



(51) International Patent Classification:

C12N 15/09 (2006.01) *C12N 15/86* (2006.01)
C12N 15/63 (2006.01) *C12N 15/864* (2006.01)

(21) International Application Number:

PCT/US2019/054539

(22) International Filing Date:

03 October 2019 (03.10.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/742,835 08 October 2018 (08.10.2018) US
 62/749,012 22 October 2018 (22.10.2018) US
 62/810,281 25 February 2019 (25.02.2019) US

(71) Applicants: **ALLEN INSTITUTE** [—/US]; 615 Westlake Ave N., Seattle, Washington 98109 (US). **SEATTLE CHILDREN'S HOSPITAL (D/B/A SEATTLE CHILDREN'S RESEARCH INSTITUTE)** [US/US]; 1900 Ninth Avenue, Seattle, Washington 98101 (US).

(72) Inventors: **TING, Jonathan**; 615 Westlake Ave N., Seattle, Washington 98109 (US). **LEVI, Boaz P.**; 615 Westlake Ave N., Seattle, Washington 98109 (US). **MICH, John K.**; 615 Westlake Ave N., Seattle, Washington 98109 (US).

LEIN, Edward Sebastian; 615 Westlake Ave N., Seattle, Washington 98109 (US). **KALUME, Franck**; 1900 Ninth Avenue, Seattle, Washington 98101 (US).

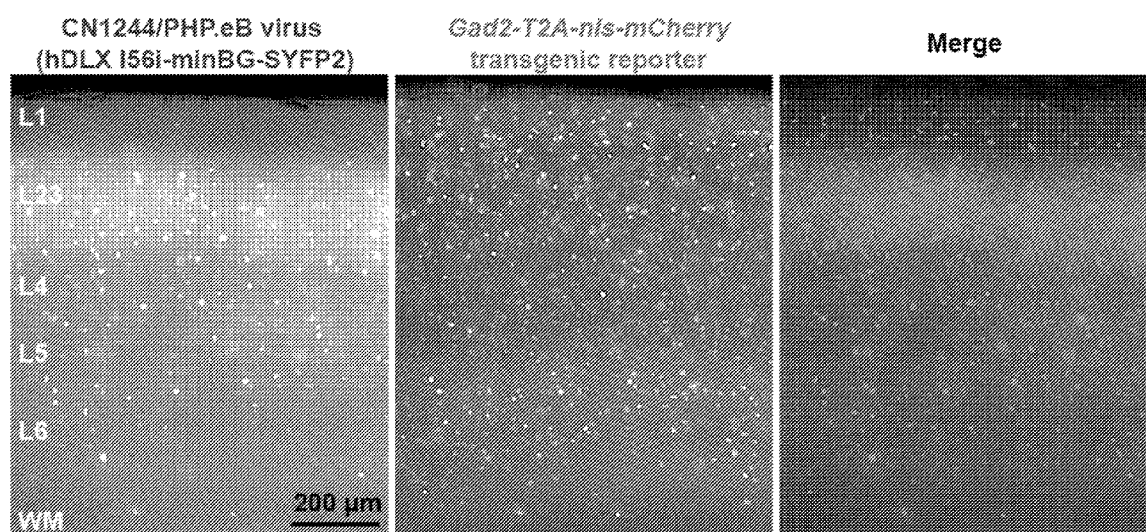
(74) Agent: **WINGER, C. Rachal** et al.; Lee & Hayes, PC, 601 W. Riverside Ave., Suite 1400, Spokane, Washington 99201 (US).

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

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

(54) Title: ARTIFICIAL EXPRESSION CONSTRUCTS FOR SELECTIVELY MODULATING GENE EXPRESSION IN INTERNEURONS

FIG. 1



(57) Abstract: Artificial expression constructs for selectively modulating gene expression in selected central nervous system cell types are described. The artificial expression constructs can be used to selectively express synthetic genes or modify gene expression in GABAergic interneurons.

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
- *with sequence listing part of description (Rule 5.2(a))*

ARTIFICIAL EXPRESSION CONSTRUCTS FOR SELECTIVELY MODULATING GENE EXPRESSION IN INTERNEURONS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to US Provisional Patent Application Nos. 62/742,835 filed October 8, 2018; 62/749,012 filed October 22, 2018; and 62/810,281 filed February 25, 2019, each of which is incorporated herein by reference in its entirety as if fully set forth herein.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

[0002] This invention was made with government support under grant RF1MH114126 awarded by the National Institutes of Health. The government has certain rights in the invention.

REFERENCE TO SEQUENCE LISTING

[0003] The Sequence Listing associated with this application is provided in text format in lieu of a paper copy and is hereby incorporated by reference into the specification. The name of the text file containing the Sequence Listing is A166-0006PCT_ST25.txt. The text file is 379 KB, was created on October 3, 2019, and is being submitted electronically via EFS-Web.

FIELD OF THE DISCLOSURE

[0004] The current disclosure provides artificial expression constructs for selectively modulating gene expression in selected central nervous system cell types. The artificial expression constructs can be used to selectively express synthetic genes or modify gene expression in GABAergic forebrain interneurons.

BACKGROUND OF THE DISCLOSURE

[0005] GABAergic interneurons play critical roles in central nervous system processing as well as development. Dysfunction of these cells can also contribute to numerous neuropsychiatric disorders, such as schizophrenia and autism. GABAergic interneurons also play a role in epilepsy.

[0006] Cell-type or cell-class specific gene delivery using non-pathogenic recombinant adeno-associated virus (rAAV) is showing increasing promise for the treatment of diverse diseases. Inclusion within rAAVs of one or more cis-acting DNA-control elements, such as specific promoters or enhancers, has been beneficial to provide specificity of expression within particular target cells, including specific cell types or cell classes in the brain.

[0007] Dimidschstein and colleagues (*Nat Neurosci* 19(12):1743-1749, 2016) developed a rAAV that permits largely selective gene expression in GABAergic interneurons within the telencephalon. This rAAV includes a 527 bp enhancer sequence (referred to as m156i or mDlx)

from the intergenic interval between the distal-less homeobox 5 and 6 genes (*Dlx5/6*), which are naturally expressed by forebrain GABAergic interneurons during embryonic development. The construct of Dimidschstein *et al.* is available on Addgene as ID# 83900 (in which the enhancer drives eGFP expression). Additional constructs which employ the murine or human I56i enhancer to drive various transgenes are available through Addgene, such as Plasmid ID #s 83899 (driving GCaMP6f expression), 83898 (driving ChR2 expression), 83895 (driving synthetic eGFP expression), 89897 (driving hM3DREADD expression), 83896 (driving hM4Di expression), and 83894 (driving synthetic tdTomato expression). See also U.S. Patent Publication No. US2018/0078658.

[0008] Additionally, the mI56i enhancer has previously been used to reliably target reporter genes in a pattern very similar to the normal patterns of *Dlx5/6* expression during embryonic development (Zerucha *et al.*, *J Neuroscience* 20:709-721, 2000; Stühmer *et al.*, *Cerebral Cortex* 12:75-85, 2002; Stenman *et al.*, *J Neuroscience* 23:167-174, 2003; Monory *et al.*, *Neuron*. 51:455-455, 2006; Miyoshi *et al.*, *J Neuroscience* 30:1532-1594, 2010).

[0009] One significant drawback to using rAAVs as a gene-delivery system is the restricted packaging limit of AAVs; this is particularly limiting to the inclusion of lengthy genetic control and expression elements. In addition, many existing interneuron-specific rAAV expression constructs can provide weak gene expression reducing their usefulness in research and therapeutic uses.

SUMMARY OF THE DISCLOSURE

[0010] The current disclosure overcomes drawbacks of the prior art by providing engineered enhancer elements that provide rapid and strong cell-specific expression of heterologous encoding sequences in forebrain GABAergic interneurons.

[0011] In particular embodiments, the artificial enhancer elements include a concatemerized core of a I56i enhancer. These artificial enhancer elements provide more rapid onset of transgene expression compared to a single full length original (native) enhancer.

[0012] In particular embodiments, the I56i enhancer core can be derived from, for example the human, murine, or zebrafish I56i enhancer (SEQ ID NOs. 1, 4, and 5 respectively). The selected cores of the I56i enhancer can include SEQ ID NO: 2 (core shared by human and mouse) or SEQ ID NO: 6 (zebrafish core). In particular embodiments, the cores are concatemerized. For example, SEQ ID NO: 3 provides a three-copy concatemer of the selected human/murine I56i core while SEQ ID NO: 7 provides a three-copy concatemer of the selected zebrafish I56i core.

[0013] Of particular interest, the synthetic 3x human/murine core (referred to herein as the 3xhI56iCore; SEQ ID NO: 3) is shorter than the original full length enhancer sequence reported

in Dimidschstein *et al.* (*Nat Neurosci* 19(12):1743-1749, 2016), despite being a 3x concatemer. Thus, this concatemerized core provides more room for cargo genes linked to the enhancer, which is highly desirable. Moreover, the peak level of transgene expression driven by the 3xh156iCore enhancer is much greater than simply three times the level of the original single full-length original enhancer.

[0014] The engineered concatemerized I56i cores disclosed herein enable new and improved gene delivery vectors that are particularly useful for achieving selective transgene expression in forebrain GABAergic interneurons in diverse animal species, including humans.

BRIEF DESCRIPTION OF THE FIGURES

[0015] Many of the drawings submitted herein are better understood in color. Applicants consider the color versions of the drawings as part of the original submission and reserve the right to present color images of the drawings in later proceedings.

[0016] FIG. 1: Virus CN1244/PHP.eB. 10^{11} genome copies delivered intravenously (IV) in adult mouse. PHP.eB encodes for a capsid originating from AAV9 that allows efficient AAV transit across the mouse blood brain barrier, which enables delivery of AAV vectors in a brain-wide fashion. This capsid differs from AAV9 such that amino acids starting at residue 586: SAQA (SEQ ID NO: 98) are changed to SDGTLAVPFKA (SEQ ID NO: 33). The *Gad2-T2A-nls-mCherry* reporter marks nearly all inhibitory neurons in the mouse brain (here shown V1 visual cortex), and the delivered CN1244/PHP.eB virus drives specific SYFP2 reporter activity in forebrain GABAergic neurons.

[0017] FIGs. 2A,2B: Comparison of CN1244 vs CN1389 vs CN1390. (FIG. 2A) Schematic representations of three vector constructs, CN1390, CN1389, and CN1244 (CN1203 scAAV). Key: h156i – full length human enhancer (black box; SEQ ID NO: 1); selected h156i core (grey box; SEQ ID NO: 2) and 3X concatemer of core (grey boxes; SEQ ID NO: 3); minBG - minimal beta globin promoter; SYFP2 - super yellow fluorescent protein 2; WPRE3 - woodchuck hepatitis virus post-transcriptional regulatory element 3; BGHpA - bovine growth hormone polyA sequence; L-ITR and R-ITR - Adeno-associated virus-2 (AAV2) inverted terminal repeats (ITRs). (FIG. 2B) Fluorograph images showing relative expression of SYFP2 from AAV vector constructs CN1244, CN1389, and CN1390. Adult wild type mice were retro-orbitally injected with $1E+11$ genome copies of the indicated viruses. Animals were maintained for 3-4 weeks, then euthanized, with brains extracted and sliced, followed by live tissue epifluorescence imaging of native fluorescence. Exposure times were matched to allow direct comparison of transgene expression levels. The first three panels are a 500 msec exposure for each of the indicated constructs; the

fourth panel is a shorter (50 msec) exposure image of CN1390. CN1390 with the engineered concatemerized core demonstrated strong and more rapid transgene expression.

[0018] FIG. 3: CN1390 retains cell type specificity for the pan-GABAergic neuron population. Cortical/hippocampal brain slice cultures were prepared from P5-10 Gad2-IRES-Cre het/Ai75 het animals. One hour after culturing, CN1390 viral suspension was pipetted onto the slice surface to transduce brain cell types. At 10 DIV/10DPI, native fluorescence was imaged in green and red channels on a Nikon inverted microscope. 10DIV/10DPI. DIV: days *in vitro*, DPI: days post infection.

[0019] FIGs. 4A, 4B: Comparison of CN1244 vs CN1390 in non-human primate *ex vivo* brain slice culture. (FIG. 4A) Fluorograph images showing relative expression of SYFP2 from AAV vector constructs CN1244 and CN1390. Neocortical slices were cultured from adult macaque brain and infected with nominally-matched titers of the indicated viruses. Brain slice cultures were maintained in the incubator for 4 days *in vitro*, 4 days post infection (4DIV/4DPI), then used for live tissue epifluorescence imaging of native fluorescence. Exposure times were matched to allow direct comparison of transgene expression levels. CN1390 with the engineered concatemerized core demonstrated strong and more rapid transgene expression. (FIG. 4B) Fluorograph images showing relative expression of SYFP2 from AAV vector constructs CN1244 and CN1390. Hippocampal slices were cultured from adult macaque brain and infected with nominally-matched titers of the indicated viruses. Brain slice cultures were maintained in the incubator for 6 days *in vitro*, 6 days post infection (6DIV/6DPI), then used for live tissue epifluorescence imaging of native fluorescence. Exposure times were matched to allow direct comparison of transgene expression levels. CN1390 with the engineered concatemerized core demonstrated strong and more rapid transgene expression.

[0020] FIGs. 5A-5E: CN1390 exhibits rapid onset of transgene expression in human *ex vivo* brain slices. Human *ex vivo* neocortical brain slice cultures were prepared from live neurosurgical specimens as described in Ting *et al.*, *Scientific Reports* 8(1):8407, 2018. One hour after culturing, CN1390 viral suspension was pipetted onto the slice surface to transduce brain cell types. At 1, 3, and 6DIV/DPI, native SYFP2 fluorescence was imaged using matched exposure times on a Nikon microscope. FIGs. 5A-5D illustrate rapid viral-genetic labeling of human neocortical interneurons for targeted patch clamp recording and analysis. (FIG. 5A) Time course of virus-mediated YFP expression following human brain slice transduction with CN1390 eB. (FIG. 5B) Expanded view of the boxed region in (FIG. 5A). (FIG. 5C) High magnification view of a virus labeled interneuron with bipolar morphology. (FIG. 5D) Example whole cell recordings from four different virus-labeled YFP+ human interneurons demonstrating diverse firing patterns to supra-

threshold current injection. (FIG. 5E) At various times in culture, slices were taken for terminal patch clamp recording analysis to establish the firing properties of labeled neurons. Functional analysis of human neocortical interneuron firing patterns and electrical properties by patch clamp recording was feasible as early as 40 hours post-infection with CN1390 AAV-PHP.eB virus.

[0021] FIG. 6: CN1390 maintains GABAergic cell class selectivity. At 7 to 34 DIV/DPI, virally transduced human organotypic slices from 4 unique human donors were transduced, dissociated, and 234 single SYFP2⁺ cells were FACS sorted from glia and debris-depleted cell suspensions, and profiled by single cell RNA-seq (SMARTer V.4). These cells were mapped to the existing MTG cell type taxonomy. Bars at the bottom of the taxonomy indicate the number of SYFP⁺ cells that mapped to the final leaf. Circles farther up the taxonomy indicate the number of cells that could only be mapped to that branch point. Note that cells from all major GABAergic classes were labeled, and no glutamatergic or glial cells were recovered. The listed cell types from top to bottom are: GABAergic types; 3 Inh L1-2 PAX-6 CDH12, 4 Inh L1-2 PAX6 TNF AIP8L3, 5 Inh L1 SST NMBR (ADARB2+), 6 Inh L1-4 LAMP5 LCP2 (rosehip), 7 Inh L1-2 LAMP5 DBP, 8 Inh L2-6 LAMP5 CA1 (lgtp), Inh L1 SST CHRNA4 (ADARB2+), 14 Inh L1-2 GAD1 MC4R (ADARB2+), 15 Inh L1-2 SST BAGE2 (ADARB2+), 17 Inh L1-3 PAX6 SYT6 (Sncg), 19 Inh L1-2 VIP TSPAN12, 20 Inh L1-4 VIP CHRNA6, 21 Inh L1-3 VIP ADAMTSL1, 22 Inh L1-4 VIP PENK, 27 Inh L2-6 VIP QPCT, 28 Inh L3-6 VIP HS3ST3A1, 29 Inh L1-2 VIP PCDH20, 31 Inh L2-5 VIP SERPINF1, 32 Inh L2-5 VIP TYR, 37 Inh L1-3 VIP CHRM2, 38 Inh L2-4 VIP CBLN1, 39 Inh L1-3 VIP CCDC184, 40 Inh L1-3 VIP GGH, 42 Inh L1-2 VIP LBH, 43 Inh L2-3 VIP CASC6, 45 Inh L2-4 VIP SPAG17, 46 Inh L1-4 VIP OPRM1, Inh L3-6 SST NPY (Chodl), 52 Inh L3-6 SST HPGD, 55 Inh L4-6 SST B3GAT2, 56 Inh L5-6 SST KLHDC8A, 57 Inh L5-6 SST NPM1P10, 58 Inh L4-6 SST GXYLT2, 59 Inh L4-5 SST STK32A, 62 Inh L1-3 SST CALB1, 63 Inh L3-5 SST ADGRG6, 64 Inh L2-4 SST FRZB, 65 Inh L5-6 SST TH, 66 Inh L5-6 GAD1 GLP1R (LHX6+), 68 Inh L5-6 PVALB LGR5, 71 Inh L4-5 PVALB MEPE, 73 Inh L2-4 PVALB WFDC2, 74 Inh L4-6 PVALB SULF1, 75 Inh L5-6 SST MIR548F2, 76 Inh L2-5 PVALB SCUBE3 (chandelier), Excitatory types; 82 Exc L2-5 LAMP5 LTK, 83 Exc L2-4 LINC00507 GLP2R, 84 Exc L2-3 LINC00507 FREM3, 85 Exc L5-6 THEMIS C1QL3, 87 Exc L3-4 RORB CARM1P1, 89 Exc L3-5 RORB ESR1, 90 Exc L3-5 RORB COL22A1, 92 Exc L3-5 RORB FILIP1L, 93 Exc L3-5 RORB TWIST2, 96 Exc L4-5 RORB FOLH1B, 98 Exc L4-6 RORB SEMA3E, 99 Exc L4-5 RORB DAPK2, 100 Exc L5-6 RORB TTC12, 101 Exc L4-6 RORB C1R, Exc L4-5 FEZF2 SCN4B (PT), 102 Exc L5-6 THEMIS DCSTAMP, 103 Exc L5-6 THEMIS CRABP1, 104 Exc L5-6 THEMIS FGF10, 105 Exc L4-6 FEZF2 IL26 (NP), 106 Exc L5-6 FEZF2 ABO, 107 Exc L6 FEZF2 SCUBE1, 108 Exc L5-6 SLC17A7 IL15, 109 Exc L6 FEZF2 OR2T8, 110 Exc L5-6 FEZF2 EFTUD1P1, Glial types; OPC L1-6 PDGFRA, Astro L1-6 FGFR3

SLC14A1, Astro L1-2 FGFR3 GFAP, Oligo L1-6 OPALIN, Endo L2-6 NOSTRIN, AND Micro L1-3 TYROBP.

[0022] FIG. 7: Fast expression from CN1390 allows assessment of human circuit connectivity. Human neocortical organotypic slice was transduced with CN1390 and AAV-hSyn1-dTomato for 2.5 days. After only two and half days in culture, GABAergic and all neuronal cells can be labeled in culture using CN1390 and AAV-hSyn1-dTomato, respectively. Human synapsin 1 (hSyn1) is a well-known pan-neuronal promoter. This allows the assessment of connectivity between prospectively virally marked patched cells (labeled by cascade blue). The fluorescent dyes listed in the bottom left corner of the fluorescent image are (from top to bottom): panGABA-SYFP, hSyn1-tdTomato, and Fill-Blue.

[0023] FIGs. 8A, 8B: All major classes of human neocortical GABAergic neurons are marked by CN1390. (FIG. 8A) Multiplexed FISH using HCR v3.0 reveals major classes of GABAergic neurons labeled by somatostatin (SST), parvalbumin (PVALB), or vasoactive intestinal peptide (VIP) genes. Labeling by CN1390 in 350 μ m thick neocortical brain slice culture is shown. Text on the left image of FIG. 8A are as follows: (top left) Pial surface; (top right) Lipofuscin, PVALB, SST, VIP, and SYFP; and (bottom left) Hu, 350 μ m Slice, Virus CN1390eB, 7DIV/DPI. (FIG. 8B) Prospective cell class marking for physiology, connectivity and morphology. Multiplexed FISH reveals molecular identities of the CN1390-labeled cell classes and some of the patched cells that were back-filled with neurobiotin and visualized by Streptavidin-BV421. The left image of FIG. 8B is labeled with SYFP, SST, VIP, PVALB, and Lipofuscin. The right image of FIG. 8B is labeled Biocytin-BV421. All these cells show GABAergic cell morphology and most were marked by expression of SYFP2 from CN1390.

[0024] FIGs. 9A-9F. AAV vector reagents to reverse Dravet Syndrome (DS) symptoms in *Scn1a*^{+/-} mice. (9A) Vectors to deliver epitope-tagged Nav genes of bacterial origin (NavBacs). The Nav genes shown here are NavMs (from *Magnetococcus marinus*), NavBp (from *Bacillus pseudofirmus*), and NavSheP-D60N (from *Shewanella putrefaciens* with an engineered D60N mutation). These examples all have N-terminal epitope tags (hexahistidine in the case of CN1367, or 3xHA for CN1498, CN1499, and CN1500). hI56i refers to the full-length I56i enhancer of SEQ ID NO: 1; 3xhI56iCore refers to the concatemerized core of the I56i enhancer (SEQ ID NO: 3); (9B) Graded expression levels from NavBac vectors. (9C) Weak but detectable expression in Pvalb interneurons from vector CN1367. (9D) Trend towards seizure protection with vector CN1367. (9E) Vector 1500 drives high-level expression in Pvalb⁺ and Pvalb⁻ interneurons throughout cortex. (9F) Abundant production of HA-tagged NavBacs in cell bodies and proximal processes with vectors 1498 and 1500, but not 1499.

[0025] FIG. 10. CN1500 rAAV vector substantially reverses febrile seizures in *Scn1a*^{+/-} mice. Febrile seizure assay shown as internal temperature where a seizure is first detected. (Top) Circles show *Scn1a*^{+/-} mice untransduced with AAVs, while the diamonds represent animal that were transduced with CN1500. The large dot and error bars represent the average $\pm$ SEM for each group of animals. (Bottom) Trends of the same data are shown as the percentage of mice in each group that remain seizure free at different temperatures using a Kaplan-Meier curve.

[0026] FIGs. 11A, 11B: Conservation of I56i enhancer sequences. (FIG. 11A) alignment of the human (SEQ ID NO: 1) I56i, murine (SEQ ID NO: 4) I56i, and zebrafish (SEQ ID NO: 5) I46i enhancer sequences. Residues shared by all three sequences are highlighted in light gray; those shared by the murine and human sequences are highlighted in dark gray. The core sequence (SEQ ID NO: 2) corresponds to positions 268-398 of the illustrated human sequence. The mouse and human I56i enhancer core sequences are exactly identical (100% sequence identity), as this is an ultraconserved enhancer sequence. It is also highly similar with the zebrafish genomic sequence, and the orthologous zebrafish enhancer (called I46i) has been used in many contexts over the years to drive transgene expression in neocortical interneurons, including for mouse neocortical interneurons. (FIG. 11B) graph illustrating the similarity between the human, murine, and zebrafish enhancer sequences. Graph shows (from right to left, and as labeled) Similarity, Absolute Complexity, and Absolute Complexity (human I56i).

[0027] FIG. 12: the sequence and indicated features of construct CN1389 pAAV-hI56i(core)-minBG-SYFP2-WPRE3-BGHpA (SEQ ID NO: 41). Select restriction endonuclease sites are indicated, as are the regions corresponding to different parts of the construct.

[0028] FIG. 13: the sequence and indicated features of construct CN1390 pAAV-3xhI56i(core)-minBG-SYFP2-WPRE3-BGHpA (SEQ ID NO: 42). Select restriction endonuclease sites are indicated, as are the regions corresponding to different parts of the construct.

[0029] FIG. 14: the sequence and indicated features of construct CN1203 scAAV-hI56i-minbGlobin-SYFP2-WPRE3-BGHpA (SEQ ID NO: 43). Select restriction endonuclease sites are indicated, as are the regions corresponding to different parts of the construct.

[0030] FIG. 15. Features of exemplary vectors disclosed herein.

[0031] FIG. 16. Artificial expression constructs within the teaching of the current disclosure. Each construct begins with a concatemerized core of the hI56i core (e.g., SEQ ID NO: 3 or 7) designated as *. The following abbreviations are also used: Beta-Globin minimal promoter (minB, referred to as minBglobin elsewhere herein), Minimal cytomegalovirus promoter (minC, referred to as minCMV elsewhere herein), Mutated minimal cytomegalovirus promoter (mut), Minimal rhodopsin promoter (minR, referred to as minRho elsewhere herein), Cytomegalovirus promoter

(CMV), Simian vacuolating virus 40 promoter (SV40), Hsp68 minimal promoter (H68, referred to as proHSP68 elsewhere herein), Rous Sarcoma Virus long-terminal repeat promoter (RSV), Fluorescent protein (FP), Blue fluorescent protein (BFP), Cyan fluorescent protein (CFP), Green fluorescent protein (GFP), Orange fluorescent protein (OPF), Red fluorescent protein (RFP), Far red fluorescent protein (fRFP), Yellow fluorescent protein (YFP), Luciferase (Luc), Enzyme (enz), Transcription factor (TF), Receptor (rec), Cellular trafficking protein (CTP), Signaling molecule (SM), Neurotransmitter (NT), Calcium reporter (CR), channel rhodopsin (ChR), Guide RNA (gRNA), Nuclease (Nuc), Woodchuck hepatitis virus post-transcriptional response element (W, referred to as WPRE3 elsewhere herein), Bovine growth hormone polyadenylation signal (bG, referred to as bGHpA elsewhere herein), Simian vacuolating virus 40 polyadenylation signal (S, referred to as SV40pA elsewhere herein), Internal ribosome entry site 2 (I2, referred to as IRES2 elsewhere herein), and 2A skipping elements (T2A, P2A, E2A, and F2A).

[0032] FIG. 17. Additional sequences supporting the disclosure: hI56i enhancer: (SEQ ID NO: 1); Core of the hI56i enhancer: (SEQ ID NO: 2); 3xhI56iCore, Triply Concatamerized Core of the hI56i enhancer: (SEQ ID NO: 3); Murine I56i Enhancer (core is the same as human): (SEQ ID NO: 4); Zebrafish I46i Enhancer: (SEQ ID NO: 5); Core of the Zebrafish I46i Enhancer: (SEQ ID NO: 6); 3x Concatamerized Core of the Zebrafish I46i Enhancer: (SEQ ID NO: 7); Beta-Globin Minimal Promoter pBGmin/minBGlobin/minBGprom): (SEQ ID NO: 8); minCMV Promoter: (SEQ ID NO: 9); Mutated minCMV Promoter (Sacl RE site removed): (SEQ ID NO: 10); minRho Promoter: (SEQ ID NO: 11); Hsp68 minimal Promoter (proHsp68): (SEQ ID NO: 12); SYFP2: (SEQ ID NO: 13); EGFP: (SEQ ID NO: 14); Optimized Flp recombinase (FlpO): (SEQ ID NO: 15); Improved Cre recombinase (iCre): (SEQ ID NO: 16); NavMs, endogenous sequence: (SEQ ID NO: 17); NavMs, codon optimized, with N-terminal 3x HA tag and linker: (SEQ ID NO: 18); NavMs, codon optimized, with N-terminal His tag and linker: (SEQ ID NO: 19); NavBp, endogenous sequence: (SEQ ID NO: 20); NavBp, codon optimized, with N-terminal 3x HA tag: (SEQ ID NO: 21); NavSheP-D60N, codon optimized, with N-terminal 3x HA tag: (SEQ ID NO: 22); NavSheP endogenous sequence: (SEQ ID NO: 23); WPRE3: (SEQ ID NO: 24); BGHpA: (SEQ ID NO: 25); P2A Encoding Sequence: (SEQ ID NO: 26); P2A: (SEQ ID NO: 27); T2A: (SEQ ID NO: 28); E2A: (SEQ ID NO: 29); F2A: (SEQ ID NO: 30); N-terminal 3XHA tag: (SEQ ID NO: 31); N-terminal 3XHA tag: (SEQ ID NO: 32); PHP.eB capsid: (SEQ ID NO: 90); AAV9 VP1 capsid protein: (SEQ ID NO: 34); tet-Transactivator version 2 (tTA2): (SEQ ID NO: 35); CN1367 - The portion between L-ITR and R-ITR: positions 142-2984: (SEQ ID NO: 36); CN1500 - The portion between L-ITR and R-ITR: positions 142-2976: (SEQ ID NO: 37); CN1498 - The portion between L-ITR and R-ITR: positions 142-2943: (SEQ ID NO: 38); CN1499 - The portion between L-ITR and R-ITR:

positions 142-2946: (SEQ ID NO: 39); CN1244 - The portion between L-ITR and R-ITR: positions 142-2042: (SEQ ID NO: 40); CN1389 - The portion between L-ITR and R-ITR corresponds to positions 142-1660: (SEQ ID NO: 41); CN1390 - The portion between L-ITR and R-ITR corresponds to positions 142-1897: (SEQ ID NO: 42); CN1203 - The portion between L-ITR and R-ITR corresponds to positions 183-2052: (SEQ ID NO: 43); Lactase (SEQ ID NO: 44); Lipase (SEQ ID NO: 45); Helicase (SEQ ID NO: 46); Amylase (SEQ ID NO: 47); α -glucosidase (SEQ ID NO: 48); Transcription factor SP1 (SEQ ID NO: 49); Transcription factor AP-1 (SEQ ID NO: 50); Heat shock factor protein 1 (SEQ ID NO: 51); CCAAT/enhancer-binding protein (C/EBP) β isoform a (SEQ ID NO: 52); Octamer-binding protein 1 (SEQ ID NO: 53); Transforming growth factor receptor β 1 (SEQ ID NO: 54); Platelet-derived growth factor receptor (SEQ ID NO: 55); Epidermal growth factor receptor (SEQ ID NO: 56); Vascular endothelial growth factor receptor (SEQ ID NO: 57); Interleukin 8 receptor α (SEQ ID NO: 58); Caveolin (SEQ ID NO: 59); Dynamin (SEQ ID NO: 60); Clathrin heavy chain 1 isoform 1 (SEQ ID NO: 61); Clathrin heavy chain 2 isoform 1 (SEQ ID NO: 62); Clathrin light chain A isoform a (SEQ ID NO: 63); Clathrin light chain B isoform a (SEQ ID NO: 64); Ras-related protein Rab-4A isoform 1 (SEQ ID NO: 65); Ras-related protein Rab-11A (SEQ ID NO: 66); Platelet-derived growth factor (SEQ ID NO: 67); Transforming growth factor- β 3 (SEQ ID NO: 68); Nerve growth factor (SEQ ID NO: 69); Epidermal growth factor (SEQ ID NO: 70); GTPase HRas (SEQ ID NO: 71); Cocaine And Amphetamine Regulated Transcript (Chain A) (SEQ ID NO: 72); Protachykinin-1 (SEQ ID NO: 73); Substance P (SEQ ID NO: 74); Oxytocin-neurophysin 1 (SEQ ID NO: 75); Oxytocin (SEQ ID NO: 76); Somatostatin (SEQ ID NO: 77); Myosin light chain kinase, Green fluorescent protein, Calmodulin chimera (Chain A) (SEQ ID NO: 78); Genetically-encoded green calcium indicator NTnC (chain A) (SEQ ID NO: 79); Calcium indicator TN-XXL (SEQ ID NO: 80); BRET-based auto-luminescent calcium indicator (SEQ ID NO: 81); Calcium indicator protein OeNL(Ca²⁺)-18u (SEQ ID NO: 82); GCaMP6m (SEQ ID NO: 99); GCaMP6s (SEQ ID NO: 100); GCaMP6f (SEQ ID NO: 101); Channelopsin 1 (SEQ ID NOs: 83 and 102); Channelrhodopsin-2 (SEQ ID NOs: 84 and 103); CRISPR-associated protein (Cas) (SEQ ID NO: 85); Cas9 (SEQ ID NO: 86); CRISPR-associated endonuclease Cpf1 (SEQ ID NO: 87); Ribonuclease 4 or Ribonuclease L (SEQ ID NO: 88); Deoxyribonuclease II β (SEQ ID NO: 89); Sodium channel protein type 1 subunit alpha (SEQ ID NO: 104); Potassium voltage-gated channel subfamily KQT member 2 (SEQ ID NO: 105); and Voltage-dependent L-type calcium channel subunit alpha-1C (SEQ ID NO: 106).

DETAILED DESCRIPTION

[0033] To fully understand the biology of the brain, different cell types need to be distinguished

and defined. To identify and/or study these different cell types, vectors that can selectively label and perturb them need to be identified. In mouse, recombinase driver lines have been used to great effect to label cell populations that share marker gene expression. However, the creation, maintenance, and use of such lines that label cell types with high specificity can be costly, frequently requiring triple transgenic crosses, which yield a low frequency of experimental animals. Furthermore, those tools require germline transgenic animals and thus are not applicable to humans, and recent advances in single-cell profiling, such as single-cell RNA-seq (Tasic et al., *Nature* 563, 72-78 (2018); Tasic 2016, *Nat Neurosci* 19, 335-346) and surveys of neural electrophysiology and morphology (Gouwens 2019, *Nat Neurosci* 22, 1182-1195), have revealed that many recombinant driver lines label heterogeneous mixtures of cell types, and often include cells from multiple subclasses. For example, the Rbp4-Cre mouse driver line, which is commonly used to label layer 5 (L5) neurons, also labels cells with drastically different connectivity patterns: L5 intratelencephalic (IT, also called cortico-cortical) and pyramidal tract (PT, also called cortico-subcortical) neurons.

[0034] Dimidschstein and colleagues (*Nat Neurosci* 19(12):1743-1749, 2016) developed a rAAV that permits largely selective gene expression in GABAergic interneurons within the telencephalon. This rAAV includes a 527 bp enhancer sequence (referred to as mI56i or mDlx) from the intergenic interval between the distal-less homeobox 5 and 6 genes (*Dlx5/6*), which are naturally expressed by forebrain GABAergic interneurons during embryonic development. The construct of Dimidschstein *et al.* is available on Addgene as ID# 83900 (in which the enhancer drives eGFP expression). Additional constructs which employ the murine or human I56i enhancer to drive various transgenes are available through Addgene, such as Plasmid ID #s 83899 (driving GCaMP6f expression), 83898 (driving ChR2 expression), 83895 (driving synthetic eGFP expression), 89897 (driving hM3DREADD expression), 83896 (driving hM4Di expression), and 83894 (driving synthetic tdTomato expression). See also U.S. Patent Publication No. US2018/0078658.

[0035] Additionally, the mI56i enhancer has previously been used to reliably target reporter genes in a pattern very similar to the normal patterns of *Dlx5/6* expression during embryonic development (Zerucha *et al.*, *J Neuroscience* 20:709-721, 2000; Stühmer *et al.*, *Cerebral Cortex* 12:75-85, 2002; Stenman *et al.*, *J Neuroscience* 23:167-174, 2003; Monory *et al.*, *Neuron* 51:455-455, 2006; Miyoshi *et al.*, *J Neuroscience* 30:1532-1594, 2010).

[0036] One significant drawback to using rAAVs as a gene-delivery system is the restricted packaging limit of AAVs; this is particularly limiting to the inclusion of lengthy genetic control and expression elements. In addition, many existing interneuron-specific rAAV expression constructs

can provide weak gene expression reducing their usefulness in research and therapeutic uses.

[0037] The current disclosure overcomes drawbacks of the prior art by providing artificial enhancer elements that include a concatemerized core of a I56i enhancer. These artificial enhancer elements provide unexpectedly strong peak transgene expression in forebrain GABAergic interneurons following viral transduction of mouse, monkey, and human brain tissue (see FIGs. 2A, 2B, 3, 4, 5A, 5E, 7, 8A, and 8B). The onset is also surprisingly rapid (see FIGs. 5A-5E), leading to faster and higher expression in direct comparison to virus packaged with, for instance, Addgene plasmid #83900. The increase in expression appears to be synergistically supra-linear and not simply three times the level driven by the original enhancer (FIG. 2B).

[0038] In particular embodiments, the I56i enhancer core can be derived from, for example the human and murine I56i enhancer, or zebrafish I46i enhancer (SEQ ID NOs. 1, 4, and 5 respectively). The selected cores of the I56i enhancer can include SEQ ID NO: 2 (core shared by human and mouse) or SEQ ID NO: 6 (zebrafish I46i core). In particular embodiments, the cores are concatemerized. For example, SEQ ID NO: 3 provides a three-copy concatemer of the selected human/murine I56i core while SEQ ID NO: 7 provides a three-copy concatemer of the selected zebrafish I46i core.

[0039] Of particular interest, the synthetic 3x human/murine core (referred to herein as the 3xhI56iCore; SEQ ID NO: 3) is shorter than the original full-length enhancer sequence reported in Dimidschstein *et al.* (*Nat Neurosci* 19(12):1743-1749, 2016), despite being a 3x concatemer. When used to construct a heterologous expression cassette, such as a recombinant adeno-associated virus (rAAV), this artificial enhancer element provides more room for cargo genes (heterologous encoding sequences) linked to the enhancer. This is highly desirable in many gene expression vectors. For instance, many functioning protein cargo genes (more generally, effector elements) are too long to fit in an AAV vector design, so space (length of sequence) is at a premium in the overall vector.

[0040] The engineered concatemerized I56i cores disclosed herein enable new and improved gene delivery vectors that are particularly useful for achieving selective transgene expression in neocortical GABAergic interneurons in diverse animal species, including humans and non-human primates. Importantly, GABAergic interneurons are highly involved in central processing and development and their dysfunction is implicated in a variety of brain disorders. As such, the herein-described enhancers and expression constructs have many immediate applications in research and clinical treatment development. The artificial enhancers can be used in experimental contexts where the original enhancer hI56i proved insufficient (e.g. retroorbital delivery of virus encoding transgenes for functional perturbation experiments).

[0041] Aspects of the disclosure are now described with the following additional options and detail: (i) Artificial Expression Constructs & Vectors for Selective Expression of Genes in Selected Cell Types; (ii) Compositions for Administration (iii) Cell Lines Including Artificial Expression Constructs; (iv) Transgenic Animals; (v) Methods of Use; (vi) Kits and Commercial Packages; (vii) Exemplary Embodiments; (viii) Experimental Examples; and (ix) Closing Paragraphs.

[0042] (i) Artificial Expression Constructs & Vectors for Selective Expression of Genes in Selected Cell Types. Artificial expression constructs disclosed herein include (i) an enhancer sequence that leads to selective expression of a coding sequence within a targeted central nervous system cell type, (ii) a coding sequence that is expressed, and (iii) a promoter. The expression construct can also include other regulatory elements if necessary or beneficial.

[0043] In particular embodiments, an “enhancer” or an “enhancer element” is a cis-acting sequence that increases the level of transcription associated with a promoter and can function in either orientation relative to the promoter and the coding sequence that is to be transcribed and can be located upstream or downstream relative to the promoter or the coding sequence to be transcribed. There are art-recognized methods and techniques for measuring function(s) of enhancer element sequences. Particular examples of enhancer sequences utilized within artificial expression constructs disclosed herein include concatemerized cores of I56i enhancers, such as concatamerizations of SEQ ID NO: 2 and/or 6 including, as examples, SEQ ID NO: 3 and 7. Additional particular examples of concatemerized cores of I56i enhancers can include SEQ ID NO: 2 and SEQ ID NO: 6 within one sequence such as SEQ ID NO: 2 - SEQ ID NO: 2 - SEQ ID NO: 6; SEQ ID NO: 2 - SEQ ID NO: 6 - SEQ ID NO: 6; SEQ ID NO: 2 - SEQ ID NO: 6 - SEQ ID NO: 2; SEQ ID NO: 6 - SEQ ID NO: 6 - SEQ ID NO: 2; SEQ ID NO: 6 - SEQ ID NO: 2 - SEQ ID NO: 2; and SEQ ID NO: 6 - SEQ ID NO: 2 - SEQ ID NO: 6.

[0044] In particular embodiments, a targeted central nervous system cell type enhancer is an enhancer that is uniquely or predominantly utilized by the targeted central nervous system cell type. A targeted central nervous system cell type enhancer enhances expression of a gene in the targeted central nervous system cell type but does not substantially direct expression of genes in other non-targeted cell types, thus having neural specific transcriptional activity.

[0045] When a coding sequence is selectively expressed in selected neural cells and is not substantially expressed in other neural cell types, the product of the coding sequence is preferentially expressed in the selected cell type. In particular embodiments, preferential expression is greater than 50% expression as compared to a reference cell type; greater than 60% expression as compared to a reference cell type; greater than 70% expression as compared to a reference cell type; greater than 80% expression as compared to a reference cell type; or

greater than 90% expression as compared to a reference cell type. In particular embodiments, a reference cell type refers to non-targeted neural cells. The non-targeted neural cells can be within the same anatomical structure as the targeted cells and/or can project to a common anatomical area. In particular embodiments, a reference cell type is within an anatomical structure that is adjacent to an anatomical structure that includes the targeted cell type. In particular embodiments, a reference cell type is a non-targeted neural cell with a different gene expression profile than the targeted cells.

[0046] In particular embodiments, the product of the coding sequence may be expressed at low levels in non-selected cell types, for example at less than 1% or 1%, 2%, 3%, 5%, 10%, 15% or 20% of the levels at which the product is expressed in selected neural cells. In particular embodiments, the targeted central nervous system cell type is the only cell type that expresses the right combination of transcription factors that bind an enhancer disclosed herein to drive gene expression. Thus, in particular embodiments, expression occurs exclusively within the targeted cell type.

[0047] In particular embodiments, targeted cell types (e.g. neural, neuronal, and/or non-neuronal) can be identified based on transcriptional profiles, such as those described in Tasic et al., 2018 Nature. For reference, the following description of neural cell types and distinguishing features is also provided:

[0048] GABAergic Interneurons: express GABA synthesis genes Gad1/GAD1 and/or Gad2/GAD2.

[0049] GABAergic Subclasses:

Lamp5: Found in many cortical layers, especially upper (L1-L2/3), and have mainly neurogliaform and single bouquet morphology.

Sncg: Found in many cortical layers, and have molecular overlaps with Lamp5 and Vip cells, but inconsistent expression of Lamp5 or Vip, with more consistent expression of Sncg. These neurons express the neurotransmitter Cck and have primarily multipolar or basket cell morphology.

Serpinf1: Found in many cortical layers, and have molecular overlaps with Sncg and Vip cells, but inconsistent expression of Sncg or Vip, with more consistent expression of Serpinf1.

Vip: Found in many cortical layers, but especially frequent in upper layers (L1-L4), and highly express the neurotransmitter vasoactive intestinal peptide (Vip).

Sst: Found in many cortical layers, but especially frequent in lower layers (L5-L6). They highly express the neurotransmitter somatostatin (Sst), and frequently block dendritic inputs to postsynaptic neurons. Included in this subclass are sleep-active horizontal-projecting Sst Chodl (or Sst Nos1) neurons that are highly distinct from other Sst neurons, but express shared marker

genes including Sst.

Pvalb: Found in many cortical layers, but especially frequent in lower layers (L5-L6). They highly express the neurotransmitter parvalbumin (Pvalb), express Tac1, and frequently dampen the output of postsynaptic neurons. Included in this subclass are chandelier cells, which have distinct, chandelier-like morphology and express the markers Cpne5 and Vipr2 in mouse, and NOG and UNC5B in human.

Meis2: A distinct subclass defined by a single type, found in L6b and subcortical white matter.

Lamp5, Sncg, Serpinf1, and VIP: Developmentally derived from precursor neurons in the caudal ganglionic eminence (CGE).

Sst and Pvalb: Developmentally derived from precursor neurons in the medial ganglionic eminence (MGE).

[0050] Glutamatergic Subclasses:

All: Express glutamate transmitters Slc17a6 and/or Slc17a7.

L2/3 IT: Primarily reside in Layer 2/3 and have mainly intratelencephalic (cortico-cortical) projections.

L4 IT: Primarily reside in Layer 4 and have mainly intratelencephalic (cortico-cortical) projections.

L5 IT: Primarily reside in Layer 5 and have mainly intratelencephalic (cortico-cortical) projections. Also called L5a.

L5 PT: Primarily reside in Layer 5 and have mainly cortico-subcortical (pyramidal tract or corticofugal) projections. Also called L5b or L5 CF. These cells are located in primary motor cortex and neighboring areas and are corticospinal projection neurons. They are associated with motor neuron/movement disorders, such as ALS.

Neocortical L5 extratelencephalic (ET)-projecting pyramidal neurons (L5 ET): thick-tufted pyramidal neurons, including distinctive subtypes found only in specialized regions, e.g. Betz cells, Meynert cells, and von Economo cells.

L5 NP: Primarily reside in Layer 5 and have mainly nearby projections.

L6 CT: Primarily reside in Layer 6 and have mainly cortico-thalamic projections.

L6 IT: Primarily reside in Layer 6 and have mainly intratelencephalic (cortico-cortical) projections. Included in this subclass are L6 IT Car3 cells, which are highly similar to intracortical-projecting cells in the claustrum.

L6b: Primarily reside in the cortical subplate (L6b), with observed projections to local regions (near the cell body), cortico-cortical projections from VISp to anterior cingulate, and cortico-subcortical projections to the thalamus.

CR: A distinct subclass defined by a single type in L1, Cajal-Retzius cells express distinct

molecular markers Lhx5 and Trp73.

[0051] Non-neuronal Subclasses:

Astrocytes: Neuroectoderm-derived glial cells which express the marker Aqp4. They have a distinct star-shaped morphology and are involved in metabolic support of other cells in the brain.

Oligodendrocytes: Neuroectoderm-derived glial cells, which express the marker Sox10. This category includes oligodendrocyte precursor cells (OPCs). Oligodendrocytes are the subclass that is primarily responsible for myelination of neurons.

VLMCs: Vascular leptomeningeal cells (VLMCs) are part of the meninges that surround the outer layer of the cortex and express the marker genes Lum and Col1a1.

Pericytes: Blood vessel-associated cells, also called mural cells, that express the marker genes Kcnj8 and Abcc9. Pericytes wrap around endothelial cells and are important for regulation of capillary blood flow and are involved in blood-brain barrier permeability.

SMCs: Blood vessel-associated cells, also called mural cells, that express the marker gene Acta2. SMCs cover arterioles in the brain and are involved in blood-brain barrier permeability.

Endothelial: Cells that line blood vessels of the brain. Endothelial cells express the markers Tek and PDGF- β .

Macrophages: Immune cells, including macrophages, which are brain-resident macrophages, and perivascular macrophages (PVMs) that may be transitionally associated with brain tissue, or included as a byproduct of brain dissection methods.

[0052] In particular embodiments, a coding sequence is a heterologous coding sequence that encodes an effector element. An effector element is a sequence that is expressed to achieve, and that in fact achieves, an intended effect. Examples of effector elements include reporter genes/proteins and functional genes/proteins.

[0053] Exemplary reporter genes/proteins include those expressed by Addgene ID#s 83894 (pAAV-hDlx-Flex-dTomato-Fishell_7), 83895 (pAAV-hDlx-Flex-GFP-Fishell_6), 83896 (pAAV-hDlx-GiDREADD-dTomato-Fishell-5), 83898 (pAAV-mDlx-ChR2-mCherry-Fishell-3), 83899 (pAAV-mDlx-GCaMP6f-Fishell-2), 83900 (pAAV-mDlx-GFP-Fishell-1), and 89897 (pcDNA3-FLAG-mTET2 (N500)). Exemplary reporter genes particularly can include those which encode an expressible fluorescent protein, or expressible biotin; blue fluorescent proteins (e.g. eBFP, eBFP2, Azurite, mKalama1, GFPuv, Sapphire, T-sapphire); cyan fluorescent proteins (e.g. eCFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan, mTurquoise); green fluorescent proteins (e.g. GFP, GFP-2, tagGFP, turboGFP, EGFP, Emerald, Azami Green, Monomeric Azami Green (mAzamigreen), CopGFP, AceGFP, avGFP, ZsGreen1, Oregon Green™ (Thermo Fisher Scientific)); Luciferase; orange fluorescent proteins (mOrange, mKO, Kusabira-Orange,

Monomeric Kusabira-Orange, mTangerine, tdTomato, dTomato); red fluorescent proteins (mKate, mKate2, mPlum, DsRed monomer, mCherry, mRuby, mRFP1, DsRed-Express, DsRed2, DsRed-Monomer, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred, Texas Red™ (Thermo Fisher Scientific)); far red fluorescent proteins (e.g., mPlum and mNeptune); yellow fluorescent proteins (e.g., YFP, eYFP, Citrine, SYFP2, Venus, YPet, PhiYFP, ZsYellow1); and tandem conjugates.

[0054] GFP is composed of 238 amino acids (26.9 kDa), originally isolated from the jellyfish *Aequorea victoria*/*Aequorea aequorea*/*Aequorea forskalea* that fluoresces green when exposed to blue light. The GFP from *A. victoria* has a major excitation peak at a wavelength of 395 nm and a minor one at 475 nm. Its emission peak is at 509 nm which is in the lower green portion of the visible spectrum. The GFP from the sea pansy (*Renilla reniformis*) has a single major excitation peak at 498 nm. Due to the potential for widespread usage and the evolving needs of researchers, many different mutants of GFP have been engineered. The first major improvement was a single point mutation (S65T) reported in 1995 in *Nature* by Roger Tsien. This mutation dramatically improved the spectral characteristics of GFP, resulting in increased fluorescence, photostability and a shift of the major excitation peak to 488 nm with the peak emission kept at 509 nm. The addition of the 37°C folding efficiency (F64L) point mutant to this scaffold yielded enhanced GFP (EGFP). EGFP has an extinction coefficient (denoted ϵ), also known as its optical cross section of 9.13×10^{-21} m²/molecule, also quoted as 55,000 L/(mol•cm). Superfolder GFP, a series of mutations that allow GFP to rapidly fold and mature even when fused to poorly folding peptides, was reported in 2006.

[0055] The "yellow fluorescent protein" (YFP) is a genetic mutant of green fluorescent protein, derived from *Aequorea victoria*. Its excitation peak is 514 nm and its emission peak is 527 nm.

[0056] Exemplary functional molecules include functioning ion transporters, cellular trafficking proteins, enzymes, transcription factors, neurotransmitters, calcium reporters, channel rhodopsins, guide RNA, nucleases, or designer receptors exclusively activated by designer drugs (DREADDs).

[0057] Ion transporters are transmembrane proteins that mediate transport of ions across cell membranes. These transporters are pervasive throughout most cell types and important for regulating cellular excitability and homeostasis. Ion transporters participate in numerous cellular processes such as action potentials, synaptic transmission, hormone secretion, and muscle contraction. Many important biological processes in living cells involve the translocation of cations, such as calcium (Ca²⁺), potassium (K⁺), and sodium (Na⁺) ions, through such ion channels. In particular embodiments, ion transporters include voltage gated sodium channels (e.g., SCN1A),

potassium channels (e.g., KCNQ2), and calcium channels (e.g. CACNA1C)).

[0058] Exemplary enzymes, transcription factors, receptors, membrane proteins, cellular trafficking proteins, signaling molecules, and neurotransmitters include enzymes such as lactase, lipase, helicase, alpha-glucosidase, amylase; transcription factors such as SP1, AP-1, Heat shock factor protein 1, C/EBP (CCAA-T/enhancer binding protein), and Oct-1; receptors such as transforming growth factor receptor beta 1, platelet-derived growth factor receptor, epidermal growth factor receptor, vascular endothelial growth factor receptor, and interleukin 8 receptor alpha; membrane proteins, cellular trafficking proteins such as clathrin, dynamin, caveolin, Rab-4A, and Rab-11A; signaling molecules such as nerve growth factor (NGF), platelet-derived growth factor (PDGF), transforming growth factor β (TGF β), epidermal growth factor (EGF), GTPase and HRas; and neurotransmitters such as cocaine and amphetamine regulated transcript, substance P, oxytocin, and somatostatin.

[0059] In particular embodiments, functional molecules include reporters of neural function and states such as calcium reporters. Intracellular calcium concentration is an important predictor of numerous cellular activities, which include neuronal activation, muscle cell contraction and second messenger signaling. A sensitive and convenient technique to monitor the intracellular calcium levels is through the genetically encoded calcium indicator (GECI). Among the GECIs, green fluorescent protein (GFP) based calcium sensors named GCaMPs are efficient and widely used tools. The GCaMPs are formed by fusion of M13 and calmodulin protein to N- and C-termini of circularly permuted GFP. Some GCaMPs yield distinct fluorescence emission spectra (Zhao et al., Science, 2011, 333(6051): 1888-1891). Exemplary GECIs with green fluorescence include GCaMP3, GCaMP5G, GCaMP6s, GCaMP6m, GCaMP6f, jGCaMP7s, jGCaMP7c, jGCaMP7b, and jGCaMP7f. Furthermore, GECIs with red fluorescence include jRGECO1a and jRGECO1b. AAV products containing GECIs are commercially available. For example, Vigene Biosciences provides AAV products including AAV8-CAG-GCaMP3 (Cat. No:BS4-CX3AAV8), AAV8-Syn-FLEX-GCaMP6s-WPRE (Cat. No:BS1-NXSAAV8), AAV8-Syn-FLEX-GCaMP6s-WPRE (Cat. No:BS1-NXSAAV8), AAV9-CAG-FLEX-GCaMP6m-WPRE (Cat. No:BS2-CXMAAV9), AAV9-Syn-FLEX-jGCaMP7s-WPRE (Cat. No:BS12-NXSAAV9), AAV9-CAG-FLEX-jGCaMP7f-WPRE (Cat. No:BS12-CXFAAV9), AAV9-Syn-FLEX-jGCaMP7b-WPRE (Cat. No:BS12-NXBAAV9), AAV9-Syn-FLEX-jGCaMP7c-WPRE (Cat. No:BS12-NXCAAV9), AAV9-Syn-FLEX-NES-jRGECO1a-WPRE (Cat. No:BS8-NXAAV9), and AAV8-Syn-FLEX-NES-jRCaMP1b-WPRE (Cat. No:BS7-NXBAAV8).

[0060] In particular embodiments calcium reporters include the genetically encoded calcium indicators GECI, NTnC; Myosin light chain kinase, GFP, Calmodulin chimera; Calcium indicator

TN-XXL; BRET-based auto-luminescent calcium indicator; and/or Calcium indicator protein OeNL(Ca²⁺)-18u).

[0061] In particular embodiments, functional molecules include modulators of neuronal activity like channel rhodopsins (e.g., channelrhodopsin-1, channelrhodopsin-2, and variants thereof). Channelrhodopsins are a subfamily of retinylidene proteins (rhodopsins) that function as light-gated ion channels. In addition to channelrhodopsin 1 (ChR1) and channelrhodopsin 2 (ChR2), several variants of channelrhodopsins have been developed. For example, Lin et al. (*Biophys J*, 2009, 96(5): 1803-14) describe making chimeras of the transmembrane domains of ChR1 and ChR2, combined with site-directed mutagenesis. Zhang et al. (*Nat Neurosci*, 2008, 11(6): 631-3) describe VChR1, which is a red-shifted channelrhodopsin variant. VChR1 has lower light sensitivity and poor membrane trafficking and expression. Other known channelrhodopsin variants include the ChR2 variant described in Nagel, et al., *Proc Natl Acad Sci USA*, 2003, 100(24): 13940-5, ChR2/H134R (Nagel, G., et al., *Curr Biol*, 2005, 15(24): 2279-84), and ChD/ChEF/ChIEF (Lin, J. Y., et al., *Biophys J*, 2009, 96(5): 1803-14), which are activated by blue light (470 nm) but show no sensitivity to orange/red light. Additional variants are described in Lin, *Experimental Physiology*, 2010, 96.1: 19-25 and Knopfel et al., *The Journal of Neuroscience*, 2010, 30(45): 14998-15004).

[0062] In particular embodiments, functional molecules include DNA and RNA editing tools such CRISPR/CAS (e.g., guide RNA and a nuclease, such as Cas, Cas9 or cpf1). Functional molecules can also include engineered Cpf1s such as those described in US 2018/0030425, US 2016/0208243, WO/2017/184768 and Zetsche et al. (2015) *Cell* 163: 759-771; single gRNA (see e.g., Jinek et al. (2012) *Science* 337:816-821; Jinek et al. (2013) *eLife* 2:e00471; Segal (2013) *eLife* 2:e00563) or editase, guide RNA molecules or homologous recombination donor cassettes.

[0063] Additional information regarding CRISPR-Cas systems and components thereof are described in, US8697359, US8771945, US8795965, US8865406, US8871445, US8889356, US8889418, US8895308, US8906616, US8932814, US8945839, US8993233 and US8999641 and applications related thereto; and WO2014/018423, WO2014/093595, WO2014/093622, WO2014/093635, WO2014/093655, WO2014/093661, WO2014/093694, WO2014/093701, WO2014/093709, WO2014/093712, WO2014/093718, WO2014/145599, WO2014/204723, WO2014/204724, WO2014/204725, WO2014/204726, WO2014/204727, WO2014/204728, WO2014/204729, WO2015/065964, WO2015/089351, WO2015/089354, WO2015/089364, WO2015/089419, WO2015/089427, WO2015/089462, WO2015/089465, WO2015/089473 and WO2015/089486, WO2016205711, WO2017/106657, WO2017/127807 and applications related thereto.

[0064] In particular embodiments, functional molecules include designer receptor exclusively activated by designer drug (DREADD). Designer receptors exclusively activated by designer drugs (DREADDs) can be used to modulate cellular functions (Rogan and Roth, *Pharmacol. Rev.* 2011, 63(2): 291-315). This family of evolved muscarinic receptors has been shown to increase (Gs-DREADD; Gq-DREADD) or decrease (Gi/o-DREADD) cellular activity following administration of an otherwise inert synthetic ligand, clozapine-n-oxide (Armbruster et al., *PNAS*, 2007, 104(12): 5163-5168). When packaged into viral vectors or expressed in transgenic mouse models, these tools allow cellular activity to be controlled in a defined spatial and temporal manner. For example, activation of hippocampal neurons by Gq-DREADD receptors amplifies γ -rhythms and increases locomotor activity and stereotypy in mice (Alexander et al., *Neuron*, 2009, 63(1): 27-39). DREADDs are formed by point mutations in the third and fifth transmembrane regions of muscarinic receptors (Y149C and A239G in hM3). In addition, the Gs-coupled DREADD contains the second and third intracellular loops of the β_1 -AR in place of those of the M3 muscarinic receptor. Some exemplary DREADDs include hM3DREADD (hM3D) and hM4DREADD (hM4D). Various plasmids containing DREADDs are commercially available. For example, on addgene, AAV plasmids containing DREADDs include: pAAV-hSyn-DIO-hM3D(Gq)-mCherry (Plasmid #44361), pAAV-hSyn-DIO-hM4d(Gi)-mCherry (Plasmid #44362), pAAV-EF1a-DIO-hM4d(Gi)-mCherry (Plasmid #50461), pAAV-GFAP-HA-hM3D(Gq)-IRES-mCitrine (Plasmid #50470), and pAAV-CaMKIIa-hM4D(Gi)-mCherry (Plasmid #50477).

[0065] Additional effector elements include Cre, iCre, dgCre, FlpO, and tTA2. iCre refers to a codon-improved Cre. dgCre refers to an enhanced GFP/Cre recombinase fusion gene with an N terminal fusion of the first 159 amino acids of the Escherichia coli K-12 strain chromosomal dihydrofolate reductase gene (DHFR or folA) harboring a G67S mutation and modified to also include the R12Y/Y100I destabilizing domain mutation. FlpO refers to a codon-optimized form of FLPe that greatly increases protein expression and FRT recombination efficiency in mouse cells. Like the Cre/LoxP system, the FLP/FRT system has been widely used for gene expression (and generating conditional knockout mice, mediated by the FLP/FRT system). tTA2 refers to tetracycline transactivator.

[0066] Exemplary expressible elements are expression products that do not include effector elements, for example, a non-functioning or defective protein. In particular embodiments, expressible elements can provide methods to study the effects of their functioning counterparts. In particular embodiments, expressible elements are non-functioning or defective based on an engineered mutation that renders them non-functioning. In these aspects, non-expressible elements are as similar in structure as possible to their functioning counterparts.

[0067] Exemplary self-cleaving peptides include the 2A peptides which lead to the production of two proteins from one mRNA. The 2A sequences are short (e.g., 20 amino acids), allowing more use in size-limited constructs. Particular examples include P2A, T2A, E2A, and F2A. In particular embodiments, the expression constructs include an internal ribosome entry site (IRES) sequence. IRES allow ribosomes to initiate translation at a second internal site on a mRNA molecule, leading to production of two proteins from one mRNA.

[0068] Coding sequences encoding molecules (e.g., RNA, proteins) described herein can be readily obtained from publicly available databases and publications. Coding sequences can further include various sequence polymorphisms, mutations, and/or sequence variants wherein such alterations do not affect the function of the encoded molecule. The term “encode” or “encoding” refers to a property of sequences of nucleic acids, such as a vector, a plasmid, a gene, cDNA, mRNA, to serve as templates for synthesis of other molecules such as proteins.

[0069] The term “gene” may include not only coding sequences but also regulatory regions such as promoters, enhancers, and termination regions. The term further can include all introns and other DNA sequences spliced from the mRNA transcript, along with variants resulting from alternative splice sites. The sequences can also include degenerate codons of a reference sequence or sequences that may be introduced to provide codon preference in a specific organism or cell type.

[0070] Promoters can include general promoters, tissue-specific promoters, cell-specific promoters, and/or promoters specific for the cytoplasm. Promoters may include strong promoters, weak promoters, constitutive expression promoters, and/or inducible promoters. Inducible promoters direct expression in response to certain conditions, signals or cellular events. For example, the promoter may be an inducible promoter that requires a particular ligand, small molecule, transcription factor or hormone protein in order to effect transcription from the promoter. Particular examples of promoters include minBglobin, CMV, minCMV, a mutated minCMV, SV40 immediately early promoter, the Hsp68 minimal promoter (proHSP68), and the Rous Sarcoma Virus (RSV) long-terminal repeat (LTR) promoter. Minimal promoters have no activity to drive gene expression on their own, but can be activated to drive gene expression when linked to a proximal enhancer element.

[0071] In particular embodiments, expression constructs are provided within vectors. The term vector refers to a nucleic acid molecule capable of transferring or transporting another nucleic acid molecule, such as an expression construct. The transferred nucleic acid is generally linked to, e.g., inserted into, the vector nucleic acid molecule. A vector may include sequences that direct autonomous replication in a cell or may include sequences that permit integration into host cell

DNA. Useful vectors include, for example, plasmids (*e.g.*, DNA plasmids or RNA plasmids), transposons, cosmids, bacterial artificial chromosomes, and viral vectors.

[0072] Viral vector is widely used to refer to a nucleic acid molecule that includes virus-derived nucleic acid elements that facilitate transfer and expression of non-native nucleic acid molecules within a cell. The term adeno-associated viral vector refers to a viral vector or plasmid containing structural and functional genetic elements, or portions thereof, that are primarily derived from AAV. The term "retroviral vector" refers to a viral vector or plasmid containing structural and functional genetic elements, or portions thereof, that are primarily derived from a retrovirus. The term "lentiviral vector" refers to a viral vector or plasmid containing structural and functional genetic elements, or portions thereof, that are primarily derived from a lentivirus, and so on. The term "hybrid vector" refers to a vector including structural and/or functional genetic elements from more than one virus type.

[0073] Adenovirus. "Adenovirus vectors" refer to those constructs containing adenovirus sequences sufficient to (a) support packaging of an expression construct and (b) to express a coding sequence that has been cloned therein in a sense or antisense orientation. A recombinant Adenovirus vector includes a genetically engineered form of an adenovirus. Knowledge of the genetic organization of adenovirus, a 36 kb, linear, double-stranded DNA virus, allows substitution of large pieces of adenoviral DNA with foreign sequences up to 7 kb. In contrast to retrovirus, the adenoviral infection of host cells does not result in chromosomal integration because adenoviral DNA can replicate in an episomal manner without potential genotoxicity. Also, adenoviruses are structurally stable, and no genome rearrangement has been detected after extensive amplification.

[0074] Adenovirus is particularly suitable for use as a gene transfer vector because of its mid-sized genome, ease of manipulation, high titer, wide target-cell range, and high infectivity. Both ends of the viral genome contain 100-200 base pair inverted repeats (ITRs), which are cis elements necessary for viral DNA replication and packaging. The early (E) and late (L) regions of the genome contain different transcription units that are divided by the onset of viral DNA replication. The E1 region (E1A and E1B) encodes proteins responsible for the regulation of transcription of the viral genome and a few cellular genes. The expression of the E2 region (E2A and E2B) results in the synthesis of the proteins for viral DNA replication. These proteins are involved in DNA replication, late gene expression, and host cell shut-off. The products of the late genes, including the majority of the viral capsid proteins, are expressed only after significant processing of a single primary transcript issued by the major late promoter (MLP). The MLP is particularly efficient during the late phase of infection, and all the mRNAs issued from this

promoter possess a 5'-tripartite leader (TPL) sequence which makes them preferred mRNAs for translation.

[0075] Other than the requirement that an adenovirus vector be replication defective, or at least conditionally defective, the nature of the adenovirus vector is not believed to be crucial to the successful practice of particular embodiments disclosed herein. The adenovirus may be of any of the 42 different known serotypes or subgroups A-F. In particular embodiments, adenovirus type 5 of subgroup C is the preferred starting material in order to obtain a conditional replication-defective adenovirus vector for use in particular embodiments, since Adenovirus type 5 is a human adenovirus about which a great deal of biochemical and genetic information is known, and it has historically been used for most constructions employing adenovirus as a vector.

[0076] As indicated, the typical vector is replication defective and will not have an adenovirus E1 region. Thus, it will be most convenient to introduce the polynucleotide encoding the gene of interest at the position from which the E1-coding sequences have been removed. However, the position of insertion of the construct within the adenovirus sequences is not critical. The polynucleotide encoding the gene of interest may also be inserted in lieu of a deleted E3 region in E3 replacement vectors or in the E4 region where a helper cell line or helper virus complements the E4 defect.

[0077] Adeno-Associated Virus (AAV) is a parvovirus, discovered as a contamination of adenoviral stocks. It is a ubiquitous virus (antibodies are present in 85% of the US human population) that has not been linked to any disease. It is also classified as a dependovirus, because its replication is dependent on the presence of a helper virus, such as adenovirus. Various serotypes have been isolated, of which AAV-2 is the best characterized. AAV has a single-stranded linear DNA that is encapsidated into capsid proteins VP1, VP2 and VP3 to form an icosahedral virion of 20 to 24 nm in diameter.

[0078] The AAV DNA is 4.7 kilobases long. It contains two open reading frames and is flanked by two ITRs. There are two major genes in the AAV genome: rep and cap. The rep gene codes for proteins responsible for viral replications, whereas cap codes for capsid protein VP1-3. Each ITR forms a T-shaped hairpin structure. These terminal repeats are the only essential cis components of the AAV for chromosomal integration. Therefore, the AAV can be used as a vector with all viral coding sequences removed and replaced by the cassette of genes for delivery. Three AAV viral promoters have been identified and named p5, p19, and p40, according to their map position. Transcription from p5 and p19 results in production of rep proteins, and transcription from p40 produces the capsid proteins.

[0079] AAVs stand out for use within the current disclosure because of their superb safety profile

and because their capsids and genomes can be tailored to allow expression in selected cell populations. scAAV refers to a self-complementary AAV. pAAV refers to a plasmid adeno-associated virus. rAAV refers to a recombinant adeno-associated virus.

[0080] Other viral vectors may also be employed. For example, vectors derived from viruses such as vaccinia virus, polioviruses and herpes viruses may be employed. They offer several attractive features for various mammalian cells.

[0081] Retrovirus. Retroviruses are a common tool for gene delivery. "Retrovirus" refers to an RNA virus that reverse transcribes its genomic RNA into a linear double-stranded DNA copy and subsequently covalently integrates its genomic DNA into a host genome. Once the virus is integrated into the host genome, it is referred to as a "provirus." The provirus serves as a template for RNA polymerase II and directs the expression of RNA molecules which encode the structural proteins and enzymes needed to produce new viral particles.

[0082] Illustrative retroviruses suitable for use in particular embodiments, include: Moloney murine leukemia virus (M-MuLV), Moloney murine sarcoma virus (MoMSV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV), gibbon ape leukemia virus (GaLV), feline leukemia virus (FLV), spumavirus, Friend murine leukemia virus, Murine Stem Cell Virus (MSCV) and Rous Sarcoma Virus (RSV) and lentivirus.

[0083] "Lentivirus" refers to a group (or genus) of complex retroviruses. Illustrative lentiviruses include: HIV (human immunodeficiency virus; including HIV type 1, and HIV type 2); visna-maedi virus (VMV); the caprine arthritis-encephalitis virus (CAEV); equine infectious anemia virus (EIAV); feline immunodeficiency virus (FIV); bovine immune deficiency virus (BIV); and simian immunodeficiency virus (SIV). In particular embodiments, HIV based vector backbones (i.e., HIV cis-acting sequence elements) can be used.

[0084] A safety enhancement for the use of some vectors can be provided by replacing the U3 region of the 5' LTR with a heterologous promoter to drive transcription of the viral genome during production of viral particles. Examples of heterologous promoters which can be used for this purpose include, for example, viral simian virus 40 (SV40) (e.g., early or late), cytomegalovirus (CMV) (e.g., immediate early), Moloney murine leukemia virus (MoMLV), Rous sarcoma virus (RSV), and herpes simplex virus (HSV) (thymidine kinase) promoters. Typical promoters are able to drive high levels of transcription in a Tat-independent manner. This replacement reduces the possibility of recombination to generate replication-competent virus because there is no complete U3 sequence in the virus production system. In particular embodiments, the heterologous promoter has additional advantages in controlling the manner in which the viral genome is transcribed. For example, the heterologous promoter can be inducible, such that transcription of

all or part of the viral genome will occur only when the induction factors are present. Induction factors include one or more chemical compounds or the physiological conditions such as temperature or pH, in which the host cells are cultured.

[0085] In particular embodiments, viral vectors include a TAR element. The term "TAR" refers to the "trans-activation response" genetic element located in the R region of lentiviral LTRs. This element interacts with the lentiviral trans-activator (tat) genetic element to enhance viral replication. However, this element is not required in embodiments wherein the U3 region of the 5' LTR is replaced by a heterologous promoter.

[0086] The "R region" refers to the region within retroviral LTRs beginning at the start of the capping group (i.e., the start of transcription) and ending immediately prior to the start of the poly(A) tract. The R region is also defined as being flanked by the U3 and U5 regions. The R region plays a role during reverse transcription in permitting the transfer of nascent DNA from one end of the genome to the other.

[0087] In particular embodiments, expression of heterologous sequences in viral vectors is increased by incorporating posttranscriptional regulatory elements, efficient polyadenylation sites, and optionally, transcription termination signals into the vectors. A variety of posttranscriptional regulatory elements can increase expression of a heterologous nucleic acid. Examples include the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE; Zufferey *et al.*, 1999, J. Virol., 73:2886); the posttranscriptional regulatory element present in hepatitis B virus (HPRE) (Smith *et al.*, Nucleic Acids Res. 26(21):4818-4827, 1998); and the like (Liu *et al.*, 1995, Genes Dev., 9:1766). In particular embodiments, vectors include a posttranscriptional regulatory element such as a WPRE or HPRE. In particular embodiments, vectors lack or do not include a posttranscriptional regulatory element such as a WPRE or HPRE.

[0088] Elements directing the efficient termination and polyadenylation of a heterologous nucleic acid transcript can increase heterologous gene expression. Transcription termination signals are generally found downstream of the polyadenylation signal. In particular embodiments, vectors include a polyadenylation sequence 3' of a polynucleotide encoding a molecule (e.g., protein) to be expressed. The term "poly(A) site" or "poly(A) sequence" denotes a DNA sequence which directs both the termination and polyadenylation of the nascent RNA transcript by RNA polymerase II. Polyadenylation sequences can promote mRNA stability by addition of a poly(A) tail to the 3' end of the coding sequence and thus, contribute to increased translational efficiency. Particular embodiments may utilize BGHpA or SV40pA. In particular embodiments, a preferred embodiment of an expression construct includes a terminator element. These elements can serve to enhance transcript levels and to minimize read through from the construct into other plasmid

sequences.

[0089] In particular embodiments, a viral vector further includes one or more insulator elements. Insulators elements may contribute to protecting viral vector-expressed sequences, e.g., effector elements or expressible elements, from integration site effects, which may be mediated by cis-acting elements present in genomic DNA and lead to deregulated expression of transferred sequences (*i.e.*, position effect; see, e.g., Burgess-Beusse *et al.*, *PNAS*, USA, 99:16433, 2002; and Zhan *et al.*, *Hum. Genet.*, 109:471, 2001). In particular embodiments, viral transfer vectors include one or more insulator elements at the 3' LTR and upon integration of the provirus into the host genome, the provirus includes the one or more insulators at both the 5' LTR and 3' LTR, by virtue of duplicating the 3' LTR. Suitable insulators for use in particular embodiments include the chicken β -globin insulator (see Chung *et al.*, *Cell* 74:505, 1993; Chung *et al.*, *PNAS USA* 94:575, 1997; and Bell *et al.*, *Cell* 98:387, 1999), SP10 insulator (Abhyankar *et al.*, *JBC* 282:36143, 2007), or other small CTCF recognition sequences that function as enhancer blocking insulators (Liu *et al.*, *Nature Biotechnology*, 33:198, 2015).

[0090] Beyond the foregoing description, a wide range of suitable expression vector types will be known to a person of ordinary skill in the art. These can include commercially available expression vectors designed for general recombinant procedures, for example plasmids that contain one or more reporter genes and regulatory elements required for expression of the reporter gene in cells. Numerous vectors are commercially available, e.g., from Invitrogen, Stratagene, Clontech, etc., and are described in numerous associated guides. In particular embodiments, suitable expression vectors include any plasmid, cosmid or phage construct that is capable of supporting expression of encoded genes in mammalian cell, such as pUC or Bluescript plasmid series.

[0091] Particular embodiments of vectors disclosed herein include:

Expression Construct Name	Features
CN1390	rAAV: 3xhI56Core-minBglobin-SYFP2-WPRE3-BGHpA
CN1244	rAAV: hI56i-minBglobin-SYFP2-WPRE3-BGHpA
CN1389	rAAV: hI56Core-minBglobin-SYFP2-WPRE3-BGHpA
CN1203	scAAV: hI56i-minBglobin-SYFP2-WPRE3-BGHpA
CN1367	rAAV: hI56i-minBglobin-His-NavMs-P2A-SYFP2-WPRE3-BGHpA
CN1498	rAAV: 3xhI56iCore-minCMV-SYFP2-P2A-3xHA-NavBp-WPRE3-BGHpA
CN1499	rAAV: 3xhI56iCore-minCMV-SYFP2-P2A-3xHA-NavMs-WPRE3-BGHpA
CN1500	rAAV:3xhI56iCore-minCMV-SYFP2-P2A-3xHA-NavSheP-D60N-WPRE3-BGHpA
CN1838	scAAV-3XzI46i-minBG-SYFP2-WPRE3-SPA

[0092] In particular embodiments, SYFP2 within CN1390, CN1244, CN1389, CN1203, CN1367, CN1498, CN1499, CN1500, and CN1838 can be replaced with a channel rhodopsin or calcium

reporter, such as ChR2 or GCaMP. In particular embodiments, SYFP2 within CN1390 is replaced with ChR2 or GCaMP. 3Xzl46i within CN1838 refers to the 3X concatemer of the zebrafish l46iCore. See also FIG. 16 which provides additional exemplary vector components and combinations of the disclosure.

[0093] In particular embodiments vectors (e.g., AAV) with capsids that cross the blood-brain barrier (BBB) are selected. In particular embodiments, vectors are modified to include capsids that cross the BBB. Examples of AAV with viral capsids that cross the blood brain barrier include AAV9 (Gombash et al., *Front Mol Neurosci.* 2014; 7:81), AAVrh.10 (Yang, et al., *Mol Ther.* 2014; 22(7): 1299-1309), AAV1R6, AAV1R7 (Albright et al., *Mol Ther.* 2018; 26(2): 510), rAAVrh.8 (Yang, et al., *supra*), AAV-BR1 (Marchio et al., *EMBO Mol Med.* 2016; 8(6): 592), AAV-PHP.S (Chan et al., *Nat Neurosci.* 2017; 20(8): 1172), AAV-PHP.B (Deverman et al., *Nat Biotechnol.* 2016; 34(2): 204), and AAV-PPS (Chen et al., *Nat Med.* 2009; 15: 1215). The PHP.eB capsid differs from AAV9 such that, using AAV9 as a reference, amino acids starting at residue 586: S-AQ-A (SEQ ID NO: 98) are changed to S-DGTLAVPFK-A (SEQ ID NO: 33).

[0094] AAV9 is a naturally occurring AAV serotype that, unlike many other naturally occurring serotypes, can cross the BBB following intravenous injection. It transduces large sections of the central nervous system (CNS), thus permitting minimally invasive treatments (Naso et al., *BioDrugs.* 2017; 31(4): 317), for example, as described in relation to clinical trials for the treatment of superior mesenteric artery (SMA) syndrome by AveXis (AVXS-101, NCT03505099) and the treatment of CLN3 gene-Related Neuronal Ceroid-Lipofuscinosis (NCT03770572).

[0095] AAVrh.10, was originally isolated from rhesus macaques and shows low seropositivity in humans when compared with other common serotypes used for gene delivery applications (Selot et al., *Front Pharmacol.* 2017; 8: 441) and has been evaluated in clinical trials LYS-SAF302, LYSOGENE, and NCT03612869.

[0096] AAV1R6 and AAV1R7, two variants isolated from a library of chimeric AAV vectors (AAV1 capsid domains swapped into AAVrh.10), retain the ability to cross the BBB and transduce the CNS while showing significantly reduced hepatic and vascular endothelial transduction.

