

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

20 August 2020 (20.08.2020)



(10) International Publication Number

WO 2020/168279 A2

(51) International Patent Classification:

A01K 67/027 (2006.01)

(21) International Application Number:

PCT/US2020/018416

(22) International Filing Date:

14 February 2020 (14.02.2020)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/806,686 15 February 2019 (15.02.2019) US

62/806,660 15 February 2019 (15.02.2019) US

62/874,859 16 July 2019 (16.07.2019) US

(71) Applicant: ALLEN INSTITUTE [US/US]; 615 Westlake Ave N., Seattle, Washington 98109 (US).

(72) Inventors: TING, Jonathan; 615 Westlake Ave N., Seattle, Washington 98109 (US). LEVI, Boaz P.; 615 Westlake Ave N., Seattle, Washington 98109 (US). TASIC, Bosiljka; 615 Westlake Ave N., Seattle, Washington 98109 (US). MICH, John K.; 615 Westlake Ave N., Seattle, Washington 98109 (US). HESS, Erik; 615 Westlake Ave N., Seattle, Washington 98109 (US). LEIN, Edward Sebastian; 615 Westlake Ave N., Seattle, Washington 98109 (US). GRAY-BUCK, Lucas T.; 615 Westlake Ave N., Seattle, Washington 98109 (US). DAIGLE, Tanya; 615 Westlake Ave N., Seattle, Washington 98109 (US). ZENG, Hongkui; 615 Westlake Ave N., Seattle, Washington 98109 (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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

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

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:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))

(54) Title: ARTIFICIAL EXPRESSION CONSTRUCTS FOR SELECTIVELY MODULATING GENE EXPRESSION IN SELECTED NEURONAL CELL POPULATIONS

(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 neurons generally; and/or GABAergic neuron cell types such as lysosomal associated membrane protein 5 (Lamp5) neurons; vasoactive intestinal polypeptide-expressing (Vip) neurons; somatostatin (Sst) neurons; and/or parvalbumin (Pvalb) neuron cell types. Certain artificial expression constructs additionally drive selective gene expression in Layer 4 and/or layer 5 intratelencephalic (IT) neurons, deep cerebellar nuclear neurons or cerebellar Purkinje cells.



WO 2020/168279 A2

ARTIFICIAL EXPRESSION CONSTRUCTS FOR
SELECTIVELY MODULATING GENE EXPRESSION IN SELECTED NEURONAL CELL
POPULATIONS

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to U.S. Provisional Patent Application No. 62/806,660 filed February 15, 2019, U.S. Provisional Patent Application No. 62/806,686 filed February 15, 2019, and U.S. Provisional Patent Application No. 62/874,859 filed July 16, 2019 each of which is incorporated 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 MH114126 and DA036909 awarded by the National Institutes of Health. The government has certain rights in the invention.

FIELD OF THE DISCLOSURE

[0003] 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 gamma-aminobutyric acid (GABA)ergic neurons generally; and/or GABAergic neuron cell subclasses such as lysosomal associated membrane protein 5 (Lamp5) neurons; vasoactive intestinal polypeptide-expressing (Vip) neurons; somatostatin (Sst) neurons; and/or parvalbumin (Pvalb) neurons. Certain artificial expression constructs additionally drive selective gene expression in Layer 4 and/or layer 5 intratelencephalic (IT) neurons, or non-neocortical neurons like deep cerebellar nuclear Pvalb-positive neurons or cerebellar Purkinje cells.

BACKGROUND OF THE DISCLOSURE

[0004] To fully understand the biology of the brain, different cell types need to be distinguished and defined and, to further study them, artificial expression constructs 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.

SUMMARY OF THE DISCLOSURE

[0005] The current disclosure provides artificial expression constructs that selectively drive gene expression in targeted central nervous system cell populations. Targeted central nervous system cell populations include: gamma-aminobutyric acid (GABA)ergic neurons generally; and/or GABAergic neuron cell subclasses such as lysosomal associated membrane protein 5 (Lamp5) neurons; vasoactive intestinal polypeptide-expressing (Vip) neurons; somatostatin (Sst) neurons; and/or parvalbumin (Pvalb) neurons. Layer 4 and/or layer 5 intratelencephalic (IT) neurons, or non-neocortical neurons like deep cerebellar nuclear Pvalb-positive neurons or cerebellar Purkinje cells can also be targeted for selective gene expression.

[0006] Particular embodiments of the artificial expression constructs utilize the following enhancers to selectively drive protein expression within targeted central nervous system cell populations as follows (enhancer / targeted cell population): Grik1_enhGad2-1 / GABAergic neurons generally; Grik1_enhGad2-2 / GABAergic neurons generally; mscRE5 / GABAergic neurons generally; mscRE8 / GABAergic neurons generally; eHGT_019h / Lamp5 neurons; eHGT_022h / Lamp5 and Vip neurons; eHGT_022m / Lamp5 and Vip neurons; eHGT_017h / Lamp5, Vip, and Sst neurons; eHGT_17m / Lamp5, Vip, and Sst neurons; eHGT_079h / parvalbumin (Pvalb) neuron cell types; eHGT_082h / Pvalb neuron cell types in cortex and deep cerebellar nucleus neurons; eHGT_086h / Pvalb neuron cell types; eHGT_128h / Pvalb neuron cell types; eHGT_140h / Pvalb neuron cell types; eHGT_064h / Pvalb and Sst neuron cell types; eHGT_023h / Pvalb cell types, L4 and L5 IT neurons, and Purkinje cells; and eHGT_359 / Pvalb cell types and cerebellar Purkinje cells.

[0007] Particular embodiments provide artificial expression constructs including the features of vectors described herein including vectors: AiP1146, AiP1113, AiP1147, AiP1147, AiP1013, AiP1012, CN1525, CN1528, CN1532, CN1621, CN1633, CN1259, CN2045, CN1255, CN1408, CN1258, CN1279, CN1253, and CN1274.

BRIEF DESCRIPTION OF THE FIGURES

[0008] Some of the drawings submitted herein are better understood in color. Applicant considers the color versions of the drawings as part of the original submission and reserves the right to present color images of the drawings in later proceedings. For example, FIGS. 3A, 3B, 5, 6A, 6B, 7A, 7B, 8A, 8B, 9A, 9C, 10A, 11A, 11B, 12A, 13A, 13B, 14A, 14B, 14D, 15A, 15B, 16A, 16B, 17A, 17B, 18A, 18B, 19A, 19B, 19C, and 20, described below, reflect color labeling assays that have been presented as black and white images.

[0009] FIG. 1: Overview of enhancer discovery for viral tools. To build cell type-specific labeling

tools, cells from adult mouse cortex were isolated and a single cell assay for transposase-accessible chromatin using sequencing (scATAC-seq) was performed. Samples were clustered and compared to single cell RNA sequencing (scRNA-seq) datasets to identify the clusters. Single cells matching the same transcriptomic types were then pooled and the genome was searched for type-specific putative enhancers. These regions were cloned upstream of a minimal promoter in an AAV genomic backbone, which was used to generate self-complementary adeno-associated viral vectors (scAAVs) or recombinant adeno-associated viral vectors (rAAVs). These viral tools were delivered retro-orbitally to label specific GABAergic neurons populations. In cells with a matching cell type, enhancers recruit their cognate transcription factors to drive cell type-specific expression. In other cells, viral genomes are present, but transcripts are not expressed.

[0010] FIGs. 2A, 2B. vAi30.0 (AiP1146) with Grik1_enhGad2-1 enhancer. (2A) Schematic diagram of the enhancer-containing viral vector with all major components denoted. ITR = inverted terminal repeat, Hsp68 = heat shock protein 68 minimal promoter, WPRE3 = woodchuck post-transcriptional regulatory element 3, BGHpA = bovine growth hormone polyA, and Grik1_enhGAD2-1 = eAi12.0 (MGT_E31) enhancer. (2B) Purified AiV1146 virus was injected into the primary visual cortex of a Gad2-IRES-Cre; Ai14 animal and expression of the transgenes was analyzed in fixed brain sections several weeks post-injection. The tdTomato labels the pan-GABAergic interneuron population. Co-labeling of EGFP with tdTomato was observed in many cells (see merge image), which confirms that this virus labels a subset of GABAergic neurons.

[0011] FIGs. 3A, 3B. vAi30.1 (AiP1113) with Grik1_enhGad2-1 enhancer. Purified AiV1113 virus was injected into the retro-orbital sinus of a C57BL/6J wild-type mice and EGFP expression was analyzed in fixed brain sections two weeks post-injection. (3A) Native fluorescence and (3B) fluorescence enhanced by staining with an anti-GFP antibody are shown. GFP positive labelled neurons were scattered throughout the cortex and exhibited the typical aspiny dendrite morphology, a hallmark of cortical interneurons.

[0012] FIG. 4. vAi31.0 (AiP1147) with Grik1_enhGad2-2 enhancer. Purified AiV1147 virus was injected into the primary visual cortex of a Gad2-IRES-Cre; Ai14 animal and expression of the transgenes was analyzed in fixed brain sections several weeks post-injection. The tdTomato labels the pan-GABAergic interneuron population. Co-labeling of EGFP with tdTomato was observed in many cells (see merge image), which confirms that this virus labels a subset of GABAergic neurons.

[0013] FIG. 5. vAi11.0 (AiP989) with mscRE5 enhancer. Purified AiV989 virus was injected into

the retro-orbital sinus of a C57BL/6J wild-type mice and EGFP expression was analyzed in fixed brain sections two weeks post-injection. GFP positive neurons were observed to be scattered throughout the cortex and exhibited the typical aspiny dendrite morphology, a hallmark of cortical interneurons. Labelled neurons of similar morphology were also observed in other subcortical brain structures.

[0014] FIGs. 6A, 6B. vAi12.0 (AiP1013) with mscRE5 enhancer. (6A) Purified AiV1013 virus was injected into the retro-orbital sinus of an Ai65F mouse (Ai65F mouse is a Flp-dependent tdTomato reporter mouse line) and tdTomato expression was analyzed in fixed brain sections two weeks post-injection. (6B) Enlarged image of boxed area indicated in FIG. 6A. tdTomato positive neurons were observed to be scattered throughout the cortex and exhibited the typical aspiny dendrite morphology, a hallmark of cortical interneurons. Labelled neurons of both similar and diverse morphology were also observed in many other subcortical brain structures.

[0015] FIGs. 7A, 7B. vAi14.0 (AiP1012) with mscRE8 enhancer. (7A) Purified AiV1012 virus was injected into the retro-orbital sinus of an Ai65F mouse (Ai65F mouse is a Flp-dependent tdTomato reporter mouse line) and tdTomato expression was analyzed in fixed brain sections two weeks post-injection. (7B) Enlarged image of boxed area indicated in FIG. 7A. tdTomato positive neurons were observed to be scattered throughout the cortex and exhibited the typical aspiny dendrite morphology, a hallmark of cortical interneurons. Labelled neurons of similar morphology were also observed in many other subcortical brain structures.

[0016] FIGs. 8A-8C. (8A) Fluorescence expression of CN1525 (eHGT_019h), in black, shown in whole mouse brain in sagittal section. (8B) Native SYFP2 fluorescence image of a live slice of V1 shows sparse cortical labeling. (8C) Single cell transcriptomic characterization of SYFP2 fluorescent cells isolated from mouse V1. After single cell gene expression analysis, cells were mapped to an existing taxonomy of mouse V1 cell types. The plot shows single cells grouped by subtype. (8D) Dendrogram shows the mapping of each single cell to the terminal branch of the mouse taxonomy, if possible. The location of each circle reflects the extent of single cell mapping (toward the terminal branch), while size of the circle reflects the number of single cells that mapped to that point in the hierarchy. Bars projecting down reflect the number of cells that map to that terminal branch of the cell type taxonomy. Note that the majority of cells are Lamp5+.

[0017] FIGs. 9A-9D. (9A) Fluorescence (white) image of CN1258 (eHGT_022h) in a live slice of mouse V1 shows sparse cortical labeling. (9B) Quantification of three replicates of the overlap of CN1258-driven SYFP2 expression with antibody markers of GABAergic neuron types Lamp5, Vip, Sst and Pvalb. (9C) CN1258 labeling of human organotypic slice tissue ex vivo shows

enrichment of SYFP2 in upper layers of neocortex indicating an enrichment in LAMP5 and VIP cells. (9D) Single cell transcriptomic characterization of SYFP2 fluorescent cells isolated from human MTG (top) and mouse V1 (bottom). After single cell gene expression analysis, cells were mapped to existing taxonomies of human MTG cell types or mouse V1 cell types. The plots show the mapping of each single cell to either taxonomy, as described in relation to FIG. 8D. Note that the majority of cells are Lamp5+ or Vip+ for both species.

[0018] FIGs. 10A, 10B. (10A) Fluorescence (white) image of CN1279 (eHGT_022m) in a live slice of mouse V1 shows sparse cortical labeling. (10B) Quantification of three replicates of the overlap of CN1259-driven SYFP2 expression with antibody markers of GABAergic neuron types Lamp5, Vip, Sst and Pvalb.

[0019] FIGs. 11A-11C. (11A) Fluorescence expression of CN1253 (eHGT_017h), in black, shown in whole mouse brain in sagittal section. (11B) High resolution images showing overlap of CN1253 SYFP2 fluorescence with GABAergic markers Gad1, Sst and Lamp5 mRNA expression. The arrows identify SYFP-labeled cells. (11C) Single cell transcriptomic characterization of SYFP2 fluorescent cells isolated from mouse V1. After single cell gene expression analysis, cells were mapped to an existing taxonomy of mouse V1 cell types, as described in relation to FIG. 8D. Note that the majority of cells are Lamp5+, Vip+, or Sst+.

[0020] FIGs. 12A, 12B. (12A) Fluorescence expression of CN1274 (eHGT_017m), in black, shown in whole mouse brain in sagittal section. (12B) Single cell transcriptomic characterization of SYFP2 fluorescent cells isolated from mouse V1. After single cell gene expression analysis, cells were mapped to an existing taxonomy of mouse V1 cell types, as described in relation to FIG. 8D. Note that the majority of cells are Lamp5+, Vip+, or Sst+.

[0021] FIGs. 13A-13C. (13A) Fluorescence expression of CN1525 (eHGT_079h), in black, shown in whole mouse brain in sagittal section. (13B) High resolution images showing overlap of CN1525 SYFP2 fluorescence with GABAergic markers Gad1 and Pvalb mRNA expression. The arrows identify SYFP2-labeled cells. (13C) Single cell transcriptomic characterization of SYFP2 fluorescent cells isolated from mouse V1. After single cell gene expression analysis, cells were mapped to an existing taxonomy of mouse V1 cell types, as described in relation to FIG. 8D. Note that nearly all cells are types of Pvalb neurons.

[0022] FIGs. 14A-14D. (14A) Fluorescence expression of CN1528 (eHGT_082h), in black, shown in whole mouse brain in sagittal section. (14B) High resolution images showing overlap of CN1528 SYFP2 fluorescence with GABAergic markers Gad1 and Pvalb mRNA expression. The arrows identify SYFP2-labeled cells. The arrows highlight several SYFP2+ cells. (14C) Single cell transcriptomic characterization of SYFP2 fluorescent cells isolated from mouse V1.

After single cell gene expression analysis, cells were mapped to an existing taxonomy of mouse V1 cell types, as described in relation to FIG. 8D. Note that nearly all cells are types of Pvalb neurons. (14D) Pvalb-positive glutamatergic (Gad1 negative) and GABAergic (Gad1 positive) cells in the deep cerebellar nucleus are labeled by SYFP2 after intravenous administration of CN1528 packaged by PHP.eB.

[0023] FIGs. 15A, 15B. (15A) Fluorescence expression of CN1532 (eHGT_086h), in black, shown in whole mouse brain in sagittal section. (15B) High resolution images showing overlap of CN1532 SYFP2 fluorescence with GABAergic markers Gad1 and Pvalb mRNA. The arrows identify SYFP2-labeled cells.

[0024] FIGs. 16A-16C. (16A) Fluorescence expression of CN1621 (eHGT_128h), in black, shown in whole mouse brain in sagittal section. (16B) High resolution images showing overlap of CN1621 SYFP2 fluorescence with GABAergic marker Pvalb mRNA expression. The arrows identify SYFP2-labeled cells. (16C) Single cell transcriptomic characterization of SYFP2 fluorescent cells isolated from mouse V1. After single cell gene expression analysis, cells were mapped to an existing taxonomy of mouse V1 cell types, as described in relation to FIG. 8D. Note that nearly all cells are types of Pvalb neurons.

[0025] FIGs. 17A-17C. (17A) Fluorescence expression of CN1633 (eHGT_140h), in black, shown in whole mouse brain in sagittal section. (17B) High resolution images showing overlap of CN1633 SYFP2 fluorescence with GABAergic markers Gad1 and Pvalb mRNA expression. The arrows identify SYFP2-labeled cells. (17C) Single cell transcriptomic characterization of SYFP2 fluorescent cells isolated from mouse V1. After single cell gene expression analysis, cells were mapped to an existing taxonomy of mouse V1 cell types, as described in relation to FIG. 8D. Note that nearly all cells are types of Pvalb neurons

[0026] FIGs. 18A-18C. (18A) Fluorescence expression of CN1408 (eHGT_064), in black, shown in whole mouse brain in sagittal section. (18B) High resolution images showing overlap of CN1408 SYFP2 fluorescence with GABAergic markers Pvalb or Sst mRNA expression. The arrows identify SYFP2-labeled cells that co-label with Pvalb or Sst, while the asterisks mark cells labeled cells that are not co-labeled by Pvalb or Sst. (18C) Single cell transcriptomic characterization of SYFP2 fluorescent cells isolated from mouse V1. After single cell gene expression analysis, cells were mapped to an existing taxonomy of mouse V1 cell types, as described in relation to FIG. 8D. Note that nearly all recovered cells are types of Pvalb or Sst neurons.

[0027] FIGs. 19A-19C. (19A) Fluorescence expression of CN1259 (eHGT_023h), in black, shown in whole mouse brain in sagittal section. Strong expression is seen in the neocortex and

non-neocortical brain regions such as the cerebellum (19B) High resolution images show overlap of CN1259 SYFP2 fluorescence with GABAergic markers Gad1, Vip and Pvalb mRNA expression. The arrows identify SYFP2-labeled cells. Note that most cells overlap with Gad1, and many cells overlap with Pvalb. (19C) Pvalb-positive Purkinje cells (Gad1 and Pvalb positive) in the cerebellum are labeled by SYFP2 after intravenous administration of CN1259 packaged with PHP.eB.

[0028] FIG. 20. Fluorescence expression of CN2045 (eHGT_359h), in black, shown in whole mouse brain in sagittal section. Expression in cortex and hippocampus indicates Pvalb expression and there is strong labeling of cerebellar Purkinje cells.

[0029] FIGs. 21A, 21B. (21A) Table describing the components included in each vector sequence to summarize the vector name and length, enhancer, promoter, product class, primary product, and other components of the vector. (21B) Cell type specificity of enhancers and vectors are summarized. Origin species is indicated where H indicates Human and M indicates mouse. Cell type specificity is indicated where S = subset of types in group and A = all types in group. The validation method is indicated by: * = tested and validated in mouse, RNA-seq, and another modality; ~ = tested and validated in mouse and primate/human, RNA-seq and another modality; ^ = tested and validated in mouse with at least one validation method; and + = tested and awaiting additional validation. The column labeled Method of Validation describes the validation methods where T indicates validation methods by tissue expression, R indicates validation methods by single cell RNAseq, I indicates validation methods by immunohistochemistry or mFISH, and TG indicates validation methods by tissue expression and genetic labeling.

[0030] FIG. 22. Sequences supporting the disclosure. Sequences for Enhancer Grik1_enhGad2-1 (eAi12.0; MGT_E31) (SEQ ID NOs: 1 and 42), Enhancer Grik1_enhGad2-2 (eAi13.0; MGT_E65) (SEQ ID NO: 2), Enhancer mscRE5 (eAi4.0; MGT_E5) (SEQ ID NO: 3), Enhancer mscRE8 (eAi5.0; MGT_E8) (SEQ ID NO: 4), Enhancer eHGT_079h (eAi115.0) (SEQ ID NO: 5), Enhancer eHGT_082h (eAi116.0) (SEQ ID NO: 6), Enhancer eHGT_086h (eAi117.0) (SEQ ID NO: 7), Enhancer eHGT_128h (eAi119.0) (SEQ ID NO: 8), Enhancer eHGT_140h (eAi120.0) (SEQ ID NO: 9), Enhancer eHGT_023h (eAi104.0) (SEQ ID NO: 10), Enhancer eHGT_359h (SEQ ID NO: 11), Enhancer eHGT_019h (eAi101.0) (SEQ ID NO: 12), Enhancer eHGT_064h (eAi110.0) (SEQ ID NO: 13), Enhancer eHGT_022h (eAi103.0) (SEQ ID NO: 14), Enhancer eHGT_022m (eAi102.0) (SEQ ID NO: 15), Enhancer eHGT_017h (eAi100.0) (SEQ ID NO: 16), Enhancer eHGT_017m (SEQ ID NO: 17), hsA2 (SEQ ID NO: 18), Beta-Globin Minimal Promoter (pBGmin/minBGlobin/minBGprom) (SEQ ID NO: 19), minCMV Promoter (SEQ ID NO:

20), Mutated minCMV Promoter (Sacl RE site removed) (SEQ ID NO: 21), minRho Promoter (SEQ ID NO: 22), minRho* Promoter (SEQ ID NO: 23), Hsp68 minimal Promoter (proHsp68) (SEQ ID NO: 24), SYFP2 (SEQ ID NO: 25), EGFP (SEQ ID NO: 26), Optimized Flp recombinase (FlpO) (SEQ ID NO: 27), Improved Cre recombinase (iCre) (SEQ ID NO: 28), SP10 insulator (SP10ins) (SEQ ID NO: 29), 3xSP10ins (SEQ ID NO: 30), WPRE3 (SEQ ID NO: 31), BGHpA (SEQ ID NO: 32), P2A (SEQ ID NO: 33), T2A (SEQ ID NO: 34), E2A (SEQ ID NO: 35), F2A (SEQ ID NO: 36), Exemplary Plasmid Backbone 1 – Left ITR (SEQ ID NO: 124), Exemplary Plasmid Backbone 1 – Right ITR (SEQ ID NO: 125), Exemplary Plasmid Backbone 2 – Left ITR (SEQ ID NO: 126), Exemplary Plasmid Backbone 2 – Right ITR (SEQ ID NO: 127), PHP.eB capsid (SEQ ID NO: 37), AAV9 VP1 capsid protein (SEQ ID NO: 38), tTA2 (SEQ ID NO: 39), Plasmid backbone 1 (SEQ ID NO: 40), Plasmid backbone 2 (SEQ ID NO: 41), AiP1146 (T502-047, vAi30.0) (SEQ ID NO: 43), AiP1113 (T502-053, vAi30.1) (SEQ ID NO: 44), AiP1147 (T502-048, vAi31.0) (SEQ ID NO: 45), AiP989 (TG989 vAi11.0,) (SEQ ID NO: 46), AiP1013 (TG1013, vAi12.0) (SEQ ID NO: 47), AiP1012 (TG1012, vAi14.0) (SEQ ID NO: 48), CN1525 (vAi115.0) (SEQ ID NO: 49), CN1528 (vAi116.0) (SEQ ID NO: 50), CN1532 (vAi117.0) (SEQ ID NO: 51), CN1621 (vAi119.0) (SEQ ID NO: 52), CN1633 (vAi120.0) (SEQ ID NO: 53), CN1259 (vAi104.0) (SEQ ID NO: 54), CN2045 (SEQ ID NO: 55), CN1255 (vAi101.0) (SEQ ID NO: 56), CN1408 (vAi110.0) (SEQ ID NO: 57), CN1258 (vAi103.0) (SEQ ID NO: 58), CN1279 (vAi102.0) (SEQ ID NO: 59), CN1253 (vAi100.0) (SEQ ID NO: 60), CN1274 (SEQ ID NO: 61), Lactase (SEQ ID NO: 62), Lipase (SEQ ID NO: 63), Helicase (SEQ ID NO: 64), Amylase (SEQ ID NO: 65), Alpha-glucosidase (SEQ ID NO: 66), Transcription factor SP1 (SEQ ID NO: 67), Transcription factor AP-1 (SEQ ID NO: 68), Heat shock factor protein 1 (SEQ ID NO: 69), CCAAT/enhancer-binding protein (C/EBP) beta isoform a (SEQ ID NO: 70), 44105 (SEQ ID NO: 71), Transforming growth factor receptor beta 1 (SEQ ID NO: 72), Platelet-derived growth factor receptor (SEQ ID NO: 73), Epidermal growth factor receptor (SEQ ID NO: 74), Vascular endothelial growth factor receptor (SEQ ID NO: 75), Interleukin 8 receptor alpha (SEQ ID NO: 76), Caveolin (SEQ ID NO: 77), Dynamin (SEQ ID NO: 78), Clathrin heavy chain 1 isoform 1 (SEQ ID NO: 79), Clathrin heavy chain 2 isoform 1 (SEQ ID NO: 80), Clathrin light chain A isoform a (SEQ ID NO: 81), Clathrin light chain B isoform a (SEQ ID NO: 82), Ras-related protein Rab-4A isoform 1 (SEQ ID NO: 83), Ras-related protein Rab-11A, UniProtKB/Swiss-Prot: P62491.3: (SEQ ID NO: 84), Platelet-derived growth factor (SEQ ID NO: 85), Transforming growth factor-beta3 (SEQ ID NO: 86), Nerve growth factor (SEQ ID NO: 87), Epidermal growth factor (EGF) (SEQ ID NO: 88), GTPase HRas (SEQ ID NO: 89), Cocaine And Amphetamine Regulated Transcript (Chain A) (SEQ ID NO: 90), Protachykinin-1 (SEQ ID NO: 91),

Protachykinin-1 (SEQ ID NO: 92), Oxytocin-neurophysin 1 (SEQ ID NO: 93), Oxytocin is position 20-28 of Oxytocin-neurophysin 1 (SEQ ID NO: 94), Somatostatin (SEQ ID NO: 95), Myosin light chain kinase, Green fluorescent protein, Calmodulin chimera (Chain A) (SEQ ID NO: 96), Genetically-encoded green calcium indicator NTnC (chain A) (SEQ ID NO: 97), Calcium indicator TN-XXL (SEQ ID NO: 98), BRET-based auto-luminescent calcium indicator (SEQ ID NO: 99), Calcium indicator protein OeNL(Ca²⁺)-18u (SEQ ID NO: 100), GCaMP6m (SEQ ID NO: 101), GCaMP6s (SEQ ID NO: 102), GCaMP6f (SEQ ID NO: 103), Channelopsin 1 (SEQ ID NOs: 104 and 105), Channelrhodopsin-2 (SEQ ID NOs: 106 and 107), CRISPR-associated protein (Cas) (SEQ ID NO: 108), Cas9 (SEQ ID NO: 109), CRISPR-associated endonuclease Cpf1 (SEQ ID NO: 110), Ribonuclease 4 (SEQ ID NO: 111), Deoxyribonuclease II beta (SEQ ID NO: 112), Sodium channel protein type 1 subunit alpha (SEQ ID NO: 113), Potassium voltage-gated channel subfamily KQT member 2 (SEQ ID NO: 114), and Voltage-dependent L-type calcium channel subunit alpha-1C (SEQ ID NO: 115).

DETAILED DESCRIPTION

[0031] To fully understand the biology of the brain, different cell types need to be distinguished and defined and, to further study them, artificial expression constructs that can selectively label and perturb them need to be identified. Tasic, *Curr. Opin. Neurobiol.* 50, 242–249 (2018); Zeng & Sanes, *Nat. Rev. Neurosci.* 18, 530–546 (2017). In mouse, recombinase driver lines have been used to great effect to label cell populations that share marker gene expression. Daigle et al., *Cell* 174, 465–480.e22 (2018); Taniguchi, et al., *Neuron* 71, 995–1013 (2011); Gong et al., *J. Neurosci.* 27, 9817–9823 (2007). 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.

[0032] The current disclosure provides artificial expression constructs that selectively drive gene expression in targeted central nervous system cell populations. Targeted central nervous system cell populations include: gamma-aminobutyric acid (GABA)ergic neurons generally; and/or GABAergic neuron cell types such as lysosomal associated membrane protein 5 (Lamp5) neurons, vasoactive intestinal polypeptide-expressing (Vip) neurons, somatostatin (Sst) neurons, and parvalbumin (Pvalb) neuron cell types. Layer 4 (L4) and/or layer 5 (L5) intratelencephalic (IT) neurons, deep cerebellar nuclear neurons, or cerebellar Purkinje cells can also be targeted for selective gene expression.

[0033] Particular embodiments of the artificial expression constructs utilize the following

enhancers to selectively drive gene expression within targeted central nervous system cell populations as follows (enhancer / targeted cell population): Grik1_enhGad2-1 / GABAergic neurons generally; Grik1_enhGad2-2 / GABAergic neurons generally; mscRE5 / GABAergic neurons generally; mscRE8 / GABAergic neurons generally; eHGT_019h / Lamp5 neurons; eHGT_022h / Lamp5 and Vip neurons; eHGT_022m / Lamp5 and Vip neurons; eHGT_017h / Lamp5, Vip, and Sst neurons; eHGT_17m / Lamp5, Vip, and Sst neurons; eHGT_079h / parvalbumin (Pvalb) neuron cell types; eHGT_082h / Pvalb neuron cell types, and deep cerebellar nuclear cells; eHGT_086h / Pvalb neuron cell types; eHGT_128h / Pvalb neuron cell types; eHGT_140h / Pvalb neuron cell types; eHGT_064h / Pvalb and Sst neuron cell types; eHGT_023h / Pvalb cell types, and L4 and L5 IT neurons, and cerebellar Purkinje cells; and eHGT_359 / Pvalb cell types and cerebellar Purkinje cells. In particular embodiments, unless otherwise specified, targeted cell types are neocortical cell types.

[0034] Particular embodiments provide artificial expression constructs including the features of vectors described herein including vectors: AiP1146, AiP1113, AiP1147, AiP1147, AiP1013, AiP1012, CN1525, CN1528, CN1532, CN1621, CN1633, CN1259, CN2045, CN1255, CN1408, CN1258, CN1279, CN1253, and CN1274.

[0035] 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.

[0036] (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 artificial expression construct can also include other regulatory elements if necessary or beneficial.

[0037] 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 Grik1_enhGad2-1, Grik1_enhGad2-2, mscRE5, mscRE8, eHGT_019h, eHGT_022h, eHGT_022m, eHGT_017h, eHGT_17m, eHGT_079h, eHGT_082h, eHGT_086h, eHGT_128h, eHGT_140h, eHGT_064h, eHGT_023h and eHGT_359.

[0038] 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.

[0039] When a coding sequence is selectively expressed in selected cells and is not substantially expressed in other 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 cells. The non-targeted 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 GABAergic cell with a different gene expression profile than the targeted cells.

[0040] 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 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.

[0041] 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., Nature 563, 72-78 (2018) and Hodge et al., Nature 573, 61-68 (2019). For reference, the following description of neural cell types and distinguishing features is also provided:

[0042] Neocortical GABAergic Subclasses:

- All: Express GABA synthesis genes Gad1/GAD1 and Gad2/GAD2.
- Lamp5, Sncg, Serpinf1, and Vip: Developmentally derived from neuronal progenitors from the caudal ganglionic eminence (CGE) or preoptic area (POA).
- Sst and Pvalb: Developmentally derived from neuronal progenitors in the medial ganglionic eminence (MGE).
- Lamp5: Found in many neocortical layers, especially upper (L1-L2/3), and have mainly neurogliaform and single bouquet morphology.
- Sncg: Found in many neocortical layers, and have molecular overlaps with Lamp5 and Vip cells, but inconsistent expression of Lamp5 or Vip, with more consistent expression of Sncg.
- Serpinf1: Found in many neocortical 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 neocortical layers, but especially frequent in upper layers (L1-L4), and highly express the neurotransmitter vasoactive intestinal peptide (Vip).
- Sst: Found in many neocortical 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 Sst Chodl neurons (which also express Nos1 and Tacr1) that are highly distinct from other Sst neurons but express some shared marker genes including Sst. In human, SST gene expression is often detected in layer 1 LAMP5+ cells.
- Pvalb: Found in many neocortical layers, but especially frequent in lower layers (L5-L6). They highly express the calcium-binding protein parvalbumin (Pvalb), express neuropeptide Tac1, and frequently dampen the output of postsynaptic neurons. Most fast-spiking GABAergic cells express Pvalb strongly. 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, only neocortical GABAergic type that expresses Meis2 gene, and does not express some other genes that are expressed by all other neocortical GABAergic types (for example, Thy1 and Scn2b). This type is found in L6b and subcortical white matter.

[0043] Neocortical Glutamatergic Subclasses:

- All: Express glutamate transmitters Slc17a6 and/or Slc17a7. They all express Snap25 and lack expression of Gad1/Gad2 and lack expression of Slc1A3.

- L2/3 IT: Primarily reside in Layer 2/3 and have mainly intratelencephalic (cortico-cortical) projections.
- L4 IT: Primarily reside in Layer 4 and mainly have either local or 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 (corticofugal) or L5 ET (extratelencephalic). This subclass includes cells that are located in the primary motor cortex and neighboring areas and are corticospinal projection neurons, which are associated with motor neuron/movement disorders, such as ALS. This subclass includes thick-tufted pyramidal neurons, including distinctive cell types 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 neocortical subplate (L6b), with local (near the cell body) projections and some 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.

[0044] Cerebellar Purkinje cells: large GABAergic neurons that are the only projection neurons and the sole output from the cerebellum. Their cell bodies form a single layer, so called 'Purkinje cell layer', and they express parvalbumin.

[0045] Deep cerebellar nuclear neurons: neurons located in the deep cerebellar nuclear structure. These include excitatory and GABAergic cells that express the gene Pvalb.

[0046] Non-neuronal Subclasses:

- Astrocytes: Neuroectoderm-derived glial cells which express the marker Aqp4 and often GFAP, but do not express neuronal marker SNAP25. They can have a distinct star-shaped morphology and are involved in metabolic support of other cells in the brain. Multiple astrocyte morphologies are observed in mouse and human
- 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 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: Specialized smooth-muscle cells which are blood vessel-associated cells that express the marker gene Acta2. SMCs cover arterioles in the brain and are involved in blood-brain barrier permeability.
- Endothelial cells: Cells that line blood vessels of the brain. Endothelial cells express the markers Tek and PDGF-B.
- Microglia: hematopoietic-derived immune cells, 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. Microglia are known to express Cx3cr1, Tmem119, and PTPRC (CD45).

[0047] 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.

[0048] 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, ZsYellowl); and tandem conjugates.

[0049] 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.

[0050] 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.

[0051] Exemplary functional molecules include functioning ion transporters, cellular trafficking proteins, enzymes, transcription factors, neurotransmitters, calcium reporters, channelrhodopsins, guide RNA, nucleases, or designer receptors exclusively activated by designer drugs (DREADDs).

[0052] 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)).

[0053] 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.

[0054] 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).

[0055] 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).

[0056] In particular embodiments, functional molecules include modulators of neuronal activity

like channelrhodopsins (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. 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).

[0057] In particular embodiments, functional molecules include DNA and RNA editing tools such as 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.

[0058] 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.

[0059] 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.

[0060] 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 artificial 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.

[0061] Coding sequences encoding molecules (e.g., RNA, proteins) described herein can be 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.

[0062] The term “gene” may include not only coding sequences but also regulatory regions such as promoters, enhancers, insulators, and/or post-regulatory elements, such as 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.

[0063] 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*, (minCMV* is minCMV with a SacI restriction site removed), minRho, minRho* (minRho* is minRho with a SacI restriction site removed), 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.

[0064] 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.

[0065] Viral vector is widely used to refer to a nucleic acid molecule that includes virus-derived components 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.

[0066] 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.

[0067] 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.

[0068] 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.

[0069] 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.

[0070] 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.

[0071] 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.

[0072] 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.

[0073] 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.

[0074] 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.

[0075] 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), Rous Sarcoma Virus (RSV), and lentivirus.

[0076] "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.

[0077] 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.

[0078] 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.

[0079] 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.

[0080] 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.

[0081] 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 signal 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.

[0082] 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).

[0083] 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.

[0084] Particular embodiments of vectors disclosed herein include:

Vector Name	Vector Features
AiP1146	rAAV-Grik1_enhGad2-1-Hsp68-EGFP-WPRE3-BGHpA
AiP1113	rAAV-Grik1_enhGad2-1-pBGmin-EGFP-WPRE3-BGHpA
AiP1147	rAAV-Grik1_enhGad2-2-Hsp68-EGFP-WPRE3-BGHpA
AiP1013	rAAV-mscRE5-pBGmin-EGFP-WPRE3-BGHpA
AiP1013	rAAV-mscRE5-pBGmin-FlpO-WPRE3-BGHpA
AiP1012	rAAV-mscRE8-pBGmin-FlpO-WPRE3-BGHpA
CN1525	rAAV-hsA2-eHGT_079h-minRho-SYFP2-WPRE3-BGHpA

CN1528	rAAV-hsA2-eHGT_082h-minRho-SYFP2-WPRE3-BGHpA
CN1532	rAAV-hsA2-eHGT_086h-minRho-SYFP2-WPRE3-BGHpA
CN1621	rAAV-hsA2-eHGT_128h-minRho-SYFP2-WPRE3-BGHpA
CN1633	rAAV-hsA2-eHGT_140h-minRho-SYFP2-WPRE3-BGHpA
CN1259	scAAV-eHGT_023h-minBGlobin-SYFP2-WPRE3-BGHpA
CN2045	rAAV-3xSP10ins-eHGT_359h-minRho*-SYFP2-WPRE3-BGHpA
CN1255	scAAV-eHGT_019h-minBGlobin-SYFP2-WPRE3-BGHpA
CN1408	rAAV-eHGT_064h-minBGlobin-SYFP2-WPRE3-BGHpA
CN1258	scAAV-eHGT_022h-minBGlobin-SYFP2-WPRE3-BGHpA
CN1279	scAAV-eHGT_022m-minBGlobin-SYFP2-WPRE3-BGHpA
CN1253	scAAV-eHGT_017h-minBGlobin-SYFP2-WPRE3-BGHpA
CN1274	scAAV-eHGT_017m-minBGlobin-SYFP2-WPRE3-BGHpA

[0085] Subcomponent sequences within the larger vector sequences can be readily identified by one of ordinary skill in the art and based on the contents of the current disclosure (see FIG. 22). Nucleotides between identifiable and enumerated subcomponents reflect restriction enzyme recognition sites used in assembly (cloning) of the constructs, and in some cases, additional nucleotides do not convey any identifiable function. These segments of complete vector sequences can be adjusted based on use of different cloning strategies and/or vectors. In general, short 6-nucleotide palindromic sequences reflect vector construction artifacts that are not important to vector function.

[0086] 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), AAV-PPS (Chen et al., Nat Med. 2009; 15: 1215), and PHP.eB. In particular embodiments, 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: 116) are changed to S-DGTLAVPFK-A (SEQ ID NO: 117). In particular embodiments, PHP.eb refers to SEQ ID NO: 37.

[0087] 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 spinal muscular atrophy (SMA) syndrome by AveXis (AVXS-101, NCT03505099)

and the treatment of CLN3 gene-Related Neuronal Ceroid-Lipofuscinosis (NCT03770572).

[0088] 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.

[0089] 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.

[0090] 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.

[0091] AAV-BR1 is an AAV2 variant displaying the NRGTEWD (SEQ ID NO: 118) 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).

[0092] 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: 119), transduces neurons in the enteric nervous system, and strongly transduces peripheral sensory afferents entering the spinal cord and brain stem.

[0093] 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: 120). It transfers genes throughout the CNS with higher efficiency than AAV9 and transduces the majority of astrocytes and neurons across multiple CNS regions.

[0094] AAV-PPS, an AAV2 variant created by insertion of the DSPAHPS (SEQ ID NO: 121) epitope into the capsid of AAV2, shows a dramatically improved brain tropism relative to AAV2.

[0095] For additional information regarding capsids that cross the blood brain barrier, see Chan et al., Nat. Neurosci. 2017 Aug; 20(8): 1172-1179.

[0096] (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.

[0097] 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.

[0098] 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.

[0099] 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).

[0100] In particular embodiments, compositions can be formulated for intravenous, intraparenchymal, intraocular, intravitreal, parenteral, subcutaneous, intracerebro-ventricular, intramuscular, intrathecal, intraspinal, intraperitoneal, oral or nasal inhalation, or by direct injection in or application to one or more cells, tissues, or organs.

[0101] Compositions may include liposomes, lipids, lipid complexes, microspheres, microparticles, nanospheres, and/or nanoparticles.

[0102] 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).

[0103] 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 polyalkyl-cyanoacrylate 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.

[0104] 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.

[0105] 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.

[0106] 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.

[0107] 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.

[0108] 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.

[0109] 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).

[0110] Supplementary active ingredients can also be incorporated into the compositions.

[0111] 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.

[0112] 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.

[0113] (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.

[0114] A variety of host cell lines can be used, but in particular embodiments, the cell is a mammalian neural cell. In particular embodiments, the artificial express construct includes an enhancer and/or a vector sequence of Grik1_enhGad2-1, Grik1_enhGad2-2, mscRE5, mscRE8, eHGT_019h, eHGT_022h, eHGT_022m, eHGT_017h, eHGT_17m, eHGT_079h, eHGT_082h, eHGT_086h, eHGT_128h, eHGT_140h, eHGT_064h, eHGT_023h, and eHGT_359 and/or a vector sequence of AiP1146, AiP1113, AiP1147, AiP1147, AiP1013, AiP1012, CN1525, CN1528, CN1532, CN1621, CN1633, CN1259, CN2045, CN1255, CN1408, CN1258, CN1279, CN1253, and CN1274, 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.

[0115] 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.

[0116] 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.

[0117] 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.

[0118] 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.)

[0119] 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

[0120] 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. 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).

[0121] 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 Grik1_enhGad2-1, Grik1_enhGad2-2, mscRE5, mscRE8, eHGT_019h, eHGT_022h, eHGT_022m, eHGT_017h, eHGT_17m, eHGT_079h, eHGT_082h, eHGT_086h, eHGT_128h, eHGT_140h, eHGT_064h, eHGT_023h, or eHGT_359.

[0122] 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 artificial expression construct already integrated into the genome. (for an example see MacKenzie & Quinn, Proc Natl Acad Sci USA 96: 15251-15255, 1999).

[0123] (iv) Transgenic Animals. Another aspect of the disclosure includes transgenic animals, the genome of which contains an artificial expression construct including Grik1_enhGad2-1; Grik1_enhGad2-1, Grik1_enhGad2-2, mscRE5, mscRE8, eHGT_019h, eHGT_022h, eHGT_022m, eHGT_017h, eHGT_17m, eHGT_079h, eHGT_082h, eHGT_086h, eHGT_128h, eHGT_140h, eHGT_064h, eHGT_023h, and/or eHGT_359 operatively linked to a heterologous coding sequence. In particular embodiments, the genome of a transgenic animal includes AiP1146, AiP1113, AiP1147, AiP1147, AiP1013, AiP1012, CN1525, CN1528, CN1532, CN1621, CN1633, CN1259, CN2045, CN1255, CN1408, CN1258, CN1279, CN1253, and/or CN1274. In particular embodiments, when a non-integrating vector is utilized, a transgenic animal includes an artificial expression construct including Grik1_enhGad2-1, Grik1_enhGad2-2, mscRE5, mscRE8, eHGT_019h, eHGT_022h, eHGT_022m, eHGT_017h, eHGT_17m, eHGT_079h, eHGT_082h, eHGT_086h, eHGT_128h, eHGT_140h, eHGT_064h, eHGT_023h, eHGT_359, AiP1146, AiP1113, AiP1147, AiP1147, AiP1013, AiP1012, CN1525, CN1528, CN1532, CN1621, CN1633, CN1259, CN2045, CN1255, CN1408, CN1258, CN1279, CN1253, and/or CN1274 within one or more of its cells.

[0124] 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.

[0125] 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.

[0126] 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.

[0127] 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); feedback-controlled delivery (U.S. Pat. No. 5,697,899), and any other delivery method available and/or described elsewhere in the disclosure.

[0128] (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.

[0129] 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.

[0130] Particular embodiments include methods of administering to a subject an artificial expression construct that includes Grik1_enhGad2-1, Grik1_enhGad2-2, mscRE5, mscRE8, eHGT_019h, eHGT_022h, eHGT_022m, eHGT_017h, eHGT_17m, eHGT_079h, eHGT_082h, eHGT_086h, eHGT_128h, eHGT_140h, eHGT_064h, eHGT_023h, eHGT_359, AiP1146, AiP1113, AiP1147, AiP1147, AiP1013, AiP1012, CN1525, CN1528, CN1532, CN1621, CN1633, CN1259, CN2045, CN1255, CN1408, CN1258, CN1279, CN1253, and/or CN1274 as described herein to drive selective expression of a gene in a selected cell type. The subject can be an isolated cell, a network of cells, a tissue slice, an experimental animal, a veterinary animal, or a human.

[0131] 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, intraparenchymal, intraspinal, retro-orbital, or intrathecal administration can be infused with from 10^6 to 10^{22} copies of the artificial expression construct.

[0132] 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.

[0133] 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.

[0134] 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, intraparenchymally, intracerebro-ventricularly, intramuscularly, intrathecally, intraspinally, 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.

[0135] (vi) Kits and Commercial Packages. Kits and commercial packages contain an artificial expression construct described herein. The artificial 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.

[0136] 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.

[0137] 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.

[0138] (vii) Exemplary Embodiments.

1. An artificial expression construct including (i) an enhancer selected from Grik1_enhGad2-1, Grik1_enhGad2-1, Grik1_enhGad2-2, mscRE5, mscRE8, eHGT_019h, eHGT_022h, eHGT_022m, eHGT_017h, eHGT_17m, eHGT_079h, eHGT_082h, eHGT_086h, eHGT_128h, eHGT_140h, eHGT_064h, eHGT_023h, and eHGT_359; (ii) a promoter; and (iii) a heterologous encoding sequence.
2. The artificial expression construct of embodiment 1, wherein the heterologous encoding sequence encodes an effector element or an expressible element.

