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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 November 2023 (16.11.2023)

WIPO | PCT

(10) International Publication Number
WO 2023/220287 A1

(51) International Patent Classification:
A61P 25/28 (2006.01) *C07K 14/005* (2006.01)
C07K 7/06 (2006.01) *C12N 15/86* (2006.01)

(21) International Application Number:
PCT/US2023/021905

(22) International Filing Date:
11 May 2023 (11.05.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/340,823 11 May 2022 (11.05.2022) US
63/342,324 16 May 2022 (16.05.2022) US

(71) Applicant: **THE CHILDREN'S HOSPITAL OF PHILADELPHIA** [US/US]; 3401 Civic Center Boulevard, Philadelphia, Pennsylvania 19146 (US).

(72) Inventors: **DAVIDSON, Beverly**; c/o The Children's Hospital of Philadelphia, 2716 South Street, 15th Floor, Philadelphia, Pennsylvania 19146 (US). **TECEDOR, Luis**; c/o The Children's Hospital of Philadelphia, 2716 South Street, 15th Floor, Philadelphia, Pennsylvania 19146 (US). **RANUM, Paul**; c/o The Children's Hospital of Philadelphia, 2716 South Street, 15th Floor, Philadelphia, Pennsylvania 19146 (US). **LEIB, David**; c/o The Children's Hospital of Philadelphia, 2716 South Street, 15th Floor, Philadelphia, Pennsylvania 19146 (US). **KEISER, Megan**; c/o The Children's Hospital of Philadelphia, 2716 South Street, 15th Floor, Philadelphia, Pennsylvania 19146 (US). **CHEN, Yonghong**; c/o The Children's Hospital of Philadelphia, 2716 South Street, 15th Floor, Philadelphia, Pennsylvania 19146 (US).

(74) Agent: **SCHNEPP, Amanda S. J.**; Parker Highlander PLLC, 1120 S. Capital of Texas Highway, Bldg. One, Suite 200, Austin, Texas 78746 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

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

(54) Title: ADENO-ASSOCIATED VIRAL VECTORS FOR TARGETING DEEP BRAIN STRUCTURES

(57) Abstract: Provided herein are targeting peptides and vectors containing a sequence that encodes the targeting peptides that deliver agents to specific substructures in the brain. Specifically, the targeting peptide is a component of a modified, sequence-specified adeno-associated virus (AAV) capsid protein further wherein the brain substructure may be the globus pallidus, putamen, internal capsule, caudate, claustrum, substantial nigra, motor cortex, insula,.temporal cortex, thalamus, hippocampus, subiculum, and deep cerebellar nuclei.

WO 2023/220287 A1

ADENO-ASSOCIATED VIRAL VECTORS FOR TARGETING DEEP BRAIN STRUCTURES

REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of priority to United States Provisional Application No. 63/340,823, filed May 11, 2022, and United States Provisional Application No. 63/342,324, filed May 16, 2022, the entire contents of each of which are hereby incorporated by reference.

REFERENCE TO A SEQUENCE LISTING

[0002] This application contains a Sequence Listing XML, which has been submitted electronically and is hereby incorporated by reference in its entirety. Said Sequence Listing XML, created on May 9, 2023, is named CHOPP0054WO_ST26.xml and is 32,221 bytes in size.

BACKGROUND

1. Field

[0003] The present disclosure relates generally to the fields of medicine and virology. More particularly, it concerns compositions and methods for delivery of molecular therapeutics to deep brain structures.

2. Description of Related Art

[0004] Different strategies have been developed to generate AAV vector variants including rational design and directed evolution. The rational design approach utilizes knowledge of AAV capsids to make targeted changes to the capsid to alter transduction efficiency or specificity, such as tyrosine mutations on the capsid surface for increasing transduction efficiency. However, no AAV variants target deep brain structures specifically or efficiently. Diseases including Huntington's Disease (HD) and Parkinson's Disease (PD) mainly affect deep brain structures that existing AAV variants are unable to target effectively. As such, AAV variants that are able to target deep brain structures are needed.

SUMMARY

[0005] Provided herein are viral vectors each comprising a modified capsid, wherein the modified capsid comprises at least one amino acid sequence that targets the viral vector to a distinct brain structure.

[0006] In one embodiment, provided are modified adeno-associated virus (AAV) capsid proteins comprising a targeting peptide that targets a viral vector comprising the modified AAV capsid protein to a distinct organ or brain structure, and the targeting peptide is three to ten amino acids in length. In some aspects, the modified AAV capsid proteins are modified AAV1 capsid proteins, modified AAV2 capsid proteins, or modified AAV9 capsid proteins.

[0007] In some aspects, the modified AAV capsid proteins are derived from an AAV1 capsid protein (see SEQ ID NO: 1), and the targeting peptide is inserted after residue 590 of the AAV1 capsid protein. In some aspects, the targeting peptide is flanked by linker sequences, and the linker sequences on each side of the targeting peptides are two or three amino acids long. In some aspects, the linker sequences are SSA on the N-terminal side of the targeting peptide and AS on the C-terminal side of the targeting peptide. In some aspects, the modified AAV1 capsid proteins have a sequence at least 95% identical to SEQ ID NO: 4.

[0008] In some aspects, the modified AAV capsid proteins are derived from an AAV2 capsid protein (see SEQ ID NO: 2), and the targeting peptide is inserted after residue 587 of the AAV2 capsid protein. In some aspects, the targeting peptide is flanked by linker sequences, and the linker sequences on each side of the targeting peptides are two or three amino acids long. In some aspects, the linker sequences are AAA on the N-terminal side of the targeting peptide and AA on the C-terminal side of the targeting peptide. In some aspects, the modified AAV2 capsid proteins have a sequence at least 95% identical to SEQ ID NO: 5.

[0009] In some aspects, the modified AAV capsid proteins are derived from an AAV9 capsid protein (see SEQ ID NO: 3), and the targeting peptide is inserted after residue 588 of the AAV9 capsid protein. In some aspects, the targeting peptide is flanked by linker sequences, and the linker sequences on each side of the targeting peptides are two or three amino acids long. In some aspects, the linker sequences are AAA on the N-terminal side of the targeting peptide and AS on the C-terminal side of the targeting peptide. In some aspects, the modified AAV9 capsid proteins have a sequence at least 95% identical to SEQ ID NO: 6.

[0010] In some aspects, the target peptide comprises a sequence up to ten amino acids in length having therein an amino acid sequence selected from the group consisting of SEQ ID NOs: 7-13 or an amino acid sequence at least 85% identical to an amino acid sequence selected from the group consisting of SEQ ID NOs: 8, 10, 12, 14, 16, and 18. In some aspects, the targeting peptide is seven amino acids in length. In some aspects, the AAV capsid protein has a sequence of any one of SEQ ID NOs: 19-24.

[0011] In some aspects, the distinct brain structure is a deep brain structure, such as the globus pallidus, putamen, internal capsule, caudate, claustrum, substantia nigra, thalamus, hippocampus, subiculum, and deep cerebellar nuclei. In some aspects, the distinct brain structure is the motor cortex, or temporal cortex. In some aspects, the distinct brain structure is layer V/VI projection neurons in the motor cortex and premotor cortex.

[0012] In one embodiment, provided herein are nucleic acids comprising a sequence encoding the modified capsid protein of any one of the present embodiments.

[0013] In one embodiment, provided herein are recombinant adeno-associated viruses (rAAV) comprising the modified capsid protein of any one of the present embodiments. In some aspects, combinations of rAAVs are provided. In some aspects, the rAAVs do not transduce liver cells.

[0014] In one embodiment, provided herein are viral vectors comprising a nucleic acid encoding the modified capsid protein of any one of the present embodiments. In some aspects, the viral vectors further comprise a nucleic acid sequence encoding a nucleic acid of interest. In some aspects, the nucleic acid of interest is a therapeutic agent. In some aspects, the therapeutic agent is an enzyme or an RNAi molecule.

[0015] In one embodiment, provided herein are cells comprising the viral vector of any one of the present embodiments. In some aspects, the cell is a mammalian cell, such as a human cell. In some aspects, the cell is in vitro or in vivo.

[0016] In one embodiment, provided herein are pharmaceutical compositions comprising the viral vector of the present embodiments and a pharmaceutically acceptable carrier.

[0017] In one embodiment, provided herein are methods to deliver an agent to a distinct brain structure of a subject, comprising administering an AAV virus encoding the agent directly to the globus pallidus. In some aspects, the AAV virus is a virus of the present embodiments.

[0018] In one embodiment, provided herein are methods to deliver an agent to a distinct brain structure of a subject, comprising administering a virus of the present embodiments to the subject. In some aspects, the distinct brain structure is the globus pallidus, putamen, internal capsule, caudate, claustrum, substantial nigra, motor cortex, insula, temporal cortex, thalamus, hippocampus, subiculum, and deep cerebellar nuclei. In some aspects, the distinct brain structure is layer V/VI projection neurons in the motor cortex and premotor cortex.

[0019] In some aspects, the agent is an siRNA, shRNA, miRNA, non-coding RNA, lncRNA, therapeutic protein, or CRISPR system. In some aspects, the administration is to the central nervous system. In some aspects, the administration is to a cisterna magna, an intraventricular space, an ependyma, a brain ventricle, a subarachnoid space, and/or an intrathecal space. In some aspects, the brain ventricle is the rostral lateral ventricle, and/or the caudal lateral ventricle, and/or the right lateral ventricle, and/or the left lateral ventricle, and/or the right rostral lateral ventricle, and/or the left rostral lateral ventricle, and/or the right caudal lateral ventricle, and/or the left caudal lateral ventricle. In some aspects, the administration is systemic.

[0020] In some aspects, a plurality of viral particles is administered. In some aspects, the virus is administered at a dose of about 1×10^{6} to about 1×10^{18} vector genomes per kilogram (vg/kg). In some aspects, the virus is administered at a dose from about $1x10^{7}$-$1x10^{17}$, about $1x10^{8}$-$1x10^{16}$, about $1x10^{9}$-$1x10^{15}$, about $1x10^{10}$-$1x10^{14}$, about $1x10^{10}$-$1x10^{13}$, about $1x10^{10}$-$1x10^{13}$, about $1x10^{10}$-$1x10^{11}$, about $1x10^{11}$-$1x10^{12}$, about $1x10^{12}$-$x10^{13}$, or about $1x10^{13}$-$1X10^{14}$ vg/kg of the patient. In some aspects, the subject is human.

[0021] In one embodiment, provided herein are methods of treating a disease in a mammal comprising administering the virus of the present embodiments to the mammal. In some aspects, the disease is a neurodegenerative disease. In some aspects, the neurodegenerative disease is Huntington's disease, amyotrophic lateral sclerosis (ALS), hereditary spastic hemiplegia, primary lateral sclerosis, spinal muscular atrophy, Kennedy's disease, Alzheimer's disease, a polyglutamine repeat disease, or Parkinson's disease. In some aspects, the mammal is human.

[0022] Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0023] The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present invention. The invention may be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

[0024] FIG. 1. Fluorescent In Situ Hybridization (FISH) images, captured by confocal microscopy, for each specific fluorophore to localize mRNA within cells of the Globus Pallidus (site of injection). Positive signal is localized to the soma and nucleus of these cells. Most cells are positive for all three capsid variants.

[0025] FIG. 2. Epifluorescence images to localize protein from each specific fluorophore within cells of the Globus Pallidus (site of injection). Positive signal is localized to both the soma and to the projections of these cells. Specifically, although mRuby3 mRNA was present in the soma of globus pallidus neurons by FISH, there is little to no positive signal for protein in this slab of tissue.

[0026] FIG. 3. Epifluorescence images, from different Globus Pallidus regions than the images in FIG. 2, localizing protein from each specific fluorophore within cells of the caudal Globus Pallidus (site of injection). Positive signal is localized to both the soma and to the projections of these cells. In this tissue section, mRuby signal is present.

[0027] FIG. 4. Fluorescent In Situ Hybridization (FISH) images, captured by confocal microscopy, for each specific fluorophore to localize mRNA within cells of the putamen. Positive mRNA signal is localized to the soma and nuclei of these cells. The capsid variant expressing mNeonGreen (mNG) shows the highest transduction of cells in the putamen.

[0028] FIG. 5. Epifluorescence images to localize protein from each specific fluorophore within cells of the putamen. Positive signal is localized to both the soma and to the projections of these cells. Similar to mRNA by FISH, the novel capsid variant expressing mNeonGreen (mNG) shows the most positive signal in both cells and projections.

[0029] FIG. 6. Epifluorescence images to localize protein from each specific fluorophore within the internal capsule. There is a striking separation of positive signal of mTFP1 and NeonGreen fluorescence.

[0030] FIG. 7. Fluorescent In Situ Hybridization (FISH) images, captured by confocal microscopy, for each specific fluorophore to localize mRNA within cells of the caudate. Positive mRNA signal is localized to the soma of these cells.

[0031] FIG. 8. Epifluorescence images to localize protein from each specific fluorophore within cells of the caudate. Positive signal is localized to both the soma and to the projections of these cells.

[0032] FIG. 9. Fluorescent In Situ Hybridization (FISH) images, captured by confocal microscopy, for each specific fluorophore to localize mRNA within cells of the claustrum. Positive mRNA signal is localized to the soma of these cells.

[0033] FIG. 10. Epifluorescence images to localize protein from each specific fluorophore within the claustrum. Again, positive signal of fluorescence is seen in both cell soma and cell projections.

[0034] FIG. 11. Fluorescent In Situ Hybridization (FISH) images, captured by confocal microscopy, for each specific fluorophore to localize mRNA within cells of the substantia nigra. Positive mRNA signal is localized to the soma and nucleus of these cells.

[0035] FIG. 12. Fluorescent In Situ Hybridization (FISH) images, captured by confocal microscopy, for each specific fluorophore to localize mRNA within cells of the substantia nigra, but from different areas with respect to the images in FIG. 11. Positive mRNA signal is localized to the soma and nucleus of these cells.

[0036] FIG. 13. Epifluorescence images to localize protein from each specific fluorophore within the substantia nigra. Positive signal of fluorescence is seen in both cell soma and cell projections.

[0037] FIG. 14. Epifluorescence images to localize protein from each specific fluorophore within the substantia nigra, but from different areas with respect to the images in FIG. 13. Positive signal of fluorescence is seen in both cell soma and cell projections.

[0038] FIG. 15. Fluorescent In Situ Hybridization (FISH) images, captured by confocal microscopy, for each specific fluorophore to localize mRNA within cells of the motor cortex. Positive mRNA signal is localized to the soma and nucleus of these cells.

[0039] FIG. 16. Epifluorescence images to localize protein from each specific fluorophore within cells of the motor cortex. Positive signal is localized to both the soma and to the projections of these cells. Strikingly, mRNA signal was present for both mTFP1 and mNeonGreen by FISH; however, it is mTFP1 that is predominantly expressed as shown by Epifluorescence.

[0040] FIG. 17. In the cingular cortex, a few positive cells for the novel capsid mRNA were detected by FISH. However, wide spread mTFP1 positive cells were detected by epifluorescence microscopy.

[0041] FIG. 18. Fluorescent In Situ Hybridization (FISH) images, captured by confocal microscopy, for each specific fluorophore to localize mRNA within cells of the insula of the temporal cortex. Positive mRNA signal is localized to the soma (cytoplasm and nucleus) of these cells.

[0042] FIG. 19. Epifluorescence images to localize protein from each specific fluorophore within cells of the insula of the temporal cortex. Positive signal is localized to both the soma and to the projections of these cells.

[0043] FIG. 20. Fluorescent In Situ Hybridization (FISH) images, captured by confocal microscopy, for each specific fluorophore to localize mRNA within cells of the thalamus. Positive mRNA signal is localized to the soma (cytoplasm and nucleus) of these cells.

[0044] FIG. 21. Epifluorescence images to localize protein from each specific fluorophore within cells of the thalamus. Positive signal is localized to both the soma and to the projections of these cells.

[0045] FIGS. 22A-D. Epifluorescence images of top three novel capsids delivered to the globus pallidus of rhesus macaque. Expression of all three is high in the putamen. A) Image showing expression of AAV1.THTD.mTFP1; B) Image showing expression of AAV2.SGGR.mNeonGreen; C) Image showing expression of AAV2.SKGQ.mRuby3; D) Merged image of A-C.

[0046] FIGS. 23A-D. Epifluorescence images of top three novel capsids delivered to the globus pallidus of rhesus macaque. Expression of AAV1.THTD.mTFP1 is extremely high in the motor cortex. A) Image showing expression of AAV1.THTD.mTFP1; B) Image showing expression of AAV2.SGGR.mNeonGreen; C) Image showing expression of AAV2.SKGQ.mRuby3; D) Merged image of A-C.

[0047] FIGS. 24A-D. Epifluorescence images of top three novel capsids delivered to the globus pallidus of rhesus macaque. Expression of all three is localized in the Hippocampus (CA1) and adjacent subiculum. A) Image showing expression of AAV1.THTD.mTFP1; B) Image showing expression of AAV2.SGGR.mNeonGreen; C) Image showing expression of AAV2.SKGQ.mRuby3; D) Merged image of A-C.

[0048] FIG. 25. Fluorescent In Situ Hybridization (FISH) images, captured by confocal microscopy, for each specific fluorophore to localize mRNA within cells of the Globus Pallidus (site of injection) of a second rhesus macaque. Positive signal is localized to the soma and nucleus of these cells.

[0049] FIG. 26. Epifluorescence images to localize protein from each specific fluorophore within cells of the Globus Pallidus (site of injection) of a second rhesus macaque. Positive signal is localized to both the soma and to the projections of these cells.

[0050] FIG. 27. Whole mounts of brains and spinal cords from mice injected with virus to the internal capsule/medial globus pallidus. Mouse numbers are on the left side, their ID numbers are color coded to reflect which hemisphere received which fluorophore-expressing capsid.

[0051] FIG. 28. Whole mounts of brains and spinal cords from mice injected with virus to the thalamus. Mouse numbers are on the left side, their ID numbers are color coded to reflect which hemisphere received which fluorophore-expressing capsid.

[0052] FIG. 29. Whole mounts of brains and spinal cords from mice injected with virus to the deep cerebellar nuclei. Mouse numbers are on the left side, their ID numbers are color coded to reflect which hemisphere received which fluorophore-expressing capsid.

[0053] FIG. 30. Four coronal brain slices from a mouse injected with AAV2.SGGR.mNeonGreen to the left Internal Capsule and Medial Globus Pallidus and AAV2.SKGO.mRuby3 to the right.

[0054] FIG. 31. Three coronal brain slices from a mouse injected with AAV2. SKGO.mRuby3 to the left Internal Capsule and Medial Globus Pallidus and AAV1.THTD.mTFP1 to the right.

[0055] FIG. 32. Four coronal brain slices from a mouse injected with AAV2.SGGR.mNeonGreen to the left thalamus and AAV2.SKGO.mRuby3 to the right. A single sagittal section of the cerebellum and hindbrain.

[0056] FIG. 33. Four coronal brain slices from a mouse injected with AAV2.SKGO.mRuby3 to the left thalamus and AAV1.THTD.mTFP1 to the right. Sagittal sections of the cerebellum and brainstem show neurons in the deep cerebellar nuclei positive for AAV1.THTD.mTFP1.

[0057] FIG. 34. Sagittal slices of the cerebellum injected with AAV1.THTD.mTFP1 to the left Deep Cerebellar Nuclei, and AAV2.SGGR.mNeonGreen to the right.

[0058] FIG. 35. Sagittal slice of the lateral cerebrum injected with AAV1.THTD.mTFP1 to the left Deep Cerebellar Nuclei, and AAV2.SGGR.mNeonGreen to the right.

[0059] FIG. 36. Sagittal slice of the entire brain injected with AAV1.THTD.mTFP1 to the left Deep Cerebellar Nuclei, and AAV2.SGGR.mNeonGreen to the right. There are neurons and projections throughout the cerebrum positive for AAV1.THTD.mTFP1.

[0060] FIG. 37. Sagittal slice of the entire brain injected with AAV2.SKGO.mRuby3 to the left Deep Cerebellar Nuclei, and AAV1.THTD.mTFP1 to the right. There are neurons and projections throughout the cerebrum positive for AAV1.THTD.mTFP1.

[0061] FIGS. 38A-38C. Heatmaps representations of UMI enrichment of AAV capsid variants across sampled brain regions. (FIG. 38A) Amplicon seq spanning the area of the AAV peptide insertion allowed us to collect unique molecular identifier (UMI) counts of delivered capsid variants across a number of sampled tissues. Amplicon in A were generated from capsid DNA recovered from transduced tissues. The % abundance of each capsid variant is represented for each indicated tissue. The capsids carried forward to fluorescence validation are indicated. (FIG. 38B) The same DNA amplicon-Seq UMI counts shown in A were normalized by calculating enrichment over input. The most highly represented capsid candidates are those that show the highest enrichment relative to their abundance in the input capsid pool. (FIG. 38C) The % abundance of UMI counts obtained from RNA by amplicon-Seq.

[0062] FIG. 39. Epifluorescence images to localize AAV1.THTD.mTFP1 to a large number of cortical layer V/VI projection neurons in the motor cortex and premotor cortex of rhesus macaque.

[0063] FIG. 40. AAV gDNA and protein expression were below the threshold of detection in liver of rhesus macaque.

DETAILED DESCRIPTION

[0064] Provided herein are viral vectors each comprising a modified capsid, wherein the modified capsid comprises at least one amino acid sequence that targets the viral vector to a deep brain structure. In certain embodiments, the brain structure is the globus pallidus, putamen, internal capsule, caudate, claustrum, substantial nigra, motor cortex, insula, temporal cortex, thalamus, hippocampus, subiculum, and deep cerebellar nuclei. In certain embodiments, the AAV is AAV1, AAV2, or AAV9. An exemplary wildtype reference AAV1 capsid protein sequence is provided in SEQ ID NO: 1. An exemplary wildtype reference AAV2 capsid protein sequence is provided in SEQ ID NO: 2. An exemplary wildtype reference AAV9 capsid protein sequence is provided in SEQ ID NO: 3. In certain aspects, the targeting peptide is inserted at position 590 of the AAV1 capsid, position 587 of the AAV2 capsid, or position 588 of the AAV9 capsid. An exemplary modified AAV1 capsid protein sequence is provided in SEQ ID NO: 4, which shows the targeting peptide insertion after position 590 as $SSAX_7AS$, where the leading SSA and the trailing AS are linker sequences and X_7 represents the targeting peptide. An exemplary modified AAV2 capsid protein sequence is provided in SEQ ID NO: 5,

which shows the targeting peptide insertion after position 587 as $AAAX_7AA$, where the leading AAA and the trailing AA are linker sequences and X_7 represents the targeting peptide. An exemplary modified AAV9 capsid protein sequence is provided in SEQ ID NO: 6, which shows the targeting peptide insertion after position 588 as $AAAX_7AS$, where the leading AAA and the trailing AS are linker sequences and X_7 represents the targeting peptide.

I. Adeno-Associated Virus (AAV) Vectors

[0065] Adeno-associated virus (AAV) is a small nonpathogenic virus of the parvoviridae family. To date, numerous serologically distinct AAVs have been identified, and more than a dozen have been isolated from humans or primates. AAV is distinct from other members of this family by its dependence upon a helper virus for replication.

[0066] AAV genomes can exist in an extrachromosomal state without integrating into host cellular genomes; possess a broad host range; transduce both dividing and non-dividing cells *in vitro* and *in vivo* and maintain high levels of expression of the transduced genes. AAV viral particles are heat stable; resistant to solvents, detergents, changes in pH, and temperature; and can be column purified and/or concentrated on CsCl gradients or by other means. The AAV genome comprises a single-stranded deoxyribonucleic acid (ssDNA), either positive- or negative-sensed. The approximately 4.7 kb genome of AAV consists of one segment of single stranded DNA of either plus or minus polarity. The ends of the genome are short inverted terminal repeats (ITRs) that can fold into hairpin structures and serve as the origin of viral DNA replication.

[0067] An AAV "genome" refers to a recombinant nucleic acid sequence that is ultimately packaged or encapsulated to form an AAV particle. An AAV particle often comprises an AAV genome packaged with AAV capsid proteins. In cases where recombinant plasmids are used to construct or manufacture recombinant vectors, the AAV vector genome does not include the portion of the "plasmid" that does not correspond to the vector genome sequence of the recombinant plasmid. This non-vector genome portion of the recombinant plasmid is referred to as the "plasmid backbone," which is important for cloning and amplification of the plasmid, a process that is needed for plasmid propagation and production, but is not itself packaged or encapsulated into viral particles. Thus, an AAV vector "genome" refers to nucleic acid that is packaged or encapsulated by AAV capsid proteins.

[0068] The AAV virion (particle) is a non-enveloped, icosahedral particle approximately 25 nm in diameter that comprises an AAV capsid. The AAV particle comprises an icosahedral symmetry comprised of three related capsid proteins, VP1, VP2 and VP3, which interact together to form the capsid. The genome of most native AAVs often contain two open reading frames (ORFs), sometimes referred to as a left ORF and a right ORF. The right ORF often encodes the capsid proteins VP1, VP2, and VP3. These proteins are often found in a ratio of 1:1:10 respectively, but may be in varied ratios, and are all derived from the right-hand ORF. The VP1, VP2 and VP3 capsid proteins differ from each other by the use of alternative splicing and an unusual start codon. Deletion analysis has shown that removal or alteration of VP1 which is translated from an alternatively spliced message results in a reduced yield of infectious particles. Mutations within the VP3 coding region result in the failure to produce any single-stranded progeny DNA or infectious particles. In certain embodiments, the genome of an AAV particle encodes one, two or all three VP1, VP2 and VP3 polypeptides.

[0069] The left ORF often encodes the non-structural Rep proteins, Rep 40, Rep 52, Rep 68 and Rep 78, which are involved in regulation of replication and transcription in addition to the production of single-stranded progeny genomes. Two of the Rep proteins have been associated with the preferential integration of AAV genomes into a region of the q arm of human chromosome 19. Rep68/78 have been shown to possess NTP binding activity as well as DNA and RNA helicase activities. Some Rep proteins possess a nuclear localization signal as well as several potential phosphorylation sites. In certain embodiments the genome of an AAV (*e.g.*, an rAAV) encodes some or all of the Rep proteins. In certain embodiments the genome of an AAV (*e.g.*, an rAAV) does not encode the Rep proteins. In certain embodiments one or more of the Rep proteins can be delivered in trans and are therefore not included in an AAV particle comprising a nucleic acid encoding a polypeptide.

[0070] The ends of the AAV genome comprise short inverted terminal repeats (ITR) which have the potential to fold into T-shaped hairpin structures that serve as the origin of viral DNA replication. Accordingly, the genome of an AAV comprises one or more (*e.g.*, a pair of) ITR sequences that flank a single stranded viral DNA genome. The ITR sequences often have a length of about 145 bases each. Within the ITR region, two elements have been described which are believed to be central to the function of the ITR, a GAGC repeat motif and the terminal resolution site (trs). The repeat motif has been shown to bind Rep when the ITR is in either a linear or hairpin conformation. This binding is thought to position Rep68/78 for

cleavage at the trs which occurs in a site- and strand-specific manner. In addition to their role in replication, these two elements appear to be central to viral integration. Contained within the chromosome 19 integration locus is a Rep binding site with an adjacent trs. These elements have been shown to be functional and necessary for locus specific integration.

[0071] The term "recombinant," as a modifier of vector, such as recombinant viral, *e.g.*, lenti- or parvo-virus (*e.g.*, AAV) vectors, as well as a modifier of sequences such as recombinant nucleic acid sequences and polypeptides, means that the compositions have been manipulated (*i.e.*, engineered) in a fashion that generally does not occur in nature. A particular example of a recombinant vector, such as an AAV, retroviral, or lentiviral vector would be where a nucleic acid sequence that is not normally present in the wild-type viral genome is inserted within the viral genome. An example of a recombinant nucleic acid sequence would be where a nucleic acid (*e.g.*, gene) encodes an inhibitory RNA cloned into a vector, with or without 5′, 3′ and/or intron regions that the gene is normally associated within the viral genome. Although the term "recombinant" is not always used herein in reference to vectors, such as viral vectors, as well as sequences such as polynucleotides, "recombinant" forms including nucleic acid sequences, polynucleotides, transgenes, etc. are expressly included in spite of any such omission.

[0072] A recombinant viral "vector" is derived from the wild type genome of a virus by using molecular methods to remove part of the wild type genome from the virus, and replacing with a non-native nucleic acid, such as a nucleic acid sequence. Typically, for example, for AAV, one or both inverted terminal repeat (ITR) sequences of the AAV genome are retained in the recombinant AAV vector. A "recombinant" viral vector (*e.g.*, rAAV) is distinguished from a viral (*e.g.*, AAV) genome, since part of the viral genome has been replaced with a non-native sequence with respect to the viral genomic nucleic acid such a nucleic acid encoding a transactivator or nucleic acid encoding an inhibitory RNA or nucleic acid encoding a therapeutic protein. Incorporation of such non-native nucleic acid sequences therefore defines the viral vector as a "recombinant" vector, which in the case of AAV can be referred to as a "rAAV vector."

[0073] In certain embodiments, an AAV (*e.g.*, a rAAV) comprises two ITRs. In certain embodiments, an AAV (*e.g.*, a rAAV) comprises a pair of ITRs. In certain embodiments, an AAV (*e.g.*, a rAAV) comprises a pair of ITRs that flank (*i.e.*, are at each 5′ and 3′ end) of a nucleic acid sequence that at least encodes a polypeptide having function or activity.

[0074] An AAV vector (*e.g.*, rAAV vector) can be packaged and is referred to herein as an "AAV particle" for subsequent infection (transduction) of a cell, *ex vivo*, *in vitro* or *in vivo*. Where a recombinant AAV vector is encapsulated or packaged into an AAV particle, the particle can also be referred to as a "rAAV particle." In certain embodiments, an AAV particle is a rAAV particle. A rAAV particle often comprises a rAAV vector, or a portion thereof. A rAAV particle can be one or more rAAV particles (*e.g.*, a plurality of AAV particles). rAAV particles typically comprise proteins that encapsulate or package the rAAV vector genome (*e.g.*, capsid proteins). It is noted that reference to a rAAV vector can also be used to reference a rAAV particle.

[0075] Any suitable AAV particle (*e.g.*, rAAV particle) can be used for a method or use herein. A rAAV particle, and/or genome comprised therein, can be derived from any suitable serotype or strain of AAV. A rAAV particle, and/or genome comprised therein, can be derived from two or more serotypes or strains of AAV. Accordingly, a rAAV can comprise proteins and/or nucleic acids, or portions thereof, of any serotype or strain of AAV, wherein the AAV particle is suitable for infection and/or transduction of a mammalian cell. Non-limiting examples of AAV serotypes include AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV-rh74, AAV-rh10 and AAV-2i8.

[0076] In certain embodiments a plurality of rAAV particles comprises particles of, or derived from, the same strain or serotype (or subgroup or variant). In certain embodiments a plurality of rAAV particles comprise a mixture of two or more different rAAV particles (*e.g.*, of different serotypes and/or strains).