[0097] rAAVrh.8, also isolated from rhesus macaques, shows a global transduction of glial and neuronal cell types in regions of clinical importance following peripheral administration and also displays reduced peripheral tissue tropism compared to other vectors.

[0098] AAV-BR1 is an AAV2 variant displaying the NRGTEWD (SEQ ID NO: 91) epitope that was isolated during in vivo screening of a random AAV display peptide library. It shows high specificity accompanied by high transgene expression in the brain with minimal off-target affinity (including for the liver) (Körbelin et al., *EMBO Mol Med.* 2016; 8(6): 609).

[0099] AAV-PHP.S (Addgene, Watertown, MA) is a variant of AAV9 generated with the CREATE method that encodes the 7-mer sequence QAVRTSL (SEQ ID NO: 92), transduces neurons in the enteric nervous system, and strongly transduces peripheral sensory afferents entering the spinal cord and brain stem.

[0100] AAV-PHP.B (Addgene, Watertown, MA) is a variant of AAV9 generated with the CREATE method that encodes the 7-mer sequence TLAVPFK (SEQ ID NO: 93). It transfers genes throughout the CNS with higher efficiency than AAV9 and transduces the majority of astrocytes and neurons across multiple CNS regions.

[0101] AAV-PPS, an AAV2 variant created by insertion of the DSPAHPS (SEQ ID NO: 94) epitope into the capsid of AAV2, shows a dramatically improved brain tropism relative to AAV2.

[0102] For additional information regarding capsids that cross the blood brain barrier, see Chan et al., Nat. Neurosci. 2017 Aug; 20(8): 1172-1179.

[0103] (ii) Compositions for Administration. Artificial expression constructs and vectors of the present disclosure (referred to herein as physiologically active components) can be formulated with a carrier that is suitable for administration to a cell, tissue slice, animal (e.g., mouse, non-human primate), or human. Physiologically active components within compositions described herein can be prepared in neutral forms, as freebases, or as pharmacologically acceptable salts.

[0104] Pharmaceutically-acceptable salts include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like.

[0105] Carriers of physiologically active components can include solvents, dispersion media, vehicles, coatings, diluents, isotonic and absorption delaying agents, buffers, solutions, suspensions, colloids, and the like. The use of such carriers for physiologically active components is well known in the art. Except insofar as any conventional media or agent is incompatible with the physiologically active components, it can be used with compositions as described herein.

[0106] The phrase "pharmaceutically-acceptable carriers" refer to carriers that do not produce an allergic or similar untoward reaction when administered to a human, and in particular embodiments, when administered intravenously (e.g. at the retro-orbital plexus).

[0107] In particular embodiments, compositions can be formulated for intravenous, intraocular, intravitreal, parenteral, subcutaneous, intracerebro-ventricular, intramuscular, intrathecal, intraspinal, oral, intraperitoneal, oral or nasal inhalation, or by direct injection in or application to

one or more cells, tissues, or organs.

[0108] Compositions may include liposomes, lipids, lipid complexes, microspheres, microparticles, nanospheres, and/or nanoparticles.

[0109] The formation and use of liposomes is generally known to those of skill in the art. Liposomes have been developed with improved serum stability and circulation half-times (see, for instance, U.S. Pat. No. 5,741,516). Further, various methods of liposome and liposome like preparations as potential drug carriers have been described (see, for instance U.S. Pat. Nos. 5,567,434; 5,552,157; 5,565,213; 5,738,868; and 5,795,587).

[0110] The disclosure also provides for pharmaceutically acceptable nanocapsule formulations of the physiologically active components. Nanocapsules can generally entrap compounds in a stable and reproducible way (Quintanar-Guerrero *et al.*, *Drug Dev Ind Pharm* 24(12):1113-1128, 1998; Quintanar-Guerrero *et al.*, *Pharm Res.* 15(7):1056-1062, 1998; Quintanar-Guerrero *et al.*, *J. Microencapsul.* 15(1):107-119, 1998; Douglas *et al.*, *Crit Rev Ther Drug Carrier Syst* 3(3):233-261, 1987). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles can be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkylcyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present disclosure. Such particles can be easily made, as described in Couvreur *et al.*, *J Pharm Sci* 69(2):199-202, 1980; Couvreur *et al.*, *Crit Rev Ther Drug Carrier Syst.* 5(1)1-20, 1988; zur Muhlen *et al.*, *Eur J Pharm Biopharm*, 45(2):149-155, 1998; Zambaux *et al.*, *J Control Release* 50(1-3):31-40, 1998; and U.S. Pat. No. 5,145,684.

[0111] Injectable compositions can include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (U.S. Pat. No. 5,466,468). For delivery via injection, the form is sterile and fluid to the extent that it can be delivered by syringe. In particular embodiments, it is stable under the conditions of manufacture and storage, and optionally contains one or more preservative compounds against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (e.g., glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion, and/or by the use of surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial and/or antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In various embodiments, the preparation will include an isotonic agent(s), for example, sugar(s) or sodium chloride. Prolonged absorption of the injectable

compositions can be accomplished by including in the compositions of agents that delay absorption, for example, aluminum monostearate and gelatin. Injectable compositions can be suitably buffered, if necessary, and the liquid diluent first rendered isotonic with sufficient saline or glucose.

[0112] Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. As indicated, under ordinary conditions of storage and use, these preparations can contain a preservative to prevent the growth of microorganisms.

[0113] Sterile compositions can be prepared by incorporating the physiologically active component in an appropriate amount of a solvent with other optional ingredients (e.g., as enumerated above), followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized physiologically active components into a sterile vehicle that contains the basic dispersion medium and the required other ingredients (e.g., from those enumerated above). In the case of sterile powders for the preparation of sterile injectable solutions, preferred methods of preparation can be vacuum-drying and freeze-drying techniques which yield a powder of the physiologically active components plus any additional desired ingredient from a previously sterile-filtered solution thereof.

[0114] Oral compositions may be in liquid form, for example, as solutions, syrups or suspensions, or may be presented as a drug product for reconstitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g., sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., almond oil, oily esters, or fractionated vegetable oils); and preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinized maize starch, polyvinyl pyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). Tablets may be coated by methods well-known in the art.

[0115] Inhalable compositions can be delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in

an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

[0116] Compositions can also include microchip devices (U.S. Pat. No. 5,797,898), ophthalmic formulations (Bourlais *et al.*, *Prog Retin Eye Res*, 17(1):33-58, 1998), transdermal matrices (U.S. Pat. No. 5,770,219 and U.S. Pat. No. 5,783,208) and feedback-controlled delivery (U.S. Pat. No. 5,697,899).

[0117] Supplementary active ingredients can also be incorporated into the compositions.

[0118] Typically, compositions can include at least 0.1% of the physiologically active components or more, although the percentage of the physiologically active components may, of course, be varied and may conveniently be between 1 or 2% and 70% or 80% or more or 0.5-99% of the weight or volume of the total composition. Naturally, the amount of physiologically active components in each physiologically-useful composition may be prepared in such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of compositions and dosages may be desirable.

[0119] In particular embodiments, for administration to humans, compositions should meet sterility, pyrogenicity, and the general safety and purity standards as required by United States Food and Drug Administration (FDA) or other applicable regulatory agencies in other countries.

[0120] (iii) Cell Lines Including Artificial Expression Constructs. The present disclosure includes cells including an artificial expression construct described herein. A cell that has been transformed with an artificial expression construct can be used for many purposes, including in neuroanatomical studies, assessments of functioning and/or non-functioning proteins, and drug screens that assess the regulatory properties of enhancers.

[0121] A variety of host cell lines can be used, but in particular embodiments, the cell is a mammalian neural cell. In particular embodiments, the enhancer sequence of the artificial expression construct is SEQ ID NO: 3 and/or 7 and/or a CN1390, CN1244, CN1389, CN1203, CN1367, CN1498, CN1499, CN1500, CN1838 or a combination of components depicted in FIG. 16, and the cell line is a human, primate, or murine neural cell. Cell lines which can be utilized for transgenesis in the present disclosure also include primary cell lines derived from living tissue such as rat or mouse brains and organotypic cell cultures, including brain slices from animals such as rats or mice. The PC12 cell line (available from the American Type Culture Collection, ATCC, Manassas, VA) has been shown to express a number of neuronal marker proteins in

response to Neuronal Growth Factor (NGF). The PC12 cell line is considered to be a neuronal cell line and is applicable for use with this disclosure. JAR cells (available from ATCC) are a platelet derived cell-line that express some neuronal genes, such as the serotonin transporter gene, and may be used with embodiments described herein.

[0122] WO 91/13150 describes a variety of cell lines, including neuronal cell lines, and methods of producing them. Similarly, WO 97/39117 describes a neuronal cell line and methods of producing such cell lines. The neuronal cell lines disclosed in these patent applications are applicable for use in the present disclosure.

[0123] In particular embodiments, a "neural cell" refers to a cell or cells located within the central nervous system, and includes neurons and glia, and cells derived from neurons and glia, including neoplastic and tumor cells derived from neurons or glia. A "cell derived from a neural cell" refers to a cell which is derived from or originates or is differentiated from a neural cell.

[0124] In particular embodiments, "neuronal" describes something that is of, related to, or includes, neuronal cells. Neuronal cells are defined by the presence of an axon and dendrites. The term "neuronal-specific" refers to something that is found, or an activity that occurs, in neuronal cells or cells derived from neuronal cells, but is not found in or occur in, or is not found substantially in or occur substantially in, non-neuronal cells or cells not derived from neuronal cells, for example glial cells such as astrocytes or oligodendrocytes.

[0125] In particular embodiments, non-neuronal cell lines may be used, including mouse embryonic stem cells. Cultured mouse embryonic stem cells can be used to analyze expression of genetic constructs using transient transfection with plasmid constructs. Mouse embryonic stem cells are pluripotent and undifferentiated. These cells can be maintained in this undifferentiated state by Leukemia Inhibitory Factor (LIF). Withdrawal of LIF induces differentiation of the embryonic stem cells. In culture, the stem cells form a variety of differentiated cell types. Differentiation is caused by the expression of tissue specific transcription factors, allowing the function of an enhancer sequence to be evaluated. (See for example Fiskerstrand *et al.*, *FEBS Lett* 458: 171-174, 1999.)

[0126] Methods to differentiate stem cells into neuronal cells include replacing a stem cell culture media with a media including basic fibroblast growth factor (bFGF) heparin, an N2 supplement (e.g., transferrin, insulin, progesterone, putrescine, and selenite), laminin and polyornithine. A process to produce myelinating oligodendrocytes from stem cells is described in Hu, *et al.*, 2009, *Nat. Protoc.* 4:1614-22. Bibel, *et al.*, 2007, *Nat. Protoc.* 2:1034-43 describes a protocol to produce glutamatergic neurons from stem cells while Chatzi, *et al.*, 2009, *Exp. Neurol.* 217:407-16 describes a procedure to produce GABAergic neurons. This procedure includes exposing stem

cells to all-trans-RA for three days. After subsequent culture in serum-free neuronal induction medium including Neurobasal medium supplemented with B27, bFGF and EGF, 95% GABA neurons develop

[0127] U.S. Publication No. 2012/0329714 describes use of prolactin to increase neural stem cell numbers while U.S. Publication No. 2012/0308530 describes a culture surface with amino groups that promotes neuronal differentiation into neurons, astrocytes and oligodendrocytes. Thus, the fate of neural stem cells can be controlled by a variety of extracellular factors. Commonly used factors include brain derived growth factor (BDNF; Shetty and Turner, 1998, *J. Neurobiol.* 35:395-425); fibroblast growth factor (bFGF; U.S. Pat. No. 5,766,948; FGF-1, FGF-2); Neurotrophin-3 (NT-3) and Neurotrophin-4 (NT-4); Caldwell, *et al.*, 2001, *Nat. Biotechnol.* 1;19:475-9); ciliary neurotrophic factor (CNTF); BMP-2 (U.S. Pat. Nos. 5,948,428 and 6,001,654); isobutyl 3-methylxanthine; leukemia inhibitory growth factor (LIF; U.S. Patent No. 6,103,530); somatostatin; amphiregulin; neurotrophins (*e.g.*, cyclic adenosine monophosphate; epidermal growth factor (EGF); dexamethasone (glucocorticoid hormone); forskolin; GDNF family receptor ligands; potassium; retinoic acid (U.S. Patent No. 6,395,546); tetanus toxin; and transforming growth factor- α and TGF- β (U.S. Pat. Nos. 5,851,832 and 5,753,506).

[0128] In particular embodiments, yeast one-hybrid systems may also be used to identify compounds that inhibit specific protein/DNA interactions, such as transcription factors for I56i enhancers, cores thereof and/or the SEQ ID NO: 3 and/or 7

[0129] Transgenic animals are described below. Cell lines may also be derived from such transgenic animals. For example, primary tissue culture from transgenic mice (*e.g.*, also as described below) can provide cell lines with the expression construct already integrated into the genome. (for an example see MacKenzie & Quinn, *Proc Natl Acad Sci USA* 96: 15251-15255, 1999).

[0130] (iv) Transgenic Animals. Another aspect of the disclosure includes transgenic animals, the genome of which contains an artificial expression construct including concatamerizations of I56i enhancer cores such as SEQ ID NO: 2 and/or 6 (*e.g.*, SEQ ID NO: 3 and/or 7) operatively linked to a heterologous coding sequence. In particular embodiments, the genome of a transgenic animal includes the CN1390, CN1244, CN1389, CN1203, CN1367, CN1498, CN1499, CN1500, CN1838 or a combination of components depicted in FIG. 16. In particular embodiments, when a non-integrating vector is utilized, a transgenic animal includes an artificial expression construct including concatamerizations of I56i enhancer cores such as SEQ ID NO: 2 and/or 6 (*e.g.*, SEQ ID NO: 3 and/or 7) and/or CN1390, CN1244, CN1389, CN1203, CN1367, CN1498, CN1499, CN1500, CN1838 or a combination of components depicted in FIG. 16 within one or more of its

cells.

[0131] Detailed methods for producing transgenic animals are described in U.S. Pat. No. 4,736,866. Transgenic animals may be of any nonhuman species, but preferably include nonhuman primates (NHPs), sheep, horses, cattle, pigs, goats, dogs, cats, rabbits, chickens, and rodents such as guinea pigs, hamsters, gerbils, rats, mice, and ferrets.

[0132] In particular embodiments, construction of a transgenic animal results in an organism that has an engineered construct present in all cells in the same genomic integration site. Thus, cell lines derived from such transgenic animals will be consistent in as much as the engineered construct will be in the same genomic integration site in all cells and hence will suffer the same position effect variegation. In contrast, introducing genes into cell lines or primary cell cultures can give rise to heterologous expression of the construct. A disadvantage of this approach is that the expression of the introduced DNA may be affected by the specific genetic background of the host animal.

[0133] As indicated above in relation to cell lines, the artificial expression constructs of this disclosure can be used to genetically modify mouse embryonic stem cells using techniques known in the art. Typically, the artificial expression construct is introduced into cultured murine embryonic stem cells. Transformed ES cells are then injected into a blastocyst from a host mother and the host embryo re-implanted into the mother. This results in a chimeric mouse whose tissues are composed of cells derived from both the embryonic stem cells present in the cultured cell line and the embryonic stem cells present in the host embryo. Usually the mice from which the cultured ES cells used for transgenesis are derived are chosen to have a different coat color from the host mouse into whose embryos the transformed cells are to be injected. Chimeric mice will then have a variegated coat color. As long as the germ-line tissue is derived, at least in part, from the genetically modified cells, then the chimeric mice be crossed with an appropriate strain to produce offspring that will carry the transgene.

[0134] In addition to the methods of delivery described above, the following techniques are also contemplated as alternative methods of delivering artificial expression constructs to target cells or selected tissues and organs of an animal, and in particular, to cells, organs, or tissues of a vertebrate mammal: sonophoresis (*e.g.*, ultrasound, as described in U.S. Pat. No. 5,656,016); intraosseous injection (U.S. Pat. No. 5,779,708); microchip devices (U.S. Pat. No. 5,797,898); ophthalmic formulations (Bourlais *et al.*, *Prog Retin Eye Res*, 17(1):33-58, 1998); transdermal matrices (U.S. Pat. No. 5,770,219 and U.S. Pat. No. 5,783,208); and feedback-controlled delivery (U.S. Pat. No. 5,697,899).

[0135] (v) Methods of Use. In particular embodiments, a composition including a physiologically

active component described herein is administered to a subject to result in a physiological effect.

[0136] In particular embodiments, the disclosure includes the use of the artificial expression constructs described herein to modulate expression of a heterologous gene which is either partially or wholly encoded in a location downstream to that enhancer in an engineered sequence. Thus, there are provided herein methods of use of the disclosed artificial expression constructs in the research, study, and potential development of medicaments for preventing, treating or ameliorating the symptoms of a disease, dysfunction, or disorder.

[0137] Particular embodiments include methods of administering to a subject an artificial expression construct that includes SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 3 and/or SEQ ID NO: 7 as described herein to drive selective expression of a gene in a selected neural cell type.

[0138] Particular embodiments include methods of administering to a subject an artificial expression construct that includes CN1390, CN1244, CN1389, CN1203, CN1367, CN1498, CN1499, CN1500, CN1838 or a combination of components depicted in FIG. 16 as described herein to drive selective expression of a gene in a selected neural cell type wherein the subject can be an isolated cell, a network of cells, a tissue slice, an experimental animal, a veterinary animal, or a human.

[0139] As is well known in the medical arts, dosages for any one subject depends upon many factors, including the subject's size, surface area, age, the particular compound to be administered, sex, time and route of administration, general health, and other drugs being administered concurrently. Dosages for the compounds of the disclosure will vary, but, in particular embodiments, a dose could be from 10^5 to 10^{100} copies of an artificial expression construct of the disclosure. In particular embodiments, a patient receiving intravenous, intraspinal, retro-orbital, or intrathecal administration can be infused with from 10^6 to 10^{22} copies of the artificial expression construct.

[0140] An "effective amount" is the amount of a composition necessary to result in a desired physiological change in the subject. Effective amounts are often administered for research purposes. Effective amounts disclosed herein can cause a statistically-significant effect in an animal model or in vitro assay.

[0141] In particular embodiments, constructs disclosed herein can be utilized to treat Dravet Syndrome. In particular embodiments, the methods reduce or prevent seizures, or symptoms thereof in a patient in need thereof. In particular embodiments, the methods provided may reduce or prevent one or more different types of seizures. Ideally, the methods of the disclosure result in a total prevention of seizures. However, the disclosure also encompasses methods in which the instances of seizures are decreased by at least 10%, at least 20%, at least 30%, at least 40%, at

least 50%, at least 60%, at least 70%, at least 80% or at least 90%.

[0142] Generally, a seizure can include convulsions, repetitive movements, unusual sensations, and combinations thereof. Seizures can be categorized as focal seizures (also referred to as partial seizures) and generalized seizures. Focal seizures affect only one side of the brain, while generalized seizures affect both sides of the brain. Specific types of focal seizures include simple focal seizures, complex focal seizures, and secondarily generalized seizures. Simple focal seizures can be restricted or focused on a particular lobe (e.g., temporal lobe, frontal lobe, parietal lobe, or occipital lobe). Complex focal seizures generally affect a larger part of one hemisphere than simple focal seizures, but commonly originate in the temporal lobe or the frontal lobe. When a focal seizure spreads from one side (hemisphere) to both sides of the brain, the seizure is referred to as a secondarily generalized seizure. Specific types of generalized seizures include absences (also referred to as petit mal seizures), tonic seizures, atonic seizures, myoclonic seizures, tonic clonic seizures (also referred to as grand mal seizures), and clonic seizures.

[0143] In particular embodiments, methods described herein may reduce the frequency of seizures, reduce the severity of seizures, change the type of seizures (e.g., from a more severe type to a less severe type), or a combination thereof in a patient after treatment compared to the absence of treatment (e.g., before treatment), or compared to treatment with an alternative conventional treatment.

[0144] The amount of expression constructs and time of administration of such compositions will be within the purview of the skilled artisan having benefit of the present teachings. It is likely, however, that the administration of effective amounts of the disclosed compositions may be achieved by a single administration, such as for example, a single injection of sufficient numbers of infectious particles to provide an effect in the subject. Alternatively, in some circumstances, it may be desirable to provide multiple, or successive administrations of the artificial expression construct compositions or other genetic constructs, either over a relatively short, or a relatively prolonged period of time, as may be determined by the individual overseeing the administration of such compositions. For example, the number of infectious particles administered to a mammal may be 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , or even higher, infectious particles/ml given either as a single dose or divided into two or more administrations as may be required to achieve an intended effect. In fact, in certain embodiments, it may be desirable to administer two or more different expression constructs in combination to achieve a desired effect.

[0145] In certain circumstances it will be desirable to deliver the artificial expression construct in suitably formulated compositions disclosed herein either by pipette, retro-orbital injection, subcutaneously, intraocularly, intravitreally, parenterally, subcutaneously, intravenously,

intracerebro-ventricularly, intramuscularly, intrathecally, intraspinally, orally, intraperitoneally, by oral or nasal inhalation, or by direct application or injection to one or more cells, tissues, or organs. The methods of administration may also include those modalities as described in U.S. Pat. No. 5,543,158; U.S. Pat. No. 5,641,515 and U.S. Pat. No. 5,399,363.

[0146] (vi) Kits and Commercial Packages. Kits and commercial packages contain an artificial expression construct described herein. The expression construct can be isolated. In particular embodiments, the components of an expression product can be isolated from each other. In particular embodiments, the expression product can be within a vector, within a viral vector, within a cell, within a tissue slice or sample, and/or within a transgenic animal. Such kits may further include one or more reagents, restriction enzymes, peptides, therapeutics, pharmaceutical compounds, or means for delivery of the compositions such as syringes, injectables, and the like.

[0147] Embodiments of a kit or commercial package will also contain instructions regarding use of the included components, for example, in basic research, electrophysiological research, neuroanatomical research, and/or the research and/or treatment of a disorder, disease or condition.

[0148] The Exemplary Embodiments and Experimental Examples below are included to demonstrate particular embodiments of the disclosure. Those of ordinary skill in the art should recognize in light of the present disclosure that many changes can be made to the specific embodiments disclosed herein and still obtain a like or similar result without departing from the spirit and scope of the disclosure.

[0149] (vii) Exemplary Embodiments.

1. A core of a I56i enhancer, a concatemerized core of a I56i enhancer, or a concatemerized I56i enhancer.
2. The I56i enhancer core, concatemerized I56i enhancer core, or concatemerized I56i enhancer of embodiment 1, wherein the I56i enhancer is human, murine, or zebrafish (I46i).
3. The I56i enhancer core, concatemerized I56i enhancer core, or concatemerized I56i enhancer of embodiment 1 or 2, wherein the concatemerized core includes SEQ ID NO: 2 or 6.
4. The I56i enhancer core, concatemerized I56i enhancer core, or concatemerized I56i enhancer of any of embodiments 1-3, wherein the concatemerized core includes 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the I56i core.
5. The I56i enhancer core, concatemerized I56i enhancer core, or concatemerized I56i enhancer of embodiment 4, including 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of SEQ ID NO: 2 and/or 6 (e.g., SEQ ID NO: 2 and SEQ ID NO: 6 within one sequence such as SEQ ID NO: 2 - SEQ ID NO: 2 - SEQ ID NO: 6; SEQ ID NO: 2 - SEQ ID NO: 6 - SEQ ID NO: 6; SEQ ID NO: 2 - SEQ ID NO: 6 - SEQ

ID NO: 2; SEQ ID NO: 6 - SEQ ID NO: 6 - SEQ ID NO: 2; SEQ ID NO: 6 - SEQ ID NO: 2 - SEQ ID NO: 2; and SEQ ID NO: 6 - SEQ ID NO: 2 - SEQ ID NO: 6).

6. The I56i enhancer core, concatemerized I56i enhancer core, or concatemerized I56i enhancer of embodiment 4 or 5, including 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of SEQ ID NO: 2.

7. The I56i enhancer core, concatemerized I56i enhancer core, or concatemerized I56i enhancer of embodiment 4 or 5, including 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of SEQ ID NO: 6.

8. The I56i enhancer core, concatemerized I56i enhancer core, or concatemerized I56i enhancer of embodiment 4 or 5, including 3 copies of SEQ ID NO: 2.

9. The I56i enhancer core, concatemerized I56i enhancer core, or concatemerized I56i enhancer of embodiment 4 or 5, including 3 copies of SEQ ID NO: 6.

10. The I56i enhancer core, concatemerized I56i enhancer core, or concatemerized I56i enhancer of embodiment 8, wherein the concatemerized core includes SEQ ID NO: 3.

11. The I56i enhancer core, concatemerized I56i enhancer core, or concatemerized I56i enhancer of embodiment 9, wherein the concatemerized core includes SEQ ID NO: 7.

12. An artificial expression construct including (i) an I56i enhancer core, concatemerized I56i enhancer core, or concatemerized I56i enhancer of any of embodiments 1-11, (ii) a promoter; and (iii) a heterologous encoding sequence.

13. The artificial expression construct of embodiment 12, wherein the heterologous encoding sequence encodes an effector element or an expressible element.

14. The artificial expression construct of embodiment 12 or 13, wherein the effector element includes a reporter protein or a functional molecule.

15. The artificial expression construct of embodiment 14, wherein the reporter protein is a fluorescent protein.

16. The artificial expression construct of embodiment 14 or 15, wherein the effector element is Cre, iCre, dgCre, FlpE, FlpO, or tTA2 or a functional molecule selected from a functional ion transporter, enzyme, a transcription factor, a receptor, a membrane protein, a cellular trafficking protein, a signaling molecule, a neurotransmitter, a calcium reporter, a channel rhodopsin, a CRISPR/CAS molecule, an editase, a guide RNA molecule, a homologous recombination donor cassette, or a designer receptor exclusively activated by designer drug (DREADD).

17. The artificial expression construct of any of embodiments 13, wherein the expressible element is a non-functional molecule.

18. The artificial expression construct of embodiment 17, wherein the non-functional molecule is a non-functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channel

rhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or a DREADD.

19. The artificial expression construct of any of embodiments 12-18, wherein the expression construct is associated with a capsid that crosses the blood brain barrier.

20. The artificial expression construct of embodiment 19, wherein the capsid includes PHP.eB, AAV-BR1, AAV-PHP.S, AAV-PHP.B, or AAV-PPS.

21. The artificial expression construct of any of embodiments 12-20, wherein the expression construct includes or encodes a skipping element.

22. The artificial expression construct of embodiment 21, wherein the skipping element includes a 2A peptide and/or an internal ribosome entry site (IRES).

23. The artificial expression construct of embodiment 22, wherein the 2A peptide is selected from T2A, P2A, E2A, or F2A.

24. The artificial expression construct of any of embodiments 12-23, wherein the expression construct includes a set of features selected from: 3Xh156Core, minBglobin, minCMV, SYFP2, His, 3xHA, NavMs, NavBp, NavSheP-D60N, WPRE3, BGHpA or a combination of features selected from a construct depicted in FIG. 16.

25. A vector including an artificial expression construct of any of embodiments 12-24.

26. A vector including a combination of components depicted in FIG. 16.

27. The vector of embodiment 26, wherein the vector is a viral vector.

28. The vector of embodiment 26 or 27, wherein the viral vector is a recombinant adeno-associated viral (AAV) vector.

29. An adeno-associated viral (AAV) vector including at least one heterologous encoding sequence, wherein the heterologous encoding sequence is under control of a promoter and an enhancer selected from SEQ ID NO: 3 and/or 7.

30. The AAV vector of embodiment 29, wherein the AAV vector is replication-competent.

31. A transgenic cell including an expression construct or vector of any of the preceding embodiments.

32. The transgenic cell of embodiment 31, wherein the transgenic cell is a GABAergic interneuron.

33. A non-human transgenic animal including an expression construct, vector, or transgenic cell of any of the preceding embodiments.

34. The non-human transgenic animal of embodiment 33 wherein the non-human transgenic animal is a mouse or a non-human primate.

35. An administrable composition including an expression construct, vector, or transgenic cell of any of the preceding embodiments.

36. A kit including an expression construct, vector, transgenic cell, transgenic animal, and/or administrable compositions of any of the preceding embodiments.
37. A method for selectively expressing a heterologous gene within a population of neural cells in vivo or in vitro, the method including providing the administrable composition of embodiment 35 in a sufficient dosage and for a sufficient time to a sample or subject including the population of neural cells thereby selectively expressing the gene within the population of neural cells.
38. The method of embodiment 37, wherein the heterologous gene encodes an effector element or an expressible element.
39. The method of embodiment 38, wherein the effector element includes a reporter protein or a functional molecule.
40. The method of embodiment 39, wherein the reporter protein is a fluorescent protein.
41. The method of embodiment 39 or 40, wherein the effector element is Cre, iCre, dgCre, FlpE, FlpO, or tTA2 or a functional molecule selected from a functional ion transporter, enzyme, a transcription factor, a receptor, a membrane protein, a cellular trafficking protein, a signaling molecule, a neurotransmitter, a calcium reporter, a channel rhodopsin, a CRISPR/CAS molecule, an editase, a guide RNA molecule, a homologous recombination donor cassette, or a DREADD.
42. The method of embodiment 38, wherein the expressible element is a non-functional molecule.
43. The method of embodiment 42, wherein the non-functional molecule is a non-functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channel rhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or DREADD.
44. The method of any of embodiments 37 - 43, wherein the providing includes pipetting.
45. The method of embodiment 44, wherein the pipetting is to a brain slice.
46. The method of embodiment 45, wherein the brain slice includes a GABAergic interneuron.
47. The method of any of embodiments 45 or 46, wherein the brain slice is murine, human, or non-human primate.
48. The method of any of embodiments 37 - 43, wherein the providing includes administering to a living subject.
49. The method of embodiment 48, wherein the living subject is a human, non-human primate, or a mouse.
50. The method of any of embodiments 48 or 49, wherein the administering to a living subject is through injection.
51. The method of embodiment 50, wherein the injection includes intravenous injection, intraparenchymal injection into brain tissue, intracerebroventricular (ICV) injection, intra-cisterna

magna (ICM) injection, or intrathecal injection.

52. An artificial expression construct consisting of or consisting essentially of a combination of features depicted in FIG. 16.

53. Any of the embodiments above including an effector element or expressible element wherein the effector element or expressible element is an ion transporter selected from a voltage gated sodium channel (e.g., SCN1A), a potassium channel (e.g., KCNQ2), or a calcium channel (e.g., CACNA1C); a cellular trafficking protein selected from clathrin, dynamin, caveolin, Rab-4A, or Rab-11A; an enzyme selected from lactase, lipase, helicase, alpha-glucosidase, and amylase); a transcription factor selected from SP1, AP-1, Heat shock factor protein 1, C/EBP (CCAAT/enhancer binding protein), and Oct-1; a receptor selected from transforming growth factor receptor beta 1, platelet-derived growth factor receptor, epidermal growth factor receptor, vascular endothelial growth factor receptor, and interleukin 8 receptor alpha; a signaling molecule selected from nerve growth factor (NGF), platelet-derived growth factor (PDGF), transforming growth factor β (TGF β), epidermal growth factor (EGF), and GTPase HRas; a neurotransmitter selected from cocaine and amphetamine regulated transcript, substance P, oxytocin, and somatostatin; a calcium reporter selected from genetically encoded calcium indicators (GEC1, NTnC, GCaMP6s, GCaMP6f, GCaMP6m, jGCaMP7s, jGCaMP7f, jGCaMP7b, jGCaMP7c, jRGECO1a, jRGECO1b), Myosin light chain kinase, Green fluorescent protein, Calmodulin chimera, Calcium indicator TN-XXL, BRET-based auto-luminescent calcium indicator, and Calcium indicator protein OeNL(Ca²⁺)-18u); a channel rhodopsin selected from channelrhodopsin-1 and channelrhodopsin-2 or a variant thereof; guide RNA; a nuclease selected from Cas, Cas9, Cpf1, ribonuclease 4, and deoxyribonuclease II beta; and/or a DREADD (e.g., hM3DREADD, hM4DREADD).

Within this disclosure I56i should be interpreted to be I46i when the context demonstrates reference to or use of the zebrafish form of the enhancers described herein

[0150] (viii) Experimental Example. Dravet syndrome (DS) is a drug-resistant and life-threatening form of epilepsy. It typically begins in the first year of life, with fever- or temperature-induced seizures that evolve into generalized clonic, tonic-clonic, and unilateral seizures. These seizures are often resistant to current anti-epileptic drugs, the first-line therapies for this syndrome; complete seizure control is typically not achieved. As the disease progresses, most affected children also suffer from comorbid conditions including developmental delays, intellectual disabilities, impaired motor control and coordination, autistic behaviors, sleep disturbances, and many die prematurely.

[0151] Heterozygous loss-of-function mutations in *SCN1A*, the gene that encodes the pore-

forming subunit of the voltage-gated sodium channel Nav1.1 are the most common cause of DS and occur in nearly 1/16,000 newborns.

[0152] A mouse model, generated by knock-out of *Scn1a*, replicates the several key phenotypic features of this epilepsy including infantile (P21)-epilepsy onset, high susceptibility to thermal seizures, ataxia, spontaneous seizures, sleep impairments, autistic behaviors, and premature death. Seizures and several comorbidities arise from impaired interneuron function in these mice.

[0153] This mouse model was used to investigate the efficacy of a new viral vector for DS. The virus was delivered by retro-orbital injection using an insulin syringe and its ability to suppress seizure was evaluated using the thermal seizure test. In this test, the mouse body core temperature is elevated slowly, using a temperature controller and a heat lamp, until a seizure occurs, or 42.5°C is attained. The temperature of seizure onset in treated and control mice are compared to determine the efficacy of the intervention. In additional tests, the efficacy of treatment on spontaneous seizure and premature mortality are assessed using video and electroencephalographic monitoring.

[0154] The viral vector is a new AAV viral vector named CN1500. This viral vector is a recombinant AAV that expresses the transgene SYFP2-P2A-NavSheP-D60N to rescue the loss of the voltage-gated sodium channel Nav1.1. NavSheP-D60N is a modified voltage-gated sodium channel of bacterial origin that has been modified to improve the kinetics and expression in mammalian cells. The transgene expression level is elevated by the addition of a WPRE3 element, and transcription is terminated with the bovine growth hormone poly adenylation sequence. Expression of the transgene is high and limited to inhibitory cells in forebrain structures including the cortex and the hippocampus, via the 3xh156iCore synthetic enhancer (SEQ ID NO: 3) directly 5' of a CMV minimal promoter. Furthermore, the therapeutic transgene NavSheP-D60N is labeled by an HA epitope tag to verify correct protein localization.

[0155] To test the efficacy of the therapeutic AAV viral vector, CN1500 package using the PHP.eB serotype was used. A cohort of postnatal day 35 *Scn1a*^{+/-} mice were either injected with 2x10¹¹ vg per animal or were left un-injected. The AAV was introduced intravenously using the retro-orbital delivery route. Two weeks after viral administration, animals from the treatment and control groups were assessed for their susceptibility to febrile seizures. As indicated previously, febrile seizures were measured by steadily raising the mouse's temperature under a heat lamp 0.5 Celsius every two minutes and measuring the internal temperature of the mouse with a rectal probe. The temperature where the mouse experienced a seizure is recorded.

[0156] The new therapeutic vector CN1500 was both highly expressed in mouse cortical and hippocampal GABAergic cells, but also raised the average temperature where *Scn1a*^{+/-} mice

experienced febrile seizures from 38.7 °C to 41 °C. These data show that CN1500 can substantially rescue the loss of *Scn1a*.

[0157] Example 1 references include: Catterall et al. (2010) *The Journal of physiology* 588:1849-1859; Cheah et al. (2012) *Proceedings of the National Academy of Sciences of the United States of America* 109:14646-14651; Kalume (2013) *Respir Physiol Neurobiol.* 189(2):324-8; Kalume et al., (2007) *J Neurosci* 27:11065-11074; Kalume et al., (2013) *The Journal of clinical investigation* 123:1798-1808; Oakley et al., (2009) *Proceedings of the National Academy of Sciences of the United States of America* 106:3994-3999.

[0158] (ix) Closing Paragraphs. Nucleic acid sequences described herein are shown using standard letter abbreviations for nucleotide bases, as defined in 37 C.F.R. §1.822. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included in embodiments where it would be appropriate.

[0159] Variants of the sequences disclosed and referenced herein are also included. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological activity can be found using computer programs well known in the art, such as DNASTAR™ (Madison, Wisconsin) software. Preferably, amino acid changes in the protein variants disclosed herein are conservative amino acid changes, i.e., substitutions of similarly charged or uncharged amino acids. A conservative amino acid change involves substitution of one of a family of amino acids which are related in their side chains.

[0160] In a peptide or protein, suitable conservative substitutions of amino acids are known to those of skill in this art and generally can be made without altering a biological activity of a resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, e.g., Watson et al. *Molecular Biology of the Gene*, 4th Edition, 1987, The Benjamin/Cummings Pub. Co., p. 224). Naturally occurring amino acids are generally divided into conservative substitution families as follows: Group 1: Alanine (Ala), Glycine (Gly), Serine (Ser), and Threonine (Thr); Group 2: (acidic): Aspartic acid (Asp), and Glutamic acid (Glu); Group 3: (acidic; also classified as polar, negatively charged residues and their amides): Asparagine (Asn), Glutamine (Gln), Asp, and Glu; Group 4: Gln and Asn; Group 5: (basic; also classified as polar, positively charged residues): Arginine (Arg), Lysine (Lys), and Histidine (His); Group 6 (large aliphatic, nonpolar residues): Isoleucine (Ile), Leucine (Leu), Methionine (Met), Valine (Val) and Cysteine (Cys); Group 7 (uncharged polar): Tyrosine (Tyr), Gly, Asn, Gln, Cys, Ser, and Thr; Group 8 (large aromatic residues): Phenylalanine (Phe), Tryptophan (Trp), and Tyr; Group 9 (non-polar): Proline (Pro), Ala, Val, Leu, Ile, Phe, Met, and Trp; Group 11 (aliphatic): Gly, Ala, Val, Leu,

and Ile; Group 10 (small aliphatic, nonpolar or slightly polar residues): Ala, Ser, Thr, Pro, and Gly; and Group 12 (sulfur-containing): Met and Cys. Additional information can be found in Creighton (1984) *Proteins*, W.H. Freeman and Company.

[0161] In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, *J. Mol. Biol.* 157(1), 105-32). Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: Ile (+4.5); Val (+4.2); Leu (+3.8); Phe (+2.8); Cys (+2.5); Met (+1.9); Ala (+1.8); Gly (−0.4); Thr (−0.7); Ser (−0.8); Trp (−0.9); Tyr (−1.3); Pro (−1.6); His (−3.2); Glutamate (−3.5); Gln (−3.5); aspartate (−3.5); Asn (−3.5); Lys (−3.9); and Arg (−4.5).

[0162] It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, i.e., still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity.

[0163] As detailed in U.S. Pat. No. 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: Arg (+3.0); Lys (+3.0); aspartate (+3.0 $\pm$ 1); glutamate (+3.0 $\pm$ 1); Ser (+0.3); Asn (+0.2); Gln (+0.2); Gly (0); Thr (−0.4); Pro (−0.5 $\pm$ 1); Ala (−0.5); His (−0.5); Cys (−1.0); Met (−1.3); Val (−1.5); Leu (−1.8); Ile (−1.8); Tyr (−2.3); Phe (−2.5); Trp (−3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

[0164] As outlined above, amino acid substitutions may be based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like.

[0165] As indicated elsewhere, variants of gene sequences can include codon optimized variants, sequence polymorphisms, splice variants, and/or mutations that do not affect the function of an encoded product to a statistically-significant degree.

[0166] Variants of the protein, nucleic acid, and gene sequences disclosed herein also include

sequences with at least 70% sequence identity, 80% sequence identity, 85% sequence identity, 90% sequence identity, 95% sequence identity, 96% sequence identity, 97% sequence identity, 98% sequence identity, or 99% sequence identity to the protein, nucleic acid, or gene sequences disclosed herein.

[0167] "% sequence identity" refers to a relationship between two or more sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between protein, nucleic acid, or gene sequences as determined by the match between strings of such sequences. "Identity" (often referred to as "similarity") can be readily calculated by known methods, including those described in: Computational Molecular Biology (Lesk, A. M., ed.) Oxford University Press, NY (1988); Biocomputing: Informatics and Genome Projects (Smith, D. W., ed.) Academic Press, NY (1994); Computer Analysis of Sequence Data, Part I (Griffin, A. M., and Griffin, H. G., eds.) Humana Press, NJ (1994); Sequence Analysis in Molecular Biology (Von Heijne, G., ed.) Academic Press (1987); and Sequence Analysis Primer (Gribskov, M. and Devereux, J., eds.) Oxford University Press, NY (1992). Preferred methods to determine identity are designed to give the best match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Sequence alignments and percent identity calculations may be performed using the Megalign program of the LASERGENE bioinformatics computing suite (DNASTAR, Inc., Madison, Wisconsin). Multiple alignment of the sequences can also be performed using the Clustal method of alignment (Higgins and Sharp CABIOS, 5, 151-153 (1989) with default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Relevant programs also include the GCG suite of programs (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, Wisconsin); BLASTP, BLASTN, BLASTX (Altschul, et al., J. Mol. Biol. 215:403-410 (1990); DNASTAR (DNASTAR, Inc., Madison, Wisconsin); and the FASTA program incorporating the Smith-Waterman algorithm (Pearson, Comput. Methods Genome Res., [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor(s): Suhai, Sandor. Publisher: Plenum, New York, N.Y.. Within the context of this disclosure it will be understood that where sequence analysis software is used for analysis, the results of the analysis are based on the "default values" of the program referenced. As used herein "default values" will mean any set of values or parameters, which originally load with the software when first initialized.

[0168] Variants also include nucleic acid molecules that hybridizes under stringent hybridization conditions to a sequence disclosed herein and provide the same function as the reference sequence. Exemplary stringent hybridization conditions include an overnight incubation at 42 °C in a solution including 50% formamide, 5XSSC (750 mM NaCl, 75 mM trisodium citrate), 50 mM

sodium phosphate (pH 7.6), 5XDenhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1XSSC at 50 °C. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, moderately high stringency conditions include an overnight incubation at 37°C in a solution including 6XSSPE (20XSSPE=3M NaCl; 0.2M NaH₂PO₄; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50 °C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5XSSC). Variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

[0169] As will be understood by one of ordinary skill in the art, each embodiment disclosed herein can comprise, consist essentially of or consist of its particular stated element, step, ingredient or component. Thus, the terms “include” or “including” should be interpreted to recite: “comprise, consist of, or consist essentially of.” The transition term “comprise” or “comprises” means includes, but is not limited to, and allows for the inclusion of unspecified elements, steps, ingredients, or components, even in major amounts. The transitional phrase “consisting of” excludes any element, step, ingredient or component not specified. The transition phrase “consisting essentially of” limits the scope of the embodiment to the specified elements, steps, ingredients or components and to those that do not materially affect the embodiment. A material effect would cause a statistically significant reduction in selective expression in the targeted cell population as determined by scRNA-Seq and the following enhancer / targeted cell population pairing: concatemerized core of the l56i enhancer (e.g., SEQ ID NO: 3) / GABAergic interneurons.

[0170] Unless otherwise indicated, all numbers expressing quantities of ingredients, properties such as molecular weight, reaction conditions, and so forth used in the specification and claims are to be understood as being modified in all instances by the term “about.” Accordingly, unless indicated to the contrary, the numerical parameters set forth in the specification and attached claims are approximations that may vary depending upon the desired properties sought to be obtained by the present invention. At the very least, and not as an attempt to limit the application of the doctrine of equivalents to the scope of the claims, each numerical parameter should at least

be construed in light of the number of reported significant digits and by applying ordinary rounding techniques. When further clarity is required, the term “about” has the meaning reasonably ascribed to it by a person skilled in the art when used in conjunction with a stated numerical value or range, i.e. denoting somewhat more or somewhat less than the stated value or range, to within a range of $\pm 20\%$ of the stated value; $\pm 19\%$ of the stated value; $\pm 18\%$ of the stated value; $\pm 17\%$ of the stated value; $\pm 16\%$ of the stated value; $\pm 15\%$ of the stated value; $\pm 14\%$ of the stated value; $\pm 13\%$ of the stated value; $\pm 12\%$ of the stated value; $\pm 11\%$ of the stated value; $\pm 10\%$ of the stated value; $\pm 9\%$ of the stated value; $\pm 8\%$ of the stated value; $\pm 7\%$ of the stated value; $\pm 6\%$ of the stated value; $\pm 5\%$ of the stated value; $\pm 4\%$ of the stated value; $\pm 3\%$ of the stated value; $\pm 2\%$ of the stated value; or $\pm 1\%$ of the stated value.

[0171] Notwithstanding that the numerical ranges and parameters setting forth the broad scope of the invention are approximations, the numerical values set forth in the specific examples are reported as precisely as possible. Any numerical value, however, inherently contains certain errors necessarily resulting from the standard deviation found in their respective testing measurements.

[0172] The terms “a,” “an,” “the” and similar referents used in the context of describing the invention (especially in the context of the following claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. Recitation of ranges of values herein is merely intended to serve as a shorthand method of referring individually to each separate value falling within the range. Unless otherwise indicated herein, each individual value is incorporated into the specification as if it were individually recited herein. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., “such as”) provided herein is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention otherwise claimed. No language in the specification should be construed as indicating any non-claimed element essential to the practice of the invention.

[0173] Groupings of alternative elements or embodiments of the invention disclosed herein are not to be construed as limitations. Each group member may be referred to and claimed individually or in any combination with other members of the group or other elements found herein. It is anticipated that one or more members of a group may be included in, or deleted from, a group for reasons of convenience and/or patentability. When any such inclusion or deletion occurs, the specification is deemed to contain the group as modified thus fulfilling the written description of all Markush groups used in the appended claims.

[0174] Certain embodiments of this invention are described herein, including the best mode known to the inventors for carrying out the invention. Of course, variations on these described embodiments will become apparent to those of ordinary skill in the art upon reading the foregoing description. The inventor expects skilled artisans to employ such variations as appropriate, and the inventors intend for the invention to be practiced otherwise than specifically described herein. Accordingly, this invention includes all modifications and equivalents of the subject matter recited in the claims appended hereto as permitted by applicable law. Moreover, any combination of the above-described elements in all possible variations thereof is encompassed by the invention unless otherwise indicated herein or otherwise clearly contradicted by context.

[0175] Furthermore, numerous references have been made to patents, printed publications, journal articles and other written text throughout this specification (referenced materials herein). Each of the referenced materials are individually incorporated herein by reference in their entirety for their referenced teaching.

[0176] In closing, it is to be understood that the embodiments of the invention disclosed herein are illustrative of the principles of the present invention. Other modifications that may be employed are within the scope of the invention. Thus, by way of example, but not of limitation, alternative configurations of the present invention may be utilized in accordance with the teachings herein. Accordingly, the present invention is not limited to that precisely as shown and described.

[0177] The particulars shown herein are by way of example and for purposes of illustrative discussion of the preferred embodiments of the present invention only and are presented in the cause of providing what is believed to be the most useful and readily understood description of the principles and conceptual aspects of various embodiments of the invention. In this regard, no attempt is made to show structural details of the invention in more detail than is necessary for the fundamental understanding of the invention, the description taken with the drawings and/or examples making apparent to those skilled in the art how the several forms of the invention may be embodied in practice.

[0178] Definitions and explanations used in the present disclosure are meant and intended to be controlling in any future construction unless clearly and unambiguously modified in the following examples or when application of the meaning renders any construction meaningless or essentially meaningless. In cases where the construction of the term would render it meaningless or essentially meaningless, the definition should be taken from Webster's Dictionary, 3rd Edition or a dictionary known to those of ordinary skill in the art, such as the Oxford Dictionary of Biochemistry and Molecular Biology (Ed. Anthony Smith, Oxford University Press, Oxford, 2004).

CLAIMS

What is claimed is:

1. An artificial expression construct comprising (i) an enhancer having the sequence as set forth in SEQ ID NO: 3; (ii) a promoter; and (iii) a heterologous encoding sequence encoding an effector element.
2. An artificial expression construct of claim 1, wherein the effector element is a functional protein selected from an ion transporter, a cellular trafficking protein, an enzyme, an endogenous or synthetic transcription factor, a neurotransmitter, a calcium reporter, a channel rhodopsin, guide RNA, a nuclease, or a designer receptor exclusively activated by designer drugs (DREADD).
3. The artificial expression construct of claim 2, wherein
 - the ion transporter is selected from a potassium channel, a calcium channel, or a voltage gated sodium channel;
 - the cellular trafficking protein is selected from clathrin, dynamin, caveolin, Rab-4A, or Rab-11A;
 - the enzyme is selected from lactase, lipase, helicase, alpha-glucosidase, and amylase;
 - the transcription factor is selected from SP1, AP-1, Heat shock factor protein 1, C/EBP (CCAAT/enhancer binding protein), and Oct-1; the receptor is selected from transforming growth factor receptor beta 1, platelet-derived growth factor receptor, epidermal growth factor receptor, vascular endothelial growth factor receptor, and interleukin 8 receptor alpha;
 - the signaling molecule is selected from nerve growth factor (NGF), platelet-derived growth factor (PDGF), transforming growth factor β (TGF β), epidermal growth factor (EGF), and GTPase HRas;
 - the neurotransmitter is selected from cocaine and amphetamine regulated transcript, substance P, oxytocin, and somatostatin;
 - the calcium reporter is selected from GCaMP6s, GCaMP6f, GCaMP6m, jGCaMP7s, jGCaMP7f, jGCaMP7b, jGCaMP7c, jRGECO1a, jRGECO1b, Myosin light chain kinase, Calmodulin chimera, calcium indicator TN-XXL, BRET-based auto-luminescent calcium indicator, and Calcium indicator protein OeNL(Ca²⁺)-18u);
 - the channel rhodopsin is selected from channelrhodopsin-1, channelrhodopsin-2, or a variant thereof;
 - the nuclease is selected from Cas, Cas9, and Cpf1 and/or
 - the DREADD is selected from hM3DREADD or hM4DREADD.
4. An artificial expression construct comprising (i) a concatemer of SEQ ID NO: 2 or SEQ ID NO: 6; (ii) a promoter; and (iii) a heterologous encoding sequence.

5. The artificial expression construct of claim 4, wherein the concatemer includes the sequence as set forth in SEQ ID NO: 3 or SEQ ID NO: 7.
6. The artificial expression construct of claim 4, wherein the heterologous encoding sequence encodes an effector element or an expressible element.
7. The artificial expression construct of claim 6, wherein the effector element includes a reporter protein or a functional molecule.
8. The artificial expression construct of claim 7, wherein the reporter protein is a fluorescent protein.
9. The artificial expression construct of claim 7, wherein the functional molecule is a functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channel rhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or DREADD.
10. The artificial expression construct of claim 6, wherein the expressible element is a non-functional molecule.
11. The artificial expression construct of claim 10, wherein the non-functional molecule is a non-functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channel rhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or DREADD.
12. The artificial expression construct of claim 4, wherein the artificial expression construct is associated with a capsid that crosses the blood brain barrier.
13. The artificial expression construct of claim 12, wherein the capsid includes PHP.eB, AAV-BR1, AAV-PHP.S, AAV-PHP.B, or AAV-PPS.
14. The artificial expression construct of claim 4, wherein the artificial expression construct includes or encodes a skipping element.
15. The artificial expression construct of claim 14, wherein the skipping element includes a 2A peptide or an internal ribosome entry site (IRES).
16. The artificial expression construct of claim 15, wherein the 2A peptide is selected from T2A, P2A, E2A, or F2A.
17. A vector comprising an artificial expression construct of claim 4.
18. The vector of claim 17, wherein the vector is a viral vector.
19. The vector of claim 18, wherein the viral vector is a recombinant adeno-associated viral (AAV) vector.
20. An adeno-associated viral (AAV) vector comprising at least one heterologous encoding

- sequence, wherein the heterologous encoding sequence is under the transcriptional control of a promoter and an enhancer having the sequence as set forth in SEQ ID NO: 3 or SEQ ID NO: 7.
21. The AAV vector of claim 20, wherein the heterologous encoding sequence encodes an effector element or an expressible element.
22. The AAV vector of claim 21, wherein the effector element includes a reporter protein or a functional molecule.
23. The AAV vector of claim 22, wherein the reporter protein is a fluorescent protein.
24. The AAV vector of claim 22, wherein the functional molecule is a functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channel rhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or DREADD.
25. The AAV vector of claim 21, wherein the expressible element is a non-functional molecule.
26. The AAV vector of claim 25, wherein the non-functional molecule is a non-functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channel rhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or DREADD.
27. The AAV vector of claim 20, wherein the AAV vector is replication-competent.
28. A transgenic cell comprising an artificial expression construct of claim 1 or 4 and/or a vector of claim 20.
29. The transgenic cell of claim 28, wherein the transgenic cell is a GABAergic interneuron.
30. The transgenic cell of claim 28, wherein the transgenic cell is a lysosomal associated membrane protein 5 (LAMP5) neuron, a vasoactive intestinal peptide (Vip) neuron, a somatostatin (Sst) neuron, or a parvalbumin (Pvalb) neuron.
31. The transgenic cell of claim 28, wherein the transgenic cell is murine, human, or non-human primate.
32. A non-human transgenic animal comprising an artificial expression construct of claim 1 or 4, a vector of claim 20, and/or a transgenic cell of claim 28.
33. The non-human transgenic animal of claim 32, wherein the non-human transgenic animal is a mouse or a non-human primate.
34. An administrable composition comprising an artificial expression construct of claim 1 or 4, a vector of claim 20, and/or a transgenic cell of claim 28.
35. A kit comprising an artificial expression construct of claim 1 or 4, a vector of claim 20, a transgenic cell of claim 28, and/or a transgenic animal of claim 32.
36. A method for selectively expressing a gene within a population of neural cells in vivo or in

vitro, the method comprising providing the administrable composition of claim 34 in a sufficient dosage and for a sufficient time to a sample or subject comprising the population of neural cells thereby selectively expressing the gene within the population of neural cells.

37. The method of claim 36, wherein the gene encodes an effector element or an expressible element

38. The method of claim 37, wherein the effector element comprises a reporter protein or a functional molecule.

39. The method of claim 38, wherein the reporter protein is a fluorescent protein.

40. The method of claim 38, wherein the functional molecule is a functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channel rhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or DREADD.

41. The method of claim 37, wherein the expressible element is a non-functional molecule.

42. The method of claim 41, wherein the non-functional molecule is a non-functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channel rhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or DREADD.

43. The method of claim 36, wherein the providing comprises pipetting.

44. The method of claim 43, wherein the pipetting is to a brain slice.

45. The method of claim 44, wherein the brain slice includes a GABAergic interneuron.

46. The method of claim 44, wherein the brain slice includes a LAMP5 neuron, a Vip neuron, a Sst neuron, or a Pvalb neuron.

47. The method of claim 44, wherein the brain slice is murine, human, or non-human primate.

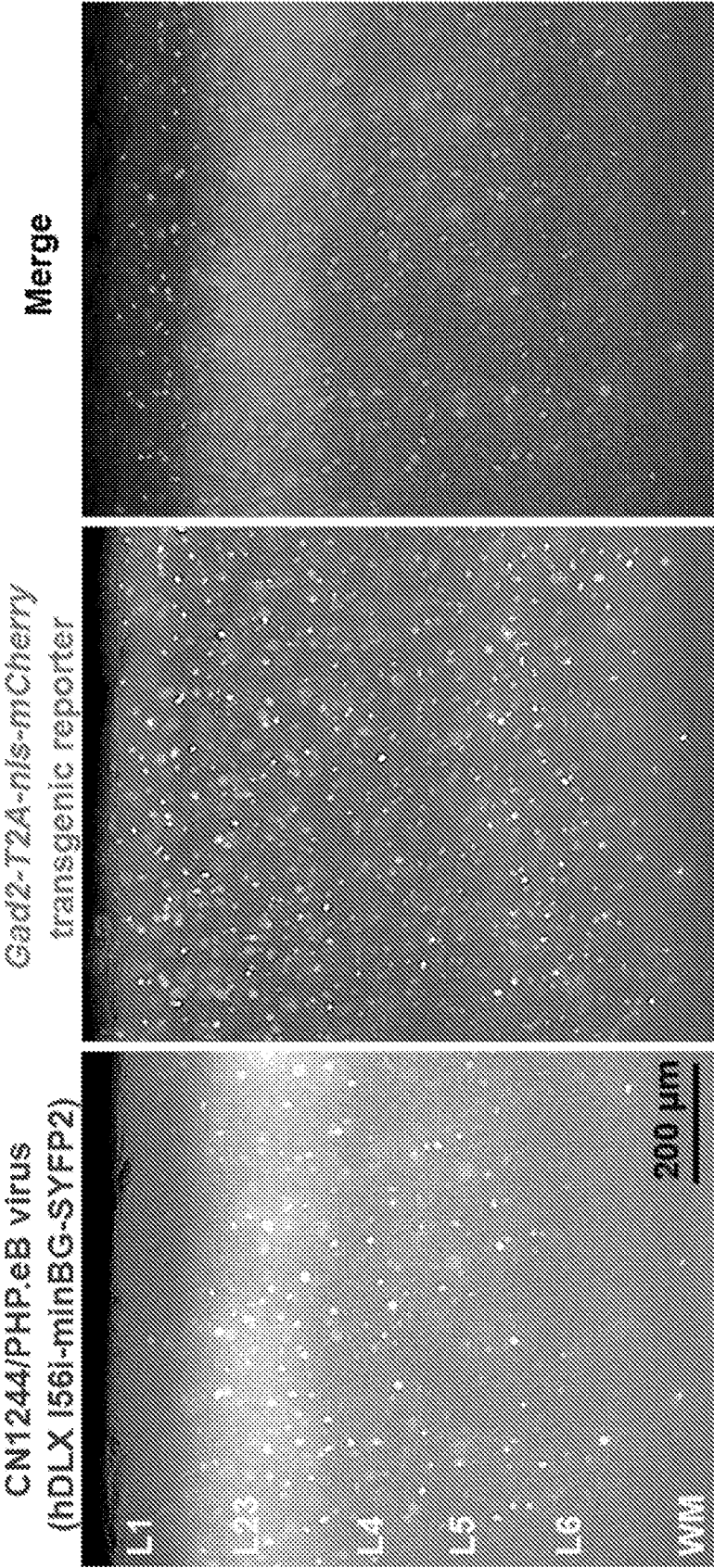
48. The method of claim 36, wherein the providing comprises administering to a living subject.

49. The method of claim 48, wherein the living subject is a human, non-human primate, or a mouse.

50. The method of claim 48, wherein the administering to a living subject is through injection.

51. The method of claim 50, wherein the injection comprises intravenous injection, intraparenchymal injection into brain tissue, intracerebroventricular (ICV) injection, intra-cisterna magna (ICM) injection, or intrathecal injection.

FIG. 1



2/104

FIG. 2A

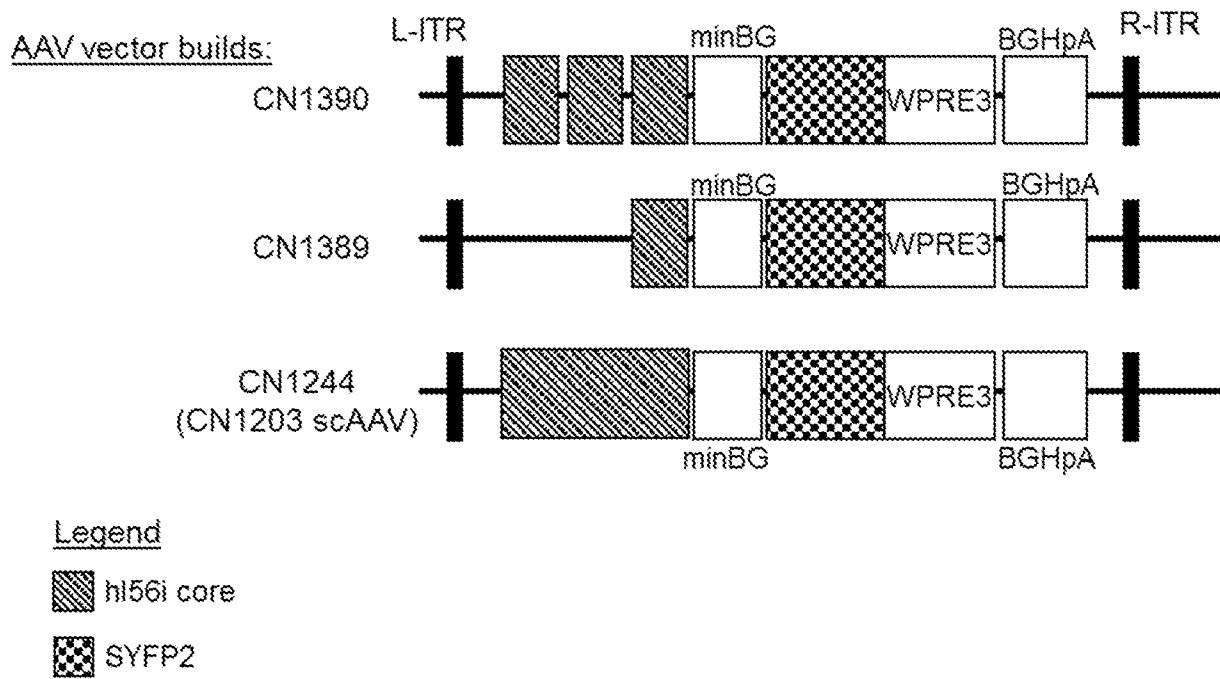


FIG. 2B

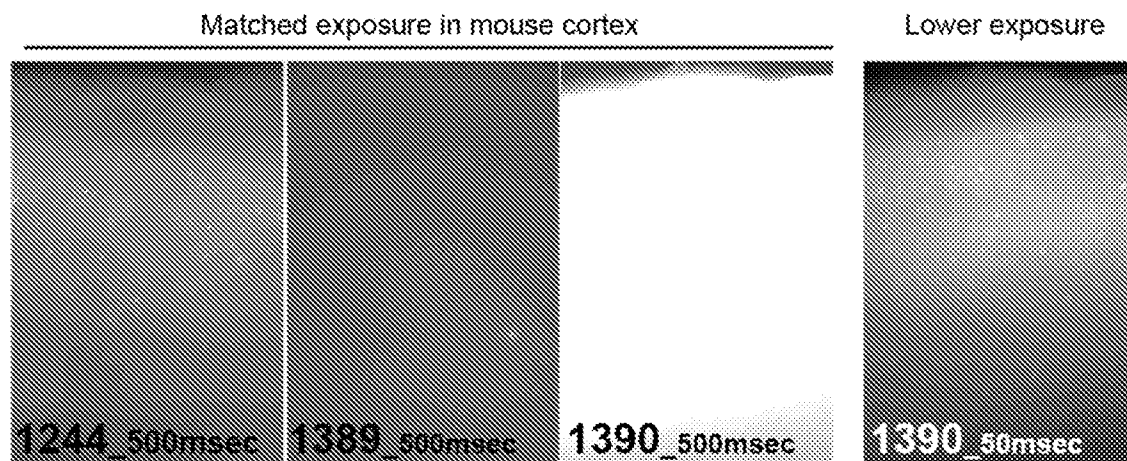
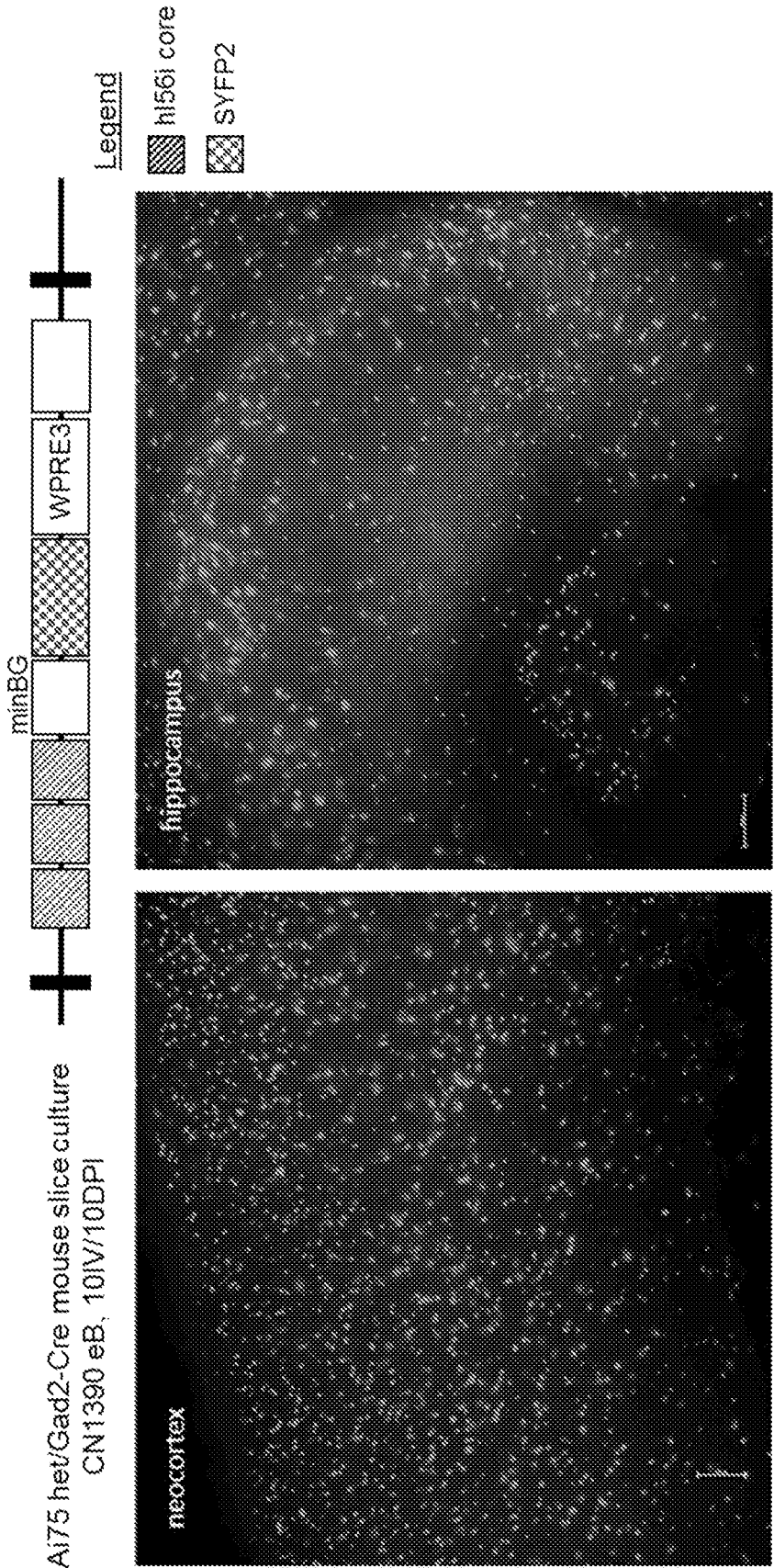


FIG. 3



4/104

FIG. 4A

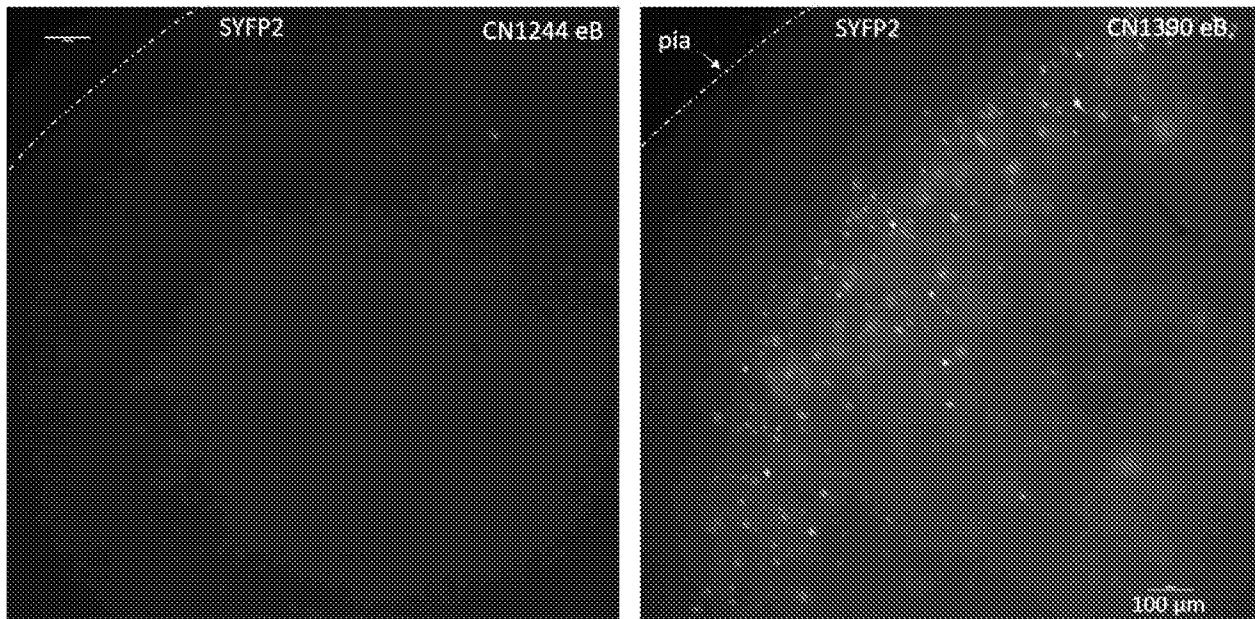
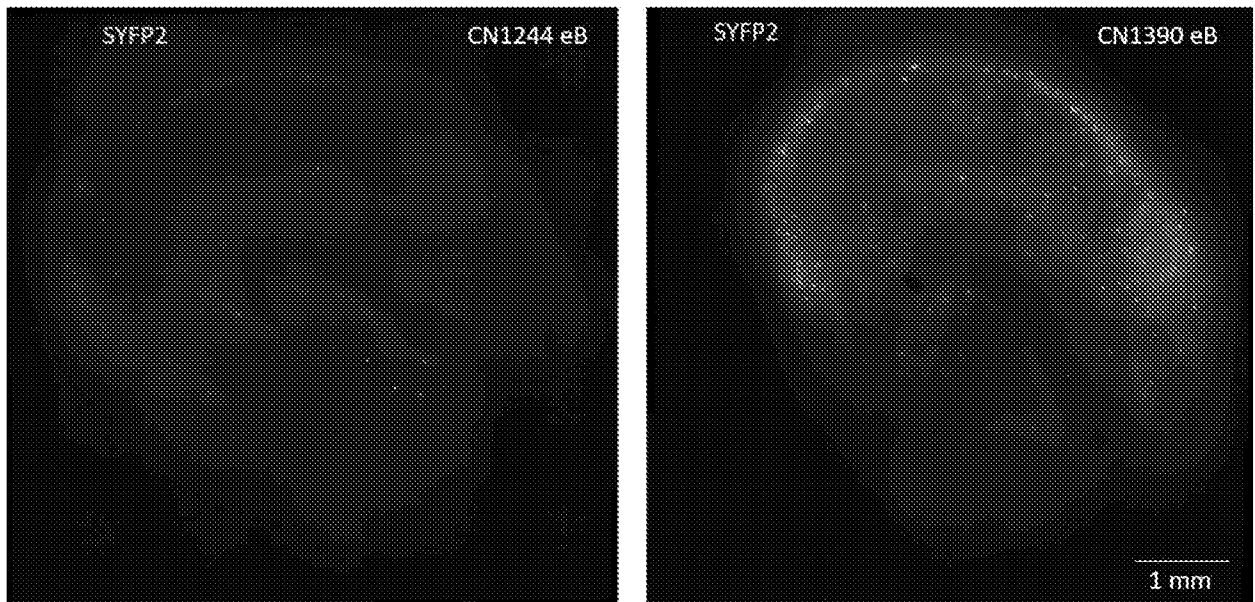
NHP neocortical slice culture, 4DIV/4DPI, matched exposure

FIG. 4B

NHP hippocampal slice culture, 6DIV/DPI, matched exposure

5/104

FIG. 5A



FIG. 5B

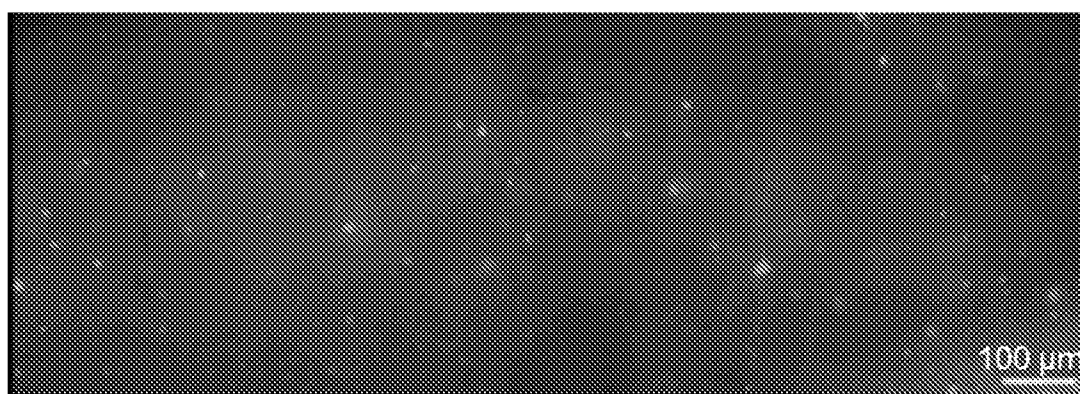


FIG. 5C

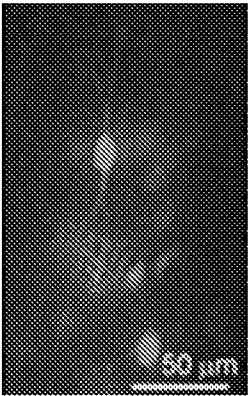


FIG. 5D

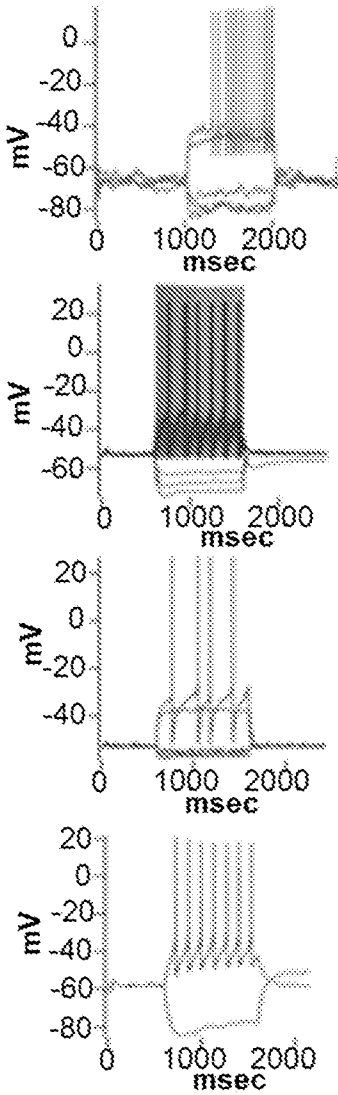
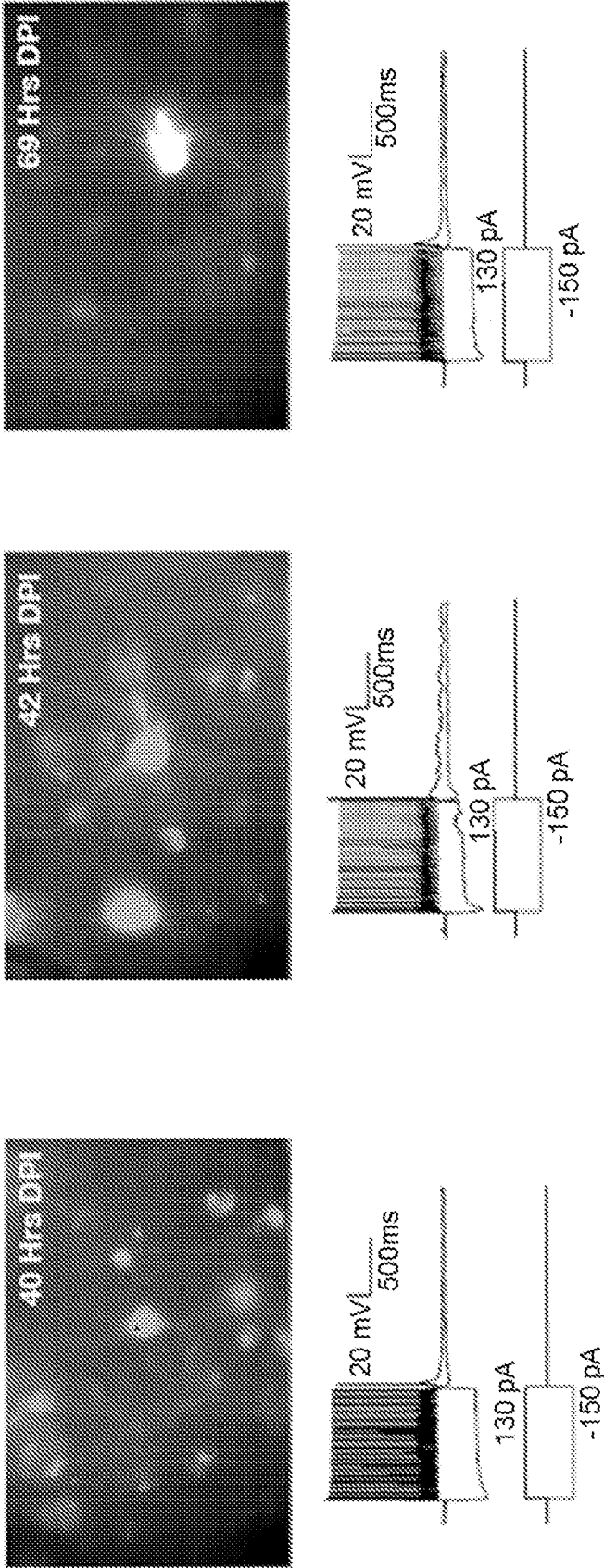


