



- (51) **International Patent Classification:**
C12N 15/86 (2006.01) *C07K 16/32* (2006.01)
C12N 7/00 (2006.01)
- (21) **International Application Number:**
PCT/US2017/027249
- (22) **International Filing Date:**
12 April 2017 (12.04.2017)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
62/322,149 13 April 2016 (13.04.2016) US
- (71) **Applicant:** SYNTHETIC GENOMICS, INC. [US/US];
11149 North Torrey Pines Road, La Jolla, CA 92037 (US).
- (72) **Inventors:** KAMRUD, Kurt, Iver; 11149 North Torrey
Pines Road, La Jolla, CA 92037 (US). WANG, Nathaniel,
Stephen; 11149 North Torrey Pines Road, La Jolla, CA
92037 (US). FELDERMAN, Martina; 11149 North Tor-
rey Pines Road, La Jolla, CA 92037 (US). CARRICO,
Nancy, C.; 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, 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).

Published:

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

(54) **Title:** RECOMBINANT ARTERIVIRUS REPLICON SYSTEMS AND USES THEREOF

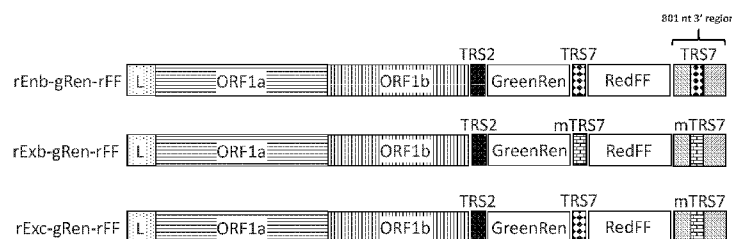


FIG. 7

(57) **Abstract:** The present disclosure generally relates to viral-based expression systems suitable for the production of molecule of interests in recombinant host cells. The disclosure particularly relates to nucleic acid constructs, such as expression vectors, containing a modified arterivirus genome or replicon RNA in which at least some of its original viral sequence has been deleted. Also included in the disclosure are viral-based expression vectors including one or more expression cassettes encoding heterologous polypeptides. In some embodiments, the expression cassettes are configured and positioned at defined locations on the viral genome so as to enable expression of the heterologous polypeptides in a tunable manner.

RECOMBINANT ARTERIVIRUS REPLICON SYSTEMS AND USES THEREOF

RELATED APPLICATIONS

[0001] The present application claims priority to U.S. Provisional Patent Application Serial No. 62/322,149; filed on April 13, 2016, the content of which is hereby 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 SGI010WO_Sequence Listing, was created on March 13, 2017 and is 240 KB. The file can be assessed using Microsoft Word on a computer that uses Windows OS.

BACKGROUND

[0003] 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 as vaccines, expression vectors, anti-tumor agents, gene therapy vectors, and drug delivery vehicles.

[0004] For example, many viral-based expression vectors have been deployed for expression of heterologous proteins in cultured recombinant cells. For example, 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.

[0005] However, there is still a need for more efficient methods and systems for expressing a gene of interest in heterologous expression systems, including multigenic expression systems for simultaneously producing of multiple heterologous polypeptides in cells and/or host organisms.

SUMMARY OF CERTAIN INVENTIVE ASPECTS

[0006] This section provides a general summary of the present disclosure, and is not comprehensive of its full scope or all its features.

[0007] In one aspect, disclosed herein is a nucleic acid molecule including a nucleotide sequence that encodes a modified arterivirus genome or replicon RNA, in which the modified genome or replicon RNA includes a sequence fragment exhibiting at least 80% sequence identity to the sequence encoding open reading frame ORF7, and wherein the modified genome or replicon RNA is devoid of the sequence encoding open reading frame ORF2a. In some embodiments, the modified arterivirus genome or replicon RNA is further devoid of at least a portion of the sequence encoding one or more of the open reading frames ORF2b, ORF3, ORF4, ORF5a, and ORF5. In some embodiments, the nucleotide sequence is devoid of the ATG start codon of ORF7. In some embodiments, the modified arterivirus genome or replicon RNA is further devoid of a portion of the sequence encoding ORF6. In some embodiments, the modified arterivirus genome or replicon RNA is devoid of the ATG start codon of ORF6. In some embodiments, the modified arterivirus genome or replicon RNA is further devoid of TRS7 or comprises a mutated TRS7.

[0008] In various embodiments disclosed herein, the nucleic acid molecule can comprise a modified arterivirus genome or replicon RNA including one or more subgenomic (*sg*) promoters at a non-native site, wherein each of the one or more *sg* promoters includes a transcriptional regulatory sequence (TRS). In some embodiments, at least one of the one or more *sg* promoters includes a sequence exhibiting at least 80% sequence identity to a sequence selected from the group consisting of *sg* promoter 1, *sg* promoter 2, *sg* promoter 3, *sg* promoter 4, *sg* promoter 5, *sg* promoter 6, *sg* promoter 7, and a variant thereof. In some embodiments, at least one of the one or more *sg* promoters is a modified *sg* promoter. In some embodiments, the modified arterivirus genome or replicon RNA includes one or more

modified *sg* promoters located at their respective native site, wherein each of the one or more modified *sg* promoters includes a TRS. In some embodiments, at least one of the one or more modified *sg* promoters is a modified *sg* promoter 7. In some embodiments, at least one of the one or more modified *sg* promoter includes one or more nucleotide modifications positioned within the primary sequence required for the formation of a secondary structure of RNA transcripts including the respective *sg* promoter sequence. In some embodiments, at least one of the one or more modified *sg* promoter includes a nucleotide modification positioned within the sequence of the TRS. In some exemplary embodiments, at least one of the one or more modified *sg* promoters includes a leader TRS or a variant thereof. In some exemplary embodiments, at least one of the one or more modified *sg* promoters includes a body TRS or a variant thereof.

[0009] In some embodiments, the modified arterivirus genome or replicon RNA disclosed herein includes one or more mutated T7 transcriptional termination signal sequences. In accordance with some exemplary embodiments, at least one of the one or more T7 mutated transcriptional termination signal sequences includes a nucleotide substitution selected from the group consisting of T9001G, T3185A, G3188A, and combinations of any two or more thereof. In some embodiments, the modified arterivirus genome or replicon RNA as disclosed herein includes one or more heterologous transcriptional termination signal sequences. In some embodiments, at least one of the one or more heterologous transcriptional termination signal sequences is a SP6 termination signal sequence, a T3 termination signal sequence, or a variant thereof. In some embodiments, at least one of the one or more heterologous transcriptional termination signal sequences is inactivated.

[0010] In some embodiments, the nucleic acid molecule disclosed herein further includes one or more spacer regions operably positioned adjacent to at least one of the one or more *sg* promoters. In some embodiments, at least one of the one or more spacer regions is positioned immediately 3' to the *sg* promoter. In some embodiments, at least one of the one or more spacer regions is positioned immediately 5' to the *sg* promoter. In some embodiments, the one or more spacer regions is about 20 to 400 nucleotides in length.

[0011] In various embodiments of this aspect and other aspects of the present disclosure, the nucleic acid molecule disclosed herein further includes one or more

expression cassettes, wherein each of the expression cassettes includes a *sg* promoter operably linked to a heterologous nucleotide sequence. In some embodiments, the *sg* promoter includes a TRS, a first flanking region positioned immediately 5' to the TRS, and second flanking region positioned immediately 3' to the TRS, wherein the first flanking region is about 5 to 400 nucleotides in length and the second flanking region is about 15 to 115 nucleotides in length. In some embodiments, the nucleic acid molecule disclosed herein includes two, three, four, five, or six expression cassettes.

[0012] In some embodiments, the heterologous nucleotide sequence as disclosed herein 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 embodiment, the secondary structure of the RNA transcript including the coding sequence of the GOI is optimized for improved RNA replication.

[0013] In some embodiments, the nucleic acid molecule disclosed herein includes a nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the arterivirus is selected from the group consisting of Equine arteritis virus (EAV), Porcine respiratory and reproductive syndrome virus (PRRSV), Lactate dehydrogenase elevating virus (LDV), and Simian hemorrhagic fever virus (SHFV). In some embodiments, the arterivirus is an Equine arteritis virus (EAV), an EAV-virulent Bucyrus strain (VBS), or a Simian hemorrhagic fever virus (SHFV).

[0014] In some particular embodiments of the application, the nucleic acid molecule disclosed herein includes a nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified arterivirus genome or replicon RNA includes a nucleotide sequence exhibiting at least 80% sequence identity to SEQ ID NO: 1 and at least 80% sequence identity to SEQ ID NO: 2, and further wherein the modified genome or replicon RNA includes a nucleotide sequence exhibiting at least 80% sequence identity to the sequence encoding open reading frame ORF7, and is devoid of the sequence encoding ORF2a. In some embodiments, the modified arterivirus genome or replicon RNA includes a nucleotide sequence exhibiting at least about 80% sequence identity to SEQ ID NO: 3. In some embodiments, the modified arterivirus genome or replicon RNA includes a nucleotide

sequence of SEQ ID NO: 3. In some particular embodiments of the application, the nucleic acid molecule disclosed herein includes a nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified arterivirus genome or replicon RNA includes a nucleotide sequence exhibiting at least 80% sequence identity to SEQ ID NO: 40 and at least 80% sequence identity to SEQ ID NO: 41, and further wherein the modified genome or replicon RNA includes a nucleotide sequence exhibiting at least 80% sequence identity to the sequence encoding open reading frame ORF7, and is devoid of the sequence encoding ORF2a. In some embodiments, the modified arterivirus genome or replicon RNA includes a nucleotide sequence exhibiting at least about 80% sequence identity to SEQ ID NO: 42. In some embodiments, the modified arterivirus genome or replicon RNA includes a nucleotide sequence of SEQ ID NO: 42.

[0015] 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 animal cell is a vertebrate animal cell or an invertebrate animal cell. In some embodiments, the recombinant cell is selected from the group consisting of a pulmonary equine artery endothelial cell, equine dermis cell, baby hamster kidney 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, and mouse 3T3 cell. In a related aspect, some embodiments disclosed herein relate to a cell culture that includes at least one recombinant cell as disclosed herein.

[0016] In one aspect, some embodiments disclosed herein relate to a method for producing a polypeptide of interest that involves culturing a host cell including a nucleic acid molecule which includes (i) a nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified genome or replicon RNA includes a sequence fragment exhibiting at least 80% sequence identity to the sequence encoding open reading frame ORF7, and wherein the modified genome or replicon RNA is devoid of the sequence encoding ORF2a; and (ii) one or more expression cassettes, wherein each of the one or more

expression cassettes includes a subgenomic (sg) promoter operably linked to a heterologous nucleotide sequence encoding a gene of interest.

[0017] In a further aspect, some embodiments disclosed herein relate to a method for producing a polypeptide of interest in a subject that involves administering to the subject a nucleic acid molecule which includes (i) a nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified genome or replicon RNA includes a sequence fragment exhibiting at least 80% sequence identity to the sequence encoding open reading frame ORF7, and wherein the modified genome or replicon RNA is devoid of the sequence encoding ORF2a; and (ii) one or more expression cassettes, wherein each of the one or more expression cassettes includes a subgenomic (sg) promoter operably linked to a heterologous nucleotide sequence encoding a gene of interest. In some embodiments, the subject is a vertebrate animal or an invertebrate animal.

[0018] Implementations of embodiments of the methods according to the present disclosure can include one or more of the following features. In some embodiments, the sg promoter includes a TRS, a first flanking region positioned immediately 5' to the TRS, and second flanking region positioned immediately 3' to the TRS, wherein the first flanking region is about 5 to 400 nucleotides in length and the second flanking region is about 15 to 115 nucleotides in length. In some embodiments, the modified arterivirus genome or replicon RNA is further devoid of a portion of the sequence encoding one or more of open reading frames ORF2b, ORF3, ORF4, ORF5a, and ORF5. In some embodiments, the modified arterivirus genome or replicon RNA includes a sequence fragment that is devoid of ATG start codon of ORF7. In some embodiments, the modified arterivirus genome or replicon RNA is further devoid of a portion of the sequence encoding ORF6. In some embodiments, the modified arterivirus genome or replicon RNA is devoid of the ATG start codon of ORF6. In some embodiments, the modified arterivirus genome or replicon RNA is further devoid of TRS7 or comprises a mutated TRS7.

[0019] In some embodiments, at least one of the one or more expression cassettes includes a modified sg promoter including at least one nucleotide modification introduced within the primary sequence required for the formation of a secondary structure of RNA transcripts including the sg promoter sequence. In some embodiments, at least one of the one

or more expression cassettes includes a modified *sg* promoter including at least one nucleotide modification introduced within the sequence of the TRS. In some embodiments, at least one of the one or more expression cassettes includes a leader TRS or a variant thereof.

[0020] In some embodiments of the methods disclosed herein, the modified arterivirus genome or replicon RNA includes one or more mutated T7 transcriptional termination signal sequences. In some embodiments, at least one of the one or more mutated T7 transcriptional termination signal sequences includes a nucleotide substitution selected from the group consisting of T9001G, T3185A, G3188A, and combinations of any two or more thereof. In some embodiments, the modified arterivirus genome or replicon RNA includes one or more heterologous transcriptional termination signal sequences. In some particular embodiments, at least one of the one or more heterologous transcriptional termination signal sequences is a SP6 termination signal sequence, a T3 termination signal sequence, or a variant thereof.

[0021] In some embodiments of the methods disclosed herein, the modified arterivirus genome or replicon RNA includes one or more spacer regions operably positioned adjacent to at least one of the one or more *sg* promoters. In some embodiments, at least one of the one or more spacer regions is positioned immediately 3' to the *sg* promoter. In some embodiments, at least one of the one or more spacer regions is positioned immediately 5' to the *sg* promoter. In some embodiments, each of the one or more spacer regions is about 20 to 400 nucleotides in length. In some embodiments, the nucleic acid molecule of the methods disclosed herein includes two, three, four, five, or six expression cassettes.

[0022] In some embodiments of the methods disclosed herein, the coding sequence of the gene of interest in one of the expression cassettes is optimized for expression at a level higher than the expression level of a reference coding sequence. In some embodiments, the secondary structure of the RNA transcript comprising coding sequence of the gene of interest is optimized for improved RNA replication.

[0023] In some embodiments of the methods disclosed herein, the nucleic acid molecule includes a nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified arterivirus genome or replicon RNA includes a nucleotide

sequence exhibiting at least 80%, 85%, 90%, 95%, 98%, 99%, or more sequence identity to SEQ ID NO: 1 and at least 80%, 85%, 90%, 95%, 98%, 99%, or more sequence identity to SEQ ID NO: 2, and further wherein the modified genome or replicon RNA includes a nucleotide sequence exhibiting at least 80%, 85%, 90%, 95%, 98%, 99%, or more sequence identity to the sequence encoding open reading frame ORF7, and is devoid of the sequence encoding ORF2a. In some embodiments, the modified arterivirus genome or replicon RNA includes a nucleotide sequence exhibiting at least about 80%, 85%, 90%, 95%, 98%, 99%, or more sequence identity to SEQ ID NO: 3. In some embodiments, the modified arterivirus genome of replicon RNA includes a nucleotide sequence of SEQ ID NO: 3. In some embodiments of the methods disclosed herein, the nucleic acid molecule includes a nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified arterivirus genome or replicon RNA includes a nucleotide sequence exhibiting at least 80%, 85%, 90%, 95%, 98%, 99%, or more sequence identity to SEQ ID NO: 40 and at least 80%, 85%, 90%, 95%, 98%, 99%, or more sequence identity to SEQ ID NO: 41, and further wherein the modified genome or replicon RNA includes a nucleotide sequence exhibiting at least 80%, 85%, 90%, 95%, 98%, 99%, or more sequence identity to the sequence encoding open reading frame ORF7, and is devoid of the sequence encoding ORF2a. In some embodiments, the modified arterivirus genome or replicon RNA includes a nucleotide sequence exhibiting at least about 80%, 85%, 90%, 95%, 98%, 99%, or more sequence identity to SEQ ID NO: 42. In some embodiments, the modified arterivirus genome of replicon RNA includes a nucleotide sequence of SEQ ID NO: 42. In some embodiments, the modified arterivirus genome of replicon RNA consists of a nucleotide sequence of SEQ ID NO: 3 or SEQ ID NO: 42. In some embodiments, the modified arterivirus genome of replicon RNA consists of a nucleotide sequence exhibit at least 80%, 85%, 90%, 95%, 98%, 99%, or more sequence identity to SEQ ID NO: 3 or SEQ ID NO: 42.

[0024] In a further aspect, some embodiments disclosed herein relate to recombinant polypeptides produced by a method in accordance with one or more embodiments described herein. In some embodiments, the polypeptide is selected from the group consisting of a therapeutic polypeptide, a prophylactic polypeptide, a diagnostic polypeptide, a nutraceutical polypeptide, an industrial enzyme, and a reporter polypeptide. In

some embodiments, the polypeptide is selected from the group consisting of an antibody, an antigen, an immune modulator, and a cytokine.

[0025] In a related aspect, some embodiments disclosed herein relate to a composition including a recombinant polypeptide as described herein and a pharmaceutically acceptable carrier.

[0026] In yet a further aspect, some embodiments disclosed herein relate to a composition including a nucleic acid molecule as disclosed herein and a pharmaceutically acceptable carrier.

[0027] In yet a further aspect, some embodiments disclosed herein relate to a composition including a recombinant cell as disclosed herein and a pharmaceutically acceptable carrier.

[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] FIGURES 1A-1D graphically summarize the results of luciferase assays and flow cytometry experiments performed to demonstrate that mutations to cryptic T7 termination sites in the original EAV replicon system Rep-EAV allow for full-length RNA replicon transcription. Potential T7 polymerase termination sequences that are endogenous to the EAV genome were mutated to allow for full-length RNA transcription. Mutations T9001G (rE), T9001G and G3188A (rE2), and T9001G and T3185A (rE3) were compared to the initial replicon system Rep-EAV for replication and protein expression of red Firefly (1A, 1C) and green Renilla (1B, 1D).

[0030] FIGURE 2 graphically depicts the structure of a base monovalent EAV replicon design, Rep-EAV (WT), which is a replicon vector with the region between TRS2 and the start of the green Renilla luciferase reporter gene magnified for details of the sequence included. The EAV sequences that flank the TRS7 element are in gray and underlined. *Xba*I is the unique restriction site engineered downstream of ORF1b stop codon;

a single nucleotide “C” separates the *Xba*I sequence from the ORF1b stop codon. L: Leader sequence; ORF1a and ORF1b: EAV nonstructural genes.

[0031] FIGURE 3 graphically depicts the structure of two initial bivalent EAV replicon designs. Schematic representation of the bivalent replicon vectors shows differences between A design and B design. The 2/7 block is highlighted in the B design.

[0032] FIGURES 4A-4B graphically summarize the results of luciferase assays performed to assess expression of a heterologous polypeptide contained in the initial bivalent replicon designs. Bulk-cell luciferase assays carried out on electroporated cells. 4A). Analysis of red firefly luciferase expression from A and B bivalent replicon designs. 4B). Analysis of green Renilla luciferase expression from A and B bivalent replicon designs.

[0033] FIGURE 5 graphically depicts the structure of exemplary EAV replicon designs containing the extended 3' nucleotide region. L: leader sequence; ORF1a and ORF1b: EAV nonstructural gene region; mTRS7: mutated TRS7 sequence.

[0034] FIGURES 6A-6B graphically summarize the results of experiments assessing effects of the incorporation of additional of 3' sequence into the monovalent base vector rE2-rFF. 6A). Introduction of 801 nucleotides of EAV sequence into the 3' region of the base vector rE2-rFF increases transfection efficiency. 6B). Introduction of 801 nucleotides of EAV sequence into the 3' region of the base vector rE2-rFF results in no change in expression level.

[0035] FIGURE 7 graphically describes the structure of three bivalent EAV replicon designs containing the extended 3' nucleotide region. L: leader sequence; ORF1a and ORF1b: EAV nonstructural gene region; mTRS7: mutated TRS7 sequence.

[0036] FIGURES 8A-8D graphically summarize the results of experiments demonstrating that the incorporation of additional of 3' UTR sequence into the bi-genic base vector rE2-gRen-rFF results in enhanced expression of genes of interest. The introduction of 801 nucleotides of EAV sequence into the 3' UTR region of the base vector rE2-gRen-rFF increases transfection efficiency (8A, 8B) and enhances luciferase production for both red Firefly (8C) and green Renilla (8D).

[0037] FIGURES 9A-9D graphically summarize the results of experiments assessing effect of various TRS7 mutations on replication and expression in replicons with

additional 3' sequences. 9A and 9B: Impact of exemplary TRS7 mutations on replication of rEnb and rExb bivalent replicons. 9C and 9D: Impact of exemplary TRS7 mutations on reporter gene expression from rEnb and rExb bivalent replicons.

[0038] FIGURES 10A-10D graphically illustrate the results of experiments assessing effect of restoring TRS7 element in rExb bivalent replicon with additional 3' sequences. 10A and 10B: Impact of restoring the TRS7 element on replication of rExb compared to rEnb bivalent replicons. 10C and 10D: Impact of restoring the TRS7 element on gene expression from rExb compared to rEnb bivalent replicons.

[0039] FIGURES 11A-11B graphically summarize the results of experiments assessing functionality of an EAV tri-genic expression design containing gRen, rFF and Cypr vs mono-genic and bi-genic replicon controls. Two types of data are shown. 11A). The top row of the figures shows the percent of cells transfected with each of the replicons; this is a measure of replication for each RNA. 11B). The lower row of the figures shows protein expression for each of the luciferase constructs from electroporated cell lysates and normalized for the amount of lysate used in the assay. The TRS used to drive each of the genes is shown for the trivalent construct. RLU: relative light units.

[0040] FIGURES 12A-12B graphically summarize the design of various spacer replicon constructs according to some exemplary embodiments described herein. 12A). Alignment of EAV sequences upstream and downstream of TRS7 subgenomic promoter. 12B). Nucleotides included in new Spacer replicon constructs.

[0041] FIGURES 13A-13B graphically summarize the results of experiments performed to demonstrate that addition of spacer sequences for TRS7 region insulation impacted replication and protein production relative to the rE2-rFF base vector. S1: Spacer 1 rE2-rFF. S2: Spacer 2 rE2-rFF, S3: Spacer 3 rE2-rFF, S4: Spacer 4 rE2-rFF, WT: rE2-rFF. The introduction of sequences upstream of the TRS7 (S1, S2) and downstream of the region (S3, S4) modified replication capacity (13A) and protein expression (13A, 13B).

[0042] FIGURES 14A-14B graphically describe the results of experiments assessing the impact on expression of bivalent EAV replicons containing added Spacer 1 sequences. S1: Spacer 1 region. 14A). Expression analysis of green Renilla luciferase in

different Spacer 1 bivalent replicons. 14B). Expression analysis of red firefly luciferase in different Spacer 1 bivalent replicons.

[0043] FIGURES 15A-15B graphically describe the results of experiments illustrating the impact of HA primary sequence on replication and expression. Flow cytometry analysis of cells electroporated with rE2 replicons coding for HA genes with different primary sequences. 15A). dsRNA and HA protein double positive cell populations detected from replicons with HA genes with differing primary sequence. 15B). HA protein expression MFI analysis (mean fluorescence intensity) of EAV replicons with HA genes having differing primary sequence.

[0044] FIGURES 16A-16B graphically describe the results of experiments illustrating the impact of F primary sequence on replication and expression. Flow cytometry analysis of cells electroporated with rE2 replicons coding for F genes with different primary sequences. 16A). dsRNA and F protein double positive cell populations detected from replicons with F genes with differing primary sequence. 16B). F protein expression MFI analysis of replicons with F genes with differing primary sequence.

[0045] FIGURES 17A-17C graphically summarize the results of experiments illustrating the robust green Renilla protein expression driven from EAV replicon. Flow cytometry analysis of cells electroporated with rE2 (EAV) and alphavirus (α) replicons coding for gREN luciferase. 17A). dsRNA and gREN protein double positive cell populations detected. 17B). gREN protein expression MFI analysis of electroporated cells. 17C). Bulk-cell luciferase assay analysis of electroporated cells.

[0046] FIGURE 18 graphically summarizes the results of experiments demonstrating the robust GFP protein expression driven from EAV replicon. Flow cytometry MFI analysis of cells electroporated with rE2 (EAV) and alphavirus (α) replicons expressing GFP reporter gene.

[0047] FIGURES 19A-19B illustrate an example of IVIS analysis in mice injected with rE2-rFF RNA. Analysis of expression from EAV replicons was carried out *in vivo* using whole body imaging analysis to detect rFF luciferase expression.

[0048] FIGURE 20 schematically depicts EAV genomic structure and genome expression strategy. The names of the replicase gene and structural protein genes are given

(references to the nomenclature of genes and proteins can be found in Snijder *et al.*, 2005). Below the genome organization, the structural relationships of the genome and *sg* mRNAs are depicted. The leader sequence and TRSs found at the 5' end of the EAV mRNAs are indicated as blue and orange boxes, respectively. The ribosomal frameshifting element (RFS) found in the genome-length mRNA1 is indicated and the translated region of each mRNA is highlighted by a green line, whereas translationally silent regions are indicated by a red line. Only the translated open reading frames are indicated for each mRNA. The right-hand panels show a typical pattern of EAV mRNAs isolated from infected cells, visualized by hybridization to a probe complementary to the 3' end of the genome and therefore recognizing all viral mRNA species.

[0049] FIGURE 21 summarizes the results of experiments analyzing luciferase expression from an EAV TRS1 replicon vector in BHK cells. BHK cells were electroporated with 3 µg of replicon RNA. The TRS1 replicon vector demonstrated robust expression that was higher than expression detected from an EAV replicon using the TRS7 subgenomic promoter.

[0050] FIGURE 22 is a plasmid map of the VBS-R-eGFP construct.

[0051] FIGURE 23 is a plasmid map of the VBS-IC construct.

[0052] FIGURE 24 is a plasmid map of the pBR322+VBS-R-eGFP construct.

[0053] FIGURE 25 is a pictorial summary of the results of experiments performed to demonstrate functionality of VBS IC construct in BHK cells. BHK cells were electroporated with 3 µg of either EAV strain 030 IC RNA (EAV030), EAV strain VS IC RNA (EAV-VBS) or no RNA (Mock). Cells were examined for the presence of CPE at 24 and 48 hours post electroporation. CPE was noted in both the EAV030 and EAV-VBS electroporated cells by 48 hours, demonstrating that both IC were functional.

[0054] FIGURE 26 schematically summarizes of the results of experiments performed to analyze eGFP expression from a VBS-based replicon vector in BHK cells. BHK cells were electroporated with 3 µg of either EAV strain EAV030 eGFP replicon RNA (rE2-GFP) or EAV strain VBS eGFP replicon RNA (VBS-R-TRS2-eGFP). Cells were examined for the relative expression of GFP by FACS analysis. The EAV VBS-based replicon was found to express GFP protein at the same level as the EAV030-based replicon.

[0055] FIGURE 27 is a plasmid map of the pBR322+VBS-R-rFF construct.

[0056] FIGURE 28 is a plasmid map of the pBR322+VBS-R-TRS7-rFF construct.

[0057] FIGURE 29 schematically summarizes of the results of experiments performed to analyze rFF expression from a VBS-based replicon vector in BHK cells. BHK cells were electroporated with 3 µg of either EAV strain 030 rFF replicon RNA (rE2-GFP) or EAV strain VBS-rFF replicon RNAs (VBS-R-TRS2-rFF or VBS-R-TRS7-rFF). Cells were examined for the expression of rFF by FACS analysis and bulk luciferase assay. Each of the EAV VBS-based replicon was observed to express rFF protein at similar levels when compared to each other.

[0058] FIGURE 30 is a schematic summary of the results of experiments performed to analyze rFF expression from a VBS-based replicon vector in Balb/c mice. In this experiment, mice were intramuscularly injected with 30, 60 or 90 µg of VBS-R-TRS2-rFF in ringers lactate. Animals were examined by whole body imaging one and three days post RNA injection.

[0059] FIGURE 31 is a plasmid map of the pW70+SHFV-R-TRS7-rFF construct.

[0060] FIGURE 32 is a schematic summary of experiments performed to analyze rFF expression from a SHFv-based replicon vector in BHK cells. BHK cells were electroporated with 3 µg of SHFv-R-TRS7-rFF replicon RNA. Cells were examined for the expression of rFF by FACS analysis and bulk luciferase assay. The SHFv-based replicon was observed to express rFF protein in BHK cells.

[0061] FIGURES 33A-33C pictorially summarize of the results of experiments performed to analyze antibody expression from EAV bivalent replicon vector in BHK cells. BHK cells were electroporated with 3 µg of rEx-herceptin (SGI-RNA-Ab) replicon RNA. FIGURE 33A schematically illustrates a molecular design of the rEx-herceptin replicon analyzed. FIGURE 33B summarizes the results of an ELISA analysis of secreted antibody from electroporated cells at ~24 hours post electroporation, which was performed to compare DNA and EAV replicon expressed antibody (mg/L). FIGURE 33C summarizes the results of experiments performed to demonstrate that rEx-herceptin expressed antibody can detect Her2 antigen

[0062] FIGURE 34 pictorially summarizes the results of experiments performed to further analyze the SHFv-R-TRS7-rFF replicon *in vivo* in Balb/c mice. In this experiment, 30 µg of RNA was injected into mice and whole body imaging was conducted. These data demonstrate the *in vivo* activity of the SHFv replicon vector and that it is equivalent to the EAV replicon.

[0063] FIGURES 35A-35D are a schematic summary of the results of experiments performed to analyze mouse IL-12 and RSV F expression from monovalent and bivalent replicon vectors in BHK cells. BHK cells were electroporated with 3 µg of each replicon RNA. Cells were examined for the expression of IL-12 or RSV F by FACS analysis using protein-specific antibodies. Percent cells transfected was determined using dsRNA-specific antibody. FIGURE 35A: percent cells transfected with bivalent IL-12-RSV F or IL-12 RNA. FIGURE 35B: mean fluorescence intensity (MFI) of IL-12 specific protein expression. FIGURE 35C: percent cells transfected with bivalent IL-12-RSV F or RSV F RNA. FIGURE 35D: mean fluorescence intensity (MFI) of RSV F specific protein expression.

[0064] FIGURES 36A-36C are a pictorial summary of the results of experiments performed to analyze cas9 expression and cutting activity from an EAV replicon. FIGURE 36A: Schematic of the EAV cas9 replicon vector. FIGURE 36B: BHK cells were electroporated with 3 µg of each replicon RNA. The percent of cells transfected was determined using dsRNA-specific antibody. The cas9 gene was synthesized using two different codon usages. FIGURE 36C: Cas9 functionality was determined *in vitro* using electroporated cell lysates combined with plasmid DNA and gRNA specific for the plasmid sequence. Specific DNA cleavage was detected in samples generated from cas9 EAV replicon RNA electroporated cell lysates.

[0065] FIGURE 37 is a schematic EAV genome.

[0066] 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 OF CERTAIN EMBODIMENTS

[0067] The present disclosure generally relates to the development of arterivirus expression systems which are suitable for expressing multiple heterologous genes in recombinant cells in a tunable manner. In some embodiments, the disclosure relates to nucleic acid molecules containing a genetically engineered arterivirus genome or replicon RNA. For example, some embodiments relate to nucleic acid molecules including a recombinant arterivirus genome or replicon RNA in which at least some of its original nucleotide sequence encoding one or more structural proteins has been removed and/or replaced with a heterologous nucleotide sequence encoding, for example, a polypeptide of interest.

[0068] As disclosed herein, monogenic or multigenic arterivirus expression systems can be generated by removing a part or the entire coding region for one or more structural proteins of the subgenomic RNAs (*sg* RNAs) of EAV, and replacing each with coding sequence of a gene of interest (GOI). Because each *sg* RNA of EAV is naturally expressed at a distinct level, the arterivirus expression systems disclosed herein allows tunable expression of a GOI in either a monogenic or multigenic design. For example, as described in further detail below, the development of a single RNA expression platform capable of expressing multiple GOIs, each of which can be produced at the same or different levels relative to one another, represents significant utility and practicality. Among other advantages, this tunability of the expression platforms disclosed herein advantageously allows for tunable expression of multiple proteins in the same cell. As an example, in immunological applications, clonal expression of multiple proteins in the same host cell offers many practical utilities, including but not limited to: 1) it allows for complex protein interactions to occur to support conformationally accurate epitopes, *in vivo* and/or *ex vivo*, 2) it supports co-expression of immunomodulatory proteins along with, *e.g.* concurrently with, vaccines and/or tumor associated antigens, which in turn may illicit and/or drive robust immune responses, and 3) it supports advanced therapeutic approaches requiring multiple

protein expression. Without being limited by any particular theory, it is believed that another non-limiting advantage of using arteriviruses as viral expression vectors is that they can direct the synthesis of large amounts of heterologous proteins in recombinant host cells.

[0069] 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.

[0070] 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 skilled in the art.

Some Definitions

[0071] 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. As used in this disclosure and the appended claims, the term “and/or” can be singular or inclusive. For example, “A and/or B” is used herein to include all of the following alternatives: “A”, “B”, “A or B”, and “A and B”.

[0072] 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.

[0073] 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.

[0074] 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.

[0075] 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.

[0076] 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 (tumor Ag), parasite (malaria), *etc.*)

[0077] 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.

[0078] Nucleic acid molecules of the present disclosure can be nucleic acid molecules of any length, including nucleic acid molecules that are preferably between about 5 Kb and about 50 Kb, for example 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.

[0079] 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.

[0080] 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.

[0081] As used herein, the term "replicon" refers to a viral nucleic acid that is capable of directing the generation of copies of itself. As used herein, the term "replicon" includes RNA as well as DNA. For example, double-stranded DNA versions of arterivirus genomes can be used to generate a single-stranded RNA transcript that constitutes an arterivirus replicon. Generally, a viral replicon contains the complete genome of the virus. "Sub-genomic replicon," as used herein, refers to a viral nucleic acid that contains something less than the full complement of genes and other features of the viral genome, yet is still capable of directing the generation of copies of itself. For example, the sub-genomic replicons of arterivirus described below contain most of the genes for the non-structural proteins of the virus, but are missing most of the genes coding for the structural proteins. Sub-genomic replicons are capable of directing the expression of all of the viral genes necessary for the replication of the viral sub- genome (replication of the sub-genomic replicon), without the production of viral particles.

[0082] A "vector" as used herein refers to any means for the transfer of a nucleic acid into a host cell. A vector may be a replicon to which another DNA segment may be attached so as to bring about the replication of the attached segment. The term "vector" includes both viral and non-viral means for introducing the nucleic acid into a cell *in vitro*, *ex vivo* or *in vivo*. Non-viral vectors include, but are not limited to plasmids, liposomes, electrically charged lipids (cytofectins), DNA-protein complexes, and biopolymers. In addition to a nucleic acid, a vector may also contain one or more regulatory regions, and/or selectable markers useful in selecting, measuring, and monitoring nucleic acid transfer results (transfer to which tissues, duration of expression, *etc.*).

[0083] 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.

[0084] In some embodiments of the methods or processes described herein, the steps can be carried out in any order, except when a temporal or operational sequence is explicitly recited. Furthermore, in some embodiments, the specified steps can be carried out concurrently unless explicit claim language recites that they be carried out separately. For example, in some embodiments a claimed step of doing X and a claimed step of doing Y can be conducted simultaneously within a single operation, and the resulting process will fall within the literal scope of the claimed method.

[0085] As used herein, “comprising” is synonymous with “including,” “containing,” or “characterized by,” and is inclusive or open-ended and does not exclude additional, unrecited elements or method steps. As used herein, “consisting of” excludes any elements, steps, or ingredients not specified in the claimed composition or method. As used herein, “consisting essentially of” does not exclude materials or steps that do not materially affect the basic and novel characteristics of the claimed composition or method. Any recitation herein of the term “comprising”, particularly in a description of components of a composition or in a description of steps of a method, is understood to encompass those compositions and methods consisting essentially of and consisting of the recited components or steps.

