

VACCINE COMPOSITIONS AND METHODS OF USE THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. provisional patent application No. 63/162,946, filed March 17, 2021, the contents of which are incorporated herein by reference in their
5 entirety.

INCORPORATION OF THE SEQUENCE LISTING

[0002] The contents of the text file submitted electronically herewith are incorporated herein by reference in their entirety: A computer readable format copy of the Sequence Listing (filename: EXCI_003_03WO_SeqList_ST25, date recorded March 17, 2022, file size 266
10 kilobytes).

BACKGROUND

[0003] Viral infectious diseases, *e.g.*, the flu, are extremely widespread, and often contagious. Viral diseases result in a wide variety of symptoms that vary in character and severity depending on the type of viral infection and other factors, including the person's age and overall
15 health. Viral infections can be treated with varying degrees of success, depending on the type of virus and other factors. Sometimes, the treatment may involve just management of symptoms.

[0004] The most effective way to combat viral infections is through vaccination, which can induce a holistic cellular and/or humoral immune response that is protective against future
20 infections. Vaccinations offer enormous public health and economic benefits by preventing the occurrence of, or minimizing, the severity of viral infections. Although vaccines are available to prevent more than 20 life-threatening diseases, including viral infections, the appearance of new infectious viruses, *e.g.*, SARS CoV2, can necessitate rapid research and development of vaccines against new viral targets. However, the development of vaccines, even with the latest
25 genome sequencing and other technological advancements, is still time-consuming and expensive.

[0005] Therefore, there is a need for methods and compositions that have the versatility to be quickly adapted for the development of vaccines targeting different viruses.

BRIEF DESCRIPTION OF FIGURES

[0006] FIGS. 1A and 1B show the plasmid maps of CoVEG2 (FIG. 1A) and CoVEG1 (FIG. 1B). Also shown are the relative positions of cytomegalovirus (CMV) enhancer and promoter and Simian virus 40 Poly A (SV40PA).

5 [0007] FIG. 2 shows the expression of SARS-CoV-2 S protein, SARS-CoV-2 M protein, SARS-CoV-2 E protein and SARS-CoV-2 N protein from CoVEG2 in HEK293 cells.

[0008] FIG. 3 shows that SARS-CoV-2 S, N, M and E proteins expressed from CoVEG2 are able to assemble into VLPs and are secreted from cells. See Example 3. FIG. 3A shows results from an SDS PAGE experiment showing S, N, M, and E protein bands in the size exclusion chromatography void sample. FIG. 3B shows the chromatogram of the size exclusion VLP isolation run, whereas the void volume peak, which contains the VLPs, is indicated with an arrow. FIG. 3C shows results from a Western Blotting experiment showing S, N, M, and E protein bands in the size exclusion chromatography void sample. E: envelope protein; M: membrane protein; N: nucleocapsid protein; S: spike protein; NT: non-transfected.

15 [0009] FIG. 4 shows the study design to test the immunogenicity of CoVEG1 and CoVEG2 plasmids in mice.

[0010] FIG. 5A-5O show the plasmid maps of each of CoVEG 3-17, respectively. FIG. 5P shows the plasmid map of the control S only plasmid. CMV: cytomegalovirus; SV40PA: Simian virus 40 Poly A.

20 [0011] FIG. 6 shows a schematic depiction of the expression cassettes in each of CoVEG 3-17. Each of the plasmids CoVEG 3-17 vary in the genes that are encoded, the order of the genetic elements and the presence or absence of regulatory elements, *e.g.*, the viral packaging signal. The boxes marked “M”, “S”, “N” and “E” indicate the genes encoding the membrane protein, spike protein, nucleocapsid protein and the envelope protein, respectively. The box marked “L” refers to the gene encoding the L enhancer protein. “S (Mut)” denotes the prefusion conformation-stabilized spike protein mutant in which an internal endogenous furin cleavage site has been mutated to comprise the following amino acid substitutions: R682G, R683S, R685S; and which further has the two amino acid substitutions, K986P and V987P.

25 [0012] FIG. 7 shows the images from anti-spike (S) protein antibody immunostaining of HEK293T cells transfected with each of the indicated plasmids or a control plasmid that expresses only the S protein without the other viral proteins (M, N or E). The images confirm the expression of the S protein in these cells.

[0013] FIG. 8 shows the results from the Western Blot analysis of a preparation of viral-like particles (VLPs) isolated from cell culture supernatants when transiently transfected with the indicated CoVEG constructs. “S” denotes transient transfection with a control plasmid that expresses only the S protein without the other viral proteins (M, N or E). Staining with anti-RBD antibody (left) and anti-N protein antibody (right) show that the tested plasmids are capable of expressing and secreting viral structural proteins as VLPs into the cell culture supernatants. The red arrow indicates the full length protein.

[0014] FIG. 9A shows the total signal values obtained from an ELISA analysis using 1:500 diluted serum samples from BALB/c mice injected with each of the indicated CoVEG plasmids after 56 days. FIG. 9B shows the endpoint titer values obtained from an ELISA analysis using serum samples from BALB/c mice injected with each of the indicated CoVEG plasmids after 56 days. Endpoint titer refers to the reciprocal maximal antibody dilution at which the ELISA signal (absorbance at 450 nm) is above 3 standard deviations of background signal.

[0015] FIG. 10 shows the percent (%) inhibition of the in vitro binding of the Spike protein receptor binding domain (RBD) to the ACE2 receptor by serum samples obtained from mice injected with each of the indicated plasmids using the commercial cPassTM neutralization assay (GenScript). The results show that the serum samples obtained from CoVEG5 and CoVEG8-injected mice have neutralizing antibodies. The positive and negative controls are the assay controls of the cPass neutralization assay kit and contain a known amount of SARS-CoV-2 neutralizing antibodies or a seronegative sample, respectively, and are used according to the manufacturer’s protocol.

[0016] FIG. 11 shows the results from the Western Blot analysis of a preparation of viral-like particles (VLPs) isolated from cell culture supernatants when transiently transfected with the indicated CoVEG constructs. Staining with anti-RBD antibody (left) and anti-N protein antibody (right) show that the tested plasmids are capable of expressing and secreting viral structural proteins as VLPs into the cell culture supernatants with varying efficiencies.

[0017] FIG. 12 shows the results from the Western Blot analysis of a preparation of viral-like particles (VLPs) isolated from cell culture supernatants when transiently transfected with the indicated CoVEG constructs. Staining with anti-RBD antibody (left) and anti-N protein antibody (right) show that the tested plasmids are capable of expressing and secreting viral structural proteins as VLPs into the cell culture supernatants with varying efficiencies.

[0018] FIG. 13 shows the total signal values obtained from an ELISA analysis from BALB/c mice injected either intradermally (ID) or intramuscularly (IM) with each of the indicated CoVEG plasmids after 15 days. “S-only” denotes administration of the Spike protein alone.

[0019] FIG. 14A shows the map of a plasmid encoding West Nile virus proteins, preM protein and envelope protein, along with the enhancer protein (EMCV L1 protein). FIG. 14B shows the map of a control plasmid encoding just the West Nile virus proteins, preM protein and envelope protein CMV: cytomegalovirus; SV40PA: Simian virus 40 Poly A.

5 [0020] FIGS. 15A and 15B show the total signal values obtained from an ELISA analysis of VLP secretion in HEK293 cells. FIG. 15A shows ELISA analysis of VLP secretion performed with anti-spike (S protein) antibodies. FIG. 15B shows ELISA analysis of VLP secretion performed with anti-nucleocapsid (N protein) antibodies.

[0021] FIG. 16 shows transmission electron micrographs (TEM) of CoVEG10 and CoVEG20 protein expression. FIG. 16, left shows TEM of CoVEG10 which contains the L regulatory protein. FIG. 16, right shows TEM of CoVEG20 which lacks the L regulatory protein.

[0022] FIG. 17 shows images from anti-nucleocapsid (N) protein antibody immunostaining of HEK293T cells transfected with each of the indicated plasmids.

15 [0023] FIG. 18 shows the results from Western Blot analysis of a preparation of viral-like particles (VLPs) isolated from cell culture supernatants when transiently transfected with the indicated CoVEG constructs. Staining with anti-RBD antibody (left) and anti-N protein antibody (right) show that the tested plasmids are capable of expressing and secreting viral structural proteins as VLPs into the cell culture supernatants with varying efficiencies.

20 [0024] FIG. 19 shows the western blot results of co-immunoprecipitation experiments of the CoVEG 10 plasmid. (Left side) The receptor binding domain (RBD) pull-down signal of the N protein in the elution indicates the N protein was retained within the particles. (Right side) A co-IP was performed without the anti-RBD antibody demonstrating that the N protein did not bind the precipitation resin non-specifically.

25 [0025] FIG. 20 shows antibody binding titers from CoVEG 5 and 9-14 plasmids, as well as a spike (S) protein only containing plasmid, after intramuscular (IM) or intradermal (ID) injection in C57BL/6 mice.

[0026] FIG. 21 shows the ELISA analysis of the neutralizing antibodies from the samples shown in FIG. 20.

30 [0027] FIG. 22 shows the antibody binding titers from CoVEG 9, 10, and 18-20 plasmids, as well as a spike (S) protein only containing plasmid, after intramuscular (IM) or intradermal (ID) injection in C57BL/6 mice.

[0028] FIG. 23 shows the ELISA analysis of neutralizing antibodies produced in response to CoVEG 9, 10 and 20 plasmids intramuscular (IM) or intradermal (ID) injection in C57BL/6 mice.

[0029] FIGS. 24A and 24B show the results of T-cell analysis of CoVEG10 and CoVEG20. FIG. 24 A shows the results of T-cell analysis of CoVEG10. FIG. 24B shows the results of T-cell analysis of CoVEG20.

[0030] FIG. 25 shows ELISA analysis of VLPs that were purified from cell culture supernatants of HEK293T cells transfected with isolated and resuspended West-Nile Virus (WNV) constructs with the enhancer protein (WNV + EG, circles) and without the enhancer protein (WNC, squares). The specificity of the ELISA was tested against a plasmid containing GFP that does not give a signal in the specific ELISA assay. The ELISA analysis revealed that the isolation of VLPs by ultracentrifugation with the addition of the enhancer protein contained more active VLPs than in the absence of the enhancer protein. This was especially surprising because the total amount of produced protein was higher in the absence of the enhancer protein, further demonstrating that the addition of an enhancer proteins increased the quality of the expressed target protein.

[0031] FIGS. 26A and 26B show the time course analysis of cell culture supernatants obtained from HEK293T cells overexpressing the supernatant containing over expressed West-Nile Virus (WNV) constructs with the enhancer protein (WNV + L) and without the enhancer protein (WNV). FIG. 26A shows ELISA analysis demonstrating that the VLP concentration in the WNV + L (left) construct peaked at 72 h after transfection and gradually decreased thereafter. This was evidence of VLP secretion from healthy cells, as the expression profile followed expected production of VLP from transient transfections and on the related VLP particle half-life. In contrast the WNV construct in the absence of the enhancer protein showed a constant increase of VLP (right) in the supernatant that is more consistent with unspecific release of protein from cells during cell death rather than active secretion. FIG. 26 B confirmed ELISA analysis with cell images. WNV +L cell images (left) revealed very little to no cell death whereas the WNV cell images revealed visible signs of cell death starting at 72 hours after transfection (right, cell death indicated by black stars).

[0032] FIGS. 27A and 27B show the analysis of neutralizing antibodies of CoVEG9, the Spike construct with the enhancer protein (Spike + L) and the Spike construct without the enhancer protein (Spike) on days 42 and 70 after immunization of mice. The presence of neutralizing antibodies was detected using the commercial cPass™ SARS-CoV-2 Surrogate Virus Neutralization Test (sVNT) Kit (GenScript). The analysis showed that although all constructs lose some neutralization ability over time, the addition of the enhancer protein prolonged the presence of the desired neutralizing antibodies in the serum of tested animals (CoVEG9 and

Spike + L). FIG. 27A shows the individual animals by group on days 42 and 70. FIG. 27B shows the median neutralizing antibody levels in each treatment group of the same data.

5 [0033] FIG. 28 shows immunofluorescence images from cells overexpressing either West-Nile Virus (WNV) constructs with the enhancer protein (WNV + L, left) or without the enhancer protein (WNV, right). As observed in other examples, the total amount of protein decreased when an enhancer protein was added, as indicated by the weaker Alexa Fluor 488 Fluor signal of the secondary antibody used in the immunostaining (WNV + L, left) compared to the WNV (right). However, the absence of the enhancer protein led to the formation of nuclei foci, consistent with the aggregation or misfolding of the expressed protein (right, arrows) indicating
10 a lower quality of the expressed protein compared to the construct with the enhancer protein (WNV + L).

15

SUMMARY

[0034] The disclosure provides compositions for use as a vaccine, comprising an expression cassette comprising a polynucleotide encoding a viral protein and a polynucleotide encoding an enhancer protein. In some embodiments, the enhancer protein is a picornavirus leader (L) protein or a functional variant thereof. In some embodiments, the amino acid sequence of the
20 enhancer protein has at least 95% identity to SEQ ID NO: 1, or at least 95% identity to SEQ ID NO: 2. In some embodiments, the amino acid sequence of the enhancer protein is SEQ ID NO: 1, or SEQ ID NO: 2. In some embodiments, the polynucleotide encoding the enhancer protein is operatively linked to a polynucleotide encoding an internal ribosome entry site (IRES). In some embodiments, the polynucleotide encoding the IRES is SEQ ID NO: 24.

25 [0035] In some embodiments, the viral protein is a viral antigen. In some embodiments, the viral protein is derived from a virus selected from the group consisting of coronavirus, influenza virus, Hepatitis B virus, Human Papilloma virus (HPV), West Nile virus, and Human Immunodeficiency Virus (HIV) virus. In some embodiments, the viral protein is derived from a coronavirus. In some embodiments, the coronavirus is a betacoronavirus. In some
30 embodiments, the betacoronavirus is severe acute respiratory syndrome (SARS) virus. In some embodiments, the SARS virus is a SARS-CoV-2 virus. In some embodiments, the betacoronavirus is Middle East respiratory syndrome (MERS) virus.

[0036] In some embodiments, the coronavirus protein is a coronavirus spike protein. In some embodiments, the spike protein shares at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 13. In some embodiments, the spike protein is SEQ ID NO: 13. In some embodiments, the coronavirus protein is a coronavirus membrane (M) protein. In some embodiments, the M protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 33. In some embodiments, the M protein is SEQ ID NO: 33. In some embodiments, the coronavirus protein is a coronavirus envelope (E) protein. In some
10 embodiments, the E protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 22. In some embodiments, the E protein is SEQ ID NO: 22. In some embodiments, the coronavirus protein is a coronavirus nucleocapsid (N) protein. In some embodiments, the N protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%,
15 at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 20. In some embodiments, the N protein is SEQ ID NO: 20. In some embodiments, the coronavirus protein forms a virus-like particle (VLP).

[0037] In some embodiments, the viral protein is derived from West Nile virus. In some embodiments, the viral protein is precursor membrane protein (preM), envelope glycoprotein (E), or a combination thereof.
20

[0038] The disclosure provides vectors for use as a vaccine, comprising an expression cassette comprising a polynucleotide, wherein the polynucleotide comprises a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to the
25 nucleic acid sequence of SEQ ID NO: 30. In some embodiments, the polynucleotide comprises the nucleic acid sequence of SEQ ID NO: 30.

[0039] The disclosure provides vectors for use as a vaccine, comprising an expression cassette comprising a polynucleotide, wherein the polynucleotide comprises a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%,
30 at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to the nucleic acid sequence of SEQ ID NO: 31. In some embodiments, the polynucleotide comprises the nucleic acid sequence of SEQ ID NO: 31.

[0040] The disclosure provides vectors for use as a vaccine, comprising a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at

least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO: 35-49, 55 and 62.

[0041] In some embodiments, the vector is a naked polynucleotide. In some embodiments, the vector is a deoxyribonucleic acid (DNA) polynucleotide. In some embodiments, the vector is a ribonucleic acid (RNA) polynucleotide. In some embodiments, the vector comprises a plasmid. In some embodiments, the vector comprises linear DNA. Vectors include plasmids, viruses, bacteriophages, pro-viruses, phagemids, transposons, artificial chromosomes, and the like, that in some cases can replicate autonomously or can integrate into a chromosome of a host cell. A vector can also be a naked RNA polynucleotide, a naked DNA polynucleotide, a polynucleotide composed of both DNA and RNA within the same strand, a poly-lysine-conjugated DNA or RNA, a peptide-conjugated DNA or RNA, a liposome-conjugated DNA, or the like, that is not autonomously replicating.

[0042] In some embodiments, the vector is an adeno-associated virus (AAV) vector, a lentivirus vector, a retrovirus vector, a replication competent adenovirus vector, a replication deficient adenovirus vector, a herpes virus vector, a baculovirus vector, a nonviral plasmid, a viral vector, a plasmid, a phage, a phagemid, a cosmid, a fosmid, a bacteriophage, an artificial chromosome, or an adenovirus vector. In some embodiments, the vector is a bacterial artificial chromosome (BAC), a plasmid, a bacteriophage P1-derived vector (PAC), a yeast artificial chromosome (YAC), or a mammalian artificial chromosome (MAC).

[0043] In some embodiments, the expression cassette comprises a promoter operatively linked to each of the polynucleotide sequences of the expression cassette. In some embodiments, the vector comprises a DNA polynucleotide, said DNA polynucleotide encoding a viral packaging signal. In some embodiments, the viral packaging signal is a RNA polynucleotide. In some embodiments, the viral packaging signal is derived from a coronavirus.

[0044] The disclosure provides vaccine compositions, comprising any one of the vectors disclosed herein, and a pharmaceutically acceptable carrier. In some embodiments, the vaccine composition comprises an adjuvant. In some embodiments, the adjuvant is alum. In some embodiments, the adjuvant is monophosphoryl lipid A (MPL).

[0045] The disclosure provides methods of expressing a viral antigen in a eukaryotic cell, comprising contacting the cell with any one of the vectors disclosed herein. In some embodiments, contacting the cell with the vector results in: (i) expression of the antigen at a higher expression level; and/or (ii) expression of the antigen for a longer period of time; and/or (iii) expression of the antigen with better protein quality, than a vector lacking the enhancer

protein. In some embodiments, contacting the cell with the vector results in: (i) expression of a virus like particle (VLP) comprising the antigen at a higher expression level; and/or (ii) expression of a VLP comprising the antigen for a longer period of time; and/or (iii) expression of a VLP comprising the antigen with better protein quality, than a vector lacking the enhancer protein.

[0046] In some embodiments, the vector comprises a polynucleotide encoding a viral packaging signal, wherein contacting the cell with the vector results in expression of the viral packaging signal, and wherein the VLPs encapsidate the viral packaging signal. In some embodiments, the vector results in the formation of a greater number of VLPs, as compared to a control vector lacking the polynucleotide encoding the viral packaging signal.

[0047] The disclosure provides methods of eliciting an immune response in a subject, comprising administering an effective amount of any one of the vaccine compositions disclosed herein to the subject. In some embodiments, tissue at an administration site of the subject expresses the antigen and/or a VLP comprising the antigen. In some embodiments, tissue at an administration site of the subject: (i) expresses the antigen and/or a VLP comprising the antigen at a higher expression level; and/or (ii) expresses the antigen and/or a VLP comprising the antigen for a longer period of time; and/or (iii) expresses the antigen and/or a VLP comprising the antigen with better protein quality, than when a vector lacking the enhancer protein is administered. In some embodiments, the vector comprises a polynucleotide encoding a viral packaging signal, wherein tissue at an administration site of the subject expresses the viral packaging signal, and wherein the VLPs encapsidate the viral packaging signal. In some embodiments, the vector results in the expression of a greater number of VLPs, as compared to a control vector lacking the polynucleotide encoding the viral packaging signal. In some embodiments, the VLPs encapsidating the viral packaging signal are more immunogenic than control VLPs comprising the antigen but lacking the viral packaging signal.

[0048] In some embodiments, the method elicits an antibody response in the subject. In some embodiments, the antibody response is a neutralizing antibody response. In some embodiments, the method elicits a cellular immune response. In some embodiments, the method elicits a prophylactic, protective and/or therapeutic immune response in the subject. In some embodiments, the administration is intradermal administration, intramuscular administration, subcutaneous administration, or intranasal administration.

[0049] The disclosure provides polynucleotides comprising an expression cassette comprising a polynucleotide encoding a coronavirus protein and a polynucleotide encoding an enhancer protein, wherein the enhancer protein is a picornavirus leader (L) protein or a functional variant

thereof. In some embodiments, the amino acid sequence of the enhancer protein has at least 95% identity to SEQ ID NO: 1, or at least 95% identity to SEQ ID NO: 2. In some embodiments, the amino acid sequence of the enhancer protein is SEQ ID NO: 1, or SEQ ID NO: 2.

5 **[0050]** In some embodiments, the polynucleotide encoding the enhancer protein is operatively linked to a polynucleotide encoding an internal ribosome entry site (IRES). In some embodiments, the polynucleotide encoding the IRES is SEQ ID NO: 24. In some embodiments, the coronavirus protein forms a virus-like particle (VLP).

[0051] The disclosure provides polynucleotides comprising a nucleic acid sequence having at
10 least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to the nucleic acid sequence of SEQ ID NO: 30. In some embodiments, the polynucleotide comprises a nucleic acid sequence of SEQ ID NO: 30. The disclosure provides polynucleotides comprising a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%,
15 at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to the nucleic acid sequence of SEQ ID NO: 31. In some embodiments, the polynucleotide comprises a nucleic acid sequence of SEQ ID NO: 31.

[0052] In some embodiments, the polynucleotide is a naked polynucleotide. In some embodiments, the polynucleotide is a deoxyribonucleic acid (DNA) polynucleotide. In some
20 embodiments, the polynucleotide is a ribonucleic acid (RNA) polynucleotide. In some embodiments, the expression cassette comprises a promoter operatively linked to each of the polynucleotide sequences of the expression cassette.

[0053] The disclosure provides kits comprising a vector, wherein the vector comprises an expression cassette comprising a polynucleotide encoding a coronavirus protein and a
25 polynucleotide encoding an enhancer protein, wherein the enhancer protein is a picornavirus leader (L) protein or a functional variant thereof.

[0054] The disclosure provides vectors, comprising an expression cassette, said expression cassette comprising a promoter linked to a target gene, wherein the vector comprises a nucleic acid sequence encoding a viral packaging element. In some embodiments, the viral packaging
30 element is a RNA polynucleotide. In some embodiments, the viral packaging element is derived from a coronavirus. In some embodiments, the viral packaging element is derived from SARS-CoV2. In some embodiments, the nucleic acid sequence encoding the viral packaging element has at least about 70% identity to the nucleic acid sequence of SEQ ID NO: 34.

[0055] The disclosure provides methods of expressing a target protein in a eukaryotic cell, comprising contacting the cell with any one of the vectors disclosed herein. In some embodiments, contacting the cell with the vector results in the formation of virus-like particles (VLPs) comprising the target protein. In some embodiments, contacting the cell with the vector results in the formation of a greater number of virus-like particles (VLPs) comprising the target protein, as compared to a control vector comprising the expression cassette but lacking the nucleic acid sequence encoding the viral packaging element. In some embodiments, the nucleic acid sequence encoding the viral packaging element has at least about 70% identity to the nucleic acid sequence of SEQ ID NO: 34.

10 [0056] The disclosure provides vectors for use as vaccines, comprising an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein, wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a first proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a second proteolytic cleavage site, a polynucleotide encoding an S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a third proteolytic cleavage site, a polynucleotide encoding an E protein, wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an IRES sequence, wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, and a polynucleotide encoding an enhancer L protein, wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto.

25 [0057] The disclosure also provides vectors for use as a vaccine, comprising an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein or an amino acid sequence at least 95% identical thereto, wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a first proteolytic cleavage site, a polynucleotide encoding an S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a second proteolytic cleavage site, a polynucleotide encoding an E protein or a polynucleotide sequence at least 95% identical thereto, wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an IRES sequence, wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical

thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, and a polynucleotide encoding a viral packaging signal, wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 95% identical thereto.

[0058] The disclosure also provides vectors for use as vaccines, comprising an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein, wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a mutated S protein, wherein the mutated S protein comprise SEQ ID NO: 51 or SEQ ID NO: 52 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein, wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an IRES sequence, wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, and a polynucleotide encoding an enhancer L protein, wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto.

[0059] The disclosure provides vectors for use as vaccines, comprising an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a first polynucleotide encoding a viral packaging signal, wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an M protein, wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a mutated S protein, wherein the S protein comprises SEQ ID NO: 51 or SEQ ID NO: 52 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein, wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an IRES sequence, wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding

an enhancer L protein, wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, and a second polynucleotide encoding a viral packaging signal, wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 95% identical thereto.

5

DETAILED DESCRIPTION

[0060] Virus-like particles (VLPs) are composed of viral structural proteins. Although VLPs are immunogenic, they are non-infectious. Therefore, VLPs have enormous potential for use in the development of vaccines. VLPs may be produced *in vitro*, and then administered to a subject in need of immunization. Alternatively, VLPs may be produced *in vivo* in the subject.

[0061] The production of VLPs is challenging primarily because it often requires the expression of more than one structural protein from more than one plasmid. In some cases, several plasmids carrying different structural proteins may need to be introduced into the host cell at defined ratios to support the formation of VLPs. This process can be unreliable and often fails to produce sufficient levels of VLPs of required quality. If the multiple structural proteins that are required for the formation of the VLPs can be expressed from a single plasmid or a single RNA transcript, that will greatly simplify the process of making VLPs and thus, provide a much-needed boost for the development of vaccines comprising VLPs.

[0062] The compositions and methods disclosed herein enable the reliable formation of high levels of VLPs *in vivo* and thus, enable a robust induction of immune response against the viral antigens on the VLPs. Furthermore, these compositions and methods may be used to induce immune response against different viruses, *e.g.*, coronaviruses (*e.g.* SARS CoV-2), influenza viruses, and West Nile virus.

Definitions

[0063] As used herein, and in the appended claims, the singular forms “a”, “an”, and “the” include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to “an antigen” can refer to one or more antigens, and reference to “the method” includes reference to equivalent steps and/or methods known to those skilled in the art, and so forth.

[0064] As used herein, the term “about” or “approximately” when preceding a numerical value indicates the value plus or minus a range of 10%. For example, “about 100” encompasses 90 and 110.

[0065] As used herein, nucleotide sequences are listed in the 5' to 3' direction, and amino acid sequences are listed in the N-terminal to C-terminal direction, unless indicated otherwise.

[0066] The terms “polypeptide,” “peptide” and “protein” are used interchangeably herein to refer to polymers of amino acids of any length. The polymer may be linear or branched, it may
5 comprise modified amino acids, and it may be interrupted by non-amino acids. The terms also encompass an amino acid polymer that has been modified; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation, *e.g.*, conjugation with a labeling component. As used herein the term “amino acid” includes natural and/or unnatural or synthetic amino acids, including glycine and both the D or L optical
10 isomers, and amino acid analogs and peptidomimetics.

[0067] As used herein, the term “subject” includes humans and other animals. Typically, the subject is a human. For example, the subject may be an adult, a teenager, a child (2 years to 14 years of age), an infant (1 month to 24 months), or a neonate (up to 1 month). In some embodiments, the adults are seniors about 65 years or older, or about 60 years or older. In
15 some embodiments, the subject is a pregnant woman or a woman intending to become pregnant. In other embodiments, subject is not a human; for example a non-human primate; for example, a baboon, a chimpanzee, a gorilla, or a macaque. In certain embodiments, the subject may be a pet, *e.g.*, a dog or cat.

[0068] As used herein, the terms “immunogen,” “antigen,” and “epitope” refer to substances
20 *e.g.*, proteins, including glycoproteins, and peptides that are capable of eliciting an immune response.

[0069] As used herein, an “immunogenic response” in a subject results in the development in the subject of a humoral and/or a cellular immune response to an antigen.

[0070] The term “effective amount” or “therapeutically effective amount” refers to the amount
25 of an agent that is sufficient to achieve an outcome, for example, to affect beneficial or desired results. The therapeutically effective amount may vary depending upon one or more of: the subject and disease condition being treated, the weight and age of the subject, the severity of the disease condition, the manner of administration and the like, which can readily be determined by one of ordinary skill in the art. The specific dose may vary depending on one or
30 more of: the particular agent chosen, the dosing regimen to be followed, whether it is administered in combination with other compounds, timing of administration, the tissue into which it is administered, and the physical delivery system in which it is carried.

[0071] As used herein, the term “virus-like particle” (VLP) refers to a structure that in at least one attribute resembles a virus but which has not been demonstrated to be infectious. Virus-

like particles in accordance with the disclosure do not carry genetic information encoding for the proteins of the virus-like particles. In general, virus-like particles lack a viral genome and, therefore, are noninfectious. In addition, virus-like particles can often be produced in large quantities by heterologous expression and can be easily purified.

- 5 [0072] As used herein, an amino acid substitution, interchangeably referred to as amino acid replacement, at a specific position on the protein sequence is denoted herein in the following manner: “one letter code of the WT amino acid residue - amino acid position - one letter code of the amino acid residue that replaces this WT residue”. For example, a Spike (S) protein which is a R682G mutant refers to an S protein in which the wild type residue at the 682nd
10 amino acid position (R or arginine) is replaced with G or glycine.

Vectors

[0073] The disclosure provides vectors comprising an expression cassette comprising a polynucleotide encoding an antigen and a polynucleotide encoding an enhancer protein. In some embodiments, the vector is used as a vaccine, or as part of a vaccine composition.

- 15 [0074] The term “vector” refers to the means by which a nucleic acid can be propagated and/or transferred between organisms, cells, or cellular components. A vector for use according to the present disclosure may comprise any vector known in the art. Vectors include plasmids, viruses, bacteriophages, pro-viruses, phagemids, transposons, artificial chromosomes, and the like, that replicate autonomously or can integrate into a chromosome of a host cell. A vector
20 can also be a naked RNA polynucleotide, a naked DNA polynucleotide, a polynucleotide composed of both DNA and RNA within the same strand, a poly-lysine-conjugated DNA or RNA, a peptide-conjugated DNA or RNA, a liposome-conjugated DNA, or the like, that is not autonomously replicating.

- [0075] In some embodiments, the vector is an adeno-associated virus (AAV) vector, a
25 lentivirus vector, a retrovirus vector, a replication competent adenovirus vector, a replication deficient adenovirus vector, a herpes virus vector, a baculovirus vector, a nonviral plasmid, a viral vector, a plasmid, a phage, a phagemid, a cosmid, a fosmid, a bacteriophage, an artificial chromosome, or an adenovirus vector. In some embodiments, the vector is a bacterial artificial chromosome (BAC), a plasmid, a bacteriophage P1-derived vector (PAC), a yeast artificial
30 chromosome (YAC), or a mammalian artificial chromosome (MAC). In some embodiments, the vector is a naked polynucleotide. In some embodiments, the vector is a deoxyribonucleic

acid (DNA) polynucleotide. In some embodiments, the vector is a ribonucleic acid (RNA) polynucleotide.

[0076] In some embodiments, the vector comprises a first polynucleotide encoding an antigen and a second polynucleotide encoding an enhancer protein. In some embodiments, the vector has a design as shown in FIG. 1A or FIG. 1B. In some embodiments, the vector is CoVEG1. In some embodiments, the vector is CoVEG2.

[0077] Table 1 shows the nucleic acid sequences of important regions of the CoVEG1 and CoVEG2, and amino acid sequences encoded by these regions.

Table 1

Name	Type of sequence	Sequence	SEQ ID NO.:
SARS-CoV-2 Spike	Amino acid	MFVFLVLLPLVSSQCVNLTTRTQLPP AYTNSFTRGVYYPDKVFRSSVLHSTQ DLFLPFFSNVTWFHAIHVSGTNGTKR FDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLIVNNATNVVIKVCE FQFCNDPFLGVYYHKNNKSWMESEF RVYSSANNCTFEYVSQPFLMDLEGKQ GNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRF QTLLALHRSYLTGDSGWTAGAAA YYVGYLQPRTFLLKYNENGTITDAVD CALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNAT RFASVYAWNKRKRISNCVADYSVLYN SASFSTFKCYGVSPTKLNDLCFTNVY ADSFVIRGDEVQRQIAPGQTGKIADYN YKL PDDFTGCVIAWNSNNLDSKVG NYYLYRLFRKSNLKPFERDISTEIQ AGSTPCNGVEGFNCYFPLQSYGFQPT NGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNENGLTGTG VLTESNKKFLPFQQFGRDIADTTDAV RDPQTLEILDITPCSFSGGVSVITPGTNT SNQVAVLYQDVNCTEVPVAIHADQL TPTWRVYSTGSNVFQTRAGCLIGAEH VNNSYECDIPIGAGICASYQTQTNSPR RARSVASQSIAYTMSLGAENSVAYS NNSIAIPTNFTISVTTEILPVSMTKTSV DCTMYICGDSTECSNLLLQYGSFCTQ LNRALTGIAVEQDKNTQEVFAQVKQI YKTPPIKDFGGFNFSQILPDPSKPSKRS FIEDLLFNKVTADAGFIKQYGDCLG DIAARDLICAQKFNGLTVLPPLLTDE MIAQYTSALLAGTITSGWTFGAGAAL	13

		QIPFAMQMAYRFNGIGVTQNVLYEN QKLIANQFNSAIGKIQDSLSTASALG KLQDVVNQNAQALNTLVKQLSSNFG AISSVLNDILSRDKVEAEVQIDRLITG RLQSLQTYVTQQLIRAAEIRASANLA ATKMSECVLGQSKRVDFCGKGYHLM SFPQSAPHGVVFLHVTYVPAQEKNT TAPAICHGDKAHFPREGVVFVSNGTHW FVTQRNFYEPQIITDNTFVSGNCDVV IGIVNNTVYDPLQPELDSFKEELDKYF KNHTSPDVDLGDIGINASVVNIQKEI DRLNEVAKNLNESLIDLQELGKYEQY IKWPWYIWLGFIAGLIAIVMVTIMLCC MTSCCCLKGCCSCGSCCKFDEDDSE PVLKGVKLHYT*	
	Nucleic acid	ATGTTTGTTCCTTGTTCCTTATTGCC ACTAGTCTCTAGTCAGTGTGTTAAT CTTACAACCAGAACTCAATTACCCC CTGCATACACTAATTCTTTCACACGT GGTGTTTATTACCCTGACAAAGTTTT CAGATCCTCAGTTTTACATTCAACTC AGGACTTGTTCTTACCTTCTTTTCC AATGTTACTTGGTTCATGCTATAC ATGTCTCTGGGACCAATGGTACTAA GAGGTTTGATAACCCTGTCCTACCA TTAAATGATGGTGTTCATTTTGCTTC CACTGAGAAGTCTAACATAATAAGA GGCTGGATTTTGGTACTACTTTAG ATTCGAAGACCCAGTCCCTACTTAT TGTTAATAACGCTACTAATGTTGTT ATTAAAGTCTGTGAATTTCAATTTTG TAATGATCCATTTTGGGTGTTTATT ACCACAAAACAACAAAAGTTGGA TGGAAAGTGAGTTCAGAGTTTATTC TAGTGCGAATAATTGCACTTTTGAA TATGTCTCTCAGCCTTTTCTTATGGA CCTTGAAGGAAAACAGGGTAATTC AAAAATCTTAGGGAATTTGTGTTTA AGAATATTGATGGTTATTTTAAAAT ATATTCTAAGCACACGCCTATTAAT TTAGTGCGTGATCTCCCTCAGGGTTT TTCGGCTTTAGAACCATTGGTAGAT TTGCCAATAGGTATTAACATCACTA GGTTTCAAACCTTACTTGCTTTACAT AGAAGTTATTTGACTCCTGGTGATT CTCTTCAGGTTGGACAGCTGGTGC TGCAGCTTATTATGTGGGTATCTTC AACCTAGGACTTTTCTATTAAAATA TAATGAAAATGGAACCATTACAGAT GCTGTAGACTGTGCACTTGACCCTC	14

		TCTCAGAAACAAAGTGTACGTTGAA ATCCTTCACTGTAGAAAAAGGAATC TATCAAACCTTCTAACTTTAGAGTCC AACCAACAGAATCTATTGTTAGATT TCCTAATATTACAAACTTGTGCCCTT TTGGTGAAGTTTTTAACGCCACCAG ATTTGCATCTGTTTATGCTTGGAACA GGAAGAGAATCAGCAACTGTGTTGC TGATTATTCTGTCCTATATAATTCCG CATCATTTTCCACTTTTAAGTGTTAT GGAGTGTCTCCTACTAAATTAATG ATCTCTGCTTTACTAATGTCTATGCA GATTCATTTGTAATTAGAGGTGATG AAGTCAGACAAATCGCTCCAGGGCA AACTGGAAAGATTGCTGATTATAAT TATAAATTACCAGATGATTTTACAG GCTGCGTTATAGCTTGGAATTCTAA CAATCTTGATTCTAAGGTTGGTGGT AATTATAATTACCTGTATAGATTGTT TAGGAAGTCTAATCTCAAACCTTTT GAGAGAGATATTTCAACTGAAATCT ATCAGGCCGGTAGCACACCTTGTA TGGTGTGAAGGTTTTAATTGTTACT TTCCTTTACAATCATATGGTTTCCAA CCCACTAATGGTGTGTTGGTTACCAAC CATACAGAGTAGTAGTACTTTCTTTT GAACTTCTACATGCACCAGCAACTG TTTGTGGACCTAAAAAGTCTACTAA TTTGGTTAAAAACAAATGTGTCAAT TTCAACTTCAATGGTTTAAACAGGCA CAGGTGTTCTTACTGAGTCTAACAA AAAGTTTCTGCCTTTCCAACAATTG GCAGAGACATTGCTGACACTACTGA TGCTGTCCGTGATCCACAGACACTT GAGATTCTTGACATTACACCATGTT CTTTTGGTGGTGTGAGTGTATAAC ACCAGGAACAAATACTTCTAACCAG GTTGCTGTTCTTTATCAGGATGTTAA CTGCACAGAAGTCCCTGTTGCTATT CATGCAGATCAACTTACTCCTACTT GGCGTGTATTATTCTACAGGTTCTAAT GTTTTTCAAACACGTGCAGGCTGTT TAATAGGGGCTGAACATGTCAACAA CTCATATGAGTGTGACATACCCATT GGTGCAGGTATATGCGCTAGTTATC AGACTCAGACTAATTCTCCTCGGCG GGCACGTAGTGTAGCTAGTCAATCC ATCATTGCCTACACTATGTCACTTG GTGCAGAAAATTGAGTTGCTTACTC TAATAACTCTATTGCCATACCCACA	
--	--	--	--

		AATTTTACTATTAGTGTTACCACAG AAATTCTACCAGTGTCTATGACCAA GACATCAGTAGATTGTACAATGTAC ATTTGTGGTGATTCAACTGAATGCA GCAATCTTTTGTTGCAATATGGCAG TTTTTGTACACAATTAAACCGTGCTT TAACTGGAATAGCTGTTGAACAAGA CAAAAACACCCAAGAAGTTTTTGCA CAAGTCAAACAAATTTACAAAACAC CACCAATTAAGATTTTGGTGGTTT TAATTTTTCACAAATATTACCAGAT CCATCAAAACCAAGCAAGAGGTCAT TTATTGAAGATCTACTTTTCAACAA AGTGACACTTGCAGATGCTGGCTTC ATCAAACAATATGGTGATTGCCTTG GTGATATTGCTGCTAGAGACCTCAT TTGTGCACAAAAGTTTAACGGCCTT ACTGTTTTGCCACCTTTGCTCACAGA TGAAATGATTGCTCAATACACTTCT GCACTGTTAGCGGGTACAATCACTT CTGGTTGGACCTTTGGTGCAGGTGC TGCATTACAAATACCATTGCTATG CAAATGGCTTATAGGTTAATGGTA TTGGAGTTACACAGAATGTTCTCTA TGAGAACCACAAAATTGATTGCCAAC CAATTTAATAGTGCTATTGGCAAAA TTCAAGACTCACTTTCTTCCACAGC AAGTGCCTTGGAAAACCTTCAAGAT GTGGTCAACCAAAATGCACAAGCTT TAAACACGCTTGTTAAACAACCTTAG CTCCAATTTTGGTGCAATTTCAAGT GTTTTAAATGATATCCTTTCACGTCT TGACAAAGTTGAGGCTGAAGTGCAA ATTGATAGGTTGATCACAGGCAGAC TTCAAAGTTTGCAGACATATGTGAC TCAACAATTAATTAGAGCTGCAGAA ATCAGAGCTTCTGCTAATCTTGCTG CTACTAAAATGTCAGAGTGTGTACT TGGACAATCAAAAAGAGTTGATTTT TGTGGAAAGGGCTATCATCTTATGT CCTTCCCTCAGTCAGCACCTCATGG TGTAAGTCTTCTTGCATGTGACTTATG TCCCTGCACAAGAAAAGAACTTCAC AACTGCTCCTGCCATTTGTCATGAT GGAAAAGCACACTTTCCTCGTGAAG GTGTCTTTGTTTCAAATGGCACACA CTGGTTTGTAAACACAAAGGAATTTT TATGAACCACAAATCATTACTACAG ACAACACATTTGTGTCTGGTAACTG TGATGTTGTAATAGGAATTGTCAAC	
--	--	---	--

		AACACAGTTTATGATCCTTTGCAAC CTGAATTAGACTCATTCAAGGAGGA GTTAGATAAATATTTTAAGAATCAT ACATCACCAGATGTTGATTTAGGTG ACATCTCTGGCATTAAATGCTTCAGTT GTAAACATTCAAAAAGAAATTGACC GCCTCAATGAGGTTGCCAAGAATTT AAATGAATCTCTCATCGATCTCCAA GAACTTGGAAAGTATGAGCAGTATA TAAAATGGCCATGGTACATTTGGCT AGGTTTATAGCTGGCTTGATTGCC ATAGTAATGGTGACAATTATGCTTT GCTGTATGACCAGTTGCTGTAGTTG TCTCAAGGGCTGTTGTTCTTGTGGAT CCTGCTGCAAATTTGATGAAGACGA CTCTGAGCCAGTGCTCAAAGGAGTC AAATTACATTACACATAA	
L protein from EMCV	Amino acid	MATTMEQETCAHSLTFECPKCSALQ YRNGFYLLKYDEEWYPEELLTDGED DVFDPELDMEVVFELQ	2
	Nucleic acid	ATGGCCACAACCATGGAACAAGAG ACTTGCGCGCACTCTCTCACTTTTGA GGAATGCCCAAATGCTCTGCTCTA CAATACCGTAATGGATTTTACCTGC TAAAGTATGATGAAGAATGGTACCC AGAGGAGTTATTGACTGATGGAGAG GATGATGTCTTTGATCCCGAATTAG ACATGGAAGTCGTTTTTCGAGTTACA G	16
2A site	Amino acid	GSGATNFSLLKQAGDVEENPGP	17
	Nucleic acid	GGAAGCGGAGCTACTAACTTCAGCC TGCTGAAGCAGGCTGGAGATGTGGA GGAGAACCCTGGACCT	18
SARS-CoV-2 M protein	Amino acid	MADSNGTITVEELKKLLEQWNLVIGF LFLTWICLLQFAYANRNRFLYIIKLIFL WLLWPVTLACFVLAAYRINWITGGI AIAMACLVGLMWLSYFIASFRLFART RSMWSFNPETNILLNVPLHGTILTRPL LESELVIGAVILRGHLRIAGHHLGRCD IKDLPKEITVATSRTLSYYKLGASQRV AGDSGFAAYSRYRIGNYKLNTDHSSS SDNIALLVQ*	33
	Nucleic acid	ATGGCAGATTCCAACGGTACTATTA CCGTTGAAGAGCTTAAAAAGCTCCT TGAACAATGGAACCTAGTAATAGGT TTCCTATTCTTACATGGATTTGTCT TCTACAATTTGCCTATGCCAACAGG	19

		AATAGGTTTTTGTATATAATTAAGTT AATTTTCCTCTGGCTGTTATGGCCAG TAACTTTAGCTTGTTTTGTGCTTGCT GCTGTTTACAGAATAAATTGGATCA CCGGTGGAATTGCTATCGCAATGGC TTGTCTTGTAGGCTTGATGTGGCTCA GCTACTTCATTGCTTCTTTTCAGACTG TTTGCGCGTACGCGTTCCATGTGGT CATTCAATCCAGAACTAACATTCT TCTCAACGTGCCACTCCATGGCACT ATTCTGACCAGACCGCTTCTAGAAA GTGAACTCGTAATCGGAGCTGTGAT CCTTCGTGGACATCTTCGTATTGCTG GACACCATCTAGGACGCTGTGACAT CAAGGACCTGCCTAAAGAAATCACT GTTGCTACATCACGAACGCTTCTT ATTACAAATTGGGAGCTTCGCAGCG TGTAGCAGGTGACTCAGGTTTTGCT GCATACAGTCGCTACAGGATTGGCA ACTATAAATTAAACACAGACCATTCT CAGTAGCAGTGACAATATTGCTTTG CTTGTACAGTAA	
SARS-CoV-2 N protein	Amino acid	MSDNGPQNQRNAPRITFGGPSDSTGS NQNGERSGARSKQRRPQGLPNNTAS WFTALTQHGKEDLKFPRGQGVPIINTN SSPDDQIGYYRRATRIRGGDGKMKD LSPRWYFYLLGTGPEAGLPYGANKD GIIWVATEGALNTPKDHIGTRNPANN AAIVLQLPQGTTLPKGFYAEGSRGGS QASSRSSRSRNSSRNSTPGSSRGTS ARMAGNGGDAALALLLDRLNQLES KMSGKGQQQGGQTVTKKSAAEASK KPRQKRTATKAYNVTQAFGRRGPEQ TQGNFGDQELIRQGTDYKHWPQIAQF APSASAFFGMSRIGMEVTPSGTWLTY TGAIKLDDKDPNFKDQVILLNKHIDA YKTFPPTEPKKDKKKKADETQALPQR QKKQQTVTLLPAADLDDFSKQLQQS MSSADSTQA	20
	Nucleic acid	ATGTCTGATAATGGACCCCAAAATC AGCGAAATGCACCCCGCATTACGTT TGGTGGACCCTCAGATTCAACTGGC AGTAACCAGAATGGAGAACGCAGT GGGGCGCGATCAAAACAACGTCGG CCCCAAGGTTTACCCAATAATACTG CGTCTTGGTTCACCGCTCTCACTCAA CATGGCAAGGAAGACCTTAAATTCC CTCGAGGACAAGGCGTTCCAATTAA CACCAATAGCAGTCCAGATGACCAA	21

		ATTGGCTACTACCGAAGAGCTACCA GACGAATTCGTGGTGGTGACGGTAA AATGAAAGATCTCAGTCCAAGATGG TATTTCTACTACCTAGGAACTGGGC CAGAAGCTGGACTTCCCTATGGTGC TAACAAAGACGGCATCATATGGGTT GCAACTGAGGGAGCCTTGAATACAC CAAAAGATCACATTGGCACCCGCAA TCCTGCTAACAATGCTGCAATCGTG CTACAACTTCCTCAAGGAACAACAT TGCCAAAAGGCTTCTACGCAGAAGG GAGCAGAGGCGGCAGTCAAGCCTCT TCTCGTTCCTCATCACGTAGTCGCA ACAGTTCAAGAAATTCAACTCCAGG CAGCAGTAGGGGAACTTCTCCTGCT AGAATGGCTGGCAATGGCGGTGATG CTGCTCTTGCTTTGCTGCTGCTTGAC AGATTGAACCAGCTTGAGAGCAAA ATGTCTGGTAAAGGCCAACAACAAC AAGGCCAACTGTCATAAGAAATC TGCTGCTGAGGCTTCTAAGAAGCCT CGGCAAAAACGTACTGCCACTAAAG CATACAATGTAACACAAGCTTTCGG CAGACGTGGTCCAGAACAACCCCA AGGAAATTTTGGGGACCAGGAACTA ATCAGACAAGGAACTGATTACAAAC ATTGGCCGCAAATTGCACAATTTGC CCCCAGCGCTTCAGCGTTCTTCGGA ATGTCGCGCATTGGCATGGAAGTCA CACCTTCGGGAACGTGGTTGACCTA CACAGGTGCCATCAAATTGGATGAC AAAGATCCAAATTTCAAAGATCAAG TCATTTTGCTGAATAAGCATATTGA CGCATACAAAACATTCCCACCAACA GAGCCTAAAAAGGACAAAAAGAAG AAGGCTGATGAACTCAAGCCTTAC CGCAGAGACAGAAAGAAACAGCAAA CTGTGACTCTTCTTCCTGCTGCAGAT TTGGATGATTTCTCCAAACAATTGC AACAATCCATGAGCAGTGCTGACTC AACTCAGGCC	
SARS-CoV-2 E protein	Amino acid	MYSFVSEETGTLIVNSVLLFLAFVVFL LVTLAILTALRLCAYCCNIVNVSLVKP SFYVYSRVKNLNSSRVPDLLV*	22
	Nucleic acid	ATGTACTCATTTCGTTTCGGAAGAGA CAGGTACGTTAATAGTTAATAGCGT ACTTCTTTTCTTGCTTTCGTGGTAT TCTTGCTAGTTACACTAGCCATCCTT ACTGCGCTTCGATTGTGTGCGTACT	23

		GCTGCAATATTGTTAACGTGAGTCT TGTAACCTTCTTTTACGTTTACT CTCGTGTTAAAAATCTGAATTCTCT AGAGTTCCTGATCTTCTGGTCTAA	
IRES	Nucleic acid	TCCCCCCCCCTAACGTTACTGGCC GAAGCCGCTTGAATAAGGCCGGTG TGCGTTTGTCTATATGTTATTTCCA CCATATTGCCGTCTTTTGGCAATGTG AGGGCCCCGAAACCTGGCCCTGTCT TCTTGACGAGCATTCTAGGGGTCT TTCCCCTCTCGCCAAAGGAATGCAA GGTCTGTTGAATGTCGTGAAGGAAG CAGTTCCTCTGGAAGCTTCTTGAAG ACAAACAACGTCTGTAGCGACCCTT TGCAGGCAGCGGAACCCCCACCTG GCGACAGGTGCCTCTGCGGCCAAAA GCCACGTGTATAAGATACACCTGCA AAGGCGGCACAACCCCAGTGCCAC GTTGTGAGTTGGATAGTTGTGGA GAGTCAAATGGCTCTCCTCAAGCGT ATTCAACAAGGGGCTGAAGGATGCC CAGAAGGTACCCATTGTATGGGAT CTGATCTGGGGCCTCGGTGCACATG CTTTACATGTGTTTAGTCGAGGTTA AAAAAACGTCTAGGCCCCCGAACC ACGGGGACGTGGTTTTCTTTGAAA AACACGATGATAAT	24
CoVEG2 polypeptide 1 (fusion of EMCV L protein, 2A site, M protein)	Amino acid	MATTMEQETCAHSLTFEECPKCSALQ YRNGFYLLKYDEEWYPEELLTDGED DVFDPELDMEVVFELQSGATNFSLL KQAGDVEENPGPMADSNGTITVEELK KLLEQWNLVIGFLFLTWICLLQFAYA NRNRFYIIKLIFLWLLWPVTLACFVL AAVYRINWITGGIAIAMACLVGLMW LSYFIASFRLFARTRSMWSFNPETNIL LNVPLHGTLTRPLLESELVIGAVILRG HLRIAGHHLGRCDIKDLPKEITVATSR TLSYYKL GASQRVAGDSGFAAYSRY RIGNYKLNTDHSSSDNIALLVQ*	25
CoVEG2 polypeptide 2 (fusion of EMCV L protein, 2A site, N protein, 2A site and E protein)	Amino acid	MATTMEQETCAHSLTFEECPKCSALQ YRNGFYLLKYDEEWYPEELLTDGED DVFDPELDMEVVFELQSGATNFSLL KQAGDVEENPGPMSDNGPQNQRNAP RITFGGPS DSTGSNQNGERSGARSKQ RRPQGLPNNTASWFTALTQHGKEDL KFPRGQGV PINTNSSPDDQIGYYRRAT RRIRGGDGKMKDLSPRWYFY YLGTG PEAGLPYGANKDGIWVATEGALNTP KDHIGTRNPANNAIIVLQLPQGTTLP	26

		KGFYAEGSRGGSQASSRSSSRN SSR NSTPGSSRGTS PARMAGNGGDAALA LLLLDRLNQLESKMSGKGQQQGGQT VTKKSAAEASKKPRQKRTATKAYNV TQAFGRRGPEQTQGNFGDQELIRQGT DYKHWPQIAQFAPSASAFFGMSRIGM EVTPSGTWLTYTGAIKLDDKDPNFKD QVILLNKHIDAYKTFPPTEPKDKKK KADETQALPQRQKKQQTVTLLPAAD LDDFSKQLQQSMSSADSTQAGSGATN FSLLKQAGDVEENPGPMYSFVSEETG TLIVNSVLLFLAFVVFLVTLAILTAL RLCAYCCNIVNVSLVKPSFYVYSRVK NLNSSRVPDLLV*	
CoVEG1 polypeptide (fusion of EMCV L protein, 2A site, M protein, 2A site and E protein)	Amino acid	MATTMEQETCAHSLTFEECPKCSALQ YRNGFYLLKYDEEWYPEELLTDGED DVFDPELDMEVVFELQSGATNFSLL KQAGDVEENPGPMADSNGTITVEELK KLEQWNLVIGFLFLTWICLLQFAYA NNRFLYIIKLIFLWLLWPVTLACFVL AAVYRINWITGGIAIAMACLVGLMW LSYFIASFRLFARTRSMWSFNPETNIL LNVPLHGTILTRP LLESELVIGAVILRG HLRIAGHHLGRCDIKDLPKEITVATSR TLSYYKLGASQRVAGDSGFAAYSRY RIGNYKLNTDHSSSDNIALLVQSGG ATNFSLLKQAGDVEENPGPMYSFVSE ETGTLIVNSVLLFLAFVVFLVTLAILT ALRLCAYCCNIVNVSLVKPSFYVYSR VKNLNSSRVPDLLV*	32
CMV enhancer	Nucleic acid	GACATTGATTATTGACTAGTTATTA ATAGTAATCAATTACGGGGTCATTA GTTCATAGCCCATATATGGAGTTCC GCGTTACATAACTTACGGTAAATGG CCCGCCTGGCTGACCGCCCAACGAC CCCC GCCCATTGACGTCAATAATGA CGTATGTTCCCATAGTAACGCCAAT AGGGACTTTCCATTGACGTCAATGG GTGGAGTATTTACGGTAAACTGCCC ACTTGGCAGTACATCAAGTGTATCA TATGCCAAGTACGCCCCCTATTGAC GTCAATGACGGTAAATGGCCCGCCT GGCATTATGCCCAGTACATGACCTT ATGGGACTTTCTACTTGGCAGTAC ATCTACGTATTAGTCATCGCTATTAC CATG	27
CMV promoter	Nucleic acid	GTGATGCGGTTTTTGGCAGTACATCA ATGGGCGTGGATAGCGGTTTGACTC ACGGGGATTTC CAAGTCTCCACCCC	28

		ATTGACGTCAATGGGAGTTTGT GGCACCAAAATCAACGGGACTTTCC AAAATGTCGTAACAACCTCCGCCCA TTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAG AGCT	
Cloning site	Nucleic acid	GGTTTAGTGAACCGTCAGATCCGCT AGCGCTACCGGACTCAGATCTCGAG CTCAAGCTTCGAATTCTGCAGTCGA CGGTACCGCGGGCCCGGGATCCACC GGTCGCCACG	29
CoVEG2 insert sequence	Nucleic acid	GACATTGATTATTGACTAGTTATTA ATAGTAATCAATTACGGGGTCATTA GTTTCATAGCCCATATATGGAGTTCC GCGTTACATAACTTACGGTAAATGG CCCGCCTGGCTGACCGCCCAACGAC CCCCGCCATTGACGTCAATAATGA CGTATGTTCCCATAGTAACGCCAAT AGGGACTTTCCATTGACGTCAATGG GTGGAGTATTTACGGTAAACTGCCC ACTTGGCAGTACATCAAGTGTATCA TATGCCAAGTACGCCCCCTATTGAC GTCAATGACGGTAAATGGCCCGCCT GGCATTATGCCCAGTACATGACCTT ATGGGACTTTCTACTTGGCAGTAC ATCTACGTATTAGTCATCGCTATTAC CATGGTGATGCGGTTTGGCAGTAC ATCAATGGGCGTGGATAGCGGTTTG ACTCACGGGGATTTCCAAGTCTCCA CCCCATTGACGTCAATGGGAGTTTG TTTTGGCACCAAAATCAACGGGACT TTCCAAAATGTCGTAACAACCTCCGC CCCATTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAA GCAGAGCTGGTTTAGTGAACCGTCA GATCCGCTAGCGCTACCGGACTCAG ATCTCGAGCTCAAGCTTCGAATTCT GCAGTCGACGGTACCGCGGGCCCGG GATCCACCGGTCGCCACGATGTTTG TTTTTCTTGTTTTATTGCCACTAGTC TCTAGTCAGTGTGTTAATCTTACAA CCAGAACTCAATTACCCCTGCATA CACTAATTCTTTCACACGTGGTGTTT ATTACCTGACAAAGTTTTTCAGATC CTCAGTTTTACATTCAACTCAGGAC TTGTTCTTACCTTTCTTTTCCAATGT TACTTGGTTCCATGCTATACATGTCT CTGGGACCAATGGTACTAAGAGGTT TGATAACCCTGTCCTACCATTTAAT	30

		GATGGTGTTTATTTTGCTTCCACTGA GAAGTCTAACATAATAAGAGGCTGG ATTTTGGTACTACTTTAGATTTCGAA GACCCAGTCCCTACTTATTGTTAAT AACGCTACTAATGTTGTTATTAAAG TCTGTGAATTTCAATTTTGTAATGAT CCATTTTGGGTGTTTATTACCACAA AAACAACAAAAGTTGGATGGAAAG TGAGTTCAGAGTTTATTCTAGTGCG AATAATTGCACCTTTGAATATGTCTC TCAGCCTTTTCTTATGGACCTTGAAG GAAAACAGGGTAATTTCAAAAATCT TAGGGAATTTGTGTTTAAGAATATT GATGGTTATTTTAAAATATATTCTA AGCACACGCCTATTAATTTAGTGCG TGATCTCCCTCAGGGTTTTTCGGCTT TAGAACCATTGGTAGATTTGCCAAT AGGTATTAACATCACTAGGTTTCAA ACTTTACTTGCTTTACATAGAAGTTA TTGACTCCTGGTGATTCTTCTTCAG GTTGGACAGCTGGTGCTGCAGCTTA TTATGTGGGTATCTTCAACCTAGG ACTTTTCTATTAATAATATAATGAAA ATGGAACCATTACAGATGCTGTAGA CTGTGCACTTGACCCTCTCTCAGAA ACAAAGTGTACGTTGAAATCCTTCA CTGTAGAAAAAGGAATCTATCAAAC TTCTAACTTTAGAGTCCAACCAACA GAATCTATTGTAGATTTTCTAATAT TACAACTTGTGCCCTTTTGGTGAA GTTTTTAACGCCACCAGATTTGCAT CTGTTTATGCTTGGAACAGGAAGAG AATCAGCAACTGTGTTGCTGATTAT TCTGTCCTATATAATTCCGCATCATT TTCCACTTTTAAGTGTTATGGAGTGT CTCCTACTAAATTAATGATCTCTG CTTTACTAATGTCTATGCAGATTCAT TTGTAATTAGAGGTGATGAAGTCAG ACAAATCGCTCCAGGGCAAACCTGGA AAGATTGCTGATTATAATTATAAAT TACCAGATGATTTTACAGGCTGCGT TATAGCTTGGAATTCTAACAATCTT GATTCTAAGGTTGGTGGTAATTATA ATTACCTGTATAGATTGTTTAGGAA GTCTAATCTCAAACCTTTTGAGAGA GATATTTCAACTGAAATCTATCAGG CCGGTAGCACACCTTGTAATGGTGT TGAAGGTTTTAATTGTTACTTTCCTT TACAATCATATGGTTTCCAACCCAC TAATGGTGTTGGTTACCAACCATAC	
--	--	---	--

		AGAGTAGTAGTACTTTCTTTTGAAC TTCTACATGCACCAGCAACTGTTTG TGGACCTAAAAAGTCTACTAATTTG GTTAAAAACAAATGTGTCAATTTCA ACTTCAATGGTTTAACAGGCACAGG TGTTCTTACTGAGTCTAACAAAAAG TTTCTGCCTTTCCAACAATTTGGCAG AGACATTGCTGACACTACTGATGCT GTCCGTGATCCACAGACACTTGAGA TTCTTGACATTACACCATGTTCTTTT GGTGGTGTGAGTGTTATAACACCAG GAACAAATACTTCTAACCAGGTTGC TGTTCTTTATCAGGATGTAACTGCA CAGAAGTCCCTGTTGCTATTCATGC AGATCAACTTACTCCTACTTGGCGT GTTTATTCTACAGGTTCTAATGTTT TCAAACACGTGCAGGCTGTTAATA GGGGCTGAACATGTCAACAACATCAT ATGAGTGTGACATAACCATTTGGTGC AGGTATATGCGCTAGTTATCAGACT CAGACTAATTCTCCTCGGCGGGCAC GTAGTGTAGCTAGTCAATCCATCAT TGCCTACACTATGTCACTTGGTGCA GAAAATTGAGTTGCTTACTCTAATA ACTCTATTGCCATACCCACAAATTTT ACTATTAGTGTTACCACAGAAATTC TACCAGTGTCTATGACCAAGACATC AGTAGATTGTACAATGTACATTTGT GGTGATTCAACTGAATGCAGCAATC TTTTGTTGCAATATGGCAGTTTTTGT ACACAATTAAACCGTGCTTTAACTG GAATAGCTGTTGAACAAGACAAAA ACACCCAAGAAGTTTTTGCACAAGT CAAACAAATTTACAAAACACCACCA ATTAAAGATTTTGGTGGTTTTAATTT TTCACAAATATTACCAGATCCATCA AAACCAAGCAAGAGGTCATTTATTG AAGATCTACTTTTCAACAAAGTGAC ACTTGCAGATGCTGGCTTCATCAAA CAATATGGTGATTGCCTTGGTGATA TTGCTGCTAGAGACCTCATTTGTGC ACAAAAGTTTAAACGGCCTTACTGTT TTGCCACCTTTGCTCACAGATGAAA TGATTGCTCAATACACTTCTGCACT GTTAGCGGGTACAATCACTTCTGGT TGGACCTTTGGTGCAGGTGCTGCAT TACAAATACCATTTGCTATGCAAAT GGCTTATAGGTTTAATGGTATTGGA GTTACACAGAATGTTCTCTATGAGA ACCAAAAATTGATTGCCAACCAATT	
--	--	--	--

