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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
07 October 2021 (07.10.2021)

WIPO | PCT

(10) International Publication Number
WO 2021/202494 A1

(51) International Patent Classification:
C07K 14/005 (2006.01) *C12N 15/86* (2006.01)

(21) International Application Number:
PCT/US2021/024837

(22) International Filing Date:
30 March 2021 (30.03.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/003,143 31 March 2020 (31.03.2020) US

(71) Applicants: **UNIVERSITY OF MASSACHUSETTS** [US/US]; One Beacon Street, 31st Floor, Boston, MA 02108 (US). **SICHUAN UNIVERISTY** [CN/CN]; Shuncheng St., 252, Qingyang District, Sichuan, Chengdu, 610017 (CN).

(72) Inventors: **GAO, Guangping**; 4 Edward Dunn Way, Westborough, MA 01581 (US). **XU, Guangchao**; 7 Pocono Road, Worcester, MA 01605 (US). **TAI, Phillip**; 51 Alvarado Ave., Apt 203, Worcester, MA 01604 (US). **WEI, Yuquan**; Gaopent Street, Keyuan Street Road 4, No. 1, Chengdu, Sichuan, 610041 (CN). **LUO, Li**; 7 Pocono Road, Worcester, MA 01605 (US).

(74) Agent: **MACDONALD, Kevin** et al.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210-2206 (US).

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

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

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

(54) Title: AAV CAPSIDS VARIANTS AND USES THEREOF

(57) Abstract: Aspects of the disclosure relate to compositions and methods for delivering a transgene (e.g., a transgene encoding one or more gene products) to a target cell. The disclosure is based, in part, on adeno-associated virus (AAV) capsid protein variants characterized by tropisms for certain cell types (e.g., neurons, muscle cells, bone cells, heart cells, etc.). In some embodiments, recombinant A A Vs (rAAVs) comprising the capsid protein variants (e.g., AAVv66, SEQ ID NO: 1) are more efficiently packaged than rAAVs having certain wild-type AAV capsid proteins. Methods of delivering an rAAV comprising the AAV capsid protein variants are also described by the disclosure.

WO 2021/202494 A1

AAV CAPSIDS VARIANTS AND USES THEREOF

BACKGROUND

Recombinant AAV adeno-associated viruses (rAAVs) are capable of driving stable and sustained transgene expression in target tissues without notable toxicity and host immunogenicity. Thus, rAAVs are promising delivery vehicles for long-term therapeutic gene expression. However, low transduction efficiency and restricted tissue tropisms by currently available rAAV vectors can limit their application as feasible and efficacious therapies. Additionally, faithful clinical translation of leading therapeutic AAV serotypes derived from non-human tissues is a concern. Accordingly, a need remains for new AAV vectors for gene delivery.

SUMMARY

Aspects of the disclosure relate to compositions and methods for delivering a transgene (*e.g.*, a transgene encoding one or more gene products) to a target cell. The disclosure is based, in part, on adeno-associated virus (AAV) capsid protein variants characterized by tropisms for certain cell types (*e.g.*, neurons, muscle cells, bone cells, heart cells, etc.). In some embodiments, recombinant AAVs (rAAVs) comprising the capsid protein variants are more efficiently packaged than rAAVs having certain wild-type AAV capsid proteins. Methods of delivering an rAAV comprising the AAV capsid protein variants are also described by the disclosure.

In some aspects, the disclosure provides a method for delivering a transgene to a target cell in a subject, the method comprising intracranially administering to the subject a recombinant adeno-associated virus (rAAV) comprising: an isolated nucleic acid comprising a transgene encoding one or more gene products of interest; and an adeno-associated acid (AAV) capsid protein having the sequence set forth in SEQ ID NO: 1.

In some embodiments, intracranial administration comprises intrahippocampal injection.

In some embodiments, a target cell is a central nervous system (CNS) cell. In some embodiments, a CNS cell is a neuron, oligodendrocyte, astrocyte, or microglial cell.

In some embodiments, a subject is a mammal. In some embodiments, a subject is a human. In some embodiments, a subject is characterized by production of anti-AAV2 antibodies. In some embodiments, administration of the rAAV does not result in a neutralizing immune response against the rAAV by the subject.

In some embodiments, an isolated nucleic acid comprises AAV inverted terminal repeats (ITRs) flanking the transgene. In some embodiments, the nucleic acid sequence encoding the one or more gene products is operably linked to a promoter. In some embodiments, the one or more gene products comprise a protein or an inhibitory nucleic acid.

In some aspects, the disclosure provides a method for delivering a transgene to a target cell in a subject, the method comprising intravenously administering to the subject a recombinant adeno-associated virus (rAAV) comprising: an isolated nucleic acid comprising a transgene encoding one or more gene products of interest; and an adeno-associated acid (AAV) capsid protein having the sequence set forth in SEQ ID NO: 1, wherein the administration results in the rAAV crossing the blood brain barrier (BBB) of the subject.

In some embodiments, a target cell is a central nervous system (CNS) cell. In some embodiments, a CNS cell is a neuron, oligodendrocyte, astrocyte, or microglial cell.

In some embodiments, a subject is a mammal. In some embodiments, a subject is a human. In some embodiments, a subject is characterized by production of anti-AAV2 antibodies. In some embodiments, administration of the rAAV does not result in a neutralizing immune response against the rAAV by the subject.

In some embodiments, an isolated nucleic acid comprises AAV inverted terminal repeats (ITRs) flanking the transgene. In some embodiments, the nucleic acid sequence encoding the one or more gene products is operably linked to a promoter. In some embodiments, the one or more gene products comprise a protein or an inhibitory nucleic acid.

In some embodiments, recombinant AAVs (rAAVs) comprising the capsid protein variants described herein (*e.g.*, AAVv66, SEQ ID NO: 1) are more efficiently packaged (*e.g.*, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 30-fold, 50-fold, 100-fold, or more) than rAAVs having certain wild-type AAV capsid proteins (*e.g.*, AAV2 capsid protein, SEQ ID NO: 2).

BRIEF DESCRIPTION OF DRAWINGS

FIGs. 1A-1D show identification of a novel proviral AAV capsid sequences from a human surgical sample. FIG. 1A shows that AAV capsid proviral sequences were first PCR amplified from a human surgical sample using primers that flank the AAV Cap ORF. Amplicons were subjected to single molecule, real-time (SMRT) sequencing and the resulting reads were analyzed by BWA-MEM alignment to contemporary AAV serotype sequences, InDelFixer to remove insertion/deletions related to PCR or SMRT sequencing errors, and de novo assembly to cluster reads of high sequence similarity. FIG. 1B shows that the cap sequence of variant

AAVv66 was found to be the most abundant in the analysis (45%). FIG. 1C shows a summary of the 13 unique residues in the AAVv66 capsid sequence that are different from AAV2. (d) Phylogenetic tree of AAV2 variants (including AAVv66) and contemporary serotypes.

FIGs. 2A-2D show transduction spread of rAAV2 and rAAVv66 following intrahippocampal injection. FIG. 2A shows native EGFP expression following rAAV2-CB6-Egfp or rAAVv66-CB6-Egfp injection via unilateral intrahippocampal administration. Scale bars = 700 μm. FIG. 2B shows quantification of EGFP-positive surface normalized to DAPI-positive surface. Data is presented as the mean ±SD; n=3. ****P<0.0001. FIG. 2C shows coronal brain schematic depicting sub-anatomical regions of interest in both contralateral and ipsilateral hemispheres. Cornu ammonis (CA1, CA2, CA3, CA4), dentate gyrus (DG), corpus callosum (CC), and cortex (CTX). FIG. 2C shows high-magnification images of rAAVv66 transduced sub-anatomical regions. Scale bars = 50 μm.

FIGs. 3A-3P show transduction of major cell types of the brain by rAAVv66. FIGs. 3A, 3E, 3I, and 3M show coronal sections of rAAVv66-CB6-Egfp transduced mouse brains. IF-stained sections with antibodies against NEUN (FIG. 3A, neurons), GFAP (FIG. 3E, astrocytes), IBA1 (FIG. 3I, microglia), or OLIG2 (FIG. 3M, oligodendrocytes) indicate the distribution of cell types across the brain. Native EGFP expression that colocalize with IF staining indicate positively transduced cell types. Scale bars = 700 μm. FIGs. 3B, 3F, 3J, and 3N show 3D rendering of sub-anatomical regions of single representative frames from dashed line rectangle boxes within coronal section views (top panels) with single-cell representations from fields defined by dashed lined square boxes (bottom three panels). Left panels, total area EGFP and cell marker IF stains; center panels, colocalized EGFP with total cell marker IF stains; right panels, colocalized EGFP and cell marker IF stains. Scale bars = 50 μm (top panels), 5 μm (bottom three panels). FIGs. 3C, 3G, 3K, and 3O show quantification of cell type-specific IF staining across indicated hippocampal regions (x axes), normalized to DAPI signal. FIGs. 3D, 3H, 3L, and 3P show quantification of cell type-specific transduction across indicated regions, normalized to total cell-type IF and DAPI signal. Data is presented as the mean ±SD; n=3. Cornu ammonis (CA1, CA2, CA3, CA4), dentate gyrus (DG), corpus callosum (CC), and cortex (CTX).

FIGs. 4A-4E show biophysical analyses of AAVv66. Heatmap displays of differential scanning fluorimetry (DSF) analyses to query capsid protein unfolding (uncoating) (FIG. 4A) and DNA accessibility (vector genome extrusion) (FIG. 4B) at pHs 7, 6, 5, and 4. Each defining amino acid residue of AAVv66 was converted to those of AAV2 by site-directed mutagenesis

and examined for changes in packaging yield (FIG. 4C), capsid stability (FIG. 4D), and genome release (FIG. 4D) at pH 7. Values represent mean ± SD. p values were determined by one-way ANOVA. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001. n ³ 3.

FIGs. 5A-5E show cryo-EM primary metrics, map reconstruction, and model generation of AAVv66. FIG. 5A shows a density map of AAVv66. Grayscale scheme demarcates the topological distance from the center (Å). FIG. 5B shows a ribbon structure of the refined AAVv66 capsid monomer. Amino acids differentiating from AAV2 are highlighted. The 2-fold (oval), 3-fold (triangle), and 5-fold (pentagon) symmetries are annotated. Part of AAVv66 electron density (dark grey mesh) and residues are shown for regions close to (FIG. 5C) L583, R487, Y533, and K532, (FIG. 5D) S446, D499, and S501, and (FIG. 5E) N407-T414.

FIG. 6 shows structural differences between AAVv66 and AAV2. At the center is the AAVv66 60-mer structure (grey). Amino acid residues unique to AAVv66 are highlighted in green, while amino acid residues for a single monomer that are in common with AAV2 are colored. Atomic models showing residue side chains of select regions with substantial difference between AAVv66 and AAV2. The alignments were made with using monomers of AAV2 (1lp3) and AAVv66, with modeled side chains from neighboring residues displayed in grey. Annotations for amino acids shown are indicated as those belonging to AAVv66, the position number, and then AAV2.

FIGs. 7A-7C show differential capsid surface electrostatics between AAV2 and AAVv66. FIG. 7A shows surface positive and negative charges are displayed for AAV2 and AAVv66 60-mer, trimer (3-fold symmetry), and pentamers (exterior and interior of the 5-fold symmetry) structures. Black arrows at the AAV2 60-mer and trimer structures indicate the approximate positions of R585 and R588 at a single 3-fold protrusion. FIG. 7B shows a zoom-in of amino acid residues at 585-588 of AAV2 and AAVv66. FIG. 7C shows bar graphs of the zeta potentials of purified vectors as measured by a zetasizer. Values represent mean ± SD, n=3.

FIG. 8 shows amino acid sequence of the AAVv66 capsid protein with mutations relative to AAV2 provided. Amino acid differences between AAV2 and AAVv66 are highlighted. Variable region (VR) residues are denoted by short bars. The aH domain is demarcated by the dotted bar, and residues forming the b-sheets are marked with black arrows. Start positions for VP1, VP2, and VP3 are marked by greater-than symbol (>). The PLA domain within VP1 is denoted by a bar.

FIG. 9 shows AAVv66 produces higher vector yields than AAV2. Crude lysate PCR assays were performed on media and cellular lysates of HEK239 cells subjected to triple-

transfection of pAAV and packaging plasmids for AAV2 or AAVv66. Values represent mean genome copies ± SD, n=3.

FIG. 10 shows AAVv66 lacks strong heparin binding. Heparin competition assay showing transduction efficiency of AAV2-CB6-FLuc and AAVv66-CB6-FLuc in HEK293 cells in the presence of increasing amounts of heparin (x-axis). Luminescence values were scaled to values obtained for wells lacking heparin and set to 1 (y-axis). Values represent mean ± SD, n=3. **, $p < 0.01$ by 2-way ANOVA.

FIG. 11 shows in vitro infection efficiencies of AAV2, AAV3b, and AAVv66 in HEK293 cells. Vectors were packaged with CB6-FLuc. Cells were lysed 48-hr post-infection to assess the infectivity of vectors via detection of luciferase activity (RLU, relative light units). Data is displayed in log-scale. Values represent mean ± SD, ***$p < 0.0001$ by one-way ANOVA, n=3.

FIGs. 12A-12D show intravenous administration of AAVv66 vector shows transduction of the liver. Systemic injection of AAVv66-CB6-Fluc resulted in the transduction of the liver. rAAV2-CB6-Fluc or AAVv66-CB6-Fluc (1.0E11 GC/mouse) was injected into mice by tail vein administration. After 14 days, mice were injected with luciferin substrate intraperitoneally and imaged (FIG. 12A). Although quantification of whole-body live bioluminescence of luciferase activity did not reveal significant differences in transduction of the liver between AAVv66-CB6-Fluc and AAV2-CB6-Fluc, isolation of liver tissues and quantification of luciferase activity and detection of vector genome copy by qPCR showed that AAVv66 is a significantly weaker transducer of liver than AAV2. Total flux of the abdomen in acquired images was recorded (FIG. 12B). Tissues were harvested and assayed for luciferase activity (FIG. 12C) and vector genome abundance by qPCR (FIG. 12D). Values represent mean ± SD, n=3. *, $p < 0.05$ by Student's t test.

FIGs. 13A-13D show intramuscular administration of AAVv66 vector shows transduction of muscle. Intramuscular injection of AAVv66 into the tibialis anterior resulted in very little difference in transduction capacity when compared with the transduction of AAV2. AAV2-CB6-FLuc or AAVv66-CB6-FLuc (4.0E10 GC/mouse) was injected into mice by intramuscular administration into one hindlimb (tibialis anterior). After 14 days, mice were injected with luciferin substrate intraperitoneally and imaged (FIG. 13A). Total flux of the injected hindlimb in acquired images was recorded (FIG. 13B). Tissues were harvested and assayed for luciferase activity (FIG. 13C) and vector genome abundance by qPCR (FIG. 13D). Values represent mean ± SD, n=3. *, $p < 0.05$ by Student's t test.

FIGs. 14A-14D show immunological characterization of AAVv66. Mice were intramuscularly administrated by AAV2-CB6-Egfp vector (1E11 GC/mouse). Four weeks after administration, sera were collected for testing neutralizing antibody (NAb) titers against AAV2 or AAVv66 infection. NAb50 values for AAV2 (FIG. 14A) and AAVv66 (FIG. 14B) are defined as the titer dilution that can block 50% of the total transduction achievable by the vector packaged with the LacZ reporter gene. Left, NAb table summaries of individual animals tested. Right, transduction efficiencies were plotted against various serum dilutions. Values represent mean ±SD. Dashed lines indicate mean NAb50 serum titers. After the four-week period, mice were intramuscularly administrated with AAV2-hA1AT or AAVv66-hA1AT (1E11 GC/mouse) on the contralateral hindlimb. Serum A1AT levels were measured by ELISA at weeks 5, 6, 7, and 8 (FIG. 14C). Values represent mean ±SD, n=3. n.s., not significant; *, $p < 0.05$; **, $p < 0.01$; and ***, $p < 0.001$ by 2-way ANOVA on cross-sectional data points. FIG. 14D shows rabbit anti-AAV serum cross-reactivity. Rabbit antisera raised against AAV serotypes was tested for NAb to AAVv66 versus the homologous AAV serotype to assess relative cross reactivity. Log2 values represent highest antibody dilution to achieve 50% inhibition of transduction.

FIGs. 15A-15B show cryo-EM primary metrics, map reconstruction, and model generation of AAVv66. FIG. 15A shows a cryo-electron micrograph of AAVv66. The scale bar represents 100 Å. FIG. 15B shows fourier shell correlation for even and odd particles (FSC_part) for AAVv66.

FIG. 16 shows RMSD (Å) statistics comparing AAVv66 to AAV2 or AAV3b. Summary of the total and regional RMSD (Å) between AAVv66 and AAV2 (1LP3) or AAV3b (3KIC) measured across all alpha-carbon pairs indicated (AAV2 numbering) calculated by the rms_cur function within PyMOL. Full capsid structures of AAV2, 3b, and AAVv66 were aligned through optimized fit within the cryo-EM density map of AAVv66. Using a custom script within PyMOL, the distance values (Å) between individual alpha-carbon pairs for either AAV2 (upper) or AAV3b (lower) were quantitatively transformed for representation as both color and radial thickness for the corresponding residues of AAVv66.

DETAILED DESCRIPTION

Aspects of the disclosure relate to compositions and methods for delivering a transgene (*e.g.*, a transgene encoding one or more gene products) to a target cell. The disclosure is based, in part, on adeno-associated virus (AAV) capsid protein variants characterized by tropisms for certain cell types (*e.g.*, neurons, muscle cells, bone cells, heart cells, etc.). In some

embodiments, recombinant AAVs (rAAVs) comprising the capsid protein variants are more efficiently packaged than rAAVs having certain wild-type AAV capsid proteins. Methods of delivering an rAAV comprising the AAV capsid protein variants are also described by the disclosure.

AAVv66 Capsid Protein

In some aspects, the disclosure provides methods for delivering a transgene to a target cell (*e.g.,* a target cell of the central nervous system (CNS)) in a subject, the methods comprising administering (*e.g.,* intracranially or intravenously) to the subject a recombinant adeno-associated virus (rAAV) comprising: an isolated nucleic acid comprising a transgene encoding one or more gene products of interest; and an adeno-associated acid (AAV) capsid protein comprising an AAVv66 capsid protein or a capsid protein having substantial homology to an AAVv66 capsid protein. In some embodiments, an AAVv66 protein comprises the amino acid sequence set forth in SEQ ID NO: 1.

In some embodiments, the AAVv66 capsid protein described herein comprises mutations relative to AAV2 selected from the group consisting of: K39Q, V151A, R447K, T450A, Q457M, S492A, E499D, F533Y, G546D, E548G, R585S, R588T, and A593T. In some embodiments, the AAVv66 capsid protein described herein comprises each of the following mutations relative to AAV2: K39Q, V151A, R447K, T450A, Q457M, S492A, E499D, F533Y, G546D, E548G, R585S, R588T, and A593T. In some embodiments, a capsid protein having substantial homology to an AAVv66 capsid protein comprises one or more mutations relative to AAV2 selected from the group consisting of: K39Q, V151A, R447K, T450A, Q457M, S492A, E499D, F533Y, G546D, E548G, R585S, R588T, and A593T. In some embodiments, an AAVv66 capsid protein or a capsid protein having substantial homology to an AAVv66 capsid protein comprises one or more mutations in its VP1, VP2, and/or VP3 regions, relative to AAV2. In some embodiments, an AAVv66 capsid protein or a capsid protein having substantial homology to an AAVv66 capsid protein comprises one or more mutations in the Variable Region (VR)-IV, VR-V, VR-VI, VT-VII, and/or VR-VIII, relative to AAV2. In some embodiments, an AAVv66 capsid protein or a capsid protein having substantial homology to an AAVv66 capsid protein comprises one or more mutations relative to AAV2 as shown in FIG. 1C.

"Homology" refers to the percent identity between two polynucleotide or two polypeptide moieties. The term "substantial homology", when referring to a nucleic acid, or

fragment thereof, indicates that, when optimally aligned with appropriate nucleotide insertions or deletions with another nucleic acid (or its complementary strand), there is nucleotide sequence identity in about 90 to 100% of the aligned sequences. When referring to a polypeptide, or fragment thereof, the term "substantial homology" indicates that, when optimally aligned with appropriate gaps, insertions or deletions with another polypeptide, there is nucleotide sequence identity in about 90 to 100% of the aligned sequences. The term "highly conserved" means at least 80% identity, preferably at least 90% identity, and more preferably, over 97% identity. In some cases, highly conserved may refer to 100% identity. Identity is readily determined by one of skill in the art by, for example, the use of algorithms and computer programs known by those of skill in the art.

As described herein, alignments between sequences of nucleic acids or polypeptides are performed using any of a variety of publicly or commercially available Multiple Sequence Alignment Programs, such as "Clustal W", accessible through Web Servers on the internet. Alternatively, Vector NTI utilities may also be used. There are also a number of algorithms known in the art that can be used to measure nucleotide sequence identity, including those contained in the programs described above. As another example, polynucleotide sequences can be compared using BLASTN, which provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences. Similar programs are available for the comparison of amino acid sequences, e.g., the "Clustal X" program, BLASTP. Typically, any of these programs are used at default settings, although one of skill in the art can alter these settings as needed. Alternatively, one of skill in the art can utilize another algorithm or computer program that provides at least the level of identity or alignment as that provided by the referenced algorithms and programs. Alignments may be used to identify corresponding amino acids between two proteins or peptides. A “corresponding amino acid” is an amino acid of a protein or peptide sequence that has been aligned with an amino acid of another protein or peptide sequence. Corresponding amino acids may be identical or non-identical. A corresponding amino acid that is a non-identical amino acid may be referred to as a variant amino acid.

In some aspects, the disclosure relates to an AAVv66 capsid protein (*e.g.*, an isolated nucleic acid encoding an AAVv66 capsid protein, a recombinant adeno-associated virus (rAAV) comprising an AAVv66 capsid protein, etc.), or a capsid protein having substantial homology to an AAVv66 capsid protein. In some embodiments, a capsid protein having substantial homology to an AAVv66 capsid protein is at least 50%, 60%, 70%, 80%, 90%, 95%, or 99%

identical to the amino acid sequence set forth in SEQ ID NO: 1. In some embodiments, a capsid protein having substantial homology to an AAVv66 capsid protein comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acid substitutions, insertions, or deletions, relative to the amino acid sequence set forth in SEQ ID NO:1. In some embodiments, a capsid protein having substantial homology to an AAVv66 capsid protein comprises more than 50 amino acid substitutions, insertions, or deletions, relative to the amino acid sequence set forth in SEQ ID NO:1.

The disclosure relates, in some aspects, to the surprising discovery that rAAVs comprising AAVv66 capsid proteins are able to be produced in higher quantities in mammalian cell lines (*e.g.*, HEK-293 cells) relative to rAAVs having certain other AAV capsid proteins (*e.g.*, AAV2 capsid proteins, AAV3B capsid proteins, etc.). In some embodiments, transduced mammalian (e.g., HEK) producer cells yield between about 1.5-fold and about 5-fold (*e.g.*, 1.5, 2, 3, 4, 5-fold) more rAAVs having AAVv66 capsid than mammalian (e.g., HEK) producer cells transduced with AAV2 capsid proteins. In some embodiments, transduced mammalian (e.g., HEK) producer cells yield between about 5% and about 50% (*e.g.*, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, etc.) more rAAVs having AAVv66 capsid than mammalian (e.g., HEK) producer cells transduced with AAV3B capsid proteins.

Aspects of the disclosure relate to the unexpectedly improved central nervous system (CNS) cell transduction efficiency of AAVv66 capsid protein (*e.g.*, rAAVs comprising AAVv66 capsid proteins) relative to rAAVs having AAV2 capsid proteins. In some embodiments, AAVv66-containing rAAVs transduce CNS cells at least 5%, 10%, 15%, 20%, 25%, 30%, 40%, 50%, 100%, 200%, 500%, 1000%, or more efficiently than AAV2-containing rAAVs. In some embodiments, the CNS cells comprise neurons, oligodendrocytes, astrocytes, or microglial cells.

Aspects of the disclosure relate to certain AAV capsid proteins (*e.g.*, AAVv66 capsid proteins) that are serologically distinct from other AAV capsid proteins (*e.g.*, AAV1, AAV2, AAV3B, AAV8, AAV9, AAVrh.8, AAVrh.10, etc.). Without wishing to be bound by any particular theory, rAAVs comprising AAVv66 capsid proteins are not subject to the neutralizing antibody response in a subject that is sero-positive for antibodies against certain other AAV capsids. Accordingly, in some embodiments, rAAVs comprising AAVv66 capsid protein may be useful as a second-line therapy for delivery of transgenes to subjects that have previously been administered AAV therapies, or that are sero-positive for certain AAV capsid neutralizing antibodies.

In some aspects, the disclosure relates to rAAV capsid proteins (*e.g.*, AAVv66 capsid protein) that exhibit increased thermostability relative to certain wild-type AAV capsid proteins (*e.g.*, AAV2 capsid protein). In some embodiments, an AAVv66 capsid protein is more thermostable than an AAV2 capsid protein at a pH ranging from about pH 4 to about pH 7. In some embodiments, thermostability is determined by calculating the melting temperature of a capsid protein. In some embodiments, an AAVv66 capsid protein is characterized by a melting temperature that is between about 5 °C and about 10 °C above the melting temperature of an AAV2 capsid protein, at a given pH (*e.g.*, between pH 4 and pH 7).

Isolated nucleic acids

In some aspects, the disclosure relates to isolated nucleic acids encoding certain AAV capsid protein variants (*e.g.*, AAVv66 capsid protein). A "nucleic acid" sequence refers to a DNA or RNA sequence. In some embodiments, the term nucleic acid captures sequences that include any of the known base analogues of DNA and RNA such as, but not limited to 4-acetylcytosine, 8-hydroxy-N6-methyladenosine, aziridinylcytosine, pseudoisocytosine, 5-(carboxyhydroxyl-methyl) uracil, 5-fluorouracil, 5-bromouracil, 5-carboxymethylaminomethyl-2-thiouracil, 5-carboxymethyl-aminomethyluracil, dihydrouracil, inosine, N6-isopentenyladenine, 1-methyladenine, 1-methylpseudo-uracil, 1-methylguanine, 1-methylinosine, 2,2-dimethyl-guanine, 2-methyladenine, 2-methylguanine, 3-methyl-cytosine, 5-methylcytosine, N6-methyladenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxy-amino-methyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarbonylmethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, oxybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, -uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, pseudouracil, queosine, 2-thiocytosine, and 2,6-diaminopurine.

In some embodiments, proteins and nucleic acids of the disclosure are isolated. As used herein, the term “isolated” means artificially obtained or produced. As used herein with respect to nucleic acids, the term “isolated” generally means: (i) amplified *in vitro* by, for example, polymerase chain reaction (PCR); (ii) recombinantly produced by cloning; (iii) purified, as by cleavage and gel separation; or (iv) synthesized by, for example, chemical synthesis. An isolated nucleic acid is one that is readily manipulable by recombinant DNA techniques well known in the art. Thus, a nucleotide sequence contained in a vector in which 5' and 3' restriction sites are known or for which polymerase chain reaction (PCR) primer sequences have been

disclosed is considered isolated but a nucleic acid sequence existing in its native state in its natural host is not. An isolated nucleic acid may be substantially purified, but need not be. For example, a nucleic acid that is isolated within a cloning or expression vector is not pure in that it may comprise only a tiny percentage of the material in the cell in which it resides. Such a nucleic acid is isolated, however, as the term is used herein because it is readily manipulable by standard techniques known to those of ordinary skill in the art. As used herein with respect to proteins or peptides, the term "isolated" generally refers to a protein or peptide that has been artificially obtained or produced (*e.g.*, by chemical synthesis, by recombinant DNA technology, *etc.*).

It should be appreciated that conservative amino acid substitutions may be made to provide functionally equivalent variants, or homologs of the capsid proteins. In some aspects the disclosure embraces sequence alterations that result in conservative amino acid substitutions. As used herein, a conservative amino acid substitution refers to an amino acid substitution that does not alter the relative charge or size characteristics of the protein in which the amino acid substitution is made. Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references that compile such methods, *e.g.*, Molecular Cloning: A Laboratory Manual, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or Current Protocols in Molecular Biology, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Conservative substitutions of amino acids include substitutions made among amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D. Therefore, one can make conservative amino acid substitutions to the amino acid sequence of the proteins and polypeptides disclosed herein.

Recombinant AAVs (rAAVs)

In some aspects, the disclosure provides isolated AAVs. As used herein with respect to AAVs, the term "isolated" refers to an AAV that has been artificially obtained or produced. Isolated AAVs may be produced using recombinant methods. Such AAVs are referred to herein as "recombinant AAVs". Recombinant AAVs (rAAVs) preferably have tissue-specific targeting capabilities, such that a transgene of the rAAV will be delivered specifically to one or more predetermined tissue(s). The AAV capsid is an important element in determining these tissue-specific targeting capabilities. Thus, an rAAV having a capsid appropriate for the tissue being targeted can be selected. In some embodiments, the rAAV comprises an AAVv66 capsid

protein. In some embodiments, the rAAV comprises a capsid protein having an amino acid sequence as set forth in SEQ ID NO: 1.

