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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 April 2018 (19.04.2018)

(10) International Publication Number
WO 2018/071831 A1

(51) International Patent Classification:
C12N 15/861 (2006.01) ***C07K 14/015*** (2006.01)
C07K 14/005 (2006.01)

(21) International Application Number:
PCT/US2017/056614

(22) International Filing Date:
13 October 2017 (13.10.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

62/408,022	13 October 2016 (13.10.2016)	US
62/417,756	04 November 2016 (04.11.2016)	US
62/486,642	18 April 2017 (18.04.2017)	US

(71) Applicant: UNIVERSITY OF MASSACHUSETTS [US/US]; One Beacon Street, 31st. Floor, Boston, MA 02108 (US).

(72) Inventors: GAO, Guangping; 4 Edward Dunn Way, Westborough, MA 01581 (US). **XU, Guangchao**; 7 Pocono Road, Worcester, MA 01605 (US). **TAI, Phillip**; 61 Frank St., Apt. 42, Worcester, MA 01604 (US). **WEI, Yuquan**; Gaopent Street, Keyuan Street Road 4, No. 1, Chengdu, 610041 (CN). **LUO, Li**; 7 Pocono Road, Worcester, MA 01605 (US).

(74) Agent: MACDONALD, Kevin; 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, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,

(54) Title: AAV CAPSID DESIGNS

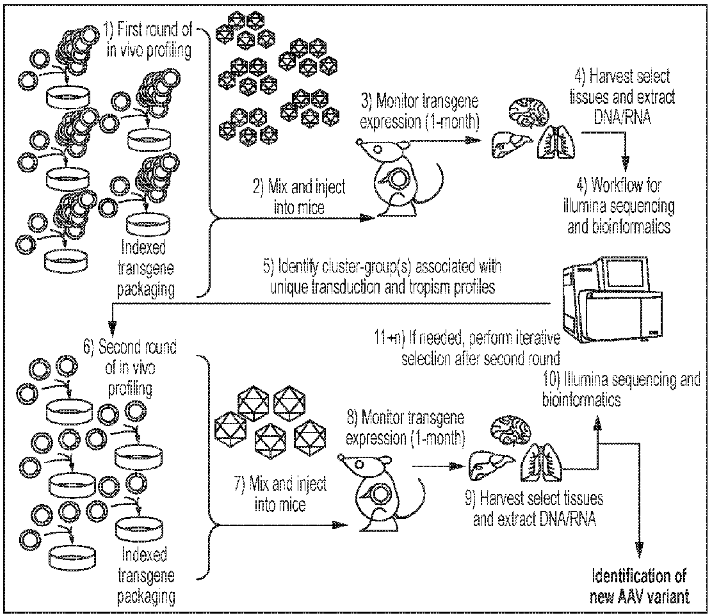


FIG. 6A

(57) Abstract: The disclosure in some aspects relates to recombinant adeno-associated viruses having distinct tissue targeting capabilities. In some aspects, the disclosure relates to gene transfer methods using the recombinant adeno-associated viruses. In some aspects, the disclosure relates to isolated AAV capsid proteins and isolated nucleic acids encoding the same.

WO 2018/071831 A1

[Continued on next page]

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))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

AAV CAPSID DESIGNS

RELATED APPLICATIONS

This application claims the benefit under 35 U.S.C. § 119(e) of the filing date of U.S. provisional application serial numbers USSN 62/486,642, filed April 18, 2017, entitled “AAV CAPSID DESIGNS”, 62/417,756, filed November 4, 2016, entitled “AAV CAPSID DESIGNS”, and 62/408,022, filed October 13, 2016, entitled “AAV CAPSID DESIGNS”, the entire contents of each application which are incorporated herein by reference.

FIELD OF THE DISCLOSURE

The disclosure in some aspects relates to isolated nucleic acids, compositions, and kits useful for identifying adeno-associated viruses in cells. In some aspects, the disclosure provides novel AAVs and methods of use thereof as well as related kits.

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

The disclosure in some aspects relates to novel AAVs for gene therapy applications. In some embodiments, AAVs described herein comprise amino acid variations in one or more capsid proteins that confer new or enhanced tissue tropism properties. According to some embodiments, variants of AAV2, AAV2/3 (*e.g.*, AAV2/3 hybrid), and AAV8 have been identified and are disclosed herein that possess useful tissue targeting properties. For example, variants of AAV8 are provided that are useful for transducing cells, such as, human hepatocytes (*e.g.*, present in liver tissue), central nervous system cells (CNS cells), and others. Variants of AAV2, AAV2/3 (*e.g.*, AAV2/3 hybrid), and AAV8 are provided that, in some embodiments, are useful for targeting cells of the ocular tissue (*e.g.*, the eye), gastrointestinal tract, respiratory system, breast tissue, pancreatic tissue, urinary tract tissue, uterine tissue, tissue associate with certain cancers (*e.g.*, breast cancer, prostate cancer, *etc.*), and other tissues. In some embodiments, the variant AAVs described herein target tissue other than the tissue targeted by their corresponding wild-type AAVs.

The disclosure in some aspects provides an isolated nucleic acid comprising a sequence encoding a polypeptide selected from the group consisting of: SEQ ID NO: 1-409, 435-868, and 1726-1988, which encodes an AAV capsid protein. In some embodiments, a fragment of the isolated nucleic acid is provided. In certain embodiments, the fragment of the isolated nucleic acid does not encode a peptide that is identical to a sequence of any one of SEQ ID NOs: 869, 870, or 871.

In some aspects, the disclosure provides a nucleic acid comprising a sequence selected from the group consisting of SEQ ID NO: 410-434, 876-1718, and 1989-2251. In some embodiments, the nucleic acid encodes an AAV capsid protein, or a variant thereof and/or an AAV assembly-activating protein (AAP), or a variant thereof. In some embodiments, the AAP is in a different open reading frame of the nucleic acid than the AAV capsid protein. In some embodiments, the AAP is AAV2 AAP (AAP-2), or variant thereof.

The disclosure in some aspects provides an isolated AAV capsid protein comprising an amino acid sequence selected from the group consisting of: SEQ ID NOs: 1-409, 435-868, and 1726-1988. In some embodiments, the isolated AAV capsid protein comprises a sequence selected from: SEQ ID NOs: 1-409, 837-852 or 1726-1814, wherein an amino acid of the sequence that is not identical to a corresponding amino acid of the sequence set forth as SEQ ID NO: 869 is replaced with a conservative substitution.

In some aspects, the disclosure provides AAV2/3 hybrid capsid proteins. In some embodiments, the isolated AAV capsid protein comprises a sequence selected from: SEQ ID NOs: 435-628 and 1815-1988, wherein an amino acid of the sequence that is not identical to a corresponding amino acid of the sequence set forth as SEQ ID NO: 869 or 870 is replaced with a conservative substitution.

In some embodiments, the isolated AAV capsid protein comprises a sequence selected from: SEQ ID NOs: 629-836 or 853-868, wherein an amino acid of the sequence that is not identical to a corresponding amino acid of the sequence set forth as SEQ ID NO: 871 is replaced with a conservative substitution.

In certain aspects of the disclosure, a composition is provided that comprises any of the foregoing isolated AAV capsid proteins. In some embodiments, the composition further comprises a pharmaceutically acceptable carrier. In some embodiments a composition of one or more of the isolated AAV capsid proteins of the disclosure and a physiologically compatible carrier is provided.

In certain aspects of the disclosure, a recombinant AAV (rAAV) is provided that comprises any of the foregoing isolated AAV capsid proteins. In some embodiments, a composition comprising the rAAV is provided. In certain embodiments, the composition comprising the rAAV further comprises a pharmaceutically acceptable carrier. A recombinant AAV is also provided, wherein the recombinant AAV includes one or more of the isolated AAV capsid proteins of the disclosure.

In some aspects of the disclosure, a host cell is provided that contains a nucleic acid that comprises a coding sequence selected from the group consisting of: SEQ ID NO: 410-434, 876-1718 and 1989-2251, that is operably linked to a promoter. In some embodiments, a composition comprising the host cell and a sterile cell culture medium is provided. In some embodiments, a composition comprising the host cell and a cryopreservative is provided.

According to some aspects of the disclosure, a method for delivering a transgene to a subject is provided. In some embodiments, the method comprises administering any of the foregoing rAAVs to a subject, wherein the rAAV comprises at least one transgene, and wherein the rAAV infects cells of a target tissue of the subject. In some embodiments, subject is selected from a mouse, a rat, a rabbit, a dog, a cat, a sheep, a pig, and a non-human primate. In one embodiment, the subject is a human.

In some embodiments, the at least one transgene is a protein coding gene. In certain embodiments, the at least one transgene encodes a small interfering nucleic acid. In certain embodiments, the small interfering nucleic acid is a miRNA. In certain embodiments, the small interfering nucleic acid is a miRNA sponge or TuD RNA that inhibits the activity of at least one miRNA in the subject. In certain embodiments, the miRNA is expressed in a cell of the target tissue In certain embodiments, the target tissue is liver, central nervous system (CNS), ocular, gastrointestinal, respiratory, breast, pancreas, urinary tract, or uterine tissue.

In some embodiments, the transgene expresses a transcript that comprises at least one binding site for a miRNA, wherein the miRNA inhibits activity of the transgene, in a tissue other than the target tissue, by hybridizing to the binding site.

In certain embodiments, the rAAV is administered to the subject intravenously, transdermally, intraocularly, intrathecally, intracererbally, orally, intramuscularly, subcutaneously, intranasally, or by inhalation.

According to some aspects of the disclosure, a method for generating a somatic transgenic animal model is provided. In some embodiments, the method comprises administering any of the foregoing rAAVs to a non-human animal, wherein the rAAV comprises at least one transgene, and wherein the rAAV infects cells of a target tissue of the non-human animal.

In some embodiments, the transgene is at least one protein coding gene. In certain embodiments, the transgene encodes at least one small interfering nucleic acid. In some embodiments, the transgene encodes at least one reporter molecule. In certain embodiments, the small interfering nucleic acid is a miRNA. In certain embodiments, the small interfering nucleic acid is a miRNA sponge or TuD RNA that inhibits the activity of at least one miRNA in the animal. In certain embodiments, the miRNA is expressed in a cell of the target tissue In certain embodiments, the target tissue is liver, central nervous system (CNS), ocular, gastrointestinal, respiratory, breast, pancreas, urinary tract, or uterine tissue.

In some embodiments, the transgene expresses a transcript that comprises at least one binding site for a miRNA, wherein the miRNA inhibits activity of the transgene, in a tissue other than the target tissue, by hybridizing to the binding site.

According to some aspects of the disclosure, methods are provided for generating a somatic transgenic animal model that comprise administering any of the foregoing rAAVs to a non-human animal, wherein the rAAV comprises at least one transgene, wherein the transgene expresses a transcript that comprises at least one binding site for a miRNA, wherein the miRNA inhibits activity of the transgene, in a tissue other than a target tissue, by hybridizing to the binding site of the transcript.

In some embodiments, the transgene comprises a tissue specific promoter or inducible promoter. In certain embodiments, the tissue specific promoter is a liver-specific thyroxin binding globulin (TBG) promoter, an insulin promoter, a glucagon promoter, a somatostatin promoter, mucin-2 promoter, a pancreatic polypeptide (PPY) promoter, a synapsin-1 (Syn) promoter, a retinoschisin promoter, a K12 promoter, a CC10 promoter, a surfactant protein C (SP-C) promoter, a PRC1 promoter, a RRM2 promoter, uroplakin 2 (UPII) promoter, or a lactoferrin promoter.

In certain embodiments, the rAAV is administered to the animal intravenously, transdermally, intraocularly, intrathecally, orally, intramuscularly, subcutaneously, intranasally, or by inhalation. According to some aspects of the disclosure, a somatic transgenic animal model is provided that is produced by any of the foregoing methods.

In other aspects of the disclosure, a kit for producing a rAAV is provided. In some embodiments, the kit comprises a container housing an isolated nucleic acid having a sequence of any one of SEQ ID NO: 410-434, 876-1718, and 1989-2251. In some embodiments, the kit comprises a container housing an isolated nucleic acid encoding a polypeptide having a sequence of any one of SEQ ID NO: 1-409, 435-868, or 1726-1988. In some embodiments, the kit further comprises instructions for producing the rAAV. In some embodiments, the kit further comprises at least one container housing a recombinant AAV vector, wherein the recombinant AAV vector comprises a transgene.

In other aspects of the disclosure, a kit is provided that comprises a container housing a recombinant AAV having any of the foregoing isolated AAV capsid proteins. In some embodiments, the container of the kit is a syringe.

In other aspects, the disclosure relates to the use of AAV based vectors as vehicles for, delivery of genes, therapeutic, prophylactic, and research purposes as well as the development of somatic transgenic animal models.

In some aspects, the disclosure relates to AAV serotypes that have demonstrated distinct tissue/cell type tropism and can achieve stable somatic gene transfer in animal tissues at levels similar to those of adenoviral vectors (*e.g.*, up to 100% *in vivo* tissue transduction depending upon target tissue and vector dose) in the absence of vector related toxicology. In other aspects, the disclosure relates to AAV serotypes having liver, central nervous system (CNS), ocular, gastrointestinal, respiratory, breast, pancreas, urinary tract, or uterine tissue-targeting capabilities. These tissues are associated with a broad spectrum of human diseases including neurological, metabolic, diabetic, ocular, respiratory, gastrointestinal, urinary tract, and reproductive diseases and certain cancers.

In some embodiments the rAAV includes at least one transgene. The transgene may be one which causes a pathological state. In some embodiments, the transgene encoding a protein that treats a pathological state.

In another aspect the novel AAVs of the disclosure may be used in a method for delivering a transgene to a subject. The method is performed by administering a rAAV of the disclosure to a subject, wherein the rAAV comprises at least one transgene. In some embodiments the rAAV targets a predetermined tissue of the subject.

In another aspect the AAVs of the disclosure may be used in a method for generating a somatic transgenic animal model. The method is performed by administering a rAAV of the disclosure to an animal, wherein the rAAV comprises at least one transgene, wherein the transgene causes a pathological state, and wherein the rAAV targets a predetermined tissue of the animal.

The transgene may express a number of genes including cancer related genes, pro-apoptotic genes and apoptosis-related genes. In some embodiments the transgene expresses a small interfering nucleic acid capable of inhibiting expression of a cancer related gene. In other embodiments the transgene expresses a small interfering nucleic acid capable of inhibiting expression of an apoptosis-related gene. The small interfering nucleic acid in other embodiments is a miRNA or shRNA. According to other embodiments the transgene expresses a toxin, optionally wherein the toxin is DTA. In other embodiments the transgene expresses a reporter gene that is optionally a reporter enzyme, such as Beta-Galactosidase or a Fluorescent protein, such as GFP or luciferase.

The transgene may express a miRNA. In other embodiments the transgene expresses a miRNA sponge, wherein miRNA sponge inhibits the activity of one or more miRNAs in the animal. The miRNA may be an endogenous miRNA or it may be expressed in a cell of a liver, central nervous system (CNS), ocular, gastrointestinal, respiratory, breast, pancreas, urinary tract, or uterine tissue, in some embodiments.

The rAAV may transduce many different types of tissue, such as neurons, squamous epithelial cells, renal proximal or distal convoluted tubular cells, mucosa gland cells, blood vessel endothelial cells, endometrial cells, retinal cells, or certain cancer cells (*e.g.*, breast cancer cells, prostate cancer cells, *etc.*).

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. The rAAV may be administered by any route. For instance it may be administered intravenously (*e.g.*, by portal vein injection) in some embodiments.

In some embodiments the transgene includes a tissue specific promoter such as a liver-specific thyroxin binding globulin (TBG) promoter, an insulin promoter, a glucagon promoter, a somatostatin promoter, mucin-2 promoter, a pancreatic polypeptide (PPY) promoter, a synapsin-1 (Syn) promoter, a retinoschisin promoter, a K12 promoter, a CC10 promoter, a surfactant protein C (SP-C) promoter, a PRC1 promoter, a RRM2 promoter, uroplakin 2 (UPII) promoter, or a lactoferrin promoter.

The somatic transgenic animal model may be a mammal, such as a mouse, a rat, a rabbit, a dog, a cat, a sheep, a pig, a non-human primate.

In some embodiments a putative therapeutic agent may be administered to the somatic transgenic animal model to determine the effect of the putative therapeutic agent on the pathological state in the animal.

In another aspect the disclosure is a somatic transgenic animal produced by the methods described herein.

A kit for producing a rAAV that generates a somatic transgenic animal having a pathological state in a predetermined tissue is provided according to another aspect of the disclosure. The kit includes at least one container housing a recombinant AAV vector, at least one container housing a rAAV packaging component, and instructions for constructing and packaging the recombinant AAV.

The rAAV packaging component may include a host cell expressing at least one rep gene and/or at least one cap gene. In some embodiments the host cell is a 293 cell. In other embodiments the host cell expresses at least one helper virus gene product that affects the production of rAAV containing the recombinant AAV vector. The at least one cap gene may encode a capsid protein from an AAV serotype that targets the predetermined tissue.

In other embodiments a rAAV packaging component includes a helper virus optionally wherein the helper virus is an adenovirus or a herpes virus.

The rAAV vector and components therein may include any of the elements described herein. For instance, in some embodiments the rAAV vector comprises a transgene, such as any of the transgenes described herein. In some embodiments the transgene expresses a miRNA inhibitor (*e.g.*, a miRNA sponge or TuD RNA), wherein miRNA inhibitor inhibits the activity of one or more miRNAs in the somatic transgenic animal.

Each of the limitations of the disclosure can encompass various embodiments of the disclosure. It is, therefore, anticipated that each of the limitations of the disclosure involving any one element or combinations of elements can be included in each aspect of the disclosure. This disclosure is not limited in its application to the details of construction and the arrangement of components set forth in the following description or illustrated in the drawings. The disclosure is capable of other embodiments and of being practiced or of being carried out in various ways.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGs. 1A-1B show workflow schematics for the identification of AAV variants. FIG. 1A depicts high-throughput detection of novel AAV variants in selected human tissues. Proviral capsid sequences are amplified using high-cycle PCR, followed by low-cycle PCR to barcode the amplicon libraries for multiplexed single-molecule, real-time (SMRT) sequencing. FIG. 1B shows a summary of the pipeline for bioinformatics analysis of sequencing data.