		TAATAGTGCTATTGGCAAATTCAA GACTCACTTTCTTCCACAGCAAGTG CACTTGGAACCTTCAAGATGTGGT CAACCAAAATGCACAAGCTTTAAAC ACGCTTGTTAAACAACCTAGCTCCA ATTTTGGTGCAATTTCAAGTGTTTA AATGATATCCTTTCACGTCTTGACA AAGTTGAGGCTGAAGTGCAAATTGA TAGGTTGATCACAGGCAGACTTCAA AGTTTGCAGACATATGTGACTCAAC AATTAATTAGAGCTGCAGAAATCAG AGCTTCTGCTAATCTTGCTGCTACTA AAATGTCAGAGTGTGTACTTGGACA ATCAAAAAGAGTTGATTTTTGTGGA AAGGGCTATCATCTTATGTCCTTCCC TCAGTCAGCACCTCATGGTGATGTC TTCTTGCAATGTGACTTATGTCCCTGC ACAAGAAAAGAACTTCACAACTGCT CCTGCCATTTGTCATGATGGAAAAG CACACTTTCCTCGTGAAGGTGTCTTT GTTTCAAATGGCACACACTGGTTTG TAACACAAAGGAATTTTTATGAACC ACAAATCATTACTACAGACAACACA TTTGTGTCTGGTAACTGTGATGTTGT AATAGGAATTGTCAACAACACAGTT TATGATCCTTTGCAACCTGAATTAG ACTCATTCAAGGAGGAGTTAGATAA ATATTTTAAGAATCATAATCACCA GATGTTGATTTAGGTGACATCTCTG GCATTAATGCTTCAGTTGTAAACAT TCAAAAAGAAATTGACCGCCTCAAT GAGGTTGCCAAGAATTTAAATGAAT CTCTCATCGATCTCCAAGAACTTGG AAAGTATGAGCAGTATATAAAATGG CCATGGTACATTTGGCTAGGTTTTAT AGCTGGCTTGATTGCCATAGTAATG GTGACAATTATGCTTTGCTGTATGA CCAGTTGCTGTAGTTGTCTCAAGGG CTGTTGTTCTTGTGGATCCTGCTGCA AATTTGATGAAGACGACTCTGAGCC AGTGCTCAAAGGAGTCAAATTACAT TACACATAATCCCCCCCCCTAACG TACTGGCCGAAGCCGCTTGGAATA AGGCCGGTGTGCGTTTGTCTATATG TTATTTTCCACCATATTGCCGTCTTT TGGCAATGTGAGGGCCCGGAAACCT GGCCCTGTCTTCTTGACGAGCATTC CTAGGGGTCTTTCCCCTCTCGCCAA AGGAATGCAAGGTCTGTTGAATGTC GTGAAGGAAGCAGTTCCTCTGGAAG	
--	--	--	--

		CTTCTTGAAGACAAACAACGTCTGT AGCGACCCTTTGCAGGCAGCGGAAC CCCCACCTGGCGACAGGTGCCTCT GCGGCCAAAAGCCACGTGTATAAG ATACACCTGCAAAGGCGGCACAACC CCAGTGCCACGTTGTGAGTTGGATA GTTGTGGAAAGAGTCAAATGGCTCT CCTCAAGCGTATTCAACAAGGGGCT GAAGGATGCCCAGAAGGTACCCCAT TGTATGGGATCTGATCTGGGGCCTC GGTGCACATGCTTTACATGTGTTTA GTCGAGGTTAAAAAACGTCTAGGC CCCCGAACCACGGGGACGTGGTTT TCCTTTGAAAAACACGATGATAATA TGGCCACAACCATGGAACAAGAGA CTTGCGCGCACTCTCTCACTTTTGAG GAATGCCCAAAATGCTCTGCTCTAC AATACCGTAATGGATTTTACCTGCT AAAGTATGATGAAGAATGGTACCCA GAGGAGTTATTGACTGATGGAGAGG ATGATGTCTTTGATCCCGAATTAGA CATGGAAGTCGTTTTTCGAGTTACAG GGAAGCGGAGCTACTAACTTCAGCC TGCTGAAGCAGGCTGGAGATGTGGA GGAGAACCCTGGACCTATGGCAGAT TCCAACGGTACTATTACCGTTGAAG AGCTTAAAAAGCTCCTTGAACAATG GAACCTAGTAATAGGTTTCCTATTCT CTTACATGGATTTGTCTTCTACAATT TGCCTATGCCAACAGGAATAGGTTT TTGTATATAATTAAGTTAATTTTCCT CTGGCTGTTATGGCCAGTAACTTTA GCTTGTTTTGTGCTTGCTGCTGTTTA CAGAATAAATTGGATCACCGGTGGA ATTGCTATCGCAATGGCTTGTCTTGT AGGCTTGATGTGGCTCAGCTACTTC ATTGCTTCTTTCAGACTGTTTGCGCG TACGCGTTCCATGTGGTCATTCAAT CCAGAAACTAACATTCTTCTCAACG TGCCACTCCATGGCACTATTCTGAC CAGACCGCTTCTAGAAAGTGAAGTC GTAATCGGAGCTGTGATCCTTCGTG GACATCTTCGTATTGCTGGACACCA TCTAGGACGCTGTGACATCAAGGAC CTGCCTAAAGAAATCACTGTTGCTA CATCACGAACGCTTTCTTATTACAA ATTGGGAGCTTCGCAGCGGTAGCA GGTGAATCAGGTTTTGCTGCATACA GTCGCTACAGGATTGGCAACTATAA ATTAAACACAGACCATTCCAGTAGC	
--	--	---	--

		AGTGACAATATTGCTTTGCTTGTAC AGTAACCCCCCCCCCTAACGTTACT GGCCGAAGCCGCTTGGAATAAGGCC GGTGTGCGTTTGTCTATATGTTATTT TCCACCATAATTGCCGTCTTTTGGCAA TGTGAGGGCCCCGGAACCTGGCCCT GTCTTCTTGACGAGCATTCCTAGGG GTCTTTCCCCTCTCGCCAAAGGAAT GCAAGGTCTGTTGAATGTCGTGAAG GAAGCAGTTCCTCTGGAAGCTTCTT GAAGACAAACAACGTCTGTAGCGA CCCTTTGCAGGCAGCGGAACCCCCC ACCTGGCGACAGGTGCCTCTGCGGC CAAAAGCCACGTGTATAAGATACAC CTGCAAAGGCGGCACAACCCCAAGTG CCACGTTGTGAGTTGGATAGTTGTG GAAAGAGTCAAATGGCTCTCCTCAA GCGTATTCAACAAGGGGCTGAAGG ATGCCCAGAAGGTACCCCATTTGTAT GGGATCTGATCTGGGGCCTCGGTGC ACATGCTTTACATGTGTTTAGTCGA GGTTAAAAAACGTCTAGGCCCCC GAACCACGGGGACGTGGTTTTCTTT TGAAAAACACGATGATAATATGGCC ACAACCATGGAACAAGAGACTTGC GCGCACTCTCTCACTTTTGAGGAAT GCCCAAATGCTCTGCTCTACAATA CCGTAATGGATTTTACCTGCTAAAG TATGATGAAGAATGGTACCCAGAGG AGTTATTGACTGATGGAGAGGATGA TGTCTTTGATCCCGAATTAGACATG GAAGTCGTTTTTCGAGTTACAGGGAA GCGGAGCTACTAACTTCAGCCTGCT GAAGCAGGCTGGAGATGTGGAGGA GAACCCTGGACCTATGTCTGATAAT GGACCCCAAAATCAGCGAAATGCA CCCCGCATTACGTTTGGTGGACCCT CAGATTCAACTGGCAGTAACCAGAA TGGAGAACGCAGTGGGGCGCGATC AAAACAACGTTCGGCCCCAAGGTTTA CCCAATAATACTGCGTCTTGGTTCA CCGCTCTCACTCAACATGGCAAGGA AGACCTTAAATTCCCTCGAGGACAA GGCGTTCCAATTAACACCAATAGCA GTCCAGATGACCAAAATTGGCTACTA CCGAAGAGCTACCAGACGAATTCGT GGTGGTGACGGTAAAATGAAAGAT CTCAGTCCAAGATGGTATTTCTACT ACCTAGGAACTGGGCCAGAAGCTG GACTTCCCTATGGTGCTAACAAAGA	
--	--	---	--

		CGGCATCATATGGGTTGCAACTGAG GGAGCCTTGAATACACCAAAAGATC ACATTGGCACCCGCAATCCTGCTAA CAATGCTGCAATCGTGCTACAACTT CCTCAAGGAACAACATTGCCAAAAG GCTTCTACGCAGAAGGGAGCAGAG GCGGCAGTCAAGCCTCTTCTCGTTC CTCATCACGTAGTCGCAACAGTTCA AGAAATTCAACTCCAGGCAGCAGTA GGGGAACCTTCTCCTGCTAGAAATGGC TGGCAATGGCGGTGATGCTGCTCTT GCTTTGCTGCTGCTTGACAGATTGA ACCAGCTTGAGAGCAAAATGTCTGG TAAAGGCCAACAACAACAAGGCCA AACTGTCACTAAGAAATCTGCTGCT GAGGCTTCTAAGAAGCCTCGGCCAAA AACGTACTGCCACTAAAGCATACAA TGTAACACAAGCTTTCGGCAGACGT GGTCCAGAACAACCCAAGGAAAT TTTGGGGACCAGGAATAATCAGAC AAGGAACTGATTACAAACATTGGCC GCAAATTGCACAATTTGCCCCCAGC GCTTCAGCGTTCTTCGGAATGTCGC GCATTGGCATGGAAGTCACACCTTC GGGAACGTGGTTGACCTACACAGGT GCCATCAAATTGGATGACAAAGATC CAAATTTCAAAGATCAAGTCATTTT GCTGAATAAGCATATTGACGCATAC AAAACATTCCCACCAACAGAGCCTA AAAAGGACAAAAAGAAGAAGGCTG ATGAAACTCAAGCCTTACCGCAGAG ACAGAAGAAACAGCAAACTGTGAC TCTTCTTCTGCTGCAGATTGATG ATTTCTCCAAACAATTGCAACAATC CATGAGCAGTGCTGACTCAACTCAG GCCGGAAGCGGAGCTACTAACTTCA GCCTGCTGAAGCAGGCTGGAGATGT GGAGGAGAACCCTGGACCTATGTAC TCATTTCGTTTCGGAAGAGACAGGTA CGTTAATAGTTAATAGCGTACTTCTT TTTCTTGCTTTCGTGGTATTCTTGCT AGTTACACTAGCCATCCTTACTGCG CTTCGATTGTGTGCGTACTGCTGCA ATATTGTTAACGTGAGTCTTGTA ACCTTCTTTTACGTTTACTCTCGTG TTAAAAATCTGAATTCTTCTAGAGT TCCTGATCTTCTGGTCTAA	
CoVEG1 insert sequence	Nucleic acid	GACATTGATTATTGACTAGTTATTA ATAGTAATCAATTACGGGGTCATTA GTTCATAGCCCATATATGGAGTTCC	31

		GCGTTACATAACTTACGGTAAATGG CCCGCCTGGCTGACCGCCCAACGAC CCCCGCCCATTTGACGTCAATAATGA CGTATGTTCCCATAGTAACGCCAAT AGGGACTTTCCATTGACGTCAATGG GTGGAGTATTTACGGTAAACTGCCC ACTTGGCAGTACATCAAGTGTATCA TATGCCAAGTACGCCCCCTATTGAC GTCAATGACGGTAAATGGCCCGCCT GGCATTATGCCCAGTACATGACCTT ATGGGACTTTCCTACTTGGCAGTAC ATCTACGTATTAGTCATCGCTATTAC CATGGTGATGCGGTTTTGGCAGTAC ATCAATGGGCGTGGATAGCGGTTTG ACTCACGGGGATTTCCAAGTCTCCA CCCCATTGACGTCAATGGGAGTTTG TTTTGGCACCAAAATCAACGGGACT TTCCAAAATGTCGTAACAACTCCGC CCCATTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAA GCAGAGCTGGTTTAGTGAACCGTCA GATCCGCTAGCGCTACCGGACTCAG ATCTCGAGCTCAAGCTTCGAATTCT GCAGTCGACGGTACCGCGGGCCCGG GATCCACCGGTCGCCACGATGTTTG TTTTTCTTGTTTTATTGCCACTAGTC TCTAGTCAGTGTGTTAATCTTACAA CCAGAACTCAATTACCCCTGCATA CACTAATTCTTTCACACGTGGTGTTT ATTACCTGACAAAGTTTTAGATC CTCAGTTTTACATTCAACTCAGGAC TTGTTCTTACCTTTCTTTTCCAATGT TACTTGGTTCCATGCTATACATGTCT CTGGGACCAATGGTACTAAGAGGTT TGATAACCCTGTCCTACCATTTAAT GATGGTGTTTATTTTGCTTCCACTGA GAAGTCTAACATAATAAGAGGCTGG ATTTTTGGTACTACTTTAGATTTCGAA GACCCAGTCCCTACTTATTGTTAAT AACGCTACTAATGTTGTTATTAAAG TCTGTGAATTTCAATTTTGTAATGAT CCATTTTTGGGTGTTTATTACCACAA AAACAACAAAAGTTGGATGGAAAG TGAGTTCAGAGTTTATTCTAGTGCG AATAATTGCACTTTTGAATATGTCTC TCAGCCTTTTCTTATGGACCTTGAAG GAAAACAGGGTAATTTCAAAAATCT TAGGGAATTTGTGTTAAGAATATT GATGGTTATTTTAAAATATATTCTA AGCACACGCCTATTAATTTAGTGCG	
--	--	---	--

		TGATCTCCCTCAGGGTTTTTCGGCTT TAGAACCATTGGTAGATTTGCCAAT AGGTATTAACATCACTAGGTTTCAA ACTTTACTTGCTTTACATAGAAGTTA TTGACTCCTGGTGATTCTTCTTCAG GTTGGACAGCTGGTGCTGCAGCTTA TTATGTGGGTTATCTTCAACCTAGG ACTTTTCTATTAATAATATAATGAAA ATGGAACCATTACAGATGCTGTAGA CTGTGCACTTGACCCTCTCTCAGAA ACAAAGTGTACGTTGAAATCCTTCA CTGTAGAAAAAGGAATCTATCAAAC TTCTAACTTTAGAGTCCAACCAACA GAATCTATTGTTAGATTTCTAATAT TACAACTTGTGCCCTTTTGGTGAA GTTTTTAACGCCACCAGATTTGCAT CTGTTTATGCTTGGAACAGGAAGAG AATCAGCAACTGTGTTGCTGATTAT TCTGTCCTATATAATTCCGCATCATT TTCCACTTTTAAGTGTTATGGAGTGT CTCCTACTAAATTAATGATCTCTG CTTTACTAATGTCTATGCAGATTCAT TTGTAATTAGAGGTGATGAAGTCAG ACAAATCGCTCCAGGGCAAACCTGGA AAGATTGCTGATTATAATTATAAAT TACCAGATGATTTTACAGGCTGCGT TATAGCTTGGAATTCTAACAATCTT GATTCTAAGGTTGGTGGTAATTATA ATTACCTGTATAGATTGTTTAGGAA GTCTAATCTCAAACCTTTTGAGAGA GATATTTCAACTGAAATCTATCAGG CCGGTAGCACACCTTGTAATGGTGT TGAAGGTTTTAATTGTTACTTTCCTT TACAATCATATGGTTTCCAACCCAC TAATGGTGTTGGTTACCAACCATAC AGAGTAGTAGTACTTTCTTTTGAAC TTCTACATGCACCAGCAACTGTTTG TGGACCTAAAAAGTCTACTAATTTG GTTAAAAACAAATGTGTCAATTTCA ACTTCAATGGTTTAAACAGGCACAGG TGTTCTTACTGAGTCTAACAAAAAG TTTCTGCCTTTCCAACAATTTGGCAG AGACATTGCTGACACTACTGATGCT GTCCGTGATCCACAGACACTTGAGA TTCTTGACATTACACCATGTTCTTTT GGTGGTGTCAGTGTTATAACACCAG GAACAAATACTTCTAACCAGGTTGC TGTTCTTTATCAGGATGTAACTGCA CAGAAGTCCCTGTTGCTATTCATGC AGATCAACTTACTCCTACTTGGCGT	
--	--	---	--

		GTTTATTCTACAGGTTCTAATGTTTT TCAAACACGTGCAGGCTGTTTAATA GGGGCTGAACATGTCAACAACATCAT ATGAGTGTGACATACCCATTGGTGC AGGTATATGCGCTAGTTATCAGACT CAGACTAATTCTCCTCGGCGGGCAC GTAGTGTAGCTAGTCAATCCATCAT TGCCTACACTATGTCACCTGGTGCA GAAAATTCAGTTGCTTACTCTAATA ACTCTATTGCCATACCCACAAATTTT ACTATTAGTGTTACCACAGAAATTC TACCAGTGTCTATGACCAAGACATC AGTAGATTGTACAATGTACATTTGT GGTGATTCAACTGAATGCAGCAATC TTTTGTTGCAATATGGCAGTTTTTGT ACACAATTAACCGTGCTTTAACTG GAATAGCTGTTGAACAAGACAAAA ACACCCAAGAAGTTTTTGCACAAGT CAAACAAATTTACAAAACACCACCA ATTAAAGATTTTGGTGGTTTTAATTT TTCACAAATATTACCAGATCCATCA AAACCAAGCAAGAGGTCATTTATTG AAGATCTACTTTTCAACAAAGTGAC ACTTGCAGATGCTGGCTTCATCAAA CAATATGGTGATTGCCTTGGTGATA TTGCTGCTAGAGACCTCATTTGTGC ACAAAAGTTTAACGGCCTTACTGTT TTGCCACCTTTGCTCACAGATGAAA TGATTGCTCAATACACTTCTGCACT GTTAGCGGGTACAATCACTTCTGGT TGGACCTTTGGTGCAGGTGCTGCAT TACAAATACCATTTGCTATGCAAAT GGCTTATAGGTTTAATGGTATTGGA GTTACACAGAATGTTCTCTATGAGA ACCAAAAATTGATTGCCAACCAATT TAATAGTGCTATTGGCAAAATTCAA GACTCACTTTCTTCCACAGCAAGTG CACTTGGAACCTTCAAGATGTGGT CAACCAAAAATGCACAAGCTTTAAAC ACGCTTGTTAAACAACCTTAGCTCCA ATTTTGGTGCAATTTCAAGTGTTTTA AATGATATCCTTTACGCTTTGACA AAGTTGAGGCTGAAGTGCAAATTGA TAGGTTGATCACAGGCAGACTTCAA AGTTTGCAGACATATGTGACTCAAC AATTAATTAGAGCTGCAGAAATCAG AGCTTCTGCTAATCTTGCTGCTACTA AAATGTCAGAGTGTGTACTTGGACA ATCAAAAAGAGTTGATTTTTGTGGA AAGGGCTATCATCTTATGTCCTTCCC	
--	--	--	--

		TCAGTCAGCACCTCATGGTGTAGTC TTCTTGCATGTGACTTATGTCCCTGC ACAAGAAAAGAACTTCACAACTGCT CCTGCCATTTGTCATGATGGAAAAG CACACTTTCCTCGTGAAGGTGTCTTT GTTTCAAATGGCACACACTGGTTTG TAACACAAAGGAATTTTATGAACC ACAAATCATTACTACAGACAACACA TTTGTGTCTGGTAACTGTGATGTTGT AATAGGAATTGTCAACAACACAGTT TATGATCCTTTGCAACCTGAATTAG ACTCATTCAAGGAGGAGTTAGATAA ATATTTTAAGAATCATAATCACCA GATGTTGATTTAGGTGACATCTCTG GCATTAATGCTTCAGTTGTAAACAT TCAAAAAGAAATTGACCGCCTCAAT GAGGTTGCCAAGAATTTAAATGAAT CTCTCATCGATCTCCAAGAACTGG AAAGTATGAGCAGTATATAAAATGG CCATGGTACATTTGGCTAGGTTTAT AGCTGGCTTGATTGCCATAGTAATG GTGACAATTATGCTTTGCTGTATGA CCAGTTGCTGTAGTTGTCTCAAGGG CTGTTGTTCTTGTGGATCCTGCTGCA AATTTGATGAAGACGACTCTGAGCC AGTGCTCAAAGGAGTCAAATTACAT TACACATAATCCCCCCCCCTAACG TACTGGCCGAAGCCGCTTGGAATA AGGCCGGTGTGCGTTTGTCTATATG TTATTTTCCACCATATTGCCGTCTTT TGGCAATGTGAGGGCCCGGAAACCT GGCCCTGTCTTCTTGACGAGCATTCT CTAGGGGTCTTTCCCTCTCGCCAA AGGAATGCAAGGTCTGTTGAATGTC GTGAAGGAAGCAGTTCCTCTGGAAG CTTCTTGAAGACAAACAACGTCTGT AGCGACCCTTTGCAGGCAGCGGAAC CCCCACCTGGCGACAGGTGCCTCT GCGGCCAAAAGCCACGTGTATAAG ATACACCTGCAAAGGCGGCACAACC CCAGTGCCACGTGTGTGAGTTGGATA GTTGTGGAAAGAGTCAAATGGCTCT CCTCAAGCGTATTCAACAAGGGGCT GAAGGATGCCCAGAAGGTACCCCAT TGTATGGGATCTGATCTGGGGCCTC GGTGCACATGCTTTACATGTGTTTA GTCGAGGTTAAAAAACGTCTAGGC CCCCGAACCACGGGGACGTGGTTT TCCTTTGAAAAACAGATGATAATA TGGCCACAACCATGGAACAAGAGA	
--	--	---	--

		CTTGCGCGCACTCTCTCACTTTTGAG GAATGCCCAAATGCTCTGCTCTAC AATACCGTAATGGATTTTACCTGCT AAAGTATGATGAAGAATGGTACCCA GAGGAGTTATTGACTGATGGAGAGG ATGATGTCTTTGATCCCGAATTAGA CATGGAAGTCGTTTTTCGAGTTACAG GGAAGCGGAGCTACTAACTTCAGCC TGCTGAAGCAGGCTGGAGATGTGGA GGAGAACCCTGGACCTATGGCAGAT TCCAACGGTACTATTACCGTTGAAG AGCTTAAAAAGCTCCTTGAACAATG GAACCTAGTAATAGGTTTCCTATTCT CTTACATGGATTTGTCTTCTACAATT TGCCTATGCCAACAGGAATAGGTTT TTGTATATAATTAAGTTAATTTTCCT CTGGCTGTTATGGCCAGTAACTTTA GCTTGTTTTGTGCTTGCTGCTGTTTA CAGAATAAATTGGATCACCGGTGGA ATTGCTATCGCAATGGCTTGTCTTGT AGGCTTGATGTGGCTCAGCTACTTC ATTGCTTCTTTCAGACTGTTTGCGCG TACGCGTTCCATGTGGTCATTCAAT CCAGAACTAACATTCTTCTCAACG TGCCACTCCATGGCACTATTCTGAC CAGACCGCTTCTAGAAAGTGAATC GTAATCGGAGCTGTGATCCTTCGTG GACATCTTCGTATTGCTGGACACCA TCTAGGACGCTGTGACATCAAGGAC CTGCCTAAAGAAATCACTGTTGCTA CATCACGAACGCTTTCTTATTACAA ATTGGGAGCTTCGCAGCGTGTAGCA GGTGACTCAGGTTTTGCTGCATACA GTCGCTACAGGATTGGCAACTATAA ATTAAACACAGACCATTCAGTAGC AGTGACAATATTGCTTTGCTTGTAC AGGGAAGCGGAGCTACTAACTTCAG CCTGCTGAAGCAGGCTGGAGATGTG GAGGAGAACCCTGGACCTATGTA CTATTTCGTTTCGGAAGAGACAGGTAC GTTAATAGTTAATAGCGTACTTCTTT TTCTTGCTTTCGTGGTATTCTTGCTA GTTACACTAGCCATCCTTACTGCGC TTCGATTGTGTGCGTACTGCTGCAA TATTGTAAACGTGAGTCTTGTA AAAACCTTCTTTTACGTTTACTCTCGTGT TAAAAATCTGAATTCTTCTAGAGTT CCTGATCTTCTGGTCTAA	
--	--	--	--

i) Expression cassettes

[0078] The vectors disclosed herein may comprise one or more expression cassettes. The phrase “expression cassette” as used herein refers to a defined segment of a nucleic acid molecule that comprises the minimum elements needed for production of another nucleic acid or protein encoded by that nucleic acid molecule. In some embodiments, the expression cassette comprises a promoter. In some embodiments, the promoter is operatively linked to each of the polynucleotide sequences of the expression cassette.

[0079] In some embodiments, a vector may comprise an expression cassette, the expression cassette comprising a first polynucleotide encoding an antigen, and a second polynucleotide encoding an enhancer protein. In some embodiments, the expression cassette comprises a first promoter, operatively linked to the first polynucleotide; and a second promoter, operatively linked to the second polynucleotide. In some embodiments, the expression cassette comprises a shared promoter operatively linked to both the first polynucleotide and the second polynucleotide.

[0080] In some embodiments, the expression cassette comprises a coding polynucleotide comprising the first polynucleotide and the second polynucleotide linked by a polynucleotide encoding a separating element (*e.g.*, a ribosome skipping site or 2A element), the coding polynucleotide operatively linked to the shared promoter.

[0081] In some embodiments, the expression cassette comprises a coding polynucleotide, the coding polynucleotide encoding an enhancer protein and an antigen linked to by a separating element (*e.g.*, a ribosome skipping site or 2A element), the coding polynucleotide operatively linked to the shared promoter.

[0082] In some embodiments, the expression cassette is configured for transcription of a single messenger RNA encoding both the antigen and the enhancer protein, linked by a separating element (*e.g.*, a ribosome skipping site or 2A element); wherein translation of the messenger RNA results in expression of the antigen and the enhancer protein (*e.g.*, the L protein) as distinct polypeptides. In some embodiments, the expression cassettes disclosed herein comprise one or more proteolytic cleavage sites, for example, 1, 2, 3, 4, or 5 proteolytic cleavage sites. In some embodiments, the proteolytic cleavage site is located between a polynucleotide encoding a first antigen, and another polynucleotide encoding a second antigen. In some embodiments, the proteolytic cleavage site is located between a polynucleotide encoding an antigen, and a polynucleotide encoding an enhancer protein. In some

embodiments, the proteolytic cleavage site comprises the nucleic acid sequence of SEQ ID NO: 50. In some embodiments, the proteolytic cleavage site is a furin cleavage site.

[0083] In some embodiments, the expression cassettes disclosed herein comprise a nucleic acid sequence encoding a viral accessory protein, for example ORF3a protein. In some embodiments, the polynucleotide encoding the ORF3 protein has a nucleic acid sequence with at least 70% sequence identity - for instance, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between – to the nucleic acid sequence of SEQ ID NO: 54. In some embodiments, the polynucleotide encoding the ORF3 protein has a nucleic acid sequence of SEQ ID NO: 54. In some embodiments, ORF3 protein has a amino acid sequence with at least 70% sequence identity - for instance, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between – to the amino acid sequence of SEQ ID NO: 53. In some embodiments, the ORF3 protein has an amino acid sequence of SEQ ID NO: 53.

[0084] In some embodiments, the vector is selected from the group consisting of CoVEG3-17. In some embodiments, the vector comprises a nucleic acid sequence having at least about 70% identity, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, at least about 99.5% or about 100% identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO: 35-49. In some embodiments, the vector comprises a nucleic acid sequence having at least 70% (for example, about 75%, about 80%, about 85%, about 90%, about 95%, about 96%, about 97%, about 98%, about 99%, about 99.5%, or about 100%) identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO: 35-49.

[0085] The nucleic acid sequence of the expression cassette of CoVEG3-17 and the genetic elements therein are listed in Table 2.

Table 2

Name of sequence	SEQ ID NO:	Sequence
CoVEG3 Expression Cassette	35	ATGGCCGACAGCAACGGCACAATCACCGTGGAAGA GCTGAAGAACTGCTGGAACAGTGGAACCTGGTCA TCGGCTTCCTGTTTCTGACCTGGATCTGTCTGCTGCA GTTTCGCTTATGCCAATCGGAACAGATTCTGTACAT CATCAAGCTGATCTTCCTGTGGCTGCTGTGGCCTGT

		GACCCTGGCTTGCTTCGTGCTGGCCGCTGTGTACCG GATCAACTGGATCACAGGCGGAATCGCCATCGCCA TGGCCTGCCTGGTGGGCCTGATGTGGCTGAGCTACT TCATCGCTTCTTTCAGACTGTTGCCAGAACCCGGA GCATGTGGTCCTTCAACCCCGAGACAAACATCCTGC TGAACGTGCCTCTGCACGGCACCATCCTGACAAGAC CTCTGCTCGAGAGCGAGCTGGTGATTGGCGCAGTGA TTCTGAGAGGCCATCTGAGGATCGCCGACACCACC TGGGCAGATGCGACATCAAGGACCTTCCAAAGGAA ATCACCGTTGCCACCAGCCGGACCCTGTCCTACTAC AAACTGGGCGCCAGCCAAAGAGTGGCCGGCGATAG CGGCTTTGCCGCCTACAGCAGATACCGCATCGGAAA TTACAAGCTCAACACCGACCACAGCAGCTCTTCTGA TAACATCGCCCTGCTGGTGCAGtaatccccccccctaacttact ggccgaagccgcttgaataaggccggtgtgcgttgtctatatgtattttccaccatat tgccgtcttttgcaatgtgagggcccgaaacctggccctgtcttctgacgagcatt cctaggggtctttcccctctcgccaaaggaatgcaaggtctgttgatgtcgtgaagga agcagttcctctggaagcttctgaagacaacaacgtctgtagcgacccttgcaggc agcggaaacccccacctggcgacaggtgcctctcgggccaaaagccacgtgtataa gatacacctgcaaaggcggcacaacccagtgccacgttgtgagttggatagttgtg gaaagagtcaaatggctctcctcaagcgtattcaacaaggggctgaaggatgccag aaggtagcccatgtatgggatctgatctggggcctcggtgcacatgctttacatgtgtt tagtcgaggttaaaaaacgtctaggccccccgaaccacggggacgtggtttcctt gaaaaacacgatgataatatggccacaacctgaacaagagactgcgcgactct ctcacttttgaggaatgccccaaatgctctgctctacaataacctgaattttacctgc taaagtatgatgaagaatggtaccagaggagttattgactgatggagaggatgatgt ctttgatccccgaattagacatggaagtcgttttcgagttacagggaagcggagctacta acttcagcctgctgaagcaggtggagatgtggaggagaacctggacctATGT TCGTGTTCTTGGTGTGCTGCTGCCCTCTGGTCAGCTCCCA GTGTGTGAACCTGACCACCAGAACCCAGCTGCCACC TGCTTATACAACTCCTTCACTCGGGGGGTATACTA CCCCGACAAGGTGTTTACAGATCTAGCGTGCTGCATTC TACACAAGACCTGTTTCTGCCCCTTCTTACAGCAACGT GACCTGGTTCCACGCCATCCACGTGTCTGGAACCAA CGGAACCAAGAGATTTCGACAACCCCGTGCTGCCTTT CAACGACGGCGTGTACTTCGCCAGCACCCGAGAAGT CCAACATCATCAGAGGATGGATTTTCGGCACACAC TGGACAGCAAAACCCAGAGCCTGCTGATCGTGAAC AACGCCACCAACGTGGTGATCAAGGTGTGCGAGTT CCAGTTCTGCAATGATCCCTTCCTGGGCGTGTACTA CCACAAGAACAACAAGTCTTGATGGAAAGCGAGT TCAGAGTGTATTCCAGCGCCAACAATTGCACCTTCG AGTACGTGAGCCAACCCCTTTCTGATGGACCTTGAAG GCAAGCAGGGCAACTTCAAAAATCTGCGAGAATTT GTGTTCAAGAACATCGACGGATACTTCAAGATCTAC TCTAAGCACACGCCAATCAACCTGGTGAGAGATCTG CCCCAGGGCTTTAGCGCTTTGGAACCTCTGGTGGAC CTGCCTATCGGAATCAACATCACCAGATTTCAAACCT CTCCTGGCCCTGCACAGATCTTATCTGACCCCTGGG GACAGTAGTAGCGGCTGGACAGCCGGCGCCGCCGC
--	--	---

		CTACTACGTGGGATACCTGCAGCCTAGAACATTCCCT GCTGAAGTACAATGAGAACGGAACAATCACAGACG CCGTGGACTGCGCCCTGGATCCTTTGAGCGAGACAA AGTGCACCCTGAAGTCGTTACCGTCGAAAAAGGC ATCTACCAGACCAGCAACTTCCGCGTGCAGCCTACG GAATCTATCGTGCGGTTCCCCAACATCACCAACCTG TGCCCTTTTCGGCGAGGTGTTTAACGCTACAAGGTT GCCAGCGTGTATGCCTGGAACAGAAAGAGAATCAG CAATTGCGTGGCCGATTATAGCGTTCTGTACAACAG CGCTTCCTTCAGCACCTTCAAGTGCTACGGCGTGCT TCCAACCAAGCTGAACGACCTCTGCTTCACCAATGT CTACGCTGACTCTTTCGTGATTAGAGGGCGATGAGGT TAGACAGATCGCACCTGGCCAGACCGGCAAAATCG CTGACTACAACCTACAAGCTGCCTGATGACTTCACAG GCTGTGTCAATTGCCTGGAACCTCAAATAACCTGGACT CTAAAGTGGGCGGCAACTACAACCTACCTGTACCGG CTGTTCCGGAAGAGCAATCTGAAACCTTTTGAGCGG GACATCTCTACAGAGATCTACCAGGCCGGCAGCAC ACCCTGCAACGGCGTTGAGGGCTTCAACTGCTACTT CCCTCTGCAGAGCTACGGCTTTCAGCCAACAAATGG AGTGGGCTACCAGCCGTACAGAGTGGTGGTGCTGA GCTTCGAACTGCTGCATGCCCCAGCCACAGTGTGTG GACCTAAGAAGTCTACCAACCTGGTGAAGAACAAG TGCGTGAACCTTTAACTTTAACGGCCTGACCGGCACA GGCGTGCTGACCGAATCCAACAAAAAGTTCTGCCC TTCCAACAGTTCGGCAGAGACATCGCCGATACAACC GATGCCGTGCGGGACCCCCAGACCTTAGAAATCCTA GATATCACCCCGTGCAGCTTCGGCGGAGTCTCTGTT ATTACTCCTGGCACCAACACCAGCAACCAAGTGGCT GTTCTGTACCAAggGTGAACCTGCACCGAAGTGCCTG TGGCTATCCACGCCGATCAGCTGACCCCAACCTGGC GGGTGTATAGCACCGGCTCTAACGTGTTCCAGACCC GGGCTGGCTGCCTGATCGGCGCCGAACACGTCAAC AACTCCTATGAATGTGACATCCCCATCGGGGCTGGC ATCTGCGCCAGTTACCAGACACAGACAAATAGCCCT AGACGGGCCAGAAGCGTGGCCTCCCAGAGTATCAT TGCCCTACACCATGAGCCTGGGCGCCGAGAACAGCG TGGCCTATTCTAACAATAGCATCGCAATCCCTACCA ACTTTACCATCTCTGTGACAACCGAGATCCTGCCTG TGAGCATGACCAAAACCAGCGTGGACTGCACGATG TACATCTGTGGCGACAGCACAGAATGCAGTAATCTG TTGCTGCAGTACGGCAGCTTTTGCACCCAGTTGAAT AGAGCCCTGACCGGAATCGCCGTAGAGCAGGACAA AAATACCCAGGAGGTGTTTCGCCCAGGTGAAACAGA TCTACAAGACACCTCCCATTAAGGACTTCGGAGGTT TTAACTTCAGCCAGATCCTGCCCCGACCTTCCAAGC CTAGCAAACGCTCCTTCATCGAGGACCTGCTCTTCA ACAAGGTGACACTGGCTGATGCCGGCTTCATCAAGC AGTACGGAGATTGTCTGGGAGACATCGCCGCTAGA GATCTGATCTGCGCCCAAAAGTTCAACGGCCTGACC
--	--	---

		GTGCTGCCTCCTCTGCTTACAGACGAGATGATCGCC CAGTACACCAGCGCCCTGCTGGCTGGCACCATCACA AGCGGCTGGACCTTCGGAGCCGGAGCCGCTCTGCA AATCCCCTTTGCCATGCAGATGGCCTACCGGTTCAA CGGCATCGGCGTGACACAGAATGTGCTGTACGAGA ACCAGAAGCTGATCGCTAACCCAGTTTAACAGCGCTA TCGGCAAGATCCAGGACTCGCTGAGTAGCACCGCCT CTGCCCTGGGCAAGCTGCAGGACGTCGTGAACCAG AACGCCCAAGCCCTGAACACACTGGTGAACAGCT GAGCAGCAACTTCGGCGCCATCAGCTCTGTGCTGAA CGATATCCTGAGCAGACTGGACAAGGTGGAAGCCG AGGTCCAGATCGACAGACTGATCACAGGAAGACTG CAGAGCCTGCAAACGTACGTGACACAGCAGCTGAT CCGGGCAGCCGAAATCCGGGGCCAGCGCCAATCTGG CCGCTACCAAGATGAGCGAGTGCGTGTTAGGCCAG AGCAAGCGGGTGGATTTCTGCGGTAAGGGATACCA CCTGATGAGCTTTCCCCAGAGCGCTCCTCACGGCGT GGTGTTTCTGCACGTGACCTACGTTCTGCCAGGA AAAGAACTTCACCACCGCCCTGCTATCTGCCACGA TGGCAAGGCCCACTTCCCTAGAGAGGGCGTTTCGT GTCTAACGGCACACACTGGTTTGTGACCCAGAGAA ACTTCTACGAGCCTCAGATCATCACCACAGACAACA CCTTTGTGAGCGGCAATTGCGACGTGGTGATCGGAA TTGTTAATAATAACCGTGTACGACCTCTGCAGCCTG AGCTCGACAGCTTCAAGGAAGAGCTGGACAAGTAC TTCAAGAACCACACCTCCCCAGATGTGGACCTGGGC GATATTTAGGCATCAACGCCTCCGTCGTGAATATC CAGAAGGAGATCGACCGGCTCAACGAGGTGGCCAA GAACCTTAACGAGAGCCTGATCGACCTGCAGGAAC TGGGCAAATATGAGCAGTACATCAAGTGGCCTTGGT ACATCTGGCTGGGCTTTATCGCAGGCCTGATCGCTA TCGTGATGGTGACCATTATGCTGTGTTGTATGACCA GCTGTTGTAGTTGTCTGAAGGGCTGCTGTTCTTGCG GCAGCTGCTGCAAGTTTCGACGAAGACGACTCAGAG CCCGTGCTGAAAGGCGTGAAGCTGCACTACACCtaac cccccccctaacgttactggccgaagccgcttgaataaggccggtgtgcgtttgtc tatatgttatttccaccatattgccgtcttttggcaatgtgagggcccgaaacctggcc ctgtcttcttgacgagcattcctaggggtctttccctctcgccaaaggaatgcaaggtc tgtgaatgtcgtgaaggaagcagttcctctggaagcttctgaagacaacaacgtct gttagcgacctttgcaggcagcggaaacccccacctggcgacaggtgcctctgcgg ccaaaagccacgtgtataagatacacctgcaaaggcggcacaaccccagtgccacg ttgtgagttggatagttgtgaaagagtcaaatggctctcctcaagcgtattcaacaag gggctgaaggatgccagaaggtagccattgtatgggatctgatctggggcctcgg tgcacatgctttacatgtgttttagtcgaggttaaaaaacgttaggcccccgaaaca cggggacgtggtttccttgaaaaacacgatgataatatggccacaaccttggaaca agagacttgcgcgcactctctcacttttgaggaaatgcccataatgctctgcttacaata ccgtaatggattttacgtgctaaagtatgatgaagaatggtagccagaggagtattgac tgatggagaggatgatgtctttgatccgaattagacatggaagtcgttttcgagttaca gggaagcggagctactaactcagcctgctgaagcaggctggagatgtggaggaga acctggacctATGAGCGACAACGGCCCTCAAACCCAGA
--	--	---

		GAAATGCCCTCGGATCACATTTGGCGGACCTAGCG ACAGCACCGGCAGCAACCAGAATGGAGAAAGAAGC GGCGCCAGATCCAAGCAGCGGAGACCTCAGGGACT GCCCAACAACACCGCTAGCTGGTTCACCGCCCTGAC CCAACACGGCAAGGAAGATCTGAAGTTCCCCAGAG GCCAGGGCGTGCCTATCAACACAACTCTTCTCCCG ACGACCAGATCGGATACTATAGACGGGCCACTCGG AGAATTCGGGGCGGGCGACGGAAAAATGAAGGACCT TTCTCCAAGATGGTACTTCTACTACCTCGGCACAGG CCCTGAGGCCGGCCTGCCTTACGGCGCCAACAAGG ATGGCATCATCTGGGTGCGCACCGAGGGCGCCCTGA ACACCCCTAAGGACCACATCGGCACAAGAAACCCC GCTAACAACGCCGCAATCGTGCTGCAGCTGCCTCAG GGCACCACCCTGCCCAAGGGCTTCTACGCCGAGGG CTCTAGAGGTGGCTCCCAGGCTTCTAGCCGCTCCTC CAGCCGCAGCAGAAACAGCAGCAGGAACAGCACCC CCGGCAGCTCCCGGGGCACCAGCCCCGCCAGAATG GCCGAAATGGCGGCGATGCCGCCCTGGCCCTGCTC CTGCTGGACAGACTGAATCAGCTGGAAAGCAAGAT GAGCGGCAAAGGACAGCAGCAGCAAGGCCAGACC GTGACCAAGAAAAGCGCTGCTGAAGCCTCCAAGAA ACCTAGACAAAAGCGGACCGCCACAAAGGCCTACA ACGTGACCCAAGCCTTTGGAAGAAGAGGCCCCGAG CAGACACAGGGCAATTCGGCGACCAGGAGCTGAT CCGGCAGGGAACCGACTACAAGCACTGGCCTCAGA TCGCCCAGTTCGCCCCTAGCGCCAGCGCCTTCTTCG GCATGAGCAGAATCGGCATGGAAGTGACCCCTTCT GGCACCTGGCTGACCTACACCGGCGCTATCAAGCTG GACGATAAGGATCCTAACTTCAAGGACCAAGTGAT CCTGCTGAACAAGCATATCGACGCCTATAAGACCTT TCCACCTACAGAGCCTAAGAAAGATAAGAAGAAGA AAGCCGACGAGACACAGGCCCTGCCTCAGAGACAG AAAAAGCAGCAGACAGTGACACTGCTGCCAGCCGC TGACCTGGATGACTTCAGCAAGCAGCTGCAGCAGA GCATGTCTTCTGCTGATAGCACCCAGGCCggaagcggag ctactaacttcagcctgctgaagcaggctggagatgtggaggagaaccctggacct ATGTATTCTTTTGTGTCCGAGGAAACCGGCACACTG ATCGTTAATAGCGTGCTGCTCTTCTGGCCTTCGTG GTGTTCTGCTGGTGACCCTGGCTATCCTGACCGCC CTGAGACTGTGTGCCTACTGCTGCAACATCGTGAAC GTGTCTCTGGTCAAGCCTAGCTTCTACGTGTACAGC CGGGTGAAGAACCTGAACAGCAGCAGAGTGCCCGA CCTGCTGGTGTGA
SARS CoV2 M gene	66	ATGGCCGACAGCAACGGCACAATCACCGTGGAAGA GCTGAAGAACTGCTGGAACAGTGGAACCTGGTCA TCGGCTTCCTGTTTCTGACCTGGATCTGTCTGCTGCA GTTTCGCTTATGCCAATCGGAACAGATTCTGTACAT CATCAAGCTGATCTTCTGTGGCTGCTGTGGCCTGT GACCCTGGCTTGTCTCGTGCTGGCCGCTGTGTACCG GATCAACTGGATCACAGGCGGAATCGCCATCGCCA

		TGGCCTGCCTGGTGGGCTGATGTGGCTGAGCTACT TCATCGCTTCTTTCAGACTGTTCCGACAGAACCCGGA GCATGTGGTCCTTCAACCCCGAGACAAACATCCTGC TGAACGTGCCTCTGCACGGCACCATCCTGACAAGAC CTCTGCTCGAGAGCGAGCTGGTGATTGGCGCAGTGA TTCTGAGAGGCCATCTGAGGATCGCCGACACCACC TGGGCAGATGCGACATCAAGGACCTTCCAAAGGAA ATCACCGTTGCCACCAGCCGGACCCTGTCCTACTAC AAACTGGGCGCCAGCCAAAGAGTGGCCGGCGATAG CGGCTTTGCCGCCTACAGCAGATACCGCATCGGAAA TTACAAGCTCAACACCGACCACAGCAGCTCTTCTGA TAACATCGCCCTGCTGGTGCAG ^{taa}
IRES encoding sequence	67	tcccccccctaacgttactggccgaagccgcttgaataaggccggtgtgcgttgg tctatatgttatttccaccatattgccgtcttttggcaatgtgagggcccgaaacctgg ccctgtcttcttgacgagcattcctaggggtcttccctctcgccaaaggaaatgcaag gtctgttgaatgtcgtgaaggaagcagttcctctggaagcttcttgaagacaacaacg tctgtagcgacctttgcaggcagcggaaacccccacctggcgacaggtgcctctgc ggccaaaagccacgtgtataagatacacctgcaaaggcggcacaacccagtgcca cgttgtgagttggatagttgtgaaaagagcaaatggctctcctcaagcgtattcaaca aggggctgaaggatgccagaaggtacccattgtatgggatctgatctggggcctc ggtgcacatgctttacatgtgtttagtcgaggttaaaaaaacgttagggcccccgaaac cacggggacgtggttttcttgaaaaacacgatgataat
L peptide from ECMV encoding sequence	68	atggccacaacctggaacaagagacttgcgcgcaactctctacttttgaggaatgcc caaaatgctctgctctacaataaccgtaatggattttacctgctaaagatgatgaagaat ggtaccagaggagttattgactgatggagaggatgatgtctttgatccgaattagac atggaagtcgttttcgagttacag
P2A skipping site encoding sequence	69	ggaaagcggagctactaacttcagcctgctgaagcaggctggagatgtggaggagaa ccctggacct
SARS CoV2 Spike gene	70	ATGTTTCGTGTTCTTGGTGCTGCTGCCTCTGGTCAGCT CCCAGTGTGTGAACCTGACCACCAGAACCCAGCTGC CACCTGCTTATACAACTCCTTCACTCGGGGGGTAT ACTACCCCGACAAGGTGTTTCAGATCTAGCGTGCTGC ATTCTACACAAGACCTGTTCTGCCCCTTCTTCAGCA ACGTGACCTGGTTCCACGCCATCCACGTGTCTGGAA CCAACGGAACCAAGAGATTCGACAACCCCGTGCTG CCTTTCAACGACGGCGTGTAATTCGCCAGCACCCGAG AAGTCCAACATCATCAGAGGATGGATTTTCGGCACCC ACACTGGACAGCAAAACCCAGAGCCTGCTGATCGT GAACAACGCCACCAACGTGGTGATCAAGGTGTGCG AGTTCCAGTTCTGCAATGATCCCTTCTGGGCGTGT ACTACCACAAGAACAACAAGTCTTGGATGGAAAGC GAGTTCAGAGTGTATTCCAGCGCCAACAATTGCACC TTCGAGTACGTGAGCCAACCCTTCTGATGGACCTT GAAGGCAAGCAGGGCAACTTCAAAAATCTGCGAGA

		ATTTGTGTTCAAGAACATCGACGGATACTTCAAGAT CTACTCTAAGCACACGCCAATCAACCTGGTGAGAG ATCTGCCCCAGGGCTTTAGCGCTTTGGAACCTCTGG TGGACCTGCCTATCGGAATCAACATCACCAGATTTC AAACTCTCCTGGCCCTGCACAGATCTTATCTGACCC CTGGGGACAGTAGTAGCGGCTGGACAGCCGGCGCC GCCGCTACTACGTGGGATACCTGCAGCCTAGAACA TTCCTGCTGAAGTACAATGAGAACGGAACAATCAC AGACGCCGTGGACTGCGCCCTGGATCCTTTGAGCGA GACAAAGTGCACCCTGAAGTCGTTACCCGTCGAAA AAGGCATCTACCAGACCAGCAACTTCCGCGTGACG CCTACGGAATCTATCGTGCGGTTCCCCAACATCACC AACCTGTGCCCTTTTCGGCGAGGTGTTTAAACGCTACA AGGTTTCGCCAGCGTGTATGCCTGGAACAGAAAAGAG AATCAGCAATTGCGTGCGCGATTATAGCGTTCTGTA CAACAGCGCTTCCTTCAGCACCTTCAAGTGCTACGG CGTGTCTCCAACCAAGCTGAACGACCTCTGCTTCAC CAATGTCTACGCTGACTCTTTCGTGATTAGAGGCGA TGAGGTTAGACAGATCGCACCTGGCCAGACCGGCA AAATCGCTGACTACAACCTACAAGCTGCCTGATGACT TCACAGGCTGTGTCATTGCCTGGAACCTCAAATAACC TGGACTCTAAAGTGGGCGGCAACTACAACCTACCTGT ACCGGCTGTTCCGGAAGAGCAATCTGAAACCTTTTG AGCGGGACATCTCTACAGAGATCTACCAGGCCGGC AGCACACCCTGCAACGGCGTTGAGGGCTTCAACTGC TACTTCCCTCTGCAGAGCTACGGCTTTCAGCCAACA AATGGAGTGGGCTACCAGCCGTACAGAGTGGTGGT GCTGAGCTTCGAACTGCTGCATGCCCCAGCCACAGT GTGTGGACCTAAGAAGTCTACCAACCTGGTGAAGA ACAAGTGCGTGAACCTTTAACTTTAACGGCCTGACCG GCACAGGCGTGCTGACCGAATCCAACAAAAAGTTC CTGCCCTTCCAACAGTTCGGCAGAGACATCGCCGAT ACAACCGATGCCGTGCGGGACCCCCAGACCTTAGA AATCCTAGATATCACCCCGTGACGTTTCGGCGGAGT CTCTGTTATTACTCCTGGCACCAACACCAGCAACCA AGTGGCTGTTCTGTACCAAggcGTGAACCTGCACCGAA GTGCCTGTGGCTATCCACGCCGATCAGCTGACCCCA ACCTGGCGGGTGTATAGCACCGGCTCTAACGTGTTT CAGACCCGGGCTGGCTGCCTGATCGGCGCCGAACA CGTCAACAACTCCTATGAATGTGACATCCCCATCGG GGCTGGCATCTGCGCCAGTTACCAGACACAGACAA ATAGCCCTAGACGGGCCAGAAGCGTGGCCTCCCGAG AGTATCATTGCCTACACCATGAGCCTGGGCGCCGAG AACAGCGTGGCCTATTCTAACAATAGCATCGCAATC CCTACCAACTTTACCATCTCTGTGACAACCGAGATC CTGCCTGTGAGCATGACCAAAAACCAGCGTGGACTG CACGATGTACATCTGTGGCGACAGCACAGAATGCA GTAATCTGTTGCTGCAGTACGGCAGCTTTTGCACCC AGTTGAATAGAGCCCTGACCGGAATCGCCGTAGAG CAGGACAAAAATACCCAGGAGGTGTTTCGCCAGGT
--	--	--

		GAAACAGATCTACAAGACACCTCCCATTAAGGACTT CGGAGGTTTTAACTTCAGCCAGATCCTGCCCCGACCC TTCCAAGCCTAGCAAACGCTCCTTCATCGAGGACCT GCTCTTCAACAAGGTGACACTGGCTGATGCCGGCTT CATCAAGCAGTACGGAGATTGTCTGGGAGACATCG CCGCTAGAGATCTGATCTGCGCCCAAAAGTTCAACG GCCTGACCGTGCTGCCTCCTCTGCTTACAGACGAGA TGATCGCCCAGTACACCAGCGCCCTGCTGGCTGGCA CCATCACAAGCGGCTGGACCTTCGGAGCCGGAGCC GCTCTGCAAATCCCCTTTGCCATGCAGATGGCCTAC CGGTTCAACGGCATCGGCGTGACACAGAATGTGCT GTACGAGAACCAGAAGCTGATCGCTAACCAGTTTA ACAGCGCTATCGGCAAGATCCAGGACTCGCTGAGT AGCACCGCCTCTGCCCTGGGCAAGCTGCAGGACGTC GTGAACCAGAACGCCCAAGCCCTGAACACACTGGT GAAACAGCTGAGCAGCAACTTCGGCGCCATCAGCT CTGTGCTGAACGATATCCTGAGCAGACTGGACAAG GTGGAAGCCGAGGTCCAGATCGACAGACTGATCAC AGGAAGACTGCAGAGCCTGCAAACGTACGTGACAC AGCAGCTGATCCGGGCAGCCGAAATCCGGGCCAGC GCCAATCTGGCCGCTACCAAGATGAGCGAGTGCGT GTTAGGCCAGAGCAAGCGGGTGGATTTCTGCGGTA AGGGATACCACCTGATGAGCTTCCCCAGAGCGCTC CTCACGGCGTGGTGTCTTGACGTGACCTACGTTT CTGCCCAGGAAAAGAACTTCACCACCGCCCTGCTA TCTGCCACGATGGCAAGGCCCACTTCCCTAGAGAGG GCGTTTTCTGTGCTAACGGCACACACTGGTTTGTGA CCCAGAGAACTTCTACGAGCCTCAGATCATACCA CAGACAACACCTTTGTGAGCGGCAATTGCGACGTG GTGATCGGAATTGTTAATAATACCGTGTACGACCCT CTGCAGCCTGAGCTCGACAGCTTCAAGGAAGAGCT GGACAAGTACTTCAAGAACCACACCTCCCCAGATGT GGACCTGGGCGATATTTTCAGGCATCAACGCCTCCGT CGTGAATATCCAGAAGGAGATCGACCGGCTCAACG AGGTGGCCAAGAACCTTAACGAGAGCCTGATCGAC CTGCAGGAACCTGGGCAAATATGAGCAGTACATCAA GTGGCCTTGGTACATCTGGCTGGGCTTTATCGCAGG CCTGATCGCTATCGTGATGGTGACCATTATGCTGTG TTGTATGACCAGCTGTTGTAGTTGTCTGAAGGGCTG CTGTTCTTGCGGCAGCTGCTGCAAGTTTCGACGAAGA CGACTCAGAGCCCCTGCTGAAAGGCGTGAAGCTGC ACTACACC _{taa}
SARS CoV2 N gene	71	ATGAGCGACAACGGCCCTCAAAACCAGAGAAATGC CCCTCGGATCACATTTGGCGGACCTAGCGACAGCAC CGGCAGCAACCAGAATGGAGAAAGAAGCGGCGCCA GATCCAAGCAGCGGAGACCTCAGGGACTGCCCAAC AACACCGCTAGCTGGTTTACCGCCCTGACCAACAC GGCAAGGAAGATCTGAAGTTCCCCAGAGGCCAGGG CGTGCCTATCAACACAACTCTTCTCCCGACGACCA GATCGGATACTATAGACGGGCCACTCGGAGAATTC

		GGGGCGGCGACGGAAAAATGAAGGACCTTTCTCCA AGATGGTACTTCTACTACCTCGGCACAGGCCCTGAG GCCGGCCTGCCTTACGGCGCCAACAAGGATGGCAT CATCTGGGTCGCCACCGAGGGCGCCCTGAACACCCC TAAGGACCACATCGGCACAAGAAACCCCGCTAACA ACGCCGCAATCGTGCTGCAGCTGCCTCAGGGCACCA CCCTGCCCCAAGGGCTTCTACGCCGAGGGCTCTAGAG GTGGCTCCCAGGCTTCTAGCCGCTCCTCCAGCCGCA GCAGAAACAGCAGCAGGAACAGCACCCCGGCAGC TCCCGGGGCACCAGCCCCGCCAGAATGGCCGGAAA TGGCGGCGATGCCGCCCTGGCCCTGCTCCTGCTGGA CAGACTGAATCAGCTGGAAAGCAAGATGAGCGGCA AAGGACAGCAGCAGCAAGGCCAGACCGTGACCAAG AAAAGCGCTGCTGAAGCCTCCAAGAAACCTAGACA AAAGCGGACCGCCACAAAGGCCTACAACGTGACCC AAGCCTTTGGAAGAAGAGGCCCGAGCAGACACAG GGCAATTTGCGCGACCAGGAGCTGATCCGGCAGGG AACCGACTACAAGCACTGGCCTCAGATCGCCAGTT CGCCCCTAGCGCCAGCGCCTTCTTCGGCATGAGCAG AATCGGCATGGAAGTGACCCCTTCTGGCACCTGGCT GACCTACACCGGCGCTATCAAGCTGGACGATAAGG ATCCTAACTTCAAGGACCAAGTGATCCTGCTGAACA AGCATATCGACGCCTATAAGACCTTTCCACCTACAG AGCCTAAGAAAGATAAAGAAGAAGAAAGCCGACGA GACACAGGCCCTGCCTCAGAGACAGAAAAAGCAGC AGACAGTGACACTGCTGCCAGCCGCTGACCTGGAT GACTTCAGCAAGCAGCTGCAGCAGAGCATGTCTTCT GCTGATAGCACCCAGGCC
SARS CoV2 Envelope gene	72	ATGTATTCTTTTGTGTCCGAGGAAACCGGCACACTG ATCGTTAATAGCGTGCTGCTCTTCCTGGCCTTCGTG GTGTTCCTGCTGGTGACCCTGGCTATCCTGACCGCC CTGAGACTGTGTGCCTACTGCTGCAACATCGTGAAC GTGTCTCTGGTCAAGCCTAGCTTCTACGTGTACAGC CGGGTGAAGAACCTGAACAGCAGCAGAGTGCCCGA CCTGCTGGTGTGA
CoVEG4 expression cassette	36	ATGGCCGACAGCAACGGCACAATCACCGTGGAAGA GCTGAAGAACTGCTGGAACAGTGGAACCTGGTCA TCGGCTTCCTGTTTCTGACCTGGATCTGTCTGCTGCA GTTTCGCTTATGCCAATCGGAACAGATTCTGTACAT CATCAAGCTGATCTTCCTGTGGCTGCTGTGGCCTGT GACCCTGGCTTGCTTCGTGCTGGCCGCTGTGTACCG GATCAACTGGATCACAGGCGGAATCGCCATCGCCA TGGCCTGCCTGGTGGGCCTGATGTGGCTGAGCTACT TCATCGCTTCTTTCAGACTGTTCGCCAGAACCCGGA GCATGTGGTCCTTCAACCCCGAGACAAACATCCTGC TGAACGTGCCTCTGCACGGCACCATCCTGACAAGAC CTCTGCTCGAGAGCGAGCTGGTGATTGGCGCAGTGA TTCTGAGAGGCCATCTGAGGATCGCCGGACACCACC TGGGCAGATGCGACATCAAGGACCTTCCAAGGAA

		ATCACCGTTGCCACCAGCCGGACCCTGTCCTACTAC AAACTGGGCGCCAGCCAAAGAGTGGCCGGCGATAG CGGCTTTGCCGCCTACAGCAGATACCGCATCGGAAA TTACAAGCTCAACACCGACCACAGCAGCTCTTCTGA TAACATCGCCCTGCTGGTGCAGCGAAAACGGCGCgg aagcggaggaagcggagctactaactcagcctgctgaagcaggctggagatgtgg aggagaaccctggacctATGTTTCGTGTTTCCTGGTGCTGCTGC CTCTGGTCAGCTCCCAGTGTGTGAACCTGACCACCA GAACCCAGCTGCCACCTGCTTATACAAACTCCTTCA CTCGGGGGGTATACTACCCCGACAAGGTGTTTCAGAT CTAGCGTGCTGCATTCTACACAAGACCTGTTCCCTGC CCTTCTTCAGCAACGTGACCTGGTTCACGCCATCC ACGTGTCTGGAACCAACGGAACCAAGAGATTTCGAC AACCCCGTGCTGCCTTTCAACGACGGCGTGTACTTC GCCAGCACCGAGAAGTCCAACATCATCAGAGGATG GATTTTCGGCACCACTGGACAGCAAAACCCAGA GCCTGCTGATCGTGAACAACGCCACCAACGTGGTG ATCAAGGTGTGCGAGTTCCAGTTCTGCAATGATCCC TTCCTGGGCGTGTACTACCACAAGAACAACAAGTCT TGGATGGAAAGCGAGTTCAGAGTGTATTCCAGCGC CAACAATTGCACCTTCGAGTACGTGAGCCAACCCTT TCTGATGGACCTTGAAGGCAAGCAGGGCAACTTCA AAAATCTGCGAGAATTTGTGTTCAAGAACATCGACG GATACTTCAAGATCTACTCTAAGCACACGCCAATCA ACCTGGTGAGAGATCTGCCCCAGGGCTTTAGCGCTT TGGAACCTCTGGTGGACCTGCCTATCGGAATCAACA TCACCAGATTTCAAACCTCTCCTGGCCCTGCACAGAT CTTATCTGACCCCTGGGGACAGTAGTAGCGGCTGGA CAGCCGGCGCCGCCGCTACTACGTGGGATACCTGC AGCCTAGAACATTCCTGCTGAAGTACAATGAGAAC GGAACAATCACAGACGCCGTGGACTGCGCCCTGGA TCCTTTGAGCGAGACAAAGTGCACCCTGAAGTCGTT CACCGTCGAAAAAGGCATCTACCAGACCAGCAACT TCCGCGTGACGCTACGGAATCTATCGTGCGGTTCC CCAACATACCAACCTGTGCCCTTTTCGGCGAGGTGT TTAACGCTACAAGGTTCCGCCAGCGTGTATGCCTGGA ACAGAAAGAGAATCAGCAATTGCGTGGCCGATTAT AGCGTTCTGTACAACAGCGCTTCCTTCAGCACCTTC AAGTGCTACGGCGTGTCTCCAACCAAGCTGAACGA CCTCTGCTTACCAATGTCTACGCTGACTCTTTCGTG ATTAGAGGCGATGAGGTTAGACAGATCGCACCTGG CCAGACCGGCAAAATCGCTGACTACAACTACAAGC TGCCTGATGACTTCACAGGCTGTGTCATTGCCTGGA ACTCAAATAACCTGGACTCTAAAGTGGGCGGCAAC TACAACTACCTGTACCGGCTGTTCCGGAAGAGCAAT CTGAAACCTTTTGAGCGGGACATCTCTACAGAGATC TACCAGGCCGGCAGCACACCCTGCAACGGCGTTGA GGGCTTCAACTGCTACTTCCCTCTGCAGAGCTACGG CTTTCAGCCAACAAATGGAGTGGGCTACCAGCCGTA CAGAGTGGTGGTGCTGAGCTTCGAACTGCTGCATGC
--	--	---

	CCCAGCCACAGTGTGTGGACCTAAGAAGTCTACCA ACCTGGTGAAGAACAAGTGCCTGAACCTTTAACTTTA ACGGCCTGACCGGCACAGGCGTGCTGACCGAATCC AACAAAAAGTTCCTGCCCTTCCAACAGTTCGGCAGA GACATCGCCGATACAACCGATGCCGTGCGGGACCC CCAGACCTTAGAAATCCTAGATATCACCCCGTGCAG CTTGCGCGGAGTCTCTGTTATTACTCCTGGCACCAA CACCAGCAACCAAGTGGCTGTTCTGTACCAAggcGTG AACTGCACCGAAGTGCCTGTGGCTATCCACGCCGAT CAGCTGACCCCAACCTGGCGGGTGTATAGCACCGG CTCTAACGTGTTCCAGACCCGGGCTGGCTGCCTGAT CGGCGCCGAACACGTCAACAACCTCCTATGAATGTG ACATCCCCATCGGGGCTGGCATCTGCGCCAGTTACC AGACACAGACAAATAGCCCTAGACGGGCCAGAAGC GTGGCCTCCCAGAGTATCATTGCCTACACCATGAGC CTGGGCGCCGAGAACAGCGTGGCCTATTCTAACAAAT AGCATCGCAATCCCTACCAACTTTACCATCTCTGTG ACAACCGAGATCCTGCCTGTGAGCATGACCAAAAC CAGCGTGGACTGCACGATGTACATCTGTGGCGACA GCACAGAATGCAGTAATCTGTTGCTGCAGTACGGCA GCTTTTGCACCCAGTTGAATAGAGCCCTGACCGGAA TCGCCGTAGAGCAGGACAAAAATACCCAGGAGGTG TTCGCCCAGGTGAAACAGATCTACAAGACACCTCCC ATTAAGGACTTCGGAGGTTTTAACTTCAGCCAGATC CTGCCCCGACCCTTCCAAGCCTAGCAAACGCTCCTTC ATCGAGGACCTGCTCTTCAACAAGGTGACACTGGCT GATGCCGGCTTCATCAAGCAGTACGGAGATTGTCTG GGAGACATCGCCGCTAGAGATCTGATCTGCGCCCA AAAGTTCAACGGCCTGACCGTGCTGCCTCCTCTGCT TACAGACGAGATGATCGCCAGTACACCAGCGCCC TGCTGGCTGGCACCATCACAAGCGGCTGGACCTTCG GAGCCGGAGCCGCTCTGCAAATCCCCTTTGCCATGC AGATGGCCTACCGGTTCAACGGCATCGGCGTGACA CAGAAATGTGCTGTACGAGAACCAGAAGCTGATCGC TAACCAGTTTAACAGCGCTATCGGCAAGATCCAGG ACTCGCTGAGTAGCACCGCCTCTGCCCTGGGCAAGC TGCAGGACGTGCTGAACCAGAACGCCCAAGCCCTG AACACACTGGTGAAACAGCTGAGCAGCAACTTCGG CGCCATCAGCTCTGTGCTGAACGATATCCTGAGCAG ACTGGACAAGGTGGAAGCCGAGGTCCAGATCGACA GACTGATCACAGGAAGACTGCAGAGCCTGCAAACG TACGTGACACAGCAGCTGATCCGGGCAGCCGAAAT CCGGGCCAGCGCCAATCTGGCCGCTACCAAGATGA GCGAGTGCGTGTTAGGCCAGAGCAAGCGGGTGGAT TTCTGCGGTAAAGGATAACCACTGATGAGCTTTCCC CAGAGCGTCTCTCACGGCGTGGTGTCTTCTGCACGTG ACCTACGTTTCTGCCAGGAAAAGAACTTCACCACC GCCCCTGCTATCTGCCACGATGGCAAGGCCCACTTC CCTAGAGAGGGCGTTTTTCGTGTCTAACGGCACACAC TGGTTTGTGACCCAGAGAACTTCTACGAGCCTCAG
--	---