[0086] 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 application, and are to be included within the spirit and purview of this application.

Arteriviruses

[0087] The arteriviruses belong to the genus *Arterivirus* in the family *Arteriviridae*, which is within the order Nidovirales, and encompass an important group of

enveloped, single-stranded, positive-sense RNA viruses which infect domestic and wild animals.

[0088] The order Nidovirales can be divided into two clades depending on the size of the genome: those with large genomes (26.3–31.7 kilobases) which included the *Coronaviridae* and *Roniviridae* (the large nidoviruses) and those with small genomes (the small nidoviruses)—a clade that includes the distantly related *Arteriviridae* (12.7–15.7 kb). The large nidoviruses encode both an 2'-O-methyltransferase and a 3'-5' exoribonuclease (ExoN)—the latter being very unusual for an RNA virus. They also encode a superfamily 1 helicase, uridylate-specific endonuclease (an enzyme unique to nidoviruses) and several proteases.

[0089] It has been well documented that although arteriviruses share a similar genome organization and replication strategy to that of members of the family *Coronaviridae* (genera *Coronavirus* and *Torovirus*), they do differ considerably in their genetic complexity, genome length, biophysical properties, size, architecture, and structural protein composition of the viral particles (*e.g.*, virion). Currently, the *Arterivirus* genus is considered to include equine arteritis virus (EAV), porcine reproductive and respiratory syndrome virus (PRRSV), lactate dehydrogenase-elevating virus (LDV) of mice, simian hemorrhagic fever virus (SHFV), and wobbly possum disease virus (WPDV). Recent studies have reported that the newly identified wobbly possum disease virus (WPDV) also belongs to the *Arterivirus* genus.

[0090] A typical arterivirus genome varies between 12.7 and 15.7 kb in length but their genome organization is relatively consistent with some minor variations. Exemplary genome organization and virion architecture of an arterivirus is shown in **FIG. 20**. The arterivirus genome is a polycistronic +RNA, with 5' and 3' non-translated regions (NTRs) that flank an array of 10–15 known ORFs. The large replicase ORFs 1a and 1b occupy the 5'-proximal three-quarters of the genome, with the size of ORF1a being much more variable than that of ORF1b. Translation of ORF1a produces replicase polyprotein (pp) 1a, whereas ORF1b is expressed by –1 programmed ribosomal frameshifting (PRF), which C-terminally extends pp1a into pp1ab. In addition, a short transframe ORF has been reported to overlap the nsp2-coding region of ORF1a in the +1 frame and to be expressed by –2 PRF. The 3'-proximal genome part has a compact organization and contains 8 to 12 relatively small genes,

most of which overlap with neighboring genes. These ORFs encode structural proteins and are expressed from a 3'-co-terminal nested set of sg mRNAs. The organization of these ORFs is conserved, but downstream of ORF1b, SHFV and all recently identified SHFV-like viruses contain three or four additional ORFs (~1.6 kb) that may be derived from an ancient duplication of ORFs 2–4. Together with the size variation in ORF1a, this presumed duplication explains the genome size differences among arteriviruses.

[0091] With regard to equine arteritis virus (EAV), the wild-type EAV genome is approximately 12.7 Kb in size. The 5' three fourths of the genome codes for two large replicase proteins 1a and 1ab; the amino acid sequences of the two proteins are N-terminally identical but due to a ribosomal frameshift the amino acid sequence of the C-terminal region of 1ab is unique. The 3' one quarter of the EAV genome codes for the virus's structural protein genes, all of which are expressed from subgenomic RNAs. The subgenomic RNAs form a nested set of 3' co-terminal RNAs that are generated via a discontinuous transcriptional mechanism. The subgenomic RNAs are made up of sequences that are not contiguous with the genomic RNA. All of the EAV subgenomic RNAs share a common 5' leader sequence (156 to 221 nt in length) that is identical to the genomic 5' sequence. The leader and body parts of the subgenomic RNAs are connected by a conserved sequence termed a transcriptional-regulatory sequence (TRS). The TRS is found on the 3' end of the leader (leader TRS) as well as in the subgenomic promoter regions located upstream of each structural protein gene (body TRS). Subgenomic RNAs are generated as the negative strand replication intermediate RNA is transcribed. As transcription occurs the replication complex pauses as it comes to each body TRS and then the nascent negative strand RNA become associated with the complementary positive strand leader TRS where negative strand RNA transcription continues. This discontinuous transcription mechanism results in subgenomic RNA with both 5' and 3' EAV conserved sequences. The negative strand subgenomic RNAs then become the template for production of the subgenomic positive sense mRNA.

[0092] Infectious cDNA clones, representing the entire genome of EAV, have been reported (van Dinten 1997; de Vries *et al.*, 2000, 2001; Glaser *et al.*, 1999) and they been used to study EAV RNA replication and transcription for nearly two decades (van Marle 1999, van Marle 1999a, Molenkamp 2000, Molenkamp 2000a, Pasternak 2000, Tijms 2001,

Pasternak 2001, Pasternak 2003, Pasternak 2004, van den Born 2005, Beerens & Snijder 2007, Tijms 2007, Kasteren 2013). In addition, infectious clones have been generated that contain the chloramphenicol acetyltransferase (CAT) gene inserted in place of ORF2 and ORF7 and CAT protein was shown to be expressed in cells electroporated with those RNAs (van Dinten 1997, van Marle 1999). Modifications of the infectious clone via site directed mutagenesis and deletion of the structural protein gene regions has been used to determine the requirement for each structural gene in support of RNA replication (Molenkamp 2000). The study reported by Molenkamp 2000 concluded that the structural genes are not required to support RNA replication. Analysis of sequence homology requirements for TRS activity in subgenomic RNA production was conducted and used to better define how discontinuous transcription mechanistically occurs (van Marle 1999, Pasternak 2000, Pasternak 2001, Pasternak 2003, van den Born 2005) and defective interfering RNAs have been used to understand the minimal genomic sequences required for replication and packaging of RNA into virus particles (Molenkamp 2000a). However, no attempt to construct a replicon vector from EAV capable of and designed specifically to efficiently express heterologous genes has been reported.

[0093] Development of an EAV replicon vector for expression of heterologous genes has not been conducted prior to the inventive work described herein because most other replicon systems have been focused on virus particle-based approaches. That is, packaging of a replicon RNA into a virus-like particle by supplying the deleted structural proteins back in-trans. Because EAV has two major and five minor structural proteins, the level of expression of each which is key to efficient virus particle production, it is considered too difficult to develop a virus particle-based system from EAV. For at least this reason, no attempt to develop an EAV replicon RNA as a vector has been conducted prior to the inventive work described herein.

[0094] The inventive work described herein is primarily based on the EAV RNA replicon and is not dependent on the formation of recombinant virus-like particle. Accordingly, the presently disclosed compositions and methods are not limited by the complexity of providing the EAV structural proteins back in order to produce a virus-like particle. In addition, because each subgenomic RNA is naturally expressed at a unique level

the system also allows tunable expression of a GOI in either a monogenic or multigenic design. There is significant utility of a single RNA capable of expressing multiple GOI, each at the same or different levels relative to one another. This capacity allows for expression of multiple proteins in the same cell. In addition to the unique design of the EAV replicon, development of methods that can be employed to tune, *e.g.* to modulate protein expression from the vector is inventive. As used herein, the term "tunable expression" refers to the ability of the compositions and methods described herein to control the level of expression of a GOI operably linked to an arterivirus replicon according to the present disclosure. This can be achieved, for example, by a number of ways. For approaches that employ use of the system launched from a DNA plasmid from the nucleus of a cell the transcription of the replicon RNA could be controlled/tuned by an inducible DNA dependent RNA polymerase promoter. For example, transcription of RNA could be controlled using Tet technology to induce production of the replicon RNA from transformed cells. Once, replicon RNA is in the cytoplasm of a cell (either by transfecting the RNA into cells or having the cell produce the RNA from an integrated DNA version of the system) a number of additional techniques can be used to tune expression from the system. The first example 1) can utilize RNA structure-seq and next generation sequencing (NGS) to understand the secondary structure surrounding TRS elements. In some embodiments, this information can provide insights into formulating approaches to tune, *e.g.* to modulate, the activity of any TRS and inform tunable GOI expression from the EAV replicon. This information combined with the 2) relative position on the genomic RNA that the GOI is placed, 3) controlling secondary structure of each GOI by utilizing sequence optimization of the primary gene sequence, and 4) by optimizing 3' sequences key to optimal replication of replicons. For example, by optimizing the 3' EAV sequences included in the replicon and by inclusion of a longer polyA region. Employing all of these approaches enables tunable protein expression. The methods, compositions, and systems described herein can be used for tunable protein expression, for example, they can be used to express one or more proteins in various expression levels. In some embodiments, the methods, compositions, and systems described herein can be used for expressing one or more proteins at about 0.1%, 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%,

100%, 150%, 200%, 300%, 400%, 500%, 10-fold, or more of a reference expression level of the protein(s).

[0095] As described in further detail below, the present disclosure also relates to significant showing of unexpected results in connection with various arterivirus expression vectors designed and evaluated by the inventors. In a study published by Molenkamp *et al.* (2000), it has been indicated that the sequence encoding open reading frame ORF2a of EAV is needed in order to retain robust TRS2 subgenomic transcription activity (Molenkamp *et al.* 2000). In fact, Molenkamp *et al.* showed that an EAV infectious clone deleted from nucleotide residues 9,756 to 12,351, which retained only 5 bases of the ORF2a sequence, exhibited a significant reduction in subgenomic RNA synthesis. This property was in contrast to the subgenomic RNA synthesis demonstrated from a different EAV infectious clone (mutant 030-2319; also referred to as EAV030 mutant); mutant 030-2319 contained the same 3' sequences as mutant 2a-2594 but maintained an intact ORF 2a sequence and this construct demonstrated wild-type robust subgenomic RNA synthesis (Molenkamp *et al.* 2000).

[0096] Surprisingly and in contrast to these teachings of Molenkamp *et al.*, in the replicon design of disclosed herein (for example Rep-EAV (WT)) and all derivative versions of the replicon, the ORF2a sequence is completely absent, yet each of the replicons exhibits robust subgenomic transcription and high expression of one or more GOIs. Furthermore, Molenkamp *et al.* 2000 define the optimal 3' terminal sequences that should be maintained for efficient replication using mutant 030-2319 as well. Molenkamp *et al.* teach that the 3' terminal 354 nt of EAV are able to support wild type replication. Surprisingly, as described and demonstrated herein, replicon vectors that code for at least two GOIs did not replicate efficiently unless significantly more 3' terminal sequences are included. In addition, significant differences in both replication and protein expression were noted from replicons coding for the same protein but having different primary GOI sequences. More than a 50-fold difference in replication activity and a 2 to 4 fold difference in protein expression have been observed in replicons coding for the same GOI with different primary nucleotide sequences. Another non-limiting unexpected aspect of the methods, composition and systems described herein for protein expression is the magnitude of protein expression that

they are capable of. It is well known in the RNA replicon field that alphavirus-based replicon systems are capable of expressing up to twenty percent of a cell's total protein content (Pushko *et al* 1997). Thus, it is surprising that the methods, arterivirus-based composition and systems described herein are capable of even higher expression levels on a per cell basis than an alphavirus replicon.

Nucleic Acid Molecules of the Disclosure

[0097] In one aspect, novel nucleic acid molecules which include a nucleotide sequence encoding a modified arterivirus genome or replicon RNA are disclosed. For example, a modified arterivirus genome or replicon RNA can comprise deletion(s), substitution(s), and/or insertion(s) in one or more of the genomic regions (*e.g.*, open reading frames (ORFs)) of the parent arterivirus genome. In some embodiments, one or more of arterivirus ORF2a, ORF2b, ORF3, ORF4, ORF5, and ORF5a are absent and/or modified in the modified arterivirus genome or replicon RNA. In some embodiments, the modified genome or replicon RNA includes a sequence fragment exhibiting at least 80%, at least 85%, preferably at least 90%, or more preferably at least 95% identity to a nucleotide sequence encoding open reading frame ORF7. In some embodiments, the modified genome or replicon RNA includes a sequence fragment exhibiting at least 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to a nucleotide sequence encoding open reading frame ORF7. In some embodiments, the modified genome or replicon RNA includes a sequence fragment exhibiting 100% sequence identity to a nucleotide sequence encoding open reading frame ORF7. In some embodiments, the modified arterivirus genome or replicon RNA is devoid of at least a portion of the sequence encoding one or more of the open reading frames ORF2b, ORF3, ORF4, ORF5, and ORF5a. For example, the modified arterivirus 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 open reading frames ORF2b, ORF3, ORF4, ORF5, and ORF5a. In some embodiments, the modified arterivirus genome or replicon RNA is devoid of the entire sequence encoding one or more of the open reading frames ORF2b, ORF3, ORF4, ORF5, and ORF5a. In some embodiments, the modified arterivirus genome or replicon RNA is devoid of a portion of or the entire sequence encoding one of the open reading frames ORF2b, ORF3, ORF4, ORF5, and ORF5a. In some

embodiments, the modified arterivirus genome or replicon RNA is devoid of a portion of or the entire sequence encoding two of the open reading frames ORF2b, ORF3, ORF4, ORF5, and ORF5a. In some embodiments, the modified arterivirus genome or replicon RNA is devoid of a portion of or the entire sequence encoding three of the open reading frames ORF2b, ORF3, ORF4, ORF5, and ORF5a. In some embodiments, the modified arterivirus genome or replicon RNA is devoid of a portion of or the entire sequence encoding four of the open reading frames ORF2b, ORF3, ORF4, ORF5, and ORF5a. In some embodiments, the modified arterivirus genome or replicon RNA is devoid of a portion of or the entire sequence encoding all the open reading frames ORF2b, ORF3, ORF4, ORF5, and ORF5a. In some embodiments, the modified arterivirus genome or replicon RNA is devoid of ORF2b, ORF3, ORF4, and ORF5. In some embodiments, the modified arterivirus genome or replicon RNA is devoid of at least a portion of ORF6, for example the first one, two, three, four, five, six, seven, eight, nine, ten, or more nucleotides of ORF6. "Fragment", as used herein with respect to a polynucleotide, refers to a clone or any part of a polynucleotide molecule, particularly a part of a polynucleotide that retains a usable, functional characteristic. For example, a "polynucleotide fragment" refers to any subsequence of a polynucleotide, typically, of at least about 9 consecutive nucleotides, for example at least about 30 nucleotides, at least about 50 nucleotides, at least about 100 nucleotides, at least about 200 nucleotides, or at least about 300 nucleotides of any of the sequences provided herein. Exemplary polynucleotide fragments are the first sixty consecutive nucleotides (*e.g.*, starting from the 5'-end or from the 3'-end) of the polynucleotides disclosed herein.

[0098] Nucleic acid fragments 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 a sequence encoding open reading frame ORF7 of an arterivirus of interest can be identified and/or isolated by using the sequences identified herein (*e.g.*, SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 40, SEQ ID NO: 41, and SEQ ID NO: 42) or any others as they are known in the art, for example, the sequences having GenBank accession numbers NC_002532 (EAV), NC_001961.1 (PRRSV), NC_003092 (SHFV), and NC_001639.1 (LDV), by genome sequence analysis, hybridization, and/or PCR with degenerate primers or gene-specific primers from sequences identified in the respective

arterivirus 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.

[0099] In some embodiments, a nucleic acid molecule disclosed herein comprises one or more nucleic acid fragments that specifically hybridize to a nucleic acid sequence encoding open reading frame ORF7 of an arterivirus; and complements of said nucleic acid sequences; and fragments of either, under low, moderate, or high stringency conditions. In a particular embodiment, nucleic acid molecules of the present application preferably include a nucleic acid sequence that hybridizes high stringency conditions, to a nucleic acid sequence encoding open reading frame ORF7 of an arterivirus; and complements of said nucleic acid sequences; and fragments of either.

[0100] In some embodiments, the nucleic acid molecules disclosed herein include a modified arterivirus genome or replicon RNA which is devoid of the sequence encoding a portion of or the entire open reading frame ORF2a. In some embodiments, the nucleic acid molecules disclosed herein includes a modified arterivirus genome or replicon RNA which is devoid of the ATG start codon of the sequence encoding open reading frame ORF7. In some embodiments, the nucleic acid molecules disclosed herein include a modified arterivirus genome or replicon RNA which is devoid of a portion of or the entire sequence encoding open reading frame ORF6. For example, the modified arterivirus genome or replicon RNA can be devoid of the ATG start codon of ORF6. In some embodiments, about 1%, 5%, 8%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of ORF6 is absent in the modified arterivirus genome or replicon RNA. In some embodiments, TRS7 within ORF6 is deleted or modified in the modified arterivirus genome or replicon RNA to reduce or abolish its activity. In some embodiments, about 1%, 5%, 8%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of ORF7 is absent in the modified arterivirus genome or replicon RNA.

[0101] The molecular techniques and methods by which these new nucleic acid molecules were constructed and characterized are described more fully in the examples herein.

[0102] In some embodiments, the nucleic acid molecules disclosed herein are recombinant nucleic acid molecules. As used herein, the term recombinant means any molecule (*e.g.* DNA, RNA, *etc.*), that is, or results, however indirect, from human manipulation of a polynucleotide. 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.

[0103] Preferably, 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.

[0104] 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.

[0105] In various embodiments disclosed herein, the nucleic acid molecule disclosed herein can include one or more of the following feature. In some embodiments, the nucleic acid molecule disclosed herein includes a modified arterivirus genome or replicon RNA including one or more subgenomic (*sg*) promoters at a non-native site, wherein each of the one or more *sg* promoters includes a transcriptional regulatory sequence (TRS).

[0106] The term "subgenomic promoter", as used herein, refers to a promoter of a subgenomic mRNA of a viral nucleic acid. As used herein, an "arterivirus subgenomic promoter" is a promoter as originally defined in a wild type arterivirus genome that directs transcription of a subgenomic messenger RNA as part of the arterivirus replication process. An arterivirus subgenomic (*sg*) promoter typically includes a conserved transcriptional regulatory sequence (TRS) with 5'- and 3' flanking sequences. Based on the particular arterivirus open reading frame (ORF) that it drives expression, the subgenomic promoter of the disclosure can be, for example, *sg* promoter 1 (which comprises arterivirus TRS1 or a variant thereof), *sg* promoter 2 (which comprises arterivirus TRS2 or a variant thereof), *sg* promoter 3 (which comprises arterivirus TRS3 or a variant thereof), *sg* promoter 4 (which

comprises arterivirus TRS4 or a variant thereof), *sg* promoter 5 (which comprises arterivirus TRS5 or a variant thereof), *sg* promoter 6 (which comprises arterivirus TRS6 or a variant thereof), *sg* promoter 7 (which comprises arterivirus TRS7 or a variant thereof), or a variant thereof. In some embodiments described herein, the nucleic acid molecules of the application can include a *sg* promoter that is essentially devoid of a 5'-flanking sequence and/or a 3'-flanking sequence. In some embodiments, the nucleic acid molecules of the disclosure can include a *sg* promoter that consists of a conserved transcriptional regulatory sequence (TRS), that is, such *sg* promoter does not include any flanking sequence. In some embodiments, a *sg* promoter can have a wild type sequence or a sequence that has been modified from wild type sequence but retains promoter activity.

[0107] Accordingly, in some embodiments, at least one of the one or more *sg* promoters includes a sequence exhibiting at least 80%, at least 85%, preferably at least 90%, or more preferably at least 95% identity sequence identity to a sequence selected from the group consisting of *sg* promoter 1, *sg* promoter 2, *sg* promoter 3, *sg* promoter 4, *sg* promoter 5, *sg* promoter 6, *sg* promoter 7, and a variant thereof. In some embodiments, at least one of the one or more *sg* promoters includes a sequence exhibiting 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to *sg* promoter 1. In some embodiments, at least one of the one or more *sg* promoters includes a sequence exhibiting 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to *sg* promoter 2. In some embodiments, at least one of the one or more *sg* promoters includes a sequence exhibiting 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to *sg* promoter 3. In some embodiments, at least one of the one or more *sg* promoters includes a sequence exhibiting 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to *sg* promoter 4. In some embodiments, at least one of the one or more *sg* promoters includes a sequence exhibiting 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to *sg* promoter 5. In some embodiments, at least one of the one or more *sg* promoters includes a sequence exhibiting 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to *sg* promoter 6. Such variants of the *sg* promoters may be naturally-occurring, including homologous polynucleotides from the same or a different species, or may be non-natural variants, for example polynucleotides synthesized using chemical synthesis methods, or generated using recombinant DNA techniques. Accordingly, in some embodiments, at least one of the one or

more *sg* promoters is a modified *sg* promoter. The modified *sg* promoter can generally be any modified *sg* promoter, and can be, for example, a modified *sg* promoter 1, a modified *sg* promoter 2, a modified *sg* promoter 3, a modified *sg* promoter 4, a modified *sg* promoter 5, a modified *sg* promoter 6, a modified *sg* promoter 7. In some particular embodiments, at least one of the one or more modified *sg* promoters is a modified *sg* promoter 7. In some embodiments, the nucleic acid molecules disclosed herein can include a modified arterivirus genome or replicon RNA which includes one or more modified *sg* promoters located at their respective native site, wherein each of the one or more modified *sg* promoters includes a TRS. In addition or alternatively, in some embodiments disclosed herein, at least one of the one or more modified *sg* promoter includes a nucleotide modification positioned within the sequence of the TRS. In some exemplary embodiments, at least one of the one or more modified *sg* promoters includes a leader TRS or a variant thereof. In some exemplary embodiments, at least one of the one or more modified *sg* promoters includes a body TRS or a variant thereof. In some embodiments, the leader TRS or a variant thereof and the body TRS or a variant thereof do not have the same sequence. In some embodiments, the nucleotide sequence of the leader TRS is not modified.

[0108] Alternatively or in addition, in some embodiments, the nucleic acid molecules disclosed herein can comprise at least one of the one or more modified *sg* promoter including one or more nucleotide modifications which are positioned within the primary sequence required for the formation of a secondary structure of RNA transcripts including the respective *sg* promoter sequence. In some embodiments, the secondary structure of RNA transcripts can include a hairpin structure. In some embodiments, the one or more nucleotide modifications are positioned within the leader TRS hairpin (LTH). In some embodiments, the nucleotide modifications positioned within the primary sequence of the hairpin structure involve in a conformational RNA switch in the 5'proximal region of the modified arterivirus genome or replicon RNA. In some embodiments, the nucleotide modifications positioned within the primary sequence of the hairpin structure modulate the production of one or more or all *sg* mRNA of the modified arterivirus genome or replicon RNA.

[0109] Further, in some embodiments disclosed herein, the modified arterivirus genome or replicon RNA can include one or more mutated T7 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. In accordance with some exemplary embodiments, at least one of the one or more T7 mutated transcriptional termination signal sequences includes a nucleotide substitution at a position selected from the group consisting of T9001, T3185, G3188, and combinations thereof. In some embodiments, the nucleotide substitution at position T9001 includes T9001G. In some embodiments, the nucleotide substitution at position T3185 includes T3185A. In some embodiments, the nucleotide substitution at position G3188 includes G3188A. In some embodiments, at least one of the one or more T7 mutated transcriptional termination signal sequences includes a nucleotide substitution at a position selected from the group consisting of T9001G, T3185A, G3188A, and combinations of any two or more thereof.

[0110] In some embodiments, the modified arterivirus genome or replicon RNA as disclosed herein includes one or more heterologous transcriptional termination signal sequences. The heterologous transcriptional termination signal sequences can generally be any heterologous transcriptional termination signal sequences, and can be, for example, SP6 termination signal sequence, a T3 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, or a variant thereof. In some embodiments, at least one of the one or more heterologous transcriptional termination signal sequences is inactivated.

[0111] In various embodiments disclosed herein, the nucleic acid molecules can include one or more spacer regions. The spacer region can generally be any spacer region, and can be, for example, a spacer region that is operably positioned adjacent to at least one of the one or more *sg* promoters. In some embodiments disclosed herein, at least one of the one or more spacer regions is positioned immediately 3' to a *sg* promoter. In some embodiments, at least one of the one or more spacer regions is positioned immediately 5' to a *sg* promoter.

In principle, the sequence of the spacer regions can be of any length, and can be, for example, about 20 to 400 nucleotides in length. In some embodiments, the sequence of the spacer regions is about 18 to 420, about 20 to 350, about 30 to 200, about 50 to 200 nucleotides in length. In some embodiments, the sequence of the spacer regions is about 100 to 300, about 200 to 350, about 300 to 350 nucleotides in length. In some embodiments, the sequence of the spacer regions is about 23 nucleotides in length. In some embodiments, the sequence of the spacer regions is about 98 nucleotides in length. In some embodiments, the sequence of the spacer regions is about 220 nucleotides in length. In some embodiments, the sequence of the spacer regions is about 343 nucleotides in length.

[0112] In some further embodiments, the nucleic acid molecules disclosed herein can include one or more 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”. The term "expression cassette" as used herein, refers to a nucleic acid construct that encodes a protein or functional RNA operably linked to expression control elements, such as a promoter, and optionally, any or a combination of other nucleic acid sequences that affect the transcription or translation of the gene.

[0113] In some embodiments, the nucleic acid molecules disclosed herein can include one or more expression cassettes, each of which includes a *sg* promoter operably linked to a heterologous nucleotide sequence. 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.

[0114] 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).

[0115] In some embodiments disclosed herein, the *sg* promoter can include a transcription regulatory sequence (TRS) and, optionally, one or more flanking regions. A flanking region can be generally of any length, and can be, for example, about 5 to 400 nucleotides in length. In some embodiments, the flanking region can be about 5 to 350, about 10 to 300, about 20 to 200, about 50 to 150, about 50 to 100 nucleotides in length. In some embodiments, the flanking region can be about 5 to 100, about 10 to 150, about 15 to 115, about 20 to 300, about 50 to 350, about 100 to 350 nucleotides in length. In some embodiments, the flanking region can be, or be about, 17 nucleotides in length. In some embodiments, the flanking region can be, or be about 17, 23, 25, 56, 73, or 112 nucleotides in length. In some embodiments, the flanking region can be, or be about, 25 nucleotides in length. In some embodiments, the flanking region can be, or be about 56 nucleotides in length. In some embodiments, the flanking region can be, or be about, 73 nucleotides in length. In some embodiments, the flanking region can be, or be about, 112 nucleotides in length. In some embodiments, the *sg* promoter can include a flanking region positioned 5' to the TRS. In some embodiments, the *sg* promoter can include a flanking region positioned immediately 5' to the TRS. In some embodiments, the *sg* promoter can include a flanking region positioned 3' to the TRS. In some embodiments, the *sg* promoter can include a flanking region positioned immediately 3' to the TRS. In some embodiments, the *sg* promoter can include 5' flanking region and a 3' flanking region. In some embodiments, the

5' flanking region and a 3' flanking region can be of the same length. In some embodiments, the 5' flanking region and a 3' flanking region can differ in their respective length. In some particular embodiments, the 5' flanking region can be, or be about, 23 nucleotides in length. In some embodiments, the 3' flanking region can be, or be about, 17, 25, 56, 73, or 112 nucleotides in length. In some particular embodiments, the 3' flanking region of the *sg* promoter 3 can be, or be about, 73 nucleotides in length. In some embodiments, the 3' flanking region of the *sg* promoter 4 can be, or be about, 17 nucleotides in length. In some embodiments, the 3' flanking region of the *sg* promoter 5 can be, or be about, 112 nucleotides in length. In some embodiments, the 3' flanking region of the *sg* promoter 6 can be, or be about, 25 nucleotides in length.

[0116] In some embodiments, the nucleic acid molecules disclosed herein can include more than one expression cassette. 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.

[0117] Accordingly, the nucleic acid molecules as provided herein can find use, for example, as an expression vector that, when 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 codon optimization. With respect to codon-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

further embodiments of the disclosure, polynucleotide sequence variants can be produced for a variety of reasons, *e.g.*, to optimize codon expression for a particular host (*e.g.*, changing codons in the arterivirus mRNA to those preferred by other organisms such as human, hamster, mice, or monkey).

[0118] In some embodiments disclosed herein, the GOI can encode amino acid sequence of a polypeptide. The polypeptide can generally any polypeptide, and can be, for example a therapeutic polypeptide, a prophylactic polypeptide, a diagnostic polypeptide, a nutraceutical polypeptide, an industrial enzyme, or 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.

[0119] Non-limiting examples of polypeptides that the GOI can encode include blood factors, such as β -globin, hemoglobin, tissue plasminogen activator, and coagulation factors; colony stimulating factors (CSF); interleukins, such as IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, etc.; growth factors, such as keratinocyte growth factor (KGF), stem cell factor (SCF), fibroblast growth factor (FGF, such as basic FGF and acidic FGF), hepatocyte growth factor (HGF), insulin-like growth factors (IGFs), bone morphogenetic protein (BMP), epidermal growth factor (EGF), growth differentiation factor-9 (GDF-9), hepatoma derived growth factor (HDGF), myostatin (GDF-8), nerve growth factor (NGF), neurotrophins, platelet-derived growth factor (PDGF), thrombopoietin (TPO), transforming growth factor alpha (TGF- α), transforming growth factor beta (TGF- β), and the like; soluble receptors, such as soluble TNF- α receptors, soluble VEGF receptors, soluble interleukin receptors (*e.g.*, soluble IL-1 receptors and soluble type II IL-1 receptors), soluble γ or δ T cell receptors, ligand-binding fragments of a soluble receptor, and the like; enzymes, such as α -glucosidase, α -glucuronidase, β -glucocerebrosidase, and α -glucuronidase; enzyme activators, such as tissue plasminogen activator; chemokines, such as IP-10, monokine induced by interferon-gamma (Mig), Gro α /IL-8, RANTES, MIP-1 α , MIP-1 β , MCP-1, PF-4, and the like; angiogenic agents, such as vascular endothelial growth factors (VEGFs, *e.g.*, VEGF121, VEGF165, VEGF-C, VEGF-2), transforming growth factor-beta, basic fibroblast growth factor, glioma-derived growth factor, angiogenin, angiogenin-2; and the like; anti-angiogenic agents, such as a soluble VEGF receptor; protein vaccine; neuroactive peptides, such as nerve

growth factor (NGF), bradykinin, cholecystokinin, gastrin, secretin, oxytocin, gonadotropin-releasing hormone, beta-endorphin, enkephalin, substance P, somatostatin, prolactin, galanin, growth hormone-releasing hormone, bombesin, dynorphin, warfarin, neurotensin, motilin, thyrotropin, neuropeptide Y, luteinizing hormone, calcitonin, insulin, glucagons, vasopressin, angiotensin II, thyrotropin-releasing hormone, vasoactive intestinal peptide, a sleep peptide, and the like; thrombolytic agents; atrial natriuretic peptide; relaxin; glial fibrillary acidic protein; follicle stimulating hormone (FSH); human alpha-1 antitrypsin; leukemia inhibitory factor (LIF); transforming growth factors (TGFs); tissue factors, luteinizing hormone; macrophage activating factors; tumor necrosis factor (TNF); neutrophil chemotactic factor (NCF); nerve growth factor; tissue inhibitors of metalloproteinases; vasoactive intestinal peptide; angiogenin; angiotropin; fibrin; hirudin; IL-1 receptor antagonists; and the like. Some other non-limiting examples of protein of interest include ciliary neurotrophic factor (CNTF); brain-derived neurotrophic factor (BDNF); neurotrophins 3 and 4/5 (NT-3 and 4/5); glial cell derived neurotrophic factor (GDNF); aromatic amino acid decarboxylase (AADC); hemophilia related clotting proteins, such as Factor VIII, Factor IX, Factor X; dystrophin or mini-dystrophin; lysosomal acid lipase; phenylalanine hydroxylase (PAH); glycogen storage disease-related enzymes, such as glucose-6-phosphatase, acid maltase, glycogen debranching enzyme, muscle glycogen phosphorylase, liver glycogen phosphorylase, muscle phosphofructokinase, phosphorylase kinase (e.g., PHKA2), glucose transporter (e.g., GLUT2), aldolase A, beta-enolase, and glycogen synthase; lysosomal enzymes (e.g., beta-N-acetylhexosaminidase A); and any variants thereof.

[0120] The peptide encoded by the GOI can be a multi-subunit protein or single-subunit protein. The peptide can be, for example, luciferases; fluorescent proteins (e.g., GFP); growth hormones (GHs) and variants thereof; insulin-like growth factors (IGFs) and variants thereof; granulocyte colony-stimulating factors (G-CSFs) and variants thereof; erythropoietin (EPO) and variants thereof; insulin, such as proinsulin, preproinsulin, insulin, insulin analogs, and the like; antibodies and variants thereof, such as hybrid antibodies, chimeric antibodies, humanized antibodies, monoclonal antibodies; antigen binding fragments of an antibody (Fab fragments), single-chain variable fragments of an antibody (scFV fragments); dystrophin and variants thereof; clotting factors and variants thereof; cystic

fibrosis transmembrane conductance regulator (CFTR) and variants thereof; and interferons and variants thereof.

[0121] In some embodiments, the secondary structure of the RNA transcript including the coding sequence of the GOI is optimized for a desired property. In some particular embodiments, the secondary structure of the RNA transcript including the coding sequence of the GOI is optimized for improved RNA replication.

[0122] The modified genome or replicon RNA disclosed herein is preferably a genome or replicon RNA of an arterivirus, such as a genome or replicon RNA of a viral species of the family *Arteriviridae*, genus *Arterivirus*.

[0123] Suitable arterivirus species includes Equine arteritis virus (EAV), Porcine respiratory and reproductive syndrome virus (PRRSV), Lactate dehydrogenase elevating virus (LDV), Simian hemorrhagic fever virus (SHFV), and wobbly possum disease virus (WPDV). In some embodiments, the modified genome or replicon RNA disclosed herein is of an Equine arteritis virus (EAV). In some embodiments, the modified genome or replicon RNA disclosed herein is of an EAV-virulent Bucyrus strain (VBS). In some embodiments, the modified genome or replicon RNA disclosed herein is of a Simian hemorrhagic fever virus (SHFV). Virulent and avirulent arterivirus strains are both suitable. Non-limiting examples of preferred arterivirus strains include, but not limited to, EAV-virulent Bucyrus strain (VBS), LDV-Plagemann, LDV-C, PRRSV-type 1, and PRRSV-type 2. Exemplary preferred EAV strains include, but not limited to, EAV VB53, EAV ATCC VR-796, EAV HK25, EAV HK116, EAV ARVAC MLV, EAV Bucyrus strain (Ohio), modified EAV Bucyrus, avirulent strain CA95, Red Mile (Kentucky), 84KY-A1 (Kentucky), Wroclaw-2 (Poland), Bibuna (Switzerland), and Vienna (Australia). Non-limiting preferred examples of PRRSV strains include PRRSV LV4.2.1, PRRSV 16244B, PRRSV HB-1(sh)/2002, PRRSV HB-2(sh)/2002, PRRSV HN1, PRRSV SD 01-08, PRRSV SD0802, PRRSV SD0803, PRRSV VR2332. Non-limiting preferred examples of SHFV strains and variants include SHFV variants SHFV-krtg1a and -krtg1b (SHFV-krtg1a/b), SHFV-krtg2a/b (GenBank accession # JX473847 to JX473850), SHFV-LVR, the SHFV prototype variant LVR 42-0/M6941 (NC_003092); SHFV-krc1 and SHFV-krc2 from Kibale red colobus (HQ845737 and HQ845738, respectively). Other non-limiting examples of preferred arteriviruses include

PRRSV-Lelystad, the European (type 1) type strain (M96262); PRRSVVR2332, the North American (type 2) type strain (U87392); EAV-Bucyrus (NC_002532); EAV-s3685 (GQ903794); LDV-P, the Plagemann strain (U15146); and LDV-C, the neurovirulent type C strain (L13298).

Recombinant Cells

[0124] 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. In some embodiments, the nucleic acid molecule is introduced into the eukaryotic cell by an electroporation procedure or a biolistic procedure.

[0125] In a related aspect, some embodiments disclosed herein 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 TALEN genome editing. 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.