Methods for obtaining recombinant AAVs having a desired capsid protein are well known in the art. (See, for example, US 2003/0138772), the contents of which are incorporated herein by reference in their entirety). Typically the methods involve culturing a host cell which contains a nucleic acid sequence encoding an AAV capsid protein (*e.g.*, a nucleic acid encoding a polypeptide having a sequence as set forth in SEQ ID NO: 1) or fragment thereof; a functional *rep* gene; a recombinant AAV vector composed of, AAV inverted terminal repeats (ITRs) and a transgene; and sufficient helper functions to permit packaging of the recombinant AAV vector into the AAV capsid proteins. In some embodiments, capsid proteins are structural proteins encoded by a *cap* gene of an AAV. In some embodiments, AAVs comprise three capsid proteins, virion proteins 1 to 3 (named VP1, VP2 and VP3), all of which may be expressed from a single *cap* gene. Accordingly, in some embodiments, the VP1, VP2 and VP3 proteins share a common core sequence. In some embodiments, the molecular weights of VP1, VP2 and VP3 are respectively about 87 kDa, about 72 kDa and about 62 kDa. In some embodiments, upon translation, capsid proteins form a spherical 60-mer protein shell around the viral genome. In some embodiments, the protein shell is primarily comprised of a VP3 capsid protein. In some embodiments, the functions of the capsid proteins are to protect the viral genome, deliver the genome and interact with the host. In some aspects, capsid proteins deliver the viral genome to a host in a tissue specific manner. In some embodiments, VP1 and/or VP2 capsid proteins may contribute to the tissue tropism of the packaged AAV. In some embodiments, the tissue tropism of the packaged AAV is determined by the VP3 capsid protein. In some embodiments, the tissue tropism of an AAV is enhanced or changed by mutations occurring in the capsid proteins.

In some embodiments, the AAV variants described herein are variants of AAV2. AAV2 is known to efficiently transduce human central nervous system (CNS) tissue, kidney tissue, ocular tissue (*e.g.*, photoreceptor cells and retinal pigment epithelium (RPE)), and other tissues. Accordingly, in some embodiments, the AAV2 variants described herein may be useful for delivering gene therapy to CNS tissue, kidney tissue, or ocular tissue. In some embodiments, AAV capsid proteins described herein are useful for targeting other tissues, for example muscle tissue, liver tissue, or cardiac tissue. In some embodiments, AAV capsid proteins described herein (*e.g.*, AAVv66 capsid protein) are capable of crossing the blood brain barrier (BBB) of a subject when delivered intravenously or systemically injected.

In some aspects, AAV variants described herein may be useful for the treatment of CNS-related disorders. As used herein, a "CNS-related disorder" is a disease or condition of the central nervous system. A CNS-related disorder may affect the spinal cord (e.g., a myelopathy), brain (e.g., a encephalopathy) or tissues surrounding the brain and spinal cord. A CNS-related disorder may be of a genetic origin, either inherited or acquired through a somatic mutation. A CNS-related disorder may be a psychological condition or disorder, e.g., Attention Deficient Hyperactivity Disorder, Autism Spectrum Disorder, Mood Disorder, Schizophrenia, Depression, Rett Syndrome, etc. A CNS-related disorder may be an autoimmune disorder. A CNS-related disorder may also be a cancer of the CNS, e.g., brain cancer. A CNS-related disorder that is a cancer may be a primary cancer of the CNS, e.g., an astrocytoma, glioblastomas, etc., or may be a cancer that has metastasized to CNS tissue, e.g., a lung cancer that has metastasized to the brain. Further non-limiting examples of CNS-related disorders, include Parkinson's Disease, Lysosomal Storage Disease, Ischemia, Neuropathic Pain, Amyotrophic lateral sclerosis (ALS), Multiple Sclerosis (MS), and Canavan disease (CD).

In some embodiments, AAV variants described herein may target liver tissue. Accordingly, in some embodiments, AAV variants described herein may be useful for treatment of hepatic disease. As used herein a "hepatic disease" is a disease or condition of the liver. A hepatic disease may be of a genetic origin, either inherited or acquired through a somatic mutation. A hepatic disease may be a cancer of the liver, including but not limited to hepatocellular carcinoma (HCC), fibrolamellar carcinoma, cholangiocarcinoma, angiosarcoma and hepatoblastoma. Further non-limiting examples of pulmonary diseases include Alagille Syndrome, Alpha 1 Anti-Trypsin Deficiency, autoimmune hepatitis, biliary atresia, cirrhosis, cystic disease of the liver, fatty liver disease, galactosemia, gallstones, Gilbert's Syndrome, hemochromatosis, liver disease in pregnancy, neonatal hepatitis, primary biliary cirrhosis, primary sclerosing cholangitis, porphyria, Reye's Syndrome, sarcoidosis, toxic hepatitis, Type 1 Glycogen Storage Disease, tyrosinemia, viral hepatitis A, B, C, Wilson Disease, and schistosomiasis.

In some embodiments, AAV variants described herein may be useful for delivering gene therapy to ocular tissue (e.g., tissue or cells of the eye). Accordingly, in some embodiments, AAV variants described herein may be useful for the treatment of ocular disorders. As used herein, an "ocular disorder" is a disease or condition of the eye. An ocular disease may affect the eye, sclera, cornea, anterior chamber, posterior chamber, iris, pupil, lens, vitreous humor, retina, or optic nerve. An ocular disorder may be of a genetic origin, either inherited or acquired

through a somatic mutation. Non-limiting examples of ocular diseases and disorders include but are not limited to: age-related macular degeneration, retinopathy, diabetic retinopathy, macular edema, glaucoma, retinitis pigmentosa and eye cancer.

The components to be cultured in the host cell to package a rAAV vector in an AAV capsid may be provided to the host cell in trans. Alternatively, any one or more of the required components (e.g., recombinant AAV vector, rep sequences, cap sequences, and/or helper functions) may be provided by a stable host cell which has been engineered to contain one or more of the required components using methods known to those of skill in the art. Most suitably, such a stable host cell will contain the required component(s) under the control of an inducible promoter. However, the required component(s) may be under the control of a constitutive promoter. Examples of suitable inducible and constitutive promoters are provided herein, in the discussion of regulatory elements suitable for use with the transgene. In still another alternative, a selected stable host cell may contain selected component(s) under the control of a constitutive promoter and other selected component(s) under the control of one or more inducible promoters. For example, a stable host cell may be generated which is derived from 293 cells (which contain E1 helper functions under the control of a constitutive promoter), but which contain the rep and/or cap proteins under the control of inducible promoters. Still other stable host cells may be generated by one of skill in the art.

The recombinant AAV vector, rep sequences, cap sequences, and helper functions required for producing the rAAV of the disclosure may be delivered to the packaging host cell using any appropriate genetic element (vector). In some embodiments, a single nucleic acid encoding all three capsid proteins (e.g., VP1, VP2 and VP3) is delivered into the packaging host cell in a single vector. In some embodiments, nucleic acids encoding the capsid proteins are delivered into the packaging host cell by two vectors; a first vector comprising a first nucleic acid encoding two capsid proteins (e.g., VP1 and VP2) and a second vector comprising a second nucleic acid encoding a single capsid protein (e.g., VP3). In some embodiments, three vectors, each comprising a nucleic acid encoding a different capsid protein, are delivered to the packaging host cell. The selected genetic element may be delivered by any suitable method, including those described herein. The methods used to construct any embodiment of this disclosure are known to those with skill in nucleic acid manipulation and include genetic engineering, recombinant engineering, and synthetic techniques. See, e.g., Sambrook et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. Similarly, methods of generating rAAV virions are well known and the selection of a suitable

method is not a limitation on the present disclosure. See, e.g., K. Fisher et al, J. Virol., 70:520-532 (1993) and U.S. Pat. No. 5,478,745.

In some embodiments, recombinant AAVs may be produced using the triple transfection method (described in detail in U.S. Pat. No. 6,001,650). Typically, the recombinant AAVs are produced by transfecting a host cell with a recombinant AAV vector (comprising a transgene) to be packaged into AAV particles, an AAV helper function vector, and an accessory function vector. An AAV helper function vector encodes the "AAV helper function" sequences (e.g., rep and cap), which function in trans for productive AAV replication and encapsidation. Preferably, the AAV helper function vector supports efficient AAV vector production without generating any detectable wild-type AAV virions (e.g., AAV virions containing functional rep and cap genes). Non-limiting examples of vectors suitable for use with the present disclosure include pHLP19, described in U.S. Pat. No. 6,001,650 and pRep6cap6 vector, described in U.S. Pat. No. 6,156,303, the entirety of both incorporated by reference herein. The accessory function vector encodes nucleotide sequences for non-AAV derived viral and/or cellular functions upon which AAV is dependent for replication (e.g., "accessory functions"). The accessory functions include those functions required for AAV replication, including, without limitation, those moieties involved in activation of AAV gene transcription, stage specific AAV mRNA splicing, AAV DNA replication, synthesis of cap expression products, and AAV capsid assembly. Viral-based accessory functions can be derived from any of the known helper viruses such as adenovirus, herpesvirus (other than herpes simplex virus type-1), and vaccinia virus.

In some aspects, the disclosure provides transfected host cells. The term "transfection" is used to refer to the uptake of foreign DNA by a cell, and a cell has been "transfected" when exogenous DNA has been introduced inside the cell (e.g., across the cell membrane). A number of transfection techniques are generally known in the art. See, e.g., Graham et al. (1973) Virology, 52:456, Sambrook et al. (1989) Molecular Cloning, a laboratory manual, Cold Spring Harbor Laboratories, New York, Davis et al. (1986) Basic Methods in Molecular Biology, Elsevier, and Chu et al. (1981) Gene 13:197. Such techniques can be used to introduce one or more exogenous nucleic acids, such as a nucleotide integration vector and other nucleic acid molecules, into suitable host cells.

A “host cell” refers to any cell that harbors, or is capable of harboring, a substance of interest. Often a host cell is a mammalian cell. A host cell may be used as a recipient of an AAV helper construct, an AAV minigene plasmid, an accessory function vector, or other transfer DNA associated with the production of recombinant AAVs. The term includes the

progeny of the original cell that has been transfected. Thus, a "host cell" as used herein may refer to a cell that has been transfected with an exogenous DNA sequence. It is understood that the progeny of a single parental cell may not necessarily be completely identical in morphology or in genomic or total DNA complement as the original parent, due to natural, accidental, or deliberate mutation.

As used herein, the term "cell line" refers to a population of cells capable of continuous or prolonged growth and division in vitro. Often, cell lines are clonal populations derived from a single progenitor cell. It is further known in the art that spontaneous or induced changes can occur in karyotype during storage or transfer of such clonal populations. Therefore, cells derived from the cell line referred to may not be precisely identical to the ancestral cells or cultures, and the cell line referred to includes such variants.

As used herein, the terms "recombinant cell" refers to a cell into which an exogenous DNA segment, such as DNA segment that leads to the transcription of a biologically-active polypeptide or production of a biologically active nucleic acid such as an RNA, has been introduced.

Cells may also be transfected with a vector (e.g., helper vector) that provides helper functions to the AAV. The vector providing helper functions may provide adenovirus functions, including, e.g., E1a, E1b, E2a, and E4ORF6. The sequences of adenovirus gene providing these functions may be obtained from any known adenovirus serotype, such as serotypes 2, 3, 4, 7, 12 and 40, and further including any of the presently identified human types known in the art. Thus, in some embodiments, the methods involve transfecting the cell with a vector expressing one or more genes necessary for AAV replication, AAV gene transcription, and/or AAV packaging.

As used herein, the term "vector" includes any genetic element, such as a plasmid, phage, transposon, cosmid, chromosome, artificial chromosome, virus, virion, etc., that is capable of replication when associated with the proper control elements and which can transfer gene sequences between cells. Thus, the term includes cloning and expression vehicles, as well as viral vectors. In some embodiments, useful vectors are contemplated to be those vectors in which the nucleic acid segment (e.g., nucleic acid sequence) to be transcribed is positioned under the transcriptional control of a promoter. A "promoter" refers to a DNA sequence recognized by the synthetic machinery of the cell, or introduced synthetic machinery, that is required to initiate the specific transcription of a gene. The phrases "operatively positioned," "under control" or "under transcriptional control" means that the promoter is in the correct

location and orientation in relation to the nucleic acid to control RNA polymerase initiation and expression of the gene. The term "expression vector or construct" means any type of genetic construct containing a nucleic acid in which part or all of the nucleic acid encoding sequence is capable of being transcribed. In some embodiments, expression includes transcription of the nucleic acid, for example, to generate a biologically-active polypeptide product or inhibitory RNA (e.g., shRNA, miRNA, miRNA inhibitor) from a transcribed gene.

In some embodiments, a promoter is a Cytomegalovirus early enhancer/chicken β actin (CB6) promoter.

In some cases, an isolated capsid gene can be used to construct and package recombinant AAVs, using methods well known in the art, to determine functional characteristics associated with the capsid protein encoded by the gene. For example, isolated capsid genes can be used to construct and package a recombinant AAV (rAAV) comprising a reporter gene (e.g., B-Galactosidase, GFP, Luciferase, etc.). The rAAV can then be delivered to an animal (e.g., mouse) and the tissue targeting properties of the novel isolated capsid gene can be determined by examining the expression of the reporter gene in various tissues (e.g., heart, liver, kidneys) of the animal. Other methods for characterizing the novel isolated capsid genes are disclosed herein and still others are well known in the art.

The foregoing methods for packaging recombinant vectors in desired AAV capsids to produce the rAAVs of the disclosure are not meant to be limiting and other suitable methods will be apparent to the skilled artisan.

rAAV Vectors

"Recombinant AAV (rAAV) vectors" of the disclosure are typically composed of, at a minimum, a transgene and its regulatory sequences, and 5' and 3' AAV inverted terminal repeats (ITRs). It is this recombinant AAV vector which is packaged into a capsid protein and delivered to a selected target cell. In some embodiments, the transgene is a nucleic acid sequence, heterologous to the vector sequences, that encodes a polypeptide, protein, functional RNA molecule (*e.g.*, miRNA, miRNA inhibitor) or other gene product, of interest. The nucleic acid coding sequence is operatively linked to regulatory components in a manner that permits transgene transcription, translation, and/or expression in a cell of a target tissue.

The AAV sequences of the vector typically comprise the cis-acting 5' and 3' inverted terminal repeat sequences (See, *e.g.*, B. J. Carter, in "Handbook of Parvoviruses", ed., P. Tijsser, CRC Press, pp. 155 168 (1990)). The ITR sequences are about 145 bp in length. Preferably,

substantially the entire sequences encoding the ITRs are used in the molecule, although some degree of minor modification of these sequences is permissible. The ability to modify these ITR sequences is within the skill of the art. (See, *e.g.*, texts such as Sambrook et al, "Molecular Cloning. A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory, New York (1989); and K. Fisher et al., J Virol., 70:520 532 (1996)). An example of such a molecule employed in the present disclosure is a "cis-acting" plasmid containing the transgene, in which the selected transgene sequence and associated regulatory elements are flanked by the 5' and 3' AAV ITR sequences. The AAV ITR sequences may be obtained from any known AAV, including presently identified mammalian AAV types.

In some embodiments, the disclosure provides a self-complementary AAV vector. As used herein, the term "self-complementary AAV vector" (scAAV) refers to a vector containing a double-stranded vector genome generated by the absence of a terminal resolution site (TR) from one of the ITRs of the AAV. The absence of a TR prevents the initiation of replication at the vector terminus where the TR is not present. In general, scAAV vectors generate single-stranded, inverted repeat genomes, with a wild-type (wt) AAV TR at each end and a mutated TR (mTR) in the middle.

In some embodiments, the rAAVs of the present disclosure are pseudotyped rAAVs. Pseudotyping is the process of producing viruses or viral vectors in combination with foreign viral envelope proteins. The result is a pseudotyped virus particle. With this method, the foreign viral envelope proteins can be used to alter host tropism or an increased/decreased stability of the virus particles. In some aspects, a pseudotyped rAAV comprises nucleic acids from two or more different AAVs, wherein the nucleic acid from one AAV encodes a capsid protein and the nucleic acid of at least one other AAV encodes other viral proteins and/or the viral genome. In some embodiments, a pseudotyped rAAV refers to an AAV comprising an inverted terminal repeat (ITR) of one AAV serotype and a capsid protein of a different AAV serotype. For example, a pseudotyped AAV vector containing the ITRs of serotype X encapsidated with the proteins of Y will be designated as AAVX/Y (*e.g.,* AAV2/1 has the ITRs of AAV2 and the capsid of AAV1). In some embodiments, pseudotyped rAAVs may be useful for combining the tissue-specific targeting capabilities of a capsid protein from one AAV serotype with the viral DNA from another AAV serotype, thereby allowing targeted delivery of a transgene to a target tissue.

In addition to the major elements identified above for the recombinant AAV vector, the vector also includes conventional control elements necessary which are operably linked to the

transgene in a manner which permits its transcription, translation and/or expression in a cell transfected with the plasmid vector or infected with the virus produced by the disclosure. As used herein, "operably linked" sequences include both expression control sequences that are contiguous with the gene of interest and expression control sequences that act in trans or at a distance to control the gene of interest.

Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation (polyA) signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (*e.g.*, Kozak consensus sequence); sequences that enhance protein stability; and when desired, sequences that enhance secretion of the encoded product. A great number of expression control sequences, including promoters that are native, constitutive, inducible and/or tissue-specific, are known in the art and may be utilized.

As used herein, a nucleic acid sequence (*e.g.*, coding sequence) and regulatory sequences are said to be "operably" linked when they are covalently linked in such a way as to place the expression or transcription of the nucleic acid sequence under the influence or control of the regulatory sequences. If it is desired that the nucleic acid sequences be translated into a functional protein, two DNA sequences are said to be operably linked if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably linked to a nucleic acid sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide. Similarly two or more coding regions are operably linked when they are linked in such a way that their transcription from a common promoter results in the expression of two or more proteins having been translated in frame. In some embodiments, operably linked coding sequences yield a fusion protein. In some embodiments, operably linked coding sequences yield a functional RNA (*e.g.*, shRNA, miRNA, miRNA inhibitor).

For nucleic acids encoding proteins, a polyadenylation sequence generally is inserted following the transgene sequences and before the 3' AAV ITR sequence. A rAAV construct useful in the present disclosure may also contain an intron, desirably located between the promoter/enhancer sequence and the transgene. One possible intron sequence is derived from

SV-40, and is referred to as the SV-40 T intron sequence. Another vector element that may be used is an internal ribosome entry site (IRES). An IRES sequence is used to produce more than one polypeptide from a single gene transcript. An IRES sequence would be used to produce a protein that contains more than one polypeptide chains. Selection of these and other common vector elements are conventional and many such sequences are available [see, *e.g.*, Sambrook et al, and references cited therein at, for example, pages 3.18 3.26 and 16.17 16.27 and Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989]. In some embodiments, a Foot and Mouth Disease Virus 2A sequence is included in polyprotein; this is a small peptide (approximately 18 amino acids in length) that has been shown to mediate the cleavage of polyproteins (Ryan, M D et al., EMBO, 1994; 4: 928-933; Mattion, N M et al., J Virology, November 1996; p. 8124-8127; Furler, S et al., Gene Therapy, 2001; 8: 864-873; and Halpin, C et al., The Plant Journal, 1999; 4: 453-459). The cleavage activity of the 2A sequence has previously been demonstrated in artificial systems including plasmids and gene therapy vectors (AAV and retroviruses) (Ryan, M D et al., EMBO, 1994; 4: 928-933; Mattion, N M et al., J Virology, November 1996; p. 8124-8127; Furler, S et al., Gene Therapy, 2001; 8: 864-873; and Halpin, C et al., The Plant Journal, 1999; 4: 453-459; de Felipe, P et al., Gene Therapy, 1999; 6: 198-208; de Felipe, P et al., Human Gene Therapy, 2000; 11: 1921-1931.; and Klump, H et al., Gene Therapy, 2001; 8: 811-817).

The precise nature of the regulatory sequences needed for gene expression in host cells may vary between species, tissues or cell types, but shall in general include, as necessary, 5' non-transcribed and 5' non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, enhancer elements, and the like. Especially, such 5' non-transcribed regulatory sequences will include a promoter region that includes a promoter sequence for transcriptional control of the operably joined gene. Regulatory sequences may also include enhancer sequences or upstream activator sequences as desired. The vectors of the disclosure may optionally include 5' leader or signal sequences. The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

Examples of constitutive promoters include, without limitation, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer) [see, *e.g.*, Boshart et al, Cell, 41:521-530 (1985)], the SV40 promoter, the dihydrofolate reductase promoter, the β-actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1α promoter [Invitrogen].

Inducible promoters allow regulation of gene expression and can be regulated by exogenously supplied compounds, environmental factors such as temperature, or the presence of a specific physiological state, *e.g.*, acute phase, a particular differentiation state of the cell, or in replicating cells only. Inducible promoters and inducible systems are available from a variety of commercial sources, including, without limitation, Invitrogen, Clontech and Ariad. Many other systems have been described and can be readily selected by one of skill in the art. Examples of inducible promoters regulated by exogenously supplied promoters include the zinc-inducible sheep metallothionine (MT) promoter, the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter, the T7 polymerase promoter system (WO 98/10088); the ecdysone insect promoter (No et al, Proc. Natl. Acad. Sci. USA, 93:3346-3351 (1996)), the tetracycline-repressible system (Gossen et al, Proc. Natl. Acad. Sci. USA, 89:5547-5551 (1992)), the tetracycline-inducible system (Gossen et al, Science, 268:1766-1769 (1995), see also Harvey et al, Curr. Opin. Chem. Biol., 2:512-518 (1998)), the RU486-inducible system (Wang et al, Nat. Biotech., 15:239-243 (1997) and Wang et al, Gene Ther., 4:432-441 (1997)) and the rapamycin-inducible system (Magari et al, J. Clin. Invest., 100:2865-2872 (1997)). Still other types of inducible promoters that may be useful in this context are those that are regulated by a specific physiological state, *e.g.*, temperature, acute phase, a particular differentiation state of the cell, or in replicating cells only.

In another embodiment, the native promoter for the transgene will be used. The native promoter may be preferred when it is desired that expression of the transgene should mimic the native expression. The native promoter may be used when expression of the transgene must be regulated temporally or developmentally, or in a tissue-specific manner, or in response to specific transcriptional stimuli. In a further embodiment, other native expression control elements, such as enhancer elements, polyadenylation sites or Kozak consensus sequences may also be used to mimic the native expression.

In some embodiments, the regulatory sequences impart tissue-specific gene expression capabilities. In some cases, the tissue-specific regulatory sequences bind tissue-specific transcription factors that induce transcription in a tissue specific manner. Such tissue-specific regulatory sequences (*e.g.*, promoters, enhancers, *etc.*) are well known in the art. Exemplary tissue-specific regulatory sequences include, but are not limited to the following tissue specific promoters: a liver-specific thyroxin binding globulin (TBG) promoter, an insulin promoter, a glucagon promoter, a somatostatin promoter, a pancreatic polypeptide (PPY) promoter, a synapsin-1 (Syn) promoter, a creatine kinase (MCK) promoter, a mammalian desmin (DES)

promoter, a α-myosin heavy chain (a-MHC) promoter, a gastrointestinal-specific mucin-2 promoter, an eye-specific retinoschisin promoter, an eye-specific K12 promoter, a respiratory tissue-specific CC10 promoter, a respiratory tissue-specific surfactant protein C (SP-C) promoter, a breast tissue-specific PRC1 promoter, a breast tissue-specific RRM2 promoter, a urinary tract tissue-specific uroplakin 2 (UPII) promoter, a uterine tissue-specific lactoferrin promoter, or a cardiac Troponin T (cTnT) promoter. Other exemplary promoters include Beta-actin promoter, hepatitis B virus core promoter, Sandig et al., Gene Ther., 3:1002-9 (1996); alpha-fetoprotein (AFP) promoter, Arbuthnot et al., Hum. Gene Ther., 7:1503-14 (1996)), bone osteocalcin promoter (Stein et al., Mol. Biol. Rep., 24:185-96 (1997)); bone sialoprotein promoter (Chen et al., J. Bone Miner. Res., 11:654-64 (1996)), CD2 promoter (Hansal et al., J. Immunol., 161:1063-8 (1998); immunoglobulin heavy chain promoter; T cell receptor α-chain promoter, neuronal such as neuron-specific enolase (NSE) promoter (Andersen et al., Cell. Mol. Neurobiol., 13:503-15 (1993)), neurofilament light-chain gene promoter (Piccioli et al., Proc. Natl. Acad. Sci. USA, 88:5611-5 (1991)), and the neuron-specific vgf gene promoter (Piccioli et al., Neuron, 15:373-84 (1995)), among others which will be apparent to the skilled artisan.

In some embodiments, a tissue-specific regulatory sequence is a CNS-specific promoter. Examples of CNS-specific promoters include but are not limited to neuron-specific enolase (NSE) promoter (Andersen et al., Cell. Mol. Neurobiol., 13:503-15 (1993)), neurofilament light-chain gene promoter (Piccioli et al., Proc. Natl. Acad. Sci. USA, 88:5611-5 (1991)), and the neuron-specific vgf gene promoter (Piccioli et al., Neuron, 15:373-84 (1995)). In some embodiments, a CNS-specific promoter is an astrocyte-specific promoter such as the glial fibrillary acidic protein promoter. In some embodiments, a CNS-specific promoter is an neuronal promoter such as the Synapsin (Syn) promoter. In some embodiments, a CNS-specific promoter is a promoter of a gene selected from: neuronal nuclei (NeuN), glial fibrillary acidic protein (GFAP), adenomatous polyposis coli (APC), and ionized calcium-binding adapter molecule 1 (Iba-1). In some embodiments, a CNS-specific promoter is as described in Kügler S. (2016) Tissue-Specific Promoters in the CNS. In: Manfredsson F. (eds) Gene Therapy for Neurological Disorders. Methods in Molecular Biology, vol 1382. Humana Press, New York, NY.

In some embodiments, one or more bindings sites for one or more of miRNAs are incorporated in a transgene of a rAAV vector, to inhibit the expression of the transgene in one or more tissues of a subject harboring the transgene (*e.g,* detargeting of transgene expression in a cell-type specific manner). The skilled artisan will appreciate that binding sites may be selected

to control the expression of a transgene in a tissue specific manner. For example, binding sites for the liver-specific miR-122 may be incorporated into a transgene to inhibit expression of that transgene in the liver. The target sites in the mRNA may be in the 5' UTR, the 3' UTR or in the coding region. Typically, the target site is in the 3' UTR of the mRNA. Furthermore, the transgene may be designed such that multiple miRNAs regulate the mRNA by recognizing the same or multiple sites. The presence of multiple miRNA binding sites may result in the cooperative action of multiple RISCs and provide highly efficient inhibition of expression. The target site sequence may comprise a total of 5-100, 10-60, or more nucleotides. The target site sequence may comprise at least 5 nucleotides of the sequence of a target gene binding site.

In some embodiments, a transgene comprises one or more (e.g., 1, 2, 3, 4, 5, or more) miRNA binding sites that de-target expression of the transgene from immune cells (*e.g.,* antigen presenting cells (APCs), such as macrophages, dendrites, etc.). Incorporation of miRNA binding sites for immune-associated miRNAs may de-target transgene expression from antigen presenting cells and thus reduce or eliminate immune responses (cellular and/or humoral) produced in the subject against products of the transgene, for example as described in US 2018/0066279, the entire contents of which are incorporated herein by reference. In some embodiments, the immune-associated miRNA is selected from: miR-15a, miR-16-1, miR-17, miR-18a, miR-19a, miR-19b-1, miR-20a, miR-21, miR-29a/b/c, miR-30b, miR-31, miR-34a, miR-92a-1, miR-106a, miR-125a/b, miR-142-3p, miR-146a, miR-150, miR-155, miR-181a, miR-223 and miR-424, miR-221, miR-222, let-7i, miR-148, and miR-152.

The composition of the transgene sequence of the rAAV vector will depend upon the use to which the resulting vector will be put. For example, one type of transgene sequence includes a reporter sequence, which upon expression produces a detectable signal. In another example, the transgene encodes a therapeutic protein or therapeutic functional RNA. In another example, the transgene encodes a protein or functional RNA that is intended to be used for research purposes, *e.g.*, to create a somatic transgenic animal model harboring the transgene, *e.g.*, to study the function of the transgene product. In another example, the transgene encodes a protein or functional RNA that is intended to be used to create an animal model of disease. Appropriate transgene coding sequences will be apparent to the skilled artisan.

Reporter sequences that may be provided in a transgene include, without limitation, DNA sequences encoding β-lactamase, β -galactosidase (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), luciferase,

and others well known in the art. When associated with regulatory elements which drive their expression, the reporter sequences, provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence or other spectrographic assays, fluorescent activating cell sorting assays and immunological assays, including enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and immunohistochemistry. For example, where the marker sequence is the LacZ gene, the presence of the vector carrying the signal is detected by assays for β-galactosidase activity. Where the transgene is green fluorescent protein or luciferase, the vector carrying the signal may be measured visually by color or light production in a luminometer. Such reporters can, for example, be useful in verifying the tissue-specific targeting capabilities and tissue specific promoter regulatory activity of an rAAV.

In some aspects, the disclosure provides rAAV vectors for use in methods of preventing or treating one or more genetic deficiencies or dysfunctions in a mammal, such as for example, a polypeptide deficiency or polypeptide excess in a mammal, and particularly for treating or reducing the severity or extent of deficiency in a human manifesting one or more of the disorders linked to a deficiency in such polypeptides in cells and tissues. The method involves administration of an rAAV vector that encodes one or more therapeutic peptides, polypeptides, siRNAs, microRNAs, antisense nucleotides, *etc.* in a pharmaceutically-acceptable carrier to the subject in an amount and for a period of time sufficient to treat the deficiency or disorder in the subject suffering from such a disorder.