3. The artificial expression construct of embodiment 1 or 2, wherein the effector element includes a reporter protein or a functional molecule.
4. The artificial expression construct of embodiment 3, wherein the reporter protein includes a fluorescent protein.
5. The artificial expression construct of embodiment 1 or 3, wherein the functional molecule includes a functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or a designer receptor exclusively activated by designer drug (DREADD).
6. The artificial expression construct of any of embodiments 1-5, wherein the expressible element includes a non-functional molecule.
7. The artificial expression construct of embodiment 6, wherein the non-functional molecule includes a non-functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or a DREADD.
8. The artificial expression construct of any of embodiments 1-7, wherein the artificial expression construct is associated with a capsid that crosses the blood brain barrier.
9. The artificial expression construct of embodiment 8, wherein the capsid includes PHP.eB, AAV-BR1, AAV-PHP.S, AAV-PHP.B, or AAV-PPS.
10. The artificial expression construct of any of embodiments 1-9, wherein the artificial expression construct includes or encodes a skipping element.
11. The artificial expression construct of embodiment 10, wherein the skipping element includes a 2A peptide and/or an internal ribosome entry site (IRES).
12. The artificial expression construct of embodiment 11, wherein the 2A peptide includes T2A, P2A, E2A, or F2A.
13. The artificial expression construct of any of embodiments 1-12, wherein the artificial expression construct includes or encodes a set of features selected from Grik1_enhGad2-1, Grik1_enhGad2-2, mscRE5, mscRE8, eHGT_019h, eHGT_022h, eHGT_022m, eHGT_017h, eHGT_17m, eHGT_079h, eHGT_082h, eHGT_086h, eHGT_128h, eHGT_140h, eHGT_064h, eHGT_023h, and eHGT_359, AAV, scAAV, rAAV, minBglobin, CMV, minCMV, minRho, minRho*, fluorescent protein (e.g., EGFP, SYFP, GFP), Cre, iCre, dgCre, FlpO, tTA2, SP10, WPRE, and/or BGHpA

14. The artificial expression construct of any of embodiments 1-13, wherein the artificial expression construct includes or encodes a set of features selected from:

Grik1_enhGad2-1-Hsp68-EGFP-WPRE3-BGHpA;

Grik1_enhGad2-1-pBGmin-EGFP-WPRE3-BGHpA;

Grik1_enhGad2-2-Hsp68-EGFP-WPRE3-BGHpA;

mscRE5-pBGmin-EGFP-WPRE3-BGHpA;

mscRE5-pBGmin-FlpO-WPRE3-BGHpA;

mscRE8-pBGmin-EGFP-WPRE3-BGHpA;

mscRE8-pBGmin-FlpO-WPRE3-BGHpA;

scAAV-eHGT_019h-minBGlobin-SYFP2-WPRE3-BGHpA;

scAAV-eHGT_022h-minBGlobin-SYFP2-WPRE3-BGHpA;

scAAV-eHGT_022m-minBGlobin-SYFP2-WPRE3-BGHpA;

scAAV-eHGT_017h-minBGlobin-SYFP2-WPRE3-BGHpA;

hsA2-eHGT_079h-minRho-SYFP2-WPRE3-BGHpA;

hsA2-eHGT_082h-minRho-SYFP2-WPRE3-BGHpA;

hsA2-eHGT_086h-minRho-SYFP2-WPRE3-BGHpA;

hsA2-eHGT_128h-minRho-SYFP2-WPRE3-BGHpA;

hsA2-eHGT_140h-minRho-SYFP2-WPRE3-BGHpA;

eHGT_064h-minBglobin-SYFP2-WPRE3-BGHpA; or

scAAV-eHGT_023h-minBGlobin-SYFP2-WPRE3-BGHpA.

15. A vector including an artificial expression construct of any of embodiments 1-14

16. The vector of embodiment 15, wherein the vector includes a viral vector.

17. The vector of embodiment 15 or 16, wherein the viral vector includes a recombinant adeno-associated viral (AAV) vector.

18. 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 Grik1_enhGad2-1, Grik1_enhGad2-2, mscRE5, mscRE8, eHGT_019h, eHGT_022h, eHGT_022m, eHGT_017h, eHGT_17m, eHGT_079h, eHGT_082h, eHGT_086h, eHGT_128h, eHGT_140h, eHGT_064h, eHGT_023h, and eHGT_359.

19. The AAV vector of embodiment 18, wherein the AAV vector is replication-competent.

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

21. The transgenic cell of embodiment 20, wherein the transgenic cell is a GABAergic neuron.
22. The transgenic cell of embodiment 20, wherein the transgenic cell is a lysosomal associated membrane protein 5 (Lamp5) neuron (e.g., neocortical); a vasoactive intestinal polypeptide-expressing (Vip) neuron (e.g., neocortical); a somatostatin (Sst) neuron (e.g., neocortical); a parvalbumin (Pvalb) neuron (e.g., neocortical); a layer 4 (L4) intratelencephalic (IT) neuron, a layer 5 (L5) IT neuron, a deep cerebellar nuclear neuron or a cerebellar Purkinje cell.
23. A non-human transgenic animal including an expression construct, vector, or transgenic cell of any of the preceding embodiments.
24. The non-human transgenic animal of embodiment 24, wherein the non-human transgenic animal is a mouse or a non-human primate.
25. An administrable composition including an expression construct, vector, or transgenic cell of any of the preceding embodiments.
26. A kit including an expression construct, vector, transgenic cell, transgenic animal, and/or administrable compositions of any of the preceding embodiments.
27. 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 25 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.
28. The method of embodiment 27, wherein the heterologous gene encodes an effector element or an expressible element.
29. The method of embodiment 28, wherein the effector element includes a reporter protein or a functional molecule.
30. The method of embodiment 29, wherein the reporter protein includes a fluorescent protein.
31. The method of embodiment 29 or 30, wherein the functional molecule includes a functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or a DREADD.
32. The method of embodiment 28, wherein the expressible element includes a non-functional molecule.
33. The method of embodiment 32, wherein the non-functional molecule includes a non-functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor

cassette, or DREADD.

34. The method of any of embodiments 27 - 33, wherein the providing includes pipetting.

35. The method of embodiment 34, wherein the pipetting is to a brain slice.

36. The method of embodiment 35, wherein the brain slice includes a GABAergic neuron.

37. The method of embodiment 35, wherein the brain slice includes a Lamp5 neuron; a Vip neuron; an Sst neuron; a Pvalb neuron; an L4 IT neuron, an L5 IT neuron, a deep cerebellar nuclear cell and/or a cerebellar Purkinje cell.

38. The method of any of embodiments 35 - 37, wherein the brain slice is murine, human, or non-human primate.

39. The method of any of embodiments 27 - 33, wherein the providing includes administering to a living subject.

40. The method of embodiment 39, wherein the living subject is a human, non-human primate, or a mouse.

41. The method of embodiment 39 or 40, wherein the administering to a living subject is through injection.

42. The method of embodiment 41, wherein the injection includes intravenous injection, intraparenchymal injection into brain tissue, intracerebroventricular (ICV) injection, intra-cisterna magna (ICM) injection, or intrathecal injection.

43. An artificial expression construct including AiP1146, AiP1113, AiP1147, AiP1147, AiP1013, AiP1012, CN1525, CN1528, CN1532, CN1621, CN1633, CN1259, CN2045, CN1255, CN1408, CN1258, CN1279, CN1253, or CN1274.

[0139] (viii) Experimental Examples. Experimental Methods for Enhancers MGT_E31 (eAi12.0), MGT_E65 (eAi13.0), MGT_E5 (eAi4.0) and MGT_E8 (eAi5.0). Viral genome cloning. Enhancers were cloned from C57Bl/6J genomic DNA using enhancer-specific primers and Phusion high-fidelity polymerase (M0530S; NEB). Individual enhancers were then inserted into an rAAV backbone that contained a minimal beta-globin promoter or the minimal Hsp68 promoter, gene, a woodchuck post-transcriptional regulatory element (WPRE) and a bovine growth hormone polyA using standard molecular cloning approaches. Plasmid integrity was verified via Sanger sequencing and restriction digests were performed to confirm intact inverted terminal repeat (ITR) sites.

[0140] Viral packaging and titering. Before transfection, 10⁵ µg of AAV viral genome plasmid, 190 µg pHelper, and 105 µg AAV-PHP.eB were mixed with 5 mL of Opti-MEM I media (Reduced Serum, GlutaMAX; ThermoFisher Scientific) and 1.1 mL of a solution of 1 mg/mL 25 kDa linear Polyethylenimine (Polysciences) in PBS at pH 4-5. This co-transfection mixture was

incubated at room temperature for 10 minutes. Recombinant AAV of the PHP.eB serotype was generated by adding 0.61 mL of this co-transfection mixture to each of ten 15-cm dishes of HEK293T cells (ATCC) at 70-80% confluence. 24 hours post-transfection, cell medium was replaced with DMEM (with high glucose, L-glutamine and sodium pyruvate; ThermoFisher Scientific) with 4% FBS (Hyclone) and 1% Antibiotic-Antimycotic solution. Cells were collected 72 hours post transfection by scraping in 5mL of medium and were pelleted at 1500 rpm at 4°C for 15 minutes. Pellets were suspended in a buffer containing 150 mM NaCl, 10 mM Tris, and 10 mM MgCl₂, pH 7.6, and were frozen in dry ice. Cell pellets were thawed quickly in a 37°C water bath, then the cell-containing medium was passed through a syringe with a 21-23G needle 5 times, followed by 3 more rounds of freeze/thaw, and a 30-minute incubation with 50 U/ml Benzonase (Sigma-Aldrich) at 37°C. The suspension was then centrifuged at 3,000 × g to pellet the cellular debris and the supernatant was further purified using a layered iodixanol step gradient (15%, 25%, 40%, and 60%) by centrifugation at 58,000 rpm in a Beckman 70Ti rotor for 90 minutes at 18°C. The virus containing fraction was purified by extraction of the full volume below the 40-60% gradient layer interface. Viruses were concentrated using Amicon Ultra-15 centrifugal filter unit by centrifugation at 3,000 rpm at 4°C, and reconstituted in PBS with 5% glycerol and 35 mM NaCl before storage at -80°C.

[0141] Virus titers were measured using quantitative PCR (qPCR) with a primer pair that recognizes a region of 117bp in the AAV2 ITRs (Forward: GGAACCCCTAGTGATGGAGTT (SEQ ID NO: 122); Reverse: CGGCCTCAGTGAGCGA (SEQ ID NO: 123). qPCR reactions were performed using QuantiTect SYBR Green PCR Master Mix (Qiagen) and 500 nM primers. To determine virus titers, a positive control AAV with known titer and newly produced viruses with unknown titers were treated with DNase I. Serial dilutions (1/10, 1/100, 1/500, 1/2500, 1/12500, and 1/62500) of both positive control and newly generated viruses were loaded on the same qPCR plate. A standard curve of virus particle concentrations vs C_q values was generated based on the positive control virus, and the titers of the new viruses were calculated based on the standard curve.

[0142] Retro-orbital injections. To introduce AAV viruses into the brain, 21 day old or older C57Bl/6J, Ai14, or Ai65F mice were briefly anesthetized by isoflurane and 1×10¹⁰ -1×10¹¹ viral genome copies (gc) were delivered into the retro-orbital sinus in a maximum volume of 50 µL or less. Madisen et al., Neuron 85, 942–958 (2015). This approach has been utilized previously to deliver AAV viruses across the blood brain barrier and into the murine brain with high efficiency. Chan et al., Nat. Neurosci. 20 1172-1179 (2017). doi:10.1038/nn.4593. For delivery of multiple AAVs, the viruses were mixed beforehand and then delivered simultaneously into the retro-

orbital sinus. Animals were allowed to recover and then sacrificed 1-3 weeks post-infection in order to analyze virally-introduced transgenes within the brain.

[0143] Stereotaxic injections. Viral DNA was packaged in a PHP.eB serotype to produce recombinant adeno-associated virus (rAAV) as described above. Each purified virus with a titer of 1.0×10^{13} gc/ml was delivered bilaterally at 250 and 50 nL or 50 and 25 nL into the primary visual cortex (VISp; coordinates: A/P: -3.8, ML: -2.5, DV: 0.6) of C57BL/6J mice or heterozygous Ai65F or Gad2-IRES-Cre;Ai14 mice heterozygous at both alleles, using a pressure injection system (Nanoject II, Drummond Scientific Company, Catalog# 3-000-204). The expression for all viruses was analyzed at 14 days post-injection. For tissue processing, mice were transcardially perfused with 4% paraformaldehyde (PFA) and post-fixed in 30% sucrose for 1-2 days. 50 μ m sections were prepared using a freezing microtome and fluorescent images of the injections were captured from mounted sections using a Nikon Eclipse TI epi-fluorescent microscope or FV3000 confocal microscope.

[0144] Experimental Methods for Enhancers eHGT_019h, eHGT_017h, eHGT_017m, eHGT_022h, eHGT_022m, eHGT_023h, eHGT_64h, eHGT_079h, eHGT_082h, eHGT_086h, eHGT_128h, eHGT_140h, eHGT_359h. Cloning enhancers. Enhancers were cloned into AAV expression vectors that are derivatives of either pscAAV-MCS (Cell Biolabs catalog # VPK-430) or pAAV-hSyn1-GCaMP6s-P2A-nls-dTomato (Addgene plasmid #51084; <https://www.addgene.org/51084/>) as the source of vector backbones including AAV ITRs,. Enhancers were amplified from male human genomic DNA, or mouse C57BL/6J genomic DNA using Pfu polymerase and inserted by standard Gibson assembly approaches, upstream of a minimal beta-globin promoter and SYFP2, a brighter EGFP alternative that is well tolerated in neurons (Kremers, et al., Biochemistry. 45, 6570-6580, 2006). NEB Stable cells (New England Biolabs # C3040I) were used for transformations. scAAV plasmids were monitored by restriction analysis and sanger sequencing for occasional (10%) recombination of the left ITR.

[0145] Virus production. Enhancer AAV plasmids were maxiprepmed and transfected with polyethyimine max into 1 plate of AAV-293 cells (Cell Biolabs catalog # AAV-100), along with helper plasmid and PHP.eB rep/cap packaging vector. The next day medium was changed to 1% FBS, and then after 5 days cells and supernatant were harvested and AAV particles released by three freeze-thaw cycles. Lysate was treated with benzonase after freeze thaw to degrade free DNA (2 μ L benzonase, 30 min at 37 degrees, MilliporeSigma catalog # E8263-25KU), and then cell debris was precleared with low-speed spin (1500 g 10 min), and finally the crude virus was concentrated over a 100 kDa molecular weight cutoff Centricon column (MilliporeSigma catalog # Z648043) to a final volume of 150 μ L. This crude virus prep was

useful in both mouse and human virus testing.

[0146] Mouse virus testing. Mice were retro-orbitally injected at P42-P49 with 10 μ L (1E11 genome copies) of crude virus prep diluted with 100 μ L PBS, then sacrificed at 18-28 days post infection. For live epifluorescence, mice were perfused with ACSF.7 and live 350 μ m physiology sections were cut with a compresstome from one hemisphere to analyze reporter expression. For antibody staining the other hemisphere was drop-fixed in 4% PFA in PBS for 4-6 hours at 4 degrees, then cryoprotected in 30% sucrose in PBS 48-72 hours, then embedded in OCT for 3 hours at room temperature, then frozen on dry ice and sectioned at 10 μ m thickness, prior to antibody stain using standard practice. Single-cell RNA-seq was accomplished as in Tasic et al., Nat Neurosci. 19, 335–346, 2016 and Tasic et al., Nature. 563, 72, 2018.

[0147] Testing cell type specificity with Hybridization Chain Reaction (HCR)-based multiplexed fluorescence in situ hybridization (mFISH). This technique was performed on mouse brain hemispheres fixed by immersion in 4% PFA in 1xPBS for 4-6 hrs at 0-4 degrees. After fixation, hemispheres were rinsed with PBS and stored them in 1xPBS at 4 degrees for 1-28 days. For sectioning, hemispheres were embedded in 1% low-melt agarose in 1xPBS and cut 50-100 μ m sagittal sections on a Leica VT1000S vibratome in cold 1xPBS buffer. After sectioning, the sagittal sections were post-fixed in 4% PFA in 1xPBS for 2 hours and rinsed in 1xPBS at room temperature. Prior to staining, the sections were dehydrated with 70% ethanol in water at 4 degrees for 1-28 days. On the day of staining, the sections were cleared with 8% SDS in 1xPBS for 2 hours at room temperature then washed three times in 2xSSC for 1 hour each. Afterwards the sections were moved to different wells containing *Hybridization Buffer (*denotes product from Molecular Instruments) before replacing with Hybridization Buffer containing *HCR Probes and hybridized overnight at 37 degrees. The next day, the hybridization mix was removed and washed with *30% Probe Wash Buffer for 1 hour at 37 degrees, then rinsed with 2xSSC. During the probe wash, fluorescently labeled *HCR Hairpins were denatured at 95 degrees for 90 seconds and then snap-cooled in a room temperature aluminum block tube holder for 30 minutes. Then the denatured hairpins were added to *Amplification Buffer before adding to tissue sections for 2 hours at room temperature in the dark. After washing out the amplification mix with PBS/0.1% Triton X-100 for 15 minutes, sections were pre-blocked with 5% normal goat serum in 1xPBS/0.1% Triton X-100 for 1 hour. Then, sections were stained with 1:1000 rabbit anti-GFP antibody (Abcam # ab290) overnight, and washed twice with 1xPBS/0.1% Triton X-100, and detected with 1:500 488-Goat anti rabbit IgG (Thermo Fisher Scientific # A11034) for 2 hours, then washed 2 times with 1xPBS/0.1%

Triton X-100, and then stained with 10 µg/mL DAPI/2xSSC for 1 hour at RT. All antibody staining and washing steps were performed at room temperature with gentle rocking agitation. Sections were mounted on Superfrost Plus slides with Prolong Glass Mounting medium (Thermo Fisher Scientific # P36980), and HCR/antibody stains were imaged with an Olympus FV3000 confocal microscope using manufacturer's software. Molecular Instruments designed probes with the following accession numbers provided to them: Rorb NM_001043354.2; Lamp5 NM_029530.2; Vip NM_011702.3; Pvalb NM_001330686.1; Sst NM_009215.1; Slc17a7 NM_182993.2; Gad1 NM_008077.5.

[0148] Human virus testing. Temporal cortex neurosurgical samples were bubbled in cold ACSF.7 and kept sterile throughout processing. Blocks of tissue were sliced at 350 µm thickness and then white matter and pial membranes were dissected away. Typically, all layers are represented in a cortical slice. Slices then underwent warm recovery (bubbled ACSF.7 at 30 degrees for 15 minutes) followed by reintroduction of sodium (bubbled ACSF.8 at room temperature for 30 minutes, recipe in Table 2; Ting et al., Scientific Reports. 8, 8407, 2018). Slices were then plated at the gas interface on Millicell PTFE cell culture inserts (MilliporeSigma # PICM03050) in a 6-well dish on 1 mL of Slice Culture Medium (recipe in Table 2). After 30 minutes, slices were infected by direct application of high-titer AAV2/PHP.eB viral prep to the surface of the slice, 1 µL per slice. Slice Culture Medium was replenished every 2 days and reporter expression was monitored.

[0149] Table 2. Buffer Recipes

Proteinase K Cleanup Buffer	EDTA	50 mM
	Sodium chloride	5 mM
	Sodium dodecyl sulfate	1.25% (w/v)
	Proteinase K (Qiagen # 19131)	5 mg/mL
Nuclei Isolation Medium	Sucrose	250 mM
	Potassium chloride	25 mM
	Magnesium chloride	5 mM
	Tris-HCl	10 mM
	pH to 8.0 and sterile filter. Store refrigerated.	
Homogenization Buffer	10 mL Nuclei Isolation Medium	
	0.1% (w/v) Triton X-100	
	One pellet Roche Mini cOmplete™ EDTA-free (Sigma catalog # 4693159001)	
	Prepare fresh on day of experiment.	
Blocking Buffer	PBS	
	BSA (catalog # A2058 from Millipore Sigma)	0.5% (w/v)
	Triton X-100	0.1% (w/v)
ACSF.7	HEPES	20 mM

	Sodium Pyruvate 3 mM Taurine 10 μ M Thiourea 2 mM D-(+)-glucose 25 mM Myo-inositol 3 mM Sodium bicarbonate 30 mM Calcium chloride dihydrate 0.5 mM Magnesium sulfate 10 mM Potassium chloride 2.5 mM Monosodium Phosphate 1.25 mM HCl 92 mM N-methyl-D-(+)-glucamine 92 mM L-ascorbic acid 5.0 mM N-acetyl-L-cysteine 12 mM Adjust pH to 7.3-7.4 with HCl, then adjust osmolarity to 295-305. Sterile filter, and then make 100 mL aliquots and freeze them. The thawed aliquot keeps 2-3 months at 4 degrees, until it turns yellow. Bubble with carbogen at least 10-15 minutes before use, and continuously while in use.
ACSF.8	HEPES 20 mM Taurine 10 μ M Thiourea 2 mM D-(+)-glucose 25 mM Myo-inositol 3 mM Sodium bicarbonate 30 mM Calcium chloride dihydrate 2.0 mM Magnesium sulfate 2.0 mM Potassium chloride 2.5 mM Monosodium Phosphate 1.25 mM Sodium chloride 92 mM L-ascorbic acid 5.0 mM N-acetyl-L-cysteine 12 mM Adjust pH to 7.3-7.4 with HCl, then adjust osmolarity to 295-305. Sterile filter, and then make 100mL aliquots and freeze them. The thawed aliquot keeps 2-3 months at 4 degrees, until it turns yellow. Bubble with carbogen at least 10-15 minutes before use, and continuously while in use.
Slice Culture Medium	MEM Eagle medium powder 1680 mg (MilliporeSigma catalog # M4642) L-ascorbic acid powder 36 mg CaCl ₂ , 2.0 M 100 μ L MgSO ₄ , 2.0 M 200 μ L HEPES, 1.0 M 6.0 mL Sodium bicarbonate, 893 mM 3.36 mL

	D-(+)-glucose, 1.11 M Pen/Strep 100x (5k U/mL) (Thermo catalog # 15070063) Tris base, 1.0 M GlutaMAX 200 mM (Thermo catalog # 35050061) Bovine Pancreas Insulin, 10 mg/mL (MilliporeSigma catalog # I0516) Heat-inactivated horse serum (Thermo catalog # 26050088) Deionized water to 250 mL Adjust pH to 7.3-7.4 with HCl, then adjust osmolarity to 300-305. Sterile filter and store refrigerated for up to 1-2 months.	2.25 mL 1.0 mL 260 µL 0.5 mL 20 µL 40 mL
ACSF.1/trehalose	HEPES Sodium Pyruvate Taurine Thiourea D-(+)-glucose Myo-inositol Sodium bicarbonate Calcium chloride dihydrate Magnesium sulfate Potassium chloride Monosodium Phosphate Trehalose dihydrate N-methyl-D-(+)-glucamine L-ascorbic acid N-acetyl-L-cysteine 1 Adjust pH to 7.3-7.4 with HCl and adjust osmolarity to 295-305. Sterile filter, and then make 100 mL aliquots and freeze them. The thawed aliquot keeps 2-3 months at 4 degrees, until it turns yellow.	20 mM 3 mM 10 µM 2 mM 25 mM 3 mM 25 mM 0.5 mM 10 mM 2.5 mM 1.25 mM 132 mM 30 mM 5.0 mM 2 mM
ACSF.1/trehalose + blockers	ACSF.1/trehalose 100 µM TTX (final 0.1 µM) 25 mM DL-AP5 (final 50 µM) 60 mM DNQX (final 20 µM) 100 mM (+)-MK801 (final 10 µM)	50 mL 50 µL 100 µL 15 µL 5 µL
ACSF.1/trehalose + blockers + papain	ACSF.1/trehalose + blockers One vial Worthington PAP2 reagent (150 U, final 10 U/mL) 10 kU/mL DNase I (Roche)	15 mL 15 µL
Low-BSA Quench buffer	ACSF.1/trehalose + blockers 10 kU/mL DNase I (Roche) 20% BSA dissolved in water (final conc. 2 mg/mL) 10 mg/mL ovomucoid inhibitor (Sigma T9253, final conc. 0.1 mg/mL)	15 mL 15 µL 150 µL 150 µL

High-BSA Quench buffer	ACSF.1/trehalose + blockers	15 mL
	10 kU/mL DNase I (Roche)	15 μ L
	20% BSA dissolved in water (final conc. 10 mg/mL)	750 μ L
	10 mg/mL ovomucoid inhibitor	150 μ L
	(Sigma T9253, final conc. 0.1 mg/mL)	
ACSF.1/trehalose + EDTA	HEPES	20 mM
	Sodium Pyruvate	3 mM
	Taurine	10 μ M
	Thiourea	2 mM
	D-(+)-glucose	25 mM
	Myo-inositol	3 mM
	Sodium bicarbonate	25 mM
	Potassium chloride	2.5 mM
	Monosodium Phosphate	1.25 mM
	Trehalose	132 mM
	HCl	2.9 mM
	EDTA	0.25 mM
	N-methyl-D-(+)-glucamine	30 mM
	L-ascorbic acid	5.0 mM
	N-acetyl-L-cysteine	12 mM
	Adjust pH to 7.3-7.4 with HCl and adjust osmolarity to 295-305.	
	Sterile filter, and then make 100 mL aliquots and freeze them (-20).	
	The thawed aliquot keeps 2-3 months at 4 degrees, until it turns yellow.	
Cell Resuspension Buffer	ACSF.1/trehalose + EDTA	50 mL
	100 μ M TTX (final 0.1 μ M)	50 μ L
	25 mM DL-AP5 (final 50 μ M)	100 μ L
	60 mM DNQX (final 20 μ M)	15 μ L
	100 mM (+)-MK801 (final 10 μ M)	5 μ L
	20% BSA dissolved in water (final conc. 2 mg/mL)	150 μ L
	4'-diamino-phenylindazole (DAPI)	1 μ g/mL

[0150] (ix) Closing Paragraphs. 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.

[0151] 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.

[0152] 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).

[0153] 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.

[0154] 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.

[0155] 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.

[0156] 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.

[0157] 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.

[0158] "% 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.

[0159] 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.

[0160] 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 pairings: Grik1_enhGad2-1 / GABAergic neurons generally; Grik1_enhGad2-2 / GABAergic neurons generally; mscRE5 / GABAergic neurons generally; mscRE8 / GABAergic neurons generally; eHGT_019h / lysosomal associated membrane protein 5 (Lamp5) neurons; eHGT_022h (also referred to herein as eHGT_022m) / Lamp5 and Vip neurons; eHGT_017h / Lamp5, Vip, and somatostatin (Sst) neurons; eHGT_17m / Lamp5, Vip, and Sst neurons; eHGT_079h / parvalbumin (Pvalb) neuron cell types; eHGT_082h / Pvalb neuron cell types and deep cerebellar nuclear neurons; eHGT_086h / Pvalb neuron cell types; eHGT_128h / Pvalb neuron cell types; eHGT_140h / Pvalb neuron cell types; eHGT_064h / Pvalb and Sst neuron cell types; eHGT_023h / Pvalb cell types and cerebellar Purkinje cells; and eHGT_359 / Pvalb cell types and cerebellar Purkinje cells.

[0161] In particular embodiments, artificial means not naturally occurring.

[0162] 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.

[0163] 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.

[0164] 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.

[0165] 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.

[0166] 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.

[0167] 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.

[0168] 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.

[0169] 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.

[0170] 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 selected from eHGT_140h, Grik1_enhGad2-1, Grik1_enhGad2-2, mscRE5, mscRE8, eHGT_019h, eHGT_022h, eHGT_022m, eHGT_017h, eHGT_17m, eHGT_079h, eHGT_082h, eHGT_086h, eHGT_128h, eHGT_064h, eHGT_023h, and eHGT_359; (ii) a promoter; and (iii) a heterologous encoding sequence.
2. The artificial expression construct of claim 1, wherein the heterologous encoding sequence encodes an effector element or an expressible element.
3. The artificial expression construct of claim 2, wherein the effector element includes a reporter protein or a functional molecule.
4. The artificial expression construct of claim 3, wherein the reporter protein comprises a fluorescent protein.
5. The artificial expression construct of claim 3, wherein the functional molecule comprises a functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or a designer receptor exclusively activated by designer drugs (DREADD).
6. The artificial expression construct of claim 2, wherein the expressible element comprises a non-functional molecule.
7. The artificial expression construct of claim 6, wherein the non-functional molecule comprises a non-functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or DREADD.
8. The artificial expression construct of claim 1, wherein the artificial expression construct is associated with a capsid that crosses the blood brain barrier.
9. The artificial expression construct of claim 8, wherein the capsid includes PHP.eB, AAV-BR1, AAV-PHP.S, AAV-PHP.B, or AAV-PPS.
10. The artificial expression construct of claim 1, wherein the artificial expression construct includes or encodes a skipping element.
11. The artificial expression construct of claim 10, wherein the skipping element includes a 2A peptide or an internal ribosome entry site (IRES).
12. The artificial expression construct of claim 11, wherein the 2A peptide comprises T2A, P2A,

E2A, or F2A.

13. The artificial expression construct of claim 1, wherein the artificial expression construct includes or encodes a set of features selected from: eHGT_140h, Grik1_enhGad2-1, Grik1_enhGad2-2, mscRE5, mscRE8, eHGT_019h, eHGT_022h, eHGT_022m, eHGT_017h, eHGT_17m, eHGT_079h, eHGT_082h, eHGT_086h, eHGT_128h, eHGT_064h, eHGT_023h, and eHGT_359, AAV, scAAV, rAAV, minBglobin, CMV, minCMV, minRho, minRho*, fluorescent protei, Cre, iCre, dgCre, FlpO, tTA2, SP10, WPRE, and/or BGHpA.

14. The artificial expression construct of claim 1, wherein the artificial expression construct includes or encodes a set of features selected from

hsA2-eHGT_140h-minRho-SYFP2-WPRE3-BGHPA;

Grik1_enhGad2-1-Hsp68-EGFP-WPRE3-BGHPA;

Grik1_enhGad2-1-pBGmin-EGFP-WPRE3-BGHPA;

Grik1_enhGad2-2-Hsp68-EGFP-WPRE3-BGHPA;

mscRE5-pBGmin-EGFP-WPRE3-BGHPA;

mscRE5-pBGmin-FlpO-WPRE3-BGHPA;

mscRE8-pBGmin-EGFP-WPRE3-BGHPA;

mscRE8-pBGmin-FlpO-WPRE3-BGHPA;

scAAV-eHGT_019h-minBGlobin-SYFP2-WPRE3-BGHPA;

scAAV-eHGT_022h-minBGlobin-SYFP2-WPRE3-BGHPA;

scAAV-eHGT_022m-minBGlobin-SYFP2-WPRE3-BGHPA;

scAAV-eHGT_017h-minBGlobin-SYFP2-WPRE3-BGHPA;

hsA2-eHGT_079h-minRho-SYFP2-WPRE3-BGHPA;

hsA2-eHGT_082h-minRho-SYFP2-WPRE3-BGHPA;

hsA2-eHGT_086h-minRho-SYFP2-WPRE3-BGHPA;

hsA2-eHGT_128h-minRho-SYFP2-WPRE3-BGHPA;

eHGT_064h-minBglobin-SYFP2-WPRE3-BGHPA; or

scAAV-eHGT_023h-minBGlobin-SYFP2-WPRE3-BGHPA.

15. A vector comprising an artificial expression construct of claim 1.

16. The vector of claim 15, wherein the vector comprises a viral vector.

17. The vector of claim 16, wherein the viral vector comprises a recombinant adeno-associated viral (AAV) vector.

18. 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 selected from eHGT_140h, Grik1_enhGad2-1, Grik1_enhGad2-2, mscRE5, mscRE8, eHGT_019h, eHGT_022h, eHGT_022m, eHGT_017h, eHGT_17m, eHGT_079h, eHGT_082h, eHGT_086h, eHGT_128h, eHGT_064h, eHGT_023h and eHGT_359.

19. The AAV vector of claim 18, wherein the heterologous encoding sequence encodes an effector element or an expressible element.

20. The AAV vector of claim 19, wherein the effector element includes a reporter protein or a functional molecule.

21. The AAV vector of claim 20, wherein the reporter protein comprises a fluorescent protein.

22. The AAV vector of claim 20, wherein the functional molecule comprises a functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or DREADD.

23. The AAV vector of claim 19, wherein the expressible element comprises a non-functional molecule.

24. The AAV vector of claim 23, wherein the non-functional molecule comprises a non-functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or DREADD.

25. The AAV vector of claim 18, wherein the AAV vector is replication-competent.

26. A transgenic cell comprising an artificial expression construct of claim 1 and/or a vector of claim 18.

27. The transgenic cell of claim 26, wherein the transgenic cell is a GABAergic neuron.

28. The transgenic cell of claim 26, wherein the transgenic cell is a parvalbumin (Pvalb) neuron, a lysosomal associated membrane protein 5 (Lamp5) neuron, a vasoactive intestinal polypeptide-expressing (Vip) neuron, a somatostatin (Sst) neuron, a layer 4 (L4) intratelencephalic (IT) neuron, a layer 5 (L5) IT neuron, a deep cerebellar nuclear neuron or a cerebellar Purkinje cell.

29. The transgenic cell of claim 26, wherein the transgenic cell is murine, human, or non-human primate.

30. A non-human transgenic animal comprising an artificial expression construct of claim 1, a

vector of claim 18, and/or a transgenic cell of claim 26.

31. The non-human transgenic animal of claim 30, wherein the non-human transgenic animal is a mouse or a non-human primate.

32. An administrable composition comprising an artificial expression construct of claim 1, a vector of claim 18, and/or a transgenic cell of claim 26.

33. A kit comprising an artificial expression construct of claim 1, a vector of claim 18, a transgenic cell of claim 26, and/or a transgenic animal of claim 30.

34. 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 32 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.

35. The method of claim 34, wherein the gene encodes an effector element or an expressible element

36. The method of claim 35, wherein the effector element comprises a reporter protein or a functional molecule.

37. The method of claim 36, wherein the reporter protein comprises a fluorescent protein.

38. The method of claim 36, wherein the functional molecule comprises a functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or DREADD.

39. The method of claim 35, wherein the expressible element comprises a non-functional molecule.

40. The method of claim 39, wherein the non-functional molecule comprises a non-functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, homologous recombination donor cassette, or DREADD.

41. The method of claim 34, wherein the providing comprises pipetting.

42. The method of claim 41, wherein the pipetting is to a brain slice.

43. The method of claim 42, wherein the brain slice comprises a GABAergic neuron.

44. The method of claim 42, wherein the brain slice comprises a Pvalb neuron, a Lamp5 neuron, a Vip neuron, an Sst neuron, an L4 IT neuron, an L5 IT neuron, a deep cerebellar nuclear neuron, and/or a cerebellar Purkinje cell.

45. The method of claim 42, wherein the brain slice is murine, human, or non-human primate.
46. The method of claim 34, wherein the providing comprises administering to a living subject.
47. The method of claim 46, wherein the living subject is a human, non-human primate, or a mouse.
48. The method of claim 46, wherein the administering to a living subject is through injection.
49. The method of claim 48, wherein the injection comprises intravenous injection, intraparenchymal injection into brain tissue, intracerebroventricular (ICV) injection, intra-cisterna magna (ICM) injection, or intrathecal injection.
50. A vector consisting of or consisting essentially of CN1633, AiP1146, AiP1113, AiP1147, AiP1147, AiP1013, AiP1012, CN1525, CN1528, CN1532, CN1621, CN1259, CN2045, CN1255, CN1408, CN1258, CN1279, CN1253, or CN1274.

FIG. 1

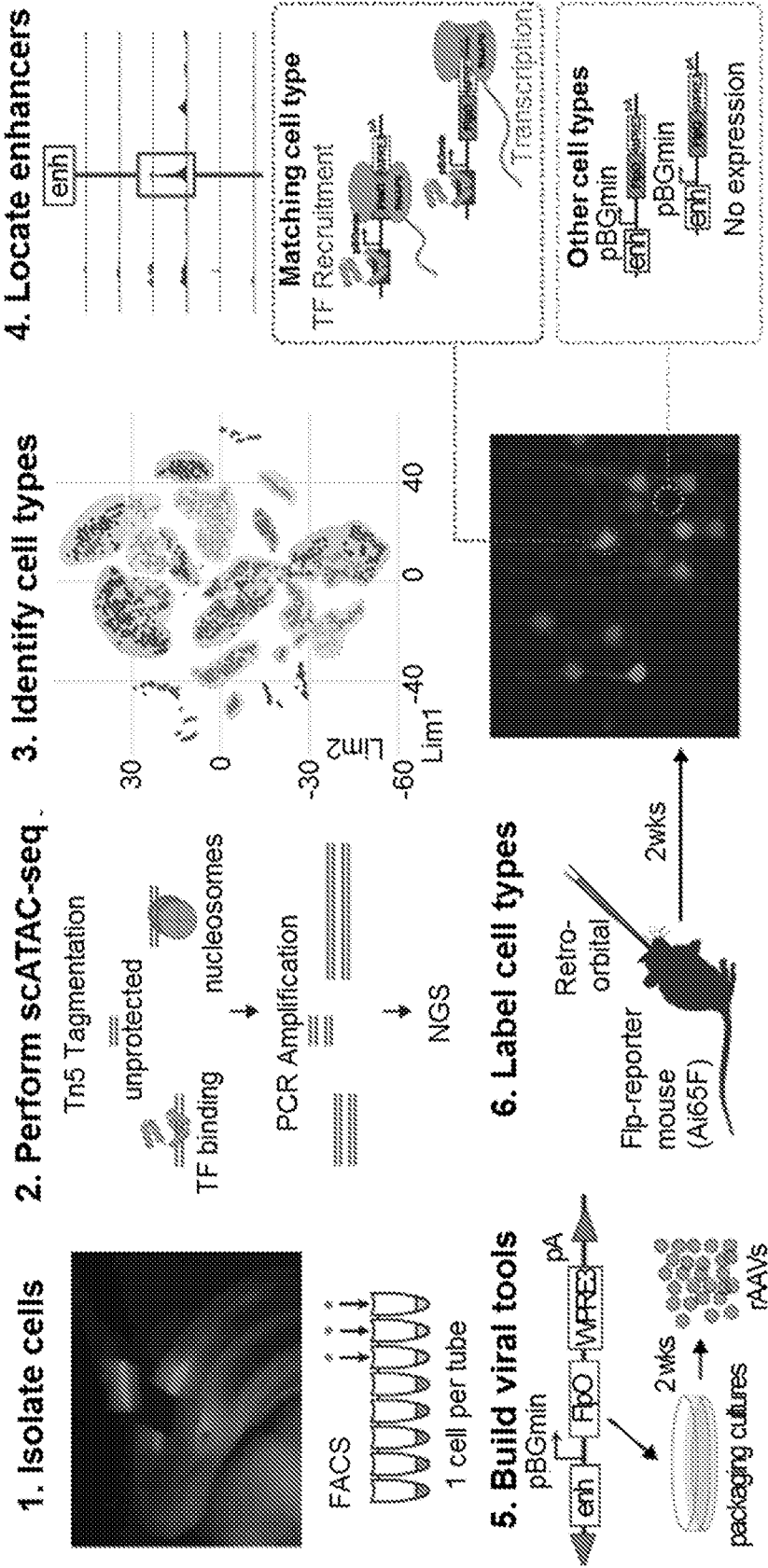


FIG. 2A

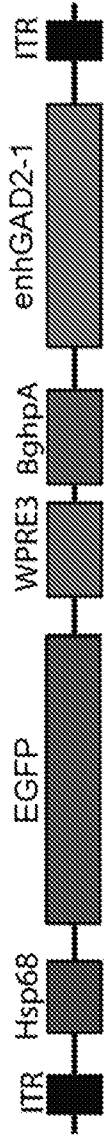


FIG. 2B

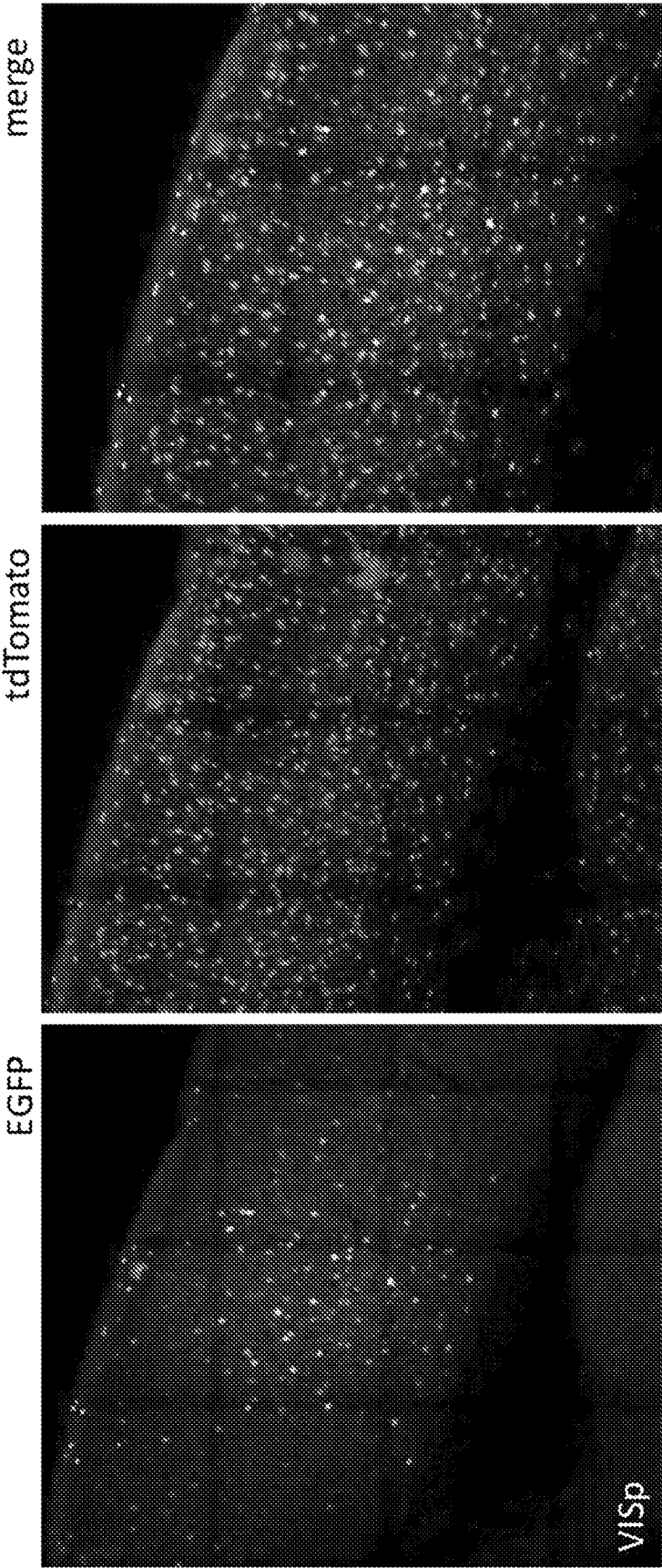
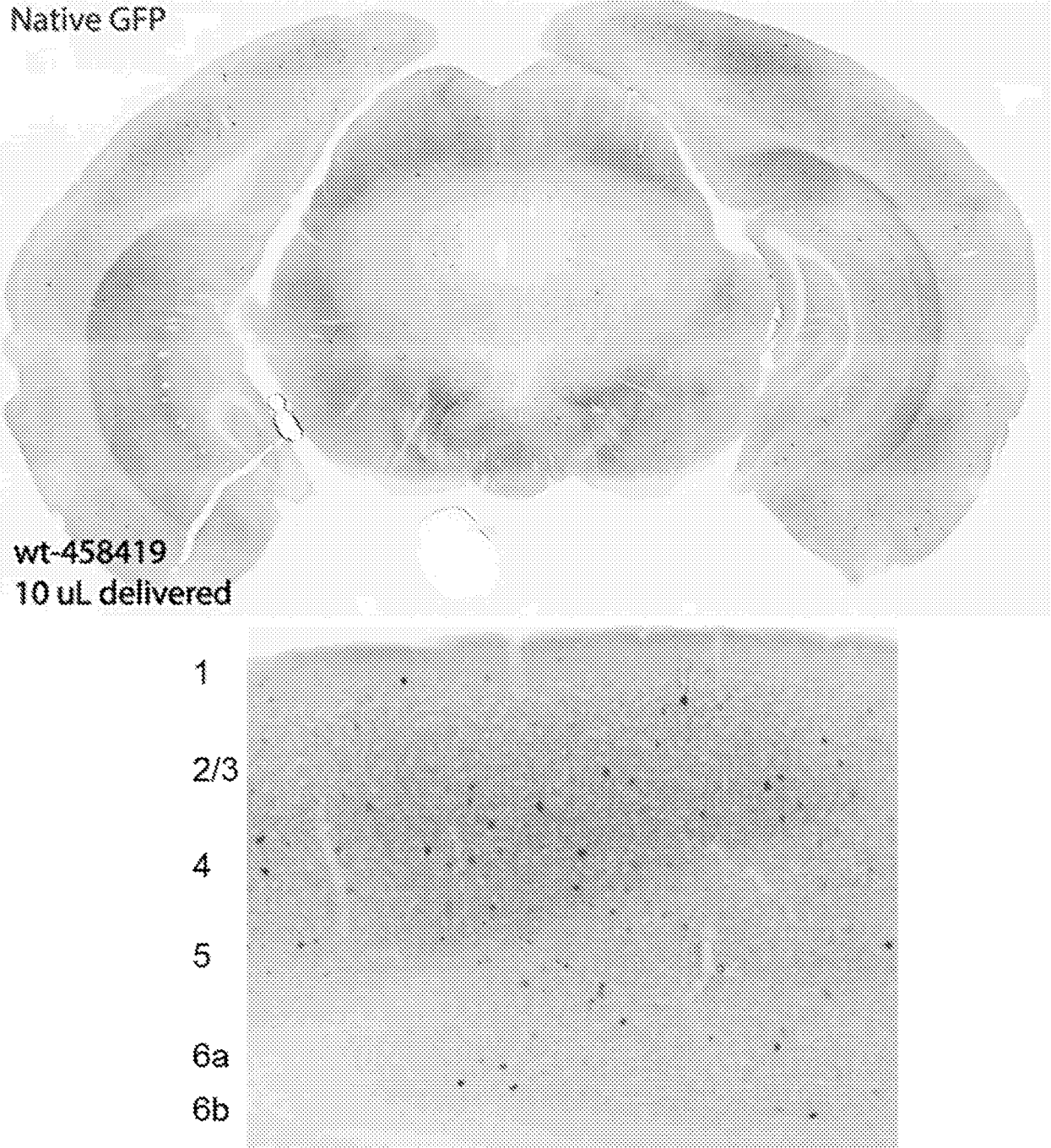