[0126] 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 arterivirus genome or replicon RNA, as well as, optionally, a selectable marker or reporter gene, can be employed to transform an appropriate host cell.

[0127] In principle, 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.

[0128] The methods and compositions disclosed herein are preferably used with 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 host cells can be cells in *ex vivo* tissues, organs, or cell cultures (*e.g.*, *ex vivo*). In some embodiments, the host cells can be cells within a living subject or organism (*e.g.*, *in vivo*). In some embodiments, the compositions and methods of the present application can be suitably used with host cells from species that are natural hosts of arteriviruses, such as horse, pig, mice, monkey, and apes. 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. Non-limiting examples of suitable animal host cells include, but not limited to, pulmonary equine artery endothelial cell, equine dermis cell, baby hamster kidney 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, and mouse 3T3 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.

[0129] 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

[0130] In some embodiments, the compositions and methods of the present disclosure can be used to achieve tissue specific expression which enables a therapeutic arterivirus vector to be delivered systemically into a patient. If the vector should infect a cell which does not express the appropriate RNA species, the vector will only be capable of expressing nonstructural proteins and not the gene of interest. Eventually, the arterivirus vector will be harmlessly degraded.

[0131] In one non-limiting example, use of the compositions and methods described herein enables virtual tissue-specific expression possible for a variety of therapeutic applications, including for example, targeting vectors for the treatment for various types of cancers. This rationale relies on specific expression of tumor-specific markers such as the carcinoembryonic tumor specific antigen (CEA) and the alpha-fetoprotein tumor marker. Briefly, utilizing such tumor-specific RNA to target specific tumors allows for the tumor-specific expression of toxic molecules, cytokines or pro-drugs discussed below. Such methods may be utilized for a wide variety of tumors, including for example, colorectal, lung, breast, ovary, bladder and prostate cancers because all these tumors express the CEA.

[0132] Briefly, CEA was one of the first tumor-specific markers to be described, along with the alpha-fetoprotein tumor marker. CEA is a normal glycoprotein in the embryonic tissue of the gut, pancreas and liver during the first two trimesters of fetal development (Pathologic Basis of Disease, 3rd edition 1984, Robbins *et al.* (eds.)). Previously, CEA was believed to be specific for adenocarcinomas of the colon, however, with the subsequent development of more sensitive radioimmunoassays it became apparent that CEA was presented in the plasma with many endodermally derived cancers, particularly pancreatic, gastric and bronchogenic.

[0133] In some embodiments, the arterivirus genome or replicon RNAs disclosed herein may be constructed to express viral antigens, ribozyme, antisense sequences or immunostimulatory factors such as gamma-interferon (γ -IFN), IL-2 or IL-5 for the targeted treatment of virus infected cell types. In particular, in order to target arterivirus vectors to specific foreign organism or pathogen-infected cells, inverted repeats of the arterivirus vector may be selected to hybridize to any pathogen-specific RNA, for instance target cells infected by pathogens such as HIV, CMV, HBV, HPV and HSV.

[0134] In some embodiments, specific organ tissues may be targeted for the treatment of tissue-specific metabolic diseases utilizing gene replacement therapies. For example, the liver is an important target tissue because it is responsible for many of the body's metabolic functions and is associated with many metabolic genetic disorders. Such diseases include many of the glycogen storage diseases, phenylketonuria, Gaucher's disease and familial hypercholesterolemia. Presently there are many liver-specific enzymes and markers which have been sequenced which may be used to engineer appropriate inverted repeats for arterivirus vectors. Such liver-specific cDNAs include sequences encoding for S-adenosylmethionine synthetase (Horikawa *et al.*, *Biochem. Int.* 25:81, 1991); lecithin: cholesterolacyl transferase (Rogne *et al.*, *Biochem. Biophys. Res. Commun.* 148:161, 1987); as well as other liver-specific cDNAs (Chin *et al.*, *Ann. N.Y. Acad. Sci.* 478:120, 1986). Such a liver-specific arterivirus vector could be used to deliver the low density lipoprotein receptor (Yamamoto *et al.*, *Cell* 39:27, 1984) to liver cells for the treatment of familial hypercholesterolemia (Wilson *et al.*, *Mol. Biol. Med.* 7:223, 1990)..

[0135] The arterivirus genome or replicon RNAs disclosed herein can be used to express one or more heterologous coding sequence(s) or functional RNA(s) of interest, also referred to herein as a heterologous RNA or heterologous sequence, which can be chosen from a wide variety of sequences derived from viruses, prokaryotes and eukaryotes. Examples of categories of heterologous sequences include, but are not limited to, sequences coding for immunogens (including native, modified or synthetic antigenic proteins, peptides, epitopes or immunogenic fragments), cytokines, toxins, therapeutic proteins, enzymes, antisense sequences, and immune response modulators.

[0136] Heterologous Nucleotide Sequences.

[0137] In accordance of some embodiments of the present disclosure, a wide variety of nucleotide sequences can be carried by the modified arterivirus genome or replicon RNA of the present disclosure. In some embodiments, the modified arterivirus genome or replicon RNA as described herein does not contain any additional heterologous nucleotide sequence. In some embodiments, the modified arterivirus genome or replicon RNA of the present disclosure contains one or more additional heterologous or foreign nucleotide sequences. In some embodiments, the heterologous nucleotide sequence comprises a heterologous nucleotide sequence of at least about 100 bases, 2 kb, 3.5 kb, 5 kb, 7 kb, or even a heterologous sequence of at least about 8 kb.

[0138] A wide variety of heterologous nucleotide sequences may be included in the modified arterivirus genome or replicon RNA of the present disclosure, including for example sequences which encode palliatives such as cytokines, toxins, prodrugs, antigens which stimulate an immune response, ribozymes, and proteins which assist or inhibit an immune response, as well as antisense sequences (or sense sequences for "antisense applications"). As noted above, within various embodiments of the disclosure the modified arterivirus genome or replicon RNAs provided herein may contain (and express, within certain embodiments) two or more heterologous nucleotide sequence.

[0139] 1). Cytokines

[0140] In some embodiments disclosed herein, the heterologous nucleotide sequence encodes a cytokine. Briefly, cytokines act to proliferate, activate, or differentiate immune effectors cells. Representative examples of cytokines include, macrophages, B lymphocytes, T lymphocytes, endothelial cells, fibroblasts, lymphokines likes gamma interferon, tumor necrosis factor, interleukin, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL- 1, IL-12, IL-13, IL-14, IL-15, GM-CSF, CSF-1 and G-CSF.

[0141] In some related embodiments, the heterologous nucleotide sequence encodes an immunomodulatory cofactor. Briefly, as utilized within the context of the present disclosure, "immunomodulatory cofactor" refers to factors which, when manufactured by one or more of the cells involved in an immune response, or when added exogenously to the cells, cause the immune response to be different in quality or potency from that which would have occurred in the absence of the cofactor. The quality or potency of a response may be

measured by a variety of assays known to one of skill in the art including, for example, in vitro assays which measure cellular proliferation (e.g., ^3H thymidine uptake), and in vitro cytotoxic assays (e.g., which measure ^{51}Cr release) (see *Warner et al.*, *AIDS Res. and Human Retroviruses* 7:645-655, 1991).

[0142] Representative examples of immunomodulatory co-factors include alpha interferon (Finter *et al.*, *Drugs* 42(5):749-765, 1991; U.S. Pat. No. 4,892,743; U.S. Pat. No. 4,966,843; WO 85/02862; Nagata *et al.*, *Nature* 284:316-320, 1980; Familletti *et al.*, *Methods in Enz.* 78:387-394, 1981; Twu *et al.*, *Proc. Natl. Acad. Sci. USA* 86:2046-2050, 1989; Faktor *et al.*, *Oncogene* 5:867-872, 1990), beta interferon (Seif *et al.*, *J. Virol.* 65:664-671, 1991), gamma interferons (Radford *et al.*, *American Society of Hepatology*:2008-2015, 1991; Watanabe *et al.*, *Proc. Natl. Acad. Sci. USA* 86:9456-9460, 1989; Gansbacher *et al.*, *Cancer Research* 50:7820-7825, 1990; Maio *et al.*, *Can. Immunol. Immunother.* 30:34-42, 1989; U.S. Pat. Nos. 4,762,791 and 4,727,138), G-CSF (U.S. Pat. Nos. 4,999,291 and 4,810,643), GM-CSF (WO 85/04188), TNFs (Jayaraman *et al.*, *J. Immunology* 144:942-951, 1990), Interleukin-2 (IL-2) (Karupiah *et al.*, *J. Immunology* 144:290-298, 1990; Weber *et al.*, *J. Exp. Med.* 166:1716-1733, 1987; Gansbacher *et al.*, *J. Exp. Med.* 172:1217-1224, 1990; U.S. Pat. No. 4,738,927), IL-4 (Tepper *et al.*, *Cell* 57:503-512, 1989; Golumbek *et al.*, *Science* 254:713-716, 1991; U.S. Pat. No. 5,017,691), IL-6 (Brakenhof *et al.*, *J. Immunol.* 139:4116-4121, 1987; WO 90/06370), IL-12, IL-15 (Grabstein *et al.*, *Science* 264:965-968, 1994; Genbank-EMBL Accession No. V03099), ICAM-1 (Altman *et al.*, *Nature* 338:512-514, 1989), ICAM-2, LFA-1, LFA-3, MHC class I molecules, MHC class II molecules, 2 - microglobulin, chaperones, CD3, B7/BB 1, MHC linked transporter proteins or analogues thereof.

[0143] The choice of which immunomodulatory cofactor to include within the modified arterivirus genome or replicon RNA of the present disclosure may be based upon known therapeutic effects of the cofactor, or experimentally determined. For example, in chronic hepatitis B infections alpha interferon has been found to be efficacious in compensating a patient's immunological deficit and thereby assisting recovery from the disease. Alternatively, a suitable immunomodulatory cofactor may be experimentally determined. Briefly, blood samples are first taken from patients with a hepatic disease.

Peripheral blood lymphocytes (PBLs) are restimulated in vitro with autologous or HLA-matched cells (*e.g.*, EBV transformed cells), and transduced with modified arterivirus genome or replicon RNA of the present disclosure which directs the expression of an immunogenic portion of a hepatitis antigen and the immunomodulatory cofactor. Stimulated PBLs are used as effectors in a CTL assay with the HLA-matched transduced cells as targets. An increase in CTL response over that seen in the same assay performed using HLA-matched stimulator and target cells transduced with a vector encoding the antigen alone, indicates a useful immunomodulatory cofactor. In some embodiments, the immunomodulatory cofactor gamma interferon is particularly preferred.

[0144] Another non-limiting example of an immunomodulatory cofactor is the B7/BB1 costimulatory factor. Briefly, activation of the full functional activity of T cells requires two signals. One signal is provided by interaction of the antigen-specific T cell receptor with peptides which are bound to major histocompatibility complex (MHC) molecules, and the second signal, referred to as costimulation, is delivered to the T cell by antigen-presenting cells. Briefly, the second signal is required for interleukin-2 (IL-2) production by T cells and appears to involve interaction of the B7/BB1 molecule on antigen-presenting cells with CD28 and CTLA-4 receptors on T lymphocytes (Linsley *et al.*, *J. Exp. Med.*, 173:721-730, 1991a, and *J. Exp. Med.*, 174:561-570, 1991). In some embodiments, B7/BB1 may be introduced into tumor cells in order to cause costimulation of CD8⁺T cells, such that the CD8⁺T cells produce enough IL-2 to expand and become fully activated. These CD8⁺T cells can kill tumor cells that are not expressing B7 because costimulation is no longer required for further CTL function. Vectors that express both the costimulatory B7/BB1 factor and, for example, an immunogenic HBV core protein, may be made utilizing methods which are described herein. Cells transduced with these vectors will become more effective antigen-presenting cells. The HBV core-specific CTL response will be augmented from the fully activated CD8⁺T cell via the costimulatory ligand B7/BB1.

[0145] 2). Toxins

[0146] In some embodiments disclosed herein, the heterologous nucleotide sequence encodes a toxin. Briefly, toxins act to directly inhibit the growth of a cell. Representative examples of toxins include ricin (Lamb *et al.*, *Eur. J. Biochem.* 148:265-270,

1985), abrin (Wood *et al.*, *Eur. J. Biochem.* 198:723-732, 1991; Evensen *et al.*, *J. of Biol. Chem.* 266:6848-6852, 1991; Collins *et al.*, *J. of Biol. Chem.* 265:8665-8669, 1990; Chen *et al.*, *Fed. of Eur. Biochem Soc.* 309:115-118, 1992), diphtheria toxin (Tweten *et al.*, *J. Biol. Chem.* 260:10392-10394, 1985), cholera toxin (Mekalanos *et al.*, *Nature* 306:551-557, 1983; Sanchez and Holmgren, *PNAS* 86:481-485, 1989), gelonin (Stirpe *et al.*, *J. Biol. Chem.* 255:6947-6953, 1980), pokeweed (Irvin, *Pharmac. Ther.* 21:371-387, 1983), antiviral protein (Barbieri *et al.*, *Biochem. J.* 203:55-59, 1982; Irvin *et al.*, *Arch. Biochem. & Biophys.* 200:418-425, 1980; Irvin, *Arch. Biochem. & Biophys.* 169:522-528, 1975), tritin, Shigella toxin (Calderwood *et al.*, *PNAS* 84:4364-4368, 1987; Jackson *et al.*, *Microb. Path.* 2:147-153, 1987), Pseudomonas exotoxin A (Carroll and Collier, *J. Biol. Chem.* 262:8707-8711, 1987), herpes simplex virus thymidine kinase (HSVTK) (Field *et al.*, *J. Gen. Virol.* 49:115-124, 1980), and *E. coli* guanine phosphoribosyl transferase.

[0147] 3). Pro-drugs

[0148] In some embodiments disclosed herein, the heterologous nucleotide sequence encodes a "pro-drug". Briefly, as utilized within the context of the present disclosure, "pro-drug" refers to a gene product that activates a compound with little or no cytotoxicity into a toxic product. Representative examples of such gene products include HSVTK and VZVTK (as well as analogues and derivatives thereof), which selectively monophosphorylate certain purine arabinosides and substituted pyrimidine compounds, converting them to cytotoxic or cytostatic metabolites. More specifically, exposure of the drugs ganciclovir, acyclovir, or any of their analogues (*e.g.*, FIAU, FIAC, DHPG) to HSVTK phosphorylates the drug into its corresponding active nucleotide triphosphate form.

[0149] Representative examples of other pro-drugs which may be utilized within the context of the present disclosure include: *E. coli* guanine phosphoribosyl transferase which converts thioxanthine into toxic thioxanthine monophosphate (Besnard *et al.*, *Mol. Cell. Biol.* 7:4139-4141, 1987); alkaline phosphatase, which will convert inactive phosphorylated compounds such as mitomycin phosphate and doxorubicin-phosphate to toxic dephosphorylated compounds; fungal (*e.g.*, *Fusarium oxysporum*) or bacterial cytosine deaminase, which will convert 5-fluorocytosine to the toxic compound 5-fluorouracil (Mullen, *PNAS* 89:33, 1992); carboxypeptidase G2, which will cleave the glutamic acid from

para-N-bis (2-chloroethyl) aminobenzoyl glutamic acid, thereby creating a toxic benzoic acid mustard; and Penicillin-V amidase, which will convert phenoxyacetabide derivatives of doxorubicin and melphalan to toxic compounds (see generally, Vrudhula *et al.*, *J. of Med. Chem.* 36(7):919-923, 1993; Kern *et al.*, *Canc. Immun. Immunother.* 31(4):202-206, 1990).

[0150] 4). Antisense sequence

[0151] In some embodiments disclosed herein, the heterologous nucleotide sequence is an antisense sequence. Briefly, antisense sequences are designed to bind to RNA transcripts, and thereby prevent cellular synthesis of a particular protein or prevent use of that RNA sequence by the cell. Representative examples of such sequences include antisense thymidine kinase, antisense dihydrofolate reductase (Maher and Dolnick, *Arch. Biochem. & Biophys.* 253:214-220, 1987; Bzik *et al.*, *PNAS* 84:8360-8364, 1987), antisense HER2 (Coussens *et al.*, *Science* 230:1132-1139, 1985), antisense ABL (Fainstein *et al.*, *Oncogene* 4:1477-1481, 1989), antisense Myc (Stanton *et al.*, *Nature* 310:423-425, 1984) and antisense ras, as well as antisense sequences which block any of the enzymes in the nucleotide biosynthetic pathway. In addition, in accordance with some embodiments disclosed herein, antisense sequences to interferon and 2 microglobulin may be utilized in order to decrease immune response.

[0152] Alternatively or in addition, in some embodiments, antisense RNA may be utilized as an anti-tumor agent in order to induce a potent Class I restricted response. Briefly, in addition to binding RNA and thereby preventing translation of a specific mRNA, high levels of specific antisense sequences are believed to induce the increased expression of interferons (including gamma-interferon) due to the formation of large quantities of double-stranded RNA. The increased expression of gamma interferon, in turn, boosts the expression of MHC Class I antigens. Preferred antisense sequences for use in this regard include actin RNA, myosin RNA, and histone RNA. Antisense RNA which forms a mismatch with actin RNA is particularly preferred.

[0153] 5). Ribozymes

[0154] In some embodiments disclosed herein, modified arterivirus genome or replicon RNAs are provided which produce ribozymes upon infection of a host cell. Briefly, ribozymes are used to cleave specific RNAs and are designed such that it can only affect one

specific RNA sequence. Generally, the substrate binding sequence of a ribozyme is between 10 and 20 nucleotides long. The length of this sequence is sufficient to allow a hybridization with target RNA and disassociation of the ribozyme from the cleaved RNA. Representative examples for creating ribozymes include those described in U.S. Pat. Nos. 5,116,742; 5,225,337 and 5,246,921.

[0155] 6). Proteins and other cellular constituents

[0156] In some embodiments disclosed herein, a wide variety of proteins or other cellular constituents can be carried by the modified arterivirus genome or replicon RNAs of the disclosure. Representative examples of such proteins include native or altered cellular components, as well as foreign proteins or cellular constituents, found in for example, viruses, bacteria, parasites, fungus or animal such as mammalian.

Methods for Producing Polypeptides

[0157] In another aspect, the present application provides methods for producing one or more polypeptide of interests. The polypeptides of interest according the present disclosure can be generally any polypeptide and can be, for example recombinant proteins and peptides suitable for pharmaceutical, nutraceutical, and/or industrial compositions. Non-limiting examples of suitable recombinant polypeptides include therapeutic polypeptides, prophylactic polypeptides, diagnostic polypeptides, nutraceutical polypeptides, industrial enzymes, and reporter polypeptides. In some embodiments, polypeptides of interest made in accordance with the present disclosure have a variety of uses including, but not limited to, use as vaccines and other therapeutic compounds, use as diagnostic agents and use as antigens in the production of polyclonal or monoclonal antibodies.

[0158] In some embodiments, the host cells can be recombinant cells in culture (*e.g.*, *ex vivo*). In some embodiments, the host cells can be recombinant cells in a living subject (*e.g.*, *in vivo*). Accordingly, in some embodiments, the method for producing a polypeptide of interest according to this aspect of the disclosure can include the cultivation of a recombinant host cell, such as a recombinant mammalian cell, including a nucleic acid molecule according to any one of the preceding aspects and embodiments. To produce one or more polypeptides of interest according the present disclosure, a recombinant cell, produced as described above, is cultured in an effective medium, using any one of cell culturing

techniques known in the art. As used herein, an effective medium refers to any medium in which the transfected cells can produce one or more polypeptides of interest according to the present disclosure. An effective medium is typically an aqueous medium comprising assimilable carbohydrate, nitrogen and phosphate sources, as well as appropriate salts, minerals, metals and other nutrients, such as vitamins, growth factors and other hormones. The medium may comprise complex nutrients or may be a defined medium. Recombinant cells of the present disclosure can be cultured in conventional fermentation bioreactors, which include, but are not limited to, batch, fed-batch, cell recycle and continuous fermenters. Culturing can also be conducted in shake flasks, test tubes, microtiter dishes and petri plates. Culturing is carried out at a temperature, pH and oxygen content appropriate for the recombinant cell. Such culturing conditions are well within the expertise of one of ordinary skill in the art. Non-limiting examples of preferred effective media and culturing conditions are included in the Examples section.

[0159] Depending on whether expression results in a polypeptide of interest having or lacking a signal segment, the resultant polypeptide may be secreted into the medium or remain within the recombinant cell. The phrase “recovering the protein” refers simply to collecting the whole fermentation medium (including cells) containing the polypeptide and can, but need not, entail additional steps of separation or purification. Polypeptides of interest of the present disclosure can be purified using a variety of standard protein purification techniques, such as, but not limited to, affinity chromatography, ion exchange chromatography, filtration, electrophoresis, hydrophobic interaction chromatography, gel filtration chromatography, reverse phase chromatography, chromatofocusing and differential solubilization.

[0160] Isolated polypeptides of interest of the present disclosure are preferably retrieved in “substantially pure” form. As used herein, “substantially pure” refers to a purity that allows for the effective use of the compound as a therapeutic composition or diagnostic. A vaccine for animals, for example, should exhibit no preferably substantial toxicity and should be capable of stimulating the production of antibodies in a vaccinated animal.

[0161] In some embodiments, the method for producing a polypeptide of interest of the present disclosure includes culturing a host cell comprising a nucleic acid as described

herein. In some embodiments, the nucleic acid includes (i) nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified genome or replicon RNA comprises a sequence fragment exhibiting at least 80% sequence identity to the sequence encoding open reading frame ORF7, and wherein the modified genome or replicon RNA is devoid of the sequence encoding ORF2a; and (ii) one or more expression cassettes, wherein each of the one or more expression cassettes comprises a subgenomic (sg) promoter operably linked to a heterologous nucleotide sequence encoding a gene of interest (GOI).

[0162] In some embodiments, the method for producing a polypeptide of interest of the present disclosure includes administering to a subject a nucleic acid as described herein. In some embodiments, the nucleic acid includes (i) nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified genome or replicon RNA comprises a sequence fragment exhibiting at least 80% sequence identity to the sequence encoding open reading frame ORF7, and wherein the modified genome or replicon RNA is devoid of the sequence encoding ORF2a; and (ii) one or more expression cassettes, wherein each of the one or more expression cassettes comprises a subgenomic (sg) promoter operably linked to a heterologous nucleotide sequence encoding a gene of interest (GOI).

[0163] 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, polypeptides produced by a method according to this aspect of the application are also within the scope of this application.

[0164] In some embodiments, the methods according to this aspect are preferably deployed in 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 methods of this aspect can be suitably deployed in animal cells. Therapeutic protein production in small and large scale is 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 some embodiments, the host cells can be of animal species that are natural hosts of arteriviruses,

such as horse, pig, mice, monkey, and apes. 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. Non-limiting examples of suitable animal host cells include, but not limited to, pulmonary equine artery endothelial cell, equine dermis cell, baby hamster kidney 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, and mouse 3T3 cell. In some embodiments, the host cell is baby hamster kidney (BHK) cell, as described in more details in Examples 10-13 and 17-18. In some embodiments, the baby hamster kidney cell is a BHK-21 cell.

Recombinant Polypeptides

[0165] In a further aspect, some embodiments disclosed herein relate to recombinant polypeptides produced by a method in accordance with one or more embodiments described in the present application.

[0166] The recombinant polypeptide according the present disclosure can be generally any polypeptide and can be, for example recombinant proteins and peptides suitable for pharmaceutical, nutraceutical, and/or industrial compositions. Non-limiting examples of suitable recombinant polypeptides include therapeutic polypeptides, prophylactic polypeptides, diagnostic polypeptides, nutraceutical polypeptides, industrial enzymes, and reporter polypeptides.

[0167] The term "therapeutic polypeptide," as used herein denotes a bioactive polypeptide that has therapeutic utility. The term encompasses any polypeptide that can be administered to a patient to produce a beneficial therapeutic or diagnostic effect through binding to and/or altering the function of a biological target molecule in the patient. The target molecule can be an endogenous target molecule encoded by the patient's genome (*e.g.*, an enzyme, receptor, growth factor, cytokine encoded by the patient's genome) or an

exogenous target molecule encoded by the genome of a pathogen (*e.g.*, an enzyme encoded by the genome of a virus, bacterium, fungus, nematode or other pathogen). Illustrative categories of therapeutic peptides suitable for practicing the compositions and methods of the present disclosure are hormones, monoclonal antibodies, vaccines, enzymes, cytokines, toxins, and the like. The term therapeutic polypeptide includes functional fragments of therapeutic polypeptides.

[0168] The term "nutraceutical polypeptide" as used herein refers to any polypeptide which may prevent, ameliorate or otherwise confer benefits against an undesirable condition, and used for its associated health benefits, to maintain the healthy condition of the consumer. The term "nutraceutical" as used herein denotes a usefulness in both the nutritional and pharmaceutical field of application. Thus, the nutraceutical polypeptides and compositions of the present disclosure can find use as supplement to food and beverages, and as pharmaceutical formulations not associated with food, suitable for consumption by an individual and usually sold in medicinal forms which may be solid formulations such as caplets, tablet, capsules, soft gel capsules, gel caps and the like, or liquid formulations, such as solutions or suspensions. As such, the term nutraceutical composition comprises food and beverages containing the nutraceutical polypeptides disclosed herein, for example protein hydrolysates which are rich in tripeptides.

[0169] A "reporter polypeptide", as used herein, is a polypeptide that is detectable or has an activity that produces a detectable product. A reporter polypeptide can include a visual marker or enzyme that produces a detectable signal. Non-limiting examples of reporter polypeptides includes cat, lacZ, uidA, xylE, an alkaline phosphatase gene, an α -amylase gene, an α -galactosidase gene, a β -glucuronidase gene, a β -lactamase gene, a horseradish peroxidase gene, a luciferin/luciferase gene, an R-locus gene, a tyrosinase gene, or a gene encoding a fluorescent protein, including but not limited to a blue, cyan, green, red, or yellow fluorescent protein, a photoconvertible, photoswitchable, or optical highlighter fluorescent protein, or any of variant thereof, including, without limitation, codon-optimized, rapidly folding, monomeric, increased stability, and enhanced fluorescence variants.

Pharmaceutical Compositions

[0170] In a further aspect, some embodiments disclosed herein relate to a composition including a recombinant polypeptide as described herein and a pharmaceutically acceptable carrier.

[0171] In yet further aspect, some embodiments disclosed herein relate to a composition including a nucleic acid molecule as disclosed herein and a pharmaceutically acceptable carrier.

[0172] In yet a further aspect, some embodiments disclosed herein relate to a composition including a recombinant cell as disclosed herein and a pharmaceutically acceptable carrier.

[0173] In some embodiments disclosed herein, the compositions of the present application can be further formulated for use as a protective composition (*e.g.*, vaccine) or therapeutic composition. In particular, protective compositions made in accordance with the present disclosure have a variety of uses including, but not limited to, use as vaccines and other therapeutic agents, use as diagnostic agents and use as antigens in the production of polyclonal or monoclonal antibodies. Thus, in the case of vaccines, the compositions of the present application can provide a method for inducing an immune response in a nucleic acid of the composition in an immunogenic amount to a subject, particles, which method comprises administering a population and / or composition, the target.

[0174] When used as vaccines, the compositions in general must be stored at low temperature, or they have to be in a freeze-dried form. Freeze-dried vaccines can be kept under moderate cooling conditions or even at room temperature. Often, the vaccine is mixed with stabilizers, *e.g.* to protect degradation-prone proteins from being degraded, to enhance the shelf-life of the vaccine, or to improve freeze-drying efficiency. Useful stabilizers include, but are not limited to, SPGA, carbohydrates such as, for example, sorbitol, mannitol, trehalose, starch, sucrose, dextran or glucose, proteins such as albumin or casein or degradation products thereof, and buffers, such as alkali metal phosphate. Accordingly, in some embodiments, vaccine according to the present disclosure is in a freeze-dried form. Alternatively or in addition, the vaccine may be suspended in a physiologically acceptable diluent and/or buffer.

[0175] In some embodiments disclosed herein, the compositions of the present application can be further formulated into a therapeutic composition capable of protecting an animal from disease caused by a parasite when the composition is administered to the animal in an effective amount. In some embodiments, the therapeutic composition is a multivalent therapeutic composition which contains multiple protective polypeptides targeting multiple targets and/or multiple parasites. Such multivalent therapeutic compositions can be produced by combining one or more protective polypeptides after production, by culturing more than one recombinant cell in a culturing reaction or by producing more than one protective polypeptides in a recombinant cell by, for example, transfecting an animal cell with one or more recombinant molecules and/or by transfecting an animal cell with a recombinant molecule containing more than one nucleic acid sequence encoding one or more protective polypeptides as disclosed herein.

[0176] In some embodiments, the therapeutic composition as described herein can also include an immunopotentiator, such as an adjuvant or a carrier. Suitable adjuvants or carriers include the adjuvants and carriers suitable for administration of recombinant polypeptides of the present disclosure. Therapeutic compositions of the present disclosure can be formulated in an excipient that the animal to be administered can tolerate. Examples of such excipients include water, saline, Ringer's solution, dextrose solution, Hank's solution, and other aqueous physiologically balanced salt solutions. Nonaqueous vehicles, such as fixed oils, sesame oil, ethyl oleate, or triglycerides may also be used. Other useful formulations include suspensions containing viscosity enhancing agents, such as sodium carboxymethylcellulose, sorbitol, or dextran. Excipients can also contain minor amounts of additives, such as substances that enhance isotonicity and chemical stability. Examples of buffers include phosphate buffer, bicarbonate buffer and Tris buffer, while examples of preservatives include thimerosal, m or o-cresol, formalin and benzyl alcohol. Standard formulations will either be liquid injectables or solids which can be taken up in a suitable liquid as a suspension or solution for injection. Thus, in a non-liquid formulation, the excipient may comprise dextrose, human serum albumin, preservatives, *etc.*, to which sterile water or saline could be added prior to administration.

[0177] 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.

[0178] 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.

[0179] 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.

[0180] 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

[0181] 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

Construction of Base Vectors

[0182] This Example describes the generation of the base arterivirus expression vector that was then further modified and subsequently used in the construction of monovalent, bivalent, and trivalent vectors.

Construction of the base vector REP-EAV(WT)

[0183] The Rep-EAV-Ren(1G)v2-N-seq vector was assembled as follows. The Renilla luciferase gene and three EAV fragments (EAV F1-F3) were synthesized.

[0184] The sequence contig of the three synthesized fragments EAV F1 (SEQ ID NO: 4), EAV F2 (SEQ ID NO: 5, and EAV F3 (SEQ ID NO: 6) included (1) an upstream portion of the reporter Renilla luciferase gene, (2) the EAV leader sequence, and (3) the coding sequence of the non-structural polypeptide pp1ab, which corresponded to nucleotide residues 1-9751 of the EAV genome (NCBI Accession Number gi|14571796). The Renilla luciferase gene (SEQ ID NO: 7) contained a 5' *Xba*I site and a 3' *Pst*I site. To include the 40 nucleotide polyA-tail, a synthetic nucleic acid design, named EAV_ultramer, was designed (SEQ ID NO: 8), containing the polyA sequence in the middle and flanking regions to the Renilla luciferase gene and a portion of the linear vector sequence respectively. In the final replicon, the sequence of each of the synthesized fragments had a 50 bp overlap to its neighboring fragment in the following order: 5' — linear vector — EAV F1 fragment — EAV F2 fragment — EAV F3 fragment — Renilla luciferase gene — EAV_ultramer — linear vector — 3'. A schematic representation of this final replicon is shown in **FIG. 2** of the disclosure. The nucleotide sequence of the final replicon assembled as described above, minus the sequence encoding the Renilla luciferase reporter is provided as SEQ ID NO: 3 in the Sequence Listing. In addition, the sequence contig that contains the 5' leader, ORF1a, ORF1b is provided as SEQ ID NO: 1 in the Sequence Listing.

Construction of backbone plasmids containing mutated T7 termination sequence

[0185] In some experiments, mutations in the T7 terminator sequences were introduced into various vectors using the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies) in accordance with the manufacturer's instructions. Mutagenesis primers were designed using the QuikChange Primer Design Tool in accordance with the manufacturer's instructions (www.genomics.agilent.com/primerDesignProgram.jsp).

[0186] The following mutations were introduced into the Rep-EAV(WT) vector, T9001G, T9001G and G3188A, and T9001G and T3185A, resulting in new constructs rE(WT)-Ren (T9001G), rE2(WT)-Ren (containing T9001G and G3188A mutations), and rE3(WT)-Ren (containing T9001G and T3185C mutations), respectively. In addition, an *XhoI* restriction enzyme site was added immediately 3' to the sequence of the Renilla gene by QuikChange mutagenesis for future cloning needs.

TRS mutations

[0187] Towards the design of tunable regulation system of gene expression, six mutations were introduced to both the leader TRS and body TRS7 present in the Rep-EAV(WT) backbone. The wild type TRS sequence was TCAACT and the sequences of the mutated TRS1-6 were as follows: TRS1-CTAACC, TRS2-CCAACC, TRS3-CCAAGC, TRS4-CCAGGC, TRS5-CCAGGT, TRS6-GGTTAG. The resulting vectors were named Rep-EAV(TRS1), Rep-EAV(TRS2), Rep-EAV(TRS3), Rep-EAV(TRS4), Rep-EAV(TRS5), and Rep-EAV(TRS6), respectively.

Vector constructs with spacers

[0188] In some experiments described below, four different spacer sequences, named Spacer 1-4, were introduced into various vectors to insulate the TRS7 region (FIG. 11). The spacer sequences are provided as SEQ ID NOs: 9-12 in the Sequence Listing. Spacer 1 (SEQ ID NO: 9) was designed to include the entire ORF6 and mutations were also introduced to eliminate two start codons and a native *XbaI* site. Spacer 2 (SEQ ID NO: 10) included a conserved AT-rich region upstream of the TRS7 sequence. Spacer 3 (SEQ ID NO: 11) included an additional sequence downstream of the TRS7 and ends near the second start codon 3' of TRS7. Spacer 4 (SEQ ID NO: 12) represents a sequence that has been extended in the 5' direction from TRS7 to include Spacer 1 and in the 3' direction to include Spacer 3. These spacers were assembled into the rE2(WT) backbone replicon with four different

reporter genes, Cypridina (Cypr; SEQ ID NO: 13), Green Renilla (gRen; SEQ ID NO: 14), Red Firefly (rFF; SEQ ID NO: 15), and Renilla (Ren; SEQ ID NO: 7).

Bivalent expression constructs

[0189] In some experiments described below, two different base designs of bivalent expression constructs were constructed. The construction of version 'A' bivalent constructs utilized the *XbaI* restriction site between TRS2 and TRS7 to introduce another reporter gene into the rE2(WT)-reporter constructs. The construction of version 'B' bivalent constructs used the *PsiI* restriction site located immediately downstream of the reporter genes to introduce a PCR amplified DNA fragment that carries another reporter gene, as well as an upstream fragment with both TRS2 and TRS7 (also referred to as 2/7 block henceforth). Bivalent constructs made according to this Example and subsequently evaluated in some Examples below include rE2(WT)-gRen-rFF-A, rE2(WT)-gRen-rFF-B, rE2(WT)-rFF-gRen-A, rE2(WT)-rFF-gRen-B.

3' UTR modifications

[0190] In some experiments described herein, 3'UTR sequences were modified to enhance expression of the genes of interest encoded in the replicon. Whereas the initial base vector contained 366 bp from the stop codon of the gene of interest to the polyA string, two different modified 3'UTR sequences were designed to contain 801 bp 3' terminal region of the EAV genome (SEQ ID NO: 41). To clone this additional 3' UTR region, the rE2(WT) vectors were digested with *XhoI* and column purified, the insert was amplified from an infectious clone sequence and gel purified, then both DNA fragments were assembled together by using the Gibson Assembly® procedure. Monovalent versions of this replicon are referred to as rEna constructs. Another monovalent version of this backbone was made to inactivate both the TRS7 used to drive a GOI and the TRS7 in the 801-nt 3' region; these vectors are subsequently referred to as rExa constructs. A bivalent vector form of the rExa replicon was generated and this vector is referred to as rExb. Because the bivalent rExb construct did not have any functional TRS7 sequences, another version of it was made so that only the TRS7 in the 801-nt region was inactivated, these vectors are referred to as rExc constructs.

