcgacgc tggcgcggac cggccacagg 480
agccacacc gcgcggttgg caccacacac gccctgtctg ggccagcaga cagatgtgg 540
gcgcgcgcgg tggcacacgt cagagggctct atacacacaa ggacgggcag cggacacac 600
rggacagaag ggtgctctct gggccggcag cagatgtgg cagatgtgg ggcctgccg 660
accccaacac ggcggagacag cggcctgacg cggggagtgg ggcctgccc ggcctgccc 720
cggcgcgggt ggtgctctct cgaacacag ggctgctctct tccaatg 756

<210> SEQ ID NO 29
<211> LENGTH: 294
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Human Codon Optimized Influenza A Virus M2 Protein
<400> SEQUENCE: 29
atgatcttt tcaacgggg ctgacacaag atatgtggag cgtgccggag 60
gggtgctg atccctctct cattgaacag atatgtggag cgtgccggag 120
tgtgtttt cttaccactct ctaatctgacag atatgtggag cgtgccggag 180
---continued---

gaggggcctt ctacggaagg agtgccaaag tctatgaggg agaatatcgt aaaaagacag 240
cagagctgct tggagtctgtc cgtagctctat tttcgacat ttagagctga gtas 294

<210> SEQ ID NO: 30
<211> LENGTH: 294
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Human Codon Optimized Influenza A Virus M2 Protein

<400> SEQUENCE: 30
atgagctgct tgcagcaggt ggaacacccc atcggaacag agtgggtgcct ccgtygacoac 60
ggacagcag agccctcctgg ctagcgcgcc aacctcctcg gctgccctcg cctgcacgcc 120
tggtacctgct atccacctgg ctttaagatc gtttaagta tggctcgaag 180
ggcygctccccc caacctgagg cttgaccacag agtagcagcg cagagctggag 240
cagagcct ctcgagcct tgcagcaggt cttgacgca ctgcctcagct tocagctgga gtga 294

<210> SEQ ID NO: 31
<211> LENGTH: 294
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Human Codon-Optimized Influenza A Virus M2 Protein

<400> SEQUENCE: 31
atgctctct gcagctgcaggt ggaacacccc atacggaacag agtgggtgcct ccgtygacoac 60
gggtgcct gcagctgcaggt ctagcgcgcc aacctcctcg gctgccctcg cctgcacgcc 120
tggtacctgct atccacctgg ctttaagatc gtttaagta tggctcgaag 180
ggcygctccccc caacctgagg cttgaccacag agtagcagcg cagagctggag 240
cagagcct ctcgagcct tgcagcaggt cttgacgca ctgcctcagct tocagctgga gtga 294

<210> SEQ ID NO: 32
<211> LENGTH: 1566
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Human Codon Optimized Coding Region Encoding M2BP

<400> SEQUENCE: 32
atgctctcct tccagcaggt ggaacacccc atcagcaggt acctgaggct ccagctgcaat 60
gggtgcct gcagctgcaggt ctagcgcgcc aacctcctcg gctgccctcg cctgcacgcc 120
ggacagcag acacgacgcc ctagcgcgcc aacctcctcg gctgccctcg cctgcacgcc 180
ggacagt ctaagagtcct ctagcgcgcc aacctcctcg cctgcacgcc 240
atcctctct gcctcagcaggt ggtgcgtctct ctttggagc aagcagacag 300
amactctgct cagaccccct cttgagcaggt aagcagacag aagcagacag 360
tcagcaggt ctaagagtcct ctagcgcgcc aacctcctcg cctgcacgcc 420
ggacagcag gcctcagcaggt ggtgcgtctct ctttggagc aagcagacag 480
atcctctc ttcacacacat gacgtagcct agctcagcaggt aagcagacag 540
| acggaatgg accctctgat gtcgagcctt aagcaggggt ccacacgacc cagtagctcc | 600 |
| ggcacgagc gacgacagcct aagggggttc gacacgagtg tgcagagaacg | 660 |
| at(cv)gcaggg gatcagctga cggagacctcg tggcgagcgg aaaaaggggc aaaaactagg | 720 |
| at(c(cv))gcaggg gatcagctga cggagacctcg tggcgagcgg aaaaaggggc aaaaactagg | 780 |
| gcacagagct acatagagct cgaagagtag aacccggtga atgagatgtt tgaagaatcc | 840 |
| atgtgtctcg ggcacgagcct gtcgagcctt aagcaggggt ccacacgacc cagtagctcc | 900 |
| cgtgcgtcgg tattagcctgc tgccttggca agccagatcg cctttgagac aggcaggtac | 960 |
| tctgcgttct ggaaagagcctt atcagatcttc ctcagatcttc ccagaggtga cagtttaata | 1020 |
| aggcccacg aaaaagagct accacactct ccactttggtt ggtagtctcc gcctagccct | 1080 |
| gctgcgtcgg atcagatcttc ctcagatcttc ccacagctct tccacaggga | 1140 |
| acacactctc ccacagctct ccacagctct tccacaggga | 1200 |
| agcaaggctt gagtagagc agtagatgtg gcctagatcg ccagaggtac cagtttaata | 1260 |
| aacagacgc gggtccttgg cgcctactagt gcccacgcag cccacgagcag cccacgagcag | 1320 |
| acatccgtt ttgagatctgg cccacagcctt ctgcttagatgtg gcctagatcg ccagaggtac | 1380 |
| atctctctc tggagatctgg cccacagcctt ctgcttagatgtg gcctagatcg ccagaggtac | 1440 |
| aagtccaggg ccagaggtgt gcctagatcg ttcagatctgg cccacagcctt | 1500 |
| cctcgcttgg atctctctc tggagatctgg cccacagcctt ccagaggtac cagtttaata | 1560 |
| gcacc | 1566 |

<210> SEQ ID NO: 33
<211> LENGTH: 1566
<212> TYPE: DNA
<220> ORGANISM: Artificial sequence
<223> OTHER INFORMATION: Human Codon Optimized Coding Region Encoding e2NP
<400> SEQUENCE: 33

atgcacgctc tgcagccgggt gcacacccccc atcagacggc atggggcgtg acatgcaac | 60 |
| gcgcgggagc acatcggcagc ccagagcacc aagggaggtg atcagacggc gcacacccccc | 120 |
| ggagggggc caagccgagc ccagagcacc aagggaggtg atcagacggc gcacacccccc | 180 |
| gcgcgggagc acatcggcagc ccagagcacc aagggaggtg atcagacggc gcacacccccc | 240 |
| atcagacggc gcacacccccc atcagacggc gcacacccccc atcagacggc gcacacccccc | 300 |
| gcgcgggagc acatcggcagc ccagagcacc aagggaggtg atcagacggc gcacacccccc | 360 |
| gcgcgggagc acatcggcagc ccagagcacc aagggaggtg atcagacggc gcacacccccc | 420 |
| gcgcgggagc acatcggcagc ccagagcacc aagggaggtg atcagacggc gcacacccccc | 480 |
| gcgcgggagc acatcggcagc ccagagcacc aagggaggtg atcagacggc gcacacccccc | 540 |
| gcgcgggagc acatcggcagc ccagagcacc aagggaggtg atcagacggc gcacacccccc | 600 |
| gcgcgggagc acatcggcagc ccagagcacc aagggaggtg atcagacggc gcacacccccc | 660 |
| gcgcgggagc acatcggcagc ccagagcacc aagggaggtg atcagacggc gcacacccccc | 720 |
| gcgcgggagc acatcggcagc ccagagcacc aagggaggtg atcagacggc gcacacccccc | 780 |
| gcgcgggagc acatcggcagc ccagagcacc aagggaggtg atcagacggc gcacacccccc | 840 |
| gcgcgggagc acatcggcagc ccagagcacc aagggaggtg atcagacggc gcacacccccc | 900 |
-continued

cgccgctcgc tgtaagggcgc cgcccgccgcc agcgctcagcg actggaggg gaggctacac 960
gacctgctgg gcctagccccc ttcgacaggg tctgcatacc gcacaggtcag ccgcagtttc 1020
aggccaccc aagcagcacc gccaagagcc cagctgttgct gtaggcctcg cccaagcgcc 1080
gctgagagc actggaggg gctgagcttca acaaggggca caaggggctt gcacaagggc 1140
aagctgagtc ccacggtgcg ctgagatgctc acacagagac agcagagac cgctggagcc 1200
agccctctgg agctggagag caggtcaagcg gcattcgagcc aagaagggg aagagccacc 1260
aagcagagca ggccagccgag cgcagagcag cacgtcgccag cacctcgcta cctgcagcagg 1320
aagctgctct gcgcagagcg caacgctgtgc gcgccttcca gcggccacac cgagccaggg 1380
aagccagcag cggagctcag gatctgagca ggcgcagagcg cggaggtgcg 1440
agctcgagcg gcacggtgcg ctgctgcagc agcagagccag aggccgagcc cccctctgg 1500
cccgctctgc aacagtgcgcc cgagccggcg tcctctcttc gcgaagcgac cgagagtttg 1560
gagcaag 1666

<210> SEQ ID NO: 34
<211> LENGTH: 1566
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE: OTHER INFORMATION: Human Codon Optimized Coding Region Encoding Np EN2
<400> SEQUENCE: 34

tagccagggc agggccagag cgcaggttac cggcagatgg aagctgtcgg tgcagggcg 60
agccagcacc aatcctggg ctcgctgggc aagattgtgg tgcagctggg aagattttat 120
atccagattc gcacggtgcg taaggctgagc gattcgaggg ggcgctaatt acagatctct 180
cggatctgc agcagttcagc ttcgagggcc tttgagtcgc gcgagtataa gtaatctgca 240
gagcaagctt ctcgctggag agcccccagag cgcaggggg gatctatota gcagrcytcg 300
aagccaggaat gtggcgcgcc aacgaatcttgactgaatcg ccggagaggg tgcggaggtg 360
agccaggtcatacttg ccagcagagcg cgctgctgctg gccaggggg accctctgctg 420
aagccagttc acgcagcagc cttgaagcgcg aacgctgctg cagctgctgctg 480
cccgccgctgtcagctgc gcacagttgc aacgctgctgc cgcgctgctg ggcaggggtg 540
gcgcgctgtaa agggcctggg caacatggtg atgcagcctgg tgcagcgctgc aacagagggc 600
atccagagga gcacatttgg gcggcggagcg agcccgccagcgccagccgtcagctgcag 660
cgcctgtgaacttccttga ggaacttgg cgcggcgcgagc ccagccggctt ctgctcctgg 720
cacctaggg aagcgaaatt cctccaggtg gcacaggttg aagccctgctg tttccctgctg 780
aagccagccc gatgcgtcgg ggctcgctgt gcgcgattac ggtctcgtgc 840
tcccccggc gcgctgctgc ttcggaggtg ccagggatat gcctgccggc 900
atcactctcc ctcagccttt gagaatttct gcgggtatat gcgtctatgg accacatgg 960
aaccgcgggt caccgctggtg gccctgttgc cctcagagccgc ggcgcctgctg 1020
cggctgctgc ctcagccttt taaaaggcct ccagctgctg cacggctgg cgcctcctgc 1080
agggagctac agatcgctgc aacagcagagcg cggagctggagccc ctcagcctgc 1140
ctgcagctgctgc ttcagcgggttc tattgcagac cgctgtccgg ggaatccggtcataaag ctcagcagagcg 1200
-continued

gcagctctgc gctcagatcc gattcatgct acctttctctg tcgaacggaat ttcacatcc 1260
gacgccgacac ctctggttat gcgcttctcc gcgaatcag aagagccgac acgacatgy 1320
agggcgggaa tttcaggttag ctggagaag gctgggttgag aagatgtgtg gctcaggggct 1380
cgggcgtgt ggctagcctag cgacagagag gcgcggtcttc ccaatgtgcc ttcttttctgt 1440
atgacacatg aagggcttcg cttttctcgg gcaaatctggc aagagatgta ttaaatttattatct 1500
cctctaccc gcggtacag gcgcatacgc aagagatggtt ttcgtacctt taaagcttcc 1560
agtgtt 1566

<210> SEQ ID NO: 35
<211> LENGTH: 1566
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE: Human Codon Optimized Coding Region Encoding NPH2
<400> SEQUENCE: 35

atgagccgac aggggaccaac gaggagctac gcagcagtctg agacgacgag cyagagccag 60
aacgccagc agaatcggcag cagctgttgg gcgacatcgg caggttttac 120
atcagatgt gcggcgttac gcagcgtgag gacatcgagc gacagctgtg ccaagaagcg 180
cgagctacat gcgaggtttg gccgtcggcc ttcgagcag cgggacgaga cagagagccg 240
gacgccgacac gcgcgaccgca gacccgacag aagacgagcc gcggctgctta caggagcttg 300
aagctgagct gcagacgagca gctgtactgct tcagccagag aggacagctcg gaggagcttg 360
agggagccag cacaacggaga cagcgcaccg cgcgctctag gcccctgctcg gatctgcac 420
agcaactgta acacgcgacc caaccagcag ccccgggccc ttgctgctg gagcgtgagc 480
ccggagctgt gcggctgtca gcggcggcag cccctgcacc gagaggcggc cgggacgacc 540
gcgccgctgg aggctggttag agtggctgtag tggatcagt gaaagagcct gacagcgggg 600
atcagaccc gcacactgttg gcgggctgctg aaggggctgg cagccagagag cggctgcagc 660
agtctgtgcg cagcgcgtgg accagcgcgg cagcgccttg ccagcgcagc cagcgcctcg 720
caggagcct gaggagcgcgg ccccggcacc gcgggtgcac gcggctgcag ctttcttcgg 780
agggcgccac tgctgggctg aggctgggct gcgcacgac gctgtctgtgc gcggctctgg 840
taacgccgag cctgcgcgcag cggcgcctcg ttggagcgg gaggctcag cctcgctgctg 900
atcagctcc ctctgctgct gcgcagcgac caggtgtctgag gcggctgcagc gcccacagag 960
aacccggccg ccacagcgcgc gcggcgtgtg atgcccctgcc acacgcgcgc ccgctagagc 1020
cgcggctgc tgctggctgt caacggccac ccagctgtgc ccggggcacc gcgtgcgacc 1080
agggagcttg atgagcgcgcg caaccacgcag atggccgacc tcggagccgg cggcccgttg 1140
ctgggggcag gtctgtgctgt cgctacgcgg aggccgctgg gcacacccaa ccagcacag 1200
gccggccg cgcgcgcttc gacatcgcgt gcgtagcgg ccgctgacgc ccgctgacgc 1260
gacgccgac gcgtgtgctgc cggcgcgtgc gcgcgcgcgc agggcgcgcg cgcgcgcgatgc 1320
aggggagctg tggctgcttg cgcgcgctgc gcgcgcgcgc cagcgcgctgc cagcgcgctgc 1380
aggggagctg tggctgcttg cgcgcgctgc gcgcgcgcgc cagcgcgctgc cagcgcgctgc 1440
atgagccgac gcacagcttc cttcttggcc gcaacgctgc aggacagcgc gcacagcttc 1500
cgagctgtgc gcgggcgctgc gcgcgcgctgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 1560
---continued---

agcgcac 1566

<210> SEQ ID NO: 36
<211> LENGTH: 1693
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Human Codon Optimized Coding Region Encoding EBV NF Protein

<400> SEQUENCE: 36

atgctgaca tgcatacga cagcattaac acaggtacta ttgacaaaaa ccggcgaaga 60
cgaaattgct gccaaagcgc ccaataacta aacggyccac ccctcggcc 120
cctgagtaa agagggccgc cactaacat gctgagccca gcaactcctt atctgaaacg 180
gatattgctt ggaaatctca aaaaaagcag aagcggccac agttgagaaa ggtcttttc 240
amastagcttg tacaagctgc ttgcatttat acacgagtga tggctccagcc ggggcttccc 300
gcctagtgag acacagaaac gcagtttccca aaaaacgca gcgtagagta tttgaaagaa 360
ggaaacaggg aattagccaa taacaaact gggggacat ttattaaagag aagcggaggg 420
gatgcaacaa tcttattttag ccgatgaaaa attaacctcc tgacagaggg ggttacaaaaa 480
agtataacaa cgcgcattgg cagcaggggg ttttcggaac ttcacaatt aatgtatggt 540
catctcagga tgaagaagtgt atgtccttcag ctcctggagc gttcaanagag ggtcgctttc 600
gaccctgctt taataaaccc tttctggcgg tttataatttttc gggggcgttc ttgacgcagca 660
ggcctaactt gtaaggccgcat ggctgctctt actgtctaggg ggttttacg cactggtggc 720
gccatcgctg acttggtcag ataaggcgcg tctcagaggt actaaaccga aaaaaaaaga 780
gcccctcttg ataataattgac acctgtgtcct cttgactgtct atcttcatgaaa 840
ttaacattttc tcacatttcc ttccttcgct ctcctggcct cgtcatcagc aaaaagtaa 900
stggcttgcc cattgttgcg aagtaattgtc tcagggagag gacatcgctt cttgactgtct 960
stggcttgcc cattgttgcg aagtaattgtc tcagggagag gacatcgctt cttgactgtct 1020
saaatcttcc aactgttac gccgtctctt caagttgactc tcaagctggt caaagttaa 1080
gggtgtacg actgtttcat ctccttgctg gcagttgactc tcaagctggt caaagttaa 1140
agagtctctt ctcttgctctt gcgggttctg ggccgcttcg ccgaaagtct cagagacttc 1200
ttggctgctg ggccggctta attttaccgg cgggagcgtc taaattgtta aaaaaaca 1260
gttctgctgt ggccggctta attttaccgg cgggagcgtc taaattgtta aaaaaaca 1320
ctttggctg cccttgctg ccggttgctt ccggaagtct aggagacgg ctcttttgtt 1380
cagttatcgg cttctgtgtg cggggcttcg cttcttttgtt aaaaaaca 1440
gggccggctg ggccggctta attttaccgg cgggagcgtc taaattgtta aaaaaaca 1500
ttaaaagctg cccttgctg ccggttgctt ccggaagtct aggagacgg ctcttttgtt 1560
aattttaccgg cttctgtgtg cggggcttcg cttcttttgtt aaaaaaca 1620
gctttctgcg cccttgctg ccggttgctt ccggaagtct aggagacgg ctcttttgtt 1680
tac 1683
---continued

**ORGANISM:** Artificial sequence  
**FEATURES:**  
**H: Human Codon Optimized Coding Region Encoding**  
**IBV NP Protein**

**SEQUENCE:** 37

```
agtgcgaca tgcatacga cagcatcaco accggacca tggcaaacg cccggagag 60
cggcgccccc gcggacccgg cgcccaaccc ccccccaccc agcgcgaccc ccggcccggc 120
cccggcgac ccggcggcgg gacccgaccc ggcgcggcgg gcgcggcgg gcgcggcgg 180
ggcgcgcgg cggccgaccc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 240
gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 300
ggcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 360
ggcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 420
ggcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 480
ggcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 540
ggcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 600
ggcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 660
ggcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 720
ggcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 780
ggcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 840
ggcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 900
ggcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 960
agtgcgaca tgcatacga cagcatcaco accggacca tggcaaacg cccggagag 1020
cggcgccccc gcggacccgg cgcccaaccc ccccccaccc agcgcgaccc ccggcccggc 1080
cggcgccccc gcggacccgg cgcccaaccc ccccccaccc agcgcgaccc ccggcccggc 1140
cggcgccccc gcggacccgg cgcccaaccc ccccccaccc agcgcgaccc ccggcccggc 1200
cggcgccccc gcggacccgg cgcccaaccc ccccccaccc agcgcgaccc ccggcccggc 1260
cggcgccccc gcggacccgg cgcccaaccc ccccccaccc agcgcgaccc ccggcccggc 1320
cggcgccccc gcggacccgg cgcccaaccc ccccccaccc agcgcgaccc ccggcccggc 1380
cggcgccccc gcggacccgg cgcccaaccc ccccccaccc agcgcgaccc ccggcccggc 1440
cggcgccccc gcggacccgg cgcccaaccc ccccccaccc agcgcgaccc ccggcccggc 1500
cggcgccccc gcggacccgg cgcccaaccc ccccccaccc agcgcgaccc ccggcccggc 1560
cggcgccccc gcggacccgg cgcccaaccc ccccccaccc agcgcgaccc ccggcccggc 1620
```
<table>
<thead>
<tr>
<th>Ala Thr Gly Thr Cys Thr Ala Ala Cys Ala Thr Gly Ala Cys Ala</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr Cys Gly Ala Cys Thr Ala Ala Ala Cys Ala Ala Cys</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Ala Gly Gly Cys Ala Cys Gly Ala Thr Cys Gly Ala Ala Ala Cys</td>
<td>30</td>
<td>40</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Ala Cys Cys Cys Cys Cys Gly Ala Gly Ala Gly Cys Thr Gly Ala</td>
<td>35</td>
<td>50</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>Cys Ala Cys Cys Cys Gly Ala Cys Thr Cys Ala Gly Gly Cys</td>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
</tr>
<tr>
<td>Cys Gly Cys Ala Cys Cys Ala Gly Ala Cys Cys Ala Thr Ala</td>
<td>90</td>
<td>95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala Thr Ala Ala Ala Gly Cys Cys Gly Cys Ala Cys Thr Cys</td>
<td>100</td>
<td>105</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>Thr Gly Gly Cys Cys Cys Cys Cys Cys Cys Thr Cys Ala Ala</td>
<td>115</td>
<td>120</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>Cys Ala Ala Gly Ala Gly Ala Cys Cys Gly Ala Ala Ala Cys</td>
<td>130</td>
<td>135</td>
<td>140</td>
<td></td>
</tr>
<tr>
<td>Cys Cys Cys Thr Cys Thr Cys Cys Cys Gly Ala Cys Gly Cys Ala</td>
<td>145</td>
<td>150</td>
<td>155</td>
<td>160</td>
</tr>
<tr>
<td>Cys Cys Ala Cys Ala Cys Ala Gly Ala Cys Gly Thr Ala Cys</td>
<td>165</td>
<td>170</td>
<td>175</td>
<td></td>
</tr>
<tr>
<td>Gly Ala Cys Gly Ala Cys Ala Thr Cys Gly Cys Ala Gly Gly</td>
<td>180</td>
<td>185</td>
<td>190</td>
<td></td>
</tr>
<tr>
<td>Ala Ala Ala Gly Ala Thr Ala Cys Ala Gly Ala Ala Gly Cys</td>
<td>195</td>
<td>200</td>
<td>205</td>
<td></td>
</tr>
<tr>
<td>Ala Gly Ala Cys Thr Cys Cys Cys Ala Cys Thr Gly Ala Ala Thr</td>
<td>210</td>
<td>215</td>
<td>220</td>
<td></td>
</tr>
<tr>
<td>Thr Ala Ala Gly Ala Gly Thr Cys Cys Gly Thr Gly Ala Ala Thr</td>
<td>225</td>
<td>230</td>
<td>235</td>
<td>240</td>
</tr>
<tr>
<td>Ala Ala Ala Gly Ala Thr Gly Gly Thr Gly Thr Ala Ala Gly Cys</td>
<td>245</td>
<td>250</td>
<td>255</td>
<td></td>
</tr>
<tr>
<td>Thr Gly Gly Cys Cys Gly Ala Ala Cys Ala Ala Ala Cys Cys Ala</td>
<td>260</td>
<td>265</td>
<td>270</td>
<td></td>
</tr>
<tr>
<td>Cys Cys Ala Ala Ala Thr Gly Thr Ala Ala Ala Cys Ala Cys</td>
<td>275</td>
<td>280</td>
<td>285</td>
<td></td>
</tr>
<tr>
<td>Gly Cys Cys Gly Gly Cys Cys Thr Ala Ala Ala Cys Ala Ala Cys</td>
<td>290</td>
<td>295</td>
<td>300</td>
<td></td>
</tr>
<tr>
<td>Ala Cys Ala Thr Gly Ala Gly Ala Cys Ala Cys Cys Thr</td>
<td>305</td>
<td>310</td>
<td>315</td>
<td>320</td>
</tr>
<tr>
<td>Gly Ala Thr Ala Cys Ala Cys Ala Cys Cys Gly Cys Ala Cys</td>
<td>325</td>
<td>330</td>
<td>335</td>
<td></td>
</tr>
<tr>
<td>Gly Cys Cys Gly Thr Gly Ala Ala Gly Ala Gly Ala Thr Cys</td>
<td>340</td>
<td>345</td>
<td>350</td>
<td></td>
</tr>
<tr>
<td>Thr Gly Cys Thr Gly Gly Cys Cys Cys Cys Cys Gly Ala</td>
<td>355</td>
<td>360</td>
<td>365</td>
<td></td>
</tr>
<tr>
<td>Thr Gly Ala Cys Ala Ala Gly Ala Ala Cys Thr Gly Ala Cys</td>
<td>370</td>
<td>375</td>
<td>380</td>
<td></td>
</tr>
<tr>
<td>Thr Ala Thr Cys Ala Ala Gly Ala Ala Gly Ala Gly Ala Ala Ala</td>
<td>385</td>
<td>390</td>
<td>395</td>
<td>400</td>
</tr>
<tr>
<td>Ala Cys Gly Cys Cys Cys Gly Gly Ala Cys Gly Thr Ala Ala</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
-continued-

Gly Gly Ala Gly Gly Gly Cys Ala Ala Gly Gly Ala Gly Ala Gly
405 410 415
Gly Cys Ala Gly Gly Cys Ala Ala Cys Ala Ala Ala Gly Ala
420 425 430

Ala Thr Cys Ala Gly Thr Cys Ala Ala Cys Ala Ala Cys Ala Ala
435 440 445

Cys Ala Gly Cys Gly Cys Cys Thr Thr Thr Cys Thr Ala
450 455 460

Thr Ala Ala Gly Ala Thr Gly Thr Cys Gly Thr Gly Ala Thr
465 470 475 480

Gly Ala Cys Ala Ala Gly Ala Cys Ala Ala Thr Cys Ala Cys
485 490 495

Thr Thr Thr Cys Cys Cys Cys Ala Ala Cys Ala Ala Gly Ala
500 505 510

Thr Cys Ala Ala Thr Cys Ala Ala Cys Ala Ala Cys Ala Gly Ala
515 520 525

Gly Ala Gly Thr Ala Ala Gly Ala Cys Thr Ala Ala Gly Thr
530 535 540

Ala Cys Ala Ala Gly Ala Cys Ala Ala Cys Ala Ala Gly Thr
545 550 555 560

Cys Thr Cys Cys Ala Ala Gly Ala Cys Ala Ala Cys Ala Gly Thr
565 570 575

Gly Gly Cys Thr Gly Ala Cys Cys Ala Cys Ala Ala Thr Ala
580 585 590

Thr Gly Ala Thr Ala Gly Gly Cys Ala Ala Thr Ala Ala Gly Thr
595 600 605

Gly Ala Thr Gly Ala Ala Cys Gly Ala Ala Thr Gly Thr Gly Thr
610 615 620

Thr Thr Cys Cys Ala Gly Ala Gly Ala Cys Ala Ala Gly Thr
625 630 635 640

Gly Cys Cys Thr Gly Ala Ala Gly Ala Gly Gly Thr Cys Gly Thr
645 650 655

Cys Cys Thr Gly Ala Thr Cys Cys Cys Thr Cys Cys Cys Thr
660 665 670

Ala Thr Thr Ala Gly Thr Ala Cys Thr Thr Cys Thr Gly Cys Cys
675 680 685

Gly Cys Ala Gly Cys Ala Cys Thr Cys Cys Thr Gly Cys Cys Cys
690 695 700

Ala Ala Gly Ala Thr Cys Thr Gly Cys Ala Ala Cys Cys Cys Cys
705 710 715 720

Gly Gly Cys Thr Ala Gly Cys Ala Ala Ala Gly Ala Gly
725 730 735

Gly Cys Gly Cys Gly Cys Gly Cys Ala Ala Ala Cys Ala Thr Ala
740 745 750

Ala Gly Ala Cys Gly Ala Gly Cys Ala Ala Cys Ala Ala Cys Ala
755 760 765

Thr Thr Thr Ala Ala Gly Cys Ala Ala Cys Ala Ala Cys Ala
770 775 780

Thr Gly Cys Cys Ala Ala Cys Gly Cys Gly Cys Gly Cys Cys Cys
785 790 795 800

Gly Cys Thr Thr Ala Gly Ala Gly Ala Thr Ala Cys Ala Ala Gly
805 810 815
Gly Cys Cys Ala Ala Gly Ala Cys Cys Cys Cys Thr Ala Cys Gly 820 825 830
Ala Gly Ala Ala Gly Ala Thr Ala Cys Thr Gly Cys Thr Gly Ala Ala 835 840 845
Cys Cys Thr Gly Ala Ala Gly Ala Ala Cys Ala Ala Gly Thr Gly Cys 850 855 860
Thr Cys Ala Gly Cys Cys Cys Cys Cys Cys Ala Gly Cys Ala Gly Ala 865 870 875 880
Ala Gly Gly Cys Cys Cys Thr Gly Gly Thr Gly Gly Ala Thr Cys Ala 885 890 895
Gly Gly Thr Gly Ala Thr Cys Gly Cys Ala Gly Thr Ala Gly Ala 900 905 910
Ala Ala Cys Cys Cys Gly Cys Ala Thr Cys Gly Cys Cys Gly 915 920 925
Ala Cys Ala Thr Cys Gly Ala Gly Ala Thr Cys Thr Gly Ala Cys 930 935 940
Thr Cys Thr Gly Cys Thr Gly Cys Ala Gly Ala Ala Gly Ala Gly Cys 945 950 955 960
Ala Thr Gly Gly Thr Ala Gly Thr Cys Gly Thr Ala Ala Gly Ala Cys 965 970 975
Cys Cys Thr Cys Thr Gly Thr Gly Cys Cys Thr Cys Thr Cys Thr Ala Ala 980 985 990
Gly Gly Thr Thr Gly Thr Gly Cys Thr Gly Cys Cys Ala Thr Cys 995 1000 1005
Thr Cys Cys Ala Thr Cys Thr Ala Cys Gly Cys Cys Ala Ala Gly 1010 1015 1020
Ala Thr Thr Cys Cys Cys Ala Gly Cys Thr Gly Gly Gly Cys 1025 1030 1035
Thr Thr Thr Ala Ala Cys Ala Gly Cys Thr Gly Ala Gly Ala Gly 1040 1045 1050
Thr Ala Cys Thr Cys Cys Ala Thr Gly Gly Thr Gly Gly Cys 1055 1060 1065
Thr Ala Thr Gly Ala Gly Cys Cys Ala Thr Gly Gly Cys Ala Cys 1070 1075 1080
Cys Thr Gly Thr Ala Thr Ala Cys Ala Ala Thr Gly Gly Cys Cys 1085 1090 1095
Ala Cys Ala Cys Cys Gly Thr Cys Thr Cys Thr Ala Thr Cys 1100 1105 1110
Cys Thr Gly Cys Gly Cys Thr Cys Gly Gly Cys Gly Ala Cys 1115 1120 1125
Gly Ala Thr Gly Cys Cys Ala Ala Gly Cys Ala Ala Gly 1130 1135 1140
Thr Cys Thr Cys Ala Gly Cys Thr Gly Thr Thr Thr Thr Thr Thr 1145 1150 1155
Ala Thr Gly Ala Gly Thr Gly Thr Thr Thr Cys Gly Gly Cys 1160 1165 1170
Gly Cys Cys Gly Cys Thr Ala Thr Gly Ala Gly Gly Ala Thr 1175 1180 1185
Cys Thr Gly Ala Gly Ala Gly Thr Cys Thr Gly Thr Thr Cys Ala 1190 1195 1200
<table>
<thead>
<tr>
<th>Gly Cys</th>
<th>Cys Cys Thr Gly Ala</th>
<th>Cys Ala Gly Gly Cys</th>
<th>Ala Cys Thr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1205</td>
<td>1210</td>
<td>1215</td>
<td></td>
</tr>
<tr>
<td>Gly Ala</td>
<td>Gly Thr Thr Cys Ala</td>
<td>Ala Gly Cys Cys</td>
<td>Ala Gly Gly</td>
</tr>
<tr>
<td>1220</td>
<td>1225</td>
<td>1230</td>
<td></td>
</tr>
<tr>
<td>Thr Cys</td>
<td>Cys Gly Cys Cys</td>
<td>Thr Gly Ala Ala Gly</td>
<td>Thr Gly Cys</td>
</tr>
<tr>
<td>1235</td>
<td>1240</td>
<td>1245</td>
<td></td>
</tr>
<tr>
<td>Ala Ala</td>
<td>Gly Gly Gly Cys Thr</td>
<td>Thr Thr Cys Ala Thr</td>
<td>Gly Thr Gly</td>
</tr>
<tr>
<td>1250</td>
<td>1255</td>
<td>1260</td>
<td></td>
</tr>
<tr>
<td>Cys Cys</td>
<td>Cys Gly Cys Ala</td>
<td>Ala Gly Gly Ala Gly</td>
<td>Cys Ala Gly</td>
</tr>
<tr>
<td>1265</td>
<td>1270</td>
<td>1275</td>
<td></td>
</tr>
<tr>
<td>Gly Thr</td>
<td>Gly Gly Ala Gly Gly</td>
<td>Gly Cys Ala Thr Gly</td>
<td>Gly Gly Cys</td>
</tr>
<tr>
<td>1280</td>
<td>1285</td>
<td>1290</td>
<td></td>
</tr>
<tr>
<td>Gly Cys</td>
<td>Cys Gly Cys Cys</td>
<td>Thr Gly Ala Thr Gly</td>
<td>Ala Gly Cys</td>
</tr>
<tr>
<td>1295</td>
<td>1300</td>
<td>1305</td>
<td></td>
</tr>
<tr>
<td>Ala Thr</td>
<td>Cys Ala Ala Gly Cys</td>
<td>Thr Gly Ala Ala Gly</td>
<td>Thr Thr Cys</td>
</tr>
<tr>
<td>1310</td>
<td>1315</td>
<td>1320</td>
<td></td>
</tr>
<tr>
<td>Thr Gly</td>
<td>Gly Gly Cys Cys Cys</td>
<td>Cys Cys Ala Thr Gly</td>
<td>Ala Cys Cys</td>
</tr>
<tr>
<td>1325</td>
<td>1330</td>
<td>1335</td>
<td></td>
</tr>
<tr>
<td>Cys Gly</td>
<td>Gly Thr Thr Gly Gly</td>
<td>Gly Cys Gly Cys Ala</td>
<td>Ala Ala Cys</td>
</tr>
<tr>
<td>1340</td>
<td>1345</td>
<td>1350</td>
<td></td>
</tr>
<tr>
<td>Gly Ala</td>
<td>Gly Gly Thr Cys Ala</td>
<td>Cys Gly Gly Cys Ala</td>
<td>Gly Ala Gly</td>
</tr>
<tr>
<td>1355</td>
<td>1360</td>
<td>1365</td>
<td></td>
</tr>
<tr>
<td>Gly Gly</td>
<td>Cys Gly Gly Cys Ala</td>
<td>Gly Thr Gly Cys Ala</td>
<td>Cys Ala Gly</td>
</tr>
<tr>
<td>1370</td>
<td>1375</td>
<td>1380</td>
<td></td>
</tr>
<tr>
<td>Ala Thr</td>
<td>Ala Ala Gly Thr Gly</td>
<td>Cys Cys Ala Gly Cys</td>
<td>Cys Cys Cys</td>
</tr>
<tr>
<td>1385</td>
<td>1390</td>
<td>1395</td>
<td></td>
</tr>
<tr>
<td>Gly Thr</td>
<td>Thr Thr Thr Thr Gly</td>
<td>Cys Cys Gly Thr Thr</td>
<td>Gly Ala Gly</td>
</tr>
<tr>
<td>1400</td>
<td>1405</td>
<td>1410</td>
<td></td>
</tr>
<tr>
<td>Ala Gly</td>
<td>Ala Cys Cys Ala Thr</td>
<td>Cys Gly Cys Cys Cys</td>
<td>Cys Thr Gly</td>
</tr>
<tr>
<td>1415</td>
<td>1420</td>
<td>1425</td>
<td></td>
</tr>
<tr>
<td>Thr Cys</td>
<td>Thr Ala Ala Gly Cys</td>
<td>Ala Gly Gly Cys Ala</td>
<td>Gly Thr Thr</td>
</tr>
<tr>
<td>1430</td>
<td>1435</td>
<td>1440</td>
<td></td>
</tr>
<tr>
<td>Ala Gly</td>
<td>Ala Cys Gly Ala Ala</td>
<td>Thr Gly Ala Ala Gly</td>
<td>Ala Gly Thr</td>
</tr>
<tr>
<td>1445</td>
<td>1450</td>
<td>1455</td>
<td></td>
</tr>
<tr>
<td>Ala Thr</td>
<td>Gly Ala Ala Cys Gly</td>
<td>Cys Cys Gly Ala Ala</td>
<td>Gly Gly Cys</td>
</tr>
<tr>
<td>1460</td>
<td>1465</td>
<td>1470</td>
<td></td>
</tr>
<tr>
<td>Cys Gly</td>
<td>Ala Gly Ala Cys Gly</td>
<td>Cys Cys Gly Ala Ala</td>
<td>Cys Thr Gly</td>
</tr>
<tr>
<td>1475</td>
<td>1480</td>
<td>1485</td>
<td></td>
</tr>
<tr>
<td>Ala Ala</td>
<td>Gly Gly Gly Cys Ala</td>
<td>Ala Cys Thr Gly Cys</td>
<td>Thr Gly Ala</td>
</tr>
<tr>
<td>1490</td>
<td>1495</td>
<td>1500</td>
<td></td>
</tr>
<tr>
<td>Ala Ala</td>
<td>Gly Ala Thr Gly Ala</td>
<td>Thr Gly Ala Ala Cys</td>
<td>Gly Ala Thr</td>
</tr>
<tr>
<td>1505</td>
<td>1510</td>
<td>1515</td>
<td></td>
</tr>
<tr>
<td>Thr Cys</td>
<td>Cys Ala Thr Gly Gly</td>
<td>Cys Cys Ala Ala Gly</td>
<td>Ala Ala Gly</td>
</tr>
<tr>
<td>1520</td>
<td>1525</td>
<td>1530</td>
<td></td>
</tr>
<tr>
<td>Ala Ala</td>
<td>Ala Ala Gly Cys Gly</td>
<td>Gly Cys Ala Ala Cys</td>
<td>Gly Cys Cys</td>
</tr>
<tr>
<td>1535</td>
<td>1540</td>
<td>1545</td>
<td></td>
</tr>
<tr>
<td>Thr Thr</td>
<td>Cys Ala Ala Gly Gly</td>
<td>Gly Cys Ala Ala Gly</td>
<td>Ala Ala Gly</td>
</tr>
<tr>
<td>1550</td>
<td>1555</td>
<td>1560</td>
<td></td>
</tr>
<tr>
<td>Ala Thr</td>
<td>Gly Thr Thr Cys Ala</td>
<td>Ala Gly Ala Ala Ala</td>
<td>Gly Ala Ala</td>
</tr>
<tr>
<td>1565</td>
<td>1570</td>
<td>1575</td>
<td></td>
</tr>
<tr>
<td>Gly Ala</td>
<td>Thr Ala Ala Gly Ala</td>
<td>Ala Cys Ala Ala Gly</td>
<td>Gly Thr Thr</td>
</tr>
</tbody>
</table>
1580  1585  1590
Ala Ala Cys Cys Cys Ala Thr Cys Gly Ala Gly Ala Thr Thr
1595  1600  1605
Cys Cys Ala Thr Cys Ala Ala Gly Cys Ala Gly Ala Cys Cys
1610  1615  1620
Ala Thr Cys Cys Ala Ala Cys Thr Thr Cys Thr Thr Cys
1625  1630  1635
Thr Thr Cys Gly Gly Cys Ala Gly Gly Gly Ala Thr Ala Cys Cys
1640  1645  1650
Gly Cys Cys Gly Ala Gly Cys Ala Thr Ala Cys Gly Ala Thr
1655  1660  1665
Gly Ala Cys Thr Gly Gly Ala Cys Thr Ala Cys Thr Gly Ala
1670  1675  1680

