Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2755257 A1 2010/09/23

(21) 2 755 257

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2010/03/16
(87) Date publication PCT/PCT Publication Date: 2010/09/23
(85) Entrée phase nationale/National Entry: 2011/09/12
(86) N° demande PCT/PCT Application No.: US 2010/027552
(87) N° publication PCT/PCT Publication No.: 2010/107847
(30) Priorité/Priority: 2009/03/16 (US61/210,305)

(51) Cl.Int./Int.Cl. *A61K 39/12* (2006.01),
C12N 7/00 (2006.01)

(71) Demandeurs/Applicants:
SANOFI PASTEUR BIOLOGICS CO., US;
SANOFI PASTEUR LIMITED, CA

(72) Inventeurs/Inventors:
PUGACHEV, KONSTANTIN V., US;
RUMYANTSEV, ALEXANDER A., US;
GIEL-MOLONEY, MARYANN, US;
PARRINGTON, MARK, CA;
ZHANG, LINONG, CA;
ZHANG, LINXNG, CA

(74) Agent: FASKEN MARTINEAU DUMOULIN LLP

(54) Titre : VECTEURS DE VACCIN ANTI-FLAVIVIRUS A REPLICATION DEFICIENTE CONTRE UN VIRUS SYNCYTIAL RESPIRATOIRE

(54) Title: REPLICATION-DEFECTIVE FLAVIVIRUS VACCINE VECTORS AGAINST RESPIRATORY SYNCYTIAL VIRUS

(57) Abrégé/Abstract:
Replication-defective vaccine vectors against respiratory syncytial virus (RSV) are disclosed. Corresponding compositions and methods employing the vaccine vectors are also disclosed.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 September 2010 (23.09.2010)

PCT

(10) International Publication Number
WO 2010/107847 A1

(51) International Patent Classification:
A61K 39/12 (2006.01) *C12N 7/00* (2006.01)

(21) International Application Number:
PCT/US2010/027552

(22) International Filing Date:
16 March 2010 (16.03.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/210,305 16 March 2009 (16.03.2009) US

(71) Applicants *(for all designated States except US)*: **SANOFI PASTEUR BIOLOGICS CO.** [US/US]; 38 Sidney Street, Cambridge, MA 02139 (US). **SANOFI PASTEUR LIMITED** [CA/CA]; 1755 Steeles Avenue West, Toronto, ON M2R 3T4 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **PUGACHEV, Konstantin, V.** [US/US]; 26 Harwood Road, Natick, MA 01760 (US). **RUMYANTSEV, Alexander, A.** [RU/US]; 154 Cedar Street, Apt. 4-1, Somerville, MA 02144 (US). **GIEL-MOLONEY, Maryann** [US/US]; 33 Brooksdale Road, Brighton, MA 02135 (US). **PARRINGTON, Mark** [CA/CA]; 214 Melbourne Drive, Bradford, ON L3Z 2Y8 (CA). **ZHANG, Linong** [CA/CA]; 73 Ascalon Drive, Maple, ON L6A 0M4 (CA).

(74) Agent: **MICHAUD, Susan, M.**; Clark & Elbing LLP, 101 Federal Street, Boston, MA 02110 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

— *with international search report (Art. 21(3))*

WO 2010/107847 A1

(54) **Title**: REPLICATION-DEFECTIVE FLAVIVIRUS VACCINE VECTORS AGAINST RESPIRATORY SYNCYTIAL VIRUS

(57) **Abstract**: Replication-defective vaccine vectors against respiratory syncytial virus (RSV) are disclosed. Corresponding compositions and methods employing the vaccine vectors are also disclosed.

REPLICATION-DEFECTIVE FLAVIVIRUS VACCINE VECTORS AGAINST RESPIRATORY SYNCYTIAL VIRUS

Cross-Reference to Related Applications

This application claims benefit of U.S. provisional application no. 61/210,305, filed March 16, 2009, the contents of which are incorporated herein by reference.

Field of the Invention

This invention relates to replication-defective flavivirus vaccine vectors against respiratory syncytial virus (RSV), and corresponding compositions and methods.

Background of the Invention

Flaviviruses are distributed worldwide and represent a global public health problem. Flaviviruses also have a significant impact as veterinary pathogens. Flavivirus pathogens include yellow fever (YF), dengue types 1-4 (DEN1-4), Japanese encephalitis (JE), West Nile (WN), tick-borne encephalitis (TBE), and other viruses from the TBE serocomplex, such as Kyasanur Forest disease (KFD) and Omsk hemorrhagic fever (OHF) viruses. Vaccines against YF [live attenuated vaccine (LAV) strain 17D], JE [inactivated vaccines (INV) and LAV], and TBE (INV) are available. No licensed human vaccines are currently available against DEN and WN. Veterinary vaccines have been in use including, for example, vaccines against WN in horses (INV, recombinant and live chimeric vaccines), JE (INV and LAV) to prevent encephalitis in horses and stillbirth in pigs in Asia, louping ill flavivirus (INV) to prevent neurologic disease in sheep in the UK, and TBE (INV) used in farm animals in Czech Republic (INV) (Monath and Heinz, Flaviviruses, in Fields et al. Eds., Fields Virology, 3rd Edition, Philadelphia, New York, Lippincott-Raven Publishers, 1996, pp. 961-1034).

Flaviviruses are small, enveloped, plus-strand RNA viruses transmitted primarily by arthropod vectors (mosquitoes or ticks) to natural hosts, which are primarily vertebrate animals, such as various mammals, including humans, and birds.

The flavivirus genomic RNA molecule is about 11,000 nucleotides (nt) in length and encompasses a long open reading frame (ORF) flanked by 5' and 3' untranslated terminal regions (UTRs) of about 120 and 500 nucleotides in length, respectively. The ORF encodes a polyprotein precursor that is cleaved co- and post-translationally to generate individual viral proteins. The proteins are encoded in the order: C-prM/M-E-NS1-NS2A/2B-NS3-NS4A/4B-NS5, where C (core/capsid), prM/M (pre-membrane/membrane), and E (envelope) are the structural proteins, i.e., the components of viral particles, and the NS proteins are non-structural proteins, which are involved in intracellular virus replication. Flavivirus replication occurs in the cytoplasm. Upon infection of cells and translation of genomic RNA, processing of the polyprotein starts with translocation of the prM portion of the polyprotein into the lumen of endoplasmic reticulum (ER) of infected cells, followed by translocation of E and NS1 portions, as directed by the hydrophobic signals for the prM, E, and NS1 proteins. Amino-termini of prM, E, and NS1 proteins are generated by cleavage with cellular signalase, which is located on the luminal side of the ER membrane, and the resulting individual proteins remain carboxy-terminally anchored in the membrane. Most of the remaining cleavages, in the nonstructural region, are carried out by the viral NS2B/NS3 serine protease. The viral protease is also responsible for generating the C-terminus of the mature C protein found in progeny virions. Newly synthesized genomic RNA molecules and the C protein form a dense spherical nucleocapsid, which becomes surrounded by cellular membrane in which the E and prM proteins are embedded. The mature M protein is produced by cleavage of prM shortly prior to virus release by cellular furin or a similar protease. E, the major protein of the envelope, is the principal target for neutralizing antibodies, the main correlate of immunity against flavivirus infection. Virus-specific cytotoxic T-lymphocyte (CTL) response is the other key attribute of immunity. Multiple CD8+ and CD4+ CTL epitopes have been characterized in various flavivirus structural and non-structural proteins. In addition, innate immune responses contribute to both virus clearance and regulating the development of adaptive immune responses and immunologic memory.

In addition to the inactivated (INV) and live-attenuated (LAV) vaccines against flaviviruses discussed above, other vaccine platforms have been developed. One example is based on chimeric flaviviruses that include yellow fever virus capsid

and non-structural sequences and prM-E proteins from other flaviviruses, to which immunity is sought. This technology has been used to develop vaccine candidates against dengue (DEN), Japanese encephalitis (JE), West Nile (WN), and St. Louis encephalitis (SLE) viruses (see, e.g., U.S. Patent Nos. 6,962,708 and 6,696,281). Yellow fever virus-based chimeric flaviviruses have yielded highly promising results in clinical trials.

Another flavivirus vaccine platform is based on the use of pseudoinfectious virus (PIV) technology (Mason et al., Virology 351:432-443, 2006; Shustov et al., J. Virol. 21:11737-11748, 2007; Widman et al., Adv. Virus. Res. 72:77-126, 2008; Suzuki et al., J. Virol. 82:6942-6951, 2008; Suzuki et al., J. Virol. 83:1870-1880, 2009; Ishikawa et al., Vaccine 26:2772-2781, 2008; Widman et al., Vaccine 26:2762-2771, 2008). PIVs are replication-defective viruses attenuated by a deletion(s). Unlike live flavivirus vaccines, they undergo a single round replication *in vivo* (or optionally limited rounds, for two-component constructs; see below), which may provide benefits with respect to safety. PIVs also do not induce viremia and systemic infection. Further, unlike inactivated vaccines, PIVs mimic whole virus infection, which can result in increased efficacy due to the induction of robust B- and T-cell responses, higher durability of immunity, and decreased dose requirements. Similar to whole viruses, PIV vaccines target antigen-presenting cells, such as dendritic cells, stimulate toll-like receptors (TLRs), and induce balanced Th1/Th2 immunity. In addition, PIV constructs have been shown to grow to high titers in substrate cells, with little or no cytopathic effect (CPE), allowing for high-yield manufacture, optionally employing multiple harvests and/or expansion of infected substrate cells.

The principles of the PIV technology are illustrated in Figs. 1 and 2. There are two variations of the technology. In the first variation, a single-component pseudoinfectious virus (s-PIV) is constructed with a large deletion in the capsid protein (C), rendering mutant virus unable to form infectious viral particles in normal cells (Fig. 1). The deletion does not remove the first ~20 codons of the C protein, which contain an RNA cyclization sequence, and a similar number of codons at the end of C, which encode a viral protease cleavage site and the signal peptide for prM. The s-PIV can be propagated, e.g., during manufacture, in substrate (helper) cell cultures in which the C protein is supplied *in trans*, e.g., in stably transfected cells

producing the C protein (or a larger helper cassette including C protein), or in cells containing an alphavirus replicon [e.g., a Venezuelan equine encephalitis virus (VEE) replicon] expressing the C protein or another intracellular expression vector expressing the C protein. Following inoculation in vivo, e.g., after immunization, the PIV undergoes a single round of replication in infected cells in the absence of trans-complementation of the deletion, without spread to surrounding cells. The infected cells produce empty virus-like particles (VLPs), which are the product of the prM-E genes in the PIV, resulting in the induction of neutralizing antibody response. A T-cell response should also be induced via MHCI presentation of viral epitopes. This approach has been applied to YF 17D virus and WN viruses and WN/JE and WN/DEN2 chimeric viruses (Mason et al., Virology 351:432-443, 2006; Suzuki et al., J. Virol. 83:1870-1880, 2009; Ishikawa et al., Vaccine 26:2772-2781, 2008; Widman et al., Vaccine 26:2762-2771, 2008; WO 2007/098267; WO 2008/137163).

In the second variation, a two-component PIV (d-PIV) is constructed (Fig. 2). Substrate cells are transfected with two defective viral RNAs, one with a deletion in the C gene and another lacking the prM-E envelope protein genes. The two defective genomes complement each other, resulting in accumulation of two types of PIVs in the cell culture medium (Shustov et al., J. Virol. 21:11737-11748, 2007; Suzuki et al., J. Virol. 82:6942-6951, 2008). Optionally, the two PIVs can be manufactured separately in appropriate helper cell lines and then mixed in a two-component formulation. The latter may offer an advantage of adjusting relative concentrations of the two components, increasing immunogenicity and efficacy. This type of PIV vaccine should be able to undergo a limited spread in vivo due to coinfection of some cells at the site of inoculation with both components. The spread is expected to be self-limiting as there are more cells in tissues than viral particles produced by initially coinfected cells. In addition, a relatively high MOI is necessary for efficient co-infection, and cells outside of the inoculation site are not expected to be efficiently coinfected (e.g., in draining lymph nodes). Cells infected with the ΔC PIV alone produce the highly immunogenic VLPs. Coinfected cells produce the two types of packaged defective viral particles, which also stimulate neutralizing antibodies. The limited infection is expected to result in a stronger neutralizing antibody response and T-cell response compared to s-PIVs. To decrease chances of recombination during

manufacture or in vivo, including with circulating flaviviruses, viral sequences can be modified in both s-PIVs and d-PIVs using, e.g., synonymous codon replacements, to reduce nucleotide sequence homologies, and mutating the complementary cyclization 5' and 3' elements.

Respiratory syncytial virus (RSV) is a negative-sense, single-stranded RNA virus of the family *Paramyxoviridae*. Its name is based on the activity of the RSV fusion or F glycoprotein, which is on the surface of the virus and causes cell membranes of infected cells to merge, resulting in the formation of syncytia. RSV infects the respiratory tract, and is the major cause of lower respiratory tract infections (including pneumonia) and hospital visits during infancy and childhood. For example, in the United States, 60% of infants are infected during their first RSV season, and nearly all children will have been infected by 2-3 years of age (Glezen et al., Am. J. Dis. Child. 140(6):543-546, 1986). Of those infected, 2-3% will develop bronchiolitis, or inflammation of the small airways in the lung, and require hospitalization (Hall et al., N. Engl. J. Med. 360(6):588–598, 2009). Further, RSV infection is increasingly being found as an infection of the elderly. Current treatment is generally focused on supportive care, including administration of fluids and oxygen.

Summary of the Invention

The invention provides replication-deficient pseudoinfectious flaviviruses that each include a flavivirus genome including (i) one or more deletions or mutations in nucleotide sequences encoding one or more proteins selected from the group consisting of capsid (C), pre-membrane (prM), envelope (E), non-structural protein 1 (NS1), non-structural protein 3 (NS3), and non-structural protein 5 (NS5), and (ii) a sequence encoding a respiratory syncytial virus (RSV) peptide or protein, or a fragment or analog thereof. As described elsewhere herein, the vectors of the invention are replication deficient due to the one or more deletions or mutations, and can be complemented *in trans* (see below for details). Any of the deletions/mutations described herein, as well as other deletions/mutations resulting in replication deficiency, can be used in the vectors of the invention.

In one embodiment, the respiratory syncytial virus (RSV) protein is the RSV F protein, or a fragment or analog thereof. In various examples, the RSV F protein

lacks a trans-membrane domain, e.g., it is truncated so that it is produced in secreted form. In other examples, the respiratory syncytial virus (RSV) protein is the RSV G protein, or a fragment or analog thereof.

In various embodiments, the one or more deletions or mutations is within capsid (C) sequences of the flavivirus genome; is within pre-membrane (prM) and/or envelope (E) sequences of the flavivirus genome; is within capsid (C), pre-membrane (prM), and envelope (E) sequences of the flavivirus genome; and/or is within non-structural protein 1 (NS1) sequences of the flavivirus genome. In other examples, the flavivirus genome includes sequences encoding a pre-membrane (prM) and/or envelope (E) protein.

The flavivirus genome of the replication-deficient pseudoinfectious flaviviruses can be, for example, selected from that of yellow fever virus, West Nile virus, tick-borne encephalitis virus, Langat virus, Japanese encephalitis virus, dengue virus (1-4), and St. Louis encephalitis virus sequences, and chimeras thereof (also see below). In certain examples, the chimeras include pre-membrane (prM) and envelope (E) sequences of a first flavivirus, and capsid (C) and non-structural sequences of a second, different flavivirus. In other examples, the genome is packaged in a particle including pre-membrane (prM) and envelope (E) sequences from a flavivirus that is the same or different from that of the genome. Further, sequences encoding the RSV protein can be inserted in the place of or in combination with the one or more deletions or mutations of the one or more proteins.

In certain examples, sequences encoding the respiratory syncytial virus peptide or protein, or a fragment or analog thereof, are inserted in the flavivirus genome within sequences encoding the envelope (E) protein, within sequences encoding the non-structural 1 (NS1) protein, within sequences encoding the pre-membrane (prM) protein, intergenically between sequences encoding the envelope (E) protein and non-structural protein 1 (NS1), intergenically between non-structural protein 2B (NS2B) and non-structural protein 3 (NS3), or as a bicistronic insertion in the 3' untranslated region of the flavivirus genome.

The invention also includes pharmaceutical compositions including one or more of the replication-deficient pseudoinfectious flaviviruses described above and

elsewhere herein. Compositions of the invention can also a pharmaceutically acceptable carrier or diluent, and, optionally, an adjuvant.

Other compositions of the invention include a first replication-deficient pseudoinfectious flavivirus, such as one of those described above and elsewhere herein, and a second, different replication-deficient pseudoinfectious flavivirus including a genome having one or more deletions or mutations in nucleotide sequences encoding one or more proteins selected from the group consisting of capsid (C), pre-membrane (prM), envelope (E), non-structural protein 1 (NS1), non-structural protein 3 (NS3), and non-structural protein 5 (NS5), wherein the one or more proteins encoded by the sequences in which the one or more deletion(s) or mutation(s) occur in the second, different replication-deficient pseudoinfectious flavivirus are different from the one or more proteins encoded by the sequences in which the one or more deletion(s) or mutation(s) occur in the first replication-deficient pseudoinfectious flavivirus.

The invention also provides methods of inducing an immune response to respiratory syncytial virus (RSV) in a subject, involving administering to the subject one or more replication-deficient pseudoinfectious flaviviruses or a composition as described above and elsewhere herein. The subject may be at risk of but not have an infection by respiratory syncytial virus (RSV), or the subject may have an infection by respiratory syncytial virus (RSV). In certain examples, the subject is an infant, young child, or elderly person. The methods of the invention can be for inducing an immune response against a protein encoded by the flavivirus genome, in addition to RSV. In such methods, the subject may be at risk of but does not have an infection by the flavivirus corresponding to the genome of the pseudoinfectious flavivirus, which includes sequences encoding a flavivirus pre-membrane and/or envelope protein. In another example, the subject has an infection by the flavivirus corresponding to the genome of the pseudoinfectious flavivirus, which includes sequences encoding a flavivirus pre-membrane and/or envelope protein.

Also included in the invention are nucleic acid molecules corresponding to the genomes of pseudoinfectious flaviviruses as described herein and complements thereof.

The invention also provides methods of making a replication-deficient pseudoinfectious flavivirus as described herein. These methods involve introducing a nucleic acid molecule as described above into a cell that expresses the protein corresponding to any sequences deleted from the flavivirus genome of the replication-deficient pseudoinfectious flavivirus. The protein can be expressed in the cell from, for example, the genome of a second, different, replication-deficient pseudoinfectious flavivirus. In various examples, the protein is expressed from a replicon (e.g., an alphavirus replicon, such as a Venezuelan Equine Encephalitis virus replicon).

By "replication-deficient pseudoinfectious flavivirus" or "PIV" is meant a flavivirus that is replication-deficient due to a deletion or mutation in the flavivirus genome. The deletion or mutation can be, for example, a deletion of a large sequence, such as most of the capsid protein, as described herein (with the cyclization sequence remaining; see below). In other examples, sequences encoding different proteins (e.g., prM, E, NS1, NS3, and/or NS5; see below) or combinations of proteins (e.g., prM-E or C-prM-E) are deleted. This type of deletion may be advantageous for use of the PIV as a vector to deliver a heterologous immunogen, as the deletion can permit insertion of sequences that may be, for example, at least up to the size of the deleted sequence. In other examples, the mutation can be, for example, a point mutation, provided that it results in replication deficiency, as discussed above. Because of the deletion or mutation, the genome does not encode all proteins necessary to produce a full flavivirus particle. The missing sequences can be provided *in trans* by a complementing cell line that is engineered to express the missing sequence (e.g., by use of a replicon; s-PIV; see below), or by co-expression of two replication-deficient genomes in the same cell, where the two replication-deficient genomes, when considered together, encode all proteins necessary for production (d-PIV system; see below).

Upon introduction into cells that do not express complementing proteins, the genomes replicate and, in some instances, generate "virus-like particles," which are released from the cells and are able to leave the cells and be immunogenic, but cannot infect other cells and lead to the generation of further particles. For example, in the case of a PIV including a deletion in capsid protein encoding sequences, after infection of cells that do not express capsid, VLPs including prM-E proteins are

released from the cells. Because of the lack of capsid protein, the VLPs lack capsid and a nucleic acid genome. In the case of the d-PIV approach, production of further PIVs is possible in cells that are infected with two PIVs that complement each other with respect to the production of all required proteins (see below).

The invention provides several advantages. For example, the PIV vectors and PIVs of the invention are highly attenuated and highly efficacious after one-to-two doses, providing durable immunity. Further, unlike inactivated vaccines, PIVs mimic whole virus infection, which can result in increased efficacy due to the induction of robust B- and T-cell responses, higher durability of immunity, and decreased dose requirements. In addition, similar to whole viruses, PIV vaccines target antigen-presenting cells, such as dendritic cells, stimulate toll-like receptors (TLRs), and induce balanced Th1/Th2 immunity. PIV constructs have also been shown to grow to high titers in substrate cells, with little or no CPE, allowing for high-yield manufacture, optionally employing multiple harvests and/or expansion of infected substrate cells. Further, the PIV vectors of the invention provide an option for developing vaccines against non-flavivirus pathogens, such as RSV, for which no vaccines are currently available.

Other features and advantages of the invention will be apparent from the following detailed description, the drawings, and the claims.

Brief Description of the Drawings

Fig. 1 is a schematic illustration of single component PIV (s-PIV) technology.

Fig. 2 is a schematic illustration of two-component PIV (d-PIV) technology.

Fig. 3 is a schematic illustration of a general experimental design for testing immunogenicity and efficacy of PIVs in mice.

Fig. 4 is a graph comparing the humoral immune response induced by PIV-WN (RV-WN) with that of YF/WN LAV (CV-WN) in mice.

Fig. 5 is a series of graphs showing the results of challenging hamsters immunized with PIV-YF (RV-YF), YF17D, PIV-WN (RV-WN), and YF/WN LAV (CVWN) with hamster-adapted Asibi (PIV-YF and YF17D vaccinees) and wild type WN-NY99 (PIV-WN and YF/WN LAV vaccinees).

Fig. 6 is a table showing YF/TBE and YF/LGT virus titers and plaque morphology obtained with the indicated chimeric flaviviruses.

Fig. 7 is a table showing WN/TBE PIV titers and examples of immunofluorescence of cells containing the indicated PIVs.

Fig. 8 is a set of graphs showing the replication kinetics of YF/TBE LAV and PIV-WN/TBE in Vero and BHK cell lines (CV-Hypr = YF/Hypr LAV; CV-LGT = YF/LGT LAV; RV-WN/TBEV = PIV-WN/TBEV).

Fig. 9 is a series of graphs showing survival of mice inoculated IC with PIV-TBE and YF/TBE LAV constructs in a neurovirulence test (3.5 week old ICR mice; RV-WN/Hypr = PIV-WN/TBE(Hypr); CV-Hypr = YF/TBE(Hypr) LAV; CV-LGT = YF/LGT LAV).

Fig. 10 is a graph showing survival of mice inoculated IP with PIV-WN/TBE(Hypr) (RV-WN/Hypr), YF/TBE(Hypr) LAV (CV-Hypr), and YF/LGT LAV (CV-LGT) constructs and YF17D in a neuroinvasiveness test (3.5 week old ICR mice).

Fig. 11 is a series of graphs illustrating morbidity in mice measured by dynamics of body weight loss after TBE virus challenge, for groups immunized with s-PIV-TBE candidates (upper left panel), YF/TBE and YF/LGT chimeric viruses (upper right panel), and controls (YF 17D, human killed TBE vaccine, and mock; bottom panel).

Fig. 12 is a schematic representation of PIV constructs expressing rabies virus G protein, as well as illustration of packaging of the constructs to make pseudoinfectious virus and immunization.

Fig. 13 is a schematic representation of insertion designs resulting in viable/expressing constructs (exemplified by rabies G).

Fig. 14 is series of images showing immunofluorescence analysis and graphs showing growth curves of cells transfected with the indicated PIV-WN constructs (ΔC-Rabies G, ΔPrM-E-Rabies G, and ΔC-PrM-E-Rabies G).

Fig. 15 is a series of images showing immunofluorescence analysis of RabG expressed on the plasma membranes of Vero cells transfected with the indicated PIV constructs (ΔC-Rabies G, ΔPrM-E-Rabies G, and ΔC-PrM-E-Rabies G).

Fig. 16 is a schematic illustration of a PIV-WN-rabies G construct and a series of images showing that this construct spreads in helper cells, but not in naïve cells.

Fig. 17 is a series of graphs showing stability of the rabies G protein gene in PIV-WN vectors.

Fig. 18 is a set of images showing a comparison of spread of single-component vs. two-component PIV-WN-rabies G variants in Vero cells.

Fig. 19 is a set of immunofluorescence images showing expression of full-length RSV F protein (strain A2) by the ΔprM-E component of d-PIV-WN in helper cells after transfection.

Fig. 20 is a schematic representation of wild-type RSV F and RSV trF.

Fig. 21 is a schematic representation of three PIV(WN)-RSVtrF (A1 strain) constructs: ΔC-RSVtrF sPIV, ΔprME-RSVtrF dPIV helper, and ΔCprME-RSVtrF. Immunofluorescence of helper cells after transfection (Day 4) is also shown.

Fig. 22 is a series of images showing titration of WNΔC-RSV trF PIV in Vero cells visualized by immunostaining.

Fig. 23 is an image showing a Western blot analysis of two ΔprME-RSVtrF stocks, 2 days post infection.

Fig. 24 is an image showing a Western blot analysis of Vero cells infected with the indicated amounts of VP2400, vFP2403, and PIV-F.

Fig. 25 is a set of graphs showing endpoint titers obtained using the indicated constructs and routes of administration in two sets of experiments (RSVi27 and RSVi32) indicating the anti-RSV-F IgG antibody titres obtained by ELISA. "F" represents vector with the F insert (truncated), while "e" represents the empty vector alone. "FI_RSV" is a formalin inactivated RSV virus, while "RSV" is a live RSV virus preparation.

Fig. 26 is a set of graphs showing serum neutralization titers obtained using the indicated constructs and routes of administration in two sets of experiments (RSVi27 and RSVi32). "F" represents vector with the F insert (truncated), while "e" represents the empty vector alone. "FI_RSV" is a formalin inactivated RSV virus, while "RSV" is a live RSV virus preparation (see Fig. 25).

Detailed Description of the Invention

The invention provides replication-defective or deficient pseudoinfectious virus (PIV) vectors including flavivirus sequences, which can be used in methods for inducing immunity against heterologous immunogens inserted into the vectors as well as, optionally, the vectors themselves. The invention also includes compositions including combinations of PIVs and/or PIV vectors, as described herein, and methods of using such compositions to induce immune responses against inserted immunogen sequences and/or sequences of the PIVs themselves. The focus of the invention is PIV vectors containing respiratory syncytial virus (RSV) immunogens, such as F or G protein immunogens, in one embodiment (see, e.g., truncated F protein, below). These vectors can be used in methods to prevent or treat RSV infection, and also in combination methods involving use of, for example, any of the other vectors described herein (such as vectors including immunogens of other pathogens and/or cancer, and/or allergy-related immunogens). The vectors, compositions, and methods of the invention are described further below.

PIV Vectors

The PIV vectors of the invention can be based on the single- or two-component PIVs described above (also see WO 2007/098267 and WO 2008/137163). Thus, for example, in the case of single component PIVs, the PIV vectors and PIVs can include a genome including a large deletion in capsid protein encoding sequences and be produced in a complementing cell line that produces capsid protein *in trans* (single component; Fig. 1 and Fig. 12). According to this approach, most of the capsid-encoding region is deleted, which prevents the PIV genome from producing infectious progeny in normal cell lines (i.e., cell lines not expressing capsid sequences) and vaccinated subjects. The capsid deletion typically does not disrupt RNA sequences required for genome cyclization (i.e., the sequence encoding amino acids in the region of positions 1-26), and/or the prM sequence required for maturation of prM to M. In specific examples, the deleted sequences correspond to those encoding amino acids 26-100, 26-93, 31-100, or 31-93 of the C protein.

Single component PIV vectors and PIVs can be propagated in cell lines that express either C or a C-prM-E cassette, where they replicate to high levels.

Exemplary cell lines that can be used for expression of single component PIV vectors and PIVs include BHK-21 (e.g., ATCC CCL-10), Vero (e.g., ATCC CCL-81), C7/10, and other cells of vertebrate or mosquito origin. The C or C-prM-E cassette can be expressed in such cells by use of a viral vector-derived replicon, such as an alphavirus replicon (e.g., a replicon based on Venezuelan Equine Encephalitis virus (VEEV), Sindbis virus, Semliki Forest virus (SFV), Eastern Equine Encephalitis virus (EEEV), Western Equine Encephalitis virus (WEEV), or Ross River virus). To decrease the possibility of productive recombination between the PIV vectors/PIVs and complementing sequences, the sequences in the replicons (encoding C, prM, and/or E) can include nucleotide mutations. For example, sequences encoding a complementing C protein can include an unnatural cyclization sequence. The mutations can result from codon optimization, which can provide an additional benefit with respect to PIV yield. Further, in the case of complementing cells expressing C protein sequences (and not a C-prM-E cassette), it may be beneficial to include an anchoring sequence at the carboxy terminus of the C protein including, for example, about 20 amino acids of prM (see, e.g., WO 2007/098267).

The PIV vectors and PIVs of the invention can also be based on the two-component genome technology described above. This technology employs two partial genome constructs, each of which is deficient in expression of at least one protein required for productive replication (capsid or prM/E) but, when present in the same cell, result in the production of all components necessary to make a PIV. Thus, in one example of the two-component genome technology, the first component includes a large deletion of C, as described above in reference to single component PIVs, and the second component includes a deletion of prM and E (Fig. 2 and Fig. 12). In another example, the first component includes a deletion of C, prM, and E, and the second component includes a deletion of NS1 (Fig. 12). Both components can include cis-acting promoter elements required for RNA replication and a complete set of non-structural proteins, which form the replicative enzyme complex. Thus, both defective genomes can include a 5'-untranslated region and at least about 60 nucleotides (Element 1) of the following, natural protein-coding sequence, which comprises an amino-terminal fragment of the capsid protein. This sequence can be followed by a protease cleavage sequence such as, for example, a ubiquitine or foot-and-mouth

disease virus (FAMDV)-specific 2A protease sequence, which can be fused with either capsid or envelope (prM-E) coding sequences. Further, artificial, codon optimized sequences can be used to exclude the possibility of recombination between the two defective viral genomes, which could lead to formation of replication-competent viruses (see, e.g., WO 2008/137163). Use of the two-component genome approach does not require the development of cell lines expressing complementing genomes, such as the cells transformed with replicons, as discussed above in reference to the single component PIV approach. Exemplary cell lines that can be used in the two-component genome approach include Vero (e.g., ATCC CCL-81), BHK-21 (e.g., ATCC CCL-10), C7/10, and other cells of vertebrate or mosquito origin.

Additional examples of d-PIV approaches that can be used in the invention are based on use of complementing genomes including deletions in NS3 or NS5 sequences. A deletion in, e.g., NS1, NS3, or NS5 proteins can be used as long as several hundred amino acids in the ORF, removing the entire chosen protein sequence, or as short as 1 amino acid inactivating protein enzymatic activity (e.g., NS5 RNA polymerase activity, NS3 helicase activity, etc.). Alternatively, point amino acid changes (as few as 1 amino acid mutation, or optionally more mutations) can be introduced into any NS protein, inactivating enzymatic activity. In addition, several ΔNS deletions can be combined in one helper molecule. The same heterologous gene (such as an RSV F or G protein (e.g., truncated RSV F protein) gene), i.e., expressed by the first d-PIV component, can be expressed in place or in combination with the NS deletion(s) in the second component, increasing the amount of expressed immunogen. Notably, the insertion capacity of the helper will increase proportionally to the size of NS deletion(s). Alternatively, a different foreign immunogen(s) can be inserted in place of deletion(s) of the helper to produce multivalent vaccines.

Further, additional approaches that can be used in making PIV vectors and PIVs for use in the present invention are described, for example, in WO 99/28487, WO 03/046189, WO 2004/108936, US 2004/0265338, US 2007/0249032, and U.S. Patent No. 7,332,322.

The PIV vectors of the invention can be comprised of sequences from a single flavivirus type (e.g., West Nile, tick-borne encephalitis (TBE, e.g., strain Hypr),

Langat (LGT), yellow fever (e.g., YF17D), Japanese encephalitis, dengue (serotype 1-4), St. Louis encephalitis, Kunjin, Rocio encephalitis, Ilheus, Central European encephalitis, Siberian encephalitis, Russian Spring-Summer encephalitis, Kyasanur Forest Disease, Omsk Hemorrhagic fever, Louping ill, Powassan, Negishi, Absettarov, Hansalova, and Apoi viruses), or can comprise sequences from two or more different flaviviruses. Sequences of some strains of these viruses are readily available from generally accessible sequence databases; sequences of other strains can be easily determined by methods well known in the art. In the case of PIV vectors and PIVs including sequences of more than one flavivirus, the sequences can be those of a chimeric flavivirus, as described above (also see, e.g., U.S. Patent No. 6,962,708; U.S. Patent No. 6,696,281; and U.S. Patent No. 6,184,024). In certain examples, the chimeras include pre-membrane and envelope sequences from one flavivirus (such as a flavivirus to which immunity may be desired), and capsid and non-structural sequences from a second, different flavivirus. In one specific example, the second flavivirus is a yellow fever virus, such as the vaccine strain YF17D. Other examples include the YF/WN, YF/TBE, YF/LGT, WN/TBE, and WN/LGT chimeras described below. Another example is an LGT/TBE chimera based on LGT virus backbone containing TBE virus prM-E proteins. A PIV vaccine based on this genetic background would have an advantage, because LGT replicates very efficiently in vitro and is highly attenuated and immunogenic for humans. Thus, a chimeric LGT/TBE PIV vaccine is expected to provide a robust specific immune response in humans against TBE, particularly due to inclusion of TBE prM-E genes.

Vectors of the invention can be based on PIV constructs or live, attenuated chimeric flaviviruses as described herein (in particular, YF/TBE, YF/LGT, WN/TBE, and WN/LGT; see below). Use of PIV constructs as vectors provides particular advantages in certain circumstances, because these constructs by necessity include large deletions, which render the constructs amenable to accommodation of insertions that are at least up to the size of the deleted sequences, without there being a loss in replication efficiency. Thus, PIV vectors in general can comprise very small insertions (e.g., in the range 6-10, 11-20, 21-100, 101-500, or more amino acid residues combined with the ΔC deletion or other deletions), as well as relatively large insertions or insertions of intermediate size (e.g., in the range 501-1000, 1001-1700,

1701-3000, or 3001-4000 or more residues). In contrast, in certain examples, it may be advantageous to express relatively short sequences in live attenuated viruses, particularly if the insertions are made in the absence of a corresponding deletion. Additional information concerning insertion sites that can be used in the invention is provided below. In addition, as discussed further below, expression of non-flavivirus immunogens in PIVs and chimeric flaviviruses of the invention can result in dual vaccines that elicit protective immunity against both a flavivirus vector virus pathogen and a target heterologous immunogen (e.g., RSV immunogens, such as those described herein).

As discussed above, the PIV vectors and PIVs of the invention can comprise sequences of chimeric flaviviruses, for example, chimeric flaviviruses including pre-membrane and envelope sequences of a first flavivirus (e.g., a flavivirus to which immunity is sought), and capsid and non-structural sequences of a second, different flavivirus, such as a yellow fever virus (e.g., YF17D; see above and also U.S. Patent No. 6,962,708; U.S. Patent No. 6,696,281; and U.S. Patent No. 6,184,024). Further, chimeric flaviviruses (as well as non-chimeric flaviviruses, e.g., West Nile virus) used in the invention, used as a source for constructing PIVs, can optionally include one or more specific attenuating mutations (e.g., E protein mutations, prM protein mutations, deletions in the C protein, and/or deletions in the 3'UTR), such as any of those described in WO 2006/116182. For example, the C protein or 3'UTR deletions can be directly applied to YF/WN, YF/TBE, or YF/LGT chimeras. Similar deletions can be designed and introduced in other chimeric LAV candidates such as based on LGT/TBE, WN/TBE, and WN/LGT genomes. With respect to E protein mutations, attenuating mutations similar to those described for YF/WN chimera in WO 2006/116182 can be designed, e.g., based on the knowledge of crystal structure of the E protein (Rey et al., Nature 375(6529):291-298, 1995), and employed. Further, additional examples of attenuating E protein mutations described for TBE virus and other flaviviruses are provided in Table 9. These can be similarly introduced into chimeric vaccine candidates.

The invention also provides new, particular chimeric flaviviruses, which can be used as a basis for the design of PIV vectors and PIVs, and as live attenuated chimeric flavivirus vectors. These chimeras include tick-borne encephalitis (TBE)

virus or related prM-E sequences. Thus, the chimeras can include prM-E sequences from, for example, the Hypr strain of TBE or Langat (LGT) virus. Capsid and non-structural proteins of the chimeras can include those from yellow fever virus (e.g., YF17D) or West Nile virus (e.g., NY99).

A central feature of these exemplary YF/TBE, YF/LGT, WN/TBE, and WN/LGT chimeras is the signal sequence between the capsid and prM proteins. As is shown in the Examples, below, we have found that, in the case of YF-based PIV chimeras, it is advantageous to use a signal sequence comprising yellow fever and TBE sequences (see below). In one example, the signal sequence includes yellow fever sequences in the amino terminal region (e.g., SHDVLTVQFLIL) and TBE sequences in the carboxy terminal region (e.g., GMLGMTIA), resulting in the sequence SHDVLTVQFLILGMLGMTIA. We have also found that, in the case of WN-based PIV chimeras, it is advantageous to use a signal sequence comprising TBE sequences (e.g., GGTDWMSWLLVIGMLGMTIA). The invention thus includes YF/TBE, YF/LGT, WN/TBE, and WN/LGT chimeras, both PIVs and LAVs, which include the above-noted signal sequences, or variants thereof having, e.g., 1-8, 2-7, 3-6, or 4-5 amino acid substitutions, deletions, or insertions, which do not substantially interfere with processing at the signal sequence. In various examples, the substitutions are "conservative substitutions," which are characterized by replacement of one amino acid residue with another, biologically similar residue. Examples of conservative substitutions include the substitution of one hydrophobic residue such as isoleucine, valine, leucine, or methionine for another, or the substitution of one polar residue for another, such as between arginine and lysine, between glutamic and aspartic acids, or between glutamine and asparagine and the like. Additional information concerning these chimeras is provided below, in the Examples.

Insertion Sites

Sequences encoding immunogens can be inserted at one or more different sites within the vectors of the invention. Relatively short peptides can be delivered on the surface of PIV or LAV glycoproteins (e.g., prM, E, and/or NS1 proteins) and/or in the context of other proteins (to induce predominantly B-cell and T-cell responses, respectively). Other inserts, including larger portions of foreign proteins (e.g., certain

RSV F or G protein sequences, as described herein), as well as complete proteins, can be expressed intergenically, at the N- and C-termini of the polyprotein, or bicistronically (e.g., within the ORF under an IRES or in the 3'UTR under an IRES; see, e.g., WO 02/102828, WO 2008/036146, WO 2008/094674, WO 2008/100464, WO 2008/115314, and below for further details). In PIV constructs, there is an additional option of inserting a foreign amino acid sequence directly in place of introduced deletion(s). Insertions can be made in, for example, ΔC, ΔprM-E, ΔC-prM-E, ΔNS1, ΔNS3, and ΔNS5. Thus, in one example, in the case of s-PIVs and the ΔC component of d-PIVs, immunogen-encoding sequences can be inserted in place of deleted capsid sequences. Immunogen-encoding sequences can also, optionally, be inserted in place of deleted prM-E sequences in the ΔprM-E component of d-PIVs. In another example, the sequences are inserted in place of or combined with deleted sequences in ΔC-prM-E constructs. Examples of such insertions are provided in the Examples section, below.

In the case of making insertions into PIV deletions, the insertions can be made with a few (e.g., 1, 2, 3, 4, or 5) additional vector-specific residues at the N- and/or C-termini of the foreign immunogen, if the sequence is simply fused in-frame (e.g., ~ 20 first a.a. and a few last residues of the C protein if the sequence replaces the ΔC deletion), or without, if the foreign immunogen is flanked by appropriate elements well known in the field (e.g., viral protease cleavage sites; cellular protease cleavage sites, such as signalase, furin, etc.; autoprotease; termination codon; and/or IRES elements).

If a protein is expressed outside of the continuous viral open reading frame (ORF), e.g., if vector and non-vector sequences are separated by an internal ribosome entry site (IRES), cytoplasmic expression of the product can be achieved or the product can be directed towards the secretory pathway by using appropriate signal/anchor segments, as desired. If the protein is expressed within the vector ORF, important considerations include cleavage of the foreign protein from the nascent polyprotein sequence, and maintaining correct topology of the foreign protein and all viral proteins (to ensure vector viability) relative to the ER membrane, e.g.,

translocation of secreted proteins into the ER lumen, or keeping cytoplasmic proteins or membrane-associated proteins in the cytoplasm/in association with the ER membrane.

In more detail, the above-described approaches to making insertions can employ the use of, for instance, appropriate vector-derived, insert-derived, or unrelated signal and anchor sequencess included at the N and C termini of glycoprotein inserts. Standard autoproteases, such as FMDV 2A autoprotease (~20 amino acids) or ubiquitin (gene ~ 500 nt), or flanking viral NS2B/NS3 protease cleavage sites can be used to direct cleavage of an expressed product from a growing polypeptide chain, to release a foreign protein from a vector polyprotein, and to ensure viability of the construct. Optionally, growth of the polyprotein chain can be terminated by using a termination codon, e.g., following a foreign gene insert, and synthesis of the remaining proteins in the constructs can be re-initiated by incorporation of an IRES element, e.g., the encephalomyocarditis virus (EMCV) IRES commonly used in the field of RNA virus vectors. Viable recombinants can be recovered from helper cells (or regular cells for d-PIV versions). Optionally, backbone PIV sequences can be rearranged, e.g., if the latter results in more efficient expression of a foreign gene. For example, a gene rearrangement has been applied to TBE virus, in which the prM-E genes were moved to the 3' end of the genome under the control of an IRES (Orlinger et al., J. Virol. 80:12197-12208, 2006). Translocation of prM-E or any other genes can be applied to PIV flavivirus vaccine candidates and expression vectors, according to the invention.

Additional details concerning different insertion sites that can be used in the invention are as follows (also see WO 02/102828, WO 2008/036146, WO 2008/094674, WO 2008/100464, WO 2008/115314, as noted above). Peptide sequences can be inserted within the envelope protein, which is the principle target for neutralizing antibodies. The sequences can be inserted into the envelope in, for example, positions corresponding to amino acid positions 59, 207, 231, 277, 287, 340, and/or 436 of the Japanese encephalitis virus envelope protein (see, e.g., WO 2008/115314 and WO 02/102828). To identify the corresponding loci in different flaviviruses, the flavivirus sequences are aligned with that of Japanese encephalitis virus. As there may not be an exact match, it should be understood that, in non-JE

viruses, the site of insertion may vary by, for example, 1, 2, 3, 4, or 5 amino acids, in either direction. Further, given the identification of such sites as being permissive in JE, they can also vary in JE by, for example, 1, 2, 3, 4, or 5 amino acids, in either direction. Additional permissive sites can be identified using methods such as transposon mutagenesis (see, e.g., WO 02/102828 and WO 2008/036146). The insertions can be made at the indicated amino acids by insertion just C-terminal to the indicated amino acids (i.e., between amino acids 51-52, 207-208, 231-232, 277-278, 287-288, 340-341, and 436-437), or in place of short deletions (e.g., deletions of 1, 2, 3, 4, 5, 6, 7, or 8 amino acids) beginning at the indicated amino acids (or within 1-5 positions thereof, in either direction).

In addition to the envelope protein, insertions can be made into other virus proteins including, for example, the membrane/pre-membrane protein and NS1 (see, e.g., WO 2008/036146). For example, insertions can be made into a sequence preceding the capsid/pre-membrane cleavage site (at, e.g., -4, -2, or –1) or within the first 50 amino acids of the pre-membrane protein (e.g., at position 26), and/or between amino acids 236 and 237 of NS1 (or in regions surrounding the indicated sequences, as described above). In other examples, insertions can be made intergenically. For example, an insertion can be made between E and NS1 proteins and/or between NS2B and NS3 proteins (see, e.g., WO 2008/100464). In one example of an intergenic insertion, the inserted sequence can be fused with the C-terminus of the E protein of the vector, after the C-terminal signal/anchor sequence of the E protein, and the insertion can include a C-terminal anchor/signal sequence, which is fused with vector NS1 sequences. In another example of an intergenic insertion, the inserted sequences, with flanking protease cleavage sites (e.g., YF 17D cleavage sites), can be inserted into a unique restriction site introduced at the NS2B/NS3 junction (WO 2008/100464).

In other examples, a sequence can be inserted in the context of an internal ribosome entry site (IRES, e.g., an IRES derived from encephalomyocarditis virus; EMCV), as noted above, such as inserted in the 3'-untranslated region (WO 2008/094674). In one example of such a vector, employing, for example, yellow fever virus sequences, an IRES-immunogen cassette can be inserted into a multiple cloning site engineered into the 3'-untranslated region of the vector, e.g., in a deletion

(e.g., a 136 nucleotide deletion in the case of a yellow fever virus-based example) after the polyprotein stop codon (WO 2008/094674).

Details concerning the insertion of rabies virus G protein and respiratory syncytial virus (RSV) F protein (including truncated F) into s-PIV and d-PIV vectors of the invention are provided below in Example 3. The information provided in Example 3 can be applied in the context of other vectors and immunogens described herein.

Immunogens

PIVs (s-PIVs and d-PIVs) based on flavivirus sequences and live, attenuated chimeric flaviviruses (e.g., YF/WN, YF/TBE, YF/LGT, WN/TBE, and WN/LGT), as described above, can be used in the invention to deliver foreign (e.g., non-flavivirus) pathogen immunogens. The focus of the invention is the delivery of RSV immunogens, such as RSV fusion or F protein (or RSV G) immunogens (e.g., truncated F proteins; see below, for example the truncated F protein sequence in Example 3). PIVs and chimeric flavivirus vectors delivering a particular RSV immunogen can, optionally, be delivered with vectors delivering one or more other RSV immunogens, or one or more immunogens from another pathogen (e.g., viral, bacterial, fungal, and parasitic pathogens), one or more immunogens from cancer, and/or allergy-related immunogens. Specific, non-limiting examples of immunogens that can be delivered according to the invention are provided as follows.

As noted above, a central focus of the invention is delivery of the RSV proteins such as, in one embodiment, the RSV fusion or F glycoprotein and, in particular, truncated forms of this protein. The RSV F glycoprotein is one of the major immunogenic proteins of the virus. It is an envelope glycoprotein that mediates both fusion of the virus to the host cell membrane, and cell-to-cell spread of the virus. The amino acid sequence of the F protein is highly conserved among RSV subgroups A and B and is a cross-protective antigen.

RSV F protein comprises an extracellular region, a trans-membrane region, and a cytoplasmic tail region. A truncated protein delivered according to the invention can be, for example, one in which the trans-membrane and cytoplasmic tail

regions of the F protein are absent (see, e.g., Example 3, below). Lack of expression of the trans-membrane region results in a secreted form of the RSV protein.

RSV F protein, as used herein, includes both full-length and truncated RSV fusion proteins, which may have the sequences described herein, or have variations in their amino acid sequences including naturally occurring in various strains of RSV and those introduced by PCR amplification of the encoding gene while retaining the immunogenic properties, a secreted form of the RSV F protein lacking a trans-membrane anchor and cytoplasmic tail, as well as fragments capable of generating antibodies which specifically react with RSV F protein and functional analogs. A first protein is a functional analog of a second protein if the first protein is immunologically related to and/or has the same function as the second protein. It may be for example, a fragment of the protein, or a substitution, addition, or deletion mutant thereof. The RSV F glycoprotein can be from, e.g., subgroup A or B (Wertz et al., Biotechnology 20:151-176, 1992).

In a further embodiment of the present invention, RSV G glycoprotein can be delivered. The G protein is a approximately 33 kDa protein and is heavily O-glycosylated, giving rise to a glycoprotein having a molecular weight of about 90 kDa (Levine, S., Kleiber-France, R., and Paradiso, P.R. (1987) J. Gen. Virol. 69, 2521-2524). The 298 amino acid residue RSV G protein belongs to the type II glycoproteins with the transmembrane domain (TM) located near the N-terminus (putative location: residues 38 to 66 underlined in Sequence Appendix 7. The RSV F and G proteins, or fragments or analogs thereof, can be from, for example, group A (e.g., A1 or A2) or B RSV.

Other examples of immunogens that can be delivered according to the invention are protective immunogens of the causative agent of Lyme disease (tick-borne spirochete *Borrelia burgdorferi*). In one example, PIVs including TBE/LGT sequences, as well as chimeric flaviviruses including TBE sequences (e.g., YF/TBE, YF/LGT, WN/TBE, LGT/TBE, and WN/LGT; in all instances where "TBE" is indicated, this includes the option of using the Hypr strain), can be used as vectors to deliver these immunogens. This combination, targeting both infectious agents (TBE and *B. burgdorferi*) is advantageous, because TBE and Lyme disease are both tick-borne diseases. The PIV approaches can be applied to chimeras (e.g., YF/TBE,

YF/LGT, WN/TBE, or WN/LGT), according to the invention, as well as to non-chimeric TBE and LGT viruses. An exemplary immunogen from *B. burgdorferi* that can be used in the invention is OspA (Gipson et al., Vaccine 21:3875-3884, 2003). Optionally, to increase safety and/or immunogenicity, OspA can be mutated to reduce chances of autoimmune responses and/or to eliminate sites for unwanted post-translational modification in vertebrate animal cells, such as N-linked glycosylation, which may affect immunogenicity of the expression product. Mutations that decrease autoimmunity can include, e.g., those described by Willett et al., Proc. Natl. Acad. Sci. U.S.A. 101:1303-1308, 2004. In one example, FTK-OspA, a putative cross-reactive T cell epitope, *Bb* $OspA_{165-173}$ (YVLEGTLTA) is altered to resemble the corresponding peptide sequence of *Borrelia afzelli* (FTLEGKVAN). In FTK-OspA, the corresponding sequence is FTLEGKLTA.

The sequence of OspA is as follows:

```
      1 mkkyllgigl ilaliackqn vssldeknsv svdlpgemkv lvskeknkdg
kydliatvdk
     61 lelkgtsdkn ngsgvlegvk adkskvklti sddlgqttle vfkedgktlv
skkvtskdks
    121 steekfnekg evsekiitra dgtrleytgi ksdgsgkake vlkgyvlegt
ltaekttlvv
    181 kegtvtlskn isksgevsve lndtdssaat kktaawnsgt stltitvnsk
ktkdlvftke
    241 ntitvqqyds ngtklegsav eitkldeikn alk
```

The full-length sequence and/or immunogenic fragments of the full-length sequence can be used. Exemplary fragments can include one or more of domains 1 (amino acids 34-41), 2 (amino acids 65-75), 3 (amino acids 190-220), and 4 (amino acids 250-270) (Jiang et al., Clin. Diag. Lab. Immun. 1(4):406-412, 1994). Thus, for example, a peptide comprising any one (or more) of the following sequences (which include sequence variations that can be included in the sequence listed above, in any combination) can be delivered: LPGE/GM/IK/T/GVL; GTSDKN/S/DNGSGV/T; N/H/EIS/P/L/A/SK/NSGEV/IS/TV/AE/ALN/DDT/SD/NS/TS/TA/Q/ RATKKTA/GA/K/TWN/DS/AG/N/KT; SN/AGTK/NLEGS/N/K/TAVEIT/KK/TLD/KEI/LKN.

In addition to *B. burgdorferi* immunogens, tick saliva proteins, such as 64TRP, Isac, and Salp20, can be expressed, e.g., to generate a vaccine candidate of trivalent-specificity (TBE+Lyme disease+ticks). Alternatively, tick saliva proteins can be expressed instead of *B. burgdorferi* immunogens in TBE sequence-containing vectors.

In addition, there are many other candidate tick saliva proteins that can be used for tick vector vaccine development according to the invention (Francischetti et al., Insect Biochem. Mol. Biol. 35:1142-1161, 2005). One or more of these immunogens can be expressed in s-PIV-TBE. However, d-PIV-TBE may also be selected, because of its large insertion capacity. In addition to PIV-TBE, other PIV vaccines can be used as vectors, e.g., to protect from Lyme disease and another flavivirus disease, such as West Nile virus. Expression of these immunogens can be evaluated in cell culture, and immunogenicity/protection examined in available animal models (e.g., as described in Gipson et al., Vaccine 21:3875-3884, 2003; Labuda et al., Pathog. 2(e27):0251-0259, 2006). Immunogens of other pathogens can be similarly expressed, in addition to Lyme disease and tick immunogens, with the purpose of making multivalent vaccine candidates. Exemplary tick saliva immunogens that can be used in the invention include the following:

64TRP (AF469170)

MKAFFVLSLL STAALTNAAR AGRLGSDLDT FGRVHGNLYA GIERAGPRGY PGLTASIGGE
VGARLGGRAG VGVSSYGYGY PSWGYPYGGY GGYGGYGGYG GYDQGFGSAY GGYPGYYGYY
YPSGYGGGYG GSYGGSYGGS YTYPNVRASA GAAA

Isac (AF270496)

MRTAFTCALL AISFLGSPCS SSEDGLEQDT IVETTTQNLY ERHYRNHSGL CGAQYRNSSH
AEAVYNCTLN HLPPVVNATW EGIRHRINKT IPQFVKLICN FTVAMPQEFY LVYMGSDGNS
DFEEDKESTG TDEDSNTGSS AAAKVTEALI IEAEENCTAH ITGWTTETPT TLEPTTESQF
EAIP

Salp20 (EU008559)

MRTALTCALL AISFLGSPCS SSEGGLEKDS RVETTTQNLY ERYYRKHPGL CGAQYRNSSH
AEAVYNCTLS LLPLSVNTTW EGIRHRINKT IPEFVNLICN FTVAMPDQFY LVYMGSNGNS
YSEEDEDGKT GSSAAVQVTE QLIIQAEENC TAHITGWTTE APTTLEPTTE TQFEAIS

Additional details concerning the TBE-related PIVs and LAVs are provided in Example 2, below.

Other PIV and LAV-vectored vaccines against other non-flavivirus pathogens, including vaccines having dual action, eliciting protective immunity against both flavivirus (as specified by the vector envelope proteins) and non-flavivirus pathogens (as specified by expressed immunologic determinant(s)) can also be used. These are similar to the example of PIV-TBE-Lyme disease-tick vector vaccines described

above. As mentioned above, such dual-action vaccines can be developed against a broad range of pathogens by expression of immunogens from, for example, viral, bacterial, fungal, and parasitic pathogens, and immunogens associated with cancer and allergy. As specific non-limiting examples, we describe herein the design and biological properties of PIV vectored-rabies and -respiratory syncytial virus (RSV) vaccine candidates constructed by expression of rabies virus G protein or RSV F protein in place of or in combination with various deletions in one- and two-component PIV vectors (see Example 3, below).

As is demonstrated in the Examples, below, s-PIV constructs may be advantageously used to stably deliver relatively short foreign immunogens (similar to Lyme disease agent OspA protein and tick saliva proteins), because insertions are combined with a relatively short ΔC deletion. Two-component PIV vectors may be advantageously used to stably express relatively large immunogens, such as rabies G protein and RSV F, as the insertions in such vectors are combined with, for example, large ΔprM-E, ΔC-prM-E, and/or ΔNS1 deletions. Some of the d-PIV components can be manufactured and used as vaccines individually, for instance, the PIV-RSV F construct described below containing a ΔC-prM-E deletion. In this case, the vaccine induces an immune response (e.g., neutralizing antibodies) predominantly against the expressed protein, but not against the flavivirus vector virus pathogen. In other examples of the invention, dual immunity is obtained by having immunity induced both to vector and insert components. Additionally, because of the large insertion capacity of PIV vectors, and the option of using two-component genomes, PIV vectors offer the opportunity to target several non-flavivirus pathogens simultaneously, e.g., by expressing foreign immunogens from two different non-flavivirus pathogens in the two components of a d-PIV.

In addition to the RSV F or G protein, rabies G protein, Lyme disease protective immunogens, and tick saliva proteins, as examples of foreign immunogens described above, other foreign immunogens can be expressed to target respective diseases including, for example, influenza virus type A and B immunogens. In these examples, a few short epitopes and/or whole genes of viral particle proteins can be used, such as the M2, HA, and NA genes of influenza A, and/or the NB or BM2 genes of influenza B. Shorter fragments of M2, NB, and BM2, corresponding for instance

to M2e, the extracellular fragment of M2, can also be used. In addition, fragments of the HA gene, including epitopes identified as HA0 (23 amino acids in length, corresponding to the cleavage site in HA) can be used. Specific examples of influenza-related sequences that can be used in the invention include PAKLLKERGFFGAIAGFLE (HA0), PAKLLKERGFFGAIAGFLEGSGC (HA0), NNATFNYTNVNPISHIRGS (NBe), MSLLTEVETPIRNEWGCRCNDSSD (M2e), MSLLTEVETPTRNEWECRCSDSSD (M2e), MSLLTEVETLTRNGWGCRCSDSSD (M2e), EVETPTRN (M2e), SLLTEVETPIRNEWGCRCNDSSD (M2e), and SLLTEVETPIRNEWGCR (M2e). Additional M2e sequences that can be used in the invention include sequences from the extracellular domain of BM2 protein of influenza B (consensus MLEPFQ, e.g., LEPFQILSISGC), and the M2e peptide from the H5N1 avian flu (MSLLTEVETLTRNGWGCRCSDSSD).

Other examples of pathogen immunogens that can be delivered in the vectors of the invention include codon-optimized SIV or HIV gag (55 kDa), gp120, gp160, SIV mac239-rev/tat/nef genes or analogs from HIV, and other HIV immunogens; immunogens from HPV viruses, such as HPV16, HPV18, etc., e.g., the capsid protein L1 which self-assembles into HPV-like particles, the capsid protein L2 or its immunodominant portions (e.g., amino acids 1-200, 1-88, or 17-36), the E6 and E7 proteins which are involved in transforming and immortalizing mammalian cells fused together and appropriately mutated (fusion of the two genes creates a fusion protein, referred to as E6E7Rb$^-$, that is about 10-fold less capable of transforming fibroblasts, and mutations of the E7 component at 2 residues renders the resulting fusion protein mutant incapable of inducing transformation (Boursnell et al., Vaccine 14:1485-1494, 1996). Other immunogens include protective immunogens from HCV, CMV, HSV2, viruses, malaria parasite, *Mycobacterium tuberculosis* causing tuberculosis, *C. difficile*, and other nosocomial infections, that are known in the art, as well as fungal pathogens, cancer immunogens, and proteins associated with allergy that can be used as vaccine targets.

Foreign immunogen inserts of the invention, such as RSV immunogens as described herein, can be modified in various ways. For instance, codon optimization is used to increase the level of expression and eliminate long repeats in nucleotide

sequences to increase insert stability in the RNA genome of PIV vectors. Further, the genes can be truncated at N- and/or C-termini, or by internal deletion(s), or modified by specific amino acid changes to increase visibility to the immune system and immunogenicity. Immunogenicity can be increased by chimerization of proteins with immunostimulatory moieties well known in the art, such as TLR agonists, stimulatory cytokines, components of complement, heat-shock proteins, etc. (e.g., reviewed in "Immunopotentiators in Modern Vaccines," Schijns and O'Hagan Eds., 2006, Elsevier Academic Press: Amsterdam, Boston).

With respect to construction of dual vaccines against rabies and other flavivirus diseases, other combinations, such as TBE + rabies, YF + rabies, etc., can be of interest both for human and veterinary use in corresponding geographical regions, and thus can be similarly generated. Possible designs of expression constructs are not limited to those described herein. For example deletions and insertions can be modified, genetic elements can be rearranged, or other genetic elements (e.g. non-flavivirus, non-rabies signals for secretion, intracellular transport determinants, inclusion of or fusion with immunostimulatory moieties such as cytokines, TLR agonists such as flagellin, multimerization components such as leucine zipper, and peptides that increase the period of protein circulation in the blood) can be used to facilitate antigen presentation and increase immunogenicity. Further, such designs can be applied to s-PIV and d-PIV vaccine candidates based on vector genomes of other flaviviruses, and expressing immunogens of other pathogens, e.g., including but not limited to pathogens described in elsewhere herein.

Other examples of PIV and LAV vectors of the invention including combination vaccines such as DEN+Chikungunya virus (CHIKV) and YF+CHIKV. CHIKV, an alphavirus, is endemic in Africa, South East Asia, Indian subcontinent and the Islands, and the Pacific Islands and shares ecological/geographical niches with YF and DEN1-4. It causes serious disease primarily associated with severe pain (arthritis, other symptoms similar to DEN) and long-lasting sequelae in the majority of patients (Simon et al., Med. Clin. North Am. 92:1323-1343, 2008; Seneviratne et al., J. Travel Med. 14:320-325, 2007). Other examples of PIV and LAV vectors of the invention include YF+Ebola or DEN+Ebola, which co-circulate in Africa.

Immunogens for the above-noted non-flavivirus pathogens, sequences of which are well known in the art, may include glycoprotein B or a pp65/IE1 fusion protein of CMV (Reap et al., Vaccine 25(42):7441-7449, 2007; and references therein), several TB proteins (reviewed in Skeiky et al., Nat. Rev. Microbiol. 4(6):469-476, 2006), malaria parasite antigens such as RTS,S (a pre-erythrocytic circumsporozoite protein, CSP) and others (e.g., reviewed in Li et al., Vaccine 25(14):2567-2574, 2007), CHIKV envelope proteins E1 and E2 (or the C-E2-E1, E2-E1 cassettes), HCV structural proteins C-E1-E2 forming VLPs (Ezelle et al., J. Virol. 76(23):12325-12334, 2002) or other proteins to induce T-cell responses, Ebola virus glycoprotein GP (Yang et al., Virology 377(2):255-264, 2008).

In addition to the immunogens described above, the vectors described herein may include one or more immunogen(s) derived from or that direct an immune response against one or more viruses (e.g., viral target antigen(s)) including, for example, a dsDNA virus (e.g., adenovirus, herpesvirus, epstein-barr virus, herpes simplex type 1, herpes simplex type 2, human herpes virus simplex type 8, human cytomegalovirus, varicella-zoster virus, poxvirus); ssDNA virus (e.g., parvovirus, papillomavirus (e.g., E1, E2, E3, E4, E5, E6, E7, E8, BPV1, BPV2, BPV3, BPV4, BPV5, and BPV6 (*In* Papillomavirus and Human Cancer, edited by H. Pfister (CRC Press, Inc. 1990)); Lancaster et al., Cancer Metast. Rev. pp. 6653-6664, 1987; Pfister et al., Adv. Cancer Res. 48:113-147, 1987)); dsRNA viruses (e.g., reovirus); (+)ssRNA viruses (e.g., picornavirus, coxsackie virus, hepatitis A virus, poliovirus, togavirus, rubella virus, flavivirus, hepatitis C virus, yellow fever virus, dengue virus, west Nile virus); (-)ssRNA viruses (e.g., orthomyxovirus, influenza virus, rhabdovirus, paramyxovirus, measles virus, mumps virus, parainfluenza virus, rhabdovirus, rabies virus); ssRNA-RT viruses (e.g., retrovirus, human immunodeficiency virus (HIV)); and dsDNA-RT viruses (e.g. hepadnavirus, hepatitis B). Immunogens may also be derived from other viruses not listed above but available to those of skill in the art.

With respect to HIV, immunogens may be selected from any HIV isolate. As is well-known in the art, HIV isolates are now classified into discrete genetic subtypes. HIV-1 is known to comprise at least ten subtypes (A, B, C, D, E, F, G, H, J, and K). HIV-2 is known to include at least five subtypes (A, B, C, D, and E).

Subtype B has been associated with the HIV epidemic in homosexual men and intravenous drug users worldwide. Most HIV-1 immunogens, laboratory adapted isolates, reagents and mapped epitopes belong to subtype B. In sub-Saharan Africa, India, and China, areas where the incidence of new HIV infections is high, HIV-1 subtype B accounts for only a small minority of infections, and subtype HIV-1 C appears to be the most common infecting subtype. Thus, in certain embodiments, it may be desirable to select immunogens from HIV-1 subtypes B and/or C. It may be desirable to include immunogens from multiple HIV subtypes (e.g., HIV-1 subtypes B and C, HIV-2 subtypes A and B, or a combination of HIV-1 and HIV-2 subtypes) in a single immunological composition. Suitable HIV immunogens include ENV, GAG, POL, NEF, as well as variants, derivatives, and fusion proteins thereof, for example.

Immunogens may also be derived from or direct an immune response against one or more bacterial species (spp.) (e.g., bacterial target antigen(s)) including, for example, Bacillus spp. (e.g., Bacillus anthracis), Bordetella spp. (e.g., Bordetella pertussis), Borrelia spp. (e.g., Borrelia burgdorferi), Brucella spp. (e.g., Brucella abortus, Brucella canis, Brucella melitensis, Brucella suis), Campylobacter spp. (e.g., Campylobacter jejuni), Chlamydia spp. (e.g., Chlamydia pneumoniae, Chlamydia psittaci, Chlamydia trachomatis), Clostridium spp. (e.g., Clostridium botulinum, Clostridium difficile, Clostridium perfringens, Clostridium tetani), Corynebacterium spp. (e.g., Corynebacterium diptheriae), Enterococcus spp. (e.g., Enterococcus faecalis, enterococcus faecum), Escherichia spp. (e.g., Escherichia coli), Francisella spp. (e.g., Francisella tularensis), Haemophilus spp. (e.g., Haemophilus influenza), Helicobacter spp. (e.g., Helicobacter pylori), Legionella spp. (e.g., Legionella pneumophila), Leptospira spp. (e.g., Leptospira interrogans), Listeria spp. (e.g., Listeria monocytogenes), Mycobacterium spp. (e.g., Mycobacterium leprae, Mycobacterium tuberculosis), Mycoplasma spp. (e.g., Mycoplasma pneumoniae), Neisseria spp. (e.g., Neisseria gonorrhea, Neisseria meningitidis), Pseudomonas spp. (e.g., Pseudomonas aeruginosa), Rickettsia spp. (e.g., Rickettsia rickettsii), Salmonella spp. (e.g., Salmonella typhi, Salmonella typhinurium), Shigella spp. (e.g., Shigella sonnei), Staphylococcus spp. (e.g., Staphylococcus aureus, Staphylococcus epidermidis, Staphylococcus saprophyticus, coagulase negative staphylococcus (e.g., U.S. Patent No. 7,473,762)), Streptococcus spp. (e.g., Streptococcus agalactiae,

Streptococcus pneumoniae, Streptococcus pyrogenes), Treponema spp. (e.g., Treponema pallidum), Vibrio spp. (e.g., Vibrio cholerae), and Yersinia spp. (Yersinia pestis). Immunogens may also be derived from or direct the immune response against other bacterial species not listed above but available to those of skill in the art.

Immunogens may also be derived from or direct an immune response against one or more parasitic organisms (spp.) (e.g., parasite target antigen(s)) including, for example, Ancylostoma spp. (e.g., A. duodenale), Anisakis spp., Ascaris lumbricoides, Balantidium coli, Cestoda spp., Cimicidae spp., Clonorchis sinensis, Dicrocoelium dendriticum, Dicrocoelium hospes, Diphyllobothrium latum, Dracunculus spp., Echinococcus spp. (e.g., E. granulosus, E. multilocularis), Entamoeba histolytica, Enterobius vermicularis, Fasciola spp. (e.g., F. hepatica, F. magna, F. gigantica, F. jacksoni), Fasciolopsis buski, Giardia spp. (Giardia lamblia), Gnathostoma spp., Hymenolepis spp. (e.g., H. nana, H. diminuta), Leishmania spp., Loa loa, Metorchis spp. (M. conjunctus, M. albidus), Necator americanus, Oestroidea spp. (e.g., botfly), Onchocercidae spp., Opisthorchis spp. (e.g., O. viverrini, O. felineus, O. guayaquilensis, and O. noverca), Plasmodium spp. (e.g., P. falciparum), Protofasciola robusta, Parafasciolopsis fasciomorphae, Paragonimus westermani, Schistosoma spp. (e.g., S. mansoni, S. japonicum, S. mekongi, S. haematobium), Spirometra erinaceieuropaei, Strongyloides stercoralis, Taenia spp. (e.g., T. saginata, T. solium), Toxocara spp. (e.g., T. canis, T. cati), Toxoplasma spp. (e.g., T. gondii), Trichobilharzia regenti, Trichinella spiralis, Trichuris trichiura, Trombiculidae spp., Trypanosoma spp., Tunga penetrans, and/or Wuchereria bancrofti. Immunogens may also be derived from or direct the immune response against other parasitic organisms not listed above but available to those of skill in the art.

Immunogens may be derived from or direct the immune response against tumor target antigens (e.g., tumor target antigens). The term tumor target antigen (TA) may include both tumor-associated antigens (TAAs) and tumor-specific antigens (TSAs), where a cancerous cell is the source of the antigen. A TA may be an antigen that is expressed on the surface of a tumor cell in higher amounts than is observed on normal cells or an antigen that is expressed on normal cells during fetal development. A TSA is typically an antigen that is unique to tumor cells and is not expressed on normal cells. TAs are typically classified into five categories according to their

expression pattern, function, or genetic origin: cancer-testis (CT) antigens (i.e., MAGE, NY-ESO-1); melanocyte differentiation antigens (e.g., Melan A/MART-1, tyrosinase, gp100); mutational antigens (e.g., MUM-1, p53, CDK-4); overexpressed 'self' antigens (e.g., HER-2/neu, p53); and viral antigens (e.g., HPV, EBV). Suitable TAs include, for example, gp100 (Cox et al., Science 264:716-719, 1994), MART-1/Melan A (Kawakami et al., J. Exp. Med., 180:347-352, 1994), gp75 (TRP-1) (Wang et al., J. Exp. Med., 186:1131-1140, 1996), tyrosinase (Wolfel et al., Eur. J. Immunol., 24:759-764, 1994), NY-ESO-1 (WO 98/14464; WO 99/18206), melanoma proteoglycan (Hellstrom et al., J. Immunol., 130:1467-1472, 1983), MAGE family antigens (e.gl, MAGE-1, 2, 3, 4, 6, and 12; Van der Bruggen et al., Science 254:1643-1647, 1991; U.S. Patent No. 6,235,525), BAGE family antigens (Boel et al., Immunity 2:167-175, 1995), GAGE family antigens (e.g., GAGE-1,2; Van den Eynde et al., J. Exp. Med. 182:689-698, 1995; U.S. Patent No. 6,013,765), RAGE family antigens (e.g., RAGE-1; Gaugler et al., Immunogenetics 44:323-330, 1996; U.S. Patent No. 5,939,526), N-acetylglucosaminyltransferase-V (Guilloux et al., J. Exp. Med. 183:1173-1183, 1996), p15 (Robbins et al., J. Immunol. 154:5944-5950, 1995), ß-catenin (Robbins et al., J. Exp. Med., 183:1185-1192, 1996), MUM-1 (Coulie et al., Proc. Natl. Acad. Sci. U.S.A. 92:7976-7980, 1995), cyclin dependent kinase-4 (CDK4) (Wolfel et al., Science 269:1281-1284, 1995), p21-*ras* (Fossum et al., Int. J. Cancer 56:40-45, 1994), BCR-*abl* (Bocchia et al., Blood 85:2680-2684, 1995), p53 (Theobald et al., Proc. Natl. Acad. Sci. U.S.A. 92:11993-11997, 1995), p185 HER2/neu (erb-B1; Fisk et al., J. Exp. Med., 181:2109-2117, 1995), epidermal growth factor receptor (EGFR) (Harris et al., Breast Cancer Res. Treat, 29:1-2, 1994), carcinoembryonic antigens (CEA) (Kwong et al., J. Natl. Cancer Inst., 85:982-990, 1995) U.S. Patent Nos. 5,756,103; 5,274,087; 5,571,710; 6,071,716; 5,698,530; 6,045,802; EP 263933; EP 346710; and EP 784483; carcinoma-associated mutated mucins (e.g., MUC-1 gene products; Jerome et al., J. Immunol., 151:1654-1662, 1993); EBNA gene products of EBV (e.g., EBNA-1; Rickinson et al., Cancer Surveys 13:53-80, 1992); E7, E6 proteins of human papillomavirus (Ressing et al., J. Immunol. 154:5934-5943, 1995); prostate specific antigen (PSA; Xue et al., The Prostate 30:73-78, 1997); prostate specific membrane antigen (PSMA; Israeli et al., Cancer Res. 54:1807-1811, 1994); idiotypic epitopes or antigens, for example,

immunoglobulin idiotypes or T cell receptor idiotypes (Chen et al., J. Immunol. 153:4775-4787, 1994); KSA (U.S. Patent No. 5,348,887), kinesin 2 (Dietz, et al., Biochem. Biophys. Res. Commun. 275(3):731-738, 2000), HIP-55, TGFβ-1 anti-apoptotic factor (Toomey et al., Br. J. Biomed. Sci. 58(3):177-183, 2001), tumor protein D52 (Bryne et al., Genomics 35:523-532, 1996), H1FT, NY-BR-1 (WO 01/47959), NY-BR-62, NY-BR-75, NY-BR-85, NY-BR-87, and NY-BR-96 (Scanlan, M. Serologic and Bioinformatic Approaches to the Identification of Human Tumor Antigens, in *Cancer Vaccines 2000*, Cancer Research Institute, New York, NY), and/or pancreatic cancer antigens (e.g., SEQ ID NOs: 1-288 of U.S. Patent No. 7,473,531). Immunogens may also be derived from or direct the immune response against include TAs not listed above but available to one of skill in the art.

