Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 3089824 A1 2019/08/08

(21) 3 089 824

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) **Date de dépôt PCT/PCT Filing Date:** 2019/02/01
(87) **Date publication PCT/PCT Publication Date:** 2019/08/08
(85) **Entrée phase nationale/National Entry:** 2020/07/28
(86) **N° demande PCT/PCT Application No.:** US 2019/016351
(87) **N° publication PCT/PCT Publication No.:** 2019/152841
(30) **Priorité/Priority:** 2018/02/01 (US62/625,150)

(51) **Cl.Int./Int.Cl.** ***A61K 38/43*** (2006.01),
A61K 35/76 (2015.01), ***A61K 48/00*** (2006.01),
A61P 3/00 (2006.01), ***C12N 15/864*** (2006.01),
C12N 5/10 (2006.01), ***C12N 9/02*** (2006.01)

(71) **Demandeur/Applicant:**
HOMOLOGY MEDICINES, INC., US

(72) **Inventeurs/Inventors:**
SEYMOUR, ALBERT BARNES, US;
AHMED, SEEMIN SEHER, US;
WRIGHT, JASON BOKE, US;
DOLLIVE, SERENA NICOLE, US;
RUBIN, HILLARD, US

(74) **Agent:** SMART & BIGGAR LLP

(54) **Titre : COMPOSITIONS DE VIRUS ADENO-ASSOCIES POUR TRANSFERT DE GENE DE PAH ET PROCEDES D'UTILISATION**
(54) **Title: ADENO-ASSOCIATED VIRUS COMPOSITIONS FOR PAH GENE TRANSFER AND METHODS OF USE THEREOF**

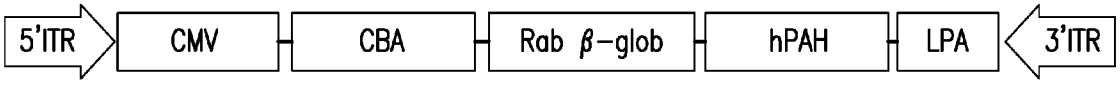

FIG.8A

(57) **Abrégé/Abstract:**
Provided herein are adeno-associated virus (AAV) compositions that can express a phenylalanine hydroxylase (PAH) polypeptide in a cell, thereby restoring the PAH gene function. Also provided are methods of use of the AAV compositions, and packaging systems for making the AAV compositions.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
08 August 2019 (08.08.2019)

WIPO | PCT

(10) International Publication Number
WO 2019/152841 A1

(51) International Patent Classification:
A61K 38/43 (2006.01) *C12N 9/02* (2006.01)
A61K 35/761 (2015.01) *C12N 5/10* (2006.01)
A61K 48/00 (2006.01) *A61P 3/00* (2006.01)
C12N 15/864 (2006.01)

(21) International Application Number:
PCT/US2019/016351

(22) International Filing Date:
01 February 2019 (01.02.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/625,150 01 February 2018 (01.02.2018) US

(71) Applicant: **HOMOLOGY MEDICINES, INC.** [US/US]; 45 Wiggins Avenue, Bedford, Massachusetts 01730 (US).

(72) Inventors: **SEYMOUR, Albert Barnes**; 1 Brickyard Lane, Westborough, Massachusetts 01581 (US). **AHMED, Seemin Seher**; 81 Park Avenue, Apartment 10, Worcester, Massachusetts 01605 (US). **WRIGHT, Jason Boke**; 57 Whits End Road, Concord, MA 01742 (US). **DOLLIVE, Serena Nicole**; 19 Everett St., Unit 1, Waltham, Massachusetts 02453 (US). **RUBIN, Hillard**; 10 Gale Street, Apt. 2, Northborough, Massachusetts 01532 (US).

(74) Agent: **WILKINS, Andrew T.** et al.; Lathrop Gage LLP, 28 State Street, Suite 700, Boston, Massachusetts 02109 (US).

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

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

Declarations under Rule 4.17:
— *as to the identity of the inventor (Rule 4.17(i))*
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

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

(54) Title: ADENO-ASSOCIATED VIRUS COMPOSITIONS FOR PAH GENE TRANSFER AND METHODS OF USE THEREOF

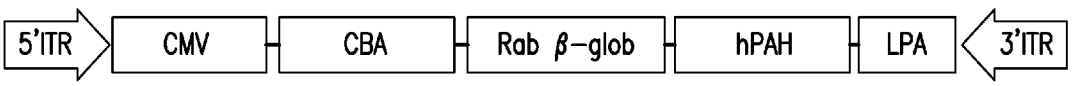

FIG.8A

(57) Abstract: Provided herein are adeno-associated virus (AAV) compositions that can express a phenylalanine hydroxylase (PAH) polypeptide in a cell, thereby restoring the PAH gene function. Also provided are methods of use of the AAV compositions, and packaging systems for making the AAV compositions.

WO 2019/152841 A1

ADENO-ASSOCIATED VIRUS COMPOSITIONS FOR PAH GENE TRANSFER AND METHODS OF USE THEREOF

RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Patent Application Serial No. 62/625,150, filed February 1, 2018, the entire disclosure of which is hereby incorporated herein by reference.

REFERENCE TO SEQUENCE LISTING SUBMITTED ELECTRONICALLY

[0002] The content of the electronically submitted sequence listing in ASCII text file (Name: HMT-025PC_SeqList_ST25.txt; Size: 34,214 bytes; and Date of Creation: January 27, 2019) is incorporated herein by reference in its entirety.

BACKGROUND

[0003] Phenylketonuria (PKU) is an autosomal recessive genetic disorder where the majority of cases are caused by mutations in the phenylalanine hydroxylase (PAH) gene. The PAH gene encodes a hepatic enzyme that catalyzes the hydroxylation of L-phenylalanine (Phe) to L-tyrosine (Tyr) upon multimerization. Reduction or loss of PAH activity leads to phenylalanine accumulation and its conversion into phenylpyruvate (also known as phenylketone). This abnormality in phenylalanine metabolism impairs neuronal maturation and the synthesis of myelin, resulting in mental retardation, seizures and other serious medical problems.

[0004] Currently, there is no cure for PKU. The standard of care is diet management by minimizing foods that contain high amounts of phenylalanine. Dietary management from birth with a low phenylalanine formula largely prevents the development of the intellectual disability of the disorder. However, even on a low-phenylalanine diet, children still suffer from growth retardation, and adults often have osteoporosis and vitamin deficiencies. Moreover, adherence to life-long dietary treatment is difficult, particularly once children reach school age.

[0005] New treatment strategies have recently emerged, including large neutral amino acid (LNAA) supplementation, cofactor tetrahydrobiopterin therapy, enzyme replacement therapy, and genetically modified probiotic therapy. However, these strategies suffer from shortcomings. The LNAA supplementation is suitable only for adults not adhering to a low Phe diet. The cofactor tetrahydrobiopterin can only be used in some mild forms of PKU. Enzyme replacement by administration of a substitute for PAH, e.g., phenylalanine ammonia-lyase (PAL), can lead to immune responses that reduce the efficacy and/or cause side effects.

As to genetically modified probiotic therapy, the pathogenicity of PAL-expressing *E. coli* has been a concern.

[0006] Gene therapy provides a unique opportunity to cure PKU. Retroviral vectors, including lentiviral vectors, are capable of integrating nucleic acids into host cell genomes, raising safety concerns due to their non-targeted insertion into the genome. For example, there is a risk of the vector disrupting a tumor suppressor gene or activating an oncogene, thereby causing a malignancy. Indeed, in a clinical trial for treating X-linked severe combined immunodeficiency (SCID) by transducing $CD34^+$ bone marrow precursors with a gammaretroviral vector, four out of ten patients developed leukemia (Hacein-Bey-Abina *et al.*, J Clin Invest. (2008) 118(9):3132-42). Non-integrating vectors, on the other hand, often suffer insufficient expression level or inadequate duration of expression *in vivo*.

[0007] Accordingly, there is a need in the art for improved gene therapy compositions and methods that can efficiently and safely restore PAH gene function in PKU patients.

SUMMARY

[0008] Provided herein are adeno-associated virus (AAV) compositions that can restore PAH gene function in cells, and methods for using the same to treat diseases associated with reduction of PAH gene function (e.g., PKU). Also provided are packaging systems for making the adeno-associated virus compositions.

[0009] Accordingly, in one aspect, the instant disclosure provides a method for expressing a PAH polypeptide in a cell, the method comprising transducing the cell with a replication-defective adeno-associated virus (AAV) comprising:

(a) an AAV capsid comprising an AAV Clade F capsid protein; and

(b) a transfer genome comprising a transcriptional regulatory element operably linked to a PAH coding sequence.

[0010] In certain embodiments, the cell is a hepatocyte, a renal cell, or a cell in the brain, pituitary gland, adrenal gland, pancreas, urinary bladder, gallbladder, colon, small intestine, or breast. In certain embodiments, the cell is in a mammalian subject and the AAV is administered to the subject in an amount effective to transduce the cell in the subject.

[0011] In another aspect, the instant disclosure provides a method for treating a subject having a disease or disorder associated with a PAH gene mutation, the method comprising administering to the subject an effective amount of a replication-defective AAV comprising:

(a) an AAV capsid comprising an AAV Clade F capsid protein; and

(b) a transfer genome comprising a transcriptional regulatory element operably linked to a PAH coding sequence.

[0012] In certain embodiments, the disease or disorder is phenylketonuria. In certain embodiments, the subject is a human subject.

[0013] In another aspect, the instant disclosure provides a replication-defective adeno-associated virus (AAV) comprising:

(a) an AAV capsid comprising an AAV Clade F capsid protein; and

(b) a transfer genome comprising a transcriptional regulatory element operably linked to a PAH coding sequence.

[0014] The following embodiments apply to each of the foregoing aspects.

[0015] In certain embodiments, the PAH coding sequence encodes an amino acid sequence set forth in SEQ ID NO: 23. In certain embodiments, the PAH coding sequence comprises the nucleotide sequence set forth in SEQ ID NO: 24. In certain embodiments, the PAH coding sequence is silently altered. In certain embodiments, the PAH coding sequence comprises the nucleotide sequence set forth in SEQ ID NO: 25.

[0016] In certain embodiments, the transcriptional regulatory element is capable of mediating transcription in a hepatocyte, a renal cell, or a cell in the brain, pituitary gland, adrenal gland, pancreas, urinary bladder, gallbladder, colon, small intestine, or breast. In certain embodiments, the transcriptional regulatory element is capable of mediating transcription in a hepatocyte or renal cell. In certain embodiments, the transcriptional regulatory element comprises one of more of the elements selected from the group consisting of a CAG promoter, a human EF-1α promoter, a human hepatic control region 1 (HCR1), a human α1-antitrypsin (hAAT) promoter, a hepatic specific regulatory module of the hAAT promoter, an SV40 intron, and a minute virus of mouse (MVM) intron. In certain embodiments, the transcriptional regulatory element comprises a nucleotide sequence at least 90% identical to a sequence selected from the group consisting of SEQ ID NOs: 28-30 and 32-41. In certain embodiments, the transcriptional regulatory element comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 28-30 and 32-41. In certain embodiments, the transcriptional regulatory element comprises from 5' to 3' the nucleotide sequences set forth in SEQ ID NOs: 29, 30, and 31. In certain embodiments, the transcriptional regulatory element comprises the nucleotide sequences set forth in SEQ ID NO: 32.

[0017] In certain embodiments, the transfer genome further comprises an intron operably linked to the PAH coding sequence. In certain embodiments, the intron comprises a nucleotide sequence at least 90% identical to the sequence set forth in SEQ ID NO: 31 or 35.

In certain embodiments, the intron comprises the nucleotide sequence set forth in SEQ ID NO: 31 or 35. In certain embodiments, the transfer genome comprises from 5' to 3': a non-coding exon, the intron, and the PAH coding sequence.

[0018] In certain embodiments, the transfer genome further comprises a polyadenylation sequence 3' to the PAH coding sequence. In certain embodiments, the polyadenylation sequence is an exogenous polyadenylation sequence. In certain embodiments, the exogenous polyadenylation sequence is an SV40 polyadenylation sequence. In certain embodiments, the SV40 polyadenylation sequence comprises a sequence selected from the group consisting of SEQ ID NOs: 42, 43, and 45.

[0019] In certain embodiments, the transfer genome comprises a sequence selected from the group consisting of SEQ ID NOs: 46-50, 61, 64, 67, 74, 76, 78, 80, 82, 84, 86, and 89.

[0020] In certain embodiments, the transfer genome further comprises a 5' inverted terminal repeat (5' ITR) nucleotide sequence 5' of the genome, and a 3' inverted terminal repeat (3' ITR) nucleotide sequence 3' of the genome. In certain embodiments, the 5' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 18, and the 3' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 19. In certain embodiments, the 5' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 20, and the 3' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 21. In certain embodiments, the 5' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 26, and the 3' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 27.

[0021] In certain embodiments, the transfer genome comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 51-55, 62, 65, 68, 75, 77, 79, 81, 83, 85, 87, and 90. In certain embodiments, the transfer genome consists of a nucleotide sequence selected from the group consisting of SEQ ID NOs: 51-55, 62, 65, 68, 75, 77, 79, 81, 83, 85, 87, and 90. In certain embodiments, the transfer genome consists of the nucleotide sequence set forth in SEQ ID NO: 52.

[0022] In certain embodiments, the AAV Clade F capsid protein comprises an amino acid sequence having at least 95% sequence identity with the amino acid sequence of amino acids 203-736 of SEQ ID NO: 2, 3, 4, 6, 7, 10, 11, 12, 13, 15, 16, or 17. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 206 of SEQ ID NO: 2 is C; the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H; the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q;

the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A; the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N; the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S; the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 590 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G or Y; the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M; the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K; the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C; or, the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G.

[0023] In certain embodiments, (a) the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G, and the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G;

(b) the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H, the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M;

(c) the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R;

(d) the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A, and the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; or

(e) the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C.

[0024] In certain embodiments, the capsid protein comprises the amino acid sequence of amino acids 203-736 of SEQ ID NO: 2, 3, 4, 6, 7, 10, 11, 12, 13, 15, 16, or 17.

[0025] In certain embodiments, the AAV Clade F capsid protein comprises an amino acid sequence having at least 95% sequence identity with the amino acid sequence of amino acids 138-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, or 17. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 151 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 160 of SEQ ID NO: 2 is D; the amino acid in the capsid protein corresponding to amino acid 206 of SEQ ID NO: 2 is C; the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H; the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q; the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A; the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N; the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S; the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 590 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G or Y; the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M; the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K; the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C; or, the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G.

[0026] In certain embodiments, (a) the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G, and the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G;

(b) the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H, the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M;

(c) the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R;

(d) the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A, and the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; or

(e) the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C.

[0027] In certain embodiments, the capsid protein comprises the amino acid sequence of amino acids 138-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, or 17.

[0028] In certain embodiments, the AAV Clade F capsid protein comprises an amino acid sequence having at least 95% sequence identity with the amino acid sequence of amino acids 1-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 2 of SEQ ID NO: 2 is T; the amino acid in the capsid protein corresponding to amino acid 65 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 68 of SEQ ID NO: 2 is V; the amino acid in the capsid protein corresponding to amino acid 77 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 119 of SEQ ID NO: 2 is L; the amino acid in the capsid protein corresponding to amino acid 151 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 160 of SEQ ID NO: 2 is D; the amino acid in the capsid protein corresponding to amino acid 206 of SEQ ID NO: 2 is C; the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H; the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q; the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A; the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N; the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S; the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 590 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G or Y; the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M; the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K; the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is

C; or, the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G.

[0029] In certain embodiments, (a) the amino acid in the capsid protein corresponding to amino acid 2 of SEQ ID NO: 2 is T, and the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q;

(b) the amino acid in the capsid protein corresponding to amino acid 65 of SEQ ID NO: 2 is I, and the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is Y;

(c) the amino acid in the capsid protein corresponding to amino acid 77 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K;

(d) the amino acid in the capsid protein corresponding to amino acid 119 of SEQ ID NO: 2 is L, and the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S;

(e) the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G, and the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G;

(f) the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H, the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M;

(g) the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R;

(h) the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A, and the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; or

(i) the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C.

[0030] In certain embodiments, the capsid protein comprises the amino acid sequence of amino acids 1-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17.

[0031] In another aspect, the instant disclosure provides a pharmaceutical composition comprising an AAV disclosed herein.

[0032] In another aspect, the instant disclosure provides a packaging system for recombinant preparation of an AAV, wherein the packaging system comprises:

(a) a Rep nucleotide sequence encoding one or more AAV Rep proteins;

(b) a Cap nucleotide sequence encoding one or more AAV Clade F capsid proteins as disclosed herein; and

(c) a transfer genome as disclosed herein, wherein the packaging system is operative in a cell for enclosing the transfer genome in the capsid to form the AAV.

[0033] In certain embodiments, the packaging system comprises a first vector comprising the Rep nucleotide sequence and the Cap nucleotide sequence, and a second vector comprising the transfer genome. In certain embodiments, the Rep nucleotide sequence encodes an AAV2 Rep protein. In certain embodiments, the AAV2 Rep protein is 78/68 or Rep 68/52. In certain embodiments, the AAV2 Rep protein comprises an amino acid sequence having a minimum percent sequence identity to the AAV2 Rep amino acid sequence of SEQ ID NO: 22, wherein the minimum percent sequence identity is at least 70% across the length of the amino acid sequence encoding the AAV2 Rep protein.

[0034] In certain embodiments, the packaging system further comprises a third vector, wherein the third vector is a helper virus vector. In certain embodiments, the helper virus vector is an independent third vector. In certain embodiments, the helper virus vector is integral with the first vector. In certain embodiments, the helper virus vector is integral with the second vector. In certain embodiments, the third vector comprises genes encoding helper virus proteins.

[0035] In certain embodiments, the helper virus is selected from the group consisting of adenovirus, herpes virus, vaccinia virus, and cytomegalovirus (CMV). In certain embodiments, the helper virus is adenovirus. In certain embodiments, the adenovirus genome comprises one or more adenovirus RNA genes selected from the group consisting of E1, E2, E4 and VA. In certain embodiments, the helper virus is herpes simplex virus (HSV). In certain embodiments, the HSV genome comprises one or more of HSV genes selected from the group consisting of UL5/8/52, ICPO, ICP4, ICP22 and UL30/UL42.

[0036] In certain embodiments, the first vector and the third vector are contained within a first transfecting plasmid. In certain embodiments, the nucleotides of the second vector and the third vector are contained within a second transfecting plasmid. In certain embodiments, the nucleotides of the first vector and the third vector are cloned into a recombinant helper

virus. In certain embodiments, the nucleotides of the second vector and the third vector are cloned into a recombinant helper virus.

[0037] In another aspect, the instant disclosure provides a method for recombinant preparation of an AAV, the method comprising introducing a packaging system as described herein into a cell under conditions operative for enclosing the transfer genome or the transfer genome in the capsid to form the AAV.

BRIEF DESCRIPTION OF THE DRAWINGS

[0038] **Figure 1A, 1B, 1C, 1D, and 1E** are vector maps of the pHMI-hPAH-TC-004, pHMI-hPAH-TC-025, pHMI-hPAH-TC-010, pHMI-hPAH-TC-011, and pHMI-hPAH-TC-012 vectors, respectively.

[0039] **Figure 2** is an image of Western blot showing the expression of human PAH from the pCOH-WT-PAH ("WT PAH"), pCOH-CO-PAH ("CO PAH pCOH"), and pHMI-CO-PAH ("CO PAH pHMI") vectors. $5x10^5$ HEK 293 cells were transfected with 1μg of vector. Lysate of the cells was collected 48 hours after transfection. The expression of human PAH was detected by Western blotting with an anti-PAH antibody (Sigma HPA031642). The amount of GAPDH protein as detected by an anti-GAPDH antibody (Millipore MAB 374) was shown as a loading control.

[0040] **Figures 3A and 3B** are graphs showing the Phe level in the serum of two *pah*$^{-/-}$ mice ("Mouse H1" and "Mouse H5") each administered with $5x10^{13}$ vector genomes of the rAAV-CBA-mPAH vector packaged in an AAVHSC capsid per kg of body weight intravenously via the tail vein. Serum samples were collected in a time course. The Phe levels were measured with a BioAssay Systems ELISA kit EPHE-100 (Figure 3A) or mass spectrometry (Figure 3B).

[0041] **Figure 4** is a graph and a table showing the numbers of vector genomes per 10^6 cells detected in major organs. The rAAV-CBA-mPAH vector packaged in an AAVHSC capsid was injected to *pah*$^{-/-}$ mice intravenously via the tail vein at a dose of $5x10^{13}$ vector genomes per kg of body weight. Organs of the mice were collected 4 weeks after the administration. The numbers of vector genomes per 10^6 cells were measured by the following method: (1) the weight/volume concentration of the vector genome in a sample was measured by Taqman PCR using a standard curve generated with serial dilutions of the vector plasmid; (2) the mass of a single vector genome was calculated based on the sequence of the vector; (3) the number/volume concentration of the vector genome in the sample was calculated; (4) the weight/volume concentration of genomic DNA in the same sample was

measured by Taqman PCR of the apolipoprotein B gene using a standard curve generated with serial dilutions of calculated amounts of genomic DNA isolated from mouse tissues; (5) the number/volume concentration of cell genome in the sample was calculated based on copies of ApoB; and (6) the number of vector genomes per 10^6 cells was calculated by dividing the number/volume concentration of the vector genome by the number/volume concentration of the cell genome and multiplying the result by 10^6.

[0042] **Figures 5A and 5B** are graphs showing the aspartate aminotransferase (AST) and alanine aminotransferase (ALT) levels in the serum of *pah*$^{-/-}$ mice administered with the rAAV-CBA-mPAH vector. The rAAV-CBA-mPAH vector packaged in an AAVHSC capsid was injected to *pah*$^{-/-}$ mice intravenously via the tail vein at a dose of $5x10^{13}$ vector genomes per kg of body weight. Serum samples were collected 4 weeks after the administration. The levels of AST (Figure 5A) and ALT (Figure 5B) were measured by ELISA using the Sigma MAK055 and Sigma MAK052 kits, respectively.

[0043] **Figures 6A-6H** are graphs showing the levels of phenylalanine (Figures 6A, 6C, 6E, and 6G) or tyrosine (Figures 6B, 6D, 6F, and 6H) in the serum of male (Figures 6A, 6B, 6E, and 6F) or female (Figures 6C, 6D, 6G, and 6H) mice administered with the indicated doses of the pHMI-hPAH-TC-004, pHMI-hPAH-TC-025, pHMI-hPAH-TC-010, pHMI-hPAH-TC-011, or pHMI-hPAH-TC-012 vector.

[0044] **Figures 6I-6J** are graphs showing the long-term efficacy on levels of phenylalanine in the serum of male (Figure 6I) or female (Figure 6J) mice administered the indicated doses of the pHMI-hPAH-TC-025 vector.

[0045] **Figures 7A-7D** are graphs showing the levels of phenylalanine (Figures 7A and 7C) or tyrosine (Figures 7B and 7D) in the serum of male (Figures 7A and 7B) or female (Figures 7C and 7D) mice administered with the indicated doses of the pHMI-hPAH-TC-025 vector.

[0046] **Figure 8A, 8B, 8C** are vector maps of pHMI-hPAH-TC-009, pHMI-hPAH-TC-013 and pHMI-hPAH-TC-017 vectors, respectively.

[0047] **Figure 9A-9B** depict the quantification of Western blots of human PAH expression, from the indicated AAV vectors, in Huh7 cells (**Figure 9A**) and HEK293 cells (**Figure 9B**).

[0048] **Figure 10A-10C** are graphs showing serum phenylalanine levels in mice that have been administered the indicated AAV vectors. **Figure 10A** is a graph showing serum Phe levels over time of male *Pah*$^{-/-}$ PAH^{enu2} mice. **Figure 10B** is a graph showing serum Phe levels over time of female *Pah*$^{-/-}$ PAH^{enu2} mice. **Figure 10C** is a graph showing the average

baseline serum Phe levels of the male and female mice in the study (55 mice per group; **** indicates $p<0.05$).

[0049] **Figure 11A, 11B, 11C, 11D, 11E, and 11F** are vector maps of pHMI-hPAH-TC-018, pHMI-hPAH-TC-019, pHMI-hPAH-TC-020, pHMI-hPAH-TC-021, pHMI-hPAH-TC-022, and pHMI-hPAH-TC-023 vectors, respectively.

[0050] **Figure 12** depicts the quantification of Western blots of human PAH expression from HEK293 cells transfected with the indicated AAV vectors under the control of a CBA promoter.

[0051] **Figure 13** is a graph showing serum phenylalanine levels over time of male *Pah*$^{-/-}$ PAH^{enu2} mice administered the indicated AAV vectors.

DETAILED DESCRIPTION

[0052] The instant disclosure provided adeno-associated virus (AAV) compositions that can restore PAH gene function in a cell. Also provide are packaging systems for making the adeno-associated virus compositions.

I. Definitions

[0053] As used herein, the term "replication-defective adeno-associated virus" refers to an AAV comprising a genome lacking Rep and Cap genes.

[0054] As used herein, the term "PAH gene" refers to the phenylalanine hydroxylase gene. The human PAH gene is identified by Entrez Gene ID 5053. An exemplary nucleotide sequence of a PAH mRNA is provided as SEQ ID NO: 24. An exemplary amino acid sequence of a PAH polypeptide is provided as SEQ ID NO: 23.

[0055] As used herein, the term "transfer genome" refers to a recombinant AAV genome comprising a coding sequence operably linked to an exogenous transcriptional regulatory element that mediates expression of the coding sequence when the transfer genome is introduced into a cell. In certain embodiments, the transfer genome does not integrate in the chromosomal DNA of the cell. The skilled artisan will appreciate that the portion of a transfer genome comprising the transcriptional regulatory element operably linked to a PAH coding sequence can be in the sense or antisense orientation relative to direction of transcription of the PAH coding sequence.

[0056] As used herein, the term "Clade F capsid protein" refers to an AAV VP1, VP2, or VP3 capsid protein that has at least 90% identity with the VP1, VP2, or VP3 amino acid

sequences set forth, respectively, in amino acids 1-736, 138-736, and 203-736 of SEQ ID NO: 1 herein.