Thus, the disclosure embraces the delivery of rAAV vectors encoding one or more peptides, polypeptides, or proteins, which are useful for the treatment or prevention of disease states in a mammalian subject. Exemplary therapeutic proteins include one or more polypeptides selected from the group consisting of growth factors, interleukins, interferons, anti-apoptosis factors, cytokines, anti-diabetic factors, anti-apoptosis agents, coagulation factors, anti-tumor factors. Other non-limiting examples of therapeutic proteins include BDNF, CNTF, CSF, EGF, FGF, G-SCF, GM-CSF, gonadotropin, IFN, IFG-1, M-CSF, NGF, PDGF, PEDF, TGF, VEGF, TGF-B2, TNF, prolactin, somatotropin, XIAP1, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-10 (187A), viral IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16 IL-17, and IL-18.

The rAAV vectors may comprise a gene to be transferred to a subject to treat a disease associated with reduced expression, lack of expression or dysfunction of the gene. In some embodiments, the rAAV vectors are used to treat a disease associated with the central nervous

system. Exemplary genes and associated disease states include, but are not limited to: glucose-6-phosphatase, associated with glycogen storage deficiency type 1A; phosphoenolpyruvate-carboxykinase, associated with Pepck deficiency; galactose-1 phosphate uridyl transferase, associated with galactosemia; phenylalanine hydroxylase, associated with phenylketonuria; branched chain alpha-ketoacid dehydrogenase, associated with Maple syrup urine disease; fumarylacetoacetate hydrolase, associated with tyrosinemia type 1; methylmalonyl-CoA mutase, associated with methylmalonic acidemia; medium chain acyl CoA dehydrogenase, associated with medium chain acetyl CoA deficiency; omithine transcarbamylase, associated with omithine transcarbamylase deficiency; argininosuccinic acid synthetase, associated with citrullinemia; low density lipoprotein receptor protein, associated with familial hypercholesterolemia; UDP-glucouronosyltransferase, associated with Crigler-Najjar disease; adenosine deaminase, associated with severe combined immunodeficiency disease; hypoxanthine guanine phosphoribosyl transferase, associated with Gout and Lesch-Nyan syndrome; biotinidase, associated with biotinidase deficiency; beta-glucocerebrosidase, associated with Gaucher disease; beta-glucuronidase, associated with Sly syndrome; peroxisome membrane protein 70 kDa, associated with Zellweger syndrome; porphobilinogen deaminase, associated with acute intermittent porphyria; alpha-1 antitrypsin for treatment of alpha-1 antitrypsin deficiency (emphysema); erythropoietin for treatment of anemia due to thalassemia or to renal failure; vascular endothelial growth factor, angiopoietin-1, and fibroblast growth factor for the treatment of ischemic diseases; thrombomodulin and tissue factor pathway inhibitor for the treatment of occluded blood vessels as seen in, for example, atherosclerosis, thrombosis, or embolisms; aromatic amino acid decarboxylase (AADC), and tyrosine hydroxylase (TH) for the treatment of Parkinson's disease; the beta adrenergic receptor, anti-sense to, or a mutant form of, phospholamban, the sarco(endo)plasmic reticulum adenosine triphosphatase-2 (SERCA2), and the cardiac adenylyl cyclase for the treatment of congestive heart failure; a tumor suppessor gene such as p53 for the treatment of various cancers; a cytokine such as one of the various interleukins for the treatment of inflammatory and immune disorders and cancers; dystrophin or minidystrophin and utrophin or miniutrophin for the treatment of muscular dystrophies; and, insulin for the treatment of diabetes.

The skilled artisan will also realize that in the case of transgenes encoding proteins or polypeptides, that mutations that results in conservative amino acid substitutions may be made in a transgene to provide functionally equivalent variants, or homologs of a protein or polypeptide. In some aspects the disclosure embraces sequence alterations that result in

conservative amino acid substitution of a transgene. In some embodiments, the transgene comprises a gene having a dominant negative mutation. For example, a transgene may express a mutant protein that interacts with the same elements as a wild-type protein, and thereby blocks some aspect of the function of the wild-type protein.

Useful transgene products also include miRNAs. miRNAs and other small interfering nucleic acids regulate gene expression via target RNA transcript cleavage/degradation or translational repression of the target messenger RNA (mRNA). miRNAs are natively expressed, typically as final 19-25 non-translated RNA products. miRNAs exhibit their activity through sequence-specific interactions with the 3' untranslated regions (UTR) of target mRNAs. These endogenously expressed miRNAs form hairpin precursors that are subsequently processed into a miRNA duplex, and further into a "mature" single stranded miRNA molecule. This mature miRNA guides a multiprotein complex, miRISC, which identifies target site, *e.g.*, in the 3' UTR regions, of target mRNAs based upon their complementarity to the mature miRNA.

The following non-limiting list of miRNA genes, and their homologues, are useful as transgenes or as targets for small interfering nucleic acids encoded by transgenes (*e.g.*, miRNA sponges, antisense oligonucleotides, TuD RNAs) in certain embodiments of the methods: hsa-let-7a, hsa-let-7a*, hsa-let-7b, hsa-let-7b*, hsa-let-7c, hsa-let-7c*, hsa-let-7d, hsa-let-7d*, hsa-let-7e, hsa-let-7e*, hsa-let-7f, hsa-let-7f-1*, hsa-let-7f-2*, hsa-let-7g, hsa-let-7g*, hsa-let-7i, hsa-let-7i*, hsa-miR-1, hsa-miR-100, hsa-miR-100*, hsa-miR-101, hsa-miR-101*, hsa-miR-103, hsa-miR-105, hsa-miR-105*, hsa-miR-106a, hsa-miR-106a*, hsa-miR-106b, hsa-miR-106b*, hsa-miR-107, hsa-miR-10a, hsa-miR-10a*, hsa-miR-10b, hsa-miR-10b*, hsa-miR-1178, hsa-miR-1179, hsa-miR-1180, hsa-miR-1181, hsa-miR-1182, hsa-miR-1183, hsa-miR-1184, hsa-miR-1185, hsa-miR-1197, hsa-miR-1200, hsa-miR-1201, hsa-miR-1202, hsa-miR-1203, hsa-miR-1204, hsa-miR-1205, hsa-miR-1206, hsa-miR-1207-3p, hsa-miR-1207-5p, hsa-miR-1208, hsa-miR-122, hsa-miR-122*, hsa-miR-1224-3p, hsa-miR-1224-5p, hsa-miR-1225-3p, hsa-miR-1225-5p, hsa-miR-1226, hsa-miR-1226*, hsa-miR-1227, hsa-miR-1228, hsa-miR-1228*, hsa-miR-1229, hsa-miR-1231, hsa-miR-1233, hsa-miR-1234, hsa-miR-1236, hsa-miR-1237, hsa-miR-1238, hsa-miR-124, hsa-miR-124*, hsa-miR-1243, hsa-miR-1244, hsa-miR-1245, hsa-miR-1246, hsa-miR-1247, hsa-miR-1248, hsa-miR-1249, hsa-miR-1250, hsa-miR-1251, hsa-miR-1252, hsa-miR-1253, hsa-miR-1254, hsa-miR-1255a, hsa-miR-1255b, hsa-miR-1256, hsa-miR-1257, hsa-miR-1258, hsa-miR-1259, hsa-miR-125a-3p, hsa-miR-125a-5p, hsa-miR-125b, hsa-miR-125b-1*, hsa-miR-125b-2*, hsa-miR-126, hsa-miR-126*, hsa-miR-1260, hsa-miR-1261, hsa-miR-1262, hsa-miR-1263, hsa-miR-1264, hsa-miR-1265, hsa-miR-1266, hsa-miR-

1267, hsa-miR-1268, hsa-miR-1269, hsa-miR-1270, hsa-miR-1271, hsa-miR-1272, hsa-miR-1273, hsa-miR-127-3p, hsa-miR-1274a, hsa-miR-1274b, hsa-miR-1275, hsa-miR-127-5p, hsa-miR-1276, hsa-miR-1277, hsa-miR-1278, hsa-miR-1279, hsa-miR-128, hsa-miR-1280, hsa-miR-1281, hsa-miR-1282, hsa-miR-1283, hsa-miR-1284, hsa-miR-1285, hsa-miR-1286, hsa-miR-1287, hsa-miR-1288, hsa-miR-1289, hsa-miR-129*, hsa-miR-1290, hsa-miR-1291, hsa-miR-1292, hsa-miR-1293, hsa-miR-129-3p, hsa-miR-1294, hsa-miR-1295, hsa-miR-129-5p, hsa-miR-1296, hsa-miR-1297, hsa-miR-1298, hsa-miR-1299, hsa-miR-1300, hsa-miR-1301, hsa-miR-1302, hsa-miR-1303, hsa-miR-1304, hsa-miR-1305, hsa-miR-1306, hsa-miR-1307, hsa-miR-1308, hsa-miR-130a, hsa-miR-130a*, hsa-miR-130b, hsa-miR-130b*, hsa-miR-132, hsa-miR-132*, hsa-miR-1321, hsa-miR-1322, hsa-miR-1323, hsa-miR-1324, hsa-miR-133a, hsa-miR-133b, hsa-miR-134, hsa-miR-135a, hsa-miR-135a*, hsa-miR-135b, hsa-miR-135b*, hsa-miR-136, hsa-miR-136*, hsa-miR-137, hsa-miR-138, hsa-miR-138-1*, hsa-miR-138-2*, hsa-miR-139-3p, hsa-miR-139-5p, hsa-miR-140-3p, hsa-miR-140-5p, hsa-miR-141, hsa-miR-141*, hsa-miR-142-3p, hsa-miR-142-5p, hsa-miR-143, hsa-miR-143*, hsa-miR-144, hsa-miR-144*, hsa-miR-145, hsa-miR-145*, hsa-miR-146a, hsa-miR-146a*, hsa-miR-146b-3p, hsa-miR-146b-5p, hsa-miR-147, hsa-miR-147b, hsa-miR-148a, hsa-miR-148a*, hsa-miR-148b, hsa-miR-148b*, hsa-miR-149, hsa-miR-149*, hsa-miR-150, hsa-miR-150*, hsa-miR-151-3p, hsa-miR-151-5p, hsa-miR-152, hsa-miR-153, hsa-miR-154, hsa-miR-154*, hsa-miR-155, hsa-miR-155*, hsa-miR-15a, hsa-miR-15a*, hsa-miR-15b, hsa-miR-15b*, hsa-miR-16, hsa-miR-16-1*, hsa-miR-16-2*, hsa-miR-17, hsa-miR-17*, hsa-miR-181a, hsa-miR-181a*, hsa-miR-181a-2*, hsa-miR-181b, hsa-miR-181c, hsa-miR-181c*, hsa-miR-181d, hsa-miR-182, hsa-miR-182*, hsa-miR-1825, hsa-miR-1826, hsa-miR-1827, hsa-miR-183, hsa-miR-183*, hsa-miR-184, hsa-miR-185, hsa-miR-185*, hsa-miR-186, hsa-miR-186*, hsa-miR-187, hsa-miR-187*, hsa-miR-188-3p, hsa-miR-188-5p, hsa-miR-18a, hsa-miR-18a*, hsa-miR-18b, hsa-miR-18b*, hsa-miR-190, hsa-miR-190b, hsa-miR-191, hsa-miR-191*, hsa-miR-192, hsa-miR-192*, hsa-miR-193a-3p, hsa-miR-193a-5p, hsa-miR-193b, hsa-miR-193b*, hsa-miR-194, hsa-miR-194*, hsa-miR-195, hsa-miR-195*, hsa-miR-196a, hsa-miR-196a*, hsa-miR-196b, hsa-miR-197, hsa-miR-198, hsa-miR-199a-3p, hsa-miR-199a-5p, hsa-miR-199b-5p, hsa-miR-19a, hsa-miR-19a*, hsa-miR-19b, hsa-miR-19b-1*, hsa-miR-19b-2*, hsa-miR-200a, hsa-miR-200a*, hsa-miR-200b, hsa-miR-200b*, hsa-miR-200c, hsa-miR-200c*, hsa-miR-202, hsa-miR-202*, hsa-miR-203, hsa-miR-204, hsa-miR-205, hsa-miR-206, hsa-miR-208a, hsa-miR-208b, hsa-miR-20a, hsa-miR-20a*, hsa-miR-20b, hsa-miR-20b*, hsa-miR-21, hsa-miR-21*, hsa-miR-210, hsa-miR-211, hsa-miR-212, hsa-miR-214, hsa-miR-214*, hsa-miR-215, hsa-miR-216a, hsa-miR-216b, hsa-miR-217,

hsa-miR-218, hsa-miR-218-1*, hsa-miR-218-2*, hsa-miR-219-1-3p, hsa-miR-219-2-3p, hsa-miR-219-5p, hsa-miR-22, hsa-miR-22*, hsa-miR-220a, hsa-miR-220b, hsa-miR-220c, hsa-miR-221, hsa-miR-221*, hsa-miR-222, hsa-miR-222*, hsa-miR-223, hsa-miR-223*, hsa-miR-224, hsa-miR-23a, hsa-miR-23a*, hsa-miR-23b, hsa-miR-23b*, hsa-miR-24, hsa-miR-24-1*, hsa-miR-24-2*, hsa-miR-25, hsa-miR-25*, hsa-miR-26a, hsa-miR-26a-1*, hsa-miR-26a-2*, hsa-miR-26b, hsa-miR-26b*, hsa-miR-27a, hsa-miR-27a*, hsa-miR-27b, hsa-miR-27b*, hsa-miR-28-3p, hsa-miR-28-5p, hsa-miR-296-3p, hsa-miR-296-5p, hsa-miR-297, hsa-miR-298, hsa-miR-299-3p, hsa-miR-299-5p, hsa-miR-29a, hsa-miR-29a*, hsa-miR-29b, hsa-miR-29b-1*, hsa-miR-29b-2*, hsa-miR-29c, hsa-miR-29c*, hsa-miR-300, hsa-miR-301a, hsa-miR-301b, hsa-miR-302a, hsa-miR-302a*, hsa-miR-302b, hsa-miR-302b*, hsa-miR-302c, hsa-miR-302c*, hsa-miR-302d, hsa-miR-302d*, hsa-miR-302e, hsa-miR-302f, hsa-miR-30a, hsa-miR-30a*, hsa-miR-30b, hsa-miR-30b*, hsa-miR-30c, hsa-miR-30c-1*, hsa-miR-30c-2*, hsa-miR-30d, hsa-miR-30d*, hsa-miR-30e, hsa-miR-30e*, hsa-miR-31, hsa-miR-31*, hsa-miR-32, hsa-miR-32*, hsa-miR-320a, hsa-miR-320b, hsa-miR-320c, hsa-miR-320d, hsa-miR-323-3p, hsa-miR-323-5p, hsa-miR-324-3p, hsa-miR-324-5p, hsa-miR-325, hsa-miR-326, hsa-miR-328, hsa-miR-329, hsa-miR-330-3p, hsa-miR-330-5p, hsa-miR-331-3p, hsa-miR-331-5p, hsa-miR-335, hsa-miR-335*, hsa-miR-337-3p, hsa-miR-337-5p, hsa-miR-338-3p, hsa-miR-338-5p, hsa-miR-339-3p, hsa-miR-339-5p, hsa-miR-33a, hsa-miR-33a*, hsa-miR-33b, hsa-miR-33b*, hsa-miR-340, hsa-miR-340*, hsa-miR-342-3p, hsa-miR-342-5p, hsa-miR-345, hsa-miR-346, hsa-miR-34a, hsa-miR-34a*, hsa-miR-34b, hsa-miR-34b*, hsa-miR-34c-3p, hsa-miR-34c-5p, hsa-miR-361-3p, hsa-miR-361-5p, hsa-miR-362-3p, hsa-miR-362-5p, hsa-miR-363, hsa-miR-363*, hsa-miR-365, hsa-miR-367, hsa-miR-367*, hsa-miR-369-3p, hsa-miR-369-5p, hsa-miR-370, hsa-miR-371-3p, hsa-miR-371-5p, hsa-miR-372, hsa-miR-373, hsa-miR-373*, hsa-miR-374a, hsa-miR-374a*, hsa-miR-374b, hsa-miR-374b*, hsa-miR-375, hsa-miR-376a, hsa-miR-376a*, hsa-miR-376b, hsa-miR-376c, hsa-miR-377, hsa-miR-377*, hsa-miR-378, hsa-miR-378*, hsa-miR-379, hsa-miR-379*, hsa-miR-380, hsa-miR-380*, hsa-miR-381, hsa-miR-382, hsa-miR-383, hsa-miR-384, hsa-miR-409-3p, hsa-miR-409-5p, hsa-miR-410, hsa-miR-411, hsa-miR-411*, hsa-miR-412, hsa-miR-421, hsa-miR-422a, hsa-miR-423-3p, hsa-miR-423-5p, hsa-miR-424, hsa-miR-424*, hsa-miR-425, hsa-miR-425*, hsa-miR-429, hsa-miR-431, hsa-miR-431*, hsa-miR-432, hsa-miR-432*, hsa-miR-433, hsa-miR-448, hsa-miR-449a, hsa-miR-449b, hsa-miR-450a, hsa-miR-450b-3p, hsa-miR-450b-5p, hsa-miR-451, hsa-miR-452, hsa-miR-452*, hsa-miR-453, hsa-miR-454, hsa-miR-454*, hsa-miR-455-3p, hsa-miR-455-5p, hsa-miR-483-3p, hsa-miR-483-5p, hsa-miR-484, hsa-miR-485-3p, hsa-miR-485-5p, hsa-miR-486-3p, hsa-miR-486-5p, hsa-miR-

487a, hsa-miR-487b, hsa-miR-488, hsa-miR-488*, hsa-miR-489, hsa-miR-490-3p, hsa-miR-490-5p, hsa-miR-491-3p, hsa-miR-491-5p, hsa-miR-492, hsa-miR-493, hsa-miR-493*, hsa-miR-494, hsa-miR-495, hsa-miR-496, hsa-miR-497, hsa-miR-497*, hsa-miR-498, hsa-miR-499-3p, hsa-miR-499-5p, hsa-miR-500, hsa-miR-500*, hsa-miR-501-3p, hsa-miR-501-5p, hsa-miR-502-3p, hsa-miR-502-5p, hsa-miR-503, hsa-miR-504, hsa-miR-505, hsa-miR-505*, hsa-miR-506, hsa-miR-507, hsa-miR-508-3p, hsa-miR-508-5p, hsa-miR-509-3-5p, hsa-miR-509-3p, hsa-miR-509-5p, hsa-miR-510, hsa-miR-511, hsa-miR-512-3p, hsa-miR-512-5p, hsa-miR-513a-3p, hsa-miR-513a-5p, hsa-miR-513b, hsa-miR-513c, hsa-miR-514, hsa-miR-515-3p, hsa-miR-515-5p, hsa-miR-516a-3p, hsa-miR-516a-5p, hsa-miR-516b, hsa-miR-517*, hsa-miR-517a, hsa-miR-517b, hsa-miR-517c, hsa-miR-518a-3p, hsa-miR-518a-5p, hsa-miR-518b, hsa-miR-518c, hsa-miR-518c*, hsa-miR-518d-3p, hsa-miR-518d-5p, hsa-miR-518e, hsa-miR-518e*, hsa-miR-518f, hsa-miR-518f*, hsa-miR-519a, hsa-miR-519b-3p, hsa-miR-519c-3p, hsa-miR-519d, hsa-miR-519e, hsa-miR-519e*, hsa-miR-520a-3p, hsa-miR-520a-5p, hsa-miR-520b, hsa-miR-520c-3p, hsa-miR-520d-3p, hsa-miR-520d-5p, hsa-miR-520e, hsa-miR-520f, hsa-miR-520g, hsa-miR-520h, hsa-miR-521, hsa-miR-522, hsa-miR-523, hsa-miR-524-3p, hsa-miR-524-5p, hsa-miR-525-3p, hsa-miR-525-5p, hsa-miR-526b, hsa-miR-526b*, hsa-miR-532-3p, hsa-miR-532-5p, hsa-miR-539, hsa-miR-541, hsa-miR-541*, hsa-miR-542-3p, hsa-miR-542-5p, hsa-miR-543, hsa-miR-544, hsa-miR-545, hsa-miR-545*, hsa-miR-548a-3p, hsa-miR-548a-5p, hsa-miR-548b-3p, hsa-miR-548b-5p, hsa-miR-548c-3p, hsa-miR-548c-5p, hsa-miR-548d-3p, hsa-miR-548d-5p, hsa-miR-548e, hsa-miR-548f, hsa-miR-548g, hsa-miR-548h, hsa-miR-548i, hsa-miR-548j, hsa-miR-548k, hsa-miR-548l, hsa-miR-548m, hsa-miR-548n, hsa-miR-548o, hsa-miR-548p, hsa-miR-549, hsa-miR-550, hsa-miR-550*, hsa-miR-551a, hsa-miR-551b, hsa-miR-551b*, hsa-miR-552, hsa-miR-553, hsa-miR-554, hsa-miR-555, hsa-miR-556-3p, hsa-miR-556-5p, hsa-miR-557, hsa-miR-558, hsa-miR-559, hsa-miR-561, hsa-miR-562, hsa-miR-563, hsa-miR-564, hsa-miR-566, hsa-miR-567, hsa-miR-568, hsa-miR-569, hsa-miR-570, hsa-miR-571, hsa-miR-572, hsa-miR-573, hsa-miR-574-3p, hsa-miR-574-5p, hsa-miR-575, hsa-miR-576-3p, hsa-miR-576-5p, hsa-miR-577, hsa-miR-578, hsa-miR-579, hsa-miR-580, hsa-miR-581, hsa-miR-582-3p, hsa-miR-582-5p, hsa-miR-583, hsa-miR-584, hsa-miR-585, hsa-miR-586, hsa-miR-587, hsa-miR-588, hsa-miR-589, hsa-miR-589*, hsa-miR-590-3p, hsa-miR-590-5p, hsa-miR-591, hsa-miR-592, hsa-miR-593, hsa-miR-593*, hsa-miR-595, hsa-miR-596, hsa-miR-597, hsa-miR-598, hsa-miR-599, hsa-miR-600, hsa-miR-601, hsa-miR-602, hsa-miR-603, hsa-miR-604, hsa-miR-605, hsa-miR-606, hsa-miR-607, hsa-miR-608, hsa-miR-609, hsa-miR-610, hsa-miR-611, hsa-miR-612, hsa-miR-613, hsa-miR-614, hsa-miR-615-3p, hsa-miR-615-5p, hsa-miR-616, hsa-

miR-616*, hsa-miR-617, hsa-miR-618, hsa-miR-619, hsa-miR-620, hsa-miR-621, hsa-miR-622, hsa-miR-623, hsa-miR-624, hsa-miR-624*, hsa-miR-625, hsa-miR-625*, hsa-miR-626, hsa-miR-627, hsa-miR-628-3p, hsa-miR-628-5p, hsa-miR-629, hsa-miR-629*, hsa-miR-630, hsa-miR-631, hsa-miR-632, hsa-miR-633, hsa-miR-634, hsa-miR-635, hsa-miR-636, hsa-miR-637, hsa-miR-638, hsa-miR-639, hsa-miR-640, hsa-miR-641, hsa-miR-642, hsa-miR-643, hsa-miR-644, hsa-miR-645, hsa-miR-646, hsa-miR-647, hsa-miR-648, hsa-miR-649, hsa-miR-650, hsa-miR-651, hsa-miR-652, hsa-miR-653, hsa-miR-654-3p, hsa-miR-654-5p, hsa-miR-655, hsa-miR-656, hsa-miR-657, hsa-miR-658, hsa-miR-659, hsa-miR-660, hsa-miR-661, hsa-miR-662, hsa-miR-663, hsa-miR-663b, hsa-miR-664, hsa-miR-664*, hsa-miR-665, hsa-miR-668, hsa-miR-671-3p, hsa-miR-671-5p, hsa-miR-675, hsa-miR-7, hsa-miR-708, hsa-miR-708*, hsa-miR-7-1*, hsa-miR-7-2*, hsa-miR-720, hsa-miR-744, hsa-miR-744*, hsa-miR-758, hsa-miR-760, hsa-miR-765, hsa-miR-766, hsa-miR-767-3p, hsa-miR-767-5p, hsa-miR-768-3p, hsa-miR-768-5p, hsa-miR-769-3p, hsa-miR-769-5p, hsa-miR-770-5p, hsa-miR-802, hsa-miR-873, hsa-miR-874, hsa-miR-875-3p, hsa-miR-875-5p, hsa-miR-876-3p, hsa-miR-876-5p, hsa-miR-877, hsa-miR-877*, hsa-miR-885-3p, hsa-miR-885-5p, hsa-miR-886-3p, hsa-miR-886-5p, hsa-miR-887, hsa-miR-888, hsa-miR-888*, hsa-miR-889, hsa-miR-890, hsa-miR-891a, hsa-miR-891b, hsa-miR-892a, hsa-miR-892b, hsa-miR-9, hsa-miR-9*, hsa-miR-920, hsa-miR-921, hsa-miR-922, hsa-miR-923, hsa-miR-924, hsa-miR-92a, hsa-miR-92a-1*, hsa-miR-92a-2*, hsa-miR-92b, hsa-miR-92b*, hsa-miR-93, hsa-miR-93*, hsa-miR-933, hsa-miR-934, hsa-miR-935, hsa-miR-936, hsa-miR-937, hsa-miR-938, hsa-miR-939, hsa-miR-940, hsa-miR-941, hsa-miR-942, hsa-miR-943, hsa-miR-944, hsa-miR-95, hsa-miR-96, hsa-miR-96*, hsa-miR-98, hsa-miR-99a, hsa-miR-99a*, hsa-miR-99b, and hsa-miR-99b*.

A miRNA inhibits the function of the mRNAs it targets and, as a result, inhibits expression of the polypeptides encoded by the mRNAs. Thus, blocking (partially or totally) the activity of the miRNA (*e.g.*, silencing the miRNA) can effectively induce, or restore, expression of a polypeptide whose expression is inhibited (derepress the polypeptide). In one embodiment, derepression of polypeptides encoded by mRNA targets of a miRNA is accomplished by inhibiting the miRNA activity in cells through any one of a variety of methods. For example, blocking the activity of a miRNA can be accomplished by hybridization with a small interfering nucleic acid (*e.g.*, antisense oligonucleotide, miRNA sponge, TuD RNA) that is complementary, or substantially complementary to, the miRNA, thereby blocking interaction of the miRNA with its target mRNA. As used herein, an small interfering nucleic acid that is substantially complementary to a miRNA is one that is capable of hybridizing with a miRNA, and blocking

the miRNA's activity. In some embodiments, a small interfering nucleic acid that is substantially complementary to a miRNA is a small interfering nucleic acid that is complementary to the miRNA at all but 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or 18 bases. In some embodiments, a small interfering nucleic acid sequence that is substantially complementary to a miRNA, or is a small interfering nucleic acid sequence that is complementary to the miRNA with at least one base.

A "miRNA Inhibitor" is an agent that blocks miRNA function, expression and/or processing. For instance, these molecules include but are not limited to microRNA specific antisense, microRNA sponges, tough decoy RNAs (TuD RNAs) and microRNA oligonucleotides (double-stranded, hairpin, short oligonucleotides) that inhibit miRNA interaction with a Drosha complex. MicroRNA inhibitors can be expressed in cells from a transgenes of a rAAV vector, as discussed above. MicroRNA sponges specifically inhibit miRNAs through a complementary heptameric seed sequence (Ebert, M.S. Nature Methods, Epub August, 12, 2007). In some embodiments, an entire family of miRNAs can be silenced using a single sponge sequence. TuD RNAs achieve efficient and long-term-suppression of specific miRNAs in mammalian cells (See, *e.g.*, Takeshi Haraguchi, et al., Nucleic Acids Research, 2009, Vol. 37, No. 6 e43, the contents of which relating to TuD RNAs are incorporated herein by reference). Other methods for silencing miRNA function (derepression of miRNA targets) in cells will be apparent to one of ordinary skill in the art.

In some embodiments, the cloning capacity of the recombinant RNA vector may limit a desired coding sequence and may require the complete replacement of the virus's 4.8 kilobase genome. Large genes may, therefore, not be suitable for use in a standard recombinant AAV vector, in some cases. The skilled artisan will appreciate that options are available in the art for overcoming a limited coding capacity. For example, the AAV ITRs of two genomes can anneal to form head to tail concatamers, almost doubling the capacity of the vector. Insertion of splice sites allows for the removal of the ITRs from the transcript. Other options for overcoming a limited cloning capacity will be apparent to the skilled artisan.

Administration

The rAAVs may be delivered to a subject in compositions according to any appropriate methods known in the art. The rAAV, preferably suspended in a physiologically compatible carrier (*e.g.*, in a composition), may be administered to a subject, *e.g.,* host animal, such as a human, mouse, rat, cat, dog, sheep, rabbit, horse, cow, goat, pig, guinea pig, hamster, chicken,

turkey, or a non-human primate (*e.g.*, Macaque). In some embodiments a host animal does not include a human.