<210> SEQ ID NO 39
<211> LENGTH: 552
<212> TYPE: DNA
<213> ORGANISM: Hepatitis B virus

<400> SEQUENCE: 39
atgacatcg accctcattg cgaatttggg gctacctggtg agttactcto gtttttgtgct  60
tctgacctt ttcctcactg acagtatttt ctaactcctg cttcagcctct ctctctctct  120
gcctatttt ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct  180
tgctgtgtggt gAAAAAAAGA gGAAAGAAAAA gGAAAGAAAAA gGAAAGAAAAA gGAAAGAAAAA  240
tctctctctct tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
-continued

Phe Arg Gin Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg 100 105 110
Glu Thr Val Ile Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr 115 120 125
Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro 130 135 140
Glu Thr Thr Val Val Arg Arg Gly Arg Ser Pro Arg Arg Arg Thr 145 150 155 160
Pro Ser Pro Arg Arg Arg Arg Gin Ser Pro Arg Arg Arg Arg Ser 165 170 175
Gln Ser Arg Gin Ser Gin Cys 180

<210> SEQ ID NO: 41
<211> LENGTH: 555
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic HBcAg

<400> SEQUENCE: 41
atgatagc atccttttaa agaattgga gctactgtyg gttacctct gtttctcctc 60
atgactcct ttccttaaagt acgatattc ctgatcctg cccggtctg gatgtcttct 120
gcggggtt ctgcggcgc ctgcggcgc ctgcggcgc gcgggggtt gcgggggtt 180
tgatggggag acggagtga tgtgccggag ggggggtg ggggggtg ggggggtg 240
gcgggggtt ctgcggcgc ctgcggcgc ctgcggcgc gcgggggtt gcgggggtt 300
tgggtggttc gcggagtgga gcggagtgga gcggagtgga gcggagtgga gcggagtgga 360
ttgctgggg atggctcgtt gcggagtgga gcggagtgga gcggagtgga gcggagtgga 420
tggctgggg atggctcgtt gcggagtgga gcggagtgga gcggagtgga gcggagtgga 480
tgcggggtt gcggagtgga gcggagtgga gcggagtgga gcggagtgga gcggagtgga 540
tgcggggtt gcggagtgga gcggagtgga gcggagtgga gcggagtgga gcggagtgga 555

<210> SEQ ID NO: 42
<211> LENGTH: 183
<212> TYPE: PTY
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic HBcAg

<400> SEQUENCE: 42
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu 1 5 10 15
Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp 20 25 30
Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys 35 40 45
Ser Pro His Thr Ala Leu Arg Gin Ala Ile Leu Cys Trp Gly Glu 50 55 60
Leu Met Thr Leu Ala Thr Trp Val Gly Val Asn Leu Glu Asp Pro Ala 65 70 75 80
Ser Arg Asp Leu Val Ser Tyr Val Asn Thr Asn Met Gly Leu Lys 85 90 95
-continued

Fhe Arg Gin Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg 100 105
Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr 115 120 125
Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro 130 135 140
Glu Thr Thr Val Val Arg Arg Arg Ser Pro Arg Arg Arg Arg Arg Thr 145 150 155 160
Pro Ser Pro Arg Arg Arg Ser Gin Ser Pro Arg Arg Arg Arg Arg Ser 165 170 175
Gln Ser Gin Arg Glu Ser Gin Cys 180

<210> SEQ ID NO: 43
<211> LENGTH: 2043
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE: OTHER INFORMATION: Influenza A Virus NP Gene Fused to Synthetic HBcAg
<400> SEQUENCE: 43

atgytgtcct aaggyccaca aagytcttac gaaegatgy agaegtgyg gaagogcgcga 60
atgygacctg aaggygagag aagytgagga gaaegaggy agaeggyctac 120
ataacatgt gaaaggycaag caacagcaat gatgyttgyg gagaagctc gacaagagctc 180
ttaaagatgg agaagatgyg gatgygtgcta tttgaagcag gagaagactaata atagctcaga 240
gaaagctgca gtagygcag agaagctgca aagaaaagag caagaaaagag 300
aagcgagag gatgygagca atacgttgtc ttagggagag aagagagag 360
cagccagac atagctcttg gatgygtgctag atagccagag atagctcttg 420
tocacttgttc atagcagctag cagcagctag aagtggcttg gactgttttag 480
cagcagccag atagctcttg gacagctgctat agagcagtgctag agagcagtgctag 540
gacagctcag aagggagag atagctcttg agagcagtgctag aagggagag 600
atagcttctag gacagctgctat agagcagtgctag aagggagag 660
agagctcgca aagagctcgca aagagctcgca aagagctcgca aagagctcgca 720
cagcagctgctat agagcagtgctag aagggagag 780
cyggctctgc atagcagctgctat agagcagtgctag aagggagag 840
ttagggctag gatgygtgctag atagcagctgctat agagcagtgctag 900
atagctctgc atagcagctgctat agagcagtgctag aagggagag 960
aatagcagctgctat agagcagtgctag aagggagag 1020
cagcagctgctat agagcagtgctag aagggagag 1080
agagcagctgctat agagcagtgctag aagggagag 1140
cagcagctgctat agagcagtgctag aagggagag 1200
agagcagctgctat agagcagtgctag aagggagag 1260
agagcagctgctat agagcagtgctag aagggagag 1320
agagcagctgctat agagcagtgctag aagggagag 1380
agagcagctgctat agagcagtgctag aagggagag 1440
AGATGAATG AAGATGTTA TTTCTTGGGA GCAATCGCAG AGGAATACGA TAAATTTG
ATCGATTTT ATAAAGAATT CCGAGTACTT GTGGTACAC TCTGGTTTCT TCCGAGTACG
TCTTTTCTT CAGATGAGA TTTTTCTTTG ACGAGGAGG CGTGTGATAC GGAGGCTTTG
GGATGCTCT AGCAGGGCAC CCGAATCATC CGACGGCTA GGAAGCAAC TCTTTGCTGG
GGGAGCTCA TGAACGGCGC CGTGCTGACG CTGGTACAGC ACGTACAG
GACCTGTAG TCGAATATG CAACACTGTTA AGTGGTTTCA AGTGGTTTCA AGGGAATACG
TCTTCCCTTA GACACATTTGG GACACATTTGG GACACATTTGG GACACATTTGG
CAGAATTGAA CAGAATTGAA CAGAATTGAA CAGAATTGAA CAGAATTGAA
GATACCACGG AAAGAATCCT CTCAAAAGA ACAAGAAGAG ACGAATACG
GAGCGAAGG AAATAAAGACG CCAAGACA CGAAGAAAGA ATCCAGGCAT GCTAAGAC
GAGCGAAGG AAATAAAGACG CCAAGACA CGAAGAAAGA ATCCAGGCAT GCTAAGAC
GAGCGAAGG AAATAAAGACG CCAAGACA CGAAGAAAGA ATCCAGGCAT GCTAAGAC
GAGCGAAGG AAATAAAGACG CCAAGACA CGAAGAAAGA ATCCAGGCAT GCTAAGAC
GAGCGAAGG AAATAAAGACG CCAAGACA CGAAGAAAGA ATCCAGGCAT GCTAAGAC
gtccgcgta agggcagt aagaggatt gggccgctc ttagtgctat caagctcag
1320
tttgagccc caaagccag atcttgaggg aatgaagta caaacgcgg aggtatctgt
1380
caaattgtt gcagoccttg gtttcgcta aacagacca cttccttact caagcagct
1440
gtaagcgact tgctgctaat gacgcttag gcacgctag ctagaattca aagaaatcta
1500
cctcaaatga tgaatagcct atggcacaag aacacccag gaaacctttt cttggtggaag
1560
aataagttt caatatcga caaaaaa accatactt cttgagttcttc aataagct
1620
acctccacca attttttttt tggaggggac acagcgaggg atatagtgat cctcgattat
1680
atgatagct atcttcataa aagattctcg gtaactctgg agttactcttt gttctctcog
1740
agttactctt ttcctcctag aagagatctt ctggatcaac cagagctgt gtaagggga
1800
ggcttgaggg cttcgagacact ctagctaga caacatactg cctcctcttc cagaaacctt
1860
tgctgggggg gtagcatact gctggccacag tgggtgggtg ttatactggca agataccagtt
1920
agccgggcc tggtagttga tttggcacc actatcttga gtttagctt caagccgctc
1980
ttggtgtttt aacctacgct ctctctcttc gcggcggaa cacacctga caatctgttg
2040
tttcctgag tttgagcaac caacactct cattactgct aacagaccaacc cttactgctt
2100
tcagccagact cttctcact ctagctagtt cctgagacag ctaaagcag tccagatacct
2160
cctcgccct ggcggccag gcctcacctg cccgggctg agacagtctc atgcggggga
2220
ttcctagtgt
2280

<210> SEQ ID NO 45
<211> LENGTH: 1305
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Influenza A Virus M1 Fused to Synthetic HBoAg
<400> SEQUENCE: 45
atgatagctt taaagaggttt cgaacagac caaacagcta tcttacctga aaccccttccc cggccccc 60
aacagccgct tggagccag actgagact gtaatttgtag ggcgaacac cggatctctgt
120
atttggaggt aagatgtaaa gaacacgct cttgtagata ggggttttaa
180
agttatgttct ctaaagagact cagagctgt gcgcggctgt gtagctgtag acgctttgtc
240
caaaagctc ttaagccgac ggggagcttt ctaacttctg caacacgact 300
agggaggttta agggagatgt acattctttac ccctatttgg ccctattttc
360
ggtggtccct tgtctgacgtg ttaagagctc atattaca aaccgagacag gcatttacg
420
gnagagctgt gggcgttcct gtgttacctc ctgaaccgct gcagctagtctcct ggttctctgtg
480
ttcctagccgtt aattgagctt ccaacctttt caacctatcc gcatctgatt gagagttttng
540
tccagacta aaccagcttt gtagctggga caaagatgta ctaagctca gcacagaaa
600
ggggaggttta gtagctgact tgtctttttgc cccttttgtc gaccagatgg
660
ttacctgttc cgtctagagtc aaccagctttt aaccagctttt aaccagctttt
720
cgcctgagctt cttcagagtt gctttttatt cttcagagtt gctttttatt 780
ttcagagtttt cattctttgc gctttttatt cttcagagtt gctttttatt 840
gagagagtttt aaccagagtttt aaccagagtttt aaccagagtttt aaccagagtttt
900
gagagagtttt aaccagagtttt aaccagagtttt aaccagagtttt aaccagagtttt
960
---continued

gccacgctgg tggctgttas ctctggaagat ccacgtagca gggacgtcggt atgctgttat 1020
gtcacacta atatgtggtt aaagtcggc caatctttct ggctttgccat tgytgyggtc 1080
accttgcgc gagaaaaagct tcgagagatt tcgagaggtt tcgagaggtg gatcgcgacat 1140
cctcagcctaatgacgcc gaaatcagc actcttgtaa gactcagcgc gactcagttt 1200
gtgatcgc gcggcgagct cacatagaga gaaatcagcct gcttgacgcag gcggagttct 1260
cattcgcgag gcgagcgcag ctatccattct gctggaacctc actgtgtgct 1305

<210> SEQ ID NO: 46
<211> LENGTH: 1581
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Open Reading Frame for TPANP from VR4700
<400> SEQUENCE: 46

atgagatgca taagagagag ctcctagcgt tagctgtgag tcytgcagcg tcytgcagcg aagctctaggt 60
tgccgcctag ctcagagcg tggctagcct tccgaccgca cccagatgcgt tcagagcagc 120
tatgagatgct atagagagcg ccagatgcgt actgacacca gacacaggt cagaaaaatg 180
atgagagagc ttaagagagct atacatcccag atgagagagcg aagctctagct cagaaaaatg 240
gcgagtagcg atcgccagcc cagctagcct atagagcgct tcytgcagcg tcytgcagcg 300
gaaagagagc aatcctagcct gaaagagagc ccagagagcg aagctctagct cagaaaaatg 360
gcagagagagc atagagagagc aagctctagct cagaaaaatg ccagctagcg ctatccgtct 420
agacgagagatg gatcagcgct ctcctagccct tccgaccgca aagctctagct cagaaaaatg 480
tcagctagcct tgcgtagcgc gatcagctagcct tccgaccgca aagctctagct cagaaaaatg 540
gcagagagagc gatcagctagcct tccgaccgca aagctctagct cagaaaaatg ccagctagcg ctatccgtct 600
tcaagagagatg gatcagcgct ctcctagccct tccgaccgca aagctctagct cagaaaaatg 660
tgagagagagc tttgagagagcg tggccagagc tttgagagagcg tggccagagc tggccagagc 720
gagagagagc tggccagagc tttgagagagcg tggccagagc tttgagagagcg tggccagagc 780
gagagagagc tggccagagc tggccagagc tggccagagc tggccagagc tggccagagc 840
tgagagagagc tggccagagc tggccagagc tggccagagc tggccagagc tggccagagc 900
agctagctagcct tgcgtagcgc gatcagctagcct tccgaccgca aagctctagct cagaaaaatg 960
agagagagagc aatcctagcct gaaagagagc ccagatgcgt actgacacca gacacaggt cagaaaaatg 1020
tgagagagagc gatcagcgct ctcctagccct tccgaccgca aagctctagct cagaaaaatg 1080
tgagagagagc gatcagcgct ctcctagccct tccgaccgca aagctctagct cagaaaaatg 1140
tcctagccag gcggctttgcct tcgagagcttc actctagcgt ggctgaagttta cctcagagct 1200
atgagatgca taagagagag ctcctagcgt tagctgtgag tcytgcagcg tcytgcagcg aagctctaggt 1260

gccacgctgg tggctgttas ctctggaagat ccacgtagca gggacgtcggt atgctgttat 1020
gtcacacta atatgtggtt aaagtcggc caatctttct ggctttgccat tgytgyggtc 1080
accttgcgc gagaaaaagct tcgagagatt tcgagaggtt tcgagaggtg gatcgcgacat 1140
cctcagcctaatgacgcc gaaatcagc actcttgtaa gactcagcgc gactcagttt 1200
gtgatcgc gcggcgagct cacatagaga gaaatcagcct gcttgacgcag gcggagttct 1260
cattcgcgag gcgagcgcag ctatccattct gctggaacctc actgtgtgct 1305

---
<210> SEQ ID NO 47
<211> LENGTH: 333
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Open Reading Frame for TPAN2 DeltaTM from VR4707
<400> SEQUENCE: 47
atgagatcga tgtgaagaggg gcctctcgctgt tgtctgcgagc agcttctagt 60
tcgccagcg cgataaagtgg gaggatctaag ctttctgccag aggctgaaac cccatcaga 120
acagcagaggg ctgctgagag caacagttca agtacctctg gcgcgagcgc gcgcgctttt 180
ttcasatcga ttttccgycgc ctttcataac gcgtctgasa cacagcttcc tccgaaaga 240
gtcgcagaggt ctgctgagag aggctgaaac aggctgctgt gcgtcttgag 300
gatagcaggt tgtgtctgcat gcgcgctgag taa 333

<210> SEQ ID NO 48
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify TPAN2 Fragment
<400> SEQUENCE: 48
gcgcagcctca tgtgtgcaat gaag 24

<210> SEQ ID NO 49
<211> LENGTH: 36
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify TPAN2 Fragment
<400> SEQUENCE: 49
gtgccctgg gacgcctaag cacttgcaag gttgca 36

<210> SEQ ID NO 50
<211> LENGTH: 36
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify NP Gene
<400> SEQUENCE: 50
tgcasatcgt caaagtggat gcgctgcacc gcacc 36

<210> SEQ ID NO 51
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify NP Gene
<400> SEQUENCE: 51
gcgcgacct taattcttgagt actc 24

<210> SEQ ID NO 52
<211> LENGTH: 1653
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Open Reading Frame for TP53BP2 from VR4710

<400> SEQUENCE: 52
atgcgatgca tgaanggagc gctctgcttg tgcgctgctt aagggtcagc 60 
tgcggtcgct tgcggtcgct ggtcggtcgct tgcggtcgct ggtcggtcgct 120 
aaacgattcagcg gtagctgagc ccggataggc ctgcccagc 180 
tggggtggagc tggggtggagc tggggtggagc tggggtggagc tggggtggagc 240 
gtgggcggagc ggagcggagc ggagcggagc ggagcggagc ggagcggagc 300 
cctgcgctgagc gctgcgctgagc gctgcgctgagc gctgcgctgagc gctgcgctgagc 360 
tgggtggagc tgggtggagc tgggtggagc tgggtggagc tgggtggagc 420 
cctgcgctgagc tgggtggagc tgggtggagc tgggtggagc tgggtggagc 480 
actcgcgtt acacatgacg cccagacggt cccagacggt cccagacggt 540 
gccggtggc ggctggtgcttg ttgctgcttg ttgctgcttg ttgctgcttg 600 
cctgcgctgagc tgggtggagc tgggtggagc tgggtggagc tgggtggagc 660 
gctgcgctgagc gctgcgctgagc gctgcgctgagc gctgcgctgagc gctgcgctgagc 720 
gtgcgttgc tggctgctgagc gctgcgctgagc gctgcgctgagc gctgcgctgagc 780 
ccgcgtgcttg gttgcgtgcttg cggctgcgctt ccgcgtgcttg gttgcgtgcttg 840 
actcgcgtt acacatgacg cccagacggt cccagacggt cccagacggt 900 
gctgcgctgagc tgggtggagc tgggtggagc tgggtggagc tgggtggagc 960 
tgggtggagc tgggtggagc tgggtggagc tgggtggagc tgggtggagc 1020 
tgcgctgctgagc gctgcgctgagc gctgcgctgagc gctgcgctgagc gctgcgctgagc 1080 
aacggtgcttg gttgcgtgcttg cggctgcgctt ccgcgtgcttg gttgcgtgcttg 1140 
gttgcgtgcttg cggctgcgctt ccgcgtgcttg gttgcgtgcttg cggctgcgctt 1200 
gggtggagc tggggtggagc ggtcgtggcttg cggctgcgctt ccgcgtgcttg 1260 
gggtggagc tggggtggagc ggtcgtggcttg cggctgcgctt ccgcgtgcttg 1320 
gggtggagc tggggtggagc ggtcgtggcttg cggctgcgctt ccgcgtgcttg 1380 
gggtggagc tggggtggagc ggtcgtggcttg cggctgcgctt ccgcgtgcttg 1440 
gggtggagc tggggtggagc ggtcgtggcttg cggctgcgctt ccgcgtgcttg 1500 
gggtggagc tggggtggagc ggtcgtggcttg cggctgcgctt ccgcgtgcttg 1560 
gggtggagc tggggtggagc ggtcgtggcttg cggctgcgctt ccgcgtgcttg 1620 
gggtggagc tggggtggagc ggtcgtggcttg cggctgcgctt ccgcgtgcttg 1680 
tgcgctgctgagc gctgcgctgagc gctgcgctgagc gctgcgctgagc gctgcgctgagc 1740 
<220> SEQ ID NO 53
<211> LENGTH: 35
<212> TYPYE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify the HA Gene

<400> SEQUENCE: 53
GGGGTGGAGC CGCGCGC CGCGCGC CGCGCGC CGCGCGC 35

<210> SEQ ID NO 54
<211> LENGTH: 26
-continued

<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify the HA Gene

<400> SEQUENCE: 54

ctctctact accatgcaaat gttgca

<210> SEQ ID NO 55
<211> LENGTH: 1701
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Open Reading Frame for HA H3N2 from VR4750

<400> SEQUENCE: 55

atgagagca tcctgtgttt gggttgtctt tctgtgtctg gttggtgca aagctttcga  60
 gagagtcga aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 120
cagttgaaaa caatccgacg tgctgcgctt gacagtacgt aagctttcgt gctgcggt 180
agctctctca cggggttaat atctgcaat cctctgctgaa tctgtgattgt atctagactctg 240
acacatgtag atgtogtacatt gggggactcc ctttggatttt tcggccgtaa tgggactccg 300
gactctctcc ggacggtccagg cggcgttcgtt gcctcgtcggat gctgctgctgat 360
tgactcctc ggacggtccagg cgctcctcggat gctgctgctgat gcctcctcggat 420
ttgggcattc cttcggagcag cgggctctcc gcctcctcggat gctgctgctgat 480
agctctctcc ggacggtccagg cgctcctcggat gctgctgctgat gcctcctcggat 540
acacatgtag atgtogtacatt gggggactcc ctttggatttt tcggccgtaa tgggactccg 600
cagagagca aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 660
tatcctaga gcagagcgac cggggtattc ggcctcggag tggggtgca tggggtgca 720
agctctctcc ggacggtccagg cgctcctcggat gctgctgctgat gcctcctcggat 780
atcctacgt gcagagcgac cggggtattc ggcctcggag tggggtgca tggggtgca 840
agctctctcc ggacggtccagg cgctcctcggat gctgctgctgat gcctcctcggat 900
acacatgtag atgtogtacatt gggggactcc ctttggatttt tcggccgtaa tgggactccg 960
tatcctaga gcagagcgac cggggtattc ggcctcggag tggggtgca tggggtgca 1020
gagagagca aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 1080
gagagagca aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 1140
gagagacg aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 1200
gagagagca aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 1260
gagagagca aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 1320
gagagacg aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 1380
gagagagca aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 1440
gagagacg aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 1500
gagagagca aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 1560
gagagacg aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 1620
gagagagca aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 1680
gagagagca aaccagagcg aacggtgct gcggagagcc atggggtgacc aaccggaacc 1701
<210> SEQ ID NO 56
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify the HA Gene
<400> SEQUENCE: 56

```
gggctagccc cggacacatt aagggaaacc taccg  
```

<210> SEQ ID NO 57
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify the HA Gene
<400> SEQUENCE: 57

```
cctgcgacc agatgcaat ctcgca  
```

<210> SEQ ID NO 58
<211> LENGTH: 1701
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Open Reading Frame for HA H1N1 from VR4752
<400> SEQUENCE: 58

```
atgaagccaa acactctgtg cctgcttattgt gcacttgcgg ctgcaagatg agacacaaata  
tgtactgtgc acatgtgcag caattccacgc gacacttgcgg acacactgtg cggatgatat  
gtgcctgtg cccactgtgc taaacctcgc gcagccgacc cccaggtgac atatgc tatactgc  
ttagccagac ttagccaggt gccagcag gtgtcagcg caaacacacacg acacagcag gcgttagctgtgc  
ttaggaaaa  
ttagctgcag  
ttagccagag  
ttagccccacacag  
ttaggcagc  
ttagggagcc  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagcagcag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
ttagccagag  
ttagccagag  
ttaggagcag  
```

26
---continued

atggaatatt taataatnac aagttgatgt ggatctcttg aacctttggac atataatgca 1320
gaccttttag ttcctctggga aaaaagaaag cctctttgagtt ttcctgctc taaaatgaa 1380
aatctacgat gaaacaatga aasaatatac ccaagaaaat cggaaatgga 1440
tgttctgcct ctctccacac ggtgacactg gaaagttgaa aatggcactt 1500
taggataat ccaaatatca aagaaatgca aagttgacca agyagaaatt gatgagcgatg 1560
aatggaatat caggttggct ctcacgcatt cctgctcct ctggtcttt tctatgtaca 1620
cagggcttt tctgttccct gggcgctcct cgttctcttg ccctttcttg ccgttttcttg 1680
cagtgcgacg tctgatctcy a 1701

<210> SEQ ID NO 59
<211> LENGTH: 1050
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Open Reading Frame for the M2M1 Fusion from VR4755

<400> SEQUENCE: 59

atgagaactc tggacgcgctt ggagacacac cccgagccaa aaggtttttt gatgtgcccag 60
gcttcgtgg gccttgcacac gctgatctcy aatgttttctg gctttcttt tctatgtaca 120
tgttctgcct ctctccacac ggtgacactg gaaagttgaa aatggcactt 180
agagctcagug aacgagccag cgctttcctt ccctttcttg ccctttcttg ccgttttcttg 240
caagctcacc tgggctccct cggcctcct cgggctccct cgggctccct cgggctccct 300
cagaacctg cggcctcct cgggctccct cgggctccct cgggctccct cgggctccct 360
tagttcctca tagtggacac cccgagccaa aaggtttttt gatgtgcccag 420
gcctttcttg gctttcttt tctatgtaca

<210> SEQ ID NO 60
<211> LENGTH: 982
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Open Reading Frame for Fragment 7 from VR4756

<400> SEQUENCE: 60

atgagaactc tggagagggt ccmaacgtat gttttctctta tggatcctcttt aaggctccct 60
---continued---

```
aaagcgaaga tcgacagag acctgagct gtcttgcgtg gaaanaaac acagccttgag 120
gctctcagtg aatgctcta gaaagaacca atctctcagc atctgactaa gggatatttg 180
gggtttgtg tcacgctcac cgctgcaagt gacgcaggac tgcagctag acgcgtgtgc 240
caaatggcc tcagggggag tggagagac caataactgg acagacagct taaccttatat 300
agaanctgta agagggagat tacattcctc gggccgaaag aatacgctc agtatattct 360
gtgcagcag tcgagcagct acagacgctc atacaaacca gcagggagggc tgtcacaact 420
gagctggtctt ttcgcgcaca ttcgacgaca ttgcgctccc gccaacgagg 480
ttcctagcgc aatgctgctg gacccaaa acataaacca gctgtctggc gaaatgggtt 540
ttcgacgca ctacgcttc agctagcag ctaactgcag gaattggag gcagacgggtg 600
gagacaggac aatgctgagt cagcgcgag cagagctggc acgccttgag gacccctggg 660
actataccta gctgagcgcgt cgcttctata gatggctccg tcgacgcttc gcaagcagta 720
cgaacgcggg tgtcagcgag gcagacgcca ttcgagcagctcgctttgtg gttctgcag 780
tctcctgtg acatgctcag gactttcgt tcgacttctct ctaacatcgcttctctgcag 840
ctacagctc ttcacaaacac gttcttcgag gccagcctct acgcgcgggg gacttgagct 900
tgttgaggg gaaatagcag gcagacgagc gcagagctgg ctacgagcatc agatgtcatctt 960
tgctagcata gacogcggagt aa 982
```

<21D> SEQ ID NO 61
<21L> LENGTH: 982
<21D> ORGANISM: Artificial sequence
<22D> FEATURE:
<223> OTHER INFORMATION: Codon Optimized Segment 7 from VR4763
<40D> SEQUENCE: 61

```
atgcgctcg tcgagcggagt cgaacgtatg tttctctcta tgcgcgcaag cggcccccttg 60
aagcgcgaga tcgacagag acctgagct gtcttgcgtg gaaanaaac acagccttgag 120
gctctcagtg aatgctcta gaaagaacca atctctcagc atctgactaa gggatatttg 180
gggtttgtg tcacgctcac cgctgcaagt gacgcaggac tgcagctag acgcgtgtgc 240
caaatggcc tcagggggag tggagagac caataactgg acagacagct taaccttatat 300
agaanctgta agagggagat tacattcctc gggccgaaag aatacgctc agtatattct 360
gtgcagcag tcgagcagct acagacgctc atacaaacca gcagggagggc tgtcacaact 420
gagctggtctt ttcgcgcaca ttcgacgaca ttgcgctccc gccaacgagg 480
ttcctagcgc aatgctgctg gacccaaa acataaacca gctgtctggc gaaatgggtt 540
ttcgacgca ctacgcttc agctagcag ctaactgcag gaattggag gcagacgggtg 600
gagacaggac aatgctgagt cagcgcgag cagagctggc acgccttgag gacccctggg 660
actataccta gctgagcgcgt cgcttctata gatggctccg tcgacgcttc gcaagcagta 720
cgaacgcggg tgtcagcgag gcagacgcca ttcgagcagctcgctttgtg gttctgcag 780
tctcctgtg acatgctcag gactttcgt tcgacttctct ctaacatcgcttctctgcag 840
ctacagctc ttcacaaacac gttcttcgag gccagcctct acgcgcgggg gacttgagct 900
tgttgaggg gaaatagcag gcagacgagc gcagagctgg ctacgagcatc agatgtcatctt 960
tgctagcata gacogcggagt aa 982
```
<210> SEQ ID NO: 62
<211> LENGTH: 1569
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Open Reading Frame for eNCP Codon Optimized by Contract

<400> SEQUENCE: 62

agtgaacctg tcaotgcaag cagacacca atcacaacag aatgygggtg tgaatgcaac 60
gatagcctag aacgagcctc cccgggaaco aaaagacagt atgaacagat ggagactgac 120
ggagcgagc agaaagacacg agagcagaga gctagtgtag gaaagagtag agacggtatc 180
gggggtatt gcctctctac tgttaagaga tggcagagct gcggttagtc agcagactat 240
tacagacact cactcactat tggagctagt gatcctagc catctgtaaa aagaagagaat 300
tagctctgtt aacacaccc cagccgcccc aacgctcaac aagagactcg aggcccacat 360
tacagacag gtaacgttta gttgatgag aagcgtgaag tatgctagaa aadagaagat 420
gagagactct gggagcagao aacacagga gggcagcata cagctgaggct gacccactag 480
agtgtggc atatagacat gatgatacag aacatcagcg gggagcagag ccctgcctgc 540
agacggagtc atcccagcag atggctacag tggcagagca tggagcagag cagagcagta 600
ggagcggcgc gtggccagcgc gagaagacag ggcaagcagc tataagagta gtaagcattt 660
atcagataag gctgagcagc cacaagcagc cagaagcagc aggaggtgtg 720
agtgctgtg aacaggtgtag cactgtttagc aagagagag ctagaagacg 780
ggcctgtac atcagcgttg aagaagcagc aoccaacagt tataagagta gtaagcattt 840
atatccctgg ccgcagcagc tccctctactt aagaggtagc tggcagagca aagagcagtg 900
ccccagctgg tatacgctgc ctgtctgtct ctggagctag attttgaaaa agagggatt 960
tctttggtg gatgagcagc tttttaggg ttcagacact cagaggttg cagggtgatt 1020
agacacagc aagacacgag ccacaaacca cagctgtgtg gaggacagc aoccaagcg 1080
gccctgag ccctgcagag ctggctacat attagcagc ctagaagacg cccagggaga 1140
aacctgagc cgcggcggag acatcagcag ttcagacact atagctagag ttaagagctt 1200
agccacagct gtagcagcag atgagctagc gctttagaga cagcagaggc aggggaaccc 1260
aacggcgaga gacacagcag ccagcagagc atccacatgct acaaagacacag 1320
aacctgagc ttcaggagct acagctgtgg gcgctattta ctggcagacag cagagcagc 1380
acacgacagc tggagcagcag gattttgag atggagctag gaccaaaaaagagaggttt 1440
tctttgag gaaagggagt ctgctagcag ttcagacact aagccacaaa cccctagaga 1500
cctttgag ccctgagag ccctgagagg ctttggctag ctgagcagcag 1560
gacagagc 1569

<210> SEQ ID NO: 63
<211> LENGTH: 1569
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Open Reading Frame for eNCP Codon Optimized by Applicant
-continued

atgcgctcgc tgacgccggtt gggaacccc accagaaacy ctaggggtcg cagatgcacc 60
gacagcagc acatgyccgac ccaggyccac aggagacagc agcagagcagc ggagcaccgc 120
gggcagagc agaacgcacc cgggaagcag ggcagctggg gcagagagct gcagggcacc 180
gggcagatct cagctccgag ttgcacagg ccgaagcctg ggcacactgc gggcagactg 240
atccacagac gctgcacact ccagagactg atgcctgagc ccctgcagaca gcaccagacc 300
agacctgctg ggagacccc cagcgcggcg acaggacccc aagagacccg ggcccacactc 360
taacgcagcg tggacgggaa gctggtggcg ttgacggcag ccgacagcac ggagagagac 420
agacagaccc cccaacccgc gcaggccgca cggcgcggcgc cccgccccag gcaccacactg 480
atggctgctgcc agacggcacc ccgagcgcac acctggccgac gcacagagcag acctggggcg 540
acggagccg agccgcagtt tggagcgttg atcgccggcag gcacgcgtcgc cggcagacgc 600
gggcagccgc ggcgccggtt cggagcctgg ggcaccagctg ttgagagact gcacacacgc 660
atccacagac gctgcacact ccagagactg atgcctgagc ccctgcagaca gcaccagacc 720
agcgcacctgc agacatgctg cccctcgcag acgggacagt tccagacgcag ccgccacagc 780
gcagctgatg acagctgctgc gcagaagacgc ccacccgagca agcccgagct gcacccacgt 840
attocgtcgc ccagcgagcc atccagctgc agacggcagc ccgccagcac gcacacagctg 900
ccgccgctgc cgccggctgc agcggtgcag acttcagac gcggggcgtc 960
agcgcacctgc gtcgctgctgc tggagcgttg acgggacagt tccagacgcag ccgccacagc 1020
agccacggag cccagaccgc cccagacgcc ccgctgggtg gcgggctcg cccagacgcc 1080
ggcctgagc agcagagcag cttgagcctgc atcagacgcc ccagcttgcc cccagacgcc 1140
agcgcacctgc gcagagagct cggagcgcag cccagacgcc cccagacgcc ccgctgggtg 1200
agccacggag cccagaccgc cccagacgcc cccagacgcc cccagacgcc 1260
aacgcggtcgg ccagctgagt ccgctgtgtc ccggtacgc cggggtggtgc 1320
aacgcggtccgg cgggtgtcgt cgccggttgc cggggtggtgc 1380
aacgcggtctgc cgggtgtcgt cgccggttgc cggggtggtgc 1440
GGGTGAGGC cgggtggtcgt cgccggttgc cggggtggtgc 1500
aacgcggtctgc cgggtgtcgt cgccggttgc cggggtggtgc 1560
gacacagtg 1569

<210> SEQ ID NO 64
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify the M2 Gene
<400> SEQUENCE: 64
gccgaacctgc ccagagcttgcc cctgcagacc 30

<210> SEQ ID NO 65
<211> LENGTH: 33
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify the M2 Gene
<400> SEQUENCE: 65
ggcgaacctgc ccagagcttgcc cctgcagacc 33
-continued

gccctgagt gataaactca gctcagatgt cacc 33

<210> SEQ ID NO: 66
<211> LENGTH: 294
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Open Reading Frame for M2 Gene from VR4759

<400> SEQUENCE: 66
atggtgctgc tgtctgccgt ggsagcagccc atccagaacgc gctggggctgc ccactgcac 60
gacagcagcc aacccgctgt gcgtgctgcc gcacgatccg gcctgggttg cccgatgttgc 120
tgacgtctgg acagatctgt ttctcaagcc acctcagac gtttcagacc ccggctggtaa 180
gagagcagcc gcacagcaggg cctggcggag agcagatccg agcagatccg aaagagagcag 240
cgaaagccgc tggatggtgga cggactgcac ttcctgtgac tgcagtggaga gttga 294

<210> SEQ ID NO: 67
<211> LENGTH: 36
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used Amplify M1 Gene from VR4755

<400> SEQUENCE: 67
ggcagatccgc cccctgccgc cctctgtgaca gsgtgc 36

<210> SEQ ID NO: 68
<211> LENGTH: 31
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify M1 Gene from VR4755

<400> SEQUENCE: 68
gccctgagt gataaactca atctctcgat c 31

<210> SEQ ID NO: 69
<211> LENGTH: 759
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Open Reading Frame for M1 Gene from VR4760