FIG. 5E



8/104

FIG. 6

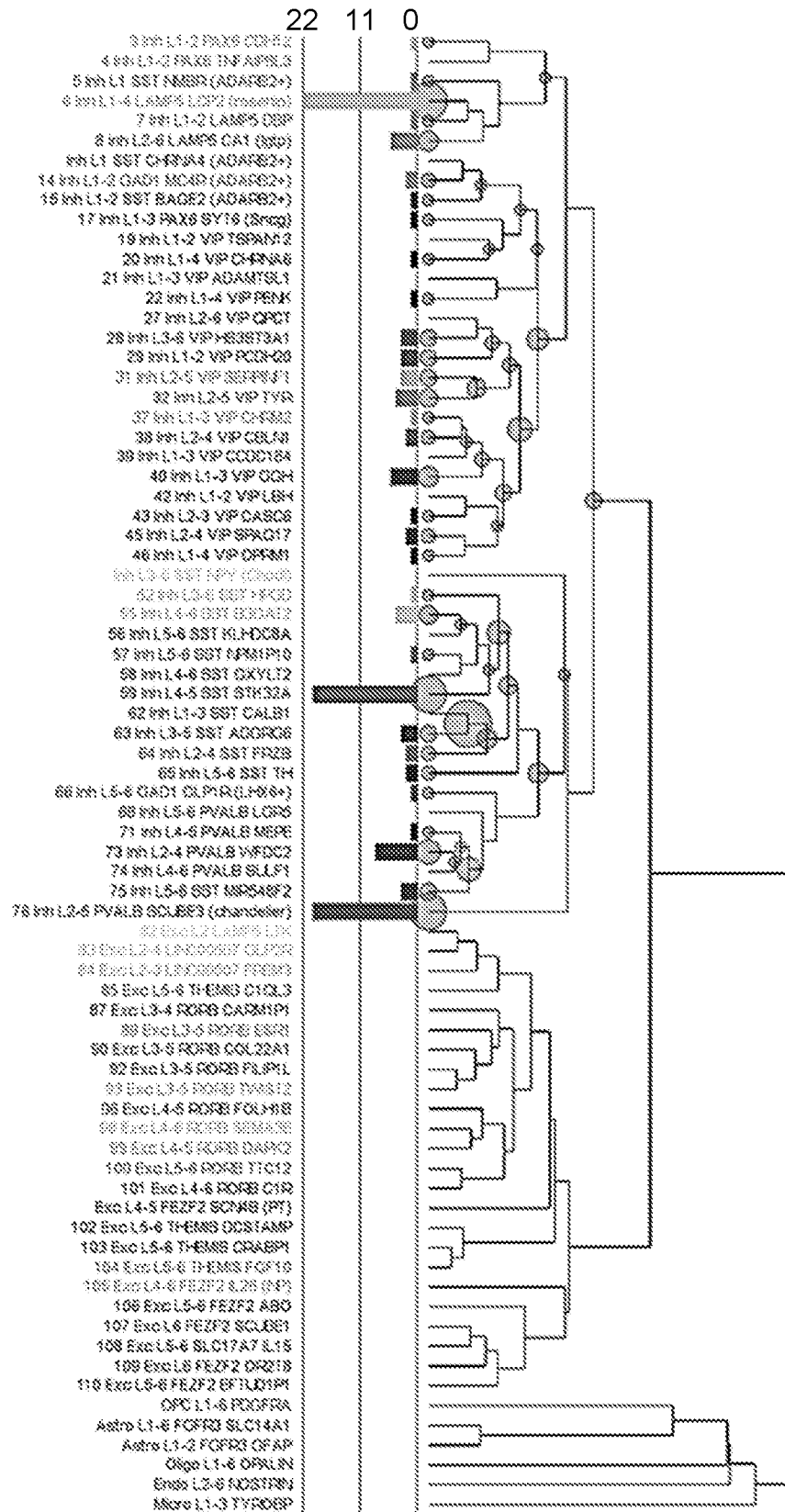


FIG. 7

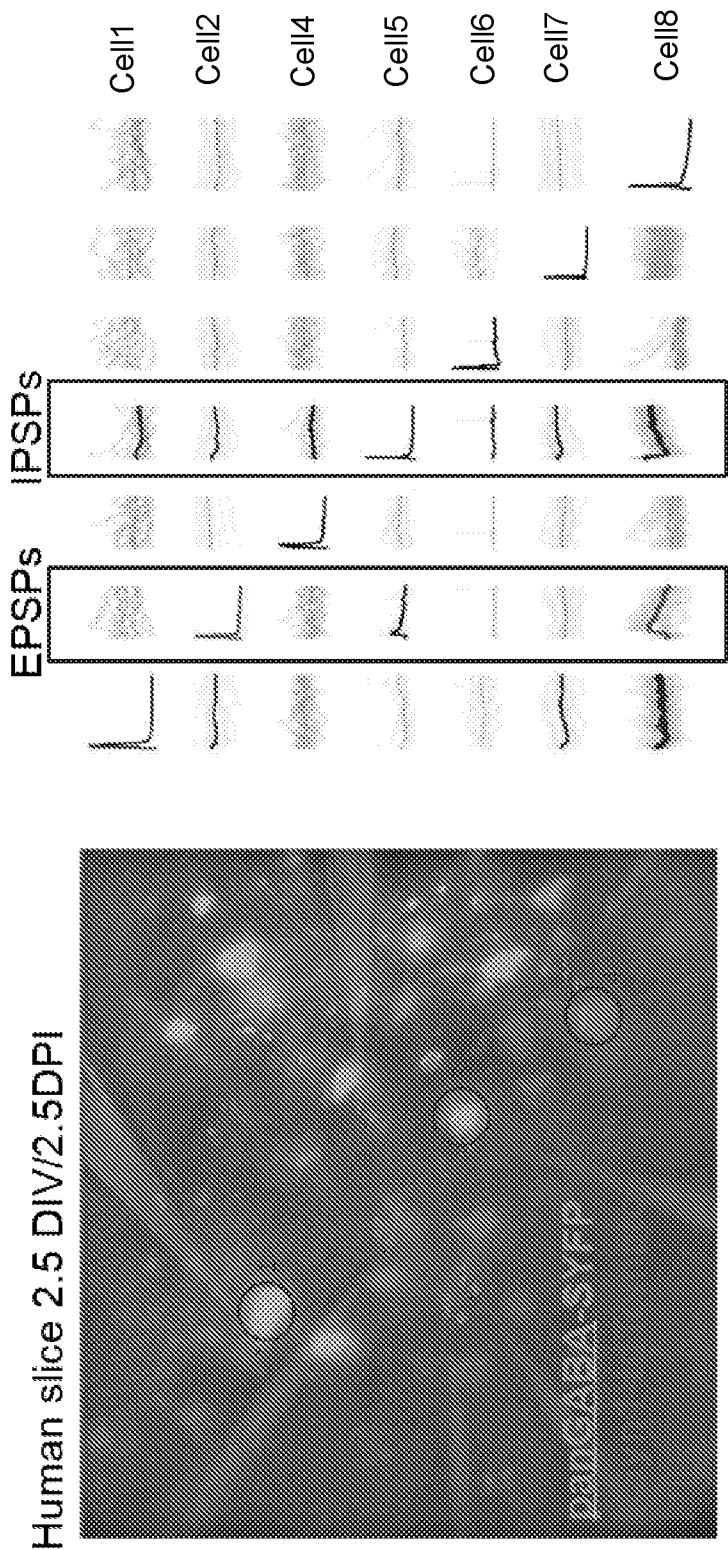


FIG. 8A

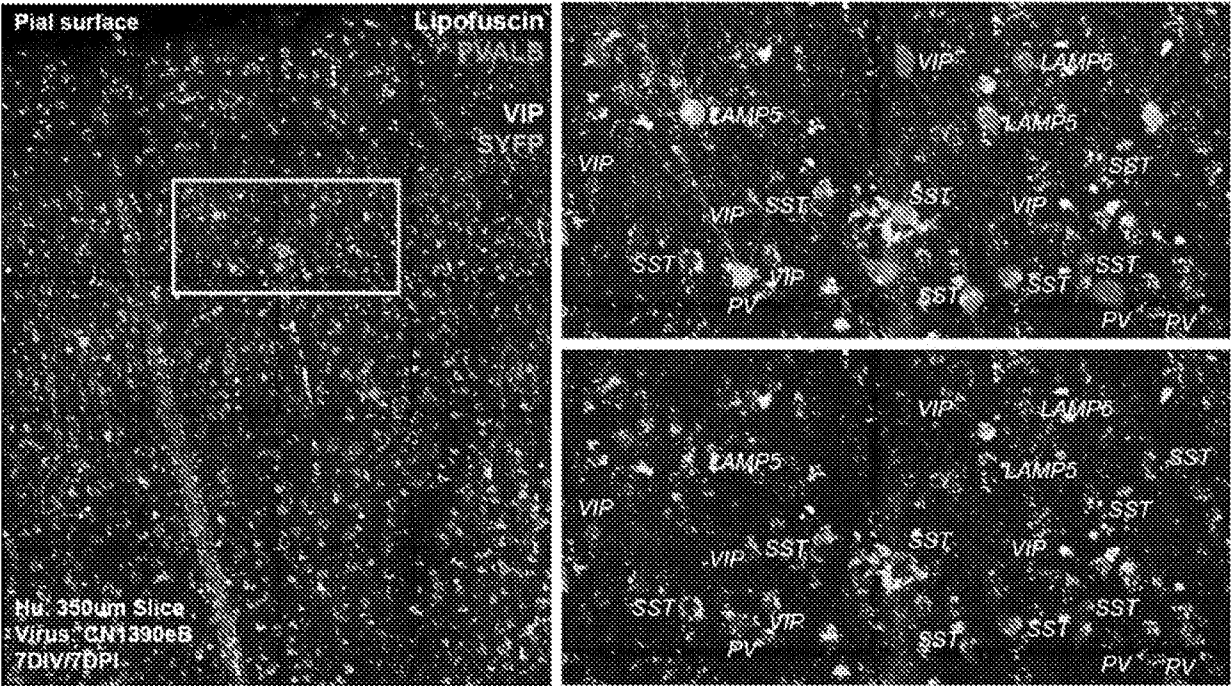


FIG. 8B

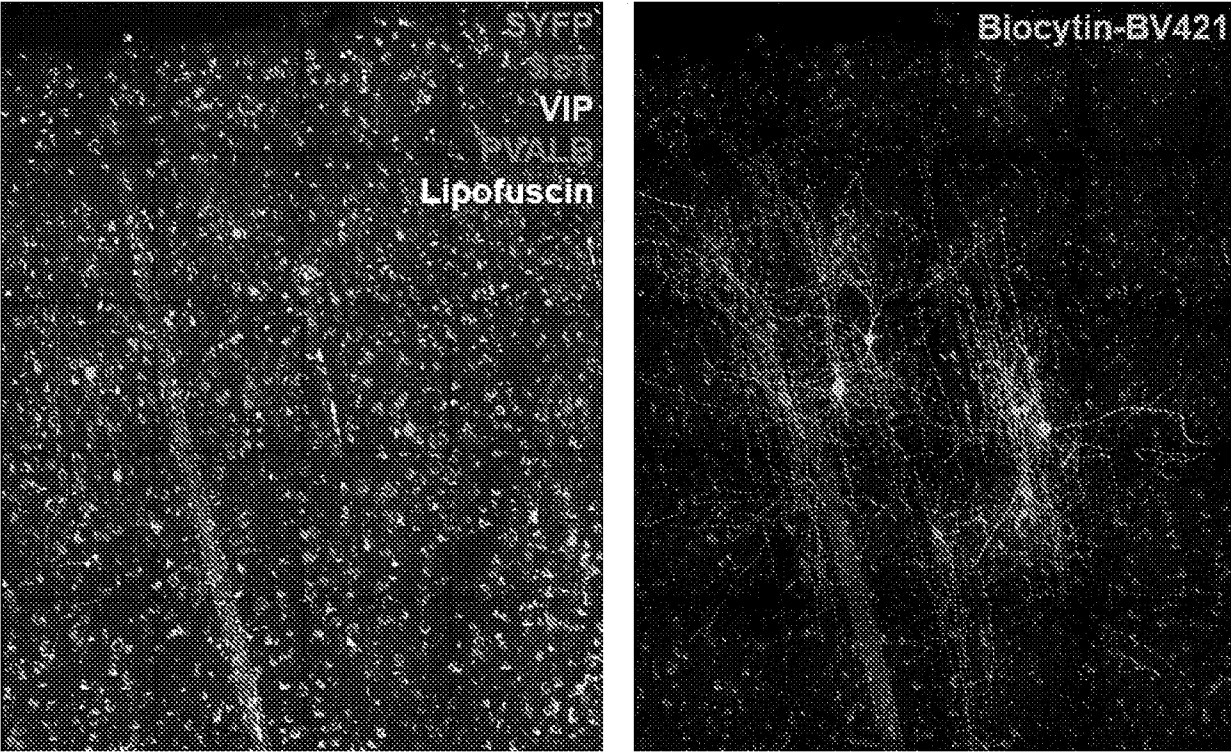


FIG. 9A

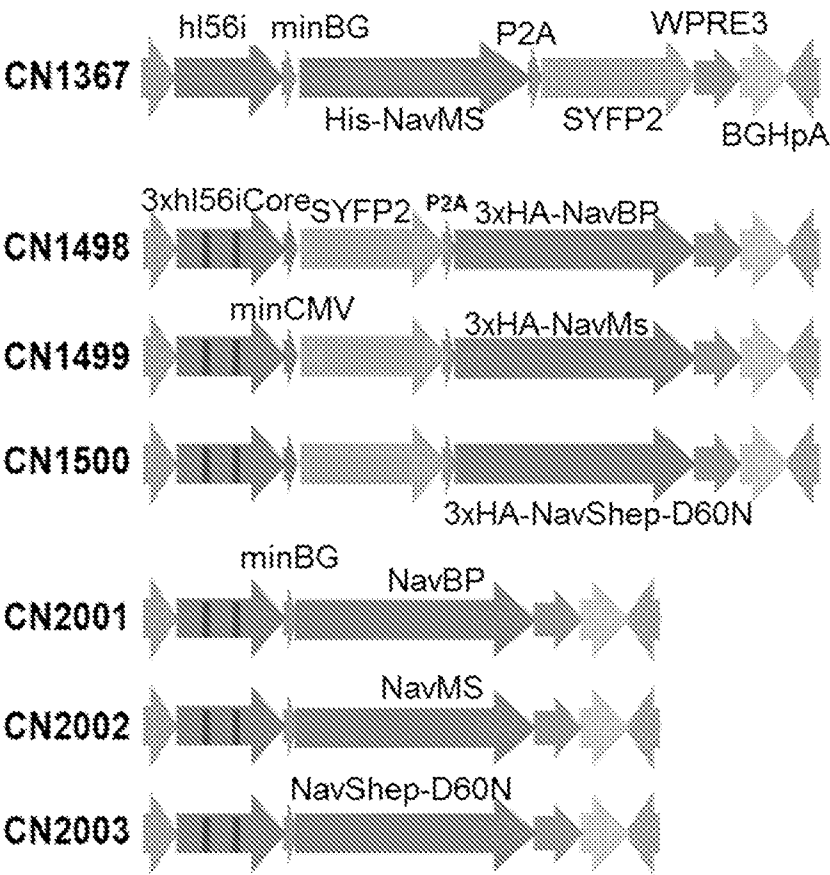
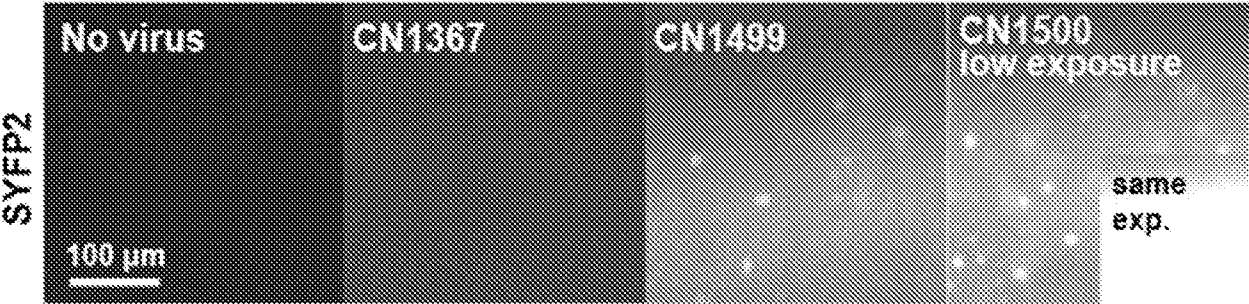
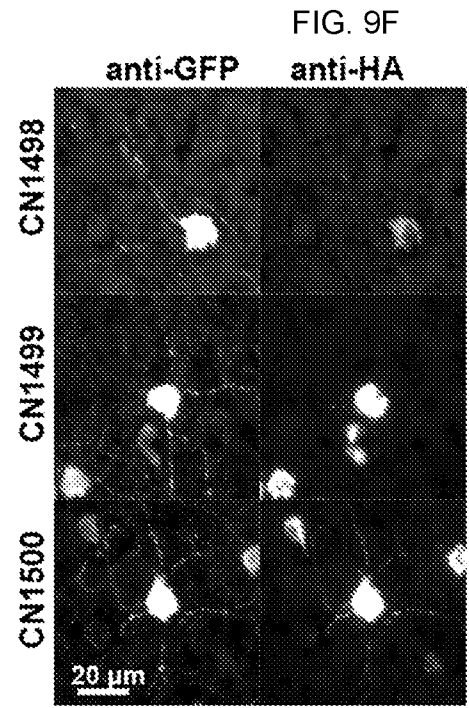
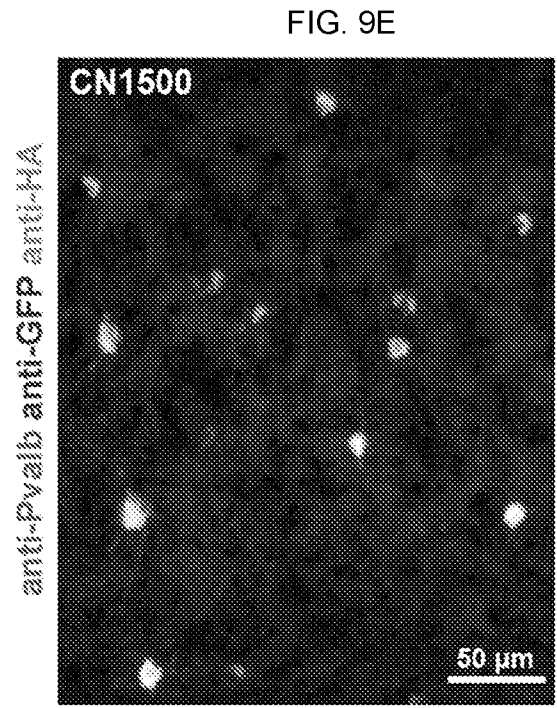
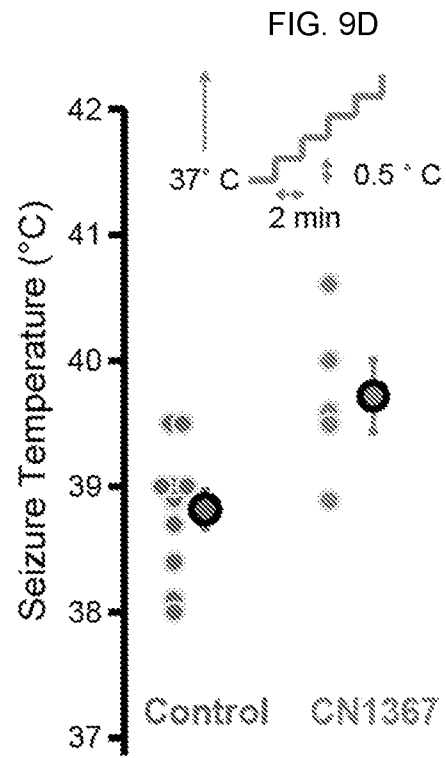
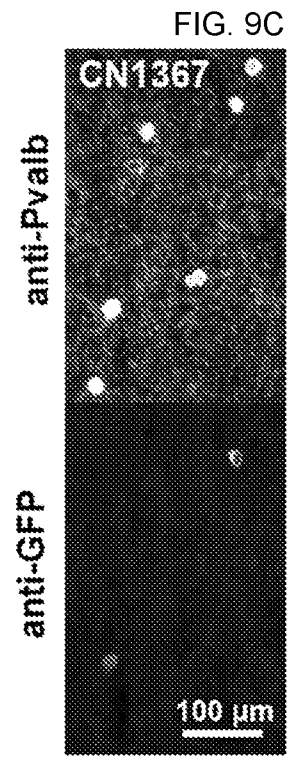


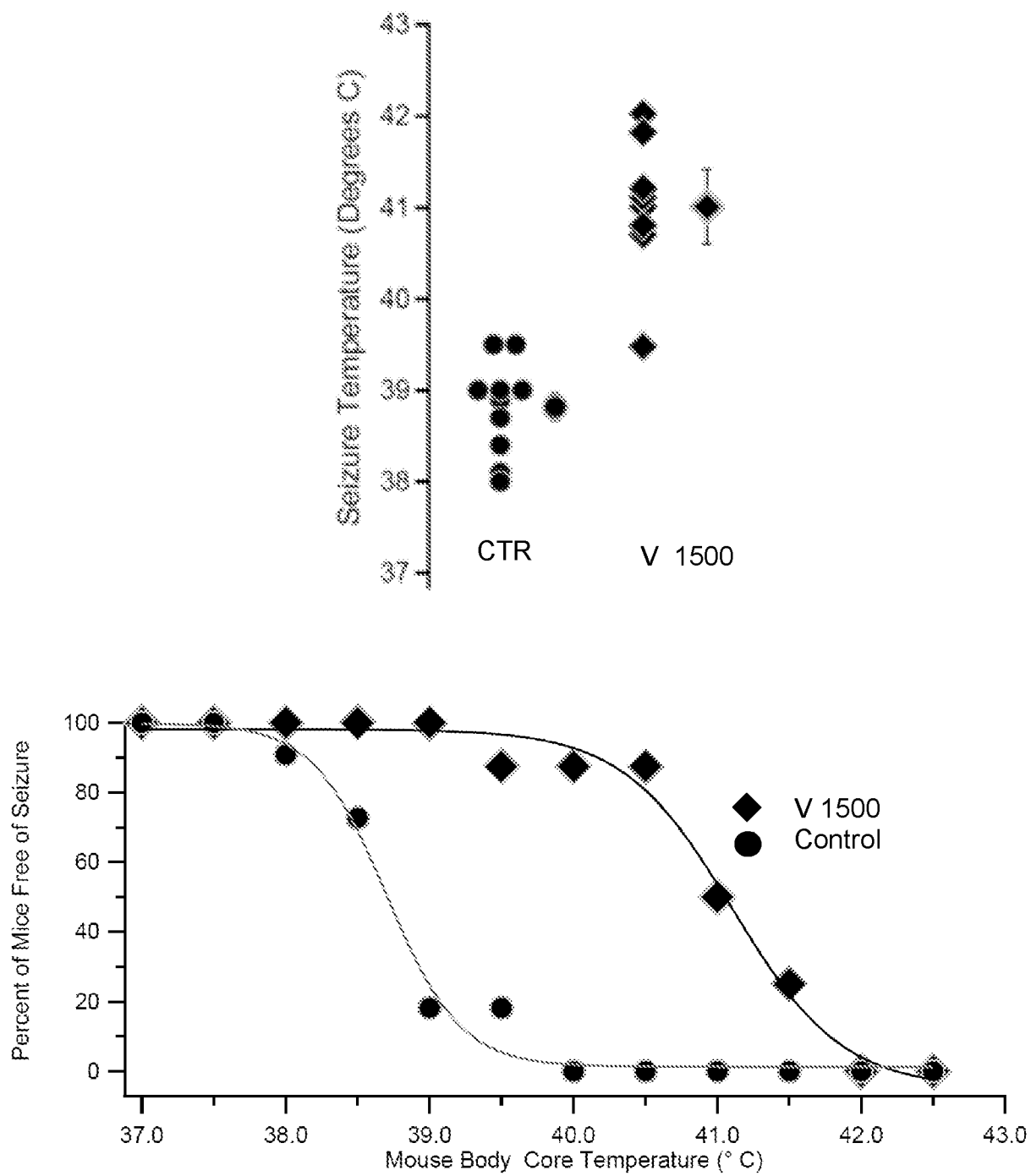
FIG. 9B





13/104

FIG. 10

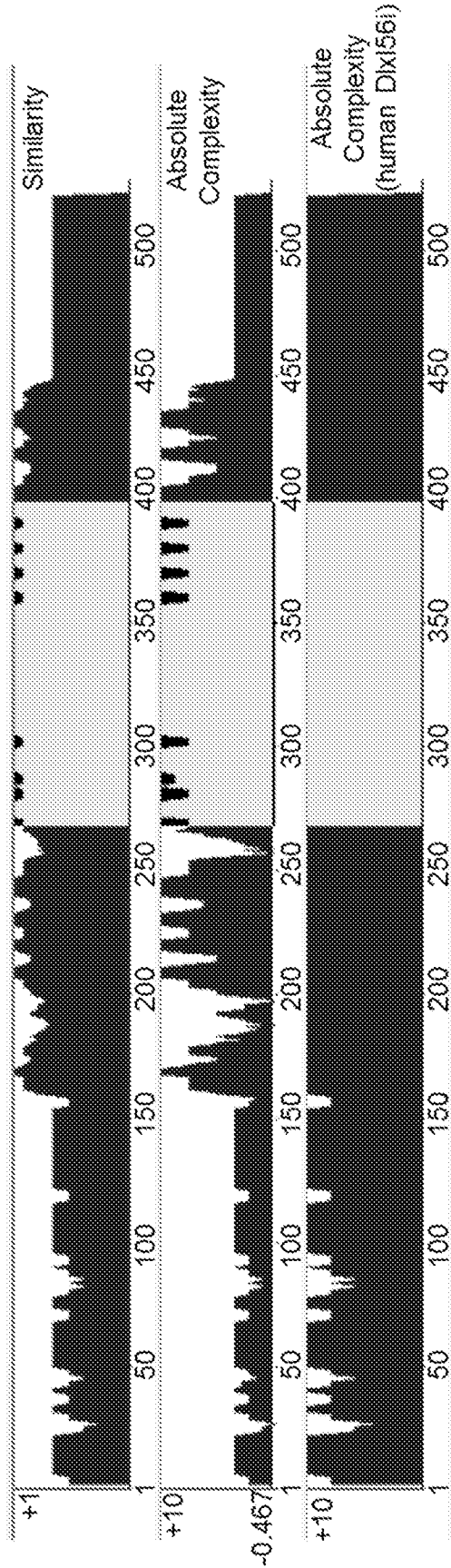


14/104

FIG. 11A

Section 1							
	(1)	1	10	20	30	40	50
human Dlx156i	(1)	TATGCACTCAGAGTGGTTGGGATGCACTGGTCAATTTTCTTAAAGAA					
mi156i (rev)	(1)	TATCACTCACTGGTTGGGATGCACTGGTCAATTTTCTTAAAGAA					
zebrafish I46i enhancer	(1)	-----					
Section 2							
	(51)	51	60	70	80	90	100
human Dlx156i	(51)	AAATTAGTGTGTTTGGAGGATGGTAAATCTCTCTAAAGCTAAATTTT					
mi156i (rev)	(50)	AAATTAGTGTGTTTGGAGGATGGTAAATCTCTCTAAAGCTAAATTTT					
zebrafish I46i enhancer	(1)	-----					
Section 3							
	(101)	101	110	120	130	140	150
human Dlx156i	(101)	AATAATTCAGCAAAAGGCGCTACAGCGTGAAGTTGTGAGGCTTTTCCA					
mi156i (rev)	(100)	AATAATTCAGCAAAAGGCGCTACAGCGTGAAGTTGTGAGGCTTTTCCA					
zebrafish I46i enhancer	(1)	-----					
Section 4							
	(151)	151	160	170	180	190	200
human Dlx156i	(151)	TCTTTTCTGACATTATTAATTTTGGTTAAATTTTCAATTITTAAGTCTACGT					
mi156i (rev)	(150)	TCTTTTCTGACATTATTAATTTTGGTTAAATTTTCAATTITTAAGTCTACGT					
zebrafish I46i enhancer	(1)	-----ACATTGGAATTTTGAATAAATATCCCMAG--CCTTCACTCTC					
Section 5							
	(201)	201	210	220	230	240	250
human Dlx156i	(201)	CTCGCAATTTTGTGATGAATAACAAATAATTTTCTCTTTTGTTTTCC					
mi156i (rev)	(200)	CTCGCAATTTTGTGATGAATAACAAATAATTTTCTCTTTTGTTTTCC					
zebrafish I46i enhancer	(40)	CTCGCAATTTTGTACATGAATAACCAATAATTTTCAAGCTTTTGTTCGT					
Section 6							
	(251)	251	260	270	280	290	300
human Dlx156i	(251)	CTTTTCTGCTTCTGAATCAAAATAAAGATGCTTTTCTAGTATTAAAAGTG					
mi156i (rev)	(250)	CTTTTCTGCTTCTGAATCAAAATAAAGATGCTTTTCTAGTATTAAAAGTG					
zebrafish I46i enhancer	(89)	CTTTGCCACCTTCAAAATCAAAATAAAGATGCTTTTCTAGTATTAAAAGTG					
Section 7							
	(301)	301	310	320	330	340	350
human Dlx156i	(301)	GAGAGAAATTACAGGTAAATTATCTTTGACGGTAAAAAAGCTGTAAATCAGC					
mi156i (rev)	(300)	GAGAGAAATTACAGGTAAATTATCTTTGACGGTAAAAAAGCTGTAAATCAGC					
zebrafish I46i enhancer	(137)	GTAGAGAAATTACAGGTAAATTATCTTTGACGGTAAAAAAGCTGTAAATCAGC					
Section 8							
	(351)	351	360	370	380	390	400
human Dlx156i	(351)	GGGCTACATCAAAAATTACCTTAATTATGCTGCAATTTATGAGAATGGAA					
mi156i (rev)	(350)	GGGCTACATCAAAAATTACCTTAATTATGCTGCAATTTATGAGAATGGAA					
zebrafish I46i enhancer	(187)	GGGCTACATCAAAAATTACCTTAATTATGCTGCAATTTATGAGAATGGAA					
Section 9							
	(401)	401	410	420	430	440	450
human Dlx156i	(401)	AAAAACCTCTCTGTGGATAAAACCTTAAATTGTCCCAATCTCTGCTTC					
mi156i (rev)	(400)	AAAAACCTCTCTGTGGATAAAACCTTAAATTGTCCCAATCTCTGCTTC					
zebrafish I46i enhancer	(237)	AAAAACCTCTCTGTGGATAAAACCTTAAATTGTCCCAATCTCTGCTTC					
Section 10							
	(451)	451	460	470	480	490	500
human Dlx156i	(451)	AAATTGGATGGCACTGCAGCTGGAGGCTTTGTTGAGAATTGATGCTGGAG					
mi156i (rev)	(450)	AAATTGGATGGCACTGCAGCTGGAGGCTTTGTTGAGAATTGATGCTGGAG					
zebrafish I46i enhancer	(282)	-----					
Section 11							
	(501)	501	510	520			
human Dlx156i	(501)	AGGTACGAACCAAAAGTTTCAAGTAG			(SEQ ID NO: 1)		
mi156i (rev)	(500)	AGGTACGAACCAAAAGTTTCAAGTAG			(SEQ ID NO: 4)		
zebrafish I46i enhancer	(282)	-----			(SEQ ID NO: 5)		

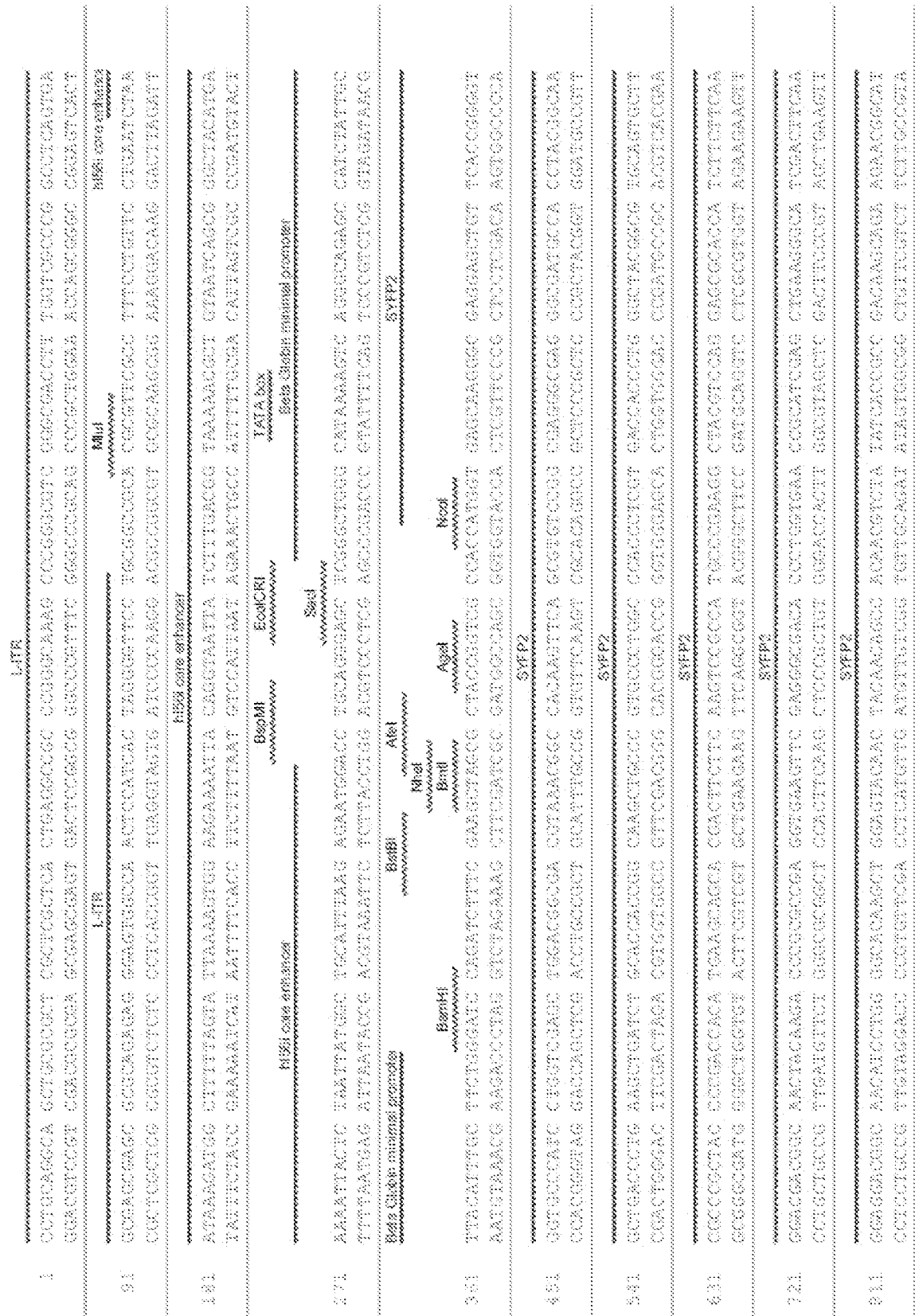
FIG. 11B

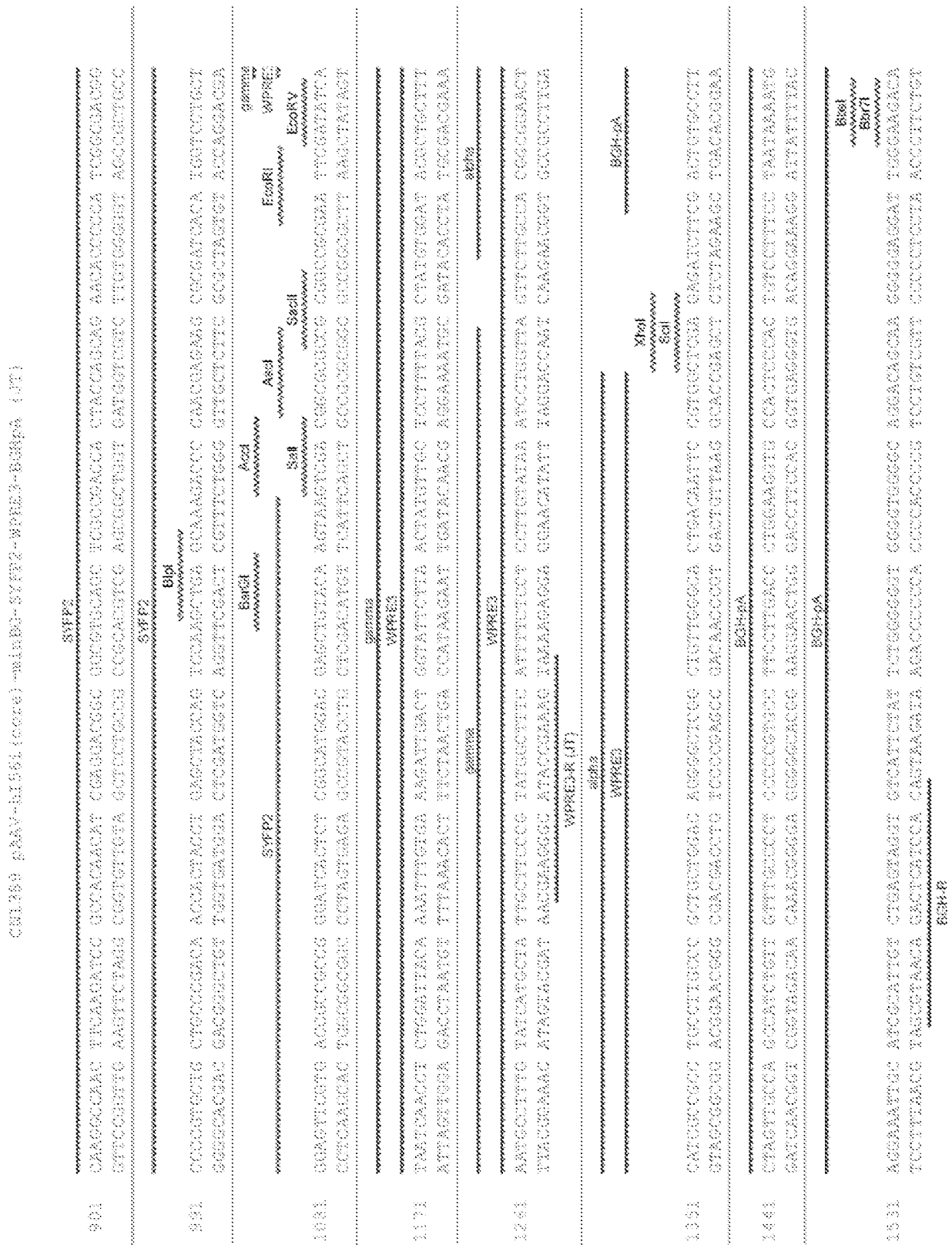


16/104

FIG. 12

CB1399 PAAV-N1561 (core)-mini-SYFP2-WRE3-BGHPA (UT)





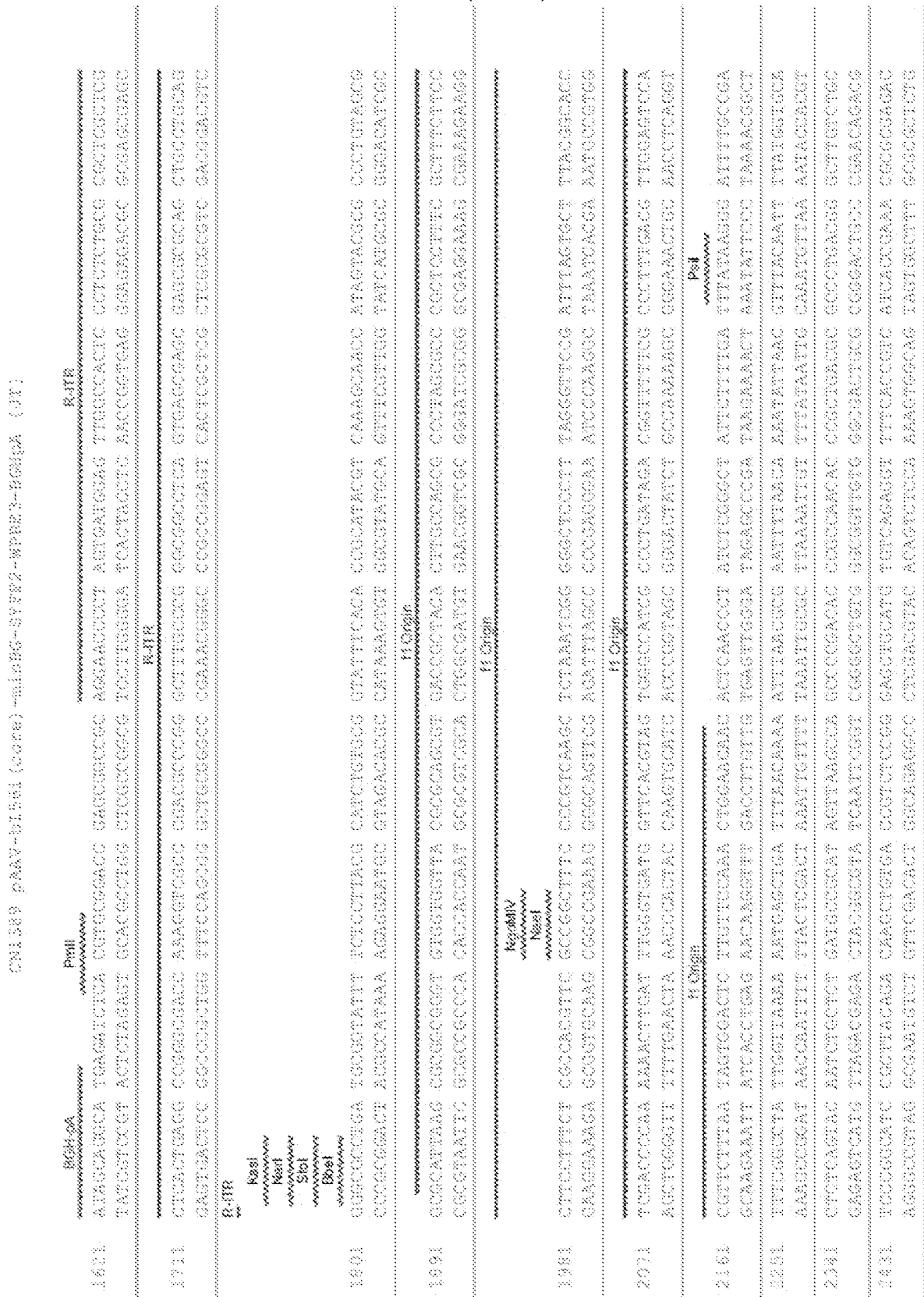


FIG. 12 (cont'd)

CN1889 PAVV-B1581 (core) -M180-SYF92-MF82-B3R1A (JF)	
2521	EcoRI1928 GAAGAGGAGCT CATTATACCG CATTATTAT AGGTTATCTT CATGATATATA AGGTTTCTT AGAGGTCAGG TGGCACTTTT CCGGGGAATG GAAGAGGAGCT CATTATACCG CATTATTAT AGGTTATCTT CATGATATATA AGGTTTCTT AGAGGTCAGG TGGCACTTTT CCGGGGAATG
2611	Zfp CTTTCCCGGA GCACTATCGG GATGAAATATA TCAATTTACA GATCTATTAT TACGAGGAA TCTGAGTCC ACCGAGAAAT CCCCCCTTAC CTTTCCCGGA GCACTATCGG GATGAAATATA TCAATTTACA GATCTATTAT TACGAGGAA TCTGAGTCC ACCGAGAAAT CCCCCCTTAC
2701	AspR TGGCGGGAAC GCGATTTCTT AAATACATTC AAATAGTAT GCGTCAAGA GACATATACC CAGATAAAG CTTCAATATAT TGGCGGGAAC GCGATTTCTT AAATACATTC AAATAGTAT GCGTCAAGA GACATATACC CAGATAAAG CTTCAATATAT
2791	AspR ACGCGCTTG GCGATGAGCA AATAGAGUS TTATAGTAG TTATACATA GCGAGTACT CIGTTATTGG GACTATTATC GAGGTTATTA ACGCGCTTG GCGATGAGCA AATAGAGUS TTATAGTAG TTATACATA GCGAGTACT CIGTTATTGG GACTATTATC GAGGTTATTA
2881	AspR ATTGAAAAG GAGGUTATG ACTATGAGC ATTTCGGT CCGCTTAT CCGCTTTTG GCGATTTG CDTTCCTGTT TTGCTTCAC ATTGAAAAG GAGGUTATG ACTATGAGC ATTTCGGT CCGCTTAT CCGCTTTTG GCGATTTG CDTTCCTGTT TTGCTTCAC
2971	AspR TACTTTTTC CTCTCATAC TCATAGCTTG TAAAGGACCA GCGGGAATTA GCGGAAATAC GCGGAAATAC GGAAGGACAA AAGAGAGUS TACTTTTTC CTCTCATAC TCATAGCTTG TAAAGGACCA GCGGGAATTA GCGGAAATAC GCGGAAATAC GGAAGGACAA AAGAGAGUS
3061	AspR CAGAAACGT GGTGAAGSHA AAGATGCTG AAGTCCAGT GGTGACAGA GTGGGTACA TCGAATGGA TCGAAGAGC GGTAGAGUC CAGAAACGT GGTGAAGSHA AAGATGCTG AAGTCCAGT GGTGACAGA GTGGGTACA TCGAATGGA TCGAAGAGC GGTAGAGUC
3151	AspR GTCTTTGGA CACCTTCAT TTCTAGAG TTCTAGTCA GCGAGTCTT CAGCCTATCT ACTTATCTT ACGTCTTUS CCACTCTAGS GTCTTTGGA CACCTTCAT TTCTAGAG TTCTAGTCA GCGAGTCTT CAGCCTATCT ACTTATCTT ACGTCTTUS CCACTCTAGS
3241	AspR TTGACAGTIT TCGCCCGAA GAGCTTTG CAGTATGAG CAUUTTTAA GTCTCTAT GTGGGCTT ATTATCTT ATTGAGGUS TTGACAGTIT TCGCCCGAA GAGCTTTG CAGTATGAG CAUUTTTAA GTCTCTAT GTGGGCTT ATTATCTT ATTGAGGUS
3331	AspR ACCTGTAAA ACGGGGCTT GTTGAAGA GTTGTACT GGTAAATTT CAGAGGATA CAGGTTTCA TAATAGGGA TAAGTGGG ACCTGTAAA ACGGGGCTT GTTGAAGA GTTGTACT GGTAAATTT CAGAGGATA CAGGTTTCA TAATAGGGA TAAGTGGG
3421	AspR GCCAGACAGA ACTCGCTCG CCGATACACT ATTCTAGAA TCACTTGGT GACTATCAC AAAGCATCTT ACCGATGGA GCCAGACAGA ACTCGCTCG CCGATACACT ATTCTAGAA TCACTTGGT GACTATCAC AAAGCATCTT ACCGATGGA
3511	AspR CCTTTCTCTT TGAACAGAG GCGATCTCA TAGAGCTCT ACTGACAGAA CCGATGCTG CTCATCTCTT TTCTCTAGAA TCGCTACCTT CCTTTCTCTT TGAACAGAG GCGATCTCA TAGAGCTCT ACTGACAGAA CCGATGCTG CTCATCTCTT TTCTCTAGAA TCGCTACCTT
3601	PvuI TGACAGTAA AGAATATAG AATGCTGGA TAACATAG TATATACACT GCGGCAACT TACTCTGAC AAGCATGGA GCAAGAGAG TGACAGTAA AGAATATAG AATGCTGGA TAACATAG TATATACACT GCGGCAACT TACTCTGAC AAGCATGGA GCAAGAGAG
3691	AspR ACTGTCTTC TCTATATAG TCAAGCTCT ATTGCTCT ACTATCTGCA CTTGCTTGA ATGAGAGCT TTGCTAGCT CCGCTCTTC ACTGTCTTC TCTATATAG TCAAGCTCT ATTGCTCT ACTATCTGCA CTTGCTTGA ATGAGAGCT TTGCTAGCT CCGCTCTTC
3781	AspR AGCTAGCGC TTCTTTGAC ATCAGTGGUS ATCAGTGGUS TCGCTTGA TCGCTTGA TGAAGCTGA TGAAGCTGA CCGAGCGG AGCTAGCGC TTCTTTGAC ATCAGTGGUS ATCAGTGGUS TCGCTTGA TCGCTTGA TGAAGCTGA TGAAGCTGA CCGAGCGG
3871	AspR TGATTTGGC AAAAAGCTG TTGATCTTC TAGTACATG ACGGCACTA GCAAGCTTG GCTGCACTT AGTGGCTT GTTTGCTGC TGATTTGGC AAAAAGCTG TTGATCTTC TAGTACATG ACGGCACTA GCAAGCTTG GCTGCACTT AGTGGCTT GTTTGCTGC
3961	AspR AGCTAGCTG CAGATGCTT GAGCAATG CAGCAAGCT GCGCAACTA TCACTGCGC ACTACTGAC TCACTCTTC CCGCAAGAT AGCTAGCTG CAGATGCTT GAGCAATG CAGCAAGCT GCGCAACTA TCACTGCGC ACTACTGAC TCACTCTTC CCGCAAGAT
4051	AspR TGACAGCTG GCTATACGA CATCTTACC GCTGTGCA CCGCTTGA CCGCTTGA TGAAGCTGA TGAAGCTGA CCGCTTGA TGACAGCTG GCTATACGA CATCTTACC GCTGTGCA CCGCTTGA CCGCTTGA TGAAGCTGA TGAAGCTGA CCGCTTGA
4141	AspR GAATAGAGT GATGAGGCG GATAGAGCT CAGAGAGCT TCTGCTGCTT GCGCTTGGC CTGGCTGCTT TATGCTGAT AAATCTGAG GAATAGAGT GATGAGGCG GATAGAGCT CAGAGAGCT TCTGCTGCTT GCGCTTGGC CTGGCTGCTT TATGCTGAT AAATCTGAG
4231	AspR ATTATCTGAC CTACTCTGC CATTTCAC GTCTGCTGA AGCTGCTGC CAGAGGCGC GAGGAGGCA ATAGGACTA TTAGAGCTC ATTATCTGAC CTACTCTGC CATTTCAC GTCTGCTGA AGCTGCTGC CAGAGGCGC GAGGAGGCA ATAGGACTA TTAGAGCTC

FIG. 12 (cont'd)

CN1389 PAAV-H1561 (core)-minBC-SYFE2-WBE2-BCHpa (JT)

	BeaI GGGCGGCGG	AmplB	BmiI GAGGAGGAG	AmplB
3421	CGGGTAGGG TGGGTCGCGG GGHATCAATG CAGCACTGGG GCGAGATGGT AAGCCCTGCC STATCTAGT TATCTACAG AGGGGAGTC			
	GGGCACTGG ACCGACAGG CCAATAGTAC CTGGTGAGCC GGGTCTAGCA TTGGGAGGCC CATACATCA ATAGATGTGC TGGCCCTGAG			
3511	AGGCAACTAT GGTGTAGCA AATAGACAGA TGGCTGAGAT AGGTGCGTCA CTGATTAGC ATGGTAACT GTGAGACCAA GTTACTACTAT			
	TTGGTIGATA CCACTTGTCT TTATCTGTCT AGCGACTCTA TCCAGCGAGT GACTAATTC TACCATATGA CACTCTGTT CAATATGATA			
3601	ATATACTTGA GATTCATTGA AAACHTCAAT TTAAATTTAA AAGGATCAAG GTGAGATCC TTTTTSATTA TGTGATGAGC AAAATGCTT			
	TATATGAAT CTAACTAAT TTGGAAGTAA AAATTAAT TTCTAGATC CACTCTTAGG AAAACTAT AGGTACTGCG TTTTAGGAAA			
3691	AGCTTBAAT TTCTTCCAG TGAGGTCAG ACCCGTAGA AAGATCAAA GATCTCTT AGATCTCTT TTCTCTGTC GTATCTGCT			
	TTGCACTGAA AAGCAAGTG ACTCGAGTC TGGGCACT TTCTAGATT CCTAGAGAA CTTTAGGAAA AAAGAGCGG CATTABAGGA			
3791	GCTTGCAAC AAAAAXXXA CCGCTACAG CAGTGTCTG TTGCGGAT CAGAGCTAC CAATCTTT TCGGAGGTA ACTGCTTCA			
	CGAACTTTS TTTTCTTGGT GCGATGATC GCGACCAAC AACGSCAA GTTCTGATG GTTGAGAAA AGCTTCTAT TGAAGAAAT			
3871	GCAGAGGCA GATACCAAT ACTGTCTTC TAGTGTAGC GTAGTATGC CACCCTTCA AGAAGCTGT AGCACGCTT ACATACCTTC			
	CGTCTGGGT CTATGTCTTA TGCAGGAG ATCATCTGG CATCATCTG GTGCTAAGT TCTTGAGACA TGTGAGGGA TGTATGAGC			
3961	CTCTGCTAAT CCTGTATACA GTGATGCTG CAGTGTGCA TAATGCTT GTTAACTGGT TGGACTGAG AGGATAGTTA CCGGATAGG			
	GAGACGATTA GACATATGCT CAGCGAGGAC GGTCAAGCT ATTCAAGCA GAATGCCCCA ACTGTAGTTC TGTATCAAT GGCCTATTCC			
4051	CGCAGCGGTC GGGCTGAGC GGGGCTGCT GACACAGCC CAGCTTGGG CGAACCACT ACCCACTCA GAGATACCTA CAGCGTGGC			
	GGCTGGCCAG CCGCACTTC CCGCCAGCA CCGTGTGCG GTGCAACTC GCTTCTGGA TGTGCTTGA CTCTATGAT GTGCACTCG			
4141	TATGAGAAAG CCGCAGGCT CCGGAGGGA GAAGGCGGA CAGGTATCG GTAGCGGCA GGTCTGAAAC AGGAGAGCGC AGGAGGAGC			
	ATACTCTTC GCGGTGCGAA GGGCTTCTT CTCTCCGCT GTCCATAGC CATTGCGCT CCGAGCTTG TCTCTGCG TGTCTCTGCG			
4231	TTGCAAGGAG AAGCGCTCG TATCTTATA GTCTCTCGG GTTTCGCCAC CTCTGACTG AGGCTGAT TTGTGATGC TGTGTAGGAG			
	AAATCTGCT TTTCGGAAC ATAGAAATAT CAGGACAGC CAAGCGGCG GAGACTGAG TGCATGTA AAACATGTC AGTATGCTGC			
4321	GGCGGAGGCT ATGGAAAAAG GCGAGGAGG CAGCTCTTT AGGCTCTG GCGCTTCTT GCGCTTTTC TGCATCT			
	CGGCTCTGGA TACCTTTTTC CAGTCTTTC GCTGGGAAA TCGCAAGGAC CGGAGAGGAG AGCTGACA (SEQ ID NO: 95)			

PflI
 (SEQ ID NO: 41)
 (SEQ ID NO: 95)

233

[illegible]

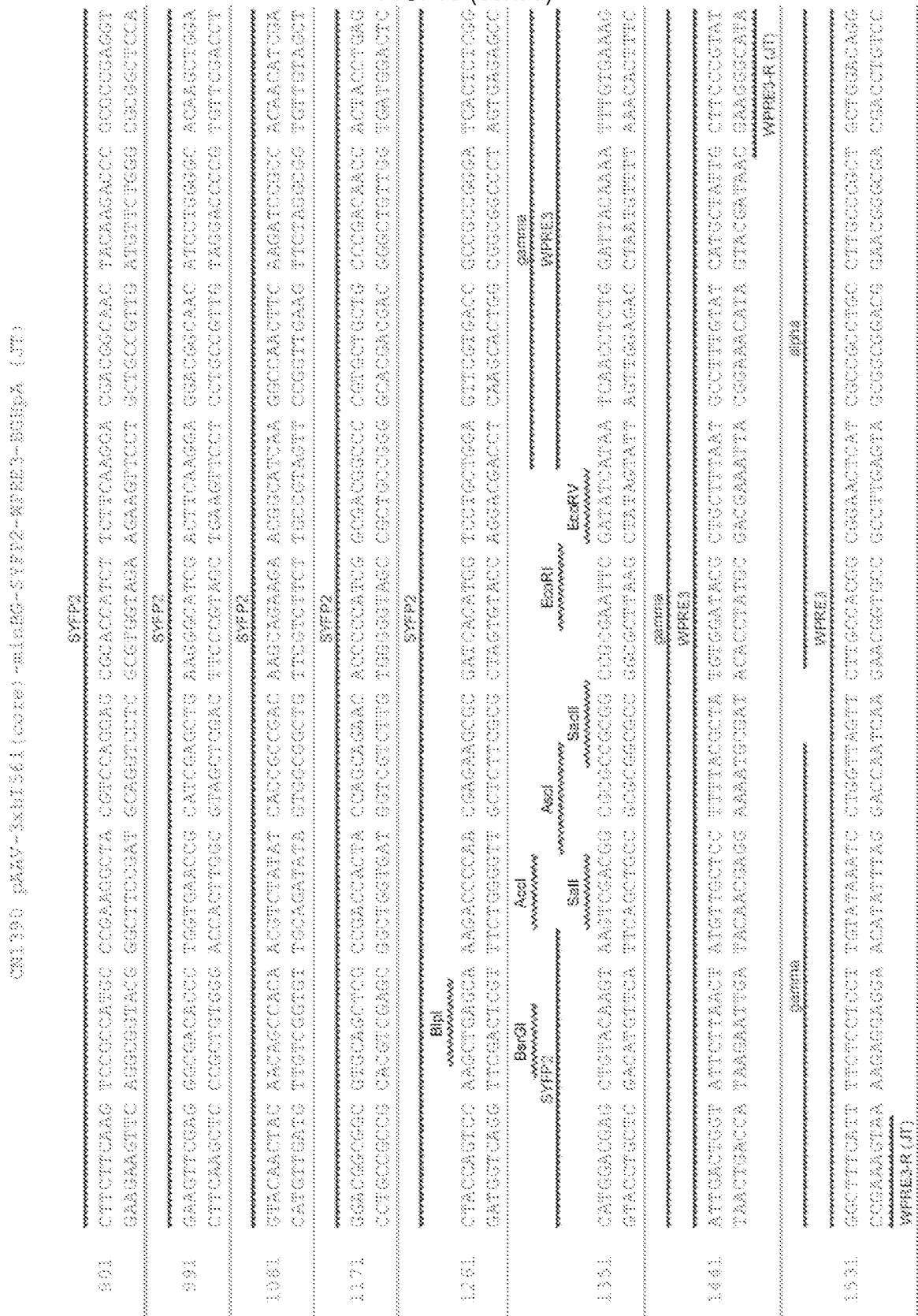


FIG. 13 (cont'd)

CH1390 pAAV-3x361561(cc.re)-mlnRG-SYFP2-NP83-BGSpA (J1)

abba WP83 B3H.pA		XbaI SbfI B3H.pA	
1621	GGCTGGGCTG TTGGGCACTG ACGATTCCCT GCTCTGAGAG ATCTTCCACT GTCCCTTCTA GTGGCAGCC AICTGTCTTT TGGCCCTGCT		
	CCGAGCCGAC ACCCGTGC ACCTGATGCG CCGAGCTCTC TAGAAGCTCA CAGCGAGAT CAGCGTGG TAGACAGAA ACCGAGAGG		
1711	CCTGGGCTTC CTGAGCTG GAAGGAGCA CCGCAGCTGT CTTTCTCTAA TAAATGAGG AATATGATC GATTTGCTG AATAGGATG		
	GACAGGAGG GAATGGGAC CTTCACGCT GAGGGAGACA GGAAGGATT ATTATGCTC TTATAGCTAG CTTACGCGAC TGAATGACAG		B3H.R
1801	ATTCTATCTT GGGGGTGG GTGGGCAAG ACAGCAAGG GAGGATGG GAGACATA GAGGCAATG GATCTACGT GGGAGCGAG		
	TACATPAGA CCGACAGCC CAGCCCTCC TTCTGTTCC CTTCTTAC CTTCTTAT CTGCTACT CTAGAGTGA CCGCTGGCTC		
1891	CGCGGCGAG ACCCTAGT GATGATTT GCACTGCT CTTCTGCGC TGGCTGCTC ACGAGCGCG GCGAGCGAA GGTGCTGGA		
	GCGGCTCC TTGAGTGA CTACTGAC CCGTCAAGG GAGACGCGG AGCGAGGAG TGGCTGCGC CCGCTGCTT CGAGGCTCT		
1981	CGCGGCGCT TTGCTGCT GCGTCAATG AGCGAGGAG CCGCAGCTG CTTGAGGCG CCGCTGCTC GGTATTTCT CTTAAGGAT		
	GCGGCGGA AGCGGCGC CCGGATGAC TGGCTGCTC GCGCTGCG CCGCTGCGC GCGCTGCTC GGTATGGA GGTATGCTA		
2071	CTGTGCTGTA TTTCACGCG CATAGCTCAA AGCAGCATA GTACGCGCC TGAAGCGG CATTAAGCG GCGGCTGCT GGTATGCTC		
	GACAGGCTAT AAGTGTGG GTATGCTAT TGTGTTGAT CAGCGCGCG ACATGCGCG GTATTGCTC CCGGCTGCG CAGCATGCG		
2161	CGAGGCTGAC CCGTCACTT CCGAGCGCG TAGCGCGCG TCTTCTGCT TCTTCTGCT CCGTCTGCG CAGCTTCTG CCGTCTGCG		
	CCTCGCACTG CGATGTCGA CCGTCTGCG ATCGGCGCG AGCAGGCG AGCAGGCG GAGAGCGG GTGCAAGCG CCGAAGCGG		
2251	GTCAGGCTCT AATCGGCG CTCCTTTAG GTTCTGAT TAGTCTTGA CCGCAGCTG CCGCAGCTG ACCGAGGAA ACTGATTTG GGTGATGCT		
	CAGTTGAGA TTATGCTTC GAGGAGATC CCAAGGCTTA ATCAGGAT GCGGCGGAG TGGGTTTTT TGAACTAAC CCACTAGCA		

FIG. 13 (cont'd)

CNI350 pXV-3ab1561(core)-minG-SYFP2-WPRE3-BURpA (JT)

2341 CACUATAGG GCGACGGCCG TGTATAGCGG PTHHHGCC TTTCAGCTTG GASTUCAGCT TCPTTAAGAG TGGACTCTTG TTCGAACTG
 GTGCTCACC CGGTGGCGGG ACTATCTGCC AAAAGSCGGG AACTGCGAAC CTGAGTTGCA AGAATTATC AACTGAGAAC AGGTTCGAC

B Origin
 PstI
 2431 GAGCAGACAT CAACCCATAC TCGGCTATT CTTHGCTTT ATACGGGATT HGGCGATT CGGCTATAG GTTAAAAAT GAGCTGATTT
 CTGGTGCGA GTTGGGATAG AGCGGATTA GAGAGCTAAA TATTCCTTAA AAGGCTAAA GCGGATAGC CATTTTTTA CTGAGCTAAA

2521 AGCAAAAGT TAACGCGAAT TTHACAAAA TATTACGTT TACAATTTA TCGTGCATC HCASTACAA CTGCTCTGAT GCGGCTATG
 TTGTTTTTAA ATTGCGCTTA AAATTTTAT ATAATTGCA ATGTTTAAAT ACCACGTGAG AGTCATGTTA GACGAGACTA CCGCTATATA

2611 TAAGCCAGCC CCGACACCGG CCAAGATAG CTGACGCGCC CTGAGGAGCT TGTCTGCTC DGGCATCGG TTACAGACAA GTTGTGACG
 ATTGCGTCCG GCGTGTGCGG GATTGTGCGG GACTGCGCGG ACAGACAGAG CCGGTAGCGG AATGTCTGTT CGACACTGAC

EcoRI
 2781 TCTCGGURAG CTGCAAGTGH CARAGTTT CACGGTCATC ACGGAACCG GCHAGACAAA AGGGCTGCG GATACGCTA TTTTTHAGG
 AGAGGCGTC GAGGTACAGA GTCTGAAA GTGGCAGTAG TGGCTTTGCG GGTCTGCTT TCCCGAGCA GTATGCGAT AAAATATAC

Zet
 2791 TTAATGTCAT GATTAATATG GTTTTCTG CCGCAGCTGG CACTTTTCCG GGAATGTC GCGGACCGG TATTTTCTTA TTTTCTTAA
 AATTACATA CTATTATATC CAAGATCT GCACTCCACC GTGAAAGCC CTTTACAGG CCGCTGCGG AAAACAAT AAAAGATT

AspR
 2881 TACATTGAAA TATGTAACG CTGATAGAC AAATAGCTG ATAAATGCT CAATAATAT GAAAGAGAA GATATGAT ATTGACATTT
 AEGTAAGTT ATACATAGG GATATCTG ITAATGGAC TATTACGAA GTTATATATA CTTTTCGCT CCAATACCA TAAGTGTGAA

AspR
 2971 TCGGTGTGG CATTATCCC TTTTTCGG CATTTCGCT TCTGTTTTT CCGACCGAG AAGCGTCTT GAACTGAAA GTTCTGTGAG
 AGGCACAGG GGAATAGCG AAAAGCGC GTAAACCGA AGACGAAA CCACTGCTC TTGCGACCA CTTTCTTT CTACGACTC

AspR
 3061 ATCAGTGGG TCGACAGTG GGTATGATG AACTGATCT CACAGCGGT AGATGCTG AGAGTTTC CCGCAGAA CTTTCTGAA
 TATGACAGC AGGCTGTC CAAATGAG TGAACGAAA GTTGTGCGA TTATAGAAC TGTCAAGG GGGCTCTT GCAAGAGTT

AspR
 3151 TCATGAGAC TTTTAAAGT CTGCTATG GCGCGTAT ATCCGTAAT GACGCGGCG AGAGCAAT CCGTCCCG ATACACTAT
 ACTATCTCTG AAANTTCGA GACGATAC CCGGCTATA TAGGCTATA CCGCTGCTG TTCTGCTGA GCGAGCGG TATGTAAG

AspR
 3241 CTCAGATGA CTGTGTTGAG TACTGCTAG TCACAGAAA GATCTTACG GATGCGTGA GATTAAGAG ATTATGAT GTTGTGATA
 GACTCTTACT GAACCAACTC ATGATGCTC AGTCTCTT CTAGATGC CTACCTACT GTATCTCT TATAGCTCA CGACGTAT

SacI

FIG. 13 (cont'd)

CU1300 5AAV-3xHIS5.1(core)-HLRG-SYTF3-W2823-BORPA (3T)

```

3431  CCAATGAGTA TAAAGATGG CGGANTATG TTCTGACAC GATCGAGCA CGGAAAGAGC TAAAGATTT TTGCGAAGC ATGGGGATC
3441  GGTACTACT ATTGCAATC CGTTGATATG AAGATGTTG CTAAATGCT GATCTCTGG ATGATGAMA AAGATGTTG TACCACTAG
3451  ATGTAAGTGG CATTGATCG TGGGATCGG AGCTGATGA AGCTATGCA ACCGATGAGC GTGACATGAC GATGCTGTA GGAATGGGA
3461  TACATTCAGC GGATCTAGCA ACCCTTGCC TGGATGACT TGGATGCTT TTGCTCTG CACTGTGTC CTACGATG GATAGGCTT
3471  CAGCGTTGCG CAAGCTATG ACTGGGCAAC TACTTACTCT AGTTTCGCG CAACATTTA TAACTGATG GAGCGGAT AAGCTGGCG
3481  GTTCGAGCG GTTGATATG TCACTGCTG AAGATGACA TGGAGGCGC GTTGTATTT ATCTGACTA CTGCGCTA TTCAAGGTC
3491  GACCACTTCC GCGCTGCGC CTTCGCTGCT GCTGCTTAT TCTGATATA TCTGAGCGC GTGAGGCTGG GTCTGGGCT ATCATGTCAG
3501  CTGCTGACA GCGAGCGG GAGGCTGAC CGAGGCAATA AGCATATTT AGACTGCGC CACTGCGC CAGATGCA TACTAGGTC
3511  CACGCGGCGC AAGGCTAGC CGCTGCTGTA TCGTATGAT CTACAGAGC GGGATCGAGC CACTATGGA TGAAGCAAT AAGCAGATCG
3521  GTGATCGCGC TCTCCATC GGGAGGCAAT AGCATGATA GATGCTGTC GCTGCTGTC GTTGATGCT ATCTGTTA TTGCTGAGC
3531  CTGAGATAG TCGCTGAGC ATTAGAGCT GGTACTGTC AAGCAATTT TACTGATATA TACTTATG TGAATTAAA CTCATTTTT
3541  GACTCTATCC AGGAGTGC TAAITGCAA CCAATGACAG TCTGCTCAA ATGATGAT ATGATGCTA ATTAATTT GAGTAAAAA
3551  AATTAAGAG GATCTAGG AGAGCTCTT TTGATATCT CATGACAAA ATCTGTCAC GTGAGCTTC GTTCACCTA GCTTACAGC
3561  TTAAATTTT CTAGATGAC TTCTAAGAA AACTATGTA GTAGTGTG TAGGCTTC GACTGAAAG GAGATGAT GAGAGCTGG
3571  CAGTAAAGA GATCAAGGA TCTCTGAG AAGCTTTT TCTGCGCTA AGTCTGCT TGAACAAA AAGACAGC CTAGAGCG
3581  GCGCTTTT CTAGTTCT AGAGAGCTC TAGGAAAAA AGAGGCAAT TAGAGGCA AGTTGTTT TTGCTGCG GATGCTGCG
3591  TCGTTGTTT GCGGATGAA GAGTATGAA GTTTTCTC GAGTAAAT GGTTCAGCA GAGGCAAT ACCAATAT GCTGCTGAT
3601  ACCAAGACA CGGCTAGT CTGATGCTT GAGAAAGG CTTCGATGA CGGATGCT CTGCTGCTA TGTATGTA GAGGAGTTC
3611  TGTGCTGTA GATGAGGAG CACTTCAGC ACTCTGAGC AGCTGATCA TACTGCTC TGTATGCT GTTACAGT GCTGCTGCA
3621  AATGCGCAT CAAGCGGCG GTGATGCT TGAAGATCG TGGGATCT ATGAGAGAG AGGATAGGA CATGCTGC GAGGCGCT
3631  GTGGGATAA GCGGCTCT ACCGCTGCT ACTCAGGCG AATATGCG GATAGGCG AGCTGCGC CAGAGCGG GATGCTGCA
3641  CAGCGTAT CAGCACAAA TGGGCGACC TGAATGCTG TATCATGCG CTATGCTGCT TATGCTGCG GATGCTGCG GAGGCTGCT

```

26/104

FIG. 13 (cont'd)

CN1390 pAAV-3xHIS6L (core)-HISG-SYEP2-WHE3-BGHPA (JT)

```

4321 CACAGCTTAG GTTGGAGCGA AGAAGCTTACA CCGAAGTGG AGAAGCTTAT GAGAAAGCGC CCGGTTTGGC GAGAGGAGA
      GTGTGGGTC GAACTGGCT TCGTGGAGT GGTGAGATC PATGATGTC GCACTGATA CTTTTCGCG GTGGAGAGG GTTGGTCTT
      pUC ori
4411 AGCGGACAG GTATCCGGA AGCGGACAG TCGGAGAGG AGGAGCTTC CAGGAGAAA CGCTGTAT CTGTAAGTC
      TCGGCTGTC CATAGGCAAT TCGCGTCCC AGCTTGTC TCTGGGTGC TCGCTGAG GTCCGCTTT GCGGACATA GAAATATGAG
      pUC ori
4501 CTTGCGGATT TCGGACCTC TCACTTGAG GTGATTTT GTGAGCTCG TCGGSSGC GGAGCTTG GAAAGCGC AGCAGCGAG
      GACAGCCCA AGCGTGGAG ACTGACTCG CAGCTAATAA CACAGAGAG ACTCCCCG CTCGGAAC CTTTTCGCG TCTTTCGAC
      pUC ori
4591 CTTTATTAG GTTCTGGCC TTTTGTGSC GTTTTGCICA CATGT
      pUC ori
      p31
      (SEQ ID NO: 42)
      (SEQ ID NO: 96)

```

27/104

FIG. 14

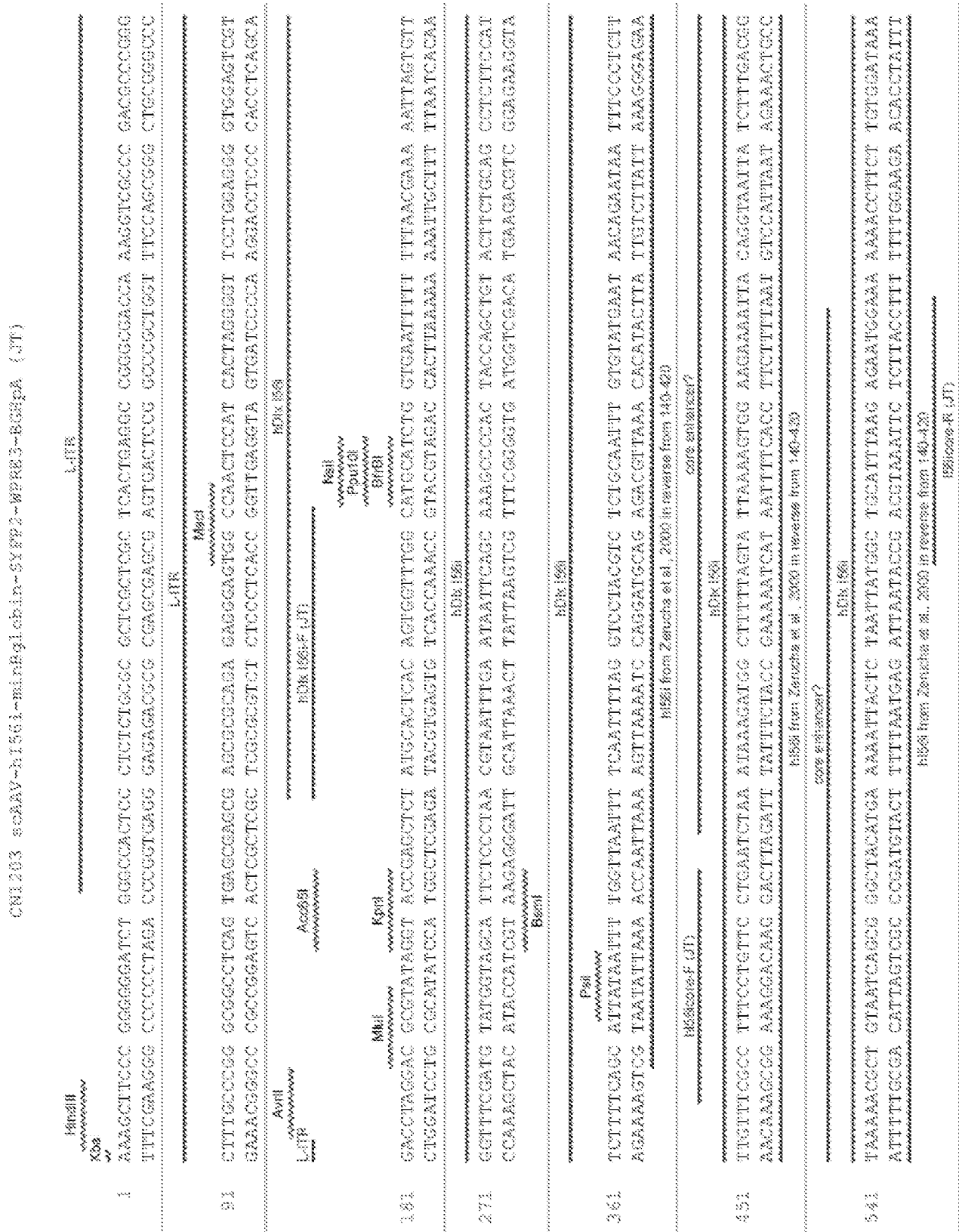


FIG. 14 (cont'd)

CH1203 scAAV-h1561-minidglobin-SYFP2-WF22-EGFPa (JT)