Delivery of the rAAVs to a mammalian subject may be by, for example, intramuscular injection or by administration into the bloodstream of the mammalian subject. Administration into the bloodstream may be by injection into a vein, an artery, or any other vascular conduit. In some embodiments, the rAAVs are administered into the bloodstream by way of isolated limb perfusion, a technique well known in the surgical arts, the method essentially enabling the artisan to isolate a limb from the systemic circulation prior to administration of the rAAV virions. A variant of the isolated limb perfusion technique, described in U.S. Pat. No. 6,177,403, can also be employed by the skilled artisan to administer the virions into the vasculature of an isolated limb to potentially enhance transduction into muscle cells or tissue. Moreover, in certain instances, it may be desirable to deliver the virions to the CNS of a subject. By "CNS" is meant all cells and tissue of the brain and spinal cord of a vertebrate. Thus, the term includes, but is not limited to, neuronal cells, glial cells, astrocytes, cereobrospinal fluid (CSF), interstitial spaces, bone, cartilage and the like. Recombinant AAVs may be delivered directly to the CNS or brain by injection into, *e.g.*, the ventricular region, as well as to the striatum (*e.g.*, the caudate nucleus or putamen of the striatum), spinal cord and neuromuscular junction, or cerebellar lobule, with a needle, catheter or related device, using neurosurgical techniques known in the art, such as by stereotactic injection (see, *e.g.*, Stein et al., J Virol 73:3424-3429, 1999; Davidson et al., PNAS 97:3428-3432, 2000; Davidson et al., Nat. Genet. 3:219-223, 1993; and Alisky and Davidson, Hum. Gene Ther. 11:2315-2329, 2000).

The compositions of the disclosure may comprise an rAAV alone, or in combination with one or more other viruses (*e.g.*, a second rAAV encoding having one or more different transgenes). In some embodiments, a composition comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more different rAAVs each having one or more different transgenes.

Suitable carriers may be readily selected by one of skill in the art in view of the indication for which the rAAV is directed. For example, one suitable carrier includes saline, which may be formulated with a variety of buffering solutions (*e.g.*, phosphate buffered saline). Other exemplary carriers include sterile saline, lactose, sucrose, calcium phosphate, gelatin, dextran, agar, pectin, peanut oil, sesame oil, and water. The selection of the carrier is not a limitation of the present disclosure.

Optionally, the compositions of the disclosure may contain, in addition to the rAAV and carrier(s), other conventional pharmaceutical ingredients, such as preservatives, or chemical

stabilizers. Suitable exemplary preservatives include chlorobutanol, potassium sorbate, sorbic acid, sulfur dioxide, propyl gallate, the parabens, ethyl vanillin, glycerin, phenol, and parachlorophenol. Suitable chemical stabilizers include gelatin and albumin.

The rAAVs are administered in sufficient amounts to transfect the cells of a desired tissue and to provide sufficient levels of gene transfer and expression without undue adverse effects. Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, direct delivery to the selected organ (*e.g.*, intraportal delivery to the liver), oral, inhalation (including intranasal and intratracheal delivery), intraocular, intravenous, intramuscular, subcutaneous, intradermal, intratumoral, intracranial (*e.g.*, intrahippocampal), and other parental routes of administration. Routes of administration may be combined, if desired.

The dose of rAAV virions required to achieve a particular "therapeutic effect," *e.g.*, the units of dose in genome copies/per kilogram of body weight (GC/kg), will vary based on several factors including, but not limited to: the route of rAAV virion administration, the level of gene or RNA expression required to achieve a therapeutic effect, the specific disease or disorder being treated, and the stability of the gene or RNA product. One of skill in the art can readily determine a rAAV virion dose range to treat a patient having a particular disease or disorder based on the aforementioned factors, as well as other factors that are well known in the art.

An effective amount of an rAAV is an amount sufficient to target infect an animal, target a desired tissue. In some embodiments, an effective amount of an rAAV is an amount sufficient to produce a stable somatic transgenic animal model. The effective amount will depend primarily on factors such as the species, age, weight, health of the subject, and the tissue to be targeted, and may thus vary between animals or tissues. For example, an effective amount of the rAAV is generally in the range of from about 1 ml to about 100 ml of solution containing from about 10^9 to 10^{16} genome copies. In some embodiments the rAAV is administered at a dose of 10^{10}, 10^{11}, 10^{12}, 10^{13}, 10^{14}, or 10^{15} genome copies per subject. In some embodiments the rAAV is administered at a dose of 10^{10}, 10^{11}, 10^{12}, 10^{13}, or 10^{14} genome copies per kg. In some cases, a dosage between about 10^{11} to 10^{12} rAAV genome copies is appropriate. In certain embodiments, 10^{12} rAAV genome copies is effective to target heart, liver, and pancreas tissues. In some cases, stable transgenic animals are produced by multiple doses of an rAAV.

In some embodiments, rAAV compositions are formulated to reduce aggregation of AAV particles in the composition, particularly where high rAAV concentrations are present (*e.g.*, ~10^{13} GC/ml or more). Methods for reducing aggregation of rAAVs are well-known in the art and, include, for example, addition of surfactants, pH adjustment, salt concentration

adjustment, *etc*. (See, *e.g.*, Wright FR, et al., Molecular Therapy (2005) 12, 171–178, the contents of which are incorporated herein by reference.)

Formulation of pharmaceutically-acceptable excipients and carrier solutions is well-known to those of skill in the art, as is the development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens.

Typically, these formulations may contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 70% or 80% or more of the weight or volume of the total formulation. Naturally, the amount of active compound in each therapeutically useful composition may be prepared is such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

In certain circumstances it will be desirable to deliver the rAAV-based therapeutic constructs in suitably formulated pharmaceutical compositions disclosed herein either subcutaneously, intraopancreatically, intranasally, parenterally, intravenously, intracranially (*e.g.*, intrahippocampally), intramuscularly, intrathecally, or orally, intraperitoneally, or by inhalation. In some embodiments, the administration modalities as described in U.S. Pat. Nos. 5,543,158; 5,641,515 and 5,399,363 (each specifically incorporated herein by reference in its entirety) may be used to deliver rAAVs. In some embodiments, a preferred mode of administration is by portal vein injection.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms. In many cases the form is sterile and fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper

fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

For administration of an injectable aqueous solution, for example, the solution may be suitably buffered, if necessary, and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the host. The person responsible for administration will, in any event, determine the appropriate dose for the individual host.

Sterile injectable solutions are prepared by incorporating the active rAAV in the required amount in the appropriate solvent with various other ingredients enumerated herein, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

The rAAV compositions disclosed herein may also be formulated in a neutral or salt form. Pharmaceutically acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such

organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, drug-release capsules, and the like.

As used herein, "carrier" includes any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Supplementary active ingredients can also be incorporated into the compositions. The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a host.

Delivery vehicles such as liposomes, nanocapsules, microparticles, microspheres, lipid particles, vesicles, and the like, may be used for the introduction of the compositions of the present disclosure into suitable host cells. In particular, the rAAV vector delivered trangenes may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like.

Such formulations may be preferred for the introduction of pharmaceutically acceptable formulations of the nucleic acids or the rAAV constructs disclosed herein. The formation and use of liposomes is generally known to those of skill in the art. Recently, liposomes were developed with improved serum stability and circulation half-times (U.S. Pat. No. 5,741,516). Further, various methods of liposome and liposome like preparations as potential drug carriers have been described (U.S. Pat. Nos. 5,567,434; 5,552,157; 5,565,213; 5,738,868 and 5,795,587).

Liposomes have been used successfully with a number of cell types that are normally resistant to transfection by other procedures. In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, drugs, radiotherapeutic agents, viruses, transcription factors and allosteric effectors into a variety of cultured cell lines and animals. In addition, several successful clinical trials examining the effectiveness of liposome-mediated drug delivery have been completed.

Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs). MLVs generally have diameters of from 25 nm to 4 μm. Sonication of MLVs results in

the formation of small unilamellar vesicles (SUVs) with diameters in the range of 200 to 500.ANG., containing an aqueous solution in the core.

Alternatively, nanocapsule formulations of the rAAV may be used. Nanocapsules can generally entrap substances in a stable and reproducible way. To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use.

In addition to the methods of delivery described above, the following techniques are also contemplated as alternative methods of delivering the rAAV compositions to a host. Sonophoresis (*i.e.*, ultrasound) has been used and described in U.S. Pat. No. 5,656,016 as a device for enhancing the rate and efficacy of drug permeation into and through the circulatory system. Other drug delivery alternatives contemplated are intraosseous injection (U.S. Pat. No. 5,779,708), microchip devices (U.S. Pat. No. 5,797,898), ophthalmic formulations (Bourlais et al., 1998), transdermal matrices (U.S. Pat. Nos. 5,770,219 and 5,783,208) and feedback-controlled delivery (U.S. Pat. No. 5,697,899).

EXAMPLES

Example 1

Recombinant adeno-associated viruses (rAAVs) have recently gained a lot of attention within the human gene therapy field, as safe and reliable gene delivery vehicles. AAV2 is currently most commonly used in preclinical and clinical studies. However, AAV2-based drug Luxturna is the only FDA approved virus-based biotherapeutic, so it is critical to improve the pharmaceutical properties of AAV.

AAV2 is known to be a "poor producer" for vector production and "underperforms" in many tissue- and cell-types. A virus variant with improved properties was isolated.

A variant named AAVv66 was identified as the most abundant pro-viral capsid variant in a clinical pancreatic neoplasm sample. The AAVv66 capsid harbors 13 residues that differ from AAV2 (mutations relative to AAV2 including: K39Q, V151A, R447K, T450A, Q457M, S492A, E499D, F533Y, G546D, E548G, R585S, R588T, and A593T). The variant exhibits favorable tropism in the CNS following intracranial (*e.g.,* subcranial) injections. Furthermore, AAVv66 demonstrates better packaging efficiencies than prototypical AAV2. Using differential scanning fluorimetry (DSF), it was observed that the melting temperature of AAVv66 is ~6°C higher than

AAV2 across a pH range spanning pH4 - pH7. Furthermore, DSF analysis shows that at pH4, AAVv66 expunges its vector DNA at higher temperatures than AAV2.

It was also observed that AAVv66 confers superior CNS transduction relative to AAV2. Cryo-EM structure at 2.9Å resolution reveals structural differences between AAV2 and AAVv66 at the 3-fold protrusions and at the interface of the 5-fold axis of symmetry, indicating that residues at these positions confer improved stability and function for vector transduction.

Example 2

AAVs have recently attracted attention as effective and proven gene therapy vectors. The current class of AAV vectors confer stable, long-term gene expression, have a broad range of tissue tropisms, and exhibit relatively low pathogenicity. To date, three serotype capsids (AAV1, AAV2, and AAV9) have gained regulatory approval for commercial use in patients. Unfortunately, the current library of discovered and engineered AAV capsids falls short for certain clinical applications that require targeting of specific tissues or cell types. Furthermore, patients may have pre-existing immunity to the vector via neutralizing antibodies that would limit therapeutic efficacy. Additionally, certain capsids are known to be problematic under standard production schemes for generating high-yield titers needed to meet therapeutic doses. In response to these shortcomings, there is a need to discover and develop novel capsids that exhibit better vector yields, can escape innate immunity, and possess unique tropism profiles.

This example describes a capsid protein variant, AAVv66 (SEQ ID NO: 1) that was identified by high-throughput single molecule real-time (SMRT) sequencing and whose properties substantially differ from those of AAV2 despite high (98%) sequence similarity. First, AAVv66 exhibits better vector yield and is more thermostable than prototypical AAV2. Second, AAVv66 has a better spread of distribution within brain tissue when administrated by intracranial injections. Finally, AAVv66 is antigenically distinct from AAV2.

To better understand how AAVv66 differs from AAV2, cryogenic electron microscopy (cryo-EM) was performed to explore the structural and functional characteristics that define AAVv66. Our 2.5-Å resolution structure of the AAVv66 capsid reveals differences from the structure of AAV2 and provides insights into the functional properties the capsids. Together, these observations describe the mechanistic properties of AAVv66.

Materials and Methods

DNA Extraction

A pancreatic neoplasm sample was acquired from a 71-year-old female patient following tumorectomy and pathology of the tissue by frozen section examination and intraoperative frozen section diagnosis. The sample was stored in liquid nitrogen until DNA extraction. To avoid AAV DNA cross-contamination, DNA extraction, and PCR procedures were performed in a sterile UV-irradiated biosafety cabinet. All surfaces and equipment were sprayed with DNA-Exitus Plus (Applichem, Cat No: A7089) and wiped clean with milli-Q water after 15 minutes. Frozen tissues were then thawed at room temperature, quickly cut to about 25 mg of tissue with disposable scalpels and placed in a 2 mL tube. Extractions of DNA from tissues were performed using the QIAamp DNA Mini Kit (Qiagen, #51306) according to manufacturer's recommended procedures.

SMRT Sequencing

Amplicon libraries were generated from genomic DNA by standard PCR procedures. To amplify AAV genomes, PCR was performed using Platinum™ PCR SuperMix High Fidelity (Invitrogen) with the following cycle conditions: 97 °C for 1 min, 46 cycles of 98 °C for 10 s, 60 °C for 15 s, and 68 °C for 2 min 30 s; and 68 °C for 10min. Correctly-sized PCR products were gel-purified with PureLink™ PCR Purification Kit (Thermo Fisher) and used for a second round of 15-cycle PCR for barcoding. The primer pairs used were:

First Round Primers: CapF 5'-GACTGCATCTTTGAACAATAAATGA-3' (SEQ ID NO: 3) and CapR 5'-GAAACGAATTAACCGGTTTATTGATTAA-3' (SEQ ID NO: 4)

Second Round Primers: EF 5'-CATCACTACGCTAGATGACTGCATCTTTGAACAATAAATGA-3' (SEQ ID NO: 5) and ER 5'-TAGTATATCGAGACTCGAAACGAATTAACCGGTTTATTGATTAA-3' (SEQ ID NO: 6)

Amplicons representing the capsid variant ORFs were subjected to standard SMRT sequencing library generation. Sequencing was performed on the RSII platform. SMRT sequencing returned 17,727 DNA reads that mapped to the AAV2 Cap ORF using the BWA-MEM algorithm. To rule out artifactual sequences, reads were then filtered to exclude those that were less than 1,800 nt and more than 2,500 nt in length, and then filtered on the quality of reads (Phred score >30). This filtering reduced the reads to 14,500. Finally, reads were processed through InDelFixer to

remove single nucleotide insertions and deletions that may result from error-prone PCR or sequencing error. In order to consider only unique capsid sequences and to rule out low-confidence variants, de novo assembly (Geneious R9) was performed on the filtered reads to cluster reads with 99% of sequence similarity. Only read clusters represented by at least 10 reads were considered unique DNA capsid sequences. DNA sequences were then translated to amino acid sequences to define the final list of unique AAV capsids.

Full AAV Cap ORFs from contemporary AAV serotypes (hu.2 used for AAV2/3) were obtained from NCBI and the predicted amino acid sequences were aligned using the MUSCLE algorithm, iterating until convergence was achieved. PhyML was then used to generate the phylogenetic tree using default parameters from within SeaView55 and then visualized via the Interactive Tree of Life online tool.

Viral Vector Production

Viruses were produced using the triple-transfection method in HEK293 cells and purified by CsCl gradient centrifugation. All vectors described were packaged with either the self-complementary AAV vector expressing enhanced green fluorescence protein (scAAV-CB6-EGFP), single-strand vector expressing Firefly luciferase (ssAAV-CB6-Fluc), single-strand vector expressing secreted human alpha1-anti-trypsin (ssAAV-CB6-hA1AT), or single-strand vector expressing LacZ. All transgenes are driven by the CMV early enhancer/chicken β actin (CB6) ubiquitous promoter.

Animals

Six- to eight-week-old male C57BL/6J mice (The Jackson Laboratory) were injected by intravenous (IV), intramuscular (IM), or intracranial administration of test vectors. Mice subjected to IV injections were administered with vectors packaged with ssAAV-CB6-Fluc transgenes (1.0E11 vg/mouse), and mice were sacrificed 14 days post-injection. Mice subjected to IM injections of the TA muscle, were administered with vectors packaged with ssAAV-CB6-Fluc transgenes (4.0E10 vg/mouse), and mice were sacrificed 28-days post-injection. Every week, up until, and at time of sacrifice, animals were injected intraperitoneally with D-luciferin substrate and sedated with isoflurane and luciferase activity was quantified using the IVIS SpectrumCT imaging platform with 1 min exposures. Image acquisition was performed using Living Image software. Mice subjected to intrahippocampal injections were administered with vectors packaged with scAAV-CB6-Egfp transgenes (3.6E9 vg/mouse). Unilateral injections

were performed in the right hemisphere using a stereotaxic frame (Stoelting Co. Wood Dale, IL), Hamilton Syringe (1207K95, Thomas Scientific), and Hamilton Needle (77602-06, Hamilton). The following relative coordinates were used for all intra-hippocampal injections: x: -1.5mm, y: -2mm, z: -2mm.

Immunostaining

Four-weeks post-injection, animals were transcardially perfused with 1X phosphate buffered saline (PBS), followed by 4% paraformaldehyde (PFA). Brains were extracted and subsequently fixed in 4% PFA overnight at 4 °C. Brains were then immersed in 30% sucrose (prepared in 1X PBS), at 4 °C, until equilibrated in sucrose mixture. Brains were embedded in a 1:2 OCT (Tissue Tek, Torrance, CA) and 30% sucrose mixture, and cryo-sectioned at 40 µm (Cryostar NX70, ThermoScientific, Waltham, MA). Sections were permeabilized in 0.5% TritonX-100 for 1 hr, blocked in 5% goat serum (10% normal goat serum, 50062Z, Life Technologies) for 1 hr, and then incubated in primary antibody (anti-NeuN, 1:1000, EMD Millipore MAB377; anti-Gfap, 1:500, EMD Millipore MAB360; anti-Olig2, 1:200, Abcam ab109186; anti-Iba1, 1:1000, Wako Chemicals NC9288364) overnight at 4 °C. Sections were washed three times in 1X PBS and incubated in secondary antibody (anti-mouse, Invitrogen A32744; or anti-rabbit, Invitrogen A32740) for 1 hour at room temperature. Sections were washed three times in 1X PBS and mounted with Vectashield containing DAPI (Vector Laboratories, Burlingame, CA).

Microscopy

Brain section images were acquired on a Leica SP8 Lightning High Resolution Confocal (Leica Microsystems, Wetzlar, Germany). Global brain images (10X tiled brain sections) and high-magnification images (63X region specific areas) were collected at the same intensity and exposure thresholds for each respective magnification. For high-magnification images, 40-50 z-stack steps were collected at a 0.29 z-size. Analysis was performed using Imaris 9.3 Software (Bitplane Inc., Zurich, Switzerland). Each image was 3D rendered and thresholds were manually established. To ensure consistency, non-biased 3D rendering of total sub-anatomical EGFP volumes colocalized with DAPI volumes and cell type-specific stains were used as proxies for cellular counts and the number of positively transduced cells. Percent quantifications of the different cell types within each ipsilateral sub-anatomical region was conducted, followed by percentage quantification of each cell type. Percent transduction was determined by normalizing

colocalized EGFP volume to total volume of cell type-specific staining within each region. Per cell-specific stain, n =3 mice were analyzed. Statistical calculations for FIG. 2B was conducted in Prism 7 (GraphPad Software, Inc., San Diego, CA) and analysis was performed using Student's unpaired t-test.

DSF Analysis

For capsid stability experiments, 5 µL SYPRO Orange 5000X (Thermo Fisher Scientific) was diluted in 495 µL PBS (Corning) to make a 50X stock. 45 µL of virus was mixed with 5 µL of 50X SYPRO Orange (final SYPRO Orange concentration was 5X). Fluorescence was quantified using a ViiA 7 real-time PCR instrument (Thermo Fisher Scientific) with the following parameters: samples were incubated at 25 °C for 2 min, followed by a temperature gradient (25 °C to 99 °C, 0.4 °C per step and held at each step for 2 min). To monitor the fluorescence of the SYPRO Orange at each temperature step, the ROX filter was used with no passive reference. To investigate the effect of pH on the melting temperature of AAV vectors, 5 µL of the virus vectors, 5 µL of 50X SYBR Orange, and 40 µL of 0.6M acetate buffer pH-adjusted from pH7 to pH4 were mixed. Tm values reported in this study is defined as the max Dsignal/Dtemp detected between 25 and 95 °C. To investigate vector genome release, the SYBRO Orange dye was switched to SYBR Gold (Thermo Fisher Scientific).

Site-directed Mutagenesis

To generate point mutations in the AAVv66 capsid ORF, the Q5 site-directed mutagenesis kit (New England Biolabs) and following pairs of mutagenesis primers were used:

Table 1:

Mutation in AAVv66	**F. Primer (lowercase = mutated bases) (SEQ ID NOs: 7-19)**	**R. Primer (SEQ ID NOs: 20-32)**
Q39K	AGAGCGGCATaagGACGACAGCA	GCGGGCTTTGGTGGTGGT
A151V	GCATTCTCCTgtgGAGCCAGACT	TCTACCGGCCTCTTTTTTCC
K447R	TTACTTGAGCagaACAAACGCTC	TACAGATACTGGTCGATC
A450T	CAAAACAAACactCCAAGCGGAAC	CTCAAGTAATACAGATACTGG
M457Q	AACCACCACGcagTCCAGGCTTC	CCGCTTGGAGCGTTTGTT
A492S	ATCAAAAACAtctGCGGATAACAACAACAGTG	ACTCGCTGCTGGCGGTAA
D499E	CAACAACAGTgaaTATTCGTGGAC	TTATCCGCAGCTGTTTTTG
Y533F	TGAAGAAAAAtttTTTCCTCAGAGCGGGGTTC	TCGTCCTTGTGGCTGGCC
D546G	TGGAAAACAAggcTCGGGAAAAA	AAGATGAGAACCCCGCTC
G548E	ACAAGACTCGgagAAAACTAATGTG	TTTCCAAAGATGAGAACC

S585R	CAACCTCCAGagaGGCAACACAC	GTAGATACAGAACCATACTGCTC
T588R	GAGCGGCAACagaCAGGCAGCCA	TGGAGGTTGGTAGATACAGAACCATACTG
T593A	GGCAGCCACCgcaGATGTCAACA	TGTGTGTTGCCGCTCTGG

Cryo-EM

AAVv66 was prepared for cryo-EM on grids with a lacey carbon support film (01824G, Ted Pella, Inc.). First, the grids were washed with acetyl acetate and allowed to dry overnight. Next, the grids were glow discharged with 20 mA current with negative polarity for 60 sec in a PELCO easiGlow glow discharge unit. 3 µL of 1E13 vg/mL AAVv66-CB6-Egfp vector in buffer (5% sorbitol, 0.001% pluronic acid F68 in PBS) was placed onto the grids loaded on a Vitrobot Mark IV (ThermoFisher) cryo-EM plunging apparatus. The grids were blotted for 6 to 6.5 seconds with Whatman #1 filter paper at 10 °C and 95% relative humidity prior to rapid freezing in liquid ethane.

A data set consisting of 2,033 movies was collected using SerialEM on a Titan Krios electron microscope (FEI) operating at 300 kV and equipped with a Gatan Image Filter (GIF) and a K2 Summit direct electron detector (Gatan Inc.) using 0.5-2.2 µm underfocus. 50 frames per movie were collected, and 34 frames were used at 1.43 e-/Å2 per frame for a total dose of 48.62 e-/Å2 on the sample. Pixel size was 1.0588 Å on the sample. Movies were imported into cisTEM and were aligned with dose filtering and CTF parameters were determined. Next, a total of 52,874 particles were automatically picked within cisTEM (characteristic and maximum radius: 130 and 140 Å). We note that both particles encapsulating vector transgenes and the small percentage of empty capsids were used to determine the final structure. Within cisTEM, an initial reference for alignment was generated from all particles using the Ab initio 3D reconstruction function. This reference and all particles were iteratively refined using auto refine to obtain a 2.95-Å resolution map as determined from the FSC_part cutoff at 0.143. One round of per-particle CTF refinement in manual mode improved map resolution to 2.62 Å. Lastly, one round of beam tilt refinement and reconstruction improved map resolution to 2.46 Å. 3D classification did not improve the maps. The final map was B-factor sharpened by applying a B-factor of -32.92 Å2 using the PHENIX auto sharpen function.

Cryo-EM structure of AAV2 (PDB ID: 1LP3) was used as a starting model for structure refinement. Variant residues were modeled using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.). The resulting AAVv66 model, containing 60 copies of VP3, was refined using PHENIX59 against the cryo-EM map. Real-space simulated annealing

and B-factor refinement in PHENIX resulted in a stereochemically optimal model. The refinement results are summarized in Table 2. The model was inspected, and figures were prepared using PyMOL.

Vectors were diluted to a concentration of ~1.0E9 vg/mL for zeta potential analysis using the Zetasizer Nano ZS system (Malvern). 500 µL of sample was added into a universal dip cell (Malvern). Before measurement, the system was stabilized for 2 min. Three measurements were recorded for each sample.

Table 2

Cryo-EM data collection, refinement and validation statistics

	#1 name (EMDB-20630) (PDB 6U3Q)
Data collection and processing	
Magnification	47,214
Voltage (kV)	300
Electron exposure (e–/Å^2)	48.62
Defocus range (µm)	0.4-5.0
Pixel size (Å)	1.059
Symmetry imposed	I
Initial particle images (no.)	52,874
Final particle images (no.)	52,874
Map resolution (Å)	2.46
FSC threshold	0.143
Refinement	
Initial model used (PDB code)	1LP3
Model resolution (Å)	2.6
FSC threshold	0.5
Map sharpening *B* factor (Å^2)	32.92
Model composition	
Non-hydrogen atoms	248,280
Protein residues	31,140
Ligands	0
B factors (Å^2)	
Protein	86.48
Ligand	0
r.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	0.603
Validation	
MolProbity score	1.66
Clashscore	2.25
Poor rotamers (%)	4.15
Ramachandran plot	
Favored (%)	96.91
Allowed (%)	4.15
Disallowed (%)	0.00

Immunological Studies

1.0E11 vg/mice of scAAV-CB6-Egfp were intramuscularly administrated into the left/right tibialis anterior of C57BL/6J mice. Four weeks later, 1.0E11 vg/mice of ssAAVv66-CB6-hA1AT or ssAAV2-CB6-hA1AT was delivered to the contralateral leg. Serum was collected at weeks 4, 5, 6, 7, and 8 by facial vein bleeds to assess neutralizing antibody titers and A1AT levels by ELISA.

Huh-7.5 (5.0E4 cells/well) were seeded onto a 96-well plate 24 hr prior to transduction at 37 °C. Ad helper virus was then added at a multiplicity of infection (MOI) of 100:1 to the cell monolayer and incubated for at least an hour. Serial dilutions of serum and ssAAV2-LacZ or ssAAVv66-LacZ mixed solution were prepared in a V-bottom 96-well plate and incubated at 37 °C for 1 hr. The serum-AAV mixed solution was then added to cells and incubated at 37 °C for 24 hr. Cells were lysed and treated with beta-galactosidase substrate using the Galacto-Star One-Step Assay System (Invitrogen). Luminescence signal was detected by Synergy HT microplate reader (BioTek, Winooski, VT).

A1AT ELISA

A 96-well plate was first coated with anti-A1AT antibody at 4 °C overnight, and wells were incubated with blocking buffer (1% non-fat milk and 0.05% Tween-20 in PBS buffer) for 1 hr at room temperature. 1/20, 1/200 and 1/2,000 serum dilutions were performed using sample buffer (0.05% Tween-20 in PBS buffer) in a 96-well plate along with positive control (100, 50, 25, 12.5, 6.25, and 3.125 ng/mL A1AT). After the plate was washed 3 times, sera were added into each well and incubated at 4 °C overnight. The plate was then washed 3 times and incubated with goat anti-trypsin-HRP antibody (1:5,500 dilution in sample buffer) for 2 hr. Before reacting with substrate, the plate was washed 6 times to remove all residual proteins. Lastly, ABTS substrates were added into wells and the signal was read by a Synergy HT microplate reader (BioTek).

Identification of novel AAV variants in a human tissue sample by long-read sequencing

To identify novel full-length capsid sequences from human tissues, SMRT sequencing was performed to obtain long DNA reads that span the entire capsid open reading frame (FIG. 1A). This method can resolve sequences of long DNA fragments without the need for sequence assembly, which is necessary in short-read sequencing approaches. In this way, capsid diversity

that is defined by both point mutations and recombination events can be assessed on individual intact molecules that span the entire capsid ORF. To explore AAV diversity, a single tissue was selected from about 800 human surgical samples. Using primers that flank the capsid ORF at conserved sequence across known serotypes, target PCR amplicons were produced for SMRT sequencing analysis. One capsid sequence that made up ~45% of all sequences identified from the single tissue was isolated (FIG. 1B). This predominant capsid, named "variant 66" (AAVv66), exhibits closest homology to AAV2 (98% sequence similarity; FIGs. 1C-1D). It was observed that AAVv66 contains 13 amino acid residues that differ from AAV2 (FIG. 1C and FIG. 8): one within the VP1u region (K39Q), one within the VP2 domain (V151A), and eleven within VP3 (R447K, T450A, Q457M, S492A, E499D, F533Y, G546D, E548G, R585S, R588T and A593T). Notably, the unique amino acid residues within VP3 are all within or near variable regions VR-IV through VR-VIII.

The VP3 region of AAVv66 was compared with those of other contemporary AAV serotypes (AAV1-AAV9). The most notable differences occur at four positions (499, 533, 585, and 588), which are highly conserved among AAV serotypes (FIG. 8). At position 499, most serotypes harbor an asparagine, while AAVv66, AAV2, AAV4, and AAV9 have a negatively charged aspartic acid or glutamic acid. The highly conserved phenylalanine at position 533 is a tyrosine in AAVv66 (also, T533 in AAV5). Finally, unlike AAV2, which harbors positively charged arginine residues that define AAV2's capacity to bind heparan sulfate proteoglycans (HSPG) at positions 585 and 58823, AAVv66 contains S585 and T588 (identical to AAV1, AAV3, AAV5, and AAV6).