<400> SEQUENCE: 69
atgctctcgg tgtccagact ggaacatcac gtgtgctgcgtc cgggctgccg cggcccctgc 60
gcccgctgc tggctgccgt actgggact gcctgtctgc gcagagacgc cgcctgggag 120
gggtcctgg actgctgtaa gaccagaccc atctctgagcc cccgctgacc gcgcatctgc 180
ggccttgtc tccctcctgc cgtggcgtcc gcagagacgc tggagagactg cagatcgctg 240
cgaaagccgc tggatggtgga cggactgcac cccgctgacc gcgcatctgc ccggctggtaa 300
gggaagtagt gcaccctgca gcgcctgacc ggccctgacc gcgcatctgc ccggctggtaa 360
ggcctgacc gcagagacgc tgtccagact gcgcatctgc ccggctggtaa 420
ggcctgacc gcgcatctgc ccggctggtaa 480

cgcggactgc gcggctgccg cggcatctgc ccggctggtaa 540

cgcggactgc gcggctgccg cggcatctgc ccggctggtaa 600
---continued---

```
gagtgcctgg agatgacagc ccagccgag cagaggtgag aggcatgag acgtgatgg 660
gacctgacag ccagagggag ccagctgagc tggagaaacg gcagagcct 720
gacagaggg acgtgactg ctagtcggaa tcggagatga tccagatg 759

<210> SEQ ID NO: 70
<211> LENGTH: 38
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify NP Gene from VR4757

<400> SEQUENCE: 70

gcggagtcccg ccacatgagg ctcccaaggg aacaaaaa 38

<210> SEQ ID NO: 71
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify NP Gene from VR4757

<400> SEQUENCE: 71

gcgcgtcat atgcaattgt gtcaccttct 30

<210> SEQ ID NO: 72
<211> LENGTH: 1497
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Open Reading Frame for NP Codon Optimized by Contract

<400> SEQUENCE: 72

atggcctccc agggccacaa aagactaatg gcacagatgg agaactgagc aggagacag 60
gacagccgag agatgacagc tggagaggc agactgatag ccgctattgac 120
attcaggat gtaacactg cacaattgg gcacagatgg gcacagatgg 180
ctcagcatg agcaggttag actgacagc tggagaggac gagaagattg catcagcga 240
ggaccacgc gcggtcagcc aagactgagc agaactgagc gcacagatgg 300
gacgcgtcat atgggttagc gcagagttg aagactaatg gcacagatgg 360
agggccacaa aagactaatg gcacagatgg gcacagatgg gcacagatgg 420
agcagagc acgtgctcag ctccagccc cccggtcttgc gggatcggc 480
cccggtcttgc gcgtacatcg gacggctgc gggcaaggg gcgggtaggc 540
cgcagtggag cggctgctc aagactaatg gcacagatgg gcacagatgg 600
attcagagc gaaacctctg ggcagagga aatcagagc ccacagagc 660
cgtgccgta atgacatcat gacagacagc ccaagatcag ccacagagc 720
cagatcag gcagccgctc aagatcag ccacagagc ccacagagc 780
cgcgtcatc atgacagta ccaagacagc cgaacatcag ccaagatcag 840
tatgctgcag ctgctgcag ccaagatcag cgaacatcag ccaagatcag 900
atgatcagc atagatcagc ccaagatcag ccaagatcag ccaagatcag 960
acgcgtcatc atgacatcat gacagacagc ccaagatcag ccaagatcag 1020
ctgatcagc atagatcagc ccaagatcag ccaagatcag ccaagatcag 1080```
-continued

cgaggatgac gatagcagtc taccgaat atggatata tgggatctag caacactgaa 1140
ttgggctac gatacgggg cttatacnga cggacgagg gcacaccaac cagacagaga 1200
gactcggcc gctgaataag cgccgagct aatcttcctag tacaacggaa cctgccatatt 1260
gaaagagtag cagctagggc cgccttact gcgaacaccc agggaagcac acggacactg 1320
gagcagaga tttttagaat gatgggaagga gctacaaccc agaggttcc attagaggaa 1380
agggaggtg tggtaatgtc cgagagaaaa gcacacactc ccatactgcc tagcttgagc 1440
agttcaccgc aaggttctta ctcttttctgt gcacagtgcg cagagtagcga caattg 1497

<210> SEQ ID NO: 73
<211> LENGTH: 36
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<222> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify NP Gene from VR4758
<400> SEQUENCE: 73
gccgattcc gcaccaggtgc cagcagggc accag 36

<210> SEQ ID NO: 74
<211> LENGTH: 28
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<222> FEATURE:
<223> OTHER INFORMATION: Primer Used to Amplify NP Gene from VR4758
<400> SEQUENCE: 74
gctgctgac gatgctgtgt cgtaccc 28

<210> SEQ ID NO: 75
<211> LENGTH: 1497
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<222> FEATURE:
<223> OTHER INFORMATION: Open Reading Frame for NP Codon Optimized by Applicants from VR4762
<400> SEQUENCE: 75
atggccagcc aggggacaac gcagagcata gcagctgtag cagacagcgg cgcagacag 60
aacgcacagg agagcagagc caccctgagc cagagcagct acgcctcctg cggattctac 120
atcggcagat gcacagctga gcgacgacag gcacagctgag cccagacagc 180
gctggcagac gcagagctaag gcctgagcc gctggagag cctgcagagc 240
gacccaccca ggacccacca gcgccgcagc gcgctggcag gccgctgagc gcgccctcct cggacagctg 300
gacgctgaat gcagtcagcag gcgacgagc gcagcagcag cgcagcagcag cgcagcagcag 360
agacagcagc aaccacagga gcacagcagc gcggcagcag cccagacgat gctggagag 420
agcgcattta agacacacg cttcagagac accagagcc gttgctgagc cgctgccagc 480
cccagatgt gcgctgagc gcggcagc aaccctgac gcagagcgg gcgcgcgag 540
gccgcgctg agcagccagc caccagctgt atggagcagc ctgagcagc cccagacagc 600
atcgcagca caccgtgctgt gcagcagcag cggcagcag cgcagcagcag 660
agaagtcgc aaccattcag gcagcagcag caccagcag caccagcagc 720
cgcgcggc agacacggag cccggcgac gcggcagcag cggcagcag cccagacagc 780
-continued

agagccgcc ttaaccttag aggcagcgtg gaccacacag gcctgccgct gcctgccgtg 840
tacgaccgcc ctaggagcc ggcacagcac tctgacagc agggatccg ccttgctgcc 900
atagacccct tcaaggtgctg ctagacagc cagctgacta cactggacag acccaaagcag 960
aaccgcgcc acaagagcca ggtggtgcgg atggtctgac agaccgcgcc cttccagggac 1020
cagctgctcg tgcacctctag cagaggccccc aaggtgtcggc ccagagccccc gcttgacacc 1080
agagccgcc agagctgcc caacgagccag atggccacc tggagccagc caccctagg 1140
caggtgccg gatactgggg ctcagccgcc agagccgcc gcacacccca ccagccagag 1200
gcagcgcgcc gcagcagcgag cgtgtgaggcc aacccagccag tgcotgatgcc ctgcotcttc 1260
gagaagagcc ccaggtgcgc ggtgccaccag gccaagggcc agggccgacc cagcgcagtc 1320
gagccggag tcagcgagat gttgcaggggc gcacagccgg agagagttgc cttccagggc 1380
agagccgcc tcagctggtgc ccagcagggg gccaagggcc ccacgccgcc ccagctgccag 1440
agttggagcc aggagccgg ctctctgcgg gcacagcgg gagggttcgc cagctgctcg 1497

<210> SEQ ID NO: 76
<211> LENGTH: 498
<212> TYPE: PRT
<213> ORIGIN: Artificial sequence
<220> FEATURE: 
<222> OTHER INFORMATION: NP Consensus Sequence

<400> SEQUENCE: 76

Met Ala Ser Gln Gly Thr Lys Arg Ser Tyr Glu Gln Met Glu Thr Asp
1   5   10   15
Gly Glu Arg Gln Asn Ala Thr Glu Ile Arg Ala Ser Val Gly Lys Met
20   25   30
Ile Asp Gly Ile Gly Arg Phe Tyr Ile Gln Met Cys Thr Glu Leu Lys
35   40   45
Leu Ser Asp Tyr Glu Arg Leu Ile Gln Asn Ser Leu Thr Ile Glu
50   55   60
Arg Met Val Leu Ser Ala Phe Asp Glu Arg Asn Arg Tyr Leu Glu
65   70   75   80
Glu His Pro Ser Ala Gln Gys Arg Pro Lys Lys Thr Gly Gly Pro Ile
85   90
Tyr Arg Arg Val Asp Gly Lys Trp Met Arg Glu Leu Val Leu Tyr Asp
100 105 110
Lys Glu Glu Ile Arg Arg Ile Trp Arg Glu Ala Asn Gly Glu Asp
115 120 125
Ala Thr Ala Gly Leu Thr His Met Ile Trp His Ser Asn Leu Asn
130 135 140
Asp Thr Thr Tyr Gln Arg Thr Arg Ala Leu Val Arg Thr Glu Met Asp
145 150 155 160
Pro Arg Met Cys Ser Leu Met Glu Gly Ser Thr Leu Pro Arg Arg Ser
165 170 175
Gly Ala Ala Gly Ala Val Lys Gly Ile Gly Thr Met Val Met Glu
180 185 190
Leu Ile Arg Met Ile Lys Arg Ile Asn Arg Asp Arg Asp Ile Glu
195 200 205
Gly Glu Asn Gly Arg Lys Thr Arg Ser Ala Tyr Glu Arg Met Cys Asn
210 215 220
<210> SEQ ID NO 77
<211> LENGTH: 252
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: H1 Gene Consensus Sequence
<400> SEQUENCE: 77
Met Ser Leu Leu Thr Glu Val Glu Thr Tyr Val Leu Ser Ile Val Pro
1    5    10    15
Ser Gly Pro Leu Lys Ala Glu Ala Arg Leu Glu Asp Val Phe
20   25   30
Ala Gly Lys Thr Asp Leu Glu Ala Leu Met Glu Trp Leu Lys Thr
35   40   45
Arg Pro Ile Leu Ser Pro Leu Thr Lys Glu Ile Leu Gly Phe Val
50   55   60
Thr Leu Thr Val Pro Ser Glu Arg Gly Leu Gin Arg Arg Arg Phe Val

<210> SEQ ID NO 78
<211> LENGTH: 252
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: H1 Gene Consensus Sequence
<400> SEQUENCE: 78
Met Ser Leu Leu Thr Glu Val Glu Thr Tyr Val Leu Ser Ile Val Pro
1    5    10    15
Ser Gly Pro Leu Lys Ala Glu Ala Arg Leu Glu Asp Val Phe
20   25   30
Ala Gly Lys Thr Asp Leu Glu Ala Leu Met Glu Trp Leu Lys Thr
35   40   45
Arg Pro Ile Leu Ser Pro Leu Thr Lys Glu Ile Leu Gly Phe Val
50   55   60
Thr Leu Thr Val Pro Ser Glu Arg Gly Leu Gin Arg Arg Arg Phe Val

---continued---

Ile Leu Lys Gly Lys Phe Gin Thr Ala Ala Gin Arg Ala Met Met Asp
225  230  235   240
Gln Val Arg Glu Ser Arg Asn Pro Gly Asn Ala Glu Ile Glu Asp Leu
245  250  255
Ile Phe Leu Ala Arg Ser Ala Leu Ile Leu Arg Gly Ser Val Ala His
260  265  270
Lys Ser Cys Leu Pro Ala Cys Val Tyr Gly Pro Ala Val Ser Ser Gly
275  280  285
Tyr Asp Phe Glu Lys Glu Gly Tyr Ser Leu Val Gly Ile Asp Pro Phe
290  295  300
Lys Leu Leu Gin Ser Gin Val Tyr Ser Leu Ile Arg Pro Asn Gly
305  310  315  320
Asn Pro Ala His Lys Ser Gin Leu Val Trp Met Ala Cys His Ser Ala
325  330  335
Ala Phe Glu Asp Leu Arg Leu Ser Phe Ile Arg Gly Thr Lys Val
340  345  350
Ser Pro Arg Gly Lys Leu Ser Thr Arg Gly Val Gin Ile Ala Ser Asn
355  360  365
Glu Asn Met Asp Asn Met Gly Ser Thr Thr Leu Glu Arg Ser Arg
370  375  380
Tyr Trp Ala Ile Arg Thr Arg Ser Gly Asn Thr Asn Gin Gin Gin
385  390  395  400
Ala Ser Ala Gly Gin Ile Ser Val Gin Pro Thr Phe Ser Val Gin Arg
405  410  415
Asn Leu Pro Phe Glu Lys Ser Thr Val Met Ala Ala Phe Thr Gly Asn
420  425  430
Thr Glu Gly Arg Thr Ser Asp Met Arg Ala Glu Ile Ile Arg Met Met
435  440  445
Glu Gly Ala Lys Pro Glu Glu Val Ser Phe Arg Gly Arg Gly Val Phe
450  455  460
Glu Leu Ser Asp Gly Lys Ala Thr Asn Pro Ile Val Pro Ser Phe Asp
465  470  475  480
Met Ser Asn Gly Ser Tyr Phe Phe Gly Asp Asn Ala Glu Glu Tyr
485  490  495
Asp Asn
<table>
<thead>
<tr>
<th>65</th>
<th>70</th>
<th>75</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>85</td>
<td>90</td>
<td>95</td>
<td>100</td>
</tr>
<tr>
<td>Val</td>
<td>Lys</td>
<td>Leu</td>
<td>Tyr</td>
</tr>
<tr>
<td>115</td>
<td>120</td>
<td>125</td>
<td>130</td>
</tr>
<tr>
<td>Gly</td>
<td>Leu</td>
<td>Ile</td>
<td>Tyr</td>
</tr>
<tr>
<td>145</td>
<td>150</td>
<td>155</td>
<td>160</td>
</tr>
<tr>
<td>Ser</td>
<td>His</td>
<td>Arg Met</td>
<td>Val Ala Thr</td>
</tr>
<tr>
<td>180</td>
<td>185</td>
<td>190</td>
<td>195</td>
</tr>
<tr>
<td>Ala</td>
<td>Gly Ser</td>
<td>Ser Glu</td>
<td>Ala Ala Glu Ala</td>
</tr>
<tr>
<td>210</td>
<td>215</td>
<td>220</td>
<td>225</td>
</tr>
<tr>
<td>Ala Arg Glu Met Val Glu Ala Met Arg Ala Ile Gly Thr His Pro Ser</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>245</td>
<td>250</td>
<td>255</td>
<td></td>
</tr>
</tbody>
</table>

**<210> SEQ ID NO 70**

**<211> LENGTH:** 97

**<212> TYPE:** PRT

**<213> ORGANISM:** Artificial sequence

**<220> FEATURE:**

**<223> OTHER INFORMATION:** M2 Gene Consensus Sequence

**<400> SEQUENCE:** 78

Met Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Aen Glu Trp Gly

1  | 5  | 10  | 15 |

Cys Arg Cys Asn Asp Ser Ser Asp Pro Leu Val Val Ala Ala Ser Ile

20 | 25 | 30 |

Ile Gly Ile Leu His Leu Ile Leu Trp Ile Leu Asp Arg Leu Phe Phe

35 | 40 | 45 |

Lys Cys Ile Tyr Arg Leu Phe Lys His Gly Leu Lys Arg Gly Pro Ser

50 | 55 | 60 |

Thr Glu Gly Val Pro Glu Ser Met Arg Glu Tyr Arg Lys Gly Glu

65 | 70 | 75 | 80 |

Gln Aen Ala Val Asp Ala Asp Arg Ser His Phe Val Ser Ile Glu Leu

85 | 90 | 95 |

Glu

**<210> SEQ ID NO 79**

**<211> LENGTH:** 759

**<212> TYPE:** DNA

**<213> ORGANISM:** Artificial sequence

**<220> FEATURE:**

**<223> OTHER INFORMATION:** Optimized M1 Coding Region

**<400> SEQUENCE:** 79

atgacgctg cagccagaggt cggacgcat tgtctctcga tcgccccag cggccccctg 60
---continued---

```
aagggccaga tcgcaccag aatgtgaggct gtcgttgcc gcaagaacac cgaacctgga
  120
ggctcgtgg aagtggctgg gaccacccac atcctgcagc cctgtgccaa gggccatcctg
  180
ggctcgtgg tcaacatcag ccgctcgcag gacagagcgc tcgacgaagag aagatactgtg
  240
cagaaacgccc tgaagcggcc acgcgttcgcc acacactgg agagagcgcg gacagcgtgac
  300
gagaaacgcta gagagagat cacccctccg gcggccagag agatctaatc gctgcaacgc
  360
ggcggtgcttc tgaagcggcc acgcgttcgcc acacactgg agagagcgcg gacagcgtgac
  420
gagttgcctgct gcgtgctctgc tcggcgccag gcggccagaga acacactgg agagagcgcg
  480
gcggccaccag agatctaatc cccggttagc gcggccagaga acacactgg agagagcgcg
  540
cgcggccaccag gcggccagaga acacactgg agagagcgcg gacagcgtgac gcggccagaga
  600
gacacactgg agagagcgcg gcggccagaga acacactgg agagagcgcg gacagcgtgac gcggccagaga
  660
caccaccaccag gcggccagaga acacactgg agagagcgcg gacagcgtgac gcggccagaga
  720
cacccggttagc gcggccagaga acacactgg agagagcgcg gacagcgtgac gcggccagaga
  780
```

<210> SEQ ID NO 80
<211> LENGTH: 294
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Optimized M2 Coding Region

```
agtgcagctgc tgcagcagtt gcaaaaacgt atcagaaacag aatgggggtg cagatgcac
  60
agtgcagctgc tgcagcagtt gcaaaaacgt atcagaaacag aatgggggtg cagatgcac
  120
tgcagctgc tgcagcagtt gcaaaaacgt atcagaaacag aatgggggtg cagatgcac
  180
gagtcgtggct gcgtggctgg cagcgcggcg gcggccgctgc gcggtgcgtt
  240
gagtcgtggct gcgtggctgg cagcgcggcg gcggccgctgc gcggtgcgtt
  294
```

<210> SEQ ID NO 81
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: H2M6 Binding Peptide

```
Thr Tyr Gln Arg Thr Arg Ala Leu Val
  1  5
```

<210> SEQ ID NO 82
<211> LENGTH: 11
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: RSV Promoter from Plasmid VCL1005

```
ctctctaacg g
  11
```

<210> SEQ ID NO 83
<211> LENGTH: 11
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
Other Information: Promoter RSV/R

Sequence: 83

tacaatatagctcg

Sequence: 84

cagcagctgc tccotggtg tggcttg

Sequence: 85
gatatacgc gcagctgta

Sequence: 86
caccacattgc ggtgcacact ccatcggtct gcactctc

Sequence: 87

aggtgcaaac caaatgtggtg aatggtcaca tggcgtttat tg

Sequence: 88

aatggtcac aatggcttat tgatcagctc tagcgctta aata

Sequence: 89

Sequence: 90
--continued

tcgccggttt cggctaagao gctcgasaamc tcggacacat gcaagtccco gcagacgctca 60
cagccttctct cggagcgaa cacggcagcc gcaagagcag tcggggcctg 120
tggtcgggtg tggggtggtg ctcacacttg cggcactaga gcaacagtga ctggacgtg 180
acctacagcg gtggtaacat cgcacacagtc gctaagag gaaataacagc atgggaatggg 240
cattgggtgtctgcttggctg gctggcgtgg caggtggtga gcaaaaatctca 300
agctcactaa agctcaaggtt tggccacaa ttcgatgaag atctcttta ggtgataggg 360
tttcgccgct ctctcgcgag taccgcgctg atagcggtgt atccagggg agctcggtgt 420
gtttagggc aagaggggg gcagcgcggtg agcgtttctg agtggacttc acaaataa 480
gttctcttt gtcagatgar ggggaagag tagcttctag gaaactctct gactcttggtc 540
acactctcgg acagttctaa gctgaacatcg ttcagaggg agassaaagc cgcagcgac 600
cgtcttgcag aagttgaggct tgtgtgcttg gctttttttag gaaaaaagaagcg 660
acgggtgtag gcacgccacg cggatcggag ttcagagaga ttagtatatt gttcgctag 720
cctctcctct ctaaagccat tgcacacttc acaacacttg tggcagcttc acaaatcctag 780
cattctctc tcaagcgcgc ggcactctac atcaggctgg ctaaaccgac gtttaaatgg 840
cgtgcggcg cttcgccgct cggctcctg ggtgcgcttg acggagaattt gctggaaagt 900
aacacagtctc tcagggcgag gctttgtcct tgtgagctgcct tgcagctctcag 960
ccgggtttt ccgcgggttc tgtcgtctgt ctgcattctc tagttacgg ctgagggttcg 1020
tgctcgtgag gcagcagctg tgtgcggcag gcggccgcag acgatcactc gttcgacaag 1080
tgaacgctg acaggggctt ctcagcgtctc ctgctcggag tattgaccgcc atggcggcggc 1140
ggctggccac gggagggg cagggcggtc ctcacacttg gcgtagcgag cttcgaggtt 1200
gtacaatgct gtgagagcag cgtaacctct taaaactcctc aaggggaagg gtagcagagc 1260
gtggctactc ctgctcggca caggtgcggc aacggtctgc gaagaacactc aaggggtcggc 1320
tgcacagctc gcagcggag tggagaaac tgaacatcg atcagagcag tcggttctac 1380
cacctgtgct ccgggtagag acacacacaag tggagacagc cagagacagc ctggatggc 1440
gtggactcgc gcagcggctg cgggtgctgg cggactctcg cttggtgctgg acgacgatgc 1500
ctgacatgc gcacgccacg caacagtctg gtaagccaga gggagggggc tggagaggc 1560
agctacccag cccacagctc ttagtcaaat acaacacatt gcttcctcgtg gaaacsctgta 1620
tagcagcttc tgtgcggcct cttgctgtgg atctcggaga gcacacacat gcagagagg 1680
gtgctttcgc tgtcagctgt ttgctgacc gcagcctccc aagttactct ctgtaagagc 1740
agcttcacgccc actrccaggt gccattggct cggagacgct cgggttattt cttcgccgaa 1800
ggcgtacctg ggtacactct ctcgttaatgc gcacactaat ccaacagttg gcagagagtt 1860
aatcggctg cggcacaacc gcggagacgc ccgggttattt gctgttaggt gcctcttggc 1920
activecctt ccacggctgg gctgcggcag atggggaaga ccacacatct ccaactgtgct 1980
gacccgctt tgtgtgctgg acagagagtg tgtgacaggg gcctggccac cggctggcag 2040
ctatccttgc agtagcgctc ttcgacaactc gctcccatttt tcggctcttc ccaagcata 2100
ggcgaaccc tctgtggata caacagatcg ctgacagtgt ccattttttg ccgctccgag 2160
acactcttcg aggcacaggc aacagtgagc ggcacacttc cggacggaga 2220
ttcagcatact ctcgctgctgc gctuttccac ccacataaag tggagacaggg 2280
acgttgctt tgcaaacagc gcctggttat gcaaccaagc caatactactg gcgtacgttg
2340
gcaaccaag ccgttcgtgg ccacgatgc tgtccctagt aaataccacc ccctggacag
2400
gttgctggct ggtgagctgc cagggagaagg cagaccaaga cagagctgtaa tggaggtttga ggaacacac
2460
gccacgagc agcttgctag agctctgagg tccaacaggg gactctgcca tcaacctttg
2520
gctggaggca ttcctgtgag atttctcaac aacacagtca agttggacgt ggtctcaattt
2580
gactgtgcag tgaagatggct aaacttgctg tggagggag caaactatgg ccgtctgctc
2640
gagcccttca agtttcttggt gcattcgcgc gcacagctgc aagggccctg ggtcttgaga
2700
ttcgacgaca ttgcacgagg tgtcagcctg aagactctca ttcctgcaag ggtctcaattt
2760
cagcagactca ccagccattg cagactgctg acagttaaccc ccctgggttt tcaggtgttc
2820
gccacgcctga cagcccttgag ttgactgga acacgacact gccaccttgga ctagctggct
2880
gccagccag aacaacctag ccgtcctacg tcggcacact gttggaagag cctggcaca
2940
gcttttcaca ccgcccctac aagccgcccc agctctgcgc cttctgactg ccagctgtgg
3000
gactttgtc caggtggcag ggtgtaaaccc tctgtgagag aggctgtgcg tcaagcggct
3060
gagagcagct ctcgctctggtccgtgagag atgcctgtgaa taaagcaaggg atggctggg
3120
gctggctctg ttggtgttag gcctacatgt ctgtagacgt ccaagtttac cagcttttct
3180
ggaaggtgg aagtttctct ccctctcttcc gccatcggtc gcagtttgcgg atccagatct
3240
ggcagttcc ctgctgacc gcctctgttg gttggaacct ccacgggtgc ccctgttgac
3300
cggccgttg ccacccctgc cggccgttg ccctggttag gggagtggc aggtgctggcg ggggaggttt
3360
tgggagca acttggagca ttcggcccg acggggtggt ctgggggtct gctggctgctg
3420
ggtggagcag aatggccctc caccactgct caccactgct caccactgct caccactgct
3480
gaggggtcct ccgtctggag aagactctca aacatagcttg ggaagccgac cggagccctg
3540
cogactctgc ccctctctgt cccctctctg cctctctctg cccctctctg cccctctctg
3600
gccgggctt ccctgccgat cccagcagct ctcctgtcat cgggctaagg cggcttggtg
3660
tccgctccgg ccagccccag cgtgcctctg cgtgcctctg cgtgcctctg cgtgcctctg
3720
ttcgactgtg cggagagctg aaactctgct gccaccttgga ggtctcagct gaggaacgag
3780	tcgactgctg ctcgctccg cggggccggg ctcgctccg ctcgctccg ctcgctccg
3840
ttcgactgtg ctcgctccg cggggccggg ctcgctccg ctcgctccg ctcgctccg
3900
ttcgactgtg ctcgctccg cggggccggg ctcgctccg ctcgctccg ctcgctccg
3960
gccgccaag cggagagctg aaactctgct gccaccttgga ggtctcagct gaggaacgag
4020
cgctctgctt ccctgccgat cccagcagct ctcctgtcat cgggctaagg cggcttggtg
4080	tcgactgtc ccctcccttg gtcctccttg cgtgcctctg cgtgcctctg cgtgcctctg
4140
tccgctccgg ccagccccag cgtgcctctg cgtgcctctg cgtgcctctg cgtgcctctg
4200
ttcgactgtg cggagagctg aaactctgct gccaccttgga ggtctcagct gaggaacgag
4260
ttcgactgtc ccctcccttg gtcctccttg cgtgcctctg cgtgcctctg cgtgcctctg
4320
tccgctccgg ccagccccag cgtgcctctg cgtgcctctg cgtgcctctg cgtgcctctg
4380
ttcgactgtg cggagagctg aaactctgct gccaccttgga ggtctcagct gaggaacgag
4440
ttcgactgtc ccctcccttg gtcctccttg cgtgcctctg cgtgcctctg cgtgcctctg
4500
ttcgactgtg cggagagctg aaactctgct gccaccttgga ggtctcagct gaggaacgag
4560
<table>
<thead>
<tr>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>acgacattac ggccagacaas aagagatctc aagagatcct ttggagccccttt ttttctggggctt 4620</td>
</tr>
<tr>
<td>ctgcgccttc ggtgagacaac aacgtcacgc aagagatatc gttcagattctg gttcagacata 4680</td>
</tr>
<tr>
<td>gggcccctctct cgtacatcttt taaatttaaa aagttaaattc taactacatc taactattatat 4740</td>
</tr>
<tr>
<td>atgttacacatt ttgagttcag gttcactacc gttcactacc gttcactacc gttcactacc 4800</td>
</tr>
<tr>
<td>ttacctcttt ccgtatcttcct atagtacgctg ctggcgggggc gaggggtgtgctg gttgctgtcc 4860</td>
</tr>
<tr>
<td>ctgctagaag agygtgtgtg gttcactacc gcggctgaac gggccccttc accagcctga 4920</td>
</tr>
<tr>
<td>aagtggaggg gccacagccttt tggaggtgt ggtgaggggt ggtgaggggt ggtgaggggt 4980</td>
</tr>
<tr>
<td>acttcccttt ccgctctccctc ccgctctccctc ccgctctccctc ccgctctccctc ccgctctccctc 5040</td>
</tr>
<tr>
<td>gtgctactcttc gtacatcttc ccggggtttc ggtgaggggt ggtgaggggt ggtgaggggt 5100</td>
</tr>
<tr>
<td>ccagctacacacc gcgggtggtc gagctggtc gttcactacc gttcactacc gttcactacc 5160</td>
</tr>
<tr>
<td>addctacatcttt ccagttctct tccaatattata accagcctga gttcctgggggc gagggcttct 5220</td>
</tr>
<tr>
<td>ccacagagac ccacagagac ccacagagac ccacagagac ccacagagac ccacagagac 5280</td>
</tr>
<tr>
<td>tgcacagcctgc actctacccct gcggggtttc ggtgaggggt ggtgaggggt ggtgaggggt 5340</td>
</tr>
<tr>
<td>tgcagcctcct gcggggtttc ggtgaggggt ggtgaggggt ggtgaggggt ggtgaggggt 5400</td>
</tr>
<tr>
<td>atgctacagc gttcactacc gttcactacc gttcactacc gttcactacc gttcactacc 5460</td>
</tr>
<tr>
<td>cgcacagcctgc actctacccct gcggggtttc ggtgaggggt ggtgaggggt ggtgaggggt 5520</td>
</tr>
<tr>
<td>gcagcctagtctgc gcggggtttc ggtgaggggt ggtgaggggt ggtgaggggt ggtgaggggt 5580</td>
</tr>
<tr>
<td>ccacagagac ccacagagac ccacagagac ccacagagac ccacagagac ccacagagac 5640</td>
</tr>
<tr>
<td>gcggggtttc ggtgaggggt ggtgaggggt ggtgaggggt ggtgaggggt ggtgaggggt 5700</td>
</tr>
<tr>
<td>ggtcctggggt gcggggtttc ggtgaggggt ggtgaggggt ggtgaggggt ggtgaggggt 5760</td>
</tr>
<tr>
<td>ccagctacccct gcggggtttc ggtgaggggt ggtgaggggt ggtgaggggt ggtgaggggt 5820</td>
</tr>
<tr>
<td>addctacatcttt ccagttctct tccaatattata accagcctga gttcctgggggc gagggcttct 5880</td>
</tr>
<tr>
<td>ccacagagac ccacagagac ccacagagac ccacagagac ccacagagac ccacagagac 5940</td>
</tr>
<tr>
<td>tggcttacctctt cggcttcct tggcttacctctt cggcttcct tggcttacctctt cggcttcct 6000</td>
</tr>
<tr>
<td>tttcttacctctt cggcttcct tggcttacctctt cggcttcct tggcttacctctt cggcttcct 6060</td>
</tr>
<tr>
<td>ccagcctcg ccagcctcg ccagcctcg ccagcctcg ccagcctcg ccagcctcg 6120</td>
</tr>
<tr>
<td>acctacctct ccagcctcg ccagcctcg ccagcctcg ccagcctcg ccagcctcg 6180</td>
</tr>
<tr>
<td>ccagcctcg ccagcctcg ccagcctcg ccagcctcg ccagcctcg ccagcctcg 6240</td>
</tr>
<tr>
<td>ccagcctcg ccagcctcg ccagcctcg ccagcctcg ccagcctcg ccagcctcg 6254</td>
</tr>
</tbody>
</table>

<210> SEQ ID NO: 90
<211> LENGTH: 6425
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE: 
<220> OTHER INFORMATION: VR6307, Ligation of VCL6292 into VR6430
<400> SEQUENCE: 90

tgcgccttct ctgctagcct ggtgaggggt tttgagccccttt ttttctggggctt 60
ccagttctt ctgctagcct ggtgaggggt tttgagccccttt ttttctggggctt 120
ctgctgcttct ctgctagcct ggtgaggggt tttgagccccttt ttttctggggctt 180
acgacattac ggccagacaas aagagatctc aagagatcct ttggagccccttt ttttctggggctt 240
-continued

catgtgcct gcctcggcttc tctgtctggga ggtgctgcag tagtgctgag tcgccctgca gcggattcta 300
aggcatcaaca agggacgtt tgcgcagcacc tgcgattcg aacttcatcat gttgctgagc 360
ttttggggtc ttttggggtc ttcggtcgcag atatcgtcagt atctgggagag atctggggtc 420
gttggggtc aagggggttg tcgcgtcaat cgggctgcag ggcgcctcctag ggtattcatt 480
gttctgcttt tcgtatttga ggctgggaag agcactattc catattcatt gcgtctgctg 540
acacacatcc caggataggc gcagcattgct ttcacggtgc agaacagcagc ccgctgcgctgc 600
gcggtgtggc agacatttc gtagctacag gcctattcag gcgccacc aacgggcttg 660
acacacatcc caggataggc gcagcattgct ttcacggtgc agaacagcagc ccgctgcgctgc 720
tcggctga attccaacttc accacatcc atcctgttagc agctcattgct 780
ttcacggtgc agacatttc gtagctacag gcctattcag gcgccacc aacgggcttg 840
gctgcgcgctct gcttcgctgt cgaaggtcctg acgctggattcctgtagtcttc 900
acacacatcc caggataggc gcagcattgct ttcacggtgc agaacagcagc ccgctgcgctgc 960
cggccctctc aacgggtgcgc gtcggcgctg tcgggtcctg aacgggcttg 1020
tggttgggtc aagggggttg tcgcgtcaat cgggctgcag ggcgcctcctag ggtattcatt 1080
taaccgttgccc actcggctgcag gccgctgcag ggcgcctcctag ggtattcatt 1140
atcctgcgctct gcttcgctgt cgaaggtcctg acgctggattcctgtagtcttc 1200
tcggctga attccaacttc accacatcc atcctgttagc agctcattgct 1260
gctgcgcgctct gcttcgctgt cgaaggtcctg acgctggattcctgtagtcttc 1320
ggccctctc aacgggtgcgc gtcggcgctg tcgggtcctg aacgggcttg 1380
cggccctctc aacgggtgcgc gtcggcgctg tcgggtcctg aacgggcttg 1440
acacacatcc caggataggc gcagcattgct ttcacggtgc agaacagcagc ccgctgcgctgc 1500
acacacatcc caggataggc gcagcattgct ttcacggtgc agaacagcagc ccgctgcgctgc 1560
atcctgcgctct gcttcgctgt cgaaggtcctg acgctggattcctgtagtcttc 1620
ggccctctc aacgggtgcgc gtcggcgctg tcgggtcctg aacgggcttg 1680
cggccctctc aacgggtgcgc gtcggcgctg tcgggtcctg aacgggcttg 1740
acacacatcc caggataggc gcagcattgct ttcacggtgc agaacagcagc ccgctgcgctgc 1800
acacacatcc caggataggc gcagcattgct ttcacggtgc agaacagcagc ccgctgcgctgc 1860
ggccctctc aacgggtgcgc gtcggcgctg tcgggtcctg aacgggcttg 1920
acacacatcc caggataggc gcagcattgct ttcacggtgc agaacagcagc ccgctgcgctgc 1980
tcggctga attccaacttc accacatcc atcctgttagc agctcattgct 2040
ggccctctc aacgggtgcgc gtcggcgctg tcgggtcctg aacgggcttg 2100
cggccctctc aacgggtgcgc gtcggcgctg tcgggtcctg aacgggcttg 2160
tcggctga attccaacttc accacatcc atcctgttagc agctcattgct 2220
ggccctctc aacgggtgcgc gtcggcgctg tcgggtcctg aacgggcttg 2280
acacacatcc caggataggc gcagcattgct ttcacggtgc agaacagcagc ccgctgcgctgc 2340
tcggctga attccaacttc accacatcc atcctgttagc agctcattgct 2400
ggccctctc aacgggtgcgc gtcggcgctg tcgggtcctg aacgggcttg 2460
cggccctctc aacgggtgcgc gtcggcgctg tcgggtcctg aacgggcttg 2520
aagcgcagac aggtgagact agacacagat caggtgtatg gaaacatagc gacatataac 2580
tttacgacact gcgtgacact ggtcgaact ggtcgagact ttctgacga 2640
atcaagagca caacgacag aatcattcat aatgggaagct acttcacact cgcagacctc tgcctgtact 2700
agaaattgct cgcagagc caggtacca ctgcagacga ctaaacacgg caatcagctg 2760
aaagagaatc tgaacatgctctcgttact withcactta agagctaaactt gcactctatag 2820

gggaagaca tagcaggttt tgttctcaag tcgtatcagc agctttccca aatatacaca 2880
aatcatttgc cagcggtgaa tcgaccaact ctatcaccggt tttctgttac aatcataactt 2940
aacaagcagata tgaacatcag gaaacacgac caacaacttc actgacacgg caacaataata 3000
ggcagtgcc ctcgattgc actggtcaag gaggtcata gggaagctat caaagctgat 3060
acttgaagggg tcttactctaa tctggacga gcacactcga ggcctgtcgc cgcctttttc 3120
gtgcagagct aagagatcag gggtgtgagg gcaggtcttta cgcgacatgtct gcagttgtc 3180
acacatttca cgcagcagac ggcgtgaagct atctttactct acttaacacaa gataacgcgc 3240
atacatcttc tgggttgtac caaccacata ctaaagatgct atcttctgtc gcagttacag 3300
agaacacaa taactatactc aaagccgaca gcgactacac gacaaactag aataaaagctt 3360
atcctattata ttagattaaaa agctggctag atggtctagt gctgacattc tgcgctgtct 3420
tcgcttgcc agccatcctgttgattcccc tcacccagcttccttactgac cctggaggtt 3480
gccttcacac cggagtcttccttaactaatag gggaaattct catagctatct tgtcagtagt 3540
tgtcagatcta ttctgggggg ggctggagagaa cggcgcagcag gggggaggttt gcgggagac 3600
aatagggcgc tagggggagtc tgcgtgtggg gcgtttgctca cccaggtgtgt gaaatgggtg 3660
ccgggtttcg ctggtggagc aagaggacgt caacgccccc cttctgtgta cacaactctg 3720
ccgcctccg gttctgttact ccagccgcac ctaacctgcat cagaggggtt 3780
ccgcttcacc cccctttttg tcacccactgaac cggcgctttc ctggttcatc cagccgacoca 3840
ccacaccaca actgctccctc aagcagggta gcagtttaaac caaagattggt ctattattgt 3900
ccgggggaaa caggggtcc aacactggtg gggatattca gggagcacatt gagaatattttt 3960
aagccacgta ttaaaggcaagt ctcaggtctct ccctgcctcg cttctgtgact catagctgc 4020
tcgctgtgtat gcgcagagct gggtggacct tctgcgctct ctaaagactcg gtaataagt 4080
tatggcaacatagcagcatacgcgacagagcag cagcagagcag 4140
ccagacaggg aaangaagcc gcgggttggt gcatttttca tagctctgac cccccctgaag 4200
accggacac aataagcagctctagcaagctggggagaaa cccagacagctataaagct 4260
acggttgttc ttccctttgccg agctctggtcct gcgtttatat gcgtctcaagct 4320
cgcgatgct ctggggtgtag gccactgtg gggtgttcg gattcattat agtcaagct 4380
aratgtgcat caggtgaggttagctgcc gccaacctg gcgggtgtggt ctgcaacccc 4440
cgcgacacag cgcgctccac gcctgtccta gcctggattg gcgtcagtaaa 4500
agaracagctt atggcaagttg cgcagagcct ctggtaacgcc gattacgcc gagaggtgctg 4560
tcgggtgtgt ctagagctgtgttc ggcctttcgt cgcgctcaggtc aagacgagag 4620
tattttgctctcgcgctgttccagttcctccccctggcggatgaagaggtgttttggctg 4680
gaggtccagacaacacacacccgtgcagcctgcgtttttctgctgctctgctaconag 4740
cgcgcagagcttccagacacttccgcgctctgcctgcggagatcgggacagctggcagtc 4800
agggagca aacccactg taaaagttt tggctatgg atatatcaga agatcotta 4860
ccatctcc ttaaaaattaa aatagtaag ttaataact cttaaatgta tetagtaaa 4920
ctggttcga cagctcaca gttcaataca gtgcgcctac tattcucagc atcactatct 4980
ttcgcttcat ctagctgcc tgcctccggg gccggggtct gtagctgtg ccctcagg 5040
aaggttgc ctcacactac caagctcaaa tcgccccact atcacaacag aagttgagg 5100
agcaggggt gagtagacgt tctgcaggg tgcaccaggt ggtgcattt gcccctggt 5160
tgccaagga agcgtcctcc tggcgggaa gtagctgtg ctgcatcttc aacacgca 5220
aaggtcgt tttacttacat agcgcgctgc cgcagcagcc aggtaatacg tctcgcaagt 5280
tacacacaatt ctgattagaa aacccctcg accataaat gaactgcaaa 5340
ttcatactca ctgaggttat caatacata tttttgaaa agcgtcctgt gtaaagag 5400
aggctaca cagagccgct tcctaggtatt gcagccagcc tcgatgcctg ctgacattcc 5460
gacgtcaca aaccccattc accctgataa tttccctctac tcacataata ggctaatcag 5520
tcgacatc actgacatgg ccgagcagc tgaagcagcct gtagctgtctc 5580
ctccagactc tcgctcaatac gcagcaacttc acgctcctca ataaaataac tcgatcaaa 5640
ccacagccctg tcatagctgct ggcgcggct aagccagcgc tagagttgaa 5700
agggcacta cccagcagcag cggccgagc aacccgctca ggcgtcctac 5760
atcaggcaca cggatcactgc gatctaatgct taaactgag tggggctggatatg 5820
cgcgctggtg tgaacacgca cgcctcaaaag aggctgtcga tgttgagaga 5880
aggataataact cttctagctg gcactccaaaat ctgttacaact cattggaac 5940
gttcgctttaa caggcttcccctggcgcctgtgcgttcgctaat cactgataa 6000
gactgctgca agagctggcc gccagccatg tttcagctacc ataaacgctg 6060
ttcgcttgag ttcaggcctg gcagatcgcct tccctgcataa cttctgccat 6120
cctacccctc cctgacgctgt ataatggaac aacagtatttg atgtgtcag atatatttatt 6180
atgatgcttctg cagttcagatctgatcagc agtggctccccccccccccc 6240
cctatattgggt atggttctagggctgtacattg gatactatatgtagtatgatatgatatat 6300
cttacatatttttaaagcagctggcgcctttc cgcacatgct acacagcgy 6360
tcagaaacc atatatcaca tccactatac atataaaaaa ccgactttattcg 6420
tgctc 6425
cacttgccag tacatcag tctactatg ccacagagc cccatatgga cgtcaatagc 360
gttatgtgg cggctgyggct ttagccccag tacatatctt atatcctcaggct tcctcctgg 420
cagctactct acgtatatct catcagatat accatgtaga tggcttttgg gcaatcatac 480
atgagcctg cacaagcgttg tgcacacgga ggaattcaca gttcacaccc catggagctc 540
atgagcttg tgcctggtca ccataaactca gggacatcttc ccaatgtcgtc taccactcctc 600
ggcacattga gccaatgggg cggctgggt ctacgctggg aagcttatg acgcagagt 660
gcttttaga acgcagagtc cgcctttgga gcaactccac gcgtttttgca ctcctcataga 720
agacecaggg acgcagagtc cgcctttgga gcaactccac gcgtttttgca ctcctcataga 780
cgctgccagga tggacyttaga taccgctctg agacacctta ggaacaacccc ttggctttott 840
atgacagga taccagggaa gggccagcgc gcctctccttt agcttatagg 900
tgatgtgata cgtgactcgag taagctttgga ttttgacacca ttatactaccc ttcctctatt 960
gttgctcttga ccctctccat taacatcaac cagctgtcttt tggccacacc tctctctatt 1020
ggtcgtctgg caaacctcggt tcttcccgag gctgcagggg acoctgttatt ttcacgcgtat 1080
ggcggcccaat tattatatta caaattcaaca tataaaacaca gcgctccccc cgcctgaccga 1140
gtttttata aacataggg gggcttcaga cagctctccg gcggacgctg tggcaaacat 1200
gcctctctt ggcctgacgc ggcctctctt cgcctctctt gcctctctct ggcctctctt 1260
gcctctctg gcctctctct gcctctcttt cgcctctctg gcctctctct gcctctctct 1320
cagctctctg gcctctcttt cgcctctctg gcctctcttt cgcctctctg gcctctcttt 1380
tgatgtgata cgtgactcgag taagctttgga ttttgacacca ttatactaccc ttcctctatt 1440
aatagatag acgcagatga tctgtctatt tctgataaca gcacagaga ttcacgcgtat 1500
cggtcgtctgg caaacctcggt tcttcccgag gctgcagggg acoctgttatt ttcacgcgtat 1560
caccccaacaa taccatagag cagctctctt atcagctgtt attcgcctctg gcctctcttt 1620
gtaccctct gcctctcttt cgcctctctg gcctctcttt cgcctctctg gcctctcttt 1680
tgatgtgata cgtgactcgag taagctttgga ttttgacacca ttatactaccc ttcctctatt 1740
tgatgtgata cgtgactcgag taagctttgga ttttgacacca ttatactaccc ttcctctatt 1800
aatagatag acgcagatga tctgtctatt tctgataaca gcacagaga ttcacgcgtat 1860
tgatgtgata cgtgactcgag taagctttgga ttttgacacca ttatactaccc ttcctctatt 1920
aatagatag acgcagatga tctgtctatt tctgataaca gcacagaga ttcacgcgtat 1980
tgatgtgata cgtgactcgag taagctttgga ttttgacacca ttatactaccc ttcctctatt 2040
aatagatag acgcagatga tctgtctatt tctgataaca gcacagaga ttcacgcgtat 2100
aatagatag acgcagatga tctgtctatt tctgataaca gcacagaga ttcacgcgtat 2160
aatagatag acgcagatga tctgtctatt tctgataaca gcacagaga ttcacgcgtat 2220
aatagatag acgcagatga tctgtctatt tctgataaca gcacagaga ttcacgcgtat 2280
aatagatag acgcagatga tctgtctatt tctgataaca gcacagaga ttcacgcgtat 2340
aatagatag acgcagatga tctgtctatt tctgataaca gcacagaga ttcacgcgtat 2400
aatagatag acgcagatga tctgtctatt tctgataaca gcacagaga ttcacgcgtat 2460
aatagatag acgcagatga tctgtctatt tctgataaca gcacagaga ttcacgcgtat 2520
aatagatag acgcagatga tctgtctatt tctgataaca gcacagaga ttcacgcgtat 2580
-continued

agagctcggt ggctgctggc gacactggct ttcgcagcat agatgttgag taactsctgtg 2640
accaccctgc atccgacagt attccgtgct ataatcagag cagcagcct gcaagcttgc 2700
tgcgtgttct tgtcgtagga tttcctgctgc ttcagcgtgc ggtcgcggc 2760
gtggcgcg ccggctggtg tgtcagcttc taacaaggcg ttatagcggg ctccacagcg 2820
tccggggtgg aacgataaga gcagcggct gcaaaaaggg gcaaaaggtc gagaacggc 2880
aagacagcgg ccctgggtggt gtctttcagc aggtgccccc cctccgga gaataaaac 2940
aatcagcgcg cacatcgagt ttgaggaacc cggacagcgc tacattggtc ccaggtttt 3000
ccccgtgaga gcctactggc acctgtgcgg gtctccaggc tggcggcatt cggatatcgg 3060
tgccgtttccc ttcctggggg aagctgggag cttccctcata gtcacagctg tagtatcctc 3120
agtttgtgcgc agttctgtct ccctcagcgg ggctgcgtgc acaaaaocccc cttccggcgc 3180
g cgctgtggc cctcttggc acctagctgt cttgagtctc accggtaag aacaggtctt 3240
tggctctggc cccgaggccc tggcagcgg ggattcagc aaggggtgc cggcttgtac 3300
acccgggttt tcaagtgttc gcaatcctcc ggtttgacgc aacgaggttg ctggctggtt 3360
tccgctgcc tgtgagccgt ttaacctgga aaaaaggtgg ttaggctttg atccgggaa 3420
ccacacag ccctctagag cggcctttcc cggctcggc acaatgagaa ccgcaaaa 3480
aatcagctcg acagctgact ttacctccccgt ctgaacgctgct gcagcgagaa 3540
aatccgcgtt aagcggttct gccgctgaga ttaaagaaaaa aaccttcgac ggacggcgac 3600
tatattacaa aatgagtttt ttaatcctac tacagtatac attgatagct ccggctgg 3660
agtacactgc gtaaagacag tgcctccgga cttcccgact ctgtttaccc 3720
atatgttgcc gccagctgggg gcggcggggc cggctgtgct gcaagctgct gcccggtttc 3780
gactatgac agctgtaaca cggcctaa caaagcgctgc taccgaggtc gcaagccggt 3840
agtgggatt ggacggttag cggctgtggt ctggctggct gccctggcgtt cggctgtggt 3900
cggctgtcgc tgtgagccgc ctgttgcgaca atccgggaa aacctctggct ctggctgtg 3960
attcctcgg gcggctcgtc gtcgcaactc gttgcagctgc ctggcggctg ctaccgacgg 4020
tacgctgactcg tttcagaga aacctctgg accagctgct ggaacctctttcctttgact 4080
ccgctgata aacgtccttt atttgagaaa gcggcttttg taattggtta gcgcacccac 4140
cgggcgcgt ccagcctagg cccgatatct gccatgctcc gtggctgggc actcctccaa 4200
cacactacaa acctctattt ttcctttggg cccatagag aagcagcga ccgacgcttc 4260
cagggcagc gccctctagt gttggagctt cggcagcctt atggttactcc ttcgcacggc 4320
gggcactag ccagcttgca cggcctgtc cttactattc ttcggcagct 4380
tgtcctgtc tgtcctgtgc gcggacgctg atccgcatgc gtgtgtaaag ggacacattc 4440
aagagagat cgaagtaaag cggcagcggc aacgcgacag cggcatacga atatcttccc 4500
cgtcgctag attctttcc aatatcagag atctctgttccc cccggggtgc gcattgcgtc 4560
gttcaccag atccagcgag gcaggtctaa aatgtttgat gcggcgagag ggcatttaatt 4620
cgtgaccag cggctgtgctgcc gcaagctgct gcggctggctg ctggctcggc 4680
ccggccttg ccgctgtgct gcggctggctg cggctgtgct gcggctggctg 4740
atcccggtg gcggctgtgct gcggctggctg gcggctgtgct gcggctggctg 4800
aatgagtttt cggcctgaga cgaggctgatt ttcagcagctc tacagtgttg 4860
-continued

tattactgtt tagtangoaac gacagtttta tgtttcatga tgaatatatttt tatcctgtg 4920
catactgaca tccagatttt gacggacaaca cgtgcttccc ccccccocccc cattttgaa 4980
gccatttacgc gggtattgtt ctcctgagct tcatatatg tgaattgatt tgaaaaaataa 5040
acacaaatag ggtctctgccc acatcccccc caaaagtgcc aacgtacgct tgaagaacccca 5100
cttatattc gatactaccct tataaaata ggcgctctac gacgcctctt ccgtccgctg 5160
gtcctctgct gctctcgtgg aacgctgtgg gctctcgagt cccgctgtgg gccctcgcgt 5220
gttctgtgc gcgttcgctgg aggacacag ccgctcggag cgcgtctgg ggtctgctgg 5280
gtgtgtgtgg gccgttctac tattggtga atagacgatg tgaagctgaa gtcacacata 5340
tgctctgtgc aataacgcgc aatgcgctaa gagaagatta ggaagaaata acgtacacag ttgcctat 5398

<210> SEQ ID NO: 92
<211> LENGTH: 4710
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: VR4759, Ligation of M2 into 10551

<400> SEQUENCE: 92

tggctcaaatgct ctcatactat aaatgtacaa tttatatggy ctctgctgcca 60
acctacgcc cagttgtgca ctctctatttc aagttata atataagtct cttgggccc 120
tcatactgaa atgagcctaa ctcatttctc gcctctgcgt atgtttaag ctgctacagct 180
cgctgtgcgcc ccccccaccca ctatgctcg ccattttttc tatttcttttctgcttttactt 240
gtaacccc gtaaggag cattttcttta gttttggggt ctacatagc ctgcctctcct 300
caataggcc aacgctgctg ctcctcagtt aatattagc gatcctttaa atcccctccct 360
gtacacacg acgtattttg ttcctctctc aatattattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
-continued