Non-reporter replicon constructs

[0191] Genes other than luciferase reporter genes that were cloned into the rE2 replicon are as follows: hemagglutinin (HA), RSV F0 precursor protein, EGFP (SEQ ID NO: 24), Cas9 (SEQ ID NO: 25 and SEQ ID NO: 26), Csy4 (SEQ ID NO: 27), neomycin resistance gene (SEQ ID NO: 28), puromycin resistance gene (SEQ ID NO: 29), anti-NP antibody (light chain-IRES-heavy chain; SEQ ID NO: 30 and SEQ ID NO: 31), Humira (anti-TNF antibody; SEQ ID NO: 33), Herceptin (anti-Her2 antibody; SEQ ID NO: 32), and GFP-ApoAI fusion gene (SEQ ID NO: 34). Some of the above genes and the following genes were cloned into the rEn replicon: Interleukin12 (IL12; SEQ ID NO: 35), EpCam (SEQ ID NO: 36), and His6 or myc-tagged peptide string (CT26; SEQ ID NO: 37). For HA and F0 proteins, 4 different sequence-optimized genes each were tested (SEQ ID Nos: 16-19 and 20-23, respectively). For Cas9, 2 sequence-optimized genes were tested (SEQ ID NO: 25 and SEQ ID NO: 26). To construct all the above genes into the rE2 base plasmid, an rE2 reporter construct was digested with *Xba*I and *Xho*I to remove the existing reporter gene, followed by gel extraction of the backbone vector. The non-reporter genes were amplified with primers that contain flanking sequences and 40-60 bp of overlapping sequence with the backbone vector, gel purified, then assembled with the backbone vector using Gibson Assembly® as described below. To clone into the rEna plasmid, the rEna-rFF plasmid was digested with *Xba*I, the backbone vector as gel extracted to remove the original insert. A neomycin resistance gene and EGFP bivalent (A) design was also cloned into rEna base vector. For this bivalent construct, the coding sequence of EGFP was sequence-optimized (SEQ ID NO: 24).

Gibson Assembly® protocol

[0192] SGI's Archetype® Software was used to design 60-bp long, overlapping oligonucleotides covering the DNA sequence of interest. The 60-bp oligonucleotides overlapped neighboring oligonucleotides by 30 bp. Oligonucleotides were ordered from Integrated Digital Technologies (IDT) at a concentration of 100 µM and then pooled to reach a target concentration of 25nM for subsequent gene assembly.

[0193] Gene assembly was performed according to the method described in Gibson *et al.* (Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345, 2009). Error correction was performed by forming heteroduplexes of

any error containing PCR products by incubation at 98°C for 2min, to 85°C at a rate of 2°C/sec, incubation for 2min, to 25°C at a rate of 0.1°C/sec, and incubation for 2min. Resulting heteroduplexes were then cleaved by adding 2.7 µL of each PCR reaction to 5.3 microliters water, 2 µL Surveyor Nuclease and 1 µL of 1:4000 diluted NEB *ExoIII*, followed by incubation at 42°C for 1 hour. A recovery PCR reaction (PCR2) was identical to the first amplification except 2.5 µL of error corrected DNA was added to 47.4 µL of mastermix. In addition, 0.12 µL of the EAV_ultramer (10 nM) was added to the PCR2 for the generation of the Renilla luciferase insert.

DNA template preparation

[0194] Plasmid DNA templates were purified (Qiagen Cat. no. 12163) from 300 mL of saturated *E. coli* TransforMax 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.

In vitro transcription

[0195] *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). 1U of DNaseI, 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.

Transfection and analysis

[0196] 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 were 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.

[0197] 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.

In vivo studies

[0198] *In vivo* bioluminescence imaging was performed using an IVIS® Spectrum optical imaging system (PerkinElmer, Inc.). In a typical experiment, animals were imaged up

to three at a time under 2% isoflurane gas anesthesia. Each mouse was injected intraperitoneally (IP) with 200mg/kg D-luciferin and RO with 5mg/kg coelenterazine and imaged in the supine position immediately following substrate injection. Coelenterazine was injected immediately prior to D-luciferin for all time points. Scans were acquired at the following wavelengths for the first 2 time points: 520, 540, 560, 620, 640, and 660nm. At all other time points, only 540nm and 640nm wavelengths were used during scanning. On Day 8, animals were scanned for 10 minutes, *in vivo* post-substrate injections then euthanized via CO2 overexposure, animals were then debrided and the thoracic cavity was opened to expose the lungs for *in situ* BLI imaging. Large binning of the CCD chip was used and the exposure time was adjusted (10 minutes to 10 seconds) to obtain at least several hundred counts from the hind limbs in each mouse in the image and to avoid saturation of the CCD chip.

EXAMPLE 2

T7 termination sequence modifications

[0199] This Example describes the results of experiments assessing impact of various point mutations introduced into the sequence of T7 RNA polymerase transcription termination signals identified in the coding sequence of EAV non-structural polypeptides.

[0200] Initial replicon designs were modified to remove cryptic T7 RNA polymerase transcription termination signals identified in the non-structural gene coding region of the EAV genome. In these experiments, the T7 terminator containing replicon vector versions were designated Rep-EAV (WT), rE, rE2 and rE3; each of these replicons contain various T7 terminator modifications by using the procedure described above. Subsequently, agarose gel electrophoresis analysis of *in vitro* transcribed RNAs derived from two exemplary replicon vectors containing modified T7 terminator sequences, rE-rFF and rE2-rFF, has been performed to demonstrate that modification of the T7 terminator sites resulted in the loss of RNA bands that are not full length.

[0201] Luciferase reporter gene was then cloned into each of these replicon vectors. Cells were electroporated with individual replicon RNAs and analyzed by both bulk-cell luciferase assay and flow cytometry technique. Flow cytometry analysis allowed both percent cell transfection determination and mean fluorescence intensity (MFI) evaluations as a measure of protein expression. An example of flow cytometry analysis with replicons

expressing either rFF or gREN luciferase genes is shown in **FIGs. 1A-1D**. All versions of both reporter gene replicons expressed similar levels of luciferase on a per cell basis (**FIGs. 1C and 1D**). Some variation in the percent of cells transfected based on the replicon vector version was observed (**FIGs. 1A and 1B**). The EAV replicon vector rE2 was selected as the base vector in subsequent experiments of replicon construction and optimization.

EXAMPLE 3

Monovalent EAV replicon design

[0202] This Example describes experiments performed to construct and evaluate a monovalent EAV replicon, which was subsequently deployed for expression of single polypeptides in recombinant cells.

[0203] As discussed above, the present disclosure is partly based on significant showing of unexpected results in connection with various arterivirus expression vectors designed and evaluated by the inventors. The first unexpected result relates to the amount of the structural protein gene ORF 2a that should be retained in a replicon to maintain TRS2 activity based on the published literature. In particular, in a study published in 2000, Molenkamp *et al.* concluded that the ORF 2a sequence is needed in order to retain robust TRS2 subgenomic transcription activity (Molenkamp *et al* 2000). This conclusion was drawn based on an observation that an EAV infectious clone deleted from nucleotide residues 9,756 to 12,351 (termed mutant 2a-2594), that retained only 5 bases of the ORF 2a sequence, exhibited a significant reduction in subgenomic RNA synthesis. This observation was in contrast to the subgenomic RNA synthesis exhibited by a different EAV infectious clone (mutant 030-2319). Mutant 030-2319 contained the same 3' sequences as mutant 2a-2594 but maintained an intact ORF 2a sequence and this construct demonstrated wild-type robust subgenomic RNA synthesis (Molenkamp *et al* 2000).

[0204] Surprisingly and in contrast to the above teachings of Molenkamp *et al* 2000, the base replicon design of the inventive work described in the present disclosure (Rep-EAV (WT)) and all subsequent versions were completely devoid of any ORF 2a sequences yet each demonstrated robust subgenomic transcription and expression of a gene of interest (GOI). A schematic of a base replicon design in accordance with some embodiments of the present disclosure is shown in **FIG. 2**. The TRS7 cassette driving expression of a gene of

interest (GOI, which was gRen in this example) has been cloned immediately downstream of the ORF1b stop codon.

EXAMPLE 4

Bivalent EAV replicon design

[0205] This Example describes experiments performed to construct and evaluate some bivalent EAV replicon, which were subsequently deployed for expression of two different polypeptides in recombinant cells. In these experiments, two initial bivalent designs were generated and they were based on iterations using the rE2 backbone described above. The construction of version 'A' bivalent replicon utilized the unique *XbaI* restriction site present between TRS2 and TRS7 to introduce a coding sequence of the gene of interest (GOI). Version 'B' bivalent constructs maintain both TRS2 and TRS7 as a tandem cassette for cloning of the GOI coding sequence (also referred to as a 2/7 block). A schematic representation of the A and B bivalent designs is shown in **FIG. 3**.

[0206] Two versions of each design were constructed so that each of the reporter genes could be tested in either the first or second position in the bivalent replicons. Cells were electroporated with both designs and luciferase expression level were compared with a monogenic replicon vector. The results of a representative bulk-cell luciferase assay carried out for all the bivalent versions generated as described above are shown in **FIGs. 4A-4B**. It was observed that the level of reporter gene expression from the bivalent vectors was lower than that detected from the monovalent version no matter which position the respective reporter gene was cloned into. Nevertheless, expression of both reporter genes from a single RNA was detected demonstrating that the bivalent EAV replicons constructed as described in this example were functional. Furthermore, the expression of the 5' most gene driven from TRS2 alone was no different than when it was driven as a TRS 2/7 block in direct contradiction to the requirement for ORF2a sequence teachings of Molenkamp *et al* 2000 described above

[0207] The second unexpected result relates to the requirement of the 3' EAV sequences important for EAV replication and transcription. Specifically, Molenkamp *et al.* 2000 define the optimal 3' terminal sequences that should be maintained for efficient replication using mutant 030-2319. Molenkamp *et al.* 2000 teach that the 3' terminal 354

nucleotides of the EAV genome are able to support wild type level of replication. However, surprisingly and unexpectedly, we found that replicon vectors described herein (especially those expressing more than one genes of interest) did not replicate efficiently unless significantly more 3' terminal sequences were included in the replicon vectors.

[0208] In particular, in some experiments, an additional 801-nt of 3' terminal sequence was included in a number of EAV replicon vectors described herein. This additional 3' terminal sequence contained all but the first two nucleotides of ORF6, a TRS7 sequence, all of the ORF7 gene and the 3' UTR sequence of an EAV genome. Accordingly, the ATG start codon of ORF6 is absent in this particular design. Two versions of a monovalent EAV replicon expressing the rFF luciferase gene were constructed. The first replicon maintained a functional TRS2, a TRS7 to drive the reporter gene, and a functional TRS7 sequence in the 801-nt region (vector termed rEna). The second replicon was the same as the rEna vector except that both of the TRS7 elements were modified to be inactivated (vector termed rExa). A schematic representation of each of the vectors is shown in **FIG. 5**.

[0209] An example of flow cytometry analysis with replicons rExa-rFF and rEna, both of which expressing the rFF luciferase gene, is shown in **FIGs. 6A-7B**. It was observed that incorporation of the additional 3' sequence had a significant positive impact on replication of both replicons containing additional 3' sequences while expression levels were not impacted. Furthermore, the rExa vector that only has a functional TRS2 element expresses an equivalent amount of protein to the rEna vector, indicating that no aspect of the ORF 2a structural gene was required for robust replication or transcription of these replicons.

[0210] In the context of a bivalent replicon expressing gREN and rFF luciferase genes, three versions of this replicon vector were also constructed. The first replicon maintained a functional TRS2 to drive the first reporter gene, a TRS7 to drive the second reporter gene and a functional TRS7 sequence in the 801-nt region (vector termed rEnb). The second replicon was the same as the rEnb vector except that both of the TRS7 elements were modified to be inactivated (vector termed rExb). The third replicon maintained a functional TRS2 to drive the first reporter gene, a TRS7 to drive the second reporter gene and an inactivated TRS7 sequence in the 801-nt region (vector termed rExc). A schematic representation of each of these vectors is shown in **FIG. 7**.

[0211] The inventors have additionally demonstrated that addition of the 801-nt 3' region was required for robust replication of replicons carrying more than one GOI. An example of flow cytometry analysis with replicons expressing both rFF and gREN luciferase genes from vectors that do or do not contain the additional 801-nt 3' region (rEnb-gRen-rFF and rE2-gRen-rFF, respectively) is shown in **FIGs. 8A-8D**. Addition of the 801-nt 3' sequence resulted in ~4 fold increase in replication (**FIGs. 8A and 9B**) and a ~10 fold increase in GOI expression (**FIGs. 8C and 8D**).

[0212] Next a comparison of the rEnb and rExb bivalent replicons was conducted. An example of flow cytometry analysis for this comparison is shown in **FIGs. 9A-9D**. The data indicate that having a functional TR7 in the 801-nt sequence is not required for the increase in replication as the transfection efficiency noted was not impacted in the rExb design when compared to the rEnb design (**FIGs. 9A and 9B**). Mutating the TRS7 controlling the rFF luciferase gene (rExb-gRen-rFF) significantly reduced the expression of the rFF reporter while the gREN expression (controlled by the TRS2 element) was not impacted in either of the replicon designs (**FIGs. 9C and 9D**).

[0213] In order to restore expression from the second reporter gene (rFF) a vector with a functional TRS2 element to drive the first gene, a functional TRS7 element to drive the second gene and an 801-nt region with a mutated TRS7 element was constructed (rExc-gRen-rFF) and it was compared to the rEnb-gRen-rFF replicon design. An example of flow cytometry analysis for this comparison is shown in **FIGs. 10A-10D**. Restoration of the TRS7 element in the rExc-gRen-rFF resulted in robust rFF expression comparable to that seen with rEnb-gRen-rFF (**FIG. 10D**).

EXAMPLE 5

Trivalent EAV replicon design

[0214] A replicon vector containing three different luciferase genes was constructed as follows. The bivalent vector rEnb-gRen-rFF was modified to include a third luciferase gene (Cypridina (Cypr)) by designing oligonucleotides to amplify the Cypridina gene from a monovalent rE2-Cypr backbone. The PCR product was designed to include TRS7 and its flanking sequences to produce a TRS7 Cyp expression cassette. The PCR product was introduced into the rEnb-gRen-rFF vector by using the Gibson Assembly®

procedure downstream of the rFF reporter gene, thereby generating the tri-genic rEnb-gRen-rFF-Cyp vector.

[0215] Cells were electroporated with individual tri-genic replicon RNAs and analyzed by both bulk-cell luciferase assay and flow cytometry technique. The results of these experiments are shown in **FIGs. 11A-11B** where two types of data are shown. FIG. 11A). The top row of the figures is showing the percent of cells transfected with each of the replicons; this is a measure of replication activity for each RNA replicon construct. FIG. 11B). The lower row is showing protein expression for each of the luciferase constructs from electroporated cell lysates and normalized for the amount of lysate used in the assay. The TRS used to drive each of the genes is also shown for the trivalent construct. RLU: relative light units.

EXAMPLE 6

EAV TRS spacer design

[0216] Incorporation of additional 3' UTR sequences represents one approach to modulate expression of a gene of interest in the RNA replicon designs. EAV vectors with shorter 3' UTR (*e.g.*, rE2 vector) express lower amounts of protein than EAV vectors with additional 3' UTR sequences (*e.g.*, rEn or rEx vectors). Development of multiple methods to attenuate or modify GOI expression levels from the EAV replicon is another key aspect of the inventive work described herein. Another approach to tune protein expression involves increasing or decreasing the amount of native sequence surrounding the body TRS elements. An example of employing this strategy is described below. Four different monogenic spacer replicons were designed that include varying amounts of EAV sequence (**FIGs. 12A-12B**). The base rE2 vector represents the starting point for the modifications. The incorporated spacers were delimited by regions of high homology separated by AT rich runs. There were two such regions 5' of TRS7 resulting in ORF6 spacers of 343 bp and 220 bp (Spacer 1 and Spacer 2, respectively). An additional 98 bp of ORF7 sequence was included; the ORF7 wild type ATG was inactivated and no additional ATG were present in the 3' spacer construct (Spacer 3). A fourth construct (Spacer 4) included both the Spacer 1 and Spacer 3 sequences.

[0217] The rFF luciferase gene was cloned into each of the spacer-containing replicons described above and each resulting RNA replicons was electroporated into cells.

The results of flow cytometry analysis of the electroporated cells are shown in **FIGs. 13A-13B**. It was observed that introduction of the spacer regions impacted expression levels and showed a range of activity relative to the rE2-rFF base vector. These results demonstrate that an effective approach to modulate GOI expression by modifying the amount of sequence included either 5' or 3' of the TRS used for subgenomic transcription.

[0218] In a subsequent experiment, the spacer regions described above were used to generate bigenic EAV replicons. For this purpose, the spacer regions were introduced either upstream of the 5' most gene, upstream of the 3' most gene or upstream of both genes. An example of the impact of the inclusion of Spacer 1 in an EAV replicon expressing both the gREN and rFF reporter genes is shown in **FIGs. 14A-14B**. In conclusion, the location of a spacer sequence, for instance Spacer 1, in the bivalent replicons impacted luciferase expression levels representing another example of protein expression modulation in EAV replicon vectors.

EXAMPLE 7

RNA structure sequence

[0219] RNA structure sequence analysis (Tijerina 2007; Ding 2014) was conducted on wild type EAV as well as on the RNA replicons of the inventive compositions and methods described herein. That analysis has revealed key non-TRS sequence elements that significantly impact subgenomic transcription levels. The result of this novel approach is that we have developed a method that can be used to rationally tune GOI expression levels.

EXAMPLE 8

Impact of the primary sequence of the gene of interest (GOI)

[0220] This example illustrates a fourth approach to modulate GOI expression from EAV replicons. The development of this approach was based on the understanding that modifying codon usage of GOI can impact RNA secondary structure.

[0221] Four different codon usage versions of the H5N1 influenza A/VietNam/1203/04 hemagglutinin (HA) gene and the respiratory syncytial virus (RSV) fusion (F) gene were generated. The nucleotide sequences of the codon-optimized fusion

glycoprotein F0 and HA genes are provided in the Sequence Listing (SEQ ID Nos: 16-19 and 20-24).

[0222] The different HA and F codon-optimized genes were each cloned into the rE2 vector. Cells were electroporated with RNA generated from each of the constructs and the cells were then analyzed by flow cytometry with protein-specific antibodies. The results of flow cytometry analysis of the different HA replicons are shown in **FIGs. 15A-15B**. Significant differences in both replication and protein expression were noted from replicons coding for the same protein but having different primary sequences. More than a 50-fold difference in replication was noted between the different HA versions (**FIG. 15A**). Modulation in protein expression (2-4 fold) was also demonstrated with the different HA gene versions (**FIG. 15B**).

[0223] Further, the results of flow cytometry analysis of the different F replicons are shown in **FIGs. 16A-16B**. Similar to what was noted with HA differences in both replication and protein expression were noted from replicons coding for the same F protein but having different primary sequences. Interestingly, replication did not always predict protein expression as the rE2-F (GA) replicon expressed as much or more F protein than the replicons yet had the lowest transfection percentage.

EXAMPLE 9

Analysis of GOI expression from EAV replicon vectors

[0224] Another non-limiting unexpected aspect of the inventive work described in the present disclosure is the magnitude of protein expression that the RNA replicons described herein are capable of. It has been previously reported in the RNA replicon field that alphavirus-based replicon systems are capable of expressing up to twenty percent of a cell's total protein content (Pushko 1997). Thus, it is surprising and unexpected that the inventive work described here is capable of even higher expression levels on a per cell basis than an alphavirus replicon based on the fact that alphaviruses grow to titers 2-3 orders of magnitude higher than EAV does (Castillo-Olivares 2003). Two examples of the EAV replicon GOI expression potential are described below. The gREN luciferase gene or green fluorescent protein (GFP) genes were cloned into the rE2 vector. The two genes were also cloned into an alphavirus replicon vector based on the TC-83 strain of Venezuelan equine

encephalitis virus (Hooper 2009). An equivalent amount of RNA *in vitro* transcribed from each replicon was electroporated into cells. Cells were analyzed by flow cytometry to determine both the percent of cells transfected as well as the GOI mean fluorescent intensity (MFI) as an assessment of protein expression. The results of a representative experiment are shown in **FIGs. 17A-17C** and **FIG. 18**. For cells transfected with gREN expressing replicons (**FIGs. 17A-17C**), approximately three times as many cells were transfected with the alphavirus gREN RNA than the rE2-gREN RNA (**FIG. 17A**) but the MFI for rE2-gREN electroporated cells was more than 1.5 times higher than the alphavirus gREN cells (**FIG. 17B**). Bulk luciferase assays performed on the cells in parallel indicate that even though three times fewer cells received replicon RNA the rE2-gREN produced an equivalent amount of luciferase.

[0225] A similar higher expression level was detected in cells transfected with GFP expressing replicons (**FIG. 18**). In these experiments, cells electroporated with rE2-GFP expressed more than 1.5 times more GFP reporter protein than the alphavirus GFP replicon electroporated cells (**FIG. 18**).

EXAMPLE 10

Molecular evolution of EAV replicons for specific phenotypes

[0226] This Example demonstrates the ability to tune protein expression levels when the replicon vector has been modified to have additional characteristics specific for the intended use of the system. If extended GOI expression time is required for a particular indication, a vector with long term protein expression would be ideal. Ultimately, the impact of EAV replicon RNA replication in a cell would result in cell death. To determine when that occurs with EAV replicons, an analysis of when cell toxicity occurs *in vitro* was carried out. Different cell types were transfected with rE2-GFP RNA, and the cells were subsequently monitored for the presence of cytopathic effects (CPE). Time course studies were conducted to determine how long the different cell types (BHK-21, CHO and HEK-293) could maintain EAV replicons before CPE eliminated the presence of green cells. GFP was detected in cells for up to four days before CPE was complete in all of the cells tested.

[0227] To generate an EAV replicon that is capable of expressing a GOI in a cell for more than four days, molecular evolution experiments were conducted by cloning a

selective marker (neomycin or puromycin) into the rE2 replicon vector (rE2-neo or rE2-pur, respectively). Cells were transfected with rE2-neo replicon RNA and 24 hours after transfection the cells were put under 400-600 µg/ml geneticin antibiotic selection. By 72 hours post antibiotic treatment all cells in control wells were dead while patches of growing cells from samples that had been transfected with the rE2-neo RNA remained for up to 12 days. In an exemplary experiment performed to assess molecular evolution of EAV replicon vectors for extended expression of a gene of interest (GOI) *in vitro*, BHK-21 cells electroporated with rE2-neo RNA were placed under geneticin antibiotic selection and the growth under selection of a patch of cells at 5 and 6 days were examined. In this experiment, a patch of BHK-21 cells were found significantly expanding from Day 5 to Day 6 while under selective pressure. In comparison, all control cells had died by Day 3. The rE2-neo vectors were molecularly evolved, by antibiotic selection, to express protein in cells for up to 3 times longer than was possible with rE2-GFP. In conclusion, the results of this experiment illustrates that the molecularly evolved EAV replicons represent new vectors with longer term protein expression capability.

EXAMPLE 11

In vivo EAV replicon expression analysis

[0228] This Example summarizes the results of experiments assessing expression from EAV replicons, which were carried out *in vivo* using whole body imaging analysis to detect rFF luciferase expression. An example of IVIS® analysis in mice injected with 30 µg of rE2-rFF RNA is shown in **FIGs. 19A-19B**. The schedule of injection and image analysis as well as representative IVIS whole body analysis study is shown in **FIGs. 19A** and **19B**, respectively. Luciferase activity was detected in the lungs of all rE2-rFF injected mice (15 out of 15 animals) between days 4 and 7 post injection. This data shows *in vivo* protein expression from the EAV replicon vector.

EXAMPLE 12

Expression of antibody in EAV replicon expression system

[0229] SGI's Archetype® Software was used to generate Herceptin codon optimized DNA sequences for expression in CHOK1 cells. These sequences were then

synthesized *de novo* from oligonucleotides, cloned into a plasmid, and their sequences were then verified. Assembly of the Herceptin expressing EAV replicon was performed as follows. The EAV replicon backbone was PCR amplified using primers specific to the 3' end of pplab, which includes the TRS2, and 5' end of ORF6 excluding the start codon. The Herceptin light chain (LC) forward primer and heavy chain (HC) reverse primer were designed with the gene specific amplification region and an additional 30bp of sequence complementary to the EAV backbone. The TRS7 sequence (84-bp), which controlled the expression of HC, was introduced on the LC reverse primer and HC forward primer. These two primers were designed with the gene specific amplification region and 65bp of the TRS7 Promoter region. The 46-bp portion of homology was shared between the LC and HC PCR products allowing for overlap extension PCR amplification to join the two products together to generate a single fragment containing the sequences of Herceptin LC and TRS7, which was followed by the sequence of Herceptin HC. Two-step Gibson Assembly® procedure as described in Example 1 was performed with the EAV replicon backbone and the Herceptin LC-TRS7-HC gene fragment. The assembly reaction was transformed into *E. coli* TransforMax EPI300 cells (Epicentre Cat. no. EC300105) and plated on selective LB agar plates. *E. coli* clones were screened using colony PCR with primers annealing within the EAV backbone just outside the assembly junctions. *E. coli* clones containing the Herceptin gene cassette could be easily identified based upon expected PCR fragment size. Positive *E. coli* clones were then verified by sequencing using Sanger and Illumina MiSeq platforms.

[0230] Analysis of antibody expressed from EAV replicon expression system

[0231] Antibody expressed from EAV replicon expression systems described in the disclosure was analyzed by using solid-phase binding assays. Media from BHK-21 cells was collected 24 h after transfection with the replicon and treated with protease inhibitor (Pierce). To measure Herceptin expression, a quantitative total IgG capture ELISA was performed. An anti-human IgG heavy chain (mouse anti-Human IgG unlabeled, # 9040-01, Southern Biotech) was coated on microtiter wells overnight. The wells were incubated with 100 µl of media, probed with an HRP-labeled anti-human IgG light chain (mouse Anti-Human Kappa-HRP, # 9230-05, Southern Biotech) and developed with TMB substrate solution (Pierce). Standard curves of a human IgG1 antibody (IgG1 Kappa-UNLB, 0151K-

01, Southern Biotech) and Herceptin were used to estimate the amount of antibody expressed by BHK-21 cells in pg/cell/day. A similar assay was used for specific antigen binding, in which recombinant Erb2 receptor (Thermo Fisher Scientific) was used to capture the Herceptin antibody present in the cell culture media. In this experiment, standard curves were based on Herceptin and were correlated with the total IgG amount.

EXAMPLE 13

Construction of EAV replicon vectors with GOI expression under control of TRS1 subgenomic promoter

[0232] In this experiment, an EAV replicon was engineered to use the TRS1 subgenomic promoter to express the red firefly (rFF) luciferase reporter. A comparison of the level of rFF expression from another EAV replicon is shown in **FIGS. 21A** and **21B**. In this experiment, BHK cells were electroporated with 3 µg of replicon RNA. The TRS1 replicon vector demonstrated robust expression that was higher than expression detected from an EAV replicon using the TRS7 subgenomic promoter. In **FIGS. 21A-21B**, the dotted line represents the amount of expression detected from a replicon engineered to use the TRS7 subgenomic promoter to drive the expression of rFF reporter. As illustrated in **FIGs. 21A-21B**, robust expression was detected from the TRS1 replicon, as indicated by transfection efficiency (**FIG. 21A**) and Luciferase expression level (**FIG. 21B**).

EXAMPLE 14

Construction of VBS-R-eGFP and VBS-IC

[0233] To extend vector development to additional EAV strains, the virulent Bucyrus strain (VBS) was selected. The VBS strain is more virulent than the highly attenuated EAV030 strain, as such, a replicon based on it may have different expression characteristics than the replicon based on the EAV030 strain. To this end, the complete genomic sequence for the VBS strain was downloaded from Genbank (Accession No. DQ846751). The sequences of a VBS replicon containing an eGFP reporter gene driven by TRS2 subgenomic promoter within the polyprotein pp1b gene (VBS-R-eGFP) and a VBS infectious clone (VBS-IC) containing overlap regions to pW70, the T7 promoter, and a polyA

tail with 125 A's, are provided in the Sequence Listing as SEQ ID NO: 46 and SEQ ID NO: 47, respectively.

[0234] The VBS-R-eGFP and VBS-IC constructs were intended to be built via homologous recombination in *S. cerevisiae* using an *E. coli*-yeast shuttle vector pW70. The VBS-R-eGFP construct was split into 7 fragments (excluding the vector), the first 6 of which are shared by the VBS-IC with two additional fragments, as shown in **FIG. 22** and **FIG. 23**, respectively.

[0235] The g-blocks for the fragments were ordered from IDT for assembly. To be used as a vector, pW70 was digested with restriction enzymes *AhdI/NotI*. The last fragment of each construct was PCR-amplified for the addition of the polyA tail using an appropriate primer set.

[0236] Yeast colonies were obtained from plating of 100 μ L for both constructs (VBS-R-eGFP and VBS-IC). Colonies were pooled for each construct, and plasmids were isolated from each pool.

[0237] The isolated plasmid pool of each construct was then transformed into *E. coli* (EPI300 from Epicenter) for screening in bacteria. The transformed culture (10-100 μ L) was plated.

[0238] *E. coli* colonies from transformations of plasmids pools isolated from yeast colonies were screened for 5' and 3' ends.

[0239] Multiple "positive" clones were found. Sanger sequencing results revealed two completely sequence-correct clones for the VBS-IC (clones 33 and 36) and 4 positive clones for VBS-R-eGFP (clones 6, 30, 33, and 47). A schematic map of the pBR322+VBS-R-eGFP is shown in **FIG. 24**.

Demonstration of the functionality of VBS IC and VBS-R-eGFP replicons

[0240] The VBS IC was tested to demonstrate functionality by producing *in vitro* transcribed RNAs from the VBS IC c33 plasmid DNA, followed by electroporating the transcribed RNAs into BHK cells. As a positive control, RNA generated from an infectious clone for EAV strain EAV030 was included in a separate electroporation. The generation of cytopathic effects (CPE) in cells electroporated with IC RNA indicated that the IC RNA was functional. As shown in **FIG. 25**, CPE was noted in both the VBS and EAV030 strain by 48

hours post electroporation (see, e.g., **FIG. 25**). This data demonstrated that the VBS IC RNA was functional.

[0241] The VBS-R-eGFP vector was tested to demonstrate functionality by producing *in vitro* transcribed RNA from the VBS-R-eGFP c30 plasmid DNA, followed by electroporation the transcribed RNA into BHK cells. As a positive control, RNA generated from an EAV strain EAV030 replicon expressing eGFP was included in a separate electroporation. BHK cells were electroporated with 3 µg of either VBS-R-eGFP RNA (VBS-R-TRS2-eGFP) or EAV030 eGFP replicon RNA (rE2-GFP). An example of flow cytometry analysis for this comparison is shown in **FIG. 26**. Analysis of the mean fluorescence intensity of GFP expressed from both vector backbones demonstrated that the EAV strain VBS and EAV strain EAV030 replicon systems were expressing protein at very similar levels and that the VBS replicon system was completely functional.

EXAMPLE 15

Construction of VBS-R-rFF replicon

[0242] To evaluate VBS-R in mice for IVIS® analysis as an alternative backbone, VBS-R-rFF was needed to be made. Therefore, the primers were designed and ordered to replace eGFP in pBR322+VBS-R-eGFP. A schematic map of pBR322+VBS-R-rFF is shown **FIG. 27**.

[0243] Two fragments, PCR-amplified using the primers RP52 and PR53 and the primers RP54 and RP20, respectively, from their appropriate templates, were gel-extracted and assembled via a fusion PCR using RP55 and RP6 as primers. The resulting assembled fragment was then gel-extracted again and cloned into the vector pBR322+VBS-R-eGFP pre-digested with restriction enzymes *PmeI* and *NotI*. One clone- Clone 7 - was sequence-confirmed to be completely correct via MiSeq procedure. The sequence of resulting construct VBS-R-rFF includes the T7 promoter and a polyA tail with 125 A's, and is set forth in the Sequence Listing as SEQ ID NO: 48.

EXAMPLE 16

Construction of VBS-R-TRS7-rFF replicon

[0244] In pBR322+VBS-R-rFF (c7), the expression of rFF reporter is driven by TRS2 embedded in the pp1b gene. To make pBR322+VBS-R-TRS7-rFF, where rFF expression would be driven by TRS7 located downstream of the pp1b gene, a g-block consisting of the following sequence was used as an insert and cloned into the vector pBR322+VBS-R-rFF (c7) pre-digested with restriction enzymes *PmeI/EcoRI* via Gibson Assembly® procedure. Clone 11 was sequence-confirmed to be completely correct via MiSeq. The sequence of g-block for construction of pBR322+VBS-R-TRS7-rFF is provided in the Sequence Listing as SEQ ID NO: 49.

[0245] A schematic map of the pBR322+VBS-R-TRS7-rFF is shown in **FIG. 28**.

[0246] The sequence of VBS-R-TRS7-rFF is provided in the Sequence Listing as SEQ ID NO: 50 with the T7 promoter and a polyA tail with 125 A's.

[0247] The VBS-R-TRS2-rFF (previously identified as pBR322+VBS-R-eGFP above) and VBS-R-TRS7-rFF vectors were tested to demonstrate functionality by producing *in vitro* transcribed RNA from each plasmid DNA (clone 7 and clone 11, respectively), followed by electroporating the transcribed RNA into BHK cells. As a positive control, RNA generated from an EAV strain EAV030 replicon expressing rFF (rE2-rFF) was included in a separate electroporation. BHK cells were electroporated with 3 µg of each RNA and cells were analyzed by FACS. An example of flow cytometry analysis for this comparison is shown in **FIG. 29**, in which both VBS-rFF vectors were observed to express luciferase protein. These data demonstrate that the VBS replicons described herein can express the rFF reporter gene and these vectors can be used for *in vivo* expression analysis.

[0248] The VBS-R-TRS2-rFF replicon was further analyzed *in vivo* in Balb/c mice. In this experiment, a range of RNA doses were tested and whole body imaging was conducted one and three days post intramuscular injection. Interestingly, animals that received the highest dose were observed to express luciferase less well than animals that received the lowest dose. A summary of the results of these experiments is presented in **FIG. 30**. These data demonstrate the *in vivo* activity of the VBS replicon vector.

EXAMPLE 17

Construction of SHFv-R-TRS7-rFF

[0249] In addition to the EAV strain EAV030 and EAV strain VBS replicon systems, an alternative backbone was constructed based on another virus within the *Arterivirus* genus. This was to extend the development of arterivirus replicon systems based on other species within the *Arteriviridae* Family of viruses. A replicon vector system derived from the Simian Hemorrhagic Fever virus (SHFv) was developed. The complete genomic sequence of SHFv was downloaded from Genbank (Accession No. AF180391 L39091 U20522 U63121). The complete SHFv nonstructural gene sequence and a portion of the 3' most structural region was divided into 9 fragments including rFF, and the corresponding g-blocks except rFF and respective primers for amplifying each fragment were ordered. In this design, rFF expression is driven by an independent TRS7 subgenomic promoter.

[0250] The reporter rFF gene was PCR-amplified. The PCR-amplified fragments were then assembled in *S. cerevisiae* following the previously described transformation protocols.

[0251] A schematic map of the pW70+SHFV-R-TRS7-rFF construct is shown in **FIG. 31**. In addition, the sequence of the SHFV replicon with rFF expression driven by TRS7 (SHFV-R-TRS7-rFF) is provided in the Sequence Listing as SEQ ID NO:51, which includes overlap regions to the vector pW70, the T7 promoter, and a polyA tail with 125 A's.