AAVv66 vector production and cell infectivity differ from those of AAV2

The strong affinity of AAV2 for heparin and its resulting strong cell-surface association is proposed to lead to the virus' relatively poor packaging titers. The limited vector yield by AAV2 is thought to result from non-productive binding and re-infection of the packaging cells by vector particles during production. Vector production and cell infectivity of AAVv66 was compared with those of AAV2 and AAV3b. Of note, AAV3b is the closest distinct cousin to AAV2 (89% sequence similarity), but uses different electrostatic surface charges at the 3-fold protrusions to weakly bind heparin. This difference between AAV3b and AAV2 likely explains AAV3b's increased packaging titers that result from transduction of HEK producer cells.

The packaging profiles of AAVv66 were compared with those of AAV2 and AAV3b by measuring the yields of encapsidated vector genomes in cell lysates. To this end, the AAVv66

capsid ORF was synthesized and cloned into a trans-plasmid expressing AAV2 Rep under the AAV2 p5 promoter (pAAV2/v66). Small-scale vector preparations of AAVv66, AAV2, and AAV3b were used to package a single-stranded vector consisting of the firefly luciferase transgene driven by the ubiquitous chicken-beta actin promoter (AAV-CB6-Fluc). Quantification of viral vector yields by crude lysate qPCR29 revealed that the yield of encapsidated, DNase-resistant genomes of AAVv66 vectors is ~2.4-fold higher than AAV2 yields and is ~30% higher than AAV3b yields (FIG. 9, "combined" samples).

Whether the higher abundance of AAVv66 in crude lysate is due to non-productive binding of particles to the packaging cells, which would be manifested by a predominance of AAVv66 particles in the media rather than in the cell lysate fraction, was investigated next. PCR analysis revealed that encapsidated genomes of AAVv66 within the media are ~3-fold more abundant than in cell lysates (FIG. 9). In contrast, very few AAV2 particles were detected in the media of packaging cells. To test whether AAVv66's ability to produce more DNase-resistant genomes is related to weak re-infectivity of packaging cells due to poor HSPG binding, a heparin competition assay was performed (FIG. 10). For this purpose, large-scale AAVv66 and AAV2 vectors, again packaging CB6-Fluc, were produced using the standard cesium chloride purification protocol. Transduction of AAVv66 is not affected by the presence of heparin, whereas 1.25 μg/well of heparin blocked AAV2 transduction by 50% and 5 μg/well of heparin completely abolished transduction. These results indicate that the improved production efficiency of AAVv66 is at least in part due to poor heparin binding.

To determine whether AAVv66's lower affinity to heparin coincides with reduced cell transduction compared to AAV2, purified AAVv66, AAV2, and AAV3b vectors were used to infect HEK293 cells. Data indicate that AAV2 exhibits greater transduction than AAVv66 (~65-fold) and AAV3b (~7.5-fold) (FIG. 11). Vectorized AAVv66 proviral capsid sequence is able to efficiently transduce cells in vitro, but its vector production and cell infectivity properties are distinct from those of its closest serotype relative, AAV2.

AAVv66 exhibits CNS transduction that is distinct from that of AAV2

AAVv66 was tested for its capacity to transduce selected target tissues via different routes of administration. To this end, biodistribution of AAVv66 was assessed in mice via multiple routes of delivery (FIGs. 12A-13D). Among all routes tested, the most striking was AAVv66's transduction profile following intracranial delivery to target cells of the central nervous system (CNS) (FIGs. 2A-2D). To determine whether AAVv66 has increased tropism in

the CNS, relative to that of AAV2, the Egfp transgene, driven by the ubiquitous chicken-beta actin promoter, was packaged into AAVv66 and AAV2 capsids. Vectors were unilaterally injected into the right hemisphere of the hippocampus at a dose of 3.6E9 vg/animal. Four-weeks post-injection, cryo-sections of treated brains showed that AAVv66 transduced ~13-fold more cells of the CNS than AAV2, as demonstrated by the enhanced spread throughout the tissue, while AAV2 tended to stay localized to the site of injection (FIGs. 2A-2B). High-magnification imaging of contralateral regions to the site of injection showed that all sub-anatomical regions of the brain (cornu ammonis [CA1, CA2, CA3, and CA4], dentate gyrus, and corpus callosum, FIG. 2C), exhibited detectable levels of EGFP expression (FIG. 2D), indicating that AAVv66 can spread efficiently throughout the hippocampal hemispheres.

The specific cell types that are transduced by AAVv66 were investigated. Antibody staining was performed with cell type-specific markers; anti-NEUN (neurons), anti-GFAP (astrocytes), anti-IBA1 (microglia), and anti-OLIG2 (oligodendrocytes) (FIGs. 3A, 3E, 3I, and 3M). 3D-volume reconstruction of sub-anatomical CNS regions demonstrates that EGFP expression colocalized with each investigated cell type (FIGs. 3B, 3J, 3F, 3N). Neurons were the predominant cell type found in the cortex and CA1 regions (FIG. 3C). Interestingly, CA2-4 regions and the dentate gyrus exhibited the greatest transduction (~20-40%). Astrocytes and microglia shared a similar distribution pattern, showing the highest enrichment in the dentate gyrus (FIGs. 3G and 3K). Astrocytes showed approximately 1-7% transduction across all regions (FIG. 3H), while microglia exhibited slightly higher transduction efficiencies (2-12%) (FIG. 3L). Oligodendrocytes were enriched in the corpus callosum (FIG. 3O) and had approximately 1-7% transduction by AAVv66 across all regions (FIG. 3P). These data indicate that AAVv66 can transduce all major cell types of the CNS following intrahippocampal injection.

AAVv66 is serologically distinct from AAV2

Neutralization of AAV by the host immune system is a major limiting factor for AAV vector transduction efficacy. Individuals harboring pre-existing antibodies against AAV serotypes that are used as capsids in therapeutic vectors are at greater risk to adverse effects and ineffectual treatment. Furthermore, patients requiring repeated administration of an AAV gene therapy, risk poorer transduction efficiencies and stronger immune responses, necessitating alternative vectors.

The question of whether AAVv66 transduction can be blocked by pre-immunization with AAV2 was investigated. To create pre-existing anti-AAV2 antibodies in circulation, AAV2-Egfp vectors (1E11 vg/mouse) were intramuscularly delivered to mice. Sera were collected after four weeks to assess neutralizing antibody (NAb) titers in vitro (FIGs. 13A-13D and FIGs. 14A-14B). Low NAb titers were needed to achieve 50% neutralization (NAb50) of AAV2 infection in Huh-7.5 cells (1/1,280~1/2,560), indicating that antibodies generated from AAV2 pre-immunization are sufficient to inhibit AAV2 transduction. By contrast, the NAb50 for AAVv66 infection with AAV2-treated mouse sera was 1/20~1/40, indicating that AAVv66 is able to infect cells despite the presence of NAbs generated against AAV2.

To test these findings with a secreted therapeutic transgene product in vivo, we re-dosed AAV2-immunized mice with AAV2 or AAVv66 packaged with the alpha-1 antitrypsin transgene (AAV2-A1AT or AAVv66-A1AT). Sera were collected at weeks 5, 6, 7, and 8, and secreted A1AT levels were quantified by ELISA31 (FIG. 14C). Low A1AT expression would suggest that NAbs generated from the first vector dose were preventing the transduction of the second vector dose. To establish a baseline of "maximal" A1AT expression, naïve mice were also treated in the same fashion. At weeks six and seven, A1AT expression in mice treated with AAV2-Egfp and then AAVv66-A1AT reached ~90% of A1AT expression as compared to naïve mice, whereas mice re-dosed with AAV2-A1AT reached only ~40% of naïve levels (FIG. 14C). These results concur with in vitro observations of robust infectivity by AAVv66 in the presence of sera from mice pre-immunized with AAV2 capsids.

Pre-immunity was also tested to investigate whether a broad range of AAV serotypes (AAV1, AAV2, AAV3b, AAV8, AAV9, AAV-DJ, AAVrh.8, and AAVrh.10) can compromise AAVv66 vector transduction. Antisera of rabbits separately pre-immunized with the eight serotypes were screened for AAVv66-vector neutralization. It was observed that, AAV1, AAV3b, and AAV-DJ exhibit about an order of magnitude difference in NAb50 titers compared to AAVv66, whereas AAV2, AAV8, AAVrh.8 and AAVrh.10 exhibit a two-order magnitude difference, and AAV9 had a three-order magnitude difference (FIG. 14D). Taken together, these data indicate that AAVv66 is serologically distinct from AAV2 and some other contemporary AAV capsids.

The AAVv66 capsid is more thermostable than AAV2 across a range of pHs

The efficient formation and structural stability of the capsid is essential to the production, purification, and storage of viral vectors. Additionally, for productive infection to

take place, vector particles must also maintain stability throughout the entry process and uncoat only under conditions in which delivery of the genomic payload can result in transduction of the cell. Although AAV vectors have been studied widely and are utilized for their strong transduction profiles in a range of tissues, the processes of intracellular trafficking, endosomal escape, and transportation of capsids into the nucleus are not fully understood. Among the presumed intracellular checkpoints impacting AAV intracellular trafficking and transduction that are dependent on capsid dynamics, endosomal escape is best understood. This process is believed to be triggered by a pH-dependent structural change of the capsid. Acidification of the endosomal lumen leads to a conformational change of the VP1 domain and exposure of the PLA2 domain within VP1, which triggers escape from the endosome compartment. In principle, vector capsids that can retain stability throughout intracellular trafficking are desirable and may exhibit high transduction capacity.

To determine the overall stability of AAVv66 capsids, differential scanning fluorimetry (DSF) analysis was used to measure the thermostability of the AAVv66 capsid across a range of physiological pHs (pH7-pH4) (FIG. 4). This range includes pH 4.5, which is observed in the lumen of late endosomes and lysosomes. In this assay, vector particles are suspended in SYPRO Orange dye, which fluoresces upon binding to hydrophobic residues in proteins. Thus, peak fluorescence signals are an indirect readout for maximally bound hydrophobic regions exposed upon protein unfolding. The melting temperatures (maximum slope values [Dsignal/Dtemp], Tm) for AAVv66 are more than five degrees higher than for AAV2 across all pH conditions tested. The most extreme difference was observed at pH 7, where the Tm of AAVv66 (75.29 ± 0.34 °C) is nearly 10 degrees higher than that of AAV2 (65.85 ± 0.18 °C) (FIG. 4A). Thus, the AAVv66 capsid is more thermally stable and resistant to pH than AAV2.

The effect of stability of AAVv66 capsid on vector genome release was investigated. Gauging vector genome release as a function of temperature range has been used as a proxy for pressure-driven DNA extrusion exerted by the nucleolar environment. The temperature dependence of AAVv66 genome release was compared with that of AAV2 under different pHs. To this end, DSF analysis was employed with SYBR Gold dye, which fluoresces upon binding to DNA. Peak fluorescence is an indirect measure of maximal accessibility of the encapsidated genomes to the dye solution. Vector genome release at pH 7 was observed to be concomitant with capsid stability, showing signal peaks at ~65 °C for AAV2 and ~74 °C for AAVv66. At lower pHs, however, peak fluorescence for dye-accessible DNA was detected at lower temperatures than peak fluorescence for unfolded capsid protein (FIG. 4B). Furthermore, DNA

accessibility for AAVv66 was more evident than that for AAV2 - where peak DNA accessibility at pHs 5 and 4 for AAV2 occurred at ~53 °C and ~42 °C, respectively; and AAVv66 exhibited peak signals at 25 °C. This surprising observation shows that in AAVv66 capsids DNA is especially accessible at low pHs (4-5) even at room temperature.

The question of whether AAVv66-specific amino acid residues contribute to the structural and functional differences observed between AAVv66 and AAV2 was then investigated. The thirteen AAVv66-defining amino acid residues were mutated to those of AAV2 and their impact on packaging vector genomes during production with HEK293 cells was tested (FIG. 4C). Additionally, thermal capsid stability and vector genome release were assessed for the mutant capsids (FIGs. 4D and 4E). All but four mutations (A151V, K447R, Y533F, and S585R) resulted in lower yields of DNase-resistant genomes, similar to or lower than those of AAV2 (FIG. 4C). Remarkably, the relatively conservative mutation D499E, which does not involve a charge change, lowered the packaging yield to ~5% of AAV2 yields. The modification also affected capsid stability, as D499E along with S585R and the S585R/T588R double mutation lowered Tm by 5.9 °C, 3.8 °C, and 5.4 °C, respectively (FIG. 4D), while other mutations affected the Tm by only 1~2 °C. Vector genome accessibilities of the same amino acid mutations exhibit lowered peak signal temperatures, while other mutations led to little or no change (FIG. 4E). Notably, the overall titers of purified vector were not drastically impacted (Table 3). Thus, the packaging yield of AAVv66 is only partially dependent on capsid stability, suggesting that partial capsid destabilization may be sufficient to facilitate genome release. Only residue D499 drastically affects both packaging and capsid stability.

Table 3

Virus titer after large-scale production and purification

Vector name	Titer GC/mL
ssAAV2-CB6-*FLuc*	1.60E+12
ssAAVv66-CB6-*FLuc*	6.00E+12
ssAAVv66-Q39K-CB6-*FLuc*	7.70E+12
ssAAVv66-A151V-CB6-*FLuc*	9.00E+12
ssAAVv66-K447R-CB6-*FLuc*	6.60E+12
ssAAVv66-A450T-CB6-*FLuc*	8.80E+12
ssAAVv66-M457Q-CB6-*FLuc*	1.00E+13
ssAAVv66-A492S-CB6-*FLuc*	1.40E+13
ssAAVv66-D499E-CB6-*FLuc*	9.00E+12
ssAAVv66-Y533F-CB6-*FLuc*	7.10E+12
ssAAVv66-D540F-CB6-*FLuc*	1.20E+13
ssAAVv66-G548E-CB6-*FLuc*	7.80E+12
ssAAVv66-S585R-CB6-*FLuc*	7.00E+12
ssAAVv66-T588R-CB6-*FLuc*	1.30E+13
ssAAVv66-T593A-CB6-*FLuc*	8.10E+12
ssAAVv66-S585R/T588R-CB6-*FLuc*	1.85E+12

Cryo-EM structure analysis of capsid differences between AAVv66 and AAV2

To characterize the structural properties of AAVv66, AAV2v66-Egfp vector was purified for cryo-EM analysis. 52,874 particle images were obtained, which yielded a cryo-EM map at 2.5 Å resolution (FIG. 5E and FIG. 16), and obtained a structural model with optimal real-space fit and stereochemical parameters (Table 2). Overall, the AAVv66 structure is similar to AAV2 (root-mean-square deviation (RMSD) of atomic coordinates = 0.456 Å) (FIG. 16). Thus, AAVv66 exhibits the characteristic features of an AAV capsid, which include the depression at two-fold axis, the three-fold symmetry that is defined by the three-fold protrusions, and the five-fold pore that is comprised of five monomers that form the interface and pore for Rep binding (FIG. 5A). Of note, VP1u and VP2 domains are each represented at approximately a twelfth of the VP3 domains for each particle, and similar to other AAV structures before, were not resolved in our symmetrized cryo-EM map. Therefore, only residues 219- 736 are definitively resolved within the cryo-EM map, including eleven of thirteen AAVv66-defining residues (FIG. 5B).

Comparison of the AAVv66 structure with that of AAV2 reveals several structural differences, which may contribute to the improved DNA packaging and/or capsid stability. Key differences occur at the interfaces between monomers of VP3 at the protrusions around the 3-fold axis. D499, whose mutation to the longer glutamate residue resulted in dramatic defects in

vector genome packaging (FIG. 4C), forms electrostatic interactions and/or hydrogen bonding with S501 (FIG. 5D). This region is tightly packed against the neighboring VP3 monomer (FIG. 5D). Here, the backbone atoms of D499 and S501 interact with the side chains of the symmetry-related N449 and T448, respectively, whereas the hydroxyl group of S501 hydrogen-bonds with the backbone carbonyl of the symmetry-related S446. The strong effect of D499 mutation is therefore likely due to the disruption of the interface between VP3 monomers, leading to destabilization of the capsid. In the same region, residues K447 and A450 of the neighboring monomer eliminate potential for electrostatic interaction between corresponding AAV2-R447 and T450 side chains (FIG. 6D). Amino acid M457 is located on the three-fold protrusion of AAVv66 at variable region IV, with the sidechain poised toward the solvent (FIG. 6E). Interestingly, this methionine is a unique feature among other serotypes (FIG. 8), suggesting at potential unique capsid interactions with cellular receptors, host factors, or antibodies. The polar hydroxyl group in AAVv66-Y533 (AAV2-F533) likely stabilizes the polar environment between the side chains of R487 and K532 and may contribute to the interaction with L583 of the symmetry-related monomer (FIG. 6A). AAVv66-D546 and G548 redistribute the surface charge conferred by AAV2-G546 and E548 (FIG. 6B) and is yet another defining feature of AAVv66.

A key functional region of AAV2 involves the positively charged arginine residues at position 585 and 588. These residues are at the surface of the three-fold protrusions and govern the capsid's interaction with HSPG receptors, which are vital to attachment and entry in many cell types. By contrast, S585 and T588 in the AAVv66 capsid are neutral charged polar residues (FIG. 5D and FIG. 6), similar to S586 and T589 of AAV3b (FIG. 8). AAV3b's physical and functional interactions with HSPG rely on electrostatic interactions conferred by residues R447 and R594 (R447 and A593 in AAV2)7, but AAVv66 also lacks these arginine residues (K447 and T593). These differences from AAV2 and AAV3b suggest that AAVv66 associates differently with the canonical cell surface receptor commonly utilized by AAV clade B and C capsids, consistent with our findings that AAVv66 lacks heparin binding.

AAV2 and AAVv66 show surface charge differences

Because electrostatic properties of the virus are important for capsid-receptor interaction7,43, how the net loss of positive charge for the AAVv66 capsid in relation to AAV2 affects the electrostatic properties of the capsid was investigated. First, the calculated electrostatic potential values for AAV2 and AAVv66 structures were compared (FIG. 7A). The

distribution of electrostatic potential on the surface of AAVv66 differs from that of AAV2. The most notable difference is at the three-fold protrusions, where the positive charge conferred by R585 and R588 in AAV2 is drastically reduced by S585 and T588 in AAVv66 (FIG. 7B).

Whether the distinct structure and surface electrostatics of AAVv66 affects the charge-dependent particle migration (zeta potential) of the capsid was then investigated (FIG. 7C). The zeta potential of AAVv66 (-10 mV) is remarkably different from that of AAV2 (-3.5 mV), consistent with differences in electrostatic potential between the capsids. To test the contributions of individual substitutions, particle migration of AAVv66 harboring single amino acid substitutions that convert residues to the corresponding residues of AAV2 was measured (FIG. 7C). Single mutations S585R and T588R resulted in the most dramatic change of the zeta potential (by ~3 mV each), bringing the zeta potential closer to that of AAV2 (FIG. 7C). These observations indicate that the electrostatic properties of AAVv66 differ from those of AAV2 and the difference is predominantly due to substitutions at positions 585 and 588. Thus, interactions of capsid AAVv66 with receptors, antibodies and other proteins likely differ substantially from those of other closely related capsids.

SELECTED SEQUENCES

AAVv66 amino acid sequence (SEQ ID NO: 1)

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHQDDSRGLVLPGYKYLGPFN
GLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQERLKEDTSFGGNL
GRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPAEPDSSSGTGKAGQQPARKRLN
FGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWH
CDSTWMGDRVITTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRF
HCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDS
EYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLR
TGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSKTNAPSGTTTMSRLQFSQ
AGASDIRDQSRNWLPGPCYRQQRVSKTAADNNNSDYSWTGATKYHLNGRDSLVNPGP
AMASHKDDEEKYFPQSGVLIFGKQDSGKTNVDIEKVMITDEEEIRTTNPVATEQYGSVS
TNLQSGNTQAATTDVNTQGVLPGMVWQDRDVYLQGPIWAKIPHTDGHFHPSPLMGGF
GLKHPPPQILIKNTPVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEI
QYTSNYNKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL

Wild-type AAV2 capsid protein amino acid sequence (SEQ ID NO: 2)

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFN
GLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQERLKEDTSFGGNL
GRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLN
FGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWH
CDSTWMGDRVITTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRF
HCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDS
EYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLR

TGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFSQ
AGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNPGP
AMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVST
NLQRGNRQAATADVNTQGVLPGMVWQDRDVYLQGPIWAKIPHTDGHFHPSPLMGGF
GLKHPPPQILIKNTPVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEI
QYTSNYNKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL

CLAIMS

What is claimed is:

1. A method for delivering a transgene to a target cell in a subject, the method comprising intracranially administering to the subject a recombinant adeno-associated virus (rAAV) comprising:

(i) an isolated nucleic acid comprising a transgene encoding one or more gene products of interest; and

(ii) an adeno-associated acid (AAV) capsid protein having the sequence set forth in SEQ ID NO: 1.

2. The method of claim 1, wherein the intracranial administration comprises intrahippocampal injection.

3. The method of claim 1 or 2, wherein the target cell is a central nervous system (CNS) cell.

4. The method of any one of claims 1 to 3, wherein the CNS cell is a neuron, oligodendrocyte, astrocyte, or microglial cell.

5. The method of any one of claims 1 to 4, wherein the subject is a mammal, optionally wherein the mammal is a human.

6. The method of any one of claims 1 to 5, wherein the subject is characterized by production of anti-AAV2 antibodies.

7. The method of claim 6, wherein after administration of the rAAV, the subject does not elicit a neutralizing immune response against the rAAV.

8. The method of any one of claims 1 to 7, wherein the isolated nucleic acid comprises AAV inverted terminal repeats (ITRs) flanking the transgene.

9. The method of any one of claims 1 to 8, wherein the nucleic acid sequence encoding the one or more gene products is operably linked to a promoter.

10. The method of any one of claims 1 to 9, wherein the one or more gene products comprise a protein or an inhibitory nucleic acid.

11. A method for delivering a transgene to a target cell in a subject, the method comprising intravenously administering to the subject a recombinant adeno-associated virus (rAAV) comprising:

(i) an isolated nucleic acid comprising a transgene encoding one or more gene products of interest; and

(ii) an adeno-associated acid (AAV) capsid protein having the sequence set forth in SEQ ID NO: 1,

wherein the administration results in the rAAV crossing the blood brain barrier (BBB) of the subject.

12. The method of claim 11, wherein the target cell is a central nervous system (CNS) cell.

13. The method of claim 12, wherein the CNS cell is a neuron, oligodendrocyte, astrocyte, or microglial cell.

14. The method of any one of claims 11 to 13, wherein the administration results in decreased transduction of liver cells relative to administration of an rAAV having an AAV2 capsid protein

15. The method of any one of claims 11 to 14, wherein the subject is a mammal, optionally wherein the mammal is a human.

16. The method of any one of claims 11 to 15, wherein the subject is characterized by production of anti-AAV2 antibodies.

17. The method of claim 16, wherein after administration of the rAAV, the subject does not elicit a neutralizing immune response against the rAAV.

18. The method of any one of claims 11 to 17, wherein the isolated nucleic acid comprises AAV inverted terminal repeats (ITRs) flanking the transgene.

19. The method of any one of claims 11 to 18, wherein the nucleic acid sequence encoding the one or more gene products is operably linked to a promoter.

20. The method of any one of claims 11 to 19, wherein the one or more gene products comprise a protein or an inhibitory nucleic acid.

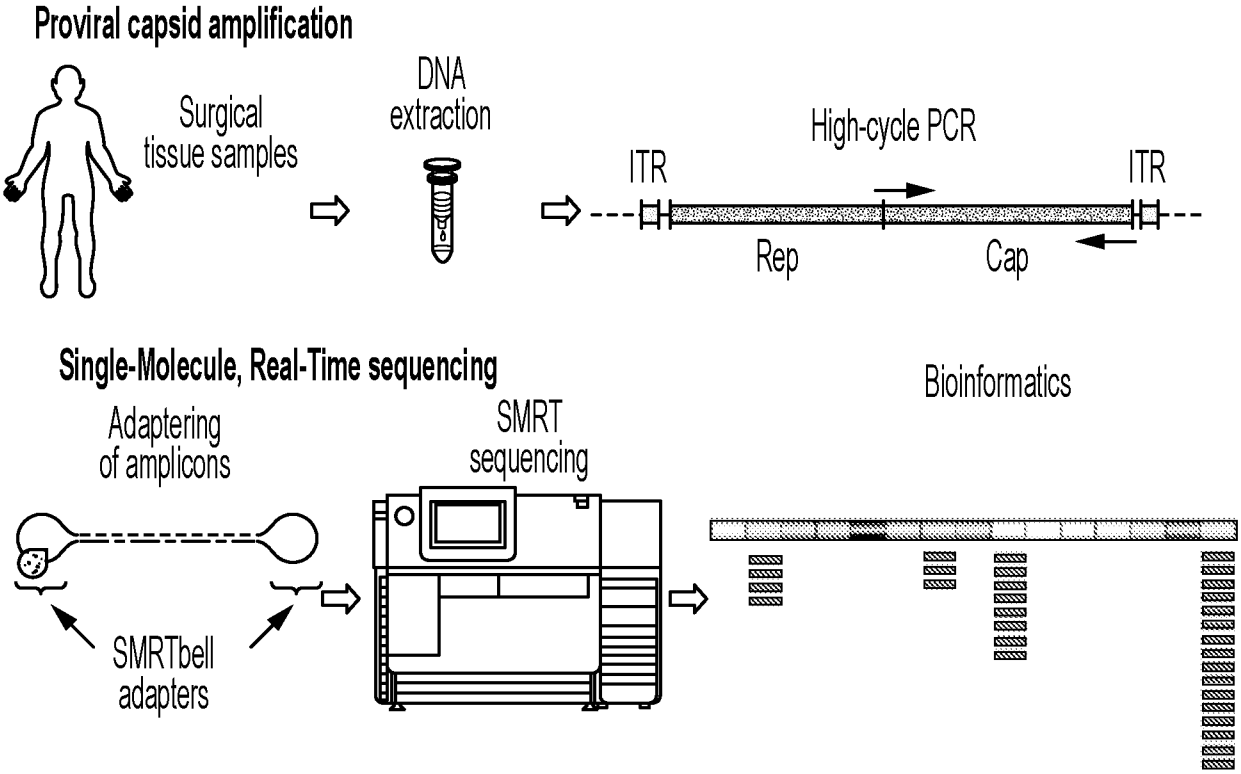

Proviral capsid amplification
Surgical tissue samples
DNA extraction
High-cycle PCR
ITR
ITR
Rep
Cap
Single-Molecule, Real-Time sequencing
Adaptering of amplicons
SMRT sequencing
Bioinformatics
SMRTbell adapters

FIG. 1A

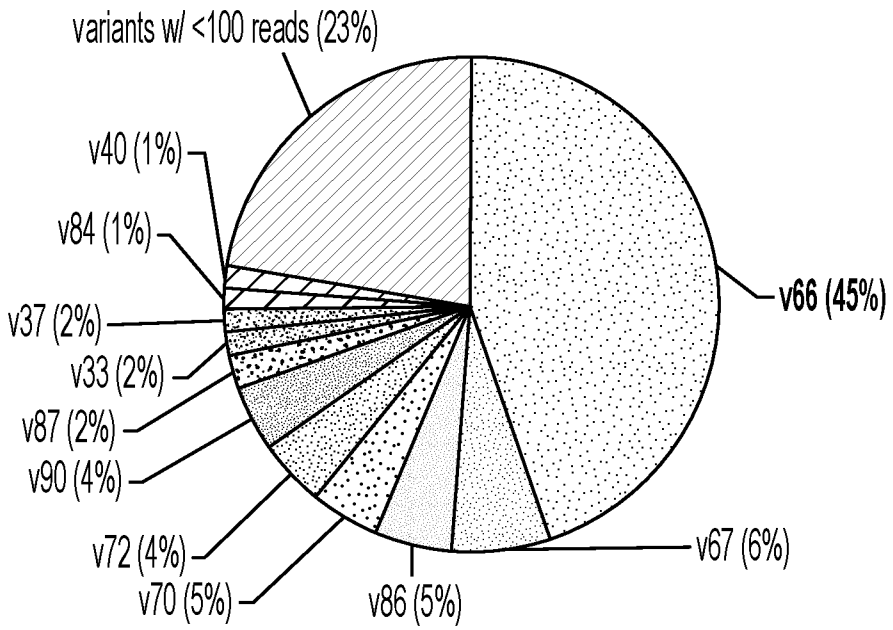

variants w/ <100 reads (23%)
v40 (1%)
v84 (1%)
v37 (2%)
v33 (2%)
v87 (2%)
v90 (4%)
v72 (4%)
v70 (5%)
v86 (5%)
v67 (6%)
v66 (45%)

FIG. 1B

Unique AAVv66 a.a.	Capsid domain	Variable region
K39Q	VP1	-
V151A	VP2	-
R447K	VP3	VR-IV
T450A		
Q457M		
S492A		VR-V
E499D		
F533Y		VR-VI
G546D		VR-VII
E548G		
R585S		VR-VIII
R588T		
A593T		