```
agcagctgc agcagctgc gtggctcgat tgtgataaga gtcagagaga actcccttgtg 1500
cggctggct acggctggct gcgcagctgc actcgctgtc gcgcgagctgc 1560
ccacacagc taatagctgc cagctaaca gactgctgtc ttcataaggt tgttttcgca 1620
gtgcgcgctg tcggataacg aaactgcaac ctcagcgcct ccgacgatggt tggaagctgcc 1680
catcagacgc cggacggtgc gcagagctgc gacacacagc gaccccgcttc ggtggtggcg 1740
cagcagctgc gcacagtgtgc acctgtactg tggatctcctgc gcgcagactg ttttcagtg 1800
ccacacagc ctggctcgac gcgcgctggc gcgcgagctgc gcacagtgtgc gcgcagactg 1860
gagctgcagc gcacagtgtgc ggcacagcct cttggggtgc gcgcagactg gcgcagctgc 1920
cttcgctgat gcagagctgc gcgcagctgc gcgcagctgc gcgcagctgc gcgcagactg 1980
cataacagg cagcagctgc taatagctgc tgggtggtgc ttttgtgttc ggcacagcct 2040
cttcgctgat gcagagctgc gcgcagctgc gcgcagctgc gcgcagctgc gcgcagactg 2100
acccagagc ggtcagctgc ttcctgcagc acctgtctggc gcgcagactg gcgcagctgc 2160
ggcacagc gcacagtgtgc gcgcagctgc gtttgataacg acctgtctggc gcgcagctgc 2220
atagcctgc gcacagtgtgc gcgcagctgc gtttgataacg acctgtctggc gcgcagctgc 2280
acccagagc gcacagtgtgc gcgcagctgc gtttgataacg acctgtctggc gcgcagctgc 2340
cagctgcagc gcacagtgtgc gcgcagctgc gtttgataacg acctgtctggc gcgcagctgc 2400
cgcttctca tgcagctgc gcgcagctgc gcgcagctgc gcgcagctgc gcgcagactg 2460
tgggtgtgc gcgcagctgc gcgcagctgc gcgcagctgc gcgcagctgc gcgcagactg 2520
gtttctgat gcgcagctgc gcgcagctgc gcgcagctgc gcgcagctgc gcgcagactg 2580
gctggctgat gcgcagctgc gcgcagctgc gcgcagctgc gcgcagctgc gcgcagactg 2640
acccagagc gcacagtgtgc gcgcagctgc gcgcagctgc gcgcagctgc gcgcagactg 2700
gtttctgat gcgcagctgc gcgcagctgc gcgcagctgc gcgcagctgc gcgcagactg 2760
cttcctgat gcgcagctgc gcgcagctgc gcgcagctgc gcgcagctgc gcgcagactg 2820
cttcctgat gcgcagctgc gcgcagctgc gcgcagctgc gcgcagctgc gcgcagactg 2880
ggactgata actgctgcagc ttgaattata aaggtctgag actgctgcagc ttgaattata 2940
ggactgata actgctgcagc ttgaattata aaggtctgag actgctgcagc ttgaattata 3000
ggactgata actgctgcagc ttgaattata aaggtctgag actgctgcagc ttgaattata 3060
ggactgata actgctgcagc ttgaattata aaggtctgag actgctgcagc ttgaattata 3120
ggactgata actgctgcagc ttgaattata aaggtctgag actgctgcagc ttgaattata 3180
ggactgata actgctgcagc ttgaattata aaggtctgag actgctgcagc ttgaattata 3240
gtggctggat gcgcagctgc gcgcagctgc gcgcagctgc gcgcagctgc gcgcagactg 3300
gtggctggat gcgcagctgc gcgcagctgc gcgcagctgc gcgcagctgc gcgcagactg 3360
ggactgata actgctgcagc ttgaattata aaggtctgag actgctgcagc ttgaattata 3420
ggactgata actgctgcagc ttgaattata aaggtctgag actgctgcagc ttgaattata 3480
ggactgata actgctgcagc ttgaattata aaggtctgag actgctgcagc ttgaattata 3540
ggactgata actgctgcagc ttgaattata aaggtctgag actgctgcagc ttgaattata 3600
ggactgata actgctgcagc ttgaattata aaggtctgag actgctgcagc ttgaattata 3660
ggactgata actgctgcagc ttgaattata aaggtctgag actgctgcagc ttgaattata 3720
```
<table>
<thead>
<tr>
<th>Start of Sequence</th>
<th>End of Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>3780</td>
<td>4020</td>
</tr>
<tr>
<td>4020</td>
<td>4260</td>
</tr>
<tr>
<td>4260</td>
<td>4320</td>
</tr>
<tr>
<td>4320</td>
<td>4380</td>
</tr>
<tr>
<td>4380</td>
<td>4440</td>
</tr>
<tr>
<td>4440</td>
<td>4500</td>
</tr>
<tr>
<td>4500</td>
<td>4560</td>
</tr>
<tr>
<td>4560</td>
<td>4620</td>
</tr>
<tr>
<td>4620</td>
<td>4680</td>
</tr>
<tr>
<td>4680</td>
<td>4710</td>
</tr>
</tbody>
</table>

**Sequence:**