[0057] As used herein, the identity between two nucleotide sequences or between two amino acid sequences is determined by the number of identical nucleotides or amino acids in alignment divided by the full length of the longer nucleotide or amino acid sequence.

[0058] As used herein, the term "a disease or disorder associated with a PAH gene mutation" refers to any disease or disorder caused by, exacerbated by, or genetically linked with mutation of a PAH gene. In certain embodiments, the disease or disorder associated with a PAH gene mutation is phenylketonuria (PKU).

[0059] As used herein, the term "coding sequence" refers to the portion of a complementary DNA (cDNA) that encodes a polypeptide, starting at the start codon and ending at the stop codon. A gene may have one or more coding sequences due to alternative splicing, alternative translation initiation, and variation within the population. A coding sequence may either be wild-type or codon-altered. An exemplary wild-type PAH coding sequence is set forth in SEQ ID NO: 24.

[0060] As used herein, the term "silently altered" refers to alteration of a coding sequence or a stuffer-inserted coding sequence of a gene (*e.g.*, by nucleotide substitution) without changing the amino acid sequence of the polypeptide encoded by the coding sequence or stuffer-inserted coding sequence. Such silent alteration is advantageous in that it may increase the translation efficiency of a coding sequence.

[0061] In the instant disclosure, nucleotide positions in a PAH gene are specified relative to the first nucleotide of the start codon. The first nucleotide of a start codon is position 1; the nucleotides 5' to the first nucleotide of the start codon have negative numbers; the nucleotides 3' to the first nucleotide of the start codon have positive numbers. An exemplary nucleotide 1 of the human PAH gene is nucleotide 5,473 of the NCBI Reference Sequence: NG_008690.1, and an exemplary nucleotide 3 of the human PAH gene is nucleotide 5,475 of the NCBI Reference Sequence: NG_008690.1. The nucleotide adjacently 5' to the start codon is nucleotide -1.

[0062] In the instant disclosure, exons and introns in a PAH gene are specified relative to the exon encompassing the first nucleotide of the start codon, which is nucleotide 5473 of the NCBI Reference Sequence: NG_008690.1. The exon encompassing the first nucleotide of the start codon is exon 1. Exons 3' to exon 1 are from 5' to 3': exon 2, exon 3, etc. Introns 3' to exon 1 are from 5' to 3': intron 1, intron 2, etc. Accordingly, the PAH gene comprises from 5' to 3': exon 1, intron 1, exon 2, intron 2, exon 3, etc. An exemplary exon 1 of the human PAH

gene is nucleotides 5001-5532 of the NCBI Reference Sequence: NG_008690.1. An exemplary intron 1 of the human PAH gene is nucleotides 5533-9704 of the NCBI Reference Sequence: NG_008690.1.

[0063] As used herein, the term "transcriptional regulatory element" or "TRE" refers to a cis-acting nucleotide sequence, for example, a DNA sequence, that regulates (e.g., controls, increases, or reduces) transcription of an operably linked nucleotide sequence by an RNA polymerase to form an RNA molecule. A TRE relies on one or more trans-acting molecules, such as transcription factors, to regulate transcription. Thus, one TRE may regulate transcription in different ways when it is in contact with different trans-acting molecules, for example, when it is in different types of cells. A TRE may comprise one or more promoter elements and/or enhancer elements. A skilled artisan would appreciate that the promoter and enhancer elements in a gene may be close in location, and the term "promoter" may refer to a sequence comprising a promoter element and an enhancer element. Thus, the term "promoter" does not exclude an enhancer element in the sequence. The promoter and enhancer elements do not need to be derived from the same gene or species, and the sequence of each promoter or enhancer element may be either identical or substantially identical to the corresponding endogenous sequence in the genome.

[0064] As used herein, the term "operably linked" is used to describe the connection between a TRE and a coding sequence to be transcribed. Typically, gene expression is placed under the control of a TRE comprising one or more promoter and/or enhancer elements. The coding sequence is "operably linked" to the TRE if the transcription of the coding sequence is controlled or influenced by the TRE. The promoter and enhancer elements of the TRE may be in any orientation and/or distance from the coding sequence, as long as the desired transcriptional activity is obtained. In certain embodiments, the TRE is upstream from the coding sequence.

[0065] As used herein, the term "polyadenylation sequence" refers to a DNA sequence that when transcribed into RNA constitutes a polyadenylation signal sequence. The polyadenylation sequence can be native (e.g., from the PAH gene) or exogenous. The exogenous polyadenylation sequence can be a mammalian or a viral polyadenylation sequence (e.g., an SV40 polyadenylation sequence).

[0066] As used herein, "exogenous polyadenylation sequence" refers to a polyadenylation sequence not identical or substantially identical to the endogenous polyadenylation sequence of a PAH gene (e.g., human PAH gene). In certain embodiments, an exogenous polyadenylation sequence is a polyadenylation sequence of a non-PAH gene in

the same species (e.g., human). In certain embodiments, an exogenous polyadenylation sequence is a polyadenylation sequence of a different species (e.g., a virus).

[0067] As used herein, the term "effective amount" in the context of the administration of an AAV to a subject refers to the amount of the AAV that achieves a desired prophylactic or therapeutic effect.

II. Adeno-Associated Virus Compositions

[0068] In one aspect, provided herein are novel replication-defective AAV compositions useful for expressing PAH polypeptide in cells with reduced or otherwise defective PAH gene function. In certain embodiments, the AAV disclosed herein comprise: an AAV capsid comprising an AAV Clade F capsid protein; and a transfer genome comprising a transcriptional regulatory element operably linked to a PAH coding sequence, allowing for extrachromosomal expression of PAH.

[0069] Any AAV Clade F capsid protein or derivative thereof can be used in the AAV compositions disclosed herein. For example, in certain embodiments, the AAV Clade F capsid protein comprises an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the amino acid sequence of amino acids 203-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17. In certain embodiments, the AAV Clade F capsid protein comprises an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the amino acid sequence of amino acids 203-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17, wherein: the amino acid in the capsid protein corresponding to amino acid 206 of SEQ ID NO: 2 is C; the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H; the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q; the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A; the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N; the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S; the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 590 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G or Y; the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M; the amino acid in the capsid protein corresponding to

amino acid 687 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K; the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C; or, the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G, and the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H, the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A, and the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C. In certain embodiments, the AAV Clade F capsid protein comprises the amino acid sequence of amino acids 203-736 of SEQ ID NO: 2, 3, 4, 6, 7, 10, 11, 12, 13, 15, 16, or 17.

[0070] For example, in certain embodiments, the AAV Clade F capsid protein comprises an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the amino acid sequence of amino acids 138-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17. In certain embodiments, the AAV Clade F capsid protein comprises an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the amino acid sequence of amino acids 138-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17, wherein: the amino acid in the capsid protein corresponding to amino acid 151 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 160 of SEQ ID NO: 2 is D; the amino acid in the capsid protein corresponding to amino acid 206 of SEQ ID NO: 2 is C; the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H; the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q; the amino acid in the capsid protein corresponding to amino acid

346 of SEQ ID NO: 2 is A; the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N; the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S; the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 590 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G or Y; the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M; the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K; the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C; or, the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G, and the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H, the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A, and the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C. In certain embodiments, the AAV Clade F capsid protein comprises the amino acid sequence of amino acids 138-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, or 17.

[0071] For example, in certain embodiments, the AAV Clade F capsid protein comprises an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the amino acid sequence of amino acids 1-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17. In certain embodiments, the AAV Clade F capsid protein comprises an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%,

90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the amino acid sequence of amino acids 1-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17, wherein: the amino acid in the capsid protein corresponding to amino acid 2 of SEQ ID NO: 2 is T; the amino acid in the capsid protein corresponding to amino acid 65 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 68 of SEQ ID NO: 2 is V; the amino acid in the capsid protein corresponding to amino acid 77 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 119 of SEQ ID NO: 2 is L; the amino acid in the capsid protein corresponding to amino acid 151 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 160 of SEQ ID NO: 2 is D; the amino acid in the capsid protein corresponding to amino acid 206 of SEQ ID NO: 2 is C; the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H; the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q; the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A; the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N; the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S; the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 590 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G or Y; the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M; the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K; the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C; or, the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 2 of SEQ ID NO: 2 is T, and the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 65 of SEQ ID NO: 2 is I, and the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is Y. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 77 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 119 of SEQ ID NO: 2 is L, and the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S. In certain embodiments, the amino acid in the capsid protein corresponding

to amino acid 626 of SEQ ID NO: 2 is G, and the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H, the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A, and the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R. In certain embodiments, the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C. In certain embodiments, the AAV Clade F capsid protein comprises the amino acid sequence of amino acids 1-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17.

[0072] In certain embodiments, the AAV capsid comprises two or more of: (a) a Clade F capsid protein comprising the amino acid sequence of amino acids 203-736 of SEQ ID NO: 2, 3, 4, 6, 7, 10, 11, 12, 13, 15, 16, or 17; (b) a Clade F capsid protein comprising the amino acid sequence of amino acids 138-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, or 17; and (c) a Clade F capsid protein comprising the amino acid sequence of amino acids 1-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17. In certain embodiments, the AAV capsid comprises: (a) a Clade F capsid protein having an amino acid sequence consisting of amino acids 203-736 of SEQ ID NO: 2, 3, 4, 6, 7, 10, 11, 12, 13, 15, 16, or 17; (b) a Clade F capsid protein having an amino acid sequence consisting of amino acids 138-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, or 17; and (c) a Clade F capsid protein having an amino acid sequence consisting of amino acids 1-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17.

[0073] In certain embodiments, the AAV capsid comprises one or more of: (a) a Clade F capsid protein comprising an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the sequence of amino acids 203-736 of SEQ ID NO: 8; (b) a Clade F capsid protein comprising an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%

sequence identity with the sequence of amino acids 138-736 of SEQ ID NO: 8; and (c) a Clade F capsid protein comprising an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the sequence of amino acids 1-736 of SEQ ID NO: 8. In certain embodiments, the AAV capsid comprises one or more of: (a) a Clade F capsid protein comprising the amino acid sequence of amino acids 203-736 of SEQ ID NO: 8; (b) a Clade F capsid protein comprising the amino acid sequence of amino acids 138-736 of SEQ ID NO: 8; and (c) a Clade F capsid protein comprising the amino acid sequence of amino acids 1-736 of SEQ ID NO: 8. In certain embodiments, the AAV capsid comprises two or more of: (a) a Clade F capsid protein comprising the amino acid sequence of amino acids 203-736 of SEQ ID NO: 8; (b) a Clade F capsid protein comprising the amino acid sequence of amino acids 138-736 of SEQ ID NO: 8; and (c) a Clade F capsid protein comprising the amino acid sequence of amino acids 1-736 of SEQ ID NO: 8. In certain embodiments, the AAV capsid comprises: (a) a Clade F capsid protein having an amino acid sequence consisting of amino acids 203-736 of SEQ ID NO: 8; (b) a Clade F capsid protein having an amino acid sequence consisting of amino acids 138-736 of SEQ ID NO: 8; and (c) a Clade F capsid protein having an amino acid sequence consisting of amino acids 1-736 of SEQ ID NO: 8.

[0074] In certain embodiments, the AAV capsid comprises one or more of: (a) a Clade F capsid protein comprising an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the sequence of amino acids 203-736 of SEQ ID NO: 11; (b) a Clade F capsid protein comprising an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the sequence of amino acids 138-736 of SEQ ID NO: 11; and (c) a Clade F capsid protein comprising an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the sequence of amino acids 1-736 of SEQ ID NO: 11. In certain embodiments, the AAV capsid comprises one or more of: (a) a Clade F capsid protein comprising the amino acid sequence of amino acids 203-736 of SEQ ID NO: 11; (b) a Clade F capsid protein comprising the amino acid sequence of amino acids 138-736 of SEQ ID NO: 11; and (c) a Clade F capsid protein comprising the amino acid sequence of amino acids 1-736 of SEQ ID NO: 11. In certain embodiments, the AAV capsid comprises two or more of: (a) a Clade F capsid protein comprising the amino acid sequence of amino acids 203-736 of SEQ ID NO: 11; (b) a Clade F capsid protein comprising the amino acid sequence of amino acids 138-

736 of SEQ ID NO: 11; and (c) a Clade F capsid protein comprising the amino acid sequence of amino acids 1-736 of SEQ ID NO: 11. In certain embodiments, the AAV capsid comprises: (a) a Clade F capsid protein having an amino acid sequence consisting of amino acids 203-736 of SEQ ID NO: 11; (b) a Clade F capsid protein having an amino acid sequence consisting of amino acids 138-736 of SEQ ID NO: 11; and (c) a Clade F capsid protein having an amino acid sequence consisting of amino acids 1-736 of SEQ ID NO: 11.

[0075] In certain embodiments, the AAV capsid comprises one or more of: (a) a Clade F capsid protein comprising an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the sequence of amino acids 203-736 of SEQ ID NO: 13; (b) a Clade F capsid protein comprising an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the sequence of amino acids 138-736 of SEQ ID NO: 13; and (c) a Clade F capsid protein comprising an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the sequence of amino acids 1-736 of SEQ ID NO: 13. In certain embodiments, the AAV capsid comprises one or more of: (a) a Clade F capsid protein comprising the amino acid sequence of amino acids 203-736 of SEQ ID NO: 13; (b) a Clade F capsid protein comprising the amino acid sequence of amino acids 138-736 of SEQ ID NO: 13; and (c) a Clade F capsid protein comprising the amino acid sequence of amino acids 1-736 of SEQ ID NO: 13. In certain embodiments, the AAV capsid comprises two or more of: (a) a Clade F capsid protein comprising the amino acid sequence of amino acids 203-736 of SEQ ID NO: 13; (b) a Clade F capsid protein comprising the amino acid sequence of amino acids 138-736 of SEQ ID NO: 13; and (c) a Clade F capsid protein comprising the amino acid sequence of amino acids 1-736 of SEQ ID NO: 13. In certain embodiments, the AAV capsid comprises: (a) a Clade F capsid protein having an amino acid sequence consisting of amino acids 203-736 of SEQ ID NO: 13; (b) a Clade F capsid protein having an amino acid sequence consisting of amino acids 138-736 of SEQ ID NO: 13; and (c) a Clade F capsid protein having an amino acid sequence consisting of amino acids 1-736 of SEQ ID NO: 13.

[0076] In certain embodiments, the AAV capsid comprises one or more of: (a) a Clade F capsid protein comprising an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the sequence of amino acids 203-736 of SEQ ID NO: 16; (b) a Clade F capsid protein comprising an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%,

85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the sequence of amino acids 138-736 of SEQ ID NO: 16; and (c) a Clade F capsid protein comprising an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with the sequence of amino acids 1-736 of SEQ ID NO: 16. In certain embodiments, the AAV capsid comprises one or more of: (a) a Clade F capsid protein comprising the amino acid sequence of amino acids 203-736 of SEQ ID NO: 16; (b) a Clade F capsid protein comprising the amino acid sequence of amino acids 138-736 of SEQ ID NO: 16; and (c) a Clade F capsid protein comprising the amino acid sequence of amino acids 1-736 of SEQ ID NO: 16. In certain embodiments, the AAV capsid comprises two or more of: (a) a Clade F capsid protein comprising the amino acid sequence of amino acids 203-736 of SEQ ID NO: 16; (b) a Clade F capsid protein comprising the amino acid sequence of amino acids 138-736 of SEQ ID NO: 16; and (c) a Clade F capsid protein comprising the amino acid sequence of amino acids 1-736 of SEQ ID NO: 16. In certain embodiments, the AAV capsid comprises: (a) a Clade F capsid protein having an amino acid sequence consisting of amino acids 203-736 of SEQ ID NO: 16; (b) a Clade F capsid protein having an amino acid sequence consisting of amino acids 138-736 of SEQ ID NO: 16; and (c) a Clade F capsid protein having an amino acid sequence consisting of amino acids 1-736 of SEQ ID NO: 16.

[0077] Transfer genomes useful in the AAV compositions disclosed herein generally comprise a transcriptional regulatory element (TRE) operably linked to a PAH coding sequence. In certain embodiments, the transfer genome comprises a 5' inverted terminal repeat (5' ITR) nucleotide sequence 5' of the TRE and PAH coding sequence, and a 3' inverted terminal repeat (3' ITR) nucleotide sequence 3' of the TRE and PAH coding sequence.

[0078] In certain embodiments, the PAH coding sequence comprises all or substantially all of a coding sequence of a PAH gene. In certain embodiments, the transfer genome comprises a nucleotide sequence encoding SEQ ID NO: 23 and can optionally further comprise an exogenous polyadenylation sequence 3' to the PAH coding sequence. In certain embodiments, the nucleotide sequence encoding SEQ ID NO: 23 is wild-type (e.g., having the sequence set forth in SEQ ID NO: 24). In certain embodiments, the nucleotide sequence encoding SEQ ID NO: 23 is codon-altered (e.g., having the sequence set forth in SEQ ID NO: 25). In certain embodiments, the nucleotide sequence encoding SEQ ID NO: 23 is codon-altered (e.g., having the sequence set forth in SEQ ID NO: 69, 70, 71, 72, or 73).

[0079] In certain embodiments, the PAH coding sequence encodes a polypeptide comprising all or substantially all of the amino acids sequence of a PAH protein. In certain

embodiments, the PAH coding sequence encodes the amino acid sequence of a wild-type PAH protein (e.g., human PAH protein). In certain embodiments, the PAH coding sequence encodes the amino acid sequence of a mutant PAH protein (e.g., human PAH protein), wherein the mutant PAH polypeptide is a functional equivalent of the wild-type PAH polypeptide, i.e., can function as a wild-type PAH polypeptide. In certain embodiments, the functionally equivalent PAH polypeptide further comprises at least one characteristic not found in the wild-type PAH polypeptide, e.g., the ability to stabilize PAH protein (e.g., dimer or tetramer), or the ability to resist protein degradation.

[0080] The transfer genome can be used to express PAH in any mammalian cells (e.g., human cells). Thus, the TRE can be active in any mammalian cells (e.g., human cells). In certain embodiments, the TRE is active in a broad range of human cells. Such TREs may comprise constitutive promoter and/or enhancer elements including cytomegalovirus (CMV) promoter/enhancer (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 58), SV40 promoter, chicken beta actin (CBA) promoter (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 59), human elongation factor 1 alpha (EF1α) promoter (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 40), minute virus of mouse (MVM) intron which comprises transcription factor binding sites (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 35), human phosphoglycerate kinase (PGK1) promoter, human ubiquitin C (Ubc) promoter, human beta actin promoter, human neuron-specific enolase (ENO2) promoter, human beta-glucuronidase (GUSB) promoter, a rabbit beta-globin element (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 60), and/or human Methyl-CpG Binding Protein 2 (MeCP2) promoter. Any of these TREs can be combined in any order to drive efficient transcription. For example, a transfer genome may comprise a CMV enhancer, a CBA promoter, and the splice acceptor from exon 3 of the rabbit beta-globin gene, collectively called a CAG promoter (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 28). For example, a transfer genome may comprise a hybrid of CMV enhancer and CBA promoter followed by a splice donor and splice acceptor, collectively called a CASI promoter region (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 63).

[0081] Alternatively, the TRE may be a tissue-specific TRE, i.e., it is active in specific tissue(s) and/or organ(s). A tissue-specific TRE comprises one or more tissue-specific promoter and/or enhancer elements, and optionally one or more constitutive promoter and/or enhancer elements. A skilled artisan would appreciate that tissue-specific promoter and/or enhancer elements can be isolated from genes specifically expressed in the tissue by methods well known in the art. In certain embodiments, the TRE is liver-specific (e.g., hepatocyte-specific). Exemplary liver-specific TREs may comprise one or more elements selected from the group consisting of human albumin promoter, human transthyretin (TTR) promoter (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 34), human APOE/C-I hepatic control region (HCR) 1 or 2 (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 29 or 37), human APOH promoter, and human SERPINA1 (hAAT) promoter (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 30 or 38) or a hepatic specific regulatory module thereof (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 33). In certain embodiments, an hAAT promoter region comprises a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 66. More liver-specific promoter elements are disclosed in WO 2009/130208 and Kramer et al. (Molecular Therapy (2003) 7, 375–385), which are incorporated by reference herein in their entirety.

[0082] In certain embodiments, the TRE is kidney-specific (e.g., renal epithelial cell-specific). Exemplary kidney-specific TREs may comprise one or more elements selected from the group consisting of human nephrin promoter, human parathyroid hormone receptor promoter, human uromodulin promoter, and human SLC12A1 promoter. In certain embodiments, the TRE is brain-specific (e.g., neuron-specific, glial cell-specific, astrocyte-specific, oligodendrocyte-specific, microglia-specific and/or central nervous system-specific). Exemplary brain-specific TREs may comprise one or more elements selected from the group consisting of human glial fibrillary acidic protein (GFAP) promoter and human synapsin 1 (SYN1) promoter. More brain-specific promoter elements are disclosed in WO 2016/100575A1, which is incorporated by reference herein in its entirety.

[0083] In certain embodiments, the transfer genome comprises two or more TREs, optionally comprising at least one of the TREs disclosed above. A skilled person in the art would appreciate that any of these TREs can be combined in any order, and combinations of a

constitutive TRE and a tissue-specific TRE can drive efficient and tissue-specific transcription. For example, in certain embodiments, the transfer genome comprises a human HCR1 (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 29 or 37) and a human EF-1α promoter (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 40), optionally wherein the human HCR1 is 5' to the human EF-1α promoter. In certain embodiments, the transfer genome comprises a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the sequence set forth in SEQ ID NO: 41. In certain embodiments, the transfer genome comprises a human HCR1 (e.g., comprising the nucleotide sequence set forth in SEQ ID NO: 29 or 37) and a human EF-1α promoter (e.g., comprising the nucleotide sequence set forth in SEQ ID NO: 40), optionally wherein the human HCR1 is 5' to the human EF-1α promoter. In certain embodiments, the transfer genome comprises the sequence set forth in SEQ ID NO: 41.

[0084] Similarly, combinations of two or more tissue-specific TREs can drive efficient and tissue-specific transcription. For example, in certain embodiments, the transfer genome comprises a human HCR1 (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 29) and a hAAT promoter (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 30), optionally wherein the human HCR1 is 5' to the hAAT promoter. In certain embodiments, the transfer genome comprises a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the sequence set forth in SEQ ID NO: 32. In certain embodiments, the transfer genome comprises a human HCR1 (e.g., comprising the nucleotide sequence set forth in SEQ ID NO: 29) and a hAAT promoter (e.g., comprising the nucleotide sequence set forth in SEQ ID NO: 30), optionally wherein the human HCR1 is 5' to the hAAT promoter. In certain embodiments, the transfer genome comprises the nucleotide sequence set forth in SEQ ID NO: 32.

[0085] In certain embodiments, the transfer genome comprises a hepatic specific regulatory module of hAAT promoter (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 33) and a human TTR promoter (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 34), optionally wherein the hepatic specific regulatory module is 5' to the human TTR promoter. In certain embodiments, the transfer genome comprises a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the sequence set forth in SEQ ID NO: 36. In certain

embodiments, the transfer genome comprises a hepatic specific regulatory module of hAAT promoter (e.g., comprising the nucleotide sequence set forth in SEQ ID NO: 33) and a human TTR promoter (e.g., comprising the nucleotide sequence set forth in SEQ ID NO: 34), optionally wherein the hepatic specific regulatory module is 5' to the human TTR promoter. In certain embodiments, the transfer genome comprises the nucleotide sequence set forth in SEQ ID NO: 36.

[0086] In certain embodiment, the transfer genome comprises a human HCR1 (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 29 or 37) and a hAAT promoter (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 30 or 38), optionally wherein the human HCR1 is 5' to the hAAT promoter. In certain embodiments, the transfer genome comprises a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the sequence set forth in SEQ ID NO: 39. In certain embodiment, the transfer genome comprises a human HCR1 (e.g., comprising the nucleotide sequence set forth in SEQ ID NO: 29 or 37) and a hAAT promoter (e.g., comprising the nucleotide sequence set forth in SEQ ID NO: 30 or 38), optionally wherein the human HCR1 is 5' to the hAAT promoter. In certain embodiments, the transfer genome comprises the nucleotide sequence set forth in SEQ ID NO: 39.

[0087] In certain embodiments, the transfer vector further comprises an intron 5' to or inserted in the PAH coding sequence. Such introns can increase transgene expression, for example, by reducing transcriptional silencing and enhancing mRNA export from the nucleus to the cytoplasm. In certain embodiments, the transfer genome comprises from 5' to 3': a non-coding exon, an intron, and the PAH coding sequence. In certain embodiments, an intron sequence is inserted in the PAH coding sequence, optionally wherein the intron is inserted at an internucleotide bond that links two native exons. In certain embodiments, the intron is inserted at an internucleotide bond that links native exon 1 and exon 2.

[0088] The intron can comprise a native intron sequence of the PAH gene, an intron sequence from a different species or a different gene from the same species, and/or a synthetic intron sequence. A skilled worker will appreciate that synthetic intron sequences can be designed to mediate RNA splicing by introducing any consensus splicing motifs known in the art (e.g., in Sibley et al., (2016) Nature Reviews Genetics, 17, 407-21, which is incorporated by reference herein in its entirety). Exemplary intron sequences are provided in Lu et al. (2013) Molecular Therapy 21(5): 954-63, and Lu et al. (2017) Hum. Gene Ther. 28(1): 125-34, which are incorporated by reference herein in their entirety. In certain embodiments, the transfer

genome comprises an SV40 intron (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 31) or a minute virus of mouse (MVM) intron (e.g., comprising a nucleotide sequence at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 35). In certain embodiments, the transfer genome comprises an SV40 intron (e.g., comprising the nucleotide sequence set forth in SEQ ID NO: 31) or a minute virus of mouse (MVM) intron (e.g., comprising the nucleotide sequence set forth in SEQ ID NO: 35).

