

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
25 June 2009 (25.06.2009)

PCT

(10) International Publication Number
WO 2009/079185 A2

(51) International Patent Classification:
C12N 15/09 (2006.01)

(21) International Application Number:
PCT/US2008/084848

(22) International Filing Date:
26 November 2008 (26.11.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/990,088 26 November 2007 (26.11.2007) US

(71) Applicant (for all designated States except US): **NOVARTIS VACCINES AND DIAGNOSTICS, INC.** [US/US];
4560 Horton Street, Emeryville, CA 94608 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **TANG, Zequn** [US/US]; C/o Novartis Vaccines And Diagnostics, Inc.,
4560 Horton Street, Emeryville, CA 94608 (US). **ZUR MEGEDE, Jan** [DE/US]; 1422 A Church Street, San
Francisco, CA 94131 (US).

(74) Agent: **HEMMENDINGER, Lisa, M.**; Banner & Wit-
coff, Ltd., 1100 13th Street, N.W., Suite 1200, Washington,
DC 20005-4051 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA,
CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE,
EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID,
IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK,
LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW,
MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT,
RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, 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, MT, NL,
NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG,
CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished
upon receipt of that report
- with sequence listing part of description published sepa-
rately in electronic form and available upon request from
the International Bureau

(54) Title: METHODS OF GENERATING ALPHAVIRUS PARTICLES

(57) Abstract: Strategies for increasing the productivity of alphavirus packaging cell lines and of reducing the possibility that
replication competent virus may be generated during large scale production of recombinant alphavirus particles.



WO 2009/079185 A2

METHODS OF GENERATING ALPHAVIRUS PARTICLES

- [01] This application claims the benefit of and incorporates by reference Serial No. 60/990,088 filed November 26, 2007.

FIELD OF THE INVENTION

- [02] The invention relates to the preparation of recombinant alphavirus particles.
- [03] This invention was supported by Contract No. HHSN266200500007C from the National Institutes of Health. The U.S. Government may have certain rights in the invention.

BACKGROUND OF THE INVENTION

- [04] Recombinant alphavirus particles (alphavirus replicon particles) have great potential for use in protein production, antigen delivery, and various therapeutic applications. Alphavirus packaging cell lines (PCL) are the most efficient and cost effective way to generate alphavirus replicon particles. One obstacle in the development of alphavirus packaging cell lines, however, is the low particle yield. On the other hand, generation of RCV (replication competent viral particles) is a potential problem when generating large numbers of recombinant alphavirus particles. The probability of recombination can be greatly reduced by dividing the defective helpers in two separate cassettes, because multiple switches would be required to produce an infectious RNA. However, it is possible that large-scale production could still generate RCV. Thus, there is a need in the art for methods of increasing the productivity of PCL and of reducing the possibility that replication competent virus may be generated during large scale production of recombinant alphavirus particles.

BRIEF DESCRIPTION OF THE FIGURES

- [05] **FIG. 1.** Configurations of several double subgenomic promoter helper cassettes. “CMV,” cytomegalovirus promoter; “5’,” 5’ untranslated region and sequences from

the N-terminus of the Nsp1 coding region that are necessary for replication; “JR,” subgenomic promoter with the adjacent sequences (junction region); “Cap/Gly,” capsid or glycoprotein; “Nsps,” non-structural proteins 1-4; “A40,” 40-mer polyA tail; “3’,” 3’ untranslated region; “Psv40,” SV40 promoter controlling transcription of neomycin resistance gene (“neo”); “IRES,” internal ribosome entry site.

- [06] **FIGS. 2A-D.** BLAST alignment showing cleavage sites in nonstructural proteins of various types of alphavirus. **FIG. 2A**, SFV Nsp2, SEQ ID NO:14; Nsp2 VCR-Chim, SEQ ID NO:15; Sindbis Nsp2, SEQ ID NO:16; EEE Nsp2, SEQ ID NO:17; consensus, SEQ ID NO:18. **FIG. 2B**, SFV Nsp2, SEQ ID NO:21; Nsp2 VCR-Chim, SEQ ID NO:22; Sindbis Nsp2, SEQ ID NO:23; EEE Nsp2, SEQ ID NO:24; consensus, SEQ ID NO:25. **FIG. 2C**, SFV Nsp2, SEQ ID NO:26; Nsp2 VCR-Chim, SEQ ID NO:27; Sindbis Nsp2, SEQ ID NO:28; EEE Nsp2, SEQ ID NO:29; consensus, SEQ ID NO:30. **FIG. 2D**, SFV Nsp2, SEQ ID NO:31; Nsp2 VCR-Chim, SEQ ID NO:32; Sindbis Nsp2, SEQ ID NO:33; EEE Nsp2, SEQ ID NO:34; consensus, SEQ ID NO:35. In the consensus sequences of **FIGS. 2A-D** provided as SEQ ID NOS:18, 25, 30, and 35 in the sequence listing, “Xaa” can be any amino acid or can be missing at the positions shown. Preferably the amino acids at positions indicated in SEQ ID NOS:18, 25, 30, and 35 with “Xaa” are selected from the amino acids shown at those positions in **FIGS. 2A-D**.
- [07] **FIG. 3.** Capsid cleavage sites, either at the 3’ end of capsid protein (Trp) or at the 5’ end of the glycoprotein serine residue (Met-Ser). 3’ end of Sindbis capsid (Scap), SEQ ID NO:11; 5’ end of Sindbis glycoprotein (Sgly), SEQ ID NO:12; RCV (replication competent viral particles), SEQ ID NO:13.
- [08] **FIG. 4.** BLAST alignment of capsid protein sequences SEQ ID NO:1 (Sindbis), SEQ ID NO:2 (SFV), SEQ ID NO:3 (EEE), and SEQ ID NO:4 (VEE). Consensus sequence, SEQ ID NO:5.
- [09] **FIGS. 5A-L.** BLAST alignment of VEE Nsp1-4 coding sequence with optimized coding sequence. VEE Nsp1-4 codon opt, SEQ ID NO:10; VEE Nsp1-4, SEQ ID NO:19; consensus, SEQ ID NO:20.

- [10] **FIG. 6.** Configuration of a double subgenomic promoter cassette. “CMV,” cytomegalovirus promoter; “5’,” 5’ untranslated region and sequences from the N-terminus of the Nsp1 coding region that are necessary for replication; “JR,” subgenomic promoter with the adjacent sequences (junction region); “Cap or Gly,” capsid or glycoprotein; “Nsp1-3,” non-structural proteins 1-3; “Nsp4,” non-structural protein 4; “3’,” 3’ untranslated region; “δR,” ribozyme cleavage site.

DETAILED DESCRIPTION OF THE INVENTION

- [11] In a split helper system, each structural protein is encoded in a separate defective helper (DH) cassette containing 5’ and 3’ cis elements necessary for replication (*see, e.g.,* US 2006/0292175). Expression of the encoded structural protein depends on the successful replication of the DH cassette by a replicase complex translated from the replicon and the subsequent transcription of subgenomic RNA. The replicase complex is translated as a single polypeptide chain which then undergoes sequential self-cleavage events, with different cleavage complexes then performing distinctive replication functions. The functional replicase complexes, particularly minus strand replicases which are necessary for the very first step of DH replication, are available for a limited time and location.
- [12] The invention provides strategies which can be used to increase the amount and availability of effective replicase complexes, thereby increasing the replication efficiency of DH transcripts and the productivity of PCL. The invention also provides strategies for minimizing the generation of replication competent viral particles (RCV). Though described below in connection with alphavirus-based packaging cell-line systems, the concepts and methods of the invention can readily be applied to other protein expression systems or viral packaging cell line systems to obtain commercially viable yields.

Double subgenomic promoter expression cassettes

- [13] In some embodiments, coding sequences for alphavirus nonstructural proteins 1-4 (nsp1-4) are placed under the control of a second subgenomic promoter in the same

expression cassette as a structural protein. Once induced by the replicon, these “double subgenomic promoter expression constructs” have the property of self-sustained replication but have virtually no expression without induction.

- [14] An expression cassette of the invention comprises two transcription units as well as a promoter and control elements needed for expression. Typical control elements include, but are not limited to, transcription promoters, transcription enhancer elements, chromatin insulator, transcription termination signals, polyadenylation sequences (located 3' to the translation stop codon), sequences for optimization of initiation of translation (located 5' to the coding sequence), translation termination sequences, 5' sequences required for nonstructural protein-mediated amplification, and 3' sequences required for nonstructural protein-mediated amplification.
- [15] Promoters for use in expression cassettes of the invention can be inducible or constitutive. Useful promoters include promoters include the CMV, MMTV, MoMLV, adenovirus VA1RNA promoters, and PolI promoters.
- [16] In some embodiments the first transcription unit is 5' to the second transcription unit. In other embodiments the first transcription unit is 3' to the second transcription unit. In either case, the first transcription unit comprises an alphavirus subgenomic promoter operably linked to a first coding sequence which encodes an alphavirus structural protein. The second transcription unit comprises another alphavirus subgenomic promoter operably linked to a second coding sequence which encodes alphavirus non-structural proteins 1-4. Elements of a transcription unit are “operably linked” when they are configured so as to perform their usual function; *i.e.*, expression of the structural protein and non-structural proteins is under the control of the subgenomic promoters.
- [17] Alphavirus subgenomic promoters (also referred to as “junction region promoters” or JR) are derived generally from the region between the nonstructural and structural protein open reading frames. Typically, an alphavirus subgenomic promoter contains a core sequence that provides most promoter-associated activity, as well as flanking regions that further enhance the promoter-associated activity. For example, the HR

strain Sindbis virus subgenomic junction region promoter typically begins at approximately nucleotide number 7579 and continues through at least nucleotide number 7612 (and possibly beyond). At a minimum, nucleotides 7579 to 7602 are believed to serve as the core sequence necessary for transcription of the subgenomic fragment.

- [18] The two subgenomic promoters in an expression cassette of the invention preferably are the same but can be derived from different alphaviruses. For example, at least one of the first and second subgenomic promoters is a Venezuelan encephalitis virus (VEE) subgenomic promoter, a Sindbis virus subgenomic promoter, an Eastern equine encephalitis virus (EEE) subgenomic promoter, or a Semliki Forest virus subgenomic promoter. In preferred embodiments both subgenomic promoters are Sindbis, VEE, SFV, or EEE promoters.
- [19] An “alphavirus structural protein” refers to either a capsid protein or a glycoprotein (which includes E1 and E2 and, where appropriate, E3). The capsid and glycoproteins can but need not be derived from the same type of alphavirus, *e.g.*, Sindbis virus, SFV, VEE, or EEE. Thus, in some expression cassettes at least one of the first and second alphaviruses is a Sindbis virus. In other expression cassettes, at least one of the first and second alphaviruses is a VEE virus.
- [20] Examples of capsid protein sequences are provided in SEQ ID NO:1 (Sindbis), SEQ ID NO:2 (SFV), SEQ ID NO:3 (EEE), and SEQ ID NO:4 (VEE). Examples of structural polyprotein sequences are provided in SEQ ID NO:36 (Sindbis; capsid, amino acids 1-264; E3, amino acids 265-328; E2, amino acids 329-751; 6K, amino acids 752-806; E1, amino acids 807-1245), SEQ ID NO:37 (SFV; capsid, amino acids 1-267; E3, amino acids 268-333; E2, amino acids 334-755; 6K, amino acids 756-815; E1, amino acids 816-1253), SEQ ID NO:38 (VEE; capsid, amino acids 1-275; E3, amino acids 276-334; E2, amino acids 335-756; 6K, amino acids 757-812; E1, amino acids 813-1254), and SEQ ID NO:39 (EEE; capsid, amino acids 1-260; E3, amino acids 261-323; E2, amino acids 324-743; 6K, amino acids 744-800; E1, amino acids 801-1241).

