



(51) International Patent Classification:
C12N 15/86 (2006.01)

(21) International Application Number:
PCT/US2017/054928

(22) International Filing Date:
03 October 2017 (03.10.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/409,228 17 October 2016 (17.10.2016) US

(71) Applicant: SYNTHETIC GENOMICS, INC. [US/US];
11149 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventor: KAMRUD, Kurt, Iver; 11149 North Torrey
Pines Road, La Jolla, CA 92037 (US).

(74) Agent: MALLON, Joseph, J.; Knobbe, Martens, Olson &
Bear, Llp, 2040 Main Street, 14th Floor, Irvine, CA 92614
(US).

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

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

(54) Title: RECOMBINANT VIRUS REPLICON SYSTEMS AND USES THEREOF

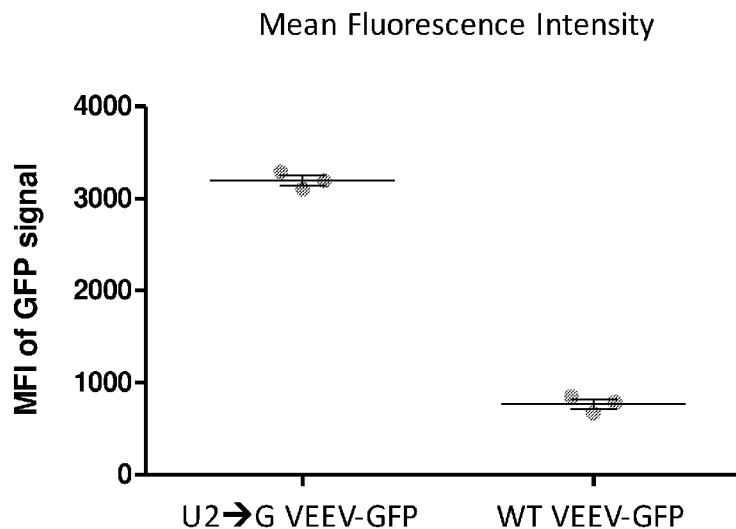


FIG. 6

(57) Abstract: The present disclosure generally relates to viral-based expression systems suitable for the production of molecules of interest. The disclosure relates to nucleic acid constructs, such as expression vectors, containing a modified replicon RNA which includes a modified 5'-untranslated region (5'-UTR) and, optionally, at least some of its original viral sequence encoding structural proteins having been deleted. Also disclosed are methods for producing polypeptides of interest.



Published:

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

RECOMBINANT VIRUS REPLICON SYSTEMS AND USES THEREOF

RELATED APPLICATIONS

[0001] The present application claims priority under 35 U.S.C. § 119(e) to U.S. Provisional Application No. 62/409228, filed on October 17, 2016, which is herein expressly incorporated by reference in its entirety.

INCORPORATION OF THE SEQUENCE LISTING

[0002] The material in the accompanying sequence listing is hereby incorporated by reference into this application. The accompanying sequence listing text file, name SGI011WO_SeqListing.txt, was created on October 3, 2017 and is 34 KB.

FIELD

[0003] The present disclosure relates to the field of molecular biology, including nucleic acid molecules comprising modified viral replicons and the use of such nucleic acid molecules for production of desired products in suitable host cells in cell culture or in a living body.

BACKGROUND

[0004] In recent years, several different groups of animal viruses have been subjected to genetic manipulation either by homologous recombination or by direct engineering of their genomes. The availability of reverse genetics systems for both DNA and RNA viruses has created new perspectives for the use of recombinant viruses, for example, as vaccines, expression vectors, anti-tumor agents, gene therapy vectors, and drug delivery vehicles.

[0005] For example, many viral-based expression vectors have been deployed for expression of heterologous proteins in cultured recombinant cells. In particular, the application of modified viral vectors for gene expression in host cells continues to expand. Recent advances in this regard include further development of techniques and systems for production of multi-subunit protein complexes, and co-expression of protein-modifying

enzymes to improve heterologous protein production. Other recent progresses regarding viral expression vector technologies include many advanced genome engineering applications for controlling gene expression, preparation of viral vectors, *in vivo* gene therapy applications, and creation of vaccine delivery vectors.

[0006] However, there is still a need for more efficient methods and systems for expressing genes of interest in recombinant expression systems.

SUMMARY

[0007] This section provides a general summary of the present application, and is not comprehensive of its full scope or all of its features.

[0008] In one aspect, disclosed herein is a nucleic acid molecule including a modified replicon RNA, in which the modified replicon RNA includes a modified 5'-UTR and is devoid of at least a portion of a nucleic acid sequence encoding viral structural proteins. In various embodiments of this aspect and other aspects of the present disclosure, the nucleic acid molecule as disclosed herein can include one or more of the following features. In some embodiments, the modified replicon RNA is a modified alphavirus replicon RNA. In some embodiments, the modified alphavirus replicon RNA includes a modified alphavirus genome. In some embodiments, the modified 5'-UTR includes one or more nucleotide substitutions at position 1, 2, 4, or a combination thereof. In some embodiments, at least one of the nucleotide substitutions is a nucleotide substitution at position 2 of the modified 5'-UTR. In some embodiments, the nucleotide substitutions at position 2 of the modified 5'-UTR is a U→G substitution.

[0009] In some embodiments, the nucleic acid molecule includes a modified replicon RNA which is devoid of a substantial portion of the nucleic acid sequence encoding viral structural proteins. In some embodiments, the modified alphavirus genome or replicon RNA as disclosed herein includes no nucleic acid sequence encoding viral structural proteins.

[0010] In various embodiments of this aspect and other aspects of the present disclosure, the nucleic acid molecule further includes one or more expression cassettes, wherein each of the expression cassettes includes a promoter operably linked to a heterologous nucleic acid sequence. In some embodiments, the nucleic acid molecule

includes at least two, three, four, five, or six expression cassettes. In some embodiments, the promoter of at least one of the expression cassettes is or comprises a 26S subgenomic promoter.

[0011] In some embodiments, the heterologous nucleic acid sequence of at least one of the expression cassettes as disclosed herein includes a coding sequence of a gene of interest (GOI). In some embodiments, the GOI encodes a polypeptide selected from the group consisting of a therapeutic polypeptide, a prophylactic polypeptide, a diagnostic polypeptide, a neutraceutical polypeptide, an industrial enzyme, and a reporter polypeptide. In some embodiments, the GOI encodes a polypeptide selected from the group consisting of an antibody, an antigen, an immune modulator, and a cytokine. In some particular embodiments, the coding sequence of the GOI is optimized for expression at a level higher than the expression level of a reference coding sequence.

[0012] In some embodiments, the nucleic acid molecule includes a modified replicon RNA comprising a modified genome or replicon RNA of a virus belonging to the *Alphavirus* genus of the *Togaviridae* family. In some embodiments, the modified genome or replicon RNA is of an alphavirus belonging to the VEEV/EEEV group, or the SF group, or the SIN group. In some embodiments, the alphavirus is selected from the group consisting of Eastern equine encephalitis virus (EEEV), Venezuelan equine encephalitis virus (VEEV), Everglades virus (EVEV), Mucambo virus (MUCV), Pixuna virus (PIXV), Middleburg virus (MIDV), Chikungunya virus (CHIKV), O’Nyong-Nyong virus (ONNV), Ross River virus (RRV), Barmah Forest virus (BF), Getah virus (GET), Sagiya virus (SAGV), Bebaru virus (BEBV), Mayaro virus (MAYV), Una virus (UNAV), Sindbis virus (SINV), Aura virus (AURAV), Whataroa virus (WHAV), Babanki virus (BABV), Kyzylagach virus (KYZV), Western equine encephalitis virus (WEEV), Highland J virus (HJV), Fort Morgan virus (FMV), Ndumu (NDUV), and Buggy Creek virus. In some embodiments, the alphavirus is Venezuelan equine encephalitis virus (VEEV).

[0013] Some embodiments provide a nucleic acid molecule which includes a modified alphavirus genome or replicon RNA that is operably linked to a heterologous regulatory element. In some embodiments, the heterologous regulatory element includes a promoter sequence. In some embodiments, the promoter sequence includes a T7 promoter

sequence. In some embodiments, the heterologous regulatory element includes a transcriptional termination sequence. In some embodiments, the transcriptional termination sequence is or comprises a T7 termination sequence.

[0014] In some embodiments, the nucleic acid molecule as disclosed herein includes a modified alphavirus genome or replicon RNA including a modified alphavirus genome or replicon RNA, wherein the nucleic acid molecule exhibits at least 80% sequence identity to the nucleic acid sequence of SEQ ID NO: 1, wherein the modified alphavirus genome or replicon RNA comprises a U→G substitution at position 2 of the 5'-untranslated region (5'-UTR) and is devoid of at least a portion of the sequence encoding viral structural proteins. In some embodiments, the nucleic acid molecule exhibits at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the nucleic acid sequence of SEQ ID NO: 1.

[0015] In some embodiments, the nucleic acid molecule as disclosed herein includes a modified alphavirus genome or replicon RNA, wherein the modified alphavirus genome or replicon RNA comprises a 5'-UTR exhibiting at least 80% sequence identity to the nucleic acid sequence of at least one of SEQ ID NOS: 2-18 and a U→G substitution at position 2 of the 5'-UTR, and wherein the modified alphavirus genome or replicon RNA is devoid of at least a portion of the sequence encoding viral structural proteins. In some embodiments, the modified alphavirus genome or replicon RNA comprises a 5'-UTR exhibiting at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the nucleic acid sequence of at least one of SEQ ID NOS: 2-18. In certain embodiments, the modified alphavirus genome or replicon RNA is devoid of a substantial portion of the nucleic acid sequence encoding viral structural proteins. In certain embodiments, the modified alphavirus genome or replicon RNA comprises no nucleic acid sequence encoding viral structural proteins.

[0016] In one aspect, some embodiments disclosed herein relate to a recombinant cell which includes a nucleic acid molecule described herein. In some embodiments, the recombinant cell is a prokaryotic cell or a eukaryotic cell. In some embodiments, the recombinant cell is an animal cell. In some embodiments, the recombinant cell is a vertebrate animal cell or an invertebrate cell. In some embodiments, the recombinant cell is selected

from the group consisting of a pulmonary equine artery endothelial cell, an equine dermis cell, a baby hamster kidney (BHK) cell, a rabbit kidney cell, a mouse muscle cell, a mouse connective tissue cell, a human cervix cell, a human epidermoid larynx cell, a Chinese hamster ovary cell (CHO), a human HEK-293 cell, a mouse 3T3 cell, a Vero cell, a Madin-Darby Canine Kidney Epithelial Cell (MDCK), primary chicken fibroblast cell, a HuT78 cell, an A549 lung cell, a HeLa cell, a PER.C6® cell, a WI-38 cell, a MRC-5 cell, a FRhL-2, and a CEM T-cell. Some embodiments disclosed herein provide a cell culture that includes at least one recombinant cell as disclosed herein.

[0017] In one aspect, some embodiments provide a method for producing a polypeptide of interest that involves culturing a host cell comprising a nucleic acid molecule which comprises a modified 5'-UTR and is devoid of at least a portion of a nucleic acid sequence encoding viral structural proteins. In some embodiments, the recombinant cell is a vertebrate cell or an invertebrate cell.

[0018] In a further aspect, some embodiments provide a method for producing a polypeptide of interest in a subject that involves administering to the subject a nucleic acid molecule which comprises a modified 5'-UTR and is devoid of at least a portion of a nucleic acid sequence encoding viral structural proteins. In some embodiments, the subject is human, horse, pig, primate, mouse, cattle, swine, sheep, rabbit, cat, dog, bird, fish, goat, donkey, hamster, or buffalo.

[0019] Implementations of embodiments of the methods according to the present disclosure can include one or more of the following features. In some embodiments, the modified replicon RNA is a modified alphavirus replicon RNA. In some embodiments, the modified alphavirus replicon RNA includes a modified alphavirus genome. In some embodiments, the modified 5'-UTR includes one or more nucleotide substitutions at position 1, 2, 4, or a combination thereof. In some embodiments, at least one of the nucleotide substitutions is a nucleotide substitution at position 2 of the modified 5'-UTR. In some embodiments, the nucleotide substitutions at position 2 of the modified 5'-UTR is a U→G substitution. In certain embodiments, the modified replicon RNA is devoid of a substantial portion of the nucleic acid sequence encoding viral structural proteins. In some

embodiments, the modified alphavirus genome or replicon RNA includes no nucleic acid sequence encoding viral structural proteins.

[0020] In some embodiments, the nucleic acid molecule as described herein further includes one or more expression cassettes, wherein each of the expression cassettes includes a promoter operably linked to a heterologous nucleic acid sequence. In some embodiments, the nucleic acid molecule includes at least two, at least three, at least four, at least five, or at least six expression cassettes. In some embodiments, the promoter of at least one of the expression cassettes includes a 26S subgenomic promoter. In some particular embodiments, the promoter of at least one of the expression cassettes includes an alphavirus 26S subgenomic promoter. Preferably, the promoter comprises a Venezuelan equine encephalitis (VEEV) 26S subgenomic promoter. In certain embodiments, the heterologous nucleic acid sequence of at least one of the expression cassettes includes a coding sequence of a gene of interest (GOI). The coding sequence of the GOI, in some embodiments, is optimized for expression at a level higher than the expression level of a reference coding sequence. In some embodiments, the promoter operably linked to the heterologous nucleic acid sequence comprises a heterologous promoter sequence. Suitable heterologous promoters include, but are not limited to, regulatory elements from internal ribosome entry site (IRES) derived from encephalomyocarditis viruses (EMCV), Bovine Viral Diarrhea Viruses (BVDV), polioviruses, Foot-and-mouth disease viruses (FMD), enterovirus 71, or hepatitis C viruses.

[0021] In some embodiments, the modified replicon RNA includes a modified genome or replicon RNA of a virus belonging to the *Alphavirus* genus of the *Togaviridae* family. In some embodiments, the modified genome or replicon RNA is of an alphavirus belonging to the VEEV/EEEV group, or the SF group, or the SIN group. In some embodiments, the alphavirus is selected from the group consisting of Eastern equine encephalitis virus (EEEV), Venezuelan equine encephalitis virus (VEEV), Everglades virus (EVEV), Mucambo virus (MUCV), Pixuna virus (PIXV), Middleburg virus (MIDV), Chikungunya virus (CHIKV), O'Nyong-Nyong virus (ONNV), Ross River virus (RRV), Barmah Forest virus (BF), Getah virus (GET), Sagiya virus (SAGV), Bebaru virus (BEBV), Mayaro virus (MAYV), Una virus (UNAV), Sindbis virus (SINV), Aura virus

(AURAV), Whataroa virus (WHAV), Babanki virus (BABV), Kyzylagach virus (KYZV), Western equine encephalitis virus (WEEV), Highland J virus (HJV), Fort Morgan virus (FMV), Ndumu (NDUV), and Buggy Creek virus. In some embodiments, the alphavirus is Venezuelan equine encephalitis virus (VEEV).

[0022] In some embodiments, the modified alphavirus genome or replicon RNA is operably linked to a heterologous regulatory element. In some embodiments, the heterologous regulatory element includes a promoter sequence. In some embodiments, the promoter sequence includes a T7 promoter sequence. In some embodiments, the heterologous regulatory element comprises a transcriptional termination sequence. In some embodiments, the transcriptional termination sequence is or comprises a T7 termination sequence.

[0023] In some embodiments, the nucleic acid molecule includes a modified alphavirus genome or replicon RNA, wherein the modified alphavirus genome or replicon RNA includes a 5'-UTR exhibiting at least 80% sequence identity to the nucleic acid sequence of SEQ ID NO: 1 and a U->G substitution at position 2 of the 5'-UTR, and wherein the modified alphavirus genome or replicon RNA is devoid of at least a portion of the sequence encoding viral structural proteins. In some embodiments, the nucleic acid molecule exhibits at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the nucleic acid sequence of SEQ ID NO: 1. In some embodiments, the nucleic acid molecule includes a modified alphavirus genome or replicon RNA, wherein the modified alphavirus genome or replicon RNA includes a 5'-UTR exhibiting at least 80% sequence identity to the nucleic acid sequence of at least one of SEQ ID NOS: 2-18 and a U->G substitution at position 2 of the 5'-UTR, and wherein the modified alphavirus genome or replicon RNA is devoid of at least a portion of the sequence encoding viral structural proteins. In some embodiments, the modified alphavirus genome or replicon RNA includes a 5'-UTR exhibiting at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the nucleic acid sequence of at least one of SEQ ID NOS: 2-18.