FIG. 1C

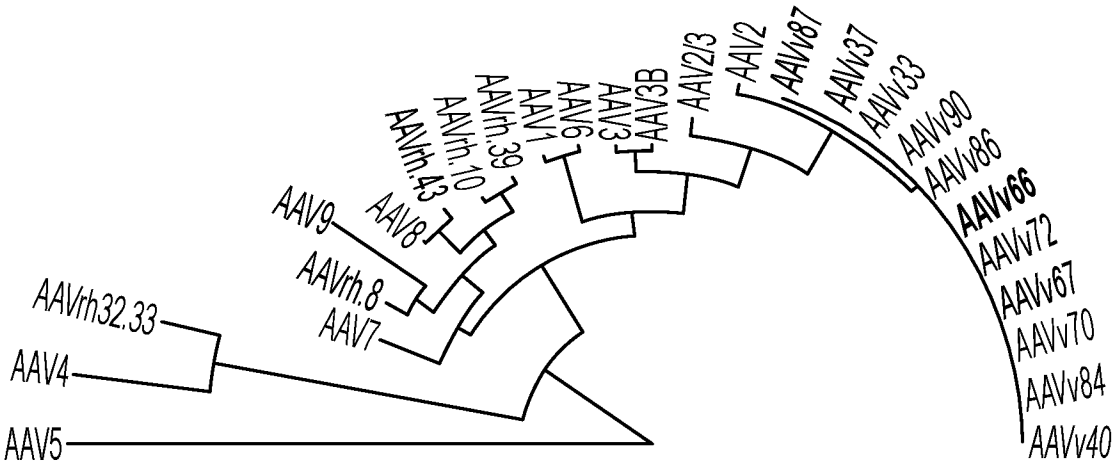


FIG. 1D

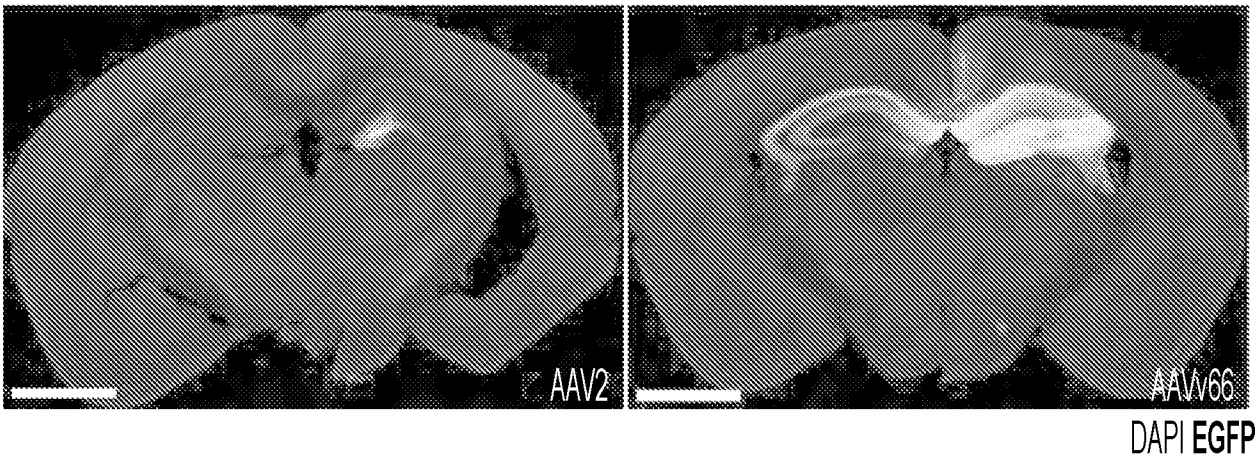


FIG. 2A

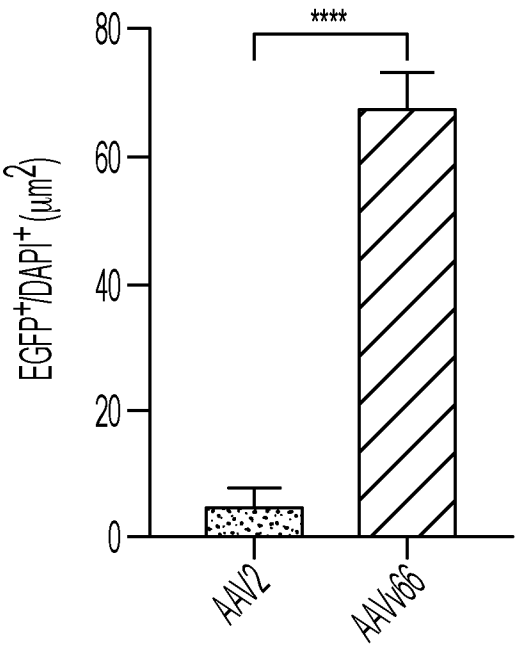


FIG. 2B

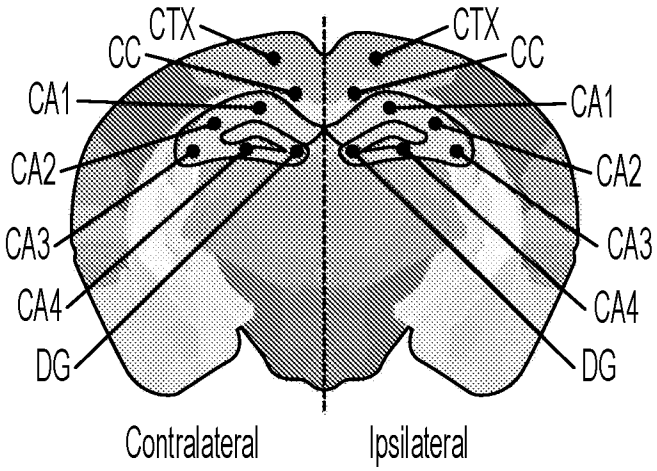

CTX
CC
CA1
CA2
CA3
CA4
DG
CTX
CC
CA1
CA2
CA3
CA4
DG
Contralateral
Ipsilateral

FIG. 2C

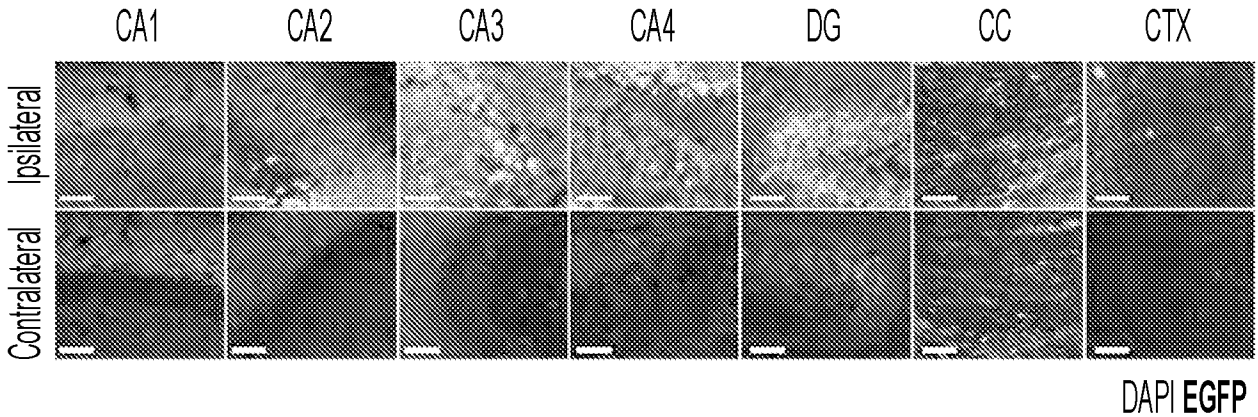

CA1
CA2
CA3
CA4
DG
CC
CTX
Ipsilateral
Contralateral
DAPI EGFP

FIG. 2D

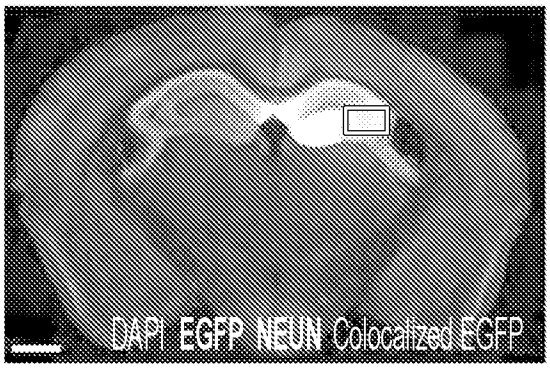


FIG. 3A

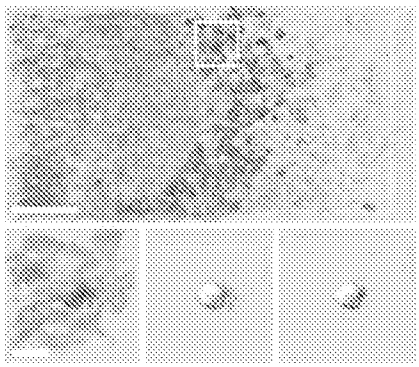


FIG. 3B

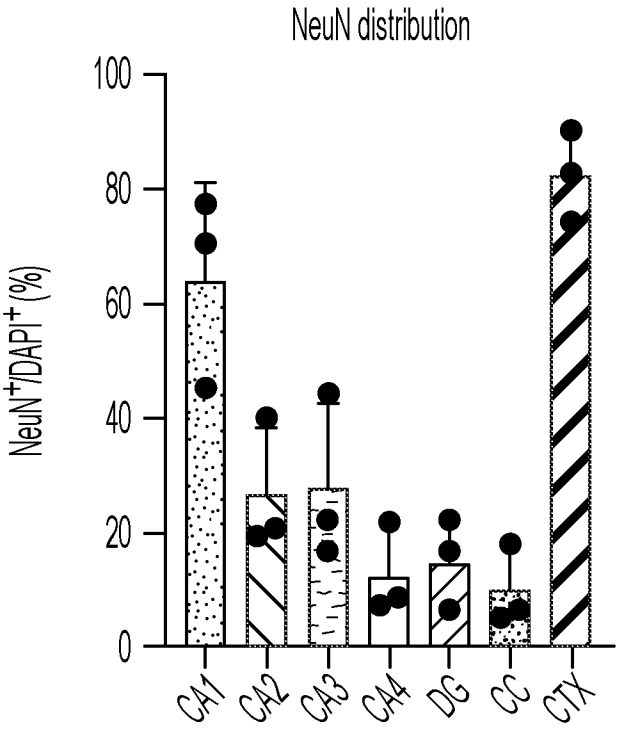


FIG. 3C

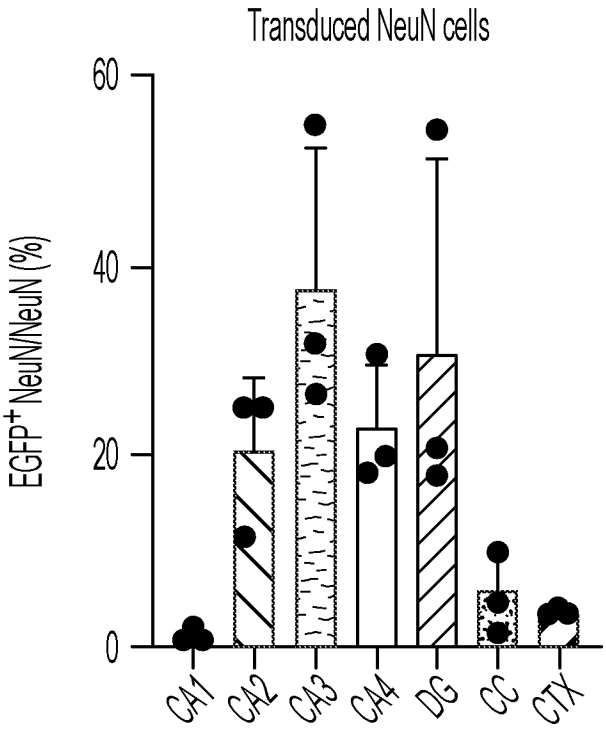


FIG. 3D

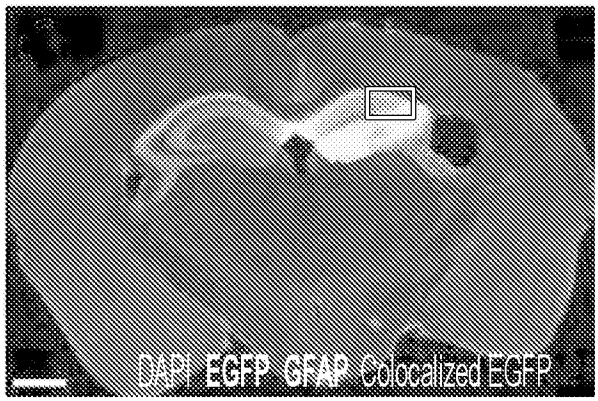


FIG. 3E

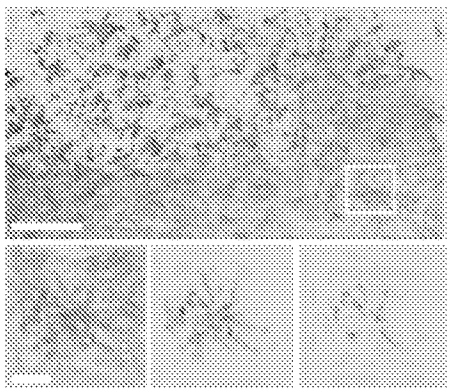


FIG. 3F

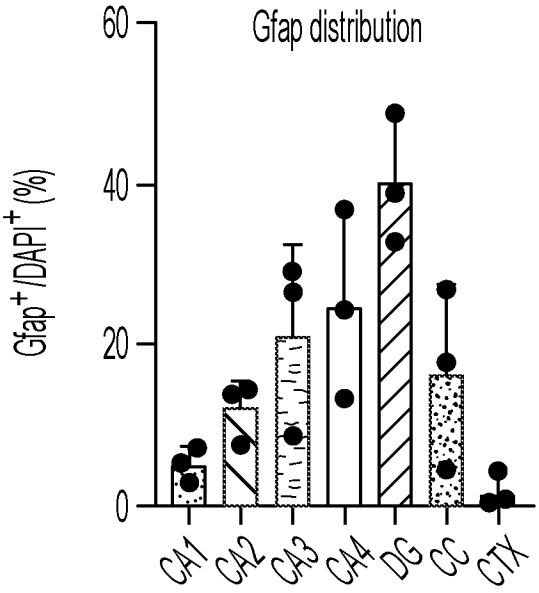


FIG. 3G

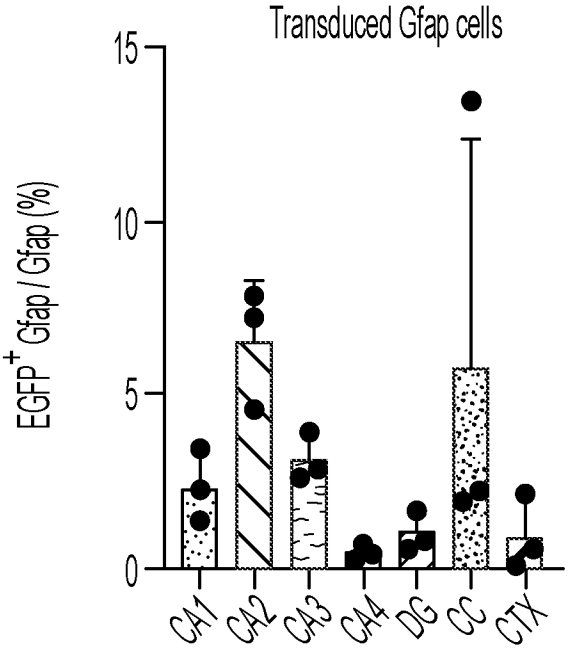


FIG. 3H

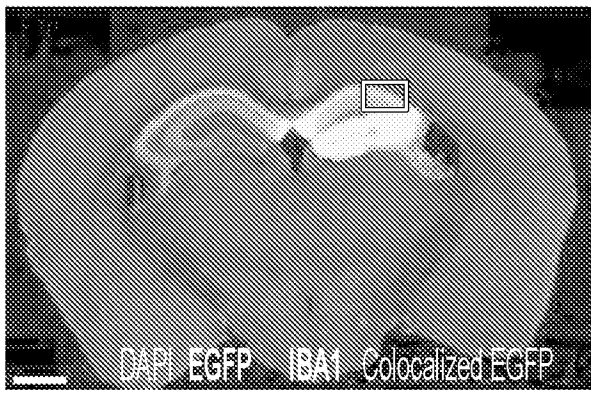


FIG. 3I

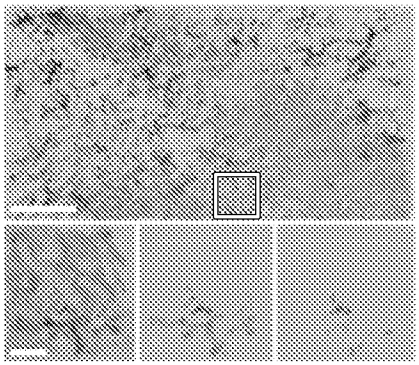


FIG. 3J

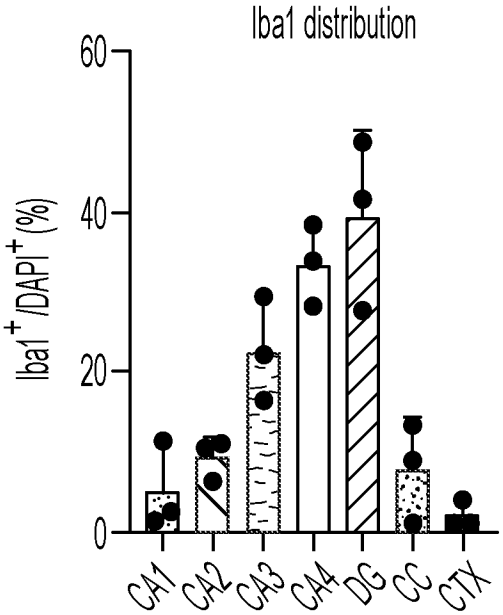


FIG. 3K

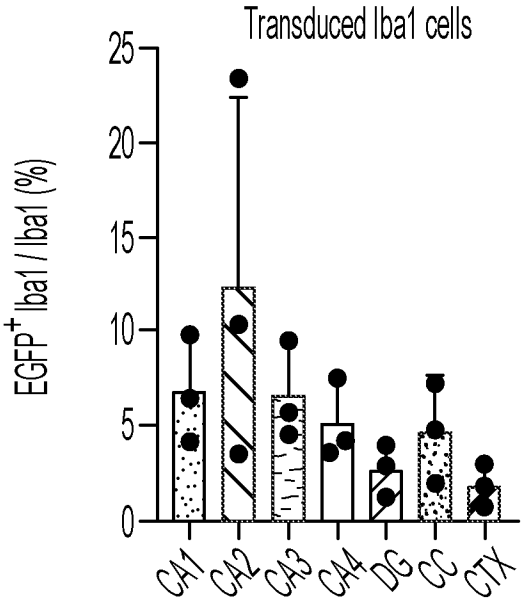


FIG. 3L

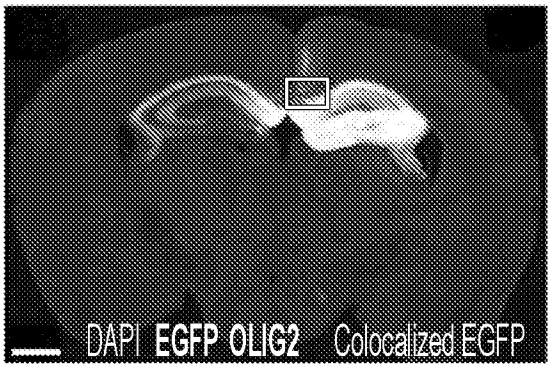


FIG. 3M

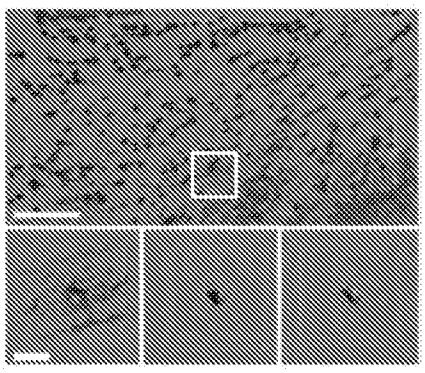


FIG. 3N

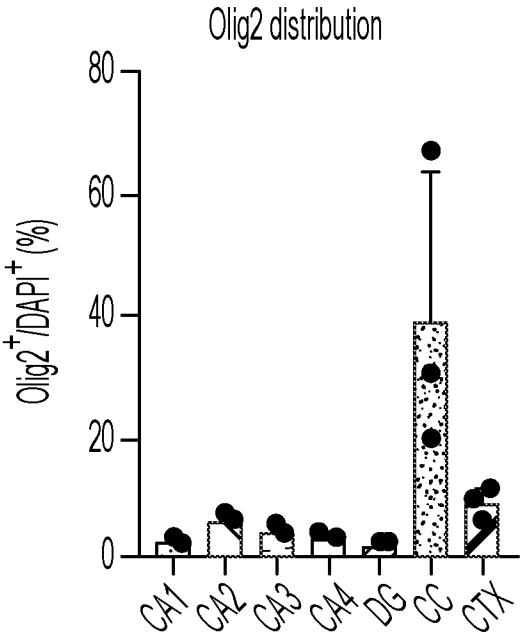


FIG. 3O

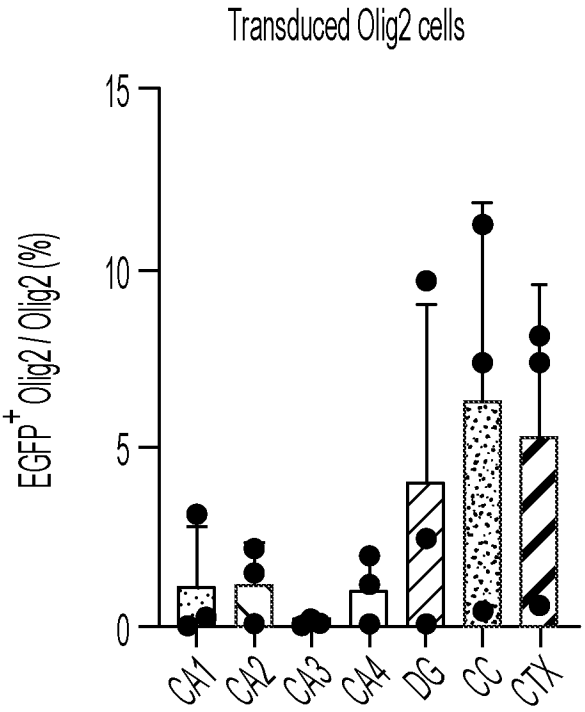

Transduced Olig2 cells
EGFP+ Olig2 / Olig2 (%)
15
10
5
0
CA1
CA2
CA3
CA4
DG
CC
CTX

FIG. 3P

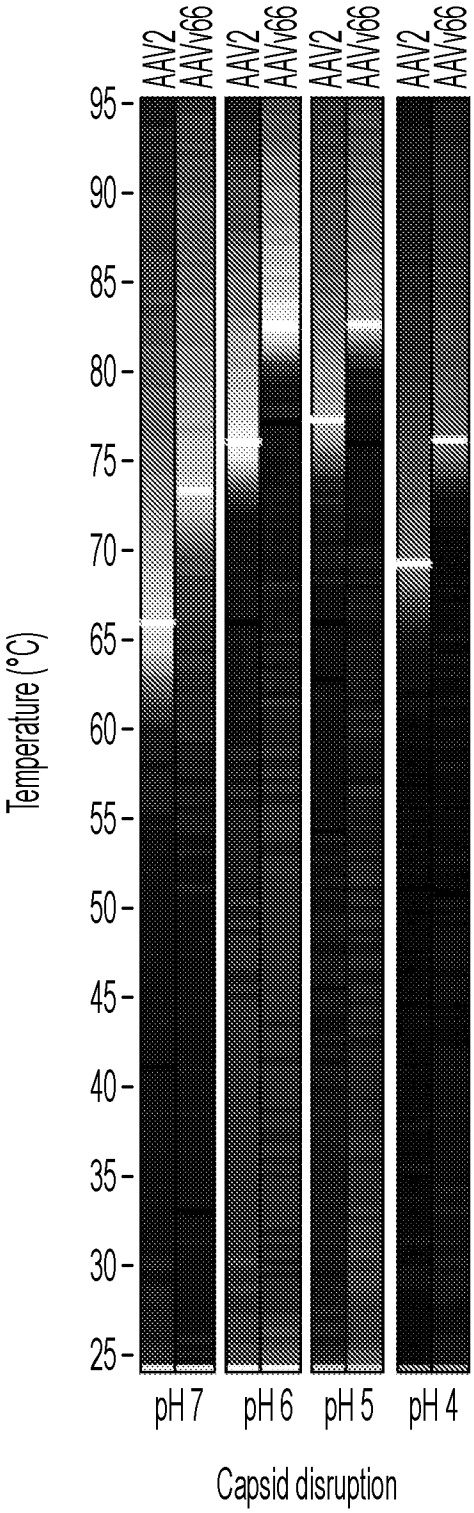

AAV2
AAVv66
AAV2
AAVv66
AAV2
AAVv66
AAV2
AAVv66
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
Temperature (°C)
pH 7
pH 6
pH 5
pH 4
Capsid disruption

FIG. 4A

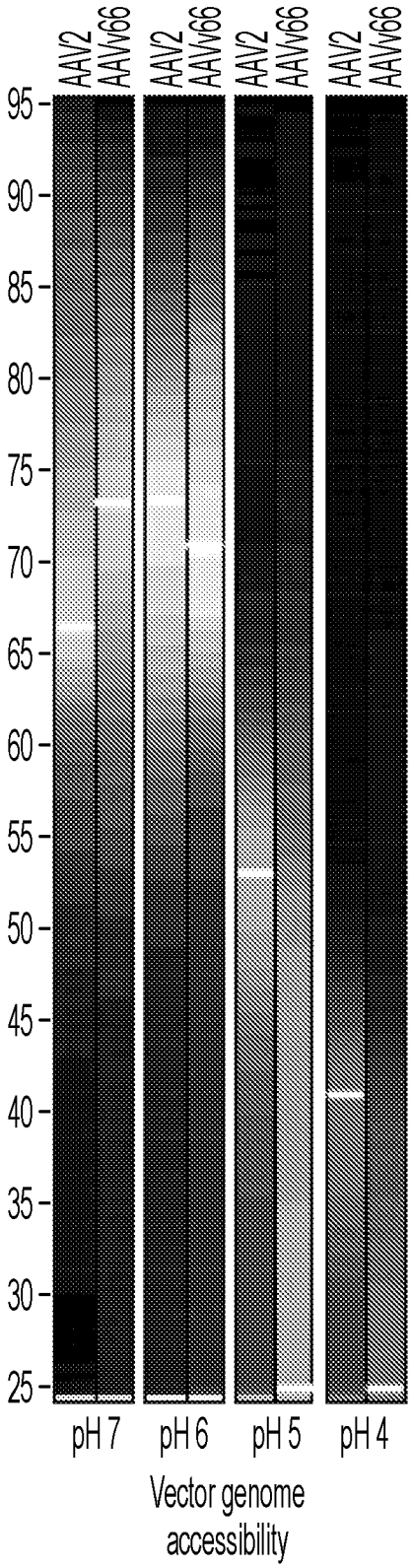

AAV2
AAVw66
AAV2
AAVw66
AAV2
AAVw66
AAV2
AAVw66
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
pH 7
pH 6
pH 5
pH 4
Vector genome accessibility