[0252] Plasmids from the yeast colonies resulting from the assembly transformation of pW70+SHFV-R-TRS7-rFF construct were isolated as a pool and transformed back into *E. coli*. A total of 12 of the resulting *E. coli* colonies were sequenced via MiSeq. Error correction and assembly of the clones were performed to correct any mutations via Gibson Assembly® procedure.

[0253] The resulting *E. coli* clones from the correcting transformation were screened by Sanger sequencing using RP77 specific for that region. Clone 2 was identified to contain the correct nucleotide, and its subsequent MiSeq sequencing result revealed that the entire replicon sequence was also found to be completely correct.

[0254] The SHFv-R-TRS7-rFF vector was tested to demonstrate functionality by producing *in vitro* transcribed RNA from the SHFv-R-TRS7-rFF plasmid DNA and electroporating it into BHK cells. BHK cells were electroporated with 3 µg of each RNA and electroporated cells were analyzed by FACS. An example of flow cytometry analysis for this

comparison is shown in **FIGs. 32A-32B**. These data demonstrate that SHFv replicon can express the rFF reporter gene and the vector can be used for *in vivo* expression analysis.

[0255] Subsequently, the SHFv-R-TRS7-rFF replicon was further analyzed *in vivo* in Balb/c mice. In this experiment, 30 µg of RNA was injected into mice and whole body imaging was conducted. A summary of the results of these experiments is presented in **FIG. 34**. These data demonstrate the *in vivo* activity of the SHFv replicon vector and that it is equivalent to the EAV replicon.

EXAMPLE 18

Antibody expression using EAV replicon expression systems

[0256] The EAV replicon has also been shown to express antibody constructs as well. The vector used for these experiments was rEx-herceptin. The light and heavy chain genes for the Herceptin monoclonal antibody were cloned into the either the first or second position within the replicon. Data collected from BHK cells electroporated with 3 µg of purified RNA from a representative rEx-herceptin construct are shown in **FIG. 33A**. The amount of antibody produced was determined by an ELISA procedure in which heavy chain capture and light chain detection to only measure complete antibody structures (**FIG. 33B**). The yields of antibody from the EAV replicon are ~10 times the amount expressed from an equivalent DNA construct. The expressed antibody was also used to detect Her2 antigen by ELISA to demonstrate activity of the antibody (**FIG. 33C**).

EXAMPLE 19

EAV RSV and cas9 expression

[0257] Several additional genes have been cloned into EAV replicons. These additional genes include mouse IL-12, respiratory syncytial virus (RSV) F and cas9. A bivalent EAV replicon expressing IL-12 and RSV F was constructed along with monovalent versions of each. These replicon RNAs were electroporated into BHK-21 cells and examined for protein expression by flow cytometry using IL-12 and RSV F specific antibodies. The results of that analysis are shown in **FIGS. 35A-35D**. Both IL-12 and RSV F protein were readily detected in electroporated cells.

[0258] A cas9 gene optimized to use two different codon usages was cloned into the EAB replicon vector. RNA was generated from both cas9 replicon vectors and electroporated into BHK-21 cells. Eighteen hours post electroporation cell lysates were generated and different amounts of cell lysate were used to treat plasmid DNA. Guide RNA (gRNA) specific for a target sequence in the plasmid was added to the cell lysate and after incubation the DNA was analyzed on an agarose gel. The results of a representative experiment are shown in **FIGS. 36A-36C**. The cas9 protein expressed from the EAV replicon was able to cleave the plasmid DNA combined with the gRNA indicating that the enzyme was active.

[0259] Throughout this disclosure, various information sources are referred to and incorporated by reference. The information sources include, for example, scientific journal articles, patent documents, textbooks, and World Wide Web browser-inactive page addresses. The reference to such information sources is solely for the purpose of providing an indication of the general state of the art at the time of filing. While the contents and teachings of each and every one of the information sources can be relied on and used by one of skill in the art to make and use the embodiments disclosed herein, any discussion and comment in a specific information source should no way be considered as an admission that such comment was widely accepted as the general opinion in the field.

[0260] 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.

CLAIMS

WHAT IS CLAIMED IS:

1. A nucleic acid molecule comprising a nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified arterivirus genome or replicon RNA comprises a sequence fragment exhibiting at least 80% sequence identity to the sequence encoding open reading frame ORF7, and wherein the modified arterivirus genome or replicon RNA is devoid of the sequence encoding ORF2a.
2. The nucleic acid molecule of claim 1, wherein the modified arterivirus genome or replicon RNA is further devoid of at least a portion of the sequence encoding one or more of the open reading frames ORF2b, ORF3, ORF4, ORF5a, and ORF5.
3. The nucleic acid molecule of claim 1 or 2, wherein the sequence fragment is devoid of the ATG start codon of ORF7.
4. The nucleic acid molecule of any one of claims 1 to 3, wherein the modified arterivirus genome or replicon RNA is further devoid of a portion of the sequence encoding ORF6.
5. The nucleic acid molecule of claim 4, wherein the modified arterivirus genome or replicon RNA is devoid of the ATG start codon of ORF6.
6. The nucleic acid molecule of claim 5, wherein the modified arterivirus genome or replicon RNA is further devoid of TRS7 or comprises a mutated TRS7.
7. The nucleic acid molecule of any one of claims 1 to 6, wherein the modified arterivirus genome or replicon RNA comprises one or more subgenomic (sg) promoters located at a non-native site, wherein each of the one or more sg promoters comprises a transcriptional regulatory sequence (TRS).

8. The nucleic acid molecule of claim 7, wherein at least one of the one or more *sg* promoters comprises a sequence exhibiting at least 80% sequence identity to a sequence selected from the group consisting of *sg* promoter 1, *sg* promoter 2, *sg* promoter 3, *sg* promoter 4, *sg* promoter 5, *sg* promoter 6, *sg* promoter 7, and a variant thereof.
9. The nucleic acid molecule of any one of claims 7 to 8, wherein at least one of the one or more *sg* promoters is a modified *sg* promoter.
10. The nucleic acid molecule of any one of claims 1 to 9, wherein the modified arterivirus genome or replicon RNA comprises one or more modified *sg* promoters located at their respective native site, wherein each of the one or more modified *sg* promoters comprises a TRS.
11. The nucleic acid molecule of claim 10, wherein at least one of the one or more modified *sg* promoters is a modified *sg* promoter 7.
12. The nucleic acid molecule of any one of claims 9 to 11, wherein at least one of the one or more modified *sg* promoter comprises one or more nucleotide modifications positioned within the primary sequence required for the formation of a secondary structure of RNA transcripts comprising the respective *sg* promoter sequence.
13. The nucleic acid molecule of any one of claims 9 to 11, wherein at least one of the one or more modified *sg* promoter comprises a nucleotide modification positioned within the sequence of the TRS.
14. The nucleic acid molecule of any one of claims 9 to 13, wherein at least one of the one or more modified *sg* promoters comprises a leader TRS or a variant thereof.
15. The nucleic acid molecule of any one of claims 9 to 14, wherein at least one of the one or more modified *sg* promoters comprises a body TRS or a variant thereof.

16. The nucleic acid molecule of any one of claims 1 to 15, wherein the modified arterivirus genome or replicon RNA comprises one or more mutated T7 transcriptional termination signal sequences.

17. The nucleic acid molecule of claim 16, wherein at least one of the one or more T7 mutated transcriptional termination signal sequences comprises a nucleotide substitution selected from the group consisting of T9001G, T3185A, G3188A, and combinations of any two or more thereof.

18. The nucleic acid molecule of any one of claims 1 to 17, wherein the modified arterivirus genome or replicon RNA comprises one or more heterologous transcriptional termination signal sequences.

19. The nucleic acid molecule of claim 18, wherein at least one of the one or more heterologous transcriptional termination signal sequences is a SP6 termination signal sequence, a T3 termination signal sequence, or a variant thereof.

20. The nucleic acid molecule of any one of claims 18 to 19, wherein at least one of the one or more heterologous transcriptional termination signal sequences is inactivated.

21. The nucleic acid molecule of any one of claims 7 to 20, further comprising one or more spacer regions operably positioned adjacent to at least one of the one or more *sg* promoters.

22. The nucleic acid molecule of claim 21, wherein at least one of the one or more spacer regions is positioned immediately 3' to the *sg* promoter.

23. The nucleic acid molecule of any one of claims 21 to 22, wherein at least one of the one or more spacer regions is positioned immediately 5' to the *sg* promoter.

24. The nucleic acid molecule of any one of claims 21 to 23, wherein the one or more spacer regions is about 20 to 400 nucleotides in length.

25. The nucleic acid molecule of any one of claims 1 to 24, further comprising one or more expression cassettes, wherein each of the expression cassettes comprises a *sg* promoter operably linked to a heterologous nucleotide sequence.

26. The nucleic acid molecule of claim 25, wherein the *sg* promoter comprises a TRS, a first flanking region positioned immediately 5' to the TRS, and second flanking region positioned immediately 3' to the TRS, wherein the first flanking region is about 5 to 400 nucleotides in length and the second flanking region is about 15 to 115 nucleotides in length.

27. The nucleic acid molecule of any one of claims 25 to 26, comprising two, three, four, five, or six expression cassettes.

28. The nucleic acid molecule of any one of claims 25 to 27, wherein the heterologous nucleotide sequence comprises a coding sequence of a gene of interest (GOI).

29. The nucleic acid molecule of claim 28, wherein the coding sequence of the GOI is optimized for expression at a level higher than the expression level of a reference coding sequence.

30. The nucleic acid molecule of any one of claims 28 to 29, wherein the secondary structure of the RNA transcript comprising the coding sequence of the GOI is optimized for improved RNA replication.

31. The nucleic acid molecule of any one of claims 1 to 30, wherein the arterivirus is selected from the group consisting of Equine arteritis virus (EAV), Porcine respiratory and reproductive syndrome virus (PRRSV), Lactate dehydrogenase elevating virus (LDV), and Simian hemorrhagic fever virus (SHFV).

32. The nucleic acid molecule of claim 31, wherein the arterivirus is an Equine arteritis virus (EAV), an EAV-virulent Bucyrus strain (VBS), or a Simian hemorrhagic fever virus (SHFV).

33. A nucleic acid comprising a nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified arterivirus genome or replicon RNA comprises a nucleotide sequence exhibiting at least 80% sequence identity to SEQ ID NO: 1 and at least 80% sequence identity to SEQ ID NO: 2,

and further wherein the modified genome or replicon RNA comprises a nucleotide sequence exhibiting at least 80% sequence identity to the sequence encoding open reading frame ORF7, and is devoid of the sequence encoding ORF2a.

34. The nucleic acid molecule of claim 33, wherein the modified arterivirus genome or replicon RNA comprises a nucleotide sequence exhibiting at least about 80% sequence identity to SEQ ID NO: 3.

35. The nucleic acid molecule of claim 33, wherein the modified arterivirus genome or replicon RNA comprises a nucleotide sequence of SEQ ID NO: 3.

36. A recombinant cell comprising a nucleic acid molecule of any one of claims 1 to 35 and 74 to 76.

37. The recombinant cell of claim 36, wherein the recombinant cell is a prokaryotic cell or a eukaryotic cell.

38. The recombinant cell of claim 36, wherein the recombinant cell is an animal cell.

39. The recombinant cell of claim 38, wherein the animal cell is a vertebrate animal cell or an invertebrate animal cell.

40. The recombinant cell of claim 36, wherein the recombinant cell is selected from the group consisting of a pulmonary equine artery endothelial cell, equine dermis cell, baby hamster kidney 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, and mouse 3T3 cell.

41. A cell culture comprising at least one recombinant cell of any one of claims 36 to 40.

42. A method for producing a polypeptide of interest, comprising culturing a host cell comprising a nucleic acid, wherein the nucleic acid comprises:

- i. a nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified genome or replicon RNA comprises a sequence fragment exhibiting at least 80% sequence identity to the sequence encoding open reading frame ORF7, and wherein the modified genome or replicon RNA is devoid of the sequence encoding ORF2a; and
- ii. one or more expression cassettes, wherein each of the one or more expression cassettes comprises a subgenomic (*sg*) promoter operably linked to a heterologous nucleotide sequence encoding a gene of interest (GOI).

43. A method for producing a polypeptide of interest in a subject, comprising administering to the subject a nucleic acid, wherein the nucleic acid comprises:

- i. a nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified genome or replicon RNA comprises a sequence fragment exhibiting at least 80% sequence identity to the sequence encoding open reading frame ORF7, and wherein the modified genome or replicon RNA is devoid of the sequence encoding ORF2a; and
- ii. one or more expression cassettes, wherein each of the one or more expression cassettes comprises a subgenomic (*sg*) promoter operably linked to a heterologous nucleotide sequence encoding a gene of interest (GOI).

44. The method of claim 43, wherein the subject is a vertebrate animal or an invertebrate animal.

45. The method of any one of claims 42 to 44, wherein the *sg* promoter comprises a TRS, a first flanking region positioned immediately 5' to the TRS, and second flanking region positioned immediately 3' to the TRS, wherein the first flanking region is about 5 to 400 nucleotides in length and the second flanking region is about 15 to 115 nucleotides in length.

46. The method of any one of claims 42 to 45, wherein the modified arterivirus genome or replicon RNA is further devoid of a portion of the sequence encoding one or more of open reading frames ORF2b, ORF3, ORF4, ORF5a, and ORF5.

47. The method of any one of claims 42 to 45, wherein the sequence fragment is devoid of ATG start codon of ORF7.

48. The method of any one of claims 42 to 47, wherein the modified arterivirus genome or replicon RNA is further devoid of a portion of the sequence encoding ORF6.

49. The method of claim 48, wherein the modified arterivirus genome or replicon RNA is devoid of the ATG start codon of ORF6.

50. The method of any one of claims 42 to 49, wherein the modified arterivirus genome or replicon RNA is further devoid of TRS7 or comprises a mutated TRS7.

51. The method of any one of claims 42 to 50, wherein at least one of the one or more expression cassettes comprises a modified *sg* promoter comprising at least one nucleotide modification introduced within the primary sequence required for the formation of a secondary structure of RNA transcripts comprising the *sg* promoter sequence.

52. The method of any one of claims 42 to 51, wherein at least one of the one or more expression cassettes comprises a modified *sg* promoter comprising at least one nucleotide modification introduced within the sequence of the TRS.

53. The method of claim 42 to 52, wherein at least one of the one or more expression cassettes comprises a leader TRS or a variant thereof.

54. The method of any one of claims 42 to 53, wherein at least one of the one or more expression cassettes comprises a body TRS or a variant thereof.

55. The method of any one of claims 42 to 54, wherein the modified arterivirus genome or replicon RNA comprises one or more mutated T7 transcriptional termination signal sequences.

56. The method of claim 55, wherein at least one of the one or more mutated T7 transcriptional termination signal sequences comprises a nucleotide substitution selected from the group consisting of T9001G, T3185A, G3188A, and combinations of any two or more thereof.

57. The method of any one of claims 42 to 56, wherein the modified arterivirus genome or replicon RNA comprises one or more heterologous transcriptional termination signal sequences.

58. The nucleic acid molecule of claim 57, wherein at least one of the one or more heterologous transcriptional termination signal sequences is a SP6 termination signal sequence, a T3 termination signal sequence, or a variant thereof.

59. The method of any one of claims 42 to 58, further comprising one or more spacer regions operably positioned adjacent to at least one of the one or more *sg* promoters.

60. The method of claim 59, wherein at least one of the one or more spacer regions is positioned immediately 3' to the *sg* promoter.

61. The method of any one of claims 59 to 60, wherein at least one of the one or more spacer regions is positioned immediately 5' to the *sg* promoter.

62. The method of any one of claims 59 to 61, wherein each of the one or more spacer regions is about 20 to 400 nucleotides in length.

63. The method of any one of claims 42 to 62, wherein the nucleic acid comprises comprising two, three, four, five, or six expression cassettes.

64. The method of any one of claims 42 to 63, wherein the coding sequence of the gene of interest is optimized for expression at a level higher than the expression level of a reference coding sequence.

65. The method of any one of claims 42 to 64, wherein the secondary structure of the RNA transcript comprising coding sequence of the gene of interest is optimized for improved RNA replication.

66. The method of any one of claims 42 to 65, wherein the arterivirus is selected from the group consisting of Equine arteritis virus (EAV), Porcine respiratory and reproductive syndrome virus (PRRSV), Lactate dehydrogenase elevating virus (LDV), and Simian hemorrhagic fever virus (SHFV).

67. The method of claim 66, wherein the arterivirus is an Equine arteritis virus (EAV), an EAV-virulent Bucyrus strain (VBS), or a Simian hemorrhagic fever virus (SHFV).

68. The method of any one of claims 42 to 67, wherein the nucleic acid comprising a nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified arterivirus genome or replicon RNA comprises a nucleotide sequence exhibiting at least 80% or more sequence identity to SEQ ID NO: 1 and at least 80% or more sequence identity to SEQ ID NO: 2, and wherein the modified genome or replicon RNA comprises a nucleotide sequence exhibiting at least 80% sequence identity to the sequence encoding open reading frame ORF7, and is devoid of the sequence encoding ORF2a.

69. The method of claim 68, wherein the modified arterivirus genome or replicon RNA comprises a nucleotide sequence exhibiting at least about 80% or more sequence identity to SEQ ID NO: 3.

70. The method of claim 69, wherein the modified arterivirus genome or replicon RNA comprises a nucleotide sequence of SEQ ID NO: 3.

71. The method of any one of claims 42 to 70, wherein the gene of interest (GOI) comprises a nucleotide sequence encoding a polypeptide.

72. The method of claim 71, wherein the polypeptide is selected from the group consisting of a therapeutic polypeptide, a prophylactic polypeptide, a diagnostic polypeptide, a nutraceutical polypeptide, an industrial enzyme, and a reporter polypeptide.

73. The method of claim 71, wherein the polypeptide is selected from the group consisting of an antibody, an antigen, an immune modulator, and a cytokine.

74. A recombinant polypeptide produced by the method of any one of claims 42 to 73.

75. A composition comprising a recombinant polypeptide of claim 74 and a pharmaceutically acceptable carrier.

76. A composition comprising a nucleic acid molecule of any one of claims 1-35 and 78 to 80, and a pharmaceutically acceptable carrier.

77. A composition comprising a recombinant cell of any one of claims 36 to 40, and a pharmaceutically acceptable carrier.

78. A nucleic acid comprising a nucleotide sequence encoding a modified arterivirus genome or replicon RNA, wherein the modified arterivirus genome or replicon RNA comprises a nucleotide sequence exhibiting at least 80% sequence identity to SEQ ID NO: 40 and at least 80% sequence identity to SEQ ID NO: 41,

and further wherein the modified genome or replicon RNA comprises a nucleotide sequence exhibiting at least 80% sequence identity to the sequence encoding open reading frame ORF7, and is devoid of the sequence encoding ORF2a.

79. The nucleic acid molecule of claim 78, wherein the modified arterivirus genome or replicon RNA comprises a nucleotide sequence exhibiting at least about 80% sequence identity to SEQ ID NO: 42.

80. The nucleic acid molecule of claim 78, wherein the modified arterivirus genome of replicon RNA comprises a nucleotide sequence of SEQ ID NO: 42.

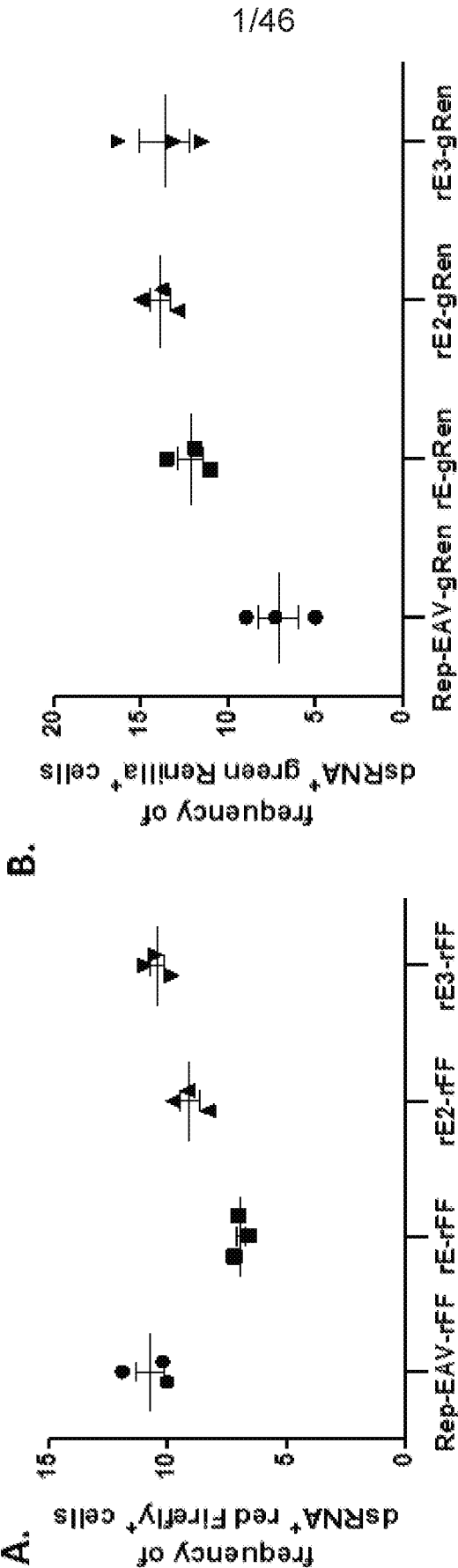


FIG. 1A

FIG. 1B

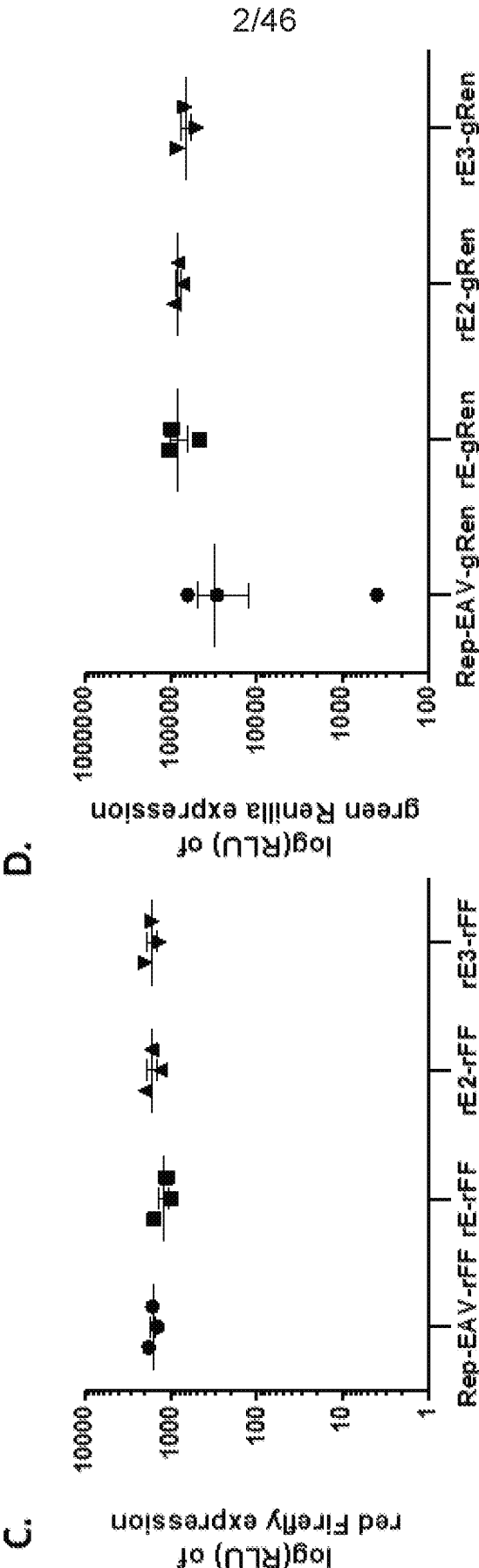


FIG. 1D

FIG. 1C

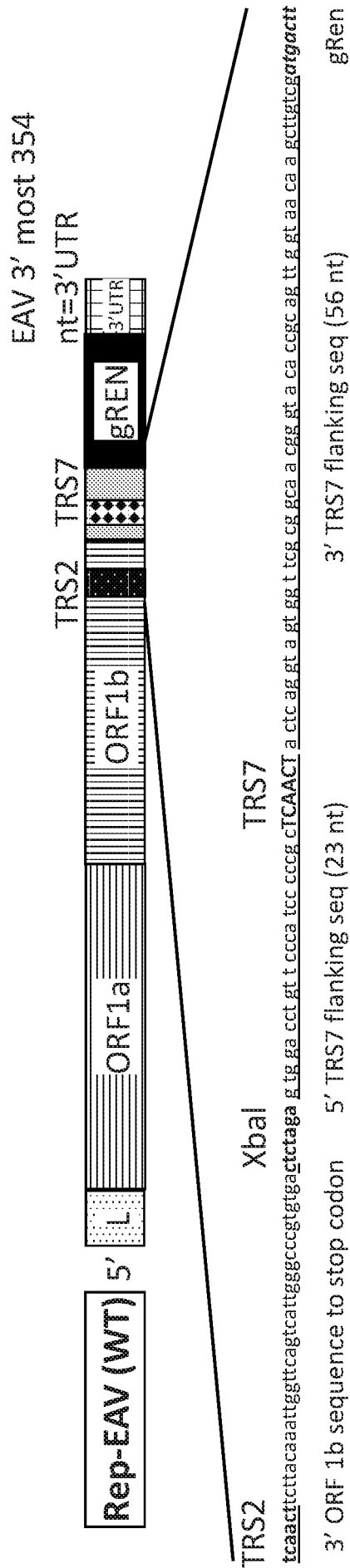


FIG. 2

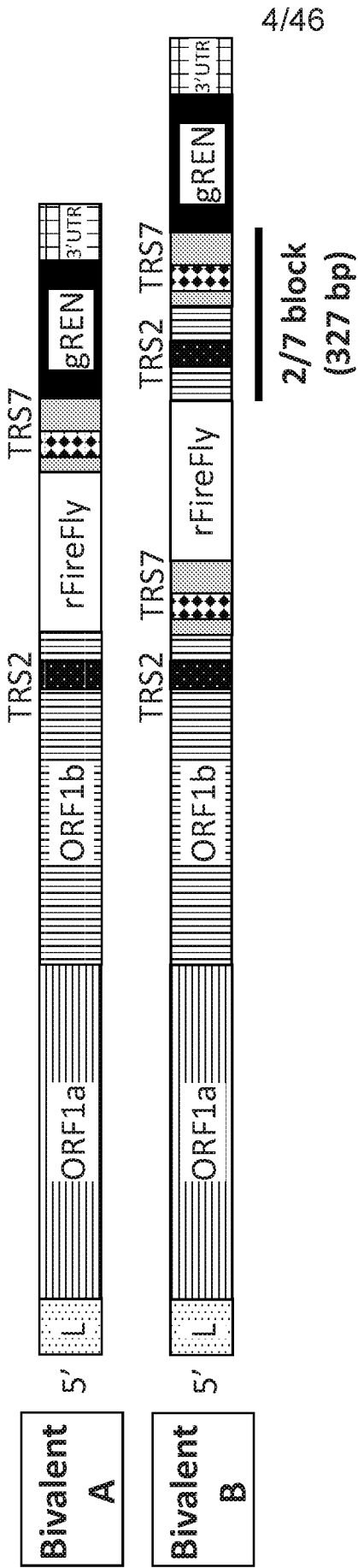
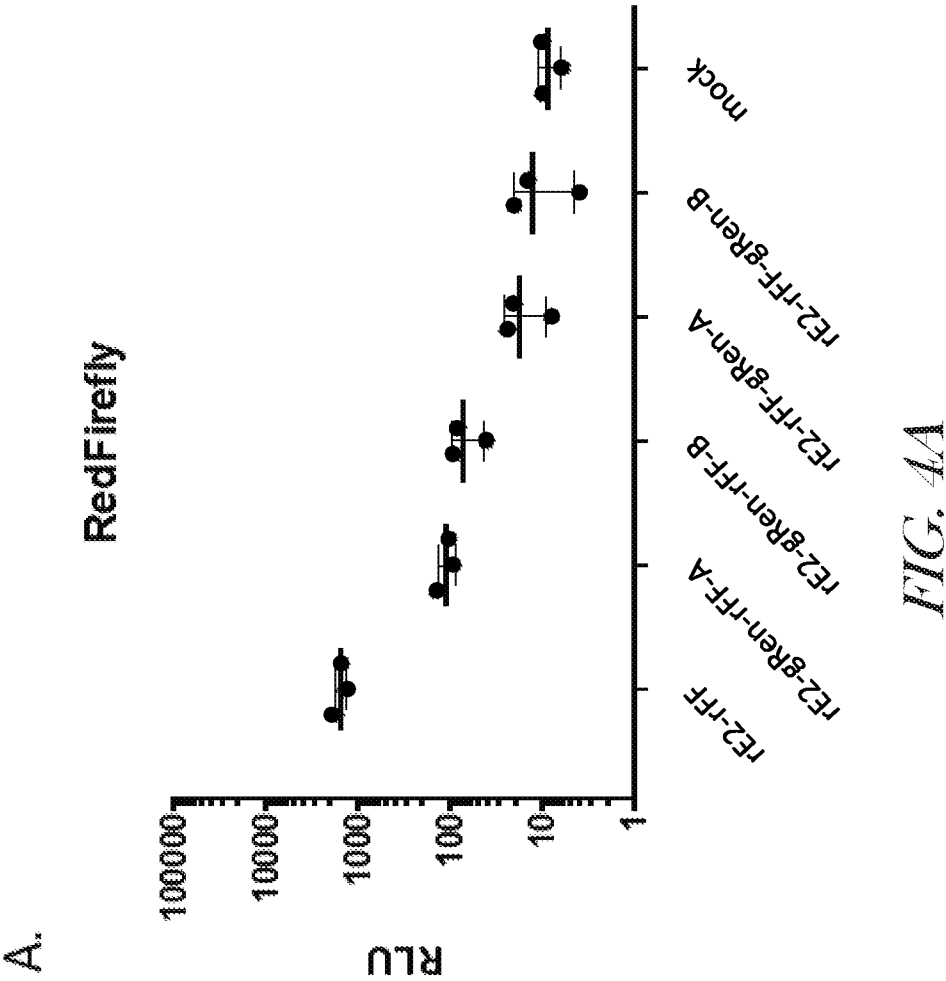
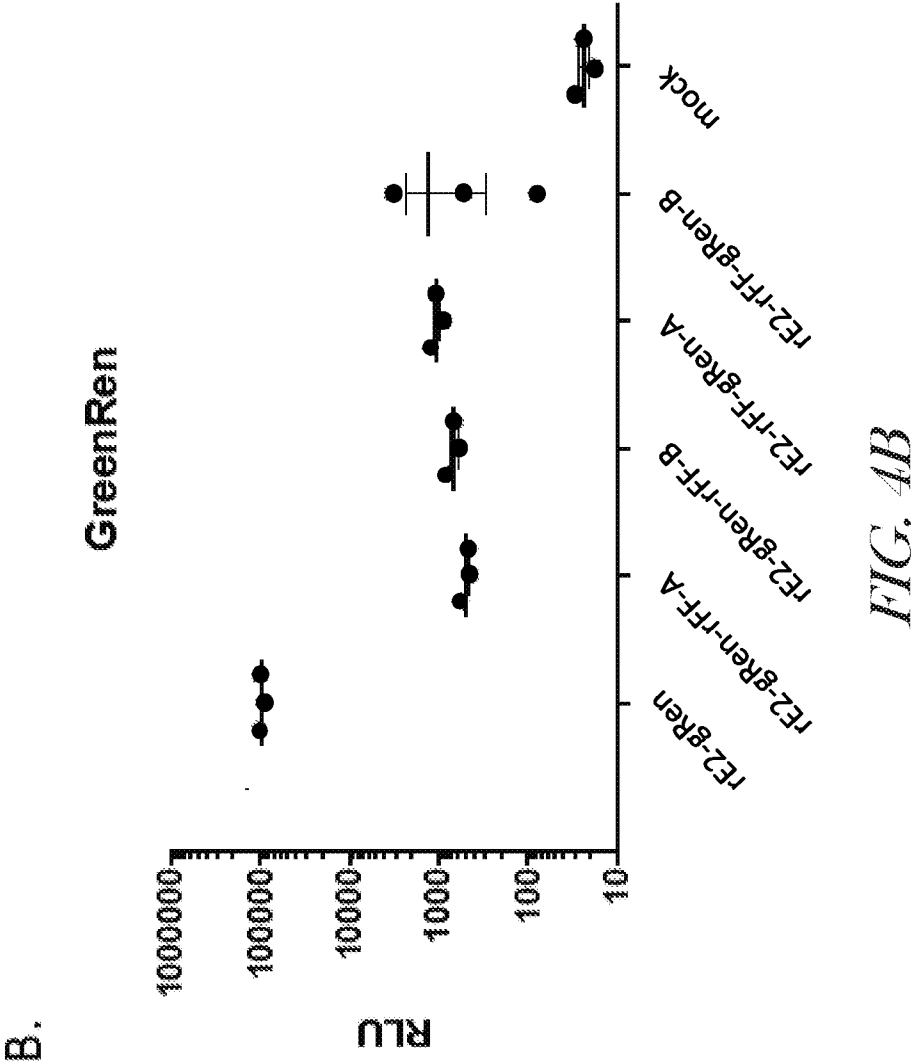