NDx 1581									
	NDx 1581		TATA box		Beta Globin minimal promoter		transcriptional start		BstBI XbaI
	NDx 1581	AatI RheI BstBI	Agel XbaI	NotI XbaI	Beta Globin minimal promoter	transcriptional start	BstBI XbaI		
631	AACCCTAAAT TCTCCCAAT CTCCTGCTCA ATTGGATGG CAGTCAGCT GAGGCTTTC TTGAGAAITG ATCTGGGGA GCTACGAACC TTGGAAITTA ACAGGGGTTA CAGACGAAT TTAACTTACC GTGACCTGGA CTTCCGAAAC AAGTCTTAACT TAGGACCTCT CCAATGCTTGG B856 from Zentgraf et al. 2000 in reverse from 142								
721	CAAAAGTTTCA CAGTAGGAG CTCGGGCTGG GCATATAAGT CAGGGGACAG CCATCTATTG CTTGCAITTC CTTCTGGAT CCGATCTCTT GTTTCAAAAT GTCTATGGCTC GAGCCCGACC CGTATTTCAT GTCCCTCTTC GGTAGATAAC GAACTATAAC GAAGACCTTA GGTCTAGAAA B856 from Zentgraf et al. 2000 in reverse from 142								
811	CGAAGCTTAC GCTACCGCTC GCGACCGATG TGAGGAGGAG CGAGGAGCTG TTACAGCGGG TGTGACCAT CTTGTGGAG CTGAGCGGGG GCTTGGATCG CGATGGCCAG CGGTGGTACC ACTGGTTCCC GCTGCTGAC AAGTGGCCC ACCAGGCTTA GACCACTCTC GACCTGGCCG SYFP2								
901	ACGTAAACGG CCACAAGTTC AGCGTGCTCG GCGAGGCTGA GGGGATGTC ACCTACGGCA AGCTTACCTT GAAGTGCAT TGCACCAACG TGCATTTCG CGCTTTCAG TCGCACAGGC CGCTCCCTCT CCGCTACCGG TGGATGCTCT TGTACTGGA CTTGCACTAG ACCTGGTGGC SYFP2								
991	GCAAGCTTCC CGTCCCTTGG CCACCTCTCG TGACCACTCT GGTGACGCT CTGAGTCTCT TCGCTGCTTA CCCCAGCCAC ATGAGACGAG CGTTGGAGGG GCAAGGAGCC GGGTGGAGC ACTGGTGGCA CCGCATTCGG CAGCTACGGA AGCGGCTCAT GGGCTGCTG TACTTGGCTG SYFP2								
1081	ACGACTTCTT CAAGTCCGCC ATGCCCGGAG CTTACGTTCA GAGGCGACC ATCTTCTTCA AGAGGACGG CCACTACAG ACCCGCGCCG TCTTGAAGAA GTTCAAGGCG TAGCGGCTTC CGATGZAGGT CTTCCGCTGG TGAAGAGAGT TCTCTCTGCC GTTCAATCTC TGGTCCGCGC SYFP2								
1171	AGGTGAAGT CGAGGGGCGC ACCCTGCTGA ACCGATGGA GCTGAGGAGG ATCGACTTCA AGAGGACGG CAAGTCTCTG GGCACACAGC TCCACTTCAA GCTCCCGGTC TGGGACCACT TGGCGTACCT CGACTTCCCG TACGTGAAT TCTCTCTGCC GTTGTAGGAC CCGGTCTTCC SYFP2								
1261	TGGAGTACAA GTACACAGAC CACAGAGCTT AATACAGGC CCAACAGCAG AAGAGACGCA TCAAGGCCAA CTTCAAGATC CCGCACACAA ACCTCATGTT GATGTTGTC GTGTGGAGA TATAGTGGC GCTGTGCTC TTCTTGCGCT AGTTCCGGTT GAATTTCTAG GCGGTCTTGT SYFP2								
1351	TCGAGGAGCG GGGGTGCG CCGCGGAC ACTAGGACA GAGACGCCC ATCGCGACG GCGGCTCT GTTGGCGGAC AAGCACTTAC AGCTCTGCC GCGGACGCTC GAGCGGCTGG TGATGGTCTT CTTTGGGAG TACGCGCTCC CCGGACAGCA CCAAGGCTG TTGCTGATGG SYFP2								
1441	TGAGCTACCA GTCCAAGCTG AGCAAGACC CCAACAGGAA GCGGATCAC ATGGTCTTGG TGGATCTGCT GAGCGCGCG GGCATCAGTC ACTCGATGCT CAGGTTCCAC TCGTTTCTCG GGTTCCTCTT CCGGCTACTG TACCAAGACC ACCTCAAGCA CTGGGCGCG CCGTAGTGAG BstBI XbaI								

FIG. 14 (cont'd)

CN1203 sCAAU-h1561-minB-globin-SYFP2-WPRE3-BGHpa (JF)	
SYFP2	WPRE3
Ascl	EcoRV
NotI	EcoRI
SacII	
BsrGI	
1531	TCGGCATGGA CGAGCTGTAC AAGTAAGTCC ACGGCGGCGC GCGGCGCGCA ATTGGAATAC ATAAATCAGC TCTGGATAC AAAATTGTG
AGCGGTACCT GCTGGACATG TTCAATTCAGC TGGCGGCGCG GCGGCGCGCT TAACTATAG TATTAGTGG AGACCTAATG TTTTAAACAC	WPRE3
1621	AAAGATTGAC TGGTATTCTT AACTATGTTG CTGCTTTTAC GCTATGTCGA TACGCTGCTT TAAAGCTTT GATATAGCT ATTGCTTCCC
TTTCTACTG ACCATAGAA TTGATACAC GAGGAAATG CGATACACT ATCGGACGA ATTACGGAA CATAGTACGA TAACGAAGCG	WPRE3
1711	GTATGGCTT CATTTCTCC TCGTTGTATA AATCTGCTT AGTCTCTGCC ACGGCGGAAC TCATGGGCGC CTGCTTGGC CGCTGCTGGA
CATACCGAA GTAAAGAGG AGGACATAT TTAGACCAA TCAAGAGCG TCGCGCTTG AGTAGGCGCG GACGGAACGG GCGAGGACCT	BGHpa
1801	CAGGGGCTCG GCTGTTGGC ACTGACAAT CGTGCTCG AGCGACTGT CTTCTAGT GCCAGCCATC TGTGTTTGC CCTGCCCGCG
GTGCGCGAGC CGACACCGG TCACTGTTAA GGCACCGAGC TCGCTGACAC GGAAGATCAA CGCTGGTAG ACACACACG GGAAGCGCGC	BGHpa
1891	TGCTTCCTT GACCTGGAA GGTGCCACTC CCACTGTCTT TTCTATATA AATGAGGAA TTGCAATCGA TTGTCTAGT AGGTGTCAIT
ACGGAAGGAA CTGGGACCTT CCAGGTTGAG GGTGACAGGA AAGGATTAIT TTACTCCTT AACGTAAGCT AACAGACTCA TCCACAGTAA	BGHpa(JF)
1981	CTATTCTGG GGTGGGGTG GGGCAGGACA GCAAGGGGGA GGTATGGGAA GACATAGCA GGCATGACTA GTCCACTCCC TCTCTGGCGG
GATAGACCC CCACACCCAC CCGTCTCTGT CGTTCCCTT CCTAACCTT CTTTATCTT CCGTACTGAT CAGGTGAGG AGAGACGCGC	RflBbs
2071	CTGCTCGCT CACTGAGGC GGGGACCAA AGGTGCGCG AGCGCGCGG TTGCGCGCG GCGGCTGACT GAGCGAGCGA GCGGCGAGAG
GACCGAGCGA GTGACTCGG CCGCTGCTT TCCAGTGGC TCGGCGCGG AACGCGCGC GCGGATGTA CTCGCTGCTT CCGGCGTCTC	RflBbs
2161	AGGACAGAT CCGGCGCGC ATGCGTCGAC AATTAAGTG CCGTCTTTT ACAAGCTCTT GACTGGGAA ACCTTGGCTT TACCAACTT
TCCGTGTCTA GCGCGCGCGG TACGAGCTG TTAAGTACC GGCAGCAAA TGTGACGA CTGACCTTT TGGGACCGCA ATGGTTGAA	
2251	AATCGGCTTG CAGCACATCC CCGTTTGGC AGTGGGGA ATAGCGAGA GCGCGGACG GATCGCTT CCCACAGAT CCGCAGCTG
TTAGCGGAC GTGCTGTAG GGAAGCGG TCAAGCGAT TATCGCTCT CCGGCGCTG CTAGCGGAA GGTGTGTCA CCGCTCGGAC	

FIG. 14 (cont'd)

CN1203 scAAV-h1561-m1aBgl1ob1n-5YFP2-WPRE3-2BglpA (JT)	
	Kozl ~~~~~ R3H ~~~~~ S181 ~~~~~ B261 ~~~~~ Kozl ~~~~~
2341	AATGGCGAAT GGGGCTTGAT GGGGTATTTT CTGCTTAGGC ATCTGTGGGG TATTTCACAC CGCATATGGT GCATCTICAG TACAATCTGC TTACCGCTTA CGCGGACTA CGCCATAAAA GAGGAATGGG TAGACACGGC ATAAATGTTG GGTATATACCA GGTGAGATGC ATGTTAGAGG
2431	TCTGATGCGG CATAGTTAAG CCAAGCCCGA CACCGGCTGA CCGGCTTGA CCGGCTTGC TCTGCGGC ATCGGCTTAC AGACTAGGCG GTATCAATTC GTCGCGGCT GTGGCGGCT GTGGCGGCT GCGCGGACT GCGCGGACT GCGCGGACT TAGGCGAATG
2521	AGACAAGCTG TGACCGTCTC CGGAGCTGC ATGTGTACA GGTTTTCACC GTCATACCG GATATAGG CAGTATGTC TTTCGCGCT CTGCTTTCC GAGCACTAT TCTGTTCGAC ACTGCGACAG GCGCTCGAG TACACACTCT CCAAAAGTGG CAGTATGTC CAGTATGTC TTTCGCGCT CTGCTTTCC GAGCACTAT
	Zet ~~~~~ AsiI ~~~~~
2611	CGCTATTTT TATAGGTTAA TGTCATGATA ATATGGGTTT CTTAGAGGTC AGGTCGACT TTTCGGGAA ATGTGGGCG AGCCCTTAT CGGGAATAAA ATATCCAAAT ACAGTACTAT TATTACCAAA GAATCTGCAG TCCACGTTGA AAAGCCCTT TACAGCGCC TTGGGGATAA
2701	TGTTTATTTT TCTAAATACA TTCAAATATG TATCGGCTCA TCAGACAATA ACCGTGATAA ATGCTTCAAT AATATTGAAA AGGGAAGAGT ACAAATAAAA AGATTATGT AGTTTATATC ATAGCGAGT ACTCTGTTAT TGGGACTATT TACGAAGTTA TTATAACTTT TTCTTTCTCA
	SspI ~~~~~
2791	ATGAGTATTC AGCATTTCCG TGTGCGCCTT ATTCCCTTTT TTGCGGCTT TTGCTTCTT ACCGGAAC GCTGCTGAAA TACTCATAG TTGTAAAGGC ACAGCGGAA TAGGGAATAA AACGCGCTAA AACGGAAGA CAAAACGAG TGGTCTTTC GAGCACTTT
2881	GTAAAGATG CTGAGATCA GTTGGGTGA CGAGTGGT ACACTGAAT GATCTCAAC ACCGTAGA TCCITGAGAG TTTCGCCCC CATTTCTAC GACTTCTAGT CAAGCCAGCT GGTACGCTAA TGTAGCTTGA CCTAGAGTTC TCGCACTCT AGGACTCTC AAGAGCGGG
	SspI ~~~~~
2971	GAAGAACGTT TTGCAATGAT GAGCACTTTT AAAGTCTGCG TAATGTGCGG GGTATATACC CGGGCAGAG GCAACTCGGT CTTCTTGCAA AAGGTTACTA CTGCTGAAA TTTCAGACAG ATACACGCG CCAATATAGG GCAATACTGC GGGCGCTCT GGTGAGCCA
3061	CGCGGCTATC ACTATTCTCA GAATGACTTG GTTGAGTACT CACCACTCAC AGAAAGGAT CTACGGATG GCATGAGCT AAGAGAATTA CGCGGCTATG TGTATAGACT CTTACTGAAC CAATCTATGA GTGGTCAGTG TCTTTTCTA GATGCTCTAC GGTACTCTCA TTCTCTTAAT
3151	TGGAGTGGTG CCAATACCAT GAGTGAUAC ACUGGGGCTA ACTTACTTCT GACAAGATC GAGGAGCGGA AGGAGCTAAC GGTCTTTTTC ACGTACGAG GGTATGGTA CTCACTATTG TGACCGGCTT TCAATGAAGA CTGTGCTAG CCTCTGCTT TCTCTGATTC GCGAATAAAC
3241	CACACATGG GGCATCATGT AACTGGCTT GAICGTTGG AACCGAGCT GAATGAGCC ATACCAACG AGGAGCTCA CACCACGATG GTGTTGTACC CCTAGTACA TTGAGCGGAA CTAGGACCC TTGGCTTGA TATGTTTC TGTCTGCTT GGTGCTCTAC

31/104

FIG. 14 (cont'd)

CH1203 scAAV-61561-minBglIobin-SYF2-WF23-BGMPA (CT)

3331	CCTGTAGCAA TGGACACAA CTTGCGGAAA CTAATTACAG GCGAAGTACT TACTCTAGCT TCCGCGCAAC AATTAATAGA CTGGATGGAG GGACATCGTT ACCGTGTTG CAACGGCTT GAAATTCAG CCCTTCATGA ATGAAATCGA AGGCGCGTGT TTAATTATCT GACCHACTC	333
3431	GCGGATAAG TTGAGAGAC ACTTCTGCGC TCGCGCCCTC CGGCTGGCTG GTTATTGCT GATAAATCTG GAGCGGCTGA GCGTGGGTCT CGCCTATTTC AACGTCCCTG TGAAGACCG AGCGCGGAG GCGGACCGAC CAAATACGA CTATTAGAC CTGCGGCACT CCGACCCAGA	334
3511	CGCGGTATCA TTGAGGCACT GGGGCCAGAT GGTAAAGGCT CCGGTATGCT AGTTATCTAC AGGAGCGGGA GTCAGGCAAC TATGGATGAA GCGGCATAGT AACGTCTGA CCGCGCTCTA CCAATCGGA GGCATAGCA TCAATAGATG TCGTGGCGCT CAGTCCGTG ATACCTACTT	335
3601	CGAATAGAC AGATCCCTGA GATAGGTCG TCACTGATTA AGCATGGTA ACTGTCAGAC CAGTTTACT CATATATACT TTAGATTCAT GCTTTATCTG TCTAGCGACT CTAATCCACG AGTCACTAT TCGTAACCAT TGCACAGCTG GTTCAATGCA GTATATATGA AATCTACTA	336
3691	TTAAACCTTC ATTATTAAAT TAAAGGATC TAGGTGAGA TCTTTTGA TAATCTCATG ACCAAATCC CTTAACGTGA GTTTGCTC AATTTGAAG TAAAAATTA AATTTCTTAG ATCCACTCT AGGAAANCT ATTAGAGTAC TGGTTTAGG GAATTGCACT CAAAGCGAAG	337
3781	CACGTAGCGT CAGACCCCGT AGAAGATC AAGGATCT CTGAGATCC TTTTTCCTG CCGGTAATCT GCTGCTGCA ACGAABAAA GTGACTCGCA GTCTGGGCA TCTTTCTAG TTTCTTAGA GAATCTTAGG AAAAAAGAC GGCATTAGA CGACGACGT TTGTTTTT	338
3871	CCACCGCTAC CAGCGTGT TGTTCGCG GATCAAGAC TACCAACTCT TTTTCGAG GTAACGCTCT TCAGCAGAGC GCAGATACCA GGTGGCGATG GTGCGACCA AACAAAGGC CTAGTTCTG ATGCTTGA AAGAGGCTTC CATTCACGA AGTGTCTCG CGTCTATGCT	339
3961	ATACTGTTT TTTAGTGA GCGTAGTTA GCGCACCACT TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT AATCTCTGTA TTATGACAG AAGATCAGAT CCGCATCAAT CCGGTGGTGA AGTTCTTAG ACATCGTGG GATGTATGG ACGGAGACGA TTAGGACAT	340
4051	CCAGTGGCTG CTGCGCTG CATAAGTGG TGTCTTAGG GGTTCGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGCTGA GGTCACCGAC GACGCTCAC GTATTTCAG ACAGAAAGGC CCAACTGAG TTCTGCTATC AATGCTCTAT TCGCGCTCG CAGCCCGACT	341
4141	ACGGGGGCTT CTTGACACA GCGGAGCTT GCGGAGAGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA AAGCGCCAGC TGGCCCCCAA GCAGGTGT GCGGTGAG CTTGCTTCTG GATCTCTATG GATCTCGAC TCGATACCTCT TTGCGGCTGC	342
4231	CTTCCGGAAG GGAGAAAGC GGACAGGAT CCGGTAGCG CAGGGTGG AAGAGGAG AGCTTCCAG GGGAAAGGCC GAAGGCTTC CTTCTTCC CTTCTCCATA GGCATTCGC TGTCTCTCTC GGTGCTCTCC TCGAAGGCTCC CCGTTTGGCG	343
4321	TGGTATCTTT ATAGTCTGT CCGGTTTCC CACCTCTGAC TTGAGCGTGG ATTTTGTGA TGTCTCTGAG GGGGCGGAG CCTATGGAAA ACCATAGAAA TATCAGACA GCGCAAGCG GTGAGAGCTG AACTCGGAGC TAAAGACACT ACGGACAGTC CCGCGGCTC GGATACCTTT	344
4411	AAGCGCAGCA ACGCGGCTT TTACGGTTC CTGGCTTTT GCTGACATG TCTTCTCTG CGTTATGCCC TGATCTGTG TTGCGGTGCT TCGCGCGGAA AATGCGGAG GACCGGAAA CAGCGGAAA ACGAGTGTAC AAGAAAGGAC GCAATAGGGG ACTAGACAC	345

32/104

FIG. 14 (cont'd)

CN1203 sCAAV-h1961-mineg10b10-SYFP2-WPRE3-BCHPA (JT)

```

4501  GATAACCGTA TTACGGCCCTT TGAGTGAAGT GATACCGGCTC GCGGCAAGCGG AACGACCGAG CCGAGCGCAAT CAGTGAAGGA GGAAGCGGAA
      CTATGGCAT AATGGCGGAA ACTGACTGGA CTATGCGGAG CCGCGTCGGC TTGCTGGCTC GCGTGGCTCA GTCACTGGCT CCTTGGCCCT
      5000
4591  GAGCGGCCAA TACGCAACCC GGTCTCTGCG GCGCGTTGGC CGATTGATTA ATGCAAGTGG CACGACAGGT TTCCCGACTG GAAAGCGGGC
      CTCGGCGGCTT ATCGCTTTGG CGGAGAGGGG CCGGCAACCG GCTAAGTAAT TACGTTGAGC GTGCTGTCCA AGGAGTGAC CTTTGGCCCG
      5500
4681  AGTGAGCGCA AGCAAAATTA TGTGAGTTAG CTCACTCATI AGGCACCGCA GAGTTTACAC TTTATGCTTC CCGCTGCTAT GTTGTTGGA
      TCACTGGCGT TCGGTTAATT ACACTCATC GAGTGAGTAA TCGGTGGGGT CCGAATGTG AATATCGGAG GCGGAGGATA CAGCACACCT
      6000
4771  ATTGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG ACCATGATTA CCGCAGGCTC TCGAGATCTA G
      TAACACTCGC CTATTGTTAA AGTGTTGCTT TTGTGATATC TGGTACTAAT GCGGTTGAG AGCTCTAGAT C
      6500

```

(SEQ ID NO: 43)

(SEQ ID NO: 97)

33/104

FIG. 15

Sequence Name	Enhancer & Promoter	Primary product(s)	Other components
CN1203	hI56i minBglobin	SYFP2	WPRES BGHpA
CN1244	hI56i minBglobin	SYFP2	WPRES BGHpA
CN1367	hI56i minBglobin	SYFP2 HisNavMs	P2A WPRES BGHpA
CN1389	hI56iCore minBglobin	SYFP2	WPRES BGHpA
CN1390	3XhI56iCore minBglobin	SYFP2	WPRES BGHpA
CN1498	3xhI56iCore minCMV	SYFP2 NavBp	P2A 3xHA WPRES BGHpA
CN1499	3xhI56iCore minCMV	SYFP2 NavMs	P2A 3xHA WPRES BGHpA
CN1500	3xhI56iCore minCMV	SYFP2 NavSheP D60N	P2A 3xHA WPRES BGHpA
CN1838	3Xzl46i minBG	SYFP2	WPRES SPA

FIG. 16

*-minB-BFP-W	*-H68-FP-T2A-enz-W	*-CMV-FP-I2-enz-W	*-mut-luc-E2A-enz-W
*-minB-CFP-W	*-H68-FP-T2A-TF-W	*-CMV-FP-I2-TF-W	*-mut-luc-E2A-TF-W
*-minB-GFP-W	*-H68-FP-T2A-rec-W	*-CMV-FP-I2-rec-W	*-mut-luc-E2A-rec-W
*-minB-OFP-W	*-H68-FP-T2A-CTP-W	*-CMV-FP-I2-CTP-W	*-mut-luc-E2A-CTP-W
*-minB-RFP-W	*-H68-FP-T2A-SM-W	*-CMV-FP-I2-SM-W	*-mut-luc-E2A-SM-W
*-minB-fRFP-W	*-H68-FP-T2A-NT-W	*-CMV-FP-I2-NT-W	*-mut-luc-E2A-NT-W
*-minB-YFP-W	*-H68-FP-T2A-CR-W	*-CMV-FP-I2-CR-W	*-mut-luc-E2A-CR-W
*-minB-Luc-W	*-H68-FP-T2A-chR-W	*-CMV-FP-I2-chR-W	*-mut-luc-E2A-chR-W
*-minC-BFP-W	*-H68-FP-T2A-gRNA-W	*-CMV-FP-I2-gRNA-W	*-mut-luc-E2A-gRNA-W
*-minC-CFP-W	*-H68-FP-T2A-Nuc-W	*-CMV-FP-I2-Nuc-W	*-mut-luc-E2A-Nuc-W
*-minC-GFP-W	*-RSV-FP-T2A-enz-W	*-SV40-FP-I2-enz-W	*-minR-luc-E2A-enz-W
*-minC-OFP-W	*-RSV-FP-T2A-TF-W	*-SV40-FP-I2-TF-W	*-minR-luc-E2A-TF-W
*-minC-RFP-W	*-RSV-FP-T2A-rec-W	*-SV40-FP-I2-rec-W	*-minR-luc-E2A-rec-W
*-minC-fRFP-W	*-RSV-FP-T2A-CTP-W	*-SV40-FP-I2-CTP-W	*-minR-luc-E2A-CTP-W
*-minC-YFP-W	*-RSV-FP-T2A-SM-W	*-SV40-FP-I2-SM-W	*-minR-luc-E2A-SM-W
*-minC-Luc-W	*-RSV-FP-T2A-NT-W	*-SV40-FP-I2-NT-W	*-minR-luc-E2A-NT-W
*-mut-BFP-W	*-RSV-FP-T2A-CR-W	*-SV40-FP-I2-CR-W	*-minR-luc-E2A-CR-W
*-mut-CFP-W	*-RSV-FP-T2A-chR-W	*-SV40-FP-I2-chR-W	*-minR-luc-E2A-chR-W
*-mut-GFP-W	*-RSV-FP-T2A-gRNA-W	*-SV40-FP-I2-gRNA-W	*-minR-luc-E2A-gRNA-W
*-mut-OFP-W	*-RSV-FP-T2A-Nuc-W	*-SV40-FP-I2-Nuc-W	*-minR-luc-E2A-Nuc-W
*-mut-RFP-W	*-minB-FP-T2A-enz-bG	*-H68-FP-I2-enz-W	*-CMV-luc-E2A-enz-W
*-mut-fRFP-W	*-minB-FP-T2A-TF-bG	*-H68-FP-I2-TF-W	*-CMV-luc-E2A-TF-W
*-mut-YFP-W	*-minB-FP-T2A-rec-bG	*-H68-FP-I2-rec-W	*-CMV-luc-E2A-rec-W
*-mut-Luc-W	*-minB-FP-T2A-CTP-bG	*-H68-FP-I2-CTP-W	*-CMV-luc-E2A-CTP-W
*-minR-BFP-W	*-minB-FP-T2A-SM-bG	*-H68-FP-I2-SM-W	*-CMV-luc-E2A-SM-W
*-minR-CFP-W	*-minB-FP-T2A-NT-bG	*-H68-FP-I2-NT-W	*-CMV-luc-E2A-NT-W
*-minR-GFP-W	*-minB-FP-T2A-CR-bG	*-H68-FP-I2-CR-W	*-CMV-luc-E2A-CR-W
*-minR-OFP-W	*-minB-FP-T2A-chR-bG	*-H68-FP-I2-chR-W	*-CMV-luc-E2A-chR-W
*-minR-RFP-W	*-minB-FP-T2A-gRNA-bG	*-H68-FP-I2-gRNA-W	*-CMV-luc-E2A-gRNA-W
*-minR-fRFP-W	*-minB-FP-T2A-Nuc-bG	*-H68-FP-I2-Nuc-W	*-CMV-luc-E2A-Nuc-W
*-minR-YFP-W	*-minC-FP-T2A-enz-bG	*-RSV-FP-I2-enz-W	*-SV40-luc-E2A-enz-W
*-minR-Luc-W	*-minC-FP-T2A-TF-bG	*-RSV-FP-I2-TF-W	*-SV40-luc-E2A-TF-W
*-CMV-BFP-W	*-minC-FP-T2A-rec-bG	*-RSV-FP-I2-rec-W	*-SV40-luc-E2A-rec-W

FIG. 16 cont'd

*-CMV-CFP-W	*-minC-FP-T2A-CTP-bG	*-RSV-FP-I2-CTP-W	*-SV40-luc-E2A-CTP-W
*-CMV-GFP-W	*-minC-FP-T2A-SM-bG	*-RSV-FP-I2-SM-W	*-SV40-luc-E2A-SM-W
*-CMV-OFP-W	*-minC-FP-T2A-NT-bG	*-RSV-FP-I2-NT-W	*-SV40-luc-E2A-NT-W
*-CMV-RFP-W	*-minC-FP-T2A-CR-bG	*-RSV-FP-I2-CR-W	*-SV40-luc-E2A-CR-W
*-CMV-fRFP-W	*-minC-FP-T2A-chR-bG	*-RSV-FP-I2-chR-W	*-SV40-luc-E2A-chR-W
*-CMV-YFP-W	*-minC-FP-T2A-gRNA-bG	*-RSV-FP-I2-gRNA-W	*-SV40-luc-E2A-gRNA-W
*-CMV-Luc-W	*-minC-FP-T2A-Nuc-bG	*-RSV-FP-I2-Nuc-W	*-SV40-luc-E2A-Nuc-W
*-SV40-BFP-W	*-mut-FP-T2A-enz-bG	*-minB-FP-I2-enz-bG	*-H68-luc-E2A-enz-W
*-SV40-CFP-W	*-mut-FP-T2A-TF-bG	*-minB-FP-I2-TF-bG	*-H68-luc-E2A-TF-W
*-SV40-GFP-W	*-mut-FP-T2A-rec-bG	*-minB-FP-I2-rec-bG	*-H68-luc-E2A-rec-W
*-SV40-OFP-W	*-mut-FP-T2A-CTP-bG	*-minB-FP-I2-CTP-bG	*-H68-luc-E2A-CTP-W
*-SV40-RFP-W	*-mut-FP-T2A-SM-bG	*-minB-FP-I2-SM-bG	*-H68-luc-E2A-SM-W
*-SV40-fRFP-W	*-mut-FP-T2A-NT-bG	*-minB-FP-I2-NT-bG	*-H68-luc-E2A-NT-W
*-SV40-YFP-W	*-mut-FP-T2A-CR-bG	*-minB-FP-I2-CR-bG	*-H68-luc-E2A-CR-W
*-SV40-Luc-W	*-mut-FP-T2A-chR-bG	*-minB-FP-I2-chR-bG	*-H68-luc-E2A-chR-W
*-H68-BFP-W	*-mut-FP-T2A-gRNA-bG	*-minB-FP-I2-gRNA-bG	*-H68-luc-E2A-gRNA-W
*-H68-CFP-W	*-mut-FP-T2A-Nuc-bG	*-minB-FP-I2-Nuc-bG	*-H68-luc-E2A-Nuc-W
*-H68-GFP-W	*-minR-FP-T2A-enz-bG	*-minC-FP-I2-enz-bG	*-RSV-luc-E2A-enz-W
*-H68-OFP-W	*-minR-FP-T2A-TF-bG	*-minC-FP-I2-TF-bG	*-RSV-luc-E2A-TF-W
*-H68-RFP-W	*-minR-FP-T2A-rec-bG	*-minC-FP-I2-rec-bG	*-RSV-luc-E2A-rec-W
*-H68-fRFP-W	*-minR-FP-T2A-CTP-bG	*-minC-FP-I2-CTP-bG	*-RSV-luc-E2A-CTP-W
*-H68-YFP-W	*-minR-FP-T2A-SM-bG	*-minC-FP-I2-SM-bG	*-RSV-luc-E2A-SM-W
*-H68-Luc-W	*-minR-FP-T2A-NT-bG	*-minC-FP-I2-NT-bG	*-RSV-luc-E2A-NT-W
*-RSV-BFP-W	*-minR-FP-T2A-CR-bG	*-minC-FP-I2-CR-bG	*-RSV-luc-E2A-CR-W
*-RSV-CFP-W	*-minR-FP-T2A-chR-bG	*-minC-FP-I2-chR-bG	*-RSV-luc-E2A-chR-W
*-RSV-GFP-W	*-minR-FP-T2A-gRNA-bG	*-minC-FP-I2-gRNA-bG	*-RSV-luc-E2A-gRNA-W
*-RSV-OFP-W	*-minR-FP-T2A-Nuc-bG	*-minC-FP-I2-Nuc-bG	*-RSV-luc-E2A-Nuc-W
*-RSV-RFP-W	*-CMV-FP-T2A-enz-bG	*-mut-FP-I2-enz-bG	*-minB-luc-E2A-enz-bG
*-RSV-fRFP-W	*-CMV-FP-T2A-TF-bG	*-mut-FP-I2-TF-bG	*-minB-luc-E2A-TF-bG
*-RSV-YFP-W	*-CMV-FP-T2A-rec-bG	*-mut-FP-I2-rec-bG	*-minB-luc-E2A-rec-bG
*-RSV-Luc-W	*-CMV-FP-T2A-CTP-bG	*-mut-FP-I2-CTP-bG	*-minB-luc-E2A-CTP-bG
*-minB-BFP-S	*-CMV-FP-T2A-SM-bG	*-mut-FP-I2-SM-bG	*-minB-luc-E2A-SM-bG
*-minB-CFP-S	*-CMV-FP-T2A-NT-bG	*-mut-FP-I2-NT-bG	*-minB-luc-E2A-NT-bG
*-minB-GFP-S	*-CMV-FP-T2A-CR-bG	*-mut-FP-I2-CR-bG	*-minB-luc-E2A-CR-bG

FIG. 16 cont'd

*-minB-OFP-S	*-CMV-FP-T2A-chR-bG	*-mut-FP-I2-chR-bG	*-minB-luc-E2A-chR-bG
*-minB-RFP-S	*-CMV-FP-T2A-gRNA-bG	*-mut-FP-I2-gRNA-bG	*-minB-luc-E2A-gRNA-bG
*-minB-fRFP-S	*-CMV-FP-T2A-Nuc-bG	*-mut-FP-I2-Nuc-bG	*-minB-luc-E2A-Nuc-bG
*-minB-YFP-S	*-SV40-FP-T2A-enz-bG	*-minR-FP-I2-enz-bG	*-minC-luc-E2A-enz-bG
*-minB-Luc-S	*-SV40-FP-T2A-TF-bG	*-minR-FP-I2-TF-bG	*-minC-luc-E2A-TF-bG
*-minC-BFP-S	*-SV40-FP-T2A-rec-bG	*-minR-FP-I2-rec-bG	*-minC-luc-E2A-rec-bG
*-minC-CFP-S	*-SV40-FP-T2A-CTP-bG	*-minR-FP-I2-CTP-bG	*-minC-luc-E2A-CTP-bG
*-minC-GFP-S	*-SV40-FP-T2A-SM-bG	*-minR-FP-I2-SM-bG	*-minC-luc-E2A-SM-bG
*-minC-OFP-S	*-SV40-FP-T2A-NT-bG	*-minR-FP-I2-NT-bG	*-minC-luc-E2A-NT-bG
*-minC-RFP-S	*-SV40-FP-T2A-CR-bG	*-minR-FP-I2-CR-bG	*-minC-luc-E2A-CR-bG
*-minC-fRFP-S	*-SV40-FP-T2A-chR-bG	*-minR-FP-I2-chR-bG	*-minC-luc-E2A-chR-bG
*-minC-YFP-S	*-SV40-FP-T2A-gRNA-bG	*-minR-FP-I2-gRNA-bG	*-minC-luc-E2A-gRNA-bG
*-minC-Luc-S	*-SV40-FP-T2A-Nuc-bG	*-minR-FP-I2-Nuc-bG	*-minC-luc-E2A-Nuc-bG
*-mut-BFP-S	*-H68-FP-T2A-enz-bG	*-CMV-FP-I2-enz-bG	*-mut-luc-E2A-enz-bG
*-mut-CFP-S	*-H68-FP-T2A-TF-bG	*-CMV-FP-I2-TF-bG	*-mut-luc-E2A-TF-bG
*-mut-GFP-S	*-H68-FP-T2A-rec-bG	*-CMV-FP-I2-rec-bG	*-mut-luc-E2A-rec-bG
*-mut-OFP-S	*-H68-FP-T2A-CTP-bG	*-CMV-FP-I2-CTP-bG	*-mut-luc-E2A-CTP-bG
*-mut-RFP-S	*-H68-FP-T2A-SM-bG	*-CMV-FP-I2-SM-bG	*-mut-luc-E2A-SM-bG
*-mut-fRFP-S	*-H68-FP-T2A-NT-bG	*-CMV-FP-I2-NT-bG	*-mut-luc-E2A-NT-bG
*-mut-YFP-S	*-H68-FP-T2A-CR-bG	*-CMV-FP-I2-CR-bG	*-mut-luc-E2A-CR-bG
*-mut-Luc-S	*-H68-FP-T2A-chR-bG	*-CMV-FP-I2-chR-bG	*-mut-luc-E2A-chR-bG
*-minR-BFP-S	*-H68-FP-T2A-gRNA-bG	*-CMV-FP-I2-gRNA-bG	*-mut-luc-E2A-gRNA-bG
*-minR-CFP-S	*-H68-FP-T2A-Nuc-bG	*-CMV-FP-I2-Nuc-bG	*-mut-luc-E2A-Nuc-bG
*-minR-GFP-S	*-RSV-FP-T2A-enz-bG	*-SV40-FP-I2-enz-bG	*-minR-luc-E2A-enz-bG
*-minR-OFP-S	*-RSV-FP-T2A-TF-bG	*-SV40-FP-I2-TF-bG	*-minR-luc-E2A-TF-bG
*-minR-RFP-S	*-RSV-FP-T2A-rec-bG	*-SV40-FP-I2-rec-bG	*-minR-luc-E2A-rec-bG
*-minR-fRFP-S	*-RSV-FP-T2A-CTP-bG	*-SV40-FP-I2-CTP-bG	*-minR-luc-E2A-CTP-bG
*-minR-YFP-S	*-RSV-FP-T2A-SM-bG	*-SV40-FP-I2-SM-bG	*-minR-luc-E2A-SM-bG
*-minR-Luc-S	*-RSV-FP-T2A-NT-bG	*-SV40-FP-I2-NT-bG	*-minR-luc-E2A-NT-bG
*-CMV-BFP-S	*-RSV-FP-T2A-CR-bG	*-SV40-FP-I2-CR-bG	*-minR-luc-E2A-CR-bG
*-CMV-CFP-S	*-RSV-FP-T2A-chR-bG	*-SV40-FP-I2-chR-bG	*-minR-luc-E2A-chR-bG
*-CMV-GFP-S	*-RSV-FP-T2A-gRNA-bG	*-SV40-FP-I2-gRNA-bG	*-minR-luc-E2A-gRNA-bG
*-CMV-OFP-S	*-RSV-FP-T2A-Nuc-bG	*-SV40-FP-I2-Nuc-bG	*-minR-luc-E2A-Nuc-bG
*-CMV-RFP-S	*-minB-FP-T2A-enz-S	*-H68-FP-I2-enz-bG	*-CMV-luc-E2A-enz-bG

FIG. 16 cont'd

*-CMV- <i>rfFP</i> -S	*-minB-FP-T2A-TF-S	*-H68-FP-I2-TF-bG	*-CMV-luc-E2A-TF-bG
*-CMV-YFP-S	*-minB-FP-T2A-rec-S	*-H68-FP-I2-rec-bG	*-CMV-luc-E2A-rec-bG
*-CMV-Luc-S	*-minB-FP-T2A-CTP-S	*-H68-FP-I2-CTP-bG	*-CMV-luc-E2A-CTP-bG
*-SV40-BFP-S	*-minB-FP-T2A-SM-S	*-H68-FP-I2-SM-bG	*-CMV-luc-E2A-SM-bG
*-SV40-CFP-S	*-minB-FP-T2A-NT-S	*-H68-FP-I2-NT-bG	*-CMV-luc-E2A-NT-bG
*-SV40-GFP-S	*-minB-FP-T2A-CR-S	*-H68-FP-I2-CR-bG	*-CMV-luc-E2A-CR-bG
*-SV40-OFP-S	*-minB-FP-T2A-chR-S	*-H68-FP-I2-chR-bG	*-CMV-luc-E2A-chR-bG
*-SV40-RFP-S	*-minB-FP-T2A-gRNA-S	*-H68-FP-I2-gRNA-bG	*-CMV-luc-E2A-gRNA-bG
*-SV40- <i>rfFP</i> -S	*-minB-FP-T2A-Nuc-S	*-H68-FP-I2-Nuc-bG	*-CMV-luc-E2A-Nuc-bG
*-SV40-YFP-S	*-minC-FP-T2A-enz-S	*-RSV-FP-I2-enz-bG	*-SV40-luc-E2A-enz-bG
*-SV40-Luc-S	*-minC-FP-T2A-TF-S	*-RSV-FP-I2-TF-bG	*-SV40-luc-E2A-TF-bG
*-H68-BFP-S	*-minC-FP-T2A-rec-S	*-RSV-FP-I2-rec-bG	*-SV40-luc-E2A-rec-bG
*-H68-CFP-S	*-minC-FP-T2A-CTP-S	*-RSV-FP-I2-CTP-bG	*-SV40-luc-E2A-CTP-bG
*-H68-GFP-S	*-minC-FP-T2A-SM-S	*-RSV-FP-I2-SM-bG	*-SV40-luc-E2A-SM-bG
*-H68-OFP-S	*-minC-FP-T2A-NT-S	*-RSV-FP-I2-NT-bG	*-SV40-luc-E2A-NT-bG
*-H68-RFP-S	*-minC-FP-T2A-CR-S	*-RSV-FP-I2-CR-bG	*-SV40-luc-E2A-CR-bG
*-H68- <i>rfFP</i> -S	*-minC-FP-T2A-chR-S	*-RSV-FP-I2-chR-bG	*-SV40-luc-E2A-chR-bG
*-H68-YFP-S	*-minC-FP-T2A-gRNA-S	*-RSV-FP-I2-gRNA-bG	*-SV40-luc-E2A-gRNA-bG
*-H68-Luc-S	*-minC-FP-T2A-Nuc-S	*-RSV-FP-I2-Nuc-bG	*-SV40-luc-E2A-Nuc-bG
*-RSV-BFP-S	*-mut-FP-T2A-enz-S	*-minB-FP-I2-enz-S	*-H68-luc-E2A-enz-bG
*-RSV-CFP-S	*-mut-FP-T2A-TF-S	*-minB-FP-I2-TF-S	*-H68-luc-E2A-TF-bG
*-RSV-GFP-S	*-mut-FP-T2A-rec-S	*-minB-FP-I2-rec-S	*-H68-luc-E2A-rec-bG
*-RSV-OFP-S	*-mut-FP-T2A-CTP-S	*-minB-FP-I2-CTP-S	*-H68-luc-E2A-CTP-bG
*-RSV-RFP-S	*-mut-FP-T2A-SM-S	*-minB-FP-I2-SM-S	*-H68-luc-E2A-SM-bG
*-RSV- <i>rfFP</i> -S	*-mut-FP-T2A-NT-S	*-minB-FP-I2-NT-S	*-H68-luc-E2A-NT-bG
*-RSV-YFP-S	*-mut-FP-T2A-CR-S	*-minB-FP-I2-CR-S	*-H68-luc-E2A-CR-bG
*-RSV-Luc-S	*-mut-FP-T2A-chR-S	*-minB-FP-I2-chR-S	*-H68-luc-E2A-chR-bG
*-minB-BFP-W-S	*-mut-FP-T2A-gRNA-S	*-minB-FP-I2-gRNA-S	*-H68-luc-E2A-gRNA-bG
*-minB-CFP-W-S	*-mut-FP-T2A-Nuc-S	*-minB-FP-I2-Nuc-S	*-H68-luc-E2A-Nuc-bG
*-minB-GFP-W-S	*-minR-FP-T2A-enz-S	*-minC-FP-I2-enz-S	*-RSV-luc-E2A-enz-bG
*-minB-OFP-W-S	*-minR-FP-T2A-TF-S	*-minC-FP-I2-TF-S	*-RSV-luc-E2A-TF-bG
*-minB-RFP-W-S	*-minR-FP-T2A-rec-S	*-minC-FP-I2-rec-S	*-RSV-luc-E2A-rec-bG
*-minB- <i>rfFP</i> -W-S	*-minR-FP-T2A-CTP-S	*-minC-FP-I2-CTP-S	*-RSV-luc-E2A-CTP-bG
*-minB-YFP-W-S	*-minR-FP-T2A-SM-S	*-minC-FP-I2-SM-S	*-RSV-luc-E2A-SM-bG

FIG. 16 cont'd

*-minB-Luc-W-S	*-minR-FP-T2A-NT-S	*-minC-FP-I2-NT-S	*-RSV-luc-E2A-NT-bG
*-minC-BFP-W-S	*-minR-FP-T2A-CR-S	*-minC-FP-I2-CR-S	*-RSV-luc-E2A-CR-bG
*-minC-CFP-W-S	*-minR-FP-T2A-chR-S	*-minC-FP-I2-chR-S	*-RSV-luc-E2A-chR-bG
*-minC-GFP-W-S	*-minR-FP-T2A-gRNA-S	*-minC-FP-I2-gRNA-S	*-RSV-luc-E2A-gRNA-bG
*-minC-OFP-W-S	*-minR-FP-T2A-Nuc-S	*-minC-FP-I2-Nuc-S	*-RSV-luc-E2A-Nuc-bG
*-minC-RFP-W-S	*-CMV-FP-T2A-enz-S	*-mut-FP-I2-enz-S	*-minB-luc-E2A-enz-S
*-minC-fRFP-W-S	*-CMV-FP-T2A-TF-S	*-mut-FP-I2-TF-S	*-minB-luc-E2A-TF-S
*-minC-YFP-W-S	*-CMV-FP-T2A-rec-S	*-mut-FP-I2-rec-S	*-minB-luc-E2A-rec-S
*-minC-Luc-W-S	*-CMV-FP-T2A-CTP-S	*-mut-FP-I2-CTP-S	*-minB-luc-E2A-CTP-S
*-mut-BFP-W-S	*-CMV-FP-T2A-SM-S	*-mut-FP-I2-SM-S	*-minB-luc-E2A-SM-S
*-mut-CFP-W-S	*-CMV-FP-T2A-NT-S	*-mut-FP-I2-NT-S	*-minB-luc-E2A-NT-S
*-mut-GFP-W-S	*-CMV-FP-T2A-CR-S	*-mut-FP-I2-CR-S	*-minB-luc-E2A-CR-S
*-mut-OFP-W-S	*-CMV-FP-T2A-chR-S	*-mut-FP-I2-chR-S	*-minB-luc-E2A-chR-S
*-mut-RFP-W-S	*-CMV-FP-T2A-gRNA-S	*-mut-FP-I2-gRNA-S	*-minB-luc-E2A-gRNA-S
*-mut-fRFP-W-S	*-CMV-FP-T2A-Nuc-S	*-mut-FP-I2-Nuc-S	*-minB-luc-E2A-Nuc-S
*-mut-YFP-W-S	*-SV40-FP-T2A-enz-S	*-minR-FP-I2-enz-S	*-minC-luc-E2A-enz-S
*-mut-Luc-W-S	*-SV40-FP-T2A-TF-S	*-minR-FP-I2-TF-S	*-minC-luc-E2A-TF-S
*-minR-BFP-W-S	*-SV40-FP-T2A-rec-S	*-minR-FP-I2-rec-S	*-minC-luc-E2A-rec-S
*-minR-CFP-W-S	*-SV40-FP-T2A-CTP-S	*-minR-FP-I2-CTP-S	*-minC-luc-E2A-CTP-S
*-minR-GFP-W-S	*-SV40-FP-T2A-SM-S	*-minR-FP-I2-SM-S	*-minC-luc-E2A-SM-S
*-minR-OFP-W-S	*-SV40-FP-T2A-NT-S	*-minR-FP-I2-NT-S	*-minC-luc-E2A-NT-S
*-minR-RFP-W-S	*-SV40-FP-T2A-CR-S	*-minR-FP-I2-CR-S	*-minC-luc-E2A-CR-S
*-minR-fRFP-W-S	*-SV40-FP-T2A-chR-S	*-minR-FP-I2-chR-S	*-minC-luc-E2A-chR-S
*-minR-YFP-W-S	*-SV40-FP-T2A-gRNA-S	*-minR-FP-I2-gRNA-S	*-minC-luc-E2A-gRNA-S
*-minR-Luc-W-S	*-SV40-FP-T2A-Nuc-S	*-minR-FP-I2-Nuc-S	*-minC-luc-E2A-Nuc-S
*-minB-BFP-bG	*-H68-FP-T2A-enz-S	*-CMV-FP-I2-enz-S	*-mut-luc-E2A-enz-S
*-minB-CFP-bG	*-H68-FP-T2A-TF-S	*-CMV-FP-I2-TF-S	*-mut-luc-E2A-TF-S
*-minB-GFP-bG	*-H68-FP-T2A-rec-S	*-CMV-FP-I2-rec-S	*-mut-luc-E2A-rec-S
*-minB-OFP-bG	*-H68-FP-T2A-CTP-S	*-CMV-FP-I2-CTP-S	*-mut-luc-E2A-CTP-S
*-minB-RFP-bG	*-H68-FP-T2A-SM-S	*-CMV-FP-I2-SM-S	*-mut-luc-E2A-SM-S
*-minB-fRFP-bG	*-H68-FP-T2A-NT-S	*-CMV-FP-I2-NT-S	*-mut-luc-E2A-NT-S
*-minB-YFP-bG	*-H68-FP-T2A-CR-S	*-CMV-FP-I2-CR-S	*-mut-luc-E2A-CR-S
*-minB-Luc-bG	*-H68-FP-T2A-chR-S	*-CMV-FP-I2-chR-S	*-mut-luc-E2A-chR-S
*-minC-BFP-bG	*-H68-FP-T2A-gRNA-S	*-CMV-FP-I2-gRNA-S	*-mut-luc-E2A-gRNA-S

FIG. 16 cont'd

*-minC-CFP-bG	*-H68-FP-T2A-Nuc-S	*-CMV-FP-I2-Nuc-S	*-mut-luc-E2A-Nuc-S
*-minC-GFP-bG	*-RSV-FP-T2A-enz-S	*-SV40-FP-I2-enz-S	*-minR-luc-E2A-enz-S
*-minC-OFP-bG	*-RSV-FP-T2A-TF-S	*-SV40-FP-I2-TF-S	*-minR-luc-E2A-TF-S
*-minC-RFP-bG	*-RSV-FP-T2A-rec-S	*-SV40-FP-I2-rec-S	*-minR-luc-E2A-rec-S
*-minC-fRFP-bG	*-RSV-FP-T2A-CTP-S	*-SV40-FP-I2-CTP-S	*-minR-luc-E2A-CTP-S
*-minC-YFP-bG	*-RSV-FP-T2A-SM-S	*-SV40-FP-I2-SM-S	*-minR-luc-E2A-SM-S
*-minC-Luc-bG	*-RSV-FP-T2A-NT-S	*-SV40-FP-I2-NT-S	*-minR-luc-E2A-NT-S
*-mut-BFP-bG	*-RSV-FP-T2A-CR-S	*-SV40-FP-I2-CR-S	*-minR-luc-E2A-CR-S
*-mut-CFP-bG	*-RSV-FP-T2A-chR-S	*-SV40-FP-I2-chR-S	*-minR-luc-E2A-chR-S
*-mut-GFP-bG	*-RSV-FP-T2A-gRNA-S	*-SV40-FP-I2-gRNA-S	*-minR-luc-E2A-gRNA-S
*-mut-OFP-bG	*-RSV-FP-T2A-Nuc-S	*-SV40-FP-I2-Nuc-S	*-minR-luc-E2A-Nuc-S
*-mut-RFP-bG	*-minB-FP-T2A-enz-W-bG	*-H68-FP-I2-enz-S	*-CMV-luc-E2A-enz-S
*-mut-fRFP-bG	*-minB-FP-T2A-TF-W-bG	*-H68-FP-I2-TF-S	*-CMV-luc-E2A-TF-S
*-mut-YFP-bG	*-minB-FP-T2A-rec-W-bG	*-H68-FP-I2-rec-S	*-CMV-luc-E2A-rec-S
*-mut-Luc-bG	*-minB-FP-T2A-CTP-W-bG	*-H68-FP-I2-CTP-S	*-CMV-luc-E2A-CTP-S
*-minR-BFP-bG	*-minB-FP-T2A-SM-W-bG	*-H68-FP-I2-SM-S	*-CMV-luc-E2A-SM-S
*-minR-CFP-bG	*-minB-FP-T2A-NT-W-bG	*-H68-FP-I2-NT-S	*-CMV-luc-E2A-NT-S
*-minR-GFP-bG	*-minB-FP-T2A-CR-W-bG	*-H68-FP-I2-CR-S	*-CMV-luc-E2A-CR-S
*-minR-OFP-bG	*-minB-FP-T2A-chR-W-bG	*-H68-FP-I2-chR-S	*-CMV-luc-E2A-chR-S
*-minR-RFP-bG	*-minB-FP-T2A-gRNA-W-bG	*-H68-FP-I2-gRNA-S	*-CMV-luc-E2A-gRNA-S
*-minR-fRFP-bG	*-minB-FP-T2A-Nuc-W-bG	*-H68-FP-I2-Nuc-S	*-CMV-luc-E2A-Nuc-S
*-minR-YFP-bG	*-minC-FP-T2A-enz-W-bG	*-RSV-FP-I2-enz-S	*-SV40-luc-E2A-enz-S
*-minR-Luc-bG	*-minC-FP-T2A-TF-W-bG	*-RSV-FP-I2-TF-S	*-SV40-luc-E2A-TF-S
*-CMV-BFP-bG	*-minC-FP-T2A-rec-W-bG	*-RSV-FP-I2-rec-S	*-SV40-luc-E2A-rec-S
*-CMV-CFP-bG	*-minC-FP-T2A-CTP-W-bG	*-RSV-FP-I2-CTP-S	*-SV40-luc-E2A-CTP-S
*-CMV-GFP-bG	*-minC-FP-T2A-SM-W-bG	*-RSV-FP-I2-SM-S	*-SV40-luc-E2A-SM-S
*-CMV-OFP-bG	*-minC-FP-T2A-NT-W-bG	*-RSV-FP-I2-NT-S	*-SV40-luc-E2A-NT-S
*-CMV-RFP-bG	*-minC-FP-T2A-CR-W-bG	*-RSV-FP-I2-CR-S	*-SV40-luc-E2A-CR-S
*-CMV-fRFP-bG	*-minC-FP-T2A-chR-W-bG	*-RSV-FP-I2-chR-S	*-SV40-luc-E2A-chR-S
*-CMV-YFP-bG	*-minC-FP-T2A-gRNA-W-bG	*-RSV-FP-I2-gRNA-S	*-SV40-luc-E2A-gRNA-S
*-CMV-Luc-bG	*-minC-FP-T2A-Nuc-W-bG	*-RSV-FP-I2-Nuc-S	*-SV40-luc-E2A-Nuc-S
*-SV40-BFP-bG	*-mut-FP-T2A-enz-W-bG	*-minB-FP-I2-enz-W-bG	*-H68-luc-E2A-enz-S
*-SV40-CFP-bG	*-mut-FP-T2A-TF-W-bG	*-minB-FP-I2-TF-W-bG	*-H68-luc-E2A-TF-S
*-SV40-GFP-bG	*-mut-FP-T2A-rec-W-bG	*-minB-FP-I2-rec-W-bG	*-H68-luc-E2A-rec-S

FIG. 16 cont'd

*-SV40-OFP-bG	*-mut-FP-T2A-CTP-W-bG	*-minB-FP-I2-CTP-W-bG	*-H68-luc-E2A-CTP-S
*-SV40-RFP-bG	*-mut-FP-T2A-SM-W-bG	*-minB-FP-I2-SM-W-bG	*-H68-luc-E2A-SM-S
*-SV40-rRFP-bG	*-mut-FP-T2A-NT-W-bG	*-minB-FP-I2-NT-W-bG	*-H68-luc-E2A-NT-S
*-SV40-YFP-bG	*-mut-FP-T2A-CR-W-bG	*-minB-FP-I2-CR-W-bG	*-H68-luc-E2A-CR-S
*-SV40-Luc-bG	*-mut-FP-T2A-chR-W-bG	*-minB-FP-I2-chR-W-bG	*-H68-luc-E2A-chR-S
*-H68-BFP-bG	*-mut-FP-T2A-gRNA-W-bG	*-minB-FP-I2-gRNA-W-bG	*-H68-luc-E2A-gRNA-S
*-H68-CFP-bG	*-mut-FP-T2A-Nuc-W-bG	*-minB-FP-I2-Nuc-W-bG	*-H68-luc-E2A-Nuc-S
*-H68-GFP-bG	*-minR-FP-T2A-enz-W-bG	*-minC-FP-I2-enz-W-bG	*-RSV-luc-E2A-enz-S
*-H68-OFP-bG	*-minR-FP-T2A-TF-W-bG	*-minC-FP-I2-TF-W-bG	*-RSV-luc-E2A-TF-S
*-H68-RFP-bG	*-minR-FP-T2A-rec-W-bG	*-minC-FP-I2-rec-W-bG	*-RSV-luc-E2A-rec-S
*-H68-rRFP-bG	*-minR-FP-T2A-CTP-W-bG	*-minC-FP-I2-CTP-W-bG	*-RSV-luc-E2A-CTP-S
*-H68-YFP-bG	*-minR-FP-T2A-SM-W-bG	*-minC-FP-I2-SM-W-bG	*-RSV-luc-E2A-SM-S
*-H68-Luc-bG	*-minR-FP-T2A-NT-W-bG	*-minC-FP-I2-NT-W-bG	*-RSV-luc-E2A-NT-S
*-RSV-BFP-bG	*-minR-FP-T2A-CR-W-bG	*-minC-FP-I2-CR-W-bG	*-RSV-luc-E2A-CR-S
*-RSV-CFP-bG	*-minR-FP-T2A-chR-W-bG	*-minC-FP-I2-chR-W-bG	*-RSV-luc-E2A-chR-S
*-RSV-GFP-bG	*-minR-FP-T2A-gRNA-W-bG	*-minC-FP-I2-gRNA-W-bG	*-RSV-luc-E2A-gRNA-S
*-RSV-OFP-bG	*-minR-FP-T2A-Nuc-W-bG	*-minC-FP-I2-Nuc-W-bG	*-RSV-luc-E2A-Nuc-S
*-RSV-RFP-bG	*-CMV-FP-T2A-enz-W-bG	*-mut-FP-I2-enz-W-bG	*-minB-luc-E2A-enz-W-bG
*-RSV-rRFP-bG	*-CMV-FP-T2A-TF-W-bG	*-mut-FP-I2-TF-W-bG	*-minB-luc-E2A-TF-W-bG
*-RSV-YFP-bG	*-CMV-FP-T2A-rec-W-bG	*-mut-FP-I2-rec-W-bG	*-minB-luc-E2A-rec-W-bG
*-RSV-Luc-bG	*-CMV-FP-T2A-CTP-W-bG	*-mut-FP-I2-CTP-W-bG	*-minB-luc-E2A-CTP-W-bG
*-minB-BFP-W-bG	*-CMV-FP-T2A-SM-W-bG	*-mut-FP-I2-SM-W-bG	*-minB-luc-E2A-SM-W-bG
*-minB-CFP-W-bG	*-CMV-FP-T2A-NT-W-bG	*-mut-FP-I2-NT-W-bG	*-minB-luc-E2A-NT-W-bG
*-minB-GFP-W-bG	*-CMV-FP-T2A-CR-W-bG	*-mut-FP-I2-CR-W-bG	*-minB-luc-E2A-CR-W-bG
*-minB-OFP-W-bG	*-CMV-FP-T2A-chR-W-bG	*-mut-FP-I2-chR-W-bG	*-minB-luc-E2A-chR-W-bG
*-minB-RFP-W-bG	*-CMV-FP-T2A-gRNA-W-bG	*-mut-FP-I2-gRNA-W-bG	*-minB-luc-E2A-gRNA-W-bG
*-minB-rRFP-W-bG	*-CMV-FP-T2A-Nuc-W-bG	*-mut-FP-I2-Nuc-W-bG	*-minB-luc-E2A-Nuc-W-bG
*-minB-YFP-W-bG	*-SV40-FP-T2A-enz-W-bG	*-minR-FP-I2-enz-W-bG	*-minC-luc-E2A-enz-W-bG
*-minB-Luc-W-bG	*-SV40-FP-T2A-TF-W-bG	*-minR-FP-I2-TF-W-bG	*-minC-luc-E2A-TF-W-bG
*-minC-BFP-W-bG	*-SV40-FP-T2A-rec-W-bG	*-minR-FP-I2-rec-W-bG	*-minC-luc-E2A-rec-W-bG
*-minC-CFP-W-bG	*-SV40-FP-T2A-CTP-W-bG	*-minR-FP-I2-CTP-W-bG	*-minC-luc-E2A-CTP-W-bG
*-minC-GFP-W-bG	*-SV40-FP-T2A-SM-W-bG	*-minR-FP-I2-SM-W-bG	*-minC-luc-E2A-SM-W-bG
*-minC-OFP-W-bG	*-SV40-FP-T2A-NT-W-bG	*-minR-FP-I2-NT-W-bG	*-minC-luc-E2A-NT-W-bG
*-minC-RFP-W-bG	*-SV40-FP-T2A-CR-W-bG	*-minR-FP-I2-CR-W-bG	*-minC-luc-E2A-CR-W-bG

FIG. 16 cont'd

*-minC-fRFP-W-bG	*-SV40-FP-T2A-chR-W-bG	*-minR-FP-I2-chR-W-bG	*-minC-luc-E2A-chR-W-bG
*-minC-YFP-W-bG	*-SV40-FP-T2A-gRNA-W-bG	*-minR-FP-I2-gRNA-W-bG	*-minC-luc-E2A-gRNA-W-bG
*-minC-Luc-W-bG	*-SV40-FP-T2A-Nuc-W-bG	*-minR-FP-I2-Nuc-W-bG	*-minC-luc-E2A-Nuc-W-bG
*-mut-BFP-W-bG	*-H68-FP-T2A-enz-W-bG	*-CMV-FP-I2-enz-W-bG	*-mut-luc-E2A-enz-W-bG
*-mut-CFP-W-bG	*-H68-FP-T2A-TF-W-bG	*-CMV-FP-I2-TF-W-bG	*-mut-luc-E2A-TF-W-bG
*-mut-GFP-W-bG	*-H68-FP-T2A-rec-W-bG	*-CMV-FP-I2-rec-W-bG	*-mut-luc-E2A-rec-W-bG
*-mut-OFP-W-bG	*-H68-FP-T2A-CTP-W-bG	*-CMV-FP-I2-CTP-W-bG	*-mut-luc-E2A-CTP-W-bG
*-mut-RFP-W-bG	*-H68-FP-T2A-SM-W-bG	*-CMV-FP-I2-SM-W-bG	*-mut-luc-E2A-SM-W-bG
*-mut-fRFP-W-bG	*-H68-FP-T2A-NT-W-bG	*-CMV-FP-I2-NT-W-bG	*-mut-luc-E2A-NT-W-bG
*-mut-YFP-W-bG	*-H68-FP-T2A-CR-W-bG	*-CMV-FP-I2-CR-W-bG	*-mut-luc-E2A-CR-W-bG
*-mut-Luc-W-bG	*-H68-FP-T2A-chR-W-bG	*-CMV-FP-I2-chR-W-bG	*-mut-luc-E2A-chR-W-bG
*-minR-BFP-W-bG	*-H68-FP-T2A-gRNA-W-bG	*-CMV-FP-I2-gRNA-W-bG	*-mut-luc-E2A-gRNA-W-bG
*-minR-CFP-W-bG	*-H68-FP-T2A-Nuc-W-bG	*-CMV-FP-I2-Nuc-W-bG	*-mut-luc-E2A-Nuc-W-bG
*-minR-GFP-W-bG	*-RSV-FP-T2A-enz-W-bG	*-SV40-FP-I2-enz-W-bG	*-minR-luc-E2A-enz-W-bG
*-minR-OFP-W-bG	*-RSV-FP-T2A-TF-W-bG	*-SV40-FP-I2-TF-W-bG	*-minR-luc-E2A-TF-W-bG
*-minR-RFP-W-bG	*-RSV-FP-T2A-rec-W-bG	*-SV40-FP-I2-rec-W-bG	*-minR-luc-E2A-rec-W-bG
*-minR-fRFP-W-bG	*-RSV-FP-T2A-CTP-W-bG	*-SV40-FP-I2-CTP-W-bG	*-minR-luc-E2A-CTP-W-bG
*-minR-YFP-W-bG	*-RSV-FP-T2A-SM-W-bG	*-SV40-FP-I2-SM-W-bG	*-minR-luc-E2A-SM-W-bG
*-minR-Luc-W-bG	*-RSV-FP-T2A-NT-W-bG	*-SV40-FP-I2-NT-W-bG	*-minR-luc-E2A-NT-W-bG
*-CMV-BFP-W-bG	*-RSV-FP-T2A-CR-W-bG	*-SV40-FP-I2-CR-W-bG	*-minR-luc-E2A-CR-W-bG
*-CMV-CFP-W-bG	*-RSV-FP-T2A-chR-W-bG	*-SV40-FP-I2-chR-W-bG	*-minR-luc-E2A-chR-W-bG
*-CMV-GFP-W-bG	*-RSV-FP-T2A-gRNA-W-bG	*-SV40-FP-I2-gRNA-W-bG	*-minR-luc-E2A-gRNA-W-bG
*-CMV-OFP-W-bG	*-RSV-FP-T2A-Nuc-W-bG	*-SV40-FP-I2-Nuc-W-bG	*-minR-luc-E2A-Nuc-W-bG
*-CMV-RFP-W-bG	*-minB-FP-T2A-enz-W-S	*-H68-FP-I2-enz-W-bG	*-CMV-luc-E2A-enz-W-bG
*-CMV-fRFP-W-bG	*-minB-FP-T2A-TF-W-S	*-H68-FP-I2-TF-W-bG	*-CMV-luc-E2A-TF-W-bG
*-CMV-YFP-W-bG	*-minB-FP-T2A-rec-W-S	*-H68-FP-I2-rec-W-bG	*-CMV-luc-E2A-rec-W-bG
*-CMV-Luc-W-bG	*-minB-FP-T2A-CTP-W-S	*-H68-FP-I2-CTP-W-bG	*-CMV-luc-E2A-CTP-W-bG
*-SV40-BFP-W-bG	*-minB-FP-T2A-SM-W-S	*-H68-FP-I2-SM-W-bG	*-CMV-luc-E2A-SM-W-bG
*-SV40-CFP-W-bG	*-minB-FP-T2A-NT-W-S	*-H68-FP-I2-NT-W-bG	*-CMV-luc-E2A-NT-W-bG
*-SV40-GFP-W-bG	*-minB-FP-T2A-CR-W-S	*-H68-FP-I2-CR-W-bG	*-CMV-luc-E2A-CR-W-bG
*-SV40-OFP-W-bG	*-minB-FP-T2A-chR-W-S	*-H68-FP-I2-chR-W-bG	*-CMV-luc-E2A-chR-W-bG
*-SV40-RFP-W-bG	*-minB-FP-T2A-gRNA-W-S	*-H68-FP-I2-gRNA-W-bG	*-CMV-luc-E2A-gRNA-W-bG
*-SV40-fRFP-W-bG	*-minB-FP-T2A-Nuc-W-S	*-H68-FP-I2-Nuc-W-bG	*-CMV-luc-E2A-Nuc-W-bG
*-SV40-YFP-W-bG	*-minC-FP-T2A-enz-W-S	*-RSV-FP-I2-enz-W-bG	*-SV40-luc-E2A-enz-W-bG

FIG. 16 cont'd

*-SV40-Luc-W-bG	*-minC-FP-T2A-TF-W-S	*-RSV-FP-I2-TF-W-bG	*-SV40-luc-E2A-TF-W-bG
*-H68-BFP-W-bG	*-minC-FP-T2A-rec-W-S	*-RSV-FP-I2-rec-W-bG	*-SV40-luc-E2A-rec-W-bG
*-H68-CFP-W-bG	*-minC-FP-T2A-CTP-W-S	*-RSV-FP-I2-CTP-W-bG	*-SV40-luc-E2A-CTP-W-bG
*-H68-GFP-W-bG	*-minC-FP-T2A-SM-W-S	*-RSV-FP-I2-SM-W-bG	*-SV40-luc-E2A-SM-W-bG
*-H68-OFP-W-bG	*-minC-FP-T2A-NT-W-S	*-RSV-FP-I2-NT-W-bG	*-SV40-luc-E2A-NT-W-bG
*-H68-RFP-W-bG	*-minC-FP-T2A-CR-W-S	*-RSV-FP-I2-CR-W-bG	*-SV40-luc-E2A-CR-W-bG
*-H68-rRFP-W-bG	*-minC-FP-T2A-chR-W-S	*-RSV-FP-I2-chR-W-bG	*-SV40-luc-E2A-chR-W-bG
*-H68-YFP-W-bG	*-minC-FP-T2A-gRNA-W-S	*-RSV-FP-I2-gRNA-W-bG	*-SV40-luc-E2A-gRNA-W-bG
*-H68-Luc-W-bG	*-minC-FP-T2A-Nuc-W-S	*-RSV-FP-I2-Nuc-W-bG	*-SV40-luc-E2A-Nuc-W-bG
*-RSV-BFP-W-bG	*-mut-FP-T2A-enz-W-S	*-minB-FP-I2-enz-W-S	*-H68-luc-E2A-enz-W-bG
*-RSV-CFP-W-bG	*-mut-FP-T2A-TF-W-S	*-minB-FP-I2-TF-W-S	*-H68-luc-E2A-TF-W-bG
*-RSV-GFP-W-bG	*-mut-FP-T2A-rec-W-S	*-minB-FP-I2-rec-W-S	*-H68-luc-E2A-rec-W-bG
*-RSV-OFP-W-bG	*-mut-FP-T2A-CTP-W-S	*-minB-FP-I2-CTP-W-S	*-H68-luc-E2A-CTP-W-bG
*-RSV-RFP-W-bG	*-mut-FP-T2A-SM-W-S	*-minB-FP-I2-SM-W-S	*-H68-luc-E2A-SM-W-bG
*-RSV-rRFP-W-bG	*-mut-FP-T2A-NT-W-S	*-minB-FP-I2-NT-W-S	*-H68-luc-E2A-NT-W-bG
*-RSV-YFP-W-bG	*-mut-FP-T2A-CR-W-S	*-minB-FP-I2-CR-W-S	*-H68-luc-E2A-CR-W-bG
*-RSV-Luc-W-bG	*-mut-FP-T2A-chR-W-S	*-minB-FP-I2-chR-W-S	*-H68-luc-E2A-chR-W-bG
*-CMV-BFP-W-S	*-mut-FP-T2A-gRNA-W-S	*-minB-FP-I2-gRNA-W-S	*-H68-luc-E2A-gRNA-W-bG
*-CMV-CFP-W-S	*-mut-FP-T2A-Nuc-W-S	*-minB-FP-I2-Nuc-W-S	*-H68-luc-E2A-Nuc-W-bG
*-CMV-GFP-W-S	*-minR-FP-T2A-enz-W-S	*-minC-FP-I2-enz-W-S	*-RSV-luc-E2A-enz-W-bG
*-CMV-OFP-W-S	*-minR-FP-T2A-TF-W-S	*-minC-FP-I2-TF-W-S	*-RSV-luc-E2A-TF-W-bG
*-CMV-RFP-W-S	*-minR-FP-T2A-rec-W-S	*-minC-FP-I2-rec-W-S	*-RSV-luc-E2A-rec-W-bG
*-CMV-rRFP-W-S	*-minR-FP-T2A-CTP-W-S	*-minC-FP-I2-CTP-W-S	*-RSV-luc-E2A-CTP-W-bG
*-CMV-YFP-W-S	*-minR-FP-T2A-SM-W-S	*-minC-FP-I2-SM-W-S	*-RSV-luc-E2A-SM-W-bG
*-CMV-Luc-W-S	*-minR-FP-T2A-NT-W-S	*-minC-FP-I2-NT-W-S	*-RSV-luc-E2A-NT-W-bG
*-SV40-BFP-W-S	*-minR-FP-T2A-CR-W-S	*-minC-FP-I2-CR-W-S	*-RSV-luc-E2A-CR-W-bG
*-SV40-CFP-W-S	*-minR-FP-T2A-chR-W-S	*-minC-FP-I2-chR-W-S	*-RSV-luc-E2A-chR-W-bG
*-SV40-GFP-W-S	*-minR-FP-T2A-gRNA-W-S	*-minC-FP-I2-gRNA-W-S	*-RSV-luc-E2A-gRNA-W-bG
*-SV40-OFP-W-S	*-minR-FP-T2A-Nuc-W-S	*-minC-FP-I2-Nuc-W-S	*-RSV-luc-E2A-Nuc-W-bG
*-SV40-RFP-W-S	*-CMV-FP-T2A-enz-W-S	*-mut-FP-I2-enz-W-S	*-minB-luc-E2A-enz-W-S
*-SV40-rRFP-W-S	*-CMV-FP-T2A-TF-W-S	*-mut-FP-I2-TF-W-S	*-minB-luc-E2A-TF-W-S
*-SV40-YFP-W-S	*-CMV-FP-T2A-rec-W-S	*-mut-FP-I2-rec-W-S	*-minB-luc-E2A-rec-W-S
*-SV40-Luc-W-S	*-CMV-FP-T2A-CTP-W-S	*-mut-FP-I2-CTP-W-S	*-minB-luc-E2A-CTP-W-S
*-H68-BFP-W-S	*-CMV-FP-T2A-SM-W-S	*-mut-FP-I2-SM-W-S	*-minB-luc-E2A-SM-W-S

FIG. 16 cont'd

*-H68-CFP-W-S	*-CMV-FP-T2A-NT-W-S	*-mut-FP-I2-NT-W-S	*-minB-luc-E2A-NT-W-S
*-H68-GFP-W-S	*-CMV-FP-T2A-CR-W-S	*-mut-FP-I2-CR-W-S	*-minB-luc-E2A-CR-W-S
*-H68-OPF-W-S	*-CMV-FP-T2A-chR-W-S	*-mut-FP-I2-chR-W-S	*-minB-luc-E2A-chR-W-S
*-H68-RFP-W-S	*-CMV-FP-T2A-gRNA-W-S	*-mut-FP-I2-gRNA-W-S	*-minB-luc-E2A-gRNA-W-S
*-H68-fRFP-W-S	*-CMV-FP-T2A-Nuc-W-S	*-mut-FP-I2-Nuc-W-S	*-minB-luc-E2A-Nuc-W-S
*-H68-YFP-W-S	*-SV40-FP-T2A-enz-W-S	*-minR-FP-I2-enz-W-S	*-minC-luc-E2A-enz-W-S
*-H68-Luc-W-S	*-SV40-FP-T2A-TF-W-S	*-minR-FP-I2-TF-W-S	*-minC-luc-E2A-TF-W-S
*-RSV-BFP-W-S	*-SV40-FP-T2A-rec-W-S	*-minR-FP-I2-rec-W-S	*-minC-luc-E2A-rec-W-S
*-RSV-CFP-W-S	*-SV40-FP-T2A-CTP-W-S	*-minR-FP-I2-CTP-W-S	*-minC-luc-E2A-CTP-W-S
*-RSV-GFP-W-S	*-SV40-FP-T2A-SM-W-S	*-minR-FP-I2-SM-W-S	*-minC-luc-E2A-SM-W-S
*-RSV-OPF-W-S	*-SV40-FP-T2A-NT-W-S	*-minR-FP-I2-NT-W-S	*-minC-luc-E2A-NT-W-S
*-RSV-RFP-W-S	*-SV40-FP-T2A-CR-W-S	*-minR-FP-I2-CR-W-S	*-minC-luc-E2A-CR-W-S
*-RSV-fRFP-W-S	*-SV40-FP-T2A-chR-W-S	*-minR-FP-I2-chR-W-S	*-minC-luc-E2A-chR-W-S
*-RSV-YFP-W-S	*-SV40-FP-T2A-gRNA-W-S	*-minR-FP-I2-gRNA-W-S	*-minC-luc-E2A-gRNA-W-S
*-RSV-Luc-W-S	*-SV40-FP-T2A-Nuc-W-S	*-minR-FP-I2-Nuc-W-S	*-minC-luc-E2A-Nuc-W-S
*-minB-enz-W	*-H68-FP-T2A-enz-W-S	*-CMV-FP-I2-enz-W-S	*-mut-luc-E2A-enz-W-S
*-minB-TF-W	*-H68-FP-T2A-TF-W-S	*-CMV-FP-I2-TF-W-S	*-mut-luc-E2A-TF-W-S
*-minB-rec-W	*-H68-FP-T2A-rec-W-S	*-CMV-FP-I2-rec-W-S	*-mut-luc-E2A-rec-W-S
*-minB-CTP-W	*-H68-FP-T2A-CTP-W-S	*-CMV-FP-I2-CTP-W-S	*-mut-luc-E2A-CTP-W-S
*-minB-SM-W	*-H68-FP-T2A-SM-W-S	*-CMV-FP-I2-SM-W-S	*-mut-luc-E2A-SM-W-S
*-minB-NT-W	*-H68-FP-T2A-NT-W-S	*-CMV-FP-I2-NT-W-S	*-mut-luc-E2A-NT-W-S
*-minB-CR-W	*-H68-FP-T2A-CR-W-S	*-CMV-FP-I2-CR-W-S	*-mut-luc-E2A-CR-W-S
*-minB-chR-W	*-H68-FP-T2A-chR-W-S	*-CMV-FP-I2-chR-W-S	*-mut-luc-E2A-chR-W-S
*-minB-gRNA-W	*-H68-FP-T2A-gRNA-W-S	*-CMV-FP-I2-gRNA-W-S	*-mut-luc-E2A-gRNA-W-S
*-minB-Nuc-W	*-H68-FP-T2A-Nuc-W-S	*-CMV-FP-I2-Nuc-W-S	*-mut-luc-E2A-Nuc-W-S
*-minC-enz-W	*-RSV-FP-T2A-enz-W-S	*-SV40-FP-I2-enz-W-S	*-minR-luc-E2A-enz-W-S
*-minC-TF-W	*-RSV-FP-T2A-TF-W-S	*-SV40-FP-I2-TF-W-S	*-minR-luc-E2A-TF-W-S
*-minC-rec-W	*-RSV-FP-T2A-rec-W-S	*-SV40-FP-I2-rec-W-S	*-minR-luc-E2A-rec-W-S
*-minC-CTP-W	*-RSV-FP-T2A-CTP-W-S	*-SV40-FP-I2-CTP-W-S	*-minR-luc-E2A-CTP-W-S
*-minC-SM-W	*-RSV-FP-T2A-SM-W-S	*-SV40-FP-I2-SM-W-S	*-minR-luc-E2A-SM-W-S
*-minC-NT-W	*-RSV-FP-T2A-NT-W-S	*-SV40-FP-I2-NT-W-S	*-minR-luc-E2A-NT-W-S
*-minC-CR-W	*-RSV-FP-T2A-CR-W-S	*-SV40-FP-I2-CR-W-S	*-minR-luc-E2A-CR-W-S
*-minC-chR-W	*-RSV-FP-T2A-chR-W-S	*-SV40-FP-I2-chR-W-S	*-minR-luc-E2A-chR-W-S
*-minC-gRNA-W	*-RSV-FP-T2A-gRNA-W-S	*-SV40-FP-I2-gRNA-W-S	*-minR-luc-E2A-gRNA-W-S

FIG. 16 cont'd

*-minC-Nuc-W	*-RSV-FP-T2A-Nuc-W-S	*-SV40-FP-I2-Nuc-W-S	*-minR-luc-E2A-Nuc-W-S
*-mut-enz-W	*-minB-FP-E2A-enz-W	*-H68-FP-I2-enz-W-S	*-CMV-luc-E2A-enz-W-S
*-mut-TF-W	*-minB-FP-E2A-TF-W	*-H68-FP-I2-TF-W-S	*-CMV-luc-E2A-TF-W-S
*-mut-rec-W	*-minB-FP-E2A-rec-W	*-H68-FP-I2-rec-W-S	*-CMV-luc-E2A-rec-W-S
*-mut-CTP-W	*-minB-FP-E2A-CTP-W	*-H68-FP-I2-CTP-W-S	*-CMV-luc-E2A-CTP-W-S
*-mut-SM-W	*-minB-FP-E2A-SM-W	*-H68-FP-I2-SM-W-S	*-CMV-luc-E2A-SM-W-S
*-mut-NT-W	*-minB-FP-E2A-NT-W	*-H68-FP-I2-NT-W-S	*-CMV-luc-E2A-NT-W-S
*-mut-CR-W	*-minB-FP-E2A-CR-W	*-H68-FP-I2-CR-W-S	*-CMV-luc-E2A-CR-W-S
*-mut-chR-W	*-minB-FP-E2A-chR-W	*-H68-FP-I2-chR-W-S	*-CMV-luc-E2A-chR-W-S
*-mut-gRNA-W	*-minB-FP-E2A-gRNA-W	*-H68-FP-I2-gRNA-W-S	*-CMV-luc-E2A-gRNA-W-S
*-mut-Nuc-W	*-minB-FP-E2A-Nuc-W	*-H68-FP-I2-Nuc-W-S	*-CMV-luc-E2A-Nuc-W-S
*-minR-enz-W	*-minC-FP-E2A-enz-W	*-RSV-FP-I2-enz-W-S	*-SV40-luc-E2A-enz-W-S
*-minR-TF-W	*-minC-FP-E2A-TF-W	*-RSV-FP-I2-TF-W-S	*-SV40-luc-E2A-TF-W-S
*-minR-rec-W	*-minC-FP-E2A-rec-W	*-RSV-FP-I2-rec-W-S	*-SV40-luc-E2A-rec-W-S
*-minR-CTP-W	*-minC-FP-E2A-CTP-W	*-RSV-FP-I2-CTP-W-S	*-SV40-luc-E2A-CTP-W-S
*-minR-SM-W	*-minC-FP-E2A-SM-W	*-RSV-FP-I2-SM-W-S	*-SV40-luc-E2A-SM-W-S
*-minR-NT-W	*-minC-FP-E2A-NT-W	*-RSV-FP-I2-NT-W-S	*-SV40-luc-E2A-NT-W-S
*-minR-CR-W	*-minC-FP-E2A-CR-W	*-RSV-FP-I2-CR-W-S	*-SV40-luc-E2A-CR-W-S
*-minR-chR-W	*-minC-FP-E2A-chR-W	*-RSV-FP-I2-chR-W-S	*-SV40-luc-E2A-chR-W-S
*-minR-gRNA-W	*-minC-FP-E2A-gRNA-W	*-RSV-FP-I2-gRNA-W-S	*-SV40-luc-E2A-gRNA-W-S
*-minR-Nuc-W	*-minC-FP-E2A-Nuc-W	*-RSV-FP-I2-Nuc-W-S	*-SV40-luc-E2A-Nuc-W-S
*-CMV-enz-W	*-mut-FP-E2A-enz-W	*-minB-luc-P2A-enz-W	*-H68-luc-E2A-enz-W-S
*-CMV-TF-W	*-mut-FP-E2A-TF-W	*-minB-luc-P2A-TF-W	*-H68-luc-E2A-TF-W-S
*-CMV-rec-W	*-mut-FP-E2A-rec-W	*-minB-luc-P2A-rec-W	*-H68-luc-E2A-rec-W-S
*-CMV-CTP-W	*-mut-FP-E2A-CTP-W	*-minB-luc-P2A-CTP-W	*-H68-luc-E2A-CTP-W-S
*-CMV-SM-W	*-mut-FP-E2A-SM-W	*-minB-luc-P2A-SM-W	*-H68-luc-E2A-SM-W-S
*-CMV-NT-W	*-mut-FP-E2A-NT-W	*-minB-luc-P2A-NT-W	*-H68-luc-E2A-NT-W-S
*-CMV-CR-W	*-mut-FP-E2A-CR-W	*-minB-luc-P2A-CR-W	*-H68-luc-E2A-CR-W-S
*-CMV-chR-W	*-mut-FP-E2A-chR-W	*-minB-luc-P2A-chR-W	*-H68-luc-E2A-chR-W-S
*-CMV-gRNA-W	*-mut-FP-E2A-gRNA-W	*-minB-luc-P2A-gRNA-W	*-H68-luc-E2A-gRNA-W-S
*-CMV-Nuc-W	*-mut-FP-E2A-Nuc-W	*-minB-luc-P2A-Nuc-W	*-H68-luc-E2A-Nuc-W-S
*-SV40-enz-W	*-minR-FP-E2A-enz-W	*-minC-luc-P2A-enz-W	*-RSV-luc-E2A-enz-W-S
*-SV40-TF-W	*-minR-FP-E2A-TF-W	*-minC-luc-P2A-TF-W	*-RSV-luc-E2A-TF-W-S
*-SV40-rec-W	*-minR-FP-E2A-rec-W	*-minC-luc-P2A-rec-W	*-RSV-luc-E2A-rec-W-S