[0089] In certain embodiments, the transfer genome disclosed herein further comprises a transcription terminator (e.g., a polyadenylation sequence). In certain embodiments, the transcription terminator is 3' to the PAH coding sequence. The transcription terminator may be any sequence that effectively terminates transcription, and a skilled artisan would appreciate that such sequences can be isolated from any genes that are expressed in the cell in which transcription of the PAH coding sequence is desired. In certain embodiments, the transcription terminator comprises a polyadenylation sequence. In certain embodiments, the polyadenylation sequence is identical or substantially identical to the endogenous polyadenylation sequence of the human PAH gene. In certain embodiments, the polyadenylation sequence is an exogenous polyadenylation sequence. In certain embodiments, the polyadenylation sequence is an SV40 polyadenylation sequence (e.g., comprising the nucleotide sequence set forth in SEQ ID NO: 42, 43, or 45, or a nucleotide sequence complementary thereto). In certain embodiments, the polyadenylation sequence comprises the sequence set forth in SEQ ID NO: 43.

[0090] In certain embodiments, the transfer genome comprises from 5' to 3': a TRE, optionally a non-coding exon and an intron, a PAH coding sequence, and a polyadenylation sequence. In certain embodiments, the TRE has at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to any one of SEQ ID NOs: 28-30 and 32-41; the intron has at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to SEQ ID NO: 31 or 35; the PAH coding sequence has at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to SEQ ID NO: 25; and/or the polyadenylation sequence has at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to any one of SEQ ID NOs: 42, 43, and 45. In certain embodiments, the TRE comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 28-30 and 32-41; the intron comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 31 and 35; the PAH coding sequence comprises the sequence set forth in SEQ ID NO: 25; and/or the polyadenylation sequence comprises a nucleotide sequence selected from the group

consisting of SEQ ID NOs: 42, 43, and 45. In certain embodiments, the TRE comprises the sequence set forth in SEQ ID NO: 28; the PAH coding sequence comprises the sequence set forth in SEQ ID NO: 25; and/or the polyadenylation sequence comprises the sequence set forth in SEQ ID NO: 42. In certain embodiments, the TRE comprises from 5' to 3' the sequence set forth in SEQ ID NO: 29, and the sequence set forth in SEQ ID NO: 30 (e.g., the TRE comprises the sequence set forth in SEQ ID NO: 32); the intron comprises the sequence set forth in SEQ ID NO: 31; the PAH coding sequence comprises the sequence set forth in SEQ ID NO: 25; and/or the polyadenylation sequence comprises the sequence set forth in SEQ ID NO: 43. In certain embodiments, the TRE comprises from 5' to 3' the sequence set forth in SEQ ID NO: 33, and the sequence set forth in SEQ ID NO: 34 (e.g., the TRE comprises the sequence set forth in SEQ ID NO: 36); the intron comprises the sequence set forth in SEQ ID NO: 35; the PAH coding sequence comprises the sequence set forth in SEQ ID NO: 25; and/or the polyadenylation sequence comprises the sequence set forth in SEQ ID NO: 45. In certain embodiments, the TRE comprises from 5' to 3' the sequence set forth in SEQ ID NO: 37, and the sequence set forth in SEQ ID NO: 38 (e.g., the TRE comprises the sequence set forth in SEQ ID NO: 39); the intron comprises the sequence set forth in SEQ ID NO: 35; the PAH coding sequence comprises the sequence set forth in SEQ ID NO: 25; and/or the polyadenylation sequence comprises the sequence set forth in SEQ ID NO: 45. In certain embodiments, the TRE comprises from 5' to 3' the sequence set forth in SEQ ID NO: 37, and the sequence set forth in SEQ ID NO: 40 (e.g., the TRE comprises the sequence set forth in SEQ ID NO: 41); and/or the polyadenylation sequence comprises the sequence set forth in SEQ ID NO: 45.

[0091] In certain embodiments, the transfer genome comprises a sequence at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 46, 47, 48, 49, 50, 61, 64, 67, 74, 76, 78, 80, 82, 84, 86, or 89. In certain embodiments, the transfer genome comprises the nucleotide sequence set forth in SEQ ID NO: 46, 47, 48, 49, 50, 61, 64, 67, 74, 76, 78, 80, 82, 84, 86, or 89. In certain embodiments, the transfer genome consists of the nucleotide sequence set forth in SEQ ID NO: 46, 47, 48, 49, 50, 61, 64, 67, 74, 76, 78, 80, 82, 84, 86, or 89. In certain embodiments, the transfer genome comprises the nucleotide sequence set forth in SEQ ID NO: 47. In certain embodiments, the transfer genome consists of the nucleotide sequence set forth in SEQ ID NO: 47.

[0092] In certain embodiments, the transfer genomes disclosed herein further comprise a 5' inverted terminal repeat (5' ITR) nucleotide sequence 5' of the TRE, and a 3' inverted terminal repeat (3' ITR) nucleotide sequence 3' of the PAH coding sequence. ITR sequences

from any AAV serotype or variant thereof can be used in the transfer genomes disclosed herein. The 5' and 3' ITR can be from an AAV of the same serotype or from AAVs of different serotypes. Exemplary ITRs for use in the transfer genomes disclosed herein are set forth in SEQ ID NOs: 18-21, 26, and 27 herein.

[0093] In certain embodiments, the 5' ITR or 3' ITR is from AAV2. In certain embodiments, both the 5' ITR and the 3' ITR are from AAV2. In certain embodiments, the 5' ITR nucleotide sequence has at least 90% (e.g., at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%) sequence identity to SEQ ID NO: 18, or the 3' ITR nucleotide sequence has at least 90% (e.g., at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%) sequence identity to SEQ ID NO: 19. In certain embodiments, the 5' ITR nucleotide sequence has at least 90% (e.g., at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%) sequence identity to SEQ ID NO: 18, and the 3' ITR nucleotide sequence has at least 90% (e.g., at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%) sequence identity to SEQ ID NO: 19. In certain embodiments, the transfer genome comprises a nucleotide sequence set forth in any one of SEQ ID NOs: 46-50, 61, 64, 67, 74, 76, 78, 80, 82, 84, 86, and 89, a 5' ITR nucleotide sequence having the sequence of SEQ ID NO: 18, and a 3' ITR nucleotide sequence having the sequence of SEQ ID NO: 19.

[0094] In certain embodiments, the 5' ITR or 3' ITR are from AAV5. In certain embodiments, both the 5' ITR and 3' ITR are from AAV5. In certain embodiments, the 5' ITR nucleotide sequence has at least 90% (e.g., at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%) sequence identity to SEQ ID NO: 20, or the 3' ITR nucleotide sequence has at least 90% (e.g., at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%) sequence identity to SEQ ID NO: 21. In certain embodiments, the 5' ITR nucleotide sequence has at least 90% (e.g., at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%) sequence identity to SEQ ID NO: 20, and the 3' ITR nucleotide sequence has at least 90% (e.g., at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%) sequence identity to SEQ ID NO: 21. In certain embodiments, the transfer genome comprises a nucleotide sequence set forth in any one of SEQ ID NOs: 46-50, a 5' ITR nucleotide sequence having the sequence of SEQ ID NO: 20, and a 3' ITR nucleotide sequence having the sequence of SEQ ID NO: 21.

[0095] In certain embodiments, the 5' ITR nucleotide sequence and the 3' ITR nucleotide sequence are substantially complementary to each other (e.g., are complementary to each other except for mismatch at 1, 2, 3, 4, or 5 nucleotide positions in the 5' or 3' ITR).

[0096] In certain embodiments, the 5' ITR or the 3' ITR is modified to reduce or abolish resolution by Rep protein ("non-resolvable ITR"). In certain embodiments, the non-resolvable ITR comprises an insertion, deletion, or substitution in the nucleotide sequence of the terminal resolution site. Such modification allows formation of a self-complementary, double-stranded DNA genome of the AAV after the transfer genome is replicated in an infected cell. Exemplary non-resolvable ITR sequences are known in the art (see e.g., those provided in U.S. Patent Nos. 7,790,154 and 9,783,824, which are incorporated by reference herein in their entirety). In certain embodiments, the 5' ITR comprises a nucleotide sequence at least 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 26. In certain embodiments, the 5' ITR consists of a nucleotide sequence at least 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 26. In certain embodiments, the 5' ITR consists of the nucleotide sequence set forth in SEQ ID NO: 26. In certain embodiments, the 3' ITR comprises a nucleotide sequence at least 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 27. In certain embodiments, the 5' ITR consists of a nucleotide sequence at least 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 27. In certain embodiments, the 3' ITR consists of the nucleotide sequence set forth in SEQ ID NO: 27. In certain embodiments, the 5' ITR consists of the nucleotide sequence set forth in SEQ ID NO: 26, and the 3' ITR consists of the nucleotide sequence set forth in SEQ ID NO: 27. In certain embodiments, the 5' ITR consists of the nucleotide sequence set forth in SEQ ID NO: 26, and the 3' ITR consists of the nucleotide sequence set forth in SEQ ID NO: 19.

[0097] In certain embodiments, the 3' ITR is flanked by an additional nucleotide sequence derived from a wild-type AAV2 genomic sequence. In certain embodiments, the 3' ITR is flanked by an additional 37 bp sequence derived from a wild-type AAV2 sequence that is adjacent to a wild-type AAV2 ITR. *See, e.g., Savy et al., Human Gene Therapy Methods (2017) 28(5): 277-289* (which is hereby incorporated by reference herein in its entirety). In certain embodiments, the additional 37 bp sequence is internal to the 3' ITR. In certain embodiments, the 37 bp sequence consists of the sequence set forth in SEQ ID NO: 56. In certain embodiments, the 3' ITR comprises a nucleotide sequence at least 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 57. In certain embodiments, the 3' ITR comprises the nucleotide sequence set forth in SEQ ID NO: 57. In certain embodiments, the nucleotide sequence of the 3' ITR consists of a nucleotide sequence at least 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 57. In certain embodiments, the nucleotide sequence of the 3' ITR consists of the nucleotide sequence set forth in SEQ ID NO: 57.

[0098] In certain embodiments, the transfer genome comprises from 5' to 3': a 5' ITR; an internal element comprising from 5' to 3': a TRE, optionally a non-coding exon and an intron, a PAH coding sequence, and a polyadenylation sequence, as disclosed herein; a non-resolvable ITR; a nucleotide sequence complementary to the internal element; and a 3' ITR. Such transfer genome can form a self-complementary, double-stranded DNA genome of the AAV after infection and before replication.

[0099] In certain embodiments, the transfer genome comprises from 5' to 3': a 5' ITR, a TRE, optionally a non-coding exon and an intron, a PAH coding sequence, a polyadenylation sequence, and a 3' ITR. In certain embodiments, the 5' ITR has at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to SEQ ID: 18, 20, or 26; the TRE has at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to any one of SEQ ID NOs: 28-30 and 32-41; the intron has at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to SEQ ID NO: 31 or 35; the PAH coding sequence has at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to SEQ ID NO: 25; the polyadenylation sequence has at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to any one of SEQ ID NOs: 42, 43, and 45; and/or the 3' ITR has at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to SEQ ID: 19, 21, or 27. In certain embodiments, the 5' ITR comprises or consists of a nucleotide sequence selected from the group consisting of SEQ ID NOs: 18, 20, and 26; the TRE comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 28-30 and 32-41; the intron comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 31 and 35; the PAH coding sequence comprises the sequence set forth in SEQ ID NO: 25; the polyadenylation sequence comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 42, 43, and 45; and/or the 3' ITR comprises or consists of a nucleotide sequence selected from the group consisting of SEQ ID NOs: 19, 21, and 27. In certain embodiments, the 5' ITR comprises or consists of the sequence set forth in SEQ ID NO: 18; the TRE comprises the sequence set forth in SEQ ID NO: 28; the PAH coding sequence comprises the sequence set forth in SEQ ID NO: 25; the polyadenylation sequence comprises the sequence set forth in SEQ ID NO: 42; and/or the 3' ITR comprises or consists of the sequence set forth in SEQ ID NO: 19. In certain embodiments, the 5' ITR comprises or consists of the sequence set forth in SEQ ID NO: 26; the TRE comprises from 5' to 3' the sequence set forth in SEQ ID NO: 29, and the sequence set forth in SEQ ID NO: 30 (e.g., the TRE comprises the sequence set forth in SEQ ID NO: 32); the intron comprises the sequence set forth in SEQ ID

NO: 31; the PAH coding sequence comprises the sequence set forth in SEQ ID NO: 25; the polyadenylation sequence comprises the sequence set forth in SEQ ID NO: 43 and/or the 3' ITR comprises or consists of the sequence set forth in SEQ ID NO: 27. In certain embodiments, the 5' ITR comprises or consists of the sequence set forth in SEQ ID NO: 18; the TRE comprises from 5' to 3' the sequence set forth in SEQ ID NO: 33, and the sequence set forth in SEQ ID NO: 34 (e.g., the TRE comprises the sequence set forth in SEQ ID NO: 36); the intron comprises the sequence set forth in SEQ ID NO: 35; the PAH coding sequence comprises the sequence set forth in SEQ ID NO: 25; the polyadenylation sequence comprises the sequence set forth in SEQ ID NO: 45; and/or the 3' ITR comprises or consists of the sequence set forth in SEQ ID NO: 19. In certain embodiments, the 5' ITR comprises or consists of the sequence set forth in SEQ ID NO: 18; the TRE comprises from 5' to 3' the sequence set forth in SEQ ID NO: 37, and the sequence set forth in SEQ ID NO: 38 (e.g., the TRE comprises the sequence set forth in SEQ ID NO: 39); the intron comprises the sequence set forth in SEQ ID NO: 35; the PAH coding sequence comprises the sequence set forth in SEQ ID NO: 25; the polyadenylation sequence comprises the sequence set forth in SEQ ID NO: 45; and/or the 3' ITR comprises or consists of the sequence set forth in SEQ ID NO: 19. In certain embodiments, the 5' ITR comprises or consists of the sequence set forth in SEQ ID NO: 18; the TRE comprises from 5' to 3' the sequence set forth in SEQ ID NO: 37, and the sequence set forth in SEQ ID NO: 40 (e.g., the TRE comprises the sequence set forth in SEQ ID NO: 41); the polyadenylation sequence comprises the sequence set forth in SEQ ID NO: 45; and/or the 3' ITR comprises or consists of the sequence set forth in SEQ ID NO: 19.

[00100] In certain embodiments, the transfer genome comprises a sequence at least 80% (e.g., at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%) identical to the sequence set forth in SEQ ID NO: 51, 52, 53, 54, 55, 62, 65, 68, 75, 77, 79, 81, 83, 85, 87, or 90. In certain embodiments, the transfer genome comprises the sequence set forth in SEQ ID NO: 51, 52, 53, 54, 55, 62, 65, 68, 75, 77, 79, 81, 83, 85, 87, or 90. In certain embodiments, the transfer genome comprises the sequence set forth in SEQ ID NO: 52. In certain embodiments, the transfer genome consists of the sequence set forth in SEQ ID NO: 51, 52, 53, 54, 55, 62, 65, 68, 75, 77, 79, 81, 83, 85, 87, or 90. In certain embodiments, the transfer genome consists of the sequence set forth in SEQ ID NO: 52.

[00101] In certain embodiments, the replication-defective AAV comprises: (a) an AAV capsid protein comprising the amino acid sequence of amino acids 203-736 of SEQ ID NO: 16, and a transfer genome comprising 5' to 3' following genetic elements: a 5' ITR element (e.g., the 5' ITR of SEQ ID NOs: 26), a human HCR1 (e.g., the HCR1 of SEQ ID

NOs: 29), an hAAT promoter (e.g., the hAAT promoter of SEQ ID NOs: 30), an SV40 intron (e.g., the SV40 intron of SEQ ID NOs: 31), a silently altered human PAH coding sequence (e.g., the PAH coding sequence of SEQ ID NOs: 25), an SV40 polyadenylation sequence (e.g., the SV40 polyadenylation sequence of SEQ ID NOs: 43), and a 3' ITR element (e.g., the 3' ITR of SEQ ID NOs: 27); (b) an AAV capsid protein comprising the amino acid sequence of amino acids 138-736 of SEQ ID NO: 16, and a transfer genome comprising 5' to 3' following genetic elements: a 5' ITR element (e.g., the 5' ITR of SEQ ID NOs: 26), a human HCR1 (e.g., the HCR1 of SEQ ID NOs: 29), an hAAT promoter (e.g., the hAAT promoter of SEQ ID NOs: 30), an SV40 intron (e.g., the SV40 intron of SEQ ID NOs: 31), a silently altered human PAH coding sequence (e.g., the PAH coding sequence of SEQ ID NOs: 25), an SV40 polyadenylation sequence (e.g., the SV40 polyadenylation sequence of SEQ ID NOs: 43), and a 3' ITR element (e.g., the 3' ITR of SEQ ID NOs: 27; and/or (c) an AAV capsid protein comprising the amino acid sequence of SEQ ID NO: 16, and a transfer genome comprising 5' to 3' following genetic elements: a 5' ITR element (e.g., the 5' ITR of SEQ ID NOs: 26), a human HCR1 (e.g., the HCR1 of SEQ ID NOs: 29), an hAAT promoter (e.g., the hAAT promoter of SEQ ID NOs: 30), an SV40 intron (e.g., the SV40 intron of SEQ ID NOs: 31), a silently altered human PAH coding sequence (e.g., the PAH coding sequence of SEQ ID NOs: 25), an SV40 polyadenylation sequence (e.g., the SV40 polyadenylation sequence of SEQ ID NOs: 43), and a 3' ITR element (e.g., the 3' ITR of SEQ ID NOs: 27.

[00102] In certain embodiments, the replication-defective AAV comprises: (a) an AAV capsid protein comprising the amino acid sequence of amino acids 203-736 of SEQ ID NO: 16, and a transfer genome comprising the nucleotide sequence set forth in any one of SEQ ID NOs: 24, 25, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 61, 62, 64, 65, 67, 68, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 89, or 90; (b) an AAV capsid protein comprising the amino acid sequence of amino acids 138-736 of SEQ ID NO: 16, and a transfer genome comprising the nucleotide sequence set forth in any one of SEQ ID NOs: 24, 25, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 61, 62, 64, 65, 67, 68, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 89, or 90; and/or (c) an AAV capsid protein comprising the amino acid sequence of SEQ ID NO: 16, and a transfer genome comprising the nucleotide sequence set forth in any one of SEQ ID NOs: 24, 25, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 61, 62, 64, 65, 67, 68, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 89, or 90.

[00103] In another aspect, provided herein is a polynucleotide comprising a nucleic acid sequence that is at least 80% (e.g., at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%) identical to the nucleic acid sequence set forth in SEQ ID NO: 88, 91, or

92. In certain embodiments, the polynucleotide comprises the nucleic acid sequence set forth in SEQ ID NO: 88, 91, or 92. In certain embodiments, the polynucleotide comprises the nucleic acid sequence set forth in SEQ ID NO: 92. In certain embodiments, the polynucleotide consists of the nucleic acid sequence set forth in SEQ ID NO: 88, 91, or 92. In certain embodiments, the polynucleotide consists of the nucleic acid sequence set forth in SEQ ID NO: 92.

[00104] In another aspect, the instant disclosure provides pharmaceutical compositions comprising an AAV as disclosed herein together with a pharmaceutically acceptable excipient, adjuvant, diluent, vehicle or carrier, or a combination thereof. A "pharmaceutically acceptable carrier" includes any material which, when combined with an active ingredient of a composition, allows the ingredient to retain biological activity and without causing disruptive physiological reactions, such as an unintended immune reaction. Pharmaceutically acceptable carriers include water, phosphate buffered saline, emulsions such as oil/water emulsion, and wetting agents. Compositions comprising such carriers are formulated by well-known conventional methods such as those set forth in Remington's Pharmaceutical Sciences, current Ed., Mack Publishing Co., Easton Pa. 18042, USA; A. Gennaro (2000) "Remington: The Science and Practice of Pharmacy", 20th edition, Lippincott, Williams, & Wilkins; Pharmaceutical Dosage Forms and Drug Delivery Systems (1999) H. C. Ansel et al, 7th ed., Lippincott, Williams, & Wilkins; and Handbook of Pharmaceutical Excipients (2000) A. H. Kibbe et al, 3rd ed. Amer. Pharmaceutical Assoc.

[00105] In another aspect, the instant disclosure provides a polynucleotide comprising a coding sequence encoding a human PAH protein or a fragment thereof, wherein the coding sequence has been codon-altered to have less than 100% (e.g., less than 95%, 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, or 50%) identical to a wild-type human PAH gene. In certain embodiments, the polynucleotide comprises the sequence set forth in SEQ ID NO: 25. In certain embodiments, the polynucleotide comprises nucleotides 4 to 1359 of the sequence set forth in SEQ ID NO: 25. The polynucleotide can comprise DNA, RNA, modified DNA, modified RNA, or a combination thereof. In certain embodiments, the polynucleotide is an expression vector.

III. Method of Use

[00106] In another aspect, the instant disclosure provides methods for expressing a PAH polypeptide in a cell. The methods generally comprise transducing the cell with a replication-defective AAV as disclosed herein. Such methods are highly efficient at restoring PAH

expression. Accordingly, in certain embodiments, the methods disclosed herein involve transducing the cell with a replication-defective AAV as disclosed herein.

[00107] The methods disclosed herein can be applied to any cell harboring a mutation in the PAH gene. The skilled worker will appreciate that cells that are active in Phe metabolism are of particular interest. Accordingly, in certain embodiments, the method is applied to cells in the liver, kidney, brain, pituitary gland, adrenal gland, pancreas, urinary bladder, gallbladder, colon, small intestine, or breast. In certain embodiments, the method is applied to hepatocytes and/or renal cells.

[00108] The methods disclosed herein can be performed *in vitro* for research purposes or can be performed *ex vivo* or *in vivo* for therapeutic purposes.

[00109] In certain embodiments, the cell to be transduced is in a mammalian subject and the AAV is administered to the subject in an amount effective to transduce the cell in the subject. Accordingly, in certain embodiments, the instant disclosure provides a method for treating a subject having a disease or disorder associated with a PAH gene mutation, the method generally comprising administering to the subject an effective amount of a replication-defective AAV as disclosed herein. The subject can be a human subject, a non-human primate subject (e.g., a cynomolgus), or a rodent subject (e.g., a mouse) with a PAH mutation, or a non-human primate subject (e.g., a cynomolgus) or a rodent subject (e.g., a mouse) containing PAH-mutant human liver cells. Suitable mouse subjects include without limitation, mice into which human liver cells (e.g., human hepatocytes) have been engrafted. Any disease or disorder associated with a PAH gene mutation can be treated using the methods disclosed herein. Suitable diseases or disorders include, without limitation, phenylketonuria.

[00110] In certain embodiments, the foregoing methods employ a replication-defective AAV comprising: (a) an AAV capsid protein comprising the amino acid sequence of amino acids 203-736 of SEQ ID NO: 16, and a transfer genome comprising 5' to 3' following genetic elements: a 5' ITR element (e.g., the 5' ITR of SEQ ID NOs: 26), a human HCR1 (e.g., the HCR1 of SEQ ID NOs: 29), an hAAT promoter (e.g., the hAAT promoter of SEQ ID NOs: 30), an SV40 intron (e.g., the SV40 intron of SEQ ID NOs: 31), a silently altered human PAH coding sequence (e.g., the PAH coding sequence of SEQ ID NOs: 25), an SV40 polyadenylation sequence (e.g., the SV40 polyadenylation sequence of SEQ ID NOs: 43), and a 3' ITR element (e.g., the 3' ITR of SEQ ID NOs: 27); (b) an AAV capsid protein comprising the amino acid sequence of amino acids 138-736 of SEQ ID NO: 16, and a transfer genome comprising 5' to 3' following genetic elements: a 5' ITR element (e.g., the 5' ITR of SEQ ID NOs: 26), a human HCR1 (e.g., the HCR1 of SEQ ID NOs: 29), an hAAT

promoter (e.g., the hAAT promoter of SEQ ID NOs: 30), an SV40 intron (e.g., the SV40 intron of SEQ ID NOs: 31), a silently altered human PAH coding sequence (e.g., the PAH coding sequence of SEQ ID NOs: 25), an SV40 polyadenylation sequence (e.g., the SV40 polyadenylation sequence of SEQ ID NOs: 43), and a 3' ITR element (e.g., the 3' ITR of SEQ ID NOs: 27; and/or (c) an AAV capsid protein comprising the amino acid sequence of SEQ ID NO: 16, and a transfer genome comprising 5' to 3' following genetic elements: a 5' ITR element (e.g., the 5' ITR of SEQ ID NOs: 26), a human HCR1 (e.g., the HCR1 of SEQ ID NOs: 29), an hAAT promoter (e.g., the hAAT promoter of SEQ ID NOs: 30), an SV40 intron (e.g., the SV40 intron of SEQ ID NOs: 31), a silently altered human PAH coding sequence (e.g., the PAH coding sequence of SEQ ID NOs: 25), an SV40 polyadenylation sequence (e.g., the SV40 polyadenylation sequence of SEQ ID NOs: 43), and a 3' ITR element (e.g., the 3' ITR of SEQ ID NOs: 27.

[00111] In certain embodiments, the foregoing methods employ a replication-defective AAV comprising: (a) an AAV capsid protein comprising the amino acid sequence of amino acids 203-736 of SEQ ID NO: 16, and a transfer genome comprising the nucleotide sequence set forth in any one of SEQ ID NOs: 24, 25, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 61, 62, 64, 65, 67, 68, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 89, or 90; (b) an AAV capsid protein comprising the amino acid sequence of amino acids 138-736 of SEQ ID NO: 16, and a transfer genome comprising the nucleotide sequence set forth in any one of SEQ ID NOs: 24, 25, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 61, 62, 64, 65, 67, 68, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 89, or 90; and/or (c) an AAV capsid protein comprising the amino acid sequence of SEQ ID NO: 16, and a transfer genome comprising the nucleotide sequence set forth in any one of SEQ ID NOs: 24, 25, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 61, 62, 64, 65, 67, 68, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 89, or 90.

[00112] The methods disclosed herein are particularly advantageous in that they are capable of expressing a PAH protein in a cell with high efficiency both *in vivo* and *in vitro*. In certain embodiments, the expression level of the PAH protein is at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of the expression level of the endogenous PAH protein in a cell of the same type that does not have a mutation in the PAH gene. In certain embodiments, the expression level of the PAH protein is at least 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 3, 4, 5, 6, 7, 8, 9, or 10 fold higher than the expression level of the endogenous PAH protein in a cell of the same type that does not have a mutation in the PAH gene. Any methods of determining the expression level of the PAH protein can be employed including, without limitation, ELISA, Western blotting, immunostaining, and mass spectrometry.