FIG. 3





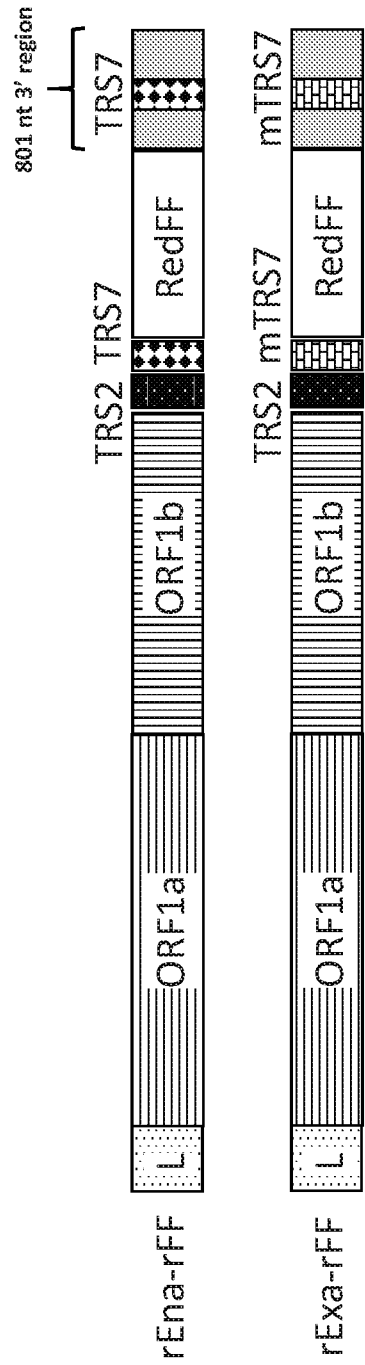


FIG. 5

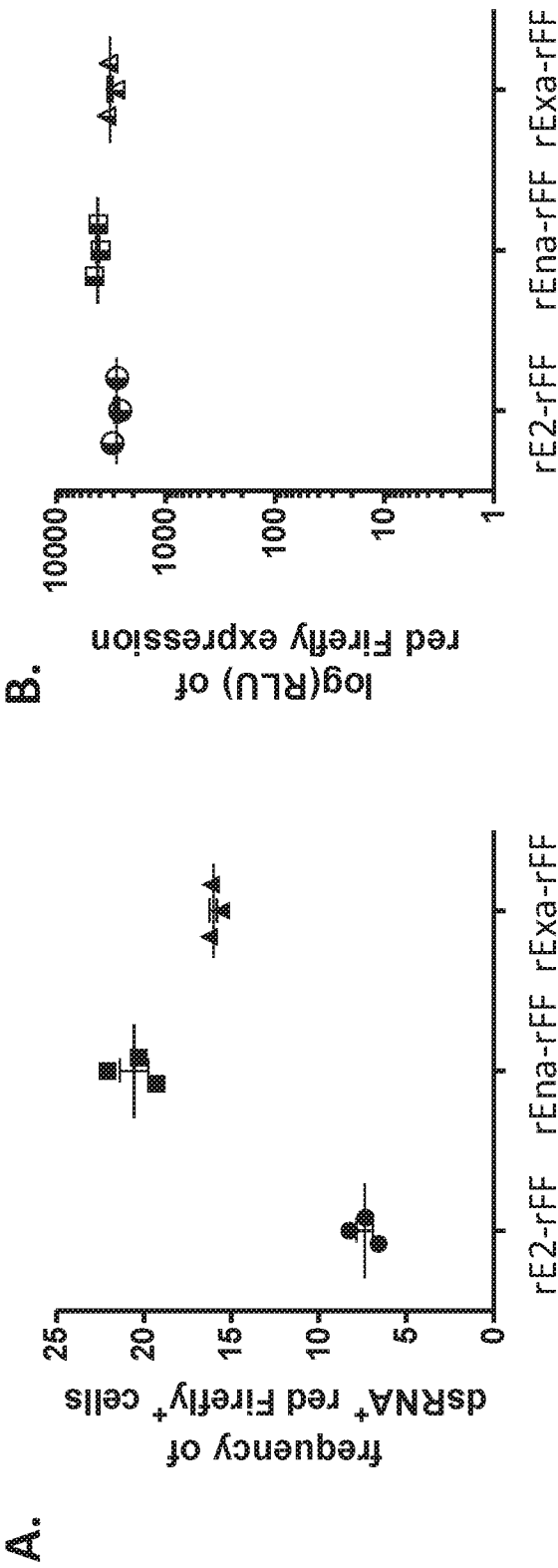


FIG. 6A

FIG. 6B

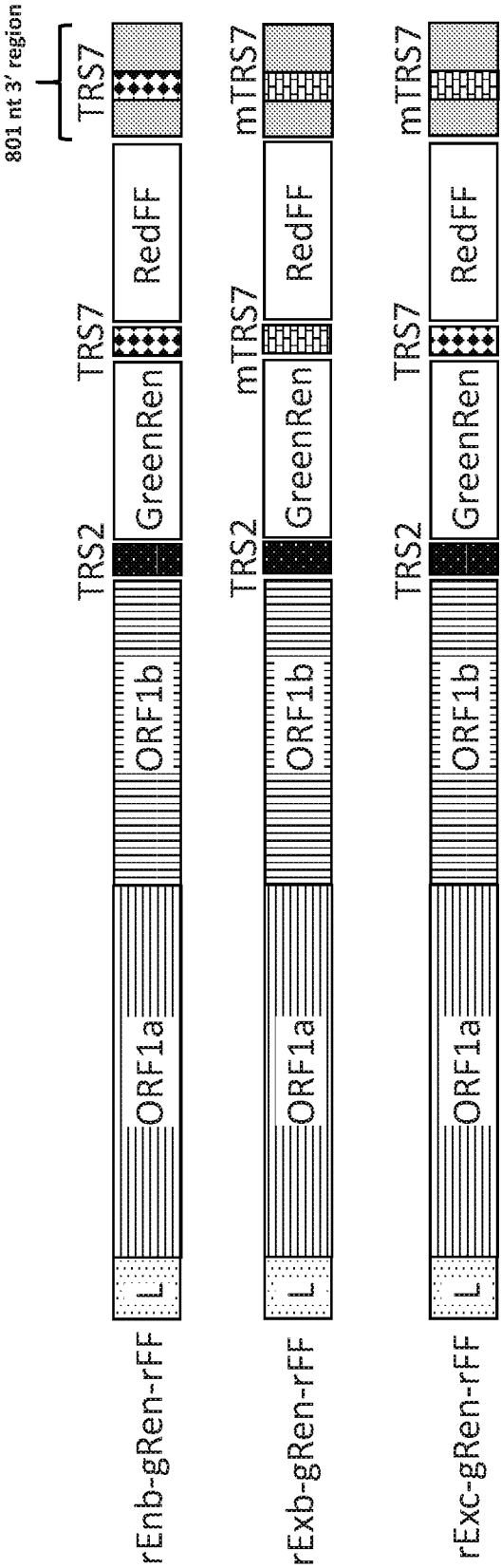
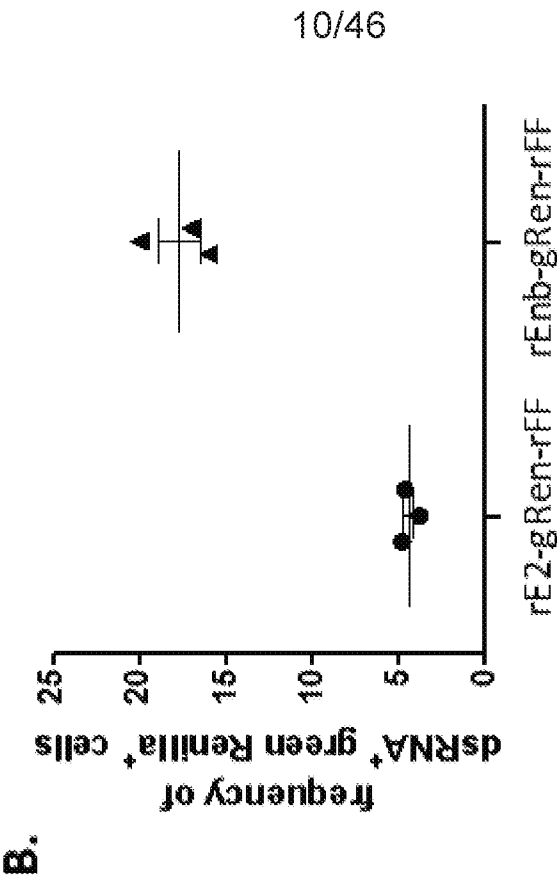
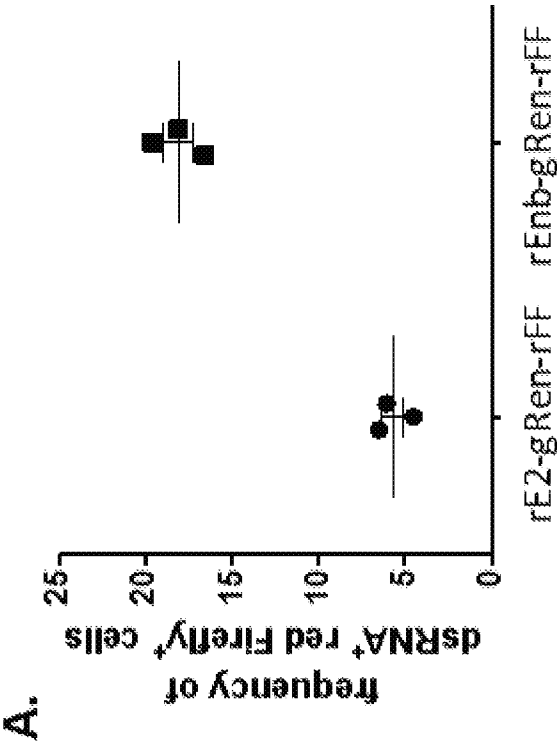


FIG. 7



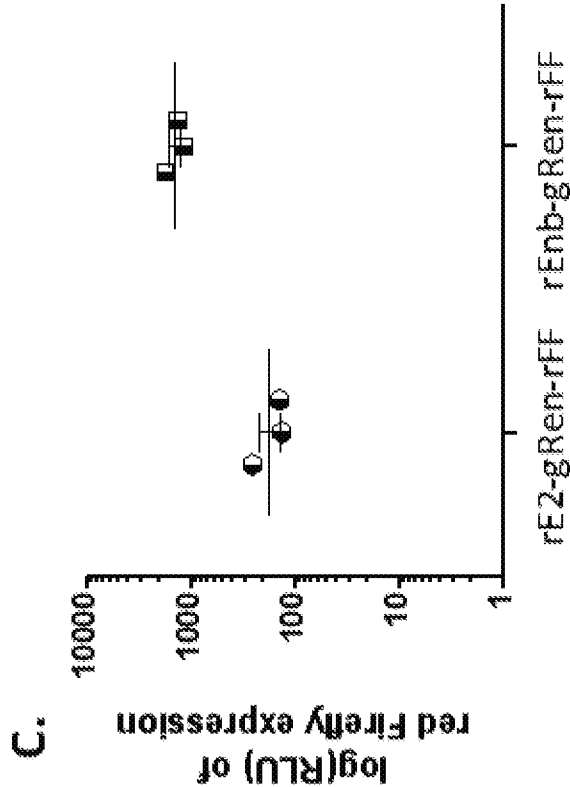
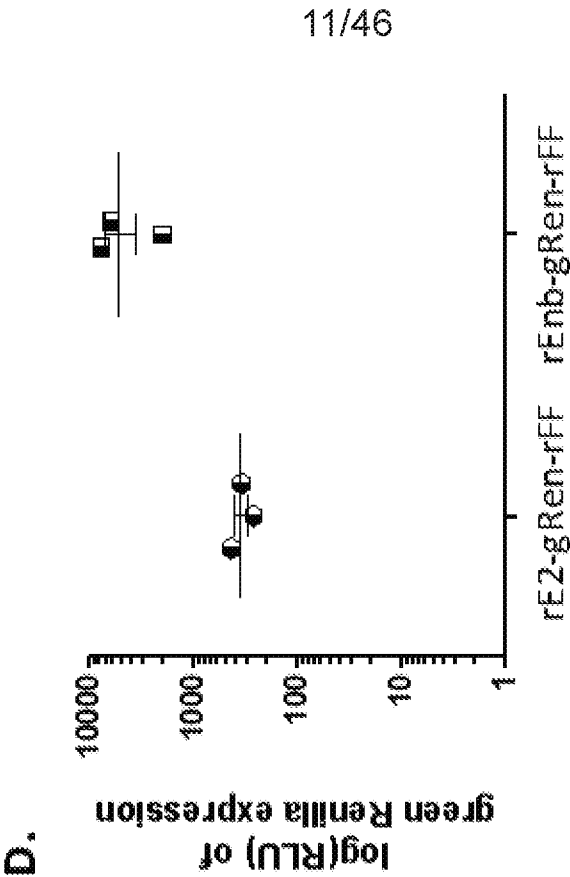


FIG. 8C



11/46

FIG. 8D

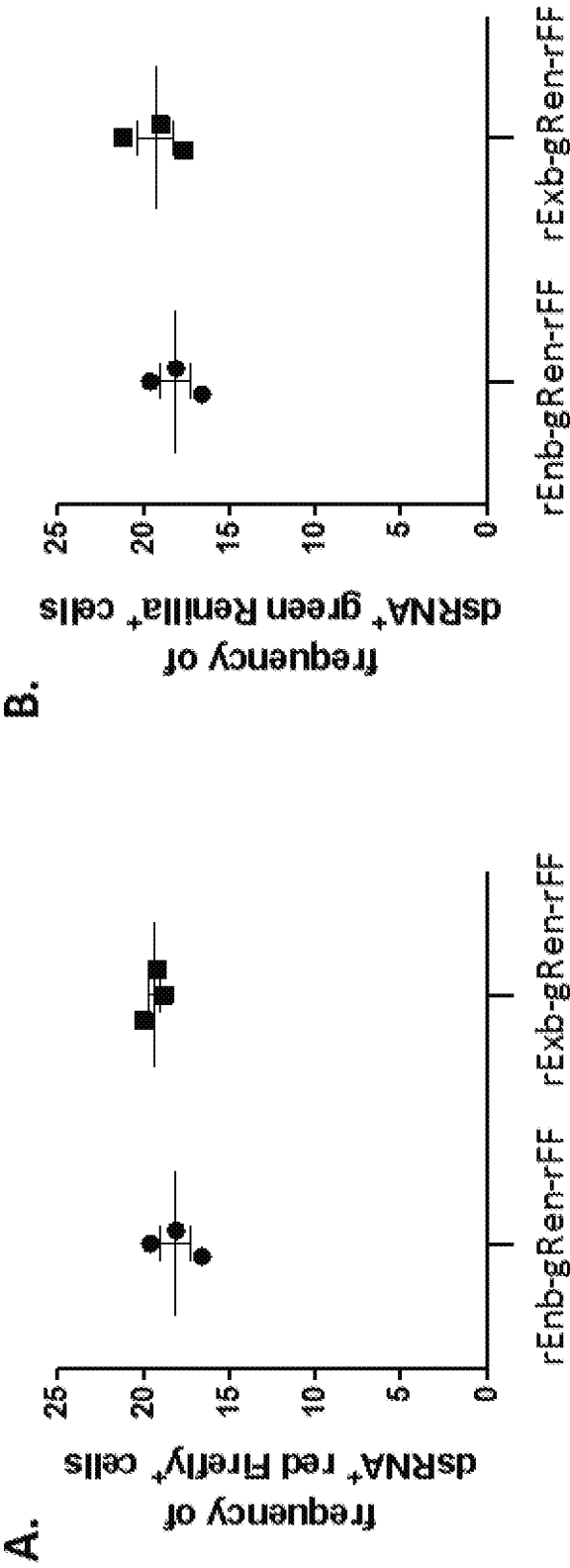


FIG. 9B

FIG. 9A

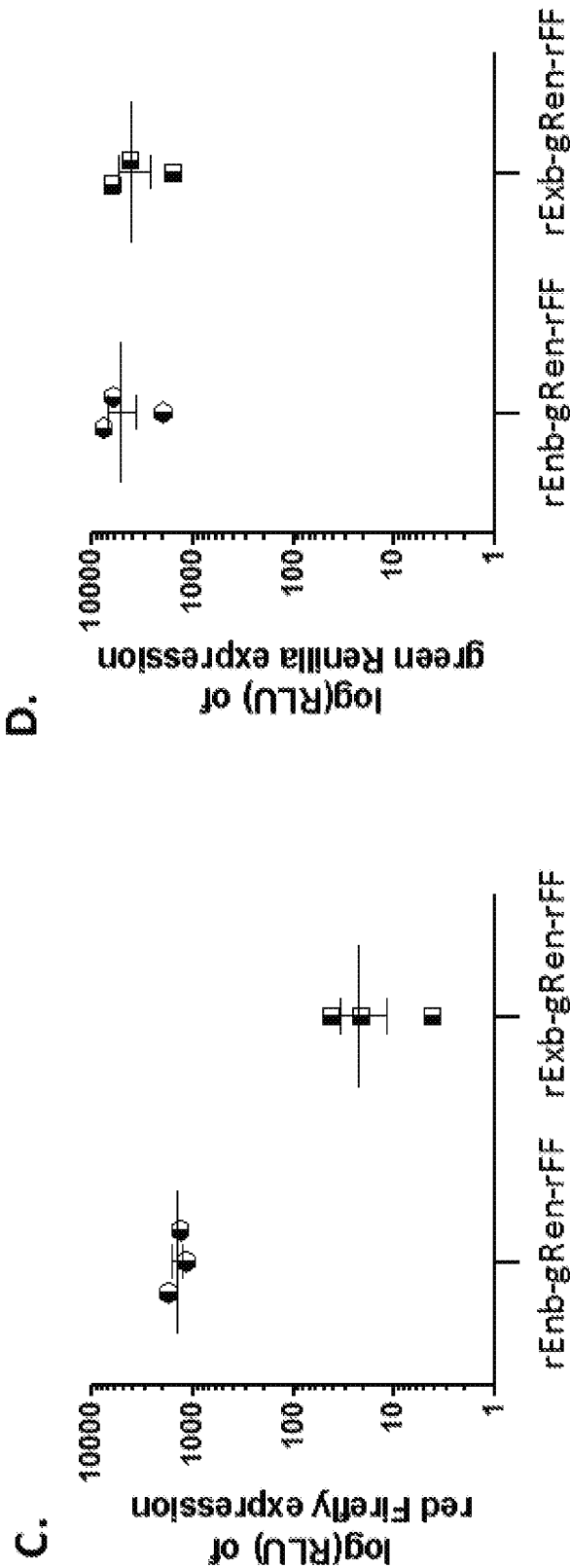


FIG. 9D

FIG. 9C

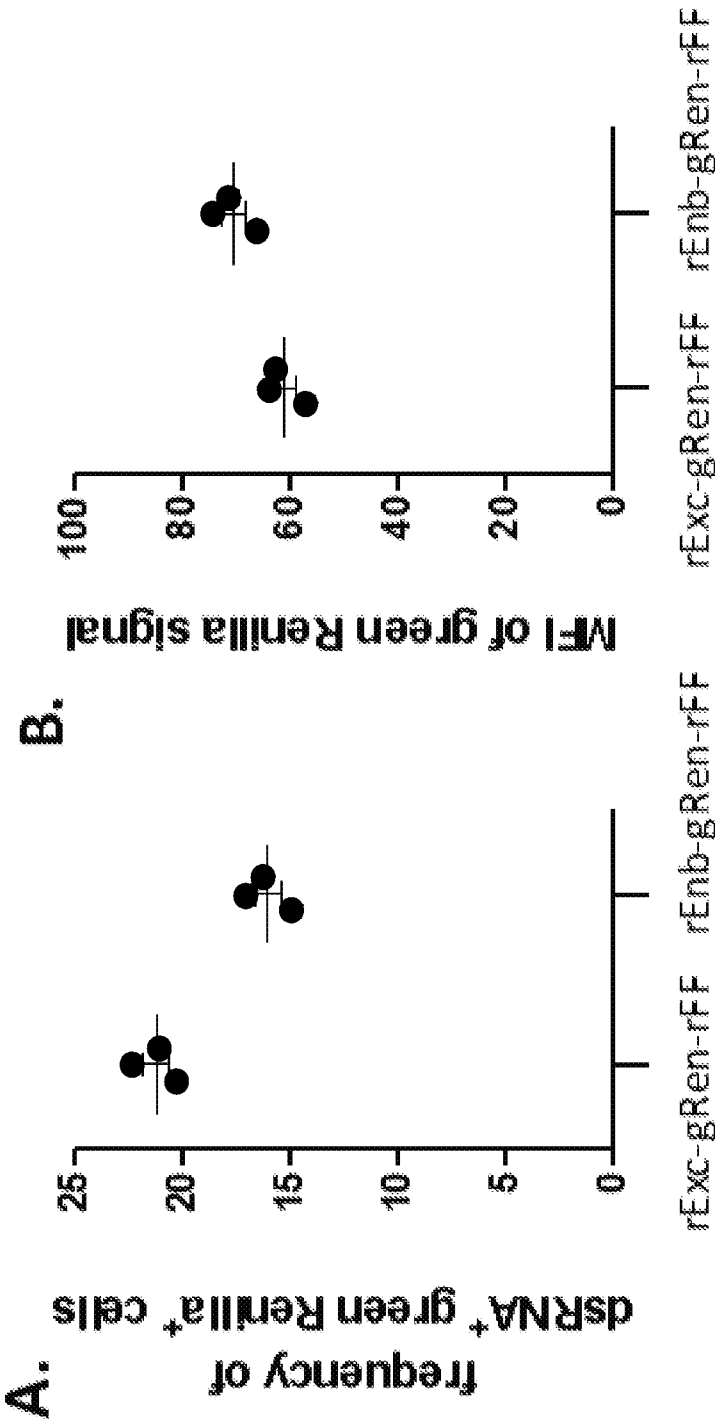


FIG. 10A

FIG. 10B

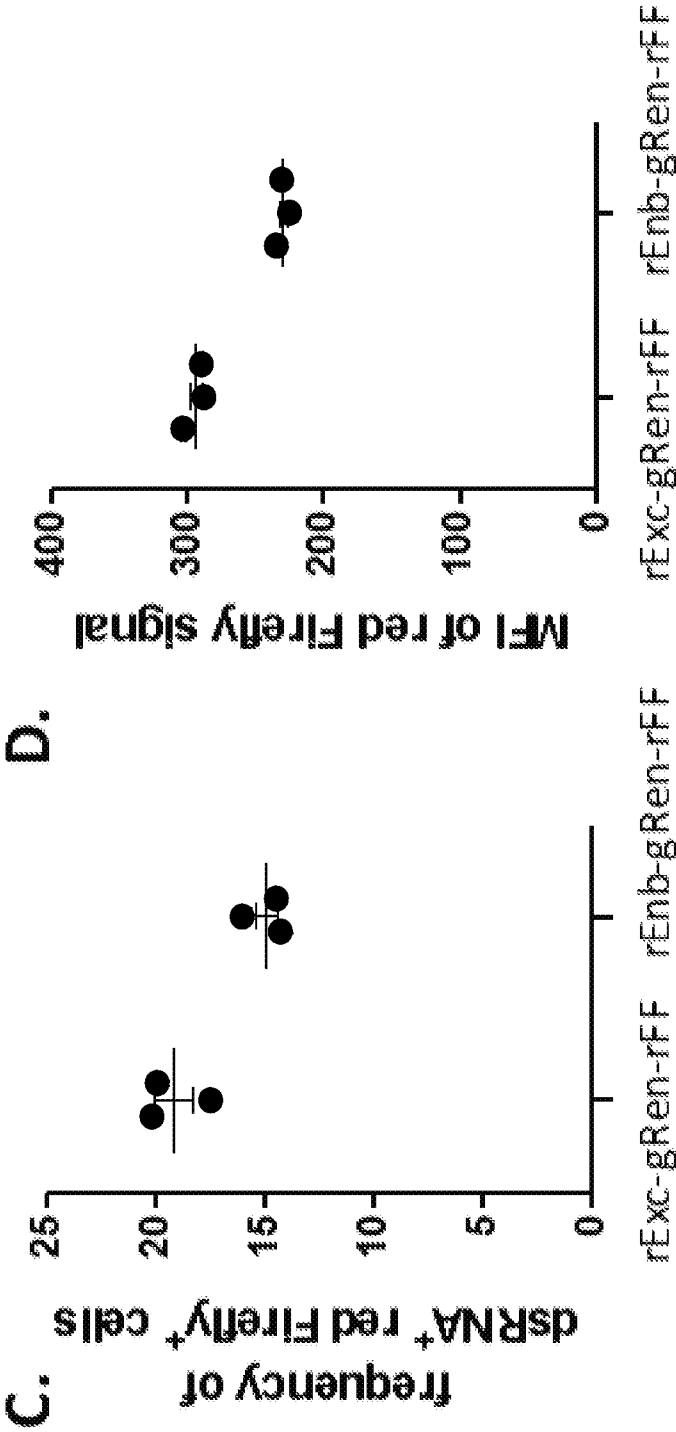


FIG. 10D

FIG. 10C

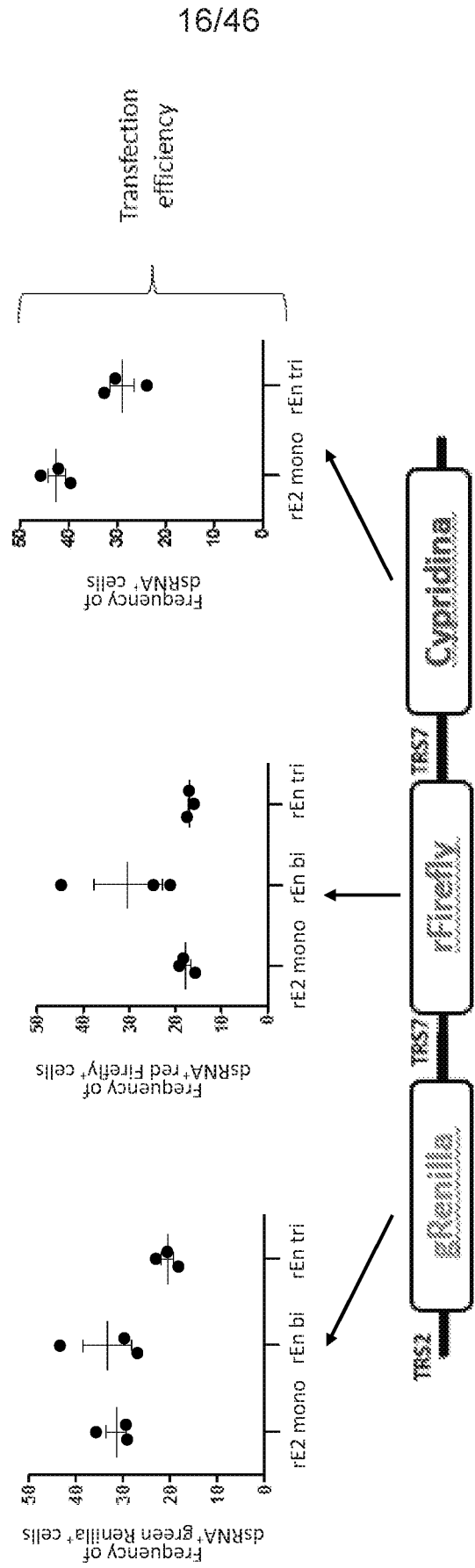


FIG. 11A

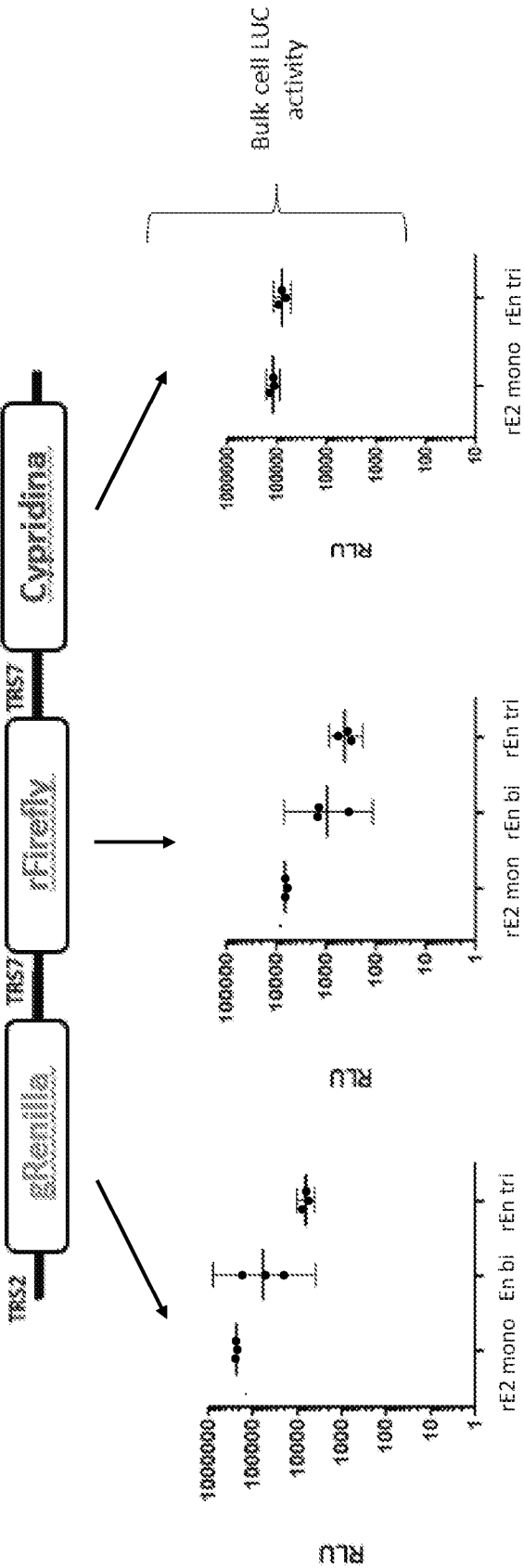
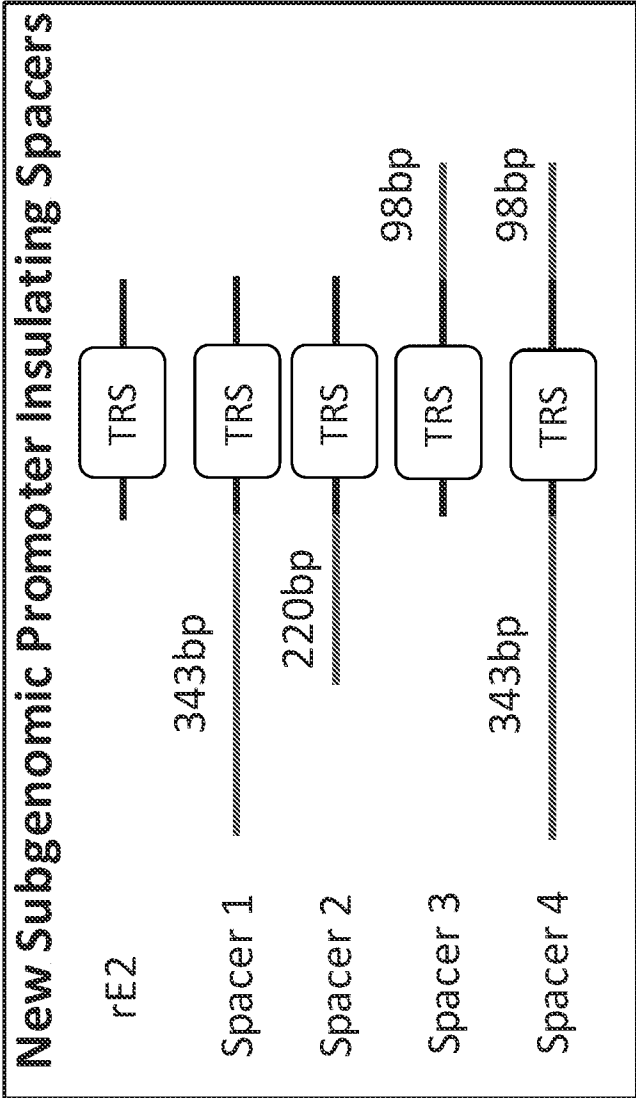