FIGs. 2A-2D show data relating to *in vivo* detection of FFLuc transgene activity with different administrations of selected AAV8 variants. FIG. 2A shows luciferase activities of different AAV8 variants were evaluated at week 6 after IV (intravenous), IM (intramuscular), or IN (intranasal) injection. FIGs. 2B-2D data relating to evaluation of FFLuc activity for each variant, B2 (FIG. 2B), B3 (FIG. 2C), and B61 (FIG. 2D), compared to AAV8 (mean±SD, n=3, *t* test).

FIGs. 3A-3B show data relating to evaluation of FFLuc transgene activity delivered by the AAV8 variant B61 compared to AAV9 at day 21 after neonatal injection. Luciferase activities and genome copies of brain (FIG. 3A) and spinal cord (FIG. 3B) were detected (mean±SD, n=5, *t* test).

FIGs. 4A-4B show data relating to *in vivo* detection of FFLuc transgene activity after right hindlimb intramuscular (IM) injection of the AAV8 variant B44 compared to AAV8. FIG. 4A shows whole animal Luciferase expression of variant B44 was evaluated at week 6 after IM injection. FIG. 4B shows evaluation of muscle (RTA, right tibialis anterior; LTA, left tibialis anterior), liver, and heart. Luciferase activities (left bar graph) and relative ratios (right bar graph) for B44 compared to AAV8 (mean±SD, n=3).

FIG. 5 shows a phylogenic comparison of AAV8 variants (B2, B3, B61) to other AAV serotypes.

FIG. 6A shows a schematic depiction of a workflow for the *in vivo* characterization of novel AAV variants by high-throughput tropism screening.

FIG. 6B shows a schematic depiction of a workflow for the NHP characterization of novel AAV variants by high-throughput tropism screening.

FIG. 7 shows a scatter plot displaying the distribution of distinct AAV2 capsid variants (409 total) and AAV2/3 variants (194 total) harboring one or more single-amino-acid variants.

FIG. 8 shows diagrams of vector constructs used in the multiplexed screening of discovered capsid variants. Unique 6-bp barcodes were cloned into transgenes and packaged into candidate capsid variants.

FIG. 9 shows a schematic of an indexed transgene and high-throughput sequencing library design to assess capsid variant tropism profiling. The indexed and adapter cassette containing a 6-bp barcode (1° barcode) and a BstEII restriction site can be cloned into vector constructs using flanking BsrGI and SacI sites. Whole crude DNA from rAAV-treated tissues containing both host genome and vector genomes was cut with BstEII enzyme. The resulting 5'- overhang was used to specifically ligate to an adapter containing a second barcode, which allows for further multiplexed sequencing and streamlining; and a 5'-biotin modification, which can be used to select for adapter-containing fragments using magnetic bead enrichment. Enriched material can then undergo PCR amplification using primers specific to adapter and transgene sequences to produce libraries for high-throughput sequencing. SEQ ID NOs.: 1719-1725 are shown from top to bottom.

DETAILED DESCRIPTION

Adeno-associated virus (AAV) is a small (~26 nm) replication-defective, non-enveloped virus that generally depends on the presence of a second virus, such as adenovirus or herpes virus, for its growth in cells. AAV is not known to cause disease and induces a very mild immune response. AAV can infect both dividing and non-dividing cells and may incorporate its genome into that of the host cell. These features make AAV a very attractive candidate for creating viral vectors for gene therapy. Prototypical AAV vectors based on serotype 2 provided a proof-of-concept for non-toxic and stable gene transfer in murine and large animal models, but exhibited poor gene transfer efficiency in many major target tissues. The disclosure in some aspects seeks to overcome this shortcoming by providing novel AAVs having distinct tissue targeting capabilities for gene therapy and research applications.

In some aspects of the disclosure new AAV capsid proteins are provided that have distinct tissue targeting capabilities. In some embodiments, an AAV capsid protein is isolated from the tissue to which an AAV comprising the capsid protein targets. In some aspects, methods for delivering a transgene to a target tissue in a subject are provided. The transgene delivery methods may be used for gene therapy (*e.g.*, to treat disease) or research (*e.g.*, to create a somatic transgenic animal model) applications.

Methods for Discovering AAVs

Much of the biology of AAV is influenced by its capsid. Consequently, methods for discovering novel AAVs have been largely focused on isolating DNA sequences for AAV capsids. A central feature of the adeno-associated virus (AAV) latent life cycle is persistence in the form of integrated and/or episomal genomes in a host cell. Methods used for isolating novel AAV include PCR based molecular rescue of latent AAV DNA genomes, infectious virus rescue of latent proviral genome from tissue DNAs *in vitro* in the presence of adenovirus helper function, and rescue of circular proviral genome from tissue DNAs by rolling-circle-linear amplification, mediated by an isothermal phage Phi-29 polymerase. All of these isolation methods take advantage of the latency of AAV proviral DNA genomes and focus on rescuing persistent viral genomic DNA.

In some aspects, the disclosure relates to the discovery that novel AAV variants with desirable tissue tropisms can be identified from *in vivo* tissues of a subject. Without wishing to be bound by any particular theory, the use of *in vivo* tissue exploits the natural reservoir of genomic diversity observed among viral genomic sequences isolated from both normal and tumor tissues of a subject. Thus in some embodiments, *in vivo* tissues act as natural incubators for viral (*e.g.*, viral capsid protein) diversity through selective pressure and/or immune evasion.

In some aspects, the disclosure relates to the discovery that PCR products resulting from amplification of AAV DNA (*e.g.*, AAV DNA isolated or extracted from a host cell or *in vivo* tissue of a subject) can be subjected to high-throughput single-molecule, real-time (SMRT) sequencing to identify novel capsid protein variants. As used herein, "single-molecule, real-time (SMRT) sequencing" refers to a parallelized single molecule sequencing method, for example as described by Roberts et al. (2013) *Genome Biology* 14:405, doi:10.1186/gb-2013-14-7-405. Without wishing to be bound by any particular theory, the use of SMRT sequencing removes the need to perform viral genome reconstruction and chimera prediction from aligned short-read fragments obtained from other conventional high-throughput genome sequencing methodologies.

Endogenous latent AAV genomes are transcriptionally active in mammalian cells (*e.g.*, cells of nonhuman primate tissues such as liver, spleen and lymph nodes). Without wishing to be bound by theory, it is hypothesized that to maintain AAV persistence in host, low levels of transcription from AAV genes could be required and the resulting cap RNA could serve as more suitable and abundant substrates to retrieve functional cap sequences for vector development. Both rep and cap gene transcripts are detected with variable abundances by RNA detection methods (*e.g.*, RT-PCR). The presence of cap gene transcripts and ability to generate cDNA of cap RNA through reverse transcription (RT) *in vitro* significantly increases abundance of templates for PCR-based rescue of novel cap sequences from tissues and enhances the sensitivity of novel AAV discovery.

Novel cap sequences may also be identified by transfecting cells with total cellular DNAs isolated from the tissues that harbor proviral AAV genomes at very low abundance, The cells may be further transfected with genes that provide helper virus function (*e.g.*, adenovirus) to trigger and/or boost AAV gene transcription in the transfected cells. In some embodiments, novel cap sequences of the disclosure may be identified by isolating cap mRNA from the transfected cells, creating cDNA from the mRNA (*e.g.*, by RT-PCR) and sequencing the cDNA.

Isolated Capsid Proteins and Nucleic Acids Encoding the Same

AAVs isolated from mammals, particularly non-human primates, are useful for creating gene transfer vectors for clinical development and human gene therapy applications. The disclosure provides in some aspects novel AAVs that have been discovered in various *in vivo* tissues (*e.g.*, liver, brain, gastric, respiratory, breast, pancreatic, rectal, prostate, urologic, and cervical tissues) using the methods disclosed herein. In some embodiments, the tissue(s) in which a novel AAV variant is discovered is a cancerous tissue (*e.g.*, a tumor or a cancer cell). In some embodiments, nucleic acids encoding capsid proteins of these novel AAVs have been discovered in viral genomic DNA isolated from the human tissues. Examples of tissues in which novel AAV capsid proteins have been discovered are described in Table 1.

Nucleic acid and protein sequences as well as other information regarding the AAVs are set forth in Tables 3-5 and 8, and in the sequence listing.

Isolated nucleic acids of the disclosure that encode AAV capsid proteins include any nucleic acid having a sequence as set forth in any one of SEQ ID NOs: 410-435, 876-1718, or 1989-2251, as well as any nucleic acid having a sequence with substantial homology thereto. In some embodiments, isolated nucleic acids of the disclosure include any nucleic acid having a sequence encoding a polypeptide having a sequence as set forth in any one of SEQ ID NOs: 1-409, 435-868, and 1726-1988. In some embodiments, the disclosure provides an isolated nucleic acid that has substantial homology with a nucleic acid having a sequence as set forth in any one of SEQ ID NOs: 410-435, 876-1718, and 1989-2251, but that does not encode a protein having an amino acid sequence as set forth in SEQ ID NOs: 869, 870, or 871.

In some embodiments, isolated AAV capsid proteins of the disclosure include any protein having an amino acid sequence as set forth in any one of SEQ ID NOs: 1-409, 837-852, or 1726-1814 as well as any protein having substantial homology thereto. In some embodiments, the disclosure provides an isolated capsid protein that has substantial homology with a protein having a sequence as set forth in any one of SEQ ID NOs 1-409, 837-852, or 1726-1814, but that does not have an amino acid sequence as set forth in SEQ ID NO: 869.

In some embodiments, isolated AAV capsid proteins of the disclosure include any protein having an amino acid sequence as set forth in any one of SEQ ID NOs: 435-628 or 1815-1988 as well as any protein having substantial homology thereto. In some embodiments, the disclosure provides an isolated capsid protein that has substantial homology with a protein having a sequence as set forth in any one of SEQ ID NOs 435-628 or 1815-1988, but that does not have an amino acid sequence as set forth in SEQ ID NO: 869 or 870.

In some embodiments, isolated AAV capsid proteins of the disclosure include any protein having an amino acid sequence as set forth in any one of SEQ ID NOs: 629-836 or 853-868 as well as any protein having substantial homology thereto. In some embodiments, the disclosure provides an isolated capsid protein that has substantial homology with a protein having a sequence as set forth in any one of SEQ ID NOs 629-836 or 853-868, but that does not have an amino acid sequence as set forth in SEQ ID NO: 871.

“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. Table 6 provides examples of variant amino acids.

Alternatively for nucleic acids, homology can be determined by hybridization of polynucleotides under conditions that form stable duplexes between homologous regions, followed by digestion with single-stranded-specific nuclease(s), and size determination of the digested fragments. DNA sequences that are substantially homologous can be identified in a Southern hybridization experiment under, for example, stringent conditions, as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art.

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.

An example of an isolated nucleic acid that encodes a polypeptide comprising an AAV capsid protein is a nucleic acid having a sequence selected from the group consisting of: SEQ ID NO: 410-434, 876-1718, and 1989-2251. A fragment of an isolated nucleic acid encoding an AAV capsid sequence may be useful for constructing a nucleic acid encoding a desired capsid sequence. Fragments may be of any appropriate length. In some embodiments, a fragment (portion) of an isolated nucleic acid encoding an AAV capsid sequence may be useful for constructing a nucleic acid encoding a desired capsid sequence. Fragments may be of any appropriate length (*e.g.*, at least 6, at least 9, at least 18, at least 36, at least 72, at least 144, at least 288, at least 576, at least 1152 or more nucleotides in length). For example, a fragment of nucleic acid sequence encoding a polypeptide of a first AAV capsid protein may be used to construct, or may be incorporated within, a nucleic acid sequence encoding a second AAV capsid sequence to alter the properties of the AAV capsid. In some embodiments, AAV capsid proteins that comprise capsid sequence fragments from multiple AAV serotypes are referred to as chimeric AAV capsids. The fragment may be a fragment that does not encode a peptide that is identical to a sequence of any one of SEQ ID NOs: 869, 870, or 871. For example, a fragment of nucleic acid sequence encoding a variant amino acid (compared with a known AAV serotype) may be used to construct, or may be incorporated within, a nucleic acid sequence encoding an AAV capsid sequence to alter the properties of the AAV capsid. In some embodiments, a nucleic acid sequence encoding an AAV variant may comprise about 1 to about 100 amino acid variants compared with a known AAV serotype (*e.g.,* AAV serotype 2, AAV2/3 (*e.g.*, AAV2/3 hybrid) or AAV8). In some embodiments, a nucleic acid sequence encoding an AAV variant may comprise about 5 to about 50 amino acid variants compared with a known AAV serotype (*e.g.,* AAV serotype 2, AAV2/3 (*e.g.*, AAV2/3 hybrid) or AAV8). In some embodiments, a nucleic acid sequence encoding an AAV variant may comprise about 10 to about 30 amino acid variants compared with a known AAV serotype (*e.g.,* AAV serotype 2, AAV2/3 (*e.g.*, AAV2/3 hybrid) or AAV8). In some embodiments, a nucleic acid sequence encoding an AAV variant may comprise 1, or 2, or 3, or 4, 5, or 6, or 7, or 8, or 9, or 10, or 11, or 12, or 13, or 14, or 15, or 16, or 17, or 18, or 19, or 20 amino acid variants compared with a known AAV serotype (*e.g.,* AAV serotype 2, AAV2/3 (*e.g.*, AAV2/3 hybrid) or AAV8). For example, a nucleic sequence encoding an AAV variant (*e.g.*, SEQ ID NO: 861 may comprise 3 amino acid variants compared with a known AAV serotype (*e.g.*, AAV8). A recombinant cap sequence may be constructed having one or more of the 3 amino acid variants by incorporating fragments of a nucleic acid sequence comprising a region encoding a variant amino acid into the sequence of a nucleic acid encoding the known AAV serotype. The fragments may be incorporated by any appropriate method, including using site directed mutagenesis. Thus, new AAV variants may be created having new properties.

In some aspects, the disclosure provides isolated nucleic acids encoding AAV assembly-activating proteins (AAPs), or variants thereof. As used herein, an “assembly activating protein” or “AAP” is a protein chaperone that functions to target newly synthesized capsid proteins (*e.g.*, VP proteins, such as AAV VP1, VP2, and VP3) to the nucleolus of a cell thereby promoting encapsidation of viral genomes. Generally, an AAP is encoded in the *cap* gene of an adeno-associated virus. For example, AAP-2 is encoded in the *cap* gene of

AAV2. Other examples of AAPs include but are not limited to AAP-1, AAP-3, AAP-4, AAP-5, AAP-8, AAP-9, AAP-11 and AAP-12, for example as described by Sonntag et al. *J. Virol.* 2011 Dec. 85(23): 12686-12697. In some embodiments, an AAP is translated from a different open reading frame (ORF) of the *cap* gene than a capsid protein (*e.g.*, VP1, VP2, VP3). For example, in some embodiments, a capsid protein (*e.g.*, AAV2 VP1, VP2, VP3) is translated from ORF 1 of a *cap* gene and an AAP (*e.g.*, AAP-2) is translated from ORF 2 of the *cap* gene. In some embodiments, an isolated nucleic acid encoding an AAP comprises or consists of a sequence selected from SEQ ID NO: 410-434 and 876-1718.

Recombinant AAVs

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 a capsid protein having an amino acid sequence as set forth in any one of SEQ ID NOs 1-409, 435-852, 859-874, or 1726-1988, or a protein having substantial homology thereto.

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 any one of SEQ ID NOs 1-409, 435-868, or 1726-1988) 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 aspects, the instant disclosure describes variants of wild-type AAV serotypes In some embodiments, the variants have altered tissue tropism. In some embodiments, the AAV variants described herein comprise amino acid variations (*e.g.*, substitution, deletion, insertion) within the *cap* gene. As discussed above, all three capsid proteins are transcribed from a single *cap* gene. Accordingly, in some embodiments, an amino acid variation within a *cap* gene is present in all three capsid proteins encoded by said *cap* gene. Alternatively, in some embodiments, an amino acid variation may not be present in all three capsid proteins. In some embodiments, an amino acid variation occurs only in the VP1 capsid protein. In some embodiments, an amino acid variation occurs only in the VP2 capsid protein. In some embodiments, an amino acid variation occurs only within the VP3 capsid protein. In some embodiments, an AAV variant comprises more than one variation in a *cap* gene. In some embodiments, the more than one variation occur within the same capsid protein (*e.g.,* within VP3). In some embodiments, the more than one variation occur within different capsid proteins (*e.g.,* at least one variation in VP2 and at least one variation in VP3).

In some embodiments, the AAV variants described herein are variants of AAV2, AAV2/3 (*e.g.*, AAV2/3 hybrid) or AAV8. 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 AAV3 variants described herein may be useful for delivering gene therapy to CNS tissue, kidney tissue, or ocular tissue. It is also known that AAV3 efficiently transduces cancerous human hepatocytes. Accordingly, in some embodiments, the AAV3 variants described herein may be useful for delivering gene therapy to cancerous and normal human hepatocytes. AAV8 is known to target tissue of the liver tissue, respiratory tissue, and the eye. Accordingly, in some embodiments, the AAV8 variants described herein may be useful for delivering gene therapy to the liver tissue, respiratory tissue or the eye.

It should be appreciated that the AAV2, AAV2/3 (*e.g.*, AAV2/3 hybrid) and AAV8 variants described herein may comprise one or more variations within the *cap* gene compared with a corresponding wild-type AAV. Therefore, in some embodiments, the AAV2, AAV2/3 (*e.g.*, AAV2/3 hybrid) and AAV8 variants described herein may have a tissue tropism useful for delivering gene therapy to additional tissue types that are not targeted by wild-type AAV2, AAV2/3 (*e.g.*, AAV2/3 hybrid) or AAV8. For example, in some embodiments, AAV8 variants described herein (*e.g.,* B61; SEQ ID NO: 865) may be useful for delivering gene therapy to the central nervous system (CNS). In some embodiments, AV2, AAV2/3 (*e.g.*, AAV2/3 hybrid), or AAV8 variants described herein may be useful for targeting cells of the kidney or cells of the liver. In some embodiments, AAV2, AAV2/3 (*e.g.*, AAV2/3 hybrid), or AAV8 variants described herein may be useful for targeting gene therapy to the liver, spleen, heart or brain.