[0077] As used herein, the term "serotype" is a distinction used to refer to an AAV having a capsid that is serologically distinct from other AAV serotypes. Serologic distinctiveness is determined on the basis of the lack of cross-reactivity between antibodies to one AAV as compared to another AAV. Such cross-reactivity differences are usually due to differences in capsid protein sequences/antigenic determinants (*e.g.*, due to VP1, VP2, and/or VP3 sequence differences of AAV serotypes). Despite the possibility that AAV variants including capsid variants may not be serologically distinct from a reference AAV or other AAV serotype, they differ by at least one nucleotide or amino acid residue compared to the reference or other AAV serotype.

[0078] In certain embodiments, a rAAV vector based upon a first serotype genome corresponds to the serotype of one or more of the capsid proteins that package the vector. For example, the serotype of one or more AAV nucleic acids (*e.g.*, ITRs) that comprises the AAV vector genome corresponds to the serotype of a capsid that comprises the rAAV particle.

[0079] In certain embodiments, a rAAV vector genome can be based upon an AAV (*e.g.*, AAV2) serotype genome distinct from the serotype of one or more of the AAV capsid proteins that package the vector. For example, a rAAV vector genome can comprise AAV2 derived nucleic acids (*e.g.*, ITRs), whereas at least one or more of the three capsid proteins are derived from a different serotype, *e.g.*, an AAV1, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, Rh10, Rh74 or AAV-2i8 serotype or variant thereof.

[0080] In certain embodiments, a rAAV particle or a vector genome thereof related to a reference serotype has a polynucleotide, polypeptide or subsequence thereof that comprises or consists of a sequence at least 60% or more (*e.g.*, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, etc.) identical to a polynucleotide, polypeptide or subsequence of an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, Rh10, Rh74 or AAV-2i8 particle. In particular embodiments, a rAAV particle or a vector genome thereof related to a reference serotype has a capsid or ITR sequence that comprises or consists of a sequence at least 60% or more (*e.g.*, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, etc.) identical to a capsid or ITR sequence of an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, Rh10, Rh74 or AAV-2i8 serotype.

[0081] In certain embodiments, a method herein comprises use, administration or delivery of an rAAV1, rAAV2, rAAV3, rAAV4, rAAV5, rAAV6, rAAV7, rAAV8, rAAV9, rAAV10, rAAV11, rAAV12, rRh10, rRh74 or rAAV-2i8 particle.

[0082] In certain embodiments, a method herein comprises use, administration or delivery of a rAAV2 particle. In certain embodiments a rAAV2 particle comprises an AAV2 capsid. In certain embodiments a rAAV2 particle comprises one or more capsid proteins (*e.g.*, VP1, VP2 and/or VP3) that are at least 60%, 65%, 70%, 75% or more identical, *e.g.*, 80%, 85%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, etc., up to 100% identical to a corresponding capsid protein of a native or wild-type AAV2 particle. In certain embodiments a rAAV2 particle comprises VP1,

VP2 and VP3 capsid proteins that are at least 75% or more identical, *e.g.*, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, etc., up to 100% identical to a corresponding capsid protein of a native or wild-type AAV2 particle. In certain embodiments, a rAAV2 particle is a variant of a native or wild-type AAV2 particle. In some aspects, one or more capsid proteins of an AAV2 variant have 1, 2, 3, 4, 5, 5-10, 10-15, 15-20 or more amino acid substitutions compared to capsid protein(s) of a native or wild-type AAV2 particle.

[0083] In certain embodiments a rAAV9 particle comprises an AAV9 capsid. In certain embodiments a rAAV9 particle comprises one or more capsid proteins (*e.g.*, VP1, VP2 and/or VP3) that are at least 60%, 65%, 70%, 75% or more identical, *e.g.*, 80%, 85%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, etc., up to 100% identical to a corresponding capsid protein of a native or wild-type AAV9 particle. In certain embodiments a rAAV9 particle comprises VP1, VP2 and VP3 capsid proteins that are at least 75% or more identical, *e.g.*, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, etc., up to 100% identical to a corresponding capsid protein of a native or wild-type AAV9 particle. In certain embodiments, a rAAV9 particle is a variant of a native or wild-type AAV9 particle. In some aspects, one or more capsid proteins of an AAV9 variant have 1, 2, 3, 4, 5, 5-10, 10-15, 15-20 or more amino acid substitutions compared to capsid protein(s) of a native or wild-type AAV9 particle.

[0084] In certain embodiments, a rAAV particle comprises one or two ITRs (*e.g.*, a pair of ITRs) that are at least 75% or more identical, *e.g.*, 80%, 85%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, etc., up to 100% identical to corresponding ITRs of a native or wild-type AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV-rh74, AAV-rh10 or AAV-2i8, as long as they retain one or more desired ITR functions (*e.g.*, ability to form a hairpin, which allows DNA replication; integration of the AAV DNA into a host cell genome; and/or packaging, if desired).

[0085] In certain embodiments, a rAAV2 particle comprises one or two ITRs (*e.g.*, a pair of ITRs) that are at least 75% or more identical, *e.g.*, 80%, 85%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, etc., up to 100% identical to corresponding ITRs of a native or wild-type AAV2

particle, as long as they retain one or more desired ITR functions (*e.g.*, ability to form a hairpin, which allows DNA replication; integration of the AAV DNA into a host cell genome; and/or packaging, if desired).

[0086] In certain embodiments, a rAAV9 particle comprises one or two ITRs (*e.g.*, a pair of ITRs) that are at least 75% or more identical, *e.g.*, 80%, 85%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, etc., up to 100% identical to corresponding ITRs of a native or wild-type AAV2 particle, as long as they retain one or more desired ITR functions (*e.g.*, ability to form a hairpin, which allows DNA replication; integration of the AAV DNA into a host cell genome; and/or packaging, if desired).

[0087] A rAAV particle can comprise an ITR having any suitable number of "GAGC" repeats. In certain embodiments an ITR of an AAV2 particle comprises 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 or more "GAGC" repeats. In certain embodiments a rAAV2 particle comprises an ITR comprising three "GAGC" repeats. In certain embodiments a rAAV2 particle comprises an ITR which has less than four "GAGC" repeats. In certain embodiments a rAAV2 particle comprises an ITR which has more than four "GAGC" repeats. In certain embodiments an ITR of a rAAV2 particle comprises a Rep binding site wherein the fourth nucleotide in the first two "GAGC" repeats is a C rather than a T.

[0088] Exemplary suitable length of DNA can be incorporated in rAAV vectors for packaging/encapsidation into a rAAV particle can about 5 kilobases (kb) or less. In particular, embodiments, length of DNA is less than about 5kb, less than about 4.5 kb, less than about 4 kb, less than about 3.5 kb, less than about 3 kb, or less than about 2.5 kb.

[0089] rAAV vectors that include a nucleic acid sequence that directs the expression of an RNAi or polypeptide can be generated using suitable recombinant techniques known in the art (*e.g.*, *see* Sambrook *et al.*, 1989). Recombinant AAV vectors are typically packaged into transduction-competent AAV particles and propagated using an AAV viral packaging system. A transduction-competent AAV particle is capable of binding to and entering a mammalian cell and subsequently delivering a nucleic acid cargo (*e.g.*, a heterologous gene) to the nucleus of the cell. Thus, an intact rAAV particle that is transduction-competent is configured to transduce a mammalian cell. A rAAV particle configured to transduce a mammalian cell is often not replication competent, and requires additional protein machinery to self-replicate.

Thus, a rAAV particle that is configured to transduce a mammalian cell is engineered to bind and enter a mammalian cell and deliver a nucleic acid to the cell, wherein the nucleic acid for delivery is often positioned between a pair of AAV ITRs in the rAAV genome.

[0090] Suitable host cells for producing transduction-competent AAV particles include but are not limited to microorganisms, yeast cells, insect cells, and mammalian cells that can be, or have been, used as recipients of a heterologous rAAV vectors. Cells from the stable human cell line, HEK293 (readily available through, *e.g.*, the American Type Culture Collection under Accession Number ATCC CRL1573) can be used. In certain embodiments a modified human embryonic kidney cell line (*e.g.*, HEK293), which is transformed with adenovirus type-5 DNA fragments, and expresses the adenoviral E1a and E1b genes is used to generate recombinant AAV particles. The modified HEK293 cell line is readily transfected, and provides a particularly convenient platform in which to produce rAAV particles. Methods of generating high titer AAV particles capable of transducing mammalian cells are known in the art. For example, AAV particle can be made as set forth in Wright, 2008 and Wright, 2009.

[0091] In certain embodiments, AAV helper functions are introduced into the host cell by transfecting the host cell with an AAV helper construct either prior to, or concurrently with, the transfection of an AAV expression vector. AAV helper constructs are thus sometimes used to provide at least transient expression of AAV rep and/or cap genes to complement missing AAV functions necessary for productive AAV transduction. AAV helper constructs often lack AAV ITRs and can neither replicate nor package themselves. These constructs can be in the form of a plasmid, phage, transposon, cosmid, virus, or virion. A number of AAV helper constructs have been described, such as the commonly used plasmids pAAV/Ad and pIM29+45 which encode both Rep and Cap expression products. A number of other vectors are known which encode Rep and/or Cap expression products.

[0092] An "expression vector" is a specialized vector that contains a gene or nucleic acid sequence with the necessary regulatory regions needed for expression in a host cell. An expression vector may contain at least an origin of replication for propagation in a cell and optionally additional elements, such as a heterologous nucleic acid sequence, expression control element (*e.g.*, a promoter, enhancer), intron, ITR(s), and polyadenylation signal.

II. Therapeutic Agents

[0093] In some embodiments, viral gene transfer methods can be used to introduce nucleic acids in mammalian cells or target tissues. Such methods can be used to administer nucleic acids encoding inhibitory RNAs, non-coding RNAs, and/or therapeutic proteins to cells in culture or in a host organism.

A. Inhibitory RNAs

[0094] "RNA interference (RNAi)" is the process of sequence-specific, post-transcriptional gene silencing initiated by siRNA. During RNAi, siRNA induces degradation of target mRNA with consequent sequence-specific inhibition of gene expression.

[0095] An "inhibitory RNA," "RNAi," "small interfering RNA" or "short interfering RNA" or "siRNA" molecule, "short hairpin RNA" or "shRNA" molecule, or "miRNA" is an RNA duplex of nucleotides that is targeted to a nucleic acid sequence of interest. As used herein, the term "siRNA" is a generic term that encompasses the subset of shRNAs and miRNAs. An "RNA duplex" refers to the structure formed by the complementary pairing between two regions of an RNA molecule. siRNA is "targeted" to a gene in that the nucleotide sequence of the duplex portion of the siRNA is complementary to a nucleotide sequence of the targeted gene. In certain embodiments, the siRNAs are targeted to the sequence encoding huntingtin. In some embodiments, the length of the duplex of siRNAs is less than 30 base pairs. In some embodiments, the duplex can be 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11 or 10 base pairs in length. In some embodiments, the length of the duplex is 19 to 25 base pairs in length. In certain embodiment, the length of the duplex is 19 or 21 base pairs in length. The RNA duplex portion of the siRNA can be part of a hairpin structure. In addition to the duplex portion, the hairpin structure may contain a loop portion positioned between the two sequences that form the duplex. The loop can vary in length. In some embodiments the loop is 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25 nucleotides in length. In certain embodiments, the loop is 18 nucleotides in length. The hairpin structure can also contain 3′ and/or 5′ overhang portions. In some embodiments, the overhang is a 3′ and/or a 5′ overhang 0, 1, 2, 3, 4 or 5 nucleotides in length.

[0096] shRNAs are comprised of stem-loop structures which are designed to contain a 5′ flanking region, siRNA region segments, a loop region, a 3′ siRNA region and a 3′ flanking region. Most RNAi expression strategies have utilized short-hairpin RNAs (shRNAs) driven

by strong polIII-based promoters. Many shRNAs have demonstrated effective knock down of the target sequences in vitro as well as in vivo, however, some shRNAs which demonstrated effective knock down of the target gene were also found to have toxicity in vivo.

[0097] miRNAs are small cellular RNAs (~22 nt) that are processed from precursor stem loop transcripts. Known miRNA stem loops can be modified to contain RNAi sequences specific for genes of interest. miRNA molecules can be preferable over shRNA molecules because miRNAs are endogenously expressed. Therefore, miRNA molecules are unlikely to induce dsRNA-responsive interferon pathways, they are processed more efficiently than shRNAs, and they have been shown to silence 80% more effectively.

[0098] A recently discovered alternative approach is the use of artificial miRNAs (pri-miRNA scaffolds shuttling siRNA sequences) as RNAi vectors. Artificial miRNAs more naturally resemble endogenous RNAi substrates and are more amenable to Pol-II transcription (*e.g.*, allowing tissue-specific expression of RNAi) and polycistronic strategies (*e.g.*, allowing delivery of multiple siRNA sequences). See U.S. Pat. No. 10,093,927, which is incorporated by reference.

[0099] The transcriptional unit of a "shRNA" is comprised of sense and antisense sequences connected by a loop of unpaired nucleotides. shRNAs are exported from the nucleus by Exportin-5, and once in the cytoplasm, are processed by Dicer to generate functional siRNAs. "miRNAs" stem-loops are comprised of sense and antisense sequences connected by a loop of unpaired nucleotides typically expressed as part of larger primary transcripts (pri-miRNAs), which are excised by the Drosha-DGCR8 complex generating intermediates known as pre-miRNAs, which are subsequently exported from the nucleus by Exportin-5, and once in the cytoplasm, are processed by Dicer to generate functional siRNAs. "Artificial miRNA" or an "artificial miRNA shuttle vector", as used herein interchangeably, refers to a primary miRNA transcript that has had a region of the duplex stem loop (at least about 9-20 nucleotides) which is excised via Drosha and Dicer processing replaced with the siRNA sequences for the target gene while retaining the structural elements within the stem loop necessary for effective Drosha processing. The term "artificial" arises from the fact the flanking sequences (~35 nucleotides upstream and ~40 nucleotides downstream) arise from restriction enzyme sites within the multiple cloning site of the siRNA. As used herein the term "miRNA" encompasses both the naturally occurring miRNA sequences as well as artificially generated miRNA shuttle vectors.

[00100] The siRNA can be encoded by a nucleic acid sequence, and the nucleic acid sequence can also include a promoter. The nucleic acid sequence can also include a polyadenylation signal. In some embodiments, the polyadenylation signal is a synthetic minimal polyadenylation signal or a sequence of six Ts.

[00101] In designing RNAi there are several factors that need to be considered, such as the nature of the siRNA, the durability of the silencing effect, and the choice of delivery system. To produce an RNAi effect, the siRNA that is introduced into the organism will typically contain exonic sequences. Furthermore, the RNAi process is homology dependent, so the sequences must be carefully selected so as to maximize gene specificity, while minimizing the possibility of cross-interference between homologous, but not gene-specific sequences. Preferably the siRNA exhibits greater than 80%, 85%, 90%, 95%, 98%, or even 100% identity between the sequence of the siRNA and the gene to be inhibited. Sequences less than about 80% identical to the target gene are substantially less effective. Thus, the greater homology between the siRNA and the gene to be inhibited, the less likely expression of unrelated genes will be affected.

[00102] In addition, the size of the siRNA is an important consideration. In some embodiments, the present invention relates to siRNA molecules that include at least about 19-25 nucleotides and are able to modulate gene expression. In the context of the present invention, the siRNA is preferably less than 500, 200, 100, 50, or 25 nucleotides in length. More preferably, the siRNA is from about 19 nucleotides to about 25 nucleotides in length.

[00103] A siRNA target generally means a polynucleotide comprising a region that encodes a polypeptide, or a polynucleotide region that regulates replication, transcription, or translation or other processes important to expression of the polypeptide, or a polynucleotide comprising both a region that encodes a polypeptide and a region operably linked thereto that regulates expression. Any gene being expressed in a cell can be targeted. Preferably, a target gene is one involved in or associated with the progression of cellular activities important to disease or of particular interest as a research object.

B. Non-Coding RNAs

[00104] As evidenced by cDNA cloning projects and genomic tiling arrays, more than 90% of the human genome undergoes transcription but does not code for proteins. These transcriptional products are referred to as non-protein coding RNAs (ncRNAs). A variety of

ncRNA transcripts, such as ribosomal RNAs, transfer RNAs, competing endogenous RNA (ceRNA), small nuclear RNA (snRNA), and small nucleolar RNA (snoRNA), are essential for cell function. Similarly, a large number of short ncRNAs such as micro-RNAs (miRNAs), endogenous short interfering RNAs (siRNAs), PIWI-interacting RNAs (piRNAs), and small nucleolar RNAs (snoRNAs) are also known to play important regulatory roles in eukaryotic cells. Recent studies have demonstrated a group of long ncRNA (lncRNA) transcripts that exhibit cell type-specific expression and localize into specific subcellular compartments. lncRNAs are also known to play important roles during cellular development and differentiation supporting the view that they have been selected during the evolutionary process.

[00105] LncRNAs appear to have many different functions. In many cases, they seem to play a role in regulating the activity or localization of proteins, or serve as organizational frameworks for subcellular structures. In other cases, lncRNAs are processed to yield multiple small RNAs or they may modulate how other RNAs are processed. The latest edition of data produced by the public research consortium GenCode (version #27) catalogs just under 16,000 lncRNAs in the human genome, producing nearly 28,000 transcripts; when other databases are included, more than 40,000 lncRNAs are known.

[00106] Interestingly, lncRNAs can influence the expression of specific target proteins at specific genomic loci, modulate the activity of protein binding partners, direct chromatin- modifying complexes to their sites of action, and are post-transcriptionally processed to produce numerous 5′-capped small RNAs. Epigenetic pathways can also regulate the differential expression of lncRNAs.

[00107] A growing body of evidence also suggests that aberrantly expressed lncRNAs play important roles in normal physiological processes as well as multiple disease states. lncRNAs are misregulated in various diseases, including ischaemia, heart disease, Alzheimer's disease, psoriasis, and spinocerebellar ataxia type 8. This misregulation has also been shown in various types of cancers, such as breast cancer, colon cancer, prostate cancer, hepatocellular carcinoma and leukemia. Several lncRNAs, *e.g.* gadd74 and lncRNA-RoR5, modulate cell cycle regulators such as cyclins, cyclin-dependent kinases (CDKs), CDK inhibitors and p53 and thus provide an additional layer of flexibility and robustness to cell cycle progression. In addition, some lncRNAs are linked to mitotic processes such as centromeric satellite RNA, which is essential for kinetochore formation and thus crucial for chromosome

segregation during mitosis in humans and flies. Another nuclear lncRNA, MA-lincl, regulates M phase exit by functioning in cis to repress the expression of its neighbouring gene Pura, a regulator of cell proliferation.

[00108] lncRNAs are a group that is commonly defined as transcripts of more than 200 nucleotides (*e.g.* about 200 to about 1200 nt, about 2500 nt, or more) that lack an extended open reading frame (ORF). The term "non-coding RNA" (ncRNA) includes lncRNA as well as shorter transcripts of, *e.g.*, less than about 200 nt, such as about 30 to 200 nt.

[00109] Thus, in some embodiments, delivery of a ncRNA, such as to a specific brain structure of interest, corrects aberrant RNA expression levels or modulates levels of disease-causing lncRNA. Accordingly, in some embodiments, the present invention provides an rAAV, wherein the viral genome is engineered to encode a therapeutic non-coding RNA (ncRNA). In some embodiments, the ncRNA is a long non-coding RNA (lncRNA) of about 200 nucleotides (nt) in length or greater. In some embodiments, the therapeutic is a ncRNA of about 25 nt or about 30 nt to about 200 nt in length. In some embodiments, the lncRNA is about 200 nt to about 1,200 nt in length. In some embodiments, the lncRNA is about 200 nt to about 1,100, about 1,000, about 900, about 800, about 700, about 600, about 500, about 400, or about 300 nt in length.

C. CRISPR Systems

[00110] Gene editing is a technology that allows for the modification of target genes within living cells. Recently, harnessing the bacterial immune system of CRISPR to perform on demand gene editing revolutionized the way scientists approach genomic editing. The Cas9 protein of the CRISPR system, which is an RNA guided DNA endonuclease, can be engineered to target new sites with relative ease by altering its guide RNA sequence. This discovery has made sequence specific gene editing functionally effective.

[00111] In general, "CRISPR system" refers collectively to transcripts and other elements involved in the expression of or directing the activity of CRISPR-associated ("Cas") genes, including sequences encoding a Cas gene, a tracr (trans-activating CRISPR) sequence (*e.g.* tracrRNA or an active partial tracrRNA), a tracr-mate sequence (encompassing a "direct repeat" and a tracrRNA-processed partial direct repeat in the context of an endogenous CRISPR system), a guide sequence (also referred to as a "spacer" in the context of an endogenous CRISPR system), and/or other sequences and transcripts from a CRISPR locus.

[00112] The CRISPR/Cas nuclease or CRISPR/Cas nuclease system can include a non-coding RNA molecule (guide) RNA, which sequence-specifically binds to DNA, and a Cas protein (*e.g.*, Cas9), with nuclease functionality (*e.g.*, two nuclease domains). One or more elements of a CRISPR system can derive from a type I, type II, or type III CRISPR system, *e.g.*, derived from a particular organism comprising an endogenous CRISPR system, such as *Streptococcus pyogenes*.

[00113] The CRISPR system can induce double stranded breaks (DSBs) at the target site, followed by disruptions as discussed herein. In other embodiments, Cas9 variants, deemed "nickases," are used to nick a single strand at the target site. Paired nickases can be used, *e.g.*, to improve specificity, each directed by a pair of different gRNAs targeting sequences such that upon introduction of the nicks simultaneously, a 5' overhang is introduced. In other embodiments, catalytically inactive Cas9 is fused to a heterologous effector domain such as a transcriptional repressor (*e.g.*, KRAB) or activator, to affect gene expression. Alternatively, a CRISPR system with a catalytically inactivate Cas9 further comprises a transcriptional repressor or activator fused to a ribosomal binding protein.

[00114] In some aspects, a Cas nuclease and gRNA (including a fusion of crRNA specific for the target sequence and fixed tracrRNA) are introduced into the cell. In general, target sites at the 5' end of the gRNA target the Cas nuclease to the target site, *e.g.*, the gene, using complementary base pairing. The target site may be selected based on its location immediately 5' of a protospacer adjacent motif (PAM) sequence, such as typically NGG, or NAG. In this respect, the gRNA is targeted to the desired sequence by modifying the first 20, 19, 18, 17, 16, 15, 14, 14, 12, 11, or 10 nucleotides of the guide RNA to correspond to the target DNA sequence. In general, a CRISPR system is characterized by elements that promote the formation of a CRISPR complex at the site of a target sequence. Typically, "target sequence" generally refers to a sequence to which a guide sequence is designed to have complementarity, where hybridization between the target sequence and a guide sequence promotes the formation of a CRISPR complex. Full complementarity is not necessarily required, provided there is sufficient complementarity to cause hybridization and promote formation of a CRISPR complex.

[00115] The target sequence may comprise any polynucleotide, such as DNA or RNA polynucleotides. The target sequence may be located in the nucleus or cytoplasm of the cell, such as within an organelle of the cell. Generally, a sequence or template that may be used

for recombination into the targeted locus comprising the target sequences is referred to as an "editing template" or "editing polynucleotide" or "editing sequence." In some aspects, an exogenous template polynucleotide may be referred to as an editing template. In some aspects, the recombination is homologous recombination.

[00116] Typically, in the context of an endogenous CRISPR system, formation of the CRISPR complex (comprising the guide sequence hybridized to the target sequence and complexed with one or more Cas proteins) results in cleavage of one or both strands in or near (*e.g.* within 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or more base pairs from) the target sequence. The tracr sequence, which may comprise or consist of all or a portion of a wild-type tracr sequence (*e.g.* about or more than about 20, 26, 32, 45, 48, 54, 63, 67, 85, or more nucleotides of a wild-type tracr sequence), may also form part of the CRISPR complex, such as by hybridization along at least a portion of the tracr sequence to all or a portion of a tracr mate sequence that is operably linked to the guide sequence. The tracr sequence has sufficient complementarity to a tracr mate sequence to hybridize and participate in formation of the CRISPR complex, such as at least 50%, 60%, 70%, 80%, 90%, 95% or 99% of sequence complementarity along the length of the tracr mate sequence when optimally aligned.

[00117] One or more vectors driving expression of one or more elements of the CRISPR system can be introduced into the cell such that expression of the elements of the CRISPR system direct formation of the CRISPR complex at one or more target sites. Components can also be delivered to cells as proteins and/or RNA. For example, a Cas enzyme, a guide sequence linked to a tracr-mate sequence, and a tracr sequence could each be operably linked to separate regulatory elements on separate vectors. The Cas enzyme may be a target gene under the control of a regulated alternative splicing event, as disclosed herein, either as a chimeric target gene minigene or as a target gene for a chimeric minigene transactivator. The gRNA may be under the control of a constitutive promoter.

[00118] Alternatively, two or more of the elements expressed from the same or different regulatory elements, may be combined in a single vector, with one or more additional vectors providing any components of the CRISPR system not included in the first vector. The vector may comprise one or more insertion sites, such as a restriction endonuclease recognition sequence (also referred to as a "cloning site"). In some embodiments, one or more insertion sites are located upstream and/or downstream of one or more sequence elements of one or more vectors. When multiple different guide sequences are used, a single expression construct may

be used to target CRISPR activity to multiple different, corresponding target sequences within a cell.

[00119] A vector may comprise a regulatory element operably linked to an enzyme-coding sequence encoding the CRISPR enzyme, such as a Cas protein. Non-limiting examples of Cas proteins include Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (also known as Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, homologs thereof, or modified versions thereof. These enzymes are known; for example, the amino acid sequence of *S. pyogenes* Cas9 protein may be found in the SwissProt database under accession number Q99ZW2.

[00120] The CRISPR enzyme can be Cas9 (*e.g.*, from *S. pyogenes* or *S. pneumonia*). The CRISPR enzyme can direct cleavage of one or both strands at the location of a target sequence, such as within the target sequence and/or within the complement of the target sequence. The vector can encode a CRISPR enzyme that is mutated with respect to a corresponding wild-type enzyme such that the mutated CRISPR enzyme lacks the ability to cleave one or both strands of a target polynucleotide containing a target sequence. For example, an aspartate-to-alanine substitution (D10A) in the RuvC I catalytic domain of Cas9 from S. pyogenes converts Cas9 from a nuclease that cleaves both strands to a nickase (cleaves a single strand). In some embodiments, a Cas9 nickase may be used in combination with guide sequence(s), *e.g.*, two guide sequences, which target respectively sense and antisense strands of the DNA target. This combination allows both strands to be nicked and used to induce NHEJ or HDR.

[00121] In some embodiments, an enzyme coding sequence encoding the CRISPR enzyme is codon optimized for expression in particular cells, such as eukaryotic cells. The eukaryotic cells may be those of or derived from a particular organism, such as a mammal, including but not limited to human, mouse, rat, rabbit, dog, or non-human primate. In general, codon optimization refers to a process of modifying a nucleic acid sequence for enhanced expression in the host cells of interest by replacing at least one codon of the native sequence with codons that are more frequently or most frequently used in the genes of that host cell while maintaining the native amino acid sequence. Various species exhibit particular bias for certain codons of a particular amino acid. Codon bias (differences in codon usage between organisms)

often correlates with the efficiency of translation of messenger RNA (mRNA), which is in turn believed to be dependent on, among other things, the properties of the codons being translated and the availability of particular transfer RNA (tRNA) molecules. The predominance of selected tRNAs in a cell is generally a reflection of the codons used most frequently in peptide synthesis. Accordingly, genes can be tailored for optimal gene expression in a given organism based on codon optimization.

[00122] In general, a guide sequence is any polynucleotide sequence having sufficient complementarity with a target polynucleotide sequence to hybridize with the target sequence and direct sequence-specific binding of the CRISPR complex to the target sequence. In some embodiments, the degree of complementarity between a guide sequence and its corresponding target sequence, when optimally aligned using a suitable alignment algorithm, is about or more than about 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or more.

[00123] Optimal alignment may be determined with the use of any suitable algorithm for aligning sequences, non-limiting example of which include the Smith-Waterman algorithm, the Needleman-Wunsch algorithm, algorithms based on the Burrows-Wheeler Transform (*e.g.* the Burrows Wheeler Aligner), Clustal W, Clustal X, BLAT, Novoalign (Novocraft Technologies, ELAND (Illumina, San Diego, Calif.), SOAP (available at soap.genomics.org.cn), and Maq (available at maq.sourceforge.net).

[00124] The CRISPR enzyme may be part of a fusion protein comprising one or more heterologous protein domains. A CRISPR enzyme fusion protein may comprise any additional protein sequence, and optionally a linker sequence between any two domains. Examples of protein domains that may be fused to a CRISPR enzyme include, without limitation, epitope tags, reporter gene sequences, and protein domains having one or more of the following activities: methylase activity, demethylase activity, transcription activation activity, transcription repression activity, transcription release factor activity, histone modification activity, RNA cleavage activity and nucleic acid binding activity. Non-limiting examples of epitope tags include histidine (His) tags, V5 tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Examples of reporter genes include, but are not limited to, glutathione-5- transferase (GST), horseradish peroxidase (HRP), chloramphenicol acetyltransferase (CAT) beta galactosidase, beta-glucuronidase, luciferase, green fluorescent protein (GFP), HcRed, DsRed, cyan fluorescent protein (CFP), yellow fluorescent protein (YFP), and autofluorescent proteins including blue

fluorescent protein (BFP). A CRISPR enzyme may be fused to a gene sequence encoding a protein or a fragment of a protein that bind DNA molecules or bind other cellular molecules, including but not limited to maltose binding protein (MBP), S-tag, Lex A DNA binding domain (DBD) fusions, GAL4A DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. Additional domains that may form part of a fusion protein comprising a CRISPR enzyme are described in US 20110059502, incorporated herein by reference.