FIG. 3A



4/93

FIG. 3B

Stained rb anti-GFP
anti-rb Alexa 488

wt--458419
10 uL delivered

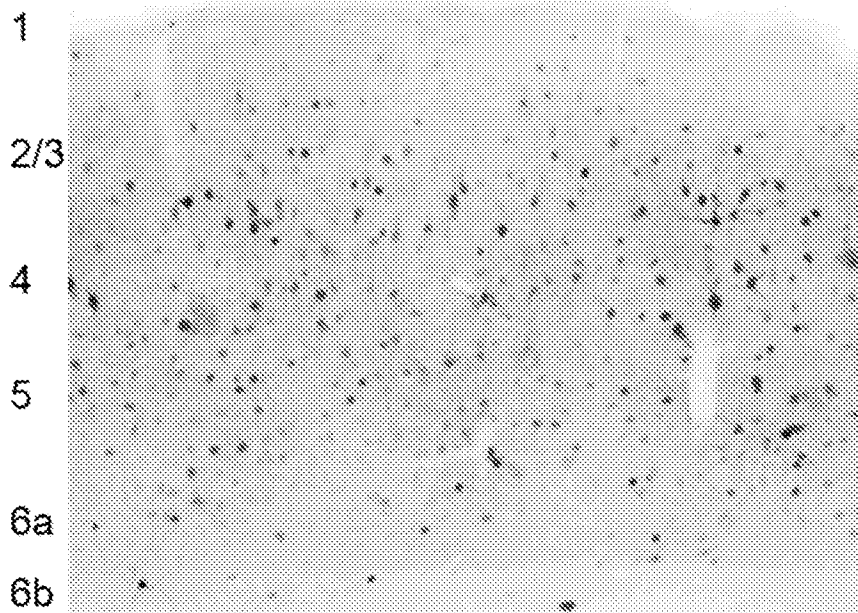


FIG. 4

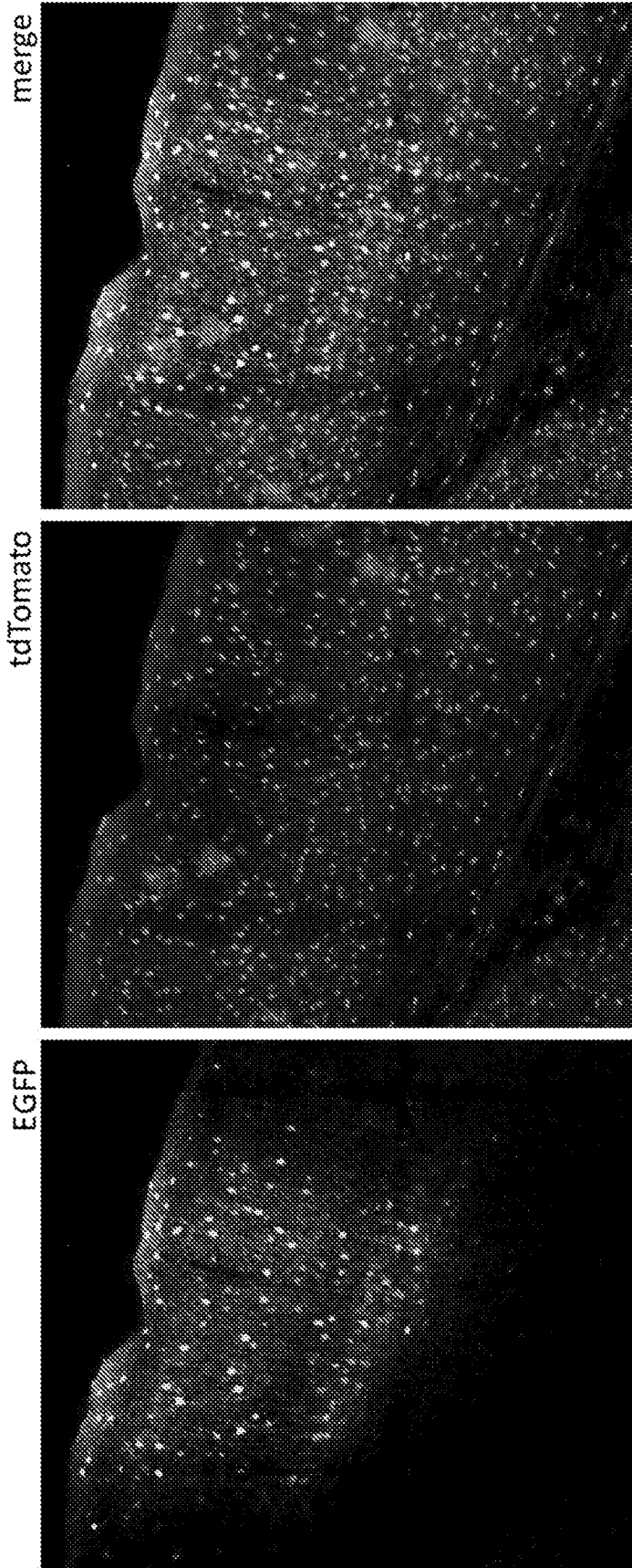


FIG. 5

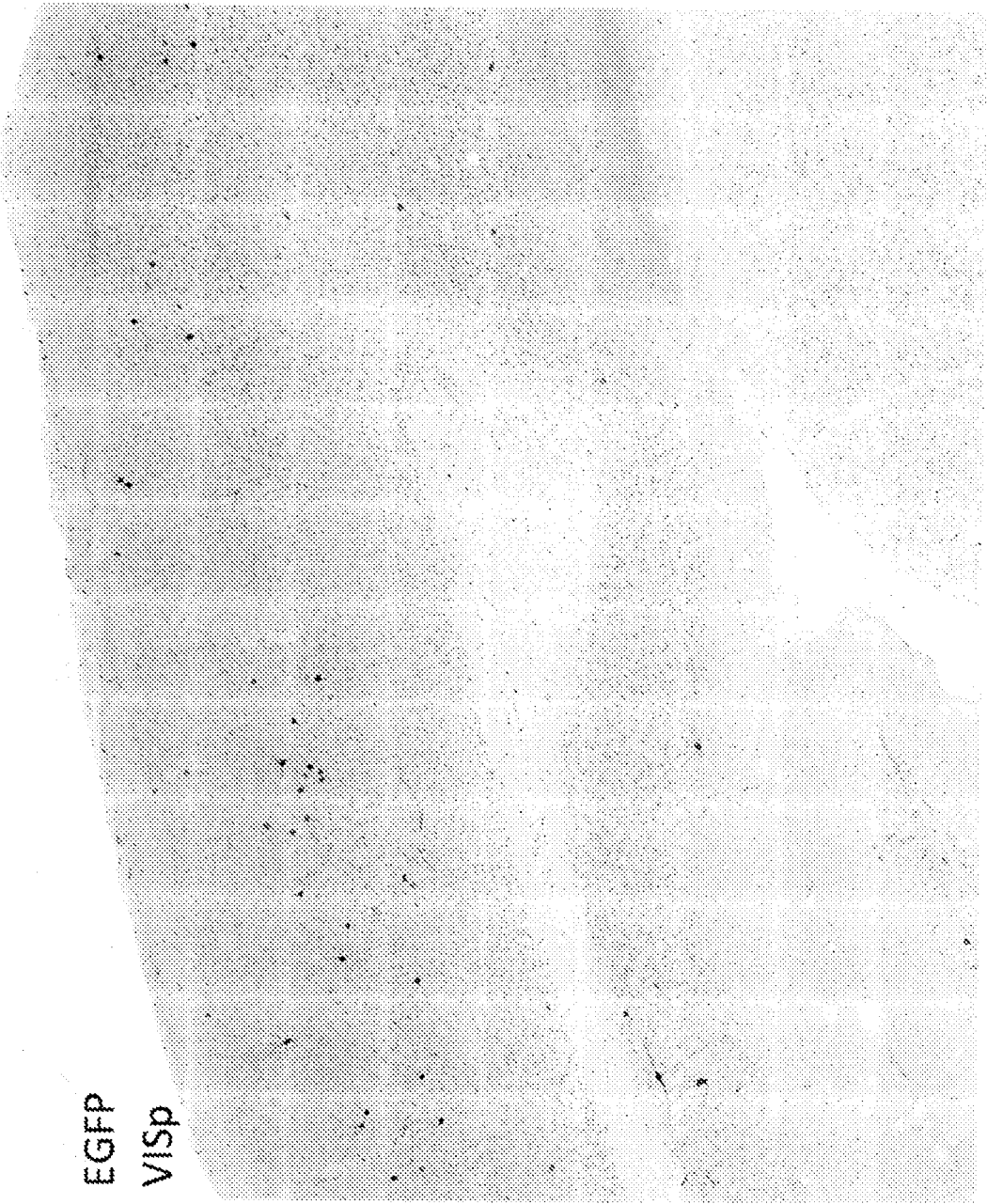


FIG. 6A

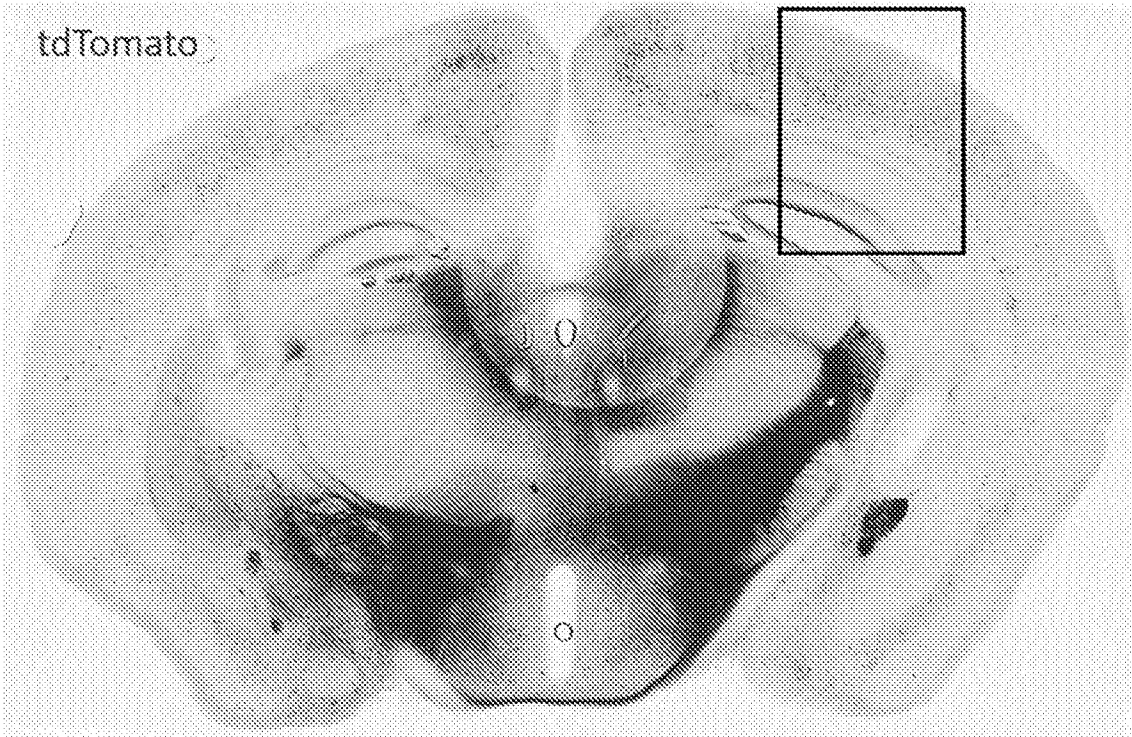


FIG. 6B

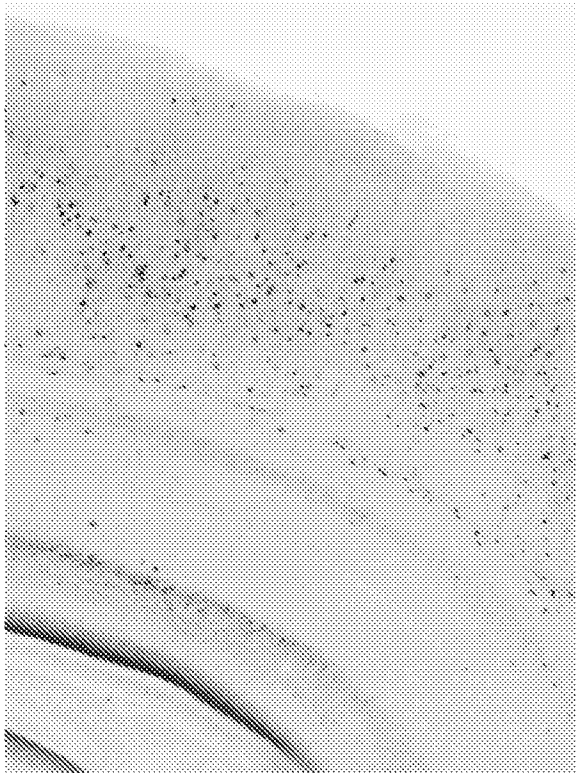


FIG. 7A

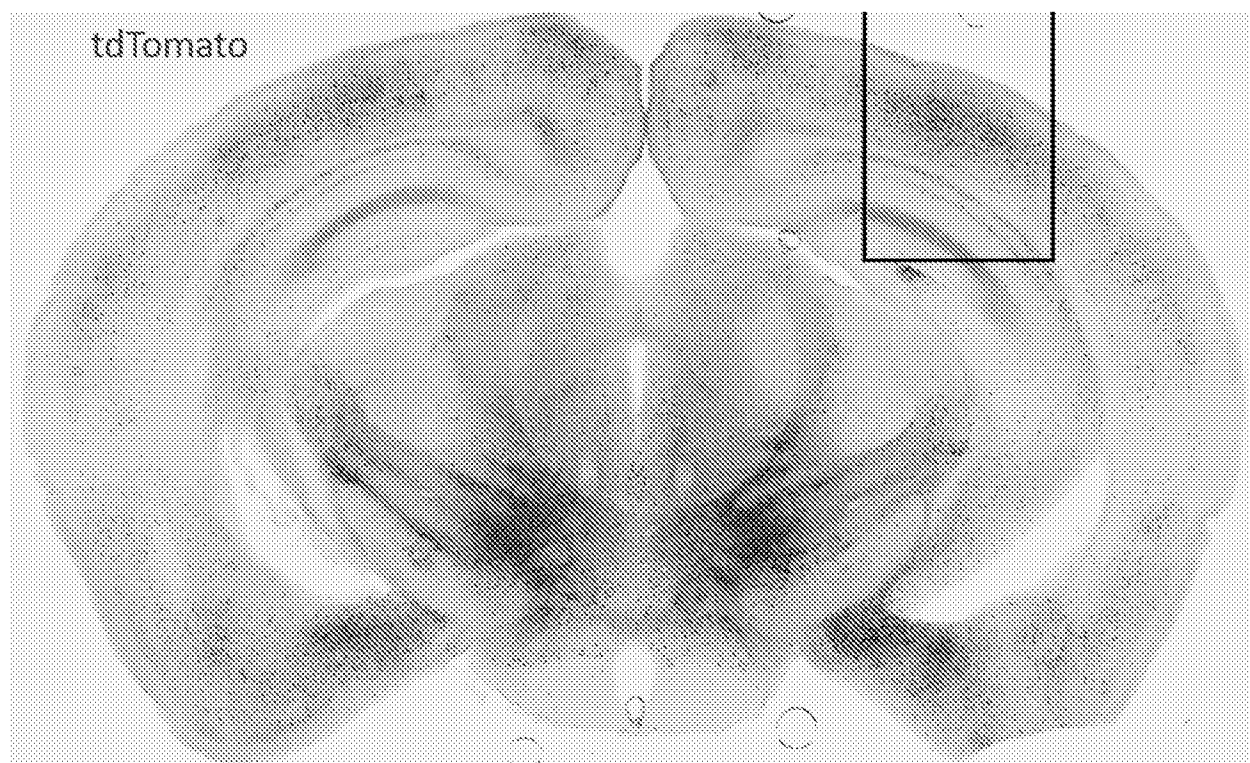


FIG. 7B

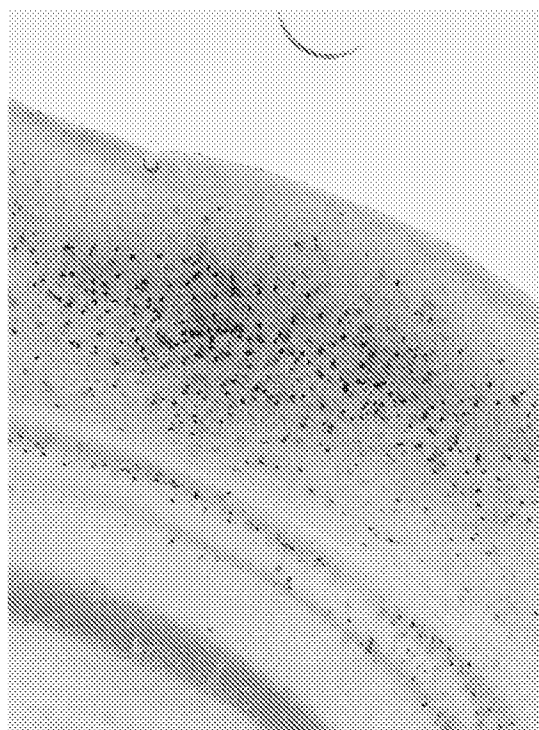


FIG. 8A



10/93

FIG. 8B

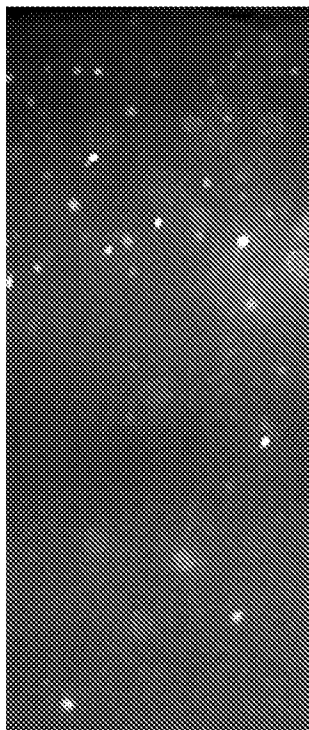


FIG. 8C

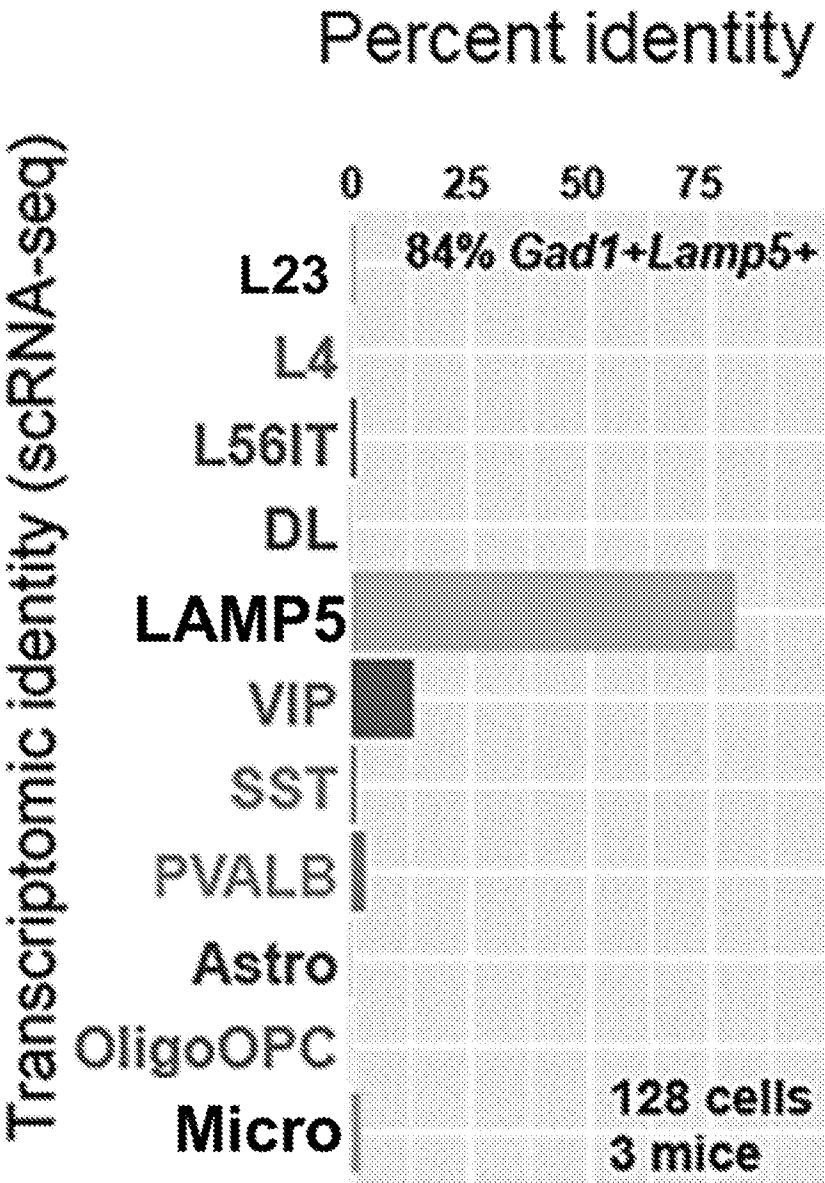


FIG. 8D

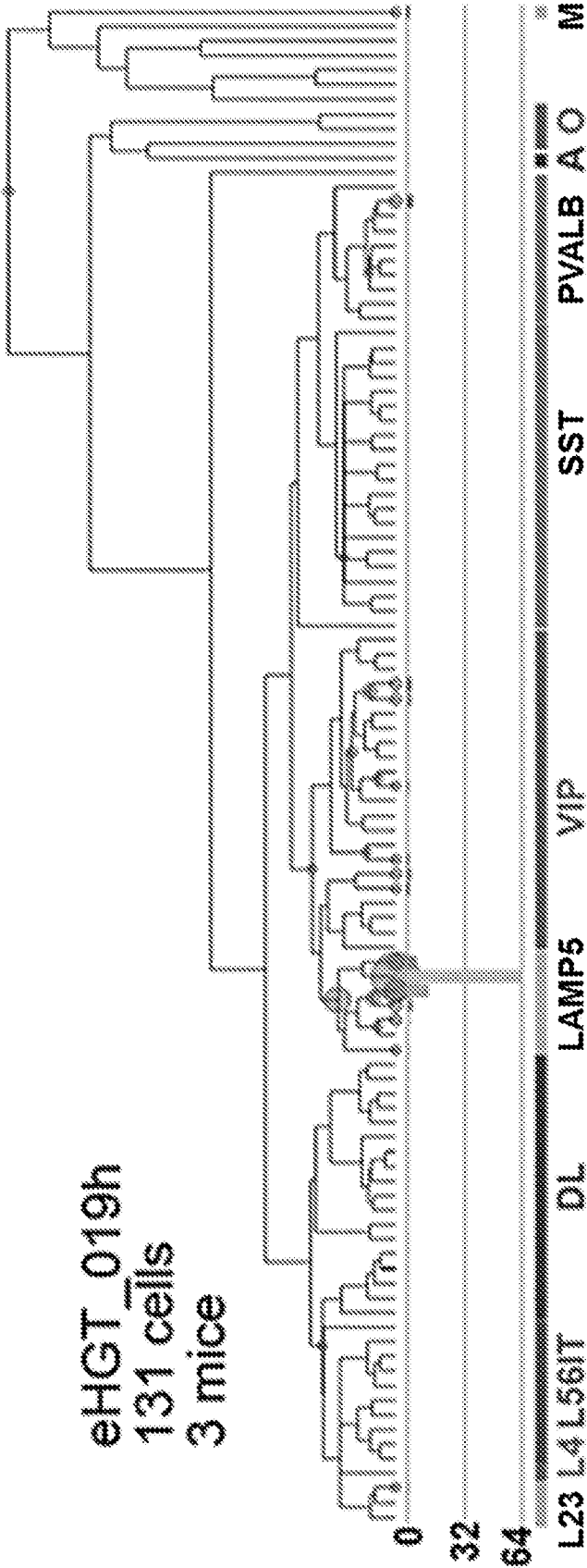


FIG. 9A

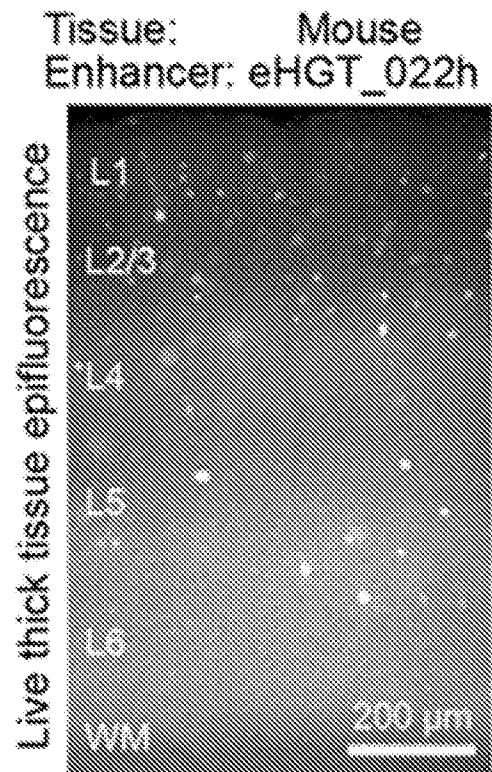


FIG. 9B

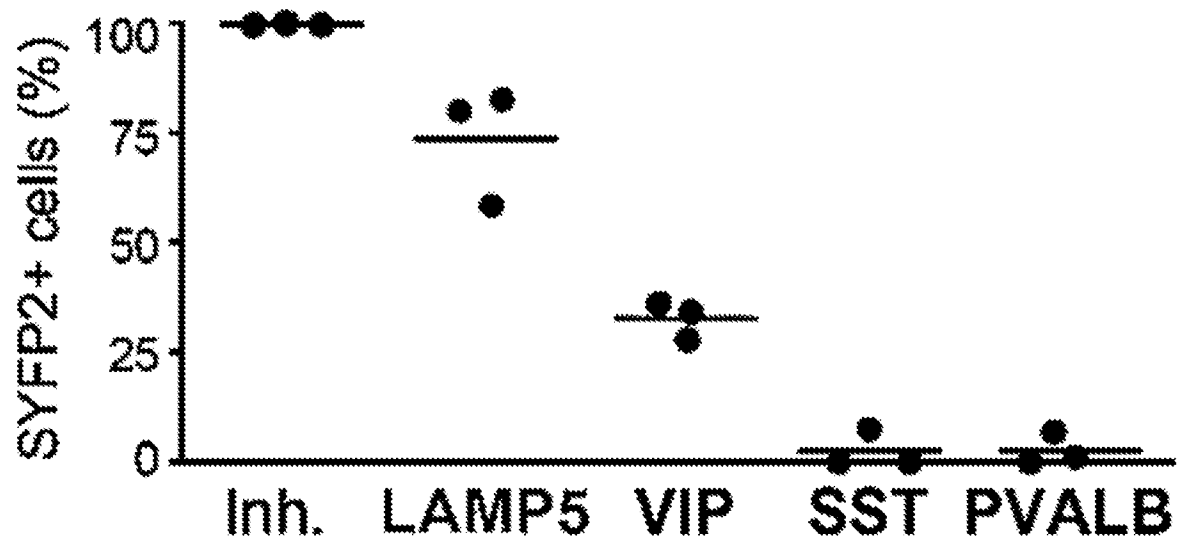


FIG. 9C

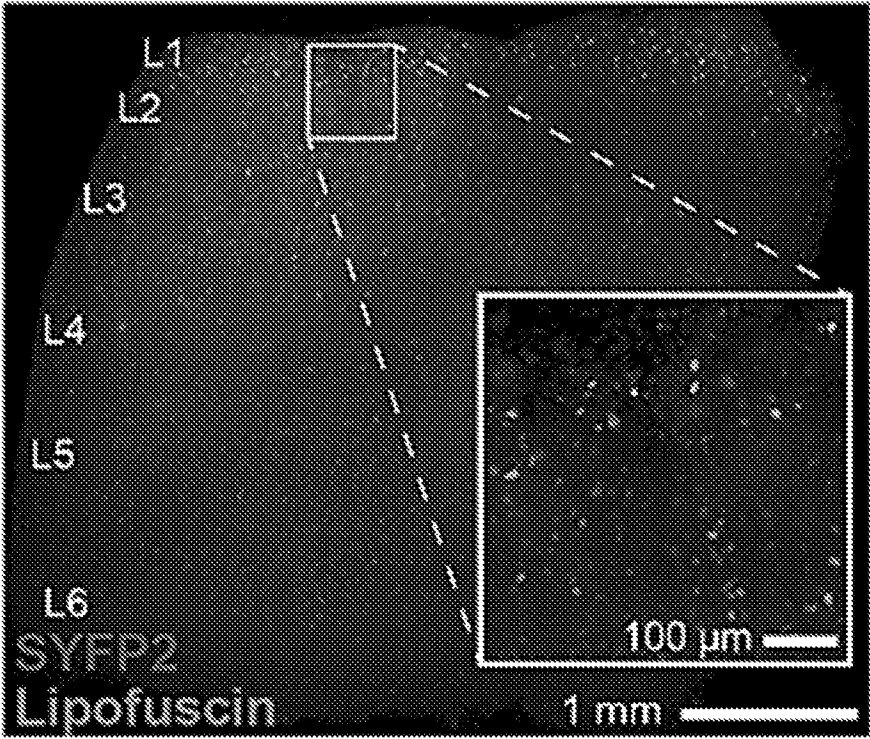


FIG. 9D

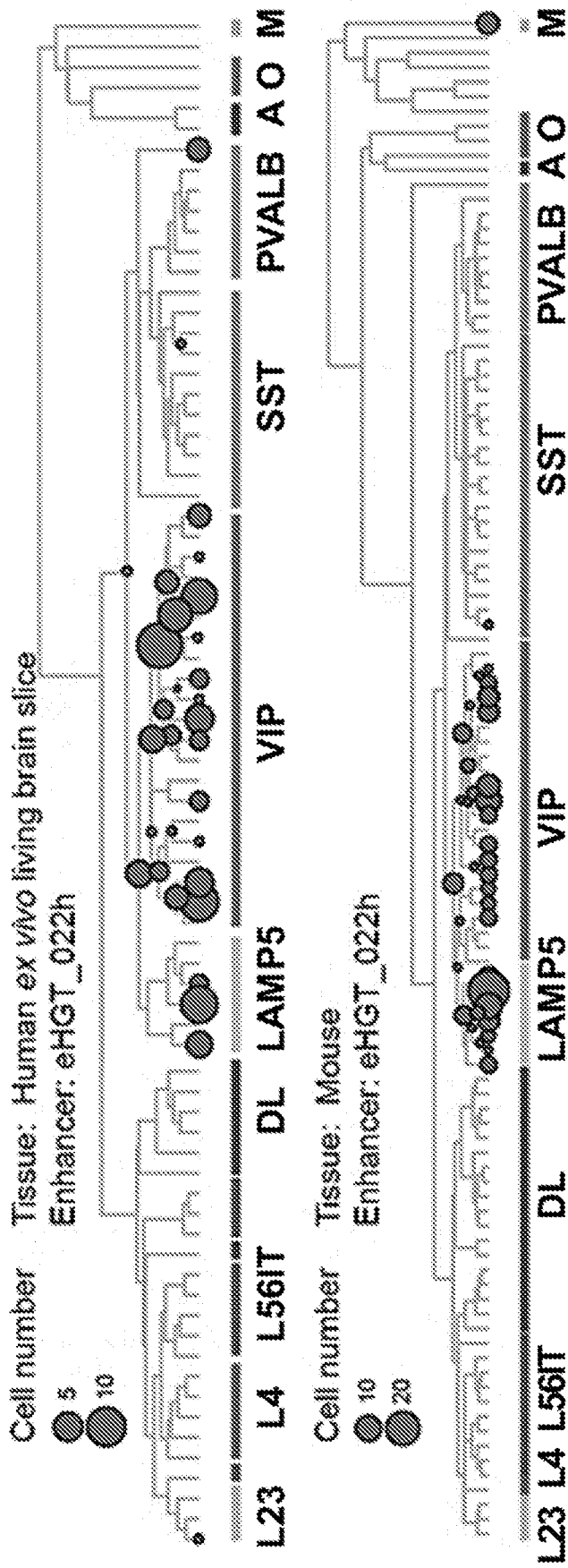


FIG. 10A

Tissue: Mouse
Enhancer: eHGT_022m

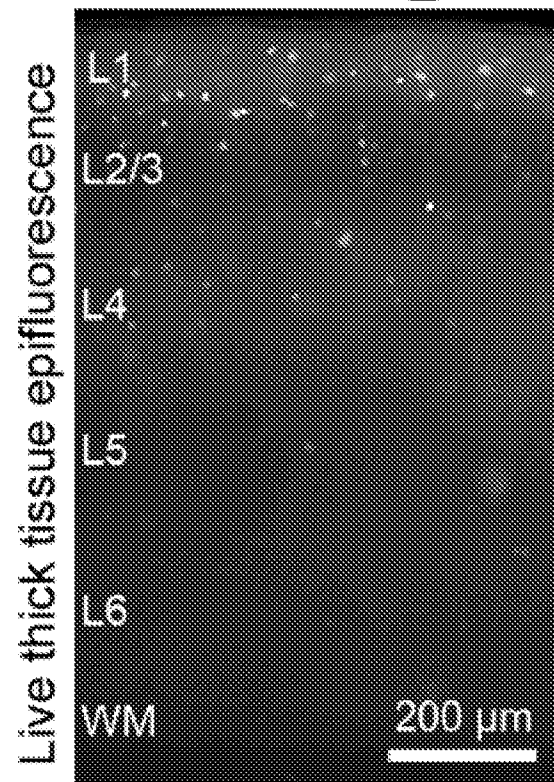


FIG. 10B

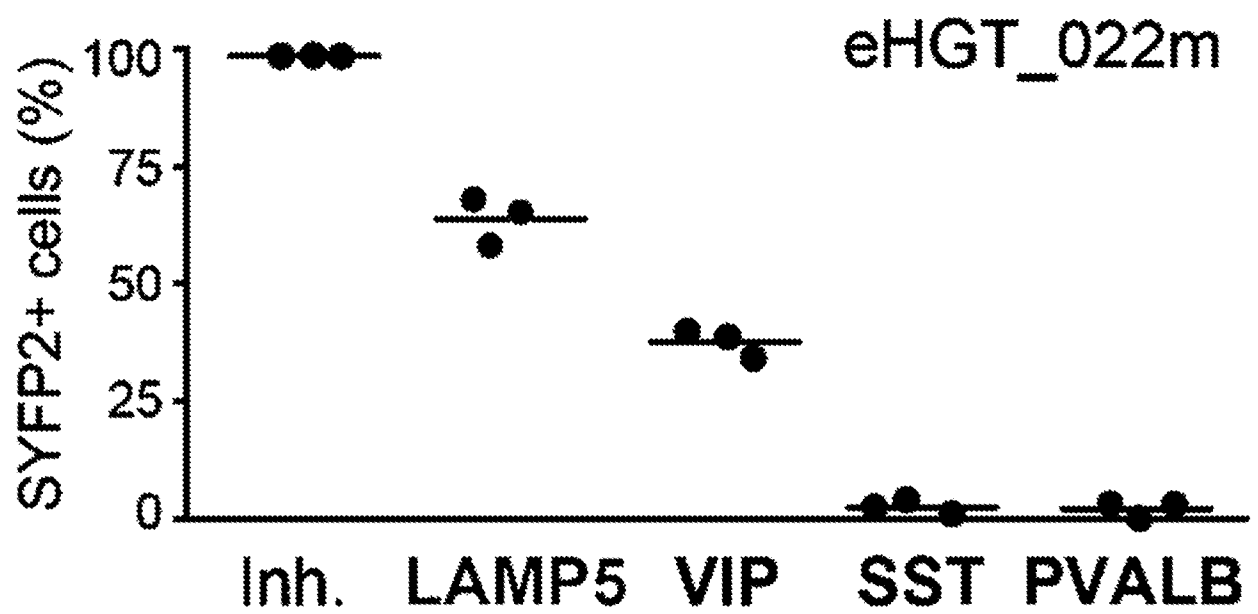


FIG. 11A



FIG. 11B

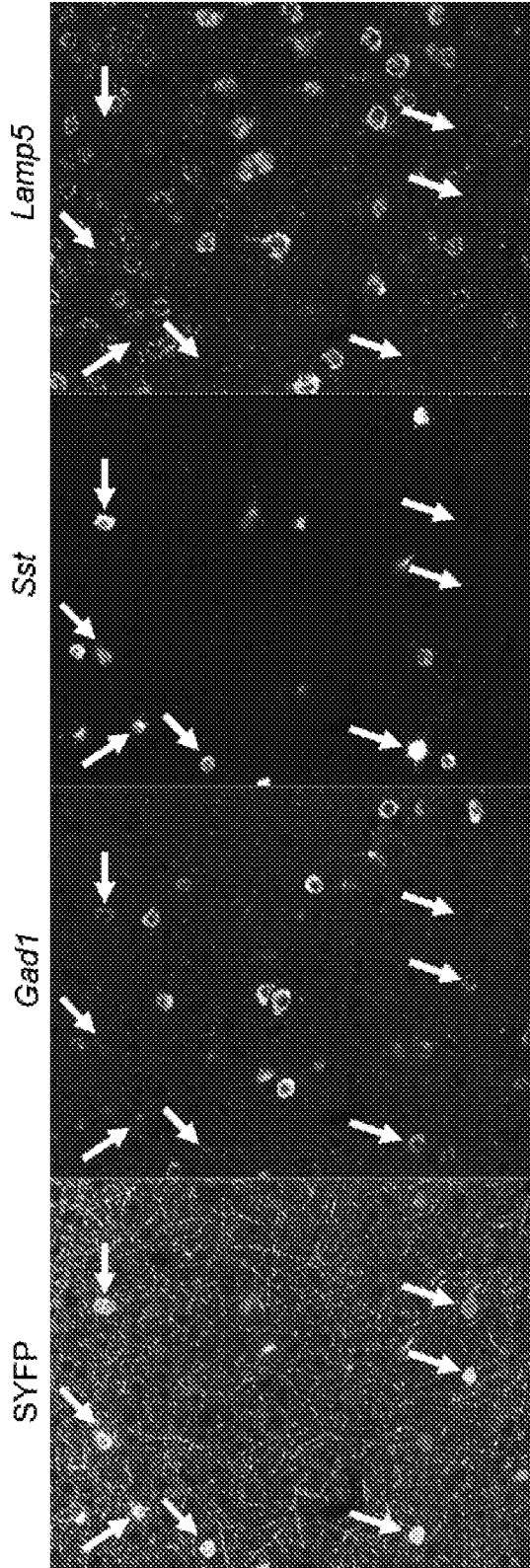


FIG. 11C

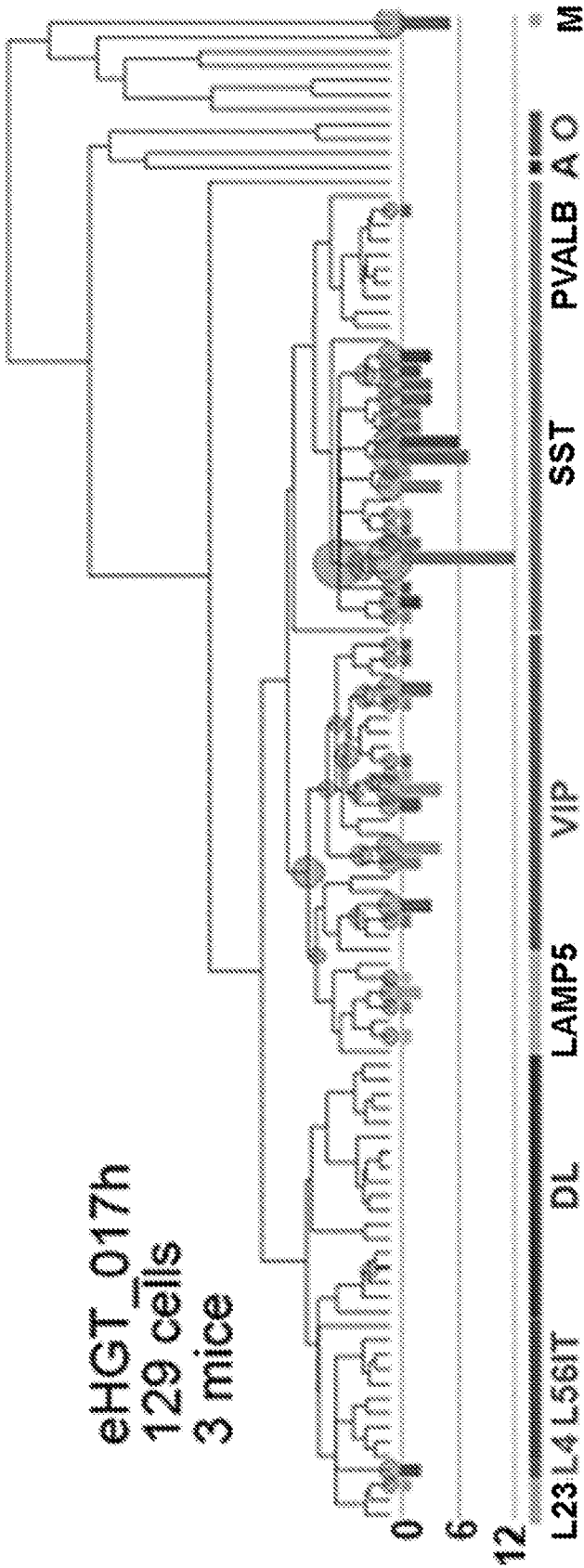


FIG. 12A



FIG. 12B

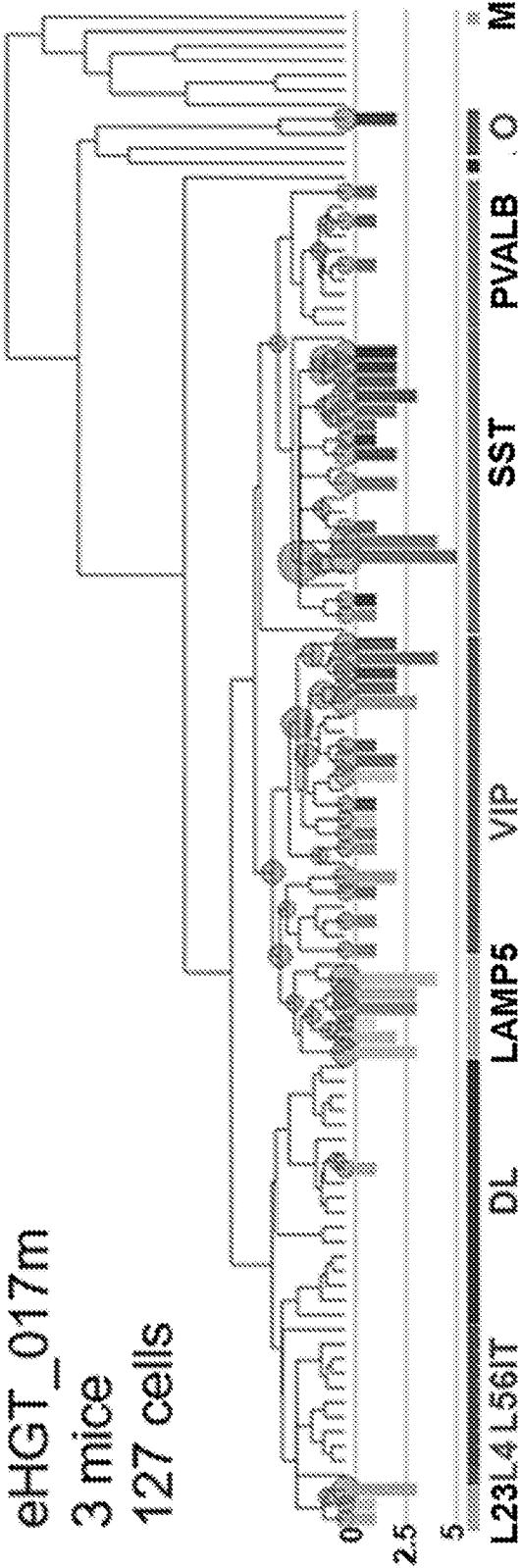


FIG. 13A



FIG. 13B

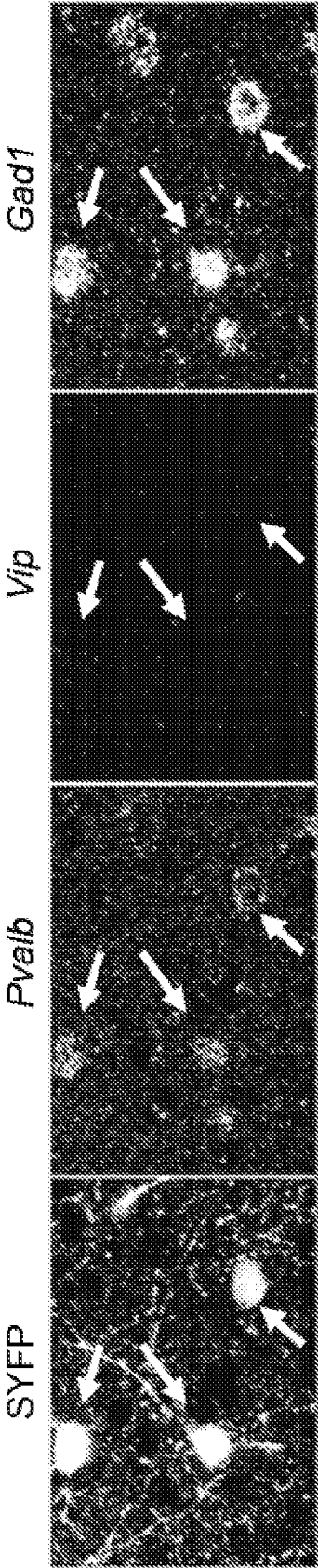


FIG. 13C

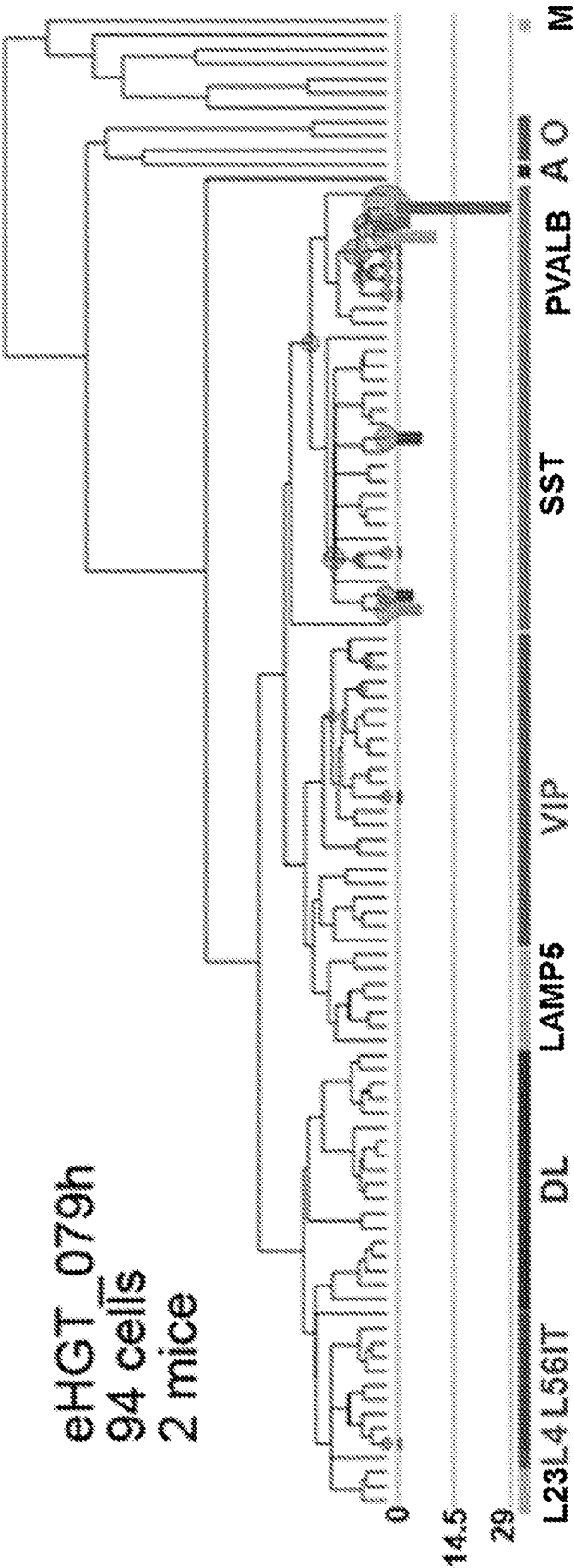


FIG. 14A



FIG. 14B

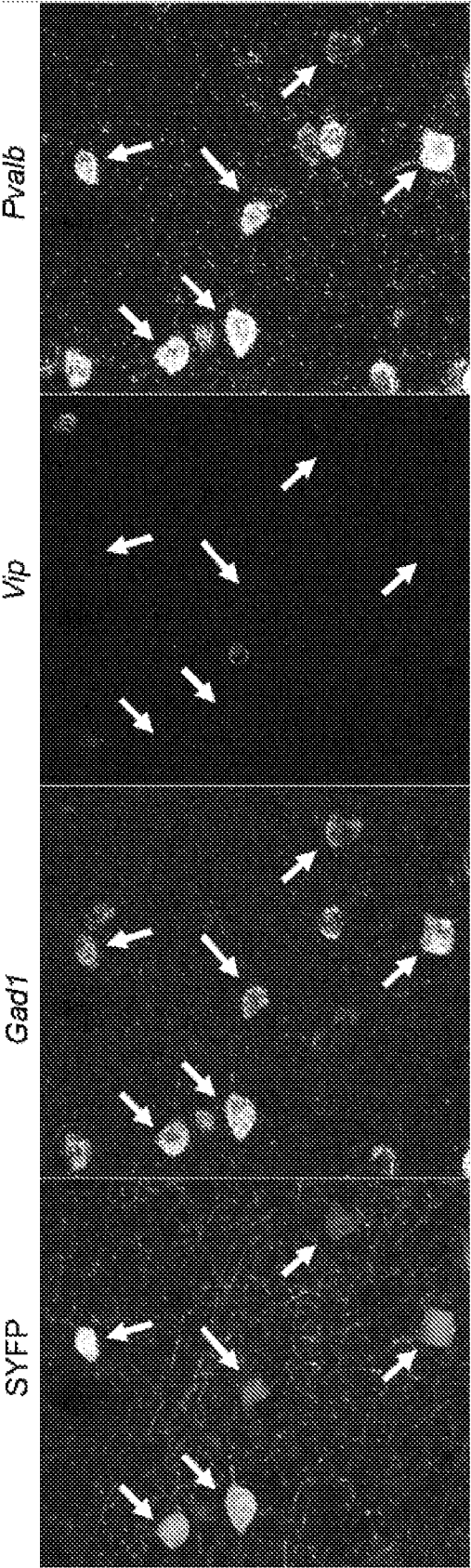


FIG. 14C

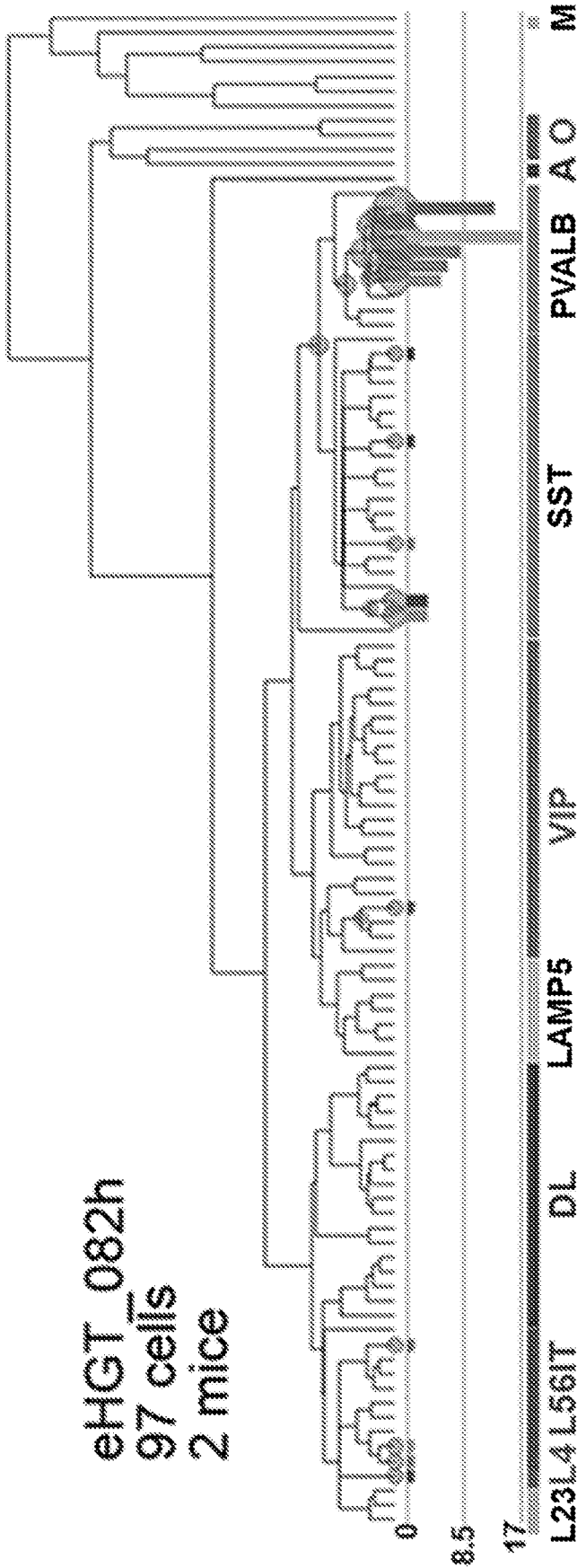


FIG. 14D

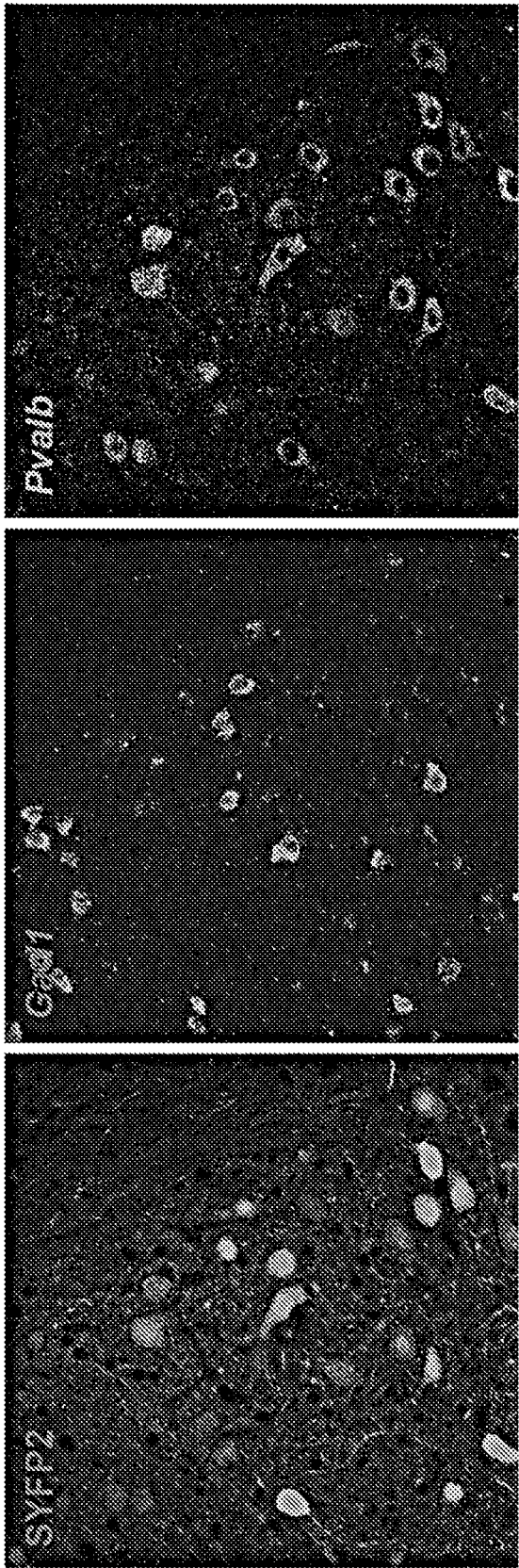


FIG. 15A



FIG. 15B

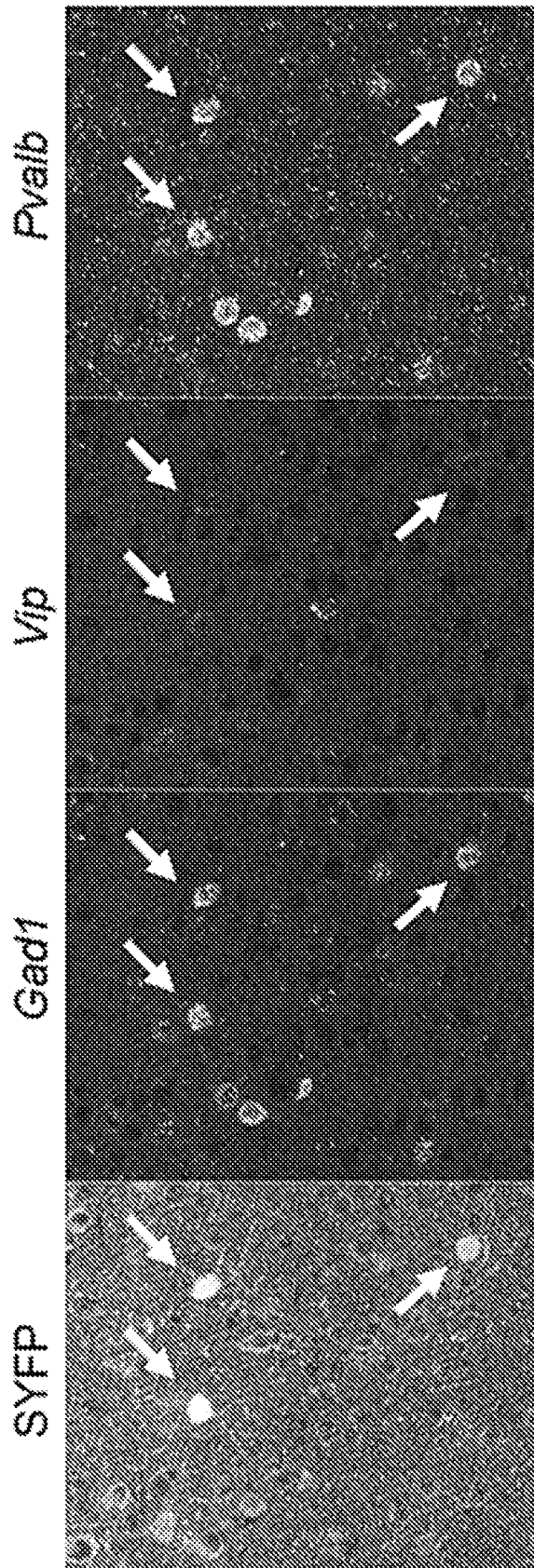


FIG. 16A



FIG. 16B

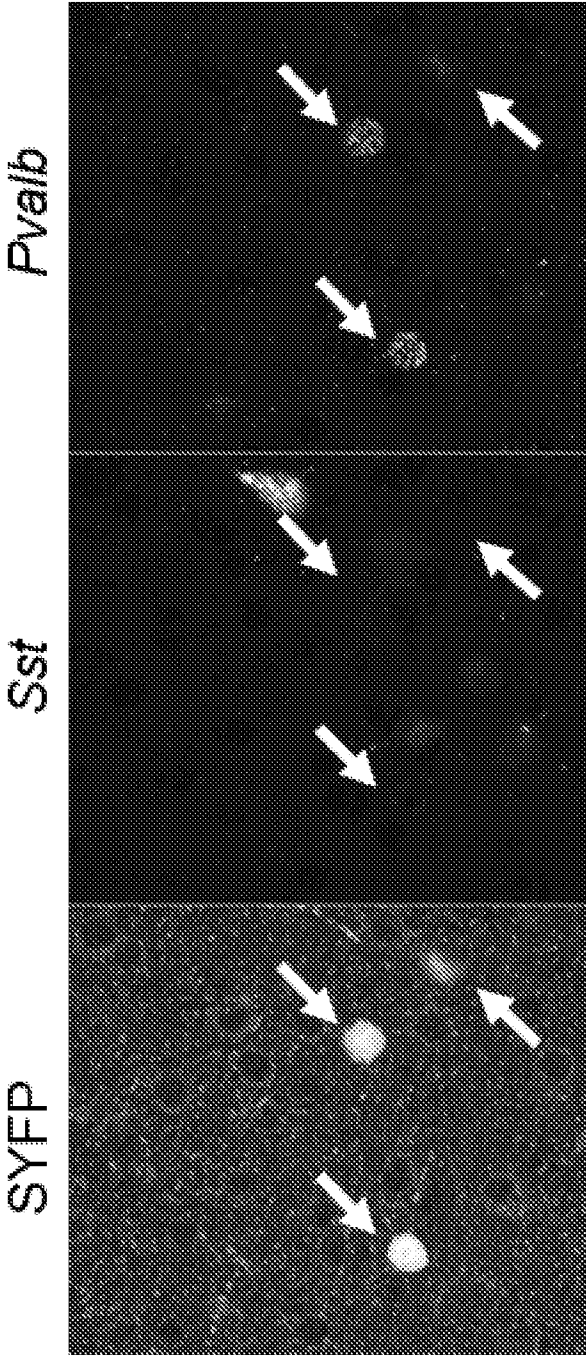


FIG. 16C

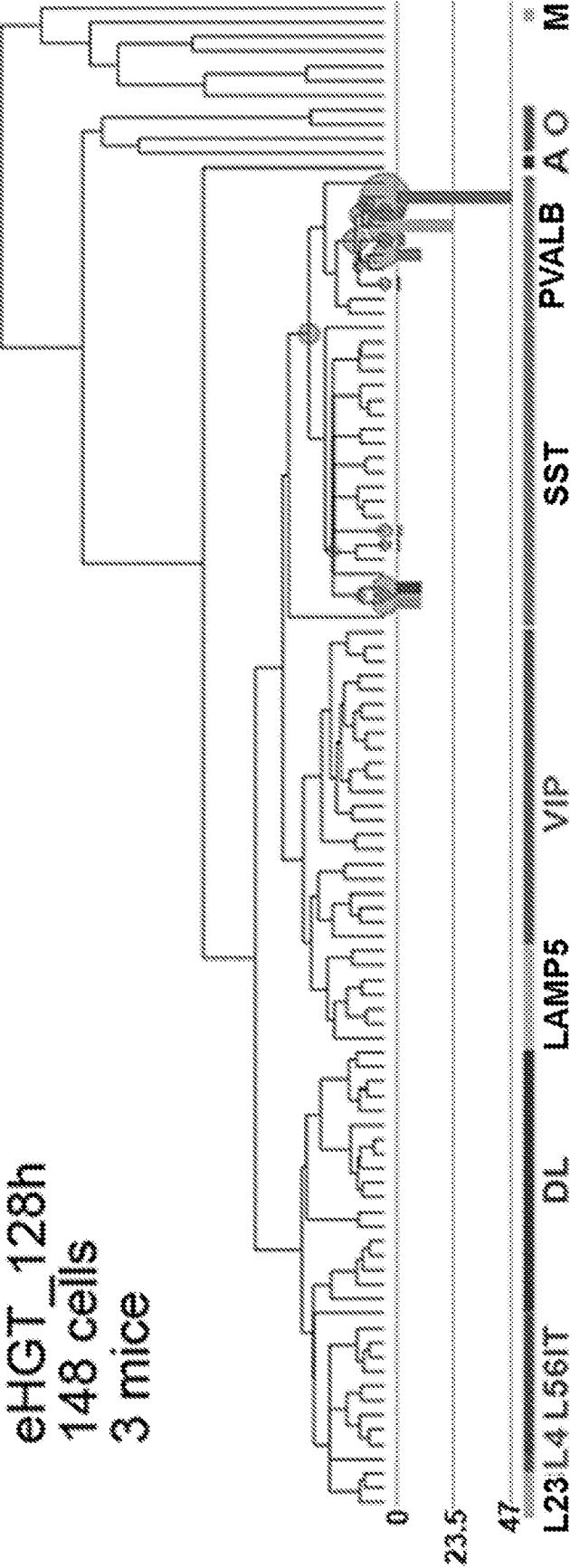


FIG. 17A



FIG. 17B

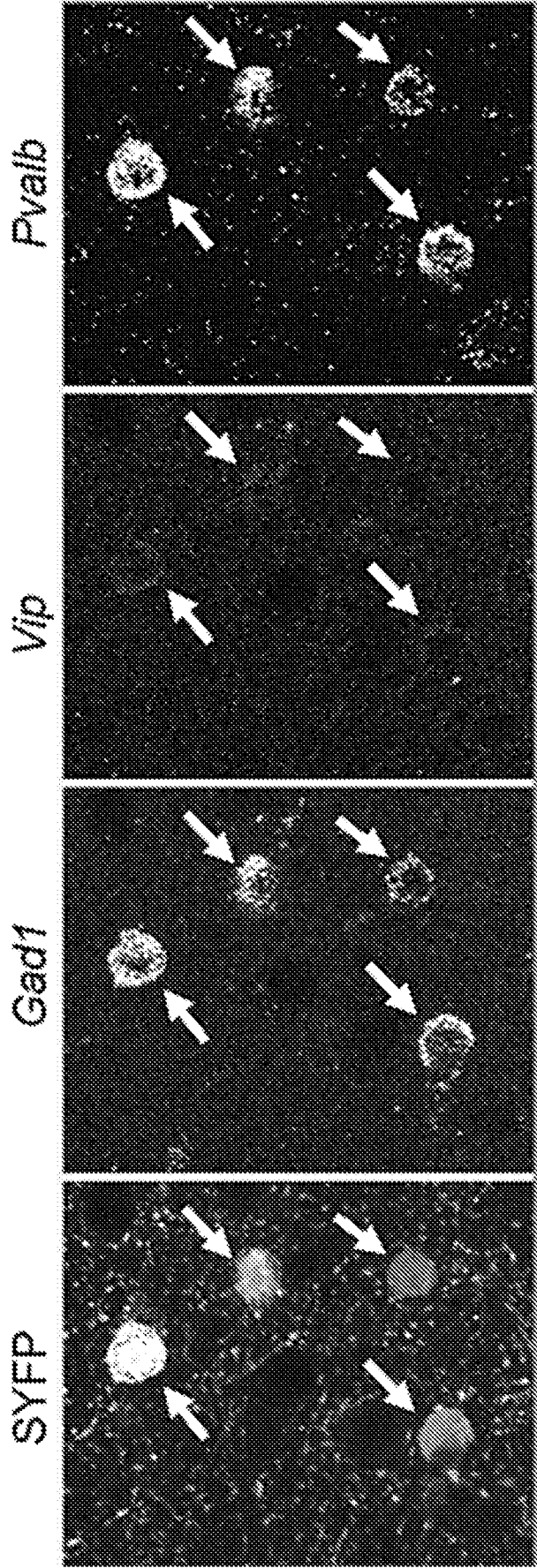


FIG. 18A



FIG. 18B

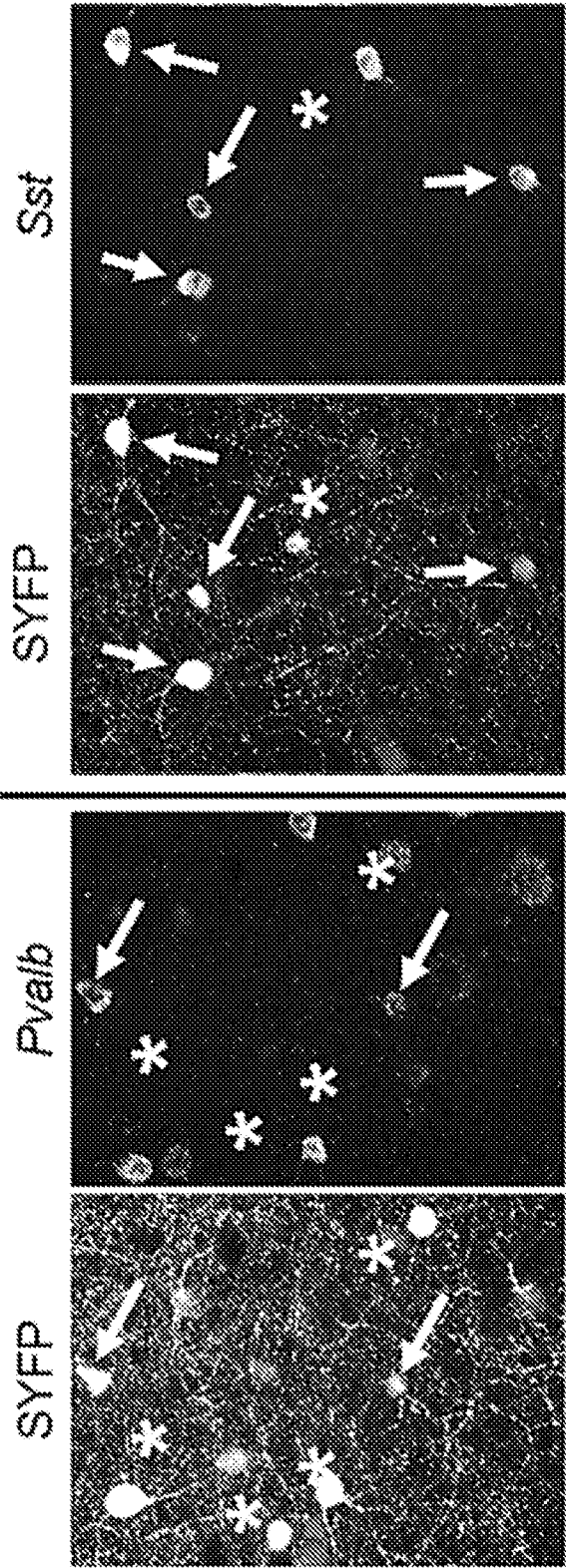


FIG. 18C

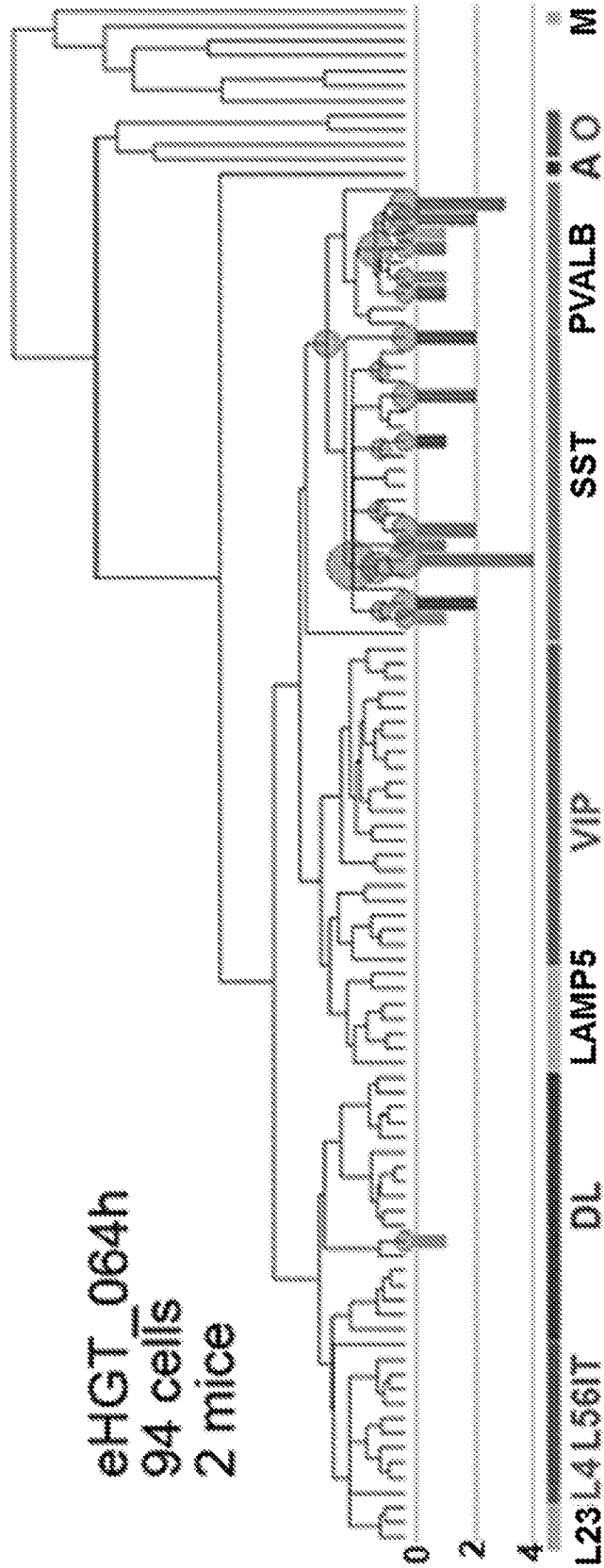


FIG. 19A



FIG. 19B

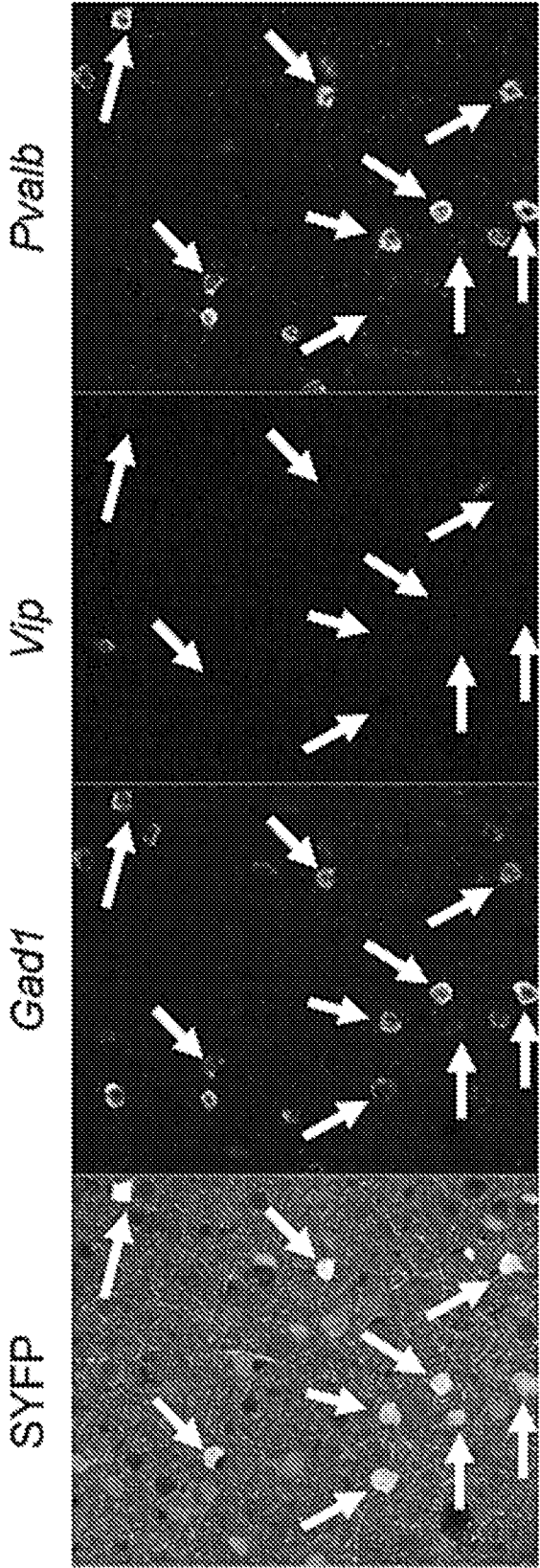


FIG. 19C

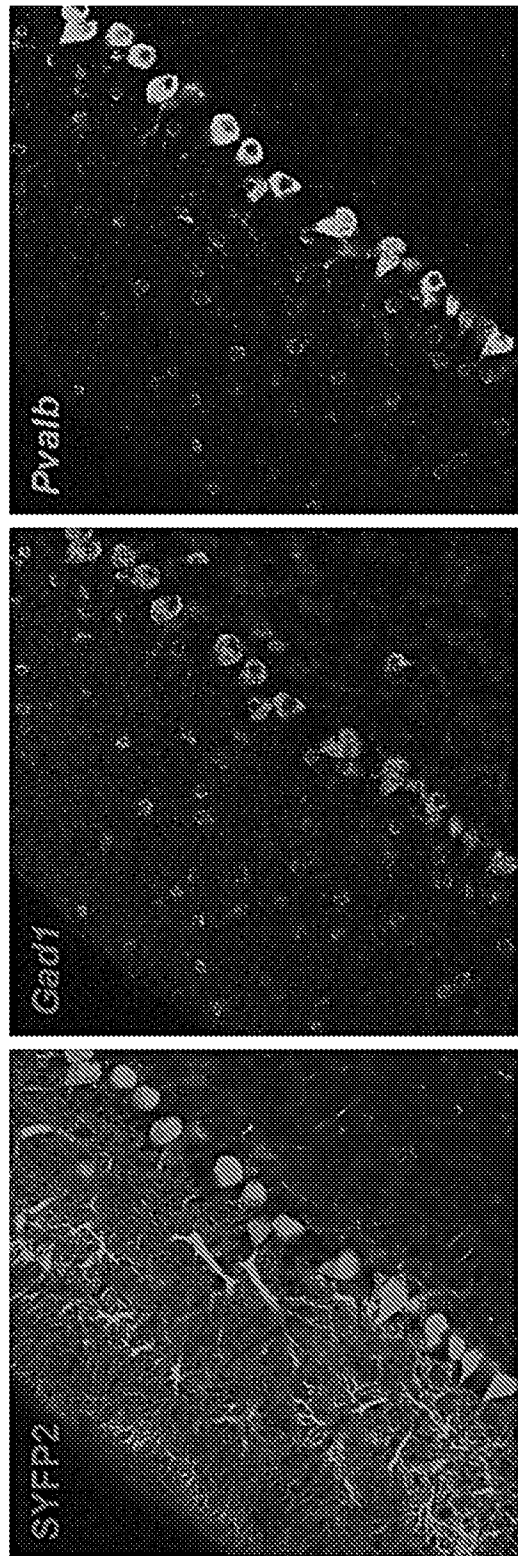


FIG. 20



FIG. 21A

Sequence	Sequence Name	Sequence Length between ITRs (bp)	Enhancer ID & Promoter	Product Class	Primary product	Other components
rAAV: Grik1_enhGad2-1-Hsp68-EGFP-WPRE3-BGHpA	AiP1146 (vAi30.0, T502-047)	3047	eAi12.0 (MGT_E31) Hsp68	Fluorophore	EGFP	WPRES BGHpA
rAAV: Grik1_enhGad2-1-pBGmin-EGFP-WPRE3-BGHpA	AiP1113 (vAi30.1, T502-053)	2883	eAi12.0 (MGT_E31) pBGmin	Fluorophore	EGFP	WPRES BGHpA
rAAV: Grik1_enhGad2-2-Hsp68-EGFP-WPRE3-BGHpA	AiP1147 (vAi31.0, T052-048)	3069	eAi13.0 (MGT_E65) Hsp68	Fluorophore	EGFP	WPRES BGHpA
rAAV: mscRE5-pBGmin-EGFP-WPRE3-BGHpA	AiP989 (vAi11.0, TG989)	2526	eAi4.0 (MGT_E5) pBGmin	Fluorophore	EGFP	WPRES BGHpA
rAAV: mscRE5-pBGmin-FlpO-WPRE3-BGHpA	AiP1013 (vAi12.0, TG1013)	3104	eAi4.0 (MGT_E5) pBGmin	Recombinase	FlpO	WPRES BGHpA
rAAV: mscRE8-pBGmin-FlpO-WPRE3-BGHpA	AiP1012 (vAi14.0, TG1012)	3083	eAi5.0 (MGT_E8) pBGmin	Recombinase	FlpO	WPRES BGHpA
rAAV: hsA2-eHGT_079h-minRho-SYFP2-WPRE3-BGHpA	CN1525 (vAi115.0)	2338	eHGT_079h (eAi115.0) minRho	Fluorophore	SYFP2	hsA2 WPRES BGHpA
rAAV: hsA2-eHGT_082h-minRho-SYFP2-WPRE3-BGHpA	CN1528 (vAi116.0)	2360	eHGT_082h (eAi116.0) minRho	Fluorophore	SYFP2	hsA2 WPRES BGHpA
rAAV: hsA2-eHGT_086h-minRho-SYFP2-WPRE3-BGHpA	CN1532 (vAi117.0)	2116	eHGT_086h (eAi117.0) minRho	Fluorophore	SYFP2	hsA2 WPRES BGHpA
rAAV: hsA2-eHGT_128h-minRho-SYFP2-WPRE3-BGHpA	CN1621 (vAi119.0)	2170	eHGT_128h (eAi119.0) minRho	Fluorophore	SYFP2	hsA2 WPRES BGHpA

FIG. 21A cont'd

Sequence	Sequence Name	Sequence Length between ITRs (bp)	Enhancer ID & Promoter	Product Class	Primary product	Other components
rAAV: hsA2-eHGT_140h-minRho-SYFP2-WPRE3-BGHpA	CN1633 (vAi120.0)	2011	eHGT_140h (eAi120.0) minRho	Fluorophore	SYFP2	hsA2 WPRES BGHpa
sCAAV: eHGT_023h-minBGlobin-SYFP2-WPRE3-BGHpA	CN1259 (vAi104.0)	1956	eHGT_023h (eAi104.0) minBGlobin	Fluorophore	SYFP2	WPRES BGHpa
rAAV: 3xSP10ins-eHGT_359h-minRho*-SYFP2-WPRE3-BGHpA	CN2045	2073	eHGT_359h minRho*	Fluorophore	SYFP2	3xSP10ins WPRES BGHpa
sCAAV-eHGT_019h-minBGlobin-SYFP2-WPRE3-BGHpA	CN1255 (vAi101.0)	1903	eHGT_019h (eAi101.0) minBGlobin	Fluorophore	SYFP2	WPRES BGHpa
rAAV: eHGT_064h-minBGlobin-SYFP2-WPRE3-BGHpA	CN1408 (vAi110.0)	1951	eHGT_064h (eAi110.0) minBGlobin	Fluorophore	SYFP2	WPRES BGHpa