[0024] In a further aspect, some embodiments disclosed herein provide recombinant polypeptides produced by a method in accordance with one or more embodiments disclosed herein.

[0025] In one aspect, some embodiments disclosed herein relate to a composition including a recombinant polypeptide as described herein. In some embodiments, the composition is a pharmaceutical composition comprising a pharmaceutical acceptable carrier. In some embodiments, the composition is a prophylactic composition, a neutraceutical composition, a pharmaceutical composition, or a combination thereof.

[0026] In a further aspect, some embodiments disclosed herein relate to a composition including a nucleic acid molecule as disclosed herein. In some embodiments, the composition is a pharmaceutical composition comprising a pharmaceutical acceptable carrier. In some embodiments, the composition is a prophylactic composition, a neutraceutical composition, a pharmaceutical composition, or a combination thereof.

[0027] In yet a further aspect, some embodiments disclosed herein relate to a composition including a recombinant cell as described herein. In some embodiments, the composition is a pharmaceutical composition comprising a pharmaceutical acceptable carrier. In some embodiments, the composition is a prophylactic composition, a neutraceutical composition, a pharmaceutical composition, or a combination thereof.

[0028] The foregoing summary is illustrative only and is not intended to be in any way limiting. In addition to the illustrative embodiments and features described herein, further aspects, embodiments, objects and features of the application will become fully apparent from the drawings and the detailed description and the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0029] **FIGURE 1A** shows a sequence alignment of the 5'untranslated regions (5'-UTRs) from representative alphaviruses: Aura virus (AURAV; SEQ ID NO: 2), Chikungunya virus (CHIKV, SEQ ID NO: 3), O'Nyong-Nyong virus (SEQ ID NO: 4, ONNV), Bebaru virus (SEQ ID NO: 5, BEBV), Semliki Forest virus (SEQ ID NO: 6, SFV), Mayaro virus (SEQ ID NO: 7, MAYV), Getah virus (SEQ ID NO: 8, GETV), Sagiyama virus (SEQ ID NO: 9, SAGV), Ndumu virus (SEQ ID NO: 10, NDUV), Middleburg virus (SEQ ID

NO: 11, MIDV), Eastern equine encephalitis virus (SEQ ID NO: 12, EEEV), Fort Morgan virus (SEQ ID NO: 13, FMV), Buggy Creek virus (SEQ ID NO: 14, Buggy), Venezuelan equine encephalitis virus (SEQ ID NO: 15, VEEV), Whataroa virus (SEQ ID NO: 16, WHAV), Sindbis virus (SEQ ID NO: 17, SINV), and Bebaru virus (SEQ ID NO: 18, BEBV). The sequence alignment of **FIG. 1A** was generated using the program MUSCLE 3.6 with default setting. In the sequence alignment shown herein, a dash in an aligned sequence, which is created by the program MUSCLE 3.6 for optimal alignment, represents a gap, *i.e.*, a lack of nucleotide at that position. As discussed in detail below, several conserved nucleotide residues have been identified in this sequence comparison analysis. Asterisks identify identical nucleotide residues among the aligned sequences. **FIG. 1B** shows a graphical representation of the consensus sequence as a weighted consensus in which the size of the letter designating a given amino acid is proportional to the conservation of the residue in the different sequences used to generate the motif (the size of the letter denotes a residue's relative frequency at that position among the aligned sequences). The size of the character reflects the information content measured in bits.

[0030] **FIGURE 2** shows a schematic illustration of the structure of a non-limiting exemplary base monovalent VEEV replicon design, Rep-Alpha, which includes a T7 promoter sequence, a 5'UTR sequence from VEEV having a U2→G substitution as described herein, coding sequence of the nonstructural polypeptides nsp1, nsp2, nsp3, and nsp4 of an alphavirus genome. The base monovalent VEEV replicon Rep-Alpha also contains a 26S subgenomic promoter sequence, 3'UTR sequence, T7 termination sequence, polyadenylation sequence PolyA, and a number of unique restriction sites engineered to facilitate insertion of additional components into the replicon.

[0031] **FIGURES 3 and 4** graphically depicts the structures of two non-limiting exemplary monovalent VEEV replicon designs, in which the gene of interest (GOI) operably incorporated into the vector was an A Vietnam 1203 HA gene (**FIG. 3**) and an enhanced green fluorescence protein (eGFP) reporter gene (**FIG. 4**), respectively.

[0032] **FIGURE 5** graphically summarizes the results of exemplary experiments illustrating that a U2→G modification at position 2 of the 5'UTR in a modified VEEV-HA replicon does not affect biological activity of the modified replicon. Flow cytometry analysis

(FACS) was performed on cells electroporated with a modified U2→G VEEV-HA replicon expressing an influenza hemagglutinin – HA gene (also see **FIG. 3** for structural organization of the replicon). A wild-type VEEV-HA replicon, *i.e.* containing a U residue at position 2, was used as control.

[0033] FIGURE 6 graphically summarizes the results of an exemplary flow cytometry analysis performed to demonstrate that expression from an alphavirus replicon carrying a U2→G modification in the 5'UTR is not restricted to any particular gene of interest (GOI). In this experiment, the modified alphavirus replicon was engineered to express a green fluorescent protein (GFP) reporter gene (also see **FIG. 4** for structural organization of the replicon). The U2→G modification in the 5'UTR of the modified VEEV-GFP replicon was shown to enhance expression of GFP gene by 3-fold relative to the expression detected from a wild-type control replicon.

[0034] FIGURE 7 graphically summarizes the results of yet another exemplary flow cytometry analysis assessing expression of a red Firefly reporter gene from a modified VEEV replicon. In this experiment, the U2→G modification in the 5'UTR of the modified VEEV-rFF replicon was shown to enhance expression of the red Firefly gene by 2-fold relative to the expression detected from a wild-type control replicon.

[0035] FIGURE 8 schematically depicts a non-limiting exemplary alphavirus genomic structure and genome expression (adapted from Strauss *et al.*, Microbiological Reviews, pp. 491-562, September 1994). Genome organization of a SIN virus is shown. The names of the nonstructural genes and structural protein genes are given. Referenced to the nomenclature of the genes and proteins can be found in Strauss *et al.*, supra, 1994. The 49S genomic RNA is illustrated schematically in the center, with its translated ORF shown as an open box. Small black boxes are conserved sequence elements; the open diamond denotes the leaky opal termination codon. The nonstructural polyproteins and their processed products are shown above. Termination at the opal codon produces P123, whose major function in replication is believed to be as a proteinase that acts in *trans* to process the polyproteins in active RNA replicases; this proteinase domain is found in the nsP2 region. Readthrough of the opal stop codon produces P1234, which can form an active replicase. The 26S subgenomic mRNA is expanded below to show the structural ORF and its

translation products. Polypeptides present in the virion are shaded. vcRNA is the minus-strand complement of the genomic RNA.

[0036] The foregoing and other features of the present disclosure will become more fully apparent from the following description and appended claims, taken in conjunction with the accompanying drawings. Understanding that these drawings depict only several embodiments in accordance with the disclosure and are not to be considered limiting of its scope; the disclosure will be described with additional specificity and detail through use of the accompanying drawings.

DETAILED DESCRIPTION

[0037] The present disclosure generally relates to viral expression systems with superior expression potential which are suitable for expressing heterologous molecules such as, for example, vaccines and therapeutic polypeptides, in recombinant cells. For example, some embodiments of the disclosure relate to nucleic acid molecules such as, *e.g.* expression constructs and vectors, containing a modified replicon RNA which includes a modified 5'-untranslated region (5'-UTR) and, optionally, at least some of its original viral sequence encoding structural proteins having been deleted. Also included in accordance with some embodiments of the disclosure are viral-based expression vectors including one or more expression cassettes encoding heterologous polypeptide. Accordingly, recombinant cells that are genetically modified to include one or more of the nucleic acid molecules disclosed herein, as well as biomaterials and recombinant products derived from such cells are also within the scope of the application. Further provided in particular aspects of the disclosure are compositions that include one or more of the molecules and/or recombinant cells disclosed herein.

[0038] Self-amplifying RNAs (replicons) based on RNA virus (*e.g.*, alphaviruses) can be used as robust expression systems. For example, modifications to the wild-type virus (*e.g.*, alphavirus) 5' untranslated region (UTR) can allow dissection of the key RNA nucleotides that comprise the promoter elements involved in both RNA replication and RNA transcription. Development of enhanced viral (*e.g.*, alphavirus) expression systems by manipulation of the 5' UTR sequence represents an important advancement in replicon

platform development. Without being limited by any particular theory, it is believed that a non-limiting advantage of using alphaviruses as viral expression vectors is that they can direct the synthesis of large amounts of heterologous proteins in recombinant host cells. In particular, among other advantages, the alphavirus replicon platform systems disclosed herein, in some embodiments, are capable of expressing up to three times the amount of protein normally expressed from an alphavirus replicon. This improvement is significant given the already naturally high expression levels noted with standard alphavirus replicon systems and that the 5' UTR mutation that imparts this capability was previously considered to be a nearly lethal mutation to alphavirus replication and transcription. For example, polypeptides such as therapeutic single chain antibodies may be most effective if expressed at high levels *in vivo*. In addition, for producing recombinant antibodies purified from cells in culture (*ex vivo*), high protein expression from a replicon RNA may increase overall yields of the antibody product. Furthermore, if the protein being expressed is a vaccine antigen, high level expression may induce the most robust immune response *in vivo*.

[0039] In the following detailed description, reference is made to the accompanying drawings, which form a part hereof. In the drawings, similar symbols typically identify similar components, unless context dictates otherwise. The illustrative alternatives described in the detailed description, drawings, and claims are not meant to be limiting. Other alternatives may be used, and other changes may be made, without departing from the spirit or scope of the subject matter presented here. It will be readily understood that the aspects, as generally described herein, and illustrated in the Figures, can be arranged, substituted, combined, and designed in a wide variety of different configurations, all of which are explicitly contemplated and make part of this application.

[0040] Unless otherwise defined, all terms of art, notations and other scientific terms or terminology used herein are intended to have the meanings commonly understood by those of skill in the art to which this application pertains. In some cases, terms with commonly understood meanings are defined herein for clarity and/or for ready reference, and the inclusion of such definitions herein should not necessarily be construed to represent a substantial difference over what is generally understood in the art. Many of the techniques

and procedures described or referenced herein are well understood and commonly employed using conventional methodology by those of skill in the art.

Some Definitions

[0041] The singular form “a”, “an”, and “the” include plural references unless the context clearly dictates otherwise. For example, the term “a cell” includes one or more cells, comprising mixtures thereof. “A and/or B” is used herein to include all of the following alternatives: “A”, “B”, “A or B”, and “A and B”.

[0042] The term “about”, as used herein, has its ordinary meaning of approximately. If the degree of approximation is not otherwise clear from the context, “about” means either within plus or minus 10% of the provided value, or rounded to the nearest significant figure, in all cases inclusive of the provided value. Where ranges are provided, they are inclusive of the boundary values.

[0043] The terms, “cells”, “cell cultures”, “cell line”, “recombinant host cells”, “recipient cells” and “host cells” as used herein, include the primary subject cells and any progeny thereof, without regard to the number of transfers. It should be understood that not all progeny are exactly identical to the parental cell (due to deliberate or inadvertent mutations or differences in environment); however, such altered progeny are included in these terms, so long as the progeny retain the same functionality as that of the originally transformed cell.

[0044] As used herein, the term “construct” is intended to mean any recombinant nucleic acid molecule such as an expression cassette, plasmid, cosmid, virus, autonomously replicating polynucleotide molecule, phage, or linear or circular, single-stranded or double-stranded, DNA or RNA polynucleotide molecule, derived from any source, capable of genomic integration or autonomous replication, comprising a nucleic acid molecule where one or more nucleic acid sequences has been linked in a functionally operative manner, *e.g.* operably linked.

[0045] The term “gene” is used broadly to refer to any segment of nucleic acid molecule that encodes a protein or that can be transcribed into a functional RNA. Genes may include sequences that are transcribed but are not part of a final, mature, and/or functional RNA transcript, and genes that encode proteins may further comprise sequences that are

transcribed but not translated, for example, 5' untranslated regions, 3' untranslated regions, introns, *etc.* Further, genes may optionally further comprise regulatory sequences required for their expression, and such sequences may be, for example, sequences that are not transcribed or translated. Genes can be obtained from a variety of sources, including cloning from a source of interest or synthesizing from known or predicted sequence information, and may include sequences designed to have desired parameters.

[0046] The term "heterologous" when used in reference to a polynucleotide, a gene, or a nucleic acid molecule refers to a polynucleotide, gene, or a nucleic acid molecule that is not derived from the host species. For example, "heterologous gene" or "heterologous nucleic acid sequence" as used herein, refers to a gene or nucleic acid sequence from a different species than the species of the host organism it is introduced into. When referring to a gene regulatory sequence or to an auxiliary nucleic acid sequence used for manipulating expression of a gene sequence (*e.g.* a 5' untranslated region, 3' untranslated region, poly A addition sequence, *etc.*) or to a nucleic acid sequence encoding a protein domain or protein localization sequence, "heterologous" means that the regulatory or auxiliary sequence or sequence encoding a protein domain or localization sequence is from a different source than the gene with which the regulatory or auxiliary nucleic acid sequence or nucleic acid sequence encoding a protein domain or localization sequence is juxtaposed in a genome. Thus, a promoter operably linked to a gene to which it is not operably linked to in its natural state (for example, in the genome of a non-genetically engineered organism) is referred to herein as a "heterologous promoter," even though the promoter may be derived from the same species (or, in some cases, the same organism) as the gene to which it is linked. For example, in some embodiments disclosed herein, a coding sequence of a heterologous gene of interest (GOI) is not linked to the EAV replicon sequence in its natural state. In some embodiments, the coding GOI sequence is derived from another organism, such as another virus, bacteria, fungi, human cell (*e.g.*, tumor Ag), parasite (*e.g.*, malaria), *etc.*)

[0047] The terms "nucleic acid molecule" and "polynucleotide" are used interchangeably herein, and refer to both RNA and DNA molecules, including nucleic acid molecules comprising cDNA, genomic DNA, synthetic DNA, and DNA or RNA molecules containing nucleic acid analogs. Nucleic acid molecules can have any three-dimensional

structure. A nucleic acid molecule can be double-stranded or single-stranded (*e.g.*, a sense strand or an antisense strand). Non-limiting examples of nucleic acid molecules include genes, gene fragments, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, siRNA, micro-RNA, tracrRNAs, crRNAs, guide RNAs, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, nucleic acid probes and nucleic acid primers. A nucleic acid molecule may contain unconventional or modified nucleotides. The terms “polynucleotide sequence” and “nucleic acid sequence” as used herein interchangeably refer to the sequence of a polynucleotide molecule. The nomenclature for nucleotide bases as set forth in 37 CFR §1.822 is used herein.

[0048] Nucleic acid molecules can be nucleic acid molecules of any length, including but not limited to, nucleic acid molecules that are between about 3 Kb and about 50 Kb, for example between about 3 Kb and about 40 Kb, between about 3 Kb and about 40 Kb, between about 3 Kb and about 30 Kb, between about 3 Kb and about 20 Kb, between 5 Kb and about 40 Kb, between about 5 Kb and about 40 Kb, between about 5 Kb and about 30 Kb, between about 5 Kb and about 20 Kb, or between about 10 Kb and about 50 Kb, for example between about 15 Kb to 30Kb, between about 20 Kb and about 50 Kb, between about 20 Kb and about 40 Kb, about 5 Kb and about 25 Kb, or about 30 Kb and about 50 Kb. The nucleic acid molecules can also be, for example, more than 50 kb.