D. Therapeutic Proteins

[00125] Some embodiments concern expression of recombinant proteins and polypeptides. Such proteins may be neuroprotective proteins, such as an anti-apoptotic protein, an anti-oxidant enzyme (e.g., those belonging to the superoxide dismutase (SOD) family), a neurotrophic/neuroprotective factor, or an anti-inflammatory protein. The therapeutic protein may be Birc1a (NAIP), Birc2 (c-IAP1/HIAP-2), Birc3 (cIAP-2/HIAP-1), Birc4 (XIAP), Birc5 (survivin), Birc6 (apollon), Birc7 (livin), Birc8 (TsIAP); members of the Bcl-2 family: Bcl-2, Bcl-XL, Bcl-w, Mcl-1, Bcl-2L10, BFL-1; endogenous inhibitors of the c-Jun N-terminus kinase (JNK) known as Jun-interacting protein (JIP), JIP-1, JIP-2, JIP-3, JIP-4; SOD1, SOD2; catalase; peroxiredoxin-1, peroxiredoxin-2, glutathione preoxidase 1 (Gpx1), Gpx2, Gpx3, or Gpx4; NGF, BDNF, CNTF, GDNF, Growth/differentiation factor-15 (GDF-15), erythropoietin or vascular endothelial growth factor (VEGF); interleukin-10 (IL-10); glutathione S-transferase, Annexin-1 (ANXA1), inhibitor of NF-κB (IκB); USH1, USH1G (a.k.a. SANS), CIB2 (calcium and integrin binding protein 2), ATOH-1, ACTG1, ATP2B2, CDH23 (cadherin 23), CLDN14, CLRN1, COCH, COL11A2, DFNA5, DFNB31, DFNB59, ESPN, EYA4, GJB2, GJB3, GJB6, KCNQ4, LHFPL5, MT-RNR1, MT-TS1, MYO1A, MYO6, MYO7A (myosin 7a), MYO15A, OTOF, PCDH15 (protocadherin 15), PDZD7, POU3F4, SLC26A4, STRC, TECTA, TMC1, TMC2, TMIE, TMPRSS3, TRIOBP, USH1C, VLGR1, WFS1, ACTG1, ADCY1, ATOFfl, ATP6V1B1, BDNF, BDP1, BSND, DATSPER2, CABP2, CD 164, CDC 14 A, CDH23, CEACAM16, CHD7, CCDC50, CIB2, CLDN14, CLIC5, CLPP, CLRNl, COCH, COL2A1, COL4A3, COL4A4, COL4A5, COL9A 1, COL9A2, COL1 1 A1, COL1 1 A2, CRYM, DCDC2, DFNA5, DFNB31, DFNB59, DIAPH1, EDN3, EDNRB, ELMOD3, EMOD3, EPS8, EPS8L2, ESPN, ESRRB, EYA1, EYA4, FAM65B, FOXI1, GIPC3, GJB2, GJB3, GJB6, GPR98, GRHL2, GPSM2, GRXCR1 , GRXCR2, HARS2, HGF, HOMER2, HSD17B4, ILDR1, KARS, KCNE1, KCNJ10, KCNQ1, KCNQ4, KITLG, LARS2, LHFPL5, LOXHD1, LRTOMT, MARVELD2, MCM2, MET, MIR183, MIRN96, MITF, MSRB3, MT-RNR1, MT-TS 1, MYH14, MYH9, MY015A, MYOIA, MY03A, MY06,

MY07A, NARS2, NDP, NF2, NT3, OSBPL2, OTOA, OTOF, OTOG, OTOGL, P2RX2, PAX3, PCDH15, PDZD7, PJVK, PNPT1, POLR1D, POLR1C, POU3F4, POU4F3, PRPS 1, PTPRQ, RDX, S1PR2, SANS, SEMA3E, SERPINB6, SLC17A8, SLC22A4, SLC26A4, SLC26A5, SIX1, SIX5, S MAC/DIABLO, SNAI2, SOX 10, STRC, SYNE4, TBC 1 D24, TCOF1, TECTA, TIMM8A, TJP2, TNC, TMC1, TMC2, TMIE, TMEM 132E, TMPRSS3, TRPN, TRIOBP, TSPEAR, USH1 C, USH1 G, USH2A, USH2D, VLGRl, WFS1, WHRN, or XIAP.

[00126] When the present application refers to the function or activity of "modified protein" or a "modified polypeptide," one of ordinary skill in the art would understand that this includes, for example, a protein or polypeptide that possesses an additional advantage over the unmodified protein or polypeptide. It is specifically contemplated that embodiments concerning a "modified protein" may be implemented with respect to a "modified polypeptide," and vice versa.

[00127] Recombinant proteins may possess deletions and/or substitutions of amino acids; thus, a protein with a deletion, a protein with a substitution, and a protein with a deletion and a substitution are modified proteins. In some embodiments, these proteins may further include insertions or added amino acids, such as with fusion proteins or proteins with linkers, for example. A "modified deleted protein" lacks one or more residues of the native protein, but may possess the specificity and/or activity of the native protein. A "modified deleted protein" may also have reduced immunogenicity or antigenicity. An example of a modified deleted protein is one that has an amino acid residue deleted from at least one antigenic region, *i.e.* a region of the protein determined to be antigenic in a particular organism, such as the organism to which the modified protein is being administered.

[00128] Substitution or replacement variants typically contain the exchange of one amino acid for another at one or more sites within the protein and may be designed to modulate one or more properties of the polypeptide, particularly its effector functions and/or bioavailability. Substitutions may or may not be conservative, that is, one amino acid is replaced with one of similar shape and charge. Conservative substitutions are well known in the art and include, for example, the changes of: alanine to serine; arginine to lysine; asparagine to glutamine or histidine; aspartate to glutamate; cysteine to serine; glutamine to asparagine; glutamate to aspartate; glycine to proline; histidine to asparagine or glutamine; isoleucine to leucine or valine; leucine to valine or isoleucine; lysine to arginine; methionine to leucine or

isoleucine; phenylalanine to tyrosine, leucine, or methionine; serine to threonine; threonine to serine; tryptophan to tyrosine; tyrosine to tryptophan or phenylalanine; and valine to isoleucine or leucine.

[00129] In addition to a deletion or substitution, a modified protein may possess an insertion of residues, which typically involves the addition of at least one residue in the polypeptide. This may include the insertion of a targeting peptide or polypeptide or simply a single residue. Terminal additions, called fusion proteins, are discussed below.

[00130] The term "biologically functional equivalent" is well understood in the art and is further defined in detail herein. Accordingly, sequences that have between about 70% and about 80%, or between about 81% and about 90%, or even between about 91% and about 99% of amino acids that are identical or functionally equivalent to the amino acids of a control polypeptide are included, provided the biological activity of the protein is maintained. A recombinant protein may be biologically functionally equivalent to its native counterpart in certain aspects.

[00131] It also will be understood that amino acid and nucleic acid sequences may include additional residues, such as additional N- or C-terminal amino acids or 5′ or 3′ sequences, and yet still be essentially as set forth in one of the sequences disclosed herein, so long as the sequence meets the criteria set forth above, including the maintenance of biological protein activity where protein expression is concerned. The addition of terminal sequences particularly applies to nucleic acid sequences that may, for example, include various non-coding sequences flanking either of the 5′ or 3′ portions of the coding region or may include various internal sequences, *i.e.*, introns, which are known to occur within genes.

[00132] As used herein, a protein or peptide generally refers, but is not limited to, a protein of greater than about 200 amino acids, up to a full-length sequence translated from a gene; a polypeptide of greater than about 100 amino acids; and/or a peptide of from about 3 to about 100 amino acids. For convenience, the terms "protein," "polypeptide," and "peptide are used interchangeably herein.

[00133] As used herein, an "amino acid residue" refers to any naturally occurring amino acid, any amino acid derivative, or any amino acid mimic known in the art. In certain embodiments, the residues of the protein or peptide are sequential, without any non-amino acids interrupting the sequence of amino acid residues. In other embodiments, the sequence

may comprise one or more non-amino acid moieties. In particular embodiments, the sequence of residues of the protein or peptide may be interrupted by one or more non-amino acid moieties.

[00134] Accordingly, the term "protein or peptide" encompasses amino acid sequences comprising at least one of the 20 common amino acids found in naturally occurring proteins, or at least one modified or unusual amino acid.

[00135] Certain embodiments of the present invention concern fusion proteins. These molecules may have a therapeutic protein linked at the N- or C-terminus to a heterologous domain. For example, fusions may also employ leader sequences from other species to permit the recombinant expression of a protein in a heterologous host. Another useful fusion includes the addition of a protein affinity tag, such as a serum albumin affinity tag or six histidine residues, or an immunologically active domain, such as an antibody epitope, preferably cleavable, to facilitate purification of the fusion protein. Non-limiting affinity tags include polyhistidine, chitin binding protein (CBP), maltose binding protein (MBP), and glutathione-S-transferase (GST).

[00136] Methods of generating fusion proteins are well known to those of skill in the art. Such proteins can be produced, for example, by *de novo* synthesis of the complete fusion protein, or by attachment of the DNA sequence encoding the heterologous domain, followed by expression of the intact fusion protein.

[00137] Production of fusion proteins that recover the functional activities of the parent proteins may be facilitated by connecting genes with a bridging DNA segment encoding a peptide linker that is spliced between the polypeptides connected in tandem. The linker would be of sufficient length to allow proper folding of the resulting fusion protein.

[00138] Expression of a transgene may be directed by the transgene's natural promoter (i.e., the promoter found naturally with the transgenic coding sequence) or expression of a transgene may be directed by a heterologous promoter (e.g., CMV promoter, Espin promoter, a PCDH15 promoter, a PTPRQ promoter and a TMHS (LHFPL5) promoter). For example, any of the transgenes described herein can be used with its natural promoter. Alternatively, any of the transgenes described herein can be used with a heterologous promoter. As used herein, a heterologous promoter refers to a promoter that does not naturally direct expression of that sequence (i.e., is not found with that sequence in nature). Representative

heterologous promoters that can be used to direct expression of any of the transgenes indicated herein include, for example, a CMV promoter, a CBA promoter, a CASI promoter, a P promoter, and a EF-1 promoter, an alpha9 nicotinic receptor promoter, a prestin promoter, a Gfil promoter, and a Vglut3 promoter. In addition, a promoter that naturally directs expression of one of the above-referenced transgenes (e.g., a KCNQ4 promoter, a Myo7a promoter, a Myo6 promoter or an Atohl promoter) can be used as a heterologous promoter to direct expression of a transgene. In other embodiments, the promoter is an Espin promoter, a PCDH15 promoter, a PTPRQ promoter and a TMHS (LHFPL5) promoter.

III. Methods of Treatment and Administration

[00139] Viral vectors in some aspects may be administered directly to patients (*in vivo*) or they can be used to treat cells *in vitro* or *ex vivo*, and then administered to patients. The term "vector" refers to small carrier nucleic acid molecule, a plasmid, virus (*e.g.*, AAV vector, retroviral vector, lentiviral vector), or other vehicle that can be manipulated by insertion or incorporation of a nucleic acid. Vectors, such as viral vectors, can be used to introduce/transfer nucleic acid sequences into cells, such that the nucleic acid sequence therein is transcribed and, if encoding a protein, subsequently translated by the cells.

[00140] Any suitable cell or mammal can be administered or treated by a method or use described herein. Typically, a mammal in need of a method described herein is suspected of having or expressing an abnormal or aberrant protein that is associated with a disease state. Alternative, the mammalian recipient may have a condition that is amenable to gene replacement therapy. As used herein, "gene replacement therapy" refers to administration to the recipient of exogenous genetic material encoding a therapeutic agent and subsequent expression of the administered genetic material in situ. Thus, the phrase "condition amenable to gene replacement therapy" embraces conditions such as genetic diseases (*i.e.*, a disease condition that is attributable to one or more gene defects), acquired pathologies (*i.e.*, a pathological condition which is not attributable to an inborn defect), cancers and prophylactic processes (*i.e.*, prevention of a disease or of an undesired medical condition). Accordingly, as used herein, the term "therapeutic agent" refers to any agent or material, which has a beneficial effect on the mammalian recipient. Thus, "therapeutic agent" embraces both therapeutic and prophylactic molecules having nucleic acid or protein components.

[00141] Non-limiting examples of mammals include humans, non-human primates (*e.g.*, apes, gibbons, chimpanzees, orangutans, monkeys, macaques, and the like),

domestic animals (*e.g.*, dogs and cats), farm animals (*e.g.*, horses, cows, goats, sheep, pigs) and experimental animals (*e.g.*, mouse, rat, rabbit, guinea pig). In certain embodiments a mammal is a human. In certain embodiments a mammal is a non-rodent mammal (*e.g.*, human, pig, goat, sheep, horse, dog, or the like). In certain embodiments a non-rodent mammal is a human. A mammal can be any age or at any stage of development (*e.g.*, an adult, teen, child, infant, or a mammal in utero). A mammal can be male or female. In certain embodiments a mammal can be an animal disease model, for example, animal models having or expressing an abnormal or aberrant protein that is associated with a disease state or animal models with insufficient expression of a protein, which causes a disease state.

[00142] Mammals (subjects) treated by a method or composition described herein include adults (18 years or older) and children (less than 18 years of age). Adults include the elderly. Representative adults are 50 years or older. Children range in age from 1-2 years old, or from 2–4, 4–6, 6–18, 8–10, 10–12, 12–15 and 15–18 years old. Children also include infants. Infants typically range from 1–12 months of age.

[00143] In certain embodiments, a method includes administering a plurality of viral particles to a mammal as set forth herein, where severity, frequency, progression or time of onset of one or more symptoms of a disease state, such as a neuro-degenerative disease, decreased, reduced, prevented, inhibited or delayed. In certain embodiments, a method includes administering a plurality of viral particles to a mammal to treat an adverse symptom of a disease state, such as a neuro-degenerative disease. In certain embodiments, a method includes administering a plurality of viral particles to a mammal to stabilize, delay or prevent worsening, or progression, or reverse and adverse symptom of a disease state, such as a neuro-degenerative disease.

[00144] In certain embodiments a method includes administering a plurality of viral particles to the central nervous system, or portion thereof as set forth herein, of a mammal and severity, frequency, progression or time of onset of one or more symptoms of a disease state, such as a neuro-degenerative disease, are decreased, reduced, prevented, inhibited or delayed by at least about 5 to about 10, about 10 to about 25, about 25 to about 50, or about 50 to about 100 days.

[00145] In certain embodiments, a symptom or adverse effect comprises an early stage, middle or late stage symptom; a behavior, personality or language symptom; swallowing,

movement, seizure, tremor or fidgeting symptom; ataxia; and/or a cognitive symptom such as memory, ability to organize.

IV. Pharmaceutical Compositions

[00146] As used herein the term "pharmaceutically acceptable" and "physiologically acceptable" mean a biologically acceptable composition, formulation, liquid or solid, or mixture thereof, which is suitable for one or more routes of administration, *in vivo* delivery or contact. A "pharmaceutically acceptable" or "physiologically acceptable" composition is a material that is not biologically or otherwise undesirable, *e.g.*, the material may be administered to a subject without causing substantial undesirable biological effects. Such composition, "pharmaceutically acceptable" and "physiologically acceptable" formulations and compositions can be sterile. Such pharmaceutical formulations and compositions may be used, for example in administering a viral particle to a subject.

[00147] Such formulations and compositions include solvents (aqueous or non-aqueous), solutions (aqueous or non-aqueous), emulsions (*e.g.*, oil-in-water or water-in-oil), suspensions, syrups, elixirs, dispersion and suspension media, coatings, isotonic and absorption promoting or delaying agents, compatible with pharmaceutical administration or *in vivo* contact or delivery. Aqueous and non-aqueous solvents, solutions and suspensions may include suspending agents and thickening agents. Supplementary active compounds (*e.g.*, preservatives, antibacterial, antiviral and antifungal agents) can also be incorporated into the formulations and compositions.

[00148] Pharmaceutical compositions typically contain a pharmaceutically acceptable excipient. Such excipients include any pharmaceutical agent that does not itself induce the production of antibodies harmful to the individual receiving the composition, and which may be administered without undue toxicity. Pharmaceutically acceptable excipients include, but are not limited to, sorbitol, Tween80, and liquids such as water, saline, glycerol and ethanol. Pharmaceutically acceptable salts can be included therein, for example, mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulfates, and the like; and the salts of organic acids such as acetates, propionates, malonates, benzoates, and the like. Additionally, auxiliary substances, such as surfactants, wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles.

[00149] Pharmaceutical compositions can be formulated to be compatible with a particular route of administration or delivery, as set forth herein or known to one of skill in the art. Thus, pharmaceutical compositions include carriers, diluents, or excipients suitable for administration or delivery by various routes.

[00150] Pharmaceutical forms suitable for injection or infusion of viral particles can include sterile aqueous solutions or dispersions which are adapted for the extemporaneous preparation of sterile injectable or infusible solutions or dispersions, optionally encapsulated in liposomes. In all cases, the ultimate form should be a sterile fluid and stable under the conditions of manufacture, use and storage. The liquid carrier or vehicle can be a solvent or liquid dispersion medium comprising, for example, water, ethanol, a polyol (for example, glycerol, propylene glycol, liquid polyethylene glycols, and the like), vegetable oils, nontoxic glyceryl esters, and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the formation of liposomes, by the maintenance of the required particle size in the case of dispersions or by the use of surfactants. Isotonic agents, for example, sugars, buffers or salts (*e.g.*, sodium chloride) can be included. Prolonged absorption of injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

[00151] Solutions or suspensions of viral particles can optionally include one or more of the following components: a sterile diluent such as water for injection, saline solution, such as phosphate buffered saline (PBS), artificial CSF, a surfactants, fixed oils, a polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), glycerin, or other synthetic solvents; antibacterial and antifungal agents such as parabens, chlorobutanol, phenol, ascorbic acid, and the like; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose.

[00152] Pharmaceutical formulations, compositions and delivery systems appropriate for the compositions, methods and uses of the invention are known in the art (*see, e.g.*, Remington: The Science and Practice of Pharmacy (2003) 20th ed., Mack Publishing Co., Easton, PA; Remington's Pharmaceutical Sciences (1990) 18th ed., Mack Publishing Co., Easton, PA; The Merck Index (1996) 12th ed., Merck Publishing Group, Whitehouse, NJ; Pharmaceutical Principles of Solid Dosage Forms (1993), Technonic Publishing Co., Inc., Lancaster, Pa.; Ansel and Stoklosa, Pharmaceutical Calculations (2001) 11th ed., Lippincott

Williams & Wilkins, Baltimore, MD; and Poznansky *et al.*, Drug Delivery Systems (1980), R. L. Juliano, ed., Oxford, N.Y., pp. 253-315).

[00153] Viral particles and their compositions may be formulated in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for an individual to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The dosage unit forms are dependent upon the number of viral particles believed necessary to produce the desired effect(s). The amount necessary can be formulated in a single dose, or can be formulated in multiple dosage units. The dose may be adjusted to a suitable viral particle concentration, optionally combined with an anti-inflammatory agent, and packaged for use.

[00154] In one embodiment, pharmaceutical compositions will include sufficient genetic material to provide a therapeutically effective amount, *i.e.*, an amount sufficient to reduce or ameliorate symptoms or an adverse effect of a disease state in question or an amount sufficient to confer the desired benefit.

[00155] A "unit dosage form" as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity optionally in association with a pharmaceutical carrier (excipient, diluent, vehicle or filling agent) which, when administered in one or more doses, is calculated to produce a desired effect (*e.g.*, prophylactic or therapeutic effect). Unit dosage forms may be within, for example, ampules and vials, which may include a liquid composition, or a composition in a freeze-dried or lyophilized state; a sterile liquid carrier, for example, can be added prior to administration or delivery *in vivo*. Individual unit dosage forms can be included in multi-dose kits or containers. Thus, for example, viral particles, and pharmaceutical compositions thereof, can be packaged in single or multiple unit dosage form for ease of administration and uniformity of dosage.

[00156] Formulations containing viral particles typically contain an effective amount, the effective amount being readily determined by one skilled in the art. The viral particles may typically range from about 1% to about 95% (w/w) of the composition, or even higher if suitable. The quantity to be administered depends upon factors such as the age, weight and physical condition of the mammal or the human subject considered for treatment. Effective

dosages can be established by one of ordinary skill in the art through routine trials establishing dose response curves.

V. Definitions

[00157] The terms "polynucleotide," "nucleic acid" and "transgene" are used interchangeably herein to refer to all forms of nucleic acid, oligonucleotides, including deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) and polymers thereof. Polynucleotides include genomic DNA, cDNA and antisense DNA, and spliced or unspliced mRNA, rRNA, tRNA and inhibitory DNA or RNA (RNAi, *e.g.*, small or short hairpin (sh)RNA, microRNA (miRNA), small or short interfering (si)RNA, trans-splicing RNA, or antisense RNA). Polynucleotides can include naturally occurring, synthetic, and intentionally modified or altered polynucleotides (*e.g.*, variant nucleic acid). Polynucleotides can be single stranded, double stranded, or triplex, linear or circular, and can be of any suitable length. In discussing polynucleotides, a sequence or structure of a particular polynucleotide may be described herein according to the convention of providing the sequence in the 5′ to 3′ direction.

[00158] A nucleic acid encoding a polypeptide often comprises an open reading frame that encodes the polypeptide. Unless otherwise indicated, a particular nucleic acid sequence also includes degenerate codon substitutions.

[00159] Nucleic acids can include one or more expression control or regulatory elements operably linked to the open reading frame, where the one or more regulatory elements are configured to direct the transcription and translation of the polypeptide encoded by the open reading frame in a mammalian cell. Non-limiting examples of expression control/regulatory elements include transcription initiation sequences (*e.g.*, promoters, enhancers, a TATA box, and the like), translation initiation sequences, mRNA stability sequences, poly A sequences, secretory sequences, and the like. Expression control/regulatory elements can be obtained from the genome of any suitable organism.

[00160] A "promoter" refers to a nucleotide sequence, usually upstream (5') of a coding sequence, which directs and/or controls the expression of the coding sequence by providing the recognition for RNA polymerase and other factors required for proper transcription. A pol II promoter includes a minimal promoter that is a short DNA sequence comprised of a TATA-box and optionally other sequences that serve to specify the site of transcription initiation, to which regulatory elements are added for control of expression. A

type 1 pol III promoter includes three cis-acting sequence elements downstream of the transcriptional start site: a) 5'sequence element (A block); b) an intermediate sequence element (I block); c) 3' sequence element (C block). A type 2 pol III promoter includes two essential cis-acting sequence elements downstream of the transcription start site: a) an A box (5' sequence element); and b) a B box (3' sequence element). A type 3 pol III promoter includes several cis-acting promoter elements upstream of the transcription start site, such as a traditional TATA box, proximal sequence element (PSE), and a distal sequence element (DSE).

[00161] An "enhancer" is a DNA sequence that can stimulate transcription activity and may be an innate element of the promoter or a heterologous element that enhances the level or tissue specificity of expression. It is capable of operating in either orientation (5'->3' or 3'->5'), and may be capable of functioning even when positioned either upstream or downstream of the promoter.

[00162] Promoters and/or enhancers may be derived in their entirety from a native gene, or be composed of different elements derived from different elements found in nature, or even be comprised of synthetic DNA segments. A promoter or enhancer may comprise DNA sequences that are involved in the binding of protein factors that modulate/control effectiveness of transcription initiation in response to stimuli, physiological or developmental conditions.

[00163] Non-limiting examples of promoters include SV40 early promoter, mouse mammary tumor virus LTR promoter; adenovirus major late promoter (Ad MLP); a herpes simplex virus (HSV) promoter, a cytomegalovirus (CMV) promoter such as the CMV immediate early promoter region (CMVIE), a rous sarcoma virus (RSV) promoter, pol II promoters, pol III promoters, synthetic promoters, hybrid promoters, and the like. In addition, sequences derived from non-viral genes, such as the murine metallothionein gene, will also find use herein. Exemplary constitutive promoters include the promoters for the following genes which encode certain constitutive or "housekeeping" functions: hypoxanthine phosphoribosyl transferase (HPRT), dihydrofolate reductase (DHFR), adenosine deaminase, phosphoglycerol kinase (PGK), pyruvate kinase, phosphoglycerol mutase, actin promoter, U6, and other constitutive promoters known to those of skill in the art. In addition, many viral promoters function constitutively in eukaryotic cells. These include: the early and late promoters of SV40; the long terminal repeats (LTRs) of Moloney Leukemia Virus and other retroviruses; and the thymidine kinase promoter of Herpes Simplex Virus, among many others.

Accordingly, any of the above-referenced constitutive promoters can be used to control transcription of a heterologous gene insert.

[00164] A "transgene" is used herein to conveniently refer to a nucleic acid sequence/polynucleotide that is intended or has been introduced into a cell or organism. Transgenes include any nucleic acid, such as a gene that encodes an inhibitory RNA or polypeptide or protein, and are generally heterologous with respect to naturally occurring AAV genomic sequences.

[00165] The term "transduce" refers to introduction of a nucleic acid sequence into a cell or host organism by way of a vector (*e.g.*, a viral particle). Introduction of a transgene into a cell by a viral particle is can therefore be referred to as "transduction" of the cell. The transgene may or may not be integrated into genomic nucleic acid of a transduced cell. If an introduced transgene becomes integrated into the nucleic acid (genomic DNA) of the recipient cell or organism it can be stably maintained in that cell or organism and further passed on to or inherited by progeny cells or organisms of the recipient cell or organism. Finally, the introduced transgene may exist in the recipient cell or host organism extra chromosomally, or only transiently. A "transduced cell" is therefore a cell into which the transgene has been introduced by way of transduction. Thus, a "transduced" cell is a cell into which, or a progeny thereof in which a transgene has been introduced. A transduced cell can be propagated, transgene transcribed and the encoded inhibitory RNA or protein expressed. For gene therapy uses and methods, a transduced cell can be in a mammal.

[00166] Transgenes under control of inducible promoters are expressed only or to a greater degree, in the presence of an inducing agent, (*e.g.*, transcription under control of the metallothionein promoter is greatly increased in presence of certain metal ions). Inducible promoters include responsive elements (REs) which stimulate transcription when their inducing factors are bound. For example, there are REs for serum factors, steroid hormones, retinoic acid and cyclic AMP. Promoters containing a particular RE can be chosen in order to obtain an inducible response and in some cases, the RE itself may be attached to a different promoter, thereby conferring inducibility to the recombinant gene. Thus, by selecting a suitable promoter (constitutive versus inducible; strong versus weak), it is possible to control both the existence and level of expression of a polypeptide in the genetically modified cell. If the gene encoding the polypeptide is under the control of an inducible promoter, delivery of the polypeptide *in situ* is triggered by exposing the genetically modified cell *in situ* to conditions

for permitting transcription of the polypeptide, *e.g.*, by intraperitoneal injection of specific inducers of the inducible promoters which control transcription of the agent. For example, *in situ* expression by genetically modified cells of a polypeptide encoded by a gene under the control of the metallothionein promoter, is enhanced by contacting the genetically modified cells with a solution containing the appropriate (*i.e.*, inducing) metal ions *in situ*.

[00167] A nucleic acid/transgene is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. A nucleic acid/transgene encoding and RNAi or a polypeptide, or a nucleic acid directing expression of a polypeptide may include an inducible promoter, or a tissue-specific promoter for controlling transcription of the encoded polypeptide. A nucleic acid operably linked to an expression control element can also be referred to as an expression cassette.

[00168] In certain embodiments, cell-type-specific or inducible promoters, enhancers and the like, are employed in the methods and uses described herein. Non-limiting examples of cell-type-specific promoters include those isolated from the genes from TMC1, TMC2, Espin, PCDH15, PTPRQ, TMHS (LHFPL5), MYO1A. Non-limiting examples of inducible promoters include DNA responsive elements for ecdysone, tetracycline, hypoxia and IFN.

[00169] In certain embodiments, an expression control element comprises a CMV enhancer. In certain embodiments, an expression control element comprises a beta actin promoter. In certain embodiments, an expression control element comprises a chicken beta actin promoter. In certain embodiments, an expression control element comprises a CMV enhancer and a chicken beta actin promoter.

[00170] As used herein, the terms "modify" or "variant" and grammatical variations thereof, mean that a nucleic acid, polypeptide or subsequence thereof deviates from a reference sequence. Modified and variant sequences may therefore have substantially the same, greater or less expression, activity or function than a reference sequence, but at least retain partial activity or function of the reference sequence. A particular type of variant is a mutant protein, which refers to a protein encoded by a gene having a mutation, *e.g.*, a missense or nonsense mutation.

[00171] A "nucleic acid" or "polynucleotide" variant refers to a modified sequence which has been genetically altered compared to wild-type. The sequence may be

genetically modified without altering the encoded protein sequence. Alternatively, the sequence may be genetically modified to encode a variant protein. A nucleic acid or polynucleotide variant can also refer to a combination sequence which has been codon modified to encode a protein that still retains at least partial sequence identity to a reference sequence, such as wild-type protein sequence, and also has been codon-modified to encode a variant protein. For example, some codons of such a nucleic acid variant will be changed without altering the amino acids of a protein encoded thereby, and some codons of the nucleic acid variant will be changed which in turn changes the amino acids of a protein encoded thereby.

[00172] The terms "protein" and "polypeptide" are used interchangeably herein. The "polypeptides" encoded by a "nucleic acid" or "polynucleotide" or "transgene" disclosed herein include partial or full-length native sequences, as with naturally occurring wild-type and functional polymorphic proteins, functional subsequences (fragments) thereof, and sequence variants thereof, so long as the polypeptide retains some degree of function or activity. Accordingly, in methods and uses of the invention, such polypeptides encoded by nucleic acid sequences are not required to be identical to the endogenous protein that is defective, or whose activity, function, or expression is insufficient, deficient or absent in a treated mammal.

[00173] Non-limiting examples of modifications include one or more nucleotide or amino acid substitutions (*e.g.*, about 1 to about 3, about 3 to about 5, about 5 to about 10, about 10 to about 15, about 15 to about 20, about 20 to about 25, about 25 to about 30, about 30 to about 40, about 40 to about 50, about 50 to about 100, about 100 to about 150, about 150 to about 200, about 200 to about 250, about 250 to about 500, about 500 to about 750, about 750 to about 1000 or more nucleotides or residues).

[00174] An example of an amino acid modification is a conservative amino acid substitution or a deletion. In particular embodiments, a modified or variant sequence retains at least part of a function or activity of the unmodified sequence (*e.g.*, wild-type sequence).

[00175] Another example of an amino acid modification is a targeting peptide introduced into a capsid protein of a viral particle. Peptides have been identified that target recombinant viral vectors, to the central nervous system, such as to distinct brain regions.

[00176] A recombinant virus so modified may preferentially bind to one type of tissue (*e.g.*, CNS tissue) over another type of tissue (*e.g.*, liver tissue). In certain embodiments,

a recombinant virus bearing a modified capsid protein may "target" brain vascular epithelia tissue by binding at level higher than a comparable, unmodified capsid protein. For example, a recombinant virus having a modified capsid protein may bind to brain vascular epithelia tissue at a level 50% to 100% greater than an unmodified recombinant virus.

[00177] A "nucleic acid fragment" is a portion of a given nucleic acid molecule. Deoxyribonucleic acid (DNA) in the majority of organisms is the genetic material while ribonucleic acid (RNA) is involved in the transfer of information contained within DNA into proteins. Fragments and variants of the disclosed nucleotide sequences and proteins or partial-length proteins encoded thereby are also encompassed by the present invention. By "fragment" or "portion" is meant a full length or less than full length of the nucleotide sequence encoding, or the amino acid sequence of, a polypeptide or protein. In certain embodiments, the fragment or portion is biologically functional (*i.e.*, retains 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or 100% of activity or function of wild-type).

