



(51) International Patent Classification:

A61K 38/16 (2006.01) A61P 43/00 (2006.01)

A61K 35/76 (2015.01) C12N 15/63 (2006.01)

A61K 48/00 (2006.01) C12N 15/86 (2006.01)

(21) International Application Number:

PCT/US2022/026101

(22) International Filing Date:

25 April 2022 (25.04.2022)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/179,642 26 April 2021 (26.04.2021) US

(71) Applicant: UNIVERSITY OF MASSACHUSETTS

[US/US]; One Beacon Street, 31st Floor, Boston, MA 02108 (US).

(72) Inventors: MUELLER, Christian; 29 Winslow Street, Concord, MA 01742 (US). BOREL, Florie; 29 Winslow St., Concord, MA 01742 (US). ZIEGER, Marina; 51 Chestnut Street, Hopkinton, MA 01748 (US).

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

(81) Designated States (unless otherwise indicated, for every kind of national protection available):

AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU,

RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available):

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

Published:

- with international search report (Art. 21(3))
- 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: AAV ENCODING HERMANSKY-PUDLAK SYNDROME 1 (HPS1) PROTEIN AND USES THEREOF

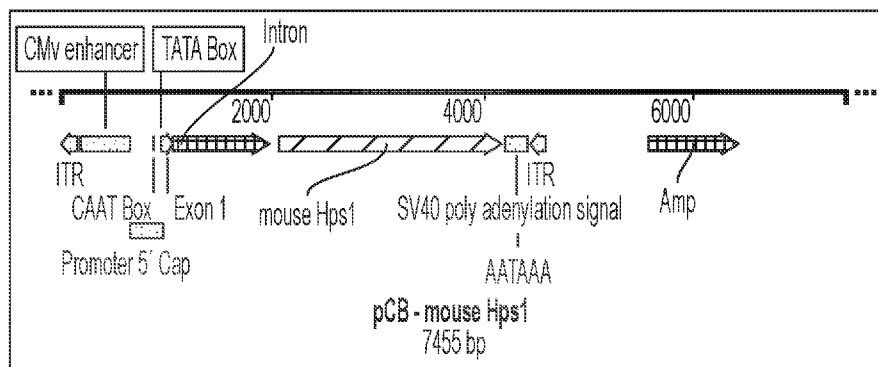


FIG. 1A

(57) Abstract: Aspects of the disclosure relates to composition and methods for treating Hermansky-Pudlak syndrome-associated pulmonary fibrosis in a subject. The disclosure is based, in part, on isolated nucleic acids and rAAVs encoding a Hermansky-Pudlak syndrome 1 (HPS1) protein. In some embodiments, the rAAV delivers the HPS1 protein to the lungs (e.g., alveolar type II cells) of the subject.

- 1 -

AAV ENCODING HERMANSKY-PUDLAK SYNDROME 1 (HPS1) PROTEIN AND USES THEREOF

RELATED APPLICATIONS

This application claims the benefit under 35 U.S.C. 119(e) of the filing date of U.S.

5 Provisional Application Serial No. 63/179,642, filed April 26, 2021, entitled “**AAV
ENCODING HERMANSKY-PUDLAK SYNDROME 1 (HPS1) PROTEIN AND USES
THEREOF**”, the entire contents of which are incorporated herein by reference.

SEQUENCE LISTING

10 The instant application contains a Sequence Listing which has been submitted in ASCII
format via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy,
created on April 20, 2022, is named U012070156WO00-SEQ-LJG and is 80,982 bytes in size.

BACKGROUND

15 Hermansky-Pudlak syndrome (HPS) is a group of various autosomal recessive
multisystemic disorders. HPS-associated pulmonary fibrosis (HPS-PF) remains the most serious
complication for patients with HPS-1, -2 and -4 types of the disease; 100% of patients with HPS
type 1 develop lethal form of pulmonary fibrosis that manifests as earlier as at 30 – 40 years of
age. Despite the extremely poor prognosis of HPS-PF, no effective treatments have been
20 developed.