FIG. 16 cont'd

*-SV40-CTP-W	*-minR-FP-E2A-CTP-W	*-minC-luc-P2A-CTP-W	*-RSV-luc-E2A-CTP-W-S
*-SV40-SM-W	*-minR-FP-E2A-SM-W	*-minC-luc-P2A-SM-W	*-RSV-luc-E2A-SM-W-S
*-SV40-NT-W	*-minR-FP-E2A-NT-W	*-minC-luc-P2A-NT-W	*-RSV-luc-E2A-NT-W-S
*-SV40-CR-W	*-minR-FP-E2A-CR-W	*-minC-luc-P2A-CR-W	*-RSV-luc-E2A-CR-W-S
*-SV40-chR-W	*-minR-FP-E2A-chR-W	*-minC-luc-P2A-chR-W	*-RSV-luc-E2A-chR-W-S
*-SV40-gRNA-W	*-minR-FP-E2A-gRNA-W	*-minC-luc-P2A-gRNA-W	*-RSV-luc-E2A-gRNA-W-S
*-SV40-Nuc-W	*-minR-FP-E2A-Nuc-W	*-minC-luc-P2A-Nuc-W	*-RSV-luc-E2A-Nuc-W-S
*-H68-enz-W	*-CMV-FP-E2A-enz-W	*-mut-luc-P2A-enz-W	*-minB-luc-F2A-enz-W
*-H68-TF-W	*-CMV-FP-E2A-TF-W	*-mut-luc-P2A-TF-W	*-minB-luc-F2A-TF-W
*-H68-rec-W	*-CMV-FP-E2A-rec-W	*-mut-luc-P2A-rec-W	*-minB-luc-F2A-rec-W
*-H68-CTP-W	*-CMV-FP-E2A-CTP-W	*-mut-luc-P2A-CTP-W	*-minB-luc-F2A-CTP-W
*-H68-SM-W	*-CMV-FP-E2A-SM-W	*-mut-luc-P2A-SM-W	*-minB-luc-F2A-SM-W
*-H68-NT-W	*-CMV-FP-E2A-NT-W	*-mut-luc-P2A-NT-W	*-minB-luc-F2A-NT-W
*-H68-CR-W	*-CMV-FP-E2A-CR-W	*-mut-luc-P2A-CR-W	*-minB-luc-F2A-CR-W
*-H68-chR-W	*-CMV-FP-E2A-chR-W	*-mut-luc-P2A-chR-W	*-minB-luc-F2A-chR-W
*-H68-gRNA-W	*-CMV-FP-E2A-gRNA-W	*-mut-luc-P2A-gRNA-W	*-minB-luc-F2A-gRNA-W
*-H68-Nuc-W	*-CMV-FP-E2A-Nuc-W	*-mut-luc-P2A-Nuc-W	*-minB-luc-F2A-Nuc-W
*-RSV-enz-W	*-SV40-FP-E2A-enz-W	*-minR-luc-P2A-enz-W	*-minC-luc-F2A-enz-W
*-RSV-TF-W	*-SV40-FP-E2A-TF-W	*-minR-luc-P2A-TF-W	*-minC-luc-F2A-TF-W
*-RSV-rec-W	*-SV40-FP-E2A-rec-W	*-minR-luc-P2A-rec-W	*-minC-luc-F2A-rec-W
*-RSV-CTP-W	*-SV40-FP-E2A-CTP-W	*-minR-luc-P2A-CTP-W	*-minC-luc-F2A-CTP-W
*-RSV-SM-W	*-SV40-FP-E2A-SM-W	*-minR-luc-P2A-SM-W	*-minC-luc-F2A-SM-W
*-RSV-NT-W	*-SV40-FP-E2A-NT-W	*-minR-luc-P2A-NT-W	*-minC-luc-F2A-NT-W
*-RSV-CR-W	*-SV40-FP-E2A-CR-W	*-minR-luc-P2A-CR-W	*-minC-luc-F2A-CR-W
*-RSV-chR-W	*-SV40-FP-E2A-chR-W	*-minR-luc-P2A-chR-W	*-minC-luc-F2A-chR-W
*-RSV-gRNA-W	*-SV40-FP-E2A-gRNA-W	*-minR-luc-P2A-gRNA-W	*-minC-luc-F2A-gRNA-W
*-RSV-Nuc-W	*-SV40-FP-E2A-Nuc-W	*-minR-luc-P2A-Nuc-W	*-minC-luc-F2A-Nuc-W
*-minB-enz-S	*-H68-FP-E2A-enz-W	*-CMV-luc-P2A-enz-W	*-mut-luc-F2A-enz-W
*-minB-TF-S	*-H68-FP-E2A-TF-W	*-CMV-luc-P2A-TF-W	*-mut-luc-F2A-TF-W
*-minB-rec-S	*-H68-FP-E2A-rec-W	*-CMV-luc-P2A-rec-W	*-mut-luc-F2A-rec-W
*-minB-CTP-S	*-H68-FP-E2A-CTP-W	*-CMV-luc-P2A-CTP-W	*-mut-luc-F2A-CTP-W
*-minB-SM-S	*-H68-FP-E2A-SM-W	*-CMV-luc-P2A-SM-W	*-mut-luc-F2A-SM-W
*-minB-NT-S	*-H68-FP-E2A-NT-W	*-CMV-luc-P2A-NT-W	*-mut-luc-F2A-NT-W
*-minB-CR-S	*-H68-FP-E2A-CR-W	*-CMV-luc-P2A-CR-W	*-mut-luc-F2A-CR-W

FIG. 16 cont'd

*-minB-chR-S	*-H68-FP-E2A-chR-W	*-CMV-luc-P2A-chR-W	*-mut-luc-F2A-chR-W
*-minB-gRNA-S	*-H68-FP-E2A-gRNA-W	*-CMV-luc-P2A-gRNA-W	*-mut-luc-F2A-gRNA-W
*-minB-Nuc-S	*-H68-FP-E2A-Nuc-W	*-CMV-luc-P2A-Nuc-W	*-mut-luc-F2A-Nuc-W
*-minC-enz-S	*-RSV-FP-E2A-enz-W	*-SV40-luc-P2A-enz-W	*-minR-luc-F2A-enz-W
*-minC-TF-S	*-RSV-FP-E2A-TF-W	*-SV40-luc-P2A-TF-W	*-minR-luc-F2A-TF-W
*-minC-rec-S	*-RSV-FP-E2A-rec-W	*-SV40-luc-P2A-rec-W	*-minR-luc-F2A-rec-W
*-minC-CTP-S	*-RSV-FP-E2A-CTP-W	*-SV40-luc-P2A-CTP-W	*-minR-luc-F2A-CTP-W
*-minC-SM-S	*-RSV-FP-E2A-SM-W	*-SV40-luc-P2A-SM-W	*-minR-luc-F2A-SM-W
*-minC-NT-S	*-RSV-FP-E2A-NT-W	*-SV40-luc-P2A-NT-W	*-minR-luc-F2A-NT-W
*-minC-CR-S	*-RSV-FP-E2A-CR-W	*-SV40-luc-P2A-CR-W	*-minR-luc-F2A-CR-W
*-minC-chR-S	*-RSV-FP-E2A-chR-W	*-SV40-luc-P2A-chR-W	*-minR-luc-F2A-chR-W
*-minC-gRNA-S	*-RSV-FP-E2A-gRNA-W	*-SV40-luc-P2A-gRNA-W	*-minR-luc-F2A-gRNA-W
*-minC-Nuc-S	*-RSV-FP-E2A-Nuc-W	*-SV40-luc-P2A-Nuc-W	*-minR-luc-F2A-Nuc-W
*-mut-enz-S	*-minB-FP-E2A-enz-bG	*-H68-luc-P2A-enz-W	*-CMV-luc-F2A-enz-W
*-mut-TF-S	*-minB-FP-E2A-TF-bG	*-H68-luc-P2A-TF-W	*-CMV-luc-F2A-TF-W
*-mut-rec-S	*-minB-FP-E2A-rec-bG	*-H68-luc-P2A-rec-W	*-CMV-luc-F2A-rec-W
*-mut-CTP-S	*-minB-FP-E2A-CTP-bG	*-H68-luc-P2A-CTP-W	*-CMV-luc-F2A-CTP-W
*-mut-SM-S	*-minB-FP-E2A-SM-bG	*-H68-luc-P2A-SM-W	*-CMV-luc-F2A-SM-W
*-mut-NT-S	*-minB-FP-E2A-NT-bG	*-H68-luc-P2A-NT-W	*-CMV-luc-F2A-NT-W
*-mut-CR-S	*-minB-FP-E2A-CR-bG	*-H68-luc-P2A-CR-W	*-CMV-luc-F2A-CR-W
*-mut-chR-S	*-minB-FP-E2A-chR-bG	*-H68-luc-P2A-chR-W	*-CMV-luc-F2A-chR-W
*-mut-gRNA-S	*-minB-FP-E2A-gRNA-bG	*-H68-luc-P2A-gRNA-W	*-CMV-luc-F2A-gRNA-W
*-mut-Nuc-S	*-minB-FP-E2A-Nuc-bG	*-H68-luc-P2A-Nuc-W	*-CMV-luc-F2A-Nuc-W
*-minR-enz-S	*-minC-FP-E2A-enz-bG	*-RSV-luc-P2A-enz-W	*-SV40-luc-F2A-enz-W
*-minR-TF-S	*-minC-FP-E2A-TF-bG	*-RSV-luc-P2A-TF-W	*-SV40-luc-F2A-TF-W
*-minR-rec-S	*-minC-FP-E2A-rec-bG	*-RSV-luc-P2A-rec-W	*-SV40-luc-F2A-rec-W
*-minR-CTP-S	*-minC-FP-E2A-CTP-bG	*-RSV-luc-P2A-CTP-W	*-SV40-luc-F2A-CTP-W
*-minR-SM-S	*-minC-FP-E2A-SM-bG	*-RSV-luc-P2A-SM-W	*-SV40-luc-F2A-SM-W
*-minR-NT-S	*-minC-FP-E2A-NT-bG	*-RSV-luc-P2A-NT-W	*-SV40-luc-F2A-NT-W
*-minR-CR-S	*-minC-FP-E2A-CR-bG	*-RSV-luc-P2A-CR-W	*-SV40-luc-F2A-CR-W
*-minR-chR-S	*-minC-FP-E2A-chR-bG	*-RSV-luc-P2A-chR-W	*-SV40-luc-F2A-chR-W
*-minR-gRNA-S	*-minC-FP-E2A-gRNA-bG	*-RSV-luc-P2A-gRNA-W	*-SV40-luc-F2A-gRNA-W
*-minR-Nuc-S	*-minC-FP-E2A-Nuc-bG	*-RSV-luc-P2A-Nuc-W	*-SV40-luc-F2A-Nuc-W
*-CMV-enz-S	*-mut-FP-E2A-enz-bG	*-minB-luc-P2A-enz-bG	*-H68-luc-F2A-enz-W

FIG. 16 cont'd

*-CMV-TF-S	*-mut-FP-E2A-TF-bG	*-minB-luc-P2A-TF-bG	*-H68-luc-F2A-TF-W
*-CMV-rec-S	*-mut-FP-E2A-rec-bG	*-minB-luc-P2A-rec-bG	*-H68-luc-F2A-rec-W
*-CMV-CTP-S	*-mut-FP-E2A-CTP-bG	*-minB-luc-P2A-CTP-bG	*-H68-luc-F2A-CTP-W
*-CMV-SM-S	*-mut-FP-E2A-SM-bG	*-minB-luc-P2A-SM-bG	*-H68-luc-F2A-SM-W
*-CMV-NT-S	*-mut-FP-E2A-NT-bG	*-minB-luc-P2A-NT-bG	*-H68-luc-F2A-NT-W
*-CMV-CR-S	*-mut-FP-E2A-CR-bG	*-minB-luc-P2A-CR-bG	*-H68-luc-F2A-CR-W
*-CMV-chR-S	*-mut-FP-E2A-chR-bG	*-minB-luc-P2A-chR-bG	*-H68-luc-F2A-chR-W
*-CMV-gRNA-S	*-mut-FP-E2A-gRNA-bG	*-minB-luc-P2A-gRNA-bG	*-H68-luc-F2A-gRNA-W
*-CMV-Nuc-S	*-mut-FP-E2A-Nuc-bG	*-minB-luc-P2A-Nuc-bG	*-H68-luc-F2A-Nuc-W
*-SV40-enz-S	*-minR-FP-E2A-enz-bG	*-minC-luc-P2A-enz-bG	*-RSV-luc-F2A-enz-W
*-SV40-TF-S	*-minR-FP-E2A-TF-bG	*-minC-luc-P2A-TF-bG	*-RSV-luc-F2A-TF-W
*-SV40-rec-S	*-minR-FP-E2A-rec-bG	*-minC-luc-P2A-rec-bG	*-RSV-luc-F2A-rec-W
*-SV40-CTP-S	*-minR-FP-E2A-CTP-bG	*-minC-luc-P2A-CTP-bG	*-RSV-luc-F2A-CTP-W
*-SV40-SM-S	*-minR-FP-E2A-SM-bG	*-minC-luc-P2A-SM-bG	*-RSV-luc-F2A-SM-W
*-SV40-NT-S	*-minR-FP-E2A-NT-bG	*-minC-luc-P2A-NT-bG	*-RSV-luc-F2A-NT-W
*-SV40-CR-S	*-minR-FP-E2A-CR-bG	*-minC-luc-P2A-CR-bG	*-RSV-luc-F2A-CR-W
*-SV40-chR-S	*-minR-FP-E2A-chR-bG	*-minC-luc-P2A-chR-bG	*-RSV-luc-F2A-chR-W
*-SV40-gRNA-S	*-minR-FP-E2A-gRNA-bG	*-minC-luc-P2A-gRNA-bG	*-RSV-luc-F2A-gRNA-W
*-SV40-Nuc-S	*-minR-FP-E2A-Nuc-bG	*-minC-luc-P2A-Nuc-bG	*-RSV-luc-F2A-Nuc-W
*-H68-enz-S	*-CMV-FP-E2A-enz-bG	*-mut-luc-P2A-enz-bG	*-minB-luc-F2A-enz-bG
*-H68-TF-S	*-CMV-FP-E2A-TF-bG	*-mut-luc-P2A-TF-bG	*-minB-luc-F2A-TF-bG
*-H68-rec-S	*-CMV-FP-E2A-rec-bG	*-mut-luc-P2A-rec-bG	*-minB-luc-F2A-rec-bG
*-H68-CTP-S	*-CMV-FP-E2A-CTP-bG	*-mut-luc-P2A-CTP-bG	*-minB-luc-F2A-CTP-bG
*-H68-SM-S	*-CMV-FP-E2A-SM-bG	*-mut-luc-P2A-SM-bG	*-minB-luc-F2A-SM-bG
*-H68-NT-S	*-CMV-FP-E2A-NT-bG	*-mut-luc-P2A-NT-bG	*-minB-luc-F2A-NT-bG
*-H68-CR-S	*-CMV-FP-E2A-CR-bG	*-mut-luc-P2A-CR-bG	*-minB-luc-F2A-CR-bG
*-H68-chR-S	*-CMV-FP-E2A-chR-bG	*-mut-luc-P2A-chR-bG	*-minB-luc-F2A-chR-bG
*-H68-gRNA-S	*-CMV-FP-E2A-gRNA-bG	*-mut-luc-P2A-gRNA-bG	*-minB-luc-F2A-gRNA-bG
*-H68-Nuc-S	*-CMV-FP-E2A-Nuc-bG	*-mut-luc-P2A-Nuc-bG	*-minB-luc-F2A-Nuc-bG
*-RSV-enz-S	*-SV40-FP-E2A-enz-bG	*-minR-luc-P2A-enz-bG	*-minC-luc-F2A-enz-bG
*-RSV-TF-S	*-SV40-FP-E2A-TF-bG	*-minR-luc-P2A-TF-bG	*-minC-luc-F2A-TF-bG
*-RSV-rec-S	*-SV40-FP-E2A-rec-bG	*-minR-luc-P2A-rec-bG	*-minC-luc-F2A-rec-bG
*-RSV-CTP-S	*-SV40-FP-E2A-CTP-bG	*-minR-luc-P2A-CTP-bG	*-minC-luc-F2A-CTP-bG
*-RSV-SM-S	*-SV40-FP-E2A-SM-bG	*-minR-luc-P2A-SM-bG	*-minC-luc-F2A-SM-bG

FIG. 16 cont'd

*-RSV-NT-S	*-SV40-FP-E2A-NT-bG	*-minR-luc-P2A-NT-bG	*-minC-luc-F2A-NT-bG
*-RSV-CR-S	*-SV40-FP-E2A-CR-bG	*-minR-luc-P2A-CR-bG	*-minC-luc-F2A-CR-bG
*-RSV-chR-S	*-SV40-FP-E2A-chR-bG	*-minR-luc-P2A-chR-bG	*-minC-luc-F2A-chR-bG
*-RSV-gRNA-S	*-SV40-FP-E2A-gRNA-bG	*-minR-luc-P2A-gRNA-bG	*-minC-luc-F2A-gRNA-bG
*-RSV-Nuc-S	*-SV40-FP-E2A-Nuc-bG	*-minR-luc-P2A-Nuc-bG	*-minC-luc-F2A-Nuc-bG
*-minB-enz-W-S	*-H68-FP-E2A-enz-bG	*-CMV-luc-P2A-enz-bG	*-mut-luc-F2A-enz-bG
*-minB-TF-W-S	*-H68-FP-E2A-TF-bG	*-CMV-luc-P2A-TF-bG	*-mut-luc-F2A-TF-bG
*-minB-rec-W-S	*-H68-FP-E2A-rec-bG	*-CMV-luc-P2A-rec-bG	*-mut-luc-F2A-rec-bG
*-minB-CTP-W-S	*-H68-FP-E2A-CTP-bG	*-CMV-luc-P2A-CTP-bG	*-mut-luc-F2A-CTP-bG
*-minB-SM-W-S	*-H68-FP-E2A-SM-bG	*-CMV-luc-P2A-SM-bG	*-mut-luc-F2A-SM-bG
*-minB-NT-W-S	*-H68-FP-E2A-NT-bG	*-CMV-luc-P2A-NT-bG	*-mut-luc-F2A-NT-bG
*-minB-CR-W-S	*-H68-FP-E2A-CR-bG	*-CMV-luc-P2A-CR-bG	*-mut-luc-F2A-CR-bG
*-minB-chR-W-S	*-H68-FP-E2A-chR-bG	*-CMV-luc-P2A-chR-bG	*-mut-luc-F2A-chR-bG
*-minB-gRNA-W-S	*-H68-FP-E2A-gRNA-bG	*-CMV-luc-P2A-gRNA-bG	*-mut-luc-F2A-gRNA-bG
*-minB-Nuc-W-S	*-H68-FP-E2A-Nuc-bG	*-CMV-luc-P2A-Nuc-bG	*-mut-luc-F2A-Nuc-bG
*-minC-enz-W-S	*-RSV-FP-E2A-enz-bG	*-SV40-luc-P2A-enz-bG	*-minR-luc-F2A-enz-bG
*-minC-TF-W-S	*-RSV-FP-E2A-TF-bG	*-SV40-luc-P2A-TF-bG	*-minR-luc-F2A-TF-bG
*-minC-rec-W-S	*-RSV-FP-E2A-rec-bG	*-SV40-luc-P2A-rec-bG	*-minR-luc-F2A-rec-bG
*-minC-CTP-W-S	*-RSV-FP-E2A-CTP-bG	*-SV40-luc-P2A-CTP-bG	*-minR-luc-F2A-CTP-bG
*-minC-SM-W-S	*-RSV-FP-E2A-SM-bG	*-SV40-luc-P2A-SM-bG	*-minR-luc-F2A-SM-bG
*-minC-NT-W-S	*-RSV-FP-E2A-NT-bG	*-SV40-luc-P2A-NT-bG	*-minR-luc-F2A-NT-bG
*-minC-CR-W-S	*-RSV-FP-E2A-CR-bG	*-SV40-luc-P2A-CR-bG	*-minR-luc-F2A-CR-bG
*-minC-chR-W-S	*-RSV-FP-E2A-chR-bG	*-SV40-luc-P2A-chR-bG	*-minR-luc-F2A-chR-bG
*-minC-gRNA-W-S	*-RSV-FP-E2A-gRNA-bG	*-SV40-luc-P2A-gRNA-bG	*-minR-luc-F2A-gRNA-bG
*-minC-Nuc-W-S	*-RSV-FP-E2A-Nuc-bG	*-SV40-luc-P2A-Nuc-bG	*-minR-luc-F2A-Nuc-bG
*-mut-enz-W-S	*-minB-FP-E2A-enz-S	*-H68-luc-P2A-enz-bG	*-CMV-luc-F2A-enz-bG
*-mut-TF-W-S	*-minB-FP-E2A-TF-S	*-H68-luc-P2A-TF-bG	*-CMV-luc-F2A-TF-bG
*-mut-rec-W-S	*-minB-FP-E2A-rec-S	*-H68-luc-P2A-rec-bG	*-CMV-luc-F2A-rec-bG
*-mut-CTP-W-S	*-minB-FP-E2A-CTP-S	*-H68-luc-P2A-CTP-bG	*-CMV-luc-F2A-CTP-bG
*-mut-SM-W-S	*-minB-FP-E2A-SM-S	*-H68-luc-P2A-SM-bG	*-CMV-luc-F2A-SM-bG
*-mut-NT-W-S	*-minB-FP-E2A-NT-S	*-H68-luc-P2A-NT-bG	*-CMV-luc-F2A-NT-bG
*-mut-CR-W-S	*-minB-FP-E2A-CR-S	*-H68-luc-P2A-CR-bG	*-CMV-luc-F2A-CR-bG
*-mut-chR-W-S	*-minB-FP-E2A-chR-S	*-H68-luc-P2A-chR-bG	*-CMV-luc-F2A-chR-bG
*-mut-gRNA-W-S	*-minB-FP-E2A-gRNA-S	*-H68-luc-P2A-gRNA-bG	*-CMV-luc-F2A-gRNA-bG

FIG. 16 cont'd

*-mut-Nuc-W-S	*-minB-FP-E2A-Nuc-S	*-H68-luc-P2A-Nuc-bG	*-CMV-luc-F2A-Nuc-bG
*-minR-enz-W-S	*-minC-FP-E2A-enz-S	*-RSV-luc-P2A-enz-bG	*-SV40-luc-F2A-enz-bG
*-minR-TF-W-S	*-minC-FP-E2A-TF-S	*-RSV-luc-P2A-TF-bG	*-SV40-luc-F2A-TF-bG
*-minR-rec-W-S	*-minC-FP-E2A-rec-S	*-RSV-luc-P2A-rec-bG	*-SV40-luc-F2A-rec-bG
*-minR-CTP-W-S	*-minC-FP-E2A-CTP-S	*-RSV-luc-P2A-CTP-bG	*-SV40-luc-F2A-CTP-bG
*-minR-SM-W-S	*-minC-FP-E2A-SM-S	*-RSV-luc-P2A-SM-bG	*-SV40-luc-F2A-SM-bG
*-minR-NT-W-S	*-minC-FP-E2A-NT-S	*-RSV-luc-P2A-NT-bG	*-SV40-luc-F2A-NT-bG
*-minR-CR-W-S	*-minC-FP-E2A-CR-S	*-RSV-luc-P2A-CR-bG	*-SV40-luc-F2A-CR-bG
*-minR-chR-W-S	*-minC-FP-E2A-chR-S	*-RSV-luc-P2A-chR-bG	*-SV40-luc-F2A-chR-bG
*-minR-gRNA-W-S	*-minC-FP-E2A-gRNA-S	*-RSV-luc-P2A-gRNA-bG	*-SV40-luc-F2A-gRNA-bG
*-minR-Nuc-W-S	*-minC-FP-E2A-Nuc-S	*-RSV-luc-P2A-Nuc-bG	*-SV40-luc-F2A-Nuc-bG
*-minB-enz-bG	*-mut-FP-E2A-enz-S	*-minB-luc-P2A-enz-S	*-H68-luc-F2A-enz-bG
*-minB-TF-bG	*-mut-FP-E2A-TF-S	*-minB-luc-P2A-TF-S	*-H68-luc-F2A-TF-bG
*-minB-rec-bG	*-mut-FP-E2A-rec-S	*-minB-luc-P2A-rec-S	*-H68-luc-F2A-rec-bG
*-minB-CTP-bG	*-mut-FP-E2A-CTP-S	*-minB-luc-P2A-CTP-S	*-H68-luc-F2A-CTP-bG
*-minB-SM-bG	*-mut-FP-E2A-SM-S	*-minB-luc-P2A-SM-S	*-H68-luc-F2A-SM-bG
*-minB-NT-bG	*-mut-FP-E2A-NT-S	*-minB-luc-P2A-NT-S	*-H68-luc-F2A-NT-bG
*-minB-CR-bG	*-mut-FP-E2A-CR-S	*-minB-luc-P2A-CR-S	*-H68-luc-F2A-CR-bG
*-minB-chR-bG	*-mut-FP-E2A-chR-S	*-minB-luc-P2A-chR-S	*-H68-luc-F2A-chR-bG
*-minB-gRNA-bG	*-mut-FP-E2A-gRNA-S	*-minB-luc-P2A-gRNA-S	*-H68-luc-F2A-gRNA-bG
*-minB-Nuc-bG	*-mut-FP-E2A-Nuc-S	*-minB-luc-P2A-Nuc-S	*-H68-luc-F2A-Nuc-bG
*-minC-enz-bG	*-minR-FP-E2A-enz-S	*-minC-luc-P2A-enz-S	*-RSV-luc-F2A-enz-bG
*-minC-TF-bG	*-minR-FP-E2A-TF-S	*-minC-luc-P2A-TF-S	*-RSV-luc-F2A-TF-bG
*-minC-rec-bG	*-minR-FP-E2A-rec-S	*-minC-luc-P2A-rec-S	*-RSV-luc-F2A-rec-bG
*-minC-CTP-bG	*-minR-FP-E2A-CTP-S	*-minC-luc-P2A-CTP-S	*-RSV-luc-F2A-CTP-bG
*-minC-SM-bG	*-minR-FP-E2A-SM-S	*-minC-luc-P2A-SM-S	*-RSV-luc-F2A-SM-bG
*-minC-NT-bG	*-minR-FP-E2A-NT-S	*-minC-luc-P2A-NT-S	*-RSV-luc-F2A-NT-bG
*-minC-CR-bG	*-minR-FP-E2A-CR-S	*-minC-luc-P2A-CR-S	*-RSV-luc-F2A-CR-bG
*-minC-chR-bG	*-minR-FP-E2A-chR-S	*-minC-luc-P2A-chR-S	*-RSV-luc-F2A-chR-bG
*-minC-gRNA-bG	*-minR-FP-E2A-gRNA-S	*-minC-luc-P2A-gRNA-S	*-RSV-luc-F2A-gRNA-bG
*-minC-Nuc-bG	*-minR-FP-E2A-Nuc-S	*-minC-luc-P2A-Nuc-S	*-RSV-luc-F2A-Nuc-bG
*-mut-enz-bG	*-CMV-FP-E2A-enz-S	*-mut-luc-P2A-enz-S	*-minB-luc-F2A-enz-S
*-mut-TF-bG	*-CMV-FP-E2A-TF-S	*-mut-luc-P2A-TF-S	*-minB-luc-F2A-TF-S
*-mut-rec-bG	*-CMV-FP-E2A-rec-S	*-mut-luc-P2A-rec-S	*-minB-luc-F2A-rec-S

FIG. 16 cont'd

*-mut-CTP-bG	*-CMV-FP-E2A-CTP-S	*-mut-luc-P2A-CTP-S	*-minB-luc-F2A-CTP-S
*-mut-SM-bG	*-CMV-FP-E2A-SM-S	*-mut-luc-P2A-SM-S	*-minB-luc-F2A-SM-S
*-mut-NT-bG	*-CMV-FP-E2A-NT-S	*-mut-luc-P2A-NT-S	*-minB-luc-F2A-NT-S
*-mut-CR-bG	*-CMV-FP-E2A-CR-S	*-mut-luc-P2A-CR-S	*-minB-luc-F2A-CR-S
*-mut-chR-bG	*-CMV-FP-E2A-chR-S	*-mut-luc-P2A-chR-S	*-minB-luc-F2A-chR-S
*-mut-gRNA-bG	*-CMV-FP-E2A-gRNA-S	*-mut-luc-P2A-gRNA-S	*-minB-luc-F2A-gRNA-S
*-mut-Nuc-bG	*-CMV-FP-E2A-Nuc-S	*-mut-luc-P2A-Nuc-S	*-minB-luc-F2A-Nuc-S
*-minR-enz-bG	*-SV40-FP-E2A-enz-S	*-minR-luc-P2A-enz-S	*-minC-luc-F2A-enz-S
*-minR-TF-bG	*-SV40-FP-E2A-TF-S	*-minR-luc-P2A-TF-S	*-minC-luc-F2A-TF-S
*-minR-rec-bG	*-SV40-FP-E2A-rec-S	*-minR-luc-P2A-rec-S	*-minC-luc-F2A-rec-S
*-minR-CTP-bG	*-SV40-FP-E2A-CTP-S	*-minR-luc-P2A-CTP-S	*-minC-luc-F2A-CTP-S
*-minR-SM-bG	*-SV40-FP-E2A-SM-S	*-minR-luc-P2A-SM-S	*-minC-luc-F2A-SM-S
*-minR-NT-bG	*-SV40-FP-E2A-NT-S	*-minR-luc-P2A-NT-S	*-minC-luc-F2A-NT-S
*-minR-CR-bG	*-SV40-FP-E2A-CR-S	*-minR-luc-P2A-CR-S	*-minC-luc-F2A-CR-S
*-minR-chR-bG	*-SV40-FP-E2A-chR-S	*-minR-luc-P2A-chR-S	*-minC-luc-F2A-chR-S
*-minR-gRNA-bG	*-SV40-FP-E2A-gRNA-S	*-minR-luc-P2A-gRNA-S	*-minC-luc-F2A-gRNA-S
*-minR-Nuc-bG	*-SV40-FP-E2A-Nuc-S	*-minR-luc-P2A-Nuc-S	*-minC-luc-F2A-Nuc-S
*-CMV-enz-bG	*-H68-FP-E2A-enz-S	*-CMV-luc-P2A-enz-S	*-mut-luc-F2A-enz-S
*-CMV-TF-bG	*-H68-FP-E2A-TF-S	*-CMV-luc-P2A-TF-S	*-mut-luc-F2A-TF-S
*-CMV-rec-bG	*-H68-FP-E2A-rec-S	*-CMV-luc-P2A-rec-S	*-mut-luc-F2A-rec-S
*-CMV-CTP-bG	*-H68-FP-E2A-CTP-S	*-CMV-luc-P2A-CTP-S	*-mut-luc-F2A-CTP-S
*-CMV-SM-bG	*-H68-FP-E2A-SM-S	*-CMV-luc-P2A-SM-S	*-mut-luc-F2A-SM-S
*-CMV-NT-bG	*-H68-FP-E2A-NT-S	*-CMV-luc-P2A-NT-S	*-mut-luc-F2A-NT-S
*-CMV-CR-bG	*-H68-FP-E2A-CR-S	*-CMV-luc-P2A-CR-S	*-mut-luc-F2A-CR-S
*-CMV-chR-bG	*-H68-FP-E2A-chR-S	*-CMV-luc-P2A-chR-S	*-mut-luc-F2A-chR-S
*-CMV-gRNA-bG	*-H68-FP-E2A-gRNA-S	*-CMV-luc-P2A-gRNA-S	*-mut-luc-F2A-gRNA-S
*-CMV-Nuc-bG	*-H68-FP-E2A-Nuc-S	*-CMV-luc-P2A-Nuc-S	*-mut-luc-F2A-Nuc-S
*-SV40-enz-bG	*-RSV-FP-E2A-enz-S	*-SV40-luc-P2A-enz-S	*-minR-luc-F2A-enz-S
*-SV40-TF-bG	*-RSV-FP-E2A-TF-S	*-SV40-luc-P2A-TF-S	*-minR-luc-F2A-TF-S
*-SV40-rec-bG	*-RSV-FP-E2A-rec-S	*-SV40-luc-P2A-rec-S	*-minR-luc-F2A-rec-S
*-SV40-CTP-bG	*-RSV-FP-E2A-CTP-S	*-SV40-luc-P2A-CTP-S	*-minR-luc-F2A-CTP-S
*-SV40-SM-bG	*-RSV-FP-E2A-SM-S	*-SV40-luc-P2A-SM-S	*-minR-luc-F2A-SM-S
*-SV40-NT-bG	*-RSV-FP-E2A-NT-S	*-SV40-luc-P2A-NT-S	*-minR-luc-F2A-NT-S
*-SV40-CR-bG	*-RSV-FP-E2A-CR-S	*-SV40-luc-P2A-CR-S	*-minR-luc-F2A-CR-S

FIG. 16 cont'd

*-SV40-chR-bG	*-RSV-FP-E2A-chR-S	*-SV40-luc-P2A-chR-S	*-minR-luc-F2A-chR-S
*-SV40-gRNA-bG	*-RSV-FP-E2A-gRNA-S	*-SV40-luc-P2A-gRNA-S	*-minR-luc-F2A-gRNA-S
*-SV40-Nuc-bG	*-RSV-FP-E2A-Nuc-S	*-SV40-luc-P2A-Nuc-S	*-minR-luc-F2A-Nuc-S
*-H68-enz-bG	*-minB-FP-E2A-enz-W-bG	*-H68-luc-P2A-enz-S	*-CMV-luc-F2A-enz-S
*-H68-TF-bG	*-minB-FP-E2A-TF-W-bG	*-H68-luc-P2A-TF-S	*-CMV-luc-F2A-TF-S
*-H68-rec-bG	*-minB-FP-E2A-rec-W-bG	*-H68-luc-P2A-rec-S	*-CMV-luc-F2A-rec-S
*-H68-CTP-bG	*-minB-FP-E2A-CTP-W-bG	*-H68-luc-P2A-CTP-S	*-CMV-luc-F2A-CTP-S
*-H68-SM-bG	*-minB-FP-E2A-SM-W-bG	*-H68-luc-P2A-SM-S	*-CMV-luc-F2A-SM-S
*-H68-NT-bG	*-minB-FP-E2A-NT-W-bG	*-H68-luc-P2A-NT-S	*-CMV-luc-F2A-NT-S
*-H68-CR-bG	*-minB-FP-E2A-CR-W-bG	*-H68-luc-P2A-CR-S	*-CMV-luc-F2A-CR-S
*-H68-chR-bG	*-minB-FP-E2A-chR-W-bG	*-H68-luc-P2A-chR-S	*-CMV-luc-F2A-chR-S
*-H68-gRNA-bG	*-minB-FP-E2A-gRNA-W-bG	*-H68-luc-P2A-gRNA-S	*-CMV-luc-F2A-gRNA-S
*-H68-Nuc-bG	*-minB-FP-E2A-Nuc-W-bG	*-H68-luc-P2A-Nuc-S	*-CMV-luc-F2A-Nuc-S
*-RSV-enz-bG	*-minC-FP-E2A-enz-W-bG	*-RSV-luc-P2A-enz-S	*-SV40-luc-F2A-enz-S
*-RSV-TF-bG	*-minC-FP-E2A-TF-W-bG	*-RSV-luc-P2A-TF-S	*-SV40-luc-F2A-TF-S
*-RSV-rec-bG	*-minC-FP-E2A-rec-W-bG	*-RSV-luc-P2A-rec-S	*-SV40-luc-F2A-rec-S
*-RSV-CTP-bG	*-minC-FP-E2A-CTP-W-bG	*-RSV-luc-P2A-CTP-S	*-SV40-luc-F2A-CTP-S
*-RSV-SM-bG	*-minC-FP-E2A-SM-W-bG	*-RSV-luc-P2A-SM-S	*-SV40-luc-F2A-SM-S
*-RSV-NT-bG	*-minC-FP-E2A-NT-W-bG	*-RSV-luc-P2A-NT-S	*-SV40-luc-F2A-NT-S
*-RSV-CR-bG	*-minC-FP-E2A-CR-W-bG	*-RSV-luc-P2A-CR-S	*-SV40-luc-F2A-CR-S
*-RSV-chR-bG	*-minC-FP-E2A-chR-W-bG	*-RSV-luc-P2A-chR-S	*-SV40-luc-F2A-chR-S
*-RSV-gRNA-bG	*-minC-FP-E2A-gRNA-W-bG	*-RSV-luc-P2A-gRNA-S	*-SV40-luc-F2A-gRNA-S
*-RSV-Nuc-bG	*-minC-FP-E2A-Nuc-W-bG	*-RSV-luc-P2A-Nuc-S	*-SV40-luc-F2A-Nuc-S
*-minB-enz-W-bG	*-mut-FP-E2A-enz-W-bG	*-minB-luc-P2A-enz-W-bG	*-H68-luc-F2A-enz-S
*-minB-TF-W-bG	*-mut-FP-E2A-TF-W-bG	*-minB-luc-P2A-TF-W-bG	*-H68-luc-F2A-TF-S
*-minB-rec-W-bG	*-mut-FP-E2A-rec-W-bG	*-minB-luc-P2A-rec-W-bG	*-H68-luc-F2A-rec-S
*-minB-CTP-W-bG	*-mut-FP-E2A-CTP-W-bG	*-minB-luc-P2A-CTP-W-bG	*-H68-luc-F2A-CTP-S
*-minB-SM-W-bG	*-mut-FP-E2A-SM-W-bG	*-minB-luc-P2A-SM-W-bG	*-H68-luc-F2A-SM-S
*-minB-NT-W-bG	*-mut-FP-E2A-NT-W-bG	*-minB-luc-P2A-NT-W-bG	*-H68-luc-F2A-NT-S
*-minB-CR-W-bG	*-mut-FP-E2A-CR-W-bG	*-minB-luc-P2A-CR-W-bG	*-H68-luc-F2A-CR-S
*-minB-chR-W-bG	*-mut-FP-E2A-chR-W-bG	*-minB-luc-P2A-chR-W-bG	*-H68-luc-F2A-chR-S
*-minB-gRNA-W-bG	*-mut-FP-E2A-gRNA-W-bG	*-minB-luc-P2A-gRNA-W-bG	*-H68-luc-F2A-gRNA-S
*-minB-Nuc-W-bG	*-mut-FP-E2A-Nuc-W-bG	*-minB-luc-P2A-Nuc-W-bG	*-H68-luc-F2A-Nuc-S
*-minC-enz-W-bG	*-minR-FP-E2A-enz-W-bG	*-minC-luc-P2A-enz-W-bG	*-RSV-luc-F2A-enz-S

FIG. 16 cont'd

*-minC-TF-W-bG	*-minR-FP-E2A-TF-W-bG	*-minC-luc-P2A-TF-W-bG	*-RSV-luc-F2A-TF-S
*-minC-rec-W-bG	*-minR-FP-E2A-rec-W-bG	*-minC-luc-P2A-rec-W-bG	*-RSV-luc-F2A-rec-S
*-minC-CTP-W-bG	*-minR-FP-E2A-CTP-W-bG	*-minC-luc-P2A-CTP-W-bG	*-RSV-luc-F2A-CTP-S
*-minC-SM-W-bG	*-minR-FP-E2A-SM-W-bG	*-minC-luc-P2A-SM-W-bG	*-RSV-luc-F2A-SM-S
*-minC-NT-W-bG	*-minR-FP-E2A-NT-W-bG	*-minC-luc-P2A-NT-W-bG	*-RSV-luc-F2A-NT-S
*-minC-CR-W-bG	*-minR-FP-E2A-CR-W-bG	*-minC-luc-P2A-CR-W-bG	*-RSV-luc-F2A-CR-S
*-minC-chR-W-bG	*-minR-FP-E2A-chR-W-bG	*-minC-luc-P2A-chR-W-bG	*-RSV-luc-F2A-chR-S
*-minC-gRNA-W-bG	*-minR-FP-E2A-gRNA-W-bG	*-minC-luc-P2A-gRNA-W-bG	*-RSV-luc-F2A-gRNA-S
*-minC-Nuc-W-bG	*-minR-FP-E2A-Nuc-W-bG	*-minC-luc-P2A-Nuc-W-bG	*-RSV-luc-F2A-Nuc-S
*-mut-enz-W-bG	*-CMV-FP-E2A-enz-W-bG	*-mut-luc-P2A-enz-W-bG	*-minB-luc-F2A-enz-W-bG
*-mut-TF-W-bG	*-CMV-FP-E2A-TF-W-bG	*-mut-luc-P2A-TF-W-bG	*-minB-luc-F2A-TF-W-bG
*-mut-rec-W-bG	*-CMV-FP-E2A-rec-W-bG	*-mut-luc-P2A-rec-W-bG	*-minB-luc-F2A-rec-W-bG
*-mut-CTP-W-bG	*-CMV-FP-E2A-CTP-W-bG	*-mut-luc-P2A-CTP-W-bG	*-minB-luc-F2A-CTP-W-bG
*-mut-SM-W-bG	*-CMV-FP-E2A-SM-W-bG	*-mut-luc-P2A-SM-W-bG	*-minB-luc-F2A-SM-W-bG
*-mut-NT-W-bG	*-CMV-FP-E2A-NT-W-bG	*-mut-luc-P2A-NT-W-bG	*-minB-luc-F2A-NT-W-bG
*-mut-CR-W-bG	*-CMV-FP-E2A-CR-W-bG	*-mut-luc-P2A-CR-W-bG	*-minB-luc-F2A-CR-W-bG
*-mut-chR-W-bG	*-CMV-FP-E2A-chR-W-bG	*-mut-luc-P2A-chR-W-bG	*-minB-luc-F2A-chR-W-bG
*-mut-gRNA-W-bG	*-CMV-FP-E2A-gRNA-W-bG	*-mut-luc-P2A-gRNA-W-bG	*-minB-luc-F2A-gRNA-W-bG
*-mut-Nuc-W-bG	*-CMV-FP-E2A-Nuc-W-bG	*-mut-luc-P2A-Nuc-W-bG	*-minB-luc-F2A-Nuc-W-bG
*-minR-enz-W-bG	*-SV40-FP-E2A-enz-W-bG	*-minR-luc-P2A-enz-W-bG	*-minC-luc-F2A-enz-W-bG
*-minR-TF-W-bG	*-SV40-FP-E2A-TF-W-bG	*-minR-luc-P2A-TF-W-bG	*-minC-luc-F2A-TF-W-bG
*-minR-rec-W-bG	*-SV40-FP-E2A-rec-W-bG	*-minR-luc-P2A-rec-W-bG	*-minC-luc-F2A-rec-W-bG
*-minR-CTP-W-bG	*-SV40-FP-E2A-CTP-W-bG	*-minR-luc-P2A-CTP-W-bG	*-minC-luc-F2A-CTP-W-bG
*-minR-SM-W-bG	*-SV40-FP-E2A-SM-W-bG	*-minR-luc-P2A-SM-W-bG	*-minC-luc-F2A-SM-W-bG
*-minR-NT-W-bG	*-SV40-FP-E2A-NT-W-bG	*-minR-luc-P2A-NT-W-bG	*-minC-luc-F2A-NT-W-bG
*-minR-CR-W-bG	*-SV40-FP-E2A-CR-W-bG	*-minR-luc-P2A-CR-W-bG	*-minC-luc-F2A-CR-W-bG
*-minR-chR-W-bG	*-SV40-FP-E2A-chR-W-bG	*-minR-luc-P2A-chR-W-bG	*-minC-luc-F2A-chR-W-bG
*-minR-gRNA-W-bG	*-SV40-FP-E2A-gRNA-W-bG	*-minR-luc-P2A-gRNA-W-bG	*-minC-luc-F2A-gRNA-W-bG
*-minR-Nuc-W-bG	*-SV40-FP-E2A-Nuc-W-bG	*-minR-luc-P2A-Nuc-W-bG	*-minC-luc-F2A-Nuc-W-bG
*-CMV-enz-W-bG	*-H68-FP-E2A-enz-W-bG	*-CMV-luc-P2A-enz-W-bG	*-mut-luc-F2A-enz-W-bG
*-CMV-TF-W-bG	*-H68-FP-E2A-TF-W-bG	*-CMV-luc-P2A-TF-W-bG	*-mut-luc-F2A-TF-W-bG
*-CMV-rec-W-bG	*-H68-FP-E2A-rec-W-bG	*-CMV-luc-P2A-rec-W-bG	*-mut-luc-F2A-rec-W-bG
*-CMV-CTP-W-bG	*-H68-FP-E2A-CTP-W-bG	*-CMV-luc-P2A-CTP-W-bG	*-mut-luc-F2A-CTP-W-bG
*-CMV-SM-W-bG	*-H68-FP-E2A-SM-W-bG	*-CMV-luc-P2A-SM-W-bG	*-mut-luc-F2A-SM-W-bG

FIG. 16 cont'd

*-CMV-NT-W-bG	*-H68-FP-E2A-NT-W-bG	*-CMV-luc-P2A-NT-W-bG	*-mut-luc-F2A-NT-W-bG
*-CMV-CR-W-bG	*-H68-FP-E2A-CR-W-bG	*-CMV-luc-P2A-CR-W-bG	*-mut-luc-F2A-CR-W-bG
*-CMV-chR-W-bG	*-H68-FP-E2A-chR-W-bG	*-CMV-luc-P2A-chR-W-bG	*-mut-luc-F2A-chR-W-bG
*-CMV-gRNA-W-bG	*-H68-FP-E2A-gRNA-W-bG	*-CMV-luc-P2A-gRNA-W-bG	*-mut-luc-F2A-gRNA-W-bG
*-CMV-Nuc-W-bG	*-H68-FP-E2A-Nuc-W-bG	*-CMV-luc-P2A-Nuc-W-bG	*-mut-luc-F2A-Nuc-W-bG
*-SV40-enz-W-bG	*-RSV-FP-E2A-enz-W-bG	*-SV40-luc-P2A-enz-W-bG	*-minR-luc-F2A-enz-W-bG
*-SV40-TF-W-bG	*-RSV-FP-E2A-TF-W-bG	*-SV40-luc-P2A-TF-W-bG	*-minR-luc-F2A-TF-W-bG
*-SV40-rec-W-bG	*-RSV-FP-E2A-rec-W-bG	*-SV40-luc-P2A-rec-W-bG	*-minR-luc-F2A-rec-W-bG
*-SV40-CTP-W-bG	*-RSV-FP-E2A-CTP-W-bG	*-SV40-luc-P2A-CTP-W-bG	*-minR-luc-F2A-CTP-W-bG
*-SV40-SM-W-bG	*-RSV-FP-E2A-SM-W-bG	*-SV40-luc-P2A-SM-W-bG	*-minR-luc-F2A-SM-W-bG
*-SV40-NT-W-bG	*-RSV-FP-E2A-NT-W-bG	*-SV40-luc-P2A-NT-W-bG	*-minR-luc-F2A-NT-W-bG
*-SV40-CR-W-bG	*-RSV-FP-E2A-CR-W-bG	*-SV40-luc-P2A-CR-W-bG	*-minR-luc-F2A-CR-W-bG
*-SV40-chR-W-bG	*-RSV-FP-E2A-chR-W-bG	*-SV40-luc-P2A-chR-W-bG	*-minR-luc-F2A-chR-W-bG
*-SV40-gRNA-W-bG	*-RSV-FP-E2A-gRNA-W-bG	*-SV40-luc-P2A-gRNA-W-bG	*-minR-luc-F2A-gRNA-W-bG
*-SV40-Nuc-W-bG	*-RSV-FP-E2A-Nuc-W-bG	*-SV40-luc-P2A-Nuc-W-bG	*-minR-luc-F2A-Nuc-W-bG
*-H68-enz-W-bG	*-minB-FP-E2A-enz-W-S	*-H68-luc-P2A-enz-W-bG	*-CMV-luc-F2A-enz-W-bG
*-H68-TF-W-bG	*-minB-FP-E2A-TF-W-S	*-H68-luc-P2A-TF-W-bG	*-CMV-luc-F2A-TF-W-bG
*-H68-rec-W-bG	*-minB-FP-E2A-rec-W-S	*-H68-luc-P2A-rec-W-bG	*-CMV-luc-F2A-rec-W-bG
*-H68-CTP-W-bG	*-minB-FP-E2A-CTP-W-S	*-H68-luc-P2A-CTP-W-bG	*-CMV-luc-F2A-CTP-W-bG
*-H68-SM-W-bG	*-minB-FP-E2A-SM-W-S	*-H68-luc-P2A-SM-W-bG	*-CMV-luc-F2A-SM-W-bG
*-H68-NT-W-bG	*-minB-FP-E2A-NT-W-S	*-H68-luc-P2A-NT-W-bG	*-CMV-luc-F2A-NT-W-bG
*-H68-CR-W-bG	*-minB-FP-E2A-CR-W-S	*-H68-luc-P2A-CR-W-bG	*-CMV-luc-F2A-CR-W-bG
*-H68-chR-W-bG	*-minB-FP-E2A-chR-W-S	*-H68-luc-P2A-chR-W-bG	*-CMV-luc-F2A-chR-W-bG
*-H68-gRNA-W-bG	*-minB-FP-E2A-gRNA-W-S	*-H68-luc-P2A-gRNA-W-bG	*-CMV-luc-F2A-gRNA-W-bG
*-H68-Nuc-W-bG	*-minB-FP-E2A-Nuc-W-S	*-H68-luc-P2A-Nuc-W-bG	*-CMV-luc-F2A-Nuc-W-bG
*-RSV-enz-W-bG	*-minC-FP-E2A-enz-W-S	*-RSV-luc-P2A-enz-W-bG	*-SV40-luc-F2A-enz-W-bG
*-RSV-TF-W-bG	*-minC-FP-E2A-TF-W-S	*-RSV-luc-P2A-TF-W-bG	*-SV40-luc-F2A-TF-W-bG
*-RSV-rec-W-bG	*-minC-FP-E2A-rec-W-S	*-RSV-luc-P2A-rec-W-bG	*-SV40-luc-F2A-rec-W-bG
*-RSV-CTP-W-bG	*-minC-FP-E2A-CTP-W-S	*-RSV-luc-P2A-CTP-W-bG	*-SV40-luc-F2A-CTP-W-bG
*-RSV-SM-W-bG	*-minC-FP-E2A-SM-W-S	*-RSV-luc-P2A-SM-W-bG	*-SV40-luc-F2A-SM-W-bG
*-RSV-NT-W-bG	*-minC-FP-E2A-NT-W-S	*-RSV-luc-P2A-NT-W-bG	*-SV40-luc-F2A-NT-W-bG
*-RSV-CR-W-bG	*-minC-FP-E2A-CR-W-S	*-RSV-luc-P2A-CR-W-bG	*-SV40-luc-F2A-CR-W-bG
*-RSV-chR-W-bG	*-minC-FP-E2A-chR-W-S	*-RSV-luc-P2A-chR-W-bG	*-SV40-luc-F2A-chR-W-bG
*-RSV-gRNA-W-bG	*-minC-FP-E2A-gRNA-W-S	*-RSV-luc-P2A-gRNA-W-bG	*-SV40-luc-F2A-gRNA-W-bG

FIG. 16 cont'd

*-RSV-Nuc-W-bG	*-minC-FP-E2A-Nuc-W-S	*-RSV-luc-P2A-Nuc-W-bG	*-SV40-luc-F2A-Nuc-W-bG
*-CMV-enz-W-S	*-mut-FP-E2A-enz-W-S	*-minB-luc-P2A-enz-W-S	*-H68-luc-F2A-enz-W-bG
*-CMV-TF-W-S	*-mut-FP-E2A-TF-W-S	*-minB-luc-P2A-TF-W-S	*-H68-luc-F2A-TF-W-bG
*-CMV-rec-W-S	*-mut-FP-E2A-rec-W-S	*-minB-luc-P2A-rec-W-S	*-H68-luc-F2A-rec-W-bG
*-CMV-CTP-W-S	*-mut-FP-E2A-CTP-W-S	*-minB-luc-P2A-CTP-W-S	*-H68-luc-F2A-CTP-W-bG
*-CMV-SM-W-S	*-mut-FP-E2A-SM-W-S	*-minB-luc-P2A-SM-W-S	*-H68-luc-F2A-SM-W-bG
*-CMV-NT-W-S	*-mut-FP-E2A-NT-W-S	*-minB-luc-P2A-NT-W-S	*-H68-luc-F2A-NT-W-bG
*-CMV-CR-W-S	*-mut-FP-E2A-CR-W-S	*-minB-luc-P2A-CR-W-S	*-H68-luc-F2A-CR-W-bG
*-CMV-chR-W-S	*-mut-FP-E2A-chR-W-S	*-minB-luc-P2A-chR-W-S	*-H68-luc-F2A-chR-W-bG
*-CMV-gRNA-W-S	*-mut-FP-E2A-gRNA-W-S	*-minB-luc-P2A-gRNA-W-S	*-H68-luc-F2A-gRNA-W-bG
*-CMV-Nuc-W-S	*-mut-FP-E2A-Nuc-W-S	*-minB-luc-P2A-Nuc-W-S	*-H68-luc-F2A-Nuc-W-bG
*-SV40-enz-W-S	*-minR-FP-E2A-enz-W-S	*-minC-luc-P2A-enz-W-S	*-RSV-luc-F2A-enz-W-bG
*-SV40-TF-W-S	*-minR-FP-E2A-TF-W-S	*-minC-luc-P2A-TF-W-S	*-RSV-luc-F2A-TF-W-bG
*-SV40-rec-W-S	*-minR-FP-E2A-rec-W-S	*-minC-luc-P2A-rec-W-S	*-RSV-luc-F2A-rec-W-bG
*-SV40-CTP-W-S	*-minR-FP-E2A-CTP-W-S	*-minC-luc-P2A-CTP-W-S	*-RSV-luc-F2A-CTP-W-bG
*-SV40-SM-W-S	*-minR-FP-E2A-SM-W-S	*-minC-luc-P2A-SM-W-S	*-RSV-luc-F2A-SM-W-bG
*-SV40-NT-W-S	*-minR-FP-E2A-NT-W-S	*-minC-luc-P2A-NT-W-S	*-RSV-luc-F2A-NT-W-bG
*-SV40-CR-W-S	*-minR-FP-E2A-CR-W-S	*-minC-luc-P2A-CR-W-S	*-RSV-luc-F2A-CR-W-bG
*-SV40-chR-W-S	*-minR-FP-E2A-chR-W-S	*-minC-luc-P2A-chR-W-S	*-RSV-luc-F2A-chR-W-bG
*-SV40-gRNA-W-S	*-minR-FP-E2A-gRNA-W-S	*-minC-luc-P2A-gRNA-W-S	*-RSV-luc-F2A-gRNA-W-bG
*-SV40-Nuc-W-S	*-minR-FP-E2A-Nuc-W-S	*-minC-luc-P2A-Nuc-W-S	*-RSV-luc-F2A-Nuc-W-bG
*-H68-enz-W-S	*-CMV-FP-E2A-enz-W-S	*-mut-luc-P2A-enz-W-S	*-minB-luc-F2A-enz-W-S
*-H68-TF-W-S	*-CMV-FP-E2A-TF-W-S	*-mut-luc-P2A-TF-W-S	*-minB-luc-F2A-TF-W-S
*-H68-rec-W-S	*-CMV-FP-E2A-rec-W-S	*-mut-luc-P2A-rec-W-S	*-minB-luc-F2A-rec-W-S
*-H68-CTP-W-S	*-CMV-FP-E2A-CTP-W-S	*-mut-luc-P2A-CTP-W-S	*-minB-luc-F2A-CTP-W-S
*-H68-SM-W-S	*-CMV-FP-E2A-SM-W-S	*-mut-luc-P2A-SM-W-S	*-minB-luc-F2A-SM-W-S
*-H68-NT-W-S	*-CMV-FP-E2A-NT-W-S	*-mut-luc-P2A-NT-W-S	*-minB-luc-F2A-NT-W-S
*-H68-CR-W-S	*-CMV-FP-E2A-CR-W-S	*-mut-luc-P2A-CR-W-S	*-minB-luc-F2A-CR-W-S
*-H68-chR-W-S	*-CMV-FP-E2A-chR-W-S	*-mut-luc-P2A-chR-W-S	*-minB-luc-F2A-chR-W-S
*-H68-gRNA-W-S	*-CMV-FP-E2A-gRNA-W-S	*-mut-luc-P2A-gRNA-W-S	*-minB-luc-F2A-gRNA-W-S
*-H68-Nuc-W-S	*-CMV-FP-E2A-Nuc-W-S	*-mut-luc-P2A-Nuc-W-S	*-minB-luc-F2A-Nuc-W-S
*-RSV-enz-W-S	*-SV40-FP-E2A-enz-W-S	*-minR-luc-P2A-enz-W-S	*-minC-luc-F2A-enz-W-S
*-RSV-TF-W-S	*-SV40-FP-E2A-TF-W-S	*-minR-luc-P2A-TF-W-S	*-minC-luc-F2A-TF-W-S
*-RSV-rec-W-S	*-SV40-FP-E2A-rec-W-S	*-minR-luc-P2A-rec-W-S	*-minC-luc-F2A-rec-W-S

FIG. 16 cont'd

*-RSV-CTP-W-S	*-SV40-FP-E2A-CTP-W-S	*-minR-luc-P2A-CTP-W-S	*-minC-luc-F2A-CTP-W-S
*-RSV-SM-W-S	*-SV40-FP-E2A-SM-W-S	*-minR-luc-P2A-SM-W-S	*-minC-luc-F2A-SM-W-S
*-RSV-NT-W-S	*-SV40-FP-E2A-NT-W-S	*-minR-luc-P2A-NT-W-S	*-minC-luc-F2A-NT-W-S
*-RSV-CR-W-S	*-SV40-FP-E2A-CR-W-S	*-minR-luc-P2A-CR-W-S	*-minC-luc-F2A-CR-W-S
*-RSV-chR-W-S	*-SV40-FP-E2A-chR-W-S	*-minR-luc-P2A-chR-W-S	*-minC-luc-F2A-chR-W-S
*-RSV-gRNA-W-S	*-SV40-FP-E2A-gRNA-W-S	*-minR-luc-P2A-gRNA-W-S	*-minC-luc-F2A-gRNA-W-S
*-RSV-Nuc-W-S	*-SV40-FP-E2A-Nuc-W-S	*-minR-luc-P2A-Nuc-W-S	*-minC-luc-F2A-Nuc-W-S
*-minB-FP-P2A-enz-W	*-H68-FP-E2A-enz-W-S	*-CMV-luc-P2A-enz-W-S	*-mut-luc-F2A-enz-W-S
*-minB-FP-P2A-TF-W	*-H68-FP-E2A-TF-W-S	*-CMV-luc-P2A-TF-W-S	*-mut-luc-F2A-TF-W-S
*-minB-FP-P2A-rec-W	*-H68-FP-E2A-rec-W-S	*-CMV-luc-P2A-rec-W-S	*-mut-luc-F2A-rec-W-S
*-minB-FP-P2A-CTP-W	*-H68-FP-E2A-CTP-W-S	*-CMV-luc-P2A-CTP-W-S	*-mut-luc-F2A-CTP-W-S
*-minB-FP-P2A-SM-W	*-H68-FP-E2A-SM-W-S	*-CMV-luc-P2A-SM-W-S	*-mut-luc-F2A-SM-W-S
*-minB-FP-P2A-NT-W	*-H68-FP-E2A-NT-W-S	*-CMV-luc-P2A-NT-W-S	*-mut-luc-F2A-NT-W-S
*-minB-FP-P2A-CR-W	*-H68-FP-E2A-CR-W-S	*-CMV-luc-P2A-CR-W-S	*-mut-luc-F2A-CR-W-S
*-minB-FP-P2A-chR-W	*-H68-FP-E2A-chR-W-S	*-CMV-luc-P2A-chR-W-S	*-mut-luc-F2A-chR-W-S
*-minB-FP-P2A-gRNA-W	*-H68-FP-E2A-gRNA-W-S	*-CMV-luc-P2A-gRNA-W-S	*-mut-luc-F2A-gRNA-W-S
*-minB-FP-P2A-Nuc-W	*-H68-FP-E2A-Nuc-W-S	*-CMV-luc-P2A-Nuc-W-S	*-mut-luc-F2A-Nuc-W-S
*-minC-FP-P2A-enz-W	*-RSV-FP-E2A-enz-W-S	*-SV40-luc-P2A-enz-W-S	*-minR-luc-F2A-enz-W-S
*-minC-FP-P2A-TF-W	*-RSV-FP-E2A-TF-W-S	*-SV40-luc-P2A-TF-W-S	*-minR-luc-F2A-TF-W-S
*-minC-FP-P2A-rec-W	*-RSV-FP-E2A-rec-W-S	*-SV40-luc-P2A-rec-W-S	*-minR-luc-F2A-rec-W-S
*-minC-FP-P2A-CTP-W	*-RSV-FP-E2A-CTP-W-S	*-SV40-luc-P2A-CTP-W-S	*-minR-luc-F2A-CTP-W-S
*-minC-FP-P2A-SM-W	*-RSV-FP-E2A-SM-W-S	*-SV40-luc-P2A-SM-W-S	*-minR-luc-F2A-SM-W-S
*-minC-FP-P2A-NT-W	*-RSV-FP-E2A-NT-W-S	*-SV40-luc-P2A-NT-W-S	*-minR-luc-F2A-NT-W-S
*-minC-FP-P2A-CR-W	*-RSV-FP-E2A-CR-W-S	*-SV40-luc-P2A-CR-W-S	*-minR-luc-F2A-CR-W-S
*-minC-FP-P2A-chR-W	*-RSV-FP-E2A-chR-W-S	*-SV40-luc-P2A-chR-W-S	*-minR-luc-F2A-chR-W-S
*-minC-FP-P2A-gRNA-W	*-RSV-FP-E2A-gRNA-W-S	*-SV40-luc-P2A-gRNA-W-S	*-minR-luc-F2A-gRNA-W-S
*-minC-FP-P2A-Nuc-W	*-RSV-FP-E2A-Nuc-W-S	*-SV40-luc-P2A-Nuc-W-S	*-minR-luc-F2A-Nuc-W-S
*-mut-FP-P2A-enz-W	*-minB-FP-F2A-enz-W	*-H68-luc-P2A-enz-W-S	*-CMV-luc-F2A-enz-W-S
*-mut-FP-P2A-TF-W	*-minB-FP-F2A-TF-W	*-H68-luc-P2A-TF-W-S	*-CMV-luc-F2A-TF-W-S
*-mut-FP-P2A-rec-W	*-minB-FP-F2A-rec-W	*-H68-luc-P2A-rec-W-S	*-CMV-luc-F2A-rec-W-S
*-mut-FP-P2A-CTP-W	*-minB-FP-F2A-CTP-W	*-H68-luc-P2A-CTP-W-S	*-CMV-luc-F2A-CTP-W-S
*-mut-FP-P2A-SM-W	*-minB-FP-F2A-SM-W	*-H68-luc-P2A-SM-W-S	*-CMV-luc-F2A-SM-W-S
*-mut-FP-P2A-NT-W	*-minB-FP-F2A-NT-W	*-H68-luc-P2A-NT-W-S	*-CMV-luc-F2A-NT-W-S
*-mut-FP-P2A-CR-W	*-minB-FP-F2A-CR-W	*-H68-luc-P2A-CR-W-S	*-CMV-luc-F2A-CR-W-S

FIG. 16 cont'd

*-mut-FP-P2A-chR-W	*-minB-FP-F2A-chR-W	*-H68-luc-P2A-chR-W-S	*-CMV-luc-F2A-chR-W-S
*-mut-FP-P2A-gRNA-W	*-minB-FP-F2A-gRNA-W	*-H68-luc-P2A-gRNA-W-S	*-CMV-luc-F2A-gRNA-W-S
*-mut-FP-P2A-Nuc-W	*-minB-FP-F2A-Nuc-W	*-H68-luc-P2A-Nuc-W-S	*-CMV-luc-F2A-Nuc-W-S
*-minR-FP-P2A-enz-W	*-minC-FP-F2A-enz-W	*-RSV-luc-P2A-enz-W-S	*-SV40-luc-F2A-enz-W-S
*-minR-FP-P2A-TF-W	*-minC-FP-F2A-TF-W	*-RSV-luc-P2A-TF-W-S	*-SV40-luc-F2A-TF-W-S
*-minR-FP-P2A-rec-W	*-minC-FP-F2A-rec-W	*-RSV-luc-P2A-rec-W-S	*-SV40-luc-F2A-rec-W-S
*-minR-FP-P2A-CTP-W	*-minC-FP-F2A-CTP-W	*-RSV-luc-P2A-CTP-W-S	*-SV40-luc-F2A-CTP-W-S
*-minR-FP-P2A-SM-W	*-minC-FP-F2A-SM-W	*-RSV-luc-P2A-SM-W-S	*-SV40-luc-F2A-SM-W-S
*-minR-FP-P2A-NT-W	*-minC-FP-F2A-NT-W	*-RSV-luc-P2A-NT-W-S	*-SV40-luc-F2A-NT-W-S
*-minR-FP-P2A-CR-W	*-minC-FP-F2A-CR-W	*-RSV-luc-P2A-CR-W-S	*-SV40-luc-F2A-CR-W-S
*-minR-FP-P2A-chR-W	*-minC-FP-F2A-chR-W	*-RSV-luc-P2A-chR-W-S	*-SV40-luc-F2A-chR-W-S
*-minR-FP-P2A-gRNA-W	*-minC-FP-F2A-gRNA-W	*-RSV-luc-P2A-gRNA-W-S	*-SV40-luc-F2A-gRNA-W-S
*-minR-FP-P2A-Nuc-W	*-minC-FP-F2A-Nuc-W	*-RSV-luc-P2A-Nuc-W-S	*-SV40-luc-F2A-Nuc-W-S
*-CMV-FP-P2A-enz-W	*-mut-FP-F2A-enz-W	*-minB-luc-T2A-enz-W	*-H68-luc-F2A-enz-W-S
*-CMV-FP-P2A-TF-W	*-mut-FP-F2A-TF-W	*-minB-luc-T2A-TF-W	*-H68-luc-F2A-TF-W-S
*-CMV-FP-P2A-rec-W	*-mut-FP-F2A-rec-W	*-minB-luc-T2A-rec-W	*-H68-luc-F2A-rec-W-S
*-CMV-FP-P2A-CTP-W	*-mut-FP-F2A-CTP-W	*-minB-luc-T2A-CTP-W	*-H68-luc-F2A-CTP-W-S
*-CMV-FP-P2A-SM-W	*-mut-FP-F2A-SM-W	*-minB-luc-T2A-SM-W	*-H68-luc-F2A-SM-W-S
*-CMV-FP-P2A-NT-W	*-mut-FP-F2A-NT-W	*-minB-luc-T2A-NT-W	*-H68-luc-F2A-NT-W-S
*-CMV-FP-P2A-CR-W	*-mut-FP-F2A-CR-W	*-minB-luc-T2A-CR-W	*-H68-luc-F2A-CR-W-S
*-CMV-FP-P2A-chR-W	*-mut-FP-F2A-chR-W	*-minB-luc-T2A-chR-W	*-H68-luc-F2A-chR-W-S
*-CMV-FP-P2A-gRNA-W	*-mut-FP-F2A-gRNA-W	*-minB-luc-T2A-gRNA-W	*-H68-luc-F2A-gRNA-W-S
*-CMV-FP-P2A-Nuc-W	*-mut-FP-F2A-Nuc-W	*-minB-luc-T2A-Nuc-W	*-H68-luc-F2A-Nuc-W-S
*-SV40-FP-P2A-enz-W	*-minR-FP-F2A-enz-W	*-minC-luc-T2A-enz-W	*-RSV-luc-F2A-enz-W-S
*-SV40-FP-P2A-TF-W	*-minR-FP-F2A-TF-W	*-minC-luc-T2A-TF-W	*-RSV-luc-F2A-TF-W-S
*-SV40-FP-P2A-rec-W	*-minR-FP-F2A-rec-W	*-minC-luc-T2A-rec-W	*-RSV-luc-F2A-rec-W-S
*-SV40-FP-P2A-CTP-W	*-minR-FP-F2A-CTP-W	*-minC-luc-T2A-CTP-W	*-RSV-luc-F2A-CTP-W-S
*-SV40-FP-P2A-SM-W	*-minR-FP-F2A-SM-W	*-minC-luc-T2A-SM-W	*-RSV-luc-F2A-SM-W-S
*-SV40-FP-P2A-NT-W	*-minR-FP-F2A-NT-W	*-minC-luc-T2A-NT-W	*-RSV-luc-F2A-NT-W-S
*-SV40-FP-P2A-CR-W	*-minR-FP-F2A-CR-W	*-minC-luc-T2A-CR-W	*-RSV-luc-F2A-CR-W-S
*-SV40-FP-P2A-chR-W	*-minR-FP-F2A-chR-W	*-minC-luc-T2A-chR-W	*-RSV-luc-F2A-chR-W-S
*-SV40-FP-P2A-gRNA-W	*-minR-FP-F2A-gRNA-W	*-minC-luc-T2A-gRNA-W	*-RSV-luc-F2A-gRNA-W-S
*-SV40-FP-P2A-Nuc-W	*-minR-FP-F2A-Nuc-W	*-minC-luc-T2A-Nuc-W	*-RSV-luc-F2A-Nuc-W-S
*-H68-FP-P2A-enz-W	*-CMV-FP-F2A-enz-W	*-mut-luc-T2A-enz-W	*-minB-luc-l2-enz-W

FIG. 16 cont'd

*-H68-FP-P2A-TF-W	*-CMV-FP-F2A-TF-W	*-mut-luc-T2A-TF-W	*-minB-luc-I2-TF-W
*-H68-FP-P2A-rec-W	*-CMV-FP-F2A-rec-W	*-mut-luc-T2A-rec-W	*-minB-luc-I2-rec-W
*-H68-FP-P2A-CTP-W	*-CMV-FP-F2A-CTP-W	*-mut-luc-T2A-CTP-W	*-minB-luc-I2-CTP-W
*-H68-FP-P2A-SM-W	*-CMV-FP-F2A-SM-W	*-mut-luc-T2A-SM-W	*-minB-luc-I2-SM-W
*-H68-FP-P2A-NT-W	*-CMV-FP-F2A-NT-W	*-mut-luc-T2A-NT-W	*-minB-luc-I2-NT-W
*-H68-FP-P2A-CR-W	*-CMV-FP-F2A-CR-W	*-mut-luc-T2A-CR-W	*-minB-luc-I2-CR-W
*-H68-FP-P2A-chR-W	*-CMV-FP-F2A-chR-W	*-mut-luc-T2A-chR-W	*-minB-luc-I2-chR-W
*-H68-FP-P2A-gRNA-W	*-CMV-FP-F2A-gRNA-W	*-mut-luc-T2A-gRNA-W	*-minB-luc-I2-gRNA-W
*-H68-FP-P2A-Nuc-W	*-CMV-FP-F2A-Nuc-W	*-mut-luc-T2A-Nuc-W	*-minB-luc-I2-Nuc-W
*-RSV-FP-P2A-enz-W	*-SV40-FP-F2A-enz-W	*-minR-luc-T2A-enz-W	*-minC-luc-I2-enz-W
*-RSV-FP-P2A-TF-W	*-SV40-FP-F2A-TF-W	*-minR-luc-T2A-TF-W	*-minC-luc-I2-TF-W
*-RSV-FP-P2A-rec-W	*-SV40-FP-F2A-rec-W	*-minR-luc-T2A-rec-W	*-minC-luc-I2-rec-W
*-RSV-FP-P2A-CTP-W	*-SV40-FP-F2A-CTP-W	*-minR-luc-T2A-CTP-W	*-minC-luc-I2-CTP-W
*-RSV-FP-P2A-SM-W	*-SV40-FP-F2A-SM-W	*-minR-luc-T2A-SM-W	*-minC-luc-I2-SM-W
*-RSV-FP-P2A-NT-W	*-SV40-FP-F2A-NT-W	*-minR-luc-T2A-NT-W	*-minC-luc-I2-NT-W
*-RSV-FP-P2A-CR-W	*-SV40-FP-F2A-CR-W	*-minR-luc-T2A-CR-W	*-minC-luc-I2-CR-W
*-RSV-FP-P2A-chR-W	*-SV40-FP-F2A-chR-W	*-minR-luc-T2A-chR-W	*-minC-luc-I2-chR-W
*-RSV-FP-P2A-gRNA-W	*-SV40-FP-F2A-gRNA-W	*-minR-luc-T2A-gRNA-W	*-minC-luc-I2-gRNA-W
*-RSV-FP-P2A-Nuc-W	*-SV40-FP-F2A-Nuc-W	*-minR-luc-T2A-Nuc-W	*-minC-luc-I2-Nuc-W
*-minB-FP-P2A-enz-bG	*-H68-FP-F2A-enz-W	*-CMV-luc-T2A-enz-W	*-mut-luc-I2-enz-W
*-minB-FP-P2A-TF-bG	*-H68-FP-F2A-TF-W	*-CMV-luc-T2A-TF-W	*-mut-luc-I2-TF-W
*-minB-FP-P2A-rec-bG	*-H68-FP-F2A-rec-W	*-CMV-luc-T2A-rec-W	*-mut-luc-I2-rec-W
*-minB-FP-P2A-CTP-bG	*-H68-FP-F2A-CTP-W	*-CMV-luc-T2A-CTP-W	*-mut-luc-I2-CTP-W
*-minB-FP-P2A-SM-bG	*-H68-FP-F2A-SM-W	*-CMV-luc-T2A-SM-W	*-mut-luc-I2-SM-W
*-minB-FP-P2A-NT-bG	*-H68-FP-F2A-NT-W	*-CMV-luc-T2A-NT-W	*-mut-luc-I2-NT-W
*-minB-FP-P2A-CR-bG	*-H68-FP-F2A-CR-W	*-CMV-luc-T2A-CR-W	*-mut-luc-I2-CR-W
*-minB-FP-P2A-chR-bG	*-H68-FP-F2A-chR-W	*-CMV-luc-T2A-chR-W	*-mut-luc-I2-chR-W
*-minB-FP-P2A-gRNA-bG	*-H68-FP-F2A-gRNA-W	*-CMV-luc-T2A-gRNA-W	*-mut-luc-I2-gRNA-W
*-minB-FP-P2A-Nuc-bG	*-H68-FP-F2A-Nuc-W	*-CMV-luc-T2A-Nuc-W	*-mut-luc-I2-Nuc-W
*-minC-FP-P2A-enz-bG	*-RSV-FP-F2A-enz-W	*-SV40-luc-T2A-enz-W	*-minR-luc-I2-enz-W
*-minC-FP-P2A-TF-bG	*-RSV-FP-F2A-TF-W	*-SV40-luc-T2A-TF-W	*-minR-luc-I2-TF-W
*-minC-FP-P2A-rec-bG	*-RSV-FP-F2A-rec-W	*-SV40-luc-T2A-rec-W	*-minR-luc-I2-rec-W
*-minC-FP-P2A-CTP-bG	*-RSV-FP-F2A-CTP-W	*-SV40-luc-T2A-CTP-W	*-minR-luc-I2-CTP-W
*-minC-FP-P2A-SM-bG	*-RSV-FP-F2A-SM-W	*-SV40-luc-T2A-SM-W	*-minR-luc-I2-SM-W