[0049] The polynucleotides of the present disclosure can be “biologically active” with respect to either a structural attribute, such as the capacity of a nucleic acid to hybridize to another nucleic acid, or the ability of a polynucleotide sequence to be recognized and bound by a transcription factor and/or a nucleic acid polymerase.

[0050] The term “recombinant” or “engineered” nucleic acid molecule as used herein, refers to a nucleic acid molecule that has been altered through human intervention. As non-limiting examples, a cDNA is a recombinant DNA molecule, as is any nucleic acid molecule that has been generated by *in vitro* polymerase reaction(s), or to which linkers have been attached, or that has been integrated into a vector, such as a cloning vector or expression vector. As non-limiting examples, a recombinant nucleic acid molecule: 1) has been synthesized or modified *in vitro*, for example, using chemical or enzymatic techniques (for example, by use of chemical nucleic acid synthesis, or by use of enzymes for the replication,

polymerization, exonucleolytic digestion, endonucleolytic digestion, ligation, reverse transcription, transcription, base modification (including, *e.g.*, methylation), or recombination (including homologous and site-specific recombination)) of nucleic acid molecules; 2) includes conjoined nucleotide sequences that are not conjoined in nature, 3) has been engineered using molecular cloning techniques such that it lacks one or more nucleotides with respect to the naturally occurring nucleic acid molecule sequence, and/or 4) has been manipulated using molecular cloning techniques such that it has one or more sequence changes or rearrangements with respect to the naturally occurring nucleic acid sequence. As non-limiting examples, a cDNA is a recombinant DNA molecule, as is any nucleic acid molecule that has been generated by *in vitro* polymerase reaction(s), or to which linkers have been attached, or that has been integrated into a vector, such as a cloning vector or expression vector.

[0051] As used herein, a “substantial portion” of a nucleic acid sequence encoding a viral structural polypeptide can comprise enough of the nucleic acid sequence encoding the viral structural polypeptide to afford putative identification of that polypeptide, either by manual evaluation of the sequence by one skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (see, for example, in “Basic Local Alignment Search Tool”; Altschul SF *et al.*, J. Mol. Biol. 215:403-410, 1993). In general, one of skill in the art will recognize that a sequence of ten or more contiguous amino acids or thirty or more nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene specific oligonucleotide probes comprising 15-30 contiguous nucleotides may be used in sequence-dependent methods of gene identification (*e.g.*, Southern hybridization) and isolation (*e.g.*, *in situ* hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 10-15 bases may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a “substantial portion” of a nucleotide sequence comprises enough of the sequence to afford specific identification and/or isolation of a nucleic acid fragment comprising the sequence. The present disclosure provides nucleic acid molecules which are devoid of partial or complete nucleic acid sequences encoding one

or more viral structural polypeptides. The skilled artisan, having the benefit of the sequences as disclosed herein, can readily use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the present application comprises the complete sequences as disclosed herein, *e.g.*, those set forth in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

[0052] As will be understood by one having ordinary skill in the art, for any and all purposes, such as in terms of providing a written description, all ranges disclosed herein also encompass any and all possible sub-ranges and combinations of sub-ranges thereof. Any listed range can be easily recognized as sufficiently describing and enabling the same range being broken down into at least equal halves, thirds, quarters, fifths, tenths, *etc.* As a non-limiting example, each range discussed herein can be readily broken down into a lower third, middle third and upper third, *etc.* As will also be understood by one skilled in the art all language such as “up to,” “at least,” “greater than,” “less than,” and the like include the number recited and refer to ranges which can be subsequently broken down into sub-ranges as discussed above. Finally, as will be understood by one skilled in the art, a range includes each individual member. Thus, for example, a group having 1-3 articles refers to groups having 1, 2, or 3 articles. Similarly, a group having 1-5 articles refers to groups having 1, 2, 3, 4, or 5 articles, and so forth.

[0053] The discussion of the general methods provided herein is intended for illustrative purposes only. Other alternative methods and alternatives will be apparent to those of skill in the art upon review of this application, and are to be included within the spirit and purview of this application.

Alphaviruses

[0054] *Alphavirus* is a genus of genetically, structurally, and serologically related viruses of the group IV *Togaviridae* family which includes at least 30 members, each having single stranded RNA genomes of positive polarity enclosed in a nucleocapsid surrounded by an envelope containing viral spike proteins. Currently, the alphavirus genus comprises among others the Sindbis virus (SIN), the Semliki Forest virus (SFV), the Ross River virus (RRV), Venezuelan equine encephalitis virus (VEEV), and Eastern equine encephalitis virus (EEEV), which are all closely related and are able to infect various vertebrates such as

mammalians, rodents, fish, avian species, and larger mammals such as humans and horses as well as invertebrates such as insects. Transmission between species and individuals occurs mainly via mosquitoes making the alphaviruses a contributor to the collection of Arboviruses – or Arthropod-Borne Viruses. In particular, the Sindbis and the Semliki Forest viruses have been widely studied and the life cycle, mode of replication, *etc.*, of these viruses are well characterized. In particular, alphaviruses have been shown to replicate very efficiently in animal cells which makes them valuable as vectors for production of protein and nucleic acids in such cells.

[0055] Alphavirus particles are enveloped, have a 70 nm diameter, tend to be spherical (although slightly pleomorphic), and have an approximately 40 nm isometric nucleocapsid. **FIG. 8** depicts a typical alphavirus genomic structure and genome expression (adapted from Strauss *et al.*, Microbiological Reviews, pp. 491-562, September 1994). *Alphavirus* genome is single-stranded RNA of positive polarity of approximately 11- 12 kb in length, comprising a 5' cap, a 3' poly-A tail, and two open reading frames with a first frame encoding the nonstructural proteins with enzymatic function and a second frame encoding the viral structural proteins (*e.g.*, the capsid protein C, E1 glycoprotein, E2 glycoprotein, E3 protein and 6K protein).

[0056] The 5' two-thirds of the alphavirus genome encodes a number of nonstructural proteins necessary for transcription and replication of viral RNA. These proteins are translated directly from the RNA and together with cellular proteins form the RNA-dependent RNA polymerase essential for viral genome replication and transcription of subgenomic RNA. Four nonstructural proteins (nsP1-4) are produced as a single polyprotein constitute the virus' replication machinery. The processing of the polyprotein occurs in a highly regulated manner, with cleavage at the P2/3 junction influencing RNA template use during genome replication. This site is located at the base of a narrow cleft and is not readily accessible. Once cleaved, nsP3 creates a ring structure that encircles nsP2. These two proteins have an extensive interface. Mutations in nsP2 that produce noncytopathic viruses or a temperature sensitive phenotypes cluster at the P2/P3 interface region. P3 mutations opposite the location of the nsP2 noncytopathic mutations prevent efficient cleavage of P2/3. This in turn can affect RNA infectivity altering viral RNA production levels.

[0057] The 3' one-third of the genome comprises subgenomic RNA which serves as a template for translation of all the structural proteins required for forming viral particles: the core nucleocapsid protein C, and the envelope proteins P62 and E1 that associate as a heterodimer. The viral membrane-anchored surface glycoproteins are responsible for receptor recognition and entry into target cells through membrane fusion. The subgenomic RNA is transcribed from the p26S subgenomic promoter present at the 3' end of the RNA sequence encoding the nsP4 protein. The proteolytic maturation of P62 into E2 and E3 causes a change in the viral surface. Together the E1, E2, and sometimes E3, glycoprotein "spikes" form an E1/E2 dimer or an E1/E2/E3 trimer, where E2 extends from the center to the vertices, E1 fills the space between the vertices, and E3, if present, is at the distal end of the spike. Upon exposure of the virus to the acidity of the endosome, E1 dissociates from E2 to form an E1 homotrimer, which is necessary for the fusion step to drive the cellular and viral membranes together. The alphaviral glycoprotein E1 is a class II viral fusion protein, which is structurally different from the class I fusion proteins found in influenza virus and HIV. The E2 glycoprotein functions to interact with the nucleocapsid through its cytoplasmic domain, while its ectodomain is responsible for binding a cellular receptor. Most alphaviruses lose the peripheral protein E3, while in Semliki viruses it remains associated with the viral surface.

[0058] Alphavirus replication has been reported to take place in the cytoplasm of the cell. In the first step of the infectious cycle, the 5' end of the genomic RNA is translated into a polyprotein (nsP1-4) with RNA polymerase activity that produces a negative strand complementary to the genomic RNA. In a second step, the negative strand is used as a template for the production of two RNAs, respectively: (1) a positive genomic RNA corresponding to the genome of the secondary viruses producing, by translation, other nsP proteins and acting as a genome for the virus; and (2) subgenomic RNA encoding the structural proteins of the virus forming the infectious particles. The positive genomic RNA/subgenomic RNA ratio is regulated by proteolytic autocleavage of the polyprotein to nsP 1, nsP 2, nsP 3 and nsP 4. In practice, the viral gene expression takes place in two phases. In a first phase, there is main synthesis of positive genomic strands and of negative strands.

During the second phase, the synthesis of subgenomic RNA is virtually exclusive, thus resulting in the production of large amount of structural protein.

[0059] Previous detailed analyses of the 5'-untranslated regions (5'-UTR) of alphaviruses have revealed the absolute importance of the extreme 5' nucleotides to support RNA replication and transcription. In particular, as illustrated in **FIG. 1**, the conservation of an AU dinucleotide at nucleotide positions 1 and 2, respectively, of the 5' UTR sequence is noted among all alphaviruses suggesting the importance of these nucleotides. As used herein, "A1" refers to the conserved A nucleotide at nucleotide position 1 of the 5'-UTR (*e.g.*, an alphavirus 5'-UTR), and "U2" refers to the conserved U nucleotide at nucleotide position 2 of the 5'-UTR (*e.g.*, an alphavirus 5'-UTR). Further, for Venezuelan equine encephalitis virus (VEEV), detailed analysis of the 5' most three nucleotides as well as the stem loop region found immediately following this sequence has been conducted. In particular, the importance of maintaining the U residue at position 2 of the 5' UTR has been determined previously (Kulasegaran-Shylini *et al.*, J. Virol. 83:17 p 8327-8339, 2009a; and Kulasegaran-Shylini *et al.* J. Virol. 83:17 p 8327-8339, 2009b). Specifically, *in vitro* transcribed RNA from a full length infectious clone designated (G2)VEE/SINV containing a single U2→G change in the 5' UTR demonstrated a loss of nearly three logs of infectivity compared to *in vitro* transcribed RNA from a wild type VEE/SINV infectious clone. This report strongly suggests that the U at position 2 is critical to RNA replication and cannot be replaced with a G. However, as described herein in details, a VEEV replicon with a U2→G change in the 5' UTR is, surprisingly and in direct contradiction to this previous report, not only completely capable of robust replication but result in three times the expression potential of a VEEV replicon as compared to a wild-type 5' UTR containing the U residue at position 2.

[0060] The extreme 5' and 3' sequences of most RNA viruses are highly constrained and little if any variation is tolerated; most modifications result in highly crippled or lethal outcomes for RNA replication. Kulasegaran-Shylini *et al.* completed an in-depth analysis of the 5' nucleotide sequences critical to RNA replication for a chimeric VEEV/SINV infectious clone, which is representative of all alphaviruses (Kulasegaran-Shylini *et al.* 2009a, *supra*). This report built on analysis carried out over the course of 25 years by many researchers that clearly supports the restriction in RNA sequence variation that

can occur at the 5' end of any particular alphavirus. The Kulasegaran-Shylini *et al.* 2009b paper (J. Virol. 83:17 p 8327-8339, 2009) specifically states/shows that changing nucleotide 2 in the 5' UTR from a U residue to a G residue (U2→G) significantly reduces the viability of that infectious clone RNA. That is, that specific change in the 5'-UTR reduced biologic activity of the infectious clone RNA by nearly 3 orders of magnitude. As disclosed herein, the change in the 5'-UTR (*e.g.*, a U2→G change) incorporated into a VEEV (strain TC-83) replicon RNA not only does not cripple the replication of the replicon but can actually increase the biological activity of the replicon. For example, the replicon comprising the U2→G substitution can, in some embodiments, leads to the expression of a protein of interest as much as three times more than a wild type replicon expressing the same protein. This result is surprising and the increased biologic activity of the replicon carrying the U2→G change could not have been predicted. This modified replicon has the potential to be a superior RNA expression platform to support both vaccine and therapeutic applications.

[0061] Conservation of the 5' most 2 nucleotides has been observed across all of the genomic RNA of alphavirus subtypes. The conserved AU dinucleotide (A1 and U2) has also been shown to be critically required for RNA replication (Kulasegaran-Shylini *et al.* 2009a and 2009b, *supra*). The demonstration that an alphavirus replicon RNA carrying an AG dinucleotide at the extreme 5' end is not only completely functional but demonstrates enhanced biologic activity is surprising and is completely contrary to the dogma in the field.

[0062] As disclosed herein, monogenic or multigenic alphavirus expression systems can be generated by using a modified replicon RNA having expression/translation enhancing activity such as, for example, a replicon RNA containing a modified 5'-UTR. In some embodiments, the viral (*e.g.*, alphavirus) expression systems as described herein are further devoid of a part or the entire coding region for one or more viral structural proteins. For example, the alphavirus expression system may be devoid of a portion of or the entire coding sequence for one or more of the viral capsid protein C, E1 glycoprotein, E2 glycoprotein, E3 protein and 6K protein. In some embodiments, modification of nucleotide at position 2 in a cDNA copy of the Venezuelan equine encephalitis virus (VEEV) 5' UTR sequence from a thymine (T) nucleotide to a guanine (G) nucleotide (T2→G mutation), in the

context of a replicon RNA, bestows the replicon with significantly higher protein expression potential compared to a VEEV replicon with a wild-type 5' UTR sequence.

[0063] In some embodiments, the level of expression and/or translation enhancement activity of the modified replicon RNAs as disclosed herein is of at least 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2 (2-fold), 3, 4, 5, 6, 7, 8, or more times, relative to the expression level detected from a corresponding unmodified replicon, *e.g.* replicon with a wild-type 5' UTR. Without being limited by any particular theory, enhanced translation can be due to an enhancement of transcription which results in an increased level of transcripts being available for translation and/or can be independent of transcription and be due to for example enhanced ribosome binding. The level of enhancement activity can be measured by any convenient methods and techniques known in the art including, but are not limited to, transcript level, amount of protein, protein activity, etc. (*see e.g.*, Examples 1, 3-5 below).

Nucleic Acid Molecules

[0064] In one aspect, novel nucleic acid molecules which include a modified replicon RNA are disclosed herein. For example, a modified replicon RNA can comprise mutation(s), deletion(s), substitution(s), and/or insertion(s) in one or more of the original genomic regions (*e.g.*, open reading frames (ORFs) and/or non-coding regions (*e.g.*, promoter sequences)) of the parent replicon RNA. In some embodiments, the modified replicon RNA includes a modified 5'-untranslated region (5'-UTR). In some embodiments, the modified 5'-UTR includes one or more nucleotide substitutions at position 1, 2, 4, or a combination thereof. In some embodiments, at least one of the nucleotide substitutions is a nucleotide substitution at position 1 of the modified 5'-UTR. In some embodiments, at least one of the nucleotide substitutions is a nucleotide substitution at position 2 of the modified 5'-UTR. In some embodiments, at least one of the nucleotide substitutions is a nucleotide substitution at position 4 of the modified 5'-UTR. In some embodiments, the nucleotide substitutions at position 2 of the modified 5'-UTR is a U→G substitution. In some embodiments, the nucleotide substitution at position 2 of the modified 5'-UTR is a U→A substitution. In some embodiments, the nucleotide substitution at position 2 of the modified 5'-UTR is a U→C substitution.