[00178] A "variant" of a molecule is a sequence that is substantially similar to the sequence of the native molecule. For nucleotide sequences, variants include those sequences that, because of the degeneracy of the genetic code, encode the identical amino acid sequence of the native protein. Naturally occurring allelic variants such as these can be identified with the use of molecular biology techniques, as, for example, with polymerase chain reaction (PCR) and hybridization techniques. Variant nucleotide sequences also include synthetically derived nucleotide sequences, such as those generated, for example, by using site-directed mutagenesis, which encode the native protein, as well as those that encode a polypeptide having amino acid substitutions. Generally, nucleotide sequence variants of the invention will have at least 40%, 50%, 60%, to 70%, *e.g.*, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, to 79%, generally at least 80%, *e.g.*, 81%-84%, at least 85%, *e.g.*, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, to 98%, sequence identity to the native (endogenous) nucleotide sequence. In certain embodiments, the variant is biologically functional (*i.e.*, retains 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% or 100% of activity or function of wild-type).

[00179] "Conservative variations" of a particular nucleic acid sequence refers to those nucleic acid sequences that encode identical or essentially identical amino acid sequences. Because of the degeneracy of the genetic code, a large number of functionally

identical nucleic acids encode any given polypeptide. For instance, the codons CGT, CGC, CGA, CGG, AGA and AGG all encode the amino acid arginine. Thus, at every position where an arginine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded protein. Such nucleic acid variations are "silent variations," which are one species of "conservatively modified variations." Every nucleic acid sequence described herein that encodes a polypeptide also describes every possible silent variation, except where otherwise noted. One of skill in the art will recognize that each codon in a nucleic acid (except ATG, which is ordinarily the only codon for methionine) can be modified to yield a functionally identical molecule by standard techniques. Accordingly, each "silent variation" of a nucleic acid that encodes a polypeptide is implicit in each described sequence.

[00180] The term "substantial identity" of polynucleotide sequences means that a polynucleotide comprises a sequence that has at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, or 79%, or at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, or 89%, or at least 90%, 91%, 92%, 93%, or 94%, or even at least 95%, 96%, 97%, 98%, or 99% sequence identity, compared to a reference sequence using one of the alignment programs described using standard parameters. One of skill in the art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning, and the like. Substantial identity of amino acid sequences for these purposes normally means sequence identity of at least 70%, at least 80%, 90%, or even at least 95%.

[00181] The term "substantial identity" in the context of a polypeptide indicates that a polypeptide comprises a sequence with at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, or 79%, or 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, or 89%, or at least 90%, 91%, 92%, 93%, or 94%, or even, 95%, 96%, 97%, 98% or 99%, sequence identity to the reference sequence over a specified comparison window. An indication that two polypeptide sequences are identical is that one polypeptide is immunologically reactive with antibodies raised against the second polypeptide. Thus, a polypeptide is identical to a second polypeptide, for example, where the two peptides differ only by a conservative substitution.

[00182] By "mechanosensation" is meant a response to a mechanical stimulus. Touch, hearing, and balance of examples of the conversion of a mechanical stimulus into a

neuronal signal. Mechanosensory input is converted into a response to a mechanical stimulus through a process termed "mechanotransduction."

[00183] By "disease" is meant any condition or disorder that damages or interferes with the normal function of a cell, tissue, or organ. Examples of diseases include genetic disorders characterized by a loss of function in a protein that functions in mechanosensory transduction that is expressed, for example, in the inner ear of a subject. In another embodiment, the disease is Usher Syndrome (e.g., USH1) or age-related hearing loss. In one embodiment, a disease is an auditory disorder associated with a genetic defect, such as a defect in TMC1, TMC2, MY07A, USCH1 C, CDH23, PCDH15, SANS, CIB2, USH2A, VLGR1, WHKN, CLRN1, PDZD7, USH1C (e.g., harmonin-a, b, or c).

[00184] The terms "treat" and "treatment" refer to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent, inhibit, reduce, or decrease an undesired physiological change or disorder, such as the development, progression or worsening of the disorder. For purposes of this invention, beneficial or desired clinical results include, but are not limited to, alleviation of symptoms, diminishment of extent of disease, stabilizing a (*i.e.*, not worsening or progressing) symptom or adverse effect of disease, delay or slowing of disease progression, amelioration or palliation of the disease state, and remission (whether partial or total), whether detectable or undetectable. "Treatment" can also mean prolonging survival as compared to expected survival if not receiving treatment. Those in need of treatment include those already with the condition or disorder as well as those predisposed (*e.g.*, as determined by a genetic assay).

[00185] As used herein, "essentially free," in terms of a specified component, is used herein to mean that none of the specified component has been purposefully formulated into a composition and/or is present only as a contaminant or in trace amounts. The total amount of the specified component resulting from any unintended contamination of a composition is therefore well below 0.05%, preferably below 0.01%. Most preferred is a composition in which no amount of the specified component can be detected with standard analytical methods.

[00186] As used herein the specification, "a" or "an" may mean one or more. As used herein in the claim(s), when used in conjunction with the word "comprising," the words "a" or "an" may mean one or more than one.

[00187] The use of the term "or" in the claims is used to mean "and/or" unless explicitly indicated to refer to alternatives only or the alternatives are mutually exclusive, although the disclosure supports a definition that refers to only alternatives and "and/or." As used herein "another" may mean at least a second or more.

[00188] Throughout this application, the term "about" is used to indicate that a value includes the inherent variation of error for the device, the inherent variation in the method being employed to determine the value, the variation that exists among the study subjects, or a value that is within 10% of a stated value.

VI. Kits

[00189] The invention provides kits with packaging material and one or more components therein. A kit typically includes a label or packaging insert including a description of the components or instructions for use *in vitro*, *in vivo*, or *ex vivo*, of the components therein. A kit can contain a collection of such components, *e.g.*, a nucleic acid, recombinant vector, and/or viral particles.

[00190] A kit refers to a physical structure housing one or more components of the kit. Packaging material can maintain the components sterilely, and can be made of material commonly used for such purposes (*e.g.*, paper, corrugated fiber, glass, plastic, foil, ampules, vials, tubes, etc.).

[00191] Labels or inserts can include identifying information of one or more components therein, dose amounts, clinical pharmacology of the active ingredient(s) including mechanism of action, pharmacokinetics and pharmacodynamics. Labels or inserts can include information identifying manufacturer, lot numbers, manufacture location and date, expiration dates. Labels or inserts can include information identifying manufacturer information, lot numbers, manufacturer location and date. Labels or inserts can include information on a disease for which a kit component may be used. Labels or inserts can include instructions for the clinician or subject for using one or more of the kit components in a method, use, or treatment protocol or therapeutic regimen. Instructions can include dosage amounts, frequency or duration, and instructions for practicing any of the methods, uses, treatment protocols or prophylactic or therapeutic regimes described herein.

[00192] Labels or inserts can include information on any benefit that a component may provide, such as a prophylactic or therapeutic benefit. Labels or inserts can include information on potential adverse side effects, complications or reactions, such as warnings to the subject or clinician regarding situations where it would not be appropriate to use a particular composition. Adverse side effects or complications could also occur when the subject has, will be or is currently taking one or more other medications that may be incompatible with the composition, or the subject has, will be or is currently undergoing another treatment protocol or therapeutic regimen which would be incompatible with the composition and, therefore, instructions could include information regarding such incompatibilities.

[00193] Labels or inserts include "printed matter," *e.g.*, paper or cardboard, or separate or affixed to a component, a kit or packing material (*e.g.*, a box), or attached to an ampule, tube or vial containing a kit component. Labels or inserts can additionally include a computer readable medium, such as a bar-coded printed label, a disk, optical disk such as CD- or DVD-ROM/RAM, DVD, MP3, or an electrical storage media such as RAM and ROM or hybrids of these such as magnetic/optical storage media, FLASH memory, hybrids and memory type cards.

VII. Examples

[00194] The following examples are included to demonstrate preferred embodiments of the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples which follow represent techniques discovered by the inventor to function well in the practice of the invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention.

Example 1 – Therapeutic delivery to the Globus Pallidus

[00195] Libraries of peptide-modified AAV capsid variants were generated by insertion of random peptides into AAV1, AAV2, and AAV9 by insertion of random targeting peptide sequences at position 590 of AAV1 capsid, position 587 of AAV2 capsid, and position 588 of AAV9 capsid, respectively. These libraries were screened in rhesus macaques for

variants that broadly transduce deep brain structures and cerebral cortex following infusion into the globus pallidus (FIGS. 38A-38C). Six variants (Table 1) were selected for further validation in a fluorescence experiment in two additional rhesus macaques. At least five capsid variants transduced deep brain structures affected in HD and PD. Transductions of AAV1.THTD.mTFP1 was widespread in caudate/putamen. AAV1.THTD.mTFP1 also transduced a large number of cortical layer V/VI projection neurons (FIG. 39), which are highly vulnerable in HD. These AAV capsids can be used to package a variety of therapeutic payloads for HD and PD. Liver transduction for these variants was below the limit of detection (FIG. 40) (AAV-BLD010 is AAV1.THTD.mTFP1; AAV-BLD011 is AAV2.SGGR.NES_mNG; AAV-BLD012 is AAV2.SKGQ.mRuby3).

Table 1. Capsid targeting peptides.

Serotype	DNA	Peptide
AAV1	TCTAGTGGGGATTCGACGAAG (SEQ ID NO: 7)	SSGDSTK (SEQ ID NO: 8)
AAV1	ACGCATACTGATCATACGCGT (SEQ ID NO: 9)	THTDDTR (SEQ ID NO: 10)
AAV2	TCTAAGGGGCAGGGGACGGGG (SEQ ID NO: 11)	SKGQGTG (SEQ ID NO: 12)
AAV2	TCGGGGGGGAGGTATGCGGAG (SEQ ID NO: 13)	SGGRYAE (SEQ ID NO: 14)
AAV9	TCTTTGGGTAGTACTAAGGTG (SEQ ID NO: 15)	SLGSTKV (SEQ ID NO: 16)
AAV9	AAGATGTTGAGTGTTGTTAAT (SEQ ID NO: 17)	KMLSVVN (SEQ ID NO: 18)

[00196] To further analyze the selected variants, the AAV1 variant with the THTDDTR targeting peptide (SEQ ID NO: 20) was packaged with a mTFP1 expression cassette (AAV1.THTD.mTFP1), the AAV2 variant with the SGGRYAE targeting peptide (SEQ ID NO: 22) was packaged with a mNeonGreen expression cassette (AAV2.SGGR.NES_mNG), the AAV2 variant with the SKGQGTG targeting peptide (SEQ ID NO: 21) was packaged with a mRuby3 expression cassette (AAV2.SKGQ.mRuby3), the AAV9 variant with the KMLSVVN targeting peptide (SEQ ID NO: 24) was packaged with a Halo-tag expression cassette (AAV9KMLS.Halo-tag), and the AAV9 variant with the SLGSTKV targeting peptide (SEQ ID NO: 23) was packaged with an mBFP2 expression cassette (AAV9SLGS.mBFP2).

[00197] Two adult rhesus macaques were injected with a dose of 7.5E10 total vg of certain capsid variants in the globus pallidus. Imaging studies were performed to determine the biodistribution and trafficking properties of each of the capsids following delivery to the globus pallidus. FISH images of 10um z-stacks of tissue were used to localize mRNA from each fluorophore specific to their respective variant capsid (FIGS. 1, 4, 7, 9, 11, 12, 15, 17, 18, 20, and 25). Epifluorescence images were used to localize the protein from each fluorophore specific to their respective variant capsid (FIGS. 2, 3, 5, 6, 8, 10, 13, 14, 16, 17, 19, 21-24, and 26).

Example 2 – Mouse injection of globus pallidus vectors

[00198] Three mice received injections of different viruses to each hemisphere of a designated anatomical structure: the thalamus, the internal capsule/medial globus pallidus (IC/MGP), or deep cerebellar nuclei (DCN) represented in Table 2. The different viruses were: RVC1023, pmAAV2.SGGR.NeonGreen; RVC1024, pmAAV2.SKGO.mRuby; and RVC1025, pmAAV1.THTD.mTFP. Mice were necropsied 22 days post-injection and assessed by epifluorescent microscopy (FIGS. 27-37).

Table 2. Mouse injection paradigm of Globus Pallidus vectors.

Mouse ID	ROA	Left	Right	vg/hemisphere
4279	Thalamus	NeonGreen	mRuby	5.00E+09
4280	Thalamus	mRuby	TFP	5.00E+09
4281	Thalamus	TFP	NeonGreen	5.00E+09
4282	IC/MGP	NeonGreen	mRuby	4.00E+09
4285	IC/MGP	mRuby	TFP	4.00E+09
4292	IC/MGP	TFP	NeonGreen	4.00E+09
4293	IC/MGP	mRuby	TFP	4.00E+09
A16	DCN	NeonGreen	mRuby	4.00E+09
A17	DCN	mRuby	TFP	4.00E+09
A18	DCN	TFP	NeonGreen	4.00E+09

* * *

[00199] All of the methods disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the methods and in the steps or in the sequence of steps of the method described herein without departing from the

concept, spirit and scope of the invention. More specifically, it will be apparent that certain agents which are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims.

WHAT IS CLAIMED IS:

1. A modified adeno-associated virus (AAV) capsid protein comprising a targeting peptide that targets a viral vector comprising the modified AAV capsid protein to a distinct organ or brain structure, wherein the targeting peptide is three to ten amino acids in length.

2. The modified AAV capsid protein of claim 1, wherein the capsid protein has a sequence according to any one of SEQ ID NOs: 20, 19, and 21-24.

3. The modified AAV capsid protein of claim 1, wherein the modified AAV capsid protein is a modified AAV1 capsid protein, a modified AAV2 capsid protein, or a modified AAV9 capsid protein.

4. The modified AAV capsid protein of claim 1, wherein the modified AAV capsid protein is derived from an AAV1 capsid protein (see SEQ ID NO: 1), wherein the targeting peptide is inserted after residue 590 of the AAV1 capsid protein.

5. The modified AAV capsid protein of claim 4, wherein the targeting peptide is flanked by linker sequences, wherein the linker sequences on each side of the targeting peptides are two or three amino acids long.

6. The modified AAV capsid protein of claim 5, wherein the linker sequences are SSA on the N-terminal side of the targeting peptide and AS on the C-terminal side of the targeting peptide.

7. The modified AAV capsid protein of claim 6, wherein the modified AAV1 capsid protein has a sequence at least 95% identical to SEQ ID NO: 4.

8. The modified AAV capsid protein of claim 1, wherein the modified AAV capsid protein is derived from an AAV2 capsid protein (see SEQ ID NO: 2), wherein the targeting peptide is inserted after residue 587 of the AAV2 capsid protein.

9. The modified AAV capsid protein of claim 8, wherein the targeting peptide is flanked by linker sequences, wherein the linker sequences on each side of the targeting peptides are two or three amino acids long.

10. The modified AAV capsid protein of claim 9, wherein the linker sequences are AAA on the N-terminal side of the targeting peptide and AA on the C-terminal side of the targeting peptide.

11. The modified AAV capsid protein of claim 10, wherein the modified AAV2 capsid protein has a sequence at least 95% identical to SEQ ID NO: 5.

12. The modified AAV capsid protein of claim 1, wherein the modified AAV capsid protein is derived from an AAV9 capsid protein (see SEQ ID NO: 3), wherein the targeting peptide is inserted after residue 588 of the AAV9 capsid protein.

13. The modified AAV capsid protein of claim 12, wherein the targeting peptide is flanked by linker sequences, wherein the linker sequences on each side of the targeting peptides are two or three amino acids long.

14. The modified AAV capsid protein of claim 13, wherein the linker sequences are AAA on the N-terminal side of the targeting peptide and AS on the C-terminal side of the targeting peptide.

15. The modified AAV capsid protein of claim 14, wherein the modified AAV9 capsid protein has a sequence at least 95% identical to SEQ ID NO: 6.

16. The modified AAV capsid protein of claim 1, wherein the target peptide comprises a sequence up to ten amino acids in length having therein an amino acid sequence selected from the group consisting of SEQ ID NOs: 8, 10, 12, 14, 16, and 18.

17. The modified AAV capsid protein of claim 16, wherein the targeting peptide is seven amino acids in length.

18. The modified AAV capsid protein of any one of claims 1-17, wherein the distinct brain structure is the globus pallidus, putamen, internal capsule, caudate, claustrum, substantial nigra, motor cortex, insula, temporal cortex, thalamus, hippocampus, subiculum, and deep cerebellar nuclei.

19. A nucleic acid comprising a sequence encoding the modified capsid protein of any one of claims 1-18.

20. A recombinant adeno-associated virus (rAAV) virus comprising the modified capsid protein of any one of claims 1-18.

21. A viral vector comprising a nucleic acid encoding the modified capsid protein of any one of claims 1-18.

22. The viral vector of claim 21, further comprising a nucleic acid sequence encoding a nucleic acid of interest.

23. The viral vector of claim 22, wherein the nucleic acid of interest is a therapeutic agent.

24. The viral vector of claim 23, wherein the therapeutic agent is an protein or an RNAi molecule.

25. A cell comprising the viral vector of any one of claims 21-24.

26. The cell of claim 25, wherein the cell is a mammalian cell.

27. The cell of claim 25, wherein the cell is a human cell.

28. The cell of claim 25, wherein the cell is in vitro.

29. The cell of claim 25, wherein the cell is in vivo.

30. A pharmaceutical composition comprising the viral vector of claim 20 and a pharmaceutically acceptable carrier.

31. A method to deliver an agent to a distinct brain structure of a subject, comprising administering the virus of claim 20 to the subject.

32. The method of claim 31, wherein the agent is an siRNA, shRNA, miRNA, non-coding RNA, lncRNA, therapeutic protein, or CRISPR system.

33. The method of any one of claims 31-32, wherein the administration is to the central nervous system.

34. The method of claim 33, wherein the administration is to a cisterna magna, an intraventricular space, an ependyma, a brain ventricle, a subarachnoid space, and/or an intrathecal space.

35. The method of claim 33, wherein the administration is to the globus pallidus.

36. The method of any one of claims 31-32, wherein the administration is systemic.

37. The method of any one of claims 31-36, wherein a plurality of viral particles are administered.

38. The method of claim 37, wherein the virus is administered at a dose of about 1×10^{6} to about 1×10^{18} vector genomes per kilogram (vg/kg).

39. The method of claim 37, wherein the virus is administered at a dose from about $1x10^{7}$-$1x10^{17}$, about $1x10^{8}$-$1x10^{16}$, about $1x10^{9}$-$1x10^{15}$, about $1x10^{10}$-$1x10^{14}$, about $1x10^{10}$-$1x10^{13}$, about $1x10^{10}$-$1x10^{13}$, about $1x10^{10}$-$1x10^{11}$, about $1x10^{11}$-$1x10^{12}$, about $1x10^{12}$-$x10^{13}$, or about $1x10^{13}$-$1X10^{14}$ vg/kg of the patient.

40. The method of any one of claims 31-39, wherein the subject is human.

41. A method of treating a disease in a mammal comprising administering the virus of claim 20 to the mammal.

42. The method of claim 41, wherein the disease is a neurodegenerative disease.

43. The method of claim 42, wherein the neurodegenerative disease is Huntington's disease, ALS, hereditary spastic hemiplegia, primary lateral sclerosis, spinal muscular atrophy, Kennedy's disease, Alzheimer's disease, a polyglutamine repeat disease, or Parkinson's disease.

44. The method of claim 41, wherein the mammal is human.

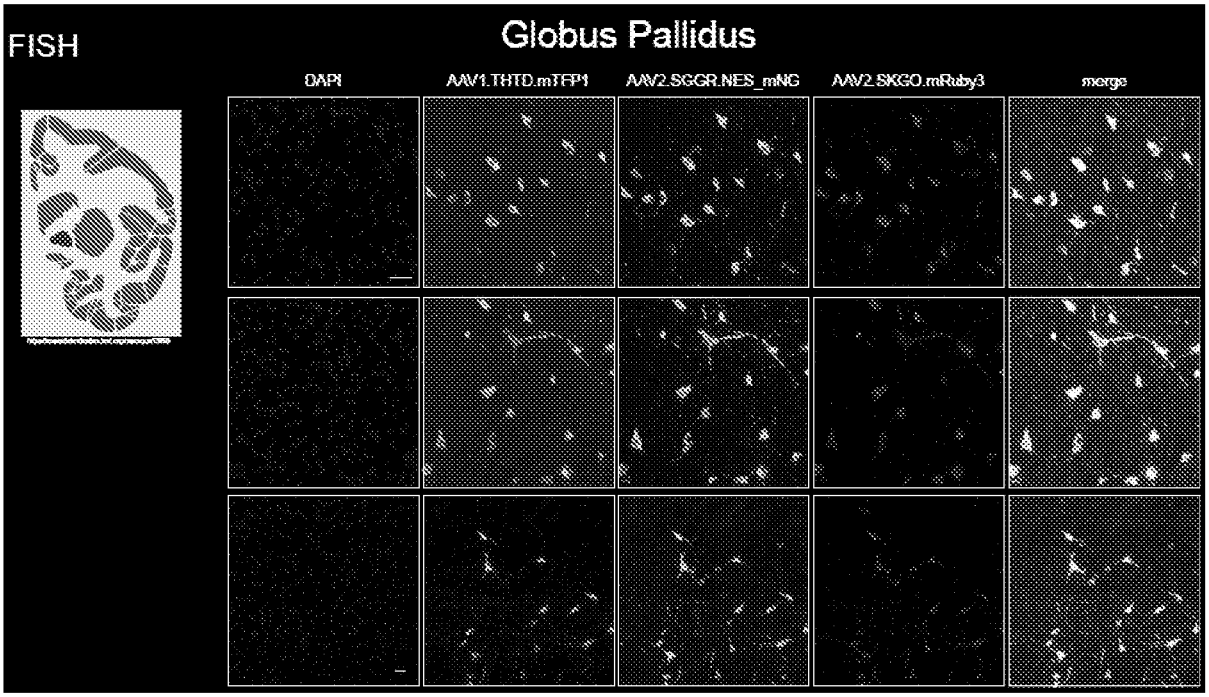

FISH
Globus Pallidus
DAPI
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 1

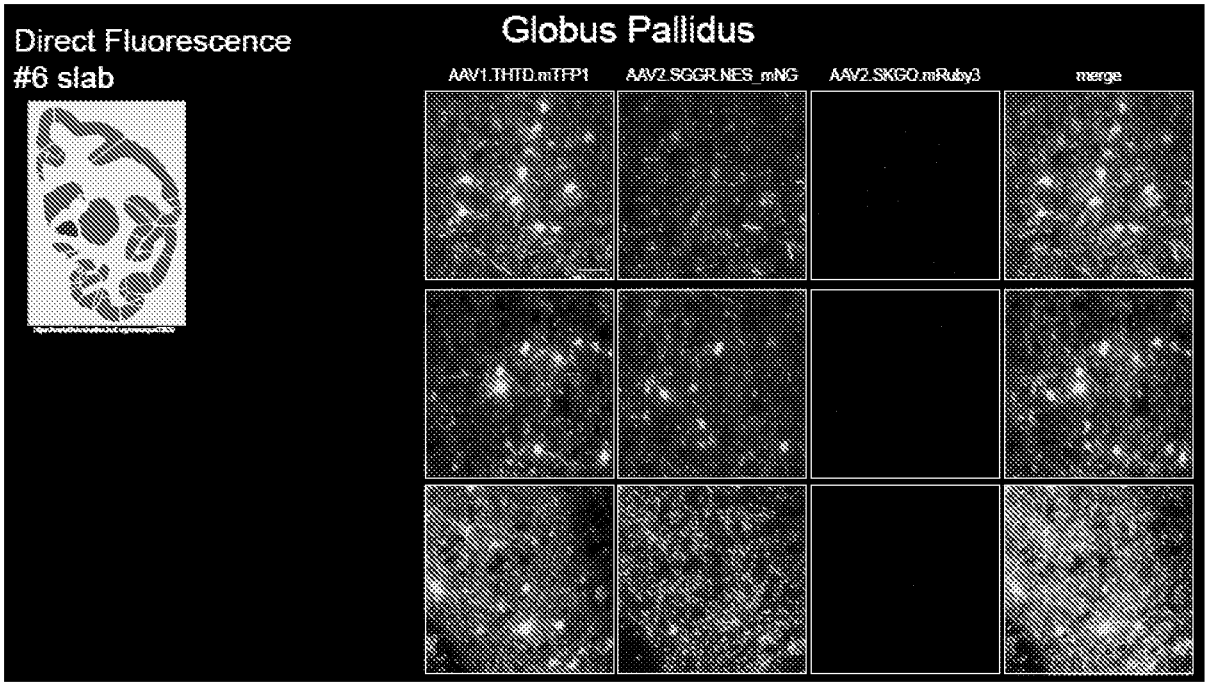

Direct Fluorescence
#6 slab
Globus Pallidus
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 2

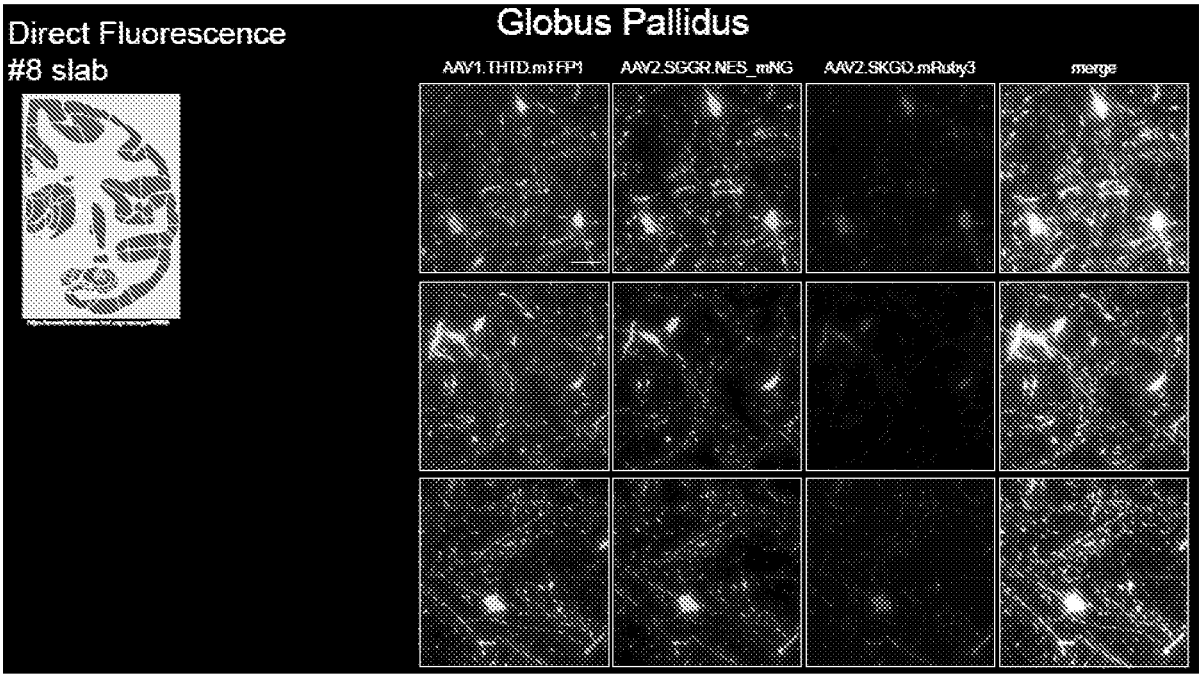

Direct Fluorescence
#8 slab
Globus Pallidus
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 3

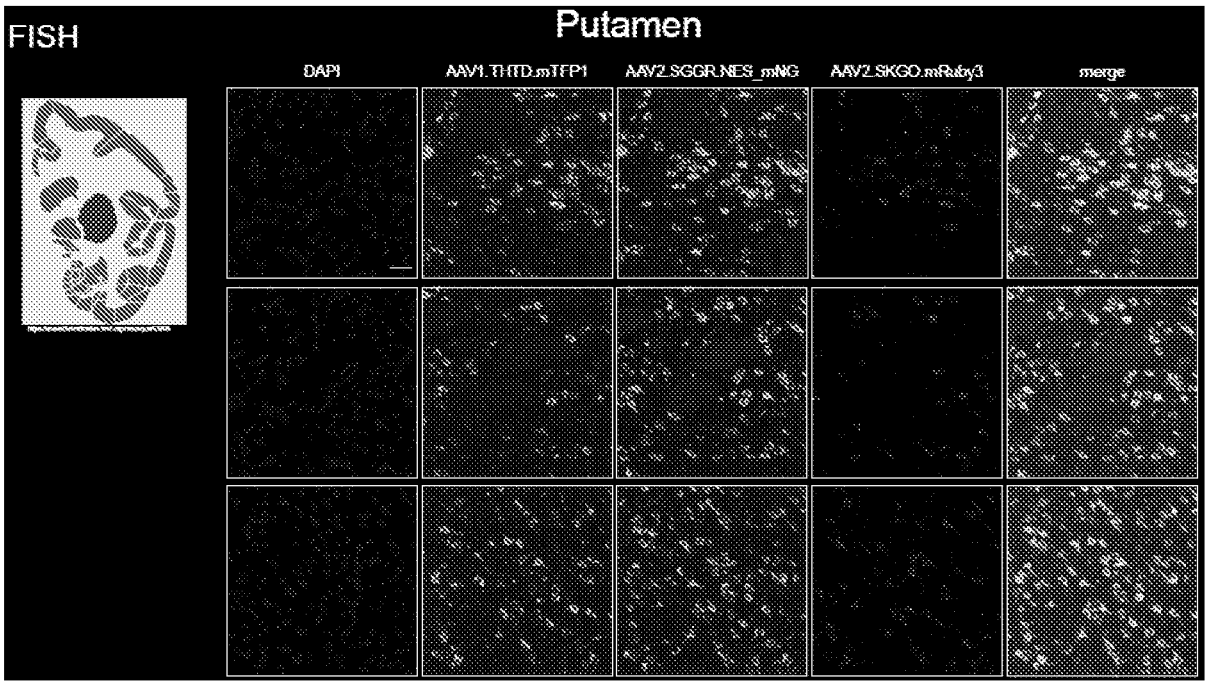

FISH
Putamen
DAPI
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 4

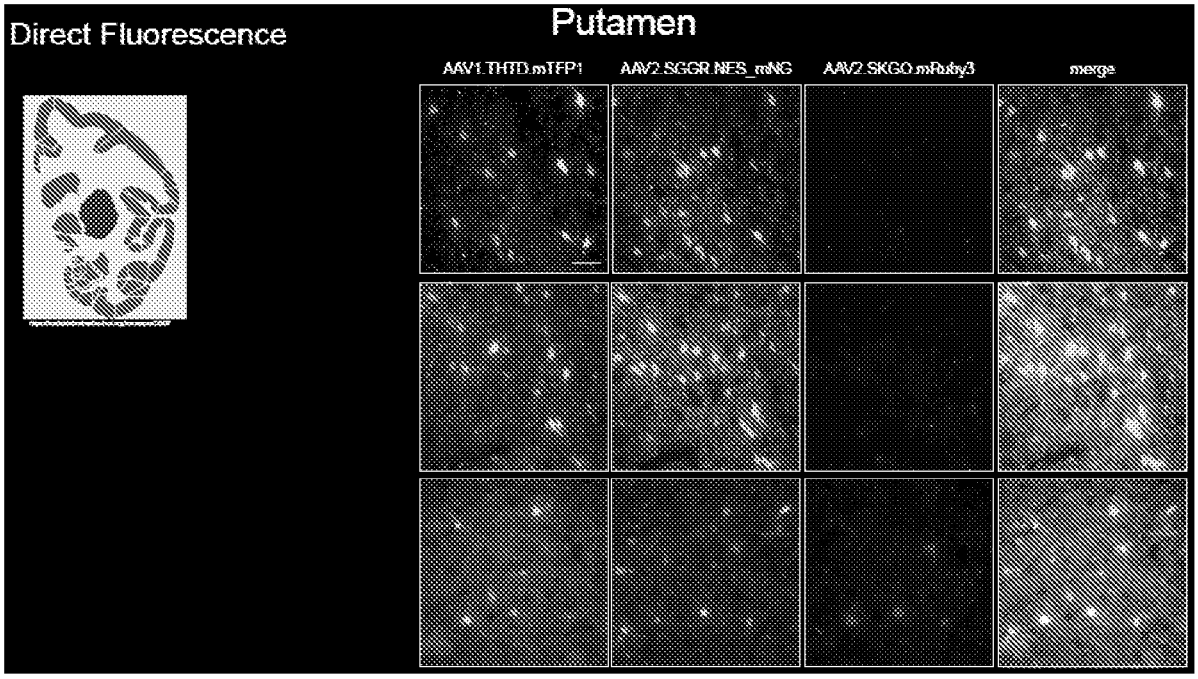

Direct Fluorescence
Putamen
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 5