		ATCATCACACAGACAACACCTTTGTGAGCGGCAAT TGCACGTGGTGATCGGAATTGTTAATAATACCGTG TACGACCTCTGCAGCCTGAGCTCGACAGCTTCAAG GAAGAGCTGGACAAGTACTTCAAGAACCACACCTC CCCAGATGTGGACCTGGGCGATATTTAGGCATCAA CGCCTCCGTCGTGAATATCCAGAAGGAGATCGACC GGCTCAACGAGGTGGCCAAGAACCTTAACGAGAGC CTGATCGACCTGCAGGAACCTGGGCAAATATGAGCA GTACATCAAGTGGCCTTGGTACATCTGGCTGGGCTT TATCGCAGGCCTGATCGCTATCGTGATGGTGACCAT TATGCTGTGTTGTATGACCAGCTGTTGTAGTTGTCTG AAGGGCTGCTGTTCTTGCGGCAGCTGCTGCAAGTTC GACGAAGACGACTCAGAGCCCCTGCTGAAAGGCGT GAAGCTGCACTACACCCGAAAACGGCGCggaagcggag gaagcggagctactaacttcagcctgctgaagcaggctggagatgtggaggagaac cctggacctATGTATTCTTTTGTGTCCGAGGAAACCGGC ACACTGATCGTTAATAGCGTGCTGCTCTTCCTGGCC TTCGTGGTGTTCCTGCTGGTGACCCTGGCTATCCTG ACCGCCCTGAGACTGTGTGCCTACTGCTGCAACATC GTGAACGTGTCTCTGGTCAAGCCTAGCTTCTACGTG TACAGCCGGGTGAAGAACCTGAACAGCAGCAGAGT GCCCACCTGCTGGTGaatccccccccctaacgttactggccgaag ccgcttgaataaggccggtgtgcgttctctatatgtatttccaccatattgccgtctt tggcaatgtgagggcccgaaacctggccctgtcttcttgacgagcattctaggggt ctttccctctcgccaaaggaatgcaaggtctgtgaatgtcgtgaaggaaagcagttcc tctggaagcttctgaagacaaacaacgtctgtacgaccccttgacggcagcgggaac ccccacctggcgacaggtgcctctcgcccaaaagccacgtgataagatacacct gcaaaggcggcacaacccagtgccacgttgtagttgatagttggaagagtc aatggctctctcaagcgtattcaacaaggggctgaaggatgccagaaggtaccc cattgtatgggatctgatctggggcctcggtgcacatgctttacatgtgttagtcgaggt taaaaaacgtctagggccccgaaccacggggacgtggttttctttgaaaaacacg atgataatatggccacaacctgaacaagagactgcgcgcactctctacattttgag gaatgccccaaatgctctgctctacaataaccgtaatggattttacctgctaaagtatgat gaagaatggtagccagaggattattgactgatggagaggatgatgtctttgatccg aattagacatggaagtcgtttcgagttacagtaa
CoVEG5 Expression cassette	37	ATGGCCGACAGCAACGGCACAAATCACCGTGGAAGA GCTGAAGAACTGCTGGAACAGTGGAACCTGGTCA TCGGCTTCTGTTTCTGACCTGGATCTGTCTGCTGCA GTTCGCTTATGCCAATCGGAACAGATTCTGTACAT CATCAAGCTGATCTTCCTGTGGCTGCTGTGGCCTGT GACCCTGGCTTGCTTCGTGCTGGCCGCTGTGTACCG GATCAACTGGATCACAGGCGGAATCGCCATCGCCA TGGCCTGCCTGGTGGGCTGATGTGGCTGAGCTACT TCATCGCTTCTTTCAGACTGTTCGCCAGAACCCGGA GCATGTGGTCCTTCAACCCCGAGACAAACATCCTGC TGAACGTGCCTCTGCACGGCACCATCCTGACAAGAC CTCTGCTCGAGAGCGAGCTGGTGATTGGCGCAGTGA TTCTGAGAGGCCATCTGAGGATCGCCGGACACCACC TGGGCAGATGCGACATCAAGGACCTTCCAAAGGAA ATCACCGTTGCCACCAGCCGGACCCTGTCCTACTAC

		AAACTGGGCGCCAGCCAAAGAGTGGCCGCGGATAG CGGCTTTGCCGCCTACAGCAGATACCGCATCGGAAA TTACAAGCTCAACACCGACCACAGCAGCTCTTCTGA TAACATCGCCCTGCTGGTGCAGCGAAAACGGCGCg aagcggaggaagcggagctactaacttcagcctgctgaagcaggctggagatgtgg aggagaaccctggacctATGAGCGACAACGGCCCTCAAAAC CAGAGAAATGCCCCTCGGATCACATTTGGCGGACCT AGCGACAGCACCGGCAGCAACCAGAATGGAGAAAAG AAGCGGCGCCAGATCCAAGCAGCGGAGACCTCAGG GACTGCCCAACAACACCGCTAGCTGGTTCACCGCCC TGACCCAACACGGCAAGGAAGATCTGAAGTTCCCC AGAGGCCAGGGCGTGCCTATCAACACAACTCTTCT CCCGACGACCAGATCGGATACTATAGACGGGCCAC TCGGAGAATTCTGGGGCGGCGACGGAAAAATGAAGG ACCTTTCTCCAAGATGGTACTTCTACTACCTCGGCA CAGGCCCTGAGGCCGGCCTGCCTTACGGCGCCAAC AAGGATGGCATCATCTGGGTGCGCACCGAGGGCGC CCTGAACACCCCTAAGGACCACATCGGCACAAGAA ACCCCGCTAACAACGCCGCAATCGTGCTGCAGCTGC CTCAGGGCACCACCCTGCCCAAGGGCTTCTACGCCG AGGGCTCTAGAGGTGGCTCCCAGGCTTCTAGCCGCT CCTCCAGCCGCAGCAGAAACAGCAGCAGGAACAGC ACCCCGGCAGCTCCCGGGGCACCAGCCCCGCCAG AATGGCCGGAATGGCGGCGATGCCGCCCTGGCCC TGCTCCTGCTGGACAGACTGAATCAGCTGGAAAGC AAGATGAGCGGCAAAGGACAGCAGCAGCAAGGCC AGACCGTGACCAAGAAAAGCGCTGCTGAAGCCTCC AAGAAACCTAGACAAAAGCGGACCGCCACAAAGGC CTACAACGTGACCCAAGCCTTTGGAAGAAGAGGGC CCGAGCAGACACAGGGCAATTCGGCGACCAGGAG CTGATCCGGCAGGGAACCGACTACAAGCACTGGCC TCAGATCGCCCAGTTCGCCCCTAGCGCCAGCGCCTT CTTCGGCATGAGCAGAAATCGGCATGGAAGTGACCC CTTCTGGCACCTGGCTGACCTACACCGGCGCTATCA AGCTGGACGATAAGGATCCTAACTTCAAGGACCAA GTGATCCTGCTGAACAAGCATATCGACGCCTATAAG ACCTTTCCACCTACAGAGCCTAAGAAAGATAAGAA GAAGAAAGCCGACGAGACACAGGCCCTGCCTCAGA GACAGAAAAAGCAGCAGACAGTGACACTGCTGCCA GCCGCTGACCTGGATGACTTCAGCAAGCAGCTGCA GCAGAGCATGTCTTCTGCTGATAGCACCCAGGCCCCG AAAACGGCGCggaagcggaggaagcggagctactaacttcagcctgctga agcaggctggagatgtggaggagaaccctggacctATGTTCTGTGTTCT GGTGCTGCTGCCTCTGGTCAGCTCCCAGTGTGTGAA CCTGACCACCAGAACCCAGCTGCCACCTGCTTATAC AAACTCCTTCACTCGGGGGGTATACTACCCCGACAA GGTGTTCAGATCTAGCGTGCTGCATTCTACACAAGA CCTGTTCCTGCCCTTCTTCAGCAACGTGACCTGGTTC CACGCCATCCACGTGTCTGGAACCAACGGAACCAA GAGATTCGACAACCCCGTGCTGCCTTTCAACGACGG
--	--	---

		CGTGTACTTCGCCAGCACCGAGAAGTCCAACATCAT CAGAGGATGGATTTTCGGCACCACTGGACAGCA AAACCCAGAGCCTGCTGATCGTGAACAACGCCACC AACGTGGTGATCAAGGTGTGCGAGTTCCAGTTCTGC AATGATCCCTTCCTGGGCGTGTACTACCACAAGAAC AACAAGTCTTGGATGGAAAGCGAGTTCAGAGTGTA TTCCAGCGCCAACAATTGCACCTTCGAGTACGTGAG CCAACCCTTTCTGATGGACCTTGAAGGCAAGCAGGG CAACTTCAAAAATCTGCGAGAATTTGTGTTCAAGAA CATCGACGGATACTTCAAGATCTACTCTAAGCACAC GCCAATCAACCTGGTGAGAGATCTGCCCCAGGGCTT TAGCGCTTTGGAACCTCTGGTGGACCTGCCTATCGG AATCAACATCACCAGATTTCAAACCTCTCCTGGCCCT GCACAGATCTTATCTGACCCCTGGGGACAGTAGTAG CGGCTGGACAGCCGGCGCCGCCGCTACTACGTGG GATACCTGCAGCCTAGAACATTCCTGCTGAAGTACA ATGAGAACGGAACAATCACAGACGCCGTGGACTGC GCCCTGGATCCTTTGAGCGAGACAAAGTGCACCCTG AAGTCGTTACCGTCGAAAAAGGCATCTACCAGAC CAGCAACTTCCGCGTGCAGCCTACGGAATCTATCGT GCGGTTCCCCAACATCACC AACCTGTGCCCTTTTCGG CGAGGTGTTTAACGCTACAAGGTTCCGCCAGCGTGTA TGCTTGGAACAGAAAGAGAATCAGCAATTGCGTGG CCGATTATAGCGTTCTGTACAACAGCGCTTCCTTCA GCACCTTCAAGTGCTACGGCGTGTCTCCAACCAAGC TGAACGACCTCTGCTTCACCAATGTCTACGCTGACT CTTTCGTGATTAGAGGCGATGAGGTTAGACAGATCG CACCTGGCCAGACCGGCAAAATCGCTGACTACAAC TACAAGCTGCCTGATGACTTCACAGGCTGTGTCAAT GCCTGGAAC TCAAATAACCTGGACTCTAAAGTGGG CGGCAACTACAAC TACCTGTACCGGCTGTTCCGGAA GAGCAATCTGAAACCTTTTGAGCGGGACATCTCTAC AGAGATCTACCAGGCCGGCAGCACACCCTGCAACG GCGTTGAGGGCTTCAACTGCTACTTCCCTCTGCAGA GCTACGGCTTTCAGCCAACAAATGGAGTGGGCTACC AGCCGTACAGAGTGGTGGTGCTGAGCTTCGAACTGC TGCATGCCCCAGCCACAGTGTGTGGACCTAAGAAGT CTACCAACCTGGTGAAGAACAAGTGCCTGAACTTTA ACTTTAACGGCCTGACCGGCACAGGCGTGCTGACCG AATCCAACAAAAAGTTCCTGCCCTTCCAACAGTTCG GCAGAGACATCGCCGATACAACCGATGCCGTGCGG GACCCCCAGACCTTAGAAATCCTAGATATCACCCCCG TGCAGCTTCGGCGGAGTCTCTGTTATTACTCCTGGC ACCAACACCAGCAACCAAGTGGCTGTTCTGTACCAA ggcGTGAACTGCACCGAAGTGCCTGTGGCTATCCACG CCGATCAGCTGACCCCAACCTGGCGGGTGTATAGCA CCGGCTCTAACGTGTTCCAGACCCGGGCTGGCTGCC TGATCGGCGCCGAACACGTCAACAACCTCTATGAAT GTGACATCCCCATCGGGGCTGGCATCTGCGCCAGTT ACCAGACACAGACAAATAGCCCTGGCAGCGCCAGC
--	--	---

		AGCGTGGCCTCCCAGAGTATCATTGCCTACACCATG AGCCTGGGCGCCGAGAACAGCGTGGCCTATTCTAA CAATAGCATCGCAATCCCTACCAACTTTACCATCTC TGTGACAACCGAGATCCTGCCTGTGAGCATGACCAA AACCAGCGTGGACTGCACGATGTACATCTGTGGCG ACAGCACAGAATGCAGTAATCTGTTGCTGCAGTACG GCAGCTTTTGCACCCAGTTGAATAGAGCCCTGACCG GAATCGCCGTAGAGCAGGACAAAAATACCCAGGAG GTGTTGCCCCAGGTGAAACAGATCTACAAGACACCT CCCATTAAGGACTTCGGAGGTTTTAACTTCAGCCAG ATCCTGCCCCGACCCTTCCAAGCCTAGCAAACGCTCC TTCATCGAGGACCTGCTCTTCAACAAGGTGACACTG GCTGATGCCGGCTTCATCAAGCAGTACGGAGATTGT CTGGGAGACATCGCCGCTAGAGATCTGATCTGCGCC CAAAAGTTCAACGGCCTGACCGTGCTGCCTCCTCTG CTTACAGACGAGATGATCGCCAGTACACCAGCGC CCTGCTGGCTGGCACCATCACAAGCGGCTGGACCTT CGGAGCCGGAGCCGCTCTGCAAATCCCCTTTGCCAT GCAGATGGCCTACCGGTTCAACGGCATCGGCGTGA CACAGAATGTGCTGTACGAGAACCAGAAGCTGATC GCTAACCAGTTTAAACAGCGCTATCGGCAAGATCCAG GACTCGCTGAGTAGCACCGCCTCTGCCCTGGGCAAG CTGCAGGACGTCGTGAACCAGAACGCCCAAGCCCT GAACACACTGGTGAAACAGCTGAGCAGCAACTTCG GCGCCATCAGCTCTGTGCTGAACGATATCCTGAGCA GACTGGACCCTcccGAAGCCGAGGTCCAGATCGACA GACTGATCACAGGAAGACTGCAGAGCCTGCAAACG TACGTGACACAGCAGCTGATCCGGGCAGCCGAAAT CCGGGCCAGCGCCAATCTGGCCGCTACCAAGATGA GCGAGTGCGTGTTAGGCCAGAGCAAGCGGGTGGAT TTCTGCGGTAAGGGATACCACTGATGAGCTTTCCC CAGAGCGTCTCTCACGGCGTGGTGTCTTCTGCACGTG ACCTACGTTTCTGCCAGGAAAAAGAACTTACCACC GCCCCTGCTATCTGCCACGATGGCAAGGCCCACTTC CCTAGAGAGGGCGTTTTTCGTGTCTAACGGCACACAC TGGTTTGTGACCCAGAGAACTTCTACGAGCCTCAG ATCATCACCACAGACAACACCTTTGTGAGCGGCAAT TGCGACGTGGTGATCGGAATTGTTAATAATACCGTG TACGACCCTCTGCAGCCTGAGCTCGACAGCTTCAAG GAAGAGCTGGACAAGTACTTCAAGAACCACACCTC CCCAGATGTGGACCTGGGCGATATTTACGGCATCAA CGCCTCCGTCGTGAATATCCAGAAGGAGATCGACC GGCTCAACGAGGTGGCCAAGAACCTTAACGAGAGC CTGATCGACCTGCAGGAAGTGGGCAAATATGAGCA GTACATCAAGTGGCCTTGGTACATCTGGCTGGGCTT TATCGCAGGCCTGATCGCTATCGTGATGGTGACCAT TATGCTGTGTTGTATGACCAGCTGTTGTAGTTGTCTG AAGGGCTGCTGTTCTTGCGGCAGCTGCTGCAAGTTC GACGAAGACGACTCAGAGCCCCTGCTGAAAGGCGT GAAGCTGCACTACACCCGAAAACGGCGCggaagcggag
--	--	--

		gaagcggagctactaacttcagcctgctgaagcaggctggagatgtggaggagaac cctggacctATGTATTCTTTTGTGTCCGAGGAAACCGGC ACACTGATCGTTAATAGCGTGCTGCTCTTCCTGGCC TTCGTGGTGTTCCTGCTGGTGACCCCTGGCTATCCTG ACCGCCCTGAGACTGTGTGCCTACTGCTGCAACATC GTGAACGTGTCTCTGGTCAAGCCTAGCTTCTACGTG TACAGCCGGGTGAAGAACCTGAACAGCAGCAGAGT GCCCCACCTGCTGGTGtaatccccccccctaacgttactggccgaag ccgcttggaataaggccggtgtgctgttctatatgtatttccaccatattgccgtttt tggcaatgtgagggcccgaacacctggccctgtcttctgacgagcattcctaggggt ctttccctctcgccaaaggaatgcaaggtctgttgatgtcgtgaaggaaagcagtcc tctggaagcttctgaagacaaacacgtctgtagcgaccttgcaggcagcggaac ccccacctggcgacaggtgcctctgcggccaaaagccacgtgtataagataacct gcaaaggcggcacaacccagtgccacgttgatgttgatagttgtgaaagagtc aatggctctctcaagcgtattcaacaaggggctgaaggatgccagaaggtacct cattgtatgggatctgatctggggcctcggtgcacatgctttacatgtgttagtcgaggt taaaaaacgtctagggccccgaaccacggggacgtggttttcttgaaaaacacg atgataatatggccacaacctggaacaagagacttgcgcgactctctcacttttgag gaatgccccaaatgctctgctctacaataaccgtaatggatttacctgctaaagtatgat gaagaatgttaccagaggagtattgactgatggagaggatgatgtctttgatccg aattagacatggaagtcgttttcgagttacagtaa
CoVEG6 expression cassette	38	ATGTTCTGTGTTCTGCTGGTGCTGCTGCCTCTGGTCAGCT CCCAGTGTGTGAACCTGACCACCAGAACCCAGCTGC CACCTGCTTATACAACTCCTTCACTCGGGGGGTAT ACTACCCCGACAAGGTGTTTCAAGATCTAGCGTGCTGC ATTCTACACAAGACCTGTTCTGCTTCTTCTCAGCA ACGTGACCTGGTTCACGCCATCCACGTGTCTGGAA CCAACGGAACCAAGAGATTCGACAACCCCGTGCTG CCTTTCAACGACGGCGTGTACTTCGCCAGCACCCGAG AAGTCCAACATCATCAGAGGATGGATTTTCGGCACC ACACTGGACAGCAAAACCCAGAGCCTGCTGATCGT GAACAACGCCACCAACGTGGTGATCAAGGTGTGCG AGTTCCAGTTCTGCAATGATCCCTTCTGGGCGTGT ACTACCACAAGAACAACAAGTCTTGATGGAAAGC GAGTTCAGAGTGTATTCCAGCGCCAACAATTGCACC TTCGAGTACGTGAGCCAACCTTTCTGATGGACCTT GAAGGCAAGCAGGGCAACTTCAAAAATCTGCGAGA ATTTGTGTTCAAGAATCGACGGATACTTCAAGAT CTACTCTAAGCACACGCCAATCAACCTGGTGAGAG ATCTGCCCCAGGGCTTTAGCGCTTTGGAACCTCTGG TGGACCTGCCTATCGGAATCAACATCACCAGATTTC AAACTCTCCTGGCCCTGCACAGATCTTATCTGACCC CTGGGGACAGTAGTAGCGGCTGGACAGCCGGCGCC GCCGCTACTACGTGGGATACCTGCAGCCTAGAACA TTCCTGCTGAAGTACAATGAGAACGGAACAATCAC AGACGCCGTGGACTGCGCCCTGGATCCTTTGAGCGA GACAAAGTGCACCCTGAAGTCGTTACCCGTCGAAA AAGGCATCTACCAGACCAGCAACTTCCGCGTGACG CCTACGGAATCTATCGTGCGGTTCCCAACATCACC AACCTGTGCCCTTTCGGCGAGGTGTTAACGCTACA

		AGGTTTCGCCAGCGTGTATGCCTGGAACAGAAAGAG AATCAGCAATTGCGTGGCCGATTATAGCGTTCTGTA CAACAGCGCTTCCTTCAGCACCTTCAAGTGCTACGG CGTGTCTCCAACCAAGCTGAACGACCTCTGCTTCAC CAATGTCTACGCTGACTCTTTCGTGATTAGAGGGCGA TGAGGTTAGACAGATCGCACCTGGCCAGACCGGCA AAATCGCTGACTACAACCTACAAGCTGCCTGATGACT TCACAGGCTGTGTCATTGCCTGGAACCTCAAATAACC TGGACTCTAAAGTGGGCGGCAACTACAACCTACCTGT ACCGGCTGTTCCGGAAGAGCAATCTGAAACCTTTTG AGCGGGACATCTCTACAGAGATCTACCAGGCCGGC AGCACACCCTGCAACGGCGTTGAGGGCTTCAACTGC TACTTCCCTCTGCAGAGCTACGGCTTTCAGCCAACA AATGGAGTGGGCTACCAGCCGTACAGAGTGGTGGT GCTGAGCTTCGAACTGCTGCATGCCCCAGCCACAGT GTGTGGACCTAAGAAGTCTACCAACCTGGTGAAGA ACAAGTGCGTGAACTTTAACTTTAACGGCCTGACCG GCACAGGCGTGCTGACCGAATCCAACAAAAAGTTC CTGCCCTTCCAACAGTTCGGCAGAGACATCGCCGAT ACAACCGATGCCGTGCGGGACCCCCAGACCTTAGA AATCCTAGATATCACCCCGTGCAGCTTCGGCGGAGT CTCTGTTATTACTCCTGGCACCAACACCAGCAACCA AGTGGCTGTTCTGTACCAAggcGTGAACCTGCACCGAA GTGCCTGTGGCTATCCACGCCGATCAGCTGACCCCA ACCTGGCGGGTGTATAGCACCGGCTCTAACGTGTTT CAGACCCGGGCTGGCTGCCTGATCGGCGCCGAACA CGTCAACAACTCCTATGAATGTGACATCCCCATCGG GGCTGGCATCTGCGCCAGTTACCAGACACAGACAA ATAGCCCTAGACGGGCCAGAAGCGTGGCCTCCAG AGTATCATTGCCTACACCATGAGCCTGGGCGCCGAG AACAGCGTGGCCTATTCTAACAATAGCATCGCAATC CCTACCAACTTTACCATCTCTGTGACAACCGAGATC CTGCCTGTGAGCATGACCAAAAACCAGCGTGGACTG CACGATGTACATCTGTGGCGACAGCACAGAATGCA GTAATCTGTTGCTGCAGTACGGCAGCTTTTGCACCC AGTTGAATAGAGCCCTGACCGGAATCGCCGTAGAG CAGGACAAAAATACCCAGGAGGTGTTGCCCCAGGT GAAACAGATCTACAAGACACCTCCCATTAAGGACTT CGGAGGTTTTAACTTCAGCCAGATCCTGCCCGACCC TTCCAAGCCTAGCAAACGCTCCTTCATCGAGGACCT GCTCTTCAACAAGGTGACACTGGCTGATGCCGGCTT CATCAAGCAGTACGGAGATTGTCTGGGAGACATCG CCGCTAGAGATCTGATCTGCGCCCAAAAGTTCAACG GCCTGACCGTGCTGCCTCCTCTGCTTACAGACGAGA TGATCGCCAGTACACCAGCGCCCTGCTGGCTGGCA CCATCACAAGCGGCTGGACCTTCGGAGCCGGAGCC GCTCTGCAAATCCCCTTTGCCATGCAGATGGCCTAC CGGTTCAACGGCATCGGCGTGACACAGAATGTGCT GTACGAGAACCAGAAGCTGATCGCTAACCAAGTTTA ACAGCGCTATCGGCAAGATCCAGGACTCGCTGAGT
--	--	--

		AGCACCGCCTCTGCCCTGGGCAAGCTGCAGGACGTC GTGAACCAGAACGCCCAAGCCCTGAACACACTGGT GAAACAGCTGAGCAGCAACTTCGGCGCCATCAGCT CTGTGCTGAACGATATCCTGAGCAGACTGGACAAG GTGGAAGCCGAGGTCCAGATCGACAGACTGATCAC AGGAAGACTGCAGAGCCTGCAAACGTACGTGACAC AGCAGCTGATCCGGGCAGCCGAAATCCGGGCCAGC GCCAATCTGGCCGCTACCAAGATGAGCGAGTGCGT GTTAGGCCAGAGCAAGCGGGTGGATTTCTGCGGTA AGGGATACCACCTGATGAGCTTTCACAGAGCGCTC CTCACGGCGTGGTGTTTCTGCACGTGACCTACGTT CTGCCCAGGAAAAGAACTTCACCACCGCCCCTGCTA TCTGCCACGATGGCAAGGCCCACTTCCTAGAGAGG GCGTTTTCTGTGTCTAACGGCACACACTGGTTTGTGA CCCAGAGAACTTCTACGAGCCTCAGATCATCACCA CAGACAACACCTTTGTGAGCGGCAATTGCGACGTG GTGATCGGAATTGTTAATAATACCGTGTACGACCCT CTGCAGCCTGAGCTCGACAGCTTCAAGGAAGAGCT GGACAAGTACTTCAAGAACCACACCTCCCCAGATGT GGACCTGGGCGATATTTACAGGCATCAACGCCTCCGT CGTGAATATCCAGAAGGAGATCGACCGGCTCAACG AGGTGGCCAAGAACCTTAACGAGAGCCTGATCGAC CTGCAGGAAGTGGGCAATATGAGCAGTACATCAA GTGGCCTTGGTACATCTGGCTGGGCTTTATCGCAGG CCTGATCGCTATCGTGATGGTGACCATTATGCTGTG TTGTATGACCAGCTGTTGTAGTTGTCTGAAGGGCTG CTGTTCTTGCGGCAGCTGCTGCAAGTTCGACGAAGA CGACTCAGAGCCCCTGCTGAAAGGCGTGAAGCTGC ACTACACCTaatccccccccctaacgttactggccgaagcgcttggaataa ggccggtgtgctgttgtctatatgttatttccaccatattgcccgtcttttgcaatgtgag ggcccggaacctggccctgtcttctgacgagcattctaggggtcttccccctctcg ccaaaggaatgcaaggtctgtgaatgtcgtgaaggaagcagttcctctggaagcttct tgaagacaaacaacgtctgtagcgaccctttgcaggcagcggaacccccacctgg cgacaggtgctctgcgccaaaagccacgtgtataagatacacctgcaaaggcgg cacaacccagtgccacgttgtgagttggatagttgtggaagagtcfaatggctctc ctcaagcgtattcaacaaggggctgaaggatgcccagaaggtacccattgtatggg atctgatctggggcctcggtgcacatgctttacatgtgttagtcgaggttaaaaaaacg tctaggccccccgaaccacggggacgtggtttcctttgaaaaacacgatgataatatg gccacaacctggaacaagagacttgccgcactctctcacttttgaggaatgccc aatgtctgtctctacaataccgtaattggattttacctgctaaagtatgatgaagaatggta cccagaggagtattgactgatggagaggatgatgtctttgatcccgaattagacatgg aagtcgttttcgagttacagggaagcggagctactaacttcagcctgctgaagcaggc tggagatgtggaggagaaccctggacctATGGCCGACAGCAACGG CACAATCACCGTGGAAGAGCTGAAGAACTGCTGG AACAGTGGAACCTGGTCATCGGCTTCCTGTTTCTGA CCTGGATCTGTCTGCTGCAGTTCGCTTATGCCAATC GGAACAGATTCTGTACATCATCAAGCTGATCTTCC TGTGGCTGCTGTGGCCTGTGACCCTGGCTTGCTTCG TGCTGGCCGCTGTGTACCGGATCAACTGGATCACAG GCGGAATCGCCATCGCCATGGCCTGCCTGGTGGGCC
--	--	---

		TGATGTGGCTGAGCTACTTCATCGCTTCTTTCAGACT GTTTCGCCAGAACCCGGAGCATGTGGTCCTTCAACCC CGAGACAAACATCCTGCTGAACGTGCCTCTGCACGG CACCATCCTGACAAGACCTCTGCTCGAGAGCGAGCT GGTGATTGGCGCAGTGATTCTGAGAGGCCATCTGAG GATCGCCGGACACCACTGGGCAGATGCGACATCA AGGACCTTCCAAAGGAAATCACCGTTGCCACACG CGGACCCTGTCCTACTACAACTGGGCGCCAGCCAA AGAGTGGCCGGCGATAGCGGCTTTGCCGCCTACAG CAGATACCGCATCGGAAATTACAAGCTCAACACCG ACCACAGCAGCTCTTCTGATAACATCGCCCTGCTGG TGCAGTaaacccccccctaacgttactggccgaagccgttgaataaggccg gtgtgcgttgtctatatgtattttccaccatattgccgtcttttgcaatgtgagggcc ggaaacctggccctgtctcttgacgagcattcctaggggtctttccctctcgcctaaa ggaatgcaaggtctgtgaatgtcgtgaaggaagcagttcctctggaagctcttgaag acaaacaacgtctgtagcgacccttgcaggcagcggaacccccacctggcgaca ggtgcctctgcggccaaaagccacgtgtataagatacacctgcaaaggcgacaaa ccccagtgccacgttgtgagttggatagttgtggaagagtcaaatggctctcctcaag cgtattcaacaaggggctgaaggatgccagaaggtacccattgtatgggatctgat ctggggcctcgggtgcacatgctttacatgtgttagtcgaggttaaaaaacgtctagg cccccgaaaccaggggacgtggtttcctttgaaaaacacgatgataataggccac aacatggaacaagagacttgcgcgactctctcacttttgaggaatcccaaatgc tctgctctacaataccgtaattgattttacctgctaaagtatgatgaagaatggtacca gaggagtattgactgatggagaggatgatgtctttgatccgaattagacatggaagt cgttttcgagttacagggaagcggagctactaacttcagcctgctgaagcaggtgga gatgtggaggagaacctggacctATGAGCGACAACGGCCCTCA AAACCAGAGAAATGCCCTCGGATCACATTTGGCG GACCTAGCGACAGCACCGGCAGCAACCAGAATGGA GAAAGAAGCGGCGCCAGATCCAAGCAGCGGAGACC TCAGGGACTGCCCAACAACACCGCTAGCTGGTTTAC CGCCCTGACCCAACACGGCAAGGAAGATCTGAAGT TCCCCAGAGGCCAGGGCGTGCCTATCAACACAAAC TCTTCTCCCGACGACCAGATCGGATACTATAGACGG GCCACTCGGAGAATTTCGGGGCGGCGACGGAAAAAT GAAGGACCTTTCTCCAAGATGGTACTTCTACTACCT CGGCACAGGCCCTGAGGCCGGCCTGCCTTACGGCG CCAACAAGGATGGCATCATCTGGGTGCCACCGAG GGCGCCCTGAACACCCCTAAGGACCACATCGGCAC AAGAAACCCCGCTAACAACGCCGCAATCGTGCTGC AGCTGCCTCAGGGCACCACTGCCCCAAGGGCTTCT ACGCCGAGGGCTCTAGAGGTGGCTCCCAGGCTTCTA GCCGCTCCTCCAGCCGAGCAGAAACAGCAGCAGG AACAGCACCCCCGGCAGCTCCCGGGGCACAGCCC CGCCAGAATGGCCGGAATGGCGGCGATGCCGCC TGGCCCTGCTCCTGCTGGACAGACTGAATCAGCTGG AAAGCAAGATGAGCGGCAAAGGACAGCAGCAGCA AGGCCAGACCGTGACCAAGAAAAGCGCTGCTGAAG CCTCCAAGAAACCTAGACAAAAGCGGACCGCCACA AAGGCCTACAACGTGACCCAAGCCTTTGGAAGAAG AGGCCCCGAGCAGACACAGGGCAATTTCCGGCGACC
--	--	--

		AGGAGCTGATCCGGCAGGGAACCGACTACAAGCAC TGGCCTCAGATCGCCCAGTTCGCCCCTAGCGCCAGC GCCTTCTTCGGCATGAGCAGAATCGGCATGGAAGTG ACCCCTTCTGGCACCTGGCTGACCTACACCGGCGCT ATCAAGCTGGACGATAAGGATCCTAACTTCAAGGA CCAAGTGATCCTGCTGAACAAGCATATCGACGCCTA TAAGACCTTTCCACCTACAGAGCCTAAGAAAGATA AGAAGAAGAAAGCCGACGAGACACAGGCCCTGCCT CAGAGACAGAAAAAGCAGCAGACAGTGACACTGCT GCCAGCCGCTGACCTGGATGACTTCAGCAAGCAGCT GCAGCAGAGCATGTCTTCTGCTGATAGCACCCAGGC Cggaagcggagctactaacttcagcctgctgaagcaggctggagatgtggaggag aaccttgacctATGTATTCTTTTGTGTCCGAGGAAACCGG CACACTGATCGTTAATAGCGTGCTGCTCTTCCTGGC CTTCGTGGTGTTCCTGCTGGTGACCCTGGCTATCCTG ACCGCCCTGAGACTGTGTGCCTACTGCTGCAACATC GTGAACGTGTCTCTGGTCAAGCCTAGCTTCTACGTG TACAGCCGGGTGAAGAACCTGAACAGCAGCAGAGT GCCCGACCTGCTGGTGTGA
CoVEG7 Expression cassette	39	ATGGCCGACAGCAACGGCACAATCACCGTGGAAGA GCTGAAGAACTGCTGGAACAGTGGAACCTGGTCA TCGGCTTCCTGTTTCTGACCTGGATCTGTCTGCTGCA GTTTCGCTTATGCCAATCGGAACAGATTCTGTACAT CATCAAGCTGATCTTCCTGTGGCTGCTGTGGCCTGT GACCCTGGCTTGCTTCGTGCTGGCCGCTGTGTACCG GATCAACTGGATCACAGGCGGAATCGCCATCGCCA TGGCCTGCCTGGTGGGCTGATGTGGCTGAGCTACT TCATCGCTTCTTTCAGACTGTTCCGCCAGAACCCGGA GCATGTGGTCCTTCAACCCCGAGACAAACATCCTGC TGAACGTGCCTCTGCACGGCACCATCCTGACAAGAC CTCTGCTCGAGAGCGAGCTGGTGATTGGCGCAGTGA TTCTGAGAGGCCATCTGAGGATCGCCGGACACCACC TGGGCAGATGCGACATCAAGGACCTTCCAAAGGAA ATCACCGTTGCCACCAGCCGGACCCTGTCCTACTAC AACTGGGCGCCAGCCAAAGAGTGGCCGGCGATAG CGGCTTTGCCGCTACAGCAGATACCGCATCGGAAA TTACAAGCTCAACACCGACCACAGCAGCTCTTCTGA TAACATCGCCCTGCTGGTGCAGCGAAAACGGCGCggg aagcggaggaagcggagctactaacttcagcctgctgaagcaggctggagatgtgg aggagaaccttgacctATGTTCGTGTTCTGCTGCTGCTGC CTCTGGTCAGCTCCCAGTGTGTGAACCTGACCACCA GAACCCAGCTGCCACCTGCTTATACAACTCCTTCA CTCGGGGGGTATACTACCCCGACAAGGTGTTTCAGAT CTAGCGTGCTGCATTCTACACAAGACCTGTTCTCTGC CCTTCTTCAGCAACGTGACCTGGTTCCACGCCATCC ACGTGTCTGGAACCAACGGAACCAAGAGATTTCGAC AACCCCGTGCTGCCTTTCAACGACGGCGTGTACTTC GCCAGCACCGAGAAGTCCAACATCATCAGAGGATG GATTTTCGGCACCACTGGACAGCAAAACCCAGA GCCTGCTGATCGTGAACAACGCCACCAACGTGGTG

		ATCAAGGTGTGCGAGTTCCAGTTCTGCAATGATCCC TTCCTGGGCGTGTACTACCACAAGAACAACAAGTCT TGGATGGAAAGCGAGTTCAGAGTGTATTCCAGCGC CAACAATTGCACCTTCGAGTACGTGAGCCAACCCTT TCTGATGGACCTTGAAGGCAAGCAGGGCAACTTCA AAAATCTGCGAGAATTTGTGTTCAAGAACATCGACG GATACTTCAAGATCTACTCTAAGCACACGCCAATCA ACCTGGTGAGAGATCTGCCCCAGGGCTTTAGCGCTT TGGAACCTCTGGTGGACCTGCCTATCGGAATCAACA TCACCAGATTTCAAACCTCTCCTGGCCCTGCACAGAT CTTATCTGACCCCTGGGGACAGTAGTAGCGGCTGGA CAGCCGGCGCCGCCGCTACTACGTGGGATACCTGC AGCCTAGAACATTCCTGCTGAAGTACAATGAGAAC GGAACAATCACAGACGCCGTGGACTGCGCCCTGGA TCCTTTGAGCGAGACAAAGTGCACCCTGAAGTCGTT CACCGTCGAAAAAGGCATCTACCAGACCAGCAACT TCCGCGTGCAGCCTACGGAATCTATCGTGCGGTTCC CCAACATCACCAACCTGTGCCCTTTCGGCGAGGTGT TTAACGCTACAAGGTTGCCAGCGTGTATGCCTGGA ACAGAAAGAGAATCAGCAATTGCGTGGCCGATTAT AGCGTTCTGTACAACAGCGCTTCCTTCAGCACCTTC AAGTGCTACGGCGTGTCTCCAACCAAGCTGAACGA CCTCTGCTTCACCAATGTCTACGCTGACTCTTTCGTG ATTAGAGGCGATGAGGTTAGACAGATCGCACCTGG CCAGACCGGCAAAATCGCTGACTACAACCTACAAGC TGCCTGATGACTTCACAGGCTGTGTCATTGCCTGGA ACTCAAATAACCTGGACTCTAAAGTGGGCGGCAAC TACAACCTGTACCGGCTGTTCCGGAAGAGCAAT CTGAAACCTTTTGAGCGGGACATCTCTACAGAGATC TACCAGGCCGGCAGCACACCCTGCAACGGCGTTGA GGGCTTCAACTGCTACTTCCCTCTGCAGAGCTACGG CTTTCAGCCAACAAATGGAGTGGGCTACCAGCCGTA CAGAGTGGTGGTGCTGAGCTTCGAACTGCTGCATGC CCCAGCCACAGTGTGTGGACCTAAGAAGTCTACCA ACCTGGTGAAGAACAAGTGCCTGAACTTTAACTTTA ACGGCCTGACCGGCACAGGCGTGCTGACCGAATCC AACAAAAAGTTCCTGCCCTTCCAACAGTTCGGCAGA GACATCGCCGATACAACCGATGCCGTGCGGGACCC CCAGACCTTAGAAATCCTAGATATCACCCCGTGACG CTTCGGCGGAGTCTCTGTTATTACTCTGGCACCAA CACCAGCAACCAAGTGGCTGTTCTGTACCAAggcGTG AACTGCACCGAAGTGCTGTGGCTATCCACGCCGAT CAGCTGACCCCAACCTGGCGGGTGTATAGCACCGG CTCTAACGTGTTCCAGACCCGGGCTGGCTGCCTGAT CGGCGCCGAACACGTCAACAACCTCCTATGAATGTG ACATCCCCATCGGGGCTGGCATCTGCGCCAGTTACC AGACACAGACAAATAGCCCTAGACGGGCCAGAAGC GTGGCCTCCCAGAGTATCATTGCCTACACCATGAGC CTGGGCGCCGAGAACAGCGTGGCCTATTCTAACAAT AGCATCGCAATCCCTACCAACTTTACCATCTCTGTG
--	--	---

		ACAACCGAGATCCTGCCTGTGAGCATGACCAAAAC CAGCGTGGACTGCACGATGTACATCTGTGGCGACA GCACAGAATGCAGTAATCTGTTGCTGCAGTACGGCA GCTTTTGCACCCAGTTGAATAGAGCCCTGACCGGAA TCGCCGTAGAGCAGGACAAAAATACCCAGGAGGTG TTCGCCCAGGTGAAACAGATCTACAAGACACCTCCC ATTAAGGACTTCGGAGGTTTTAACTTCAGCCAGATC CTGCCCCGACCCTTCCAAGCCTAGCAAACGCTCCTTC ATCGAGGACCTGCTCTTCAACAAGGTGACACTGGCT GATGCCGGCTTCATCAAGCAGTACGGAGATTGTCTG GGAGACATCGCCGCTAGAGATCTGATCTGCGCCCA AAAGTTCAACGGCCTGACCGTGCTGCCTCCTCTGCT TACAGACGAGATGATCGCCAGTACACCAGCGCCC TGCTGGCTGGCACCATCACAAGCGGCTGGACCTTCG GAGCCGGAGCCGCTCTGCAAATCCCCTTTGCCATGC AGATGGCCTACCGGTTCAACGGCATCGGCGTGACA CAGAATGTGCTGTACGAGAACCAGAAGCTGATCGC TAACCAGTTTAACAGCGCTATCGGCAAGATCCAGG ACTCGCTGAGTAGCACCGCCTCTGCCCTGGGCAAGC TGCAGGACGTCTGAACCAGAACGCCCAAGCCCTG AACACACTGGTGAAACAGCTGAGCAGCAACTTCGG CGCCATCAGCTCTGTGCTGAACGATATCCTGAGCAG ACTGGACAAGGTGGAAGCCGAGGTCCAGATCGACA GACTGATCACAGGAAGACTGCAGAGCCTGCAAACG TACGTGACACAGCAGCTGATCCGGGCAGCCGAAAT CCGGGCCAGCGCCAATCTGGCCGCTACCAAGATGA GCGAGTGCGTGTTAGGCCAGAGCAAGCGGGTGGAT TTCTGCGGTAAGGGATACCACCTGATGAGCTTTCCC CAGAGCGCTCCTCACGGCGTGGTGTCTTCTGCACGTG ACCTACGTTTCTGCCAGGAAAAGAACTTACCACC GCCCCTGCTATCTGCCACGATGGCAAGGCCCACTTC CCTAGAGAGGGCGTTTTCTGTGTCTAACGGCACACAC TGGTTTGTGACCCAGAGAACTTCTACGAGCCTCAG ATCATCACCACAGACAACACCTTTGTGAGCGGCAAT TGCGACGTGGTGATCGGAATTGTTAATAATACCGTG TACGACCCTCTGCAGCCTGAGCTCGACAGCTTCAAG GAAGAGCTGGACAAGTACTTCAAGAACCACACCTC CCCAGATGTGGACCTGGGCGATATTTACGGCATCAA CGCCTCCGTCGTGAATATCCAGAAGGAGATCGACC GGCTCAACGAGGTGGCCAAGAACCTTAACGAGAGC CTGATCGACCTGCAGGAACTGGGCAAATATGAGCA GTACATCAAGTGGCCTTGGTACATCTGGCTGGGCTT TATCGCAGGCCTGATCGCTATCGTGATGGTGACCAT TATGCTGTGTTGTATGACCAGCTGTTGTAGTTGTCTG AAGGGCTGCTGTTCTTGCGGCAGCTGCTGCAAGTTC GACGAAGACGACTCAGAGCCCCTGCTGAAAGGCGT GAAGCTGCACTACACCCGAAAACGGCGCggaagcggag gaagcggagctactaactcagcctgctgaagcaggctggagatgtggaggagaac cctggacctATGTATTCTTTTGTGTCCGAGGAAACCGGC ACACTGATCGTTAATAGCGTGCTGCTCTTCCTGGCC
--	--	---

		TTCGTGGTGTTCCTGCTGGTGACCCTGGCTATCCTG ACCGCCCTGAGACTGTGTGCCTACTGCTGCAACATC GTGAACGTGTCTCTGGTCAAGCCTAGCTTCTACGTG TACAGCCGGGTGAAGAACCTGAACAGCAGCAGAGT GCCCCACCTGCTGGTGaatccccccccctaactgtactggccgaag ccgcttgaataaggccggtgtgctgttctatattgtatttccaccatattgccgtctt tggcaatgtgagggcccgaaacctggccctgtcttctgacgagcattcctaggggt ctttccctctcgccaaaggaatgcaaggtctgttgatgtcgtgaaggagcagttcc tctggaagcttctgaagacaaacaacgtctgtagcgaccttgcaggcagcggaac ccccacctggcgacaggtgcctctgcggccaaaagccacgtgtataagatacacct gcaaaggcggcacaacccagtgccacgttgtgagttgatatgttggaagagtc aatggctctcctcaagcgtattcaacaaggggctgaaggatgccagaaggtaccc cattgtatgggatctgatctggggcctcgtgcacatgctttacatgtgttagtcgaggt taaaaaacgtctagggcccccgaaccacggggacgtggtttcttgaaaaacacg atgataatatggccacaacctggaacaagagacttgcgcgactctctcacttttgag gaatgccccaaatgctctgctctacaataaccgtaattgatttacctgctaaagtatgat gaagaatggtaccagaggagtattgactgatggagaggatgatgtctttgatccg aattagacatggaagtcgtttcaggtacagggaagcggagctactaacttcagcctg ctgaagcaggctggagatgtggaggagaacctggacctATGAGCGACA ACGGCCCTCAAAACCAGAGAAATGCCCTCGGATC ACATTTGGCGGACCTAGCGACAGCACCGGCAGCAA CCAGAATGGAGAAAGAAGCGGCGCCAGATCCAAGC AGCGGAGACCTCAGGGACTGCCCAACAACACCGCT AGCTGGTTCACCGCCCTGACCCAACACGGCAAGGA AGATCTGAAGTTCCCCAGAGGCCAGGGCGTGCCTAT CAACACAACTCTTCTCCCGACGACCAGATCGGATA CTATAGACGGGCCACTCGGAGAATTTCGGGGCGGCG ACGGAAAAATGAAGGACCTTTCTCCAAGATGGTAC TTCTACTACCTCGGCACAGGCCCTGAGGCCGGCCTG CCTTACGGCGCCAACAAGGATGGCATCATCTGGGTC GCCACCGAGGGCGCCCTGAACACCCCTAAGGACCA CATCGGCACAAGAAACCCCGCTAACAACGCCGCAA TCGTGCTGCAGCTGCCTCAGGGCACCAACCTGCCCA AGGGCTTCTACGCCGAGGGCTCTAGAGGTGGCTCCC AGGCTTCTAGCCGCTCCTCCAGCCGCAGCAGAAACA GCAGCAGGAACAGCACCCCGGCAGCTCCCGGGGC ACCAGCCCCGCCAGAATGGCCGGAAATGGCGGCGA TGCCGCCCTGGCCCTGCTCCTGCTGGACAGACTGAA TCAGCTGGAAAGCAAGATGAGCGGCAAAGGACAGC AGCAGCAAGGCCAGACCGTGACCAAGAAAAGCGCT GCTGAAGCCTCCAAGAAACCTAGACAAAAGCGGAC CGCCACAAAGGCCTACAACGTGACCCAAGCCTTTG GAAGAAGAGGCCCGAGCAGACACAGGGCAATTTT GGCGACCAGGAGCTGATCCGGCAGGGAACCGACTA CAAGCACTGGCCTCAGATCGCCCAGTTTCGCCCTAG CGCCAGCGCCTTCTTCGGCATGAGCAGAATCGGCAT GGAAGTGACCCCTTCTGGCACCTGGCTGACCTACAC CGGCGCTATCAAGCTGGACGATAAGGATCCTAACTT CAAGGACCAAGTGATCCTGCTGAACAAGCATATCG ACGCCTATAAGACCTTTCCACCTACAGAGCCTAAGA
--	--	--

		AAGATAAGAAGAAGAAAGCCGACGAGACACAGGC CCTGCCTCAGAGACAGAAAAAGCAGCAGACAGTGA CACTGCTGCCAGCCGCTGACCTGGATGACTTCAGCA AGCAGCTGCAGCAGAGCATGTCTTCTGCTGATAGCA CCCAGGCC
CoVEG8 expression cassette	40	ATGGCCGACAGCAACGGCACAATCACCGTGGAAGA GCTGAAGAACTGCTGGAACAGTGGAACCTGGTCA TCGGCTTCCTGTTTCTGACCTGGATCTGTCTGCTGCA GTTTCGCTTATGCCAATCGGAACAGATTCTGTACAT CATCAAGCTGATCTTCCTGTGGCTGCTGTGGCCTGT GACCCTGGCTTGCTTCGTGCTGGCCGCTGTGTACCG GATCAACTGGATCACAGGCGGAATCGCCATCGCCA TGGCCTGCCTGGTGGGCTGATGTGGCTGAGCTACT TCATCGCTTCTTTCAGACTGTTCCGCCAGAACC CGA GCATGTGGTCCTTCAACCCCGAGACAAACATCCTGC TGAACGTGCCTCTGCACGGCACCATCCTGACAAGAC CTCTGCTCGAGAGCGAGCTGGTGATTGGCGCAGTGA TTCTGAGAGGCCATCTGAGGATCGCCGGACACCACC TGGGCAGATGCGACATCAAGGACCTTCCAAAGGAA ATCACCGTTGCCACCAGCCGGACCCTGTCCTACTAC AACTGGGCGCCAGCCAAAGAGTGCCCGGCGATAG CGGCTTTGCCGCCTACAGCAGATACCGCATCGGAAA TTACAAGCTCAACACCGACCACAGCAGCTCTTCTGA TAACATCGCCCTGCTGGTGCAGCGAAAACGGCGCgg aagcggaggaagcggagctactaacttcagcctgctgaagcaggtggagatgtg aggagaaccctggacctATGTTTCGTGTTCTTGGTGTCTGCTGC CTCTGGTCAGCTCCCAGTGTGTGAACCTGACCACCA GAACCCAGCTGCCACCTGCTTATACAAACTCCTTCA CTCGGGGGGTATACTACCCCGACAAGGTGTTCAAGT CTAGCGTGCTGCATTCTACACAAGACCTGTTCTCTGC CCTTCTTCAGCAACGTGACCTGGTTCACGCCATCC ACGTGTCTGGAACCAACGGAACCAAGAGATTTCGAC AACCCCGTGCTGCCTTTCAACGACGGCGTGTACTTC GCCAGCACCGAGAAGTCCAACATCATCAGAGGATG GATTTTCGGCACCACTGGACAGCAAAACCCAGA GCCTGCTGATCGTGAACAACGCCACCAACGTGGTG ATCAAGGTGTGCGAGTTCCAGTTCTGCAATGATCCC TTCCTGGGCGTGTACTACCACAAGAACAACAAGTCT TGGATGGAAAGCGAGTTCAGAGTGTATTCCAGCGC CAACAATTGCACCTTCGAGTACGTGAGCCAACCCTT TCTGATGGACCTTGAAGGCAAGCAGGGCAACTTCA AAAATCTGCGAGAATTTGTGTTCAAGAACATCGACG GATACTTCAAGATCTACTCTAAGCACACGCCAATCA ACCTGGTGAGAGATCTGCCCCAGGGCTTTAGCGCTT TGGAACCTCTGGTGGACCTGCCTATCGGAATCAACA TCACCAGATTTCAAACTCTCCTGGCCCTGCACAGAT CTTATCTGACCCCTGGGGACAGTAGTAGCGGCTGGA CAGCCGGCGCCGCCGCTACTACGTGGGATACCTGC AGCCTAGAACATTCCTGCTGAAGTACAATGAGAAC

		GGAACAATCACAGACGCCGTGGACTGCGCCCTGGA TCCTTTGAGCGAGACAAAGTGCACCCTGAAGTCGTT CACCGTCGAAAAAGGCATCTACCAGACCAGCAACT TCCGCGTGCAGCCTACGGAATCTATCGTGCGGTTCC CCAACATCACCAACCTGTGCCCTTTCGGCGAGGTGT TTAACGCTACAAGGTTTCGCCAGCGTGTATGCCTGGA ACAGAAAGAGAATCAGCAATTGCGTGGCCGATTAT AGCGTTCTGTACAACAGCGCTTCCTTCAGCACCTTC AAGTGCTACGGCGTGTCTCCAACCAAGCTGAACGA CCTCTGCTTCACCAATGTCTACGCTGACTCTTTCGTG ATTAGAGGCGATGAGGTTAGACAGATCGCACCTGG CCAGACCGGCAAAATCGCTGACTACAACCTACAAGC TGCCTGATGACTTCACAGGCTGTGTCAATTGCCTGGA ACTCAAATAACCTGGACTCTAAAGTGGGCGGCAAC TACAACCTACCTGTACCGGCTGTTCCGGAAGAGCAAT CTGAAACCTTTTGAGCGGGACATCTCTACAGAGATC TACCAGGCCGGCAGCACACCCTGCAACGGCGTTGA GGGCTTCAACTGCTACTTCCCTCTGCAGAGCTACGG CTTTCAGCCAACAAATGGAGTGGGCTACCAGCCGTA CAGAGTGGTGGTGCTGAGCTTCGAACTGCTGCATGC CCCAGCCACAGTGTGTGGACCTAAGAAGTCTACCA ACCTGGTGAAGAACAAGTGCCTGAACCTTAACTTTA ACGGCCTGACCGGCACAGGCGTGCTGACCGAATCC AACAAAAAGTTCCTGCCCTTCCAACAGTTCGGCAGA GACATCGCCGATACAACCGATGCCGTGCGGGACCC CCAGACCTTAGAAATCCTAGATATCACCCCGTGAG CTTCGGCGGAGTCTCTGTTATTACTCTGGCACCAA CACCAGCAACCAAGTGGCTGTTCTGTACCAAgcGTG AACTGCACCGAAGTGCTGTGGCTATCCACGCCGAT CAGCTGACCCCAACCTGGCGGGTGTATAGCACCGG CTCTAACGTGTTCCAGACCCGGGCTGGCTGCCTGAT CGGCGCCGAACACGTCAACAACCTCCTATGAATGTG ACATCCCCATCGGGGCTGGCATCTGCGCCAGTTACC AGACACAGACAAATAGCCCTAGACGGGCCAGAAGC GTGGCCTCCCAGAGTATCATTGCCTACACCATGAGC CTGGGCGCCGAGAACAGCGTGGCCTATTCTAACAAT AGCATCGCAATCCCTACCAACTTTACCATCTCTGTG ACAACCGAGATCCTGCCTGTGAGCATGACCAAAAC CAGCGTGGACTGCACGATGTACATCTGTGGCGACA GCACAGAATGCAGTAATCTGTTGCTGCAGTACGGCA GCTTTTGCACCCAGTTGAATAGAGCCCTGACCGGAA TCGCCGTAGAGCAGGACAAAAATACCCAGGAGGTG TTCGCCCAGGTGAAACAGATCTACAAGACACCTCCC ATTAAGGACTTCGGAGGTTTTAACTTCAGCCAGATC CTGCCCCGACCCTTCCAAGCCTAGCAAACGCTCCTTC ATCGAGGACCTGCTCTTCAACAAGGTGACACTGGCT GATGCCGGCTTCATCAAGCAGTACGGAGATTGTCTG GGAGACATCGCCGCTAGAGATCTGATCTGCGCCCA AAAGTTCAACGGCCTGACCGTGCTGCCTCCTCTGCT TACAGACGAGATGATCGCCAGTACACCAGCGCCC
--	--	--

		TGCTGGCTGGCACCATCACAAAGCGGCTGGACCTTCG GAGCCGGAGCCGCTCTGCAAATCCCCTTTGCCATGC AGATGGCCTACCGGTTCAACGGCATCGGCCGTGACA CAGAATGTGCTGTACGAGAACCAGAAGCTGATCGC TAACCAGTTTAACAGCGCTATCGGCAAGATCCAGG ACTCGCTGAGTAGCACCGCCTCTGCCCTGGGCAAGC TGCAGGACGTCGTGAACCAGAACGCCCAAGCCCTG AACACACTGGTGAAACAGCTGAGCAGCAACTTCGG CGCCATCAGCTCTGTGCTGAACGATATCCTGAGCAG ACTGGACAAGGTGGAAGCCGAGGTCCAGATCGACA GACTGATCACAGGAAGACTGCAGAGCCTGCAAACG TACGTGACACAGCAGCTGATCCGGGCAGCCGAAAT CCGGGCCAGCGCCAATCTGGCCGCTACCAAGATGA GCGAGTGCGTGTTAGGCCAGAGCAAGCGGGTGGAT TTCTGCGGTAAGGGATAACACCTGATGAGCTTTCCC CAGAGCGCTCCTCACGGCGTGGTGTCTTCTGCACGTG ACCTACGTTTCTGCCACGATGGCAAGGCCCACTTC GCCCCTGCTATCTGCCACGATGGCAAGGCCCACTTC CCTAGAGAGGGCGTTTTCTGTGTCTAACGGCACACAC TGGTTTGTGACCCAGAGAACTTCTACGAGCCTCAG ATCATCACCACAGACAACACCTTTGTGAGCGGCAAT TGCGACGTGGTGATCGGAATTGTTAATAATACCGTG TACGACCCTCTGCAGCCTGAGCTCGACAGCTTCAAG GAAGAGCTGGACAAGTACTTCAAGAACCACACCTC CCCAGATGTGGACCTGGGCGATATTTAGGCATCAA CGCCTCCGTCGTGAATATCCAGAAGGAGATCGACC GGCTCAACGAGGTGGCCAAGAACCTTAACGAGAGC CTGATCGACCTGCAGGAACCTGGGCAAATATGAGCA GTACATCAAGTGGCCTTGGTACATCTGGCTGGGCTT TATCGCAGGCCTGATCGCTATCGTGATGGTGACCAT TATGCTGTGTTGTATGACCAGCTGTTGTAGTTGTCTG AAGGGCTGCTGTTCTTGCGGCAGCTGCTGCAAGTTC GACGAAGACGACTCAGAGCCCCGTGCTGAAAGGCGT GAAGCTGCACTACACCCGAAAACGGCGCggaagcggag gaagcggagctactaactcagcctgctgaagcaggctggagatgtggaggagAAC cctggacctATGTATTCTTTTGTGTCCGAGGAAACCGGC ACACTGATCGTTAATAGCGTGCTGCTCTTCTGACC TTCTGTTGTTCTGCTGGTGACCCTGGCTATCCTG ACCGCCCTGAGACTGTGTGCCTACTGCTGCAACATC GTGAACGTGTCTCTGGTCAAGCCTAGCTTCTACGTG TACAGCCGGGTGAAGAACCTGAACAGCAGCAGAGT GCCCCGACCTGCTGGTGtaatccccccccctaactgttactggccgaag ccgcttggataaggccggtgtgcgtttgtctatatgttatttccaccatattgccgtcttt tggcaatgtgaggggcccgaaacctggccctgtcttcttgacgagcattcctagggggt cttccctctcgccaaaggaatgcaaggtctgttgatgtcgtgaagggaagcagttcc tctggaagctcttgaagacaacaacgtctgtagcgaccttgcaggcagcggaaac ccccacctggcgacaggtgcctctgcggccaaaagccacgtgtataagataacct gcaaaggcggcacaacccagtgccacgttgtgagttggatagttgtgaaagagtc aatggctctcctcaagcgtattcaacaaggggctgaaggatgccagaaggtaccc cattgtatgggatctgatctggggcctcggtgcacatgctttacatgtgttagtcgaggt
--	--	--

	taaaaaaacgtctaggccccccgaaccacggggacgtggttttcctttgaaaaacacg atgataatatggccacaacccatggaacaagagacttgcgcgcactctcacttttgag gaatgccccaaatgctctgctctacaataccgtaatggattttacctgctaaagtatgat gaagaatggtaccagaggaggtattgactgatggagaggatgatgtctttgatcccg aattagacatggaagtcgttttcgagttacaggggaagcggagctactaacttcagcctg ctgaagcaggctggagatgtggaggagaaccctggacctATGAGCGACA ACGGCCCTCAAAACCAGAGAAATGCCCTCGGATC ACATTTGGCGGACCTAGCGACAGCACCGGCAGCAA CCAGAATGGAGAAAGAAGCGGCGCCAGATCCAAGC AGCGGAGACCTCAGGGACTGCCCAACAACACCGCT AGCTGGTTCACCGCCCTGACCCAACACGGCAAGGA AGATCTGAAGTTCCCCAGAGGCCAGGGCGTGCCTAT CAACACAAACTCTTCTCCCGACGACCAGATCGGATA CTATAGACGGGCCACTCGGAGAATTTCGGGGCGGCG ACGGAAAAATGAAGGACCTTTCTCCAAGATGGTAC TTCTACTACCTCGGCACAGGCCCTGAGGCCGGCCTG CCTTACGGCGCCAACAAGGATGGCATCATCTGGGTC GCCACCGAGGGCGCCCTGAACACCCCTAAGGACCA CATCGGCACAAGAAACCCCGCTAACAACGCCGCAA TCGTGCTGCAGCTGCCTCAGGGCACCACCCTGCCCA AGGGCTTCTACGCCGAGGGCTCTAGAGGTGGCTCCC AGGCTTCTAGCCGCTCCTCCAGCCGCAGCAGAAACA GCAGCAGGAACAGCACCCCGGCAGCTCCCGGGGC ACCAGCCCCGCCAGAATGGCCGGAATGGCGGCGA TGCCGCCCTGGCCCTGCTCCTGCTGGACAGACTGAA TCAGCTGGAAAGCAAGATGAGCGGCAAAGGACAGC AGCAGCAAGGCCAGACCGTGACCAAGAAAAGCGCT GCTGAAGCCTCCAAGAAACCTAGACAAAAGCGGAC CGCCACAAAGGCCTACAACGTGACCCAAGCCTTTG GAAGAAGAGGCCCCGAGCAGACACAGGGCAATTTT GGCGACCAGGAGCTGATCCGGCAGGGAACCGACTA CAAGCACTGGCCTCAGATCGCCCAGTTTCGCCCTAG CGCCAGCGCCTTCTTCGGCATGAGCAGAATCGGCAT GGAAGTGACCCCTTCTGGCACCTGGCTGACCTACAC CGGCGCTATCAAGCTGGACGATAAGGATCCTAACTT CAAGGACCAAGTGATCCTGCTGAACAAGCATATCG ACGCCTATAAGACCTTTCCACCTACAGAGCCTAAGA AAGATAAGAAGAAGAAAGCCGACGAGACACAGGC CCTGCCTCAGAGACAGAAAAAGCAGCAGACAGTGA CACTGCTGCCAGCCGCTGACCTGGATGACTTCAGCA AGCAGCTGCAGCAGAGCATGTCTTCTGCTGATAGCA CCCAGGCCtaa GTGTCaAAGcGCGAGGAACTGTTCACCGGAGTttTGC CCATCCTGGTCGAGCTGGACGGCGATGTGAACGGC CACAAAGTTCAGCGTTTCTGGCGAGGGtgagctttgggctaag cgcaacattaaaccagtaccagaggtgaaaatactcaataatttgggtgtggacattgc tgctaatactgtgatctgggactacaaaagagatgtccagcacatatctactattgg tgtttgtctatgactgacatagccaagaaaccaactgaaacgatttgtgcaccactcac tgtctttttgatggttagttaggtcaagtagacttattagaaatgcccgtaatggtg ttctattacagaaggtagtgtfaaagggtttacaacctctgtaggtcccaacaagcta
--	--

		gtcttaatggagtcacattaattggagaagccgtaaaacacagttcaattattataaga aagttgatgggtgtgtccaacaattacgtgaacttactttactcagagtagaatttaca agaatttaaaccaggagtgcaaatggaaattgatttcttagaattagctatggatgaattc attgaacgggtataaattagaaggctatgccttcgaacatatcgtttatggagattttagtc ata
CoVEG9 expression cassette	41	ATGGCCGACAGCAACGGCACAATCACCGTGGAAGA GCTGAAGAACTGCTGGAACAGTGGAACCTGGTCA TCGGCTTCCTGTTTCTGACCTGGATCTGTCTGCTGCA GTTTCGCTTATGCCAATCGGAACAGATTCTGTACAT CATCAAGCTGATCTTCCTGTGGCTGCTGTGGCCTGT GACCCTGGCTTGCTTCGTGCTGGCCGCTGTGTACCG GATCAACTGGATCACAGGCGGAATCGCCATCGCCA TGGCCTGCCTGGTGGGCCTGATGTGGCTGAGCTACT TCATCGCTTCTTTCAGACTGTTCCGCCAGAACC CGGA GCATGTGGTCCTTCAACCCCGAGACAAACATCCTGC TGAACGTGCCTCTGCACGGCACCATCCTGACAAGAC CTCTGCTCGAGAGCGAGCTGGTGATTGGCGCAGTGA TTCTGAGAGGCCATCTGAGGATCGCCGGACACCACC TGGGCAGATGCGACATCAAGGACCTTCCAAAGGAA ATCACCGTTGCCACCAGCCGGACCCTGTCCTACTAC AACTGGGCGCCAGCCAAAGAGTGCCCGGCGATAG CGGCTTTGCCGCCTACAGCAGATACCGCATCGGAAA TTACAAGCTCAACACCGACCACAGCAGCTCTTCTGA TAACATCGCCCTGCTGGTGCAGCGAAAACGGCGCgg aagcggaggaagcggagctactaacttcagcctgctgaagcaggctggagatgtgg aggagaaccctggacctATGAGCGACAACGGCCCTCAAAAC CAGAGAAATGCCCCTCGGATCACATTTGGCGGACCT AGCGACAGCACCGGCAGCAACCAGAATGGAGAAAG AAGCGGCGCCAGATCCAAGCAGCGGAGACCTCAGG GACTGCCCAACAACACCGCTAGCTGGTTACCGCCC TGACCCAACACGGCAAGGAAGATCTGAAGTTCCCC AGAGGCCAGGGCGTGCCTATCAACACAACTCTTCT CCCGACGACCAGATCGGATACTATAGACGGGCCAC TCGGAGAATTTCGGGGCGGCGACGGAAAAATGAAGG ACCTTCTCCAAGATGGTACTTCTACTACCTCGGCA CAGGCCCTGAGGCCGGCCTGCCTTACGGCGCCAAC AAGGATGGCATCATCTGGGTGCCACCGAGGGCGC CCTGAACACCCCTAAGGACCACATCGGCACAAGAA ACCCCGCTAACAACGCCGCAATCGTGCTGCAGCTGC CTCAGGGCACCAACCCTGCCCAAGGGCTTCTACGCCG AGGGCTCTAGAGGTGGCTCCCAGGCTTCTAGCCGCT CCTCCAGCCGCAGCAGAAACAGCAGCAGGAACAGC ACCCCGGCAGCTCCCGGGGCACCAGCCCCGCCAG AATGGCCGGAATGGCGGCGATGCCGCCCTGGCCC TGCTCCTGCTGGACAGACTGAATCAGCTGGAAAGC AAGATGAGCGGCAAAGGACAGCAGCAGCAAGGCC AGACCGTGACCAAGAAAAGCGCTGCTGAAGCCTCC AAGAAACCTAGACAAAAGCGGACCGCCACAAAGGC CTACAACGTGACCCAAGCCTTTGGAAGAAGAGGCC