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 be useful for delivering gene therapy to cardiac cells (*e.g.,* heart tissue). Accordingly, in some embodiments, AAV variants described herein may be useful for the treatment of cardiovascular disorders. As used herein, a "cardiovascular disorder" is a disease or condition of the cardiovascular system. A cardiovascular disease may affect the heart, circulatory system, arteries, veins, blood vessels and/or capillaries. A cardiovascular disorder may be of a genetic origin, either inherited or acquired through a somatic mutation. Non-limiting examples of cardiovascular disorders include rheumatic heart disease, valvular heart disease, hypertensive heart disease, aneurysm, atherosclerosis, hypertension (*e.g.,* high blood pressure), peripheral arterial disease (PAD), ischemic heart disease, angina, coronary heart disease, coronary artery disease, myocardial infarction, cerebral vascular disease, transient ischemic attack, inflammatory heart disease, cardiomyopathy, pericardial disease, congenital heart disease, heart failure, stroke, and myocarditis due to Chagas disease.

In some embodiments, AAV variants described herein may target the lung and/or tissue of the pulmonary system (*e.g.*, respiratory system). Accordingly, in some embodiments, AAV variants described herein may be useful for treatment of pulmonary disease. As used herein a "pulmonary disease" is a disease or condition of the pulmonary system. A pulmonary disease may affect the lungs or muscles involved in breathing. A pulmonary disease may be of a genetic origin, either inherited or acquired through a somatic mutation. A pulmonary disease may be a cancer of the lung, including but not limited to, non-small cell lung cancer, small cell lung cancer, and lung carcinoid tumor. Further non-limiting examples of pulmonary diseases include acute bronchitis, acute respiratory distress syndrome (ARDS), asbestosis, asthma, bronchiectasis, bronchiolitis, bronchiolitis obliterans organizing pneumonia (BOOP), bronchopulmonary dysplasia, byssinosis, chronic bronchitis, coccidioidomycosis (Cocci), chronic obstructive pulmonary disorder (COPD), cryptogenic organizing pneumonia (COP), cystic fibrosis, emphysema, Hantavirus Pulmonary Syndrome, histoplasmosis, Human Metapneumovirus, hypersensitivity pneumonitis, influenza, lymphangiomatosis, mesothelioma, Middle Eastern Respiratory Syndrome, non-tuberculosis Mycobacterium, Pertussis, Pneumoconiosis (Black Lung Disease), pneumonia, primary ciliary dyskinesia, primary pulmonary hypertension, pulmonary arterial hypertension, pulmonary fibrosis, pulmonary vascular disease, Respiratory Syncytial Virus (RSV), sarcoidosis, Severe Acute Respiratory Syndrome (SARS), silicosis, sleep apnea, Sudden Infant Death Syndrome (SIDS), and tuberculosis.

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 target kidney tissue. Accordingly, in some embodiments, AAV variants described herein may be useful for treatment of kidney disease. As used herein a “kidney disease” is a disease or condition of the liver. A kidney disease may be of a genetic origin, either inherited or acquired through a somatic mutation. A kidney disease may be a cancer of the kidney, including but not limited to renal cell cancer, clear cell cancer, papillary cancer type 1, papillary cancer type 2, chromophobe cancer, oncocytic cell cancer, collecting duct cancer, transitional cell cancer of the renal pelvis and Wilm’s tumor. Further non-limiting examples of kidney disease include Abderhalden–Kaufmann–Lignac syndrome (Nephropathic Cystinosis), Acute Kidney Failure/Acute Kidney Injury, Acute Lobar Nephronia, Acute Phosphate Nephropathy, Acute Tubular Necrosis, Adenine Phosphoribosyltransferase Deficiency, Adenovirus Nephritis, Alport Syndrome, Amyloidosis, Angiomyolipoma, Analgesic Nephropathy, Angiotensin Antibodies and Focal Segmental Glomerulosclerosis, Antiphospholipid Syndrome, Anti-TNF-α Therapy-related Glomerulonephritis, APOL1 Mutations, Apparent Mneralocorticoid Excess Syndrome, Aristolochic Acid Nephropathy, Balkan Endemic Nephropathy, Bartter Syndrome, Beeturia, β-Thalassemia Renal Disease, Bile Cast Nephropathy, BK Polyoma, C1q Nephropathy, Cardiorenal syndrome, CFHR5 nephropathy, Cholesterol Emboli, Churg–

Strauss syndrome, Chyluria, Collapsing Glomerulopathy, Collapsing Glomerulopathy Related to CMV, Congenital Nephrotic Syndrome, Conorenal syndrome (Mainzer-Saldino Syndrome or Saldino-Mainzer Disease), Contrast Nephropathy, Copper Sulfate Intoxication, Cortical Necrosis, Cryoglobuinemia, Crystal-Induced Acute Kidney injury, Cystic Kidney Disease, Acquired, Cystinuria, Dense Deposit Disease (MPGN Type 2), Dent Disease (X-linked Recessive Nephrolithiasis), Dialysis Disequilibrium Syndrome, Diabetic Kidney Disease, Diabetes Insipidus, EAST syndrome, Ectopic Ureter, Edema, Erdheim-Chester Disease, Fabry's Disease, Familial Hypocalciuric Hypercalcemia, Fanconi Syndrome, Fraser syndrome, Fibronectin Glomerulopathy, Fibrillary Glomerulonephritis and Immunotactoid Glomerulopathy, Fraley syndrome, Focal Segmental Glomerulosclerosis, Focal Sclerosis, Focal Glomerulosclerosis, Galloway Mowat syndrome, Gitelman Syndrome, Glomerular Diseases, Glomerular Tubular Reflux, Glycosuria, Goodpasture Syndrome, Hemolytic Uremic Syndrome (HUS), Atypical Hemolytic Uremic Syndrome (aHUS), Hemophagocytic Syndrome, Hemorrhagic Cystitis, Hemosiderosis related to Paroxysmal Nocturnal Hemoglobinuria and Hemolytic Anemia, Hepatic Veno-Occlusive Disease, Sinusoidal Obstruction Syndrome, Hepatitis C-Associated Renal Disease, Hepatorenal Syndrome, HIV-Associated Nephropathy (HIVAN), Horseshoe Kidney (Renal Fusion), Hunner's Ulcer, Hyperaldosteronism, Hypercalcemia, Hyperkalemia, Hypermagnesemia, Hypernatremia, Hyperoxaluria, Hyperphosphatemia, Hypocalcemia, Hypokalemia, Hypokalemia-induced renal dysfunction, Hypomagnesemia, Hyponatremia, Hypophosphatemia, IgA Nephropathy, IgG4 Nephropathy, Interstitial Cystitis, Painful Bladder Syndrome, Interstitial Nephritis, Ivemark's syndrome, Kidney Stones, Nephrolithiasis, Leptospirosis Renal Disease, Light Chain Deposition Disease, Monoclonal Immunoglobulin Deposition Disease, Liddle Syndrome, Lightwood-Albright Syndrome, Lipoprotein Glomerulopathy, Lithium Nephrotoxicity, LMX1B Mutations Cause Hereditary FSGS, Loin Pain Hematuria, Lupus, Systemic Lupus Erythematosis, Lupus Kidney Disease, Lupus Nephritis, Lyme Disease-Associated Glomerulonephritis, Malarial Nephropathy, Malignant Hypertension, Malakoplakia, Meatal Stenosis, Medullary Cystic Kidney Disease, Medullary Sponge Kidney, Megaureter, Melamine Toxicity and the Kidney, Membranoproliferative Glomerulonephritis, Membranous Nephropathy, MesoAmerican Nephropathy, Metabolic Acidosis, Metabolic Alkalosis, Microscopic Polyangiitis, Milk-alkalai syndrome, Minimal Change Disease, Multicystic dysplastic kidney, Multiple Myeloma, Myeloproliferative

Neoplasms and Glomerulopathy, Nail-patella Syndrome, Nephrocalcinosis, Nephrogenic Systemic Fibrosis, Nephroptosis (Floating Kidney, Renal Ptosis), Nephrotic Syndrome, Neurogenic Bladder, Nodular Glomerulosclerosis, Non-Gonococcal, Nutcracker syndrome, Orofaciodigital Syndrome, Orthostatic Hypotension, Orthostatic Proteinuria, Osmotic Diuresis, Page Kidney, Papillary Necrosis, Papillorenal Syndrome (Renal-Coloboma Syndrome, Isolated Renal Hypoplasia), The Peritoneal-Renal Syndrome, Posterior Urethral Valve, Post-infectious Glomerulonephritis, Post-streptococcal Glomerulonephritis, Polyarteritis Nodosa, Polycystic Kidney Disease, Posterior Urethral Valves, Preeclampsia, Proliferative Glomerulonephritis with Monoclonal IgG Deposits (Nasr Disease), Proteinuria (Protein in Urine), Pseudohyperaldosteronism, Pseudohypoparathyroidism, Pulmonary-Renal Syndrome, Pyelonephritis (Kidney Infection), Pyonephrosis, Radiation Nephropathy, Refeeding syndrome, Reflux Nephropathy, Rapidly Progressive Glomerulonephritis, Renal Abscess, Peripnephric Abscess, Renal Agenesis, Renal Artery Aneurysm, Renal Artery Stenosis, Renal Cell Cancer, Renal Cyst, Renal Hypouricemia with Exercise-induced Acute Renal Failure, Renal Infarction, Renal Osteodystrophy, Renal Tubular Acidosis, Reset Osmostat, Retrocaval Ureter, Retroperitoneal Fibrosis, Rhabdomyolysis, Rhabdomyolysis related to Bariatric Sugery, Rheumatoid Arthritis-Associated Renal Disease, Sarcoidosis Renal Disease, Salt Wasting, Renal and Cerebral, Schimke immuno-osseous dysplasia, Scleroderma Renal Crisis, Serpentine Fibula-Polycystic Kidney Syndrome, Exner Syndrome, Sickle Cell Nephropathy, Silica Exposure and Chronic Kidney Disease, Kidney Disease Following Hematopoietic Cell Transplantation, Kidney Disease Related to Stem Cell Transplantation, Thin Basement Membrane Disease, Benign Familial Hematuria, Trigonitis, Tuberous Sclerosis, Tubular Dysgenesis, Tumor Lysis Syndrome, Uremia, Uremic Optic Neuropathy, Ureterocele, Urethral Caruncle, Urethral Stricture, Urinary Incontinence, Urinary Tract Infection, Urinary Tract Obstruction, Vesicointestinal Fistula, Vesicoureteral Reflux, Von Hippel-Lindau Disease, Warfarin-Related Nephropathy, Wegener’s Granulomatosis, Granulomatosis with Polyangiitis, and Wunderlich syndrome.

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.

In some embodiments, AAV variants described herein may be useful for delivering gene therapy to gastrointestinal tissue (*e.g.*, tissue of the gastrointestinal tract). Accordingly, in some embodiments, AAV variants described herein may be useful for the treatment of gastrointestinal tract disorders. As used herein, a "gastrointestinal tract disorder" is a disease or condition of the gastrointestinal tract. A gastrointestinal disease may affect the mucosa (*e.g.*, epithelium, lamina propria, muscularis mucosae, *etc.*), submucosa (*e.g.*, submucous plexus, enteric nervous plexis, *etc.*), muscular layer of the gastrointestinal tract, the serosa and/or adventitia, oral cavity, esophagus, pylorus, stomach duodenum, small intestine, caecum, appendix, colon, anal canal, or rectum. A gastrointestinal tract disorder may be of a genetic origin, either inherited or acquired through a somatic mutation. Non-limiting examples of gastrointestinal tract diseases and disorders include but are not limited to: inflammatory bowel disease (IBD), Crohn's disease, ulcerative colitis, irritable bowel syndrome, Celiac disease, gastroesophageal reflux disease (GERD), achakasua, diverticulitus, diarrhea, and certain cancers (*e.g.*, bowel cancer, stomach cancer, colon cancer, rectal cancer, *etc.*).

In some embodiments, AAV variants described herein may be useful for delivering gene therapy to breast tissue (*e.g.*, tissue of the breast). Accordingly, in some embodiments, AAV variants described herein may be useful for the treatment of breast disorders. As used herein, a "breast disorder" is a disease or condition of the breast. A breast disease may affect the fibrous tissue, fatty tissue, lobules, or ducts of the breast. A breast disorder may be of a genetic origin, either inherited or acquired through a somatic mutation. Non-limiting examples of breast diseases and disorders include but are not limited to: mastitis, breast calcification, fat necrosis, fibroadenoma, fibrosis and simple cysts, galactorrhea, hyperplasia and breast cancer.

In some embodiments, AAV variants described herein may be useful for delivering gene therapy to pancreatic tissue (*e.g.*, tissue of the pancreas). Accordingly, in some embodiments, AAV variants described herein may be useful for the treatment of pancreatic disorders. As used herein, a “pancreatic disorder” is a disease or condition of the pancreas. A pancreatic disease may affect the head of the pancreas, neck of the pancreas, body of the pancreas, tail of the pancreas, pancreatic islets (*e.g.*, islets of Langerhans), acini, or columnar epithelium. A pancreatic disorder may be of a genetic origin, either inherited or acquired through a somatic mutation. Non-limiting examples of pancreatic diseases and disorders include but are not limited to: diabetes (*e.g.*, diabetes mellitus type 1 and diabetes mellitus type 2), pancreatitis (*e.g.*, acute pancreatitis, chronic pancreatitis), and pancreatic cancer.

In some embodiments, AAV variants described herein may be useful for delivering gene therapy to urinary tract tissue (*e.g.*, tissue of the urinary tract, such as bladder tissue). Accordingly, in some embodiments, AAV variants described herein may be useful for the treatment of urinary tract disorders. As used herein, a “urinary tract disorder” is a disease or condition of the urinary tract. A urinary tract disease may affect the bladder, ureters, urethera, or prostate. A urinary tract disorder may be of a genetic origin, either inherited or acquired through a somatic mutation. Non-limiting examples of urinary tract diseases and disorders include but are not limited to: urinary tract infections, kidney stones, bladder control problems (*e.g.*, urinary retention, urinary incontinence, *etc.*), cystitis, and bladder cancer.

In some embodiments, AAV variants described herein may be useful for delivering gene therapy to uterine tissue (*e.g.*, tissue of the uterus). Accordingly, in some embodiments, AAV variants described herein may be useful for the treatment of uterine disorders. As used herein, a “uterine disorder” is a disease or condition of the uterus. A uterine disease may affect the cervix, cervical canal, body of the uterus (fundus), endometrium, myometrium, or perimetrium. A uterine disorder may be of a genetic origin, either inherited or acquired through a somatic mutation. Non-limiting examples of uterine diseases and disorders include but are not limited to: adenomyosis, endometriosis, endometrial hyperplasia, Asherman’s syndrome, and endometrial 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 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.

Recombinant AAV 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, 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 an subject harboring the transgene. 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.

Recombinant AAV Vector: Transgene Coding Sequences

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. 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.

In some embodiments, the disclosure relates to an AAV comprising a nucleic acid encoding a protein or functional RNA useful for the treatment of a condition, disease or disorder associated with the central nervous system (CNS). The following is a non-limiting list of genes associated with CNS disease: DRD2, GRIA1, GRIA2,GRIN1, SLC1A1, SYP, SYT1, CHRNA7, 3Rtau/4rTUS, APP, BAX, BCL-2, GRIK1, GFAP, IL-1, AGER, associated with Alzheimer's Disease; UCH-L1, SKP1, EGLN1, Nurr-1, BDNF, TrkB, gstm1, S106β, associated with Parkinson's Disease; IT15, PRNP, JPH3, TBP, ATXN1, ATXN2, ATXN3, Atrophin 1, FTL, TITF-1, associated with Huntington's Disease; FXN, associated with Freidrich's ataxia; ASPA, associated with Canavan's Disease; DMD, associated with muscular dystrophy; and SMN1, UBE1, DYNC1H1 associated with spinal muscular atrophy. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more of the foregoing genes or fragments thereof. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more functional RNAs that inhibit expression of one or more of the foregoing genes.

In some embodiments, the disclosure relates to a nucleic acid encoding a protein or functional RNA useful for the treatment of a condition, disease or disorder associated with the cardiovascular system. The following is a non-limiting list of genes associated with cardiovascular disease: VEGF, FGF, SDF-1, connexin 40, connexin 43, SCN4a, HIF1α, SERCa2a, ADCY1, and ADCY6. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more of the foregoing genes or fragments thereof. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more functional RNAs that inhibit expression of one or more of the foregoing genes.

In some embodiments, the disclosure relates to an AAV comprising a nucleic acid encoding a protein or functional RNA useful for the treatment of a condition, disease or disorder associated with the pulmonary system. The following is a non-limiting list of genes associated with pulmonary disease: TNFα, TGFβ1, SFTPA1, SFTPA2, SFTPB, SFTPC, HPS1, HPS3, HPS4, ADTB3A, IL1A, IL1B, LTA, IL6, CXCR1, and CXCR2. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more of the foregoing genes or fragments thereof. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more functional RNAs that inhibit expression of one or more of the foregoing genes.

In some embodiments, the disclosure relates to an AAV comprising a nucleic acid encoding a protein or functional RNA useful for the treatment of a condition, disease or disorder associated with the liver. The following is a non-limiting list of genes associated with liver disease: α1-AT, HFE, ATP7B, fumarylacetoacetate hydrolase (FAH), glucose-6-phosphatase, NCAN, GCKR, LYPLAL1, and PNPLA3. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more of the foregoing genes or fragments thereof. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more functional RNAs that inhibit expression of one or more of the foregoing genes.