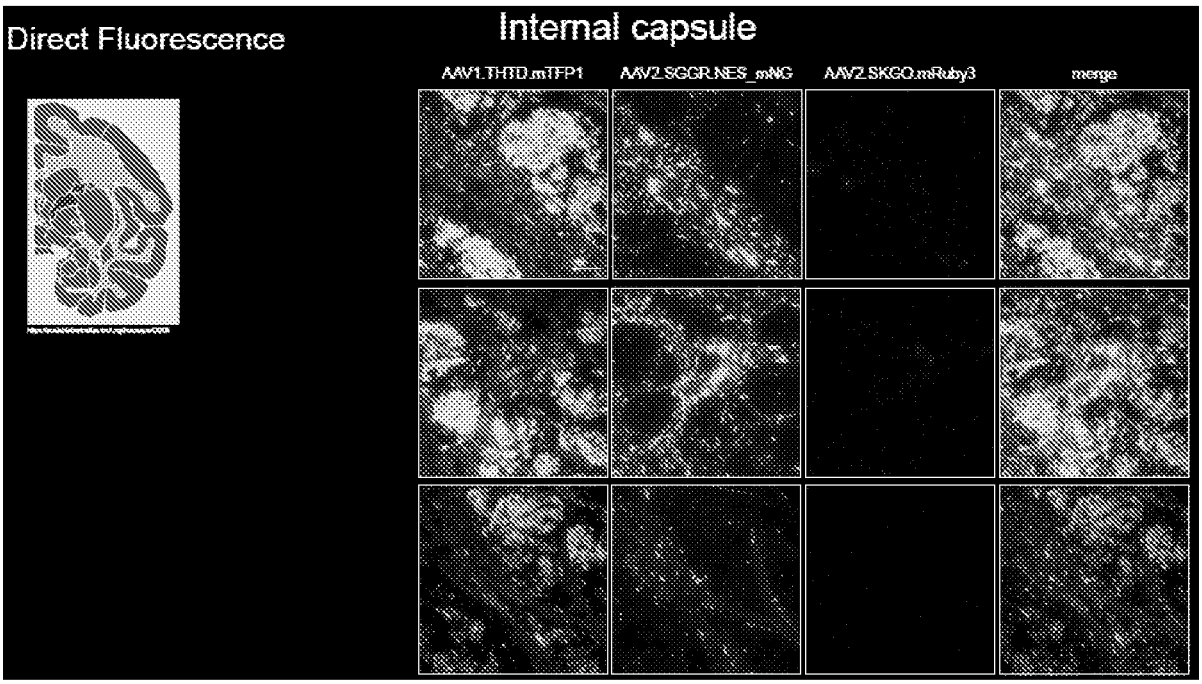

Direct Fluorescence
Internal capsule
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 6

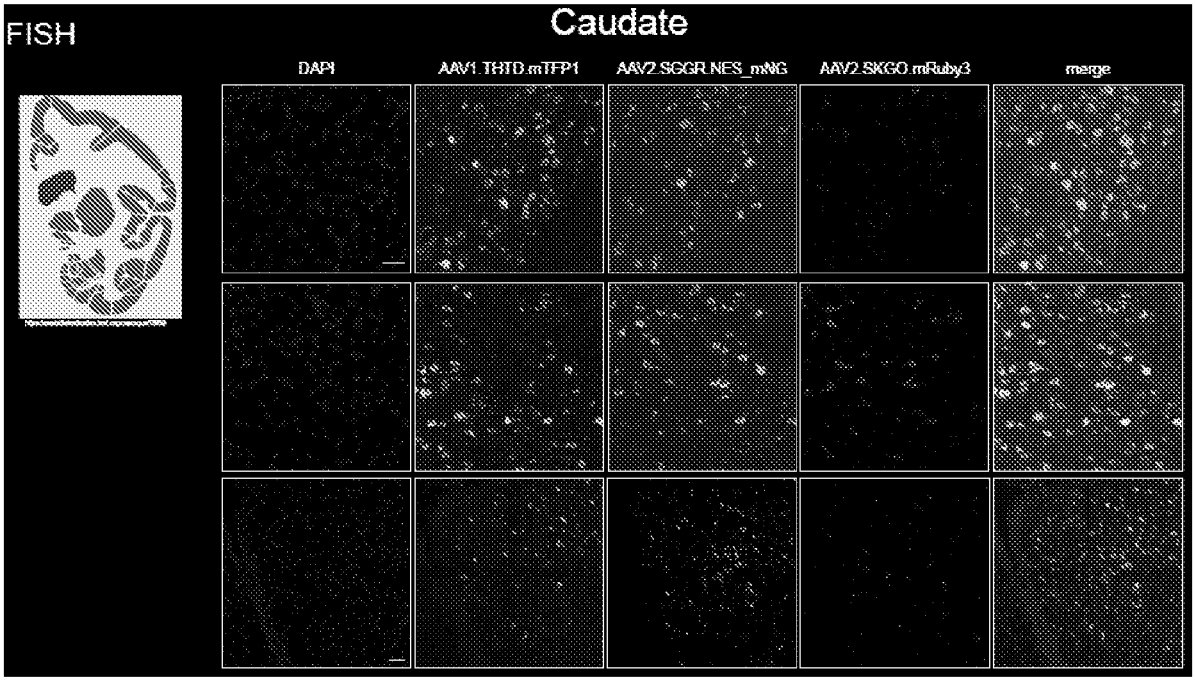

FISH
Caudate
DAPI
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 7

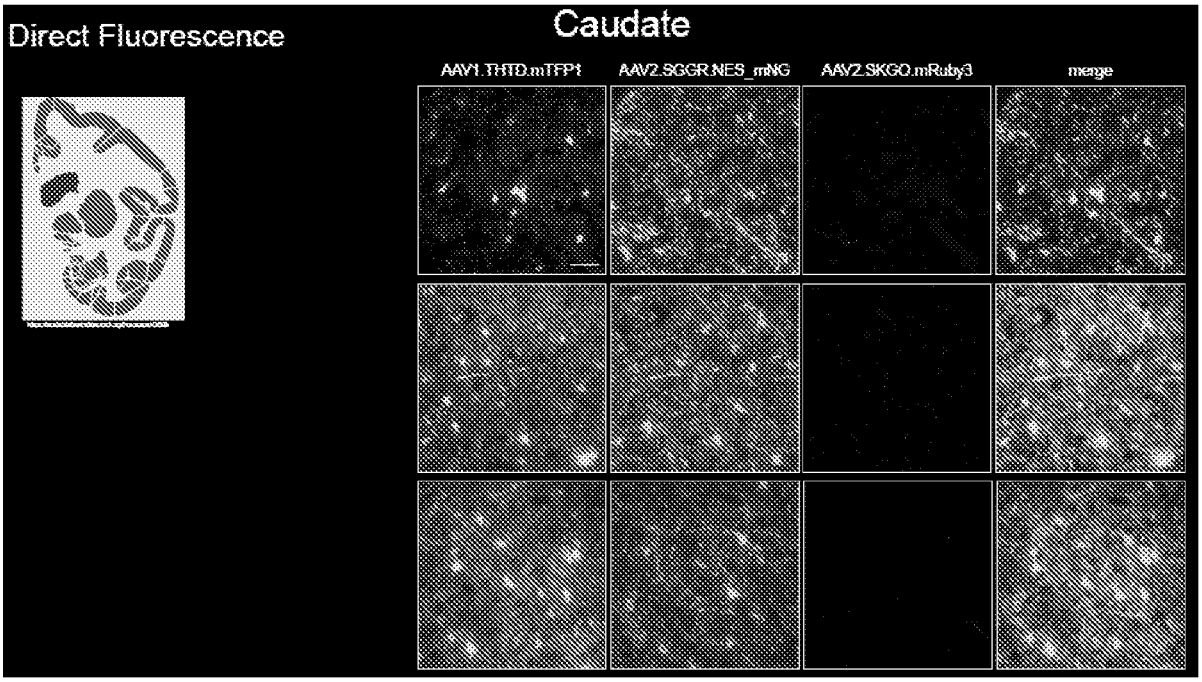

Direct Fluorescence
Caudate
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 8

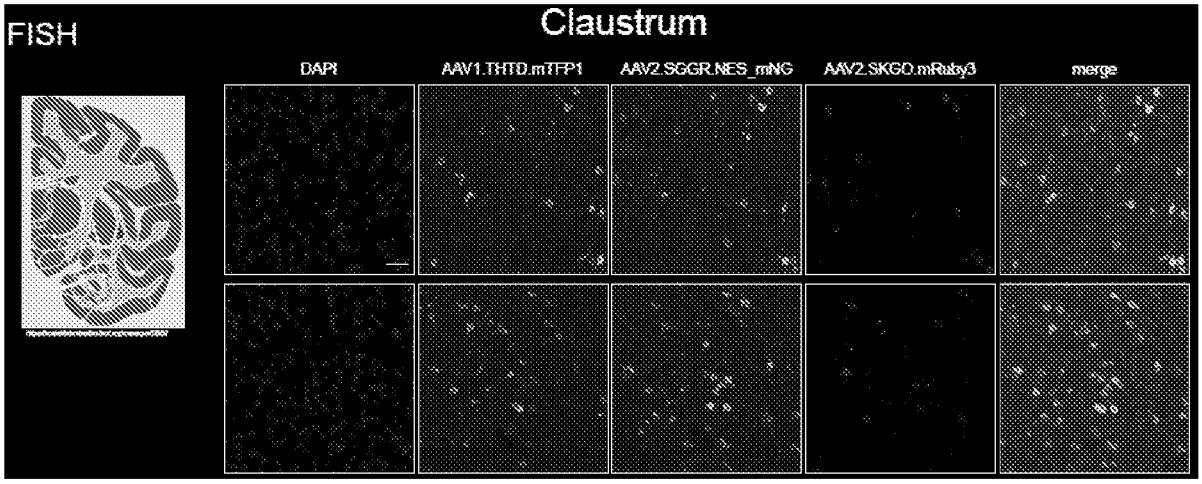

FISH
Claustrum
DAPI
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 9

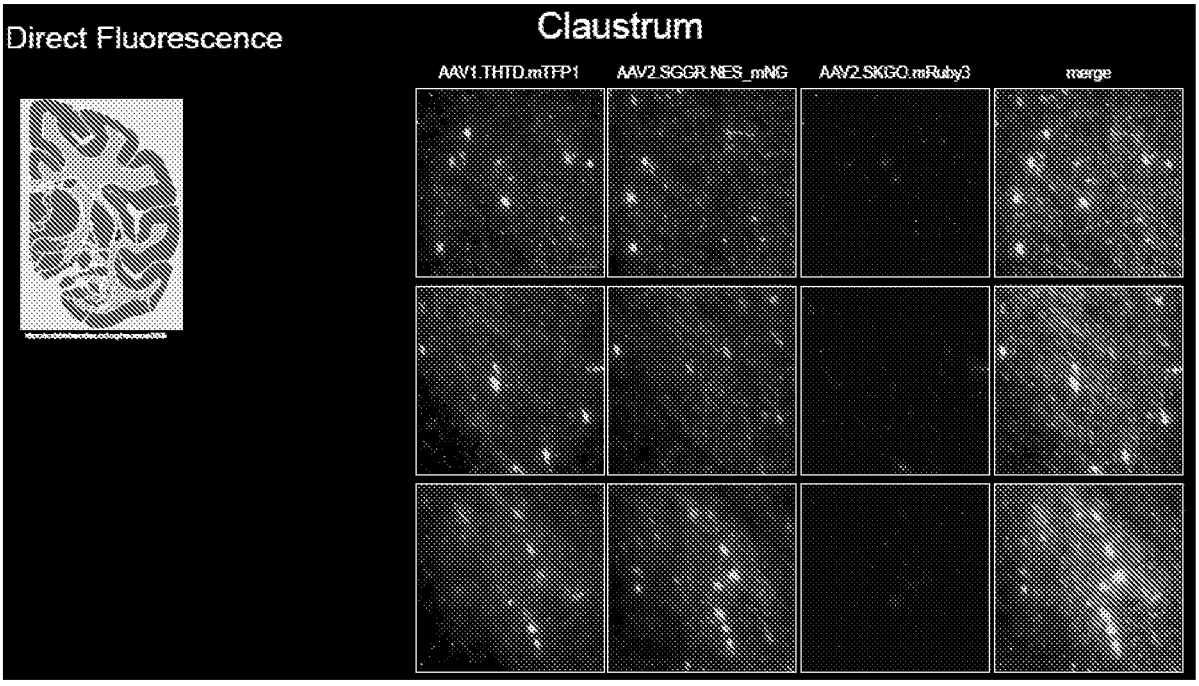

Direct Fluorescence
Claustrum
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 10

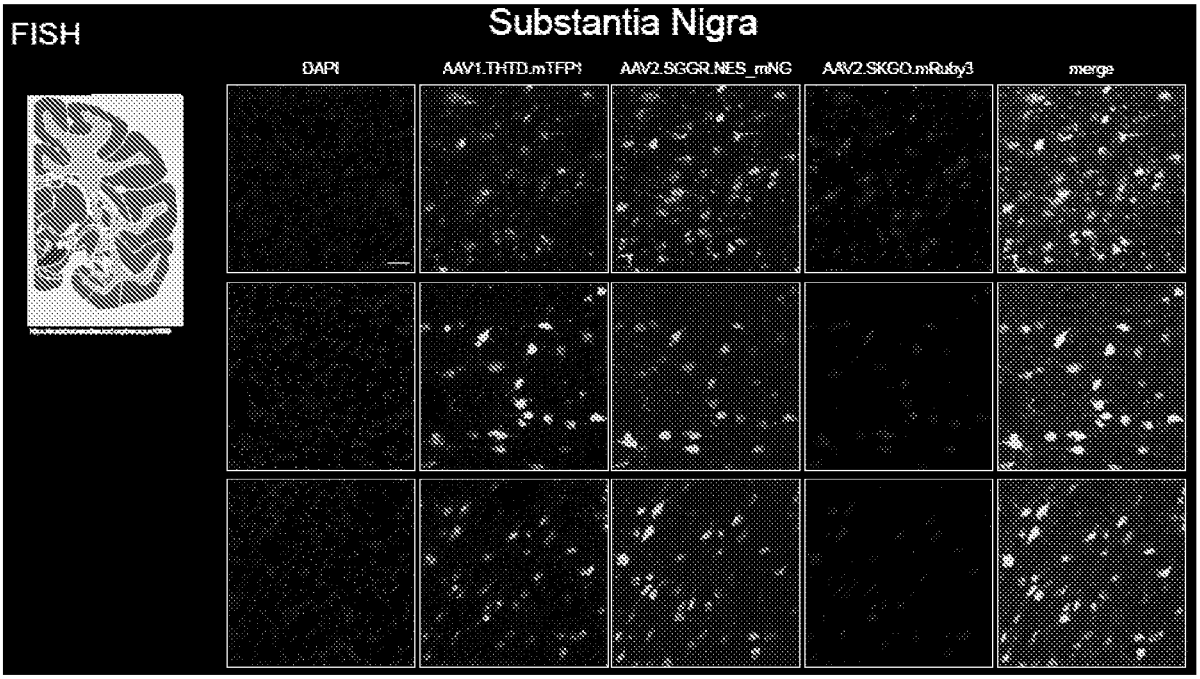

FISH
Substantia Nigra
DAPI
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 11

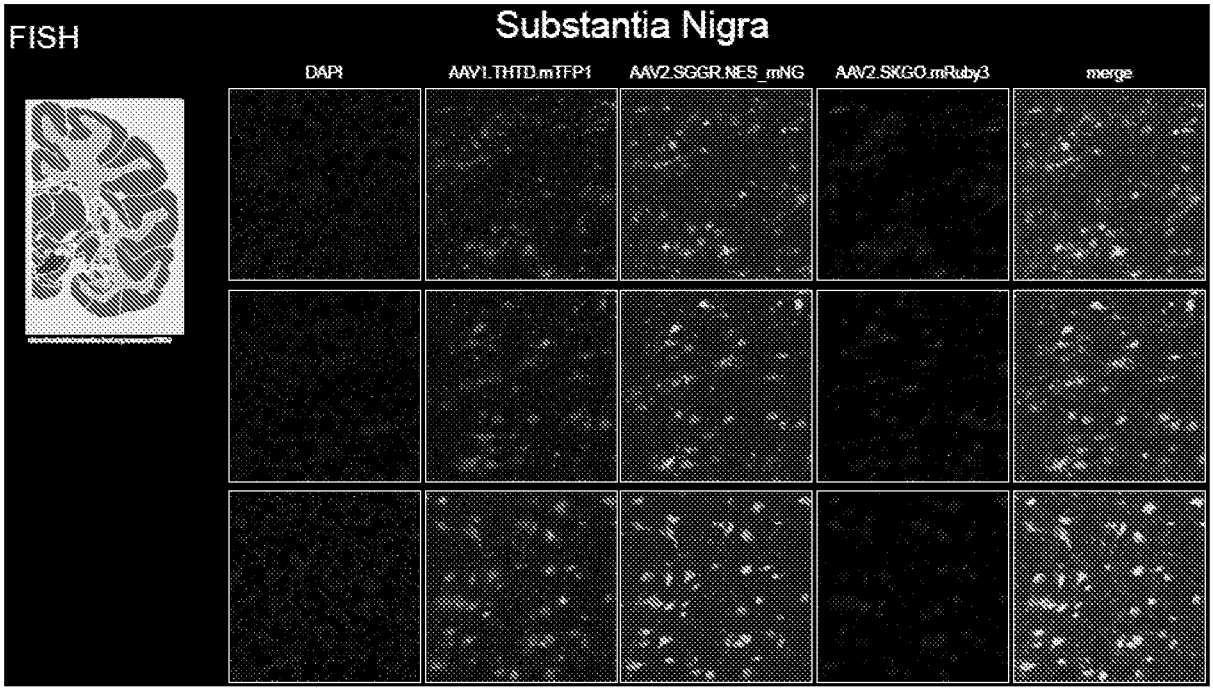

FISH
Substantia Nigra
DAPI
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 12

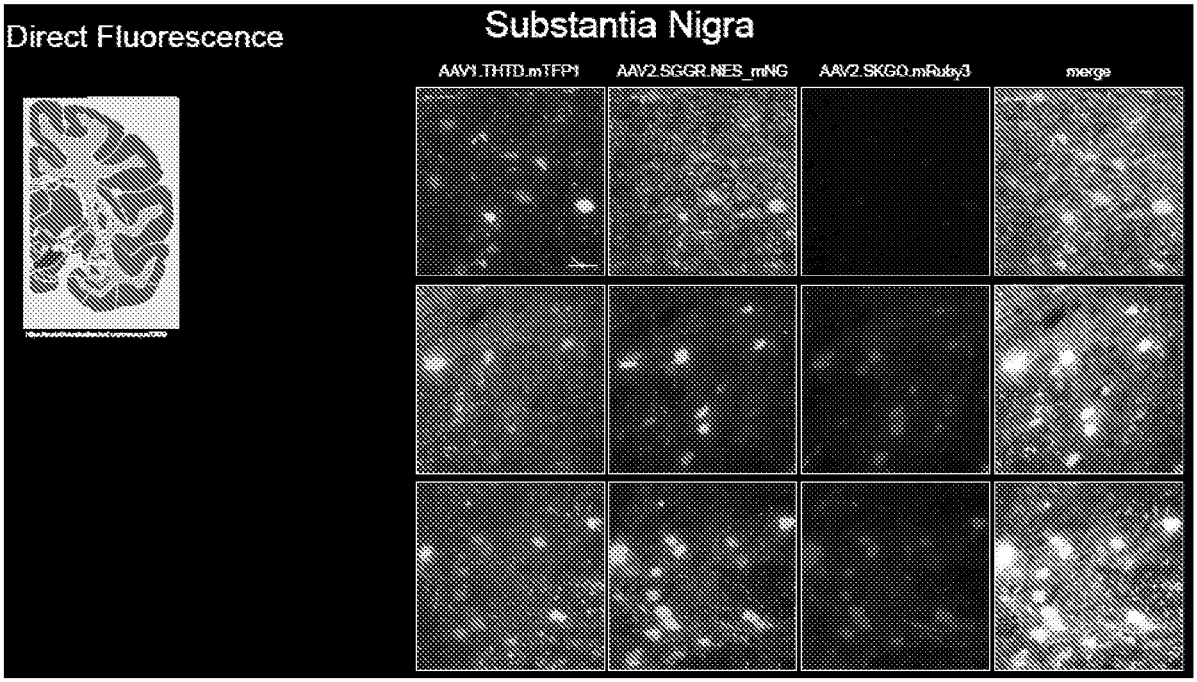

Direct Fluorescence
Substantia Nigra
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 13

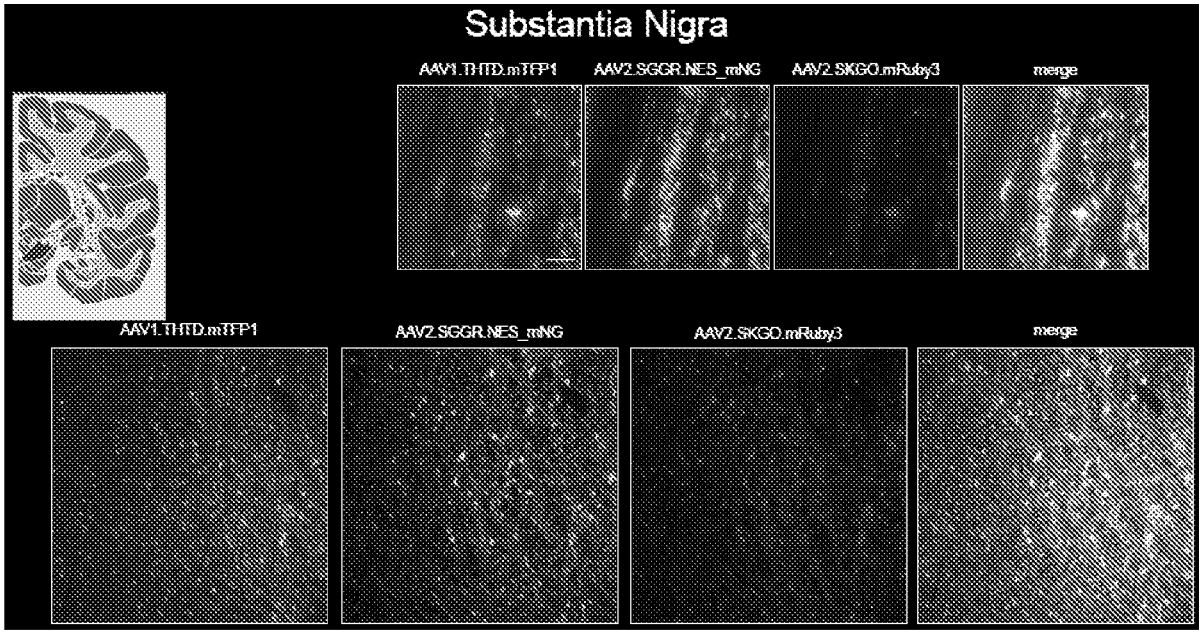

Substantia Nigra
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 14

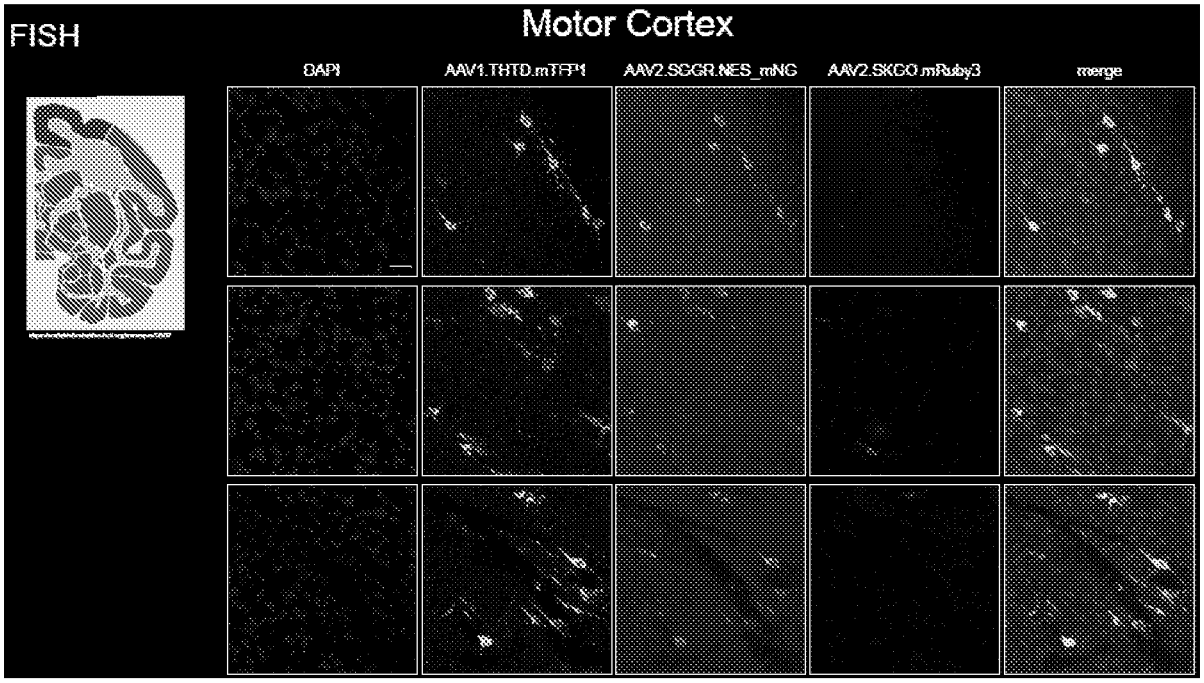

FISH
Motor Cortex
DAPI
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 15

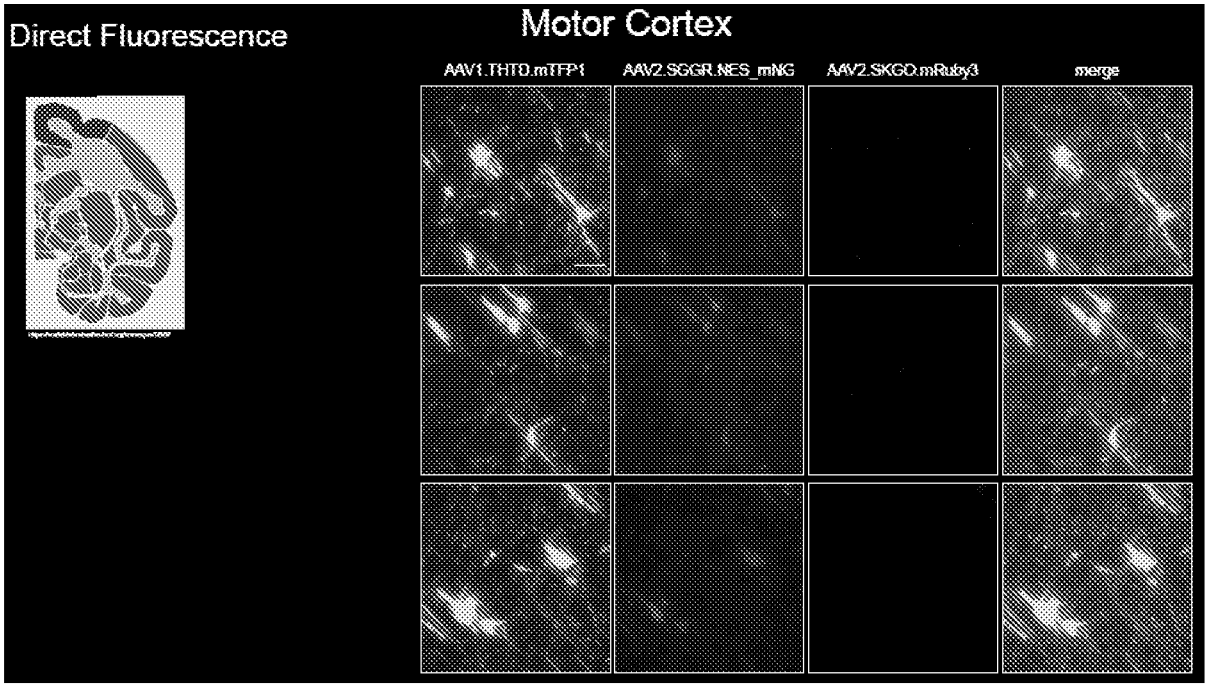

Direct Fluorescence
Motor Cortex
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 16