		CCGAGCAGACACAGGGCAATTTCCGGCGACCAGGAG CTGATCCGGCAGGGAACCGACTACAAGCACTGGCC TCAGATCGCCCAGTTCGCCCCTAGCGCCAGCGCCTT CTTCGGCATGAGCAGAATCGGCATGGAAGTGACCC CTTCTGGCACCTGGCTGACCTACACCGGCGCTATCA AGCTGGACGATAAGGATCCTAACTTCAAGGACCAA GTGATCCTGCTGAACAAGCATATCGACGCCTATAAG ACCTTTCCACCTACAGAGCCTAAGAAAAGATAAGAA GAAGAAAGCCGACGAGACACAGGCCCTGCCTCAGA GACAGAAAAAGCAGCAGACAGTGACACTGCTGCCA GCCGCTGACCTGGATGACTTCAGCAAGCAGCTGCA GCAGAGCATGTCTTCTGCTGATAGCACCCAGGCCCG AAAACGGCGCggaagcggaggaagcggagctactaacttcagcctgctga agcaggctggagatgtggaggagaaccctggacctATGTTTCGTGTTCTT GGTGCTGCTGCCTCTGGTCAGCTCCCAAGTGTGTGAA CCTGACCACCAGAACCCAGCTGCCACCTGCTTATAC AACTCCTTCACTCGGGGGGTATACTACCCCGACAA GGTGTTCAGATCTAGCGTGCTGCATTCTACACAAGA CCTGTTCTGCCCCTTCTTCAGCAACGTGACCTGGTTC CACGCCATCCACGTGTCTGGAACCAACGGAACCAA GAGATTTCGACAACCCCGTGCTGCCTTTCAACGACGG CGTGTACTTCGCCAGCACCGAGAAGTCCAACATCAT CAGAGGATGGATTTTCGGCACCACTGGACAGCA AAACCCAGAGCCTGCTGATCGTGAACAACGCCACC AACGTGGTGATCAAGGTGTGCGAGTTCAGTTCTGC AATGATCCCTTCCTGGGCGTGTACTACCACAAGAAC AACAAGTCTTGATGGAAAGCGAGTTCAGAGTGTA TTCCAGCGCCAACAATTGCACCTTCGAGTACGTGAG CCAACCCTTTCTGATGGACCTTGAAGGCAAGCAGGG CAACTTCAAAAATCTGCGAGAATTTGTGTTCAAGAA CATCGACGGATACTTCAAGATCTACTCTAAGCACAC GCCAATCAACCTGGTGAGAGATCTGCCCCAGGGCTT TAGCGCTTTGGAACCTCTGGTGGACCTGCCTATCGG AATCAACATCACCAGATTTCAAACCTCTCCTGGCCCT GCACAGATCTTATCTGACCCCTGGGGACAGTAGTAG CGGCTGGACAGCCGGCGCCGCCGCTACTACGTGG GATACCTGCAGCCTAGAACATTCCTGCTGAAGTACA ATGAGAACGGAACAATCACAGACGCCGTGGACTGC GCCCTGGATCCTTTGAGCGAGACAAAGTGACCCCTG AAGTCGTTACCGTCGAAAAAGGCATCTACCAGAC CAGCAACTTCCGCGTGACGCCTACGGAATCTATCGT GCGGTTCCCCAACATCACCACCTGTGCCCTTTCCGG CGAGGTGTTTAAACGCTACAAGGTTCCGCCAGCGTGTA TGCTTGGAACAGAAAGAGAATCAGCAATTGCGTGG CCGATTATAGCGTTCTGTACAACAGCGCTTCCTTCA GCACCTTCAAGTGCTACGGCGTGCTCCAACCAAGC TGAACGACCTCTGCTTCACCAATGTCTACGCTGACT CTTTCGTGATTAGAGGCGATGAGGTTAGACAGATCG CACCTGGCCAGACCGGCAAAATCGCTGACTACAAC TACAAGCTGCCTGATGACTTCACAGGCTGTGTCATT
--	--	---

		GCCTGGAACTCAAATAACCTGGACTCTAAAGTGGG CGGCAACTACAACCTACCTGTACCGGCTGTTCCGGAA GAGCAATCTGAAACCTTTTGAGCGGGACATCTCTAC AGAGATCTACCAGGCCGGCAGCACACCCTGCAACG GCGTTGAGGGCTTCAACTGCTACTTCCCTCTGCAGA GCTACGGCTTTCAGCCAACAAATGGAGTGGGCTACC AGCCGTACAGAGTGGTGGTGCTGAGCTTCGAACTGC TGCATGCCCCAGCCACAGTGTGTGGACCTAAGAAGT CTACCAACCTGGTGAAGAACAAGTGCCTGAACTTTA ACTTTAACGGCCTGACCGGCACAGGCGTGCTGACCG AATCCAACAAAAAGTTCCTGCCCTTCCAACAGTTCG GCAGAGACATCGCCGATACAACCGATGCCGTGCGG GACCCCCAGACCTTAGAAATCCTAGATATCACCCCG TGCAGCTTCGGCGGAGTCTCTGTTATTACTCTGGC ACCAACACCAGCAACCAAGTGGCTGTTCTGTACCAA ggcGTGAACTGCACCGAAGTGCCTGTGGCTATCCACG CCGATCAGCTGACCCCAACCTGGCGGGTGTATAGCA CCGGCTCTAACGTGTTCCAGACCCGGGCTGGCTGCC TGATCGGCGCCGAACACGTCAACAACCTCCTATGAAT GTGACATCCCCATCGGGGCTGGCATCTGCGCCAGTT ACCAGACACAGACAAATAGCCCTGGCAGCGCCAGC AGCGTGGCCTCCCAGAGTATCATTGCCTACCCATG AGCCTGGGCGCCGAGAACAGCGTGGCCTATTCTAA CAATAGCATCGCAATCCCTACCAACTTTACCATCTC TGTGACAACCGAGATCCTGCCTGTGAGCATGACCAA AACCAGCGTGGACTGCACGATGTACATCTGTGGCG ACAGCACAGAATGCAGTAATCTGTTGCTGCAGTACG GCAGCTTTTGCACCCAGTTGAATAGAGCCCTGACCG GAATCGCCGTAGAGCAGGACAAAAATACCCAGGAG GTGTTGCGCCAGGTGAAACAGATCTACAAGACACCT CCCATTAAAGGACTTCGGAGGTTTTAACTTCAGCCAG ATCCTGCCCCGACCCTTCCAAGCCTAGCAAACGCTCC TTCATCGAGGACCTGCTCTTCAACAAGGTGACACTG GCTGATGCCGGCTTCATCAAGCAGTACGGAGATTGT CTGGGAGACATCGCCGCTAGAGATCTGATCTGCGCC CAAAAGTTCAACGGCCTGACCGTGCTGCCTCCTCTG CTTACAGACGAGATGATCGCCAGTACACCAGCGC CCTGCTGGCTGGCACCATCACAAGCGGCTGGACCTT CGGAGCCGGAGCCGCTCTGCAAATCCCCTTTGCCAT GCAGATGGCCTACCGGTTCAACGGCATCGGCGTGA CACAGAAATGTGCTGTACGAGAACCAGAAGCTGATC GCTAACCAGTTTAAACAGCGCTATCGGCAAGATCCAG GACTCGCTGAGTAGCACCGCCTCTGCCCTGGGCAAG CTGCAGGACGTCGTGAACCAGAACGCCAAGCCCT GAACACACTGGTGAAACAGCTGAGCAGCAACTTCG GCGCCATCAGCTCTGTGCTGAACGATATCCTGAGCA GACTGGACCCTcccGAAGCCGAGGTCCAGATCGACA GACTGATCACAGGAAGACTGCAGAGCCTGCAAACG TACGTGACACAGCAGCTGATCCGGGCAGCCGAAAT CCGGGCCAGCGCCAATCTGGCCGCTACCAAGATGA
--	--	---

		GCGAGTGC GTGTTAGGCCAGAGCAAGCGGGTGGAT TTCTGCGGTAAGGGATACCACTGATGAGCTTTCCC CAGAGCGCTCCTCACGGCGTGGTGTCTGACGTG ACCTACGTTCTGCCCAGGAAAAGAACTTCACCACC GCCCCTGCTATCTGCCACGATGGCAAGGCCACTTC CCTAGAGAGGGCGTTTTTCGTGTCTAACGGCACACAC TGGTTTGTGACCCAGAGAACTTCTACGAGCCTCAG ATCATCACACAGACAACACCTTTGTGAGCGGCAAT TGCGACGTGGTGATCGGAATTGTTAATAATACCGTG TACGACCCTCTGCAGCCTGAGCTCGACAGCTTCAAG GAAGAGCTGGACAAGTACTTCAAGAACCACACCTC CCCAGATGTGGACCTGGGCGATATTTTCAGGCATCAA CGCCTCCGTCGTGAATATCCAGAAGGAGATCGACC GGCTCAACGAGGTGGCCAAGAACCTTAACGAGAGC CTGATCGACCTGCAGGAACCTGGGCAAATATGAGCA GTACATCAAGTGGCCTTGGTACATCTGGCTGGGCTT TATCGCAGGCCTGATCGCTATCGTGATGGTGACCAT TATGCTGTGTTGTATGACCAGCTGTTGTAGTTGTCTG AAGGGCTGCTGTTCTTGCGGCAGCTGCTGCAAGTTC GACGAAGACGACTCAGAGCCCGTGCTGAAAGGCGT GAAGCTGCACTACACCCGAAAACGGCGCggaagcggag gaagcggagctactaactcagcctgctgaagcaggctggagatgtggaggagaac cctggacctATGTATTCTTTTGTGTCCGAGGAAACCGGC ACACTGATCGTTAATAGCGTGCTGCTCTTCCTGGCC TTCGTGGTGTTCCTGCTGGTGACCCTGGCTATCCTG ACCGCCCTGAGACTGTGTGCCTACTGCTGCAACATC GTGAACGTGTCTCTGGTCAAGCCTAGCTTCTACGTG TACAGCCGGGTGAAGAACCTGAACAGCAGCAGAGT GCCCAGCTGCTGGTGaatccccccccctaacttgccgaag ccgcttgaataaggcgggtgtgctttgtctatatgtatttccaccatattgccgtcttt tggaatgtgagggcccgaaacctggccctgtcttcttgacgagcattcctaggggt cttccctctcgccaaaggaatgaaggtctgttgatgtcgtaaggaaagcagttcc tctggaagcttctgaagacaaacagctctgtagcgaccttgcaggcagcgggaac ccccacctggcgacaggtgcctctcgggccaaaagccacgtgataagataacct gcaaaggcggcacaacccagtgccacgttgtagttgtagttgtggaaagagtc aatggctctctcaagcgtattcaacaaggggctgaaggatgccagaaggtacc cattgtatgggatctgatctggggcctcggtgcacatgctttacatgttttagtcgaggt taaaaaacgtctagggccccgaaccacggggacgtggttttctttgaaaaacacg atgataatattggccacaacctggaacaagagacttgcgcgactctctcacttttag gaatgccccaaatgctctgctctacaataaccgtaattggattttacctgctaaagtatgat gaagaatggtaccagaggagttattgactgatggagaggatgatgtctttgatccg aattagacatggaagtcgttttcgagttacagtaa
CoVEG10 expression cassette	42	ATGGCCGACAGCAACGGCACAATCACCGTGGAAGA GCTGAAGAACTGCTGGAACAGTGGAACCTGGTCA TCGGCTTCCTGTTTCTGACCTGGATCTGTCTGCTGCA GTTCGCTTATGCCAATCGGAACAGATTCTGTACAT CATCAAGCTGATCTTCCTGTGGCTGCTGTGGCCTGT GACCCTGGCTTGCTTCGTGCTGGCCGCTGTGTACCG GATCAACTGGATCACAGGCGGAATCGCCATCGCCA

		TGGCCTGCCTGGTGGGCCTGATGTGGCTGAGCTACT TCATCGCTTCTTTCAGACTGTTCCGACAGAACCCGGA GCATGTGGTCCTTCAACCCCGAGACAAACATCCTGC TGAACGTGCCTCTGCACGGCACCATCCTGACAAGAC CTCTGCTCGAGAGCGAGCTGGTGATTGGCGCAGTGA TTCTGAGAGGCCATCTGAGGATCGCCGGACACCACC TGGGCAGATGCGACATCAAGGACCTTCCAAGGAA ATCACCGTTGCCACCAGCCGGACCCTGTCTACTAC AAACTGGGCGCCAGCCAAAGAGTGGCCGGCGATAG CGGCTTTGCCGCCTACAGCAGATACCGCATCGGAAA TTACAAGCTCAACACCGACCACAGCAGCTCTTCTGA TAACATCGCCCTGCTGGTGCAGCGAAAACGGCGCgg aagcggaggaagcggagctactaacttcagcctgctgaagcaggctggagatgtgg aggagaaccctggacctATGAGCGACAACGGCCCTCAAAAC CAGAGAAATGCCCCTCGGATCACATTTGGCGGACCT AGCGACAGCACCGGCAGCAACCAGAATGGAGAAAG AAGCGGCGCCAGATCCAAGCAGCGGAGACCTCAGG GACTGCCCAACAACACCGCTAGCTGGTTCACCGCCC TGACCCAACACGGCAAGGAAGATCTGAAGTTCCCC AGAGGCCAGGGCGTGCCTATCAACACAACTCTTCT CCCGACGACCAGATCGGATACTATAGACGGGCCAC TCGGAGAATTTCGGGGCGGCGACGGAAAAATGAAGG ACCTTTCTCCAAGATGGTACTTCTACTACCTCGGCA CAGGCCCTGAGGCCGGCCTGCCTTACGGCGCCAAC AAGGATGGCATCATCTGGGTGCGCACCGAGGGCGC CCTGAACACCCCTAAGGACCACATCGGCACAAGAA ACCCCGCTAACAACGCCGCAATCGTGCTGCAGCTGC CTCAGGGCACCAACCCTGCCCCAAGGGCTTCTACGCCG AGGGCTCTAGAGGTGGCTCCCAGGCTTCTAGCCGCT CCTCCAGCCGCAGCAGAAACAGCAGCAGGAACAGC ACCCCGGCAGCTCCCGGGGCACCAGCCCCGCCAG AATGGCCGGAATGGCGGCGATGCCGCCCTGGCCC TGCTCCTGCTGGACAGACTGAATCAGCTGGAAAGC AAGATGAGCGGCAAAGGACAGCAGCAGCAAGGCC AGACCGTGACCAAGAAAAGCGCTGCTGAAGCCTCC AAGAAACCTAGACAAAAGCGGACCGCCACAAAGGC CTACAACGTGACCCAAGCCTTTGGAAGAAGAGGGCC CCGAGCAGACACAGGGCAATTTGCGCGACCAGGAG CTGATCCGGCAGGGAACCGACTACAAGCACTGGCC TCAGATCGCCCAGTTTCGCCCCTAGCGCCAGCGCCTT CTTCGGCATGAGCAGAAATCGGCATGGAAGTGACCC CTTCTGGCACCTGGCTGACCTACACCGGCGCTATCA AGCTGGACGATAAGGATCCTAACTTCAAGGACCAA GTGATCCTGCTGAACAAGCATATCGACGCCTATAAG ACCTTTCCACCTACAGAGCCTAAGAAAGATAAGAA GAAGAAAGCCGACGAGACACAGGCCCTGCCTCAGA GACAGAAAAAGCAGCAGACAGTGACACTGCTGCCA GCCGCTGACCTGGATGACTTCAGCAAGCAGCTGCA GCAGAGCATGTCTTCTGCTGATAGCACCCAGGCCCG AAAACGGCGCggaagcggaggaagcggagctactaacttcagcctgctga
--	--	---

	agcaggctggagatgtggaggagaaccctggacctATGTTCTGTTTCCT GGTGCTGCTGCCTCTGGTCAGCTCCCAGTGTGTGAA CCTGACCACCAGAACCCAGCTGCCACCTGCTTATAC AAACTCCTTCACTCGGGGGGTATACTACCCCGACAA GGTGTTAGATCTAGCGTGCTGCATTCTACACAAGA CCTGTTCTGCCCCTTCTTCAGCAACGTGACCTGGTTC CACGCCATCCACGTGTCTGGAACCAACGGAACCAA GAGATTCGACAACCCCGTGCTGCCTTTCAACGACGG CGTGTAATTTCGCCAGCACCGAGAAGTCCAACATCAT CAGAGGATGGATTTTCGGCACCACTGGACAGCA AAACCCAGAGCCTGCTGATCGTGAACAACGCCACC AACGTGGTGATCAAGGTGTGCGAGTTCAGTTCTGC AATGATCCCTTCCTGGGCGTGTACTACCACAAGAAC AACAAGTCTTGATGGAAAGCGAGTTCAGAGTGTA TTCCAGCGCCAACAATTGCACCTTCGAGTACGTGAG CCAACCCTTTCTGATGGACCTTGAAGGCAAGCAGGG CAACTTCAAAAATCTGCGAGAATTTGTGTTCAAGAA CATCGACGGATACTTCAAGATCTACTCTAAGCACAC GCCAATCAACCTGGTGAGAGATCTGCCCCAGGGCTT TAGCGCTTTGGAACCTCTGGTGGACCTGCCTATCGG AATCAACATCACCAGATTTCAAACCTCTCCTGGCCCT GCACAGATCTTATCTGACCCCTGGGGACAGTAGTAG CGGCTGGACAGCCGGCGCCGCCGCTACTACGTGG GATACCTGCAGCCTAGAACATTCCTGCTGAAGTACA ATGAGAACGGAACAATCACAGACGCCGTGGACTGC GCCCTGGATCCTTTGAGCGAGACAAAGTGCACCCTG AAGTCGTTACCGTCGAAAAAGGCATCTACCAGAC CAGCAACTTCCGCGTGACGCTACGGAATCTATCGT GCGGTTCCCCAACATCACC AACCTGTGCCCTTTCCG CGAGGTGTTTAACGCTACAAGGTTCCGCCAGCGTGTA TGCTTGGAACAGAAAGAGAATCAGCAATTGCGTGG CCGATTATAGCGTTCTGTACAACAGCGCTTCCTTCA GCACCTTCAAGTGCTACGGCGTGCTCCAACCAAGC TGAACGACCTCTGCTTCACCAATGTCTACGCTGACT CTTTCGTGATTAGAGGCGATGAGGTTAGACAGATCG CACCTGGCCAGACCGGCAAAATCGCTGACTACAAC TACAAGCTGCCTGATGACTTCACAGGCTGTGTGATT GCCTGGAACCTCAAATAACCTGGACTCTAAAGTGGG CGGCAACTACAACCTGTACCGGCTGTTCCGGAA GAGCAATCTGAAACCTTTTGAGCGGGACATCTCTAC AGAGATCTACCAGGCCGGCAGCACACCCTGCAACG GCGTTGAGGGCTTCAACTGCTACTTCCCTCTGCAGA GCTACGGCTTTCAGCCAACAAATGGAGTGGGCTACC AGCCGTACAGAGTGGTGGTGCTGAGCTTCGAAGTGC TGATGCCCCAGCCACAGTGTGTGGACCTAAGAAGT CTACCAACCTGGTGAAGAACAAGTGCGTGAACCTTA ACTTTAACGGCCTGACCGGCACAGGCGTGCTGACCG AATCCAACAAAAAGTTCCTGCCCTTCCAACAGTTCG GCAGAGACATCGCCGATACAACCGATGCCGTGCGG GACCCCCAGACCTTAGAAATCCTAGATATCACCCCG
--	--

		TGCAGCTTCGGCGGAGTCTCTGTTATTACTCCTGGC ACCAACACCAGCAACCAAGTGGCTGTTCTGTACCAA ggcGTGAACTGCACCGAAGTGCCTGTGGCTATCCACG CCGATCAGCTGACCCCAACCTGGCGGGTGTATAGCA CCGGCTCTAACGTGTTCCAGACCCGGGCTGGCTGCC TGATCGGCGCCGAACACGTCAACAACCTCTATGAAT GTGACATCCCCATCGGGGCTGGCATCTGCGCCAGTT ACCAGACACAGACAAATAGCCCTGGCAGCGCCAGC AGCGTGGCCTCCCAGAGTATCATTGCCTACACCATG AGCCTGGGCGCCGAGAACAGCGTGGCCTATTCTAA CAATAGCATCGCAATCCCTACCAACTTTACCATCTC TGTGACAACCGAGATCCTGCCTGTGAGCATGACCAA AACCAGCGTGGACTGCACGATGTACATCTGTGGCG ACAGCACAGAATGCAGTAATCTGTTGCTGCAGTACG GCAGCTTTTGCACCCAGTTGAATAGAGCCCTGACCG GAATCGCCGTAGAGCAGGACAAAAATACCCAGGAG GTGTTGCGCCAGGTGAAACAGATCTACAAGACACCT CCCATTAAGGACTTCGGAGGTTTTAACTTCAGCCAG ATCCTGCCCCGACCCTTCCAAGCCTAGCAAACGCTCC TTCATCGAGGACCTGCTCTTCAACAAGGTGACACTG GCTGATGCCGGCTTCATCAAGCAGTACGGAGATTGT CTGGGAGACATCGCCGCTAGAGATCTGATCTGCGCC CAAAAGTTCAACGGCCTGACCGTGCTGCCTCCTCTG CTTACAGACGAGATGATCGCCAGTACACCAGCGC CCTGCTGGCTGGCACCATCACAAGCGGCTGGACCTT CGGAGCCGGAGCCGCTCTGCAAATCCCCTTTGCCAT GCAGATGGCCTACCGGTTCAACGGCATCGGCGTGA CACAGAATGTGCTGTACGAGAACCAGAAGCTGATC GCTAACCAGTTTAAACAGCGCTATCGGCAAGATCCAG GACTCGCTGAGTAGCACCGCCTCTGCCCTGGGCAAG CTGCAGGACGTGCTGAACCAGAACGCCCAAGCCCT GAACACACTGGTGAAACAGCTGAGCAGCAACTTCG GCGCCATCAGCTCTGTGCTGAACGATATCCTGAGCA GACTGGACCCTcccGAAGCCGAGGTCCAGATCGACA GACTGATCACAGGAAGACTGCAGAGCCTGCAAACG TACGTGACACAGCAGCTGATCCGGGCAGCCGAAAT CCGGGCCAGCGCCAATCTGGCCGCTACCAAGATGA GCGAGTGCGTGTTAGGCCAGAGCAAGCGGGTGGAT TTCTGCGGTAAGGGATACCACCTGATGAGCTTTCCC CAGAGCGCTCCTCACGGCGTGGTGTCTTCTGCACGTG ACCTACGTTTCTGCCAGGAAAAGAACTTCAACCACC GCCCCTGCTATCTGCCACGATGGCAAGGCCCACTTC CCTAGAGAGGGCGTTTTTCGTGTCTAACGGCACACAC TGGTTTGTGACCCAGAGAACTTCTACGAGCCTCAG ATCATCACCACAGACAACACCTTTGTGAGCGGCAAT TGCGACGTGGTGATCGGAATTGTTAATAATACCGTG TACGACCCTCTGCAGCCTGAGCTCGACAGCTTCAAG GAAGAGCTGGACAAGTACTTCAAGAACCACACCTC CCCAGATGTGGACCTGGGCGATATTTCAAGGCATCAA CGCCTCCGTCGTGAATATCCAGAAGGAGATCGACC
--	--	--

		GGCTCAACGAGGTGGCCAAGAACCTTAACGAGAGC CTGATCGACCTGCAGGAAGTGGGCAAATATGAGCA GTACATCAAGTGGCCTTGGTACATCTGGCTGGGCTT TATCGCAGGCCTGATCGCTATCGTGATGGTGACCAT TATGCTGTGTTGTATGACCAGCTGTTGTAGTTGTCTG AAGGGCTGCTGTTCTTGCGGCAGCTGCTGCAAGTTC GACGAAGACGACTCAGAGCCCCGTGCTGAAAGGCGT GAAGCTGCACTACACCCGAAAACGGCGCggaagcggag gaagcggagctactaacttcagcctgctgaagcaggctggagatgtggaggagaac cctggacctATGTATTCTTTTGTGTCCGAGGAAACCGGC ACACTGATCGTTAATAGCGTGCTGCTCTTCCTGGCC TTCGTGGTGTTCCTGCTGGTGACCCTGGCTATCCTG ACCGCCCTGAGACTGTGTGCCTACTGCTGCAACATC GTGAACGTGTCTCTGGTCAAGCCTAGCTTCTACGTG TACAGCCGGGTGAAGAACCTGAACAGCAGCAGAGT GCCCACCTGCTGGTGtaatccccccccctaactgactggccgaag ccgcttggaataaggccggtgtgcgtttgtctatatgtatttccaccatattgccgtctt tggcaatgtgagggcccgaaacctggccctgtcttctgacgagcattcctaggggt ctttccctctcgccaaaggaatgcaaggctgttgatgtcgtgaaggaaagcagtcc tctggaagcttctgaagacaaacagctgttagcgacctttgaggcagcgggaac ccccacctggcgacaggtgctctcgcccaaaagccacgtgtataagataacct gcaaaggcggcacaacccagtgccacgttgtagttgtagatgttggaagagtc aatggctctctcaagcgtattcaacaaggggctgaaggatgccagaaggtacc cattgtatgggatctgatctggggcctcggtgcacatgctttacatgtgttagtcgaggt taaaaaacgtctagggccccgaaccacgggacgtggttttctttgaaaaacacg atgataatatggccacaacctggaacaagagactgcgcgactctctcacttttgag gaatgcccaaatgctctgctctacaataaccgtaattgatttacctgctaaagtatgat gaagaatggtaccagaggagtattgactgatggagaggatgatgtctttgatcccg aattagacatggaagtcgtttcgagttacagtaaGTGTCaAAGcGCGAG GAACTGTTCACCGGAGTtTGCCCATCCTGGTCGAGC TGGACGGCGATGTGAACGGCCACAAGTTCAGCGTTT CTGGCGAGGGtgagctttgggctaagcgcaacattaaaccagtaccagagg tgaaaatactcaataattgggtgtggacattgctgctaatactgtgatctgggactaca aaagagatgtccagcacatatatctactattgggtttgttctatgactgacatagccaa gaaaccaactgaaacgatttgtgcaccactcactgtctttttgatggtagagttgatggt caagtagacttattagaatgcccgaatgggtgtcttattacagaaggtagtgttaaag gtttacaacctctgtaggtccaaacaagctagtcttaattggagtcacattaattggag aagccgtaaaaacacagttcaattattataagaaagttgatggtgttccaacaattac ctgaaacttactttactcagagtagaaatttacaagaatttaaacccaggagtcfaatgg aaattgatttcttagaattagctatggatgaattcattgaacggtataaattagaaggctat gccttcgaacatatcgtttatggagattttagtcata
CoVEG11 expression cassette	43	aaaaaaaaATTAAAGGTTTATACCTTCCCAGGTAACA AACCAACCAACTTTCGATCTCTTGTAGATCTGTTCTC TAAACGAACCTTTAAAATCTGTGTGGCTGTCACCTCGG CTGCATGCTTAGTGCACCTCACGCAGTATAATTAATA ACTAATTACTGTCGTTGACAGGACACGAGTAACTCG TCTATCTTCTGCAGGCTGCTTACGGTTTCGTCCGTGT TGCAGCCGATCATCAGCACATCTAGGTTTCGTCCGG GTGTGACCGAAAGGTAAgATGGCCGACAGCAACGG

		CACAATCACCGTGGAAGAGCTGAAGAACTGCTGG AACAGTGGAACCTGGTCATCGGCTTCCTGTTTCTGA CCTGGATCTGTCTGCTGCAGTTCGCTTATGCCAATC GGAACAGATTCTGTACATCATCAAGCTGATCTTCC TGTGGCTGCTGTGGCCTGTGACCCTGGCTTGCTTCG TGCTGGCCGCTGTGTACCGGATCAACTGGATCACAG GCGGAATCGCCATCGCCATGGCCTGCCTGGTGGGCC TGATGTGGCTGAGCTACTTCATCGCTTCTTTTCAGACT GTTTCGCCAGAACCCGGAGCATGTGGTCCTTCAACCC CGAGACAAACATCCTGCTGAACGTGCCTCTGCACGG CACCATCCTGACAAGACCTCTGCTCGAGAGCGAGCT GGTGATTGGCGCAGTGATTCTGAGAGGCCATCTGAG GATCGCCGGACACCACCTGGGCAGATGCGACATCA AGGACCTTCCAAAGGAAATCACCGTTGCCACCAGC CGGACCCTGTCCTACTACAACTGGGCGCCAGCCAA AGAGTGGCCGGCGATAGCGGCTTTGCCGCCTACAG CAGATACCGCATCGGAAATTACAAGCTCAACACCG ACCACAGCAGCTCTTCTGATAACATCGCCCTGCTGG TGCAGCGAAAACGGCGCggaagcggaggaagcggagctactaact tcagcctgctgaagcaggtgagatgtggaggagaacctggacctATGAGC GACAAACGGCCCTCAAAACCAGAGAAATGCCCTCG GATCACATTTGGCGGACCTAGCGACAGCACCGGCA GCAACCAGAAATGGAGAAAGAAGCGGCGCCAGATCC AAGCAGCGGAGACCTCAGGGACTGCCCAACAACAC CGTAGCTGGTTACCGCCCTGACCCAACACGGCAA GGAAGATCTGAAGTTCCCCAGAGGCCAGGGCGTGC CTATCAACACAACTCTTCTCCCGACGACCAGATCG GATACTATAGACGGGCCACTCGGAGAATTCGGGGC GGCGACGGAAAAATGAAGGACCTTTCTCCAAGATG GTACTTCTACTACCTCGGCACAGGCCCTGAGGCCGG CCTGCCTTACGGCGCCAACAAGGATGGCATCATCTG GGTCGCCACCGAGGGCGCCCTGAACACCCCTAAGG ACCACATCGGCACAAGAAACCCCGCTAACAACGCC GCAATCGTGCTGCAGCTGCCTCAGGGCACCACCCTG CCCAAGGGCTTCTACGCCGAGGGCTCTAGAGGTGG CTCCCAGGCTTCTAGCCGCTCCTCCAGCCGCAGCAG AAACAGCAGCAGGAACAGCACCCCCGGCAGCTCCC GGGGCACCAGCCCCGCCAGAATGGCCGGAATGGC GGCGATGCCGCCCTGGCCCTGCTCCTGCTGGACAGA CTGAATCAGCTGGAAAGCAAGATGAGCGGCAAAGG ACAGCAGCAGCAAGGCCAGACCGTGACCAAGAAAA GCGCTGCTGAAGCCTCCAAGAAACCTAGACAAAAAG CGGACCGCCACAAAGGCCTACAACGTGACCCAAGC CTTTGGAAGAAGAGGCCCCGAGCAGACACAGGGCA ATTTGCGGACCAAGGAGCTGATCCGGCAGGGAACC GACTACAAGCACTGGCCTCAGATCGCCAGTTTCGCC CCTAGCGCCAGCGCCTTCTTCGGCATGAGCAGAATC GGCATGGAAGTGACCCCTTCTGGCACCTGGCTGACC TACACCGGCGCTATCAAGCTGGACGATAAGGATCCT AACTTCAAGGACCAAGTGATCCTGCTGAACAAGCA
--	--	--

		TATCGACGCCTATAAGACCTTTCCACCTACAGAGCC TAAGAAAGATAAGAAGAAGAAAGCCGACGAGACA CAGGCCCTGCCTCAGAGACAGAAAAAGCAGCAGAC AGTGACACTGCTGCCAGCCGCTGACCTGGATGACTT CAGCAAGCAGCTGCAGCAGAGCATGTCTTCTGCTGA TAGCACCCAGGCCCGAAAACGGCGCggaagcggaggaagc ggagctactaacttcagcctgctgaagcaggctggagatgtggaggagaacctgg acctATGTTTCGTGTTCTTGGTGCTGCTGCCTCTGGTCA GCTCCCAGTGTGTGAACCTGACCACCAGAACCCAGC TGCCACCTGCTTATACAACTCCTTCACTCGGGGGG TATACTACCCCGACAAGGTGTTTCAAGATCTAGCGTGC TGCATTCTACACAAGACCTGTTTCTGCCCTTCTTCA CAACGTGACCTGGTTCACGCCATCCACGTGTCTGG AACCAACGGAACCAAGAGATTGACAACCCCGTGC TGCCTTTCAACGACGGCGTGTACTTCGCCAGCACCG AGAAGTCCAACATCATCAGAGGATGGATTTTCGGC ACCACACTGGACAGCAAAACCCAGAGCCTGCTGAT CGTGAACAACGCCACCAACGTGGTGATCAAGGTGT GCGAGTTCCAGTTCTGCAATGATCCCTTCTGGGCG TGTAATAACACAAGAACAACAAGTCTTGGATGGAA AGCGAGTTCAGAGTGTATTCCAGCGCCAACAATTGC ACCTTCGAGTACGTGAGCCAACCCCTTTCTGATGGAC CTTGAAGGCAAGCAGGGCAACTTCAAAAATCTGCG AGAATTTGTGTTCAAGAACATCGACGGATACTTCAA GATCTACTCTAAGCACACGCCAATCAACCTGGTGAG AGATCTGCCCCAGGGCTTTAGCGCTTTGGAACCTCT GGTGGACCTGCCTATCGGAATCAACATCACCAGATT TCAAACCTCTCCTGGCCCTGCACAGATCTTATCTGAC CCCTGGGGACAGTAGTAGCGGCTGGACAGCCGGCG CCGCCGCCTACTACGTGGGATACCTGCAGCCTAGAA CATTCCTGCTGAAGTACAATGAGAACGGAACAATC ACAGACGCCGTGGACTGCGCCCTGGATCCTTTGAGC GAGACAAAGTGCACCTGAAGTCGTTACCGTTCGA AAAAGGCATCTACCAGACCAGCAACTTCCGCGTGC AGCCTACGGAATCTATCGTGCGGTTCCCCAACATCA CCAACCTGTGCCCTTTCGGCGAGGTGTTAACGCTA CAAGGTTTCGCCAGCGTGTATGCCTGGAACAGAAAG AGAATCAGCAATTGCGTGGCCGATTATAGCGTTCTG TACAACAGCGCTTCCTTCAAGTGTCTAC GGCGTGTCTCCAACCAAGCTGAACGACCTCTGCTTC ACCAATGTCTACGCTGACTCTTTCGTGATTAGAGGC GATGAGGTTAGACAGATCGCACCTGGCCAGACCGG CAAAATCGCTGACTACAACCTACAAGCTGCCTGATGA CTTCACAGGCTGTGTCATTGCCTGGAACCTCAAATAA CCTGGACTCTAAAGTGGGCGGCAACTACAACCTACCT GTACCGGCTGTTCCGGAAGAGCAATCTGAAACCTTT TGAGCGGGACATCTCTACAGAGATCTACCAGGCCG GCAGCACACCCTGCAACGGCGTTGAGGGCTTCAACT GCTACTTCCCTCTGCAGAGCTACGGCTTTCAGCCAA CAAATGGAGTGGGCTACCAGCCGTACAGAGTGGTG
--	--	---

		GTGCTGAGCTTCGAACTGCTGCATGCCCCAGCCACA GTGTGTGGACCTAAGAAGTCTACCAACCTGGTGAA GAACAAGTGCCTGAACTTTAACTTTAACGGCCTGAC CGGCACAGGCGTGCTGACCGAATCCAACAAAAAGT TCCTGCCCTTCCAACAGTTCGGCAGAGACATCGCCG ATACAACCGATGCCGTGCGGGACCCCCAGACCTTA GAAATCCTAGATATCACCCCGTGCAGCTTCGGCGGA GTCTCTGTTATTACTCCTGGCACCAACACCAGCAAC CAAGTGGCTGTTCTGTACCAAggcGTGAACTGCACCG AAGTGCCTGTGGCTATCCACGCCGATCAGCTGACCC CAACCTGGCGGGTGTATAGCACCGGCTCTAACGTGT TCCAGACCCGGGCTGGCTGCCTGATCGGCGCCGAAC ACGTCAACAACCTCCTATGAATGTGACATCCCCATCG GGGCTGGCATCTGCGCCAGTTACCAGACACAGACA AATAGCCCTGGCAGCGCCAGCAGCGTGGCCTCCCA GAGTATCATTGCCTACACCATGAGCCTGGGCGCCGA GAACAGCGTGGCCTATTCTAACAATAGCATCGCAAT CCCTACCAACTTTACCATCTCTGTGACAACCGAGAT CCTGCCTGTGAGCATGACCAAAACCAGCGTGGACT GCACGATGTACATCTGTGGCGACAGCACAGAATGC AGTAATCTGTTGCTGCAGTACGGCAGCTTTTGCACC CAGTTGAATAGAGCCCTGACCGGAATCGCCGTAGA GCAGGACAAAAATACCCAGGAGGTGTTGCCCCAGG TGAAACAGATCTACAAGACACCTCCCATTAAAGACT TCGGAGGTTTTAACTTCAGCCAGATCCTGCCCCGACC CTTCCAAGCCTAGCAAACGCTCCTTCATCGAGGACC TGCTCTTCAACAAGGTGACACTGGCTGATGCCGGCT TCATCAAGCAGTACGGAGATTGTCTGGGAGACATC GCCGCTAGAGATCTGATCTGCGCCCCAAAAGTTCAAC GGCCTGACCGTGCTGCCTCCTCTGCTTACAGACGAG ATGATCGCCCAGTACACCAGCGCCCTGCTGGCTGGC ACCATCACAAAGCGGCTGGACCTTCGGAGCCGGAGC CGCTCTGCAAATCCCCTTTGCCATGCAGATGGCCTA CCGGTTCAACGGCATCGGCGTGACACAGAATGTGCT GTACGAGAACCAGAAGCTGATCGCTAACCAGTTTA ACAGCGCTATCGGCAAGATCCAGGACTCGCTGAGT AGCACCGCCTCTGCCCTGGGCAAGCTGCAGGACGTC GTGAACCAGAACGCCCAAGCCCTGAACACACTGGT GAAACAGCTGAGCAGCAACTTCGGCGCCATCAGCT CTGTGCTGAACGATATCCTGAGCAGACTGGACCCTc ccGAAGCCGAGGTCCAGATCGACAGACTGATCACAG GAAGACTGCAGAGCCTGCAAACGTACGTGACACAG CAGCTGATCCGGGCAGCCGAAATCCGGGGCCAGCGC CAATCTGGCCGCTACCAAGATGAGCGAGTGCGTGTT AGGCCAGAGCAAGCGGGTGGATTTCTGCGGTAAGG GATACCACCTGATGAGCTTTCCCCAGAGCGCTCCTC ACGGCGTGGTGTCTTCTGCACGTGACCTACGTTCTCTG CCCAGGAAAAGAACTTCACCACCGCCCCTGCTATCT GCCACGATGGCAAGGCCCCACTTCCCTAGAGAGGGC GTTTTCTGTGCTAACGGCACACACTGGTTTGTGACC
--	--	--

		CAGAGAACTTCTACGAGCCTCAGATCATCACCACA GACAACACCTTTGTGAGCGGCAATTGCGACGTGGTG ATCGGAATTGTTAATAATACCGTGTACGACCCTCTG CAGCCTGAGCTCGACAGCTTCAAGGAAGAGCTGGA CAAGTACTTCAAGAACCACACCTCCCCAGATGTGGA CCTGGGCGATATTTTCAGGCATCAACGCCTCCGTCGT GAATATCCAGAAGGAGATCGACCGGCTCAACGAGG TGGCCAAGAACCTTAACGAGAGCCTGATCGACCTG CAGGAACTGGGCAAATATGAGCAGTACATCAAGTG GCCTTGGTACATCTGGCTGGGCTTTATCGCAGGCCT GATCGCTATCGTGATGGTGACCATTATGCTGTGTTG TATGACCAGCTGTTGTAGTTGTCTGAAGGGCTGCTG TTCTTGCGGCAGCTGCTGCAAGTTCGACGAAGACGA CTCAGAGCCCCTGCTGAAAGGCGTGAAGCTGCACT ACACCCGAAAACGGCGCggaagcggaggaagcggagtactaact tcagcctgctgaagcaggctggagatgtggaggagaacctggacctATGTAT TCTTTTGTGTCCGAGGAAACCGGCACACTGATCGTT AATAGCGTGCTGCTCTTCTGGCCTTCGTGGTGTTC TGCTGGTGACCCTGGCTATCCTGACCGCCCTGAGAC TGTGTGCCTACTGCTGCAACATCGTGAACGTGTCTC TGGTCAAGCCTAGCTTCTACGTGTACAGCCGGGTGA AGAACCTGAACAGCAGCAGAGTGCCCCGACCTGCTG GTGaatccccccccctaactgtactggccgaagccgcttggaataagccgggtg tgcgtttgtctatatgttattttccaccatattgccgtcttttgcaatgtgagggcccgga aacctggccctgtcttctgacgagcattcctaggggtctttccctctcgccaaagga atgcaaggtctgtgaatgtcgtgaaggaagcagttcctctggaagcttctgaagaca aacaacgtctgtagcgacccttgcaggcagcggaaacccccacctggcgacaggt gcctctgcggccaaaagccacgtgtataagatacacctgcaaaagcgggcacaaacc cagtgccacgttgtagttgtagttgtgaaagagtcfaatggctctctcaagcgt attcaacaaggggctgaaggatcccagaaggtacccattgtatgggatctgatctg gggcctcggtgcacatgctttacatgtgttagtcgagggttaaaaaacgtctaggccc ccgaaccacggggacgtggttttctttgaaaaacacgatgataatatggccacaac catggaacaagagacttgcgcgactctctcacttttgaggaatgccccaaatgctctg ctctacaataccgtaattggtttacctgctaaagtatgatgaagaatggtaccagag gagttattgactgatggagaggatgatgtctttgatcccgaattagacatggaagtcgtt ttcagttacagtaaGTGTtacctgttaattgtagcatttgagctttgggctaagcgca acattaaaccagtaccagaggtgaaaatactcaataattgggtgtggacattgctgct aatactgtgatctgggactacaaaagagatgctccagcacatatatctactattgggttt gttctatgactgacatagccaagaaccaactgaaacgattgtgcaccactcactgtc tttttgatggtagagttgatggtcaagtagactatttagaaatgccgtaattgggtgttctt attacagaaggtagtgttaaagggttacaacctctgtagggtcccaacaagctagtctt aatggagtcacattaattggagaagccgtaaaaacacagttcaattattataagaaagtt gatgggtgtgtccaacaattacctgaaacttactttactcagagtagaaattacaagaat ttaaaccaggagtcfaatggaaattgatttcttagaattagctatggatgaattcattga acggtataaattagaaggctatgccttcgaacatatcgtttatggagatttttagtcatagt cagttaggtggtGCGAaattgtgtgtt
CoVEG12 expression cassette	44	ATGGCCGACAGCAACGGCACAATCACCGTGGAAGA GCTGAAGAACTGCTGGAACAGTGGAACCTGGTCA TCGGCTTCCTGTTTCTGACCTGGATCTGTCTGCTGCA

		GTTTCGCTTATGCCAATCGGAACAGATTCTGTACAT CATCAAGCTGATCTTCCTGTGGCTGCTGTGGCCTGT GACCCTGGCTTGCTTCGTGCTGGCCGCTGTGTACCG GATCAACTGGATCACAGGCGGAATCGCCATCGCCA TGGCCTGCCTGGTGGGCCTGATGTGGCTGAGCTACT TCATCGCTTCTTTCAGACTGTTTCGCCAGAACCCGGA GCATGTGGTTCCTTCAACCCCGAGACAAACATCCTGC TGAACGTGCCTCTGCACGGCACCATCCTGACAAGAC CTCTGCTCGAGAGCGAGCTGGTGATTGGCGCAGTGA TTCTGAGAGGCCATCTGAGGATCGCCGACACCACC TGGGCAGATGCGACATCAAGGACCTTCCAAAGGAA ATCACCGTTGCCACCAGCCGGACCCTGTCTACTAC AAACTGGGCGCCAGCCAAAGAGTGGCCGGCGATAG CGGCTTTGCCGCCTACAGCAGATACCGCATCGGAAA TTACAAGCTCAACACCGACCACAGCAGCTCTTCTGA TAACATCGCCCTGCTGGTGCAGCGAAAACGGCGCgg aagcggaggaagcggagctactaacttcagcctgctgaagcaggctggagatgtgg aggagaaccctggacctATGAGCGACAACGGCCCTCAAAAC CAGAGAAATGCCCCTCGGATCACATTTGGCGGACCT AGCGACAGCACCGGCAGCAACCAGAATGGAGAAAG AAGCGGCGCCAGATCCAAGCAGCGGAGACCTCAGG GACTGCCCAACAACACCGCTAGCTGGTTCACCGCCC TGACCCAACACGGCAAGGAAGATCTGAAGTTCCCC AGAGGCCAGGGCGTGCCTATCAACACAACTCTTCT CCCGACGACCAGATCGGATACTATAGACGGGCCAC TCGGAGAATTTCGGGGCGGCGACGGAAAAATGAAGG ACCTTTCTCCAAGATGGTACTTCTACTACCTCGGCA CAGGCCCTGAGGCCGGCCTGCCTTACGGCGCCAAC AAGGATGGCATCATCTGGGTGCGCACCGAGGGCGC CCTGAACACCCCTAAGGACCACATCGGCACAAGAA ACCCCGCTAACAACGCCGCAATCGTGCTGCAGCTGC CTCAGGGCACCACCCTGCCCCAAGGGCTTCTACGCCG AGGGCTCTAGAGGTGGCTCCCAGGCTTCTAGCCGCT CCTCCAGCCGCAGCAGAAACAGCAGCAGGAACAGC ACCCCGGCAGCTCCCGGGGCACCAGCCCCGCCAG AATGGCCGGAATGGCGGCGATGCCGCCCTGGCCC TGCTCCTGCTGGACAGACTGAATCAGCTGGAAAGC AAGATGAGCGGCAAAGGACAGCAGCAGCAAGGCC AGACCGTGACCAAGAAAAGCGCTGCTGAAGCCTCC AAGAAACCTAGACAAAAGCGGACCGCCACAAAGGC CTACAACGTGACCCAAGCCTTTGGAAGAAGAGGCC CCGAGCAGACACAGGGCAATTTCCGGCGACCAGGAG CTGATCCGGCAGGGAACCGACTACAAGCACTGGCC TCAGATCGCCCAGTTTCGCCCCTAGCGCCAGCGCCTT CTTCGGCATGAGCAGAATCGGCATGGAAGTGACCC CTTCTGGCACCTGGCTGACCTACACCGGCGCTATCA AGCTGGACGATAAGGATCCTAACTTCAAGGACCAA GTGATCCTGCTGAACAAGCATATCGACGCCTATAAG ACCTTTCCACCTACAGAGCCTAAGAAAGATAAGAA GAAGAAAGCCGACGAGACACAGGCCCTGCCTCAGA
--	--	--

		GACAGAAAAAGCAGCAGACAGTGACACTGCTGCCA GCCGCTGACCTGGATGACTTCAGCAAGCAGCTGCA GCAGAGCATGTCTTCTGCTGATAGCACCCAGGCCCCG AAAACGGCGCggaagcggagggaagcggagctactaacttcagcctgctga agcaggctggagatgtggaggagaaccctggacctATGTTTCGTGTTTCCT GGTGCTGCTGCCTCTGGTCAGCTCCCAGTGTGTGAA CCTGACCACCAGAACCCAGCTGCCACCTGCTTATAC AAACTCCTTCACTCGGGGGGTATACTACCCCGACAA GGTGTTCAGATCTAGCGTGCTGCATTCTACACAAGA CCTGTTCCCTGCCCTTCTTCAGCAACGTGACCTGGTTC CACGCCATCCACGTGTCTGGAACCAACGGAACCAA GAGATTGACAACCCCGTGCTGCCTTTCAACGACGG CGTGTAATTCGCCAGCACCGAGAAGTCCAACATCAT CAGAGGATGGATTTTCGGCACCACTGGACAGCA AAACCCAGAGCCTGCTGATCGTGAACAACGCCACC AACGTGGTGATCAAGGTGTGCGAGTTCAGTTCTGC AATGATCCCTTCCTGGGCGTGTACTACCACAAGAAC AACAAGTCTTGATGGAAAGCGAGTTCAGAGTGTA TTCCAGCGCCAACAATTGCACCTTCGAGTACGTGAG CCAACCCTTTCTGATGGACCTTGAAGGCAAGCAGGG CAACTTCAAAAATCTGCGAGAATTTGTGTTCAAGAA CATCGACGGATACTTCAAGATCTACTCTAAGCACAC GCCAATCAACCTGGTGAGAGATCTGCCCCAGGGCTT TAGCGCTTTGGAACCTCTGGTGGACCTGCCTATCGG AATCAACATCACCAGATTTCAAACCTCTCCTGGCCCT GCACAGATCTTATCTGACCCCTGGGGACAGTAGTAG CGGCTGGACAGCCGGCGCCGCCGCTACTACGTGG GATACCTGCAGCCTAGAACATTCCTGCTGAAGTACA ATGAGAACGGAACAATCACAGACGCCGTGGACTGC GCCCTGGATCCTTTGAGCGAGACAAAGTGCACCCTG AAGTCGTTACCGTCGAAAAAGGCATCTACCAGAC CAGCAACTTCCGCGTGACGCCTACGGAATCTATCGT GCGGTTCCCCAACATCACCACCTGTGCCCTTTCCG CGAGGTGTTTAACGCTACAAGGTTCGCCAGCGTGTA TGCTTGGAACAGAAAGAGAATCAGCAATTGCGTGG CCGATTATAGCGTTCTGTACAACAGCGCTTCCTTCA GCACCTTCAAGTGCTACGGCGTGTCTCCAACCAAGC TGAACGACCTCTGCTTCACCAATGTCTACGCTGACT CTTTCGTGATTAGAGGCGATGAGGTTAGACAGATCG CACCTGGCCAGACCGGCAAAATCGCTGACTACAAC TACAAGCTGCCTGATGACTTCACAGGCTGTGTGATT GCCTGGAACCTCAAATAACCTGGACTCTAAAGTGGG CGGCAACTACAACCTGTACCGGCTGTTCCGGAA GAGCAATCTGAAACCTTTTGAGCGGGACATCTCTAC AGAGATCTACCAGGCCGGCAGCACACCCTGCAACG GCGTTGAGGGCTTCAACTGCTACTTCCCTCTGCAGA GCTACGGCTTTCAGCCAACAAATGGAGTGGGCTACC AGCCGTACAGAGTGGTGGTGTGCTGAGCTTCGAACTGC TGCATGCCCCAGCCACAGTGTGTGGACCTAAGAAGT CTACCAACCTGGTGAAGAACAAGTGCCTGAACCTTA
--	--	---

		ACTTTAACGGCCTGACCGGCACAGGCGTGCTGACCG AATCCAACAAAAAGTTCCTGCCCTTCCAACAGTTCG GCAGAGACATCGCCGATACAACCGATGCCGTGCGG GACCCCCAGACCTTAGAAATCCTAGATATCACCCCG TGCAGCTTCGGCGGAGTCTCTGTTATTACTCCTGGC ACCAACACCAGCAACCAAGTGGCTGTTCTGTACCAA ggcGTGAACTGCACCGAAGTGCCTGTGGCTATCCACG CCGATCAGCTGACCCCAACCTGGCGGGTGTATAGCA CCGGCTCTAACGTGTTCCAGACCCGGGCTGGCTGCC TGATCGGGCGCCGAACACGTCAACAACCTCCTATGAAT GTGACATCCCCATCGGGGCTGGCATCTGCGCCAGTT ACCAGACACAGACAAATAGCCCTAGACGGGCCAGA AGCGTGGCCTCCCAGAGTATCATTGCCTACCCATG AGCCTGGGCGCCGAGAACAGCGTGGCCTATTCTAA CAATAGCATCGCAATCCCTACCAACTTTACCATCTC TGTGACAACCGAGATCCTGCCTGTGAGCATGACCAA AACCAGCGTGGACTGCACGATGTACATCTGTGGCG ACAGCACAGAATGCAGTAATCTGTTGCTGCAGTACG GCAGCTTTTGCACCCAGTTGAATAGAGCCCTGACCG GAATCGCCGTAGAGCAGGACAAAAATACCCAGGAG GTGTTGCGCCAGGTGAAACAGATCTACAAGACACCT CCCATTAAGGACTTCGGAGGTTTTAACTTCAGCCAG ATCCTGCCCAGCCCTTCCAAGCCTAGCAAACGCTCC TTCATCGAGGACCTGCTCTTCAACAAGGTGACACTG GCTGATGCCGGCTTCATCAAGCAGTACGGAGATTGT CTGGGAGACATCGCCGCTAGAGATCTGATCTGCGCC CAAAAGTTCAACGGCCTGACCGTGCTGCCTCCTCTG CTTACAGACGAGATGATCGCCAGTACACCAGCGC CCTGCTGGCTGGCACCATCACAAGCGGCTGGACCTT CGGAGCCGGAGCCGCTCTGCAAATCCCCTTTGCCAT GCAGATGGCCTACCGGTTCAACGGCATCGGCGTGA CACAGAATGTGCTGTACGAGAACCAGAAGCTGATC GCTAACCAGTTTAAACAGCGCTATCGGCAAGATCCAG GACTCGCTGAGTAGCACCGCCTCTGCCCTGGGCAAG CTGCAGGACGTCTGTGAACCAGAACGCCCAAGCCCT GAACACACTGGTGAAACAGCTGAGCAGCAACTTCG GCGCCATCAGCTCTGTGCTGAACGATATCCTGAGCA GACTGGACAAGGTGGAAGCCGAGGTCCAGATCGAC AGACTGATCACAGGAAGACTGCAGAGCCTGCAAAAC GTACGTGACACAGCAGCTGATCCGGGCAGCCGAAA TCCGGGCCAGCGCCAATCTGGCCGCTACCAAGATG AGCGAGTGCGTGTTAGGCCAGAGCAAGCGGGTGGA TTTCTGCGGTAAGGGATACCACCTGATGAGCTTTCC CCAGAGCGCTCCTCACGGCGTGGTGTCTGACAGT GACCTACGTTTCCTGCCCAGGAAAAGAACTTACCAC CGCCCCTGCTATCTGCCACGATGGCAAGGCCCACTT CCCTAGAGAGGGCGTTTTCTGTCTAACGGCACACA CTGGTTTGTGACCCAGAGAACTTCTACGAGCCTCA GATCATCACACAGACAACACCTTTGTGAGCGGCA ATTGCGACGTGGTGATCGGAATTGTTAATAATACCG
--	--	---

		TGTACGACCCTCTGCAGCCTGAGCTCGACAGCTTCA AGGAAGAGCTGGACAAGTACTTCAAGAACCACACC TCCCCAGATGTGGACCTGGGCGATATTTTCAGGCATC AACGCCTCCGTCGTGAATATCCAGAAGGAGATCGA CCGGCTCAACGAGGTGGCCAAGAACCTTAACGAGA GCCTGATCGACCTGCAGGAACCTGGGCAAATATGAG CAGTACATCAAGTGGCCTTGGTACATCTGGCTGGGC TTTATCGCAGGCCTGATCGCTATCGTGATGGTGACC ATTATGCTGTGTTGTATGACCAGCTGTTGTAGTTGTC TGAAGGGCTGCTGTTCTTGCGGCAGCTGCTGCAAGT TCGACGAAGACGACTCAGAGCCCGTGCTGAAAGGC GTGAAGCTGCACTACACCCGAAAACGGCGCggaagcgg aggaagcggagctactaactcagcctgctgaagcaggctggagatgtggaggaga accctggacctATGTATTCTTTTGTGTCCGAGGAAACCGG CACACTGATCGTTAATAGCGTGCTGCTCTTCCTGGC CTTCGTGGTGTTCCTGCTGGTGACCCTGGCTATCCTG ACCGCCCTGAGACTGTGTGCCTACTGCTGCAACATC GTGAACGTGTCTCTGGTCAAGCCTAGCTTCTACGTG TACAGCCGGGTGAAGAACCTGAACAGCAGCAGAGT GCCCACCTGCTGGTGaatccccccccctaactgactggccgaag ccgcttggaataaggccggtgtgcgtttgtctatatgtattttccaccatattgccgtcttt tggcaatgtgagggcccgaaacctggccctgtcttcttgacgagcattcctaggggt ctttccctctcgcaaagggaatgaaggctctgttgatgtcgtgaaggaaagcagttcc tctggaagcttctgaagacaaacaacgtctgtacgaccccttgacggcagcgggaac ccccacctggcgacaggtgcctctcgggccaaaagccacgtgataagataacct gcaaaggcggcacaaacccagtgccacgttgtagttgatagttgtaaagagtc aatggctctctcaagcgtattcaacaagggctgaaggatgccagaaggtaccc cattgtatgggatctgatctggggcctcggtgcacatgctttacatgtgttagtcaggt taaaaaacgtctagggccccgaaccacggggacgtggttttctttgaaaaacacg atgataatattggccacaacctggaacaagagactgcgcgcactctctcacttttag gaatgccccaaatgctctgctctacaataaccgtaattgatttacctgctaaagtatgat gaagaatgggtaccagaggagtattgactgatggagaggatgatgtctttgatccg aattagacatggaagtcgttttcgagttacagtaaGTGTtacctgttaattgtagcattt gagctttgggctaagcgcaacattaaccagtagcagaggtgaaatactcaataattt gggtgtggacattgctgctaatactgtgatctgggactacaaaagagatgtccagca cataatctactattggtgtttgtctatgactgacatagccaagaacaaactgaacga ttgtgcaccactcactgtctttttgatggtagagttgatgtcaagtagactatttagaa atgcccgtaatggtgttcttattacagaaggtagtgttaaaggtttacaacctctgtagg tcccaacaagctagtcttaattggagtcacattaattggagaagccgtaaaaacacagt tcaattattataagaaagttgatggtgtgtccaacaattacctgaaacttactttactcag agtagaaatttacaagaatttaaaccaggagtcfaatggaaattgatttcttagaatta gctatggatgaattcattgaacggtataaattagaaggctatgccttcgaacatacgttt atggagatttttagtcagtagtgtaggtggtGCGAaattgtgtgtt
CoVEG13 expression cassette	45	ATTAAAGGTTTATACCTTCCCAGGTAACAAACCAAC CAACTTTTCGATCTCTTGTAGATCTGTTCTCTAAACGA ACTTTAAAATCTGTGTGGCTGTCACTCGGCTGCATG CTTAGTGCACCTCACGCAGTATAATTAATAACTAATT ACTGTCTGTTGACAGGACACGAGTAACCTCGTCTATCT TCTGCAGGCTGCTTACGGTTTCGTCCGTGTTGCAGC