SUMMARY

Aspects of the disclosure relate to composition and methods for treating Hermansky-
Pudlak syndrome-associated pulmonary fibrosis in a subject. The disclosure is based, in part, on
25 isolated nucleic acids encoding a Hermansky-Pudlak syndrome 1 (HPS1) protein, that are
amendable to be packaged in recombinant adeno-associated viruses (rAAV). In some
embodiments, the rAAV delivers the HPS1 protein to the lungs (e.g., alveolar type II cells) of
the subject.

In some aspects, the disclosure provides an isolated nucleic acid comprising an
30 expression cassette comprising a nucleotide sequence encoding a Hermansky-Pudlak syndrome

- 2 -

1 (HPS1) protein, wherein the expression cassette is flanked by adeno-associated virus (AAV) inverted terminal repeats (ITRs).

In some embodiments, the Hermansky–Pudlak syndrome 1 (HPS1) protein comprises an amino acid sequence at least 80% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 1-8. In some embodiments, the nucleotide sequence encoding the HPS1 protein comprises a sequence at least 80% identical to the nucleic acid sequence set forth in any one of SEQ ID NOs: 9-16.

In some embodiments, the expression cassette further comprises a promoter operably linked to the nucleotide sequence encoding the HPS1 protein. In some embodiments, the promoter is a cytomegalovirus immediate – early enhancer/chicken β -actin hybrid promoter.

In some embodiments, the isolated nucleic acid further comprises a poly-adenylation (polyA) signal. In some embodiments, the polyA signal is SV40 polyadenylation signal.

In some embodiments, the ITRs are AAV ITRs of a serotype selected from the group consisting of AAV1 ITR, AAV2 ITR, AAV3 ITR, AAV4 ITR, AAV5 ITR, and AAV6 ITR. In some embodiments, the ITRs are AAV2 ITRs, or a variant thereof. In some embodiments, the AAV2 ITR variant is a mutated AAV2 ITR (e.g., a deltaITR or a mutant ITR (mTR)).

In some aspects, the disclosure also provides a vector comprising the isolated nucleic acid described herein. In some embodiments, the vector is a rAAV vector.

In some aspects, the present disclosure also provides a recombinant adeno-associated virus (rAAV) vector comprising a nucleic acid comprising, in 5' to 3' order, (a) a 5' AAV2 ITR; (b) a cytomegalovirus immediate – early enhancer/chicken β -actin hybrid promoter; (c) a nucleotide sequence encoding an Hermansky–Pudlak syndrome 1 (HPS1) protein; (d) a SV40 polyA signal; and (e) a 3' AAV2 ITR.

In some aspects, the present disclosure provides a recombinant adeno-associated virus (rAAV) comprising: (i) an AAV capsid protein; and (ii) the isolated nucleic acid or the vector as described herein. In some embodiments, the capsid protein is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9 capsid protein, or a variant thereof. In some embodiments, the capsid protein is an AAV5 capsid protein, or a variant thereof. In some embodiments, the capsid protein is an AAV8 capsid protein, or a variant thereof.

In some embodiments, the capsid protein is capable of delivering the HPS1 protein into lung cells. In some embodiments, the lung cells are alveolar type II epithelial (ATII) cells.

- 3 -

In some aspects, the disclosure also provides a pharmaceutical composition comprising the isolated nucleic acid, the vector, or the rAAV as described herein. In some embodiments, the pharmaceutical composition further comprising a pharmaceutically acceptable carrier.

5 In some aspects, the present disclosure provides a host cell comprising the isolated nucleic acid, the vector, the rAAV, or the pharmaceutical composition as described herein.

In some aspects, the disclosure provides a method for delivering a Hermansky–Pudlak syndrome 1 (HPS1) protein to a cell or a subject, the method comprising delivering to the cell or the subject the isolated nucleic acid, the vector, the rAAV, the pharmaceutical composition, or the host cell as described herein.

10 In some aspects, the present disclosure also provides a method for treating Hermansky–Pudlak syndrome-associated pulmonary fibrosis in a subject, the method comprising administering to the subject an effective amount of the isolated nucleic acid, the vector, the rAAV, the pharmaceutical composition, or the host cells described herein.