In addition to the specific immunogen sequences listed above, the invention also includes the use of analogs of the sequences. Such analogs include sequences that are, for example, at least 80%, 90%, 95%, or 99% identical to the reference sequences, or fragments thereof. The analogs can include one or more substitutions or deletions, e.g., substitutions of conservative amino acids as described herein. The analogs also include fragments of the reference sequences that include, for example, one or more immunogenic epitopes of the sequences. Further, the analogs include truncations or expansions of the sequences (e.g., insertion of additional/repeat immunodominant/helper epitopes) by, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11-20, etc., amino acids on either or both ends. Truncation may remove immunologically unimportant or interfering sequences, e.g., within known structural/immunologic domains, or between domains; or whole undesired domains can be deleted; such modifications can be in the ranges 21-30, 31-50, 51-100, 101-400, etc. amino acids. The ranges also include, e.g., 20-400, 30-100, and 50-100 amino acids.

Cocktails

The invention also includes compositions including mixtures of two or more PIVs and/or PIV vectors, as described herein. As discussed above, use of such mixtures or cocktails may be particularly advantageous when induction of immunity to more than one immunogen and/or pathogen is desired. This may be useful, for example, in vaccination against different flaviviruses that may be endemic to the

region in which the vaccine recipient resides. This may also be useful in the context of administration of multiple immunogens against the same target.

Non-limiting examples of PIV cocktails included in the invention are those including PIV-JE + PIV-DEN, and PIV-YF + PIV-DEN. In both of these examples, the PIVs for either or both components can be single or dual component PIVs, as described above. In addition, in the case of the PIV-DEN, the PIV can include sequences of just one dengue serotype selected from the group consisting of dengue serotypes 1-4, or the cocktail can include PIVs expressing sequences from two, three, or all four of the serotypes. Further, the TBE/*Borrelia burgdorferi*/tick saliva protein (e.g., 64TRP, Isac, Salp20) vaccines described herein can be based on including the different immunogens within a single PIV or live attenuated flavivirus, or can be based on mixtures of PIVs (or LAVs), which each include one or more of the immunogens. The cocktails of the invention can be formulated as such or can be mixed just prior to administration.

Use, Formulation, and Administration

The invention includes the PIV and LAV vectors, as well as corresponding nucleic acid molecules, pharmaceutical or vaccine compositions, and methods of their use and preparation. The PIV and LAV vectors of the invention can be used, for example, in vaccination methods to induce an immune response to RSV and/or the flavivirus vector, and/or another expressed immunogen, as described herein. These methods can be prophylactic, in which case they are carried out on subjects (e.g., human subjects or other mammalian subjects) not having, but at risk of developing infection or disease caused by RSV or flavivirus and/or a pathogen from which another expressed immunogen is derived. Such methods include instances in which a subject becomes infected by RSV, but is able to ward off the infection and significant symptomatic disease, because of the treatment according to the invention. The methods can also be therapeutic, in which they are carried out on subjects already having an infection by one or more of the relevant pathogens, such as RSV. Such methods include the amelioration of one or more symptoms of the infection, whether partial or complete. Further, the viruses and vectors can be used individually or in combination with one another or other vaccines. The subjects treated according to the

methods of the invention include humans, as well as non-human mammals (e.g., livestock, such as, cattle, pigs, horses, sheep, and goats, and domestic animals, including dogs and cats). Of particular interest with respect to vaccination against RSV are infants and young children, including pre-mature infants, as well as middle aged and elderly people. Thus, for example, human patients age 1 day to five years (e.g., 2 months to 3 years, or 4 months to two years), or age 50 to 65 and above.

Formulation of the PIV and LAV vectors of the invention can be carried out using methods that are standard in the art. Numerous pharmaceutically acceptable solutions for use in vaccine preparation are well known and can readily be adapted for use in the present invention by those of skill in this art (see, e.g., *Remington's Pharmaceutical Sciences* (18th edition), ed. A. Gennaro, 1990, Mack Publishing Co., Easton, PA). In two specific examples, the PIV vectors, PIVs, LAV vectors, and LAVs are formulated in Minimum Essential Medium Earle's Salt (MEME) containing 7.5% lactose and 2.5% human serum albumin or MEME containing 10% sorbitol. However, the PIV and LAV vectors can simply be diluted in a physiologically acceptable solution, such as sterile saline or sterile buffered saline.

The PIV and LAV vectors of the invention can be administered using methods that are well known in the art, and appropriate amounts of the viruses and vectors to be administered can readily be determined by those of skill in the art. What is determined to be an appropriate amount of virus to administer can be determined by consideration of factors such as, e.g., the size and general health of the subject to whom the virus is to be administered. For example, in the case of live, attenuated viruses of the invention, the viruses can be formulated as sterile aqueous solutions containing between 10^2 and 10^8, e.g., 10^3 to 10^7, infectious units (e.g., plaque-forming units or tissue culture infectious doses) in a dose volume of 0.1 to 1.0 ml. PIVs can be administered at similar doses and in similar volumes; PIV titers however are usually measured in, e.g., focus-forming units determined by immunostaining of foci, as these defective constructs tend not to form virus-like plaques. Doses can range between 10^2 and 10^8 FFU and administered in volumes of 0.1 to 1.0 ml.

All viruses and vectors of the invention can be administered by, for example, intradermal, subcutaneous, intramuscular, intraperitoneal, intranasal (e.g., by inhalation or nose drops), intravenous, or oral routes. In specific examples, dendritic

cells are targeted by intradermal or transcutaneous administration, by use of, for example, microneedles or microabrasion devices. Further, the vaccines of the invention can be administered in a single dose or, optionally, administration can involve the use of a priming dose followed by a booster dose that is administered, e.g., 2-6 months later, as determined to be appropriate by those of skill in the art. Optionally, PIV vaccines can be administered via DNA or RNA immunization using methods known to those skilled in the art (Chang et al., Nat. Biotechnol. 26:571-577, 2008; Kofler et al., Proc. Natl. Acad. Sci. U.S.A. 101:1951-1956, 2004).

Optionally, adjuvants that are known to those skilled in the art can be used in the administration of the viruses and vectors of the invention. Adjuvants that can be used to enhance the immunogenicity of the viruses include, for example, liposomal formulations, synthetic adjuvants, such as (e.g., QS21), muramyl dipeptide, monophosphoryl lipid A, polyphosphazine, CpG oligonucleotides, or other molecules that appear to work by activating Toll-like Receptor (TLR) molecules on the surface of cells or on nuclear membranes within cells. Although these adjuvants are typically used to enhance immune responses to inactivated vaccines, they can also be used with live or replication-defective vaccines. Both agonists of TLRs or antagonists may be useful in the case of live or replication-defective vaccines. The vaccine candidates can be designed to express TLR agonists. In the case of a virus delivered via a mucosal route, for example, orally, mucosal adjuvants such as the heat-labile toxin of *E. coli* (LT) or mutant derivations of LT can be used as adjuvants. In addition, genes encoding cytokines that have adjuvant activities can be inserted into the vaccine candidates. Thus, genes encoding desired cytokines, such as GM-CSF, IL-2, IL-12, IL-13, IL-5, etc., can be inserted together with foreign immunogen genes to produce a vaccine that results in enhanced immune responses, or to modulate immunity directed more specifically towards cellular, humoral, or mucosal responses (e.g., reviewed in "Immunopotentiators in Modern Vaccines", Schijns and O'Hagan Eds., 2006, Elsevier Academic Press: Amsterdam, Boston, etc.). Optionally, a patch containing a layer of an appropriate toxin-derived adjuvant, can be applied over the injection site. Toxin promotes local inflammation attracting lymphocytes, which leads to a more robust immune response.

Examples

Additional details concerning the invention are provided in the Examples, below. In the Examples, experiments are described in which PIVs based on WN, JE, and YF viruses (see, e.g., WO 2007/098267 and WO 2008/137163) were tested. Firstly, we demonstrated that the constructs are significantly more attenuated in a sensitive suckling mouse neurovirulence model (zero mortality at all tested doses) as compared to available LAV controls (YF 17D, YF/JE LAV, and YF/WN LAV). We demonstrated for the first time that d-PIV constructs were avirulent in this model and thus that two-component PIVs do not undergo uncontrolled (unlimited) spread in vivo and cannot cause clinical signs. Secondly, we performed comparisons of the immunogenicity and efficacy of the PIVs and the LAVs, and demonstrated that PIV vaccines can induce immune response comparable to LAVs and be equally efficacious (e.g., as observed for PIV-WN and YF/WN LAV pair of vaccines). In one pair examined, YF 17D LAV was significantly more immunogenic than PIV-YF. Thus, production of VLPs can vary between different, similarly designed PIV constructs. Specifically, we propose that PIV-YF does not generate a large amount of YF VLPs compared to PIV-WN (WN VLPs), and that increased production of VLPs can be achieved by genetic modifications at the C/prM junction in suboptimal PIV constructs. Specifically, the C/prM junction is an important location in the flavivirus polyprotein orchestrating the formation of viral envelope and synthesis of viral proteins (Yamshchikov and Compans, Virology 192:38-51, 1993; Amberg and Rice, J. Virol. 73:8083-8094, 1999; Stocks and Lobigs, J. Virol. 72:2141-2149, 1998). We propose that secretion of VLPs in PIV infected cells (in contrast to production of viral particles in whole viruses) can be increased by uncoupling of the viral protease and signalase cleavages at the junction, or use of a strong heterologous signal peptide (tPA, etc.) in place of the signal for prM, or by mutagenesis of the signal for prM. The efficiency of signalase cleavage at the C/prM junction of flaviviruses is low (Stocks and Lobigs, J. Virol. 72:2141-2149, 1998), e.g., as predicted by SignalP 3.0 on-line program. It is expected that more efficient cleavage efficiency can be achieved by analysis of specific amino acid substitutions near the cleavage site with SignalP 3.0 (e.g., as described in application WO 2008/100464), followed by incorporation of chosen mutation(s) into PIV genomes, recovery of PIV progeny and

measuring VLP secretion. Non-flavivirus signals are inserted by methods standard in the art. Uncoupling between the viral protease and signalase cleavages can be achieved by ablating the viral cleavage site by any non-conservative mutation (e.g., RRS in YF17D C to RRA or GRS or RSS, etc.), or deletion of the entire site or some of its 3 residues. If necessary, formation of free N-terminus of the signal of foreign protein can be achieved by using such elements as autoprotease, or termination codon followed by an IRES. Alternatively, the native AUG initiation codon of C can be ablated (in constructs where C protein sequence is unnecessary, e.g., ΔC PIV) and AUG placed in front of foreign gene. Optimization of vector signal can be performed by random mutagenesis, e.g., by insertion of synthetic randomized sequence followed by identification of viable PIV variants with increased VLP secretion.

We also discovered that PIV constructs were substantially more immunogenic in hamsters when administered by the IP route, as compared to the subcutaneous route. We concluded that this was most likely due to better targeting of antigen presenting cells in lymphoid tissues, which are abundant in the abdomen, but not abundant in tissues underlying the skin. Based on these observations, we concluded that efficient targeting of PIVs to dendritic cells, abundant in the skin, can be achieved by cutaneous inoculation, e.g., via skin microabrasion or intradermal injection using microneedles (Dean et al., Hum Vaccin. 1:106-111, 2005).

Further, we have carried out experiments to show the feasibility of administering mixtures, or cocktails, of different PIVs, such as those described herein (e.g., JE+DEN and YF+DEN). In order to administer cocktails, it is important to verify that there is no interference between co-administered components, and that a balanced immune response is induced. Several PIV mixtures were used to immunize rodents and immune responses were compared to PIV constructs administered individually. No interference was observed in mixtures, and thus cocktail PIV vaccines are feasible. Such formulations may be of particular significance in geographical regions where different flaviviruses co-circulate. This could be also used to simultaneously administer several PIV-based vaccines against non-flavivirus pathogens.

Further, we have demonstrated that no neutralizing antibody response is induced against packaging envelope after at least two doses of PIV (and thus

antibodies are elicited against VLPs secreted from infected cells). This was demonstrated using the helper (ΔprM-E) component of a d-PIV (see in Fig. 2) packaged individually, or by measuring neutralizing antibodies to heterologous packaging envelopes (e.g., to the WN envelope used to package PIV-JE in helper cells providing WN-specific C-prM-E proteins *in trans*). The latter observations support sequential use of different PIV vaccines manufactured in a universal helper packaging cells line, and sequential use of different recombinant PIV-vectored vaccines in the same individual, as discussed above. In addition, we confirmed previous observations that PIV constructs can be stably propagated to high yields in vitro, and that no recombination restoring whole virus occurs after prolonged passaging in substrate cells (Mason et al., Virology 351:432-443, 2006; Shustov et al., J. Virol. 21:11737-11748, 2007).

These and other aspects of the invention are further described in the Examples, below.

Example 1. Pseudoinfectious virus platform development studies

Attenuation in suckling mouse neurovirulence (NV) model

Materials used in the studies described below are described in Table 1 and the references cited therein. These include s-PIV-WN (based on wt WN virus strain NY99 sequences), s-PIV-JE, s-PIV-WN/JE (based on wt WN virus backbone and prM-E genes from wt JE virus Nakayama strain), s-PIV-YF/WN (YF 17D backbone and prM-E genes from WN virus), and s-PIV-YF (based on YF 17D sequences). Additional materials include d-PIV-YF (YF d-PIV, grown in regular BHK cells (Shustov et al., J. Virol. 21:11737-11748, 2007), and two-component d-PIV-WN (grown in regular Vero cells; Suzuki et al., J. Virol. 82:6942-6951, 2008).

Attenuation of these PIV prototypes was compared to LAVs YF 17D, a chimeric YF/JE virus, and a chimeric YF/WN virus in suckling mouse NV test (IC inoculation) using highly susceptible 5-day old ICR mice (the chimeric viruses include yellow fever capsid and non-structural sequences, and JE or WN prM-E sequences). None of the animals that received PIV constructs showed clinical signs or died, while mortality was observed in animals inoculated with LAVs (Table 2). The

YF 17D virus is neurovirulent for mice of all ages, while the chimeric vaccines are not neurovirulent for adult mice, but can cause dose-dependent mortality in more sensitive suckling mice (Guirakhoo et al., Virology 257:363-372, 1999; Arroyo et al., J. Virol. 78:12497-12507, 2004). Accordingly, 90%-100% of suckling mice that received doses as low as 1 PFU of YF 17D died. YF/JE and YF/WN LAVs caused partial mortality at much higher doses (> 2 $\log_{10}$ PFU and 3 $\log_{10}$ PFU, respectively), with longer average survival time (AST) of animals that died, as expected. Thus, PIV constructs are completely avirulent in this sensitive model (at least 20,000 - 200,000 times less neurovirulent than the licensed YF 17D vaccine).

The YF d-PIV and WN d-PIV caused no mortality or clinical signs. Thus, the two-component PIV variants that theoretically could spread within brain tissue from cells co-infected by both of their components did not cause disease. Moreover, we tried to detect the d-PIVs in the brains of additional animals in this experiment, sacrificed on day 6 post-inoculation by titration, and detected none (brain tissues from 10 and 11 mice that received 4 $\log_{10}$ FFU of YF d-PIV and WN d-PIV, respectively, were homogenized and used for titration). Thus, the d-PIVs did not cause spreading infection characteristic of whole virus. YF/JE LAV has been shown to replicate in the brain of adult ICR mice inoculated by the IC route with a peak titer of ~ 6 $\log_{10}$ PFU/g on day 6, albeit without clinical signs (Guirakhoo et al., Virology 257:363-372, 1999). Co-infection of cells with components of a d-PIV is clearly a less efficient process than infection with whole virus. The data show that d-PIV replication in vivo is quickly brought under control by innate immune responses (and adaptive responses in older animals).

Immunogenicity/efficacy in mice and hamsters

Immunogenicity/efficacy of the PIV prototypes described above was compared to that of chimeric LAV counterparts and YF 17D in mice and Syrian hamsters. The general experiment design is illustrated in Fig. 3 (mice, IP immunization). Experiments in hamsters were performed similarly (plus-minus a few days, SC or IP inoculation with doses indicated below). 3.5-week old ICR mice (for s-PIV-WN and -YF, YF/WN LAV, and YF 17D groups) or C57/BL6 mice (for s-PIV-JE and YF/JE LAV groups) were immunized IP with graded doses of PIV constructs (4-6 $\log_{10}$

FFU/dose) or chimeric LAV and YF 17D LAV controls (4 log_{10} PFU). Select PIV-WN, -JE and -YF groups were boosted on day 21 with 5 log_{10} FFU of corresponding constructs (Table 3). Neutralizing antibody responses were determined in animal sera by standard $PRNT_{50}$ against YF/WN or /JE LAVs, or YF 17D viruses. PIV-WN induced very high WN-specific neutralizing antibody responses in all groups, with or without boost, as evidenced by $PRNT_{50}$ titers determined in pools of sera from immunized animals on days 20 and 34, which was comparable to that in the YF/WN LAV control group. Accordingly, animals immunized with both PIV-WN and YF/WN LAV were protected from lethal challenge on day 35 with wt WN virus (IP, 270 LD_{50}), but not mock-immunized animals (Table 3). When WN neutralizing antibodies were measured in sera from individual mice, high uniformity of immune responses was observed (Fig. 4). Thus, single-round PIV vaccines can be as immunogenic and efficacious as corresponding LAVs. PIV-JE was also highly immunogenic (black mice), while immunogenicity of PIV-YF was significantly lower compared to the YF 17D control (ICR mice). Yet, dose-dependent protection of PIV-YF immunized animals (but not mock-immunized animals) was observed following a severe lethal IC challenge with wt YF strain Asibi virus (500 LD_{50}) (Table 3), which is in agreement with the knowledge that neutralizing antibody titers as low 1:10 are protective against flavivirus infections.

The YF 17D control virus was highly immunogenic (e.g., $PRNT_{50}$ titer 1:1,280 on day 34), and thus it is able to infect cells and replicate efficiently in vivo, and its envelope is a strong immunogen. Therefore, it is unlikely that low immunogenicity of PIV-YF was due to its inability to infect cells or replicate efficiently in infected cells in vivo. We believe that the low immunogenicity of PIV-YF (e.g., compared to PIV-WN) was most likely due to a low-level production of YF-specific VLPs in PIV-YF infected cells (while VLP secretion is high in PIV-WN infected cells). As discussed above, we propose that immunogenicity of PIV-YF can be significantly increased, e.g., by appropriate modifications at the C/prM junction, e.g., by uncoupling the two protease cleavages that occur at this junction (viral protease and signalase cleavages), and/or by using a strong heterologous signal [e.g., rabies virus G protein signal, or eukaryotic tissue plasminogen activator (tPA) signal (Malin et al., Microbes and Infection, 2:1677-1685, 2000), etc.] in place of the YF signal for prM.

A similar experiment was performed in ~ 4.5-week old Syrian hamsters, to compare immunogenicity of PIV constructs to LAV controls in this model. Animals were immunized SC with graded doses of the test articles (Table 4). PIV-WN was highly immunogenic, e.g., WN-specific $PRNT_{50}$ titers on day 38 (pre-challenge) were 1:320, 1:640, and 1:1280 in groups that received 5, 6, and 6 (prime)+5 (boost) log_{10} FFU doses, respectively. This was somewhat lower compared to YF/WN LAV 4 log_{10} PFU control (≥ 1:2560). PIV-JE and –YF induced detectable specific neutralizing antibody responses, albeit with lower titers compared to YF/JE LAV and YF 17D controls. All animals immunized with PIV-WN and YF/WN were solidly protected from lethal challenge with wt WN virus as evidenced by the absence of mortality and morbidity (e.g., loss of body weight after challenge), as well as absence or a significant reduction of postchallenge WN virus viremia. Mock-immunized animals were not protected (Table 4). PIV-JE and –WN protected animals from respective challenge in dose-dependent fashion. Protective efficacy in this experiment is additionally illustrated in Fig. 5. For example, high post-challenge YF virus (hamster adapted Asibi strain) viremia was observed in mock immunized animals, peaking on day 3 at a titer of> 8 log_{10} PFU/ml (upper left panel); all of the animals lost weight, and 1 out of 4 died (upper right panel). In contrast, viremia was significantly reduced or absent in hamsters immunized with PIV-YF (two doses; despite relatively low neutralizing titers) or YF 17D; none of these animals lost weight. Similarly, animals immunized with PIV-WN or YF/WN LAV were significantly or completely protected in terms of post-challenge WN virus viremia and body weigh loss/mortality, in contrast to mock controls (compare in bottom panels). Thus, high immunogenicity/efficacy of PIV was demonstrated in a second animal model.

In another hamster experiment, animals were immunized with PIV constructs by the IP route, with two doses. Table 5 compares neutralizing immune responses (specific for each vaccine) determined in pooled sera of hamsters in the above-described experiment (SC inoculation) to those after IP immunization, for PIV-WN, -YF/WN, –WN/JE, and –YF after the first dose (days 20-21) and second dose (days 34-38). A clear effect of the immunization route was observed both after the 1st and 2nd doses. For instance, for PIV-WN after 1st dose, SC immunization resulted in

WN-specific PRNT50 titer of 1:40, while IP inoculation resulted in much higher titer 1:320 (and after the 2nd dose, titers were similar). A more pronounced effect was observed for other constructs after both the 1st and 2nd doses. Interestingly, PIV-YF/WN was very highly immunogenic by IP route (titer 1:320 after 1st IP dose vs. 1:20 by SC, and 1:1,280 after 2nd dose vs. 1:160 by SC). Similarly, immunogenicity of PIV-JE was significantly increased (e.g., JE-specific titer of 1:640 after two IP poses). Thus, better targeting of lymphoid cells, specifically antigen-presenting cells (which are more abundant in the abdomen as opposed to tissues under the skin), is an important consideration for use of PIV vaccines. In humans, efficient targeting of dendritic cells of the skin, increasing the magnitude of immune response, can be achieved by intradermal delivery, which we thus propose for a route for PIV immunization of humans.

In the above-described experiments, we also determined whether a neutralizing antibody response was induced against packaging envelopes (as opposed to response to VLPs encoded by PIV constructs and secreted by infected cells). No WN-specific neutralizing antibodies were detected by $PRNT_{50}$ in animals immunized with 5 $\log_{10}$ FFU of the second component of WN d-PIV, containing the ΔC-prM-E deletion and thus not encoding VLPs, but packaged into the WN envelope in BHK-CprME(WN) helper cells, and no YF-specific neutralizing activity was found in sera from animals immunized with 4 $\log_{10}$ FFU of the second component of YF d-PIV packaged in YF envelope. No YF-specific neutralizing response was induced by two doses of PIV-YF/WN packaged into YF envelope, and similarly, no WN-specific response was induced by two doses of PIV-JE packaged into WN envelope. The absence of neutralizing response against packaging envelopes permits manufacturing different PIV vaccines in one (universal) manufacturing helper cell line, or immunization of one individual with different recombinant vaccines based on the same vector, according to the present invention.

PIV cocktails

Because PIVs undergo a single (optionally several, but limited) round(s) of replication in vivo, we considered that mixtures of different PIV vaccines can be administered without interference between individual constructs in the mixture

(cocktail). To elucidate whether PIV vaccines can be used in cocktail formulations, immune responses in mice and hamsters to several PIV constructs given as mixtures were compared to the same constructs given individually. Similar results were obtained in both animal models. Results of mouse experiments are shown in Table 6. Similar anti-JE neutralizing antibody titers were observed in pools of sera from animals that were given one or two doses of either PIV-JE + PIV-WN mixture or PIV-JE alone (1:20 vs. 1:80 and 1:640 vs. 1:160, for one and two doses, respectively). Similarly, WN-specific titers against PIV-JE + PIV-WN mixture and PIV-WN alone were similar (1:320 vs. 1:640 and 1:5,120 vs. 1:5,120 for one and 2 doses, respectively). No or little cross-specific response was induced by either PIV-JE or -WN. The result was also confirmed by measuring $PRNT_{50}$ titers in sera from individual animals. Thus, it is clear that PIV vaccines can be efficiently administered as cocktails, inducing immunity against two or more flavivirus pathogens. In addition, as discussed above, various cocktails can be made between non-flavivirus PIV vaccines, or between any of flavivirus and non-flavivirus PIV vaccines.

In vitro studies

Different PIV prototypes were serially passaged up to 10 times in helper BHK cells, for s-PIVs, or in regular Vero cells, for d-PIVs. Samples harvested after each passage were titrated in Vero cells by immunostaining. Constructs grew to high titers, and no recombination restoring whole virus was observed. For instance, PIV-WN consistently grew to titers 7-8 $\log_{10}$ FFU/ml in BHK-CprME(WN) helper cells (containing a VEE replicon expressing the WN virus C-prM-E proteins), and WN d-PIV grew to titers exceeding 8 $\log_{10}$ FFU/ml in Vero cells, without recombination.

Example 2. PIV-TBE

PIV-TBE vaccine candidates can be assembled based entirely on sequences from wt TBE virus or the closely serologically related Langat (LGT) virus (naturally attenuated virus, e.g., wt strain TP-21 or its empirically attenuated variant, strain E5), or based on chimeric sequences containing the backbone (capsid and non-structural sequences) from YF 17D or other flaviviruses, such as WN virus, and the prM-E envelope protein genes from TBE, LGT, or other serologically related flaviviruses

from the TBE serocomplex. YF/TBE LAV candidates are constructed based on the backbone from YF 17D and the prM-E genes from TBE or related viruses (e.g., the E5 strain of LGT), similar to other chimeric LAV vaccines.

Construction of PIV-TBE and YF/TBE LAV vaccine prototypes was performed by cloning of appropriate genetic elements into plasmids for PIV-WN (Mason et al., Virology 351:432-443, 2006; Suzuki et al., J. Virol. 82:6942-6951, 2008), or plasmids for chimeric LAVs (e.g., pBSA-AR1, a single-plasmid version of infectious clone of YF/JE LAV; WO 2008/036146), respectively, using standard methods in the art of reverse genetics. The prM-E sequences of TBE virus strain Hypr (GenBank accession number U39292) and LGT strain E5 (GenBank accession number AF253420) were first computer codon-optimized to conform to the preferential codon usage in the human genome, and to eliminate nucleotide sequence repeats longer than 8 nt to ensure high genetic stability of inserts (if determined to be necessary, further shortening of nt sequence repeats can be performed). The genes were chemically synthesized and cloned into plasmids for PIV-WN and YF/JE LAV, in place of corresponding prM-E genes. Resulting plasmids were in vitro transcribed and appropriate cells (Vero for chimeric viruses, and helper BHK cells for PIV) were transfected with RNA transcripts to generate virus/PIV samples.

YF/TBE LAV constructs

In YF/TBE constructs containing either the TBE Hypr (plasmids p42, p45, and p59) or LGT E5 (plasmid P43) prM-E genes, two different types of the C/prM junction were first examined (see in Fig. 6; C/prM junctions only are shown in Sequence Appendix 1, and complete 5'-terminal sequences covering the 5'UTR-C-prM-E-beginning of NS1 region are shown in Sequence Appendix 2). The p42-derived YF17D/Hypr chimera contained a hybrid YF17D/Hypr signal peptide for the prM protein, while the p45-derived YF17D/Hypr chimera contained a hybrid YF17D/WN signal peptide for prM (Sequence Appendix 1). The former chimeric virus produced very high titers at both P0 (immediately after transfection) and P1 (the next passage in Vero cells), up to 7.9 $\log_{10}$ PFU/ml, which were 0.5 $\log_{10}$ times higher, compared to the latter virus; in addition it formed significantly larger plaques in Vero cells (Fig. 6). Thus, use of TBE-specific residues in the signal peptide for

prM conferred a significant growth advantage over the signal containing WN-specific residues. The p43-derived YF17D/LGT chimera had the same prM signal as the p42-derived virus; it also produced very high titers at P0 and P1 passages (up to 8.1 $\log_{10}$ PFU/ml) and formed large plaques. A derivative of the p42-derived virus was also produced from plasmid p59, which contained a strong attenuating mutation characterized previously in the context of a YF/WN LAV vaccine virus, specifically, a 3-a.a. deletion in the YF17D-specific C protein (PSR, residues 40-42 in the beginning of α-Helix I; WO 2006/116182). As expected, the p59 virus grew to lower titers (5.6 and 6.5 $\log_{10}$ PFU/ml at P0 and P1, respectively), and formed small plaques (determined in a separate titration experiment and thus not shown in Fig. 6), compared to the parent p42-derived chimera. These initial observations of growth properties of YF/TBE LAV prototypes, and correlation of replication in vitro with plaque morphologies, have been confirmed in growth curve experiments (Fig. 8).

PIV-TBE constructs

PIV-WN/TBE variants were constructed, and packaged PIV samples were derived from plasmids p39 and p40 (Fig. 7; Sequence Appendix 1 for C/prM junction sequences, and Sequence Appendix 3 for complete 5'UTR-ΔC-prM-E-beginning of NS1 sequences). These contained complete Hypr or WN prM signals, respectively. Both PIVs were successfully recovered and propagated in BHK-CprME(WN) or BHK-C(WN) helper cells (Mason et al., Virology 351:432-443, 2006; Widman et al., Vaccine 26:2762-2771, 2008). The P0 and P1 sample titers of the p39 variant were 0.2 – 1.0 $\log_{10}$ times, higher than p40 variant. In addition, Vero cells infected with p39 variant were stained brighter in immunofluorescence assay using a polyclonal TBE-specific antibody, compared to p40, indicative of more efficient replication (Fig. 7). The higher rate of replication of the p39 candidate than p40 candidate was confirmed in a growth curve experiment (Fig. 8). In the latter experiment, both candidates appeared to grow better in the BHK-C(WN) helper cells compared to BHK-CprME(WN), with the p39 variant reaching titer of ~ 7 $\log_{10}$ PFU/ml on day 5 (note that peak titers have not been reached). The discovery of the effect of prM signal on replication rates of both PIV and chimeric LAV vaccine candidates, and head-to-head comparison of different signals to generate the most efficiently

replicating and immunogenic (see above) construct, are a distinguishing feature of our approach. As discussed above, the invention also includes the use of other flavivirus signals, including with appropriate mutations, the uncoupling the viral protease and signalase cleavages at the C/prM junction, e.g., by mutating or deleting the viral protease cleavage site at the C-terminus of C preceding the prM signal, the use of strong non-flavivirus signals (e.g., tPA signal, etc.) in place of prM signal, as well as optimization of sequences downstream from the signalase cleavage site.

Other PIV-TBE variants based entirely on wt TBE (Hypr strain) and LGT virus (TP21 wild type strain or attenuated E5 strain), and chimeric YF 17D backbone/prM-E (TBE or LGT) sequences are also included in the invention. Helper cells providing appropriate C, C-prM-E, etc., proteins (e.g., TBE-specific) for trans-complementation can be constructed by means of stable DNA transfection or through the use of an appropriate vector, e.g., an alphavirus replicon, such as based on VEE strain TC-83, with antibiotic selection of replicon-containing cells. Vero and BHK21 cells can be used in practice of the invention. The former are an approved substrate for human vaccine manufacture; any other cell line acceptable for human and/or veterinary vaccine manufacturing can be also used. In addition to s-PIV constructs, d-PIV constructs can also be assembled. To additionally ascertain safety for vaccinees and the environment, appropriate modifications can be employed, including the use of degenerate codons and complementary mutations in the 5' and 3' CS elements, to minimize chances of recombination that theoretically could result in viable virus. Following construction, all vaccine candidates can be evaluated in vitro for manufacturability/stability, and in vivo for attenuation and immunogenicity/efficacy, in available pre-clinical animal models, such as those used in development and quality control of TBE and YF vaccines.

Neurovirulence and neuroinvasiveness in mice of PIV-TBE and YF/TBE LAV constructs

Young adult ICR mice (~ 3.5 week-old), were inoculated with graded doses of PIV-TBE and YF/TBE LAV candidates by the IC route to measure neurovirulence, or IP route to measure neuroinvasiveness (and later immunogenicity/efficacy). Animals that received 5 $\log_{10}$ FFU of PIV-Hypr (p39 and p40) variants by both routes survived

and showed no signs of sickness, similar to mock-inoculated animals (Table 7), and thus PIV-TBE vaccines are completely avirulent. Mice inoculated IC with YF 17D control (1 – 3 $\log_{10}$ PFU) showed dose-dependent mortality, while all animals inoculated IP (5 $\log_{10}$ PFU) survived, in accord with the knowledge that YF 17D virus is not neuroinvasive. All animals that received graded IC doses (2 – 4 $\log_{10}$ PFU) of YF/TBE LAV prototypes p42, p45, p43, and p59 died (moribund animals were humanely euthanized). These variants appear to be less attenuated than YF 17D, e.g., as evidenced by complete mortality and shorter AST at the 2 $\log_{10}$ PFU dose, the lowest dose tested for YF/TBE LAV candidates. The non-neurovirulent phenotype of PIV-TBE, virulent phenotype of YF/TBE LAV and intermediate-virulence phenotype of YF 17D are also illustrated in Fig. 9, showing survival curves of mice after IC inoculation. It should be noted that the p43 (LGT prM-E genes) and p59 (the dC2 deletion variant of YF/Hypr LAV) were less neurovirulent than p42 and p45 YF/Hypr LAV constructs as evidenced by larger AST values for corresponding doses (Table 7). In addition, p43 and p59 candidates were non-neuroinvasive, while p42 and p45 caused partial mortality after IP inoculation (5 $\log_{10}$ PFU/dose) (Table 7; Fig. 10). It should be noted however that all the YF/TBE LAV constructs were significantly attenuated as compared to wt TBE viruses, e.g., compared to wt TBE Hypr virus, which is uniformly highly virulent for mice, both at very low IC ($LD_{50} \sim 0.1$ PFU) and IP ($LD_{50} \leq 10$ PFU) doses (Wallner et al., J. Gen. Virol. 77:1035-1042, 1996; Mandl et al., J. Virol. 72:2132-2140, 1998; Mandl et al., J. Gen. Virol. 78:1049-1057, 1997

Immunogenicity/efficacy of PIV-TBE and YF/TBE LAV constructs in mice

TBE-specific neutralizing antibody responses in mice immunized IP with one or two doses of the PIV-TBE or YF/TBE LAV variants described above, or a human formalin-inactivated TBE vaccine control (1:30 of human dose) are being measured. Animals have been challenged with a high IP dose (500 PFU) of wt Hypr TBE virus; morbidity (e.g., weight loss), and mortality after challenge are monitored.

Immunogenicity/efficacy of PIV-TBE and YF/TBE LAV constructs in mice

TBE-specific neutralizing antibody responses in mice immunized IP with one or two doses of the PIV-TBE or YF/TBE LAV variants described above (from

experiment in Table 7), or a human formalin-inactivated TBE vaccine control (1:20 of human dose; one or two doses), or YF 17D and mock controls, were measured on day 20 by $PRNT_{50}$ against wt TBE Hypr virus (Table 8; second dose of indicated test articles was given on day 14). [Titers were determined in individual sera, or pooled sera from two animals in most cases, or pooled sera from 4 animals for the YF17D and Mock negative controls]. Titers in individual test samples as well as GMTs for each group are provided in Table 8. Titers in test samples were similar within each group, e.g., in groups immunized with PIVs, indicating high uniformity of immune response in animals. As expected, no TBE-specific neutralizing antibodies were detected in negative control groups (YF 17D and Mock; GMTs < 1:10); accordingly, animals in these groups were not protected from challenge on day 21 post-immunization with a high IP dose (500 PFU) of wt Hypr TBE virus. Mortalities from partial observation (on day 9 post-challenge; observation being continued) are provided in Table 8, and dynamics of average post-challenge body weights indicative of morbidity are shown in Fig. 11. Neutralizing antibodies were detected in killed vaccine controls, which were particularly high after two doses (GMT 1:1,496); animals in the 2-dose group were completely protected in that there was no mortality or body weight loss (but not animals in the 1-dose group). Animals that received both one and two doses of PIV-Hypr p39 had very high antibody titers (GMTs 1:665 and 1:10,584) and were solidly protected, demonstrating that robust protective immunity can be induced by s-PIV-TBE defective vaccine. The two animals that survived immunization with YF/Hypr p42 chimera (see in Table 7) also had high antibody titers (GMT 1:6,085) and were protected (Table 8; Fig. 11). Interestingly, PIV-Hypr p40 and YF/Hypr p45 were poorly immunogenic (GMTs 1:15 and 1:153 for one and two doses, and 1:68, respectively). As discussed above, these contained WN-specific sequences in the signal for prM, while the highly immunogenic PIV-Hypr p39 and YF/Hypr p42 constructs contained TBE-specific signal sequences. In agreement with discussion above, this result demonstrates the importance of choosing the right prM signal, e.g., the TBE-specific signal, to achieve high-level replication/VLP secretion, which in this experiment in vivo resulted in drastically different immune responses. Immunogenicity of YF/LGT p43 and YF/Hypr dC2 p59 chimeras was relatively low which could be expected, because of the use of a heterologous envelope (LGT,

different from challenge TBE virus) and high attenuating effect of the dC2 deletion, respectively.

Example 3. Foreign gene expression

In the examples of recombinant PIV constructs described below, genes of interest were codon optimized (e.g., for efficient expression in a target vaccination host) and to eliminate long nt sequence repeats to increase insert stability (≥ 8 nt long; additional shortening of repeats can be performed if necessary), and then chemically synthesized. The genes were cloned into PIV-WN vector plasmids using standard methods of molecular biology well known in the art, and packaged PIVs were recovered following in vitro transcription and transfection of appropriate helper (for s-PIVs) or regular (for d-PIVs) cells.

Expression of rabies virus G protein in WN s-PIV and d-PIV

Rabies virus, *Rhabdoviridae* family, is a significant human and veterinary pathogen. Despite the availability of several (killed) vaccines, improved vaccines are still needed for both veterinary and human use (e.g. as an inexpensive pre-exposure prophylactic vaccines). Rabies virus glycoprotein G mediates entry of the virus into cells and is the main immunogen. It has been expressed in other vectors with the purpose of developing veterinary vaccines (e.g., Pastoret and Brochier, Epidemio. Infect. 116:235-240, 1996; Li et al., Virology 356:147-154, 2006).

Full length rabies virus G protein (original Pasteur virus isolate, GenBank accession number NC_001542) was codon-optimized, chemically synthesized, and inserted adjacent to the ΔC, ΔprM-E and ΔC-prM-E deletions in PIV-WN vectors (Fig. 12). The sequences of constructs are provided in Sequence Appendix 4. General designs of the constructs are illustrated in Fig. 13. The entire G protein containing its own signal peptide was inserted in-frame downstream from the WN C protein either with the ΔC deletion (ΔC and ΔC-prM-E constricts) or without (ΔprM-E) and a few residues from the prM signal. Foot and mouth disease virus (FMDV) 2A autoprotease was placed downstream from the transmembrane C-terminal anchor of G to provide cleavage of C-terminus of G from the viral polyprotein during translation. The FMDV 2A element is followed by WN-specific

signal for prM and prM-E-NS1-5 genes in the ΔC construct, or signal for NS1 and NS1-5 genes in ΔprM-E and ΔC-prM-E constructs.

Packaged WN(ΔC)-rabiesG, WN(ΔprME)-rabiesG, and WN(ΔCprME)-rabiesG PIVs were produced by transfection of helper BHK cells complementing the PIV vector deletion [containing a Venezuelan equine encephalitis virus (strain TC-83) replicon expressing WN virus structural proteins for trans-complementation]. Efficient replication and expression of rabies G protein was demonstrated for the three constructs by transfection/infection of BHK-C(WN) and/or BHK-C-prM-E(WN) helper cells, as well as regular BHK cells, by immunostaining and immunofluorescence assay (IFA) using anti-Rabies G monoclonal antibody (RabG-Mab) (Fig. 14). Titers were determined in Vero cells by immunostaining with the Mab or an anti-WN virus polyclonal antibody. Growth curves of the constructs in BHK-CprME(WN) cells after transfection with in vitro RNA transcripts are shown in Fig. 14, bottom panels. The PIVs grew efficiently to titers ~ 6 to >7 $\log_{10}$ FFU/ml. Importantly, nearly identical titers were detected by both RabG-Mab and WN-antibody staining, which was the first evidence of genetic stability of the insert. In PIV-infected Vero cells, which were fixed but not permeabilized, strong membrane staining was observed by RabG-Mab staining, demonstrating that the product was efficiently delivered to the cell surface (Fig. 15). The latter is known to be the main prerequisite for high immunogenicity of expressed G. Individual packaged PIVs can spread following infection of helper BHK cells, but cannot spread in regular cells as illustrated for WN(ΔC)-rabiesG PIV in Fig. 16. The fact that there is no spread in naïve BHK cells demonstrates that the recombinant RNA genomes cannot be non-specifically packaged into membrane vesicles containing the G protein, if produced by PIV infected cells. An identical result was obtained with the G protein of another rhabdovirus, Vesicular stomatitis virus (VSV), contrary to previous observations of non-specific packaging of Semliki Forest virus (SFV) replicon expressing VSV G protein (Rolls et al., Cell 79:497- 506, 1994). The latter is a desired safety feature. [Alternatively, some non-specific packaging could result in a limited spread of PIV in vivo, potentially enhancing anti-rabies immune response. The latter could be also a beneficial feature, given that such PIV is demonstrated to be safe]. The stability of the rabies G insert in the three PIVs was demonstrated by serial passages in helper BHK-CprME(WN) cells at high or low MOI (0.1 or 0.001 FFU/cell). At each passage, cell supernatants were harvested and titrated in

regular cells (e.g., Vero cells) using immunostaining with an anti-WN polyclonal antibody to determine total PIV titer, or anti-rabies G monoclonal antibody to determine titer of particles containing the G gene (illustrated for MOI 0.1 in Fig. 17; similar results were obtained at MOI 0.001). The WN(ΔC)-rabiesG PIV was stable for 5 passages, while the titer of insert-containing PIV started declining at passage 6, indicative of insert instability. This could be expected, because in this construct, large G gene insert (~ 1500 nt) is combined with a small ΔC deletion (~ 200 nt), significantly increasing the overall size of the recombinant RNA genome. In contrast, in WN(ΔprME)-rabiesG, and WN(ΔCprME)-rabiesG PIVs, the insert is combined with a much larger deletion (~ 2000 nt). Therefore, these constructs stably maintained the insert for all 10 passages examined (Fig. 17). Further, it can be seen in Fig. 17 that at some passages, titers as high as 8 $\log_{10}$ FFU/ml, or higher, were attained for all three PIVs, additionally demonstrating that PIVs can be easily propagated to high yields.

Following inoculation in vivo individually, the WN(ΔC)-rabiesG s-PIV is expected to induce strong neutralizing antibody immune responses against both rabies and WN viruses, as well as T-cell responses. The WN(ΔprME)-rabiesG and WN(ΔCprME)-rabiesG PIVs will induce humoral immune response only against rabies because they do not encode the WN prM-E genes. WN(ΔC)-rabiesG s-PIV construct can be also co-inoculated with WN(ΔprME)-rabiesG construct in a d-PIV formulation (see in Fig. 12), increasing the dose of expressed G protein, and with enhanced immunity against both pathogens due to limited spread. As an example of spread, titration results in Vero cells of a s-PIV sample, WN(ΔprME)-rabiesG, and a d-PIV sample, WN(ΔprME)-rabiesG + WN(ΔC) PIV (the latter did not encode rabies G protein), are shown in Fig. 18. Infection of naïve Vero cells with s-PIV gave only individual cells stainable with RabG-Mab (or small clusters formed due to division of cells). In contrast, large foci were observed following infection with the d-PIV sample (Fig. 18, right panel) that were products of coinfection with the two PIV types.

The WN(ΔCprME)-rabiesG construct can be also used in a d-PIV formulation, if it is co-inoculated with a helper genome providing C-prM-E *in trans* (see in Fig. 12). For example it can be a WN virus genome containing a deletion of one of the NS proteins, e.g., NS1, NS3, or NS5, which are known to be trans-complementable (Khromykh et al., J. Virol. 73:10272-10280, 1999; Khromykh et al., J. Virol. 74:3253-3263, 2000). We

have constructed a WN-ΔNS1 genome (sequence provided in Sequence Appendix 4) and obtained evidence of co-infection with WN(ΔprME)-rabiesG or WN(ΔCprME)-rabiesG constructs, and spread in vitro, by immunostaining. In the case of such d-PIVs, rabies G protein can be also inserted and expressed in helper genome, e.g., WN-ΔNS1 genome, to increase the amount of expressed rabies G protein resulting in an increased anti-rabies immune response. As with any dPIV versions, one immunogen can be from one pathogen (e.g., rabies G) and the other from a second pathogen, resulting in three antigenic specificities of vaccine. As discussed above, ΔNS1 deletions can be replaced with or used in combination with ΔNS3 and/or ΔNS5 deletions/mutations, in other examples.

Expression of RSV F protein in WN s-PIV and d-PIV

Respiratory syncytial virus (RSV), member of *Paramyxoviridae* family, is the leading cause of severe respiratory tract disease in young children worldwide (Collins and Crowe, Respiratory Syncytial Virus and Metapneumovirus, In: Knipe et al. Eds., Fields Virology, 5th ed., Philadelphia: Wolters Kluwer/Lippincott Williams and Wilkins, 2007:1601-1646). Fusion protein F of the virus is a lead viral antigen for developing a safe and effective vaccine. To avoid post-vaccination exacerbation of RSV infection observed previously with a formalin-inactivated vaccine candidate, a balanced Th1/Th2 response to F is required which can be achieved by better TLR stimulation, a prerequisite for induction of high-affinity antibodies (Delgado et al., Nat. Med. 15:34-41, 2009), which should be achievable through delivering F in a robust virus-based vector. We have previously demonstrated the capacity of yellow fever virus-based chimeric LAV vectors to induce a strong, balanced Th1/Th2 response in vivo against an influenza antigen (WO 2008/036146). In the present invention, both yellow fever virus-based chimeric LAVs and PIV vectors are used for delivering RSV F to induce optimal immune response profile. Other LAVs and PIV vectors described herein can also be used for this purpose.

Full-length RSV F protein of A2 strain of the virus (GenBank accession number P03420) was codon optimized as described above, synthesized, and cloned into plasmids for PIV-WN s-PIV and d-PIV, using the insertion schemes shown in Fig. 12 and 13 for rabies G protein, by applying standard methods of molecular biology. Exact sequences of the insertions and surrounding genetic elements are provided in Sequence Appendix 5. In vitro RNA transcripts of resulting WN(ΔC)-RSV F, WN(ΔprME)- RSV F, and

WN(ΔCprME)- RSV F PIV constructs were used to transfect helper BHK-CprME(WN) cells. Efficient replication and expression of RSV F protein was first demonstrated by immunostaining of transfected cells with an anti-RSV F Mab, as illustrated for the WN(ΔprME)- RSV F construct in Fig. 19. The presence of packaged PIVs in the supernatants from transfected cells (titer as high as 7 log10 FFU/ml) was determined by titration in Vero cells with immunostaining. Additionally, similar constructs can be used that contain a modified F protein gene. Specifically, the N-terminal native signal peptide of F is replaced in modified F protein with the one from rabies virus G protein. The modification is intended to elucidate whether the use of a heterologous signal can increase the rate of F protein synthesis and/or replication of PIVs.

It has been demonstrated that a C-terminally truncated, secreted form of RSV F could be more immunogenic than full-length protein (Li et al., J. Exp. Med. 188:681-688, 1998). Therefore, we also cloned the available truncated RSV F gene (see Fig. 20) into the WN PIV vectors. Insertion designs were as in Fig. 13, with the only exception that the gene did not contain the sequence encoding C-terminal F protein anchor, to produce soluble form of truncated F (trF) in the lumen of the ER, which should be efficiently secreted from cells. Resulting WN(ΔC)-RSV trF, WN(ΔprME)- RSV trF, and WN(ΔCprME)- RSV trF PIVs (see Sequence Appendix 6 for the sequences of these constructs) were recovered in helper BHK-CprME(WN) cells. Results of IFA for transfected cells, performed with anti-RSV F Mab, are shown in Fig. 21. Efficient expression of trF product was observed, also demonstrating that all defective recombinant viruses were viable. Titers of PIVs as high as $2x10^6$ FFU/ml were observed in cell supernatants immediately after transfection; these are expected to further increase with passages. Importantly, similar numbers of foci were detected by both anti-RSVF and anti-WN antibodies in titration experiments in Vero cells (Fig. 22), and intensities of staining with both antibodies were comparable indicative of high-level expression of trF product. Western analysis of two ΔprME-RSVtrF stocks, two days post-infection of Vero and BHK (WNV/C-prM-E) helper cells, is shown in Fig. 23. These PIVs can be evaluated further for immunogenicity/efficacy in available animal models for RSV disease, in both s-PIV and d-PIV formulations (see below for further evaluation of ΔprME-RSVtrF).

Set forth below is the RSV amino acid sequence of the truncated construct. The chimeric West Nile/RSV-F signal peptide (**ggktgiavi/*melpiikanaittiliavtfcfass***) is

designed to be cleaved by signal protease after "...fass", releasing N-terminus of F2 "qnitee....". At the C-terminus is the sequence of autoprotease FMDV 2A fused to RSVF (nfdllklagdvesnpg). This sequence and/or only the RSV F protein portion thereof can be used in any of the vectors described herein. Further, this sequence (and/or only the RSV F protein portion thereof) can be the basis for derivation of analogs and fragments for use in the invention. Thus, sequences having percentage identities to this sequence, as described above, or fragments, as described above, can be used in the invention.

ggktgiavimelpiikanaittiliavtfcfassqniteefyqstcsavskgylsalrtgwytsvitie
lsnikenkcngtdakvklikqeldkyknavtelqllmqstpaannrarrelprfmnytlnnakktnvtl
skkrkrrflgfllgvgsaiasgiavskvlhlegevnkiksallstnkavvslsngvsvltskvldlkny
idkqllpivnkqscsisnietviefqqknnrlleitrefsvnagvttpvstymltnsellslindmpit
ndqkklmsnnvqivrqqsysimsiikeevlayvvqlplygvidtpcwklhtsplcttntkegsnicltr
tdrgwycnnagsvsffpladtckvqsnrvfcdtmnsltlpsevnlcnidifnpkydckimtsktdvsss
vitslgaivscygktkctasnknrgiiktfsngcdyvsnkgvdtvsvgntlyyvnkqegkslyvkgepi
infydplvfpsdefdasisqvnekinqslafirksdellhnvnagksttnimnfdllklagdvesnpg

Recombinant poxviruses expressing RSV F

Based on the premise that protection can be obtained using a very limited, but focused antibody response, we have shown that a live vector expressing a codon optimized anchorless RSV F confers protection against RSV infection in an appropriate animal model and thus can be suitable for an infant vaccine.

NYVAC is a highly attenuated vaccinia strain with a series of deletion of virulence-associated or host-range genes of the Copenhagen strain (Tartaglia et al., Dev. Biol. Stand. 84:159-163, 1995). It has been used in a variety of pre-clinical and clinical studies and shown to be promising. Therefore, NYVAC has been included as a delivery vehicle for a comparative vaccine evaluation.

To generate the recombinant NYVAC expressing codon-optimized anchorless RSV F, IVR (in vitro recombination) was performed with CEF cells infected by parental NYVAC at M.O.I. of 10 and transfected with donor plasmid pLNZ16. Subsequently, IVR reaction products were serially diluted ten-fold from $1:10^3$ to $1:10^6$, and plated on CEF cells in 100-mm plates overlaid with medium-agarose without Blue-Gal. Three days after the first overlay, a second overlay containing Blue-Gal and Neutral Red was added. Blue plaques were picked and plaque purification continued until a white plaque was available for amplification. The isolated plaque went through three amplification steps, i.e., P1, P2, and P3.

The recombinant was fully characterized at P2 to confirm identity and purity. The NYVAC recombinant was designated vP2400.

Fowlpox is a member of the avipoxvirus genus and can cause disease in chickens and turkeys. Transmission of fowlpox virus is limited to avian species, with replication in mammalian cells resulting in abortive replication. The inability of fowlpox to produce infectious virus in mammalian cells renders fowlpox a very attractive vector for human vaccine development. The safety and efficacy of fowlpox-based vaccines have been investigated in a number of clinical trials for diseases such as cancer, HIV, and malaria. Preliminary results indicate that fowlpox vaccines are safe and well tolerated, and have demonstrated both immune and clinical efficacy. This vector was also used to compare delivery systems that express the RSV F gene product, and to allow a thorough evaluation of both immune efficacy and safety in relevant animal model systems.

To generate recombinant fowlpox expressing codon-optimized anchorless RSV F, IVR was performed with CEF cells infected by a parental fowlpox at M.O.I. of 10 and transfected with the donor plasmid pLNZ15 (Paoletti, Proc. Natl. Acad. Sci. U.S.A. 93:11349-11353, 1996). The rest of the steps are the same as above. The fowlpox recombinant was designated vFP2403.

Western blot analysis of PIV-mediated expression of RSV F (See Fig. 24)

Vero cells (~1.5 x 10^6) were infected at an MOI of 10 with: Lanes 2 and 3, vP2400 (NYVAC-RSV F); Lanes 4 and 5, vFP2403 (fowlpox-RSV F); Lanes 6 and 7, PIV-F (ΔprME-RSVtrF); and Lanes 8 and 9, mock infected cells. All recombinant viruses express a codon optimized anchorless RSV F. Cell supernatants were harvested at 24 (Lanes 2, 4, 6, and 8) and 48 (Lanes 3, 5, 7, and 9) hours after infection. Equal amounts of the supernatant samples were analyzed by SDS-PAGE and the amount of RSV F present in each sample was determined using primary antibody, i.e., a mouse anti-RSV F (5353C75), followed by a goat anti-mouse IgG-horseradish peroxidase (HRP) conjugate as secondary antibody. The level of RSV F present in each sample was measured by comparison to a purified preparation of protein F from RSV-infected cells (2.5 ng, Lane 10) measured using a Kodak Imager Station 4000MM Pro. The results demonstrate that the amount of RSV F expressed in

PIV-F infected Vero cells was significantly greater than that expressed by NYVAC and similar to that expressed by fowlpox.

Immunization studies using PIV-F

For intramuscular immunization, Balb/c (6-8 weeks old) were injected bilaterally with 2 x 50 µl of PBS solution containing viral vectors expressing RSV F protein at two doses – either 10^6 or 10^7 PFU. Animals were boosted 4 weeks later with the same dose of the vaccine. Mice in control groups were immunized intranasally with 10^6 PFU RSV-Long strain or intramuscularly with an FI-RSV vaccine (100 µl) prepared according to the procedures used for the 1960's trials. Four weeks after boost, mice were challenged intranasally with either 2.2 x 10^6 PFU RSV-A2 (for RSVi27) or 10^7 PFU RSV-A2 (for RSVi32).

ELISA analysis of sera derived from vaccinated mice (See Fig. 25)

Immune sera were analyzed for anti-RSV-F IgG antibody titers using ELISA, which was performed with an immunoaffinity-purified full-length RSV protein (50 ng/ml) by two-fold dilutions of immune sera. Goat anti-mouse F(ab)2 IgG (H+L) conjugated to horseradish peroxidase was used as secondary antibody. The titer is a reciprocal of the last dilution at which the OD450 was greater than 0.1 and at least twice that of a control, to which no sample was added. It can be seen from Fig. 25 that both i.m. and i.p. immunization with PIV-F generated the highest titers of IgG of the vectors tested.

Neutralization assay (See Fig. 26)

Vero cells were seeded onto 24-well plates (1.5 x 10^5 per well), incubated at 37°C for two days. The neutralization reaction mixtures (serial diluted sera + virus + complement) were prepared in DMEM and incubated for 1 hour in a 37°C shaker. The neutralization mixtures were added to the Vero cells. After a 2 hour incubation in a 37°C shaker, the mixtures were removed and overlay media (methyl cellulose/DMEM) was added to each well. The infected Vero cells were incubated for 4 days at 37°C, then fixed with 80% methanol and stained with a primary antibody, i.e., a mouse anti-RSV F antibody (5353C75), followed by a goat anti-mouse IgG-

horseradish peroxidase (HRP) conjugate as secondary antibody. The plaques were counted by eye and neutralizing titers were expressed as the dilution that caused 60% plaque reduction.

It can be seen from Fig. 26 that both i.m. and i.p. immunization with PIV-F showed the highest neutralization titers of all vectors tested.

Table 1. PIV prototype constructs used in platform development studies

Construct	Genetic composition	Packaged in
PIV-WN	wt NY99 WN virus	WN envelope; BHK-CprME(WN) or BHK-C(WN) helper cells (Mason et al., Virology 2006, 351:432-43; Widman et al., Vaccine 2008, 26:2762-71)
PIV-YF/WN	Envelope (VLP): wt WN NY99 Backbone: YF 17D	YF 17D envelope; BHK-CprME(YF) helper cells (Widman et al., Adv Virus Res. 2008, 72:77-126)
PIV-WN/JE	Envelope (VLP): wt JE Nakayama Backbone: wt WN NY99	JE or WN envelope; BHK-C(WN) or BHK-CprME(WN) helper cells (Ishikawa et al., Vaccine 2008, 26:2772-8)
PIV-YF	YF 17D	YF 17D envelope; BHK-CprME(YF) or BHK-C(YF) helper cells (Mason et al., Virology 2006, 351:432-43)

Table 2. Safety: Suckling mouse neurovirulence[1]

Construct	Doses (log_{10})	Mortality (%)	AST (days)[2]
PIV-YF	1 - 4	0/10 (0%)	na
PIV-WN	2 - 5	0/10 (0%)	na
PIV-WN/JE	1 - 4	0/11 (0%)	na
PIV-YF/WN	1 - 4	0/10-11 (0%)	na
WN d-PIV	1 – 4	0/10-11 (0%)	na
YF d-PIV	1 - 4	0/10 (0%)	na
YF17D	2 1 0 -1	10/10 (100%) 10/10 (100%) 9/10 (90%) 3/10 (30%)	7.6 9.3 9.9 9.6
YF/JE	4 3 2 1	9/11 (82%) 7/10 (70%) 3/11 (27%) 0/11 (0%)	9.7 12.3 12 na
YF/WN	3 0-2	2/11 (18%) 0/10-11 (0%)	12.5 na

[1]Single dose, IC inoculation, ICR 5-day old mice, graded log doses administered.
[2]AST for mice that died; na, not applicable.

Table 3. PIV highly immunogenic and efficacious in mice[1]

Group	Dose	PRNT Day 20	PRNT Day 34	Post-challenge mortality (%)
PIV-WN	10^5	640	1280	0/8 (0%)
	10^6	1280	2560	1/8 (12.5%)
	$10^6 + 10^5$	2560	2560	0/6 (0%)
YF/WN control	10^4	1280	2560	1/8 (12.5%)
PIV-WN/JE	10^4	10	20	N/D
	10^5	20	20	N/D
	10^5+10^5	20	160	N/D
YF/JE control	10^4	160	320	N/D
PIV-YF	10^4	<10	<10	8/8 (100%)
	10^5	<10	<10	5/7 (71%)
	10^5+10^5	10	10	2/5 (40%)
YF17D control	10^4	640	1280	0/7 (0%)
Mock control – WN challenge	Diluent	N/D	0	7/7 (100%)
– YF challenge	Diluent	N/D	0	8/8 (100%)

[1]IP immunization (d0 prime, and d21 boost in select groups); challenge on d35: wt WN NY99, 3 $\log_{10}$ PFU IP, 270 LD50; wt YF Asibi, 3 $\log_{10}$ PFU IC, 500 LD50; N/D, not determined.

Table 4. PIV are immunogenic in hamsters and protect against challenge[1]

Group	Dose(s)	PRNT Day 38	POST-CHALLENGE		
			Mortality	Morbidity	Peak viremia (log)
PIV-WN	10^5	320	0/5 (0%)	0/5 (0%)	2.3
	10^6	640	0/5 (0%)	0/5 (0%)	1.8
	10^6+10^5	1280	0/5 (0%)	0/5 (0%)	<1.3
YF/WN control	10^4	≥2560	0/5 (0%)	0/5 (0%)	<1.3
PIV-WN/JE	10^4	20	2/5 (40%)	2/5 (40%)	2.2
	10^5+10^5	40	0/5 (0%)	0/5 (0%)	<1.3
YF/JE control	10^4	2560	0/5 (0%)	0/5 (0%)	1.3
PIV-YF	10^4	<10	1/3 (33%)	3/3 (100%)	8.3
	10^5	<10	1/5 (20%)	4/5 (80%)	8.3
	10^5+10^5	20	0/4 (0%)	0/4 (0%)	2.5
YF17D control	10^4	≥2560	0/4 (0%)	0/4 (0%)	<1.3
Mock control – WN challenge	Diluent	<10	3/4 (75%)	4/4 (100%)	4.0
– YF challenge	Diluent	<10	1/4 (25%)	4/4 (100%)	8.4
– JE challenge	Diluent	<10	2/5 (40%)	2/5 (40%)	3.0

[1]Syrian hamsters, SC inoculation (d0, and d21 in select groups); challenge (d39): wt WN NY385/99 6 $\log_{10}$ PFU IP, wt JE Nakayama 5.8 $\log_{10}$ PFU IC, or hamster-adapted YF Asibi 7 $\log_{10}$ PFU IP (McArthur et al., J. Virol. 77:1462-1468, 2003; McArthur et al., Virus Res. 110:65-71, 2005).

Table 5. Immunization of hamsters with PIV: comparison of SC and IP routes

Inoculums	PRNT Day 20-21		Boost (log_{10})	PRNT Day 34-38	
	SC	IP		SC	IP
PIV-WN	40	**320**	5	1280	1280
PIV-YF/WN	**10**	**320**	5	**160**	**1280**
PIV-WN/JE	**10**	**80**	5	**40**	**640**
PIV-YF	<10	10	5	**20**	**80**

Table 6. Immune responses to PIV cocktails (mice)[1]

Group	Dose	PRNT Day 20		PRNT Day 34	
		Anti-JE	Anti-WN	Anti-JE	Anti-WN
PIV-WN/JE + RV-WN	10^5 +10^5	**20**	320	**640**	5120
PIV-WN/JE alone	10^5	**80**	<10	**160**	20
PIV-WN alone	10^5	<10	640	<10	5120
Mock	-	<10	<10	<10	<10

[1] C57/BL6 mice, IP inoculations on days 0 and 21; pooled serum PRNT titers.

Table 7. Neurovirulence (IC inoculation) and neuroinvasiveness (IP inoculation) of PIV-TBE and YF/TBE vaccine constructs in adult ICR mice

Construct	Neurovirulence (IC route)			Neuroinvasiveness (IP route)		
	Dose(s) (log_{10})	Mortality (%)	AST, days[1]	Dose(s) (log_{10})	Mortality (%)	AST, days[1]
PIV-Hypr p39	5	0/7 (0%)	na	5	0/16 (0%)	na
PIV-Hypr p40	5	0/6 (0%)	na	5	0/16 (0%)	na
YF/Hypr p42	4	8/8 (100%)	6.3	5	6/8 (75%)	13.3
	3	8/8 (100%)	6.4			
	2	8/8 (100%)	7.4			
YF/LGT p43	4	8/8 (100%)	7.9	5	0/8 (0%)	na
	3	8/8 (100%)	7.6			
	2	8/8 (100%)	8.4			
YF/Hypr p45	4	8/8 (100%)	6.1	5	5/8 (62.5%)	11.2
	3	8/8 (100%)	6.6			
	2	8/8 (100%)	6.8			
YF/Hypr dC2 p59	4	8/8 (100%)	6.6	5	0/8 (0%)	na
	3	8/8 (100%)	7.4			
	2	8/8 (100%)	8.1			
YF 17D	3	8/8 (100%)	9	5	0/8 (0%)	na
	2	7/8 (87.5%)	9.6			
	1	4/8 (50%)	10			
Mock (diluent)	none	0/8 (0%)	na	none	0/8 (0%)	na

[1]AST for mice that died.

Table 8. Neutralizing antibody titers ($PRNT_{50}$) in mice immunized IP (determined against wt TBE virus Hypr), and protection from challenge (postchallenge observation, day 9)

Immunogen	**Dose(s), log_{10}**	**$PRNT_{50}$ titer, individ. samples[1]**	**$PRNT_{50}$ GMT**	**Postchallenge mortality (%) on day 9[2]**
PIV-Hypr p39, 1 dose	5	1746 (2) 1187 (2) 164 (2) 574 (2)	**665**	**0/8 (0%)**
PIV-Hypr p39, 2 doses	5+5	16229 (2) 12928 (2) 12927 (2) 4627 (2)	**10,584**	**0/8 (0%)**
PIV-Hypr p40, 1 dose	5	<10 (2) <10 (2) 18 (2) 33 (2)	**15**	**6/8 (75%)**
PIV-Hypr p40, 2 doses	5+5	169 (2) 638 (2) 26 (2) 192 (2)	**153**	**1/8 (12.5%)**
YF/Hypr p42	5	9210 (1) 4020 (1)	**6,085**	**0/2 (0%)**
YF/LGT p43	5	123 (2) 32 (2) 96 (2) 45 (2)	**64**	**1/8 (12.5%)**
YF/Hypr p45	5	292 (2) 16 (1)	**68**	**0/3 (0%)**
YF/Hypr dC2 p59	5	194 (2) 93 (2) 45 (2) 26 (2)	**68**	**0/8 (0%)**
Killed human TBE vaccine, 1 dose (at 1/20 of human dose)	1/20	19 (2) <10 (2) 13 (2) <10 (2)	**12**	**1/8 (12.5%)**
Killed human TBE vaccine, 2 doses (each at 1/20 of human dose)	1/20+1/20	3435 (2) 1267 (2) 770 (2)	**1,496**	**0/6 (0%)**
YF 17D control	5	<10 (4) 11 (4)	**<10**	**5/8 (62.5%)**
Mock	none	<10 (4) <10 (4)	**<10**	**4/8 (50%)**

[1]Numbers in parenthesis correspond to number of mice in each pooled serum sample tested.
[2]Mortalities on day 9 are shown.

Table 9. Examples of published attenuating E protein mutations that can be used for attenuation of chimeric TBE LAV candidates

Residue	Domain	Comments	Attenuation in	Reference
N52R	II	DI- DII hinge, possibly involved in hinge motion required for fusion activation	JE, YF	Hasegawa et al, 1992, Schlesinger et al, 1996
E84K	II	conserved, E in TBE, K/R in others, attenuated by passage in *Ixodes ricinus* ticks, DII contains flavivirus cross reactive epitopes	TBE	Labuda et al, 1994
E85K	II	conserved, E in TBE, K/R in others, attenuation obtained as plaque variants in Vero cells, DII contains flavivirus cross reactive epitopes	JE	Wu et al, 1997
H104K	II	within highly conserved fusion peptide (aa 98-113), H in TBE, G in others	TBE	Rey et al, 1995
L107F	II	within highly conserved fusion peptide (aa 98-113), L in all flaviviruses, F in attenuated JE	TBE, JE, WN	Rey et al, 1995, Arroyo et al, 1999, 2004
T123K	II	DI- DII hinge, T in TBE, A in KFD	TBE	Holzmann et al, 1997
K126E	II	DI- DII hinge, K in TBE, E in D-2	DEN2	Bray, 98
K136E	II	DI- DII hinge, K in TBE and JE, E in D-2	JE	
N154L(Y)	I	glycosylation site, packed with conserved H 104, involved in fusion.	DEN2, DEN4, YF	Guirakhoo et al, 1993, Pletnev et al, 1993, Kawano et al, 1993, Jennings et al, 1994
K171E	I	external edge of DI, involved in fusion	TBE	Mandl, 1989, Holzmann, 1997
I173T		external edge of DI, involved in fusion	YF	Chambers and Nickells 2001
D181Y		DI- DII hinge	TBE	Holzmann et al, 1997
K204R		Lining Hydrophobic pocket, involve in fusion	DEN1, DEN3	Guirakhoo et al, 2004
P272S	II	highly conserved, junction of one the of 2 alpha helices	JE	Cecilia et al, 1991
G308N	III	cell attachment, DKT in TBE, EGS in KFD, T-X in others, change to N produced glycosylation site in LI and reduced virulence, N-X-T/S glycosylation motif	LI	Jiang et al, 1993, Gao et al, 1994
S310K	III	putative cell attachment, change from E to G in JE reduced virulence	JE	Jiang et al, 1993, Gao et al, 1994, Wu et al, 1997
K311E	III	highly conserved, putative cell attachment	TBE, YF	Rey et al, 1995, Jennings, 1994
T333L	III	putative cell attachment	YF, LGT	Raynman et al, 1998
G334K	III	putative cell attachment	YF	Chambers and Nickells, 2001
S335K	III	putative cell attachment	JE	Wu et al, 1997
K336D	III	putative cell attachment	JE	Cecilia and Gould, 1991
P337D	III	putative cell attachment	JE	Cecilia and Gould, 1991
G368R	III	putative cell attachment	TBE, JE	Holzman et al 1997, Hasegawa et al 1992
Y384H	III	change to H attenuated TBE, putative cell attachment, -3 position to deleted RGD in TBE	TBE	Holzmann et al, 1990
V385R	III	conserved, -2 position to deleted RGD in TBE, putative cell attachment	D2	Hiramatsu et al, 1996, Lobigs, 90
G386R	III	highly conserved, -1 position to deleted RGD in TBE, putative cell attachment	D2, MVE	Hiramatsu, 96, Lobigs et al, 1990
E387R	III	conserved, +2 position to deleted RGD in TBE, putative cell attachment	D2, MVE	Hiramatsu, 1996, Lobigs et al, 1990
F403K	none	highly conserved, C-terminal region not included in crystal structure sE	D-2, D-4	Kawano et al, 1993, Bray et al, 1998
H438Y	None	highly conserved, C-terminal region not included in crystal structure sE	LGT	Campbell and Pletnev 2000
H496R	none	highly conserved, C-terminal region not included in crystal structure sE	TBE	Gritsun et al, 2001

References: Hasegawa et al., Virology 191(1):158-165; Schlesinger et al., J. Gen. Virol. 1996, 77 (Pt 6):1277-1285, 1996; Labuda et al., Virus Res. 31(3):305-315, 1994; Wu et al., Virus Res. 51(2):173-181, 1997; Holzmann et al., J. Gen. Virol. 78 (Pt 1):31-37, 1997; Bray et al., J. Virol. 72(2):1647-1651, 1998; Guirakhoo et al., Virology 194(1):219-223, 1993; Pletnev et al., J. Virol. 67(8):4956-4963, 1993; Kawano et al., J. Virol. 67(11):6567-6575, 1993; Jennings et al., J. Infect. Dis. 169(3):512-518, 1994; Mandl et al., J. Virol. 63(2):564-571, 1989; Chambers et al., J. Virol. 75(22):10912-10922, 2001; Cecilia et al., Virology 181(1):70-77, 1991; Jiang et al., J. Gen. Virol. 74 (Pt 5):931-935, 1993; Gao et al., J. Gen. Virol. 75 (Pt 3):609-614, 1994; Holzmann et al., J. Virol. 64(10):5156-5159, 1990; Hiramatsu et al., Virology 224(2):437-445, 1996; Lobigs et al., Virology 176(2):587-595, 1990; Campbell et al., Virology 269(1):225-237, 2000; Gritsun et al., J. Gen. Virol. 82(Pt 7):1667-1675, 2001.

Other Embodiments

All publications, patent applications, and patents mentioned in this specification are incorporated herein by reference in their entirety as if each individual publication, patent application, or patent were specifically and individually indicated to be incorporated by reference.

Various modifications and variations of the described viruses, vectors, compositions, and methods of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention that are obvious to those skilled in the fields of medicine, pharmacology, or related fields are intended to be within the scope of the invention. Use of singular forms herein, such as "a" and "the," does not exclude indication of the corresponding plural form, unless the context indicates to the contrary. Similarly, use of plural terms does not exclude indication of a corresponding singular form. Other embodiments are within the scope of the following claims.