		CGATCATCAGCACATCTAGGTTTCGTCCGGGTGTGA CCGAAAGGTAAATGGCCGACAGCAACGGCACAATC ACCGTGGAAGAGCTGAAGAACTGCTGGAACAGTG GAACCTGGTCATCGGCTTCCTGTTTCTGACCTGGAT CTGTCTGCTGCAGTTCGCTTATGCCAATCGGAACAG ATTCTGTACATCATCAAGCTGATCTTCCTGTGGCT GCTGTGGCCTGTGACCCTGGCTTGCTTCGTGTGGC CGCTGTGTACCGGATCAACTGGATCACAGGCGGAA TCGCCATCGCCATGGCCTGCCTGGTGGGCCTGATGT GGCTGAGCTACTTCATCGCTTCTTTCAGACTGTTTCG CAGAACCCGGAGCATGTGGTCCTTCAACCCCGAGA CAAACATCCTGCTGAACGTGCCTCTGCACGGCACCA TCCTGACAAGACCTCTGCTCGAGAGCGAGCTGGTGA TTGGCGCAGTGATTCTGAGAGGCCATCTGAGGATCG CCGGACACCACCTGGGCAGATGCGACATCAAGGAC CTTCCAAAGGAAATCACCGTTGCCACCAGCCGGACC CTGTCCTACTACAACTGGGCGCCAGCCAAAGAGT GGCCGGCGATAGCGGCTTTGCCGCCTACAGCAGAT ACCGCATCGGAAATTACAAGCTCAACACCGACCAC AGCAGCTCTTCTGATAACATCGCCCTGCTGGTGCAG CGAAAACGGCGCggaagcggaggaagcggagctactaactcagcctg ctgaagcaggctggagatgtggaggagaaccctggacctATGAGCGACA ACGGCCCTCAAAACCAGAGAAATGCCCTCGGATC ACATTTGGCGGACCTAGCGACAGCACCGGCAGCAA CCAGAATGGAGAAAGAAGCGGCGCCAGATCCAAGC AGCGGAGACCTCAGGGACTGCCCAACAACACCGCT AGCTGGTTCACCGCCCTGACCCAACACGGCAAGGA AGATCTGAAGTTCCCCAGAGGCCAGGGCGTGCCTAT CAACACAACTCTTCTCCCGACGACCAGATCGGATA CTATAGACGGGCCACTCGGAGAATTGCGGGCGGCG ACGGAAAAATGAAGGACCTTCTCCAAGATGGTAC TTCTACTACCTCGGCACAGGCCCTGAGGCCGGCCTG CCTTACGGCGCCAACAAGGATGGCATCATCTGGGTC GCCACCGAGGGCGCCCTGAACACCCCTAAGGACCA CATCGGCACAAGAAACCCCGCTAACAACGCCGCAA TCGTGCTGCAGCTGCCTCAGGGCACCAACCTGCCCA AGGGCTTCTACGCCGAGGGCTCTAGAGGTGGCTCCC AGGCTTCTAGCCGCTCCTCCAGCCGCAGCAGAAACA GCAGCAGGAACAGCACCCCGGCAGCTCCCGGGGC ACCAGCCCCGCCAGAATGGCCGGAAATGGCGGCGA TGCCGCCCTGGCCCTGCTCCTGCTGGACAGACTGAA TCAGCTGGAAAGCAAGATGAGCGGCAAAGGACAGC AGCAGCAAGGCCAGACCGTGACCAAGAAAAGCGCT GCTGAAGCCTCCAAGAAACCTAGACAAAAGCGGAC CGCCACAAAGGCCTACAACGTGACCCAAGCCTTTG GAAGAAGAGGCCCCGAGCAGACACAGGGCAATTTG GGCGACCAGGAGCTGATCCGGCAGGGAACCGACTA CAAGCACTGGCCTCAGATCGCCCAGTTGCCCCCTAG CGCCAGCGCCTTCTTCGGCATGAGCAGAATCGGCAT GGAAGTGACCCCTTCTGGCACCTGGCTGACCTACAC
--	--	--

		CGGCGCTATCAAGCTGGACGATAAGGATCCTAACTT CAAGGACCAAGTGATCCTGCTGAACAAGCATATCG ACGCCTATAAGACCTTTCCACCTACAGAGCCTAAGA AAGATAAGAAGAAGAAAGCCGACGAGACACAGGC CCTGCCTCAGAGACAGAAAAAGCAGCAGACAGTGA CACTGCTGCCAGCCGCTGACCTGGATGACTTCAGCA AGCAGCTGCAGCAGAGCATGTCTTCTGCTGATAGCA CCCAGGCCCCGAAAACGGCGCggaagcggaggaagcggagctac taacttcagcctgctgaagcaggctggagatgtggaggagaaccctggacctATG TTCGTGTTCTGGTGCTGCTGCCTCTGGTCAGCTCCC AGTGTGTGAACCTGACCACCAGAACCCAGCTGCCA CCTGCTTATACAACTCCTTCACTCGGGGGGTATAC TACCCCGACAAGGTGTTTCAGATCTAGCGTGCTGCAT TCTACACAAGACCTGTTCCCTGCCCTTCTTCAGCAAC GTGACCTGGTTCACGCCATCCACGTGTCTGGAACC AACGGAACCAAGAGATTTCGACAACCCCGTGCTGCC TTTCAACGACGGCGTGACTTCGCCAGCACCCGAGAA GTCCAACATCATCAGAGGATGGATTTTCGGCACCAC ACTGGACAGCAAAACCCAGAGCCTGCTGATCGTGA ACAACGCCACCAACGTGGTGATCAAGGTGTGCGAG TTCCAGTTCTGCAATGATCCCTTCCTGGGCGTGACT ACCACAAGAACAACAAGTCTTGATGGAAAGCGAG TTCAGAGTGTATTCCAGCGCCAACAATTGCACCTTC GAGTACGTGAGCCAACCCTTTCTGATGGACCTTGAA GGCAAGCAGGGCAACTTCAAAAATCTGCGAGAATT TGTGTTCAAGAACATCGACGGATACTTCAAGATCTA CTCTAAGCACACGCCAATCAACCTGGTGAGAGATCT GCCCCAGGGCTTTAGCGCTTTGGAACCTCTGGTGGA CCTGCCTATCGGAATCAACATCACCAAGATTTCAAAC TCTCCTGGCCCTGCACAGATCTTATCTGACCCCTGG GGACAGTAGTAGCGGCTGGACAGCCGGCGCCGCCG CCTACTACGTGGGATACCTGCAGCCTAGAACATTCC TGCTGAAGTACAATGAGAACGGAACAATCACAGAC GCCGTGGACTGCGCCCTGGATCCTTTGAGCGAGACA AAGTGCACCCTGAAGTCGTTACCGTCGAAAAAGG CATCTACCAGACCAGCAACTTCCGCGTGCAGCCTAC GGAATCTATCGTGCGGTTCCCCAACATCACCAACCT GTGCCCTTTTCGGCGAGGTGTTTAAACGCTACAAGGTT CGCCAGCGTGTATGCCTGGAACAGAAAGAGAATCA GCAATTGCGTGGCCGATTATAGCGTTCTGTACAACA GCGCTTCCTTCAGCACCTTCAAGTGCTACGGCGTGT CTCCAACCAAGCTGAACGACCTCTGCTTCACCAATG TCTACGCTGACTCTTTCGTGATTAGAGGCGATGAGG TTAGACAGATCGCACCTGGCCAGACCGGCAAAATC GCTGACTACAACTACAAGCTGCCTGATGACTTCACA GGCTGTGTCAATTGCCTGGAACCTCAAATAACCTGGAC TCTAAAGTGGGCGGCAACTACAACCTGTACCGG CTGTTCCGGAAGAGCAATCTGAAACCTTTTGAGCGG GACATCTCTACAGAGATCTACCAGGCCGGCAGCAC ACCCTGCAACGGCGTTGAGGGCTTCAACTGCTACTT
--	--	---

		CCCTCTGCAGAGCTACGGCTTTCAGCCAACAAATGG AGTGGGCTACCAGCCGTACAGAGTGGTGGTGCTGA GCTTCGAACTGCTGCATGCCCCAGCCACAGTGTGTG GACCTAAGAAGTCTACCAACCTGGTGAAGAACAAG TGCGTGAACTTTAACTTTAACGGCCTGACCGGCACA GGCGTGCTGACCGAATCCAACAAAAAGTTCCTGCCC TTCCAACAGTTCGGCAGAGACATCGCCGATACAACC GATGCCGTGCGGGACCCCCAGACCTTAGAAATCCTA GATATCACCCCGTGCAGCTTCGGCGGAGTCTCTGTT ATTACTCCTGGCACCAACACCAGCAACCAAGTGGCT GTTCTGTACCAAggcGTGAACTGCACCGAAGTGCCTG TGGCTATCCACGCCGATCAGCTGACCCCAACCTGGC GGGTGTATAGCACCGGCTCTAACGTGTTCCAGACCC GGGCTGGCTGCCTGATCGGCGCCGAACACGTCAAC AACTCCTATGAATGTGACATCCCCATCGGGGCTGGC ATCTGCGCCAGTTACCAGACACAGACAAATAGCCCT AGACGGGCCAGAAAGCGTGGCCTCCAGAGTATCAT TGCCTACACCATGAGCCTGGGCGCCGAGAACAGCG TGGCCTATTCTAACAATAGCATCGCAATCCCTACCA ACTTTACCATCTCTGTGACAACCGAGATCCTGCCTG TGAGCATGACCAAAACCAGCGTGGACTGCACGATG TACATCTGTGGCGACAGCACAGAATGCAGTAATCTG TTGCTGCAGTACGGCAGCTTTTGCACCCAGTTGAAT AGAGCCCTGACCGGAATCGCCGTAGAGCAGGACAA AAATACCCAGGAGGTGTTTCGCCAGGTGAAACAGA TCTACAAGACACCTCCCATTAAGGACTTCGGAGGTT TTAACTTCAGCCAGATCCTGCCCGACCCTTCCAAGC CTAGCAAACGCTCCTTCATCGAGGACCTGCTCTTCA ACAAGGTGACACTGGCTGATGCCGGCTTCATCAAGC AGTACGGAGATTGTCTGGGAGACATCGCCGCTAGA GATCTGATCTGCGCCCAAAAGTTCAACGGCCTGACC GTGCTGCCTCCTCTGCTTACAGACGAGATGATCGCC CAGTACACCAGCGCCCTGCTGGCTGGCACCATCACA AGCGGCTGGACCTTCGGAGCCGGAGCCGCTCTGCA AATCCCCTTTGCCATGCAGATGGCCTACCGGTTCAA CGGCATCGGCGTGACACAGAATGTGCTGTACGAGA ACCAGAAGCTGATCGCTAACAGTTTAAACAGCGCTA TCGGCAAGATCCAGGACTCGCTGAGTAGCACCGCCT CTGCCCTGGGCAAGCTGCAGGACGTCGTGAACCA AACGCCCAAGCCCTGAACACACTGGTGAACAGCT GAGCAGCAACTTCGGCGCCATCAGCTCTGTGCTGAA CGATATCCTGAGCAGACTGGACAAGGTGGAAGCCG AGGTCCAGATCGACAGACTGATCACAGGAAGACTG CAGAGCCTGCAAACGTACGTGACACAGCAGCTGAT CCGGGCAGCCGAAATCCGGGCCAGCGCCAATCTGG CCGCTACCAAGATGAGCGAGTGCGTGTTAGGCCAG AGCAAGCGGGTGGATTTCTGCGGTAAGGGATACCA CCTGATGAGCTTTCCCCAGAGCGCTCCTACGGCGT GGTGTCTTCTGCACGTGACCTACGTTCTGCCAGGA AAAGAACTTCAACCACCGCCCCTGCTATCTGCCACGA
--	--	---

		TGGCAAGGCCCACTTCCCTAGAGAGGGCGTTTTCTGT GTCTAACGGCACACACTGGTTTGTGACCCAGAGAA ACTTCTACGAGCCTCAGATCATCACCACAGACAACA CCTTTGTGAGCGGCAATTGCGACGTGGTGATCGGAA TTGTTAATAATACCGTGTACGACCCTCTGCAGCCTG AGCTCGACAGCTTCAAGGAAGAGCTGGACAAGTAC TTCAAGAACCACACCTCCCCAGATGTGGACCTGGGC GATATTTTCAGGCATCAACGCCTCCGTCTGTAATATC CAGAAGGAGATCGACCGGCTCAACGAGGTGGCCAA GAACCTTAACGAGAGCCTGATCGACCTGCAGGAAC TGGGCAAATATGAGCAGTACATCAAGTGGCCTTGGT ACATCTGGCTGGGCTTTATCGCAGGCCTGATCGCTA TCGTGATGGTGACCATTATGCTGTGTTGTATGACCA GCTGTTGTAGTTGTCTGAAGGGCTGCTGTTCTTGCG GCAGCTGCTGCAAGTTCGACGAAGACGACTCAGAG CCCGTGCTGAAAGGCGTGAAGCTGCACTACACCCG AAAACGGCGCggaagcggaggaagcggagctactaacttcagcctgctga agcaggctggagatgtggaggagaacctggacctATGTATTCTTTTGT GTCCGAGGAAACCGGCACACTGATCGTTAATAGCG TGCTGCTCTTCTGGCCTTCGTGGTGTTCCTGCTGGT GACCCTGGCTATCCTGACCGCCCTGAGACTGTGTGC CTACTGCTGCAACATCGTGAACGTGTCTCTGGTCAA GCCTAGCTTCTACGTGTACAGCCGGGTGAAGAACCT GAACAGCAGCAGAGTGCCCCGACCTGCTGGTGaatcccc ccccctaacttactggccgaagccgcttgaataaggccggtgtgcgtttgtctata tgttattttccaccatattgccgtcttttgcaatgtgagggcccgaaacctggccctgt cttcttgacgagcattctaggggtctttccctctcgccaaaggaatgcaaggtctgtt gaatgtcgtgaaggaaagcagttcctctggaagcttctgaagacaacaacgtctgta gcgacctttgcaggcagcggaacccccacctggcgacaggtgcctctcgcgcca aaagccacgtgtataagatacacctgcaaaggcggcacaacccagtgccacgttgt gagttggatagttgtgaaagagtcaaatggctctctcaagcgtattcaacaagggg ctgaaggatcccagaaggtacccattgtatgggatctgatctggggcctcggtgca catgctttacatgtgttagtcgaggttaaaaaacgtctaggccccccgaaccacggg gacgtggtttccittgaaaaacacgatgataatatggccacaacctggaacaagag acttgcgcgactctctcacttttgagggaatgccccaaatgctctgctctacaataccgt aatggattttacctgctaaagtatgatgaagaatggtacccagaggagttattgactgat ggagaggatgatgtctttgatcccgaattagacatggaagtcgttttcgagttacagtaa GTGTtacctgttaatgtagcatttgagctttgggctaagcgcaacattaaaccagta ccagaggtgaaaatactcaataattgggtgtggacattgctgctaatactgtgatctgg gactacaaaagagatgtccagcacatatatctactattgggtgtttgttctatgactgaca tagccaagaacaaactgaaacgatttgtgcaccactcactgtctttttgatggtagag ttgatggtaagtagacttatttagaaatgcccgtaatgggtgttctattacagaaggttag tgttaaagggtttacaacctctgtaggtcccaacaagctagtcttaattggagtcacatt aattggagaagccgtaaaaacacagttcaattattataagaaagttgatggtgtgtcca acaattacctgaaacttacttactcagagtagaaattacaagaatttaaaaccaggag tcaaatggaaattgatttcttagaattagctatggatgaattcattgaacggtataaattag aaggctatgccttcgaacatatcgttatggagattttagtcatagtcagttaggtgtG CGAaattgtgtgtt
--	--	---

CoVEG14 expression cassette	46	ATTAAAGGTTTATACCTTCCCAGGTAACAAACCAAC CAACTTTCGATCTCTTGTAGATCTGTTCTCTAAACGA ACTTTAAAATCTGTGTGGCTGTCACTCGGCTGCATG CTTAGTGCACCTCACGCAGTATAATTAATAACTAATT ACTGTCGTTGACAGGACACGAGTAACTCGTCTATCT TCTGCAGGCTGCTTACGGTTTCGTCCGTGTTGCAGC CGATCATCAGCACATCTAGGTTTCGTCCGGGTGTGA CCGAAAGGTAAATGGCCGACAGCAACGGCACAAATC ACCGTGGAAGAGCTGAAGAACTGCTGGAACAGTG GAACCTGGTCATCGGCTTCCTGTTTCTGACCTGGAT CTGTCTGCTGCAGTTCGCTTATGCCAATCGGAACAG ATTCCTGTACATCATCAAGCTGATCTTCCTGTGGCT GCTGTGGCCTGTGACCCTGGCTTGCTTCGTGCTGGC CGCTGTGTACCGGATCAACTGGATCACAGGCGGAA TCGCCATCGCCATGGCCTGCCTGGTGGGCCTGATGT GGCTGAGCTACTTCATCGCTTCTTTCAGACTGTTTCG CAGAACCCGGAGCATGTGGTCCTTCAACCCCGAGA CAAACATCCTGCTGAACGTGCCTCTGCACGGCACCA TCCTGACAAGACCTCTGCTCGAGAGCGAGCTGGTGA TTGGCGCAGTGATTCTGAGAGGCCATCTGAGGATCG CCGGACACCACCTGGGCAGATGCGACATCAAGGAC CTTCCAAAGGAAATCACCGTTGCCACCAGCCGGACC CTGTCCTACTACAACTGGGCGCCAGCCAAAGAGT GGCCGGCGATAGCGGCTTTGCCGCCTACAGCAGAT ACCGCATCGGAAATTACAAGCTCAACACCGACCAC AGCAGCTCTTCTGATAACATCGCCCTGCTGGTGCAG CGAAAACGGCGCggaagcggaggaagcggagctactaactcagcctg ctgaagcaggctggagatgtggaggagaaccctggacctATGAGCGACA ACGGCCCTCAAAACCAGAGAAATGCCCTCGGATC ACATTTGGCGGACCTAGCGACAGCACCGGCAGCAA CCAGAATGGAGAAAGAAGCGGCGCCAGATCCAAGC AGCGGAGACCTCAGGGACTGCCCCAACACACCGCT AGCTGGTTCACCGCCCTGACCCAACACGGCAAGGA AGATCTGAAGTTCCCCAGAGGCCAGGGCGTGCCTAT CAACACAACTCTTCTCCCGACGACCAGATCGGATA CTATAGACGGGCCACTCGGAGAATTCTGGGGCGGCG ACGGAAAAATGAAGGACCTTTCTCCAAGATGGTAC TTCTACTACCTCGGCACAGGCCCTGAGGCCGGCCTG CTTACGGCGCCAACAAGGATGGCATCATCTGGGTC GCCACCGAGGGCGCCCTGAACACCCCTAAGGACCA CATCGGCACAAGAAACCCCGCTAACAACGCCGCAA TCGTGCTGCAGCTGCCTCAGGGCACCAACCTGCCCA AGGGCTTCTACGCCGAGGGCTCTAGAGGTGGCTCCC AGGCTTCTAGCCGCTCCTCCAGCCGCAGCAGAAACA GCAGCAGGAACAGCACCCCGGCAGCTCCCGGGGC ACCAGCCCCGCCAGAATGGCCGGAAATGGCGGCGA TGCCGCCCTGGCCCTGCTCCTGCTGGACAGACTGAA TCAGCTGGAAAGCAAGATGAGCGGCAAAGGACAGC AGCAGCAAGGCCAGACCGTGACCAAGAAAAGCGCT GCTGAAGCCTCCAAGAAACCTAGACAAAAGCGGAC
-----------------------------------	----	--

		CGCCACAAAGGCCTACAACGTGACCCAAGCCTTTG GAAGAAGAGGCCCGAGCAGACACAGGGCAATTTTC GGCGACCAGGAGCTGATCCGGCAGGGAACCGACTA CAAGCACTGGCCTCAGATCGCCCAGTTCGCCCTAG CGCCAGCGCCTTCTTCGGCATGAGCAGAATCGGCAT GGAAGTGACCCCTTCTGGCACCTGGCTGACCTACAC CGGCGCTATCAAGCTGGACGATAAGGATCCTAACTT CAAGGACCAAGTGATCCTGCTGAACAAGCATATCG ACGCCTATAAGACCTTTCCACCTACAGAGCCTAAGA AAGATAAGAAGAAGAAAGCCGACGAGACACAGGC CCTGCCTCAGAGACAGAAAAAGCAGCAGACAGTGA CACTGCTGCCAGCCGCTGACCTGGATGACTTCAGCA AGCAGCTGCAGCAGAGCATGTCTTCTGCTGATAGCA CCCAGGCCCGAAAACGGCGCggaagcggaggaagcggagctac taacttcagcctgtgaagcaggctggagatgtggaggagaacctggacctATG TTCGTGTTCTGGTGCTGCTGCCTCTGGTCAGCTCCC AGTGTGTGAACCTGACCACCAGAACCCAGCTGCCA CCTGCTTATACAACTCCTTCACTCGGGGGGTATAC TACCCCGACAAGGTGTTTCAGATCTAGCGTGCTGCAT TCTACACAAGACCTGTTCTGCCCCTTCTTCAGCAAC GTGACCTGGTTCACGCCATCCACGTGTCTGGAACC AACGGAACCAAGAGATTTCGACAACCCCGTGCTGCC TTTCAACGACGGCGTGACTTCGCCAGCACCCGAGAA GTCCAACATCATCAGAGGATGGATTTTCGGCACCAC ACTGGACAGCAAAACCCAGAGCCTGCTGATCGTGA ACAACGCCACCAACGTGGTGATCAAGGTGTGCGAG TTCCAGTTCTGCAATGATCCCTTCCTGGGCGTGACT ACCACAAGAACAACAAGTCTTGATGGAAGCGAG TTCAGAGTGTATTCCAGCGCCAACAATTGCACCTTC GAGTACGTGAGCCAACCTTTCTGATGGACCTTGAA GGCAAGCAGGGCAACTTCAAAAATCTGCGAGAATT TGTGTTCAAGAACATCGACGGATACTTCAAGATCTA CTCTAAGCACACGCCAATCAACCTGGTGAGAGATCT GCCCCAGGGCTTTAGCGCTTTGGAACCTCTGGTGGA CCTGCCTATCGGAATCAACATCACCAAGATTTCAAAC TCTCCTGGCCCTGCACAGATCTTATCTGACCCCTGG GGACAGTAGTAGCGGCTGGACAGCCGGCGCCGCCG CCTACTACGTGGGATACCTGCAGCCTAGAACATTCC TGCTGAAGTACAATGAGAACGGAACAATCACAGAC GCCGTGGACTGCGCCCTGGATCCTTTGAGCGAGACA AAGTGACCCCTGAAGTCGTTACCGTCGAAAAAGG CATCTACCAGACCAGCAACTTCCGCGTGCAGCCTAC GGAATCTATCGTGCGGTTCCCCAACATCACCAACCT GTGCCCTTTTCGGCGAGGTGTTTAACGCTACAAGGTT CGCCAGCGTGTATGCCTGGAACAGAAAGAGAATCA GCAATTGCGTGGCCGATTATAGCGTTCTGTACAACA GCGCTTCCTTCAGCACCTTCAAGTGCTACGGCGTGT CTCCAACCAAGCTGAACGACCTCTGCTTCACCAATG TCTACGCTGACTCTTTCGTGATTAGAGGCGATGAGG TTAGACAGATCGCACCTGGCCAGACCGGCAAAATC
--	--	---

		GCTGACTACAACACTACAAGCTGCCTGATGACTTCACA GGCTGTGTCATTGCCTGGAACCTCAAATAACCTGGAC TCTAAAGTGGGCGGCAACTACAACCTACCTGTACCGG CTGTTCCGGAAGAGCAATCTGAAACCTTTTGAGCGG GACATCTCTACAGAGATCTACCAGGCCGGCAGCAC ACCCTGCAACGGCGTTGAGGGCTTCAACTGCTACTT CCCTCTGCAGAGCTACGGCTTTCAGCCAACAATGG AGTGGGCTACCAGCCGTACAGAGTGGTGGTGCTGA GCTTCGAACTGCTGCATGCCCCAGCCACAGTGTGTG GACCTAAGAAGTCTACCAACCTGGTGAAGAACAAG TGCGTGAACTTTAACTTTAACGGCCTGACCGGCACA GGCGTGCTGACCGAATCCAACAAAAAGTTCTGCCC TTCCAACAGTTCGGCAGAGACATCGCCGATACAACC GATGCCGTGCGGGACCCCCAGACCTTAGAAATCCTA GATATCACCCCGTGCAGCTTCGGCGGAGTCTCTGTT ATTACTCCTGGCACCAACACCAGCAACCAAGTGGCT GTTCTGTACCAAggcGTGAACTGCACCGAAGTGCCTG TGGCTATCCACGCCGATCAGCTGACCCCAACCTGGC GGGTGTATAGCACCGGCTCTAACGTGTTCCAGACCC GGGCTGGCTGCCTGATCGGCGCCGAACACGTCAAC AACTCCTATGAATGTGACATCCCCATCGGGGCTGGC ATCTGCGCCAGTTACCAGACACAGACAAATAGCCCT AGACGGGCCAGAAGCGTGGCCTCCAGAGTATCAT TGCCTACACCATGAGCCTGGGCGCCGAGAACAGCG TGGCCTATTCTAACAATAGCATCGCAATCCCTACCA ACTTTACCATCTCTGTGACAACCGAGATCCTGCCTG TGAGCATGACCAAAACCAGCGTGGACTGCACGATG TACATCTGTGGCGACAGCACAGAATGCAGTAATCTG TTGCTGCAGTACGGCAGCTTTTGACCCAGTTGAAT AGAGCCCTGACCGGAATCGCCGTAGAGCAGGACAA AAATACCCAGGAGGTGTTGCGCCAGGTGAAACAGA TCTACAAGACACCTCCCATTAAGGACTTCGGAGGTT TTAACTTCAGCCAGATCCTGCCCCGACCCTTCCAAGC CTAGCAAACGCTCCTTCATCGAGGACCTGCTCTTCA ACAAGGTGACACTGGCTGATGCCGGCTTCATCAAGC AGTACGGAGATTGTCTGGGAGACATCGCCGCTAGA GATCTGATCTGCGCCCAAAGTTCAACGGCCTGACC GTGCTGCCTCCTCTGCTTACAGACGAGATGATCGCC CAGTACACCAGCGCCCTGCTGGCTGGCACCATCACA AGCGGCTGGACCTTCGGAGCCGGAGCCGCTCTGCA AATCCCCTTTGCCATGCAGATGGCCTACCGGTTCAA CGGCATCGGCGTGACACAGAATGTGCTGTACGAGA ACCAGAAGCTGATCGCTAACCAGTTTAAACAGCGCTA TCGGCAAGATCCAGGACTCGCTGAGTAGCACCGCCT CTGCCCTGGGCAAGCTGCAGGACGTCGTGAACCA AACGCCCAAGCCCTGAACACACTGGTGAACAGCT GAGCAGCAACTTCGGCGCCATCAGCTCTGTGCTGAA CGATATCCTGAGCAGACTGGACAAGGTGGAAGCCG AGGTCCAGATCGACAGACTGATCACAGGAAGACTG CAGAGCCTGCAAACGTACGTGACACAGCAGCTGAT
--	--	---

		CCGGGCAGCCGAAATCCGGGCCAGCGCCAATCTGG CCGCTACCAAGATGAGCGAGTGCGTGTTAGGCCAG AGCAAGCGGGTGGATTCTGCGGTAAGGGATACCA CCTGATGAGCTTTCCCCAGAGCGCTCCTCACGGCGT GGTGTTTCTGCACGTGACCTACGTTCTGCCAGGA AAAGAACTTCACCACCGCCCCTGCTATCTGCCACGA TGGCAAGGCCCACTTCCCTAGAGAGGGCGTTTTCGT GTCTAACGGCACACACTGGTTTGTGACCCAGAGAA ACTTCTACGAGCCTCAGATCATCACCACAGACAACA CCTTTGTGAGCGGCAATTGCGACGTGGTGATCGGAA TTGTTAATAATACCGTGTACGACCCTCTGCAGCCTG AGCTCGACAGCTTCAAGGAAGAGCTGGACAAGTAC TTCAAGAACCACACCTCCCCAGATGTGGACCTGGGC GATATTTAGGCATCAACGCCTCCGTCGTGAATATC CAGAAGGAGATCGACCGGCTCAACGAGGTGGCCAA GAACCTTAACGAGAGCCTGATCGACCTGCAGGAAC TGGGCAAATATGAGCAGTACATCAAGTGGCCTTGGT ACATCTGGCTGGGCTTTATCGCAGGCCTGATCGCTA TCGTGATGGTGACCATTATGCTGTGTTGTATGACCA GCTGTTGTAGTTGTCTGAAGGGCTGCTGTTCTTGCG GCAGCTGCTGCAAGTTCGACGAAGACGACTCAGAG CCCGTGCTGAAAGGCGTGAAGCTGCACTACACCCG AAAACGGCGCGggaagcggaggaagcggagctactaacttcagcctgctga agcaggctggagatgtggaggagaacctggacctATGTATTCTTTTGT GTCCGAGGAAACCGGCACACTGATCGTTAATAGCG TGCTGCTCTTCTGGCCTTCGTGGTGTTCCTGCTGGT GACCCTGGCTATCCTGACCGCCCTGAGACTGTGTGC CTACTGCTGCAACATCGTGAACGTGTCTCTGGTCAA GCCTAGCTTCTACGTGTACAGCCGGGTGAAGAACCT GAACAGCAGCAGAGTGCCCCGACCTGCTGGTGtaatcccc cccccctaacgttactggccgaagccgcttgaataaggccggtgtcggtttgtctata tgtttttccaccatattgcccgttttggcaatgtgagggcccggaacctggccctgt cttcttgacgagcattctaggggtcttcccctctcgccaaaggaatgcaaggctgtt gaatgtcgtgaagggaagcagttcctctggaagcttctgaagacaaacacgtctgta gcgacctttgcaggcagcgggaacccccacctggcgacaggtgcctctgcggcca aaagccacgtgtataagatacacctgcaaaggcggcacaacccagtgccacgttgt gagttggatagttgtgaaagagtcaaatggctctcctcaagcgtattcaacaagggg ctgaaggatcccagaaggtacccattgtatgggatctgatctggggcctcggtgca catgctttacatgtgttagtcgaggttaaaaaacgtctagggcccccgaaccacggg gacgtggtttctttgaaaaacacgatgataatatggccacaacctggaacaagag acttgcgcgactctctcacttttgagggaatgccccaaatgctctgctctacaataccgt aatggattttacctgctaaagtatgatgaagaatggtacccagaggagttattgactgat ggagaggatgatgtctttgatcccgaattagacatggaagtcgttttcgagttacaggg aagcggagctactaacttcagcctgctgaagcaggctggagatgtggaggagaacc ctggacctATGGACCTGTTCATGAGAATCTTCACCATCG GCACCGTGACACTGAAGCAGGGCGAGATCAAGGAT GCCACCCCTAGCGACTTCGTGAGAGCCACCGCCACA ATTCCTATCCAGGCTAGCCTGCCTTTTGGATGGCTG ATCGTGGGCGTCGCCCTGCTCGCCGTGTTCCAGAGC GCCTCTAAGATCATTACACTGAAGAAAAGATGGCA
--	--	---

		GCTGGCCCTCTCCAAAGGCGTGCACTTCGTGTGTAA TCTGCTGCTGCTTTTTGTGACAGTGTACAGCCACCT GCTGCTGGTTGCTGCTGGCCTGGAAGCCCCTTTCCT GTACCTGTACGCCCTGGTCTACTTCCTGCAGTCTATC AACTTCGTGCGGATCATCATGCGGCTGTGGCTGTGC TGGAAGTGCAGAAGCAAGAACCCACTGCTGTACGA CGCCAATTACTTCCTGTGTTGGCACACCAACTGCTA CGACTACTGCATCCCCTACAACAGCGTGACCAGCAG CATCGTGATCACCTCTGGCGACGGAACAACCAGCCC TATCAGCGAGCATGATTACCAGATCGGCGGATATAC AGAGAAGTGGGAGAGCGGCGTGAAGGACTGCGTGG TGCTGCACAGCTACTTTACCTCCGATTACTACCAAC TGTATTCTACCCAGCTGAGCACCGACACCGGCGTGG AACACGTGACCTTCTTCATCTACAACAAGATCGTGG ACGAGCCTGAGGAACACGTGCAGATCCACACTATC GACGGCAGCTCTGGCGTTGTGAACCCTGTGATGGAA CCCATCTACGATGAGCCCACCACAACAACCTCCGTG CCCCTGTaaGTGTttacctgttaatgtagcatttgagctttgggctaagcgcaa cattaaaccagtaccagaggtgaaaataactcaataatttgggtgtggacattgctgcta atactgtgatctgggactacaaaagagatgctccagcacatatactactattggtgttg ttctatgactgacatagccaagaacaaactgaaacgatttgtgcaccactcactgtctt tttgatggttagagttgatggtcaagtagacttattagaatgccgtaattggtgttcta ttacagaaggtagtgttaaggtttacaaccatctgtaggtcccaacaagctagtctta atggagtcacattaattggagaagccgtaaaaacacagttaattattataagaagttg atggtgtgtccaacaattacctgaaacttacttactcagagtagaaatttacaagaatt aaaccaggagtgcaaatggaattgatttcttagaattagctatggatgaattcattgaa cggataaattagaaggctatgccttcgaacatacgtttatggagatttttagtcatagtc agttaggtggtGCGAaattgtgtgtt
CoVEG15 expression cassette	47	ATTAAAGGTTTATACCTTCCCAGGTAACAAACCAAC CAACTTTCGATCTCTTGTAGATCTGTTCTCTAAACGA ACTTTAAATCTGTGTGGCTGTCACTCGGCTGCATG CTTAGTGCACTCACGCAGTATAATTAATAACTAATT ACTGTCGTTGACAGGACACGAGTAACTCGTCTATCT TCTGCAGGCTGCTTACGGTTTCGTCCGTGTTGCAGC CGATCATCAGCACATCTAGGTTTCGTCCGGGTGTGA CCGAAAGGTAAATGGCCGACAGCAACGGCACAATC ACCGTGGAAGAGCTGAAGAACTGCTGGAACAGTG GAACCTGGTCATCGGCTTCCTGTTTCTGACCTGGAT CTGTCTGCTGCAGTTCGCTTATGCCAATCGGAACAG ATTCCTGTACATCATCAAGCTGATCTTCCTGTGGCT GCTGTGGCCTGTGACCCTGGCTTGCTTCGTGCTGGC CGCTGTGTACCGGATCAACTGGATCACAGGCGGAA TCGCCATCGCCATGGCCTGCCTGGTGGGCCTGATGT GGCTGAGCTACTTCATCGCTTCTTTCAGACTGTTTCG CAGAACCCGGAGCATGTGGTCCTTCAACCCCGAGA CAAACATCCTGCTGAACGTGCCTCTGCACGGCACCA TCCTGACAAGACCTCTGCTCGAGAGCGAGCTGGTGA TTGGCGCAGTGATTCTGAGAGGCCATCTGAGGATCG CCGGACACCACCTGGGCAGATGCGACATCAAGGAC

		CTTCCAAAGGAAATCACCGTTGCCACCAGCCGGACC CTGTCCTACTACAAACTGGGCGCCAGCCAAAGAGT GGCCGGCGATAGCGGCTTTGCCGCCTACAGCAGAT ACCGCATCGGAAATTACAAGCTCAACACCGACCAC AGCAGCTCTTCTGATAACATCGCCCTGCTGGTGCAG CGAAAACGGCGCggaagcggaggaagcggagctactaactcagcctg ctgaagcaggctggagatgtggaggagaaccctggacctATGAGCGACA ACGGCCCTCAAAACCAGAGAAATGCCCTCGGATC ACATTTGGCGGACCTAGCGACAGCACCGGCAGCAA CCAGAATGGAGAAAGAAGCGGCGCCAGATCCAAGC AGCGGAGACCTCAGGGACTGCCCAACAACACCGCT AGCTGGTTCACCGCCCTGACCCAACACGGCAAGGA AGATCTGAAGTTCCCCAGAGGCCAGGGCGTGCCTAT CAACACAAACTCTTCTCCCGACGACCAGATCGGATA CTATAGACGGGCCACTCGGAGAATTGCGGGCGGCG ACGGAAAAATGAAGGACCTTTCTCCAAGATGGTAC TTCTACTACCTCGGCACAGGCCCTGAGGCCGGCCTG CCTTACGGCGCCAACAAGGATGGCATCATCTGGGTC GCCACCGAGGGCGCCCTGAACACCCCTAAGGACCA CATCGGCACAAGAAACCCCGCTAACAACGCCGCAA TCGTGCTGCAGCTGCCTCAGGGCACCACCCTGCCCA AGGGCTTCTACGCCGAGGGCTCTAGAGGTGGCTCCC AGGCTTCTAGCCGCTCCTCCAGCCGCAGCAGAAACA GCAGCAGGAACAGCACCCCGGCAGCTCCCGGGGC ACCAGCCCCGCCAGAATGGCCGGAATGGCGGCGA TGCCGCCCTGGCCCTGCTCCTGCTGGACAGACTGAA TCAGCTGGAAAGCAAGATGAGCGGCAAAGGACAGC AGCAGCAAGGCCAGACCGTGACCAAGAAAAGCGCT GCTGAAGCCTCCAAGAAACCTAGACAAAAGCGGAC CGCCACAAAGGCCTACAACGTGACCCAAGCCTTTG GAAGAAGAGGCCCGAGCAGACACAGGGCAATTTT GGCGACCAGGAGCTGATCCGGCAGGGAACCGACTA CAAGCACTGGCCTCAGATCGCCCAGTTGCCCCCTAG CGCCAGCGCCTTCTTCGGCATGAGCAGAATCGGCAT GGAAGTGACCCCTTCTGGCACCTGGCTGACCTACAC CGGCGCTATCAAGCTGGACGATAAGGATCCTAACTT CAAGGACCAAGTGATCCTGCTGAACAAGCATATCG ACGCCTATAAGACCTTTCCACCTACAGAGCCTAAGA AAGATAAGAAGAAGAAAGCCGACGAGACACAGGC CCTGCCTCAGAGACAGAAAAAGCAGCAGACAGTGA CACTGCTGCCAGCCGCTGACCTGGATGACTTCAGCA AGCAGCTGCAGCAGAGCATGTCTTCTGCTGATAGCA CCCAGGCCCGAAAACGGCGCggaagcggaggaagcggagctac taactcagcctgctgaagcaggctggagatgtggaggagaaccctggacctATG TTCGTGTTCTGGTGCTGCTGCCTCTGGTCAGCTCCC AGTGTGTGAACCTGACCACCAGAACCCAGCTGCCA CCTGCTTATACAAACTCCTTCACTCGGGGGGTATAC TACCCCGACAAGGTGTTTCAAGATCTAGCGTGCTGCAT TCTACACAAGACCTGTTTCTGCCCTTCTTACAGCAAC GTGACCTGGTTCACGCCATCCACGTGTCTGGAACC
--	--	--

		AACGGAACCAAGAGATTTCGACAACCCCGTGCTGCC TTTCAACGACGGCGTGTACTTCGCCAGCACCCGAGAA GTCCAACATCATCAGAGGATGGATTTTCGGCACCAC ACTGGACAGCAAAACCCAGAGCCTGCTGATCGTGA ACAACGCCACCAACGTGGTGATCAAGGTGTGCGAG TTCCAGTTCTGCAATGATCCCTTCCTGGGCGTGTACT ACCACAAGAACAACAAGTCTTGGATGGAAAGCGAG TTCAGAGTGTATTCCAGCGCCAACAATTGCACCTTC GAGTACGTGAGCCAACCCTTTCTGATGGACCTTGAA GGCAAGCAGGGCAACTTCAAAAATCTGCGAGAATT TGTGTTCAAGAACATCGACGGATACTTCAAGATCTA CTCTAAGCACACGCCAATCAACCTGGTGAGAGATCT GCCCCAGGGCTTTAGCGCTTTGGAACCTCTGGTGGA CCTGCCTATCGGAATCAACATCACCAGATTTCAAAC TCTCCTGGCCCTGCACAGATCTTATCTGACCCCTGG GGACAGTAGTAGCGGCTGGACAGCCGGCGCCGCCG CCTACTACGTGGGATACCTGCAGCCTAGAACATTCC TGCTGAAGTACAATGAGAACGGAACAATCACAGAC GCCGTGGACTGCGCCCTGGATCCTTTGAGCGAGACA AAGTGCACCCTGAAGTCGTTACCGTCGAAAAAGG CATCTACCAGACCAGCAACTTCCGCGTGCAGCCTAC GGAATCTATCGTGCGGTTCCCCAACATCACCAACCT GTGCCCTTTTCGGCGAGGTGTTTAAACGCTACAAGGT CGCCAGCGTGTATGCCTGGAACAGAAAGAGAATCA GCAATTGCGTGGCCGATTATAGCGTTCTGTACAACA GCGCTTCCTTCAGCACCTTCAAGTGCTACGGCGTGT CTCCAACCAAGCTGAACGACCTCTGCTTCACCAATG TCTACGCTGACTCTTTCGTGATTAGAGGCGATGAGG TTAGACAGATCGCACCTGGCCAGACCGGCAAAATC GCTGACTACAACCTACAAGCTGCCTGATGACTTCACA GGCTGTGTCAATTGCCTGGAACCTCAAATAACCTGGAC TCTAAAGTGGGCGGCAACTACAACCTGTACCGG CTGTTCCGGAAGAGCAATCTGAAACCTTTTGAGCGG GACATCTCTACAGAGATCTACCAGGCCGGCAGCAC ACCCTGCAACGGCGTTGAGGGCTTCAACTGCTACTT CCCTCTGCAGAGCTACGGCTTTCAGCCAACAAATGG AGTGGGCTACCAGCCGTACAGAGTGGTGGTGCTGA GCTTCGAACTGCTGCATGCCCCAGCCACAGTGTGTG GACCTAAGAAGTCTACCAACCTGGTGAAGAACAAG TGCGTGAACCTTTAACTTTAACGGCCTGACCGGCACA GGCGTGCTGACCGAATCCAACAAAAAGTTCTGCCC TTCCAACAGTTCGGCAGAGACATCGCCGATACAACC GATGCCGTGCGGGACCCCCAGACCTTAGAAATCCTA GATATCACCCCGTGACGCTTCGGCGGAGTCTCTGTT ATTACTCCTGGCACCAACACCAGCAACCAAGTGGCT GTTCTGTACCAAggcGTGAACTGCACCGAAGTGCCTG TGGCTATCCACGCCGATCAGCTGACCCCAACCTGGC GGGTGTATAGCACCGGCTCTAACGTGTTCCAGACCC GGGCTGGCTGCCTGATCGGCGCCGAACACGTCAAC AACTCCTATGAATGTGACATCCCCATCGGGGCTGGC
--	--	--

	ATCTGCGCCAGTTACCAGACACAGACAAATAGCCCT AGACGGGCCAGAAAGCGTGGCCTCCCAGAGTATCAT TGCCTACACCATGAGCCTGGGCGCCGAGAACAGCG TGGCCTATTCTAACAATAGCATCGCAATCCCTACCA ACTTTACCATCTCTGTGACAACCGAGATCCTGCCTG TGAGCATGACCAAAACCAGCGTGGACTGCACGATG TACATCTGTGGCGACAGCACAGAATGCAGTAATCTG TTGCTGCAGTACGGCAGCTTTTGCACCCAGTTGAAT AGAGCCCTGACCGGAATCGCCGTAGAGCAGGACAA AAATACCCAGGAGGTGTTGCCCCAGGTGAAACAGA TCTACAAGACACCTCCCATTAAGGACTTCGGAGGTT TTAACTTCAGCCAGATCCTGCCCCAGCCCTTCCAAGC CTAGCAAACGCTCCTTCATCGAGGACCTGCTCTTCA ACAAGGTGACACTGGCTGATGCCGGCTTCATCAAGC AGTACGGAGATTGTCTGGGAGACATCGCCGCTAGA GATCTGATCTGCGCCAAAAGTTCAACGGCCTGACC GTGCTGCCTCCTCTGCTTACAGACGAGATGATCGCC CAGTACACCAGCGCCCTGCTGGCTGGCACCATCACA AGCGGCTGGACCTTCGGAGCCGGAGCCGCTCTGCA AATCCCCTTTGCCATGCAGATGGCCTACCGGTTCAA CGGCATCGGCGTGACACAGAATGTGCTGTACGAGA ACCAGAAGCTGATCGCTAACCAGTTTAAACAGCGCTA TCGGCAAGATCCAGGACTCGCTGAGTAGCACC GCCT CTGCCCTGGGCAAGCTGCAGGACGTCGTGAACCA AACGCCCAAGCCCTGAACACACTGGTGAACAGCT GAGCAGCAACTTCGGCGCCATCAGCTCTGTGCTGAA CGATATCCTGAGCAGACTGGACAAGGTGGAAGCCG AGGTCCAGATCGACAGACTGATCACAGGAAGACTG CAGAGCCTGCAAACGTACGTGACACAGCAGCTGAT CCGGGCAGCCGAAATCCGGGCCAGCGCCAATCTGG CCGCTACCAAGATGAGCGAGTGCGTGTTAGGCCAG AGCAAGCGGGTGGATTTCTGCGGTAAGGGATACCA CCTGATGAGCTTTCCCCAGAGCGCTCCTCACGGCGT GGTGTCTTCTGCACGTGACCTACGTTCTGCCCAGGA AAAGAACTTCACCACCGCCCCTGCTATCTGCCACGA TGGCAAGGCCCACTTCCCTAGAGAGGGCGTTTTCGT GTCTAACGGCACACACTGGTTTGTGACCCAGAGAA ACTTCTACGAGCCTCAGATCATCACACAGACAACA CCTTTGTGAGCGGCAATTGCGACGTGGTGATCGGAA TTGTTAATAATACCGTGTACGACCTCTGCAGCCTG AGCTCGACAGCTTCAAGGAAGAGCTGGACAAGTAC TTCAAGAACCACACCTCCCCAGATGTGGACCTGGGC GATATTTACAGGCATCAACGCCTCCGTGCGTAATATC CAGAAGGAGATCGACCGGCTCAACGAGGTGGCCAA GAACCTTAACGAGAGCCTGATCGACCTGCAGGAAC TGGGCAAAATATGAGCAGTACATCAAGTGGCCTTGGT ACATCTGGCTGGGCTTTATCGCAGGCCTGATCGCTA TCGTGATGGTGACCATTATGCTGTGTTGTATGACCA GCTGTTGTAGTTGTCTGAAGGGCTGCTGTTCTTGCG GCAGCTGCTGCAAGTTCGACGAAGACGACTCAGAG
--	--

		CCCGTGCTGAAAGGCGTGAAGCTGCACTACACCCG AAAACGGCGCggaagcggaggaagcggagctactaacttcagcctgctga agcaggctggagatgtggaggagaaccctggacctATGTATTCTTTTGT GTCCGAGGAAACCGGCACACTGATCGTTAATAGCG TGCTGCTCTTCCTGGCCTTCGTGGTGTTCCTGCTGGT GACCCTGGCTATCCTGACCGCCCTGAGACTGTGTGC CTACTGCTGCAACATCGTGAACGTGTCTCTGGTCAA GCCTAGCTTCTACGTGTACAGCCGGGTGAAGAACCT GAACAGCAGCAGAGTGCCCGACCTGCTGGTGaatcccc ccccctaacgttactggccgaagccgcttgaataaggccggtgtgcgtttgtctata tgtattttccaccatattgccgtcttttgcaatgtgagggcccgaaacctggccctgt cttcttgacgagcattcctaggggtctttccctctcgcaaaggatgaaggtctgtt gaatgtcgtgaaggagcagttcctctggaagcttctgaagacaaacaacgtctgta gcgacccttgcaggcagcggaacccccacctggcgacaggtgcctctcgggcca aaagccacgtgtataagatacacctgcaaaggcggcacaacccagtgccacgttgt gagttggatagttgtgaaagagtc aaatggctctcctcaagcgtattcaacaagggg ctgaaggatgccagaaggtacccattgtatgggatctgatctggggcctcggtgca catgctttacatgtgtttagtcgaggttaaaaaacgtctagccccccgaaccacggg gacgtgggtttccttgaaaaacacgatgataaGTGTtacctgttaatgtagcatttg agctttgggctaagcgcaacattaaaccagtagcagaggtgaaaatactcaataattg ggtgtggacattgctgctaatactgtgatctgggactacaaaagagatgctccagcac atatatctactattggtgtttgtctatgactgacatagccaagaaaccaactgaaacgat ttgtgcaccactactgtctttttgatggtagagttgatggtcaagtagacttatttagaa atgcccgtaatggtgttctattacagaaggtagtgtaaagggttaacaacctctgtagg tcccaaacaagctagtctaatggagtcacattaattggagaagccgtaaaaacacagt tcaattattataagaaagttgatggtgtgtccaacaattacctgaaacttactttactcag agtagaaatttacaagaatttaaacccaggagtc aaatggaaattgatttcttagaatta gctatggatgaattcattgaacggtataaattagaaggctatgccttcgaacatagcttt atggagattttagtcatagtcagtagtggtGCGAaattgtgtgtt
CoVEG16 expression cassette	48	ATGGCCGACAGCAACGGCACAAATCACCGTGGAAGA GCTGAAGAAACTGCTGGAACAGTGGAACCTGGTCA TCGGCTTCCTGTTTCTGACCTGGATCTGTCTGCTGCA GTTTCGCTTATGCCAATCGGAACAGATTCTGTACAT CATCAAGCTGATCTTCCTGTGGCTGCTGTGGCCTGT GACCCTGGCTTGCTTCGTGCTGGCCGCTGTGTACCG GATCAACTGGATCACAGGCGGAATCGCCATCGCCA TGGCCTGCCTGGTGGGCTGATGTGGCTGAGCTACT TCATCGCTTCTTTCAGACTGTTTCGCCAGAACCCGGA GCATGTGGTCCTTCAACCCCGAGACAAACATCCTGC TGAACGTGCCTCTGCACGGCACCATCCTGACAAGAC CTCTGCTCGAGAGCGAGCTGGTGATTGGCGCAGTGA TTCTGAGAGGCCATCTGAGGATCGCCGACACCACC TGGGCAGATGCGACATCAAGGACCTTCCAAGGAA ATCACCGTTGCCACCAGCCGGACCCTGTCCTACTAC AACTGGGCGCCAGCCAAAGAGTGGCCGCGGATAG CGGCTTTGCCGCTACAGCAGATACCGCATCGGAAA TTACAAGCTCAACACCGACCACAGCAGCTCTTCTGA TAACATCGCCCTGCTGGTGCAGCGAAAACGGCGCgg aagcggaggaagcggagctactaacttcagcctgctgaagcaggctggagatgtgg

	aggagaaccctggacctATGAGCGACAACGGCCCTCAAAAC CAGAGAAATGCCCCCTCGGATCACATTTGGCGGACCT AGCGACAGCACCGGCAGCAACCAGAATGGAGAAAG AAGCGGCGCCAGATCCAAGCAGCGGAGACCTCAGG GACTGCCCAACAACACCGCTAGCTGGTTCACCGCCC TGACCCAACACGGCAAGGAAGATCTGAAGTTCCCC AGAGGCCAGGGCGTGCCTATCAACACAACTCTTCT CCCGACGACCAGATCGGATACTATAGACGGGCCAC TCGGAGAATTCTGGGGCGGCGACGGAAAAATGAAGG ACCTTTCTCCAAGATGGTACTTCTACTACCTCGGCA CAGGCCCTGAGGCCGGCCTGCCTTACGGCGCCAAC AAGGATGGCATCATCTGGGTGCGCACCGAGGGCGC CCTGAACACCCCTAAGGACCACATCGGCACAAGAA ACCCCGCTAACAACGCCGCAATCGTGCTGCAGCTGC CTCAGGGCACCACCCTGCCCAAGGGCTTCTACGCCG AGGGCTCTAGAGGTGGCTCCCAGGCTTCTAGCCGT CCTCCAGCCGCAGCAGAAACAGCAGCAGGAACAGC ACCCCGGCAGCTCCCGGGGCACCAGCCCCGCCAG AATGGCCGGAAATGGCGGCGATGCCGCCCTGGCCC TGCTCCTGCTGGACAGACTGAATCAGCTGGAAAGC AAGATGAGCGGCAAAGGACAGCAGCAGCAAGGCC AGACCGTGACCAAGAAAAGCGCTGCTGAAGCCTCC AAGAAACCTAGACAAAAGCGGACCGCCACAAAGGC CTACAACGTGACCCAAGCCTTTGGAAGAAGAGGCC CCGAGCAGACACAGGGCAATTCGGCGACCAGGAG CTGATCCGGCAGGGAACCGACTACAAGCACTGGCC TCAGATCGCCCAGTTCGCCCCTAGCGCCAGCGCCTT CTTCGGCATGAGCAGAATCGGCATGGAAGTGACCC CTTCTGGCACCTGGCTGACCTACACCGGCGCTATCA AGCTGGACGATAAGGATCCTAACTTCAAGGACCAA GTGATCCTGCTGAACAAGCATATCGACGCCTATAAG ACCTTTCCACCTACAGAGCCTAAGAAAGATAAGAA GAAGAAAGCCGACGAGACACAGGCCCTGCCTCAGA GACAGAAAAAGCAGCAGACAGTGACACTGCTGCCA GCCGCTGACCTGGATGACTTCAGCAAGCAGCTGCA GCAGAGCATGTCTTCTGCTGATAGCACCCAGGCCCCG AAAACGGCGCggaagcggaggaagcggagctactaacttcagcctgctga agcaggctggagatgtggaggagaaccctggacctATGTTCTGTGTTCTT GGTGCTGCTGCCTCTGGTCAGCTCCCAGTGTGTGAA CCTGACCACCAGAACCCAGCTGCCACCTGCTTATAC AAACTCCTTCACTCGGGGGGTATACTACCCCGACAA GGTGTTTCAGATCTAGCGTGCTGCATTCTACACAAGA CCTGTTCTGCCCCTTCTTCAGCAACGTGACCTGGTTC CACGCCATCCACGTGTCTGGAACCAACGGAACCAA GAGATTCGACAACCCCGTGCTGCCTTTCAACGACGG CGTGTACTTCGCCAGCACCGAGAAGTCCAACATCAT CAGAGGATGGATTTTCGGCACCACTGGACAGCA AAACCCAGAGCCTGCTGATCGTGAACAACGCCACC AACGTGGTGATCAAGGTGTGCGAGTTCCAGTTCTGC AATGATCCCTTCTGCGGCGTGTACTACCACAAGAAC
--	--

		AACAAGTCTTGGATGGAAAGCGAGTTCAGAGTGTATCCAGCGCCAACAATTGCACCTTCGAGTACGTGAGCCAACCCTTTCTGATGGACCTTGAAGGCAAGCAGGGCAACTTCAAAAATCTGCGAGAATTTGTGTTCAAGAACATCGACGGATACTTCAAGATCTACTCTAAGCACACGCCAATCAACCTGGTGAGAGATCTGCCCCAGGGCTTAGCGCTTTGGAACCTCTGGTGGACCTGCCTATCGGAATCAACATCACCAGATTTCAAACCTCTCCTGGCCCTGCACAGATCTTATCTGACCCCTGGGGACAGTAGTAGCGGCTGGACAGCCGGCGCCGCCGCTACTACGTGGATACCTGCAGCCTAGAACATTCCTGCTGAAGTACAATGAGAACGGAACAATCACAGACGCCGTGGACTGCGCCCTGGATCCTTTGAGCGAGACAAAGTGCACCCTGAAGTCGTTACCGTTCGAAAAAGGCATCTACCAGACAGCAACTTCCGCGTGCAGCCTACGGAATCTATCGTGCGGTTCCCCAACATCACCAACCTGTGCCCTTTTCGCGAGGTGTTTAACGCTACAAGGTTCCGCCAGCGTGATGCCTGGAACAGAAAGAGAATCAGCAATTGCGTGGCCGATTATAGCGTTCTGTACAACAGCGCTTCCTTCAGCACCTTCAAGTGCTACGGCGTGTCTCCAACCAAGCTGAACGACCTCTGCTTCACCAATGTCTACGCTGACTCTTTCGTGATTAGAGGCGATGAGGTTAGACAGATCGCACCTGGCCAGACCGGCAAAATCGCTGACTACAAC TACAAGCTGCCTGATGACTTCACAGGCTGTGTCAATGCCTGGAAC TCAAATAACCTGGACTCTAAAGTGGGCGGCAACTACAAC TACCTGTACCGGCTGTTCCGGAA GAGCAATCTGAAACCTTTTGAGCGGGACATCTCTACAGAGATCTACCAGGCCGGCAGCACACCCTGCAACG GCGTTGAGGGCTTCAACTGCTACTTCCCTCTGCAGAGCTACGGCTTTCAGCCAACAAATGGAGTGGGCTACCAGCCGTACAGAGTGGTGGTGTGCTGAGCTTCGAACTGCTGCATGCCCCAGCCACAGTGTGTGGACCTAAGAAGTCTACCAACCTGGTGAAGAACAAGTGCGTGAAC TTTAACTTTAACGGCCTGACCGGCACAGGCGTGTGACCGAATCCAACAAAAAGTTCCTGCCCTTCCAACAGTTCG GCAGAGACATCGCCGATACAACCGATGCCGTGCGG GACCCCCAGACCTTAGAAATCCTAGATATCACCCCCG TGCAGCTTCGGCGGAGTCTCTGTTATTACTCCTGGC ACCAACACCAGCAACCAAGTGGCTGTTCTGTACCAAGgcGTGAAC TGCACCGAAGTGCCTGTGGCTATCCACG CCGATCAGCTGACCCCAACCTGGCGGGTGTATAGCA CCGGCTCTAACGTGTTCCAGACCCGGGCTGGCTGCC TGATCGGCGCCGAACACGTCAACAACCTCCTATGAATGTGACATCCCCATCGGGGCTGGCATCTGCGCCAGTTACCAGACACAGACAAATAGCCCTGGCAGCGCCAGCAGCGTGGCCTCCCAGAGTATCATTGCCTACACCATGAGCCTGGGCGCCGAGAACAGCGTGGCCTATTCTAAC AATAGCATCGCAATCCCTACCAACTTTACCATCTCTGTGACAACCGAGATCCTGCCTGTGAGCATGACCAA AACCAGCGTGGACTGCACGATGTACATCTGTGGCG
--	--	---

		ACAGCACAGAATGCAGTAATCTGTTGCTGCAGTACG GCAGCTTTTGCACCCAGTTGAATAGAGCCCTGACCG GAATCGCCGTAGAGCAGGACAAAAATACCCAGGAG GTGTTTCGCCCAGGTGAAACAGATCTACAAGACACCT CCCATTAAAGGACTTCGGAGGTTTTAACTTCAGCCAG ATCCTGCCCCGACCCTTCCAAGCCTAGCAAACGCTCC TTCATCGAGGACCTGCTCTTCAACAAGGTGACACTG GCTGATGCCGGCTTCATCAAGCAGTACGGAGATTGT CTGGGAGACATCGCCGCTAGAGATCTGATCTGCGCC CAAAAGTTCAACGGCCTGACCGTGCTGCCTCCTCTG CTTACAGACGAGATGATCGCCAGTACACCAGCGC CCTGCTGGCTGGCACCATCACAAGCGGCTGGACCTT CGGAGCCGGAGCCGCTCTGCAAATCCCCTTTGCCAT GCAGATGGCCTACCGGTTCAACGGCATCGGCGTGA CACAGAATGTGCTGTACGAGAACCAGAAGCTGATC GCTAACCAGTTTAAACAGCGCTATCGGCAAGATCCAG GACTCGCTGAGTAGCACCGCCTCTGCCCTGGGCAAG CTGCAGGACGTCGTGAACCAGAACGCCCAAGCCCT GAACACACTGGTGAAACAGCTGAGCAGCAACTTCG GCGCCATCAGCTCTGTGCTGAACGATATCCTGAGCA GACTGGACCCTcccGAAGCCGAGGTCCAGATCGACA GACTGATCACAGGAAGACTGCAGAGCCTGCAAACG TACGTGACACAGCAGCTGATCCGGGCAGCCGAAAT CCGGGCCAGCGCCAATCTGGCCGCTACCAAGATGA GCGAGTGCGTGTTAGGCCAGAGCAAGCGGGTGGAT TTCTGCGGTAAGGGATACCACCTGATGAGCTTTCCC CAGAGCGCTCCTCACGGCGTGGTGTCTGCACGTG ACCTACGTTTCTGCCAGGAAAAGAACTTACCACC GCCCCTGCTATCTGCCACGATGGCAAGGCCCACTTC CCTAGAGAGGGCGTTTTTCGTGTCTAACGGCACACAC TGGTTTGTGACCCAGAGAACTTCTACGAGCCTCAG ATCATCACCACAGACAACACCTTTGTGAGCGGCAAT TGCGACGTGGTGATCGGAATTGTTAATAATACCGTG TACGACCCTCTGCAGCCTGAGCTCGACAGCTTCAAG GAAGAGCTGGACAAGTACTTCAAGAACCACACCTC CCCAGATGTGGACCTGGGCGATATTTAGGCATCAA CGCCTCCGTCGTGAATATCCAGAAGGAGATCGACC GGCTCAACGAGGTGGCCAAGAACCTTAACGAGAGC CTGATCGACCTGCAGGAACTGGGCAAATATGAGCA GTACATCAAGTGGCCTTGGTACATCTGGCTGGGCTT TATCGCAGGCCTGATCGCTATCGTGATGGTGACCAT TATGCTGTGTTGTATGACCAGCTGTTGTAGTTGTCTG AAGGGCTGCTGTTCTTGCGGCAGCTGCTGCAAGTTC GACGAAGACGACTCAGAGCCCCTGCTGAAAGGCGT GAAGCTGCACTACACCCGAAAACGGCGCggaagcggag gaagcggagctactaactcagcctgctgaagcaggctggagatgtggaggagaac cctggacctATGTATTCTTTTGTGTCCGAGGAAACCGGC ACACTGATCGTTAATAGCGTGCTGCTCTTCCTGGCC TTCGTGGTGTTCCTGCTGGTGACCCTGGCTATCCTG ACCGCCCTGAGACTGTGTGCCTACTGCTGCAACATC
--	--	--