FIG. 16 cont'd

*-minC-FP-P2A-NT-bG	*-RSV-FP-F2A-NT-W	*-SV40-luc-T2A-NT-W	*-minR-luc-I2-NT-W
*-minC-FP-P2A-CR-bG	*-RSV-FP-F2A-CR-W	*-SV40-luc-T2A-CR-W	*-minR-luc-I2-CR-W
*-minC-FP-P2A-chR-bG	*-RSV-FP-F2A-chR-W	*-SV40-luc-T2A-chR-W	*-minR-luc-I2-chR-W
*-minC-FP-P2A-gRNA-bG	*-RSV-FP-F2A-gRNA-W	*-SV40-luc-T2A-gRNA-W	*-minR-luc-I2-gRNA-W
*-minC-FP-P2A-Nuc-bG	*-RSV-FP-F2A-Nuc-W	*-SV40-luc-T2A-Nuc-W	*-minR-luc-I2-Nuc-W
*-mut-FP-P2A-enz-bG	*-minB-FP-F2A-enz-bG	*-H68-luc-T2A-enz-W	*-CMV-luc-I2-enz-W
*-mut-FP-P2A-TF-bG	*-minB-FP-F2A-TF-bG	*-H68-luc-T2A-TF-W	*-CMV-luc-I2-TF-W
*-mut-FP-P2A-rec-bG	*-minB-FP-F2A-rec-bG	*-H68-luc-T2A-rec-W	*-CMV-luc-I2-rec-W
*-mut-FP-P2A-CTP-bG	*-minB-FP-F2A-CTP-bG	*-H68-luc-T2A-CTP-W	*-CMV-luc-I2-CTP-W
*-mut-FP-P2A-SM-bG	*-minB-FP-F2A-SM-bG	*-H68-luc-T2A-SM-W	*-CMV-luc-I2-SM-W
*-mut-FP-P2A-NT-bG	*-minB-FP-F2A-NT-bG	*-H68-luc-T2A-NT-W	*-CMV-luc-I2-NT-W
*-mut-FP-P2A-CR-bG	*-minB-FP-F2A-CR-bG	*-H68-luc-T2A-CR-W	*-CMV-luc-I2-CR-W
*-mut-FP-P2A-chR-bG	*-minB-FP-F2A-chR-bG	*-H68-luc-T2A-chR-W	*-CMV-luc-I2-chR-W
*-mut-FP-P2A-gRNA-bG	*-minB-FP-F2A-gRNA-bG	*-H68-luc-T2A-gRNA-W	*-CMV-luc-I2-gRNA-W
*-mut-FP-P2A-Nuc-bG	*-minB-FP-F2A-Nuc-bG	*-H68-luc-T2A-Nuc-W	*-CMV-luc-I2-Nuc-W
*-minR-FP-P2A-enz-bG	*-minC-FP-F2A-enz-bG	*-RSV-luc-T2A-enz-W	*-SV40-luc-I2-enz-W
*-minR-FP-P2A-TF-bG	*-minC-FP-F2A-TF-bG	*-RSV-luc-T2A-TF-W	*-SV40-luc-I2-TF-W
*-minR-FP-P2A-rec-bG	*-minC-FP-F2A-rec-bG	*-RSV-luc-T2A-rec-W	*-SV40-luc-I2-rec-W
*-minR-FP-P2A-CTP-bG	*-minC-FP-F2A-CTP-bG	*-RSV-luc-T2A-CTP-W	*-SV40-luc-I2-CTP-W
*-minR-FP-P2A-SM-bG	*-minC-FP-F2A-SM-bG	*-RSV-luc-T2A-SM-W	*-SV40-luc-I2-SM-W
*-minR-FP-P2A-NT-bG	*-minC-FP-F2A-NT-bG	*-RSV-luc-T2A-NT-W	*-SV40-luc-I2-NT-W
*-minR-FP-P2A-CR-bG	*-minC-FP-F2A-CR-bG	*-RSV-luc-T2A-CR-W	*-SV40-luc-I2-CR-W
*-minR-FP-P2A-chR-bG	*-minC-FP-F2A-chR-bG	*-RSV-luc-T2A-chR-W	*-SV40-luc-I2-chR-W
*-minR-FP-P2A-gRNA-bG	*-minC-FP-F2A-gRNA-bG	*-RSV-luc-T2A-gRNA-W	*-SV40-luc-I2-gRNA-W
*-minR-FP-P2A-Nuc-bG	*-minC-FP-F2A-Nuc-bG	*-RSV-luc-T2A-Nuc-W	*-SV40-luc-I2-Nuc-W
*-CMV-FP-P2A-enz-bG	*-mut-FP-F2A-enz-bG	*-minB-luc-T2A-enz-bG	*-H68-luc-I2-enz-W
*-CMV-FP-P2A-TF-bG	*-mut-FP-F2A-TF-bG	*-minB-luc-T2A-TF-bG	*-H68-luc-I2-TF-W
*-CMV-FP-P2A-rec-bG	*-mut-FP-F2A-rec-bG	*-minB-luc-T2A-rec-bG	*-H68-luc-I2-rec-W
*-CMV-FP-P2A-CTP-bG	*-mut-FP-F2A-CTP-bG	*-minB-luc-T2A-CTP-bG	*-H68-luc-I2-CTP-W
*-CMV-FP-P2A-SM-bG	*-mut-FP-F2A-SM-bG	*-minB-luc-T2A-SM-bG	*-H68-luc-I2-SM-W
*-CMV-FP-P2A-NT-bG	*-mut-FP-F2A-NT-bG	*-minB-luc-T2A-NT-bG	*-H68-luc-I2-NT-W
*-CMV-FP-P2A-CR-bG	*-mut-FP-F2A-CR-bG	*-minB-luc-T2A-CR-bG	*-H68-luc-I2-CR-W
*-CMV-FP-P2A-chR-bG	*-mut-FP-F2A-chR-bG	*-minB-luc-T2A-chR-bG	*-H68-luc-I2-chR-W
*-CMV-FP-P2A-gRNA-bG	*-mut-FP-F2A-gRNA-bG	*-minB-luc-T2A-gRNA-bG	*-H68-luc-I2-gRNA-W
*-CMV-FP-P2A-Nuc-bG	*-mut-FP-F2A-Nuc-bG	*-minB-luc-T2A-Nuc-bG	*-H68-luc-I2-gRNA-W

FIG. 16 cont'd

*-CMV-FP-P2A-Nuc-bG	*-mut-FP-F2A-Nuc-bG	*-minB-luc-T2A-Nuc-bG	*-H68-luc-I2-Nuc-W
*-SV40-FP-P2A-enz-bG	*-minR-FP-F2A-enz-bG	*-minC-luc-T2A-enz-bG	*-RSV-luc-I2-enz-W
*-SV40-FP-P2A-TF-bG	*-minR-FP-F2A-TF-bG	*-minC-luc-T2A-TF-bG	*-RSV-luc-I2-TF-W
*-SV40-FP-P2A-rec-bG	*-minR-FP-F2A-rec-bG	*-minC-luc-T2A-rec-bG	*-RSV-luc-I2-rec-W
*-SV40-FP-P2A-CTP-bG	*-minR-FP-F2A-CTP-bG	*-minC-luc-T2A-CTP-bG	*-RSV-luc-I2-CTP-W
*-SV40-FP-P2A-SM-bG	*-minR-FP-F2A-SM-bG	*-minC-luc-T2A-SM-bG	*-RSV-luc-I2-SM-W
*-SV40-FP-P2A-NT-bG	*-minR-FP-F2A-NT-bG	*-minC-luc-T2A-NT-bG	*-RSV-luc-I2-NT-W
*-SV40-FP-P2A-CR-bG	*-minR-FP-F2A-CR-bG	*-minC-luc-T2A-CR-bG	*-RSV-luc-I2-CR-W
*-SV40-FP-P2A-chR-bG	*-minR-FP-F2A-chR-bG	*-minC-luc-T2A-chR-bG	*-RSV-luc-I2-chR-W
*-SV40-FP-P2A-gRNA-bG	*-minR-FP-F2A-gRNA-bG	*-minC-luc-T2A-gRNA-bG	*-RSV-luc-I2-gRNA-W
*-SV40-FP-P2A-Nuc-bG	*-minR-FP-F2A-Nuc-bG	*-minC-luc-T2A-Nuc-bG	*-RSV-luc-I2-Nuc-W
*-H68-FP-P2A-enz-bG	*-CMV-FP-F2A-enz-bG	*-mut-luc-T2A-enz-bG	*-minB-luc-I2-enz-bG
*-H68-FP-P2A-TF-bG	*-CMV-FP-F2A-TF-bG	*-mut-luc-T2A-TF-bG	*-minB-luc-I2-TF-bG
*-H68-FP-P2A-rec-bG	*-CMV-FP-F2A-rec-bG	*-mut-luc-T2A-rec-bG	*-minB-luc-I2-rec-bG
*-H68-FP-P2A-CTP-bG	*-CMV-FP-F2A-CTP-bG	*-mut-luc-T2A-CTP-bG	*-minB-luc-I2-CTP-bG
*-H68-FP-P2A-SM-bG	*-CMV-FP-F2A-SM-bG	*-mut-luc-T2A-SM-bG	*-minB-luc-I2-SM-bG
*-H68-FP-P2A-NT-bG	*-CMV-FP-F2A-NT-bG	*-mut-luc-T2A-NT-bG	*-minB-luc-I2-NT-bG
*-H68-FP-P2A-CR-bG	*-CMV-FP-F2A-CR-bG	*-mut-luc-T2A-CR-bG	*-minB-luc-I2-CR-bG
*-H68-FP-P2A-chR-bG	*-CMV-FP-F2A-chR-bG	*-mut-luc-T2A-chR-bG	*-minB-luc-I2-chR-bG
*-H68-FP-P2A-gRNA-bG	*-CMV-FP-F2A-gRNA-bG	*-mut-luc-T2A-gRNA-bG	*-minB-luc-I2-gRNA-bG
*-H68-FP-P2A-Nuc-bG	*-CMV-FP-F2A-Nuc-bG	*-mut-luc-T2A-Nuc-bG	*-minB-luc-I2-Nuc-bG
*-RSV-FP-P2A-enz-bG	*-SV40-FP-F2A-enz-bG	*-minR-luc-T2A-enz-bG	*-minC-luc-I2-enz-bG
*-RSV-FP-P2A-TF-bG	*-SV40-FP-F2A-TF-bG	*-minR-luc-T2A-TF-bG	*-minC-luc-I2-TF-bG
*-RSV-FP-P2A-rec-bG	*-SV40-FP-F2A-rec-bG	*-minR-luc-T2A-rec-bG	*-minC-luc-I2-rec-bG
*-RSV-FP-P2A-CTP-bG	*-SV40-FP-F2A-CTP-bG	*-minR-luc-T2A-CTP-bG	*-minC-luc-I2-CTP-bG
*-RSV-FP-P2A-SM-bG	*-SV40-FP-F2A-SM-bG	*-minR-luc-T2A-SM-bG	*-minC-luc-I2-SM-bG
*-RSV-FP-P2A-NT-bG	*-SV40-FP-F2A-NT-bG	*-minR-luc-T2A-NT-bG	*-minC-luc-I2-NT-bG
*-RSV-FP-P2A-CR-bG	*-SV40-FP-F2A-CR-bG	*-minR-luc-T2A-CR-bG	*-minC-luc-I2-CR-bG
*-RSV-FP-P2A-chR-bG	*-SV40-FP-F2A-chR-bG	*-minR-luc-T2A-chR-bG	*-minC-luc-I2-chR-bG
*-RSV-FP-P2A-gRNA-bG	*-SV40-FP-F2A-gRNA-bG	*-minR-luc-T2A-gRNA-bG	*-minC-luc-I2-gRNA-bG
*-RSV-FP-P2A-Nuc-bG	*-SV40-FP-F2A-Nuc-bG	*-minR-luc-T2A-Nuc-bG	*-minC-luc-I2-Nuc-bG
*-minB-luc-I2-enz-bG	*-H68-FP-F2A-enz-bG	*-CMV-luc-T2A-enz-bG	*-mut-luc-I2-enz-bG
*-minB-luc-I2-TF-bG	*-H68-FP-F2A-TF-bG	*-CMV-luc-T2A-TF-bG	*-mut-luc-I2-TF-bG
*-minB-luc-I2-rec-bG	*-H68-FP-F2A-rec-bG	*-CMV-luc-T2A-rec-bG	*-mut-luc-I2-rec-bG

FIG. 16 cont'd

*-minB-FP-P2A-CTP-S	*-H68-FP-F2A-CTP-bG	*-CMV-luc-T2A-CTP-bG	*-mut-luc-I2-CTP-bG
*-minB-FP-P2A-SM-S	*-H68-FP-F2A-SM-bG	*-CMV-luc-T2A-SM-bG	*-mut-luc-I2-SM-bG
*-minB-FP-P2A-NT-S	*-H68-FP-F2A-NT-bG	*-CMV-luc-T2A-NT-bG	*-mut-luc-I2-NT-bG
*-minB-FP-P2A-CR-S	*-H68-FP-F2A-CR-bG	*-CMV-luc-T2A-CR-bG	*-mut-luc-I2-CR-bG
*-minB-FP-P2A-chR-S	*-H68-FP-F2A-chR-bG	*-CMV-luc-T2A-chR-bG	*-mut-luc-I2-chR-bG
*-minB-FP-P2A-gRNA-S	*-H68-FP-F2A-gRNA-bG	*-CMV-luc-T2A-gRNA-bG	*-mut-luc-I2-gRNA-bG
*-minB-FP-P2A-Nuc-S	*-H68-FP-F2A-Nuc-bG	*-CMV-luc-T2A-Nuc-bG	*-mut-luc-I2-Nuc-bG
*-minC-FP-P2A-enz-S	*-RSV-FP-F2A-enz-bG	*-SV40-luc-T2A-enz-bG	*-minR-luc-I2-enz-bG
*-minC-FP-P2A-TF-S	*-RSV-FP-F2A-TF-bG	*-SV40-luc-T2A-TF-bG	*-minR-luc-I2-TF-bG
*-minC-FP-P2A-rec-S	*-RSV-FP-F2A-rec-bG	*-SV40-luc-T2A-rec-bG	*-minR-luc-I2-rec-bG
*-minC-FP-P2A-CTP-S	*-RSV-FP-F2A-CTP-bG	*-SV40-luc-T2A-CTP-bG	*-minR-luc-I2-CTP-bG
*-minC-FP-P2A-SM-S	*-RSV-FP-F2A-SM-bG	*-SV40-luc-T2A-SM-bG	*-minR-luc-I2-SM-bG
*-minC-FP-P2A-NT-S	*-RSV-FP-F2A-NT-bG	*-SV40-luc-T2A-NT-bG	*-minR-luc-I2-NT-bG
*-minC-FP-P2A-CR-S	*-RSV-FP-F2A-CR-bG	*-SV40-luc-T2A-CR-bG	*-minR-luc-I2-CR-bG
*-minC-FP-P2A-chR-S	*-RSV-FP-F2A-chR-bG	*-SV40-luc-T2A-chR-bG	*-minR-luc-I2-chR-bG
*-minC-FP-P2A-gRNA-S	*-RSV-FP-F2A-gRNA-bG	*-SV40-luc-T2A-gRNA-bG	*-minR-luc-I2-gRNA-bG
*-minC-FP-P2A-Nuc-S	*-RSV-FP-F2A-Nuc-bG	*-SV40-luc-T2A-Nuc-bG	*-minR-luc-I2-Nuc-bG
*-mut-FP-P2A-enz-S	*-minB-FP-F2A-enz-S	*-H68-luc-T2A-enz-bG	*-CMV-luc-I2-enz-bG
*-mut-FP-P2A-TF-S	*-minB-FP-F2A-TF-S	*-H68-luc-T2A-TF-bG	*-CMV-luc-I2-TF-bG
*-mut-FP-P2A-rec-S	*-minB-FP-F2A-rec-S	*-H68-luc-T2A-rec-bG	*-CMV-luc-I2-rec-bG
*-mut-FP-P2A-CTP-S	*-minB-FP-F2A-CTP-S	*-H68-luc-T2A-CTP-bG	*-CMV-luc-I2-CTP-bG
*-mut-FP-P2A-SM-S	*-minB-FP-F2A-SM-S	*-H68-luc-T2A-SM-bG	*-CMV-luc-I2-SM-bG
*-mut-FP-P2A-NT-S	*-minB-FP-F2A-NT-S	*-H68-luc-T2A-NT-bG	*-CMV-luc-I2-NT-bG
*-mut-FP-P2A-CR-S	*-minB-FP-F2A-CR-S	*-H68-luc-T2A-CR-bG	*-CMV-luc-I2-CR-bG
*-mut-FP-P2A-chR-S	*-minB-FP-F2A-chR-S	*-H68-luc-T2A-chR-bG	*-CMV-luc-I2-chR-bG
*-mut-FP-P2A-gRNA-S	*-minB-FP-F2A-gRNA-S	*-H68-luc-T2A-gRNA-bG	*-CMV-luc-I2-gRNA-bG
*-mut-FP-P2A-Nuc-S	*-minB-FP-F2A-Nuc-S	*-H68-luc-T2A-Nuc-bG	*-CMV-luc-I2-Nuc-bG
*-minR-FP-P2A-enz-S	*-minC-FP-F2A-enz-S	*-RSV-luc-T2A-enz-bG	*-SV40-luc-I2-enz-bG
*-minR-FP-P2A-TF-S	*-minC-FP-F2A-TF-S	*-RSV-luc-T2A-TF-bG	*-SV40-luc-I2-TF-bG
*-minR-FP-P2A-rec-S	*-minC-FP-F2A-rec-S	*-RSV-luc-T2A-rec-bG	*-SV40-luc-I2-rec-bG
*-minR-FP-P2A-CTP-S	*-minC-FP-F2A-CTP-S	*-RSV-luc-T2A-CTP-bG	*-SV40-luc-I2-CTP-bG
*-minR-FP-P2A-SM-S	*-minC-FP-F2A-SM-S	*-RSV-luc-T2A-SM-bG	*-SV40-luc-I2-SM-bG
*-minR-FP-P2A-NT-S	*-minC-FP-F2A-NT-S	*-RSV-luc-T2A-NT-bG	*-SV40-luc-I2-NT-bG
*-minR-FP-P2A-CR-S	*-minC-FP-F2A-CR-S	*-RSV-luc-T2A-CR-bG	*-SV40-luc-I2-CR-bG

FIG. 16 cont'd

*-minR-FP-P2A-chR-S	*-minC-FP-F2A-chR-S	*-RSV-luc-T2A-chR-bG	*-SV40-luc-l2-chR-bG
*-minR-FP-P2A-gRNA-S	*-minC-FP-F2A-gRNA-S	*-RSV-luc-T2A-gRNA-bG	*-SV40-luc-l2-gRNA-bG
*-minR-FP-P2A-Nuc-S	*-minC-FP-F2A-Nuc-S	*-RSV-luc-T2A-Nuc-bG	*-SV40-luc-l2-Nuc-bG
*-CMV-FP-P2A-enz-S	*-mut-FP-F2A-enz-S	*-minB-luc-T2A-enz-S	*-H68-luc-l2-enz-bG
*-CMV-FP-P2A-TF-S	*-mut-FP-F2A-TF-S	*-minB-luc-T2A-TF-S	*-H68-luc-l2-TF-bG
*-CMV-FP-P2A-rec-S	*-mut-FP-F2A-rec-S	*-minB-luc-T2A-rec-S	*-H68-luc-l2-rec-bG
*-CMV-FP-P2A-CTP-S	*-mut-FP-F2A-CTP-S	*-minB-luc-T2A-CTP-S	*-H68-luc-l2-CTP-bG
*-CMV-FP-P2A-SM-S	*-mut-FP-F2A-SM-S	*-minB-luc-T2A-SM-S	*-H68-luc-l2-SM-bG
*-CMV-FP-P2A-NT-S	*-mut-FP-F2A-NT-S	*-minB-luc-T2A-NT-S	*-H68-luc-l2-NT-bG
*-CMV-FP-P2A-CR-S	*-mut-FP-F2A-CR-S	*-minB-luc-T2A-CR-S	*-H68-luc-l2-CR-bG
*-CMV-FP-P2A-chR-S	*-mut-FP-F2A-chR-S	*-minB-luc-T2A-chR-S	*-H68-luc-l2-chR-bG
*-CMV-FP-P2A-gRNA-S	*-mut-FP-F2A-gRNA-S	*-minB-luc-T2A-gRNA-S	*-H68-luc-l2-gRNA-bG
*-CMV-FP-P2A-Nuc-S	*-mut-FP-F2A-Nuc-S	*-minB-luc-T2A-Nuc-S	*-H68-luc-l2-Nuc-bG
*-SV40-FP-P2A-enz-S	*-minR-FP-F2A-enz-S	*-minC-luc-T2A-enz-S	*-RSV-luc-l2-enz-bG
*-SV40-FP-P2A-TF-S	*-minR-FP-F2A-TF-S	*-minC-luc-T2A-TF-S	*-RSV-luc-l2-TF-bG
*-SV40-FP-P2A-rec-S	*-minR-FP-F2A-rec-S	*-minC-luc-T2A-rec-S	*-RSV-luc-l2-rec-bG
*-SV40-FP-P2A-CTP-S	*-minR-FP-F2A-CTP-S	*-minC-luc-T2A-CTP-S	*-RSV-luc-l2-CTP-bG
*-SV40-FP-P2A-SM-S	*-minR-FP-F2A-SM-S	*-minC-luc-T2A-SM-S	*-RSV-luc-l2-SM-bG
*-SV40-FP-P2A-NT-S	*-minR-FP-F2A-NT-S	*-minC-luc-T2A-NT-S	*-RSV-luc-l2-NT-bG
*-SV40-FP-P2A-CR-S	*-minR-FP-F2A-CR-S	*-minC-luc-T2A-CR-S	*-RSV-luc-l2-CR-bG
*-SV40-FP-P2A-chR-S	*-minR-FP-F2A-chR-S	*-minC-luc-T2A-chR-S	*-RSV-luc-l2-chR-bG
*-SV40-FP-P2A-gRNA-S	*-minR-FP-F2A-gRNA-S	*-minC-luc-T2A-gRNA-S	*-RSV-luc-l2-gRNA-bG
*-SV40-FP-P2A-Nuc-S	*-minR-FP-F2A-Nuc-S	*-minC-luc-T2A-Nuc-S	*-RSV-luc-l2-Nuc-bG
*-H68-FP-P2A-enz-S	*-CMV-FP-F2A-enz-S	*-mut-luc-T2A-enz-S	*-minB-luc-l2-enz-S
*-H68-FP-P2A-TF-S	*-CMV-FP-F2A-TF-S	*-mut-luc-T2A-TF-S	*-minB-luc-l2-TF-S
*-H68-FP-P2A-rec-S	*-CMV-FP-F2A-rec-S	*-mut-luc-T2A-rec-S	*-minB-luc-l2-rec-S
*-H68-FP-P2A-CTP-S	*-CMV-FP-F2A-CTP-S	*-mut-luc-T2A-CTP-S	*-minB-luc-l2-CTP-S
*-H68-FP-P2A-SM-S	*-CMV-FP-F2A-SM-S	*-mut-luc-T2A-SM-S	*-minB-luc-l2-SM-S
*-H68-FP-P2A-NT-S	*-CMV-FP-F2A-NT-S	*-mut-luc-T2A-NT-S	*-minB-luc-l2-NT-S
*-H68-FP-P2A-CR-S	*-CMV-FP-F2A-CR-S	*-mut-luc-T2A-CR-S	*-minB-luc-l2-CR-S
*-H68-FP-P2A-chR-S	*-CMV-FP-F2A-chR-S	*-mut-luc-T2A-chR-S	*-minB-luc-l2-chR-S
*-H68-FP-P2A-gRNA-S	*-CMV-FP-F2A-gRNA-S	*-mut-luc-T2A-gRNA-S	*-minB-luc-l2-gRNA-S
*-H68-FP-P2A-Nuc-S	*-CMV-FP-F2A-Nuc-S	*-mut-luc-T2A-Nuc-S	*-minB-luc-l2-Nuc-S
*-RSV-FP-P2A-enz-S	*-SV40-FP-F2A-enz-S	*-minR-luc-T2A-enz-S	*-minC-luc-l2-enz-S

FIG. 16 cont'd

*-RSV-FP-P2A-TF-S	*-SV40-FP-F2A-TF-S	*-minR-luc-T2A-TF-S	*-minC-luc-I2-TF-S
*-RSV-FP-P2A-rec-S	*-SV40-FP-F2A-rec-S	*-minR-luc-T2A-rec-S	*-minC-luc-I2-rec-S
*-RSV-FP-P2A-CTP-S	*-SV40-FP-F2A-CTP-S	*-minR-luc-T2A-CTP-S	*-minC-luc-I2-CTP-S
*-RSV-FP-P2A-SM-S	*-SV40-FP-F2A-SM-S	*-minR-luc-T2A-SM-S	*-minC-luc-I2-SM-S
*-RSV-FP-P2A-NT-S	*-SV40-FP-F2A-NT-S	*-minR-luc-T2A-NT-S	*-minC-luc-I2-NT-S
*-RSV-FP-P2A-CR-S	*-SV40-FP-F2A-CR-S	*-minR-luc-T2A-CR-S	*-minC-luc-I2-CR-S
*-RSV-FP-P2A-chR-S	*-SV40-FP-F2A-chR-S	*-minR-luc-T2A-chR-S	*-minC-luc-I2-chR-S
*-RSV-FP-P2A-gRNA-S	*-SV40-FP-F2A-gRNA-S	*-minR-luc-T2A-gRNA-S	*-minC-luc-I2-gRNA-S
*-RSV-FP-P2A-Nuc-S	*-SV40-FP-F2A-Nuc-S	*-minR-luc-T2A-Nuc-S	*-minC-luc-I2-Nuc-S
*-minB-FP-P2A-enz-W-bG	*-H68-FP-F2A-enz-S	*-CMV-luc-T2A-enz-S	*-mut-luc-I2-enz-S
*-minB-FP-P2A-TF-W-bG	*-H68-FP-F2A-TF-S	*-CMV-luc-T2A-TF-S	*-mut-luc-I2-TF-S
*-minB-FP-P2A-rec-W-bG	*-H68-FP-F2A-rec-S	*-CMV-luc-T2A-rec-S	*-mut-luc-I2-rec-S
*-minB-FP-P2A-CTP-W-bG	*-H68-FP-F2A-CTP-S	*-CMV-luc-T2A-CTP-S	*-mut-luc-I2-CTP-S
*-minB-FP-P2A-SM-W-bG	*-H68-FP-F2A-SM-S	*-CMV-luc-T2A-SM-S	*-mut-luc-I2-SM-S
*-minB-FP-P2A-NT-W-bG	*-H68-FP-F2A-NT-S	*-CMV-luc-T2A-NT-S	*-mut-luc-I2-NT-S
*-minB-FP-P2A-CR-W-bG	*-H68-FP-F2A-CR-S	*-CMV-luc-T2A-CR-S	*-mut-luc-I2-CR-S
*-minB-FP-P2A-chR-W-bG	*-H68-FP-F2A-chR-S	*-CMV-luc-T2A-chR-S	*-mut-luc-I2-chR-S
*-minB-FP-P2A-gRNA-W-bG	*-H68-FP-F2A-gRNA-S	*-CMV-luc-T2A-gRNA-S	*-mut-luc-I2-gRNA-S
*-minB-FP-P2A-Nuc-W-bG	*-H68-FP-F2A-Nuc-S	*-CMV-luc-T2A-Nuc-S	*-mut-luc-I2-Nuc-S
*-minC-FP-P2A-enz-W-bG	*-RSV-FP-F2A-enz-S	*-SV40-luc-T2A-enz-S	*-minR-luc-I2-enz-S
*-minC-FP-P2A-TF-W-bG	*-RSV-FP-F2A-TF-S	*-SV40-luc-T2A-TF-S	*-minR-luc-I2-TF-S
*-minC-FP-P2A-rec-W-bG	*-RSV-FP-F2A-rec-S	*-SV40-luc-T2A-rec-S	*-minR-luc-I2-rec-S
*-minC-FP-P2A-CTP-W-bG	*-RSV-FP-F2A-CTP-S	*-SV40-luc-T2A-CTP-S	*-minR-luc-I2-CTP-S
*-minC-FP-P2A-SM-W-bG	*-RSV-FP-F2A-SM-S	*-SV40-luc-T2A-SM-S	*-minR-luc-I2-SM-S
*-minC-FP-P2A-NT-W-bG	*-RSV-FP-F2A-NT-S	*-SV40-luc-T2A-NT-S	*-minR-luc-I2-NT-S
*-minC-FP-P2A-CR-W-bG	*-RSV-FP-F2A-CR-S	*-SV40-luc-T2A-CR-S	*-minR-luc-I2-CR-S
*-minC-FP-P2A-chR-W-bG	*-RSV-FP-F2A-chR-S	*-SV40-luc-T2A-chR-S	*-minR-luc-I2-chR-S
*-minC-FP-P2A-gRNA-W-bG	*-RSV-FP-F2A-gRNA-S	*-SV40-luc-T2A-gRNA-S	*-minR-luc-I2-gRNA-S
*-minC-FP-P2A-Nuc-W-bG	*-RSV-FP-F2A-Nuc-S	*-SV40-luc-T2A-Nuc-S	*-minR-luc-I2-Nuc-S
*-mut-FP-P2A-enz-W-bG	*-minB-FP-F2A-enz-W-bG	*-H68-luc-T2A-enz-S	*-CMV-luc-I2-enz-S
*-mut-FP-P2A-TF-W-bG	*-minB-FP-F2A-TF-W-bG	*-H68-luc-T2A-TF-S	*-CMV-luc-I2-TF-S
*-mut-FP-P2A-rec-W-bG	*-minB-FP-F2A-rec-W-bG	*-H68-luc-T2A-rec-S	*-CMV-luc-I2-rec-S
*-mut-FP-P2A-CTP-W-bG	*-minB-FP-F2A-CTP-W-bG	*-H68-luc-T2A-CTP-S	*-CMV-luc-I2-CTP-S
*-mut-FP-P2A-SM-W-bG	*-minB-FP-F2A-SM-W-bG	*-H68-luc-T2A-SM-S	*-CMV-luc-I2-SM-S

FIG. 16 cont'd

*-mut-FP-P2A-NT-W-bG	*-minB-FP-F2A-NT-W-bG	*-H68-luc-T2A-NT-S	*-CMV-luc-I2-NT-S
*-mut-FP-P2A-CR-W-bG	*-minB-FP-F2A-CR-W-bG	*-H68-luc-T2A-CR-S	*-CMV-luc-I2-CR-S
*-mut-FP-P2A-chR-W-bG	*-minB-FP-F2A-chR-W-bG	*-H68-luc-T2A-chR-S	*-CMV-luc-I2-chR-S
*-mut-FP-P2A-gRNA-W-bG	*-minB-FP-F2A-gRNA-W-bG	*-H68-luc-T2A-gRNA-S	*-CMV-luc-I2-gRNA-S
*-mut-FP-P2A-Nuc-W-bG	*-minB-FP-F2A-Nuc-W-bG	*-H68-luc-T2A-Nuc-S	*-CMV-luc-I2-Nuc-S
*-minR-FP-P2A-enz-W-bG	*-minC-FP-F2A-enz-W-bG	*-RSV-luc-T2A-enz-S	*-SV40-luc-I2-enz-S
*-minR-FP-P2A-TF-W-bG	*-minC-FP-F2A-TF-W-bG	*-RSV-luc-T2A-TF-S	*-SV40-luc-I2-TF-S
*-minR-FP-P2A-rec-W-bG	*-minC-FP-F2A-rec-W-bG	*-RSV-luc-T2A-rec-S	*-SV40-luc-I2-rec-S
*-minR-FP-P2A-CTP-W-bG	*-minC-FP-F2A-CTP-W-bG	*-RSV-luc-T2A-CTP-S	*-SV40-luc-I2-CTP-S
*-minR-FP-P2A-SM-W-bG	*-minC-FP-F2A-SM-W-bG	*-RSV-luc-T2A-SM-S	*-SV40-luc-I2-SM-S
*-minR-FP-P2A-NT-W-bG	*-minC-FP-F2A-NT-W-bG	*-RSV-luc-T2A-NT-S	*-SV40-luc-I2-NT-S
*-minR-FP-P2A-CR-W-bG	*-minC-FP-F2A-CR-W-bG	*-RSV-luc-T2A-CR-S	*-SV40-luc-I2-CR-S
*-minR-FP-P2A-chR-W-bG	*-minC-FP-F2A-chR-W-bG	*-RSV-luc-T2A-chR-S	*-SV40-luc-I2-chR-S
*-minR-FP-P2A-gRNA-W-bG	*-minC-FP-F2A-gRNA-W-bG	*-RSV-luc-T2A-gRNA-S	*-SV40-luc-I2-gRNA-S
*-minR-FP-P2A-Nuc-W-bG	*-minC-FP-F2A-Nuc-W-bG	*-RSV-luc-T2A-Nuc-S	*-SV40-luc-I2-Nuc-S
*-CMV-FP-P2A-enz-W-bG	*-mut-FP-F2A-enz-W-bG	*-minB-luc-T2A-enz-W-bG	*-H68-luc-I2-enz-S
*-CMV-FP-P2A-TF-W-bG	*-mut-FP-F2A-TF-W-bG	*-minB-luc-T2A-TF-W-bG	*-H68-luc-I2-TF-S
*-CMV-FP-P2A-rec-W-bG	*-mut-FP-F2A-rec-W-bG	*-minB-luc-T2A-rec-W-bG	*-H68-luc-I2-rec-S
*-CMV-FP-P2A-CTP-W-bG	*-mut-FP-F2A-CTP-W-bG	*-minB-luc-T2A-CTP-W-bG	*-H68-luc-I2-CTP-S
*-CMV-FP-P2A-SM-W-bG	*-mut-FP-F2A-SM-W-bG	*-minB-luc-T2A-SM-W-bG	*-H68-luc-I2-SM-S
*-CMV-FP-P2A-NT-W-bG	*-mut-FP-F2A-NT-W-bG	*-minB-luc-T2A-NT-W-bG	*-H68-luc-I2-NT-S
*-CMV-FP-P2A-CR-W-bG	*-mut-FP-F2A-CR-W-bG	*-minB-luc-T2A-CR-W-bG	*-H68-luc-I2-CR-S
*-CMV-FP-P2A-chR-W-bG	*-mut-FP-F2A-chR-W-bG	*-minB-luc-T2A-chR-W-bG	*-H68-luc-I2-chR-S
*-CMV-FP-P2A-gRNA-W-bG	*-mut-FP-F2A-gRNA-W-bG	*-minB-luc-T2A-gRNA-W-bG	*-H68-luc-I2-gRNA-S
*-CMV-FP-P2A-Nuc-W-bG	*-mut-FP-F2A-Nuc-W-bG	*-minB-luc-T2A-Nuc-W-bG	*-H68-luc-I2-Nuc-S
*-SV40-FP-P2A-enz-W-bG	*-minR-FP-F2A-enz-W-bG	*-minC-luc-T2A-enz-W-bG	*-RSV-luc-I2-enz-S
*-SV40-FP-P2A-TF-W-bG	*-minR-FP-F2A-TF-W-bG	*-minC-luc-T2A-TF-W-bG	*-RSV-luc-I2-TF-S
*-SV40-FP-P2A-rec-W-bG	*-minR-FP-F2A-rec-W-bG	*-minC-luc-T2A-rec-W-bG	*-RSV-luc-I2-rec-S
*-SV40-FP-P2A-CTP-W-bG	*-minR-FP-F2A-CTP-W-bG	*-minC-luc-T2A-CTP-W-bG	*-RSV-luc-I2-CTP-S
*-SV40-FP-P2A-SM-W-bG	*-minR-FP-F2A-SM-W-bG	*-minC-luc-T2A-SM-W-bG	*-RSV-luc-I2-SM-S
*-SV40-FP-P2A-NT-W-bG	*-minR-FP-F2A-NT-W-bG	*-minC-luc-T2A-NT-W-bG	*-RSV-luc-I2-NT-S
*-SV40-FP-P2A-CR-W-bG	*-minR-FP-F2A-CR-W-bG	*-minC-luc-T2A-CR-W-bG	*-RSV-luc-I2-CR-S
*-SV40-FP-P2A-chR-W-bG	*-minR-FP-F2A-chR-W-bG	*-minC-luc-T2A-chR-W-bG	*-RSV-luc-I2-chR-S
*-SV40-FP-P2A-gRNA-W-bG	*-minR-FP-F2A-gRNA-W-bG	*-minC-luc-T2A-gRNA-W-bG	*-RSV-luc-I2-gRNA-S

FIG. 16 cont'd

*-SV40-FP-P2A-Nuc-W-bG	*-minR-FP-F2A-Nuc-W-bG	*-minC-luc-T2A-Nuc-W-bG	*-RSV-luc-I2-Nuc-S
*-H68-FP-P2A-enz-W-bG	*-CMV-FP-F2A-enz-W-bG	*-mut-luc-T2A-enz-W-bG	*-minB-luc-I2-enz-W-bG
*-H68-FP-P2A-TF-W-bG	*-CMV-FP-F2A-TF-W-bG	*-mut-luc-T2A-TF-W-bG	*-minB-luc-I2-TF-W-bG
*-H68-FP-P2A-rec-W-bG	*-CMV-FP-F2A-rec-W-bG	*-mut-luc-T2A-rec-W-bG	*-minB-luc-I2-rec-W-bG
*-H68-FP-P2A-CTP-W-bG	*-CMV-FP-F2A-CTP-W-bG	*-mut-luc-T2A-CTP-W-bG	*-minB-luc-I2-CTP-W-bG
*-H68-FP-P2A-SM-W-bG	*-CMV-FP-F2A-SM-W-bG	*-mut-luc-T2A-SM-W-bG	*-minB-luc-I2-SM-W-bG
*-H68-FP-P2A-NT-W-bG	*-CMV-FP-F2A-NT-W-bG	*-mut-luc-T2A-NT-W-bG	*-minB-luc-I2-NT-W-bG
*-H68-FP-P2A-CR-W-bG	*-CMV-FP-F2A-CR-W-bG	*-mut-luc-T2A-CR-W-bG	*-minB-luc-I2-CR-W-bG
*-H68-FP-P2A-chR-W-bG	*-CMV-FP-F2A-chR-W-bG	*-mut-luc-T2A-chR-W-bG	*-minB-luc-I2-chR-W-bG
*-H68-FP-P2A-gRNA-W-bG	*-CMV-FP-F2A-gRNA-W-bG	*-mut-luc-T2A-gRNA-W-bG	*-minB-luc-I2-gRNA-W-bG
*-H68-FP-P2A-Nuc-W-bG	*-CMV-FP-F2A-Nuc-W-bG	*-mut-luc-T2A-Nuc-W-bG	*-minB-luc-I2-Nuc-W-bG
*-RSV-FP-P2A-enz-W-bG	*-SV40-FP-F2A-enz-W-bG	*-minR-luc-T2A-enz-W-bG	*-minC-luc-I2-enz-W-bG
*-RSV-FP-P2A-TF-W-bG	*-SV40-FP-F2A-TF-W-bG	*-minR-luc-T2A-TF-W-bG	*-minC-luc-I2-TF-W-bG
*-RSV-FP-P2A-rec-W-bG	*-SV40-FP-F2A-rec-W-bG	*-minR-luc-T2A-rec-W-bG	*-minC-luc-I2-rec-W-bG
*-RSV-FP-P2A-CTP-W-bG	*-SV40-FP-F2A-CTP-W-bG	*-minR-luc-T2A-CTP-W-bG	*-minC-luc-I2-CTP-W-bG
*-RSV-FP-P2A-SM-W-bG	*-SV40-FP-F2A-SM-W-bG	*-minR-luc-T2A-SM-W-bG	*-minC-luc-I2-SM-W-bG
*-RSV-FP-P2A-NT-W-bG	*-SV40-FP-F2A-NT-W-bG	*-minR-luc-T2A-NT-W-bG	*-minC-luc-I2-NT-W-bG
*-RSV-FP-P2A-CR-W-bG	*-SV40-FP-F2A-CR-W-bG	*-minR-luc-T2A-CR-W-bG	*-minC-luc-I2-CR-W-bG
*-RSV-FP-P2A-chR-W-bG	*-SV40-FP-F2A-chR-W-bG	*-minR-luc-T2A-chR-W-bG	*-minC-luc-I2-chR-W-bG
*-RSV-FP-P2A-gRNA-W-bG	*-SV40-FP-F2A-gRNA-W-bG	*-minR-luc-T2A-gRNA-W-bG	*-minC-luc-I2-gRNA-W-bG
*-RSV-FP-P2A-Nuc-W-bG	*-SV40-FP-F2A-Nuc-W-bG	*-minR-luc-T2A-Nuc-W-bG	*-minC-luc-I2-Nuc-W-bG
*-minB-FP-P2A-enz-W-S	*-H68-FP-F2A-enz-W-bG	*-CMV-luc-T2A-enz-W-bG	*-mut-luc-I2-enz-W-bG
*-minB-FP-P2A-TF-W-S	*-H68-FP-F2A-TF-W-bG	*-CMV-luc-T2A-TF-W-bG	*-mut-luc-I2-TF-W-bG
*-minB-FP-P2A-rec-W-S	*-H68-FP-F2A-rec-W-bG	*-CMV-luc-T2A-rec-W-bG	*-mut-luc-I2-rec-W-bG
*-minB-FP-P2A-CTP-W-S	*-H68-FP-F2A-CTP-W-bG	*-CMV-luc-T2A-CTP-W-bG	*-mut-luc-I2-CTP-W-bG
*-minB-FP-P2A-SM-W-S	*-H68-FP-F2A-SM-W-bG	*-CMV-luc-T2A-SM-W-bG	*-mut-luc-I2-SM-W-bG
*-minB-FP-P2A-NT-W-S	*-H68-FP-F2A-NT-W-bG	*-CMV-luc-T2A-NT-W-bG	*-mut-luc-I2-NT-W-bG
*-minB-FP-P2A-CR-W-S	*-H68-FP-F2A-CR-W-bG	*-CMV-luc-T2A-CR-W-bG	*-mut-luc-I2-CR-W-bG
*-minB-FP-P2A-chR-W-S	*-H68-FP-F2A-chR-W-bG	*-CMV-luc-T2A-chR-W-bG	*-mut-luc-I2-chR-W-bG
*-minB-FP-P2A-gRNA-W-S	*-H68-FP-F2A-gRNA-W-bG	*-CMV-luc-T2A-gRNA-W-bG	*-mut-luc-I2-gRNA-W-bG
*-minB-FP-P2A-Nuc-W-S	*-H68-FP-F2A-Nuc-W-bG	*-CMV-luc-T2A-Nuc-W-bG	*-mut-luc-I2-Nuc-W-bG
*-minC-FP-P2A-enz-W-S	*-RSV-FP-F2A-enz-W-bG	*-SV40-luc-T2A-enz-W-bG	*-minR-luc-I2-enz-W-bG
*-minC-FP-P2A-TF-W-S	*-RSV-FP-F2A-TF-W-bG	*-SV40-luc-T2A-TF-W-bG	*-minR-luc-I2-TF-W-bG
*-minC-FP-P2A-rec-W-S	*-RSV-FP-F2A-rec-W-bG	*-SV40-luc-T2A-rec-W-bG	*-minR-luc-I2-rec-W-bG

FIG. 16 cont'd

*-minC-FP-P2A-CTP-W-S	*-RSV-FP-F2A-CTP-W-bG	*-SV40-luc-T2A-CTP-W-bG	*-minR-luc-I2-CTP-W-bG
*-minC-FP-P2A-SM-W-S	*-RSV-FP-F2A-SM-W-bG	*-SV40-luc-T2A-SM-W-bG	*-minR-luc-I2-SM-W-bG
*-minC-FP-P2A-NT-W-S	*-RSV-FP-F2A-NT-W-bG	*-SV40-luc-T2A-NT-W-bG	*-minR-luc-I2-NT-W-bG
*-minC-FP-P2A-CR-W-S	*-RSV-FP-F2A-CR-W-bG	*-SV40-luc-T2A-CR-W-bG	*-minR-luc-I2-CR-W-bG
*-minC-FP-P2A-chR-W-S	*-RSV-FP-F2A-chR-W-bG	*-SV40-luc-T2A-chR-W-bG	*-minR-luc-I2-chR-W-bG
*-minC-FP-P2A-gRNA-W-S	*-RSV-FP-F2A-gRNA-W-bG	*-SV40-luc-T2A-gRNA-W-bG	*-minR-luc-I2-gRNA-W-bG
*-minC-FP-P2A-Nuc-W-S	*-RSV-FP-F2A-Nuc-W-bG	*-SV40-luc-T2A-Nuc-W-bG	*-minR-luc-I2-Nuc-W-bG
*-mut-FP-P2A-enz-W-S	*-minB-FP-F2A-enz-W-S	*-H68-luc-T2A-enz-W-bG	*-CMV-luc-I2-enz-W-bG
*-mut-FP-P2A-TF-W-S	*-minB-FP-F2A-TF-W-S	*-H68-luc-T2A-TF-W-bG	*-CMV-luc-I2-TF-W-bG
*-mut-FP-P2A-rec-W-S	*-minB-FP-F2A-rec-W-S	*-H68-luc-T2A-rec-W-bG	*-CMV-luc-I2-rec-W-bG
*-mut-FP-P2A-CTP-W-S	*-minB-FP-F2A-CTP-W-S	*-H68-luc-T2A-CTP-W-bG	*-CMV-luc-I2-CTP-W-bG
*-mut-FP-P2A-SM-W-S	*-minB-FP-F2A-SM-W-S	*-H68-luc-T2A-SM-W-bG	*-CMV-luc-I2-SM-W-bG
*-mut-FP-P2A-NT-W-S	*-minB-FP-F2A-NT-W-S	*-H68-luc-T2A-NT-W-bG	*-CMV-luc-I2-NT-W-bG
*-mut-FP-P2A-CR-W-S	*-minB-FP-F2A-CR-W-S	*-H68-luc-T2A-CR-W-bG	*-CMV-luc-I2-CR-W-bG
*-mut-FP-P2A-chR-W-S	*-minB-FP-F2A-chR-W-S	*-H68-luc-T2A-chR-W-bG	*-CMV-luc-I2-chR-W-bG
*-mut-FP-P2A-gRNA-W-S	*-minB-FP-F2A-gRNA-W-S	*-H68-luc-T2A-gRNA-W-bG	*-CMV-luc-I2-gRNA-W-bG
*-mut-FP-P2A-Nuc-W-S	*-minB-FP-F2A-Nuc-W-S	*-H68-luc-T2A-Nuc-W-bG	*-CMV-luc-I2-Nuc-W-bG
*-minR-FP-P2A-enz-W-S	*-minC-FP-F2A-enz-W-S	*-RSV-luc-T2A-enz-W-bG	*-SV40-luc-I2-enz-W-bG
*-minR-FP-P2A-TF-W-S	*-minC-FP-F2A-TF-W-S	*-RSV-luc-T2A-TF-W-bG	*-SV40-luc-I2-TF-W-bG
*-minR-FP-P2A-rec-W-S	*-minC-FP-F2A-rec-W-S	*-RSV-luc-T2A-rec-W-bG	*-SV40-luc-I2-rec-W-bG
*-minR-FP-P2A-CTP-W-S	*-minC-FP-F2A-CTP-W-S	*-RSV-luc-T2A-CTP-W-bG	*-SV40-luc-I2-CTP-W-bG
*-minR-FP-P2A-SM-W-S	*-minC-FP-F2A-SM-W-S	*-RSV-luc-T2A-SM-W-bG	*-SV40-luc-I2-SM-W-bG
*-minR-FP-P2A-NT-W-S	*-minC-FP-F2A-NT-W-S	*-RSV-luc-T2A-NT-W-bG	*-SV40-luc-I2-NT-W-bG
*-minR-FP-P2A-CR-W-S	*-minC-FP-F2A-CR-W-S	*-RSV-luc-T2A-CR-W-bG	*-SV40-luc-I2-CR-W-bG
*-minR-FP-P2A-chR-W-S	*-minC-FP-F2A-chR-W-S	*-RSV-luc-T2A-chR-W-bG	*-SV40-luc-I2-chR-W-bG
*-minR-FP-P2A-gRNA-W-S	*-minC-FP-F2A-gRNA-W-S	*-RSV-luc-T2A-gRNA-W-bG	*-SV40-luc-I2-gRNA-W-bG
*-minR-FP-P2A-Nuc-W-S	*-minC-FP-F2A-Nuc-W-S	*-RSV-luc-T2A-Nuc-W-bG	*-SV40-luc-I2-Nuc-W-bG
*-CMV-FP-P2A-enz-W-S	*-mut-FP-F2A-enz-W-S	*-minB-luc-T2A-enz-W-S	*-H68-luc-I2-enz-W-bG
*-CMV-FP-P2A-TF-W-S	*-mut-FP-F2A-TF-W-S	*-minB-luc-T2A-TF-W-S	*-H68-luc-I2-TF-W-bG
*-CMV-FP-P2A-rec-W-S	*-mut-FP-F2A-rec-W-S	*-minB-luc-T2A-rec-W-S	*-H68-luc-I2-rec-W-bG
*-CMV-FP-P2A-CTP-W-S	*-mut-FP-F2A-CTP-W-S	*-minB-luc-T2A-CTP-W-S	*-H68-luc-I2-CTP-W-bG
*-CMV-FP-P2A-SM-W-S	*-mut-FP-F2A-SM-W-S	*-minB-luc-T2A-SM-W-S	*-H68-luc-I2-SM-W-bG
*-CMV-FP-P2A-NT-W-S	*-mut-FP-F2A-NT-W-S	*-minB-luc-T2A-NT-W-S	*-H68-luc-I2-NT-W-bG
*-CMV-FP-P2A-CR-W-S	*-mut-FP-F2A-CR-W-S	*-minB-luc-T2A-CR-W-S	*-H68-luc-I2-CR-W-bG

FIG. 16 cont'd

*-CMV-FP-P2A-chR-W-S	*-mut-FP-F2A-chR-W-S	*-minB-luc-T2A-chR-W-S	*-H68-luc-l2-chR-W-bG
*-CMV-FP-P2A-gRNA-W-S	*-mut-FP-F2A-gRNA-W-S	*-minB-luc-T2A-gRNA-W-S	*-H68-luc-l2-gRNA-W-bG
*-CMV-FP-P2A-Nuc-W-S	*-mut-FP-F2A-Nuc-W-S	*-minB-luc-T2A-Nuc-W-S	*-H68-luc-l2-Nuc-W-bG
*-SV40-FP-P2A-enz-W-S	*-minR-FP-F2A-enz-W-S	*-minC-luc-T2A-enz-W-S	*-RSV-luc-l2-enz-W-bG
*-SV40-FP-P2A-TF-W-S	*-minR-FP-F2A-TF-W-S	*-minC-luc-T2A-TF-W-S	*-RSV-luc-l2-TF-W-bG
*-SV40-FP-P2A-rec-W-S	*-minR-FP-F2A-rec-W-S	*-minC-luc-T2A-rec-W-S	*-RSV-luc-l2-rec-W-bG
*-SV40-FP-P2A-CTP-W-S	*-minR-FP-F2A-CTP-W-S	*-minC-luc-T2A-CTP-W-S	*-RSV-luc-l2-CTP-W-bG
*-SV40-FP-P2A-SM-W-S	*-minR-FP-F2A-SM-W-S	*-minC-luc-T2A-SM-W-S	*-RSV-luc-l2-SM-W-bG
*-SV40-FP-P2A-NT-W-S	*-minR-FP-F2A-NT-W-S	*-minC-luc-T2A-NT-W-S	*-RSV-luc-l2-NT-W-bG
*-SV40-FP-P2A-CR-W-S	*-minR-FP-F2A-CR-W-S	*-minC-luc-T2A-CR-W-S	*-RSV-luc-l2-CR-W-bG
*-SV40-FP-P2A-chR-W-S	*-minR-FP-F2A-chR-W-S	*-minC-luc-T2A-chR-W-S	*-RSV-luc-l2-chR-W-bG
*-SV40-FP-P2A-gRNA-W-S	*-minR-FP-F2A-gRNA-W-S	*-minC-luc-T2A-gRNA-W-S	*-RSV-luc-l2-gRNA-W-bG
*-SV40-FP-P2A-Nuc-W-S	*-minR-FP-F2A-Nuc-W-S	*-minC-luc-T2A-Nuc-W-S	*-RSV-luc-l2-Nuc-W-bG
*-H68-FP-P2A-enz-W-S	*-CMV-FP-F2A-enz-W-S	*-mut-luc-T2A-enz-W-S	*-minB-luc-l2-enz-W-S
*-H68-FP-P2A-TF-W-S	*-CMV-FP-F2A-TF-W-S	*-mut-luc-T2A-TF-W-S	*-minB-luc-l2-TF-W-S
*-H68-FP-P2A-rec-W-S	*-CMV-FP-F2A-rec-W-S	*-mut-luc-T2A-rec-W-S	*-minB-luc-l2-rec-W-S
*-H68-FP-P2A-CTP-W-S	*-CMV-FP-F2A-CTP-W-S	*-mut-luc-T2A-CTP-W-S	*-minB-luc-l2-CTP-W-S
*-H68-FP-P2A-SM-W-S	*-CMV-FP-F2A-SM-W-S	*-mut-luc-T2A-SM-W-S	*-minB-luc-l2-SM-W-S
*-H68-FP-P2A-NT-W-S	*-CMV-FP-F2A-NT-W-S	*-mut-luc-T2A-NT-W-S	*-minB-luc-l2-NT-W-S
*-H68-FP-P2A-CR-W-S	*-CMV-FP-F2A-CR-W-S	*-mut-luc-T2A-CR-W-S	*-minB-luc-l2-CR-W-S
*-H68-FP-P2A-chR-W-S	*-CMV-FP-F2A-chR-W-S	*-mut-luc-T2A-chR-W-S	*-minB-luc-l2-chR-W-S
*-H68-FP-P2A-gRNA-W-S	*-CMV-FP-F2A-gRNA-W-S	*-mut-luc-T2A-gRNA-W-S	*-minB-luc-l2-gRNA-W-S
*-H68-FP-P2A-Nuc-W-S	*-CMV-FP-F2A-Nuc-W-S	*-mut-luc-T2A-Nuc-W-S	*-minB-luc-l2-Nuc-W-S
*-RSV-FP-P2A-enz-W-S	*-SV40-FP-F2A-enz-W-S	*-minR-luc-T2A-enz-W-S	*-minC-luc-l2-enz-W-S
*-RSV-FP-P2A-TF-W-S	*-SV40-FP-F2A-TF-W-S	*-minR-luc-T2A-TF-W-S	*-minC-luc-l2-TF-W-S
*-RSV-FP-P2A-rec-W-S	*-SV40-FP-F2A-rec-W-S	*-minR-luc-T2A-rec-W-S	*-minC-luc-l2-rec-W-S
*-RSV-FP-P2A-CTP-W-S	*-SV40-FP-F2A-CTP-W-S	*-minR-luc-T2A-CTP-W-S	*-minC-luc-l2-CTP-W-S
*-RSV-FP-P2A-SM-W-S	*-SV40-FP-F2A-SM-W-S	*-minR-luc-T2A-SM-W-S	*-minC-luc-l2-SM-W-S
*-RSV-FP-P2A-NT-W-S	*-SV40-FP-F2A-NT-W-S	*-minR-luc-T2A-NT-W-S	*-minC-luc-l2-NT-W-S
*-RSV-FP-P2A-CR-W-S	*-SV40-FP-F2A-CR-W-S	*-minR-luc-T2A-CR-W-S	*-minC-luc-l2-CR-W-S
*-RSV-FP-P2A-chR-W-S	*-SV40-FP-F2A-chR-W-S	*-minR-luc-T2A-chR-W-S	*-minC-luc-l2-chR-W-S
*-RSV-FP-P2A-gRNA-W-S	*-SV40-FP-F2A-gRNA-W-S	*-minR-luc-T2A-gRNA-W-S	*-minC-luc-l2-gRNA-W-S
*-RSV-FP-P2A-Nuc-W-S	*-SV40-FP-F2A-Nuc-W-S	*-minR-luc-T2A-Nuc-W-S	*-minC-luc-l2-Nuc-W-S
*-minB-FP-T2A-enz-W	*-H68-FP-F2A-enz-W-S	*-CMV-luc-T2A-enz-W-S	*-mut-luc-l2-enz-W-S

FIG. 16 cont'd

*-minB-FP-T2A-TF-W	*-H68-FP-F2A-TF-W-S	*-CMV-luc-T2A-TF-W-S	*-mut-luc-I2-TF-W-S
*-minB-FP-T2A-rec-W	*-H68-FP-F2A-rec-W-S	*-CMV-luc-T2A-rec-W-S	*-mut-luc-I2-rec-W-S
*-minB-FP-T2A-CTP-W	*-H68-FP-F2A-CTP-W-S	*-CMV-luc-T2A-CTP-W-S	*-mut-luc-I2-CTP-W-S
*-minB-FP-T2A-SM-W	*-H68-FP-F2A-SM-W-S	*-CMV-luc-T2A-SM-W-S	*-mut-luc-I2-SM-W-S
*-minB-FP-T2A-NT-W	*-H68-FP-F2A-NT-W-S	*-CMV-luc-T2A-NT-W-S	*-mut-luc-I2-NT-W-S
*-minB-FP-T2A-CR-W	*-H68-FP-F2A-CR-W-S	*-CMV-luc-T2A-CR-W-S	*-mut-luc-I2-CR-W-S
*-minB-FP-T2A-chR-W	*-H68-FP-F2A-chR-W-S	*-CMV-luc-T2A-chR-W-S	*-mut-luc-I2-chR-W-S
*-minB-FP-T2A-gRNA-W	*-H68-FP-F2A-gRNA-W-S	*-CMV-luc-T2A-gRNA-W-S	*-mut-luc-I2-gRNA-W-S
*-minB-FP-T2A-Nuc-W	*-H68-FP-F2A-Nuc-W-S	*-CMV-luc-T2A-Nuc-W-S	*-mut-luc-I2-Nuc-W-S
*-minC-FP-T2A-enz-W	*-RSV-FP-F2A-enz-W-S	*-SV40-luc-T2A-enz-W-S	*-minR-luc-I2-enz-W-S
*-minC-FP-T2A-TF-W	*-RSV-FP-F2A-TF-W-S	*-SV40-luc-T2A-TF-W-S	*-minR-luc-I2-TF-W-S
*-minC-FP-T2A-rec-W	*-RSV-FP-F2A-rec-W-S	*-SV40-luc-T2A-rec-W-S	*-minR-luc-I2-rec-W-S
*-minC-FP-T2A-CTP-W	*-RSV-FP-F2A-CTP-W-S	*-SV40-luc-T2A-CTP-W-S	*-minR-luc-I2-CTP-W-S
*-minC-FP-T2A-SM-W	*-RSV-FP-F2A-SM-W-S	*-SV40-luc-T2A-SM-W-S	*-minR-luc-I2-SM-W-S
*-minC-FP-T2A-NT-W	*-RSV-FP-F2A-NT-W-S	*-SV40-luc-T2A-NT-W-S	*-minR-luc-I2-NT-W-S
*-minC-FP-T2A-CR-W	*-RSV-FP-F2A-CR-W-S	*-SV40-luc-T2A-CR-W-S	*-minR-luc-I2-CR-W-S
*-minC-FP-T2A-chR-W	*-RSV-FP-F2A-chR-W-S	*-SV40-luc-T2A-chR-W-S	*-minR-luc-I2-chR-W-S
*-minC-FP-T2A-gRNA-W	*-RSV-FP-F2A-gRNA-W-S	*-SV40-luc-T2A-gRNA-W-S	*-minR-luc-I2-gRNA-W-S
*-minC-FP-T2A-Nuc-W	*-RSV-FP-F2A-Nuc-W-S	*-SV40-luc-T2A-Nuc-W-S	*-minR-luc-I2-Nuc-W-S
*-mut-FP-T2A-enz-W	*-minB-FP-I2-enz-W	*-H68-luc-T2A-enz-W-S	*-CMV-luc-I2-enz-W-S
*-mut-FP-T2A-TF-W	*-minB-FP-I2-TF-W	*-H68-luc-T2A-TF-W-S	*-CMV-luc-I2-TF-W-S
*-mut-FP-T2A-rec-W	*-minB-FP-I2-rec-W	*-H68-luc-T2A-rec-W-S	*-CMV-luc-I2-rec-W-S
*-mut-FP-T2A-CTP-W	*-minB-FP-I2-CTP-W	*-H68-luc-T2A-CTP-W-S	*-CMV-luc-I2-CTP-W-S
*-mut-FP-T2A-SM-W	*-minB-FP-I2-SM-W	*-H68-luc-T2A-SM-W-S	*-CMV-luc-I2-SM-W-S
*-mut-FP-T2A-NT-W	*-minB-FP-I2-NT-W	*-H68-luc-T2A-NT-W-S	*-CMV-luc-I2-NT-W-S
*-mut-FP-T2A-CR-W	*-minB-FP-I2-CR-W	*-H68-luc-T2A-CR-W-S	*-CMV-luc-I2-CR-W-S
*-mut-FP-T2A-chR-W	*-minB-FP-I2-chR-W	*-H68-luc-T2A-chR-W-S	*-CMV-luc-I2-chR-W-S
*-mut-FP-T2A-gRNA-W	*-minB-FP-I2-gRNA-W	*-H68-luc-T2A-gRNA-W-S	*-CMV-luc-I2-gRNA-W-S
*-mut-FP-T2A-Nuc-W	*-minB-FP-I2-Nuc-W	*-H68-luc-T2A-Nuc-W-S	*-CMV-luc-I2-Nuc-W-S
*-minR-FP-T2A-enz-W	*-minC-FP-I2-enz-W	*-RSV-luc-T2A-enz-W-S	*-SV40-luc-I2-enz-W-S
*-minR-FP-T2A-TF-W	*-minC-FP-I2-TF-W	*-RSV-luc-T2A-TF-W-S	*-SV40-luc-I2-TF-W-S
*-minR-FP-T2A-rec-W	*-minC-FP-I2-rec-W	*-RSV-luc-T2A-rec-W-S	*-SV40-luc-I2-rec-W-S
*-minR-FP-T2A-CTP-W	*-minC-FP-I2-CTP-W	*-RSV-luc-T2A-CTP-W-S	*-SV40-luc-I2-CTP-W-S
*-minR-FP-T2A-SM-W	*-minC-FP-I2-SM-W	*-RSV-luc-T2A-SM-W-S	*-SV40-luc-I2-SM-W-S

FIG. 16 cont'd

*-minR-FP-T2A-NT-W	*-minC-FP-I2-NT-W	*-RSV-luc-T2A-NT-W-S	*-SV40-luc-I2-NT-W-S
*-minR-FP-T2A-CR-W	*-minC-FP-I2-CR-W	*-RSV-luc-T2A-CR-W-S	*-SV40-luc-I2-CR-W-S
*-minR-FP-T2A-chR-W	*-minC-FP-I2-chR-W	*-RSV-luc-T2A-chR-W-S	*-SV40-luc-I2-chR-W-S
*-minR-FP-T2A-gRNA-W	*-minC-FP-I2-gRNA-W	*-RSV-luc-T2A-gRNA-W-S	*-SV40-luc-I2-gRNA-W-S
*-minR-FP-T2A-Nuc-W	*-minC-FP-I2-Nuc-W	*-RSV-luc-T2A-Nuc-W-S	*-SV40-luc-I2-Nuc-W-S
*-CMV-FP-T2A-enz-W	*-mut-FP-I2-enz-W	*-minB-luc-E2A-enz-W	*-H68-luc-I2-enz-W-S
*-CMV-FP-T2A-TF-W	*-mut-FP-I2-TF-W	*-minB-luc-E2A-TF-W	*-H68-luc-I2-TF-W-S
*-CMV-FP-T2A-rec-W	*-mut-FP-I2-rec-W	*-minB-luc-E2A-rec-W	*-H68-luc-I2-rec-W-S
*-CMV-FP-T2A-CTP-W	*-mut-FP-I2-CTP-W	*-minB-luc-E2A-CTP-W	*-H68-luc-I2-CTP-W-S
*-CMV-FP-T2A-SM-W	*-mut-FP-I2-SM-W	*-minB-luc-E2A-SM-W	*-H68-luc-I2-SM-W-S
*-CMV-FP-T2A-NT-W	*-mut-FP-I2-NT-W	*-minB-luc-E2A-NT-W	*-H68-luc-I2-NT-W-S
*-CMV-FP-T2A-CR-W	*-mut-FP-I2-CR-W	*-minB-luc-E2A-CR-W	*-H68-luc-I2-CR-W-S
*-CMV-FP-T2A-chR-W	*-mut-FP-I2-chR-W	*-minB-luc-E2A-chR-W	*-H68-luc-I2-chR-W-S
*-CMV-FP-T2A-gRNA-W	*-mut-FP-I2-gRNA-W	*-minB-luc-E2A-gRNA-W	*-H68-luc-I2-gRNA-W-S
*-CMV-FP-T2A-Nuc-W	*-mut-FP-I2-Nuc-W	*-minB-luc-E2A-Nuc-W	*-H68-luc-I2-Nuc-W-S
*-SV40-FP-T2A-enz-W	*-minR-FP-I2-enz-W	*-minC-luc-E2A-enz-W	*-RSV-luc-I2-enz-W-S
*-SV40-FP-T2A-TF-W	*-minR-FP-I2-TF-W	*-minC-luc-E2A-TF-W	*-RSV-luc-I2-TF-W-S
*-SV40-FP-T2A-rec-W	*-minR-FP-I2-rec-W	*-minC-luc-E2A-rec-W	*-RSV-luc-I2-rec-W-S
*-SV40-FP-T2A-CTP-W	*-minR-FP-I2-CTP-W	*-minC-luc-E2A-CTP-W	*-RSV-luc-I2-CTP-W-S
*-SV40-FP-T2A-SM-W	*-minR-FP-I2-SM-W	*-minC-luc-E2A-SM-W	*-RSV-luc-I2-SM-W-S
*-SV40-FP-T2A-NT-W	*-minR-FP-I2-NT-W	*-minC-luc-E2A-NT-W	*-RSV-luc-I2-NT-W-S
*-SV40-FP-T2A-CR-W	*-minR-FP-I2-CR-W	*-minC-luc-E2A-CR-W	*-RSV-luc-I2-CR-W-S
*-SV40-FP-T2A-chR-W	*-minR-FP-I2-chR-W	*-minC-luc-E2A-chR-W	*-RSV-luc-I2-chR-W-S
*-SV40-FP-T2A-gRNA-W	*-minR-FP-I2-gRNA-W	*-minC-luc-E2A-gRNA-W	*-RSV-luc-I2-gRNA-W-S
*-SV40-FP-T2A-Nuc-W	*-minR-FP-I2-Nuc-W	*-minC-luc-E2A-Nuc-W	*-RSV-luc-I2-Nuc-W-S
*-minB-IT-W	*-minB-IT-bG	*-minB-IT-S	*-minB-IT-W-bG
*-minC-IT-W	*-minC-IT-bG	*-minC-IT-S	*-minC-IT-W-bG
*-mut-IT-W	*-mut-IT-bG	*-mut-IT-S	*-mut-IT-W-bG
*-minR-IT-W	*-minR-IT-bG	*-minR-IT-S	*-minR-IT-W-bG
*-CMV-IT-W	*-CMV-IT-bG	*-CMV-IT-S	*-CMV-IT-W-bG
*-SV40-IT-W	*-SV40-IT-bG	*-SV40-IT-S	*-SV40-IT-W-bG
*-H68-IT-W	*-H68-IT-bG	*-H68-IT-S	*-H68-IT-W-bG
*-RSV-IT-W	*-RSV-IT-bG	*-RSV-IT-S	*-RSV-IT-W-bG
*-minB-IT-W-S	*-minB-luc-P2A-IT-W	*-minB-luc-P2A-IT-bG	*-minB-luc-P2A-IT-S

FIG. 16 cont'd

*-minC-IT-W-S	*-minC-luc-P2A-IT-W	*-minC-luc-P2A-IT-bG	*-minC-luc-P2A-IT-S
*-mut-IT-W-S	*-mut-luc-P2A-IT-W	*-mut-luc-P2A-IT-bG	*-mut-luc-P2A-IT-S
*-minR-IT-W-S	*-minR-luc-P2A-IT-W	*-minR-luc-P2A-IT-bG	*-minR-luc-P2A-IT-S
*-CMV-IT-W-S	*-CMV-luc-P2A-IT-W	*-CMV-luc-P2A-IT-bG	*-CMV-luc-P2A-IT-S
*-SV40-IT-W-S	*-SV40-luc-P2A-IT-W	*-SV40-luc-P2A-IT-bG	*-SV40-luc-P2A-IT-S
*-H68-IT-W-S	*-H68-luc-P2A-IT-W	*-H68-luc-P2A-IT-bG	*-H68-luc-P2A-IT-S
*-RSV-IT-W-S	*-RSV-luc-P2A-IT-W	*-RSV-luc-P2A-IT-bG	*-RSV-luc-P2A-IT-S
*-minB-luc-P2A-IT-W-bG	*-minB-luc-P2A-IT-W-S	*-minB-luc-P2A-IT-W	*-minB-luc-P2A-IT-bG
*-minC-luc-P2A-IT-W-bG	*-minC-luc-P2A-IT-W-S	*-minC-luc-P2A-IT-W	*-minC-luc-P2A-IT-bG
*-mut-luc-P2A-IT-W-bG	*-mut-luc-P2A-IT-W-S	*-mut-luc-P2A-IT-W	*-mut-luc-P2A-IT-bG
*-minR-luc-P2A-IT-W-bG	*-minR-luc-P2A-IT-W-S	*-minR-luc-P2A-IT-W	*-minR-luc-P2A-IT-bG
*-CMV-luc-P2A-IT-W-bG	*-CMV-luc-P2A-IT-W-S	*-CMV-luc-P2A-IT-W	*-CMV-luc-P2A-IT-bG
*-SV40-luc-P2A-IT-W-bG	*-SV40-luc-P2A-IT-W-S	*-SV40-luc-P2A-IT-W	*-SV40-luc-P2A-IT-bG
*-H68-luc-P2A-IT-W-bG	*-H68-luc-P2A-IT-W-S	*-H68-luc-P2A-IT-W	*-H68-luc-P2A-IT-bG
*-RSV-luc-P2A-IT-W-bG	*-RSV-luc-P2A-IT-W-S	*-RSV-luc-P2A-IT-W	*-RSV-luc-P2A-IT-bG
*-minB-FP-P2A-IT-S	*-minB-FP-P2A-IT-W-bG	*-minB-FP-P2A-IT-W-S	*-minB-luc-T2A-IT-W
*-minC-FP-P2A-IT-S	*-minC-FP-P2A-IT-W-bG	*-minC-FP-P2A-IT-W-S	*-minC-luc-T2A-IT-W
*-mut-FP-P2A-IT-S	*-mut-FP-P2A-IT-W-bG	*-mut-FP-P2A-IT-W-S	*-mut-luc-T2A-IT-W
*-minR-FP-P2A-IT-S	*-minR-FP-P2A-IT-W-bG	*-minR-FP-P2A-IT-W-S	*-minR-luc-T2A-IT-W
*-CMV-FP-P2A-IT-S	*-CMV-FP-P2A-IT-W-bG	*-CMV-FP-P2A-IT-W-S	*-CMV-luc-T2A-IT-W
*-SV40-FP-P2A-IT-S	*-SV40-FP-P2A-IT-W-bG	*-SV40-FP-P2A-IT-W-S	*-SV40-luc-T2A-IT-W
*-H68-FP-P2A-IT-S	*-H68-FP-P2A-IT-W-bG	*-H68-FP-P2A-IT-W-S	*-H68-luc-T2A-IT-W
*-RSV-FP-P2A-IT-S	*-RSV-FP-P2A-IT-W-bG	*-RSV-FP-P2A-IT-W-S	*-RSV-luc-T2A-IT-W
*-minB-luc-T2A-IT-bG	*-minB-luc-T2A-IT-S	*-minB-luc-T2A-IT-W-bG	*-minB-luc-T2A-IT-W-S
*-minC-luc-T2A-IT-bG	*-minC-luc-T2A-IT-S	*-minC-luc-T2A-IT-W-bG	*-minC-luc-T2A-IT-W-S
*-mut-luc-T2A-IT-bG	*-mut-luc-T2A-IT-S	*-mut-luc-T2A-IT-W-bG	*-mut-luc-T2A-IT-W-S
*-minR-luc-T2A-IT-bG	*-minR-luc-T2A-IT-S	*-minR-luc-T2A-IT-W-bG	*-minR-luc-T2A-IT-W-S
*-CMV-luc-T2A-IT-bG	*-CMV-luc-T2A-IT-S	*-CMV-luc-T2A-IT-W-bG	*-CMV-luc-T2A-IT-W-S
*-SV40-luc-T2A-IT-bG	*-SV40-luc-T2A-IT-S	*-SV40-luc-T2A-IT-W-bG	*-SV40-luc-T2A-IT-W-S
*-H68-luc-T2A-IT-bG	*-H68-luc-T2A-IT-S	*-H68-luc-T2A-IT-W-bG	*-H68-luc-T2A-IT-W-S
*-RSV-luc-T2A-IT-bG	*-RSV-luc-T2A-IT-S	*-RSV-luc-T2A-IT-W-bG	*-RSV-luc-T2A-IT-W-S
*-minB-FP-T2A-IT-W	*-minB-FP-T2A-IT-bG	*-minB-FP-T2A-IT-S	*-minB-FP-T2A-IT-W-bG
*-minC-FP-T2A-IT-W	*-minC-FP-T2A-IT-bG	*-minC-FP-T2A-IT-S	*-minC-FP-T2A-IT-W-bG
*-mut-FP-T2A-IT-W	*-mut-FP-T2A-IT-bG	*-mut-FP-T2A-IT-S	*-mut-FP-T2A-IT-W-bG

FIG. 16 cont'd

*-minR-FP-T2A-IT-W	*-minR-FP-T2A-IT-bG	*-minR-FP-T2A-IT-S	*-minR-FP-T2A-IT-W-bG
*-CMV-FP-T2A-IT-W	*-CMV-FP-T2A-IT-bG	*-CMV-FP-T2A-IT-S	*-CMV-FP-T2A-IT-W-bG
*-SV40-FP-T2A-IT-W	*-SV40-FP-T2A-IT-bG	*-SV40-FP-T2A-IT-S	*-SV40-FP-T2A-IT-W-bG
*-H68-FP-T2A-IT-W	*-H68-FP-T2A-IT-bG	*-H68-FP-T2A-IT-S	*-H68-FP-T2A-IT-W-bG
*-RSV-FP-T2A-IT-W	*-RSV-FP-T2A-IT-bG	*-RSV-FP-T2A-IT-S	*-RSV-FP-T2A-IT-W-bG
*-minB-FP-T2A-IT-W-S	*-minB-luc-E2A-IT-W	*-minB-luc-E2A-IT-bG	*-minB-luc-E2A-IT-S
*-minC-FP-T2A-IT-W-S	*-minC-luc-E2A-IT-W	*-minC-luc-E2A-IT-bG	*-minC-luc-E2A-IT-S
*-mut-FP-T2A-IT-W-S	*-mut-luc-E2A-IT-W	*-mut-luc-E2A-IT-bG	*-mut-luc-E2A-IT-S
*-minR-FP-T2A-IT-W-S	*-minR-luc-E2A-IT-W	*-minR-luc-E2A-IT-bG	*-minR-luc-E2A-IT-S
*-CMV-FP-T2A-IT-W-S	*-CMV-luc-E2A-IT-W	*-CMV-luc-E2A-IT-bG	*-CMV-luc-E2A-IT-S
*-SV40-FP-T2A-IT-W-S	*-SV40-luc-E2A-IT-W	*-SV40-luc-E2A-IT-bG	*-SV40-luc-E2A-IT-S
*-H68-FP-T2A-IT-W-S	*-H68-luc-E2A-IT-W	*-H68-luc-E2A-IT-bG	*-H68-luc-E2A-IT-S
*-RSV-FP-T2A-IT-W-S	*-RSV-luc-E2A-IT-W	*-RSV-luc-E2A-IT-bG	*-RSV-luc-E2A-IT-S
*-minB-luc-E2A-IT-W-bG	*-minB-luc-E2A-IT-W-S	*-minB-FP-E2A-IT-W	*-minB-FP-E2A-IT-bG
*-minC-luc-E2A-IT-W-bG	*-minC-luc-E2A-IT-W-S	*-minC-FP-E2A-IT-W	*-minC-FP-E2A-IT-bG
*-mut-luc-E2A-IT-W-bG	*-mut-luc-E2A-IT-W-S	*-mut-FP-E2A-IT-W	*-mut-FP-E2A-IT-bG
*-minR-luc-E2A-IT-W-bG	*-minR-luc-E2A-IT-W-S	*-minR-FP-E2A-IT-W	*-minR-FP-E2A-IT-bG
*-CMV-luc-E2A-IT-W-bG	*-CMV-luc-E2A-IT-W-S	*-CMV-FP-E2A-IT-W	*-CMV-FP-E2A-IT-bG
*-SV40-luc-E2A-IT-W-bG	*-SV40-luc-E2A-IT-W-S	*-SV40-FP-E2A-IT-W	*-SV40-FP-E2A-IT-bG
*-H68-luc-E2A-IT-W-bG	*-H68-luc-E2A-IT-W-S	*-H68-FP-E2A-IT-W	*-H68-FP-E2A-IT-bG
*-RSV-luc-E2A-IT-W-bG	*-RSV-luc-E2A-IT-W-S	*-RSV-FP-E2A-IT-W	*-RSV-FP-E2A-IT-bG
*-minB-FP-E2A-IT-S	*-minB-FP-E2A-IT-W-bG	*-minB-FP-E2A-IT-W-S	*-minB-luc-F2A-IT-W
*-minC-FP-E2A-IT-S	*-minC-FP-E2A-IT-W-bG	*-minC-FP-E2A-IT-W-S	*-minC-luc-F2A-IT-W
*-mut-FP-E2A-IT-S	*-mut-FP-E2A-IT-W-bG	*-mut-FP-E2A-IT-W-S	*-mut-luc-F2A-IT-W
*-minR-FP-E2A-IT-S	*-minR-FP-E2A-IT-W-bG	*-minR-FP-E2A-IT-W-S	*-minR-luc-F2A-IT-W
*-CMV-FP-E2A-IT-S	*-CMV-FP-E2A-IT-W-bG	*-CMV-FP-E2A-IT-W-S	*-CMV-luc-F2A-IT-W
*-SV40-FP-E2A-IT-S	*-SV40-FP-E2A-IT-W-bG	*-SV40-FP-E2A-IT-W-S	*-SV40-luc-F2A-IT-W
*-H68-FP-E2A-IT-S	*-H68-FP-E2A-IT-W-bG	*-H68-FP-E2A-IT-W-S	*-H68-luc-F2A-IT-W
*-RSV-FP-E2A-IT-S	*-RSV-FP-E2A-IT-W-bG	*-RSV-FP-E2A-IT-W-S	*-RSV-luc-F2A-IT-W
*-minB-luc-F2A-IT-bG	*-minB-luc-F2A-IT-S	*-minB-luc-F2A-IT-W-bG	*-minB-luc-F2A-IT-W-S
*-minC-luc-F2A-IT-bG	*-minC-luc-F2A-IT-S	*-minC-luc-F2A-IT-W-bG	*-minC-luc-F2A-IT-W-S
*-mut-luc-F2A-IT-bG	*-mut-luc-F2A-IT-S	*-mut-luc-F2A-IT-W-bG	*-mut-luc-F2A-IT-W-S
*-minR-luc-F2A-IT-bG	*-minR-luc-F2A-IT-S	*-minR-luc-F2A-IT-W-bG	*-minR-luc-F2A-IT-W-S
*-CMV-luc-F2A-IT-bG	*-CMV-luc-F2A-IT-S	*-CMV-luc-F2A-IT-W-bG	*-CMV-luc-F2A-IT-W-S