In some embodiments, the disclosure relates to an AAV comprising a nucleic acid encoding a protein or functional RNA useful for the treatment of a condition, disease or disorder associated with the kidney. The following is a non-limiting list of genes associated with kidney disease: PKD1, PKD2, PKHD1, NPHS1, NPHS2, PLCE1, CD2AP, LAMB2, TRPC6, WT1, LMX1B, SMARCAL1, COQ2, PDSS2, SCARB3, FN1, COL4A5, COL4A6, COL4A3, COL4A4, FOX1C, RET, UPK3A, BMP4, SIX2, CDC5L, USF2, ROBO2, SLIT2, EYA1, MYOG, SIX1, SIX5, FRAS1, FREM2, GATA3, KAL1, PAX2, TCF2, and SALL1. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more of the foregoing genes or fragments thereof. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more functional RNAs that inhibit expression of one or more of the foregoing genes.

In some embodiments, the disclosure relates to an AAV comprising a nucleic acid encoding a protein or functional RNA useful for the treatment of a condition, disease or disorder associated with the eye. The following is a non-limiting list of genes associated with ocular disease: CFH, C3, MT-ND2, ARMS2, TIMP3, CAMK4, FMN1, RHO, USH2A, RPGR, RP2, TMCO, SIX1, SIX6, LRP12, ZFPM2, TBK1, GALC, myocilin, CYP1B1, CAV1, CAV2, optineurin and CDKN2B. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more of the foregoing genes or fragments thereof. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more functional RNAs that inhibit expression of one or more of the foregoing genes.

In some embodiments, the disclosure relates to an AAV comprising a nucleic acid encoding a protein or functional RNA useful for the treatment of a condition, disease or disorder associated with breast. The following is a non-limiting list of genes associated with breast disease: BRCA1, BRCA2, Tp53, PTEN, HER2, BRAF, and PARP1. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more of the foregoing genes or fragments thereof. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more functional RNAs that inhibit expression of one or more of the foregoing genes.

In some embodiments, the disclosure relates to an AAV comprising a nucleic acid encoding a protein or functional RNA useful for the treatment of a condition, disease or disorder associated with the gastrointestinal tract. The following is a non-limiting list of genes associated with gastrointestinal disease: CYP2C19, CCL26, APC, IL12, IL10, and IL-18. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more of the foregoing genes or fragments thereof. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more functional RNAs that inhibit expression of one or more of the foregoing genes.

In some embodiments, the disclosure relates to an AAV comprising a nucleic acid encoding a protein or functional RNA useful for the treatment of a condition, disease or disorder associated with the pancreas. The following is a non-limiting list of genes associated with pancreatic disease: PRSS1, SPINK1, STK11, MLH1, KRAS2, p16, p53, and BRAF. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more of the foregoing genes or fragments thereof. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more functional RNAs that inhibit expression of one or more of the foregoing genes.

In some embodiments, the disclosure relates to an AAV comprising a nucleic acid encoding a protein or functional RNA useful for the treatment of a condition, disease or disorder associated with the urinary tract. The following is a non-limiting list of genes associated with urinary tract disease: HSPA1B, CXCR1 & 2, TLR2, TLR4, TGF-1, FGFR3, RB1, HRAS, TP53, and TSC1. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more of the foregoing genes or fragments thereof. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more functional RNAs that inhibit expression of one or more of the foregoing genes.

In some embodiments, the disclosure relates to an AAV comprising a nucleic acid encoding a protein or functional RNA useful for the treatment of a condition, disease or disorder associated with the uterus. The following is a non-limiting list of genes associated with ocular disease: DN-ER, MLH1, MSH2, MSH6, PMS1, and PMS2. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more of the foregoing genes or fragments thereof. In some embodiments, the disclosure relates to recombinant AAVs comprising nucleic acids that express one or more functional RNAs that inhibit expression of one or more of the foregoing genes.

The rAAVs of the disclosure can be used to restore the expression of genes that are reduced in expression, silenced, or otherwise dysfunctional in a subject (*e.g.*, a tumor suppressor that has been silenced in a subject having cancer). The rAAVs of the disclosure can also be used to knockdown the expression of genes that are aberrantly expressed in a subject (*e.g.*, an oncogene that is expressed in a subject having cancer). In some embodiments, an rAAV vector comprising a nucleic acid encoding a gene product associated with cancer (*e.g.*, tumor suppressors) may be used to treat the cancer, by administering a rAAV harboring the rAAV vector to a subject having the cancer. In some embodiments, an rAAV vector comprising a nucleic acid encoding a small interfering nucleic acid (*e.g.*, shRNAs, miRNAs) that inhibits the expression of a gene product associated with cancer (*e.g.*, oncogenes) may be used to treat the cancer, by administering a rAAV harboring the rAAV vector to a subject having the cancer. In some embodiments, an rAAV vector comprising a nucleic acid encoding a gene product associated with cancer (or a functional RNA that inhibits the expression of a gene associated with cancer) may be used for research purposes, *e.g.*, to study the cancer or to identify therapeutics that treat the cancer. The following is a non-limiting list of exemplary genes known to be associated with the development of cancer (*e.g.*, oncogenes and tumor suppressors): AARS, ABCB1, ABCC4, ABI2, ABL1, ABL2, ACK1, ACP2, ACY1, ADSL, AK1, AKR1C2, AKT1, ALB, ANPEP, ANXA5, ANXA7, AP2M1, APC, ARHGAP5, ARHGEF5, ARID4A, ASNS, ATF4, ATM, ATP5B, ATP5O, AXL, BARD1, BAX, BCL2, BHLHB2, BLMH, BRAF, BRCA1, BRCA2, BTK, CANX, CAP1, CAPN1, CAPNS1, CAV1, CBFB, CBLB, CCL2, CCND1, CCND2, CCND3, CCNE1, CCT5, CCYR61, CD24, CD44, CD59, CDC20, CDC25, CDC25A, CDC25B, CDC2L5, CDK10, CDK4, CDK5, CDK9, CDKL1, CDKN1A, CDKN1B, CDKN1C, CDKN2A, CDKN2B, CDKN2D, CEBPG, CENPC1, CGRRF1, CHAF1A, CIB1, CKMT1, CLK1, CLK2, CLK3, CLNS1A, CLTC, COL1A1, COL6A3, COX6C, COX7A2, CRAT, CRHR1, CSF1R, CSK, CSNK1G2, CTNNA1, CTNNB1, CTPS, CTSC, CTSD, CUL1, CYR61, DCC, DCN, DDX10, DEK, DHCR7, DHRS2, DHX8, DLG3, DVL1, DVL3, E2F1, E2F3, E2F5, EGFR, EGR1, EIF5, EPHA2, ERBB2, ERBB3, ERBB4, ERCC3, ETV1, ETV3, ETV6, F2R, FASTK, FBN1, FBN2, FES, FGFR1, FGR, FKBP8, FN1, FOS, FOSL1, FOSL2, FOXG1A, FOXO1A, FRAP1, FRZB, FTL, FZD2, FZD5, FZD9, G22P1, GAS6, GCN5L2, GDF15, GNA13, GNAS, GNB2, GNB2L1, GPR39, GRB2, GSK3A, GSPT1, GTF2I, HDAC1, HDGF, HMMR, HPRT1, HRB, HSPA4, HSPA5, HSPA8, HSPB1, HSPH1, HYAL1, HYOU1, ICAM1, ID1, ID2, IDUA, IER3, IFITM1, IGF1R, IGF2R, IGFBP3, IGFBP4, IGFBP5, IL1B, ILK, ING1, IRF3, ITGA3, ITGA6, ITGB4, JAK1, JARID1A, JUN, JUNB, JUND, K-ALPHA-1, KIT, KITLG, KLK10, KPNA2, KRAS2, KRT18, KRT2A, KRT9, LAMB1, LAMP2, LCK, LCN2, LEP, LITAF, LRPAP1, LTF, LYN, LZTR1, MADH1, MAP2K2, MAP3K8, MAPK12, MAPK13, MAPKAPK3, MAPRE1, MARS, MAS1, MCC, MCM2, MCM4, MDM2, MDM4, MET, MGST1, MICB, MLLT3, MME, MMP1, MMP14, MMP17, MMP2, MNDA, MSH2, MSH6, MT3, MYB, MYBL1, MYBL2, MYC, MYCL1, MYCN, MYD88, MYL9, MYLK, NEO1, NF1, NF2, NFKB1, NFKB2, NFSF7, NID, NINJ1, NMBR, NME1, NME2, NME3, NOTCH1, NOTCH2, NOTCH4,

NPM1, NQO1, NR1D1, NR2F1, NR2F6, NRAS, NRG1, NSEP1, OSM, PA2G4, PABPC1, PCNA, PCTK1, PCTK2, PCTK3, PDGFA, PDGFB, PDGFRA, PDPK1, PEA15, PFDN4, PFDN5, PGAM1, PHB, PIK3CA, PIK3CB, PIK3CG, PIM1, PKM2, PKMYT1, PLK2, PPARD, PPARG, PPIH, PPP1CA, PPP2R5A, PRDX2, PRDX4, PRKAR1A, PRKCBP1, PRNP, PRSS15, PSMA1, PTCH, PTEN, PTGS1, PTMA, PTN, PTPRN, RAB5A, RAC1, RAD50, RAF1, RALBP1, RAP1A, RARA, RARB, RASGRF1, RB1, RBBP4, RBL2, REA, REL, RELA, RELB, RET, RFC2, RGS19, RHOA, RHOB, RHOC, RHOD, RIPK1, RPN2, RPS6KB1, RRM1, SARS, SELENBP1, SEMA3C, SEMA4D, SEPP1, SERPINH1, SFN, SFPQ, SFRS7, SHB, SHH, SIAH2, SIVA, SIVA TP53, SKI, SKIL, SLC16A1, SLC1A4, SLC20A1, SMO, SMPD1, SNAI2, SND1, SNRPB2, SOCS1, SOCS3, SOD1, SORT1, SPINT2, SPRY2, SRC, SRPX, STAT1, STAT2, STAT3, STAT5B, STC1, TAF1, TBL3, TBRG4, TCF1, TCF7L2, TFAP2C, TFDP1, TFDP2, TGFA, TGFB1, TGFBI, TGFBR2, TGFBR3, THBS1, TIE, TIMP1, TIMP3, TJP1, TK1, TLE1, TNF, TNFRSF10A, TNFRSF10B, TNFRSF1A, TNFRSF1B, TNFRSF6, TNFSF7, TNK1, TOB1, TP53, TP53BP2, TP53I3, TP73, TPBG, TPT1, TRADD, TRAM1, TRRAP, TSG101, TUFM, TXNRD1, TYRO3, UBC, UBE2L6, UCHL1, USP7, VDAC1, VEGF, VHL, VIL2, WEE1, WNT1, WNT2, WNT2B, WNT3, WNT5A, WT1, XRCC1, YES1, YWHAB, YWHAZ, ZAP70, and ZNF9.

A rAAV vector may comprise as a transgene, a nucleic acid encoding a protein or functional RNA that modulates apoptosis. The following is a non-limiting list of genes associated with apoptosis and nucleic acids encoding the products of these genes and their homologues and encoding small interfering nucleic acids (*e.g.*, shRNAs, miRNAs) that inhibit the expression of these genes and their homologues are useful as transgenes in certain embodiments of the disclosure: RPS27A, ABL1, AKT1, APAF1, BAD, BAG1, BAG3, BAG4, BAK1, BAX, BCL10, BCL2, BCL2A1, BCL2L1, BCL2L10, BCL2L11, BCL2L12, BCL2L13, BCL2L2, BCLAF1, BFAR, BID, BIK, NAIP, BIRC2, BIRC3, XIAP, BIRC5, BIRC6, BIRC7, BIRC8, BNIP1, BNIP2, BNIP3, BNIP3L, BOK, BRAF, CARD10, CARD11, NLRC4, CARD14, NOD2, NOD1, CARD6, CARD8, CARD9, CASP1, CASP10, CASP14, CASP2, CASP3, CASP4, CASP5, CASP6, CASP7, CASP8, CASP9, CFLAR, CIDEA, CIDEB, CRADD, DAPK1, DAPK2, DFFA, DFFB, FADD, GADD45A, GDNF, HRK, IGF1R, LTA, LTBR, MCL1, NOL3, PYCARD, RIPK1, RIPK2, TNF, TNFRSF10A, TNFRSF10B, TNFRSF10C, TNFRSF10D, TNFRSF11B, TNFRSF12A, TNFRSF14,

TNFRSF19, TNFRSF1A, TNFRSF1B, TNFRSF21, TNFRSF25, CD40, FAS, TNFRSF6B, CD27, TNFRSF9, TNFSF10, TNFSF14, TNFSF18, CD40LG, FASLG, CD70, TNFSF8, TNFSF9, TP53, TP53BP2, TP73, TP63, TRADD, TRAF1, TRAF2, TRAF3, TRAF4, TRAF5 DRD2, GRIA1, GRIA2,GRIN1, SLC1A1, SYP, SYT1, CHRNA7, 3Rtau/4rTUS, APP, BAX, BCL-2, GRIK1, GFAP, IL-1, AGER, UCH-L1, SKP1, EGLN1, Nurr-1, BDNF, TrkB,gstm1, S106β, IT15, PRNP, JPH3, TBP, ATXN1, ATXN2, ATXN3, Atrophin 1, FTL, TITF-1, FXN, ASPA, DMD, and SMN1, UBE1, DYNC1H1.

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.

Somatic Transgenic Animal Models Produced Using rAAV-Based Gene Transfer

The disclosure also relates to the production of somatic transgenic animal models of disease using recombinant Adeno-Associated Virus (rAAV) based methods. The methods are based, at least in part, on the observation that AAV serotypes and variants thereof mediate efficient and stable gene transfer in a tissue specific manner in adult animals. The rAAV elements (capsid, promoter, transgene products) are combined to achieve somatic transgenic animal models that express a stable transgene in a time and tissue specific manner. The somatic transgenic animal produced by the methods of the disclosure can serve as useful models of human disease, pathological state, and/or to characterize the effects of gene for which the function (*e.g.*, tissue specific, disease role) is unknown or not fully understood. For example, an animal (*e.g.*, mouse) can be infected at a distinct developmental stage (*e.g.*, age) with a rAAV comprising a capsid having a specific tissue targeting capability (*e.g.*, liver, heart, pancreas) and a transgene having a tissue specific promoter driving expression of a gene involved in disease. Upon infection, the rAAV infects distinct cells of the target tissue and produces the product of the transgene.

In some embodiments, the sequence of the coding region of a transgene is modified. The modification may alter the function of the product encoded by the transgene. The effect of the modification can then be studied *in vivo* by generating a somatic transgenic animal model using the methods disclosed herein. In some embodiments, modification of the sequence of coding region is a nonsense mutation that results in a fragment (*e.g.*, a truncated version). In other cases, the modification is a missense mutation that results in an amino acid substitution. Other modifications are possible and will be apparent to the skilled artisan.

In some embodiments, the transgene causes a pathological state. A transgene that causes a pathological state is a gene whose product has a role in a disease or disorder (*e.g.*, causes the disease or disorder, makes the animal susceptible to the disease or disorder) and/or may induce the disease or disorder in the animal. The animal can then be observed to evaluate any number of aspects of the disease (*e.g.*, progression, response to treatment, *etc.*). These examples are not meant to be limiting, other aspects and examples are disclosed herein and described in more detail below.

The disclosure in some aspects, provide methods for producing somatic transgenic animal models through the targeted destruction of specific cell types. For example, models of type 1 diabetes can be produced by the targeted destruction of pancreatic Beta-islets. In other examples, the targeted destruction of specific cell types can be used to evaluate the role of specific cell types on human disease. In this regard, transgenes that encode cellular toxins (*e.g.*, diphtheria toxin A (DTA)) or pro-apoptotic genes (NTR, Box, *etc.*) can be useful as transgenes for functional ablation of specific cell types. Other exemplary transgenes, whose products kill cells are embraced by the methods disclosed herein and will be apparent to one of ordinary skill in the art.

The disclosure, in some aspects, provides methods for producing somatic transgenic animal models to study the long-term effects of over-expression or knockdown of genes. The long term over expression or knockdown (*e.g.*, by shRNA, miRNA, miRNA inhibitor, *etc.*) of genes in specific target tissues can disturb normal metabolic balance and establish a pathological state, thereby producing an animal model of a disease, such as, for example, cancer. The disclosure in some aspects, provides methods for producing somatic transgenic animal models to study the long-term effects of over-expression or knockdown of gene of potential oncogenes and other genes to study tumorigenesis and gene function in the targeted tissues. Useful transgene products include proteins that are known to be associated with cancer and small interfering nucleic acids inhibiting the expression of such proteins. Other suitable transgenes may be readily selected by one of skill in the art provided that they are useful for creating animal models of tissue-specific pathological state and/or disease.

Recombinant AAV Administration Methods

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, 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, 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).

Kits and Related Compositions

The agents described herein may, in some embodiments, be assembled into pharmaceutical or diagnostic or research kits to facilitate their use in therapeutic, diagnostic or research applications. A kit may include one or more containers housing the components of the disclosure and instructions for use. Specifically, such kits may include one or more agents described herein, along with instructions describing the intended application and the proper use of these agents. In certain embodiments agents in a kit may be in a pharmaceutical formulation and dosage suitable for a particular application and for a method of administration of the agents. Kits for research purposes may contain the components in appropriate concentrations or quantities for running various experiments.

The kit may be designed to facilitate use of the methods described herein by researchers and can take many forms. Each of the compositions of the kit, where applicable, may be provided in liquid form (*e.g.*, in solution), or in solid form, (*e.g.*, a dry powder). In certain cases, some of the compositions may be constitutable or otherwise processable (*e.g.*, to an active form), for example, by the addition of a suitable solvent or other species (for example, water or a cell culture medium), which may or may not be provided with the kit. As used herein, “instructions” can define a component of instruction and/or promotion, and typically involve written instructions on or associated with packaging of the disclosure. Instructions also can include any oral or electronic instructions provided in any manner such that a user will clearly recognize that the instructions are to be associated with the kit, for example, audiovisual (*e.g.*, videotape, DVD, *etc.*), Internet, and/or web-based communications, *etc.* The written instructions may be in a form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which instructions can also reflects approval by the agency of manufacture, use or sale for animal administration.