		GTGAACGTGTCTCTGGTCAAGCCTAGCTTCTACGTG TACAGCCGGGTGAAGAACCTGAACAGCAGCAGAGT GCCCCACCTGCTGGTGtaatccccccccctaactgtactggccgaag ccgcttggaataaggccggtgtgctgttctatgttatttccaccatattgccgtctt tggcaatgtgagggcccgaaacctggccctgtcttcttgacgagcattcctaggggt ctttccctctcgcaaaaggaatgcaaggtctgttgatgtcgtgaaggaaagcagtcc tctggaagcttctgaagacaaacaacgtctgtagcgacccttgcaggcagcggaaac ccccacctggcgacaggtgcctctgcggccaaaagccacgtgtataagatacacct gcaaaggcggcacaacccagtgccacgttgtgagttgtagattgtggaagagtc aatggctctcctcaagcgtattcaacaaggggctgaaggatgccagaaggtaccc cattgtatgggatctgatctggggcctcgggtgcacatgctttacatgtgtttatgcgaggt taaaaaacgtctaggccccccgaaccacggggacgtggtttccttgaaaaacacg atgataatatggccacaacctgaacaagagacttgcgcgactctctcacttttgag gaatgccccaaatgctctgctctacaataaccgtaattgatttacctgctaaagtatgat gaagaatggtaccagaggagtattgactgatggagaggatgatgtcttggatcccg aattagacatggaagtcgtttcgagtacagggaagcggagctactaacttgcacctg ctgaagcaggctggagatgtggaggagaaccctggacctATGGACCTGTT CATGAGAATCTTCACCATCGGCACCGTGACACTGAA GCAGGGCGAGATCAAGGATGCCACCCCTAGCGACT TCGTGAGAGCCACCGCCACAATTCCTATCCAGGCTA GCCTGCCTTTTGGATGGCTGATCGTGGGCGTCGCCC TGCTCGCCGTGTTCCAGAGCGCCTCTAAGATCATTA CACTGAAGAAAAGATGGCAGCTGGCCCTCTCCAAA GGCGTGCACTTCGTGTGTAATCTGCTGCTGCTTTTTG TGACAGTGTACAGCCACCTGCTGCTGGTTGCTGCTG GCCTGGAAGCCCCCTTCCTGTACCTGTACGCCCTGG TCTACTTCCTGCAGTCTATCAACTTCGTGCGGATCAT CATGCGGCTGTGGCTGTGCTGGAAGTGCAGAAGCA AGAACCCACTGCTGTACGACGCCAATTACTTCCTGT GTTGGCACACCAACTGCTACGACTACTGCATCCCCCT ACAACAGCGTGACCAGCAGCATCGTGATCACCTCTG GCGACGGAACAACCAGCCCTATCAGCGAGCATGAT TACCAGATCGGCGGATATACAGAGAAGTGGGAGAG CGGCGTGAAGGACTGCGTGGTGCTGCACAGCTACTT TACCTCCGATTACTACCAACTGTATTCTACCCAGCT GAGCACCGACACCGGCGTGGAACACGTGACCTTCTT CATCTACAACAAGATCGTGGACGAGCCTGAGGAAC ACGTGCAGATCCACACTATCGACGGCAGCTCTGGCG TTGTGAACCTGTGATGGAACCCATCTACGATGAGC CCACCACAACAACCTCCGTGCCCCCTGTaaGTGTttacctg ttaatgtagcatttgagcttgggctaagcgcaacattaaaccagtaggaggtgaaa atactcaataatttgggtgtggacattgtgctaatactgtgatctgggactcaaaaaga gatgtccagcacatatatctactattgggtgttgttctatgactgacatagccaagaaac caactgaaacgattgtgcaccactcactgtctttttgatggtagagttgatggtaagt agacttatttagaatgcccgaatgggtgttcttattacagaaggtaggttaaagggttac aaccatctgtagggtcccaacaagctagtcttaattggagtcacattaattggagaagcc gtaaaacacagttcaattattataagaagttgatgggtgttccaacaattacctgaaa cttacttactcagagtagaaattacaagaattaaaccaggaggtcaaatggaattg atttctagaattagctatggatgaattcattgaacgggtataaattagaaggctatgccttc
--	--	--

		gaacatatcgtttatggagatttttagtcatagtcagttaggtggtGCGAaattgtgtgtggt
CoVEG17 expression cassette	49	ATTAAAGGTTTATACCTTCCCAGGTAACAAACCAAC CAACTTTCGATCTCTTGTAGATCTGTTCTCTAAACGA ACTTTAAAATCTGTGTGGCTGTCACTCGGCTGCATG CTTAGTGCACCTCACGCAGTATAATTAATAACTAATT ACTGTCGTTGACAGGACACGAGTAACCTCGTCTATCT TCTGCAGGCTGCTTACGGTTTCGTCCGTGTTGCAGC CGATCATCAGCACATCTAGGTTTCGTCCGGGTGTGA CCGAAAGGTAAATGGCCGACAGCAACGGCACAATC ACCGTGGAAGAGCTGAAGAACTGCTGGAACAGTG GAACCTGGTCATCGGCTTCCTGTTTCTGACCTGGAT CTGTCTGCTGCAGTTCGCTTATGCCAATCGGAACAG ATTCCTGTACATCATCAAGCTGATCTTCCTGTGGCT GCTGTGGCCTGTGACCCTGGCTTGCTTCGTGCTGGC CGCTGTGTACCGGATCAACTGGATCACAGGCGGAA TCGCCATCGCCATGGCCTGCCTGGTGGGCCTGATGT GGCTGAGCTACTTCATCGCTTCTTTCAGACTGTTTCG CAGAACCCGGAGCATGTGGTCCTTCAACCCCGAGA CAAACATCCTGCTGAACGTGCCTCTGCACGGCACCA TCCTGACAAGACCTCTGCTCGAGAGCGAGCTGGTGA TTGGCGCAGTGATTCTGAGAGGCCATCTGAGGATCG CCGGACACCACCTGGGCAGATGCGACATCAAGGAC CTTCCAAAGGAAATCACCGTTGCCACCAGCCGGACC CTGTCCTACTACAACTGGGCGCCAGCCAAAGAGT GGCCGGCGATAGCGGCTTTGCCGCCTACAGCAGAT ACCGCATCGGAAATTACAAGCTCAACACCGACCAC AGCAGCTCTTCTGATAACATCGCCCTGCTGGTGCAG CGAAAACGGCGCggaagcggaggaagcggagctactaacttcagcctg ctgaagcaggctggagatgtggaggagaaccctggacctATGAGCGACA ACGGCCCTCAAACCCAGAGAAATGCCCTCGGATC ACATTTGGCGGACCTAGCGACAGCACCGGCAGCAA CCAGAATGGAGAAAGAAGCGGCGCCAGATCCAAGC AGCGGAGACCTCAGGGACTGCCCAACAACACCGCT AGCTGGTTCACCGCCCTGACCCAACACGGCAAGGA AGATCTGAAGTTCCCCAGAGGCCAGGGCGTGCCTAT CAACACAACTCTTCTCCCGACGACCAGATCGGATA CTATAGACGGGCCACTCGGAGAATTCGGGGCGGCG ACGGAAAAATGAAGGACCTTTCTCCAAGATGGTAC TTCTACTACCTCGGCACAGGCCCTGAGGCCGGCCTG CCTTACGGCGCCAACAAGGATGGCATCATCTGGGTC GCCACCGAGGGCGCCCTGAACACCCCTAAGGACCA CATCGGCACAAGAAACCCCGCTAACAACGCCGCAA TCGTGCTGCAGCTGCCTCAGGGCACCAACCTGCCCA AGGGCTTCTACGCCGAGGGCTCTAGAGGTGGCTCCC AGGCTTCTAGCCGCTCCTCCAGCCGCAGCAGAAACA GCAGCAGGAACAGCACCCCGGCAGCTCCCGGGGC ACCAGCCCCGCCAGAATGGCCGGAAATGGCGGCGA TGCCGCCCTGGCCCTGCTCCTGCTGGACAGACTGAA

	TCAGCTGGAAAGCAAGATGAGCGGCAAAGGACAGC AGCAGCAAGGCCAGACCGTGACCAAGAAAAGCGCT GCTGAAGCCTCCAAGAAACCTAGACAAAAGCGGAC CGCCACAAAGGCCTACAACGTGACCCAAGCCTTTG GAAGAAGAGGCCCGAGCAGACACAGGGCAATTTC GGCGACCAGGAGCTGATCCGGCAGGGAACCGACTA CAAGCACTGGCCTCAGATCGCCCAGTTCGCCCTAG CGCCAGCGCCTTCTTCGGCATGAGCAGAATCGGCAT GGAAGTGACCCCTTCTGGCACCTGGCTGACCTACAC CGGCGCTATCAAGCTGGACGATAAGGATCCTAACTT CAAGGACCAAGTGATCCTGCTGAACAAGCATATCG ACGCCTATAAGACCTTTCCACCTACAGAGCCTAAGA AAGATAAGAAGAAGAAAGCCGACGAGACACAGGC CCTGCCTCAGAGACAGAAAAAGCAGCAGACAGTGA CACTGCTGCCAGCCGCTGACCTGGATGACTTCAGCA AGCAGCTGCAGCAGAGCATGTCTTCTGCTGATAGCA CCCAGGCCCGAAAACGGCGCggaagcggaggaagcggagctac taacttcagcctgctgaagcaggctggagatgtggaggagaaccctggacctATG TTCGTGTTCTGGTGCTGCTGCCTCTGGTCAGCTCCC AGTGTGTGAACCTGACCACCAGAACCCAGCTGCCA CCTGCTTATACAACTCCTTCACTCGGGGGGTATAC TACCCCGACAAGGTGTTTCAGATCTAGCGTGCTGCAT TCTACACAAGACCTGTTCTGCTGCTTCTTTCAGCAAC GTGACCTGGTTCACGCCATCCACGTGTCTGGAACC AACGGAACCAAGAGATTTCGACAACCCCGTGCTGCC TTTCAACGACGGCGTGACTTCGCCAGCACCCGAGAA GTCCAACATCATCAGAGGATGGATTTTCGGCACCAC ACTGGACAGCAAAACCCAGAGCCTGCTGATCGTGA ACAACGCCACCAACGTGGTGATCAAGGTGTGCGAG TTCCAGTTCTGCAATGATCCCTTCCTGGGCGTGACT ACCACAAGAACAACAAGTCTTGATGGAAGCGAG TTCAGAGTGTATTCCAGCGCCAACAATTGCACCTTC GAGTACGTGAGCCAACCCTTTCTGATGGACCTTGAA GGCAAGCAGGGCAACTTCAAAAATCTGCGAGAATT TGTGTTCAAGAACATCGACGGATACTTCAAGATCTA CTCTAAGCACACGCCAATCAACCTGGTGAGAGATCT GCCCCAGGGCTTTAGCGCTTTGGAACCTCTGGTGGA CCTGCCTATCGGAATCAACATCACCAGATTTCAAAC TCTCCTGGCCCTGCACAGATCTTATCTGACCCCTGG GGACAGTAGTAGCGGCTGGACAGCCGGCGCCGCCG CCTACTACGTGGGATACCTGCAGCCTAGAACATTCC TGCTGAAGTACAATGAGAACGGAACAATCACAGAC GCCGTGGACTGCGCCCTGGATCCTTTGAGCGAGACA AAGTGCACCCTGAAGTCGTTACCGTCGAAAAAGG CATCTACCAGACCAGCAACTTCCGCGTGACGCTAC GGAATCTATCGTGCGGTTCCCCAACATCACCAACCT GTGCCCTTTTCGGCGAGGTGTTTAACGCTACAAGGTT CGCCAGCGTGTATGCCTGGAACAGAAAGAGAATCA GCAATTGCGTGGCCGATTATAGCGTTCTGTACAACA GCGCTTCCTTCAGCACCTTCAAGTGCTACGGCGTGT
--	---

		CTCCAACCAAGCTGAACGACCTCTGCTTCACCAATG TCTACGCTGACTCTTTCGTGATTAGAGGCGATGAGG TTAGACAGATCGCACCTGGCCAGACCGGCAAAATC GCTGACTACAACCTACAAGCTGCCTGATGACTTCACA GGCTGTGTCATTGCCTGGAACCTCAAATAACCTGGAC TCTAAAGTGGGCGGCAACTACAACCTACCTGTACCGG CTGTTCCGGAAGAGCAATCTGAAACCTTTTGAGCGG GACATCTCTACAGAGATCTACCAGGCCGGCAGCAC ACCTTGCAACGGCGTTGAGGGCTTCAACTGCTACTT CCCTCTGCAGAGCTACGGCTTTCAGCCAACAAATGG AGTGGGCTACCAGCCGTACAGAGTGGTGGTGCTGA GCTTCGAACTGCTGCATGCCCCAGCCACAGTGTGTG GACCTAAGAAGTCTACCAACCTGGTGAAGAACAAG TGCGTGAACTTTAACTTTAACGGCCTGACCGGCACA GGCGTGCTGACCGAATCCAACAAAAAGTTCCTGCCC TTCCAACAGTTCGGCAGAGACATCGCCGATACAACC GATGCCGTGCGGGACCCCCAGACCTTAGAAATCCTA GATATCACCCCGTGCAGCTTCGGCGGAGTCTCTGTT ATTACTCCTGGCACCAACACCAGCAACCAAGTGGCT GTTCTGTACCAAggcGTGAACTGCACCGAAGTGCCTG TGGCTATCCACGCCGATCAGCTGACCCCAACCTGGC GGGTGTATAGCACCGGCTCTAACGTGTTCCAGACCC GGGCTGGCTGCCTGATCGGCGCCGAACACGTCAAC AACTCCTATGAATGTGACATCCCCATCGGGGCTGGC ATCTGCGCCAGTTACCAGACACAGACAAATAGCCCT GGCAGCGCCAGCAGCGTGGCCTCCCAGAGTATCATT GCCTACACCATGAGCCTGGGCGCCGAGAACAGCGT GGCCTATTCTAACAATAGCATCGCAATCCCTACCAA CTTTACCATCTCTGTGACAACCGAGATCCTGCCTGT GAGCATGACCAAAACCAGCGTGGACTGCACGATGT ACATCTGTGGCGACAGCACAGAATGCAGTAATCTGT TGCTGCAGTACGGCAGCTTTTGCACCCAGTTGAATA GAGCCCTGACCGGAATCGCCGTAGAGCAGGACAAA AATACCCAGGAGGTGTTTCGCCAGGTGAAACAGAT CTACAAGACACCTCCCATTAAGGACTTCGGAGGTTT TAACTTCAGCCAGATCCTGCCCCGACCCTTCCAAGCC TAGCAAACGCTCCTTCATCGAGGACCTGCTCTTCAA CAAGGTGACACTGGCTGATGCCGGCTTCATCAAGCA GTACGGAGATTGTCTGGGAGACATCGCCGCTAGAG ATCTGATCTGCGCCCCAAAAGTTCAACGGCCTGACCG TGCTGCCTCCTCTGCTTACAGACGAGATGATCGCCC AGTACACCAGCGCCCTGCTGGCTGGCACCATCACAA GCGGCTGGACCTTCGGAGCCGGAGCCGCTCTGCAA ATCCCCTTTGCCATGCAGATGGCCTACCGGTTCAAC GGCATCGGCGTGACACAGAATGTGCTGTACGAGAA CCAGAAGCTGATCGCTAACCAAGTTTAACAGCGCTAT CGGCAAGATCCAGGACTCGCTGAGTAGCACCGCCT CTGCCCTGGGCAAGCTGCAGGACGTCGTGAACAG AACGCCCAAGCCCTGAACACACTGGTGAACAGCT GAGCAGCAACTTCGGCGCCATCAGCTCTGTGCTGAA
--	--	--

		CGATATCCTGAGCAGACTGGACCCTcccGAAGCCGA GGTCCAGATCGACAGACTGATCACAGGAAGACTGC AGAGCCTGCAAACGTACGTGACACAGCAGCTGATC CGGGCAGCCGAAATCCGGGCCAGCGCCAATCTGGC CGTACCAAGATGAGCGAGTGCGTGTAGGCCAGA GCAAGCGGGTGGATTTCTGCGGTAAGGGATACCAC CTGATGAGCTTTCCCCAGAGCGCTCCTCACGGCGTG GTGTTTCTGCACGTGACCTACGTTCTGCCAGGAA AAGA ACTTCACCACCGCCCCCTGCTATCTGCCACGAT GGCAAGGCCCACTTCCCTAGAGAGGGCGTTTTCTGTG TCTAACGGCACACACTGGTTTGTGACCCAGAGAAAC TTCTACGAGCCTCAGATCATCACCACAGACAACACC TTTGTGAGCGGCAATTGCGACGTGGTGATCGGAATT GTTAATAATACCGTGTACGACCCTCTGCAGCCTGAG CTCGACAGCTTCAAGGAAGAGCTGGACAAGTACTT CAAGAACCACACCTCCCCAGATGTGGACCTGGGCG ATATTTCAAGCATCAACGCCTCCGTCGTGAATATCC AGAAGGAGATCGACCGGCTCAACGAGGTGGCCAAG AACCTTAACGAGAGCCTGATCGACCTGCAGGAAC GGGCAAATATGAGCAGTACATCAAGTGGCCTTGGT ACATCTGGCTGGGCTTTATCGCAGGCCTGATCGCTA TCGTGATGGTGACCATTATGCTGTGTTGTATGACCA GCTGTTGTAGTTGTCTGAAGGGCTGCTGTTCTTGCG GCAGCTGCTGCAAGTTTCGACGAAGACGACTCAGAG CCCGTGCTGAAAGGCGTGAAGCTGCACTACACCCG AAAACGGCGCGggaagcggaggaagcggagctactaacttcagcctgctga agcaggctggagatgtggaggagaaccctggacctATGTATTCTTTTGT GTCCGAGGAAACCGGCACACTGATCGTTAATAGCG TGCTGCTCTTCTGGCCTTCGTGGTGTTCCTGCTGGT GACCCTGGCTATCCTGACCGCCCTGAGACTGTGTGC CTACTGCTGCAACATCGTGAACGTGTCTCTGGTCAA GCCTAGCTTCTACGTGTACAGCCGGGTGAAGAACCT GAACAGCAGCAGAGTGCCCCGACCTGCTGGTGaatcccc ccccctaactgtactggccgaagccgcttggataagccggtgtgcgtttgtctata tgtattttccaccatattgccgtcttttggcaatgtgagggcccggaacctggccctgt cttcttgacgagcattctaggggtctttccctctcgccaaaggaatgcaaggctgtt gaatgtcgtgaaggaagcagttcctctggaagcttctgaagacaaacaacgtctgta gagacctttgcaggcagcgggaacccccacctggcgacaggtgcctctgcggcca aaagccacgtgtataagatacacctgcaaaggcggcacaacccccagtgccacgttgt gagttggatagttgtgaaagagtcaaatggctctcctcaagcgtattcaacaagggg ctgaaggatgccagaaggtacccattgtatgggatctgatctggggcctcggtgca catgctttacatgtgttagtcgaggttaaaaaacgtctagggccccccgaaccacggg gacgtgggtttcctttgaaaaacacgatgataatatggccacaacctggaacaagag acttgcgcgactctctcacttttgagggaatgccccaaatgctctgctctacaataccgt aatggattttacctgctaaagtatgatgaagaatggtaaccagaggagttattgactgat ggagaggatgatgtcttgatcccgaattagacatggaagtcgttttcgagttacaggg aagcggagctactaacttcagcctgctgaagcaggctggagatgtggaggagaacc ctggacctATGGACCTGTTTCATGAGAATCTTCACCATCG GCACCGTGACACTGAAGCAGGGCGAGATCAAGGAT GCCACCCCTAGCGACTTCGTGAGAGCCACCGCCACA
--	--	---

		ATTCCTATCCAGGCTAGCCTGCCTTTTGGATGGCTG ATCGTGGGCGTCGCCCTGCTCGCCGTGTTCCAGAGC GCCTCTAAGATCATTACACTGAAGAAAAGATGGCA GCTGGCCCTCTCCAAAGGCGTGCACTTCGTGTGTAA TCTGCTGCTGCTTTTTGTGACAGTGTACAGCCACCT GCTGCTGGTTGCTGCTGGCCTGGAAGCCCCCTTCCT GTACCTGTACGCCCTGGTCTACTTCCTGCAGTCTATC AACTTCGTGCGGATCATCATGCGGCTGTGGCTGTGC TGGAAGTGCAGAAGCAAGAACCCACTGCTGTACGA CGCCAATTACTTCCTGTGTTGGCACACCAACTGCTA CGACTACTGCATCCCCTACAACAGCGTGACCAGCAG CATCGTGATCACCTCTGGCGACGGAACAACCAGCCC TATCAGCGAGCATGATTACCAGATCGGCGGATATAC AGAGAAGTGGGAGAGCGGCGTGAAGGACTGCGTGG TGCTGCACAGCTACTTTACCTCCGATTACTACCAAC TGTATTCTACCCAGCTGAGCACCGACACCGGCGTGG AACACGTGACCTTCTTCATCTACAACAAGATCGTGG ACGAGCCTGAGGAACACGTGCAGATCCACACTATC GACGGCAGCTCTGGCGTTGTGAACCCTGTGATGGAA CCCATCTACGATGAGCCCACCACAACAACCTCCGTG CCCCTGTaaGTGTttacctgttaatgtagcatttgagctttgggctaagcgcaa cattaacagtagcagaggtgaaaataactcaataatttgggtgtggacattgctgcta atactgtgatctgggactacaaaagagatgctccagcacatatctactatttgggtttg ttctatgactgacatagccaagaacaaactgaaacgatttgtgcaccactcactgtctt tttgatggtagagttgatggtcaagtagactatttagaaatgcccgtaatggtgtctta ttacagaaggtagtgtaaaaggtttacaaccatctgtaggtcccaacaagctagctta atggagtacattaattggagaagccgtaaaaacacagttcaattattataagaagttg atggtgtgtccaacaattacctgaaacttacttactcagagtagaaatttacaagaatt aaaccaggagtagcaaatggaattgatttcttagaattagctatggatgaattcattgaa cggataaaattagaaggctatgccttcgaacatatcgtttatggagattttagtcatagtc agtttagtggtGCGAaattgtgtgtt
Furin cleavage site	50	CGAAAACGGCGC
Viral packaging signal (CoVEG8)	34	GTGTCaAAGcGCGAGGAACTGTTCACCGGAGTtTGC CCATCCTGGTCGAGCTGGACGGCGATGTGAACGGC CACAAGTTCAGCGTTTCTGGCGAGGGtgagctttgggctaag cgcaacattaaaccagtagcagaggtgaaaataactcaataatttgggtgtggacattgc tgctaatactgtgatctgggactacaaaagagatgctccagcacatatctactattgg tgtttgtctatgactgacatagccaagaacaaactgaaacgatttgtgcaccactcac tgtctttttgatggtagagttgatggtcaagtagactatttagaaatgcccgtaatggtg ttcttattacagaaggtagtgtaaaaggtttacaaccatctgtaggtcccaacaagcta gtcttaattggagtacattaattggagaagccgtaaaaacacagttcaattattataaga aagttgatggtgtgtccaacaattacctgaaacttacttactcagagtagaaatttaca agaatttaaacccaggagtagcaaatggaattgatttcttagaattagctatggatgaattc attgaacggtataaaattagaaggctatgccttcgaacatatcgtttatggagattttagtc ata
Mutant protein	S 51	MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVY YPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNG

		TKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSK TQSLIVNNATNVVIKVCEFCNDPFLGVYYHKNNK SWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNF KNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEP LVDLPIGINITRFQTLALHRSYLTTPGDSSSGWTAGAA AyyVGYLQPRTEFLKYNENGTITDAVDCALDPLSETK CTLKSFTVEKGiyQTSNFRVQPTESIVRFPNITNLCPF EVFNATRFASVYAWNRKRISNCVADYSVLYNSASFST FKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPG QTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNY NYLYRLFRKSNLKPFERDISTEiyQAGSTPCNGVEGFN CYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATV CGPKKSTNLVKNKCVNFNFENGLTGTGVLTESNKKFLP FQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITP GTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVY STGSNVFQTRAGCLIGAEHVNNSEYCDIPGAGICASY QTQTNSPGSASSVASQSIIAYTMSLGAENSVAYSNNNSI AIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNS LLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQI YKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVT ADAGFIKQYGDCLGDIAARDLICAQKFENGLTVLPPLT DEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMA YRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLST ASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLN DILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAE IRASANLAATKMSECVLGQSKRVDFCGKGYHLSFP QSAPHGVVFLHVTYVPAQEKNFITTAICHGDKAHFP REGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVD LGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCC SCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT
Mutant gene (CoVEG9)	S 52	ATGTTCTGTTCTCTGGTGCTGCTGCCTCTGGTCAGCT CCCAGTGTGTGAACCTGACCACCAGAACCCAGCTGC CACCTGCTTATACAACTCCTTCACTCGGGGGGTAT ACTACCCCGACAAGGTGTTCAAGATCTAGCGTGCTGC ATTCTACACAAGACCTGTTCTCTGCCCTTCTTCAGCA ACGTGACCTGGTTCACGCCATCCACGTGTCTGGAA CCAACGGAACCAAGAGATTCGACAACCCCGTGCTG CCTTTCAACGACGGCGTGTACTTCGCCAGCACCCGAG AAGTCCAACATCATCAGAGGATGGATTTTCGGCACC ACACTGGACAGCAAAACCCAGAGCCTGCTGATCGT GAACAACGCCACCAACGTGGTGATCAAGGTGTGCG AGTTCCAGTTCTGCAATGATCCCTTCTGGGCGTGT ACTACCACAAGAACAACAAGTCTTGATGGAAAGC GAGTTCAGAGTGTATTCCAGCGCCAACAATTGCACC TTCGAGTACGTGAGCCAACCCTTTCTGATGGACCTT GAAGGCAAGCAGGGCAACTTCAAAAATCTGCGAGA ATTTGTGTTCAAGAACATCGACGGATACTTCAAGAT CTACTCTAAGCACACGCCAATCAACCTGGTGAGAG

		ATCTGCCCCAGGGCTTTAGCGCTTTGGAACCTCTGG TGGACCTGCCTATCGGAATCAACATCACCAGATTTC AAACTCTCCTGGCCCTGCACAGATCTTATCTGACCC CTGGGGACAGTAGTAGCGGCTGGACAGCCGGCGCC GCCGCTACTACGTGGGATACCTGCAGCCTAGAACA TTCCTGCTGAAGTACAATGAGAACGGAACAATCAC AGACGCCGTGGACTGCGCCCTGGATCCTTTGAGCGA GACAAAGTGCACCCTGAAGTCGTTACCCGTCGAAA AAGGCATCTACCAGACCAGCAACTTCCGCGTGCAG CCTACGGAATCTATCGTGCGGTTCCCCAACATCACC AACCTGTGCCCTTTTCGGCGAGGTGTTTAAACGCTACA AGGTTCCGCAGCGTGTATGCCTGGAACAGAAAAGAG AATCAGCAATTGCGTGGCCGATTATAGCGTTCTGTA CAACAGCGCTTCCTTCAGCACCTTCAAGTGCTACGG CGTGTCTCCAACCAAGCTGAACGACCTCTGCTTCAC CAATGTCTACGCTGACTCTTTCGTGATTAGAGGCGA TGAGGTTAGACAGATCGCACCTGGCCAGACCGGCA AAATCGCTGACTACAACCTACAAGCTGCCTGATGACT TCACAGGCTGTGTCATTGCCTGGAACCTCAAATAACC TGGACTCTAAAGTGGGCGGCAACTACAACCTACCTGT ACCGGCTGTTCCGGAAGAGCAATCTGAAACCTTTTG AGCGGGACATCTCTACAGAGATCTACCAGGCCGGC AGCACACCCTGCAACGGCGTTGAGGGCTTCAACTGC TACTTCCCTCTGCAGAGCTACGGCTTTCAGCCAACA AATGGAGTGGGCTACCAGCCGTACAGAGTGGTGGT GCTGAGCTTCGAACTGCTGCATGCCCCAGCCACAGT GTGTGGACCTAAGAAGTCTACCAACCTGGTGAAGA ACAAGTGCGTGAACCTTTAACTTTAACGGCCTGACCG GCACAGGCGTGCTGACCGAATCCAACAAAAAGTTC CTGCCCTTCCAACAGTTCGGCAGAGACATCGCCGAT ACAACCGATGCCGTGCGGGACCCCCAGACCTTAGA AATCCTAGATATCACCCCGTGCAGCTTCGGCGGAGT CTCTGTTATTACTCCTGGCACCAACACCAGCAACCA AGTGGCTGTTCTGTACCAAggcGTGAACCTGCACCGAA GTGCCTGTGGCTATCCACGCCGATCAGCTGACCCCA ACCTGGCGGGTGTATAGCACCGGCTCTAACGTGTTT CAGACCCGGGCTGGCTGCCTGATCGGCGCCGAACA CGTCAACAACTCCTATGAATGTGACATCCCCATCGG GGCTGGCATCTGCGCCAGTTACCAGACACAGACAA ATAGCCCTGGCAGCGCCAGCAGCGTGGCCTCCAG AGTATCATTGCCTACACCATGAGCCTGGGCGCCGAG AACAGCGTGGCCTATTCTAACAATAGCATCGCAATC CCTACCAACTTTACCATCTCTGTGACAACCGAGATC CTGCCTGTGAGCATGACCAAAAACCAGCGTGGACTG CACGATGTACATCTGTGGCGACAGCACAGAATGCA GTAATCTGTTGCTGCAGTACGGCAGCTTTTGCACCC AGTTGAATAGAGCCCTGACCGGAATCGCCGTAGAG CAGGACAAAAATACCCAGGAGGTGTTTCGCCAGGT GAAACAGATCTACAAGACACCTCCCATTAAGGACTT CGGAGGTTTTAACTTCAGCCAGATCCTGCCCCAGCC
--	--	--

		TTCCAAGCCTAGCAAACGCTCCTTCATCGAGGACCT GCTCTTCAACAAGGTGACACTGGCTGATGCCGGCTT CATCAAGCAGTACGGAGATTGTCTGGGAGACATCG CCGCTAGAGATCTGATCTGCGCCCAAAAGTTCAACG GCCTGACCGTGCTGCCTCCTCTGCTTACAGACGAGA TGATCGCCCAGTACACCAGCGCCCTGCTGGCTGGCA CCATCACAAGCGGCTGGACCTTCGGAGCCGGAGCC GCTCTGCAAATCCCCTTTGCCATGCAGATGGCCTAC CGGTTCAACGGCATCGGCGTGACACAGAATGTGCT GTACGAGAACCAGAAGCTGATCGCTAACCAGTTTA ACAGCGCTATCGGCAAGATCCAGGACTCGCTGAGT AGCACCGCCTCTGCCCTGGGCAAGCTGCAGGACGTC GTGAACCAGAACGCCCAAGCCCTGAACACACTGGT GAAACAGCTGAGCAGCAACTTCGGCGCCATCAGCT CTGTGCTGAACGATATCCTGAGCAGACTGGACCCTc ccGAAGCCGAGGTCCAGATCGACAGACTGATCACAG GAAGACTGCAGAGCCTGCAAACGTACGTGACACAG CAGCTGATCCGGGCAGCCGAAATCCGGGCCAGCGC CAATCTGGCCGCTACCAAGATGAGCGAGTGCGTGTT AGGCCAGAGCAAGCGGGTGGATTTCTGCGGTAAGG GATACCACCTGATGAGCTTTCCCCAGAGCGCTCCTC ACGGCGTGGTGTCTTCTGCACGTGACCTACGTTCTTG CCCAGGAAAAGAACTTCACCACCGCCCCTGCTATCT GCCACGATGGCAAGGCCCACTTCCCTAGAGAGGGC GTTTTCTGTGTCTAACGGCACACACTGGTTTGTGACC CAGAGAACTTCTACGAGCCTCAGATCATCACCACA GACAACACCTTTGTGAGCGGCAATTGCGACGTGGTG ATCGGAATTGTTAATAATACCGTGTACGACCCTCTG CAGCCTGAGCTCGACAGCTTCAAGGAAGAGCTGGA CAAGTACTTCAAGAACCACACCTCCCCAGATGTGGA CCTGGGCGATATTTCAAGGCATCAACGCCTCCGTCGT GAATATCCAGAAGGAGATCGACCGGCTCAACGAGG TGGCCAAGAACCTTAACGAGAGCCTGATCGACCTG CAGGAACTGGGCAAATATGAGCAGTACATCAAGTG GCCTTGGTACATCTGGCTGGGCTTTATCGCAGGCCCT GATCGCTATCGTGATGGTGACCATTATGCTGTGTTG TATGACCAGCTGTTGTAGTTGTCTGAAGGGCTGCTG TTCTTGCGGCAGCTGCTGCAAGTTCGACGAAGACGA CTCAGAGCCCGTGCTGAAAGGCGTGAAAGCTGCACT ACACC
ORF3 protein	53	MDLFMRIFTIGTVTLKQGEIKDATPSDFVRATATIPQA SLPFGWLIVGVALLAVFQSASKIITLKKRWQLALSKGV HFVCNLLLLFVTVYSHLLLVAAGLEAPFLYLYALVYF LQSINFVRIIMRLWLCWKCRSKNPLLYDANYFLCWHT NCYDYCIPYNSVTSSIVITSGDGTTSPISEHDYQIGGYT EKWESGVKDCVVLHSYFTSDYYQLYSTQLSTDGVE HVTFFIYNKIVDEPEEHVQIHTIDGSSGVVNPVMEPIYD EPTTTTSVPL

ORF3 gene (CoVEG14)	54	ATGGACCTGTTTCATGAGAATCTTCACCATCGGCACC GTGACACTGAAGCAGGGCGAGATCAAGGATGCCAC CCCTAGCGACTTCGTGAGAGCCACCGCCACAATTCC TATCCAGGCTAGCCTGCCTTTTGGATGGCTGATCGT GGGCGTCGCCCTGCTCGCCGTGTTCCAGAGCGCCTC TAAGATCATTACACTGAAGAAAAGATGGCAGCTGG CCCTCTCCAAAGGCGTGCACCTTCGTGTGTAATCTGC TGCTGCTTTTTGTGACAGTGTACAGCCACCTGCTGC TGGTTGCTGCTGGCCTGGAAGCCCCCTTCCTGTACC TGTACGCCCTGGTCTACTTCCTGCAGTCTATCAACTT CGTGCGGATCATCATGCGGCTGTGGCTGTGCTGGAA GTGCAGAAGCAAGAACCCACTGCTGTACGACGCCA ATTACTTCCTGTGTTGGCACACCAACTGCTACGACT ACTGCATCCCCTACAACAGCGTGACCAGCAGCATCG TGATCACCTCTGGCGACGGAACAACCAGCCCTATCA GCGAGCATGATTACCAGATCGGCGGATATACAGAG AAGTGGGAGAGCGGCGTGAAGGACTGCGTGGTGCT GCACAGCTACTTTACCTCCGATTACTACCAACTGTA TTCTACCCAGCTGAGCACCGACACCGGCGTGGAAC ACGTGACCTTCTTCATCTACAACAAGATCGTGGACG AGCCTGAGGAACACGTGCAGATCCACACTATCGAC GGCAGCTCTGGCGTTGTGAACCCTGTGATGGAACCC ATCTACGATGAGCCCACCACAACAACCTCCGTGCC CTGTaa
Control only plasmid expression cassette	S 65	ATGTTTCGTGTTTCCTGGTGCTGCTGCCTCTGGTCAGCT CCCAGTGTGTGAACCTGACCACCAGAACCCAGCTGC CACCTGCTTATACAAACTCCTTCACTCGGGGGGTAT ACTACCCCGACAAGGTGTTTCAGATCTAGCGTGCTGC ATTCTACACAAGACCTGTTTCCTGCCCTTCTTCAGCA ACGTGACCTGGTTCACGCCATCCACGTGTCTGGAA CCAACGGAACCAAGAGATTTCGACAACCCCGTGCTG CCTTTCAACGACGGCGTGTACTTCGCCAGCACCGAG AAGTCCAACATCATCAGAGGATGGATTTTCGGCACC ACACTGGACAGCAAAACCCAGAGCCTGCTGATCGT GAACAACGCCACCAACGTGGTGATCAAGGTGTGCG AGTTCCAGTTCTGCAATGATCCCTTCTGGGCGTGT ACTACCACAAGAACAACAAGTCTTGATGGAAAGC GAGTTCAGAGTGTATTCCAGCGCCAACAATTGCACC TTCGAGTACGTGAGCCAACCCTTTCTGATGGACCTT GAAGGCAAGCAGGGCAACTTCAAAAATCTGCGAGA ATTTGTGTTCAAGAACATCGACGGATACTTCAAGAT CTACTCTAAGCACACGCCAATCAACCTGGTGAGAG ATCTGCCCCAGGGCTTTAGCGCTTTGGAACCTCTGG TGGACCTGCCTATCGGAATCAACATCACCAGATTTC AAACTCTCCTGGCCCTGCACAGATCTTATCTGACCC CTGGGGACAGTAGTAGCGGCTGGACAGCCGGCGCC GCCGCCTACTACGTGGGATACCTGCAGCCTAGAACA TTCCTGCTGAAGTACAATGAGAACGGAACAATCAC AGACGCCGTGGACTGCGCCCTGGATCCTTTGAGCGA GACAAAGTGCACCCTGAAGTCGTTACCGTCGAAA

		AAGGCATCTACCAGACCAGCAACTTCCGCGTGCGAG CCTACGGAATCTATCGTGCGGTTCCCCAACATCACC AACCTGTGCCCTTTTCGGCGAGGTGTTTAACGCTACA AGGTTCGCCAGCGTGTATGCCTGGAACAGAAAGAG AATCAGCAATTGCGTGGCCGATTATAGCGTTCTGTA CAACAGCGCTTCCTTCAGCACCTTCAAGTGCTACGG CGTGTCTCCAACCAAGCTGAACGACCTCTGCTTCAC CAATGTCTACGCTGACTCTTTCGTGATTAGAGGCGA TGAGGTTAGACAGATCGCACCTGGCCAGACCGGCA AAATCGCTGACTACAACACTACAAGCTGCCTGATGACT TCACAGGCTGTGTCATTGCCTGGAACCTCAAATAACC TGGACTCTAAAGTGGGCGGCAACTACAACCTACCTGT ACCGGCTGTTCCGGAAGAGCAATCTGAAACCTTTTG AGCGGGACATCTCTACAGAGATCTACCAGGCCGGC AGCACACCCTGCAACGGCGTTGAGGGCTTCAACTGC TACTTCCCTCTGCAGAGCTACGGCTTTCAGCCAACA AATGGAGTGGGCTACCAGCCGTACAGAGTGGTGGT GCTGAGCTTCGAACTGCTGCATGCCCCAGCCACAGT GTGTGGACCTAAGAAGTCTACCAACCTGGTGAAGA ACAAGTGCGTGAACTTTAACTTTAACGGCCTGACCG GCACAGGCGTGCTGACCGAATCCAACAAAAAGTTC CTGCCCTTCCAACAGTTCGGCAGAGACATCGCCGAT ACAACCGATGCCGTGCGGGACCCCCAGACCTTAGA AATCCTAGATATCACCCCGTGCAGCTTCGGCGGAGT CTCTGTTATTACTCCTGGCACCAACACCAGCAACCA AGTGGCTGTTCTGTACCAAggcGTGAACCTGCACCGAA GTGCCTGTGGCTATCCACGCCGATCAGCTGACCCCA ACCTGGCGGGTGTATAGCACCGGCTCTAACGTGTTT CAGACCCGGGCTGGCTGCCTGATCGGCGCCGAACA CGTCAACAACTCCTATGAATGTGACATCCCCATCGG GGCTGGCATCTGCGCCAGTTACCAGACACAGACAA ATAGCCCTGGCAGCGCCAGCAGCGTGGCCTCCCAG AGTATCATTGCCTACACCATGAGCCTGGGCGCCGAG AACAGCGTGGCCTATTCTAACAATAGCATCGCAATC CCTACCAACTTTACCATCTCTGTGACAACCGAGATC CTGCCTGTGAGCATGACCAAAACCAGCGTGGACTG CACGATGTACATCTGTGGCGACAGCACAGAATGCA GTAATCTGTTGCTGCAGTACGGCAGCTTTTGCACCC AGTTGAATAGAGCCCTGACCGGAATCGCCGTAGAG CAGGACAAAAATACCCAGGAGGTGTTTCGCCAGGT GAAACAGATCTACAAGACACCTCCCATTAAGGACTT CGGAGGTTTTAACTTCAGCCAGATCCTGCCCGACCC TTCCAAGCCTAGCAAACGCTCCTTCATCGAGGACCT GCTCTTCAACAAGGTGACACTGGCTGATGCCGGCTT CATCAAGCAGTACGGAGATTGTCTGGGAGACATCG CCGCTAGAGATCTGATCTGCGCCCAAAAGTTCAACG GCCTGACCGTGCTGCCTCCTCTGCTTACAGACGAGA TGATCGCCAGTACACCAGCGCCCTGCTGGCTGGCA CCATCACAAGCGGCTGGACCTTCGGAGCCGGAGCC GCTCTGCAAATCCCCCTTTGCCATGCAGATGGCCTAC
--	--	---

	CGGTTCAACGGCATCGGCGTGACACAGAATGTGCT GTACGAGAACCAGAAGCTGATCGCTAACCAGTTTA ACAGCGCTATCGGCAAGATCCAGGACTCGCTGAGT AGCACCGCCTCTGCCCTGGGCAAGCTGCAGGACGTC GTGAACCAGAACGCCCCAAGCCCTGAACACACTGGT GAAACAGCTGAGCAGCAACTTCGGCGCCATCAGCT CTGTGCTGAACGATATCCTGAGCAGACTGGACCCTc ccGAAGCCGAGGTCCAGATCGACAGACTGATCACAG GAAGACTGCAGAGCCTGCAAACGTACGTGACACAG CAGCTGATCCGGGCAGCCGAAATCCGGGCCAGCGC CAATCTGGCCGCTACCAAGATGAGCGAGTGCGTGTT AGGCCAGAGCAAGCGGGTGGATTTCTGCGGTAAGG GATACCACCTGATGAGCTTCCCCAGAGCGCTCCTC ACGGCGTGTTTCTGCACGTGACCTACGTTTCTG CCCAGGAAAAGAACTTCACCACCGCCCCTGCTATCT GCCACGATGGCAAGGCCCACTTCCCTAGAGAGGGC GTTTTCGTGTCTAACGGCACACACTGGTTTGTGACC CAGAGAACTTCTACGAGCCTCAGATCATCACCACA GACAACACCTTTGTGAGCGGCAATTGCGACGTGGTG ATCGGAATTGTTAATAATACCGTGTACGACCCTCTG CAGCCTGAGCTCGACAGCTTCAAGGAAGAGCTGGA CAAGTACTTCAAGAACCACCTCCCCAGATGTGGA CCTGGGCGATATTTCAAGCATCAACGCCTCCGTCGT GAATATCCAGAAGGAGATCGACCGGCTCAACGAGG TGGCCAAGAACCTTAACGAGAGCCTGATCGACCTG CAGGAACTGGGCAAATATGAGCAGTACATCAAGTG GCCTTGGTACATCTGGCTGGGCTTTATCGCAGGCCT GATCGCTATCGTGATGGTGACCATTATGCTGTGTTG TATGACCAGCTGTTGTAGTTGTCTGAAGGGCTGCTG TTCTTGCGGCAGCTGCTGCAAGTTCGACGAAGACGA CTCAGAGCCCGTGCTGAAAGGCGTGAAGCTGCACT ACACCTAAccccccccctaacgttactggccgaagccgcttgaataagg ccggtgtgcgttgtctatatgttattttccacatattgccgtctttggcaatgtgaggg cccggaacctggccctgtctcttgacgagcattcctaggggtctttccccctcgcg aaaggaatgaaggtctgttgaatgtcgtgaaggaagcagttcctctggaagcttctg aagacaacaacgtctgtagcgaccctttgcaggcagcgggaacccccacctggcg acaggtgcctctgcggccaaaagccacgtgtataagatacacctgcaaggcgca caacccagtgccacgttgtgagttggatagttgtggaaagagtaaatggctctcctc aagcgtattcaacaaggggctgaaggatgccagaaggtagccattgtatgggatc tgatctggggcctcggtgcacatgctttacatgtgttagtcgaggttaaaaaacgtct agggccccgaaccacggggacgtgggtttcctttgaaaaacacgatgataatatggc cacaacctgaacaagagacttgcgcgactctctcacttttgaggaatgccccaaa tgctctgctctacaataaccgtaattgattttacgtgctaaagtatgatgaagaatggtagc cagaggagtattgactgatggagaggatgatgtctttgatccgaattagacatggaa gtcgttttcgagttacagtaa
--	---

[0086] In some embodiments, the vector disclosed herein comprises the polynucleotide from any one of CoVEG 1-17 that encodes the SARS CoV2 Spike protein. In some embodiments,

the vector disclosed herein comprises the polynucleotide from any one of CoVEG 1-17 that encodes the SARS CoV2 membrane protein. In some embodiments, the vector disclosed herein comprises the polynucleotide from any one of CoVEG 1-17 that encodes the SARS CoV2 envelope protein. In some embodiments, the vector disclosed herein comprises the polynucleotide from any one of CoVEG 1-17 that encodes the SARS CoV2 nucleocapsid protein. In some embodiments, the vector disclosed herein comprises the polynucleotide from any one of CoVEG 1-17 that encodes the EMCV L protein. In some embodiments, the vector disclosed herein comprises the polynucleotide from any one of CoVEG 1-17 that encodes the internal ribosome entry site (IRES). In some embodiments, the vector disclosed herein comprises the polynucleotide from any one of CoVEG 1-17 that encodes the viral packaging signal.

ii) Polynucleotides

[0087] Polynucleotides of the present disclosure may include DNA, RNA, and DNA-RNA hybrid molecules. In some embodiments, polynucleotides are isolated from a natural source; prepared *in vitro*, using techniques *e.g.*, PCR amplification or chemical synthesis; prepared *in vivo*, *e.g.*, via recombinant DNA technology; or prepared or obtained by any appropriate method. In some embodiments, polynucleotides are of any shape (linear, circular, etc.) or topology (single-stranded, double-stranded, linear, circular, supercoiled, torsional, nicked, etc.). Polynucleotides may also comprise nucleic acid derivatives *e.g.*, peptide nucleic acids (PNAs) and polypeptide-nucleic acid conjugates; nucleic acids having at least one chemically modified sugar residue, backbone, internucleotide linkage, base, nucleotide, nucleoside, or nucleotide analog or derivative; as well as nucleic acids having chemically modified 5' or 3' ends; and nucleic acids having two or more of such modifications. Not all linkages in a polynucleotide need to be identical.

[0088] A polynucleotide is said to “encode” a protein when it comprises a nucleic acid sequence that is capable of being transcribed and translated (*e.g.*, DNA→RNA→protein) or translated (RNA→protein) in order to produce an amino acid sequence corresponding to the amino acid sequence of said protein. *In vivo* (*e.g.*, within a eukaryotic cell) transcription and/or translation is performed by endogenous or exogenous enzymes. In some embodiments, transcription of the polynucleotides of the disclosure is performed by the endogenous polymerase II (polII) of the eukaryotic cell. In some embodiments, an exogenous RNA polymerase is provided on the same or a different vector. In some embodiments, viral polymerases may alternatively or additionally be used. In some embodiments, a viral promoter

is used in combination with one or more viral polymerase. In some embodiments, the RNA polymerase is selected from a T3 RNA polymerase, a T5 RNA polymerase, a T7 RNA polymerase, an H8 RNA polymerase, EMCV RNA polymerase, HIV RNA polymerase, Influenza RNA polymerase, SP6 RNA polymerase, CMV RNA polymerase, T3 RNA polymerase, T1 RNA polymerase, SP01 RNA polymerase, SP2 RNA polymerase, Phi15 RNA polymerase, and the like. Viral polymerases are RNA priming or capping polymerases. In some embodiments, IRES elements are used in conjunction with viral polymerases.

[0089] The polynucleotides disclosed herein may encode one or more antigens; and/or one or more enhancer proteins. In some embodiments, the polynucleotide encodes one antigen. In some embodiments, the polynucleotide encodes one enhancer protein. In some embodiments, the polynucleotide encodes more than one antigen; more than one enhancer protein, and/or one or more separating elements.

[0090] In some embodiments, the polynucleotide may encode a polypeptide that is not antigenic. In some embodiments, the polypeptide that is not antigenic may form a part of a VLP. Thus, the present disclosure provides vectors that comprise polynucleotides that encode one or more antigens, and/or polynucleotides that encode one or more non-antigenic polypeptides, and/or polynucleotides that encode one or more enhancer proteins. In some embodiments, the one or more antigens and the one or more non-antigenic polypeptides are capable of forming a virus like particle (VLP). In some embodiments, the one or more antigens may be derived from one or more proteins of a first virus, and the one or more non-antigenic polypeptides may be derived from one or more proteins of a second virus.

iii) Separating elements

[0091] In some embodiments, antigen(s) and enhancer protein(s) according to the present disclosure are encoded on the same vector. In some embodiments, antigen(s) and enhancer protein(s) according to the present disclosure are encoded on separate vectors. In some embodiments, if nucleic acid sequences encoding one or more antigens and one or more enhancer proteins are present in the same vector, the vector may comprise a separating element for separate expression of the proteins. In some embodiments, the vector is a bicistronic vector or a polycistronic vector. The separating element may be an internal ribosomal entry site (IRES) or 2A element. In some embodiments, a vector may comprise a nucleic acid encoding a 2A element, or a nucleic acid encoding an IRES.

[0092] In some embodiments, the first polynucleotide or the second polynucleotide, or both, are operatively linked to a polynucleotide encoding a 2A element. In some embodiments, the

polynucleotide encoding the enhancer protein and/or the polynucleotide encoding the antigen are operatively linked to a polynucleotide encoding an a 2A element. Non-limiting examples of 2A elements include P2A, E2A, F2A, and T2A. In some embodiments, the amino acid sequence of the 2A peptide has at least 80% sequence identity (for instance, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between) to SEQ ID NO: 17. In some embodiments, the amino acid sequence of the 2A peptide is SEQ ID NO: 17.

[0093] In some embodiments, the nucleic acid sequence encoding the 2A peptide has at least 80% sequence identity (for instance, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between) to SEQ ID NO: 18 or 69. In some embodiments, the nucleic acid sequence encoding the 2A peptide is SEQ ID NO: 18 or 69.

[0094] In some embodiments, the first polynucleotide or the second polynucleotide, or both, are operatively linked to a polynucleotide encoding an internal ribosome entry site (IRES). In some embodiments, the polynucleotide encoding the enhancer protein and/or the polynucleotide encoding the antigen are operatively linked to a polynucleotide encoding an IRES. In some embodiments, the polynucleotide encoding the IRES has a nucleic acid sequence with at least 80% sequence identity (for instance, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between) to the nucleic acid sequence of SEQ ID NO: 24 or 67. In some embodiments, the polynucleotide encoding the IRES has a nucleic acid sequence of SEQ ID NO: 24 or 67.

[0095] In some embodiments, the antigen, and the enhancer protein are comprised in a single fusion protein. In some embodiments, the fusion protein may comprise a linking element. In some embodiments, the linking element may comprise a cleavage site (*e.g.* a furin, a cathepsin or an intein cleavage site) for enzymatic cleavage in *cis* or in *trans*. In other embodiments, the fusion protein or the linking element does not comprise a cleavage site and the expressed fusion protein comprises both the target protein and the enhancer protein. In some embodiments, the linking element is a 2A element.

iv) Promoters

[0096] Vectors according to the present disclosure may comprise one or more promoters. The term “promoter” refers to a region or sequence located upstream or downstream from the start of transcription which is involved in recognition and binding of RNA polymerase and other

proteins to initiate transcription. The polynucleotide(s) or vector(s) according to the present disclosure may comprise one or more promoters. The promoters may be any promoter known in the art. The promoter may be a forward promoter or a reverse promoter. In some embodiments, the promoter is a mammalian promoter. In some embodiments, one or more promoters are native promoters. In some embodiments, one or more promoters are non-native promoters. In some embodiments, one or more promoters are non-mammalian promoters. Non-limiting examples of RNA promoters for use in the disclosed compositions and methods include U1, human elongation factor-1 alpha (EF-1 alpha), cytomegalovirus (CMV), human ubiquitin, spleen focus-forming virus (SFFV), U6, H1, tRNA^{Lys}, tRNA^{Ser} and tRNA^{Arg}, CAG, PGK, TRE, UAS, UbC, SV40, T7, Sp6, lac, araBad, trp, and Ptac promoters.

[0097] The term “operatively linked” as used herein refers to elements or structures in a nucleic acid sequence that are linked by operative ability and not physical location. The elements or structures are capable of, or characterized by, accomplishing a desired operation. It is recognized by one of ordinary skill in the art that it is not necessary for elements or structures in a nucleic acid sequence to be in a tandem or adjacent order to be operatively linked.

[0098] In some embodiments, a promoter comprised by a vector according to the present disclosure is an inducible promoter.

[0099] In some embodiments, vectors according to the present disclosure may further comprise a polynucleotide sequence encoding a polymerase. In some embodiments, the polymerase is a viral polymerase. In some embodiments, the vectors disclosed herein comprises a polynucleotide sequence encoding a T7 RNA polymerase. In some embodiments, for example, a vector may comprise a T7 promoter configured for transcription of either or both of the polynucleotide encoding an antigen, and the second polynucleotide encoding the enhancer protein by a T7 RNA polymerase.

25 **Antigens**

[0100] In some embodiments, the expression or quality of the antigen is significantly improved by expression according to the disclosed methods, *e.g.*, in conjunction with one or more enhancer proteins. In some embodiments, the antigen is derived from a single protein. In some embodiments, the antigen is derived from multiple proteins. In some embodiments, the antigen is a chimeric antigen comprising amino acid sequences from one or more proteins.

In some embodiments, the antigen is a viral antigen. The viral antigen may comprise the whole or part of an amino acid sequence derived from any viral protein, without limitation. In some

embodiments, the viral antigen is the viral protein. In some embodiments, the amino acid sequence of the viral protein is the whole or part of a structural protein or multiple structural proteins of a virus. In some embodiments, the antigen or antigens assemble into VLPs and are released from the expressing cells.

5 [0101] In some embodiments, the viral antigen comprises the whole or an antigen fragment of any coronavirus protein, without limitation. In some embodiments, the coronavirus is a betacoronavirus. In some embodiments, the betacoronavirus is severe acute respiratory syndrome (SARS) virus. In some embodiments, the betacoronavirus is Middle East respiratory syndrome (MERS) virus, OC43, or HKU1. In some embodiments, the SARS virus is SARS-CoV-1. In some embodiments, the SARS virus is SARS-CoV-2.

[0102] In some embodiments, the viral antigen comprises the whole or an antigen fragment of any one or more of the following proteins: coronavirus spike protein, coronavirus M protein, coronavirus N protein, and coronavirus E protein.

15 [0103] In some embodiments, the coronavirus spike protein is selected from the group consisting of a SARS-Cov-2 spike protein, a Middle East respiratory syndrome (MERS) spike protein, and SARS-CoV spike protein. In some embodiments, the coronavirus M protein is selected from SARS-Cov-2 M protein, MERS M protein and SARS-CoV M protein. In some embodiments, the coronavirus N protein is selected from SARS-Cov-2 N protein, MERS N protein, and SARS-CoV N protein. In some embodiments, the coronavirus E protein is selected from SARS-Cov-2 E protein, MERS E protein, and SARS-CoV E protein.

25 [0104] In some embodiments, the polynucleotide encoding the coronavirus spike protein has a nucleic acid sequence with at least 70% sequence identity - for instance, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between - to the nucleic acid sequence of SEQ ID NO: 14 or 70. In some embodiments, the polynucleotide encoding the coronavirus spike protein has a nucleic acid sequence of SEQ ID NO: 14 or 70.

30 [0105] In some embodiments, the amino acid sequence of the coronavirus spike protein has at least 70% sequence identity - for instance, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between - to the amino acid sequence of SEQ ID NO: 13. In some embodiments, the amino acid sequence of the coronavirus spike protein is SEQ ID NO: 13.

[0106] In some embodiments, the SARS-Cov-2 spike protein is a mutant S protein (also denoted as “S (Mut)”) that comprises one or more amino acid mutations, as compared to SEQ ID NO: 13. In some embodiments, the mutant S protein is expressed at a higher level, as compared to the wild type S protein. In some embodiments, the mutant S protein is prefusion conformation-stabilized spike protein. In some embodiments, the mutation in the S protein stabilizes the trimeric state of the S protein. In some embodiments, the mutant S protein comprises one or more mutations in the internal endogenous proteolytic cleavage site of the S protein. In some embodiments, the mutant S protein comprises a deletion of the internal endogenous proteolytic cleavage site of the S protein. In some embodiments, the one or more mutations in the proteolytic cleavage site of the S protein inhibit the cleavage of the S protein during the assembly process. In some embodiments, a VLP comprising any one or more of the mutant S proteins disclosed herein is more immunogenic than a VLP comprising a wild type S protein, *e.g.*, an S protein comprising an amino acid sequence of SEQ ID NO: 13.

[0107] In some embodiments, the mutant S protein comprises a modification (*e.g.* a substitution) of at least one amino acid residue selected from the group consisting of R682, R683, A684, R685, K986, and V987 in SEQ ID NO: 13. In some embodiments, the mutant S protein comprises at least one amino acid substitution selected from the group consisting of R682G, R683S, R685S, K986P, and V987P in SEQ ID NO: 13. In some embodiments, the mutant S protein comprises the amino acid substitutions, R682G, R683S, R685S, K986P, and V987P in SEQ ID NO: 13. In some embodiments, the mutant S protein comprises the following amino acid substitutions in an internal endogenous furin cleavage site: R682G, R683S, R685S. That is, in some embodiments, the mutant S protein comprises the following amino acids at an internal endogenous furin cleavage site: G at amino acid residue 682, S at amino acid residue 683, A at amino acid residue 684, and S at amino acid residue 685.

[0108] In some embodiments, the mutant S protein has at least 70% sequence identity - for instance, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between – to the amino acid sequence of SEQ ID NO: 51. In some embodiments, the amino acid sequence of the mutant S protein is SEQ ID NO: 51.

[0109] In some embodiments, the polynucleotide encoding the mutant S protein has a nucleic acid sequence with at least 70% sequence identity - for instance, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between – to the nucleic acid

sequence of SEQ ID NO: 52. In some embodiments, the polynucleotide encoding the coronavirus spike protein has a nucleic acid sequence of SEQ ID NO: 52.

[0110] In some embodiments, the polynucleotide encoding the coronavirus M protein has a nucleic acid sequence with at least 80% sequence identity - for instance, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between – to the nucleic acid sequence of SEQ ID NO: 19 or 66. In some embodiments, the polynucleotide encoding the coronavirus M protein has a nucleic acid sequence of SEQ ID NO: 19 or 66.

[0111] In some embodiments, the amino acid sequence of the coronavirus M protein has at least 80% sequence identity - for instance, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between – to the amino acid sequence of SEQ ID NO: 33. In some embodiments, the amino acid sequence of the coronavirus M protein is SEQ ID NO: 33.

[0112] In some embodiments, the polynucleotide encoding the coronavirus N protein has a nucleic acid sequence with at least 80% sequence identity - for instance, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between – to the nucleic acid sequence of SEQ ID NO: 21 or 71. In some embodiments, the polynucleotide encoding the coronavirus N protein has a nucleic acid sequence of SEQ ID NO: 21 or 71.

[0113] In some embodiments, the amino acid sequence of the coronavirus N protein has at least 80% sequence identity - for instance, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between – to the amino acid sequence of SEQ ID NO: 20. In some embodiments, the amino acid sequence of the coronavirus N protein is SEQ ID NO: 20.

[0114] In some embodiments, the polynucleotide encoding the coronavirus E protein has a nucleic acid sequence with at least 80% sequence identity - for instance, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between – to the nucleic acid sequence of SEQ ID NO: 23 or 72. In some embodiments, the polynucleotide encoding the coronavirus E protein has a nucleic acid sequence of SEQ ID NO: 23 or 72.

[0115] In some embodiments, the amino acid sequence of the coronavirus E protein has at least 80% sequence identity - for instance, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that

lie there between – to the amino acid sequence of SEQ ID NO: 22. In some embodiments, the amino acid sequence of the coronavirus E protein is SEQ ID NO: 22.

[0116] In some embodiments, the viral protein is derived from the any one of Groups I, II, III, IV, V, VI, or VII of viruses according to the Baltimore classification. In some embodiments, the viral protein is derived from an enveloped negative-sense, single stranded, segmented RNA virus (*e.g.* Influenza virus). In some embodiments, the viral protein is derived from an enveloped DNA virus (*e.g.* Hepatitis B virus). In some embodiments, the viral protein is derived from a non-enveloped DNA virus (*e.g.* Human Papillomavirus). In some embodiments, the viral protein is derived from a positive strand enveloped RNA virus (*e.g.* a coronavirus, *e.g.*, SARS CoV2, and flaviviruses, *e.g.*, West Nile virus). In some embodiments, the viral antigen comprises the whole or an antigen fragment of any protein derived from a virus selected from the group consisting of SARS-CoV-1, MERS-CoV, chikungunya virus, African Swine Fever virus, Dengue virus, Zika virus, Influenza virus (*e.g.*, A, B, C), Human Immunodeficiency Virus (HIV), Ebola virus, Hepatitis virus (*e.g.*, Hepatitis A, Hepatitis B, Hepatitis C, Hepatitis D, and Hepatitis E), herpes simplex virus type 1 (HSV-1), herpes simplex virus type 2 (HSV-2), West Nile virus, and Human Papillomavirus.

[0117] In some embodiments, the viral antigen comprises the whole or an antigen fragment of any protein derived from West Nile virus. In some embodiments, the West Nile viral protein is the precursor membrane (prM), the envelope glycoprotein (E), or a combination thereof. In some embodiments, the vector encoding one or more West Nile virus proteins, *e.g.*, prM and/or E protein is West Nile Virus Minimal plasmid (WNV minimal plasmid), as depicted in FIG. 14A or West Nile Virus Standard plasmid (WNV standard plasmid), as depicted in FIG. 14B. In some embodiments, the vector encoding one or more West Nile virus proteins, *e.g.* prM and/or E protein comprises a nucleic acid sequence having at least 70% (for example, about 75%, about 80%, about 85%, about 90%, about 95%, about 96%, about 97%, about 98%, about 99%, about 99.5%, or about 100%) identity to SEQ ID NO: 55. In some embodiments, the vector encoding one or more West Nile virus proteins, *e.g.* prM and/or E protein comprises an expression cassette with a nucleic acid sequence having at least 70% (for example, about 75%, about 80%, about 85%, about 90%, about 95%, about 96%, about 97%, about 98%, about 99%, about 99.5%, or about 100%) identity to SEQ ID NO: 64.

[0118] In some embodiments, the polynucleotide encoding the West Nile virus E protein has a nucleic acid sequence with at least 80% sequence identity - for instance, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between – to the nucleic acid sequence of

SEQ ID NO: 60. In some embodiments, the polynucleotide encoding the West Nile virus E protein has a nucleic acid sequence of SEQ ID NO: 60.

[0119] In some embodiments, the polynucleotide encoding the West Nile virus prM protein has a nucleic acid sequence with at least 80% sequence identity - for instance, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, including any values and subranges that lie there between – to the nucleic acid sequence of SEQ ID NO: 59. In some embodiments, the polynucleotide encoding the West Nile virus prM protein has a nucleic acid sequence of SEQ ID NO: 59.

[0120] The nucleic acid sequence of the vectors encoding West Nile viral antigens disclosed herein, and the genetic elements therein are listed in Table 3.