[0065] As used herein, the terms “replicon RNA” refers to RNA which contains all of the genetic information required for directing its own amplification or self-replication within a permissive cell. To direct its own replication, the RNA molecule 1) encodes polymerase, replicase, or other proteins which may interact with viral or host cell-derived proteins, nucleic acids or ribonucleoproteins to catalyze the RNA amplification process; and 2) contain *cis*-acting RNA sequences required for replication and transcription of the subgenomic replicon-encoded RNA. These sequences may be bound during the process of replication to its self-encoded proteins, or non-self-encoded cell-derived proteins, nucleic acids or ribonucleoproteins, or complexes between any of these components. For the purpose of the present disclosure, an alphavirus-derived replicon RNA molecule typically contains the following ordered elements: 5' viral or defective-interfering RNA sequence(s) required in *cis* for replication, sequences coding for biologically active alphavirus nonstructural proteins (*e.g.*, nsP1, nsP2, nsP3, and nsP4), promoter for the subgenomic RNA, 3' viral sequences required in *cis* for replication, and a polyadenylate tract. Further, the term replicon RNA generally refers to a molecule of positive polarity, or “message” sense, and the replicon RNA may be of length different from that of any known, naturally-occurring alphavirus. In some embodiments of the present disclosure, the replicon RNA does not contain the sequences of at least one of structural viral protein; sequences encoding structural genes can be substituted with heterologous sequences. In those instances, where the replicon RNA is to be packaged into a recombinant alphavirus particle, it must contain one or more sequences, so-called packaging signals, which serve to initiate interactions with alphavirus structural proteins that lead to particle formation.

[0066] As used herein, “subgenomic RNA” refers to a RNA molecule of a length or size which is smaller than the genomic RNA from which it was derived. The alphavirus subgenomic RNA should be transcribed from an internal promoter, whose sequences reside within the genomic RNA or its complement. Transcription of an alphavirus subgenomic RNA may be mediated by viral-encoded polymerase(s) associated with host cell-encoded proteins, ribonucleoprotein(s), or a combination thereof. In some embodiments of the present disclosure, the subgenomic RNA is produced from a modified replicon RNA as disclosed herein and encodes or expresses one or more gene of interest (GOI). Instead of the native

subgenomic promoter, the subgenomic RNA can be placed under control of internal ribosome entry site (IRES) derived from encephalomyocarditis viruses (EMCV), Bovine Viral Diarrhea Viruses (BVDV), polioviruses, Foot-and-mouth disease viruses (FMD), enterovirus 71, or hepatitis C viruses.

[0067] Accordingly, in some embodiments, a part or the entire coding sequence for one or more viral structural proteins are absent and/or modified in the nucleic acid molecules disclosed herein. Thus, in some particular embodiments, the modified replicon RNA as disclosed herein includes a modified 5'-UTR and is devoid of at least a portion of a nucleic acid sequence encoding one or more viral structural proteins, for example, devoid of the first one, two, three, four, five, six, seven, eight, nine, ten, or more nucleotides of the nucleic acid sequence encoding the viral structural proteins. In some embodiments, the modified alphavirus genome or replicon RNA can be devoid of about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or more of the sequence encoding one or more of the structural polypeptides E1, E2, E3, 6K, and capsid protein C. In some embodiments, the modified alphavirus genome or replicon RNA is devoid of a substantial portion of or the entire sequence encoding one of or more of the structural polypeptides E1, E2, E3, 6K, and capsid protein C. As used herein, a "substantial portion" of a nucleic acid sequence encoding a viral structural protein comprises enough of the nucleic acid sequence encoding the viral structural protein to afford putative identification of that protein, either by manual evaluation of the sequence by one skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (see, for example, Altschul SF *et al.* 1993, *supra*). In some embodiments, the modified alphavirus genome or replicon RNA is devoid of the entire sequence encoding one or more of the structural polypeptides E1, E2, E3, 6K, and capsid protein C.

[0068] In some particular embodiments of the application, the nucleic acid molecule as disclosed herein includes a modified alphavirus genome or replicon RNA including a modified alphavirus genome or replicon RNA, wherein the nucleic acid molecule exhibits at least 80% sequence identity to the nucleic acid sequence of SEQ ID NO: 1, wherein the modified alphavirus genome or replicon RNA comprises a U→G substitution at position 2 of the 5'-untranslated region (5'-UTR) and is devoid of at least a portion of the

sequence encoding viral structural proteins. In some embodiments, the nucleic acid molecule exhibits at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the nucleic acid sequence of SEQ ID NO: 1. In some embodiments, the nucleic acid molecule exhibits 100% sequence identity to the nucleic acid sequence of SEQ ID NO: 1.

[0069] In some embodiments, the nucleic acid molecule as disclosed herein includes a modified alphavirus genome or replicon RNA, wherein the modified alphavirus genome or replicon RNA comprises a 5'-UTR exhibiting at least 80% sequence identity to the nucleic acid sequence of at least one 5'-UTR disclosed herein and a U→G substitution at position 2 of the 5'-UTR, and wherein the modified alphavirus genome or replicon RNA is devoid of at least a portion of the sequence encoding viral structural proteins. In some embodiments, the modified alphavirus genome or replicon RNA comprises a 5'-UTR exhibiting at least 80% sequence identity to at least one of the sequences set forth in SEQ ID NOS: 2-18. In some embodiments, the modified alphavirus genome or replicon RNA comprises a 5'-UTR exhibiting at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to at least one of the sequences set forth in SEQ ID NOS: 2-18. In some embodiments, the modified alphavirus genome or replicon RNA comprises a 5'-UTR exhibiting 100% sequence identity to at least one of the sequences set forth in SEQ ID NOS: 2-18 of the Sequence Listing.

[0070] Nucleic acid molecules having a high degree of sequence identity (*e.g.*, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% to the nucleic acid sequence of a 5'UTR disclosed herein can be identified and/or isolated by using the sequences identified herein (*e.g.*, SEQ ID NOS: 1-18) or any others alphavirus 5'UTR as they are known in the art, for example, the sequences having GenBank/NCBI accession numbers J02363, NC_001547, U38305, L04653, NC_001449, U38304, X04129, NC_003215, and the TR339 genomic sequence (Klimstra *et al.*, J. Virol. 72:7357, 1988; McKnight *et al.*, J. Virol. 70:1981, 1996), by genome sequence analysis, hybridization, and/or PCR with degenerate primers or gene-specific primers from sequences identified in the respective alphavirus genome. As used herein "sequence identity" refers to the extent to which two optimally aligned polynucleotide are invariant throughout a window

of alignment of components, *e.g.*, nucleotides. An "identity fraction" for aligned segments of a test sequence and a reference sequence is the number of identical components which are shared by the two aligned sequences divided by the total number of components in reference sequence segment, *e.g.*, the entire reference sequence or a smaller defined part of the reference sequence.

[0071] Some embodiments disclosed herein relate to nucleic acid molecules comprising a modified replicon RNA, wherein the modified replicon RNA comprises a modified alphavirus base sequence such as, *e.g.* a 5'UTR, having translation enhancing activity. Such modified replicon RNAs can be used to achieve enhanced levels of expression of a heterologous nucleic acid sequence (*e.g.*, DNA or cDNA) coding for a desired product. In some embodiments, the modified replicon RNAs are used to achieve enhanced levels of expression of a heterologous nucleic acid sequence (*e.g.*, DNA or cDNA) coding for a desired product after introduction of the modified replicons in a cell which can be, for example, a cell in cell culture or can be a cell in a living body.

[0072] In addition, in some embodiments, the nucleic acid molecules can include a modified alphavirus genome or replicon RNA containing one or more attenuating mutations so as to increase the safety of virus manipulation and/or administration. The phrase "attenuating mutation" as used herein means a nucleotide mutation or an amino acid encoded in view of such mutation which result in a decreased probability of causing disease in its host (*i.e.*, a loss of virulence), in accordance with standard terminology in the art, whether the mutation be a substitution mutation or an in-frame deletion or insertion mutation. Attenuating mutations may be in the coding or non-coding regions (*e.g.* 5'UTR) of the alphavirus genome. As known by those skilled in the art, the phrase "attenuating mutation" excludes mutations or combinations of mutations which would be lethal to the virus. Further, those skilled in the art will appreciate that some attenuating mutations may be lethal in the absence of a second-site suppressor mutation(s).

[0073] The molecular techniques and methods by which these new nucleic acid molecules were constructed and characterized are described more fully in the Examples herein of the present application. In the Examples section, the Venezuelan equine

encephalitis virus (VEEV) has been used to illustrate the compositions and methods disclosed herein.

[0074] In some embodiments, the nucleic acid molecules are recombinant nucleic acid molecules. As used herein, the term recombinant means any molecule (*e.g.* DNA, RNA, polypeptide), that is, or results, however indirect, from human manipulation. As non-limiting examples, a cDNA is a recombinant DNA molecule, as is any nucleic acid molecule that has been generated by *in vitro* polymerase reaction(s), or to which linkers have been attached, or that has been integrated into a vector, such as a cloning vector or expression vector. As non-limiting examples, a recombinant nucleic acid molecule: 1) has been synthesized or modified *in vitro*, for example, using chemical or enzymatic techniques (for example, by use of chemical nucleic acid synthesis, or by use of enzymes for the replication, polymerization, exonucleolytic digestion, endonucleolytic digestion, ligation, reverse transcription, transcription, base modification (including, *e.g.*, methylation), or recombination (including homologous and site-specific recombination) of nucleic acid molecules; 2) includes conjoined nucleotide sequences that are not conjoined in nature; 3) has been engineered using molecular cloning techniques such that it lacks one or more nucleotides with respect to the naturally occurring nucleotide sequence; and/or 4) has been manipulated using molecular cloning techniques such that it has one or more sequence changes or rearrangements with respect to the naturally occurring nucleotide sequence.

[0075] In some embodiments, the nucleic acid molecules disclosed herein are produced using recombinant DNA technology (*e.g.*, polymerase chain reaction (PCR) amplification, cloning, *etc.*) or chemical synthesis. Nucleic acid molecules as disclosed herein include natural nucleic acid molecules and homologs thereof, including, but not limited to, natural allelic variants and modified nucleic acid molecules in which one or more nucleotide residues have been inserted, deleted, and/or substituted, in such a manner that such modifications provide the desired property in effecting a biological activity as described herein.

[0076] A nucleic acid molecule, including a variant of a naturally-occurring nucleic acid sequence, can be produced using a number of methods known to those skilled in the art (see, for example, Sambrook *et al.*, In: Molecular Cloning, A Laboratory Manual, 2nd

Edition, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (1989)). The sequence of a nucleic acid molecule can be modified with respect to a naturally-occurring sequence from which it is derived using a variety of techniques including, but not limited to, classic mutagenesis techniques and recombinant DNA techniques, such as but not limited to site-directed mutagenesis, chemical treatment of a nucleic acid molecule to induce mutations, restriction enzyme cleavage of a nucleic acid fragment, ligation of nucleic acid fragments, PCR amplification and/or mutagenesis of selected regions of a nucleic acid sequence, recombinational cloning, and chemical synthesis, including chemical synthesis of oligonucleotide mixtures and ligation of mixture groups to "build" a mixture of nucleic acid molecules, and combinations thereof. Nucleic acid molecule homologs can be selected from a mixture of modified nucleic acid molecules by screening for the function of the protein or the replicon encoded by the nucleic acid molecule and/or by hybridization with a wild-type gene or fragment thereof, or by PCR using primers having homology to a target or wild-type nucleic acid molecule or sequence.

[0077] In various embodiments disclosed herein, the nucleic acid molecule disclosed herein can include one or more of the following feature. In some embodiments, the modified replicon RNA is a modified alphavirus replicon RNA. In some embodiments, the modified alphavirus replicon RNA includes a modified alphavirus genome. In some embodiments, the modified 5'-UTR includes one or more nucleotide substitutions at position 1, 2, 4, or a combination thereof. In certain embodiments, at least one of the nucleotide substitutions is a nucleotide substitution at position 2 of the modified 5'-UTR. In some particular embodiments, the nucleotide substitutions at position 2 of the modified 5'-UTR is a U→G substitution.

[0078] In some embodiments disclosed herein, the modified alphavirus genome or replicon RNA is operably linked to a heterologous regulatory element. As used herein, "regulatory element", "regulatory sequence", or "regulatory element sequence" refers to a nucleotide sequence located upstream (5'), within, or downstream (3') of a coding sequence such as, for example, a polypeptide-encoding sequence or a functional RNA-encoding sequence. Transcription of the coding sequence and/or translation of an RNA molecule resulting from transcription of the coding sequence are typically affected by the presence or

absence of the regulatory element. These regulatory elements may comprise promoters, *cis*-elements, enhancers, terminators, or introns. One of skill in the art will appreciate that the regulatory elements described herein may be present in a chimeric or hybrid regulatory expression element. In some embodiments, the heterologous regulatory element is, or comprises, a promoter sequence. The heterologous promoter sequence can be any heterologous promoter sequence, for example, a SP6 promoter, a T3 promoter, or a T7 promoter, or a combination thereof. In some particular embodiments, the promoter sequence includes a T7 promoter sequence.

[0079] Further, in some embodiments, the modified alphavirus genome or replicon RNA can include one or more heterologous transcriptional termination signal sequences. The term “transcriptional termination signal”, “terminator” or “terminator sequence” or “transcription terminator”, as used interchangeably herein, refers to a regulatory section of genetic sequence that causes RNA polymerase to cease transcription. The heterologous transcriptional termination signal sequences can generally be any heterologous transcriptional termination signal sequences, and for example, SP6 termination signal sequence, a T3 termination signal sequence, a T7 termination signal sequence, or a variant thereof. Accordingly, the nucleic acid molecules according to some embodiments of the disclosure can include at least one of the one or more heterologous transcriptional termination signal sequences selected from the group consisting of a SP6 termination signal sequence, a T3 termination signal sequence, a T7 termination signal sequence, or a variant thereof. In some particular embodiments, the transcriptional termination sequence includes a T7 termination signal sequence.

[0080] In some embodiments, the nucleic acid molecules disclosed herein can include one or more expression cassettes. In principle, the nucleic acid molecules disclosed herein can generally include any number of expression cassettes. In some particular embodiments, the nucleic acid molecules disclosed herein can include at least two, at least three, at least four, at least five, or at least six expression cassettes. As used herein, the term “expression cassette” refers to a construct of genetic material that contains coding sequences and enough regulatory information to direct proper transcription and/or translation of the coding sequences in a recipient cell, *in vivo* and/or *ex vivo*. The expression cassette may be

inserted into a vector for targeting to a desired host cell and/or into a subject. Further, the term expression cassette may be used interchangeably with the term "expression construct". In some embodiments, the term "expression cassette" refers to a nucleic acid construct that includes a gene encoding a protein or functional RNA operably linked to regulatory elements such as, for example, a promoter and/or a termination signal, and optionally, any or a combination of other nucleic acid sequences that affect the transcription or translation of the gene.

[0081] The term "operably linked", as used herein, denotes a functional linkage between two or more sequences. For example, an operably linkage between a polynucleotide of interest and a regulatory sequence (for example, a promoter) is functional link that allows for expression of the polynucleotide of interest. In this sense, the term "operably linked" refers to the positioning of a regulatory region and a coding sequence to be transcribed so that the regulatory region is effective for regulating transcription or translation of the coding sequence of interest. In some embodiments disclosed herein, the term "operably linked" denotes a configuration in which a regulatory sequence is placed at an appropriate position relative to a sequence that encodes a polypeptide or functional RNA such that the control sequence directs or regulates the expression or cellular localization of the mRNA encoding the polypeptide, the polypeptide, and/or the functional RNA. Thus, a promoter is in operable linkage with a nucleic acid sequence if it can mediate transcription of the nucleic acid sequence. Operably linked elements may be contiguous or non-contiguous.

[0082] The basic techniques for operably linking two or more sequences of DNA together are familiar to the skilled worker, and such methods have been described in a number of texts for standard molecular biological manipulation (see, for example, Maniatis *et al.*, "Molecular Cloning: A Laboratory Manual" 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; and Gibson *et al.*, Nature Methods 6:343-45, 2009).