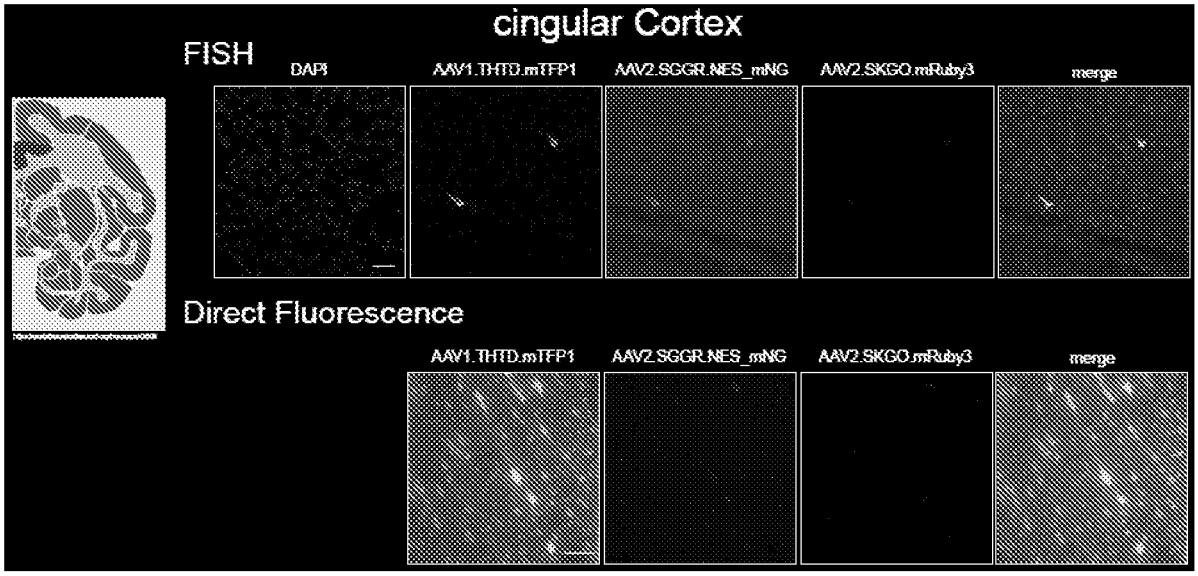

cingular Cortex
FISH
DAPI
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge
Direct Fluorescence
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 17

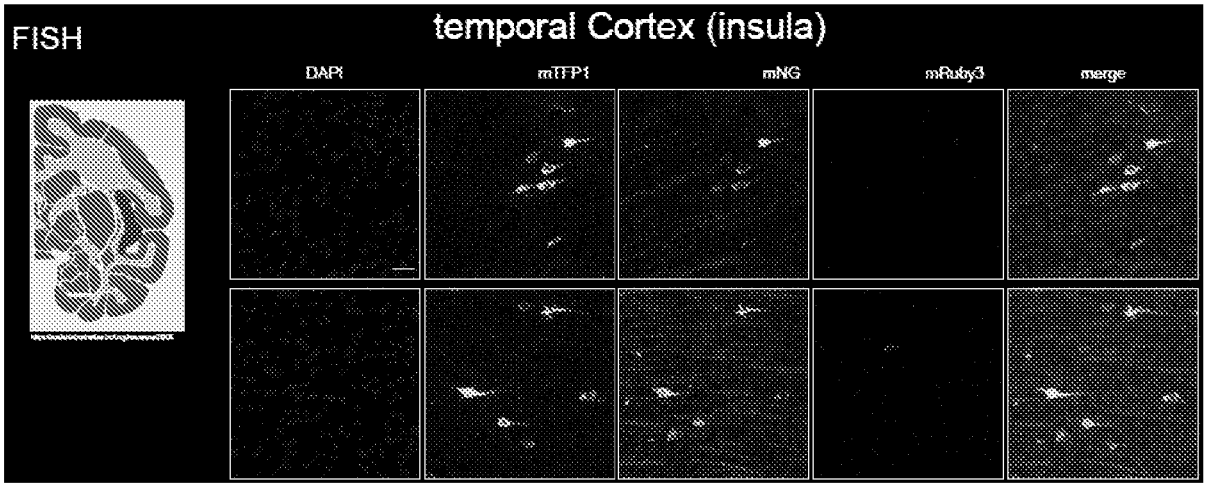

FISH
temporal Cortex (insula)
DAPI
mTFP1
mNG
mRuby3
merge

FIG. 18

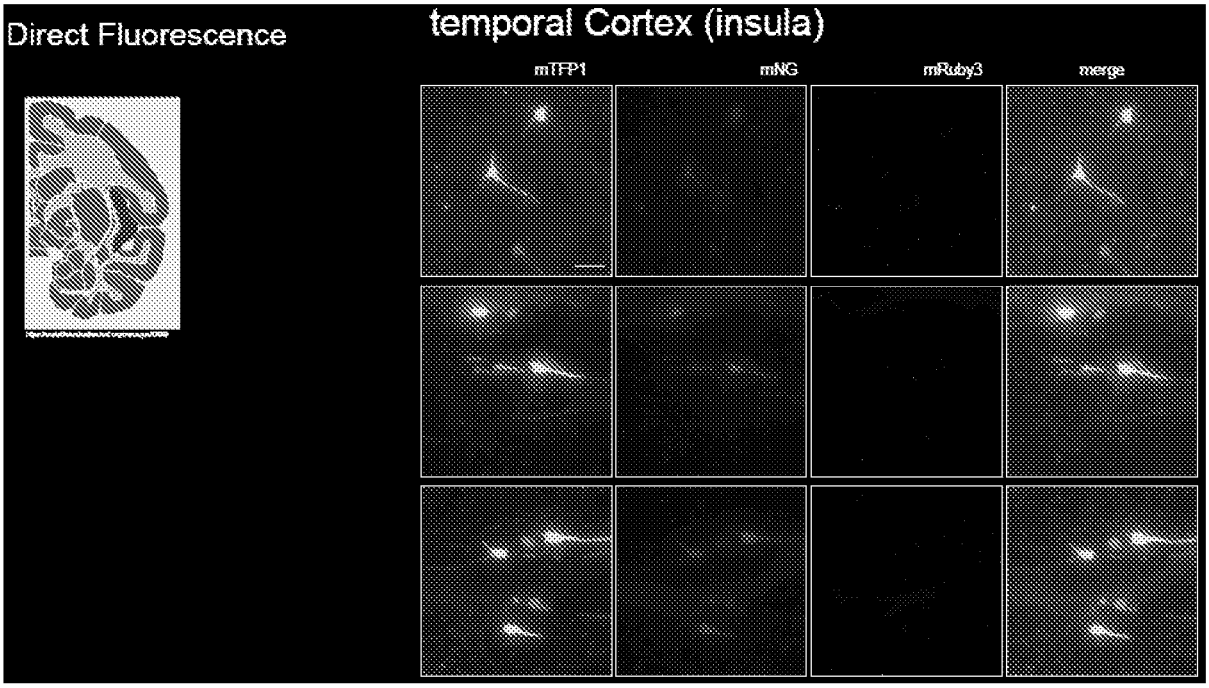

Direct Fluorescence
temporal Cortex (insula)
mTFP1
mNG
mRuby3
merge

FIG. 19

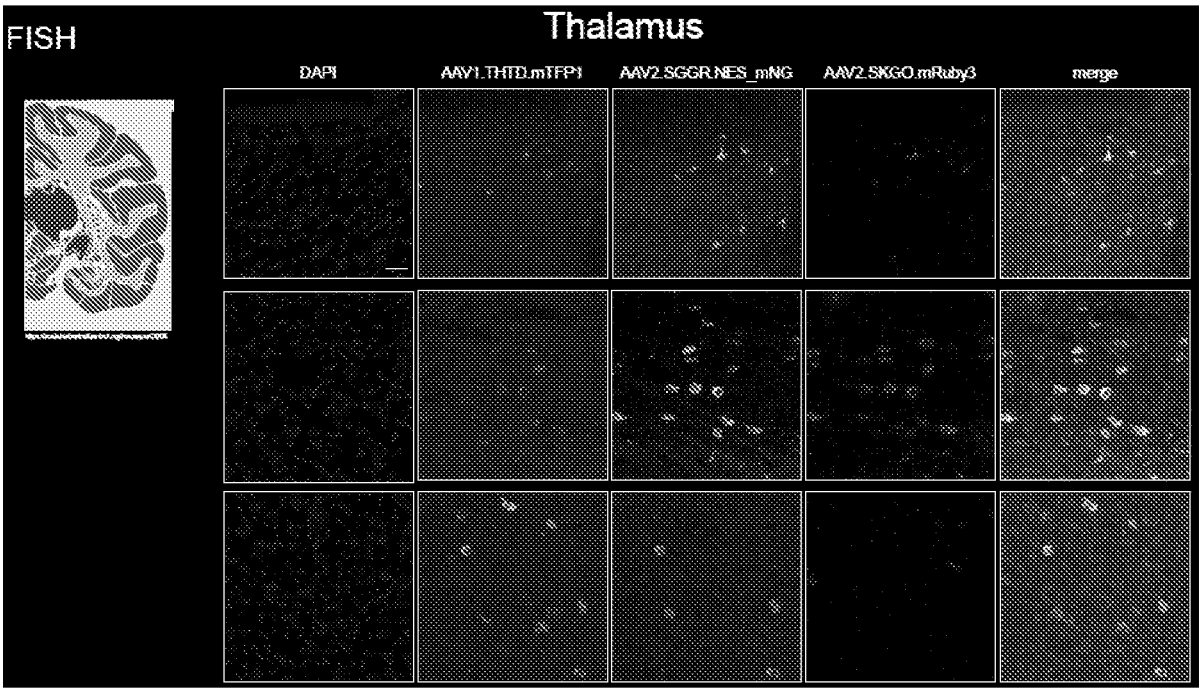

FISH
Thalamus
DAPI
AAV1.THTD.mTFP1
AAV2.SGGR.NES_mNG
AAV2.SKGO.mRuby3
merge

FIG. 20

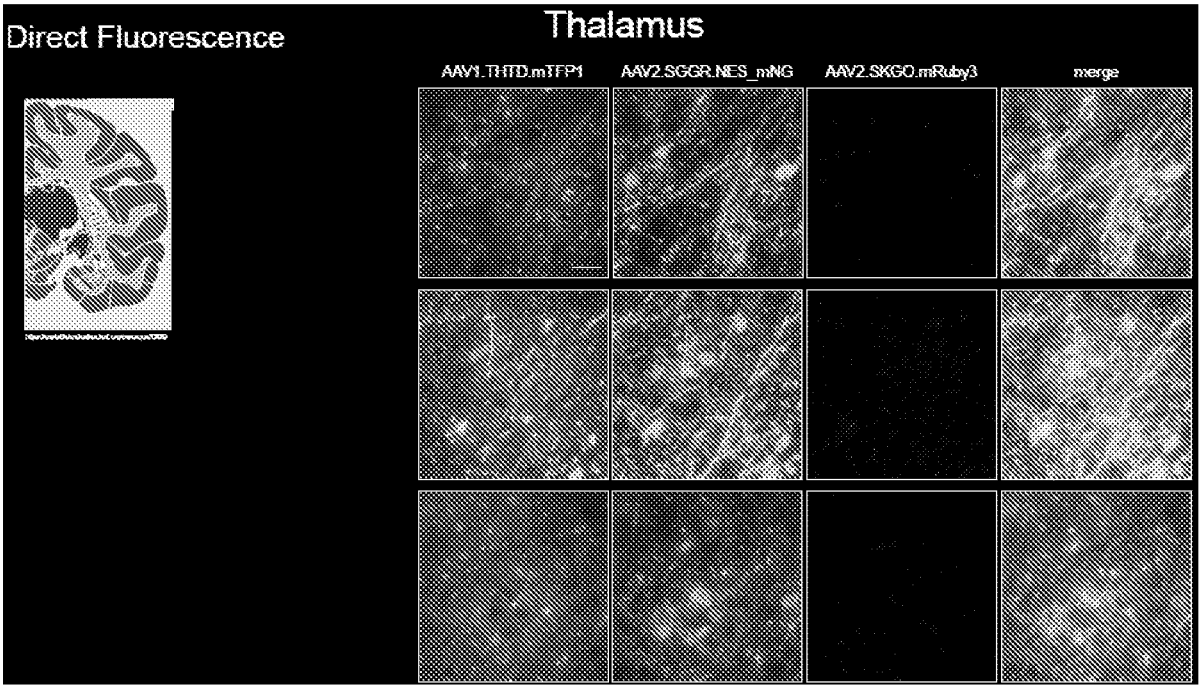
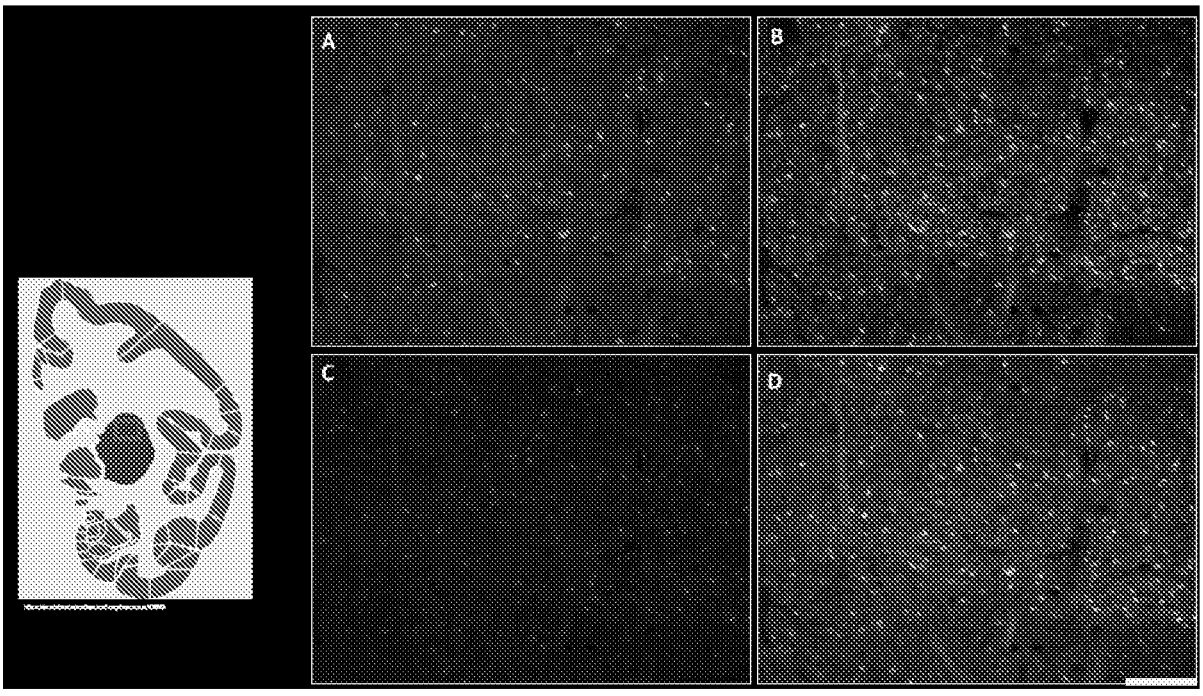
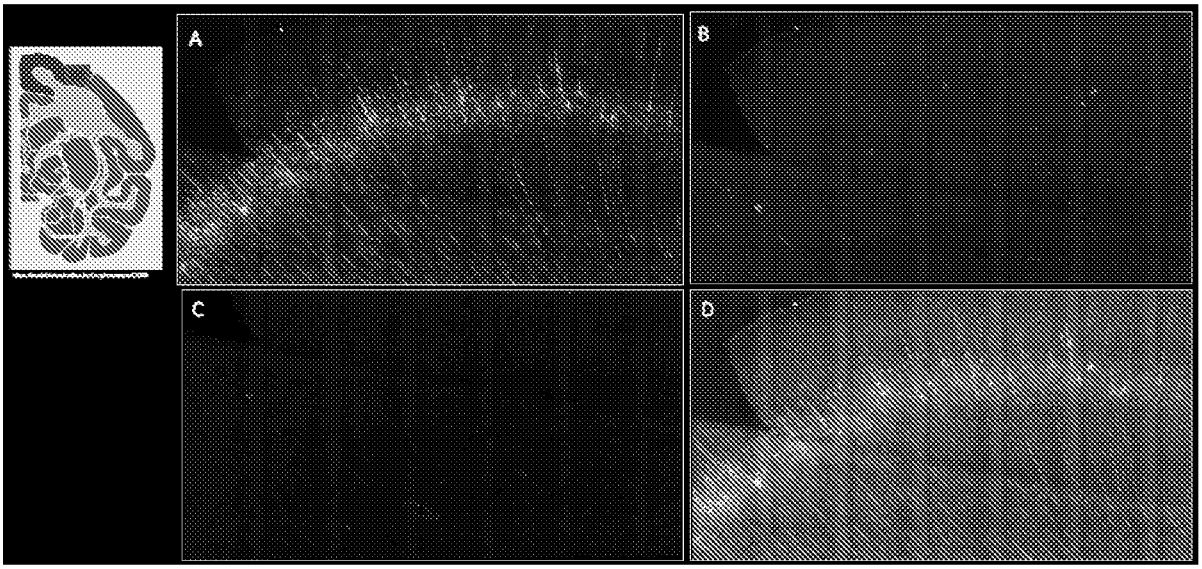


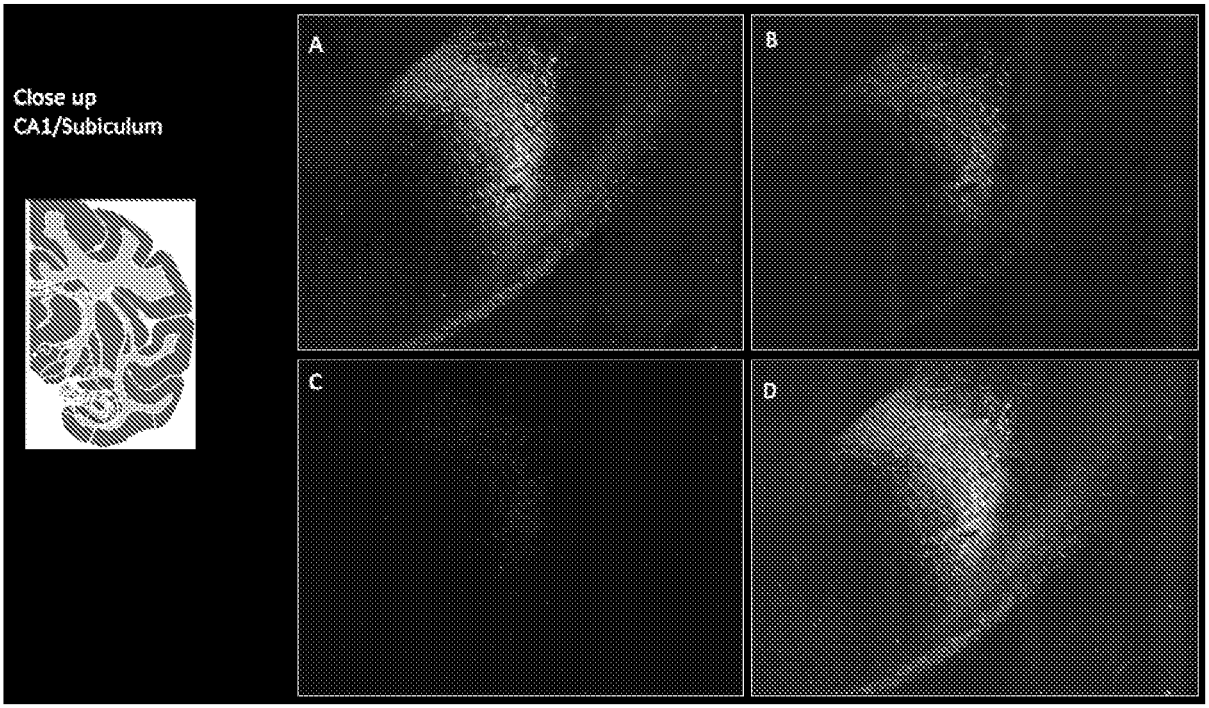
FIG. 21



FIGS. 22A-D



FIGS. 23A-D



FIGS. 24A-D

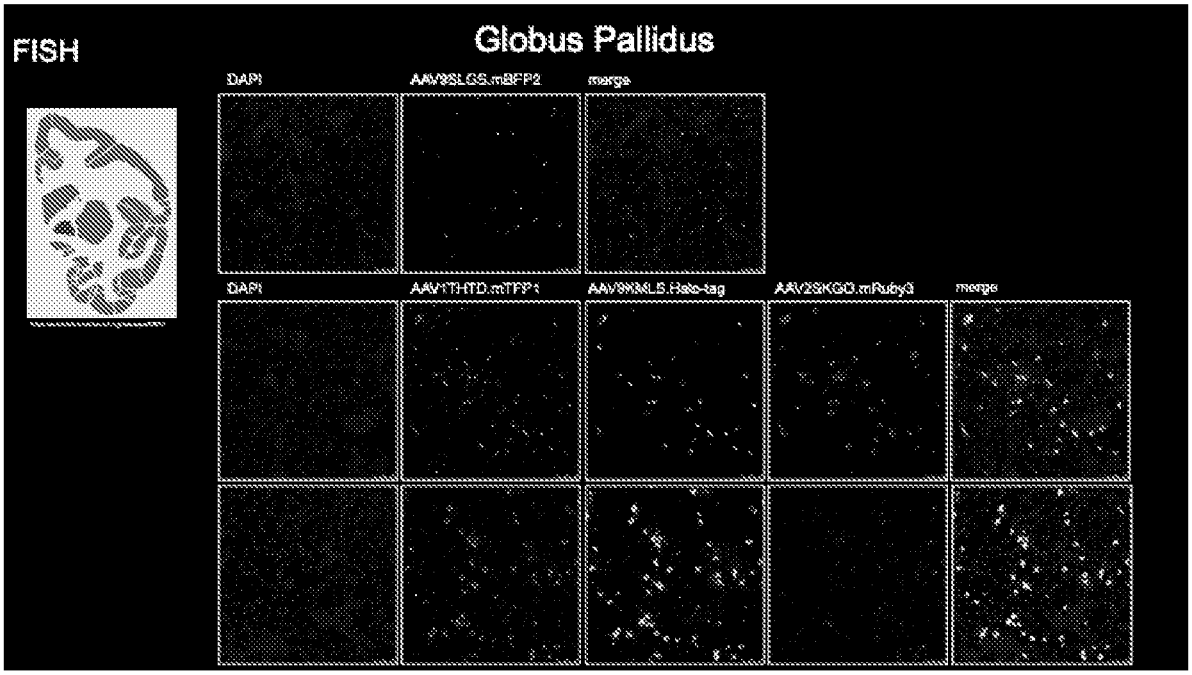

FISH
Globus Pallidus
DAPI
AAV9SLGS.mBFP2
merge
DAPI
AAV1THTD.mTFP1
AAV9KGS.Halo-tag
AAV2SKGD.mRuby3
merge

FIG. 25

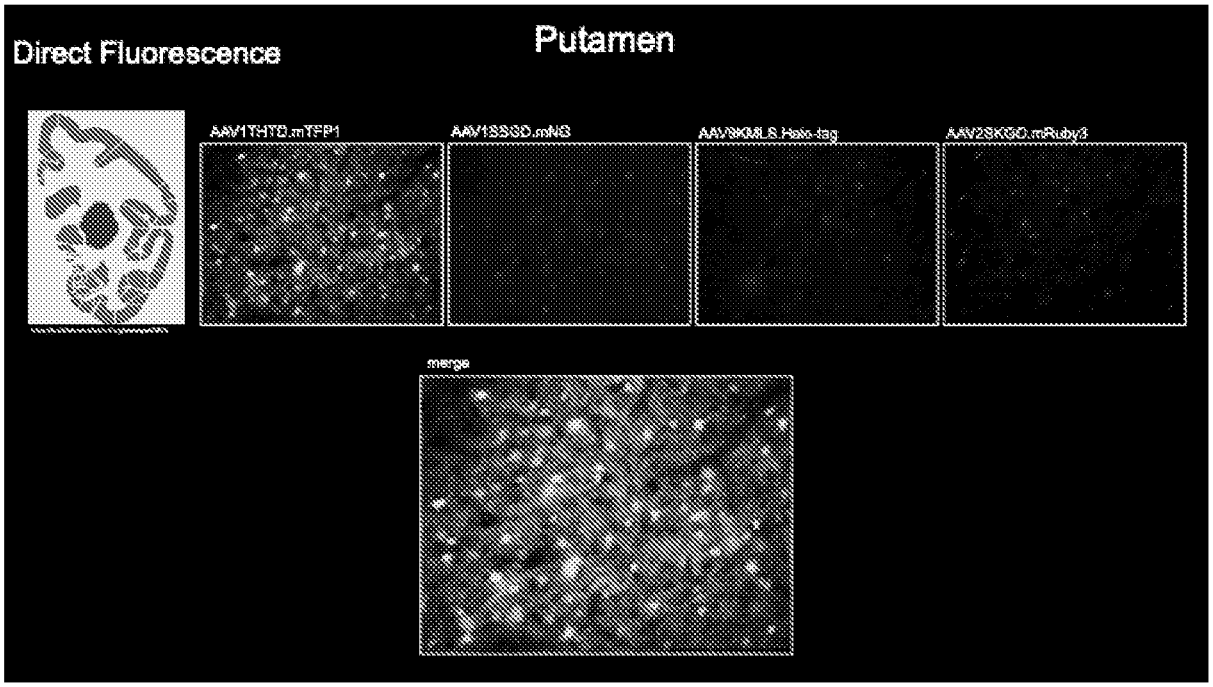

Direct Fluorescence
Putamen
AAV1THTD.mTFP1
AAV1SSGD.mNS
AAV9KGS.Halo-tag
AAV2SKGD.mRuby3
merge