FIG. 16 cont'd

*-SV40-luc-F2A-IT-bG	*-SV40-luc-F2A-IT-S	*-SV40-luc-F2A-IT-W-bG	*-SV40-luc-F2A-IT-W-S
*-H68-luc-F2A-IT-bG	*-H68-luc-F2A-IT-S	*-H68-luc-F2A-IT-W-bG	*-H68-luc-F2A-IT-W-S
*-RSV-luc-F2A-IT-bG	*-RSV-luc-F2A-IT-S	*-RSV-luc-F2A-IT-W-bG	*-RSV-luc-F2A-IT-W-S
*-minB-FP-F2A-IT-W	*-minB-FP-F2A-IT-bG	*-minB-FP-F2A-IT-S	*-minB-FP-F2A-IT-W-bG
*-minC-FP-F2A-IT-W	*-minC-FP-F2A-IT-bG	*-minC-FP-F2A-IT-S	*-minC-FP-F2A-IT-W-bG
*-mut-FP-F2A-IT-W	*-mut-FP-F2A-IT-bG	*-mut-FP-F2A-IT-S	*-mut-FP-F2A-IT-W-bG
*-minR-FP-F2A-IT-W	*-minR-FP-F2A-IT-bG	*-minR-FP-F2A-IT-S	*-minR-FP-F2A-IT-W-bG
*-CMV-FP-F2A-IT-W	*-CMV-FP-F2A-IT-bG	*-CMV-FP-F2A-IT-S	*-CMV-FP-F2A-IT-W-bG
*-SV40-FP-F2A-IT-W	*-SV40-FP-F2A-IT-bG	*-SV40-FP-F2A-IT-S	*-SV40-FP-F2A-IT-W-bG
*-H68-FP-F2A-IT-W	*-H68-FP-F2A-IT-bG	*-H68-FP-F2A-IT-S	*-H68-FP-F2A-IT-W-bG
*-RSV-FP-F2A-IT-W	*-RSV-FP-F2A-IT-bG	*-RSV-FP-F2A-IT-S	*-RSV-FP-F2A-IT-W-bG
*-minB-FP-F2A-IT-W-S	*-minB-luc-I2-IT-W	*-minB-luc-I2-IT-bG	*-minB-luc-I2-IT-S
*-minC-FP-F2A-IT-W-S	*-minC-luc-I2-IT-W	*-minC-luc-I2-IT-bG	*-minC-luc-I2-IT-S
*-mut-FP-F2A-IT-W-S	*-mut-luc-I2-IT-W	*-mut-luc-I2-IT-bG	*-mut-luc-I2-IT-S
*-minR-FP-F2A-IT-W-S	*-minR-luc-I2-IT-W	*-minR-luc-I2-IT-bG	*-minR-luc-I2-IT-S
*-CMV-FP-F2A-IT-W-S	*-CMV-luc-I2-IT-W	*-CMV-luc-I2-IT-bG	*-CMV-luc-I2-IT-S
*-SV40-FP-F2A-IT-W-S	*-SV40-luc-I2-IT-W	*-SV40-luc-I2-IT-bG	*-SV40-luc-I2-IT-S
*-H68-FP-F2A-IT-W-S	*-H68-luc-I2-IT-W	*-H68-luc-I2-IT-bG	*-H68-luc-I2-IT-S
*-RSV-FP-F2A-IT-W-S	*-RSV-luc-I2-IT-W	*-RSV-luc-I2-IT-bG	*-RSV-luc-I2-IT-S
*-minB-luc-I2-IT-W-bG	*-minB-luc-I2-IT-W-S	*-minB-FP-I2-IT-W	*-minB-FP-I2-IT-bG
*-minC-luc-I2-IT-W-bG	*-minC-luc-I2-IT-W-S	*-minC-FP-I2-IT-W	*-minC-FP-I2-IT-bG
*-mut-luc-I2-IT-W-bG	*-mut-luc-I2-IT-W-S	*-mut-FP-I2-IT-W	*-mut-FP-I2-IT-bG
*-minR-luc-I2-IT-W-bG	*-minR-luc-I2-IT-W-S	*-minR-FP-I2-IT-W	*-minR-FP-I2-IT-bG
*-CMV-luc-I2-IT-W-bG	*-CMV-luc-I2-IT-W-S	*-CMV-FP-I2-IT-W	*-CMV-FP-I2-IT-bG
*-SV40-luc-I2-IT-W-bG	*-SV40-luc-I2-IT-W-S	*-SV40-FP-I2-IT-W	*-SV40-FP-I2-IT-bG
*-H68-luc-I2-IT-W-bG	*-H68-luc-I2-IT-W-S	*-H68-FP-I2-IT-W	*-H68-FP-I2-IT-bG
*-RSV-luc-I2-IT-W-bG	*-RSV-luc-I2-IT-W-S	*-RSV-FP-I2-IT-W	*-RSV-FP-I2-IT-bG
*-minB-FP-I2-IT-S	*-minB-FP-I2-IT-W-bG	*-minB-FP-I2-IT-W-S	*-CMV-FP-I2-IT-S
*-minC-FP-I2-IT-S	*-minC-FP-I2-IT-W-bG	*-minC-FP-I2-IT-W-S	*-SV40-FP-I2-IT-S
*-mut-FP-I2-IT-S	*-mut-FP-I2-IT-W-bG	*-mut-FP-I2-IT-W-S	*-H68-FP-I2-IT-S
*-minR-FP-I2-IT-S	*-minR-FP-I2-IT-W-bG	*-minR-FP-I2-IT-W-S	*-RSV-FP-I2-IT-S
*-CMV-FP-I2-IT-S	*-CMV-FP-I2-IT-W-bG	*-CMV-FP-I2-IT-W-S	*-H68-FP-I2-IT-W-S
*-SV40-FP-I2-IT-W-bG	*-SV40-FP-I2-IT-W-S	*-RSV-FP-I2-IT-W-bG	*-RSV-FP-I2-IT-W-S

72/104

FIG. 17

hDLX I56i enhancer:

TATGCACTCACAGTGGTTTGGCATGCATCTGGTGAATTTTTTTTAACGAAAAATTAGTGTTG
GTTTCGATGTATGGTAGCATTCTCCCTAACGTAATTTGAATAATTCAGCAAAGCCCCACTAC
CAGCTGTACTTCTGCAGCCTCTTCCATTCTTTTCAGCATTATAATTTTGGTTAATTTTCAATT
TTAGGTCCTACGTCTCTGCAATTTGTGTATGAATAACAGAATAATTTCCCTCTTTTGTTCGC
CTTTCCTGTTTCTGAATCTAAATAAAGATGGCTTTTTTAGTATTAAAAGTGGAAGAAAATTACA
GGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAAT
TATGGCTGCATTTAAGAGAATGGAAAAAACCTTCTTGTGGATAAAAAACCTTAAATTGTCCC
CAATGTCTGCTTCAAATTGGATGGCACTGCAGCTGGAGGCTTTGTTTCAGAATTGATCCTGG
GGAGCTACGAACCCAAAGTTTCACAGTAGG (SEQ ID NO: 1)

Core of the hDLX I56i enhancer:

CTAAATAAAGATGGCTTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACG
GTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGA
GAATGG (SEQ ID NO: 2)

3xhI56iCore, Triply Concatamerized Core of the hDLX I56i enhancer:

CTAAATAAAGATGGCTTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACG
GTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGA
GAATGGCTAAATAAAGATGGCTTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATC
TTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCA
TTTAAGAGAATGGCTAAATAAAGATGGCTTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGT
AATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTAT
GGCTGCATTTAAGAGAATGG (SEQ ID NO: 3)

Murine I56i Enhancer (core is the same as human):

TATACACTCACAGTGGTTTGGCATATATTTGGTGAATTTTTTAAGGAAAAATTAGTGTTGGT
TTCGATATATGGTAGCTTTTTCTCTAACATAATTTGAATAATTCAGCAAAGCCCTACTACCAG
CTGTACTTCTGCAGCCTCTTCCATTCTTTCCAGCATTATAATTTTGGTTAATTTTCAATTTTA
GGTCCTACGTCTCTGCAATTTGTGTATGAATAACAGAATAATTTCCCTCTTTTGTTCGCCTT
TCCTGTTTCTGAATCTAAATAAAGATGGCTTTTTTAGTATTAAAAGTGGAAGAAAATTACAGG
TAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTA
TGGCTGCATTTAAGAGAATGGAAAAAACCTTCTTGTGGATAAAAAACCTTAAATTGTCCCCA
ATGTCTGCTTCAAATTGGATGGCACTGCAGCTGGAGGCTTTGTTTCAGAATTGATCCTGGGG
AGCTACGAACCCAAAGTTTCACAGTAGG (SEQ ID NO: 4)

Zebrafish I46i Enhancer:

ACATTGTAATTTTAGATAATATCCCAAGCGTTCACTCTCCTCGGCAATTTGTACATGAATAAC
CGAATAATTTTATCTTTTGTTCGTCTTTGCCACTTCAAATCCAAATAAAGATGCCTTTTAGT
ATTAAAAGTGGTAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGG
GCTACATCAAAAATTACCCTAATTATGTCTGCATTTATGAGAATGGAAAAAACCTCTCTT
GGATAAAACCCATAAATTGTCCCAAATATCT (SEQ ID NO: 5)

Core of the Zebrafish I46i Enhancer:

CCAAATAAAGATGCCTTTTTAGTATTAAAAGTGGTAGAAAATTACAGGTAATTATCTTTGACG
GTAAAAACGCTGTAATCAGCGGGCTACATCAAAAATTACCCTAATTATGTCTGCATTTATGA
GAATGG (SEQ ID NO: 6)

FIG. 17 cont'd

3x Concatamerized Core of the Zebrafish I46i Enhancer:

CCAAATAAAGATGCCTTTTAGTATTAAAAGTGGTAGAAAATTACAGGTAATTATCTTTGACG
GTAAAAACGCTGTAATCAGCGGGCTACATCAAAAATTACCCTAATTATGTCTGCATTTATGA
GAATGGCCAAATAAAGATGCCTTTTAGTATTAAAAGTGGTAGAAAATTACAGGTAATTATCT
TTGACGGTAAAAACGCTGTAATCAGCGGGCTACATCAAAAATTACCCTAATTATGTCTGCAT
TTATGAGAATGGCCAAATAAAGATGCCTTTTAGTATTAAAAGTGGTAGAAAATTACAGGTAA
TTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATCAAAAATTACCCTAATTATGT
CTGCATTTATGAGAATGG (SEQ ID NO: 7)

Beta-Globin Minimal Promoter (pBGmin/minBGlobin/minBGprom):

GGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTG (SEQ ID
NO: 8)

minCMV Promoter:

GAGGTAGGCGTGACGGTGGGAGGCCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGA
TCGCCTGG (SEQ ID NO: 9)

Mutated minCMV Promoter (Sacl RE site removed):

GAGGTAGGCGTGACGGTGGGAGGCCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGA
TCGCCTGG (SEQ ID NO: 10)

minRho Promoter:

GATTCAGCCGGGAGCTTAGGGAGGGGAGGTCACTTCATAAGGGCCTGGGGGGGGAGTTG
GAGCCACGAGTCGTCCAGCCGGAGCCCCGTGTGGCTGAGCTCCGGCCTCAGAAAGCATCC
CC (SEQ ID NO: 11)

Hsp68 minimal Promoter (proHsp68):

CAGGAACATCCAAACTGAGCAGCCGGGGTCCCCCCCCACCCCCACCCCCGCCCCACGCGG
CAACTTTGAGCCTGTGCTGGGACAGAGCCTCTAGTTCCTAAATTAGTCCATGAGGTCAGAG
GCAGCACTGCCATTGTAACGCGATTGGAGAGGATCACGTACCCGGACACGCCCCCAGGC
ATCTCCCTGGGTCTCCTAAACTTGGCGGGGAGAAGTTTTAGCCCTTAAGTTTTAGCCTTTAA
CCCCCATATTCAGAACTGTGCGAGTTGGCGAAACCCACAAATCACAACTGTACACA
ACACCGAGCTAGAGGTGATCTTTCTTGTCATTCCACACAGGCCTTAGTAATGCGTCGCCA
TAGCAACAGTGTCAGTAGTAGCACCAGCACTTCCCCACACCCTCCCCCTCAGGAATCCGTA
CTCTCCAGTGAACCCAGAAACCTCTGGAGAGTTCTGGACAAGGGCGGAACCCACAACCTC
CGATTACTCAAGGGAGGCGGGGAAGCTCCACCAGACGCGAACTGCTGGAAGATTCTG
GCCCCAAGGCCTCCTCCGGCTCGCTGATTGGCCAGCGGAGAGTGGGCGGGGCGCGTG
AAGACTCCTTAAAGGCGCAGGGCGGCGAGCAGGTCACCAGACGCTGACAGCTACTCAGA
ACCAAATCTGGTTCCATCCAGAGACAAGCGAAGACAAGAGAAGCAGAGCGAGCGGCGCGT
TCCCGATCCTCGGCCAGGACCAGCCTTCCCCAGAGCATCCCTGCCGCGGAGCGCAACCT
TCCAGGAGCATCCCTGCCGCGGAGCGCAACTTTCCCCGAGCATCCACGCCGCGGAGC
GCAGCCTTCCAGAAAGCAGAGCGCGGCGGCC (SEQ ID NO: 12)

FIG. 17 cont'd

SYFP2:

ATGGTGAGCAAGGGCGAGGAGCTGTTACACGGGGTGGTGCCCATCCTGGTTCGAGCTGGA
CGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCT
ACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCC
ACCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGCCCCGCTACCCCGACCAT
GAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCAT
CTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACA
CCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAG
AAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGGCGTGCA
GCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCG
ACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGT
ACAAGTAA (SEQ ID NO: 13)

EGFP:

ATGGTGAGCAAGGGCGAGGAGCTGTTACACGGGGTGGTGCCCATCCTGGTTCGAGCTGGA
CGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCT
ACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCA
CCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCATGA
AGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCT
TCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACC
CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGG
GCACAAGCTGGAGTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAA
GAACGGCATCAAGGTGAAGTTCAGATCCGCCACAACATCGAGGACGGCAGCGTGACGCT
CGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACA
ACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACA
TGGTCCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACA
AGTAA (SEQ ID NO: 14)

Optimized Flp recombinase (FlpO):

ATGGCTCCTAAGAAGAAGAGGAAGGTGATGAGCCAGTTCGACATCCTGTGCAAGACCCCC
CCCAAGGTGCTGGTGCAGGAGTTCGTGGAGAGATTCGAGAGGCCAGCGGCGAGAAGAT
CGCCAGCTGTGCCGCCGAGCTGACCTACCTGTGCTGGATGATCACCCACAACGGCACCG
CCATCAAGAGGGCCACCTTCATGAGCTACAACACCATCATCAGCAACAGCCTGAGCTTCGA
CATCGTGAACAAGAGCCTGCAGTTCAAGTACAAGACCCAGAAGGCCACCATCCTGGAGGC
CAGCCTGAAGAAGCTGATCCCCGCCTGGGAGTTACCATCATCCCTTACAACGGCCAGAA
GCACCAGAGCGACATCACCGACATCGTGTCCAGCCTGCAGCTGCAGTTCGAGAGCAGCG
AGGAGGCCGACAAGGGCAACAGCCACAGCAAGAAGATGCTGAAGGCCCTGCTGTCCGAG
GGCGAGAGCATCTGGGAGATCACCGAGAAGATCCTGAACAGCTTCGAGTACACCAGCAGG
TTCACCAAGACCAAGACCCTGTACCAGTTCCTGTTCTTGCCACATTCATCAACTGCGGCA
GGTTCAGCGACATCAAGAACGTGGACCCCAAGAGCTTCAAGCTGGTGCAGAACAAGTACC
TGGGCGTGATCATTAGTGCCTGGTGACCGAGACCAAGACAAGCGTGTCCAGGCACATCT
ACTTTTTAGCGCCAGAGGCAGGATCGACCCCTGGTGTACCTGGACGAGTTCCTGAGGA
ACAGCGAGCCCGTGCTGAAGAGAGTGAACAGGACCGGCAACAGCAGCAGCAACAAGCAG
GAGTACCAGCTGCTGAAGGACAACCTGGTGCAGCTACAACAAGGCCCTGAAGAAGAAC
GCCCCCTACCCCATCTTCGCTATCAAGAACGGCCCTAAGAGCCACATCGGCAGGCACCTG
ATGACCAGCTTTCTGAGCATGAAGGGCCTGACCGAGCTGACAAACGTGGTGGGCAACTGG
AGCGACAAGAGGGCCTCCGCCGTGGCCAGGACCACTACACCCACCAGATCACCGCCAT

FIG. 17 cont'd

CCCCGACCACTACTTCGCCCTGGTGTCCAGGTACTACGCCTACGACCCCATCAGCAAGGA
GATGATCGCCCTGAAGGACGAGACCAACCCCATCGAGGAGTGGCAGCACATCGAGCAGC
TGAAGGGCAGCGCCGAGGGCAGCATCAGATACCCCGCCTGGAACGGCATCATCAGCCAG
GAGGTGCTGGACTACCTGAGCAGCTACATCAACAGGCGGATCTGA (SEQ ID NO: 15)

Improved Cre recombinase (iCre):

ATGGTGCCCAAGAAGAAGAGGAAAGTCTCCAACCTGCTGACTGTGCACCAAAACCTGCCT
GCCCTCCCTGTGGATGCCACCTCTGATGAAGTCAGGAAGAACCTGATGGACATGTTTCAGG
GACAGGCAGGCCTTCTCTGAACACACCTGGAAGATGCTCCTGTCTGTGTGCAGATCCTGG
GCTGCCTGGTGAAGCTGAACAACAGGAAATGGTTCCCTGCTGAACCTGAGGATGTGAGG
GACTACCTCCTGTACCTGCAAGCCAGAGGCCTGGCTGTGAAGACCATCCAACAGCACCTG
GGCCAGCTCAACATGCTGCACAGGAGATCTGGCCTGCCTCGCCCTTCTGACTCCAATGCT
GTGTCCCTGGTGTGATGAGGAGAATCAGAAAGGAGAATGTGGATGCTGGGGAGAGAGCCAA
GCAGGCCCTGGCCTTTGAACGCACTGACTTTGACCAAGTCAGATCCCTGATGGAGAACTC
TGACAGATGCCAGGACATCAGGAACCTGGCCTTCTGGGCATTGCCTACAACACCCTGCT
GCGCATTGCCGAAATTGCCAGAATCAGAGTGAAGGACATCTCCCGCACCGATGGTGGGAG
AATGCTGATCCACATTGGCAGGACCAAGACCCTGGTGTCCACAGCTGGTGTGGAGAAGGC
CCTGTCCCTGGGGGTTACCAAGCTGGTGGAGAGATGGATCTCTGTGTCTGGTGTGGCTGA
TGACCCCAACAACACTACCTGTTCTGCCGGGTGAGAAAGAAATGGTGTGGCTGCCCTTCTGC
CACCTCCCAACTGTCCACCCGGGCCCTGGAAGGGATCTTTGAGGCCACCCACCGCCTGAT
CTATGGTGCCAAGGATGACTCTGGGCAGAGATACCTGGCCTGGTCTGGCCACTCTGCCAG
AGTGGGTGCTGCCAGGGACATGGCCAGGGCTGGTGTGTCCATCCCTGAAATCATGCAGG
CTGGTGGCTGGACCAATGTGAACATTGTGATGAACTACATCAGAAACCTGGACTCTGAGAC
TGGGGCCATGGTGAAGGCTGCTCGAGGATGGGGACTAA (SEQ ID NO: 16)

NavMs, endogenous sequence:

ATGTCACGCAAAATAAGAGATTTAATCGAATCCAAACGCTTTCAAACGTCATCACCGCCAT
TATTGTGCTCAATGGCGCTGTGCTGGGTCTGCTGACCGATAACAACCCTATCGGCCTCCAG
CCAAAACCTGCTGGAGCGTGTGGATCAACTTTGTCTGACTATCTTTATTGTTGAAATATCCC
TGAAAATATACGCCTATGGCGTGCGAGGCTTTTTCCGCAGCGGCTGGAATCTGTTTGATTT
TGTGATTGTGGCCATCGCGCTTATGCCCGCCCAGGGTAGCCTATCGGTGCTGCGAACCTT
CCGTATATTCCGCGTCATGCGGCTCGTATCGGTGATACCAACCATGCGAAGAGTGGTGCA
AGGCATGCTCTTGGCACTGCCCGGCGTGGGATCGGTAGCGGCACTGTTGACGGTGGTCT
TCTATATTGCGGCTGTCATGGCCACCAATCTCTACGGGGCAACCTTCCCTGAATGGTTTGG
TGATCTTAGCAAGAGCCTGTACACACTATTTTCAGGTGATGACCTTAGAGTCATGGTCTATG
GGCATTGTGCGTCCAGTGATGAACGTTTCATCCCAACGCATGGGTTTTTTTCATCCCCTTCA
TCATGCTCACCACTTTACCGTGCTCAACCTGTTTATTGGCATTATTGTAGATGCCATGGCC
ATCACCAAGGAACAGGAGGAAGAGGCCAAAACCGGCCACCACCAAGAGCCTATTAGCCAA
ACATTGCTCCATCTGGGAGATCGCCTAGATAGGATCGAAAAGCAGCTTGCAGCAAAACAAC
GAGCTCTTACAACGACAACAGCCGCAAAAAAATAG (SEQ ID NO: 17)

NavMs, codon optimized, with N-terminal 3x HA tag and linker:

ATGGTTTATCCGTATGATGTTTCTGACTATGCAGGATCCTATCCTTATGATGTTCCCGATTA
CGCTGGTTCTTACCCTTACGATGTTCCCGATTATGCCAGTTCTGGATTGGTGCCACGAGGC
AGCCACATGAGCCGGAAGATCAGAGATCTTATCGAATCTAAGAGATTTCAGAATGTTATTAC
CGCGATAATCGTACTCAACGGGGCGGTGCTCGGTCTCCTCACCGATAACCATTTGAGCGC
TTCTAGCCAGAACCTGCTCGAAAGGGTTGACCAACTGTGCCTGACAATTTTTATCGTGGA
ATTAGCTTGAATAATTTACGCCTACGGCGTTTCGCGGTTTTTTCCGGAGCGGTTGGAATCTTT
TGACTTCGTTATCGTTGCCATCGCGCTCATGCCCGCACAGGGTTCTTTGTCTGTGTTGAGG

FIG. 17 cont'd

ACATTCCGAATATTTTCGCGTGATGCGCTTGGTATCCGTGATCCCTACGATGCGCCGCGTTCG
TACAAGGAATGTTGCTGGCTCTCCCCGGCGTCGGGAGCGTTGCTGCCCTCCTTACCGTGG
TATTTTACATAGCGGCGGTTATGGCTACTAATCTTTACGGAGCTACCTTCCCGGAGTGGTT
CGGGGATTTGTCCAAGAGCCTCTATACATTGTTTCAAGTTATGACCCTGGAGTCCTGGTCT
ATGGGCATTGTCCGGCCCGTAATGAACGTACACCCAAATGCGTGCGGTGTTTTTCATTCCAT
TCATCATGCTGACTACCTTTACCGTGTGAACTTGTTTCATTGGGATTATCGTGGATGCGATG
GCCATCACTAAGGAGCAAGAAGAAGAGGCTAAAACTGGCCACCACCAAGAGCCAATTTCT
CAAACCCTCTTGCACTCTCGGGGACCGACTGGACCGCATTGAGAAGCAACTCGCGCAGAAC
AATGAGCTGTTGCAGCGACAGCAACCTCAAAAAAATAA (SEQ ID NO: 18)

NavMs, codon optimized, with N-terminal His tag and linker:

ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCA
TATGTCACGCAAAATCCGCGATTTAATCGAATCCAAACGCTTTCAAACGTCATCACCGCCA
TTATTGTGCTCAATGGCGCTGTGCTGGGTCTGCTGACCGATACAACCCTGTCGGCCTCCA
GCCAAAACCTGCTGGAGCGTGTGGATCAACTTTGTCTGACTATCTTTATTGTTGAAATCTCC
CTGAAAATCTACGCCTATGGCGTGCGCGGCTTTTTCCGCAGCGGCTGGAATCTGTTTGATT
TTGTGATTGTGGCCATCGCGCTTATGCCGGCCAGGGTAGCCTGTCGGTGCTGCGTACCT
TCCGTATCTTCCGCGTCATGCGCCTCGTATCGGTCATCCCAACCATGCGCCGTGTGGTGC
AAGGCATGCTCTTGGCACTGCCGGGCGTGGGCTCGGTAGCGGCACTGTTGACGGTGGTC
TTCTATATTGCGGCTGTCATGGCCACCAATCTCTACGGGGCAACCTTCCCTGAATGGTTTG
GTGATCTTAGCAAGAGCCTGTACACACTGTTTCAGGTGATGACCTTAGAGTCATGGTCTAT
GGGCATTGTGCGTCCAGTGATGAACGTTTCATCCGAACGCATGGGTTTTTTTCATCCCGTTC
ATCATGCTCACCACCTTTACCGTGTCAACCTGTTTATTGGCATTATTGTAGATGCAATGGC
AATCACCAAGGAACAGGAGGAAGAGGCCAAAACCGGTCACCATCAAGAACCTATTTCTCAA
ACTCTTCTTCATCTTGGTGATCGTCTTGATCGTATTGAAAAACAACCTTGCTCAAAATAATGAA
CTTCTTCAACGTCAACAACCTCAAAAAAATAA (SEQ ID NO: 19)

NavBp, endogenous sequence:

ATGGAAAACAATCCAGCCGAACAACAAGTTCCACCATTAGTAGCCTTAGCTCAGCGTATCG
TCTTTCATAAGGCCTTTACCCCAACTATTATTACCTTGATTATCATTAAATGCCATTATTGTAG
GCCTTGAAACATATCCTACTGTTTATCAAGGTTATAATGATTGGTTCTACGCAGCAGATTTA
GCCTTACTTTGGATTTTTACAATTGAGATTACACTGCGTTTTATCGCAGCGAGACCGACTAA
ATCTTTTTTTTAAAGCAGCTGGAAGTGGTTTATTATTAATCGTTCTTGCCGGTTCATGTCTT
TGCCGGTGCTCATTTTGTAAACGGTTCTTCGTATCCTGCGCGTTCTTCGCGTATTACGTGCC
ATTTCTGTCAATTCCTTCTCTGCGTCGTTTAGTCGATGCTTTGCTGATGACCATCCCGGCTTT
AGGAAACATTATGATCCTGATGGGAATTATTTTCTATATTTTCGCTGTGATTGGAACGATGT
TATTTGCTTCTGTAGCACCTGAGTACTTTGGTAACTTACAGCTTTCTTTATTAACATTATTCC
AAGTTGTTACACTTGAATCTTGGGCAAGCGGTGTCATGAGGCCGATTTTTGCAGAGGTTTG
GTGGTCTTGGATTTATTTTGTATCTTTATTTTAGTAGGGACATTTATTGTCTTTAACTTATTT
ATCGGTGTTATCGTTAATAACGTTGAAAAAGCAAACGAAGAAGAACTCAAATCAGAATTAGA
TGATAAAGAGGCAGATACAAAAGAAGAGCTTGCTTCTCTGCGTAATGAAGTAGCAGAGATG
AAAGACCTCATTAACAAATGCATAAACAGCAAACAAAAAAGGGTAA (SEQ ID NO: 20)

NavBp, codon optimized, with N-terminal 3x HA tag:

ATGGTTTACCCGTATGATGTCCCGGATTACGCTGGCAGCTACCCATACGATGTACCCGACT
ATGCCGGCAGTTATCCCTACGACGTCCCTGACTACGCAGAAAACAACCCAGCCGAACAGC
AAGTCCCACCCCTCGTGGCGCTCGCCCAACGCATAGTATTTTACAAGGCGTTTACGCCGA
CGATAATCACCCCTCATCATTATTAATGCGATCATTGTGGGACTCGAGACATACCCAACGGTT
TACCAGGGTTACAATGATTGGTTCTATGCTGCCGACCTTGCTTTGTTGTGGATATTTCACTAT

77/104

FIG. 17 cont'd

TGAAATCACGCTCCGATTCATCGCCGCCGACCGACGAAGAGTTTCTTCAAGTCTAGCTGG
AACTGGTTTGATCTGCTTATCGTATTGGCGGGCCACGTCTTCGCTGGCGCCCATTTTGTTA
CGGTGCTTAGGATCCTCCGCGTCCTGAGGGTCCTCAGAGCTATCTCAGTCATACCCAGTC
TCCGGCGGCTGGTTGACGCACTTTTGATGACAATCCCAGCACTCGGTAACATCATGATACT
GATGGGGATTATTTTTTACATATTCGCGGTTATCGGGACGATGCTCTTTGCATCAGTAGCG
CCAGAATACTTTGGCAATTTGCAGCTGTCTCTGCTTACACTGTTCCAAGTGGTTACGCTGG
AAAGTTGGGCTAGTGGGGTTATGCGACCTATTTTTGCCGAAGTCTGGTGGTCTTGGATCTA
TTTTGTAATCTTTATTCTCGTGGGAACCTTTCATAGTATTTAACCTTTTCATTGGCGTCATCGT
GAACAATGTGAAAAAGCTAACGAAGAGGAACTGAAAAGCGAACTGGATGATAAAGAGGC
TGATACAAAAGAAGAACTGGCATCATTGCGAAACGAGGTGGCAGAAATGAAGGATCTCATA
AAACAGATGCATAAACAGCAAACAAAAAAGGGTTAA (SEQ ID NO: 21)

NavSheP-D60N, codon optimized, with N-terminal 3x HA tag:

ATGGTTTACCCGTATGATGTCCCGATTACGCTGGCAGCTACCCATACGATGTACCCGACT
ATGCCGGCAGTTATCCCTACGACGTCCCTGACTACGCATCTACGTCCCTTTTGAATGCGCC
TACCGGCCTTCAAGCTAGAGTCATTAATCTCGTCGAACAAAACCTGGTTTGGACACTTTATAC
TGA CTCTCATACTCATTAATGCTGTGCAGCTTGGAATGGAACTAGCGCCAGCCTCATGGC
ACAATATGGCGCGCTGCTTATGTCCTTGAATAAGGTCCTTCTCTCTGTGTTCTGTGTCGAA
CTGCTGCTCCGGATTTATGCGTATCGGGGCAAGTTTTTTAAGGACCCGTGGAATGTGTTTG
ACTTCACTGTTATTGTTATTGCTCTGATTCTGTCATCTGGCCCATTTGGCTGTCCTCCGCTCC
CTCCGAGTTCTCCGCGTCTTGAGGGTTCTGACGATTGTCCCAGCATGAAAAGAGTAGTGT
CAGCACTGCTTGGGAGCTTGCCCGGGTTGGCCTCCATTGCAACCGTGCTTCTGTTGATCT
ATTACGTTTTTCGCTGTGATCGCCACTAAAATTTTCGGGGATGCTTTTCCGGAATGGTTCGG
GACGATAGCGGACTCCTTCTATACCCTTTTTCAAATTATGACCTTGGAAGTTGGTCTATGG
GGATCTCTAGGCCAGTGATGGAGGTGTACCCTTACGCTTGGGTATTCTTTGTGCCCTTTAT
TCTTGTTGCTACTTTTACCATGCTTAACCTTTTCATCGCCATCATAGTGAATACTATGCAGAC
ATTCTCTGACGAGGAACATGCTCTGGAGCGAGAGCAAGATAAACAGATCTTGGAACAGGA
GCAGAGACAAATGCACGAGGAACTGAAGGCCATTGCACTCGAGCTTCAGCAACTCCAAAC
CCTTTTGCAGAAATGCGGCTGGGGACTCCTCCAATGTCTCCACAAAGGGCAATATCGGCTC
AGACTAA (SEQ ID NO: 22)

NavSheP endogenous sequence:

ATGAGTACATCTTTACTTAAACGCGCCAACGGGTTTGCAGGCACGAGTGATTAACCTTGGTTG
AGCAAACTGGTTTGGTCATTTTATTTTACATTGATTTTAATCAACGCGGTGCAGTTAGGT
ATGGAGACCTCAGCCAGCCTGATGGCGCAATACGGTGCTTTGTTGATGAGTCTTGATAAG
GTGCTGCTGAGTGTATTTGTGGTGGAGTTATTGCTGCGGATTTATGCCTACAGGGGGAAAT
TTTTTAAAGACCCTTGGAACGTGTTGATTTTACCGTGATAGTGATAGCACTGATCCCTGCA
TCTGGGCCATTGGCTGTCTGCGTTGCTCAGGGTATTGCGGGTGCTGAGAGTGTTAACA
ATTGTGCCATCAATGAAACGGGTGGTGTCTGCGCTGTTGGGATCACTTCCTGGATTGGCAT
CGATCGCCACAGTATTACTGCTGATTTATTATGTGTTTGCGGTGATCGCTACCAAAATTTT
GGCGATGCATTCCCTGAATGGTTTGGCACTATTGCTGACTCATTTTATACCCTATTTCAAAT
AATGACGCTTGAAAGCTGGTCTATGGGAATTTGCGGCCAGTGATGGAAGTCTACCCTTAT
GCTTGGGTATTTTTCGTACCATTATTTCTGGTAGCGACTTTCACAATGCTAAATTTGTTTATT
GCGATTATCGTCAATACCATGCAAACCTTCAGCGACGAAGAGCATGCATTAGAGCGTGAGC
AAGACAAACAAATCTTAGAGCAGGAACAAAGACAAATGCACGAGGAGTTGAAAGCCATCAG
ACTCGAGCTACAACAATTACAAACCTTGCTGCGCAATGCTGCTGGTGATTCTTCTAATGTGT
CGACAAAGGGAAACATTGGTTCTGACTAA (SEQ ID NO: 23)

WPRES3:

FIG. 17 cont'd

ATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACATATGTTGCTC
CTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATG
GCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGC
CGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGG
(SEQ ID NO: 24)

BGHpA:

CGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGAC
CCTGGAAGGTGCCACTCCCACTGTCTTTCTAATAAAATGAGGAAATTGCATCGCATTGT
CTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGA
TTGGGAAGACAATAGCAGGCATG (SEQ ID NO: 25)

P2A Encoding Sequence:

GGCAGCGGCGCCACCAACTTCAGCCTGCTGAAGCAGGCCGGCGACGTGGAGGAGAACCC
CGGCCCCGGAGCTAGCGGA (SEQ ID NO: 26)

P2A:

(GSG)ATNFSLLKQAGDVEENPGPGASG (SEQ ID NO: 27)

T2A:

(GSG)EGRGSLLTCGDVEENPGP (SEQ ID NO: 28)

E2A:

(GSG)QCTNYALLKLAGDVESNPGPP (SEQ ID NO: 29)

F2A:

(GSG)VKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 30)

N-terminal 3XHA tag (DNA):

ATGGTTTACCCGTATGATGTCCCGGATTACGCTGGCAGCTACCCATACGATGTACCCGACT
ATGCCGGCAGTTATCCCTACGACGTCCCTGACTACGCA (SEQ ID NO: 31)

N-terminal 3XHA tag (Protein):

MVYPYDVPDYAGSYPYDVPDYAGSYPYDVPDYA (SEQ ID NO: 32)

PHP.eB capsid:

AAV9 capsid except that amino acids starting at residue 586: SAQA (SEQ ID NO: 98) are changed to SDGTLAVPFA (SEQ ID NO: 33) to form

MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLPGNGLDK
GEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKK
RLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTEVPDP
QPIGEPPAAPSGVGS LTMASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGD RVITTSTRT
WALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDNRFHCHFS PRDWQRLINNNWG
FRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADV
FMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGN NFQFSYEFENVPFHSSYAHSSQLD
RLMNPLIDQYLYLSKTINGSGQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQN
NNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADK
VMITNEEEIKTTNPVATESYGQVATNHQSDGTLAVPFAQAQGTGWVQNQGILPGMVWQDRDV
YLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPQILIKNTVPADPPTAFNKDKLNSFITQYST

FIG. 17 cont'd

GQVSVEIEWELQKENS KRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL
(SEQ ID NO: 90)

AAV9 VP1 capsid protein:

MAADGYLPDWLEDNLSEGIREW WALKPGAPQPKANQQHQDNARGLVLPGYKYLGPNGLDK
GEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADA E FQERLKEDTSFGGNLGRAV FQAKK
RLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGD TESVDPD
QPIGEPPAAPSGVGS LTMASGGGAPVADNNEGADGVGSSSGNWHCD SQWLGD RVITTSTRT
WALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDNRFHCHFS PRDWQRLINNNWG
FRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADV
FMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAH SQSLD
RLMNPLIDQYLYLSKTINGSGQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQN
NNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADK
VMITNEEEIKTTNPVATESY GQVATNHQSAQAQAQTGWVQNQGILPGMVWQDRDVYLQGP IW
AKIPHTDGNFHPSPLMGGFGMKHPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQVSVEIE
WELQKENS KRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL (SEQ ID NO:
34)

tet-Transactivator version 2 (tTA2):

ATGTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATGAAGTCG
GTATCGAAGGCCTGACGACAAGGAACTCGCTCAAAGCTGGGAGTTGAGCAGCCTACCC
TGTA CTGGCACGTGAAGAACAAAGCGGGCCCTGCTCGATGCCCTGGCAATCGAGATGCTGG
ACAGGCATCATACCCACTTCTGCCCCCTGGAAGGCGAGTCATGGCAAGACTTTCTGCGGA
ACAACGCCAAGTCATTCCGCTGTGCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCATCT
CGGCACCCGCCCAACAGAGAAACAGTACGAAACCCTGGAAAATCAGCTCGCGTTTCTGTG
TCAGCAAGGCTTCTCCCTGGAGAACGCACTGTACGCTCTGTCCGCCGTGGGCCACTTTAC
ACTGGGCTGCGTATTGGAGGATCAGGAGCATCAAGTAGCAAAAGAGGAAAGAGAGACACC
TACCACCGATTCTATGCCCCCACTTCTGAGACAAGCAATTGAGCTGTTGACCATCAGGGA
GCCGAACCTGCCTTCTTTTGGCCTGGA ACTAATCATATGTGGCCTGGAGAAACAGCTAA
AGTGCGAAAGCGGCGGGCCGGCCGACGCCCTTGACGATTTTGACTTAGACATGCTCCCAG
CCGATGCCCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTGA
CCTTGACATGCTCCCCGGGTAA (SEQ ID NO: 35)

CN1367 - The portion between L-ITR and R-ITR: positions 142-2984:

GCGGCCGCACGCGTATAGGTACCGAGCTCTATGCACTCACAGTGGTTTGGCATGCATCTG
GTGAATTTTTTTTAAACGAAAAATTAGTGTTGGTTTCGATGTATGGTAGCATTCTCCCTAACGT
AATTTGAATAATTCAGCAAAGCCCCACTACCAGCTGTACTTCTGCAGCCTCTTCCATTCTTT
TCAGCATTATAATTTTGGTTAATTTTCAATTTTAGGTCCTACGTCTCTGCAATTTGTGTATGA
ATAACAGAATAATTTCCCTCTTTTGTTCGCCTTTCCTGTTCTCTGAATCTAAATAAAGATGGC
TTTTTAGTATTAAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAA
TCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGAAAAAACCC
TTCTTGTGGATAAAACCTTAAATTGTCCCAATGTCTGCTTCAAATTGGATGGCACTGCAG
CTGGAGGCTTTGTT CAGAATTGATCCTGGGGAGCTACGAACCCAAAGTTT CACAGTAGGGA
GCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGGGAT
CCAGATCTTTCGAAGCTAGCGCTACCGGTCGCCACCATGGGCAGCAGCCATCATCATCAT
CATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGTCACGCAAAATCCGCGATTTA
ATCGAATCCAAACGCTTTCAAACGTCATCACCGCCATTATTGTGCTCAATGGCGCTGTGC
TGGGTCTGCTGACCGATACAACCCTGTCGGCCTCCAGCCAAACCTGCTGGAGCGTGTGG

FIG. 17 cont'd

ATCAACTTTGTCTGACTATCTTTATTGTTGAAATCTCCCTGAAAATCTACGCCTATGGCGTG
CGCGGCTTTTTCCGCAGCGGCTGGAATCTGTTTGATTTTGTGATTGTGGCCATCGCGCTTA
TGCCGGCCCAGGGTAGCCTGTCGGTGCTGCGTACCTTCCGTATCTTCCGCGTCATGCGCC
TCGTATCGGTCATCCCAACCATGCGCCGTGTGGTGCAAGGCATGCTCTTGGCACTGCCGG
GCGTGGGCTCGGTAGCGGCACTGTTGACGGTGGTCTTCTATATTGCGGCTGTCATGGCCA
CCAATCTCTACGGGGCAACCTTCCCTGAATGGTTTGGTGATCTTAGCAAGAGCCTGTACAC
ACTGTTTCAGGTGATGACCTTAGAGTCATGGTCTATGGGCATTGTGCGTCCAGTGATGAAC
GTTTCATCCGAACGCATGGGTTTTTTTCATCCCGTTTCATCATGCTCACCACCTTTACCGTGCT
CAACCTGTTTATTGGCATTATTGTAGATGCAATGGCAATCACCAAGGAACAGGAGGAAGAG
GCCAAAACCGGTCACCATCAAGAACCTATTTCTCAAACCTCTTCTTCATCTTGGTGATCGTCT
TGATCGTATTGAAAAACAACCTTGCTCAAAATAATGAACTTCTTCAACGTCAACAACCTCAAA
AAAAAGGCAGCGGCGCCACCAACTTCAGCCTGCTGAAGCAGGCCGGCGACGTGGAGGAG
AACCCCGGCCCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCT
GGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGGCGAG
GGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCC
CGTGCCCTGGCCACCCTCGTGACCACCCTGGGCTACGGCGTGCAAGTGTCTTCGCCCGCT
ACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCC
AGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGT
TCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGAC
GGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCACC
GCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGAC
GGCGGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGT
GCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGA
GAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGGATCACTCTCGGCAT
GGACGAGCTGTACAAGTAAGTCGACGGCGCGCCGCGGCGCGAATTCGATATCATAATCA
ACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAAGTATGTTGCTCCTTTTAC
GCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCA
TTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTG
CCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGCTCGAG
AGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTT
CCTTGACCCTGGAAGGTGCCACTCCCCTGTCTTTTCTAATAAAATGAGGAAATTGCATC
GCATTGTCTGAGTAGGTGTCACTTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGG
GGGAGGATTGGGAAGACAATAGCAGGCATGCACGTGCGGACCGAGCGGCCGC (SEQ ID
NO: 36)

CN1500 - The portion between L-ITR and R-ITR: positions 142-2976:

GCGGCCGCACGCGTGGTACCCTAAATAAAGATGGCTTTTTAGTATTAAGTGAAGAAAA
TTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACT
CTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAAGTGG
AAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAA
AATTACTCTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAA
AAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTA
CATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGAGCTCGGGCTGGTCGACAC
AATTGGAGGTAGGCGTGTACGGTGGGAGGCCTATATAAGCAGAGCTCGTTTAGTGAACCG
TCAGATCGCCTGGAGGATCCTTCGAAAAGCTTGCTACCGGTCGCCACCATGGTCAGCAAG
GGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTCAA
TGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGA
CCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCTCGTGACC
ACCCTGGGCTACGGCGTGCAAGTGTCTCGCCCGCTACCCCGACCACATGAAGCAGCACGA

FIG. 17 cont'd

CTTCTTCAAGTCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAAGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAA
GGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGCAGCTCGCCGACCACT
ACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTG
AGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTG
GAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAAGGCAGCGG
CGCCACCAACTTCAGCCTGCTGAAGCAGGCCGGCGACGTGGAGGAGAACCCCGGCCCCG
GAGCTAGCGGAATGGTTTACCCGTATGATGTCCCGGATTACGCTGGCAGCTACCCATACG
ATGTACCCGACTATGCCGGCAGTTATCCCTACGACGTCCCTGACTACGCATCTACGTCCCT
TTTGAATGCGCCTACCGGCCTTCAAGCTAGAGTCATTAATCTCGTGAACAAAAGTGGTTT
GGACACTTTATACTGACTCTCATACTCATTAAATGCTGTGCAGCTTGGAATGGAACTAGCG
CCAGCCTCATGGCACAATATGGCGCGCTGCTTATGTCCTTGAATAAGGTCCTTCTCTCTGT
GTTCTGTGTCGAAGTGTGCTCCGGATTTATGCGTATCGGGGCAAGTTTTTAAAGGACCCG
TGGAATGTGTTTGACTTCACTGTTATTGTTATTGCTCTGATTCTGCATCTGGCCCATTTGGC
TGTCCTCCGCTCCCTCCGAGTTCTCCGCGTCTTGAGGGTTCTGACGATTGTCCCCAGCAT
GAAAAGAGTAGTGTGAGCACTGCTTGGGAGCTTGCCCGGGTTGGCCTCCATTGCAACCGT
GCTTCTGTTGATCTATTACGTTTTCTGCTGTGATCGCCACTAAAATTTTCGGGGATGCTTTTC
CGGAATGGTTCGGGACGATAGCGGACTCCTTCTATACCCTTTTTCAAATTATGACCTTGGA
AAGTTGGTCTATGGGGATCTCTAGGCCAGTGATGGAGGTGTACCCTTACGCTTGGGTATTCT
TTTGTGCCCTTTATTCTTGTGCTACTTTTACCATGCTTAACCTTTTCATCGCCATCATAGTG
AATACTATGCAGACATTCTCTGACGAGGAACATGCTCTGGAGCGAGAGCAAGATAAACAGA
TCTTGGAACAGGAGCAGAGACAAATGCACGAGGAACTGAAGGCCATTCTGACTCGAGCTTC
AGCAACTCCAAACCCTTTTGCGAATGCGGCTGGGGACTCCTCCAATGTCTCCACAAAGG
GCAATATCGGCTCAGACTAATGACCGCGGCCGCGAATTCGATATCATAATCAACCTCTGGA
TTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACCTATGTTGCTCCTTTTACGCTATGTGG
ATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTTCATTTTCTCCTC
CTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACCTCATCGCCGCTGCTTGGCCGCG
TGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGCTCGAGAGATCTTCGAC
TGTGCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCTTACCCCTG
GAAGGTGCCACTCCCCTGCTCTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGA
GTAGGTGTCACTTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGG
GAAGACAATAGCAGGCATGAGATCTCACGTGCGGACCGAGCGGCCGC (SEQ ID NO: 37)

CN1498 - The portion between L-ITR and R-ITR: positions 142-2943:

GCGGCCGCACGCGTGGTACCCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAA
TTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACT
CTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGG
AAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAA
AATTACTCTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAA
AAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTA
CATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGAGCTCGGGCTGGTCGACAC
AATTGGAGGTAGGCGTGTACGGTGGGAGGCCTATATAAGCAGAGCTCGTTTAGTGAACCG
TCAGATCGCCTGGAGGATCCTTCGAAAAGCTTGCTACCGGTCGCCACCATGGTCAGCAAG
GGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTCAA
TGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGA
CCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACC
ACCCTGGGCTACGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGA

FIG. 17 cont'd

CTTCTTCAAGTCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAAGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAA
GGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGCAGCTCGCCGACCACT
ACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTG
AGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTG
GAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAAGGCAGCGG
CGCCACCAACTTCAGCCTGCTGAAGCAGGCGGCGACGTGGAGGAGAACCCCGGCCCCG
GAGCTAGCGGAATGGTTTACCCGTATGATGTCCCGGATTACGCTGGCAGCTACCCATACG
ATGTACCCGACTATGCCGGCAGTTATCCCTACGACGTCCCTGACTACGCAGAAAACAACCC
AGCCGAACAGCAAGTCCCACCCCTCGTGCGCTCGCCCAACGCATAGTATTTTACAAGGC
GTTTACGCCGACGATAATCACCCCTCATCATTATTAATGCGATCATTGTGGGACTCGAGACAT
ACCCAACGGTTTACCAGGGTTACAATGATTGGTTCTATGCTGCCGACCTTGCTTTGTTGTG
GATATTCATCTATTGAAATCACGCTCCGATTCATCGCCGCCCGACCGACGAAGAGTTTCTTC
AAGTCTAGCTGGAAGTGGTTTGATCTGCTTATCGTATTGGCGGGCCACGTCTTCGCTGGCG
CCCATTTTGTACGGTGCTTAGGATCCTCCGCGTCTTGAGGGTCTCAGAGCTATCTCAGT
CATACCCAGTCTCCGGCGGCTGTTTGACGCACTTTTGATGACAATCCCAGCACTCGGTAA
CATCATGATACTGATGGGGATTATTTTTTACATATTCGCGGTTATCGGGACGATGCTCTTTG
CATCAGTAGCGCCAGAATACTTTGGCAATTTGCAGCTGTCTCTGCTTACACTGTTCCAAGT
GGTTACGCTGGAAAGTTGGGCTAGTGGGGTTATGCGACCTATTTTTGCCGAAGTCTGGTG
GTCTTGATCTATTTTGTAACTTTTCTCGTGGAACTTTCATAGTATTTAACCTTTTCATT
GGCGTCATCGTGAACAATGTGGAAAAAGCTAACGAAGAGGAACTGAAAAGCGAACTGGAT
GATAAAGAGGCTGATACAAAAGAAGAACTGGCATCATTGCGAAACGAGGTGGCAGAAATG
AAGGATCTCATAAACAGATGCATAAACAGCAAAACAAAAAAGGGTTAATGACCGCGGCCGC
GAATTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTA
ACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATT
GCTTCCCGTATGGCTTTTCATTTTCTCCTCCTTGATATAAATCCTGGTTAGTTCTTGCCACGGC
GGAATCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTG
ACAATTCGCTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTT
GCCCCCTCCCCCGTGCCTTCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCTTAATA
AAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCAATTCTATTCTGGGGGGTGGGGTG
GGGCAGGACAGCAAGGGGGGAGGATTGGGAAGACAATAGCAGGCATGAGATCTCACGTGC
GGACCGAGCGGCCGC (SEQ ID NO: 38)

CN1499 - The portion between L-ITR and R-ITR: positions 142-2946:

GCGGCCGCACGCGTGGTACCCTAAATAAAGATGGCTTTTTAGTATTAAAAAGTGGAAGAAAA
TTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACT
CTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAAAAAGTGG
AAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAA
AATTACTCTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAA
AAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTA
CATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGAGCTCGGGCTGGTCGACAC
AATTGGAGGTAGGCGTGTACGGTGGGAGGCCTATATAAGCAGAGCTCGTTTAGTGAACCG
TCAGATCGCCTGGAGGATCCTTCGAAAAGCTTGCTACCGGTCGCCACCATGGTCAGCAAG
GGCGAGGAGCTGTTACCCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTCAA
TGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGA
CCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACC
ACCCTGGGCTACGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGA

FIG. 17 cont'd

CTTCTTCAAGTCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAAGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAA
GGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGCAGCTCGCCGACCACT
ACCAGCAGAACACCCCCATCGGCCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTG
AGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTG
GAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAAGGCAGCGG
CGCCACCAACTTCAGCCTGCTGAAGCAGGCGGCGACGTGGAGGAGAACCCCGGCCCCG
GAGCTAGCGGAATGGTTTATCCGTATGATGTTCCCTGACTATGCAGGATCCTATCCTTATGAT
GTTCCCGATTACGCTGGTTCTTACCCTTACGATGTTCCCGATTATGCCAGTTCTGGATTGGT
GCCACGAGGCAGCCACATGAGCCGGAAGATCAGAGATCTTATCGAATCTAAGAGATTTC
GAATGTTATTACCGCGATAATCGTACTCAACGGGGCGGTGCTCGGTCTCCTCACCATAACC
ACATTGAGCGCTTCTAGCCAGAACCTGCTCGAAAGGGTTGACCAACTGTGCCTGACAATTT
TTATCGTGGAATTAGCTTGAAAATTTACGCCTACGGCGTTTCGCGGTTTTTTCCGGAGCGG
TTGGAATCTTTTTGACTTCGTTATCGTTGCCATCGCGCTCATGCCCGCACAGGGTTCTTTGT
CTGTGTTGAGGACATTCCGAATATTTTCGCGTGATGCGCTTGGTATCCGTGATCCCTACGAT
GCGCCGCGTCTGACAAGGAATGTTGCTGGCTCTCCCCGGCGTCGGGAGCGTTGCTGCC
TCCTTACCCTGGTATTTTACATAGCGCGGTTATGGCTACTAATCTTTACGGAGCTACCTTC
CCGGAGTGGTTCGGGGATTTGTCCAAGAGCCTCTATACATTGTTTCAAGTTATGACCCTGG
AGTCCTGGTCTATGGGCATTGTCCGGCCCGTAATGAACGTACACCCAAATGCGTGGGTGT
TTTTCAATCCATTCATCATGCTGACTACCTTTACCGTGCTGAACTTGTTTCATTGGGATTATCG
TGGATGCGATGGCCATCACTAAGGAGCAAGAAGAAGAGGCTAAACTGGCCACCACCAAG
AGCCAATTTCTCAAACCTCTTGCATCTCGGGGACCGACTGGACCGCATTGAGAAGCAACT
CGCGCAGAACAAATGAGCTGTTGCAGCGACAGCAACCTCAAAAAAATAATGACCGCGGCC
GCGAATTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTC
TTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTA
TTGCTTCCCGTATGGCTTTTCAATTTCTCCTCCTTGATAAATCCTGGTTAGTTCTTGCCACG
GCGGAACTCATCGCCGCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCAC
TGACAATTCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGT
TTGCCCTCCCCCGTGCCTTCTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCTAA
TAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGG
TGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGAGATCTCACGT
GCGGACCGAGCGGCCGC (SEQ ID NO: 39)

CN1244 - The portion between L-ITR and R-ITR: positions 142-2042:

GCGGCCGCACGCGTATAGGTACCGAGCTCTATGCACTCACAGTGGTTTGGCATGCATCTG
GTGAATTTTTTTTAAACGAAAAATTAGTGTTGGTTTCGATGTATGGTAGCATTCTCCCTAACGT
AATTTGAATAATTCAGCAAAGCCCCACTACCAGCTGTACTTCTGCAGCCTCTTCCATTCTTT
TCAGCATTATAATTTTGGTTAATTTTCAATTTTAGGTCCTACGTCTCTGCAATTTGTGTATGA
ATAACAGAATAATTTCCCTCTTTTGTTCGCCTTTCCTGTTTCTGAATCTAAATAAAGATGGC
TTTTTAGTATTAAGAGTGGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAA
TCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGAAAAAACCC
TTCTTGTGGATAAAAACTTAAATTGTCCCAATGTCTGCTTCAAATTGGATGGCACTGCAG
CTGGAGGCTTTGTTTCAGAATTGATCCTGGGGAGCTACGAACCCAAAGTTTTCACAGTAGGGA
GCTCGGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGGGAT
CCAGATCTTTCGAAGCTAGCGCTACCGGTGCGCCACCATGGTGAGCAAGGGCGAGGAGCT
GTTACCGGGGGTGGTGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGT
TCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTG

FIG. 17 cont'd

ATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGGGCTA
CGGCGTGCAAGTGTTCGCCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTC
CGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTA
CAAGACCCGCGCCGAGGTGAAGTTCGAGGGGCGACACCCTGGTGAACCGCATCGAGCTGA
AGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACA
ACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCA
AGATCCGCCACAACATCGAGGACGGCGGCGTGAGCTCGCCGACCACTACCAGCAGAAC
ACCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCACTG
CAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGAC
CGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACGGCGCGCCGC
GGCCGCGAATTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGT
ATTCTTAAGTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCAT
GCTATTGCTTCCCGTATGGCTTTTCAATTTCTCCTCCTTGATAAATCCTGGTTAGTTCTTGCC
ACGGCGGAACATCGCCGCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGG
CACTGACAATTCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGT
TGTTTGCCCCCTCCCCGTGCCTTCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC
TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTG
GGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGAGATCTCA
CGTGCGGACCGAGCGGCCGC (SEQ ID NO: 40)

CN1389 - The portion between L-ITR and R-ITR corresponds to positions 142-1897:

CCTGCAGGCAGCTGCGCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCG
TCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTG
GCCAACTCCATCACTAGGGGTTCTGCGGCCGCACGCGTTTCGCTTTCTGTTCTGAAT
CTAAATAAAGATGGCTTTTTAGTATTAAGTGAAGAAAATTACAGGTAATTATCTTTGACG
GTAAAAACGCTGTAAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGA
GAATGGACCTGCAGGGAGCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCT
TACATTTGCTTCTGGGATCCAGATCTTTCGAAGCTAGCGCTACCGGTGCGCCACCATGGTGA
GCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGAC
GTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAA
GCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCG
TGACCACCTGGGCTACGGCGTGAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAG
CACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTC
AAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGT
GAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACA
AGCTGGAGTACAACACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAAG
GCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGAGCTCGCC
GACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCA
CTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGT
CCTGCTGGAGTTCGTGACCGCCGCGGGGATCACTCTCGGCATGGACGAGCTGTACAAGTA
AGTCGACGGCGCGCCGCGGCCGCGAATTCGATATCATAATCAACCTCTGGATTACAAAATT
TGTGAAAGATTGACTGGTATTCTTAAGTATGTTGCTCCTTTTACGCTATGTGGATACGCTGC
TTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCAATTTCTCCTCCTTGATAA
ATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCTGCCTTGCCCGCTGCTGGAC
AGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGCTCGAGAGATCTTCGACTGTGCCTT
CTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCGTGCCTTCTTGACCCTGGAAGGTGC
CACTCCCACTGTCCTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTC
ATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAAT
AGCAGGCATGAGATCTCACGTGCGGACCGAGCGGCCGCAGGAACCCCTAGTGATGGAGT

FIG. 17 cont'd

TGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCC
CGACGCCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGC
AGGGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATACG
TCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGTGTGGTGGTT
ACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTC
CCTTCCTTTCTCGCCACGTTTCGCCGGCTTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTT
TAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCCAAAAAAGTTGATTTGGGTGATGG
TTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACG
TTCTTTAATAGTGGACTCTTGTTCCAACTGGAACAACACTCAACCCTATCTCGGGCTATTC
TTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACA
AAAATTTAACGCGAATTTTAACAAAAATATTAACGTTTACAATTTTATGGTGCCTCTCAGTAC
AATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGC
GCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGG
GAGCTGCATGTGTGTCAGAGGTTTTACCGTCAACACCGAAACGCGCGAGACGAAAGGGCCT
CGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTG
GCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAAT
ATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAG
TATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCTTCCTGT
TTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGA
GTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAG
AACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATT
GACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTGAG
TACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTG
CTGCCATAACCATGAGTGATAAAGCTGCGGCCAACTTACTTCTGACAACGATCGGAGGACC
GAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAAGTTCGCCTTGATCGTTGG
GAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCA
ATGGCAACAACGTTGCGCAAACTATTAAGTGGCGAACTACTTACTCTAGCTTCCCGGCAAC
AATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTC
CGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCA
TTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGA
GTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAA
GCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACCTTCATTT
TTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACG
TGAGTTTTCTGTTCCACTGAGCGTCAGACCCCGTAGAAAAAGATCAAAGGATCTTCTTGAGAT
CCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAACAAAAAACCACCGCTACCAGCGGTGG
TTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGC
GCAGATACCAAATACTGTCCTTCTAGTGAGCCGTAGTTAGGCCACCACTTCAAGAACTCT
GTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAAGTGGCTGCTGCCAGTGGC
GATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGG
TCGGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGA
ACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGC
GGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAG
GGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGTTTTGCCACCTCTGACTTGAGCGTC
GATTTTTGTGATGCTCGTCAGGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCT
TTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGT (SEQ ID NO: 41)

CN1390 - The portion between L-ITR and R-ITR corresponds to positions 142-1660:

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCCGGGCG
TCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTG

FIG. 17 cont'd

GCCAACTCCATCACTAGGGGTTCTCGCGCCGCACGCGTGGTACCCTAAATAAAGATGGC
TTTTTAGTATTAAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAA
TCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAA
GATGGCTTTTTAGTATTAAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAAC
GCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGCT
AAATAAAGATGGCTTTTTAGTATTAAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGT
AAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGA
ATGGAGCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCT
GGGATCCAGATCTTTCGAAGCTAGCGCTACCGGTGCGCCACCATGGTGAGCAAGGGCGAG
GAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCA
CAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGA
AGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTG
GGCTACGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTC
AAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGC
AACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGA
GCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACA
ACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCA
ACTTCAAGATCCGCCACAACATCGAGGACGGCGGGCTGCAGCTCGCCGACCACTACCAG
CAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTA
CCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTT
CGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACGGCG
CGCCGCGGCCGCGAATTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATT
GACTGGTATTCTTAACATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTT
TGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAG
TTCTTGCCACGGCGGAACATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGG
CTGTTGGGCACTGACAATTCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAG
CCATCTGTTGTTTGCCCCCTCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTG
TCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTG
GGGGGTGGGGTGGGGCAGGACAGCAAGGGGGGAGGATTGGGAAGACAATAGCAGGCATG
AGATCTCACGTGCGGACCGAGCGGCCGAGGAACCCCTAGTGATGGAGTTGGCCACTCC
CTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGTCCGCCGACGCCCGG
GCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGGGGCGCCT
GATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATACGTCAAAGCAAC
CATAGTACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGC
GTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTGCTTTCTTCCCTTCTTTT
TCGCCACGTTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCC
GATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTTGGGTGATGGTTCACGTAG
TGGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAAT
AGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGGCTATTCTTTTGATTT
ATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTA
ACGCGAATTTTAACAAAATATTAACGTTTACAATTTTATGGTGCACTCTCAGTACAATCTGCT
CTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGA
CGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGC
ATGTGTCAGAGGTTTTACCGTTCATACCGAAACGCGCGAGACGAAAGGGCCTCGTGATA
CGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTT
CGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCC
GCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTA
TTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTTCGGGCATTTTGCCTTCTGTTTTTGCTC
ACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTT

FIG. 17 cont'd

ACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTT
TCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCC
GGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGTTGAGTACTCAC
CAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCAT
AACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGA
GCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCG
GAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCA
ACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAAT
AGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTG
GCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAG
CACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTG
GTAACGTGCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAACTTTCATTTTTAATTT
AAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAATCCCTTAACGTGAGTT
TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTT
TTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGT
GCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATA
CCAAATACTGTCCTTCTAGTGAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAC
CGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTC
GTGTCTTACCGGGTTGGAAGCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCCGGGCTG
AACGGGGGGTTCGTGCACACAGCCAGCTTGAGCGAACGACCTACACCGAACTGAGATA
CCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGT
ATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAA
CGCCTGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTG
TGATGCTCGTCAGGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACG
GTTCTGCGCCTTTTTGCTGGCCTTTTTGCTCACATGT (SEQ ID NO: 42)

CN1203 - The portion between L-ITR and R-ITR corresponds to positions 183-2052:

AAAGCTTCCCGGGGGGATCTGGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGG
CCGGGCGACCAAAGGTGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGA
GCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCTTGAGGGGTGGAG
TCGTGACCTAGGACGCGTATAGGTACCGAGCTCTATGCACTCACAGTGGTTTGGCATGCAT
CTGGTGAATTTTTTTTAAACGAAAAATTAGTGTTGGTTTCGATGTATGGTAGCATTCTCCCTAA
CGTAATTTGAATAATTCAGCAAAGCCCCACTACCAGCTGTACTTCTGCAGCCTCTTCCATTC
TTTTCAGCATTATAATTTTGGTTAATTTTCAATTTTAGGTCCTACGTCTCTGCAATTTGTGTAT
GAATAACAGAATAATTTCCCTCTTTTGTTCGCCTTTCTGTTCTGAATCTAAATAAAGATG
GCTTTTTAGTATTAAGAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTG
TAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGAAAAAA
ACCTTCTTGTGGATAAAAACCTTAAATTGTCCCAATGTCTGCTTCAAATTGGATGGCACTG
CAGCTGGAGGCTTTGTTTCAGAATTGATCCTGGGGAGCTACGAACCCAAAGTTTCACAGTAG
GGAGCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTG
GGATCCAGATCTTTCGAAGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAGG
AGCTGTTTACCGGGGTGGTGCCCATCCTGGTTCGAGCTGGACGGCGACGTAAACGGCCAC
AAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAA
GCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCTCGTGACCACCTGG
GCTACGGCGTGAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCA
AGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCA
ACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAG
CTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAAC

FIG. 17 cont'd

TACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAAC
TTCAAGATCCGCCACAACATCGAGGACGGCGGGCGTGCAGCTCGCCGACCACTACCAGCA
GAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCTACC
AGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCCG
TGACCGCCGCGGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACGGCGCG
CCGCGGGCCGCGAATTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGA
CTGGTATTCTTAACATATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGT
ATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTC
TTGCCACGGCGGAACATCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTG
TTGGGCACTGACAATTCCGTGGCTCGAGCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTG
TTTGGCCCTCCCCCGTGCTTCCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCCTA
ATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGG
GTGGGGCAGGACAGCAAGGGGGGAGGATTGGGAAGACAATAGCAGGCATGACTAGTCCAC
TCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGCC
CGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGACAGATCCG
GGCCCGCATGCGTCGACAATTCCTGCGCGTCTGTTTTACAACGTCGTGACTGGGAAAACC
CTGGCGTTACCCAACCTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAG
CGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGC
GCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACCGCATATGGTGCAC
TCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACC
CGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC
CGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTACCGTCATCACCGAAACGCGCGAGACG
AAAGGGCCTCGTGATACGCCTATTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGA
CGTCAGGTGGCACTTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTATTTTTCTAAATA
CATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAA
AGGAAGAGTATGAGTATTCAACATTTCCGTGTGCGCCCTTATTCCTTTTTTTCGGGCATTTTG
CCTTCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTG
GGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTG
GCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATT
ATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACTACTATTCTCAGAATGA
CTTGTTGAGTACTACCAAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA
TTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGA
TCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCGCC
TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGA
TGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTCTAGC
TTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCG
CTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTC
TCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTAC
ACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCC
TCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTA
AACTTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATGACCAA
ATCCCTTAACGTGAGTTTTTCTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT
CTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACACCGCTA
CCAGCGGTGGTTTGTGCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCT
TCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGAGCCGTAGTTAGGCCACCACTT
CAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCT
GCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGAAGACGATAGTTACCGGATAAG
GCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGAGCGAACGA
CCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAG

FIG. 17 cont'd

GGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAG
GGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCCGGGTTTCGCCACCTCTG
ACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAG
CAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCTG
CGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCG
CCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCA
ATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGG
TTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATT
AGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGG
ATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCTCTCGAGATCTAG
(SEQ ID NO: 43)

Lactase [Homo sapiens], GenBank: EAX11622.1:

MELSWHVFIALLSFSCWGSWESDRNFISTAGPLTNDLLHNLSGLLGDQSSNFVAGDKDMYV
CHQPLPTFLPEYFSSLHASQITHYKVFLSWAQLLPAGSTQNPDEKTVQCYRRLKALKTARLQP
MVILHHQTLPLASTLRREAFADLFADYATFAFHSFGDLVGIWFTFSDLEEVIKELPHQESRASQL
QTLSDAHRKAYEIYHESYAFQGGKLSVVLRAEDIPELLLEPPISALAQDVTDFLSLDLSYECQNE
ASLRQKLSKLQTIIEPKVKVIFNLKLPDCPSTMKNPASLLFSLFEAINKDQVLTIGFDINEFLSCSS
SSKKSMSCSLTGSLALQPDQQQDHETDSSPASAYQRVWEAFANQSRAERDAFLQDTFPEGF
LWGASTGAFNVEGGWAEGGRGVSIWDPRRPLNTTEGQATLEVASDSYHKVASDVALLCGLR
AQVYKFSISWSRIFPMGHGSSPSLPGVAYYNKLIDRLQDAGIEPMATLFHWDLPQALQDHGGW
QNESVVDADFIDYAAFCSTFGDRVKLWTFHEPWMSYAGYGTGQHPPGISDPGVASFVAH
LVLKAHARTWHHYNSHHRPQQQGHVGIVLNSDWAEPSPERPEDLRASERFLHFMLGWFAH
PVFVDGDYPATLRTQIQQMNRQCSHPVAQLPEFTEAEKQLLKGSADFLGLSHYTSRLISNAPQ
NTCIPSYDTIGGFSQHVNHWVPQTSSSWIRVVPWGIRLLQFVSLEYTRGKVPYILAGNGMPIG
ESENLFDDSLRVDYFNQYINEVLKAIKEDSVDVRSYIARSLIDGFEGPSGYSQRFGLHHVNFSD
SSKSRTPRKSAYFFTSIIEKNGFLT KGAKRLLPPNTVNLPSKVRAFTFPSEVPSKAKVWWEKFSS
QPKFERDLFYHGTFRDDFLWGVSSSAYQIEGAWDADGKGPSIWDNFTHTPGSNVKDNATGDI
ACDSYHQLDADLNLMLRALVKAYRFSISWSRIFPTGRNSSINSHGVDDYNNRLINGLVASNIFPM
VTLFHWDLQPALQDIGGWENPALIDLFDSDYADFCFQTFGDRVKFWMTFNEPMYLAWLGYGSG
EFPPGVKDPGWAPYRIAHAVIKAHARVYHTYDEKYRQEKGVISLSLSTHWAEPKSPGVPRDV
EAADRMLQFSLGWFAHPIFRNGDYPDTMKWKVGNRSELQHLATSRLPSFTEEEKRFIRATADV
FCLNTYYSRIVQHKTPLRNPPSYEDDQEMAEEDPSWPSTAMNRAAPWGTRLLNWIKEEYG
DIPIYITENGVLNTPNTEDTDRIFYHKTYINEALKAYRLDGLRGLGYVAWSLMDNFEWLNGYTV
KFGLYHVDNFNTNRPRATARASARYYTEVITNNGMPLAREDEFYGRFPEGFIWSAASAAAYQIE
GAWRADGKGLSIWDTFSHTPLRVENDAIGDVACDSYHKIAEDLVTLQNLGVSHYRFSISWSRIL
PDGTRYINEAGLNYYVRLIDTLLAASIQPQVTIYHWDLPQTLQDVGWENETIVQRFKEYADVL
FQRLGDKVKFWITLNEPFVIAYQGYGYGTAAPGVSNRPGTAPYIVGHNLKAHAEAHLYNDVY
RASQGGVISITISSDWAEPDRPSNQEDVEAARRYVQFMGGWFAHPIFKNGDYNEVMKTRIRDR
SLAAGLNKSRLEPTESEKRRINGTYDFFGFNHYTTVLAYNLNYATAISSFDADRGVASIADRS
WPDSGSFWLKMTPFGFRILNLWLKEEYNDPPIYVTENGVSQRETDLNDTARIYYLRITYNEAL
KAVQDKVDLRGYTVWSAMDNFEWATGFSEFGLHFVNYS DPSLPRIKASAKFYASVVRNG
FPDPATGPHACLHQPDAGPTISPVRQEEVQFLGLMLGTTEAQTALYVLFSLVLLGVCGLAFLSY
KYCKRSKQGKTQRSQQELSPVSSF (SEQ ID NO: 44)

Lipase [Homo sapiens], GenBank: AAA60129.1:

MLPLWTLSSLGAVAGKEVCYERLGCFSDDSPWSGITERPLHILPWSPKDVNTRFLLYTNENPN
NFQEVAADSSSISGSNFKTRNRKTRFIIHGFIDKGEENWLANVCKNLFKVESVNCICVDWKGGSR
TGYTQASQNIIRIVGAEVAYFVEFLQSAFGYSPSNVHVIGHSLGAHAAGEAGRRTNGTIGRITGL

FIG. 17 cont'd

DPAEPCFQGTPELVRLDPSDAKFVDVIHTDGAPIVPNLGFGMSQVVGHLDFFPNGGVEMPGC
KKNILSQIVDIDIGIWEGTRDFAACNHLRSYKYTTDSIVNPDGFAGFPCASYNVFTANKCFPCPS
GGCPQMGGHYADRYPGKTNDVGQKFYLDTGDA SNFARWRYKVS VTL SGKKVTGHILVSLFGNK
GNSKQYEIFKGTLPDSTHSNEFSDSDVDVGD LQMV KFIWYNNVINPTLPRVGASKIIVETNVGK
QNFNCSPETVREEVLLTLTPC (SEQ ID NO:45)

Helicase [Homo sapiens], GenBank: AMD82207.1:

MSISLFGGRYDNKFLNMSSAPKIELIVDKVASLSEGRLEGRLPEDWFRHIMDPETEFNSEFA
DALCIGIDEFAQPLPFLPFKALLVTGTAGAGKTNISIQTLAANLDCIVTATTSIAAQNLSVVLNRSKS
AQVKTIFKTFGFNSSHVSMSESRQSYIANDERSIQIQKQDL SIYWNVISDIADRALGAVACKTKE
LPDLCESSVIVIDEAGVILRHILHTVVFYWFYNALYKTPLYEDGIVPCIVCVGSPTQSNALVTSFN
PLTQNKDVKRGIDVLSALICDDVLSKYCEVDNNWIFVNNKRCADHAFGDFLKHIEFGLPLKPELI
EYVDQFVKPASYIRNPMNEIETTRLFLSHNEVKNYFRSLHEQVEVTNRNNLFVFPVYFLIKNKTF
EDYKSEIGNFSLEIEPWFKSNIHRLNTYSQFADQDLSKTVQLEEIVLEDGSVEETLITCHLKHIRN
SSIGVTSKIKASTVGFSGTYEKFVELLQSDLFIEKTSCDQTIHAYSFLSGLMFGGMYSFCCSKFT
TPEVLMEIKNIKMPSEIFLESEMSRMSPDVQTVETDERYDFGLVDDGLSDVDLLEIDPCGDPFFT
RYSKLPLTNSLSFEEISLLYTTFKDIFISRFAILQKHTKGKFGKTLLVTYNRNNVSRKQCGEIYSHL
KSFYGM LTYAIPANNYTLEGYTNDNVVHLGTDKQLPQILYKKGLPRLVIKDEMGFISVLDNNVSK
FIDVVNGQSFHLCTTVDYATVSKVSMTITKSQGLSIQKVAIDFGSDPKNLKLSSIYVGMSRVTD
PNNLIMNVNPLRLNYENDNFIAPHIVKALKNENTMLIF (SEQ ID NO: 46)

Amylase [Homo sapiens], GenBank: AAA51724.1:

MKFFLLLFTIGFCWAQYSPNTQQGRTSIVHLFEWRWVDIALECERYLAPKGFGGVQVSPPNEN
VAIYNPFRPWWERYQPVS YKLCTRSGNEDEFNMVTRCNVGVRIYVDAVINHMCNAVSAG
TSSTCGSYFNPGSRDFPAVPYSGWDFNDGKCKTGSGDIENYNDATQVRDCRLTGLLDLALEK
DYVRSKIAEYMNHLIDIGVAGFRLDASKHMPGDIAKILDKLHNLNSNWFPAGSKPFIYQEVIDL
GGEPIKSSDYFGNGRVTEFKYGAKLGTVIRKWNGEKMSYLKNWGEGWGFVPSDRALVFVDN
HDNQRGHGAGGASILTFWDARLYKMAVGFM LAHPYGFTRVMSSYRWPRQFQNGNDVNDWW
GPPNNNGVIKEVTINPDTTCGNDWWCEHRWRQIRNMVIFRNVVDGQPFTNWDNGSNQVAF
GRGNRGFIVFNDDWSFSLTLQTGLPAGTYCDVISGD KINGNCTGIKIYVSDDGKAHFSISNSA
EDPFIAIHAESKL (SEQ ID NO: 47)

Alpha-glucosidase [Homo sapiens], GenBank: ABI53718.1:

MGVRHPPCSHRLLAVALVSLATAALLGHILLHDFLLVPRELSGSSPVLEETHPAHQQGASRPG
PRDAQAHPRPRAVPTQCDVPPNSRFDCAPDKAITQE QCEARGCCYIPAKQGLQGAQMGP
WCFPPSYPSYKLENLSSEMGYTATLTRTPTFFPKDILTLRLDVM METENRLHFTIKDPANR
RYEVPLETPRVHSRAPSPLYSVEFSEEPFGVIVHRQLDGRVLLNTTVAPLFFADQFLQLSTSLP
SQYITGLAEHLSPLMLSTSWTRITLWNRDLAPTPGANLYGSHPFYLALEDGGS AHGVFLLNSNA
MDVVLQPSPALSWRSTGGILDVYIFLGPEPKSVVQQYLDVVGYPFMPYWG LGFHL CRWGY
STAITRQVVENMTRAHFPLDVQWNDLDYMDSRRDFTFNKDGFRDFPAMVQELHQGGRRYMM
IVDPAISSSGPAGSYRLYDEGLRRGVFITNETGQPLIGKVWPGSTAFP DFTNPTALAWWEDMV
AEFHDQVPFDGMWIDMNEPSNFIRGSEDGCPNNELENPPYVPGVVG GTLQAATICASSHQFL
STHYNLHNL YGLTEAIA SHRALVKARGTRPFVISRSTFAGHGRYAGHWTGDVWSSWEQLASS
VPEILQFNLLGVPLVGADVCGFLGNTSEELCVRWTQLGAFYPFMRNHNSLLSLPQEPYSFSEP
AQQAMRKALT LRYALLPHLYTLFHQAHVAGETVARPLFLEFPKDSSTWTVDHQLLWGEALLITP
VLQAGKAEVTGYFPLGTWYDLQTVPIEALGSLPPPPAAPREPAIHSEGQWWTLPAPLDTINVHL
RAGYIIPLQGPGLTTTESRQQPMALAVALT KGGGEARGELFWDDGESLEVLERGAYTQVIFLARN
NTIVNELVRVTSEGAGLQLQKVTVLGVATAPQQVLSNGVPVSNFTYSPDTKVLDICVSLLMGEQ
FLVSWC (SEQ ID NO: 48)