- [21] In some embodiments, described in more detail below, the capsid protein comprises a capsid protein which comprises one or more mutations which reduce autoproteolytic activity of the capsid protein (*e.g.*, His141Ala, Asp147Ala, Asp163A, Ser215Ala, and combinations thereof, numbered according to SEQ ID NO:1).
- [22] In some embodiments, the capsid protein and/or the glycoprotein are “hybrid” proteins. A hybrid protein contains at least one functional domain derived from a first alphavirus while the remaining portion of the protein is derived from one or more additional alphaviruses. For example, a hybrid capsid protein can comprise an RNA binding domain from the first alphavirus and an envelope interaction domain from a second alphavirus. Hybrid capsid proteins and glycoproteins are described in more detail in US 2006/0292175.
- [23] As is known in the art, nonstructural proteins include nsP1, nsP2, nsP3, and nsP4. Examples of nonstructural protein sequences are provided as SEQ ID NOS:6-9, respectively. A DNA sequence encoding VEE Nsp1-4 using optimized codons is provided in SEQ ID NO:10. One of ordinary skill in the art will realize that a wide variety of sequences which encode alphavirus nonstructural proteins, in addition to those disclosed herein, may be used in the present invention, and are therefore deemed to fall within the scope of the phrase “alphavirus nonstructural proteins.” For example, within one embodiment of the invention, due to the degeneracy of the genetic code, more than one codon may code for a given amino acid. Therefore, a wide variety of nucleic acid sequences which encode alphavirus nonstructural proteins may be generated. Within other embodiments of the invention, a variety of other nonstructural protein derivatives may be made, including for example, various substitutions, insertions, or deletions, the net result of which do not alter the biological activity of the alphavirus nonstructural proteins. Within the context of the present invention, alphavirus nonstructural proteins are deemed to be biologically active *in toto* if they promote the self-replication or trans-replication of the vector construct. Self-replication or trans-replication, which refers to replication of viral vector nucleic acids may be readily determined by metabolic labeling or RNase protection assays performed over a course of time.

- [24] Similarly, the capsid and glycoprotein proteins discussed above are not limited to polypeptides having the exact sequences disclosed herein. Alphaviral genomes are often in flux and contain several variable domains that exhibit relatively high degrees of variability between species and isolates. The terms “capsid,” “glycoprotein,” and “nonstructural protein(s)” encompass such proteins from any of the identified alphaviruses, as well as newly identified isolates, and subtypes of these isolates. In addition, amino acid sequences can be modified, particularly those in regions exhibiting high sequence homology.
- [25] Various nucleotide sequences can be used to encode the structural and nonstructural proteins. Optionally, as described below, sequences encoding nsP1-4 can be optimized to reduce the possibility of co-packaging into recombinant particles and to prevent recombination that could generate replication competent virus (RCVs).
- [26] In some embodiments expression cassettes of the invention comprise a selectable marker, such as Neo, SV2 Neo, hygromycin, puromycin, phleomycin, histidinol, or DHFR, which can be located at various points in the expression cassette as long as function of the transcription units is not disrupted.
- [27] Some expression cassettes of the invention comprise an internal ribosome entry site (IRES). The IRES can be placed between the 5' cis-replication element and subgenomic promoter, between two subgenomic promoters, or between subgenomic coding region and the 3' cis-replication element.
- [28] In another embodiment of the invention, all four non-structural proteins are produced from a single expression cassette, which has the advantage of more efficient assembly of replication complexes and increased expression of capsid and glycoproteins. *See Vokova et al., Virology 344, 315-27, 2006; and U.S. Patent 7,332,322.* In some embodiments of the invention, transcription of nsP1-3 is under the control of an inducible or constitutive promoter as described above, transcription of the capsid or the glycoprotein is under the control of a first subgenomic promoter, and transcription of nsP4 is under the control of a second subgenomic promoter. Optionally, Nsp1-4

sequences are codon-optimized (*see, e.g.*, SEQ ID NO:10). In some embodiments the capsid cassette has a puromycin marker and the glycoprotein cassette has no marker.

- [29] Examples of expression cassettes according to the invention are shown in **FIG. 1** and **FIG. 6**.

Host cells and packaging cell lines

- [30] Expression cassettes of the invention can be introduced into host cells. In some cases, the host cell comprises a first expression cassette, which comprises (a) a first transcription unit comprising a first alphavirus subgenomic promoter operably linked to a first coding sequence which encodes a structural protein of a first alphavirus; and (b) a second transcription unit comprising a second alphavirus subgenomic promoter operably linked to a second coding sequence which encodes non-structural proteins 1-4 of a second alphavirus. Some host cells contain two such expression cassettes; in these embodiments the first expression cassette encodes a capsid protein and the second expression cassette encodes the glycoprotein. Such host cells can be used as packaging cells, which can be used to make recombinant alphavirus particles.
- [31] Host cells can be any eukaryotic cell which is suitable for recombinant protein production. These include avian cells, insect cells (*e.g.*, C6/36, SF9), vertebrate, and mammalian cells. Examples of useful mammalian cell lines include Vero, MDBK, MDCK, MRC, NIH-3T3, BHK, PERC.6[®] (available from Crucell; *see* WO 01/38362 and WO 02/40665), EB cell lines, and HEK293 cells. Sources of avian cells include, but are not limited to, embryonic stem cells such as EBX[®] cells (Vivalis, FR), embryonic fibroblasts, and embryonic germ cells. Useful avian cells include the duck cell line AGE1.CR (ProBioGen). Other avian cell lines are disclosed, *e.g.*, in U.S. Patent 5,340,740; U.S. Patent 5,656,479; U.S. Patent 5,830,510; U.S. Patent 6,114,168; U.S. Patent 6,500,668; U.S. Patent 6,872,561; EP 0787180B; EP03291813.8; WO 03/043415; and WO 03/076601.
- [32] Expression cassettes of the invention can be introduced into host cells using methods well known in the art, including, but not limited to, microinjection, liposome-

mediated transfection, electroporation, and calcium phosphate precipitation. Alternatively, expression constructs of the invention can be incorporated into a polynucleotide delivery vehicle, such as a plasmid or a viral-based vector.

- [33] Once recombinant host cells, or “packaging cells,” have been constructed they can be used to produce recombinant alphavirus particles upon introduction of a replicon comprising an alphavirus packaging signal and encoding a protein of interest. The protein of interest is typically an antigen. Antigens can be derived from any of several known viruses, bacteria, parasites and fungi, as well as any of the various tumor antigens or any other antigen to which an immune response is desired. Furthermore, for purposes of the present invention, an “antigen” refers to a protein that includes modifications, such as deletions, additions and substitutions (generally conservative in nature), to the native sequence, so long as the protein maintains the ability to elicit an immunological response. These modifications may be deliberate, as through site-directed mutagenesis, or may be accidental, such as through mutations of hosts that produce the antigens. See US 2006/0292175.
- [34] Non-limiting examples of bacterial pathogens from which antigens can be derived include diphtheria, staphylococcus, cholera, tuberculosis, tetanus, *S. pneumoniae*, *S. agalactiae*, *S. pyogenes*, pertussis, meningitis, *N. gonorrhoeae*, *H. pyloris*, *H. influenza*, and *P. gingivalis*.
- [35] Non-limiting examples of viral pathogens include meningitis virus, influenza virus, rhinovirus, respiratory syncytial virus, parainfluenza virus, Picornaviruses, human Papilloma virus, retroviruses, and hepatitis viruses.
- [36] Tumor antigens include, but are not limited to, MART-1, gp100, tyrosinase, tyrosinase related proteins 1 and 2, β -catenin, MUM-1, CDK-4, caspase-8, KIA 0205, HLA-A2-R1701; MAGE-1, MAGE-2, MAGE-3, MAGE-12, BAGE, GAGE, NY-ESO-1, alpha-fetoprotein, telomerase catalytic protein, G-250, MUC-1, carcinoembryonic antigen, p53, Her-2-neu, triosephosphate isomerase, CDC-27, and LDLR-FUT). See also WO 91/02062, U.S. Pat. No. 6,015,567, WO 01/08636, WO 96/30514, U.S. Pat. No. 5,846,538 and U.S. Pat. No. 5,869,445.

Sequential amplification of DH cassettes involving mutant replicase complexes

- [37] Other embodiments involve sequential amplification of DH cassettes. These embodiments take advantage of cell lines which constitutively express VEE nonstructural proteins and various alphavirus nonstructural protein mutants that have specific defects in subgenomic transcription but not in DH/replicon replication. Thus, these mutant nonstructural protein replicase complexes can be constitutively expressed to amplify the DH, but will not produce subgenomic transcripts coding alphavirus structural protein. Upon induction of the replicon, the amplified DH RNA is further amplified by wild type nsps from replicon. The wild type replicase complexes also produce subgenomic transcripts and lead to the expression of structural proteins. Several of these mutants show over hundreds-fold decrease in subgenomic RNA transcription or particle production compared with wild type nonstructural proteins, providing a powerful inducible system.
- [38] One useful mutant is the nsP2 cleavage mutant. Alphavirus minus strand replication requires uncleaved P123 together with correctly cleaved nsP4 and is shut off approximately 4 hours after infection (Kaariainen and Ahola, *Prog. Nucleic Acid Res. Mol. Biol.* 71, 187-222, 2002). Thus, mutations at well-conserved alphavirus nsps cleavage sites will not be cleaved and the mutant replicase should be available for a longer time compared with wild type replicase. In addition, in Sindbis and SFV such mutants have very low level of subgenomic RNA transcription (Lemm *et al.*, *EMBO J.* 13(12), 2925-34, 1994, Shirako & Strauss, *J Virol.* 68(3), 1874-85, 1994, Kim *et al.*, *Virology* 323(1), 153-63, 2004). These cleavage sites are well conserved in VEE virus (**FIG. 2**), and several different mutants (such as mutations at nsP1/nsP2 and nsP2/nsP3 cleavage sites) can be made.
- [39] The other mutations include R331A and R332A mutations in the Sindbis Nsp4 protein (Li & Stollar, *Proc. Natl. Acad. Sci. USA* 101, 9429-34, 2004), which abolish the subgenomic promoter binding/transcription ability of replicase complexes but retain the ability to amplify viral/DH genome. These mutations are highly conserved

among different alphavirus families. Alternatively, deletions or other substitutions at R331, R332 (numbered according to SEQ ID NO:9) or both can be used. These mutant Nsp1-4 replicase complexes can be expressed from same DH transcript (such as linked to an IRES sequence) or can be expressed in cell substrate from a separate transcript cassette. Suitable substitutions include:

at R331: glutamine, leucine, serine, asparagine, glutamic acid, lysine, threonine, glycine, methionine, tryptophan, aspartic acid, histidine, phenylalanine, tyrosine, cysteine, isoleucine, proline, alanine, or valine; or

at R332: glutamine, leucine, serine, glutamic acid, lysine, threonine, glycine, methionine, tryptophan, aspartic acid, histidine, phenylalanine, tyrosine, cysteine, isoleucine, proline, alanine, or valine.