FIG. 4B

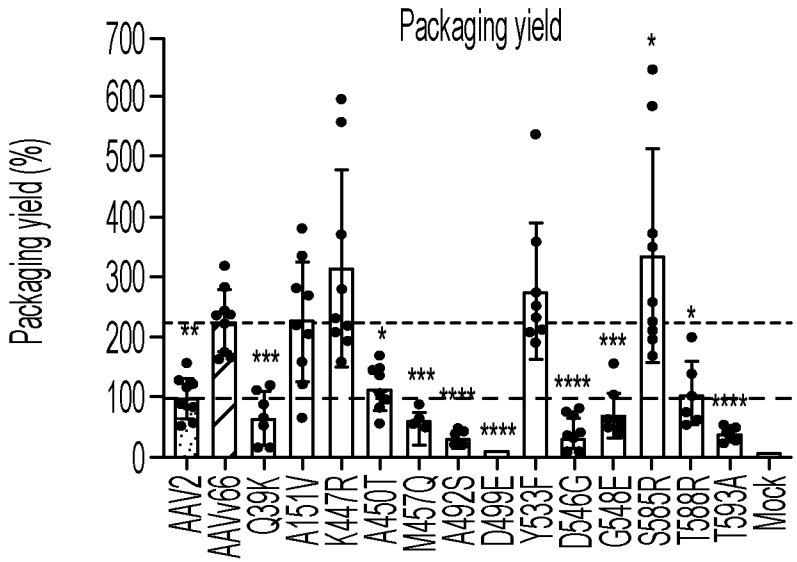


FIG. 4C

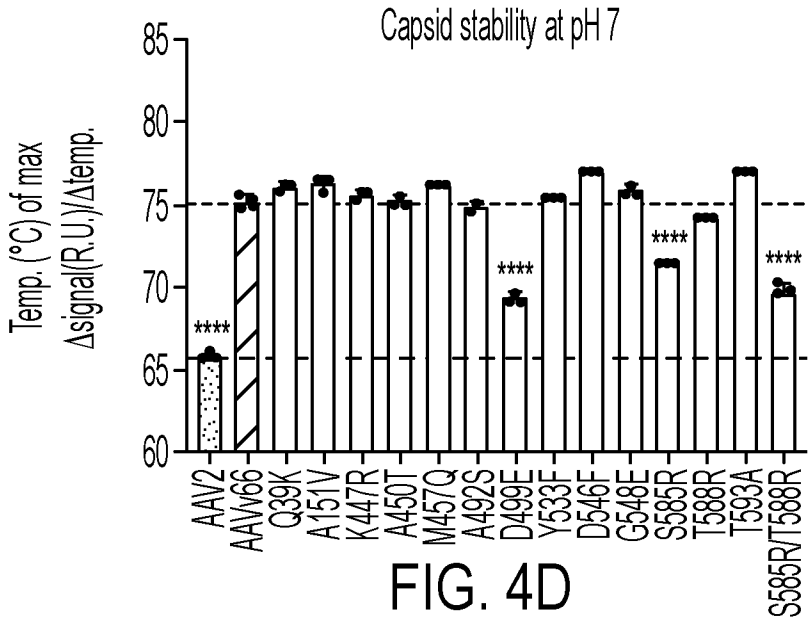


FIG. 4D

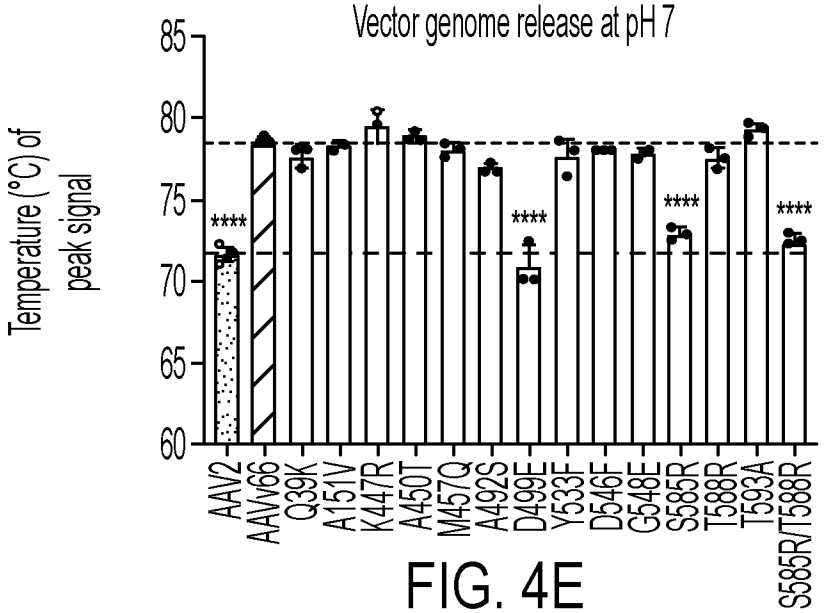


FIG. 4E

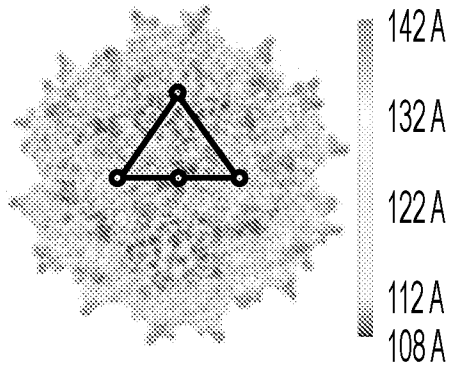


FIG. 5A

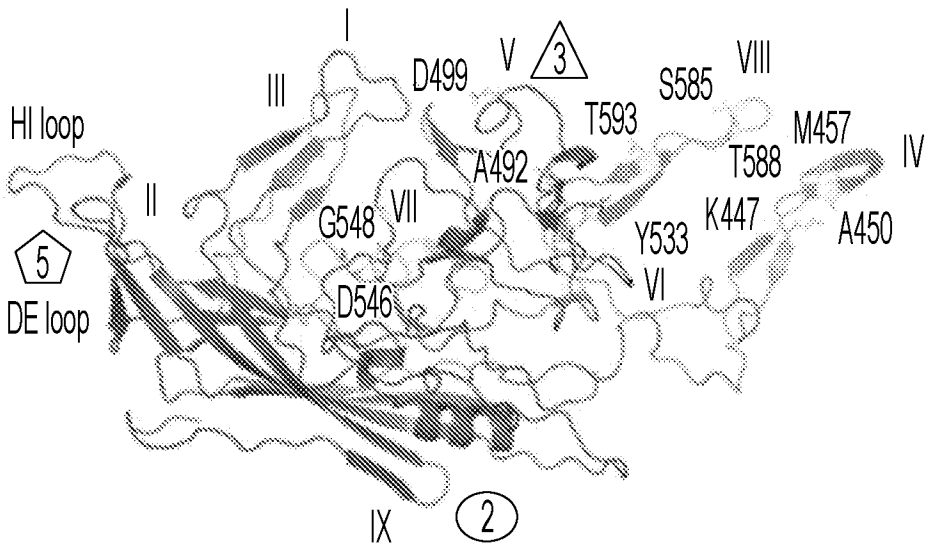


FIG. 5B

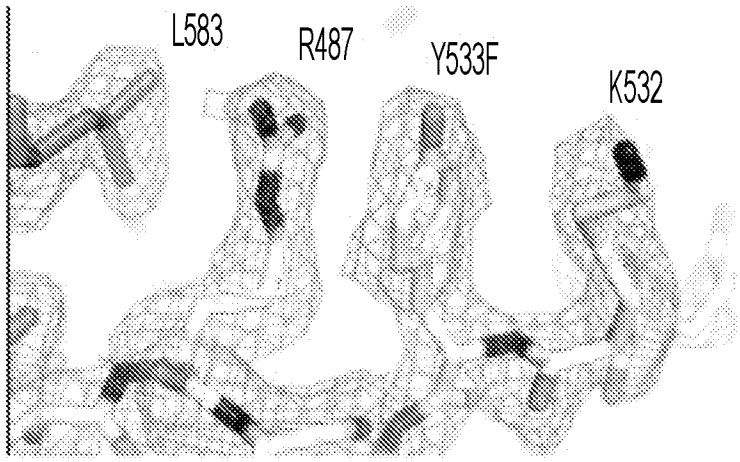


FIG. 5C

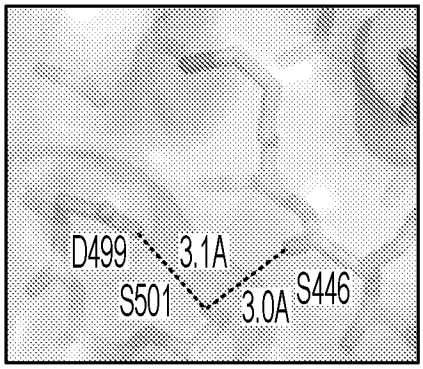

D499
3.1A
S501
3.0A
S446

FIG. 5D

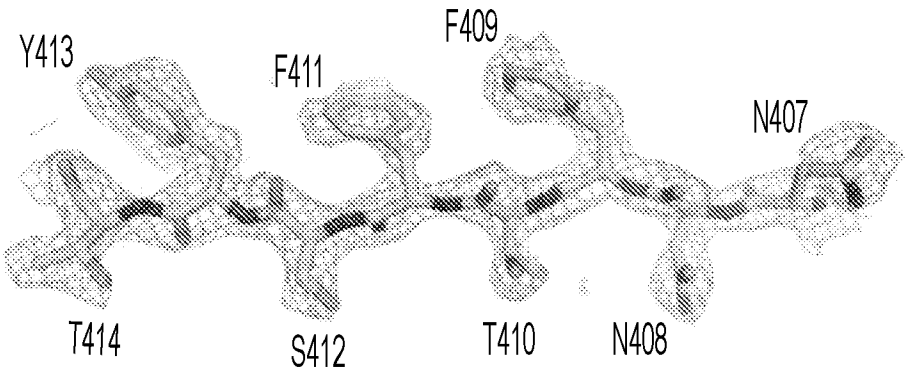

Y413
F411
F409
N407
T414
S412
T410
N408

FIG. 5E

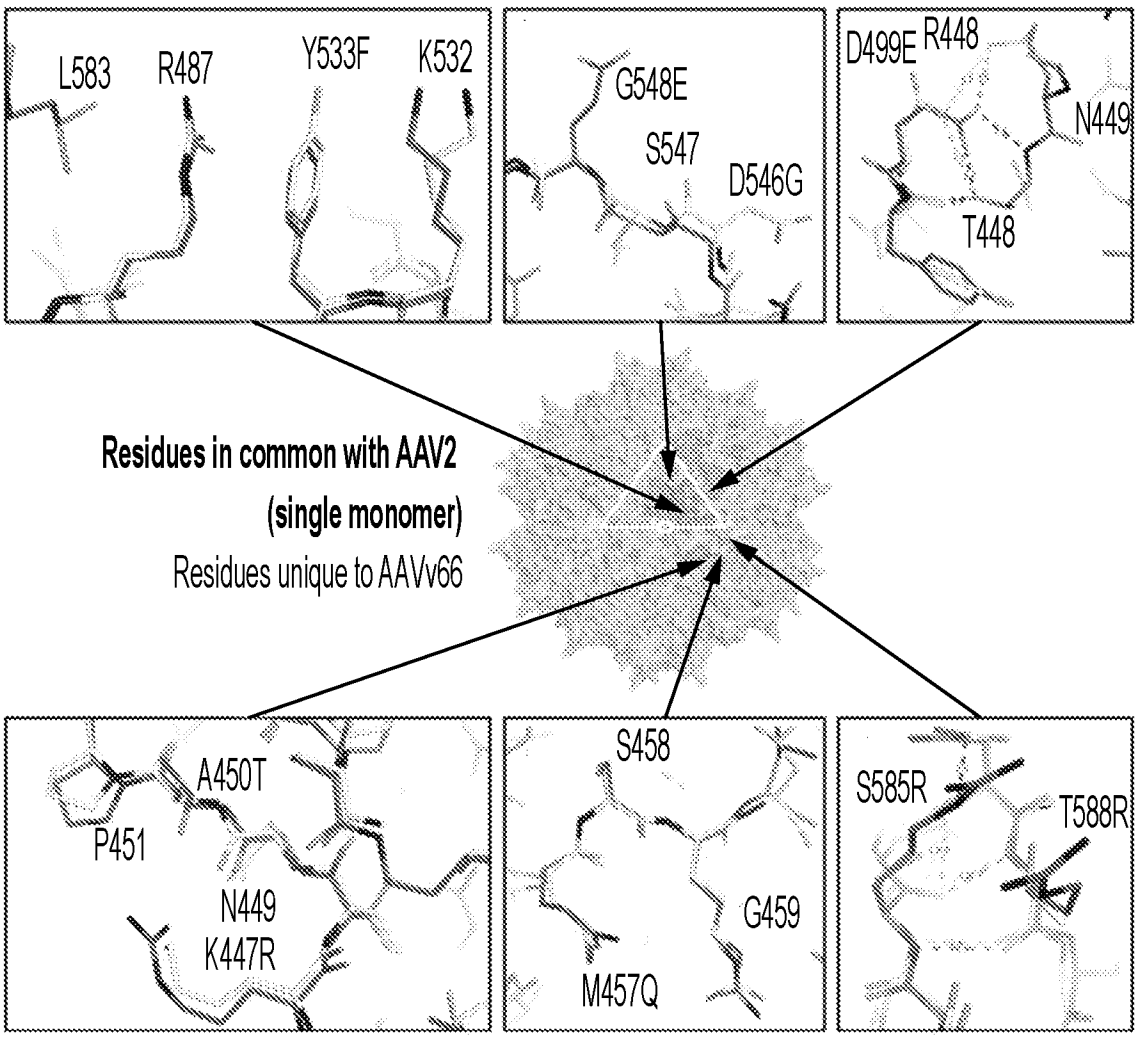

L583
R487
Y533F
K532
G548E
S547
D546G
D499E
R448
N449
T448
Residues in common with AAV2
(single monomer)
Residues unique to AAVv66
A450T
P451
N449
K447R
S458
G459
M457Q
S585R
T588R

FIG. 6

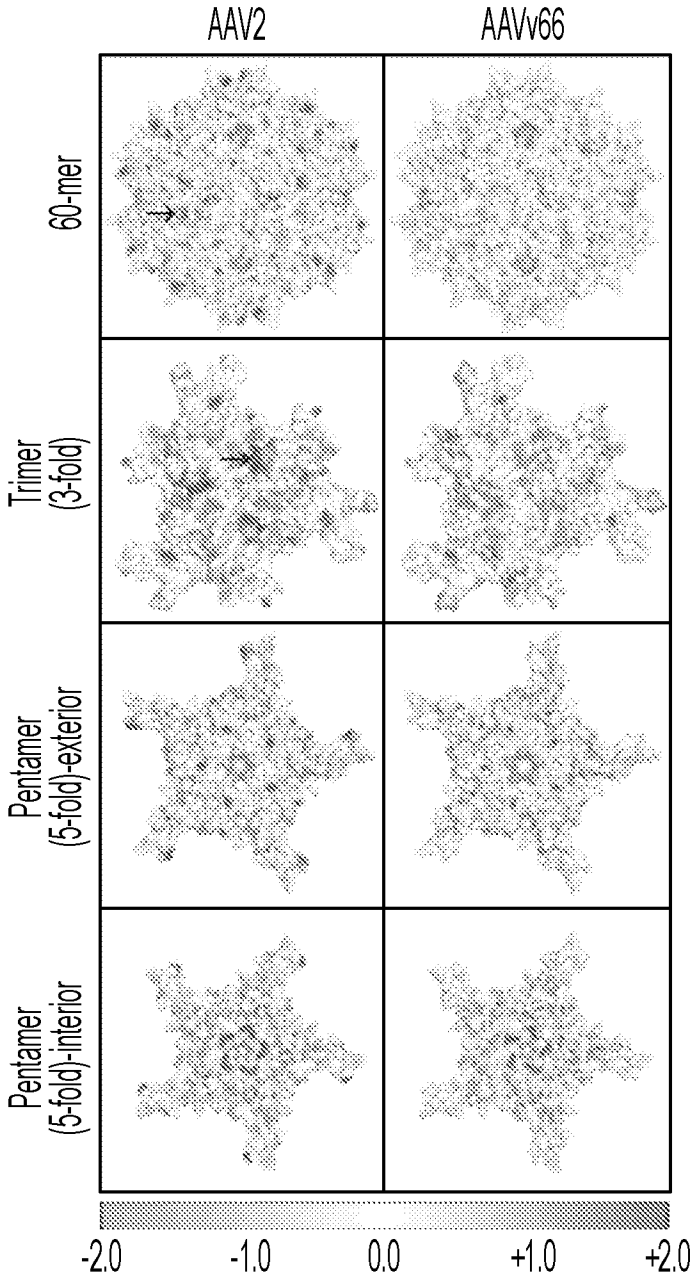

AAV2
AAVv66
60-mer
Trimer (3-fold)
Pentamer (5-fold)-exterior
Pentamer (5-fold)-interior
-2.0
-1.0
0.0
+1.0
+2.0

FIG. 7A

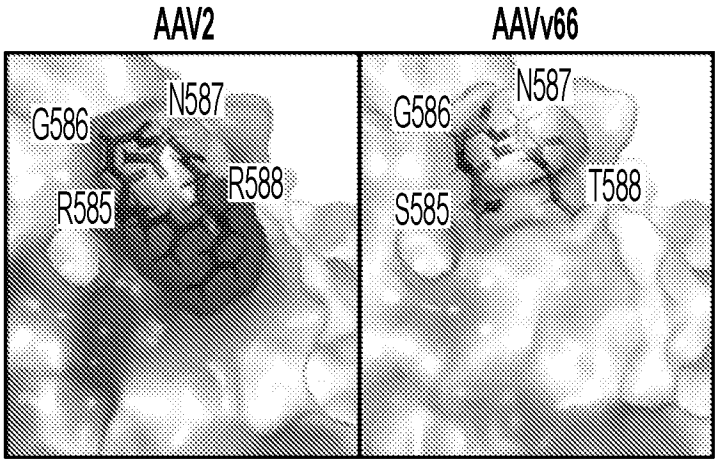

AAV2
AAVv66
N587
G586
R588
R585
N587
G586
T588
S585

FIG. 7B

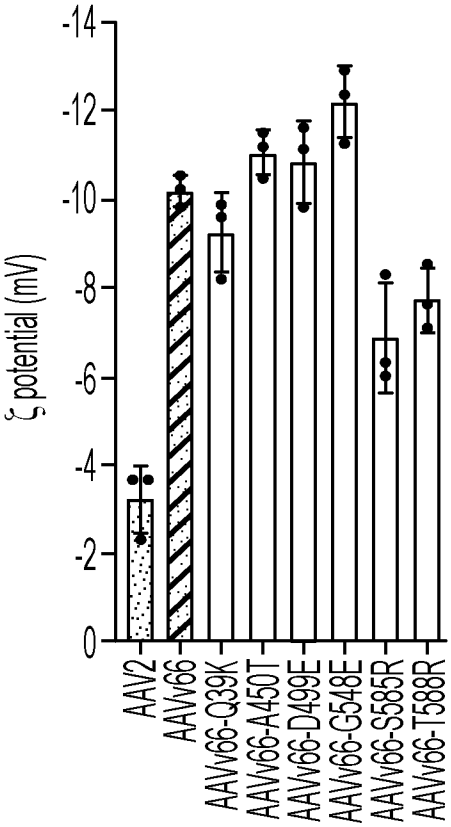

ζ potential (mV)
-14
-12
-10
-8
-6
-4
-2
0
AAV2
AAVv66
AAVv66-Q39K
AAVv66-A450T
AAVv66-D499E
AAVv66-G548E
AAVv66-S585R
AAVv66-T588R

FIG. 7C

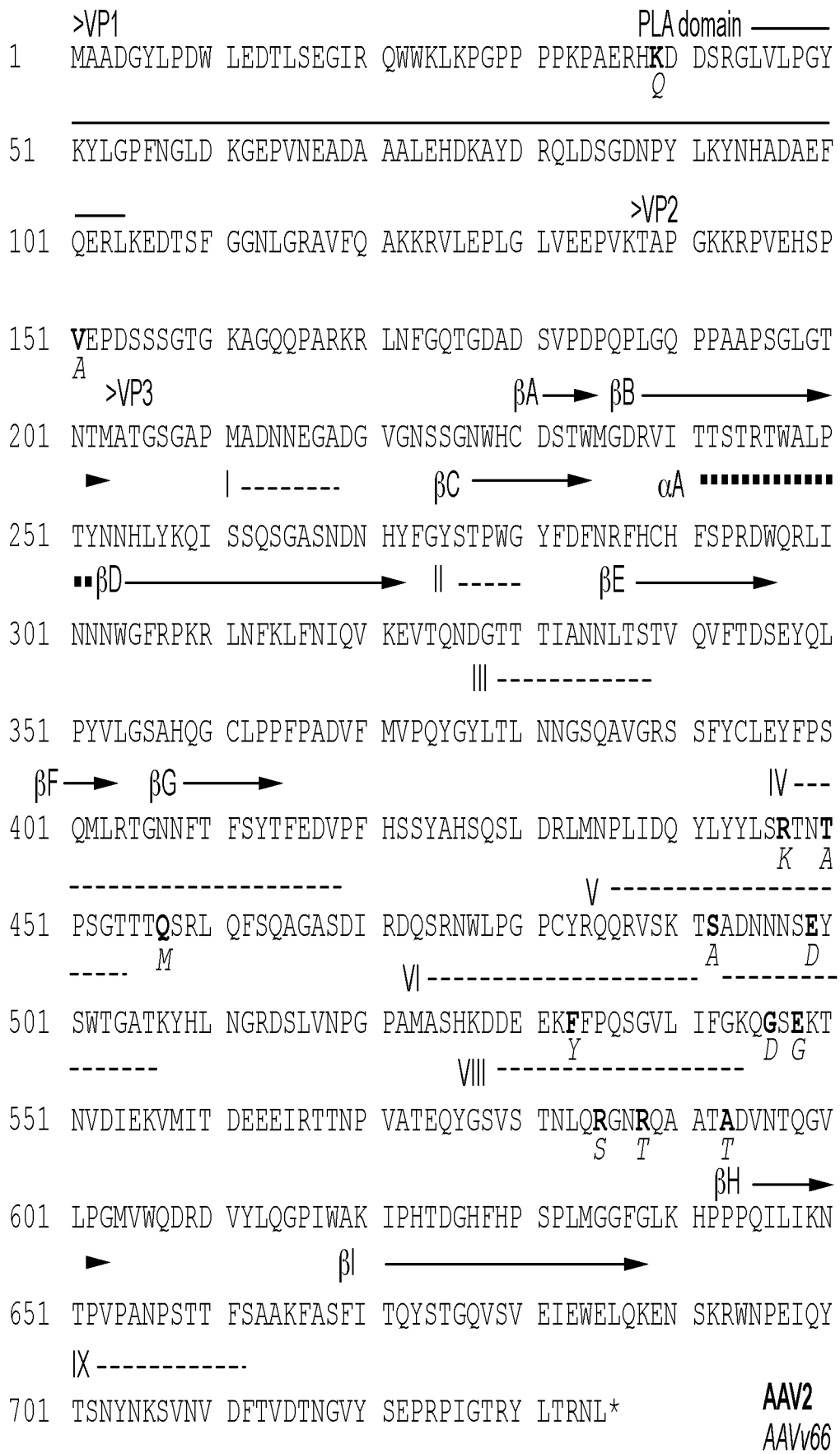

>VP1
PLA domain
1 MAADGYLPDW LEDTLSEGIR QWWKLKPGPP PPKPAERHKD DSRGLVLPGY
Q
51 KYLGPFNGLD KGEPVNEADA AALEHDKAYD RQLDSGDNPY LKYNHADAEF
>VP2
101 QERLKEDTSF GGNLGRAVFQ AKKRVLEPLG LVEEPVKTAP GKKRPVEHSP
151 VEPDSSSGTG KAGQQPARKR LNFGQTGDAD SVPDPQPLGQ PPAAPSGLGT
A
>VP3
βA βB
201 NTMATGSGAP MADNNEGADG VGNSSGNWHC DSTWMGDRVI TTSTRTWALP
I
βC αA
251 TYNNHLYKQI SSQSGASNDN HYFGYSTPWG YFDFNRFHCH FSPRDWQRLI
βD II
βE
301 NNNWGFRPKR LNFKLFNIQV KEVTQNDGTT TIANNLTSTV QVFTDSEYQL
III
351 PYVLGSAHQG CLPPFPADVF MVPQYGYLTL NNGSQAVGRS SFYCLEYFPS
βF βG IV
401 QMLRTGNNFT FSYTFEDVPF HSSYAHSQSL DRLMNPLIDQ YLYYLSRTNT
K A
V
451 PSGTTTQSRL QFSQAGASDI RDQSRNWLPG PCYRQQRVSK TSADNNNSEY
M A D
VI VII
501 SWTGATKYHL NGRDSLVNPG PAMASHKDDE EKFFPQSGVL IFGKQGSEKT
Y D G
VIII
551 NVDIEKVMIT DEEEIRTTNP VATEQYGSVS TNLQRGNRQA ATADVNTQGV
S T T
βH
601 LPGMVWQDRD VYLQGPIWAK IPHTDGHFHP SPLMGGFGLK HPPPQILIKN
βI
651 TPVPANPSTT FSAAKFASFI TQYSTGQVSV EIEWELQKEN SKRWNPEIQY
IX
701 TSNYNKSVNV DFTVDTNGVY SEPRPIGTRY LTRNL*
AAV2
AAVv66

FIG. 8

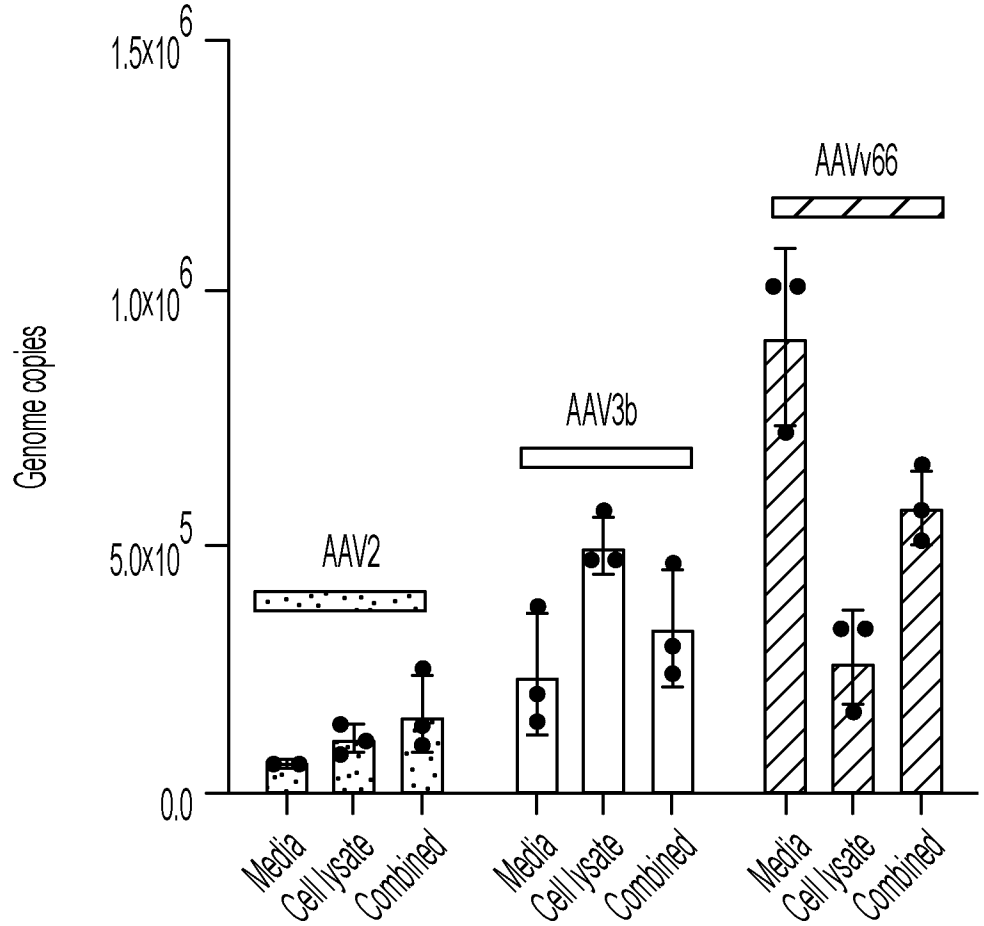

1.5x10^6
1.0x10^6
5.0x10^5
0.0
Genome copies
AAV2
AAV3b
AAVv66
Media
Cell lysate
Combined
Media
Cell lysate
Combined
Media
Cell lysate
Combined

FIG. 9

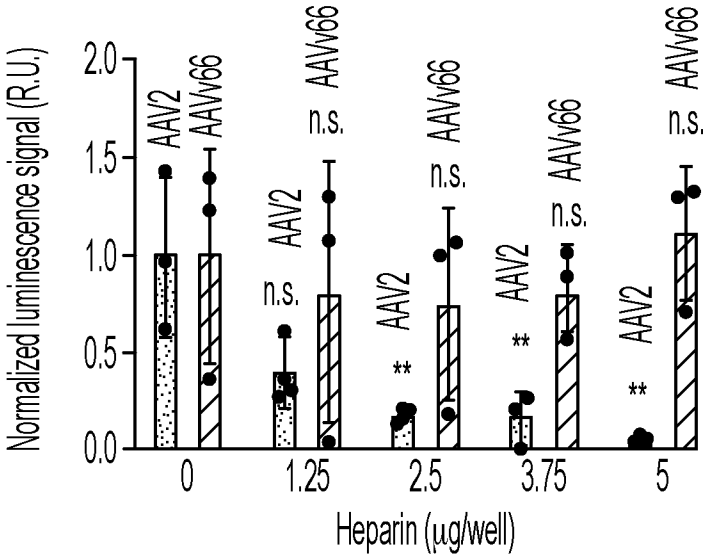

Normalized luminescence signal (R.U.)
2.0
1.5
1.0
0.5
0.0
AAV2
AAVv66
n.s.
**
0
1.25
2.5
3.75
5
Heparin (µg/well)