The kit may contain any one or more of the components described herein in one or more containers. As an example, in one embodiment, the kit may include instructions for mixing one or more components of the kit and/or isolating and mixing a sample and applying to a subject. The kit may include a container housing agents described herein. The agents may be in the form of a liquid, gel or solid (powder). The agents may be prepared sterilely, packaged in syringe and shipped refrigerated. Alternatively it may be housed in a vial or other container for storage. A second container may have other agents prepared sterilely. Alternatively the kit may include the active agents premixed and shipped in a syringe, vial, tube, or other container. The kit may have one or more or all of the components required to administer the agents to an animal, such as a syringe, topical application devices, or iv needle tubing and bag, particularly in the case of the kits for producing specific somatic animal models.

The kit may have a variety of forms, such as a blister pouch, a shrink wrapped pouch, a vacuum sealable pouch, a sealable thermoformed tray, or a similar pouch or tray form, with the accessories loosely packed within the pouch, one or more tubes, containers, a box or a bag. The kit may be sterilized after the accessories are added, thereby allowing the individual accessories in the container to be otherwise unwrapped. The kits can be sterilized using any appropriate sterilization techniques, such as radiation sterilization, heat sterilization, or other sterilization methods known in the art. The kit may also include other components, depending on the specific application, for example, containers, cell media, salts, buffers, reagents, syringes, needles, a fabric, such as gauze, for applying or removing a disinfecting agent, disposable gloves, a support for the agents prior to administration *etc.*

The instructions included within the kit may involve methods for detecting a latent AAV in a cell. In addition, kits of the disclosure may include, instructions, a negative and/or positive control, containers, diluents and buffers for the sample, sample preparation tubes and a printed or electronic table of reference AAV sequence for sequence comparisons.

EXAMPLES

Example 1: Isolation Of Transcriptionally Active, Novel AAV Capsid Sequences With Desired Tissue Tropisms and Properties From Human Tissues.

This example describes novel AAV capsid sequences isolated by the following steps: 1) PCR amplification of wtAAV genomes present in normal and diseased human tissues; 2) high-throughput single-molecule, real-time (SMRT) sequencing of PCR amplicon libraries; 3) variant identification/profiling by bioinformatic analyses; and 4) the selection of high-confidence ORFs that can be translated into full length capsid proteins. Schematic depictions of workflows used in this example are shown in FIGs. 1A-1B.

This approach exploits the natural pool of genomic diversity observed among viral genomes isolated from both normal and tumor tissues. Conceptually, *in vivo* tissues act as natural incubators for viral genomic diversity through selective pressure and/or immune evasion. Thus, the discovery of inter- and intra-tissue variability, as well as inter-patient diversity benefit from methods that are able to profile the full spectrum of AAV variants found among tissues and organs of human origin.

PCR Amplification of AAV genomes from human tissues

To isolate diverse AAV variants with the potential for identifying new serotypes with unique tropisms, 844 human surgical specimens from 455 patients were collected from West China Hospital, Sichuan University, Chengdu, China. These tissues encompass a wide-range of tissue/organ types, as well as various tumors types (Table 1). In particular, AAV variants were identified from nine normal liver tissues, 7 liver tumors, four enlarged prostate tissues, two normal lung tissues, one pancreatic tumor tissue, one breast cancer tissue, one normal breast tissue, one gastric cancer tissue, one normal gastric tissue, one brain tissue and one glioma sample.

Total genomic DNA was extracted from human tissues and subjected to PCR amplification of AAV capsid sequence. PCR primers used in this example are described in Table 2. Briefly, either panAAV primers for the amplification of 4.1-kb AAV rep-cap sequence (*e.g.*, RepF318, AV2cas), or panAAV primers for amplification of 2.3-kb AAV cap sequence (*e.g.*, CapF, CapR) were used for PCR.

Table 1: Clinical specimens for wtAAV genome amplification

Organ	**Tissue quantity**	
	Normal tissue	**Tumor tissue**
Liver	100	101
Brain	4	50

Gastric	37	37
Lung	100	100
Breast	52	57
Pancreatic	NA	45
Rectal	50	50
Prostate	34	NA
Urologic	3	12
Cervical	2	10
Sum	**378**	**466**

Table 2: PCR primer sequences

Primer	**Sequence (5'-3')**	**SEQ ID NO:**
RepF318	GCCATGCCGGGGTTCTACGAGAT	872
AV2cas	ACAGGAGACCAAAGTTCAACTGAAACGA	873
CapF	GACTGCATCTTTGAACAATAAATGA	874
CapR	GAAACGAATTAACCGGTTTATTGATTAA	875

High-throughput sequencing of AAV PCR products and bioinformatics analysis

AAV PCR products were subjected to high-throughput single-molecule, real-time (SMRT) sequencing. This approach removes the need to perform viral genome reconstruction and chimera prediction from aligned short-read fragments obtained from other conventional high-throughput genome sequencing methodologies.

Using variant analysis pipelines developed from open source bioinformatic tools more than 600 previously undescribed, high-confidence AAV2, AAV2/3 hybrid, and AAV8 capsid sequence variants were identified. Specifically, 224 AAV8 variants (harboring 1 to 10 single amino acid variants); 425 AAV2 variants (harboring 1 to 20 single amino acid variants); and 194 AAV2/3 hybrid variants (harboring 10 to 50 single amino acid variants) were identified. Tables 3, 4 and 5 summarize the unique capsid protein variants. For purposes of comparison, wild-type AAV2, AAV3, and AAV8 capsid amino acid sequences are described in SEQ ID NOs: 869, 870, and 871, respectively. FIG. 7 is a scatter plot displaying the distribution of distinct AAV2 capsid variants and AAV2/3 variants harboring one or more single-amino-acid variants.

Table 3: Unique AAV2 and AAV2/3 hybrid variants (amino acid sequences) identified by SMRT sequencing and bioinformatics analyses.

Unique AAV2 variants					
Sample Source	Patient No.	Size of DNA (kb)	Unique variants (a.a.)	SEQ ID NOs:	Total unique variants (a.a.)
Liver	7927N	2.3 kb (cap)	85	325-409	**409**
Liver Tumor	37HCC		3	322-324	
Breast	18B		26	118-143	
Breast cancer	19B		21	211-231	
Lung	18L		55	144-198	
Prostate	5		24	1-24	
	17		12	106-117	
	18		12	199-210	
	27		90	232-321	
Pancreatic cancer	10		81	25-105	
Liver	1178N	4.1 kb (rep+cap)	4	410-414; 837-840	**16**
	9955N		3	429-434; 850-852	
Liver tumor	9955C		9	415-428; 841-849	
Unique AAV2/3 variants					
Sample Source	Patient No.	Size of DNA (kb)	Unique variants (a.a.)		Total unique variants (a.a.)
Liver	42	2.3 kb (cap)	6	512-517	**194**
	74		11	543-553	
Liver Tumor	37HCC		6	506-511	
	65		4	539-542	
	7449C		15	554-568	
Breast	18B		23	435-457	
Breast cancer	19B		44	462-505	
Prostate	5		60	569-628	
	17				
	18		4	458-461	
Gastric cancer	17G		N/A (420 in DNA)	-	
Gastric	50G		21	518-538	

DNA sequences are provided for 4.1 kb libraries.

Table 4: Unique AAV8 variants (amino acid sequences) identified by SMRT sequencing and bioinformatics analyses.

Sample Source	Patient No.	Size of DNA(kb)	Unique variants (a.a.)	SEQ ID NOs:	Total unique variants(a.a.)
Liver	0067N	2.3kb(cap)	12	647-658	**208**
	3522N		73	674-746	
	5110N		3	747-749	
	7427N		6	750-755	
Liver Tumor	0067C		9	638-646	
	7803C		9	756-764	
	8818C		63	765-827	
Brain	G5		9	828-836	
Glioma	2236		14	659-672	
Lung	24		10	629-637; 673	

Table 5: Additional AAV8 variant capsid proteins

AAV8 Variant Name	SEQ ID NO:
B1	853
B2	854
B3	855
B4	856
B12	857
B18	858
B24	859
B41	860
B44	861
B45	862
B46	863
B60	864
B61	865
B62	866
B63	867
B64	868

Example 2: Identification of AAV8 variants with improved *in vivo* tropism.

A subset of candidate AAV8 variants (*e.g.*, B2, B3, B44 and B61) were cloned into AAV packaging vectors by standard molecular cloning methods, and packaged with luciferase reporter genes driven by the CB6 promoter. Produced vectors were injected into mice and *in vivo* levels of luciferase transgene expression were analyzed by whole animal imaging and quantification of relative luminescence. It was observed that the B2 (SEQ ID NO: 854) and B3 (SEQ ID NO: 855) variants have higher expression in liver after intramuscular injection (FIGs. 2A-2D), while after IV injection in neonatal mice, the B61 (SEQ ID NO: 865) variant has higher transduction efficiencies in brain and spinal cord compared to AAV9 (FIGs 3A-3B). This is notable since wild-type AAV8 has been observed to cross the blood brain barrier less than AAV9. One AAV8 variant, B44 (SEQ ID NO: 861) has better ability transduced to liver after IM injection compared to AAV8 (FIGs. 4A-4B).

Phylogenic analysis was performed to compare AAV8 capsid variants B2, B3, B44, and B61 to other AAV serotypes. Briefly, amino acid sequences of AAV8 variants were aligned with other published AAV sequences using ClustalW and Phylogenetic trees were inferred using the Minimum Evolution method in MEGA6.06. Results of the bioinformatics analysis indicate that B2, B3, B44, and B61 sequences are related to Clade E [AAV8] capsid proteins. FIG. 5. Representative amino acid substitutions in AAV8 variants are shown in Table 6.

Table 6: Representative amino acid substitutions in AAV8 variants relative to wild-type AAV8

AAV Variant	Representative Substitutions (relative to wt AAV8)
B2	E63G
B3	K259R
B44	L91Q, T234A, M374T
B61	M374T, M561V

Example 3: *In vitro* assessment of rAAV genome packaging efficiency and initial characterization of candidate capsid variants.

Molecular cloning of packaging plasmid constructs containing selected AAV capsid variants

AAV2 and AAV2/3 hybrid capsid variants identified by SMRT sequencing are cloned into packaging plasmids by replacing the conventional viral capsid genes using standard molecular cloning strategies (*e.g.*, site-directed mutagenesis of parental AAV2 or AAV2/3 capsid expression plasmids, PCR-based cloning and Gibson Assembly, or synthesized by outsourcing). FIG. 8 shows vector constructs to be used in multiplexed screening of discovered capsid variants. A summary of the proposed transgene cassettes to be used for various diagnostic strategies is shown in Table 7.

Table 7: Transgene cassettes for various diagnostic strategies

Promoter	**Transgene**	**Reporter/therapeutic gene analysis**
CMV enhancer Chicken β-actin	EGFP	Tissue or cell-type specific transduction efficiency
CMV enhancer Chicken β-actin	Luciferase	Whole-animal tropism profiling and individual tissue quantification
Thyroxin binding globulin	Factor IX	Liver-specific transduction of secreted factors. Pre-clinical testing

Multiplex assessment of packaging efficiency by high-throughput small-scale vector production and titration for vector genomes

Quantification of vector genomes of rAAV in crude lysate is used to directly test rAAV variant packaging efficiency of both first-generation (single-strand AAV) and second-generation (self-complementary AAV) vectors directly following triple-transfection of HEK 293 packaging cells. This provides a streamlined alternative to performing the full workflow for small-scale vector production followed by silver staining and conventional PCR titration of vector genomes to assess virus quality for all discovered variants. Since this method can be scaled for performance in 96-well formats, it us used to quickly identify variants that produce high titer vectors.

Serological evaluation of novel AAV variants

Candidate variants with high packaging efficiency are screened for antibody cross-reactivity to current AAVs by standard means, such as capsid immunology assays to test novel rAAVs against serum from AAV- immunized rabbits. In addition, pooled human IgG

(IVIG) neutralizing assays are performed for each candidate variant to determine the potential for pre-existing humoral immunity in the human population.

Example 4: *In vivo* analyses of rAAV2 and rAAV2/3 variants to study vector transduction biology, prevalence of pathotoxicity, tissue/organ tropism, and bio-distribution profiles.

Mouse studies

Candidate capsid variants are grouped based on tissue distribution, and prioritized by organs of interest. Groups of candidate variants are subjected to clustered-indexing (FIG. 6A), whereby multiple packaging plasmids expressing candidate capsid variants are mixed and expressed to package uniquely DNA barcoded transgenes by triple-transfection (*e.g.*, F9 coagulation factor IX (F.IX), to assess liver targeting and expression efficacy of secreted factors; EGFP, to assess bio-distribution and the extent of tissue-specific transduction via organ/tissue sectioning and comparative immunofluorescence microscopy; or Luciferase (Luc), to assess the quality of CNS and liver transduction via live animal imaging

For studies that gauge the capacity of rAAV variants for liver-targeted transgene expression and secretion, rAAV constructs comprising the thyroxine-binding globulin (TGB), a liver specific promoter, are designed. For studies that profile whole-animal vector transduction, constructs comprising the CMV-enhancer, chicken β-actin promoter (CB6) regulatory cassette are designed.

Vectors encapsidating indexed transgenes are injected into adult and newborn mice by different routes of administration, and screened for secreted F.IX expression, EGFP expression, or Luc expression in 1-month longitudinal studies to profile AAV variant-mediated transgene expression. Routes of administration for the CNS/brain include peripheral intravascular (IV, to test transduction across the blood-brain barrier), intracerebroventricular (ICV), intraparenchymal, and intrathecal. Administration for retina is performed via subretinal injection. In some embodiments, IV injections also target the liver.

Animals that exhibit unique transgene expression compared to control animals (*e.g.*, transgenes delivered by AAV2, AAV2/3, or AAV8) are sacrificed and harvested for organs. Individual organs are assayed for the presence and abundance of barcoded transgenes by conventional PCR amplification of bulk DNA extracts or cDNA libraries containing transgene message, followed by Illumina sequencing to trace barcoded transgenes enriched in each tissue. FIG. 9 outlines the general design strategy for transgene indexing. The abundance and tissue/organ distribution of detectable barcoded transgenes reflects candidate rAAV variant tropism and transduction efficacy of each group. Highly efficacious candidate groups with desirable vector properties are selected. Individual candidate variants from selected groups are used to package barcoded transgenes for a second round of screening for the purpose of identifying individual, highly efficacious variants. Clustered-indexing can be carried out iteratively in multiple rounds of hierarchical selection to reduce the workload.

Non-human primate (NHP) studies

Candidate rAAV variants are screened for bio-distribution in non-human primates by a modality similar to the clustered-indexing methodology outlined for mouse studies (FIG. 6B). The transduction efficiencies to target organs via different routes of administration are re-assessed in NHPs to validate rAAV variant profiles observed in precursor mouse studies.

Immunogenicity, prevalence of neutralization antibody in human populations, capacity for genotoxicity, and general aspects of pathogenicity are gauged alongside primary assessments, for example histopathology of multiple tissues and organs to scrutinize T-cell or neutrophil infiltrates, monitoring hepatotoxicity by ALT/AST activity, and analyzing inflammation by examination of histological sections, to determine transduction profiles in non-human primate (NHP) animals.

Example 5: Isolation Of Novel AAV Capsid Sequences.

Additional AAV capsid sequences were isolated. Using variant analysis pipelines developed from bioinformatic tools, an additional 263 previously undescribed, high-confidence AAV2 and AAV2/3 hybrid capsid sequence variants were identified. For purposes of comparison, wild-type AAV2 and AAV3 capsid amino acid sequences are described in SEQ ID NOs: 869 and 870, respectively.

Table 8: Additional unique AAV2 and AAV2/3 hybrid variants (amino acid sequences) identified by SMRT sequencing and bioinformatics analyses.

Unique AAV2 variants				
Sample Source	Size of DNA (kb)	Unique variants (a.a.)	SEQ ID NOs (aa):	Total unique variants (a.a.)
Breast Cancer	2.2 kb	8	1726-1733	**89**

Gastric Tumor		15	1734-1748	
Glioma		2	1749-1750	
Liver		25	1751-1775	
Liver Tumor		36	1776-1811	
Lung Tumor		3	1812-1814	
Unique AAV2/3 variants				
Sample Source	Size of DNA (kb)	Unique variants (a.a.)		Total unique variants (a.a.)
Breast Cancer	2.2 kb	18	1815-1832	**174**
Gastric		17	1833-1849	
Liver		117	1850-1966	
Liver Tumor		22	1967-1988	

The corresponding DNA sequences are provided for all libraries. The nucleic acid sequences for the AAV2 capsid variants correspond to SEQ ID NOs: 1989-2077. The nucleic acid sequences for the AAV2/3 capsid variants correspond to SEQ ID NOs: 2078-2251.

This disclosure is not limited in its application to the details of construction and the arrangement of components set forth in this description or illustrated in the drawings. The disclosure is capable of other embodiments and of being practiced or of being carried out in various ways. Also, the phraseology and terminology used herein is for the purpose of description and should not be regarded as limiting. The use of “including,” “comprising,” or “having,” “containing,” “involving,” and variations thereof herein, is meant to encompass the items listed thereafter and equivalents thereof as well as additional items.

Having thus described several aspects of at least one embodiment of this disclosure, it is to be appreciated various alterations, modifications, and improvements will readily occur to those skilled in the art. Such alterations, modifications, and improvements are intended to be part of this disclosure, and are intended to be within the spirit and scope of the disclosure. Accordingly, the foregoing description and drawings are by way of example only.

CLAIMS

What is claimed is:

1. A recombinant expression vector comprising a nucleic acid encoding a polypeptide having a sequence selected from the group consisting of: SEQ ID NO: 1 to 409, 435-868, or 1726-1988, or a fragment thereof that does not encode a peptide that is identical to a sequence of any one of SEQ ID NOs: 869, 870, or 871.

2. An isolated AAV capsid protein comprising an amino acid sequence selected from the group consisting of: SEQ ID NOs: 1 to 409, 435-868 and 1726-1988, or fragment thereof.