[0083] Accordingly, the nucleic acid molecules as provided herein can find use, for example, as an expression vector that, when including a regulatory element operably linked to a heterologous nucleic acid sequence, can affect expression of the heterologous nucleic acid sequence. In some embodiments, the heterologous nucleotide sequence includes a coding sequence of a gene of interest (GOI). In some embodiments, the coding sequence of

the GOI is optimized for expression at a level higher than the expression level of a reference coding sequence. In some embodiments, the reference coding sequence is a sequence that has not been optimized. In some embodiments, the optimization of the GOI coding sequence can include sequence optimization. With respect to sequence-optimization of nucleotide sequences, degeneracy of the genetic code provides the possibility to substitute at least one base of the protein encoding sequence of a gene with a different base without causing the amino acid sequence of the polypeptide produced from the gene to be changed. Hence, the nucleic acid molecules of the present application may also have any base sequence that has been changed from any polynucleotide sequence disclosed herein by substitution in accordance with degeneracy of the genetic code. References describing codon usage are readily publicly available. In some embodiments, polynucleotide sequence variants can be produced for a variety of reasons, *e.g.*, to optimize expression for a particular host (*e.g.*, changing codon usage in the alphavirus mRNA to those preferred by other organisms such as human, hamster, mice, or monkey).

[0084] The polypeptide encoded by a GOI can generally be any polypeptide, and can be, for example a therapeutic polypeptide, a prophylactic polypeptide, a diagnostic polypeptide, a neutraceutical polypeptide, or an industrial enzyme. In some embodiments, the GOI encodes a polypeptide selected from the group consisting of an antibody, an antigen, an immune modulator, and a cytokine.

[0085] In some embodiments, the coding sequence of the GOI is optimized for a desired property. In some embodiments, the coding sequence of the GOI is optimized for expression at a level higher than the expression level of a reference coding sequence.

[0086] In some embodiments, the modified genome or replicon RNA disclosed herein is a genome or replicon RNA of an alphavirus, such as a genome or replicon RNA of a viral species belonging to the *Alphavirus* genus of the *Togaviridae* family. In some embodiments, the modified genome or replicon RNA is of an alphavirus belonging to the VEEV/EEEV group, or the SF group, or the SIN group (for review, see, *e.g.* Strauss and Strauss. *Microbiol. Rev.* 58:3 p 492-562, 1994). Non-limiting examples of SF group alphaviruses include Semliki Forest virus, O'Nyong-Nyong virus, Ross River virus, Middelburg virus, Chikungunya virus, Barmah Forest virus, Getah virus, Mayaro virus,

Sagiyama virus, Bebaru virus, and Una virus. Non-limiting examples of SIN group alphaviruses include Sindbis virus, Girdwood S.A. virus, South African Arbovirus No. 86, Ockelbo virus, Aura virus, Babanki virus, Whataroa virus, and Kyzylagach virus. Non-limiting examples of VEEV/EEEV group alphaviruses include Eastern equine encephalitis virus (EEEV), Venezuelan equine encephalitis virus (VEEV), Everglades virus (EVEV), Mucambo virus (MUCV), Pixuna virus (PIXV), Middleburg virus (MIDV), Chikungunya virus (CHIKV), O’Nyong-Nyong virus (ONNV), Ross River virus (RRV), Barmah Forest virus (BF), Getah virus (GET), Sagiyama virus (SAGV), Bebaru virus (BEBV), Mayaro virus (MAYV), and Una virus (UNAV).

[0087] Non-limiting examples of alphavirus species includes Eastern equine encephalitis virus (EEEV), Venezuelan equine encephalitis virus (VEEV), Everglades virus (EVEV), Mucambo virus (MUCV), Pixuna virus (PIXV), Middleburg virus (MIDV), Chikungunya virus (CHIKV), O’Nyong-Nyong virus (ONNV), Ross River virus (RRV), Barmah Forest virus (BF), Getah virus (GET), Sagiyama virus (SAGV), Bebaru virus (BEBV), Mayaro virus (MAYV), Una virus (UNAV), Sindbis virus (SINV), Aura virus (AURAV), Whataroa virus (WHAV), Babanki virus (BABV), Kyzylagach virus (KYZV), Western equine encephalitis virus (WEEV), Highland J virus (HJV), Fort Morgan virus (FMV), Ndumu (NDUV), and Buggy Creek virus. Virulent and avirulent alphavirus strains are both suitable. In some particular embodiments, the modified genome or replicon RNA is of a Sindbis virus (SIN), a Semliki Forest virus (SFV), a Ross River virus (RRV), a Venezuelan equine encephalitis virus (VEEV), or an Eastern equine encephalitis virus (EEEV). In some embodiments, the modified genome or replicon RNA is of a Venezuelan equine encephalitis virus (VEEV).

Recombinant Cells

[0088] In one aspect, some embodiments disclosed herein relate to a method of transforming a cell that includes introducing into a host cell, such as an animal cell, a nucleic acid molecule as provided herein, and selecting or screening for a transformed cell. The terms “host cell” and “recombinant host cell” are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding

generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein. In some embodiments, the nucleic acid molecule is introduced into a host cell by an electroporation procedure or a biolistic procedure.

[0089] In a related aspect, some embodiments relate to recombinant host cells, for example, recombinant animal cells that include a nucleic acid molecule described herein. The nucleic acid molecule can be stably integrated in the host genome, or can be episomally replicating, or present in the recombinant host cell as a mini-circle expression vector for a stable or transient expression. Accordingly, in some embodiments disclosed herein, the nucleic acid molecule is maintained and replicated in the recombinant host cell as an episomal unit. In some embodiments, the nucleic acid molecule is stably integrated into the genome of the recombinant cell. Stable integration can be completed using classical random genomic recombination techniques or with more precise genome editing techniques such as using guide RNA directed CRISPR/Cas9, or DNA-guided endonuclease genome editing NgAgo (*Natronobacterium gregoryi* Argonaute), or TALEN genome editing (transcription activator-like effector nucleases). In some embodiments, the nucleic acid molecule present in the recombinant host cell as a mini-circle expression vector for a stable or transient expression.

[0090] In some embodiments, host cells can be genetically engineered (*e.g.* transduced or transformed or transfected) with, for example, a vector construct of the present application that can be, for example, a vector for homologous recombination that includes nucleic acid sequences homologous to a portion of the genome of the host cell, or can be an expression vector for the expression of any or a combination of the genes of interest. The vector can be, for example, in the form of a plasmid, a viral particle, a phage, *etc.* In some embodiments, a vector for expression of a polypeptide of interest can also be designed for integration into the host, *e.g.*, by homologous recombination. The vector containing a polynucleotide sequence as described herein, *e.g.*, nucleic acid molecule comprising a modified alphavirus genome or replicon RNA, as well as, optionally, a selectable marker or reporter gene, can be employed to transform an appropriate host cell.

[0091] The methods and compositions disclosed herein may be deployed for genetic engineering of any species, including, but not limited to, prokaryotic and eukaryotic species. Suitable host cells to be modified using the compositions and methods according to the present disclosure can include, but not limited to, algal cells, bacterial cells, heterokonts, fungal cells, chytrid cells, microfungi, microalgae, and animal cells. In some embodiments, the animal cells are invertebrate animal cells. In some embodiments, the vertebrate animal cells are mammalian cells. Host cells can be either untransformed cells or cells that have already been transfected with at least one nucleic acid molecule.

[0092] The methods and compositions disclosed herein can be used, for example, with subject and/or host cells that are important or interesting for aquaculture, agriculture, animal husbandry, and/or for therapeutic and medicinal applications, including production of polypeptides used in the manufacturing of vaccine, pharmaceutical products, industrial products, chemicals, and the like. In some embodiments, the compositions and methods disclosed herein can be used with host cells from species that are natural hosts of alphaviruses, such as rodents, mice, fish, birds, and larger mammals such as humans, horses, pig, monkey, and apes as well as invertebrates. Particularly preferred species, in some embodiments of the application, are vertebrate animal species and invertebrate animal species. In principle, any animal species can be generally used and can be, for example, human, dog, bird, fish, horse, pig, primate, mouse, cattle, swine, sheep, rabbit, cat, goat, donkey, hamster, or buffalo. Non-limiting examples of suitable bird species include chicken, duck, goose, turkey, ostrich, emu, swan, peafowl, pheasant, partridge, and guinea fowl. In some particular embodiments, the fish species is a salmon species. Primary mammalian cells and continuous/immortalized cells types are also suitable. Non-limiting examples of suitable animal host cells include, but not limited to, pulmonary equine artery endothelial cell, equine dermis cell, baby hamster kidney (BHK) cell, rabbit kidney cell, mouse muscle cell, mouse connective tissue cell, human cervix cell, human epidermoid larynx cell, Chinese hamster ovary cell (CHO), human HEK-293 cell, mouse 3T3 cell, Vero cell, Madin-Darby Canine Kidney Epithelial Cell (MDCK), primary chicken fibroblast cell, a HuT78 cell, A549 lung cell, HeLa cell, PER.C6® cell, WI-38 cell, MRC-5 cell, FRhL-2, and CEM T-cell. In some

embodiments, the host cell is baby hamster kidney cell. In some embodiments, the baby hamster kidney cell is a BHK-21 cell.

[0093] Techniques for transforming a wide variety of the above-mentioned host cells and species are known in the art and described in the technical and scientific literature. Accordingly, cell cultures including at least one recombinant cell as disclosed herein are also within the scope of this application. Methods and systems suitable for generating and maintaining cell cultures are known in the art.

Methods for Producing Polypeptides

[0094] The host cells of the present disclosure, such as a prokaryotic or eukaryotic host cell, can be used to produce (*i.e.*, express) a molecule of interest such as, *e.g.*, a polypeptide, encoded in an open reading frame of a gene of interest (GOI) as disclosed herein. Thus, the present application further provides methods for producing a molecule of interest such as, *e.g.*, a polypeptide, using the host cells of the disclosure, which can be, for example, cells in cell culture or can be cells in a living body.

[0095] Accordingly, some embodiments disclosed herein provides methods for producing a polypeptide of interest in a host cell. Such method includes the cultivation of a recombinant host cell, including a nucleic acid molecule according to any one of the preceding aspects and embodiments. In some embodiments, the methods includes culturing the host cell of invention (into which a recombinant expression vector encoding the molecule of interest has been introduced) in a suitable medium such that the molecule of interest is produced. In some embodiments, the methods further include isolating the molecule of interest from the medium or the host cell.

[0096] In another aspect, some embodiments relate to methods for producing a polypeptide of interest in a subject, including administering to the subject a nucleic acid molecule according to any one of the preceding aspects and embodiments.

[0097] Suitable host cells and/or subjects for use in the methods and compositions disclosed herein include, but are not limited to, prokaryotic and eukaryotic species. Suitable host cells to be modified using the compositions and methods according to the present disclosure can include, but not limited to, algal cells, bacterial cells, heterokonts, fungal cells, chytrid cells, microfungi, microalgae, and animal cells. In some embodiments, the animal

cells are invertebrate animal cells. In some embodiments, the vertebrate animal cells are mammalian cells. Host cells can be either untransformed cells or cells that have already been transfected with at least one nucleic acid molecule. Accordingly, biological samples, biomass, and progeny of a recombinant cell according to any one of the preceding aspects and embodiments are also within the scope of the present application. Thus, as discussed in more detail below, polypeptides produced by a method according to this aspect of the application are also within the scope of this application.

[0098] In some embodiments, the recombinant cell is an animal cell. Therapeutic protein production in small and large scale is an important field of development in pharmaceutical industry, because proteins produced in animal cells are believed to generally have proper processing, post-translational modification and therefore have adequate activity for treatment of the physiological condition. In principle, any animal species can be generally used and can be, for example, human, dog, bird, fish, horse, pig, primate, mouse, cattle, swine, sheep, rabbit, cat, goat, donkey, hamster, or buffalo. Non-limiting examples of suitable bird species include chicken, duck, goose, turkey, ostrich, emu, swan, peafowl, pheasant, partridge, and guinea fowl. In some particular embodiments, the fish species is a salmon species. Primary mammalian cells and continuous/immortalized cells types are also suitable. Non-limiting examples of suitable animal host cells include, but not limited to, pulmonary equine artery endothelial cell, equine dermis cell, baby hamster kidney (BHK) cell, rabbit kidney cell, mouse muscle cell, mouse connective tissue cell, human cervix cell, human epidermoid larynx cell, Chinese hamster ovary cell (CHO), human HEK-293 cell, mouse 3T3 cell, Vero cell, Madin-Darby Canine Kidney Epithelial Cell (MDCK), primary chicken fibroblast cell, a HuT78 cell, A549 lung cell, HeLa cell, PER.C6® cell, WI-38 cell, MRC-5 cell, FRhL-2, and CEM T-cell. In some embodiments, the host cell is baby hamster kidney cell. In some embodiments, the baby hamster kidney cell is a BHK-21 cell.

Recombinant Polypeptides

[0099] Some embodiments disclosed herein relate to recombinant polypeptides produced by a method in accordance with one or more embodiments described herein. The recombinant polypeptides of the present application generally can be any recombinant polypeptides and can be, for example, one or more of therapeutic polypeptides, prophylactic

polypeptides, diagnostic polypeptides, neutraceutical polypeptides, industrial enzymes, and reporter polypeptides. In some embodiments, the recombinant polypeptides can be one or more of antibodies, antigens, immune modulators, and cytokines. In some embodiments, the polypeptide of interest may have therapeutic or prophylactic activity.

Compositions

[0100] Some embodiments disclosed herein relate to a composition comprising any of the recombinant polypeptides described herein. The composition can be, for example, a neutraceutical composition, a prophylactic composition, a pharmaceutical composition comprising a pharmaceutically acceptable carrier, or a mixture thereof. In some embodiments, the compositions of the present application can be used as a vaccine.

[0101] Some embodiments disclosed herein relate to a composition including any of the nucleic acid molecules described herein. The composition can be, for example, a neutraceutical composition, a prophylactic composition, a pharmaceutical composition comprising a pharmaceutically acceptable carrier, or a mixture thereof. In some embodiments, the compositions of the present application can be used as a vaccine.

[0102] Some embodiments disclosed herein relate to a composition including any of the recombinant cells described herein. The composition can be, for example, a neutraceutical composition, a prophylactic composition, a pharmaceutical composition comprising a pharmaceutically acceptable carrier, or a mixture thereof. In some embodiments, the compositions of the present application can be used as a vaccine.

[0103] As used herein, the term “pharmaceutically-acceptable carrier” means a carrier that is useful in preparing a pharmaceutical composition or formulation that is generally safe, non-toxic, and neither biologically nor otherwise undesirable, and includes a carrier that is acceptable for veterinary use as well as human pharmaceutical use. In some embodiments, a pharmaceutically acceptable carrier as simple as water, but it can also include, for example, a solution of physiological salt concentration. In some embodiments, a pharmaceutically acceptable carrier can be, or may include, stabilizers, diluents and buffers. Suitable stabilizers are for example SPGA, carbohydrates (such as dried milk, serum albumin or casein) or degradation products thereof. Suitable buffers are for example alkali metal phosphates. Diluents include water, aqueous buffers (such as buffered saline), alcohols and

polyols (such as glycerol). For administration to animals or humans, the composition according to the present application can be given *inter alia* intranasally, by spraying, intradermally, subcutaneously, orally, by aerosol or intramuscularly.

[0104] All publications and patent applications mentioned in this disclosure are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

[0105] No admission is made that any reference cited herein constitutes prior art. The discussion of the references states what their authors assert, and the inventors reserve the right to challenge the accuracy and pertinence of the cited documents. It will be clearly understood that, although a number of information sources, including scientific journal articles, patent documents, and textbooks, are referred to herein; this reference does not constitute an admission that any of these documents forms part of the common general knowledge in the art.

[0106] The discussion of the general methods given herein is intended for illustrative purposes only. Other alternative methods and alternatives will be apparent to those of skill in the art upon review of this disclosure, and are to be included within the spirit and purview of this application.

EXAMPLES

[0107] Additional alternatives are disclosed in further detail in the following examples, which are not in any way intended to limit the scope of the claims.