FIG. 26

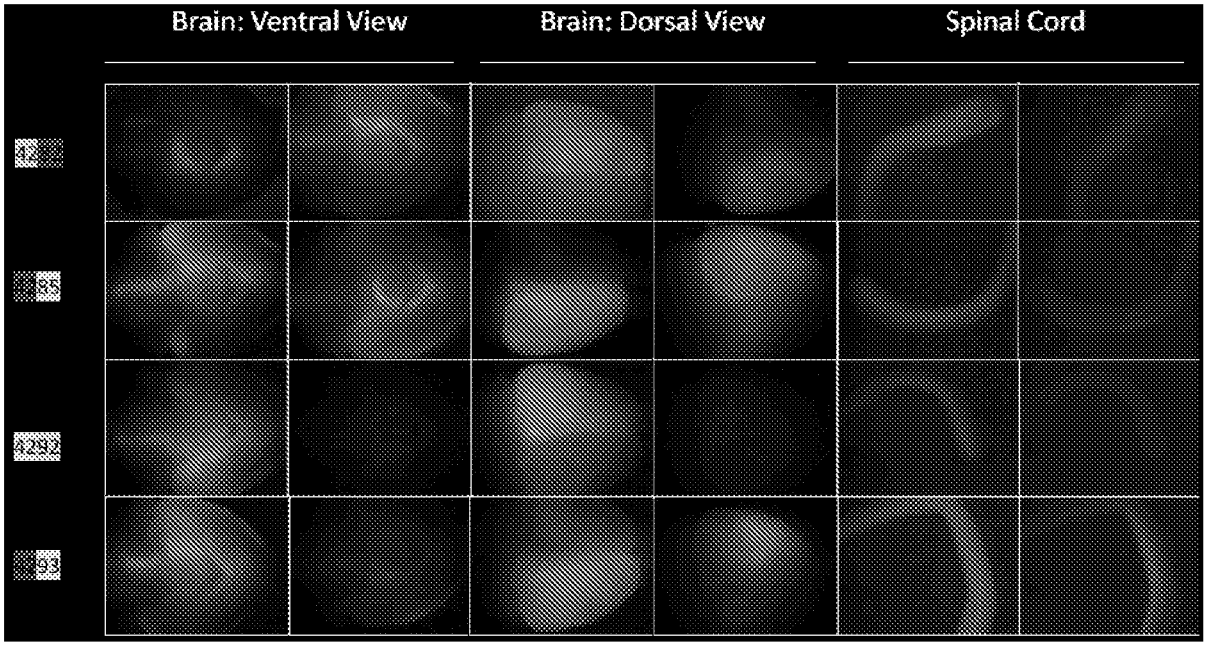

Brain: Ventral View
Brain: Dorsal View
Spinal Cord
42
95
4292
93

FIG. 27

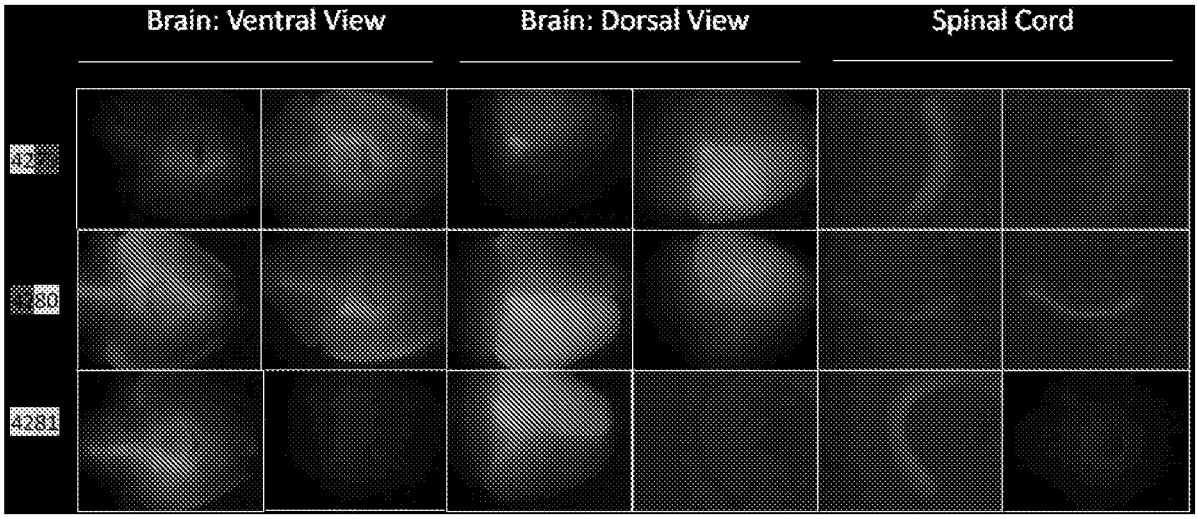

Brain: Ventral View
Brain: Dorsal View
Spinal Cord
42
80
4231

FIG. 28

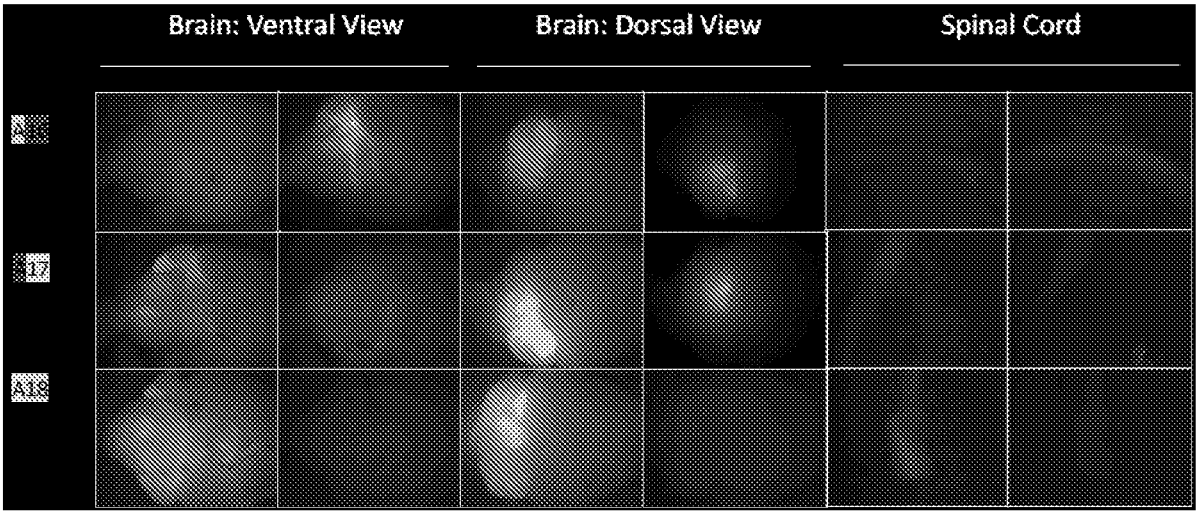


FIG. 29

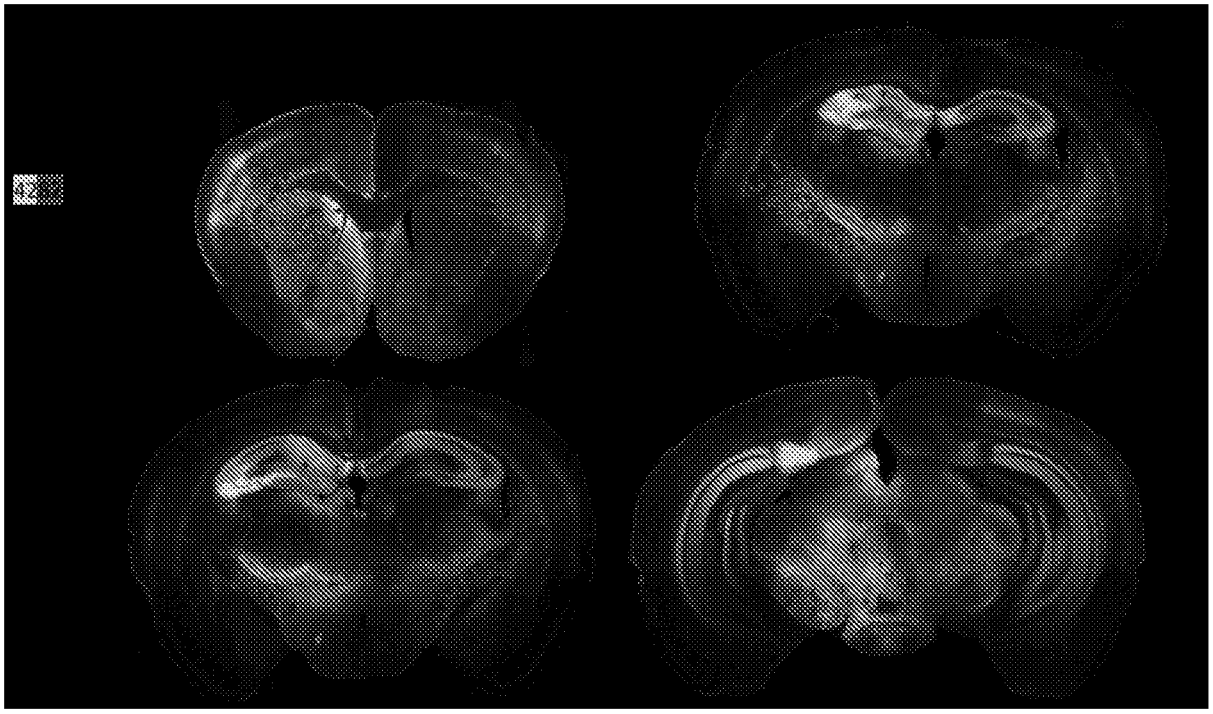


FIG. 30

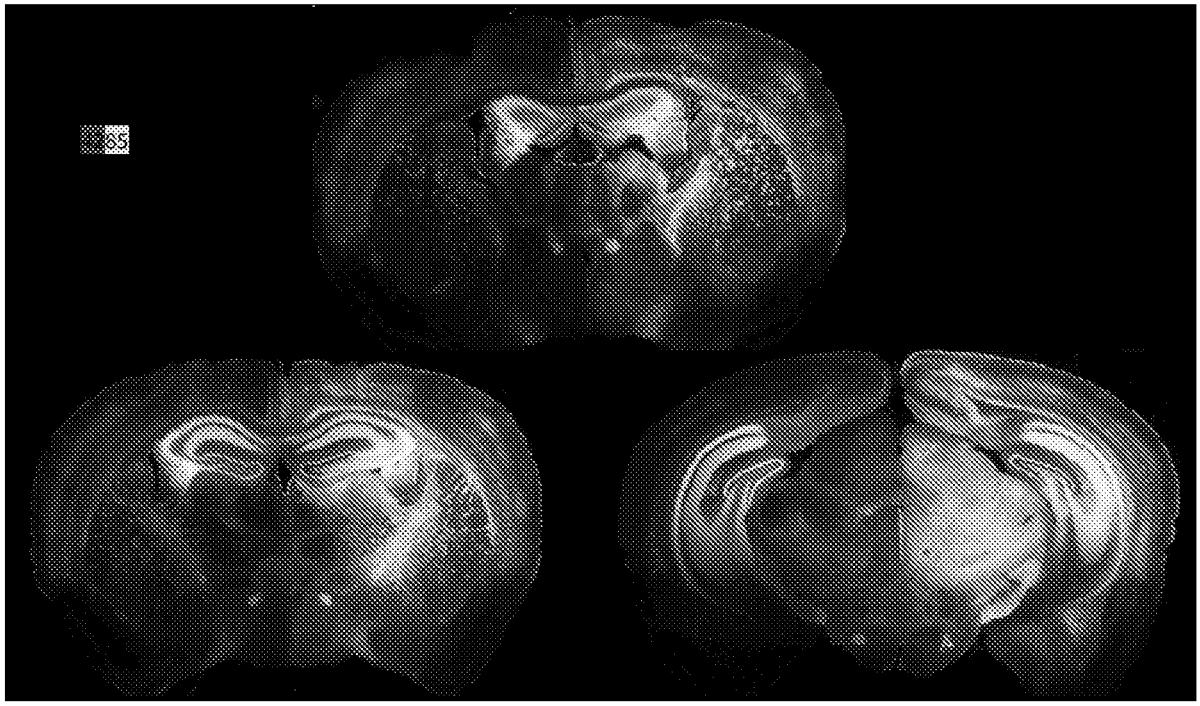


FIG. 31

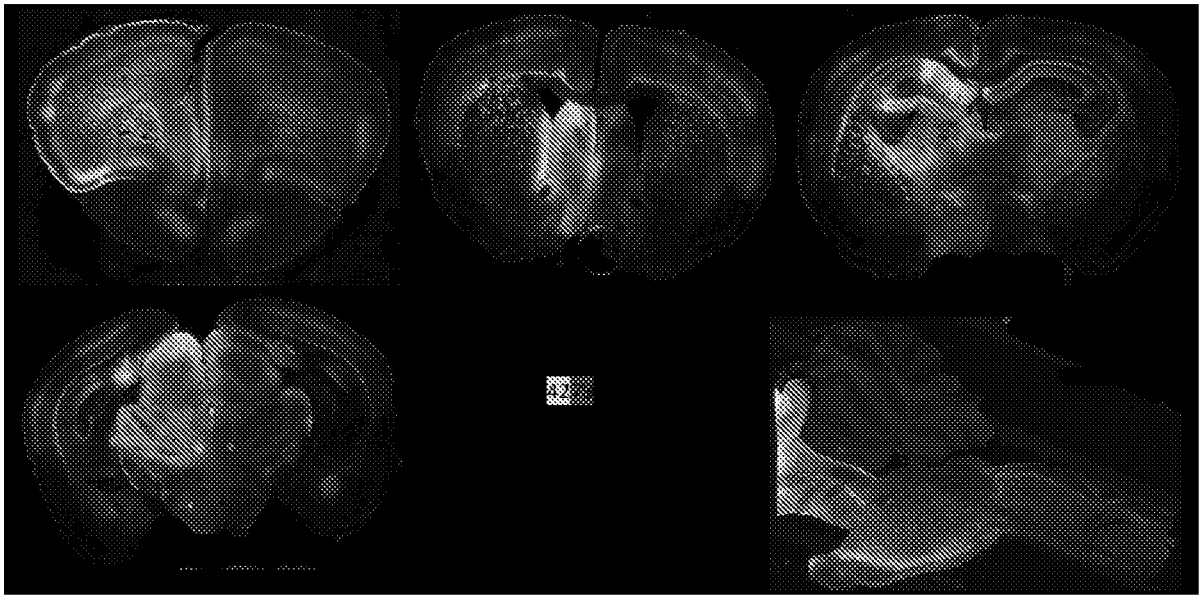


FIG. 32

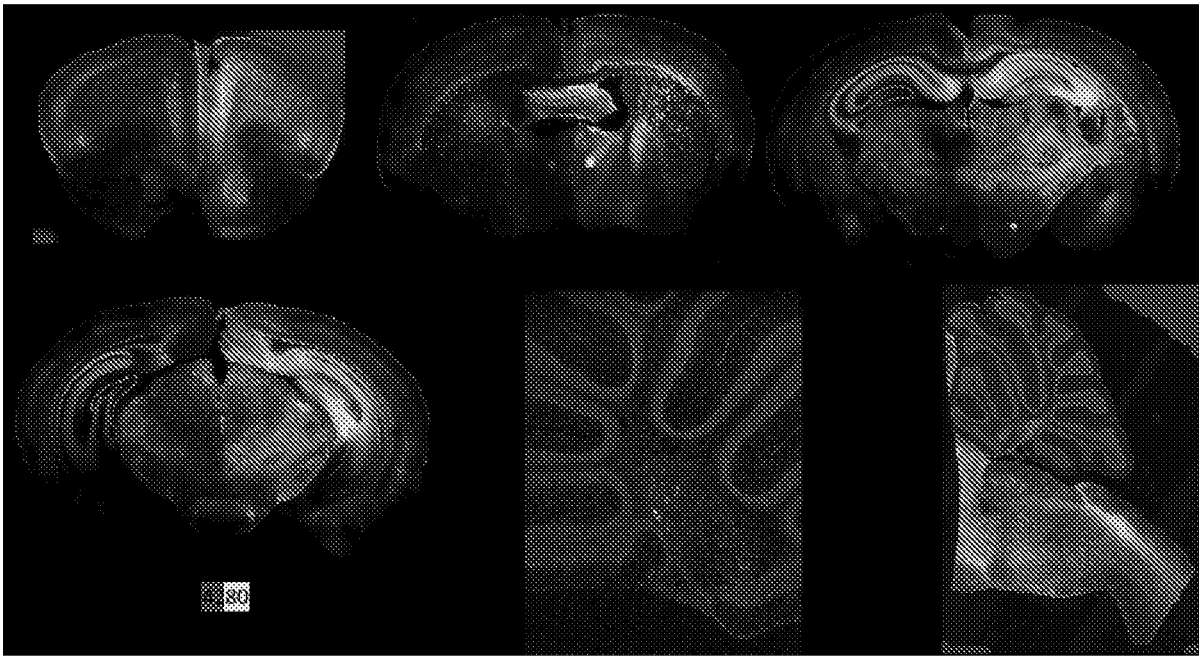


FIG. 33

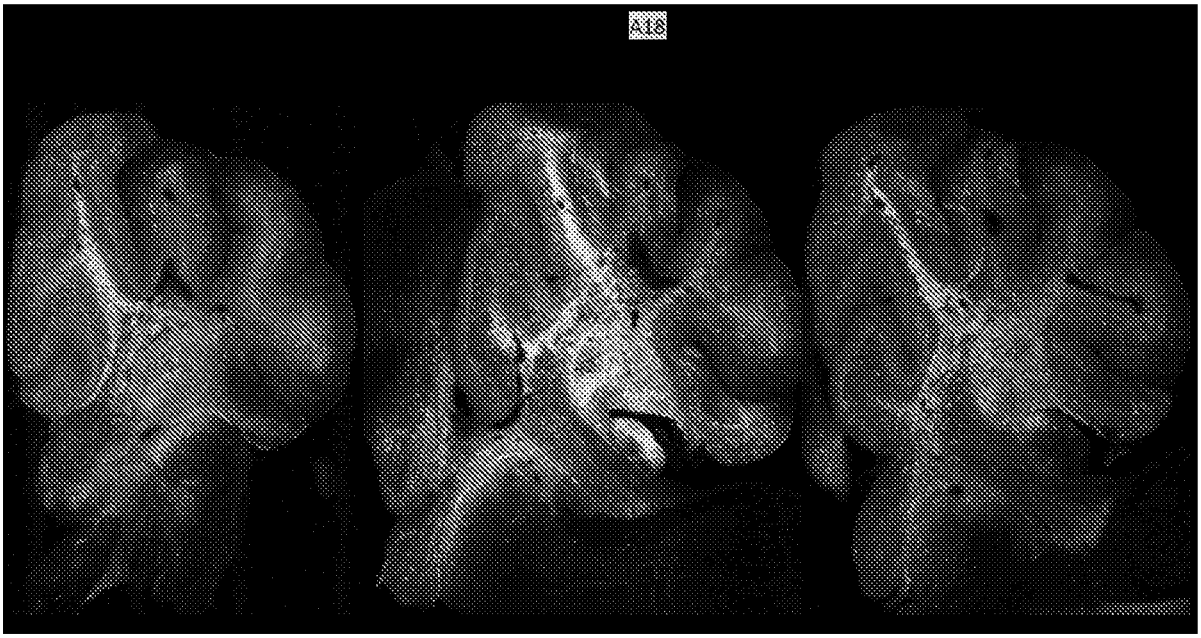


FIG. 34

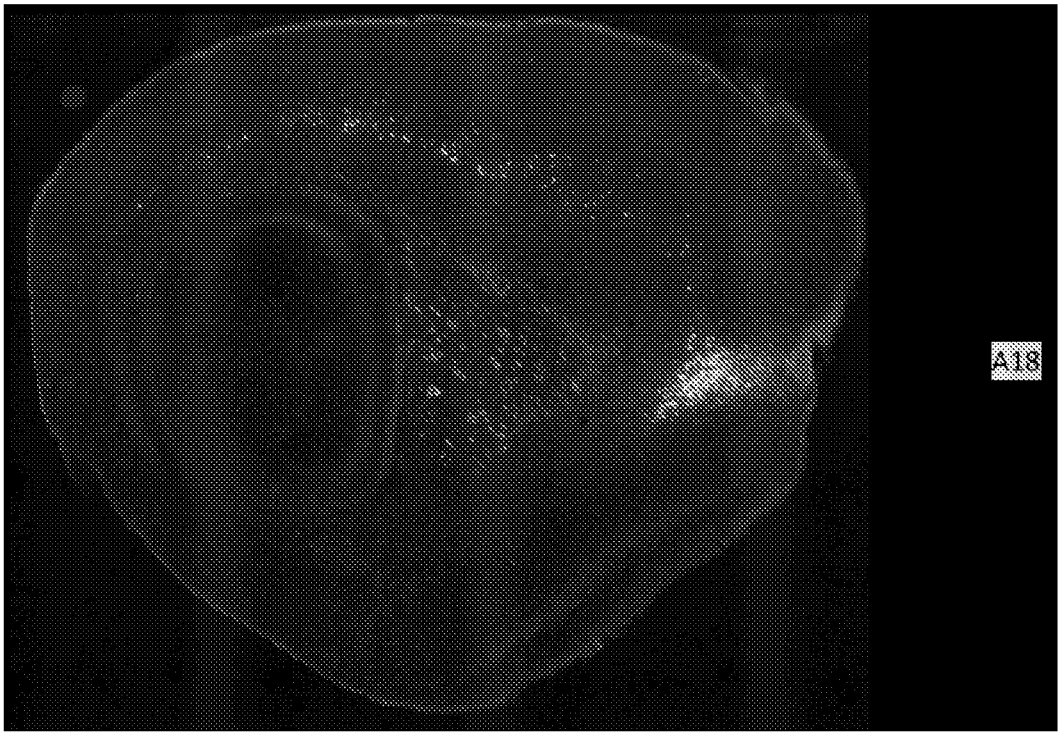


FIG. 35

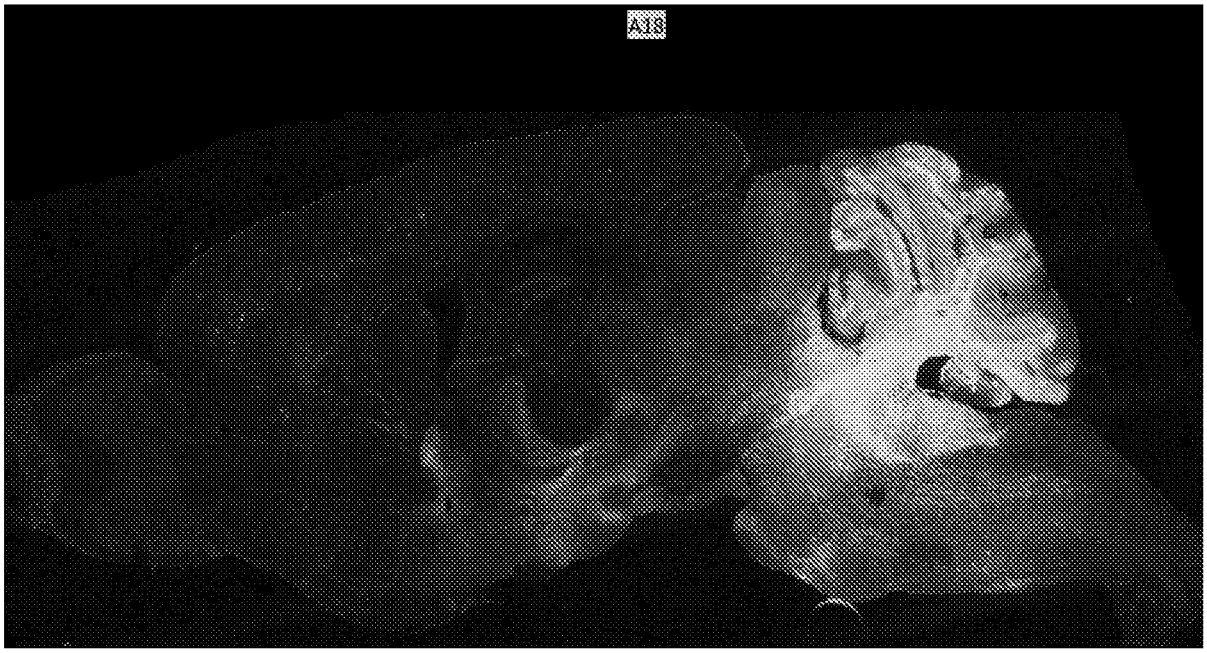


FIG. 36



FIG. 37

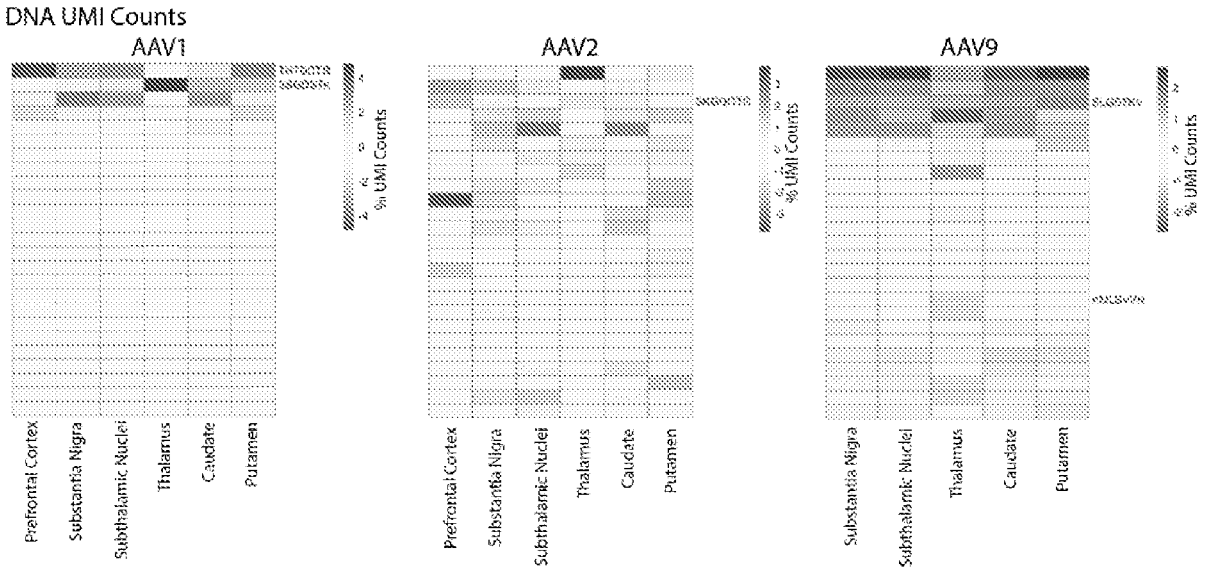


FIG. 38A

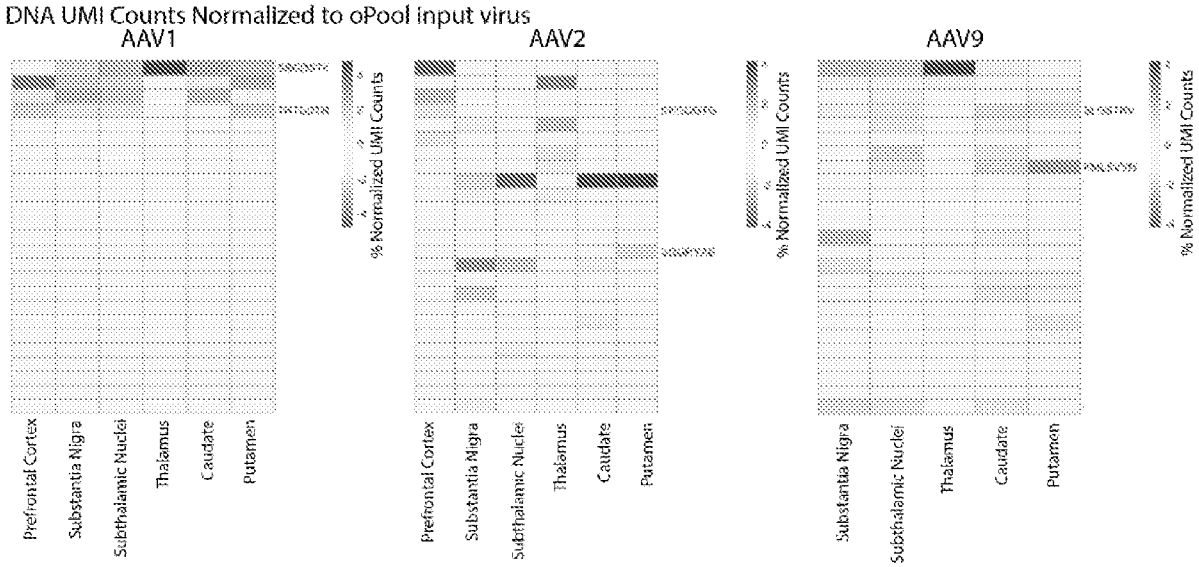


FIG. 38B

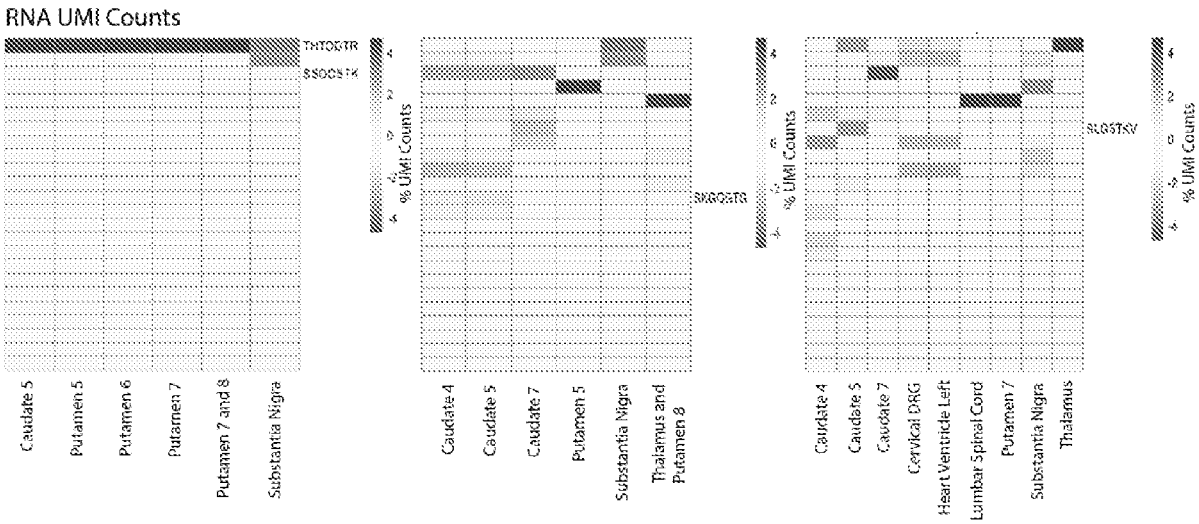


FIG. 38C

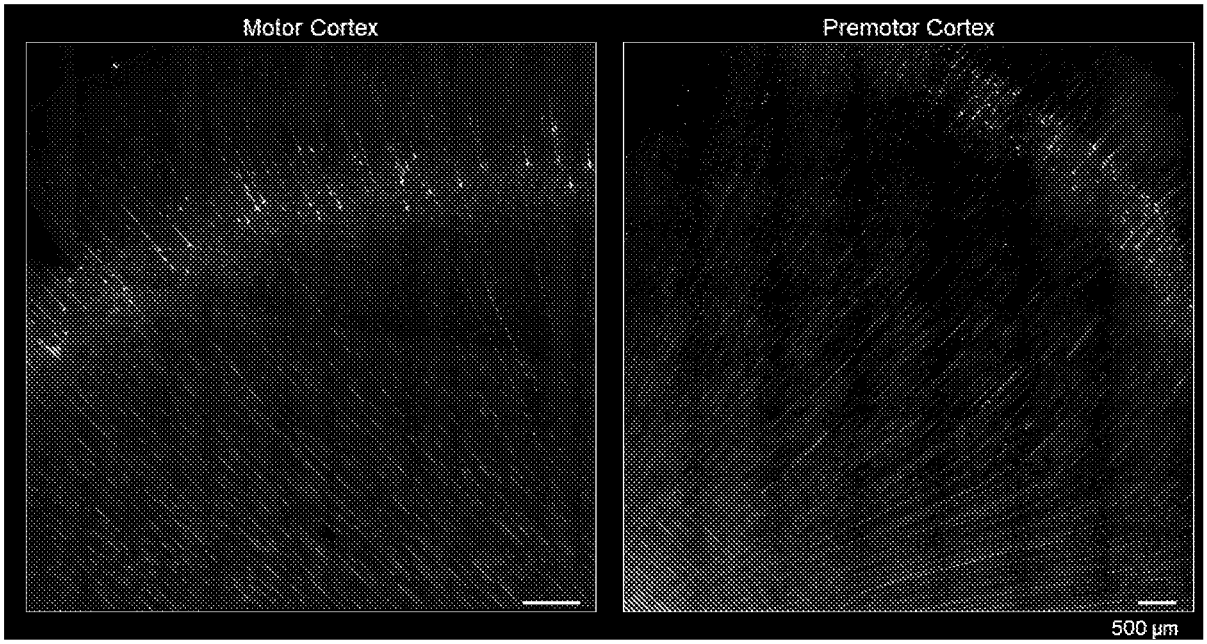

Motor Cortex
Premotor Cortex
500 µm

FIG. 39

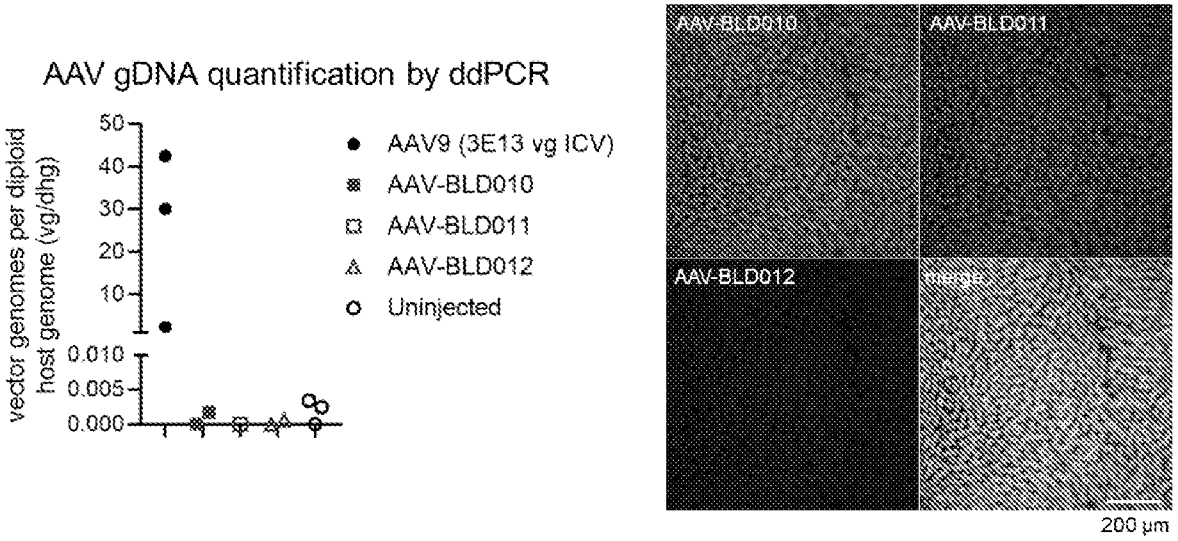

AAV gDNA quantification by ddPCR
vector genomes per diploid host genome (vg/dhg)
50
40
30
20
10
0.010
0.005
0.000
AAV9 (3E13 vg ICV)
AAV-BLD010
AAV-BLD011
AAV-BLD012
Uninjected
AAV-BLD010
AAV-BLD011
AAV-BLD012
200 µm

FIG. 40

Sequence Listing

1	**Sequence Listing Information**	
1-1	File Name	CHOPP0054WO_ST26.xml
1-2	DTD Version	V1_3
1-3	Software Name	WIPO Sequence
1-4	Software Version	2.2.0
1-5	Production Date	2023-05-09
1-6	Original free text language code	en
1-7	Non English free text language code	
2	**General Information**	
2-1	Current application: IP Office	
2-2	Current application: Application number	
2-3	Current application: Filing date	
2-4	Current application: Applicant file reference	CHOP.P0054WO
2-5	Earliest priority application: IP Office	US
2-6	Earliest priority application: Application number	63/340,823
2-7	Earliest priority application: Filing date	2022-05-11
2-8en	Applicant name	The Children's Hospital of Philadelphia
2-8	Applicant name: Name Latin	
2-9en	Inventor name	
2-9	Inventor name: Name Latin	
2-10en	Invention title	ADENO-ASSOCIATED VIRAL VECTORS FOR TARGETING DEEP BRAIN STRUCTURES
2-11	Sequence Total Quantity	24