[00113] In certain embodiments, transduction of a cell with an AAV composition disclosed herein can be performed as provided herein or by any method of transduction known to one of ordinary skill in the art. In certain embodiments, the cell may be contacted with the AAV at a multiplicity of infection (MOI) of 50,000; 100,000; 150,000; 200,000; 250,000; 300,000; 350,000; 400,000; 450,000; or 500,000, or at any MOI that provides for optimal transduction of the cell.

[00114] An AAV composition disclosed herein can be administered to a subject by any appropriate route including, without limitation, intravenous, intraperitoneal, subcutaneous, intramuscular, intranasal, topical or intradermal routes. In certain embodiments, the composition is formulated for administration via intravenous injection or subcutaneous injection.

IV. AAV Packaging Systems

[00115] In another aspect, the instant disclosure provides packaging systems for recombinant preparation of a replication–defective AAV disclosed herein. Such packaging systems generally comprise: a Rep nucleotide sequence encoding one or more AAV Rep proteins; a Cap nucleotide sequence encoding one or more AAV Clade F capsid proteins as disclosed herein; and a transfer genome for expression of the PAH gene as disclosed herein, wherein the packaging system is operative in a cell for enclosing the transfer genome in the capsid to form the AAV.

[00116] In certain embodiments, the packaging system comprises a first vector comprising the Rep nucleotide sequence and the Cap nucleotide sequence, and a second vector comprising the transfer genome. As used in the context of a packaging system as described herein, a "vector" refers to a nucleic acid molecule that is a vehicle for introducing nucleic acids into a cell (e.g., a plasmid, a virus, a cosmid, an artificial chromosome, etc.).

[00117] Any AAV Rep protein can be employed in the packaging systems disclosed herein. In certain embodiments of the packaging system, the Rep nucleotide sequence encodes an AAV2 Rep protein. Suitable AAV2 Rep proteins include, without limitation, Rep 78/68 or Rep 68/52. In certain embodiments of the packaging system, the nucleotide sequence encoding the AAV2 Rep protein comprises a nucleotide sequence that encodes a protein having a minimum percent sequence identity to the AAV2 Rep amino acid sequence of SEQ ID NO: 22, wherein the minimum percent sequence identity is at least 70% (e.g., at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, or 100%) across the length of the amino acid sequence of the AAV2 Rep protein. In certain embodiments of the

packaging system, the AAV2 Rep protein has the amino acid sequence set forth in SEQ ID NO: 22.

[00118] In certain embodiments of the packaging system, the packaging system further comprises a third vector, e.g., a helper virus vector. The third vector may be an independent third vector, integral with the first vector, or integral with the second vector. In certain embodiments, the third vector comprises genes encoding helper virus proteins.

[00119] In certain embodiments of the packaging system, the helper virus is selected from the group consisting of adenovirus, herpes virus (including herpes simplex virus (HSV)), poxvirus (such as vaccinia virus), cytomegalovirus (CMV), and baculovirus. In certain embodiments of the packaging system, where the helper virus is adenovirus, the adenovirus genome comprises one or more adenovirus RNA genes selected from the group consisting of El, E2, E4 and VA. In certain embodiments of the packaging system, where the helper virus is HSV, the HSV genome comprises one or more of HSV genes selected from the group consisting of UL5/8/52, ICPO, ICP4, ICP22 and UL30/UL42.

[00120] In certain embodiments of the packaging system, the first, second, and/or third vector are contained within one or more transfecting plasmids. In certain embodiments, the first vector and the third vector are contained within a first transfecting plasmid. In certain embodiments the second vector and the third vector are contained within a second transfecting plasmid.

[00121] In certain embodiments of the packaging system, the first, second, and/or third vector are contained within one or more recombinant helper viruses. In certain embodiments, the first vector and the third vector are contained within a recombinant helper virus. In certain embodiments, the second vector and the third vector are contained within a recombinant helper virus.

[00122] In a further aspect, the disclosure provides a method for recombinant preparation of an AAV as described herein, wherein the method comprises transfecting or transducing a cell with a packaging system as described under conditions operative for enclosing the transfer genome in the capsid to form the AAV as described herein. Exemplary methods for recombinant preparation of an AAV include transient transfection (e.g., with one or more transfection plasmids containing a first, and a second, and optionally a third vector as described herein), viral infection (e.g. with one or more recombinant helper viruses, such as a adenovirus, poxvirus (such as vaccinia virus), herpes virus (including HSV, cytomegalovirus, or baculovirus, containing a first, and a second, and optionally a third vector as described herein), and stable producer cell line transfection or infection (e.g., with a stable producer cell,

such as a mammalian or insect cell, containing a Rep nucleotide sequence encoding one or more AAV Rep proteins and/or a Cap nucleotide sequence encoding one or more AAV Clade F capsid proteins as described herein, and with a transfer genome as described herein being delivered in the form of a transfecting plasmid or a recombinant helper virus).

V. Examples

[00123] The recombinant AAV vectors disclosed herein mediate highly efficient gene transfer *in vitro* and *in vivo*. The following examples demonstrate the efficient restoration of the expression of the PAH gene, which is mutated in certain human diseases, such as phenylketonuria, using an AAV-based vector as disclosed herein. These examples are offered by way of illustration, and not by way of limitation.

Example 1: Human PAH transfer vector

[00124] This example provides human PAH transfer vectors pHMI-hPAH-TC-004, pHMI-hPAH-TC-025, pHMI-hPAH-TC-010, pHMI-hPAH-TC-011, and pHMI-hPAH-TC-012 for expression of human PAH in a human or mouse cell.

a) pHMI-hPAH-TC-004

[00125] PAH transfer vector pHMI-hPAH-TC-004, as shown in Figure 1A, comprises 5' to 3' the following genetic elements: a 5' ITR element, a CAG promoter, a silently altered human PAH coding sequence, an SV40 polyadenylation sequence, and a 3' ITR element. The sequences of these elements are set forth in Table 1. This vector is capable of expressing a human PAH protein in a cell (e.g., a human cell or a mouse cell) to which the vector is transduced.

Table 1: Genetic elements in human PAH transfer vector pHMI-hPAH-TC-004

Genetic Element	SEQ ID NO
5' ITR element	18
CAG promoter	28
codon-altered human PAH coding sequence	25
SV40 polyadenylation sequence	42
3' ITR element	19
Transfer genome (from promoter to polyadenylation sequence)	46
Transfer genome (from 5' ITR to 3' ITR)	51

b) pHMI-hPAH-TC-025

[00126] PAH transfer vector pHMI-hPAH-TC-025, as shown in Figure 1B, comprises 5' to 3' the following genetic elements: a truncated 5' ITR element, a human hepatic control region 1 (HCR1), a human α1-antitrypsin (hAAT) promoter, an SV40 intron, a silently altered human PAH coding sequence, an SV40 polyadenylation sequence, and a modified 3' ITR element. The sequences of these elements are set forth in Table 2. The truncated 5' ITR allows the vector to form a double-stranded AAV genome after transduction into cells. This vector is capable of expressing a human PAH protein in a human hepatocyte to which the vector is transduced.

Table 2: Genetic elements in human PAH transfer vector pHMI-hPAH-TC-025

Genetic Element	SEQ ID NO
truncated 5' ITR element	26
human HCR1	29
human α1-antitrypsin (hAAT) promoter	30
SV40 intron	31
transcriptional regulatory region comprising the human HCR1 and hAAT promoter	32
codon-altered human PAH coding sequence	25
SV40 polyadenylation sequence	43
modified 3' ITR element	27
Transfer genome (from HCR1 to polyadenylation sequence)	47
Transfer genome (from 5' ITR to 3' ITR)	52
Full sequence of transfer vector	92

c) pHMI-hPAH-TC-010

[00127] PAH transfer vector pHMI-hPAH-TC-010, as shown in Figure 1C, comprises 5' to 3' the following genetic elements: a 5' ITR element, a hepatic specific regulatory module of hAAT promoter, a human TTR promoter, a modified minute virus of mouse (MVM) intron, a silently altered human PAH coding sequence, an SV40 polyadenylation sequence, and a 3' ITR element. The sequences of these elements are set forth in Table 3. This vector is capable of expressing a human PAH protein in a cell (e.g., a human cell or a mouse cell) to which the vector is transduced, particularly at a high level in a hepatocyte.

Table 3: Genetic elements in human PAH transfer vector pHMI-hPAH-TC-010

Genetic Element	SEQ ID NO
5' ITR element	18
hepatic specific regulatory module of hAAT promoter	33
human TTR promoter	34
modified minute virus of mouse (MVM) intron	35
transcriptional regulatory region comprising the hepatic specific regulatory module (HSRM) and human TTR promoter	36
codon-altered human PAH coding sequence	25
SV40 polyadenylation sequence	45
3' ITR element	19
Transfer genome (from HSRM to polyadenylation sequence)	48
Transfer genome(from 5' ITR to 3' ITR)	53

d) pHMI-hPAH-TC-011

[00128] PAH transfer vector pHMI-hPAH-TC-011, as shown in Figure 1D, comprises 5' to 3' the following genetic elements: a 5' ITR element, a human HCR1, a human α1-antitrypsin (hAAT) promoter, an modified MVM intron, a silently altered human PAH coding sequence, an SV40 polyadenylation sequence, and a 3' ITR element. The sequences of these elements are set forth in Table 4. This vector is capable of expressing a human PAH protein in a human hepatocyte to which the vector is transduced.

Table 4: Genetic elements in human PAH transfer vector pHMI-hPAH-TC-011

Genetic Element	SEQ ID NO
truncated 5' ITR element	18
human HCR1	37
human α1-antitrypsin (hAAT) promoter	38
modified minute virus of mouse (MVM) intron	35
transcriptional regulatory region comprising the human HCR1 and hAAT promoter	39
codon-altered human PAH coding sequence	25
SV40 polyadenylation sequence	45

Genetic Element	SEQ ID NO
modified 3' ITR element	19
Transfer genome (from HCR1 to polyadenylation sequence)	49
Transfer genome(from 5' ITR to 3' ITR)	54

e) pHMI-hPAH-TC-012

[00129] PAH transfer vector pHMI-hPAH-TC-012, as shown in Figure 1E, comprises 5' to 3' the following genetic elements: a 5' ITR element, a human hepatic control region 1 (HCR1), a modified human EF-1α promoter, a silently altered human PAH coding sequence, an SV40 polyadenylation sequence, and a 3' ITR element. The sequences of these elements are set forth in Table 5. This vector is capable of expressing a human PAH protein in a human hepatocyte to which the vector is transduced.

Table 5: Genetic elements in human PAH transfer vector pHMI-hPAH-TC-012

Genetic Element	SEQ ID NO
truncated 5' ITR element	18
human hepatic control region 1 (HCR1)	37
modified human EF-1α promoter	40
transcriptional regulatory region comprising the human HCR1 and modified human EF-1α promoter	41
codon-altered human PAH coding sequence	25
SV40 polyadenylation sequence	45
modified 3' ITR element	19
Transfer genome (from HCR1 to polyadenylation sequence)	50
Transfer genome (from 5' ITR to 3' ITR)	55

[00130] The silent alteration significantly improves the expression of the PAH protein, as demonstrated by comparison of expression vectors pCOH-WT-PAH, pCOH-CO-PAH, and pHMI-CO-PAH. The pCOH-WT-PAH vector comprises a CAG promoter operably linked to a wild-type PAH coding sequence set forth in SEQ ID NO: 24. The pCOH-CO-PAH and pHMI-CO-PAH vectors each comprise a CAG promoter operably linked to a codon-altered human PAH coding sequence as set forth in SEQ ID NO: 25. The pCOH-CO-PAH and pHMI-CO-PAH vectors are highly similar. Each vector was transfected in HEK 293 cells

which is naturally deficient in PAH. As shown in Figure 2, VG-GT-CO-PAH ("CO-hPAH") gave rise to an expression level of human PAH several fold higher than VG-GT-PAH ("WT-hPAH").

[00131] The vectors disclosed herein can be packaged in an AAV clade F capsid, such as an AAVHSC5, AAVHSC7, AAVHSC15 or AAVHSC17 capsid. The packaged viral particles can be administered to a wild-type animal, a PAH deficient animal, or a reconstituted animal having human hepatocytes obtained from a patient with phenylketonuria caused by a PAH mutation. The gene transfer efficiency can be measured by collecting liver samples and quantifying the percentage of PAH-positive cells (e.g., cells that have a unique nucleotide sequence from the vector, cells that express a wild-type PAH protein, or cells with a higher PAH activity than in cells from a control animal not receiving the PAH expression vector). The restoration of phenylalanine metabolism, which indicates the efficacy of the PAH expression vectors, can be assessed by measuring the Phe level in the blood and by observing the coat color of the mouse. Safety of the viral particle administration can be evaluated by measuring the aspartate aminotransferase (AST) and alanine aminotransferase (ALT) levels in serum.

Example 2: Mouse PAH gene transfer in a mouse model

[00132] This example provides a mouse PAH transfer vector rAAV-CBA-mPAH that is similar to the human PAH transfer vector pHMI-hPAH-TC-004 described in Example 1 except that a wild-type mouse PAH coding sequence is substituted for the codon-altered human PAH coding sequence. This vector is capable of expressing a mouse PAH protein in a cell (e.g., a human cell or a mouse cell) to which the vector is transduced.

[00133] Briefly, *Pah*$^{-/-}$ (*PAH*^{enu2}) mice were housed in clear polycarbonate cages with contact bedding in an isolator. Picolab Mouse Diet 5058 was provided to the animals *ad libitum*. Spring or tap water acidified with 1N HCl to a targeted pH of 2.5 - 3.0 was provided *ad libitum*. Vectors packaged in AAVHSC15 capsid were prepared in PBS (with Ca and Mg), supplemented with 35 mM NaCl, 1% sucrose, and 0.05% Pluronic F-68. The formulation was injected intravenously via the tail vein.

[00134] Blood samples were collected every week after the administration of the PAH transfer vector (0 week: prior to administration) by facial vein puncture or tail snip. The samples were allowed to clot at room temperature for at least 30 minutes, centrifuged at ambient temperature at minimum 1000 x g for 10 minutes and the serum samples were extracted. Serum samples were stored at -70 °C. Serum phenylalanine and tyrosine levels were measured by tandem mass spectrometry.

[00135] For collection of tissue samples, the animals underwent cardiac perfusion with saline. Liver (caudate lobe), kidney (left), brain, heart, and muscle (quadriceps) tissues were snap frozen in liquid nitrogen and stored at -70 °C. The snap frozen tissues were ground into powder in liquid nitrogen in a mortar and pestle and divided in to aliquots to test for PAH expression for vector genome biodistribution by qPCR.

[00136] The safety of the rAAV-CBA-mPAH vector was assessed by measuring the aspartate aminotransferase (AST) and alanine aminotransferase (ALT) levels in the treated animals. Serum samples were collected pre-dose and one week after administration of the viral particles. The levels of AST and ALT were measured by the Sigma MAK055 and Sigma MAK052 ELISA kits.

[00137] The *pah*$^{-/-}$ mice manifested phenylketonuria and had lighter coat color than wild-type mice. As shown in Figures 3A and 3B, the administration of the rAAV-CBA-mPAH vector lead to significant reduction of Phe levels in the serum within one week, and the Phe levels remained low for four weeks. The coat color also changed from brown to black within one week.

[00138] Expression of mPAH was also observed in tissue samples. As shown in Figure 4, DNA of the rAAV-CBA-mPAH vector was detectable in many organs, wherein the numbers of viral genomes per 10^6 cells was the highest in liver, heart, and kidney.

[00139] With respect to the safety of the AAV administration, the AST and ALT levels remained low after administration (Figures 5A and 5B), suggesting that the rAAV-CBA-mPAH vector was not toxic to the liver.

Example 3: Human PAH gene transfer in a mouse model

[00140] This example demonstrates that the PAH transfer vectors described in Example 1 effectively reversed the phenotype caused by PAH gene deficiency in a mouse model. The mouse model, AAV packaging and formulation, and methods for examining gene transfer efficiency were identical to the ones described in Example 2.

[00141] To examine the efficacy of the five PAH transfer vectors in reversing the phenotypes, a single dose of 2.6×10^{13} vector genomes per kg of body weight for male mice, or a dose of 6×10^{13} vector genomes per kg of body weight for female mice. The *pah*$^{-/-}$ mice manifested increased level of phenylalanine (Phe) and reduced level of tyrosine (Tyr) in the serum. As shown in Figures 6A-6H, the administration of any one of the five vectors led to significant reduction of Phe levels and increase of Tyr levels within one week. The efficacy lasted for at least 12 weeks in male mice, and at least 6 weeks in female mice. Other than

pHMI-hPAH-TC-004, all the vectors maintained complete reduction of serum Phe levels during the time examined.

[00142] To examine the long-term efficacy of pHMI-hPAH-TC-025 in reversing the phenotype caused by PAH gene deficiency, a single dose of 2.6×10^{13} vector genomes per kg of body weight was administered to male mice, or a single dose of 6×10^{13} vector genomes per kg of body weight was administered to female mice. As shown in Figures 6I and 6J, the administration of the pHMI-hPAH-TC-025 vector led to significant reduction of Phe levels within one week. This reduction persisted for at least 48 weeks in male mice, and at least 46 weeks in female mice. Additionally, within two weeks post administration of the AAV, the coat color of the mice administered with pHMI-hPAH-TC-004 changed from brown to black. An increase of PAH mRNA was observed by ddPCR in the liver samples of these mice collected 4 weeks post injection relative to the mice not administered with AAV vectors. An increase of the PAH enzymatic activity was also detected in liver samples by mass spectrometry.

[00143] The efficacy of different doses of the pHMI-hPAH-TC-025 vector was further assessed. A single dose of 2.6×10^{11}, 2.6×10^{12}, or 2.6×10^{13} vector genomes per kg of body weight was administered to male mice and female mice, and the serum levels of Phe and Tyr were measured. As shown in Figures 7A-7D, the dose of 2.6×10^{13} vector genomes per kg of body weight reduced the Phe levels and increased the Tyr levels more significantly than the two lower doses, and maintained complete reduction of serum Phe levels during the time examined in both male and female subjects.

Example 3: Additional human PAH transfer vectors

[00144] This example provides human PAH transfer vectors pHMI-hPAH-TC-009, pHMI-hPAH-TC-013, and pHMI-hPAH-TC-017 for expression of human PAH in a human or mouse cell. Vector maps are shown in **Figure 8A**, **8B**, and **8C**, respectively.

a) pHMI-hPAH-TC-009

[00145] PAH transfer vector pHMI-hPAH-TC-009, as shown in **Figure 8A**, comprises 5' to 3' the following genetic elements: a 5' ITR element, a CMV enhancer, a CBA promoter, a rabbit β-globin element, a human PAH coding sequence, a polyadenylation sequence, and a 3' ITR element. The sequences of these elements are set forth in Table 6. This vector is capable of expressing a human PAH protein in a cell (e.g., a human cell or a mouse cell) to which the vector is transduced.

Table 6: Genetic elements in human PAH transfer vector pHMI-hPAH-TC-009

Genetic Element	SEQ ID NO
5' ITR element	18
CMV enhancer	58
CBA promoter	59
Rabbit β-globin element	60
codon-altered human PAH coding sequence	25
Polyadenylation sequence	45
3' ITR element	19
Transfer genome (from CMV to polyadenylation sequence)	61
Transfer genome (from 5' ITR to 3' ITR)	62

b) pHMI-hPAH-TC-013

[00146] PAH transfer vector pHMI-hPAH-TC-013, as shown in **Figure 8B**, comprises 5' to 3' the following genetic elements: a 5' ITR element, a CASI promoter region (comprising a CMV enhancer, a CASI promoter, and a ubiquitin C enhancer element (hUBC exon)), a human PAH coding sequence, a polyadenylation sequence, and a 3' ITR element. The sequences of these elements are set forth in Table 7. This vector is capable of expressing a human PAH protein in a cell (e.g., a human cell or a mouse cell) to which the vector is transduced.

Table 7: Genetic elements in human PAH transfer vector pHMI-hPAH-TC-013

Genetic Element	SEQ ID NO
5' ITR element	18
CASI promoter region	63
Human PAH coding sequence	25
Polyadenylation sequence	45
3' ITR element	19
Transfer genome (from promoter region to polyadenylation sequence)	64
Transfer genome (from 5' ITR to 3' ITR)	65

f) pHMI-hPAH-TC-017

[00147] PAH transfer vector pHMI-hPAH-TC-017, as shown in **Figure 8C**, comprises 5' to 3' the following genetic elements: a 5' ITR element, an hAAT promoter region (comprising an ABMP enhancer (an enhancer region adjacent to a gene on chromosome 9 that expresses highly in liver, 5' to the ATG), a TTR enhancer, an hAAT promoter, and an MVM intron), a human PAH coding sequence, a polyadenylation sequence, and a 3' ITR element. The sequences of these elements are set forth in Table 8. This vector is capable of expressing a human PAH protein in a cell (e.g., a human cell or a mouse cell) to which the vector is transduced.

Table 8: Genetic elements in human PAH transfer vector pHMI-hPAH-TC-017

Genetic Element	SEQ ID NO
5' ITR element	18
hAAT promoter region	66
Human PAH coding sequence	25
Polyadenylation sequence	45
3' ITR element	19
Transfer genome (from promoter region to polyadenylation sequence)	67
Transfer genome (from 5' ITR to 3' ITR)	68

[00148] The vectors described in this example were tested for expression in two different cell lines. 5 x 10^5 HEK293 cells (kidney; non-liver) and 5 x 10^5 Huh7 cells (liver) were transfected with 1 ug each of the following vectors: pHMI-hPAH-TC-004 (PAH-004); pHMI-hPAH-TC-009 (PAH-009); pHMI-hPAH-TC-010 (PAH-010); pHMI-hPAH-TC-011 (PAH-011); pHMI-hPAH-TC-012 (PAH-012); pHMI-hPAH-TC-013 (PAH-013); pHMI-hPAH-TC-025 (LP1); pHMI-hPAH-TC-017 (PAH-017). Lysate of the cells was collected 48 hours after transfection. The expression of human PAH was detected by Western blotting with an anti-PAH antibody (Sigma HPA031642). The amount of GAPDH protein as detected by an anti-GAPDH antibody (Millipore MAB 374) was used as a loading control. PAH expression levels of all vectors were normalized to pHMI-hPAH-TC-004 expression level; data was collected from multiple independent transfections and plotted in **Figure 9**. **Figure 9A** shows the normalized PAH expression level of the indicated vectors in Huh7 cells. **Figure 9B** shows the normalized PAH expression level of the indicated vectors in HEK293 cells.

[00149] **Figure 10A and 10B** are graphs showing the serum phenylalanine levels over time of male and female homozygous *Pah*$^{-/-}$ PAH^{enu2} mice respectively. Male and female mice were dosed at 2e13 vg/kg and 6e13 vg/kg respectively with pHMI-hPAH-TC-010 (hPAH-TC-010), pHMI-hPAH-TC-025 (hPAH-TC-025), pHMI-hPAH-TC-004 (hPAH-TC-004), pHMI-hPAH-TC-011 (hPAH-TC-011), or pHMI-hPAH-TC-012 (hPAH-TC-012) vectors packaged in AAVHSC15 capsid. Serum samples were collected weekly then biweekly after the administration. Serum phenylalanine concentrations were assessed by LC-MS/MS. **Figure 10C** is a graph showing the average baseline serum phenylalanine level for the male and female homozygous *Pah*$^{-/-}$ PAH^{enu2} mice in the study. The data represents a total of 55 mice per group.

[00150] As shown in **Figure 10**, the administration of certain vectors led to significant reduction of Phe levels within one week of administration, and this reduction persisted for at least 45 weeks. **Figure 10** demonstrates that some of the PAH transfer vectors effectively reversed the phenotype caused by PAH gene deficiency in a mouse model. The mouse model, AAV packaging and formulation, and methods for examining gene transfer efficiency were identical to those described in Example 2 herein. The sizes of the AAV vectors were as follows: pHMI-hPAH-TC-010 (hPAH-TC-010): 2391 bp; pHMI-hPAH-TC-025 (hPAH-TC-025): 2351 bp; pHMI-hPAH-TC-004 (hPAH-TC-004): 3781 bp; pHMI-hPAH-TC-011 (hPAH-TC-011): 3158 bp; and pHMI-hPAH-TC-012 (hPAH-TC-012): 3799 bp.

Example 4: Additional human PAH transfer vectors

[00151] This example examines the effect of PAH gene CpG content on PAH protein expression, using the PAH transfer vectors pHMI-hPAH-TC-018, pHMI-hPAH-TC-019, pHMI-hPAH-TC-020, pHMI-hPAH-TC-021, pHMI-hPAH-TC-022, and pHMI-hPAH-TC-023. Vector maps are shown in **Figure 11A**, **11B**, **11C**, **11D**, **11E**, and **11F**, respectively. These PAH transfer vectors comprise the sequences and elements set forth in Table 9.

Table 9: Genetic elements in PAH transfer vectors

	pHMI-hPAH-TC-XXX Vector					
	SEQ ID NO					
Genetic Element	**-018**	**-019**	**-020**	**-021**	**-022**	**-023**
5' ITR element	26	26	26	26	26	26
HCR1	29	29	29	29	29	29
hAAT promoter	30	30	30	30	30	30
SV40 intron	31	31	31	31	31	31

	pHMI-hPAH-TC-XXX Vector					
	SEQ ID NO					
Genetic Element	**-018**	**-019**	**-020**	**-021**	**-022**	**-023**
PAH coding sequence	69	70	71	72	73	24
Late SV40 polyadenylation sequence	45	45	45	45	45	45
3' ITR element	27	27	27	27	27	27
Transfer genome (from HCR1 to polyadenylation sequence)	74	76	78	80	82	84
Transfer genome (from 5' ITR to 3' ITR)	75	77	79	81	83	85

[00152] The vectors described in this example were tested for expression in HEK293 cells but under the control of a CBA promoter. 5 x 10^5 HEK293 cells were transfected with 1 ug each of the following vectors: pHMI-hPAH-TC-004 (PAH-TC-004); pHMI-hPAH-TC-009 (PAH-TC-009); pHMI-hPAH-TC-018 (PAH-TC-018); pHMI-hPAH-TC-019 (PAH-TC-019); pHMI-hPAH-TC-020 (PAH-TC-020); pHMI-hPAH-TC-021 (PAH-TC-021); pHMI-hPAH-TC-022 (PAH-TC-022); pHMI-hPAH-TC-023 (PAH-TC-023). Lysate of the cells was collected 48 hours after transfection. The expression of human PAH was detected by Western blotting with an anti-PAH antibody (Sigma HPA031642). The amount of GAPDH protein as detected by an anti-GAPDH antibody (Millipore MAB 374) was used as a loading control. PAH expression levels of all vectors were normalized to pHMI-hPAH-TC-004 expression level; data was plotted in **Figure 12**.