EXAMPLE 1

General Experimental Procedures

In vitro transcription

[0108] *In vitro* transcription (IVT) reactions were performed using 1µg of DNA template prepared as described above, in a 20 µl reaction over a one hour incubation at 37°C (NEB cat. no. E2065S). 1 Unit of DNase I, provided by the supplier was then added directly to the IVT reaction, and incubated at 37°C for an additional 15 mins. Reactions were then placed on ice, and purified using the manufactures suggested method (Qiagen Cat. no. 74104). Purified RNA was then quantified using a NanoDrop 2000c UV-Vis

Spectrophotometer. RNA was visualized by electrophoresis through 0.8% Agarose gels (Life Technologies Cat. no. G5018-08) and compared with Millennium RNA Marker (Ambion Cat. No. AM7150), prior to proceeding with electroporation.

[0109] Plasmid DNA templates were purified (Qiagen Cat. no. 12163) from 300 mL of saturated *E. coli* TransformMax Epi300 (Epicentre Cat. no. EC300105) cultures grown in LB broth media (Teknova Cat. no. L8000 06) supplemented with 50ng/ml carbamicillin (Teknova Cat. no. NC9730116). Plasmid DNA was linearized by *Not*-I digestion (New England Biolabs NEB cat. no. R3189S) for one hour at 37°C. Linearized template DNA was then re-purified (Zymo Cat. no. D4003), and analyzed by 0.8% agarose gel (Life Technologies Cat. no. G5018-08) against a commercial 2-log DNA ladder (New England Biolabs, NEB Cat. no. N3200S). The presence of a single band was confirmed in each sample, corresponding to the expected fragment size of the linear DNA template, prior to proceeding with *in vitro* transcription.

Transfection and analysis

[0110] In a typical cell transfection experiment, replicon RNA was introduced into BHK-21 cells by electroporation using the SF Cell Line Nucleofector™ kit for the 4D-Nucleofector™ System (Lonza). BHK-21 cells were harvested using 0.25% trypsin and washed once with cold PBS. Cells were resuspended in SF Buffer at a cell density of 1×10^6 cells per 20 μ L electroporation reaction. Three micrograms of RNA was electroporated into cells in triplicate in a 16-well cuvette strip and incubated at room temperature for 10 minutes. Electroporated cells were recovered into plates containing Dulbecco's Modified Eagle Medium containing 10% fetal bovine serum, followed by incubation for 16 – 18 h at standard cell culture conditions.

[0111] Intracellular analyses of replicon transfection efficiency and protein production were performed by flow cytometry. Transfected BHK-21 cells were fixed and permeabilized using fix/perm concentrate and permeabilization buffer (eBioscience). Cells were incubated with antibodies for double-stranded RNA production (J2 anti-dsRNA IgG2A monoclonal antibody, English & Scientific Company) conjugated with R-Phycoerythrin (Innova Biosciences). Antigen production was assessed by additional incubation with antigen-specific antibodies conjugated with PE-Cy5 (Innova Biosciences) (*e.g.* antibodies for

green Renilla, red Firefly, HA, or RSV-F0 (Abcam)). Cells were then washed once and analyzed using a FACS Aria™ Fusion Cell Sorter (BD Biosciences) or FACS Aria™ II Cell Sorter (BD Biosciences). Transfected BHK-21 cells stained with single colors for compensation controls were run prior to sample collection. Data was collected using FACSDiva (BD Biosciences) and further analyzed using FlowJo software. Initial gating was performed to exclude dead cells and debris using forward and side scatter plots. Further gating was conducted to identify cell populations that were positive for both dsRNA (R-PE-positive) and protein expression (PE-Cy5-positive or FITC-positive for GFP expression). Frequencies and mean fluorescence intensities were collected and utilized for construct comparison and optimization.

EXAMPLE 2

5'-UTR sequence modifications

[0112] This Example describes the results of experiments where 5'-UTR sequences were modified to enhance expression of the genes of interest encoded in a VEEV replicon RNA vector. Site directed mutagenesis (SDM) was carried out on a plasmid containing a cDNA copy of a VEEV replicon vector. SDM primers were designed to change the thymine (T) residue at position 2 of the 5' UTR to a G residue

[0113] In these experiments, VEEV replicons each expressing the rFF luciferase gene, the A Vietnam 1203 influenza hemagglutinin gene (HA) or the green fluorescent protein (GFP) gene, were used as templates to carry out site-directed mutagenesis. The nucleotide T at position 2 in the 5' UTR of the wild-type VEEV (strain TC-83) sequence was changed to G. Primers "VEE 5' T->G nt 2 F" and "VEE 5' T->G nt 2 R" were used to introduce the nucleotide 2 change via a site-directed mutagenesis kit from Agilent. The primers were designed through the Agilent's website. Positive clones were identified and a representative clones expressing the rFF, HA and GFP were sequence-confirmed to be completely correct.

[0114] SDM primers to make Alpha-R-T2G replicons are as follows.

Forward primer: VEE 5' T->G nt 2 F (SEQ ID NO: 23): cgactcactatagaGaggcggcgcgatgag.

Reverse primer: VEE 5' T->G nt 2 R (SEQ ID NO: 24): ctcatgcgcgcctCtctatagtgagtcg.

[0115] After sequence confirmation of T2→G change in the VEEV replicon cDNA, RNA was generated by *in vitro* transcription using T7 RNA polymerase on linearized plasmid DNA. The *in vitro* transcribed RNA was purified and used to electroporate BHK-21 cells. Both replication and GOI expression were monitored by FACS using anti-dsRNA specific and GOI-specific antibodies, respectively. Replication efficiency and GOI expression from U2→G VEEV replicons were compared directly to wild type VEEV replicons expressing the same GOI. The sequence of Alpha-R-rFF-T2G replicon comprising a red Firefly reporter gene is provided as SEQ ID NO: 19 in the Sequence Listing with the T7 promoter and a polyA tail with 40 A residues. The mutated nucleotide at position 2 following the T7 promoter sequence is also indicated.

EXAMPLE 3

U2→G substitution in the 5'UTR does not affect biologic activity of modified VEEV-HA replicon

[0116] This Example describes the results of experiments assessing impact of an U2→G substitution in the 5' UTR of a modified alphavirus replicon on expression of an A Vietnam 1203 influenza hemagglutinin gene (HA) reporter gene.

[0117] In order to demonstrate that an alphavirus replicon containing a U2→G change in the 5' UTR can indeed express protein, replicon RNA was transcribed *in vitro* from a vector carrying a U2→G change in the 5' UTR (U2→G VEEV-HA). As a positive control for expression, RNA was transcribed *in vitro* from a vector carrying a wild-type 5' UTR (WT VEEV-HA). Baby hamster kidney (BHK-21) cells were electroporated with 3 µg of either U2→G VEEV-HA RNA or wild-type VEEV-HA RNA. An example of flow cytometry analysis for this comparison is shown in **FIG. 1**. The cells were analyzed by FACS with an HA-specific antibody to demonstrate both the presence of expressed HA protein and the relative amount of HA expressed on a per cell basis (mean fluorescence intensity – MFI). There is no loss of biologic activity from a replicon RNA carrying the U2→G 5' UTR change relative to the activity detected from a replicon with a wild-type 5' UTR.

[0118] The experimental data presented in this Example indicate that not only can the U2→G VEEV-HA replicon express HA protein but the expression level is equivalent to that of the wild-type VEEV-HA replicon. This result is unexpected considering that the

VEEV/SINV infectious clone carrying the same U2→G 5' UTR change had nearly a three log loss in biological activity; here we show no reduction at all in biologic activity for a replicon carrying the U2→G 5' UTR change.

EXAMPLE 4

U2→G substitution in the 5'UTR enhances expression of GFP reporter by 3 folds

[0119] This Example describes the results of experiments assessing impact of an U2→G substitution in the 5' UTR of a modified alphavirus replicon on expression of a green fluorescence protein (GFP) reporter gene.

[0120] In order to demonstrate that expression from an alphavirus replicon containing a U2→G change in the 5' UTR is not restricted to any particular GOI, replicon vectors expressing the GFP gene were compared in a similar manner. Replicon RNA was transcribed *in vitro* from a vector carrying a U2→G change in the 5' UTR (U2→G VEEV-GFP). As a positive control for expression, RNA was transcribed *in vitro* from a vector carrying a wild-type 5' UTR (WT VEEV-GFP). BHK cells were electroporated with 3 µg of either U2→G VEEV-GFP RNA or wild-type VEEV-GFP RNA. An example of flow cytometry analysis for this comparison is shown in **FIG. 2**. The cells were analyzed for GFP expression by FACS to demonstrate both the presence of expressed GFP protein and the relative amount of GFP expressed on a per cell basis (mean fluorescence intensity – MFI). Not only is there is no loss of biologic activity from a replicon RNA carrying the U2→G 5' UTR but the U2→G 5' UTR change actually enhanced expression by 3 fold relative to the expression detected from a replicon with a wild-type 5' UTR.

[0121] Once again the experimental data presented in this Example indicate that a replicon vector carrying the U2→G 5' UTR change can express protein (this time GFP). Perhaps even more unexpected than simple protein expression is that the U2→G VEEV-GFP replicon expressed three times more GFP than that of the wild-type VEEV-GFP replicon. Once more, the expected loss in biologic activity anticipated due to the U2→G change in the 5' UTR was not realized and this result demonstrated that the U2→G change in the 5' UTR can actually significantly enhance GOI replicon expression.

EXAMPLE 5

U2→G substitution in the 5'UTR enhances expression of rFF reporter by 2 folds

[0122] This Example describes the results of experiments assessing impact of an U2→G substitution in the 5' UTR of a modified alphavirus replicon on expression of a red Firefly (rFF) reporter gene.

[0123] In these experiments, another example of expression from an alphavirus replicon containing a U2→G change in the 5' UTR, replicon vectors expressing the rFF gene were compared in a similar manner. Replicon RNA was transcribed *in vitro* from a vector carrying a U2→G change in the 5' UTR (U2→G VEEV-rFF). As a positive control for expression, RNA was transcribed *in vitro* from a vector carrying a wild-type 5' UTR (WT VEEV-rFF). BHK cells were electroporated with 3 µg of either U2→G VEEV-rFF RNA or wild-type VEEV-rFF RNA. An example of luciferase protein expression is shown in **FIG. 3**. In this experiment, the ability of an alphavirus replicon modified to carry a U2→G change in the 5' UTR (U2→G alpha rFF) was compared to an alphavirus replicon that had a wild-type 5' UTR (Alpha rFF). BHK cells were electroporated with an equivalent amount of *in vitro* transcribe RNA from either replicon and then the cells were analyzed for rFF luciferase expression. The amount of luciferase (expressed as relative light units (RLU)) expressed on a per cell basis is presented. Not only is there is no loss of biologic activity from a replicon RNA carrying the U2→G 5' UTR but the U2→G 5' UTR change actually enhanced expression by approximately 2 fold relative to the expression detected from a replicon with a wild-type 5' UTR.

[0124] Once again the experimental data presented in this Example indicate that a replicon vector carrying the U2→G 5' UTR change can express protein (this time rFF). Perhaps even more unexpected than simple protein expression is that the U2→G VEEV-rFF replicon expressed ~ two times more rFF than that of the wild-type VEEV-rFF replicon. Once more, the expected loss in biologic activity anticipated due to the U2→G change in the 5' UTR was not realized and this result demonstrated that the U2→G change in the 5' UTR can actually significantly enhance GOI replicon expression.

EXAMPLE 6

Multivalent VEEV replicon designs

[0125] This Example describes experiments performed to construct and evaluate multivalent VEEV replicons, which are subsequently deployed for expression of at least two different polypeptides in recombinant cells. In some experiments, the multivalent VEEV replicon includes in 5' to 3' order (i) a 5' sequence required for nonstructural protein-mediated amplification, (ii) a nucleotide sequence encoding VEEV nonstructural proteins nsP1, nsP2, nsP3, and nsP4, (iii) at least two promoters each of which is operably linked to a heterologous nucleic acid sequence, wherein the heterologous nucleic acid sequence replaces one or all of the VEEV structural protein genes, (iv) a 3' sequence required for nonstructural protein-mediated amplification, and (v) a polyadenylate tract.

[0126] While particular alternatives of the present disclosure have been disclosed, it is to be understood that various modifications and combinations are possible and are contemplated within the true spirit and scope of the appended claims. There is no intention, therefore, of limitations to the exact abstract and disclosure herein presented.

WHAT IS CLAIMED IS:

1. A nucleic acid molecule comprising a modified replicon RNA, wherein the modified replicon RNA comprises a modified 5'-UTR and is devoid of at least a portion of a nucleic acid sequence encoding viral structural proteins.
2. The nucleic acid molecule of claim 1, wherein the modified replicon RNA is a modified alphavirus replicon RNA.
3. The nucleic acid molecule of claim 2, wherein the modified alphavirus replicon RNA comprises a modified alphavirus genome.
4. The nucleic acid molecule of any one of claims 1 to 3, wherein the modified 5'-UTR comprises one or more nucleotide substitutions at position 1, 2, 4, or a combination thereof.
5. The nucleic acid molecule of claim 4, wherein at least one of the nucleotide substitutions is a nucleotide substitution at position 2 of the modified 5'-UTR.
6. The nucleic acid molecule of claim 5, wherein the nucleotide substitutions at position 2 of the modified 5'-UTR is a U→G substitution.
7. The nucleic acid molecule of any one of claims 1 to 6, wherein the modified replicon RNA is devoid of a substantial portion of the nucleic acid sequence encoding viral structural proteins.
8. The nucleic acid molecule of any one of claims 1 to 7, wherein the modified alphavirus genome or replicon RNA comprises no nucleic acid sequence encoding viral structural proteins.

9. The nucleic acid molecule of any one of claims 1 to 8, further comprising one or more expression cassettes, wherein each of the expression cassettes comprises a promoter operably linked to a heterologous nucleic acid sequence.
10. The nucleic acid molecule of any one of claims 1 to 9, wherein the modified replicon RNA comprises a modified genome or replicon RNA of a virus belonging to the *Alphavirus* genus of the *Togaviridae* family.
11. The nucleic acid molecule of claim 10, wherein the modified genome or replicon RNA is of an alphavirus belonging to the VEEV/EEEV group, or the SF group, or the SIN group.
12. The nucleic acid molecule of claim 10, wherein the alphavirus is Venezuelan equine encephalitis virus (VEEV).
13. The nucleic acid molecule of any one of claims 1 to 12, wherein the modified alphavirus genome or replicon RNA is operably linked to a heterologous regulatory element.
14. A nucleic acid molecule comprising a modified alphavirus genome or replicon RNA, wherein the nucleic acid molecule exhibits at least 80% sequence identity to the nucleic acid sequence of SEQ ID NO: 1, wherein the modified alphavirus genome or replicon RNA comprises a U→G substitution at position 2 of the 5'-untranslated region (5'-UTR) and is devoid of at least a portion of the sequence encoding viral structural proteins.
15. The nucleic acid molecule of claim 14, wherein the nucleic acid molecule exhibits at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the nucleic acid sequence of SEQ ID NO: 1.
16. A nucleic acid molecule comprising a modified alphavirus genome or replicon RNA, wherein the modified alphavirus genome or replicon RNA comprises a 5'-UTR exhibiting at

least 80% sequence identity to the nucleic acid sequence of at least one of SEQ ID NOS: 2-18 and a U→G substitution at position 2 of the 5'-UTR, and wherein the modified alphavirus genome or replicon RNA is devoid of at least a portion of the sequence encoding viral structural proteins.

17. The nucleic acid molecule of claim 16, wherein the modified alphavirus genome or replicon RNA comprises a 5'-UTR exhibiting at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the nucleic acid sequence of at least one of SEQ ID NOS: 2-18.

18. The nucleic acid molecule of any one of claims 14 to 17, wherein the modified alphavirus genome or replicon RNA is devoid of a substantial portion of the nucleic acid sequence encoding viral structural proteins.

19. The nucleic acid molecule of claims 14 to 18, wherein the modified alphavirus genome or replicon RNA comprises no nucleic acid sequence encoding viral structural proteins.