What is claimed is:

Sequence Appendix 1

CV-TBEV Hypr or CV-LGT E5 with YFV/TBEV chimeric signal (p42, p59, and p43 constructs)

YF17D partial signal | TBEV partial signal | C protein YF17D | Hypr or LGT E5 prM protein

```
                                                
```

Sequence Appendix 2

CV-TBEV Hypr with YFV/TBEV chimeric signal (p42)

```
                                                    5' UTR
     ----------------------------------------------------------------------------------------------------
   1 AGTAAATCCT GTGTGCTAAT TGAGGTGCAT TGGTCTGCAA ATCGAGTTGC TAGGCAATAA ACACATTTGG ATTAATTTTA ATCGTTCGTT GAGCGATTAG
     TCATTTAGGA CACACGATTA ACTCCACGTA ACCAGACGTT TAGCTCAACG ATCCGTTATT TGTGTAAACC TAATTAAAAT TAGCAAGCAA CTCGCTAATC
           5' UTR
     ----------
                                                          C protein
                                 ----------------------------------------------------------------------------
                                   M  S  G  R  K  A  Q  G  K  T  L  G  V  N  N  V  R  R  G  V  R  S  L  S  N  K  I  K
 101 CAGAGAACTG ACCAGAACAT GTCTGGTCGT AAAGCTCAGG GAAAAACCCT GGGCGTCAAT ATGGTACGAC GAGGAGTTCG CTCCTTGTCA AACAAATAA
     GTCTCTTGAC TGGTCTTGTA CAGACCAGCA TTTCGAGTCC CTTTTTGGGA CCCGCAGTTA TACCATGCTG CTCCTCAAGC GAGGAACAGT TTGTTTTATT
                                                C protein
     ----------------------------------------------------------------------------------------------------
     .  Q  K  T  K  Q  I  G  N  R  P  G  P  S  R  G  V  Q  G  F  I  F  F  F  L  F  N  I  L  T  G  K  K  I
 201 AACAAAAAAC AAAACAAATT GGAAACAGAC CTGGACCTTC AAGAGGTGTT CAAGGATTTA TCTTTTTCTT TTGTTCAAC ATTTTGACTG GAAAAAGAT
     TTGTTTTTTG TTTTGTTTAA CCTTTGTCTG GACCTGGAAG TTCTCCACAA GTTCCTAAAT AGAAAAAGAA AAACAAGTTG TAAAACTGAC CTTTTTCTA
                                                C protein
     ----------------------------------------------------------------------------------------------------
     .  T  A  H  L  K  R  L  W  K  M  L  D  P  R  Q  G  L  A  V  L  R  K  V  K  R  V  V  A  S  L  M  R  G
 301 CACAGCCCAC CTAAAGAGGT TGTGGAAAAT GCTGGACCCA AGACAAGGCT TGGCTGTTCT AAGGAAAGTC AAGAGAGTGG TGGCCAGTTT GATGAGAGGA
     GTGTCGGGTG GATTTCTCCA ACACCTTTTA CGACCTGGGT TCTGTTCCGA ACCGACAAGA TTCCTTTCAG TTCTCTCACC ACCGGTCAAA CTACTCTCCT
                                          YF17D partial signal
                                 ----------------------------------------
                                                                  TBEV partial signal
                                                                 ------------------------
           C protein                                                                  prM protein
     ----------------------                                                        ----------------
     L  S  S  R  K  R  R  S  H  D  V  L  T  V  Q  F  L  I  L  G  M  L  G  M  T  I  A  A  T  V  R  K  E  R
 401 TTGTCCTCAA GGAAACGCCG TTCCCATGAT GTTCTGACTG TGCAATTCCT AATTTTGGGC ATGCTGGGCA TGACAATCGC AGCTACGGTT CGCAAGGAAA
     AACAGGAGTT CCTTTGCGGC AAGGGTACTA CAAGACTGAC ACGTTAAGGA TTAAAACCCG TACGACCCGT ACTGTTAGCG TCGATGCCAA GCGTTCCTTT
                                                prM protein
     ----------------------------------------------------------------------------------------------------
     .  D  G  S  T  V  I  R  A  E  G  K  D  A  A  T  Q  V  R  V  E  N  G  T  C  V  I  L  A  T  D  M  G  S
 501 GAGACGGCAG TACGGTCATA CGCGCGGAAG GTAAGGATGC CGCTACCCAA GTGAGAGTGG AAAATGGTAC CTGCGTCATT CTGGCCACCG ACATGGGCTC
     CTCTGCCGTC ATGCCAGTAT GCGCGCCTTC CATTCCTACG GCGATGGGTT CACTCTCACC TTTTACCATG GACGCAGTAA GACCGGTGGC TGTACCCGAG
                                                prM protein
     ----------------------------------------------------------------------------------------------------
     .  W  C  D  D  S  L  S  Y  E  C  V  T  I  D  Q  G  E  E  P  V  D  V  D  C  F  C  R  N  V  D  G  V  Y
 601 TTGGTGTGAT GATAGCCTTT CTTATGAGTG CGTAACCATA GATCAAGGTG AGGAACCTGT TGACGTTGAT TGCTTCTGCC GAAACGTGGA TGGGGTGTAT
     AACCACACTA CTATCGGAAA GAATACTCAC GCATTGGTAT CTAGTTCCAC TCCTTGGACA ACTGCAACTA ACGAAGACGG CTTTGCACCT ACCCCACATA
                                                prM protein
     ----------------------------------------------------------------------------------------------------
     L  E  Y  G  R  C  G  K  Q  E  G  S  R  T  R  R  S  V  L  I  P  S  H  A  Q  G  E  L  T  G  R  G  H  K
 701 CTCGAATATG GACGGTGTGG TAAACAAGAA GGAAGCAGAA CCAGACGCTC AGTGCTTATA CCCTCCCACG CTCAAGGAGA GCTGACCGGA CGGGGACATA
     GAGCTTATAC CTGCCACACC ATTTGTTCTT CCTTCGTCTT GGTCTGCGAG TCACGAATAT GGGAGGGTGC GAGTTCCTCT CGACTGGCCT GCCCCTGTAT
```

prM protein

· W L E G D S L R T H L T R V E G W V W K N R L L A L A M V T V V N ·
801 AATGGTTGGA GGGCGACTCA CTCCGAACAC ATTTGACCCG CGTCGAGGGC TGGGTCTGGA AAAATCGGCT GTTGGCCCTC GCTATGGTGA CAGTCGTTTG
TTACCAACCT CCCGCTGAGT GAGGCTTGTG TAAACTGGGC GCAGCTCCCG ACCCAGACCT TTTTAGCCGA CAACCGGGAG CGATACCACT GTCAGCAAAC

Hypr E protein

prM protein

· L T L E S V V T R V A V L V V L L C L A P V Y A S R C T H L E N R
901 GCTCACGCTG GAGTCTGTGG TTACTCGCGT GGCAGTGCTG GTGGTGCTCC TCTGTCTTGC CCCTGTCTAC GCGTCCAGGT GTACTCATTT GGAAAACAGA
CGAGTGCGAC CTCAGACACC AATGAGCGCA CCGTCACGAC CACCACGAGG AGACAGAACG GGGACAGATG CGCAGGTCCA CATGAGTAAA CCTTTTGTCT

Hypr E protein

D F V T G T Q G T T R V T L V L E L G G C V T I T A E G K P S M D V ·
1001 GATTTTGTCA CCGGCACCCA GGGGACGACT CGGGTAACCC TGGTGCTTGA ACTGGGTGGT TGCGTTACTA TTACCGCTGA GGGCAAACCC TCTATGGATG
CTAAAACAGT GGCCGTGGGT CCCCTGCTGA GCCCATTGGG ACCACGAACT TGACCCACCA ACGCAATGAT AATGGCGACT CCCGTTTGGG AGATACCTAC

Hypr E protein

· W L D A I Y Q E N P A Q T R E Y C L H A K L S D T K V A A R C P T ·
1101 TGTGGCTGGA TGCAATCTAT CAGGAGAATC CCGCACAAAC CAGGGAATAT TGCCTTCACG CAAAGCTGTC CGATACAAAG GTCGCGGCTA GGTGCCCAAC
ACACCGACCT ACGTTAGATA GTCCTCTTAG GGCGTGTTTG GTCCCTTATA ACGGAAGTGC GTTTCGACAG GCTATGTTTC CAGCGCCGAT CCACGGGTTG

Hypr E protein

· M G P A T L A E E H Q G G T V C K R D Q S D R G W G N H C G L F G
1201 AATGGGACCG GCCACCCTGG CGGAGGAACA TCAGGGAGGT ACAGTGTGCA AACGGGACCA GAGTGATAGA GGCTGGGGTA ATCACTGCGG CCTGTTCGGC
TTACCCTGGC CGGTGGGACC GCCTCCTTGT AGTCCCTCCA TGTCACACGT TTGCCCTGGT CTCACTATCT CCGACCCCAT TAGTGACGCC GGACAAGCCG

Hypr E protein

K G S I V A C V K A A C E A K K K A T G H V Y D A N K I V Y T V K V ·
1301 AAAGGAAGTA TTGTCGCTTG CGTCAAGGCA GCCTGTGAGG CCAAAAAGAA GGCTACTGGG CACGTCTATG ACGCCAACAA GATCGTTTAT ACAGTGAAAG
TTTCCTTCAT AACAGCGAAC GCAGTTCCGT CGGACACTCC GGTTTTTCTT CCGATGACCC GTGCAGATAC TGCGGTTGTT CTAGCAAATA TGTCACTTTC

Hypr E protein

· E P H T G D Y V A A N E T H S G R K T A S F T V S S E K T I L T N ·
1401 TGGAACCACA CACAGGGGAT TACGTGGCGG CCAACGAGAC TCATTCCGGT CGCAAAACGG CCAGCTTCAC CGTGTCATCC GAAAAGACCA TCCTCACTAT
ACCTTGGTGT GTGTCCCCTA ATGCACCGCC GGTTGCTCTG AGTAAGGCCA GCGTTTTGCC GGTCGAAGTG GCACAGTAGG CTTTTCTGGT AGGAGTGATA

Hypr E protein

· G E Y G D V S L L C R V A S G V D L A Q T V I L E L D K T V E H L
1501 GGGGGAGTAT GGCGACGTTT CTCTGCTCTG CCGGGTGGCT AGCGGAGTCG ACCTGGCCCA GACAGTCATC CTGGAACTGG ATAAAACAGT TGAGCATCTG
CCCCCTCATA CCGCTGCAAA GAGACGAGAC GGCCCACCGA TCGCCTCAGC TGGACCGGGT CTGTCAGTAG GACCTTGACC TATTTTGTCA ACTCGTAGAC

Hypr E protein

P T A W Q V H E D W F N D L A L P W K H E G A R N W N N A E R L V E ·
1601 CCTACCGCTT GGCAGGTGCA CAGGGATTGG TTTAACGACC TTGCCCTGCC ATGGAAACAT GAAGGAGCGA GAAACTGGAA TAATGCAGAG CGACTCGTAG
GGATGGCGAA CCGTCCACGT GTCCCTAACC AAATTGCTGG AACGGGACGG TACCTTTGTA CTTCCTCGCT CTTTGACCTT ATTACGTCTC GCTGAGCATC

Hypr E protein

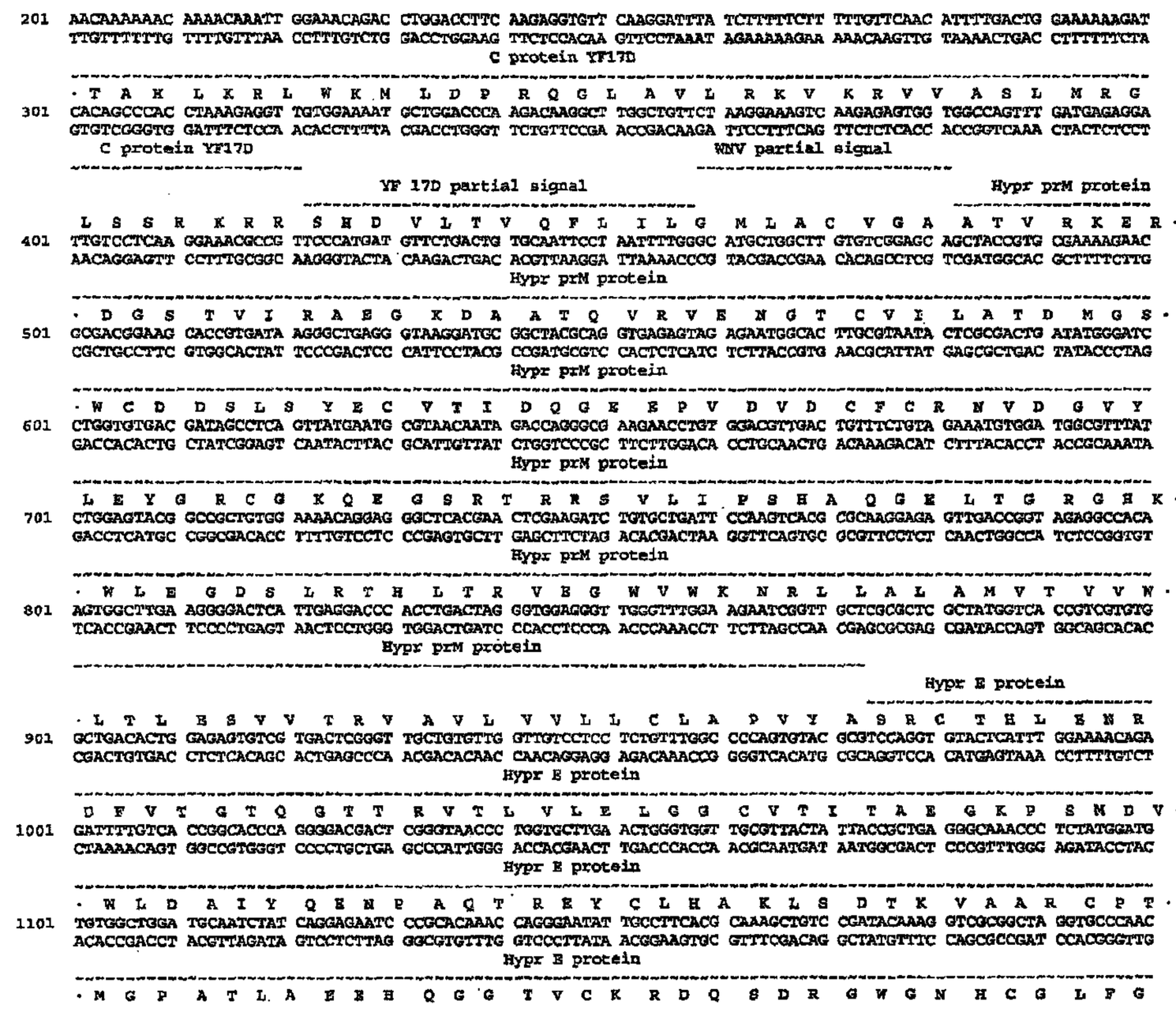
· F G A P H A V K M D V T N L G D Q T G V L L K A L A G V P V A H I ·
1701 AATTCGGTGC CCCTCATGCC GTGAAGATGG ACGTCTACAA TCTGGGTGAT CAGACCGGCG TTCTCCTTAA AGCTCTCGCT GGCGTACCAG TTGCCCACAT
TTAAGCCACG GGGAGTACGG CACTTCTACC TGCAGATGTT AGACCCACTA GTCTGGCCGC AAGAGGAATT TCGAGAGCGA CCGCATGGTC AACGGGTGTA

Hypr E protein

```
                 E  G  T  K  Y  H  L  K  S  G  H  V  T  C  E  V  G  L  E  K  L  K  M  K  G  L  T  Y  T  M  C  D  K
1801  CGAAGGAACG AAGTACCACC TGAAGTCAGG CCATGTAACT TGCGAGGTGG GCCTGGAGAA GTTGAAAATG AAAGGTCTTA CGTACACAAT GTGTGACAAG
      GCTTCCTTGC TTCATGGTGG ACTTCAGTCC GGTACATTGA ACGCTCCACC CGGACCTCTT CAACTTTTAC TTTCCAGAAT GCATGTGTTA CACACTGTTC
                                                Hypr E protein
     ---------------------------------------------------------------------------------------------------------------
                 T  K  P  T  W  K  R  A  P  T  D  S  G  H  D  T  V  V  M  E  V  T  F  S  G  T  K  P  C  R  I  P  V  R
1901  ACCAAGTTCA CATGGAAGAG GGCCCCCACA GATAGCGCCC ACGATACTGT GGTGATGGAG GTGACCTTTT CTGGAACAAA ACCCTGCAGA ATACCCGTGC
      TGGTTCAAGT GTACCTTCTC CCGGGGGTGT CTATCGCCGG TGCTATGACA CCACTACCTC CACTGGAAAA GACCTTGTTT TGGGACGTCT TATGGGCACG
                                                Hypr E protein
     ---------------------------------------------------------------------------------------------------------------
                 A  V  A  H  G  S  P  D  V  N  V  A  M  L  I  T  P  N  P  T  I  E  N  N  G  G  G  F  I  E  M  Q  L
2001  GGGCTGTAGC TCACGGATCT CCCGATGTCA ATGTTGCTAT GCTGATTACA CCTAACCCTA CCATCGAGAA TAACGGTGGT GGTTTTATTG AGATGCAGCT
      CCCGACATCG AGTGCCTAGA GGGCTACAGT TACAACGATA CGACTAATGT GGATTGGGAT GGTAGCTCTT ATTGCCACCA CCAAAATAAC TCTACGTCGA
                                                Hypr E protein
     ---------------------------------------------------------------------------------------------------------------
                 P  P  G  D  N  I  I  Y  V  G  E  L  S  Y  Q  W  F  Q  K  G  S  S  I  G  R  V  F  Q  K  T  K  K  G
2101  TCCGCCAGGC GATAACATCA TCTACGTGGG CGAACTCTCT TACCAGTGGT TTCAGAAAGG GAGTTCAATT GGGCGGGTCT TCCAAAAAAC GAAGAAGGGA
      AGGCGGTCCG CTATTGTAGT AGATGCACCC GCTTGAGAGA ATGGTCACCA AAGTCTTTCC CTCAAGTTAA CCCGCCCAGA AGGTTTTTTG CTTCTTCCCT
                                                Hypr E protein
     ---------------------------------------------------------------------------------------------------------------
                 I  E  R  L  T  V  I  G  E  H  A  W  D  F  G  S  A  G  G  F  L  S  S  I  G  K  A  L  H  T  V  L  G  G
2201  ATCGAACGAT TGACGGTTAT CGGCGAGCAC GCATGGGATT TTGGTTCCGC AGGGGGATTC CTGTCTTCTA TTGGTAAGGC ACTGCATACC GTGCTGGGGG
      TAGCTTGCTA ACTGCCAATA GCCGCTCGTG CGTACCCTAA AACCAAGGCG TCCCCCTAAG GACAGAAGAT AACCATTCCG TGACGTATGG CACGACCCCC
                                                Hypr E protein
     ---------------------------------------------------------------------------------------------------------------
                 A  F  N  S  I  F  G  G  V  G  F  L  P  K  L  L  L  G  V  A  L  A  W  L  G  L  N  M  R  N  P  T  M
2301  GCGCATTCAA TTCTATTTTC GGGGGCGTGG GGTTCCTGCC TAAACTCCTG CTGGGAGTAG CCCTGGCCTG GTTGGGACTG AATATGCGGA ATCCGACGAT
      CGCGTAAGTT AAGATAAAAG CCCCCGCACC CCAAGGACGG ATTTGAGGAC GACCCTCATC GGGACCCGAC CAACCCTGAC TTATACGCCT TAGGCTGCTA
                                Hypr E protein
     ---------------------------------------------------------------
                                                                                NS1 gene of YF17D
                                                                      ---------------------------------------------
                 S  M  S  F  L  L  A  G  V  L  V  L  A  M  T  L  G  V  G  A  D  Q  G  C  A  I  N  F  G  K  R  E  L
2401  GTCCATGTCA TTCCTCTTGG CCGGCGTGCT TGTACTGGCC ATGACACTGG GCGTTGGCGC CGATCAAGGA TGCGCCATCA ACTTTGGCAA GAGAGAGCTC
      CAGGTACAGT AAGGAGAACC GGCCGCACGA ACATGACCGG TACTGTGACC CGCAACCGCG GCTAGTTCCT ACGCGGTAGT TGAAACCGTT CTCTCTCGAG
```

CV-TBEV Hypr with YFV/WNV chimeric signal (p45)

```
                                                S' UTR
     ---------------------------------------------------------------------------------------------------------------
1     AGTAAATCCT GTGTGCTAAT TGAGGTGCAT TGGTCTGCAA ATCGAGTTGC TAGGCAATAA ACACATTTGG ATTAATTTTA ATCGTTCGTT GAGCGATTAG
      TCATTTAGGA CACACGATTA ACTCCACGTA ACCAGACGTT TAGCTCAACG ATCCGTTATT TGTGTAAACC TAATTAAAAT TAGCAAGCAA CTCGCTAATC
            S' UTR
     ---------------------
                                                C protein YF17D
                                      ---------------------------------------------------------------------------
                       M  S  G  R  K  A  Q  G  K  T  L  G  V  N  M  V  R  R  G  V  R  S  L  S  N  K  I  K
101   CAGAGAACTG ACCAGAACAT GTCTGGTCGT AAAGCTCAGG GAAAAACCCT GGGCGTCAAT ATGGTACGAC GAGGAGTTCG CTCCTTGTCA AACAAAATAA
      GTCTCTTGAC TGGTCTTGTA CAGACCAGCA TTTCGAGTCC CTTTTTGGGA CCCGCAGTTA TACCATGCTG CTCCTCAAGC GAGGAACAGT TTGTTTTATT
                                                C protein YF17D
     ---------------------------------------------------------------------------------------------------------------
                 Q  K  T  K  Q  I  G  N  R  P  G  P  S  R  G  V  Q  G  F  I  F  F  F  L  Y  N  I  L  T  G  K  K  I
```


C protein YF17D
WNV partial signal
YF 17D partial signal
Hypr prM protein
Hypr E protein

```
1201  AATGGGACCG GCCACCCTGG CGGAGGAACA TCAGGGAGGT ACAGTGTGCA AACGGGACCA GAGTGATAGA GGCTGGGGTA ATCACTGCGG CCTGTTCGGC
      TTACCCTGGC CGGTGGGACC GCCTCCTTGT AGTCCCTCCA TGTCACACGT TTGCCCTGGT CTCACTATCT CCGACCCCAT TAGTGACGCC GGACAAGCCG
                                                  Hypr E protein
      ------------------------------------------------------------------------------------------------------------
       K  G  S  I  V  A  C  V  K  A  A  C  E  A  K  K  K  A  T  G  H  V  Y  D  A  N  K  I  V  Y  T  V  K  V
1301  AAAGGAAGTA TTGTCGCTTG CGTCAAGGCA GCCTGTGAGG CCAAAAAGAA GGCTACTGGG CACGTCTATG ACGCCAACAA GATCGTTTAT ACAGTGAAAG
      TTTCCTTCAT AACAGCGAAC GCAGTTCCGT CGGACACTCC GGTTTTTCTT CCGATGACCC GTGCAGATAC TGCGGTTGTT CTAGCAAATA TGTCACTTTC
                                                  Hypr E protein
      ------------------------------------------------------------------------------------------------------------
       E  P  H  T  G  D  Y  V  A  A  N  E  T  H  S  G  R  K  T  A  S  F  T  V  S  S  E  K  T  I  L  T  M
1401  TGGAACCACA CACAGGGGAT TACGTGGCGG CCAACGAGAC TCATTCCGGT CGCAAAACGG CCAGCTTCAC CGTGTCATCC GAAAAGACCA TCCTCACTAT
      ACCTTGGTGT GTGTCCCCTA ATGCACCGCC GGTTGCTCTG AGTAAGGCCA GCGTTTTGCC GGTCGAAGTG GCACAGTAGG CTTTTCTGGT AGGAGTGATA
                                                  Hypr E protein
      ------------------------------------------------------------------------------------------------------------
       G  E  Y  G  D  V  S  L  L  C  R  V  A  S  G  V  D  L  A  Q  T  V  I  L  E  L  D  K  T  V  E  H  L
1501  GGGGGAGTAT GGCGACGTTT CTCTGCTCTG CCGGGTGGCT AGCGGAGTCG ACCTGGCCCA GACAGTCATC CTGGAACTGG ATAAAACAGT TGAGCATCTG
      CCCCCTCATA CCGCTGCAAA GAGACGAGAC GGCCCACCGA TCGCCTCAGC TGGACCGGGT CTGTCAGTAG GACCTTGACC TATTTTGTCA ACTCGTAGAC
                                                  Hypr E protein
      ------------------------------------------------------------------------------------------------------------
       P  T  A  W  Q  V  H  R  D  W  F  N  D  L  A  L  P  W  K  H  E  G  A  R  N  W  N  N  A  E  R  L  V  E
1601  CCTACCGCTT GGCAGGTGCA CAGGGATTGG TTTAACGACC TTGCCCTGCC ATGGAAACAT GAAGGAGCGA GAAACTGGAA TAATGCAGAG CGACTCGTAG
      GGATGGCGAA CCGTCCACGT GTCCCTAACC AAATTGCTGG AACGGGACGG TACCTTTGTA CTTCCTCGCT CTTTGACCTT ATTACGTCTC GCTGAGCATC
                                                  Hypr E protein
      ------------------------------------------------------------------------------------------------------------
       P  G  A  P  H  A  V  K  M  D  V  Y  N  L  G  D  Q  T  G  V  L  L  K  A  L  A  G  V  P  V  A  H  I
1701  AATTCGGTGC CCCTCATGCC GTGAAGATGG ACGTCTACAA TCTGGGTGAT CAGACCGGCG TTCTCCTTAA AGCTCTCGCT GGCGTACCAG TTGCCCACAT
      TTAAGCCACG GGGAGTACGG CACTTCTACC TGCAGATGTT AGACCCACTA GTCTGGCCGC AAGAGGAATT TCGAGAGCGA CCGCATGGTC AACGGGTGTA
                                                  Hypr E protein
      ------------------------------------------------------------------------------------------------------------
       E  G  T  K  Y  H  L  K  S  G  H  V  T  C  E  V  G  L  E  K  L  K  M  K  G  L  T  Y  T  M  C  D  K
1801  CGAAGGAACG AAGTACCACC TGAAGTCAGG CCATGTAACT TGCGAGGTGG GCCTGGAGAA GTTGAAAATG AAAGGTCTTA CGTACACAAT GTGTGACAAG
      GCTTCCTTGC TTCATGGTGG ACTTCAGTCC GGTACATTGA ACGCTCCACC CGGACCTCTT CAACTTTTAC TTTCCAGAAT GCATGTGTTA CACACTGTTC
                                                  Hypr E protein
      ------------------------------------------------------------------------------------------------------------
       T  K  F  T  W  K  R  A  P  T  D  S  G  H  D  T  V  V  M  E  V  T  F  S  G  T  K  P  C  R  I  P  V  R
1901  ACCAAGTTCA CATGGAAGAG GGCCCCACA GATAGCGGCC ACGATACTGT GGTGATGGAG GTGACCTTTT CTGGAACAAA ACCCTGCAGA ATACCCGTGC
      TGGTTCAAGT GTACCTTCTC CCGGGGGTGT CTATCGCCGG TGCTATGACA CCACTACCTC CACTGGAAAA GACCTTGTTT TGGGACGTCT TATGGGCACG
                                                  Hypr E protein
      ------------------------------------------------------------------------------------------------------------
       A  V  A  H  G  S  P  D  V  N  V  A  M  L  I  T  P  N  P  T  I  E  N  N  G  G  G  F  I  E  M  Q  L
2001  GGGCTGTAGC TCACGGATCT CCCGATGTCA ATGTTGCTAT GCTGATTACA CCTAACCCTA CCATCGAGAA TAACGGTGGT GGTTTTATTG AGATGCAGCT
      CCCGACATCG AGTGCCTAGA GGGCTACAGT TACAACGATA CGACTAATGT GGATTGGGAT GGTAGCTCTT ATTGCCACCA CCAAAATAAC TCTACGTCGA
                                                  Hypr E protein
      ------------------------------------------------------------------------------------------------------------
       P  P  G  D  N  I  I  Y  V  G  E  L  S  Y  Q  W  F  Q  K  G  S  S  I  G  R  V  F  Q  K  T  K  K  G
2101  TCCGCCAGGC GATAACATCA TCTACGTGGG CGAACTCTCT TACCAGTGGT TTCAGAAAGG GAGTTCAATT GGGCGGGTCT TCCAAAAAAC GAAGAAGGGA
      AGGCGGTCCG CTATTGTAGT AGATGCACCC GCTTGAGAGA ATGGTCACCA AAGTCTTTCC CTCAAGTTAA CCCGCCCAGA AGGTTTTTTG CTTCTTCCCT
                                                  Hypr E protein
      ------------------------------------------------------------------------------------------------------------
       I  E  R  L  T  V  I  G  E  H  A  W  D  F  G  S  A  G  G  F  L  S  S  I  G  K  A  L  H  T  V  L  G  G
2201  ATCGAACGAT TGACGGTTAT CGGCGAGCAC GCATGGGATT TTGGTTCGGC AGGGGGATTC CTGTCTTCTA TTGGTAAGGC ACTGCATACC GTGCTGGGGG
      TAGCTTGCTA ACTGCCAATA GCCGCTCGTG CGTACCCTAA AACCAAGGCG TCCCCCTAAG GACAGAAGAT AACCATTCCG TGACGTATGG CACGACCCCC
                                                  Hypr E protein
      ------------------------------------------------------------------------------------------------------------
```

```
                                                
```

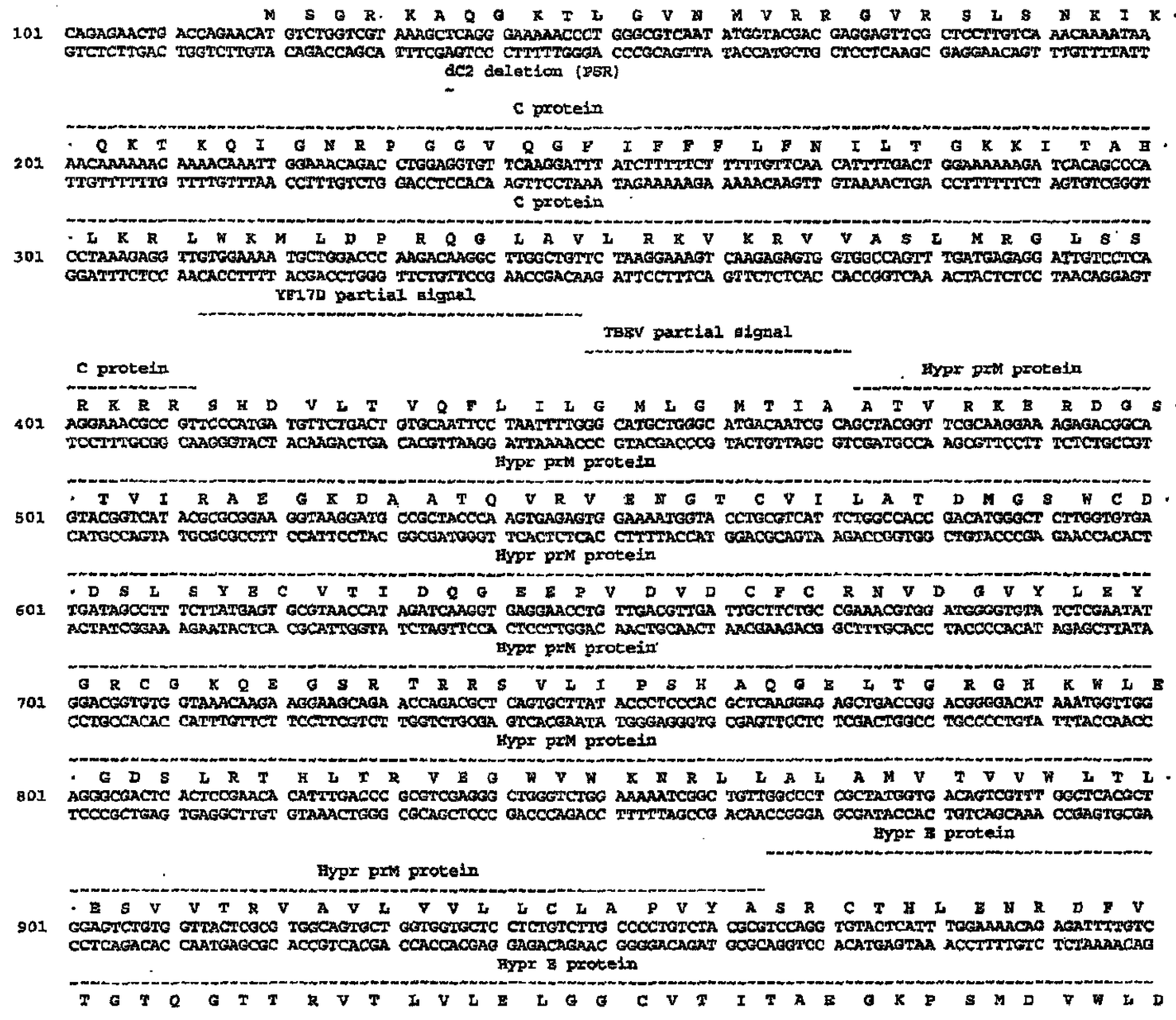
prM protein Langat E5

```
 L E Y G  R C G  R R E  G S R S  R R S  V L I  P S H A  Q R D  L T G  R G H Q
701 CTGGAATATG GACGATGTGG CCGGCGAGAA GGCTCCAGGA GTCGGAGATC CGTGTTGATC CCTTCACATG CGCAGCGCGA TCTGACAGGG AGGGGTCACC
    GACCTTATAC CTGCTACACC GGCCGCTCTT CCGAGGTCCT CAGCCTCTAG GCACAACTAG GGAAGTGTAC GCGTCGCGCT AGACTGTCCC TCCCCAGTGG
                                    prM protein Langat E5

 W L E  G E A  V K A  H L T  R V E  G W V  W K N  K L F  T L S  L V M  V A W
801 AGTGGCTCGA AGGCGAAGCA GTCAAGGCCC ATCTGACTCG CGTTGAAGGC TGGGTGTGGA AAAACAAACT CTTTACCCTT AGCCTGGTGA TGGTCGCGTG
    TCACCGAGCT TCCGCTTCGT CAGTTCCGGG TAGACTGAGC GCAACTTCCG ACCCACACCT TTTTGTTTGA GAAATGGGAA TCGGACCACT ACCAGCGCAC
                                    prM protein Langat E5

                                                                    E protein Langat E5
 L M V  D G L L  P R I  L I V  V V A L  A L A  P A Y  A S R C  T H L  E N R
901 GCTGATGGTA GACGGACTCC TTCCCCGCAT TCTCATTGTT GTGGTGGCTC TCGCGCTCGC CCCTGCATAC GCGTCCAGGT GTACGCACCT CGAAAATCGA
    CGACTACCAT CTGCCTGAGG AAGGGGCGTA AGAGTAACAA CACCACCGAG AGCGCGAGCG GGGACGTATG CGCAGGTCCA CATGCGTGGA GCTTTTAGCT
                                    E protein Langat E5

 D F V T  G V Q  G T T  R L T L  V L E  L G G  C V T V  T A D  G K P  S L D V
1001 GATTTCGTCA CAGGCGTCCA AGGTACTACC CGGCTCACCC TCGTGCTGGA GCTGGGAGGC TGTGTCACTG TTACAGCCGA CGGAAAACCT AGTCTGGATG
     CTAAAGCAGT GTCCGCAGGT TCCATGATGG GCCGAGTGGG AGCACGACCT CGACCCTCCG ACACAGTGAC AATGTCGGCT GCCTTTTGGA TCAGACCTAC
                                    E protein Langat E5

 W L D  S I Y  Q E S  P A Q  T R E  Y C L  H A K  L T G  T K V  A A R  C P T
1101 TGTGGCTGGA CTCCATCTAT CAGGAGAGCC CGGCACAGAC CAGGGAGTAC TGCCTCCACG CTAAGCTGAC TGGGACAAAG GTAGCCGCAA GATGTCCCAC
     ACACCGACCT GAGGTAGATA GTCCTCTCGG GCCGTGTCTG GTCCCTCATG ACGGAGGTGC GATTCGACTG ACCCTGTTTC CATCGGCGTT CTACAGGGTG
                                    E protein Langat E5

 M G P  A T L P  E E H  Q S G  T V C K  R D Q  S D R  G W G N  H C G  L F G
1201 AATGGGGCCT GCCACCTTGC CCGAGGAACA CCAATCCGGT ACGGTATGCA AGCGAGATCA GTCTGATCGC GGATGGGGGA ATCATTGCGG CCTCTTCGGT
     TTACCCCGGA CGGTGGAACG GGCTCCTTGT GGTTAGGCCA TGCCATACGT TCGCTCTAGT CAGACTAGCG CCTACCCCCT TAGTAACGCC GGAGAAGCCA
                                    E protein Langat E5

 K G S  I V T C  V K V  T C E  D K K K  A T G  H V Y  D V N K  I T Y  T I K V
1301 AAAGGCAGCA TTGTCACTTG CGTGAAGGTG ACATGCGAGG ACAAGAAGAA GGCCACAGGT CATGTATATG ATGTGAACAA AATCACATAT ACCATTAAGG
     TTTCCGTCGT AACAGTGAAC GCACTTCCAC TGTACGCTCC TGTTCTTCTT CCGGTGTCCA GTACATATAC TACACTTGTT TTAGTGTATA TGGTAATTCC
                                    E protein Langat E5

 E P H  T G E  F V A  A N E T  H S G  R K S  A S P T  V S S  E K T I  L T L
1401 TAGAACCACA TACAGGGGAA TTCGTGGCAC CAAACGAGAC TCATAGCGGA CGAAAGTCCG CCTCCTTCAC CGTCTCCTCC GAGAAAACAA TCCTGACCCT
     ATCTTGGTGT ATGTCCCCTT AAGCACCGTC GTTTGCTCTG AGTATCGCCT GCTTTCAGGC GGAGGAAGTG GCAGAGGAGG CTCTTTTGTT AGGACTGGGA
                                    E protein Langat E5

 G D Y  G D V S  L L C  R V A  S G V D  L A Q  T V V  L A L D  K T H  E H L
1501 CGGAGACTAC GGCGACGTAT CTTTGCTGTG CAGGGTGGCC AGCGGCGTGG ACCTTGCTCA GACAGTCGTG TTGGCCCTGG ACAAGACACA TGAGCACTTG
     GCCTCTGATG CCGCTGCATA GAAACGACAC GTCCCACCGG TCGCCGCACC TGGAACGAGT CTGTCAGCAC AACCGGGACC TGTTCTGTGT ACTCGTGAAC
                                    E protein Langat E5

 P T A W  Q V H  R D N  F N D L  A L P  W K H  D G A E  A W N  E A G  R L V E
1601 CCAACAGCCT GGCAGGTGCA CAGGGACTGG TTTAACGACC TGGCGCTCCC GTGGAAACAT GACGGCGCTG AAGCATGGAA TGAGGCAGGG AGACTGGTGG
     GGTTGTCGGA CCGTCCACGT GTCCCTGACC AAATTGCTGG ACCGCGAGGG CACCTTTGTA CTGCCGCGAC TTCGTACCTT ACTCCGTCCC TCTGACCACC
                                    E protein Langat E5
```

```
      · F G T  P H A  V K M D  V F N  L G D  Q T G V  L L K  S L A  G V P V  A S I ·
1701  AATTTGGAAC CCCACACGCC GTAAAGATGG ACGTTTTCAA TCTTGGTGAC CAGACAGGGG TGCTCCTGAA ATCACTGGCG GGCGTGCCTG TAGCCAGCAT
      TTAAACCTTG GGGTGTGCGG CATTTCTACC TGCAAAAGTT AGAACCACTG GTCTGTCCCC ACGAGGACTT TAGTGACCGC CCGCACGGAC ATCGGTCGTA
                                        E protein Langat E5
      ------------------------------------------------------------------------------------------------------------
      · E G T  K Y H L  K S G  H V T  C E V G  L E K  L K M  K G L T  Y T V  C D K
1801  CGAGGGCACA AAGTATCACC TGAAGTCTGG GCATGTAACC TGCGAAGTGG GCCTGGAAAA GCTGAAGATG AAAGGACTTA CGTACACTGT TTGTGATAAG
      GCTCCCGTGT TTCATAGTGG ACTTCAGACC CGTACATTGG ACGCTTCACC CGGACCTTTT CGACTTCTAC TTTCCTGAAT GCATGTGACA AACACTATTC
                                        E protein Langat E5
      ------------------------------------------------------------------------------------------------------------
       T K F T  W K R  A P T  D S G H  D T V  V M E  V G F S  G T R  P C R  I P V R ·
1901  ACCAAGTTTA CATGGAAGCG AGCCCCAACG GATTCCGGCC ATGATACCGT CGTGATGGAG GTTGGTTTCT CCGGCACCAG ACCATGTAGA ATACCAGTGA
      TGGTTCAAAT GTACCTTCGC TCGGGGTTGC CTAAGGCCGG TACTATGGCA GCACTACCTC CAACCAAAGA GGCCGTGGTC TGGTACATCT TATGGTCACT
                                        E protein Langat E5
      ------------------------------------------------------------------------------------------------------------
      · A V A  H G V  P E V N  V A M  L I T  P N P T  M E N  N G G  G F I E  M Q L ·
2001  GAGCTGTCGC CCACGGTGTA CCCGAGGTAA ACGTGGCCAT GCTGATTACA CCGAATCCCA CTATGGAGAA CAATGGCGGA GGGTTCATCG AAATGCAGCT
      CTCGACAGCG GGTGCCACAT GGGCTCCATT TGCACCGGTA CGACTAATGT GGCTTAGGGT GATACCTCTT GTTACCGCCT CCCAAGTAGC TTTACGTCGA
                                        E protein Langat E5
      ------------------------------------------------------------------------------------------------------------
      · P P C  D N I I  Y V G  D L D  H Q W F  Q K G  S S I  G R V L  Q K T  R K G
2101  GCCGCCTGGA GACAACATCA TTTATGTCGG CGACCTCGAT CATCAATGGT TCCAGAAAGG GTCTTCCATC GGCCGCGTCC TTCAGAAGAC ACGAAAAGGC
      CGGCGGACCT CTGTTGTAGT AAATACAGCC GCTGGAGCTA GTAGTTACCA AGGTCTTTCC CAGAAGGTAG CCGGCGCAGG AAGTCTTCTG TGCTTTTCCG
                                        E protein Langat E5
      ------------------------------------------------------------------------------------------------------------
       I E R L  T V L  G E H  A W D F  G S V  G G V  M T S I  G R A  M H T  V L G G ·
2201  ATTGAAAGAC TTACAGTCCT GGGCGAACAT GCCTGGGACT TCGGGTCAGT TGGCGGGGTA ATGACAAGCA TAGGCAGAGC TATGCACACC GTTCTCGGTG
      TAACTTTCTG AATGTCAGGA CCCGCTTGTA CGGACCCTGA AGCCCAGTCA ACCGCCCCAT TACTGTTCGT ATCCGTCTCG ATACGTGTGG CAAGAGCCAC
                                        E protein Langat E5
      ------------------------------------------------------------------------------------------------------------
      · A F N  T L L  G G V G  F L P  E I L  L G V A  M A W  L G L  N M R N  P T L ·
2301  GGGCATTTAA TACTCTGTTG GGTGGCGTGG GTTTTCTTCC GAAAATCCTG CTCGGTGTCG CAATGGCCTG GCTTGGACTG AATATGCGCA ATCCTACACT
      CCCGTAAATT ATGAGACAAC CCACCGCACC CAAAAGAAGG CTTTTAGGAC GAGCCACAGC GTTACCGGAC CGAACCTGAC TTATACGCGT TAGGATGTGA
                                        E protein Langat E5
      ------------------------------------------------------------------------------------------------------------
                                                                                  NS1 gene of YF17D
                                                                                  ------------------------------------
      · S M G  F L L S  G G L  V L A  M T L G  V G A  D Q G  C A I N  F G K  R E L
2401  GAGTATGGGG TTTCTTCTGT CAGGAGGCCT GGTCCTGGCA ATGACTCTGG GAGTGGGCGC CGATCAAGGA TGCGCCATCA ACTTTGGCAA GAGAGAGCTC
      CTCATACCCC AAAGAAGACA GTCCTCCGGA CCAGGACCGT TACTGAGACC CTCACCCGCG GCTAGTTCCT ACGCGGTAGT TGAAACCGTT CTCTCTCGAG
```

CV-TBEV Hypr with YFV/TBEV chimeric signal and dC2 deletion in C protein (p59)

```
                                                  5' UTR
      ------------------------------------------------------------------------------------------------------------
    1 AGTAAATCCT GTGTGCTAAT TGAGGTGCAT TGGTCTGCAA ATCGAGTTGC TAGGCAATAA ACACATTTGG ATTAATTTTA ATCGTTCGTT GAGCGATTAG
      TCATTTAGGA CACACGATTA ACTCCACGTA ACCAGACGTT TAGCTCAACG ATCCGTTATT TGTGTAAACC TAATTAAAAT TAGCAAGCAA CTCGCTAATC
           5' UTR
      ----------
                                                  C protein
                                                  ------------------------------------------------------------------
```


dC2 deletion (PSR)
C protein
C protein
YF17D partial signal
TBEV partial signal
C protein
Hypr prM protein
Hypr prM protein
Hypr prM protein
Hypr prM protein
Hypr prM protein
Hypr E protein
Hypr prM protein
Hypr E protein

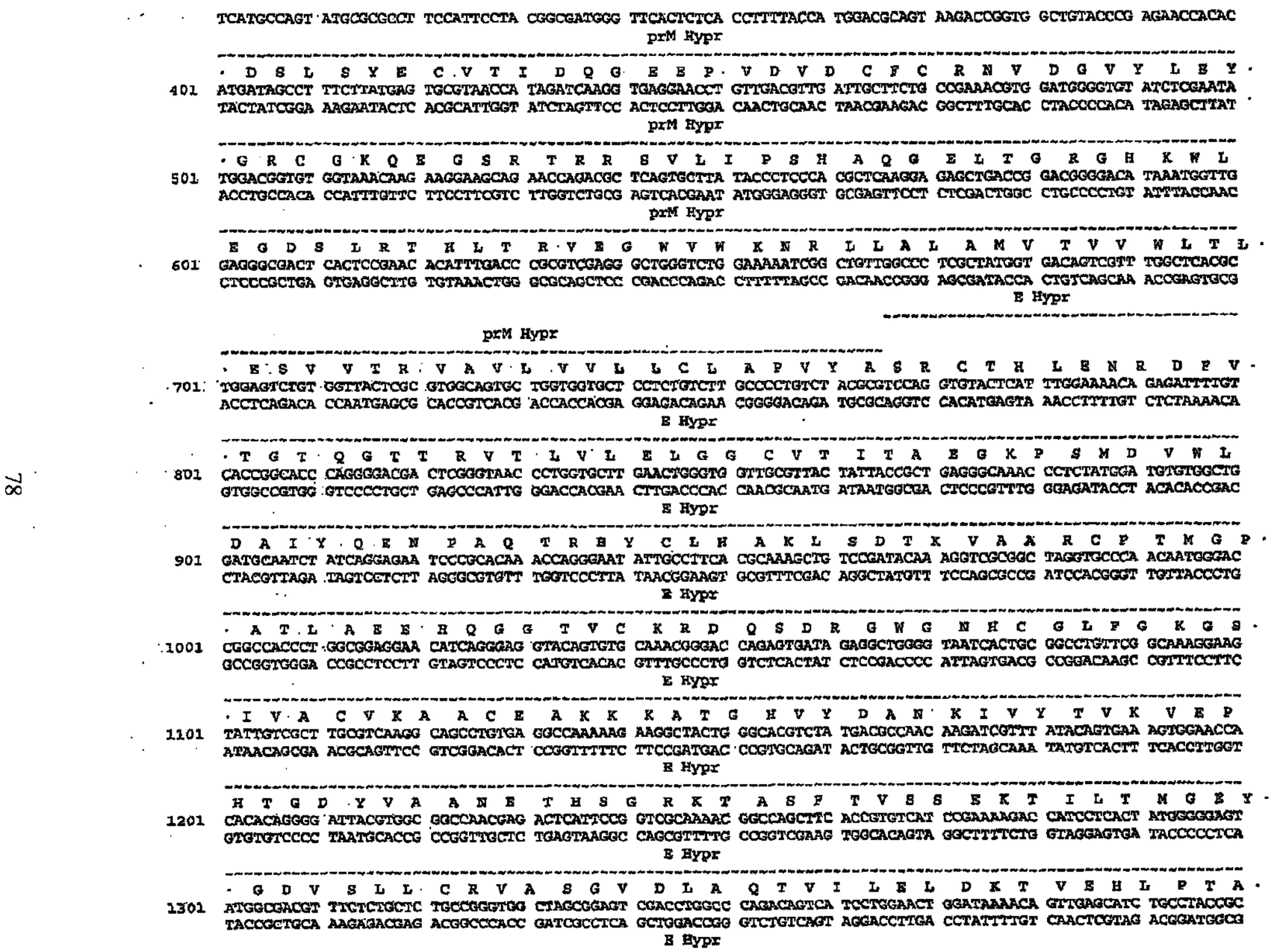
```
1001  ACCGGCACCC AGGGGACGAC TCGGGTAACC CTGGTGCTTG AACTGGGTGG TTGCGTTACT ATTACCGCTG AGGGCAAACC CTCTATGGAT GTGTGGCTGG
      TGGCCGTGGG TCCCCTGCTG AGCCCATTGG GACCACGAAC TTGACCCACC AACGCAATGA TAATGGCGAC TCCCGTTTGG GAGATACCTA CACACCGACC
                                                Hypr E protein
      ------------------------------------------------------------------------------------------------------------
      · A  I  Y  Q  E  N  P  A  Q  T  R  E  Y  C  L  H  A  K  L  S  D  T  K  V  A  A  R  C  P  T  M  G  P ·
1101  ATGCAATCTA TCAGGAGAAT CCCGCACAAA CCAGGGAATA TTGCCTTCAC GCAAAGCTGT CCGATACAAA GGTCGCGGCT AGGTGCCCAA CAATGGGACC
      TACGTTAGAT AGTCCTCTTA GGGCGTGTTT GGTCCCTTAT AACGGAAGTG CGTTTCGACA GGCTATGTTT CCAGCGCCGA TCCACGGGTT GTTACCCTGG
                                                Hypr E protein
      ------------------------------------------------------------------------------------------------------------
      · A  T  L  A  E  E  H  Q  G  G  T  V  C  K  R  D  Q  S  D  R  G  W  G  N  H  C  G  L  F  G  K  G  S
1201  GGCCACCCTG GCGGAGGAAC ATCAGGGAGG TACAGTGTGC AAACGGGACC AGAGTGATAG AGGCTGGGGT AATCACTGCG GCCTGTTCGG CAAAGGAAGT
      CCGGTGGGAC CGCCTCCTTG TAGTCCCTCC ATGTCACACG TTTGCCCTGG TCTCACTATC TCCGACCCCA TTAGTGACGC CGGACAAGCC GTTTCCTTCA
                                                Hypr E protein
      ------------------------------------------------------------------------------------------------------------
        I  V  A  C  V  K  A  A  C  E  A  K  K  K  A  T  G  H  V  Y  D  A  W  K  I  V  Y  T  V  K  V  E  P  H ·
1301  ATTGTCGCTT GCGTCAAGGC AGCCTGTGAG GCCAAAAAGA AGGCTACTGG GCACGTCTAT GACGCCAACA AGATCGTTTA TACAGTGAAA GTGGAACCAC
      TAACAGCGAA CGCAGTTCCG TCGGACACTC CGGTTTTTCT TCCGATGACC CGTGCAGATA CTGCGGTTGT TCTAGCAAAT ATGTCACTTT CACCTTGGTG
                                                Hypr E protein
      ------------------------------------------------------------------------------------------------------------
      · T  G  D  Y  V  A  A  N  E  T  H  S  G  R  K  T  A  S  F  T  V  S  S  E  K  T  I  L  T  M  G  E  Y ·
1401  ACACAGGGGA TTACGTGGCG GCCAACGAGA CTCATTCCGG TCGCAAAACG GCCAGCTTCA CCGTGTCATC CGAAAAGACC ATCCTCACTA TGGGGAGTTA
      TGTGTCCCCT AATGCACCGC CGGTTGCTCT GAGTAAGGCC AGCGTTTTGC CGGTCGAAGT GGCACAGTAG GCTTTTCTGG TAGGAGTGAT ACCCCCTCAT
                                                Hypr E protein
      ------------------------------------------------------------------------------------------------------------
      · G  D  V  S  L  L  C  R  V  A  S  G  V  D  L  A  Q  T  V  I  L  E  L  D  K  T  V  E  H  L  P  T  A
1501  TGGCGACGTT TCTCTGCTCT GCCGGGTGGC TAGCGGAGTC GACCTGGCCC AGACAGTCAT CCTGGAACTG GATAAAACAG TTGAGCATCT GCCTACCGCT
      ACCGCTGCAA AGAGACGAGA CGGCCCACCG ATCGCCTCAG CTGGACCGGG TCTGTCAGTA GGACCTTGAC CTATTTTGTC AACTCGTAGA CGGATGGCGA
                                                Hypr E protein
      ------------------------------------------------------------------------------------------------------------
        W  Q  V  H  R  D  W  F  N  D  L  A  L  P  W  K  H  E  G  A  R  N  W  N  N  A  E  R  L  V  E  F  G  A ·
1601  TGGCAGGTGC ACAGGGATTG GTTTAACGAC CTTGCCCTGC CATGGAAACA TGAAGGAGCG AGAAACTGGA ATAATGCAGA GCGACTCGTA GAATTCGGTG
      ACCGTCCACG TGTCCCTAAC CAAATTGCTG GAACGGGACG GTACCTTTGT ACTTCCTCGC TCTTTGACCT TATTACGTCT CGCTGAGCAT CTTAAGCCAC
                                                Hypr E protein
      ------------------------------------------------------------------------------------------------------------
      · P  H  A  V  K  M  D  V  Y  N  L  G  D  Q  T  G  V  L  L  K  A  L  A  G  V  P  V  A  H  I  E  G  T ·
1701  CCCCTCATGC CGTGAAGATG GACGTCTACA ATCTGGGTGA TCAGACCGGC GTTCTCCTTA AAGCTCTCGC TGGCGTACCA GTTGCCCACA TCGAAGGAAC
      GGGGAGTACG GCACTTCTAC CTGCAGATGT TAGACCCACT AGTCTGGCCG CAAGAGGAAT TTCGAGAGCG ACCGCATGGT CAACGGGTGT AGCTTCCTTG
                                                Hypr E protein
      ------------------------------------------------------------------------------------------------------------
      · K  Y  H  L  K  S  G  H  V  T  C  E  V  G  L  E  K  L  K  M  K  G  L  T  Y  T  M  C  D  K  T  K  F
1801  GAAGTACCAC CTGAAGTCAG GCCATGTAAC TTGCGAGGTG GGCCTGGAGA AGTTGAAAAT GAAAGGTCTT ACGTACACAA TGTGTGACAA GACCAAGTTC
      CTTCATGGTG GACTTCAGTC CGGTACATTG AACGCTCCAC CCGGACCTCT TCAACTTTTA CTTTCCAGAA TGCATGTGTT ACACACTGTT CTGGTTCAAG
                                                Hypr E protein
      ------------------------------------------------------------------------------------------------------------
        T  W  K  R  A  P  T  D  S  G  H  D  T  V  V  M  E  V  T  F  S  G  T  K  P  C  R  I  P  V  R  A  V  A ·
1901  ACATGGAAGA GGGCCCCCAC AGATAGCGGC CACGATACTG TGGTGATGGA GGTGACCTTT TCTGGAACAA AACCCTGCAG AATACCCGTG CGGGCTGTAG
      TGTACCTTCT CCCGGGGGTG TCTATCGCCG GTGCTATGAC ACCACTACCT CCACTGGAAA AGACCTTGTT TTGGGACGTC TTATGGGCAC GCCCGACATC
                                                Hypr E protein
      ------------------------------------------------------------------------------------------------------------
      · H  G  S  P  D  V  N  V  A  M  L  I  T  P  N  P  T  I  E  N  N  G  G  G  F  I  E  M  Q  L  P  P  G ·
2001  CTCACGGATC TCCCGATGTC AATGTTGCTA TGCTGATTAC ACCTAACCCT ACCATCGAGA ATAACGGTGG TGGTTTTATT GAGATGCAGC TTCCGCCAGG
      GAGTGCCTAG AGGGCTACAG TTACAACGAT ACGACTAATG TGGATTGGGA TGGTAGCTCT TATTGCCACC ACCAAAATAA CTCTACGTCG AAGGCGGTCC
                                                Hypr E protein
      ------------------------------------------------------------------------------------------------------------
```

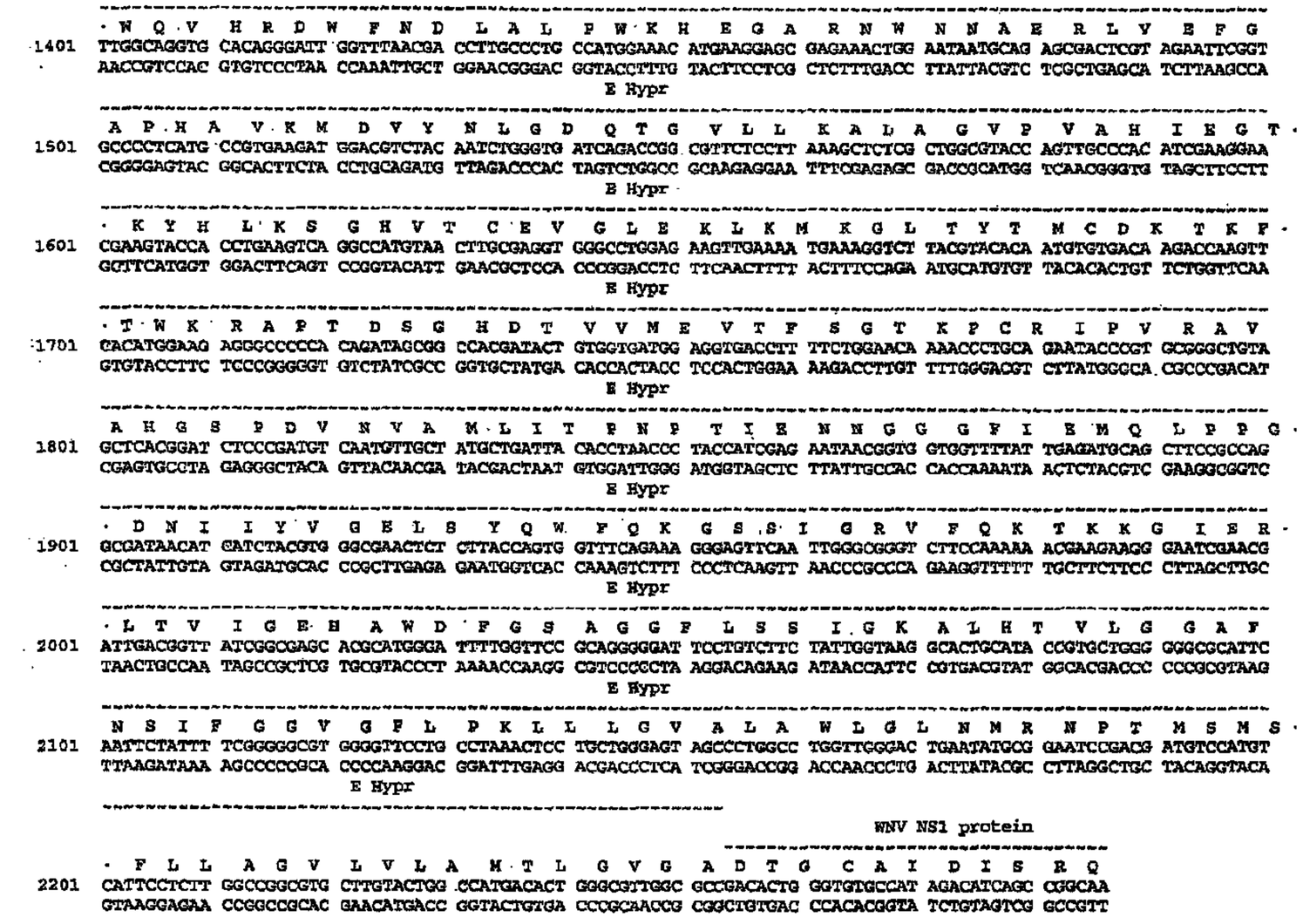
```
     · D  N  I   I  Y  V  G   E  L  S   Y  Q  W   F  Q  K  G   S  S  I   G  R  V   F  Q  K  T   K  K  G   I  E  R
2101 CGATAACATC ATCTACGTGG GCGAACTCTC TTACCAGTGG TTTCAGAAAG GGAGTTCAAT TGGGCGGGTC TTCCAAAAAA CGAAGAAGGG AATCGAACGA
     GCTATTGTAG TAGATGCACC CGCTTGAGAG AATGGTCACC AAAGTCTTTC CCTCAAGTTA ACCCGCCCAG AAGGTTTTTT GCTTCTTCCC TTAGCTTGCT
                                                Hypr E protein
     ------------------------------------------------------------------------------------------------------------
      L  T  V  I   G  E  H   A  W  D   F  G  S  A   G  G  F   L  S  S   I  G  K  A   L  H  T   V  L  G   G  A  F  N ·
2201 TTGACGGTTA TCGGCGAGCA CGCATGCGAT TTTGGTTCCG CAGGGGGATT CCTGTCTTCT ATTGGTAAGG CACTGCATAC CGTGCTGGGG GGCGCATTCA
     AACTGCCAAT AGCCGCTCGT GCGTACCCTA AAACCAAGGC GTCCCCCTAA GGACAGAAGA TAACCATTCC GTGACGTATG GCACGACCCC CCGCGTAAGT
                                                Hypr E protein
     ------------------------------------------------------------------------------------------------------------
     · S  I  F   G  G  V   G  F  L  P   K  L  L   L  G  V   A  L  A  W   L  G  L   N  M  R   N  P  T  M   S  M  S ·
2301 ATTCTATTTT CGGGGGCGTG GGGTTCCTGC CTAAACTCCT GCTGGGAGTA GCCCTGGCCT GGTTGGGACT GAATATGCGG AATCCGACGA TGTCCATGTC
     TAAGATAAAA GCCCCCGCAC CCCAAGGACG GATTTGAGGA CGACCCTCAT CGGGACCGGA CCAACCCTGA CTTATACGCC TTAGGCTGCT ACAGGTACAG
                        Hypr E protein
     ------------------------------------------------------
                                                                    NS1 gene of YF17D
                                                                ------------------------------------------
     · F  L  L   A  G  V  L   V  L  A   M  T  L   G  V  G  A   D  Q  G   C  A  I   N  F  G  K   R  E  L
2401 ATTCCTCTTG GCCGGCGTGC TTGTACTGGC CATGACACTG GGCGTTGGCG CCGATCAAGG ATGCGCCATC AACTTTGGCA AGAGAGAGCT C
     TAAGGAGAAC CGGCCGCACG AACATGACCG GTACTGTGAC CCGCAACCGC GGCTAGTTCC TACGCGGTAG TTGAAACCGT TCTCTCTCGA G
```

Sequence Appendix 3

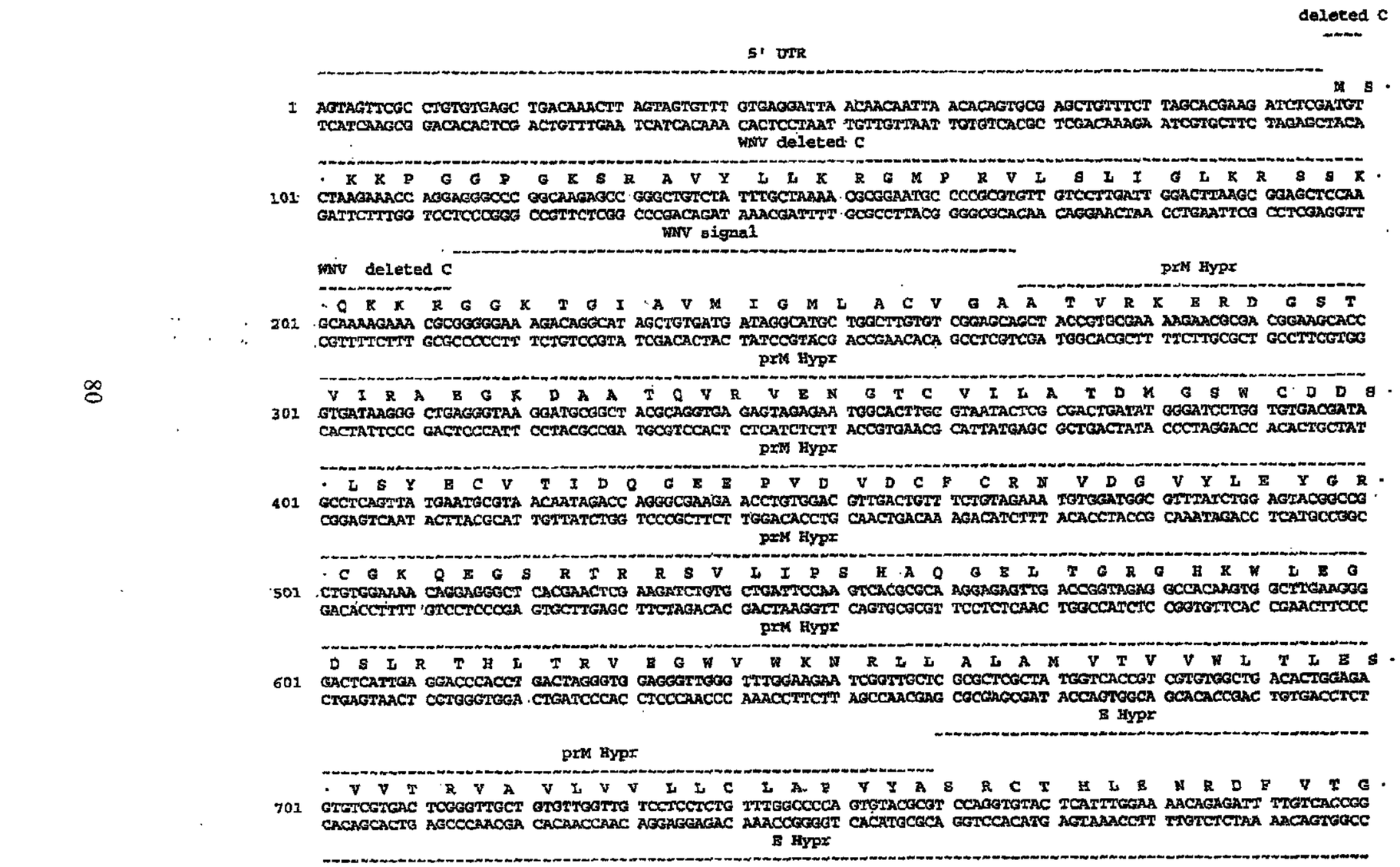
PIV-WN/TBEV Hypr with TBEV signal (p39)

```
                                                                                                    deleted C
                                                                                                    ----
                                                  5' UTR
     ------------------------------------------------------------------------------------------------------
                                                                                                    M  S ·
  1  AGTAGTTCGC CTGTGTGAGC TGACAAACTT AGTAGTGTTT GTGAGGATTA ACAACAATTA ACACAGTGCG AGCTGTTTCT TAGCACGAAG ATCTCGATGT
     TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCTAAT TGTTGTTAAT TGTGTCACGC TCGACAAAGA ATCGTGCTTC TAGAGCTACA
                                                 WNV deleted C protein
     ------------------------------------------------------------------------------------------------------
     · K  K  P   G  G  P   G  K  S  R   A  V  Y   L  L  K   R  G  M  P   R  V  L   S  L  I   G  L  K  R   S  S  K ·
101  CTAAGAAACC AGGAGGGCCC GGCAAGAGCC GGGCTGTCTA TTTGCTAAAA CGCGGAATGC CCCGCGTGTT GTCCTTGATT GGACTTAAGC GGAGCTCCAA
     GATTCTTTGG TCCTCCCGGG CCGTTCTCGG CCCGACAGAT AAACGATTTT GCGCCTTACG GGGCGCACAA CAGGAACTAA CCTGAATTCG CCTCGAGGTT
                                          TBEV signal
                                                                                                 ------------
      deleted C                                                                                  prM Hypr
     -------------                                                                              ------------------
     · Q  K  K   R  G  G  T   D  W  M   S  W  L   L  V  I  G   M  L  G   M  T  I   A  A  T  V   R  K  E   R  D  G
201  ACAAAAGAAA CGGGGGGGAA CAGACTGGAT GAGCTGGCTG CTCGTAATCG GCATGCTGGG CATGACAATC GCAGCTACGG TTCGCAAGGA AAGAGACGGC
     TGTTTTCTTT GCCCCCCCTT GTCTGACCTA CTCGACCGAC GAGCATTAGC CGTACGACCC GTACTGTTAG CGTCGATGCC AAGCGTTCCT TTCTCTGCCG
                                                 prM Hypr
     ------------------------------------------------------------------------------------------------------
      S  T  V  I   R  A  E   G  K  D   A  A  T  Q   V  R  V   E  N  G   T  C  V  I   L  A  T   D  M  G   S  W  C  D ·
301  AGTACGGTCA TACGCGCGGA AGGTAAGGAT GCCGCTACCC AAGTGAGAGT GGAAAATGGT ACCTGCGTCA TTCTGGCCAC CGACATGGGC TCTTGGTGTG
```

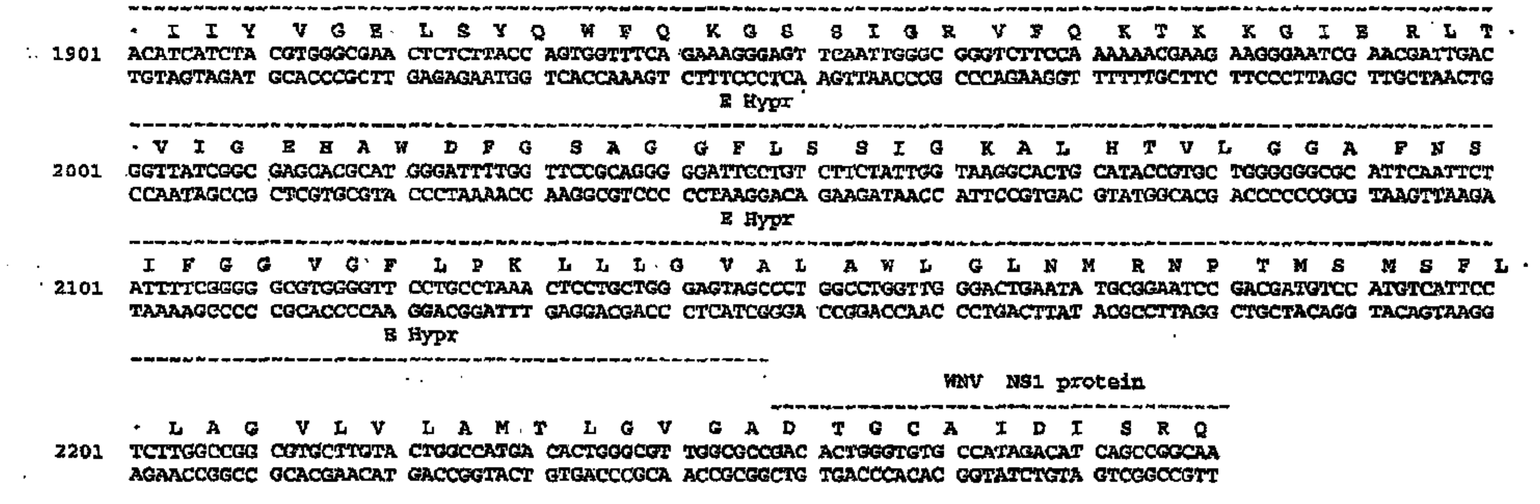

TCATGCCAGT ATGCGCGCCT TCCATTCCTA CGGCGATGGG TTCACTCTCA CCTTTTACCA TGGACGCAGT AAGACCGGTG GCTGTACCCG AGAACCACAC
prM Hypr
· D S L S Y E C V T I D Q G E E P V D V D C F C R N V D G V Y L E Y ·
401 ATGATAGCCT TTCTTATGAG TGCGTAACCA TAGATCAAGG TGAGGAACCT GTTGACGTTG ATTGCTTCTG CCGAAACGTG GATGGGGTGT ATCTCGAATA
TACTATCGGA AAGAATACTC ACGCATTGGT ATCTAGTTCC ACTCCTTGGA CAACTGCAAC TAACGAAGAC GGCTTTGCAC CTACCCCACA TAGAGCTTAT
prM Hypr
· G R C G K Q E G S R T R R S V L I P S H A Q G E L T G R G H K W L
501 TGGACGGTGT GGTAAACAAG AAGGAAGCAG AACCAGACGC TCAGTGCTTA TACCCTCCCA CGCTCAAGGA GAGCTGACCG GACGGGGACA TAAATGGTTG
ACCTGCCACA CCATTTGTTC TTCCTTCGTC TTGGTCTGCG AGTCACGAAT ATGGGAGGGT GCGAGTTCCT CTCGACTGGC CTGCCCCTGT ATTTACCAAC
prM Hypr
E G D S L R T H L T R V E G W V W K N R L L A L A M V T V V W L T L ·
601 GAGGGCGACT CACTCCGAAC ACATTTGACC CGCGTCGAGG GCTGGGTCTG GAAAAATCGG CTGTTGGCCC TCGCTATGGT GACAGTCGTT TGGCTCACGC
CTCCCGCTGA GTGAGGCTTG TGTAAACTGG GCGCAGCTCC CGACCCAGAC CTTTTTAGCC GACAACCGGG AGCGATACCA CTGTCAGCAA ACCGAGTGCG
E Hypr
prM Hypr
· E S V V T R V A V L V V L L C L A P V Y A S R C T H L E N R D F V ·
701 TGGAGTCTGT GGTTACTCGC GTGGCAGTGC TGGTGGTGCT CCTCTGTCTT GCCCCTGTCT ACGCGTCCAG GTGTACTCAT TTGGAAAACA GAGATTTTGT
ACCTCAGACA CCAATGAGCG CACCGTCACG ACCACCACGA GGAGACAGAA CGGGGACAGA TGCGCAGGTC CACATGAGTA AACCTTTTGT CTCTAAAACA
E Hypr
· T G T Q G T T R V T L V L E L G G C V T I T A E G K P S M D V W L
801 CACCGGCACC CAGGGGACGA CTCGGGTAAC CCTGGTGCTT GAACTGGGTG GTTGCGTTAC TATTACCGCT GAGGGCAAAC CCTCTATGGA TGTGTGGCTG
GTGGCCGTGG GTCCCCTGCT GAGCCCATTG GGACCACGAA CTTGACCCAC CAACGCAATG ATAATGGCGA CTCCCGTTTG GGAGATACCT ACACACCGAC
E Hypr
· D A I Y Q E N P A Q T R H Y C L H A K L S D T K V A A R C P T M G P ·
901 GATGCAATCT ATCAGGAGAA TCCCGCACAA ACCAGGGAAT ATTGCCTTCA CGCAAAGCTG TCCGATACAA AGGTCGCGGC TAGGTGCCCA ACAATGGGAC
CTACGTTAGA TAGTCCTCTT AGGGCGTGTT TGGTCCCTTA TAACGGAAGT GCGTTTCGAC AGGCTATGTT TCCAGCGCCG ATCCACGGGT TGTTACCCTG
E Hypr
· A T L A E E H Q G G T V C K R D Q S D R G W G N H C G L F G K G S ·
1001 CGGCCACCCT GGCGGAGGAA CATCAGGGAG GTACAGTGTG CAAACGGGAC CAGAGTGATA GAGGCTGGGG TAATCACTGC GGCCTGTTCG GCAAAGGAAG
GCCGGTGGGA CCGCCTCCTT GTAGTCCCTC CATGTCACAC GTTTGCCCTG GTCTCACTAT CTCCGACCCC ATTAGTGACG CCGGACAAGC CGTTTCCTTC
E Hypr
· I V A C V K A A C E A K K K A T G H V Y D A N K I V Y T V K V E P
1101 TATTGTCGCT TGCGTCAAGG CAGCCTGTGA GGCCAAAAAG AAGGCTACTG GGCACGTCTA TGACGCCAAC AAGATCGTTT ATACAGTGAA AGTGGAACCA
ATAACAGCGA ACGCAGTTCC GTCGGACACT CCGGTTTTTC TTCCGATGAC CCGTGCAGAT ACTGCGGTTG TTCTAGCAAA TATGTCACTT TCACCTTGGT
E Hypr
· H T G D Y V A A N E T H S G R K T A S F T V S S E K T I L T M G E Y ·
1201 CACACAGGGG ATTACGTGGC GGCCAACGAG ACTCATTCCG GTCGCAAAAC GGCCAGCTTC ACCGTGTCAT CCGAAAAGAC CATCCTCACT ATGGGGGAGT
GTGTGTCCCC TAATGCACCG CCGGTTGCTC TGAGTAAGGC CAGCGTTTTG CCGGTCGAAG TGGCACAGTA GGCTTTTCTG GTAGGAGTGA TACCCCCTCA
E Hypr
· G D V S L L C R V A S G V D L A Q T V I L E L D K T V E H L P T A ·
1301 ATGGCGACGT TTCTCTGCTC TGCCGGGTGG CTAGCGGAGT CGACCTGGCC CAGACAGTCA TCCTGGAACT GGATAAAACA GTTGAGCATC TGCCTACCGC
TACCGCTGCA AAGAGACGAG ACGGCCCACC GATCGCCTCA GCTGGACCGG GTCTGTCAGT AGGACCTTGA CCTATTTTGT CAACTCGTAG ACGGATGGCG
E Hypr