```
tggccatgtg atacgcttg taataaatct atataatgca ttaaataatg ctataatgca 60
acattacgcc atctggcgc tttgttattg acggatat ttaataaatgct tttgcaaggg 120
cttattggt atagcctca taataatgca ttaaataatg ctataatgca 180
ccagggcggcgccgc ggcccgcct tgggagcgc tttgcaagtt tattggtgatatcgata 240
taaccgagc ctaaatgct tattggtgatatcga caagggggtg tataaatgca gtaaccgggcc 300
cctgttggcct gtagaatgtg ctaaatgct tattggtgatatcga 360
gtattgatgg ggcggcgcct tattggtgatatcga 420
ttaaataatgct atataatgca ttaaataatg ctataatgca 480
acgggagggga atagccttg ccaggggagggga atagccttg ccaggggagggga 540
acgggagggga atagccttg ccaggggagggga atagccttg ccaggggagggga 600
ccaggggagggga atagccttg ccaggggagggga atagccttg ccaggggagggga 660
ccaggggagggga atagccttg ccaggggagggga atagccttg ccaggggagggga 720
ccaggggagggga atagccttg ccaggggagggga atagccttg ccaggggagggga 780
ccaggggagggga atagccttg ccaggggagggga atagccttg ccaggggagggga 840
acgggagggga atagccttg ccaggggagggga atagccttg ccaggggagggga 900
ttaaataatgct atataatgca ttaaataatg ctataatgca 960
```
---continued

ggtacgtat cccttccctta ctattocata acaagctctt tggccacacac tattctcttatt 1020
ggttatactgc caaactcctg gctcctcgag aacctacaag aactgttatt ttggctagatt 1080
gggttcccat ttatatattc caaactccta tatccacaac cgcgcttctcc gtcggtcgcga 1140
gttttattgg aacgtaacgt ggatctccca cgccagctct ccggatcgtg tccggacactg 1200
ggttcttcct cctggcacgg ccgtcctctta ccctccagcc ctggtccctat gctcctcagc 1260
gtctagtcgc gctggtcgcag tcttgcgcc taaaacgtgga gccggtcgtt aggccacgca 1320
cacgccgcc caccaagctgt gcggtcccat aacggtctgc aagcgcctcg aacgccagag aacgagctag 1380
tgacggtggt agaattggtggt cacgacagctg aacgccagag aacgagctag aacgagctag 1440
caagactagtc aggccggttc gttgtggtat tctgtaagca gtcggtacag aacctcggtg 1500
cgcggtcgttta aacggtggtg ccggatcgtg tccggacactg aacctcggtg cgccgctgagc 1560
cacccagca taaaacgtgga gtcgttcctg ttgctcgggt ttttcttgca 1620
gtctcgggct cctgtctacgg aactggaccc cctgcggtcgc ggcggcacc aagagactga 1680
cacgacaggt ggcggcgtcg gcggccagca ggcgctccgg ggcgctccgg ccaacgtgtg 1740
caagactagtc aggcttgctg gtcggtacag aacctcggtg tccggacactg aacctcggtg 1800
cgcggtcgttta aacggtggtg ccggatcgtg tccggacactg aacctcggtg cgccgctgagc 1860
cttggcagc aacgagctg gatccgctgg ccggtccgtg ggcggcacc aagagactga 1920
ggcggcgtcg gcggccagca ggcgctccgg ggcgctccgg ggcgctccgg ggcgctccgg 1980
gtctcgggct cctgtctacgg aactggaccc cctgcggtcgc ggcggcacc aagagactga 2040
cttggcagc aacgagctg gatccgctgg ccggtccgtg ggcggcacc aagagactga 2100
aaccgcgtcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 2160
aaccgcgtcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 2220
gtctcgggct cctgtctacgg aactggaccc cctgcggtcgc ggcggcacc aagagactga 2280
aaccgcgtcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 2340
caccgcgtcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 2400
caccgcgtcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 2460
ggcggcgtcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 2520
cgcggtcgtcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 2580
cgcggtcgtcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 2640
cgcggtcgtcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 2700
ccggtctcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 2760
ccggtctcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 2820
ccggtctcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 2880
ccggtctcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 2940
ccggtctcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 3000
ccggtctcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 3060
ccggtctcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 3120
ccggtctcct cctgtcggga cggcgttcgg aaccgcgtct ggcggcgttg tgcggtcgtg 3180
tggtcagtaa aagatcagag ctcttcggag tgcggtcgtg tgcggtcgtg 3240
-continued

gacgggttc atatattggat gtattttaga asataaaccg ataggggcttc cgacgccatt 5580
tccccggaa gtcggacgct acgttaagaa accattatca atctgacatc taacctatataa 5640

aatagggcct atcgaggggtc cctctctgct cgccggtttc gggtgatgcag gttgaaaccct 5700
tgccactctg cgcttccggg agacgcttc acgttgctgt tgaaggtgatg ccgqgagcc 5760

aacgcggcct cagccggcgct acgctggggtgt tggggggtgc cgagccggtgc taaaattgctg 5820
ggcctcgag cagatggtgca tcgagctgg ccatacggg ccagggccagty cgcaggyt 5880
cgcagagga asatacggcga tcagatggygct t 5913

<210> SEQ ID NO: 94
<211> LENGTH: 3017
<212> ORGANIZATION: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: VR16682

<400> SEQUENCE: 94

tgcgtgcgtt cggctgatgc ggtgaaaccg tctgacccat gcaatccgg cggacgggtca 60
cagcttgtc gcacgctgc gcggcggcca gcaagccggc tcgagggcgc tcgaggggtg 120
tgcgcttt cgtcctgcgt ttcattaatcg ccggacgccg caacctttg cttgcgggtc 180

acctttgtt ccacccgctgc tccagcgcgc tctgccgcg gcctgcgcgc ccctcctgtg 240
tgcgcttt cggctgatgc ggtgaaaccg tctgacccat gcaatccgg cggacgggtca 300
cgccggcccc tgcggcccttc tgcagctgg ccagtggccg cctgcgggcttc gcctgcgcgc 360

tgcgcttt cggctgatgc ggtgaaaccg tctgacccat gcaatccgg cggacgggtca 420
cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 480
tgcgcttt cggctgatgc ggtgaaaccg tctgacccat gcaatccgg cggacgggtca 540
cgccggcccc tgcggcccttc tgcagctgg ccagtggccg cctgcgggcttc gcctgcgcgc 600
cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 660
g_gcgccggcccc tgcggcccttc tgcagctgg ccagtggccg cctgcgggcttc gcctgcgcgc 720
tgcgcttt cggctgatgc ggtgaaaccg tctgacccat gcaatccgg cggacgggtca 780
cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 840
ggcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 900
ggcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 960

tgcgcttt cggctgatgc ggtgaaaccg tctgacccat gcaatccgg cggacgggtca 1020
ggcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 1080
ggcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 1140
ggcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 1200
ggcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 1260
ggcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 1320
ggcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 1380
ggcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 1440
ggcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 1500
ggcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt cgcctggtt 1560
-continued

ggcttttgc ctaggtgcc cgcocctcga cgacacact acaaaatgac gctcaagtca 1620

gagttggcy caaaccyacag acctataacg ttccccctgg gaaaccttcg 1680
cctgctct ccctgccgac cctgcgcgt taccgcgct cacgcgccttc tccctcttc 1740
ggcaagctg gccttccctc taagctccag ctctggtatct tctagctgctgctcttc 1800
tgctccac agagccagcg cccctctcac cccgcgcctgc gcgctttc 1860
cagcatactt cgtgtgctgt ccacccgctg gacacactac gctctccag cggccagctg 1920
cctgctctac aggattcgca gaggcgagtt tggaggctgt gtctaccagcttctcagctg 1980
gttgctacac tcagcggctgt ctagctgaac gatatttgcg atgctgctgc tctgcggct 2040
aggtccttct gaaaaagact tggctgctgt ctgcgtgccc gaaaaacacc cggctgctgtc 2100
cggggttcctg tttttgctgc agcgcgacat tccgcgcgac aaaaaatgac cttcgggct 2160
tccctgtgc tttccctctg ggtcctgcag tgcgtgacgc gaaacctacac gtttggtct 2220
ttttgctgtgc agatctctac aacgactctt caccctgtac ctctattatt aaaaattgac 2280	ttttacttc atctatgagtt cacagtgctg atcttgttgctgctgctgct 2340
caggtgaccg cctatctgctg ctgctgtgctg atctgctgctg ccctacagctgctgctgct 2400
ggggggggct gcggctgc tggctgtgc gagaagctgt gctgccagcc ggtgggttgac 2460
aactcccaaca ctctcagctg ctgctgtgctg ccctacagctgctgctgct 2520
ggcttcag cctgtgctgt cggactcctgcc atccctgcag ttttctctctg gttgggttgac 2580
aagacgtcgtg atctgctgtct ttcacagcttc aaaaaatgac gagaagccctgc aacagctgctgctgctgct 2640
tccctctgctg ctagcctcag ccaacctgcg tggctgcag tgcgtgacgc gaaacctacac 2700
aaaaacctt gccactaccc caaatcctg ccctactattct gttgggttgac 2760
tattggacaa aacccggttct ctcacctgcg cccgctgctgctgctgct 2820
atggcaagctccag cctgcgcggtgt cccgctgctgctgctgct 2880
aatttctgct cttactgttc atcgctggggt cgggctgctgctgctgct 2940
tgctgctgtccat cccgctgccct tggctgcagctgc tggctgcag tggctgcag 3000
cttggctctc tagctgctgtc atgcctgctggt cgggctgctgctgctgct 3060
tggctgcag gaaaacctgc atctgctgctg ccctacagctgctgctgct 3120
aactcccaaca ctctcagctg ctgctgtgctg ccctacagctgctgctgct 3180
tcttacttc ggaagctgctg ctcctcctgcgtt cgggctgctgctgctgct 3240
ggcttcag cctctgctgctg cccgctgctgctgctgct 3300
tccctctgctg ctagcctcag cccgctgctgctgctgct 3360
tgctgccact cggctgctgctg cccgctgctgctgctgct 3420
tccgtgcagctg cggctgctgctg cccgctgctgctgctgct 3480
gagacgcagcc cccgctgctgctg cccgctgctgctgctgct 3540
ggctgcagctg cggctgctgctg cccgctgctgctgctgct 3600
ggctgcagctg cggctgctgctg cccgctgctgctgctgct 3660
tgctgctctgc cggctgctgctg cccgctgctgctgctgct 3720
cttggctct cggctgctgctg cccgctgctgctgctgct 3780
aactcccaaca ctctcagctg ctgctgtgctg ccctacagctgctgctgct 3817
---continued---

SEQ ID NO: 95
LENGTH: 4022
TYPE: DNA
ORGANISM: Artificial sequence
FEATURES: 
OTHER INFORMATION: VR4764, Ligation of VR4756 RV-EaLI into VR10682 RV

SEQUENCE: 95

tcgcgtgttt cgtcgagag ggtgaaacc ctgtgacact gcagtcgcccg gaagagctca 60
cagctttct gcagcgtgat gcgggagaga gcacagcccg tcagggcgcg tcaggggggt 120
tgcgcgggtcg tgcggctcgc tctgaccat gcggactcct gcagcatgtga ctgcagctgc 180
acctattgt gcctctctcg tctgtgacgc tctgtgagct gcctgttatg ccatatctg 240
cgtctgctgg tgtgtgtgag ggtcgtctgg tagctgtgca gacaaactta agctacaca 300
agcgaagcttg cagcagacac tgtcctgag aacatgcttc gggtagggcg ttggagctgtg 360
cctcggcatt taccagagcg actacagctg atctcagggg acatggttat gtctgagcga 420
aagcggagcc tccgtgtgta cgcgcttccg aagccctccg gaattataca gttctcggtttc 480
tgccataag ggggggatg tccttcatgt cacatccttc gcctctctcg gcagttgtat 540
cgcggggtag gacagctcgg tcacagagg ccagagctcg ccggtgtggg tagtttgtg 600
aagcttcatgc gtcaggaaag cacagcggctg cagctggcttg aatctgctttc 660
gacaacac gtaactgagat ttcacgagag tattgatattt aagttgcggag tctgataccto 720
tagcgcactt tcgcaagctc accacatggtgc ttcggacttc gacgcagcag tcagctgctg 780
ccgcacgcgt gcagcctcgcc accagtcgctg ttcgcaagctc acgtgacgcc 840
ttcatctct tcatgtgcgc catcggccggt ctctacagcgc accagcgcg ggcagctgtg 900
gatcgtctttt cgcggggacct ccgagctctg ggcgtctcga ttcagggctg aacagagaga 960
cccagcctctt ctcagctgtc gcagcggatgt ttcgggttcc tgcagggctgc acgcggtgcc 1020
agtgcgagag agctgggagac gcagctgcct gcagcgtcttt gcacagatgt gctgattcc 1080
cccactagcac ggcagcacgt cacagcggctg ccagcgttgctg ctctacgagc gtc­gatacctg 1140
cgggggacg aagcaacctg actgctgttc ttcgctctgt cttgcagcgc 1200
ttcctgatc acggttggg ggcggttacac cttcaggtgg ccggttgcgt gcagcgttg 1260
cagctgcagag cagctgctgg gcctcgagcc gcagactcata gcggaatggt gcaacagacc 1320
aacagcatc aaggtccagt gtcggcgcg cggggtcgag cggagcaacgc gcagctgtgc 1380
agcggagac ctagcagcgc gcgggagaga gcggaagagc tgcgagtcg gtcgagtcg 1440
agcagcagac gcagcggctg gccgggctg gcagcgtgctg ctcgctgctgc 1500
aacaggtgt ccgctggtctc gcagctgctgc gcaggtctgc ggcaggtctgc 1560
cgtctcagt ccagcggctg ccagcggctg gcagctgctgc gcaggtctgc 1620
cttggtcttt ttcggtggct gcctcagcgc gcagcggctg gcagcgtctgc gcctggtcttc 1680
agcggagac ttcggtggct gcctcagcgc gcagcgtctgc gcctggtcttc gcctcagcgc 1740
gagctgtgct gcagcagact gcagcggctg ctcgctgcc gcagctgctgc gcctggtcttc 1800
cttggtcttt cgggctgctgc gcctcagcgc gcagcgtctgc gcctggtcttc gcctcagcgc 1860
cttggtcttt cgggctgctgc gcctcagcgc gcagcgtctgc gcctggtcttc gcctcagcgc 1920
actgctcatt cctctatcag ctcgctgc gcctcagcgc gcagcgtctgc gcctggtcttc 1980
GTTTGGGTTG CICTGCTTTT ACTTCTACCA CTGCTTCTCG GCATAGCTGA

GACAAAAGAC GCTGCTGCG TTTTTTTCCT CTGCTCTGCG GCTGCTGCG

TATTTTTTTT TTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG

CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG CTGCTGCTGCTG
tcgggttgc aggagagag ctcacatcag gctaacgctg ttagctagac cacatcagc
4320
tgaacacttt cgacagctgt acctttttcgt tcttttctgc aaactctctg ccacagcgc
4380
tttccctccaca tcttttctgc ttagctagac cacatcagc
4440
taccctata aacacattgcat gttttttgct ggggactcag tggaggtttc
4500
tggtgtag ctacagttgatta ctttagttg ttaaacagc
4560
tgctgtag atatatccac gttttttgct ggggactcag tggaggtttc
4620
tggctccccaa cccccccccc ccttttttacctct gttttttgct ccacagcgc
4680
tacatcactg agagccagt gaaccacact gtttttttct tttttttgct ggggactcag
4740
taacagcactc cttctctctc ggggactcag tggaggtttc
4800
tgctgtagc ggccagcccc ccacagcgc
4862

<210> SEQ ID NO: 96
<211> LENGTH: 5941
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE: OTHER INFORMATION: VR765, Ligation of NP from 4762 into VR10682
<400> SEQUENCE: 96

tcgggttgc aggagagag ctcacatcag gctaacgctg ttagctagac cacatcagc
gccgctct gcagccggag cacacccctg caagccggag cacacccctg
tttccctccaca tcttttctgc ttagctagac cacatcagc
accccaatgc ctttagttg ttaaacagc
accccaatgc ctttagttg ttaaacagc
accccaatgc ctttagttg ttaaacagc
tgcagctg ccacagcgc
agagccagt gaaccacact gtttttttct tttttttgct ggggactcag
agagccagt gaaccacact gtttttttct tttttttgct ggggactcag
taacagcactc cttctctctc ggggactcag tggaggtttc
agagccagt gaaccacact gtttttttct tttttttgct ggggactcag
taacagcactc cttctctctc ggggactcag tggaggtttc
gtgtcggagc tgcgtnaat gatcaagaga ggcatacagc acagaanctt cttggagagc
1440
ggcgacgga caaaacagc agagcgaca ggaagatgt ggcacactct gaagggcag
1500
ttcacagcgc cgcagcgatc agctacgtag gacagggtcc ggaagaagcg aaaaaaagcg
1560
aaagcgagga tgtagacct gactctccg ggcgaagaggg ctcgaccct gaggagcag
1620
tgaggcctcg aagagctgct ggcgacgtgc cttacgacg gcacgagggag cagctgtcat
1680
acctgcgca gaggagggct cagcctgggt cgcagcagcc cttacggacc ggctgcagac
1740
agccgcgtgct acagctgtat caccccaaco gagaaccccg ccaccaagag cccgttggtg
1800
tcgcttgctcg gcgcacagtgc gcctccgcag acctgtggtc acctgagctt ctcacagagc
1860
acccgtgtgc ccacgagggc ggcagccagt gcgccggacgc acacgggacc ggcagaacgc
1920
aacatgcgca acacgagggc gagcgcgagt gcctcgtgag gcgcacagcg gcgccgagca
1980
aacacgaac gcgcgagggc aggcaacgag aggcagccgt ggcgcagcag gcgcacagcg
2040
ccacacca gcctcgtgag aacagccgac ctcgacgaag gcaaacaggg gcacagcagc
2100
acccgacca ccaggagggc aacagccgat gctagcaggg aagctcagat aaggtgtgag
2160
ggccgcacc ccaccagagt ctctccgcc gcgcacgagc gcctggagcg gtcggcagcg
2220
acccgcacca aaccagaggt ggcgctcgtc gcgcagcagc acgcggaggc gctacggctc
2280
gggcacggag aaggagggcg cacccgcgtg ggcgcgagtgc gcgcacagcg gcgcacagcg
2340
ggcgctggct gacacctggt gctagtgtgc gacacgtgccc gctgctgccc
2400
tccacgtgcc ctccgctgcc ccgcaaggt gcacccctca ctgcctcct ccacctcaaaat
2460
ggcgacgcc catgctgtct tctggagcc gcctcctcca tctgggcttg gcgcgggctgg
2520
cccggcgcgc ggccggggc cttgggtgca agtgctgggc cctgcggagc
2580
tctctccgc ccagctgtgc gaaatggtg gcgcggtcct gcggggccag aagagccag
2640
ggcacccgtct tgttgttgga cacccctgtg ccacgggtcc ggttttctgt gcgcgccccaa
2700
ctaggcggca actacagctg cccgggtggt ccgcctcaag tacacccccg taaagctgct
2760
ggcagggtcct ctccgccccg ccacggcctta ccaaccctaa cccagtgact ggcgtggga
2820
agcgacccgg ccaatccggt cctacaagtg cagcgggag gacacgctcg cccactcgctg
2880
agcgacccgg ccagacagatc atacggtgct ccctgcctgc ctcgtcgtcg acgcgtcgc
2940
tctgggtcg ccggtggccgc gcgtgattccc cggcggcgct ggcgggtgct gcgcggtgtc
3000
cacgacgatc ggcgagggc aaaaagggc gtaggtgtctg gggagttaa aacgcgtatt
3060
gaacccggtg tacccggggg gtaggtgtctg gggagttaa aacgcgtatt
3120
tccacacta gcacgtctac gcgcgggtgc gcgcggtgct gcgcggtgct gcgcggtgct
3180
ggctttccct ctctggtggt ctcgtcggt gcgcggtgct gcgcggtgct gcgcggtgct
3240
atatagctgg cttgccgtcc cttcgagag cttgcytgcttctcttgagct ctcgcctgc
3300
gttctctgtc tgtctgggtc ggcgcgctgc gcgcggtggc gcgcggtggc gcgcggtggc
3360
ttcacagcgc cgcagcgatc agctacgtag gacagggtcc ggaagaagcg aaaaaaagcg
3420
caacgcgtgc ccacgagggc ggcagccagt gcgccggacgc acacgggacc ggcagaacgc
3480
gccttacg ctcctgggtc gcgcggtgct gcgcggtgct gcgcggtgct gcgcggtgct
3540
tctgcgctgc gcctggtgct gcgcggtgct gcgcggtgct gcgcggtgct gcgcggtgct
3600
ccgggcccc aacacagcggt gtcgctggcg ttttgggtgt gtcgctggc gcgcggtgct
3660
gtasocgcaac tagggaactt ccatggaagt caasgggctg agatatttag ctaaactgoc 300
cactgagcag tacatctagc gatctcatat cacagtcgcc ccocatattg cctcctatag 360
gtataactgc cctcctgacc tacatgact tataagcttc tootactctag 420
cagctactct agctattatt ccctgctatt accatgctga tgcttcttat gccttacaag 480
aatggctgct gatagctgct ctactccgct actccacccc agtactccgct tataagcctc 540
aatggctgct gatagctgct ctactccgct actccacccc agtactccgct tataagcctc 600
gcccacacgg cgacatggg gctctagcgc gctaaggcct gctgcttct 660
gttaagctgag acctgacact cctctctggga cctctcttctt gctcttctttgaa 720
agacccgccg agacaggtcgc cctcctggg ggcctacggt gctctctttt catctcttct 780
cgctgctagc cgctgctagc actgagcctt ccacccctctg gctcctctgtctt 840
tactagctgag acctgagcct cctctctggga cctctcttctt gctcttctttgaa 900
gtgatgtatg gctctctttt gctctctggga cctctcttctt gctcttctttgaa 960
gtgatgtatg gctctctttt gctctctggga cctctcttctt gctcttctttgaa 1020
ggctgatctc ccctctttt ccctctctttg gctctctttt catctcttct 1080
ccacccctctg gctctctttt gctctctttt catctcttct 1140
tactagctgag acctgagcct cctctctggga cctctcttctt gctcttctttgaa 1200
ctctctctttt gctctctttt gctctctttt catctcttct 1260
ctctctctttt gctctctttt gctctctttt catctcttct 1320
cacccctctc ccctctctttt gctctctttt catctcttct 1380
agatctgctg atgtgctgtt cctctctttt gctctctttt catctcttct 1440
acacccctctc ccctctctttt gctctctttt catctcttct 1500
ctctctctttt gctctctttt catctcttct 1560
ctctctctttt gctctctttt catctcttct 1620
ctctctctttt gctctctttt catctcttct 1680
ctctctctttt gctctctttt catctcttct 1740
ctctctctttt gctctctttt catctcttct 1800
ctctctctttt gctctctttt catctcttct 1860
ctctctctttt gctctctttt catctcttct 1920
ctctctctttt gctctctttt catctcttct 1980
ctctctctttt gctctctttt catctcttct 2040
ctctctctttt gctctctttt catctcttct 2100
ctctctctttt gctctctttt catctcttct 2160
ctctctctttt gctctctttt catctcttct 2220
ctctctctttt gctctctttt catctcttct 2280
ctctctctttt gctctctttt catctcttct 2340
ctctctctttt gctctctttt catctcttct 2400
ctctctctttt gctctctttt catctcttct 2460
ctctctctttt gctctctttt catctcttct 2520
--continued

ttgcgagag gagggtttgct gtcggagccc ttccgaggtc tggcgagagc 2580
cagaggtgctg accggctgct gaccacccc gactgcttag tcagagaggc 2640
gtgcggctg tggcgcgtgc gcggcaacc ctggagcggc ctggcccctg 2700
cagaggctg agggcagacg cagcgagagc cagaagcgctc gcgcggagac 2760
cctgcaacc acggccggag ggcggcctg gttggagagc aggcttgctg 2820
caggcagagcg gcggcagacc gcggagggcg gcgttcgagc gcggagaggc 2880
cactttca gggagagagc acggtcttg ctggagcagc gagaagagc 2940
cggcgggac gcggagcagc ggctttcttc cgctccttcg cggctttctt 3000
cggcgggac gcggagcagc ggctttcttc cgctccttcg cggctttctt 3060
ggcggagagcg gcggagcagc ggctttcttc cgctccttcg cggctttctt 3120
cggcggagagcg gcggagcagc ggctttcttc cgctccttcg cggctttctt 3180
tggcagacta gatcctagag cttccgaggt ccgccggagc gcggagaggc 3240
cggcctcttc gcgtccttg ccgccggagc gcggagaggc 3300
cggcctcttc gcgtccttg ccgccggagc gcggagaggc 3360
tggcagacta gatcctagag cttccgaggt ccgccggagc gcggagaggc 3420
tggcagacta gatcctagag cttccgaggt ccgccggagc gcggagaggc 3480
ggcggagagcg gcggagcagc ggctttcttc cgctccttcg cggctttctt 3540
cggcctcttc gcgtccttg ccgccggagc gcggagaggc 3600
tggcagacta gatcctagag cttccgaggt ccgccggagc gcggagaggc 3660
tggcagacta gatcctagag cttccgaggt ccgccggagc gcggagaggc 3720
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 3780
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 3840
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 3900
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 3960
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4020
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4080
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4140
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4200
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4260
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4320
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4380
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4440
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4500
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4560
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4620
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4680
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4740
gtgctgtgg aggggagcct tcggcgttgc gcgtccttcg cggctttctt 4800
gaatggcaca acottttatgc cacacatcttc ctggggttcgacatc ctcgctgcctc 4860
gtttcgcaacgcaaaaacttgctttaggctgctgcttgctggaggtcagggcttc 4920
ggatggcttcgacgctgccagggctcgggcttcatggtggaggtccgctggcc 4980
cacaagtgcgagcgcagcgttgctgtggggggaggtggggcggcgctggggggggggggc 5040
cgcgttcggctgctggggttcgtggggttcggcgggttcgcttgctggcgg 5100
cgcgctgtgggggttccctgcttgctggggttcgtggggttcgcttgctggcgg 5160
cgcgctgtggggttccctgcttgctggggttcgtggggttcgcttgctggcgg 5220
ctggggttcgtggggttcgctggggttcgtggggttcgcttgctggcgg 5280
ctggggttcgtggggttcgctggggttcgtggggttcgcttgctggcgg 5340
cgcgctgtgggggttccctgcttgctggggttcgtggggttcgcttgctggcgg 5400
cgcgctgtgggggttccctgcttgctggggttcgtggggttcgcttgctggcgg 5460
ctggggttcgtggggttcgctggggttcgtggggttcgcttgctggcgg 5520
tggggggttcgtggggttcgctggggttcgtggggttcgcttgctggcgg 5580
ctggggttcgtggggttcgctggggttcgtggggttcgcttgctggcgg 5640
tggggggttcgtggggttcgctggggttcgtggggttcgcttgctggcgg 5700
ctggggttcgtggggttcgctggggttcgtggggttcgcttgctggcgg 5760
ctggggttcgtggggttcgctggggttcgtggggttcgcttgctggcgg 5820
gatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 5880
gatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 5940
ctggggttcgtggggttcgctggggttcgctggggttcgcttgctggcgg 6000
gatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6060
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6120
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6180
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6240
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6300
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6360
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6420
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6480
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6540
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6600
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6660
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6720
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6780
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6840
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6900
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 6960
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 7020
ggatggcttcgacatctacgctgtgggttcgctggggttcgctggggttcgctggggttcgctggtgcggcgg 7080
-continued

gaccacatgc cgccggtctc agacacggg ccgtggtaata gatctggttgc gcgtcctag ctggtggtcc 7140
tgccgctct tgggttggtt ccccctcctcc gcctggctct cggcatctgg gcgcggcact 7200
cggcattct cttctctatata aatgagggaa ttgctagctg attgtctgtag agatgtctat 7260
tctatttttt cgggcggggg gggcagggg agacaggggg aggttggagga gagaatagcc 7320
agcgtatagc gggtatgcggtt ggcgttattgg ggtgggttct ccccccccc cattattgaa 7380
gacacatca ggtgattttt ctcctatagc gataatatc tgaatcttt ctagaaatata 7440
ataacaaaggg gtcgtagggc accctcccc gaaaaggtgcc acgtgaacgct taagaaacca 7500
atatttatg gacatcatac tataaaaata ggcgtatccag ggcgtctcttt cgcgtcgcgc 7560
gtcgtgctgta tcggtctggaa acctctctac acctgcatgc cccgagggac gtcctcggct 7620
gtcgtaagc ggtatcgcgg cagcagaacc cccgtcgcgg cgcgttagcgc ggtttggcg 7680
gttggtgggg gttcgttatac tgtgcgtgctc catggcagct tgaatgacggt gtcacgata 7740
tcgcgtgtga aatacgcac agatgctgta gggagaaaaa ccgcattcaca tggcttgtat 7798

<210> SEQ ID NO 98
<211> LENGTH: 7790
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<222> OTHER INFORMATION: VR4767, Ligation of Inverted RSV3eg7 Into VR4762
<400> SEQUENCE: 98

tgcccattctg aatacgtggta tcctattcatt aatattcata tttatttttg tctgtggtca 60
acattcaggg cagttggttca tcatttttgtc atcaattttt aataggattc catccayggg 120
tcaatgattc atagcacaata tagaagcttc cgtctctatc aactctaccgt aatagtgcgg 180
cggtcctgac ccgccgaca ttcctctctc tgcgtccttt tcaatacata tggcctcata 240
gtaagcgccaa taggtttctt ccttgtgagt tgaagggggg agtttaaag gcctaatgcg 300
ccttggtggtc tattattgtc gctctgcttt aactatagttc accgtcttgtc ctcttcgc 360
gtaataagcg ccctgtcttc tttgatctgatat catttctgtc ctcttcgcc tatttggc 420
cattcatactt aactattagtt catggctaat cccagtggggc tgcgctttg tttctcatc 480
aattggggtt gatacgtcga ccctcccctc ggcgattcagg tcatgctcct catctccgc 540
aattgggggg tggggttttt gacccactaca cccataaacc ccgcatttcg taatcatac 600
gcccattctg aatacgtgggt gctatgtggt cttcctgtggt gcgttggtgg gcacgtcact 660
cgttattctg acgtcctgat ccctgtggtg ccgctcttcact gcctgtgtgtt tctctccata 720
agccacggag acgttggggtt ccctcctgcc gaagacgtt gcctgacggc gcgggttccc 780
ggtgcccaag ggcgtctttc aactctctct gcccaccccc cttcggttct tgcgttttta 840
agctgatcct agcttgggtt ggtcctggtgt tcatctctct ccagcttctcc cgtctttatc 900
tgcctgtgtct ggtcctggttt tcttctccct cctctataa aacccaccgtc gatcatttgt 960
ggtgcata cttccttcct tataactata cctctgctcttg cgggtgttcc 1020
gcgtctcttc atgcctcttc atgcctcttc tctcgcttctt ctgctctcttctc 1080
gcgttcgatt ccctcctctt ccctcctct tccgctccct tgtctccctt cgtgctctct 1140
gttttttta aacatacgtgc gcggcctcct ccgtgatctg ctggatgatg 120
| ggtccttcct cctgagcgcg gcagacctca cactcacaac cgtgcccctt gctcagcagc | 1260 |
| gctcctggtt gctgcccagc cttcgtgctt tacatgtaa ggcacagctt aagcacacga | 1320 |
| cactgcaccc cacaccacatg gcctgcccag gcagctggct ggctagctgg cggctgtgaa | 1380 |
| atgggctttg tagctgctgc cgcagcctgt gccagacagc aagacttaag gcagagccag | 1440 |
| aagcgagcag gcacgagcgg gctgtgcttt tctctacgtt tcagatgact gctagagttg | 1500 |
| cgtgaggtt cagctgtgag gcctgtgctt tcagatgact aaccttgctt gcagagcagc | 1560 |
| caacccacaca taaacccata gagctttttt ttttcacgcttttttta | 1620 |
| gtgcccagttc tgtgatcagc aacttcccgtt gactgccttc ccagggcagc agaagagta | 1680 |
| cgacagcggag gcagagagtc gcagagacct gcagagacct gcagagacct gcagagacct | 1740 |
| caagagctttc ggcagagttt tcagagagtc gcagagagtc gcagagagtc gcagagagtc | 1800 |
| cgactgagag gcacgagcgg cagccagcag cagccagcag cagccagcag cagccagcag | 1860 |
| ctctcctggtt agaagaggtg gcagtgctgg gcagagcagc gcagagagtc gcagagagtc | 1920 |
| gcagagtcag gcagagtcag gcagagtcag gcagagtcag gcagagtcag gcagagtcag | 1980 |
| gtacacactc gcagagagtc gcagagagtc gcagagagtc gcagagagtc gcagagagtc | 2040 |
| cgccagcctt aacccctgag gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 2100 |
| aacccgagcag cagctctggc ggcagagcag gcccagcttt gcagagcagc gcagagcagc | 2160 |
| cccccgccg gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 2220 |
| gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 2280 |
| gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 2340 |
| ccacccagcgc gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 2400 |
| cgccagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 2460 |
| ggccacagcgc gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 2520 |
| cttgccgctt gcgctgggct cctgccgctt cctgccgctt cctgccgctt cctgccgctt | 2580 |
| ccacccatgac gcgctgctgt gggccacgtgc gcgctgctgt gcgctgctgt gcgctgctgt | 2640 |
| gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 2700 |
| gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 2760 |
| gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 2820 |
| ccacccctgc gcgctgctgt gcgctgctgt gcgctgctgt gcgctgctgt gcgctgctgt | 2880 |
| ccacccctgc gcgctgctgt gcgctgctgt gcgctgctgt gcgctgctgt gcgctgctgt | 2940 |
| gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 3000 |
| gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 3060 |
| gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 3120 |
| gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc gcagagcagc | 3180 |
| cgctgcgtga caagctgtcgt gcgtgctgt gcgtgctgt gcgtgctgt gcgtgctgt | 3240 |
| gcctgcgtga caagctgtcgt gcgtgctgt gcgtgctgt gcgtgctgt gcgtgctgt | 3300 |
| cgctgcgtga caagctgtcgt gcgtgctgt gcgtgctgt gcgtgctgt gcgtgctgt | 3360 |
| cgctgcgtga caagctgtcgt gcgtgctgt gcgtgctgt gcgtgctgt gcgtgctgt | 3420 |
| cgctgcgtga caagctgtcgt gcgtgctgt gcgtgctgt gcgtgctgt gcgtgctgt | 3480 |
gaasccogcg ccagccattaa cagatatccg aggctgatgg ccggatccccc ccggctgtgtc ctgcggcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg
---continued

caaaaacct gctgtcagcc taccaagcct ttcgtgcctg ttacagtat tacccctcca 5820
tcgcacttc tgcggccac cttccttccg aocctgtttc aagcgtcagc 5880
agctgtttt gagaaaaaga gctaaagaa taccaaaatt caaagttcag atcccaatga 5940
taaccaagc aaccaacagc ggtctccttg aacttgagca tctgcacccc atcctgtttc 6000
tgtcttgct gcoaatacct aagaagaaoc ttctctgtgc gcagcagtgtgaca 6060
gtcggcagcc ctctactgtgc gtcgacatc tgccgtgcct gccagctagct tccagagc 6120
tcacagcct gccctcagtc tacccactag cttcagcttg gttggctgcg 6180
aacaacactc tggccagctcg ctactctaccg gtcgctgcgt tcggtcggac aagcagc 6240
gacgctgtag ggaacgcaag aactcgcttc catggtgcac atacaggtgaa aagcagc 6300
tacgctgatta cagccggccag tcgtcgctag atggcggccag tgcacagctgc aagcagc 6360
gcaagactac tgcgtctgct ttctgggacc cccagagctg ttaactctcc cttggtttt 6420
catctagtt ttcgacgctc gctcagcctc ttgctgtgttct cccctctttt 6480
tggcagcagc gtcctcctgc gatcgctttgc tctctcgtgc gcagcgtgcagc 6540
ccaccaacct cctctctgcag aagcagtgaag actgcggctg tttgagccagt ttcagcagca 6600
gtcggagat cttcggtttt ccagggccagc aacccctcgt gcctgctgcc gatcctggct 6660
ntcggcggtc atctgagcag gatcgagata aacccctcgt gcctgctgcc gatcctggct 6720
atgctggcag aacccctcgt gcctgctgcc gatcctggct 6780
gcggcaacact cttcaggtgc gtcctgctgcc gatcctggct 6840
ccaccaacct cctctctgcag aagcagtgaag actgcggctg tttgagccagt ttcagcagca 6900
cctctgctc aagcagcagc cttccagcgct atccctcagc aagcagc 6960
tttcctcgct ttcggtcct gcggcatacg cttcaggtgc aagcagc 7020
ttacgctaca cttctctgctc ttaccaagcctc gcctgctgcc gatcctggct 7080
gcgcaacact cttctctgctc ccagggccagc aacccctcgt gcctgctgcc gatcctggct 7140
cctctgctc aagcagcagc cttccagcgct atccctcagc aagcagc 7200
agctctgct aacccctcgt gcctgctgcc gatcctggct 7260
ccaccaacct cctctctgcag aagcagtgaag actgcggctg tttgagccagt ttcagcagca 7320
tggcggagac ttcagcagcagc aacccctcgt gcctgctgcc gatcctggct 7380
gcttccggtc ccagggccagc aacccctcgt gcctgctgcc gatcctggct 7440
ccaccaacct cctctctgcag aagcagtgaag actgcggctg tttgagccagt ttcagcagca 7500
tttacgctaca cttctctgctc ccagggccagc aacccctcgt gcctgctgcc gatcctggct 7560
gttccggtc ccagggccagc aacccctcgt gcctgctgcc gatcctggct 7620
ccaccaacct cctctctgcag aagcagtgaag actgcggctg tttgagccagt ttcagcagca 7680
ccaccaacct cctctctgcag aagcagtgaag actgcggctg tttgagccagt ttcagcagca 7740
tggcggagac ttcagcagcagc aacccctcgt gcctgctgcc gatcctggct 7798

<210> SEQ ID NO: 99
<211> LENGTH: 7790
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE: 
<223> OTHER INFORMATION: VR4768, Ligation of RSVPN into VR4756
-continued

<400> SEQUENCE: 99

tgacccaggc atacgttagc tctcatcatc aatagcta tttatatgg ccctatgca 60
aacattacgc catgctgaca tttatatagc acatatgcc aatagctggg 120
tctattgtc atagccatata tattaggctc ccgcttcctat cacttacgct 180
cccggctgac cggccacgcct taccgcttac acaggagta tggctccgta 240
gtacccacaga taggcacatt tctctgagct ctaggtggct atgatttgcc 300
cacttgagc cactcaagt gcattaatgc ccaatgacg ccccctagc 360
gtacttacgc cgggcaagcag tcaagctatgc tattaggctc ttaaaccgct 420
cagccacatc cctcgctgtg acatggattt aatagaggtg ccgttttctg 480
aatttggg agctaggcct ggttccccgt gccctcctag cattcgcctc 540
aatggaatg tttttggtg ccagccctgtt ccacccgctt cagactcggggt 600
gccgcatcgc gcggggtatg cgggacgtgc atggctgtggc cgctgtggccc 660
cgggctgtgc gcggggtatg cgggacgtgc atggctgtggc cgctgtggccc 720
agacacgctg gacatgct gcggggcct gagctggtgct ggtttggttg 780
cggggcctg gttttggtg ccagccctgtt ccacccgctt cagactcggggt 840
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 900
tgatgctata gctgtggtgg ctatggacca tttatgctctt gatgctata 960
ggggtacgtg atccggc cttccctttg ctctagcata cccatgccgct 1020
ggggtaagc atccggc cttccctttg ctctagcata cccatgccgct 1080
ggggtaagc atccggc cttccctttg ctctagcata cccatgccgct 1140
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1200
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1260
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1320
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1380
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1440
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1500
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1560
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1620
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1680
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1740
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1800
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1860
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1920
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 1980
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 2040
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 2100
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 2160
atggtcagta tctggttctt gcgggtgggg ctaataaccgc cggctgctctt atgttctga 2220

tcataata gagctacgag aacaagttgt tttgccccag atcagatgac aagagttggc 2280
agagttggc 2340

gcsaatgct ggtcagatg agccggcago ggggcccagt ggatattgta gcgggagcag 2280
gcattagtct cggcagcag gcggcagcag gcagcagcag gcggcagcag gcggcagcag 2360
gtacatgtct ttgtcattat ctttttttaaa acgtctat ctatttctat ctatttctat 2470
atctaatggt gacgctgcga cggctcctat ctgcgattag gccttgtctt tgtctctaat 2500
ctgcttctct tcttcacact ctctaactct ctctctactct actttctactct actttctactct 2520
gacagctgatt gatgtgtgtag ctgttatttt gactactact ctctctctct ctctctctct 2580
gaatgctgtg gatgtgtgtag ctgttatttt gactactact ctctctctct ctctctctct 2640
acgcgctgatt atccatagct atctctgtgatt atattatagact caggtactact gcagctgtg 2700
tgtgtgggttt tgtgtgggttt tgtgtgggttt tgtgtgggttt tgtgtgggttt tgtgtgggttt 2760
ggctgtgtgtg gatgtgtgtag ctgttatttt gactactact ctctctctct ctctctctct 2820
tgaccccact ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 2880
ctactatggtt gttcattatc gttcattatc gttcattatc gttcattatc gttcattatc 2940
actcagact ctactatggtt ggttccaggc cagccgcgtc ttttaacaggg ttttaacaggg 3000
ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 3060
ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 3120
agctgtgtgatt atccatagct atctctgtgatt atattatagact caggtactact gcagctgtg 3180
gagcgattatc ggttccaggc cagccgcgtc ttttaacaggg ttttaacaggg 3240
tgatcaaat gatgtgtgtag ctgttatttt gactactact ctctctctct ctctctctct 3300
actctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 3360
tgatcaaat gatgtgtgtag ctgttatttt gactactact ctctctctct ctctctctct 3420
ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 3480
ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 3540
actctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 3600
ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 3660
ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 3720
actctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 3780
ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 3840
actctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 3900
ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 3960
ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 4020
actctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 4080
cagctctac agactcagcag gcagccttaatt ggtcactagc agactcagcag gcagccttaatt 4140
cagctctac agactcagcag gcagccttaatt ggtcactagc agactcagcag gcagccttaatt 4200
cagctctac agactcagcag gcagccttaatt ggtcactagc agactcagcag gcagccttaatt 4260
cagctctac agactcagcag gcagccttaatt ggtcactagc agactcagcag gcagccttaatt 4320
cagctctac agactcagcag gcagccttaatt ggtcactagc agactcagcag gcagccttaatt 4380
cagctctac agactcagcag gcagccttaatt ggtcactagc agactcagcag gcagccttaatt 4440
cagctctac agactcagcag gcagccttaatt ggtcactagc agactcagcag gcagccttaatt 4500
-continued

gagcggggc agatacgtg ggaccccaac ttgaagcttg agaagaaccc gcacccctgg 6840
gggacctgg tgaattcggca ctctagccgg cacggagacc gcgatctgag 6900
ggagagcc ctgagatgtg ggaggggccc agacccagaag aggctgtctt cagagcag 6960
gggcagctgg agctggaggcc gagaagggcc ccacaacccca tgcgggttag ctctagctag 7020
ggcacggag gcagcatctt ctccggccac agccgagg agtacacacc ctcgaagctc 7080
ggcacggag gcagcctgct cagccagggc ctcggatcgg gatctgtgct gcgtctgctg 7140
tcgagctgct cgtggtggtt cccctcccccc gcgtctccttg tgcaccccgg cgcgctcagct 7200
tccagtctgct ttccatgtaa aaccaggaga atgagcatctg atgtgtctgg tagctgtcat 7260
tctttttgct ggggtggggtt ggagcgagcc cagacaagggg aggatgggga agccaaagag 7320
ggcagctgct ggggtggggtt ggagcgagcc cagacaagggg aggatgggga agccaaagag 7380
ggcagctgct ggggtggggtt ggagcgagcc cagacaagggg aggatgggga agccaaagag 7440
aacgacgagag gttccagcgg cgcgagcgg ccgaggtgcg cagccagctg ctcagactgac 7500
ncgtctctcg gtatctgctg ttataatcata ttctgacgg cgtatcgtct cgtatcgtct 7560
gtctctcgtc tgaactgtgac cccgatcgcc cgggcgctgct cctgctgctg ccctgccg 7620
gtctctcgtc tgaactgtgac cccgatcgcc cgggcgctgct cctgctgctg ccctgccg 7680
gtctctcgtc tgaactgtgac cccgatcgcc cgggcgctgct cctgctgctg ccctgccg 7740
tgcgctgctc tgcgctgctc tgcgctgctc tgcgctgctc tgcgctgctc tgcgctgctc 7800

<210> SEQ ID NO 100
<211> LENGTH: 7798
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: VR4769, Ligation of Inverted HP into VR4756
<400> SEQUENCE: 100

tggcctctgga atagcttggag tccatacttgta atatgtgaccc ttttaattgg cctatgyccaa 60
acattacccag catattggca tctaatgttag actattgtact atatacatc attaagtggg 120
cttactgtgagcc agctggacat caggttaatg cttaagccgtt aaaaatggcacc 180
cagtctgac cggaccaacga cccgactccaa cgtcgagtc caaagttgact gtttgcaccat 240
gtcggcac tgagcacac ccctgagcgtt cactgctt cccgactccaa cgtcgagtc caaagttgact 300
cattcggccatt cattcggccatt cattcggccatt cattcggccatt cattcggccatt 360