FIG. 11B



B

FIG. 12B

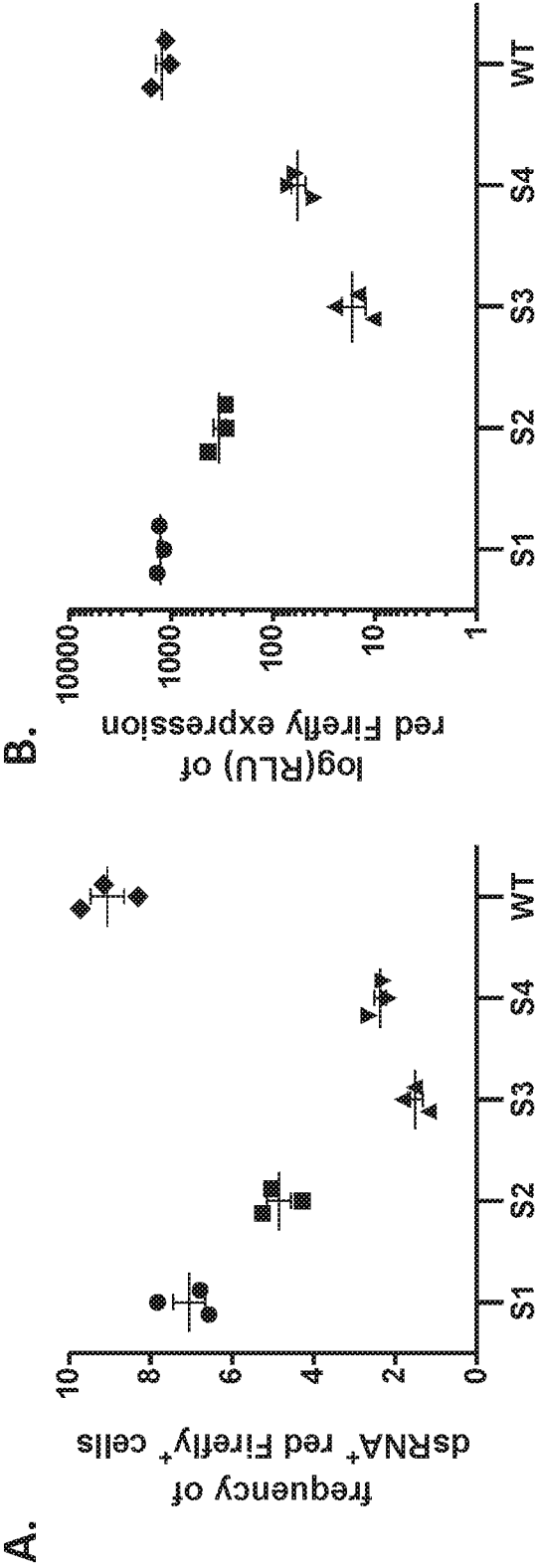


FIG. 13B

FIG. 13A

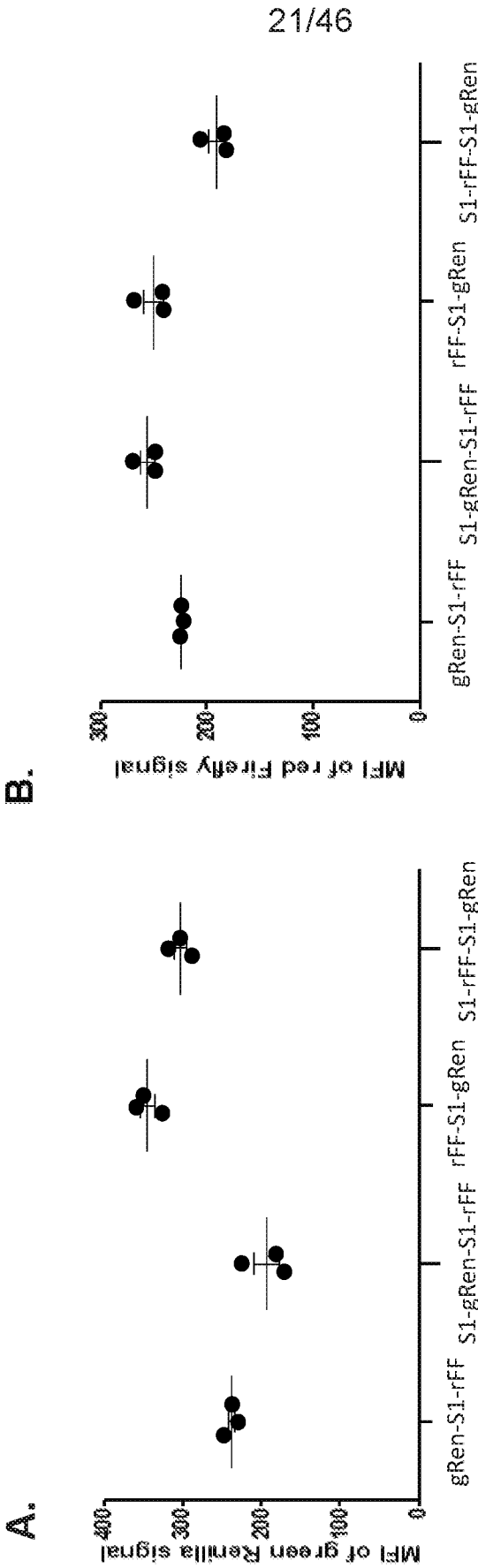


FIG. 14A

FIG. 14B

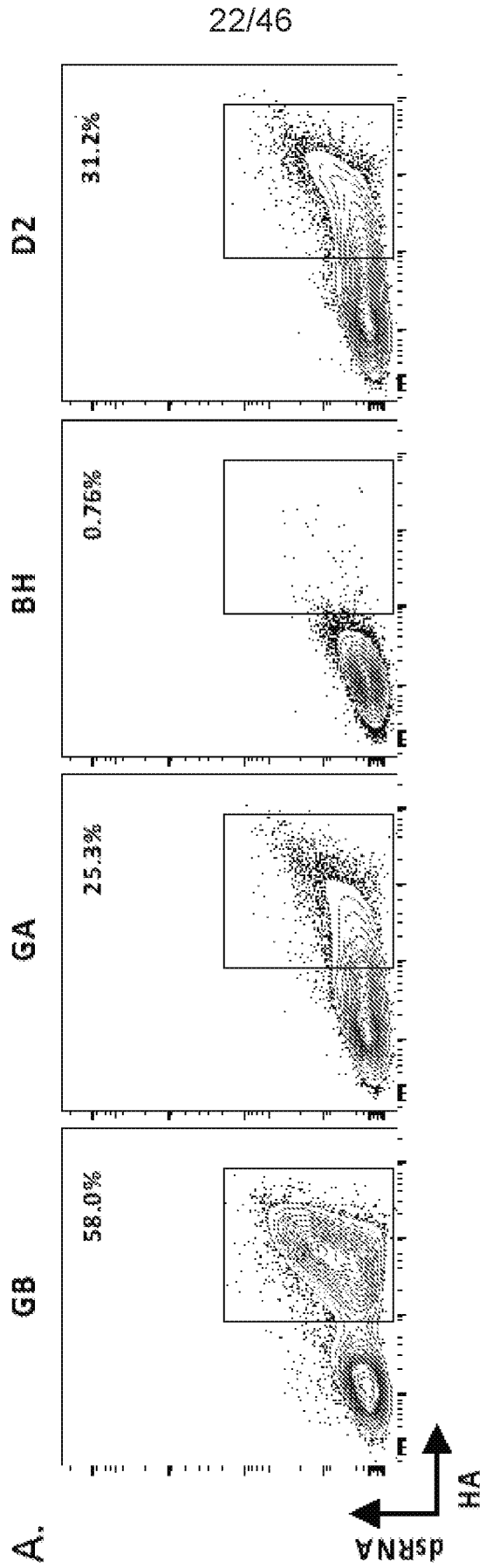


FIG. 15A

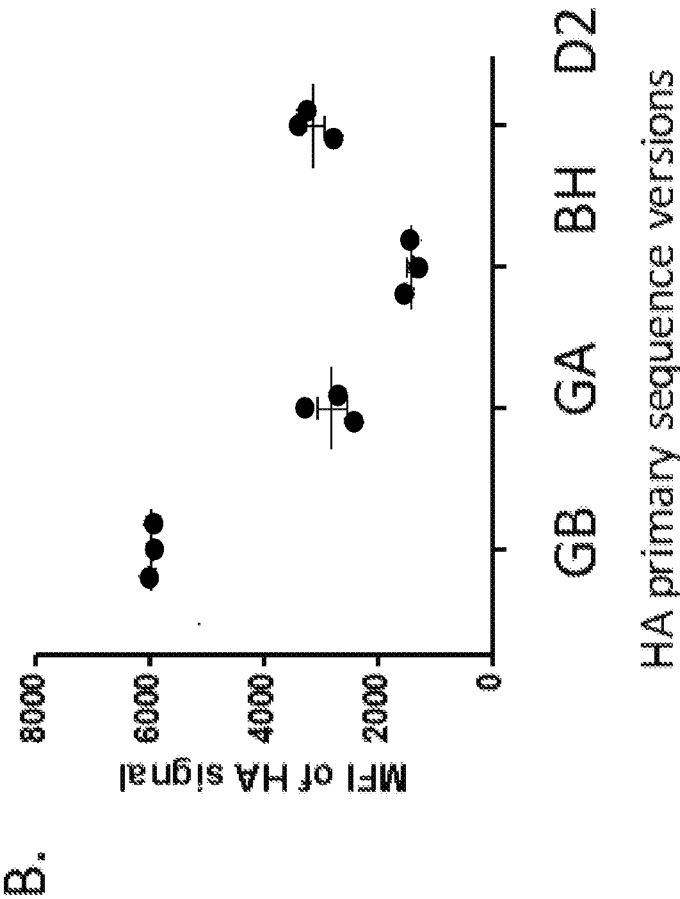


FIG. 15B

24/46

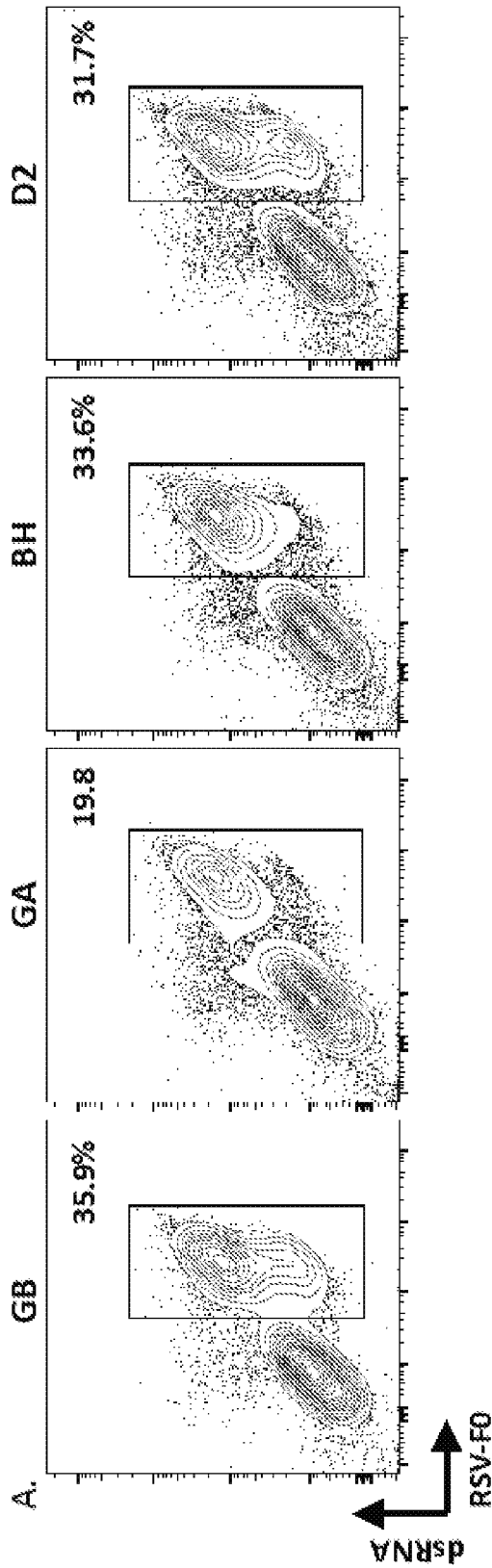


FIG. 16A

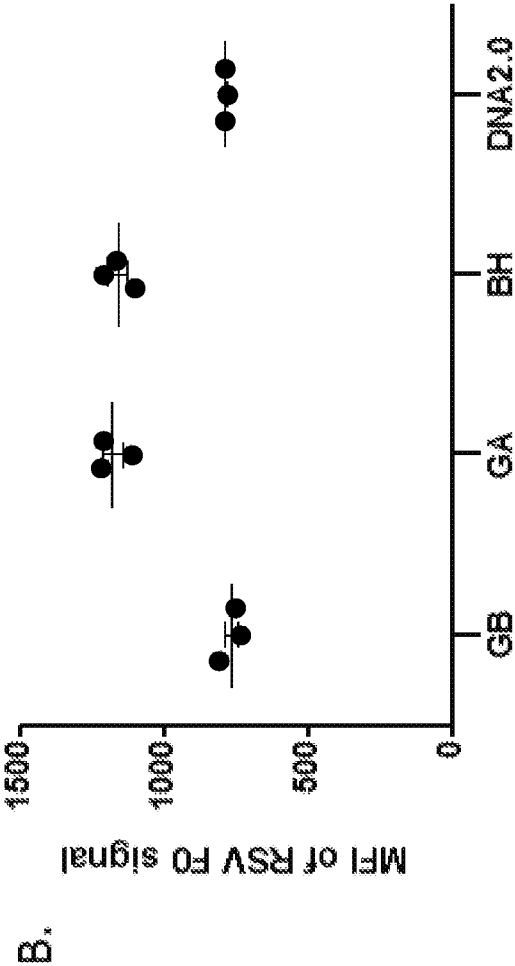


FIG. 16B

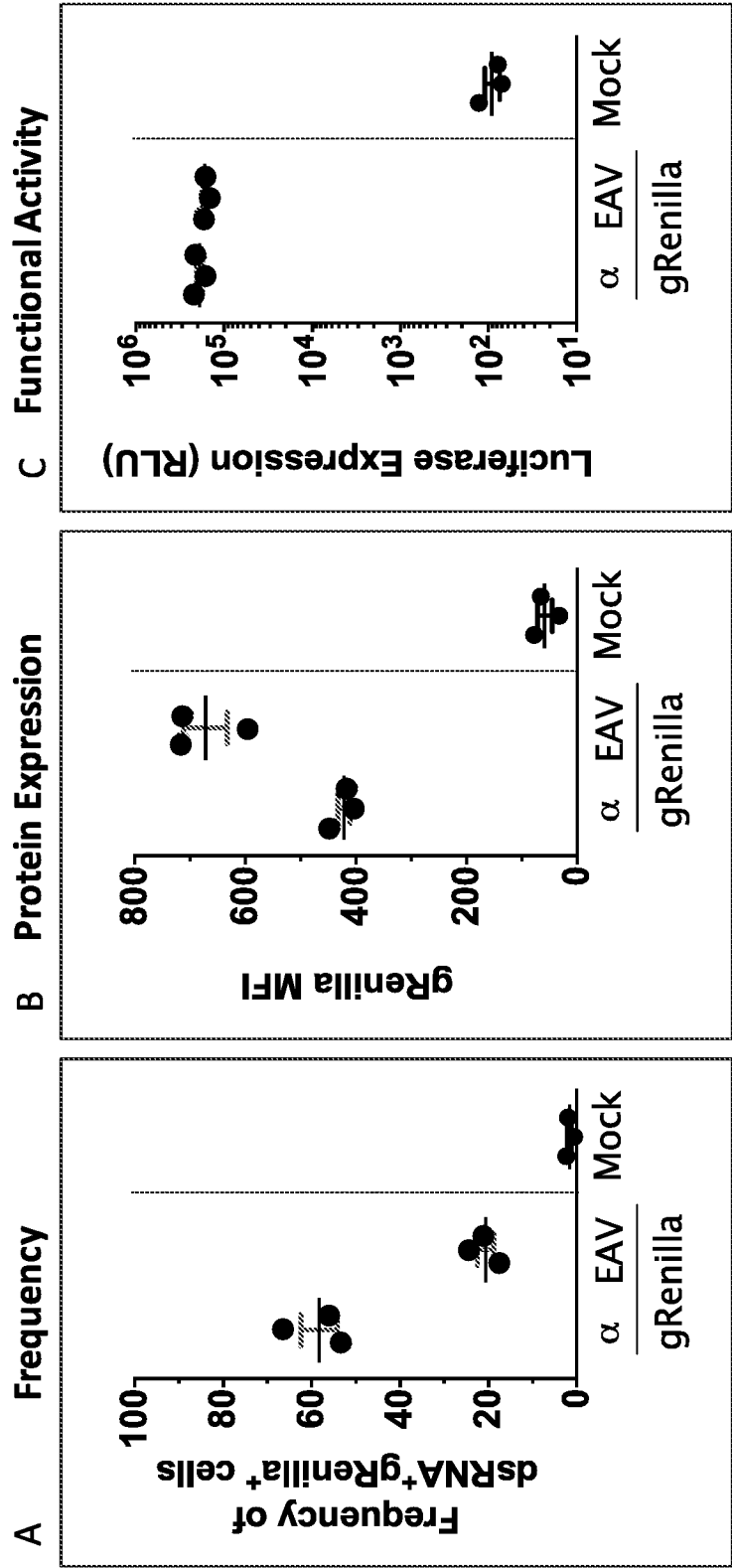


FIG. 17A

FIG. 17B

FIG. 17C

27/46

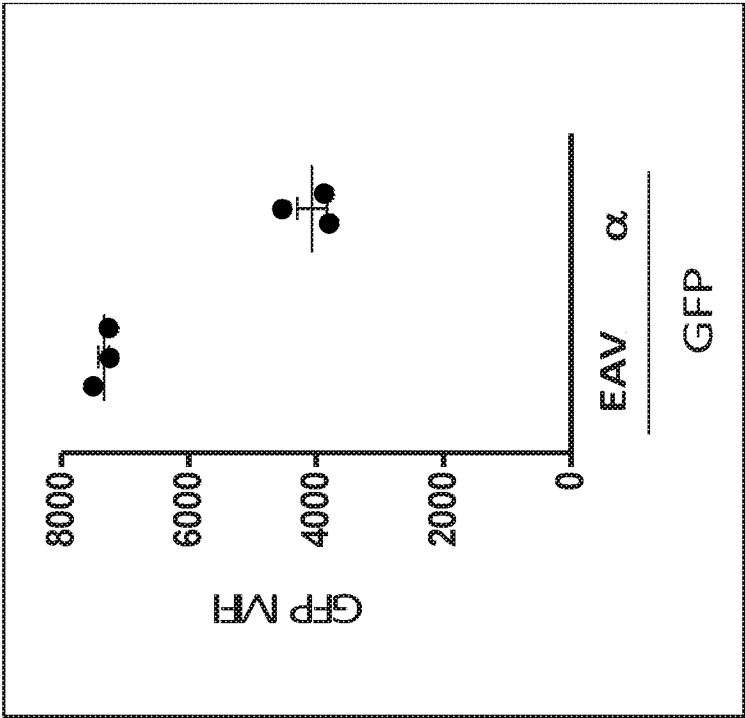
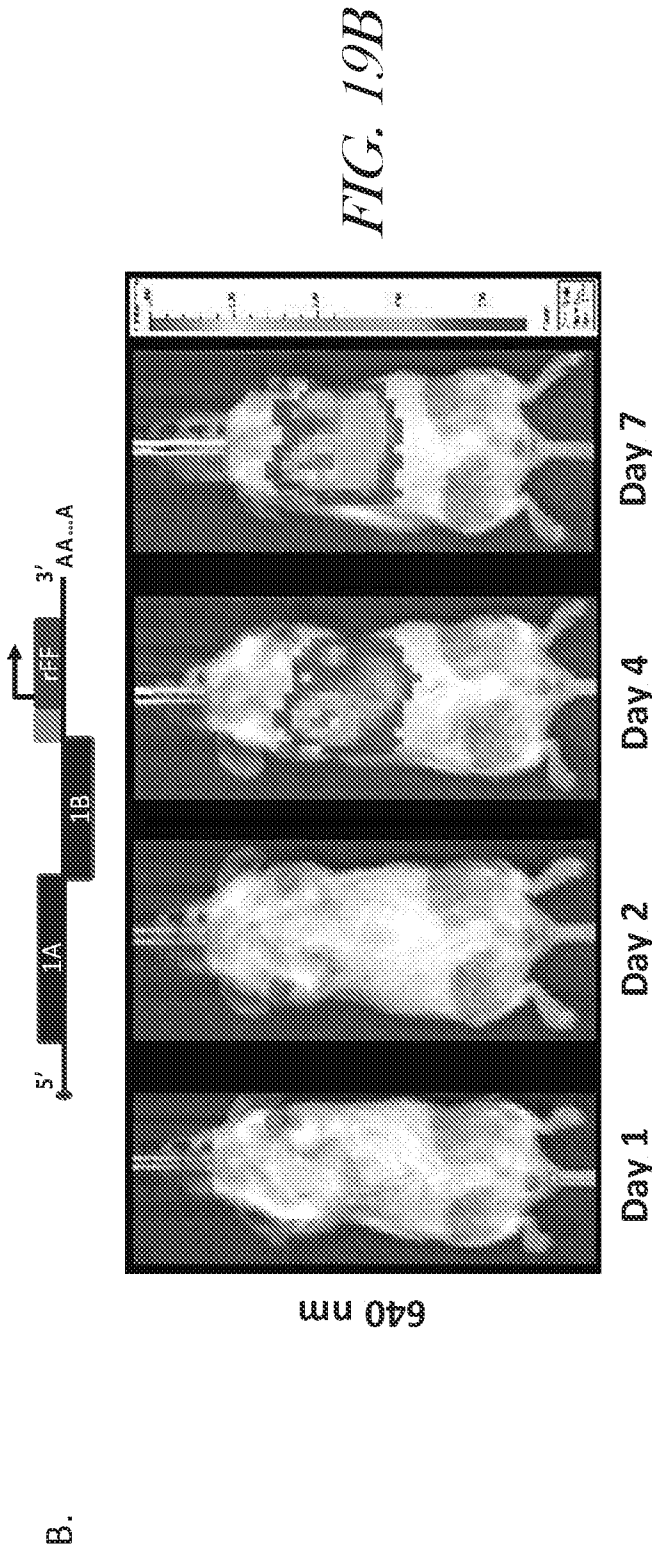
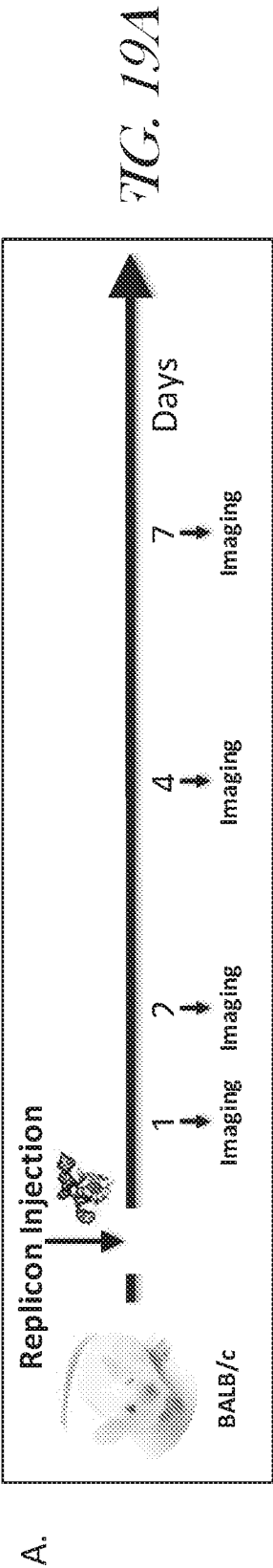