Table 3

Name of sequence	SEQ ID NO:	Sequence
WNV Minimal plasmid sequence (FIG. 14A)	55	GACTCTTCGCGATGTACGGGCCAGATATACGCGTTGA CATTGATTATTGACTAGTTATTAATAGTAATCAATTA CGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCG CGTTACATAACTTACGGTAAATGGCCCCGCTGGCTGA CCGCCAACGACCCCCGCCATTGACGTCAATAATGA CGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCA TTGACGTCAATGGGTGGACTATTTACGGTAACTGCC CACTTGGCAGTACATCAAGTGTATCATATGCCAAGTA CGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCG CTGGCATTATGCCCAGTACATGACCTTATGGGACTTT CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTA TTACCATGGTGATGCGGTTTTGGCAGTACATCAATGG GCGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGT CTCCACCCCATTGACGTCAATGGGAGTTTGTGTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAACA ACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCTggtttagtaaccg tcagatccgctagcgctaccggactcagatctcgagctcaagcttcgaattctgcagtcg acggtagcgcgggcccgatccaccgggtcgccacATGGGCGGCAAGA CAGGCATCGCCGTGATGATCGGCCTGATCGCCTCCGT GGGCGCCGTGACCTGAGCAACTTCCAGGGCAAGGT GATGATGACAGTGAACGCCACAGATGTTACCGATGTT ATCACAATCCCTACCGCCGCTGGAAAGAACTTGTGCA TCGTACGGGCCATGGATGTGGGGTACATGTGCGACG ACACCATCACCTACGAGTGCCCTGTGCTGAGCGCCGG CAATGACCCCGAGGACATCGACTGCTGGTGCACCAA

		GTCTGCCGTTTACGTGAGATATGGCAGGTGTACCAA ACCAGACACAGCAGGAGATCTCGGAGAAGCCTGACC GTGCAAACACACGGCGAGTCCACCCTGGCCAACAAG AAGGGCGCATGGATGGACAGCACCAAGGCCACTCGG TACCTGGTGAAGACCGAGAGCTGGATCCTGAGAAAC CCTGGATACGCCCTGGTGGCCGCCGTGATTGGCTGGA TGCTGGGCTCTAACACCATGCAGAGAGTGGTGTTCGT GGTGCTACTGCTCCTAGTGGCTCCTGCTTACAGCTTC AACTGCCTGGGCATGTCTAACCGGGACTTCCTGGAAG GCGTGTCCGGCGCTACATGGGTGGACCTGGTGCTCGA GGGAGATAGCTGCGTGACCATCATGTCAAAGGACAA GCCCACCATCGACGTGAAAATGATGAACATGGAAGC TGCTAATCTGGCCGAGGTCAGATCTTACTGCTACCTG GCCACAGTGAGTGATCTGAGCACAAAGGCCGCCTGC CCCACCATGGGCGAGGCCCACAACGATAAGCGGGCC GATCCTGCCTTCGTGTGTAGACAGGGCGTGGTGGACC GGGGATGGGGCAACGGCTGCGGGCTGTTCCGGCAAGG GCAGCATCGATACCTGTGCCAAATTCGCCTGTAGCAC CAAGGCCATCGGCCGGACCATTCTGAAAGAAAACAT CAAGTACGAGGTGGCTATCTTCGTGCATGGCCCTACC ACCGTCGAGAGCCACGGCAACTACTCCACACAGGTG GGCGCCACACAGGCCGGCCGATTTTCTATCACACCTG CCGCCCCCAGCTATACACTGAAACTGGGCGAGTACG GCGAAGTGACAGTGGATTGCGAGCCTAGAAGCGGCA TCGACACTAACGCCTACTACGTGATGACCGTGGGCAC AAAAACCTTCCTGGTTCACAGAGAGTGGTTCATGGAC CTGAACCTGCCTTGGTCCAGCGCCGGCAGCACCGTGT GGCGCAATAGAGAGACACTGATGGAATTCGAGGAAC CTCACGCCACCAAGCAGAGCGTGATCGCCCTCGGTAG CCAGGAGGGCGCCCTGCACCAGGCCCTGGCTGGCGC CATCCCCGTGGAATTCTCTAGCAACACCGTGAACTG ACCAGCGGCCACCTGAAGTGCAGAGTGAAGATGGAA AAGCTCCAAGTGAAGGGAACAACCTTACGGCGTCTGC AGCAAGGCTTTCAAGTTCCTGGGCACCCCTGCCGACA CCGGACACGGAACCGTCGTGCTGGAAGTGCAGTACA CCGGCACAGACGGCCCATGTAAAGTGCCTATCAGCA GCGTGGCCTCCCTGAACGACCTGACACCAAGTGGGCA GACTGGTGACCGTTAATCCTTTCGTCAGCGTGGCTAC TGCCAATGCCAAGGTGCTGATCGAGCTGGAACCCCCC TTCGGCGACTCTTATATCGTGGTGGGAAGAGGAGAAC AACAGATCAACCACCACTGGCACAAGAGCGGTTCGT CTATCGGAAAGGCTTTTACCACCACACTGAAGGGCGC TCAGCGGCTGGCCGCCCTGGGCGACACAGCCTGGGA CTTCGGCAGCGTGGGCGGAGTATTTACGTCCGTCGGC AAGGCCGTCCATCAGGTGTTTGGAGGAGCCTTTCGGA GCCTGTTCGGAGGCATGAGCTGGATCACCCAGGGCCT
--	--	---

		GCTGGGCGCGCTGCTGCTGTGGATGGGAATTAACGCC AGAGATAGAAGCATCGCCCTGACATTCTGGCCGTGG GCGGCGTGCTGCTGTTTCTGTCTGTGAACGTGCACGC Gtaacccccccccctaactgtactggccgaagccgcttggaataaggccggtgctgctt tgtctatatgttattttccaccatattgccgtcttttggcaatgtgagggcccggaacctgg ccctgtctcttgacgagcattcctaggggtcttccctctcgccaaaggaatgcaaggt ctgttgatgtcgtgaaggaagcagttcctctggaagcttctgaagacaaacaacgtctg tagcgacccttgcaggcagcgggaacccccacctggcgacaggtgcctctcgggcca aaagccacgtgtataagatacacctgcaaaggcggcacaaacccagtgccacgttgtg agttggatagttgtggaagagtcaaatggctctcctcaagcgattcaacaaggggctg aaggatgcccagaaggtacccattgtatgggatctgatctggggcctcggtgcacatg ctttacatgtgttagtcgaggttaaaaaacgtctaggccccgaaccacggggacgt ggtttcttggaaaaacacgatgataatatggccacaacatggaacaagagactgcg cgactctctcacttttaggaatgccccaaatgctctgctctacaataaccgtaatggattt acctgctaaagtatgatgaagaatggtaccagaggagtattgactgatggagaggatg atgtctttgatcccgaattagacatggaagtcgttttcgagttacagtaaGCGAaattgtt gttgtaactgtttattgcagcttataatggttacaataaagcaatagcatcacaaattca caaataaagcatttttctactgcattctagtgtgtgttgcctcaaacatcaatgtattctaa ggcgtCTTCTACTGGGCGGTTTTATGGACAGCAAGCGAA CCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTG GGAAGCCCTGCAAAGTAACTGGATGGCTTTCTCGCC GCCAAGGATCTGATGGCGCAGGGGATCAAGCTCTGA TCAAGAGACAGGATGAGGATCGTTTCGCATGATTGA ACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGG GTGGAGAGGCTATTCGGCTATGACTGGGCACAACAG ACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGT CAGCGCAGGGGCGCCCGGTTCTTTTGTCAAGACCGA CCTGTCCGGTGCCCTGAATGAACTGCAAGACGAGGC AGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCT TGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAA GGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGG ATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGT ATCCATCATGGCTGATGCAATGCGGCGGCTGCATACG CTTGATCCGGCTACCTGCCCATTGACCAACCAAGCGA AACATCGCATCGAGCGAGCACGTA CTGGATGGAAG CCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCA TCAGGGGCTCGCGCCAGCCGAAGTTCGCCAGGCTC AAGGCGAGCATGCCCGACGGCGAGGATCTCGTCGTG ACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGG AAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCG GCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTG GCTACCCGTGATATTGCTGAAGAGCTTGCGGGCGAAT GGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGC TCCCGATTTCGAGCGCATCGCCTTCTATCGCCTTCTTG ACGAGTTCTTCTGAATTATTAACGCTTACAATTTCCTG ATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTC
--	--	---

		ACACCGCATACAGGTGGCACTTTTCGGGGAAATGTGC GCGGAACCCCTATTTGTTTATTTTCTAAATACATTCA AATATGTATCCGCTCATGAGACAATAACCCTGATAAA TGCTTCAATAATAGCACGTGCTAAAACTTCATTTTAA ATTTAAAAGGATCTAGGTGAAGATCCTTTTTTGATAAT CTCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCC ACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGC TTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTT GTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAA GGTAACTGGCTTCAGCAGAGCGCAGATACCAAATAC TGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTC AAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGC TAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAA GTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTA CCGGATAAGGCGCAGCGGTGCGGGCTGAACGGGGGGT TCGTGCACACAGCCCAGCTTGGAGCGAACGACCTAC ACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTA TGGAAAAACGCCAGCAACGCGGCCTTTTACGGTTCC TGGGCTTTTGCTGGCCTTTTGCTCACATGTTCTT
Origin	56	TTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGC AAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTT GCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTA ACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCC TTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAA CTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATC CTGTTACCAGTGGCTGCTGCCAGTGGCGATAA
CMV promoter and enhancer sequence	57	GACATTGATTATTGACTAGTTATTAATAGTAATCAAT TACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC CGCGTTACATAACTTACGGTAAATGGCCCCGCTGGCT GACCGCCCAACGACCCCCGCCATTGACGTCAATAAT GACGTATGTTCCCATAGTAACGCCAATAGGGACTTTC CATTGACGTCAATGGGTGGACTATTTACGGTAAACTG CCCACTTGGCAGTACATCAAGTGTATCATATGCCAAG TACGCCCCCTATTGACGTCAATGACGGTAAATGGCCC GCCTGGCATTATGCCCAGTACATGACCTTATGGGACT TTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCT ATTACCATGGTGATGCGGTTTTGGCAGTACATCAATG GGCGTGGATAGCGGTTTGACTCACGGGGATTTCGAAG TCTCCACCCCATGACGTCAATGGGAGTTTGTGTTGG CACCAAATCAACGGGACTTTCCAAAATGTCGTAACA

		ACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCT
Signal sequence	58	GGCGGCAAGACAGGCATCGCCGTGATGATCGGCCTG ATCGCCTCCGTGGGCGCC
prM gene	59	GTGACCCTGAGCAACTTCCAGGGCAAGGTGATGATG ACAGTGAACGCCACAGATGTTACCGATGTTATCACA TCCCTACCGCCGCTGGAAAGAACTTGTGCATCGTACG GGCCATGGATGTGGGGTACATGTGCGACGACACCAT CACCTACGAGTGCCCTGTGCTGAGCGCCGGCAATGAC CCCGAGGACATCGACTGCTGGTGCACCAAGTCTGCCG TTTACGTGAGATATGGCAGGTGTACCAAAACCAGAC ACAGCAGGAGATCTCGGAGAAGCCTGACCGTGCAAA CACACGGCGAGTCCACCCTGGCCAACAAGAAGGGCG CATGGATGGACAGCACCAAGGCCACTCGGTACCTGG TGAAGACCGAGAGCTGGATCCTGAGAAACCCTGGAT ACGCCCTGGTGGCCGCCGTGATTGGCTGGATGCTGGG CTCTAACACCATGCAGAGAGTGGTGTTCGTGGTGCTA CTGCTCCTAGTGGCTCCTGCTTACAGC
Envelope gene	60	TTCAACTGCCTGGGCATGTCTAACCGGGACTTCCTGG AAGGCGTGTCCGGCGCTACATGGGTGGACCTGGTGCT CGAGGGAGATAGCTGCGTGACCATCATGTCAAAGGA CAAGCCCACCATCGACGTGAAAATGATGAACATGGA AGCTGCTAATCTGGCCGAGGTCAGATCTTACTGCTAC CTGGCCACAGTGAGTGATCTGAGCACAAAGGCCGCC TGCCCCACCATGGGCGAGGCCCAACGATAAGCGG GCCGATCCTGCCTTCGTGTGTAGACAGGGCGTGGTGG ACCGGGGATGGGGCAACGGCTGCGGGCTGTTCCGGCA AGGGCAGCATCGATACCTGTGCCAAATTCGCCTGTAG CACCAAGGCCATCGGCCGGACCATTCTGAAAGAAAA CATCAAGTACGAGGTGGCTATCTTCGTGCATGGCCCT ACCACCGTCGAGAGCCACGGCAACTACTCCACACAG GTGGGCGCCACACAGGCCGGCCGATTTTCTATCACAC CTGCCGCCCCCAGCTATACTGAAACTGGGCGAGTA CGGCGAAGTGACAGTGGATTGCGAGCCTAGAAGCGG CATCGACACTAACGCCTACTACGTGATGACCGTGGGC ACAAAAACCTTCCTGGTTCACAGAGAGTGGTTCATGG ACCTGAACCTGCCTTGGTCCAGCGCCGGCAGCACCGT GTGGCGCAATAGAGAGACACTGATGGAATTCGAGGA ACCTCACGCCACCAAGCAGAGCGTGATCGCCCTCGGT AGCCAGGAGGGCGCCCTGCACCAGGCCCTGGCTGGC GCCATCCCCGTGGAATTCTCTAGCAACACCGTGAAAC TGACCAGCGGCCACCTGAAGTGCAGAGTGAAGATGG AAAAGCTCCAAGTGAAGGGAACAACCTACGGCGTCT GCAGCAAGGCTTTCAAGTTCCTGGGCACCCCTGCCGA CACCGGACACGGAACCGTCGTGCTGGAACCTGCAGTA CACCGGCACAGACGGCCCATGTAAAGTGCCTATCAG CAGCGTGGCCTCCCTGAACGACCTGACACCAGTGGGC AGACTGGTGACCGTTAATCCTTTCGTCAGCGTGGCTA CTGCCAATGCCAAGGTGCTGATCGAGCTGGAACCCCC

		CTTCGGCGACTCTTATATCGTGGTGGGAAGAGGAGAA CAACAGATCAACCACCACTGGCACAAGAGCGGTTTCG TCTATCGGAAAGGCTTTTACCACCACACTGAAGGGCG CTCAGCGGCTGGCCGCCCTGGGCGACACAGCCTGGG ACTTCGGCAGCGTGGGCGGAGTATTTACGTCCGTCGG CAAGGCCGTCCATCAGGTGTTTGGAGGAGCCTTTCGG AGCCTGTTTCGGAGGCATGAGCTGGATCACCCAGGGC CTGCTGGGCGCGCTGCTGCTGTGGATGGGAATTAACG CCAGAGATAGAAGCATCGCCCTGACATTCCTGGCCGT GGGCGGCGTGCTGCTGTTTCTGTCTGTGAACGTGCAC GCGtaa
IRES encoding sequence	61	ccccccccctaacgttactggccgaagccgcttggaataaggccggtgtgcgtttgtct atatgttatttccaccatattgccgtcttttgcaatgtgagggcccgaaacctggccct gtcttctgacgagcattcctagggtctttccctctcgcaaaggaatcaaggtctgtt gaatgtcgtgaagggaagcagttcctctggaagcttctgaagacaaacaacgtctgtacg gacctttgcaggcagcggaacccccacctggcgacaggtgctctcgggccaaaa gccacgtgtataagatacacctgcaaaggcggcacaaccccgatgccacgtgtgagtt ggatagttgtgaaagagtcaaatggctctcctcaagcgtattcaacaaggggtgaag gatgccagaaggtacccattgtatgggatctgatctggggcctcggtgcacatgcttta catgtgttagtcgaggttaaaaaaacgtctaggccccccgaaccacggggacgtggtt tcctttgaaaaacacgatgataa
EMCV L protein encoding sequence	62	atggccacaacatggaacaagagactgcgcgcactctctcatttgaggaatgccc aaatgctctgctctacaataaccgtaattggattttacctgctaaagtatgaagaatggt cccagaggagtattgactgatggagaggatgatgtctttgatccgaattagacatggaa gtcgttttcgagttacagtaa
SV40 poly A encoding sequence	63	aactgtttattgcagcttataatggttacaataaagcaatagcatcacaatttcacaaat aaagcattttttcactgcattctagtgtgtgtttgtccaaactcatcaatgtatctta
Expression cassette of WNV standard plasmid (FIG. 11B)	64	ATGGGCGGCAAGACAGGCATCGCCGTGATGATCGGC CTGATCGCCTCCGTGGGCGCCGTGACCCTGAGCAACT TCCAGGGCAAGGTGATGATGACAGTGAACGCCACAG ATGTTACCGATGTTATCACAATCCCTACCGCCGCTGG AAAGAACTTGTGCATCGTACGGGCCATGGATGTGGG GTACATGTGCGACGACACCATCACCTACGAGTGCCCT GTGCTGAGCGCCGGCAATGACCCCGAGGACATCGAC TGCTGGTGCACCAAGTCTGCCGTTTACGTGAGATATG GCAGGTGTACCAAAACCAGACACAGCAGGAGATCTC GGAGAAGCCTGACCGTGCAAACACACGGCGAGTCCA CCCTGGCCAACAAGAAGGGCGCATGGATGGACAGCA CCAAGGCCACTCGGTACCTGGTGAAGACCGAGAGCT GGATCCTGAGAAACCCTGGATACGCCCTGGTGGCCGC CGTGATTGGCTGGATGCTGGGCTCTAACACCATGCAG AGAGTGGTGTTCGTGGTGCTACTGCTCCTAGTGGCTC CTGCTTACAGCTTCAACTGCCTGGGCATGTCTAACCG GGACTTCCTGGAAGGCGTGTCCGGCGCTACATGGGTG GACCTGGTGCTCGAGGGAGATAGCTGCGTGACCATC ATGTCAAAGGACAAGCCCACCATCGACGTGAAAATG

	ATGAACATGGAAGCTGCTAATCTGGCCGAGGTCAGA TCTTACTGCTACCTGGCCACAGTGAGTGATCTGAGCA CAAAGGCCGCCTGCCCCACCATGGGCGAGGCCACA ACGATAAGCGGGCCGATCCTGCCTTCGTGTGTAGACA GGGCGTGGTGGACCGGGGATGGGGCAACGGCTGCGG GCTGTTTCGGCAAGGGCAGCATCGATACCTGTGCCAAA TTCGCCTGTAGCACCAAGGCCATCGGCCGGACCATT TGAAAGAAAACATCAAGTACGAGGTGGCTATCTTCGT GCATGGCCCTACCACCGTCGAGAGCCACGGCAACTA CTCCACACAGGTGGGCGCCACACAGGCCGGCCGATTT TCTATCACACCTGCCGCCCCCAGCTATACTGAAAC TGGGCGAGTACGGCGAAGTGACAGTGATTGCGAGC CTAGAAGCGGCATCGACACTAACGCCTACTACGTGAT GACCGTGGGCACAAAAACCTTCCTGGTTCACAGAGA GTGGTTCATGGACCTGAACCTGCCTTGGTCCAGCGCC GGCAGCACCGTGTGGCGCAATAGAGAGACACTGATG GAATTCGAGGAACCTCACGCCACCAAGCAGAGCGTG ATCGCCCTCGGTAGCCAGGAGGGCGCCCTGCACCAG GCCCTGGCTGGCGCCATCCCCGTGGAATTCTCTAGCA ACACCGTGAAACTGACCAGCGGCCACCTGAAGTGCA GAGTGAAGATGGAAAAGCTCCAAGTGAAGGGAACAA CTTACGGCGTCTGCAGCAAGGCTTTCAAGTTCCTGGG CACCCCTGCCGACACCGGACACGGAACCGTCGTGCTG GAACTGCAGTACACCGGCACAGACGGCCCATGTAAA GTGCCTATCAGCAGCGTGGCCTCCCTGAACGACCTGA CACCAGTGGGCAGACTGGTGACCGTTAATCCTTTCGT CAGCGTGGCTACTGCCAATGCCAAGGTGCTGATCGAG CTGGAACCCCCCTTCGGCGACTCTTATATCGTGGTGG GAAGAGGAGAACAACAGATCAACCACCACTGGCACA AGAGCGGTTCGTCTATCGGAAAGGCTTTTACCACCAC ACTGAAGGGCGCTCAGCGGCTGGCCGCCCTGGGCGA CACAGCCTGGGACTTCGGCAGCGTGGGCGGAGTATTT ACGTCCGTTCGGCAAGGCCGTCCATCAGGTGTTTGGAG GAGCCTTTTCGGAGCCTGTTTCGGAGGCATGAGCTGGAT CACCCAGGGCCTGCTGGGCGCGCTGCTGCTGTGGATG GGAATTAACGCCAGAGATAGAAGCATCGCCCTGACA TTCCTGGCCGTGGGCGGCGTGCTGCTGTTTCTGTCTGT GAACGTGCACGCGtaaGCGAaattgtgtgttaactgtttattgcagctta taatggttacaaataaagcaatagcatcacaatttcacaaataaagcattttttcactgcat tctagttgtggtttgtccaaactcatcaatgtatctta
--	---

[0121] In some embodiments, the viral antigen comprises the whole or an antigen fragment of any protein derived from the Influenza virus. The strain of the Influenza virus is not limited, and may be any strain that is currently known or later discovered, *e.g.*, for example, H1N1,

- H3N2, or an Influenza B strain. In some embodiments, the Influenza viral protein is the HA protein, NA protein, M1 protein, M2 protein, or any combination thereof. In some embodiments, the viral antigen comprises the whole or an antigen fragment of any protein derived from the Hepatitis B virus. In some embodiments, the Hepatitis B viral protein is the sAg (S protein), sAg (M protein), sAg (L protein), preS1, preS2, cAg (core antigen), or any combination thereof. In some embodiments, the viral antigen comprises the whole or an antigen fragment of any protein derived from the Human Papilloma virus. In some embodiments, Human Papilloma viral protein is the L1 protein of HPV 6, L1 protein of HPV 11, L1 protein of HPV 16, L1 protein of HPV 18, or any combination thereof.
- 5 [0122] In some embodiments, the viral antigen comprises the whole or an antigen fragment of any one or more of the proteins derived from each of the viruses listed below in Table 4. For instance, in some embodiments, the viral antigen may comprise the whole or an antigen fragment of any protein derived from the avian Influenza virus (H5N3). Table 4

Virus	Viral protein
Avian influenza (H5N3)	HA, NA and M1
BFDV	VP1
BTV	VP3 and VP7
Ebola	VP40 and glycoprotein
Enterovirus 71	P1 and 3CD
GHPV	VP1 and VP2
HBV	sAg (S protein)
HBV	sAg (S protein) and VSPalphaS
HBV	sAg (M protein)
HBV	Core antigen
HBV	Surface and core antigens
HBV	sAg (S protein), preS1 and preS2
HCV	Core protein, E1 and E2
HDV	HBsAg and L-HDAg
HEV	Capsid protein
HIV	Pr55gag
HIV	Pr160gag-pol
HIV	Gag protein
HIV	Pr55gag and RT

HIV	Pr55gag and TN
HPV11	L1 protein
HPV16	L1 protein
IBDV	VP2 and VP3
Influenza A	M1 and ESAT6-HA
Influenza A	HA (H1N1) and M1 (H3N2)
Influenza A	HA (H1N1) and M1 (H3N2)
Influenza A	HA (H3N2) and M1 (H1N1)
Influenza A H1N1	HA and M1
Influenza A H3N2	HA and M1
IPCV	Coat protein
JC polyomavirus	VP1
Marburg	VP40 and glycoprotein
MS2	Coat protein
NDV	HN, F, NP and MP
No	Capsid protein
No	VP1
Nv	Capsid protein
PhMV	Coat protein
PhMV	Coat protein, CPV epitopes and F protein (CDV)
Polyomavirus	VP1
PPV	VP2
RHDV	VP60
Rotavirus	VP2, VP6 and VP7
Rotavirus	VP2 and VP6
SARS	SP, EP and MP
SIV	Pr55gag and envelope protein
SV40	VP1
SVDV	P1 and 3CD

Enhancer proteins

[0123] Without being bound by any theory, it is thought that the co-expression of the enhancer proteins with an antigen, may improve one or more aspects of antigen expression, including but not limited to yield, quality, folding, posttranslational modification, activity, localization, and downstream activity, or may reduce one or more of misfolding, altered activity, incorrect posttranslational modifications, and/or toxicity.

[0124] In some embodiments, the enhancer protein is a picornavirus leader (L) protein, or a functional variant thereof. In some embodiments, the picornavirus leader (L) protein is capable of blocking the nuclear pore, thereby inhibiting nucleocytoplasmic transport (“NCT”). As used herein, the term “functional variant” refers to a protein that is homologous to the picornavirus leader (L) protein and/or shares substantial sequence similarity to the picornavirus leader (L) protein (*e.g.*, more than 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, or 99% sequence identity). In some embodiments, the functional variant shares one or more functional characteristics of the picornavirus leader (L) protein. For example, in some embodiments, a functional variant of the picornavirus leader (L) protein retains the ability to inhibit NCT.

[0125] In some embodiments, the picornavirus leader (L) protein is an L protein from the Cardiovirus, Hepatovirus, or Aphthovirus genera. For example, the enhancer protein may be from Bovine rhinitis A virus, Bovine rhinitis B virus, Equine rhinitis A virus, Foot-and-mouth disease virus, Hepatovirus A, Hepatovirus B, Marmota himalayana hepatovirus, Phopivirus, Cardiovirus A, Cardiovirus B, Theiler's Murine encephalomyelitis virus (TMEV), Vilyuisk human encephalomyelitis virus (VHEV), Theiler-like rat virus (TRV), or Saffold virus (SAF-V).

[0126] In some embodiments, the picornavirus leader (L) protein is the L protein of Theiler's virus or a functional variant thereof. In some embodiments, the L protein shares at least 90% identity to SEQ ID NO: 1. In some embodiments, the enhancer protein may comprise or consist of SEQ ID NO: 1. In some embodiments, the enhancer protein may share at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, at least 99% or 100% identity to SEQ ID NO: 1.

[0127] In some embodiments, the picornavirus leader (L) protein is the L protein of Encephalomyocarditis virus (EMCV) or a functional variant thereof. In some embodiments, the L protein may share at least 90% identity to SEQ ID NO: 2. In some embodiments, the enhancer protein may comprise or consist of SEQ ID NO: 2. In some embodiments, the

enhancer protein may share at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, at least 99% or 100% identity to SEQ ID NO: 2.

[0128] In some embodiments, the nucleic acid sequence encoding the enhancer protein may comprise or consist of SEQ ID NO: 68. In some embodiments, the nucleic acid sequence encoding the enhancer protein may share at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, at least 99% or 100% identity to SEQ ID NO: 68.

[0129] In some embodiments, the picornavirus leader (L) protein is selected from the group consisting of the L protein of poliovirus, the L protein of HRV16, the L protein of mengo virus, and the L protein of Saffold virus 2 or a functional variant thereof.

[0130] In some embodiments, the picornavirus leader (L) protein is selected from the proteins listed in Table 5 or functional variants thereof. The polynucleotide encoding the picornavirus leader (L) protein may encode an amino acid sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, at least 99% or 100% identical to an amino acid sequence listed in Table 2. The amino acid sequence of the picornavirus leader (L) protein may be at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, at least 99% or 100% identical to an amino acid sequence listed in Table 2. The amino acid sequence of the picornavirus leader (L) protein may be at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, at least 99% or 100% identical to the amino acid sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, or 12. In some embodiments, an enhancer protein may have an amino acid sequence comprising, or consisting of, one of the amino acid sequences listed in Table 2. In some embodiments, an enhancer protein may have an amino acid sequence comprising or consisting of the amino acid sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, or 12.

Table 5: Illustrative enhancer proteins

Nuclear pore blocking viral protein	Origin	Family	Amino acid sequence
Leader protein	Theiler's virus	Picornaviridae	MACKHGYPDVCPICATAVDAT PGFEYLLMADGEWYPTDLLC VDLDDDVFWPSDTSNQSQTM DWTDVPLIRDIVMEPQ (SEQ ID NO: 12)

Nuclear pore blocking viral protein	Origin	Family	Amino acid sequence
Leader protein	Theiler's-like virus	Picornaviridae	MACKHGYPLMCPLCTALDK TSDGLFTLLFDNEWYPTDLLT VDLEDEVFYPDDPHMEWTDL PLIQDIEMEPQ (SEQ ID NO: 1)
Leader protein	EMCV	Picornaviridae	MATTMEQETCAHSLTFEECP KCSALQYRNGFYLLKYDEEW YPEELLTDGEDDVFDPELDM EVVFELQ (SEQ ID NO: 2)
Leader protein	Poliovirus (Enterovirus C)	Picornaviridae	NYHLATQDDLQNAVNVMS RDLLVTESRAQGTDSIARCNC NAGVYYCESRRKYYPVSFVG PTFQYMEANNYPARYQSH MLIGHGFASPGDCGGILRCHH GVIGIITAGGEGLVAFSDIRDL YAYEE (SEQ ID NO: 3)
Leader protein	Equine rhinitis B virus 1	Picornaviridae	MVTMAGNMICNVFAGLATEI CSPKQGPLLDNELPLPLELAE FPNKDNNCWVAALSHYYTL CDVTNHVTKVPTTSGIRYYL TAWQSILQTDLFNGYYPAAF AVETGLCHGPFPMQQHGYVR NATSHPYNFCLCSEPVPGEDY WHAVVKVDLSRTEARVDKW LCIDDDRMVLSGPPTRVKLAS SYKIPTWIESLAQFCLQLHPV QHRRTLANSRNEQCR (SEQ ID NO: 4)
Leader protein	Mengo virus (Cardiovirus)	Picornaviridae	MATTMEQEICAHSMTFEECP KCSALQYRNGFYLLKYDEEW YPEESLTDGEDDVFDPLDM EVVFETQ (SEQ ID NO: 5)
Leader protein	Saffold virus 2 (Cardiovirus)	Picornaviridae	MACKHGYPLCPLCTAIDTT HDGSFTLLIDNEWYPTDLLTV DLDDDVHFPDDSVMEWTDL PLIQDVVMEPQ (SEQ ID NO: 6)

[0131] The antigens, enhancer proteins, and/or fusion proteins, or the polynucleotides encoding such, may be modified to comprise one or more markers, labels, or tags. For example, in some embodiments, a protein of the present disclosure may be labeled with any label that will allow its detection, *e.g.*, a radiolabel, a fluorescent agent, biotin, a peptide tag, an enzyme
5 fragment, or the like. The proteins may comprise an affinity tag, *e.g.*, a His-tag, a GST-tag, a Strep-tag, a biotin-tag, an immunoglobulin binding domain, *e.g.*, an IgG binding domain, a calmodulin binding peptide, and the like. In some embodiments, polynucleotides of the present disclosure comprise a selectable marker, *e.g.*, an antibiotic resistance marker.

Viral Packaging Signal

10 [0132] In some embodiments, the vectors disclosed herein comprise a polynucleotide sequence encoding a viral packaging signal (interchangeably referred to herein as “viral packaging sequence” or packaging signal” or “psi sequence”). In some embodiments, the polynucleotide sequence encoding a viral packaging signal is a DNA polynucleotide, an RNA polynucleotide, or a combination thereof. In some embodiments, the viral packaging signal is an RNA
15 polynucleotide. In some embodiments, the vectors comprise more than one copy of the polynucleotide sequence encoding a viral packaging signal, for example, 2, 3, 4 or 5 copies of the polynucleotide sequence.

[0133] The viral packaging signal may be derived from any virus. In some embodiments, the viral packaging signal is derived from the same virus as the antigens that are expressed from
20 the vector. In some embodiments, the viral packaging signal is derived from a different virus as the antigens that are expressed from the vector. In some embodiments, the viral packaging signal is derived from a virus selected from the group consisting of SARS-CoV-2, SARS-CoV-1, MERS-CoV, chikungunya virus, African Swine Fever virus, Dengue virus, Zika virus, Influenza virus (*e.g.*, A, B, C), Human Immunodeficiency Virus (HIV), Ebola virus, Hepatitis
25 virus (*e.g.*, Hepatitis A, Hepatitis B, Hepatitis C, Hepatitis D, and Hepatitis E), herpes simplex virus type 1 (HSV-1), herpes simplex virus type 2 (HSV-2), West Nile virus, and Human Papillomavirus.

[0134] In some embodiments, the polynucleotide encoding the viral packaging element has at least about 70% identity (for example, about 75%, about 80%, about 85%, about 90%, about
30 95%, about 96%, about 97%, about 98%, about 99%, about 99.5%, or about 100%), including all values and subranges that lie therebetween, to the polynucleotide of SEQ ID NO: 34.

[0135] The location of the polynucleotide encoding the viral packaging signal on the vector is not limited. In some embodiments, the location of the polynucleotide encoding the viral packaging signal on the vector may be 5' to all the nucleic acid sequences encoding the viral antigens. In some embodiments, the location of the polynucleotide encoding the viral packaging signal on the vector may be 3' to all the nucleic acid sequences encoding the viral antigens. In some embodiments, the location of one copy of the polynucleotide encoding the viral packaging signal on the vector is 5' to all the nucleic acid sequences encoding the viral antigens, and the location of the other copy of the polynucleotide encoding the viral antigen is 3' to all the nucleic acid sequences encoding the viral antigens. Further, the size of the viral packaging signal is not limited and may be in the range of about 50 bps to about 3 kb, for example, about 100 bps, about 200 bps, about 300 bps, about 400 bps, about 500 bps, about 550 bps, about 600 bps, about 650 bps, about 700 bps, about 800 bps, about 900 bps, about 1 kb, about 2 kb, or about 3 kb, including all values and subranges that lie therebetween. In some embodiments, the size of the viral packaging signal is about 600 to about 700 bps, for example, about 650 bps. In some embodiments, the size of the viral packaging signal is about 661 bps.

[0136] The disclosure further provides vectors, comprising an expression cassette, said expression cassette comprising a promoter linked to a target gene, wherein the vector comprises a polynucleotide encoding any one of the viral packaging elements disclosed herein.

Order of the Genetic Elements in the Expression Cassette

[0137] In the vectors disclosed herein, the polynucleotides sequences encoding one or more viral antigens, and the polynucleotide sequence encoding the enhancer, and/or one or more regulatory elements (*e.g.*, a polynucleotide encoding the IRES sequence, the CMV protein, a polynucleotide encoding the viral packaging signal, and a polynucleotide encoding the proteolytic cleavage site) may be ordered in any possible combination. For instance, the order of elements in the expression cassette may be as depicted for any one of the plasmids CoVEG 3-17 in FIG. 6. Without being bound by a theory, it is thought that the order of elements in the expression cassette might be related to the expression of antigens encoded by the vector, and/or formation of VLPs. Furthermore, it is thought that when the expression cassette comprises the genes in the following order from 5' to 3' – M, N, S, and E – it might result in higher protein expression and more stable VLP formation.

[0138] In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG3. In some embodiments, the vector comprises

an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a first polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a second polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a second polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an N protein wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, and a polynucleotide encoding an E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto.

[0139] In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG4. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto.

[0140] In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG5. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ

ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, and a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto.

10 **[0141]** In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG6. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, and a polynucleotide encoding a E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto.

30 **[0142]** In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG7. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein wherein the E protein

comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto.

[0143] In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG8. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, and a polynucleotide encoding a viral packaging signal wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 9% identical thereto.

[0144] In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG9. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, wherein the S protein comprises SEQ ID NO: 51 or SEQ ID NO: 52 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a

polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto.

[0145] In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG10. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a mutated S protein wherein the S protein comprises SEQ ID NO: 51 or SEQ ID NO: 52 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, and a polynucleotide encoding a viral packaging wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 9% identical thereto.

[0146] In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG11. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a first polynucleotide encoding a viral packaging wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 9% identical thereto, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a mutated S protein wherein the S protein comprises SEQ ID NO: 51 or SEQ ID NO: 52 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ

ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, and a second polynucleotide encoding a viral packaging signal wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 9% identical thereto.

[0147] In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG12. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, and a polynucleotide encoding a viral packaging signal wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 9% identical thereto.

[0148] In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG13. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a first polynucleotide encoding a viral packaging signal (wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 9% identical thereto, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein

wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, and a second polynucleotide encoding a viral packaging signal wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 9% identical thereto.

[0149] In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG14. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a first polynucleotide encoding a viral packaging wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 9% identical thereto, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding ORF3a wherein the ORF3a encodes SEQ ID NO: 53, or an amino acid sequence with at least 95% identity thereto, and a second polynucleotide encoding a viral packaging signal wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 9% identical thereto.

[0150] In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG15. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a first polynucleotide encoding a viral packaging signal wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic

cleavage site, a polynucleotide encoding an N protein wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a mutated S wherein the S protein comprises SEQ ID NO: 51 or SEQ ID NO: 52 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, and a second polynucleotide encoding a viral packaging signal (wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 9% identical thereto.

[0151] In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG16. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding ORF3a encodes SEQ ID NO: 53, or an amino acid sequence with at least 95% identity thereto, and a polynucleotide encoding a viral packaging signal, wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 95% identical thereto..

[0152] In some embodiments, the vector comprises an expression cassette, comprising the elements in the same 5' to 3' order as CoVEG17. In some embodiments, the vector comprises an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a first polynucleotide encoding a viral packaging signal wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 9% identical thereto, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an

amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a mutated S wherein the S protein comprises SEQ ID NO: 51 or SEQ ID NO: 52 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding ORF3a wherein the ORF3a protein comprises SEQ ID NO: 53 and a second polynucleotide encoding a viral packaging signal wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide signal with at least 95% identity thereto.

15 **Expression of Antigens and VLPs in Cells**

[0153] The disclosure provides methods of expressing an antigen in a eukaryotic cell, comprising contacting the cell with any one of the vectors disclosed herein. In some embodiments, the vector is contacted with the cell *in vitro*, *ex vivo* or *in vivo*. In some embodiments, the vector is contacted with the cell (*in vivo*) in a subject.

20 **[0154]** In some embodiments, the expression of one or more antigens results in the formation of a virus like particle (VLP). In some embodiments, a VLP is immunogenic. In some embodiments, a VLP is capable of eliciting an immune response in a subject. In some embodiments, the VLP is enveloped. In some embodiments, the VLP is non-enveloped. The number of antigens present in a VLP is not limited. In some embodiments, a VLP comprises

25 one antigen, two antigens, three antigens, four antigens, five antigens, six antigens, seven antigens, eight antigens, nine antigens, ten antigens, or a higher number of antigens. In some embodiments, the VLP comprises three antigens. In some embodiments, the VLP comprises four antigens. In some embodiments, the structural proteins that form a VLP and the immunogenic viral antigens that are a part of the VLP are derived from the same virus (i.e., a native VLP).

30 In some embodiments, the structural viral proteins that form a VLP are derived from one virus and the immunogenic viral antigens that get incorporated to that said VLP are derived from another virus (i.e., a chimeric VLP). In some embodiments, the viral proteins are

mutated to enhance VLP assembly, VLP secretion and/or loading of the immunogenic antigen or antigens to the said VLP.

[0155] In some embodiments, the vector comprises a DNA polynucleotide encoding a viral packaging signal, such that contacting the cell with the vector results in expression of the viral packaging signal. In some embodiments, the VLPs encapsidate the viral packaging signal. In some embodiments, the expression of the viral packaging signal increases or promotes the formation of VLPs. In some embodiments, a greater number of VLPs are formed in the presence of a viral packaging signal, as compared to in the absence of a viral packaging signal. In some embodiments, contacting the cell with any one of disclosed vectors encoding the viral packaging signal results in the expression of a greater number of VLPs, as compared to a control vector lacking the DNA polynucleotide encoding the viral packaging signal. In some embodiments, contacting the cell with any one of disclosed vectors encoding the viral packaging signal results in the packaging of the viral packaging signal within the VLPs, which in turn leads to enhanced immune response due to an improved adjuvating characteristics or other mechanisms. In some embodiments, the packaging signals and proteins are derived from the same virus from which the VLP is formed (i.e., native packaging). In some embodiments, the packaging signals and proteins are derived from another virus with a known packaging mechanism (i.e., chimeric packaging).

[0156] In some embodiments, the expression cassette comprises a polynucleotide sequence encoding a first antigen, a second antigen, a third antigen, a fourth antigen, or a combination thereof. In some embodiments, the expression cassette comprises a polynucleotide sequence encoding a first antigen, a second antigen, and a third antigen. In some embodiments, the expression cassette comprises a polynucleotide sequence encoding a first antigen, a second antigen, a third antigen, and a fourth antigen.

[0157] In some embodiments, the first antigen is a coronavirus spike protein, the second antigen is a coronavirus membrane (M) protein, and the third antigen is a coronavirus envelope (E) protein. In some embodiments, wherein the first antigen is a coronavirus spike protein, the second antigen is a coronavirus membrane (M) protein, the third antigen is a coronavirus envelope (E) protein and the fourth antigen is a coronavirus nucleocapsid (N) protein.

[0158] In some embodiments, the vector causes: (i) expression of the antigen at a higher expression level; and/or (ii) expression of the antigen for a longer period of time; and/or (iii) expression of the antigen with better protein quality, as compared to a vector lacking the enhancer protein. In some embodiments, the vector causes: (i) expression of a virus like particle (VLP) comprising the antigen at a higher expression level; and/or (ii) expression of a VLP

comprising the antigen for a longer period of time; and/or (iii) expression of a VLP comprising the antigen with better protein quality, than a vector lacking the enhancer protein. As used herein, “protein quality” might refer to without limitation, protein folding, posttranslational modification, functional activity, localization, and downstream activity. Thus, in some
5 embodiments, the antigen which is co-expressed with an enhancer protein using any of the methods or vectors or compositions disclosed herein may have improved protein folding, improved posttranslational modification, improved functional activity, improved localization, and improved downstream activity, as compared to the antigen which is not co-expressed with an enhancer protein.

- 10 **[0159]** As used herein, the terms “transfection,” “transduction,” and “transformation” refer to the process of introducing nucleic acids into cells (*e.g.*, eukaryotic cells). The vectors disclosed herein may be introduced into a cell (*e.g.*, a eukaryotic cell) using any method known in the art. For example, the vector can be introduced into a cell using chemical, physical, biological, or viral means. Methods of introducing a vector into a cell include, but are not limited to, the
15 use of calcium phosphate, dendrimers, cationic polymers, lipofection, fugene, cell-penetrating peptides, peptide dendrimers, electroporation, cell squeezing, sonoporation, optical transfection, protoplast fusion, impalefection, hydrodynamic delivery, gene gun, magnetofection, particle bombardment, nucleofection, viral transduction, injection, transformation, transfection, direct uptake, projectile bombardment, and liposomes. Other non-
20 limiting examples of methods include viral transfection, direct uptake, projectile bombardment, direct injection with or without electroporation/sonoporation while using or not using cationic polymers, lipids, lipid formulations, and jet-gene devices. Antigens and enhancer proteins can be stably or transiently expressed in cells using expression vectors. Techniques of expression in eukaryotic cells are well known to those in the art. (See Current Protocols in Human
25 Genetics: Chapter 12 “Vector Therapy” & Chapter 13 “Delivery Systems for Gene Therapy”).
- [0160]** In some embodiments, vectors can be introduced into a host cell by insertion into the genome using standard methods to produce stable cell lines, optionally through the use of lentiviral transfection, baculovirus gene transfer into mammalian cells (BacMam), retroviral transfection, CRISPR/Cas9, and/or transposons. In some embodiments, polynucleotides or
30 vectors can be introduced into a host cell for transient transfection. In some embodiments, transient transfection may be effected through the use of viral vectors, helper lipids, *e.g.*, PEI, Lipofectamine, and/or Fectamine 293. The genetic elements can be encoded as DNA on *e.g.* a vector or as RNA from *e.g.* PCR. The genetic elements can be separated in different or combined on the same vector.

[0161] The host cell used to express the antigen and enhancer protein is not limited, and may include a prokaryotic host (*e.g.*, *E. coli*) or a eukaryotic host (*e.g.*, *Saccharomyces cerevisiae*, insect cells, *e.g.*, Sf21 cells, or mammalian cell lines and primary cells, *e.g.*, NIH 3T3, HeLa, COS cells). Such cells are available from a wide range of sources (*e.g.*, the American Type Culture Collection, Rockland, Md.; also, see, *e.g.*, Ausubel et al., *supra*). Non limiting
5 examples of insect cells are, *Spodoptera frugiperda* (SI) cells, *e.g.* Sf9, Sf21, *Trichoplusia ni* cells, *e.g.* High Five cells, and *Drosophila* S2 cells. Examples of fungi (including yeast) host cells are *S. cerevisiae*, *Kluyveromyces lactis* (K lactis), species of *Candida* including *C. albicans* and *C. glabrata*, *Aspergillus nidulans*, *Schizosaccharomyces pombe* (*S. pombe*),
10 *Pichia pastoris*, and *Yarrowia lipolytica*. Examples of mammalian cells are COS cells, baby hamster kidney cells, mouse L cells, LNCaP cells, Chinese hamster ovary (CHO) cells, human embryonic kidney (HEK) cells, African green monkey cells, CV1 cells, HeLa cells, MDCK cells, Vero and Hep-2 cells. *Xenopus laevis* oocytes, or other cells of amphibian origin, may also be used. Prokaryotic host cells include bacterial cells, for example, *E. coli*, *B. subtilis*, and
15 mycobacteria.

Vaccine Compositions

[0162] The disclosure provides vaccine compositions comprising any one of the vectors disclosed herein, and at least one pharmaceutically acceptable carrier, excipient, and/or vehicle, for example, solvents, buffers, solutions, dispersion media, coatings, antibacterial and
20 antifungal agents, isotonic and absorption delaying agents. As used herein, the term “pharmaceutically acceptable” means being approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia, European Pharmacopeia or other generally recognized pharmacopeia for use in mammals, and more particularly in humans. In some embodiments, the pharmaceutically acceptable carrier, excipient, and/or vehicle may
25 comprise saline, buffered saline, dextrose, water, glycerol, sterile isotonic aqueous buffer, and combinations thereof. In some embodiments, the pharmaceutically acceptable carrier, excipient, and/or vehicle comprises phosphate buffered saline, sterile saline, lactose, sucrose, calcium phosphate, dextran, agar, pectin, peanut oil, sesame oil, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium
30 carbonate, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol, and the like) or suitable mixtures thereof. In some embodiments, the compositions disclosed herein further comprise minor amounts of emulsifying or wetting agents, or pH buffering agents.

[0163] In some embodiments, the composition is in a solid form, *e.g.* a lyophilized powder suitable for reconstitution, a liquid solution, suspension, emulsion, tablet, pill, capsule, sustained release formulation, or powder. In some embodiments, delivery vehicles *e.g.* liposomes, nanocapsules, nanoparticles, microparticles, microspheres, lipid particles, vesicles, polymers, peptides, and the like, may be used for the introduction of the vectors and vaccine compositions disclosed herein into suitable host cells. In some embodiments, the vectors and vaccine compositions disclosed herein may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like.

[0164] In some embodiments, the compositions disclosed herein comprise other conventional pharmaceutical ingredients, *e.g.* preservatives, or chemical stabilizers, *e.g.* chlorobutanol, potassium sorbate, sorbic acid, sulfur dioxide, propyl gallate, the parabens, ethyl vanillin, glycerin, phenol, parachlorophenol or albumin. In some embodiments, the compositions disclosed herein comprise antibacterial and antifungal agents, *e.g.*, parabens, chlorobutanol, phenol, sorbic acid or thimerosal; isotonic agents, *e.g.*, sugars or sodium chloride and/or agents delaying absorption, *e.g.*, aluminum monostearate and gelatin.

[0165] In some embodiments, the vaccine composition comprises an adjuvant. As used herein, the term “adjuvant” refers to a compound that, when used in combination with an immunogen, augments or otherwise alters or modifies the immune response induced against the immunogen. Modification of the immune response may include intensification or broadening the specificity of either or both antibody and cellular immune responses.

[0166] In some embodiments, the adjuvant is alum. In some embodiments, the adjuvant is monophosphoryl lipid A (MPL). In some embodiments, other adjuvants may be used in addition or as an alternative. The inclusion of any adjuvant described in Vogel et al., "A Compendium of Vaccine Adjuvants and Excipients (2nd Edition)," herein incorporated by reference in its entirety for all purposes, is envisioned within the scope of this disclosure. Other adjuvants include complete Freund's adjuvant (a non-specific stimulator of the immune response containing killed *Mycobacterium tuberculosis*), incomplete Freund's adjuvants and aluminum hydroxide adjuvant, GMCSP, BCG, MDP compounds, *e.g.* thur-MDP and nor-MDP, CGP (MTP-PE), lipid A, and monophosphoryl lipid A (MPL), MF-59, RIBI, which contains three components extracted from bacteria, MPL, trehalose dimycolate (TDM) and cell wall skeleton (CWS) in a 2% squalene/Tween® 80 emulsion. In some embodiments, the adjuvant may be a paucilamellar lipid vesicle; for example, Novasomes®. Novasomes® are paucilamellar nonphospholipid vesicles ranging from about 100 nm to about 500 nm. They comprise Brij 72, cholesterol, oleic acid and squalene. Novasomes have been shown to be an

effective adjuvant (see, U.S. Pat. Nos. 5,629,021, 6,387,373, and 4,911,928). In some embodiments, the compositions may be free of added adjuvant. Alum-free compositions that induce robust immune responses are especially useful in adults about 60 and older.

Methods of eliciting an immune response in a subject

5 [0167] The disclosure further provides methods of eliciting an immune response in a subject, comprising administering an effective amount of any one of the vaccine compositions disclosed herein to the subject. In some embodiments, tissue at an administration site of the subject expresses the antigen and/or a VLP comprising the antigen. In some embodiments, tissue at an administration site of the subject: (i) expresses the antigen and/or a VLP comprising the antigen
10 at a higher expression level; and/or (ii) expresses the antigen and/or a VLP comprising the antigen for a longer period of time; and/or (iii) expresses the antigen and/or a VLP comprising the antigen with better protein quality, as compared to when a vector lacking the enhancer protein is administered.

[0168] In some embodiments, the method elicits an antibody response in the subject. In some
15 embodiments, the antibody response is a neutralizing antibody response. In some embodiments, the method elicits a cellular immune response. In some embodiments, the method elicits a prophylactic, protective and/or therapeutic immune response in the subject.

[0169] In some embodiments, the vector comprises a DNA polynucleotide encoding a viral packaging signal, such that the tissue at an administration site of the subject expresses the viral
20 packaging signal. In some embodiments, the VLPs encapsidate the viral packaging signal. In some embodiments, the VLPs encapsidate a polynucleotide comprising the viral packaging signal. In some embodiments, the VLPs encapsidate a polynucleotide consisting of the viral packaging signal. In some embodiments, the VLPs encapsidating the viral packaging signal are more immunogenic than control VLPs comprising the antigen but lacking the viral packaging
25 signal. Without being bound by a theory, it is thought that a greater number of VLPs may be formed in the presence of a viral packaging signal, as compared to in the absence of the viral packaging signal. Thus, in some embodiments, the disclosed vectors encoding a viral packaging signal promote the formation of a greater number of VLPs, as compared to a control vector which does not encode the viral packaging signal. Without being bound by a theory, it
30 is also thought that the RNA viral packaging signals may act as an adjuvant by acting as an agonist of Toll-like Receptors (TLRs).

[0170] Methods of administering any one of the compositions or vectors disclosed herein include, but are not limited to, parenteral administration (*e.g.*, intradermal, intramuscular, intravenous and subcutaneous), epidural, and mucosal (*e.g.*, intranasal and oral or pulmonary routes or by suppositories), subdermal, and intraperitoneal. In some embodiments, compositions of the present invention are administered intramuscularly, intravenously, subcutaneously, transdermally or intradermally. The compositions or vectors may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (*e.g.*, oral mucous, colon, conjunctiva, nasopharynx, oropharynx, vagina, urethra, urinary bladder and intestinal mucosa, etc.) and may be administered together with other biologically active agents. In some embodiments, the administration is intradermal administration. In some embodiments, the administration is intramuscular administration. In some embodiments, the administration is subcutaneous administration. In some embodiments, the administration is intranasal administration. In some embodiments, the compositions or vectors disclosed herein are administered by injection. In some embodiments, the injection is performed using a needle, a syringe, a microneedle, or a needle-less injection device. In some embodiments, the compositions or vectors disclosed herein are administered intranasally, either by drops, large particle aerosol (greater than about 10 microns), or spray into the upper respiratory tract or small particle aerosol (less than 10 microns) or spray into the lower respiratory tract. In some embodiments, the injection is followed by electroporation.

[0171] The vectors or vaccine compositions disclosed herein may be administered on a single dose schedule or a multiple dose schedule. Multiple doses may be used in a primary immunization schedule or in a booster immunization schedule. In a multiple dose schedule the various doses may be given by the same or different routes *e.g.*, a parenteral prime and mucosal boost, a mucosal prime and parenteral boost, etc. In some aspects, a follow-on boost dose is administered within a time period of about 1 hour to about several years (for example, about 12 hours, about 1 day, about 2 days, about 5 days, about 1 week, about 2 weeks, about 3 weeks, about 4 weeks, about 5 weeks, about 6 weeks, about 1 month, about 6 months, about 1 year, about 2 years, including all values and subranges that lie there between) after the prior dose.

30 Immunogenic effects

[0172] In some embodiments, inclusion of the enhancer protein in a polynucleotide encoding one or more viral antigen proteins increases functional viral-like particle (VLP) production relative to a polynucleotide without an enhancer protein. In some embodiments, inclusion of

the enhancer protein increases functional VLP production by about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 100%, about 125%, about 150%, about 175%, about 200%, about 250%, about 300%, about 350%, about 400%, about 500%, or about 1000% relative to a vector without an enhancer protein.

- 5 Functional VLP production as used herein may be measured by method known in the art, including but not limited to: the level of protein aggregation, the titer of neutralizing antibodies in vivo, induced Th1 response, the amount of VLPs over time relative to VLP half-life, and/or cell death associated with mis-folded VLPs.

- [0173] In some embodiments, inclusion of the enhancer protein in a polynucleotide encoding one or more viral antigen proteins increases the duration or the amount of neutralizing antibodies in a subject relative to a vaccine composition without an enhancer protein. In some embodiments inclusion of the enhancer protein increases the duration or the amount of neutralizing antibodies in a subject by about 2-fold, about 3-fold, about 4-fold, about 5-fold, about 6-fold, about 7-fold, about 8-fold, about 9-fold, or about 10-fold relative to a vaccine composition without an enhancer protein.

- [0174] In some embodiments, inclusion of the enhancer protein increases Th1 cellular response relative to a vaccine composition without an enhancer protein. In some embodiments, inclusion of the enhancer protein increases Th1 response by about 25%, about 50%, about 100%, about 200%, about 300%, about 400%, about 500%, or about 1000% relative to a vaccine composition without an enhancer protein.

Kits and Articles of Manufacture

- [0175] The disclosure provides kits comprising any one or more of the vectors disclosed herein. The disclosure further provides kits comprising any one or more of the polynucleotides disclosed herein. The disclosure also provides kits comprising any one or more of the vaccine compositions disclosed herein. The kits disclosed herein are useful for performing, or aiding in the performance of, the disclosed methods. In some embodiments, the kits comprise a pharmaceutically acceptable carrier. In some embodiments, the kits comprise instructions for proper use and safety information of the product or formulation. In some embodiments, the kits comprise dosage information based on the application and method of administration as determined by a doctor.

[0176] The present application also provides articles of manufacture comprising any one of the vaccine compositions or kits described herein. Examples of an article of manufacture include vials (*e.g.* sealed sterile vials).

5 [0177] In some embodiments, the kits comprise one or more containers or vials filled with one or more of the ingredients of the vaccine compositions disclosed herein. In some embodiments, the kit comprises two containers, one containing the vector, or polynucleotide, or vaccine composition disclosed herein, and the other containing an adjuvant. In some embodiments, the kits further comprise a notice reflecting approval by a governmental agency for manufacture, use or sale for human administration.

10 [0178] The inventions are further illustrated by the following additional examples that should not be construed as limiting. Those of skill in the art, in light of the present disclosure, would be able to appreciate that many changes can be made to the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the inventions.

15

EXAMPLES

Example 1: Construction of polynucleotides encoding SARS-CoV-2 and L protein

[0179] CoVEG1 and CoVEG2 plasmids encode SARS-CoV-2 and the L enhancer protein.

20 [0180] Plasmid CoVEG1 comprises polynucleotides encoding viral proteins of full-length S protein (SEQ ID NO: 14), M protein (SEQ ID NO: 19), and E protein (SEQ ID NO: 23) of SARS-CoV-2. Plasmid CoVEG2 comprises polynucleotides encoding viral proteins of full-length S protein, M protein, N protein (SEQ ID NO: 21) and E protein of SARS-CoV-2. The backbone of CoVEG1 and CoVEG2 plasmids is shown in FIG. 1. In addition, the CoVEG1 and CoVEG2 plasmids also comprise a polynucleotide encoding the L protein from EMCV (SEQ ID NO: 16).

25 [0181] The nucleic acid sequence of the complete insert in CoVEG2 is represented by SEQ ID NO: 30. See Table 1. The expression of this construct gives rise to three polypeptides: the SARS-CoV-2 Spike protein having amino acid sequence of SEQ ID NO: 13, CoVEG2 polypeptide 1 having amino acid sequence of SEQ ID NO: 25, and CoVEG2 polypeptide 2 having amino acid sequence of SEQ ID NO: 26. The nucleic acid sequence of the insert in
30 CoVEG1 is represented by SEQ ID NO: 31. The expression of this construct gives rise to two polypeptides: the SARS-CoV-2 Spike protein having amino acid sequence of SEQ ID NO: 13, and CoVEG1 polypeptide having amino acid sequence of SEQ ID NO: 32. See Table 1.

[0182] The plasmid backbone (based on the design principles of the pVax1 plasmid) and insert for both the plasmids were generated using gene synthesis and do not contain any animal or human source material. The plasmid backbone consists of a Kanamycin resistance gene, the ColE1 origin of replication, the Human cytomegalovirus immediate-early promoter and Simian virus (SV40) Poly A signal. Polynucleotides encoding viral proteins were cloned in between the CMV promoter and the SV40 PolyA signal. After gene synthesis and plasmid preparation, the plasmid was transformed into *E. coli* for cloning and then screened using kanamycin. A representative colony was selected, and its plasmid sequence verified and used as source plasmid for further development. After transcription, the viral proteins were expressed from a single polycistronic mRNA.

[0183] Without being bound by any particular theory, it is thought that when co-expressed, the S, E and M proteins assemble into VLPs, and are secreted by expressing cells; and that the VLP secretion is significantly increased when N protein is also expressed together with the S, E and M proteins.

Example 2: Expression of SARS-CoV-2 S, E, M, and N proteins in eukaryotic cells observed by immunofluorescence

[0184] HEK 293 (eukaryotic) cells were transfected with the pCoVEG2 plasmid. Twenty-four hours later, cells were fixed, permeabilized and analyzed by immunocytochemistry using commercial Alexa Fluor 568 fluorescently labelled secondary antibodies for detection. FIG. 2 shows the expression of S, M, N and E proteins in HEK 293 cells, demonstrating that pCoVEG2 disclosed herein expresses the viral antigens in cells.

Example 3: Expression of SARS-CoV-2 S, E, M, and N proteins in eukaryotic cells observed by SDS-PAGE and western blot

[0185] HEK 293 cells were transfected with the pCoVEG2 plasmid and incubated for 96 hours. Thereafter, cell culture supernatant was harvested and concentrated. The concentrate was run over Superose 6 GL resin packed in the Tricorn 10/300 column using PBS as eluant. The void fraction, which contains secreted VLPs, was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and/or western blotting using monoclonal antibodies against S, N, M or E to demonstrate the presence of S, N, M and E proteins. See FIG. 3.

[0186] These data demonstrate that the DNA vector CoVEG2 disclosed herein expresses all SARS-CoV-2 viral structural proteins (S, M, E, and N proteins) in HEK 293 cells and that secreted VLPs components can be detected in cell culture supernatants. These results suggest that CoVEG1 and CoVEG2 plasmids could potentially be used as highly effective DNA
5 vaccines against SARS-CoV-2.

Example 4: Immunogenicity of polynucleotide vaccine

[0187] To determine the immunogenicity of CoVEG1 and CoVEG2, these plasmids are injected intradermally into 6 weeks old BALB/c mice in 2 week intervals, for a total of 3 injections at Day 1, 15, and 29. The elicited humoral immune response [the titer of anti-S
10 antibody using a respective enzyme linked immunosorbent assay (ELISA)] as well as cellular immune response [the presence of antigen reactive T cells using a respective IFN- γ and IL-4 enzyme-linked immune absorbent Spot (Dual color ELISpot) assay] is measured. To measure the neutralizing versus total antibodies, in vitro viral neutralization assays are performed. For this, isolated serum from day 43 is diluted and incubated with SARS-CoV-2 life virus before
15 adding to VERO cells. Virus isolation is determined by the absence of successful infection of the cells compared to the native virus.

[0188] Anti-SARS-CoV-2 antibody analysis comprising anti-S protein antibody ELISA assay is performed based on commercially available materials. Alternatively, in-house developed cell-based and VLP-based ELISA assays is used. For ELISpot analysis, spleen is collected and
20 T cells are isolated. ELISpot assessment is performed by priming the T cells with Miltenyi Biotec Peptivator SARS-CoV-2 peptide pools to activate SARS-CoV-2 reactive T cells. In addition, the toxicokinetic and pharmacodynamic characteristics of the plasmids are determined. See FIG. 4.

[0189] Female BALB/c mice (6-8 weeks of age) weighing 15 to 25 grams are randomly
25 assigned to 4 groups with each group containing 10 animals. Mice are dosed intradermally with either the vehicle -PBS, a reference item EG-BB, which encodes the enhancer protein(s) under the control a CMV promoter, and two doses of CoVEG1 and CoVEG2 at 1 and 25 μ g. Mice are evaluated twice daily for mortality and moribundity. Clinical observations and body weights are collected weekly starting Week -1 and thereafter at least every 2 weeks during the
30 study period.

[0190] Dosed mice are bled at pre-defined timepoints before dosing and serum are separated by centrifugation. The obtained serum samples are then analyzed for antibodies against the full length recombinant S protein (S1+S2) using a quantitative ELISA, as shown below in Table 3.

Table 3: Anti-Vaccine Antibody sample collection

Group Nos.	Time Points		
	Day 15 ^a	Day 29 ^a	Day 43
1-4	X	X	X
Method/Comments:	Jugular venipuncture (Day 15 and 29) or from the abdominal aorta under isoflurane anesthesia (Day 43 and at unscheduled euthanasia)		
Target Volume ^b (mL):	0.12 mL (Day 15 and 29) or as much as possible (Day 43 or unscheduled euthanasia)		
Anticoagulant:	None, in SST		
Special Requirements:	None		
Processing:	Serum		
X = Sample collected; SST = serum separator tube			
^a Sample collected before dosing.			
^b Additional blood samples obtained (e.g., due to sample quality) if permissible sampling frequency and blood volume are not exceeded.			

5

[0191] For the Day 43 time point, the resultant serum is split into 2 approximately equal aliquots; the first aliquot will be used for anti-vaccine antibody (AVA) analysis and the second aliquot kept for testing for neutralizing antibodies. The aliquots are frozen immediately over dry ice or in a freezer set to maintain -80°C.

10 [0192] At the end of the study, all animals are euthanized. Spleens are collected using cell culture clean procedures for IFN- γ and IL-4 evaluation by ELISpot.

[0193] For evaluation of T-cell mediated toxicity, a quantitative assessment is performed using ELISpot assay. Splenocytes from harvested spleens are stimulated with Miltenyi Biotec's

SARS-CoV-2 PepTivator Peptide Pools which covers the sequence of S, M and N SARS-CoV-2 proteins. Splenocytes are tested at 2 concentrations of 3 different SARS-CoV-2 peptide pools in addition to a negative (medium) and positive control (Phorbol Myristate Acetate/Ionomycin).

5 Example 5: Immunogenicity of polynucleotide vaccine in humans

[0194] To assess the safety, reactogenicity and immunogenicity of CoVEG1 and CoVEG2, an open-label, multi-center, dose-ranging study is conducted in males and non-pregnant females, starting at 18 years of age, inclusive, who are in good health and meet all eligibility criteria. Approximately 45 subjects are enrolled into one of three cohorts (1, 25, and 200 µg). Subjects receive an intradermal injection (100 µl) of CoVEG1 and CoVEG2 on Days 1 and 29 and are followed through 12 months post second vaccination (Day 394). Follow-up visits occur in 1, 2 and 4 weeks post each vaccination (Days 8, 15, 29, 36 and 57), as well as 3, 6- and 12-months post second vaccination (Days 119, 209 and 394).

[0195] The safety and reactogenicity of 2-dose vaccination schedule of CoVEG1 and CoVEG2 administered as intradermal injection, given 28 days apart, across 2 dosages in healthy adults is evaluated based on the percentage of Participants with Adverse Events (AEs), percentage of Participants with Administration (Injection) Site Reactions, and percentage of Participants with Adverse Events of Special Interest (AESIs).

[0196] To evaluate immunogenicity, the following parameters are assessed following a 2-dose vaccination schedule of CoVEG1 and CoVEG2, at Day 15, Day 29 (before the second dose) and at Day 57:

- (a) IgG, IgM and/or IgA ELISA to the S antibody by a validated ELISA method;
- (b) Immune memory indications are assessed by immune phenotyping PBMCs by flow cytometry and validated ELISpot assay focusing on CD49b+T-bet+ resting memory Th cell precursors and CXCR4+ S1P1+ memory plasma cell precursors cells on Day 15, Day 29, Day 57, Day 180, Day 270 and/or Day 394. The measurements include: Geometric mean fold rise (GMFR) in IgG, IgM and/or IgA titer from baseline (Day 1 to Day 394); Geometric mean titer (GMT) of antibody (Day 15, Day 29, Day 57, Day 180, Day 270 and Day 394); and Percentage of subjects who seroconverted (Day 1 to Days 15, 29, 57, 180, 270 and 394). Seroconversion is defined as 4-fold change in antibody titer from baseline; and

(c) IFN- γ response as a measure of CD8 T cell response and phenotype of memory immune cells will be measured in PBMC isolated from subjects by a validated ELISpot assays. PBMC will be isolated on Days 15, 29, 57, 180, 270 and 394.

Example 6: Design of further polynucleotides for expression of SARS-CoV-2 S, E, M, and N proteins and mutants with and without the enhancer L protein

[0197] Plasmids CoVEG 3-17 comprise expression cassettes encoding different viral proteins in the order indicated in FIG. 6. The plasmid backbone (based on the design principles of the pVax1 plasmid) and insert for the plasmids were generated using gene synthesis. The plasmid backbone consists of a Kanamycin resistance gene, the ColE1 origin of replication, the Human cytomegalovirus immediate-early promoter and Simian virus (SV40) Poly A signal. Polynucleotides encoding viral proteins were cloned in between the CMV promoter and the SV40 PolyA signal. After gene synthesis and plasmid preparation, the plasmid was transformed into *E. coli* for cloning and then screened using kanamycin. A representative colony was selected, and its plasmid sequence verified and used as source plasmid for further development. After transcription, the viral proteins were expressed from a single polycistronic mRNA.