ggctctctc tggctctctc tggctctctc tggctctctc tggctctctc 420
cgtcctccct atcactcctc ctcatcctc ctcactcctc ctcactcctc ctcactcctc 480
aatggcctgag cggctgtggag taaaatggcacc gcgctatggtt tggctctctc tggctctctc 540
aatggcctgag cggctgtggag taaaatggcacc gcgctatggtt tggctctctc tggctctctc 600
gcccttattata ctcacattata ctcgtctggtct ctttaaatgct gggcttgg 660
ggttcttcg acgtccttcg acgtccttcg acgtccttcg acgtccttcg acgtccttcg 720
agggccttattata ctcacattata ctcgtctggtct ctttaaatgct gggcttgg 780
cgccttcg acgtccttcg acgtccttcg acgtccttcg acgtccttcg acgtccttcg 840
attcttcg acgtccttcg acgtccttcg acgtccttcg acgtccttcg acgtccttcg 900
tggctctctc tggctctctc tggctctctc tggctctctc tggctctctc tggctctctc 960

tgcgctgctc tgcgctgctc tgcgctgctc tgcgctgctc tgcgctgctc tgcgctgctc
---continued

gttacaacata cttctccatta ctatcata cctatgtct ttcgcaacaa tacatcatt 1020
gctatctag cccactctctg tccctctcag cctgacacgc acctgtctat ttcataaggt 1080
ggggcccct tttattttta ccaactcaca tataacacag cggcgtccccc cggcatcgcag 1140
gttttttata acacatggtt cggcaatcoca cgcgcaacctt ggtgacgatt tcggagcactg 1200
gctccttttc cgygtaaggg ggcaccctca ctcggcagcc cttgcatcag ctcgagagcg 1260
gctcaggygc cgcggccgag cccccctcag ttcagcagt gcgcgacgac ccgacagcag 1320
cagcacacac cccaccacagtt gtcggtcaca ggcggtgtcc gcgggttgcg gagctgtgaa 1380
gtggcctgtc aggatgtgtgg cgaagctggc agcgacgacac aagacttacta gcggggtcag 1440
aagagagttc acaggggtga agggggtcgt gctttttactgtctgttttga tctcgaggtta acctccttgg 1500
cggcctggtt acagcggttc cggcagaggt tcggacagtt ctcgagctgt gcgcgagcag 1560
ccccacacac tacatcagtt gcgcgacagc ttcagtgctc ttcagcagct cttttcagta 1620
gtcggcgtcttc agctgggtgc cgcgagcgcc ctcgagccag tcggagcactg 1680
tgcctcctct aatcgggtcc cctgctcttg ccaccggcgc atctcctggc agtctcaaga 1740
tgcctcctct gcggaaccacc cccaccacagtt gcgcgagcagc gcgcggttgc cagctcagct 1800
actacgctgta acagtcgctg gagaggttc cggggttcgg gcgacgagtc aagcagacgc 1860
tgcggcgttc agtctccttg ccaccggcgc cccaccacagtt gcgcgagcagc gcgcagtcgg 1920
aataacgctc gcacgcggtg ctacgctcata ttaaaacactt aagggggtca cccaccttc 1980
tggggtccac gccgaacagc tcggcggctc gtcggttcct gcgggtggtc gcagcgttt 2040
cattccacac cggactcggcc cttgagctcc cttggttggc tggggtcgc gcgaccagtc 2100
atgctcggcc agtctccttg ccaccggcgc cccaccacagtt gcgcgagcagc gcgcagtcgg 2160
tctccttcata acagacgacg ctcggtccttc gcgggtggtc gcgacgagtc aagcagacgc 2220
gcagactcggc ggcacacgct ccaccggcgc cccaccacagtt gcgcgagcagc gcgcagtcgg 2280
gcagactcggc ggcacacgct ccaccggcgc cccaccacagtt gcgcgagcagc gcgcagtcgg 2340
agtagctctg gttcgccagt gcggggttcg gcgggtgttc gcgggtggtc gcgacgagtc 2400
attacgctg gttcgccagt gcggggttcg gcgggtgttc gcgggtggtc gcgacgagtc 2460
agctcacgtg ttcgctcggt ttcggtccttc gcgggtgcac ggagacggtt ccgggtggtc 2520
aggagcgtcc ctcagcggct gcggggttcg gcgggtgttc gcgggtggtc gcgacgagtc 2580
agtagctctg gttcgccagt gcggggttcg gcgggtgttc gcgggtggtc gcgacgagtc 2640
aconctcctaaaa ckgcgcacg cctcagcggc gcgggtgttc gcgggtggtc gcgacgagtc 2700
tgcggcgcag ccggggttcc ctcagctggt ctcggtccttc gcgggtggtc gcgacgagtc 2760
tgcggcgcag ccggggttcc ctcagctggt ctcggtccttc gcgggtggtc gcgacgagtc 2820
tcggcgcag ccggggttcc ctcagctggt ctcggtccttc gcgggtggtc gcgacgagtc 2880
agationtgc acagtcgctt ctcggtccttc gcgggtggtc gcgacgagtc 2940
acagtcgctt ctcggtccttc gcgggtggtc gcgacgagtc 3000
ccgtcgggt gcgacgagtc ctcggtccttc gcgggtggtc gcgacgagtc 3060
tgcggggtt ctcagctggt ctcggtccttc gcgggtggtc gcgacgagtc 3120
agtagctctg gttcgccagt gcggggttcg gcgggtgttc gcgggtggtc gcgacgagtc 3180
aggagcgtcc ctcagcggct gcggggttcg gcgggtgttc gcgggtggtc gcgacgagtc 3240
-continued

tggcagcagg cacagacacg tgtgtaacgg attagcaag cgggtgtatg agcggtcgct 3300
agagagatct tgaagtggtg gctatcaac cggacagca gaaacagtgt atggagtctc 3360
	tgctgctgc tcgacagct aatctctga gaaagctggt atggagttt gtcgacgaac 3420
caaacacgcc ctggtaaggc tgggtatcgt tgtgacagc aggacaggac gcgcaagaa 3480
agnacactgc agagagagat cttgatgtttct ctcaggggtg ctcgacgca gcggacgaa 3540
aatctcgytt aagggacttt ccgctagaga ttctgaaaaa ggccttccaa ctagatcct 3600
	taaattsttas aataggttt ctaactca tcagatatt ctagatcag ctagatccgtc 3660
agctaccaag tgtttacgtg tggcagcagc accatcaagc ctagatcag tggcagcagtcc 3720
aatgtagcct gagccagggg gggggggtgc tgggcttgct ctcgctagaga aagtgagtgtc 3780
gacctacact cggcttgggt ccaccacata ccacgacagc aagtgaggg ccacggtcctt 3840
atgtgagcct tgtttgtagt gcagctagtt gtattttttgt actttttgtt ctcgctagaga 3900
cgcgtgctgt tgcggagag agttctgtgct tgccttca gctacgacgag ttcctggttt 3960
attacaactc ggccttttct cgcctttct ctggatcgtg tataagctgcc 4020
taaccactcg tggtagaaaaa aatctacgca cgtatcagct aataacgtgct tataacgtgct 4080
cagggtagc aataacacat tattaagaaaaa ggccttccaa cgaacacgcc 4140
cagggtagc coactacagt cgcctttct gttcgatgct tcgctgcttaccttccaa 4200
cacactcaac actattctgt ctcctctgtg cagaaataag ttgcacagtct gcgacacggcc 4260
cagggtagc gcgtagacatgc gtcgtcttgcc ggccttccaa aataacacat ctagatttctt 4320
gagcaacag ccgcagcgatt cagacgctc ctaactacag ctaactacag 4380
tacactctgct cggagacag cctctccgtg gctgccttca cgcctttct ctcgcaagtct 4440
aagctttcag ccagctgttct gatccgggc ccacggttct ctcgcaagtct 4500
tgacgacag cctctccgtg gctgccttca cgcctttct ctcgcaagtct 4560
tgacgacag cctctccgtg gctgccttca cgcctttct ctcgcaagtct 4620
cgcgctttct cttccacgcc cgggctttct ctcgcaagtct 4680
catcctcct ggccttctc cggagacag cctctccgtg gctgccttca cgcctttct 4740
tgcctctggtt cacactcaac ctaactacag ctaactacag 4800
taccactcaac actattctgt ctcctctgtg cagaaataag ttgcacagtct gcgacacggcc 4860
tatcctcagc cagcctaggcc cccctttctg atgccttggt ctcctctgtg cagaaataag 4920
tatcctcagc cagcctaggcc cccctttctg atgccttggt ctcctctgtg cagaaataag 4980
tatcctcagc cagcctaggcc cccctttctg atgccttggt ctcctctgtg cagaaataag 5040
tgcctctggtt cacactcaac ctaactacag ctaactacag 5100
tgacgacag cctctccgtg gctgccttca cgcctttct ctcgcaagtct 5160
tgcctctggtt cacactcaac ctaactacag ctaactacag 5220
tgcctctggtt cacactcaac ctaactacag ctaactacag 5280
tgcctctggtt cacactcaac ctaactacag ctaactacag 5340
tgcctctggtt cacactcaac ctaactacag ctaactacag 5400
tgcctctggtt cacactcaac ctaactacag ctaactacag 5460
tgcctctggtt cacactcaac ctaactacag ctaactacag 5520
-continued

gcccctctgt gcgtgctgtt gcggcgcggt ctggtgctcg agggcgcaga tggcgcagt ctggtctcgc 5580
agccgacgg cgcgtgcgct gcggcgcggt ctggtgctcg agggcgcggt ctggtctcgc 5640
cgcgtgcgct gcgtgctgc gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 5700
agccgacgg cgcgtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 5760
tgcttcgg gcgtgctgc gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 5820
agccgacgg cgcgtgcgct gcgtgctgc gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 5880
tacggcgg cggcggcgc gcgtgctgc gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 5940
ctggtgcgct gcgtgctgc gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6000
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6060
ctggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6120
atgctgctgt gcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6180
ggcggcgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6240
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6300
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6360
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6420
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6480
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6540
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6600
atgctgctgt gcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6660
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6720
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6780
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6840
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6900
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 6960
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7020
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7080
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7140
tgcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7200
atgctgctgt gcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7260
atgctgctgt gcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7320
atgctgctgt gcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7380
atgctgctgt gcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7440
atgctgctgt gcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7500
atgctgctgt gcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7560
atgctgctgt gcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7620
atgctgctgt gcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7680
atgctgctgt gcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7740
atgctgctgt gcggtgcgct gcggtgcgct gcggcgcggt ctggtgcgtgc agggcgcggt ctggtctcgc 7798
---continued

tccggcttt cyggtgatag gcggaaaacc tgtgacacat gcagtcctcc gcaggcttca 60
cagtcttct ctagcgcatt gccgggacca gcacagcctg tcagggcccg tcagcgggtg 120
tgcggcgtt tcgctggtcct ctttaacctg cgccattcgc gcagattgta ctgsgagctgc 180
acactatgc tggtaatata ccgacccatat gtaatgcagg aaaaatcggc atcaatggtag 240
catgctgtca ctctctctct tttgctgtag gtcgctgtag tagcgcagc gaaaaattta 300
agctaacca aagcagagct tgacccagac tgcgtgagc aatctgctta ggcttgagggg 360
cttctctgcttt tcgagctccag atataccgtg atctgagggg aatgctcttg 420
gtggagtgg cagagcgcttg ccgctgcttc agcctcctga gcgaacagta 480
httctgtcctcc tgcgatagcag ggggaaaaat gattccttatt ctaactatct ttgctcttc 540
acagctctga cgtgctgtta gcacactggc tttcacgagc aaaaaacctg cgtgctctcg 600
cgctgtgctt aagcagagct gttctgcttg gttctgcttg gttctgcttg gttctgcttg 660
acagcctgcc gcacacgagt tgcagcagc cgagcctgcc gcagcctgcc gcagcctgcc 720
cogccaaaa cgaacactggc tggcgacagc tggcagcagc tggcagcagc tggcagcagc 780
cagctctctgctctc tgcgataggc ggcgctgctt ctaagcctgc cttgcagcctgc gttggtgctg 840
cgctgtgctt ccgctgcttg ctctctctct ctctctctct ctctctctct ctctctctct 900
acagtggcct tggcagctgg ccgctgcttt ctcgctgtct gtcgctgctt gcagcctgcc 960
cgcgcctctc gcacgcagct cgccctcggt ttcgctgctt tcgctgctt gcagcctgcc 1020
tgtctctgct cttctctctc tgggctctct ctcgctgctt cttctctctc cttctctctc 1080
atctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 1140
gcgcgctt ctctctctct cgcgctgctt aagcctgcct ctcgagctct ccgctgctt 1200
acagcctgcc gcacacgagt tgcagcagc cgagcctgcc gcagcctgcc gcagcctgcc 1260
ccagcctgcc gcacagcttg cattctgtgg ccgctgcttt ctcgctgctt gcagcctgcc 1320
atctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 1380
acagcctgcc gcacacgagt tgcagcagc cgagcctgcc gcagcctgcc gcagcctgcc 1440
atctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct 1500
acagcctgcc gcacacgagt tgcagcagc cgagcctgcc gcagcctgcc gcagcctgcc 1560
acagcctgcc gcacacgagt tgcagcagc cgagcctgcc gcagcctgcc gcagcctgcc 1620
acagcctgcc gcacacgagt tgcagcagc cgagcctgcc gcagcctgcc gcagcctgcc 1680
acagcctgcc gcacacgagt tgcagcagc cgagcctgcc gcagcctgcc gcagcctgcc 1740
acagcctgcc gcacacgagt tgcagcagc cgagcctgcc gcagcctgcc gcagcctgcc 1800
acagcctgcc gcacacgagt tgcagcagc cgagcctgcc gcagcctgcc gcagcctgcc 1860
acagcctgcc gcacacgagt tgcagcagc cgagcctgcc gcagcctgcc gcagcctgcc 1920
acagcctgcc gcacacgagt tgcagcagc cgagcctgcc gcagcctgcc gcagcctgcc 1980

-continued

atgcatctct acgaatttta acaacgcttg gaaagaggg ctctaatag gaggatacc
2040
tgacctctag agggaagat atcgaagaag acaacgcttg gcttgctgct ctggagccc
2100
ttcattttttgct acgtagagcc tgtgcaattc acgtgctagc atggcaatta 2160
gtggcagcgc atctgctgttatgtgctgggct gctgcgttcc gccagcggc gaaggtgcag
2220
cattctccttg ctctctctctt cattcataa aaaaagaggg aacaagccttg aagctttgctc
2280
attcgattttc gggggtcagg gttgcttcc ccggtaaccc cgagttggtc ggggaaga
2340
gtaagccgtgc ggctggtgtg tggctttaa ggggtcagc ggctgcatag aatgtgaccc
2400
gtcctctcttg gcgtgctagc ataccccctctt ctggcaacca cccctgcccag
2460
goctgcttcat ccgcaacacac gacgacaggg gaagagaggg gaggcagttt gggcaaatag
2520
cctccttcgg aacgctacga gctctctcag gctgctctcc acggacaccc
2580
gaaacacac aacgcactagg ccgtgcaagg ccgaagaggg gagtcgacag taacagccac
2640
ggagagagggg ccacgac accagcctt cattggttac cgctctcag cccagtccgg
2700
ccttggattt aagcttaccc gccgggtggt tgtcgttccg tgcacagtctg atcgcttggc
2760
tgcgctgctg acggctctcc aagcccttctgttattt cctgcttgcc
2820
ccggagttaa gggagagacc gggagagccc ccgctgctgct cccagccgg cgggagacg
2880
gaaaggggcc aagggctgcct gtcgtgttct ttcacgctg tggagctggc aagcagcctt
2940
tggaccacc cccgcagttt aacgctgtgg accgagactt ccgtgctgctgcc
3000
gctgccggg ccggagcttcc ctgctctctt ctgggaggt cggcggttcc gggcagggg
3060
cattcgctgc acgtctcctct ccgagttcgg ccggagcttct gctggcgctt ctggctggag
3120
tcggctgtgct gctgtcgttc ctgcagcctgg ctggctggggt gcggcccccct gcgtggtggc
3180
tggccggcag ccgctgtgct tgggctgggct gctggctggct acggagccgc acggagccgc
3240
cgtctcctgc ccgctgtgct tgggctgggct gctggctggct acggagccgc acggagccgc
3300
gggtgctgtgct ccgctgtgct tgggctgggct gctggctggct acggagccgc acggagccgc
3360
gttcattctgt ctggccggcc gcgggagcctt ccgctgctgctt ccgctgctgcttt
3420
gcgggacacc cccgacttgg gcggggggtt ttcacgctg tggagctggc aagcagcctt
3480
gcgggggctt ccgctgctgctt ccgctgctgctt ccgctgctgctt ccgctgctgctt
3540
gactagccgctt ccgctgctgctt ccgctgctgctt ccgctgctgctt ccgctgctgctt
3600
gtcgctttc aacgctttc cccggctttc gcgggagcctt ccgctgctgctt ccgctgctgctt
3660
gcgggagcctt ccgctgctgctt ccgctgctgctt ccgctgctgctt ccgctgctgctt
3720
gtcgctttc aacgctttc cccggctttc gcgggagcctt ccgctgctgctt ccgctgctgctt
3780
gtcgctttc aacgctttc cccggctttc gcgggagcctt ccgctgctgctt ccgctgctgctt
3840
ccgggagcctt ccgctgctgctt ccgctgctgctt ccgctgctgctt ccgctgctgctt
3900
cggagctttgt tgggctgggct ccgggagcctt ccgctgctgctt ccgctgctgctt
3960
tcgctttc aacgctttc cccggctttc gcgggagcctt ccgctgctgctt ccgctgctgctt
4020
gtcgctttc aacgctttc cccggctttc gcgggagcctt ccgctgctgctt ccgctgctgctt
4080
gtcgctttc aacgctttc cccggctttc gcgggagcctt ccgctgctgctt ccgctgctgctt
4140
gtcgctttc aacgctttc cccggctttc gcgggagcctt ccgctgctgctt ccgctgctgctt
4200
gtcgctttc aacgctttc cccggctttc gcgggagcctt ccgctgctgctt ccgctgctgctt
4260
-continued

eaatccacct gaggagaac gtagcaggt gagaagggca aagccttag catttttttc cgttctctgg ctacccagga gcattacag cgtctctggt 4320
caggctttgt caacagggga gcaattacag tcgggctgac aacagta cacagcagc 4380
cattcatcct tcctcctact gcctctgg caagcagcag gagaagctg ctgtggcttc 4440
ccttcacaaa ccagagctca atggaaagct tcgttctctc gcttcatc 4500
tgctcctact atgctgactt gccgtcttca actgctcttc gctgggtctt 4560
gtgctgtcc atcgcctcag cattactgtt gccggcctt tcgcttcttt 4620
ctgttgtagc ccctcgttgt ttagcgaag ttagtactggt atagctttttc ggtggaagtt 4680
cgcttctgtc cggctgctt gagttgggt gcctcagcag cctcagcctt 4740
ctgtctggc cttctgcttg cttccaggag ctgctgtcttt gcctgctgct 4800
ctgttcgta ttctcctgtt cttctcctgtt cggctgtctt gcctgtgtctt 4860
cctctctgtg tattgtatgttt tttcttttgttt tttcttttgttt tttcttttgttt 4920
tcttgctgtt cttcctgttg ttttcttttgtt ttttcttttgtt ttttcttttgtt 4980
ttgcctgtg cctgtctgtg cctgtctgtg cctgtctgtg cctgtctgtg 5040
aatctcttc tgctctagct cgctctctct ctctctctct ctctctctct ctctctctct 5100
gaacatcgtc cttctcctgtt cttctcctgtt cttctcctgtt cttctcctgtt 5160

5161

<210> SEQ ID NO 102
<211> LENGTH: 5684
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<222> OTHER INFORMATION: VR4771, NF Insert Repacing WHV Insert in VR6430
<400> SEQUENCE: 102
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 60
cgttcgggt gttcgtcgt cttctggtgt cggctgctg cttctggtgt cggctgctg 120
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 180
cgttcgggt gttcgtcgt cttctggtgt cggctgctg cttctggtgt cggctgctg 240
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 300
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 360
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 420
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 480
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 540
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 600
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 660
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 720
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 780
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 840
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 900
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 960
tcgggtacg ctgctacgct cttctggtgt cggctgctg cttctggtgt cggctgctg 1020
<table>
<thead>
<tr>
<th>Base Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGATGCTGA GACGACTGAG TGGCTGACG GGGCCGACC AGACTAATA GCTGACAGAC</td>
<td>1080</td>
</tr>
<tr>
<td>TACCACTGAA TCTCTTTTCC TGGTCTTTT CTGCTGCTAC GCTGCGAGA TCTGAAATTCC</td>
<td>1140</td>
</tr>
<tr>
<td>GCGCGCTGG CCCCGGGG TCCAAAGG AAGCCGCGA AGATGAGAAC CGCGGGGAG</td>
<td>1200</td>
</tr>
<tr>
<td>AGACGAGCA CCGCAGAG GAGCGCAGAG GTGGCCAGA TGATGAGGG CATGCGCAAGA</td>
<td>1260</td>
</tr>
<tr>
<td>TTCTTCTCC ACTGTCGAG CTAGCGTAAG CTGACCGACT AGAAGGGGAG AGTACGACAG</td>
<td>1320</td>
</tr>
<tr>
<td>AACGCTGTA CCCTGCGAG ATGCTCTTAC AGCGCCCTG AGCGAGAAC AAGACGATAAC</td>
<td>1380</td>
</tr>
<tr>
<td>CGGCGAGGAC CCAGAAAAAG CCAAGAAG GGGGGGCCC CACATCAGAA</td>
<td>1440</td>
</tr>
<tr>
<td>AGATGCGGAC GCGAGGTTG AGCGGCGTT TGCTGTGACG ACAAGGGG AGACACAGAA</td>
<td>1500</td>
</tr>
<tr>
<td>ATCTGAGGC GCAGCCGCG GGCGGGGCG GCTGCGCGG GCCGAGCCAG CTGTGATGTTG</td>
<td>1560</td>
</tr>
<tr>
<td>TGCCGCAGCA AGCTGGCGAC CCAACCCG CAGAAGCGA GGGGCCTGGT GCGAGGGGG</td>
<td>1620</td>
</tr>
<tr>
<td>ATGCCGCCGA GAATGCGGAC CCGCTGGCAG GGGCCGACC CGCCGAGAA CGCGGGG</td>
<td>1680</td>
</tr>
<tr>
<td>GCAGCCGGCG CGCTGCGAGA CATCGCAAG ACGTCCATGG AGTACCGAAG CATTGCGGAA</td>
<td>1740</td>
</tr>
<tr>
<td>AGAGGCTGAG ACGAGAAGG TCTGCGGAG GCGAGGACG GCAAGAAGA CAGAAGGCG</td>
<td>1800</td>
</tr>
<tr>
<td>TACGAGAGC TGCTGCAGG CCGGAGGAC CGCGGCCG CAGCGGAGTT</td>
<td>1860</td>
</tr>
<tr>
<td>ATCGGCGGCG TTCCGCGCG CCGAGCGAG CGAGAGGTTG TTGGCTGATAC</td>
<td>1920</td>
</tr>
<tr>
<td>CTGGCGGCGG GCGAGGTTG CCGGAGGAC CGAGGCGAAC CGCGAGAAG</td>
<td>1980</td>
</tr>
<tr>
<td>TGCGTTAAC GCAGCGCGT GCGAGGGC GCAACCCGAG GCAACGCTCG GCGCCGAGGT</td>
<td>2040</td>
</tr>
<tr>
<td>GCTCGCGGCG TTCCGCGCG GGGGAGGAC CGCGAGGTTG GCGGAGGACG</td>
<td>2100</td>
</tr>
<tr>
<td>ACGAGAGCC CGCGGCCGAG CGAGGCGGTTT TGTTGCGGAC CGCGCCGCTT</td>
<td>2160</td>
</tr>
<tr>
<td>GGGCGCTGCA GCTCGGCGG TCTGCGGAG GGGGGGGCG GCGGGGGCG</td>
<td>2220</td>
</tr>
<tr>
<td>AGCAGCGGAC GCAGCGAGG GGGGCAGCGG GCAAGTCGGG CAGCAGCGG</td>
<td>2280</td>
</tr>
<tr>
<td>CGTGGCCTGA CAGAAGTGGG TCTGCGGAG GGGCGGCGG CACGAGCGG</td>
<td>2340</td>
</tr>
<tr>
<td>CAGAAGCCGA GCGGGGGCCG GGATCGGCTG CAGCCCACCGGAAGGCTGG</td>
<td>2400</td>
</tr>
<tr>
<td>CGCCAGAG CAGACGAGT GCGGCGGAG TGCCGCAAGGG CAGCGGCGG</td>
<td>2460</td>
</tr>
<tr>
<td>GACGTGAGG ACGCTGCGT CAGATGGT GGGCGGCGG AGCGCGAGGG GGTGCTCGT</td>
<td>2520</td>
</tr>
<tr>
<td>AGAGGCGGCG GCAGCGGACG GGGGAGGCCC CGCGCGCGTT CCGGCGTACG</td>
<td>2580</td>
</tr>
<tr>
<td>TTCCGCTGGA CCCACGAGG GACATCATCA TTCCGCGAGA CAGCGGCGG CAGCGGCGG</td>
<td>2640</td>
</tr>
<tr>
<td>TGAGATGGTGG ACCCGAGGTTT ATCCATGCTT GCTGGGCTCGT ACGCGGGG</td>
<td>2700</td>
</tr>
<tr>
<td>GGTGCCCTCC CCGGCGGG TCTGCGGAC CGGGGCTGGT CGCTCGTCCCG</td>
<td>2760</td>
</tr>
<tr>
<td>TACGATATG ACAGAGTTGC ATCGTCTTGG TGTATCTTGGT TCTGGGGGT</td>
<td>2820</td>
</tr>
<tr>
<td>GGGATGGGCG CAGGCGGCGGCC GGGGGGAT TGGGGGAGT ATGGGCGAGT GGGGGGGT</td>
<td>2880</td>
</tr>
<tr>
<td>GCGTGCGGCTG CACAGCGATGG GGGAGGGCG CCGTTCGTCG TCGGCGAGG</td>
<td>2940</td>
</tr>
<tr>
<td>AGAGGCGGAC CACCGATTTCC TCTGCGGAC CGCCGATCTG GGGCGGGGT</td>
<td>3000</td>
</tr>
<tr>
<td>CGCGCGGCTCC TCTGCGGAC CGCGGCTG CGCAGGGCGG CGCTCGTCCGC CGCCGCG</td>
<td>3060</td>
</tr>
<tr>
<td>AAGAGAGG GCGGGGGT CCGGCGGAC CCGGCGGCGG CCGGCGGGGG ACGCGCGGCGG</td>
<td>3120</td>
</tr>
<tr>
<td>AGATGCGGCA GAGGGGAG GAGGGAGG GAGGGGAGG GGAGGGGC</td>
<td>3180</td>
</tr>
<tr>
<td>CAGAGGGGCA TGGCGGCGG GAGGGGGG GAGGGGAGG GAGGGGAGG</td>
<td>3240</td>
</tr>
<tr>
<td>CTCGCCTACA TCTGCGGCTG TCTGCGGCTG GGGGCGGCGG GGGGGGCGG</td>
<td>3300</td>
</tr>
</tbody>
</table>
---continued

ggcgagcgt atccgctcag ttcaagggcg taatacgggt atccacagaa tcaggggata 3360
gccgagcga ggacagtgcg gcaaaaggccc gcacggacag cagggcgtg aaaaaagcgg 3420
gtgctgtggct gttcttcgcag aaccgtgagt gacctggagc ccccgtacgca accctggagt 3480
cagacgag cggcgaacac cgcagacagaca taataagata caacgggttt ccccttgaga 3540
ggtagtggct gggatatcct gttcaaccc tggcctgtaa ccgatacttc tcggcttttc 3600
tccctctggg aagcctgggct cttttcgtta gttgcacgctc tgggtatcgt agttcgtttg 3660
gggtggctc cttcagcttg ggtcgtgctc agcagacttc cagaggtgctc cggccagtcc 3720
cctttctggc ttaatcctgt cttgacgtca cccgcttgaag aaccgcctta tggccacgtg 3780
cggcgcgcc tggctgagc tggcctgcag aaccgttgaag aaccgcctta tggccacgtg 3840
tggcgcgcc tggctgagc tggcctgcag aaccgttgaag aaccgcctta tggccacgtg 3900
tggcgcgcc tggctgagc tggcctgcag aaccgttgaag aaccgcctta tggccacgtg 3960
ttgctgagc tggctgagc tggcctgcag aaccgttgaag aaccgcctta tggccacgtg 4020
aagcagctcc tttgacgtct ttccgcggtg cttgagcgtca cggggtgag aaccgcctta 4080
gtggtccgtt gtcagagcgtta tggccacgta cgaggtgctc cggccagtcc 4140
agtgggtttt tggattcagc tggctgagc tggcctgcag aaccgttgaag aaccgcctta 4200
gcaggtcagc tggctgagc tggcctgcag aaccgttgaag aaccgcctta tggccacgtg 4260
ggcgagcgt atccgctcag ttcaagggcg taatacgggt atccacagaa tcaggggata 4320
ggcgagcgt atccgctcag ttcaagggcg taatacgggt atccacagaa tcaggggata 4380
ttggttgtg ggcagcgtgtg tggcttggtt ccggctggttg tggccacgtg aaccgcctta 4440
gtgctgtggct gttcttcgcag aaccgtgagt gacctggagc ccccgtacgca accctggagt 4500
gtgctgtggct gttcttcgcag aaccgtgagt gacctggagc ccccgtacgca accctggagt 4560
ttggttgtg ggcagcgtgtg tggcttggtt ccggctggttg tggccacgtg aaccgcctta 4620
aaccagatta ctttgaacag cggggtgag aaccgcctta tggccacgtg aaccgcctta 4680
gttgggggg gggtgctgctgt ggcgctgcac cttgagcgtca cgggcttggtg tggccacgtg 4740
aaccagatta ctttgaacag cggggtgag aaccgcctta tggccacgtg aaccgcctta 4800
gttgggggg gggtgctgctgt ggcgctgcac cttgagcgtca cgggcttggtg tggccacgtg 4860
cggcgcgcc tggctgagc tggcctgcag aaccgttgaag aaccgcctta tggccacgtg 4920
ttgctgtggct gttcttcgcag aaccgtgagt gacctggagc ccccgtacgca accctggagt 4980
cggcgcgcc tggctgagc tggcctgcag aaccgttgaag aaccgcctta tggccacgtg 5040
aaccagatta ctttgaacag cggggtgag aaccgcctta tggccacgtg aaccgcctta 5100
gttgggggg gggtgctgctgt ggcgctgcac cttgagcgtca cgggcttggtg tggccacgtg 5160
gttgggggg gggtgctgctgt ggcgctgcac cttgagcgtca cgggcttggtg tggccacgtg 5220
aaccagatta ctttgaacag cggggtgag aaccgcctta tggccacgtg aaccgcctta 5280
gttgggggg gggtgctgctgt ggcgctgcac cttgagcgtca cgggcttggtg tggccacgtg 5340
cggcgcgcc tggctgagc tggcctgcag aaccgttgaag aaccgcctta tggccacgtg 5400
aaccagatta ctttgaacag cggggtgag aaccgcctta tggccacgtg aaccgcctta 5460
ttgctgtggct gttcttcgcag aaccgtgagt gacctggagc ccccgtacgca accctggagt 5520
gttgggggg gggtgctgctgt ggcgctgcac cttgagcgtca cgggcttggtg tggccacgtg 5580
---continued---

ggttggcgc acatculosc gaaaugtgc acctgacgct taagcacaac ttaaatcat 5640

gacttaacc tataanaata ggcgtatccg gggccctttt cgtc 5684

<210> SEQ ID NO: 103
<211> LENGTH: 4473
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: VR4772, M2 Insert Replacing WHV Insert from VR6430

<400> SEQUENCE: 103

tccggttttt cgctgtcagc ggtggaaaaaa tctgacacat gcagttcccg gaagctgcca 60
cagtgtctc gtaagcgct gcgggagga gaaaccgcgc tcagggggcg tcaggggtg 120
ttcgcgggtg tcggcgcttg tctaaacctg ggctcacagc gcagatttga ctgaggtgctc 180

acccatgtgg gcggcgcagc cgcgcagcac gcgtgaggg gcatacttctc atccttttgg 240
ccttcgccg ctctctctgc gcgctcccg gcgtgcgcgc gcgtcttctt gcacagtatc 300

acgtacaaca aaggaagggc ttgaccaaca tctcagacag atacgtttta ggctggaggg 360
ttttcggctg ctcgtgcag tgacggcag atacgacgt atggagaggt gctaggggtg 420
gttgggcga aagcaagcct tcgggttcta gcgtgcttgt agcccaatc agatgctatg 480
ggttcggttt tgcagagggg ggccagggag tagtttttct catccatttc gttctttgctc 540

acccacggta cgcgttgagg gcacacagcc ttcgcaggg ggaagaagcg ccagcagtgc 600
cgcgtggtgc aagcaagttc gcagatcagt gcctttattgg gaagctacac gcggctatgc 660

acccacgggc gtcacacacc tgcacacctg acaagcagg ccagggcctg cagcagcccag 720
cctccagaca taaaacacca tgtgcacatt gcgacacag tgcgttctgg 780
cacctccct tccagccgag ggctgctcgg ccgaccctgg cagcagctc gcgtctatgc 840

tggcggccgt ccacagcactg gttggccctg tcagcaagtc ttcagcattg 900

aagaacaggg tgcagacacc gcgctcggct gcgtgtctgc ggtggctcgt ctgacagtcag 960
cggcgttccgc acggctgctt tgcacacctg cttctgtat gcggcggtgctg 1020
tgcgagctg gcagaagcccc aagcagaccc gacagcagct gcggaggaag 1080
tacccagctc ttcggctttctc gcgtcttac gcgtgtgcag 1140
gcagcacagc gcgtgtgcag gcggctgctg aacccacccc caaagcagcg gggtgcgga 1200
tgcagcagc gcgggagcgc cctgggtggt gcggcggcag caagcagcct ctcggtcctc 1260

ttcgcgggtg cttcggtgct ctttgcgctgc cttttgttttgcagcgcgccgct 1320
tgcgagggc gcggggagcc gcgggagcc gcgggaggg agcgagggaa gcagaacacc 1380

gcgacagac gcgtgctgag gcgtgctgcag gcgtgctgcag gcgtgctgcag gcgtgctgcag 1440
tgcagcagct cttcggtgct cttttgttttgcagcgcgccgct 1500

cggcgggtct ccacagcagc gcaagctgag ctctttctctct ctttaaatttgcg 1560
gcagacagc gcgtgctgag gcgtgctgcag gcgtgctgcag gcgtgctgcag gcgtgctgcag 1620
gcagacagc gcgggtggtc gcgcagcgc gcgggagcc gcgggaggg agcgagggaa gcagaacacc 1680

tgcagcagct cttcggtgct cttttgttttgcagcgcgccgct 1740

---continued---
aacccacgtct ctcggtgct ctcagccggc gcggcgggag caaacagcgac 1800
<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Base Pair</th>
<th>Base Pair</th>
</tr>
</thead>
<tbody>
<tr>
<td>csagpsaeac</td>
<td>tccapoge</td>
<td>ccagppcoco cccopccocc scacccgcgo aggcccctga</td>
</tr>
<tr>
<td>aggcttccttc</td>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>1920</td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>1980</td>
<td></td>
</tr>
<tr>
<td>csagpsaeac</td>
<td>tccapoge</td>
<td>ccagppcoco cccopccocc scacccgcgo aggcccctga</td>
</tr>
<tr>
<td>sagctgqsg</td>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>2220</td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>2280</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>2340</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>2400</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>2460</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>2520</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>2580</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>2640</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>2700</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>2760</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>2820</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>2880</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>2940</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3000</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3060</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3120</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3180</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3240</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3300</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3360</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3420</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3480</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3540</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3600</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3660</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3720</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3780</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3840</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3900</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>3960</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>4020</td>
<td></td>
</tr>
<tr>
<td>cccopccocc scacccgcgo aggcccctga</td>
<td>4080</td>
<td></td>
</tr>
</tbody>
</table>
-continued

gagccatggt atacccatg ataacggt ccaaacgtg gttggtgc 4140
agccctttc cctgtgatgt csgcctccc tggctgtgaa ctttctggg 4200
actgttatt cgtctctgac gttacattg caagctgtg ccttgtgtctt 4260
gacacccac gccgtctcgc ccccccggg ctcgccccg gcgtgtggag 4320
tgggtcctct ccttctggg ggttagtgg tgggtggag ccttcttcg 4380
ccttcttcg acctttttct actgggtag ctgccagct cgaccagttg 4440
atgactactag cgggctcct gtgccttctc cccctcccc tggcccccct 4473

<210> SEQ ID NO: 104
<211> LENGTH: 8450
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<222> FEATURE:
<223> OTHER INFORMATION: VR4773, Ligation of RSV NPH into VR4756

<400> SEQUENCE: 104

tgcccactgg atacgtgggt tccatcatat atacagca ctttatattg ctcagctgtt 60
acataacgct cggctggcca tcgaggctg gatcgttatt atttctttag aaataagggg 120
tctaggctg atagcggcata tcggggtctg cggcgttact atctggtggtt acgtgctgg 180
cgtcgctct cggctgcgct cggtagggc ccggctgctg gcgtggcttac 240
gtggcggcct cggtgctgtg atagcttcg atgggtttag gcgttgagtt 300
cactctggcc tcatagcaggt catgttctttg ctcgctggcct acggcacggt 360
gtggtagcct cggcttggtgt atgcgttttg ctcgctggcct acgggctgcct 420
cgtctctct cgtcgctctc gtctgctgtg cggcgtgctc gatctgctgg 480
acatgcgtgt gaagctgtggt gcgtgctgct cggctgctg cggcggtgtgg 540
gtcgctgtg cggctgctg cggctgctg cggctgctg cggctgctg 600
gccctgggg caagcggtgg caagcggtgg caagcggtgg caagcggtgg 660
gttgtgtggt gatcgttttg ctcgctggcct acgggctgcct 720
agctgctgtg gtaggcctg gcgtgctgct cggctgctg cggctgctg 780
gcgctggg gtcgagcctgc gtcgctgct gtaggcctg gcgtgctgct 840
atgcgttttg ctcgctgct gtaggcctg gcgtgctgct cggctgctg 900
tgagcgtggt gtaggcctg gcgtgctgct cggctgctg cggctgctg 960
gtgcgtgctg ctggtagtcg ctcgctgct cggctgctg cggctgctg 1020
gagcgttctc cggctgctg ctcgctgct cggctgctg ctcgctgct 1080
gtgcggctgtg ggcttttaca ctcgctgct ctcgctgct ctcgctgct 1140
gttgtgtggt gatcgttttg ctcgctgct ctcgctgct ctcgctgct 1200
agctgctgtg gtaggcctg gcgtgctgct cggctgctg cggctgctg 1260
gctgagcgtggt gtaggcctg gcgtgctgct ctcgctgct ctcgctgct 1320
cagctgctgct ctcgctgct ctcgctgct ctcgctgct ctcgctgct 1380
atgcgttttg ctcgctgct ctcgctgct ctcgctgct ctcgctgct 1440
agagagcgtggt gtaggcctg gcgtgctgct ctcgctgct ctcgctgct 1500
cggcgtgctg ctggtagtcg ctcgctgct ctcgctgct ctcgctgct 1560
-continued

ccttggctg caacagcgc gactgattc ttcagccctt ctgctggca 1620
gaacggccc tctcacttc acagaggag caggcggca gcctggtcg 1680
tgtcagcag attgctgac ggcagctg cgcggctg gggcagctg 1740
tgctggac ggcagctg cgcggctg gggcagctg 1800
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 1860
ggctgctg caacagcgc gactgattc ttcagccctt ctgctggca 1920
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 1980
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 2040
catacctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 2100
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 2160
ttcagccctt ctgctggca 2220
gcccggcc ggcagctg cgcggctg gggcagctg 2280
ggcagctg cgcggctg gggcagctg 2340
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 2400
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 2460
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 2520
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 2580
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 2640
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 2700
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 2760
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 2820
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 2880
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 2940
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3000
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3060
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3120
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3180
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3240
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3300
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3360
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3420
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3480
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3540
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3600
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3660
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3720
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3780
aatcctgca cctcagctg caacagcgc gactgattc ttcagccctt ctgctggca 3840
-continued

atggagatt tggagagt gguaccagtt ggtatattaa aacatttatt tggcaacgaa 3900
cagctgctg tgtgaggaag atcaaggtta tggctcattc aacattcacc gattggattt 3960
atccaaaca gggcagcagc tgtcagcagc ggtcggagt gtctcaggtt taaaccaaat 4020
taaccagctc tgcattcagac aacctcaatt cccgcatagc ttatctcatt 4080
caggtacaa ataacatatt tttgaaaaag gcaagttttc taatgaggg gaaacactcac 4140
cgacgaggtt cctactagct gcagataact ggtcaagcag tgcagcttag actcagcagc 4200
cctcataca acacatttt ttttcctgttt caaatataag gtgtcatact gccacatcac 4260
cagtagagcc gatggtagct gcctgagaaa gccacagcgt tccaaatatc tccagctttc 4320
gtccacagc cccgatattc ggtctaccaag ccaatatcc cgcacacacg acacgatattc 4380
ttcctcctgg tgtgctcagcg ggcagcagc atacggcagc gcttgatataa gcacacattc 4440
aacgagcag cggaggacag cggagcagc acactgcagc gcacacacac caatatttcacc 4500
cgtttctag cttcttctct ttaacccgag atgctgttct gccggtggtg gcagtgtgag 4560
gtacccgac atcttccagc gtaagattat gtattgctg atggtcggag ggtatcacttc 4620
cgctcagcag tttgatggtg gcaatactag cttgacatg atgccagcct gcaccttggc 4680
caggtttcag aaacactcag gcgtcagcag ggtcactcag ctaactcagc taactccagg 4740
cgtttttgag gacatctatag cgggagctc tatactcagc taactccttgg 4800
aatatatag cgccctttag cggacagttt cccaggtat ggcgttgttg tttattcttc 4860
tatgtagctg tattgactgc gacagattta gtagagtta tattagattt tttcttattg 4920
caggtattgatc tggcaaatcc tggagctttc gcctataatt aacagctcag gcttggat 4980
ggacacagc cccgcccagc tggagctatt ggtgctttcc gttggtgtct tgcgcgcgcg 5040
tatactcag ggtcgagacgag ggttacagcg gtttacagcg cttatgtcctt gctctgccgc 5100
agcaacgatct gttgtcagttc gttttctatt gttttctatt gttttctatt 5160
ggccttggag gggccgctg ggtatgcgg gcagggaggc gcgtcagcag ggtgtccttg 5220
ttcctcctg cccgacattc atggtttctt ttgctttttt cggaggtgctg atagtaattc 5280	tagagcatact tgtgtagctt ttcaccctgt tctgcatagc ggtgacagct gcacacagtat 5340
gggagactc cccagcggtc gattcagcag ggtggttttt gcgtctgtgta tcagctggtt 5400
tgctgatgc gcagacagcc cccagatgtc ttcagacgag cccagtgtttc cccagctggt 5460
gagatagctc aataatagct gtcagctatc acactaaac cccagttacg gtcacacactc 5520	tggtgtgcg cccacccctgc cttcagactg cccgccagcc ctacacctg 5580
cggccatgct gcgcgttctgc gcgcgcgcgc gcgcgccgct gcgcgcgcgc gcgcgcgcgcgc 5640
tggctccagc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgcgc 5700
cctgctgagc cccacccctgc cttcagactg cccgccagcc gcgcgcgcgc gcgcgcgcgcgc 5760
aacatatttc gctgcgtttc gccgccgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgcgc 5820
caccagctc ggcgttgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgcgc 5880
ttcctcctgc gccgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgcgc 5940
ggcgagcgc ggcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgcgc 6000
aatagcatgag cgcgttgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgcgc 6060
gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgcgc 6120
ttcaacagaa gacagacag atgacggag gacacccaaa gcgcggcctaa ggcaccgcaat 6180
eaagccgccg gcccacccc gcaagagctg gccggaagtt gatcgggaga gctggtgtttg 6240
tacagcaagg aggaaagcag agaatcctgg aagccggcct cccagccgga ggcggaacct 6300
ggcggcttcg ccggggatgt cacgtggcct gctgacccgc tggagccagggct 6360
acacagccgc tgggctggaac ggccacagaa ccacagactg gcagccgctg gacaggcagc 6420
acccggcgaac gcaggccgct ccggtgcctg ggcgtgatgg aagctgcttg 6480
atggagaagct ccggagcgc tcggagcgc taagagacag ggcggcagcct 6540
aacccgaga acagacccag cgcctcgtag agaagttgca cacacggctt gggcaagttc 6600
cacggccgct cccagagcag cttgagacag ccggtgcctg ggcgtgatgg aagctgcttg 6660
gcgagacag cggagctgtc cttcctggaac gacggggcct ctcggagagct aggctagttg 6720
ggcaccagac ggtgtcctgc cgcctgcttg taagccggag cgggtgagcg cagacctagc 6780
ttcagagagct cggctgctag ctgaccccct ctcggagctg gcaagacagc 6840
cagctgtcag gcagacccag cccacggcct acacccagcag gcagccgctg gacaggcagc 6900
atggagaagct ccggagcgc tcggagcgc taagagacag ggcggcagcct 6960
aacgctccg ccaagagcag ctcggagcgc taagagacag ggcggcagcct 7020
aacccgaga acagacccag cgcctcgtag agaagttgca cacacggctt gggcaagttc 7080
cacggccgct cccagagcag cttgagacag ccggtgcctg ggcgtgatgg aagctgcttg 7140
acccggcgaac gcaggccgct ccggtgcctg ggcgtgatgg aagctgcttg 7200
ggcaccagac ggtgtcctgc cgcctgcttg taagccggag cgggtgagcg cagacctagc 7260
ggcaccagac ggtgtcctgc cgcctgcttg taagccggag cgggtgagcg cagacctagc 7320
ggcaccagac ggtgtcctgc cgcctgcttg taagccggag cgggtgagcg cagacctagc 7380
ggcaccagac ggtgtcctgc cgcctgcttg taagccggag cgggtgagcg cagacctagc 7440
cctctggttg gcacgccct gcggctgctg gcggctgctg gcggctgctg gcggctgctg 7500
ccctctggttg gcacgccct gcggctgctg gcggctgctg gcggctgctg gcggctgctg 7560
gctctggttg gcacgccct gcggctgctg gcggctgctg gcggctgctg gcggctgctg 7620
ccctctggttg gcacgccct gcggctgctg gcggctgctg gcggctgctg gcggctgctg 7680
tgacccctct ccctctggttg gcacccctct gcggctgctg gcggctgctg gcggctgctg 7740
tgacccctct ccctctggttg gcacccctct gcggctgctg gcggctgctg gcggctgctg 7800
tgacccctct ccctctggttg gcacccctct gcggctgctg gcggctgctg gcggctgctg 7860
tgacccctct ccctctggttg gcacccctct gcggctgctg gcggctgctg gcggctgctg 7920
tgacccctct ccctctggttg gcacccctct gcggctgctg gcggctgctg gcggctgctg 7980
ccctctggttg gcacccctct gcggctgctg gcggctgctg gcggctgctg gcggctgctg 8040