- [40] Several Sindbis and SFV temperature sensitive mutants show specific defects in subgenomic RNA synthesis (Lulla, *Virology* 80(6), 3108-11, 2006; Lastarza, *J. Virol.* 68(9), 5781-91, 1994). Such mutants also are useful for making Sindbis-, VEE-, and SFV-based PCL.
- [41] Optionally, alphavirus mutant nsp1-4 codons can be optimized to reduce the possibility of co-packaging into recombinant particles and to prevent recombination that could generate replication competent virus (RCVs). A DNA sequence encoding VEE Nsp1-4 using optimized codons is shown in SEQ ID NO:13.
- [42] Each of the strategies described above can be used in conjunction with one or more of the strategies described below.

Minimizing the risk of generating RCV using capsid autoproteolysis mutants

- [43] Generation of RCV (replication competent viral particles) is a potential problem for the application of alphavirus based replicon particles. The probability of recombination is greatly reduced by dividing the defective helpers into two separate

cassettes because multiple switches would be required to produce an infectious RNA. However, it is conceivable that during large-scale production, RCV could be generated. The invention provides capsid autoproteolytic mutants which can be used to further reduce the possibility of generating RCV, providing an additional safeguard for the production of alphavirus based replicon particles. Using this strategy it is virtually impossible to generate wild type RCV.

- [44] Alphavirus structural proteins are translated *in vivo* from a 26S subgenomic RNA as a polyprotein that is processed both cotranslationally and posttranslationally. The capsid is postulated to be a serine protease that release itself from the N terminus of the nascent polyprotein by autoproteolysis. Several Sindbis virus autoproteolysis mutants have been identified (*e.g.*, His141, Asp147, and Ser215) and all were lethal to the virus (Hahn & Strauss, 1990, *J. Virol.* 64, 3069-73, 1990). In a double helper system, the capsid is artificially separated from structural polyprotein, and the autoproteolysis function is probably not critical for alphavirus particle production. Thus, capsid autoprotease mutations can be used to minimize the risk of generating RCV. These mutations include changes at His141 (*e.g.*, His141Ala), Asp147 (*e.g.*, Asp147Ala), Asp163 (*e.g.*, Asp163Ala), Ser215 (*e.g.*, Ser215Ala), numbered according to SEQ ID NO:1, and combinations thereof. Other substitutions include:

at His141: glutamine, leucine, serine, arginine, glutamic acid, lysine, threonine, glycine, methionine, tryptophan, aspartic acid, histidine, phenylalanine, tyrosine, cysteine, isoleucine, proline, or valine;

at Asp147: glutamine, leucine, serine, arginine, glutamic acid, lysine, threonine, glycine, methionine, tryptophan, aspartic acid, histidine, phenylalanine, tyrosine, cysteine, isoleucine, proline, or valine;

at Asp163: glutamine, leucine, serine, arginine, glutamic acid, lysine, threonine, glycine, methionine, tryptophan, aspartic acid,

histidine, phenylalanine, tyrosine, cysteine, isoleucine, proline, or valine; or

at Ser215: glutamine, leucine, serine, arginine, glutamic acid, lysine, threonine, glycine, methionine, tryptophan, aspartic acid, histidine, phenylalanine, tyrosine, cysteine, isoleucine, proline, or valine.

- [45] Changes also include deletions (*e.g.*, Δ His141, Δ Asp147, Δ Asp163, Δ Ser215, and Δ Trp264 and insertions. Capsid proteins for use in the invention can comprise one, two, three or more such mutations.
- [46] In some embodiments, mutations are introduced at the capsid cleavage sites, either at the 3' end of capsid protein (Trp) or at the 5' end of the glycoprotein serine residue (Met-Ser) or in combinations (see **FIG. 3**). Deletions of key residues (*e.g.*, capsid W264 and Gly Ser2) also can be made. Because capsid autocatalytic sites are conserved among different strains, this strategy can be used for a variety of alphavirus-based systems (*e.g.*, Sindbis, SFV, and VEE; see **FIG. 4**).
- [47] All patents, patent applications, and references cited in this disclosure are expressly incorporated herein by reference. The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples, which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

EXAMPLE 1

Improved expression of heterologous protein under the control of a DH cassette comprising two subgenomic promoters

- [48] DH expression cassettes which encode alphavirus nsP1-4 under a subgenomic promoter have the property of self-sustained replication once induced by replicon. Preliminary results using green fluorescent protein (GFP) as reporter system showed that such constructs have provide a 3-4 fold increase in the percentage of GFP

positive cells and a similar fold of increase in mean fluorescence intensity (see Table 1).

Table 1.

vector	mean GFP value	% cells positive
single subgenomic promoter	85	0.8
double subgenomic promoter (nsp1-neo fusion protein)	307	2.68
double subgenomic promoter (IRESneo)	206	2.12

EXAMPLE 2

Capsid mutants

- [49] Site-directed mutagenesis was used to generate the following capsid mutants: His141Ala, Asp147Ala, Ser215Ala, Trp264Ala and various of compound mutants. Mutagenesis was confirmed by sequencing. Mutations were incorporated into two split cassette RNAs (VCR-DH-Scap, VCR-DH-Sgly) to test whether they interfere with recombinant alphavirus particle production. *In vitro* Sp6-transcribed RNAs (wild-type or mutant capsid RNA, glycoprotein RNA, and green fluorescent protein (GFP) replicon RNA) were electroporated into BHK-v cells. Twenty hours later the supernatants were harvested and used to infect naïve BHK cells. Eighteen hours later, FACS analysis was performed to determine the titer of replicon GFP particles. The results are shown in Tables 2 and 3.

Table 2.

capsid	Replicon particle titer (IU/ml)
wild-type capsid	1.43E8
capsid H141A mutant	1.63E8
capsid D147A mutant	9.65E7
capsid S215A mutant	1.14E8
capsid W264A mutant	2.48E6

Table 3.

capsid	Replicon particle titer (IU/ml)
wild-type capsid	3.17E8
capsid H141A+D163A	3.08E8
capsid H141+S215A	3.84E8
capsid D163+S215A	3.76E8
capsid H141A +D163A+S215A	4.89E8

- [50] The results show that the H141A, D147A, and S215A mutations do not affect the replicon particle titer, and the various compound mutants have a comparable level of particle production compared with wild type.

EXAMPLE 3

Comparison of single and double subgenomic promoter expression constructs

- [51] Using green fluorescent protein (GFP) as the protein of interest, this example demonstrates that double subgenomic promoter expression constructs of the invention produce more protein of interest than constructs that employ only one subgenomic promoter.
- [52] BHK-v cells were propagated on 6-well plates and maintained in Dulbecco's modified Eagle medium (DMEM) (Cellgro, Vermont, VA) supplemented with 10% gamma-irradiated fetal bovine serum, 1% antibiotic (penicillin and streptomycin), and 1% sodium pyruvate (Cellgro). VEE defective helper plasmid DNAs (VCP-nf3.1-

GFP which codes for single subgenomic GFP transcript, and VCP-PsubGFP-PsubNsp1-4 which codes for double subgenomic transcripts GFP and Nsp1-4) were transfected into BHK-v cells using LT1 transfection agent (Mirus Bio) at 2 µg per well. Cells were expanded 48 hours post-transfection, and Geneticin (G418 sulfate, a neomycin sulfate analog; Cellgro) was added at 600 µg/ml in growth medium for selection and maintenance of stable recombinant BHK-v cell lines. Pools were collected from both transfections and propagated on 6-well plates. VCR-Chim2.1-gp120 replicon particles were used to infect the pool at MOI 5, and cells were collected 24 hours after infection. FACS analysis was performed to determine the GFP positive ratio. The results of duplicate (or quadruplicate) experiments for each construct are shown in Table 3.

- [53] The construct “VCP-nf3.1-GFP” listed in Table 3 contains only one subgenomic promoter. The construct “VCP-Pgfp-Pnsp-IRESneo” is the middle construct in FIG. 1. “VCP” stands for VEE CMV promoter plasmid; “Pgfp” stands for subgenomic promoter with GFP coding region; “Pnsp” stands for subgenomic promoter with Nsp1-4 coding region; and “IRES” stands for EMCV IRES driven neomycin.

Table 3.

Construct	Mean (not-induced) /SD / % (+)	Mean (induced) / SD / %
VCP-nf3.1-GFP	26.45 / 6.68 / 0.17	50.36 / 83.38 / 0.61
	28.34 / 8.64 / 0.17	149.60 / 476.88 / 0.61
	1.15 / 0.31 / 0	77.07 / 148.31 / 1.52
	1.32 / 0.58 / 0	64.56 / 87.55 / 0.60
VCP-Pgfp-Pnsp-IRESneo	235.52 / 977.35 / 0.71	249.03 / 450.47 / 1.92
	115.11 / 163.18 / 1.0	240.43 / 449.26 / 2.91
	105.35 / 158.29 / 0.78	281.36 / 540.27 / 2.65
	151.60 / 368.25 / 0.80	324.42 / 541.99 / 1.86
	266 / 397 / 0.23	394 / 1116 / 2.55
	253 / 208 / 1.03	319 / 447 / 4.65
	436 / 810 / 0.27	198 / 380 / 1.34
	471 / 513 / 1.13	451 / 541 / 3.59
VCP-nf3.1Pgfp-Pnsp	130 / 281 / 0.21	115 / 422 / 1.29
	448 / 837 / 0.5	299 / 989 / 3.65
	54 / 47 / 0.17	144 / 711 / 1.11
	247 / 492 / 0.41	269 / 789 / 2.45

CLAIMS

1. An expression cassette comprising a first and a second transcription unit, wherein:
 - (a) the first transcription unit comprises a first alphavirus subgenomic promoter operably linked to a first coding sequence which encodes a structural protein of a first alphavirus; and
 - (b) the second transcription unit comprises a second alphavirus subgenomic promoter operably linked to a second coding sequence which encodes non-structural proteins 1-4 of a second alphavirus.
2. The expression cassette of claim 1 wherein the first transcription unit is 5' to the second transcription unit.
3. The expression cassette of claim 1 wherein the first transcription unit is 3' to the second transcription unit.
4. The expression cassette of claim 1 wherein the structural protein is a glycoprotein.
5. The expression cassette of claim 1 wherein the structural protein is a capsid protein.
6. The expression cassette of claim 1 wherein the structural protein is a hybrid capsid protein.
7. The expression cassette of claim 1 wherein the structural protein is a capsid protein which comprises a mutation which reduces autoproteolytic activity of the capsid protein.