sCAAV: eHGT_022h-minBGlobin-SYFP2-WPRE3-BGHpA	CN1258 (vAi103.0)	1743	eHGT_022h (eAi103.0) minBGlobin	Fluorophore	SYFP2	WPRES BGHpa
sCAAV: eHGT_022m-minBGlobin-SYFP2-WPRE3-BGHpA	CN1279 (vAi102.0)	1708	eHGT_022m (eAi102.0) minBGlobin	Fluorophore	SYFP2	WPRES BGHpa
sCAAV: eHGT_017h-minBGlobin-SYFP2-WPRE3-BGHpA	CN1253 (vAi100.0)	1751	eHGT_017h (eAi100.0) minBGlobin	Fluorophore	SYFP2	WPRES BGHpa
sCAAV: eHGT_017m-minBGlobin-SYFP2-WPRE3-BGHpA	CN1274	1635	eHGT_017m minBGlobin	Fluorophore	SYFP2	WPRES BGHpa

FIG. 21B

Vector Name (Plasmid ID)	Enhancer Name (Enhancer ID)	Origin Species	GABAergic	Lamp5	Vip	Sst	Pvalb	L4 IT	L5 IT	Method of Validation
AIp1146 (vAi30.0)	Grik1_enhGad2- 1 (eAi12.0)	M	S ⁺							TG
AIp1113 (vAi30.1)	Grik1_enhGad2- 1 (eAi12.0)	M	S ⁺							TG
AIp1147 (vAi31.0)	Grik1_enhGad2- 2 (eAi13.0)	M	S ⁺							TG
AIp989 (vAi11.0)	mscRE5 (eAi4.0)	M	S ⁺							T
AIp1013 (vAi12.0)	mscRE5 (eAi4.0)	M	S ⁺							T
AIp1012 (vAi14.0)	mscRE8 (eAi5.0)	M	S ⁺							T
CN1525 (vAi115.0)	eHGT_079h (eAi115.0)	H	S [~]				S [~]			T, R, I
CN1528 (vAi116.0)	eHGT_082h (eAi116.0)	H	S [~]				S [~]			T, R, I
CN1532 (vAi117.0)	eHGT_086h (eAi117.0)	H	S ⁺				S ⁺			T, R, I
CN1621 (vAi119.0)	eHGT_128h (eAi119.0)	H	S [~]				S [~]			T, R, I
CN1633 (vAi120.0)	eHGT_140h (eAi120.0)	H	S ⁺				S ⁺			T, R, I
CN1259 (vAi104.0)	eHGT_023h (eAi104.0)	H	S ⁺				S ⁺	S ⁺	S ⁺	T, I
CN2045	eHGT_359h	H	S ⁺				S ⁺			T
CN1255 (vAi101.0)	eHGT_019h (eAi101.0)	H	S ⁺	A ⁺						T, R, I
CN1408 (vAi110.0)	eHGT_064h (eAi110.0)	H	S ⁺			S ⁺	S ⁺			T, R, I
CN1258 (vAi103.0)	eHGT_022h (eAi103.0)	H	S [~]	A [~]	A [~]					T, R, I
CN1279 (vAi102.0)	eHGT_022m (eAi102.0)	M	S ⁺	A ⁺	A ⁺					T, R, I

FIG. 21B cont'd

Vector Name (Plasmid ID)	Enhancer Name (Enhancer ID)	Origin Species	GABAergic	Lamp5	Vip	Sst	Pvalb	L4 IT	L5 IT	Method of Validation
CN1253 (vAi100.0)	eHGT_017h (eAi100.0)	H	S*	A*	A*	A*				T, R, I
CN1274	eHGT_017m	M	S*	A*	A*	A*				T, R, I

FIG. 22

Enhancer eAi12.0 (MGT_E31, Grik1_enhGad2-1) nucleotide sequence 1 (923 bp):

CTAGGACACTCATTACACCTTTGGAGTTTTCTCACCCTCACATTTTCATCCAGAAAAAAA
AAAATCCTGCAGTGCCTAACATTTCTCAAGGTATTGGATCTCCCTGAGACGTGTCCTTTTCA
AGGTGACCTGATGCAAGACCTCGCTTGGTTGGCTTGCCAAATCTGGTGGCTTTGCCAAATC
CCAGTCCTGCAAAGCATTCTGTCTCTAACATTTCTGCTTGGAGTACAAAGAACTGAAAGT
TCATTTCCCAGAGGGTGGGGGGAACTGGAGCTCTGTCTAAACTGGATTTTATTGGGTGA
CGCTTCAAAGCACTCACTTAGCATTTCTGCTAACAGAGCTGAACGTTTACGCCAAAATATTG
GAAAAGGAAATGCTCAGGCAGCATGAAAAGACAGAGGAGGCTGGCAAAGCAGGGAAGCC
AAGCCAAGGGTGACAAAATCGAAGGGGAATGTTCCAGAAAAAGGCATGAAAAGGAACTG
TTTCAAGCACAGCGGTTAGAGCCTTATCCCACAGACACATCCACATCTTGAAAATAAGTAG
CGTTGGAGTTAAGAGTGAGTGTTATTCTCTGATATACTCACTCACAGTCTTATGCCAGTGCA
TCACAGATTTTGGCAGCCCTTCTGAACATGATAGGCTCATCCCACAACATGGCAGTTAGCA
ATTGATGGCAGGAATCTGATGAGCTTGGACCTCGAAACCCACGGTGGGGATCCTGTGAAT
ATTTTATCCCTCCAGTTCAGGGGGCTAGAAATAAAACAGCTTTAGTGTCCACACAGCGAAC
CACTATATATTCTGAGAAGATATTCTTATTGAGCATAGTTGTCATCTGAAACATAAGTCTGG
GTGTATTTTAGGTTCTACGACGTGGCTTTGAGAATGTACTTATCTGTCTCCCACCAGGTTTC
CACGTG (SEQ ID NO: 1)

Enhancer eAi12.0 (MGT_E31, Grik1_enhGad2-1) nucleotide sequence 2 (930 bp):

ACGCGTACACTCATTACACCTTTGGAGTTTTCTCACCCTCACATTTTCATCCAGAAAAAAA
AAAAATCCTGCAGTGCCTAACATTTCTCAAGGTATTGGATCTCCCTGAGACGTGTCCTTTCA
CAGGTGACCTGATGCAAGACCTCGCTTGGTTGGCTTGCCAAATCTGGTGGCTTTGCCAAAT
CCCAGTCCTGCAAAGCATTCTGTCTCTAACATTTCTGCTTGGAGTACAAAGAACTGAAAG
TTCATTTCCCAGAGGGTGGGGGGAACTGGAGCTCTGTCTAAACTGGATTTTATTGGGTG
ACGCTTCAAAGCACTCACTTAGCATTTCTGCTAACAGAGCTGAACGTTTACGCCAAAATATTG
GAAAAGGAAATGCTCAGGCAGCATGAAAAGACAGAGGAGGCTGGCAAAGCAGGGAAGCC
AAGCCAAGGGTGACAAAATCGAAGGGGAATGTTCCAGAAAAAGGCATGAAAAGGAACTG
TTTCAAGCACAGCGGTTAGAGCCTTATCCCACAGACACATCCACATCTTGAAAATAAGTAG
CGTTGGAGTTAAGAGTGAGTGTTATTCTCTGATATACTCACTCACAGTCTTATGCCAGTGCA
TCACAGATTTTGGCAGCCCTTCTGAACATGATAGGCTCATCCCACAACATGGCAGTTAGCA
ATTGATGGCAGGAATCTGATGAGCTTGGACCTCGAAACCCACGGTGGGGATCCTGTGAAT
ATTTTATCCCTCCAGTTCAGGGGGCTAGAAATAAAACAGCTTTAGTGTCCACACAGCGAAC
CACTATATATTCTGAGAAGATATTCTTATTGAGCATAGTTGTCATCTGAAACATAAGTCTGG
GTGTATTTTAGGTTCTACGACGTGGCTTTGAGAATGTACTTATCTGTCTCCCACCAGGTTTC
CTTAAGGAGCTC (SEQ ID NO: 42)

Enhancer eAi13.0 (MGT_E65, Grik1_enhGad2-2) nucleotide sequence (945 bp):