[00153] **Figure 13** is a graph showing the serum phenylalanine levels over time of male homozygous *Pah*$^{-/-}$ PAH^{enu2} mice. Male mice have been dosed at 2e13 vg/kg with pHMI-hPAH-TC-018 (hPAH-TC-018); pHMI-hPAH-TC-019 (hPAH-TC-019); pHMI-hPAH-TC-020 (hPAH-TC-020); pHMI-hPAH-TC-021 (hPAH-TC-021); pHMI-hPAH-TC-022 (hPAH-TC-022); pHMI-hPAH-TC-023 (hPAH-TC-023); and pHMI-hPAH-TC-025 (hPAH-TC-025) vectors packaged in AAVHSC15 capsid. Serum samples were collected weekly after the administration. Serum phenylalanine concentration was assessed by LC-MS/MS.

[00154] As shown in **Figure 13**, the administration of certain vectors led to significant reduction of Phe levels within one week of administration, and this reduction persisted for at least 25 weeks. The mouse model, AAV packaging and formulation, and methods for examining gene transfer efficiency were identical to those previously described in Example 2 herein. The CpG content of the vectors were as follows: pHMI-hPAH-TC-018 (hPAH-TC-018): 2; pHMI-hPAH-TC-019 (hPAH-TC-019): 7; pHMI-hPAH-TC-020 (hPAH-TC-020): 22; pHMI-hPAH-TC-021 (hPAH-TC-021): 10; pHMI-hPAH-TC-022 (hPAH-TC-022): 7; pHMI-hPAH-TC-023 (hPAH-TC-023): 23; and pHMI-hPAH-TC-025 (hPAH-TC-025): 60.

Example 5: Alternative ITR human PAH transfer vectors

[00155] This example provides human PAH transfer vectors pHMI-01004 and pHMI-01008 for expression of human PAH in a human or mouse cell. Vector maps are shown in **Figure 14A** and **14B**, respectively. These PAH transfer vectors comprise the sequences and elements set forth in Table 10.

Table 10: Genetic elements in PAH transfer vectors pHMI-01004 and pHMI-01008

	pHMI-01004	**pHMI-01008**
Genetic Element	**SEQ ID NO**	**SEQ ID NO**
5' ITR element	26	26
HCR1	29	29
hAAT promoter region	30	30
SV40 intron	31	31
Human PAH coding sequence	25	25
Polyadenylation sequence	43	43
3' ITR element	27	57
Transfer genome (from HCR1 to polyadenylation sequence)	86	89
Transfer genome (from 5' ITR to 3' ITR)	87	90
Full sequence of transfer vector	88	91

* * *

[00156] The invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described will become apparent to those skilled in the art from the foregoing description and

accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

[00157] All references (e.g., publications or patents or patent applications) cited herein are incorporated herein by reference in their entirety and for all purposes to the same extent as if each individual reference (e.g., publication or patent or patent application) was specifically and individually indicated to be incorporated by reference in its entirety for all purposes. Other embodiments are within the following claims.

We claim:

1. A method for expressing a PAH polypeptide in a cell, the method comprising transducing the cell with a replication-defective adeno-associated virus (AAV) comprising:
(a) an AAV capsid comprising an AAV Clade F capsid protein; and
(b) a transfer genome comprising a transcriptional regulatory element operably linked to a PAH coding sequence.

2. The method of claim 1, wherein the cell is a hepatocyte, a renal cell, or a cell in the brain, pituitary gland, adrenal gland, pancreas, urinary bladder, gallbladder, colon, small intestine, or breast.

3. The method of claim 1 or 2, wherein the cell is in a mammalian subject and the AAV is administered to the subject in an amount effective to transduce the cell in the subject.

4. A method for treating a subject having a disease or disorder associated with a PAH gene mutation, the method comprising administering to the subject an effective amount of a replication-defective adeno-associated virus (AAV) comprising:
(a) an AAV capsid comprising an AAV Clade F capsid protein; and
(b) a transfer genome comprising a transcriptional regulatory element operably linked to a PAH coding sequence.

5. The method of claim 4, wherein the disease or disorder is phenylketonuria.

6. The method of claim 4 or 5, wherein the subject is a human subject.

7. The method of any one of the preceding claims, wherein the PAH coding sequence encodes an amino acid sequence set forth in SEQ ID NO: 23.

8. The method of claim 7, wherein the PAH coding sequence comprises the nucleotide sequence set forth in SEQ ID NO: 24.

9. The method of claim 7, wherein the PAH coding sequence is silently altered.

10. The method of claim 9, wherein the PAH coding sequence comprises the nucleotide sequence set forth in SEQ ID NO: 25.

11. The method of any one of the preceding claims, wherein the transcriptional regulatory element is capable of mediating transcription in a hepatocyte, a renal cell, or a cell in the brain, pituitary gland, adrenal gland, pancreas, urinary bladder, gallbladder, colon, small intestine, or breast.

12. The method of claim 11, wherein the transcriptional regulatory element comprises one of more of the elements selected from the group consisting of a CAG promoter, a human EF-1α promoter, a human hepatic control region 1 (HCR1), a human α1-antitrypsin (hAAT) promoter, a hepatic specific regulatory module of hAAT promoter, an SV40 intron, and a minute virus of mouse (MVM) intron.

13. The method of claim 12, wherein the transcriptional regulatory element comprises a nucleotide sequence at least 90% identical to a sequence selected from the group consisting of SEQ ID NOs: 28-30 and 32-41.

14. The method of claim 13, wherein the transcriptional regulatory element comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 28-30 and 32-41.

15. The method of claim 14, wherein the transcriptional regulatory element comprises from 5' to 3' the nucleotide sequences set forth in SEQ ID NOs: 29, 30, and 31.

16. The method of claim 15, wherein the transcriptional regulatory element comprises the nucleotide sequences set forth in SEQ ID NO: 32.

17. The method of any one of the preceding claims, wherein the transfer genome further comprises an intron operably linked to the PAH coding sequence.

18. The method of claim 17, wherein the intron comprises a nucleotide sequence at least 90% identical to the sequence set forth in SEQ ID NO: 31 or 35.

19. The method of claim 18, wherein the intron comprises the nucleotide sequence set forth in SEQ ID NO: 31 or 35.

20. The method of any one of claims 17-19, wherein the transfer genome comprises from 5' to 3': a non-coding exon, the intron, and the PAH coding sequence.

21. The method of any one of the preceding claims, wherein the transfer genome further comprises a polyadenylation sequence 3' to the PAH coding sequence.

22. The method of claim 21, wherein the polyadenylation sequence is an exogenous polyadenylation sequence.

23. The method of claim 22, wherein the exogenous polyadenylation sequence is an SV40 polyadenylation sequence.

24. The method of claim 23, wherein the SV40 polyadenylation sequence comprises a sequence selected from the group consisting of SEQ ID NOs: 42, 43, and 45.

25. The method of any one of the preceding claims, wherein the transfer genome comprises a sequence selected from the group consisting of SEQ ID NOs: 46-50, 61, 64, 67, 74, 76, 78, 80, 82, 84, 86, and 89.

26. The method of any one of the preceding claims, wherein the transfer genome further comprises a 5' inverted terminal repeat (5' ITR) nucleotide sequence 5' of the genome, and a 3' inverted terminal repeat (3' ITR) nucleotide sequence 3' of the genome.

27. The method of claim 26, wherein the 5' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 18, and the 3' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 19.

28. The method of claim 26, wherein the 5' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 20, and the 3' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 21.

29. The method of claim 26, wherein the 5' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 26, and the 3' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 27.

30. The method of any one of the preceding claims, wherein the transfer genome comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 51-55, 62, 65, 68, 75, 77, 79, 81, 83, 85, 87, and 90.

31. The method of any one of the preceding claims, wherein the transfer genome consists of a nucleotide sequence selected from the group consisting of SEQ ID NOs: 51-55, 62, 65, 68, 75, 77, 79, 81, 83, 85, 87, and 90.

32. The method of any one of the preceding claims, wherein the transfer genome consists of the nucleotide sequence set forth in SEQ ID NO: 52.

33. The method of any one of the preceding claims, wherein the AAV Clade F capsid protein comprises an amino acid sequence having at least 95% sequence identity with the amino acid sequence of amino acids 203-736 of SEQ ID NO: 2, 3, 4, 6, 7, 10, 11, 12, 13, 15, 16, or 17.

34. The method of claim 33, wherein: the amino acid in the capsid protein corresponding to amino acid 206 of SEQ ID NO: 2 is C; the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H; the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q; the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A; the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N; the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S; the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 590 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G or Y; the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M; the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K; the amino acid in the capsid protein corresponding

to amino acid 706 of SEQ ID NO: 2 is C; or, the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G.

35. The method of claim 34, wherein:

(a) the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G, and the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G;

(b) the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H, the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M;

(c) the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R;

(d) the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A, and the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; or

(e) the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C.

36. The method of claim 34, wherein the capsid protein comprises the amino acid sequence of amino acids 203-736 of SEQ ID NO: 2, 3, 4, 6, 7, 10, 11, 12, 13, 15, 16, or 17.

37. The method of any one of the preceding claims, wherein the AAV Clade F capsid protein comprises an amino acid sequence having at least 95% sequence identity with the amino acid sequence of amino acids 138-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, or 17.

38. The method of claim 37, wherein: the amino acid in the capsid protein corresponding to amino acid 151 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 160 of SEQ ID NO: 2 is D; the amino acid in the capsid protein corresponding

to amino acid 206 of SEQ ID NO: 2 is C; the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H; the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q; the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A; the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N; the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S; the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 590 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G or Y; the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M; the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K; the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C; or, the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G.

39. The method of claim 38, wherein:

(a) the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G, and the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G;

(b) the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H, the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M;

(c) the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R;

(d) the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A, and the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; or

(e) the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R,

and the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C.

40. The method of claim 38, wherein the capsid protein comprises the amino acid sequence of amino acids 138-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, or 17.

41. The method of any one of the preceding claims, wherein the AAV Clade F capsid protein comprises an amino acid sequence having at least 95% sequence identity with the amino acid sequence of amino acids 1-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17.

42. The method of claim 41, wherein: the amino acid in the capsid protein corresponding to amino acid 2 of SEQ ID NO: 2 is T; the amino acid in the capsid protein corresponding to amino acid 65 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 68 of SEQ ID NO: 2 is V; the amino acid in the capsid protein corresponding to amino acid 77 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 119 of SEQ ID NO: 2 is L; the amino acid in the capsid protein corresponding to amino acid 151 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 160 of SEQ ID NO: 2 is D; the amino acid in the capsid protein corresponding to amino acid 206 of SEQ ID NO: 2 is C; the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H; the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q; the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A; the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N; the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S; the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 590 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G or Y; the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M; the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K; the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C; or, the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G.

43. The method of claim 42, wherein:

(a) the amino acid in the capsid protein corresponding to amino acid 2 of SEQ ID NO: 2 is T, and the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q;

(b) the amino acid in the capsid protein corresponding to amino acid 65 of SEQ ID NO: 2 is I, and the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is Y;

(c) the amino acid in the capsid protein corresponding to amino acid 77 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K;

(d) the amino acid in the capsid protein corresponding to amino acid 119 of SEQ ID NO: 2 is L, and the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S;

(e) the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G, and the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G;

(f) the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H, the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M;

(g) the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R;

(h) the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A, and the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; or

(i) the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C.

44. The method of claim 42, wherein the capsid protein comprises the amino acid sequence of amino acids 1-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17.

45. A replication-defective adeno-associated virus (AAV) comprising:
(a) an AAV capsid comprising an AAV Clade F capsid protein; and
(b) a transfer genome comprising a transcriptional regulatory element operably linked to a PAH coding sequence.

46. The AAV of claim 45, wherein the PAH coding sequence encodes an amino acid sequence set forth in SEQ ID NO: 23.

47. The AAV of claim 46, wherein the PAH coding sequence comprises the nucleotide sequence set forth in SEQ ID NO: 24.

48. The AAV of claim 46, wherein the PAH coding sequence is codon-altered and comprises the sequence set forth in SEQ ID NOs: 25 and 69-73.

49. The AAV of claim 46, wherein the PAH coding sequence is codon-altered and comprises the sequence set forth in SEQ ID NO: 25.

50. The AAV of any one of claims 45-49, wherein the transcriptional regulatory element is capable of mediating transcription in a hepatocyte, a renal cell, or a cell in the brain, pituitary gland, adrenal gland, pancreas, urinary bladder, gallbladder, colon, small intestine, or breast.

51. The AAV of claim 50, wherein the transcriptional regulatory element comprises one of more of the elements selected from the group consisting of a CAG promoter, a human EF-1α promoter, a human HCR1, a human α1-antitrypsin (hAAT) promoter, a hepatic specific regulatory module of hAAT promoter, an SV40 intron, and a minute virus of mouse (MVM) intron.

52. The AAV of claim 51, wherein the transcriptional regulatory element comprises a nucleotide sequence at least 90% identical to a sequence selected from the group consisting of SEQ ID NOs: 28-30 and 32-41.

53. The AAV of claim 52, wherein the transcriptional regulatory element comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 28-30 and 32-41.

54. The AAV of claim 53, wherein the transcriptional regulatory element comprises from 5' to 3' the nucleotide sequences set forth in SEQ ID NOs: 29, 30, and 31.

55. The AAV of claim 54, wherein the transcriptional regulatory element comprises the nucleotide sequences set forth in SEQ ID NO: 32.

56. The AAV of any one of claims 45-55, wherein the transfer genome further comprises an intron operably linked to the PAH coding sequence.

57. The AAV of claim 56, wherein the intron comprises a nucleotide sequence at least 90% identical to the sequence set forth in SEQ ID NO: 31 or 35.

58. The AAV of claim 57, wherein the intron comprises the nucleotide sequence set forth in SEQ ID NO: 31 or 35.

59. The AAV of any one of claims 56-58, wherein the transfer genome comprises from 5' to 3': a non-coding exon, the intron, and the PAH coding sequence.

60. The AAV of any one of claims 45-59, wherein the transfer genome further comprises a polyadenylation sequence 3' to the PAH coding sequence.

61. The AAV of claim 60, wherein the polyadenylation sequence is an exogenous polyadenylation sequence.

62. The AAV of claim 61, wherein the exogenous polyadenylation sequence is an SV40 polyadenylation sequence.

63. The AAV of claim 62, wherein the SV40 polyadenylation sequence comprises a sequence selected from the group consisting of SEQ ID NOs: 42, 43, and 45.

64. The AAV of any one of claims 45-63, wherein the transfer genome comprises a sequence selected from the group consisting of SEQ ID NOs: 46-50, 61, 64, 67, 74, 76, 78, 80, 82, 84, 86, and 89.

65. The AAV of any one of claims 45-64, wherein the transfer genome further comprises a 5' inverted terminal repeat (5' ITR) nucleotide sequence 5' of the genome, and a 3' inverted terminal repeat (3' ITR) nucleotide sequence 3' of the genome.

66. The AAV of claim 65, wherein the 5' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 18, and the 3' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 19.

67. The AAV of claim 65, wherein the 5' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 20, and the 3' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 21.

68. The AAV of claim 65, wherein the 5' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 26, and the 3' ITR nucleotide sequence has at least 95% sequence identity to SEQ ID NO: 27.

69. The AAV of any one of claims 45-68, wherein the transfer genome comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs: 51-55, 62, 65, 68, 75, 77, 79, 81, 83, 85, 87, and 90.

70. The AAV of any one of claims 45-69, wherein the transfer genome consists of a nucleotide sequence selected from the group consisting of SEQ ID NOs: 51-55, 62, 65, 68, 75, 77, 79, 81, 83, 85, 87, and 90.

71. The AAV of any one of claims 45-70, wherein the transfer genome consists of the nucleotide sequence set forth in SEQ ID NO: 52.

72. The AAV of any one of claims 44-71, wherein the AAV Clade F capsid protein comprises an amino acid sequence having at least 95% sequence identity with the amino acid sequence of amino acids 203-736 of SEQ ID NO: 2, 3, 4, 6, 7, 10, 11, 12, 13, 15, 16, or 17.

73. The AAV of claim 72, wherein: the amino acid in the capsid protein corresponding to amino acid 206 of SEQ ID NO: 2 is C; the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H; the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q; the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A; the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N; the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S; the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 590 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G or Y; the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M; the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K; the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C; or, the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G.

74. The AAV of claim 73, wherein:

(a) the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G, and the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G;

(b) the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H, the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M;

(c) the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R;

(d) the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A, and the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; or

(e) the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C.

75. The AAV of claim 73, wherein the capsid protein comprises the amino acid sequence of amino acids 203-736 of SEQ ID NO: 2, 3, 4, 6, 7, 10, 11, 12, 13, 15, 16, or 17.

76. The AAV of any one of claims 45-75, wherein the AAV Clade F capsid protein comprises an amino acid sequence having at least 95% sequence identity with the amino acid sequence of amino acids 138-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, or 17.

77. The AAV of claim 76, wherein: the amino acid in the capsid protein corresponding to amino acid 151 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 160 of SEQ ID NO: 2 is D; the amino acid in the capsid protein corresponding to amino acid 206 of SEQ ID NO: 2 is C; the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H; the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q; the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A; the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N; the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S; the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 590 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G or Y; the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M; the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K; the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C; or, the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G.

78. The AAV of claim 77, wherein:

(a) the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G, and the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G;

(b) the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H, the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M;

(c) the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R;

(d) the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A, and the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; or

(e) the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C.

79. The AAV of claim 77, wherein the capsid protein comprises the amino acid sequence of amino acids 138-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, 16, or 17.

80. The AAV of any one of claims 45-79, wherein the AAV Clade F capsid protein comprises an amino acid sequence having at least 95% sequence identity with the amino acid sequence of amino acids 1-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17.

81. The AAV of claim 80, wherein: the amino acid in the capsid protein corresponding to amino acid 2 of SEQ ID NO: 2 is T; the amino acid in the capsid protein corresponding to amino acid 65 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 68 of SEQ ID NO: 2 is V; the amino acid in the capsid protein corresponding to amino acid 77 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 119 of SEQ ID NO: 2 is L; the amino acid in the capsid protein corresponding to amino acid 151 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to

amino acid 160 of SEQ ID NO: 2 is D; the amino acid in the capsid protein corresponding to amino acid 206 of SEQ ID NO: 2 is C; the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H; the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q; the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A; the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N; the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S; the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I; the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 590 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G or Y; the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M; the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R; the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K; the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C; or, the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G.

82. The AAV of claim 81, wherein:

(a) the amino acid in the capsid protein corresponding to amino acid 2 of SEQ ID NO: 2 is T, and the amino acid in the capsid protein corresponding to amino acid 312 of SEQ ID NO: 2 is Q;

(b) the amino acid in the capsid protein corresponding to amino acid 65 of SEQ ID NO: 2 is I, and the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is Y;

(c) the amino acid in the capsid protein corresponding to amino acid 77 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 690 of SEQ ID NO: 2 is K;

(d) the amino acid in the capsid protein corresponding to amino acid 119 of SEQ ID NO: 2 is L, and the amino acid in the capsid protein corresponding to amino acid 468 of SEQ ID NO: 2 is S;

(e) the amino acid in the capsid protein corresponding to amino acid 626 of SEQ ID NO: 2 is G, and the amino acid in the capsid protein corresponding to amino acid 718 of SEQ ID NO: 2 is G;

(f) the amino acid in the capsid protein corresponding to amino acid 296 of SEQ ID NO: 2 is H, the amino acid in the capsid protein corresponding to amino acid 464 of SEQ ID NO: 2 is N, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 681 of SEQ ID NO: 2 is M;

(g) the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 687 of SEQ ID NO: 2 is R;

(h) the amino acid in the capsid protein corresponding to amino acid 346 of SEQ ID NO: 2 is A, and the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R; or

(i) the amino acid in the capsid protein corresponding to amino acid 501 of SEQ ID NO: 2 is I, the amino acid in the capsid protein corresponding to amino acid 505 of SEQ ID NO: 2 is R, and the amino acid in the capsid protein corresponding to amino acid 706 of SEQ ID NO: 2 is C.

83. The AAV of claim 81, wherein the capsid protein comprises the amino acid sequence of amino acids 1-736 of SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, or 17.

84. A pharmaceutical composition comprising an AAV of any one of claims 45-83.

85. A packaging system for recombinant preparation of a replication-defective AAV, wherein the packaging system comprises:

(a) a Rep nucleotide sequence encoding one or more AAV Rep proteins;

(b) a Cap nucleotide sequence encoding one or more AAV Clade F capsid proteins as set forth in any one of claims 72-83; and

(c) a transfer genome as set forth in any one of claims 45-71, wherein the packaging system is operative in a cell for enclosing the transfer genome in the capsid to form the AAV.

86. The packaging system of claim 85, wherein the packaging system comprises a first vector comprising the Rep nucleotide sequence and the Cap nucleotide sequence, and a second vector comprising the transfer genome.

87. The packaging system of claims 85 or 86, wherein the Rep nucleotide sequence encodes an AAV2 Rep protein.

88. The packaging system of claim 87, wherein the AAV2 Rep protein is 78/68 or Rep 68/52.

89. The packaging system of claim 87 or 88, wherein the AAV2 Rep protein comprises an amino acid sequence having a minimum percent sequence identity to the AAV2 Rep amino acid sequence of SEQ ID NO: 22, wherein the minimum percent sequence identity is at least 70% across the length of the amino acid sequence encoding the AAV2 Rep protein.

90. The packaging system of any one of claims 85-89, further comprising a third vector, wherein the third vector is a helper virus vector.

91. The packaging system of claim 90, wherein the helper virus vector is an independent third vector.

92. The packaging system of claim 90, wherein the helper virus vector is integral with the first vector.

93. The packaging system of claim 90, wherein the helper virus vector is integral with the second vector.

94. The packaging system of any one of claims 90-93, wherein the third vector comprises genes encoding helper virus proteins.

95. The packaging system of any one of claims 90-94, wherein the helper virus is selected from the group consisting of adenovirus, herpes virus, vaccinia virus, and cytomegalovirus (CMV).

96. The packaging system of claim 95, wherein the helper virus is adenovirus.

97. The packaging system of claim 96, wherein the adenovirus genome comprises one or more adenovirus RNA genes selected from the group consisting of E1, E2, E4 and VA.

98. The packaging system of claim 95, wherein the helper virus is herpes simplex virus (HSV).

99. The packaging system of claim 98, wherein the HSV genome comprises one or more of HSV genes selected from the group consisting of UL5/8/52, ICPO, ICP4, ICP22 and UL30/UL42.

100. The packaging system of any one of claims 90-99, wherein the first vector and the third vector are contained within a first transfecting plasmid.

101. The packaging system of any one of claims 90-99, wherein the nucleotides of the second vector and the third vector are contained within a second transfecting plasmid.

102. The packaging system of any one of claims 90-99, wherein the nucleotides of the first vector and the third vector are cloned into a recombinant helper virus.

103. The packaging system of any one of claims 90-99, wherein the nucleotides of the second vector and the third vector are cloned into a recombinant helper virus.

104. A method for recombinant preparation of a replication-defective AAV, the method comprising introducing the packaging system of any one of claims 85-103 into a cell under conditions operative for enclosing the transfer genome in the capsid to form the AAV.