8. The expression cassette of claim 7 wherein the mutation is selected from the group consisting of His141Ala, Asp147Ala, Asp163A, Ser215Ala, Δ His141, Δ Asp147, Δ Asp163, Δ Ser215, and Δ Trp264 and combinations thereof, numbered according to SEQ ID NO:1.

9. The expression cassette of claim 1 wherein at least one of the first and second alphaviruses is a Sindbis virus.

10. The expression cassette of claim 1 wherein at least one of the first and second alphaviruses is a Venezuelan equine encephalitis (VEE) virus.

11. The expression cassette of claim 1 wherein at least one of the first and second subgenomic promoters is a VEE subgenomic promoter, a Sindbis virus subgenomic promoter, an Eastern equine encephalitic (EEE) subgenomic promoter, or a Semliki Forest virus subgenomic promoter.

12. The expression cassette of claim 1 further comprising a selectable marker.

13. The expression cassette of claim 1 further comprising an internal ribosome entry site (IRES).

14. The expression cassette of claim 1 wherein the second coding sequence comprises SEQ ID NO:2.

15. The expression cassette of claim 1 wherein the second transcription unit comprises a mutation in at least one of R331 or R332 in Nsp4 numbered according to SEQ ID NO:9.

16. A host cell comprising a first expression cassette, wherein the first expression cassette comprises:

(a) a first transcription unit comprising a first alphavirus subgenomic promoter operably linked to a first coding sequence which encodes a structural protein of a first alphavirus; and

(b) a second transcription unit comprising a second alphavirus subgenomic promoter operably linked to a second coding sequence which encodes non-structural proteins 1-4 of a second alphavirus.

17. The host cell of claim 16 further comprising a second expression cassette, wherein the second expression cassette comprises:

(c) a third transcription unit comprising a third alphavirus subgenomic promoter operably linked to a third coding sequence which encodes a second structural protein of a third alphavirus; and

(d) a fourth transcription unit comprising a fourth alphavirus subgenomic promoter operably linked to a fourth coding sequence which encodes non-structural proteins 1-4 of a fourth alphavirus,

wherein the first structural protein is a capsid protein and the second structural protein is a glycoprotein.

18. The host cell of claim 16 wherein at least one of the first and second alphaviruses is a VEE virus, a Sindbis virus, or a Semliki Forest virus.

19. The host cell of claim 16 further comprising a replicon which comprises a packaging signal and which encodes a protein of interest.

20. A method of making a packaging cell for producing recombinant alphavirus particles, comprising steps of:

(a) transfecting a first expression cassette into a suitable host cell, wherein the first expression cassette comprises:

(i) a first transcription unit comprising a first alphavirus subgenomic promoter operably linked to a first coding sequence which encodes a structural protein of a first alphavirus; and

(ii) a second transcription unit comprising a second alphavirus subgenomic promoter operably linked to a second coding sequence which encodes non-structural proteins 1-4 of a second alphavirus; and

(b) transfecting a second expression cassette into the suitable host cell, wherein the second expression cassette comprises:

(i) a third transcription unit comprising a third alphavirus subgenomic promoter operably linked to a third coding sequence which encodes a second structural protein of a third alphavirus; and

(ii) a fourth transcription unit comprising a fourth alphavirus subgenomic promoter operably linked to a fourth coding sequence which encodes non-structural proteins 1-4 of a fourth alphavirus.

21. A method of producing recombinant alphavirus particles, comprising introducing into a host cell a replicon which comprises a packaging signal and which encodes a protein of interest, wherein the host cell comprises two expression cassettes, wherein:

(a) a first expression cassette comprises:

(i) a first transcription unit comprising a first alphavirus subgenomic promoter operably linked to a first coding sequence which encodes a structural protein of a first alphavirus; and

(ii) a second transcription unit comprising a second alphavirus subgenomic promoter operably linked to a second coding sequence which encodes non-structural proteins 1-4 of a second alphavirus; and

(b) a second expression cassette comprises:

(i) a third transcription unit comprising a third alphavirus subgenomic promoter operably linked to a third coding sequence which encodes a second structural protein of a third alphavirus; and

(ii) a fourth transcription unit comprising a fourth alphavirus subgenomic promoter operably linked to a fourth coding sequence which encodes non-structural proteins 1-4 of a fourth alphavirus.

22. An expression cassette comprising three transcription units, wherein:

(a) a first transcription unit comprises a promoter operably linked to a coding sequence which encodes alphavirus structural proteins 1-3;

(b) a second transcription unit comprises a first alphavirus subgenomic promoter operably linked to a coding sequence which encodes an alphavirus structural protein; and

(c) a third transcription unit comprises a second alphavirus subgenomic promoter operably linked to a coding sequence which encodes an alphavirus nonstructural protein 4.

23. The expression cassette of claim 22 wherein the structural protein is a glycoprotein.

24. The expression cassette of claim 22 wherein the structural protein is a capsid protein.

25. The expression cassette of claim 22 wherein the structural protein is a hybrid capsid protein.

26. The expression cassette of claim 22 wherein the structural protein is a capsid protein which comprises a mutation which reduces autoproteolytic activity of the capsid protein.

27. The expression cassette of claim 26 wherein the mutation is selected from the group consisting of His141Ala, Asp147Ala, Asp163A, Ser215Ala, Δ His141, Δ Asp147, Δ Asp163, Δ Ser215, and Δ Trp264 and combinations thereof, numbered according to SEQ ID NO:1.

28. A host cell comprising a first expression cassette of claim 22.

29. The host cell of claim 28 further comprising a second expression cassette of claim 22, wherein the first expression cassette encodes an alphavirus glycoprotein and the second expression cassette encodes an alphavirus capsid protein.

30. The host cell of claim 29 further comprising a replicon which comprises a packaging signal and which encodes a protein of interest.

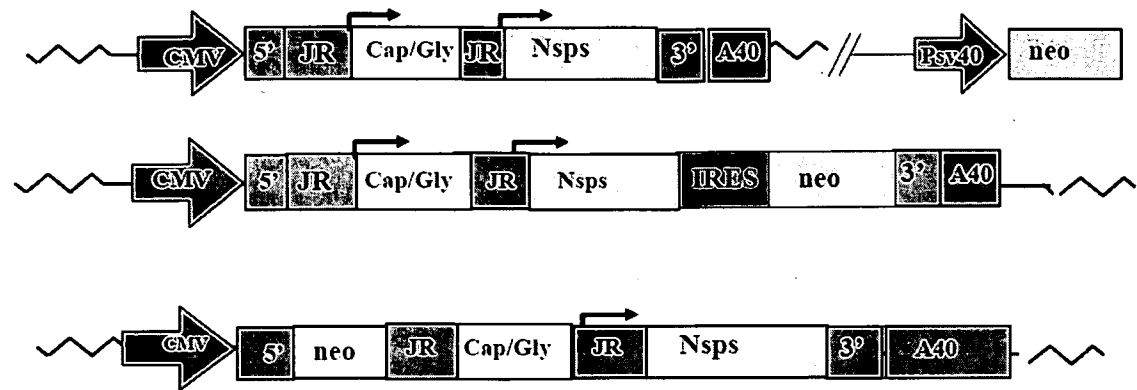
31. A method of making a packaging cell for producing recombinant alphavirus particles, comprising steps of:

(a) transfecting a first expression cassette of claim 22 into a suitable host cell, wherein the first expression cassette encodes an alphavirus glycoprotein; and

(b) transfecting a second expression cassette of claim 22 into the host cell, wherein the second expression cassette encodes an alphavirus capsid protein.

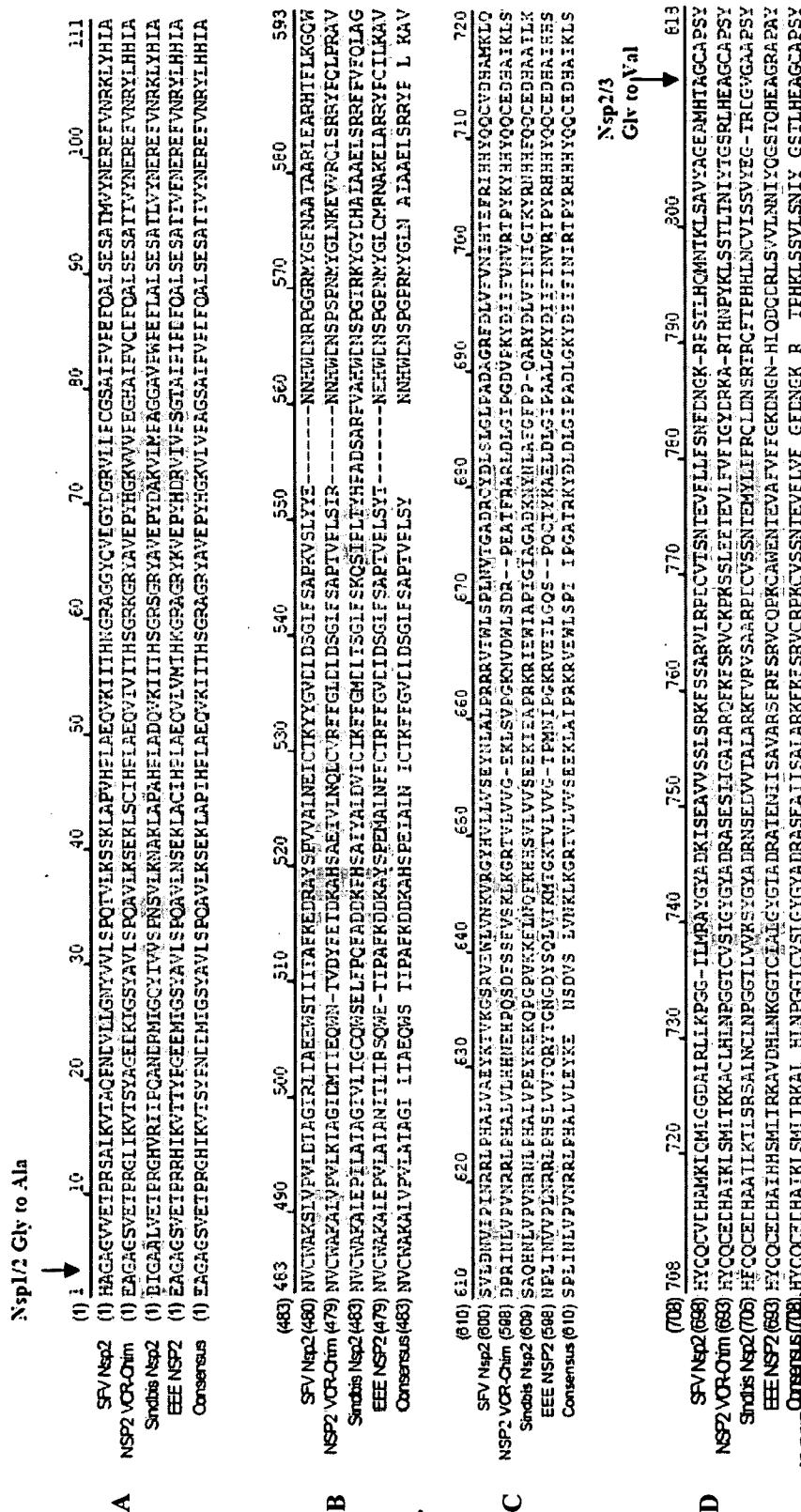
32. A method of producing recombinant alphavirus particles, comprising introducing into the host cell of claim 29 a replicon which comprises a packaging signal and which encodes a protein of interest.