E Hypr
WNV NS1 protein

PIV-WN/TBEV Hypr with WNV signal (p40)

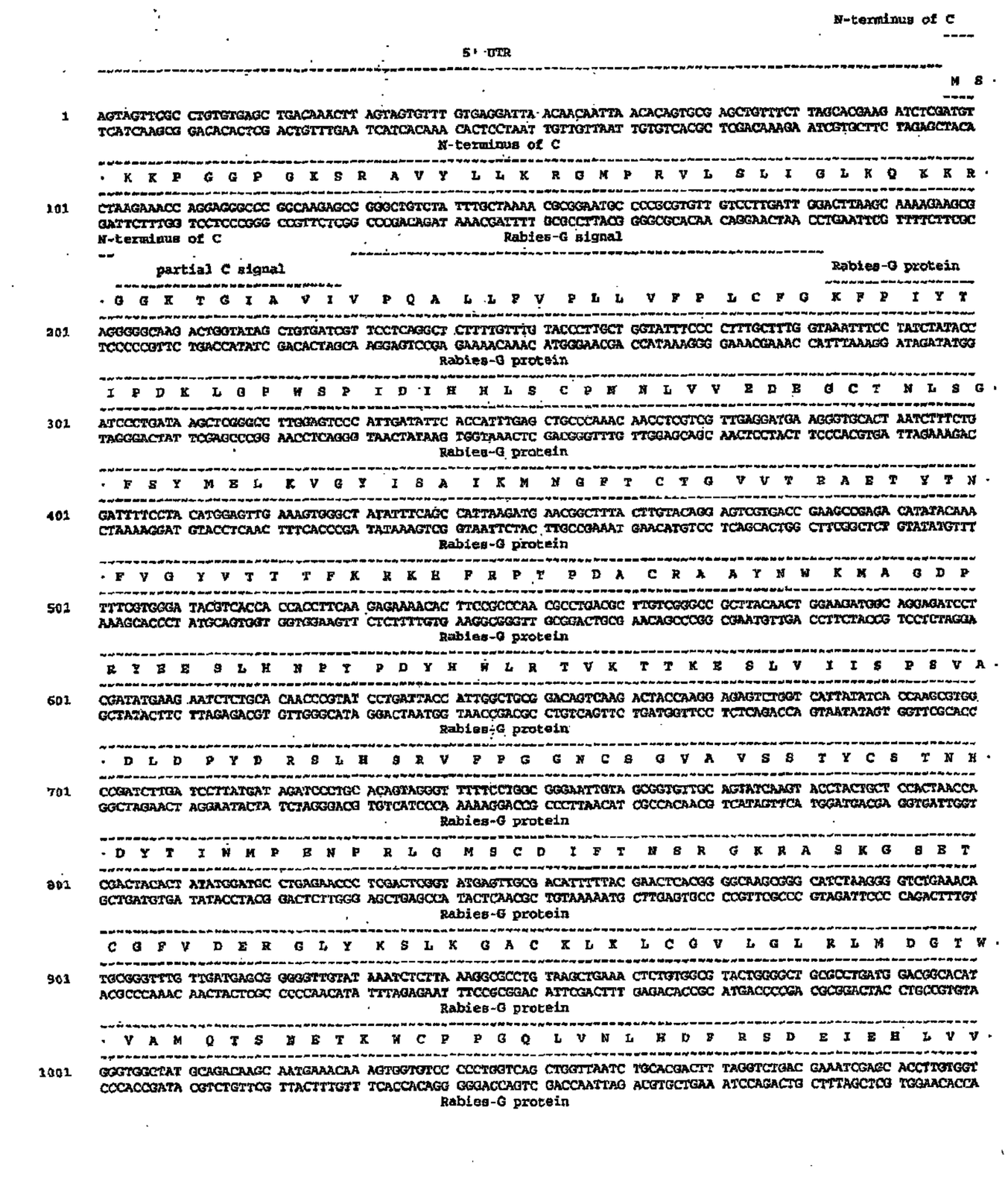


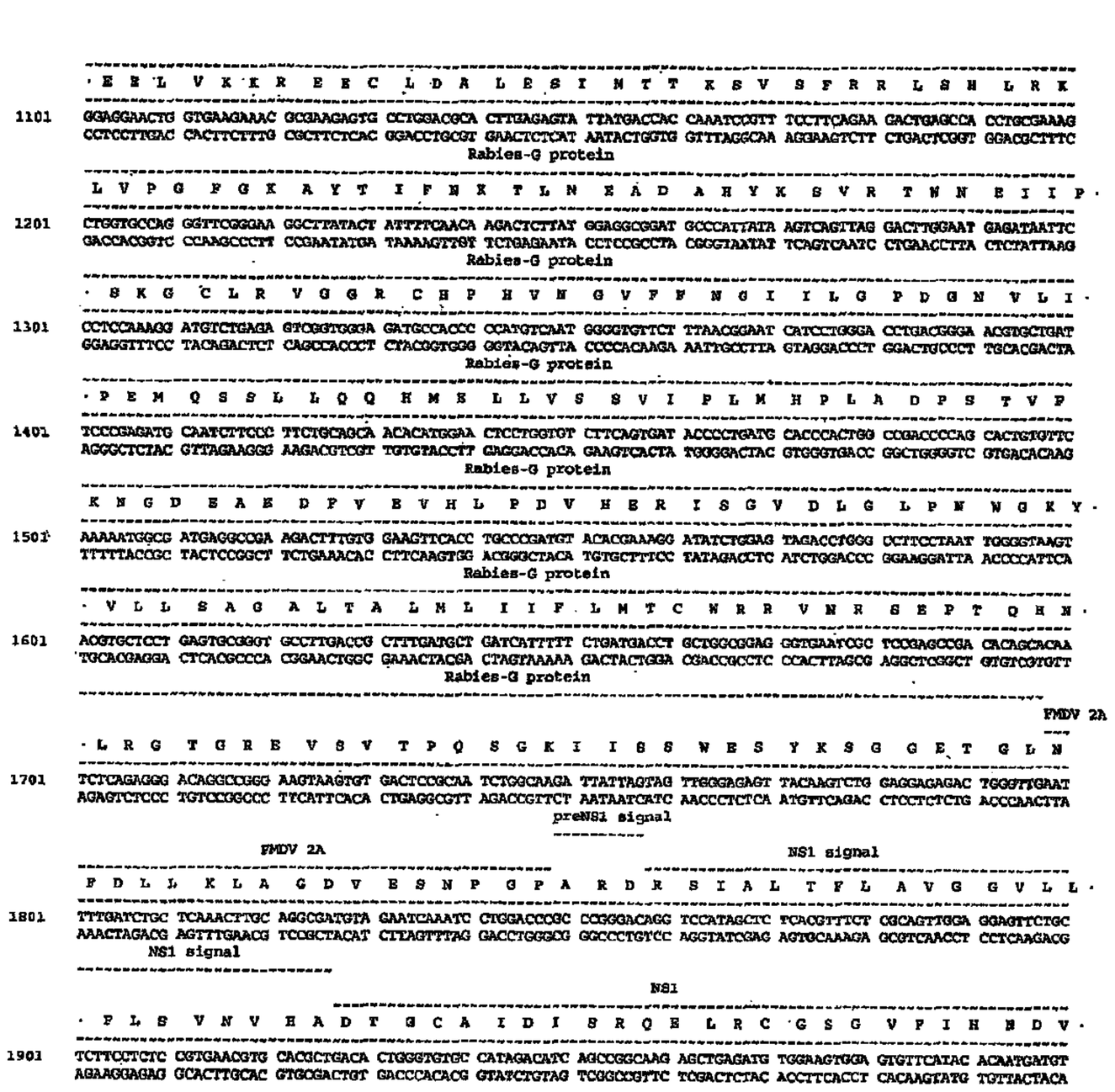
```
     T  Q  G  T  T  R  V  T  L  V  L  E  L  G  G  C  V  T  I  T  A  E  G  K  P  S  M  D  V  W  L  D  A
801  CACCCAGGGG ACGACTCGGG TAACCCTGGT GCTTGAACTG GGTGGTTGCG TTACTATTAC CGCTGAGGGC AAACCCTCTA TGGATGTGTG GCTGGATGCA
     GTGGGTCCCC TGCTGAGCCC ATTGGGACCA CGAACTTGAC CCACCAACGC AATGATAATG GCGACTCCCG TTTGGGAGAT ACCTACACAC CGACCTACGT
                                          E Hypr
     ---------------------------------------------------------------------------------------------------------
     I  Y  Q  E  N  F  A  Q  T  R  E  Y  C  L  H  A  K  L  S  D  T  K  V  A  A  R  C  P  T  M  G  P  A  T
901  ATCTATCAGG AGAATCCCGC ACAAACCAGG GAATATTGCC TTCACGCAAA GCTGTCCGAT ACAAAGGTCG CGGCTAGGTG CCCAACAATG GGACCGGCCA
     TAGATAGTCC TCTTAGGGCG TGTTTGGTCC CTTATAACGG AAGTGCGTTT CGACAGGCTA TGTTTCCAGC GCCGATCCAC GGGTTGTTAC CCTGGCCGGT
                                          E Hypr
     ---------------------------------------------------------------------------------------------------------
     L  A  E  E  H  Q  G  G  T  V  C  K  R  D  Q  S  D  R  G  W  G  N  H  C  G  L  F  G  K  G  S  I  V
1001 CCCTGGCGGA GGAACATCAG GGAGGTACAG TGTGCAAACG GGACCAGAGT GATAGAGGCT GGGGTAATCA CTGCGGCCTG TTCGGCAAAG GAAGTATTGT
     GGGACCGCCT CCTTGTAGTC CCTCCATGTC ACACGTTTGC CCTGGTCTCA CTATCTCCGA CCCCATTAGT GACGCCGGAC AAGCCGTTTC CTTCATAACA
                                          E Hypr
     ---------------------------------------------------------------------------------------------------------
     A  C  V  K  A  A  C  E  A  K  K  K  A  T  G  H  V  Y  D  A  N  K  I  V  Y  T  V  K  V  E  P  H  T
1101 CGCTTGCGTC AAGGCAGCCT GTGAGGCCAA AAAGAAGGCT ACTGGGCACG TCTATGACGC CAACAAGATC GTTTATACAG TGAAAGTGGA ACCACACACA
     GCGAACGCAG TTCCGTCGGA CACTCCGGTT TTTCTTCCGA TGACCCGTGC AGATACTGCG GTTGTTCTAG CAAATATGTC ACTTTCACCT TGGTGTGTGT
                                          E Hypr
     ---------------------------------------------------------------------------------------------------------
     G  D  Y  V  A  A  N  E  T  K  S  G  R  K  T  A  S  F  T  V  S  S  E  K  T  I  L  T  M  G  E  Y  G  D
1201 GGGGATTACG TGGCGGCCAA CGAGACTCAT TCCGGTCGCA AAACGGCCAG CTTCACCGTG TCATCCGAAA AGACCATCCT CACTATGGGG GAGTATGGCG
     CCCCTAATGC ACCGCCGGTT GCTCTGAGTA AGGCCAGCGT TTTGCCGGTC GAAGTGGCAC AGTAGGCTTT TCTGGTAGGA GTGATACCCC CTCATACCGC
                                          E Hypr
     ---------------------------------------------------------------------------------------------------------
     V  S  L  L  C  R  V  A  S  G  V  D  L  A  Q  T  V  I  L  E  L  D  K  T  V  E  H  L  P  T  A  W  Q
1301 ACGTTTCTCT GCTCTGCCGG GTGGCTAGCG GAGTCGACCT GGCCCAGACA GTCATCCTGG AACTGGATAA AACAGTTGAG CATCTGCCTA CCGCTTGGCA
     TGCAAAGAGA CGAGACGGCC CACCGATCGC CTCAGCTGGA CCGGGTCTGT CAGTAGGACC TTGACCTATT TTGTCAACTC GTAGACGGAT GGCGAACCGT
                                          E Hypr
     ---------------------------------------------------------------------------------------------------------
     V  H  R  D  W  F  N  D  L  A  L  P  W  K  H  E  G  A  R  N  W  N  N  A  E  R  L  V  E  F  G  A  P
1401 GGTGCACAGG GATTGGTTTA ACGACCTTGC CCTGCCATGG AAACATGAAG GAGCGAGAAA CTGGAATAAT GCAGAGCGAC TCGTAGAATT CGGTGCCCCT
     CCACGTGTCC CTAACCAAAT TGCTGGAACG GGACGGTACC TTTGTACTTC CTCGCTCTTT GACCTTATTA CGTCTCGCTG AGCATCTTAA GCCACGGGGA
                                          E Hypr
     ---------------------------------------------------------------------------------------------------------
     H  A  V  K  M  D  V  Y  N  L  G  D  Q  T  G  V  L  L  K  A  L  A  G  V  P  V  A  H  I  E  G  T  K  Y
1501 CATGCCGTGA AGATGGACGT CTACAATCTG GGTGATCAGA CCGGCGTTCT CCTTAAAGCT CTCGCTGGCG TACCAGTTGC CCACATCGAA GGAACGAAGT
     GTACGGCACT TCTACCTGCA GATGTTAGAC CCACTAGTCT GGCCGCAAGA GGAATTTCGA GAGCGACCGC ATGGTCAACG GGTGTAGCTT CCTTGCTTCA
                                          E Hypr
     ---------------------------------------------------------------------------------------------------------
     H  L  K  S  G  H  V  T  C  E  V  G  L  E  K  L  K  M  K  G  L  T  Y  T  M  C  D  K  T  K  F  T  W
1601 ACCACCTGAA GTCAGGCCAT GTAACTTGCG AGGTGGGCCT GGAGAAGTTG AAAATGAAAG GTCTTACGTA CACAATGTGT GACAAGACCA AGTTCACATG
     TGGTGGACTT CAGTCCGGTA CATTGAACGC TCCACCCGGA CCTCTTCAAC TTTTACTTTC CAGAATGCAT GTGTTACACA CTGTTCTGGT TCAAGTGTAC
                                          E Hypr
     ---------------------------------------------------------------------------------------------------------
     K  R  A  P  T  D  S  G  H  D  T  V  V  M  E  V  T  F  S  G  T  K  P  C  R  I  P  V  R  A  V  A  H
1701 GAAGAGGGCC CCCACAGATA GCGGCCACGA TACTGTGGTG ATGGAGGTGA CCTTTTCTGG AACAAAACCC TGCAGAATAC CCGTGCGGGC TGTAGCTCAC
     CTTCTCCCGG GGGTGTCTAT CGCCGGTGCT ATGACACCAC TACCTCCACT GGAAAAGACC TTGTTTTGGG ACGTCTTATG GGCACGCCCG ACATCGAGTG
                                          E Hypr
     ---------------------------------------------------------------------------------------------------------
     G  S  P  D  V  N  V  A  M  L  I  T  P  N  P  T  I  E  N  N  G  G  G  F  I  E  M  Q  L  P  P  G  D  N
1801 GGATCTCCCG ATGTCAATGT TGCTATGCTG ATTACACCTA ACCCTACCAT CGAGAATAAC GGTGGTGGTT TTATTGAGAT GCAGCTTCCG CCAGGCGATA
     CCTAGAGGGC TACAGTTACA ACGATACGAC TAATGTGGAT TGGGATGGTA GCTCTTATTG CCACCACCAA AATAACTCTA CGTCGAAGGC GGTCCGCTAT
                                          E Hypr
```


E Hypr
E Hypr
E Hypr
WNV NS1 protein

Sequence Appendix 4. WN PIV constructs expressing rabies virus G protein.

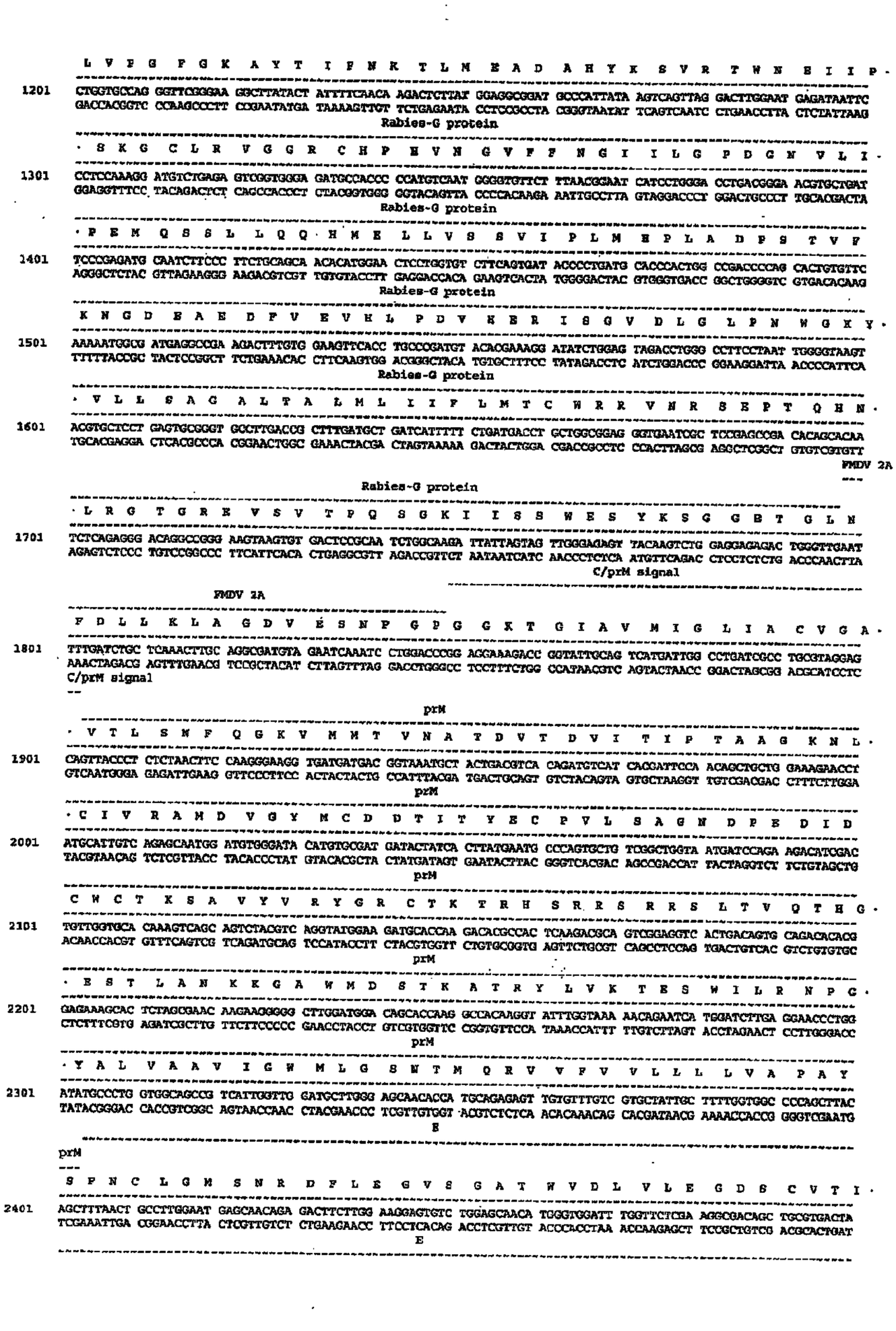
WN (ΔCprME)-Rabies PIV sequence (partial)

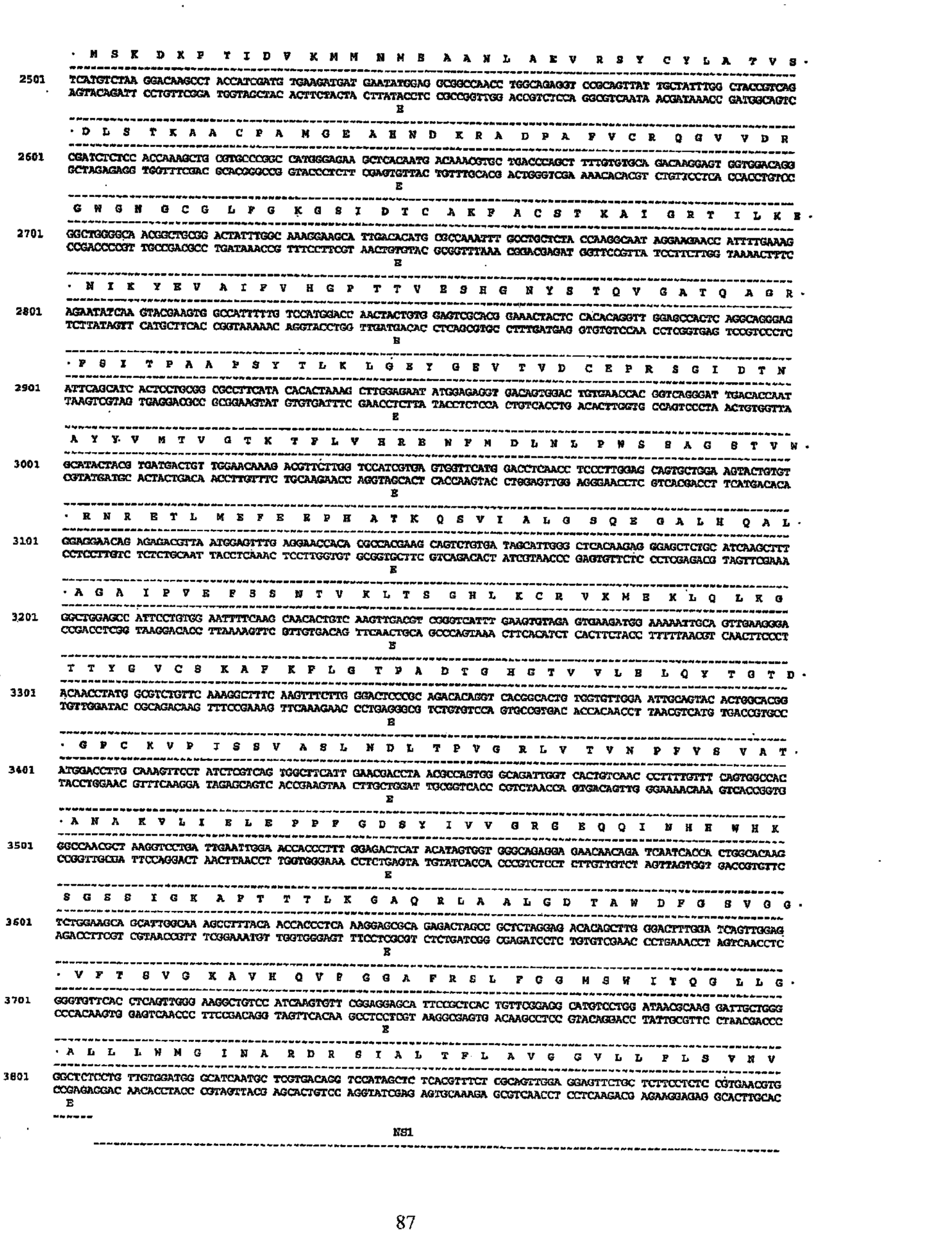



E E L V K K R E E C L D A L E S I M T T K S V S F R R L S H L R K
1101 GGAGGAACTG GTGAAGAAAC GCGAAGAGTG CCTGGACGCA CTTGAGAGTA TTATGACCAC CAAATCGGTT TCCTTCAGAA GACTGAGCCA CCTGCGAAAG
CCTCCTTGAC CACTTCTTTG CGCTTCTCAC GGACCTGCGT GAACTCTCAT AATACTGGTG GTTTAGGCAA AGGAAGTCTT CTGACTCGGT GGACGCTTTC
Rabies-G protein
L V P G F G K A Y T I F N K T L N E A D A H Y K S V R T W N E I I P
1201 CTGGTGCCAG GGTTCGGGAA GGCTTATACT ATTTTCAACA AGACTCTTAT GGAGGCGGAT GCCCATTATA AGTCAGTTAG GACTTGGAAT GAGATAATTC
GACCACCGGTC CCAAGCCCTT CCGAATATGA TAAAAGTTGT TCTGAGAATA CCTCCGCCTA CGGGTAATAT TCAGTCAATC CTGAACCTTA CTCTATTAAG
Rabies-G protein
S K G C L R V G G R C H P H V N G V F F N G I I L G P D G N V L I
1301 CCTCCAAAGG ATGTCTGAGA GTCGGTGGGA GATGCCACCC CCATGTCAAT GGGGTGTTCT TTAACGGAAT CATCCTGGGA CCTGACGGGA ACGTGCTGAT
GGAGGTTTCC TACAGACTCT CAGCCACCCT CTACGGTGGG GGTACAGTTA CCCCACAAGA AATTGCCTTA GTAGGACCCT GGACTGCCCT TGCACGACTA
Rabies-G protein
P E M Q S S L L Q Q H M E L L V S S V I P L M H P L A D P S T V F
1401 TCCCGAGATG CAATGTTCCC TTCTGCAGCA ACACATGGAA CTCCTGGTGT CTTCAGTGAT ACCCCTGATG CACCCACTGG CCGACCCCAG CACTGTGTTC
AGGGCTCTAC GTTAGAAGGG AAGACGTCGT TGTGTACCTT GAGGACCACA GAAGTCACTA TGGGGACTAC GTGGGTGACC GGCTGGGGTC GTGACACAAG
Rabies-G protein
K N G D E A E D P V E V H L P D V H E R I S G V D L G L P N W G K Y
1501 AAAAATGGCG ATGAGGCCGA AGACTTTGTG GAAGTTCACC TGCCCGATGT ACACGAAAGG ATATCTGGAG TAGACCTGGG CCTTCCTAAT TGGGGTAAGT
TTTTTACCGC TACTCCGGCT TCTGAAACAC CTTCAAGTGG ACGGGCTACA TGTGCTTTCC TATAGACCTC ATCTGGACCC GGAAGGATTA ACCCCATTCA
Rabies-G protein
V L L S A G A L T A L M L I I F L M T C W R R V N R S E P T Q H N
1601 ACGTGCTCCT GAGTGCGGGT GCCTTGACCG CTTTGATGCT GATCATTTTT CTGATGACCT GCTGGCGGAG GGTGAATCGC TCCGAGCCGA CACAGCACAA
TGCACGAGGA CTCACGCCCA CGGAACTGGC GAAACTACGA CTAGTAAAAA GACTACTGGA CGACCGCCTC CCACTTAGCG AGGCTCGGCT GTGTCGTGTT
Rabies-G protein
FMDV 2A
L R G T G R E V S V T P Q S G K I I S S W E S Y K S G G E T G L N
1701 TCTCAGAGGG ACAGGCCGGG AAGTAAGTGT GACTCCGCAA TCTGGCAAGA TTATTAGTAG TTGGGAGAGT TACAAGTCTG GAGGAGAGAC TGGGTTGAAT
AGAGTCTCCC TGTCCGGCCC TTCATTCACA CTGAGGCGTT AGACCGTTCT AATAATCATC AACCCTCTCA ATGTTCAGAC CTCCTCTCTG ACCCAACTTA
preNS1 signal
FMDV 2A
NS1 signal
F D L L K L A G D V E S N P G P A R D R S I A L T F L A V G G V L L
1801 TTTGATCTGC TCAAACTTGC AGGCGATGTA GAATCAAATC CTGGACCCGC CCGGGACAGG TCCATAGCTC TCACGTTTCT CGCAGTTGGA GGAGTTCTGC
AAACTAGACG AGTTTGAACG TCCGCTACAT CTTAGTTTAG GACCTGGGCG GGCCCTGTCC AGGTATCGAG AGTGCAAAGA GCGTCAACCT CCTCAAGACG
NS1 signal
NS1
F L S V N V H A D T G C A I D I S R Q E L R C G S G V F I H N D V
1901 TCTTCCTCTC CGTGAACGTG CACGCTGACA CTGGGTGTGC CATAGACATC AGCCGGCAAG AGCTGAGATG TGGAAGTGGA GTGTTCATAC ACAATGATGT
AGAAGGAGAG GCACTTGCAC GTGCGACTGT GACCCACACG GTATCTGTAG TCGGCCGTTC TCGACTCTAC ACCTTCACCT CACAAGTATG TGTTACTACA

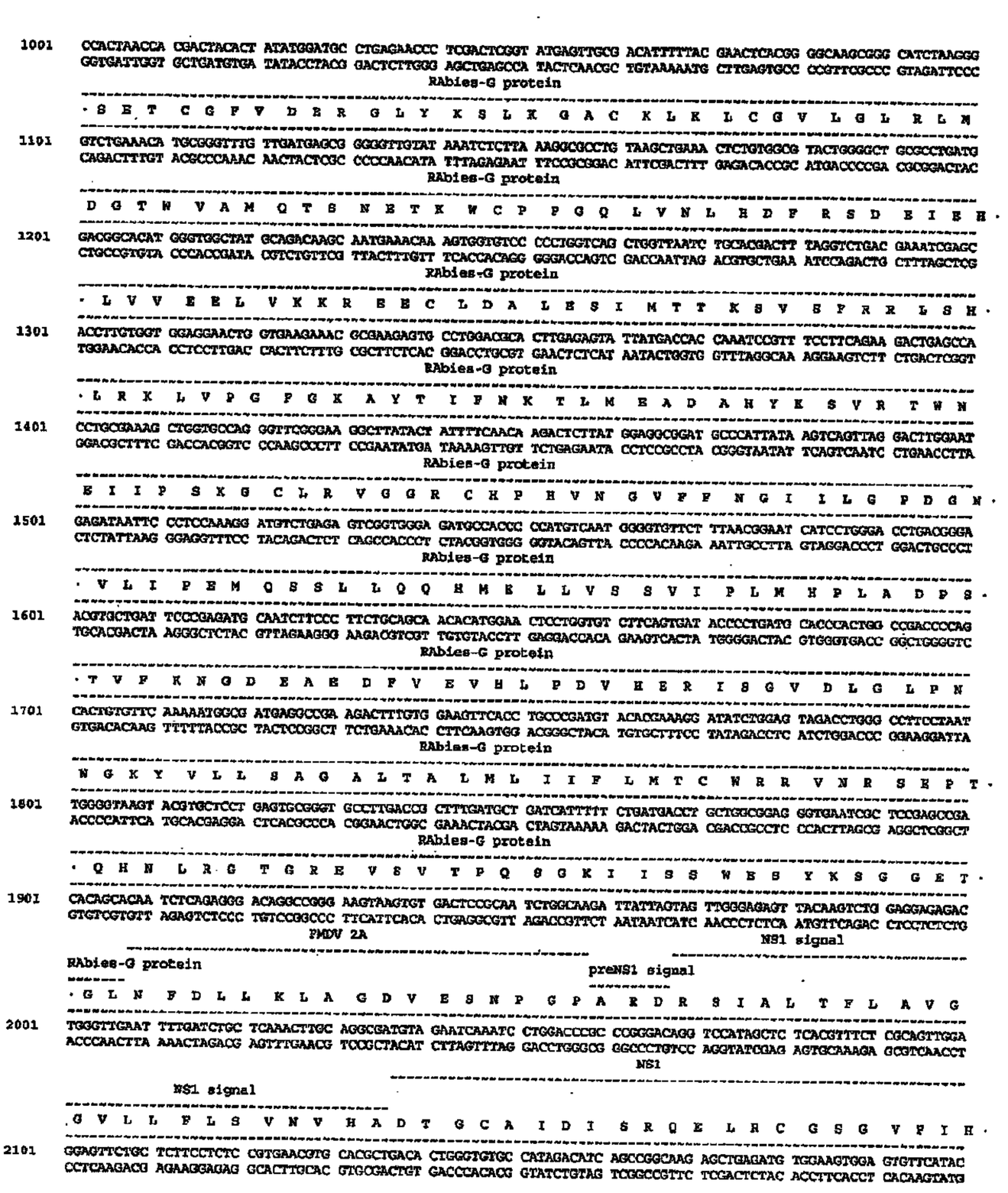
WN (ΔC)-Rabies G PIV sequence (partial).

```
                                                                                                                5'UTR
                                                                                                                ----------
terminus of C                                                                                                                
```


L V F G F G K A Y T I F M K T L M E A D A H Y K S V R T W N E I I P
1201 CTGGTGCCAG GGTTCGGGAA GGCTTATACT ATTTTCAACA AGACTCTTAT GGAGGCGGAT GCCCATTATA AGTCAGTTAG GACTTGGAAT GAGATAATTC
GACCACGGTC CCAAGCCCTT CCGAATATGA TAAAAGTTGT TCTGAGAATA CCTCCGCCTA CGGGTAATAT TCAGTCAATC CTGAACCTTA CTCTATTAAG
Rabies-G protein
S K G C L R V G G R C H P M V N G V F F N G I I L G P D G N V L I
1301 CCTCCAAAGG ATGTCTGAGA GTCGGTGGGA GATGCCACCC CCATGTCAAT GGGGTGTTCT TTAACGGAAT CATCCTGGGA CCTGACGGGA ACGTGCTGAT
GGAGGTTTCC TACAGACTCT CAGCCACCCT CTACGGTGGG GGTACAGTTA CCCCACAAGA AATTGCCTTA GTAGGACCCT GGACTGCCCT TGCACGACTA
Rabies-G protein
P E M Q S S L L Q Q H M E L L V S S V I P L M H P L A D P S T V F
1401 TCCCGAGATG CAATCTTCCC TTCTGCAGCA ACACATGGAA CTCCTGGTGT CTTCAGTGAT ACCCCTGATG CACCCACTGG CCGACCCCAG CACTGTGTTC
AGGGCTCTAC GTTAGAAGGG AAGACGTCGT TGTGTACCTT GAGGACCACA GAAGTCACTA TGGGGACTAC GTGGGTGACC GGCTGGGGTC GTGACACAAG
Rabies-G protein
K N G D E A E D F V E V H L P D V H E R I S G V D L G L P N W G K Y
1501 AAAAATGGCG ATGAGGCCGA AGACTTTGTG GAAGTTCACC TGCCCGATGT ACACGAAAGG ATATCTGGAG TAGACCTGGG CCTTCCTAAT TGGGGTAAGT
TTTTTACCGC TACTCCGGCT TCTGAAACAC CTTCAAGTGG ACGGGCTACA TGTGCTTTCC TATAGACCTC ATCTGGACCC GGAAGGATTA ACCCCATTCA
Rabies-G protein
V L L S A G A L T A L M L I I F L M T C W R R V N R S E P T Q H N
1601 ACGTGCTCCT GAGTGCGGGT GCCTTGACCG CTTTGATGCT GATCATTTTT CTGATGACCT GCTGGCGGAG GGTGAATCGC TCCGAGCCGA CACAGCACAA
TGCACGAGGA CTCACGCCCA CGGAACTGGC GAAACTACGA CTAGTAAAAA GACTACTGGA CGACCGCCTC CCACTTAGCG AGGCTCGGCT GTGTCGTGTT
FMDV 2A
Rabies-G protein
L R G T G R E V S V T P Q S G K I I S S W E S Y K S G G E T G L N
1701 TCTCAGAGGG ACAGGCCGGG AAGTAAGTGT GACTCCGCAA TCTGGCAAGA TTATTAGTAG TTGGGAGAGT TACAAGTCTG GAGGAGAGAC TGGGTTGAAT
AGAGTCTCCC TGTCCGGCCC TTCATTCACA CTGAGGCGTT AGACCGTTCT AATAATCATC AACCCTCTCA ATGTTCAGAC CTCCTCTCTG ACCCAACTTA
C/prM signal
FMDV 2A
F D L L K L A G D V E S N P G P G G K T G I A V M I G L I A C V G A
1801 TTTGATCTGC TCAAACTTGC AGGCGATGTA GAATCAAATC CTGGACCCGG AGGAAAGACC GGTATTGCAG TCATGATTGG CCTGATCGCC TGCGTAGGAG
AAACTAGACG AGTTTGAACG TCCGCTACAT CTTAGTTTAG GACCTGGGCC TCCTTTCTGG CCATAACGTC AGTACTAACC GGACTAGCGG ACGCATCCTC
C/prM signal
prM
V T L S M F Q G K V M M T V N A T D V T D V I T I P T A A G K N L
1901 CAGTTACCCT CTCTAACTTC CAAGGGAAGG TGATGATGAC GGTAAATGCT ACTGACGTCA CAGATGTCAT CACGATTCCA ACAGCTGCTG GAAAGAACCT
GTCAATGGGA GAGATTGAAG GTTCCCTTCC ACTACTACTG CCATTTACGA TGACTGCAGT GTCTACAGTA GTGCTAAGGT TGTCGACGAC CTTTCTTGGA
prM
C I V R A M D V G Y M C D D T I T Y E C P V L S A G N D P E D I D
2001 ATGCATTGTC AGAGCAATGG ATGTGGGATA CATGTGCGAT GATACTATCA CTTATGAATG CCCAGTGCTG TCGGCTGGTA ATGATCCAGA AGACATCGAC
TACGTAACAG TCTCGTTACC TACACCCTAT GTACACGCTA CTATGATAGT GAATACTTAC GGGTCACGAC AGCCGACCAT TACTAGGTCT TCTGTAGCTG
prM
C W C T K S A V Y V R Y G R C T K T R H S R R S R R S L T V Q T H G
2101 TGTTGGTGCA CAAAGTCAGC AGTCTACGTC AGGTATGGAA GATGCACCAA GACACGCCAC TCAAGACGCA GTCGGAGGTC ACTGACAGTG CAGACACACG
ACAACCACGT GTTTCAGTCG TCAGATGCAG TCCATACCTT CTACGTGGTT CTGTGCGGTG AGTTCTGCGT CAGCCTCCAG TGACTGTCAC GTCTGTGTGC
prM
E S T L A N K K G A W M D S T K A T R Y L V K T E S W I L R N P G
2201 GAGAAAGCAC TCTAGCGAAC AAGAAGGGGG CTTGGATGGA CAGCACCAAG GCCACAAGGT ATTTGGTAAA AACAGAATCA TGGATCTTGA GGAACCCTGG
CTCTTTCGTG AGATCGCTTG TTCTTCCCCC GAACCTACCT GTCGTGGTTC CGGTGTTCCA TAAACCATTT TTGTCTTAGT ACCTAGAACT CCTTGGGACC
prM
Y A L V A A V I G W M L G S N T M Q R V V F V V L L L L V A P A Y
2301 ATATGCCCTG GTGGCAGCCG TCATTGGTTG GATGCTTGGG AGCAACACCA TGCAGAGAGT TGTGTTTGTC GTGCTATTGC TTTTGGTGGC CCCAGCTTAC
TATACGGGAC CACCGTCGGC AGTAACCAAC CTACGAACCC TCGTTGTGGT ACGTCTCTCA ACACAAACAG CACGATAACG AAAACCACCG GGGTCGAATG
E
prM
S F N C L G M S N R D F L E G V S G A T W V D L V L E G D S C V T I
2401 AGCTTTAACT GCCTTGGAAT GAGCAACAGA GACTTCTTGG AAGGAGTGTC TGGAGCAACA TGGGTGGATT TGGTTCTCTA AGGCGACAGC TGCGTGACTA
TCGAAATTGA CGGAACCTTA CTCGTTGTCT CTGAAGAACC TTCCTCACAG ACCTCGTTGT ACCCACCTAA ACCAAGAGCT TCCGCTGTCG ACGCACTGAT
E


· M S K D K P T I D V K M M N M S A A N L A E V R S Y C Y L A T V S ·
2501 TCATGTCTAA GGACAAGCCT ACCATCGATG TGAAGATGAT GAATATGGAG GCGGCCAACC TGGCAGAGGT CGGCAGTTAT TGCTATTTGG CTACCGTCAG
AGTACAGATT CCTGTTCGGA TGGTAGCTAC ACTTCTACTA CTTATACCTC CGCCGGTTGG ACCGTCTCCA GGCGTCAATA ACGATAAACC GATGGCAGTC
E
· D L S T K A A C P A M G E A H N D K R A D P A F V C R Q G V V D R
2601 CGATCTCTCC ACCAAAGCTG CGTGCCCGGC CATGGGAGAA GCTCACAATG ACAAACGTGC TGACCCAGCT TTTGTGTGCA GACAAGGAGT GGTGGACAGG
GCTAGAGAGG TGGTTTCGAC GCACGGGCCG GTACCCTCTT CGAGTGTTAC TGTTTGCACG ACTGGGTCGA AAACACACGT CTGTTCCTCA CCACCTGTCC
E
G W G N G C G L F G K G S I D T C A K F A C S T K A I G R T I L K E ·
2701 GGCTGGGGCA ACGGCTGCGG ACTATTTGGC AAAGGAAGCA TTGACACATG CGCCAAATTT GCCTGCTCTA CCAAGGCAAT AGGAAGAACC ATTTTGAAAG
CCGACCCCGT TGCCGACGCC TGATAAACCG TTTCCTTCGT AACTGTGTAC GCGGTTTAAA CGGACGAGAT GGTTCCGTTA TCCTTCTTGG TAAAACTTTC
E
· N I K Y E V A I F V H G P T T V E S H G N Y S T Q V G A T Q A G R ·
2801 AGAATATCAA GTACGAAGTG GCCATTTTTG TCCATGGACC AACTACTGTG GAGTCGCACG GAAACTACTC CACACAGGTT GGAGCCAGTC AGGCAGGGAG
TCTTATAGTT CATGCTTCAC CGGTAAAAAC AGGTACCTGG TTGATGACAC CTCAGCGTGC CTTTGATGAG GTGTGTCCAA CCTCGGTGAG TCCGTCCCTC
E
· F S I T P A A P S Y T L K L G E Y G E V T V D C E P R S G I D T N
2901 ATTCAGCATC ACTCCTGCGG CGCCTTCATA CACACTAAAG CTTGGAGAAT ATGGAGAGGT GACAGTGGAC TGTGAACCAC GGTCAGGGAT TGACACCAAT
TAAGTCGTAG TGAGGACGCC GCGGAAGTAT GTGTGATTTC GAACCTCTTA TACCTCTCCA CTGTCACCTG ACACTTGGTG CCAGTCCCTA ACTGTGGTTA
E
A Y Y V M T V G T K T F L V H R E N F N D L N L P W S S A G S T V N ·
3001 GCATACTACG TGATGACTGT TGGAACAAAG ACGTTCTTGG TCCATCGTGA GTGGTTCATG GACCTCAACC TCCCTTGGAG CAGTGCTGGA AGTACTGTGT
CGTATGATGC ACTACTGACA ACCTTGTTTC TGCAAGAACC AGGTAGCACT CACCAAGTAC CTGGAGTTGG AGGGAACCTC GTCACGACCT TCATGACACA
E
· R N R E T L M E F E E P H A T K Q S V I A L G S Q E G A L H Q A L ·
3101 GGAGGAACAG AGAGACGTTA ATGGAGTTTG AGGAACCACA CGCCACGAAG CAGTCTGTGA TAGCATTGGG CTCACAAGAG GGAGCTCTGC ATCAAGCTTT
CCTCCTTGTC TCTCTGCAAT TACCTCAAAC TCCTTGGTGT GCGGTGCTTC GTCAGACACT ATCGTAACCC GAGTGTTCTC CCTCGAGACG TAGTTCGAAA
E
· A G A I P V E F S S N T V K L T S G H L K C R V K M E K L Q L K G
3201 GGCTGGAGCC ATTCCTGTGG AATTTTCAAG CAACACTGTC AAGTTGACGT CGGGTCATTT GAAGTGTAGA GTGAAGATGG AAAAATTGCA GTTGAAGGGA
CCGACCTCGG TAAGGACACC TTAAAAGTTC GTTGTGACAG TTCAACTGCA GCCCAGTAAA CTTCACATCT CACTTCTACC TTTTTAACGT CAACTTCCCT
E
T T Y G V C S K A F K F P L G T P A D T G H G T V V L E L Q Y T G T D ·
3301 ACAACCTATG GCGTCTGTTC AAAGGCTTTC AAGTTTCTTG GGACTCCCGC AGACACAGGT CACGGCACTG TGGTGTTGGA ATTGCAGTAC ACTGGCACGG
TGTTGGATAC CGCAGACAAG TTTCCGAAAG TTCAAAGAAC CCTGAGGGCG TCTGTGTCCA GTGCCGTGAC ACCACAACCT TAACGTCATG TGACCGTGCC
E
· G P C K V P I S S V A S L N D L T P V G R L V T V N P F V S V A T ·
3401 ATGGACCTTG CAAAGTTCCT ATCTCGTCAG TGGCTTCATT GAACGACCTA ACGCCAGTGG GCAGATTGGT CACTGTCAAC CCTTTTGTTT CAGTGGCCAC
TACCTGGAAC GTTTCAAGGA TAGAGCAGTC ACCGAAGTAA CTTGCTGGAT TGCGGTCACC CGTCTAACCA GTGACAGTTG GGAAAACAAA GTCACCGGTG
E
· A N A K V L I E L E P P F G D S Y I V V G R G E Q Q I N H H W H K
3501 GGCCAACGCT AAGGTCCTGA TTGAATTGGA ACCACCCTTT GGAGACTCAT ACATAGTGGT GGGCAGAGGA GAACAACAGA TCAATCACCA CTGGCACAAG
CCGGTTGCGA TTCCAGGACT AACTTAACCT TGGTGGGAAA CCTCTGAGTA TGTATCACCA CCCGTCTCCT CTTGTTGTCT AGTTAGTGGT GACCGTGTTC
E
S G S S I G K A P T T T L K G A Q R L A A L G D T A W D F G S V G G ·
3601 TCTGGAAGCA GCATTGGCAA AGCGTTTACA ACCACCCTCA AAGGAGCGCA GAGACTAGCC GCTCTAGGAG ACACAGCTTG GGACTTTGGA TCAGTTGGAG
AGACCTTCGT CGTAACCGTT TCGGAAATGT TGGTGGGAGT TTCCTCGCGT CTCTGATCGG CGAGATCCTC TGTGTCGAAC CCTGAAACCT AGTCAACCTC
E
· V F T S V G K A V H Q V F G G A F R S L F G G M S W I T Q G L L G ·
3701 GGGTGTTCAC CTCAGTTGGG AAGGCTGTCC ATCAAGTGTT CGGAGGAGCA TTCCGCTCAC TGTTCGGAGG CATGTCCTGG ATAACGCAAG GATTGCTGGG
CCCACAAGTG GAGTCAACCC TTCCGACAGG TAGTTCACAA GCCTCCTCGT AAGGCGAGTG ACAAGCCTCC GTACAGGACC TATTGCGTTC CTAACGACCC
E
· A L L L W M G I N A R D R S I A L T F L A V G G V L L F L S V N V
3801 GGCTCTCCTG TIGTGGATGG GCATCAATGC TCGTGACAGG TCCATAGCTC TCACGTTTCT CGCAGTTGGA GGAGTTCTGC TCTTCCTCTC CGTGAACGTG
CCGAGACGAC AACACCTACC CGTAGTTACG AGCACTGTCC AGGTATCGAG AGTGCAAAGA GCGTCAACCT CCTCAAGACG AGAAGGAGAG GCACTTGCAC
E
NS1

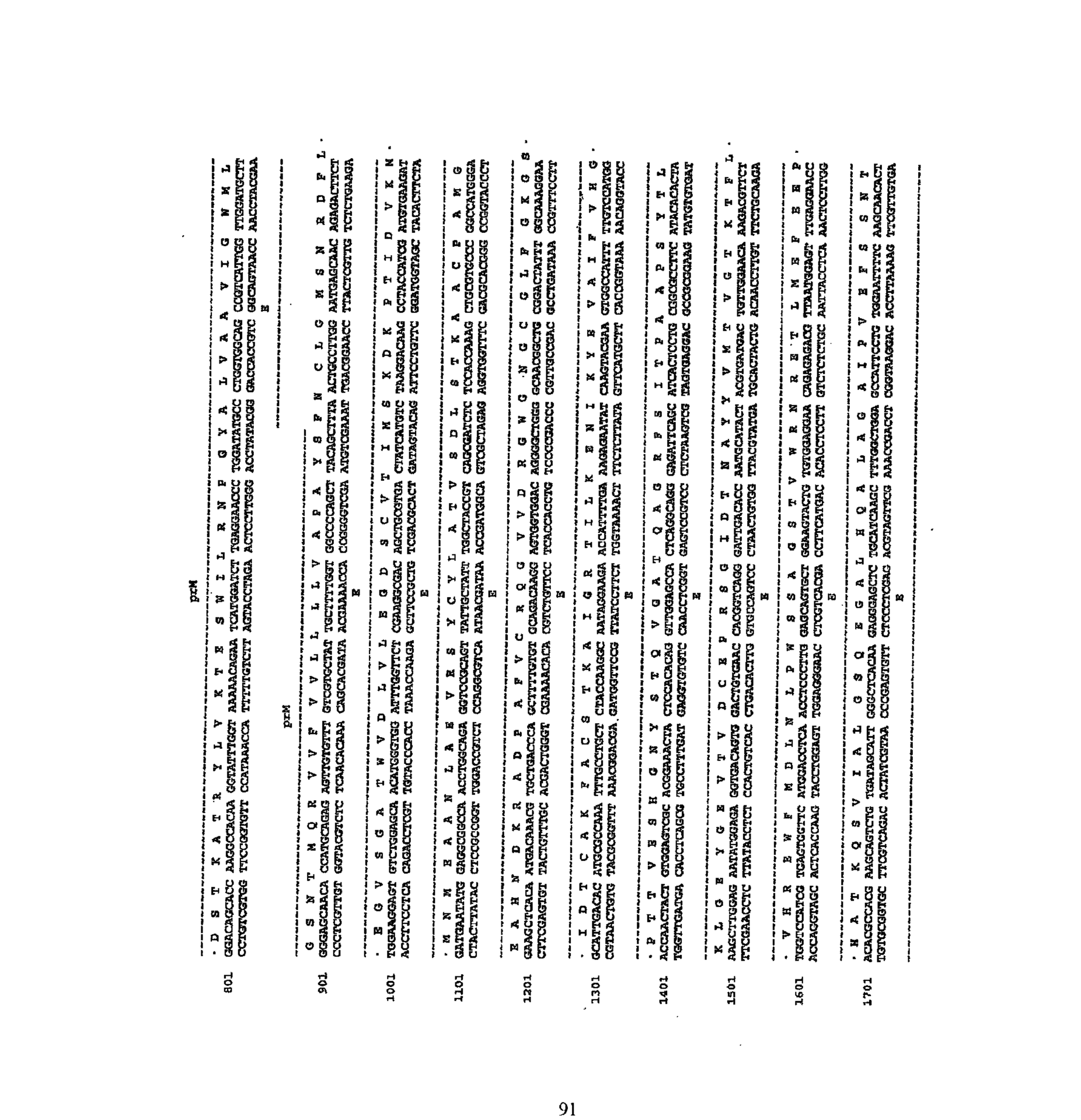
```
                                                
```

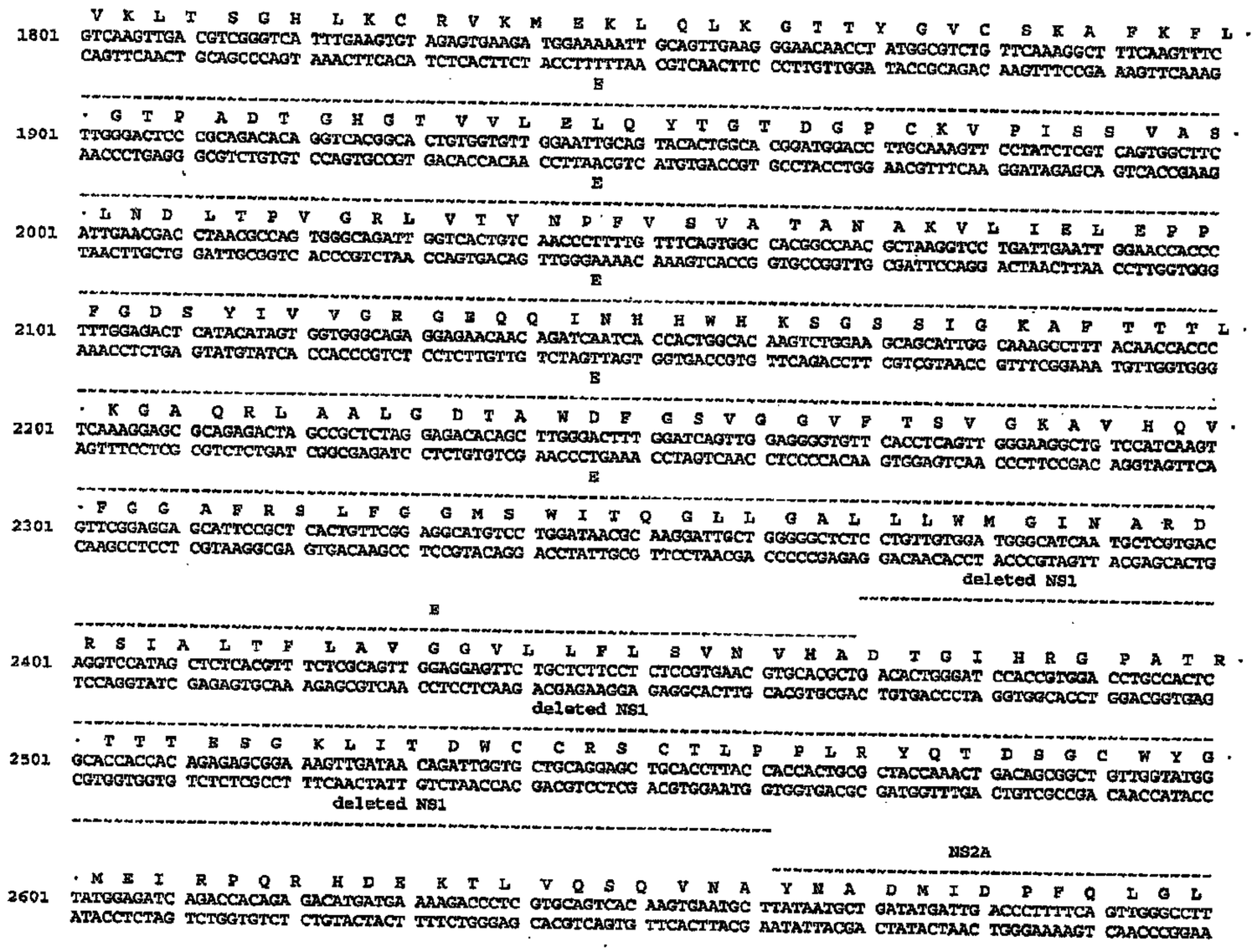

1001 CCACTAACCA CGACTACACT ATATGGATGC CTGAGAACCC TCGACTCGGT ATGAGTTGCG ACATTTTTAC GAACTCACGG GGCAAGCGGG CATCTAAGGG
GGTGATTGGT GCTGATGTGA TATACCTACG GACTCTTGGG AGCTGAGCCA TACTCAACGC TGTAAAAATG CTTGAGTGCC CCGTTCGCCC GTAGATTCCC
RAbies-G protein
· S E T C G F V D E R G L Y K S L K G A C K L K L C G V L G L R L M
1101 GTCTGAAACA TGCGGGTTTG TTGATGAGCG GGGGTTGTAT AAATCTCTTA AAGCGCCTG TAAGCTGAAA CTCTGTGGCG TACTGGGGCT GCGCCTGATG
CAGACTTTGT ACGCCCAAAC AACTACTCGC CCCCAACATA TTTAGAGAAT TTCCGCGGAC ATTCGACTTT GAGACACCGC ATGACCCCGA CGCGGACTAC
RAbies-G protein
D G T W V A M Q T S N E T K W C P P G Q L V N L H D F R S D E I E H ·
1201 GACGGCACAT GGGTGGCTAT GCAGACAAGC AATGAAACAA AGTGGTGTCC CCCIGGTCAG CTGGTTAATC TGCACGACTT TAGGTCTGAC GAAATCGAGC
CTGCCGTGTA CCCACCGATA CGTCTGTTCG TTACTTTGTT TCACCACAGG GGGACCAGTC GACCAATTAG ACGTGCTGAA ATCCAGACTG CTTTAGCTCG
RAbies-G protein
· L V V E E L V K K R E E C L D A L E S I M T T K S V S F R R L S H ·
1301 ACCTTGTGGT GGAGGAACTG GTGAAGAAAC GCGAAGAGTG CCTGGACGCA CTTGAGAGTA TTATGACCAC CAAATCCGTT TCCTTCAGAA GACTGAGCCA
TGGAACACCA CCTCCTTGAC CACTTCTTTG CGCTTCTCAC GGACCTGCGT GAACTCTCAT AATACTGGTG GTTTAGGCAA AGGAAGTCTT CTGACTCGGT
RAbies-G protein
· L R K L V P G F G K A Y T I F N K T L M E A D A H Y K S V R T W N
1401 CCTGCGAAAG CTGGTGCCAG GGTTCGGGAA GGCTTATACT ATTTTCAACA AGACTCTTAT GGAGGCGGAT GCCCATTATA AGTCAGTTAG GACTTGGAAT
GGACGCTTTC GACCACGGTC CCAAGCCCTT CCGAATATGA TAAAAGTTGT TCTGAGAATA CCTCCGCCTA CGGGTAATAT TCAGTCAATC CTGAACCTTA
RAbies-G protein
E I I P S K G C L R V G G R C H P H V N G V F F N G I I L G P D G N ·
1501 GAGATAATTC CCTCCAAAGG ATGTCTGAGA GTCGGTGGGA GATGCCACCC CCATGTCAAT GGGGTGTTCT TTAACGGAAT CATCCTGGGA CCTGACGGGA
CTCTATTAAG GGAGGTTTCC TACAGACTCT CAGCCACCCT CTACGGTGGG GGTACAGTTA CCCCACAAGA AATTGCCTTA GTAGGACCCT GGACTGCCCT
RAbies-G protein
· V L I P E M Q S S L L Q Q H M E L L V S S V I P L M H P L A D P S ·
1601 ACGTGCTGAT TCCCGAGATG CAATCTTCCC TTCTGCAGCA ACACATGGAA CTCCTGGTGT CTTCAGTGAT ACCCCTGATG CACCCACTGG CCGACCCCAG
TGCACGACTA AGGGCTCTAC GTTAGAAGGG AAGACGTCGT TGTGTACCTT GAGGACCACA GAAGTCACTA TGGGGACTAC GTGGGTGACC GGCTGGGGTC
RAbies-G protein
· T V F K N G D E A E D F V E V H L P D V H E R I S G V D L G L P N
1701 CACTGTGTTC AAAAATGGCG ATGAGGCCGA AGACTTTGTG GAAGTTCACC TGCCCGATGT ACACGAAAGG ATATCTGGAG TAGACCTGGG CCTTCCTAAT
GTGACACAAG TTTTTACCGC TACTCCGGCT TCTGAAACAC CTTCAAGTGG ACGGGCTACA TGTGCTTTCC TATAGACCTC ATCTGGACCC GGAAGGATTA
RAbies-G protein
W G K Y V L L S A G A L T A L M L I I F L M T C W R R V N R S E P T ·
1801 TGGGGTAAGT ACGTGCTCCT GAGTGCGGGT GCCTTGACCG CTTTGATGCT GATCATTTTT CTGATGACCT GCTGGCGGAG GGTGAATCGC TCCGAGCCGA
ACCCCATTCA TGCACGAGGA CTCACGCCCA CGGAACTGGC GAAACTACGA CTAGTAAAAA GACTACTGGA CGACCGCCTC CCACTTAGCG AGGCTCGGCT
RAbies-G protein
· Q H N L R G T G R E V S V T P Q S G K I I S S W E S Y K S G G E T ·
1901 CACAGCACAA TCTCAGAGGG ACAGGCCGGG AAGTAAGTGT GACTCCGCAA TCTGGCAAGA TTATTAGTAG TTGGGAGAGT TACAAGTCTG GAGGAGAGAC
GTGTCGTGTT AGAGTCTCCC TGTCCGGCCC TTCATTCACA CTGAGGCGTT AGACCGTTCT AATAATCATC AACCCTCTCA ATGTTCAGAC CTCCTCTCTG
FMDV 2A NS1 signal
RAbies-G protein preNS1 signal
· G L N F D L L K L A G D V E S N P G P A E D R S I A L T F L A V G
2001 TGGGTTGAAT TTTGATCTGC TCAAACTTGC AGGCGATGTA GAATCAAATC CTGGACCCGC CGGGACAGG TCCATAGCTC TCACGTTTCT CGCAGTTGGA
ACCCAACTTA AAACTAGACG AGTTTGAACG TCCGCTACAT CTTAGTTTAG GACCTGGGCG GGCCCTGTCC AGGTATCGAG AGTGCAAAGA GCGTCAACCT
NS1
NS1 signal
G V L L F L S V N V H A D T G C A I D I S R Q E L R C G S G V F I H ·
2101 GGAGTTCTGC TCTTCCTCTC CGTGAACGTG CACGCTGACA CTGGGTGTGC CATAGACATC AGCCGGCAAG AGCTGAGATG TGGAAGTGGA GTGTTCATAC
CCTCAAGACG AGAAGGAGAG GCACTTGCAC GTGCGACTGT GACCCACACG GTATCTGTAG TCGGCCGTTC TCGACTCTAC ACCTTCACCT CACAAGTATG

Sequence Appendix 4 (continued)

PIV-WNV helper ΔNS1

```
                                                                                                                                C
                                                                                                                               -----
                                                                5' UTR
-------------------------------------------------------------------------------------------------------------------------------
                                                                                                                           M  S
  1 AGTAGTTCGC CTGTGTGAGC TGACAAACTT AGTAGTGTTT GTGAGGATTA ACAACAATTA ACACAGTGCG AGCTGTTTCT TAGCACGAAG ATCTCGATGT
    TCATCAAGCG GACACACTCG ACTGTTTGAA TCATCACAAA CACTCCTAAT TGTTGTTAAT TGTGTCACGC TCGACAAAGA ATCGTGCTTC TAGAGCTACA
                                                                C
-------------------------------------------------------------------------------------------------------------------------------
 · K  K  P  G  G  P  G  K  S  R  A  V  N  M  L  K  R  G  M  P  R  V  L  S  L  I  G  L  K  R  A  M  L ·
101 CTAAGAAACC AGGAGGGCCC GGCAAGAGCC GGGCTGTCAA TATGCTAAAA CGCGGAATGC CCCGCGTGTT GTCCTTGATT GGACTTAAGA GGGCTATGTT
    GATTCTTTGG TCCTCCCGGG CCGTTCTCGG CCCGACAGTT ATACGATTTT GCGCCTTACG GGGCGCACAA CAGGAACTAA CCTGAATTCT CCCGATACAA
                                                                C
-------------------------------------------------------------------------------------------------------------------------------
 · S  L  I  D  G  K  G  P  I  R  F  V  L  A  L  L  A  F  F  R  F  T  A  I  A  P  T  R  A  V  L  D  R
201 GAGCCTGATC GACGGCAAGG GGCCAATACG ATTTGTGTTG GCTCTCTTGG CGTTCTTCAG GTTCACAGCA ATTGCTCCGA CCCGAGCAGT GCTGGATCGA
    CTCGGACTAG CTGCCGTTCC CCGGTTATGC TAAACACAAC CGAGAGAACC GCAAGAAGTC CAAGTGTCGT TAACGAGGCT GGGCTCGTCA CGACCTAGCT
                                                                C
-------------------------------------------------------------------------------------------------------------------------------
 W  R  G  V  N  K  Q  T  A  M  K  H  L  L  S  F  K  K  E  L  G  T  L  T  S  A  I  N  R  R  S  S  K  Q ·
301 TGGAGAGGTG TGAACAAACA AACAGCGATG AAACACCTTC TGAGTTTCAA GAAGGAACTA GGGACCTTGA CCAGTGCTAT CAATCGGCGG AGCTCAAAAC
    ACCTCTCCAC ACTTGTTTGT TTGTCGCTAC TTTGTGGAAG ACTCAAAGTT CTTCCTTGAT CCCTGGAACT GGTCACGATA GTTAGCCGCC TCGAGTTTTG
                                                Signal peptide·
                                                ---------------------------------------------
      C
      -----------                                                                  prM
                                                                                   ---------------------------------------
 · K  K  R  G  G  K  T  G  I  A  V  M  I  G  L  I  A  S  V  G  A  V  T  L  S  N  F  Q  G  K  V  M  M ·
401 AAAAGAAAAG AGGAGGAAAG ACCGGAATTG CAGTCATGAT TGGCCTGATC GCCAGCGTAG GAGCAGTTAC CCTCTCTAAC TTCCAAGGGA AGGTGATGAT
    TTTTCTTTTC TCCTCCTTTC TGGCCTTAAC GTCAGTACTA ACCGGACTAG CGGTCGCATC CTCGTCAATG GGAGAGATTG AAGGTTCCCT TCCACTACTA
                                                prM
-------------------------------------------------------------------------------------------------------------------------------
 · T  V  N  A  T  D  V  T  D  V  I  T  I  P  T  A  A  G  K  N  L  C  I  V  R  A  M  D  V  G  Y  M  C
501 GACGGTAAAT GCTACTGACG TCACAGATGT CATCACGATT CCAACAGCTG CTGGAAAGAA CCTATGCATT GTCAGAGCAA TGGATGTGGG ATACATGTGC
    CTGCCATTTA CGATGACTGC AGTGTCTACA GTAGTGCTAA GGTTGTCGAC GACCTTTCTT GGATACGTAA CAGTCTCGTT ACCTACACCC TATGTACACG
                                                prM
-------------------------------------------------------------------------------------------------------------------------------
 D  D  T  I  T  Y  E  C  P  V  L  S  A  G  N  D  P  E  D  I  D  C  W  C  T  K  S  A  V  Y  V  R  Y  G ·
601 GATGATACTA TCACTTATGA ATGCCCAGTG CTGTCGGCTG GTAATGATCC AGAAGACATC GACTGTTGGT GCACAAAGTC AGCAGTCTAC GTCAGGTATG
    CTACTATGAT AGTGAATACT TACGGGTCAC GACAGCCGAC CATTACTAGG TCTTCTGTAG CTGACAACCA CGTGTTTCAG TCGTCAGATG CAGTCCATAC
                                                prM
-------------------------------------------------------------------------------------------------------------------------------
 · R  C  T  K  T  R  H  S  R  R  S  R  R  S  L  T  V  Q  T  H  G  E  S  T  L  A  N  K  K  G  A  W  M ·
701 GAAGATGCAC CAAGACACGC CACTCAAGAC GCAGTCGGAG GTCACTGACA GTGCAGACAC ACGGAGAAAG CACTCTAGCG AACAAGAAGG GGGCTTGGAT
    CTTCTACGTG GTTCTGTGCG GTGAGTTCTG CGTCAGCCTC CAGTGACTGT CACGTCTGTG TGCCTCTTTC GTGAGATCGC TTGTTCTTCC CCCGAACCTA
```




E
deleted NS1
NS2A

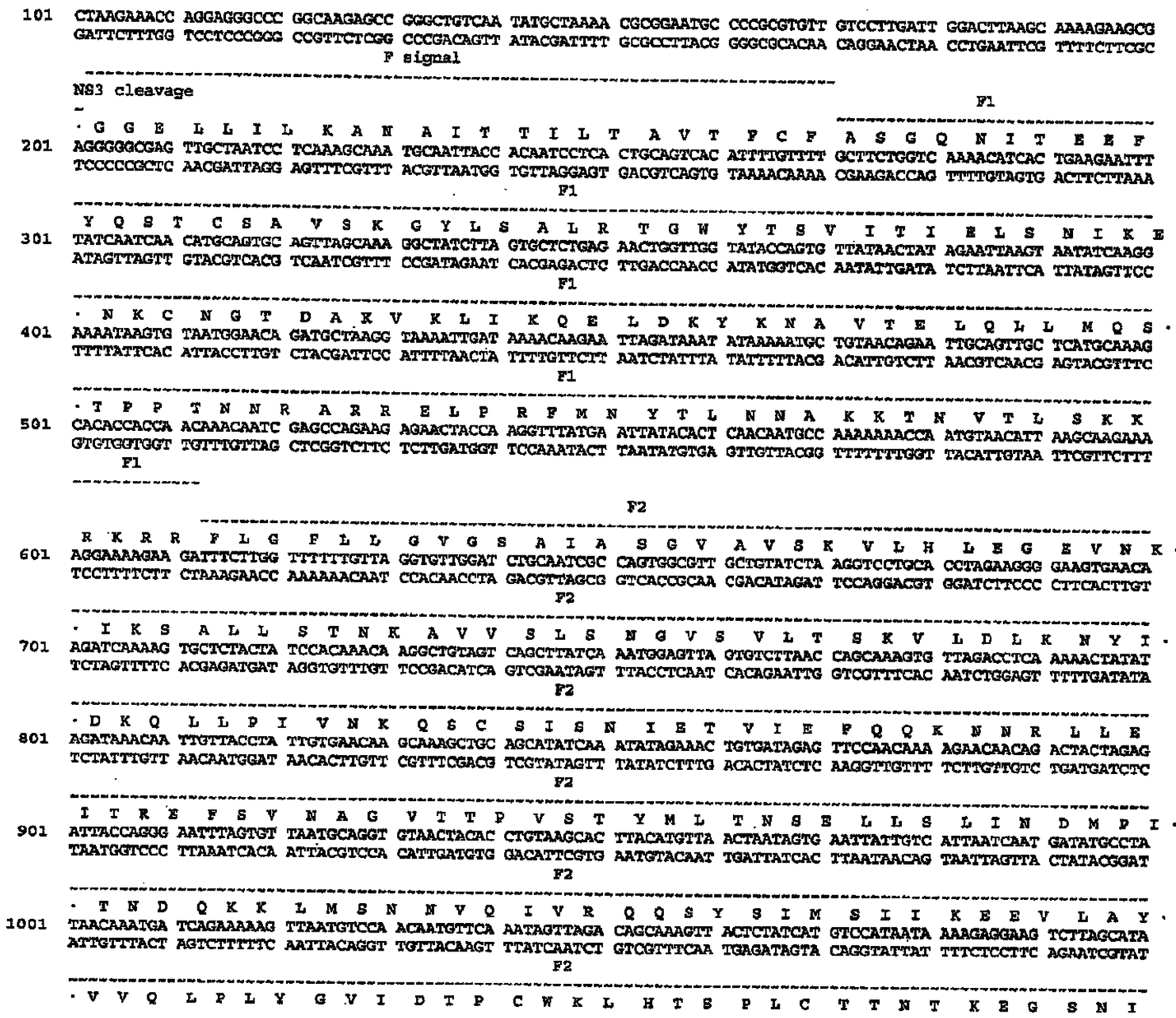
Sequence Appendix 5

PIV-WNV(ΔprME)/RSV-F

```
                                                
```

```
                                                
```