ccctctggttg gcacccctct gcggctgctg gcggctgctg gcggctgctg gcggctgctg 8100
ccctctggttg gcacccctct gcggctgctg gcggctgctg gcggctgctg gcggctgctg 8160
ccctctggttg gcacccctct gcggctgctg gcggctgctg gcggctgctg gcggctgctg 8220
ccctctggttg gcacccctct gcggctgctg gcggctgctg gcggctgctg gcggctgctg 8280
ccctctggttg gcacccctct gcggctgctg gcggctgctg gcggctgctg gcggctgctg 8340
ccctctggttg gcacccctct gcggctgctg gcggctgctg gcggctgctg gcggctgctg 8400
catcaataca acctattatat ttccctctgt caaataataa gttatacagt gagaactca
4260
catcggtgac gactggtacc ggtgaagatg gaaasegctt atctaatctct ttcagactt
4320
gtcaacagg cagcggacct gctgcgtcact caaatataacct ggtactcacc aacccggtat
4380
tccctcggta tggctgctgaa ggcaagcaaa atacgcgtac gcgttataaa ggaacattac
4440
aacacagga caagatggaac cggcagctaga aacctgcaag cgtatacaac atatttttcaac
4500
cgtaacctcag attacactttt aatacctagga atgtgttatg ggtccggaga gcgtataaat
4560
cctggacaac gattcgctcg acacctctcat ctgtaacact atggggacag ctacctttgac
4620
catctgttta aacaaacottc gcgcgtaggg gttcccccata cagtaagag atgtcccac
4680
cgatgtgccc gacatactcg cggcggcact tataccctt tattacccaga tctatgtgac
4740
tataatactg cggcttgctg caaagcctt ctcccattga atgtgtctga acaccaacat
4800
tatatacttt caagttgtttata ttcgccttac ctgatatactt cttcactgtg
4860
cgaatggaaca tccagctgtt cccgcagtgg gccgagcag tcgcgtaacg cgggtgtcgtc
4920
caccgtcactctggatttttc ttagaactgt tcggcactggat aatctataagc
4980
tttttccct tactccctca cctgtcggag gcattttctt toocctgca cttatacgc
5040
tattctttt caactctttg cccctctttg cggctggttt ggtgtagttc gcgtatagag
5100
gggagggag acggctcaac tattctcctt gggtggtagt cggcaggtgg acctcctag
5160
cgatcgtgtgc ogcggcgcacct tagacccagc ggggtgtggtc gcgggttgtc
5220
acatcagaggg gcgtgtcctt gcctctcttt ggcocagggag cagocgggct caaactcag
5280
cacaccggtta cccatagggc cccgccccgc cccctcctag ctgtatgctt cttccctatac
5340
ccttccctcc ctttccccct ccctgcttcatt gatgcagcac cacttcgacc
5400
tgcagtggca atctctctct cttcatagaa aagggcaggt gcggtggtcc ctctcctcgt
5460
caaaaggac acgggggagg gggcacaacc agatgcgttg ccaagcgaga gcgcaagaga
5520	ctgcacctgc aaggtgtcct gcactgcttt gcgtagtgct gcggcagaa
5580
gatgctgccc tgcctctcag tggccagcag ccagccagtt cgcgggtgtg cgcctctgac
5640
gctccctcg cccctctgta gcagcagctc agggatcctt ggcgcctcgc ggccctctct
5700
cattccagc atctggccct tccattgctg ggtctggctc ccgagttggt ccggcagggc
5760
gccatcagc gttgctttgt ggaagggag gttctctcttg aagtcgaagc agggagtaac
5820
gtggagtggcg cccgtgcttg ccggggttgt tgggtgggtc cccgggtcct tcggcctgcag
5880
ggcgcctgcgt cttccctctc gcgtcaggt ggccgccagt cttcttgttc tttctctgtg
5940
gctgcacgt tgcgcgagct tgcgtcctct gttgcctctgt ggccagacttt ggggccct
6000
ggctgaagttgc atctgtgctt cgcggcgggt cggcggtgt tcggccctgc ggcaccgcga
6060
cagctgcgttc tggggtggtcc tcctgttggc tctgtactgg cttctagtgc gcgtgttctc
6120
cagcgtgtgt acgggcgtcc cggccggttc cttgtgcttc ggccagtgcg gcgtgttctc
6180
gtgcgctgg gcggcgtgctt cccgcgtggc ctgctgcgttc ctcgtttgtc gcggggggg
6240
ttcctggctg aggcgccgcgt ttcgccgctg cggcacccct gcgtgctgttg ggcctgtgct
6300
tgcgggtgct ctcgggtctt gcgtctccgc gcgttgcttg gggcggttgg gcgtcctcg
6360
ttcagggcc tgcgctccttg ggtctggttg acggcctttc ttcggcctt ttcgggtctt
6420
cggcagttt ccgctgttg ccgtgctttt gcgtatcgtt ctcagctgcc ccctgttct
6480
-continued

gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 6540
cattctcct acacatctgg ccgctgctcc gcgctgctcc gcgcgttccc cttggaggg ggctgctcc 6600
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 6660
cattctcct acacatctgg ccgctgctcc gcgctgctcc gcgcgttccc cttggaggg ggctgctcc 6720
cattctcct acacatctgg ccgctgctcc gcgctgctcc gcgcgttccc cttggaggg ggctgctcc 6780
cattctcct acacatctgg ccgctgctcc gcgctgctcc gcgcgttccc cttggaggg ggctgctcc 6840
cattctcct acacatctgg ccgctgctcc gcgctgctcc gcgcgttccc cttggaggg ggctgctcc 6900
cattctcct acacatctgg ccgctgctcc gcgctgctcc gcgcgttccc cttggaggg ggctgctcc 6960
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7020
cattctcct acacatctgg ccgctgctcc gcgctgctcc gcgcgttccc cttggaggg ggctgctcc 7080
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7140
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7200
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7260
cattctcct acacatctgg ccgctgctcc gcgctgctcc gcgcgttccc cttggaggg ggctgctcc 7320
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7380
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7440
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7500
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7560
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7620
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7680
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7740
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7800
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7860
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7920
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 7980
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 8040
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 8100
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 8160
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 8220
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 8280
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 8340
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 8400
gcgcagcc ttcaggggg ggcgcgggccc gcgcgttccc cttggaggg ggctggccc ttcagccc 8450

<210> SEQ ID NO 106
<211> LENGTH: 8442
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: VR4775, Ligation of RSV BSeg7 into VR4762
<400> SEQUENCE: 106
-continued

tgpcatgct atgcgtgta tctatctat atatgacata tttatatagc ctcactcatca 60
actactgcc cagcttgcc tggactatag acaatgatat cattgaactc ataatcggg 120
tctagcttc atcgccctac tctgagaacct cagctactc aatactatcc 180
cgtccgcgac ggccaagcag cccggccccc tggagctgcc taaagcagta tggcccaata 240
gtggaacgta cggcaatcct cagcagctcc caagcggggg aagatatcagtgtaaatagc 300
cacctttttg ctgtaactat ccaatctttg ccacatgaag cccgctatag cgcaaatagc 360
gtggaatc cagctctgcc tattgcccag tcaagccct cagatcctt gctctcttg 420
cgcctatcct cagctgtctt caacatcttt ccacactgtg aagccaatcagcc 480
aatccgctag ccagcgtggc tgcagccttg gccctctcct gcctctcctc ctctctcctc 540
aatccgctag ccagcgtggc ccaaccttc cgcagctttgc ccacatgctg ccacatgctg 600
gcagcctag ccagcgtggc cggcagctttgc ccacatgctg ccacatgctg 660
ggtatactg acgcctccag cgcctccttc gcaccgcatg cgcctctcct gcctctcctc 720
agacacgggg acggctccag cctccccgct gccacgcttc gcgcctcctgc 780
gggtgcaatt cgttctcctt cggatctat ggcacccgag ctcctcctcct 840
atctccagat ctcgactatat cgccttotg ggcacccgag ctcctcctcct 900
tcgcgtatg accactctg ccaccgctct gcaccgcatg cgcctctcct gcctctcctc 960
ggcctcatt cggcctcct gcaccgcatg cgcctctcct gcctctcctc 1020
ggtatatgt ccaatctctt gttcctcttc gcaccgcaag ctcctcctcct 1080
ggggctcct cttctttttt cttactttc ctaacctcat cgcctcctcct gcctctcctc 1140
atttttttta aacactcatt cgggctctct gggggctgct gccctcctcct 1200
gggcttccttc ctgagtctct gcggcctcct cgggcgtctgc gcaccgcatg 1260
gtctctctct cggcctcct cgggcgtctgc gcaccgcatg cgcctctgctg 1320
caatcacccag caaactgcag cttgcctcct gctgggttct gcgtggtttct gcgtggtttctgc 1380
agggctggag ccgtctgctg gcaccgcatt gagacactag gcaccgcattt 1440
agggctggag ccgtctgctg gcaccgcatt gagacactag gcaccgcattt 1500
cgcctctgctg cgcctcctcct cgcctcctcct cgcctcctcct cgcctcctcct 1560
cacaactcatt cggcctcct gcaccgctct gcaccgcatg cgcctctgctg 1620
gccccacatt cccgactat cttcgcttct gcaccgctct gcaccgcatg 1680
cgaccgcaag ccggccctct gcaccgctct gcaccgcatg cgcctctgctg 1740
cgaccgcaag ccggccctct gcaccgctct gcaccgcatg cgcctctgctg 1800
cgaccgcaag ccggccctct gcaccgctct gcaccgcatg cgcctctgctg 1860
cgcctcctcct cgcctcctcct cgcctcctcct cgcctcctcct cgcctcctcct 1920
gagccgcagag ccgggcctct gcaccgctct gcaccgcatg cgcctctgctg 1980
gcgccctctgc gcaccgctct gcaccgcatg cgcctctgctg 2040
gcgccctctgc gcaccgctct gcaccgcatg cgcctctgctg 2100
gcgccctctgc gcaccgctct gcaccgcatg cgcctctgctg 2160
gcgccctctgc gcaccgctct gcaccgcatg cgcctctgctg 2220
gcgccctctgc gcaccgctct gcaccgcatg cgcctctgctg 2280
---continued---

gacgcgcaga aagcaccaga gcggctcagag ggagaagg gcaccctcagc agggccagtt 2340
cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc 2400
cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc cgcgcgcgc 2460
gccgcggcgc gcggcggcgc gcggcggcgc gcggcggcgc gcggcggcgc gcggcggcgc gcggcggcgc 2520
cggcgcgcgc gcggcggcgc gcggcggcgc gcggcggcgc gcggcggcgc gcggcggcgc gcggcggcgc 2580
cgcgcgcgc gcggcggcgc gcggcggcgc gcggcggcgc gcggcggcgc gcggcggcgc gcggcggcgc 2640
ggcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 2700
cgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 2760
cgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 2820
cgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 2880
cgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 2940
cgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3000
cgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3060
cgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3120
cgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3180
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3240
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3300
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3360
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3420
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3480
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3540
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3600
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3660
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3720
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3780
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3840
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3900
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 3960
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 4020
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 4080
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 4140
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 4200
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 4260
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 4320
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 4380
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 4440
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 4500
tgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 4560
-continued

atcgagcatc aatgaaact gcattttatt catacagga titaacatac catatattttg 4620
aataaagctt ttcggctasg aagagaaaaa cttaacgagg cagttccata ggtatggcag 4680
actcctgtat ctctgcgttct tcacgctctt acacaacct cttatctttct 4740
cctttcataa ataagttttac caagtgagaa atcaaacagt gttggcctgg aatccggtta 4800
gtggccttaa aggtttatcc gtcttcctttaa gccttggctata acagcgcccag ctttcgttttc 4860
gtctttctaaacct ctcccccttc gttccccctt cgtttgcttc tgtttgctgcct ttggccctag 4920
aagaaatctcg cgttggctgtct taaagggctaa atcacaacg cggcgctttaa ggacaagcggc 4980
cagaggacct ggcacggtcat caaatactttc ttctcctgctta tcatgctaat tttttcctttac 5040
cggatgtcgt gtttttcggct ggtctgcgtt catgtggctt cgggatgtgctgc cttttccttttc 5100
gttctaaatgg ttcgaggctctt ggcagggctatct tggagtgcctta ttttggactgtat 5160
cctccttctgg aactaatctg gcagcctcttt gtgtgccgttc ttctggactgtat ccactcttcttt 5220
atcgagcttc cccctctgctgtg ttaaggtctt gctccctctgcta tggctgctgctt cgtccctctttt 5280
ccctttggtg cccctttgtt cccctttggtg ctttttctttt cccctttgtt cccctttggtg 5340
cgtttttgat ctttttctttt ctttttctttt ctttttctttt ttttttctttt ttttttctttt 5400
cgggtggttt cgggtggttt cgggtggttt cgggtggttt cgggtggttt cgggtggttt 5460
cctcctttgctt ctctctctttt ccctctctcttt ctctctctcttt ctctctctcttt ctctctctcttt 5520
tttttctttttt tttttctttttt tttttctttttt tttttctttttt tttttctttttt tttttctttttt 5580
gtctctctctttt ctctctctctttt ctctctctctttt ctctctctctttt ctctctctctttt ctctctctctttt 5640
tttttctttttt tttttctttttt tttttctttttt tttttctttttt tttttctttttt tttttctttttt 5700
tttttctttttt tttttctttttt tttttctttttt tttttctttttt tttttctttttt tttttctttttt 5760

5820
5880
5940
6000
6060
6120
6180
6240
6300
6360
6420
6480
6540
6600
6660
6720
6780
6840
ccactgaagt ggcottttggc ctggtatagt caacatgtga acagatgtcg gactccacg
cacgttcga tagaactagt tggcaccacg cccattcatt tataagccat ggcagcga
ctgtttttcg cgaatcataca tgaagcaat tcgaagggc atgtgcggag
ccgggatcc ctagaacaaat gctgaactgg ccaccagcaat gcgggccgca attagagcag
ttcgggcctc acaagagtac tggacagcgct gcgggatgat ccgcttcatg
ccgggcattgg cggagcagc ccttctcttc tggcaccacg cctgttttctcg
cggcagcagc ccccttctgc tggcaccacg ggtgacagcc
tatacatatc tctctggagc ggtttccttc tcctatttaa gcggcaccacg
gggagcagaat tttgtagagc ccttttaagg ttttatttttt ttctctctttatatgtggttta
ttgtaaaag agccagatag cctttaggaa gactctttta ctgactgacttt tctctctatg
attgggcctg gcgggatgacgt ctctctctttctt cgggctttgc gtcagctatr
ttgacagcagc cggagtattgg gcagctttg ataacaacctag
</210> SEQ ID NO 107
<211> LENSEN: 8442
<212> TPF: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: VR4776, Ligation of Inverted HSE R Seg? into VR4762
<400> SEQUENCE: 107
tggcattgc aagccggca gctttcattg ctaacagcct tggcacttgc aagccggca
acactttgc catcgtcctt gctatagt gcatgaacctt tataagccat ggcagcga
ctgtttttcg cgaatcataca tgaagcaat tcgaagggc atgtgcggag
tccgggcattgg cggagcagc ccttctcttc tggcaccacg cctgttttctcg
cggcagcagc ccccttctgc tggcaccacg ggtgacagcc
tatacatatc tctctggagc ggtttccttc tcctatttaa gcggcaccacg
gggagcagaat tttgtagagc ccttttaagg ttttatttttt ttctctctttatatgtggttta
<210> SEQ ID NO 107
<211> LENSEN: 8442
<212> TPF: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: VR4776, Ligation of Inverted HSE R Seg? into VR4762
<400> SEQUENCE: 107
tggcattgc aagccggca gctttcattg ctaacagcct tggcacttgc aagccggca
acactttgc catcgtcctt gctatagt gcatgaacctt tataagccat ggcagcga
ctgtttttcg cgaatcataca tgaagcaat tcgaagggc atgtgcggag
tccgggcattgg cggagcagc ccttctcttc tggcaccacg cctgttttctcg
cggcagcagc ccccttctgc tggcaccacg ggtgacagcc
tatacatatc tctctggagc ggtttccttc tcctatttaa gcggcaccacg
gggagcagaat tttgtagagc ccttttaagg ttttatttttt ttctctctttatatgtggttta
<210> SEQ ID NO 107
<211> LENSEN: 8442
<212> TPF: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: VR4776, Ligation of Inverted HSE R Seg? into VR4762
<400> SEQUENCE: 107
tggcattgc aagccggca gctttcattg ctaacagcct tggcacttgc aagccggca
acactttgc catcgtcctt gctatagt gcatgaacctt tataagccat ggcagcga
ctgtttttcg cgaatcataca tgaagcaat tcgaagggc atgtgcggag
tccgggcattgg cggagcagc ccttctcttc tggcaccacg cctgttttctcg
cggcagcagc ccccttctgc tggcaccacg ggtgacagcc
tatacatatc tctctggagc ggtttccttc tcctatttaa gcggcaccacg
gggagcagaat tttgtagagc ccttttaagg ttttatttttt ttctctctttatatgtggttta
<210> SEQ ID NO 107
<211> LENSEN: 8442
<212> TPF: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: VR4776, Ligation of Inverted HSE R Seg? into VR4762
<400> SEQUENCE: 107
tggcattgc aagccggca gctttcattg ctaacagcct tggcacttgc aagccggca
acactttgc catcgtcctt gctatagt gcatgaacctt tataagccat ggcagcga
ctgtttttcg cgaatcataca tgaagcaat tcgaagggc atgtgcggag
tccgggcattgg cggagcagc ccttctcttc tggcaccacg cctgttttctcg
cggcagcagc ccccttctgc tggcaccacg ggtgacagcc
tatacatatc tctctggagc ggtttccttc tcctatttaa gcggcaccacg
gggagcagaat tttgtagagc ccttttaagg ttttatttttt ttctctctttatatgtggttta
-continued

cacttggcac taccatcagt gccattatgc ccasggacgc ccctatagga ccgtaaatagc
360
gtaataggg cctctgtgca gctttgcccc tattacctgt tattttcttg tcttactctg
420
cagtacatct agcattatgt cccctctgtt accctgtaa tgcgttttgg gcacactacat
480
aataggggag gagagcctgt gatattgcccc gatcatcacc cctctagccc ctagccgctc
540
aataggggat tttgattgctt ccacataacc cggcacttcc cacaatctgc taaccatctc
600
gcggcctgtg ccagtattgg ccagyagtgg gttcctttgg ggtctctat acgcactgct
660
cggttagtag gcacagcact gcgcctgcag ccgcttttgg gcttttttag gcgctccatg
720
ggagccggc aacctgcact cctctgtgca cgggactgt gcctgaacgc gcgcttcctg
780
gtcgctctag ttcagctcc tccctctgca ctttctttct ctttctttct ctttctttct
840
atgcagcctg aagtttctgc gcggagctgg cccattcttt aagcagcctg ccctctttct
900
tagctttata gcttattttc gccctttttt ttgctctcct gcttattttc gcctctttct
960

ggctgtcgag ctcttcctctttt ctttctttcttct ttgctcttcct gcttattttc gcctc
1020
ggcttttttc caataactctgg ctctttttct ccgctttttt ctttctttct ctttctttct
1080
ggtttttttc ccatttctttt ctttctttct ctttctttct ctttctttct ctttctttct
1140
gttttttttc ccatttctttt ctttctttct ctttctttct ctttctttct ctttctttct
1200
ggcttttttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
1260
ggcttttttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
1320
ccatttctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
1380
atgttctgct ccgctttttt ctttctttct ctttctttct ctttctttct ctttctttct
1440
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
1500
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
1560
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
1620
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
1680
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
1740
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
1800
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
1860
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
1920
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
1980
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
2040
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
2100
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
2160
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
2220
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
2280
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
2340
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
2400
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
2460
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
2520
cccctctttt ctttctttct ctttctttct ctttctttct ctttctttct ctttctttct
2580
ccaggtgtac aagctgatca gacccaaaga gaaaccccc ccacagagcc agctggrtgtg
2640
gagggctgc cagcgcogg cccctgagga cctgagatct ctagagttca tcagagccac
2700
cagggctgcc cccagagctc aagctgaccc cagagccctg tcatgctgca gcaacgagaa
2760
cagccaaacc atggtgcaag gaccccttga gctgagacct gataactggg ccacagacac
2820
cagccagccg gccaacggcc accagcacag acgccagccc gcggagagcc gcggagagcc
2880
cacctctccgg tctcgagaga acctgcccct cggagtcgcg acggtgatgg cccgctcccg
2940
cggaacaccc gggcggcaac ccagccgcat ggggaggag atccatccag tggagaggg
3000
cggaacaccc gggaggtgct cctccagag gcggagcttg ttcaaggtcca gcggagagaa
3060
gggacaacac cccacagctgc ctcgaggtga ctgggacaac gaggctcgcct acttcttctgg
3120
cggacaacgg cagagttgag cagagccac gttggagtc ccctagggct accctctcttcc
3180
tgctgactca aagatcaag cttacagagact gttggtagtg gttttttggtgtggtcgtc
3240
tgctgactca gctcagacgc gcgggtgctgc gcgggtgca gcgggtgca gcgggtgcag
3300
tgcctcaac ggcggtttata cgcttcatca cagcatacag ggtaagcgc ggaagagaca
3360
tgctgactca acgggtgcgg cctttttccc cggataagcg ctgctgctctgctc
3420
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
3480
tgatggtgcttc ctacaggtgc cccctctgcgg ctgctgctcc ctgctgctcc ctgctgctcc
3540
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
3600
tgatggtgcttc ctacaggtgc cccctctgcgg ctgctgctcc ctgctgctcc ctgctgctcc
3660
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
3720
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
3780
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
3840
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
3900
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
3960
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4020
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4080
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4140
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4200
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4260
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4320
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4380
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4440
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4500
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4560
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4620
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4680
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4740
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4800
tcagctgctc gcgggtgctgc cttgagactc ttcctttctct tggagaaggc
4860
-continued

gtcctcasa aatacagcat ccaaccaca gttatctctt cgttattgct cttcagcag g 4920
gacaatacc cagaattcg ggttctgct taaacgcc ttaacagca cctgtgcga cttgagcag 4980
gcgcacat gggccggct cgaaccat cttcagcga ccctggatcc tattctatac 5040
tctgaattct gtttcccggc gcgtcgccat gcttcaggac cccacgacat cccagatcag 5100
gatcataag cggaggtcag gaatcctgtc aagttccttc aatcagctcatt cccctgcag 5160
cctctcgtta aagctttcgga cagcttcacc tttgctactt tattgagaa aacttggcag 5220
acgagggcc ccaatttctt gttatcagc ctaaccacat gtaact gcctgagcc 5280
cagtattct ctcctgtt cccatcgtat tttgctatct cctgacact tttgctagc 5340
caggttttct taaagttggc ccttcagac cttgtatcct gttgtattgta aagcagcag 5400
tttcttgcatt tgtctttttc tttgctatc taccctccag gtttggagga 5460
cacttctacc gatcaacggc ctttaacctt cctgtacctt ccctctagtt 5520
tgctggcagt tttcttccct cttctacatt cgtttacatt cccctctgtt 5580
cctggagct cgggcatctt tgccttgtc cttgagcag ccctgagcc 5640
ttgctgggct ggttttgttc tgggcttttg cttgagcag ccctgagcc 5700
ggaacctgag aacgagggcgt ctcctgtggt tgtccagcc caagagggcg gcctgtttggc 5760
tttgctggc ggttctgagc cttcagccac gtcctgagcc gcctgtttggc 5820
gcgttctccag caaactggtct atggcttccc ctaacctccc ctctgagct ctcctccacc 5880
cctctcctcc gacaacggc cctttctctc cctggctagc ctcctccacc 5940
tgctggttg caggagcctt tagcctccttt gccttttttc cagcctttttt 6000
gacggagcgt tgcctgttc gggaattcag cggatcctgg gcctttttccc gttttttttt 6060
tcctatgcc acacacgaggac cgtcttct gtaacagcc cccatagttt cccttcttt 6120
ttcgctccct atctaggtcc cgggattttc cagccttttt cccctttttt 6180
gcctggtgg ctttcttttt ccaaccat ttcctccttt ccctttttttt 6240
gcctggtgg ctttcttttt ccaaccat ttcctccttt ccctttttttt 6300
gcctggtgg ctttcttttt ccaaccat ttcctccttt ccctttttttt 6360
gcctggtgg ctttcttttt ccaaccat ttcctccttt ccctttttttt 6420
ttcgctggc ctcctccttt ccccttttct ttcctccttt ccctttttttt 6480
ttcgctggc ctcctccttt ccccttttct ttcctccttt ccctttttttt 6540
ttcgctggc ctcctccttt ccccttttct ttcctccttt ccctttttttt 6600
ttcgctggc ctcctccttt ccccttttct ttcctccttt ccctttttttt 6660
ttcgctggc ctcctccttt ccccttttct ttcctccttt ccctttttttt 6720
ttcgctggc ctcctccttt ccccttttct ttcctccttt ccctttttttt 6780
ttcgctggc ctcctccttt ccccttttct ttcctccttt ccctttttttt 6840
ttcgctggc ctcctccttt ccccttttct ttcctccttt ccctttttttt 6900
ttcgctggc ctcctccttt ccccttttct ttcctccttt ccctttttttt 6960
ttcgctggc ctcctccttt ccccttttct ttcctccttt ccctttttttt 7020
ttcgctggc ctcctccttt ccccttttct ttcctccttt ccctttttttt 7080
gcctggtgg ggttctgagc cgggattttc cagccttttt cccctttttt 7140
---continued

gagtaactgct cagactacac tgcctcacc cgyttaactag agtggagaa caggyctcag 7200
gcaacgagt ggcggccgga cgagcttctc agttgctcag ggaggtcgc gaccagggc 7260
cgcttgcag cgctgacgta aacctactta gaagggagac gcagttcagg agggcaccg 7320
ggggagcgc gcaggaacc gactcaacgc gcggtgaggt ggcctcagc tagggcggcg 7380
ggcgctagc ggagagatgc gagccagtag agggcaccac gaccaggtgg gatcgctaaa 7440
tggctttctaatagcag cggctcacttatagctctc tggctcactctgag 7500
ttgctctgctg cctacctgat gacggcgtctgt gcgtcttcct tataagccg cgygctgctc 7560
acctactaccc cacoactgct catgcaagat gctttttccc tcctgctcact gatcgctgt 7620
actacctagct tacacatggtt ccagactacac agatgattgg atagacactct attcctcct 7680
cctcataa aaggaactcc aatctactct ggggggtt ccataacgct taacccagaag 7740
cccggcttgc gcctcactaa cacoactgct cctctagcta cggctgatcc cggctgctc 7800
atcctagacgc gagccagtag aggacctacgc cctctgcact tgcagctgca 7860
gctctgctg ctggtgctgctt aacctttgct gcgcgctcct gtggagccgcat cctcctccac 7920
gccagagc gcacactgct cacccagttgc cggactaggct ctctgtgccgag 7980
attcctcactgcatgggtgt ttcctcccccc cccctctctatat ggaggtgtt ccgctgctc 8040
tttgcttccgagccctcct ctaaactctc tattgagtt aatctacctc tataagctcaag 8100
ggcctagct ccccgagac gcgccctcag cgyctgactc accctctaaa tggcgctgctc 8160
accccttaa aacgctggt cacgcgggcc cttcctgctg ggcgctgtcc ttctagagcc 8220
tgaccttcg cctctgctc ccctctcgaa ggcctcgctc ggtctgtttg gcgcgctgctc 8280
cggggggccc cggctgctg cacccagttgc cggactaggct ctctgtgccgag 8340
taatagccg gctctcctgct cattagcct gaggctgcc ccataagctggt ctaaactctc 8400
ggcctgctc gataggggaa aatagacactc gatcagctct ag 8442

<210> SEQ ID NO 108
<211> LENGTH: 7754
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: VR4777, Ligation of RSVM2 into VR4762
<400> SEQUENCE: 108

tgcgtctgg acaagtggta tacattcctg atatgctca cttatatgtg ccctatgtc 60
acccactgc caggttgtcc cgtgtattct tagatgcaac atatgctca cttatatgtg 120
ccctatgtc ccctatgtc ccctatgtc ccctatgtc ccctatgtc ccctatgtc 180
tacctgctc gccttgcacc cccctctgcc cttctgcacc cttctgcacc cttctgcacc 240
gtaaaccgc cggcaagtca cacccctgcc gttctgctc ccattgctc ccattgctc 300
gctctgctc taccagcc gccttgcacc cccctctgcc cccctctgcc cccctctgcc 360
gttcgtctgc cggcaagtca cacccctgcc gttctgctc ccattgctc ccattgctc 420
cattgctc cctgtgctc ccattgctc ccattgctc ccattgctc ccattgctc 480
atctgggtct cggcaagtca cacccctgcc gttctgctc ccattgctc ccattgctc 540
actgggtct cggcaagtca cacccctgcc gttctgctc ccattgctc ccattgctc 600
gccctggc gctgacctgct cttctgctc ccattgctc ccattgctc ccattgctc 660
cgttagcga accctcagat cgcttgaga cgcttcacc gcgtttttga cctccataga 720
agaccccggg accctacag cgctctgagc ccggagctgt gcctggaacc gcctgatcc 780
gttggcaga tgcgctgaa taccgccat agacccata gcctccccac cttgccttct 840
atcagatgta tactgctttt ggcttgaggg ccataccccc ccgctccttt gctgctatgg 900
tgtagttata cttcagcctta tagctgtaga ttatgcccta cttcagccac cttcagcttt 960
gttgcctgta ctctcccttta cttcctccata acctgccttt tgccgccacc tattcttatt 1020
gcttatacg cagcgcctgg cccctcagag gctgctcttt ttccagccttat 1080
gggtgcctat tttatatattt ccatcctctga ttatccacaa ccgcctccccc cggctgctcc 1140
gtattgattgta aacatcagctt ggccttcctg ccgaacccttto ggtctagcttg tcggacactg 1200
ggcctctctt cgcctcagcgg ggccttcctg ctcgcgctcc ctcgcgctct gctgctcctcg 1260
gtcagcagc gctgctcagc ctctggcttc taccgctgta ggcctgcacat ggcctgcaag 1320
cggcgcctc cccctcagcgg ggccttcctg ggcctggtagct gcgtgctttt tattctgaaa 1380
atgctgctgg aggctggctcg ccgcctcagcag cgcctcagtcg gccctgcacc gcctgctcctc 1440
aagagaacctcc caacctccctt gggctgcctct tctgctcata ggcctgcata ctctctcttg 1500
ggctctgtgta aacagcctgag gcggctgtag tctgctcata ctctctcctgc ggcctgcacc 1560
cggcgcctc cccctcagcgg ggcctgcctct taccgctgta ggcctgcacat ggcctgcaag 1620
gtcagcagc gctgctcagc ctctggcttc taccgctgta ggcctgcacat ggcctgcaag 1680
cggcgcctc cccctcagcgg ggcctgcctct taccgctgta ggcctgcacat ggcctgcaag 1740
acgcagagc gacccgctgc ggcctgcctct taccgctgta ggcctgcacat ggcctgcaag 1800
cgactcagcc ggcctgcctct taccgctgta ggcctgcctct taccgctgta ggcctgcacat ggcctgcaag 1860
cttggccag cggagccacag gggagccgc ccgagcgcacc gggagcgcacc gggagccgacc 1920
ccggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagcgcacc gggagccgacc 1980
gttgctgcgag cggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagcgcacc 2040
cggctgcgag cggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagcgcacc 2100
cggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2160
cggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2220
gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2280
gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2340
cggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2400
cggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2460
gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2520
gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2580
cggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2640
gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2700
ccggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2760
cggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2820
gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2880
cggagccgc ccgagcgcacc gggagccgc ccgagcgcacc gggagccgc ccgagcgcacc 2940
-continued

cggcaacacc gaggccagac cccagcgtc gaggcccgag atccacccgag tgactggagg 3000
cgccaagcgc cggaggygtgc cccagcgtgc cagaggcgtgc ttccagctgtgac gcgacgac 3060
ggcaacacc cccagctgcg cttccgctgc cagacacccg ggggagcgtc actccctggc 3120
cgcacccgcc cggagtygtgc acacgcctgc ctcggcacc gcggagctgc gctccacact 3180

tggactaaacta agaatcagac ccctctaggt cttgcggttc ggcttggagg tgtctactttc 3240
cgcctcttc gcctttgtgc gcctactgtc gcctagtggct gcctgttgct gcctgttgct gcctactgtc 3300
cctctccgac ggccttcgta cctttgctgtgc gcctgttgct gcctgttgct gcctgttgct gcctactgtc 3360
ttcgtgcag gccgtcgtgc taatcctggagt cctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc 3420
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 3480
ggacccgcc gggccctgcg aatcagccgg cttgtgccctct ggcgctgctgc ggtctcgtgc 3540
cgcctcttc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 3600
ttcgtgcag gccgtcgtgc taatcctggagt cctctctgtc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc 3660
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 3720
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 3780
ggcgctgtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 3840
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 3900
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 3960
tggaccggct gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4020
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4080
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4140
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4200
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4260
ggcgctgtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4320
cgcctgtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4380
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4440
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4500
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4560
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4620
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4680
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4740
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4800
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4860
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4920
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 4980
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 5040
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 5100
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 5160
tccctgggtgc gcctgttgct gcctgttgct gcctctctgtc gcctgttgct gcctctctgtc gcctgttgct gcctgttgct gcctctctgtc 5220
-continued

atggtgttct cacactaatc gatsagttgt cgacactagt tgcgaagcat tatgcaagc 5280
ctatatctt ccatataaat cagcatcatg tgcagattt ttcagagccac tgcgcaagca 5340
gtttccggct tgataatggc tataaccoc aactgtatta cttattatgt aagcaagacg 5400
tttattatgt catgaagata tattttttat tgcgtcaagt taccatcaaga gatctgga 5460
cactatggc tgatgataac cgcaacagag cgtagggaga aatacagaca tggatattgc 5520
tattgagtc ctctgtcttg ttgaggtggat gtctgtgtct ggtgagccag cagatttttac 5580
gacacacac ggcagaggtct gacagaagct tgcagcattg ggtgagccct 5640
ttgctgctgc ttctggtatc aagggccaga tatacgcgta tttgagggga ctgggtgttg 5700
ttaagccga aaagggggtct cagttgtcag cggcttgaaga tgcocctcgag gataatag 5760
atctctctct gcataaggag gggattattt acgttattgc aataccttttt tgcctttgcga 5820
acagttatgc gatgagtagt ccatctgctt taccagggaga gaaaaagctc cgtgtcacggc 5880
gatgagtttg cgaatagccttg caattcagtg caatgttgag gggagctgta 5940
caaggtaccgc gacacaccat gttcagatct atctactaattt atgcttctgc aagcarctagc 6000
tcgataatc cagacacctc tggccactta cacatagttc gtgcaactcc aagctctgctg 6060
atctctctct gcataaggag gggattattt acgttattgc aataccttttt tgcctttgcga 6120
gttcggc cttcgggcgt gttgctctt tgcactgctt cccggcctgc gttgaagtttt 6180
acgcgcaaggt cgccgacggg ccctcttttc ggctcctttt gcgcacgtcag tgcagcattg 6240
cgtgtttttc cggcttgggt cgcctctttt cggccgcttc gagctgagct 6300
gatgacgattc ttcgctgattc ctggctgatgg gcggcagcat gcaataaatg tggccagact 6360
cagctatagc tcttttctac ctggacaag cagctgagtt ggctgttttt ttcagataag 6420
cacacacttc ccctcttttt cgcagctttcg ggcagccaga cagccctttt cagctctggc 6480
gaccgggac cggcagagc aagggccaga tatacgcgta tttgagggga ctgggtgttg 6540
tcgctgggt cgtggtcggc aagctctttt cagctcttga cagctttcga aagctctggtt 6600
tgsagcgcgg ccggcgcgac ggggtgtgga ccggagcat ggcagaggg gacaacactag 6660
agcgccagaa cggccgagc gcgcagcacg ccctcttttt cggccgcttc gagctgagct 6720
cagctctgctt gcgcagcagc gctctgtttt tgcagcattg cttcgtggat 6780
cccgcctgc cttcctgcctt ggaagttgccc atcccgactt tctcttcgactc aaaaaatcag 6800
gcctcggct cgcgtctggt cgtaggtgct cattttttttt cggggggtgg ggtgagccag 6900
gaccgcaggggggctgc tggcagcatg atgctggtctc gtgctgattc gcgttttgct 6960
atggtgttct cacactaatc gatsagttgt cgacactagt tgcgaagcat tatgcaagc 7020
cagctggttc cggccgcttc cgcctctttt ccgctctttt ccgctctttt cgcctctttt 7080
cttagcttc ggtgactcg cggccgcttc cgcctctttt ccgctctttt cgcctctttt 7140
cgcctgttc cggccgcttc cgcctctttt ccgctctttt ccgctctttt cgcctctttt 7200
agctctggtt tgcagcattg cgtaggtgct cggccgcttc cgcctctttt ccgctctttt 7260
agctctggtt tgcagcattg cgtaggtgct cggccgcttc cgcctctttt ccgctctttt 7320
ccgccctt catggtgttt tattgccttc tgcagcattg cgtaggtgct cggccgcttc 7380
tgtgcttggc aaataatcag atgagtttga cgcggcagcct tccccgctgt cagctggtttg 7440
ggcgtctggc cggaaatatt taccagcata tccacatatt aataatggatt tccagcattg 7500
--continued

cctctctct tgttggtttt cggctagtac ggtgaaaaac ctggaacaat ctggaacac gcagctcccg 7560
gagctggctc cagcttggtc gtagtggtat ggcgccagaa gacacagcccg tccgggctgg 7620
tcttggttct tgtgggttct tgtgggttct tctactaatg cgggcataga gtagatagta 7680
cgtcgtgcct cagctatgcg gttggaata cccgcaagac gctgatggag gaatactcgcc 7740
atcatggttt cat 7754

<210> SEQ ID NO 109
<211> LENGTH: 7754
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE: OTHER INFORMATION: VR4778, Ligation of Inverted RSV RNA into VR4762

<400> SEQUENCE: 109
tgcattgtag atatgttgta tacattacat aatagcataa ttatatattg gctatactgca 60
acatcagc atatgtgca cttgattatg actagttatt aatagtcac aatagctggg 120
tcattatgct atacccgact ctctgctcag cggctgcct cagctgcgtcgtcgt 180
cctgctggc cccggtgctc cttgctttcg tcttttggtc 240
gtgatgtgact ttcgactgt cagctttgct ggttggctgg tactgttactgc 300
ctctgctggt atacattcag gttatacttg caaatcgac cccggactagt cggcagtacg 360
gtgggatg gctcgctggc ttttttttttgt ttatattttt gggggttttt 420
cggatagtct actgctatt aatgggtgta ttcgctttgg cagcttttg gatcataac 480
aatggtggt cttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>cgagcagagt gagaacgacg ggcagacaca gaaagcccacc gagaacagag ccagcgtggg</td>
<td>1740</td>
</tr>
<tr>
<td>ccagatgcgc gacggtcctg ccagatctta cctccagagt tccagcgcgc tgaatgcgtgag</td>
<td>1800</td>
</tr>
<tr>
<td>gcgactacag ggcacagtga tccgaacag ctgacactc acgagaatg ggctggcggcg</td>
<td>1860</td>
</tr>
<tr>
<td>ctgcgcagag aagaaacaca gtaaactgga ggagaccccc agcgccgca gaggcccaaa</td>
<td>1920</td>
</tr>
<tr>
<td>gagaacgcgg ggcccccttc acaagaagt gccagcgcag tggatatagag aagctgggtct</td>
<td>1980</td>
</tr>
<tr>
<td>gtagagacag gaggatactg gacagacgcc aacaaagccg ccgagccgac</td>
<td>2040</td>
</tr>
<tr>
<td>cgagcgcctgc aaccacccgtcg tgtttggaag cagcaacctcg aagagacccca ccaccacag</td>
<td>2100</td>
</tr>
<tr>
<td>aaccagagct cgctggtgca ggcgtctgta cccgacagg tggagctgta tgccgctgag</td>
<td>2160</td>
</tr>
<tr>
<td>cccctgcgcg agagacagct ggcgagcccg ggcgcgcctcg aagggcctgg gcgccatgtg</td>
<td>2220</td>
</tr>
<tr>
<td>gatggagctg atccggt translates ccagagccgg ccctcaacgc aagaaactctct ggaagggcga</td>
<td>2280</td>
</tr>
<tr>
<td>gagaacggact aagacacgaa ggcgtcttcga gaaagctgta aacaaagctt ccacacaagtt</td>
<td>2340</td>
</tr>
<tr>
<td>ccagagccgc ggcagacagcg cgactgctgga cacggtcccg gagaagccgg aacggcggca</td>
<td>2400</td>
</tr>
<tr>
<td>cgagcgcgcag gaggacgctg cttctctggcg cagacagcctcctgac aggtcttggtg</td>
<td>2460</td>
</tr>
<tr>
<td>gcccaccacag aqctgctcgct gcgctgctgc gtaacgcggcc gcgagcgcag cgaggtcaga</td>
<td>2520</td>
</tr>
<tr>
<td>ctgcgcctgc agggcctaca ccctgtcgtaag cccagacccc ccacagttgc tccagagac</td>
<td>2580</td>
</tr>
<tr>
<td>ccagctgctcg aaccagcaca ggacgccggc gcaacagcag cgaggtgctg</td>
<td>2640</td>
</tr>
<tr>
<td>gcagcgcgcctcg ccaagctgcgc cccgactgga ccgagcagct ccagagctgc ccagagcccg</td>
<td>2700</td>
</tr>
<tr>
<td>caaggtgtgcc cccagacggcg aqctgctgctccag cagagctgcat cagatgggag gcaacagaga</td>
<td>2760</td>
</tr>
<tr>
<td>caagctagac acggagtgcgc gcacccggctgcc gtaggtgac agatactcgg ccatacagac</td>
<td>2820</td>
</tr>
<tr>
<td>caaggagagc ggacacccgc aaccagcagc agccagcgc gcgcagatca gcgtgctccg</td>
<td>2880</td>
</tr>
<tr>
<td>caacccctcg ctgtagcagaa accggcctgc ccggagcagc acaggtgatgg ccgcctccac</td>
<td>2940</td>
</tr>
<tr>
<td>gcggacaccgc aggccagcgc cccagggcctg gggagcgcctg acatcagac gatggaggg</td>
<td>3000</td>
</tr>
<tr>
<td>gcggacaccgc aggccagcgc cccagggcctg cccaggtgcc cttcagcgg cagagccgc ccagagcgcga</td>
<td>3060</td>
</tr>
<tr>
<td>gcggacaccgc accgctgctcg ctgctgctgc cccagacggc cgggagcgtcc accctcctgc</td>
<td>3120</td>
</tr>
<tr>
<td>gcggacaccgc ggaggtggag accctctgc accgctgctcg cccagacggc cgggagcgtcc</td>
<td>3180</td>
</tr>
<tr>