3. An isolated AAV capsid protein comprising a sequence selected from SEQ ID NOs: 1-409, 837-852, or 1726-1814, wherein an amino acid of the sequence that is not identical to a corresponding amino acid of the sequence set forth as SEQ ID NO: 869 is replaced with a conservative substitution.

4. An isolated AAV capsid protein comprising a sequence selected from SEQ ID NOs: 435-628 or 1815-1988, wherein an amino acid of the sequence that is not identical to a corresponding amino acid of the sequence set forth as SEQ ID NO: 869 or 870 is replaced with a conservative substitution.

5. An isolated AAV capsid protein comprising a sequence selected from SEQ ID NOs: 629-836 or 853-868, wherein an amino acid of the sequence that is not identical to a corresponding amino acid of the sequence set forth as SEQ ID NO: 871 is replaced with a conservative substitution.

6. A peptide fragment of the isolated AAV capsid protein of any one of claims 2 to 5 that is not identical to a sequence of any one of SEQ ID NOs: 869, 870, or 871.

7. An isolated AAV capsid protein comprising the peptide fragment of claim 6.

8. An recombinant expression vector comprising a nucleic acid sequence encoding the isolated AAV capsid protein of any one of claims 2 to 5.

9. A composition comprising the isolated AAV capsid protein of any one of claims 2 to 5.

10. A composition comprising the isolated AAV capsid protein of any one of claims 2 to 5 and a pharmaceutically acceptable carrier.

11. A recombinant AAV (rAAV) comprising the isolated AAV capsid protein of any one of claims 2 to 5.

12. A composition comprising the recombinant rAAV of claim 11.

13. The composition of claim 12 further comprising a pharmaceutically acceptable carrier.

14. A host cell containing a nucleic acid that comprises a coding sequence of a polypeptide selected from the group consisting of: SEQ ID NO: 1-409, 435-868, and 1726-1988, that is operably linked to a promoter.

15. A composition comprising the host cell of claim 14 and a sterile cell culture medium.

16. A composition comprising the host cell of claim 15 and a cryopreservative.

17. A method for delivering a transgene to a subject comprising

administering a rAAV of claim 11 to a subject, wherein the rAAV comprises at least one transgene, and wherein the rAAV infects cells of a target tissue of the subject.

18. A method for generating a somatic transgenic animal model comprising administering a recombinant rAAV of claim 11 to a non-human animal, wherein the rAAV comprises at least one transgene, and wherein the rAAV infects cells of a target tissue of the non-human animal.

19. The method of claim 17, wherein the at least one transgene is a protein coding gene.

20. The method of claim 17, wherein the at least one transgene encodes a small interfering nucleic acid.

21. The method of claim 20, wherein the small interfering nucleic acid is a miRNA.

22. The method of claim 20, wherein the small interfering nucleic acid is a miRNA sponge or TuD RNA that inhibits the activity of at least one miRNA in the subject or animal.

23. The method of claim 22, wherein the miRNA is expressed in a cell of the target tissue

24. The method of claim of claim 17, wherein the target tissue is liver, central nervous system (CNS), ocular, gastrointestinal, respiratory, breast, pancreas, urinary tract, or uterine tissue.

25. The method of claim 18, wherein the transgene expresses a transcript that comprises at least one binding site for a miRNA, wherein the miRNA inhibits activity of the transgene, in a tissue other than the target tissue, by hybridizing to the binding site.

26. A method for generating a somatic transgenic animal model comprising administering a rAAV of claim 23 to a non-human animal, wherein the rAAV comprises at least one transgene, wherein the transgene expresses a transcript that comprises at least one

binding site for a miRNA, wherein the miRNA inhibits activity of the transgene, in a tissue other than a target tissue, by hybridizing to the binding site of the transcript.

27. The method of claim 26, wherein the transgene comprises a tissue specific promoter or inducible promoter.

28. The method of claim 27, wherein the tissue specific promoter is a liver-specific thyroxin binding globulin (TBG) promoter, an insulin promoter, a glucagon promoter, a somatostatin promoter, mucin-2 promoter, a pancreatic polypeptide (PPY) promoter, a synapsin-1 (Syn) promoter, a retinoschisin promoter, a K12 promoter, a CC10 promoter, a surfactant protein C (SP-C) promoter, a PRC1 promoter, a RRM2 promoter, uroplakin 2 (UPII) promoter, or a lactoferrin promoter.

29. The method of claim 17, wherein the rAAV is administered intravenously, transdermally, intraocularly, intrathecally, orally, intramuscularly, subcutaneously, intranasally, or by inhalation.

30. The method of claim 17, wherein the subject is selected from a mouse, a rat, a rabbit, a dog, a cat, a sheep, a pig, and a non-human primate.

31. The method of claim 17, wherein the subject is a human.

32. A somatic transgenic animal model produced by the method of claim 18.

33. A kit for producing a rAAV, the kit comprising:

a container housing an isolated nucleic acid encoding a polypeptide having a sequence of any one of SEQ ID NO: 1 to 409, 435-868, or 1726-1988.

34. The kit of claim 33 further comprising instructions for producing the rAAV.

35. The kit of claim 34 further comprising at least one container housing a recombinant AAV vector, wherein the recombinant AAV vector comprises a transgene.

36. A kit comprising:

a container housing a recombinant AAV having an isolated AAV capsid protein of any one of claims 2 to 5.

37. The kit of claim 36, wherein the container is a syringe.

38. The isolated AAV capsid protein of any one of claims 2-5, wherein the capsid protein is a VP1 capsid protein.

39. The isolated AAV capsid protein of any one of claims 2-5, wherein the capsid protein is a VP2 capsid protein.

40. The isolated capsid protein of any one of claims 2-5, wherein the capsid protein is a VP3 capsid protein.

41. A pseudotyped AAV comprising the capsid protein of any one of claims 2 to 5, or 7.

42. The recombinant expression vector of claim 1, wherein the nucleic acid encodes a V1 capsid protein.

43. The recombinant expression vector of claim 1, wherein the nucleic acid encodes a V2 capsid protein.

44. The recombinant expression vector of claim 1, wherein the nucleic acid encodes a V3 capsid protein.

45. A nucleic acid comprising a sequence selected from the group consisting of SEQ ID NO: 410-434, 876-1718, and 1989-2251.

46. The nucleic acid of claim 45, wherein the nucleic acid is engineered to express an AAV capsid protein, or a variant thereof and/or an AAV assembly-activating protein (AAP), or a variant thereof.

47. The nucleic acid of claim 45, wherein the AAP is in a different open reading frame of the nucleic acid than the AAV capsid protein.

48. The nucleic acid of claim 46, wherein the AAP is AAV2 AAP (AAP-2), or variant thereof.

49. An isolated protein encoded by the nucleic acid of claim 45.

50. A recombinant AAV comprising the isolated protein of claim 49.

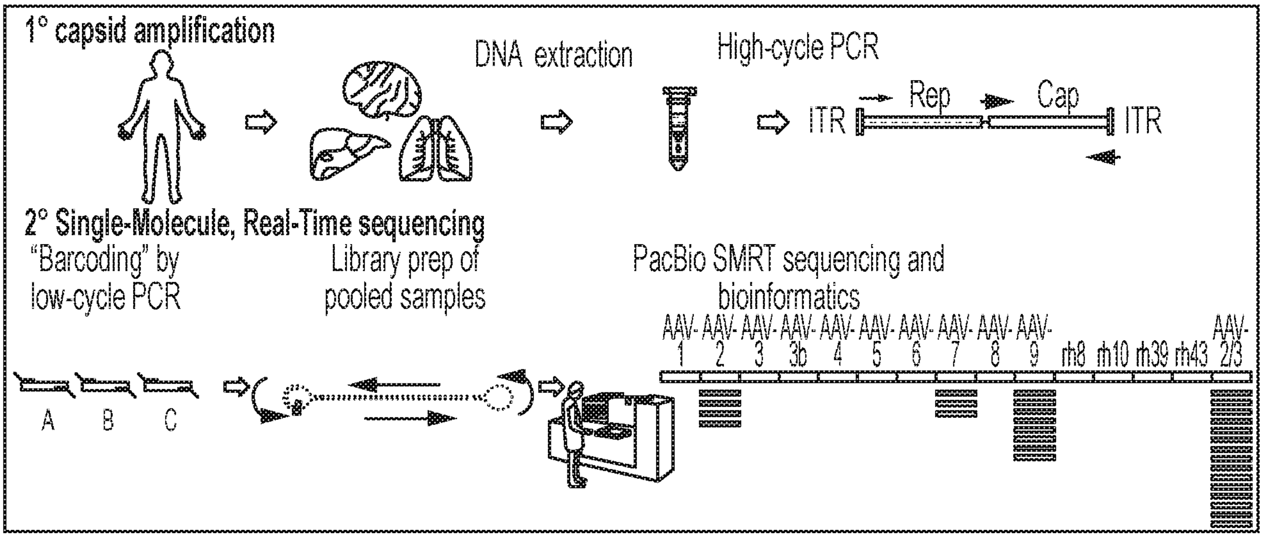

1° capsid amplification
DNA extraction
High-cycle PCR
ITR
Rep
Cap
ITR
2° Single-Molecule, Real-Time sequencing
"Barcoding" by low-cycle PCR
Library prep of pooled samples
PacBio SMRT sequencing and bioinformatics
A
B
C
AAV-1
AAV-2
AAV-3
AAV-3b
AAV-4
AAV-5
AAV-6
AAV-7
AAV-8
AAV-9
rh8
rh10
rh39
rh43
AAV-2/3

FIG. 1A

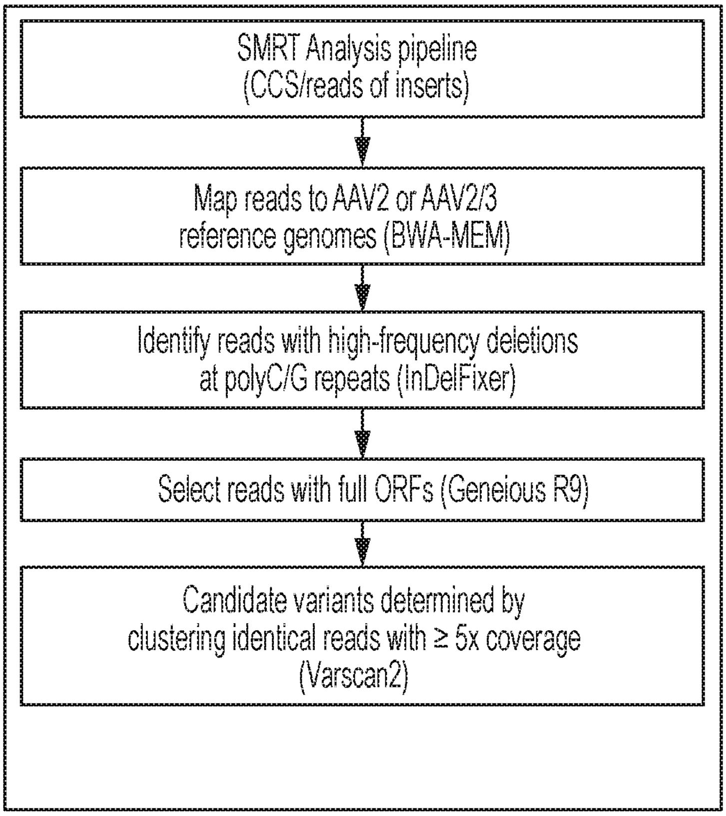

SMRT Analysis pipeline (CCS/reads of inserts)
Map reads to AAV2 or AAV2/3 reference genomes (BWA-MEM)
Identify reads with high-frequency deletions at polyC/G repeats (InDelFixer)
Select reads with full ORFs (Geneious R9)
Candidate variants determined by clustering identical reads with ≥ 5x coverage (Varscan2)