15 In some embodiments, the subject is a human or a non-human mammal. In some embodiments, the administration results in delivery of the HPS1 protein to the lungs of the subject. In some embodiments, the administration results in delivery of the HPS1 protein to alveolar type II epithelial (ATII) cells of the subject.

20 In some embodiments, administration results in an at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or 100% increase in HPS1 expression level or activity in the ATII cells of the subject.

In some embodiments, the administration comprises systemic administration. In some embodiments, the systemic administration is intravenous administration, intramuscular administration, subcutaneous administration, or intraperitoneal administration.

25 In some embodiments, the administration comprises local administration to the lungs. In some embodiments, the local administration to the lungs is Intratracheal instillation, oro-tracheal injection, intranasal administration, oropharyngeal aspiration, or nebulization administration.

BRIEF DESCRIPTION OF DRAWINGS

30 FIGs. 1A-1D show schematics of pCB-mHps1 and pCB-hHPS1 vector constructs described herein. FIGs. 1A-1B show vector maps for AAV- pCB-mHps1 (mouse HPS1). FIGs. 1C-1D show vector maps for AAV- pCB-hHPS1 (human HPS1).

- 4 -

FIGs. 2A-2I show AAVs effectively deliver reporter GFP gene to control muscle and target lung cells. FIGs. 2A-2D show representative anti-GFP protein immunolabeling of control tongue muscle. FIGs. 2E-2G are representative images of GFP protein mRNA hybridization (RNAscope) in tongue muscle control. FIG. 2F is a representative image of GFP protein mRNA hybridization (RNAscope) in lung control. FIGs. 2H-2I are representative images of GFP protein mRNA hybridization (RNAscope) in lung epithelial lining cells.

FIGs. 3A-3D show that the mouse model of Hermansky-Pudlak type 1 deficiency recapitulates human alveolar type 2 histopathology. FIGs. 3A-3B show representative images of Masson Trichrome staining of lung sections showing enlarged foamy alveolar type 2 cells filled with lamellar bodies in *Hps1* knockout (KO) mouse. FIG. 3C is an images of Hematoxylin & eosin staining of the autopsied lung of human HPS1 patient showing foamy swelling of alveolar type 2 cell and interstitial inflammatory cell infiltration. FIG. 3D shows that numerous giant lamellar bodies are observed in the cytoplasm by electron microscopy.

FIGs. 4A-4G show *Hps1* KO mice were susceptible to inflammation. FIG. 4A shows respiratory mechanics parameters from the lung function assessment with the flexiVent system, in LPS-treated *Hps1* KO mice. FIG. 4B shows respiratory mechanics parameters from the lung function assessment with the flexiVent system in LPS-treated wild-type mice. FIGs. 4C-4D show notable pulmonary decline in *Hps1* KO evidenced by reduced static compliance (FIG. 4C) and stiffening of the lung parenchyma (FIG. 4D) 14 days post-challenge. Static compliance showed decreased values, indicating reduced lung compliance and distensibility, which is indicative of structural changes and damage in the lungs. FIG. 4E shows, upon bleomycin challenge, damage to the lung tissue was evident from shift of PV loops downward characteristic to restrictive pulmonary fibrosis phenotype. FIGs. 4F-4G shows increase of tissue airway resistance (airway constriction) (FIG. 4F) and decrease of total lung capacity (FIG. 4G), reflecting the contribution of the conducting airways. The results indicate disease induction in the alveoli and lung parenchyma (peripheral tissues). These findings are consistent with the histopathological findings in mice as well as the pulmonary function in humans, with pulmonary fibrosis and respiratory failure. Results from this study show that *Hps1* KO mouse model is a relevant model to further investigate the pathogenesis of severe pulmonary fibrosis in Hermansky-Pudlak syndrome type 1.

FIGs. 5A-5C show AAV.CB-mHps1 treated mice were protected from pro-fibrotic exacerbations in LPS challenged *Hps1*KO mice by respiratory mechanics parameters from the

- 5 -

lung function assessment with the flexiVent system. FIG. 5A shows significant preservation of pulmonary function was evidenced by shift of pressure volume loops upward. FIG. 5B shows significant preservation of pulmonary function was evidenced by increased static compliance. FIG. 5C shows significant preservation of pulmonary function was evidenced by total lung capacity.