FIG. 17 cont'd

Transcription factor SP1 [Homo sapiens], UniProtKB/Swiss-Prot: P08047.3:

MSDQDHSMDEMTAVVKIEKGVGGNNGGNGNGGGAFSQARSSSTGSSSSTGGGGQESQPSP
 LALLAATCSRIESPNENSNNSQGQPSQSGGTGELDLTATQLSQGANGWQIISSSSGATPTSKEQ
 SGSSTNGSNGSESSKNRTVSGGQYVVAAPNLQNQQVLTGLPGVMPNIQYQVIPQFQTVDDGQ
 QLQFAATGAQVQQDGSGQIQIIPGANQQIITNRSGSGNIIAAMPNLLQQAVPLQGLANNVLSGQ
 TQYVTNVPVALNGNITLLPVNSVSAATLTPSSQAVTISSSGSQESGSQPVTSGTTISSASLVSSQ
 ASSSSFFTANANSYSTTTTTSNMGIMNFTTSGSSGTNSQGQTPQRVSGLQGS DALNIQQNQTS
 GGS LQAGQQKEGEQNQQTQQQQILIQPQLVQGGQALQALQAAPLSGQTFTTQAISQETLQNL
 QLQAVPNSGPIIIRTPTVGPNGQVSWQTLQLQNLQVQNPQAQTITLAPMQGVSLGQTSSSNTT
 LTPIASAASIPAGTVTVNAAQLSSMPGLQTINLSALGTSGIQVHPIQGLPLAIANAPGDHGAQLGL
 HGAGGDGIHDDTAGGEEGENSPDAQPQAGRRTREACTCPYCKDSEGRSGSDPGKKKQHIC
 HIQGCCKVYGKTSHLRAHLRWHTGERPFMCTWSYCGKRFRTRDELQRHKRHTHTGEKKFACP
 ECPKRFMRSDHLSKHIKTHQNKKGPGVALSVGTLPLDSGAGSESGSGTATPSALITTNMVAME
 AICPEGIARLANSGINVMQVADLQSNISGNGF (SEQ ID NO: 49)

Transcription factor AP-1 [Homo sapiens], NP_002219.1:

MTAKMETTFYDDALNASFLPSESGPYGYSNPKILKQSMTLNLADPVGSLKPHLRKNSDLLTSP
 DVGLLLKLASPELERLIIQSSNGHITTTPTPTQFLCPKNVTDEQEGFAEGFVRALAEHLSQNTLPS
 VTSAAPVNGAGMVAPAVASVAGGSGSGGFSASLHSEPPVYANLSNFNPGALSSGGGAPSY
 GAAGLAFPAQPQQQQPPHLPQQMPVQHPRQLALKEEPQTVPEMPGETPPLSPIDMESQE
 RIKAEKRMNRNRIAASKCRKRKLERIARLEEKVKTLKAQNSELASTANMLREQVAQLKQKVMN
 HVNSGCQLMLTQQLQTF (SEQ ID NO: 50)

Heat shock factor protein 1, UniProtKB/Swiss-Prot: Q00613.1:

MDLPVGPGAAGPSNVPAFLTKLWTLVSDPDALICWSPSGNSFHVFDQGGQFAKEVLPKYFK
 HNNMASFVRQLNMYGFRKVVHIEQGGLVKPERDDTEFQHPCFLRGQEQLLENIKRKVTSVSTL
 KSEDIKIRQDSVTKLLTDVQLMKGKQECMDSKLLAMKHENEALWREVASLRQKHAQQQKVVN
 KLIQFLISLVQSNRILGVKRKIPLMLNDSGSAHSMPKYSRQFSLEHVHSGSPYSAPSPAYSSSS
 YAPDAVASSGPIISDITELAPASPMASPGGSIDERPLSSSPLVRVKEEPPSPPPQSPRVEEASPGR
 PSSVDTLTSPALIDSILRESEPPASVTALTDARGHTDTEGRPPSPPTSTPEKCLSVACLDKN
 ELSDHLDAMDNDNLQTMSSSHGFSVDTSALLDLFSPSVTPDMSLPDLSSLASIQELLSPQ
 EPPRPPEAENSSPDGKQLVHYTAQPLFLDPGSVDTSNDLPVLFELGEGSYFSEGDGFAED
 PTISLLTGSEPPKAKDPTVS (SEQ ID NO: 51)

CCAAT/enhancer-binding protein (C/EBP) beta isoform a [Homo sapiens], NP_005185.2:

MQRLVAWDPACLPLPPPPAFKSMEVANFYEADCLAAAYGGKAAPAAPPAARPGPRPPAGE
 LGSIGDHERAIDFSPYLEPLGAPQAPAPATATDTFEAAPAPAPAPASSGQHHDFLSDLFSDDY
 GGKNCKKPAEYGYVSLGRLGAAKGALHPGCFAPLHPPPPPPPPPAELKAEPGFEPADCKRKE
 EAGAPGGGAGMAAGFPYALRAYLGYQAVPSGSSGSLSTSSSSSPPGTPSPADAKAPPTACYA
 GAAPAPSQVKSKAKKTVDKHSDEYKIRRERNNIIVRKS RDKAKMRNLETQHKVLELTAENERL
 QKKVEQLSRELSTLRNLFKQLPEPLASSGHC (SEQ ID NO: 52)

Octamer-binding protein 1 (Oct-1), UniProtKB/Swiss-Prot: P14859.2:

MNNPSETSKPSMESGDGNTGTQTNGLDFQKQVPVVGGAISTAQAQAFGLHLHQVQLAGTSLQ
 AAAQSLNVQSKSNEESGDSQQPSQPSQPSVQAAIPQTQLMLAGGQITGLTLTPAQQQQLLQ
 QAQAQAQLLAAAVQQHSASQQHSAAGATISASAATPMTQIPLSQPIQIAQDLQQLQQLQQQNL
 NLQQFVLVHPTTNLQPAQFIISQTPQGQQGQLLQAQNLTLQLPQQSQANLLQSQPSITLTSQPAT
 PTRTIAATPIQTL PQSQSTPKRIDTPSLEEPSDLEEELEQFAKTFKQRRIKLGFTQGDVGLAMGKL

FIG. 17 cont'd

YGNDFSQTTISRFEALNLSFKNMCKLKLPLEKWLND AENLSSDSSLSSPSALNSPGIEGLSRRR
KKRTSIETNIRVALEKSFLNQKPTSEEITMIADQLNMEKEVIRVWF CNRRQKEKRINPPSSGGT
SSSPIKAIFPSPTSLVATTPSLVTSSAATTLTVSPVLPLTSAAVTNLSVTGTSDTTSNNTATVISTA
PPASSAVTSPSLSPSPSASASTSEASSASETSTTQTTSTPLSSPLGTSQVMVTASGLQTAAAAA
LQGAAQLPANASLAAMAAAAGLNPSLMA PSQFAAGGALLSLNPGTSLGALSPALMSNSTLATI
QALASGGSPLITSLDATGNLVFANAGGAPNIVTAPLFLNPQNLSLLTSNPVSLVSAAAASAGNSA
PVASLHATSTSAESI QNSLFTVASASGAAS TTTTASKAQ (SEQ ID NO: 53)

Transforming growth factor receptor beta 1 [Homo sapiens], GenBank: CAF02096.2:

MEAAVAAPRPRLLLLVLAAAAAALLPGATALQCFCHLCTKDNFTCVTDGLCFVSVTETTDK
VIHNSMCIAEIDLIPDRPFVFCAPSSKTGSVTTTYCCNQDHCNKIELPTTGPFVSVKSSPGLGPVE
LAAVIAGPVCFCISLMLMVYICHNRTVIHHRVPNEEDPSLDRPFISEGTTLKDLIYDMTTSGSGS
GLPLL VQRTIARTIVLQESIGKGRFGEVWRGKWRGEEVAVKIFSSREERSWFFREAEIYQTVMLR
HENILGFIAADNKDNGTWTQLWLVS DYHEHGS LFDYLNRYTVTVEGMIKLALSTASGLAHLHME
IVGTQGKPAIAHRDLKSKNILVKKNGTCC IADLGLAVRHDSATDTIDIAPNHRVGT KRYMAPEVL
DDSINMKHFESFKRADIYAMGLVFWEIARRCSIGGIHEDYQLPYYDLVPSDPSVEEMRKVVCEQ
KL RPNIPNRWQSCEALRVMKIMREC WYANGAARLTALRIKKTLSQLSQQEGIKM (SEQ ID NO:
54)

Platelet-derived growth factor receptor [Homo sapiens], GenBank: AAA60049.1:

MRLPGAMPALALKGELLLL SLLLLLEPQISQGLVVT PPGPELV LNVSSTFVLTCSGSAPVWVER
MSQEPPQEMAKAQDGT FSSVLTLTNLTGLDTGEYFCTHNSR GLETDERKRLYIFVPDPTVGF
LPND AEELFIFLTEITEITIPCRVTD PQLVVT LHEKKGDVALPVPYDHQRGFSGIFEDRSYICKTTI
GDREVDSDAYYVYRLQVSSINVS VNAVQTVVRQGENITLMCIVIGNEVVNF EWTPRKESGRL
VEPVTD FLLDMPYHIRSILHIPS AELEDSGYTCNVTESVNDHQDEKAINITV VESGYVRLLGEV
GTLQFAELHRSRTLQVVF EAYPPPTVLWFKDNRTLGDSSAGEIALSTRNVSETRYVSELT LVRV
KVAEAGHYTMRAF HEDAEVQLSFQLQINVPVRVLELSESHPD SGEQTVRCRGRGMPQPNIWS
ACRDLKRCPRELPPTLLGNSSEEE SQLETNVTYWEEEQEF EVVSTLRLQHVD RPLSVRCTLRN
AVGQDTQEVIVVPHSLPFKVVISAILALVLT IISLIILIMLWQKKPRYEIRWKVIESVSSD GHEYIY
VDPMQLPYDSTWELPRDQLVLGRTL GSGAFGQVVEATAHGLSHSQATMKVAVKMLKSTARSS
EKQALMSELKIMSHLGPLN VVNLLGACTKGGPIYIITEYCRYGDLVDYLHRNKHTFLQH HSDK
RRPPSAELYSNALPVGLPLPSHVSLTGESDGGYMDMSKDESVDYVPMLDMKGDVKYADIESS
NYMAPYDNYVPSAPERTCRATLINESPVLSYMDLVGFSYQVANGMEFLASKNCVHRDLAARN
VLICEGKLVKICDFGLARDIMRDSNYISKGSTFLPLKWMAPESIFNSLYTTLSDVWSFGILLWEIF
TLGGTPYPELPMNEQFYNAIKRGYRMAQPAHASDEIYEIMQKC WEEKFEIRPPFSQLVLLLERL
LGEGYKKKYQQVDEEFLRSDHPAILRSQARLP GFHGLRSP LDTSSVLYTAVQPNEGDN DYIIP L
PDPKPEVADEGPLEGSPSLASSTLNEVNTSSTISCDSPLEPQDEPEPEPQLELQVEPEPELEQL
PDSGCPAPRAEAEDSFL (SEQ ID NO: 55)

Epidermal growth factor receptor [Homo sapiens], GenBank: CAA25240.1:

MRPSGTAGAALLALLAALCPASRALEEKKVCQGT SNKLTQLGTFEDHFLSLQRMFNNCEVVLG
NLEITYVQRNYDLSFLKTIQEVAGYVLIALNTVERIPL ENLQIIRGNMYYENSYALAVLSNYDANKT
GLKELPMRNLQEILHGAVRFSNNPALCNVESIQWRDIVSSDFLSNM SMDFQNH LGSCQKCDPS
CPNGSCWGAGEENCQKLTKIICAQQCSGR CRGKSPSDCCHNQCAAGCTGPRES DCLVCRKF
RDEATCKDTC PPLMLYNPTTYQMDVNPEGKYSFGATCVKK CPRNYVVDH GSCVRACGADS
YEMEEDGVRKCKKCEGPCRKVCNGIGIGEFKDSLSINATNIKHFKNCT SISGDLHILPVAFRGDS
FTHTPPLDPQELDILKTVKEITGFLLIQAWPENRTDLHAFENLEIRGR TKQHGGQFSLAVVSLNITS
LGLRSLKEISDGDVIIISGNKNLCYANTINWKKLFGTSGQKTKIISNRGENSCKATGQVCHALCSP
EGCWGPEPRDCVSCRNVSRGRECVDKCKLLEGEPREFVENSECIQCHPECLPQAMNITCTGR

FIG. 17 cont'd

GPDNCIQCAHYIDGPHCVKTCPAGVMGENNTLVWKYADAGHVCHLCHPNCTYGCTGPGLEG
CPTNGPKIPSIATGMVGALLLLLVVALGIGLFMRRRHIVRKRTLRRLLQERELVEPLTPSGEAPN
QALLRILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELREATSPKANKEILDEAYVM
ASVDNPHVCRLLGICLTSTVQLITQLMPFGCLLDYVREHKDNIGSQYLLNWCVQIAKGMNYLED
RRLVHRDLAARNVLVKTPQHVKITDFGLAKLLGAEEKEYHAEGGKVPIKWMMALESILHRIYTHQS
DVWSYGVTWELMTFGSKPYDGIPASEISSILEKGERLPQPPICTIDVYMIMVKCWMIDADSRP
KFRELIIEFKSMARDPQRYLVIQGDERMHLPSPTDSNFYRALMDEEDMDDVDADEYLIPQQGF
FSSPSTSRTPLLSSLSATSNNSTVACIDRNLQSCPIKEDSFLQRYSSDPTGALTEDSIDDTFLP
VPEYINQSVPKRPAGSVQNPVYHNQPLNPAPSRDPHYQDPHSTAVGNPEYLNVTQPTCVNST
FDSPAHWAKQKSHQISLDNPDYQQDFFPKEAKPNGIFKGSTAENAAYLRVAPQSSEFIGA
(SEQ ID NO: 56)

Vascular endothelial growth factor receptor [Homo sapiens], GenBank: AAC16449.2:

MVSYWDTGVLLCALLSCLLLTGSSSGSKLKDPESLKGQTQHIMQAGQTLHLQCRGEAAHKWSL
PEMVSKESERLSITKSACGRNGKQFCSTLTNLTAQANHTGFYSCKYLAVPTSKKKETESAIYIFI
SDTGRPFVEMYSEIPEIIHMTGRELVIPCRVTSPNITVTLKKFPLDTLIPDGKRIIWDSRKGFIIISN
ATYKEIGLLTCEATVNGHLYKTNYLTHRQNTIIDVQISTPRPVKLLRGHTLVNLCTATTPLNTRV
QMTWSYPDEKNKRASVRRRIDQSNSHANIFYSVLTIDKMQNKDKGLYTCRVRSGPSFKSVNTS
VHIYDKAFITVKHRKQQVLETVAGKRSYRLSMKVKAFFSPEVVWLKDGLPATEKSARYLTRGYS
LIIKDVTEEDAGNYTILLSIKQSNVFNLTATLIVNVKPIYEKAVSSFPDPALYPLGSRQILTCTAY
GIPQPTIKWFWHPCNNHSEARCDFCSNNEESSILDADSNMGNRIESITQRMALIEGKNKMAST
LVVADSRISGIYICIASNKVGTVGRNISFYITDVPNGFHVNLKMPTEGEDLKLSCTVNKFYLRDV
TWILLRTVNNRTMHYSISQKQMAITKEHSITLNLTIMNVSLQDSGTACRARNVYTGEELQKKEI
TIRDQEAPYLLRNLSDHVAISSSTTLDCANGVPEPQITWFKNNHKIQQEPGILGPGSSSTLFIE
RVTEEDEGVYHCKATNQKGSVESSAYLTVQGTSDKSNLELITLTCTCVAATLFWLLLTFLIRKMK
RSSSEIKTDYLSIIMDPDEVPLDEQCERLPYDASKWEFARERLKLGLSLGRGAFGKVQASAFG
IKKSPTCRTVAVKMLKEGATASEYKALMTELKILTHIGHHLNVNLLGACTKQGGPLMVIVEYCK
YGNLSNYLKSQRDLFFLNKDAALHMEPKKEKMEPGLEQGKKPRLDSVTSSSEFASSGFQEDKS
LSDVEEEEDSDGFYKEPITMEDLISYSFQVARGMEFLSSRKCIHRDLAARNILLSENNVMKICDF
GLARDIYKNPDYVRKGDTRLPLKWMAPESIFDKIYSTKSDVWSYGVLLWEIFSLGGSPYPGVQ
MDEDFCSRLREGMRMRAPEYSTPEIYQIMLDCWHRDPKERPRFAELVEKLGDLLQANVQQDG
KDYIPINAILTGNSGFTYSTPAFSEDFKESISAPKFNSGSSDDVRYVNAFKFMSLERIKTFEELL
PNATSMFDDYQGDSSSTLLASPMLKRFTWTDSPKASLKIDLRVTSKSKESGLSDVSRPSFCHS
SCGHVSEGKRRFTYDHAELERKIACCSPPPDYNSVVLSTPPI (SEQ ID NO: 57)

Interleukin 8 receptor alpha [Homo sapiens], GenBank: AAB59436.1:

MSNITDPQMWDFFDDLNFTGMPPADEDYSPCMLETETLNKYVVIIAYALVFLLSLLGNSLVMLVIL
YSRVGRSVTDVYLLNLALADLLFALTLPiWAASKVNGWIFGTFLCKVVSLLKEVNFYSGILLACI
SVDRYLAIVHATRTLTKRHLVKFVCLGCWGLSMNLSLPFFLFRQAYHPNNSSPVCYEVLGND
TAKWRMVLRLPHTFGFIVPLFVMLFCYGFTRLTLFKAHMGQKHRAMRVIFAVVLIFLLCWLPLYN
LVLLADTLMRQTQVIQESCERRNIGRALDATEILGFLHSLNPIYAFIQNFRHGFLKILAMHGLV
SKEFLARHRVTSYTSSSVNVSSNL (SEQ ID NO: 58)

Caveolin [Homo sapiens], GenBank: CAA79476.1:

FIG. 17 cont'd

MSGGKYVDSEGHLYTVPIREQGNIYKPNNKAMADELSEKQVYDAHTKEIDLVNRPKHLNDDV
VKIDFEDVIAEPEGTHSFHGIWKASFTTFTVTKYWFYRLLSALFGIPMALIWGIYFAILSFLHIWAV
VPCIKSFLIEIQCTSRVYSIYVHTVCDPLFEAVGKIFSNVRINLQKEI (SEQ ID NO: 59)

Dynamin [Homo sapiens], GenBank: AAA88025.1:

MGNRGMEELIPLVNLQDAFSSIGQSSCHLDLPQIAVVGGQSAGKSSVLENFVGRDFLPRGSGI
VTRRPLILQLIFSKTEHAFLHCKSKKFTDFDEVQRQEIAETDRVTGTNKGISPVPINLRVYSPHV
LNLTIDLPGITKVPVGDQPPDIEYRVKDMILQFISRESSLILAVTPANMDLANSDALKLAKEVDP
QGLRTIGVITKLDLMDEGTDARDVLENKLLPLRRGYIGVNNRSQKDIEGKKDIRAALAAERKFFL
SHPAYRHMADRMGTPHLQKTLNQQLTNHIRESLPALRSKLQSQLLSLEKEVEEYKIFRPDDPTP
KTKALLQMVGQFGVDFEKRIEGSGDQVDTLELSGGARINRIFHERFPFELVKMEFDEKDLRREI
SYAIKNIHGVRTGLFTPDLAFAIVKKQVVKLKEPCLKCVDLVIQELINTVRQCTSKLSSYPRLRE
ETERIVTTYIREREGRTKDQILLIDIEQSYINTNHEDFIGFANAQQRSTQLNKKRAIPNQVIRRGW
LTINNISLMKGGKEYWFLTAESLSWYKDEEEKEKKYMLPLDNLKIRDVEKGFMSNKHVFAIF
NTEQRNVYKDLRQIELACDSQEDVDSWKASFLRAGVYPEKDQAENEDGAQENTFSMDPQLER
QVETIRNLVDSYVAIINKSIRDLMPKTIMHLMINNTKAFIHHELLAYLYSSADQSSLMEESADQAQ
RRDDMLRMYHALKEALNIIGDISTSTVSTPVPVDDTWLQSASSHSPTPQRRPVSSIHPPGRP
PAVRGPTPGPPLIPVPGAAASFSAAPPISRPQSVFANSDFPAPPQIPSRPVRIPPGIPPGV
PSRRPPAAPSRPTIIRPAEPSLLD (SEQ ID NO: 60)

Clathrin heavy chain 1 isoform 1 [Homo sapiens], NP_004850.1:

MAQILPIRFQEHLLQLNLGINPANIGFSTLTMESDKFICIREKVGGEQAQVVIIDMNDPSNPIRRPIS
ADSAIMNPASKVIALKAGKTLQIFNIEMKSKMKAHTMTDDVTFWKWISLNTVALVTDNAVYHWS
MEGESQPVKMFDRHSSLAGCQIINYRTDAKQKWLLLTGISAQQNRVVGAMQLYSVDRKVSQPI
EGHAASFAQFKMEGNAAESTLFCFAVRGQAGGKLHIEVGTPTGNQPFPPKAVDVFFPPEAQ
NDFPVAMQISEKHVVFLITKYGYIHLYDLETGTCIYMNRISETIFVTAPHEATAGIIGVNRKGQ
VLSVCVEEENIIPYITNVLQNPDLALRMAVRNNLAGAEELFARKFNALFAQGNYSAAKVAANAP
KGILRTPDTIRRFQSVPAQPGQTSPLLQYFGILLDQGGQLNKYESLELCRPVLQQGRKQKLEKWL
KEDKLECSEELGDLVKSVDPTLALSVYLRANVPNKVIQCFAETGQVQKIVLYAKKVGYPDWIFL
LRNVMRISPDQGGQFAQMLVQDEEPLADITQIVDFVMEYNLIQQCTAFLLDALKNNRPSEGPLQ
TRLLEMNLMHAPQVADAILGNQMFTHYDRAHIAQLCEKAGLLQRALEHFTDLYDIKRAVVHHL
LNPEWLNVNYFGSLSVEDSLECLRAMLSANIRQNLQICVQVASKYHEQLSTQSLIELFESFKSFE
GLFYFLGSIVNFSQDPDVHFYKIQAACTGQIKEVERICRESNCYDPERVKNFLKEAKLTDQLPL
IIVCDRFDVFVHDLVLYLYRNLLQKYIEIYVQKVNPSRLPVVIGGLLDVDCSEDEVKNLILVVRGQFS
TDELVAEVEKRNRLKLLLPWLEARIHEGCEEPATHNALAKIYDSNNNPERFLRENPPYYDSRVV
GKYCEKRDPHLACVAYERGQCDLELINVCNENSLFKSLSRYLVRKDPPELWGSVLLESNPYRR
PLIDQVVQTALSETQDPPEEVSVTVKAFMTADLPNELIELLEKIVLDNSVFSEHRNLQNLLILTAIKA
DRTRVMEYINRLDNYDAPDIANIAISNELFEEAFAIFRKFDVNTSAVQVLIHIGNLDRAEYFAER
CNEPAVWSQLAKAQLQKGMVKEAIDSYIKADDPSSYMEVVQAANTSGNWEELVKYLQMARKK
ARESYVETELIFALAKTNRLAELEEFINGPNNAHIQQVGDRCYDEKMYDAAKLLYNNVSNFGR
ASTLVHLGEYQAAVDGARKANSTRTWKEVCACVDGKEFRLAQMCGLHIVVHADELEELINYY
QDRGYFEELITMLEAALGLERAHMGMTLAILYSKFKPQKMREHLELFWSRVNIPKVLRAAEQ
AHLWAEVLFLYDKYEEYDNAIITMMNHPTDAWKEGQFKDIITKVANVELYYRAIQFYLEFKPLLL
NDLLMVLSPRLDHTRAVNYFSKVKQLPLVKPYLRSVQNHNNKSVNESLNNLFITEEDYQALRTS
IDAYDNFDNISLAQRLEKHELIEFRRIAAYLFKGNNRWKQSVELCKKDSLYKDAMQYASESKDT
ELAEELLQWFLQEEKRECFGACLFTCYDLLRPDVVLETAWRHNIMDFAMPYFIQVMKEYLTKV
DKLDASESLRKEEEQATETQPIVYGGPQLMLTAGPSVAVPPQAPFGYGYTAPPYGGPQPGFG
YSM (SEQ ID NO: 61)

FIG. 17 cont'd

Clathrin heavy chain 2 isoform 1 [Homo sapiens], NP_009029.3:

MAQILPVRFQEHFQLQNLGINPANIGFSTLTMESDKFICIREKVGEQAQVTIIDMSDPMAPIRRPI
SAESAIMNPASKVIALKAGKTLQIFNIEMKSKMKAHTMAEEVIFWKWWSVNTVALVTETAVYHW
SMEGDSQPMKMFDRHTSLVGCQVIHYRTDEYQKWLLLVGISAQQNRVVGAMQLYSVDRKVS
QPIEGHAAAFAEFKMEGNAKPATLFCFAVRNPTGGKLHIIIEVGQPAAGNQPFVKKAVDVFFPPE
AQNDPVMQIGAKHGVYILITKYGYLHLYDLESGVCICMNRISADTIFVTAPHKPTSGIIGVNNK
GQVLSVCVEEDNIVNYATNVLQNPDLGLRLAVRSNLAGAEKLFVRKFNTLFAQGSYAEAAKVA
ASAPKGILRTRETQKFSIPAQSGQASPLLQYFGILLDQGQLNKLESLELCHLVLQQGRKQLL
EKWLKEDKLECSEELGDLVKTTPMLALSVELRANVPSKVIQCFAETGQFQKIVLYAKKVGYP
DWIFLLRGVMKISPEQGLQFSRMLVQDEEPLANISQIVDFMENSIIQQCTSFLLDALKNNRPAE
GLLQTWLLEMNLVHAPQVADAILGNKMFTHYDRAHIAQLCEKAGLLQQALEHYTDLYDIKRAVV
HTHLLNPEWLVNFFGSLSVEDSVECLHAMLNANIRQNLQLCVQVASKYHEQLGTQALVELFES
FKSYKGLFYFLGSIVNFSQDPDVHLKYIAACKTGQIKEVERICRESSCYNPERVKNFLKEAKLT
DQLPLIIVCDRFGFVHDLVLYLRNNLQRYIEIYVQKVNPSRTPAVIGGLLDVDCSEEVIKHLIMAV
RGQFSTDELVAEVEKRNRLKLLPWLESQIQEGCEEPATHNALAKIYIDSNNSPECFLRENAYY
DSSVVGRYCEKRDPHLACVAYERGQCDLELIKVCNENSLFKSEARYLVCRKDPELWAHVLEET
NPSRRQLIDQVVQTALSETRDPEEISVTVKAFMTADLPNELIELLEKIVLDNSVFSEHRNLQNLII
LTAIKADRTRVMEYISRLDNYDALDIASIAVSSALYEEAFTVFHKFDMNASAIQVLIHIGNLDRAV
EFAERCNEPAVWSQLAQALQKDLVKEAINSIRGDDPSSYLEVVQSASRSNNWEDLVKFLQ
MARKKGRESYIETELIFALAKTSRVSELEDINGPNNAHIQVGDRCYEEGMYEAAKLLYSNVS
NFARLASTLVHLGEYQAAVDNSRKASSTRTWKEVCFACMDGQEFRAQLCGLHIVIHADLEE
LMCYYQDRGYFEELILLLEAALGLERAHMGMTTELAILYSKFKPKMLEHLELFWSRVNIPKVL
AAEQAHLWAEVFLYDKYEEYDNAVLTMMSHPTAEWKEGQFKDITKVANVELCYRALQFYLD
YKPLLINDLLLVLSPRLDHTWTVSFFSKAGQLPLVKPYLRSVQSHNNKSNEALNHLTEEDY
QGLRASIDAYDNFDNISLAQQLEKHQLMEFRCIAAYLYKGNNWWAQSVELCKKDHLYKDAMQ
HAAESRDAELAQLLQWFLEEGKRECFAACLFCTCYDLLRPDMVLELAWRHNLDLAMPYFIQV
MREYLSKVDKLDALLESRLKQEEHVTEPAPLVDFDFDGHE (SEQ ID NO: 62)

Clathrin light chain A isoform a [Homo sapiens], NP_001824.1:

MAELDPFGAPAGAPGGPALGNGVAGAGEEDPAAAFQAQSEIAGIENDEAFAILDGGAPGPQ
PHGEPPGGPDAVDGVMNGEYYQESNGPTDSYAAISQVDRQLQSEPESIRKWREEQMERLEAL
DANSRKQEAWEKAIKELEEWYARQDEQLQKTKANNRAAEEAFVNDIDESSPGTEWERVER
LCDFNPKSSKQAKDVSRLSVLMSLKQAPLVH (SEQ ID NO: 63)

Clathrin light chain B isoform a [Homo sapiens], NP_001825.1:

MADDFGFFSSSESGAPEAAEEDPAAAFQAQSEIAGIENDEGFGAPAGSHAAPAQPGPTSGA
GSEDMGTTVNGDVFEANGPADGYAAIAQADRLTQEPESIRKWREEQRKRLQELDAASKVTE
QEWREKAKKDLEEWNRQSEQVEKNKINNRASEEAFVKESKEETPGTEWEKVAQLCDFNPK
SSKQCKDVSRLSVLMSLKQTPSLR (SEQ ID NO: 64)

Ras-related protein Rab-4A isoform 1 [Homo sapiens], NP_004569.2:

MSQTAMSETYDFLFKFLVIGNAGTGKSCLLHQFIEKKFKDDSNHTIGVEFGSKIINVGGKYVKLQI
WDTAGQERFRSVTRSYRGAAGALLVYDITSRETYNALTNLWLTDMRLASQNIIVILCGNKKDL
DADREVTFLASRFAQENELMFLETSALTGENVEEAFVQCARKILNKIESGELDPERMGSGIQY
GDAALRQLRSPRAQAPNAQECGC (SEQ ID NO: 65)

Ras-related protein Rab-11A, UniProtKB/Swiss-Prot: P62491.3:

MGTRDDEYDYLKVVLLIGDSGVGKSNLLSRFTRNEFNLESKSTIGVEFATRSIQVDGKTIKAQIW
DTAGQERYRAITSAYYRGAVGALLVYDIAKHLTYENVERWLKELRDHADSNIIVIMLVGNKSDLR

FIG. 17 cont'd

HLRAVPTDEARAFAEKNGLSFIETSALDSTNVEAAFQILTEIYRIVSQKQMSDRRENDMSPSNN
VVPIHVPPTTENKPKVQCCQNI (SEQ ID NO: 66)

Platelet-derived growth factor [Homo sapiens], GenBank: AAA60552.1:

MNRCWALFLSLCCYLRLVSAEGDPIPEELYEMLSDHSIRSFDLQRLHGDGPGEEDGAELDLN
MTRSHSGGELESLARGRRSLGSLTIAEPAMIAECKTRTEVFESISRLIDRTNANFLVWPPCDEV
QRCSGCCNNRNVQCRPTQVQLRPVQVRKIEIVRKKPIFKKATVTLEDHLACKCETVAAARPVT
RSPGGSQEQRATPQTRVTIRTVRVRRPPKGKHKRKFHHTDKTALKETLGA (SEQ ID NO: 67)

Transforming growth factor-beta3 [Homo sapiens], GenBank: AAA61161.1:

MKMHLQRALVVLALLNFATVSLSLSTCTTLDGHIKKRVEAIRGQILSKRLTSPPEPTVMTHV
PYQVLALYNSTRELLEEMHGEREEGCTQENTESEYYAKEIHKFDMIQGLAEHNELAVCPKGITS
KVFRFNVSSVEKNRTNLFRAEFRVLVVPNPSSKRNEQRIELFQILRPDEHIAKQRYIGGNLPT
GTAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENIHEVMEIKFKGVDNEDD
HGRGDLGRLKKQKDHNNPHLILMMIPPHRLDNPGQGGQKRRALDTNYCFRNLEENCCVRPL
YIDFRQDLGWKWWHEPKGYANFCSGPCPYLRADTTHSTVLGLYNTLNPEASASPCCVPQD
LEPLTILYYVGRTPKVEQLSNMVVKSCCKS (SEQ ID NO: 68)

Nerve growth factor [Homo sapiens], GenBank: CAA37703.1:

MSILFYVIFLAYLRGIQGNMDQRSLPEDSLNSLIKLIQADILKNKLSKQMVDVKENYQSTLPKA
EAPREPERGGPAKSAFQPVAMDTELLRQQRRYNsprVLLSDSTPLEPPPLYLMEDYVGSPVV
ANRTSRRKRYAEHKSHRGEYSVCDSESLWTDKSSAIDIRGHQVTVLGEIKTGNSPVKQYFYE
TRCKEARPVKNGCRGIDDKHWSQCKTSQTYVRALTSENNKLVGWRWIRIDTSCVCALSRKIG
RT (SEQ ID NO: 69)

Epidermal growth factor (EGF) [Homo sapiens], GenBank: CAA34902.2:

MRPSGTAGAALLALLAALCPASRALEEKKGKGVSRRLPRRPRIAPRTPQPAQPRTGAPARARA
PARPFLFP (SEQ ID NO: 70)

GTPase HRas [Homo sapiens]:

MTEYKLVVVGAGGVGKSALTIQLIQNHVDEYDPTIEDSYRKQVVIDGETCLLDILDITAGQEEYS
AMRDQYMRTGEGFLCVFAINNTKSFEDIHQYREQIKRVKDSDDVPMVLVGNKCDLAARTVESR
QAQDLARSYGIPYIETSAKTRQGVEDAFYTLVREIRQHKLRLNPPDESGPGCMSCKCVLS
(SEQ ID NO: 71)

Cocaine And Amphetamine Regulated Transcript (Chain A) [Homo sapiens], PDB: 1HY9_A:

YGQVPMCDAGEQCAVRKGARIGKLCDCPRGTSCNSFLLKCL (SEQ ID NO: 72)

Protachykinin-1 [Homo sapiens], UniProtKB - P20366:

MKILVALAVFFLVSTQLFAEEIGANDDLNYWSDWYDSDQIKEELPEPFEHLLQRIARRPKPQQFF
GLMGKRDADSSIEKQVALLKALYGHGQISHKRHKTDSVFGLMGKRALNSVAYERSAMQNYER
RR (SEQ ID NO: 73)

Substance P is position 58-68 of Protachykinin-1 [Homo sapiens]:

RPKPQQFFGLM (SEQ ID NO: 74)

Oxytocin-neurophysin 1 [Homo sapiens], UniProtKB - P01178:

97/104

FIG. 17 cont'd

MAGPSLACCLLGLLALTSACYIQNCPLGGKRAAPDLVVRKCLPCGPGGKGRCFGPNICCAEEL
GCFVGTAEALRCQEENYLPSPCQSGQKACGSGGRCAVLGLCCSPDGCHADPACDAEATFSQ
R (SEQ ID NO: 75)

Oxytocin is position 20-28 of Oxytocin-neurophysin 1 [Homo sapiens]:
CYIQNCPLG (SEQ ID NO: 76)

Somatostatin [Homo sapiens], GenBank: AAH32625.1:
MLSCRLQCALAALSIVLALGCVTGAPSDPRLRQFLQKSLAAAAGKQELAKYFLAELLSEPNQTE
NDALEPEDLSQAAEQDEMRLQLRSANSNPAMAPRERKAGCKNFFWKFTFTSC (SEQ ID NO:
77)

Myosin light chain kinase, Green fluorescent protein, Calmodulin chimera (Chain A)
[synthetic construct], PDB: 3EKJ_A:

MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDLATMVDSSRRKWNKTGHAVRAIGRLSSL
ENVYIMADKQKNGIKANFKIRHNIEDGGVQLAYHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDP
NEKRDHMLVLEFVTAAGITLGMDELYKGGTGGSMVSKGEELFTGVVPILVELDGDVNGHKFSV
SGEGEGDATYGKLTCLKFICTTGKLPVPWPTLVTTLXVQCFSRYPDHMKQHDFFKSAMPEGYIQ
ERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNTRDQLTEEQIAEFKE
AFSLFDKDGDDGITTQKLGTVMRSLGQNPTAEALQDMINEVGADGNGTIDFPQFLTMMARKMK
DTDSEEEIREAFRVFGKDGNGYISAAQLRHVMTNLGEKLTDEEVDEMIREAGIDGDGQVNYEQ
FVQMMTAK (SEQ ID NO: 78)

Genetically-encoded green calcium indicator NTnC (chain A) [synthetic construct], PDB:
5MWC_A:

MVSKGEEDNMAASLPATHELHIFGSINGVDFDMVGQGSNGPNVGYEELNLKSTKGDQLQFSPWIL
VPHIXFHQYLPYPDGMSPFQAAMVDGSGYQVHRTVQFEDGASLTVNYRYTYEGSHIKGEAQV
KGTGFPADGPMANSLTAMVPSEEEELSECFRTFDKDGDDGFIDREEFGGIIRLTGEQLTDEDPD
EIFGDSDDTKNGRIDFDEFLKMVENVQLSMADWCRSKMACPNDKTLISTLKWSTYTTGNGKRYR
STARTTYTFAKPMAANYLKNQPMYVFRKTELKHSKTELNFKEWQKAFTDVMGMDELYK (SEQ
ID NO: 79)

Calcium indicator TN-XXL [synthetic construct], GenBank: ACF93133.1:

MVSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTCLKFICTTGKLPVPWPTLVTT
LTWGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELK
GIDFKEDGNILGHKLEYNYISHNVYITADKQKNGIKAHFKIRHNIEDGSVQLADHYQQNTPIGDG
PVLLPDNHYLSTQSKLSKDPNEKRDHMLVLEFVTAARMLSEEELANCFRIFDKDANGFIDIEELG
EILRATGEHVTEEDIEDLMKDSKNDGRIDFDEFLKMMEGVQGTSEEELANCFRIFDKDANGF
IDIEELGEILRATGEHVTEEDIEDLMKDSKNDGRIDFDEFLKMMEGVQELMGGVQLADHYQQ
NTPIGDGPVLLPDNHYLSYQSKLSKDPNEKRDHMLVLEFVTAAGITLGMDELYKGGTGGSMVS
KGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTCLKFICTTGKLPVPWPTLVTTLG
GLMCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGID
FKEDGNILGHKLEYNYNSHNVIYIMADKQKNGIKANFKIRHNIEDGGVQLADHYQQNTPIGDG
(SEQ ID NO: 80)

BRET-based auto-luminescent calcium indicator [synthetic construct], GenBank:
ADF42668.1:

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKICTTGKLPVPWPTLVTT
LGYGLQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELK
GIDFKEDGNILGHKLEYNYNSHNVIYITADKQKNGIKANFKIRHNIEDGGVQLADHYQQNTPIGDG

FIG. 17 cont'd

PVLLPDNHLYSYQSKLSKDPNEKRDHMLLEFVTAAGITLGMDELYKRMHDQLTEEQIAEFKEA
FSLFDKDGDTITTKELGTVMRSLGQNPTAEELQDMINEVDADGNGTIYFPEFLTMMARKMKD
TDSEEEIREAFRVFDKDGNGYISAAQLRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFV
QMMTAKGGKRRWKKNFIAVSAANRFKKISSSGALELMTSKVYDPEQRKRMITGPQWWARCK
QMNVLDSFINYYDSEKHAENAVIFLHGNATSSYLWRHVPHIEPVARCIIPDLIGMGKSGKSGN
GSYRLLDHYKYLTAWFELLNLPKKIIFVGHWDGAALAFHYAYEHQDRIKAIVHMESVVDVIESW
DEWPDIEEDIALIKSEEGEKMLENNFFVETVLPSPKIMRKLEPEEFAAYLEPFKEKGEVRRPTLS
WPREIPLVKGGKPDVVQIVRNYNAYLRASDDLPLKFIESDPGFFSNAIVEGAKKFPNTEFVKVK
GLHFLQEDAPDEMCKYIKSFVERVLKNEQ (SEQ ID NO: 81)

Calcium indicator protein OeNL(Ca²⁺)-18u [synthetic construct], GenBank: BBB18812.1:
MVSVIKPEMKMRYMDGSGVNGHEFTIEGEGTGRPYEGHQEMTLRVTMAEGGPMPFAFDLVS
HVFCYGHRVFTKYPEEIPDYFKQAFPEGLSWERSLEFEDGGSASVSAHISLRGNTFYHKSFT
GVNFPADGPIMQNSVDWEPSTEKITASDGVKGDVTMYLKLEGGGNHKCQFKTTYKAAKEIL
EMPGDHYIGHRLVRKTEGNITEQVEDAVAHSGTLEDFVGDWRQTAGYNLDQVLEQGGVSSLF
QNLGVSVTPIQRIVLSGENGLKIDHVIIPYEGPMMHDQLTEEQIAEFKEAFSLFDKDGDTITTK
ELGTVMRSLGQNPTAEELQDMINEVDADGNGTIYFPDFLTMMARKMKD TDSEEEIREAFRVFD
KDGNGYISAADLRHVMTNLGEKLTDEEVDEMIREADIDGEGQVNYEEFVQMMTAKGGKRRWK
KNFIAVSAANRFKKISSSGALELLSGDQMGQIEKIFKVVPVDDHHFKVILHYGTLVIDGVTPNMI
DYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSLLFRVTINGVTGWRLCERILA (SEQ
ID NO: 82)

GCaMP6m

ATGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGG
GTCGGGATCTGTACGACGATGACGATAAGGATCTCGCCACCATGGTCGACTCATCACGTC
GTAAGTGAATAAGACAGGTCACGCAGTCAGAGCTATAGGTCGGCTGAGCTCACTCGAGA
ACGTCTATATCAAGGCCGACAAGCAGAAGAACGGCATCAAGGCGAACTTCAAGATCCGCC
ACAACATCGAGGACGGCGGCGTGCAGCTCGCCTACCACTACCAGCAGAACACCCCCATC
GGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCGTGCAGTCCAAACTTTTCG
AAAGACCCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGCG
GATCACTCTCGGCATGGACGAGCTGTACAAGGGCGGTACCGGAGGGAGCATGGTGAGCA
AGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGTGAGGGCGATGCCACCTACGGCAAGCT
GACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGAC
CACCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCATGAAGCAGCACGA
CTTCTTCAAGTCCGCCATGCCCGAAGGCTACATCCAGGAGCGCACCATCTTCTTCAAGGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACCTGCCGGACCAACTGACTGAAGAGCAGATCGCAGAATTTAAAGAGGCTTTCTC
CCTATTTGACAAGGACGGGGATGGGACAATAACAACCAAGGAGCTGGGGACGGTGATGC
GGTCTCTGGGGCAGAACCCACAGAAAGCAGAGCTGCAGGACATGATCAATGAAGTAGATG
CCGACGGTGACGGCACAATCGACTTCCCTGAGTTCCTGACAATGATGGCAAGAAAAGGGA
GCTACAGGGACACGGAAGAAGAAATTAGAGAAGCGTTCGGTGTGTTTGATAAGGATGGCA
ATGGCTACATCAGTGACGAGAGCTTCGCCACGTGATGACAAACCTTGGAGAGAAAGTTAA
CAGATGAAGAGGTTGATGAAATGATCAGGGAAGCAGACATCGATGGGGATGGTCAGGTAA
ACTACGAAGAGTTTGTACAAATGATGACAGCGAAGTGA (SEQ ID NO: 99)

GCaMP6s

FIG. 17 cont'd

ATGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGG
GTCGGGATCTGTACGACGATGACGATAAGGATCTCGCCACCATGGTCGACTCATCACGTC
GTAAGTGGAATAAGACAGGTCACGCAGTCAGAGCTATAGGTCGGCTGAGCTCACTCGAGA
ACGTCTATATCAAGGCCGACAAGCAGAAGAACGGCATCAAGGCCGAACCTCCACATCCGCC
ACAACATCGAGGACGGCGGCGTGCAGCTCGCCTACCACTACCAGCAGAACACCCCCATC
GGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCGTGCAGTCCAAACTTTTCG
AAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGGGCGGTACCGGAGGGAGCATGGTGAGCA
AGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGTGAGGGCGATGCCACCTACGGCAAGCT
GACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCACCCTCGTGAC
CACCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGA
CTTCTTCAAGTCCGCCATGCCCGAAGGCTACATCCAGGAGCGCACCATCTTCTTCAAGGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACCTGCCGGACCAACTGACTGAAGAGCAGATCGCAGAATTTAAAGAGGCTTTCTC
CCTATTTGACAAGGACGGGGATGGGACAATAACAACCAAGGAGCTGGGGACGGTGATGC
GGTCTCTGGGGCAGAACCCACAGAAAGCAGAGCTGCAGGACATGATCAATGAAGTAGATG
CCGACGGTGACGGCACAATCGACTTCCCTGAGTTCCTGACAATGATGGCAAGAAAAATGA
AATACAGGGACACGGAAGAAGAAATTAGAGAAGCGTTCGGTGTGTTTGATAAGGATGGCA
ATGGCTACATCAGTGCAGCAGAGCTTCGCCACGTGATGACAAACCTTGGAGAGAAGTTAA
CAGATGAAGAGGTTGATGAAATGATCAGGGAAGCAGACATCGATGGGGATGGTCAGGTAA
ACTACGAAGAGTTTGTACAAATGATGACAGCGAAGTGA (SEQ ID NO: 100)

GCaMP6f

ATGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGG
GTCGGGATCTGTACGACGATGACGATAAGGATCTCGCCACCATGGTCGACTCATCACGTC
GTAAGTGGAATAAGACAGGTCACGCAGTCAGAGCTATAGGTCGGCTGAGCTCACTCGAGA
ACGTCTATATCAAGGCCGACAAGCAGAAGAACGGCATCAAGGCCGAACCTCAAGATCCGCC
ACAACATCGAGGACGGCGGCGTGCAGCTCGCCTACCACTACCAGCAGAACACCCCCATC
GGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCGTGCAGTCCAAACTTTTCG
AAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGGGCGGTACCGGAGGGAGCATGGTGAGCA
AGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGTGAGGGCGATGCCACCTACGGCAAGCT
GACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCACCCTCGTGAC
CACCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGA
CTTCTTCAAGTCCGCCATGCCCGAAGGCTACATCCAGGAGCGCACCATCTTCTTCAAGGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACCTGCCGGACCAACTGACTGAAGAGCAGATCGCAGAATTTAAAGAGGAATTCTC
CCTATTTGACAAGGACGGGGATGGGACAATAACAACCAAGGAGCTGGGGACGGTGATGC
GGTCTCTGGGGCAGAACCCACAGAAAGCAGAGCTGCAGGACATGATCAATGAAGTAGATG
CCGACGGTGACGGCACAATCGACTTCCCTGAGTTCCTGACAATGATGGCAAGAAAAATGA
AATACAGGGACACGGAAGAAGAAATTAGAGAAGCGTTCGGTGTGTTTGATAAGGATGGCA
ATGGCTACATCAGTGCAGCAGAGCTTCGCCACGTGATGACAAACCTTGGAGAGAAGTTAA
CAGATGAAGAGGTTGATGAAATGATCAGGGAAGCAGACATCGATGGGGATGGTCAGGTAA
ACTACGAAGAGTTTGTACAAATGATGACAGCGAAGTGA (SEQ ID NO: 101)

100/104

FIG. 17 cont'd

Channelopsin 1 [Mesostigma viride], UniProtKB - F8UVI5:

MSPPTSPTPDTGHDTPTDGHDTGGHGAVEICFAPCEEDCVTIRYFVENDFEGCIPGHFDQYSS
HGSLHDIVKAALYICMVISILQILFYGFQWWRKTCGWEVWFVACIETSIYIIAITSEADSPFTLYLT
NGQISPLRYMEWLMTCPVILIALSNITGMAEEYNKRTMTLLTSDVCCIVLGMMSAASKPRLKGI
LYAVGWAFGAWTYWTALQVYRDAHKAVPKPLAWYVRAMGYVFFTSWLTFFPGWFLLGPEGLE
VVTGTVSTLMHACSDLISKNLWGFMDWHLRVLVARHHRKLFKAEEEHALKKGQTLPEGMPRS
TSFVRGLGDDVEIDPSYELYRLKRQNHPEYFLSPAQTTPRRGPSFDKRTSFEMDGGKNGMLQM
MPVTGMGMGMGMGMGGGKTVLFLDYTGGGYVSFFEQQLSNMGMVNVTKCWSDDDDMYNTAG
VANVKQLFHFAMIPNNALGGQMVMMDLRGTGLLVVAYGPEPPMPGMGQDEFVPLQMPGVYPD
ESILHNLVMRHAITQGLGMNGMQGNMGQQQQMMGMQGNMNGMQGNMNGMQGNMNGMQ
GNMSGMQGNMNGMQGNSGMNQGWNNQGFTNTGAFGY (SEQ ID NO: 83)

Channelopsin 1 [Chlamydomonas yellowstonensis], GenBank: AER58217.1

MDTLAWWARELLSSGHGTDATDSGHGTDTSGGHDSHDAVAHNVTLLIAPPHAGGHAGPTD
TSQQITGIDGWAIAPAGDCYCAGWYVSHGSSFEATFAHVCQWSIFAVCVLSLLWYAYQYWKAT
CGWEEVYVCCIELVFICFELYHEFDSPCSLYLSTSNVNVWLRYSEWLLCCPVILIHLSNVTGLSD
DYGRRTMGLLVSDIATIVFGVTAAMLVNWPKIIFYLIGFTMCCYTFFLAAKVLIESFHQVPKGICR
HLVKAMAITYFVGWSFFPLIFLFGQSGFKKISPYADVIASSFGDLISKNAFGMLGHFLRVKIHEHIL
KHGDIRKTTHLRIAGEEKEVETFVEEDEDATAKHSTKELANRGSFIVMRDKMKEQGIDVRASLD
MDEDEEARTGKGKGAGATSLVPGRVILAVPDISMVDFFHDHFAHLGASIELVPALGVENTLLLV
QQAMQLGGLDFVLVHPEFLRDRSQNGLVSRCLKMTGHGVCAFGWWPSGPMREIIESAGVDGW
LDGPSFGTGIDQEQLIELIGYMQAKRKFGMRFGGGGASKAGYSSDGGFGGKGMLEMQPSMS
QGSQVPLLQQNNSMMRAPPSPMGNMANNGMMNPMMSMNNPMMGGGAVMMTSMGSMQQ
AANPLYGAPPSPLSSQPGAGMYGAPAPQPMGSQGSMSHGSMSYGGSSQQQHQQPQQAAAAAPA
AADGGSEAEMLKQLMSEINRLKAELGES (SEQ ID NO: 102)

Channelrhodopsin-2 [Volvox carteri f. nagariensis], UniProtKB - B4Y105:

MDHPVARSLIGSSYTNLNGSIVIPSDACFCMKWLKSKGSPVALKMANALQWAAFALSVIILYY
AYATWRTTCGWEEVYVCCVELTKVIEFFHEFDEPGMLYLANGNRVLWLRYGEWLLTCPVILI
HLSNLTGLKDDYNKRTMRLLVSDVGTIVWGATAAMSTGYIKVIFLLGCMYGANTFFHAAKVYIE
SYHTVPKGLCRQLVRAMAWLFFVSWGMFPVLFLLGPEGFGHLSVYGSTIGHTIIDLLSKNCWG
LLGHFLRLKIHEHILLYGDIRKVQKIRVAGEELEVETLMTEEAPDVTKKSTAQYANRESFLTMRD
KLKEKGFEVRASLDNSGIDAVINHHNNYNNALANAAAAGVKGPMELSKLDHVAANAAGMGGIA
DHVATTSGAISPGRVILAVPDISMVDYFREQFAQLPVQYEVVPALGADNAVQLVVQAAGLGCC
DFVLLHPEFLRDKSSTSLPARLRSIGQRVAAGWSPVGPVRDLIESAGLDGWLEGPSFGLGISL
PNLASLVLRMQHARKMAAMLGGMGMLGSLNLSGSGGVGLMGAGSPGGGGGAMGVGMTG
MGMVGTNAMGRGAVGNSVANASMGGGSAGMGMGMGMVGAGVGGQQQMGANGMGPTS
FQLGSNPLYNTAPSPPLSSQPGGDASAAAAAATGAASNSMNAMQAGGSVRNSGILAGG
LGSMMPGPPGAPAPTAAATAAPAVTMGAPGGGGAAASEAEMLQQLMAEINRLKSELGE (SEQ
ID NO: 84)

Channel rhodopsin 2 [synthetic construct], GenBank: ABO64386.1

MDYGGALSAVGRELLFVTNPVVVNGSVLPEDQCYCAGWIESRGTNGAQTASNVLQWLAAGF
SILLMFYAYQTWKSTCGWEEIYVCAIEMVKVILEFFFEKNPSMLYLATGHRVQWLYAEWLL
TCPVILIHLSNLTGLSNDYSRRTMGLLVSDIGTIVWGATSAMATGYVKVIFCLGLCYGANTFFH
AAKAYIEGYHTVPKGRCRQVVTGMAWLFFVSWGMFPILFILGPEGFGVLSVYGSTVGHITIDLM
SKNCWGLLGHYLRVLIHEHILIHGDIRKTKLNIGGTEIEVETLVEDEAEAGAVP (SEQ ID NO:
103)

FIG. 17 cont'd

CRISPR-associated protein (Cas) [*Streptococcus pyogenes*], GenBank: AKG27598.1:

MLEHKIDFMVTLEVKEANANGDPLNGNMPRTDAKGYGVMSDVSIRKIRNRLQDMGKSIFVQA
NERIEDDFRSLEKRFSQHFTAKTPDKEIEEKANALWFDVRAFGQVFTYLKKSIGVRGPVSISMA
KSLEPIVISSLQITRSTNGMEAKNNSGRSSDTMGTKHFVDYGVYVLKGSINAYFAEKTGFSQED
AEAIKEVLVSLFENDASSARPEGSMRVCEVFWFTHSSKLGNVSSARVFDLLEYHQSIEEKSTYD
AYQIHLNQEKLAKEYEAKGLTLEILEGL (SEQ ID NO: 85)

Cas9 [synthetic construct], GenBank: AST09977.1:

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLK
RTARRRYTRRKNRICYLQEFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYH
EKYPTIYHLRKKLV DSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQL
FEENPINASGVDAKILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED
AKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDE
HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVK
LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPLKDNREKIEKILTFRIPIYVVGPLARG
NSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQ
GDSLHEHIANLAGSPAIIKKGILQTVKVVDLVKVMGRHKPENIVIEMARENQTTQKGQKNSRER
MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLLQNGRDMYVDQELDINRLSDYDVDHIVPQ
SFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG
LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFY
KVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFF
YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTG
GFSKESILPKRNSDKLIARKKDWD PKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITI
MERSSF EKNPIDFLEAKGYKEVKKDLIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
NFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIEIQISEFSKRVLADANLDKVL SAYNKHR
DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLG
GD (SEQ ID NO: 86)

CRISPR-associated endonuclease Cpf1 [*Acidaminococcus* sp. BV3L6], UniProtKB/Swiss-Prot: U2UMQ6.1:

MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARN DHYKELKPIIDRIYKTYADQCL
QLVQLDWENLSAIDSYRKEKTEETR NALIEEQATYRNAIHDFIGRTDNLTDAINKRHAEIYKGL
FKAELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYENRKNVFS AEDISTAIPHRIVQDN
FPKFKENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLYNQLLGGI
SREAGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNTLSFILEEFKSDEEVIQSF
CKYKTLLRNENVLETA EALFNELNSIDLTHIFISHKKLETISSALCDHWDTLRNALYERRISELTGK
ITKSAKEKVQRSLKHEDINLQEII SAAGKELSEAFKQKTSEILSHAHAA LDQPLPTTLKKQEEKEIL
KSQLD SLLGLYHLLDWFAVDESNEVDPEFSARLTGIKLEMPSLSFY NKARNYATKKPYSVEKF
KLNFMPTLASGWDVNKEKNNGAILFVKNGLYYLGIMPKQKG RYKALSFEPT EKTSEGFDKMY
YDYFPDAAKMIPKCSTQLKAVTAHFQTH TTPILLSNNFIEPLEITKEIYDLNNPEKEPKKFQTAYA
KKTGDQKGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYHISFQR
IAEKEIMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKTSIKLNGQAELFYR
PKSRMKRMAHRLGEKMLNKKLKDQKTPIDTLYQELYDYVNHRLSHDLSDEARALLPNVITKEV
SHEIKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEHPETPIIGIDRGERNLIYITVIDS
TGKILEQRSLNTIQQFDYQKKLDNREK RVAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQ
AVVVLENLNF GFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYP AEKVGGLNPYQLTDQFTS

FIG. 17 cont'd

FAKMGTQSGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHESRKHFLLEGFDLHYDVKTGDFIL
HFKMNRNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPVIEHRFTGRYRDLYPA
NELIALLEEKGIVFRDGSNILPKLLENDSDHAIDTMVALIRSVLQMRNSNAATGEDYINSPVRDLN
GVCFDSRFQNPPEWPMADANGAYHIALKGQLLNHLKESKDLKLQNGISNQDWLAYIQELRN
(SEQ ID NO: 87)

Ribonuclease 4 or Ribonuclease L [Homo sapiens], UniProtKB/Swiss-Prot: Q05823.2:

MESRDHNNPQEGPTSSSGRRAAVEDNHLLIKAVQNEDVDLVQQLLEGGANVNFQEEEGGWT
PLHNAVQMSREDIVELLRLRHGADPVLRRKNGATPFILAAIAGSVKLLKFLSKGADVNECDFYGF
TAFMEAAYVGKVKALKFLYKRGANVNLRRKTKEDQERLRKGGATALMDAAEKGHVEVLKILLD
EMGADVNAACDNMGRNALIHALLSSDDSDVEAITHLLLDHGADVNVRRGERGKTPLILAVEKKHLG
LVQRLLQEHEIENDTSDGKTALLLAVELKLKIAELLCKRGASTDCGDLVMTARRNYDHSVLK
VLLSHGAKEDFHPPAEDWKPQSSHWGAALKDLHRIYRPMIGKLFIDEKYKIADTSEGGIYLG
FYEKQEVAVKTFCEGSPRAQREVSLQSSRENSHLVTFYGSSEHRGHLFVCVTLCEQTLEACL
DVHRGEDVNEEDEFARNVLSSIFKAVQELHLSCGYTHQDLQPQNILIDSKKAAHLADFDKSIK
WAGDPQEVKRDLEDLGRVLVYVKKGSISFEDLKAQSNEEVVQLSPDEETKDLIHLRHFHPGEH
VRDCLSDLLGHPFFWTWESRYRTLNRVGNESDIKTRKSESEILRLLQPGPSEHSKSFDKWTTKI
NECVMKKMNKFYEKRGNFYQNTVGDLLKFIKRLGEHIDEKHKMKLKGDPSTLYFQKTFPDLV
IYVYTKLQNTYRKHFQTHSPNKPQCDGAGGASGLASPGC (SEQ ID NO: 88)

Deoxyribonuclease II beta [Homo sapiens], GenBank: AAF76893.1:

MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFTFYKLPKRQNKESGETGLELYLDS
TTRSWRKSEQLMNDTKSVLGRTLQQLYEAYASKSNNTAYLIYNDGVKPKVNYSRKYGHTKGLL
LWNRVQGFWLIHSIPQFPPIPEEGYDYPPTGRRNGQSGICITFKYNQYEAIDSQLLVCNPNVYS
CSIPATFHQELIHMPQLCTRASSSEIPGRLLTTLQSAQGQKFLHFAKSDSFLDDIFAAWMAQRLK
THLLTETWQRKRQELPSNCSLPYHVYNIKAIKLSRHSYFSSYQDHAKWCISQKGTKNRWTCIG
DLNRSPHQAFRSGGFICTQNWQIYQAFQGLVLYYESCK (SEQ ID NO: 89)

Sodium channel protein type 1 subunit alpha [Homo sapiens], UniProtKB - P35498

MEQTVLVPPGPDSFNFFTTRESLAAIERIAEEKAKNPDPKDDDDENGPKPNSDLEAGKNLPFI
YGDIPPEMVSEPLEDLPYYINKKTFIVLNKGKAIFRFSATSALYILTPFNPLRKIAIKILVHSLFSML
IMCTILTNCVFMTMSNPDPWTKNVEYFTFTGIYTFESLIKIIARGFCLEDFTFLRDPWNWLDFTVIT
FAYVTEFVDLGNVSALRTFRVLRALKTISVIPGLKTIVGALIQSVKKLSVDMILTVFCLSVFALIGLQ
LFMGNLRNKCQWPPTNASLEEHSIEKNITVNYNGTLINETVFEFDWKSYYQDSRYHYFLEGFLD
ALLCGNSSDAGQCPEGYMCVKAGRNPNGYTSFDTFSWAFSLFRLMTQDFWENLYQLTLRA
AGKTYMIFVFLVIFLGSFYLINLILAVVAMAYEEQNQATLEEAQKEAEFQQMIEQLKKQEEAAQ
QAATATASEHSREPSAAGRLSDSSSEASKLSSSAKERRNRRKKRKQKEQSGGEEKDEDEFQ
KSESEDSIRRKGRFRFSIEGNRLTYEKRYSSPHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDVG
SENDFADDEHSTFEDNESRRDSLFPVRRHGERRNSNLSQTSRSSRMLAVFPANGKMHSTVD
CNGVVSLVGGPSVPTSPVGQLLPEVIIDKPATDDNGTTTETEMRKRRSSSFHVSMDFLEDPSQ
RQRAMSIASILTNTVEELEESRQKCPWCWYKFSNIFLIWDCSPYWLKVKHVVNLVMDPFVDLA
ITICIVLNTLFMAMEHYPMTHFNNVLTVGNLVFTGIFTAEMFLKIIAMPYFFQEGWNIFDGFIV
TSLSLVELGLANVEGLSVLRSFRLLRVFKLAKSWPTLNMLIKIIGNSVGALGNLTLVLAIVFIFAVVG
MQLFGKSYKDCVCKIASDCQLPRWHMNDFFHSFLIVFRVLCGEWIETMWDCMEVAGQAMCLT
VFMMVMVIGNLVVLNLFLALLSSFSADNLAATDDDNEMNQLQIAVDRMHKGVAYVVKRIYEFI
QQSFIRKQKILDEIKPLDDLNNKDKSCMSNHTAEIGKDLVDYLDVNGTTSGIGTGSSVEKYIDES
DYMSFINNPSLTVTVPIAVGESDFENLNTEDFSSESLEESKEKLNESSSSSSEGSTVDIGAPVEE
QPVVEPEETLEPEACFTEGCVQRFKCCQINVEEGRGKQWWNLRRTCFRIVEHNWFETFIVFMI
LLSSGALAFEDIYIDQRKTIKTMLEYADKVFTYIFILEMLLKWWAYGYQTYFTNAWCWLDLIVDV

FIG. 17 cont'd

SLVSLTANALGYSELGAIKSLRTRLRALRPLRALS RFEGMRVVVNALLGAIPSIMNVLLVCLIFWLIF
SIMGVNLFAGKFYHCINTTTGDRFDIEDVNNHTDCLKLIERNETARWKNVKVNFNDNVGFGYLSL
LQVATFKGWMDIMYAAVDSRNVELQPKYEESLYMYLYFVIFIFGSFFTLNLFIVGVIIDNFNQKK
KFGGQDIFMTEEQKKYYNAMKKLGSKKPQKPIPRPGNKFQGMVFDVTRQVFDISIMILICLNM
VTMMVETDDQSEYVTTILSRINLVFIVLFTGECVLKLSLRHYFTIGWNIFDFVVLVLSIVGMFLAE
LIEKYFVSPTLFRVIRLARIGRILRLIKGAKGIRTLLFALMMSLPALFNIGLLLFLVMFIYAIFGMSNF
AYVKREVGIDDMFNFFETFGNSMICLFQITTSAGWDGLLAPILNSKPPDCDPNKVNPSSSVKGDC
GNPSVGIFFFVSYIIISFLVVVNMYIAVILENFSVATEESAEPSEDDFEMFYEVWEKFDPDATQF
MEFEKLSQFAAALEPPLNLPQPNKLQLIAMDLPMSVSGDRIHCLDILFAFTKRVLGESGEMDALRI
QMEERFMASNP SKVSYPITTTTLKRKQEEVSAVIIQRAYRRHLLKRTVKQASFTYNKNKIKGGA
NLLIKEDMIIDRINENSITEKDTLTMSTAACPPSYDRVT KPIVEKHEQEGKDEKAKGK (SEQ ID
NO: 104)

Potassium voltage-gated channel subfamily KQT member 2 [Homo sapiens], UniProtKB - O43526

MVQKSRNGGVYPGPSGEKKLVGFVGLDPGAPDSTRDGALLIAGSEAPKRGSIKSKPRAGGA
GAGKPPKRNAFYRKLQNFLYNVLERPRGWAFIYHAYVFLLVFSLVSVFSTIKEYESSEGAL
YILEIVTIVVFGVEYFVRIWAAGCCCRYRGWRGRLKFARKPFCVIDIMVLIAVLAAGSQGNVF
ATSALRSLRFLQILRMIRMDRRGGTWKLLGSVYAHSKELVTAWYIGFLCLILASFLVYLAEKGE
NDHFDTYADALWWGLITLTIGYGDYKYPQTWNGRLLAATFTLIGVSFFALPAGILGSGFALKVQE
QHRQKHFEKRRNPAAGLIQSAWRFYATNLSRTDLHSTWQYYERTVTVP MYSSQTQTYGASRL
IPPLNQLELLRNLSKSKGLAFRKDPPPEPSPSKSGSPCRGPLCGCCPGRSSQKVSLKDRVFSSP
RGVAAKGKGSPPQAQT VRRSPSADQSLEDSPSKVPKSWSF GDRSRARQAFRIKGAASRQNSE
EASLPGEDIVDDKSCPCFEVTEDLTPGLKVSIRAVCVMRFLVSKRKFKESLRPYDVM DVIEQYS
AGHLDMLSRIKSLQSRVDQIVGRGPAITDKDRTKGPAAEALPEDPSMMGR LGKVEKQVLSMEK
KLDFLVNIYMQRMGIPPTETEAYFGAKEPEPAPPYHSPEDSREHVDRHGCIVKIVRSSSSTGQK
NFSAPPAAPPVQCPPSTSWQPQSHPRQGHGTSPVGDHGS LVRIPPPPAHERSLSAYGGGNR
ASMEFLRQEDTPGCRPPEGNLRDSDTSISIPSV DHEELERSFSGFSISQSKENLDALNSCYAAV
APCAKVRPYIAEGESD TSDLC TPCGPPPRSATGEGPFGDVGWAGPRK (SEQ ID NO: 105)

Voltage-dependent L-type calcium channel subunit alpha-1C [Homo sapiens], UniProtKB - Q13936

MVNENTRMYIPEENHQGSNYGSPRPAHANMNANAAAGLAPEHIPTPGAALSWQAAIDAARQA
KLMGSAGNATISTVSSTQRKRQQYGKPKKQGSTTATRP RALLCLTLKNPIRRACISIVEWKPF
EIIILLTIFANCVALAIYIPFPEDDSNATNSNLERVEYLFLLIIFTVEAFLKVIAYGLLFHPNAYLRNGW
NLLDFIIVVVG LFS AILEQATKADGANALGGKGAGFDVKALRAFRVLRPLRLVSGVPSLQVVLNSI
IKAMVPLLHIALLVLFVIIYAIIGLELFMGKMHKTCYNQEGIADVPAEDDPSPCALETGHGRQCQN
GTVCKPGWDGPKHGITNFDNFAFAMLT VFQCITMEGWTDVLYWVNDAVGRDWPWIYFVTLIII
GSFFVLNLVLGVLSGEFSKEREKAKARGDFQKLREKQQLEEDLKGYLDWITQAEDIDPENED
GMDEEKPRNMSMPTSETESVNTENVAGGDIEGENCARLAHRISKSKFSRYWRRWNRFCRR
KCRAAVKSNVFYWLVI FLVFLNTLTIASEHYNQPNWLTEVQDTANKALLALFTAEMLLKMYSLGL
QAYFVSLFNRFD CFVVC GGILETILVETKIMSPLGISVLRCVRLLRIFKITRYWNSLSNLVASLLNS
VRSIASLLLLLFLFIIFSL LGMQLFGGKFN FDEMQTRRSTFDNFPQSLLTVFQILTGEDWNSVMY
DGIMAYGGPSFPGMLVCIYFIILFICGNYILLNVFLAIAVDNLADAESL TSAQKEEEEEKERKKLAR
TASPEKKQELVEKPAVGESKEEKIELKSITADGESPPATKINMDDLQPNENEDKSPYPNPETT
EEDDEEPEMPVGPRPRPLSELHLKEKAVPMPEASAFFIFSSNNRFR LQCHRIVNDTIFTNLILFFI
LLSSISLAAEDPVQHTSFRNHILFYFDIVFTTIFTIEIALKILGNADYVFTSIFTLEIILKMTAYGAFLH
KGSFCRNYFNILDLLVSVSLISFGIQSSAINVVKILRVLRVLRPLRAINRAKGLKHVVQC VFVAIR
TIGNIVIVTLLQFMFACIGVQLFKGKLYTCSDSSKQTEAECKGNYITYKDGEVDHP IIQPRSWEN

104/104

FIG. 17 cont'd

SKFDFDNVLAAMMALFTVSTFEGWPELLYRSIDSHTEDKGPIYNYRVEISIFFIYIIIIIAFFMMNIFV
GFVIVTFQEQGEQEYKNCELDKNQRQCVEYALKARPLRRYIPKNQHQYKVWYVNVNSTYFEYL
MFVLILLNTICLAMQHYGQSCLFKIAMNILNMLFTGLFTVEMILKLIAFKPKGYFSDPWNVFDLIV
IGSIIDVILSETNHYFCDAWNTFDALIVVGSIVDIAITEVNPAEHTQCSPSMNAEENSRSISITFFRLF
RVMRLVKLLSRGEGIRTLLWTFIKSFQALPYVALLIVMLFFIYAVIGMQVFGKIALNDTTEINRNNN
FQTFPQAVLLLFRCATGEAWQDIMLACMPGKKCAPESEPSNSTEGETPCGSSFAVFYFISFYM
LCAFLIINLFAVIMDNFDYLTRDWSILGPHHLDEFKRIWAEYDPEAKGRIKHLDVVTLLRRIQPPL
GFGKLCPHRVACKRLVSMNMPLNSDGTVMFNATLFALVRTALRIKTEGNLEQANEELRAIIKI
WKRTSMKLLDQVPPAGDDEVTVGKFYATFLIQEYFRKFKKRKEQGLVGKPSQRNALS LQAGL
RTLHDIGPEIRRAISGDLTAEELDKAMKEAVSAASEDDIFRRAGGLFGNHVSYYQSDGRSAFP
QTFTTQRPLHINKAGSSQGDTESPSHEKLVDSTFTPSSYSSTGSNANINNANNTALGRLPRPA
GYPSTVSTVEGHGPPLSPAIRVQEVAWKLSSNRERHVPMCEDLELRRDSGSAGTQAHCLLLR
KANPSRCHSRESQAAMAGQEETSQDETYEVKMNHDTEACSEPSLLSTEMLSYQDDENRQLTL
PEEDKRDIRQSPKRGFLRSASLGRRASFHLECLKRQKDRGGDISQKTVLPLHLVHHQALAVAG
LSPLLQRSHSPASFPRPFATPPATPGSRGWPPQPVP TLRLEGVESSEKLNSSFPSIHCGSWAE
TTPGGGGSSAARRVRPVSLMVPSQAGAPGRQFHGSASSLVEAVLISEGLGQFAQDPKFIEVTT
QELADACDMTIEEMESAADNILSGGAPQSPNGALLPFVNCRDAGQDRAGGEEDAGCVRARGR
PSEELQDSRVYVSSL (SEQ ID NO: 106)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US19/54539

A. CLASSIFICATION OF SUBJECT MATTER

IPC - C12N 15/09, 15/63, 15/86, 15/864 (2020.01)

CPC - C12N 15/09, 15/63, 15/86

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)

See Search History document

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

See Search History document

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

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2004/0087028 A1 (CUNNINGHAM) 06 May 2004; paragraphs [0012], [0018], [0060], [0051], [0070], [0071]	1-7, 9, 12-22, 24, 27, 28/1, 28/4, 28/20, 29/28/1, 29/28/4, 29/28/20, 30/28/1, 30/28/4, 30/28/20, 31/28/1, 31/28/4, 31/28/20
A	US 2018/0078658 A1 (FISHELL et al.) 22 March 2018; paragraph [0039]	1-7, 9, 12-22, 24, 27, 28/1, 28/4, 28/20, 29/28/1, 29/28/4, 29/28/20, 30/28/1, 30/28/4, 30/28/20, 31/28/1, 31/28/4, 31/28/20
A	'PREDICTED: Panthera pardus uncharacterized LOC109263725 (LOC109263725), transcript variant X2, ncRNA, Accession XR_002078354'; Publication [online]. 05 December 2016; [retrieved 27 January 2020]. Retrieved from the Internet: <URL: https://www.ncbi.nlm.nih.gov/nucleotide/XR_002078354.1?report=genbank&log\$=nucldp&blast_rank=17&RI=D=30VXJA38016 >; pp 1-2; page 1	1-7, 9, 12-22, 24, 27, 28/1, 28/4, 28/20, 29/28/1, 29/28/4, 29/28/20, 30/28/1, 30/28/4, 30/28/20, 31/28/1, 31/28/4, 31/28/20

☐ 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

"D" document cited by the applicant in the international application

"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

03 February 2020 (03.02.2020)

Date of mailing of the international search report

14 FEB 2020

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US19/54539

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 32-51
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

-Please See Supplemental Page-

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-3, 5-7, 9, 12-22, 24, 27, 28/1, 28/4, 28/20, 29/28/1, 29/28/4, 29/28/20, 30/28/1, 30/28/4, 30/28/20, 31/28/1, 31/28/4, 31/28/20; SEQ ID NOS.: 2, 3

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US19/54539

-***-Continued from Box No. III Observations where unity of invention is lacking: -***-

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Groups I+, Claims 1-31, SEQ ID NO: 2 (enhancer sequence), SEQ ID NO: 3 (enhancer concatemer sequence); and an ion transporter (heterologous encoding sequence) are directed toward an artificial expression construct, an AAV vector comprising the expression construct; and transgenic cells associated therewith.

The expression constructs, vectors, and cells will be searched to the extent they encompass SEQ ID NO: 2 (first exemplary enhancer sequence), SEQ ID NO: 3 (first exemplary enhancer concatemer sequence) and an ion transporter (heterologous encoding sequence). Applicant is invited to elect additional enhancer sequence(s), and associated concatemers thereof, with specified SEQ ID NO: for each, or with specified substitution(s) at specified site(s) of a SEQ ID NO:; such that the sequence of each elected species is fully specified (i.e. no optional or variable bases or substituents), and/or heterologous encoding sequence(s), where available as an option within at least one searchable claim, to be searched. Additional enhancer sequence(s) and/or heterologous encoded sequence(s) will be searched upon the payment of additional fees. It is believed that claims 1, 2-7 (each in-part), 9 (in-part), 12-22 (each in-part), 24 (in-part), and 27-31 (each in-part) encompass this first named invention and thus these claims will be searched without fee to the extent that they encompass SEQ ID NO: 2 (enhancer sequence), SEQ ID NO: 3 (enhancer concatemer sequence); and a fluorescent reporter protein (heterologous encoding sequence). Applicants must specify the searchable claims that encompass any additionally elected enhancer sequence(s) and/or heterologous encoding sequence(s). Applicants must further indicate, if applicable, the claims which encompass the first named invention, if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined. An exemplary election would be SEQ ID NO: 6 (enhancer sequence); and SEQ ID NO: 7 (enhancer concatemer sequence).

No technical features are shared between the enhancer sequences and/or heterologous sequences of Groups I+ and, accordingly, these groups lack unity a priori.

Groups I+ share the technical features including: an artificial expression construct comprising (i) an enhancer having a sequence, or a concatemer thereof; (ii) a promoter; and (iii) a heterologous encoding sequence encoding an effector element; an adeno-associated viral (AAV) vector comprising the artificial expression construct; wherein the heterologous encoding sequence is under the transcriptional control of a promoter and the enhancer; and a transgenic cell comprising the artificial expression construct or vector.

However, these shared technical features are previously disclosed by US 2004/0087028 A1 (CUNNINGHAM).

Cunningham discloses an artificial expression construct (a recombinant expression vector (an artificial expression construct); paragraphs [0018], [0060]) comprising (i) an enhancer having a sequence, or a concatemer thereof (an enhancer comprising three or more SP1 elements (an enhancer having a sequence, or a concatemer thereof); paragraph [0071]); (ii) a promoter (a promoter; paragraph [0070]); and (iii) a heterologous encoding sequence encoding an effector element (a heterologous encoding sequence encoding an effector element; paragraphs [0012], [0018], [0060]); an adeno-associated viral (AAV) vector comprising the artificial expression construct (an adeno-associated viral (AAV) vector comprising the artificial expression construct; paragraph [0060]); wherein the heterologous encoding sequence is under the transcriptional control of a promoter and the enhancer (wherein the heterologous encoding sequence is under the transcriptional control of a promoter and the enhancer; paragraphs [0051], [0070], [0071]); and a transgenic cell comprising the artificial expression construct or vector (a transgenic cell comprising the artificial expression construct or vector; paragraphs [0018], [0060]).

Since none of the special technical features of the Groups I+ inventions is found in more than one of the inventions, and since all of the shared technical features are previously disclosed by the Cunningham reference, unity of invention is lacking.