CTAGGAGAGGAGGCAATATTTGGGATGTTAATAAATAAAGTAATTAATTAATAAAAAATAGAT
ACTATAAAACAACTACAAGCAAACAAACCTAGAACTACAAAGGCCTAAGGGTGTGCTTTG
GAAGAGCTAACCAAAGAAAGTAACTTTCAAAGGTTTTGAAAGTAGAAGGCAAGGCCTTTG
AGACGCCAGGGGGTAATGCGAGGAGGTAATATGAATTCTTGGAACCAGCATGTTTATTAGA
GTTGTATCAAATCAGGGAGTGAGAAGAGATGTTTCTATTTATGTGGAGGGAGGAATAATT
GGAAGAGAGATTTTAAACCTGTAATTAGGAACCTTGAATTGATACAAAGCAGAATGGGAGA
ATCGATTGATGTGAATTAACAACAGCAGTGATGAGTCCCACTGCTGGGCGGGGATGGATT
GGCAGGCCTGGGCCTTTTGGCTGTCTACAAGGCAGGCAGAGTGGTACCAGCCAGATCAGT

FIG. 22 cont'd

TGCCAGCAGTGGTATACTGTGCCAAGTCAAAGAAATGTGACTGAGATAGGATCCACCAATC
TGAGCAATGGCTCCAGCCTCCTCTCCAGCCTTCCAAGAAACCCCATGAGTAATCTCCTAAA
ACCCTTTTGTCTTTCTTCCGTCTCTATTCTTGAAACATCGTTTTTTCCAGTAGAACTGCATGT
GATATATTTGGACATGAAGTAGTGATTGTTTTCTCACTAAATCATCTAGAGCCAAATGAAAG
CGGGTGTTATTGAAAAGAAATAAATTATTTTAACATCAAATAAATTATATAAATATATAGTTTG
AGGGGAGAAACATTAAAGATAATCCTGCCATTTATCACACTATGAAACTGGACTTAGAATGA
TTCTGTTTCATCTGCATTTTATTCGGCATAAGTATGAATCCTTTCTATATATAATGTGGCTAAC
TTCAGGAAAATCACCACGTG (SEQ ID NO: 2)

Enhancer eAi4.0 (MGT_E5, mscRE5) nucleotide sequence (501 bp):

AGTTTTATTCTTTTCCTCATTTGTATCCAGCTTCAGTTTGAAAGCATGGATTGCATGCTTTCT
GTGGATGGCAGAAGATGCAACTCTGTGACAGTGTGCTCCTTATTAATTCCAACATTTCCAC
ATTGCTGTGCTGAGTCGTACACATGGGCTTGGCATATGAGCTGTTTACACATTGAAACCAG
TTCTGAAAACGTAGCAAAGTGACAGGTGCACATCCAGAAGGATATGCAGATTAATTTTTTTTG
TGCAAATACCTTTGTGTTTTCTCCACAAATTCCTCCCTAACGTCGTACATCAAGCTACACCA
CCCAGTTAATTTGCTATTATGCCTTGCCTCAGAAATCCTCACCTTTTTTAGTTGCTTGATTGAT
GTGAACTCACACTATTTGGAAACAGCACAAATACACCATTTGTGCTTTTCCAGGTCTTTAGC
TCTTCCGTTTGCACTTTTAAAGATTTTTCTTACATAAAAGTCAGGATATTCAAATATTCATCC
AAGT (SEQ ID NO: 3)

Enhancer eAi5.0 (MGT_E8, mscRE8) nucleotide sequence (501 bp):

AGACACCTGGAGATTCAAGCTTACATGCAACAACAAAGGAAAGTACTTAAAATTCTAGCAA
GTTAACAAATAGCCATTTAATCCCTTGCTAGCATAATTTCTTATTACATCTAAGATGCCAAT
TTCAATCATATAAATATAAGAAATAAAGAGATTTCTATTCTCAAGGCATCATAACTCCTTCTC
TAGACCTTTGGTGATGTAAATGAAGTCTCTGAGACTGTCATACTGAGGAATGATGTTACAC
TCCATTGTGTGAAATCTGGGCTCTACATTTCTGATATTTATGATGCAAATGCTAGCAGGCAA
AGTGCAAGTGAAATCAATATGATTTTATTTAGTATTCAGGCAACAGGCTCCATCTGTACAAT
ATTCTCCAAGATGATATTTAACAGAACTATAGAACTAATGACTGCTCAGATATAAATGACA
CTTGATTTTAACTGAAAAGCAATCTTACAACACCAAAGAAGTGCCAGTCATTTTCATTCTAAA
GAA (SEQ ID NO: 4)

Enhancer eHGT_079h nucleotide sequence (602 bp):

AGGTGAATAAGATACCCAGGATATGAAATTTAAAGAGGCATTCACTTTCAGAGTCCTGCAAA
TGCACTTACATAAGTGAGTGCCTCTTTTAATTTTGTACTCTAATTGACTCACTTCATCTCATT
CTCGTCCACTGCTAGCTTTGAAGGCAGAGAAAACGAGCCATGAGCCAAAGATTGTGGGTG
GCCTCCGAAAGCTATATATCCTGAGCTGACAGCCAGCAAGGAAATGAGGACCTCAGTAGA
ATAACCACATGGAAGTGAATTCTACCAATAGCAAGAGTGATATTGGAAGTGGATTCATTCTT
AGAACTCCCAGAAAGGAATACAGCCTGCCTGGCACTTCTATTTCAATACTGTGGCACTCTA
AGCAGAGGACTCCACTGAACCACACCGTACCCATAAACAGGAAGTCTTTTATGCAGCTGGT
ATTTATGGGAATTGGAAATTTTGACCTACACAGCTGTGAGATAATAAATGGGTGCTGTTTTA
AGCCACTAAAATTTTTTTTATTTTTTAAATTTATTTTTAATTTCCATAGGTTTTTGGGGAACAGG
GGTATTTGGTTACATGAGTAAGTTCTTTAGTGGTGATTGGTGAGATT (SEQ ID NO: 5)

FIG. 22 cont'd

Enhancer eHGT_082h nucleotide sequence (699 bp):

CCTTTGGCCGGCAGGAGCAGCCTGATTTTAGTGA CTCTCAGAAAGGGAGGTGCAGCTGG
AAGGTTGTGTTTATTAAAGCCCAAGTGACTGGAGGTAAATGCAGGTGTGGTCAGGAAATGT
AGCTGGTTCACCTCTCAGTAGGCACAAGGCCAAAGTCAGCCGACTTGCAGATTGTAATGC
CGGTTCTCCACCCAGAGGTTATTTGAGTTCATCCAGCACTGCAGCTGAGCAGAATGCTGC
AGTCAGGGTGACACAATGCAAACCTGGAGCCATGTCCTCAAGTGCATGCCAATCAAACCAAC
TTTAACCTTCGCTGGTTTGTGGCTGCTGACCTAGAGAGGAATTTGTTATATGGCTGAGGTC
CTGTCCTACCTAAGCCAGACTCCAGTGCAAAAAAAAAATCCTGTCAGATGCGCCTCTTCTT
CTCTGTCTCCATCTCCCTCTCTGCCTCCCGTTTCCCCTTTACTCTTTACTGTTAGTTGGCCT
GGGGTTCCTGGGCTGAAGTGGAGTTTCCTCCTGTCTCCGTGTGTTTGTCTGCTCTGGTTTGC
TCTTCTTCCAGGGCCCCCTTGGTAAAAGAAGAGCTGCTGACAGGGTAGAGTGGAGGTGGGG
AAGTGACCACCAGACTGGGAGTGATGTAGGTTGTGGTCATCAATGATGTAGGTTGCTGTCA
CCAGGACACTCAAATCAGCTGTGACCAGAG (SEQ ID NO: 6)

Enhancer eHGT_086h nucleotide sequence (355 bp):

TCCCCTGTCTCCTTGGCCGAGCCATCCTCTGTCACCTTCACTGTTTCCACTAGGTTATTGA
GGCCTTTTGGTCATGAGCTCCTGCCACCAAATCAGTGCCTGAGGCTGCTGCATAACAATCT
CCTTCCTTGATTAGAGCATCTTAATTATGCTCCAGGAATTTGCTTAAACAGATTAATCATTCT
GGAGGACAGATTTAGAAGCTGTGCTAGCTGAGGGATCTGCTCTGAGTGTCCCCGGGGAGA
TTCTGGTGACAACCGCCTGTTGGGCCCTTGGCAAGGCCTTGGGTGCCTCGGAGACTCGGA
CACTGGCTGGAAGCTCTACGCACCCCCGAAGGCTTCTCTTTGCTCCTCTG (SEQ ID NO: 7)

Enhancer eHGT_128h nucleotide sequence (428 bp):

CTAGTATCATGAGGAGGATCTCAAAGCAGGGAAGGGAAGCCAGCTGAGGGACCCCTGAG
AGGAGATCAGGGGATAAATACTTTGCAACAGAATCCCCATGGGAAACATGACAAAATTAGT
AACAGGAATGGAGCCTCCCTCAGAGGATACATGAATCGTTGAGTTGAAAATAACCACAAGG
TCATCAAATCCAGCTGCCACCAACTGCCTGAATATTTTCAACAAATTTTGAACAAGCAGTC
CTCCAGCCTCCTCCAGCTGCATCCAGCAGCTCAGTGCTCCACCTGACAGCCTCTTCTGATA
GCGTGAAGGCTCAGAAAGCCCCTGTGGGGCCCAGCACATCCTGGGGCAGGCCTCTAGTG
GGCAGAGGGCAGCATACTGGGGTCTGGATCTAACACACTCCTCCATCCAGGTGGTATATT
TATAG (SEQ ID NO: 8)

Enhancer eHGT_140h nucleotide sequence (304 bp):

TGTGTGTGAAGATTAAGCCCTTTCTTGCTCTATGTATTTGAGCCTAGAAAAAGACACACTCC
CAAGAAATTAGAATAAAGGGAAAAATGACTTTGAGATGTACTTTCTTAAATAACTGAAGTA
ATGACCGCATGAGCCAGCTGATGGGTTTTTAAATGGATACGTTTCTATCAGCCTGCTGCTT
GATGCCAGAGCCAAATATATATGGAGTGTATTATATCATATCCTGTAGGCAGGAGACTGTT
GGGGCATTGGGGGACTTAGAGAGGTGAAGTGGCAGGCTTGGCACAGGAATTAACAGCA
(SEQ ID NO: 9)

Enhancer eHGT_023h nucleotide sequence (629 bp):

CCCTAAGCTGAGCTGAGACGTTTTAGTTATAAATGGCTTCTGTGCTTAGCAATCTGCTCTTT
TTATTCCCGTGTGGACTTTCCCTAGCTCTGGCCTTATGCTGCACTAGAAAAGATTTAGCAAG
GGGAGAGGAAGCAGCCTTCCTTACATAACTGGCCTCTTGTGAAAGGGAGCAGCTGCTTGG
TGAAAAAAGACATTCCCCTCCATGCATCCCTCTCCTTCTGCCTCTGGGGGTTGCAGCTTGA

FIG. 22 cont'd

GTCAGAACCCAGGACCATTTAACTCCAACCTTTGAGGAAGAGACGCACCCTGGCCCCAGCC
ACGCCTGTTAGAATCTTCCTAGCTGAGTGACACAGTGACACTCAGCCTCAGTTTCTCTGTA
AACAAAATGAAGATAAGAGAGCCGACGAGGATGAAATGGAATAACACACCCGTGCGGTGCC
TGGTACAGAGTGAGCCCCAGAACTGTTGACGCGGCCTCCTTGTGGCTGTCTGGCTTGACC
CGGAGTGACTCTGCCTCCCAGTTCTCCGGGATGGGAAGGTGATCCCTGTTTGAGATACCA
ATTTATAAGAAACCGAGCCCGGGAGCTATTTAGAGGTGAGGTGATAAACCAGGAGGCCGG
CTCCTTCATCCCGGTCATCACGA (SEQ ID NO: 10)

Enhancer eHGT_359h nucleotide sequence (485 bp):

TCACAGCTCCTTGGTCTGAACTGGTACCCAGGATCAAGATGTCCCTCCCATGATTCTAATC
TCCAGCCCCCTTACACCATAGTCACTTCCAACTCTGGATTTCTTAACATTGCAGTTTGCTT
CCAAAAAATTAATAAAAAAACCTCATCTGGGCATTAGAAGTCATTCTGAAGAGGCTGTC
CTATATCTCAGTGCAATTTCTACTTATCCACTCCCGACTATACCCTGCTTTCATCAGATTG
CAACTGACCTAATTTGTAAAGACTTCAGTGAAAAGAGCCTCATTAAAAGCCTTCATCTCTTT
TGCAGCCTCTCTTCTCCCTCAAATGATCCAGGTGAGAATGGCCCAATAAAAATAGCTGG
GAAAATTACAAAGTCACTCAAGACTCTGCCAGGAGTAGAGATTTTAAGAATTAGTGACAGA
CATATACTGTCTAGGAGTAGGGTGGGAGTGGGTTGCATGGCTTAGGGACATAGA (SEQ ID
NO: 11)

Enhancer eHGT_019h nucleotide sequence (576 bp):

TCCATCCACAGTCTGTCTCTTCCCCTTGTACAGATTGGCTGCCCTCCCCCACTTGAAATC
ACTATTAACTGCAGAGATAACTGCCCGTCTTAGGGCTGCAGTCCTCCACATAACCTGAA
GCCACTGCCATTTGTCAACAACAGATCATTTCTTGAAGTCTGACAGCCAGCCTTGGAAGCAG
GAAGCATGCTCATTTTCCTCACCTGCCGGAGCTTCAGAAAGAACTCAGCTCAAAATAGAC
GCTATTGAAGCAGGAGACACATGAGGAGGTACACAGCCTATCAATACCTCAACACTGGCCT
CTCCTCGCCCTCCAATCCTATTGACTGCATGAGAGCCACTGCTGAAGGCTTTGTAATCCAA
AACCATTAAGTATTCATGGGGCTGGCAACCACAAATAGAAAATGAATTGGATTGCTCCCTTT
CAAATTAGTATAATGACATTATCTGGGAAATGATGATTTAAAAAATAGTTCAAAACATGGGTT
TGAATGTTCTTCTACTGTTCTGCCAGGCTCCTCTCTGTTTTTTACAATAGAGAAGAGCCC
TCCTGGTGCTGTGGAGACATTCC (SEQ ID NO: 12)

Enhancer eHGT_064h nucleotide sequence (567 bp):

TTTGGCCTAGCTGCTTTGACTAACCCCTCTTCTATTTTCAGTTATGCGGCAAGTTGCATACT
CAGGTGCCCTTCTGACTACTTGAATACTTTCCCTGTGATGTAAGAAGTGTTTTCAATTGGT
AAAGTTGTGGTATATAATTACAATTGAAGTCTCTTGTACTTGCCTCTTTTACAAAAATTCTCT
CCTAGCAGAACGTAGTGTGAGTCATCTACACAGCTGTTTTTCTGATTATTGGAATTTTCTTTT
GACATGAAGGAAGTATCTCATTGACAGAACTGCGTTGTGAAGGAGTGCTAACTGTAGCATA
AAATACAAAATTGGATTTTTAGATTGCAAAATACAGTAAAGCTTTGAAAAGTATTTGGCATGA
CATTTAACTCAATACATTTTGCCTAAAAAATATTAGCCAAGAACCCTATCAACTTGTTTTTG
AATAAACTTCTGTATGGACCTTAAATTCATGCTGAGTTTGACCGCATTTTCTTGCACTGGT
AGCATTTTCCCTCTGAGTCATCCTCATTTCCCTTCTACTTTCTCACATGACTAGGTAAAGATAA
CTCAT (SEQ ID NO: 13)

FIG. 22 cont'd

Enhancer eHGT_022h nucleotide sequence (416 bp):

AGTCCTGTGGATACCATCTGAGCCTGTGCCCTTTTCTCTCTACTCTTGAGCAAGTGCCCAT
CAGGAGCCACATTTGAGTCAGGACCTGCCAAGAGCATTTTCATCTCATCCTCAGTTCCCT
TTTGTCTGGAACCACTCGCTATGAGTTTCTATGAGATCGTAGCCAGGTTTGACCAAACCAC
TCCTGCACACATGTTGTTTTGAAAGATTAGCACAGAATCCAACACAACATTTCCCTTGATT
CATCTTTCAAGCCAACAGAGAAAGAGCTATTTCCAAAAAATAACACATTATTAACCAATGA
ATAAGAGCCAAGAAAACAATGTAGCCAGCATATTTATTCCAAACAGCTTGTTTGAATGCAA
ATTGCAAATACATCTATTTCCACCCGGCTCTTTGGGGAACCATGCTG (SEQ ID NO: 14)

Enhancer eHGT_022m nucleotide sequence (381 bp):

GTGTCCTCCAGAGTCCGTGGGTCCTGTGATCCCTTGCTCCATTCAATGTGCTCATCAGAAG
CCATATTTGAAGCAGCATCTACCATCCATGGTCCACCTTCTGTCTGGATCCACCCATTTAG
GAAACCATATCCTTGTTTGACCAGATCACTCCTGCACACGTGTTGCTTTGAATTCTAGCACA
GGATTGAGCATGACATTTCCCTTGATTCTCTTCTGAGCCAACAGAGTGAGGGCTATTTTC
CCAAAAAACAACACATTATTAACAAATGAATAAGAGCCAAGAAAACAATACAGCCAGCATA
TTAGTCCACGCAAGCTTGTTTAGAATACAAATCGCAAATGTATTTAATTCACAAGGCTCTT
GGGAAACTTTC (SEQ ID NO: 15)

Enhancer eHGT_017h nucleotide sequence (424 bp):

GGTAACAGCCTGAGGGCTGCAGGTACACCCACAGGGCCAAGGGGTGGGCAGCATGGAG
GTGCAGGGTGGCAGGGACCCTGGGCGGGGCGGCACAGCTGTGCGGGAGGCTGGGCTGC
TGGCATCAGCAGGCGCCCCTCCTCCCCACCTCGCTAAACAATCATTGCACAAAATATGCAA
ATGGTATAATTACTGTTATTTGTTTTGCTGATATAAGTGTTTGAAATGCAAATGTCAAGTTTG
GGCGCCTTCATTTTTCCAACCCTCTCACCCGGACATTTGCAAGTTGATGAGTTGTTCTTCAT
CCTGGAAGGAGGAGGAGGAGCTCCCCCAACCGCCAGGGTGCCAGGGGAGTGAGTCCA
GCGTGGCAGCCGCCACTGCCTGCCCGAGGGCACTGCTGGGCCCCCTTCCGACGGCACA
CAGT (SEQ ID NO: 16)

Enhancer eHGT_017m nucleotide sequence (308 bp):

TGTGTGTGGTCTTAGGGACAGGGGGTGGCTGGCAGGCAGGCAGTGGCGGGAGGCACAG
CTGTGCAGGAGGACGGACTGCTGGCATCAGCAAGCGCCCCTCCTCCCCCCCACACTAAAC
AATCACGGCACAATATGCAAATGGTATAATTACTGTTATTTGTTTTGCTGATATAAGTGTTT
GAAATGCAAATGTCAACTCTGGGCGCTGTTGGTTTTTCCACCCCTCTCTCCCGGACATTT
GCAAGTCGATCAGTGAGCCCTCATCCTGGGAACACTAGGGGTGCCTCCTTCTGACAGACT
TGGAGC (SEQ ID NO: 17)

hsA2:

TTAGAACAATGGCTGGCCCATAGTAAATGCCGTGTTAGTGTGTTAGTTGCTGTTCTTCCAC
GTCAGAAGAGGCACAGACAAATTACCACCAGGTGGCGCTCAGAGTCTGCGGAGGCATCAC
AACAGCCCTGAATTTGAATCCTGCTCTGCCACTGCCTAGTTGAGACCTTTTACTACCTGACT
AGCT (SEQ ID NO: 18)

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

GGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTG (SEQ ID
NO: 19)

FIG. 22 cont'd

minCMV Promoter:

GAGGTAGGCGTGTACGGTGGGAGGCCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGG (SEQ ID NO: 20)

Mutated minCMV Promoter (Sacl RE site removed):

GAGGTAGGCGTGTACGGTGGGAGGCCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCGCCTGG (SEQ ID NO: 21)

minRho Promoter:

GATTCAGCCGGGAGCTTAGGGAGGGGAGGTCACTTCATAAGGGCCTGGGGGGGGAGTTGAGGCCACGAGTCGTCCAGCCGGAGCCCCGTGTGGCTGAGCTCCGGCCTCAGAAGCATCC (SEQ ID NO: 22)

minRho* Promoter:

GATTCAGCCGGGAGCTTAGGGAGGGGAGGTCACTTCATAAGGGCTTGGGGGGGGAGTTGAGGCCACGAGTCGTCCAGCCGGAGCCCCGTGTGGCTGTGCTCCGGCCTCAGAAGCATCC (SEQ ID NO: 23)

Hsp68 minimal Promoter (proHsp68):

CAGGAACATCCAACTGAGCAGCCGGGGTCCCCCCCCACCCCCACCCCCGCCCCACGCGGCAACTTTGAGCCTGTGCTGGGACAGAGCCTCTAGTTCCTAAATTAGTCCATGAGGTCAGAGGCAGCACTGCCATTGTAACGCGATTGGAGAGGATCACGTCACCGGACACGCCCCCAGGCATCTCCCTGGGTCTCCTAACTTGGCGGGGAGAAGTTTTAGCCCTTAAGTTTTAGCCTTTAACCCCATATTCAGAACTGTGCGAGTTGGCGAAACCCACAAATCACAACAACTGTACACAACACCGAGCTAGAGGTGATCTTTCTTGCTCATTCCACACAGGCCTTAGTAATGCGTCGCCATAGCAACAGTGTCACTAGTAGCACCAGCACTTCCCCACACCCTCCCCCTCAGGAATCCGTACTCTCCAGTGAACCCCAAGAACTCTGGAGAGTTCTGGACAAGGGCGGAACCCACAACCTCGATTACTCAAGGGAGGCGGGGAAGCTCCACCAGACGCGAACTGCTGGAAGATTCTTGCCCCAAGGCCTCCTCCGGCTCGCTGATTGGCCAGCGGAGAGTGGGCGGGGCGCGGTGAAGACTCCTTAAAGGCGCAGGGCGGCGAGCAGGTACCAGACGCTGACAGCTACTCAGAACCAAATCTGGTTCCATCCAGAGACAAGCGAAGACAAGAGAAGCAGAGCGAGCGGCGCGTTCCCGATCCTCGGCCAGGACCAGCCTTCCCCAGAGCATCCCTGCCGCGGAGCGCAACCTTCCCAGGAGCATCCCTGCCGCGGAGCGCAACTTTCCCCGGAGCATCCACGCCGCGGAGCGCAGCCTTCCAGAAGCAGAGCGCGGCGCC (SEQ ID NO: 24)

SYFP2:

ATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGA CGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGCCCCGCTACCCCGACCATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCCTTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAG

FIG. 22 cont'd

AAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGGCGTGCA
GCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCG
ACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGT
ACAAGTAA (SEQ ID NO: 25)

EGFP:

ATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTTCGAGCTGGA
CGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCT
ACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCA
CCCTCGTGACCACCCTGACCTACGGCGTGCAAGTTCAGCCGCTACCCCGACCATATGA
AGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCT
TCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACC
CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGG
GCACAAGCTGGAGTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAA
GAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCT
CGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACA
ACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACA
TGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACA
AGTAA (SEQ ID NO: 26)

Optimized Flp recombinase (FlpO):

ATGGCTCCTAAGAAGAAGAGGAAGGTGATGAGCCAGTTCGACATCCTGTGCAAGACCCCC
CCCAAGGTGCTGGTGCGGCAGTTCGTGGAGAGATTCGAGAGGCCAGCGGCGAGAAGAT
CGCCAGCTGTGCCGCCGAGCTGACCTACCTGTGCTGGATGATCACCCACAACGGCACCG
CCATCAAGAGGGGCCACCTTCATGAGCTACAACACCATCATCAGCAACAGCCTGAGCTTCGA
CATCGTGAACAAGAGCCTGCAGTTCAAGTACAAGACCCAGAAGGCCACCATCCTGGAGGC
CAGCCTGAAGAAGCTGATCCCCGCCTGGGAGTTCACCATCATCCCTTACAACGGCCAGAA
GCACCAGAGCGACATCACCGACATCGTGTCCAGCCTGCAGCTGCAGTTCGAGAGCAGCG
AGGAGGGCCGACAAGGGCAACAGCCACAGCAAGAAGATGCTGAAGGCCCTGCTGTCCGAG
GGCGAGAGCATCTGGGAGATCACCGAGAAGATCCTGAACAGCTTCGAGTACACCAGCAGG
TTCACCAAGACCAAGACCCTGTACCAGTTCCTGTTCTGACCACATTCATCAACTGCGGCA
GGTTCAGCGACATCAAGAACGTGGACCCCAAGAGCTTCAAGCTGGTGCAGAACAAGTACC
TGGGCGTGATCATTAGTGCCTGGTGACCGAGACCAAGACAAGCGTGTCCAGGCACATCT
ACTTTTTAGCGCCAGAGGCAGGATCGACCCCTGGTGTACCTGGACGAGTTCCTGAGGA
ACAGCGAGCCCCGTGCTGAAGAGAGTGAACAGGACCGGCAACAGCAGCAGCAACAAGCAG
GAGTACCAGCTGCTGAAGGACAACCTGGTGCGCAGCTACAACAAGGCCCTGAAGAAGAAC
GCCCCCTACCCCATCTTCGCTATCAAGAACGGCCCTAAGAGCCACATCGGCAGGCACCTG
ATGACCAGCTTTCTGAGCATGAAGGGCCTGACCGAGCTGACAAACGTGGTGGGCAACTGG
AGCGACAAGAGGGCCTCCGCCGTGGCCAGGACCACCTACACCCACCAGATCACCGCCAT
CCCCGACCACTACTTCGCCCTGGTGTCCAGGTACTACGCCTACGACCCCATCAGCAAGGA
GATGATCGCCCTGAAGGACGAGACCAACCCCATCGAGGAGTGGCAGCACATCGAGCAGC
TGAAGGGCAGCGCCGAGGGCAGCATCAGATACCCCGCCTGGAACGGCATCATCAGCCAG
GAGGTGCTGGACTACCTGAGCAGCTACATCAACAGGCGGATCTGA (SEQ ID NO: 27)

FIG. 22 cont'd

Improved Cre recombinase (iCre):

ATGGTGCCCAAGAAGAAGAGGAAAGTCTCCAACCTGCTGACTGTGCACCAAAACCTGCCT
GCCCTCCCTGTGGATGCCACCTCTGATGAAGTCAGGAAGAACCTGATGGACATGTTTCAGG
GACAGGCAGGCCTTCTCTGAACACACCTGGAAGATGCTCCTGTCTGTGTGCAGATCCTGG
GCTGCCTGGTGCAAGCTGAACAACAGGAAATGGTTCCCTGCTGAACCTGAGGATGTGAGG
GACTACCTCCTGTACCTGCAAGCCAGAGGCCTGGCTGTGAAGACCATCCAACAGCACCTG
GGCCAGCTCAACATGCTGCACAGGAGATCTGGCCTGCCTCGCCCTTCTGACTCCAATGCT
GTGTCCCTGGTGATGAGGAGAATCAGAAAGGAGAATGTGGATGCTGGGGAGAGAGCCAA
GCAGGCCCTGGCCTTTGAACGCACTGACTTTGACCAAGTCAGATCCCTGATGGAGAACTC
TGACAGATGCCAGGACATCAGGAACCTGGCCTTCCTGGGCATTGCCTACAACACCCTGCT
GCGCATTGCCGAAATTGCCAGAATCAGAGTGAAGGACATCTCCCGCACCGATGGTGGGAG
AATGCTGATCCACATTGGCAGGACCAAGACCCTGGTGTCCACAGCTGGTGTGGAGAAGGC
CCTGTCCCTGGGGGTTACCAAGCTGGTGGAGAGATGGATCTCTGTGTCTGGTGTGGCTGA
TGACCCCAACAACACTACCTGTTCTGCCGGGTCAGAAAGAATGGTGTGGCTGCCCTTCTGC
CACCTCCCAACTGTCCACCCGGGCCCTGGAAGGGATCTTTGAGGCCACCCACCGCCTGAT
CTATGGTGCCAAGGATGACTCTGGGCAGAGATACCTGGCCTGGTCTGGCCACTCTGCCAG
AGTGGGTGCTGCCAGGGACATGGCCAGGGCTGGTGTGTCCATCCCTGAAATCATGCAGG
CTGGTGGCTGGACCAATGTGAACATTGTGATGAACTACATCAGAAACCTGGACTCTGAGAC
TGGGGCCATGGTGAGGCTGCTCGAGGATGGGGACTAA (SEQ ID NO: 28)

SP10 insulator (SP10ins):

GAAGCTACCCCTAACACACTATTCTACACACAGAAAATGCTCTTCACTAG (SEQ ID NO: 29)

3xSP10ins:

GAAGCTACCCCTAACACACTATTCTACACACAGAAAATGCTCTTCACTAGGAAGCTACCCC
TAACACACTATTCTACACACAGAAAATGCTCTTCACTAGGAAGCTACCCCTAACACACTATT
CTACACACAGAAAATGCTCTTCACTAG (SEQ ID NO: 30)

WPRE3:

ATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTC
CTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATG
GCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACATCGC
CGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGG
(SEQ ID NO: 31)

BGHpA:

CGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGAC
CCTGGAAGGTGCCACTCCCACTGTCTTTCTAATAAAATGAGGAAATTGCATCGCATTGT
CTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGA
TTGGGAAGACAATAGCAGGCATG (SEQ ID NO: 32)

P2A:

GGCAGCGGCGCCACCAACTTCAGCCTGCTGAAGCAGGCCGGCGACGTGGAGGAGAACCC
CGGCCCCGGAGCTAGCGGA (SEQ ID NO: 33)

FIG. 22 cont'd

T2A:

(GSG)EGRGSLTCDGVEENPGP (SEQ ID NO: 34)

E2A:

(GSG)QCTNYALLKLAGDVESNPGRP (SEQ ID NO: 35)

F2A:

(GSG)VKQTLNFDLLKLAGDVESNPGRP (SEQ ID NO: 36)

Exemplary Plasmid Backbone 1 – Left ITR:

GGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTGCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCTTGAGGGGTGGAGTCGTGA (SEQ ID NO: 124)

Exemplary Plasmid Backbone 1 – Right ITR:

CCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTGCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGA (SEQ ID NO: 125)

Exemplary Plasmid Backbone 2 – Left ITR:

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGCCAACTCCATCACTAGGGGTTCT (SEQ ID NO: 126)

Exemplary Plasmid Backbone 2 – Right ITR:

AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTGCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG (SEQ ID NO: 127)

PHP.eB capsid:

MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLPGNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTEVPDPQPIGEPPAAPSGVGSMTASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDREVTTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDNRFHCHFSRDPWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYLSKTINGSGQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGVATNHQSDGTLAVPFKAQAQTGWVQNQGILPGMVWQDRDYYLQGPWAKIPHDTGNFHPSPLMGGFGMKHPPQILIKNTVPADPPTAFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL (SEQ ID NO: 37)

FIG. 22 cont'd

AAV9 VP1 capsid protein:

MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLPGNGLDK
 GEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKK
 RLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTEVPDP
 QPIGEPPAAPSGVGSGLTMASSGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRTVTTSTRT
 WALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWG
 FRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADV
 FMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNMFQFSYEFENVPFHSSYAHSQSLD
 RLMNPLIDQYLYLSKTINGSGQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQN
 NNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADK
 VMITNEEEIKTTNPVATESYQVATNHQSAQAQAQTGWVQNQGILPGMVWQDRDVYLQGPW
 AKIPHTDGNFHPSPMLGGFGMKHPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQVSVEIE
 WELQKENS KRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL (SEQ ID NO:
 38)

tet-Transactivator version 2 (tTA2):

ATGTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATGAAGTCG
 GTATCGAAGGCCTGACGACAAGGAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCC
 TGTACTGGCACGTGAAGAACAAGCGGGCCCTGCTCGATGCCCTGGCAATCGAGATGCTGG
 ACAGGCATCATACCCACTTCTGCCCCCTGGAAGGCGAGTCATGGCAAGACTTTCTGCGGA
 ACAACGCCAAGTCATTCCGCTGTGCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCATCT
 CGGCACCCGCCAACAGAGAAACAGTACGAAACCCTGGAAAATCAGCTCGCGTTTCTGTG
 TCAGCAAGGCTTCTCCCTGGAGAACGCACTGTACGCTCTGTCCGCCGTGGGCCACTTTAC
 ACTGGGCTGCGTATTGGAGGATCAGGAGCATCAAGTAGCAAAAGAGGAAAGAGAGACACC
 TACCACCGATTCTATGCCCCCACTTCTGAGACAAGCAATTGAGCTGTTGACCATCAGGGA
 GCCGAACCTGCCTTCTTTTCGGCCTGGAACATAATCATATGTGGCCTGGAGAAACAGCTAA
 AGTGCGAAAGCGGCGGGCCGGCCGACGCCCTTGACGATTTTGACTTAGACATGCTCCCAG
 CCGATGCCCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTGA
 CCTTGACATGCTCCCCGGGTAA (SEQ ID NO: 39)

Plasmid backbone 1 (including ITRs; R-ITRΔtrs is Bold (start of sequence). L-ITR is underlined (end of sequence); construct sequences are inserted into Plasmid backbone 1 between the L-ITR and the R-ITR):

CCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGA
CGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGACAG
 ATCCGGGCCCCGCATGCGTCGACAATCACTGGCCGTCGTTTTACAACGTCGTGACTGGGA
 AAACCCTGGCGTTACCCAACTTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGT
 AATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAA
 TGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATATGGT
 GCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAA
 CACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTG
 TGACCGTCTCCGGGAGCTGCATGTGTGAGAGGTTTTACCGTCATACCGAAACGCGCGA
 GACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGTTTTCT
 TAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCT

FIG. 22 cont'd

AAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATT
GAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTTCGGCA
TTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCA
GTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAG
TTTTCGCCCCGAAGAACGTTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCG
GTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGA
ATGACTTGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAG
AGAATTATGCAGTGCTGCCATAACCATGAGTGATAAAGTGGCGGCAACTTACTTCTGACA
ACGATCGGAGGACCGAAGGAGCTAACCCTTTTTTGCACAACATGGGGGATCATGTAAGT
CGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACC
ACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAAGTGGCGAACTACTTACTC
TAGCTTCCCGGCAACAATTAAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCT
GCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGG
GTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTAT
CTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGG
TGCCTCACTGATTAAGCATTGGTAAGTGTGACACCAAGTTTACTCATATATACTTTAGATTG
ATTTAAAACTTCATTTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATGA
CCAAAATCCCTTAACGTGAGTTTTTCTGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAA
AGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCAC
CGCTACCAGCGGTGGTTTGTGGCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAC
TGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCAC
CACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAAGTGG
CTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTGGACTCAAGACGATAGTTACCGG
ATAAGGCGCAGCGGTGCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGA
ACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCC
GAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCCGAACAGGAGAGCGCA
CGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGGTTTCGCCACC
TCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACG
CCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTT
CCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCG
CTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCG
CCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGCACGA
CAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCAC
TCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTG
AGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCTCTCGAGAT
CTAGAAAGCTTCCCGGGGGGATCTGGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTG
AGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAG
CGAGCGAGCGCGCAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCTGGAGGGGTG
GAGTCGTGA (SEQ ID NO: 40)

FIG. 22 cont'd

Plasmid backbone 2 (including ITRs; R-ITR is Bold (start of sequence). L-ITR is underlined (end of sequence); construct sequences are inserted into Plasmid backbone 2 between the L-ITR and the R-ITR):

AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAG
GCCGGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCG
AGCGAGCGCGCAGCTGCCTGCAGGGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGT
GCGGTATTTTACACCCGCATACGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATT
AAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAG
CGCCCCGCTCCTTTTCGCTTTCTTCCCTTCCCTTTCTCGCCACGTTTCGCCGGCTTTCCCCGTCA
AGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCC
AAAAAACTTGATTTGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTTC
GCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAACTGGAACAAC
ACTCAACCCTATCTCGGGCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATT
GGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTA
CAATTTTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCG
ACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGCTGCTCCCGGCATCCGCTTA
CAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGTCAGAGGTTTTACCCGTCATCACC
GAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATA
ATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTT
GTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGC
TTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCGCCCTTATTCCC
TTTTTTGCGGCATTTTGCCTTCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGA
TGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAA
GATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGC
TATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATAC
ACTATTCTCAGAATGACTTGTTGAGTACTACCAAGTCACAGAAAAGCATCTTACGGATGG
CATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAAC
TACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGACAACATGGGG
GATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGAC
GAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGC
GAACTACTTACTCTAGCTTCCCGGCAACAATTAAGACTGGATGGAGGCGGATAAAGTTG
CAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAG
CCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCC
CGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGA
TCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATAT
ATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTTG
ATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTA
GAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAAC
AAAAAAACCACCGCTACCAGCGGTGGTTTGTGGCCGGATCAAGAGCTACCAACTCTTTTT
CCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGAGCCGT
AGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCT
GTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACG
ATAGTTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCA
GCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCG
CCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTTCGGAACA

FIG. 22 cont'd

GGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGG
GTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCT
ATGGA AAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCT
CACATGTCCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCC
CGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAG
GGAGTGGCCAACTCCATCACTAGGGGTTTCT (SEQ ID NO: 41)

Nucleotide sequences for constructs are below with enhancer sequence in bold

Construct vAi30.0 nucleotide sequence: rAAV: Grik1_enhGad2-1-Hsp68-EGFP-WPRE3-BGHpA (sequence between ITRs):

GCGGCCGCACGCGTTCTAGACAGGAACATCCAACTGAGCAGCCGGGGTCCCCCCCCACC
CCCCACCCCGCCCCACGCGGCAACTTTGAGCCTGTGCTGGGACAGAGCCTCTAGTTCCTA
AATTAGTCCATGAGGTCAGAGGCAGCACTGCCATTGTAACGCGATTGGAGAGGATCACGT
CACCGGACACGCCCCCAGGCATCTCCCTGGGTCTCCTAACTTGGCGGGGAGAAGTTTTA
GCCCTTAAGTTTTAGCCTTTAACCCCCATATTCAGAACTGTGCGAGTTGGCGAAACCCAC
AAATCACAACAACTGTACACAACACCGAGCTAGAGGTGATCTTTCTTGTCCATTCCACACA
GGCCTTAGTAATGCGTCGCCATAGCAACAGTGTCACTAGTAGCACCAGCACTTCCCCACAC
CCTCCCCCTCAGGAATCCGTACTCTCCAGTGAACCCAGAAACCTCTGGAGAGTTCTGGA
CAAGGGCGGAACCCACAACCTCCGATTACTCAAGGGAGGCGGGGAAGCTCCACCAGACGC
GAACTGCTGGAAGATTCTGCCCCAAGGCCTCCTCCGGCTCGCTGATTGGCCCAGCG
GAGAGTGGGCGGGGCGCGTGAAGACTCCTTAAAGGCGCAGGGCGGCGAGCAGGTCACC
AGACGCTGACAGCTACTCAGAACCAATCTGGTTCCATCCAGAGACAAGCGAAGACAAGA
GAAGCAGAGCGAGCGGGCGCGTTCCCGATCCTCGGCCAGGACCAGCCTTCCCCAGAGCAT
CCCTGCCGCGGAGCGCAACCTTCCAGGAGCATCCCTGCCGCGGAGCGCAACTTTCCCC
GGAGCATCCACGCGCGGAGCGCAGCCTTCCAGAAGCAGAGCGCGGCCACATATGCC
GCCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGT
CGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGC
GATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTG
CCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGAGTGCTTCAGCCGCTACCCC
GACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAG
CGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAG
GGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAA
CATCCTGGGGCACAAGCTGGAGTACAACCTACAACAGCCACAACGTCTATATCATGGCCGA
CAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAG
CGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGC
TGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGC
GCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGAC
GAGCTGTACAAGTAAGTTAATTAATCTCATAATCAACCTCTGGATTACAAAATTTGTGAAAG
ATTGACTGGTATTCTTAACCTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGC
CTTTGTATCATGCTATTGCTTCCCGTATGGCTTTTCAATTTCTCCTCCTTGTATAAATCCTGGT
TAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGGCTC
GGCTGTTGGGCACTGACAATTCCGTGGCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT
GTTTGCCCTCCCCCGTGCTTCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCT
AATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGG

FIG. 22 cont'd

GGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGACTTGCTGC
TATCCCCCTAGGACACTCATTACACCTTTGGAGTTTTCTCACCCTCACATTTTCATCCAGA
AAAAAAAAAAAAATCCTGCAGTGCCTAACATTTCTCAAGGTATTGGATCTCCCTGAGACGTG
TCCTTTACAGGTGACCTGATGCAAGACCTCGCTTGGTTGGCTTGCCAAATCTGGTGGCT
TTGCCAAATCCCAGTCCTGCAAAGCATTCTGTCTCTAACATTTCTGCTTGGAGTACAAAG
AAACTGAAAGTTCATTTCCCAGAGGGTGGGGGGAAACTGGAGCTCTGTCTAAAACCTGGA
TTTTATTGGGTGACGCTTCAAAGCACTCACTTAGCATTCTGCTAACAGAGCTGAACGTTG
AGCCAAAATATTGGAAAAGGAAATGCTCAGGCAGCATGAAAAGACAGAGGAGGCTGGC
AAAGCAGGGAAGCCAAGCCAAGGGTGACAAAATCGAAGGGGAATGTTCCAGAAAAAGG
CATGAAAAGGAACTGTTTCAAGCACAGCGGTTAGAGCCTTATCCCACAGACACATCCA
CATCTTGAAAATAAGTAGCGTTGGAGTTAAGAGTGAGTGTTATTCTCTGATATACTCACTC
ACAGTCTTATGCCAGTGATCACAGATTTTGCCAGCCCTTCTGAACATGATAGGCTCATC
CCACAACATGGCAGTTAGCAATTGATGGCAGGAATCTGATGAGCTTGGACCTCGAAACC
CACGGTGGGGATCCTGTGAATATTTTATCCCTCCAGTTCAGGGGGCTAGAAATAAAACA
GCTTTAGTGTCCACACAGCGAACCCTATATATTCTGAGAAGATATTCTTATTGAGCATA
GTTGTCTCTGAAACATAAGTCTGGGTGTATTTTAGGTTCTACGACGTGGCTTTGAGAAT
GTACTTATCTGTCTCCCACCAGGTTTCCACGTGGTGCGGACCGAGCGGCCGC (SEQ ID
NO: 43)

Construct vAi30.1 nucleotide sequence: rAAV: Grik1_enhGad2-1-pBGmin-EGFP-WPRE3-BGHPA (sequence between ITRs):

GCGGCCGCACGCGTACACTCATTACACCTTTGGAGTTTTCTCACCCTCACATTTTCATC
CAGAAAAAAAAAAAAATCCTGCAGTGCCTAACATTTCTCAAGGTATTGGATCTCCCTGAGA
CGTGTCTTTTCACAGGTGACCTGATGCAAGACCTCGCTTGGTTGGCTTGCCAAATCTGGT
GGCTTTGCCAAATCCCAGTCCTGCAAAGCATTCTGTCTCTAACATTTCTGCTTGGAGTAC
AAAGAACTGAAAGTTCATTTCCCAGAGGGTGGGGGGAAACTGGAGCTCTGTCTAAAAC
TGGATTTTATTGGGTGACGCTTCAAAGCACTCACTTAGCATTCTGCTAACAGAGCTGAA
CGTTCAGCCAAAATATTGGAAAAGGAAATGCTCAGGCAGCATGAAAAGACAGAGGAGG
CTGGCAAAGCAGGGAAGCCAAGCCAAGGGTGACAAAATCGAAGGGGAATGTTCCAGAA
AAAGGCATGAAAAGGAACTGTTTCAAGCACAGCGGTTAGAGCCTTATCCCACAGACAC
ATCCACATCTTGAAAATAAGTAGCGTTGGAGTTAAGAGTGAGTGTTATTCTCTGATATACT
CACTCACAGTCTTATGCCAGTGATCACAGATTTTGCCAGCCCTTCTGAACATGATAGGC
TCATCCCACAACATGGCAGTTAGCAATTGATGGCAGGAATCTGATGAGCTTGGACCTCG
AAACCCACGGTGGGGATCCTGTGAATATTTTATCCCTCCAGTTCAGGGGGCTAGAAATAA
AACAGCTTTAGTGTCCACACAGCGAACCCTATATATTCTGAGAAGATATTCTTATTGAG
CATAGTTGTCTGATCTGAAACATAAGTCTGGGTGTATTTTAGGTTCTACGACGTGGCTTTGAG
AATGTACTTATCTGTCTCCCACCAGGTTTCTTAAGGAGCTCGGGCTGGGCATAAAAGTCA
GGGCAGAGCCATCTATTGCTTACATTTGCTTCTGGGATCCGCCACCATGGTGAGCAAGGG
CGAGGAGCTGTTACCGGGGGTGGTGCCATCCTGGTCGAGCTGGACGGCGACGTAAACG
GCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACC
CTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCTCGTGACCACC
CTGACCTACGGCGTGCAAGTCTCAGCCGCTACCCCGACCATGAAGCAGCACGACTTC
TTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGAC
GGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCAT
CGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGT

FIG. 22 cont'd

ACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGT
GAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCA
GCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCA
CCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGT
TCGTGACCGCCGCGGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGAATTCGATA
TCAAGCTTATCGATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTA
ACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATT
GCTTCCCGTATGGCTTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGA
GGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAAC
CCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTTCGCTTTCCC
CCTCCCTATTGCCACGGCGGAACATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGG
CTCGGCTGTTGGGCACTGACAATTCCGTGGTGTGTCGGGGAAATCATCGTCCTTTCTTG
GCTGCTCGCCTATGTTGCCACCTGGATTCTGCGCGGGACGTCTTCTGCTACGTCCCTTC
GGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCC
GCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCATCG
ATACCGAGCGCTGCTCGAGAGATCTACGGGTGGCATCCCTGTGACCCCTCCCCAGTGCCT
CTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCACCAGCCTTGTCTAATAAAATTAAGTT
GCATCATTTTTGTCTGACTAGGTGTCTTCTATAATATTATGGGGTGGAGGGGGGTGGTATG
GAGCAAGGGGCAAGTTGGGAAGACAACCTGTAGGGCCTGCGGGGTCTATTGGGAACCAA
GCTGGAGTGCAGTGGCACAATCTTGGCTCACTGCAATCTCCGCCTCCTGGGTTCAAGCGA
TTCTCCTGCCTCAGCCTCCCGAGTTGTTGGGATTCCAGGCATGCATGACCAGGCTCAGCT
AATTTTTGTTTTTTTGGTAGAGACGGGGTTTACCATATTGGCCAGGCTGGTCTCCAACTCC
TAATCTCAGGTGATCTACCCACCTTGGCCTCCCAAATTGCTGGGATTACAGGCGTGAACCA
CTGCTCCCTTCCCTGTCTTCTGATTTTGTAGGTAACCACGTGCGGACCGAGCGGCCGC
(SEQ ID NO: 44)

Construct vAi31.0 nucleotide sequence: rAAV: Grik1_enhGad2-2-Hsp68-EGFP-WPRE3-BGHpA (sequence between ITRs):

GCGGCCGCACGCGTTCTAGACAGGAACATCCAACTGAGCAGCCGGGGTCCCCCCCCACC
CCCCACCCCGCCCCACGCGGCAACTTTGAGCCTGTGCTGGGACAGAGCCTCTAGTTCCCTA
AATTAGTCCATGAGGTCAGAGGCAGCACTGCCATTGTAACGCGATTGGAGAGGATCACGT
CACCGGACACGCCCCCAGGCATCTCCCTGGGTCTCCTAACTTGGCGGGGAGAAGTTTTA
GCCCTTAAGTTTTAGCCTTTAACCCCCATATTCAGAACTGTGCGAGTTGGCGAAACCCAC
AAATCACAACAACTGTACACAACACCGAGCTAGAGGTGATCTTTCTTGTCCATTCCACACA
GGCCTTAGTAATGCGTCGCCATAGCAACAGTGTCACTAGTAGCACCAGCACTTCCCCACAC
CCTCCCCCTCAGGAATCCGTACTCTCCAGTGAACCCCAAGAACTCTGGAGAGTTCTGGA
CAAGGGCGGAACCCACAACCTCCGATTACTCAAGGGAGGCGGGGAAGCTCCACCAGACGC
GAACTGCTGGAAGATTCTTGGCCCCAAGGCCTCCTCCGGCTCGCTGATTGGCCCAGCG
GAGAGTGGGCGGGGCGGTGAAGACTCCTTAAAGGCGCAGGGCGGCGAGCAGGTCACC
AGACGCTGACAGCTACTCAGAACCAATCTGGTTCCATCCAGAGACAAGCGAAGACAAGA
GAAGCAGAGCGAGCGGCGCGTTCCCGATCCTCGGCCAGGACCAGCCTTCCCCAGAGCAT
CCCTGCCGCGGAGCGCAACCTTCCAGGAGCATCCCTGCCGCGGAGCGCAACTTTCCCC
GGAGCATCCACGCCGCGGAGCGCAGCCTTCCAGAAGCAGAGCGCGGCGCCACATATGCC
GCCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGT
CGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGGCGAGGGC

FIG. 22 cont'd

GATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTG
CCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAAGTTCAGCCGCTACCCC
GACCACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCCGAAGGCTACGTCCAGGAG
CGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAG
GGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAA
CATCCTGGGGCACAAGCTGGAGTACAACACAAGCCACAACGTCTATATCATGGCCGA
CAAGCAGAAGAACGGCATCAAGGTGAAGTTCAGATCCGCCACAACATCGAGGACGGCAG
CGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGC
TGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGC
GCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGAC
GAGCTGTACAAGTAAGTTAATTAATCTCATAATCAACCTCTGGATTACAAAATTTGTGAAAG
ATTGACTGGTATTCTTAAGTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGC
CTTTGTATCATGCTATTGCTTCCCGTATGGCTTTTCAATTTCTCCTCCTTGTATAAATCCTGGT
TAGTTCTTGCCACGGCGGAAGTCAATCGCCGCTGCTTGGCCGCTGCTGGACAGGGGCTC
GGCTGTTGGGCACTGACAATTCGCTGGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTT
GTTTGCCCTCCCCCGTGCCTTCTTGACCCTGGAAGGTGCCACTCCCCTGTCTTTCT
AATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCTATTCTATTTCTGGGGGGTGG
GGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGACTTGCTGC
TATCCCCTAGGAGAGGAGGCAATATTTGGGATGTTAATAAATAAAGTAATTAATTAATAA
AAATAGATACTATAAAACAACTACAAGCAAACAAACCTAGAACTACAAAGGCCTAAGG
GTGTTGCTTTGGAAGAGCTAACCAAGAAAGTAAGTTCAAAAGGTTTTGAAAGTAGAAG
GCAAGGCCTTTGAGACGCCAGGGGGTAATGCGAGGAGGTAATATGAATTCTTGAACCA
GCATGTTTCATTAGAGTTGTATCAAAATCAGGGAGTGAGAAGAGATGTTTCTATTTATGTG
GAGGGAGGAATAATTGGAAGAGAGATTTTAAACCTGTAATTAGGAACTTGGAATTGATAC
AAAGCAGAATGGGAGAATCGATTGATGTGAATTAAACAACAGCAGTGATGAGTCCCCT
GCTGGGCGGGGATGGATTGGCAGGCCTGGGCCTTTTGGCTGTCTACAAGGCAGGCAGA
GTGGTACCAGCCAGATCAGTTGCCAGCAGTGGTATACTGTGCCAAGTCAAAGAAATGTG
ACTGAGATAGGATCCACCAATCTGAGCAATGGCTCCAGCCTCCTCTCCAGCCTTCCAAG
AAACCCCATGAGTAATCTCCTAAAACCTTTTGTCTTTCTTCCGTCTCTATTCTTGAAACAT
CGTTTTTTCCAGTAGAACTGCATGTGATATATTTGGACATGAAGTAGTGATTGTTTTCTCA
CTAAATCATCTAGAGCCAAATGAAAGCGGGTGTTATTGAAAAGAAATAAATTATTTAAC
ATCAAATAAATTATATAAATATATAGTTTGAGGGGAGAAACATTAAAGATAATCCTGCCA
TTTATCACACTATGAACTGGACTTAGAATGATTCTGTTTCTGCAATTTTATTCGGCATA
AGTATGAATCCTTTCTATATATAATGTGGCTAACTTCAGGAAAATCACACGTTGGTGCGG
ACCGAGCGGCCGC (SEQ ID NO: 45)

Construct vAi11.0 nucleotide sequence: rAAV: mscRE5-pBGmin-EGFP-WPRE3-BGHpA
(sequence between ITRs):

GCGGCCGCACGCGTTTGGCAACAGAGAAGCACTCAGTTTTATTCTTTTCTCATTGTATC
CAGCTTCAGTTTGAAAGCATGGATTGCATGCTTTCTGTGGATGGCAGAAGATGCAACTCT
GTGACAGTGTGCTCCTTATTAATTCCAACATTTCCACATTGCTGTGCTGAGTCGTACACAT
GGGCTTGGCATATGAGCTGTTTACACATTGAAACCAGTTCTGAAAACGTAGCAAAGTGAC
AGGTGCACATCCAGAAGGATATGCAGATTAATTTTTTTGTGCAAATACCTTTGTGTTTTCT
CCACAAATTCCTCCCTAACGTCGTACATCAAGCTACACCACCCAGTTAATTTGCTATTATG
CCTTGCGTCAGAATCCTCACCTTTTTTAGTTGCTTGATTGATGTGAACTCACACTATTTGG

FIG. 22 cont'd

AAACAGCACAAATACACCATTTGTGCTTTTCCAGGTCTTTAGCTCTTCCGTTTGCACCTTT
AAAGATTTTTCTTACATAAAAGTCAGGATATTCAAATATTCATCCAAGTTTTATAGTCTCTT
GGTTGGCTTATCTTTCACCTCATCCTGGGAGCTCGGGCTGGGCATAAAAGTCAGGGCAGAG
CCATCTATTGCTTACATTTGCTTCTGGGATCCGCCACCATGGTGAGCAAGGGCGAGGAGC
TGTTCAACGGGGTGGTGCCCATCCTGGTTCGAGCTGGACGGCGACGTAAACGGCCACAAG
TTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTT
CATCTGCACCACCGGCAAGCTGCCCCTGCCCTGGCCCACCCTCGTGACCACCCTGACCTA
CGGCGTGCAAGTCTTACGCCGCTACCCCGACCATGAAGCAGCACGACTTCTTCAAGTC
CGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTA
CAAGACCCGCGCCGAGGTGAAGTTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGA
AGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACCTACA
ACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAA
GATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACA
CCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCC
GCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACC
GCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGAATTCGATATCAAGCTT
ATCGATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTT
GCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCG
TATGGCTTTTCATTTTCTCCTCCTTGATAAATCCTGGTTGCTGTCTTTTATGAGGAGTTGTG
GCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCCACTGG
TTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTTCGCTTTCCCCCTCCCTATT
GCCACGGCGGAACCTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTT
GGGCACTGACAATTCCGTGGTGTGTCGGGGAAATCATCGTCCTTTCTTGGCTGCTCGC
CTATGTTGCCACCTGGATTCTGCGCGGGACGTCTTCTGCTACGTCCCTTCGGCCCTCAAT
CCAGCGGACCTTCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCG
CCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCATCGATAACGAGC
GCTGCTCGAGAGATCTACGGGTGGCATCCCTGTGACCCCTCCCCAGTGCCTCTCCTGGCC
CTGGAAGTTGCCACTCCAGTGCCACCAAGCCTTGTCTAATAAAATTAAGTTGCATCATTTT
GTCTGACTAGGTGTCCTTCTATAATATTATGGGGTGGAGGGGGGTGGTATGGAGCAAGGG
GCAAGTTGGGAAGACAACCTGTAGGGCCTGCGGGGTCTATTGGGAACCAAGCTGGAGTG
CAGTGGCACAATCTTGGCTCACTGCAATCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCC
TCAGCCTCCCGAGTTGTTGGGATTCCAGGCATGCATGACCAGGCTCAGCTAATTTTTGTTT
TTTTGGTAGAGACGGGGTTTTACCATATTGGCCAGGCTGGTCTCCAACCTCTAATCTCAGG
TGATCTACCCACCTTGGCCTCCCAAATTGCTGGGATTACAGGCGTGAACCACTGCTCCCTT
CCCTGTCCTTCTGATTTTGTAGGTAACCACGTGCGGACCGAGCGGCCGC (SEQ ID NO: 46)

Construct vAi12.0 nucleotide sequence: rAAV: mscRE5-pBGmin-FlpO-WPRE3-BGHpA
(sequence between ITRs):

GCGGCCGCACGCGTTTGGCAACAGAGAAGCACTCAGTTTTATTCTTTTCTCATTGTATC
CAGCTTCAGTTTGAAAGCATGGATTGCATGCTTTCTGTGGATGGCAGAAGATGCAACTCT
GTGACAGTGTGCTCCTTATTAATTCCAACATTTCCACATTGCTGTGCTGAGTCGTACACAT
GGGCTTGGCATATGAGCTGTTTACACATTGAAACCAGTTCTGAAAACGTAGCAAAGTGAC
AGGTGCACATCCAGAAGGATATGCAGATTAATTTTTTTGTGCAAATACCTTTGTGTTTTCT
CCACAAATTCCTCCCTAACGTCGTACATCAAGCTACACCACCCAGTTAATTTGCTATTATG
CCTTGCCTCAGAATCCTCACCTTTTTTAGTTGCTTGATTGATGTGAACCTCACACTATTTGG

FIG. 22 cont'd

AAACAGCACAAATACACCATTTGTGCTTTTCCAGGTCTTTAGCTCTTCCGTTTGCACTTTT
AAAGATTTTTCTTACATAAAAGTCAGGATATTCAAATATTCATCCAAGTTTTATAGTCTCTT
GGTTGGCTTATCTTTCACTCATCCTGGGAGCTCGGGCTGGGCATAAAAGTCAGGGCAGAG
CCATCTATTGCTTACATTTGCTTCTGGCGTGGCCACCATGGCTCCTAAGAAGAAGAGGAAG
GTGATGAGCCAGTTCGACATCCTGTGCAAGACCCCCCCCCAAGGTGCTGGTGCGGCAGTTC
GTGGAGAGATTGAGAGGGCCAGCGGCGAGAAGATCGCCAGCTGTGCCGCCGAGCTGAC
CTACCTGTGCTGGATGATCACCCACAACGGCACCGCCATCAAGAGGGGCCACCTTCATGAG
CTACAACACCATCATCAGCAACAGCCTGAGCTTCGACATCGTGAACAAGAGCCTGCAGTTC
AAGTACAAGACCCAGAAGGCCACCATCCTGGAGGGCCAGCCTGAAGAAGCTGATCCCCGCC
TGGGAGTTACCATCATCCCTTACAACGGCCAGAAGCACCAGAGCGACATCACCGACATC
GTGTCCAGCCTGCAGCTGCAGTTCGAGAGCAGCGAGGAGGCCGACAAGGGCAACAGCCA
CAGCAAGAAGATGCTGAAGGCCCTGCTGTCCGAGGGCGAGAGCATCTGGGAGATCACCG
AGAAGATCCTGAACAGCTTCGAGTACACCAGCAGGTTACCAAGACCAAGACCCTGTACC
AGTTCCTGTTCTGCGCCACATTCATCAACTGCGGCAGGTTACGCGACATCAAGAACGTGGA
CCCCAAGAGCTTCAAGCTGGTGCAGAACAAAGTACCTGGGCGTGATCATTCAAGTGCCTGGT
GACCGAGACCAAGACAAGCGTGTCCAGGCACATCTACTTTTTACGCGCCAGAGGCAGGAT
CGACCCCTGGTGTACCTGGACGAGTTCCTGAGGAACAGCGAGCCCGTGCTGAAGAGAG
TGAACAGGACCGGCAACAGCAGCAGCAACAAGCAGGAGTACCAGCTGCTGAAGGACAAC
CTGGTGCAGCTACAACAAGGCCCTGAAGAAGAACGCCCCCTACCCCATCTTCGCTATC
AAGAACGGCCCTAAGAGCCACATCGGCAGGCACCTGATGACCAGCTTTCTGAGCATGAAG
GGCCTGACCGAGCTGACAAACGTGGTGGGCAACTGGAGCGACAAGAGGGCCTCCGCCGT
GGCCAGGACCACCTACACCCACCAGATCACCGCCATCCCCGACCACTACTTCGCCCTGGT
GTCCAGGTACTACGCCTACGACCCCATCAGCAAGGAGATGATCGCCCTGAAGGACGAGAC
CAACCCCATCGAGGAGTGGCAGCACATCGAGCAGCTGAAGGGCAGCGCCGAGGGCAGCA
TCAGATACCCCGCCTGGAACGGCATCATCAGCCAGGAGGTGCTGGACTACCTGAGCAGCT
ACATCAACAGGCGGATCTGAGAATTGATATCAAGCTTATCGATAATCAACCTCTGGATTAC
AAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATAC
GCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTTCAATTTCTCCTCCTTG
TATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCG
TGGTGTGCACTGTGTTTGTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTC
AGCTCCTTTCCGGGACTTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACCTCATCGCCG
CCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCTGTTGT
TGTCGGGGAAATCATCGTCCTTTCTTGGCTGCTCGCCTATGTTGCCACCTGGATTCTGCG
CGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGG
CCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCCTCAGACGAGTCGGAT
CTCCCTTTGGGCGCCTCCCCGCATCGATACCGAGCGCTGCTCGAGAGATCTACGGGTG
GCATCCCTGTGACCCCTCCCCAGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCC
CACCAGCCTTGTCTAATAAAATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTTCTATAA
TATTATGGGGTGGAGGGGGGTGGTATGGAGCAAGGGGCAAGTTGGGAAGACAACCTGTA
GGGCCTGCGGGGTCTATTGGGAACCAAGCTGGAGTGCAGTGGCACAATCTTGGCTCACTG
CAATCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTTGTTGGGATT
CCAGGCATGCATGACCAGGCTCAGCTAATTTTTGTTTTTTGGTAGAGACGGGGTTTTACC
ATATTGGCCAGGCTGGTCTCCAACCTCCTAATCTCAGGTGATCTACCCACCTTGGCCTCCCA
AATTGCTGGGATTACAGGCGTGAACCACTGCTCCCTTCCCTGTCCTTCTGATTTTGTAGGT
AACCACGTGCGGACCGAGCGGCCGC (SEQ ID NO: 47)

FIG. 22 cont'd

Construct vAi14.0 nucleotide sequence: rAAV: mscRE8-pBGmin-FlpO-WPRE3-BGHpA

(sequence between ITRs):

GCGGCCGCACGCGTCCTGAGGCTTGAGCTAGACACCTGGAGATTCAAGCTTACATGCAA
CAACAAAGGAAAGTACTTAAAATTCTAGCAAGTTAAACAAATAGCCATTTAATCCCTTGCT
AGCATAATTTCTATTACATCTAAGATGCCAATTTCAATCATATAAATATAAGAAATAAA
GAGATTTCTATTCTCAAGGCATCATAACTCCTTCTCTAGACCTTTGGTGATGTAAATGAAG
TCTCTGAGACTGTCATACTGAGGAATGATGTTCACTCCATTGTGTGAAATCTGGGCTC
TACATTTCTGATATTTATGATGCAAATGCTAGCAGGCAAAGTGCAAGTGAAATCAATATG
ATTTTATTTAGTATTCAGGCAACAGGCTCCATCTGTACAATATTCTCCAAGATGATATTTA
ACAGAACTATAGAACTAATGACTGCTCAGATATAAATGACACTTGATTTTAACTGAAAA
GCAATCTTACAACACCAAAGAAGTGCCAGTCATTTTCATTCTAAAGAATGCCCTGTGTATT
TAAAGCACCAAGAGCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACA
TTTGCTTCTGGCGTGGCCACCATGGCTCCTAAGAAGAAGAGGAAGGTGATGAGCCAGTTC
GACATCCTGTGCAAGACCCCCCAAGGTGCTGGTGCGGCAGTTCGTGGAGAGATTCTGA
GAGGCCCAGCGGCGAGAAGATCGCCAGCTGTGCCGCCGAGCTGACCTACCTGTGCTGGA
TGATCACCCACAACGGCACCGCCATCAAGAGGGGCCACCTTCATGAGCTACAACACCATCA
TCAGCAACAGCCTGAGCTTCGACATCGTGAACAAGAGCCTGCAGTTCAAGTACAAGACCC
AGAAGGCCACCATCCTGGAGGCCAGCCTGAAGAAGCTGATCCCCGCCTGGGAGTTCACC
ATCATCCCTTACAACGGCCAGAAGCACCGAGCGACATCACCGACATCGTGTCCAGCCTG
CAGCTGCAGTTCGAGAGCAGCGAGGAGGCCGACAAGGGCAACAGCCACAGCAAGAAGAT
GCTGAAGGCCCTGCTGTCCGAGGGCGAGAGCATCTGGGAGATCACCGAGAAGATCCTGA
ACAGCTTCGAGTACACCAGCAGGTTACCAAGACCAAGACCCTGTACCAGTTCCTGTTCT
GGCCACATTCATCAACTGCGGCAGGTTTCAGCGACATCAAGAACGTGGACCCCAAGAGCTT
CAAGCTGGTGCAGAACAAGTACCTGGGCGTGATCATTAGTGCCTGGTGACCGAGACCAA
GACAAGCGTGTCCAGGCACATCTACTTTTTTCAGCGCCAGAGGCAGGATCGACCCCTGGT
GTACCTGGACGAGTTCCTGAGGAACAGCGAGCCCGTGCTGAAGAGAGTGAACAGGACCG
GCAACAGCAGCAGCAACAAGCAGGAGTACCAGCTGCTGAAGGACAACCTGGTGCAGCAGC
TACAACAAGGCCCTGAAGAAGAACGCCCCCTACCCCATCTTCGCTATCAAGAACGGCCCT
AAGAGCCACATCGGCAGGCACCTGATGACCAGCTTTCTGAGCATGAAGGGCCTGACCGAG
CTGACAAACGTGGTGGGCAACTGGAGCGACAAGAGGGCCTCCGCCGTGGCCAGGACCAC
CTACACCCACCAGATCACCGCCATCCCCGACCACTACTTCGCCCTGGTGTCCAGGTACTA
CGCCTACGACCCCATCAGCAAGGAGATGATCGCCCTGAAGGACGAGACCAACCCCATCGA
GGAGTGGCAGCACATCGAGCAGCTGAAGGGCAGCGCCGAGGGCAGCATCAGATACCCCG
CCTGGAACGGCATCATCAGCCAGGAGGTGCTGGACTACCTGAGCAGCTACATCAACAGGC
GGATCTGAGAATTCGATATCAAGCTTATCGATAATCAACCTCTGGATTACAAAATTTGTGAA
AGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAAT
GCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTTCAATTTCTCCTCCTTGTATAAATCCTG
GTTGCTGTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCAC
TGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCC
GGGACTTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACCTCATCGCCGCCTGCCTTGCC
CGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCGCTGGTGTGTCGGGGAA
ATCATCGTCCCTTTCTTGGCTGCTCGCCTATGTTGCCACCTGGATTCTGCGCGGGACGTCC
TTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCG
GCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGG

FIG. 22 cont'd

GCCGCCTCCCCGCATCGATACCGAGCGCTGCTCGAGAGATCTACGGGTGGCATCCCTGT
GACCCCTCCCCAGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCCACCAGCCTT
GTCCTAATAAAATTAAGTTGCATCATTTTTGTCTGACTAGGTGTCCTTCTATAATATTATGGGG
TGGAGGGGGGTGGTATGGAGCAAGGGGCAAGTTGGGAAGACAACCTGTAGGGCCTGCGG
GGTCTATTGGGAACCAAGCTGGAGTGCAGTGGCACAATCTTGGCTCACTGCAATCTCCGC
CTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTTGTTGGGATTCCAGGCATG
CATGACCAGGCTCAGCTAATTTTTGTTTTTTGGTAGAGACGGGGTTTCACCATATTGGCCA
GGCTGGTCTCCAACCTCTAATCTCAGGTGATCTACCCACCTTGGCCTCCCAAATTGCTGGG
ATTACAGGCGTGAACCACTGCTCCCTTCCCTGTCTTCTGATTTTGTAGGTAACCACGTGC
GGACCGAGCGGCCGC (SEQ ID NO: 48)

**Construct CN1525 nucleotide sequence: rAAV: hsA2-eHGT_079h-minRho-SYFP2-
WPRE3-BGHpA (sequence between ITRs):**

GCGGCCGCAACGCGTTTAGAACAATGGCTGGCCCATAGTAAATGCCGTGTTAGTGTGTTA
GTTGCTGTTCTTCCACGTCAGAAGAGGGCACAGACAAATTACCACCAGGTGGCGCTCAGAG
TCTGCGGAGGCATCACAACAGCCCTGAATTTGAATCCTGCTCTGCCACTGCCTAGTTGAGA
CCTTTTACTACCTGACTAGCTGTTTGTGTATTTTAGGTGTTTGTCTGCAGTGGGGTCAGG
AGTAAGTGAGGTGAATAAGATACCCAGGATATGAAATTTAAAGAGGGCATTCACTTTCAGA
GTCCTGCAAATGCACTTACATAAGTGAGTGCCTCTTTTAATTTTGTACTCTAATTGACTCA
CTTCATCTCATTCTCGTCCACTGCTAGCTTTGAAGGCAGAGAAAACGAGCCATGAGCCAA
AGATTGTGGGTGGCCTCCGAAAGCTATATATCCTGAGCTGACAGCCAGCAAGGAAATGA
GGACCTCAGTAGAATAACCACATGGAAGTGAATTCTACCAATAGCAAGAGTGATATTGG
AAGTGGATTCAATCTTAGAACTCCCAGAAAGGAATACAGCCTGCCTGGCACTTCTATTTT
AATACTGTGGCACTCTAAGCAGAGGACTCCACTGAACCACACCGTACCCATAAACAGGA
AGTCTTTTATGCAGCTGGTATTTATGGGAATTGGAAATTTTGACCTACACAGCTGTGAGAT
AATAAATGGGTGCTGTTTTAAGCCACTAAAATTTTTTATTTTTTAAATTTATTTTTTAATTTT
CATAGGTTTTTGGGGAACAGGGGTATTTGGTTACATGAGTAAGTTCTTTAGTGGTGATTG
GTGAGATTTTGGTGACCTATCACCCAAGCAGTATACACTGAACCCAATTTGTAGTCTTTTA
TCCCTCACTCCCCTCCTACGAATTCGATATCATAATCAACCATAGGTACCGAGCTCGGGAT
TCAGCCGGGAGCTTAGGGAGGGGAGGTCACTTCATAAGGGCCTGGGGGGGGAGTTGGAG
CCACGAGTCGTCCAGCCGGAGCCCCGTGTGGCTGAGCTCCGGCCTCAGAAGCATCCCCG
GGTTGGATCCTTCGAAGCTAGCGCTACCGGTGCGCCACCATGGTGAGCAAGGGCGAGGAG
CTGTTACCGGGGTGGTGCCCATCCTGGTTCGAGCTGGACGGCGACGTAAACGGCCACAA
GTTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGC
TGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGC
TACGGCGTGCACTGCTTCGCCCCGTACCCCGACCACATGAAGCAGCACGACTTCTTCAAG
TCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAAC
TACAAGACCCGCGCCGAGGTGAAGTTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCT
GAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAATA
CAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTT
CAAGATCCGCCACAACATCGAGGACGGCGGCGTGCAGCTCGCCGACCACTACCAGCAGA
ACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCACT
CCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGA
CCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACATCATAATCA
ACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACCTATGTTGCTCCTTTTAC

FIG. 22 cont'd

GCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCA
TTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACATCGCCGCCTG
CCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGCTCGAG
AGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTT
CCTTGACCCTGGAAGGTGCCACTCCCCTGTCTTTTCTAATAAAATGAGGAAATTGCATC
GCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGG
GGGAGGATTGGGAAGACAATAGCAGGCATGCACGTGCGGACCGAGCGGCCGC (SEQ ID
NO: 49)

**Construct CN1528 nucleotide sequence: rAAV: hsA2-eHGT_082h-minRho-SYFP2-
WPRE3-BGHpA (sequence between ITRs):**

GCGGCCGCAACGCGTTTAGAACAATGGCTGGCCCATAGTAAATGCCGTGTTAGTGTGTTA
GTTGCTGTTCTTCCACGTCAGAAGAGGCACAGACAAATTACCACCAGGTGGCGCTCAGAG
TCTGCGGAGGCATCACAACAGCCCTGAATTTGAATCCTGCTCTGCCACTGCCTAGTTGAGA
CCTTTTACTACCTGACTAGCTGTTTGTGTATTTTAGGTGTTTGTGTACTCCTTTGGCCGG
CAGGAGCAGCCTGATTTTAGTGACTCGTTCAGAAGGGAGGTGCAGCTGGAAGGTTGTGT
TTATTAAGCCCAAGTGACTGGAGGTAAATGCAGGTGTGGTCAGGAAATGTAGCTGGTT
CACCTCTCAGTAGGCACAAGGCCAAAGTCAGCCGACTTGCAAGATTGTAATGCCGGTTCT
CCCACCCAGAGGTTATTTGAGTTCATCCAGCACTGCAGCTGAGCAGAATGCTGCAGTCA
GGGTGACACAATGCAAACCTGGAGCCATGTCCTCAAGTGCATGCCAATCAAACCAACTTT
AACCTTCGCTGGTTTGTGGCTGCTGACCTAGAGAGGAATTTGTTATATGGCTGAGGTCCT
GTCCTACCTAAGCCAGACTCCAGTGCAAAAAAAAAATCCTGTCAGATGCGCCTCTTCTTC
TCTGTCTCCATCTCCCTCTCTGCCTCCCGTTTCCCCTTTACTCTTTACTGTTAGTTGGCCTG
GGGTTCTTGGGCTGAAGTGGAGTTTCTCCTGTCTCCGTGTGTTTGCTGCTCTGGTTTGCT
CTTCTTCCAGGGGCCCTTGGTAAAGAAGAGCTGCTGACAGGGTAGAGTGGAGGTGGG
GAAGTGACCACCAGACTGGGAGTGATGTAGGTTGTGGTCATCAATGATGTAGGTTGCTG
TCACCAGGACACTCAAATCAGCTGTGACCAGAGCCTGCAATGACAACGTGGAATTCGATA
TCATAATCAACCATAGGTACCGAGCTCGGGATTACGCCGGGAGCTTAGGGAGGGGAGGTC
ACTTCATAAGGGCCTGGGGGGGGAGTTGGAGCCACGAGTCGTCCAGCCGGAGCCCCGTG
TGGCTGAGCTCCGGCCTCAGAAGCATCCCCGGGTTGGATCCTTCGAAGCTAGCGCTACCG
GTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGT
CGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGC
GATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTG
CCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGCCCGCTACCC
CGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGA
GCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTGCA
GGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCA
ACATCCTGGGGCACAAGCTGGAGTACAACATAACAGCCACAACGTCTATATCACCGCCG
ACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCG
GCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTG
CTGCCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAG
CGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGA
CGAGCTGTACAAGTAAGTCGACATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATT
GACTGGTATTCTTAACATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTT
TGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAG

FIG. 22 cont'd

TTCTTGCCACGGCGGAACATCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGG
CTGTTGGGCACTGACAATTCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAG
CCATCTGTTGTTTGCCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTG
TCCTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTG
GGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATG
CACGTGCGGACCGAGCGGCCGC (SEQ ID NO: 50)

**Construct CN1532 nucleotide sequence: rAAV: hsA2-eHGT_086h-minRho-SYFP2-
WPRE3-BGHpA (sequence between ITRs):**

GCGGCCGCAACGCGTTTAGAACAATGGCTGGCCCATAGTAAATGCCGTGTTAGTGTGTTA
GTTGCTGTTCTTCCACGTCAGAAGAGGACACAGACAAATTACCACCAGGTGGCGCTCAGAG
TCTGCGGAGGCATCACAACAGCCCTGAATTTGAATCCTGCTCTGCCACTGCCTAGTTGAGA
CCTTTTACTACCTGACTAGCTGTTTGTGTATTTTAGGTGTTTGTCTTCTGACCTCCCCGTG
GCCTGCCTTCCCCTGTCTCCTTGGCCGAGCCATCCTCTGTACCTTCACTGTTTCCACTAG
GTTATTGAGGCCTTTTGGTCATGAGCTCCTGCCACCAAATCAGTGCCTGAGGCTGCTGCA
TAACAATCTCCTTCTTGATTAGAGCATCTTAATTATGCTCCAGGAATTTGCTTAAACAGA
TTAATCATTCTGGAGGACAGATTTAGAAGCTGTGCTAGCTGAGGGATCTGCTCTGAGTGT
CCCCGGGGAGATTCTGGTGACAACCGCCTGTTGGGCCTTTGCCAAGGCCTTGGGTGCCT
CGGAGACTCGGACACTGGCTGGAAGCTCTACGCACCCCCGAAGGCTTCTCTTTGCTCCT
CTGGGCCTGTAGAGCAGGCTCTGAAACTCTTGGATGTGAAACGTGGCAGGAGTTTGAGAG
TAAGGAGGCCCTGGTGTCCCCTTTCCCTTCCAGCGGACTCGGAATTCGATATCATAATCAA
CCATAGGTACCGAGCTCGGGATTCAGCCGGGAGCTTAGGGAGGGGAGGTCACTTCATAA
GGGCCTGGGGGGGGAGTTGGAGCCACGAGTCGTCCAGCCGGAGCCCCGTGTGGCTGAG
CTCCGGCCTCAGAAGCATCCCCGGGTGGATCCTTCGAAGCTAGCGCTACCGGTCGCCAC
CATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGG
ACGGCGACGTAAACGGCCACAAGTTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACC
TACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCC
CACCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACAT
GAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCAT
CTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACA
CCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACACAGCCACAACGTCTATATCACCGCCGACAAGCAG
AAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGCA
GCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCG
ACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGT
ACAAGTAAGTCGACATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTA
TTCTTAATATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATG
CTATTGCTTCCCGTATGGCTTTTCAATTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCA
CGGCGGAACATCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGC
ACTGACAATTCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTT
GTTTGCCCTCCCCCGTGCCTTCTTGACCCTGGAAGGTGCCACTCCCCTGTCTTTCT
AATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGG
GGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCACGTGCG
GACCGAGCGGCCGC (SEQ ID NO: 51)

FIG. 22 cont'd

Construct CN1621 nucleotide sequence: rAAV: hsA2-eHGT_128h-minRho-SYFP2-WPRE3-BGHpA (sequence between ITRs):

GCGGCCGCAACGCGTTTGAACAATGGCTGGCCCATAGTAAATGCCGTGTTAGTGTGTTA
GTTGCTGTTCTTCCACGTCAGAAGAGGGCACAGACAAATTACCACCAGGTGGCGCTCAGAG
TCTGCGGAGGCATCACACAGCCCTGAATTTGAATCCTGCTCTGCCACTGCCTAGTTGAGA
CCTTTTACTACCTGACTAGCTGTTTGTGTATTTTAGGTGTTTGTTCATCAAGGTCAGGGGC
CTTACCTCTCCTAGTATCATGAGGAGGATCTCAAAGCAGGGAAGGGAAGCCAGCTGAGG
GACCCCTGAGAGGAGATCAGGGGATAAATACTTTGCAACAGAATCCCCATGGGAAACAT
GACAAAATTAGTAACAGGAATGGAGCCTCCCTCAGAGGATACATGAATCGTTGAGTTGA
AAATAACCACAAGGTCATCAAATCCAGCTGCCACCAACTGCCTGAATATTTTCAACAAAT
TTTTGAACAAGCAGTCCTCCAGCCTCCTCCAGCTGCATCCAGCAGCTCAGTGCTCCACCT
GACAGCCTCTTCTGATAGCGTGAAGGCTCAGAAAGCCCCCTGTGGGGGCCAGCACATCCT
GGGGCAGGCCTCTAGTGGGCAGAGGGCAGCATACTGGGGTCTGGATCTAACACACTCC
TCCATCCAGGTGGTATATTTATAGCCCTGTGCTTTACAGAAAAGACTTCTGATGAGATCAA
GACTTTAGGCTGTTTATGTGGCTTCTTTCATGAGCCAACTGAATTGCATATCATAATCAACC
ATAGGTACCGAGCTCGGGATTCAGCCGGGAGCTTAGGGAGGGGAGGTCACTTCATAAGG
GCCTGGGGGGGGGAGTTGGAGCCACGAGTCGTCCAGCCGGAGCCCCGTGTGGCTGAGCT
CCGGCCTCAGAAGCATCCCCGGGTTGGATCCTTCGAAGCTAGCGCTACCGGTGCGCCACCA
TGGTGAGCAAGGGGCGAGGAGCTGTTCAACGGGGTGGTGCCCATCCTGGTCGAGCTGGAC
GGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTA
CGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCA
CCCTCGTGACCACCCTGGGCTACGGCGTGCAAGTGTTCGCCCGCTACCCCGACCACATG
AAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATC
TTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGGCGACAC
CCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGG
GGCACAAGCTGGAGTACAACACAGCCACAACGTCTATATCACCGCCGACAAGCAGA
AGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGA
CAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCA
CATGGTCTGCTGGAGTTCGTGACCGCCGCGGGGATCACTCTCGGCATGGACGAGCTGTA
CAAGTAAGTCGACATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTAT
TCTTAACATATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGC
TATTGCTTCCCGTATGGCTTTTCAATTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCAC
GGCGGAACATCATCGCCGCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCA
CTGACAATTCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTG
TTTGCCCTCCCCCGTGCTTCCCTTGACCTGGAAGGTGCCACTCCCACTGTCTTTCTTA
ATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGG
GTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCACGTGCGGA
CCGAGCGGCCGC (SEQ ID NO: 52)

Construct CN1633 nucleotide sequence: rAAV: hsA2-eHGT_140h-minRho-SYFP2-WPRE3-BGHpA (sequence between ITRs):

GCGGCCGCAACGCGTTTGAACAATGGCTGGCCCATAGTAAATGCCGTGTTAGTGTGTTA
GTTGCTGTTCTTCCACGTCAGAAGAGGGCACAGACAAATTACCACCAGGTGGCGCTCAGAG

FIG. 22 cont'd

TCTGCGGAGGCATCACAACAGCCCTGAATTTGAATCCTGCTCTGCCACTGCCTAGTTGAGACCTTTTACTACCTGACTAGCTGTTTGTGTATTTTAGGTGTTTGTATCTATACCTCTGAGAAAGCCATTTTTTTTAAACATGAAGAAGTATAAAATATTGTGTGTGAAGATTAAGCCCTTTCTTGCTCTATGTATTTGAGCCTAGAAAAAGACACACTCCCAAGAAATTAGAATAAAGGGAAAAATGACTTTGAGATGTACTTTCTTAAATAACTGAAGTAATGACCGCATGAGCCAGCTGATGGGTTTTTAAATGGATACGTTTCTATCAGCCTGCTGCTTGATGCCAGAGCCAAATATATATGGAGTGTATTATATCATATCCTGTAGGCAGGAGACTGTTGGGGCATTGGGGGACTTAGAGAGGTGAAGTGGCAGGCTTGGCACAGGAATTAACAGCATGGCTACCGTCAACCAGAATTCGATATCATAATCAACCATAGGTACCGAGCTCGGGATTACGCCGGGAGCTTAGGGAGGGGAGGTCACCTTCATAAGGGCCTGGGGGGGGAGTTGGAGCCACGAGTCGTCCAGCCGGAGCCCGGTGTGGCTGAGCTCCGGCCTCAGAAGCATCCCCGGGTTGGATCCTTCGAAGCTAGCGCTACCGGTGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGCCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGGCTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACATCATAATCAACCTCTGGATTACAAAATTTGTGAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTTCATTTTCTCCTCCTTGATAAATCCTGGTTAGTTCTTGCCACGGCGGAACATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGCTTCCTTGACCCTGGAAGGTGCCACTCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCACGTGCGGACCGAGCGGCCGC (SEQ ID NO: 53)

Construct CN1259 nucleotide sequence: scAAV: eHGT_023h-minBGlobin-SYFP2-WPRE3-BGHpA (sequence between ITRs):

CCTAGGACGCGTCCCTAAGCTGAGCTGAGACGTTTTAGTTATAAATGGCTTCTGTGCTTAGCAATCTGCTCTTTTTATTCCCGTGTGGACTTTCCTAGCTCTGGCCTTATGCTGCACTAGAAAAGATTTAGCAAGGGGAGAGGAAGCAGCCTTCCTTACATAACTGGCCTCTTGTAAGGGAGCAGCTGCTTGGTGGAAAAAGACATTCCCCTCCATGCATCCCTCTCCTTCTGCCTCTGGGGGTTGCAGCTTGAGTCAGAACCAGGACCAATTAACCTCAACCTTTGAGGAAGAGACGCACCCTGGCCCCAGCCACGCCTGTTAGAATCTTCCTAGCTGAGTGACACAGTGACACTCAGCCTCAGTTTCTCTGTAAACAAAATGAAGATAAGAGAGCCGACGAGGATGAAATGGAATAACACACCGTGCGGTGCCTGGTACAGAGTGAGCCCCAGAAGTGTGACGCGGCCTCTTGTTGGCTGTCTGGCTTGACCCGGAGTGACTCTGCCTCCAGTTCTCCGGGATGGGAAGGTGATCCCTGTTTGAGATACCAATTTATAAGAAACCGAGCCCGGGAGCTATTTAGAGGT

FIG. 22 cont'd

GAGGTGATAAACCAGGAGGCCGGCTCCTTCATCCCGGTCATCACGAGAGCTCGGGCTG
GGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGGGATCCAGATCTTT
CGAAGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTACACGGG
GTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTC
CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCA
CCGGCAAGCTGCCCCTGCCCTGGCCACCCCTCGTGACCACCCTGGGCTACGGCGTGACG
TGCTTCGCCCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCC
GAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGC
GCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGA
CTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAA
CGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCA
CAACATCGAGGACGGCGGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCATCG
GCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGC
AAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACGGCGCGCCGCGGCCGCGAA
TTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACT
ATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTT
CCCGTATGGCTTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAA
CTCATCGCCGCTGCTTGGCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAA
TTCCGTGGCTCGAGCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCC
GTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCTTTCTTAATAAAATGAGGAAA
TTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACA
GCAAGGGGGGAGGATTGGGAAGACAATAGCAGGCATGACTAGT (SEQ ID NO: 54)

**Construct CN2045 nucleotide sequence: rAAV: 3xSP10ins-eHGT_359h-minRho*-SYFP2-
WPRE3-BGHpA (sequence between ITRs):**

GCGGCCGCACGCGCCGGTACCGAAGCTACCCCTAACACACTATTCTACACACAGAAAATG
CTCTTCACTAGGAAGCTACCCCTAACACACTATTCTACACACAGAAAATGCTCTTCACTAGG
AAGCTACCCCTAACACACTATTCTACACACAGAAAATGCTCTTCACTAGACGCGTT**CACAG**
CTCCTTGGTCTGAACTGGTACCCAGGATCAAGATGTCCCTCCCATGATTCTAATCTCCAG
CCCCCTTACACCATAGTCACTTCCAACTCTGGATTTCTTAACATTGCAGTTTGCTTCAA
AAAAAAATTAAAAAAAACCTCATCTGGGCATTAGAAGTCATTCTGAAGAGGCTGTCCT
ATATCTCAGTGCAATTTCTACTTATCCACTCCCGACTATACCCTGCTTTCATCAGATTGC
AACTGACCTAATTTGTAAAGACTTCAGTGAAAAGAGCCTCATTAAGCCTTCATCTCTT
TGCAGCCTCTTTTCTCCCTCAAATGATCCAGGTCAGAATGGCCCAATAAAAAATAGCTG
GGAAAATTACAAAGTCACTCAAGACTCTGCCAGGAGTAGAGATTTTAAGAATTAGTGACA
GACATATACTGTCTAGGAGTAGGGTGGGAGTGGGTTGCATGGCTTAGGGACATAGAGAG
CTCGATTCAAGCCGGGAGCTTAGGGAGGGGAGGTCACTTCATAAGGGCTTGGGGGGGGAG
TTGGAGCCACGAGTCGTCCAGCCGGAGCCCCGTGTGGCTGTGCTCCGGCCTCAGAAGCA
TCCCCGGATCCAGATCTTTCGAAGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGC
GAGGAGCTGTTACACGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGG
CCACAAGTTCAGCGTGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCC
TGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACC
CTGGGCTACGGCGTGACGTGCTTCGCCCCGCTACCCCGACCACATGAAGCAGCAGCACTTC
TTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGAC

FIG. 22 cont'd

GGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCAT
CGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGT
ACAACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGG
CCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGCAGCTCGCCGACCACTACC
AGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGC
TACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAG
TTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACGGC
GCGCCGCGGCGCGAATTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGAT
TGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCT
TTGTATCATGCTATTGCTTCCCGTATGGCTTTTCAATTTCTCCTCCTTGATAAATCCTGGTTA
GTTCTTGCCACGGCGGAATCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCG
GCTGTTGGGCACTGACAATTCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCA
GCCATCTGTTGTTTGGCCCTCCCCCGTGCTTCCCTTGACCCTGGAAGGTGCCACTCCCACT
GTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCT
GGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCAT
GAGATCTCACGTGCGGACCGAGCGGCCGC (SEQ ID NO: 55)

Construct CN1255 nucleotide sequence: scAAV: eHGT_019h-minBGlobin-SYFP2-WPRE3-BGHpA (sequence between ITRs):

CCTAGGACGCGTTCCATCCACAGTCTGTCATCTTCCCCTTGACAGATTGGCTGCCCTCC
CCCCTTGAAATCACTATTAACTGCAGAGATAACTGCCCGTCTTAGGGCTGCAGTCCTC
CCACATAACCTGAAGCCACTGCCATTTGTCACAACAGATCATTTCTTGAATCTGACAGC
CAGCCTTGGAAGCAGGAAGCATGCTCATTTTCTCACCTGCCGGAGCTTCAGAAAGAA
CTCAGCTCAAAATAGACGCTATTGAAGCAGGAGACACATGAGGAGGTACACAGCCTATC
AATACCTCAACACTGGCCTCTCCTCGCCCTCCAATCCTATTGACTGCATGAGAGCCACTG
CTGAAGGCTTTGTAATCCAAAACCATTAAGTATTCATGGGGCTGGCAACCACAAATAGAA
AATGAATTGGATTGCTCCCTTTCAAATTAGTATAATGACATTATCTGGGAAATGATGATTT
AAAAAATAGTTCAAACATGGGTTTGAATGTTCTTCTACTGTTCTGCCCAGGCTCCTCTC
TGTTTTTTACAATAGAGAAGAGCCCTCCTGGTGCTGTGGAGACATTCCGAGCTCGGGCTG
GGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGGGATCCAGATCTTT
CGAAGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTACCCGGG
GTGGTGCCCATCCTGGTTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTC
CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCA
CCGGCAAGCTGCCCCTGCCCTGGCCACCCCTCGTGACCACCCCTGGGCTACGGCGTGACG
TGCTTCGCCCCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCC
GAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGC
GCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGA
CTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAA
CGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCA
CAACATCGAGGACGGCGGCGTGAGCTCGCCGACCACTACCAGCAGAACACCCCCATCG
GCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGC
AAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGCG
GATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACGGCGCGCCGCGGCGCGAA
TTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACT
ATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTT

FIG. 22 cont'd

CCCGTATGGCTTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAA
CTCATCGCCGCTGCTTGGCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAA
TTCCGTGGCTCGAGCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCC
GTGCCTTCTTGACCCTGGAAGGTGCCACTCCCACTGTCTTTCCTAATAAAATGAGGAAA
TTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACA
GCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGACTAGT (SEQ ID NO: 56)

Construct CN1408 nucleotide sequence: rAAV: eHGT_064h-minBglobin-SYFP2-WPRE3-BGHpA (sequence between ITRs):

GCGGCCGCACGCGTATGAACTCTTAAATCTCAAATGTGTTTGCCCTAGCTGCTTTGACTAA
CCCCCTCTTCTATTTTCAGTTATGCGGCAAGTTGCATACTCAGGTGCCCCTTCTGACTACTT
GAATACTTTCCCTGTGATGTAAGAAGTGTTCATTGGTAAAGTTGTGGTATATAATTAC
AATTGAACTCTCTTGTACTTGCCTCTTTTACAAAAATTCTCTCCTAGCAGAACGTAGTGTG
AGTCATCTACACAGCTGTTTTTCTGATTATTGGAATTTTCTTTTGACATGAAGGAAGTATCT
CATTGACAGAAGCTGCGTTGTGAAGGAGTGCTAACTGTAGCATAAAATACAAAATTGGATT
TTTAGATTGCAAAAATACAGTAAAGCTTTGAAAAGTATTTGGCATGACATTTAACTCAATAC
ATTTTGCTTAAAAAATATTAGCCAAGAACCCTATCAACTTGTTTTTGAATAAACTTCTGT
ATGGACCTTAAATTCATGCTGAGTTTGACCGCATTTTCTTGCACTGGTAGCATTTTCCCT
CTGAGTCATCCTCATTTCTTCTACTTTCTCACATGACTAGGTAAAGATAACTCATGTATTT
GCGAGCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCT
GGGATCCAGATCTTTTGAAGCTAGCGCTACCGGTGCGCCACCATGGTGAGCAAGGGCGAG
GAGCTGTTACCGGGGTGGTGCCCATCCTGGTTCGAGCTGGACGGCGACGTAAACGGCCA
CAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGA
AGCTGATCTGCACCACCGGCAAGCTGCCCCTGCCCTGGCCCACCCTCGTGACCACCCTG
GGCTACGGCGTGCAAGTCTTCCGCCGCTACCCCGACCATGAAGCAGCAGCACTTCTTC
AAGTCCGCCATGCCCAGAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGC
AACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGA
GCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACA
ACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCA
ACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGCAGCTCGCCGACCACTACCAG
CAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCTA
CCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTT
CGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACGGCG
CGCCGCGGCGCGAATTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATT
GACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTT
TGTATCATGCTATTGCTTCCCGTATGGCTTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAG
TTCTTGCCACGGCGGAACATCGCCGCTGCTTGGCCGCTGCTGGACAGGGGCTCGG
CTGTTGGGCACTGACAATTCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAG
CCATCTGTTGTTTGGCCCTCCCCGTGCCTTCTTGACCCTGGAAGGTGCCACTCCCACTG
TCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTG
GGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATG
CACGTGCGGACCGAGCGGCCGC (SEQ ID NO: 57)

Construct CN1258 nucleotide sequence: scAAV: eHGT_022h-minBglobin-SYFP2-WPRE3-BGHpA (sequence between ITRs):

FIG. 22 cont'd

CCTAGGACGCGTAGTCCTGTGGATACCATCTGAGCCTGTGCCCTTTTCTCTCTACTCTTGA
GCAAGTGCCCATCAGGAGCCACATTTGAGTCAGGACCTGCCAAGAGCATTTTCATCTCAT
CCTCAGTTCCCCTTTTGTCTGGAACCACTCGCTATGAGTTTCTATGAGATCGTAGCCAGG
TTTGACCAAACCACTCCTGCACACATGTTGTTTTGAAAGATTAGCACAGAATCCAACACA
ACATTTCCCTTGGATTTCATCTTTCAAGCCAACAGAGAAAGAGCTATTTCCAAAAAATAA
CACATTATTAACCAATGAATAAGAGCCAAGAAAACAATGTAGCCAGCATATTTATTCCAA
ACAGCTTGTTTAGAATGCAAATTGCAAATACATCTATTTCCACCCGGCTCTTTGGGGAAC
CATGCTGGAGCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTG
CTTCTGGGATCCAGATCTTTTGAAGCTAGCGCTACCGGTGCGCCACCATGGTGAGCAAGGG
CGAGGAGCTGTTACCGGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACG
GCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACC
CTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCAC
CCTGGGCTACGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCAGCACTT
CTTCAAGTCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGA
CGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCA
TCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAG
TACAACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAG
GCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGGCGTGCAGCTCGCCGACCACTA
CCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGA
GCTACCAGTCCAAGCTGAGCAAAGACCCCCAACGAGAAGCGCGATCACATGGTCCTGCTGG
AGTTCGTGACCGCCGCGGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACG
GCGCGCCGCGGCCGCGAATTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAG
ATTGACTGGTATTCTTAAGTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGC
CTTTGTATCATGCTATTGCTTCCCGTATGGCTTTTCAATTTCTCCTCCTTGTATAAATCCTGGT
TAGTTCTTGCCACGGCGGAACATCGCCGCTGCCTTGCCCGCTGCTGGACAGGGGGCTC
GGCTGTTGGGCACTGACAATTCGCTGGCTCGAGCGACTGTGCCTTCTAGTTGCCAGCCAT
CTGTTGTTTGCCCTCCCCCGTGCCTTCTTGACCCTGGAAGGTGCCACTCCCACTGTCCT
TTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTATTCTATTCTGGGGG
GTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGACTAG
T (SEQ ID NO: 58)

**Construct CN1279 nucleotide sequence: scAAV: eHGT_022m-minBGlobin-SYFP2-
WPRE3-BGHpA (sequence between ITRs):**

CCTAGGACGCGTGTGTCCTCCAGAGTCCGTGGGTCCTGTGATCCCTTGCTCCATTCAATG
TGCTCATCAGAAGCCATATTTGAAGCAGCATCTACCATCCATGGTCCACCTTCTGTCTGG
ATCCACCCATTTAGGAAACCATATCCTTGTTTGACCAGATCACTCCTGCACACGTGTTGC
TTTGAATTCTAGCACAGGATTCAGCATGACATTTCCCTTGGATTCTTCTGAGCCAACA
GAGTGAGGGCTATTTCCCAAAAAAACAACACATTATTAACAAATGAATAAGAGCCAAGA
AAACAATACAGCCAGCATATTTAGTCCACGCAAGCTTGTTTAGAATACAAATCGCAAATG
TATTTAATTCACAAGGCTCTTGGGGAAACTTTTCGAGCTCGGGCTGGGCATAAAAGTCAGG
GCAGAGCCATCTATTGCTTACATTTGCTTCTGGGATCCAGATCTTTTGAAGCTAGCGCTAC
CGGTGCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGGTGGTGCCCATCCTG
GTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGG
GCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCC
GTGCCCTGGCCACCCCTCGTGACCACCCCTGGGCTACGGCGTGCAGTGCTTCGCCCGCTA

FIG. 22 cont'd

CCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAGGCTACGTCCA
GGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTT
CGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACG
GCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCACCG
CCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACG
GCGGCGTGAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTG
CTGCTGCCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAG
AAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGGATCACTCTCGGCATG
GACGAGCTGTACAAGTAAGTCGACGGCGCGCCGCGGGCGCAATTCGATATCATAATCAA
CCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACATATGTTGCTCCTTTTACG
CTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTTAT
TTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGC
CTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGCTCGAGC
GACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCCTTGACC
CTGGAAGGTGCCACTCCCACTGTCCTTTCTAATAAAATGAGGAAATTGCATCGCATTGTC
TGAGTAGGTGTATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGAT
TGGAAGACAATAGCAGGCATGACTAGT (SEQ ID NO: 59)

Construct CN1253 nucleotide sequence: scAAV: eHGT_017h-minBGlobin-SYFP2-WPRE3-BGHpA (sequence between ITRs):

CCTAGGACGCGTGGTAACAGCCTGAGGGCTGCAGGTCACACCCACAGGGCCAAGGGGT
GGGCAGCATGGAGGTGCAGGGTGGCAGGGACCCTGGGCGGGGCGGCACAGCTGTGCG
GGAGGCTGGGCTGCTGGCATCAGCAGGCGCCCTCCTCCCCACCTCGCTAAACAATCAT
TGCACAAAATATGCAAATGGTATAATTACTGTTATTTGTTTTGCTGATATAAGTGTTTGAA
ATGCAAATGTCAAGTTTGGGCGCCTTCATTTTTCCAACCCTCTCACCCGGACATTTGCAA
GTTGATGAGTTGTTCTTCATCCTGGAAGGAGGAGGAGGAGCTCCCCCAACCGCCAGGG
TGCCAGGGGAGTGAGTCCAGCGTGGCAGCCGCCACTGCCTGCCCCGAGGGCACTGCTGG
GCCCCCTTCCGACGGCACACAGTGAGCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCC
ATCTATTGCTTACATTTGCTTCTGGGATCCAGATCTTTCGAAGCTAGCGCTACCGGTGCGC
ACCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCT
GGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCC
ACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCCGTGCCCTG
GCCACCCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGCCCCGCTACCCCGACC
ACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCA
CCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCG
ACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCC
TGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATCACCGCCGACAAGC
AGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTG
CAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCC
CGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGA
TCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGGATCACTCTCGGCATGGACGAGCT
GTACAAGTAAGTCGACGGCGCGCCGCGGCGCGAATTCGATATCATAATCAACCTCTGGA
TTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACATATGTTGCTCCTTTTACGCTATGTGG
ATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTTATTTTCTCCTC
CTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCCG

FIG. 22 cont'd

TGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGCTCGAGCGACTGTGCC
TTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGT
GCCACTCCCCTGTCTTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGT
GTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGAC
AATAGCAGGCATGACTAGT (SEQ ID NO: 60)

Construct CN1274 nucleotide sequence: scAAV: eHGT_017m-minBGlobin-SYFP2-WPRE3-BGHPA (sequence between ITRs):

CCTAGGACGCGTTGTGTGTGGTCTTAGGGACAGGGGGTGGCTGGCAGGCAGGCAGTGG
CGGGAGGCACAGCTGTGCAGGAGGACGGACTGCTGGCATCAGCAAGCGCCCCCTCCTCC
CCCCACACTAAACAATCACGGCACAAATATGCAAATGGTATAATTACTGTTATTTGTTTT
GCTGATATAAGTGTGTTGAAATGCAAATGTCAACTCTGGGCGCTGTTGGTTTTTCCCACCC
CTCTCTCCCGGACATTTGCAAGTCGATCAGTGAGCCCTCATCCTGGGAACACTAGGGGT
GCCTCCTTCTGACAGACTTGGAGCGAGCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCC
ATCTATTGCTTACATTTGCTTCTGGGATCCAGATCTTTCGAAGCTAGCGCTACCGGTGCC
ACCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCT
GGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCC
ACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCCTGCCCTG
GCCACCCCTCGTGACCACCCTGGGCTACGGCGTGCAAGTCTTCGCCCGCTACCCCGACC
ACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCA
CCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCG
ACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCC
TGGGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCACCGCCGACAAGC
AGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTG
CAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCC
CGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGA
TCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGGATCACTCTCGGCATGGACGAGCT
GTACAAGTAAGTCGACGGCGCGCCGCGGCGCGAATTCGATATCATAATCAACCTCTGGA
TTACAAAATTTGTGAAAGATTGACTGGTATTCTTAAGTATGTTGCTCCTTTTACGCTATGTGG
ATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTTCATTTTCTCCTC
CTTGATATAATCCTGGTTAGTTCTTGCCACGGCGGAACATCGCCGCCTGCCTTGCCCGC
TGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGCTCGAGCGACTGTGCC
TTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGT
GCCACTCCCCTGTCTTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGT
GTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGAC
AATAGCAGGCATGACTAGT (SEQ ID NO: 61)

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

MELSWHVVFIALLSFSCWGSWESDRNFISTAGPLTNDLLHNLSGLLGDQSSNFVAGDKDMYV
CHQPLPTFLPEYFSSLHASQITHYKVFLSWAQLLPAGSTQNPDEKTVQCYRRLKALKTARLQP
MVILHHQTLPLASTLRRTAFADLFADYATFAFHSGDLVGIWFTFSDLEEVIKELPHQESRASQL
QTLSDAHRKAYEIIYHESYAFQGGKLSVVLRAEDIPELLLEPPISALAQDVTDFLSLDLSYECQNE
ASLRQKLSKLQTIKPKVKVIFNLKLPDCPSTMKNPASLLFSLFEINKDQVLTIGFDINEFLSCSS
SSKKSMSCSLTGLALQPDQQQDHETDSSPASAYQRVWEAFANQSRAERDAFLQDTFPEGF
LWGASTGAFNVEGGWAEGGRGVSIWDPRRPLNTTEGQATLEVASDSYHKVASDVALLCGLR

FIG. 22 cont'd

AQVYKFSISWSRIFPMGHGSSPSLPGVAYYNKLIDRLQDAGIEPMATLFHWDLPQALQDHGGW
QNESVVDADFLDYAAFCFSTFGDRVKLWTFHEPWMSYAGYGTGQHPPGISDPGVASFKVAH
LVLKAHARTWHHYNHSHRPQQQGHVGIVLNSDWAEPLSPERPEDLRASERFLHFMLGWFAH
PVFVDGDYPATLRTQIQQMNRQCSPVAQLPEFTEAEKQLLKGSADFLGLSHYTSRLISNAPQ
NTCIPSYDTIGGFSQHVNHVWPQTSSSWIRVVPWVGIRLLQFVSLEYTRGKVPYIYLAGNGMPIG
ESENLFDDSLRVDYFNQYINEVLKAIKEDSDVRSYIARSLIDGFEGPSGYSQRFGLHHVNFSD
SSKSRTPRKSAYFFTSIIKNGFLT KGAKRLLPNTVNLPSKVRAFTFPSEVPSKAKVVWEKFSS
QPKFERDLFYHGTFRDDFLWGVSSSAYQIEGAWDADGKGPSIWDNFTHTPGSNVKDNATGDI
ACDSYHQLDADLNLMLRALVKAYRFSISWSRIFPTGRNSSINSHGVDDYNNRLINGLVASNIFPM
VTLFHWDLPPQALQDIGGWENPALIDLFDSDYADFCFQTFGDRVKFWMTFNEMMYLAWLGYGSG
EFPPGVKDPGWAPYRIAHAVIKAHARVYHTYDEKYRQEKGVISLSLSTHWAEPKSPGVPRDV
EAADRMLQFSLGWFAHPIFRNGDYPDTMKWKVGNRSELQHLATSRLPSFTEEEKRFIRATADV
FCLNTYYSRIVQHKTPRLNPPSYEDDQEMAEEDPSWPSTAMNRAAPWGTRLLNWIKEEYG
DIPIYITENGVLNTPNTEDTDRIFYHKTYINEALKAYRLDGDIDLRGYVAWSLMDNFEWLNGYTV
KFGLYHVDNFNNTNRPR TARASARYYTEVITNNGMPLAREDEFYGRFPEGFIWSAASAAAYQIE
GAWRADGKGLSIWDTFSHTPLRVENDAIGDVACDSYHKIAEDLVTLQNLGVSHYRFSISWSRIL
PDGTTYRINEAGLNYVRLIDTLAASIQQVVTIYHWDLPQTLQDVGWENETIVQRFKEYADVL
FQRLGDKVKFWITLNEPFVIAYQGYGYGTAAPGVSNRPGTAPYIVGHNLKAHAEAHWLYNDVY
RASQGGVISITISSDWAEPDPSNQEDVEAARRYVQFMGGWFAHPIFKNGDYNEVMKTRIRDR
SLAAGLNKSRSLPEFTESEKRRINGTYDFFGFNHYTTVLAYNLNYATAISSFDADRGVASIADRS
WPDSGSFWLKMTPFGFRILNLWLKEEYNDPPIYVTENGVSQREETDLNDTARIYYLRTYINEAL
KAVQDKVDLRGYTVWSAMDNFEWATGFSERFGLHFVNYS DPSLPRIPKASAKFYASVVRONG
FPDPATGPHACLHQPDAGPTISPVRQEEVQFLGLMLGTTEAQTALYVLFSLVLLGVCGLAFLSY
KYCKRSKQKGTQRSQQELSPVSSF (SEQ ID NO: 62)

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

MLPLWTL SLLLGA VAGKEVCYERLGCFSDDSPWSGITERPLHILPWSPKDVNTRFLLYTNENPN
NFQEVAADSSSISGSNFKTNRKTRFIIHGFIDKGEENWLANVCKNLFKVESVNCICVDWKGGSR
TGYTQASQNI RIVGA EVAYFVEFLQSAFGYSPSNVHVIGHSLGAHAAGEAGRRTNGTIGRITGL
DPAEPCFQGTPELVRLDPSDAKFVDVIHTDGAPIVPNLGFGMSQVVGHLDFFPNGGVEMPGC
KKNILSQIVDIDGIWEGTRDFAACNHLRSYKYTTDSIVNPDGFAGFPCASYNVFTANKCFPCPS
GGCPQMGMHYADRYPGKTNVDVGQKFYLDTGDA SNFARWRYKVSVTLSGKKVTGHILVSLFGNK
GNSKQYEIFKGT LKPDSTHSNEFDSDDVDVGD LQMVKFIWYNNVINPTLPRVGASKIIVETNVGK
QFNFCSPETVREEVLLTLTPC (SEQ ID NO: 63)

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

MSISLFGGRYDNKFLNMSSAPKIELIVDKVASLSEGRLEGRLPEDWFRHIMDPETEFNSEFA
DALCIGIDEFAQPLPFLPFKALLVTGTAGAGKTN SIQT LANLDCIVTATT SIAAQNLSVVLNRSKS
AQVKTIFKTGFGFNSSHVSMSE RQSYIANDERSIQIQKQDLSIYWNVISDIADRALGAVACKTKE
LPDLCESSVIVIDEAGVILRHILHTVVFYWFYNALYKTPLYEDGIVPCIVCVGSPTQSNALVTSFN
PLTQNKDVKRGIDVLSALICDDVLSKYCEVDNNWIFVNNKRCADHAFGDFLKHIEFGLPLKPELI
EYVDQFVKPASYIRNPMNEIETTRLFLSHNEVKNYFRSLHEQVEVTNRNNLNFVFPVYFLIKNKT
EDYKSEIGNFSLEIEPWFKSNIHRLNTYSQFADQDLSKTVQLEEVLEDGSVEETLITCHLKHIRN
SSIGVTSKIKASTVGFSGTYEKFVELLQSDLFIEKTSCDQTIHAYSFLSGLMFGGMYSFCCSKFT
TPEVLMEIKNIKMP SIEFLESEMSRMSPDVQTVETDERYDFGLVDDGLSDVDLLEIDPCGDPFFT

FIG. 22 cont'd

RYSKLPLTNSLSFEEISLLYTTFKDIFISRFAILQKHTKGKFGKTLLVTYNRNNVSRKQCGEIYSHL
KSFYGMITYAIPANNYTLEGYTNDNVVHLGTDKQLPQILYKKGLPRLVIKDEMGFISVLDNNVSK
FIDVVNGQSFHLCTTVDYATVSKVSMTITKSQGLSIQKVAIDFGSDPKNLKLSSIVGMSRVTD
NNLIMNVNPLRLNYENDNFIAPHIVKALKNENTMLIF (SEQ ID NO: 64)

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

MKFFLLFTIGFCWAQYSPNTQQGRTSIVHLFEWRWVDIALECERYLAPKGFGGVQVSPPNEN
VAIYNPFRPWWERYQPVSYKLCRSGNEDEFNMVTRCANNVGVRIYVDAVINHMCGNAVSAG
TSSTCGSYFNPGSRDFPAVPYSGWDFNDGKCKTGSGDIENYNDATQVRDCRLTGLLDLALEK
DYVRSKIAEYMNHLIDIGVAGFRLDASKHMPGDIKAILDKLHNLNSNWFPAGSKPFIYQEVIDL
GGEPIKSSDYFGNGRVTEFKYGAKLGTVIRKWNGEKMSYLNWGEWGFVPSDRALVFVDN
HDNQRGHGAGGASILTFWDARLYKMAVGFMALHPYGFTRVMSSYRWPRQFQNGNDVNDWW
GPPNNNGVIKEVTINPDTTCGNDWVCEHRWRQIRNMVIFRNVVDGQPFTNWDYDNGSNQVAF
GRGNRGFIVFNDDWSFSLTLQTGLPAGTYCDVISGDKINGNCTGIKIYVSDDGKAHFSISNSA
EDPFIAHAESKL (SEQ ID NO: 65)

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

MGVRHPPCSHRLLAVALVSLATAALLGHILLHDFLLVPRELSGSSPVLEETHPAHQQGASRPG
PRDAQAHPRPRAVPTQCDVPPNSRFDCAPDKAITQECCARGCCYIPAKQGLQGAQMGP
WCFPPSPSYKLENLSSEMGTATLTRTPTFFPKDILTRLDDVMMETENRLHFTIKDPANR
RYEVPLETPRVHSRAPSPLYSVEFSEEPFGVIVHRQLDGRVLLNTTVAPLFFADQFLQLSTSLP
SQYITGLAEHLSPLMLSTSWTRITLWNRDLAPTPGANLYGSHPFYLALEDGGSAGHVFLLSNA
MDVVLQPSPALSWRSTGGILDVYIFLGPPEKSVVQQYLDVVGYPFMPYWGFGHLCRWGYS
STAITRQVVENMTRAHFPLDVQWNDLDYMDSRDFTFNKDGFRDFPAMVQELHQGGRRYMM
IVDPAISSSGPAGSYRLYDEGLRRGVFITNETGQPLIGKVWPGSTAFPDTNPTALAWWEDMV
AEFHDQVPFDGMWIDMNEPSNFIRGSEDGCPNNELENPPYVPGVVGGLQAATICASSHQFL
STHYNLHNLGLTEAIAASHRALVKARGTRPFVISRSTFAGHGRYAGHWTGDVWSSWEQLASS
VPEILQFNLLGVPLVGADVCGFLGNTSEELCVRWTQLGAFYPMRNHNSLLSLPQEPYSFSEP
AQQAMRKALTLRYALLPHLYTLFHQAHVAGETVARPLFLEFPKDSSTWTVDHQLLWGEALLITP
VLQAGKAEVTGYFPLGTWYDLQTVPIEALGSLPPPPAAPREPAIHSEGQWWTLPAPLDTINVHL
RAGYIPLQGPGLTTTESRQQPMALAVALTKGGEARGELFWDDGESLEVLERGAYTQVIFLARN
NTIVNELVRVTSEGAGLQLQKVTVLG VATAPQQVLSNGVPVSNFTYSPDTKVLDICVSLLMGEQ
FLVSWC (SEQ ID NO: 66)

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

MSDQDHSMDEMTAVVKIEKGVGGNNGGNGGGGAFSQRSSSTGSSSSTGGGGQESQSPSP
LALLAATCSRIESPNENSNNNSQGSPSQSGGTGELDLTATQLSQGANGWQIISSSSGATPTSKEQ
SGSSTNGSNGSESSKNRTVSGGQYVVAAPNLQNNQVLTGLPGVMPNIQYQVIPQFQTVDGQ
QLQFAATGAQVQQDGSGQIQIIPGANQQIITNRGSGGNIIAAMPNLLQQAVPLQGLANNVLSGQ
TQYVTNVPVALNGNITLLPVNSVSAATLTPSSQAVTISSSGSQESGSQPVTSGTTISSASLVSSQ
ASSSSFFTANANSYSTTTTTSNMGIMNFTTSGSSGTNSQGQTPQRVSGLQGS DALNIQQNQTS
GGSLQAGQQKEGEQNQQTQQQQILIQPQLVQGGQALQALQAAPLSGQTFTTQAISQETLQNL
QLQAVPNSGPIIRTPTVGPNGQVSWQTLQLQNLQVQNPQAQTITLAPMQGVSLGQTSSSNTT
LTPIASAASIPAGTVTVNAAQLSSMPGLQTINLSALGTSGIQVHPIQGLPLAIANAPGDHGAQLGL
HGAGGDGIHDDTAGGEEGENSPDAQPPQAGRRTREACTCPYCKDSEGRSGDPPGKKKQHIC

FIG. 22 cont'd

HIQGCGKVYGKTSHLRAHLRWHTGERPFMCTWSYCGKRFTSRDELQRHKRTHHTGEKKFACP
ECPKRFMRSDHLSKHIKTHQNKKGPGVALSVGTLPLDSGAGSESGTATPSALITTNMVAME
AICPEGIARLANSGINVMQVADLQSNISGNGF (SEQ ID NO: 67)

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

MTAKMETTFYDDALNASFLPSESGPYGYSNPKILKQSMTLNLADPVGSLKPHLRKNSDLLTSP
DVGLLLKLASPELERLIQSSNGHITTTPTPTQFLCPKNVTDEQEGFAEGFVRALAEHSQNTLPS
VTSAAPVNGAGMVAPAVASVAGGSGSGGFSASLHSEPPVYANLSNPNPGALSSGGGAPSY
GAAGLAFPAQPQQQQPPHLPQQMPVQHPRQLALKEEPQTVPEMPGETPPLSPIDMESQE
RIKAERKMRNRNIAASKCRKRLERARLEEKVKTLKAQNSELASTANMLREQVAQLKQKVMN
HVNSGCQLMLTQQLQTF (SEQ ID NO: 68)

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

MDLPVGPGAAGPSNVPAFLTKLWTLVSDPDTDALICWSPSGNSFHVFDQGGQFAKEVLPKYFK
HNNMASFVRQLNMYGFRKVVHIEQGGLVKPERDDTEFQHPCFLRGQEQLLENIKRKVTSVSTL
KSEDIKIRQDSVTKLLTDVQLMKGKQECMDSKLLAMKHENEALWREVASLRQKHAQQQKVVN
KLIQFLISLVQSNRILGVKRKIPLMLNDSGSAHSMKYSRQFSLEHVHSGPYSA P SPAYSSSSL
YAPDAVASSGPIISDITELAPASPMASPGGSIDERPLSSSPLVRVKEEPPSPPPQSPRVEEASPGR
PSSVDTLTSLPTALIDSILRESEPPASVTALTDARGHTDTEGRPPSPPTSTPEKCLSVACLDKN
ELSDHLDAMDNDNLQTM LSSHGFSVDTSALLDLFSPSVTPDMSLPDL DSSLASIQELLSPQ
EPPRPPEAENSSPD SGKQLVHYTAQPLFLD PGSVDTGSNDLPVLFELGEGSYFSEG DGFAED
PTISLLTGSEPPKAKDPTVS (SEQ ID NO: 69)

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

MQRLVAWDPACLPLPPPPPAFKSMEVANFYEADCLAAAYGGKAAPAAPPAARPGPRPPAGE
LGSIGDHERAIDFSPYLEPLGAPQAPAPATATDTFEAAPAPAPAPASSGQHHDFLSDLFSDDY
GGKNCKKPAEYGYVSLGRLGAAKGALHPGCFAPLHPPPPPPPPPAELKAEPGFEPADCKRKE
EAGAPGGGAGMAAGFPYALRAYLG YQAVPSGSSGSLSTSSSSSPGTPSPADAKAPPTACYA
GAAPAPSQVKSKAKKTVDKHSDEYKIRRRERNNI AVRKSRDKAKMRNLETQHKVLELTAENERL
QKKVEQLSRELSTLRNLFKQLPEPLLASSGHC (SEQ ID NO: 70)

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

MNNPSETSKPSMESGDGNTGTQTNGLDFQKQPVVGGAIATAQAQAFGLHGHVQLAGTSLQ
AAAQSLNVQSKSNEESGDSQQPSQPSQQPSVQAAIPQTQLMLAGGQITGLTLTPAQQLLLQ
QAQAQAQLLAAAVQQHSASQQHSAAGATISASAATPMTQIPLSQPIQIAQDLQQLQQLQQNL
NLQQFVLVHPTTNLQPAQFIISQTPQGQQGLLQAQNLTLQLPQQSQANLLQSQPSITLTSQPAT
PTRTIAATPIQTL PQSQSTPKRIDTPSLEEPSDLEELEQFAKTFKQRRIKLGFTQGDVGLAMGKL
YGNDFSQTTISRFEALNLSFKNMCKLKLPLEKWLND AENLSSDSSLSSPSALNSPGIEGLSRRR
KKRTSIETNIRVALEKSFLENQKPTSEEITMIADQLNMEKEVIRVWFCNRRQKEKRINPPSSGGT
SSSPIKAIFPSPTSLVATTPSLVTSSAATTLTVSPVPLTSAAVTNLSVTGTSDTTSNNTATVISTA
PPASSAVTSPSLSPSPSASASTSEASSASETSTTQTTSTPLSSPLGTSQVMVTASGLQTAAAAA
LQGAAQLPANASLAAMAAAAGLNPSLMAPSQFAAGGALLSLNPGTL SGALSPALMSNSTLATI
QALASGGSPLITSLDATGNLVFANAGGAPNIVTAPLFLNPQNLSLLTSNPVSLVSAAAASAGNSA
PVASLHATSTSAESI QNSLFTVASASGAASTTTTASKAQ (SEQ ID NO: 71)

FIG. 22 cont'd

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

MEAAVAAPRPRLLLLVAAAAAAAAAALLPGATALQCFCCHLCTKDNFTCVTDGLCFVSVTETTDK
VIHNSMCIAEIDLIPDRPFVFCAPSSKTGSVTTTTYCCNQDHCNKIELPTTGPFSVKSSPGLGPVE
LAAVIAGPVCFCISLMLMVYICHNRTVIHHRVPNEEDPSLDRPFISEGTTLKDLIYDMTTSGSGS
GLPLLQRTIARTIVLQESIGKGRFGEVWRGKWRGEEVAVKIFSSREERSWFREAEIYQTVMLR
HENILGFIAADNKDNGTWTQLWLVS DYHEHGS LFDYLNRYTVTVEGMIKLALSTASGLAHLHME
IVGTQGKPAIAHRDLKSKNILVKNGTCCIALDLGLAVRHDSATDTIDIAPNHRVGTKRYMAPEVL
DDSINMKHFESFKRADIYAMGLVFWEIARRCSIGGIHEDYQLPYYDLVPSDPSVEEMRKVVCEQ
KL RPNIPNRWQSCALRVMKIMREC WYANGAARLTALRIKKTLSQLSQQEGIKM (SEQ ID NO:
72)

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

MRLPGAMPALALKGELLLLLLLLLLEPQISQGLVVTTPGPPELVNVSSTFVLTCSGSAPVWVER
MSQEPPQEMAKAQDGTFFSVLTLTNLTGLDTGEYFCTHND SRGLETDERKRLYIFVPDPTVGF
LPND AEELFIFLTEITEITIPCRVTD PQLVVT LHEKKGDVALPVYDHQRGFSGIFEDRSYICKTTI
GDREVDSDAYYVYRLQVSSINVS VNAVQTVVRQGENITLMCIVIGNEVVNFEWYTPRKESGRL
VEPVTDFLLDMPYHIRSILHIPS AELEDSGTYTCNVTESVNDHQDEKAINITVVESGYVRLLGEV
GTLQFAELHRSRTLQVVFEAYPPPTVLWFKDNRTLGDSSAGEIALSTRNVSETRYVSELTLVRV
KVAEAGHYTMRAFHEDA EVQLSFQLQINVPVRVLESESHPDSGEQTVRCRGRGMPQPNIIWS
ACRDLKRCPRELPPTLLGNSSEEEESQLETNVTYWEEEQEF EVVSTLRLQHVD RPLSVRCTLRN
AVGQDTQEVIVPHSLPFKVVISAILALVVLTIISLIILIMLWQKKPRYEIRWKVIESVSSDGHEYIY
VDPMQLPYDSTWELPRDQLVLGRTLGS GAFGQVVEATAHGLSHSQATMKVAVKMLKSTARSS
EKQALMSELKIMSHLGPLN VVNLLGACTKGGPIYIITEYCRYGDLVDYLHRNKHTFLQHHS DK
RRPPSAELYSNALPVGLPLPSHVS LTGESDGGYMDMSKDESVDYVPMLDMKGDVKYADIESS
NYMAPYDNYVPSAPERTCRATLINESPVLSYMDLVGFSYQVANGMEFLASKNCVHRDLAARN
VLICEGKLVKICDFGLARDIMRDSNYISKGSTFLPLKWMAPESIFNSLYTTLSDVWSFGILLWEIF
TLGGTPYPELPMNEQFYNAIKRGYRMAQPAHASDEIYEIMQKC WEEKFEIRPPFSQLVLLLERL
LGEYK KKYQQVDEEFLRSDHPAILRSQARLPGFHGLRSP LDTSSVLYTAVQPNEGDN DYIIP L
PDPKPEVADEGPLEGSPSLASSTLNEVNTSSTISCDSPLEPQDEPEPEPQLELQVEPEPELEQL
PDSGCPAPRAEAEDSFL (SEQ ID NO: 73)

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

MRPSGTAGAALLALLAALCPASRALEEKVCQGT SNKLTQLGTFEDHFLSLQRMFN NCEVVLG
NLEITYVQRNYDLSFLKTIQEVAGYVLIALNTVERIPL ENLQIIRGNMY YENS YALAVLSNYDANKT
GLKELPMRNLQEILHGAVRFSNNPALCNVESIQWRDIVSSDFLSNM SMDFQNH LGSCQKCDPS
CPNGSCWGAGEENCQKLTKIICAQQCSGR CRGKSPSDCCHNQCAAGCTGPRES DCLVCRKF
RDEATCKDTC PPLMLYNPTTYQMDVNPEGKYSFGATCVKK CPRNYVVDH GSCVRACGADS
YEMEEDGVRKCKKEGPCRKVCNGIGIGEFKDSLSINATNIKHFKNCT SISGDLHILPVAFRGDS
FTHTPPLDPQELDILKTVKEITGFLLIQAWPENRTDLHAFENLEIRGR TKQHGGFSLAVVSLNITS
LGLRSLKEISDGDVIIISGNKNLCYANTINWKKLFGTSGQKTKIISNRGENSCKATGQVCHALCSP
EGCWGPEPRDCVSCRNVSRGRECVDKCKLLEGEPREFVENSECIQCHPECLPQAMNITCTGR
GPDNCIQCAHYIDGPHCVKTCPAGVMGENNTLVWKYADAGHVCHLCHPNCTY GCTGPGLEG
CPTNGPKIPSIATGMVGALLLLLVALGIGLFMRRRHIVRKRTLRRLLQERELVEPLTPSGEAPN
QALLRILKETEFKKIKVLGSGAFGT VYKGLWIPEGEKV KIPVAIKELREATSPKANKEILDEAYVM
ASVDNPHVCRLLGICLTSTVQLITQLMPFGCLLDYVREHKDNIGS QYLLNWC VQIAKGMNYLED

FIG. 22 cont'd

RRLVHRDLAARNVLVKTPQHVKITDFGLAKLLGAEKEYHAEGGKVPIKWMMALESILHRIYTHQS
DVWSYGVTVWELMTFGSKPYDGIPASEISSILEKGERLPQPPICTIDVYMIMVKCWMIDADSRP
KFRELIIEFKSMARDPQRYLVIQGDERMHLPSPTDSNFYRALMDEEDMDDVVDADAYLIPQQGF
FSSPSTSRTPLLSSLSATSNNSTVACIDRNLQSCPIKEDSFLQRYSSDPTGALTEDSIDDTF
VPEYINQSVPKRPAGSVQNPVYHNQPLNPAPSRDPHYQDPHSTAVGNPEYLNTVQPTCVNST
FDSPAHWAKGSHQISLDNPDYQQDFFPKEAKPNGIFKGSTAENAEYLRVAPQSSEFIGA
(SEQ ID NO: 74)

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

MVSYWDTGVLLCALLSCLLLTGSSSGSKLKDPESLKGQTQHIMQAGQTLHLQCRGEAAHKWSL
PEMVSKESERLSITKSACGRNGKQFCSTLTNTAQANHTGFYSCYLA VPTS KKKETESA IYFI
SDTGRPFVEMYSEIPEIIHMTGRELVIPCRVTSPNITVTLKKFPLDTLPDGKRIIWD SRKGFIISN
ATYKEIGLLTCEATVNGHLYKTNYLTHRQTNTIIDVQISTPRPVKLLRGHTLVLNCTATTPLNTRV
QMTWSYPDEKNKRASVRRRIDQSN SHANIFYSVLTIDKMQNKDKGLYTCRVRSGPSFKSVNTS
VHIYDKAFITVKHRKQQVLETVAGKRSYRLSMKVKAFFSPEVVWLKDGLPATEKSARYLTRGYS
LIIKDVTEEDAGNYTILLSIKQSNVFNLTATLIVNVKQIYEKAVSSFPDPALYPLGSRQILTCTAY
GIPQPTIKWFWHPCNNHSEARCD FCSNNEESSILDADSNMGNRIESITQRMAIIEGKNKMAST
LVVADSRISGIYICIASNKVGTVGRNISFYITDVPNGFHVNL EKMPTEGEDLKL SCTV NKFLYRDV
TWILLRTVNNRTMHYSISKQKMAITKEHSITLNLTIMNVSLQDSGTYACRARNVYTGEI LQKKEI
TIRDQEAPYLLRNLS DHTVAISSSTTL DCHANGVPEPQITWFKNNHKIQQEPGIILGPSSTLFIE
RVTEEDEGVYHCKATNQKGSVESSAYLTVQGTSDKSNLELITLTCTCVAATLFWLLLT L FIRKMK
RSSSEIKTDYLSIIMDPDEVPLDEQCERLPYDASKWEFARERLKL GKS LGRAFGKV VQASAFG
IKKSPTCRTLAVKMLKEGATASEYKALMTELKILTHIGHHLNVVNLLGACTKQGGPLMVIVEYCK
YGNLSNYLKS KRD LFFLNKDAALHMEPKKEKMEPGLEQGGKPRLD SVTSSSEFASSGFQEDKS
LSDVEEEEDSDGFYKEPITMEDLISYSFQVARGMEFLSSRKCIHRDLAARNILLS ENNVVKICDF
GLARDIYKNPDYVRKGDTRLPLKWMAPESIFDKIYSTKSDVWSYGVLLWEIFSLGGSPYPGVQ
MDEDFCSRLREGMRMRAP EYSTPEIYQIMLDCWHRDPKERPRFAELVEKLGDLLQANVQQDG
KDYIPINAILTGN SGFTYSTPAFSEDFKESISAPKFNSGSSDDVRYVNAFKFMSLERIKTFEELL
PNATSMFDDYQGDSS TLLASPM LKRFTWTD SKPKASLKIDLRVTSKSKESGLSDVSRPSFCHS
SCGHVSEGKRRFTYDHAELERKIACCSPPPDYNSVVLYSTPPI (SEQ ID NO: 75)

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

MSNITDPQMWFDFDDL NFTGMPPADEDYSPCMLETETLNKYVVIAYALVFLLSLLGNSLVMLVIL
YSRVGRSVTDVYLLNLALADLLFALTLP IWAASKVNGWIFGTFLCKVVSLLKEVNFYSGILL LACI
SVDRYLAIVHATRTL TQKRHLVKFVCLGCWGLSMNLSLPFFLFRQAYHPNNSSPVCYEVLGND
TAKWRMVLRLPHTFGFIVPLFVMLFCYGFTLR TL FKAHMGQKHRAMRVIFAVVLIFLLCWL PYN
LVLLADTL MRTQVIQESCERRN NIGRALDATEILGFLHSCLNPIYAFIQNFRHGFLKILAMHGLV
SKEFLARHRVTSYTSSSVNVSSNL (SEQ ID NO: 76)

FIG. 22 cont'd

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

MSGGKYVDSEGHLYTVPIREQGNIYKPNNKAMADELSEKQVYDAHTKEIDLVNRPKHLNDDV
VKIDFEDVIAEPEGTHSFHGIWKASFTTFTVTKYWFYRLLSALFGIPMALIWGIYFAILSFLHIWAV
VPCIKSFLIEIQCTSRVYSIYVHTVCDPLFEAVGKIFSNVRINLQKEI (SEQ ID NO: 77)

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

MGNRGMEELIPLVNKLQDAFSSIGQSCHLDLPQIAVVGGQSAGKSSVLENFVGRDPLPRGSGI
VTRRPLILQLIFSKTEHAEFLHCKSKKFTDFDEVRQEIEAETDRVTGTNKGISPVPINLRVYSPHV
LNLTLDLPGITKVPVGDQPPDIEYRVKDMILQFISRESSLILAVTPANMDLANSDALKLAKEVDP
QGLRTIGVITKLDLMDEGTDARDVLENKLLPLRRGYIGVNNRSQKDIEGKKDIRAALAAERKFFL
SHPAYRHMADRMGTPHLQKTLNQQLTNHIRESLPALRSKLQSQLLSLEKEVEEYKIFRPDDPTP
KTKALLQMVQQFGVDFEKRIEGSGDQVDTLELSGGARINRIFHERFPFELVKMEFDEKDLRREI
SYAIKNIHGVRTGLFTPDLAFEIVKKQVVKLKEPCLKCVDLVIQELINTVRQCTSKLSSYPRLRE
ETERIVTTYIREREGRTKDQILLIDIEQSYINTNHEDFIGFANAQQRSTQLNKKRAIPNQVIRRGW
LTINNISLMKGGKEYWFLTAESLSWYKDEEEKEKKYMLPLDNLKIRDVEKGFMSNKHVFAIF
NTEQRNVYKDLRQIELACDSQEDVDSWKASFLRAGVYPEKDQAENEDGAQENTFSMDPQLER
QVETIRNLVDSYVAIINKSIRDLMPTIMHLMINNTKAFIHHELLAYLYSSADQSSLMEESADQAQ
RRDDMLRMYHALKEALNIIGDISTSTVSTPVPVDDTWLQSASSHSPTPQRRPVSSIHPPGRP
PAVRGPTPGPPLIPVPVGAAASFAPPIPSRPGPQSVFANSDLFPAPPQIPSRPVRIPPGIPPGV
PSRRPPAAPSRPTIIRPAEPSLLD (SEQ ID NO: 78)

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

MAQILPIRFQEHLQLQNLGINPANIGFSTLTMESDKFICIREKVGEEQAQVVIIDMNDPSNPIRRPIS
ADSAIMNPASKVIALKAGKTLQIFNIEMKSKMKAHTMTDDVTFWKWISLNTVALVTDNAVYHWS
MEGESQPVKMFDRHSSLAGCQIINYRTDAKQKWLLLTGISAQQNRVVGAMQLYSVDRKVSQPI
EGHAASFAQFKMEGNAAESTLFCFAVRGQAGGKLHIIEVGTPPTGNQPFPPKAVDVFFPPEAQ
NDFPVAMQISEKHVVFLITKYGYIHLYDLETGTCIYMNRISETIFVTAPHEATAGIIGVNRKGQ
VLSVCVEEENIIPYITNVLQNPDLALRMAVRNNLAGAEELFARKFNALFAQGNYSAAKVAANAP
KGILRTPDTIRRFQSVPAQPGQTSPLLQYFGILLDQGQLNKYESLELCRPVLQQGRKQLLEKWL
KEDKLECSEELGDLVKSVDPTLALSVYLRANVPNKVIQCFEAETGQVQKIVLYAKKVGYPDWIFL
LRNVMRISPDQGGQFAQMLVQDEEPLADITQIVDVFMENLIQQCTAFLLDALKNNRPSEGPLQ
TRLLEMNLMHAPQVADAILGNQMFTHYDRAHIAQLCEKAGLLQRALEHFTDLYDIKRAVVHTHL
LNPEWLVNYFGSLSVEDSLECLRAMLSANIRQNLQICVQVASKYHEQLSTQSLIELFESFKSFE
GLFYFLGSIVNFSQDPDVHFKYIQAACKTGQIKEVERICRESNCYDPERVKNFLKEAKLTDQLPL
IIVCDRDFDFVHDLVLYLRNNLQKYIEIYVQKVNPSRLPVVIGGLLDVDCSEDVIKNLILVVRGQFS
TDELVAEVEKRNRLKLLLPWLEARIHEGCEEPATHNALAKIYDSNNNPERFLRENPPYDSRVV
GKYCEKRDPHLACVAYERGQCDLELINVCNENSLFKSLSRYLVRRKDPELWGSVLLESNPYRR
PLIDQVVQTALSETQDPEEVSVTVKAFMTADLPNELIELLEKIVLDNSVFSEHRNLQNLILLTAIKA
DRTRVMEYINRLDNYDAPDIANIAISNELFEEAFAIFRKFVDNTSAVQVLIEHIGNLDRAVEFAER
CNEPAVWSQLAKAQLQKGMVKEAIDSYIKADDPSSYMEVVQAANTSGNWEELVKYLQMARKK
ARESYVETELIFALAKTNRLAELEEFINGPNNNAHIQQVGDRCYDEKMYDAAKLLYNNVSNFGR
ASTLVHLGEYQAAVDGARKANSTRTWKEVCFACVDGKEFRLAQMCGLHIVVHADELEELINYY
QDRGYFEELITMLEAALGLERAHMGMFTELAILYSKFKPQKMREHLELFWSRVNIPKVLRAAEQ
AHLWAEVLFLYDKYEEYDNAIITMMNHPTDAWKEGQFKDIITKVANVELYYRAIQFYLEFKPLLL
NDLLMVLSPRLDHTRAVNYFSKVQKQLPLVKPYLRSVQNHNNKSVNESLNNLFITEEDYQALRTS

FIG. 22 cont'd

IDAYDNFDNISLAQRLEKHELIEFRRIAAYLFKGNNRWKQSVELCKKDSLYKDAMQYASESKDT
ELAEELLQWFLQEEKRECFGACLFTCYDLLRPDVVLETAWRHNIMDFAMPYFIQVMKEYLTKV
DKLDASESLRKEEEQATETQPIVYGQPQLMLTAGPSVAVPPQAPFGYGYTAPPYGGPQPGFG
YSM (SEQ ID NO: 79)