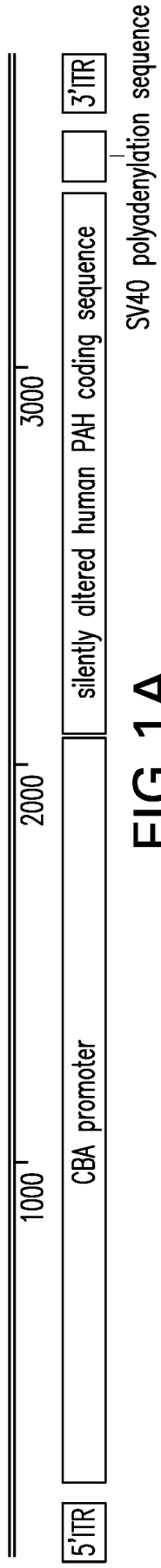

FIG. 1A

FIG. 1B

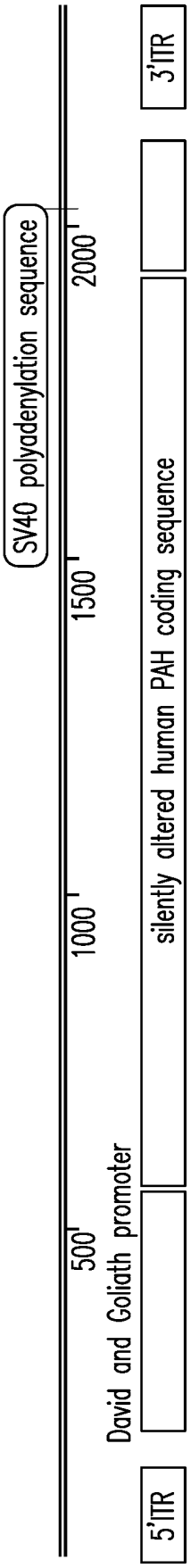

FIG. 1C

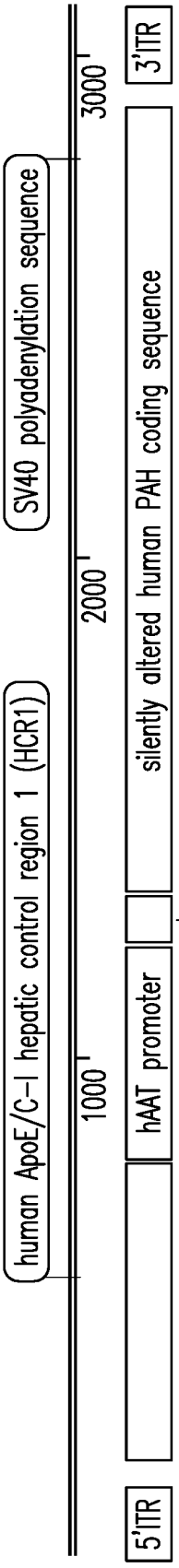
human ApoE/C-I hepatic control region 1 (HCR1)
1000
2000
3000
SV40 polyadenylation sequence
5'ITR
hAAT promoter
modified MVM intron
silently altered human PAH coding sequence
3'ITR

FIG.1D

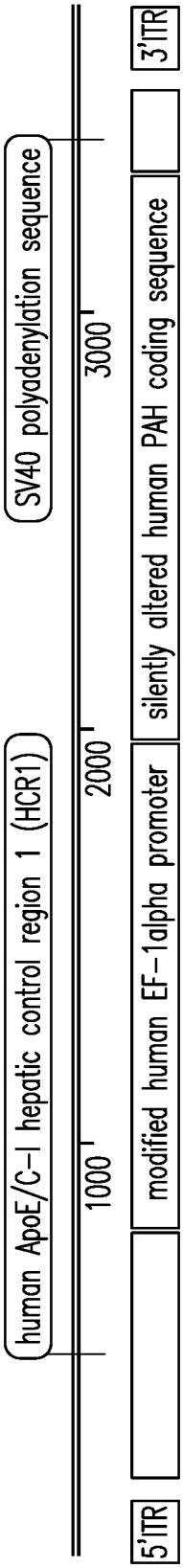
human ApoE/C-I hepatic control region 1 (HCR1)
1000
2000
3000
SV40 polyadenylation sequence
5'ITR
modified human EF-1alpha promoter
silently altered human PAH coding sequence
3'ITR

FIG.1E

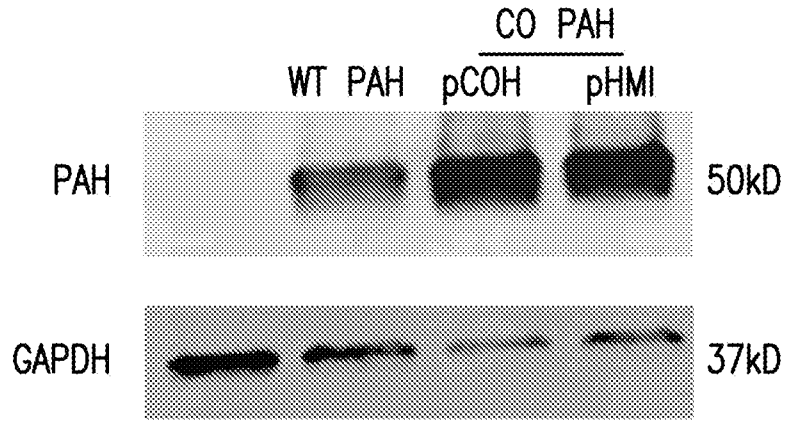

FIG.2

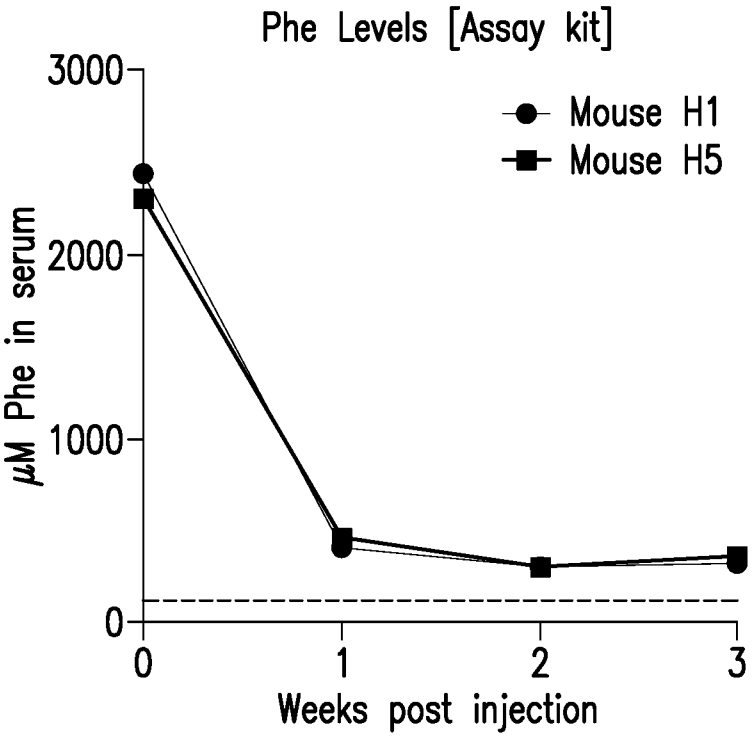
Phe Levels [Assay kit]
Mouse H1
Mouse H5
3000
2000
1000
0
μM Phe in serum
0
1
2
3
Weeks post injection

FIG.3A

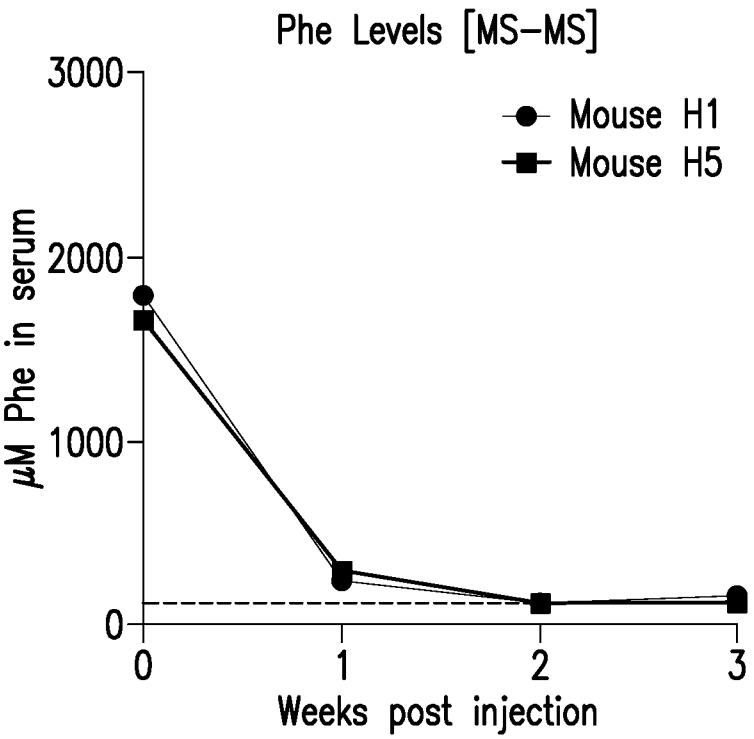
Phe Levels [MS-MS]
Mouse H1
Mouse H5
3000
2000
1000
0
μM Phe in serum
0
1
2
3
Weeks post injection

FIG.3B

Average of 2 mice

	vg/1e6 cells
Brain	52237
Heart	203161
Kidney	493460
Liver	89832875
Muscle	13157

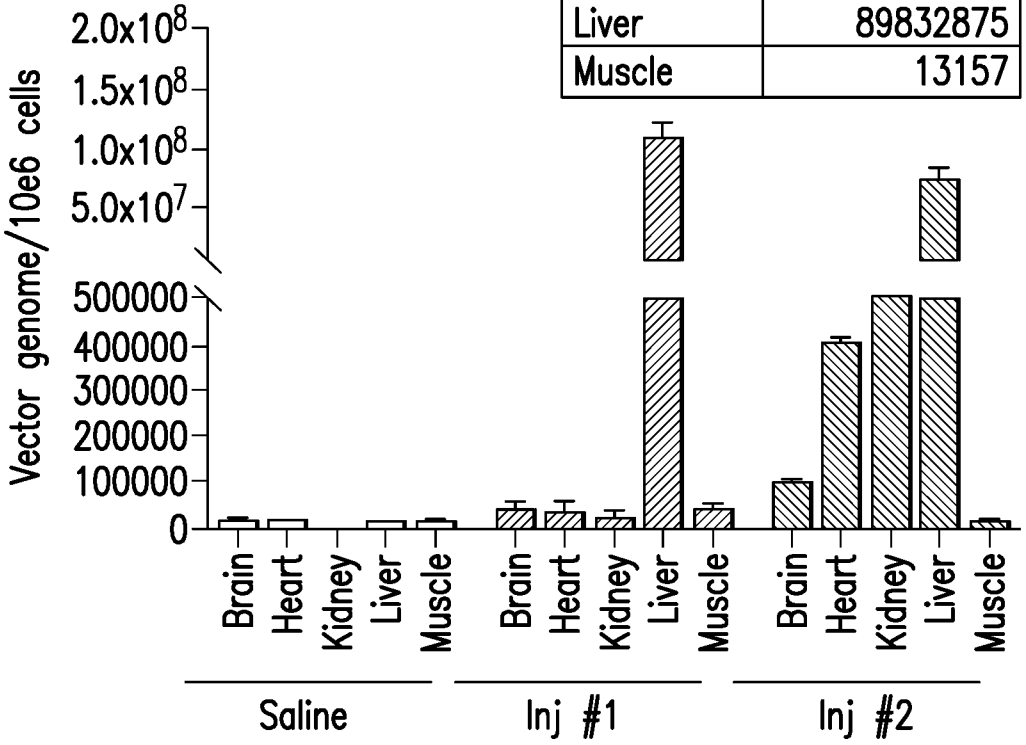

FIG.4

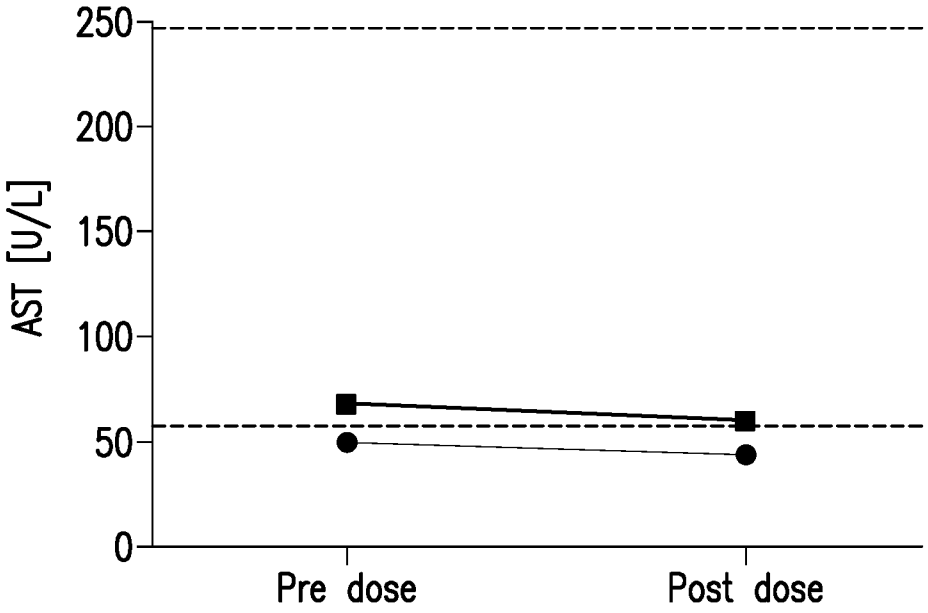
250
200
150
100
50
0
AST [U/L]
Pre dose
Post dose

FIG.5A

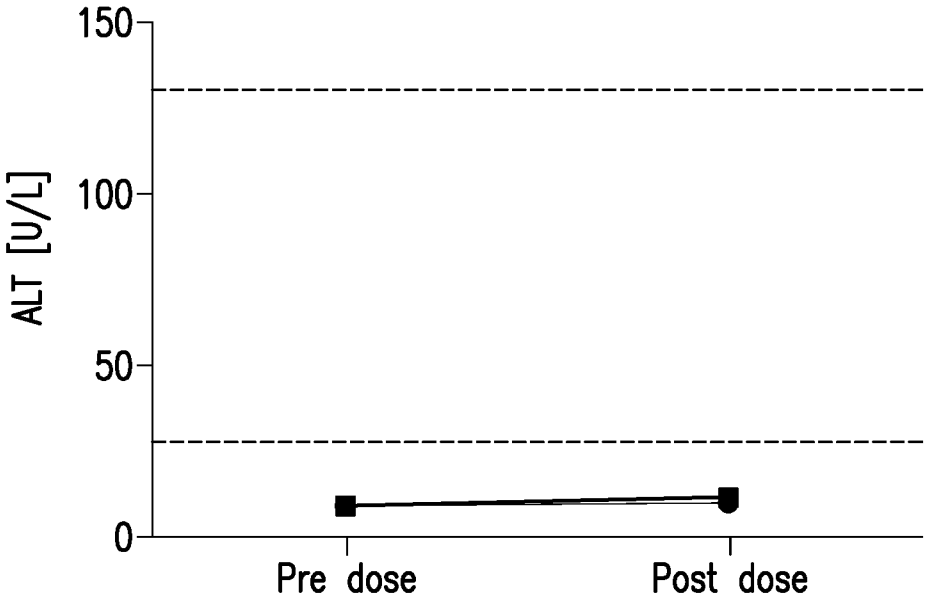
150
100
50
0
ALT [U/L]
Pre dose
Post dose

FIG.5B

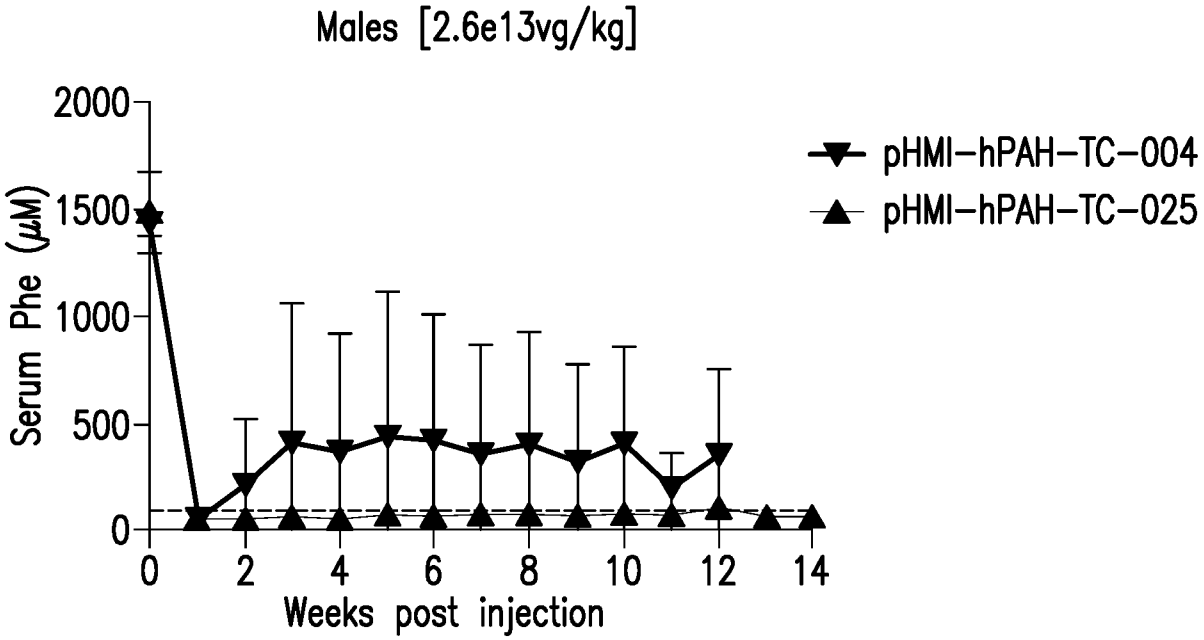
Males [2.6e13vg/kg]
Serum Phe (μM)
2000
1500
1000
500
0
0
2
4
6
8
10
12
14
Weeks post injection
pHMI-hPAH-TC-004
pHMI-hPAH-TC-025

FIG.6A

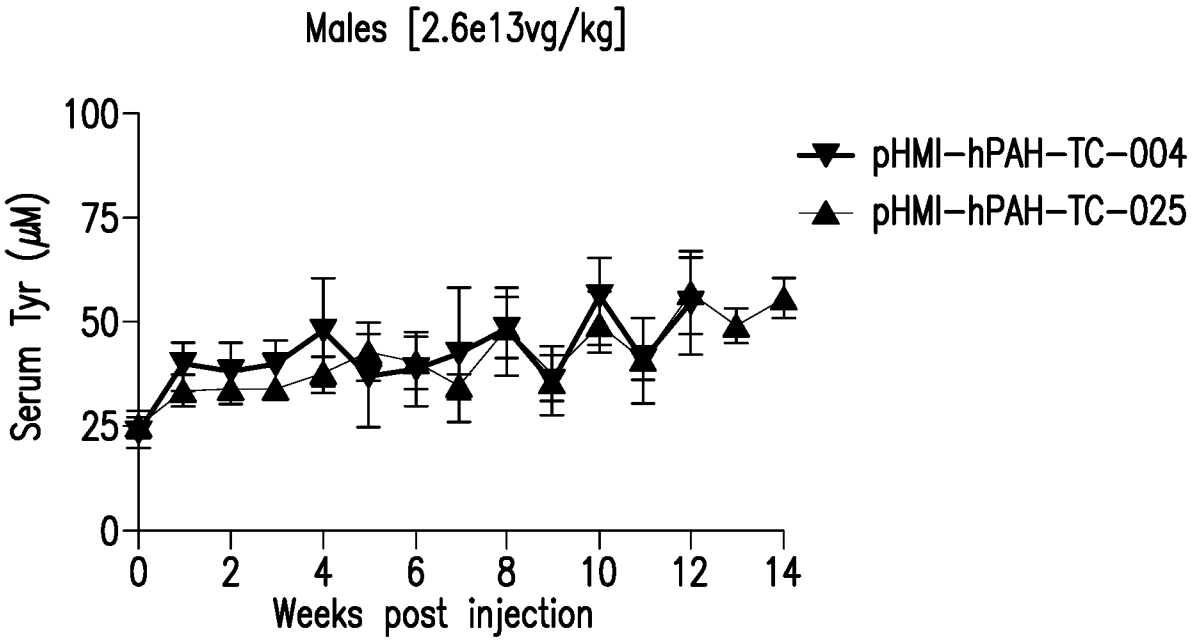
Males [2.6e13vg/kg]
Serum Tyr (μM)
100
75
50
25
0
0
2
4
6
8
10
12
14
Weeks post injection
pHMI-hPAH-TC-004
pHMI-hPAH-TC-025

FIG.6B

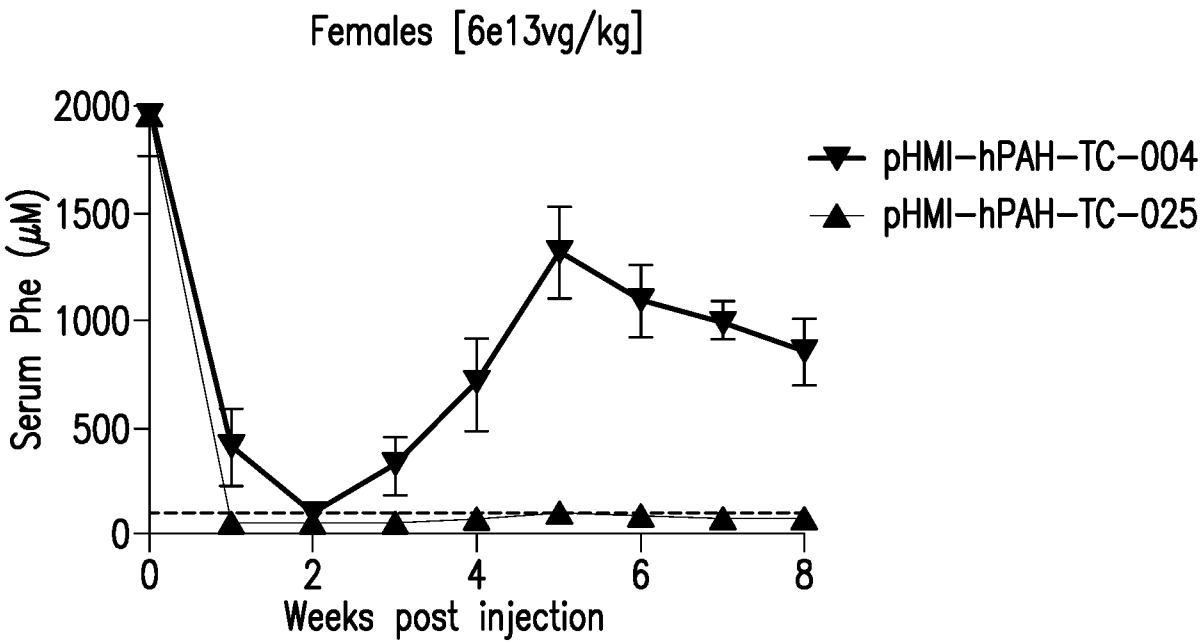
Females [6e13vg/kg]
Serum Phe (μM)
2000
1500
1000
500
0
0
2
4
6
8
Weeks post injection
pHMI-hPAH-TC-004
pHMI-hPAH-TC-025

FIG.6C

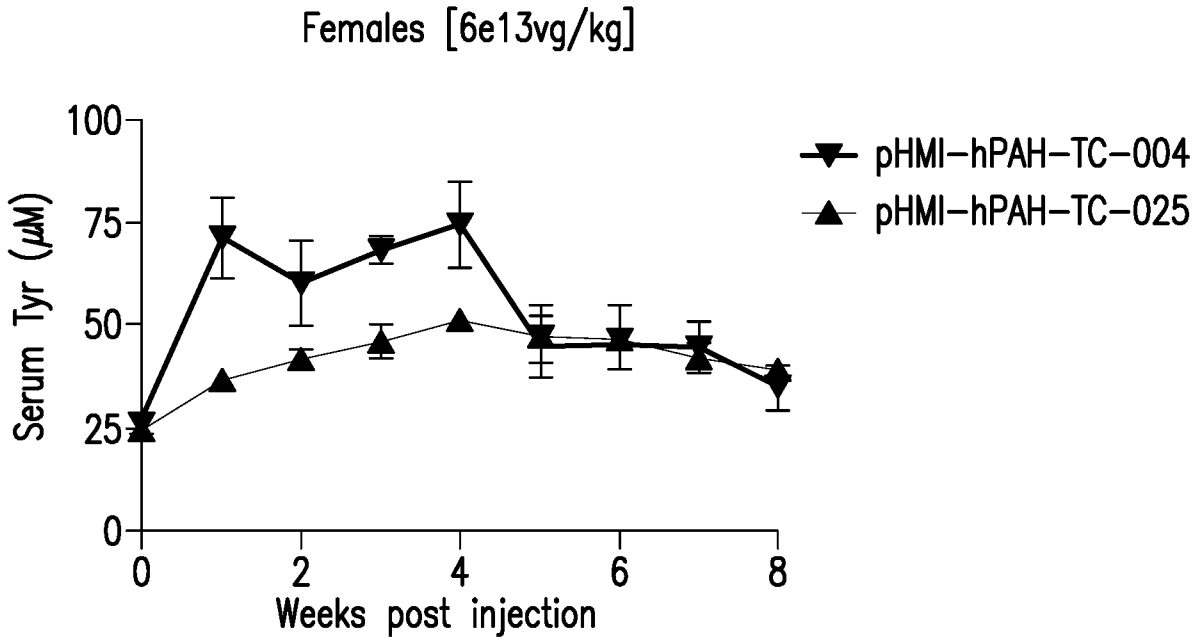
Females [6e13vg/kg]
Serum Tyr (μM)
100
75
50
25
0
0
2
4
6
8
Weeks post injection
pHMI-hPAH-TC-004
pHMI-hPAH-TC-025

FIG.6D

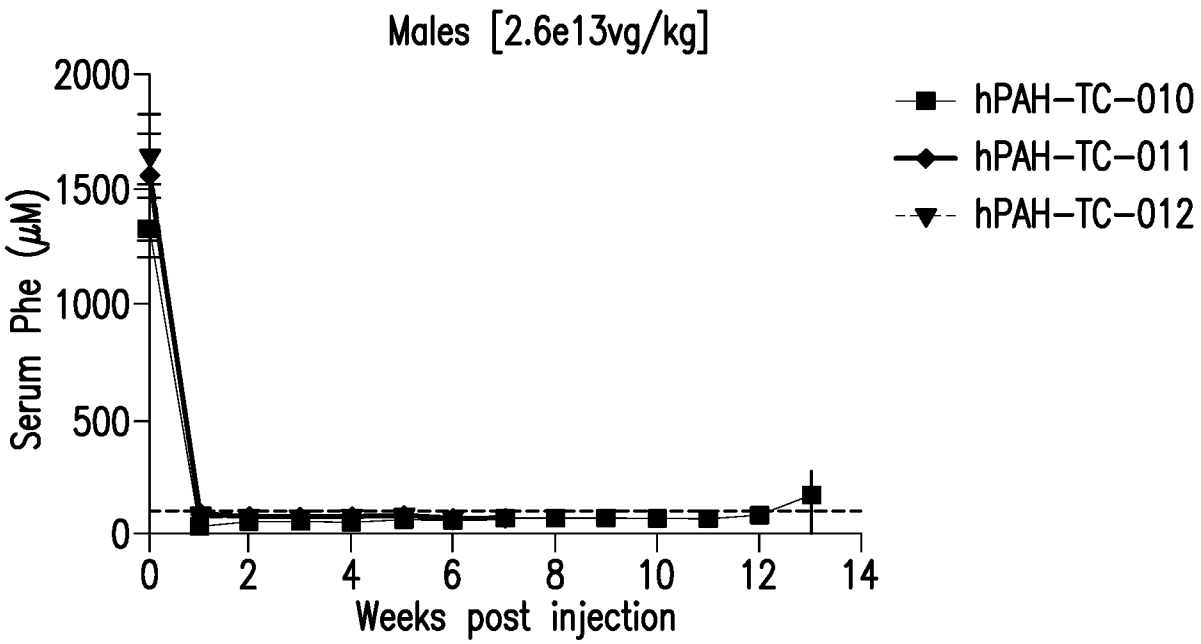
Males [2.6e13vg/kg]
Serum Phe (μM)
2000
1500
1000
500
0
0
2
4
6
8
10
12
14
Weeks post injection
hPAH-TC-010
hPAH-TC-011
hPAH-TC-012