20. A recombinant cell comprising a nucleic acid molecule of any one of claims 1 to 19.

21. The recombinant cell of claim 20, wherein the recombinant cell is a prokaryotic cell or a eukaryotic cell.

22. The recombinant cell of claim 20, wherein the recombinant cell is an animal cell.

23. A method for producing a polypeptide of interest, comprising culturing a host cell comprising a nucleic acid molecule which comprises a modified replicon RNA, wherein the modified replicon RNA comprises a modified 5'-UTR and is devoid of at least a portion of a nucleic acid sequence encoding viral structural proteins.

24. A method for producing a polypeptide of interest in a subject, comprising administering to the subject a nucleic acid molecule comprising a modified replicon, RNA wherein the modified replicon RNA comprises a modified 5'-UTR and is devoid of at least a portion of a nucleic acid sequence encoding viral structural proteins.
25. The method of claim 24, wherein the subject is a human, horse, pig, primate, mouse, cattle, swine, sheep, rabbit, cat, dog, bird, fish, goat, donkey, hamster, or buffalo.
26. The method of any one of claims 23 to 25, wherein the modified replicon RNA is a modified alphavirus replicon RNA.
27. The method of claim 26, wherein the modified alphavirus replicon RNA comprises a modified alphavirus genome.
28. The method of any one of claims 23 to 27, wherein the modified 5'-UTR comprises one or more nucleotide substitutions at position 1, 2, 4, or a combination thereof.
29. The method of claim 28, wherein at least one of the nucleotide substitutions is a nucleotide substitution at position 2 of the modified 5'-UTR.
30. The method of claim 29, wherein the nucleotide substitutions at position 2 of the modified 5'-UTR is a U→G substitution.
31. The method of any one of claims 23 to 30, wherein the modified replicon RNA is devoid of a substantial portion of the nucleic acid sequence encoding viral structural proteins.
32. The method of any one of claims 23 to 31, wherein the modified alphavirus genome or replicon RNA comprises no nucleic acid sequence encoding viral structural proteins.

33. The method of any one of claims 23 to 32, further comprising one or more expression cassettes, wherein each of the expression cassettes comprises a promoter operably linked to a heterologous nucleic acid sequence.
34. The method of any one of claims 23 to 33, wherein the modified replicon RNA comprises a modified genome or replicon RNA of a virus belonging to the *Alphavirus* genus of the *Togaviridae* family.
35. The method of claim 34, wherein the modified genome or replicon RNA is of an alphavirus belonging to the VEEV/EEEV group, or the SF group, or the SIN group.
36. The method of claim 34, wherein the alphavirus is Venezuelan equine encephalitis virus (VEEV).
37. The method of any one of claims 23 to 36, wherein the modified alphavirus genome or replicon RNA is operably linked to a heterologous regulatory element.
38. The method of claims 23 or 24, wherein the nucleic acid molecule comprising a modified alphavirus genome or replicon RNA, wherein the modified alphavirus genome or replicon RNA comprises a 5'-UTR exhibiting at least 80% sequence identity to the nucleic acid sequence of SEQ ID NO: 1 and a U->G substitution at position 2 of the 5'-UTR, and wherein the modified alphavirus genome or replicon RNA is devoid of at least a portion of the sequence encoding viral structural proteins.
39. The method of claim 38, wherein the nucleic acid molecule exhibits at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the nucleic acid sequence of SEQ ID NO: 1.
40. The method of claims 23 or 24, wherein the nucleic acid molecule comprising a modified alphavirus genome or replicon RNA, wherein the modified alphavirus genome or replicon RNA comprises a 5'-UTR exhibiting at least 80% sequence identity to the nucleic

acid sequence of at least one of SEQ ID NOS: 2-18 and a U->G substitution at position 2 of the 5'-UTR, and wherein the modified alphavirus genome or replicon RNA is devoid of at least a portion of the sequence encoding viral structural proteins.

41. The method of claim 40, wherein the modified alphavirus genome or replicon RNA comprises a 5'-UTR exhibiting at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the nucleic acid sequence of at least one of SEQ ID NOS: 2-18.

42. A recombinant polypeptide produced by the method of any one of claims 23 to 41.

43. A composition comprising a recombinant polypeptide of claim 42, and a pharmaceutically acceptable carrier.

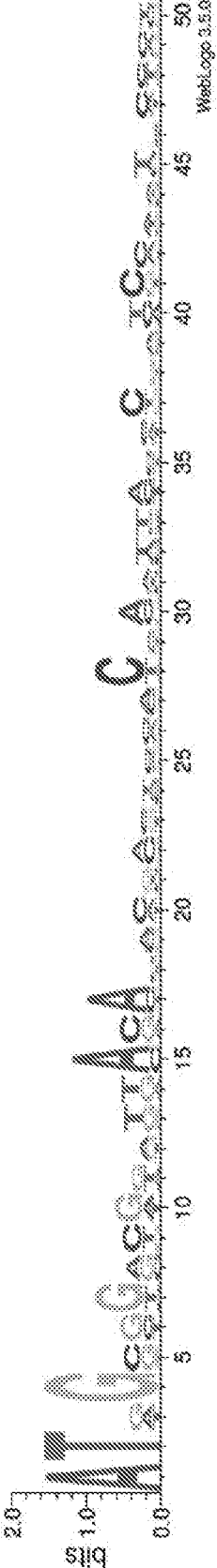
44. A composition comprising a nucleic acid molecule of any one of claims 1 to 19, and a pharmaceutically acceptable carrier.

45. A composition comprising a recombinant cell of any one of claims 20 to 22, and a pharmaceutically acceptable carrier.

FIG. 1A

SEQ ID NO: 2 (AURAV)	ATAGCGGACGGACTAGTACTTGTACTACAGAAATTAAC-TGCCGGTGTGCCGC-----
SEQ ID NO: 3 (CHIKV)	ATGGCTG-CGTGAGA-----CACACGTAGCCTACCAGTTTCTTACTGCTCTACTCT-
SEQ ID NO: 4 (ONNV)	ATAGCTG-CGTGATA-----CACACACGACGCTTACGGGTTTCATACIGCTCTACT-
SEQ ID NO: 5 (BEBV)	ATGGCGGCTGTGTGA-CACACGAGCGGTGATTTCAA-CCTTCTTGCTCCCT-----
SEQ ID NO: 6 (SEV)	ATGGCGGATGTGTGA-CATACACGACGCCAAAGATT-TTGTTCAGCTCCT-----
SEQ ID NO: 7 (MAYV)	ATGGCGGGCAAGTGA-CACTGTTCGGCCGGTCGTCT-----CTAAGCTCTCCCTC-----
SEQ ID NO: 8 (GETV)	ATGGCGGACGTGTGACATCACCGTTCGGCTCTTCTAG-GATCCTTTGCTAC-----
SEQ ID NO: 9 (SAGV)	ATGGCGGACGTGTGACATCACCGTTCGGCTCTTCTAG-GATCCTTTGCTAC-----
SEQ ID NO: 10 (NDUV)	ATGG--TGCGGAGTT-GA--GAGACGA-AGCACCAA-ACAACTACGGGGCTCACC-AT-----
SEQ ID NO: 11 (MIDV)	ATTGGTGGTTACGTA--CACGTGCCACCCACCCCCCA-CCCTCCAAAGCGATCCA-----
SEQ ID NO: 12 (EEEV)	ATAGGGTACGGTGTA-----GAGGCAACCAACCTAT-----TTCCACCTATCCAAATGG-----
SEQ ID NO: 13 (FMV)	ATAGGGTATGGTTA-----GAGGCGCTACCCCTAC-----TTAACCGATCCAAACATGG-----
SEQ ID NO: 14 (Buggy)	ATAGGGTATGGTTA-----GAGGCGCTACCCCTAC-----TTAACCGATCCAAACATGG-----
SEQ ID NO: 15 (VEEV)	ATGGGCGGCGCAAGA-----GAGAAGCCCAAAACCAA-----TTACCTACCCAAATGGAG-----
SEQ ID NO: 16 (WHAV)	ATTGGCGGCATAGTA-----CATACTATATAAAGAAACAGCCGACCAATTGCAC-----
SEQ ID NO: 17 (SINV)	ATTGACGGCGTAGTA-----CACACTAT-TGAATCAAAACAGCCGACCAATTGCAC-----
SEQ ID NO: 18 (BABV)	ATTGGCGGCGTAGTA-----CACACTAT-TGAATCAAAACAGCCGACCAATTGCAC-----