DETAILED DESCRIPTION

Aspects of the disclosure relate to compositions and methods for treating Hermansky-Pudlak syndrome-associated pulmonary fibrosis in a subject. The disclosure is based, in part, on isolated nucleic acid encoding a Hermansky-Pudlak syndrome 1 (HPS1) protein that are amendable to be packaged in recombinant adeno-associated viruses (rAAV). In some embodiments, the rAAV delivers the HPS1 protein to the lungs (e.g., alveolar type II cells) of the subject.

Hermansky-Pudlak syndrome is a disorder characterized by a condition called oculocutaneous albinism, which causes abnormally light coloring (pigmentation) of the skin, hair, and eyes. Affected individuals typically have fair skin and white or light-colored hair. Patients with this disorder typically have a higher than average risk of skin damage and skin cancers caused by long-term sun exposure. Oculocutaneous albinism reduces pigmentation of the colored part of the eye (iris) and the light-sensitive tissue at the back of the eye (retina). Reduced vision, rapid and involuntary eye movements (nystagmus), and increased sensitivity to light (photophobia) are also common in oculocutaneous albinism. In Hermansky-Pudlak syndrome, these vision problems usually remain stable after early childhood. An abnormality of the formation or movement of lysosome-like vesicles may be responsible for the development of the disease.

Some patients with Hermansky-Pudlak syndrome develop breathing problems due to a lung disease called pulmonary fibrosis, which causes scar tissue to form in the lungs. The symptoms of pulmonary fibrosis usually appear during an individual's early thirties and rapidly worsen. Individuals with Hermansky-Pudlak syndrome who develop pulmonary fibrosis often do not live for more than a decade after they begin to experience breathing problems.

There are different types of Hermansky-Pudlak syndrome, which can be distinguished by their signs and symptoms and underlying genetic cause. Mutations in one of 10 genes (e.g., HPS1, AP3B1, HPS3, HPS4, HPS5, HPS6, DTNBP1, BLOC1S3, PLDN, and AP3D1) are

- 6 -

responsible for this disorder. Ten subtypes have been described. Patients with HPS-1, HPS-2, and HPS-4 tend to develop pulmonary fibrosis. In some embodiments, HPS-1 is associated with mutations in the HPS1 gene. In some embodiments, the disclosure relates to gene therapy for (e.g., isolated nucleic acid and rAAVs) to treat Hermansky-Pudlak syndrome associated
5 pulmonary fibrosis.

Isolated Nucleic Acids

A "nucleic acid" sequence refers to a DNA or RNA sequence. In some embodiments, proteins and nucleic acids of the disclosure are isolated. As used herein, the term "isolated" means artificially produced. As used herein, with respect to nucleic acids, the term "isolated" means: (i) amplified in vitro by, for example, polymerase chain reaction (PCR); (ii) recombinantly produced by cloning; (iii) purified, as by cleavage and gel separation; or (iv) synthesized by, for example, chemical synthesis. An isolated nucleic acid is one which is readily manipulable by recombinant DNA techniques well known in the art. Thus, a nucleotide
10 sequence contained in a vector in which 5' and 3' restriction sites are known or for which polymerase chain reaction (PCR) primer sequences have been disclosed is considered isolated but a nucleic acid sequence existing in its native state in its natural host is not. An isolated nucleic acid may be substantially purified, but need not be. For example, a nucleic acid that is isolated within a cloning or expression vector is not pure in that it may comprise only a tiny
15 percentage of the material in the cell in which it resides. Such a nucleic acid is isolated, however, as the term is used herein because it is readily manipulable by standard techniques known to those of ordinary skill in the art. As used herein with respect to proteins or peptides, the term "isolated" refers to a protein or peptide that has been isolated from its natural environment or artificially produced (e.g., by chemical synthesis, by recombinant DNA
20 technology, etc.).