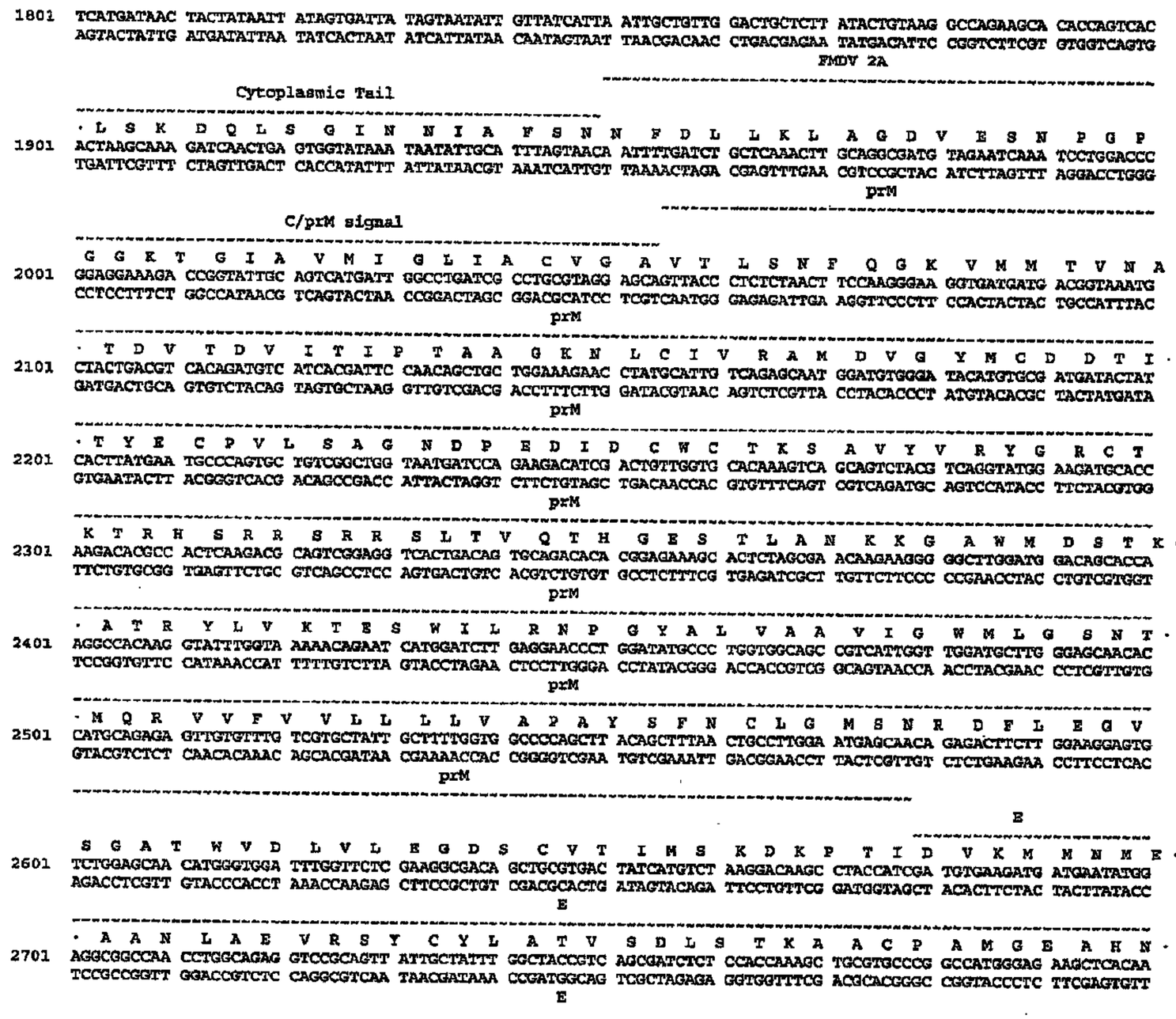
```
                                                
```


F signal
NS3 cleavage
F1
F2

```
1101  TGTAGTACAA TTACCACTAT ATGGTGTTAT AGATACACCC TGTTGGAAAC TACACACATC CCCTCTATGT ACAACCAACA CAAAAGAAGG GTCCAACATC
      ACATCATGTT AATGGTGATA TACCACAATA TCTATGTGGG ACAACCTTTG ATGTGTGTAG GGGAGATACA TGTTGGTTGT GTTTTCTTCC CAGGTTGTAG
                                                F2
      ------------------------------------------------------------------------------------------------------------
       C  L  T  R  T  D  R  G  W  Y  C  D  N  A  G  S  V  S  F  F  P  Q  A  E  T  C  K  V  Q  S  N  R  V  F
1201  TGTTTAACAA GAACTGACAG AGGATGGTAC TGTGACAATG CAGGATCAGT ATCTTTCTTC CCACAAGCTG AAACATGTAA AGTTCAATCA AATCGAGTAT
      ACAAATTGTT CTTGACTGTC TCCTACCATG ACACTGTTAC GTCCTAGTCA TAGAAAGAAG GGTGTTCGAC TTTGTACATT TCAAGTTAGT TTAGCTCATA
                                                F2
      ------------------------------------------------------------------------------------------------------------
       C  D  T  M  N  S  L  T  L  P  S  E  I  N  L  C  N  V  D  I  F  N  P  K  Y  D  C  K  I  M  T  S  K
1301  TTTGTGACAC AATGAACAGT TTAACATTAC CAAGTGAAAT AAATCTCTGC AATGTTGACA TATTCAACCC CAAATATGAT TGTAAAATTA TGACTTCAAA
      AAACACTGTG TTACTTGTCA AATTGTAATG GTTCACTTTA TTTAGAGACG TTACAACTGT ATAAGTTGGG GTTTATACTA ACATTTTAAT ACTGAAGTTT
                                                F2
      ------------------------------------------------------------------------------------------------------------
       T  D  V  S  S  S  V  I  T  S  L  G  A  I  V  S  C  Y  G  K  T  K  C  T  A  S  N  K  N  R  G  I  I
1401  AACAGATGTA AGCAGCTCCG TTATCACATC TCTAGGAGCC ATTGTGTCAT GCTATGGCAA AACTAAATGT ACAGCATCCA ATAAAAATCG TGGAATCATA
      TTGTCTACAT TCGTCGAGGC AATAGTGTAG AGATCCTCGG TAACACAGTA CGATACCGTT TTGATTTACA TGTCGTAGGT TATTTTTAGC ACCTTAGTAT
                                                F2
      ------------------------------------------------------------------------------------------------------------
       K  T  F  S  N  G  C  D  Y  V  S  N  K  G  N  D  T  V  S  V  G  N  T  L  Y  Y  V  N  K  Q  E  G  K  S
1501  AAGACATTTT CTAACGGGTG CGATTATGTA TCAAATAAAG GGATGGACAC TGTGTCTGTA GGTAACACAT TATATTATGT AAATAAGCAA GAAGGTAAAA
      TTCTGTAAAA GATTGCCCAC GCTAATACAT AGTTTATTTC CCTACCTGTG ACACAGACAT CCATTGTGTA ATATAATACA TTTATTCGTT CTTCCATTTT
                                                F2
      ------------------------------------------------------------------------------------------------------------
       L  Y  V  K  G  E  P  I  I  N  F  Y  D  P  L  V  F  P  S  D  E  F  D  A  S  I  S  Q  V  N  E  K  I
1601  GTCTCTATGT AAAAGGTGAA CCAATAATAA ATTTCTATGA CCCATTAGTA TTCCCCTCTG ATGAATTTGA TGCATCAATA TCTCAAGTCA ACGAGAAGAT
      CAGAGATACA TTTTCCACTT GGTTATTATT TAAAGATACT GGGTAATCAT AAGGGGAGAC TACTTAAACT ACGTAGTTAT AGAGTTCAGT TGCTCTTCTA
                                                F2
      ------------------------------------------------------------------------------------------------------------
                                                                                                TM Domain
                                                                                               ----------
       N  Q  S  L  A  F  I  R  K  S  D  E  L  L  H  N  V  N  A  G  K  S  T  T  N  I  M  I  T  T  I  I  I
1701  TAACCAGAGC CTAGCATTTA TTCGTAAATC CGATGAATTA TTACATAATG TAAATGCTGG TAAATCCACC ACAAATATCA TGATAACTAC TATAATTATA
      ATTGGTCTCG GATCGTAAAT AAGCATTTAG GCTACTTAAT AATGTATTAC ATTTACGACC ATTTAGGTGG TGTTTATAGT ACTATTGATG ATATTAATAT
                TM Domain
      ----------
                                                                                Cytoplasmic Tail
                                                                               ----------
       V  I  I  V  I  L  L  S  L  I  A  V  G  L  L  L  Y  C  K  A  R  S  T  P  V  T  L  S  K  D  Q  L  S  G
1801  GTGATTATAG TAATATTGTT ATCATTAATT GCTGTTGGAC TGCTCTTATA CTGTAAGGCC AGAAGCACAC CAGTCACACT AAGCAAAGAT CAACTGAGTG
      CACTAATATC ATTATAACAA TAGTAATTAA CGACAACCTG ACGAGAATAT GACATTCCGG TCTTCGTGTG GTCAGTGTGA TTCGTTTCTA GTTGACTCAC
                                                FMDV 2A
      ------------------------------------------------------------------------------------------------------------
                                                                                                membrane domain of WNV E (split)
                                                                                               ----------
                                                                                                pre E/NS1 signal
                                                                                               ----------
      Cytoplasmic Tail
      ----------
       I  N  N  I  A  F  S  N  N  F  D  L  L  K  L  A  G  D  V  E  S  N  P  G  P  A  R  D  R  S  I  A  L
1901  GTATAAATAA TATTGCATTT AGTAACAATT TTGATCTGCT CAAACTTGCA GGCGATGTAG AATCAAATCC TGGACCCGCC CGGGACAGGT CCATAGCTCT
      CATATTTATT ATAACGTAAA TCATTGTTAA AACTAGACGA GTTTGAACGT CCGCTACATC TTAGTTTAGG ACCTGGGCGG GCCCTGTCCA GGTATCGAGA
             Transmembrane domain of WNV E (split)
      ------------------------------------------------------------
                                                                                                NS1
                                                                                               ----------
```

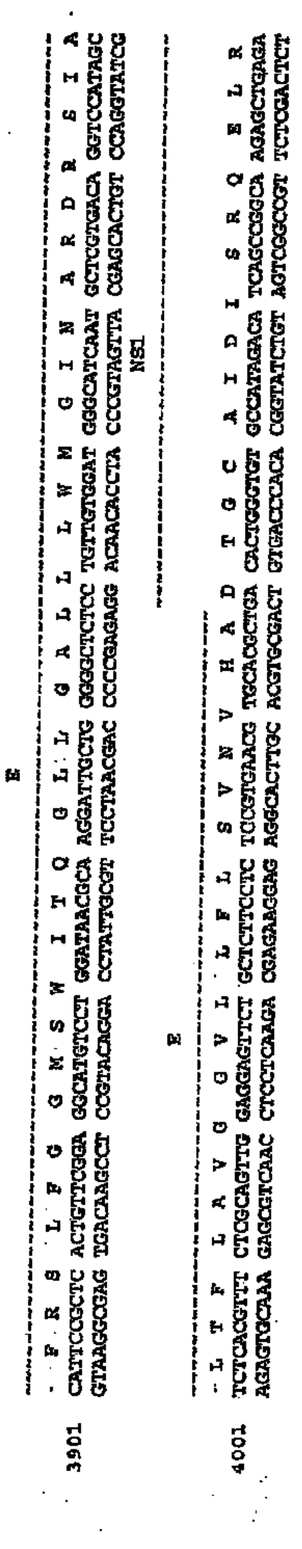
```
                 T  F  L  A  V  G  G  V  L  L  P  L  S  V  N  V  H  A  D  T  G  C  A  I  D  I  S  R  Q  E  L  R
2001  CACGTTTCTC GCAGTTGGAG GAGTTCTGCT CTTCCTCTCC GTGAACGTGC ACGCTGACAC TGGGTGTGCC ATAGACATCA GCCGGCAAGA GCTGAGA
      GTGCAAAGAG CGTCAACCTC CTCAAGACGA GAAGGAGAGG CACTTGCACG TGCGACTGTG ACCCACACGG TATCTGTAGT CGGCCGTTCT CGACTCT

PIV-WNV(ΔC)/RSV-F

   1  GATCCTAATA CGACTCACTA TAGAGTAGTT CGCCTGTGTG AGCTGACAAA CTTAGTAGTG TTTGTGAGGA TTAACAACAA TTAACACAGT GCGAGCTGTT
      CTAGGATTAT GCTGAGTGAT ATCTCATCAA GCGGACACAC TCGACTGTTT GAATCATCAC AAACACTCCT AATTGTTGTT AATTGTGTCA CGCTCGACAA
                                                    N-terminus of C
                        ------------------------------------------------------------------------------------------
                          M  S  K  K  P  G  G  P  G  K  S  R  A  V  N  M  L  K  R  G  N  P  R  V  L  S  L
101   TCTTAGCACG AAGATCTCGA TGTCTAAGAA ACCAGGAGGG CCCGGCAAGA GCCGGGCTGT CAATATGCTA AAACGCGGAA TGCCCCGCGT GTTGTCCTTG
      AGAATCGTGC TTCTAGAGCT ACAGATTCTT TGGTCCTCCC GGGCCGTTCT CGGCCCGACA GTTATACGAT TTTGCGCCTT ACGGGGCGCA CAACAGGAAC
      N-terminus of C                                F signal
      ----------  ------------------------------------------------------------------------------------------
                  NS3 cleavage                                                                          F1
                  ----------------                                                                      -------
                I  G  L  K  Q  X  K  R  G  G  E  L  L  I  L  K  A  N  A  I  T  T  I  L  T  A  V  T  F  C  F  A  S  G
201   ATTGGACTTA AGCAAAAGAA GCGAGGGGGC GAGTTGCTAA TCCTCAAAGC AAATGCAATT ACCACAATCC TCACTGCAGT CACATTTTGT TTTGCTTCTG
      TAACCTGAAT TCGTTTTCTT CGCTCCCCCG CTCAACGATT AGGAGTTTCG TTTACGTTAA TGGTGTTAGG AGTGACGTCA GTGTAAAACA AAACGAAGAC
                                                    F1
      ------------------------------------------------------------------------------------------------------
                  Q  N  I  T  E  E  F  Y  Q  S  T  C  S  A  V  S  K  G  Y  L  S  A  L  R  T  G  W  Y  T  S  V  I  T
301   GTCAAAACAT CACTGAAGAA TTTTATCAAT CAACATGCAG TGCAGTTAGC AAAGGCTATC TTAGTGCTCT GAGAACTGGT TGGTATACCA GTGTTATAAC
      CAGTTTTGTA GTGACTTCTT AAAATAGTTA GTTGTACGTC ACGTCAATCG TTTCCGATAG AATCACGAGA CTCTTGACCA ACCATATGGT CACAATATTG
                                                    F1
      ------------------------------------------------------------------------------------------------------
                I  E  L  S  N  I  K  E  N  K  C  N  G  T  D  A  K  V  K  L  I  K  Q  E  L  D  K  Y  K  N  A  V  T
401   TATAGAATTA AGTAATATCA AGGAAAATAA GTGTAATGGA ACAGATGCTA AGGTAAAATT GATAAAACAA GAATTAGATA AATATAAAAA TGCTGTAACA
      ATATCTTAAT TCATTATAGT TCCTTTTATT CACATTACCT TGTCTACGAT TCCATTTTAA CTATTTTGTT CTTAATCTAT TTATATTTTT ACGACATTGT
                                                    F1
      ------------------------------------------------------------------------------------------------------
                E  L  Q  L  L  M  Q  S  T  P  P  T  N  N  R  A  R  R  E  L  P  R  F  M  N  Y  T  L  N  N  A  K  K  T
501   GAAITGCAGT TGCTCATGCA AAGCACACCA CCAACAAACA ATCGAGCCAG AAGAGAACTA CCAAGGTTTA TGAATTATAC ACTCAACAAT GCCAAAAAAA
      CTTAACGTCA ACGAGTACGT TTCGTGTGGT GGTTGTTTGT TAGCTCGGTC TTCTCTTGAT GGTTCCAAAT ACTTAATATG TGAGTTGTTA CGGTTTTTTT
                                                    F2
                        ------------------------------------------------------------------------------------------
                F1
      ----------------
                N  V  T  L  S  K  K  R  K  R  R  F  L  G  F  L  L  G  V  G  S  A  I  A  S  G  V  A  V  S  K  V  L
601   CCAATGTAAC ATTAAGCAAG AAAAGGAAAA GAAGATTTCT TGGTTTTTTG TTAGGTGTTG GATCTGCAAT CGCCAGTGGC GTTGCTGTAT CTAAGGTCCT
      GGTTACATTG TAATTCGTTC TTTTCCTTTT CTTCTAAAGA ACCAAAAAAC AATCCACAAC CTAGACGTTA GCGGTCACCG CAACGACATA GATTCCAGGA
                                                    F2
      ------------------------------------------------------------------------------------------------------
                H  L  E  G  E  V  N  K  I  K  S  A  L  L  S  T  N  K  A  V  V  S  L  S  N  G  V  S  V  L  T  S  K
701   GCACCTAGAA GGGGAAGTGA ACAAGATCAA AAGTGCTCTA CTATCCACAA ACAAGGCTGT AGTCAGCTTA TCAAATGGAG TTAGTGTCTT AACCAGCAAA
      CGTGGATCTT CCCCTTCACT TGTTCTAGTT TTCACGAGAT GATAGGTGTT TGTTCCGACA TCAGTCGAAT AGTTTACCTC AATCACAGAA TTGGTCGTTT
                                                    F2
      ------------------------------------------------------------------------------------------------------
                V  L  D  L  K  N  Y  I  D  K  Q  L  L  P  I  V  N  K  Q  S  C  S  I  S  N  I  E  T  V  I  E  F  Q  Q
```

```
801  GTGTTAGACC TCAAAAACTA TATAGATAAA CAATTGTTAC CTATTGTGAA CAAGCAAAGC TGCAGCATAT CAAATATAGA AACTGTGATA GAGTTCCAAC
     CACAATCTGG AGTTTTTGAT ATATCTATTT GTTAACAATG GATAACACTT GTTCGTTTCG ACGTCGTATA GTTTATATCT TTGACACTAT CTCAAGGTTG
                                                F2
     ---------------------------------------------------------------------------------------------------------
     · K N N  R L L  E I T  R E F S  V N A  G V T T  F V S  T Y M  L T N S  E L L ·
901  AAAAGAACAA CAGACTACTA GAGATTACCA GGGAATTTAG TGTTAATGCA GGTGTAACTA CACCTGTAAG CACTTACATG TTAACTAATA GTGAATTATT
     TTTTCTTGTT GTCTGATGAT CTCTAATGGT CCCTTAAATC ACAATTACGT CCACATTGAT GTGGACATTC GTGAATGTAC AATTGATTAT CACTTAATAA
                                                F2
     ---------------------------------------------------------------------------------------------------------
     · S L I  N D M P  I T N  D Q K  K L M S  N N V  Q I V  R Q Q S  Y S I  M S I
1001 GTCATTAATC AATGATATGC CTATAACAAA TGATCAGAAA AAGTTAATGT CCAACAATGT TCAAATAGTT AGACAGCAAA GTTACTCTAT CATGTCCATA
     CAGTAATTAG TTACTATACG GATATTGTTT ACTAGTCTTT TTCAATTACA GGTTGTTACA AGTTTATCAA TCTGTCGTTT CAATGAGATA GTACAGGTAT
                                                F2
     ---------------------------------------------------------------------------------------------------------
      I K E  E V L A  Y V V  Q L P L  Y G V  I D T  P C W K  L H T  S P L  C T T N ·
1101 ATAAAAGAGG AAGTCTTAGC ATATGTAGTA CAATTACCAC TATATGGTGT TATAGATACA CCCTGTTGGA AACTACACAC ATCCCCTCTA TGTACAACCA
     TATTTTCTCC TTCAGAATCG TATACATCAT GTTAATGGTG ATATACCACA ATATCTATGT GGGACAACCT TTGATGTGTG TAGGGGAGAT ACATGTTGGT
                                                F2
     ---------------------------------------------------------------------------------------------------------
     · T K E  G S N  I C L T  R T D  R G W  Y C D N  A G S  V S F  F P Q A  E T C ·
1201 ACACAAAAGA AGGGTCCAAC ATCTGTTTAA CAAGAACTGA CAGAGGATGG TACTGTGACA ATGCAGGATC AGTATCTTTC TTCCCACAAG CTGAAACATG
     TGTGTTTTCT TCCCAGGTTG TAGACAAATT GTTCTTGACT GTCTCCTACC ATGACACTGT TACGTCCTAG TCATAGAAAG AAGGGTGTTC GACTTTGTAC
                                                F2
     ---------------------------------------------------------------------------------------------------------
     · K V Q  S N R V  F C D  T M N  S L T L  P S E  I N L  C N V D  I F N  P K Y
1301 TAAAGTTCAA TCAAATCGAG TATTTTGTGA CACAATGAAC AGTTTAACAT TACCAAGTGA AATAAATCTC TGCAATGTTG ACATATTCAA CCCCAAATAT
     ATTTCAAGTT AGTTTAGCTC ATAAAACACT GTGTTACTTG TCAAATTGTA ATGGTTCACT TTATTTAGAG ACGTTACAAC TGTATAAGTT GGGGTTTATA
                                                F2
     ---------------------------------------------------------------------------------------------------------
      D C K I  M T S  K T D  V S S S  V I T  S L G  A I V S  C Y G  K T K  C T A S ·
1401 GATTGTAAAA TTATGACTTC AAAAACAGAT GTAAGCAGCT CCGTTATCAC ATCTCTAGGA GCCATTGTGT CATGCTATGG CAAAACTAAA TGTACAGCAT
     CTAACATTTT AATACTGAAG TTTTTGTCTA CATTCGTCGA GGCAATAGTG TAGAGATCCT CGGTAACACA GTACGATACC GTTTTGATTT ACATGTCGTA
                                                F2
     ---------------------------------------------------------------------------------------------------------
     · N K N  R G I  I K T F  S N G  C D Y  V S N K  G M D  T V S  V G N T  L Y Y ·
1501 CCAATAAAAA TCGTGGAATC ATAAAGACAT TTTCTAACGG GTGCGATTAT GTATCAAATA AAGGGATGGA CACTGTGTCT GTAGGTAACA CATTATATTA
     GGTTATTTTT AGCACCTTAG TATTTCTGTA AAAGATTGCC CACGCTAATA CATAGTTTAT TTCCCTACCT GTGACACAGA CATCCATTGT GTAATATAAT
                                                F2
     ---------------------------------------------------------------------------------------------------------
     · V N K  Q E G K  S L Y  V K G  E P I  I N F Y  D P L  V F P S  D E F  D A S
1601 TGTAAATAAG CAAGAAGGTA AAAGTCTCTA TGTAAAAGGT GAACCAATAA TAAATTTCTA TGACCCATTA GTATTCCCCT CTGATGAATT TGATGCATCA
     ACATTTATTC GTTCTTCCAT TTTCAGAGAT ACATTTTCCA CTTGGTTATT ATTTAAAGAT ACTGGGTAAT CATAAGGGGA GACTACTTAA ACTACGTAGT
                                                F2
     ---------------------------------------------------------------------------------------------------------
                                                                                                    TM Domain
                                                                                                    -
      I S Q V  N E K  I N Q  S L A F  I R K  S D E  L L H N  V N A  G K S  T T N  I ·
1701 ATATCTCAAG TCAACGAGAA GATTAACCAG AGCGTAGCAT TTATTCGTAA ATCCGATGAA TTATTACATA ATGTAAATGC TGGTAAATCC ACCACAAATA
     TATAGAGTTC AGTTGCTCTT CTAATTGGTC TCGGATCGTA AATAAGCATT TAGGCTACTT AATAATGTAT TACATTTACG ACCATTTAGG TGGTGTTTAT
                                    TM Domain
     ---------------------------------------------------------------
                                                                  Cytoplasmic Tail
                                                                  ---------------------------------
     · M I T  T I I  I V I I  V I L  L S L  I A V G  L L L  Y C K  A R S T  P V T ·
```


Cytoplasmic Tail
FMDV 2A
C/prM signal
prM
E

CA 02755225 2011-09-12

```
                                                
```

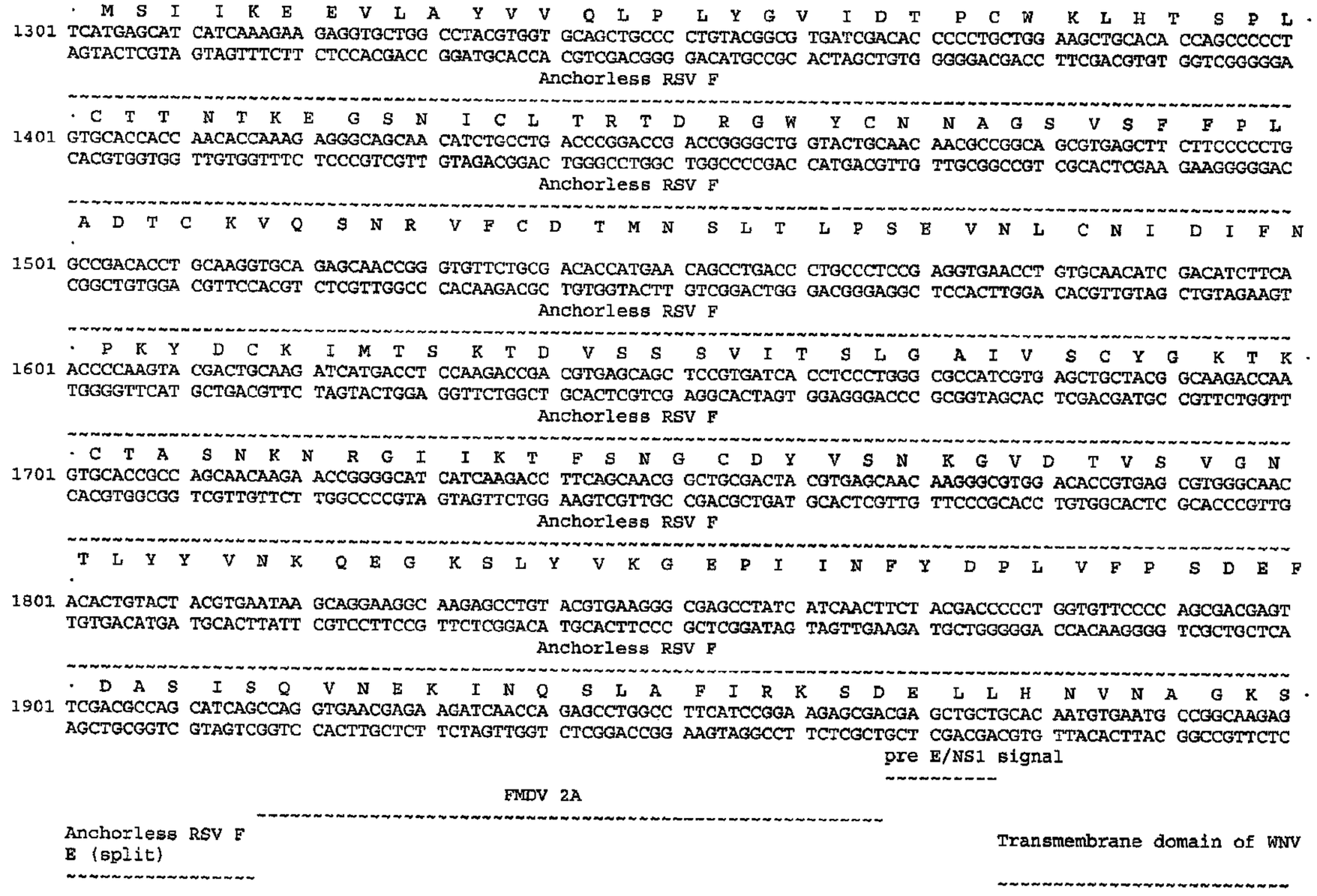


SEQUENCE APPENDIX 6

RepliVax WN - Anchorless F inserted in place of ΔprM-E.
F insert starts at nucleotide position 439bp and ends at 2016bp.

```
                                                                                                                  C
                                                                                                                  ----
                                                5' UTR
     ~~~
```

```
501 CTTCGCCAGC AGCCAGAACA TCACCGAGGA ATTCTACCAG AGCACCTGCA GCGCCGTGAG CAAGGGCTAC CTGAGCGCCC TGCGGACCGG CTGGTACACC
    GAAGCGGTCG TCGGTCTTGT AGTGGCTCCT TAAGATGGTC TCGTGGACGT CGCGGCACTC GTTCCCGATG GACTCGCGGG ACGCCTGGCC GACCATGTGG
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     S  V  I  T  I  E  L  S  N  I  K  E  N  K  C  N  G  T  D  A  K  V  K  L  I  K  Q  E  L  D  K  Y  K  N
    .
601 AGCGTGATCA CCATCGAGCT GTCCAACATC AAAGAAAACA AGTGCAACGG CACCGACGCC AAGGTGAAAC TGATCAAGCA GGAACTGGAC AAGTACAAGA
    TCGCACTAGT GGTAGCTCGA CAGGTTGTAG TTTCTTTTGT TCACGTTGCC GTGGCTGCGG TTCCACTTTG ACTAGTTCGT CCTTGACCTG TTCATGTTCT
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    .  A  V  T  E  L  Q  L  L  M  Q  S  T  P  A  A  N  N  R  A  R  R  E  L  P  R  F  M  N  Y  T  L  N  N  .
701 ACGCCGTGAC CGAGCTGCAG CTGCTGATGC AGAGCACCCC TGCCGCCAAC AACCGGGCCA GACGCGAGCT GCCCCGGTTC ATGAACTACA CCCTGAACAA
    TGCGGCACTG GCTCGACGTC GACGACTACG TCTCGTGGGG ACGGCGGTTG TTGGCCCGGT CTGCGCTCGA CGGGGCCAAG TACTTGATGT GGGACTTGTT
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    .  A  K  K  T  N  V  T  L  S  K  K  R  K  R  R  F  L  G  F  L  L  G  V  G  S  A  I  A  S  G  I  A  V
801 CGCCAAGAAA ACCAACGTGA CCCTGAGCAA GAAGCGGAAG CGGCGGTTCC TGGGCTTCCT GCTGGGCGTG GGCAGCGCCA TCGCCAGCGG CATCGCCGTG
    GCGGTTCTTT TGGTTGCACT GGGACTCGTT CTTCGCCTTC GCCGCCAAGG ACCCGAAGGA CGACCCGCAC CCGTCGCGGT AGCGGTCGCC GTAGCGGCAC
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     S  K  V  L  H  L  E  G  E  V  N  K  I  K  S  A  L  L  S  T  N  K  A  V  V  S  L  S  N  G  V  S  V  L
    .
901 TCCAAGGTGC TGCACCTGGA AGGCGAGGTG AACAAGATCA AGTCCGCCCT GCTGTCCACC AACAAGGCCG TGGTGTCCCT GAGCAACGGC GTGAGCGTGC
    AGGTTCCACG ACGTGGACCT TCCGCTCCAC TTGTTCTAGT TCAGGCGGGA CGACAGGTGG TTGTTCCGGC ACCACAGGGA CTCGTTGCCG CACTCGCACG
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    .  T  S  K  V  L  D  L  K  N  Y  I  D  K  Q  L  L  P  I  V  N  K  Q  S  C  S  I  S  N  I  E  T  V  I  .
1001 TGACCAGCAA GGTGCTGGAT CTGAAGAACT ACATCGACAA GCAGCTGCTG CCCATCGTGA ACAAGCAGAG CTGCAGCATC AGCAACATCG AGACCGTGAT
    ACTGGTCGTT CCACGACCTA GACTTCTTGA TGTAGCTGTT CGTCGACGAC GGGTAGCACT TGTTCGTCTC GACGTCGTAG TCGTTGTAGC TCTGGCACTA
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    .  E  F  Q  Q  K  N  N  R  L  L  E  I  T  R  E  F  S  V  N  A  G  V  T  T  P  V  S  T  Y  M  L  T  N
1101 CGAGTTCCAG CAGAAGAACA ACCGGCTGCT GGAAATCACC CGGGAGTTCA GCGTGAACGC CGGCGTGACC ACCCCCGTGA GCACCTACAT GCTGACCAAC
    GCTCAAGGTC GTCTTCTTGT TGGCCGACGA CCTTTAGTGG GCCCTCAAGT CGCACTTGCG GCCGCACTGG TGGGGGCACT CGTGGATGTA CGACTGGTTG
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     S  E  L  L  S  L  I  N  D  M  P  I  T  N  D  Q  K  K  L  M  S  N  N  V  Q  I  V  R  Q  Q  S  Y  S  I
    .
1201 AGCGAGCTGC TGTCCCTGAT CAATGACATG CCCATCACCA ACGACCAGAA GAAACTGATG AGCAACAACG TGCAGATCGT GCGGCAGCAG AGCTACTCCA
    TCGCTCGACG ACAGGGACTA GTTACTGTAC GGGTAGTGGT TGCTGGTCTT CTTTGACTAC TCGTTGTTGC ACGTCTAGCA CGCCGTCGTC TCGATGAGGT
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```


M S I I K E E V L A Y V V Q L P L Y G V I D T P C W K L H T S P L
1301 TCATGAGCAT CATCAAAGAA GAGGTGCTGG CCTACGTGGT GCAGCTGCCC CTGTACGGCG TGATCGACAC CCCCTGCTGG AAGCTGCACA CCAGCCCCCT
AGTACTCGTA GTAGTTTCTT CTCCACGACC GGATGCACCA CGTCGACGGG GACATGCCGC ACTAGCTGTG GGGGACGACC TTCGACGTGT GGTCGGGGGA
Anchorless RSV F
C T T N T K E G S N I C L T R T D R G W Y C N N A G S V S F F P L
1401 GTGCACCACC AACACCAAAG AGGGCAGCAA CATCTGCCTG ACCCGGACCG ACCGGGGCTG GTACTGCAAC AACGCCGGCA GCGTGAGCTT CTTCCCCCTG
CACGTGGTGG TTGTGGTTTC TCCCGTCGTT GTAGACGGAC TGGGCCTGGC TGGCCCCGAC CATGACGTTG TTGCGGCCGT CGCACTCGAA GAAGGGGGAC
Anchorless RSV F
A D T C K V Q S N R V F C D T M N S L T L P S E V N L C N I D I F N
1501 GCCGACACCT GCAAGGTGCA GAGCAACCGG GTGTTCTGCG ACACCATGAA CAGCCTGACC CTGCCCTCCG AGGTGAACCT GTGCAACATC GACATCTTCA
CGGCTGTGGA CGTTCCACGT CTCGTTGGCC CACAAGACGC TGTGGTACTT GTCGGACTGG GACGGGAGGC TCCACTTGGA CACGTTGTAG CTGTAGAAGT
Anchorless RSV F
P K Y D C K I M T S K T D V S S S V I T S L G A I V S C Y G K T K
1601 ACCCCAAGTA CGACTGCAAG ATCATGACCT CCAAGACCGA CGTGAGCAGC TCCGTGATCA CCTCCCTGGG CGCCATCGTG AGCTGCTACG GCAAGACCAA
TGGGGTTCAT GCTGACGTTC TAGTACTGGA GGTTCTGGCT GCACTCGTCG AGGCACTAGT GGAGGGACCC GCGGTAGCAC TCGACGATGC CGTTCTGGTT
Anchorless RSV F
C T A S N K N R G I I K T F S N G C D Y V S N K G V D T V S V G N
1701 GTGCACCGCC AGCAACAAGA ACCGGGGCAT CATCAAGACC TTCAGCAACG GCTGCGACTA CGTGAGCAAC AAGGGCGTGG ACACCGTGAG CGTGGGCAAC
CACGTGGCGG TCGTTGTTCT TGGCCCCGTA GTAGTTCTGG AAGTCGTTGC CGACGCTGAT GCACTCGTTG TTCCCGCACC TGTGGCACTC GCACCCGTTG
Anchorless RSV F
T L Y Y V N K Q E G K S L Y V K G E P I I N F Y D P L V F P S D E F
1801 ACACTGTACT ACGTGAATAA GCAGGAAGGC AAGAGCCTGT ACGTGAAGGG CGAGCCTATC ATCAACTTCT ACGACCCCCT GGTGTTCCCC AGCGACGAGT
TGTGACATGA TGCACTTATT CGTCCTTCCG TTCTCGGACA TGCACTTCCC GCTCGGATAG TAGTTGAAGA TGCTGGGGGA CCACAAGGGG TCGCTGCTCA
Anchorless RSV F
D A S I S Q V N E K I N Q S L A F I R K S D E L L H N V N A G K S
1901 TCGACGCCAG CATCAGCCAG GTGAACGAGA AGATCAACCA GAGCCTGGCC TTCATCCGGA AGAGCGACGA GCTGCTGCAC AATGTGAATG CCGGCAAGAG
AGCTGCGGTC GTAGTCGGTC CACTTGCTCT TCTAGTTGGT CTCGGACCGG AAGTAGGCCT TCTCGCTGCT CGACGACGTG TTACACTTAC GGCCGTTCTC
pre E/NS1 signal
FMDV 2A
Anchorless RSV F
E (split)
Transmembrane domain of WNV

```
     · T  T  N  I  M  N  F  D  L  L  K  L  A  G  D  V  E  S  N  P  G  P  A  R  D  R  S  I  A  L  T  F  L
2001 CACCACCAAT ATCATGAATT TTGATCTGCT CAAACTTGCA GGCGATGTAG AATCAAATCC TGGACCCGCC CGGGACAGGT CCATAGCTCT CACGTTTCTC
     GTGGTGGTTA TAGTACTTAA AACTAGACGA GTTTGAACGT CCGCTACATC TTAGTTTAGG ACCTGGGCGG GCCCTGTCCA GGTATCGAGA GTGCAAAGAG
                                                                                 NS1
                                                  ~~~
```

```
     TGCCGACAGT TCTTGTTGAA CCGCTAGGTG TCACTGGACA GGATAACCTA ACTTTCGTCC GAGTTACTAT GCACCTTCGA ACTTTCCCGT CAAGACCCAC
                                                 NS1
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · V K S  C T W  P E T H  T L W  G D G  I L E S  D L I  I P V  T L A G  P R S ·
2801 AAGTCAAATC ATGTACGTGG CCTGAGACGC ATACCTTGTG GGGCGATGGA ATCCTTGAGA GTGACTTGAT AATACCAGTC ACACTGGCGG GACCACGAAG
     TTCAGTTTAG TACATGCACC GGACTCTGCG TATGGAACAC CCCGCTACCT TAGGAACTCT CACTGAACTA TTATGGTCAG TGTGACCGCC CTGGTGCTTC
                                                 NS1
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · N H N  R R P G  Y K T  Q N Q  G P W D  E G R  V E I  D F D Y  C P G  T T V
2901 CAATCACAAT CGGAGACCTG GGTATAAGAC ACAAAACCAG GGCCCATGGG ACGAAGGCCG GGTAGAGATT GACTTCGATT ACTGCCCAGG AACTACGGTC
     GTTAGTGTTA GCCTCTGGAC CCATATTCTG TGTTTTGGTC CCGGGTACCC TGCTTCCGGC CCATCTCTAA CTGAAGCTAA TGACGGGTCC TTGATGCCAG
                                                 NS1
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      T L S E  S C G  H R G  P A T R  T T T  E S G  K L I T  D W C  C R S  C T L P
     .
3001 ACCCTGAGTG AGAGCTGCGG ACACCGTGGA CCTGCCACTC GCACCACCAC AGAGAGCGGA AAGTTGATAA CAGATTGGTG CTGCAGGAGC TGCACCTTAC
     TGGGACTCAC TCTCGACGCC TGTGGCACCT GGACGGTGAG CGTGGTGGTG TCTCTCGCCT TTCAACTATT GTCTAACCAC GACGTCCTCG ACGTGGAATG
                                                 NS1
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · P L R  Y Q T  D S G  C W Y G  M E I  R P Q R  H D E  K T L  V Q S Q  V N A ·
3101 CACCACTGCG CTACCAAACT GACAGCGGCT GTTGGTATGG TATGGAGATC AGACCACAGA GACATGATGA AAAGACCCTC GTGCAGTCAC AAGTGAATGC
     GTGGTGACGC GATGGTTTGA CTGTCGCCGA CAACCATACC ATACCTCTAG TCTGGTGTCT CTGTACTACT TTTCTGGGAG CACGTCAGTG TTCACTTACG
     NS1
     ~
                                                 NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · Y N A  D M I D  P F Q  L G L  L V V F  L A T  Q E V  L R K R  W T A  K I S
3201 TTATAATGCT GATATGATTG ACCCTTTTCA GTTGGGCCTT CTGGTCGTGT TCTTGGCCAC CCAGGAGGTC CTTCGCAAGA GGTGGACAGC CAAGATCAGC
     AATATTACGA CTATACTAAC TGGGAAAAGT CAACCCGGAA GACCAGCACA AGAACCGGTG GGTCCTCCAG GAAGCGTTCT CCACCTGTCG GTTCTAGTCG
                                                 NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      M P A I  L I A  L L V  L V F G  G I T  Y T D  V L R Y  V I L  V G A  A F A E
     .
3301 ATGCCAGCTA TACTGATTGC TCTGCTAGTC CTGGTGTTTG GGGGCATTAC TTACACTGAT GTGTTACGCT ATGTCATCTT GGTGGGGGCA GCTTTCGCAG
     TACGGTCGAT ATGACTAACG AGACGATCAG GACCACAAAC CCCCGTAATG AATGTGACTA CACAATGCGA TACAGTAGAA CCACCCCCGT CGAAAGCGTC
                                                 NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · S N S  G G D  V V H L  A L M  A T F  K I Q P  V F M  V A S  F L K A  R W T ·
3401 AATCTAATTC GGGAGGAGAC GTGGTACACT TGGCGCTCAT GGCGACCTTC AAGATACAAC CAGTGTTTAT GGTGGCATCG TTTCTTAAAG CGAGATGGAC
     TTAGATTAAG CCCTCCTCTG CACCATGTGA ACCGCGAGTA CCGCTGGAAG TTCTATGTTG GTCACAAATA CCACCGTAGC AAAGAATTTC GCTCTACCTG
                                                 NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

CA 02755255 2011-09-12

```
    · N  Q  E  N  I  L  L  M  L  A  A  V  F  F  Q  M  A  Y  H  D  A  R  Q  I  L  L  W  E  I  P  D  V  L
3501 CAACCAGGAG AACATTTTGT TGATGTTGGC GGCTGTTTTC TTTCAAATGG CTTATCACGA TGCCCGCCAA ATTCTGCTCT GGGAGATCCC TGATGTGTTG
     GTTGGTCCTC TTGTAAAACA ACTACAACCG CCGACAAAAG AAAGTTTACC GAATAGTGCT ACGGGCGGTT TAAGACGAGA CCCTCTAGGG ACTACACAAC
                                                 NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      N  S  L  A  I  A  W  M  I  L  R  A  I  T  F  T  T  T  S  N  V  V  V  P  L  L  A  L  L  T  P  G  L  R
     -
3601 AATTCACTGG CAATAGCTTG GATGATACTG AGAGCCATAA CATTCACAAC GACATCAAAC GTGGTTGTTC CGCTGCTAGC CCTGCTAACA CCCGGGCTGA
     TTAAGTGACC GTTATCGAAC CTACTATGAC TCTCGGTATT GTAAGTGTTG CTGTAGTTTG CACCAACAAG GCGACGATCG GGACGATTGT GGGCCCGACT
                                                 NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · C  L  N  L  D  V  Y  R  I  L  L  L  M  V  G  I  G  S  L  I  R  E  K  R  S  A  A  A  K  K  K  G  A ·
3701 GATGCTTGAA TCTGGATGTG TACAGGATAC TGCTGTTGAT GGTCGGAATA GGCAGCTTGA TCAGGGAGAA GAGGAGCGCA GCTGCAAAAA AGAAAGGAGC
     CTACGAACTT AGACCTACAC ATGTCCTATG ACGACAACTA CCAGCCTTAT CCGTCGAACT AGTCCCTCTT CTCCTCGCGT CGACGTTTTT TCTTTCCTCG
                                                                                                 NS2B
                                                                                                 ~~~~~~
                                                 NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · S  L  L  C  L  A  L  A  S  T  G  L  F  N  P  M  I  L  A  A  G  L  I  A  C  D  P  N  R  K  R  G  W
3801 AAGTCTGCTA TGCTTGGCTC TAGCCTCAAC AGGACTCTTC AACCCCATGA TCCTTGCTGC TGGACTGATT GCATGTGATC CCAACCGTAA ACGCGGGTGG
     TTCAGACGAT ACGAACCGAG ATCGGAGTTG TCCTGAGAAG TTGGGGTACT AGGAACGACG ACCTGACTAA CGTACACTAG GGTTGGCATT TGCGCCCACC
                                                 NS2B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      P  A  T  E  V  M  T  A  V  G  L  M  F  A  I  V  G  G  L  A  E  L  D  I  D  S  M  A  I  P  M  T  I  A
     .
3901 CCCGCAACTG AAGTGATGAC AGCTGTCGGC CTAATGTTTG CCATCGTCGG AGGGCTGGCA GAGCTTGACA TTGACTCCAT GGCCATTCCA ATGACTATCG
     GGGCGTTGAC TTCACTACTG TCGACAGCCG GATTACAAAC GGTAGCAGCC TCCCGACCGT CTCGAACTGT AACTGAGGTA CCGGTAAGGT TACTGATAGC
                                                 NS2B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · G  L  M  F  A  A  F  V  I  S  G  K  S  T  D  M  W  I  E  R  T  A  D  I  S  W  E  S  D  A  E  I  T ·
4001 CGGGGCTCAT GTTTGCTGCT TTCGTGATTT CTGGGAAATC AACAGATATG TGGATTGAGA GAACGGCGGA CATTTCCTGG GAAAGTGATG CAGAGATTAC
     GCCCCGAGTA CAAACGACGA AAGCACTAAA GACCCTTTAG TTGTCTATAC ACCTAACTCT CTTGCCGCCT GTAAAGGACC CTTTCACTAC GTCTCTAATG
                                                 NS2B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · G  S  S  E  R  V  D  V  R  L  D  D  D  G  N  F  Q  L  M  N  D  P  G  A  P  W  K  I  W  M  L  R  M
4101 AGGCTCGAGC GAAAGAGTTG ATGTGCGGCT TGATGATGAT GGAAACTTCC AGCTCATGAA TGATCCAGGA GCACCTTGGA AGATATGGAT GCTCAGAATG
     TCCGAGCTCG CTTTCTCAAC TACACGCCGA ACTACTACTA CCTTTGAAGG TCGAGTACTT ACTAGGTCCT CGTGGAACCT TCTATACCTA CGAGTCTTAC
                                                                                                 NS3
                                                                                                 ~~~~~~~~~~~~~~~
                                                 NS2B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

```
     V  C  L  A  I  S  A  Y  T  P  W  A  I  L  P  S  V  V  G  F  W  I  T  L  Q  Y  T  K  R  G  G  V  L  W
     .
4201 GTCTGTCTCG CGATTAGTGC GTACACCCCC TGGGCAATCT TGCCCTCAGT AGTTGGATTT TGGATAACTC TCCAATACAC AAAGAGAGGA GGCGTGTTGT
     CAGACAGAGC GCTAATCACG CATGTGGGGG ACCCGTTAGA ACGGGAGTCA TCAACCTAAA ACCTATTGAG AGGTTATGTG TTTCTCTCCT CCGCACAACA
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     ·  D  T  P  S  P  K  E  Y  K  K  G  D  T  T  T  G  V  Y  R  I  M  T  R  G  L  L  G  S  Y  Q  A  G  A  ·
4301 GGGACACTCC CTCACCAAAG GAGTACAAAA AGGGGGACAC GACCACCGGC GTCTACAGGA TCATGACTCG TGGGCTGCTC GGCAGTTATC AAGCAGGAGC
     CCCTGTGAGG GAGTGGTTTC CTCATGTTTT TCCCCCTGTG CTGGTGGCCG CAGATGTCCT AGTACTGAGC ACCCGACGAG CCGTCAATAG TTCGTCCTCG
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     ·  G  V  M  V  E  G  V  F  H  T  L  W  H  T  T  K  G  A  A  L  M  S  G  E  G  R  L  D  P  Y  W  G  S
4401 AGGCGTGATG GTTGAAGGTG TTTTCCACAC CCTTTGGCAT ACAACAAAAG GAGCCGCTTT GATGAGCGGA GAGGGCCGCC TGGACCCATA CTGGGGCAGT
     TCCGCACTAC CAACTTCCAC AAAAGGTGTG GGAAACCGTA TGTTGTTTTC CTCGGCGAAA CTACTCGCCT CTCCCGGCGG ACCTGGGTAT GACCCCGTCA
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     V  K  E  D  R  L  C  Y  G  G  P  W  K  L  Q  H  K  W  N  G  Q  D  E  V  Q  M  I  V  V  E  P  G  K  N
     .
4501 GTCAAGGAGG ATCGACTTTG TTACGGAGGA CCCTGGAAAT TGCAGCACAA GTGGAACGGG CAGGATGAGG TGCAGATGAT TGTGGTGGAA CCTGGCAAGA
     CAGTTCCTCC TAGCTGAAAC AATGCCTCCT GGGACCTTTA ACGTCGTGTT CACCTTGCCC GTCCTACTCC ACGTCTACTA ACACCACCTT GGACCGTTCT
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     ·  V  K  N  V  Q  T  K  P  G  V  F  K  T  P  E  G  E  I  G  A  V  T  L  D  F  P  T  G  T  S  G  S  P  ·
4601 ACGTTAAGAA CGTCCAGACG AAACCAGGGG TGTTCAAAAC ACCTGAAGGA GAAATCGGGG CCGTGACTTT GGACTTCCCC ACTGGAACAT CAGGCTCACC
     TGCAATTCTT GCAGGTCTGC TTTGGTCCCC ACAAGTTTTG TGGACTTCCT CTTTAGCCCC GGCACTGAAA CCTGAAGGGG TGACCTTGTA GTCCGAGTGG
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     ·  I  V  D  K  N  G  D  V  I  G  L  Y  G  N  G  V  I  M  P  N  G  S  Y  I  S  A  I  V  Q  G  E  R  M
4701 AATAGTGGAC AAAAACGGTG ATGTGATTGG GCTTTATGGC AATGGAGTCA TAATGCCCAA CGGCTCATAC ATAAGCGCGA TAGTGCAGGG TGAAAGGATG
     TTATCACCTG TTTTTGCCAC TACACTAACC CGAAATACCG TTACCTCAGT ATTACGGGTT GCCGAGTATG TATTCGCGCT ATCACGTCCC ACTTTCCTAC
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     D  E  P  I  P  A  G  F  E  P  E  M  L  R  K  K  Q  I  T  V  L  D  L  H  P  G  A  G  K  T  R  R  I  L
     .
4801 GATGAGCCAA TCCCAGCCGG ATTCGAACCT GAGATGCTGA GGAAAAAACA GATCACTGTA CTGGATCTCC ATCCCGGCGC CGGTAAAACA AGGAGGATTC
     CTACTCGGTT AGGGTCGGCC TAAGCTTGGA CTCTACGACT CCTTTTTTGT CTAGTGACAT GACCTAGAGG TAGGGCCGCG GCCATTTTGT TCCTCCTAAG
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     ·  P  Q  I  I  K  E  A  I  N  R  R  L  R  T  A  V  L  A  P  T  R  V  V  A  A  E  M  A  E  A  L  R  G  ·
4901 TGCCACAGAT CATCAAAGAG GCCATAAACA GAAGACTGAG AACAGCCGTG CTAGCACCAA CCAGGGTTGT GGCTGCTGAG ATGGCTGAAG CACTGAGAGG
     ACGGTGTCTA GTAGTTTCTC CGGTATTTGT CTTCTGACTC TTGTCGGCAC GATCGTGGTT GGTCCCAACA CCGACGACTC TACCGACTTC GTGACTCTCC
                                                 NS3
```

```
                                                                                                                    
     · L  P  I  R  Y  Q  T  S  A  V  P  R  E  H  N  G  N  E  I  V  D  V  M  C  H  A  T  L  T  H  R  L  M
5001 ACTGCCCATC CGGTACCAGA CATCCGCAGT GCCCAGAGAA CATAATGGAA ATGAGATTGT TGATGTCATG TGTCATGCTA CCCTCACCCA CAGGCTGATG
     TGACGGGTAG GCCATGGTCT GTAGGCGTCA CGGGTCTCTT GTATTACCTT TACTCTAACA ACTACAGTAC ACAGTACGAT GGGAGTGGGT GTCCGACTAC
                                                 NS3
                                                                                                                    
      S  P  H  R  V  P  N  Y  N  L  F  V  M  D  E  A  H  F  T  D  P  A  S  I  A  A  R  G  Y  I  S  T  K  V
     .
5101 TCTCCTCACA GGGTGCCGAA CTACAACCTG TTCGTGATGG ATGAGGCTCA TTTCACCGAC CCAGCTAGCA TTGCAGCAAG AGGTTACATT TCCACAAAGG
     AGAGGAGTGT CCCACGGCTT GATGTTGGAC AAGCACTACC TACTCCGAGT AAAGTGGCTG GGTCGATCGT AACGTCGTTC TCCAATGTAA AGGTGTTTCC
                                                 NS3
                                                                                                                    
     · E  L  G  E  A  A  A  I  F  M  T  A  T  P  P  G  T  S  D  P  F  P  E  S  N  S  P  I  S  D  L  Q  T ·
5201 TCGAGCTAGG GGAGGCGGCG GCAATATTCA TGACAGCCAC CCCACCAGGC ACTTCAGATC CATTCCCAGA GTCCAATTCA CCAATTTCCG ACTTACAGAC
     AGCTCGATCC CCTCCGCCGC CGTTATAAGT ACTGTCGGTG GGGTGGTCCG TGAAGTCTAG GTAAGGGTCT CAGGTTAAGT GGTTAAAGGC TGAATGTCTG
                                                 NS3
                                                                                                                    
     · E  I  P  D  R  A  W  N  S  G  Y  E  W  I  T  E  Y  T  G  K  T  V  W  F  V  P  S  V  K  M  G  N  E
5301 TGAGATCCCG GATCGAGCTT GGAACTCTGG ATACGAATGG ATCACAGAAT ACACCGGGAA GACGGTTTGG TTTGTGCCTA GTGTTAAGAT GGGGAATGAG
     ACTCTAGGGC CTAGCTCGAA CCTTGAGACC TATGCTTACC TAGTGTCTTA TGTGGCCCTT CTGCCAAACC AAACACGGAT CACAATTCTA CCCCTTACTC
                                                 NS3
                                                                                                                    
      I  A  L  C  L  Q  R  A  G  K  K  V  V  Q  L  N  R  K  S  Y  E  T  E  Y  P  K  C  K  N  D  D  W  D  F
     .
5401 ATTGCCCTTT GCCTACAACG TGCTGGAAAG AAAGTAGTCC AATTGAACAG AAAGTCGTAC GAGACGGAGT ACCCAAAATG TAAGAACGAT GATTGGGACT
     TAACGGGAAA CGGATGTTGC ACGACCTTTC TTTCATCAGG TTAACTTGTC TTTCAGCATG CTCTGCCTCA TGGGTTTTAC ATTCTTGCTA CTAACCCTGA
                                                 NS3
                                                                                                                    
     · V  I  T  T  D  I  S  E  M  G  A  N  F  K  A  S  R  V  I  D  S  R  K  S  V  K  P  T  I  I  T  E  G ·
5501 TTGTTATCAC AACAGACATA TCTGAAATGG GGGCTAACTT CAAGGCGAGC AGGGTGATTG ACAGCCGGAA GAGTGTGAAA CCAACCATCA TAACAGAAGG
     AACAATAGTG TTGTCTGTAT AGACTTTACC CCCGATTGAA GTTCCGCTCG TCCCACTAAC TGTCGGCCTT CTCACACTTT GGTTGGTAGT ATTGTCTTCC
                                                 NS3
                                                                                                                    
     · E  G  R  V  I  L  G  E  P  S  A  V  T  A  A  S  A  A  Q  R  R  G  R  I  G  R  N  P  S  Q  V  G  D
5601 AGAAGGGAGA GTGATCCTGG GAGAACCATC TGCAGTGACA GCAGCTAGTG CCGCCCAGAG ACGTGGACGT ATCGGTAGAA ATCCGTCGCA AGTTGGTGAT
     TCTTCCCTCT CACTAGGACC CTCTTGGTAG ACGTCACTGT CGTCGATCAC GGCGGGTCTC TGCACCTGCA TAGCCATCTT TAGGCAGCGT TCAACCACTA
                                                 NS3
                                                                                                                    
     E  Y  C  Y  G  G  H  T  N  E  D  D  S  N  F  A  H  W  T  E  A  R  I  M  L  D  N  I  N  M  P  N  G  L
     .
5701 GAGTACTGTT ATGGGGGGCA CACGAATGAA GACGACTCGA ACTTCGCCCA TTGGACTGAG GCACGAATCA TGCTGGACAA CATCAACATG CCAAACGGAC
     CTCATGACAA TACCCCCCGT GTGCTTACTT CTGCTGAGCT TGAAGCGGGT AACCTGACTC CGTGCTTAGT ACGACCTGTT GTAGTTGTAC GGTTTGCCTG
```

```
                                                NS3
     ~~~
```

```
     ~~~
```

NS5

NS4B

```
                                                
```

```
8001 CACGCACGAG ATGTATTGGG TGAGTCGAGC TTCAGGCAAT GTGGTACATT CAGTGAATAT GACCAGCCAG GTGCTCCTAG GAAGAATGGA AAAAAGGACC
     GTGCGTGCTC TACATAACCC ACTCAGCTCG AAGTCCGTTA CACCATGTAA GTCACTTATA CTGGTCGGTC CACGAGGATC CTTCTTACCT TTTTTCCTGG
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      W  K  G  P   Q  Y  E   E  D  V   N  L  G  S   G  T  R   A  V  G   K  P  L  L   N  S  D   T  S  K   I  K  N  R
     .
8101 TGGAAGGGAC CCCAATACGA GGAAGACGTA AACTTGGGAA GTGGAACCAG GGCGGTGGGA AAACCCCTGC TCAACTCAGA CACCAGTAAA ATCAAGAACA
     ACCTTCCCTG GGGTTATGCT CCTTCTGCAT TTGAACCCTT CACCTTGGTC CCGCCACCCT TTTGGGGACG AGTTGAGTCT GTGGTCATTT TAGTTCTTGT
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · I  E  R   L  R  R   E  Y  S   S  T  W  H   H  D  E   N  H  P   Y  R  T  W   N  Y  H   G  S  Y   D  V  K  P ·
8201 GGATTGAACG ACTCAGGCGT GAGTACAGTT CGACGTGGCA CCACGATGAG AACCACCCAT ATAGAACCTG GAACTATCAT GGCAGTTATG ATGTGAAGCC
     CCTAACTTGC TGAGTCCGCA CTCATGTCAA GCTGCACCGT GGTGCTACTC TTGGTGGGTA TATCTTGGAC CTTGATAGTA CCGTCAATAC TACACTTCGG
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · T  G  S   A  S  S   L  V  N   G  V  V  R   L  L  S   K  P  W   D  T  I  T   N  V  T   T  M  A   M  T  D  T
8301 CACAGGCTCC GCCAGTTCGC TGGTCAATGG AGTGGTCAGG CTCCTCTCAA AACCATGGGA CACCATCACG AATGTTACCA CCATGGCCAT GACTGACACT
     GTGTCCGAGG CGGTCAAGCG ACCAGTTACC TCACCAGTCC GAGGAGAGTT TTGGTACCCT GTGGTAGTGC TTACAATGGT GGTACCGGTA CTGACTGTGA
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      T  P  F  G   Q  Q  R   V  F  K   E  K  V  D   T  K  A   P  E  P   P  E  G  V   K  Y  V   L  N  E   T  T  N  W
     .
8401 ACTCCCTTCG GGCAGCAGCG AGTGTTCAAA GAGAAGGTGG ACACGAAAGC TCCTGAACCG CCAGAAGGAG TGAAGTACGT GCTCAACGAG ACCACCAACT
     TGAGGGAAGC CCGTCGTCGC TCACAAGTTT CTCTTCCACC TGTGCTTTCG AGGACTTGGC GGTCTTCCTC ACTTCATGCA CGAGTTGCTC TGGTGGTTGA
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · L  W  A   F  L  A   R  E  K  R   P  R  M   C  S  R   E  E  F  I   R  K  V   N  S  N   A  A  L  G   A  M  P ·
8501 GGTTGTGGGC GTTTTTGGCC AGAGAAAAAC GTCCCAGAAT GTGCTCTCGA GAGGAATTCA TAAGAAAGGT CAACAGCAAT GCAGCTTTGG GTGCCATGTT
     CCAACACCCG CAAAAACCGG TCTCTTTTTG CAGGGTCTTA CACGAGAGCT CTCCTTAAGT ATTCTTTCCA GTTGTCGTTA CGTCGAAACC CACGGTACAA
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · E  E  Q   N  Q  W   R  S  A   R  E  A  V   E  D  P   K  F  W   E  M  V  D   E  E  R   E  A  H   L  R  G  E
8601 TGAAGAGCAG AATCAATGGA GGAGCGCCAG AGAAGCAGTT GAAGATCCAA AATTTTGGGA AATGGTGGAT GAGGAGCGCG AGGCACATCT GCGGGGGGAA
     ACTTCTCGTC TTAGTTACCT CCTCGCGGTC TCTTCGTCAA CTTCTAGGTT TTAAAACCCT TTACCACCTA CTCCTCGCGC TCCGTGTAGA CGCCCCCCTT
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      C  H  T  C   I  Y  N   M  M  G   K  R  E  K   K  P  G   E  F  G   K  A  K  G   S  R  A   I  W  F   M  W  L  G
     .
8701 TGTCACACTT GCATTTACAA CATGATGGGA AAGAGAGAGA AAAAACCCGG AGAGTTCGGA AAGGCCAAGG GAAGCAGAGC CATTTGGTTC ATGTGGCTCG
     ACAGTGTGAA CGTAAATGTT GTACTACCCT TTCTCTCTCT TTTTTGGGCC TCTCAAGCCT TTCCGGTTCC CTTCGTCTCG GTAAACCAAG TACACCGAGC
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

```
     · A R F  L E F  E A L G  F L N  E D H  W L G R  K N S  G G G  V E G L  G L Q ·
8801 GAGCTCGCTT TCTGGAGTTC GAGGCTCTGG GTTTTCTCAA TGAAGACCAC TGGCTTGGAA GAAAGAACTC AGGAGGAGGT GTCGAGGGCT TGGGCCTCCA
     CTCGAGCGAA AGACCTCAAG CTCCGAGACC CAAAAGAGTT ACTTCTGGTG ACCGAACCTT CTTTCTTGAG TCCTCCTCCA CAGCTCCCGA ACCCGGAGGT
                                                 NS5
     ~~~
```

```
                                                
```

```
      CTCCTGACCC AATTGTTTCC GTTTGGTTGC GGGGTGCGCC GGGTTCGGGG CCATTACCAC AATTGGTCCC GCTTTCCTGA TCTCCAATCT CCTCTGGGGC
                                                    3' UTR
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
10501 CGGTTTAAAG TGCACGGCCC AGCCTGGCTG AAGCTGTAGG TCAGGGGAAG GACTAGAGGT TAGTGGAGAC CCCGTGCCAC AAAACACCAC AACAAAACAG
      GCCAAATTTC ACGTGCCGGG TCGGACCGAC TTCGACATCC AGTCCCCTTC CTGATCTCCA ATCACCTCTG GGGCACGGTG TTTTGTGGTG TTGTTTTGTC
                                                    3' UTR
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
10601 CAAATAGACA CCTGGGATAG ACTAGGAGAT CTTCTGCTCT GCACAACCAG CCACACGGCA CAGTGCGCCG ACAATGGTGG CTGGTGGTGC GAGAACACAG
      GTTTATCTGT GGACCCTATC TGATCCTCTA GAAGACGAGA CGTGTTGGTC GGTGTGCCGT GTCACGCGGC TGTTACCACC GACCACCACG CTCTTGTGTC
      3' UTR
      ~~~~~
10701 GATCT
      CTAGA
```

RepliVax WN - Anchorless F inserted in place of ΔC.
F insert starts at nucleotide position 229bp and ends at 1806bp.

```
                                                                5' UTR
      ~~~
```

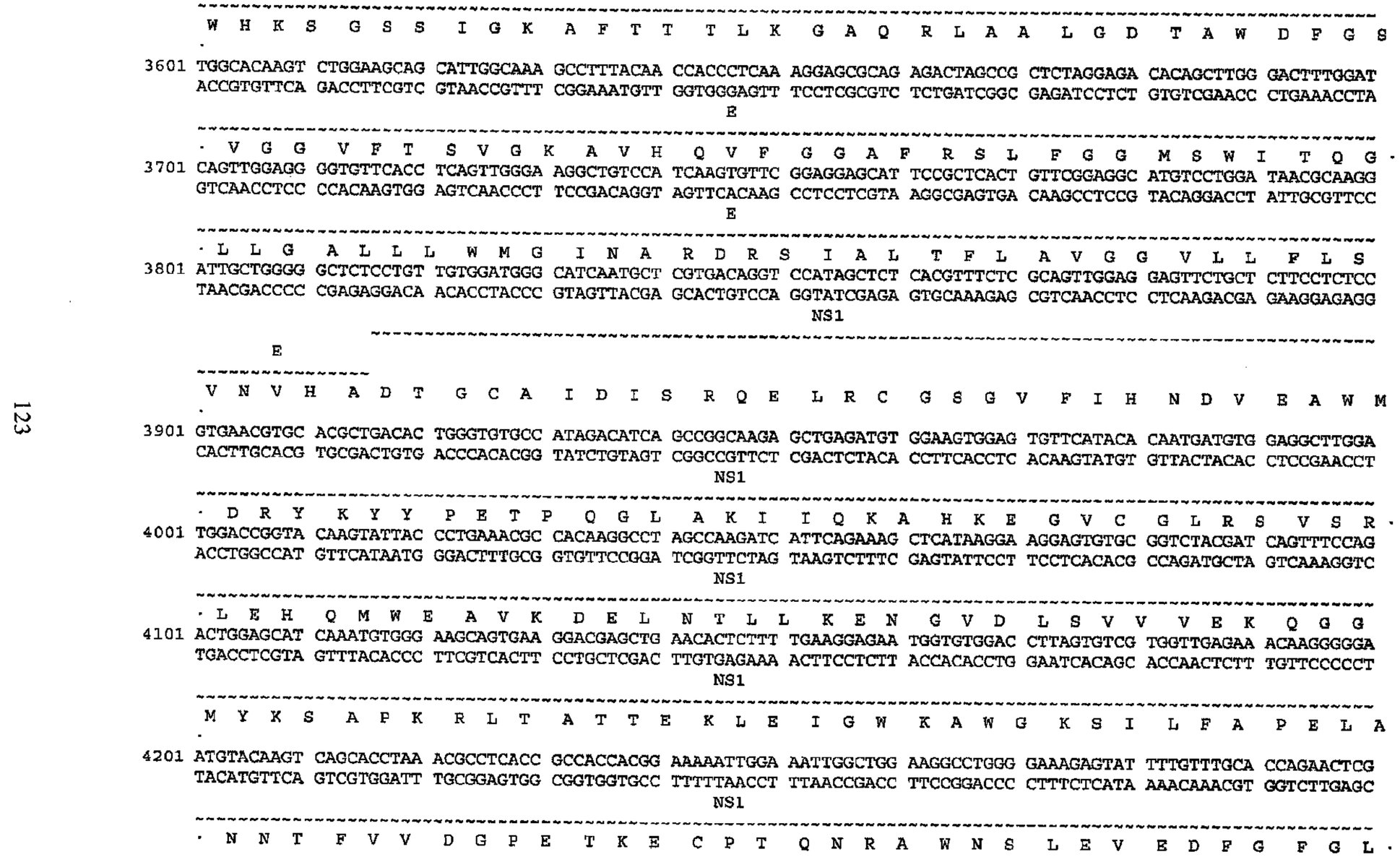
```
    GCTCGACGTC GACGACTACG TCTCGTGGGG ACGGCGGTTG TTGGCCCGGT CTGCGCTCGA CGGGGCCAAG TACTTGATGT GGGACTTGTT GCGGTTCTTT
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     T  N  V  T  L  S  K  K  R  K  R  R  F  L  G  F  L  L  G  V  G  S  A  I  A  S  G  I  A  V  S  K  V  L
    .
601 ACCAACGTGA CCCTGAGCAA GAAGCGGAAG CGGCGGTTCC TGGGCTTCCT GCTGGGCGTG GGCAGCGCCA TCGCCAGCGG CATCGCCGTG TCCAAGGTGC
    TGGTTGCACT GGGACTCGTT CTTCGCCTTC GCCGCCAAGG ACCCGAAGGA CGACCCGCAC CCGTCGCGGT AGCGGTCGCC GTAGCGGCAC AGGTTCCACG
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    · H  L  E  G  E  V  N  K  I  K  S  A  L  L  S  T  N  K  A  V  V  S  L  S  N  G  V  S  V  L  T  S  K ·
701 TGCACCTGGA AGGCGAGGTG AACAAGATCA AGTCCGCCCT GCTGTCCACC AACAAGGCCG TGGTGTCCCT GAGCAACGGC GTGAGCGTGC TGACCAGCAA
    ACGTGGACCT TCCGCTCCAC TTGTTCTAGT TCAGGCGGGA CGACAGGTGG TTGTTCCGGC ACCACAGGGA CTCGTTGCCG CACTCGCACG ACTGGTCGTT
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    · V  L  D  L  K  N  Y  I  D  K  Q  L  L  P  I  V  N  K  Q  S  C  S  I  S  N  I  E  T  V  I  E  F  Q
801 GGTGCTGGAT CTGAAGAACT ACATCGACAA GCAGCTGCTG CCCATCGTGA ACAAGCAGAG CTGCAGCATC AGCAACATCG AGACCGTGAT CGAGTTCCAG
    CCACGACCTA GACTTCTTGA TGTAGCTGTT CGTCGACGAC GGGTAGCACT TGTTCGTCTC GACGTCGTAG TCGTTGTAGC TCTGGCACTA GCTCAAGGTC
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     Q  K  N  N  R  L  L  E  I  T  R  E  F  S  V  N  A  G  V  T  T  P  V  S  T  Y  M  L  T  N  S  E  L  L
    .
901 CAGAAGAACA ACCGGCTGCT GGAAATCACC CGGGAGTTCA GCGTGAACGC CGGCGTGACC ACCCCCGTGA GCACCTACAT GCTGACCAAC AGCGAGCTGC
    GTCTTCTTGT TGGCCGACGA CCTTTAGTGG GCCCTCAAGT CGCACTTGCG GCCGCACTGG TGGGGGCACT CGTGGATGTA CGACTGGTTG TCGCTCGACG
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    · S  L  I  N  D  M  P  I  T  N  D  Q  K  K  L  M  S  N  N  V  Q  I  V  R  Q  Q  S  Y  S  I  M  S  I ·
1001 TGTCCCTGAT CAATGACATG CCCATCACCA ACGACCAGAA GAAACTGATG AGCAACAACG TGCAGATCGT GCGGCAGCAG AGCTACTCCA TCATGAGCAT
    ACAGGGACTA GTTACTGTAC GGGTAGTGGT TGCTGGTCTT CTTTGACTAC TCGTTGTTGC ACGTCTAGCA CGCCGTCGTC TCGATGAGGT AGTACTCGTA
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    · I  K  E  E  V  L  A  Y  V  V  Q  L  P  L  Y  G  V  I  D  T  P  C  W  K  L  H  T  S  P  L  C  T  T
1101 CATCAAAGAA GAGGTGCTGG CCTACGTGGT GCAGCTGCCC CTGTACGGCG TGATCGACAC CCCCTGCTGG AAGCTGCACA CCAGCCCCCT GTGCACCACC
    GTAGTTTCTT CTCCACGACC GGATGCACCA CGTCGACGGG GACATGCCGC ACTAGCTGTG GGGGACGACC TTCGACGTGT GGTCGGGGGA CACGTGGTGG
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     N  T  K  E  G  S  N  I  C  L  T  R  T  D  R  G  W  Y  C  N  N  A  G  S  V  S  F  F  P  L  A  D  T  C
    .
1201 AACACCAAAG AGGGCAGCAA CATCTGCCTG ACCCGGACCG ACCGGGGCTG GTACTGCAAC AACGCCGGCA GCGTGAGCTT CTTCCCCCTG GCCGACACCT
    TTGTGGTTTC TCCCGTCGTT GTAGACGGAC TGGGCCTGGC TGGCCCCGAC CATGACGTTG TTGCGGCCGT CGCACTCGAA GAAGGGGGAC CGGCTGTGGA
                                          Anchorless RSV F
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    · K  V  Q  S  N  R  V  F  C  D  T  M  N  S  L  T  L  P  S  E  V  N  L  C  N  I  D  I  F  N  P  K  Y ·
```

```
1301 GCAAGGTGCA GAGCAACCGG GTGTTCTGCG ACACCATGAA CAGCCTGACC CTGCCCTCCG AGGTGAACCT GTGCAACATC GACATCTTCA ACCCCAAGTA
     CGTTCCACGT CTCGTTGGCC CACAAGACGC TGTGGTACTT GTCGGACTGG GACGGGAGGC TCCACTTGGA CACGTTGTAG CTGTAGAAGT TGGGGTTCAT
                                          Anchorless RSV F
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · D  C  K  I  M  T  S  K  T  D  V  S  S  S  V  I  T  S  L  G  A  I  V  S  C  Y  G  K  T  K  C  T  A
1401 CGACTGCAAG ATCATGACCT CCAAGACCGA CGTGAGCAGC TCCGTGATCA CCTCCCTGGG CGCCATCGTG AGCTGCTACG GCAAGACCAA GTGCACCGCC
     GCTGACGTTC TAGTACTGGA GGTTCTGGCT GCACTCGTCG AGGCACTAGT GGAGGGACCC GCGGTAGCAC TCGACGATGC CGTTCTGGTT CACGTGGCGG
                                          Anchorless RSV F
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      S  N  K  N  R  G  I  I  K  T  F  S  N  G  C  D  Y  V  S  N  K  G  V  D  T  V  S  V  G  N  T  L  Y  Y
     .
1501 AGCAACAAGA ACCGGGGCAT CATCAAGACC TTCAGCAACG GCTGCGACTA CGTGAGCAAC AAGGGCGTGG ACACCGTGAG CGTGGGCAAC ACACTGTACT
     TCGTTGTTCT TGGCCCCGTA GTAGTTCTGG AAGTCGTTGC CGACGCTGAT GCACTCGTTG TTCCCGCACC TGTGGCACTC GCACCCGTTG TGTGACATGA
                                          Anchorless RSV F
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · V  N  K  Q  E  G  K  S  L  Y  V  K  G  E  P  I  I  N  F  Y  D  P  L  V  F  P  S  D  E  F  D  A  S ·
1601 ACGTGAATAA GCAGGAAGGC AAGAGCCTGT ACGTGAAGGG CGAGCCTATC ATCAACTTCT ACGACCCCCT GGTGTTCCCC AGCGACGAGT TCGACGCCAG
     TGCACTTATT CGTCCTTCCG TTCTCGGACA TGCACTTCCC GCTCGGATAG TAGTTGAAGA TGCTGGGGGA CCACAAGGGG TCGCTGCTCA AGCTGCGGTC
                                          Anchorless RSV F
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · I  S  Q  V  N  E  K  I  N  Q  S  L  A  F  I  R  K  S  D  E  L  L  H  N  V  N  A  G  K  S  T  T  N
1701 CATCAGCCAG GTGAACGAGA AGATCAACCA GAGCCTGGCC TTCATCCGGA AGAGCGACGA GCTGCTGCAC AATGTGAATG CCGGCAAGAG CACCACCAAT
     GTAGTCGGTC CACTTGCTCT TCTAGTTGGT CTCGGACCGG AAGTAGGCCT TCTCGCTGCT CGACGACGTG TTACACTTAC GGCCGTTCTC GTGGTGGTTA
                                     FMDV 2A
          ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     Anchorless RSV F                                                                C/prM signal
     ~~~~~~                                                                          ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      I  M  N  F  D  L  L  K  L  A  G  D  V  E  S  N  P  G  P  G  G  K  T  G  I  A  V  M  I  G  L  I  A  C
     .
1801 ATCATGAATT TTGATCTGCT CAAACTTGCA GGCGATGTAG AATCAAATCC TGGACCCGGA GGAAAGACCG GTATTGCAGT CATGATTGGC CTGATCGCCT
     TAGTACTTAA AACTAGACGA GTTTGAACGT CCGCTACATC TTAGTTTAGG ACCTGGGCCT CCTTTCTGGC CATAACGTCA GTACTAACCG GACTAGCGGA
                                                                     prM
               ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     C/prM signal
     ~~~~~~~~~~~~
     · V  G  A  V  T  L  S  N  F  Q  G  K  V  M  M  T  V  N  A  T  D  V  T  D  V  I  T  I  P  T  A  A  G ·
1901 GCGTAGGAGC AGTTACCCTC TCTAACTTCC AAGGGAAGGT GATGATGACG GTAAATGCTA CTGACGTCAC AGATGTCATC ACGATTCCAA CAGCTGCTGG
     CGCATCCTCG TCAATGGGAG AGATTGAAGG TTCCCTTCCA CTACTACTGC CATTTACGAT GACTGCAGTG TCTACAGTAG TGCTAAGGTT GTCGACGACC
                                                   prM
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · K  N  L  C  I  V  R  A  M  D  V  G  Y  M  C  D  D  T  I  T  Y  E  C  P  V  L  S  A  G  N  D  P  E
2001 AAAGAACCTA TGCATTGTCA GAGCAATGGA TGTGGGATAC ATGTGCGATG ATACTATCAC TTATGAATGC CCAGTGCTGT CGGCTGGTAA TGATCCAGAA
```