FIG. 1B

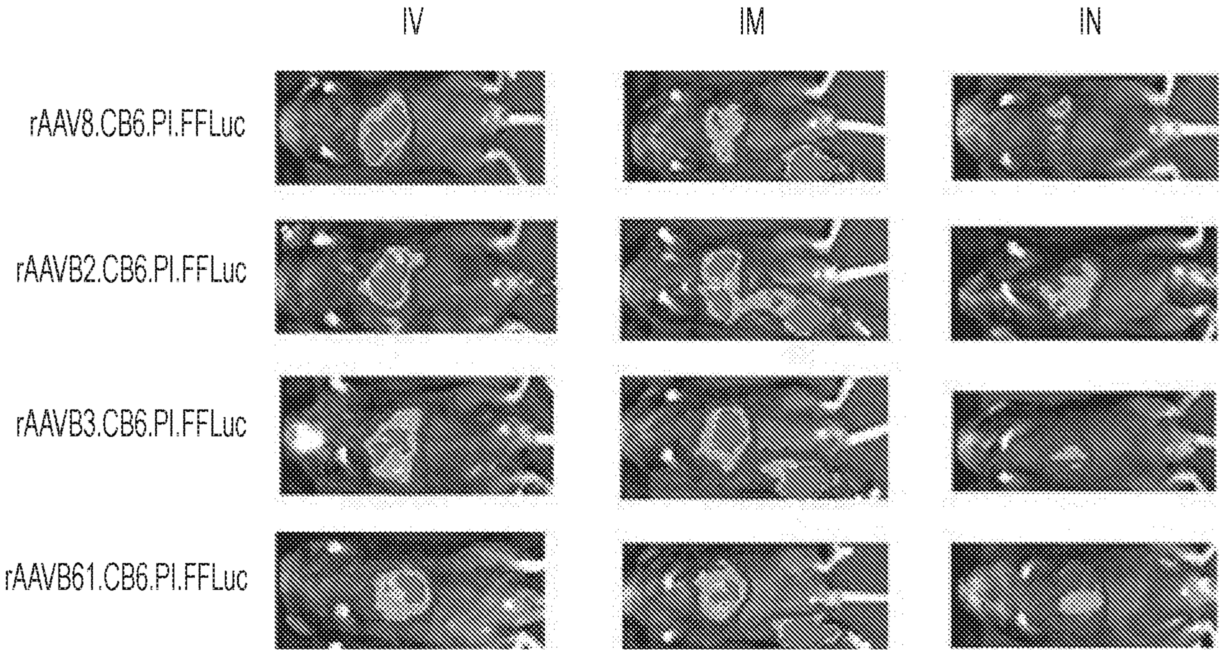

IV
IM
IN
rAAV8.CB6.PI.FFLuc
rAAVB2.CB6.PI.FFLuc
rAAVB3.CB6.PI.FFLuc
rAAVB61.CB6.PI.FFLuc

FIG. 2A

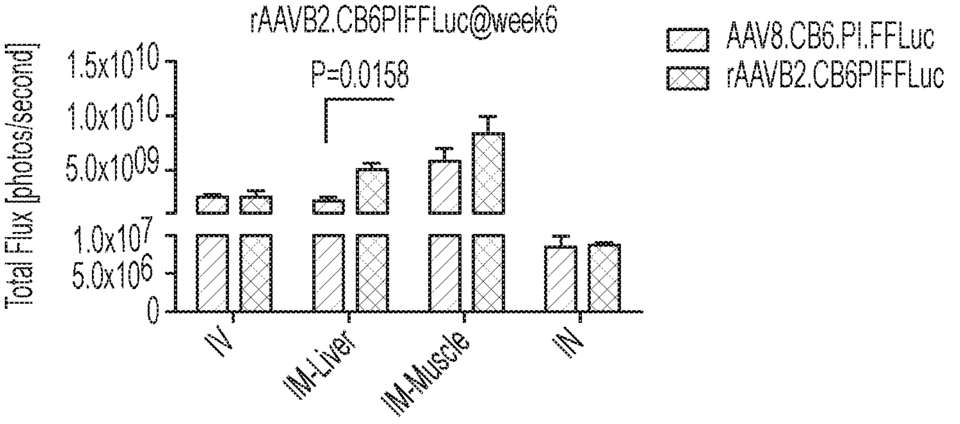


FIG. 2B

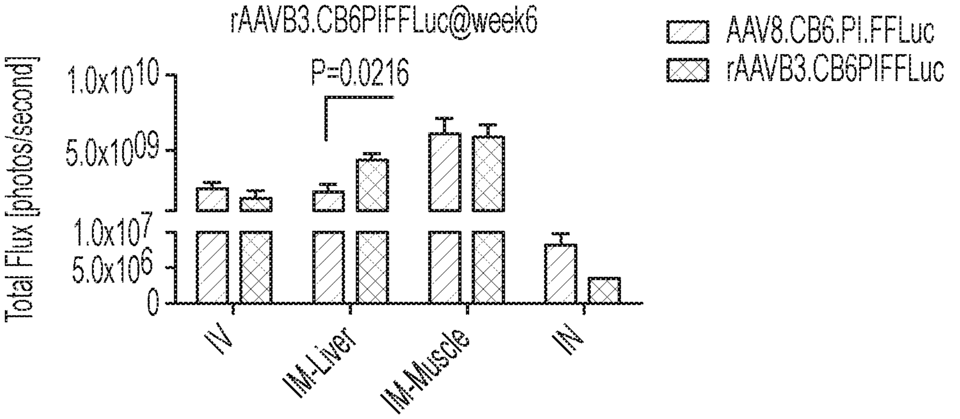


FIG. 2C

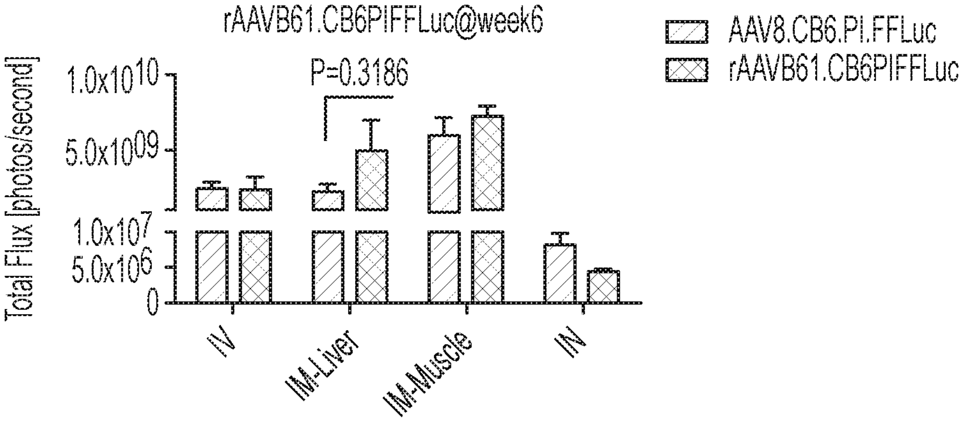


FIG. 2D

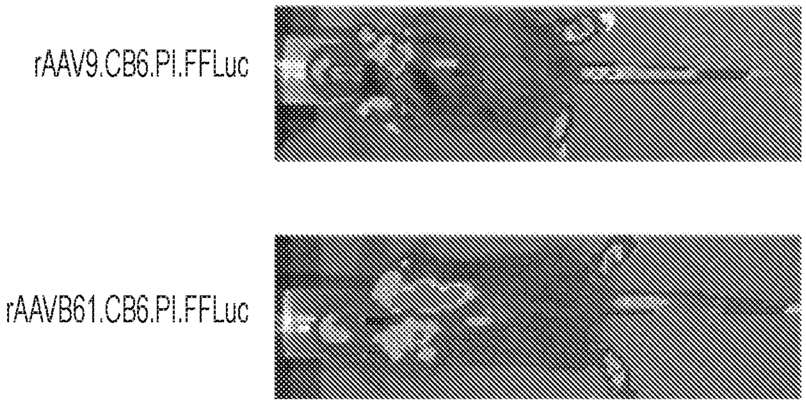

rAAV9.CB6.PI.FFLuc
rAAVB61.CB6.PI.FFLuc

FIG. 3A

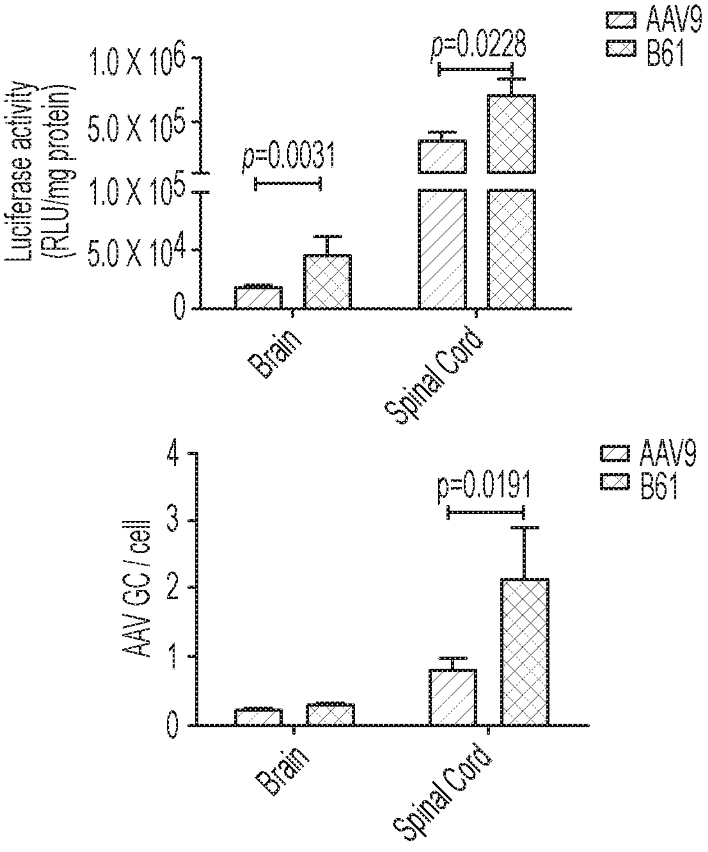

Luciferase activity (RLU/mg protein)
1.0 X 10^6
5.0 X 10^5
1.0 X 10^5
5.0 X 10^4
0
p=0.0031
p=0.0228
Brain
Spinal Cord
AAV9
B61
AAV GC / cell
4
3
2
1
0
p=0.0191
Brain
Spinal Cord
AAV9
B61

FIG. 3B

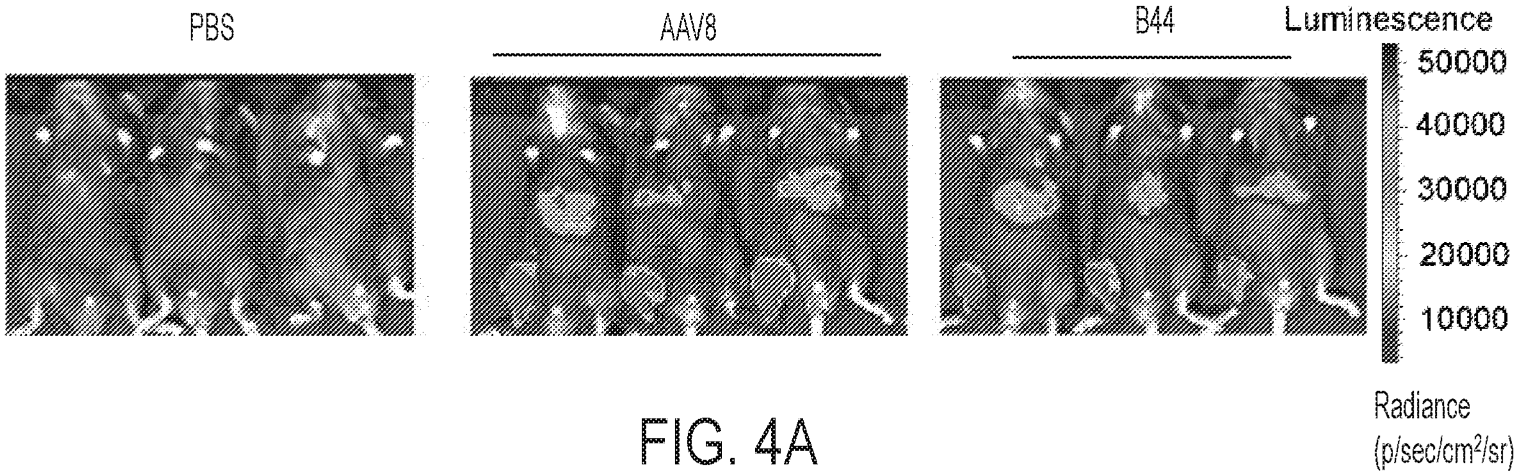

PBS
AAV8
B44
Luminescence
50000
40000
30000
20000
10000
Radiance
(p/sec/cm²/sr)

FIG. 4A

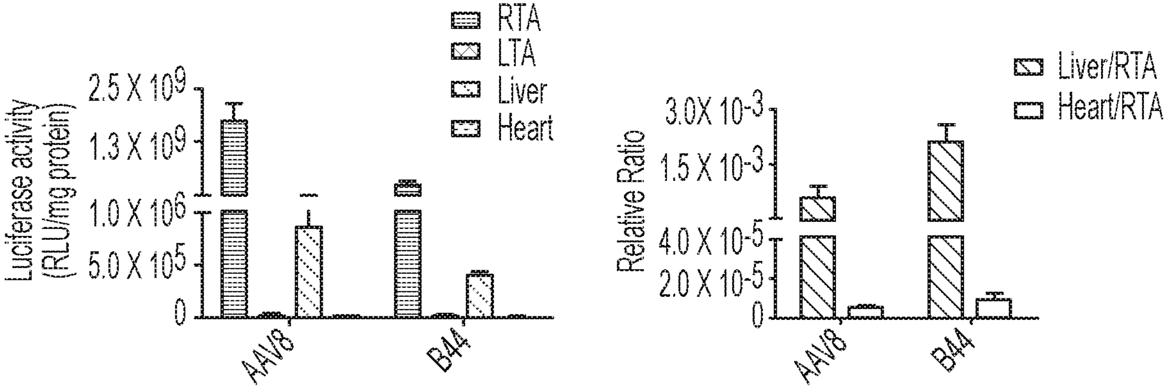

Luciferase activity
(RLU/mg protein)
2.5 X 10⁹
1.3 X 10⁹
1.0 X 10⁶
5.0 X 10⁵
0
AAV8
B44
RTA
LTA
Liver
Heart
Relative Ratio
3.0X 10-3
1.5 X 10-3
4.0 X 10-5
2.0 X 10-5
0
AAV8
B44
Liver/RTA
Heart/RTA

FIG. 4B

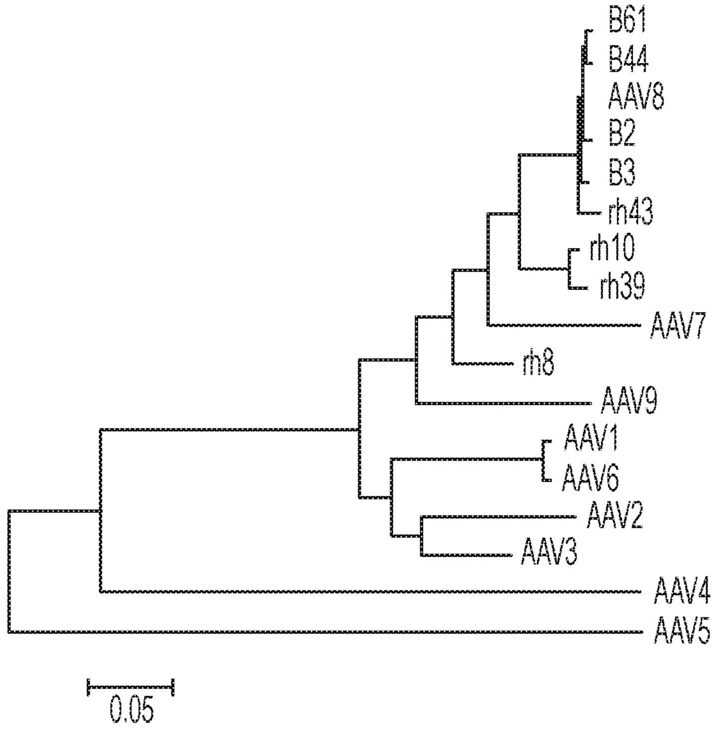

B61
B44
AAV8
B2
B3
rh43
rh10
rh39
AAV7
rh8
AAV9
AAV1
AAV6
AAV2
AAV3
AAV4
AAV5
0.05

FIG. 5

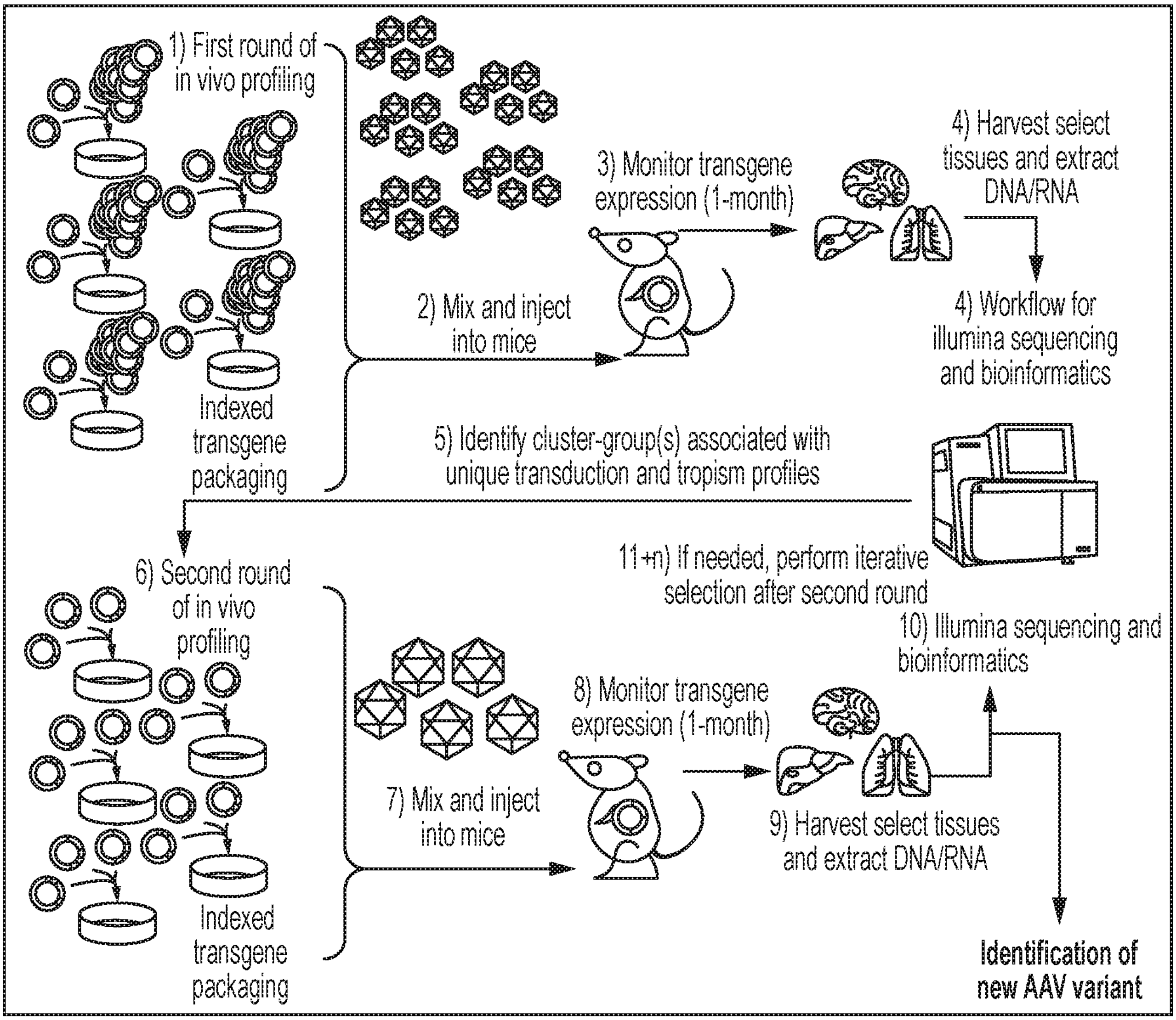

1) First round of in vivo profiling
Indexed transgene packaging
2) Mix and inject into mice
3) Monitor transgene expression (1-month)
4) Harvest select tissues and extract DNA/RNA
4) Workflow for illumina sequencing and bioinformatics
5) Identify cluster-group(s) associated with unique transduction and tropism profiles
6) Second round of in vivo profiling
Indexed transgene packaging
7) Mix and inject into mice
8) Monitor transgene expression (1-month)
9) Harvest select tissues and extract DNA/RNA
10) Illumina sequencing and bioinformatics
11+n) If needed, perform iterative selection after second round
Identification of new AAV variant

FIG. 6A

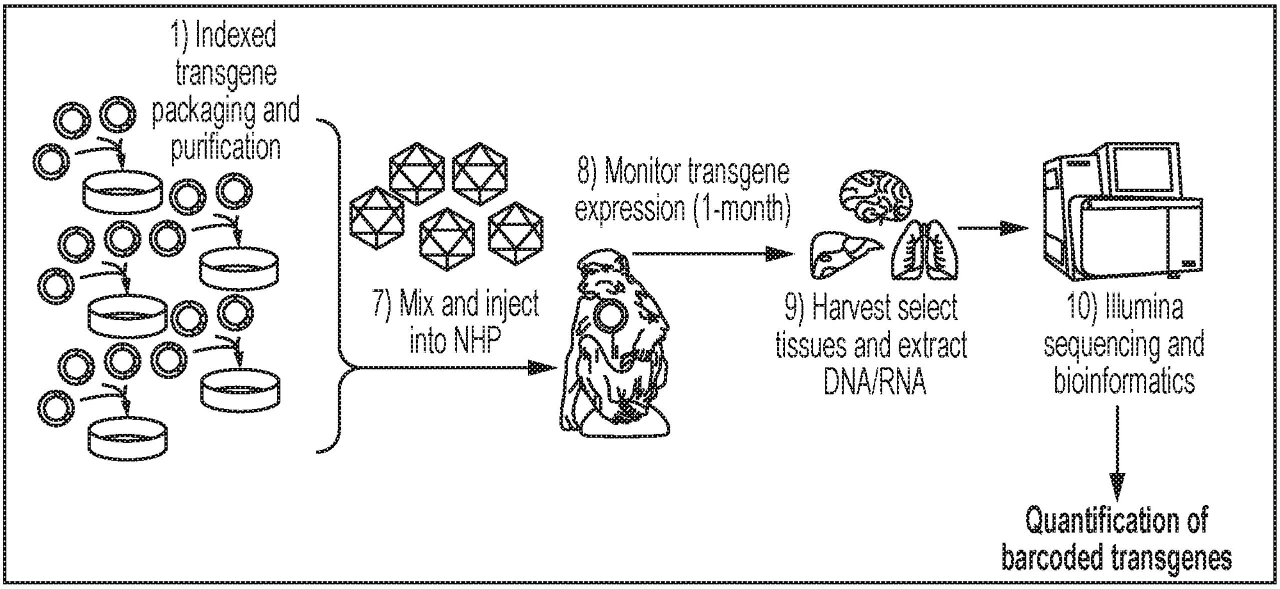
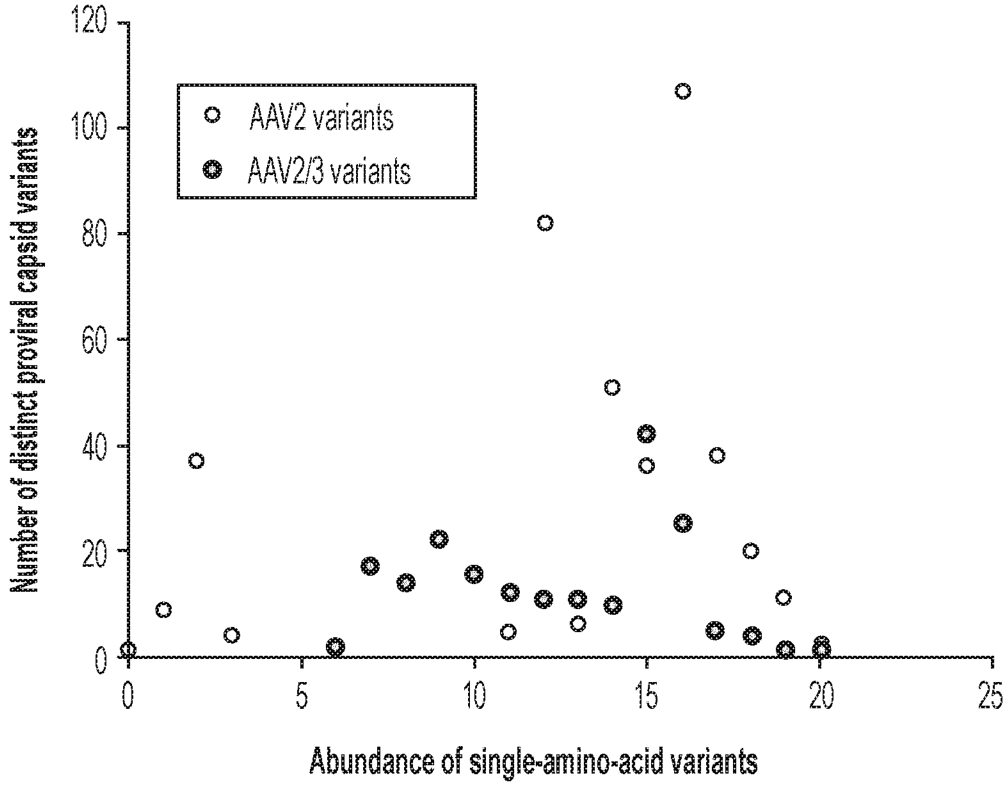