Example 7: Expression of SARS-CoV-2 S and N proteins observed by immunofluorescence

[0198] HEK293T cells were seeded at 40,000 cells / well in a 24 well plate 24h prior to transfection. Cells were transfected with the pCoVEG 3-20 plasmids using PEI complexes following manufacturers description. Media was changed 12 hours after transfection and cells were incubated at 37°C, 5% CO₂ for 48 h. Cell media was removed, and cells were fixed with 10% neutral buffered formalin for 10minutesfollowing permeabilization with 0.2 % Triton X-100 in PBS for 10 min. Unspecific binding was blocked by adding EZ block (SCYTEK) before immunostaining was performed. Stain with an anti-spike (S) protein antibody that binds to the receptor binding domain (RBD) – also referred to herein as “anti-RBD” - was added at a dilution of 1:500 and incubated for 1 hour at room temperature. The stain was removed, cells were washed and secondary antibody (Alexa Fluor 568 fluorescently labelled secondary anti-Rabbit, 1:1000 dilution) was added. The stain was incubated for 1 hour at room temperature before removal of the stain and washing. Cells were imaged using a EVOS cell imaging system. FIG. 7 shows the expression of the S protein in HEK293 cells, demonstrating that all tested CoVEG plasmids were capable of expressing the spike protein.

[0199] Additionally, to check for the expression of VLPs, cells were analyzed for expression of the nucleocapsid (N) protein using immunofluorescence staining. For this experiment, HEK293T cells were seeded at 40,000 cells / well in a 24 well plate 24 hours prior to transfection. Cells were transfected with the pCoVEG 5, 9-12, and 14-20 plasmids using PEI complexes following the manufacturers description. The media was changed 12 hours after transfection and cells were incubated at 37°C, 5% CO₂ for 48 hours. The cell media was removed, and cells were fixed with 10% neutral buffered formalin for 10 minutes following permeabilization with 0.2% Triton X-100 in PBS for 10 minutes. Unspecific binding was blocked by adding EZ block (SCYTEK) before immunostaining. Stain with anti-nucleocapsid (N) protein antibody was added at a dilution of 1:1000 and incubated for 1 hour at room temperature. The stain was removed, cells were washed and secondary antibody (Alexa Fluor 488 fluorescently labelled secondary anti-mouse, 1:1000 dilution) was added. The secondary antibody was incubated for 1 hour at room temperature before removal of the stain and washing. Cells were imaged using a EVOS cell imaging system. FIG. 17 shows the expression of the N protein in HEK293 cells, demonstrating that all tested CoVEG plasmids were capable of expressing the nucleocapsid protein.

Example 8: L protein required for detectable SARS-CoV-2 VLP formation

[0200] To isolate intact viral-like particles (VLPs), 4 x 10⁶ HEK293 cells were transfected with the pCoVEG 3-20 plasmids in a 150 mm dish using PEI complexes following manufacturers description. The media was changed 12 hours after transfection and cells were incubated at 37 °C, 5% CO₂ for 72 hours. VLP containing supernatants were harvested, spun down (1,500 x g, 15 min) and concentrated using an Amicon centrifugal filter unit (100 kDa cut off). Concentrate was spun down (4,500 x g, 15 minutes) to remove precipitate and VLPs were pelleted (100,000 x g, 1.5 hours) through a 20 % sucrose cushion. VLPs were resuspended in PBS and analyzed by western blot, as shown in FIG. 8. FIG. 8 shows that CoVEG 3-8 plasmids were capable of expressing the S protein and CoVEG 3 and 5-8 plasmids were capable of expressing the N protein. These data demonstrate that the DNA vectors CoVEG 3, and 5-8 disclosed herein express SARS-CoV-2 viral structural proteins necessary for the formation of VLPs in HEK 293 cells and that secreted VLPs components can be detected in cell culture supernatants.

[0201] FIG. 11 shows the expression of S protein and N protein from CoVEG 5, 9, 11, 16, 10, and 15 plasmids. The results showed that expressing the mutant S protein (from CoVEG 9, 11, 10, and 15) increased the amount of Spike protein expressed and presented on VLPs.

Expression of ORF3 protein (from CoVEG 16) appeared to decrease the amount of S and N proteins in the VLPs. Absence of the enhancer L protein upon expression of the CoVEG 15 plasmid resulted in a similar amount of S protein, but a far greater amount of N protein. Without being bound by a theory, it is thought that in the absence of the enhancer L protein, a higher amount of N protein is expressed resulting in unbalanced VLP formation.

[0202] FIG. 12 shows the expression of S protein and N protein from CoVEG 5, 12, 14, 13, 10, 9 and 11 plasmids. These data demonstrate that expressing the mutant S protein (from CoVEG 9, 10, and 11 plasmids) increased the amount of Spike protein expressed and presented on VLPs, as compared to the comparator plasmids that expressed wild type S protein, while the amount of N protein appeared constant across plasmids.

[0203] FIG 18 shows the expression of S protein and N protein from CoVEG 5, 8, 9, 10, 15, 16, 17 and 20 plasmids. Notably, the presence of the enhancer L protein resulted in different S and N protein expression ratios, *e.g.* as shown by the CoVEG 10 (L protein) versus the CoVEG 20 (no L protein) S and N western blotting in FIG. 18.

[0204] The presence of secreted VLPs were also confirmed by ELISA. For this experiment, 24,000 HEK293 cells were transfected with pCoVEG 5 and 9-14 plasmids or plasmids containing either the Spike protein or the Nucleocapsid protein as controls. Experiments were performed in a 24 well plate using PEI complexes following the manufacturer's descriptions. The media was changed 12 hours after transfection and cells were incubated at 37°C and 5% CO₂ for 72 hours. VLP containing supernatants were harvested, spun down (1,500 x g, 15 min) and 75 µl of the cleared supernatant was used to coat ELISA plates over night at 4°C. After incubation, the plates were washed twice with 0.05% Tween-20 in PBS and wells were blocked using EZ block (2 hours at 37°C). The plates were washed twice with 0.05 % Tween-20 in PBS. To detect VLPs in the coating material, anti-RBD (Sino Biological, mouse anti-RBD SARS-CoV-2 (2019-nCoV) Spike Neutralizing Antibody, Mouse Mab, 40592-MM57 SARS-CoV-2, 1:500 dilution in EZ Block) or anti-N (Novus Biologicals, Mouse anti-SARS-CoV-2 Nucleocapsid, Clone: B3449M, N2787B09, 1:1000 dilution in EZ Block) were added to the wells. Antibodies were incubated for 1 hour at room temperature before washing three times with 0.05% Tween-20 in PBS and adding 75 µl of secondary antibody (Goat-Anti-mouse, HRP-conjugate, 1:2,000 dilution, Southern Biotech, Goat anti-Mouse IgG(H+L), horseradish peroxidase (HRP) , Polyclonal, OB103405) and incubating for 1 hour at room temperature. Wells were thoroughly washed (5x with 0.05% Tween-20 in PBS), and binding was developed using 75 µl 3,3',5,5'-tetramethylbenzidine (TMB) substrate (Surmodisc Inc TMB One

Component HRP Microwell Substrate). The reaction was carried out for 30 minutes with 75 µl Stop Solution (Surmodisc Inc 450 NMLIQ STOP REAGENT) and Absorbance was measured at 450 nm.

[0205] FIG 15 shows ELISA results from VLP secretion of CoVEG 5 and 9-14 plasmids compared with single protein S and N expressing vectors. Both Spike and Nucleocapsid proteins secreted from HEK293 cells. However, while the S protein demonstrated high ELISA VLP signal relative to single protein expression, the N protein demonstrated a notably lower VLP signal relative to single protein expression. It may be that N signal in VLPs is lower than the S signal in VLPs because the N protein is on the interior of the VLP and not accessible to the antibody.

[0206] To further ensure that the expressed structural proteins from SARS-CoV2 were forming intact VLP, 20×10^6 HEK293 cells were transfected with the pCoVEG 10 plasmid in 5 -150 mm dishes using PEI complexes following manufacturers description. The media was changed 12 hours after transfection and cells were incubated at 37°C, 5% CO₂ for 72 hours. VLP containing supernatants were harvested, spun down (1,500 x g, 15 minutes) and concentrated using an Amicon centrifugal filter unit (100 kDa cut off). Concentrate was spun down (4,500 x g, 15min) to remove precipitate and VLPs were pelleted (100,000 x g, 1.5 h) through a 20% sucrose cushion. VLPs were resuspended in PBS and used for co-Immunoprecipitation (co-IP). For the co-IP resuspended VLPs were incubated with anti-S RBD antibody (Sino Biological, mouse anti-RBD SARS-CoV-2 (2019-nCoV) Spike Neutralizing Antibody, Mouse Mab, 40592-MM57 SARS-CoV-2) for 60 minutes before adding 100 µl washed Protein A/G agarose resin (Thermo Fisher scientific). Resin was incubated for 120 minutes before eluting with 0.1M glycine pH 2. Eluates were immediately neutralized by adding 5 times volume of 1M Tris pH 8.0. Fractions were analyzed for the presence of N protein by western blot using anti-N antibody (Novus Biologicals, Mouse anti-SARS-CoV-2 Nucleocapsid, Clone: B3449M, N2787B09).

[0207] FIG. 19 shows the western blot result of the co-IP experiments of the CoVEG 10 plasmid. The N-protein is packed in intact VLPs as demonstrated by the presence of an N-signal in the elution fraction after incubation with RBD (left side, arrow). This signal, compared with the absence of signal in the washing fraction, indicates the N protein is retained within the particles. As a control for the possibility of non-specific N protein binding to the outside of particles, the co-IP was run in parallel without the anti-RBD antibody (right side). The N

protein signal was not detectable in the elution fraction, demonstrating that the N protein did not bind the resin non-specifically.

[0208] To visualize the secreted VLPs, 20×10^6 HEK293 cells were transfected with the pCoVEG 10 and 20 plasmids in 5 x 150 mm dishes using PEI complexes and following the manufacturer's description. The media was changed 12 hours after transfection and the cells were incubated at 37°C, 5% CO₂ for 72 hours. VLP containing supernatants were harvested, spun down (1,500 x g, 15 minutes) and concentrated using an Amicon 100 kDa centrifugal filter unit. The concentrate was spun down (4,500 x g, 15 minutes) to remove precipitate and VLPs were pelleted (100,000 x g, 1.5 hours) through a 20% sucrose cushion. VLPs were resuspended in PBS, flash frozen, and stored at -80°C until used for transmission electron microscopy (TEM).

[0209] For the TEM experiments, VLPs were ultracentrifuged for 2 hours at 25000 g on a 20% sucrose cushion using a TLS-55 (Optima TLX Ultracentrifuge, Beckman). 10 µl were put on a microscopy copper grid (Sigma Aldrich) and fixed with 2% (v/v) paraformaldehyde for 5 minutes. Samples were then negatively stained with 5 mL of phosphotungstic acid (Sigma Aldrich). The grid was examined with a Hitachi HT7700 TEM operating at 100 KeV.

[0210] FIG. 16 shows the presence of VLPs in the isolated material for CoVEG 10. The presence of a larger particle with a clear Spike trimerization surface could be observed for CoVEG10 (see zoom inset). CoVEG 20 which is identical to CoVEG10 apart from the presence of the L regulatory protein, failed to generate recognizable VLPs.

[0211] Intact and immunogenic VLPs are highly dependent on the ratio of all VLP forming proteins. Herein it was demonstrated that the L protein controlled expression of all VLP forming proteins and the correct formation of the VLPs.

25 **Example 9: L protein increased neutralizing antibodies in vivo to SARS -CoV-2 proteins and was required for Th1 response**

[0212] To determine the immunogenicity of plasmids CoVEG 3-8, the plasmids were diluted to 1 mg/ml in PBS and 50 µl was injected intramuscularly into 6 week old BALB/c mice in 2 week intervals, for a total of 2 injections at day 1 and 15. Blood was collected on days 14, day 28, day 42 and day 56, and the serum was isolated and snap frozen in the presence of an anti-coagulant.

[0213] To determine the immunogenicity of the plasmids CoVEG 5, 8 and 9-14 as well as the S-only plasmid, the plasmids were diluted to 2 mg/ml or 0.5 mg/ml in PBS and 50 µl was injected intramuscularly (2 mg/ml) or intradermally (0.5 mg/ml) into 6 week old BALB/c mice in 2 week intervals, for a total of 2 injections at day 1 and 15. Blood was collected on day 14, day 28, day 42 and day 56 and the serum was isolated and snap frozen in the presence of an anti-coagulant.

[0214] To determine immunogenicity of plasmids CoVEG 9, 10 and 20, as well as the S only plasmid with and without the enhancer protein, the plasmids were diluted to 2 mg/ml or 0.2 mg/ml in PBS and 50 µl was injected intramuscularly into 6 week old C57BL/6 mice in 2 week intervals, for a total of 1-3 injections at day 1, 15 and 29. Blood was collected on day 0, day 14, day 28, day 42, day 56 and day 70 and the serum was isolated and snap frozen in the presence of an anti-coagulant. To measure the binding antibody concentration, *i.e.*, the elicited humoral immune response, enzyme linked immunosorbent assays (ELISA) were performed using purified SARS-CoV-2 Spike RBD protein (Creative Diagnostics® DAGC149 Recombinant SARS-CoV-2 Spike Protein Receptor Binding Domain [His]) as a coating material. For this experiment, high-binding 96-well plates were coated with 75 µl of a 2 µg/ml SARS-CoV-2-Spike RBD solution, and plates were incubated over-night at 4°C. After incubation, plates were washed twice with 0.05% Tween-20 in PBS, and wells were blocked using EZ block (2 hours at 37°C). Plates were washed twice with 0.05% Tween-20 in PBS. Serum was collected from the mice after 56 days and added to the wells (1:500 dilution for binding antibody detection, 1:100 – 1:7812500 for Endpoint Titer measurement). Serum was incubated for 1 hour at room temperature before washing thrice with 0.05% Tween-20 in PBS and adding 75 µl of secondary antibody (Goat-Anti-rabbit, HRP-conjugate, 1:4,000 dilution) and incubating for 1 hour at room temperature. Wells were thoroughly washed (5x with 0.05% Tween-20 in PBS) and binding was developed using 75 µl 3,3',5,5'-tetramethylbenzidine (TMB) substrate (Surmodisc Inc TMB One Component HRP Microwell Substrate). The reaction was carried out for 30 minutes before stopping with 75 µl Stop Solution (Surmodisc Inc 450 NM LIQ STOP REAGENT). The Absorbance was measured at 450 nm.

[0215] FIG. 9A shows the total binding antibody measured using ELISA, and FIG. 9B shows the measured endpoint titers. These results demonstrate that CoVEG 3-8 plasmids are capable of eliciting a strong immune response when injected into mice and thus, producing anti-SARS CoV2 antibodies. CoVEG 8 was identified as being particularly superior in being able to induce

a strong immune response in these mice. These results demonstrate that the plasmids disclosed herein are highly effective DNA vaccines against SARS-CoV-2.

[0216] FIG. 13 shows ELISA results from injection of CoVEG 5, 9, 11, 10, 12, 13, 8 and 14 plasmids, either intramuscularly (IM) or intradermally (ID). Remarkably, the intramuscular injection of CoVEG10 induced a much higher immune response as compared to the injection of the Spike protein alone. Intramuscular injections of CoVEG5, CoVEG9, and CoVEG11 also induced a higher immune response as compared to the injection of the S protein alone. These results further demonstrate that the plasmids disclosed herein are highly effective DNA vaccines against SARS CoV2 that perform better than vaccines that express the Spike protein alone.

[0217] FIGS. 20 and 22 show antibody binding titers from CoVEG 5 and 9-14, 18, 19, and 20 plasmids, as well as S only, on day 42 after IM or ID injection. Interestingly, all of the tested CoVEGs were capable of inducing an immune response, with CoVEG 9 the most efficacious. Notably, the VLP forming CoVEG 9 performed better than the spike protein alone.

[0218] Additionally, the cellular immune response was measured using the presence of antigen reactive T cells using IFN- γ and IL-4 enzyme-linked immune absorbent Spot (ELISpot) assays. For ELISpot analysis, spleen was collected and T cells were isolated. ELISpot assessment was performed by priming the T cells with Miltenyi Biotec Peptivator SARS-CoV-2 peptide pools to activate SARS-CoV-2 reactive T cells. For the analysis, Mouse IL-4 Single color ELISPOT and Mouse INF- γ Single color ELISPOT (Immunospot, Cellular Technology Limited) were used according to manufacturer's instructions. In short, 96 well PVDF membrane plates were coated with IL-4 or INF- γ capture antibody and incubated over night at 4°C. After washing, 150,000 splenocytes in 100 μ l CTL test medium, seeded on pre-coated plates, and incubated for 15 minutes at 37 °C and 4% CO₂. Cells were activated with either PMA/Ionophore (as a positive control) or 0.6 μ l of reconstituted Miltenyi Biotec Peptivator SARS-CoV-2 peptide pools (S, N or M) per well. Reactions were incubated for 24 hours before developing and counting using an ImmunoSpot analyzer (CTL).

[0219] FIG. 24 shows the result of the T-cell analysis of CoVEG10 and CoVEG20. Importantly, to generate a successful vaccine against respiratory viruses *e.g.*, SARS-CoV-2, a Th1 preference (INF- γ) or at least a balanced Th1/Th2 respond is needed. Surprisingly, the addition of the L regulatory protein in CoVEG10 influenced T-cellular immune response in favor of the INF- γ (Th1) response. Notably, the absence of the regulatory protein in CoVEG20

reversed the cellular response to IL4 (Th2). Th1 responses are necessary to generate a long-lasting immune response to a virus.

[0220] To test whether the serum has neutralizing antibodies against SARS CoV2 that bind to the Spike protein, *in vitro* viral neutralization assays using the cPass™ neutralization assay (GenScript) were performed according to manufacturer's instructions. The cPass™ allows the detection of total neutralizing antibodies in a sample by mimicking the interaction between the virus and the host cell *in vitro*. In the assay, if neutralizing antibodies are present in the sample being tested, then the binding of the receptor binding domain (RBD) to host cell membrane receptor, ACE2 is inhibited. However, if neutralizing antibodies are absent in the sample, then the RBD is able to bind to ACE2. FIG 10 shows the percent (%) inhibition of RBD binding to the ACE2 receptor. If the inhibition of RBD binding to ACE2 is more than 30% (red dotted line), then the sample is identified as having neutralizing antibodies. As shown in FIG. 10, serum samples obtained from mice injected with COVEG5 and COVEG8 show a high percentage of inhibition, and therefore, generated neutralizing antibodies. These results further underline that the disclosed plasmids are highly effective DNA vaccines against SARS-CoV-2.

[0221] FIG 21 shows the neutralizing antibodies of the samples shown in FIG. 20. These data indicate that not all generated antibodies possess neutralizing capacity. A lower signal in this assay confirmed neutralization as defined by a signal less than 70% of the negative control (dashed line). Signals lower than 30% of the negative control were confirmed to be strongly neutralizing. The tested CoVEGs performed differently depending on the design and the injection site. Surprisingly, ID injection did not induce strong neutralizing antibodies, whereas IM injections did induce strong neutralizing antibodies. Additionally, CoVEG13 and 14 did not meet the criteria for neutralization. However, CoVEG9 showed the highest neutralization of all tested constructs including the S spike protein only constructs, again demonstrating that the VLP approach provides a more potent vaccine than the sub-unit S spike only vaccines.

[0222] FIG. 23 shows neutralizing antibodies of CoVEG 9, 10 and 20 plasmids, in which plasmid 20 is the control for the absence of the enhancer L protein. Although CoVEG20 showed neutralizing potential, Th2 overhang as demonstrated in T-cell analysis, it is not a viable option for a vaccine.

[0223] Further, the neutralization capacity with and without the enhancer protein over time was analyzed by cPass™. SARS-CoV-2 Surrogate Virus Neutralization Test (sVNT) Kit. For this, serum samples from immunized animals (immunized with CoVEG9, Spike + enhancer

protein L and Spike without enhancer) were collected on day 42 and day 70 and cPassTM was performed as described above.

[0224] FIG.27 A shows the individual values of the analyzed serum samples and FIG. 27 B shows the median of the same data as summary. Interestingly, the addition of the enhancer protein clearly showed a benefit of neutralizing antibodies over the time. Firstly, both constructs with the enhancer protein L (CoVEG9 and Spike + L) showed higher median neutralization values (FIG. 27B B, CoVEG9 circles, Spike + L squares). Secondly, the level of neutralizing antibodies remained high, even after 70 days post injection compared to the construct without the enhancer protein (FIG. 27B Spike triangles).

[0225] This further demonstrated the advantage of the addition of the enhancer protein to the vaccine candidates.

Example 10: L protein increased functional West Nile Virus (WNV) VLP production when co-expressed with WNV E and M proteins

[0226] A plasmid encoding the precursor membrane protein (prM), the envelope glycoprotein (E) of NY99 strain of WNV and an enhancer protein was constructed as described in Example 6 (see FIG. 14A, SEQ ID NO: 55). Also, a control plasmid encoding just the precursor membrane protein (prM), and the envelope glycoprotein (E) of NY99 strain of WNV was constructed (FIG. 11B). HEK293T cells were cultured in DMEM supplemented with 10% FBS at 37°C at 5% CO₂. On Day 1, the cells were seeded on 24-well plates at 20,000 cells per well and grown overnight. On Day 2, cells seeded to the 24-well plates were transiently transfected with a plasmid encoding the precursor membrane protein (prM), the envelope glycoprotein (E) of NY99 strain of WNV and an enhancer protein. A control plasmid was used in all experiments, which encodes just the precursor membrane protein (prM) and the envelope glycoprotein (E) of NY99 strain of WNV, and not the enhancer protein.

[0227] Each well of a 24-well plate was transfected using plasmid/PEI complexes, which were formed using 0.5 ug of the corresponding plasmid and 1 ug of PEI in 50 µl Opti-MEM. The complexes were formed by incubating plasmid/PEI mixture at room temperature for 30 min. Cell medium in 24-well plates was replaced by fresh Opti-MEM and complexes were added to the wells. On Day 3, the complexes were removed from transfected cells and replaced with fresh Opti-MEM.

[0228] On Day 4, cell culture supernatants were collected, removed from cell debris by centrifugation at 500xg for 5 minutes and saved for downstream analysis by ELISA. Cells were

fixed using 250 µl of 10% neutral buffered formalin (10 minutes at room temperature), and permeabilized using 0.2% Triton-X 100 (10 minutes at room temperature) and washed.

[0229] Fluorescence microscopy was used to visualize protein expression in cells as followed. Cells were stained using mouse anti-WNV_E and rabbit anti-WNV_M primary antibodies
5 (1:500 dilution in PBS, 1 h at room temperature), washed, developed with goat anti-mouse Alexa Fluor 488 secondary antibodies (1:1000 dilution in PBS, 1 h at room temperature), washed, and imaged using fluorescence microscopy.

[0230] The results of the immunostaining experiments are shown in FIG. 28. As observed in other experiments described herein, the quantity of the WNV + enhancer protein (EG) was
10 lower compared to the quantity of the WNV without it. However, the quality seemed to be higher when the enhancer protein was present as observed by the formation of nuclei (FIG. 28, right, as indicated by arrows). These data demonstrate the higher quality of proteins expressed with the compositions and methods provided herein.

[0231] ELISA assays were used to demonstrate the secretion of expressed antigens. For this, supernatant from transfected cells were collected on days 4 (48 hours after transfection), 5 (72 hours), 6 (96 hours), 7 (120 hours) and 8 (144 hours). High-binding 96-well plates were coated with the cell culture supernatants using 75 µl of cell culture supernatant per well and incubated at +4°C overnight. The next day, the coated wells were washed using PBST buffer and blocked using 200 µl of EZ Block™ reagent (Scytek Laboratories) per well for 2 h at +37°C. The wells
20 were washed 3 times with PBST and incubated with a primary antibody (mouse anti-WNV_E, diluted 1:1000 in EZ Block, 75 µl per well) for 1 hour at room temperature. The wells were then washed 3 times with PBST and incubated with the goat anti-mouse HRP secondary antibody diluted 1:1000 in EZ Block reagent, 75 µl per well, for 1 hour at room temperature. The wells were then washed 5 times using PBST and 75 µl of TMB substrate was added to
25 each well and incubated 30 minutes at room temperature, followed by the addition of 75 µl of Stop Solution, and absorbance measured at 450 nm using a plate reader. Additionally, to demonstrate that the VLP secretion was not caused by cell death and the unspecific release of intracellular protein, cells were imaged every day and ELISA results were compared to the images.

30 [0232] FIG. 26 A illustrates the secretion of VLPs from transfected cells over time as measured by ELISA. Whereas the WNV construct with the enhancer protein showed the highest secretion 72 hours after transfection as was expected for intact VLPs and healthy cells, the WNV construct without the enhancer protein showed a steady increase in secreted material over time. The latter was more consistent with increased cell death and unspecific release of protein

material that most likely are not fully formed VLPs. The results were confirmed by analysis of cell images taken at the time of harvest from the supernatant. Whereas the WNV with the enhancer protein showed little to no cell death (as indicated by stars) over the time of the experiment, WNV without the enhancer protein showed visible cell death (as indicated by stars), after 72 hours this became increasingly more pronounced over the time of the experiment. The cell images are further proof that the released material from the WNV without the enhancer protein was most likely due to the release of protein from cell death rather from controlled secretion of VLPs.

[0233] This example further demonstrates that the methods and compositions of the disclosure improve the quality of produced antigen.

Example 11: L protein increased total West Nile Virus (WNV) VLP production when co-expressed with WNV- E and M proteins

[0234] To isolate intact VLPs, 4×10^6 HEK293 cells in a 150 mm dish were transfected with a plasmid encoding the precursor membrane protein (prM) and the envelope glycoprotein (E) of NY99 strain of WNV, and an enhancer protein. A control plasmid was used in all experiments, which encodes just the precursor membrane protein (prM) and the envelope glycoprotein (E) of NY99 strain of WNV, and not the enhancer protein. The transfections were conducted using PEI complexes following the manufacturers description using 40 μ g plasmid and 80 μ g PEI per 150 mm dish. Media was changed 12 hours after transfection and cells were incubated at 37°C, 5% CO₂ for 72 hours. VLP containing supernatants were harvested, spun down (1,500 x g, 15 minutes) and concentrated using an Amicon Ultra centrifugal filter unit (100 kDa cut off). Concentrate was spun down (4,500 x g, 15 minutes) to remove precipitate and VLPs were pelleted through a 20% sucrose cushion at 100,000 x g for 1.5 hours. VLPs were resuspended in PBS and analyzed by ELISA.

[0235] ELISAs were performed as described above, and as known in the art, for instance, as described in *Cold Spring Harb Protoc*; doi:10.1101/pdb.prot093708. Briefly, high-binding 96-well plates were coated using VLPs resuspended in PBS in serial dilutions from 1:20 to 1:72,000 to visualize the difference of expression quantity between the constructs with and without the enhancer protein and incubated overnight at 4°C. Plates were washed and blocked with EZ block (Scytex Laboratories) for 2 hours at 37°C. Anti-West Nile Virus Antibody, clone E16, was diluted in EZ block (1 : 5,000) and plates were incubated for 1 hour at RT. Wells were washed and goat anti-mouse HRP labeled detection antibody (Southern Biotech) was

added for detection, followed by washes and the incubation with TMB substrate and the stop solution. Signal was read out as absorbance at 450 nm using a plate reader.

[0236] FIG. 25 demonstrates the difference between the total expression and secretion of West Nile virus constructs with and without the enhancer protein. As shown, the addition of the enhancer protein (circles) led to a higher expression of West Nile virus particles compared to the construct lacking the enhancer protein (squares). This demonstrates that the compositions and methods of the disclosure are beneficial for WNV VLP vaccine production.

Example 12: Immunogenicity of West Nile Virus proteins co-expressed with the L enhancer protein in mice

[0237] The ability to evoke immune responses *in vivo* upon vaccination with a plasmid encoding the precursor membrane (prM), the envelope glycoprotein (E) of WNV, and the enhancer protein, is evaluated using BALB/c mice as follows. 6-week-old female BALB/c mice are randomized into groups based on body weight. Mice are dosed with the plasmids using intradermal or intramuscular injections on Day 1 and Day 21. Mouse serum samples are collected on Day 1 (pre-vaccination), on Day 21 (prior to boost) and on 42. On day 42, mice are sacrificed and splenocytes are isolated.

[0238] The elicited humoral immune response is measured by evaluating the titer of anti-M and anti-E antibodies a respective enzyme linked immunosorbent assay (ELISA). Additionally, cellular immune response is measured by evaluating the presence of antigen reactive T cells using a respective IFN- γ and IL-4 enzyme-linked immune absorbent Spot (Dual color ELISpot) assay.

ELISAs are performed as described here, and as known in the art, for instance, as described in *Cold Spring Harb Protoc*; doi:10.1101/pdb.prot093708. Briefly, high-binding 96-well plates are coated using recombinant prM and E proteins (Abcam) at 2 μ g/ml concentration and blocked. Serum samples are serially diluted in EZ Block reagent and added to pre-coated wells, washed and detected using goat anti-mouse HRP labeled detection antibody, followed by washes and the incubation with TMB substrate and the stop solution. Endpoint titer is defined as the reciprocal maximal antibody dilution at which the ELISA signal (absorbance at 450 nm) is above 3 standard deviations of background signal.

[0239] Dual color ELISpot assay is conducted as described here, and as known in the art, for instance, as described in *Cold Spring Harb Protoc* 2010 doi:10.1101/pdb.prot5369. Briefly, splenocytes are isolated on Day 42, stimulated with respective prM or E peptide arrays

(Biodefense and Emerging Infections Research Resources Repository) and added to the pre-prepared ELISpot microplates. Negative (medium) and positive controls (Phorbol Myristate Acetate/Ionomycin) are included in the assay. The number of antigen-reactive IFN-gamma and IL-4 secreting T cells are counted using an ELISpot reader.

- 5 [0240] Finally, the presence of WNV neutralizing antibodies in mouse sera isolated at different time points (see above) is evaluated as described herein, and as described in *The Journal of Infectious Diseases*, Volume 196, Issue 12, 15 December 2007, Pages 1732–1740, and *Virology*, Volume 346, Issue 1, 1 March 2006, Pages 53-65. Briefly, WNV reporter-virus particles (RVPs) are generated in HEK293T cells by transiently transfecting WNV prM and E
- 10 proteins (to form virus-like particles also known as subviral particles), complemented with transiently transfected reporter-replicon (luciferase) and transiently transfected capsid protein. Isolated RVPs are incubated with mouse serum samples at different serial dilutions and added to pre-plated PHK-21 cells and incubated for 2 days, after which the reporter gene activity is measured using a microplate reader. The reduction in the reporter gene activity reflects the
- 15 level of WNV neutralizing antibodies in mouse sera.

Example 13: Expression and immunogenicity of additional polynucleotide constructs

- [0241] Construction of plasmids encoding viral proteins derived from other viruses, *e.g.*, Influenza viral proteins (*e.g.*, HA, NA, M1, M2, or any combination thereof), Hepatitis B viral proteins (*e.g.*, sAg (S protein), sAg (M protein), sAg (L protein), preS1, preS2, cAg (core
- 20 antigen), or any combination thereof), Human Papillomavirus (*e.g.*, L1 protein of HPV 6, L1 protein of HPV 11, L1 protein of HPV 16, L1 protein of HPV 18, or any combination thereof) is performed using the methods described in Example 6. Expression of these proteins in different combinations in HEK293T cells and isolation of the VLPs is performed using methods described in Examples 7, 8, 10 and 11. Finally, the immunogenicity of the plasmids
- 25 encoding these proteins is tested using the methods described in Example 9 and 12.

INCORPORATION BY REFERENCE

- [0242] All references, articles, publications, patents, patent publications, and patent applications cited herein are incorporated by reference in their entireties for all purposes. However, mention of any reference, article, publication, patent, patent publication, and patent
- 30 application cited herein is not, and should not be taken as an acknowledgment or any form of

suggestion that they constitute valid prior art or form part of the common general knowledge in any country in the world.

NUMBERED EMBODIMENTS

[0243] Embodiment 1. A vector for use as a vaccine, comprising an expression cassette comprising a polynucleotide encoding a viral protein and a polynucleotide encoding an enhancer protein, wherein the enhancer protein is a picornavirus leader (L) protein or a functional variant thereof.

[0244] Embodiment 2. The vector of embodiment 1, wherein the amino acid sequence of the enhancer protein has at least 95% identity to SEQ ID NO: 1, or at least 95% identity to SEQ ID NO: 2.

[0245] Embodiment 3. The vector of embodiment 1 or embodiment 2, wherein the amino acid sequence of the enhancer protein is SEQ ID NO: 1, or SEQ ID NO: 2.

[0246] Embodiment 4. The vector of any one of embodiments 1-3, wherein the polynucleotide encoding the enhancer protein is operatively linked to a polynucleotide encoding an internal ribosome entry site (IRES).

[0247] Embodiment 5. The vector of embodiment 4, wherein the polynucleotide encoding the IRES is SEQ ID NO: 24.

[0248] Embodiment 6. The vector of any one of embodiments 1-5, wherein the viral protein is a viral antigen.

[0249] Embodiment 6.1 The vector of any one of embodiments 1-6, wherein the viral protein is derived from a virus selected from the group consisting of coronavirus, influenza virus, Hepatitis B virus, Human Papilloma virus (HPV), West Nile virus, and Human Immunodeficiency Virus (HIV) virus.

[0250] Embodiment 6.2 The vector of embodiment 6.1, wherein the viral protein is derived from a coronavirus.

[0251] Embodiment 7. The vector of any one of embodiments 1-6.2, wherein the coronavirus is a betacoronavirus.

[0252] Embodiment 8. The vector of embodiment 7, wherein the betacoronavirus is severe acute respiratory syndrome (SARS) virus.

[0253] Embodiment 9. The vector of embodiment 8, wherein the SARS virus is a SARS-CoV-2 virus.

[0254] Embodiment 10. The vector of embodiment 7, wherein the betacoronavirus is Middle East respiratory syndrome (MERS) virus.

[0255] Embodiment 11. The vector of any one of embodiments 1-10, wherein the coronavirus protein is a coronavirus spike protein.

- [0256] Embodiment 12. The vector of embodiment 11, wherein the spike protein shares at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 13.
- 5 [0257] Embodiment 13. The vector of embodiment 12, wherein the spike protein is SEQ ID NO: 13.
- [0258] Embodiment 13.1 The vector of embodiment 11, wherein the spike protein is a mutant spike protein.
- [0259] Embodiment 13.2 The vector of embodiment 13.1, wherein the mutant spike protein
10 comprises the amino acid substitutions, R682G, R683S, R685S, K986P, and V987P, in SEQ ID NO: 13.
- [0260] Embodiment 13.3 The vector of embodiment 13.1, wherein the mutant spike protein comprises an amino acid sequence of SEQ ID NO: 51.
- [0261] Embodiment 14. The vector of any one of embodiments 1-13.3, wherein the
15 coronavirus protein is a coronavirus membrane (M) protein.
- [0262] Embodiment 15. The vector of embodiment 14, wherein the M protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 33.
- [0263] Embodiment 16. The vector of embodiment 14 or embodiment 15, wherein the M
20 protein is SEQ ID NO: 33.
- [0264] Embodiment 17. The vector of any one of embodiments 1-16, wherein the coronavirus protein is a coronavirus envelope (E) protein.
- [0265] Embodiment 18. The vector of embodiment 17, wherein the E protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least
25 99%, at least 99.5% or 100% identity to SEQ ID NO: 22.
- [0266] Embodiment 19. The vector of embodiment 17 or embodiment 18, wherein the E protein is SEQ ID NO: 22.
- [0267] Embodiment 20. The vector of any one of embodiments 1-19, wherein the coronavirus protein is a coronavirus nucleocapsid (N) protein.
- 30 [0268] Embodiment 21. The vector of embodiment 20, wherein the N protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 20.
- [0269] Embodiment 22. The vector of embodiment 20 or embodiment 21, wherein the N protein is SEQ ID NO: 20.

- [0270] Embodiment 23. The vector of any one of embodiments 1-22, wherein the coronavirus protein forms a virus-like particle (VLP).
- [0271] Embodiment 23.1 The vector of embodiment 6.1, wherein the viral protein is derived from West Nile virus.
- 5 [0272] Embodiment 23.2 The vector of embodiment 23.1, wherein the viral protein is precursor membrane protein (preM), envelope glycoprotein (E), or a combination thereof.
- [0273] Embodiment 24. A vector for use as a vaccine, comprising an expression cassette comprising a polynucleotide, wherein the polynucleotide comprises a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%,
- 10 at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to the nucleic acid sequence of SEQ ID NO: 30.
- [0274] Embodiment 25. The vector of embodiment 24, wherein the polynucleotide comprises the nucleic acid sequence of SEQ ID NO: 30.
- [0275] Embodiment 26. A vector for use as a vaccine, comprising an expression cassette
- 15 comprising a polynucleotide, wherein the polynucleotide comprises a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to the nucleic acid sequence of SEQ ID NO: 31.
- [0276] Embodiment 27. The vector of embodiment 26, wherein the polynucleotide comprises
- 20 the nucleic acid sequence of SEQ ID NO: 31.
- [0277] Embodiment 27.1 A vector for use as a vaccine, comprising a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO: 35-49, and 55.
- 25 [0278] Embodiment 28. The vector of any one of embodiments 1-27.1, wherein the vector is a naked polynucleotide.
- [0279] Embodiment 29. The vector of any one of embodiments 1-28, wherein the vector is a deoxyribonucleic acid (DNA) polynucleotide.
- [0280] Embodiment 30. The vector of any one of embodiments 1-28, wherein the vector is a
- 30 ribonucleic acid (RNA) polynucleotide.
- [0281] Embodiment 31. The vector of any one of embodiments 1-30, wherein the vector comprises a plasmid.
- [0282] Embodiment 32. The vector of any one of embodiments 1-30, wherein the vector comprises linear DNA.

[0283] Embodiment 33. The vector of any one of embodiments 1-32, wherein the expression cassette comprises a promoter operatively linked to each of the polynucleotide sequences of the expression cassette.

5 [0284] Embodiment 33.1 The vector of any one of embodiments 1-33, wherein the vector comprises a DNA polynucleotide, said DNA polynucleotide encoding a viral packaging signal.

[0285] Embodiment 33.2 The vector of embodiment 33.1, wherein the viral packaging signal is a RNA polynucleotide.

[0286] Embodiment 33.3 The vector of embodiment 33.2, wherein the viral packaging signal is derived from a coronavirus.

10 [0287] Embodiment 34. A vaccine composition, comprising the vector of any one of embodiments 1 to 33.4 and a pharmaceutically acceptable carrier.

[0288] Embodiment 35. The vaccine composition of embodiment 34, wherein the vaccine composition comprises an adjuvant.

15 [0289] Embodiment 36. The vaccine composition of embodiment 35, wherein the adjuvant is alum.

[0290] Embodiment 37. The vaccine composition of embodiment 35, wherein the adjuvant is monophosphoryl lipid A (MPL).

[0291] Embodiment 38. A method of expressing a viral antigen in a eukaryotic cell, comprising contacting the cell with the vector of any one of embodiments 1 to 33.4.

20 [0292] Embodiment 39. The method of embodiment 38, wherein contacting the cell with the vector results in: (i) expression of the antigen at a higher expression level; and/or (ii) expression of the antigen for a longer period of time; and/or (iii) expression of the antigen with better protein quality, than a vector lacking the enhancer protein.

[0293] Embodiment 40. The method of embodiment 38 or embodiment 39, wherein contacting
25 the cell with the vector results in: (i) expression of a virus like particle (VLP) comprising the antigen at a higher expression level; and/or (ii) expression of a VLP comprising the antigen for a longer period of time; and/or (iii) expression of a VLP comprising the antigen with better protein quality, than a vector lacking the enhancer protein.

[0294] Embodiment 40.1 The method of embodiment 40, wherein the vector comprises a DNA
30 polynucleotide encoding a viral packaging signal, wherein contacting the cell with the vector results in expression of the viral packaging signal, and wherein the VLPs encapsidate the viral packaging signal.

[0295] Embodiment 40.2 The method of embodiment 40.1, wherein the vector results in the formation of a greater number of VLPs, as compared to a control vector lacking the DNA polynucleotide encoding the viral packaging signal.

5 [0296] Embodiment 41. A method of eliciting an immune response in a subject, comprising administering an effective amount of the vaccine composition of any one of embodiments 34 to 37 to the subject.

[0297] Embodiment 42. The method of embodiment 41, wherein tissue at an administration site of the subject expresses the antigen and/or a VLP comprising the antigen.

10 [0298] Embodiment 43. The method of embodiment 42, wherein tissue at an administration site of the subject: (i) expresses the antigen and/or a VLP comprising the antigen at a higher expression level; and/or (ii) expresses the antigen and/or a VLP comprising the antigen for a longer period of time; and/or (iii) expresses the antigen and/or a VLP comprising the antigen with better protein quality, than when a vector lacking the enhancer protein is administered.

15 [0299] Embodiment 43.1 The method of any one of embodiments 41-43, wherein the vector comprises a DNA polynucleotide encoding a viral packaging signal, wherein tissue at an administration site of the subject expresses the viral packaging signal, and wherein the VLPs encapsidate the viral packaging signal.

20 [0300] Embodiment 43.2 The method of embodiment 43 or 43.1, wherein the vector results in the expression of a greater number of VLPs, as compared to a control vector lacking the DNA polynucleotide encoding the viral packaging signal.

[0301] Embodiment 43.3 The method of embodiment 43-43.2, wherein the VLPs encapsidating the viral packaging signal are more immunogenic than control VLPs comprising the antigen but lacking the viral packaging signal.

25 [0302] Embodiment 44. The method of any one of embodiments 41 to 43, wherein the method elicits an antibody response in the subject.

[0303] Embodiment 45. The method of embodiment 44, wherein the antibody response is a neutralizing antibody response.

[0304] Embodiment 46. The method of any one of embodiments 41 to 43, wherein the method elicits a cellular immune response.

30 [0305] Embodiment 47. The method of any one of embodiments 41 to 46, wherein the method elicits a prophylactic, protective and/or therapeutic immune response in the subject.

[0306] Embodiment 48. The method of any one of embodiments 41 to 47, wherein the administration is intradermal administration, intramuscular administration, subcutaneous administration, or intranasal administration.

- [0307] Embodiment 49. A polynucleotide comprising an expression cassette comprising a polynucleotide encoding a coronavirus protein and a polynucleotide encoding an enhancer protein, wherein the enhancer protein is a picornavirus leader (L) protein or a functional variant thereof.
- 5 [0308] Embodiment 50. The polynucleotide of embodiment 49, wherein the amino acid sequence of the enhancer protein has at least 95% identity to SEQ ID NO: 1, or at least 95% identity to SEQ ID NO: 2.
- [0309] Embodiment 51. The polynucleotide of embodiment 49 or embodiment 50, wherein the amino acid sequence of the enhancer protein is SEQ ID NO: 1, or SEQ ID NO: 2.
- 10 [0310] Embodiment 52. The polynucleotide of any one of embodiments 49-51, wherein the polynucleotide encoding the enhancer protein is operatively linked to a polynucleotide encoding an internal ribosome entry site (IRES).
- [0311] Embodiment 53. The polynucleotide of embodiment 52, wherein the polynucleotide encoding the IRES is SEQ ID NO: 24.
- 15 [0312] Embodiment 54. The polynucleotide of any one of embodiments 49-53, wherein the coronavirus protein is a coronavirus antigen.
- [0313] Embodiment 55. The polynucleotide of any one of embodiments 49-54, wherein the coronavirus is a betacoronavirus.
- [0314] Embodiment 56. The polynucleotide of embodiment 55, wherein the betacoronavirus
- 20 is severe acute respiratory syndrome (SARS) virus.
- [0315] Embodiment 57. The polynucleotide of embodiment 56, wherein the SARS virus is a SARS-CoV-2 virus.
- [0316] Embodiment 58. The polynucleotide of embodiment 55, wherein the betacoronavirus is Middle East respiratory syndrome (MERS) virus.
- 25 [0317] Embodiment 59. The polynucleotide of any one of embodiments 49-58, wherein the coronavirus protein is a coronavirus spike protein.
- [0318] Embodiment 60. The polynucleotide of embodiment 59, wherein the spike protein shares at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ
- 30 ID NO: 13.
- [0319] Embodiment 61. The polynucleotide of embodiment 59 or embodiment 60, wherein the spike protein is SEQ ID NO: 13.
- [0320] Embodiment 62. The polynucleotide of any one of embodiments 49-61, wherein the coronavirus protein is a coronavirus membrane (M) protein.

- [0321] Embodiment 63. The polynucleotide of embodiment 62, wherein the M protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 33.
- 5 [0322] Embodiment 64. The polynucleotide of embodiment 62 or embodiment 63, wherein the M protein is SEQ ID NO: 33.
- [0323] Embodiment 65. The polynucleotide of any one of embodiments 49-64, wherein the coronavirus protein is a coronavirus envelope (E) protein.
- [0324] Embodiment 66. The polynucleotide of embodiment 65, wherein the E protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%,
10 at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 22.
- [0325] Embodiment 67. The polynucleotide of embodiment 65 or embodiment 66, wherein the E protein is SEQ ID NO: 22.
- [0326] Embodiment 68. The polynucleotide of any one of embodiments 49-67, wherein the coronavirus protein is a coronavirus nucleocapsid (N) protein.
- 15 [0327] Embodiment 69. The polynucleotide of embodiment 68, wherein the N protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 20.
- [0328] Embodiment 70. The polynucleotide of embodiment 68 or embodiment 69, wherein the N protein is SEQ ID NO: 20.
- 20 [0329] Embodiment 71. The polynucleotide of any one of embodiments 49-70, wherein the coronavirus protein forms a virus-like particle (VLP).
- [0330] Embodiment 72. A polynucleotide comprising a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to the nucleic acid
25 sequence of SEQ ID NO: 30.
- [0331] Embodiment 73. The polynucleotide of embodiment 72, wherein the polynucleotide comprises a nucleic acid sequence of SEQ ID NO: 30.
- [0332] Embodiment 74. A polynucleotide comprising a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%,
30 at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to the nucleic acid sequence of SEQ ID NO: 31.
- [0333] Embodiment 75. The polynucleotide of embodiment 74, wherein the polynucleotide comprises a nucleic acid sequence of SEQ ID NO: 31.

- [0334] Embodiment 76. The polynucleotide of any one of embodiments 49-75, wherein the polynucleotide is a naked polynucleotide.
- [0335] Embodiment 77. The polynucleotide of any one of embodiments 49-76, wherein the polynucleotide is a deoxyribonucleic acid (DNA) polynucleotide.
- 5 [0336] Embodiment 78. The polynucleotide of any one of embodiments 49-76, wherein the polynucleotide is a ribonucleic acid (RNA) polynucleotide.
- [0337] Embodiment 79. The polynucleotide of any one of embodiments 49-71 and 76-78, wherein the expression cassette comprises a promoter operatively linked to each of the polynucleotide sequences of the expression cassette.
- 10 [0338] Embodiment 80. A kit comprising a vector, wherein the vector comprises an expression cassette comprising a polynucleotide encoding a coronavirus protein and a polynucleotide encoding an enhancer protein, wherein the enhancer protein is a picornavirus leader (L) protein or a functional variant thereof.
- [0339] Embodiment 81. The kit of embodiment 80, wherein the amino acid sequence of the enhancer protein has at least 95% identity to SEQ ID NO: 1, or at least 95% identity to SEQ ID NO: 2.
- 15 [0340] Embodiment 82. The kit of embodiment 80 or embodiment 81, wherein the polynucleotide encoding the enhancer protein is operatively linked to a polynucleotide encoding an internal ribosome entry site (IRES).
- 20 [0341] Embodiment 83. The kit of embodiment 82, wherein the polynucleotide encoding the IRES is SEQ ID NO: 24.
- [0342] Embodiment 84. The kit of any one of embodiments 80-83, wherein the coronavirus protein is a coronavirus antigen.
- [0343] Embodiment 85. The kit of any one of embodiments 80-84, wherein the coronavirus is a betacoronavirus.
- 25 [0344] Embodiment 86. The kit of embodiment 85, wherein the betacoronavirus is severe acute respiratory syndrome (SARS) virus.
- [0345] Embodiment 87. The kit of embodiment 86, wherein the SARS virus is a SARS-CoV-2 virus.
- 30 [0346] Embodiment 88. The kit of embodiment 85, wherein the betacoronavirus is Middle East respiratory syndrome (MERS) virus.
- [0347] Embodiment 89. The kit of any one of embodiments 80-88, wherein the coronavirus protein is a coronavirus spike protein.

[0348] Embodiment 90. The kit of embodiment 89, wherein the spike protein shares at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 13.

5 **[0349]** Embodiment 91. The kit of embodiment 90, wherein the spike protein is SEQ ID NO: 13.

[0350] Embodiment 92. The kit of any one of embodiments 80-91, wherein the coronavirus protein is a coronavirus membrane (M) protein.

10 **[0351]** Embodiment 93. The kit of embodiment 92, wherein the M protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 33.

[0352] Embodiment 94. The kit of embodiment 92 or embodiment 93, wherein the M protein is SEQ ID NO: 33.

[0353] Embodiment 95. The kit of any one of embodiments 80-94, wherein the coronavirus protein is a coronavirus envelope (E) protein.

15 **[0354]** Embodiment 96. The kit of embodiment 95, wherein the E protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 22.

[0355] Embodiment 97. The kit of embodiment 95 or embodiment 96, wherein the E protein is SEQ ID NO: 22.

20 **[0356]** Embodiment 98. The kit of any one of embodiments 80-97, wherein the coronavirus protein is a coronavirus nucleocapsid (N) protein.

[0357] Embodiment 99. The kit of embodiment 98, wherein the N protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to SEQ ID NO: 20.

25 **[0358]** Embodiment 100. The kit of embodiment 98 or embodiment 99, wherein the N protein is SEQ ID NO: 20.

30 **[0359]** Embodiment 101. The kit of embodiment 80, wherein the expression cassette comprises a polynucleotide, comprising a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5% or 100% identity to the nucleic acid sequence of SEQ ID NO: 30.

[0360] Embodiment 102. The kit of embodiment 80, wherein the expression cassette comprises a polynucleotide, comprising a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%,

at least 98%, at least 99%, at least 99.5% or 100% identity to the nucleic acid sequence of SEQ ID NO: 31.

[0361] Embodiment 103. The kit of any one of embodiments 80-102, wherein the kit comprises a pharmaceutically acceptable carrier.

- 5 **[0362]** Embodiment 104. A vector, comprising an expression cassette, said expression cassette comprising a promoter linked to a target gene, wherein the vector comprises a nucleic acid sequence encoding a viral packaging element.

[0363] Embodiment 105. The vector of embodiment 104, wherein the viral packaging element is a RNA polynucleotide.

- 10 **[0364]** Embodiment 106. The vector of embodiment 104 or 105, wherein the viral packaging element is derived from a coronavirus.

[0365] Embodiment 107. The vector of embodiment 106, wherein the viral packaging element is derived from SARS-CoV2.

- 15 **[0366]** Embodiment 108. The vector of any one of embodiments 104-107, wherein the nucleic acid sequence encoding the viral packaging element has at least about 70% identity to the nucleic acid sequence of SEQ ID NO: 34.

[0367] Embodiment 109. The method of expressing a target protein in a eukaryotic cell, comprising contacting the cell with the vector of any one of embodiments 104-108.

- 20 **[0368]** Embodiment 110. The method of embodiment 109, wherein contacting the cell with the vector results in the formation of virus-like particles (VLPs) comprising the target protein.

[0369] Embodiment 111. The method of embodiment 110, wherein contacting the cell with the vector results in the formation of a greater number of virus-like particles (VLPs) comprising the target protein, as compared to a control vector comprising the expression cassette but lacking the nucleic acid sequence encoding the viral packaging element.

- 25 **[0370]** Embodiment 112. The vector of any one of embodiments 33.1-33.3, or the method of any one of embodiments 40.1, 40.2, 43.1-43.3, wherein the nucleic acid sequence encoding the viral packaging element has at least about 70% identity to the nucleic acid sequence of SEQ ID NO: 34.

- 30 **[0371]** Embodiment 113. A vector for use as a vaccine, comprising an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding SEQ ID NO: 33 (M protein), a polynucleotide encoding a first proteolytic cleavage site, a polynucleotide encoding SEQ ID NO: 20 (N protein), a polynucleotide encoding a second proteolytic cleavage site, a polynucleotide encoding SEQ ID NO: 13 (S protein), a polynucleotide encoding a third proteolytic cleavage site, a polynucleotide encoding SEQ ID

NO: 22 (E protein), polynucleotide encoding SEQ ID NO: 24 (IRES), and a polynucleotide encoding SEQ ID NO: 2 (enhancer L protein).

[0372] Embodiment 114. A vector for use as a vaccine, comprising an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding
5 SEQ ID NO: 33 (M protein), a polynucleotide encoding a first proteolytic cleavage site, a polynucleotide encoding SEQ ID NO: 13 (S protein), a polynucleotide encoding a second proteolytic cleavage site, a polynucleotide encoding SEQ ID NO: 22 (E protein), polynucleotide encoding SEQ ID NO: 24 (IRES), a polynucleotide encoding SEQ ID NO: 2 (enhancer L protein), a polynucleotide encoding SEQ ID NO: 20 (N protein), and a
10 polynucleotide encoding SEQ ID NO: 34 (viral packaging signal).

[0373] Embodiment 115. A vector for use as a vaccine, comprising an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a
15 polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a mutated S protein wherein the S protein comprises SEQ ID NO: 51 or SEQ ID NO: 52 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein
20 wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, and a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto.

[0374] Embodiment 116. A vector for use as a vaccine, comprising an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a first polynucleotide encoding a viral packaging signal wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 9% identical thereto, a polynucleotide encoding an M protein wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least
30 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a mutated S protein wherein the S protein comprises SEQ ID NO: 51 or SEQ ID NO: 52 or an amino acid sequence at least 95% identical thereto, a polynucleotide

- encoding a proteolytic cleavage site, a polynucleotide encoding a E protein wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding IRES wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer
- 5 L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, and a second polynucleotide encoding a viral packaging signal, wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 9% identical thereto.

Claims:

1. A polynucleotide for use in a vaccine encoding a viral antigen protein and an enhancer protein, wherein the enhancer protein is a picornavirus leader (L) protein or a functional variant thereof.
2. The polynucleotide of claim 1, wherein the polynucleotide encodes an enhancer protein amino acid sequence with at least 95% identity to SEQ ID NO: 1, or at least 95% identity to SEQ ID NO: 2.
3. The polynucleotide of any one of claims 1-2, wherein the polynucleotide encodes an enhancer protein amino acid sequence of SEQ ID NO: 1 or SEQ ID NO: 2.
4. The polynucleotide of any one of claims 1-3, wherein the polynucleotide encodes an internal ribosome entry site (IRES) operatively linked to the enhancer protein.
5. The polynucleotide of claim 4, wherein the polynucleotide sequence encoding the IRES is SEQ ID NO: 24.
6. The polynucleotide of any one of claims 1-5, wherein the viral antigen protein is a structural protein.
7. The polynucleotide of claim any one of claims 1-6, wherein the viral antigen protein is derived from a coronavirus.
8. The polynucleotide of claim 7, wherein the coronavirus is a betacoronavirus.
9. The polynucleotide of claim 8, wherein the betacoronavirus is severe acute respiratory syndrome (SARS-CoV-2) virus.
10. The polynucleotide of claim 7, wherein the betacoronavirus is Middle East respiratory syndrome (MERS) virus.
11. The polynucleotide of any one of claims 1-10, wherein the viral antigen protein is a spike (S) protein.
12. The polynucleotide of claim 11, wherein the spike protein comprises SEQ ID NO: 13 or an amino acid sequence with at least 70%, at least 75%, at least 80%, at least 85%,

at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity thereto.

13. The polynucleotide of claim 11, wherein the spike protein comprises the mutant spike protein sequence amino acid sequence of SEQ ID NO: 51 or an amino acid sequence with at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity thereto.

14. The polynucleotide of any one of claims 1-13, wherein the viral antigen protein is a membrane (M) protein.

15. The polynucleotide of claim 14, wherein the M protein comprises the sequence of SEQ ID NO: 33 or an amino acid sequence with at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity thereto.

16. The polynucleotide of any one of claims 1-15, wherein the viral antigen protein is an envelope (E) protein.

17. The polynucleotide of claim 16, wherein the E protein comprises the amino acid sequence of SEQ ID NO: 22 or an amino acid sequence with at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity thereto.

18. The polynucleotide of any one of claims 1-17, wherein the viral antigen protein is a nucleocapsid (N) protein.

19. The polynucleotide of claim 18, wherein the N protein comprises the amino acid sequence of SEQ ID NO: 20 or an amino acid sequence with at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity thereto.

20. The polynucleotide of any one of claims 1-19, wherein the viral antigen protein forms a virus-like particle (VLP).

21. The polynucleotide of claim 6, wherein the viral antigen protein is a West Nile virus protein.

22. The polynucleotide of claim 21, wherein the viral antigen protein is precursor membrane protein (preM) and/or envelope glycoprotein (E).

23. A polynucleotide for use as a vaccine, comprising the sequence of SEQ ID NO: 30, or a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99%, sequence identity thereto.

24. A polynucleotide for use as a vaccine comprising the sequence of SEQ ID NO: 31, or a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99%, sequence identity thereto.

25. A polynucleotide for use as a vaccine comprising any of the sequences of SEQ ID NO: 35-49, and 55, or a nucleic acid sequence having at least 70% identity, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity thereto.

26. The polynucleotide of any one of claims 1-25, wherein an expression cassette comprises a promoter operatively linked to each of the polynucleotide sequences of the expression cassette.

27. The polynucleotide of claim 6, wherein the viral antigen protein is derived from Influenza viral proteins.

28. The polynucleotide of claim 27, wherein the viral antigen protein is one or more of HA, NA, M1, and M2.

29. The polynucleotide of claim 6, wherein the viral antigen protein is derived from Hepatitis B viral proteins

30. The polynucleotide of claim 29, wherein the viral antigen protein is one or more of sAg (S protein), sAg (M protein), preS1, preS2, and cAg (core antigen).

31. The polynucleotide of any one of claims 1-30, wherein the polynucleotide encodes a viral packaging signal.

32. The polynucleotide of claim 31, wherein the viral packaging signal is derived from a coronavirus.

33. A vector comprising the polynucleotide of any one of claims 1-32.

34. A vaccine composition, comprising the vector of claim 33 and a pharmaceutically acceptable carrier and/or an adjuvant.

35. The vaccine composition of claim 34, wherein the adjuvant is alum and/or monophosphoryl lipid A (MPL).

36. A method of expressing a viral antigen protein in a eukaryotic cell, comprising contacting a cell with the vector of claim 33.

37. The method of claim 36, wherein inclusion of the enhancer protein increases functional viral-like particle (VLP) production relative to a vector without an enhancer protein.

38. The method of claim 37, wherein inclusion of the enhancer protein increases functional VLP production by about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 100%, about 125%, about 150%, about 175%, about 200%, about 250%, about 300%, about 350%, about 400%, about 500%, or about 1000% relative to a vector without an enhancer protein.

39. A method of eliciting an immune response in a subject, comprising administering an effective amount of the vaccine composition of claims 34-35 to the subject.

40. The method of claim 39, wherein inclusion of the enhancer protein increases the duration of neutralizing antibodies in the subject relative to a vaccine composition without an enhancer protein.

41. The method of claim 40, wherein inclusion of the enhancer protein increases the duration of neutralizing antibodies in a subject by about 2-fold, about 3-fold, about 4-fold, about 5-fold, about 6-fold, about 7-fold, about 8-fold, about 9-fold, or about 10-fold relative to a vaccine composition without an enhancer protein.

42. The method of claims 39-41, wherein inclusion of the enhancer protein increases Th1 cellular response relative to a vaccine composition without an enhancer protein.

43. The method of claim 42, wherein inclusion of the enhancer protein increases Th1 response by about 25%, about 50%, about 100%, about 200%, about 300%, about 400%, about 500%, or about 1000% relative to a vaccine composition without an enhancer protein.

44. The method of any one of claims 39-43, wherein the method elicits a prophylactic or preventative immune response in the subject.

45. The method of any one of claims 39-44, wherein the method elicits a therapeutic immune response in the subject.

46. The method of any one of claims 39-45, wherein the administration is intradermal administration, intramuscular administration, subcutaneous administration, or intranasal administration.

47. A kit comprising the vaccine composition of claims 34-35, one or more vials, and instructions for use thereof.

48. A vector for use as a vaccine, comprising an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein, wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a first proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a second proteolytic cleavage site, a polynucleotide encoding an S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a third proteolytic cleavage site, a polynucleotide encoding an E protein, wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an IRES sequence, wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, and a polynucleotide encoding an enhancer L protein, wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto.

49. A vector for use as a vaccine, comprising an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein or an amino acid sequence at least 95% identical thereto, wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding

a first proteolytic cleavage site, a polynucleotide encoding an S protein wherein the S protein comprises SEQ ID NO: 13 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a second proteolytic cleavage site, a polynucleotide encoding an E protein or a polynucleotide sequence at least 95% identical thereto, wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an IRES sequence, wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, and a polynucleotide encoding a viral packaging signal, wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 95% identical thereto.

50. A vector for use as a vaccine, comprising an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a polynucleotide encoding an M protein, wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a mutated S protein, wherein the mutated S protein comprise SEQ ID NO: 51 or SEQ ID NO: 52 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein, wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an IRES sequence, wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, and a polynucleotide encoding an enhancer L protein, wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto.

51. A vector for use as a vaccine, comprising an expression cassette, comprising the following elements in the 5' to 3' order: a promoter, a first polynucleotide encoding a viral packaging signal, wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an M protein, wherein the M protein comprises SEQ ID NO: 33 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide

encoding an N protein, wherein the N protein comprises SEQ ID NO: 20 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a mutated S protein, wherein the S protein comprises SEQ ID NO: 51 or SEQ ID NO: 52 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding a proteolytic cleavage site, a polynucleotide encoding a E protein, wherein the E protein comprises SEQ ID NO: 22 or an amino acid sequence at least 95% identical thereto, a polynucleotide encoding an IRES sequence, wherein the IRES sequence comprises SEQ ID NO: 24 or a polynucleotide sequence at least 95% identical thereto, a polynucleotide encoding an enhancer L protein, wherein the L protein comprises SEQ ID NO: 2 or an amino acid sequence at least 95% identical thereto, and a second polynucleotide encoding a viral packaging signal, wherein the viral packaging signal comprises SEQ ID NO: 34 or a polynucleotide sequence at least 95% identical thereto.

52. A vaccine composition, comprising the vector of any one of claims 48-51 and a pharmaceutically acceptable carrier and/or an adjuvant.

53. A method of eliciting an immune response in a subject, comprising administering an effective amount of the vaccine composition of claim 52 to the subject.

54. The method of claim 53, wherein inclusion of the enhancer protein increases the duration of neutralizing antibodies in the subject relative to a vaccine composition without an enhancer protein.

55. The method of claim 54, wherein inclusion of the enhancer protein increases the duration of neutralizing antibodies in a subject by about 2-fold, about 3-fold, about 4-fold, about 5-fold, about 6-fold, about 7-fold, about 8-fold, about 9-fold, or about 10-fold relative to a vaccine composition without an enhancer protein.

56. The method of claim 53, wherein inclusion of the enhancer protein increases Th1 cellular response relative to a vaccine composition without an enhancer protein.

57. The method of claim 56, wherein inclusion of the enhancer protein increases Th1 response by about 25%, about 50%, about 100%, about 200%, about 300%, about 400%, about 500%, or about 1000% relative to a vaccine composition without an enhancer protein.

58. The method of any one of claims 53-57, wherein the method elicits a prophylactic, protective and/or therapeutic immune response in the subject.

59. The method of any one of claims 53-57, wherein the administration is intradermal administration, intramuscular administration, subcutaneous administration, or intranasal administration.