FIG. 1B



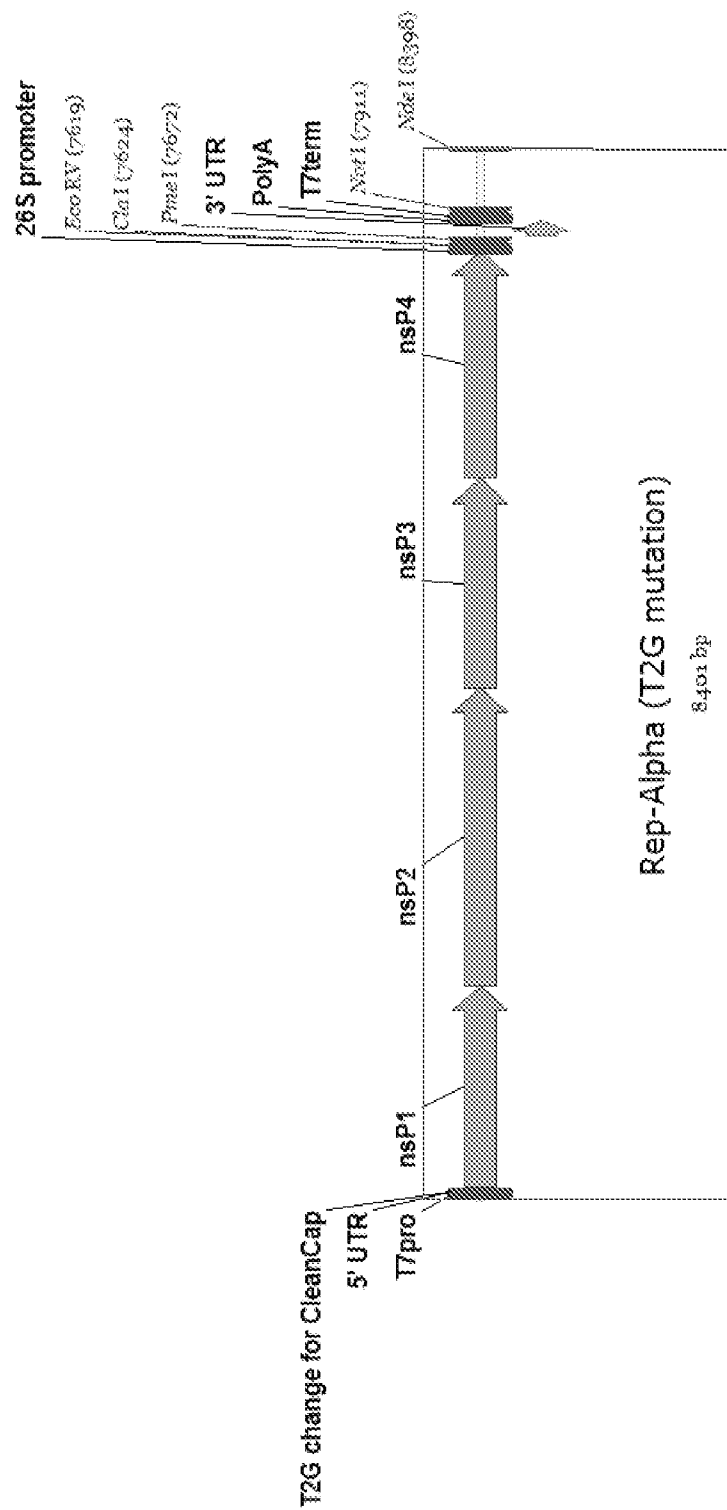


FIG. 2

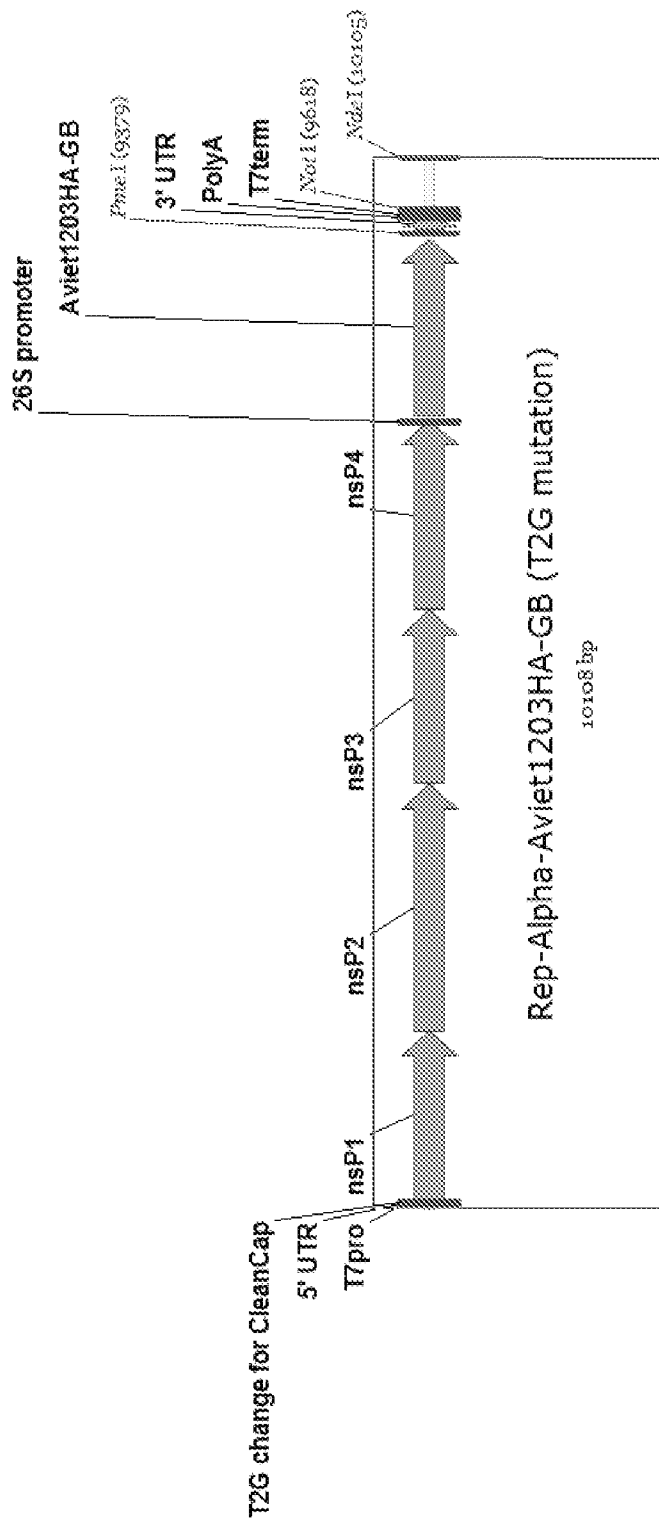


FIG. 3

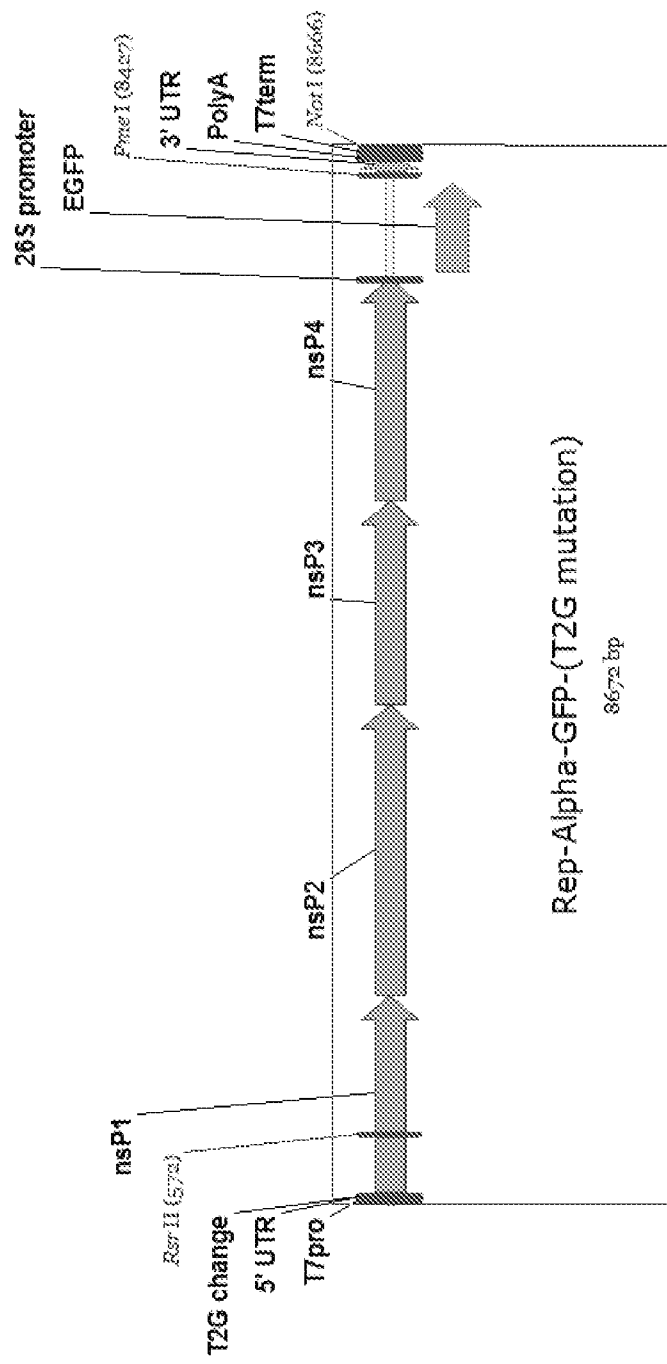


FIG. 4

Mean Fluorescence Intensity

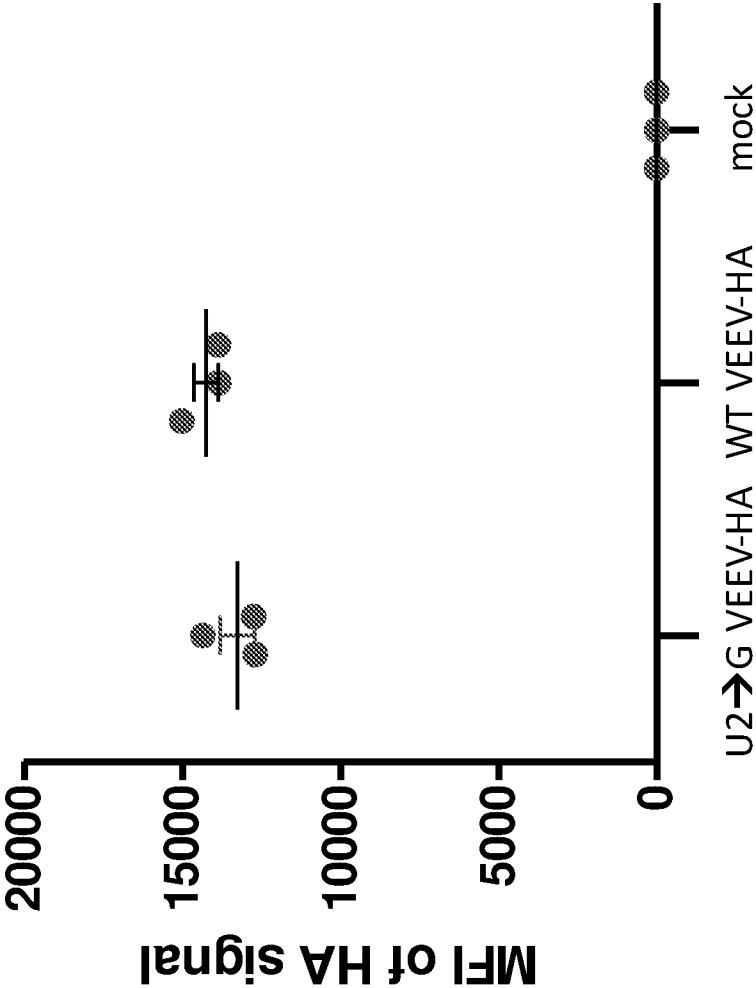


FIG. 5

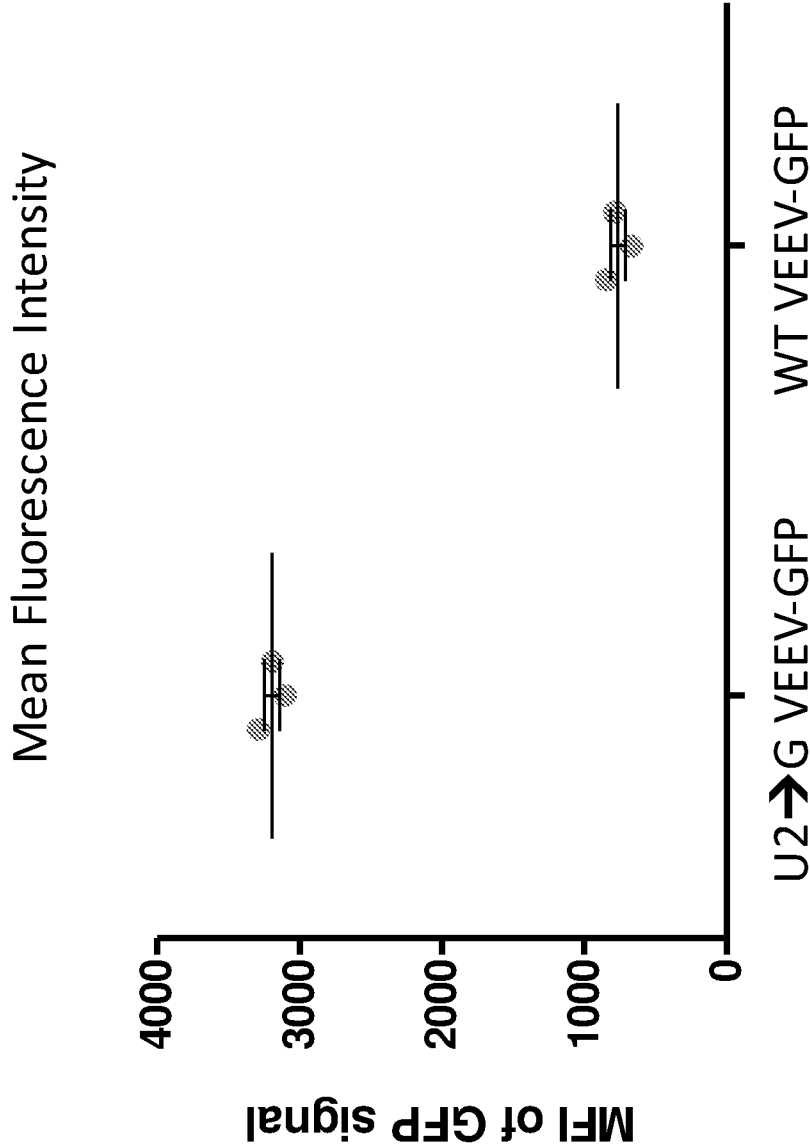


FIG. 6

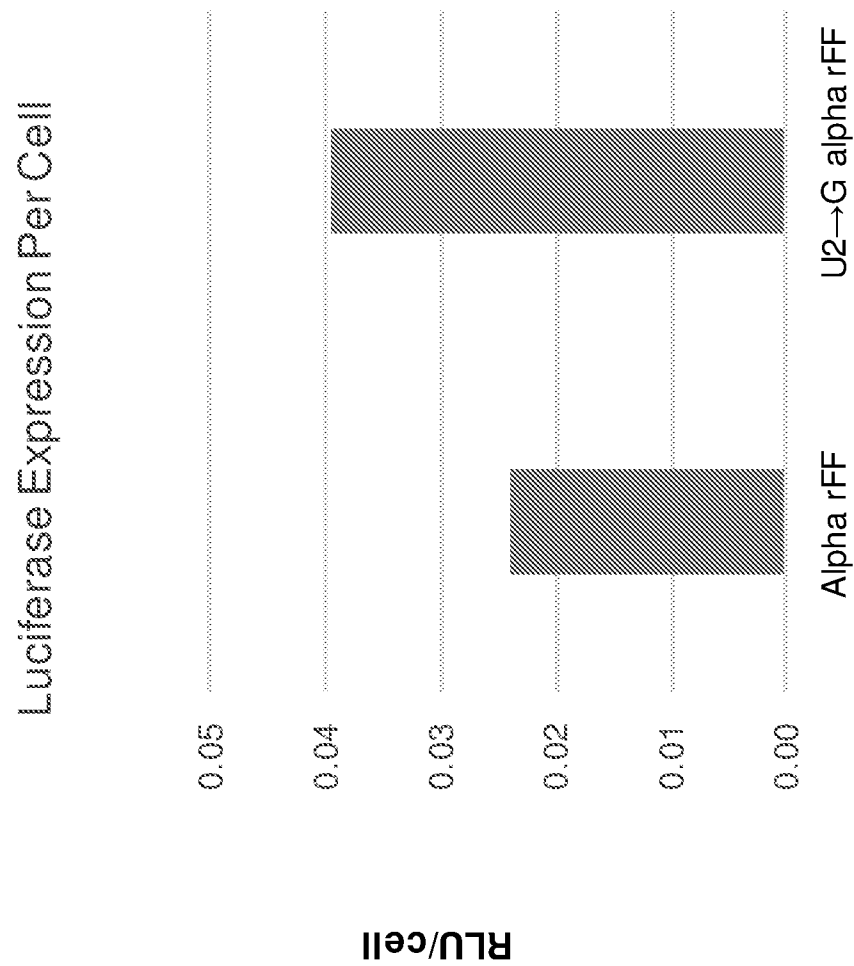


FIG. 7

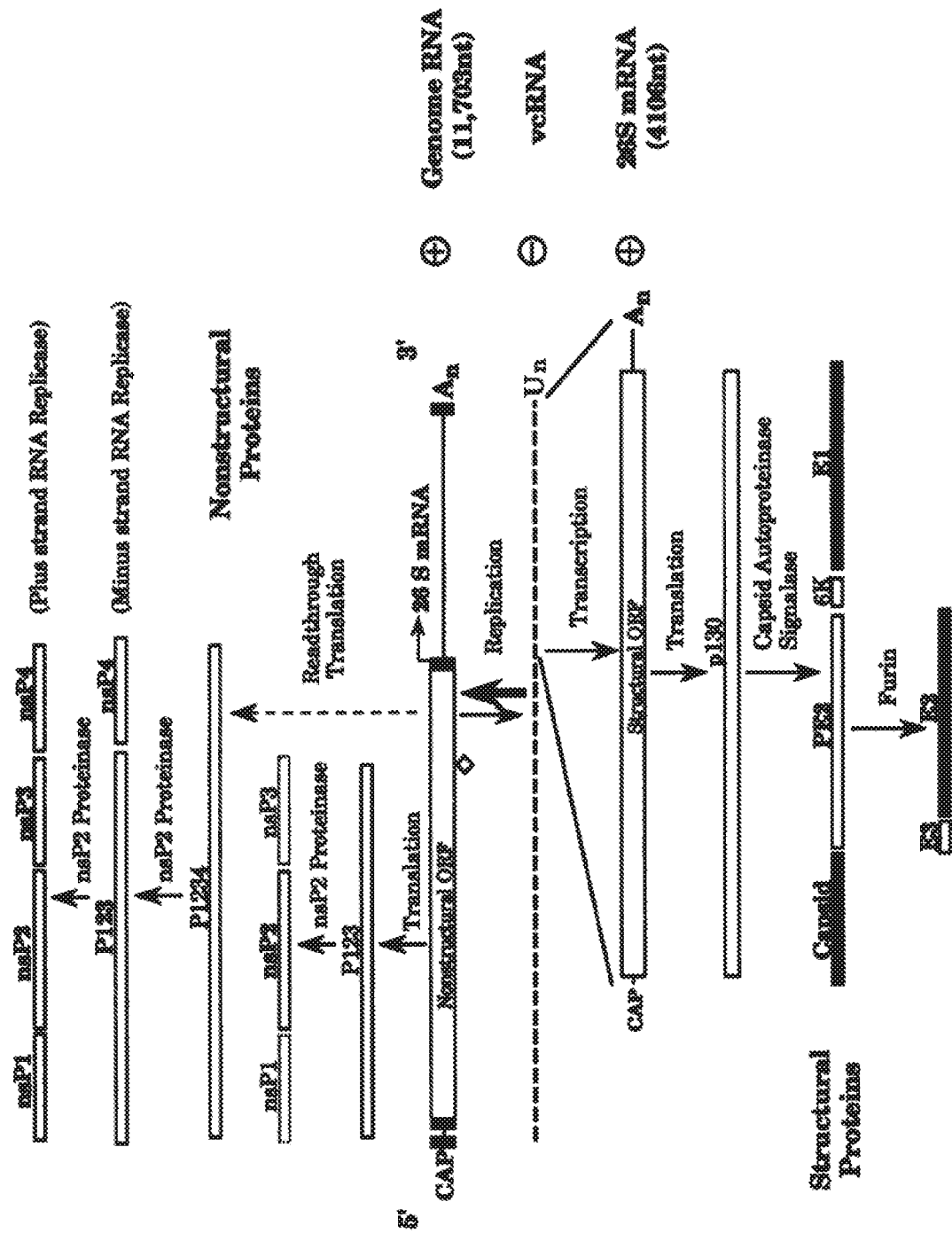


FIG. 8

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2017/054928A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/86
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N A61K

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

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

EPO-Internal, Sequence Search, WPI Data, BIOSIS, CHEM ABS Data, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	R. KULASEGARAN-SHYLINI ET AL: "Structural and Functional Elements of the Promoter Encoded by the 5' Untranslated Region of the Venezuelan Equine Encephalitis Virus Genome", JOURNAL OF VIROLOGY., vol. 83, no. 17, September 2009 (2009-09), pages 8327-8339, XP055427201, US ISSN: 0022-538X, DOI: 10.1128/JVI.00586-09 cited in the application abstract page 8328, right-hand column, line 21 - line 34 page 8330, left-hand column, last paragraph - right-hand column, paragraph 2 figures 1,8 ----- -/--	1-23,42



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

22 November 2017

Date of mailing of the international search report

01/12/2017

Name and mailing address of the ISA/

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

Authorized officer

Lewis, Birgit

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2017/054928

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

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

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2017/054928

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>WO 2014/170493 A2 (NOVARTIS AG) 23 October 2014 (2014-10-23)</p> <p>paragraphs [0025], [0026] example 3; table 1 figures 1,4A,5A,5B,6A,6B</p> <p>-----</p>	<p>1-3, 7-13, 20-29, 31-37, 42-45</p>
X	<p>MARUGGI GIULIETTA ET AL: "Engineered alphavirus replicon vaccines based on known attenuated viral mutants show limited effects on immunogenicity", VIROLOGY, vol. 447, no. 1, 5 October 2013 (2013-10-05), pages 254-264, XP028754361, ISSN: 0042-6822, DOI: 10.1016/J.VIROL.2013.07.021 abstract figures 1,2,3</p> <p>-----</p>	<p>1-3, 7-11,13, 20-27, 31-35, 42-44</p>
X	<p>US 2014/079734 A1 (FROLOV ILYA ET AL) 20 March 2014 (2014-03-20)</p> <p>paragraphs [0024], [0025], [0046], [0055], [0129]; figures 3,4</p> <p>-----</p>	<p>1-3, 7-11, 42-44</p>
X	<p>D. Y. KIM ET AL: "Enhancement of protein expression by alphavirus replicons by designing self-replicating subgenomic RNAs", PROCEEDINGS NATIONAL ACADEMY OF SCIENCES PNAS, vol. 111, no. 29, 7 July 2014 (2014-07-07) , pages 10708-10713, XP055427078, US ISSN: 0027-8424, DOI: 10.1073/pnas.1408677111 abstract figures 1,5</p> <p>-----</p>	<p>42-44</p>

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2017/054928

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2014170493	A2	23-10-2014	NONE
US 2014079734	A1	20-03-2014	NONE