1) Indexed transgene packaging and purification
7) Mix and inject into NHP
8) Monitor transgene expression (1-month)
9) Harvest select tissues and extract DNA/RNA
10) Illumina sequencing and bioinformatics
Quantification of barcoded transgenes

FIG. 6B



120
100
80
60
40
20
0
0
5
10
15
20
25
Number of distinct proviral capsid variants
Abundance of single-amino-acid variants
○ AAV2 variants
◍ AAV2/3 variants

FIG. 7

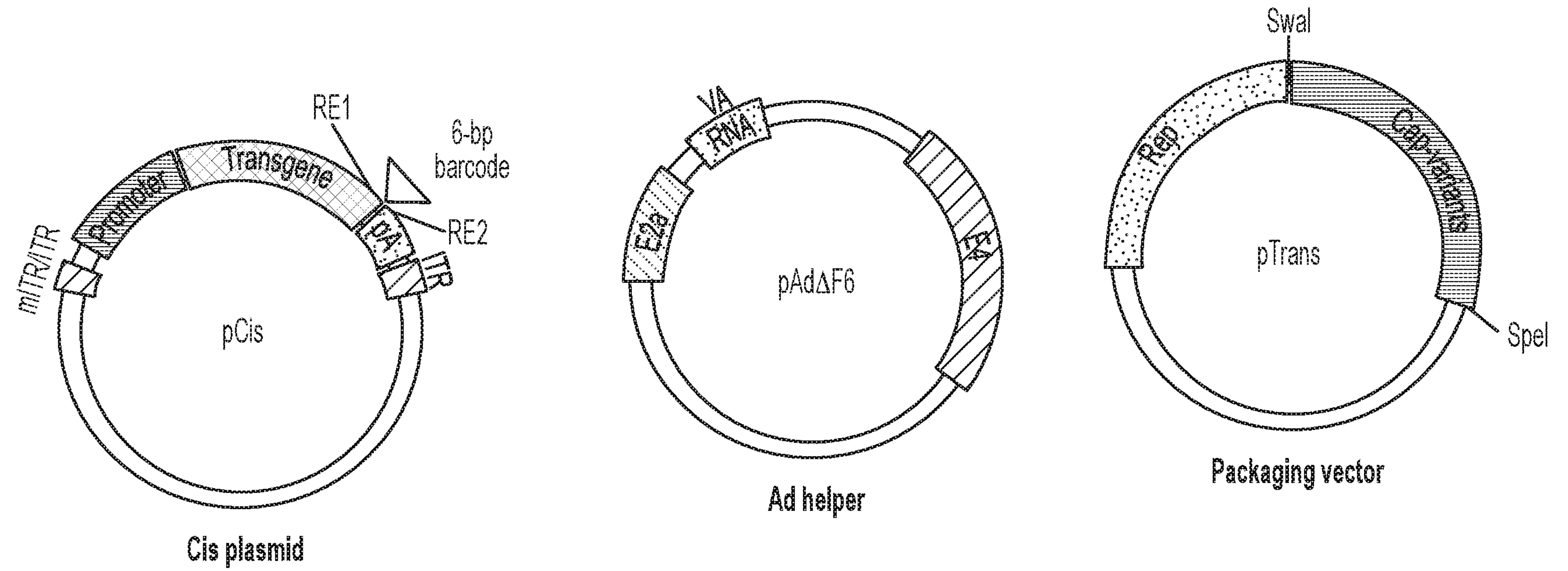

m/TR/ITR
Promoter
Transgene
RE1
6-bp
barcode
RE2
pA
ITR
pCis
Cis plasmid
VA
RNA
E2a
E4
pAdΔF6
Ad helper
SwaI
Rep
Cap-variants
pTrans
SpeI
Packaging vector

FIG. 8

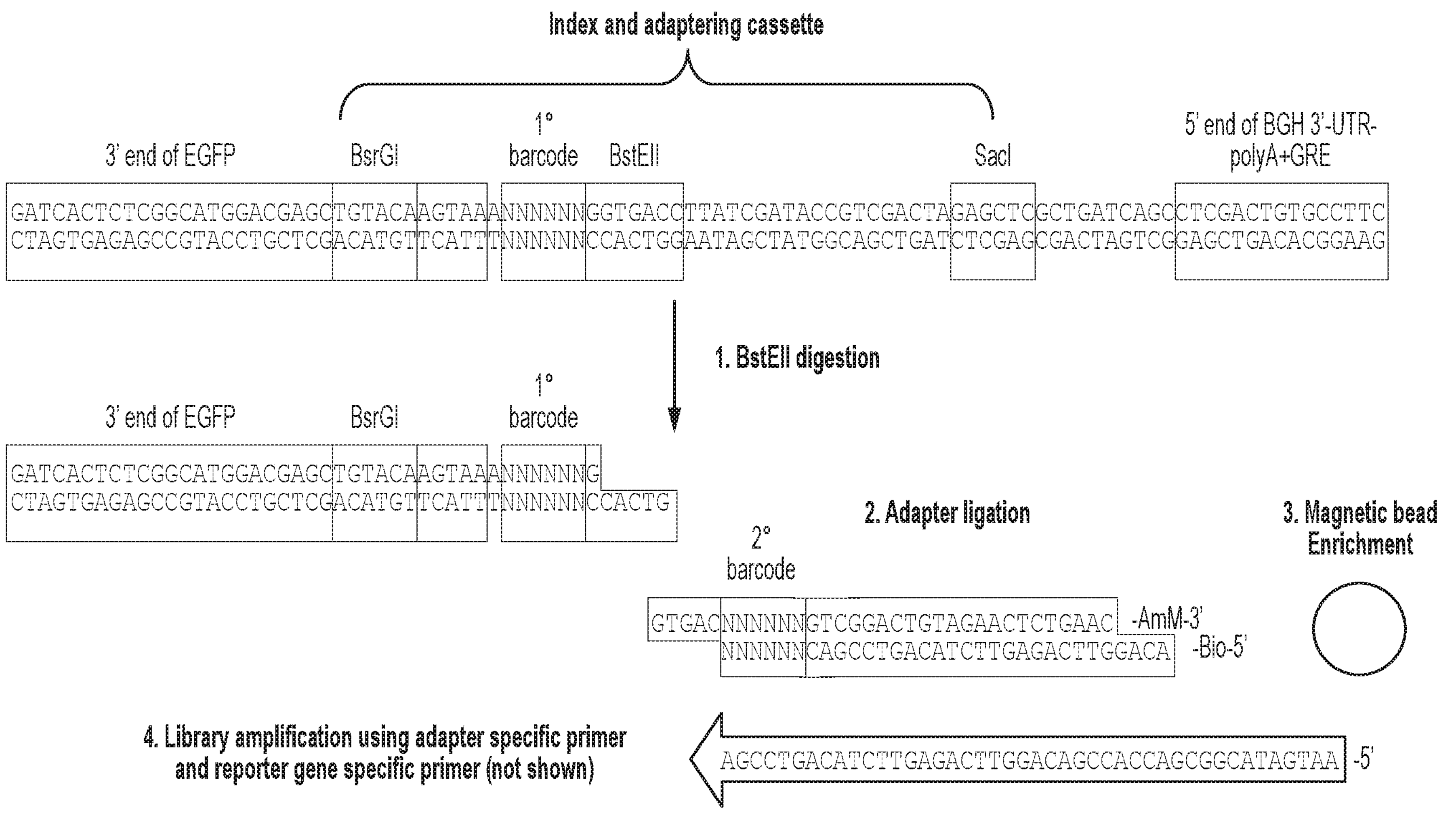

Index and adaptering cassette
3' end of EGFP
BsrGI
1°
barcode
BstEII
SacI
5' end of BGH 3'-UTR-
polyA+GRE
GATCACTCTCGGCATGGACGAGCTGTACAAGTAAANNNNNNGGTGACCTTATCGATACCGTCGACTAGAGCTCGCTGATCAGCCTCGACTGTGCCTTC
CTAGTGAGAGCCGTACCTGCTCGACATGTTCATTTNNNNNNCCACTGGAATAGCTATGGCAGCTGATCTCGAGCGACTAGTCGGAGCTGACACGGAAG
1. BstEII digestion
3' end of EGFP
BsrGI
1°
barcode
GATCACTCTCGGCATGGACGAGCTGTACAAGTAAANNNNNNG
CTAGTGAGAGCCGTACCTGCTCGACATGTTCATTTNNNNNNCCACTG
2. Adapter ligation
2°
barcode
GTGACNNNNNNGTCGGACTGTAGAACTCTGAAC
-AmM-3'
NNNNNNCAGCCTGACATCTTGAGACTTGGACA
-Bio-5'
3. Magnetic bead
Enrichment
4. Library amplification using adapter specific primer
and reporter gene specific primer (not shown)
AGCCTGACATCTTGAGACTTGGACAGCCACCAGCGGCATAGTAA
-5'

FIG. 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 17/56614

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12N 15/861; C07K 14/005, 14/015 (2018.01)
CPC - C12N 2750/14111, 2750/14122, 2750/14143, 2750/14141, 15/86, 15/8645; C07K 14/005, 14/015

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
See Search History Document

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- A	WO 2015/121501 A1 (King's College, London, et al.) 20 August 2015 (20.08.2015). Especially pg 2 para 3, pg 8 para 5, pg 9 para 1, pg 17 para 2, pg 32 para 3, pg 39 para 4, pg 41 para 1, pg 43 para 1.	1, 2, (6-13)/2, (36-40)/2, 42-44 --- 3, (6-13)/3, 14-16, 33-35, (36-40)/3
A	WO 2016/065001 A1 (University of Massachusetts) 28 April 2016 (28.04.2016). Especially pg 2 ln 21 continued to pg 3 ln 7, claims 1-16	1-3, (6-13)(2,3), 14-16, 33 -35, (36-40)/(2,3), 42-44
A	US 2007/036760 A1 (Wilson et al.) 15 February 2007 (15.02.2007). Especially para [0013], [0014], sheets 2-29 Fig 2A-Z, AA, AB	1-3, (6-13)(2,3), 14-16, 33 -35, (36-40)/(2,3), 42-44

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:
"A"	document defining the general state of the art which is not considered to be of particular relevance
"E"	earlier application or patent but published on or after the international filing date
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O"	document referring to an oral disclosure, use, exhibition or other means
"P"	document published prior to the international filing date but later than the priority date claimed
"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&"	document member of the same patent family

Date of the actual completion of the international search 04 Janaury 2018	Date of mailing of the international search report 0 9 MAR 2018
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 17/56614

Box No. I **Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)**

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. ☒ forming part of the international application as filed:

      ☒ in the form of an Annex C/ST.25 text file.

      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☐ furnished subsequent to the international filing date for the purposes of international search only:

      ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 17/56614

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

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

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 41
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

This International Searching Authority found multiple inventions in this international application, as follows:
------go to Extra Sheet for continuation-------

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-3, (6-13)(in part), 14-16, 33-37, (36-40)(in part), 42-44 limited to SEQ ID NO: 1

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

----continued from previous sheet-----

As to common technical feature #2, King's College teaches an isolated capsid AAV capsid protein, or an isolated AAV capsid protein having a conservative amino acid substitution (pg 4 para 5; "In another aspect, the present invention provides an isolated variant AAV capsid protein, wherein the variant AAV capsid protein comprises at least one amino acid substitution with respect to a wild type AAV capsid protein; wherein the at least one amino acid substitution is present at one or more of the following positions in an AAV2 capsid protein sequence: 125, 151, 162, 205, 312, 457, 492, 499, 533, 546, 548, 585, 588 and/or 593; or at one or more corresponding positions in an alternative AAV capsid protein sequence"; claim 4: amino acid mutations compared to the wild-type AAV2 [conservative mutations]: V125I, A162S, T205S, S492A, E499D, F533Y, and A593S")

As to common technical feature #3, King's College teaches a vector encoding a nucleic acid encoding an AAV capsid protein (claim 1), or a host comprising said nucleic acid (pg 8 para 7; "In another aspect, the present invention provides an isolated host cell comprising a nucleic acid as defined above") operably linked to a promoter (pg 32 para 2; "The term "promoters" or "promoter" as used herein can refer to a DNA sequence that is located adjacent to a DNA sequence that encodes a recombinant product").

As to common technical feature #4, concerning a kit for producing rAAV, King's College does not teach a kit. However, one of ordinary skill in the art would have known how to assemble the claimed kit and provide instructions for its use, with the aim of simplifying administration of the claimed composition.

As to common technical feature #5, King's College teaches (pg 40 para 3; "Assembly associated protein (AAP) is the product of an open reading frame within the cap gene, and may also be required for packaging").

As the common technical features were known in the art at the time of the invention, they cannot be considered common special technical features that would otherwise unify the groups. The inventions lack unity with one another.

Therefore, Groups I+ and II and III+ lack unity of invention under PCT Rule 13 because they do not share a same or corresponding special technical feature.

Note concerning item 4: Claim 41 is a multiple dependent claim and is not drafted according to the second and third sentences of PCT Rule 6.4(a).

continuation of Box III: Observations where unity of invention is lacking

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

Group I+: Claims 1-16, 33-40, 42-44 drawn to an isolated adeno-associated virus capsid protein, a recombinant vector comprising a nucleic acid encoding it, a recombinant AAV (rAAV; virus) comprising the capsid protein, or a kit comprising the capsid protein.
The isolated AAV capsid protein or vector comprising a nucleic acid encoding it will be searched to the capsid protein is the first named AAV capsid protein SEQ ID NO: 1. It is believed that claims 1-3, (6-13)(in part), 14-16, 33-37, (38-40)(in part), 42-44 read on this first named invention and thus these claims will be searched without fee to the extent that they encompass SEQ ID NO: 1. Additional AAV capsid proteins will be searched upon payment of additional fees. Applicant must specify the claims that encompass any additional elected AAV capsid proteins. Applicants must further indicate, if applicable, the claims which read on the first named invention if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined. An exemplary election would be: SEQ ID NO: 435 (claims 1, 2, 4, (6-13)(in part), 14-16, 33-37, (38-40)(in part), 42-44).

Group II: Claims 17-32, drawn to a method involving administrating a recombinant adeno-associated virus (rAAV)

.

Group III+, Claims 45-50, drawn to a nucleic acid sequence.
Group III+ will be searched upon payment of additional fee(s). The composition may be searched, for example, to the extent that the nucleic acid sequence encompasses SEQ ID NO: 1989, for an additional fee and election as such. It is believed that claims 45-50 read on this exemplary invention. Additional nucleic acid sequences will be searched upon the payment of additional fees. Applicants must indicate, if applicable, which claims read on this named invention if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the '+' group(s) will result in only the first named invention to be searched/examined. An exemplary election would be the the nucleic acid sequence SEQ ID NO: 2251 (claims 45-50).

The inventions listed as Groups I+, II, III+ do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Technical Features:

Group I+ has the special technical feature an isolated AAV capsid protein, a recombinant vector comprising a nucleic acid encoding it, an rAAV comprising the capsid protein, or a kit comprising the capsid protein, not required by Group II.

Group II has the special technical feature of a method involving administrating a recombinant rAAV.

Group III+ has the special technical feature of a specified nucleic acid sequence, where the sequence may encode an AAV-assembly activation protein, not required by Groups I+ or II.

No technical features are shared between the amino acid sequences of AAV capsid protein of Groups I+ and, accordingly, these groups lack unity a priori.

No technical features are shared between the nucleic acid sequences of Group III+, and, accordingly, these groups lack unity a priori.

Additionally, even if Groups I+, II and III+ were considered to share the technical features of:
1. a recombinant adeno-associated virus (rAAV).
2. an isolated capsid AAV capsid protein, or an isolated AAV capsid protein having a conservative amino acid substitution.
3. a vector encoding a nucleic acid encoding an AAV capsid protein, or a host comprising said nucleic acid operably linked to a promoter.
4. a kit for producing rAAV.
5. an AAV-assembly activation protein.
these shared technical features are previously disclosed by WO 2015/121501 A1 to King's College London (hereinafter King's College).

As to common technical feature #1, King's College teaches a recombinant adeno-associated virus (rAAV) (claim 1; 1. A recombinant adeno-associated virus (AAV) vector comprising:
(a) a variant AAV2 capsid protein, wherein the variant AAV2 capsid protein comprises at least four amino acid substitutions with respect to a wild type AAV2 capsid protein; wherein the at least four amino acid substitutions are present at the following positions in an AAV2 capsid protein sequence: 457, 492, 499 and 533; and (b) a heterologous nucleic acid comprising a nucleotide sequence encoding a gene product").

-----continued on next sheet-----