FIG. 1



2/15

FIG. 2



3/15

FIG. 3

3' end DH-Scap: KGKTIKTTPEGTEEW↓
 5' end of DH-Sgly: MSAAPLVTAMCLLGNVSF
 RCV: KGKTIKTTPEGTEEW↓SAAPLVTAMCLLGNVSF

FIG. 4

	(1)	1	10	20	30	40	50	60	76	
Sindbis Virus Capsid	(1)	-MNRGFENMLGRRPFAPATAMWRPRRRRCAPMPARNGLASCIQQLTAVSALVIGQATRECFERFRPPR-----								
SFV Capsid	(1)	MNYIPTQTFFYGRWRPFPAAAFWFLQATFVAVVP-DFQAQCMQQLISAVNALTRQNAIAFAPFPKPKKKK-----								
EEE Capsid	(1)	MEFYPTLNYPFMAPINPMAYRDPNPFRRRRWRPFRP--PLAAQIEDLRRIANLTLKQRAFNFEAGFPAPKKK-----								
VEE Capsid	(1)	-----MEFPQPMYPMQPMFYRNPFAPRRFWFPRTPDPELANQVQELTRSMANLTFKQRRDAEFEGFPAPKKPKREAP								
Consensus	(1)	M PTFNF PRRPIPPAYR FP RRR AFMRP FLAAQIQQLTRAVANLTIKQRA AEF GEPFKKKK								

	(77)	77	90	100	110	120	130	140	152	
Sindbis Virus Capsid	(71)	-----CKKQAFKQFPKPKPKTKQEKKKKQPAKPK-----PGKRQRMALKLEADRLEFVKNEDGDEVIGHALAMEGK								
SFV Capsid	(72)	-----TTKEPKTKQPKKINGKTQQQKKKQADKKKKKPGKFRMCMKIENDCIFEVKHEG-KVTGYACLVGDK								
EEE Capsid	(70)	-----KFAESLSLRRKKRPFPAPAKKQKRPK-----PGKRQRMCMKLESEKTFPIMLNG-QVNGYACVVGGK								
VEE Capsid	(72)	QKQKGGGQCKKKKNQKKKAKTGFPNPKAQSGNKKKPNKKPGKRQRMVMKLESEKTFPIMLEG-KINGYACVVGGK								
Consensus	(77)	Q KEKEKQ FKKKK KTCNEKKKQKNKPK KPGKRQRMCMKLESEKTFPIMLEG KVNGYACVVGGK								

		H141	D147		D163				
	(153)	153	160	170	180	190	200	210	228
Sindbis Virus Capsid	(136)	VMKPLHVKGTTIDHPVLSKLKFTKSSAYDMEFAQLFVNMRSEAFYTSSEHPEGFYNNWHGAVQYSGGRFTIPRGVGG							
SFV Capsid	(140)	VMKPAHVKGVIDNADLAKLAFKRSKYDLECAQIFVHMRSDASKYTHEKPEGHYNNWHGAVQYSGGRFTIPTGAGK							
EEE Capsid	(132)	VFKPLHVEGRIDNEQLAAIKLKASIIDLEYGDFVQCMKSDTLQYTSKPPGFFYNWHGAVQYENNRFTVPRGVGG							
VEE Capsid	(147)	LFRPMHVEGRIDNDVLAALKKASKYDLEYADVPQNMRAITFKYTHEKPPQGYSSWHGAVQYENGRFTVPRGVGA							
Consensus	(153)	VMKPLHVKGKIDNDVLAALKFKKASKYDLEYAQVFNMRSDTFKYTSEKPEGFYNNWHGAVQYSGGRFTIPRGVGG							

			S215			281		W264
	229							
Sindbis Virus Capsid	(212)	EGDSGRPIPDNSGRVVAIVLGGADEGTRTALS VVTWNSKGKTIKTTPEGTEEW						
SFV Capsid	(216)	EGDSGRPIFDNKGRVVAIVLGGANEGSRTALS VVTWN-KDMVTRVTPEGSEEW						
EEE Capsid	(208)	EGDSGRPILDNKGRVVAIVLGGVNEGSRTALS VVTWNQKGVTVKDTPEGSEPW						
VEE Capsid	(223)	KGDSGRPILDNQGRVVAIVLGGVNEGSRTALS VVMWNEKGVTVKTYTPENCEQW						
Consensus	(229)	KGDSGRPILDNKGRVVAIVLGGVNEGSRTALS VVTWN KGVTVK TPEGSEEW						

[illegible]

VEE Nsp1-4 Codon Opt(1262) GGGCTTCAGGCGCCACAAAATCACAAGTATCTACAAAAGGCTTGACACGCGAAACAAATAAAGTGAATTCGACATTCCTTTTGTCTGCGCAAGAAAT
VEE Nsp1-4(1262) GGGCTTTTAGAGGCACAGATAACATCTATTTATAGCGCCCGATACCCAAACCATCATCAAGTGAACAGCGATTTCCACTCATCTGCTGCTGCCCAGGAT
Consensus(1262) GGGC TT AG G CACAA AT ACA TAT TA AA G CC GA AC CAAAC AT AT AAGTGA CGA TT CACTC TT GT CTGCC AG AT

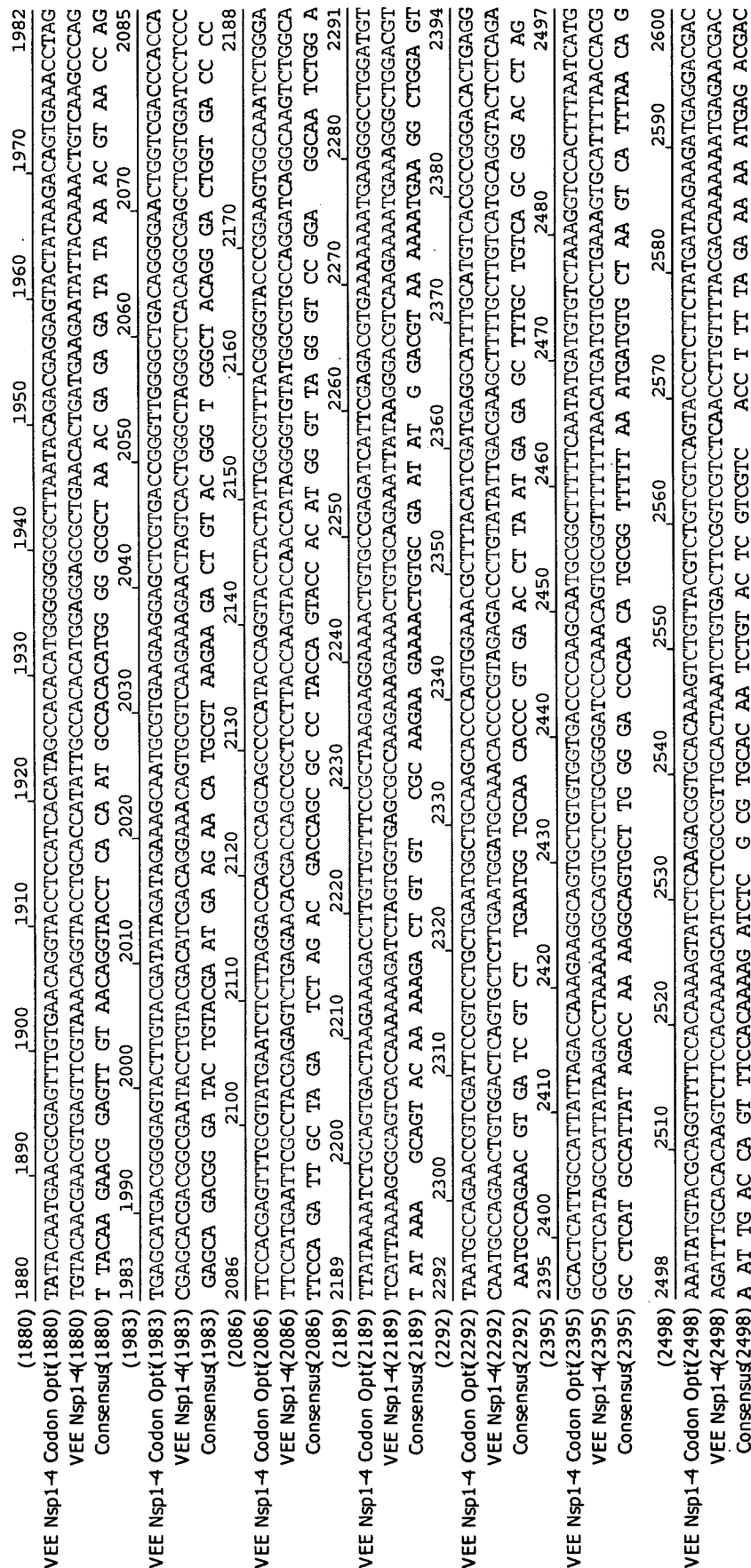
VEE Nsp1-4 Codon Opt(1365) AGGTAGCAACACTCTGGAATTCGGCTCAGGACCGAATACGAAAATGCTCAAGAAACACAAGGAACCCCTCTCTTTTGATCACGGCAGAGGACGTGCAGGAA
VEE Nsp1-4(1365) AGCAGTAACACATTTGGATCGGGCTGAGAACAAAGATCAGGAAAATGTTAGAGGAGCAAGGACCGTCACTCTCATTACCCGCGAGGACGTACAGAA
Consensus(1365) AGG AG AACAC TGA ATCGGGCT AG AC AGAAT G AAAATG T GA GA CACAAGGA CC TC CCT T AT AC GC GAGGACGT CA GAA