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

MAQILPVRFQEHFQLQNLGINPANIGFSTLTMESDKFICIREKVGEQAQVTIIDMSDPMAPIRRP
SAESAIMNPASKVIALKAGKTLQIFNIEMKSKMKAHTMAEEVIFWKWWSVNTVALVTETAVYHW
SMEGDSQPMKMFDRHTSLVGCQVIHYRTDEYQKWLLLVGISAQQNRVVGAMQLYSVDRKVS
QPIEGHAAFAEFKMEGNAKPATLFCFAVRNPTGGKLHIIEVGQPAAGNQPFVKKAVDVFFPPE
AQNDFPVAMQIGAKHGVYIITKYGYLHLYDLESGVCICMNRISADTIFVTAPHKPTSGIIGVNKK
GQVLSVCVEEDNIVNYATNVLQNPDLGLRLAVRSNLAGAEKLFVRKFNTLFAQGSYAEAAKVA
ASAPKGILRTRETQKFQSIQAQSGQASPLLQYFGILLDQGQLNKLESLELCHLVLQQGRKQLL
EKWLKEDKLECSEELGDLVKTTPMLALSVELRANVPSKVIQCFATGQFQKIVLYAKKVGYP
DWIFLLRGVMKISPEQGLQFSRMLVQDEEPLANISQIVDIFMENSIIQQCTSFLLDALKNNRPAE
GLLQTWLLEMNLVHAPQVADAILGNKMFTHYDRAHIAQLCEKAGLLQQALEHYTDLYDIKRAV
HTHLLNPEWLVNFFGSLSVEDSVECLHAMLSANIRQNLQLCVQVASKYHEQLGTQALVELFES
FKSYKGLFYFLGSIVNFSQDPDVHLKYIAACKTGQIKEVERICRESSCYNPERVKNFLKEAKLT
DQLPLIIVCDRFGFVHDLVLYLRNNLQRYIEIYVQKVNPSRTPAVIGGLLDVDCSEEVIKHLIMAV
RGQFSTDELVAEVEKRNRLKLLLPWLESQIQEGCEEPATHNALAKIYIDSNNSPECFLRENAYY
DSSVVGRYCEKRDPHLACVAYERGQCDLELIKVCNENSLFKSEARYLVCRKDPELWAHVLEET
NPSRRQLIDQVVQTALSETRDPEEISVTVKAFMTADLPNELIELLEKIVLDNSVFSEHRNLQNLLI
LTAIKADRTRVMEYISRLDNYDALDIASIAVSSALYEEAFTVFHKFDMNASAIQVLIEHIGNLDRAY
EFAERCNEPAVWSQLAQALQKDLVKEAINSIRGDDPSSYLEVVQSASRSNNWEDLVKFLQ
MARKKGRESYIETELIFALAKTSRVSELEDFINGPNNAHIQQVGDRCYEEGMYEAAKLLYSNVS
NFARLASTLVHLGEYQAAVDNSRKASSTRTWKEVCFACMDGQEFRAQLCGLHIVIHADLEEL
LMCYQDRGYFEELILLLEAALGLERAHMGMFTELAILYSKFKPQKMLEHLELFWSRVNIPKVL
AAEQAHLWAEVLFLYDKYEEYDNAVLTMMSHPTAEWKEGQFKDIITKVANVELCYRALQFYLD
YKPLLINDLLLVLSPRLDHTWTVSFFSKAGQLPLVKPYLRSVQSHNNKSVNEALNHLLTEEDY
QGLRASIDAYDNFDNISLAQQLEKHQLMEFRCIAAYLYKGNNWWAQSVELCKKDHLYKDAMQ
HAAESRDAELAQLLQWFLEEGKRECFAACLFTCYDLLRPDMVLELAWRHNLDLAMPYFIQV
MREYLSKVDKLDALSLRKQEEHVTEPAPLVDFDFDGHE (SEQ ID NO: 80)

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

MAELDPFGAPAGAPGGPALGNGVAGAGEEDPAAAFQAQSEIAGIENDEAFAILDGGAPGPQ
PHGEPPGGPDAVDGVMNGEYYQESNGPTDSYAAISQVDRLQSEPESESIRKWREEQMERLEAL
DANSRKQEAWEKAIKELEEYARQDEQLQKTKANNRAAEEAFVNDIDESSPGTEWERVAR
LCDFNPKSSKQAKDVSRMRSVLISLKQAPLVH (SEQ ID NO: 81)

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

MADDFGFFSSSESGAPEAAEEDPAAAFQAQSEIAGIENDEGFGAPAGSHAAPAQPGPTSGA
GSEDMGTTVNGDVFQEANGPADGYAAIAQADRLTQEPESIRKWREEQRKRLQELDAASKVTE
QEWREKAKKDLEEWNRQSEQVEKNKINNRASEEAFVKESKEETPGTEWEKVAQLCDFNPK
SSKQCKDVSRLRSVLSLQKQTPSLR (SEQ ID NO: 82)

FIG. 22 cont'd

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

MSQTAMSETYDFLFKFLVIGNAGTGKSCLLHQFIEKKFKDDSNHTIGVEFGSKIINVGGKYVKLQI
WDTAGQERFRSVTRSYRGAAGALLVYDITSRETYNALTNWLT DARM LASQNIVILCGNKKDL
DADREVT FLEASRFAQENELMFLETSALTGENVEEAFVQCARKILNKIESGELDPERMGSGIQY
GDAALRQLRSPRRAQAPNAQECGC (SEQ ID NO: 83)

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

MGTRDDEYDYLFKVVLIGDSGVGKSNLLSRFTRNEFNLESKSTIGVEFATRSIQVDGKTIKAQIW
DTAGQERYRAITSAYYRGAVGALLVYDIAKHLTYENVERWLKELRDHADSNIVIMLVGNKSDLR
HLRAVPTDEARAF AEKNGLSFIETSALDSTNVEAAFQITLTIYRIVSQKQMSDRRENDMSPSNN
VVPIHVPPTTENKPKVQCCQNI (SEQ ID NO: 84)

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

MNRCWALFLSLCCYLRLVSAEGDPIPEELYEMLSDHSIRSFDDLQRL LHGDPGEEDGAELDLN
MTRSHSGGELESLARGRRSLGSLTIAEPAMIAECKTRTEVF EISRRLIDRTNANFLVWPPCVEV
QRCSGCCNNRNVQCRPTQVQLRPVQVRKIEIVRKKPIFKKATVTLEDHLACKCETVAAARPVT
RSPGGSQEQRATPQTRVTIRTVRVRPPKGKHKRKFHHTDKTALKETLGA (SEQ ID NO: 85)

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

MKMH LQRALVVLALLNFATVSLSLSTCTTLD FGHKKRVEAIRGQILSKLRLTSPPEPTVMTHV
PYQV LALYNSTRELLEEMHGEREEGCTQENTESEYYAKEIHKFDMIQGLAEHNELAVCPKGITS
KVFRFNVSSVEKNRTNLFRAEFRVL RVPNPSSKRNEQRIELFQILRPDEHIAKQRYIGGKNLPTR
GTAEWLSFDVTDTVREWLLRRESNLGLEISIHCPCHTFQPNGDILENIHEVMEIKFKGVDNEDD
HGRGDLGRLKKQKDHHNPHLILMMIPPHRLDNPGQGGQ RKKRALDTNYCFRNLEENCCVRPL
YIDFRQDLGWKWWHEPKGYANFCSGPCPYLRSADTTHSTVLGLYNTLNPEASASPCCPVD
LEPLTILYYVGRTPKVEQLSNMVVKSCCKS (SEQ ID NO: 86)

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

MSILFYVIFLAYLRGIQGNM DQRSLPEDSLNSLIKLIQADILKNKLSKQMVDVKENYQSTLPKA
EAPREPERGGPAKSAFQPVIA MDTELLRQRRYNSPRVLLSDSTPLEPPPLYLMEDYVGSPVV
ANRTSRRKRYAEHKSHRGEYSVCDSESLWVTDKSSAIDIRGHQVTVLGEIKTGNSPVKQYFYE
TRCKEARPVKNGCRGIDDKHWSQCKTSQTYVRALTSENNKLVGWRWIRIDTSCVCALSRKIG
RT (SEQ ID NO: 87)

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

MRPSGTAGAALLALLAALCPASRALEEKKGKGVSRRLPRRPRIAPRTPQPAQPRTGAPARARA
PARPFLFP (SEQ ID NO: 88)

GTPase HRas [Homo sapiens]:

MTEYKL VVVGAGGVGKSALT IQLIQNHVDEYDPTIEDSYRKQVVIDGETCLLDILD TAGQEEYS
AMRDQYMRTGEGFLCVFAINNTKSFEDIHQYREQIKRVKDSDDVPMVLVGNKCDLAARTVESR
QAQDLARSYGIPYIETSAKTRQGVEDAFYTLVREIRQHKLRLNPPDESGPGCMSCKCVLS
(SEQ ID NO: 89)

FIG. 22 cont'd

Cocaine and Amphetamine Regulated Transcript (Chain A) [Homo sapiens], PDB: 1HY9_A:
YGQVPMCDAGEQCAVRKGARIGKLCDCPRGTSCNSFLLKCL (SEQ ID NO: 90)

Protachykinin-1 [Homo sapiens], UniProtKB - P20366:
MKILVALAVFFLVSTQLFAEEIGANDDLNYWSDWYDSDQIKEELPEPFEHLLQRIARRPKPQQFF
GLMGKRDADSSIEKQVALLKALYGHGQISHKRHKTDSFVGLMGKRALNSVAYERSAMQNYER
RR (SEQ ID NO: 91)

Substance P is position 58-68 of Protachykinin-1 [Homo sapiens]:
RPKPQQFFGLM (SEQ ID NO: 92)

Oxytocin-neurophysin 1 [Homo sapiens], UniProtKB - P01178:
MAGPSLACCLLGLLALTSACYIQNCPLGGKRAAPDLVRKCLPCGPGGKGRCFGPNICCAEEL
GCFVGTAEALRCQEENYLPSPCQSGQKACGSGGRCVGLGCCSPDGCHADPACDAEATFSQ
R (SEQ ID NO: 93)

Oxytocin is position 20-28 of Oxytocin-neurophysin 1 [Homo sapiens]:
CYIQNCPLG (SEQ ID NO: 94)

Somatostatin [Homo sapiens], GenBank: AAH32625.1:
MLSCRLQCALAALSIVLALGCVTGAPSDPRLRQFLQKSLAAAAGKQELAKYFLAELLSEPNQTE
NDALEPEDLSQAAEQDEMRLQLRSANSNPAMAPRERKAGCKNFFWKTFTSC (SEQ ID NO:
95)

Myosin light chain kinase, Green fluorescent protein, Calmodulin chimera (Chain A) [synthetic construct], PDB: 3EKJ_A:
MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDLATMVDSSRRKWNKTGHAVRAIGRLSSL
ENVYIMADKQKNGIKANFKIRHNIEDGGVQLAYHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDP
NEKRDHMLLEFVTAAGITLGMDELYKGGTGGSMVSKGEELFTGVVPILVELDGDVNGHKFSV
SGEGEGDATYGKLTCLKFICTTGKLPVPWPTLVTTLXVQCFSRYPDHMKQHDFFKSAMPEGYIQ
ERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNTRDQLTEEQIAEFKE
AFSLFDKDGDDGITTQKLGTVMRSLGQNPTAEALQDMINEVGADGNGTIDFPQFLTMMAKMK
DTDSEEEIREAFRVFGKDGNGYISAAQLRHVMTNLGEKLTDEEVDEMIREAGIDGDGQVNYEQ
FVQMMTAK (SEQ ID NO: 96)

Genetically-encoded green calcium indicator NTnC (chain A) [synthetic construct], PDB: 5MWC_A:
MVSKGEEDNMA SLPATHELHIFGSINGVDFDMVGQSGGNPNVGYEELNLKSTKGDLQFSPWIL
VPHIXFHQYLPYPDGMSPFQAAMVDGSGYQVHRTVQFEDGASLTVNYRYTYEGSHIKGEAQV
KGTGFPADGPVMANSLTAMVPSEEEELSECFRTFDKDGDFIDREEFGGIIRLTGEQLTDEDPD
EIFGDSDDTKNGRIDFDEFLKMVENVQLSMADWCRSKMACPNDKTLISTLKWSYTTGNGKRYR
STARTTYTFAKPMAANYLKNQPMYVFRKTELKHSKTELNFKEWQKAFTDVMGMDELYK (SEQ
ID NO: 97)

FIG. 22 cont'd

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

MVSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTCLKICTTGKLPVPWPTLVTT
 LTWGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELK
 GIDFKEDGNILGHKLEYNYISHNVYITADKQKNGIKAHFKIRHNIEDGSVQLADHYQQNTPIGDG
 PVLLPDNHLYSTQSKLSKDPNEKRDHMLLEFVTAARMLSEEELANCFRIFDKDANGFIDIEELG
 EILRATGEHVTEEDIEDLMKDSKNDGRIDFDEFLKMMEGVQGTSEEELANCFRIFDKDANGF
 IDIEELGEILRATGEHVTEEDIEDLMKDSKNDGRIDFDEFLKMMEGVQELMGGVQLADHYQQ
 NTPIGDGPVLLPDNHLYSYQSKLSKDPNEKRDHMLLEFVTAAGITLGMDELYKGGTGGSMVS
 KGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTCLKICTTGKLPVPWPTLVTTLG
 GLMCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKID
 FKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKIRHNIED (SEQ ID NO: 98)

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

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGLTCLKICTTGKLPVPWPTLVTT
 LYGGLQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELK
 GIDFKEDGNILGHKLEYNYNSHNVYITADKQKNGIKANFKIRHNIEDGGVQLADHYQQNTPIGDG
 PVLLPDNHLYSYQSKLSKDPNEKRDHMLLEFVTAAGITLGMDELYKRMHDQLTEEQIAEFKEA
 FSLFDKDGDTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTIYFPEFLTMMARKMKD
 TDSEEEIREAFRVFDKDGNGYISAAQLRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFV
 QMMTAKGGKRRWKKNFIAVSAANRFKKISSSGALELMTSKVYDPEQRKRMITGPQWWARCK
 QMNVLDSFINYYDSEKHAENAVIFLHGNATSSYLWRHVPHIEPVARCIIPDLIGMGKSGKSGN
 GSYRLLDHYKYLTAWFELLNLPKKIIFVGHDWGAALAFHYAYEHQDRIKAIVHMESVVDVIESW
 DEWPDI EEDIALIKSEEGEKMLENNFFVETVLPSPKIMRKLEPEEFAAYLEPFKEKGEVRRPTLS
 WPREIPLVKGKPKDQVQIVRNYNAYLRASDDLPLKFIESDPGFFSNAIVEGAKKFPNTEFVKVK
 GLHFLQEDAPDEMKGKYSFVERVLKNEQ (SEQ ID NO: 99)

Calcium indicator protein OeNL(Ca²⁺)-18u [synthetic construct], GenBank: BBB18812.1:

MVSVIKPEMKMRYMDGSVNGHEFTIEGEGTGRPYEGHQEMTLRVTMAEGGPMPFAFDLVS
 HVFCYGHRVFTKYPEEIPDYFKQAFPEGLSWERSLEFEDGGSASVSAHISLRGNTFYHKSFT
 GVNFPADGPIMQNSVDWEPSTEKITASDGVLGKGDVTMYLKLEGGGNHKCQFKTTYKAAKEIL
 EMPGDHYIGHRLVRKTEGNITEQVEDAVAHSGTLEDFVGDWRQTAGYNLDQVLEQGGVSSLF
 QNLGVSVTPIQRIVLSGENGLKIDHVIIPYEGPVMHDQLTEEQIAEFKEAFSLFDKDGDTITTK
 ELGTVMRSLGQNPTEAELQDMINEVDADGNGTIYFPDFLTMMARKMKD TDSEEEIREAFRVFD
 KDGNGYISAADLRHVMTNLGEKLTDEEVDEMIREADIDGEGQVNYEEFVQMMTAKGGKRRWK
 KNFIAVSAANRFKKISSSGALELLSGDQMGQIEKIFKVVPVDDHHFKVILHYGTLVIDGVTPNMI
 DYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSLLFRVTINGVTGWRLCERILA (SEQ
 ID NO: 100)

GCaMP6m:

ATGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGG
 GTCGGGATCTGTACGACGATGACGATAAGGATCTCGCCACCATGGTCTGACTCATCACGTC
 GTAAGTGGAATAAGACAGGTCACGCAGTCAGAGCTATAGGTCTGGCTGAGCTCACTCGAGA
 ACGTCTATATCAAGGCCGACAAGCAGAAGAACGGCATCAAGGCCGAACCTTCAAGATCCGCC
 ACAACATCGAGGACGGCGGCGTGCAGCTCGCCTACCACTACCAGCAGAACACCCCCATC

FIG. 22 cont'd

GGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCGTGCAGTCCAAACTTTTCG
AAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGGGCGGTACCGGAGGGAGCATGGTGAGCA
AGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGTGAGGGCGATGCCACCTACGGCAAGCT
GACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCACCCTCGTGAC
CACCTGACCTACGGCGTGCACTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGA
CTTCTTCAAGTCCGCCATGCCCCGAAGGCTACATCCAGGAGCGCACCATCTTCTTCAAGGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACCTGCCGGACCAACTGACTGAAGAGCAGATCGCAGAATTTAAAGAGGCTTTCTC
CCTATTTGACAAGGACGGGGATGGGACAATAACAACCAAGGAGCTGGGGACGGTGATGC
GGTCTCTGGGGCAGAACCCACAGAAGCAGAGCTGCAGGACATGATCAATGAAGTAGATG
CCGACGGTGACGGCACAATCGACTTCCCTGAGTTCCTGACAATGATGGCAAGAAAAGGGA
GCTACAGGGACACGGAAGAAGAAATTAGAGAAGCGTTCGGTGTGTTTGATAAGGATGGCA
ATGGCTACATCAGTGCAGCAGAGCTTCGCCACGTGATGACAAACCTTGGAGAGAAGTTAA
CAGATGAAGAGGTTGATGAAATGATCAGGGAAGCAGACATCGATGGGGATGGTCAGGTAA
ACTACGAAGAGTTTGTACAAATGATGACAGCGAAGTGA (SEQ ID NO: 101)

GCaMP6s:

ATGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGG
GTCGGGATCTGTACGACGATGACGATAAGGATCTCGCCACCATGGTCGACTCATCACGTC
GTAAGTGGAATAAGACAGGTCACGCAGTCAGAGCTATAGGTCGGCTGAGCTCACTCGAGA
ACGTCTATATCAAGGCCGACAAGCAGAAGAACGGCATCAAGGCGAACTTCCACATCCGCC
ACAACATCGAGGACGGCGGCGTGCAGCTCGCCTACCACTACCAGCAGAACACCCCCATC
GGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCGTGCAGTCCAAACTTTTCG
AAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGGGCGGTACCGGAGGGAGCATGGTGAGCA
AGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGTGAGGGCGATGCCACCTACGGCAAGCT
GACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCACCCTCGTGAC
CACCTGACCTACGGCGTGCACTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGA
CTTCTTCAAGTCCGCCATGCCCCGAAGGCTACATCCAGGAGCGCACCATCTTCTTCAAGGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACCTGCCGGACCAACTGACTGAAGAGCAGATCGCAGAATTTAAAGAGGCTTTCTC
CCTATTTGACAAGGACGGGGATGGGACAATAACAACCAAGGAGCTGGGGACGGTGATGC
GGTCTCTGGGGCAGAACCCACAGAAGCAGAGCTGCAGGACATGATCAATGAAGTAGATG
CCGACGGTGACGGCACAATCGACTTCCCTGAGTTCCTGACAATGATGGCAAGAAAATGA
AATACAGGGACACGGAAGAAGAAATTAGAGAAGCGTTCGGTGTGTTTGATAAGGATGGCA
ATGGCTACATCAGTGCAGCAGAGCTTCGCCACGTGATGACAAACCTTGGAGAGAAGTTAA
CAGATGAAGAGGTTGATGAAATGATCAGGGAAGCAGACATCGATGGGGATGGTCAGGTAA
ACTACGAAGAGTTTGTACAAATGATGACAGCGAAGTGA (SEQ ID NO: 102)

ATGGGTTCTCATCATCATCATCATCATGATGGTATGGCTAGCATGACTGGTGGACAGCAAAATGG
GTCGGGATCTGTACGACGATGACGATAAGGATCTCGCCACCATGGTTCGACTCATCACGTC
GTAAGTGGAATAAGACAGGTCACGCAGTCAGAGCTATAGGTCGGCTGAGCTCACTCGAGA
ACGTCTATATCAAGGCCGACAAGCAGAAAGAACGGCATCAAGGCCAACTTCAAGATCCGCC
ACAACATCGAGGACGGCGGCGTGCAGCTCGCCTACCACTACCAGCAGAACACCCCCATC
GGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCGTGCAGTCCAAACTTTTCG
AAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCTGTACCGCCGCGCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGGGCGGTACCGGAGGGAGCATGGTGAGCA
AGGGCGAGGAGCTGTTCAACGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGTGAGGGCGATGCCACCTACGGCAAGCT
GACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGAC
CACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGA
CTTCTTCAAGTCCGCCATGCCCCGAAGGCTACATCCAGGAGCGCACCATCTTCTTCAAGGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACCTGCCGGACCAACTGACTGAAGAGCAGATCGCAGAATTTAAAGAGGAATTCTC
CCTATTTGACAAGGACGGGGATGGGACAATAACAACCAAGGAGCTGGGGACGGTGATGC
GGTCTCTGGGGCAGAACCCACAGAAGCAGAGCTGCAGGACATGATCAATGAAGTAGATG
CCGACGGTGACGGCACAATCGACTTCCCTGAGTTCCTGACAAATGATGGCAAGAAAAATGA
AATACAGGGACACGGAAGAAGAAATTAGAGAAGCGTTCGGTGTGTTTGATAAGGATGGCA
ATGGCTACATCAGTGCAGCAGAGCTTCGCCACGTGATGACAAACCTTGGAGAGAAGTTAA
CAGATGAAGAGGTTGATGAAATGATCAGGGAAGCAGACATCGATGGGGATGGTCAGGTAA
ACTACGAAGAGTTTGTACAAATGATGACAGCGAAGTGA (SEQ ID NO: 103)

MSPPTSPTPDTGHDTPDTGHD TGHGAVEICFAPCEEDCVTIRYFVENDFEGCIPGHFDQYSS
HGSLHDIVKAALYICMVISILQILFYGFQWWRKTCGWVWVWVACIETSIYIIAITSEADSPFTLYLT
NGQISPQLRYMEWLMTCPVILIALSNITGMAEEYNKRTMTLLTSDVCCIVLGMMSAASKPRLKGI
LYAVGWAFGAWTYWTALQVYRDAHKAVPKPLAWYVRAMGYVFFTSWLTFFPGWFLLGPEGLE
VVTGTVSTLMHACSDLISKNLWGFMDWHLRVLVARHHRKLFKAE EEHALKKKGQTLEPGMPRS
TSFVRGLGDDVEIDPSYELYRLKRQNHPEYFLSPAQTPRRGPSFDKRTSFEMDGGKNGMLQM
MPVTGMGMGMGMGMGGGKTVLFLDYTG GGYVSFFEQQLSNMGVNVTKCWSDDDDMYNTAG
VANVKQLFHFAMIPNNALGGQMVMDLRGTGLLVVAYGPEPPMPGMGQDEFVPLQMPGPVYD
ESILHNLVMRHAITQGLGMNGMQGNMGQQQQMMGMQGNMNGMQGNMNGMQGNMNGMQ
GNMSGMQGNMNGMQGNSGMNQGWNNQGFTNTGAFGY (SEQ ID NO: 104)

MDTLAWVARELLSSGHGTDATDSGHGTDTSGGHDSSSHDAVAHNVTLLIAPPHAGGHAGPTD
TSQQITGIDGWAIAPAGDCYCAGWYVSHGSSFEATFAHVCQWSIFAVCVLSLLWYAYQYWKAT
CGWEEVYVCCIELVFICFELYHEFDSPCSLYLSTSNVVNWLRYSWLLCCPVILHLSNVTGLSD
DYGRRTMGLLVSDIATIVFGVTAAMLVNWPKIIFYLIGFTMCCYTFFLAAKVLIESFHQVPKGICR
HLVKAMAITYFVGWSFFPLIFLFGQSGFKKISPYADVIASSFGDLISKNAFGMLGHFLRVKIHEHIL
KHGDIRKTTHLRIAGEEKEVETFVEEDEDATAKHSTKELANRGSFIVMRDKMKEQGIDVRASLD
MDEDEEARTGKGKGAGATSLVPGRVILAVPDISMVDFHFDHFAHLGASIELVPALGVENTLLLV

FIG. 22 cont'd

QQAMQLGGLDFVLVHPEFLRDRSQNGLVSRLKMTGHGVCAFGWVPSGPMREIIESAGVDGW
LDGPSFGTGIDQEQLIELIGYMQAKRKFGMRFGGGGASKAGYSSDGGFGGKGMLEMQPSMS
QGSQVPLLQQNNSMMRAPPSPMGNMANNMNMNPMMSMNNPMMGGGAVMMTSMGSMQQ
AANPLYGAPPSPLSSQPGAGMYGAPAPQMGSQGSMSHGSMSYGGSSQQQHQQPQQAAAAAPA
AADGGSEAEMLKQLMSEINRLKAELGES (SEQ ID NO: 105)

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

MDHPVARSLIGSSYTNLNNGSIVIPSDACFCMKWLKSKGSPVALKMANALQWAAFALSVIILIYY
AYATWRTTCGWEEVYVCCVELTKVIEFFHEFDEPGMLYLANGNRVLWRLRYGEWLLTCPVILI
HLSNLTGLKDDYNKRTMRLLVSDVGTIVWGATAAMSTGYIKVIFLLGCMYGANTFFHAAKVYIE
SYHTVPKGLCRQLVRAMAWLFFVSWGMFPVLFLLGPEGFGHLSVYGSTIGHTIIDLLSKNCWG
LLGHFLRLKIHILLYGDIRKVQKIRVAGEELEVETLMTEEAPDTVKKSTAQYANRESFLTMRD
KLKEKGFEVRASLDNSGIDAVINHNHNNYNALANAAAAVGKPGMELSKLDHVAANAAGMGGIA
DHVATTSGAISPGRVILAVPDISMVDYFREQFAQLPVQYEVVPALGADNAVQLVVQAAGLGGC
DFVLLHPEFLRDKSSTSLPARLRSIGQRVAAGWSPVGPVRDLIESAGLDGWLEGPSFGLGISL
PNLASLVLRMQHARKMAAMLGGMGGMLGSNLMSSGGVGLMGAGSPGGGGGGAMGVGMTG
MGMVGTNAMGRGAVGNSVANASMGGGSAGMGMGMGMVGAGVGGQQQMGANGMGPTS
FQLGSNPLYNTAPSPPLSSQPGGDASAAAAAAAAAAAAATGAASNSMNAMQAGGSVRNSGILAGG
LGSMMPGPPGAPAAPTAAATAAPAVTMGAPGGGGAAASEAEMLQQLMAEINRLKSELGE (SEQ
ID NO: 106)

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

MDYGGALSAVGRELLFVTNPVVVNGSVLPEDQCYCAGWIESRGTNGAQTASNVLQWLAAGF
SILLMFYAYQTWKSTCGWEEIYVCAIEMVKVILEFFFEFKNPSMLYLATGHRVQWRLRYAEWLL
TCPVILHLSNLTGLSNDYSRRTMGLLVSDIGTIVWGATSAMATGYVKVIFCLGLCYGANTFFH
AAKAYIEGYHTVPKGRCRQVVTGMAWLFFVSWGMFPILFILGPEGFGVLSVYGSTVGHITIDLM
SKNCWGLLGHYLRVLIHEHILIHGDIRKTTKLNIGGTEIEVETLVEDEAEAGAVP (SEQ ID NO:
107)

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

MLEHKIDFMVTLEVKEANANGDPLNGNMPRTDAKGYGVMSDVSIKRKIRNRLQDMGKSIFVQA
NERIEDDFRSLEKRFSQHFTAKTPDKEIEEKANALWFDVRAFGQVFTYLKKSIGVRGPVSISMA
KSLEPIVISSLQITRSTNGMEAKNNSGRSSDTMGTKHFVDYGVYVLKGSINAYFAEKTGFSQED
AEAIEVLVSLFENDASSARPEGSMRVCEVFWFTHSSKLGNVSSARVFDLLEYHQSIEEKSTYD
AYQIHLNQEKLAKYEAKGLTLEILEGL (SEQ ID NO: 108)

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

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLK
RTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYH
EKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQL
FEENPINASGVDAKAILSARLSKSRRENLIQAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAED
AKLQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDE
HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVK
LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARG
NSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTVY

FIG. 22 cont'd

NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
 FNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLK
 RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIKKGILQTVKVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRER
 MKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQ
 SFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGG
 LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFY
 KVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSESEQEIGKATAKYFF
 YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTG
 GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITI
 MERSSFENPIDFLEAKGYKEVKKDLIIKPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKH
 RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSI TGLYETRIDLSQLG
 GD (SEQ ID NO: 109)

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

MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNNDHYKELKPIIDRIYKTYADQCL
 QLVQLDWNLSAAIDSYRKEKTEETRNLALIEEQATYRNAIHDFYFIGRTDNLTDANKRHA EIYKGL
 FKAELFNGKVLKQLGTVTTEHENALLRSFDKFTTYFSGFYENRKNVFS AEDISTAIPHRIVQDN
 FPKFKENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLYNQLLGGI
 SREAGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNTLSFILEEFKSDEEVIQSF
 CKYKTLLRNENVLETA EALFNELNSIDLTHIFISHKKLETISSALCDHWDTLRNALYERRISELTGK
 ITKSAKEKVQRSLKHEDINLQEII SAAGKELSEAFKQKTSEILSHAHAA LDQPLPTTLKKQEEKEIL
 KSQQLDSSLGLYHLLDWF AVDESNEVDPEFSARLTGIKLEMEPSLSFY NKARNYATKKPYSVEKF
 KLNFMPTLASGWDVNKEKNNGAILFVKNGLYYLGIMPKQKG RYKALSFEPT EKTSEGFDKMY
 YDYFPDAAKMIPKCSTQLKAVTAHFQTH TTPILLSNFI EPLEITKEIYDLNNPEKEPKKFQTAYA
 KKTGDQKGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYHISFQR
 IAEKEIMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKTSIKLNGQAELFYR
 PKSRMKRMAHRLGEKMLNKKLDQKTPIDTLYQELYDYVNHRLSHDLSDEARALLPNVITKEV
 SHEIHKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYLKEHPETPIIGIDRGERNLIYITVIDS
 TGKILEQRSLNTIQQFDYQKKLDNREKERVAA RQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQ
 AVVVLENLNF GFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYP AEKVGGLNPNYQLTDQFTS
 FAKMGTQSGFLFYVPAPYTSKIDPLTG FVDPFVWKTIKNHESRKH FLEGFDLHYDVKTGDFIL
 HFKMNRNLSFQRGLPGFMPAWDIVFEKNETQF DAKGTPFIAGKRIVPV IENHRFTGRYRDLYPA
 NELIALLEEKGIVFRDGSNILPKLLEND DSHAIDTMVALIRSVLQMRNSNAATGEDYINSPVRDLN
 GVCFDSRFQNP EWPMADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYIQELRN
 (SEQ ID NO: 110)

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

MESRDHNNPQEGPTSSSGRRRAVEDNHLLIKAVQNEDVDLVQQLLEGGANVNFQEEEGGWT
 PLHNAVQMSREDIVELLRLRHGADPVL RKKNGATPFILAAIAGSVKLLKLFLSKGADVNECDFYGF
 TAFMEA AYGKVKALKFLYKRGANVNLRRKTKEDQERLRKGGATALMDAAEKGHVEVLKILLD
 EMGADV NACDNMGRNALIHALLSSDDSDVEAITHLLLDHGADV NVRGERGKTPILAVEKKHLG
 LVQRLLQEHEIENDTSDSGKTALLLAVELKLKKIAELLCKRGASTDCGDLVMTARRNYDHS LVK

FIG. 22 cont'd

VLLSHGAKEDFHPPAEDWKPQSSHWGAALKDLHRIYRPMIGKLGFFIDEKYKIADTSEGGIYLG
FYEKQEVAVKTFCEGSPRAQREVSCSQSSRENSHLVTFYGSSESHRGHLFVCVTLCEQTLEACL
DVHRGEDVENEDEFARNVLSSIFKAVQELHLSCGYTHQDLQPQNILIDSKKAAHLADFDKSIK
WAGDPQEVKRDLEDLGRVLVYVVKKGSISFEDLKAQSNEEVVQLSPDEETKDLIHRLFHPGEH
VRDCLSDLLGHPFFWTWESRYRTLNRVGNESDIKTRKSESEILRLLQPGPSEHSKSFDKWTTKI
NECVMKKMNKFYEKRGNFYQNTVGDLLKFIRNLGEHIDEKHKMKLKGDPSTLYFQKTFPDV
IYVYTKLQNTYRKHFPQTHSPNKPQCDGAGGASGLASPGC (SEQ ID NO: 111)

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

MMARLLRTSFALLFLGLFGVLGAATISCRNEEGKAVDWFTFYKLPKRQNKESGETGLELYLYDS
TTRSWRKSEQLMNDTKSVLGRTLQQLYEAYASKSNNTAYLIYNDGVPKPVNYSRKYGHTKGLL
LWNRVQGFVLIHSIPQFPPIPEEGYDYPPTGRRNGQSGICITFKYNQYEAIDSQLLVCNPNVYS
CSIPATFHQELIHMPQLCTRASSSEIPGRLLTTLQSAQGQKFLHFAKSDSFLDDIFAAWMAQRLK
THLLTETWQRKRQELPSNCSLPYHVYNIKAIKLSRHSYFSSYQDHAKWCISQKGTKNRWTCIG
DLNRSPHQAFRSGGFICTQNWQIYQAFQGLVLYYESCK (SEQ ID NO: 112)

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

MEQTVLVPPGPDSFNFFTTRESLAAIERRIAEEKAKNPCKPDKKDDENGPKPNSDLEAGKNLPFI
YGDIPPEMVSEPLEDLDPPYINKKTFIVLNKGKAIFRFSATSALYILTPFNPLRKIAIKILVHSLFSML
IMCTILTNCVFMMSNPDPWTKNVEYFTFTGIYTFESLIKIIARGFCLEDFTLRDPWNWLDFTVIT
FAYVTEFVDLGNVSALRTRFVLRALKTISVIPGLKTIVGALIQSVKKLSDVMILTVFCLSVFALIGLQ
LFMGNLRNKCQWPPPTNASLEEHSIEKNITVNYNGTLINETVFEFDWKSIIQDSRYHYFLEGFLD
ALLCGNSSDAGQCPEGYMCVKAGRNPNGYTSFDTFSWAFLSLFRLMTQDFWENLYQLTLRA
AGKTYMIFVFLVFLGSFYLINLILAVVAMAYEEQNQATLEEAQEKEAEFQQMIEQLKKQEEAAQ
QAATATASEHSREPSAAGRLSDSSSEASKLSSSKAKERRNRRKKRKQKEQSGGEEKDEDEFEQ
KSESEDSIRRKGFRRFSIEGNRLTYEKRYSSPHQSLLSIRGSLFSRRNSRSTLSFSFRGRAKDVG
SENDFADDEHSTFEDNESRRDSLFPVRRHGERRNSNLSQTSRSSRMLAVFPANGKMHSTVD
CNGVVS LVGGPSVPTSPVGQLLPEVIIDKPATDDNGTTTETEMRKRSSSFHVSMDFLDPSQ
RQRAMSIASILTNTVEELESRQKCPWCWYKFSNIFLIWDCSPYWLKVHVNLVMDPFVDLA
ITICIVLNTLFMAMEHYPMTHFNNVLTGVNLVFTGIFTAEMFLKIIAMPDYYYFQEGWNIFDGFIV
TSLSLVELGLANVEGLSVLRSFRLLRVFKLAKSWPTLNMLIKIIGNSVGALGNLTLVLAIVFIFAVVG
MQLFGKSYKDCVCKIASDCQLPRWHMNDFFHSFLIVFRVLCGEWIETMWDCMEVAGQAMCLT
VFMMVMVIGNLVVLNLFALLSSFSADNLAATDDDNEMNNLQIAVDRMHKGVAYVVKRIYEFI
QQSFIRKQKILDEIKPLDDLNNKKDSCMSNHTAEIGKDLDYLDVNGTTSGIGTGSSVEKYIDES
DYMSFINNPSLTVTVPIAVGESDFENLNTEDFSSESDEESKEKLNESSSSSEGSTVDIGAPVEE
QPVVEPEETLEPEACFTEGCVQRFKCCQINVEEGRGKQWWNLRRTCFRIVEHNWFETFIVFMI
LLSSGALAFEDIYIDQRKTIKTMLEYADKVFTYIFILEMLLKWWAYGYQTYFTNAWCWLDLIVDV
SLVSLTANALGYSELGAIKSLRTLRLRPLRALSREFGMRVVVNALLGAIPSIMNVLLVCLIFWLIF
SIMGVNLFAGKFYHCINTTTGDRFDIEDVNNHTDCLKLIERNETARWKNVKVNFNDVGFYLSL
LQVATFKGWM DIMYAAVDSRNVELQPKYEESLYMYLYFVIFIFGSFFTLNLFIGVVIDNFNQKK
KFGGQDIFMTEEQKKYYNAMKKLGSKPKQPIPRPGNKFGQGMVDFVTRQVDFDISIMILICLNM
VTMMVETDDQSEYVTTILSRINLVFIVLFTGECVLKLISLRHYFTIGWNIFDFVVLISIVGMFLAE
LIEKYFVSPTLFRVIRLARIGRILRLIKGAKGIRTLLFALMMSLPALFNIGLLFLVMFIYAIFGMSNF
AYVKREVGIDDMFNFETFGNSMICLFQITTSAGWDGLLAPILNSKPPDCDPNKVNPSSSVKGDC
GNPSVGIFFFVSYIIISFLVVNMYIAVILENFSVATEESAEPLEDDFEMFYEVWEKFDPDATQF

FIG. 22 cont'd

MEFEKLSQFAAALEPPLNLPQPNKLQLIAMDLPMVSGDRIHCLDILFAFTKRVLGESGEMDALRI
QMEERFMASNP SKVSYQPITTTTLKRKQEEVSAVIIQRAYRRHLLKRTVKQASFTYNKNKIKGGA
NLLIKEDMIIDRINENSITEKTDLTMTAAACPPSYDRVTKPIVEKHEQEKGDEKAKGK (SEQ ID
NO: 113)

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

MVQKSRNGGVYPGPSGEKKLVGFVGLDPGAPDSTRDGALLIAGSEAPKRGSI LSKPRAGGA
GAGKPPKRNAFYRKLNFLYNVLERPRGWAFIYHAYVFLLVFSCVLVSFSTIKEYEKSSEGAL
YILEIVTIVVFGVEYFVRIWAAGCCCRYRGWRGRLKFARKPFCVIDIMVLIAVLAAGSQGNVF
ATSALRSLRFLQILRMIRMDRRGGTWKLLGSVVYAHSKELVTAWYIGFLCLILASFLVYLAEKGE
NDHFDTYADALWWGLITLTIGYGD KYPQTWNGRLLAATFTLIGVSFFALPAGILGSGFALKVQE
QHRQKHFEKRRNPAAGLIQSAWRFYATNLSRTDLHSTWQYYERTVTVPMYSSQTQTYGASRL
IPPLNQLELLRN LKSKSGLAFRKDPPPEPSPSKGSPCRGPLCGCCPGRSSQKVSLKDRVFSSP
RGVAAKGKGSPPQAQT VRRSPSADQSLEDSPSKVPKSWSF GDRSRARQA FRIKGAASRQNSE
EASLPGEDIVDDKSCPCFEVTEDLTPGLKVSIRAVCVMRFLVSKRKFKESLRPYDVM DVIEQYS
AGHLDMLSRIKSLQSRVDQIVGRGPAITDKDRTKGPAEAE LPEDP SMMGR LGKVEKQVLSMEK
KLDLFVNIYMQRMGIPPTETEAYFGAKEPEPAPPYHSPEDSREHVDRHGCIVKIVRSSSSTGQK
NFSAPPAAPPVQCPPSTSWQPQSHPRQGHGTSPVGDHGS LVRIPPPPAHERSL SAYGGGNR
ASMEFLRQEDTPGCRPPEGNLRDSDTSISIPSDHEELERSFSGFSISQSKENLDALNSCYAAV
APCAKVRPYIAEGESD TDSDLCTPCGPPPRSATGEGPFGDVGWAGPRK (SEQ ID NO: 114)

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

MVNENTRMYIPEENHQGSNYGSPRPAHANMNANAAAGLAPEHIPTPGAALSWQAAIDAARQA
KLMGSAGNATISTVSSTQRKRQQYGKPKKQGSTTATRPPRALLCLTLKNPIRRACISIVEWKPF
EIIILLTIFANCVALAIYIPFPEDDSNATNSNLERVEYLFLLIIFTVEAFLKVIAYGLLFHPNAYLRNGW
NLLDFIIVVVG LFS AILEQATKADGANALGGKGAGFDVKALRAFRVLRPLRLVSGVPSLQVVLNSI
IKAMVPLLHIALLVLFVIIYAIIGLELFMGKMHKTCYNQEG IADVPAEDDPSPCALETGHGRQCQN
GTVCKPGWDGPKHGITNFDNFAFAMLT VFQCITMEGWTDVLYWVND AVGRDWPWIYFVTLIII
GSFFVLNLVLGVLSGEFSKEREKAKARGDFQKLREKQQLEEDLKGYLDWITQAEDIDPENED E
GMDEEKPRNMSMPTSETESVNTENVAGGDIEGENCGARLAHRISKSKFSRYWRRWNRFCRR
KCRAAVKSNV FYWL VIFLVFLNTLTIASEHYNQPNWLTEVQDTANKALLALFTAEMLLKMYSLGL
QAYFVSLFNRFDCFVVC GGILETILVETKIMSPLGISVLRCVRLLRIFKITRYWNSLSNLVASLLNS
VRSIASLLLLLFLFIIF SLLGMQLFGGKFNFDEMQTRRSTFDNFPQSLLTVFQILTGEDWNSVMY
DGIMAYGGPSFPGMLVCIYFIILFICGNYILLNVFLAIAVDNLADAESL TSAQKEEEEEKERKKLAR
TASPEKKQELVEKPAVGESKEEKIELKSITADGESPPATKINMDDLQPNENEDKSPYPNPETT G
EEDEEEPEMPVGPRPRPLSELHLKEKAVPMPEASAFFIFSSNNRFR LQCHRIVNDTIFTNLILFFI
LLSSISLAAEDPVQHTSFRNHILFYFDIVFTTIFTIEIALKILGNADYVFTSIFTLEIILKMTAYGAFLH
KGSFCRNYFNILDLLVSVSLISFGIQSSAINVVKILRVLRVLRPLRAINRAKGLKHVVQCVFVAIR
TIGNIVIVTTLLQFMFACIGVQLFKGKLYTCS DSSKQTEAECKGN YITYKDGEVDHP IIQPRSWEN
SKFDFDNVLAAMMALFTVSTFEGWPELLYRSIDSHTEDKGPIYNRYVEISIFFIYIIIIAFFMMNIFV
GFVIVTFQEQQEQEYKNCELDKNQRQCVEYALKARPLRRYIPKNQH QYKVWYVNVNSTYFEYL
MFVLILLNTICLAMQHYGQSCLFKIAMNILNMLFTGLFTVEMILKLI AFKPKGYFSDPWNVDFDLIV
IGSIIDVILSETNH YFCDAWNTFDALIVVGSIVDIAITEVNPAEHTQCSPSMNAEENSRSISITFFRLF

FIG. 22 cont'd

RVMRLVKLLSRGEGIRTLTWTFIKSFQALPYVALLIVMLFFIYAVIGMQVFGKIALNDTTEINRNNN
FQTFPQAVLLLFRCATGEAWQDIMLACMPGKKCAPESEPSNSTEGETPCGSSFAVFYFISFYM
LCAFLIINLFVAVIMDNFDYLTRDWSILGPHHLDEFKRIWAEYDPEAKGRIKHLDVVTLRRIQPPL
GFGKLCPHRVACKRLVSMNMPLNSDGTVMFNATLFALVRTALRIKTEGNLEQANEELRAIIKKI
WKRTSMKLLDQVVPPAGDDEVTVGKFYATFLIQEYFRKFKKRKEQGLVGKPSQRNALSLQAGL
RTLHDIGPEIRRAISGDLTAEELDKAMKEAVSAASEDDIFRRAGGLFGNHVSYYQSDGRSAFP
QTFTTQRPLHINKAGSSQGDTESPSHEKLVDSTFTPSSYSSTGSNANINNANNTALGRLPRPA
GYPSTVSTVEGHGPPLSPAIRVQEVAWKLSSNRERHVPMCEDLELRRDSGSAGTQAHCLLLR
KANPSRCHSRESQAAMAGQEETSQDETYEVKMNHDTACSEPSLLSTEMLSYQDDENRQLTL
PEEDKRDIRQSPKRGFLRSASLGRRASFHLECLKRQKDRGGDISQKTVLPLHLVHHQALAVAG
LSPLLQRSHSPASFPRPFATPPATPGSRGWPPQPVPTLRLEGVESSEKLNSSFPSIHCGSWAE
TTPGGGGSSAARRVRPVSLMVPSQAGAPGRQFHGSASSLVEAVLISEGLGQFAQDPKFIEVTT
QELADACDMTIEEMESAADNILSGGAPQSPNGALLPFVNCRDAGQDRAGGEEDAGCVRARGR
PSEELQDSRVYVSSL (SEQ ID NO: 115)