FIG.6E

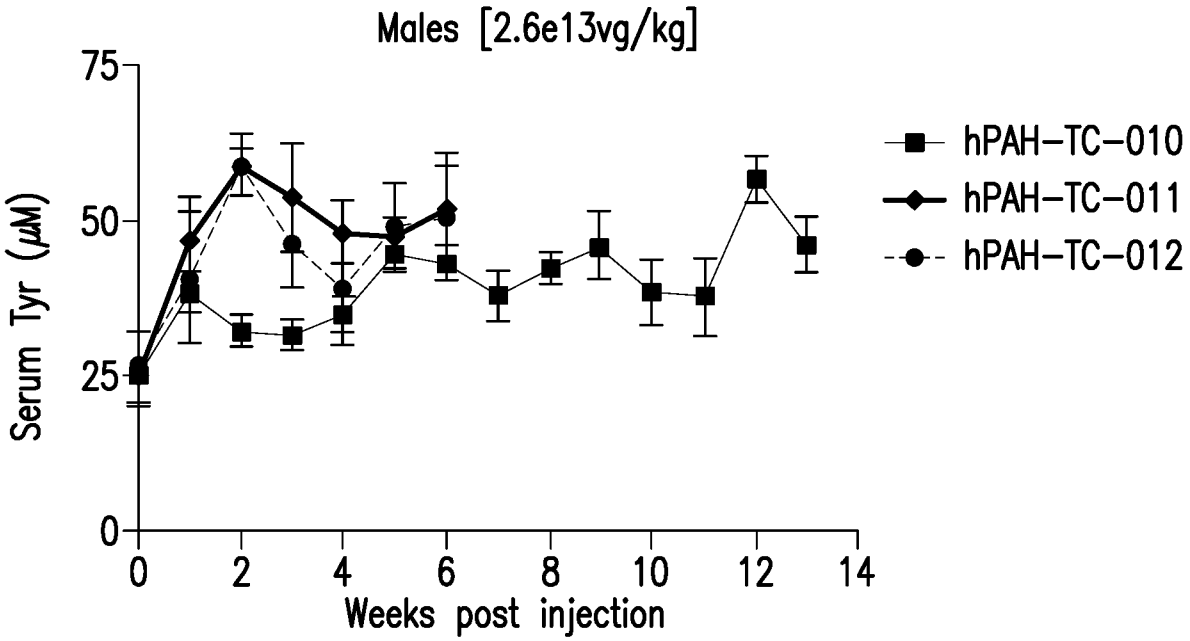
Males [2.6e13vg/kg]
Serum Tyr (μM)
75
50
25
0
0
2
4
6
8
10
12
14
Weeks post injection
hPAH-TC-010
hPAH-TC-011
hPAH-TC-012

FIG.6F

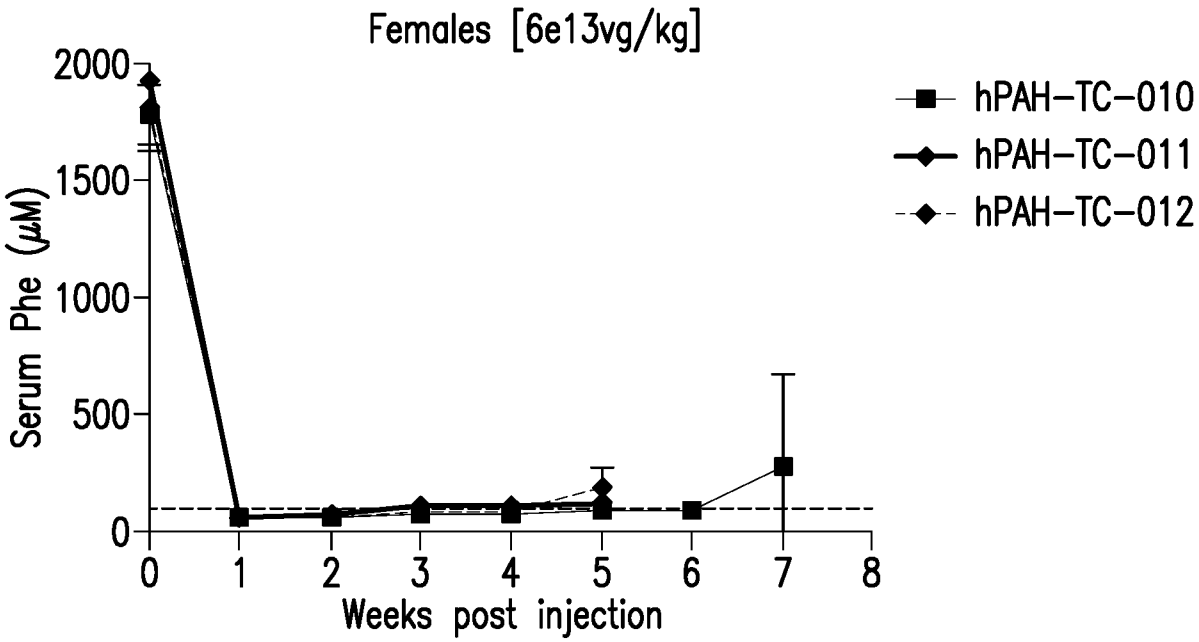
Females [6e13vg/kg]
Serum Phe (μM)
2000
1500
1000
500
0
0 1 2 3 4 5 6 7 8
Weeks post injection
hPAH-TC-010
hPAH-TC-011
hPAH-TC-012

FIG.6G

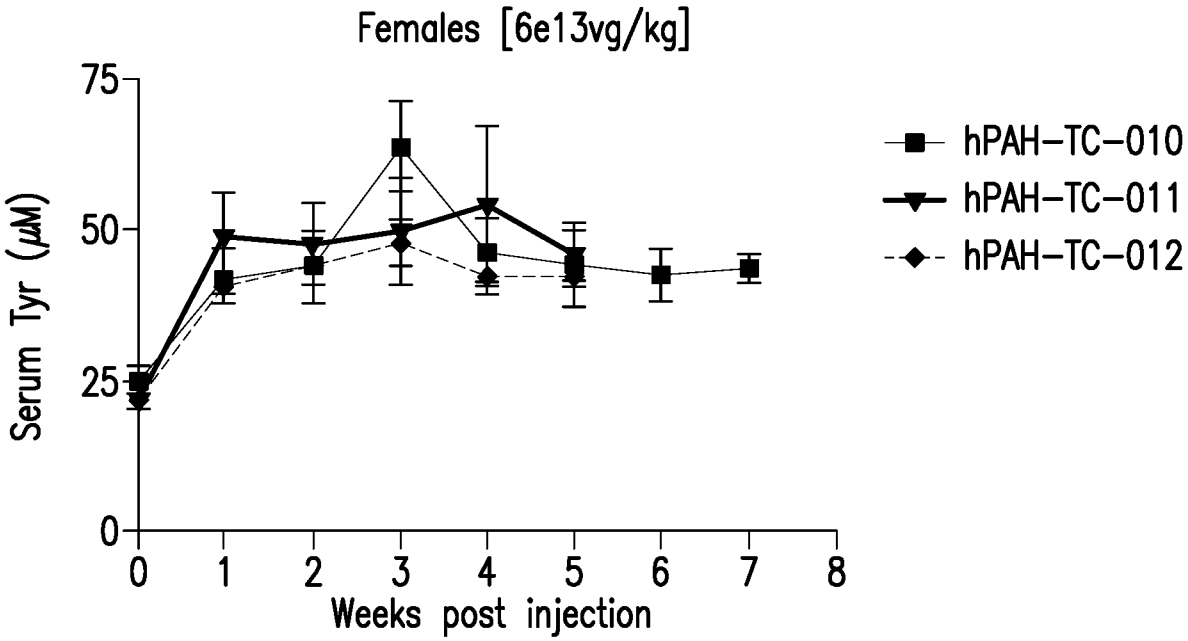
Females [6e13vg/kg]
Serum Tyr (μM)
75
50
25
0
0 1 2 3 4 5 6 7 8
Weeks post injection
hPAH-TC-010
hPAH-TC-011
hPAH-TC-012

FIG.6H

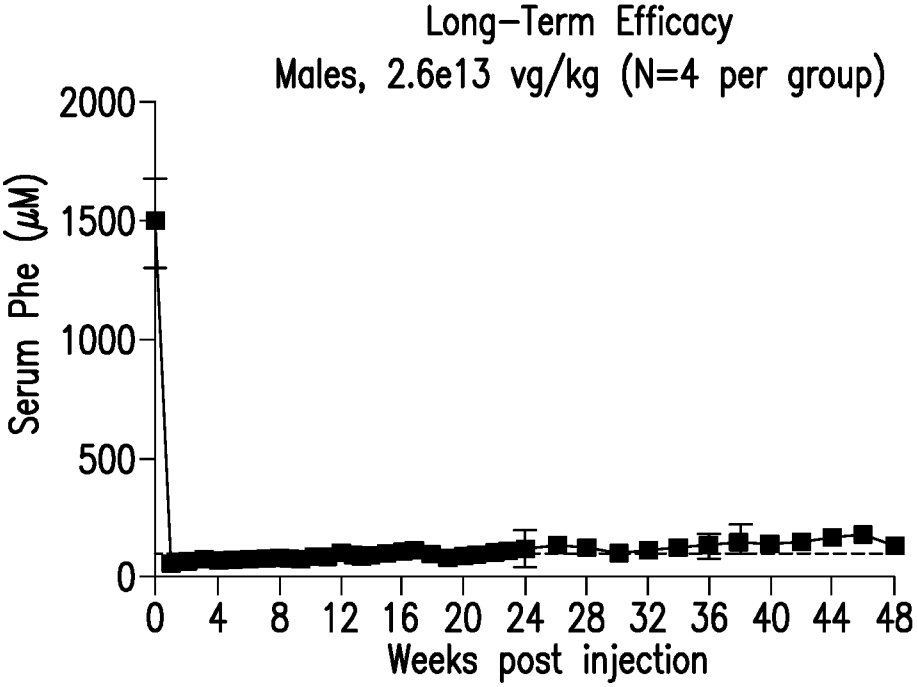
Long-Term Efficacy
Males, 2.6e13 vg/kg (N=4 per group)
Serum Phe (μM)
2000
1500
1000
500
0
0 4 8 12 16 20 24 28 32 36 40 44 48
Weeks post injection

FIG.6I

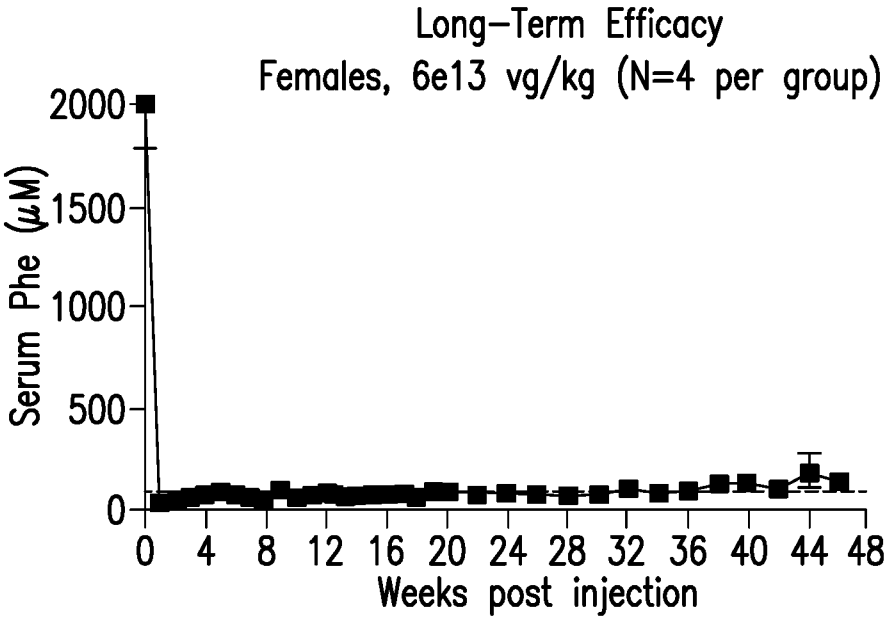
Long-Term Efficacy
Females, 6e13 vg/kg (N=4 per group)
Serum Phe (μM)
2000
1500
1000
500
0
0 4 8 12 16 20 24 28 32 36 40 44 48
Weeks post injection

FIG.6J

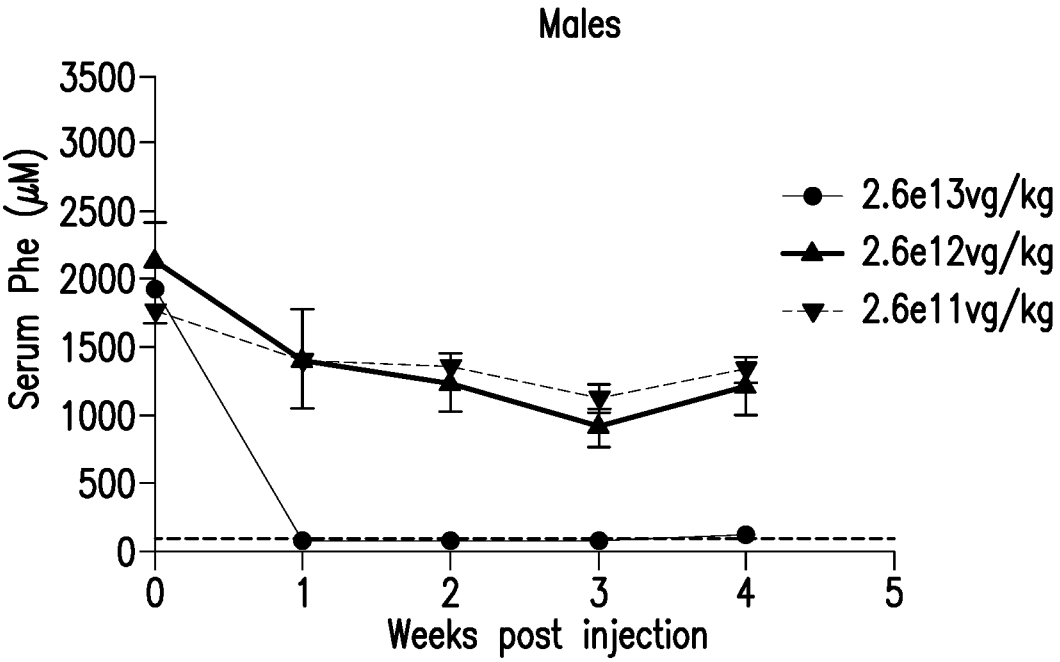
Males
Serum Phe (μM)
3500
3000
2500
2000
1500
1000
500
0
0
1
2
3
4
5
Weeks post injection
2.6e13vg/kg
2.6e12vg/kg
2.6e11vg/kg

FIG.7A

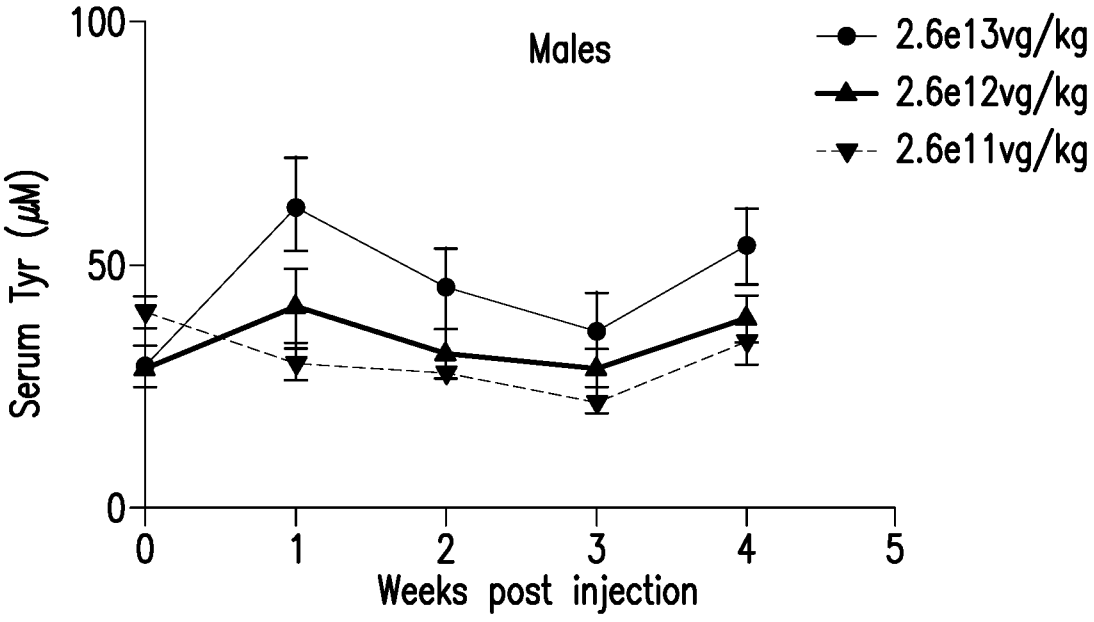
Males
Serum Tyr (μM)
100
50
0
0
1
2
3
4
5
Weeks post injection
2.6e13vg/kg
2.6e12vg/kg
2.6e11vg/kg

FIG.7B

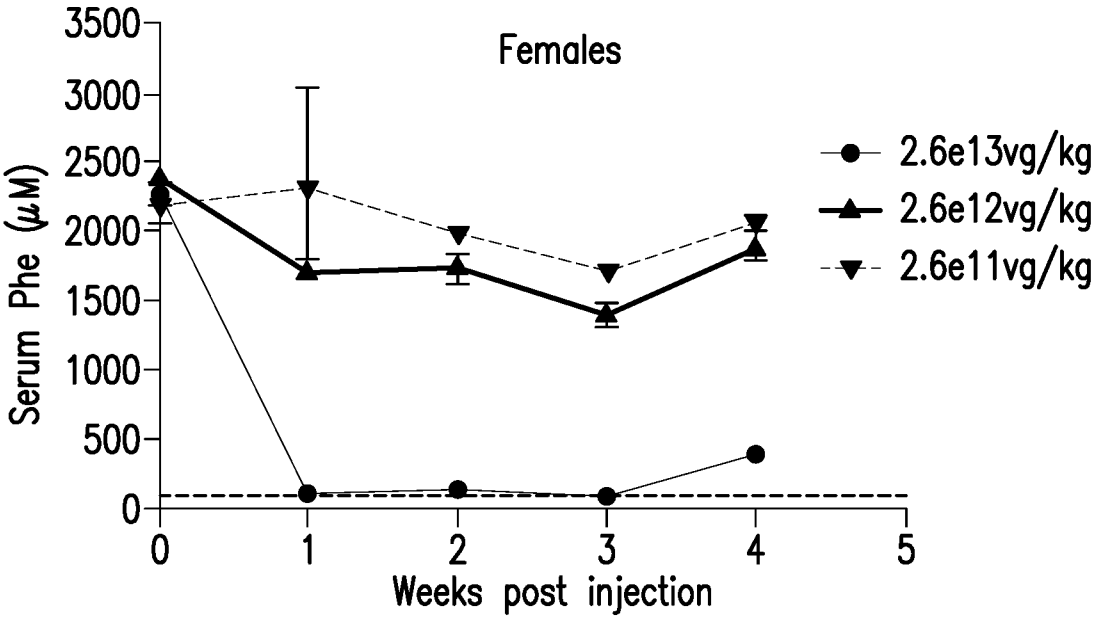
Females
Serum Phe (μM)
3500
3000
2500
2000
1500
1000
500
0
0
1
2
3
4
5
Weeks post injection
2.6e13vg/kg
2.6e12vg/kg
2.6e11vg/kg

FIG.7C

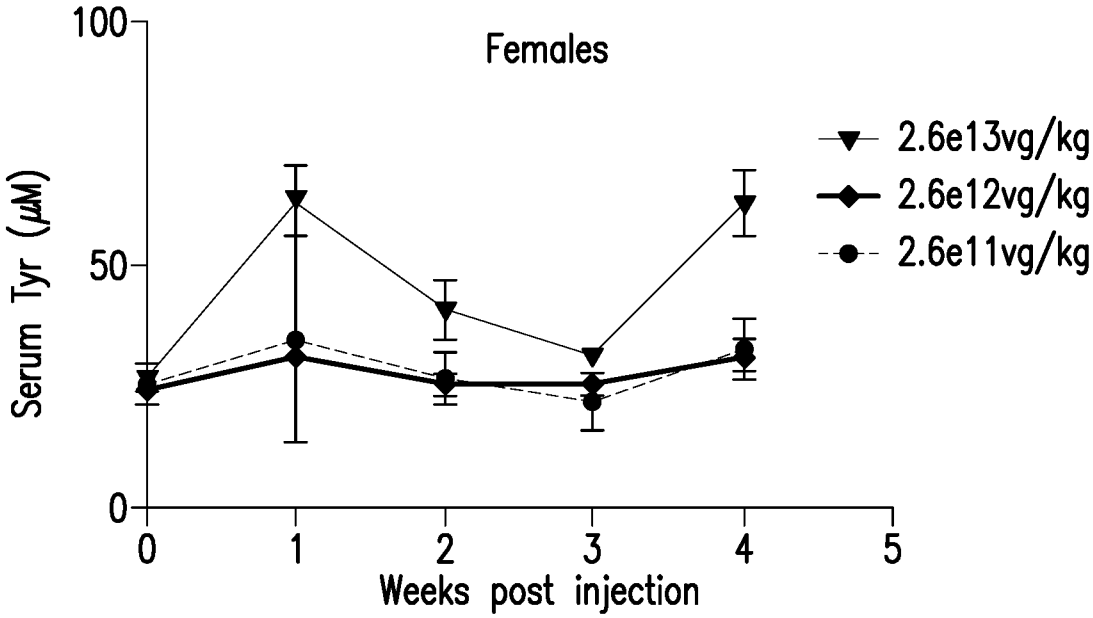
Females
Serum Tyr (μM)
100
50
0
0
1
2
3
4
5
Weeks post injection
2.6e13vg/kg
2.6e12vg/kg
2.6e11vg/kg

FIG.7D

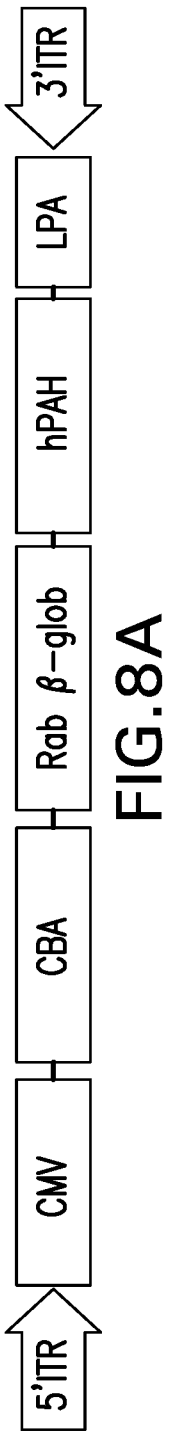

FIG.8A

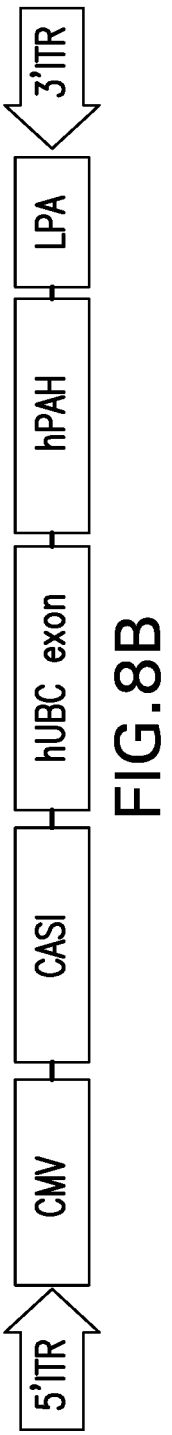

FIG.8B

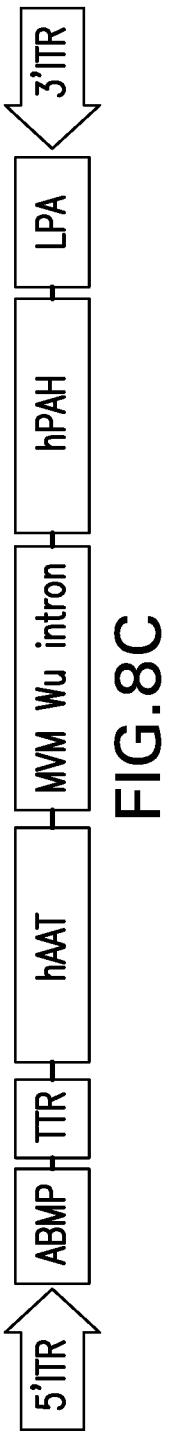

FIG.8C

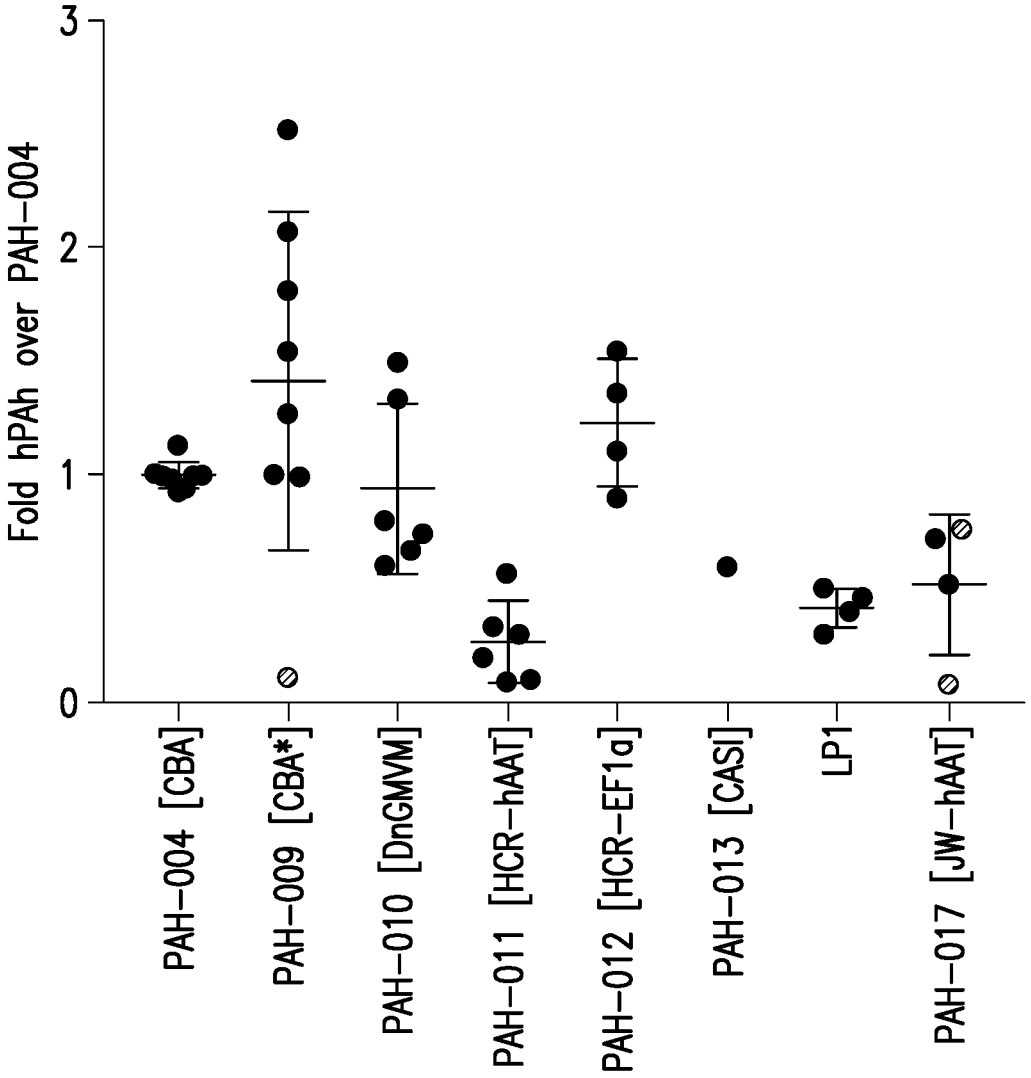
Fold hPAh over PAH-004
3
2
1
0
PAH-004 [CBA]
PAH-009 [CBA*]
PAH-010 [DnGMVM]
PAH-011 [HCR-hAAT]
PAH-012 [HCR-EF1a]
PAH-013 [CASI]
LP1
PAH-017 [JW-hAAT]