VEE Nsp1-4 Codon Opt(1468) GC AAAATGCGCGCAGACGAAAGCTAAGAAGTCCGGGAGCGGAAGAACTGCGAGCGGCTCTGCCACCCCTGGCGGTGACGTGAGGAACCCACCTTGGAGG
VEE Nsp1-4(1468) GCTAAGTGGCAGCCGATCAGGCTAAGAGGTGCGTGAAGCCGAGGAGTTGCGGCGAGCTTACCACTTTGGCAGCTGATGTTGAGGAGCCCACTCTTGGAA
Consensus(1468) GC AA TGGCG GC GA GA GCTAA GA GT CG GA GC GA GA TGGC GC GCTCT CCACC TGGC GCTGA GT GAGGA CCCAC CTGGA G
(1571) 1571 1580 1590 1600 1610 1620 1630 1640 1650 1660 1673

VEE Nsp1-4 Codon Opt(1571) CAGATGTGATCTGATGCTTCAGGAAGCGGGAGCCGGTCCGTGGAGACCCCGAGGGTTGATCAAAGTGAACATCCTATGCAGGAAGACAAAATCGGGAG
VEE Nsp1-4(1571) CCGATGTCGACTTGATGTTACAAGAGGCTGGGCGGCTCAGTGGAGACACCTCGCTGGCTTGATAAAGTTACCAGTCACTGCTGGCAGGACAAGATCGGCTC
Consensus(1571) C GATGT GA TGATG T CA GA GC GG GCGCG TC GTGAGAC CC G GG TTGAT AA GT AC CTA GC GG GA GACAA ATCGG
(1674) 1674 1680 1690 1700 1710 1720 1730 1740 1750 1760 1776

VEE Nsp1-4 Codon Opt(1674) CTATGCCGTTTTGTCCCCACAGGCTGTGCTTAATCTGAGAAACTCTCTGCATACATCCCTCGCGGAACAGGTCAITGTGATCACACATAGCGGACGGAAG
VEE Nsp1-4(1674) TTACGCTGTGCTTCTCCGAGGCTGTACTCAAGAGTGAATAATTTATCTTGATCCACCCCTCTGCTGAACAAGTCAAGTATGATACACACTCTGCGCCGAAAA
Consensus(1674) TA GC GT T TC CC CAGGCTGT CT AA TGA AAA T TC TGAT CA CC CTCGC GAACA GTCAT GTGAT ACACA GG CG AA
(1777) 1777 1790 1800 1810 1820 1830 1840 1850 1860 1879

VEE Nsp1-4 Codon Opt(1777) GGCAGGTATGCTGTGGAACCCCTATCAGGCAAAAGTAGTCTGTGCCGAGGGACAGCCCATTCGGTGCAGATTTCCAGGCACTCAGCGAATCCGCCCAATCG
VEE Nsp1-4(1777) GGGCGTTATGCGGTGGAACCATACCATGCTAAAGTAGTGTGCCAGAGGACATGCAATACCGTCCAGGACTTTCAAGCTCTGAGTGAAGTGCACCAATG
Consensus(1777) GG G TAIGC GTGGAACC TA CA GG AAAGTAGT GTGCC GAGGGACA GC AT CC GT CA GA TT CA GC CT AG GAA GCCAC AT G



VEE Nsp1-4	Codon Opt(2601)	2601	2602	2630	2640	2650	2660	2670	2680	2690	2703
VEE Nsp1-4(2601)	TAA	CCCAAGAAACCAAGATCGTGA	TGACACCACTGGAGTACAAAACCTAAACAGGATGACCTGATTTCTGACTGCTTTAGGGCTGGGTTAAGCAACTT								
VEE Nsp1-4(2601)	GAA	TCCGAAAGAGACTAAGATTGTGATTGACACTACCGGCAGTACCAAACTTAAGCAGGAGATCTCATTTCTACTTGTTCAGAGGGTGGGTGAAGCAGTTG									
Consensus(2601)	AA	CC	AAAGA AC AAGAT GTGATTGACAC AC GG AGTAC AAACCTAA CAGGA GA CT ATTCT AC TG TT AG GG TGGGT AAGCA T								
(2704)	2704	2710	2720	2730	2740	2750	2760	2770	2780	2790	2806
VEE Nsp1-4	Codon Opt(2704)	CAGATCGATTATAAAGGAAACGAGATATGACTGCGCTGCCAGCCAGGCGCTGACACCGAAAGGTGTGACGTGTGCGATACAAAAGTGAACGAGAACCCCC									
VEE Nsp1-4(2704)	CAAATAGATTACAAAGGCAACGAAATATGACGGCAGCTGCCTCTCAAGGGCTGACCCCTAAAGGTGTGATGCCGTTCGGTACAAAGGTGAATGAAATCCCTC										
Consensus(2704)	CA	AT	GAITTA AAAGG AACGA AT ATGAC GC GCTGCC CA GG CTGAC CG AAAGGTGTGTA GC GT CG TACAA GTGAA GA AA CC C								
(2807)	2807	2820	2830	2840	2850	2860	2870	2880	2890		2909
VEE Nsp1-4	Codon Opt(2807)	TCATATGCCCCCTAGCTCTGAGCAGCTCAATGTCTGTGTGACAAAGCACTGAGATCGAATCTGTGTGAAGACATTGGCGGGGATCCCTGGATTAAGACTCTCAC									
VEE Nsp1-4(2807)	TGTACGCACCCACCTCAGAACATGTGAAGTCCCTACTGACCCGACGGAGGACCGCATCTGTGTGAAAAACACTAGCCGGCGACCATGGATGATAAAAACTGAC										
Consensus(2807)	T	TA	GC CC AC TC GA CA GT AA GT CT TGAC G AC GAGGA CG ATCTGTGGAA ACA T GCGGG GA CC TGGAT AA AC CT AC								
(2910)	2910	2920	2930	2940	2950	2960	2970	2980	2990	3000	3012
VEE Nsp1-4	Codon Opt(2910)	CGCTAAGTATCAGGCAACTTTACTGCAACCATCAGGAGTGGCAGGCCGAGCACGATGCTATTATGCGACACATTTCTTGAGCGGCCCGACCCCTACGGATGTG									
VEE Nsp1-4(2910)	TGCCAAGTACCTGGGAATTCATGTCCACGATAGAGGAGTGGCAAGCAGAGCATGATGCATCATGAGGCACATCTTGGAGAGACCGGACCTACCGACGTC										
Consensus(2910)	GC	AAGTA CC GG AA TT	ACTGC AC AT GAGGAGTGGCA GC GAGCA GATGC AT ATG G CACAT T	GAG G CC GACCTAC GA GT							
(3013)	3013	3020	3030	3040	3050	3060	3070	3080	3090	3100	3115
VEE Nsp1-4	Codon Opt(3013)	TTTCAAAACRAGGCCAATGTCTGCTGGGGAAGGCACTGGTGCCTGTCTCTGAAGACTGCGGGCATTCGACATGACACCGAGCAGTGGAAATACGGTTGACTACT									
VEE Nsp1-4(3013)	TTCCAGAATAAGGCAACGTGTGTTGGGCCAAGCTTTAGTCCCGGTGCTGAAGACCGCTGGCATAGACATGACCATGAACATGGAACACTGTGGATTATT										
Consensus(3013)	TT	CA	AA	AAGC AA GT TG TGGC AAGGC T GTGCC GT CTGAAGAC GC GGCAT GACATGACCAC GA CA TGGAA AC GT GA TA T							
(3116)	3116	3130	3140	3150	3160	3170	3180	3190	3200		3218
VEE Nsp1-4	Codon Opt(3116)	TTGAGACTGATAAGGCCACAGTGCAGAGATTGTTTGAATCAGCTGTGTGAGATTCTTCGGACTGGATCTGGATAGTGGCCTGTTTCTGCACCTACCGT									
VEE Nsp1-4(3116)	TTGAAACGGACAAAGCTCACTACGACAGATAGTATTGAACCAACTATGCGTAGGTTCTTTGGACTTCGACTTCGACTTCGGTCTATTCTTCGCCACTGT										
Consensus(3116)	TTGA	AC GA AA GC CAC GCAGAGAT GT TTGAA CA CT TG GIGAG TICTT GGACT GATCTGGA GG CT TTTCTGCACC AC GT									
(3219)	3219	3230	3240	3250	3260	3270	3280	3290	3300	3310	3321
VEE Nsp1-4	Codon Opt(3219)	TCCGCTGTCCATCAGAAACAAATCATTGGGACAACAGTCCATCCCCCAATATGTTATGTTCTGGAATAAGGAGGTGGTGGCGCAGCTGTCCCGCGGTATCCACAG									
VEE Nsp1-4(3219)	TCCGTTATCCATTAGGAATAATCACTGGGATAACTCCCGTCGCTTAACATGTACGGGTGGAATAAAGAAGTGGTCGTCAGCTCTCTCGCAGGTACCCACAA										
Consensus(3219)	TCCG	T	TCCAT AG AA AATCA TGGGA AAC CC TC CC AA	ATGTA GG CTGAATAA GA GTGGT CG CAGCT TC CG GGTA CCACA							

[illegible]