In some aspects, the present disclosure provides an isolated nucleic acid comprising an expression cassette comprises a nucleotide sequence encoding a Hermansky-Pudlak syndrome 1 (HPS1) protein. HPS1 protein plays a role in organelle biogenesis associated with melanosomes, platelet dense granules, and lysosomes. HPS1 is a component of three different protein
30 complexes termed biogenesis of lysosome-related organelles complex (BLOC)-3, BLOC4, and BLOC5. In some embodiments, a human HPS1 protein comprises the amino acid sequence set forth in any one of NCBI Reference Sequence Accession Number: NP_000186.2,

- 7 -

NP_001298274.1, NP_001309405.1, NP_001309406.1, or NP_001309407.1. In some embodiments, a human HPS1 protein is encoded by an mRNA transcript having the sequence set forth in any one of NCBI Reference Sequence Accession Numbers: NM_000195.5, NM_001311345.2, NM_001322476.2, NM_001322477.2, or NM_001322478.2. In some

5 embodiments, a mouse HPS1 protein comprises the amino acid sequence set forth in any one of NCBI Reference Sequence Accession Number: NP_001333632.1, NP_062297.1, or NP_001349339.1. In some embodiments, a human HPS1 protein is encoded by an mRNA transcript having the sequence set forth in any one of NCBI Reference Sequence Accession Numbers: NM_001346703.2, NM_019424.3, or NM_001362410.1.

10 An exemplary human HPS1 protein amino acid sequence as set forth in SEQ ID NO: 1 (NP_000186.2):

MKCVLVATEGAEVLFFYWTDQEFEEESLRLKFGQSENEEEELPALEDQLSTLLAPVIISSMTMLEKLSDTYT
CFSTENGFLYVLHLFGCECLFIAINGDHTSESGDLRRKLYVLKYLFEVHFGLVTVDGHLIRKELRPPDLA
15 QRVQLWEHFQSLWLYSRLREQEQCFAVEALERLIHPQLCELCIEALERHVIQAVNTSPERGGEALHAF
LLVHSKLLAFYSSHSASSLRPADLLALILLVQDLYPSESTAEDDIQSPRRARSSQNIPVQQAWSPHSTG
PTGGSSAETETDSFSLPEEYFTAPSPGDSGSGTIWLEGGTPPMDALQIAEDTLQTLVPHCPVPSGPRR
IFLDANVKESYCPLVPHTMYCLPLWQGINLVLLTRSPSAPLALVLSQLMDGFSMLEKKLKEGPEPGASLR
SQPLVGDRLRQMDKFKVKNRGAQEIQSTWLEFKAKAFSKSEPGSSWELLQACGKLKRQLCAIYRLNFLT
20 PSRGGPHLPQHLQDQVQRLMREKLTWDKDFLLVKSRRNITMVSYLEDFPGLVHFIYVDRTTGQMVAPSLN
CSQKTSSELGKGPLAAAFVKTKVWSLIQLARRYLQKGYTTLLFQEGDFYCSYFLWFENDMGYKLQMIIEVPV
LSDDSVPIGMLGGDYRKLLRYYSKNRPTEAVRCYELLALHLSVIPTDLLVQQAGQLARRLWEASRIPLL

25 An exemplary human HPS1 protein amino acid sequence as set forth in SEQ ID NO: 2 (NP_001298274.1):

MDALQIAEDTLQTLVPHCPVPSGPRRIFLDANVKESYCPLVPHTMYCLPLWQGINLVLLTRSPSAPLALVLSQLMDG
FSMLEKKLKEGPEPGASLRSQPLVGDRLRQMDKFKVKNRGAQEIQSTWLEFKAKAFSKSEPGSSWELLQACGKLKRQL
CAIYRLNFLTTPASRGGPHLPQHLQDQVQRLMREKLTWDKDFLLVKSRRNITMVSYLEDFPGLVHFIYVDRTTGQMV
APSLNCSQKTSSELGKGPLAAAFVKTKVWSLIQLARRYLQKGYTTLLFQEGDFYCSYFLWFENDMGYKLQMIIEVPVLS
30 DDSVPIGMLGGDYRKLLRYYSKNRPTEAVRCYELLALHLSVIPTDLLVQQAGQLARRLWEASRIPLL

 An exemplary human HPS1 protein amino acid sequence as set forth in SEQ ID NO: 3 (NP_001309405.1):