CA 02755255 2011-09-12

```
     TTTCTTGGAT ACGTAACAGT CTCGTTACCT ACACCCTATG TACACGCTAC TATGATAGTG AATACTTACG GGTCACGACA GCCGACCATT ACTAGGTCTT
                                                  prM
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      D I D C  W C T  K S A  V Y V R  Y G R  C T K  T R H S  R R S  R R S  L T V Q
     .
2101 GACATCGACT GTTGGTGCAC AAAGTCAGCA GTCTACGTCA GGTATGGAAG ATGCACCAAG ACACGCCACT CAAGACGCAG TCGGAGGTCA CTGACAGTGC
     CTGTAGCTGA CAACCACGTG TTTCAGTCGT CAGATGCAGT CCATACCTTC TACGTGGTTC TGTGCGGTGA GTTCTGCGTC AGCCTCCAGT GACTGTCACG
                                                  prM
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · T H G  E S T  L A N K  K G A  W M D  S T K A  T R Y  L V K  T E S W  I L R ·
2201 AGACACACGG AGAAAGCACT CTAGCGAACA AGAAGGGGGC TTGGATGGAC AGCACCAAGG CCACAAGGTA TTTGGTAAAA ACAGAATCAT GGATCTTGAG
     TCTGTGTGCC TCTTTCGTGA GATCGCTTGT TCTTCCCCCG AACCTACCTG TCGTGGTTCC GGTGTTCCAT AAACCATTTT TGTCTTAGTA CCTAGAACTC
                                                  prM
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · N P G  Y A L V  A A V  I G W  M L G S  N T M  Q R V  V F V V  L L L  L V A
2301 GAACCCTGGA TATGCCCTGG TGGCAGCCGT CATTGGTTGG ATGCTTGGGA GCAACACCAT GCAGAGAGTT GTGTTTGTCG TGCTATTGCT TTTGGTGGCC
     CTTGGGACCT ATACGGGACC ACCGTCGGCA GTAACCAACC TACGAACCCT CGTTGTGGTA CGTCTCTCAA CACAAACAGC ACGATAACGA AAACCACCGG
        prM
     ~~~~~~~~~~~~~
                                                  E
                    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      P A Y S  F N C  L G M  S N R D  F L E  G V S  G A T W  V D L  V L E  G D S C
     .
2401 CCAGCTTACA GCTTTAACTG CCTTGGAATG AGCAACAGAG ACTTCTTGGA AGGAGTGTCT GGAGCAACAT GGGTGGATTT GGTTCTCGAA GGCGACAGCT
     GGTCGAATGT CGAAATTGAC GGAACCTTAC TCGTTGTCTC TGAAGAACCT TCCTCACAGA CCTCGTTGTA CCCACCTAAA CCAAGAGCTT CCGCTGTCGA
                                                  E
                    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · V T I  M S K  D K P T  I D V  K M M  N M E A  A N L  A E V  R S Y C  Y L A ·
2501 GCGTGACTAT CATGTCTAAG GACAAGCCTA CCATCGATGT GAAGATGATG AATATGGAGG CGGCCAACCT GGCAGAGGTC CGCAGTTATT GCTATTTGGC
     CGCACTGATA GTACAGATTC CTGTTCGGAT GGTAGCTACA CTTCTACTAC TTATACCTCC GCCGGTTGGA CCGTCTCCAG GCGTCAATAA CGATAAACCG
                                                  E
                    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · T V S  D L S T  K A A  C P A  M G E A  H N D  K R A  D P A F  V C R  Q G V
2601 TACCGTCAGC GATCTCTCCA CCAAAGCTGC GTGCCCGGCC ATGGGAGAAG CTCACAATGA CAAACGTGCT GACCCAGCTT TTGTGTGCAG ACAAGGAGTG
     ATGGCAGTCG CTAGAGAGGT GGTTTCGACG CACGGGCCGG TACCCTCTTC GAGTGTTACT GTTTGCACGA CTGGGTCGAA AACACACGTC TGTTCCTCAC
                                                  E
                    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      V D R G  W G N  G C G  L F G K  G S I  D T C  A K F A  C S T  K A I  G R T I
     .
2701 GTGGACAGGG GCTGGGGCAA CGGCTGCGGA CTATTTGGCA AAGGAAGCAT TGACACATGC GCCAAATTTG CCTGCTCTAC CAAGGCAATA GGAAGAACCA
     CACCTGTCCC CGACCCCGTT GCCGACGCCT GATAAACCGT TTCCTTCGTA ACTGTGTACG CGGTTTAAAC GGACGAGATG GTTCCGTTAT CCTTCTTGGT
                                                  E
```

```
                                                                                                                    
    · L K E  N I K  Y E V A  I F V  H G P  T T V E  S H G  N Y S  T Q V G  A T Q ·
2801 TTTTGAAAGA GAATATCAAG TACGAAGTGG CCATTTTTGT CCATGGACCA ACTACTGTGG AGTCGCACGG AAACTACTCC ACACAGGTTG GAGCCACTCA
     AAAACTTTCT CTTATAGTTC ATGCTTCACC GGTAAAAACA GGTACCTGGT TGATGACACC TCAGCGTGCC TTTGATGAGG TGTGTCCAAC CTCGGTGAGT
                                                 E
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    · A G R  F S I T  P A A  P S Y  T L K L  G E Y  G E V  T V D C  E P R  S G I
2901 GGCAGGGAGA TTCAGCATCA CTCCTGCGGC GCCTTCATAC ACACTAAAGC TTGGAGAATA TGGAGAGGTG ACAGTGGACT GTGAACCACG GTCAGGGATT
     CCGTCCCTCT AAGTCGTAGT GAGGACGCCG CGGAAGTATG TGTGATTTCG AACCTCTTAT ACCTCTCCAC TGTCACCTGA CACTTGGTGC CAGTCCCTAA
                                                 E
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     D T N A  Y Y V  M T V  G T K T  F L V  H R E  W F M D  L N L  P W S  S A G S
    ·
3001 GACACCAATG CATACTACGT GATGACTGTT GGAACAAAGA CGTTCTTGGT CCATCGTGAG TGGTTCATGG ACCTCAACCT CCCTTGGAGC AGTGCTGGAA
     CTGTGGTTAC GTATGATGCA CTACTGACAA CCTTGTTTCT GCAAGAACCA GGTAGCACTC ACCAAGTACC TGGAGTTGGA GGGAACCTCG TCACGACCTT
                                                 E
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    · T V W  R N R  E T L M  E F E  E P H  A T K Q  S V I  A L G  S Q E G  A L H ·
3101 GTACTGTGTG GAGGAACAGA GAGACGTTAA TGGAGTTTGA GGAACCACAC GCCACGAAGC AGTCTGTGAT AGCATTGGGC TCACAAGAGG GAGCTCTGCA
     CATGACACAC CTCCTTGTCT CTCTGCAATT ACCTCAAACT CCTTGGTGTG CGGTGCTTCG TCAGACACTA TCGTAACCCG AGTGTTCTCC CTCGAGACGT
                                                 E
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    ·Q A L  A G A I  P V E  F S S  N T V K  L T S  G H L  K C R V  K M E  K L Q
3201 TCAAGCTTTG GCTGGAGCCA TTCCTGTGGA ATTTTCAAGC AACACTGTCA AGTTGACGTC GGGTCATTTG AAGTGTAGAG TGAAGATGGA AAAATTGCAG
     AGTTCGAAAC CGACCTCGGT AAGGACACCT TAAAAGTTCG TTGTGACAGT TCAACTGCAG CCCAGTAAAC TTCACATCTC ACTTCTACCT TTTTAACGTC
                                                 E
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     L K G T  T Y G  V C S  K A F K  F L G  T P A  D T G H  G T V  V L E  L Q Y T
    ·
3301 TTGAAGGGAA CAACCTATGG CGTCTGTTCA AAGGCTTTCA AGTTTCTTGG GACTCCCGCA GACACAGGTC ACGGCACTGT GGTGTTGGAA TTGCAGTACA
     AACTTCCCTT GTTGGATACC GCAGACAAGT TTCCGAAAGT TCAAAGAACC CTGAGGGCGT CTGTGTCCAG TGCCGTGACA CCACAACCTT AACGTCATGT
                                                 E
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    · G T D  G P C  K V P I  S S V  A S L  N D L T  P V G  R L V  T V N P  F V S ·
3401 CTGGCACGGA TGGACCTTGC AAAGTTCCTA TCTCGTCAGT GGCTTCATTG AACGACCTAA CGCCAGTGGG CAGATTGGTC ACTGTCAACC CTTTTGTTTC
     GACCGTGCCT ACCTGGAACG TTTCAAGGAT AGAGCAGTCA CCGAAGTAAC TTGCTGGATT GCGGTCACCC GTCTAACCAG TGACAGTTGG GAAAACAAAG
                                                 E
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    ·V A T  A N A K  V L I  E L E  P P F G  D S Y  I V V  G R G E  Q Q I  N H H
3501 AGTGGCCACG GCCAACGCTA AGGTCCTGAT TGAATTGGAA CCACCCTTTG GAGACTCATA CATAGTGGTG GGCAGAGGAG AACAACAGAT CAATCACCAC
     TCACCGGTGC CGGTTGCGAT TCCAGGACTA ACTTAACCTT GGTGGGAAAC CTCTGAGTAT GTATCACCAC CCGTCTCCTC TTGTTGTCTA GTTAGTGGTG
                                                 E
```


W H K S G S S I G K A F T T T L K G A Q R L A A L G D T A W D F G S
3601 TGGCACAAGT CTGGAAGCAG CATTGGCAAA GCCTTTACAA CCACCCTCAA AGGAGCGCAG AGACTAGCCG CTCTAGGAGA CACAGCTTGG GACTTTGGAT
ACCGTGTTCA GACCTTCGTC GTAACCGTTT CGGAAATGTT GGTGGGAGTT TCCTCGCGTC TCTGATCGGC GAGATCCTCT GTGTCGAACC CTGAAACCTA
E
· V G G V F T S V G K A V H Q V F G G A F R S L F G G M S W I T Q G ·
3701 CAGTTGGAGG GGTGTTCACC TCAGTTGGGA AGGCTGTCCA TCAAGTGTTC GGAGGAGCAT TCCGCTCACT GTTCGGAGGC ATGTCCTGGA TAACGCAAGG
GTCAACCTCC CCACAAGTGG AGTCAACCCT TCCGACAGGT AGTTCACAAG CCTCCTCGTA AGGCGAGTGA CAAGCCTCCG TACAGGACCT ATTGCGTTCC
E
· L L G A L L L W M G I N A R D R S I A L T F L A V G G V L L F L S
3801 ATTGCTGGGG GCTCTCCTGT TGTGGATGGG CATCAATGCT CGTGACAGGT CCATAGCTCT CACGTTTCTC GCAGTTGGAG GAGTTCTGCT CTTCCTCTCC
TAACGACCCC CGAGAGGACA ACACCTACCC GTAGTTACGA GCACTGTCCA GGTATCGAGA GTGCAAAGAG CGTCAACCTC CTCAAGACGA GAAGGAGAGG
NS1
E
V N V H A D T G C A I D I S R Q E L R C G S G V F I H N D V E A W M
3901 GTGAACGTGC ACGCTGACAC TGGGTGTGCC ATAGACATCA GCCGGCAAGA GCTGAGATGT GGAAGTGGAG TGTTCATACA CAATGATGTG GAGGCTTGGA
CACTTGCACG TGCGACTGTG ACCCACACGG TATCTGTAGT CGGCCGTTCT CGACTCTACA CCTTCACCTC ACAAGTATGT GTTACTACAC CTCCGAACCT
NS1
· D R Y K Y Y P E T P Q G L A K I I Q K A H K E G V C G L R S V S R ·
4001 TGGACCGGTA CAAGTATTAC CCTGAAACGC CACAAGGCCT AGCCAAGATC ATTCAGAAAG CTCATAAGGA AGGAGTGTGC GGTCTACGAT CAGTTTCCAG
ACCTGGCCAT GTTCATAATG GGACTTTGCG GTGTTCCGGA TCGGTTCTAG TAAGTCTTTC GAGTATTCCT TCCTCACACG CCAGATGCTA GTCAAAGGTC
NS1
· L E H Q M W E A V K D E L N T L L K E N G V D L S V V V E K Q G G
4101 ACTGGAGCAT CAAATGTGGG AAGCAGTGAA GGACGAGCTG AACACTCTTT TGAAGGAGAA TGGTGTGGAC CTTAGTGTCG TGGTTGAGAA ACAAGGGGGA
TGACCTCGTA GTTTACACCC TTCGTCACTT CCTGCTCGAC TTGTGAGAAA ACTTCCTCTT ACCACACCTG GAATCACAGC ACCAACTCTT TGTTCCCCCT
NS1
M Y K S A P K R L T A T T E K L E I G W K A W G K S I L F A P E L A
4201 ATGTACAAGT CAGCACCTAA ACGCCTCACC GCCACCACGG AAAAATTGGA AATTGGCTGG AAGGCCTGGG GAAAGAGTAT TTTGTTTGCA CCAGAACTCG
TACATGTTCA GTCGTGGATT TGCGGAGTGG CGGTGGTGCC TTTTTAACCT TTAACCGACC TTCCGGACCC CTTTCTCATA AAACAAACGT GGTCTTGAGC
NS1
· N N T F V V D G P E T K E C P T Q N R A W N S L E V E D F G F G L ·

```
4301 CCAACAACAC CTTTGTGGTT GATGGTCCGG AGACCAAGGA ATGTCCGACT CAGAATCGCG CTTGGAATAG CTTAGAAGTG GAGGATTTTG GATTTGGTCT
     GGTTGTTGTG GAAACACCAA CTACCAGGCC TCTGGTTCCT TACAGGCTGA GTCTTAGCGC GAACCTTATC GAATCTTCAC CTCCTAAAAC CTAAACCAGA
                                                 NS1
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · T  S  T   R  M  F  L   K  V  R   E  S  N   T  T  E  C   D  S  K   I  I  G   T  A  V  K   N  N  L   A  I  H
4401 CACCAGCACT CGGATGTTCC TGAAGGTCAG AGAGAGCAAC ACAACTGAAT GTGACTCGAA GATCATTGGA ACGGCTGTCA AGAACAACTT GGCGATCCAC
     GTGGTCGTGA GCCTACAAGG ACTTCCAGTC TCTCTCGTTG TGTTGACTTA CACTGAGCTT CTAGTAACCT TGCCGACAGT TCTTGTTGAA CCGCTAGGTG
                                                 NS1
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      S  D  L   S  Y  W  I   E  S  R   L  N  D   T  W  K  L   E  R  A   V  L  G   E  V  K   S  C  T  W   P  E  T  H
     ·
4501 AGTGACCTGT CCTATTGGAT TGAAAGCAGG CTCAATGATA CGTGGAAGCT TGAAAGGGCA GTTCTGGGTG AAGTCAAATC ATGTACGTGG CCTGAGACGC
     TCACTGGACA GGATAACCTA ACTTTCGTCC GAGTTACTAT GCACCTTCGA ACTTTCCCGT CAAGACCCAC TTCAGTTTAG TACATGCACC GGACTCTGCG
                                                 NS1
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · T  L  W   G  D  G   I  L  E   S  D  L   I  T  P  V   T  L  A   G  P  R   S  N  H   N  R  R   P  G  Y   K  T  ·
4601 ATACCTTGTG GGGCGATGGA ATCCTTGAGA GTGACTTGAT AATACCAGTC ACACTGGCGG GACCACGAAG CAATCACAAT CGGAGACCTG GGTATAAGAC
     TATGGAACAC CCCGCTACCT TAGGAACTCT CACTGAACTA TTATGGTCAG TGTGACCGCC CTGGTGCTTC GTTAGTGTTA GCCTCTGGAC CCATATTCTG
                                                 NS1
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · Q  N  Q   G  P  W  D   E  G  R   V  E  I   D  F  D  Y   C  P  G   T  T  V   T  L  S  E   S  C  G   H  R  G
4701 ACAAAACCAG GGCCCATGGG ACGAAGGCCG GGTAGAGATT GACTTCGATT ACTGCCCAGG AACTACGGTC ACCCTGAGTG AGAGCTGCGG ACACCGTGGA
     TGTTTTGGTC CCGGGTACCC TGCTTCCGGC CCATCTCTAA CTGAAGCTAA TGACGGGTCC TTGATGCCAG TGGGACTCAC TCTCGACGCC TGTGGCACCT
                                                 NS1
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      P  A  T   R  T  T  T   E  S  G   K  L  I   T  D  W  C   C  R  S   C  T  L   P  P  L   R  Y  Q   T  D  S  G   C
     ·
4801 CCTGCCACTC GCACCACCAC AGAGAGCGGA AAGTTGATAA CAGATTGGTG CTGCAGGAGC TGCACCTTAC CACCACTGCG CTACCAAACT GACAGCGGCT
     GGACGGTGAG CGTGGTGGTG TCTCTCGCCT TTCAACTATT GTCTAACCAC GACGTCCTCG ACGTGGAATG GTGGTGACGC GATGGTTTGA CTGTCGCCGA
                                                                                 NS2A
                                                                                 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                        NS1
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · W  Y  G   M  E  I   R  P  Q   R  H  D   E  K  T  L   V  Q  S   Q  V  N   A  Y  N   A  D  M   I  D  P   F  Q  ·
4901 GTTGGTATGG TATGGAGATC AGACCACAGA GACATGATGA AAAGACCCTC GTGCAGTCAC AAGTGAATGC TTATAATGCT GATATGATTG ACCCTTTTCA
     CAACCATACC ATACCTCTAG TCTGGTGTCT CTGTACTACT TTTCTGGGAG CACGTCAGTG TTCACTTACG AATATTACGA CTATACTAAC TGGGAAAAGT
                                                 NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · L  G  L   L  V  V  F   L  A  T   Q  E  V   L  R  K  R   W  T  A   K  I  S   M  P  A  I   L  I  A   L  L  V
5001 GTTGGGCCTT CTGGTCGTGT TCTTGGCCAC CCAGGAGGTC CTTCGCAAGA GGTGGACAGC CAAGATCAGC ATGCCAGCTA TACTGATTGC TCTGCTAGTC
     CAACCCGGAA GACCAGCACA AGAACCGGTG GGTCCTCCAG GAAGCGTTCT CCACCTGTCG GTTCTAGTCG TACGGTCGAT ATGACTAACG AGACGATCAG
                                                 NS2A
```

```
                                                                                                                    
     L  V  F  G  G  I  T  Y  T  D  V  L  R  Y  V  I  L  V  G  A  A  F  A  E  S  N  S  G  G  D  V  V  H  L
     .
5101 CTGGTGTTTG GGGGCATTAC TTACACTGAT GTGTTACGCT ATGTCATCTT GGTGGGGGCA GCTTTCGCAG AATCTAATTC GGGAGGAGAC GTGGTACACT
     GACCACAAAC CCCCGTAATG AATGTGACTA CACAATGCGA TACAGTAGAA CCACCCCCGT CGAAAGCGTC TTAGATTAAG CCCTCCTCTG CACCATGTGA
                                                 NS2A
     ----------------------------------------------------------------------------------------------------------
     .  A  L  M  A  T  F  K  I  Q  P  V  F  M  V  A  S  F  L  K  A  R  W  T  N  Q  E  N  I  L  L  M  L  A  .
5201 TGGCGCTCAT GGCGACCTTC AAGATACAAC CAGTGTTTAT GGTGGCATCG TTTCTTAAAG CGAGATGGAC CAACCAGGAG AACATTTTGT TGATGTTGGC
     ACCGCGAGTA CCGCTGGAAG TTCTATGTTG GTCACAAATA CCACCGTAGC AAAGAATTTC GCTCTACCTG GTTGGTCCTC TTGTAAAACA ACTACAACCG
                                                 NS2A
     ----------------------------------------------------------------------------------------------------------
     .  A  V  F  F  Q  M  A  Y  H  D  A  R  Q  I  L  L  W  E  I  P  D  V  L  N  S  L  A  I  A  W  M  I  L
5301 GGCTGTTTTC TTTCAAATGG CTTATCACGA TGCCCGCCAA ATTCTGCTCT GGGAGATCCC TGATGTGTTG AATTCACTGG CAATAGCTTG GATGATACTG
     CCGACAAAAG AAAGTTTACC GAATAGTGCT ACGGGCGGTT TAAGACGAGA CCCTCTAGGG ACTACACAAC TTAAGTGACC GTTATCGAAC CTACTATGAC
                                                 NS2A
     ----------------------------------------------------------------------------------------------------------
     R  A  I  T  F  T  T  T  S  N  V  V  V  P  L  L  A  L  L  T  P  G  L  R  C  L  N  L  D  V  Y  R  I  L
     .
5401 AGAGCCATAA CATTCACAAC GACATCAAAC GTGGTTGTTC CGCTGCTAGC CCTGCTAACA CCCGGGCTGA GATGCTTGAA TCTGGATGTG TACAGGATAC
     TCTCGGTATT GTAAGTGTTG CTGTAGTTTG CACCAACAAG GCGACGATCG GGACGATTGT GGGCCCGACT CTACGAACTT AGACCTACAC ATGTCCTATG
                                                 NS2A
     ----------------------------------------------------------------------------------------------------------
     .  L  L  M  V  G  I  G  S  L  I  R  E  K  R  S  A  A  A  K  K  K  G  A  S  L  L  C  L  A  L  A  S  T  .
5501 TGCTGTTGAT GGTCGGAATA GGCAGCTTGA TCAGGGAGAA GAGGAGCGCA GCTGCAAAAA AGAAAGGAGC AAGTCTGCTA TGCTTGGCTC TAGCCTCAAC
     ACGACAACTA CCAGCCTTAT CCGTCGAACT AGTCCCTCTT CTCCTCGCGT CGACGTTTTT TCTTTCCTCG TTCAGACGAT ACGAACCGAG ATCGGAGTTG
                                                                                    NS2B
                                                                                    ----------------------------
                                                 NS2A
                                                 ----------------------------------------
     .  G  L  F  N  P  M  I  L  A  A  G  L  I  A  C  D  P  N  R  K  R  G  W  P  A  T  E  V  M  T  A  V  G
5601 AGGACTCTTC AACCCCATGA TCCTTGCTGC TGGACTGATT GCATGTGATC CCAACCGTAA ACGCGGGTGG CCCGCAACTG AAGTGATGAC AGCTGTCGGC
     TCCTGAGAAG TTGGGGTACT AGGAACGACG ACCTGACTAA CGTACACTAG GGTTGGCATT TGCGCCCACC GGGCGTTGAC TTCACTACTG TCGACAGCCG
                                                 NS2B
     ----------------------------------------------------------------------------------------------------------
     L  M  F  A  I  V  G  G  L  A  E  L  D  I  D  S  M  A  I  P  M  T  I  A  G  L  M  F  A  A  F  V  I  S
     .
5701 CTAATGTTTG CCATCGTCGG AGGGCTGGCA GAGCTTGACA TTGACTCCAT GGCCATTCCA ATGACTATCG CGGGGCTCAT GTTTGCTGCT TTCGTGATTT
     GATTACAAAC GGTAGCAGCC TCCCGACCGT CTCGAACTGT AACTGAGGTA CCGGTAAGGT TACTGATAGC GCCCCGAGTA CAAACGACGA AAGCACTAAA
                                                 NS2B
     ----------------------------------------------------------------------------------------------------------
     .  G  K  S  T  D  M  W  I  E  R  T  A  D  I  S  W  E  S  D  A  E  I  T  G  S  S  E  R  V  D  V  R  L  .
```

```
5801 CTGGGAAATC AACAGATATG TGGATTGAGA GAACGGCGGA CATTTCCTGG GAAAGTGATG CAGAGATTAC AGGCTCGAGC GAAAGAGTTG ATGTGCGGCT
     GACCCTTTAG TTGTCTATAC ACCTAACTCT CTTGCCGCCT GTAAAGGACC CTTTCACTAC GTCTCTAATG TCCGAGCTCG CTTTCTCAAC TACACGCCGA
                                                NS2B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · D  D  D   G  N  F  Q   L  M  N   D  P  G   A  P  W  K   I  W  M   L  R  M   V  C  L  A   I  S  A   Y  T  P
5901 TGATGATGAT GGAAACTTCC AGCTCATGAA TGATCCAGGA GCACCTTGGA AGATATGGAT GCTCAGAATG GTCTGTCTCG CGATTAGTGC GTACACCCCC
     ACTACTACTA CCTTTGAAGG TCGAGTACTT ACTAGGTCCT CGTGGAACCT TCTATACCTA CGAGTCTTAC CAGACAGAGC GCTAATCACG CATGTGGGGG
                                                                                NS3
                                                                                ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                NS2B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      W  A  I  L   P  S  V   V  G  F   W  I  T  L   Q  Y  T   K  R  G   G  V  L  W   D  T  P   S  P  K   E  Y  K  K
     ·
6001 TGGGCAATCT TGCCCTCAGT AGTTGGATTT TGGATAACTC TCCAATACAC AAAGAGAGGA GGCGTGTTGT GGGACACTCC CTCACCAAAG GAGTACAAAA
     ACCCGTTAGA ACGGGAGTCA TCAACCTAAA ACCTATTGAG AGGTTATGTG TTTCTCTCCT CCGCACAACA CCCTGTGAGG GAGTGGTTTC CTCATGTTTT
                                                NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · G  D  T   T  T  G   V  Y  R  I   M  T  R   G  L  L   G  S  Y  Q   A  G  A   G  V  M   V  E  G  V   F  H  T ·
6101 AGGGGGACAC GACCACCGGC GTCTACAGGA TCATGACTCG TGGGCTGCTC GGCAGTTATC AAGCAGGAGC AGGCGTGATG GTTGAAGGTG TTTTCCACAC
     TCCCCCTGTG CTGGTGGCCG CAGATGTCCT AGTACTGAGC ACCCGACGAG CCGTCAATAG TTCGTCCTCG TCCGCACTAC CAACTTCCAC AAAAGGTGTG
                                                NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · L  W  H   T  T  K  G   A  A  L   M  S  G   E  G  R  L   D  P  Y   W  G  S   V  K  E  D   R  L  C   Y  G  G
6201 CCTTTGGCAT ACAACAAAAG GAGCCGCTTT GATGAGCGGA GAGGGCCGCC TGGACCCATA CTGGGGCAGT GTCAAGGAGG ATCGACTTTG TTACGGAGGA
     GGAAACCGTA TGTTGTTTTC CTCGGCGAAA CTACTCGCCT CTCCCGGCGG ACCTGGGTAT GACCCCGTCA CAGTTCCTCC TAGCTGAAAC AATGCCTCCT
                                                NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      P  W  K  L   Q  H  K   W  N  G   Q  D  E  V   Q  M  I   V  V  E   P  G  K  N   V  K  N   V  Q  T   K  P  G  V
     ·
6301 CCCTGGAAAT TGCAGCACAA GTGGAACGGG CAGGATGAGG TGCAGATGAT TGTGGTGGAA CCTGGCAAGA ACGTTAAGAA CGTCCAGACG AAACCAGGGG
     GGGACCTTTA ACGTCGTGTT CACCTTGCCC GTCCTACTCC ACGTCTACTA ACACCACCTT GGACCGTTCT TGCAATTCTT GCAGGTCTGC TTTGGTCCCC
                                                NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · F  K  T   P  E  G   E  I  G  A   V  T  L   D  F  P   T  G  T  S   G  S  P   I  V  D   K  N  G  D   V  I  G ·
6401 TGTTCAAAAC ACCTGAAGGA GAAATCGGGG CCGTGACTTT GGACTTCCCC ACTGGAACAT CAGGCTCACC AATAGTGGAC AAAAACGGTG ATGTGATTGG
     ACAAGTTTTG TGGACTTCCT CTTTAGCCCC GGCACTGAAA CCTGAAGGGG TGACCTTGTA GTCCGAGTGG TTATCACCTG TTTTTGCCAC TACACTAACC
                                                NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · L  Y  G   N  G  V  I   M  P  N   G  S  Y   I  S  A  I   V  Q  G   E  R  M   D  E  P  I   P  A  G   F  E  P
6501 GCTTTATGGC AATGGAGTCA TAATGCCCAA CGGCTCATAC ATAAGCGCGA TAGTGCAGGG TGAAAGGATG GATGAGCCAA TCCCAGCCGG ATTCGAACCT
     CGAAATACCG TTACCTCAGT ATTACGGGTT GCCGAGTATG TATTCGCGCT ATCACGTCCC ACTTTCCTAC CTACTCGGTT AGGGTCGGCC TAAGCTTGGA
                                                NS3
```

```
                                                                                                                    
     E  M  L  R  K  K  Q  I  T  V  L  D  L  H  P  G  A  G  K  T  R  R  I  L  P  Q  I  I  K  E  A  I  N  R
     .
6601 GAGATGCTGA GGAAAAAACA GATCACTGTA CTGGATCTCC ATCCCGGCGC CGGTAAAACA AGGAGGATTC TGCCACAGAT CATCAAAGAG GCCATAAACA
     CTCTACGACT CCTTTTTTGT CTAGTGACAT GACCTAGAGG TAGGGCCGCG GCCATTTTGT TCCTCCTAAG ACGGTGTCTA GTAGTTTCTC CGGTATTTGT
                                                 NS3
                                                                                                                    
     .  R  L  R  T  A  V  L  A  P  T  R  V  V  A  A  E  M  A  E  A  L  R  G  L  P  I  R  Y  Q  T  S  A  V  .
6701 GAAGACTGAG AACAGCCGTG CTAGCACCAA CCAGGGTTGT GGCTGCTGAG ATGGCTGAAG CACTGAGAGG ACTGCCCATC CGGTACCAGA CATCCGCAGT
     CTTCTGACTC TTGTCGGCAC GATCGTGGTT GGTCCCAACA CCGACGACTC TACCGACTTC GTGACTCTCC TGACGGGTAG GCCATGGTCT GTAGGCGTCA
                                                 NS3
                                                                                                                    
     .  P  R  E  H  N  G  N  E  I  V  D  V  M  C  H  A  T  L  T  H  R  L  M  S  P  H  R  V  P  N  Y  N  L
6801 GCCCAGAGAA CATAATGGAA ATGAGATTGT TGATGTCATG TGTCATGCTA CCCTCACCCA CAGGCTGATG TCTCCTCACA GGGTGCCGAA CTACAACCTG
     CGGGTCTCTT GTATTACCTT TACTCTAACA ACTACAGTAC ACAGTACGAT GGGAGTGGGT GTCCGACTAC AGAGGAGTGT CCCACGGCTT GATGTTGGAC
                                                 NS3
                                                                                                                    
     F  V  M  D  E  A  H  F  T  D  P  A  S  I  A  A  R  G  Y  I  S  T  K  V  E  L  G  E  A  A  A  I  F  M
     .
6901 TTCGTGATGG ATGAGGCTCA TTTCACCGAC CCAGCTAGCA TTGCAGCAAG AGGTTACATT TCCACAAAGG TCGAGCTAGG GGAGGCGGCG GCAATATTCA
     AAGCACTACC TACTCCGAGT AAAGTGGCTG GGTCGATCGT AACGTCGTTC TCCAATGTAA AGGTGTTTCC AGCTCGATCC CCTCCGCCGC CGTTATAAGT
                                                 NS3
                                                                                                                    
     .  T  A  T  P  P  G  T  S  D  P  F  P  E  S  N  S  P  I  S  D  L  Q  T  E  I  P  D  R  A  W  N  S  G  .
7001 TGACAGCCAC CCCACCAGGC ACTTCAGATC CATTCCCAGA GTCCAATTCA CCAATTTCCG ACTTACAGAC TGAGATCCCG GATCGAGCTT GGAACTCTGG
     ACTGTCGGTG GGGTGGTCCG TGAAGTCTAG GTAAGGGTCT CAGGTTAAGT GGTTAAAGGC TGAATGTCTG ACTCTAGGGC CTAGCTCGAA CCTTGAGACC
                                                 NS3
                                                                                                                    
     .  Y  E  W  I  T  E  Y  T  G  K  T  V  W  F  V  P  S  V  K  M  G  N  E  I  A  L  C  L  Q  R  A  G  K
7101 ATACGAATGG ATCACAGAAT ACACCGGGAA GACGGTTTGG TTTGTGCCTA GTGTTAAGAT GGGGAATGAG ATTGCCCTTT GCCTACAACG TGCTGGAAAG
     TATGCTTACC TAGTGTCTTA TGTGGCCCTT CTGCCAAACC AAACACGGAT CACAATTCTA CCCCTTACTC TAACGGGAAA CGGATGTTGC ACGACCTTTC
                                                 NS3
                                                                                                                    
     K  V  V  Q  L  N  R  K  S  Y  E  T  E  Y  P  K  C  K  N  D  D  W  D  F  V  I  T  T  D  I  S  E  M  G
     .
7201 AAAGTAGTCC AATTGAACAG AAAGTCGTAC GAGACGGAGT ACCCAAAATG TAAGAACGAT GATTGGGACT TTGTTATCAC AACAGACATA TCTGAAATGG
     TTTCATCAGG TTAACTTGTC TTTCAGCATG CTCTGCCTCA TGGGTTTTAC ATTCTTGCTA CTAACCCTGA AACAATAGTG TTGTCTGTAT AGACTTTACC
                                                 NS3
                                                                                                                    
     .  A  N  F  K  A  S  R  V  I  D  S  R  K  S  V  K  P  T  I  I  T  E  G  E  G  R  V  I  L  G  E  P  S  .
7301 GGGCTAACTT CAAGGCGAGC AGGGTGATTG ACAGCCGGAA GAGTGTGAAA CCAACCATCA TAACAGAAGG AGAAGGGAGA GTGATCCTGG GAGAACCATC
     CCCGATTGAA GTTCCGCTCG TCCCACTAAC TGTCGGCCTT CTCACACTTT GGTTGGTAGT ATTGTCTTCC TCTTCCCTCT CACTAGGACC CTCTTGGTAG
```

```
                                                    NS3
     ~~~
```

```
     M  T  M  G  V  F  F  L  L  M  Q  R  K  G  I  G  K  I  G  L  G  G  A  V  L  G  V  A  T  F  F  C  W  M
     .
8101 ATGACCATGG GAGTATTCTT CCTCCTCATG CAGCGGAAGG GCATTGGAAA GATAGGTTTG GGAGGCGCTG TCTTGGGAGT CGCGACCTTT TTCTGTTGGA
     TACTGGTACC CTCATAAGAA GGAGGAGTAC GTCGCCTTCC CGTAACCTTT CTATCCAAAC CCTCCGCGAC AGAACCCTCA GCGCTGGAAA AAGACAACCT
                                                 NS4A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     .  A  E  V  P  G  T  K  I  A  G  M  L  L  L  S  L  L  L  M  I  V  L  I  P  E  P  E  K  Q  R  S  Q  T  .
8201 TGGCTGAAGT TCCAGGAACG AAGATCGCCG GAATGTTGCT GCTCTCCCTT CTCTTGATGA TTGTGCTAAT TCCTGAGCCA GAGAAGCAAC GTTCGCAGAC
     ACCGACTTCA AGGTCCTTGC TTCTAGCGGC CTTACAACGA CGAGAGGGAA GAGAACTACT AACACGATTA AGGACTCGGT CTCTTCGTTG CAAGCGTCTG
                                                                                    NS4B
                                                                  ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                  NS4A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     .  D  N  Q  L  A  V  F  L  I  C  V  M  T  L  V  S  A  V  A  A  N  E  M  G  W  L  D  K  T  K  S  D  I
8301 AGACAACCAG CTAGCCGTGT TCCTGATATG TGTCATGACC CTTGTGAGCG CAGTGGCAGC CAACGAGATG GGTTGGCTAG ATAAGACCAA GAGTGACATA
     TCTGTTGGTC GATCGGCACA AGGACTATAC ACAGTACTGG GAACACTCGC GTCACCGTCG GTTGCTCTAC CCAACCGATC TATTCTGGTT CTCACTGTAT
                                                 NS4B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     .  S  S  L  F  G  Q  R  I  E  V  K  E  N  F  S  M  G  E  F  L  L  D  L  R  P  A  T  A  W  S  L  Y  A  V
     .
8401 AGCAGTTTGT TTGGGCAAAG AATTGAGGTC AAGGAGAATT TCAGCATGGG AGAGTTTCTT CTGGACTTGA GGCCGGCAAC AGCCTGGTCA CTGTACGCTG
     TCGTCAAACA AACCCGTTTC TTAACTCCAG TTCCTCTTAA AGTCGTACCC TCTCAAAGAA GACCTGAACT CCGGCCGTTG TCGGACCAGT GACATGCGAC
                                                 NS4B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     .  T  T  A  V  L  T  P  L  L  K  H  L  I  T  S  D  Y  I  N  T  S  L  T  S  I  N  V  Q  A  S  A  L  F  .
8501 TGACAACAGC GGTCCTCACT CCACTGCTAA AGCATTTGAT CACGTCAGAT TACATCAACA CCTCATTGAC CTCAATAAAC GTTCAGGCAA GTGCACTATT
     ACTGTTGTCG CCAGGAGTGA GGTGACGATT TCGTAAACTA GTGCAGTCTA ATGTAGTTGT GGAGTAACTG GAGTTATTTG CAAGTCCGTT CACGTGATAA
                                                 NS4B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     .  T  L  A  R  G  F  P  F  V  D  V  G  V  S  A  L  L  L  A  A  G  C  W  G  Q  V  T  L  T  V  T  V  T
8601 CACACTCGCG CGAGGCTTCC CCTTCGTCGA TGTTGGAGTG TCGGCTCTCC TGCTAGCAGC CGGATGCTGG GGACAAGTCA CCCTCACCGT TACGGTAACA
     GTGTGAGCGC GCTCCGAAGG GGAAGCAGCT ACAACCTCAC AGCCGAGAGG ACGATCGTCG GCCTACGACC CCTGTTCAGT GGGAGTGGCA ATGCCATTGT
                                                 NS4B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     A  A  T  L  L  F  C  H  Y  A  Y  M  V  P  G  W  Q  A  E  A  M  R  S  A  Q  R  R  T  A  A  G  I  M  K
     .
8701 GCGGCAACAC TCCTTTTTTG CCACTATGCC TACATGGTTC CCGGTTGGCA AGCTGAGGCA ATGCGCTCAG CCCAGCGGCG GACAGCGGCC GGAATCATGA
     CGCCGTTGTG AGGAAAAAAC GGTGATACGG ATGTACCAAG GGCCAACCGT TCGACTCCGT TACGCGAGTC GGGTCGCCGC CTGTCGCCGG CCTTAGTACT
                                                 NS4B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     .  N  A  V  V  D  G  I  V  A  T  D  V  P  E  L  E  R  T  T  P  I  M  Q  K  K  I  G  Q  I  M  L  I  L  .
8801 AGAACGCTGT AGTGGATGGC ATCGTGGCCA CGGACGTCCC AGAATTAGAG CGCACCACAC CCATCATGCA GAAGAAAATT GGACAGATCA TGCTGATCTT
```

```
     TCTTGCGACA TCACCTACCG TAGCACCGGT GCCTGCAGGG TCTTAATCTC GCGTGGTGTG GGTAGTACGT CTTCTTTTAA CCTGTCTAGT ACGACTAGAA
                                                 NS4B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · V  S  L  A  A  V  V  V  N  P  S  V  K  T  V  R  E  A  G  I  L  I  T  A  A  A  V  T  L  W  E  N  G
8901 GGTGTCTCTA GCTGCAGTAG TAGTGAACCC GTCTGTGAAG ACAGTACGAG AAGCCGGAAT TTTGATCACG GCCGCAGCGG TGACGCTTTG GGAGAATGGA
     CCACAGAGAT CGACGTCATC ATCACTTGGG CAGACACTTC TGTCATGCTC TTCGGCCTTA AAACTAGTGC CGGCGTCGCC ACTGCGAAAC CCTCTTACCT
                                                 NS4B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      A  S  S  V  W  N  A  T  T  A  I  G  L  C  H  I  M  R  G  G  W  L  S  C  L  S  I  T  W  T  L  I  K  N
      ·
9001 GCAAGCTCTG TTTGGAACGC AACAACTGCC ATCGGACTCT GCCACATCAT GCGTGGGGGT TGGTTGTCAT GTCTATCCAT AACATGGACA CTCATAAAGA
     CGTTCGAGAC AAACCTTGCG TTGTTGACGG TAGCCTGAGA CGGTGTAGTA CGCACCCCCA ACCAACAGTA CAGATAGGTA TTGTACCTGT GAGTATTTCT
                                                                 NS5
                                 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
              NS4B
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · M  E  K  P  G  L  K  R  G  G  A  K  G  R  T  L  G  E  V  W  K  E  R  L  N  Q  M  T  K  E  E  F  T  ·
9101 ACATGGAAAA ACCAGGACTA AAAAGAGGTG GGGCAAAAGG ACGCACCTTG GGAGAGGTTT GGAAAGAAAG ACTCAACCAG ATGACAAAAG AAGAGTTCAC
     TGTACCTTTT TGGTCCTGAT TTTTCTCCAC CCCGTTTTCC TGCGTGGAAC CCTCTCCAAA CCTTTCTTTC TGAGTTGGTC TACTGTTTTC TTCTCAAGTG
                                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · R  Y  R  K  E  A  I  I  E  V  D  R  S  A  A  K  H  A  R  K  E  G  N  V  T  G  G  H  P  V  S  R  G
9201 TAGGTACCGC AAAGAGGCCA TCATCGAAGT CGATCGCTCA GCGGCAAAAC ACGCCAGGAA AGAAGGCAAT GTCACTGGAG GGCATCCAGT CTCTAGGGGC
     ATCCATGGCG TTTCTCCGGT AGTAGCTTCA GCTAGCGAGT CGCCGTTTTG TGCGGTCCTT TCTTCCGTTA CAGTGACCTC CCGTAGGTCA GAGATCCCCG
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      T  A  K  L  R  W  L  V  E  R  R  F  L  E  P  V  G  K  V  I  D  L  G  C  G  R  G  G  W  C  Y  Y  M  A
      ·
9301 ACAGCAAAAC TGAGATGGCT GGTCGAACGG AGGTTTCTCG AACCGGTCGG AAAAGTGATT GACCTTGGAT GTGGAAGAGG CGGTTGGTGT TACTATATGG
     TGTCGTTTTG ACTCTACCGA CCAGCTTGCC TCCAAAGAGC TTGGCCAGCC TTTTCACTAA CTGGAACCTA CACCTTCTCC GCCAACCACA ATGATATACC
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · T  Q  K  R  V  Q  E  V  R  G  Y  T  K  G  G  P  G  H  E  E  P  Q  L  V  Q  S  Y  G  W  N  I  V  T  ·
9401 CAACCCAAAA AAGAGTCCAA GAAGTCAGAG GGTACACAAA GGGCGGTCCC GGACATGAAG AGCCCCAACT AGTGCAAAGT TATGGATGGA ACATTGTCAC
     GTTGGGTTTT TTCTCAGGTT CTTCAGTCTC CCATGTGTTT CCCGCCAGGG CCTGTACTTC TCGGGGTTGA TCACGTTTCA ATACCTACCT TGTAACAGTG
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · M  K  S  G  V  D  V  F  Y  R  P  S  E  C  C  D  T  L  L  C  D  I  G  E  S  S  S  S  A  E  V  E  E
9501 CATGAAGAGT GGAGTGGATG TGTTCTACAG ACCTTCTGAG TGTTGTGACA CCCTCCTTTG TGACATCGGA GAGTCCTCGT CAAGTGCTGA GGTTGAAGAG
     GTACTTCTCA CCTCACCTAC ACAAGATGTC TGGAAGACTC ACAACACTGT GGGAGGAAAC ACTGTAGCCT CTCAGGAGCA GTTCACGACT CCAACTTCTC
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

```
     H  R  T  I  R  V  L  E  M  V  E  D  W  L  H  R  G  P  R  E  F  C  V  K  V  L  C  P  Y  M  P  K  V  I
.
9601 CATAGGACGA TTCGGGTCCT TGAAATGGTT GAGGACTGGC TGCACCGAGG GCCAAGGGAA TTTTGCGTGA AGGTGCTCTG CCCCTACATG CCGAAAGTCA
     GTATCCTGCT AAGCCCAGGA ACTTTACCAA CTCCTGACCG ACGTGGCTCC CGGTTCCCTT AAAACGCACT TCCACGAGAC GGGGATGTAC GGCTTTCAGT
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · E  K  M  E  L  L  Q  R  R  Y  G  G  G  L  V  R  N  P  L  S  R  N  S  T  H  E  M  Y  W  V  S  R  A ·
9701 TAGAGAAGAT GGAGCTGCTC CAACGCCGGT ATGGGGGGGG ACTGGTCAGA AACCCACTCT CACGGAATTC CACGCACGAG ATGTATTGGG TGAGTCGAGC
     ATCTCTTCTA CCTCGACGAG GTTGCGGCCA TACCCCCCCC TGACCAGTCT TTGGGTGAGA GTGCCTTAAG GTGCGTGCTC TACATAACCC ACTCAGCTCG
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · S  G  N  V  V  H  S  V  N  M  T  S  Q  V  L  L  G  R  M  E  K  R  T  W  K  G  P  Q  Y  E  E  D  V
9801 TTCAGGCAAT GTGGTACATT CAGTGAATAT GACCAGCCAG GTGCTCCTAG GAAGAATGGA AAAAAGGACC TGGAAGGGAC CCCAATACGA GGAAGACGTA
     AAGTCCGTTA CACCATGTAA GTCACTTATA CTGGTCGGTC CACGAGGATC CTTCTTACCT TTTTTCCTGG ACCTTCCCTG GGGTTATGCT CCTTCTGCAT
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     N  L  G  S  G  T  R  A  V  G  K  P  L  L  N  S  D  T  S  K  I  K  N  R  I  E  R  L  R  R  E  Y  S  S
.
9901 AACTTGGGAA GTGGAACCAG GGCGGTGGGA AAACCCCTGC TCAACTCAGA CACCAGTAAA ATCAAGAACA GGATTGAACG ACTCAGGCGT GAGTACAGTT
     TTGAACCCTT CACCTTGGTC CCGCCACCCT TTTGGGGACG AGTTGAGTCT GTGGTCATTT TAGTTCTTGT CCTAACTTGC TGAGTCCGCA CTCATGTCAA
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · T  W  H  H  D  E  N  H  P  Y  R  T  W  N  Y  H  G  S  Y  D  V  K  P  T  G  S  A  S  S  L  V  N  G ·
10001 CGACGTGGCA CCACGATGAG AACCACCCAT ATAGAACCTG GAACTATCAT GGCAGTTATG ATGTGAAGCC CACAGGCTCC GCCAGTTCGC TGGTCAATGG
      GCTGCACCGT GGTGCTACTC TTGGTGGGTA TATCTTGGAC CTTGATAGTA CCGTCAATAC TACACTTCGG GTGTCCGAGG CGGTCAAGCG ACCAGTTACC
                                                  NS5
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      · V  V  R  L  L  S  K  P  W  D  T  I  T  N  V  T  T  M  A  M  T  D  T  T  P  F  G  Q  Q  R  V  F  K
10101 AGTGGTCAGG CTCCTCTCAA AACCATGGGA CACCATCACG AATGTTACCA CCATGGCCAT GACTGACACT ACTCCCTTCG GGCAGCAGCG AGTGTTCAAA
      TCACCAGTCC GAGGAGAGTT TTGGTACCCT GTGGTAGTGC TTACAATGGT GGTACCGGTA CTGACTGTGA TGAGGGAAGC CCGTCGTCGC TCACAAGTTT
                                                  NS5
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      E  K  V  D  T  K  A  P  E  P  P  E  G  V  K  Y  V  L  N  E  T  T  N  W  L  W  A  F  L  A  R  E  K  R
.
10201 GAGAAGGTGG ACACGAAAGC TCCTGAACCG CCAGAAGGAG TGAAGTACGT GCTCAACGAG ACCACCAACT GGTTGTGGGC GTTTTTGGCC AGAGAAAAAC
      CTCTTCCACC TGTGCTTTCG AGGACTTGGC GGTCTTCCTC ACTTCATGCA CGAGTTGCTC TGGTGGTTGA CCAACACCCG CAAAAACCGG TCTCTTTTTG
                                                  NS5
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      · P  R  M  C  S  R  E  E  F  I  R  K  V  N  S  N  A  A  L  G  A  M  F  E  E  Q  N  Q  W  R  S  A  R ·
10301 GTCCCAGAAT GTGCTCTCGA GAGGAATTCA TAAGAAAGGT CAACAGCAAT GCAGCTTTGG GTGCCATGTT TGAAGAGCAG AATCAATGGA GGAGCGCCAG
      CAGGGTCTTA CACGAGAGCT CTCCTTAAGT ATTCTTTCCA GTTGTCGTTA CGTCGAAACC CACGGTACAA ACTTCTCGTC TTAGTTACCT CCTCGCGGTC
                                                  NS5
```

31

```
                                                                                                                    
     · E  A  V  E  D  P  K  F  W  E  M  V  D  E  E  R  E  A  H  L  R  G  E  C  H  T  C  I  Y  N  M  M  G
10401 AGAAGCAGTT GAAGATCCAA AATTTTGGGA AATGGTGGAT GAGGAGCGCG AGGCACATCT GCGGGGGGAA TGTCACACTT GCATTTACAA CATGATGGGA
      TCTTCGTCAA CTTCTAGGTT TTAAAACCCT TTACCACCTA CTCCTCGCGC TCCGTGTAGA CGCCCCCCTT ACAGTGTGAA CGTAAATGTT GTACTACCCT
                                                 NS5
                                                                                                                    
      K  R  E  K  K  P  G  E  F  G  K  A  K  G  S  R  A  I  W  F  M  W  L  G  A  R  F  L  E  F  E  A  L  G
     ·
10501 AAGAGAGAGA AAAAACCCGG AGAGTTCGGA AAGGCCAAGG GAAGCAGAGC CATTTGGTTC ATGTGGCTCG GAGCTCGCTT TCTGGAGTTC GAGGCTCTGG
      TTCTCTCTCT TTTTTGGGCC TCTCAAGCCT TTCCGGTTCC CTTCGTCTCG GTAAACCAAG TACACCGAGC CTCGAGCGAA AGACCTCAAG CTCCGAGACC
                                                 NS5
                                                                                                                    
     · F  L  N  E  D  H  W  L  G  R  K  N  S  G  G  G  V  E  G  L  G  L  Q  K  L  G  Y  I  L  R  E  V  G ·
10601 GTTTTCTCAA TGAAGACCAC TGGCTTGGAA GAAAGAACTC AGGAGGAGGT GTCGAGGGCT TGGGCCTCCA AAAACTGGGT TACATCCTGC GTGAAGTTGG
      CAAAAGAGTT ACTTCTGGTG ACCGAACCTT CTTTCTTGAG TCCTCCTCCA CAGCTCCCGA ACCCGGAGGT TTTTGACCCA ATGTAGGACG CACTTCAACC
                                                 NS5
                                                                                                                    
     · I  R  P  G  G  K  I  Y  A  D  D  T  A  G  W  D  T  R  I  T  R  A  D  L  E  N  E  A  K  V  L  E  L
10701 CATCCGGCCT GGGGGCAAGA TCTATGCTGA TGACACAGCT GGCTGGGACA CCCGCATCAC GAGAGCTGAC TTGGAAAATG AAGCTAAGGT GCTTGAGCTG
      GTAGGCCGGA CCCCCGTTCT AGATACGACT ACTGTGTCGA CCGACCCTGT GGGCGTAGTG CTCTCGACTG AACCTTTTAC TTCGATTCCA CGAACTCGAC
                                                 NS5
                                                                                                                    
      L  D  G  E  H  R  R  L  A  R  A  I  I  E  L  T  Y  R  H  K  V  V  K  V  M  R  P  A  A  D  G  R  T  V
     ·
10801 CTTGATGGGG AACATCGGCG TCTTGCCAGG GCCATCATTG AGCTCACCTA TCGTCACAAA GTTGTGAAAG TGATGCGCCC GGCTGCTGAT GGAAGAACCG
      GAACTACCCC TTGTAGCCGC AGAACGGTCC CGGTAGTAAC TCGAGTGGAT AGCAGTGTTT CAACACTTTC ACTACGCGGG CCGACGACTA CCTTCTTGGC
                                                 NS5
                                                                                                                    
     · M  D  V  I  S  R  E  D  Q  R  G  S  G  Q  V  V  T  Y  A  L  N  T  F  T  N  L  A  V  Q  L  V  R  M ·
10901 TTATGGATGT TATCTCCAGA GAAGATCAGA GGGGGAGTGG ACAAGTTGTC ACCTACGCCC TAAACACTTT CACCAACCTG GCTGTCCAGC TGGTGAGGAT
      AATACCTACA ATAGAGGTCT CTTCTAGTCT CCCCCTCACC TGTTCAACAG TGGATGCGGG ATTTGTGAAA GTGGTTGGAC CGACAGGTCG ACCACTCCTA
                                                 NS5
                                                                                                                    
     · M  E  G  E  G  V  I  G  P  D  D  V  E  K  L  T  K  G  K  G  P  K  V  R  T  W  L  F  E  N  G  E  E
11001 GATGGAAGGG GAAGGAGTGA TTGGCCCAGA TGATGTGGAG AAACTCACAA AAGGGAAAGG ACCCAAAGTC AGGACCTGGC TGTTTGAGAA TGGGGAAGAA
      CTACCTTCCC CTTCCTCACT AACCGGGTCT ACTACACCTC TTTGAGTGTT TTCCCTTTCC TGGGTTTCAG TCCTGGACCG ACAAACTCTT ACCCCTTCTT
                                                 NS5
                                                                                                                    
      R  L  S  R  M  A  V  S  G  D  D  C  V  V  K  P  L  D  D  R  F  A  T  S  L  H  F  L  N  A  M  S  K  V
     ·
11101 AGACTCAGCC GCATGGCTGT CAGTGGAGAT GACTGTGTGG TAAAGCCCCT GGACGATCGC TTTGCCACCT CGCTCCACTT CCTCAATGCT ATGTCAAAGG
      TCTGAGTCGG CGTACCGACA GTCACCTCTA CTGACACACC ATTTCGGGGA CCTGCTAGCG AAACGGTGGA GCGAGGTGAA GGAGTTACGA TACAGTTTCC
```

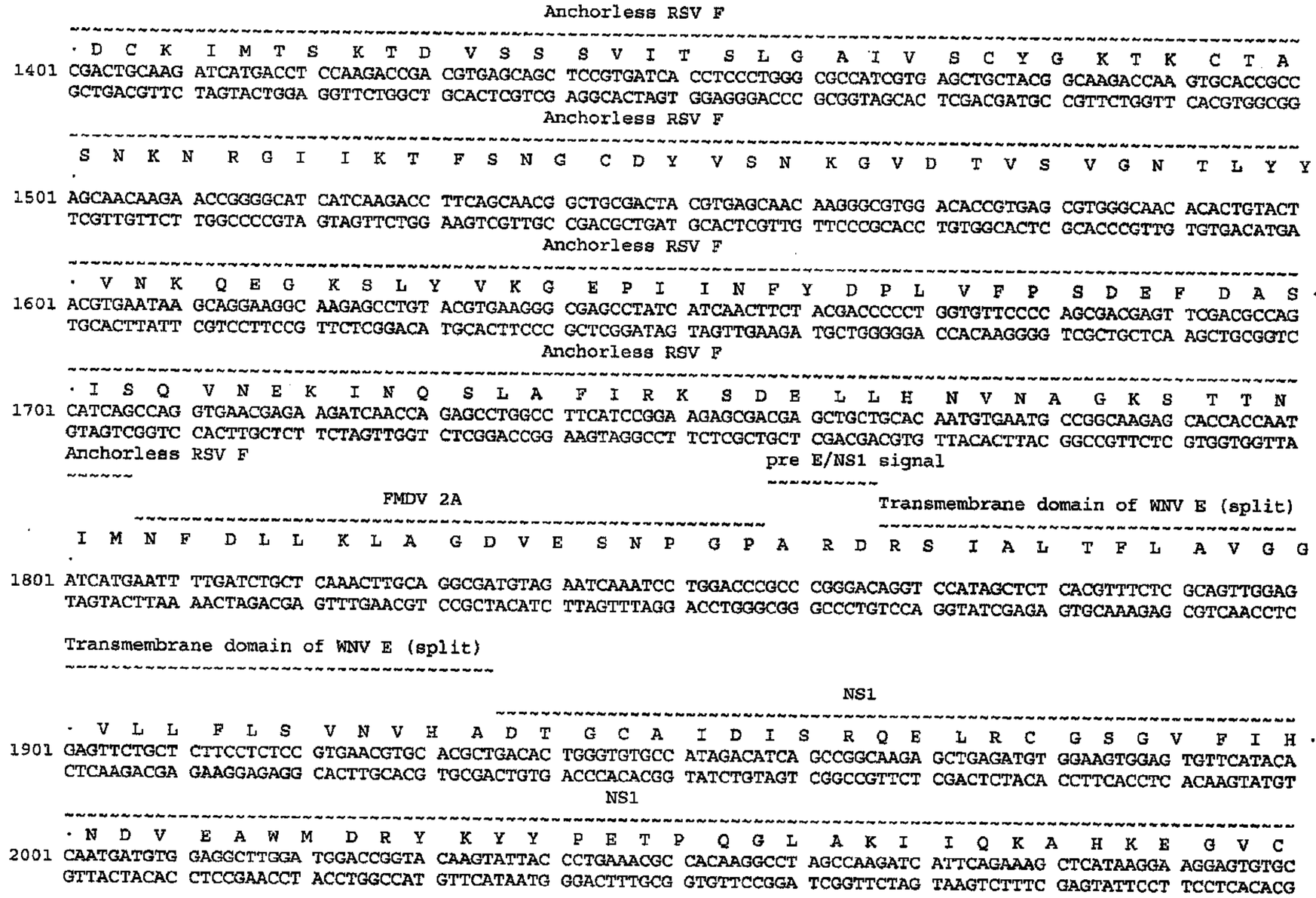
```
                                                                                                                    NS5
                                                                                                                    ~~~
```

```
      CATCATAAAT CACCACAATC ACATTTATCA ATTCTTTTAG AACTCCTCTT TCAGTCCGGC CCTTCAAGGG CGGTGGCCTT CAACTCATCT GCCACGACGG
                                                 3' UTR
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
12001 TGCGACTCAA CCCCAGGAGG ACTGGGTGAA CAAAGCCGCG AAGTGATCCA TGTAAGCCCT CAGAACCGTC TCGGAAGGAG GACCCCACAT GTTGTAACTT
      ACGCTGAGTT GGGGTCCTCC TGACCCACTT GTTTCGGCGC TTCACTAGGT ACATTCGGGA GTCTTGGCAG AGCCTTCCTC CTGGGGTGTA CAACATTGAA
                                                 3' UTR
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
12101 CAAAGCCCAA TGTCAGACCA CGCTACGGCG TGCTACTCTG CGGAGAGTGC AGTCTGCGAT AGTGCCCCAG GAGGACTGGG TTAACAAAGG CAAACCAACG
      GTTTCGGGTT ACAGTCTGGT GCGATGCCGC ACGATGAGAC GCCTCTCACG TCAGACGCTA TCACGGGGTC CTCCTGACCC AATTGTTTCC GTTTGGTTGC
                                                 3' UTR
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
12201 CCCCACGCGG CCCAAGCCCC GGTAATGGTG TTAACCAGGG CGAAAGGACT AGAGGTTAGA GGAGACCCCG CGGTTTAAAG TGCACGGCCC AGCCTGGCTG
      GGGGTGCGCC GGGTTCGGGG CCATTACCAC AATTGGTCCC GCTTTCCTGA TCTCCAATCT CCTCTGGGGC GCCAAATTTC ACGTGCCGGG TCGGACCGAC
                                                 3' UTR
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
12301 AAGCTGTAGG TCAGGGGAAG GACTAGAGGT TAGTGGAGAC CCCGTGCCAC AAAACACCAC AACAAAACAG CATATTGACA CCTGGGATAG ACTAGGAGAT
      TTCGACATCC AGTCCCCTTC CTGATCTCCA ATCACCTCTG GGGCACGGTG TTTTGTGGTG TTGTTTTGTC GTATAACTGT GGACCCTATC TGATCCTCTA
                                        3' UTR
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
12401 CTTCTGCTCT GCACAACCAG CCACACGGCA CAGTGCGCCG ACAATGGTGG CTGGTGGTGC GAGAACACAG GATCT
      GAAGACGAGA CGTGTTGGTC GGTGTGCCGT GTCACGCGGC TGTTACCACC GACCACCACG CTCTTGTGTC CTAGA
```

RepliVax WN - Anchorless F inserted in place of AC-prM-E.
F insert starts at nucleotide position 229bp and ends at 1806bp.