FIG. 10

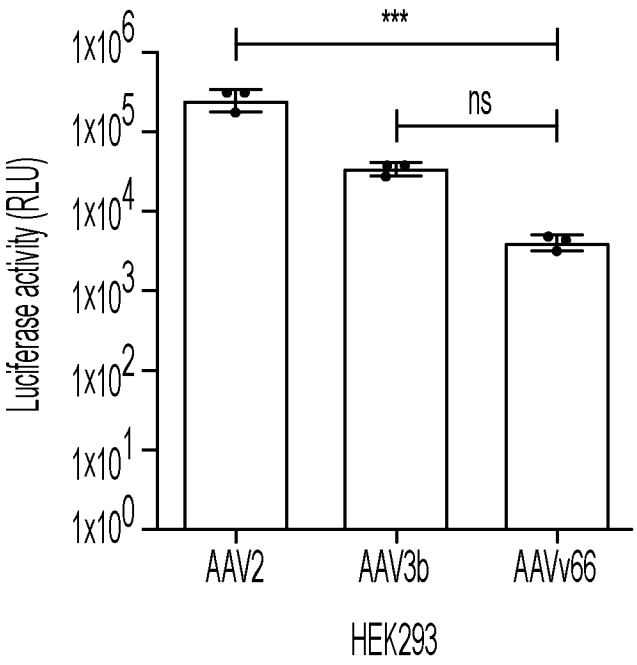


FIG. 11

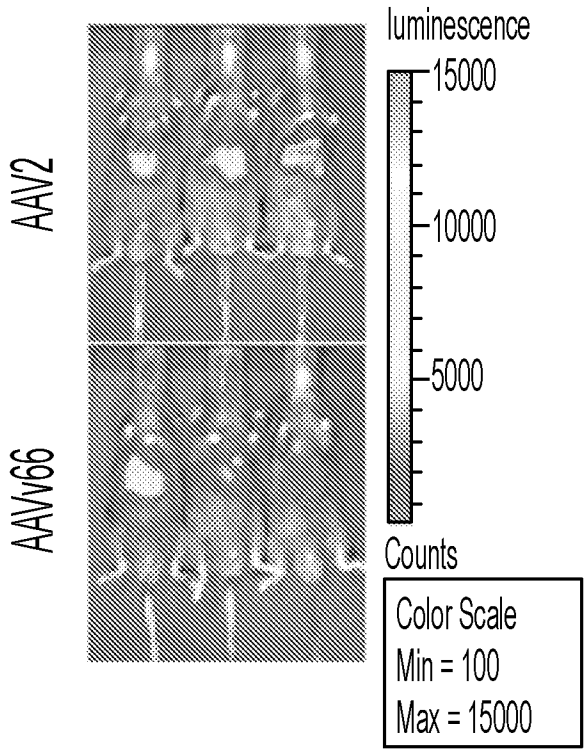


FIG. 12A

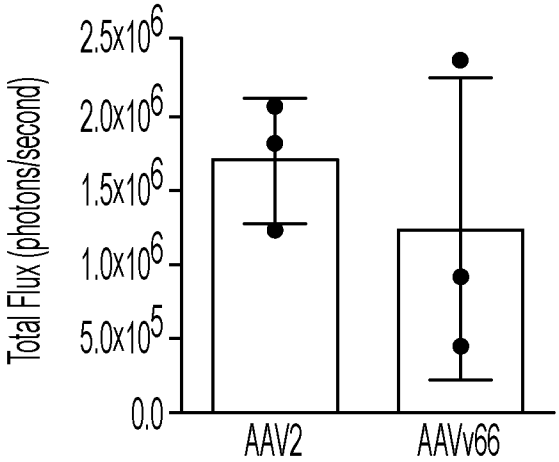


FIG. 12B

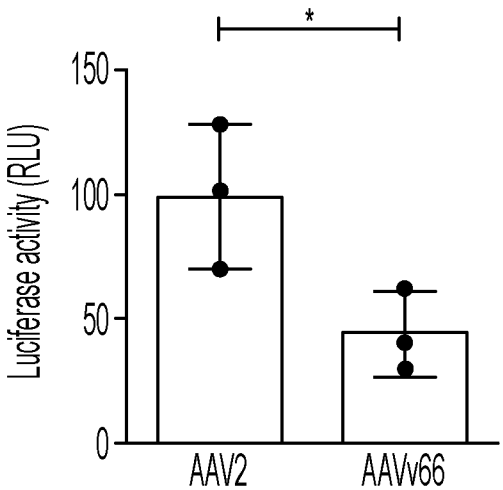


FIG. 12C

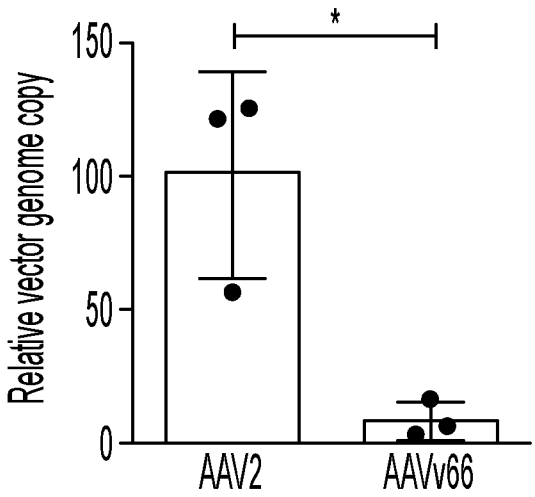


FIG. 12D

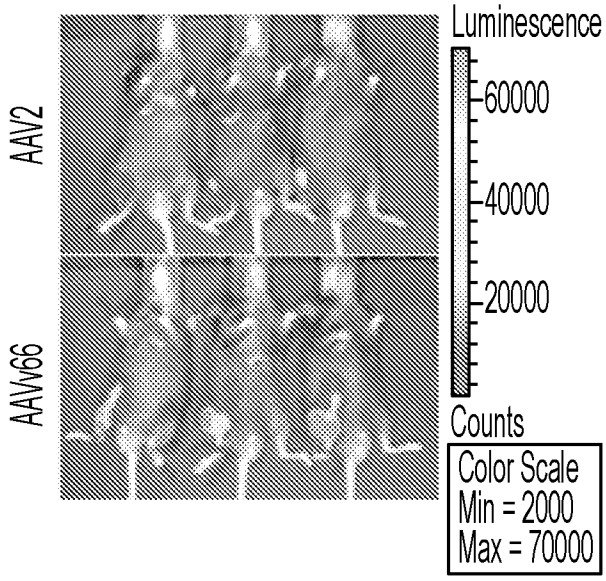


FIG. 13A

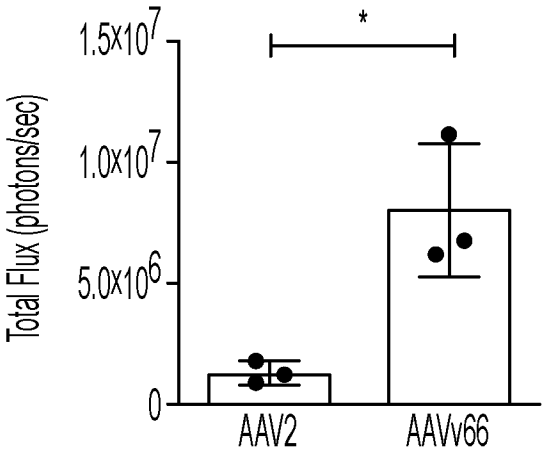


FIG. 13B

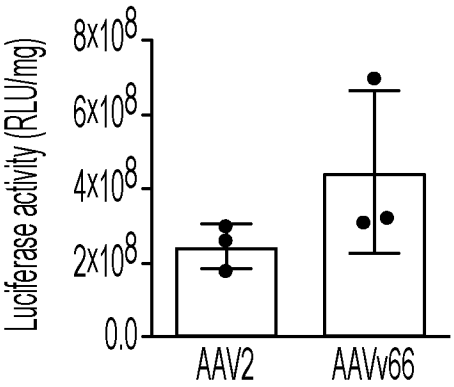


FIG. 13C

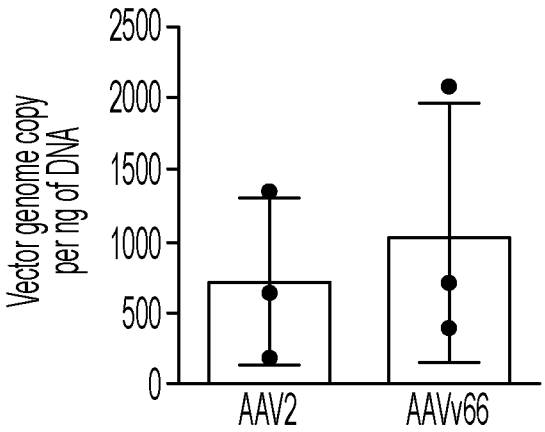
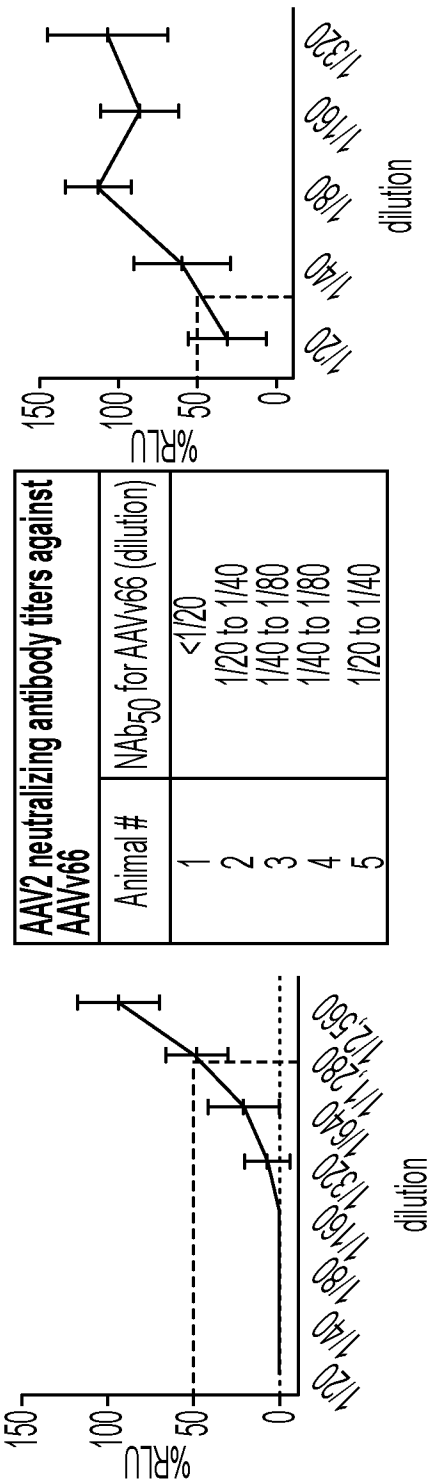


FIG. 13D

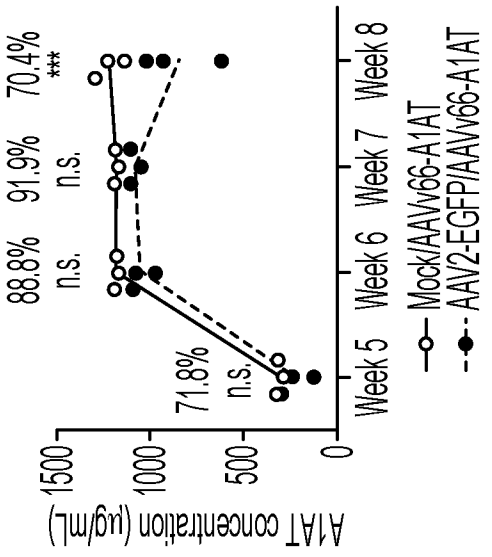


AAV2 neutralizing antibody titers against AAV2	
Animal #	NAb_{50} for AAV2 (dilution)
1	1/1,280 to 1/2,560
2	1/1,280 to 1/2,560
3	1/1,280 to 1/2,560
4	1/1,280 to 1/2,560
5	1/2,560 to 1/5,120

FIG. 14A

AAV2 neutralizing antibody titers against AAVv66	
Animal #	NAb_{50} for AAVv66 (dilution)
1	<1/20
2	1/20 to 1/40
3	1/40 to 1/80
4	1/40 to 1/80
5	1/20 to 1/40

FIG. 14B



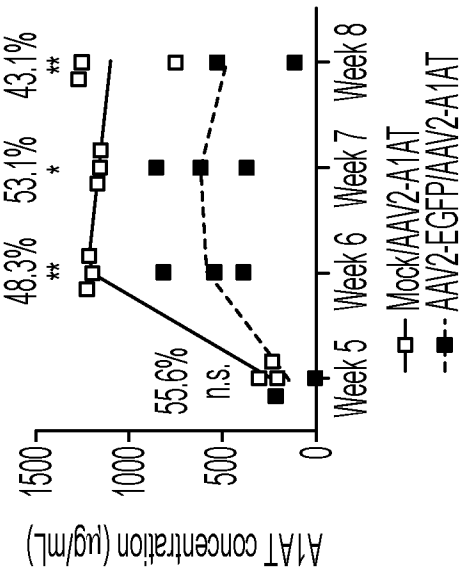


FIG. 14C

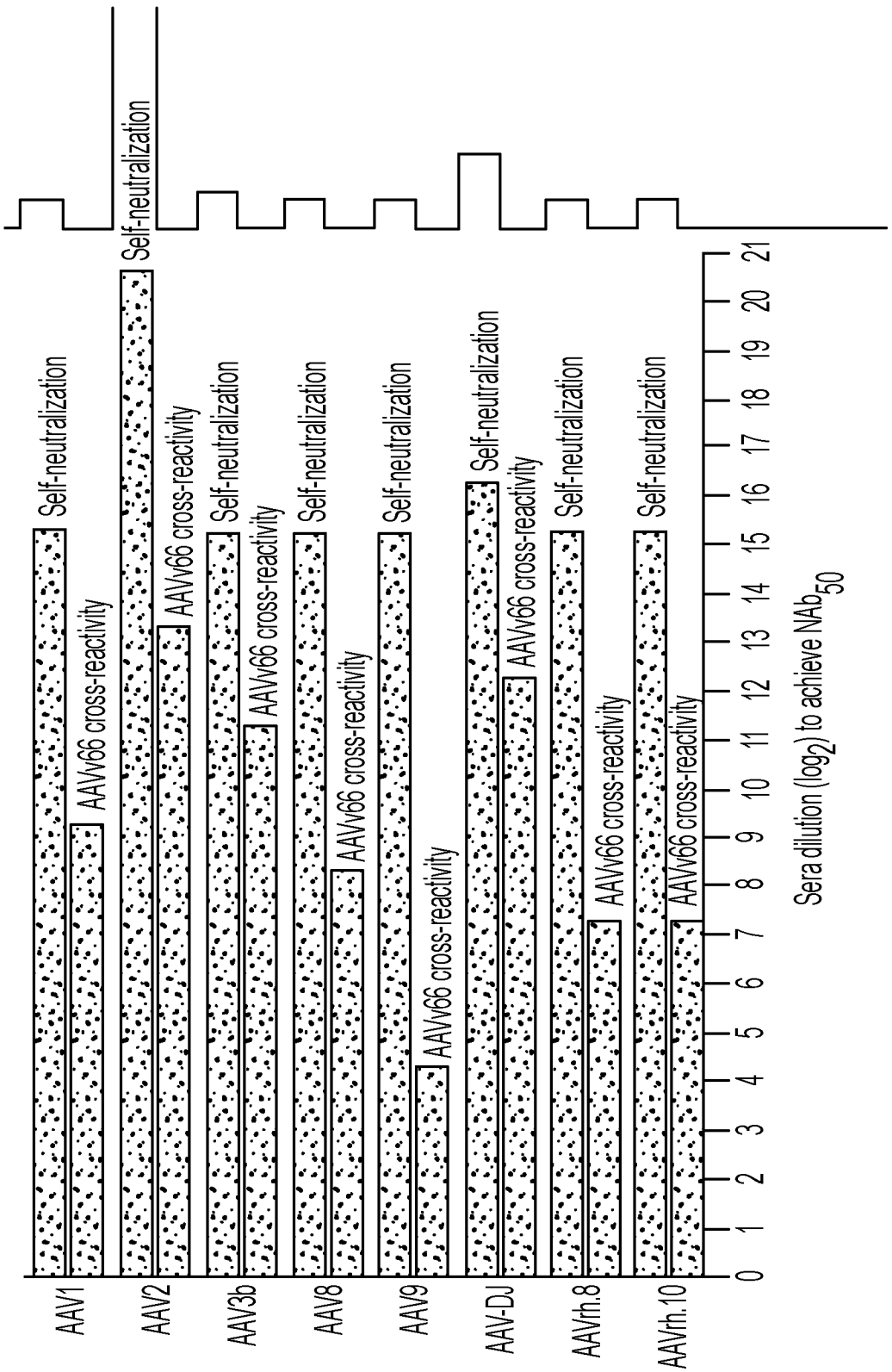

Self-neutralization
AAVv66 cross-reactivity
Self-neutralization
AAVv66 cross-reactivity
Self-neutralization
AAVv66 cross-reactivity
Self-neutralization
AAVv66 cross-reactivity
Self-neutralization
AAVv66 cross-reactivity
Self-neutralization
AAVv66 cross-reactivity
Self-neutralization
AAVv66 cross-reactivity
Self-neutralization
AAVv66 cross-reactivity
AAV1
AAV2
AAV3b
AAV8
AAV9
AAV-DJ
AAVrh.8
AAVrh.10
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Sera dilution (log_2) to achieve NAb_{50}

FIG. 14D

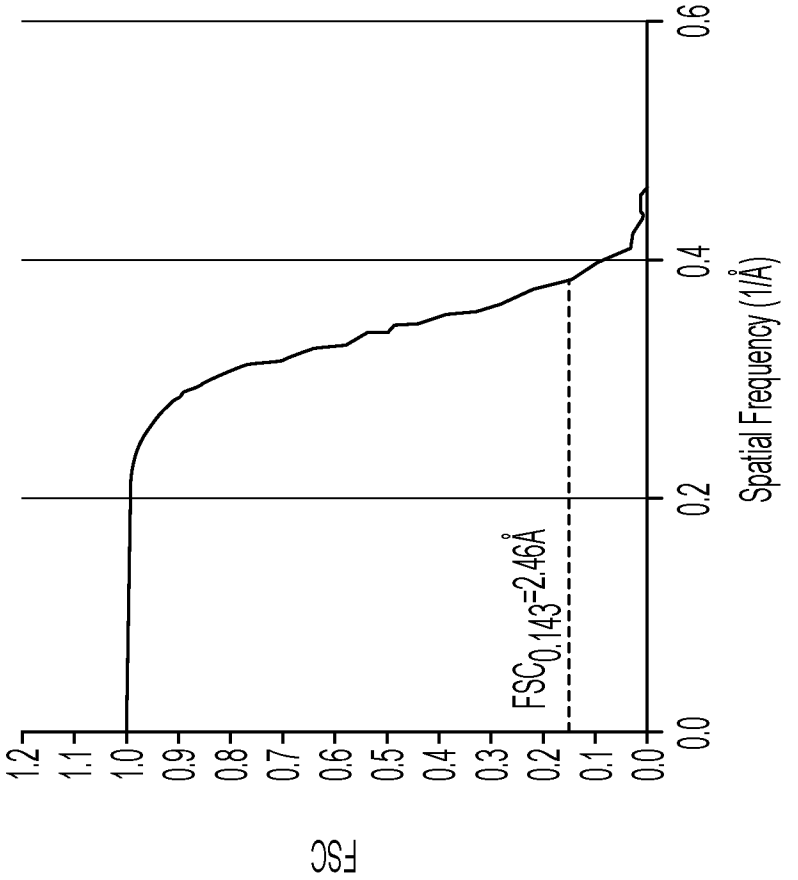


FIG. 15B

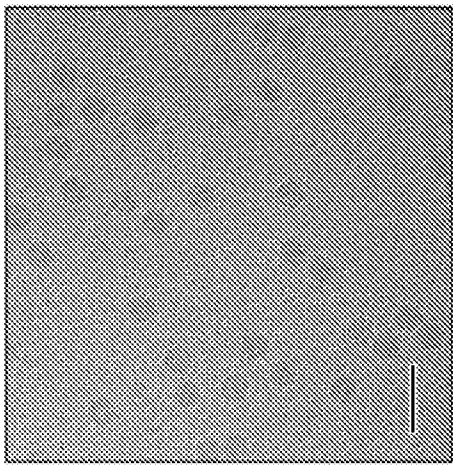


FIG. 15A

AAVv66 Structural Comparisons, RMSDc-α (Å)

Region:	VP3	VR I	VR II	VR III	VR IV	VR V	VR VI	VR VII	VR VIII	HI Loop	VR IX
Residues:	217-735	262-268	326-330	380-388	449-468	487-504	525-541	544-556	579-594	436-454	704-711
AAV2 (1LP3)	0.46	0.79	1.14	0.47	0.67	0.46	0.43	0.50	0.51	0.45	0.40
AAV3b (3KIC)	0.76	1.68	1.05	0.52	2.73	0.89	0.68	0.63	0.75	0.60	0.87

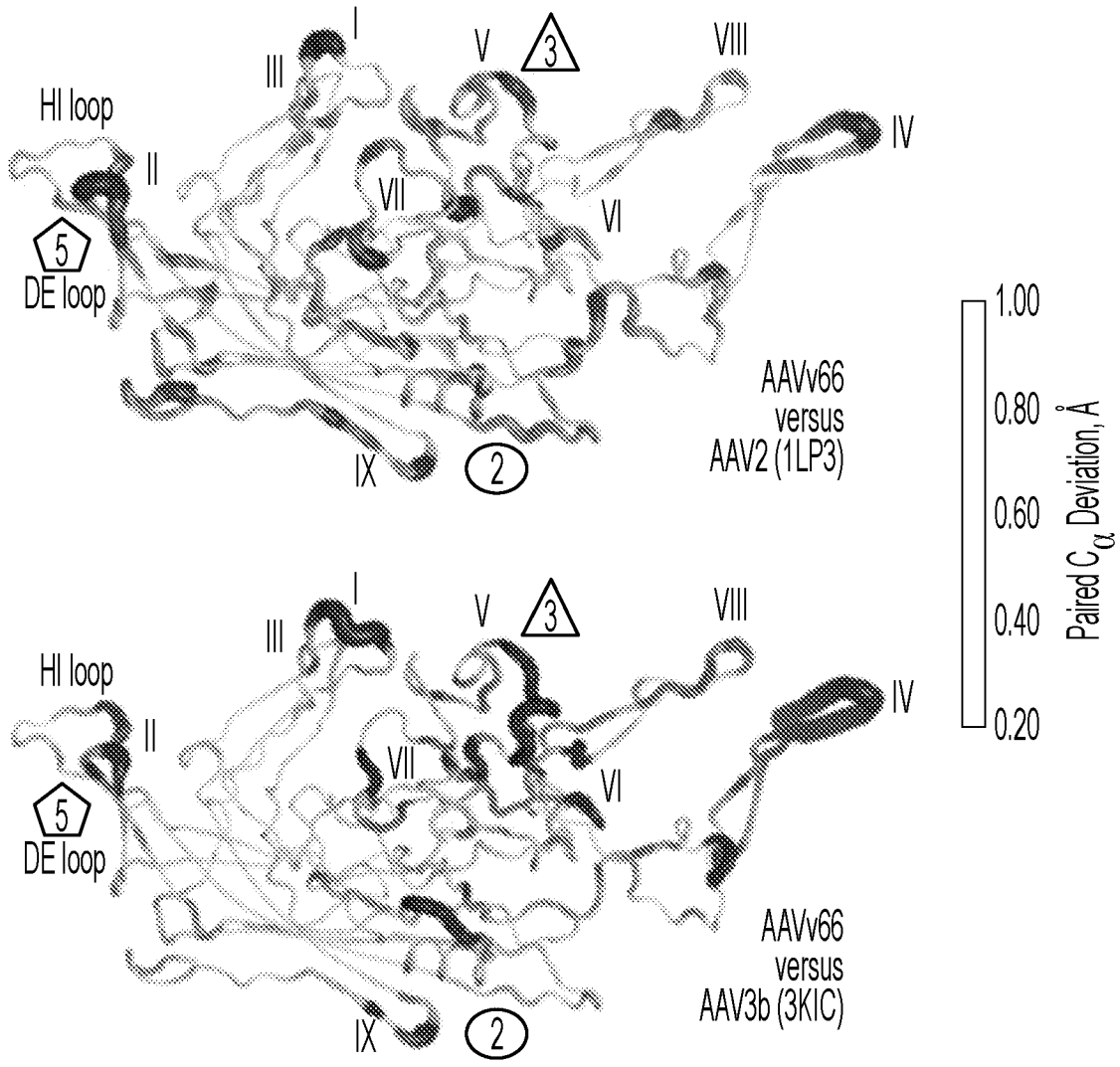


FIG. 16

SEQUENCE LISTING

<110> University of Massachusetts, Sichuan University

<120> CAPSID VARIANTS AND USES THEREOF

<130> U0120.70140WO00

<140> Not Yet Assigned
<141> Concurrently Herewith

<150> US 63/003,143
<151> 2020-03-31

<160> 32

<170> PatentIn version 3.5

<210> 1
<211> 735
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1

```
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1               5                   10                  15


Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
            20                  25                  30


Lys Pro Ala Glu Arg His Gln Asp Asp Ser Arg Gly Leu Val Leu Pro
        35                  40                  45


Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
    50                  55                  60


Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65                  70                  75                  80


Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
                85                  90                  95


Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110


Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125


Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140


Pro Val Glu His Ser Pro Ala Glu Pro Asp Ser Ser Ser Gly Thr Gly
145                 150                 155                 160
```

```
Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
            180                 185                 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
        195                 200                 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
    210                 215                 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225                 230                 235                 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
                245                 250                 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
            260                 265                 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
        275                 280                 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
    290                 295                 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305                 310                 315                 320

Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
                325                 330                 335

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
            340                 345                 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
        355                 360                 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
    370                 375                 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385                 390                 395                 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
                405                 410                 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
            420                 425                 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Lys Thr
```

```
        435                 440                 445

Asn Ala Pro Ser Gly Thr Thr Thr Met Ser Arg Leu Gln Phe Ser Gln
    450                 455                 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465                 470                 475                 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ala Ala Asp Asn Asn
                485                 490                 495

Asn Ser Asp Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
            500                 505                 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
        515                 520                 525

Asp Glu Glu Lys Tyr Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
    530                 535                 540

Gln Asp Ser Gly Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545                 550                 555                 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
                565                 570                 575

Gly Ser Val Ser Thr Asn Leu Gln Ser Gly Asn Thr Gln Ala Ala Thr
            580                 585                 590

Thr Asp Val Asn Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp
        595                 600                 605

Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr
    610                 615                 620

Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys
625                 630                 635                 640

His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn
                645                 650                 655

Pro Ser Thr Thr Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln
            660                 665                 670

Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys
        675                 680                 685

Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr
    690                 695                 700

Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr
705                 710                 715                 720
```

```
Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
                725                 730                 735


<210>  2
<211>  735
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  2

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1               5                   10                  15


Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
            20                  25                  30


Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
        35                  40                  45


Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
    50                  55                  60


Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65                  70                  75                  80


Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
                85                  90                  95


Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110


Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125


Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140


Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145                 150                 155                 160


Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175


Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
            180                 185                 190


Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
        195                 200                 205


Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
```

```
    210                 215                 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225                 230                 235                 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
                245                 250                 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
            260                 265                 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
        275                 280                 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
    290                 295                 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305                 310                 315                 320

Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
                325                 330                 335

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
            340                 345                 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
        355                 360                 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
    370                 375                 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385                 390                 395                 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
                405                 410                 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
            420                 425                 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
        435                 440                 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
    450                 455                 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465                 470                 475                 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
                485                 490                 495
```

```
Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
            500                 505                 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
        515                 520                 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
    530                 535                 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545                 550                 555                 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
                565                 570                 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Arg Gln Ala Ala Thr
            580                 585                 590

Ala Asp Val Asn Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp
        595                 600                 605

Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr
    610                 615                 620

Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys
625                 630                 635                 640

His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn
                645                 650                 655

Pro Ser Thr Thr Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln
            660                 665                 670

Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys
        675                 680                 685

Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr
    690                 695                 700

Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr
705                 710                 715                 720

Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
                725                 730                 735
```

<210> 3
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 3
gactgcatct ttgaacaata aatga 25

<210> 4
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 4
gaaacgaatt aaccggttta ttgattaa 28

<210> 5
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 5
catcactacg ctagatgact gcatctttga acaataaatg a 41

<210> 6
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 6
tagtatatcg agactcgaaa cgaattaacc ggtttattga ttaa 44

<210> 7
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 7
agagcggcat aaggacgaca gca 23

<210> 8
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 8
gcattctcct gtggagccag act 23

<210> 9
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 9
ttacttgagc agaacaaacg ctc 23

<210> 10
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 10
caaaacaaac actccaagcg gaac 24

<210> 11
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 11
aaccaccacg cagtccaggc ttc 23

<210> 12
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 12
atcaaaaaca tctgcggata acaacaacag tg 32

<210> 13
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 13
caacaacagt gaatattcgt ggac 24

<210> 14
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 14
tgaagaaaaa ttttttcctc agagcggggt tc 32

<210> 15

<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 15
tggaaaacaa ggctcgggaa aaa 23

<210> 16
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 16
acaagactcg gagaaaacta atgtg 25

<210> 17
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 17
caacctccag agaggcaaca cac 23

<210> 18
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 18
gagcggcaac agacaggcag cca 23

<210> 19
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 19
ggcagccacc gcagatgtca aca 23

<210> 20
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 20
gcgggctttg gtggtggt 18

<210> 21
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 21
tctaccggcc tctttttcc 20

<210> 22
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 22
tacagatact ggtcgatc 18

<210> 23
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 23
ctcaagtaat acagatactg g 21

<210> 24
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 24
ccgcttggag cgtttgtt 18

<210> 25
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 25
actcgctgct ggcggtaa 18

<210> 26
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 26
ttatccgcag ctgtttttg 19

<210> 27
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 27
tcgtccttgt ggctggcc 18

<210> 28
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 28
aagatgagaa ccccgctc 18

<210> 29
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 29
tttccaaaga tgagaacc 18

<210> 30
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 30
gtagatacag aaccatactg ctc 23

<210> 31
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 31
tggaggttgg tagatacaga accatactg 29

<210> 32
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 32
tgtgtgttgc cgctctgg 18