FIG.9A

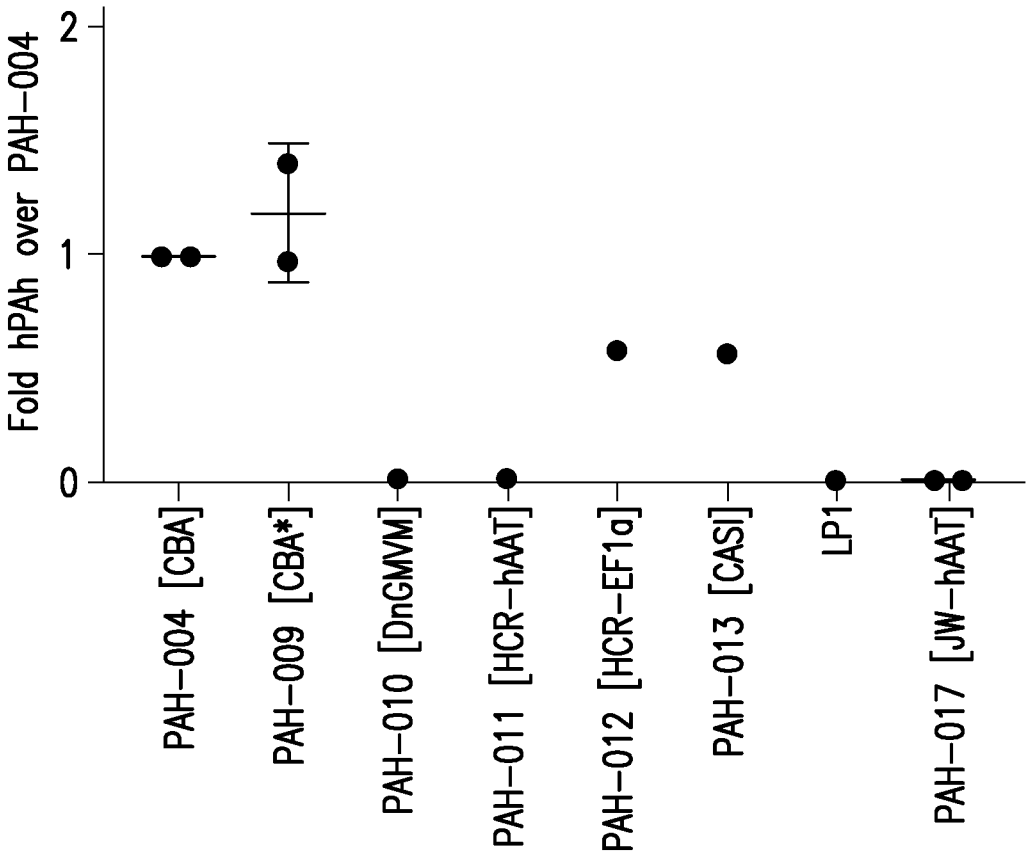
Fold hPAh over PAH-004
2
1
0
PAH-004 [CBA]
PAH-009 [CBA*]
PAH-010 [DnGMVM]
PAH-011 [HCR-hAAT]
PAH-012 [HCR-EF1a]
PAH-013 [CASI]
LP1
PAH-017 [JW-hAAT]

FIG.9B

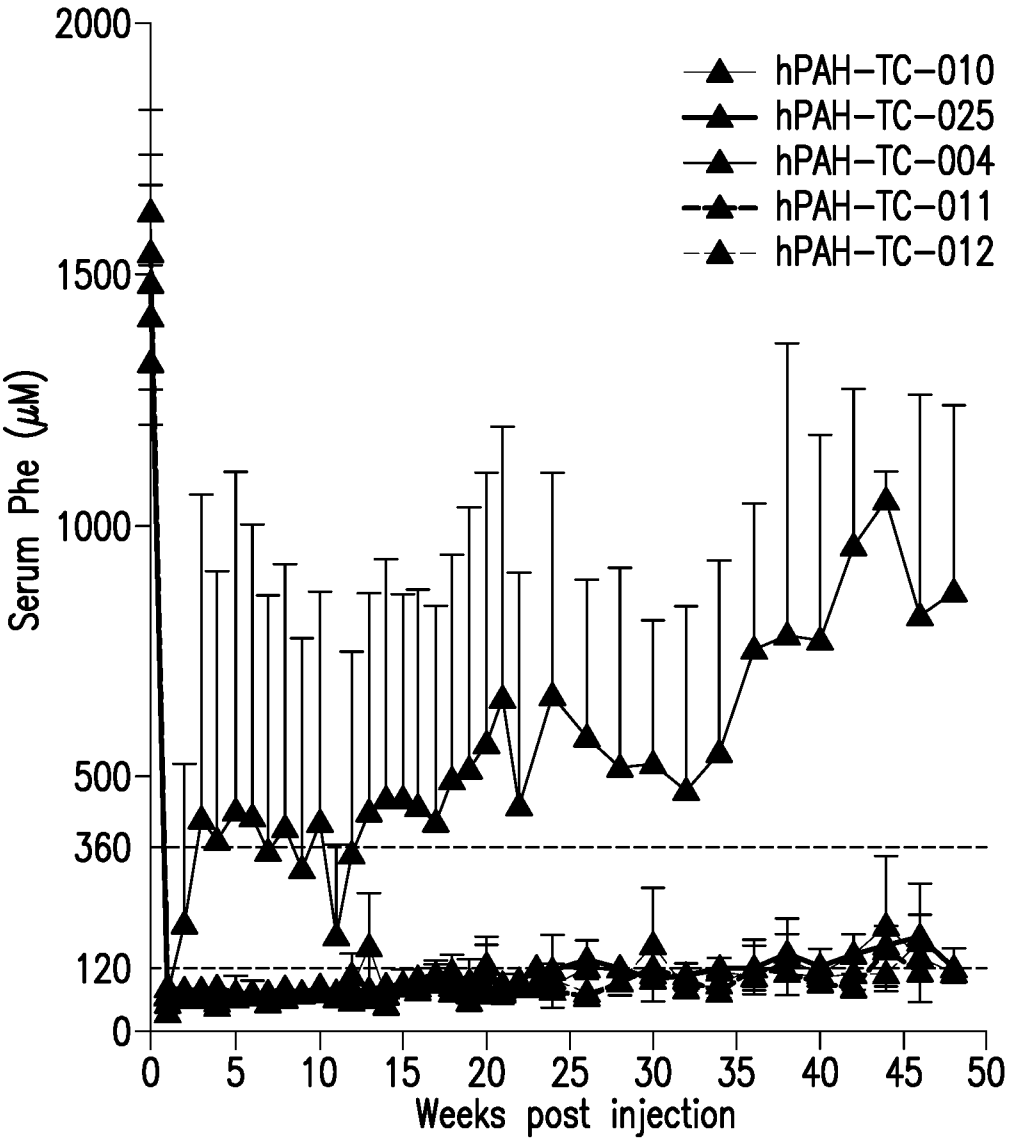
2000
1500
1000
500
360
120
0
Serum Phe (μM)
0
5
10
15
20
25
30
35
40
45
50
Weeks post injection
hPAH-TC-010
hPAH-TC-025
hPAH-TC-004
hPAH-TC-011
hPAH-TC-012

FIG.10A

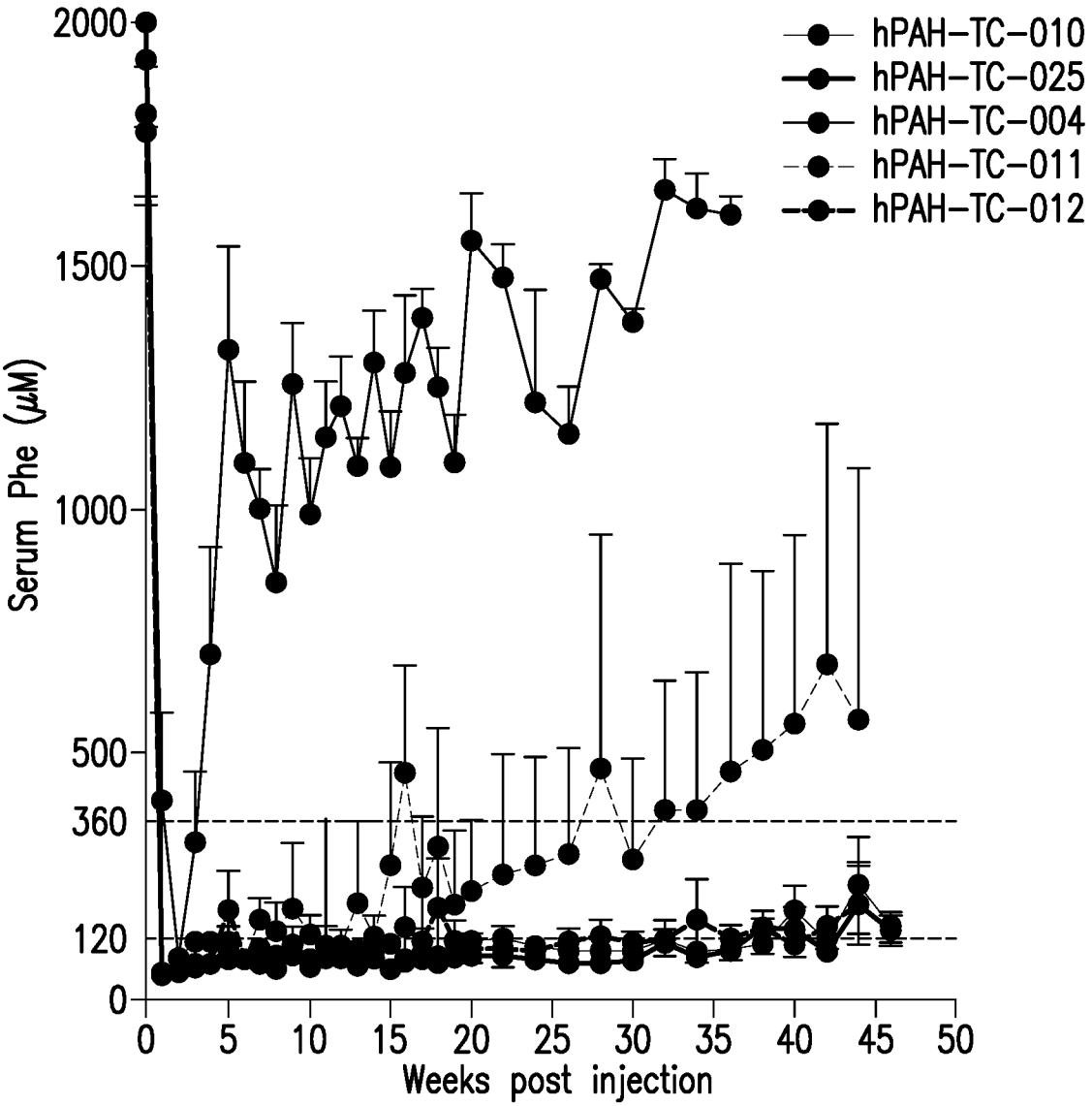
hPAH-TC-010
hPAH-TC-025
hPAH-TC-004
hPAH-TC-011
hPAH-TC-012
Serum Phe (μM)
2000
1500
1000
500
360
120
0
0
5
10
15
20
25
30
35
40
45
50
Weeks post injection

FIG.10B

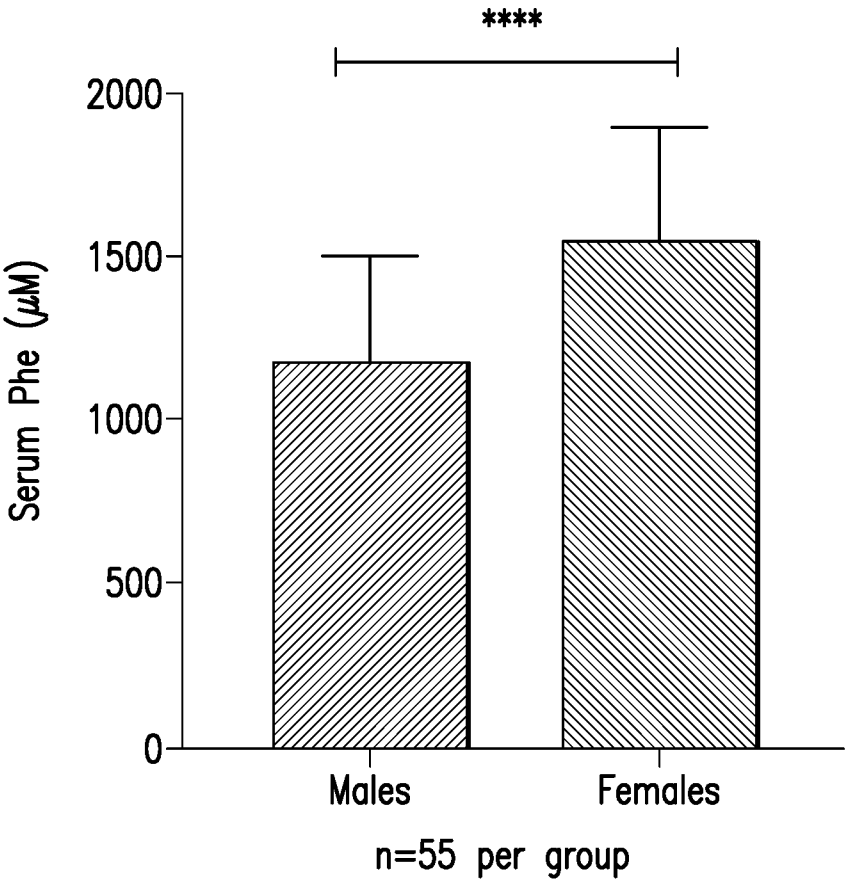

FIG.10C

FIG. 11A

FIG. 11B

FIG. 11C

FIG.11D

FIG.11E

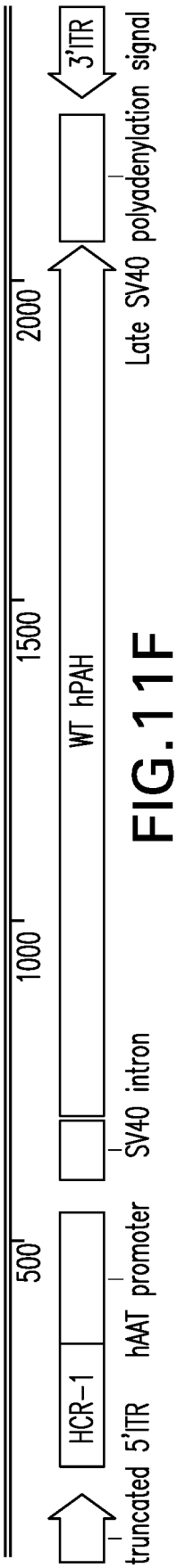

FIG.11F

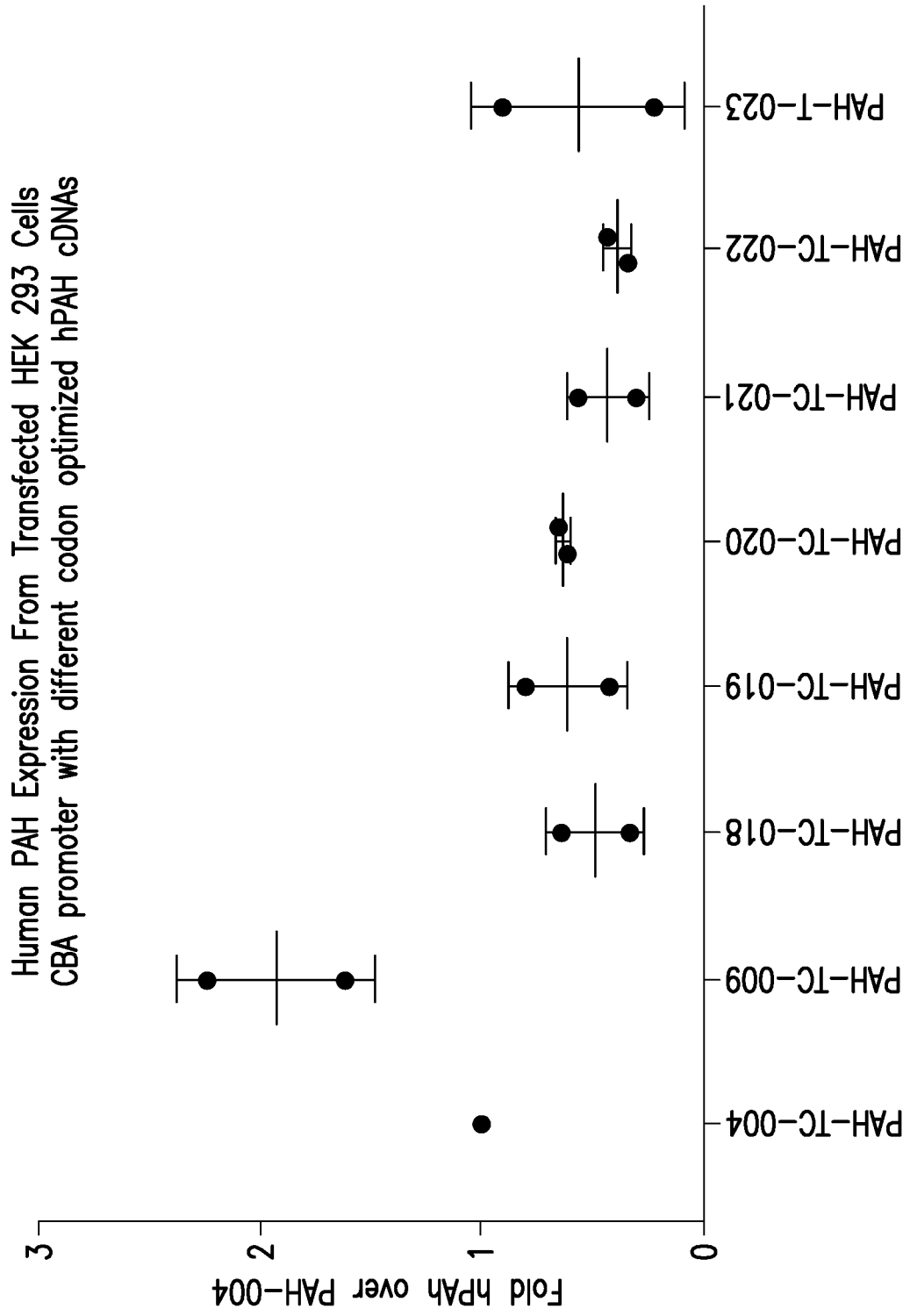
Human PAH Expression From Transfected HEK 293 Cells
CBA promoter with different codon optimized hPAH cDNAs
Fold hPAh over PAH-004
0
1
2
3
PAH-TC-004
PAH-TC-009
PAH-TC-018
PAH-TC-019
PAH-TC-020
PAH-TC-021
PAH-TC-022
PAH-T-023

FIG.12

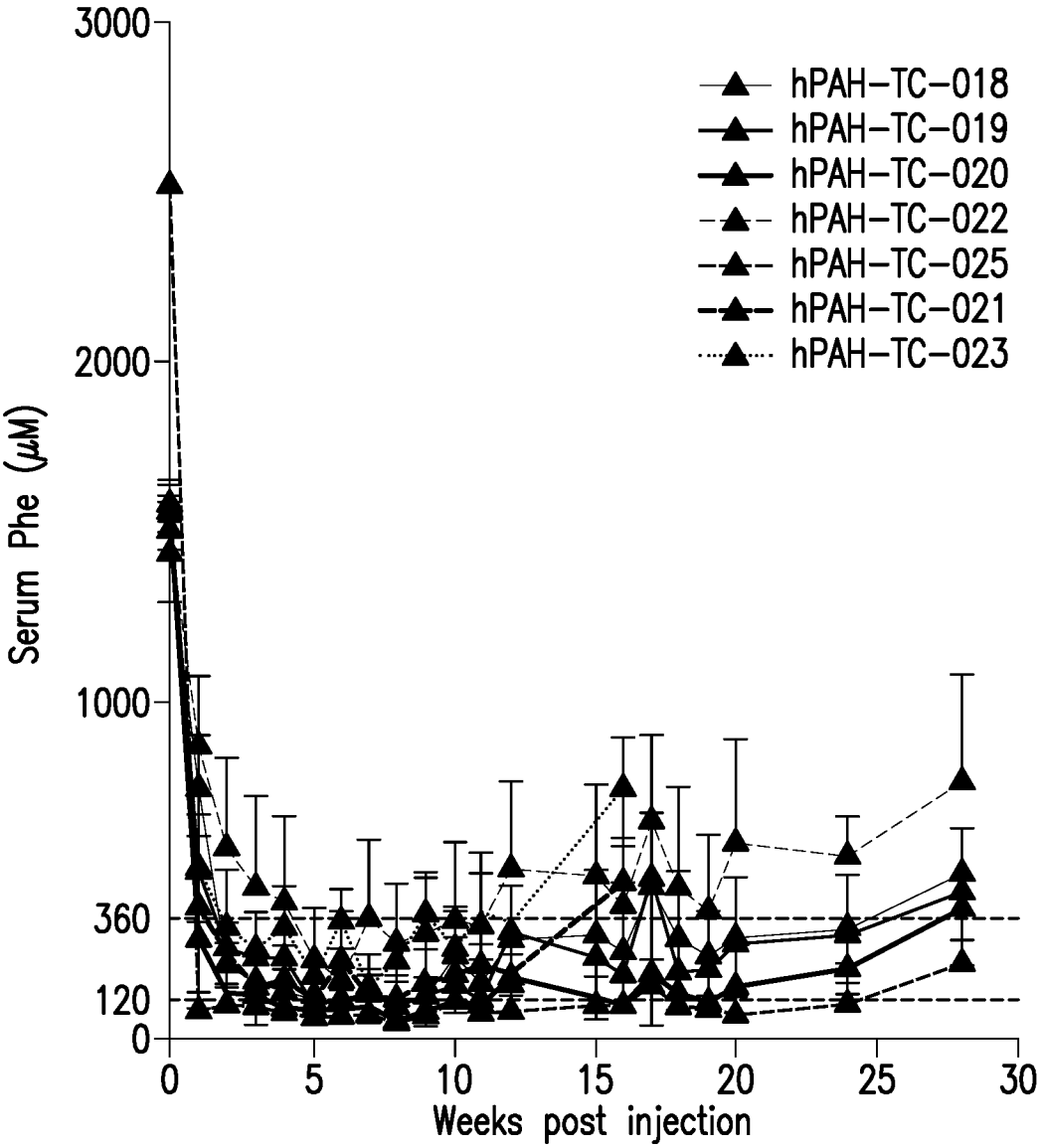
3000
2000
1000
360
120
0
Serum Phe (μM)
0
5
10
15
20
25
30
Weeks post injection
hPAH-TC-018
hPAH-TC-019
hPAH-TC-020
hPAH-TC-022
hPAH-TC-025
hPAH-TC-021
hPAH-TC-023

FIG.13

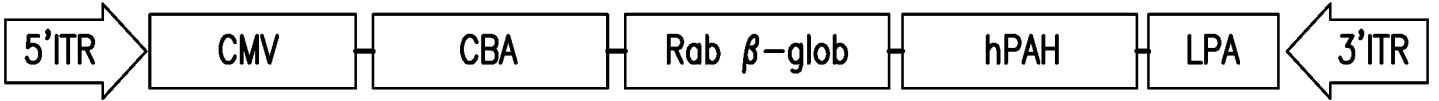

FIG.8A