```
                                                                                                                    C
                                                                                                                    ~~~~
                                                5' UTR
   ~~~~
```

```
                                                                                                                    T  N  V  T  L  S  K  K  R  K  R  R  F  L  G  F  L  L  G  V  G  S  A  I  A  S  G  I  A  V  S  K  V  L
                                                                                                                    .
601 ACCAACGTGA CCCTGAGCAA GAAGCGGAAG CGGCGGTTCC TGGGCTTCCT GCTGGGCGTG GGCAGCGCCA TCGCCAGCGG CATCGCCGTG TCCAAGGTGC
    TGGTTGCACT GGGACTCGTT CTTCGCCTTC GCCGCCAAGG ACCCGAAGGA CGACCCGCAC CCGTCGCGGT AGCGGTCGCC GTAGCGGCAC AGGTTCCACG
                                                                                                                                                Anchorless RSV F
                                                                                                                    ~~~
```


Anchorless RSV F
· D C K I M T S K T D V S S S V I T S L G A I V S C Y G K T K C T A
1401 CGACTGCAAG ATCATGACCT CCAAGACCGA CGTGAGCAGC TCCGTGATCA CCTCCCTGGG CGCCATCGTG AGCTGCTACG GCAAGACCAA GTGCACCGCC
GCTGACGTTC TAGTACTGGA GGTTCTGGCT GCACTCGTCG AGGCACTAGT GGAGGGACCC GCGGTAGCAC TCGACGATGC CGTTCTGGTT CACGTGGCGG
Anchorless RSV F
S N K N R G I I K T F S N G C D Y V S N K G V D T V S V G N T L Y Y
1501 AGCAACAAGA ACCGGGGCAT CATCAAGACC TTCAGCAACG GCTGCGACTA CGTGAGCAAC AAGGGCGTGG ACACCGTGAG CGTGGGCAAC ACACTGTACT
TCGTTGTTCT TGGCCCCGTA GTAGTTCTGG AAGTCGTTGC CGACGCTGAT GCACTCGTTG TTCCCGCACC TGTGGCACTC GCACCCGTTG TGTGACATGA
Anchorless RSV F
· V N K Q E G K S L Y V K G E P I I N F Y D P L V F P S D E F D A S ·
1601 ACGTGAATAA GCAGGAAGGC AAGAGCCTGT ACGTGAAGGG CGAGCCTATC ATCAACTTCT ACGACCCCCT GGTGTTCCCC AGCGACGAGT TCGACGCCAG
TGCACTTATT CGTCCTTCCG TTCTCGGACA TGCACTTCCC GCTCGGATAG TAGTTGAAGA TGCTGGGGGA CCACAAGGGG TCGCTGCTCA AGCTGCGGTC
Anchorless RSV F
· I S Q V N E K I N Q S L A F I R K S D E L L H N V N A G K S T T N
1701 CATCAGCCAG GTGAACGAGA AGATCAACCA GAGCCTGGCC TTCATCCGGA AGAGCGACGA GCTGCTGCAC AATGTGAATG CCGGCAAGAG CACCACCAAT
GTAGTCGGTC CACTTGCTCT TCTAGTTGGT CTCGGACCGG AAGTAGGCCT TCTCGCTGCT CGACGACGTG TTACACTTAC GGCCGTTCTC GTGGTGGTTA
Anchorless RSV F
pre E/NS1 signal
FMDV 2A
Transmembrane domain of WNV E (split)
I M N F D L L K L A G D V E S N P G P A R D R S I A L T F L A V G G
1801 ATCATGAATT TTGATCTGCT CAAACTTGCA GGCGATGTAG AATCAAATCC TGGACCCGCC CGGGACAGGT CCATAGCTCT CACGTTTCTC GCAGTTGGAG
TAGTACTTAA AACTAGACGA GTTTGAACGT CCGCTACATC TTAGTTTAGG ACCTGGGCGG GCCCTGTCCA GGTATCGAGA GTGCAAAGAG CGTCAACCTC
Transmembrane domain of WNV E (split)
NS1
· V L L F L S V N V H A D T G C A I D I S R Q E L R C G S G V F I H ·
1901 GAGTTCTGCT CTTCCTCTCC GTGAACGTGC ACGCTGACAC TGGGTGTGCC ATAGACATCA GCCGGCAAGA GCTGAGATGT GGAAGTGGAG TGTTCATACA
CTCAAGACGA GAAGGAGAGG CACTTGCACG TGCGACTGTG ACCCACACGG TATCTGTAGT CGGCCGTTCT CGACTCTACA CCTTCACCTC ACAAGTATGT
NS1
· N D V E A W M D R Y K Y Y P E T P Q G L A K I I Q K A H K E G V C
2001 CAATGATGTG GAGGCTTGGA TGGACCGGTA CAAGTATTAC CCTGAAACGC CACAAGGCCT AGCCAAGATC ATTCAGAAAG CTCATAAGGA AGGAGTGTGC
GTTACTACAC CTCCGAACCT ACCTGGCCAT GTTCATAATG GGACTTTGCG GTGTTCCGGA TCGGTTCTAG TAAGTCTTTC GAGTATTCCT TCCTCACACG

```
                                                
```

```
2801 AGAGCTGCGG ACACCGTGGA CCTGCCACTC GCACCACCAC AGAGAGCGGA AAGTTGATAA CAGATTGGTG CTGCAGGAGC TGCACCTTAC CACCACTGCG
     TCTCGACGCC TGTGGCACCT GGACGGTGAG CGTGGTGGTG TCTCTCGCCT TTCAACTATT GTCTAACCAC GACGTCCTCG ACGTGGAATG GTGGTGACGC
                                                                                                      NS2A
                                                                                                      ~~~~~~~~~
                                          NS1
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · Y  Q  T  D  S  G  C  W  Y  G  M  E  I  R  P  Q  R  H  D  E  K  T  L  V  Q  S  Q  V  N  A  Y  N  A
2901 CTACCAAACT GACAGCGGCT GTTGGTATGG TATGGAGATC AGACCACAGA GACATGATGA AAAGACCCTC GTGCAGTCAC AAGTGAATGC TTATAATGCT
     GATGGTTTGA CTGTCGCCGA CAACCATACC ATACCTCTAG TCTGGTGTCT CTGTACTACT TTTCTGGGAG CACGTCAGTG TTCACTTACG AATATTACGA
                                               NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     D  M  I  D  P  F  Q  L  G  L  L  V  V  F  L  A  T  Q  E  V  L  R  K  R  W  T  A  K  I  S  M  P  A  I
     ·
3001 GATATGATTG ACCCTTTTCA GTTGGGCCTT CTGGTCGTGT TCTTGGCCAC CCAGGAGGTC CTTCGCAAGA GGTGGACAGC CAAGATCAGC ATGCCAGCTA
     CTATACTAAC TGGGAAAAGT CAACCCGGAA GACCAGCACA AGAACCGGTG GGTCCTCCAG GAAGCGTTCT CCACCTGTCG GTTCTAGTCG TACGGTCGAT
                                               NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · L  I  A  L  L  V  L  V  F  G  G  I  T  Y  T  D  V  L  R  Y  V  I  L  V  G  A  A  F  A  E  S  N  S  ·
3101 TACTGATTGC TCTGCTAGTC CTGGTGTTTG GGGGCATTAC TTACACTGAT GTGTTACGCT ATGTCATCTT GGTGGGGGCA GCTTTCGCAG AATCTAATTC
     ATGACTAACG AGACGATCAG GACCACAAAC CCCCGTAATG AATGTGACTA CACAATGCGA TACAGTAGAA CCACCCCCGT CGAAAGCGTC TTAGATTAAG
                                               NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · G  G  D  V  V  H  L  A  L  M  A  T  F  K  I  Q  P  V  F  M  V  A  S  F  L  K  A  R  W  T  N  Q  E
3201 GGGAGGAGAC GTGGTACACT TGGCGCTCAT GGCGACCTTC AAGATACAAC CAGTGTTTAT GGTGGCATCG TTTCTTAAAG CGAGATGGAC CAACCAGGAG
     CCCTCCTCTG CACCATGTGA ACCGCGAGTA CCGCTGGAAG TTCTATGTTG GTCACAAATA CCACCGTAGC AAAGAATTTC GCTCTACCTG GTTGGTCCTC
                                               NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     N  I  L  L  M  L  A  A  V  F  F  Q  M  A  Y  H  D  A  R  Q  I  L  L  W  E  I  P  D  V  L  N  S  L  A
     ·
3301 AACATTTTGT TGATGTTGGC GGCTGTTTTC TTTCAAATGG CTTATCACGA TGCCCGCCAA ATTCTGCTCT GGGAGATCCC TGATGTGTTG AATTCACTGG
     TTGTAAAACA ACTACAACCG CCGACAAAAG AAAGTTTACC GAATAGTGCT ACGGGCGGTT TAAGACGAGA CCCTCTAGGG ACTACACAAC TTAAGTGACC
                                               NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · I  A  W  M  I  L  R  A  I  T  F  T  T  T  S  N  V  V  V  P  L  L  A  L  L  T  P  G  L  R  C  L  N  ·
3401 CAATAGCTTG GATGATACTG AGAGCCATAA CATTCACAAC GACATCAAAC GTGGTTGTTC CGCTGCTAGC CCTGCTAACA CCCGGGCTGA GATGCTTGAA
     GTTATCGAAC CTACTATGAC TCTCGGTATT GTAAGTGTTG CTGTAGTTTG CACCAACAAG GCGACGATCG GGACGATTGT GGGCCCGACT CTACGAACTT
                                               NS2A
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · L  D  V  Y  R  I  L  L  L  M  V  G  I  G  S  L  I  R  E  K  R  S  A  A  A  K  K  K  G  A  S  L  L
3501 TCTGGATGTG TACAGGATAC TGCTGTTGAT GGTCGGAATA GGCAGCTTGA TCAGGGAGAA GAGGAGCGCA GCTGCAAAAA AGAAAGGAGC AAGTCTGCTA
     AGACCTACAC ATGTCCTATG ACGACAACTA CCAGCCTTAT CCGTCGAACT AGTCCCTCTT CTCCTCGCGT CGACGTTTTT TCTTTCCTCG TTCAGACGAT
                                                                                                      NS2B
```

```
                                                                                                                  ~~~~~~~~~~~~~~~~~
                                                NS2A
     ~~~~~~~~~~~~~~~~~
```

```
                                                                                                                    
     · R  L  C  Y  G  G  P  W  K  L  Q  H  K  W  N  G  Q  D  E  V  Q  M  I  V  V  E  P  G  K  N  V  K  N ·
4301 ATCGACTTTG TTACGGAGGA CCCTGGAAAT TGCAGCACAA GTGGAACGGG CAGGATGAGG TGCAGATGAT TGTGGTGGAA CCTGGCAAGA ACGTTAAGAA
     TAGCTGAAAC AATGCCTCCT GGGACCTTTA ACGTCGTGTT CACCTTGCCC GTCCTACTCC ACGTCTACTA ACACCACCTT GGACCGTTCT TGCAATTCTT
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     ·V  Q  T  K  P  G  V  F  K  T  P  E  G  E  I  G  A  V  T  L  D  F  P  T  G  T  S  G  S  P  I  V  D
4401 CGTCCAGACG AAACCAGGGG TGTTCAAAAC ACCTGAAGGA GAAATCGGGG CCGTGACTTT GGACTTCCCC ACTGGAACAT CAGGCTCACC AATAGTGGAC
     GCAGGTCTGC TTTGGTCCCC ACAAGTTTTG TGGACTTCCT CTTTAGCCCC GGCACTGAAA CCTGAAGGGG TGACCTTGTA GTCCGAGTGG TTATCACCTG
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      K  N  G  D  V  I  G  L  Y  G  N  G  V  I  M  P  N  G  S  Y  I  S  A  I  V  Q  G  E  R  M  D  E  P  I
     ·
4501 AAAAACGGTG ATGTGATTGG GCTTTATGGC AATGGAGTCA TAATGCCCAA CGGCTCATAC ATAAGCGCGA TAGTGCAGGG TGAAAGGATG GATGAGCCAA
     TTTTTGCCAC TACACTAACC CGAAATACCG TTACCTCAGT ATTACGGGTT GCCGAGTATG TATTCGCGCT ATCACGTCCC ACTTTCCTAC CTACTCGGTT
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · P  A  G  F  E  P  E  M  L  R  K  K  Q  I  T  V  L  D  L  H  P  G  A  G  K  T  R  R  I  L  P  Q  I ·
4601 TCCCAGCCGG ATTCGAACCT GAGATGCTGA GGAAAAAACA GATCACTGTA CTGGATCTCC ATCCCGGCGC CGGTAAAACA AGGAGGATTC TGCCACAGAT
     AGGGTCGGCC TAAGCTTGGA CTCTACGACT CCTTTTTTGT CTAGTGACAT GACCTAGAGG TAGGGCCGCG GCCATTTTGT TCCTCCTAAG ACGGTGTCTA
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     ·I  K  E  A  I  N  R  R  L  R  T  A  V  L  A  P  T  R  V  V  A  A  E  M  A  E  A  L  R  G  L  P  I
4701 CATCAAAGAG GCCATAAACA GAAGACTGAG AACAGCCGTG CTAGCACCAA CCAGGGTTGT GGCTGCTGAG ATGGCTGAAG CACTGAGAGG ACTGCCCATC
     GTAGTTTCTC CGGTATTTGT CTTCTGACTC TTGTCGGCAC GATCGTGGTT GGTCCCAACA CCGACGACTC TACCGACTTC GTGACTCTCC TGACGGGTAG
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      R  Y  Q  T  S  A  V  P  R  E  H  N  G  N  E  I  V  D  V  M  C  H  A  T  L  T  H  R  L  M  S  P  H  R
     ·
4801 CGGTACCAGA CATCCGCAGT GCCCAGAGAA CATAATGGAA ATGAGATTGT TGATGTCATG TGTCATGCTA CCCTCACCCA CAGGCTGATG TCTCCTCACA
     GCCATGGTCT GTAGGCGTCA CGGGTCTCTT GTATTACCTT TACTCTAACA ACTACAGTAC ACAGTACGAT GGGAGTGGGT GTCCGACTAC AGAGGAGTGT
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · V  P  N  Y  N  L  F  V  M  D  E  A  H  F  T  D  P  A  S  I  A  A  R  G  Y  I  S  T  K  V  E  L  G ·
4901 GGGTGCCGAA CTACAACCTG TTCGTGATGG ATGAGGCTCA TTTCACCGAC CCAGCTAGCA TTGCAGCAAG AGGTTACATT TCCACAAAGG TCGAGCTAGG
     CCCACGGCTT GATGTTGGAC AAGCACTACC TACTCCGAGT AAAGTGGCTG GGTCGATCGT AACGTCGTTC TCCAATGTAA AGGTGTTTCC AGCTCGATCC
                                                 NS3
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · E  A  A  A  I  F  M  T  A  T  P  P  G  T  S  D  P  F  P  E  S  N  S  P  I  S  D  L  Q  T  E  I  P
5001 GGAGGCGGCG GCAATATTCA TGACAGCCAC CCCACCAGGC ACTTCAGATC CATTCCCAGA GTCCAATTCA CCAATTTCCG ACTTACAGAC TGAGATCCCG
     CCTCCGCCGC CGTTATAAGT ACTGTCGGTG GGGTGGTCCG TGAAGTCTAG GTAAGGGTCT CAGGTTAAGT GGTTAAAGGC TGAATGTCTG ACTCTAGGGC
                                                 NS3
```

```
                                                                                                                    
     D  R  A  W  N  S  G  Y  E  W  I  T  E  Y  T  G  K  T  V  W  F  V  P  S  V  K  M  G  N  E  I  A  L  C
                                                                                                                    
5101 GATCGAGCTT GGAACTCTGG ATACGAATGG ATCACAGAAT ACACCGGGAA GACGGTTTGG TTTGTGCCTA GTGTTAAGAT GGGGAATGAG ATTGCCCTTT
     CTAGCTCGAA CCTTGAGACC TATGCTTACC TAGTGTCTTA TGTGGCCCTT CTGCCAAACC AAACACGGAT CACAATTCTA CCCCTTACTC TAACGGGAAA
                                                NS3                                                                 
                                                                                                                    
     ·  L  Q  R  A  G  K  K  V  V  Q  L  N  R  K  S  Y  E  T  E  Y  P  K  C  K  N  D  D  W  D  F  V  I  T  ·
5201 GCCTACAACG TGCTGGAAAG AAAGTAGTCC AATTGAACAG AAAGTCGTAC GAGACGGAGT ACCCAAAATG TAAGAACGAT GATTGGGACT TTGTTATCAC
     CGGATGTTGC ACGACCTTTC TTTCATCAGG TTAACTTGTC TTTCAGCATG CTCTGCCTCA TGGGTTTTAC ATTCTTGCTA CTAACCCTGA AACAATAGTG
                                                NS3                                                                 
                                                                                                                    
     ·  T  D  I  S  E  M  G  A  N  F  K  A  S  R  V  I  D  S  R  K  S  V  K  P  T  I  I  T  E  G  E  G  R
5301 AACAGACATA TCTGAAATGG GGGCTAACTT CAAGGCGAGC AGGGTGATTG ACAGCCGGAA GAGTGTGAAA CCAACCATCA TAACAGAAGG AGAAGGGAGA
     TTGTCTGTAT AGACTTTACC CCCGATTGAA GTTCCGCTCG TCCCACTAAC TGTCGGCCTT CTCACACTTT GGTTGGTAGT ATTGTCTTCC TCTTCCCTCT
                                                NS3                                                                 
                                                                                                                    
     V  I  L  G  E  P  S  A  V  T  A  A  S  A  A  Q  R  R  G  R  I  G  R  N  P  S  Q  V  G  D  E  Y  C  Y
                                                                                                                    
5401 GTGATCCTGG GAGAACCATC TGCAGTGACA GCAGCTAGTG CCGCCCAGAG ACGTGGACGT ATCGGTAGAA ATCCGTCGCA AGTTGGTGAT GAGTACTGTT
     CACTAGGACC CTCTTGGTAG ACGTCACTGT CGTCGATCAC GGCGGGTCTC TGCACCTGCA TAGCCATCTT TAGGCAGCGT TCAACCACTA CTCATGACAA
                                                NS3                                                                 
                                                                                                                    
     ·  G  G  H  T  N  E  D  D  S  N  F  A  H  W  T  E  A  R  I  M  L  D  N  I  N  M  P  N  G  L  I  A  Q  ·
5501 ATGGGGGGCA CACGAATGAA GACGACTCGA ACTTCGCCCA TTGGACTGAG GCACGAATCA TGCTGGACAA CATCAACATG CCAAACGGAC TGATCGCTCA
     TACCCCCCGT GTGCTTACTT CTGCTGAGCT TGAAGCGGGT AACCTGACTC CGTGCTTAGT ACGACCTGTT GTAGTTGTAC GGTTTGCCTG ACTAGCGAGT
                                                NS3                                                                 
                                                                                                                    
     ·  F  Y  Q  P  E  R  E  K  V  Y  T  M  D  G  E  Y  R  L  R  G  E  E  R  K  N  F  L  E  L  L  R  T  A
5601 ATTCTACCAA CCAGAGCGTG AGAAGGTATA TACCATGGAT GGGGAATACC GGCTCAGAGG AGAAGAGAGA AAAAACTTTC TGGAACTGTT GAGGACTGCA
     TAAGATGGTT GGTCTCGCAC TCTTCCATAT ATGGTACCTA CCCCTTATGG CCGAGTCTCC TCTTCTCTCT TTTTTGAAAG ACCTTGACAA CTCCTGACGT
                                                NS3                                                                 
                                                                                                                    
     D  L  P  V  W  L  A  Y  K  V  A  A  A  G  V  S  Y  H  D  R  R  W  C  F  D  G  P  R  T  N  T  I  L  E
                                                                                                                    
5701 GATCTGCCAG TTTGGCTGGC TTACAAGGTT GCAGCGGCTG GAGTGTCATA CCACGACCGG AGGTGGTGCT TTGATGGTCC TAGGACAAAC ACAATTTTAG
     CTAGACGGTC AAACCGACCG AATGTTCCAA CGTCGCCGAC CTCACAGTAT GGTGCTGGCC TCCACCACGA AACTACCAGG ATCCTGTTTG TGTTAAAATC
                                                NS3                                                                 
                                                                                                                    
     ·  D  N  N  E  V  E  V  I  T  K  L  G  E  R  K  I  L  R  P  R  W  I  D  A  R  V  Y  S  D  H  Q  A  L  ·
5801 AAGACAACAA CGAAGTGGAA GTCATCACGA AGCTTGGTGA AAGGAAGATT CTGAGGCCGC GCTGGATTGA CGCCAGGGTG TACTCGGATC ACCAGGCACT
     TTCTGTTGTT GCTTCACCTT CAGTAGTGCT TCGAACCACT TTCCTTCTAA GACTCCGGCG CGACCTAACT GCGGTCCCAC ATGAGCCTAG TGGTCCGTGA
```

```
                                                
```

```
                                                                                                                    
     V  Q  A  S  A  L  F  T  L  A  R  G  F  P  F  V  D  V  G  V  S  A  L  L  L  A  A  G  C  W  G  Q  V  T
     .
6601 GTTCAGGCAA GTGCACTATT CACACTCGCG CGAGGCTTCC CCTTCGTCGA TGTTGGAGTG TCGGCTCTCC TGCTAGCAGC CGGATGCTGG GGACAAGTCA
     CAAGTCCGTT CACGTGATAA GTGTGAGCGC GCTCCGAAGG GGAAGCAGCT ACAACCTCAC AGCCGAGAGG ACGATCGTCG GCCTACGACC CCTGTTCAGT
                                                 NS4B
                                                                                                                    
     ·  L  T  V  T  V  T  A  A  T  L  L  F  C  H  Y  A  Y  M  V  P  G  W  Q  A  E  A  M  R  S  A  Q  R  R  ·
6701 CCCTCACCGT TACGGTAACA GCGGCAACAC TCCTTTTTTG CCACTATGCC TACATGGTTC CCGGTTGGCA AGCTGAGGCA ATGCGCTCAG CCCAGCGGCG
     GGGAGTGGCA ATGCCATTGT CGCCGTTGTG AGGAAAAAAC GGTGATACGG ATGTACCAAG GGCCAACCGT TCGACTCCGT TACGCGAGTC GGGTCGCCGC
                                                 NS4B
                                                                                                                    
     ·  T  A  A  G  I  M  K  N  A  V  V  D  G  I  V  A  T  D  V  P  E  L  E  R  T  T  P  I  M  Q  K  K  I
6801 GACAGCGGCC GGAATCATGA AGAACGCTGT AGTGGATGGC ATCGTGGCCA CGGACGTCCC AGAATTAGAG CGCACCACAC CCATCATGCA GAAGAAAATT
     CTGTCGCCGG CCTTAGTACT TCTTGCGACA TCACCTACCG TAGCACCGGT GCCTGCAGGG TCTTAATCTC GCGTGGTGTG GGTAGTACGT CTTCTTTTAA
                                                 NS4B
                                                                                                                    
     G  Q  I  M  L  I  L  V  S  L  A  A  V  V  V  N  P  S  V  K  T  V  R  E  A  G  I  L  I  T  A  A  A  V
     .
6901 GGACAGATCA TGCTGATCTT GGTGTCTCTA GCTGCAGTAG TAGTGAACCC GTCTGTGAAG ACAGTACGAG AAGCCGGAAT TTTGATCACG GCCGCAGCGG
     CCTGTCTAGT ACGACTAGAA CCACAGAGAT CGACGTCATC ATCACTTGGG CAGACACTTC TGTCATGCTC TTCGGCCTTA AAACTAGTGC CGGCGTCGCC
                                                 NS4B
                                                                                                                    
     ·  T  L  W  E  N  G  A  S  S  V  W  N  A  T  T  A  I  G  L  C  H  I  M  R  G  G  W  L  S  C  L  S  I  ·
7001 TGACGCTTTG GGAGAATGGA GCAAGCTCTG TTTGGAACGC AACAACTGCC ATCGGACTCT GCCACATCAT GCGTGGGGGT TGGTTGTCAT GTCTATCCAT
     ACTGCGAAAC CCTCTTACCT CGTTCGAGAC AAACCTTGCG TTGTTGACGG TAGCCTGAGA CGGTGTAGTA CGCACCCCCA ACCAACAGTA CAGATAGGTA
                                                                                 NS5
                                                                                                                    
                                 NS4B
                                                                                                                    
     ·  T  W  T  L  I  K  N  M  E  K  P  G  L  K  R  G  G  A  K  G  R  T  L  G  E  V  W  K  E  R  L  N  Q
7101 AACATGGACA CTCATAAAGA ACATGGAAAA ACCAGGACTA AAAAGAGGTG GGGCAAAAGG ACGCACCTTG GGAGAGGTTT GGAAAGAAAG ACTCAACCAG
     TTGTACCTGT GAGTATTTCT TGTACCTTTT TGGTCCTGAT TTTTCTCCAC CCCGTTTTCC TGCGTGGAAC CCTCTCCAAA CCTTTCTTTC TGAGTTGGTC
                                                 NS5
                                                                                                                    
     M  T  K  E  E  F  T  R  Y  R  K  E  A  I  I  E  V  D  R  S  A  A  K  H  A  R  K  E  G  N  V  T  G  G
     .
7201 ATGACAAAAG AAGAGTTCAC TAGGTACCGC AAAGAGGCCA TCATCGAAGT CGATCGCTCA GCGGCAAAAC ACGCCAGGAA AGAAGGCAAT GTCACTGGAG
     TACTGTTTTC TTCTCAAGTG ATCCATGGCG TTTCTCCGGT AGTAGCTTCA GCTAGCGAGT CGCCGTTTTG TGCGGTCCTT TCTTCCGTTA CAGTGACCTC
                                                 NS5
                                                                                                                    
     ·  H  P  V  S  R  G  T  A  K  L  R  W  L  V  E  R  R  F  L  E  P  V  G  K  V  I  D  L  G  C  G  R  G  ·
```

```
7301 GGCATCCAGT CTCTAGGGGC ACAGCAAAAC TGAGATGGCT GGTCGAACGG AGGTTTCTCG AACCGGTCGG AAAAGTGATT GACCTTGGAT GTGGAAGAGG
     CCGTAGGTCA GAGATCCCCG TGTCGTTTTG ACTCTACCGA CCAGCTTGCC TCCAAAGAGC TTGGCCAGCC TTTTCACTAA CTGGAACCTA CACCTTCTCC
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · G  W  C  Y  Y  M  A  T  Q  K  R  V  Q  E  V  R  G  Y  T  K  G  G  P  G  H  E  E  P  Q  L  V  Q  S
7401 CGGTTGGTGT TACTATATGG CAACCCAAAA AAGAGTCCAA GAAGTCAGAG GGTACACAAA GGGCGGTCCC GGACATGAAG AGCCCCAACT AGTGCAAAGT
     GCCAACCACA ATGATATACC GTTGGGTTTT TTCTCAGGTT CTTCAGTCTC CCATGTGTTT CCCGCCAGGG CCTGTACTTC TCGGGGTTGA TCACGTTTCA
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      Y  G  W  N  I  V  T  M  K  S  G  V  D  V  F  Y  R  P  S  E  C  C  D  T  L  L  C  D  I  G  E  S  S  S
     .
7501 TATGGATGGA ACATTGTCAC CATGAAGAGT GGAGTGGATG TGTTCTACAG ACCTTCTGAG TGTTGTGACA CCCTCCTTTG TGACATCGGA GAGTCCTCGT
     ATACCTACCT TGTAACAGTG GTACTTCTCA CCTCACCTAC ACAAGATGTC TGGAAGACTC ACAACACTGT GGGAGGAAAC ACTGTAGCCT CTCAGGAGCA
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · S  A  E  V  E  E  H  R  T  I  R  V  L  E  M  V  E  D  W  L  H  R  G  P  R  E  F  C  V  K  V  L  C ·
7601 CAAGTGCTGA GGTTGAAGAG CATAGGACGA TTCGGGTCCT TGAAATGGTT GAGGACTGGC TGCACCGAGG GCCAAGGGAA TTTTGCGTGA AGGTGCTCTG
     GTTCACGACT CCAACTTCTC GTATCCTGCT AAGCCCAGGA ACTTTACCAA CTCCTGACCG ACGTGGCTCC CGGTTCCCTT AAAACGCACT TCCACGAGAC
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · P  Y  M  P  K  V  I  E  K  M  E  L  L  Q  R  R  Y  G  G  G  L  V  R  N  P  L  S  R  N  S  T  H  E
7701 CCCCTACATG CCGAAAGTCA TAGAGAAGAT GGAGCTGCTC CAACGCCGGT ATGGGGGGGG ACTGGTCAGA AACCCACTCT CACGGAATTC CACGCACGAG
     GGGGATGTAC GGCTTTCAGT ATCTCTTCTA CCTCGACGAG GTTGCGGCCA TACCCCCCCC TGACCAGTCT TTGGGTGAGA GTGCCTTAAG GTGCGTGCTC
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
      M  Y  W  V  S  R  A  S  G  N  V  V  H  S  V  N  M  T  S  Q  V  L  L  G  R  M  E  K  R  T  W  K  G  P
     .
7801 ATGTATTGGG TGAGTCGAGC TTCAGGCAAT GTGGTACATT CAGTGAATAT GACCAGCCAG GTGCTCCTAG GAAGAATGGA AAAAAGGACC TGGAAGGGAC
     TACATAACCC ACTCAGCTCG AAGTCCGTTA CACCATGTAA GTCACTTATA CTGGTCGGTC CACGAGGATC CTTCTTACCT TTTTTCCTGG ACCTTCCCTG
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · Q  Y  E  E  D  V  N  L  G  S  G  T  R  A  V  G  K  P  L  L  N  S  D  T  S  K  I  K  N  R  I  E  R ·
7901 CCCAATACGA GGAAGACGTA AACTTGGGAA GTGGAACCAG GGCGGTGGGA AAACCCCTGC TCAACTCAGA CACCAGTAAA ATCAAGAACA GGATTGAACG
     GGGTTATGCT CCTTCTGCAT TTGAACCCTT CACCTTGGTC CCGCCACCCT TTTGGGGACG AGTTGAGTCT GTGGTCATTT TAGTTCTTGT CCTAACTTGC
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     · L  R  R  E  Y  S  S  T  W  H  H  D  E  N  H  P  Y  R  T  W  N  Y  H  G  S  Y  D  V  K  P  T  G  S
8001 ACTCAGGCGT GAGTACAGTT CGACGTGGCA CCACGATGAG AACCACCCAT ATAGAACCTG GAACTATCAT GGCAGTTATG ATGTGAAGCC CACAGGCTCC
     TGAGTCCGCA CTCATGTCAA GCTGCACCGT GGTGCTACTC TTGGTGGGTA TATCTTGGAC CTTGATAGTA CCGTCAATAC TACACTTCGG GTGTCCGAGG
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

```
     A  S  S  L  V  N  G  V  V  R  L  L  S  K  P  W  D  T  I  T  N  V  T  T  M  A  M  T  D  T  T  P  F  G
     .
8101 GCCAGTTCGC TGGTCAATGG AGTGGTCAGG CTCCTCTCAA AACCATGGGA CACCATCACG AATGTTACCA CCATGGCCAT GACTGACACT ACTCCCTTCG
     CGGTCAAGCG ACCAGTTACC TCACCAGTCC GAGGAGAGTT TTGGTACCCT GTGGTAGTGC TTACAATGGT GGTACCGGTA CTGACTGTGA TGAGGGAAGC
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     .  Q  Q  R  V  F  K  E  K  V  D  T  K  A  P  E  P  P  E  G  V  K  Y  V  L  N  E  T  T  N  W  L  W  A  .
8201 GGCAGCAGCG AGTGTTCAAA GAGAAGGTGG ACACGAAAGC TCCTGAACCG CCAGAAGGAG TGAAGTACGT GCTCAACGAG ACCACCAACT GGTTGTGGGC
     CCGTCGTCGC TCACAAGTTT CTCTTCCACC TGTGCTTTCG AGGACTTGGC GGTCTTCCTC ACTTCATGCA CGAGTTGCTC TGGTGGTTGA CCAACACCCG
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     .  F  L  A  R  E  K  R  P  R  M  C  S  R  E  E  F  I  R  K  V  N  S  N  A  A  L  G  A  M  F  E  E  Q
8301 GTTTTTGGCC AGAGAAAAAC GTCCCAGAAT GTGCTCTCGA GAGGAATTCA TAAGAAAGGT CAACAGCAAT GCAGCTTTGG GTGCCATGTT TGAAGAGCAG
     CAAAAACCGG TCTCTTTTTG CAGGGTCTTA CACGAGAGCT CTCCTTAAGT ATTCTTTCCA GTTGTCGTTA CGTCGAAACC CACGGTACAA ACTTCTCGTC
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     N  Q  W  R  S  A  R  E  A  V  E  D  P  K  F  W  E  M  V  D  E  E  R  E  A  H  L  R  G  E  C  H  T  C
     .
8401 AATCAATGGA GGAGCGCCAG AGAAGCAGTT GAAGATCCAA AATTTTGGGA AATGGTGGAT GAGGAGCGCG AGGCACATCT GCGGGGGGAA TGTCACACTT
     TTAGTTACCT CCTCGCGGTC TCTTCGTCAA CTTCTAGGTT TTAAAACCCT TTACCACCTA CTCCTCGCGC TCCGTGTAGA CGCCCCCCTT ACAGTGTGAA
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     .  I  Y  N  M  M  G  K  R  E  K  K  P  G  E  F  G  K  A  K  G  S  R  A  I  W  F  M  W  L  G  A  R  F  .
8501 GCATTTACAA CATGATGGGA AAGAGAGAGA AAAAACCCGG AGAGTTCGGA AAGGCCAAGG GAAGCAGAGC CATTTGGTTC ATGTGGCTCG GAGCTCGCTT
     CGTAAATGTT GTACTACCCT TTCTCTCTCT TTTTTGGGCC TCTCAAGCCT TTCCGGTTCC CTTCGTCTCG GTAAACCAAG TACACCGAGC CTCGAGCGAA
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     .  L  E  F  E  A  L  G  F  L  N  E  D  H  W  L  G  R  K  N  S  G  G  G  V  E  G  L  G  L  Q  K  L  G
8601 TCTGGAGTTC GAGGCTCTGG GTTTTCTCAA TGAAGACCAC TGGCTTGGAA GAAAGAACTC AGGAGGAGGT GTCGAGGGCT TGGGCCTCCA AAAACTGGGT
     AGACCTCAAG CTCCGAGACC CAAAAGAGTT ACTTCTGGTG ACCGAACCTT CTTTCTTGAG TCCTCCTCCA CAGCTCCCGA ACCCGGAGGT TTTTGACCCA
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     Y  I  L  R  E  V  G  I  R  P  G  G  K  I  Y  A  D  D  T  A  G  W  D  T  R  I  T  R  A  D  L  E  N  E
     .
8701 TACATCCTGC GTGAAGTTGG CATCCGGCCT GGGGGCAAGA TCTATGCTGA TGACACAGCT GGCTGGGACA CCCGCATCAC GAGAGCTGAC TTGGAAAATG
     ATGTAGGACG CACTTCAACC GTAGGCCGGA CCCCCGTTCT AGATACGACT ACTGTGTCGA CCGACCCTGT GGGCGTAGTG CTCTCGACTG AACCTTTTAC
                                                 NS5
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     .  A  K  V  L  E  L  L  D  G  E  H  R  R  L  A  R  A  I  I  E  L  T  Y  R  H  K  V  V  K  V  M  R  P  .
8801 AAGCTAAGGT GCTTGAGCTG CTTGATGGGG AACATCGGCG TCTTGCCAGG GCCATCATTG AGCTCACCTA TCGTCACAAA GTTGTGAAAG TGATGCGCCC
     TTCGATTCCA CGAACTCGAC GAACTACCCC TTGTAGCCGC AGAACGGTCC CGGTAGTAAC TCGAGTGGAT AGCAGTGTTT CAACACTTTC ACTACGCGGG
                                                 NS5
```

```
                                                                                                                    
     ·A A D  G R T V  M D V  I S R  E D Q R  G S G  Q V V  T Y A L  N T F  T N L
8901 GGCTGCTGAT GGAAGAACCG TTATGGATGT TATCTCCAGA GAAGATCAGA GGGGGAGTGG ACAAGTTGTC ACCTACGCCC TAAACACTTT CACCAACCTG
     CCGACGACTA CCTTCTTGGC AATACCTACA ATAGAGGTCT CTTCTAGTCT CCCCCTCACC TGTTCAACAG TGGATGCGGG ATTTGTGAAA GTGGTTGGAC
                                                NS5
                                                                                                                    
      A V Q L  V R M  M E G  E G V I  G P D  D V E  K L T K  G K G  P K V  R T W L
     ·
9001 GCTGTCCAGC TGGTGAGGAT GATGGAAGGG GAAGGAGTGA TTGGCCCAGA TGATGTGGAG AAACTCACAA AAGGGAAAGG ACCCAAAGTC AGGACCTGGC
     CGACAGGTCG ACCACTCCTA CTACCTTCCC CTTCCTCACT AACCGGGTCT ACTACACCTC TTTGAGTGTT TTCCCTTTCC TGGGTTTCAG TCCTGGACCG
                                                NS5
                                                                                                                    
     · F E N  G E E  R L S R  M A V  S G D  D C V V  K P L  D D R  F A T S  L H F ·
9101 TGTTTGAGAA TGGGGAAGAA AGACTCAGCC GCATGGCTGT CAGTGGAGAT GACTGTGTGG TAAAGCCCCT GGACGATCGC TTTGCCACCT CGCTCCACTT
     ACAAACTCTT ACCCCTTCTT TCTGAGTCGG CGTACCGACA GTCACCTCTA CTGACACACC ATTTCGGGGA CCTGCTAGCG AAACGGTGGA GCGAGGTGAA
                                                NS5
                                                                                                                    
     ·L N A  M S K V  R K D  I Q E  W K P S  T G W  Y D W  Q Q V P  F C S  N H F
9201 CCTCAATGCT ATGTCAAAGG TTCGCAAAGA CATCCAAGAG TGGAAACCGT CAACTGGATG GTATGATTGG CAGCAGGTTC CATTTTGCTC AAACCATTTC
     GGAGTTACGA TACAGTTTCC AAGCGTTTCT GTAGGTTCTC ACCTTTGGCA GTTGACCTAC CATACTAACC GTCGTCCAAG GTAAAACGAG TTTGGTAAAG
                                                NS5
                                                                                                                    
      T E L I  M K D  G R T  L V V P  C R G  Q D E  L V G R  A R I  S P G  A G W N
     ·
9301 ACTGAATTGA TCATGAAAGA TGGAAGAACA CTGGTGGTTC CATGCCGAGG ACAGGATGAA TTGGTAGGCA GAGCTCGCAT ATCTCCAGGG GCCGGATGGA
     TGACTTAACT AGTACTTTCT ACCTTCTTGT GACCACCAAG GTACGGCTCC TGTCCTACTT AACCATCCGT CTCGAGCGTA TAGAGGTCCC CGGCCTACCT
                                                NS5
                                                                                                                    
     · V R D  T A C  L A K S  Y A Q  M W L  L L Y F  H R R  D L R  L M A N  A I C ·
9401 ACGTCCGCGA CACTGCTTGT CTGGCTAAGT CTTATGCCCA GATGTGGCTG CTTCTGTACT TCCACAGAAG AGACCTGCGG CTCATGGCCA ACGCCATTTG
     TGCAGGCGCT GTGACGAACA GACCGATTCA GAATACGGGT CTACACCGAC GAAGACATGA AGGTGTCTTC TCTGGACGCC GAGTACCGGT TGCGGTAAAC
                                                NS5
                                                                                                                    
     ·S A V  P V N W  V P T  G R T  T W S I  H A G  G E W  M T T E  D M L  E V W
9501 CTCCGCTGTC CCTGTGAATT GGGTCCCTAC CGGAAGAACC ACGTGGTCCA TCCATGCAGG AGGAGAGTGG ATGACAACAG AGGACATGTT GGAGGTCTGG
     GAGGCGACAG GGACACTTAA CCCAGGGATG GCCTTCTTGG TGCACCAGGT AGGTACGTCC TCCTCTCACC TACTGTTGTC TCCTGTACAA CCTCCAGACC
                                                NS5
                                                                                                                    
      N R V W  I E E  N E W  M E D K  T P V  E K W  S D V P  Y S G  K R E  D I W C
     ·
9601 AACCGTGTTT GGATAGAGGA GAATGAATGG ATGGAAGACA AAACCCCAGT GGAGAAATGG AGTGACGTCC CATATTCAGG AAAACGAGAG GACATCTGGT
     TTGGCACAAA CCTATCTCCT CTTACTTACC TACCTTCTGT TTTGGGGTCA CCTCTTTACC TCACTGCAGG GTATAAGTCC TTTTGCTCTC CTGTAGACCA
```

```
                                                
```

Sequence Appendix 7. Sequence of RSV G

```
          10          19          28          37          46          55
TGCAAAC ATG TCC AAA AAC AAG GAC CAA CGC ACC GCT AAG ACA CTA GAA AAG ACC
        MET Ser Lys Asn Lys Asp Gln Arg Thr Ala Lys Thr Leu Glu Lys Thr

          64          73          82          91         100         109
TGG GAC ACT CTC AAT CAT TTA TTA TTC ATA TCA TCG GGC TTA TAT AAG TTA AAT
Trp Asp Thr Leu Asn His Leu Leu Phe Ile Ser Ser Gly Leu Tyr Lys Leu Asn

         118         127         136         145         154         163
CTT AAA TCT GTA GCA CAA ATC ACA TTA TCC ATT CTG GCA ATG ATA ATC TCA ACT
Leu Lys Ser Val Ala Gln Ile Thr Leu Ser Ile Leu Ala MET Ile Ile Ser Thr

         172         181         190         199         208         217
TCA CTT ATA ATT ACA GCC ATC ATA TTC ATA GCC TCG GCA AAC CAC AAA GTC ACA
Ser Leu Ile Ile Thr Ala Ile Ile Phe Ile Ala Ser Ala Asn His Lys Val Thr

         226         235         244         253         262         271
CTA ACA ACT GCA ATC ATA CAA GAT GCA ACA AGC CAG ATC AAG AAC ACA ACC CCA
Leu Thr Thr Ala Ile Ile Gln Asp Ala Thr Ser Gln Ile Lys Asn Thr Thr Pro

         280         289         298         307         316         325
ACA TAC CTC ACT CAG GAT CCT CAG CTT GGA ATC AGC TTC TCC AAT CTG TCT GAA
Thr Tyr Leu Thr Gln Asp Pro Gln Leu Gly Ile Ser Phe Ser Asn Leu Ser Glu

         334         343         352         361         370         379
ATT ACA TCA CAA ACC ACC ACC ATA CTA GCT TCA ACA ACA CCA GGA GTC AAG TCA
Ile Thr Ser Gln Thr Thr Thr Ile Leu Ala Ser Thr Thr Pro Gly Val Lys Ser

         388         397         406         415         424         433
AAC CTG CAA CCC ACA ACA GTC AAG ACT AAA AAC ACA ACA ACA ACC CAA ACA CAA
Asn Leu Gln Pro Thr Thr Val Lys Thr Lys Asn Thr Thr Thr Thr Gln Thr Gln

         442         451         460         469         478         487
CCC AGC AAG CCC ACT ACA AAA CAA CGC CAA AAC AAA CCA CCA AAC AAA CCC AAT
Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn Lys Pro Asn

         496         505         514         523         532         541
AAT GAT TTT CAC TTC GAA GTG TTT AAC TTT GTA CCC TGC AGC ATA TGC AGC AAC
Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys Ser Asn

         550         559         568         577         586         595
AAT CCA ACC TGC TGG GCT ATC TGC AAA AGA ATA CCA AAC AAA AAA CCA GGA AAG
Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro Asn Lys Lys Pro Gly Lys


         604         613         622         631         640         649
AAA ACC ACC ACC AAG CCT ACA AAA AAA CCA ACC TTC AAG ACA ACC AAA AAA GAT
Lys Thr Thr Thr Lys Pro Thr Lys Lys Pro Thr Phe Lys Thr Thr Lys Lys Asp

         658         667         676         685         694         703
CTC AAA CCT CAA ACC ACT AAA CCA AAG GAA GTA CCC ACC ACC AAG CCC ACA GAA
Leu Lys Pro Gln Thr Thr Lys Pro Lys Glu Val Pro Thr Thr Lys Pro Thr Glu

         712         721         730         739         748         757
GAG CCA ACC ATC AAC ACC ACC AAA ACA AAC ATC ACA ACT ACA CTG CTC ACC AAC
Glu Pro Thr Ile Asn Thr Thr Lys Thr Asn Ile Thr Thr Thr Leu Leu Thr Asn


         766         775         784         793         802         811
AAC ACC ACA GGA AAT CCA AAA CTC ACA AGT CAA ATG GAA ACC TTC CAC TCA ACC
Asn Thr Thr Gly Asn Pro Lys Leu Thr Ser Gln MET Glu Thr Phe His Ser Thr
```

```
        820         829         838         847         856         865
TCC TCC GAA GGC AAT CTA AGC CCT TCT CAA GTC TCC ACA ACA TCC GAG CAC CCA
Ser Ser Glu Gly Asn Leu Ser Pro Ser Gln Val Ser Thr Thr Ser Glu His Pro