3-1	**Sequences**	
3-1-1	Sequence Number [ID]	1
3-1-2	Molecule Type	AA
3-1-3	Length	736
3-1-4	Features Location/	**source 1..736**
	Qualifiers	mol_type=protein
		note=adeno-associated virus 1
		organism=Adeno-associated dependoparvovirus A
	NonEnglishQualifier Value	
3-1-5	Residues	MAADGYLPDW LEDNLSEGIR EWWDLKPGAP KPKANQQKQD DGRGLVLPGY KYLGPFNGLD 60 KGEPVNAADA AALEHDKAYD QQLKAGDNPY LRYNHADAEF QERLQEDTSF GGNLGRAVFQ 120 AKKRVLEPLG LVEEGAKTAP GKKRPVEQSP QEPDSSSGIG KTGQQPAKKR LNFGQTGDSE 180 SVPDPQPLGE PPATPAAVGP TTMASGGGAP MADNNEGADG VGNASGNWHC DSTWLGDRVI 240 TTSTRTWALP TYNNHLYKQI SSASTGASND NHYFGYSTPW GYFDFNRFHC HFSPRDWQRL 300 INNNWGFRPK RLNFKLFNIQ VKEVTTNDGV TTIANNLTST VQVFSDSEYQ LPYVLGSAHQ 360 GCLPPFPADV FMIPQYGYLT LNNGSQAVGR SSFYCLEYFP SQMLRTGNNF TFSYTFEEVP 420 FHSSYAHSQS LDRLMNPLID QYLYYLNRTQ NQSGSAQNKD LLFSRGSPAG MSVQPKNWLP 480 GPCYRQQRVS KTKTDNNNSN FTWTGASKYN LNGRESIINP GTAMASHKDD EDKFFPMSGV 540 MIFGKESAGA SNTALDNVMI TDEEEIKATN PVATERFGTV AVNFQSSSTD PATGDVHAMG 600 ALPGMVWQDR DVYLQGPIWA KIPHTDGHFH PSPLMGGFGL KNPPPQILIK NTPVPANPPA 660 EFSATKFASF ITQYSTGQVS VEIEWELQKE NSKRWNPEVQ YTSNYAKSAN VDFTVDNNGL 720 YTEPRPIGTR YLTRPL 736
3-2	**Sequences**	
3-2-1	Sequence Number [ID]	2
3-2-2	Molecule Type	AA
3-2-3	Length	735
3-2-4	Features Location/	**source 1..735**
	Qualifiers	mol_type=protein
		note=adeno-associated virus 2
		organism=Adeno-associated dependoparvovirus A
	NonEnglishQualifier Value	
3-2-5	Residues	MAADGYLPDW LEDTLSEGIR QWWKLKPGPP PPKPAERHKD DSRGLVLPGY KYLGPFNGLD 60 KGEPVNEADA AALEHDKAYD RQLDSGDNPY LKYNHADAEF QERLKEDTSF GGNLGRAVFQ 120 AKKRVLEPLG LVEEPVKTAP GKKRPVEHSP VEPDSSSGTG KAGQQPARKR LNFGQTGDAD 180 SVPDPQPLGQ PPAAPSGLGT NTMATGSGAP MADNNEGADG VGNSSGNWHC DSTWMGDRVI 240 TTSTRTWALP TYNNHLYKQI SSQSGASNDN HYFGYSTPWG YFDFNRFHCH FSPRDWQRLI 300 NNNWGFRPKR LNFKLFNIQV KEVTQNDGTT TIANNLTSTV QVFTDSEYQL PYVLGSAHQG 360 CLPPFPADVF MVPQYGYLTL NNGSQAVGRS SFYCLEYFPS QMLRTGNNFT FSYTFEDVPF 420 HSSYAHSQSL DRLMNPLIDQ YLYYLSRTNT PSGTTTQSRL QFSQAGASDI RDQSRNWLPG 480 PCYRQQRVSK TSADNNNSEY SWTGATKYHL NGRDSLVNPG PAMASHKDDE EKFFPQSGVL 540 IFGKQGSEKT NVDIEKVMIT DEEEIRTTNP VATEQYGSVS TNLQRGNRQA ATADVNTQGV 600 LPGMVWQDRD VYLQGPIWAK IPHTDGHFHP SPLMGGFGLK HPPPQILIKN TPVPANPSTT 660 FSAAKFASFI TQYSTGQVSV EIEWELQKEN SKRWNPEIQY TSNYNKSVNV DFTVDTNGVY 720 SEPRPIGTRY LTRNL 735
3-3	**Sequences**	
3-3-1	Sequence Number [ID]	3
3-3-2	Molecule Type	AA
3-3-3	Length	736
3-3-4	Features Location/	**source 1..736**
	Qualifiers	mol_type=protein
		note=adeno-associated virus 9
		organism=Adeno-associated dependoparvovirus A
	NonEnglishQualifier Value	
3-3-5	Residues	MAADGYLPDW LEDNLSEGIR EWWALKPGAP QPKANQQHQD NARGLVLPGY KYLGPGNGLD 60 KGEPVNAADA AALEHDKAYD QQLKAGDNPY LKYNHADAEF QERLKEDTSF GGNLGRAVFQ 120 AKKRLLEPLG LVEEAAKTAP GKKRPVEQSP QEPDSSAGIG KSGAQPAKKR LNFGQTGDTE 180 SVPDPQPIGE PPAAPSGVGS LTMASGGGAP VADNNEGADG VGSSSGNWHC DSQWLGDRVI 240 TTSTRTWALP TYNNHLYKQI SNSTSGGSSN DNAYFGYSTP WGYFDFNRFH CHFSPRDWQR 300 LINNNWGFRP KRLNFKLFNI QVKEVTDNNG VKTIANNLTS TVQVFTDSDY QLPYVLGSAH 360 EGCLPPFPAD VFMIPQYGYL TLNDGSQAVG RSSFYCLEYF PSQMLRTGNN FQFSYEFENV 420 PFHSSYAHSQ SLDRLMNPLI DQYLYYLSKT INGSGQNQQT LKFSVAGPSN MAVQGRNYIP 480 GPSYRQQRVS TTVTQNNNSE FAWPGASSWA LNGRNSLMNP GPAMASHKEG EDRFFPLSGS 540 LIFGKQGTGR DNVDADKVMI TNEEEIKTTN PVATESYGQV ATNHQSAQAQ AQTGWVQNQG 600 ILPGMVWQDR DVYLQGPIWA KIPHTDGNFH PSPLMGGFGM KHPPPQILIK NTPVPADPPT 660

		AFNKDKLNSF ITQYSTGQVS VEIEWELQKE NSKRWNPEIQ YTSNYYKSNN VEFAVNTEGV 720 YSEPRPIGTR YLTRNL 736
3-4	**Sequences**	
3-4-1	Sequence Number [ID]	4
3-4-2	Molecule Type	AA
3-4-3	Length	748
3-4-4	Features Location/	**source 1..748**
	Qualifiers	mol_type=protein
		organism=synthetic construct
	NonEnglishQualifier Value	
3-4-5	Residues	MAADGYLPDW LEDNLSEGIR EWWDLKPGAP KPKANQQKQD DGRGLVLPGY KYLGPFNGLD 60 KGEPVNAADA AALEHDKAYD QQLKAGDNPY LRYNHADAEF QERLQEDTSF GGNLGRAVFQ 120 AKKRVLEPLG LVEEGAKTAP GKKRPVEQSP QEPDSSSGIG KTGQQPAKKR LNFGQTGDSE 180 SVPDPQPLGE PPATPAAVGP TTMASGGGAP MADNNEGADG VGNASGNWHC DSTWLGDRVI 240 TTSTRTWALP TYNNHLYKQI SSASTGASND NHYFGYSTPW GYFDFNRFHC HFSPRDWQRL 300 INNNWGFRPK RLNFKLFNIQ VKEVTTNDGV TTIANNLTST VQVFSDSEYQ LPYVLGSAHQ 360 GCLPPFPADV FMIPQYGYLT LNNGSQAVGR SSFYCLEYFP SQMLRTGNNF TFSYTFEEVP 420 FHSSYAHSQS LDRLMNPLID QYLYYLNRTQ NQSGSAQNKD LLFSRGSPAG MSVQPKNWLP 480 GPCYRQQRVS KTKTDNNNSN FTWTGASKYN LNGRESIINP GTAMASHKDD EDKFFPMSGV 540 MIFGKESAGA SNTALDNVMI TDEEEIKATN PVATERFGTV AVNFQSSSTD SSAXXXXXXX 600 ASPATGDVHA MGALPGMVWQ DRDVYLQGPI WAKIPHTDGH FHPSPLMGGF GLKNPPPQIL 660 IKNTPVPANP PAEFSATKFA SFITQYSTGQ VSVEIEWELQ KENSKRWNPE VQYTSNYAKS 720 ANVDFTVDNN GLYTEPRPIG TRYLTRPL 748
3-5	**Sequences**	
3-5-1	Sequence Number [ID]	5
3-5-2	Molecule Type	AA
3-5-3	Length	747
3-5-4	Features Location/	**source 1..747**
	Qualifiers	mol_type=protein
		organism=synthetic construct
	NonEnglishQualifier Value	
3-5-5	Residues	MAADGYLPDW LEDTLSEGIR QWWKLKPGPP PPKPAERHKD DSRGLVLPGY KYLGPFNGLD 60 KGEPVNEADA AALEHDKAYD RQLDSGDNPY LKYNHADAEF QERLKEDTSF GGNLGRAVFQ 120 AKKRVLEPLG LVEEPVKTAP GKKRPVEHSP VEPDSSSGTG KAGQQPARKR LNFGQTGDAD 180 SVPDPQPLGQ PPAAPSGLGT NTMATGSGAP MADNNEGADG VGNSSGNWHC DSTWMGDRVI 240 TTSTRTWALP TYNNHLYKQI SSQSGASNDN HYFGYSTPWG YFDFNRFHCH FSPRDWQRLI 300 NNNWGFRPKR LNFKLFNIQV KEVTQNDGTT TIANNLTSTV QVFTDSEYQL PYVLGSAHQG 360 CLPPFPADVF MVPQYGYLTL NNGSQAVGRS SFYCLEYFPS QMLRTGNNFT FSYTFEDVPF 420 HSSYAHSQSL DRLMNPLIDQ YLYYLSRTNT PSGTTTQSRL QFSQAGASDI RDQSRNWLPG 480 PCYRQQRVSK TSADNNNSEY SWTGATKYHL NGRDSLVNPG PAMASHKDDE EKFFPQSGVL 540 IFGKQGSEKT NVDIEKVMIT DEEEIRTTNP VATEQYGSVS TNLQRGNAAA XXXXXXXAAR 600 QAATADVNTQ GVLPGMVWQD RDVYLQGPIW AKIPHTDGHF HPSPLMGGFG LKHPPPQILI 660 KNTPVPANPS TTFSAAKFAS FITQYSTGQV SVEIEWELQK ENSKRWNPEI QYTSNYNKSV 720 NVDFTVDTNG VYSEPRPIGT RYLTRNL 747
3-6	**Sequences**	
3-6-1	Sequence Number [ID]	6
3-6-2	Molecule Type	AA
3-6-3	Length	748
3-6-4	Features Location/	**source 1..748**
	Qualifiers	mol_type=protein
		organism=synthetic construct
	NonEnglishQualifier Value	
3-6-5	Residues	MAADGYLPDW LEDNLSEGIR EWWALKPGAP QPKANQQHQD NARGLVLPGY KYLGPGNGLD 60 KGEPVNAADA AALEHDKAYD QQLKAGDNPY LKYNHADAEF QERLKEDTSF GGNLGRAVFQ 120 AKKRLLEPLG LVEEAAKTAP GKKRPVEQSP QEPDSSAGIG KSGAQPAKKR LNFGQTGDTE 180 SVPDPQPIGE PPAAPSGVGS LTMASGGGAP VADNNEGADG VGSSSGNWHC DSQWLGDRVI 240 TTSTRTWALP TYNNHLYKQI SNSTSGGSSN DNAYFGYSTP WGYFDFNRFH CHFSPRDWQR 300 LINNNWGFRP KRLNFKLFNI QVKEVTDNNG VKTIANNLTS TVQVFTDSDY QLPYVLGSAH 360 EGCLPPFPAD VFMIPQYGYL TLNDGSQAVG RSSFYCLEYF PSQMLRTGNN FQFSYEFENV 420 PFHSSYAHSQ SLDRLMNPLI DQYLYYLSKT INGSGQNQQT LKFSVAGPSN MAVQGRNYIP 480 GPSYRQQRVS TTVTQNNNSE FAWPGASSWA LNGRNSLMNP GPAMASHKEG EDRFFPLSGS 540 LIFGKQGTGR DNVDADKVMI TNEEEIKTTN PVATESYGQV ATNHQSAQAA AXXXXXXXAS 600 AQAQTGWVQN QGILPGMVWQ DRDVYLQGPI WAKIPHTDGN FHPSPLMGGF GMKHPPPQIL 660 IKNTPVPADP PTAFNKDKLN SFITQYSTGQ VSVEIEWELQ KENSKRWNPE IQYTSNYYKS 720 NNVEFAVNTE GVYSEPRPIG TRYLTRNL 748

3-7	**Sequences**		
3-7-1	Sequence Number [ID]	7	
3-7-2	Molecule Type	DNA	
3-7-3	Length	21	
3-7-4	Features Location/	**source 1..21**	
	Qualifiers	mol_type=other DNA	
		organism=synthetic construct	
	NonEnglishQualifier Value		
3-7-5	Residues	tctagtgggg attcgacgaa g	21
3-8	**Sequences**		
3-8-1	Sequence Number [ID]	8	
3-8-2	Molecule Type	AA	
3-8-3	Length	7	
3-8-4	Features Location/	**source 1..7**	
	Qualifiers	mol_type=protein	
		organism=synthetic construct	
	NonEnglishQualifier Value		
3-8-5	Residues	SSGDSTK	7
3-9	**Sequences**		
3-9-1	Sequence Number [ID]	9	
3-9-2	Molecule Type	DNA	
3-9-3	Length	21	
3-9-4	Features Location/	**source 1..21**	
	Qualifiers	mol_type=other DNA	
		organism=synthetic construct	
	NonEnglishQualifier Value		
3-9-5	Residues	acgcatactg atcatacgcg t	21
3-10	**Sequences**		
3-10-1	Sequence Number [ID]	10	
3-10-2	Molecule Type	AA	
3-10-3	Length	7	
3-10-4	Features Location/	**source 1..7**	
	Qualifiers	mol_type=protein	
		organism=synthetic construct	
	NonEnglishQualifier Value		
3-10-5	Residues	THTDDTR	7
3-11	**Sequences**		
3-11-1	Sequence Number [ID]	11	
3-11-2	Molecule Type	DNA	
3-11-3	Length	21	
3-11-4	Features Location/	**source 1..21**	
	Qualifiers	mol_type=other DNA	
		organism=synthetic construct	
	NonEnglishQualifier Value		
3-11-5	Residues	tctaaggggc aggggacggg g	21
3-12	**Sequences**		
3-12-1	Sequence Number [ID]	12	
3-12-2	Molecule Type	AA	
3-12-3	Length	7	
3-12-4	Features Location/	**source 1..7**	
	Qualifiers	mol_type=protein	
		organism=synthetic construct	
	NonEnglishQualifier Value		
3-12-5	Residues	SKGQGTG	7
3-13	**Sequences**		
3-13-1	Sequence Number [ID]	13	
3-13-2	Molecule Type	DNA	
3-13-3	Length	21	
3-13-4	Features Location/	**source 1..21**	
	Qualifiers	mol_type=other DNA	
		organism=synthetic construct	
	NonEnglishQualifier Value		
3-13-5	Residues	tcgggggga ggtatgcgga g	21
3-14	**Sequences**		

3-14-1	Sequence Number [ID]	14	
3-14-2	Molecule Type	AA	
3-14-3	Length	7	
3-14-4	Features Location/	**source 1..7**	
	Qualifiers	mol_type=protein	
		organism=synthetic construct	
	NonEnglishQualifier Value		
3-14-5	Residues	SGGRYAE	7
3-15	**Sequences**		
3-15-1	Sequence Number [ID]	15	
3-15-2	Molecule Type	DNA	
3-15-3	Length	21	
3-15-4	Features Location/	**source 1..21**	
	Qualifiers	mol_type=other DNA	
		organism=synthetic construct	
	NonEnglishQualifier Value		
3-15-5	Residues	tctttgggta gtactaaggt g	21
3-16	**Sequences**		
3-16-1	Sequence Number [ID]	16	
3-16-2	Molecule Type	AA	
3-16-3	Length	7	
3-16-4	Features Location/	**source 1..7**	
	Qualifiers	mol_type=protein	
		organism=synthetic construct	
	NonEnglishQualifier Value		
3-16-5	Residues	SLGSTKV	7
3-17	**Sequences**		
3-17-1	Sequence Number [ID]	17	
3-17-2	Molecule Type	DNA	
3-17-3	Length	21	
3-17-4	Features Location/	**source 1..21**	
	Qualifiers	mol_type=other DNA	
		organism=synthetic construct	
	NonEnglishQualifier Value		
3-17-5	Residues	aagatgttga gtgttgttaa t	21
3-18	**Sequences**		
3-18-1	Sequence Number [ID]	18	
3-18-2	Molecule Type	AA	
3-18-3	Length	7	
3-18-4	Features Location/	**source 1..7**	
	Qualifiers	mol_type=protein	
		organism=synthetic construct	
	NonEnglishQualifier Value		
3-18-5	Residues	KMLSVVN	7
3-19	**Sequences**		
3-19-1	Sequence Number [ID]	19	
3-19-2	Molecule Type	AA	
3-19-3	Length	748	
3-19-4	Features Location/	**source 1..748**	
	Qualifiers	mol_type=protein	
		note=AAV1.SSGD	
		organism=synthetic construct	
	NonEnglishQualifier Value		
3-19-5	Residues	MAADGYLPDW LEDNLSEGIR EWWDLKPGAP KPKANQQKQD DGRGLVLPGY KYLGPFNGLD 60	
		KGEPVNAADA AALEHDKAYD QQLKAGDNPY LRYNHADAEF QERLQEDTSF GGNLGRAVFQ 120	
		AKKRVLEPLG LVEEGAKTAP GKKRPVEQSP QEPDSSSGIG KTGQQPAKKR LNFGQTGDSE 180	
		SVPDPQPLGE PPATPAAVGP TTMASGGGAP MADNNEGADG VGNASGNWHC DSTWLGDRVI 240	
		TTSTRTWALP TYNNHLYKQI SSASTGASND NHYFGYSTPW GYFDFNRFHC HFSPRDWQRL 300	
		INNNWGFRPK RLNFKLFNIQ VKEVTTNDGV TTIANNLTST VQVFSDSEYQ LPYVLGSAHQ 360	
		GCLPPFPADV FMIPQYGYLT LNNGSQAVGR SSFYCLEYFP SQMLRTGNNF TFSYTFEEVP 420	
		FHSSYAHSQS LDRLMNPLID QYLYYLNRTQ NQSGSAQNKD LLFSRGSPAG MSVQPKNWLP 480	
		GPCYRQQRVS KTKTDNNNSN FTWTGASKYN LNGRESIINP GTAMASHKDD EDKFFPMSGV 540	
		MIFGKESAGA SNTALDNVMI TDEEEIKATN PVATERFGTV AVNFQSSSTD SSASSGDSTK 600	
		ASPATGDVHA MGALPGMVWQ DRDVYLQGPI WAKIPHTDGH FHPSPLMGGF GLKNPPPQIL 660	

		IKNTPVPANP PAEFSATKFA SFITQYSTGQ VSVEIEWELQ KENSKRWNPE VQYTSNYAKS 720 ANVDFTVDNN GLYTEPRPIG TRYLTRPL 748
3-20	**Sequences**	
3-20-1	Sequence Number [ID]	20
3-20-2	Molecule Type	AA
3-20-3	Length	748
3-20-4	Features Location/	**source 1..748**
	Qualifiers	mol_type=protein
		note=AAV1.THTD
		organism=synthetic construct
	NonEnglishQualifier Value	
3-20-5	Residues	MAADGYLPDW LEDNLSEGIR EWWDLKPGAP KPKANQQKQD DGRGLVLPGY KYLGPFNGLD 60 KGEPVNAADA AALEHDKAYD QQLKAGDNPY LRYNHADAEF QERLQEDTSF GGNLGRAVFQ 120 AKKRVLEPLG LVEEGAKTAP GKKRPVEQSP QEPDSSSGIG KTGQQPAKKR LNFGQTGDSE 180 SVPDPQPLGE PPATPAAVGP TTMASGGGAP MADNNEGADG VGNASGNWHC DSTWLGDRVI 240 TTSTRTWALP TYNNHLYKQI SSASTGASND NHYFGYSTPW GYFDFNRFHC HFSPRDWQRL 300 INNNWGFRPK RLNFKLFNIQ VKEVTTNDGV TTIANNLTST VQVFSDSEYQ LPYVLGSAHQ 360 GCLPPFPADV FMIPQYGYLT LNNGSQAVGR SSFYCLEYFP SQMLRTGNNF TFSYTFEEVP 420 FHSSYAHSQS LDRLMNPLID QYLYYLNRTQ NQSGSAQNKD LLFSRGSPAG MSVQPKNWLP 480 GPCYRQQRVS KTKTDNNNSN FTWTGASKYN LNGRESIINP GTAMASHKDD EDKFFPMSGV 540 MIFGKESAGA SNTALDNVMI TDEEEIKATN PVATERFGTV AVNFQSSSTD SSATHTDDTR 600 ASPATGDVHA MGALPGMVWQ DRDVYLQGPI WAKIPHTDGH FHPSPLMGGF GLKNPPPQIL 660 IKNTPVPANP PAEFSATKFA SFITQYSTGQ VSVEIEWELQ KENSKRWNPE VQYTSNYAKS 720 ANVDFTVDNN GLYTEPRPIG TRYLTRPL 748
3-21	**Sequences**	
3-21-1	Sequence Number [ID]	21
3-21-2	Molecule Type	AA
3-21-3	Length	747
3-21-4	Features Location/	**source 1..747**
	Qualifiers	mol_type=protein
		note=AAV2.SKGQ
		organism=synthetic construct
	NonEnglishQualifier Value	
3-21-5	Residues	MAADGYLPDW LEDTLSEGIR QWWKLKPGPP PPKPAERHKD DSRGLVLPGY KYLGPFNGLD 60 KGEPVNEADA AALEHDKAYD RQLDSGDNPY LKYNHADAEF QERLKEDTSF GGNLGRAVFQ 120 AKKRVLEPLG LVEEPVKTAP GKKRPVEHSP VEPDSSSGTG KAGQQPARKR LNFGQTGDAD 180 SVPDPQPLGQ PPAAPSGLGT NTMATGSGAP MADNNEGADG VGNSSGNWHC DSTWMGDRVI 240 TTSTRTWALP TYNNHLYKQI SSQSGASNDN HYFGYSTPWG YFDFNRFHCH FSPRDWQRLI 300 NNNWGFRPKR LNFKLFNIQV KEVTQNDGTT TIANNLTSTV QVFTDSEYQL PYVLGSAHQG 360 CLPPFPADVF MVPQYGYLTL NNGSQAVGRS SFYCLEYFPS QMLRTGNNFT FSYTFEDVPF 420 HSSYAHSQSL DRLMNPLIDQ YLYYLSRTNT PSGTTTQSRL QFSQAGASDI RDQSRNWLPG 480 PCYRQQRVSK TSADNNNSEY SWTGATKYHL NGRDSLVNPG PAMASHKDDE EKFFPQSGVL 540 IFGKQGSEKT NVDIEKVMIT DEEEIRTTNP VATEQYGSVS TNLQRGNAAA SKGQGTGAAR 600 QAATADVNTQ GVLPGMVWQD RDVYLQGPIW AKIPHTDGHF HPSPLMGGFG LKHPPPQILI 660 KNTPVPANPS TTFSAAKFAS FITQYSTGQV SVEIEWELQK ENSKRWNPEI QYTSNYNKSV 720 NVDFTVDTNG VYSEPRPIGT RYLTRNL 747
3-22	**Sequences**	
3-22-1	Sequence Number [ID]	22
3-22-2	Molecule Type	AA
3-22-3	Length	747
3-22-4	Features Location/	**source 1..747**
	Qualifiers	mol_type=protein
		note=AAV2.SGGR
		organism=synthetic construct
	NonEnglishQualifier Value	
3-22-5	Residues	MAADGYLPDW LEDTLSEGIR QWWKLKPGPP PPKPAERHKD DSRGLVLPGY KYLGPFNGLD 60 KGEPVNEADA AALEHDKAYD RQLDSGDNPY LKYNHADAEF QERLKEDTSF GGNLGRAVFQ 120 AKKRVLEPLG LVEEPVKTAP GKKRPVEHSP VEPDSSSGTG KAGQQPARKR LNFGQTGDAD 180 SVPDPQPLGQ PPAAPSGLGT NTMATGSGAP MADNNEGADG VGNSSGNWHC DSTWMGDRVI 240 TTSTRTWALP TYNNHLYKQI SSQSGASNDN HYFGYSTPWG YFDFNRFHCH FSPRDWQRLI 300 NNNWGFRPKR LNFKLFNIQV KEVTQNDGTT TIANNLTSTV QVFTDSEYQL PYVLGSAHQG 360 CLPPFPADVF MVPQYGYLTL NNGSQAVGRS SFYCLEYFPS QMLRTGNNFT FSYTFEDVPF 420 HSSYAHSQSL DRLMNPLIDQ YLYYLSRTNT PSGTTTQSRL QFSQAGASDI RDQSRNWLPG 480 PCYRQQRVSK TSADNNNSEY SWTGATKYHL NGRDSLVNPG PAMASHKDDE EKFFPQSGVL 540 IFGKQGSEKT NVDIEKVMIT DEEEIRTTNP VATEQYGSVS TNLQRGNAAA SGGRYAEAAR 600

		QAATADVNTQ GVLPGMVWQD RDVYLQGPIW AKIPHTDGHF HPSPLMGGFG LKHPPPQILI 660 KNTPVPANPS TTFSAAKFAS FITQYSTGQV SVEIEWELQK ENSKRWNPEI QYTSNYNKSV 720 NVDFTVDTNG VYSEPRPIGT RYLTRNL 747
3-23	**Sequences**	
3-23-1	Sequence Number [ID]	23
3-23-2	Molecule Type	AA
3-23-3	Length	748
3-23-4	Features Location/	**source 1..748**
	Qualifiers	mol_type=protein note=AAV9.SLGS organism=synthetic construct
	NonEnglishQualifier Value	
3-23-5	Residues	MAADGYLPDW LEDNLSEGIR EWWALKPGAP QPKANQQHQD NARGLVLPGY KYLGPGNGLD 60 KGEPVNAADA AALEHDKAYD QQLKAGDNPY LKYNHADAEF QERLKEDTSF GGNLGRAVFQ 120 AKKRLLEPLG LVEEAAKTAP GKKRPVEQSP QEPDSSAGIG KSGAQPAKKR LNFGQTGDTE 180 SVPDPQPIGE PPAAPSGVGS LTMASGGGAP VADNNEGADG VGSSSGNWHC DSQWLGDRVI 240 TTSTRTWALP TYNNHLYKQI SNSTSGGSSN DNAYFGYSTP WGYFDFNRFH CHFSPRDWQR 300 LINNNWGFRP KRLNFKLFNI QVKEVTDNNG VKTIANNLTS TVQVFTDSDY QLPYVLGSAH 360 EGCLPPFPAD VFMIPQYGYL TLNDGSQAVG RSSFYCLEYF PSQMLRTGNN FQFSYEFENV 420 PFHSSYAHSQ SLDRLMNPLI DQYLYYLSKT INGSGQNQQT LKFSVAGPSN MAVQGRNYIP 480 GPSYRQQRVS TTVTQNNNSE FAWPGASSWA LNGRNSLMNP GPAMASHKEG EDRFFPLSGS 540 LIFGKQGTGR DNVDADKVMI TNEEEIKTTN PVATESYGQV ATNHQSAQAA ASLGSTKVAS 600 AQAQTGWVQN QGILPGMVWQ DRDVYLQGPI WAKIPHTDGN FHPSPLMGGF GMKHPPPQIL 660 IKNTPVPADP PTAFNKDKLN SFITQYSTGQ VSVEIEWELQ KENSKRWNPE IQYTSNYYKS 720 NNVEFAVNTE GVYSEPRPIG TRYLTRNL 748
3-24	**Sequences**	
3-24-1	Sequence Number [ID]	24
3-24-2	Molecule Type	AA
3-24-3	Length	748
3-24-4	Features Location/	**source 1..748**
	Qualifiers	mol_type=protein note=AAV9.KMLS organism=synthetic construct
	NonEnglishQualifier Value	
3-24-5	Residues	MAADGYLPDW LEDNLSEGIR EWWALKPGAP QPKANQQHQD NARGLVLPGY KYLGPGNGLD 60 KGEPVNAADA AALEHDKAYD QQLKAGDNPY LKYNHADAEF QERLKEDTSF GGNLGRAVFQ 120 AKKRLLEPLG LVEEAAKTAP GKKRPVEQSP QEPDSSAGIG KSGAQPAKKR LNFGQTGDTE 180 SVPDPQPIGE PPAAPSGVGS LTMASGGGAP VADNNEGADG VGSSSGNWHC DSQWLGDRVI 240 TTSTRTWALP TYNNHLYKQI SNSTSGGSSN DNAYFGYSTP WGYFDFNRFH CHFSPRDWQR 300 LINNNWGFRP KRLNFKLFNI QVKEVTDNNG VKTIANNLTS TVQVFTDSDY QLPYVLGSAH 360 EGCLPPFPAD VFMIPQYGYL TLNDGSQAVG RSSFYCLEYF PSQMLRTGNN FQFSYEFENV 420 PFHSSYAHSQ SLDRLMNPLI DQYLYYLSKT INGSGQNQQT LKFSVAGPSN MAVQGRNYIP 480 GPSYRQQRVS TTVTQNNNSE FAWPGASSWA LNGRNSLMNP GPAMASHKEG EDRFFPLSGS 540 LIFGKQGTGR DNVDADKVMI TNEEEIKTTN PVATESYGQV ATNHQSAQAA AKMLSVVNAS 600 AQAQTGWVQN QGILPGMVWQ DRDVYLQGPI WAKIPHTDGN FHPSPLMGGF GMKHPPPQIL 660 IKNTPVPADP PTAFNKDKLN SFITQYSTGQ VSVEIEWELQ KENSKRWNPE IQYTSNYYKS 720 NNVEFAVNTE GVYSEPRPIG TRYLTRNL 748