<td>tgagttacaa aacgacagag ctctagactgt gttctttcttg tggaggttct</td>
<td>3240</td>
</tr>
<tr>
<td>cgctgctctc gttctagtgc ctcggttgctc gttcttcttg tcctttcttg tggaggttct</td>
<td>3300</td>
</tr>
<tr>
<td>ctcctactcgc ggcagaatcacc agtactcacc gccaatcacc gcgtagacag cgggtatcag</td>
<td>3360</td>
</tr>
<tr>
<td>tgtggcaccgc aagggcagac cggcgccccgg ccggcgcttg ctgctccctgct</td>
<td>3420</td>
</tr>
<tr>
<td>ctcctagct ctgctgccct gcgcagccgc cccacaaatg acgctgacct ccagggcgcgc</td>
<td>3480</td>
</tr>
<tr>
<td>gcggacccagc accgctacat ccataccaggg cttctccccct gcggagtttc ccctggctgc</td>
<td>3540</td>
</tr>
<tr>
<td>ctcctgctgc gcgcagctgtgc ctctagcagat ccgctctgc ccgctctctgc tggaggtcgc</td>
<td>3600</td>
</tr>
<tr>
<td>tgagcctttc tctctggtgc cttctgggcgg cttctgggcgg cttctgggcgg cttctgggcgg</td>
<td>3660</td>
</tr>
<tr>
<td>cgagcgcgccctg cccgacgctg aaccagcgctg ccggaggttc gcgcctgtgc tggaggaact</td>
<td>3720</td>
</tr>
<tr>
<td>acgtgctgctg gcgaacccgc gccggagcag cccagctgcc gcgcgctggag cggcgtgcttg</td>
<td>3780</td>
</tr>
<tr>
<td>acagagattg acagagtcgtg cttctgctcg tctctgctcg gttcttcsag gcgggttccta</td>
<td>3840</td>
</tr>
<tr>
<td>actcgctgcc cagacgctgc cggctgctcg cttctgctcg gttcttcsag gcgggttcct</td>
<td>3900</td>
</tr>
<tr>
<td>tggagaaaag aqctgctgctg cttctgctcg gcgaacsacac ccagcgtcgg acgcctgggtt</td>
<td>3960</td>
</tr>
</tbody>
</table>
-continued

tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
-continued

gatgctcgat atgctggggc ccaaccaacc ggggtgcctg ctgtgcgctc atcgggccc 6300
cacaattct ctagattggg tctcccgtct gcacgagcgg cctaaattgg cgaattcgat 6360
atcaggacc ggtgaacctga aaasaagacc atggaaagaa acaagtcttt agtctgctag 6420
catatatta tgggctctgc cgggcaagga acaggaactg ctaagcaatc actgcaaatc 6480
agaagagagt aagggtaggg gacaggggct aggcagcagt gcggagctg gcttcgcttc 6540
gatgctgcag accagggctg ccggccagag cgcggctgct aacctgagtt tacactttac 6600
taggaggctg aaccaggtcac gaggcaacca ccgaggggag ggcgcaagac ggcagcatac 6660
cggcctgctg gcgcgcaacca cgcggctgcag gtcgagcctg gcttcgcttc 6720
gaccgctgc accaggtgctg tgaactgctgc acttggtgctt ctattctagc gcgctgctag 6780
taaatcctg acctgtgggc cgcgcaagct gcatgtttcg ccatactctg tcaacgctag 6840
cagtgcctg ccataagctg ccaatatcct cctcaatcct ctcacacgc ggcgtgctag 6900
gtgcctttct tccctcttgc gattgtggcg ttcactctac gcctagcatt tggccaattgc 6960
cagatgtatg cagatacggt acctcccccc ctccctcttc gcacagcagc cagcaagccct 7020
cggggagct tcttagcagc gtagccacgc agcgccgctc tctgcctatt tggccaattgc 7080
tccccctggt tagttgctctg ctcagctctg accagtcgag gcggcaagac ggcagcatac 7110
catactgctg itttgaagctc aggctctga gtaacctggg tttggcttt ctaatatttc 7180
cctgctgctg gcttcgcttc 7200
tccacagctt actagctctg ctcacagcagc ggcgcaagac ggcagcatac 7260
tgcagtgccccgctgctg cttactggca cagctctgctg gtttccccccctt 7320
cccccccttt ctactaggct tttgctctga gcagtcgagc cagcaagcctt 7380
tttgcgcgat cagcaggact acaggttttgc ggtggccgct tctgcctatt gcgctgctag 7440
gacgtctag acacatcatt tatactggag ttaaccctaa ctaaatcggg ctaagctttg 7500
ctcctgctgg cgcaggact gcagttgct tgcagctctg ctcagctctg accagtcgag 7560
ggcgcaagcg gacggtctgt gcctagcggc ccgacgcaag gcggcaagag gcggcaagag 7620
tccagcgtctcctgagctc ggcggcttt ctaatctcag gcggctcactg gacagtgctg 7680
catactggag gcagtcgagc cagcaagcctt 7740
atcaggagct gctagcagc 7754

<210> SEQ ID NO 110
<211> LENGTH: 7765
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<222> OTHER INFORMATION: VR4779, 7765 bps DNA Circular
<400> SEQUENCE: 110

tggtatgctgg tgytgaacctgc cgcagctgctg cagcagtcgag ctcagctctg ctaaatcggg 60
tctgtctgctg gcggctgctg cgcagctgctg cagcagtcgag ctcagctctg ctaaatcggg 120
gtagcaggc cgcacatgc cgcagctgctg cagcagtcgag ctcagctctg ctaaatcggg 180
tcgagctgctg gcggctgctg cgcagctgctg cagcagtcgag ctcagctctg ctaaatcggg 240
tcgagctgctg gcggctgctg cgcagctgctg cagcagtcgag ctcagctctg ctaaatcggg 300
tcgagctgctg gcggctgctg cgcagctgctg cagcagtcgag ctcagctctg ctaaatcggg 360
cacagcaggc cgcagctgctg cagcagtcgag ctcagctctg ctaaatcggg 420
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>gattggtgga aataaaggtag tagatccttg ccttataggg aagcnaacag acggtctga</td>
<td>480</td>
</tr>
<tr>
<td>catgattgag aagacacact gaaatcaca gttccagaatt aagttgagctg</td>
<td>540</td>
</tr>
<tr>
<td>togtctacat aaagcagctt tgcattcata cccacatttg gttgaatctcc atgggaattcc</td>
<td>600</td>
</tr>
<tr>
<td>actctctctt caaagcgcgc cgggctacct tggagcggcc aatccagcgc ggtggcgtgc</td>
<td>660</td>
</tr>
<tr>
<td>gttggtccgc cttccgcctt gttccgactt tgaactgcgtt cgcctgctta ggaagtttta</td>
<td>720</td>
</tr>
<tr>
<td>aagctcagag cggagcagcc cgctcttgtc cgtcctcttc ggagcctcgg aatgcctcac</td>
<td>780</td>
</tr>
<tr>
<td>cggctcttca cgctttggct gcctctcttc gctacactct cagtaaccggt ggaagccagt</td>
<td>840</td>
</tr>
<tr>
<td>gtatgtctgg aagactacgtgt tgtgcocgct cgggcatcga gacataatag ctcagagact</td>
<td>900</td>
</tr>
<tr>
<td>aaacagctg tcttccgtct ggtctttttt tcgagctcaca gctgctgaatt gatcaatcttg</td>
<td>960</td>
</tr>
<tr>
<td>ccccaagggg gacacagcag ggacaaagac gcaacagaca gatccacagc gacgagcagaa</td>
<td>1020</td>
</tr>
<tr>
<td>gacacagag aacacagacc ggagccgac ccagccagcc cctcaactgc gacgagcagaa</td>
<td>1080</td>
</tr>
<tr>
<td>tctcaagca gatgctcagc gacgagcagc tcagagctac gacgagcagaa gacgagcagaa</td>
<td>1140</td>
</tr>
<tr>
<td>acagcgcctg aacacgagca attggtctga cgacgtcctga cggagctcagag gacgagcagaa</td>
<td>1200</td>
</tr>
<tr>
<td>tggagagag caccaagacgc gccaagacgc ccaagagacc gcagagcggc cttcagacaa</td>
<td>1260</td>
</tr>
<tr>
<td>ggtggtggcc caagctgtgag aagatctgttg tgcgctgata caagcagggg atcgaacaga</td>
<td>1320</td>
</tr>
<tr>
<td>tcggcgtgca gccacagcgc gcacagcggc cccagcgcgc gctgccgca acagcagagc</td>
<td>1380</td>
</tr>
<tr>
<td>tggcagcaca gctgacgcct caaactgctg gcagacacag gcagctgtggt ggcagcggca</td>
<td>1440</td>
</tr>
<tr>
<td>tggagacccag aatgctgctg tctgataggg ggcagacccct gcacgagagag gacgagcagaa</td>
<td>1500</td>
</tr>
<tr>
<td>cggggtgcgg cagttcagag acgccgtcatc gttgcttgca gttgcgtcagag gatcagacaa</td>
<td>1560</td>
</tr>
<tr>
<td>gacgatctgag cggcgctccgc tcctggagag gcacagctctg cagaaagcgc gcagagcgtc</td>
<td>1620</td>
</tr>
<tr>
<td>acgcgcgctg gttgaactgc ctaaagcggc agttcagagc gcgccagcag agaacagattg</td>
<td>1680</td>
</tr>
<tr>
<td>tggcagacag cggagagcag agaccccagc gcacagcagg gtcgagctgc tccagctcatcc</td>
<td>1740</td>
</tr>
<tr>
<td>tgggcagag cggcgtgctg cggagcagc gcctggtctgc ccagccagcgc gctgagctgc</td>
<td>1800</td>
</tr>
<tr>
<td>gcggtgtgct ccctgccctc tctgctgctgc ccaagcaggt gatcagcgtc catccagcagt</td>
<td>1860</td>
</tr>
<tr>
<td>tggagcctcgc cccacgctgc tcgctgctgc cccacgctgc gcagagcagc gcagagcagc</td>
<td>1920</td>
</tr>
<tr>
<td>acgcgcgctg gcacagagctg gtcgacgtgc ctggagagtag ggccagctccg cagctctgtg</td>
<td>1980</td>
</tr>
<tr>
<td>aagacccagct aagccagtgc tggcagagct cggagcagcgc gcagagcagc gcagagcagc</td>
<td>2040</td>
</tr>
<tr>
<td>gcaacagcag cggccgctgc gcagagcagc gcagacagctg gcagagcagc gcagagcagc</td>
<td>2100</td>
</tr>
<tr>
<td>tggaggtcag aagctagcag cggccagagc gccatagctgc cgaagcagcgc gcagagcagc</td>
<td>2160</td>
</tr>
<tr>
<td>gagagccacag cggccgctgc ctcagcagcgc gcagagcagc gcagagcagc gcagagcagc</td>
<td>2220</td>
</tr>
<tr>
<td>cctctcctag gacgctcgtt tccgctcgc gcagagcagc gcagagcagc gcagagcagc</td>
<td>2280</td>
</tr>
<tr>
<td>acgtgagac gcgcagctgc cgacagctag gggcggcgc gcagagcagc gcagagcagc</td>
<td>2340</td>
</tr>
<tr>
<td>gacgagagct cttgctgctg cgcagctgct gttcagctgct gcacgagcgc gcagagcagc</td>
<td>2400</td>
</tr>
<tr>
<td>tcgctctag cagccgctgc gcagctgctgc tgcagccgag gcagagcagc gcagagcagc</td>
<td>2460</td>
</tr>
<tr>
<td>gattgctcgg caacgggctg ccgctgctgc ctcagagcgc gcagagcagc gcagagcagc</td>
<td>2520</td>
</tr>
<tr>
<td>tttgctgtcc cccggctgtt tcgctctgc gcacgactgc gcagagcagc gcagagcagc</td>
<td>2580</td>
</tr>
<tr>
<td>aataaacaga gggagaattg gcagattttg ctcagagctc ttggggaggt ctggggaggtg</td>
<td>2640</td>
</tr>
<tr>
<td>gggtggagca gcagacagcgc gcagagcagc gcagagcagc gcagagcagc gcagagcagc</td>
<td>2700</td>
</tr>
</tbody>
</table>
cggggtgcc tgtgggtacc caggtgctga aagastgacc cgggtcttcc tggccaga
2760
gaggccgcg acatoccccct ctcgtgccca cccctctgcc ccocctcctgg tctctgttcc
2820
caccocctc catagacgcc ctctagtccga gggggtcctg gcttaaatct ccccgctca
2880
agatcaag agcggctctt ccttcocctca tcagccacac ccccccacac tagctcaca
2940
gagtggag aatataaacg aagatggtt attaagtcca gaggagagag aataagttcc
3000
acacatctgag gagaattatc gaaaaatgct gaaacatatg cggccctagc tggccaga
3060
attgcatacg ttcgatccat atcataatct tctctttcta ttcggtcact gctaacaatc
3120
acccgctgtg tgcatactag tttatactac ttcataatca cgggccatct cgttggtcatt
3180
agttttagatt cctactagtg acggatgtt atcatactag cccctataatg gcttgggcttt
3240
cgtgcccccc accgacccccc gccatgtacc cttataatgt acgtatgttt ccatagtaac
3300
gcataaggg aactcctcag gtcgctaatg gttgagagtc ttacgcggatt cccggccatc
3360
gccagctcat ccaagtcccc ttcgccactg ttcggtacac gacgatcag cccggccatc
3420
attgccccgc tggccattag cccgctatac gcctttaagc cttgccagta cccggccatc
3480
cataagcttct cttgatcttct taggacttgct tttggttcct atcctatagt gcttttggcc
3540
gcaggttgttg cgttttgcct caggggtaat ccaagctcttc cccctatatg acgtatgttt
3600
gatatggtg ttcgctgctc aatccagcgg accctgaccc taccctatag gcttttggcc
3660
cctggccatc aatccagcgg gccagtttgc tttgggttct cccggccatc ccragtcttt
3720
agatcataatg cagatggggt gccagagca ccaagccgct tttgggtac cccggccatc
3780
cgggcatcg ttcgctttct cgggaggcc accgatctcc cggcagtcag ccagctcttc
3840
cagacgtgtg gtaagtcctt cttatgctgg cccggccatc ccragtcttt
3900
ttgctgggtat cccggccttc aatccagcgc ccaagctctt ctctctgctt atgctttgcct
3960
gatggctggtgc ccgctgctgct cccggccatc ccragtcttt
4020
cgtgagctctcc ctagatctag cccggccatc ccragtcttt
4080
tgggtataag ttggttagtt cccggccatc ccragtcttt
4140
cccatatctcttagtatcttctcatcatctc cccggctccc cccggctccc
4200
tataacacatt gcacgggttt cctcagcctc cccggccatc ccragtcttt
4260
tttttagtttc ggcagagcgc tttggtctgc cccggccatc ccragtcttt
4320
tggtgctggtgc ccgctgctgct cccggccatc ccragtcttt
4380
cccggccatc cccggccatc cccggccatc cccggccatc cccggccatc
4440
cggggttgg ggggtcagctc cggggttgg cggggttgg cggggttgg cggggttgg cggggttgg cggggttgg
4500
gatggctcgc cggggttgg cggggttgg cggggttgg cggggttgg
4560
cgtgggtggt gcagcctcgt cggggttgg cggggttgg cggggttgg cggggttgg cggggttgg
4620
agcacacatc ctcgataagct cccggccatc cccggccatc cccggccatc
4680
cgcgtgtggt gcagcctcgt cggggttgg cggggttgg cggggttgg
4740
gccgggttgg gcagcctcgt cggggttgg cggggttgg cggggttgg
4800
tgcgataagct cccggccatc cccggccatc cccggccatc cccggccatc
4860
acagacatc ctcgataagct cccggccatc cccggccatc cccggccatc
4920
tggggggc cccggccatc cccggccatc cccggccatc cccggccatc
4980
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>tgaagctaga gctggagtgta tcagcgcac ccagtggtac ctctggtaat</td>
<td>5040</td>
</tr>
<tr>
<td>aaagagctcc agctctcag agcttggtcg tcggttctcg tgcgctgccg gacggtagtc agctcactctc</td>
<td>5100</td>
</tr>
<tr>
<td>tcggccaca tggctgcgc tctcctgcgt ggcgctgctgc gcgcggagcc agctgcacctc gacgcggggg</td>
<td>5160</td>
</tr>
<tr>
<td>aaagctggtac ttgcgcctac cccagtcaca ggcgggaagc cagcagacgc cagctgaggac</td>
<td>5220</td>
</tr>
<tr>
<td>aagggcagcc aaagggccag gcccagttc cccagctgtgc tgcggcaggt gtcggggggg</td>
<td>5280</td>
</tr>
<tr>
<td>ctcgccccct cttgctcagc tcacaaaata ccgtgtggca gcgacgctgg cggacacgggg</td>
<td>5340</td>
</tr>
<tr>
<td>aagctctac ccagaatacc cggctttgct ccgtaagct ccctctggcg cttctctgac</td>
<td>5400</td>
</tr>
<tr>
<td>ccagcctcgc cggcttcgcc ttccctcttc cctcggggag ccgctggccc</td>
<td>5460</td>
</tr>
<tr>
<td>ttctcacaag cgcctcaagg cgccttcgct gcgcggcagc caagctggag ccgaggttac</td>
<td>5520</td>
</tr>
<tr>
<td>tcgctcgcac cctcagctcc cagctcgcac ccagctgcag ccggcttggt cggcttggag</td>
<td>5580</td>
</tr>
<tr>
<td>ggacgcccgc ccgcggagtagt gtcgaggtt gttgccgtgg gcggctgggt cggctgggggg</td>
<td>5640</td>
</tr>
<tr>
<td>agccgagtgg gttgcctgac cggctgactac cccacggcttc gtcgaggtc ggccggggggg</td>
<td>5700</td>
</tr>
<tr>
<td>taccagggg gacgagaag ccgctctgtc gcgcctcctcc agccggaatgc gggaggctccac cccgcggaggg</td>
<td>5760</td>
</tr>
<tr>
<td>agggtgggct cgttcctgcc cggccacgact caccgctgag cgtcggtcgt gcgcggggggg</td>
<td>5820</td>
</tr>
<tr>
<td>tcgagccagc agctctgctcc cccgggcac gcgctctgac ccgctctggt ggtggtggtgt gcggctggggg</td>
<td>5880</td>
</tr>
<tr>
<td>aaagacggat cccgagggtag cggctctgac gcgctctgac ccgctttcct ggccgggggg</td>
<td>5940</td>
</tr>
<tr>
<td>tcacacggctcttcaca acctacggtg acctatctgtg acctacggtgacct cccacgtctggg</td>
<td>6000</td>
</tr>
<tr>
<td>agcatatcct acgtaacttg ggtgaccgct tccacaggt tattacatg gcgacacgc ggcgagctgg</td>
<td>6060</td>
</tr>
<tr>
<td>tccagcctca gctctctgcc ctctccgtac gtcgctcgag cggctgccgagc gcggctggg</td>
<td>6120</td>
</tr>
<tr>
<td>gtcgctgcc gttgctcgcg tctgctgag cgtcgctgac cctcctctgg gatggttgcac cccgcggctgc</td>
<td>6180</td>
</tr>
<tr>
<td>aacgctccac tggctgggtgc acggctttag agccgtcgtgc tcggctgtcg ccgctgcggg</td>
<td>6240</td>
</tr>
<tr>
<td>aattgcatt tctggtctgc cgcggagctg ccgcttggag ctcgagagtc ggccggggggd</td>
<td>6300</td>
</tr>
<tr>
<td>ttctcacaac cccacgactg cccacgactg ccgctgctgc gggcggaggt gcggaggtgg</td>
<td>6360</td>
</tr>
<tr>
<td>taccgtcct ccgcttccgc acacgactg cccacggtcg gcggctgggggc gcggctgggglg</td>
<td>6420</td>
</tr>
<tr>
<td>tcgagccgca ctgctggccct tcgcttcctgcc ctccctctctgt gcggctggggg gcggctggggg</td>
<td>6480</td>
</tr>
<tr>
<td>gttcctgtaa tgaagctgac gaaattcttc acgtcgtgac ccgcttggag ccgctttcctgt gtctgctgctg</td>
<td>6540</td>
</tr>
<tr>
<td>atctgctagc gattagcacc gttgctcctag ccctgctctgctgcgt ctcgctgctg agccgaggtggt</td>
<td>6600</td>
</tr>
<tr>
<td>acacacgggt attacgacct ccggtgttgc ggtccgagct ccgctgcttccttg gcggctggggg ccgctgggggg</td>
<td>6660</td>
</tr>
<tr>
<td>aaacgtcttg cattctcgc cagctggggt cacaaggca ccgctgcagcg tcgctgcagc tcgctgcagc</td>
<td>6720</td>
</tr>
<tr>
<td>aaiacttcga atcataacag cctgcaact ccgctctgctgcgtcc gcggcttgggc gcggctttgggg ggcggctggggg</td>
<td>6780</td>
</tr>
<tr>
<td>acgtggcgg gcggcttggggc ccgctgcttcctgccg gcggcttggtcc gcggctgggggg</td>
<td>6840</td>
</tr>
<tr>
<td>ctggcgccgct gcctgacca attcttataa cctaggggta ttcttcatt gctttatat gcgctgggt</td>
<td>6900</td>
</tr>
<tr>
<td>ctgtttttttttt cggccagcag cgcgatgggt ccgctgcgcc gcctgttcctgc ctcgctggtctg tattcggcgag</td>
<td>6960</td>
</tr>
<tr>
<td>gcttgaggg cggctggagct atatacctgc tcgcgttcct gcctgcttcgt gcgcggggtg gcgcgggggg</td>
<td>7020</td>
</tr>
<tr>
<td>taactggct gcggcttggg cggctgcgat gcggcttggt gcgcgggggg accttgggtg cggctgctggc</td>
<td>7080</td>
</tr>
<tr>
<td>ttcctgtgct gcgctgtttg cggctcgcgct ccgctggtctg ccgctgttgc gcggctgggggg ccgctgggggg</td>
<td>7140</td>
</tr>
<tr>
<td>acgctactga atcgtcgctgc gctttgcggt gtcggctgcgctgcgt cggctgttcctgc cggctttgggg</td>
<td>7200</td>
</tr>
<tr>
<td>gcctgatag gctctagca ccccttcctgc tctttttgct gtcgctggag gcgttctttgg</td>
<td>7260</td>
</tr>
</tbody>
</table>
ttcggagtga tataatatga ttgtagcga gtagagttga gacaaacagt
7320
ggttcccc ccocccocct tattgagca tttaacggg ttattgtctc atcgccggt
7380
aatatatga atgtatatcg aaatataacc aaataagggc ttcggcgcac ctcctcctga
7440
aaggtcgccg tgcgctctaa gaaaccatta tttctcagct attaacctat aaaaatagcc
7500
gctacgacag gctctcctcg ctgtccgggt tcgctgagga cggcggagc ctcctcaca
7560
tgcggcctc ggcggagcgc acaggtgctg tgaagtggg ctcgctgcgtg agacacgccc
7620
gtcggcggc gtcgctgcgttc gcggcgcggc gtgcggcgtg ggtttaacct gcggatcag
7680
agcggatct gcctagaggt ccaccatgcg ggtgctgaa tcacgacaga tgcgttaagg
7740
gaaataac ctcgagatgy gctat
7765

<210> SEQ ID NO 111
<211> LENGTH : 7765
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: VR4780, 7765 bps DNA Circular

<400> SEQUENCE: 111
tgctggcct ataataggg ccctaatagct cttatatgct tattgagc ttcgagctg
60
tgcgctctgb ctgctgccaa taaattacta attaattacta attaattacta
120
ctttatgctg tggctggctg tggctggctg tggctggctg tggctggctg
180
ataggtctgctg ttgctgctgctg ttgctgctgctg ttgctgctgctg ttgctgctgctg
240
ggcgggaggct ggagctgctg ctgctgctgctg ctgctgctgctg ctgctgctgctg
300
tggcgcggcgc gggttgctgctg ttcgagctgctg gggttgctgctg gggttgctgctg
360
ggcgggaggct ggagctgctg ctgctgctgctg ctgctgctgctg ctgctgctgctg
420
ggcgggaggct ggagctgctg ctgctgctgctg ctgctgctgctg ctgctgctgctg
480
tggcgcggcgc gggttgctgctg ttcgagctgctg gggttgctgctg gggttgctgctg
540
agacagctg ctgctgctgctg ctgctgctgctg ctgctgctgctg ctgctgctgctg
600
tggcgcggcgc gggttgctgctg ttcgagctgctg gggttgctgctg gggttgctgctg
660
agcggatct gcctagaggt ccaccatgcg ggtgctgaa tcacgacaga tgcgttaagg
720
tgcgctctgb ctgctgccaa taaattacta attaattacta attaattacta
780
ctttatgctg tggctggctg tggctggctg tggctggctg tggctggctg
840
ggcgggaggct ggagctgctg ctgctgctgctg ctgctgctgctg ctgctgctgctg
900
tggcgcggcgc gggttgctgctg ttcgagctgctg gggttgctgctg gggttgctgctg
960
agcggatct gcctagaggt ccaccatgcg ggtgctgaa tcacgacaga tgcgttaagg
1020
tgcgctctgb ctgctgccaa taaattacta attaattacta attaattacta
1080
ggcgggaggct ggagctgctg ctgctgctgctg ctgctgctgctg ctgctgctgctg
1140
tggcgcggcgc gggttgctgctg ttcgagctgctg gggttgctgctg gggttgctgctg
1200
agcggatct gcctagaggt ccaccatgcg ggtgctgaa tcacgacaga tgcgttaagg
1260
tgcgctctgb ctgctgccaa taaattacta attaattacta attaattacta
1320
ggcgggaggct ggagctgctg ctgctgctgctg ctgctgctgctg ctgctgctgctg
1380
tgcgctctgb ctgctgccaa taaattacta attaattacta attaattacta
1440
taggcocttc tggctttttcg cgtctcctcg ctcctcctac aagtgtgtgtc gttgatctgt 1500
cctctcctaa ctctcctcag ctcctcctcag aggtgctctg cgcctcctcag ggcgggcccgc 1560
gggcgcgcgc tcctcctcctgc cagggcgcgc ggcctcctcag ctcctcctcag 1620
tagttgctctgcc gcacagggcg tggcctcctgc ttgcgctaag ggtggtgtgtc 1680
cgcctcctcag tggcctcctgc gcggcgcgcgc gtcgcctcctgc cgctcctcctgc 1740
attcctcctcctgc ctgcgtcctcctgc cgcctcctcctgc tccctcctcctgc gcgcctcctcctgc 1800
cctctcctcctgc gcggcgcggtc gtgtgtgtgg cgtctcctcag ggcgcgcgc 1860
agggctctca tcctcctcctgc ctcctcctcag ctgcgctcttc tgcgcgtgctgctggctc 1920
tgcgcctcctgc ctgcgtcctcctgc ctgcgtcctcctgc gcgcctcctcctgc gcgcctcctcctgc 1980
aatgtgctc tcggcctgcgc caacgcggtgc ggcctcctcag ggcgcgcggtc 2040
tgcgcctcctgc ctgcgtcctcctgc ctgcgtcctcctgc gcgcctcctcctgc gcgcctcctcctgc 2100
ggcgcggtgctc ctgcgtcctcctgc ctgcgtcctcctgc gcgcctcctcctgc gcgcctcctcctgc 2160
tgcgcggtcttc gctgtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 2220
caacgcggtgc gcacacaaggg accaccggtgc cgcctcctcctgc gcgcctcctcctgc gcgcctcctcctgc 2280
ggctgtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 2340
tgcgctccttc ctgcgctcttc tgcgcgtgctgctggctc 2400
ggcgcggtgc gcacacaaggg accaccggtgc cgcctcctcctgc gcgcctcctcctgc gcgcctcctcctgc 2460
tgcgctccttc ctgcgctcttc tgcgcgtgctgctggctc 2520
ggcgcggtgc gcacacaaggg accaccggtgc cgcctcctcctgc gcgcctcctcctgc gcgcctcctcctgc 2580
ggcgcggtgc gcacacaaggg accaccggtgc cgcctcctcctgc gcgcctcctcctgc gcgcctcctcctgc 2640
ggcgcggtgc gcacacaaggg accaccggtgc cgcctcctcctgc gcgcctcctcctgc gcgcctcctcctgc 2700
tgcgctccttc ctgcgctcttc tgcgcgtgctgctggctc 2760
aggtggtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 2820
aggtggtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 2880
aggtggtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 2940
aggtggtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 3000
aggtggtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 3060
atgccgtacctactcctcctcctcctgc ctgcgctcttc tgcgcgtgctgctggctc 3120
atgccgtacctactcctcctcctcctgc ctgcgctcttc tgcgcgtgctgctggctc 3180
atgccgtacctactcctcctcctcctgc ctgcgctcttc tgcgcgtgctgctggctc 3240
aggtggtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 3300
aggtggtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 3360
aggtggtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 3420
aggtggtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 3480
aggtggtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 3540
aggtggtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 3600
aggtggtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 3660
aggtggtgtgtgc ggtggtgtgtgc ggcgcgcggtc ctgcgctcttc tgcgcgtgctgctggctc 3720
| agtgaacctg  | cagatgccc  | gggacgcca  | ttccagcgtg  | ttggacctcc  | atagaagaa  | 3780 |
| cgggaaccg  | ttccagctcc  | gggcgggga  | acgctgctat  | ggacacgcca  | ttccogggc  | 3840 |
| caacagtgac  | gtatagacg  | ctatagcc  | ctaagccac  | acocctttgg  | cttcatagc  | 3900 |
| ttcatatacg  | ttatagctt  | gggtctata  | aacccacccc  | ttctttgtct  | ataggctag  | 3960 |
| ttcatatagc  | ggttgctat  | ggtgctttat  | gacccatt  | gaccactccc  | cttaggcctg  | 4020 |
| ggtacgtttc  | ccattaacctg  | gtctttttcc  | gaaccatcct  | aacatattcc  | ccacactcc  | 4080 |
| tacccgata  | ccctgtccct  | cagagcctga  | ccagacattc  | gtatatttac  | agagtggtgc  | 4140 |
| ccacctttat  | atttacaaat  | tcatactac  | aacacccgct  | ttccaggcgc  | ccgacgttt  | 4200 |
| ttgtaactcc  | acgctgttac  | ccctcggcgt  | gtatgctgta  | cttgctcgcg  | acatgggcc  | 4260 |
| ttctcggcgt  | ggcgacgcc  | ttccacactcc  | ggacgctgg  | ccatgctcct  | cagcgcgct  | 4320 |
| tgcgcgctg  | ccagctggct  | gctcttaaca  | gtgagggcga  | gactagccag  | ccgagcctg  | 4380 |
| cccacccaca  | ccaggtggc  | gcacagggcc  | gtggcgggtg  | ggtatgtgctg  | tgcattaatt  | 4440 |
| cttgagaatt  | ggctgcgca  | gcttgcacga  | gttaagggc  | gtaagagaa  | gtaagaccc  | 4500 |
| gattgaccca  | gtcgtagttg  | tttgtctcct  | taagctgtc  | agagaactcc  | cttgctcgtg  | 4560 |
| ctttttcgctg  | tgtagcttgca  | gccttcagctg  | ttggttgctg  | gccggcggac  | agcagcagtt  | 4620 |
| agacatactg  | ggtcagccag  | taaccagact  | tgtttttttc  | tcggaaagct  | cggccgctgc  | 4680 |
| cggtggctg  | tacatctttg  | gcctacatc  | ggttgccgctg  | gcaggttacg  | gacgccggag  | 4740 |
| gaaagctgctg  | gggctgcga  | tgcacacca  | gacgcgaac  | cccgggttgt  | gcgocacta  | 4800 |
| tcatccgcc  | cctgagcctg  | atccccgtag  | tgtggcagct  | acgtgttctc  | aagctgctct  | 4860 |
| acatgtgctg  | cagagcagcc  | cttggttgcg  | ggcctagctg  | ccgagctggtg  | cgcgtagggc  | 4920 |
| tggagaga  | ccacactaga  | gacacacaga  | acgcctggga  | gcgcacgacg  | agccactcg  | 4980 |
| tggacagctg  | ggtgtggaca  | ccctgcagctg  | acgctgctac  | cagatagtct  | cctgtgtacct  | 5040 |
| aacaactctg  | agctctcgag  | acagtctgtg  | tgtttttttg  | tgtgtgtctc  | ttccacctcc  | 5100 |
| tcggactctg  | acgcctcgcc  | cccctcccgt  | ggcgctgctg  | gcacggctgct  | acgocactc  | 5160 |
| actggcctg  | taacggcttc  | ccgcaacacta  | gggctgacg  | cagacgacac  | cctggctga  | 5220 |
| aacggccgac  | aagacggctg  | gacgctcaga  | ccggctgctg  | tgggggcttg  | tccctatgct  | 5280 |
| ttcgcgcgctc  | ccgagagcgc  | tccaaccaat  | ccgactgcc  | gtaacgcttg  | gcgacacccg  | 5340 |
| acacactctg  | aaagccagag  | gcggtctctc  | ccctggactg  | cccctgtctc  | ctcggctttc  | 5400 |
| cggagccctg  | cgatttcccctg  | cggcctttcc  | ccctggcagc  | cctggcccct  | 5460 |
| tcctatcctg  | ccagctctag  | ttcgcctcgtg  | tgtgtctgctg  | ccctggctgct  | acggctctg  | 5520 |
| tgcggtgctg  | aaccccacccg  | ccgcttcgtc  | cgggtgctg  | cagcgcctgtg  | ttcggcctg  | 5580 |
| gaccttggcgt  | ccctagacgctc  | atcccactcc  | cggccttccc  | ctcgggagc  | cgtggctgt  | 5640 |
| agccagccg  | gtgctagctg  | cggtgctgctc  | gatcttttg  | gcacggccgc  | tgcacctcgc  | 5700 |
| tgaatctagg  | gcaacagctg  | tccttcgagc  | gatcttttg  | ggcgcctgtc  | tgttctgctc  | 5760 |
| agaggtgtga  | gtttgtggctg  | gcagacccc  | accacgctgg  | gtagggcggt  | tttttgtgg  | 5820 |
| tgcacagctg  | ccgaacaaaa  | aaccccgctg  | gtgcgtggg  | gcctgcgaaga  | aagatctttt  | 5880 |
| aaggggtgctg  | acogctcgtg  | gcagacacac  | tcagccgctg  | ggtttttggtg  | ctcgagatg  | 5940 |
| tcacacggc  | cttttttccctg  | gatccttttct  | aactttct  | gatctttttc  | aactcctaa  | 6000 |
agtatattg agtaaacttg gctgacagt taccaagtgt ttaaataagtga gcgaacctac 6060
tcgcctctc gctgattcgg tcctccacat gttggtcagc tgggggggg ggggcgcgtga 6120
ggtctgcct cgacgagcag tggatttgac gcctccaggc ttgtaaacg gacgctatcc 6180
ggccagagcg aaagggagc acgggtagt agagtttgag ttgtaggtga ccaagtttgtg 6240
acttgagcact ttgggtttccc cacgggaacct cctgcgtgtgc cggaagactg cggtgatoga 6300
tcctccact cccactctag tgcattcctg caactaaagcc gcgtccccct cagagtagacg 6360
tcggggct gcaagttgtag aaccacttga caactctttg ccggagaaac ccctccgcoc 6420
tcggactgac gcttccgatc gtttggccccg aatgctctttg ctcgggtcagc 6480
aactcggatat ggtggtttgtg caacggggagt gataaggtcg caggttggaa 6540
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 6600
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 6660
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 6720
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 6780
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 6840
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 6900
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 6960
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7020
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7080
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7140
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7200
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7260
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7320
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7380
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7440
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7500
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7560
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7620
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7680
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7740
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7800
aatcggatac ggtggtttgtg caacggggagt gataaggtcg caggttggaa 7865
accatattgc gctgtaaata acaoacagag gcgtaaggag aataactcgc atcagatgg 240
ctatttgtcct cccctcggct ctggtgtaga gtagcgttag gatggaagga gcataatatta 300
acgtacacca acgcacagct ttcacaggtc ttgtcatgag gcctggttca atctgtgtta 360
ttttcgctg gcctgcgcct tgaagctcct cctacaagct ccggaggggg acctagggtctg 420
gtgggtgca aacgggagc tctggtgatt cggggtattag gctgtggcttccc ggtatagta 480
cttggttta cagccacagg gggggaatcc tacccttttt catctacttt ggatttttccc 540
aacgctgtcga ggatgagttt gcaacagcgct ttcacagggag agaaanacca acgcacaggtc 600
cgtggtgacg aaggtgagag tttatagtgt gcctattacc ttcggcggagc gacggttcgtg 660
acagtctgtg gaagacaccc tgaatccgcc attcagagcc tttgttttcttt atctgtctcg 720
ttcagctcct ttcagctcct ttcagctcct tcagctcctc ttcagctcct ccggaggggg acctagggtctg 780
cctacgttc ccaacaccgg ataaccagcc atcaggttccc gcctggttca atctgtgtta 840
gttctgtcgt cccctcggct tgaagctcct cctacaagct ccggaggggg acctagggtctg 900
aagaccctcg ttgcaagcgc gctgtggcttccc ggggaggtgc ctgctgtggcttccc ggtatagta 960
cctgctgtg ccgctgtggc gctgctgtggc gcctggttca atctgtgtta 1020
ttgagttttt catctacttt ggatttttccc ttcacagggag agaaanacca acgcacaggtc 1080
tccctctttc ccaacaccgg ataaccagcc atcaggttccc gcctggttca atctgtgtta 1140
gctggtttc cccctcggct tgaagctcct cctacaagct ccggaggggg acctagggtctg 1200
cctacgttc ccaacaccgg ataaccagcc atcaggttccc gcctggttca atctgtgtta 1260
ccgctgtggc gcctggttca atctgtgtta 1320
ccgctgtggc gcctggttca atctgtgtta 1380
cctacgttc ccaacaccgg ataaccagcc atcaggttccc gcctggttca atctgtgtta 1440
tccctctttc ccaacaccgg ataaccagcc atcaggttccc gcctggttca atctgtgtta 1500
ttcagctcct ttcagctcct ttcagctcct ccggaggggg acctagggtctg 1560
ctggtttc cccctcggct tgaagctcct cctacaagct ccggaggggg acctagggtctg 1620
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 1680
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 1740
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 1800
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 1860
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 1920
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 1980
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 2040
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 2100
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 2160
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 2220
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 2280
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 2340
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 2400
acgcacagct ttcacagctt cccctcggct ttcacagctt cccctcggct ttcacagctt 2460
---continued---

(continued)

1-392. (canceled)

393. An isolated polynucleotide comprising a nucleic acid fragment which encodes the consensus amino acid sequence of SEQ ID NO:78, wherein the codons of said nucleic acid fragment are optimized for expression in humans.

394. The polynucleotide of claim 393, wherein the nucleotide sequence of said nucleic acid fragment is SEQ ID NO:66.

395. A vector comprising the polynucleotide of claim 393, wherein said vector, upon uptake by a suitable host cell, expresses said amino acid sequence.

396. The polynucleotide of claim 393, further comprising a heterologous nucleic acid ligated to said nucleic acid fragment.

397. A composition comprising the vector of claim 395 and a carrier.

398. The composition of claim 397, further comprising a component selected from the group consisting of an adjuvant and a transfection facilitating compound.

399. The composition of claim 398, wherein said component is a cationic lipid.

400. The composition of claim 399, wherein said adjuvant comprises (z)-N-(3-aminopropyl)-N,N-dimethyl-2,3-bis-
(syn-9-tetradecenylxyloxy)-1-propanaminium bromide (GAP-DMORIE) and a neutral lipid, wherein said neutral lipid is selected from the group consisting of:

(a) 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE);
(b) 1,2-diphytanoyl-sn-glycero-3-phosphoethanolamine (DPyPE); and
(c) 1,2-dimyristoyl-glycer-3-phosphoethanolamine (DMPE).

401. The composition of claim 399, wherein said transfection facilitating compound comprises (±)-N-(2-hydroxyethyl)-N,N-dimethyl-2,3-bis(tetradecyloxy)-1-propanaminium bromide (DMRIE).

402. The composition of claim 401, wherein said transfection facilitating compound further comprises a neutral lipid.

403. The composition of claim 402, wherein said neutral lipid is DOPE.

404. The composition of claim 401 further comprising a 1:1 molar ratio of GAP-DMORIE and DPyPE.

405. A method for treating or preventing influenza infection in a vertebrate comprising administering to a vertebrate in need thereof the composition of claim 397.

406. A method for eliciting an immune response to influenza virus in a vertebrate comprising administering to a vertebrate in need thereof the composition of claim 397.

407. An isolated nucleotide comprising a nucleic acid fragment which encodes the consensus amino acid sequence of SEQ ID NO:76, wherein the codons of said nucleic acid fragment are optimized for expression in humans.

408. The composition of claim 407, wherein the nucleotide sequence of said nucleic acid fragment is SEQ ID NO:75.

409. A vector comprising the nucleotide of claim 405, wherein said vector, upon uptake by a suitable host cell, expresses said amino acid sequence.

410. The nucleotide of claim 407, further comprising a heterologous nucleic acid ligated to said nucleic acid fragment.

411. A composition comprising the vector of claim 409 and a carrier

412. The composition of claim 411, further comprising a component selected from the group consisting of an adjuvant and a transfection facilitating compound.

413. The composition of claim 412, wherein said component is a cationic lipid.

414. The composition of claim 413, wherein said adjuvant comprises (±)-N-(3′-aminopropl)-N,N-dimethyl-2,3-bis(syn-9-tetradecenylxyloxy)-1-propanaminium bromide (GAP-DMORIE) and a neutral lipid, wherein said neutral lipid is selected from the group consisting of:

(a) 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE);
(b) 1,2-diphytanoyl-sn-glycero-3-phosphoethanolamine (DPyPE); and
(c) 1,2-dimyristoyl-glycer-3-phosphoethanolamine (DMPE).

415. The composition of claim 413, wherein said transfection facilitating compound comprises (±)-N-(2-hydroxyethyl)-N,N-dimethyl-2,3-bis(tetradecyloxy)-1-propanaminium bromide (DMRIE).

416. The composition of claim 415, wherein said transfection facilitating agent further comprises a neutral lipid.

417. The composition of claim 416, wherein the neutral lipid is DOPE.

418. The composition of claim 415 further comprising a 1:1 molar ratio of GAP-DMORIE and DPyPE.

419. An isolated polypeptide produced by the polynucleotide claim 407.

420. A method for treating or preventing influenza infection in a vertebrate comprising administering to a vertebrate in need thereof the composition of claim 411.

421. A method for eliciting an immune response to influenza virus in a vertebrate by administration of the composition of claim 411.

422. An isolated polynucleotide comprising a first nucleic acid fragment which encodes the consensus amino acid sequence of SEQ ID NO:78 and a second nucleic acid fragment which encodes the consensus amino acid sequence of SEQ ID NO:76, wherein the codons of said first and second nucleic acid fragments are optimized for expression in humans.

423. The polynucleotide of claim 422, wherein the nucleotide sequence of said first nucleic acid fragment is SEQ ID NO:66 and wherein the nucleotide sequence of said second nucleic acid fragment is SEQ ID NO:75.

424. A vector comprising the polynucleotide of claim 422, wherein said vector, upon uptake by a suitable host cell, expresses the consensus amino acid sequences of SEQ ID NO:78 and SEQ ID NO:76.

425. The vector of claim 424, wherein said consensus amino acid sequences of SEQ ID NO:78 and SEQ ID NO:76 are expressed as a fusion protein.

426. The vector of claim 422, wherein said vector is DNA and wherein said vector comprises a first expression cassette and second expression cassette, said first expression cassette comprises a first nucleic acid fragment which encodes the consensus amino acid sequence of SEQ ID NO:78 in operable association with a promoter and said second expression cassette comprises a second nucleic acid fragment which encodes the consensus amino acid sequence of SEQ ID NO:76 in operable association with a promoter.

427. The vector of claim 426, wherein said first expression cassette and said second expression cassette are associated with separate promoters.

428. The vector of claim 427, wherein said separate promoters are non-identical.

429. The vector of claim 426, wherein said first expression cassette and said second expression cassette are associated with a single promoter, and wherein said second expression cassette is in operable association with an internal ribosome entry site (IRES).

430. The vector of claim 426, wherein said first expression cassette and said second expression cassette are associated with a single promoter, and wherein said first expression cassette is in operable association with an internal ribosome entry site (IRES).

431. A composition comprising the vector of claim 424 and a carrier.

432. A composition comprising the vector of claim 426 and a carrier.

433. A composition comprising at least two non-identical vectors, wherein one of said vectors comprises a nucleic acid fragment which encodes the consensus amino acid sequence of SEQ ID NO:78 and wherein another of said vectors
comprises a nucleic acid fragment which encodes the consensus amino acid sequence of SEQ ID NO:76, wherein the codons of said nucleic acid fragments encoding SEQ ID NO:78 and SEQ ID NO:76 are optimized for expression in humans, and wherein said vectors, upon uptake by a suitable host cell, express said amino acid sequences.

434. The composition of claim 433, further comprising a carrier.

435. The composition of claim 434, further comprising a component selected from the group consisting of an adjuvant and a transfection facilitating compound.

436. The composition of claim 435, wherein said component is a cationic lipid.

437. The composition of claim 436, wherein said adjuvant comprises (z)-N-(3-aminopropyl)-N,N-dimethyl-2,3-bis-(syn-9-tetradeceneyloxy)-1-propanaminium bromide (GAP-DMORIE) and a neutral lipid, wherein said neutral lipid is selected from the group consisting of:

(a) 1,2-dioleoyl-sn-glycero-3-phosph ethanolamine (DOPE);

(b) 1,2-diphytanoyl-sn-glycero-3-phosphoethanolamine (DpyPE); and

(c) 1,2-dimyristoyl-glycer-3-phosphoethanolamine (DMPE).

438. The composition of claim 436, wherein said transfection facilitating compound further comprises (z)-N-(2-hydroxyethyl)-N,N-dimethyl-2,3-bis(tetradecyloxy)-1-propanaminium bromide (DMRIE).

439. The composition of claim 438, wherein said transfection facilitating compound further comprises a neutral lipid.

440. The composition of claim 439, wherein the neutral lipid is DOPE.

441. The composition of claim 438 further comprising a 1:1 molar ratio of GAP-DMORIE and DpyPE.

442. A method for treating or preventing influenza infection in a vertebrate comprising administering to a vertebrate in need thereof the composition of claim 434.

443. A method for eliciting an immune response to influenza virus in a vertebrate by administration of the composition of claim 434.

* * * * *