        874         883         892         901             914
TCA CAA CCC TCA TCT CCA CCC AAC ACA ACA CGC CAG TAG TTATTAAAAA AAAAAA
Ser Gln Pro Ser Ser Pro Pro Asn Thr Thr Arg Gln  .
```

1. A replication-deficient pseudoinfectious flavivirus comprising a flavivirus genome comprising (i) one or more deletions or mutations in nucleotide sequences encoding one or more proteins selected from the group consisting of capsid (C), pre-membrane (prM), envelope (E), non-structural protein 1 (NS1), non-structural protein 3 (NS3), and non-structural protein 5 (NS5), and (ii) a sequence encoding a respiratory syncytial virus (RSV) peptide or protein, or a fragment or analog thereof.

2. The replication-deficient pseudoinfectious flavivirus of claim 1, wherein said respiratory syncytial virus (RSV) protein is the RSV F protein, or a fragment or analog thereof.

3. The replication-deficient pseudoinfectious flavivirus of claim 2, wherein said RSV F protein lacks a trans-membrane domain.

4. The replication-deficient pseudoinfectious flavivirus of claim 3, wherein said RSV F protein is truncated so that it is produced in secreted form.

5. The replication-deficient pseudoinfectious flavivirus of claim 1, wherein said respiratory syncytial virus (RSV) protein is the RSV G protein, or a fragment or analog thereof.

6. The replication-deficient pseudoinfectious flavivirus of any of claims 1-5, wherein said one or more deletions or mutations is within capsid (C) sequences of the flavivirus genome.

7. The replication-deficient pseudoinfectious flavivirus of any of claims 1-5, wherein said one or more deletions or mutations is within pre-membrane (prM) and/or envelope (E) sequences of the flavivirus genome.

8. The replication-deficient pseudoinfectious flavivirus of any of claims 1-5, wherein said one or more deletions or mutations is within capsid (C), pre-membrane (prM), and envelope (E) sequences of the flavivirus genome.

9. The replication-deficient pseudoinfectious flavivirus of any of claims 1-5, wherein said one or more deletions or mutations is within non-structural protein 1 (NS1) sequences of the flavivirus genome.

10. The replication-deficient pseudoinfectious flavivirus of any of claims 1-9, wherein said flavivirus genome comprises sequences encoding a pre-membrane (prM) and/or envelope (E) protein.

11. The replication-deficient pseudoinfectious flavivirus of any of claims 1-10, wherein the flavivirus genome is selected from that of yellow fever virus, West Nile virus, tick-borne encephalitis virus, Langat virus, Japanese encephalitis virus, dengue virus, and St. Louis encephalitis virus sequences, and chimeras thereof.

12. The replication-deficient pseudoinfectious flavivirus of claim 11, wherein said chimera comprises pre-membrane (prM) and envelope (E) sequences of a first flavivirus, and capsid (C) and non-structural sequences of a second, different flavivirus.

13. The replication-deficient pseudoinfectious flavivirus of any of claims 1-12, wherein said genome is packaged in a particle comprising pre-membrane (prM) and envelope (E) sequences from a flavivirus that is the same or different from that of the genome.

14. The replication-deficient pseudoinfectious flavivirus of any of claims 1-13, wherein sequences encoding said respiratory syncytial virus peptide or protein, or a fragment or analog thereof are inserted in the place of or in combination with the one or more deletions or mutations of the one or more proteins.

15. The replication-deficient pseudoinfectious flavivirus of any of claims 1-14, wherein sequences encoding said respiratory syncytial virus peptide or protein, or a fragment or analog thereof are inserted in the flavivirus genome within sequences encoding the envelope (E) protein, within sequences encoding the non-structural 1 (NS1) protein, within sequences encoding the pre-membrane (prM) protein, intergenically between sequences encoding the envelope (E) protein and non-structural protein 1 (NS1), intergenically between non-structural protein 2B (NS2B) and non-structural protein 3 (NS3), or as a bicistronic insertion in the 3' untranslated region of the flavivirus genome.

16. A composition comprising a first replication-deficient pseudoinfectious flavivirus of any of claims 1-15 and a second, different replication-deficient pseudoinfectious flavivirus comprising a genome comprising one or more deletions or mutations in nucleotide sequences encoding one or more proteins selected from the group consisting of capsid (C), pre-membrane (prM), envelope (E), non-structural protein 1 (NS1), non-structural protein 3 (NS3), and non-structural protein 5 (NS5), wherein the one or more proteins encoded by the sequences in which the one or more deletion(s) or mutation(s) occur in the second, different replication-deficient pseudoinfectious flavivirus are different from the one or more proteins encoded by the sequences in which the one or more deletion(s) or mutation(s) occur in the first replication-deficient pseudoinfectious flavivirus.

17. A method of inducing an immune response to respiratory syncytial virus (RSV) in a subject, the method comprising administering to the subject one or more replication-deficient pseudoinfectious flaviviruses of any of claims 1-15 and/or a composition of claim 16 to the subject.

18. The method of claim 17, wherein the subject is at risk of but does not have an infection by respiratory syncytial virus (RSV).

19. The method of claim 17, wherein the subject has an infection by respiratory syncytial virus (RSV).

20. The method of any of claims 17-19, wherein the subject is an infant, young child, or elderly person.

21. The method of any of claims 17-20, wherein the method is for inducing an immune response against a protein encoded by the flavivirus genome, in addition to respiratory syncytial virus.

22. The method of claim 21, wherein the subject is at risk of but does not have an infection by the flavivirus corresponding to the genome of the pseudoinfectious flavivirus, which comprises sequences encoding a flavivirus pre-membrane and/or envelope protein.

23. The method of claim 21, wherein the subject has an infection by the flavivirus corresponding to the genome of the pseudoinfectious flavivirus which comprises sequences encoding a flavivirus pre-membrane and/or envelope protein.

24. A pharmaceutical composition comprising a pseudoinfectious flavivirus of any of claims 1-15, or the composition of claim 16, and a pharmaceutically acceptable carrier or diluent.

25. The pharmaceutical composition of claim 24, further comprising an adjuvant.

26. A nucleic acid molecule corresponding to the genome of a pseudoinfectious flavivirus of any of claims 1-15 or the complement thereof.

27. A method of making a replication-deficient pseudoinfectious flavivirus of any of claims 1-15, the method comprising introducing a nucleic acid molecule of claim 26 into a cell that expresses the protein corresponding to any sequences deleted from the flavivirus genome of the replication-deficient pseudoinfectious flavivirus.

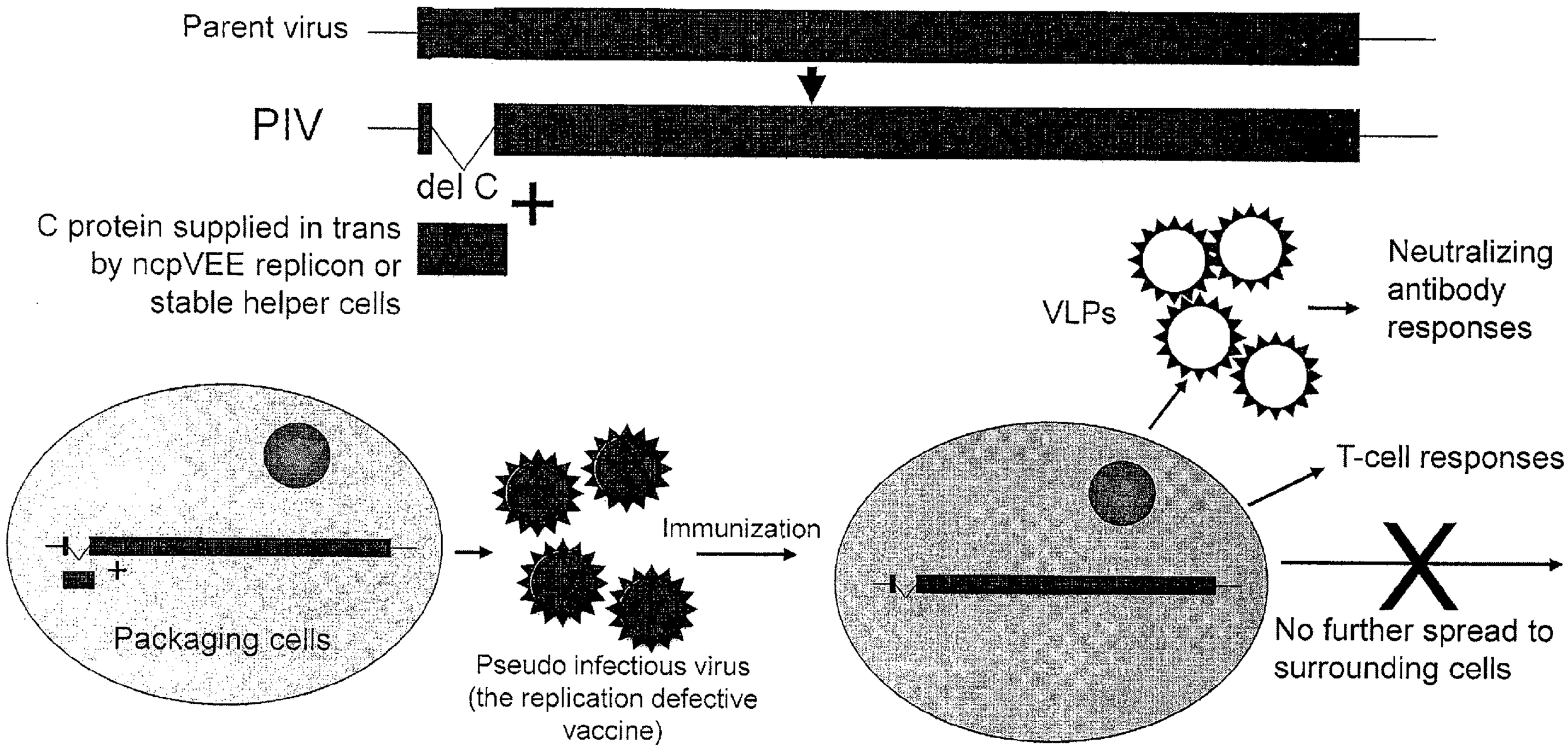
28. The method of claim 27, wherein the protein is expressed in the cell from the genome of a second, different, replication-deficient pseudoinfectious flavivirus.

29. The method of claim 27, wherein the protein is expressed from a replicon.

30. The method of claim 29, wherein the replicon is an alphavirus replicon.

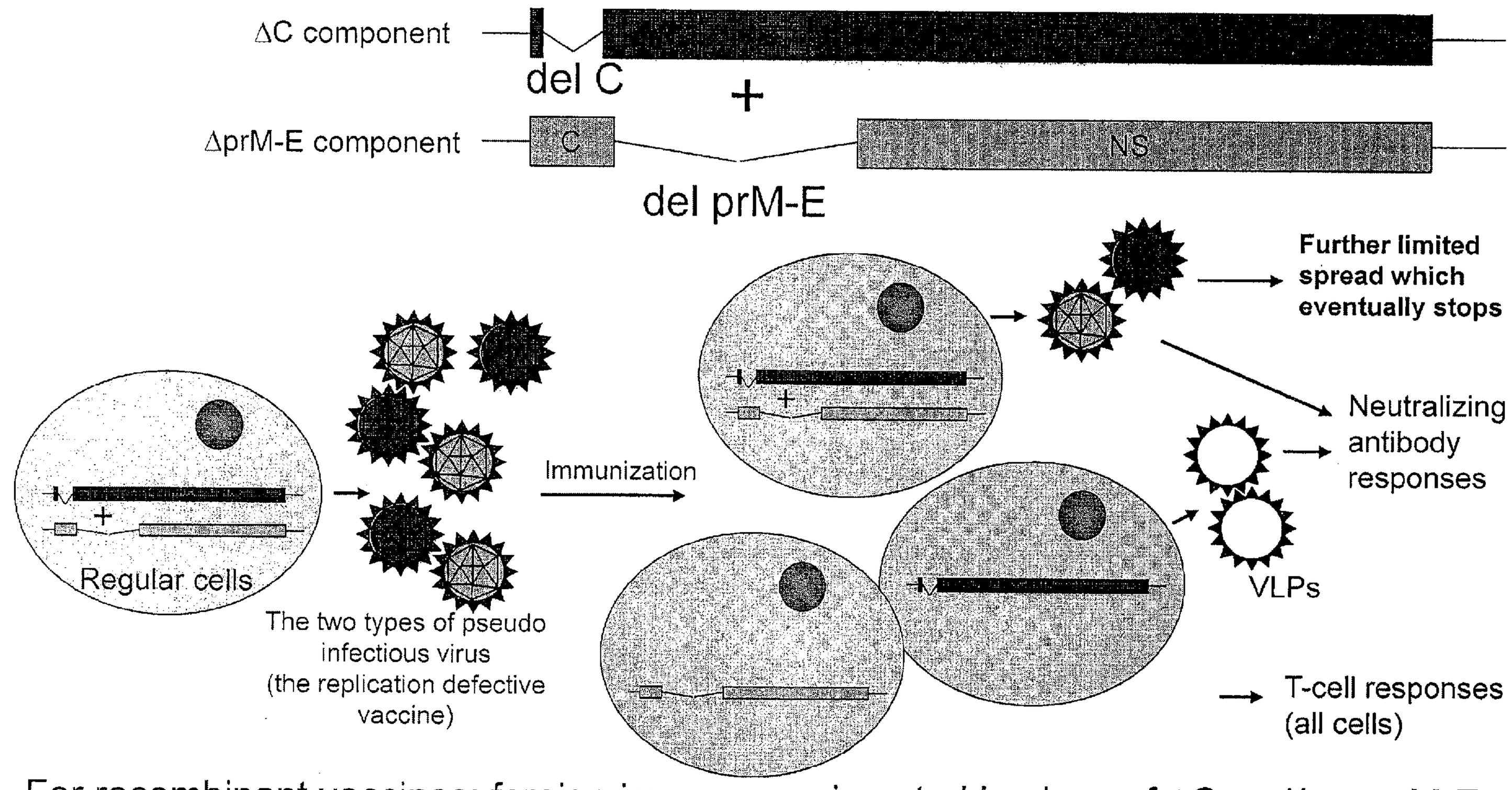
31. The method of claim 30, wherein the alphavirus is a Venezuelan Equine Encephalitis virus.

Fig. 1. Principle of single-component PIV (s-PIV; single-round replication in vivo)



For recombinant vaccines: foreign immunogen inserted in place of the ΔC deletion, or elsewhere, e.g. intergenically etc.

Fig. 2. Principle of two-component PIV (d-PIV; limited spread in vivo)



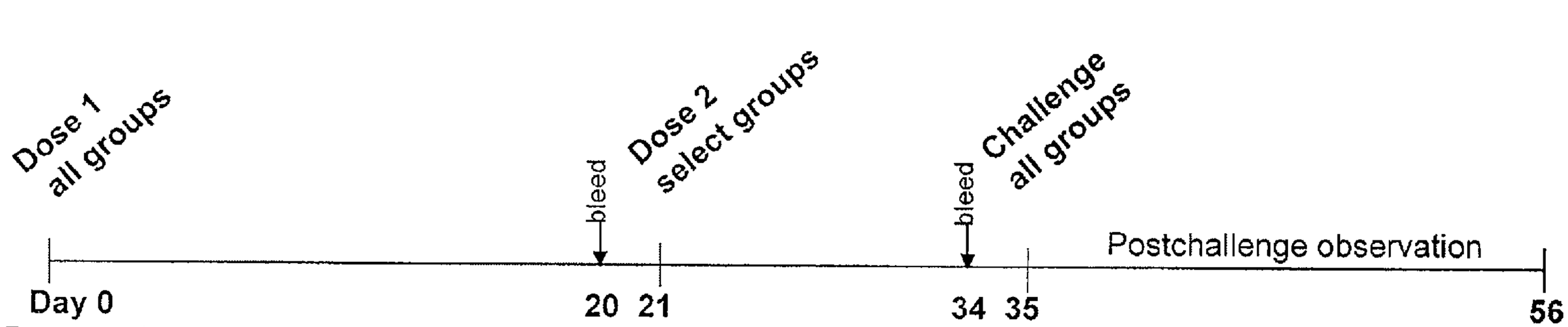
For recombinant vaccines: foreign immunogen inserted in place of ΔC and/or ΔprM-E, or elsewhere.

WO 2010/107847
CA 027552 2011-09-12
PCT/US2010/027552

Fig. 3. Immunogenicity/efficacy: general experiment design (mice)

WO 2010/107847 PCT/US2010/027552 CA 02755257 2011-09-12

Experiment design in mice



3.5 w.o. mice
IP route

ICR mice for WN and YF groups
C57/BL6 mice for JE groups

Fig. 4. PIV-WN induces uniform humoral immunity in mice

ICR mice, IP inoculation

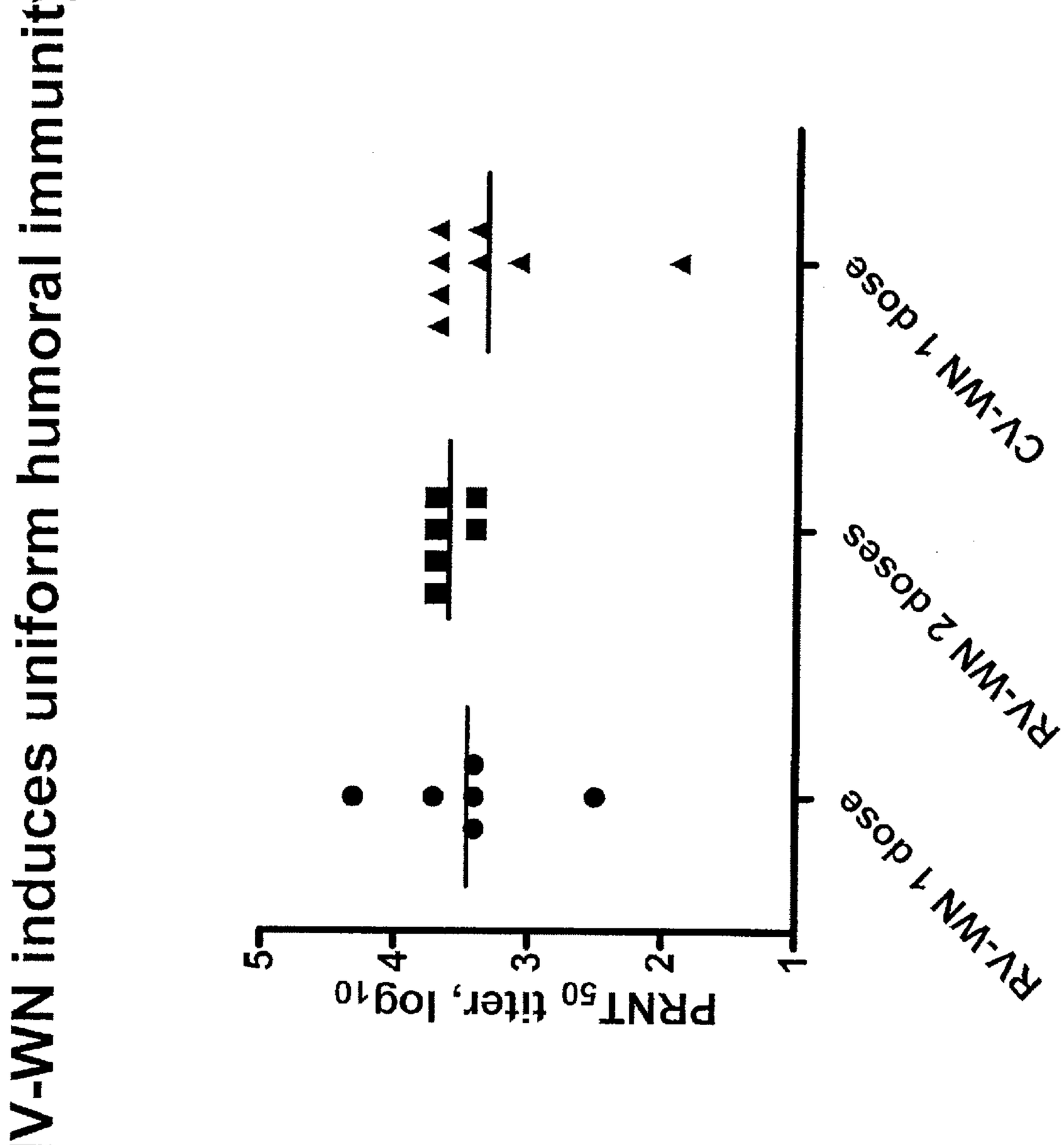
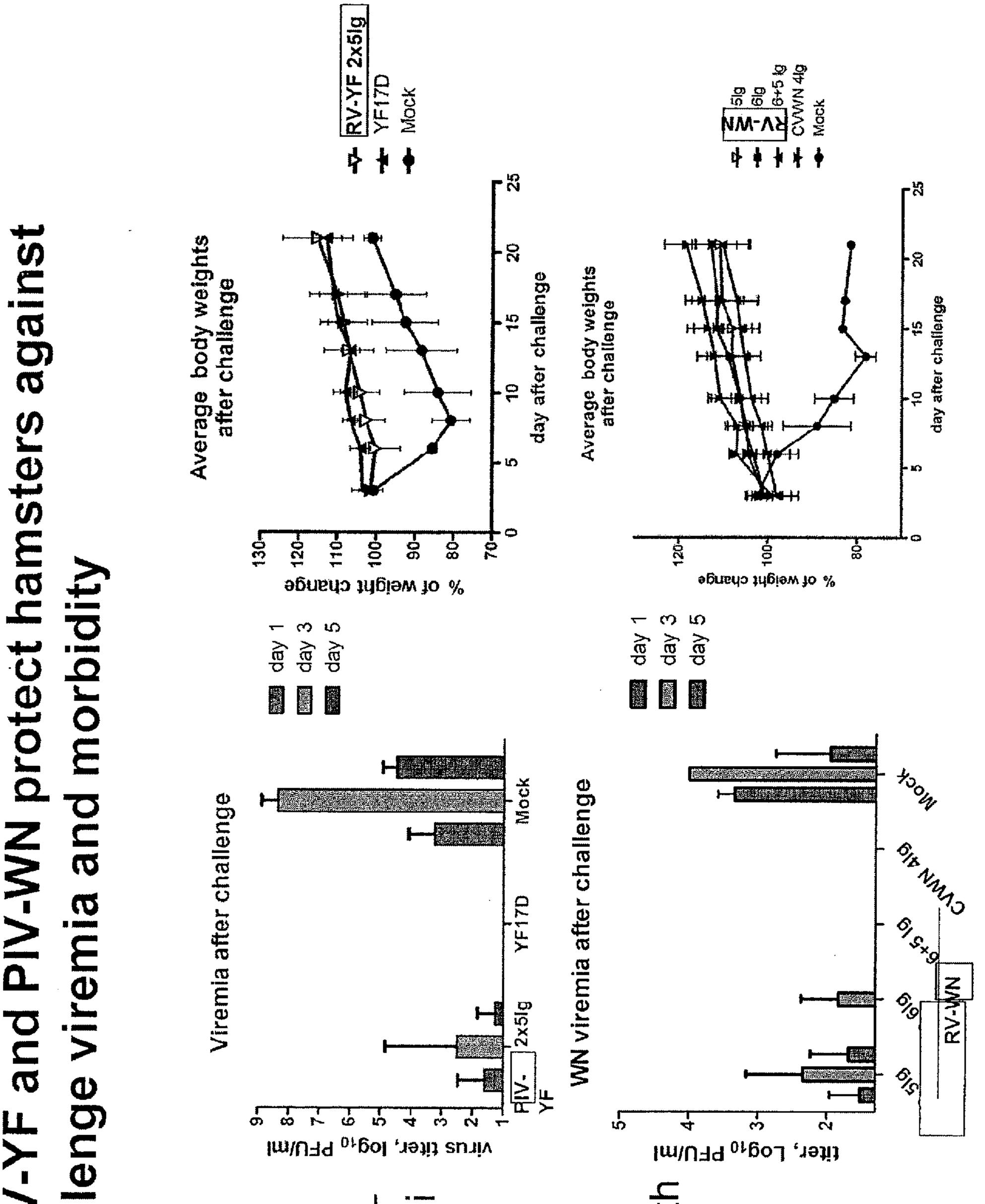


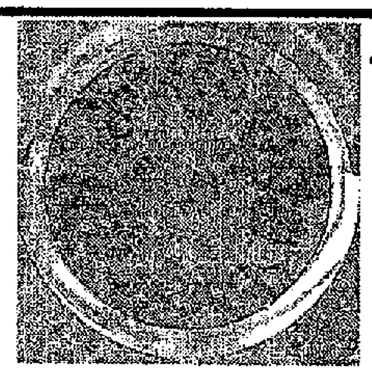
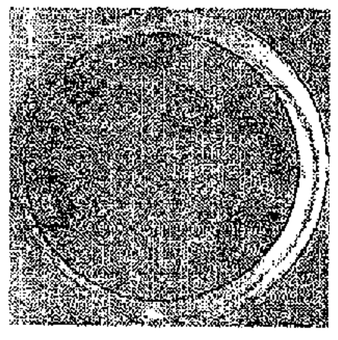
Fig. 5. PIV-YF and PIV-WN protect hamsters against post-challenge viremia and morbidity

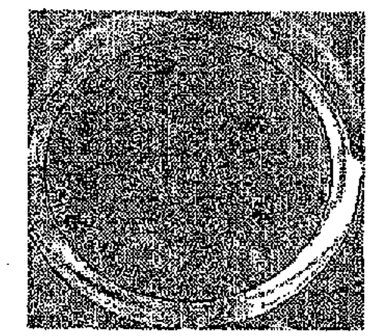


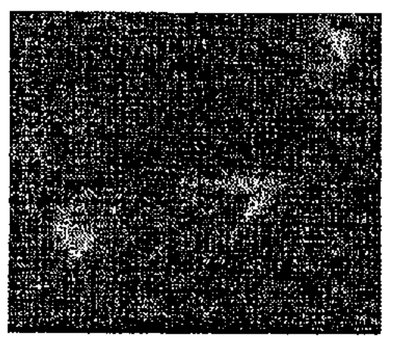
Challenge with hamster-adapted Asibi

Challenge with wt WN-NY99

Fig. 6. YF/TBE viruses

prM-E genes		P0 virus titer log_{10} PFU/ml	P1 virus titer log_{10} PFU/ml	Immuno-staining αRSSE mHIAF	
Hypr	TBEV prM-SP (p42)	**7.6**	**7.9**		**<p42**
	WNV prM-SP (**p45**)	**7.1**	**7.4**		**p45>**
Hypr + dC2 "CΔ3aa" (p59)		**5.6**	**6.5**		**p59>**
LGT/E5 (p43)		**7.8**	**8.1**		**<p43**

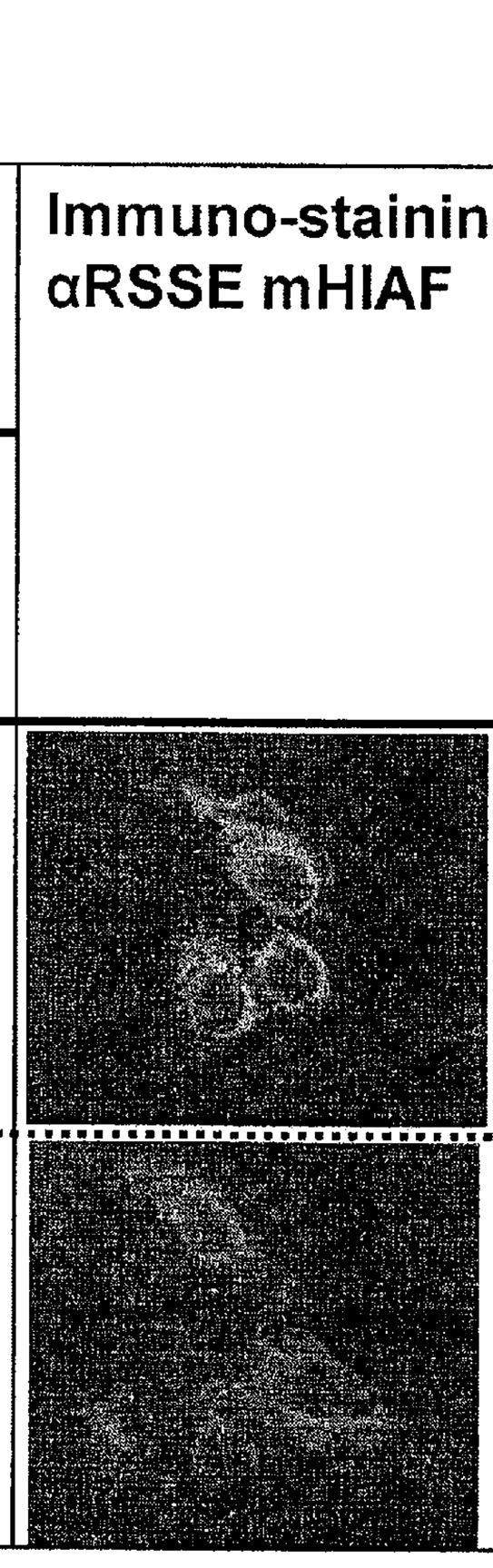




NOTES:
- p42, p45, p59, and p43 are designations of plasmids
- plaque morphology for p59-derived chimera was determined in a separate titration experiment (not shown; result of immunofluorescence assay shown)

WO 2010/107847 PCT/US2010/027552 CA 02755257 2011-09-12

Fig. 7. PIV-WNV/TBE constructs

prM-E genes	P0 titer log_{10} FFU/ml		P1 titer log_{10} FFU/ml		Immuno-staining αRSSE mHIAF
	C helper cells	CprME helper cells	C helper cells	CprME helper cells	
Hypr (TBEV signal) **(p39)**	7.2	6.7	6.9	7 7.1 . 1	
Hypr (WNV signal) **(p40)**	6.7	6.0	5.9	6 6.9 . 9	

WO 2010/107847 PCT/US2010/027552 CA 02755257 2011-09-12

Fig. 8. Replication kinetics of live YF/TBE and replication defective PIV-WNV/TBE variants in cell substrates

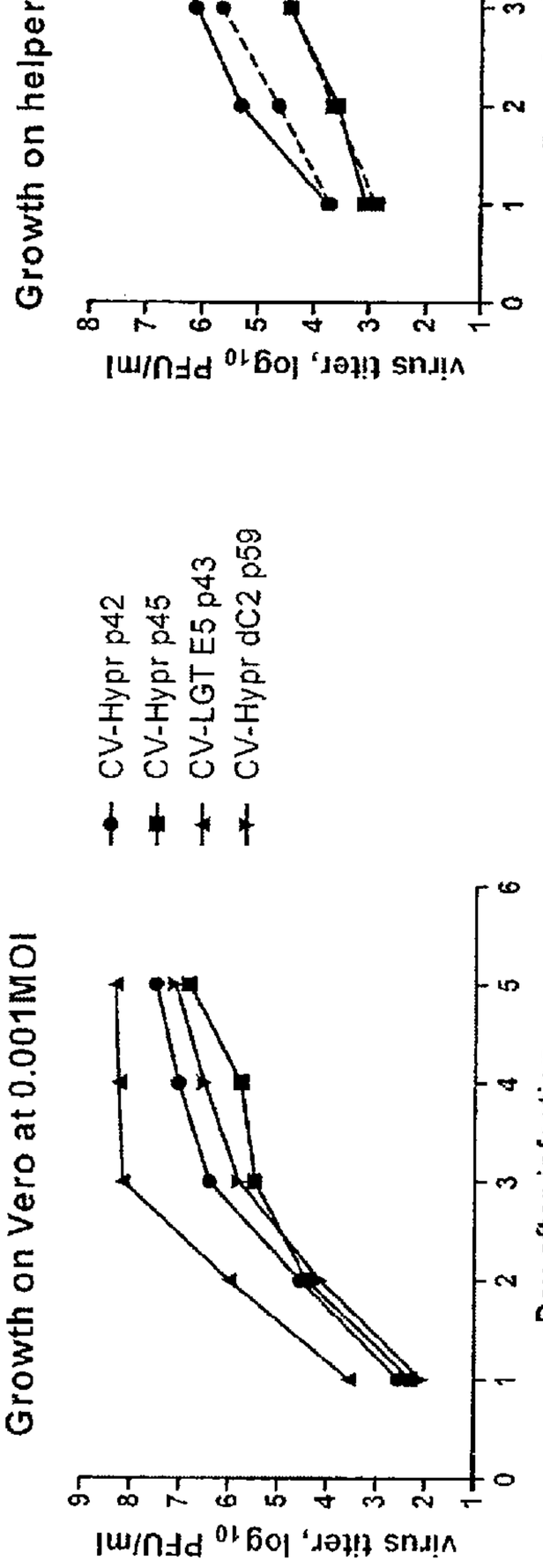


Fig. 9. Survival of mice inoculated IC with PIV-TBE and YF/TBE constructs in the neurovirulence test (3.5 week-old ICR mice)

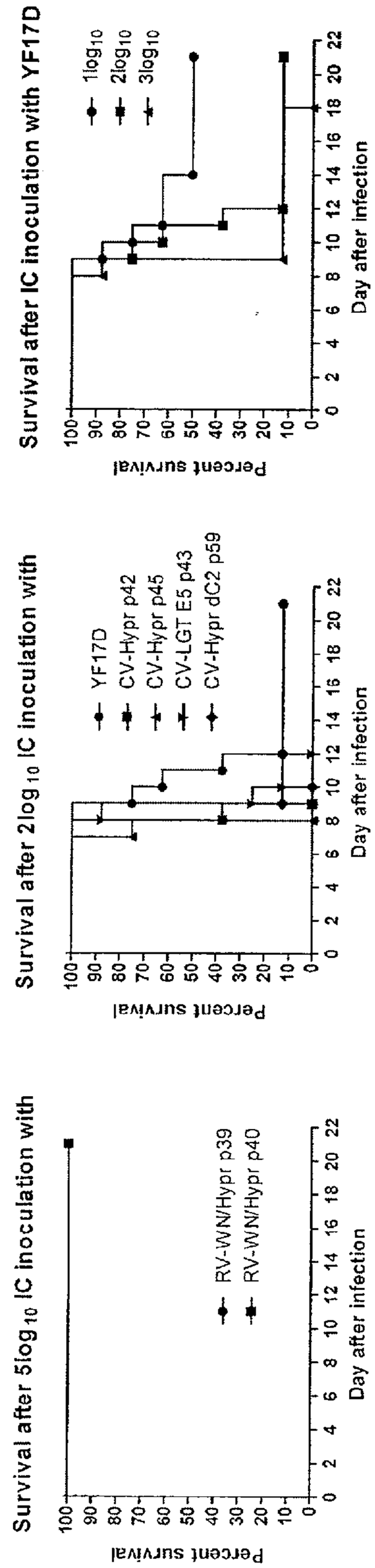


Fig. 10. Survival of mice inoculated IP with PIV-TBE and YF/TBE constructs in a neuroinvasiveness test (3.5 week-old ICR mice)

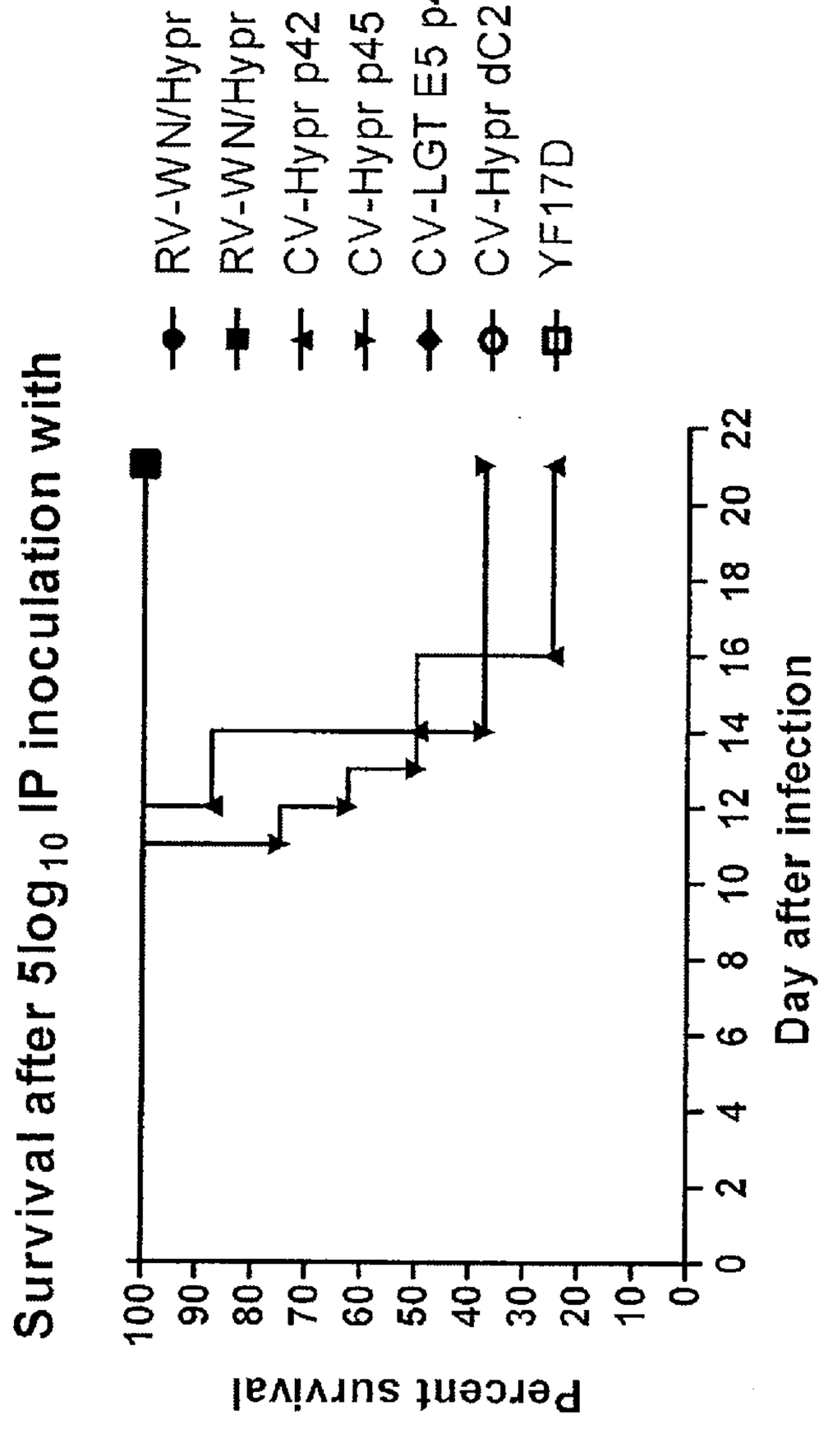
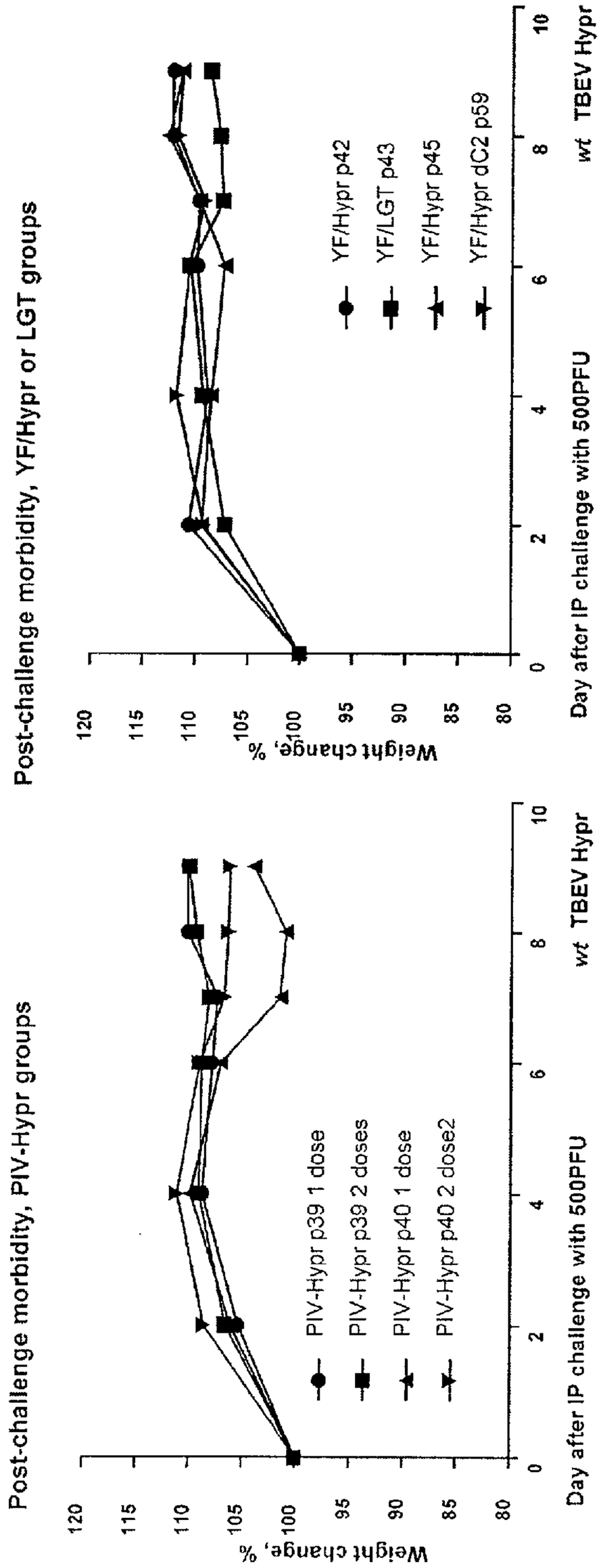


Fig. 11. Post-TBE-challenge morbidity (weight loss); day 9 post-challenge



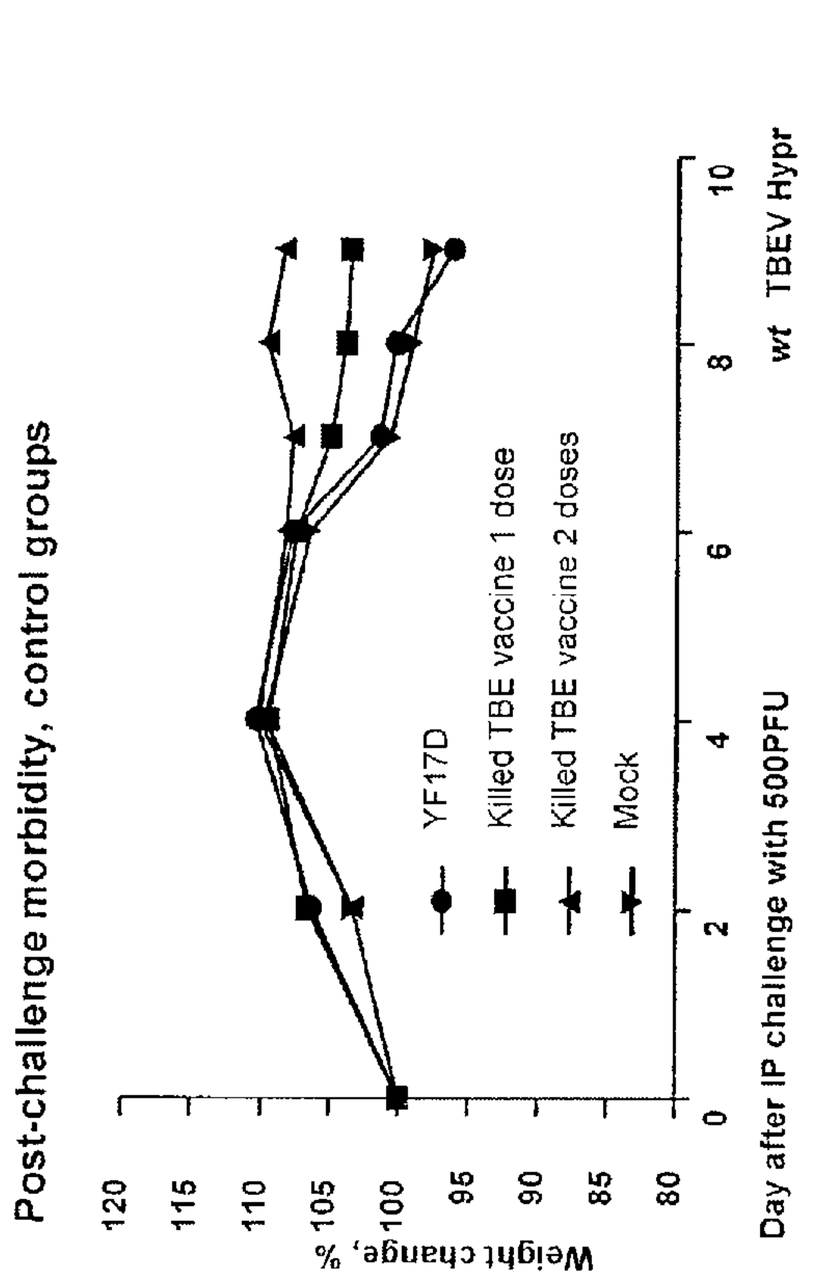
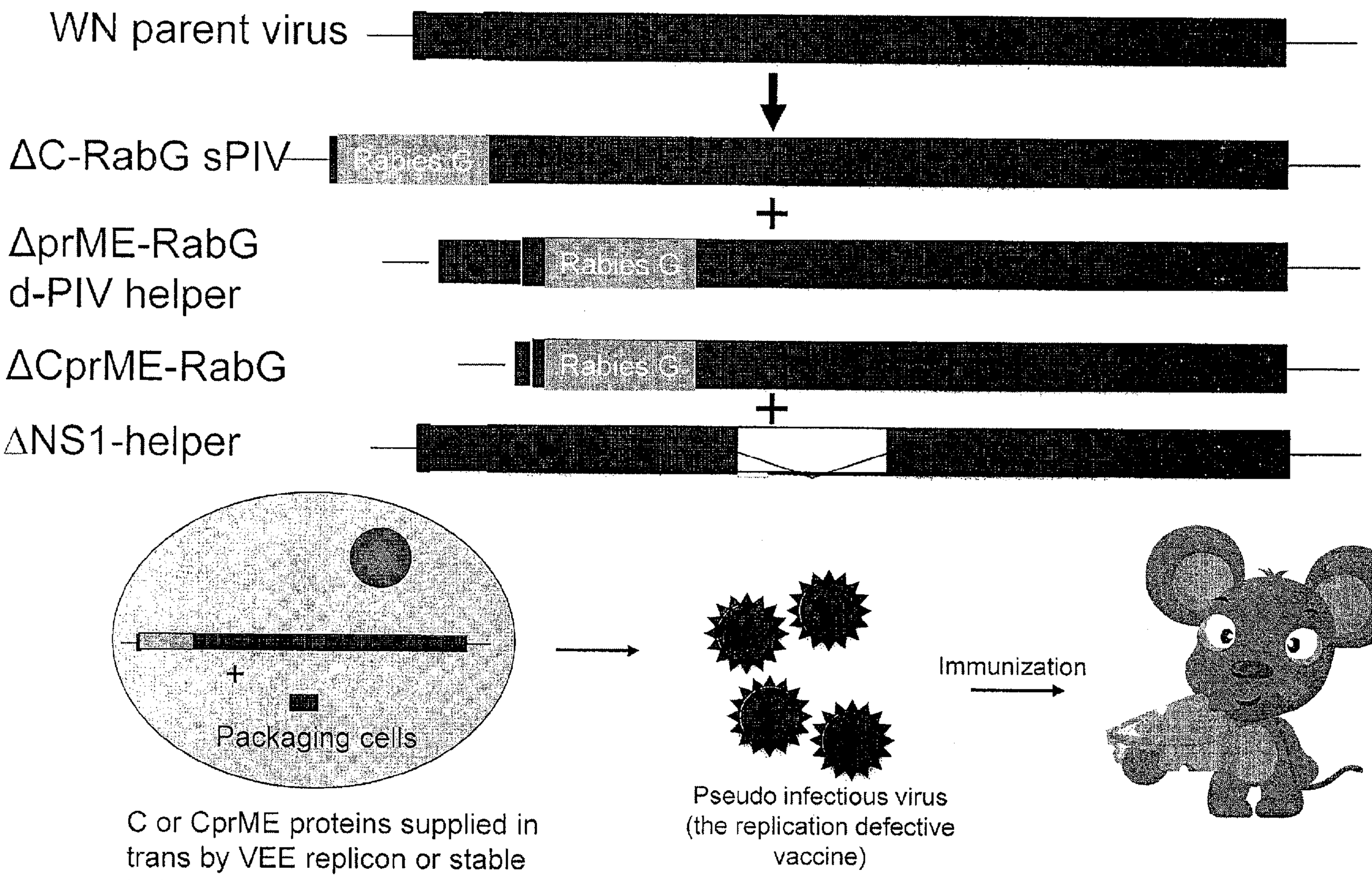


Fig. 12. Examples of PIV constructs expressing foreign antigens (rabies G)



NOTE: possible d-PIV formulations are shown (+).

Fig. 13. Schematic representation of insertion designs resulting in viable/expressing constructs (exemplified by rabies G)

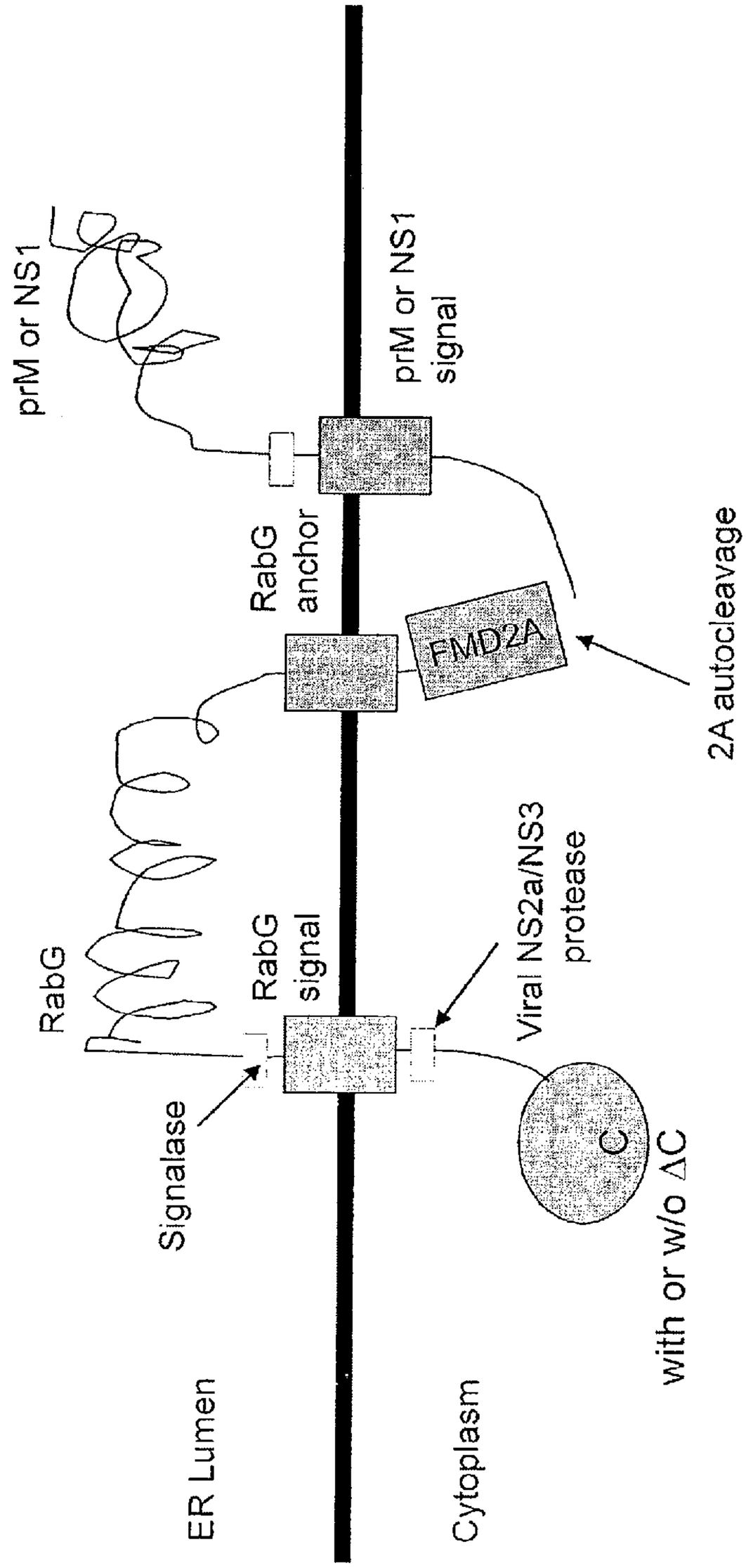


Fig. 14. PIV-WN/Rabies G PIVs: immunofluorescence of transfected cells and growth curves after transfection

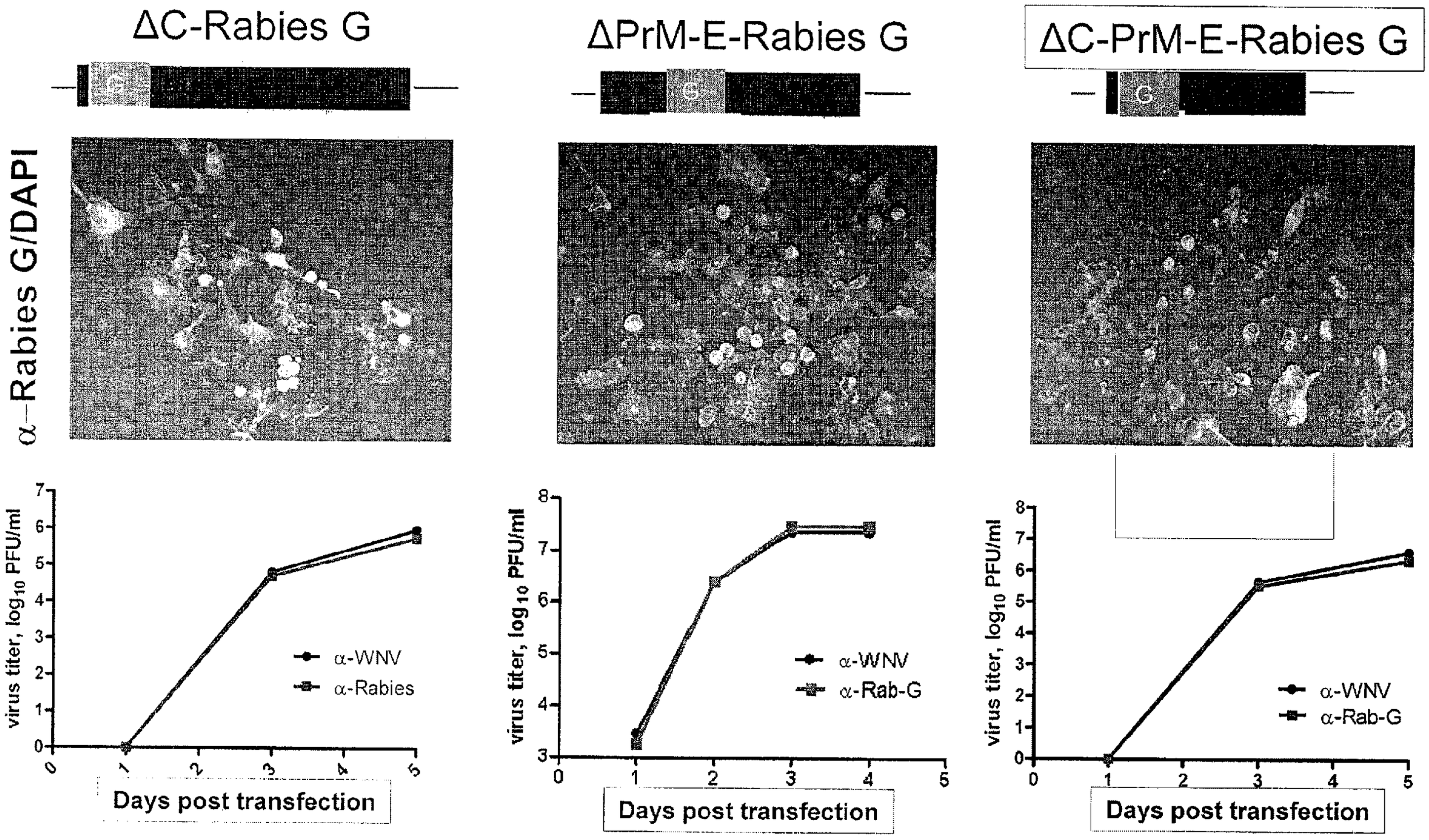
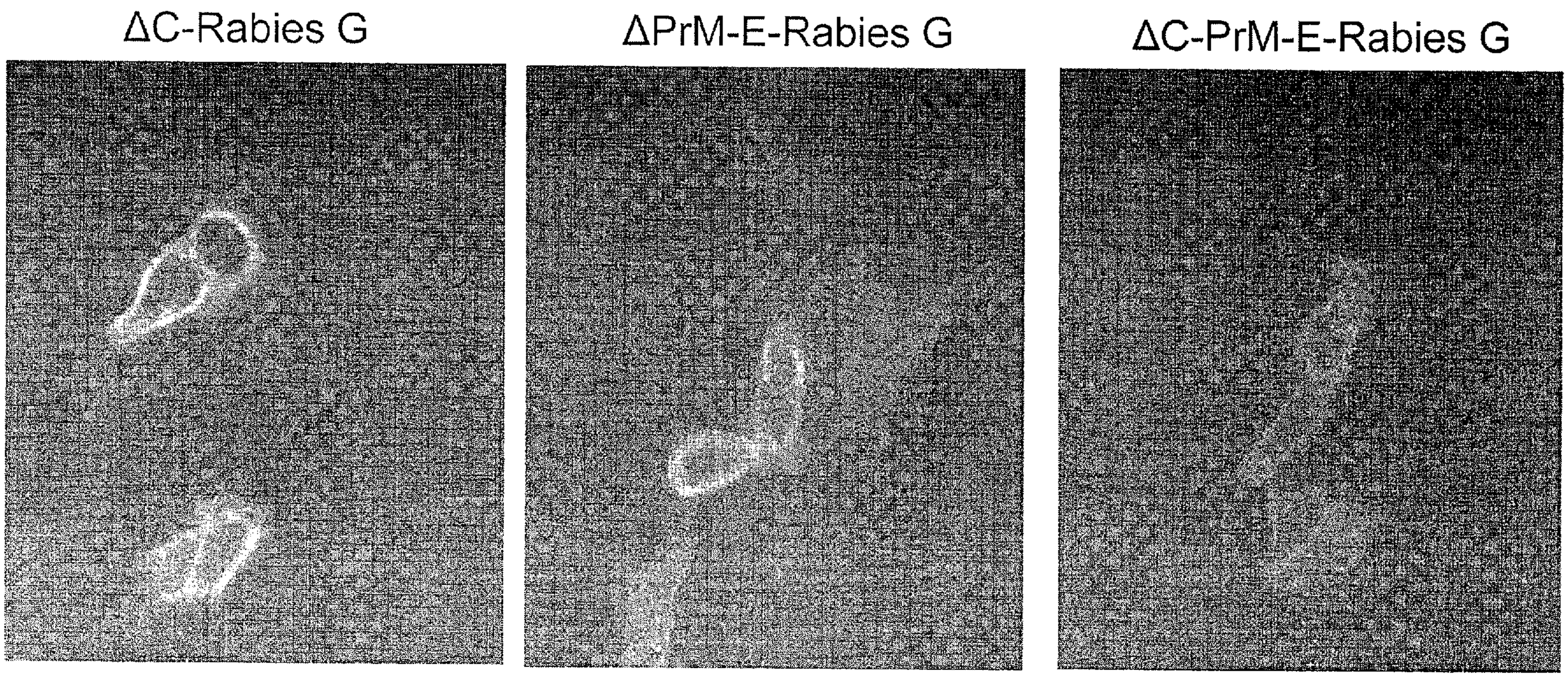


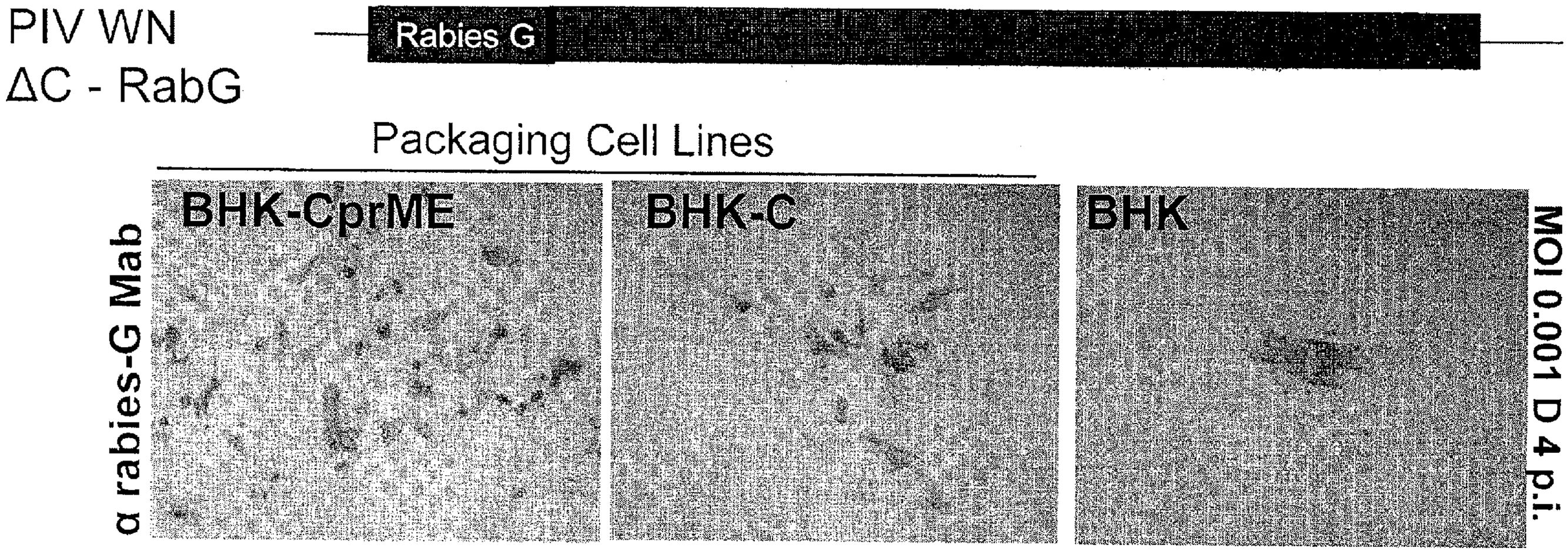
Fig. 15. Efficient expression of RabG on the plasma membrane of Vero cells

MOI 0.1, day 2 post infection, 4% PFA Fixed non-permeabilized



1Ab (Abcam) Anti Rabies 1:500 2-3hrs RT, 2Ab Anti-Mouse IgG 1:1000 1hr RT
Images are 40X and exposure time is 40ms

Fig. 16. PIV - Rabies G spread in helper cells, no spread in naïve cells

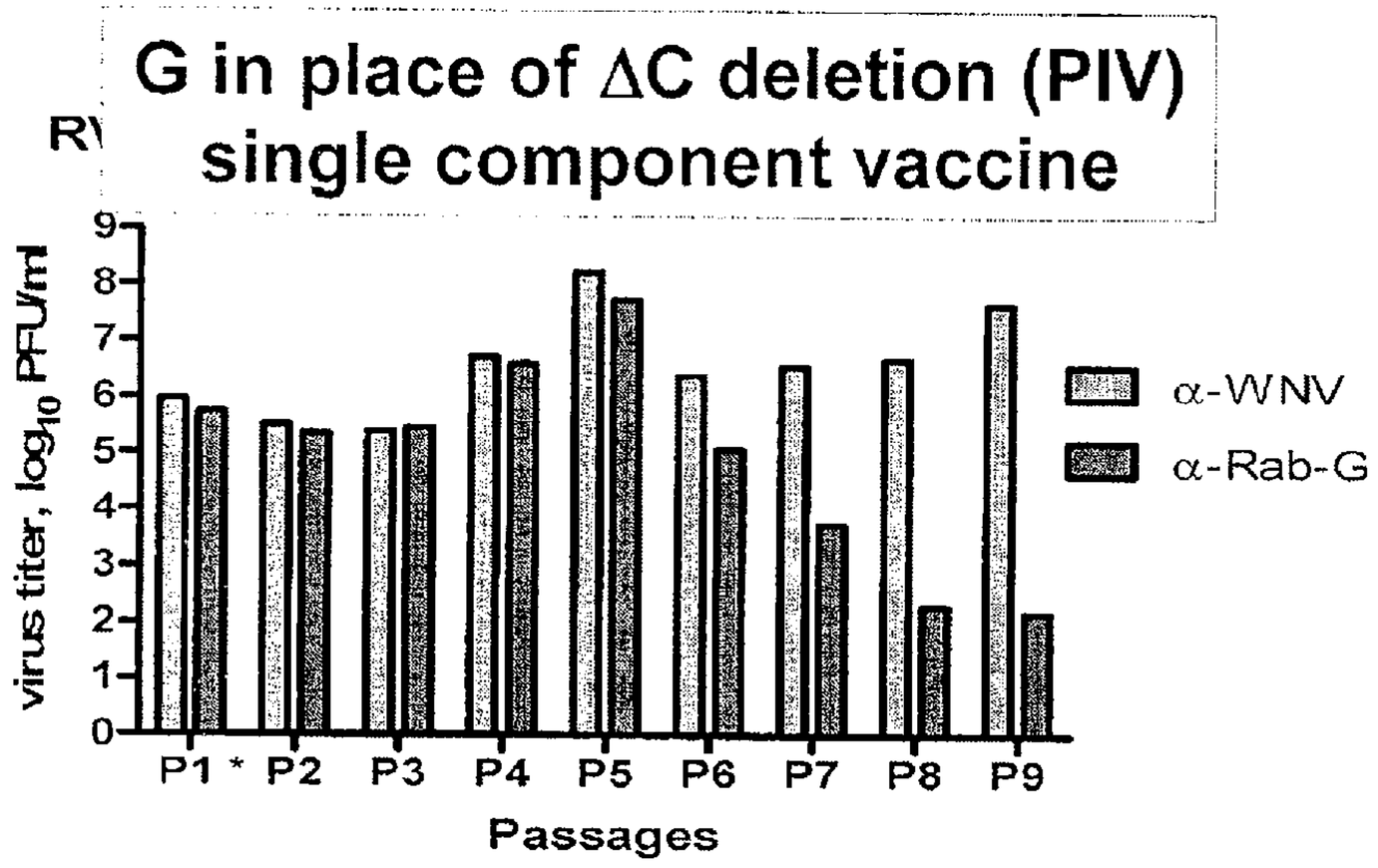


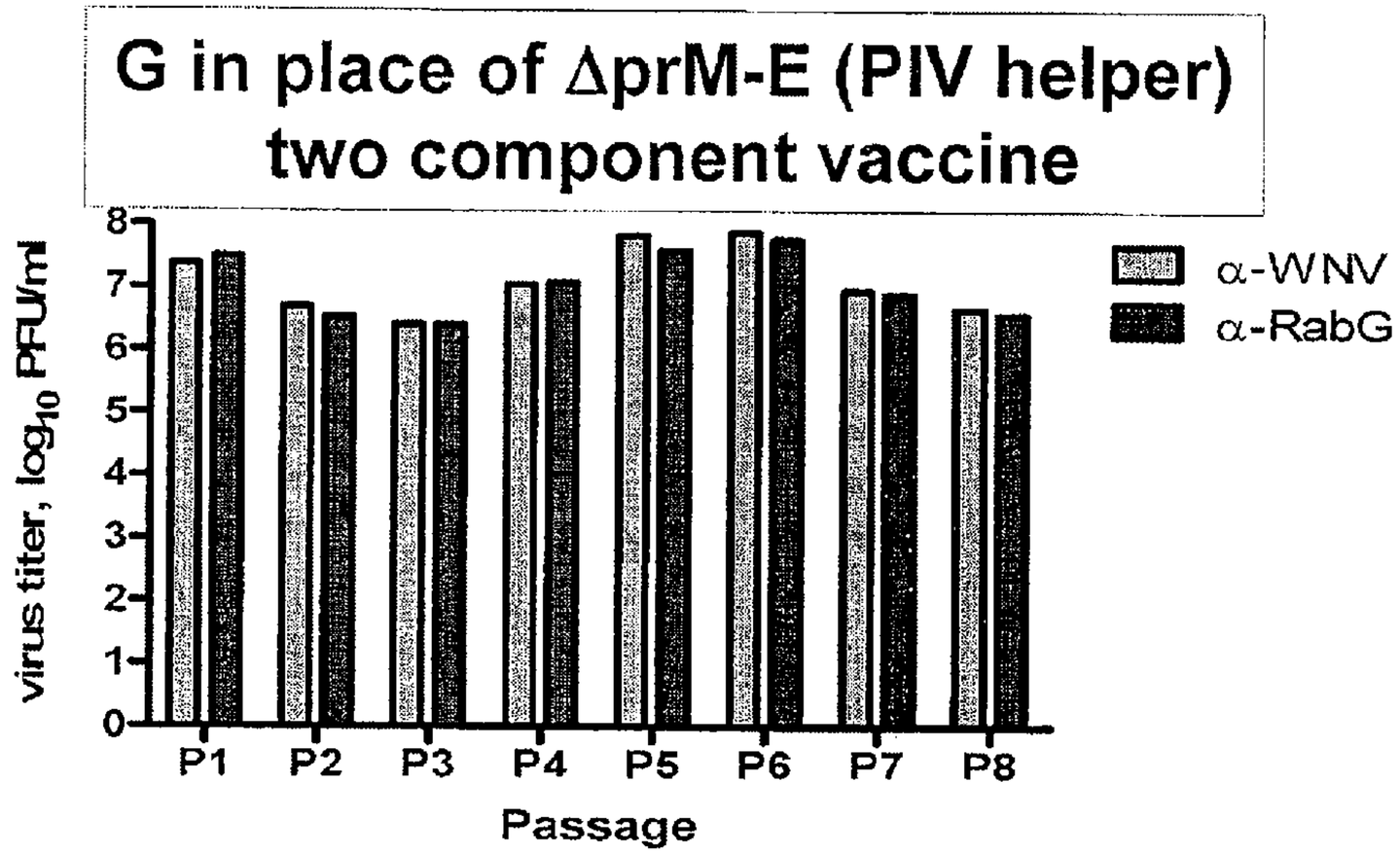
PIVs spread in helper cells and not naïve cells.
No infectious material in supernatants from BHK cells.

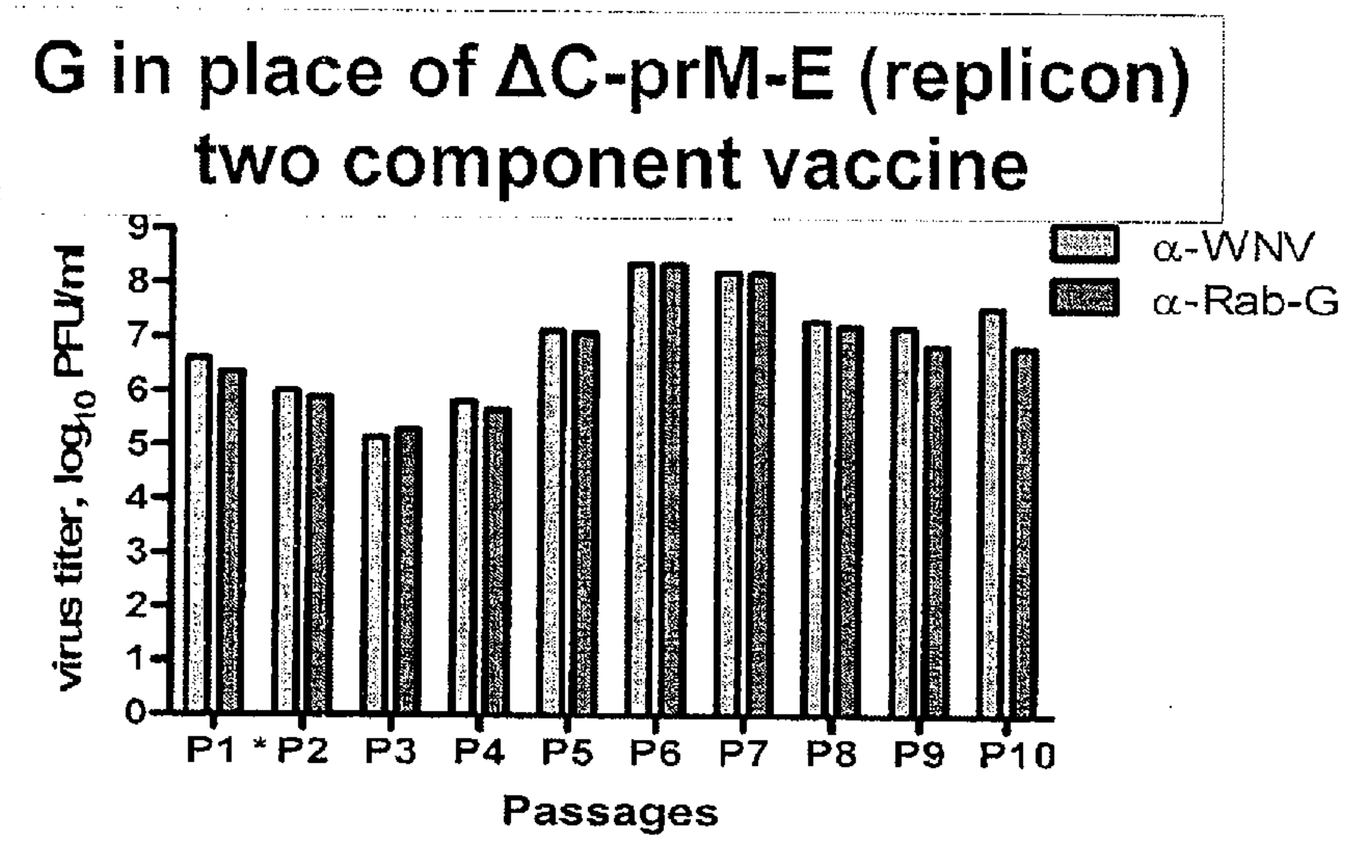
Similarly, VSV G was expressed, and no spread in regular cells (Vero, BHK) was observed and no infectious material detected in the supernatants

Fig. 17. Stability of rabies G protein gene in PIV-WN vectors

Passages in BHK-C-prM-E helper cells, MOI 0.1; titration in Vero by immunostaining

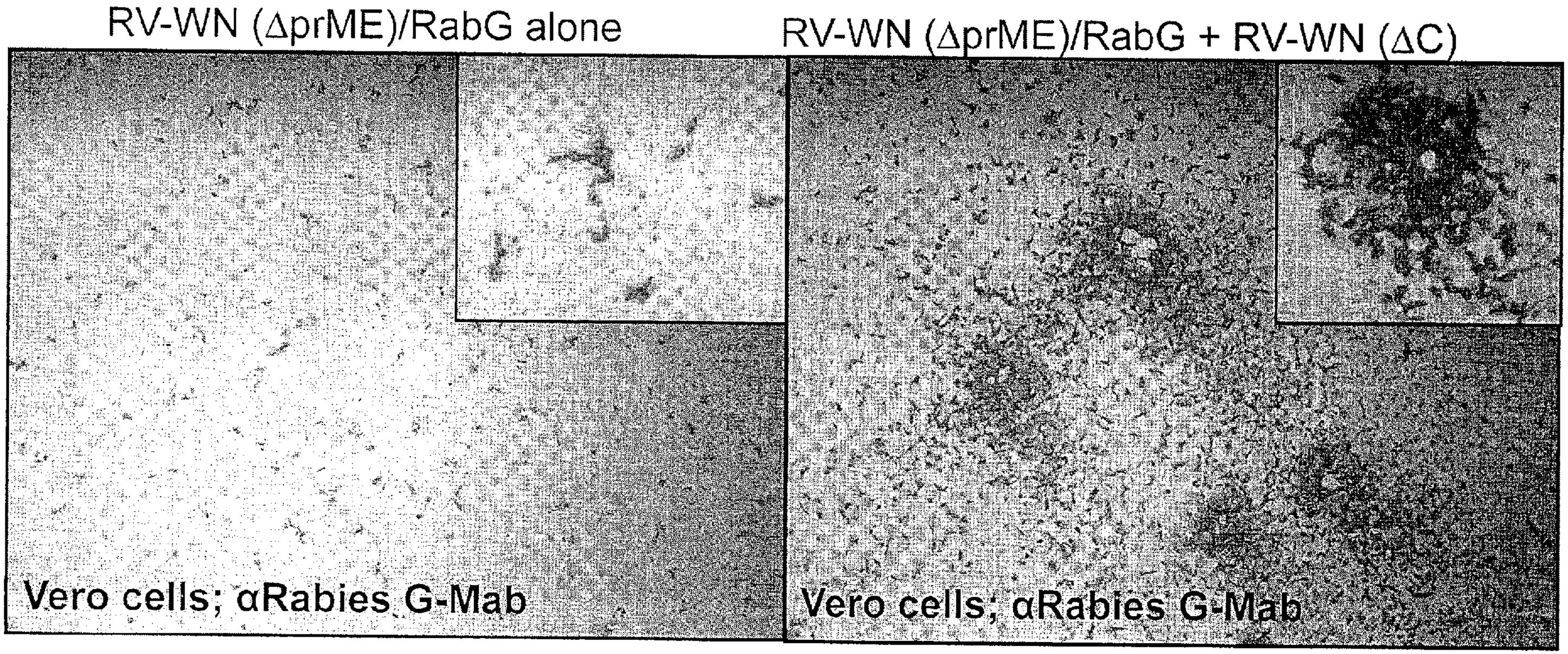






WO 2010/107847 PCT/US2010/027552
CA 02755257 2011-09-12

Fig. 18. Comparison of spread in Vero cells of single-component vs. two-component PIV-Rabies variants



Titer on helper BHK cells: **2.0x10^8 FFU/ml**

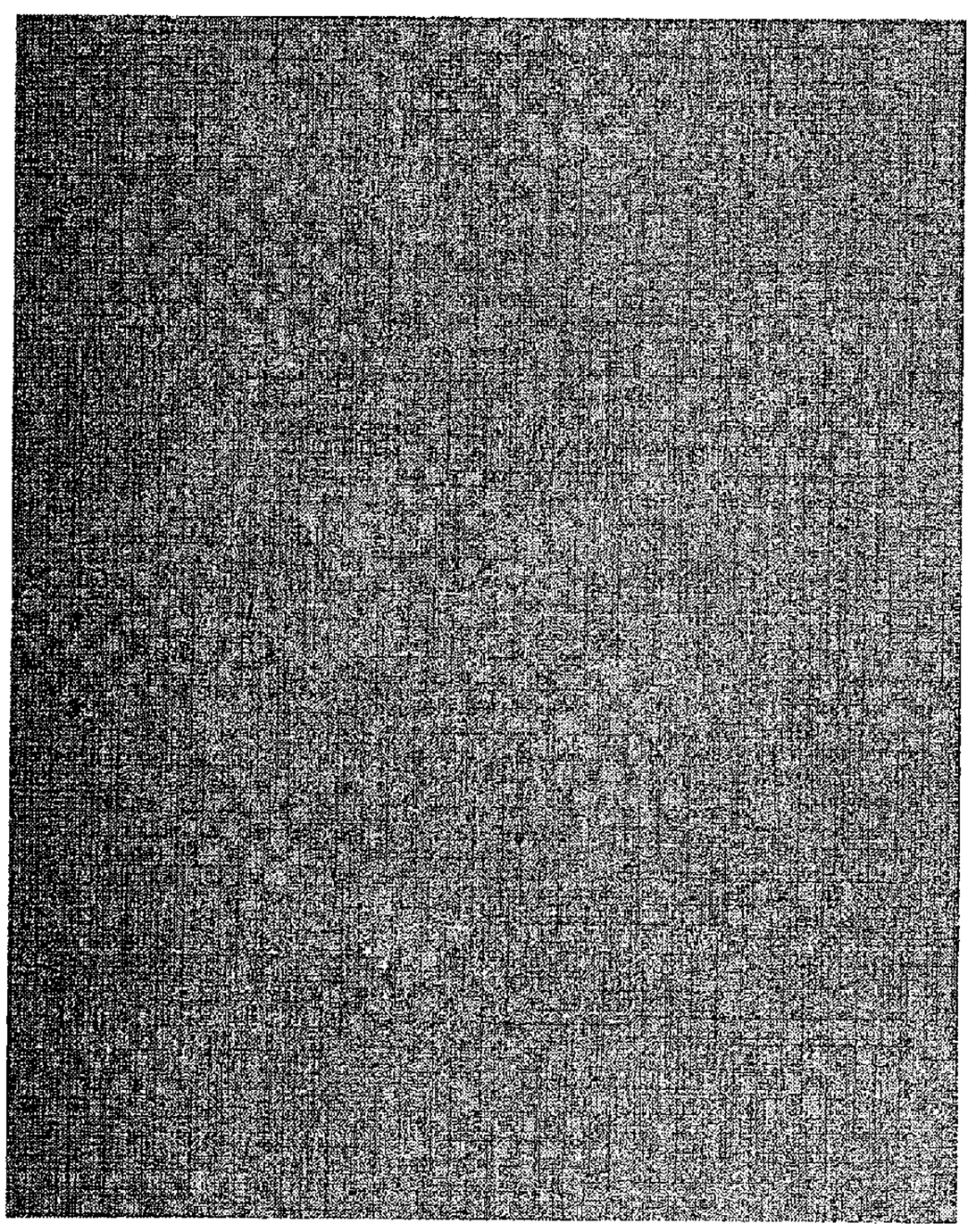
Titer on naïve Vero cells after transfection (P0): **2.0x10^8 FFU/ml**

WO 2010/107847 PCT/US2010/027552 CA 02755257 2011-09-12

Fig. 19. Example of expression of full-length RSV F protein (strain A2): immunostaining of helper cells after transfection

PIV-WNΔprME-RSV F, 3 days post-transfection

Control BHK helper cells (WN C-prM-E)



Cells stained with anti-RSVF Mab, DAPI

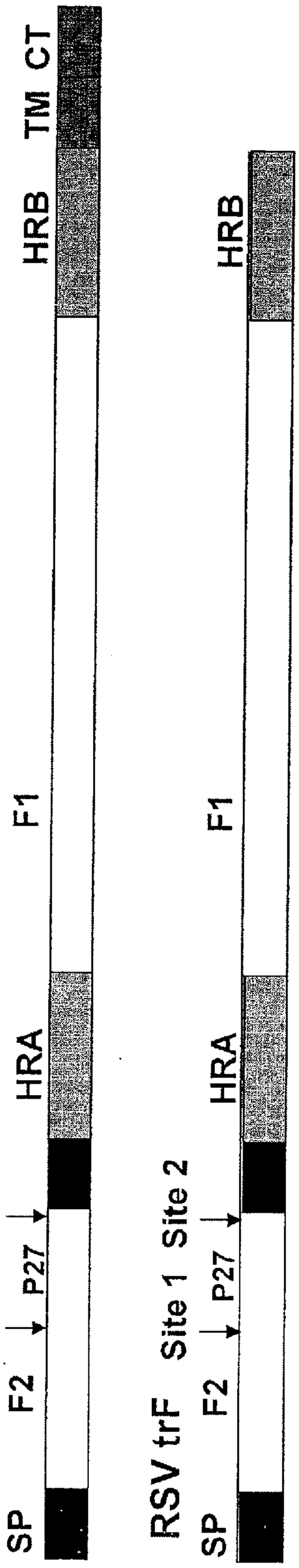
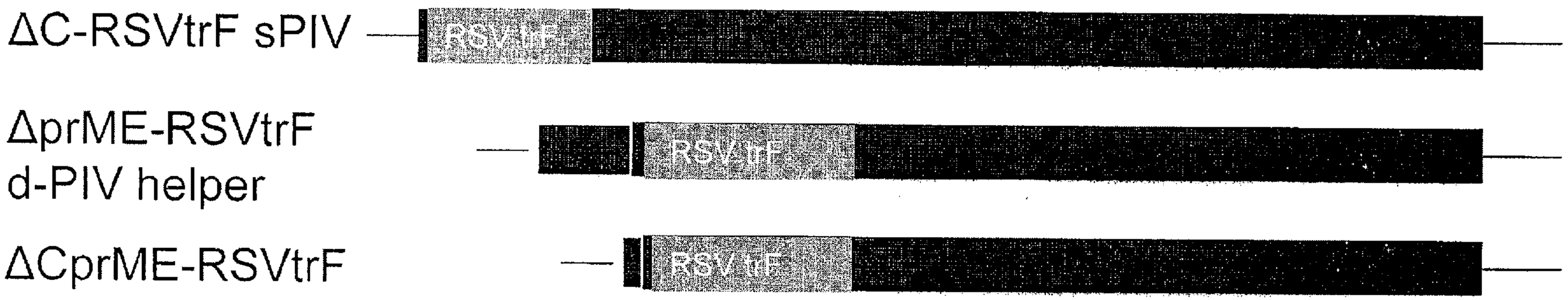

Wild type RSV F
SP
F2
Site 1
P27
Site 2
HRA
F1
HRB
TM
CT
RSV trF
SP
F2
Site 1
P27
Site 2
HRA
F1
HRB

Fig. 20

Fig. 21

Truncated F (A1 strain) expressed in RV-WN vectors:



Immunofluorescence of helper cells after transfection (Day 4)

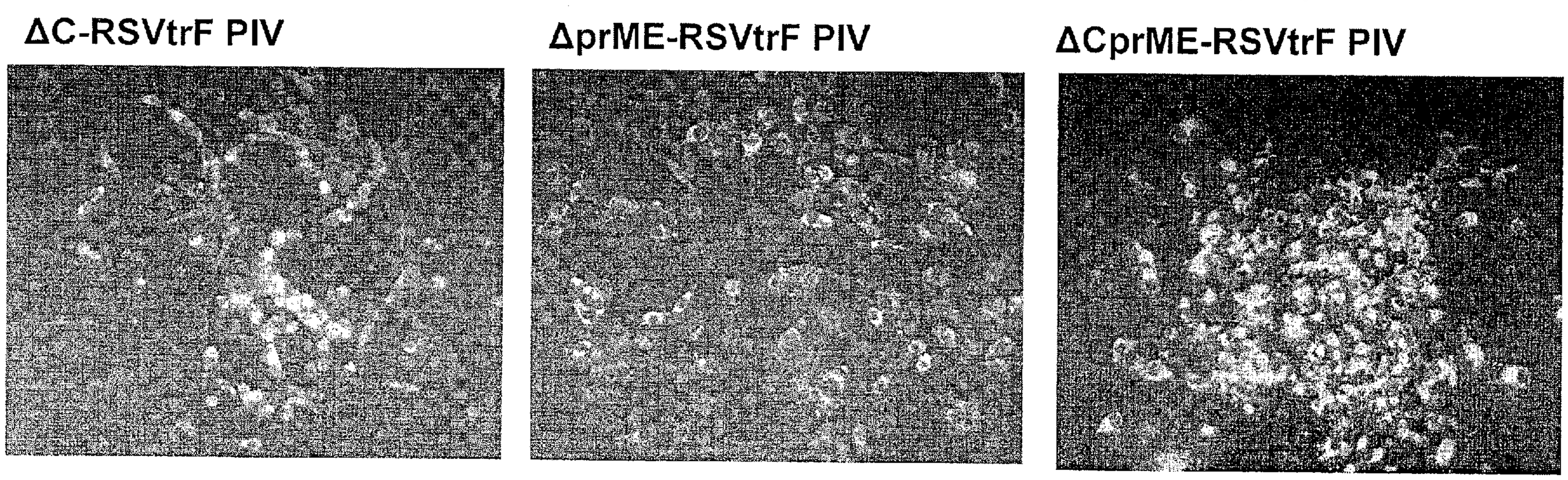
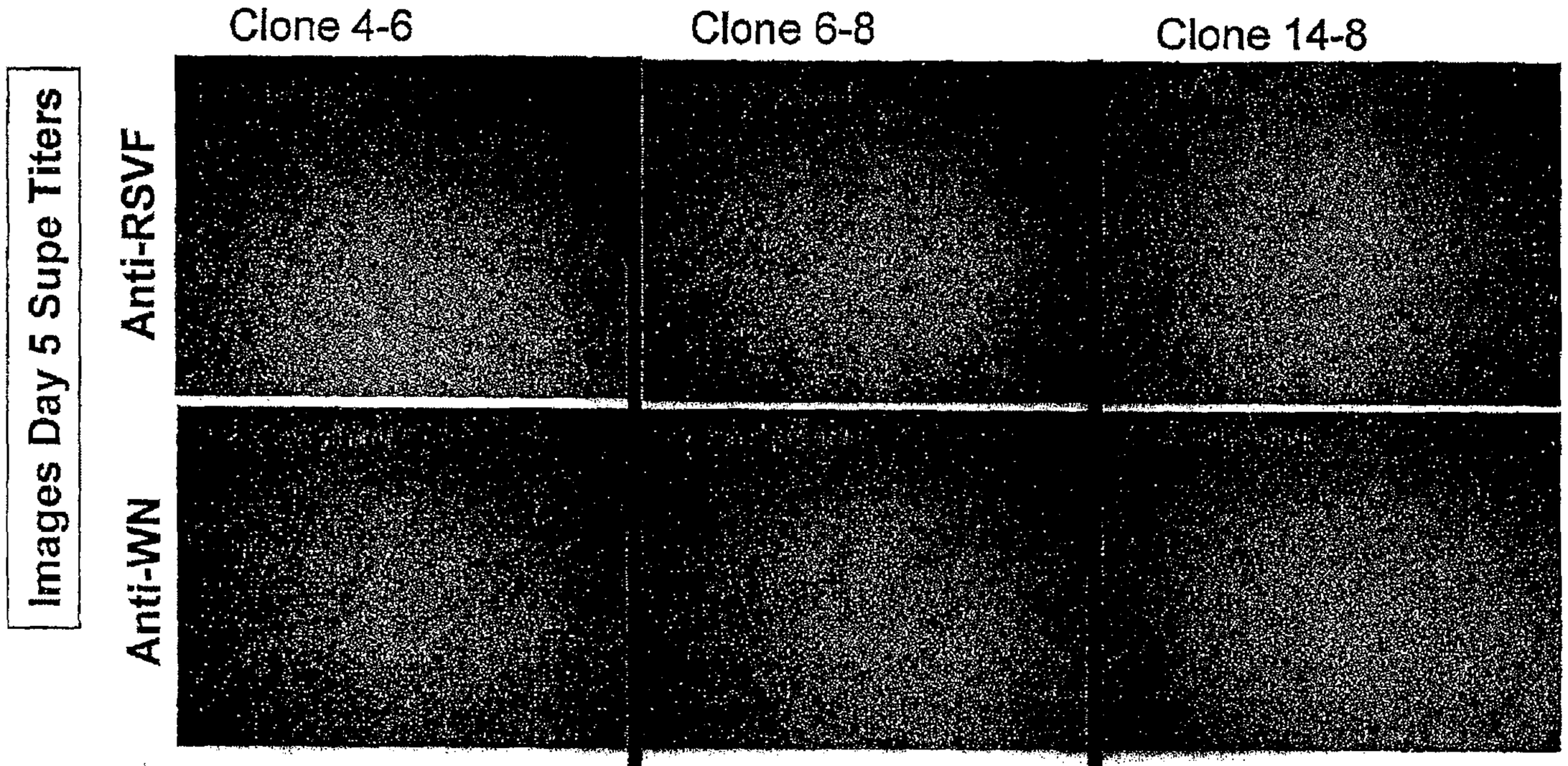


FIG. 22 **Titration of ΔC-RSV trF PIV in Vero cells using immunostaining (samples from day 4 & 5 post-transfection obtained from 3 different plasmid clones)**

Infect Vero cells 2 days, methanol fix, stain Anti-RSVF Mab 1:000 or Mab Anti-WN 1:1000)



	Clone 4-6		Clone 6-8		Clone 14-8	
	Anti-RSVF	Anti-WN	Anti-RSVF	Anti-WN	Anti-RSVF	Anti-WN
D4	2.4 X 10^5	3.4 X 10^5	5.6 X 10^4	6.5 X 10^4	2.6 X 10^5	4.8 X 10^5
D5	1.9 X 10^5	2.2 X 10^5	2.2 X 10^5	3.7 X 10^5	1 X 10^6	2.1 X 10^6

CA 02755257 2011-09-12

Fig. 23

Serum-free stock of **ΔprME-RSVtrF P3**: titer 6.7×10^8 FFU/ml

Westerns of two ΔprME-RSVtrF stocks, 2 days postinfection

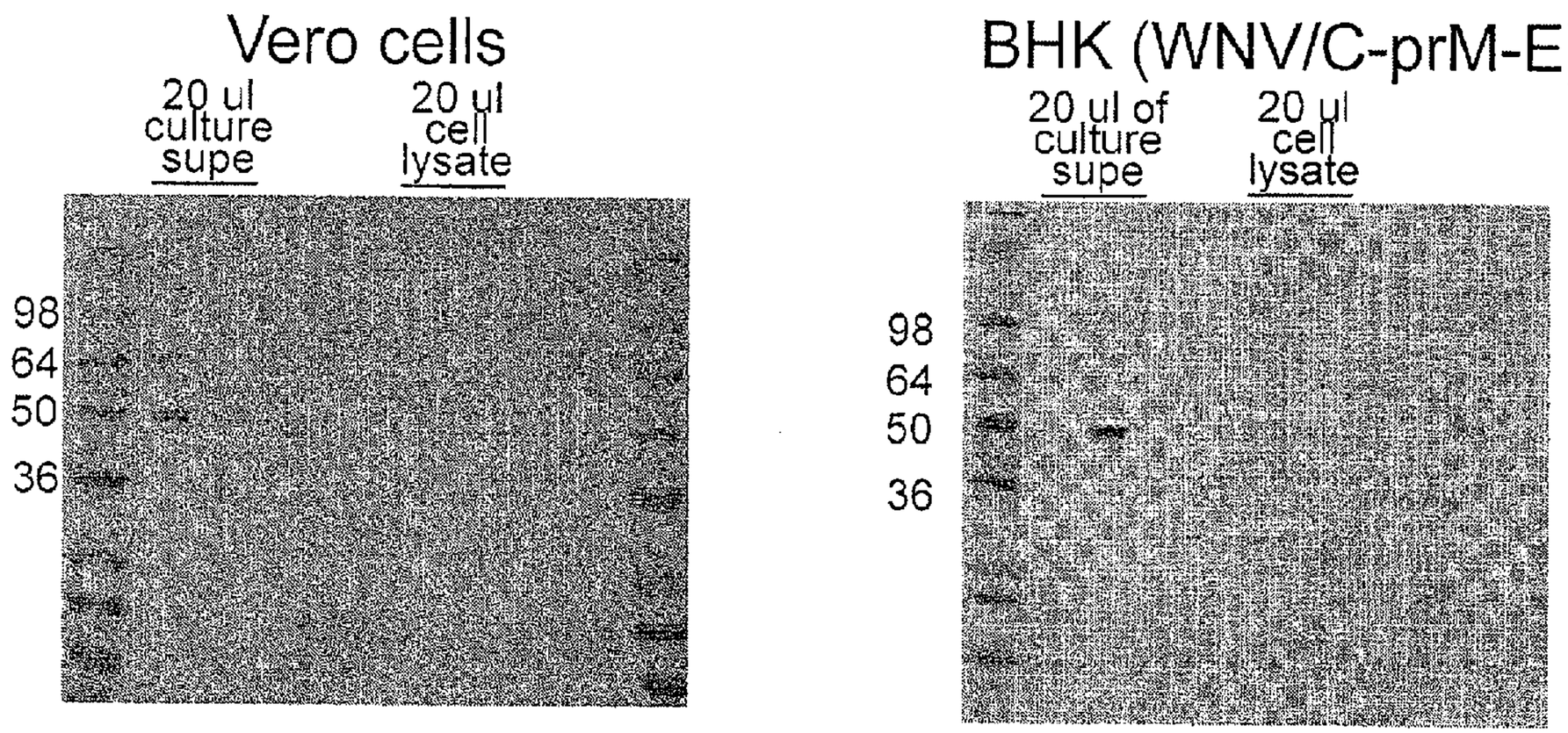


Fig. 24

Lane 1. standard marker
Lane 2. vP2400, 10 MOI, Day 1
Lane 3. vP2400, 10 MOI, Day 2
Lane 4. vFP2403, 10 MOI, Day 1
Lane 5. vFP2403, 10 MOI, Day 2
Lane 6. PIV-F, 10 MOI, Day 1
Lane 7. PIV-F, 10 MOI, Day 2
Lane 8. Media control, Day 1
Lane 9. Media control, Day 2
Lane 10. RSV-F positive control

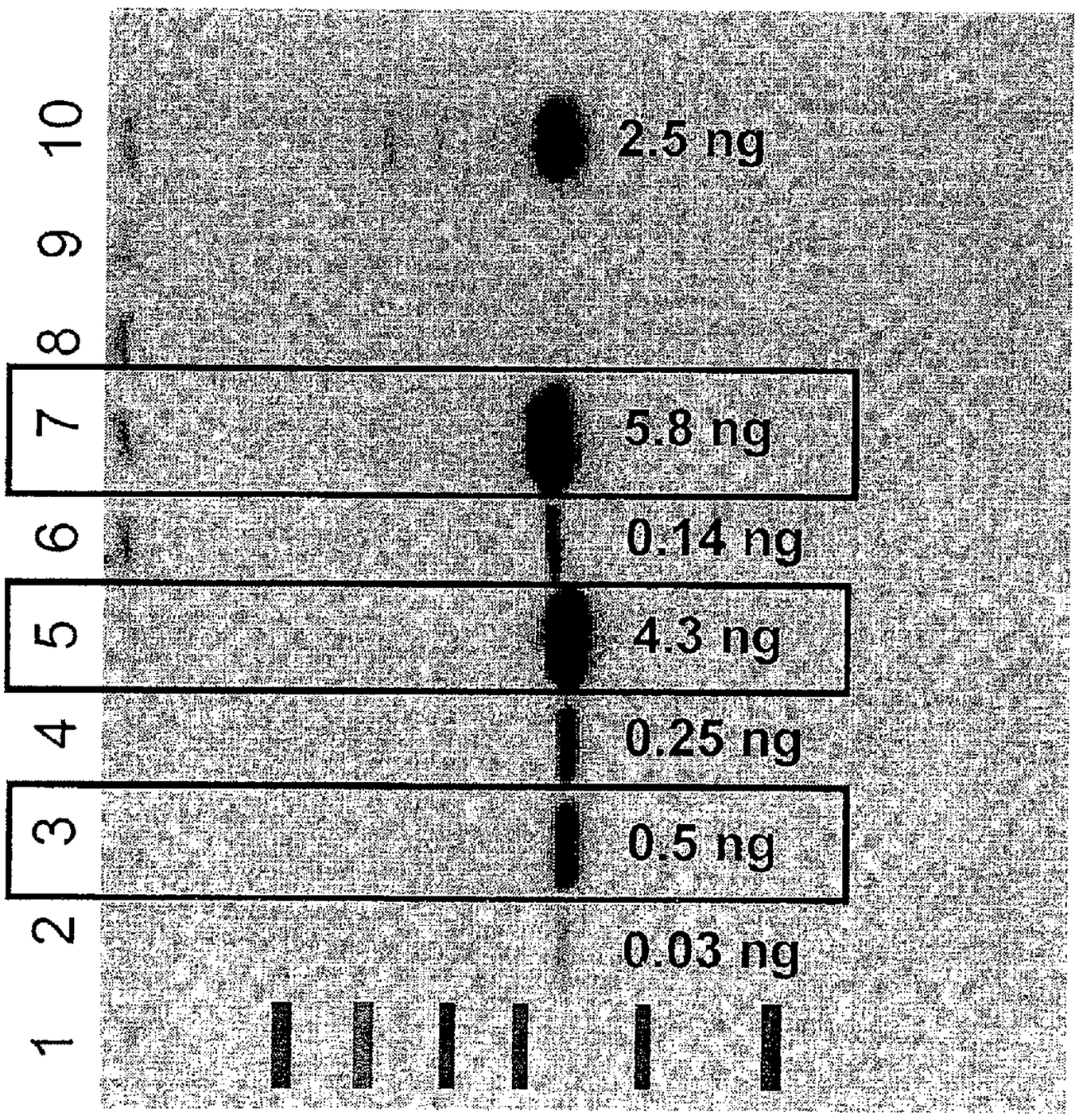
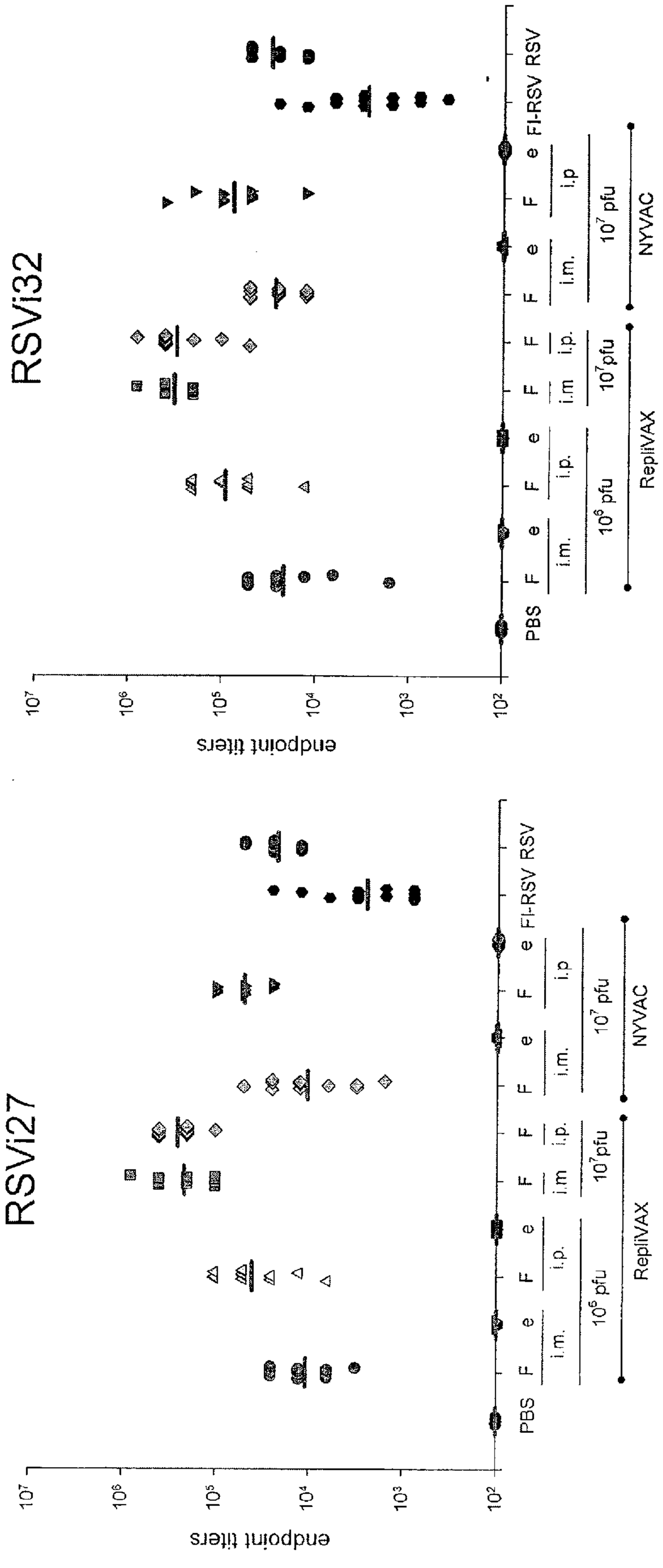


Fig. 25


RSVi32
RSVi27
endpoint titers
10^7
10^6
10^5
10^4
10^3
10^2
PBS
F e i.m.
F e i.p.
10^6 pfu
F e i.m.
F e i.p.
10^7pfu
RepliVAX
F e i.m.
F e i.p.
10^7 pfu
NYVAC
F e FI-RSV RSV

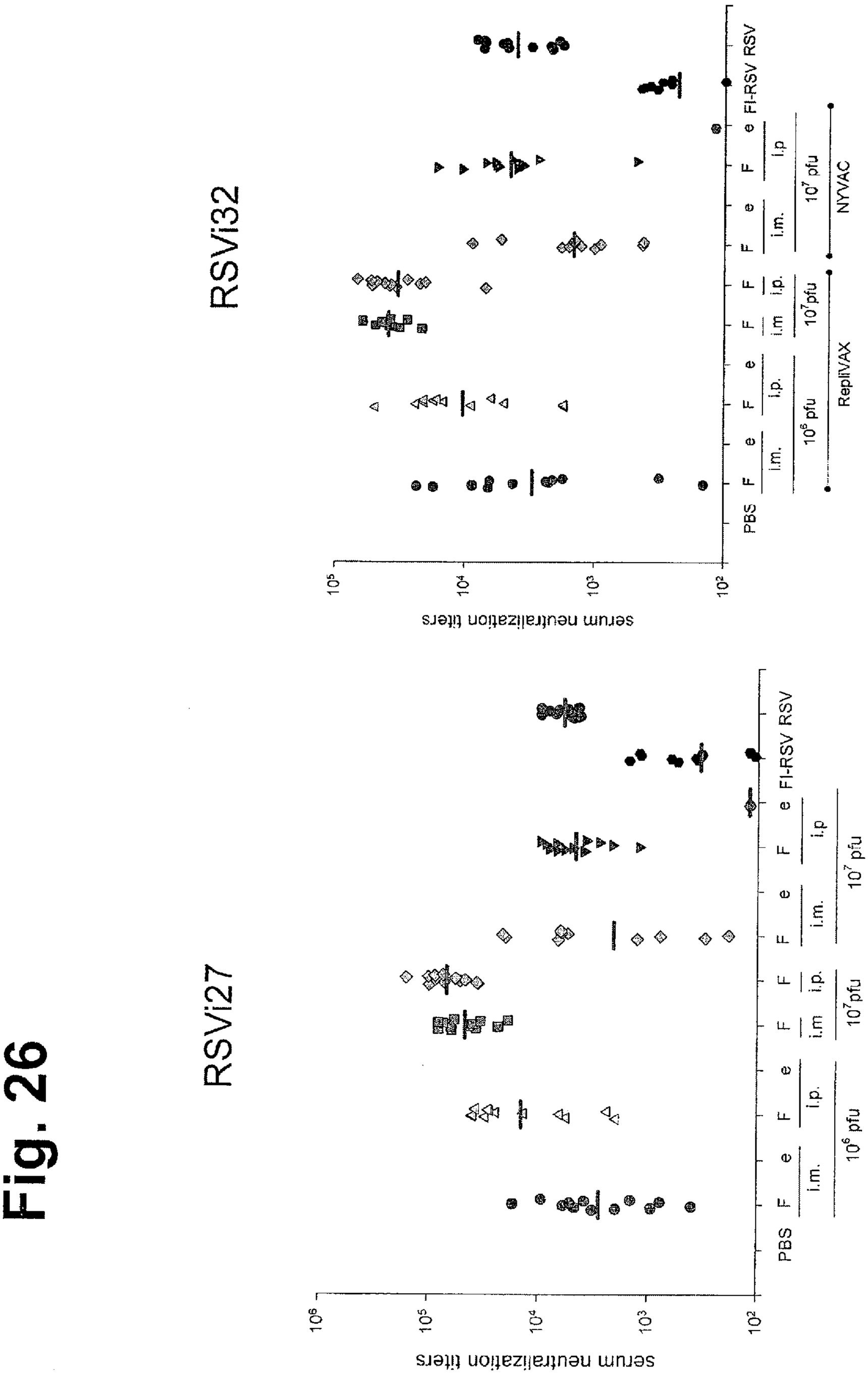

RSVi32
serum neutralization titers
10^5
10^4
10^3
10^2
PBS
F e i.m.
F e i.p.
10^6 pfu
F e i.m.
F e i.p.
10^7 pfu
RepliVAX
F e i.m.
F e i.p.
10^7 pfu
NYVAC
RSVi27
serum neutralization titers
10^6
10^5
10^4
10^3
10^2
PBS
F e i.m.
F e i.p.
10^6 pfu
F e i.m.
F e i.p.
10^7 pfu
RepliVAX
F e i.m.
F e i.p.
10^7 pfu
NYVAC

Fig. 26