VEE Nsp1-4	Codon Opt(3940)	ACACTGACCAATATTTACACGGGAAGCGCCTTCATGAGGCTGGGTGTGCTCCAGTTATCACGTGTGAGGGGAGATATTGCAACTGCAACTGAGGGGGTCA	(3940)	3940	3950	3960	3970	3980	3990	4000	4010	4020	4030	4040
VEE Nsp1-4	Consensus(3940)	ACCTTGACCAACATTTATACAGGTTCCAGACTCCAGAAAGCCGATGTGACCTCATATCATGTGTGCGAGGGGATATTGCCAGGCCACCCGAAAGGATGA	(3940)	3940	3950	3960	3970	3980	3990	4000	4010	4020	4030	4040
VEE Nsp1-4	Codon Opt(4043)	TTATPAAACGCCGCAACTCAAGGGCCAAACCGGGCGGTGAGTGTGCGGTGCACTCTACAAAAAGTTTCCAGAGAGTTTCAGACCTTCAGCCCTATTGAGGTAGG	(4043)	4043	4050	4060	4070	4080	4090	4100	4110	4120	4130	4145
VEE Nsp1-4	Consensus(4043)	TTATPAAATGCTGCTAAACAGCAAGGACAACTCTGGCGAGGGGTGTGCGAGCGCTGTATAGAAATTTCCGGAAAGCTTCGATTTCAGTTCAGCCGATGCGAAAGTAGG	(4043)	4043	4050	4060	4070	4080	4090	4100	4110	4120	4130	4145
VEE Nsp1-4	Codon Opt(4146)	CAAAGCCCGCTGGTGAAGGCGCTGCAAGGCACATAATCCATGCGAGTGGGACCGAACTTCAACAAGTTAGCGAGGTGGAGGGTGATATAACAGTCGCGCGAG	(4146)	4146	4160	4170	4180	4190	4200	4210	4220	4230	4248	
VEE Nsp1-4	Consensus(4146)	AAAAGCGCGACTGGTCAAAGGTGCACTAAACATATCATTCATGCGTAGGACCAAACTTCAACAAAGTTTCGGAGGTGAAGGTGACACAAACAGTTGGCAGAG	(4146)	4146	4160	4170	4180	4190	4200	4210	4220	4230	4248	
VEE Nsp1-4	Codon Opt(4249)	GCGTATGAATCATTGCCAAGATAGTTAATGACAATAACTAATCCGTAGTATACCTTTGCTCTCTACGGGTATATTTCAGCGGTAAATAAAGATCGCCTGA	(4249)	4249	4260	4270	4280	4290	4300	4310	4320	4330	4340	4351
VEE Nsp1-4	Consensus(4249)	GCTTATGATCCATCGCTAAGATTGTCAACGATAACAATTACAAGTCAGTAGGATTCCTACTGTGTCCACCGCATCTTTCCGGGAACAAGATCGACTAA	(4249)	4249	4260	4270	4280	4290	4300	4310	4320	4330	4340	4351
VEE Nsp1-4	Codon Opt(4352)	CCCAAAGCCTGAAACCATCTGCTTACCGCTCTGGACAAACCGATGTCAGATGTGGCCATTTATTTCGCCGACAAAAAAGTGGGAGATGACACATGAAGGAGGCCGT	(4352)	4352	4360	4370	4380	4390	4400	4410	4420	4430	4440	4454
VEE Nsp1-4	Consensus(4352)	CCCAATCATTTGAACCATTTGCTGACAGCTTTAGACACCACTGATGACAGATGTAGCCATATATCTGAGGACAAGAATGGGAAATGACTCTCAAGGAAGCAGT	(4352)	4352	4360	4370	4380	4390	4400	4410	4420	4430	4440	4454
VEE Nsp1-4	Codon Opt(4455)	TGCGACAGCGGAGGCCGTAGAGGAGATCTGTATCAGTGATGACAGTTCTGTGACCGAGCCGACGCTGAACCTAGTTCAGGTTTCACCTAAATCTAGTCTGGCC	(4455)	4455	4460	4470	4480	4490	4500	4510	4520	4530	4540	4557
VEE Nsp1-4	Consensus(4455)	GGCTAGGAGAGAAGCAGTGGAGGAGATGATGCATATCCGACGACTCTTTCAGTGACAGAACTGATGACAGCTGTGAGGGTGCATCCGAAAGAGTTCTTTGGCT	(4455)	4455	4460	4470	4480	4490	4500	4510	4520	4530	4540	4557
VEE Nsp1-4	Codon Opt(4558)	GGAAGAAAGGGCTACTCTACCGAGACGGAAAGACCTTTTCTTACTCTGGAGGAAACAAAGTTCCACCAGCGGCGAAGGACATCGCCGAGATCAACGCAATGT	(4558)	4558	4570	4580	4590	4600	4610	4620	4630	4640	4650	4660
VEE Nsp1-4	Consensus(4558)	GGAAGGAAGGGCTACAGCAACAGCGATGGCAAACTTCTCATATTGGAAGGGACCAAGTTTCACCAAGCGGCCAAGGATATAGCAGAAATTAATGCCCATGT	(4558)	4558	4570	4580	4590	4600	4610	4620	4630	4640	4650	4660
VEE Nsp1-4	Codon Opt(4558)	ACAGCGA GG AA AC TT TC TA TGGA GG AC AAGTT CACCGCGCGC AAGGAAT GC GA AT AA GC ATGT	(4558)	4558	4570	4580	4590	4600	4610	4620	4630	4640	4650	4660
VEE Nsp1-4	Consensus(4558)	ACAGCGA GG AA AC TT TC TA TGGA GG AC AAGTT CACCGCGCGC AAGGAAT GC GA AT AA GC ATGT	(4558)	4558	4570	4580	4590	4600	4610	4620	4630	4640	4650	4660

[illegible]

FIG. 51

VEE Nsp1-4 Codon Opt(5279)	5279	5290	5300	5310	5320	5330	5340	5350	5360	5370	5381
VEE Nsp1-4(5279)	AGGGTGCCTCCGTAACCTCTGGCGCCACACAGTGCAGACCAACAGCTATTTCCGCAATCAATGGAATTTTGGCAAGGCCAGTGCCTCCCGGACCGT										
Consensus(5279)	AGGG GC CGT ACC	GG GC AC	GCCGAGAC AAC	TA TTGC AA	ATGGA TTCTGGC	G CC GTGCTGC CC CG AC GT					
(5382) 5382	5390	5400	5410	5420	5430	5440	5450	5460	5470	5484	
VEE Nsp1-4 Codon Opt(5382)	CTTCAGAAACCTCCGCATCCGCATCCGCACCTCGACCCGACACCAAGCTTGGCACCAATCCCGGCTGTCTCCGGAATAACTGGCGAGACAGTCCGTTACGCC										
VEE Nsp1-4(5382)	ATTCAGGAACCTCCACATCCGCTCCGCGACAAAGAACCGCTCACTTGCAACCGAGGCTGCTCGAGAGGATCAGCGGAGAACCGTGGGATACGG										
Consensus(5382)	TTTCAG AACCTCC	CATCCGC CC CG AC	G ACACC	T GCACC	C GGGCTG TC	G GG AT AC GG GA AC GT GG TACGC					
(5485) 5485	5490	5500	5510	5520	5530	5540	5550	5560	5570	5587	
VEE Nsp1-4 Codon Opt(5485)	GTAACACAAATCCGAAGGTTTTCGTTTGAAGTGAACCGACACTGTGAAGGGCGAGAGTGTCAATTTCCCGTGTGTACTTATATCCAGCCACCATT										
VEE Nsp1-4(5485)	GTACACACAAATAGCGAGGCTTCTGCTATGCAAGTTACTGACACAGTAAAGGAGAACGGGTATCGTTCCTGTGTGACGTACATCCCGGCCACCATAA										
Consensus(5485)	GT AC CACAAT	CGA GG TT TTGCT TGCAA	GT AC GACAC	GT AA GG GA	G GT TC TT CC GTGTG	AC TA ATCCC GCCACAT A					
(5588) 5588	5600	5610	5620	5630	5640	5650	5660	5670	5680	5690	
VEE Nsp1-4 Codon Opt(5588)	ACTCCAGAACCGCCTGGTCTCCAAACCGCCAGGCGTAAATAGGGTGATTAACAAGAGAGGAGTTTGAGGCGTTTGTAGCACAAACAATGACGGTTTGATGC										
VEE Nsp1-4(5588)	ACTCGAGAACCGCCTGGTCTCCAAACCGCCAGGCGTAAATAGGGTGATTAACAAGAGAGGAGTTTGAGGCGTTTGTAGCACAAACAATGACGGTTTGATGC										
Consensus(5588)	ACTC AGAACCGCCTGGTCTCCAAACCGCCAGGCGTAAATAGGGTGATTAACAAGAGAGGAGTTTGAGGCGTTTGTAGCACAAACAATGACGGTTTGATGC										
(5691) 5691	5700	5710	5720	5730	5740	5750	5760	5770	5780	5793	
VEE Nsp1-4 Codon Opt(5691)	GGGTGCATACATCTTTTCTCCGACACCGGTCAAGGGCATTTACAAACAAAATCAGTAAGGCAACCGTCTCTGTCGAGGTGTACTTGGAGAGGACAGAACTC										
VEE Nsp1-4(5691)	GGGTGCATACATCTTTTCTCCGACACCGGTCAAGGGCATTTACAAACAAAATCAGTAAGGCAACCGTCTCTGTCGAGGTGTACTTGGAGAGGACAGAACTC										
Consensus(5691)	GGGTGCATACATCTTTTCTCCGACACCGGTCAAGGGCATTTACAAACAAAATCAGTAAGGCAACCGTCTCTGTCGAGGTGTACTTGGAGAGGACAGAACTC										
(5794) 5794	5800	5810	5820	5830	5840	5850	5860	5870	5880	5896	
VEE Nsp1-4 Codon Opt(5794)	GAAATCTCATACGACCCAGGCTGGACCGAGAGAGGAAGAACTCTTGGGAAAAAGCTCCAGCTCAACCCCACTCTGCCAATAGGAGTCGCTATCAAGTC										
VEE Nsp1-4(5794)	GAGATTTGATGCCCCGGCTCGACCAAGAAAAAGAGAAATTAAGTAAATCCACACCTGCTAACAGAGCAGATACCAAGTCCA										
Consensus(5794)	GA AT TC TA GC CC G CT GACCA	GA AA GAAGAA T T CG AA AA	T CAG T AA CC AC CTTGC	AA AG AG G TA CA							
(5897) 5897	5910	5920	5930	5940	5950	5960	5970	5980	5990	5999	
VEE Nsp1-4 Codon Opt(5897)	GAAAAGTTGAAAATATGAAGGCTATTACAGCTCGACGAAATTTGCAAGGCTCGGGCACTACCTCAAGGCCGAGGCAAGTTGAAATGTTATAGAACACTTCA										
VEE Nsp1-4(5897)	GGAAGTTGGAGAACATGAAGAGCCATAACAGCTAGACGTATCTGCAAGGCTAGGGCATTTATTGAAGGCGAGAGGAAAAAGTGGAGTGTCTACCGAACCCCTGCA										
Consensus(5897)	G AA GT GA AA ATGAA GC AT ACAGCT	GACG ATT TGCAAGGCTT	GGGCA TA T AAGGC	GA GG AA GT GA TG TA GAAC CT CA							