FIG. 18



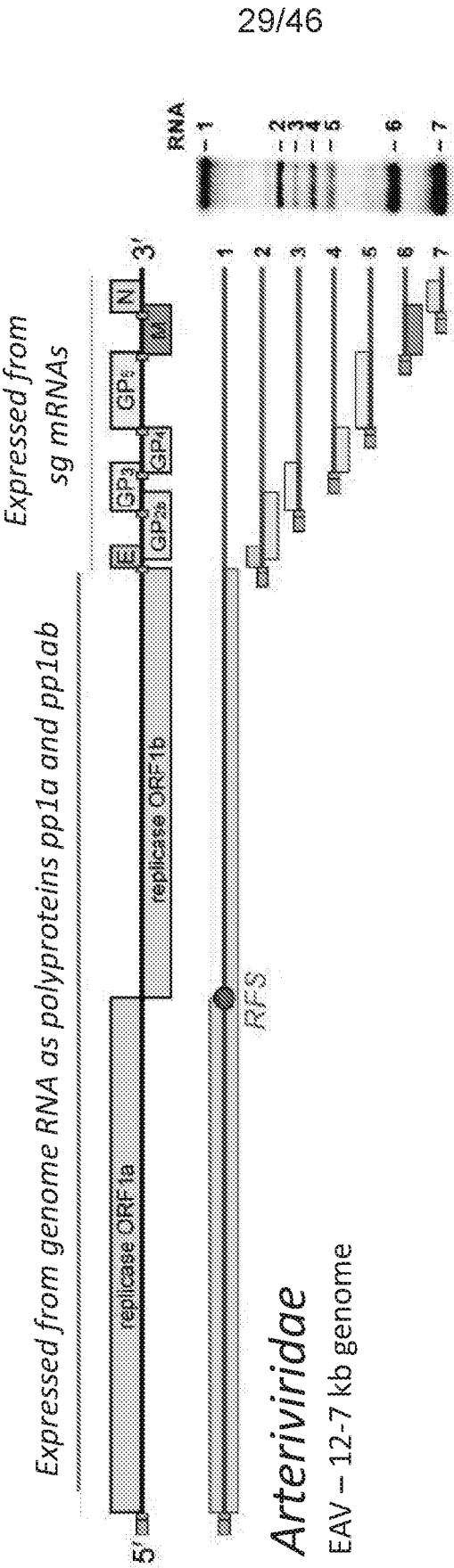


FIG. 20

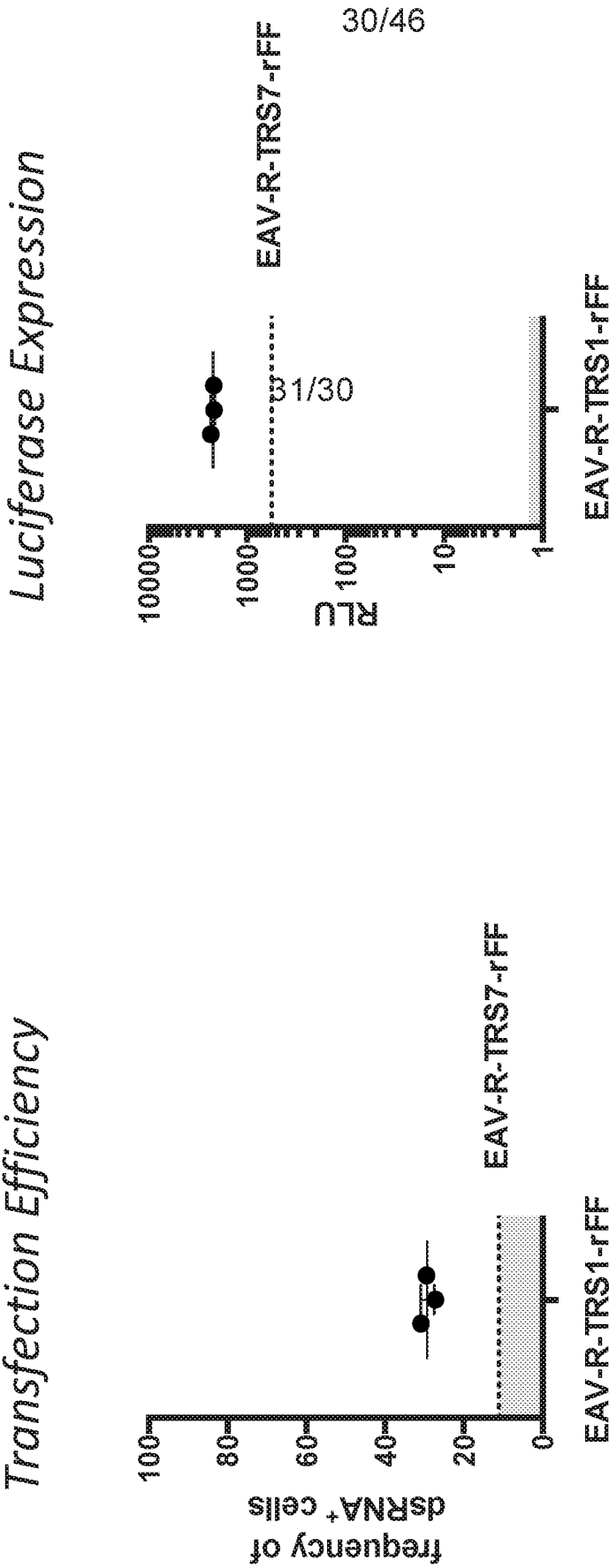


FIG. 21

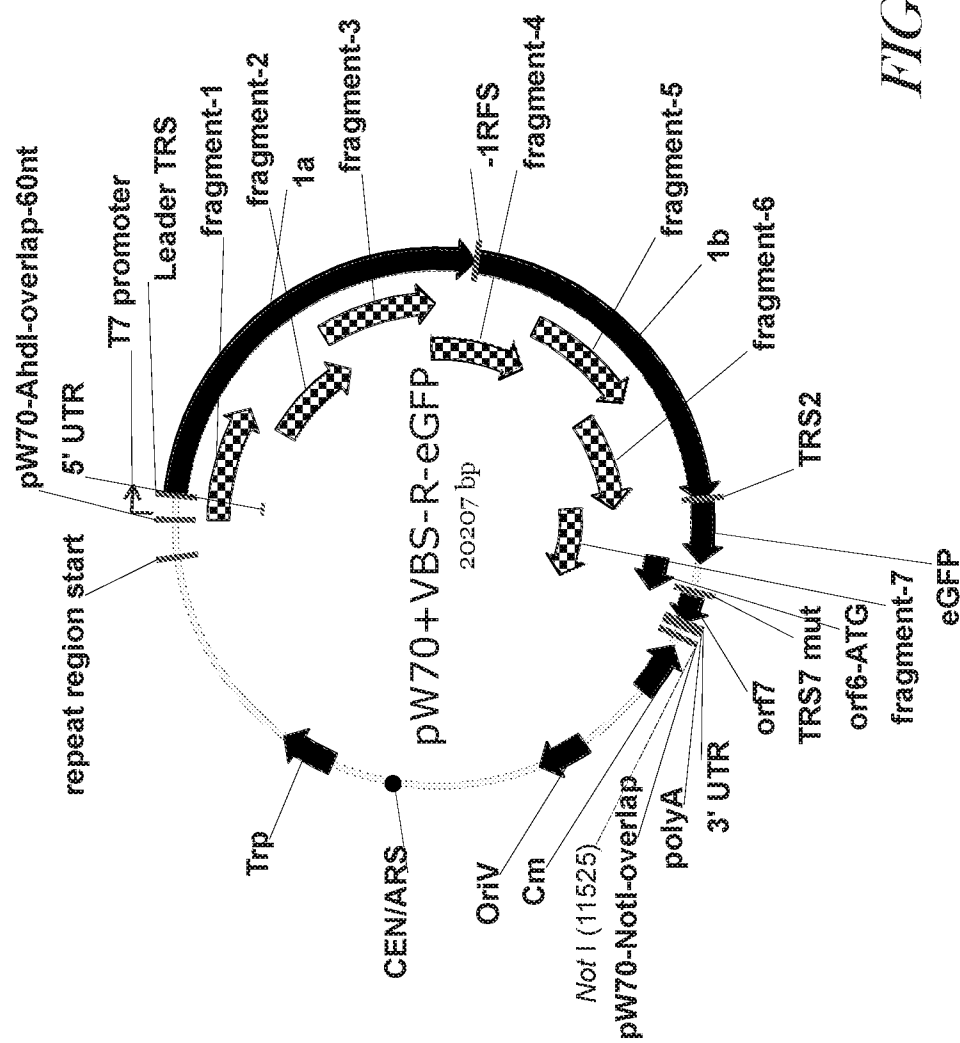


FIG. 22

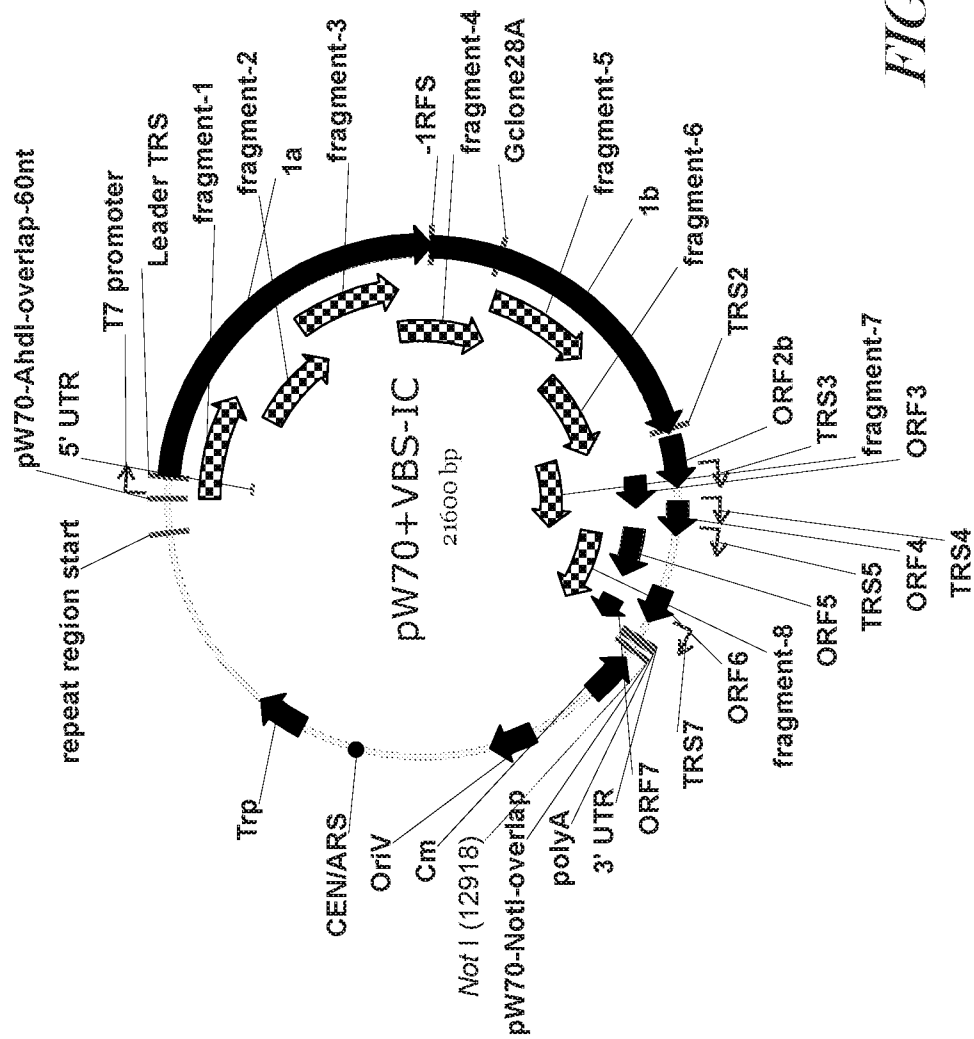


FIG. 23

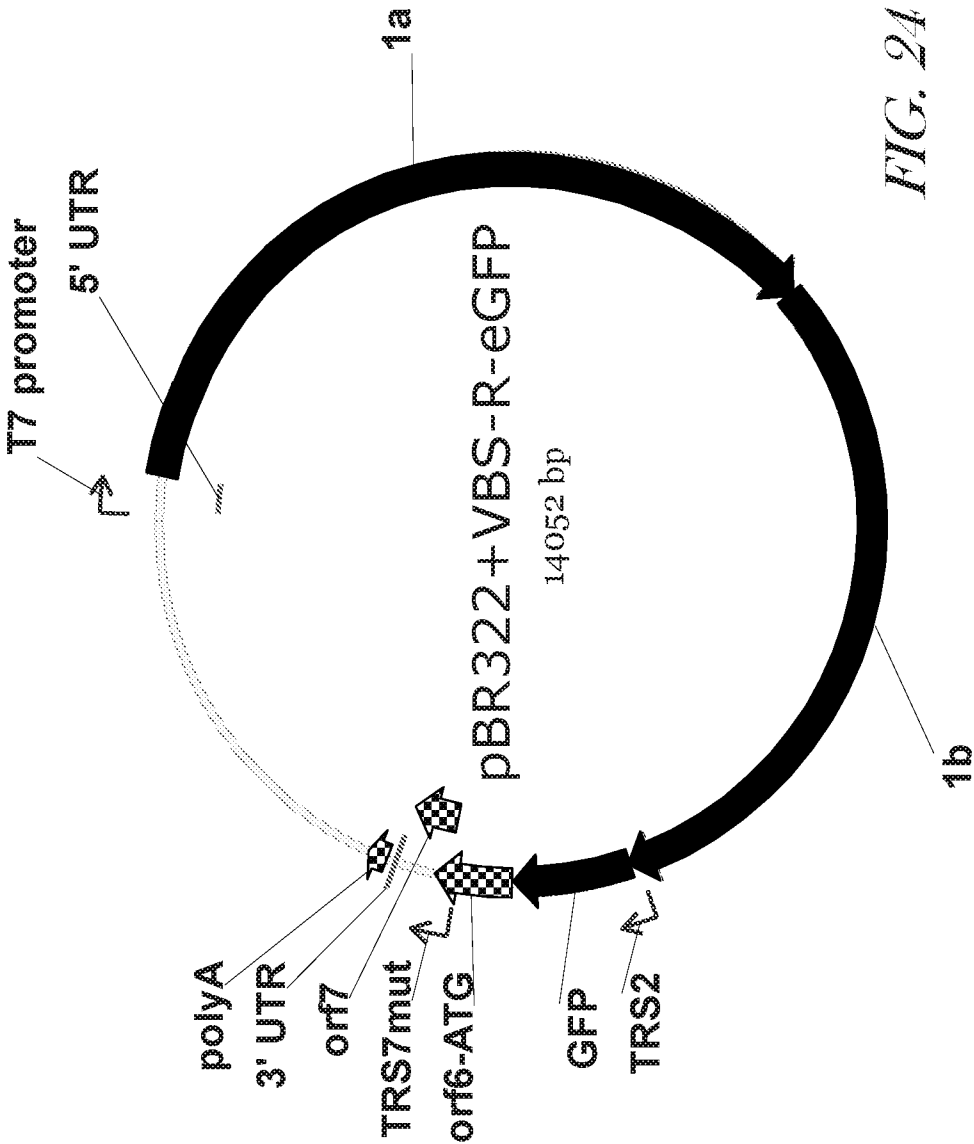


FIG. 24

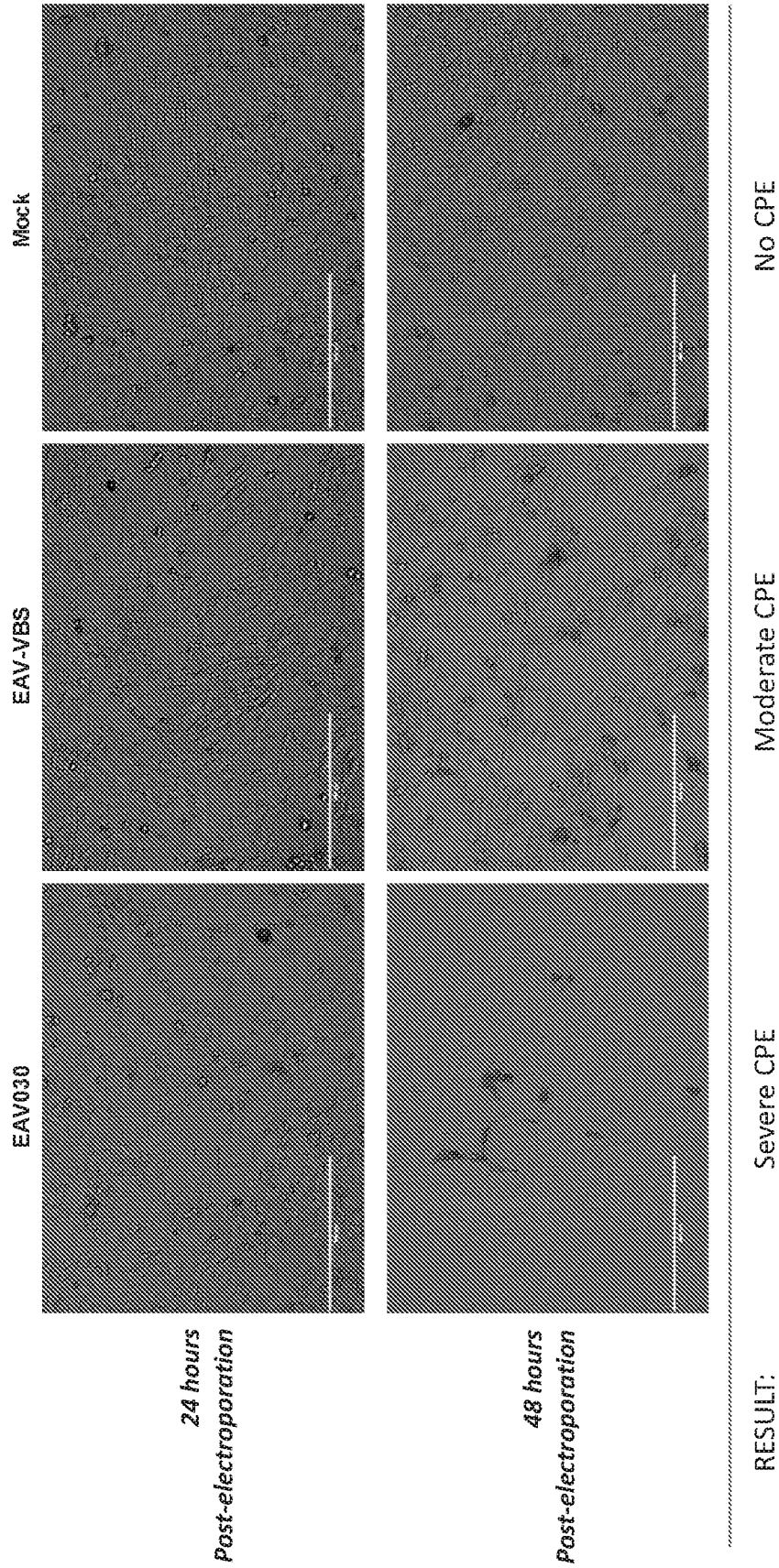


FIG. 25

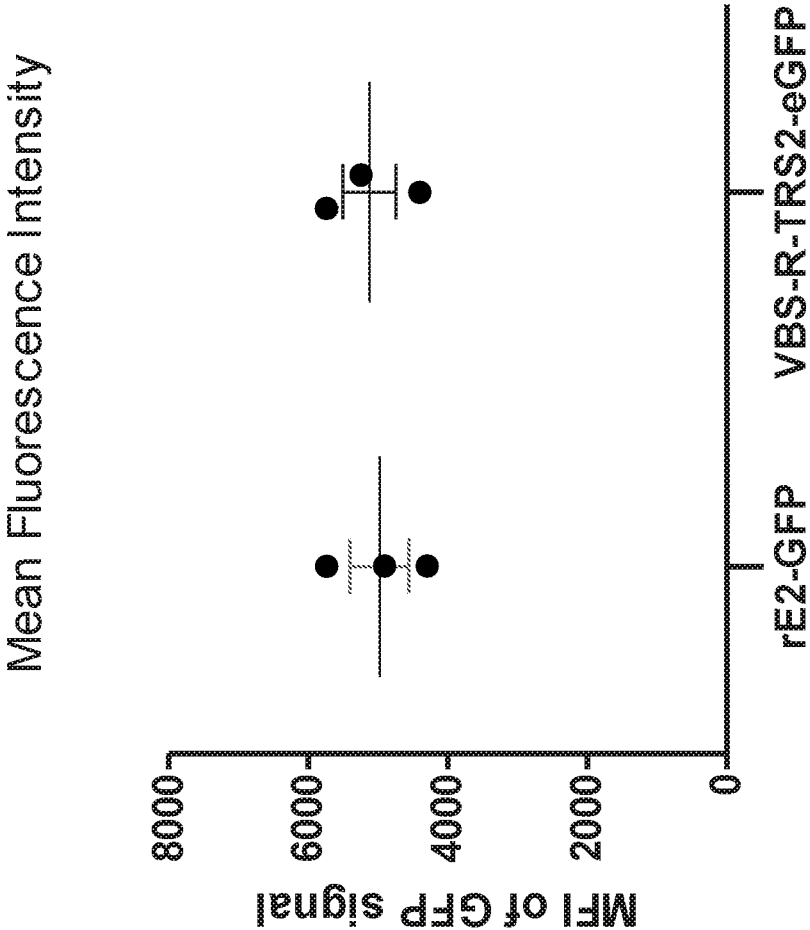


FIG. 26

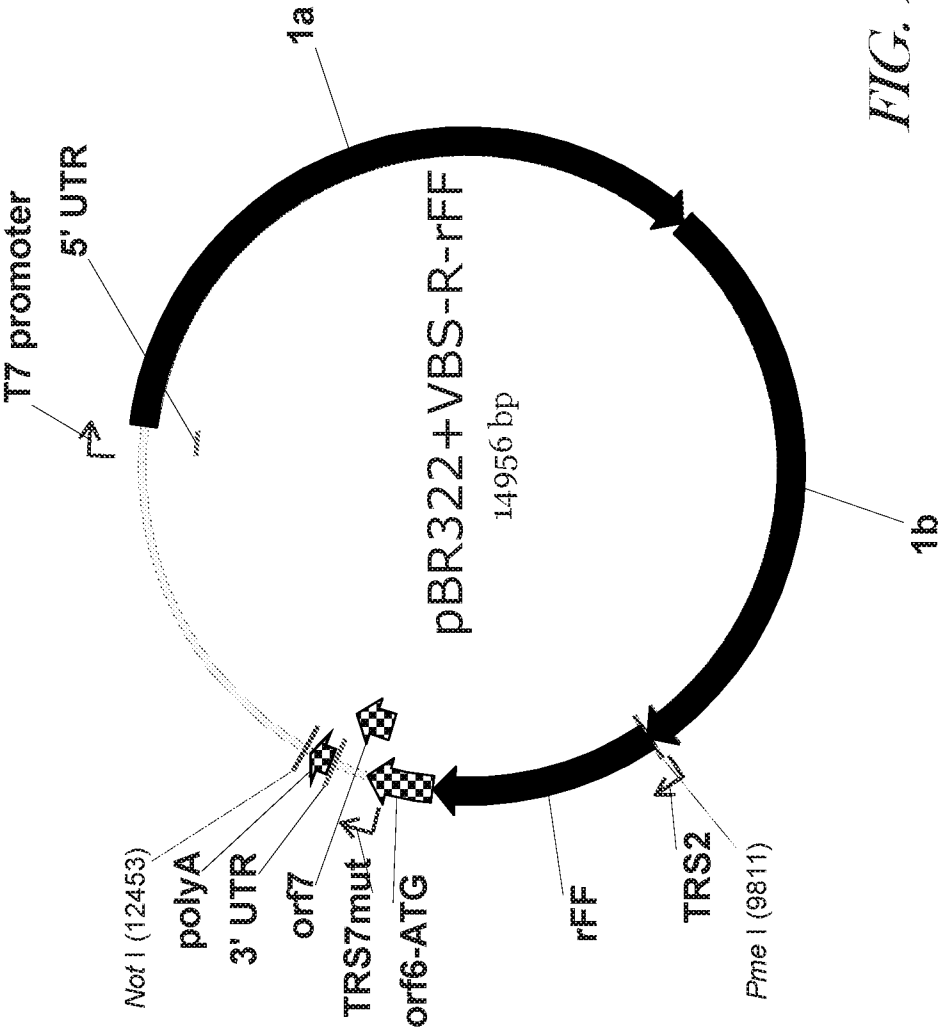


FIG. 27

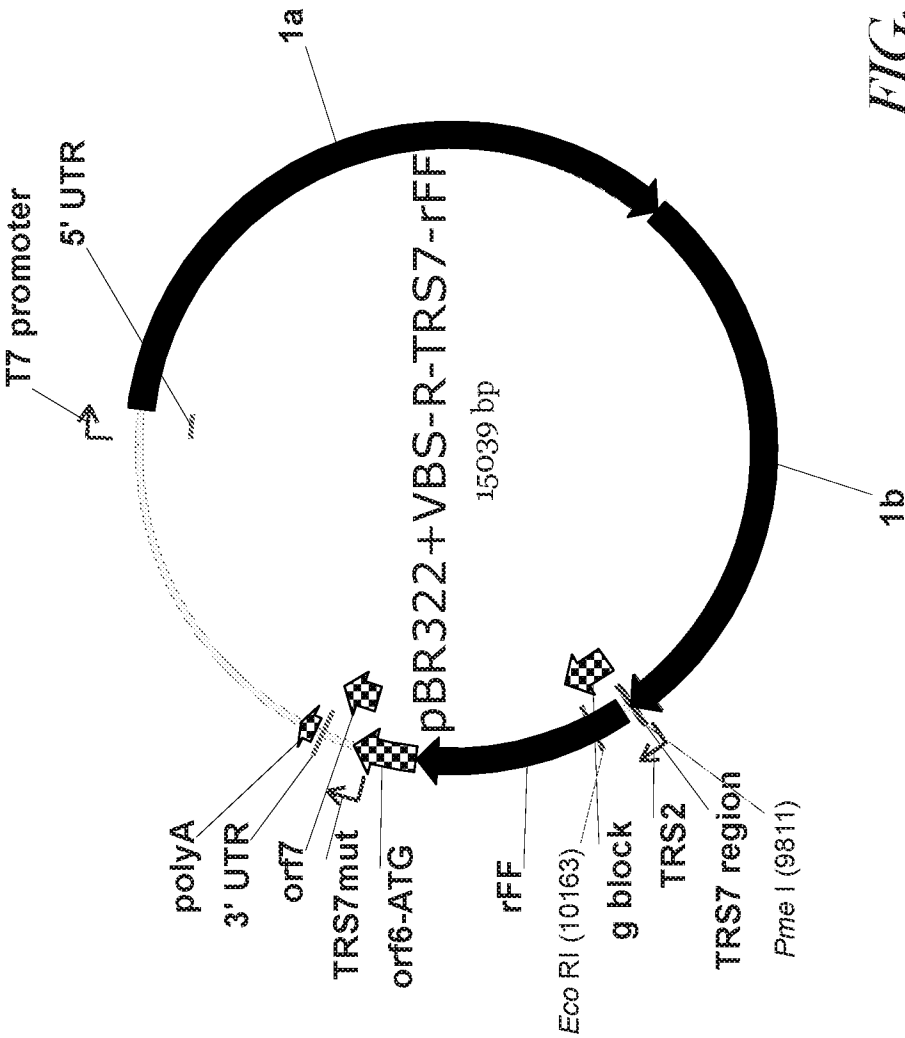


FIG. 28

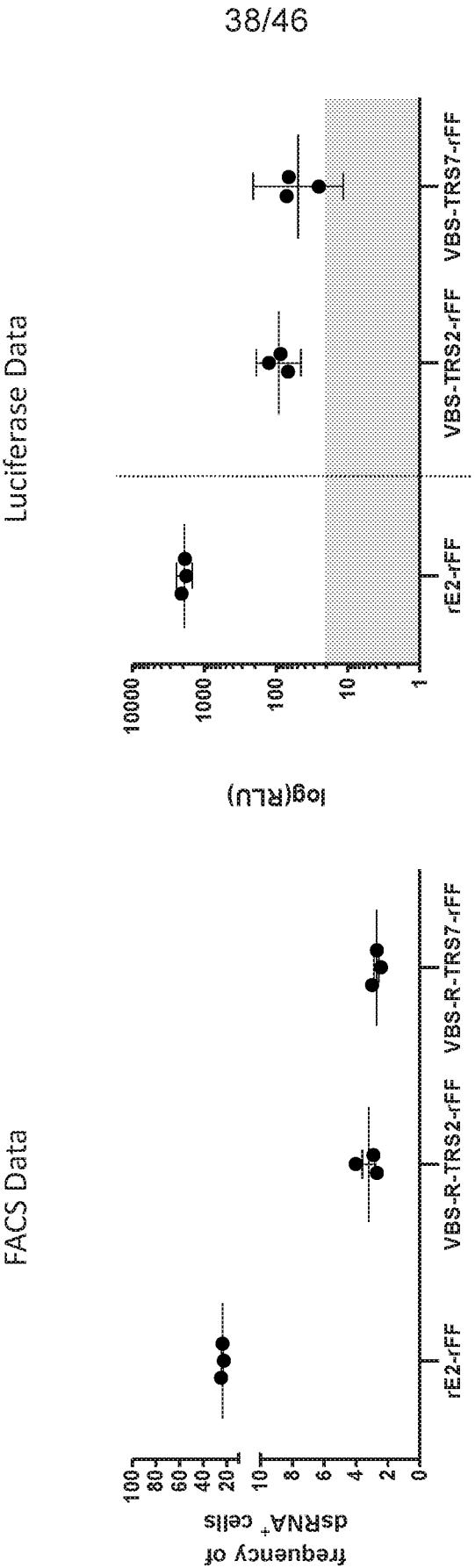


FIG. 29

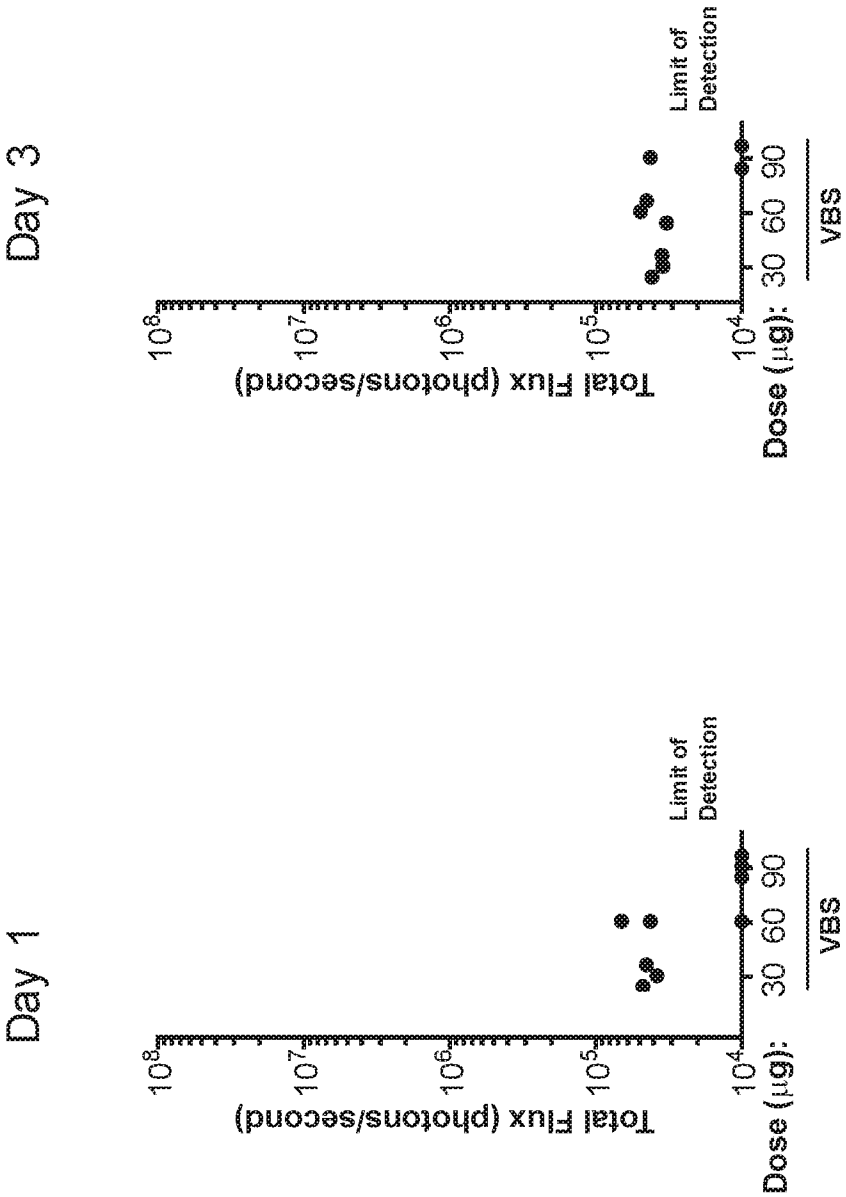


FIG. 30

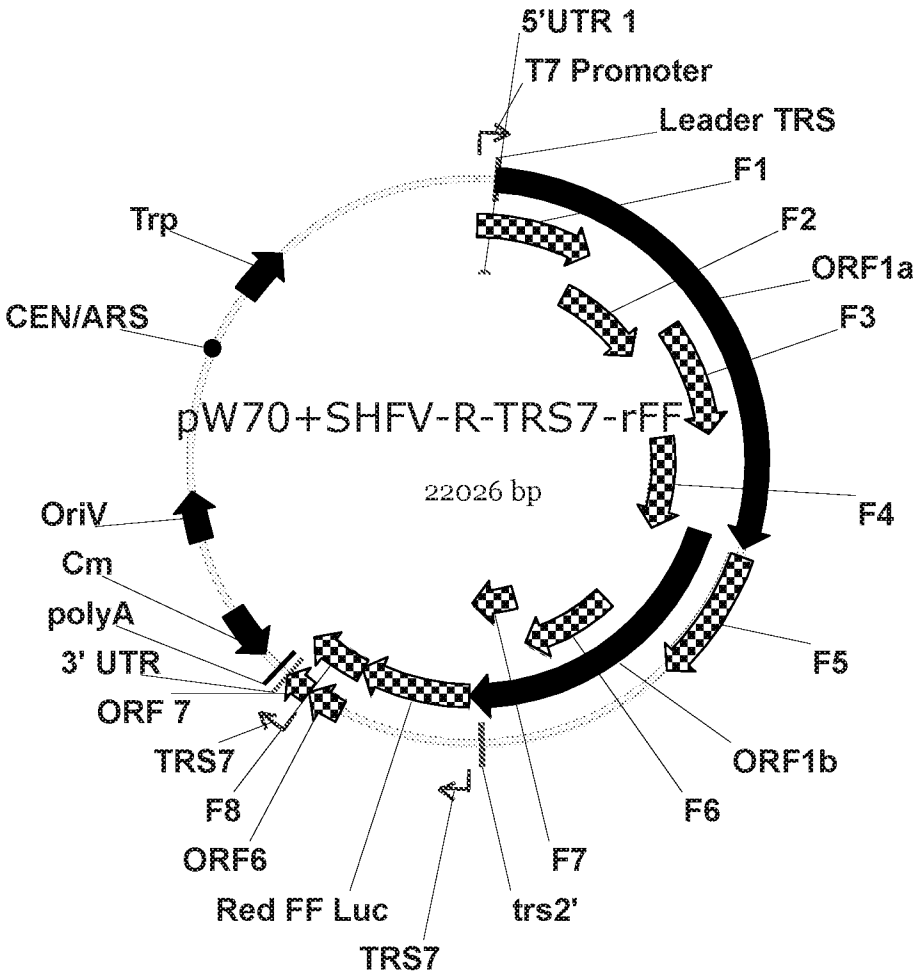


FIG. 31

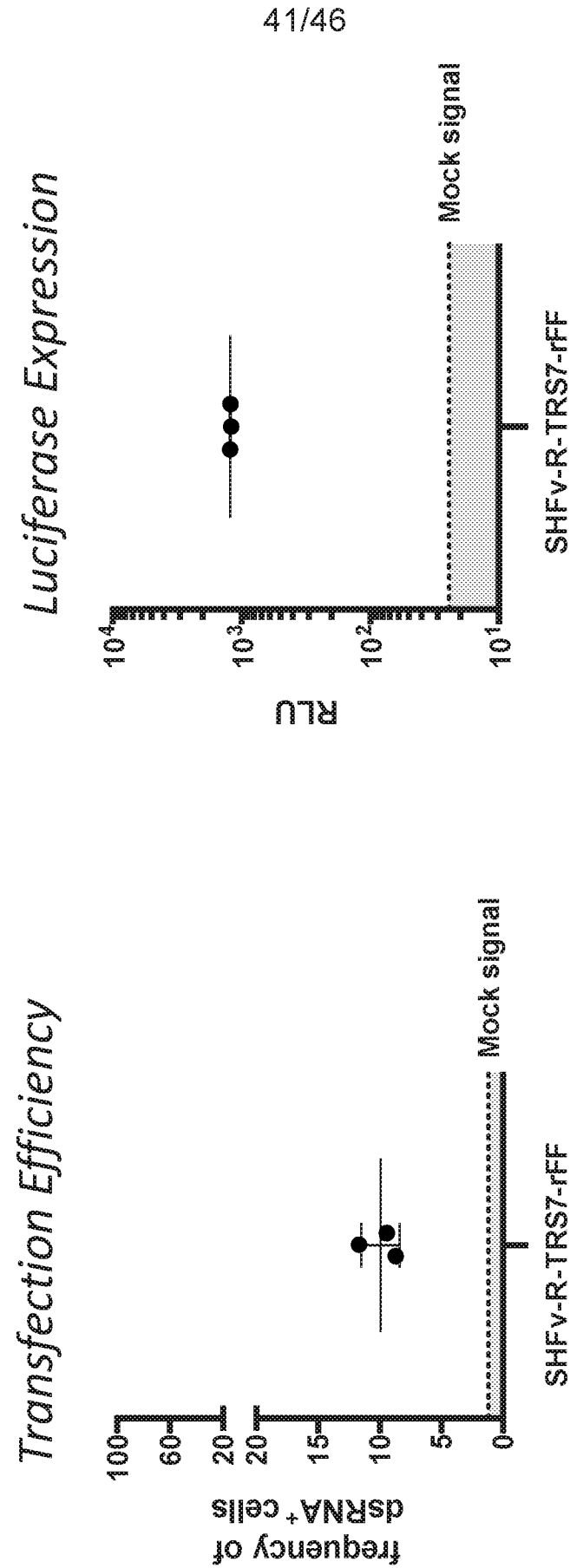
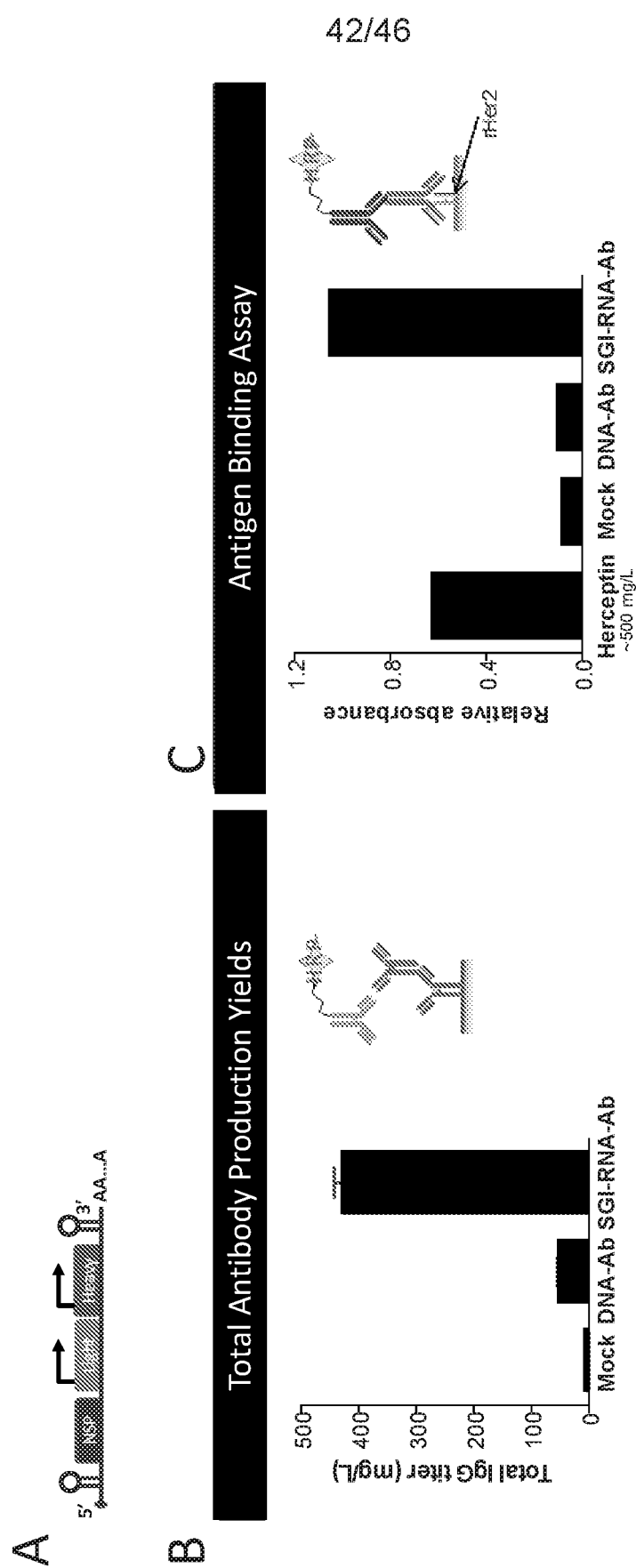


FIG. 32



FIGS. 33A-C

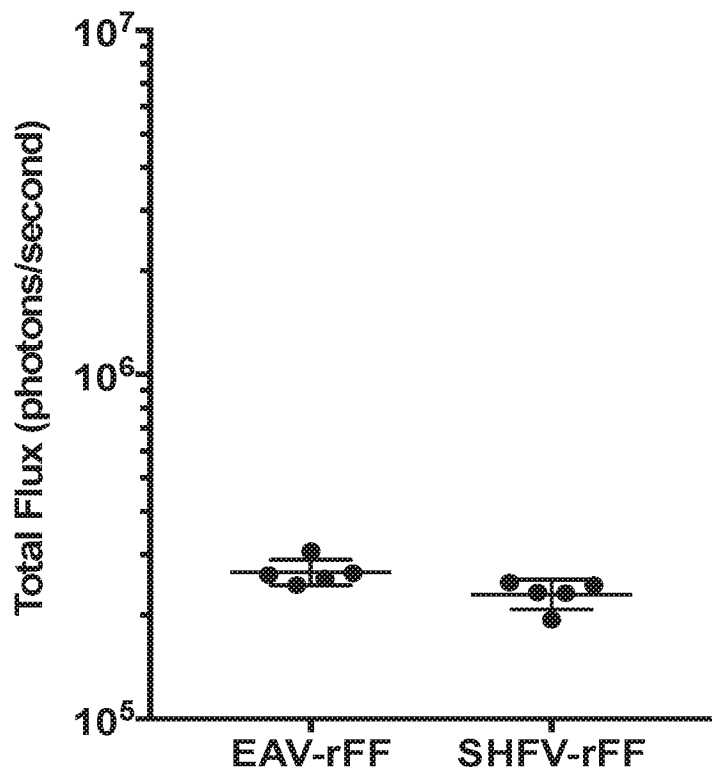
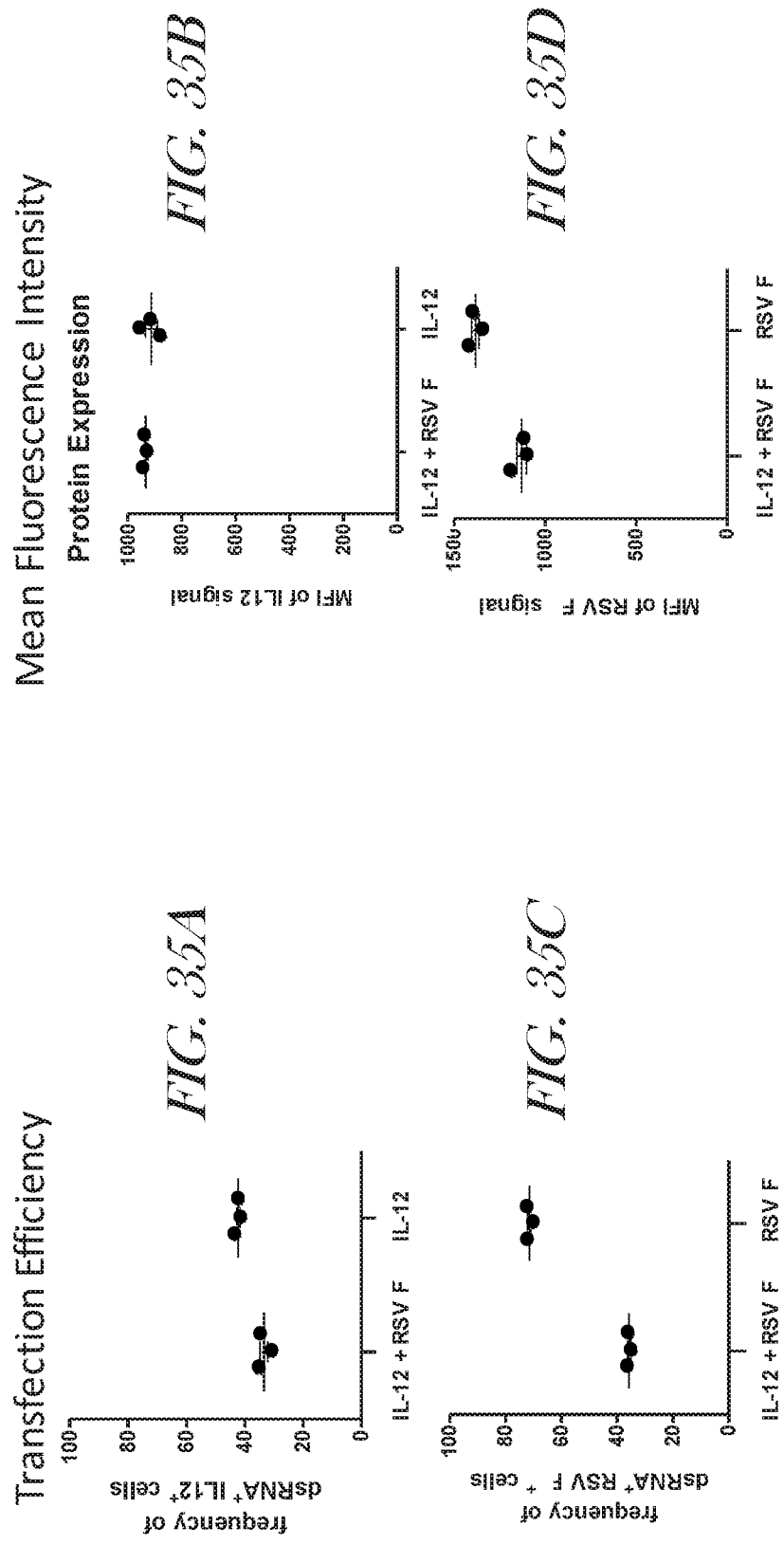
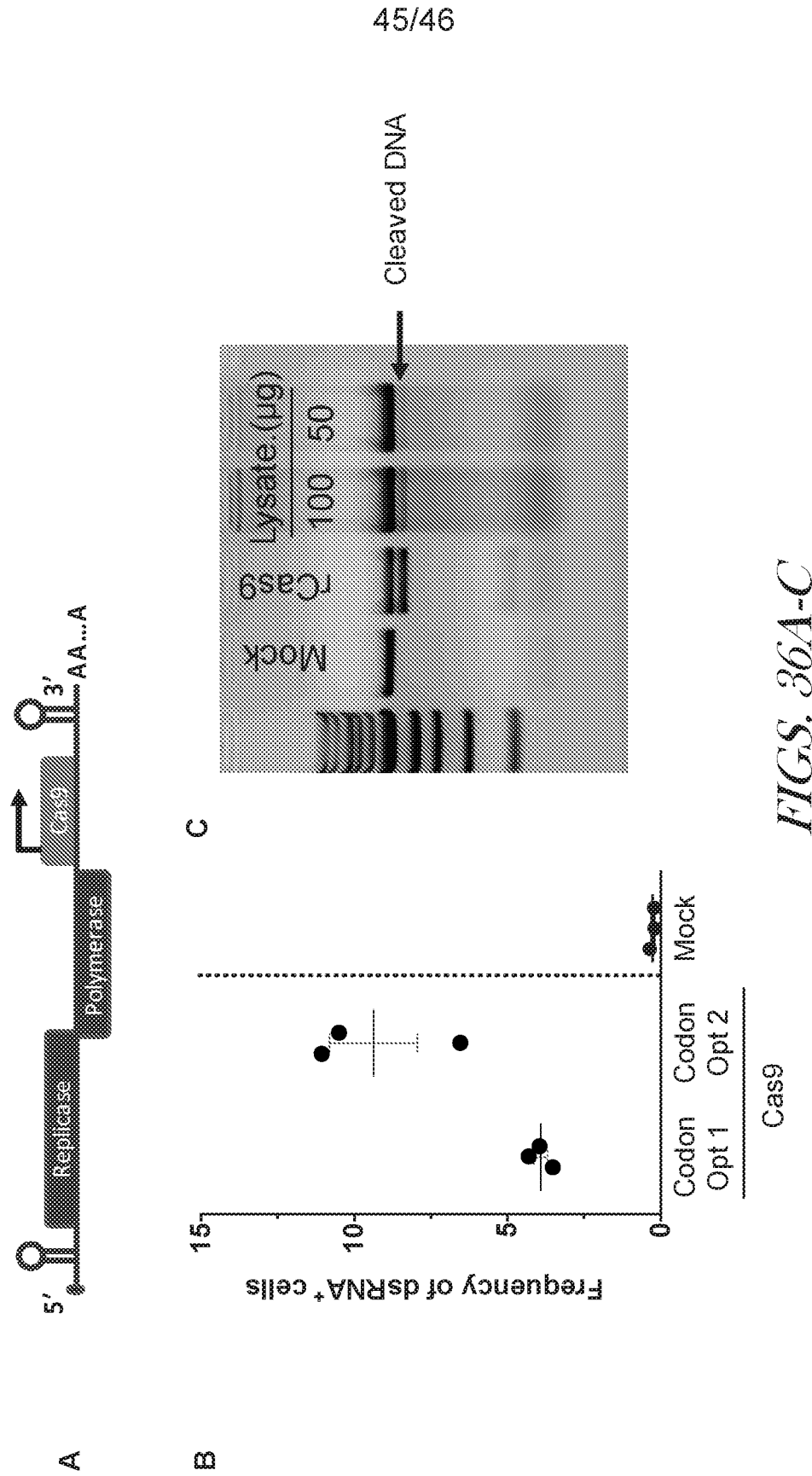


FIG. 34





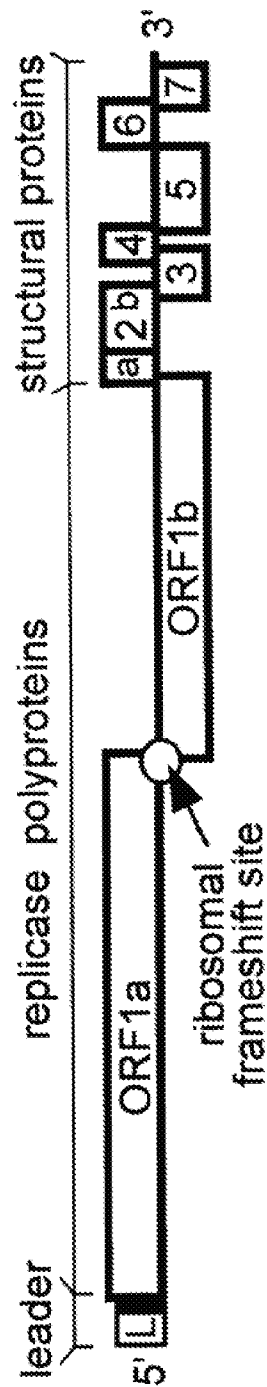


FIG. 37

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2017/027249

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/86 C12N7/00 C07K16/32
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 C07K

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, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	VAN DEN BORN ERWIN ET AL: "An infectious recombinant equine arteritis virus expressing green fluorescent protein from its replicase gene.", THE JOURNAL OF GENERAL VIROLOGY APR 2007, vol. 88, no. Pt 4, April 2007 (2007-04), pages 1196-1205, XP002771364, ISSN: 0022-1317 page 1196 abstract page 1197; figure 1 ----- -/--	1-80



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

26 June 2017

Date of mailing of the international search report

10/07/2017

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

Sitch, David

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2017/027249

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DE VRIES A A ET AL: "Recombinant equine arteritis virus as an expression vector.", VIROLOGY 05 JUN 2001, vol. 284, no. 2, 5 June 2001 (2001-06-05), pages 259-276, XP002771365, ISSN: 0042-6822 page 259 abstract page 260; figure 1	74,75
A	----- MOLENKAMP R ET AL: "The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription.", THE JOURNAL OF GENERAL VIROLOGY OCT 2000, vol. 81, no. Pt 10, October 2000 (2000-10), pages 2491-2496, XP002771366, ISSN: 0022-1317 Page 2493, table 1; Page 2493, paragraph bridging left and right hand columns; Page 2494, figure 2, page 2495, first full paragraph	1
A	----- SNIJDER E J ET AL: "Identification of a novel structural protein of arteriviruses.", JOURNAL OF VIROLOGY AUG 1999, vol. 73, no. 8, August 1999 (1999-08), pages 6335-6345, XP002771367, ISSN: 0022-538X Page 6340, RH column, first full paragraph - page 6341, first partial paragraph page 6335, paragraph 1	1
A	----- TIAN DEBIN ET AL: "Arterivirus minor envelope proteins are a major determinant of viral tropism in cell culture.", JOURNAL OF VIROLOGY APR 2012, vol. 86, no. 7, April 2012 (2012-04), pages 3701-3712, XP002771368, ISSN: 1098-5514 RH column, last 5 lines; page 3701 Paragraph bridging right and left columns; page 3702 page 3704; figure 1 -----	1