(6000)	6000	6010	6020	6030	6040	6050	6060	6070	6080	6090	6100
VEE Nsp1-4	Codon Opt(6000)	CCAGTCCCACCTATAGCAGCTC-TGTAAACC	GCGGCTTTTCAGGCCCTAAAGTGCGAGGCTGC	AATGCGATGCTGAAGAGAATTTCCTACTGT							
VEE Nsp1-4(6000)	TCCGTGTTCCCTTTGTAAT-CATCTAGTGAACCG	CGCTTTTCAAGCCCCAAGGTGCGAGTGGAAAGCC	TGTAAAGCCCATGTTGAAGAGAACITTCGACGT								
Consensus(6000)	CC GT CC T TAT CA CT TGT AACCG GC	TTTTT AGCCC AA GT GCACT GA GC TG AA	GC ATG TGAA GAGAA TT CC ACTGT								
(6103)	6103	6110	6120	6130	6140	6150	6160	6170	6180	6190	6205
VEE Nsp1-4	Codon Opt(6102)	GGCGTCATACTGCATTATACCGGAGTATGCTT	TGACATGGTGACGGCGCAAGTTGCTGCCTCGAC	ACTGCTAGTTCCTGCTCCCGCAAGCTCCCG							
VEE Nsp1-4(6102)	GGCTTCTTACTGTPATTATTCAGAGTAGCATGCC	TATTGGACATGGTTGACATGCTTCAAGAGCTTCA	TGCTGCTTAGACACTGCCAGTTCCTGCCCCGCA	AGCTTCG							
Consensus(6103)	GGC TC TACTG ATTAT CC GAGTA GATGC TA	TTGGACATGGT GACGG GC TGCTGC T GAC	ACTGC AGTTT TG CC GC AAGCT CGC								
(6206)	6206	6220	6230	6240	6250	6260	6270	6280	6290	6300	6308
VEE Nsp1-4	Codon Opt(6205)	TCATTCCCTAAAAGCATAGTTACTTACGTAGCC	CAACTATTAGGCTGCGGTACCTTCTGCTATTCA	GAAACACTGCAGAAATGTTGGCGGCGCTACCA	AGA						
VEE Nsp1-4(6205)	AGCTTTCCAAGAAACACTCCTATTGGAAACCCA	ACAATACGATCGGCAGTGCCTTCAGCGATCCAG	AACAGCTCCAGAAAGTCTGGCAGCTGCCACAAA	AAAA							
Consensus(6206)	TT CC AA AA CA TA TT GA CC AC AT G TC	GC GT CCTTC GC AT CAGAACAC CT CAGAA	GT TGGC GC GC AC AA A								
(6309)	6309	6320	6330	6340	6350	6360	6370	6380	6390	6400	6411
VEE Nsp1-4	Codon Opt(6308)	GAAACTGCAATGTGACACAGATGAGGGAAC	TCTGTTCTCGACTCGGACCTTCAACGTTGAGT	GTTCGAAGAGTACGCCTGCAATAACGAATACT	GGGA						
VEE Nsp1-4(6308)	GAAATTGCAATGTCAACGAAATGAGAGAA	TGGCGGCTTTAATGTTGAATGCTTCAAGAAAT	TGCGTGTATAATGAATATTTGGGA								
Consensus(6309)	GAAA TGCAATGT AC CA ATGAG GAA T CC	GT T GA TC GC GCCIT AA GT GA TG TTC	AAGAA TA GC TG AATAA GAATA TGGGA								
(6412)	6412	6420	6430	6440	6450	6460	6470	6480	6490	6500	6514
VEE Nsp1-4	Codon Opt(6411)	GACTTCAAGGAGAAATCCATCCGCTGACAGA	GAGAAACGTTTGAATTACATCACATAACTGA	AGGGCCCCCAAAGCGCCGCAATGTTTCGT	AAAACTCAC						
VEE Nsp1-4(6411)	AACGTTTAAAGAAACCCCATCAGGCTTACTGA	AGAAACGTTGTTAAATTACATTACCAATTA	AAAGGACCAAAAGCTGCTCTTTTTCGGA	AGACACAT							
Consensus(6412)	AC TT AA GA AA CCCATC G CT AC GAAGA	AACGT GT AATTACAT AC AAA T AA GG CC	AAAGC GC GC T TT GC AA AC CA								
(6515)	6515	6520	6530	6540	6550	6560	6570	6580	6590	6600	6617
VEE Nsp1-4	Codon Opt(6514)	AACCTTAACATGCTGCAAGATATTC	CCATGGATCCCGTTGCTGATGGATCTCA	AAACGAGACGTCGAAGGTTACACCTTGG	CAACCAACATACGGAGGAACGCCCGCA						
VEE Nsp1-4(6514)	AATTTGAATATGTTCAGACATACCAATGG	ACAGGTTTGTATGGAACTTAAAGAGAGAC	GCTGAAAGTGACTCCAGAACAAACATP	ACTGAAGAACGGCCCCA							
Consensus(6515)	AA T AA ATG TGCA GA AT CC ATGGA	GGTT GT ATGGA T AA GAGACGT AA	GT AC CC GG AC AAACATAC GA	GAACG CC A							
(6618)	6618	6630	6640	6650	6660	6670	6680	6690	6700	6710	6720
VEE Nsp1-4	Codon Opt(6617)	AGGTACAGGTCAATCAGGCAGCGACCCCT	TGCCCAGGCTTATCTCTGCGGATTCAC	AGAGAGCTGGTTAGACGACTCAACG	CAGTTCCTTCGCCGAATAT						
VEE Nsp1-4(6617)	AGGTACAGGTGATCCAGGCTGCCGATCCG	CTAGCAACAGCGTATCTGTGCGGAATCC	CCAGAGAGCTGGTTAGGAGATTAAT	TCGGTCTGCTTCGCAACAT							
Consensus(6618)	AGGTACAGGT AT CAGGC GCCGA CC CT	GC AC GC TATCT TGCGG AT CAC	GAGAGCTGGTTAG GA T AA GC	GT CT CT CCGAA AT							

VEE Nsp1-4	Codon Opt(6720)	ACACACTCTCTTCGACATGTCAGCCGAGGATTTCCGATGCCATCATCGCCGAGCACTTTCACCCAGGAGATTTGTCTCTGGAGACGATATAGCATCATTTGAT	6721	6730	6740	6750	6760	6770	6780	6790	6800	6810	6820
VEE Nsp1-4	Consensus(6721)	TCATACACTGTTTGATATGTCGGCTGAAGACTTTGACCGCTATTATAGCCGAGCACTTCCAGCTGGGATTTGTTCTGGAAACTACATCGCTGCTTTGAT	6824	6830	6840	6850	6860	6870	6880	6890	6900	6910	6920
VEE Nsp1-4	Codon Opt(6823)	AAGAAGTGAAGGAGATGCGATGGCCCTTACCGCCCTTATGATACTTGAAGACCTGGGTGTCGATGCCGAGCTTCTGACTCTCATCGAGGCTGCCCTTCGGAGAAA	6824	6830	6840	6850	6860	6870	6880	6890	6900	6910	6920
VEE Nsp1-4	Consensus(6824)	AA AGTGAGGACGA GC ATGGC CT ACCGC T ATGAT CTGGAAGAC T GGTGT GA GC GAGCT TGAC CT AT GAGGC GC TTCGG GAAA	6927	6930	6940	6950	6960	6970	6980	6990	7000	7010	7020
VEE Nsp1-4	Codon Opt(6926)	TCAGCTCCATCCACTGCCACGAAGCAAGTTTCAAGTTTGTGTGGATGATGAAGTCCGGAATGTTTCTGACGCTGTTGCTTAAATACAGTAATCAATATAGT	6926	6930	6940	6950	6960	6970	6980	6990	7000	7010	7020
VEE Nsp1-4	Consensus(6927)	TTTTCATCAATACATTTGCCACTAAACTAAATTTAAATTCGGAGCCATGATGAATCTCGAATGTTCTCACACTGTTGTGAACACAGTCATTAACCAITGT	7030	7040	7050	7060	7070	7080	7090	7100	7110	7120	7130
VEE Nsp1-4	Codon Opt(7029)	TATAGCTTCACGGTCTCGCGGAGAGACTCACTGGAAGTCCCTCGCGCTTTCATCGGGACGATAACATTGTTTAAGGTGTTAAGTCAGATAAACTTATG	7029	7030	7040	7050	7060	7070	7080	7090	7100	7110	7120
VEE Nsp1-4	Consensus(7030)	AATCGCAAGCAGAGTGTGAGAGAACGGCTAACCGCATCCACTGTGCAGCAITTCATTGGAGATGACATATCGTGAAGGAGTCAAATCGGACAAATTAATG	7133	7140	7150	7160	7170	7180	7190	7200	7210	7220	7230
VEE Nsp1-4	Codon Opt(7132)	GGGACCGCTGTGCTACATGCTGTAACATGGAGGTGAAAATAATTGACGCACTGCTCGCGGAGAGGCAACCGTACTCTTGTGTGGATTTATCCTCTGCGATT	7132	7140	7150	7160	7170	7180	7190	7200	7210	7220	7230
VEE Nsp1-4	Consensus(7133)	GCAGACAGGTGCGCCACCTGTTGTAATGGAATCAAGATTATAGATGCTGTGTGGCGAGAAAGCGCTTATTTCTGTGGAGGTTTATTTGTGTGACT	7133	7140	7150	7160	7170	7180	7190	7200	7210	7220	7230
VEE Nsp1-4	Codon Opt(7235)	CCGTACAGGCACGGCATGCGGGTTCGCGATCCCTCAAGAGGTGTTCAAGTGGGCAAGCCTCTCGCTGCAGATGAACACGACGACGCGCGCGG	7235	7240	7250	7260	7270	7280	7290	7300	7310	7320	7330
VEE Nsp1-4	Consensus(7236)	CCGTACAGGCACGGCATGCGGGTTCGCGATCCCTCAAGAGGTGTTCAAGTGGGCAAGCCTCTCGCTGCAGATGAACACGACGACGCGCGCGG	7236	7240	7250	7260	7270	7280	7290	7300	7310	7320	7330
VEE Nsp1-4	Codon Opt(7338)	CGCACTGCACGAGGAATCAACTAGGTGGAACAGAGTGGGAATCCTGTCTGAATGTCAGGCTGTGCGAATCCAGATACGAAACTGTGGGACATCCATCATC	7338	7340	7350	7360	7370	7380	7390	7400	7410	7420	7430
VEE Nsp1-4	Consensus(7339)	GGCATTGCAAGAGTCAACACGCTGGAACCGAGTGGGTATTTTCAGAGCTGTGCAAGCAGTAGAATCAAGGTATGAACCGTAGGAATTCATCATCAT	7339	7340	7350	7360	7370	7380	7390	7400	7410	7420	7430
VEE Nsp1-4	Codon Opt(7339)	GCA TGCA GA GA TCAAC G TGAAC GAGTGGG AT CT TC GA CTGTGCAAGC GT GAATC AG TA GAAAC GT GG AC TCCATCAT	7339	7340	7350	7360	7370	7380	7390	7400	7410	7420	7430

FIG. 5L

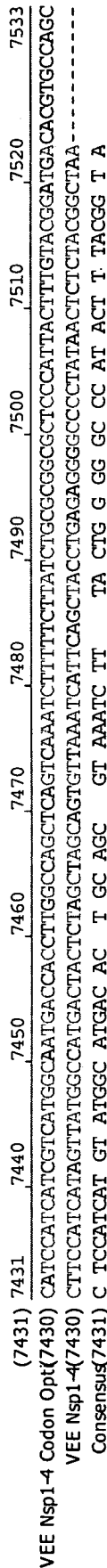


FIG. 6