MKCVLVATEGAEVLFFYWTDQEFEEESLRLKFGQSENEEEELPALEDQLSTLLAPVIISSMTMLEKLSDTYT
35 CFSTENGFLYVLHLFGCECLFIAINGDHTSESGDLRRKLYVLKYLFEVHFGLVTVDGHLIRKELRPPDLA
QRVQLWEHFQSLWLYSRLREQEQCFAVEALERLIHPQLCELCIEALERHVIQAVNTSPERGGEALHAFLLVH
SKLLAFYSSHSASSLRPADLLALILLVQDLYPSESTAEDDIQSPRRARSSQNIPVQQAWSPHSTGPTGGSSA
ETETDSFSLPEEYFTAPSPGDQSSGSGTIWLEGGTPPMDALQIAEDTLQTLVPHCPVPSGPRRIFLDANV
KESYCPLVPHTMYCLPLWQGINLVLLTRSPSAPLALVLSQLMDGFSMLEKKLKEGPEPGASLRSQPLVGD
40 RLRQMDKFKVKNRGAQEIQSTWLEFKAKAFSKSEPGSSWELLQACGKLKRQLCAIYRLNFLTTPASR
GGPHLPQHLQDQVQRLMREKLTWDKDFLLVKSRRNITMVSYLEDFPGLVHFIYVDRTTGQMVAPSLNCSQ
KTSSELGKGPLAAAFVKTKVWSLIQLARRYLQKGYTTLLFQEGDFYCSYFLWFE
NDMGYKLQMIIEVPVLSDDSVPIGMLGGDYRKLLRYYSKNRPTEAVRCYELLALHLSVIPTDLLVQQAGQL
ARRLWEASRIPLL

- 8 -

An exemplary human HPS1 protein amino acid sequence as set forth in SEQ ID NO: 4 (NP_001309406.1):

5 MKCVLVATEGAEVLFYWTDQEFEEESLRLKFGQSENEEEELPALEDQLSTLLAPVIISSMTMLEKLSDTYTCFSTENG
NFLYVLHLFGECFLFIAINGDHTSESEGLRRKLYVLKYLFEVHFGLVTVDGHLIRKELRPPDLAQRVQLWEHFQSLW
TYSRLREQEQCFAVEALERLIHPQLCELCEALERHVIQAVNTSPERGGEALHAFLLVHSKLLAFYSSHSASSLRP
ADLLALILLVQDLYPSESTAEDDIQSPRRARSSQNIPVQQAWSPHSTGPTGGSSAETETDSFSLPEEYFTAPSPG
10 DQSSGSTIWLEGGTPPMDALQIAEDTLQTLVPHCPVPSGPRRIFLDANVKESYCPLVPHTMYCLPLWQGINLVLLTR
SPSAPLALVLSQLMDGFSMLEKKLKEGPEPGASLRSQPLVGDRLRQMDKFVKNRGAQEIQSTWLEFKAKAFSKSEPG
SSWELLQACGKLKRQLCAIYRLNFLTTPASRGPHLPQHLQDQVQRLMREKLTWDKDFLLVKSRRNITMVSYLEDFP
GLVHFIYVDRTTGQMVAPSLNCSQKTSSELGKGPLAAAFVKTKVWSLIQLARRYLQKGYTTLLFQEGDFYCSYFLWFE
NDMGYKLQMI EVPVLSDDSVPIGMLGGDYRKLRLRYYSKNRPTEAVRCYELLALHLSVIPDILLVQQAGQLARRLWE
15 ASRIPLL

An exemplary human HPS1 protein amino acid sequence as set forth in SEQ ID NO: 5 (NP_001309407.1):

MKCVLVATEGAEVLFYWTDQEFEEESLRLKFGQSENEEEELPALEDQLSTLLAPVIISSMTMLEKLSDTYTCFSTENG
NFLYVLHLFGECFLFIAINGDHTSESEGLRRKLYVLKYLFEVHFGLVTVDGHLIRKELRPPDLAQRVQLWEHFQSLW
20 TYSRLREQEQCFAVEALERLIHPQLCELCEALERHVIQAVNTSPERGGEALHAFLLVHSKLLAFYSSHSASSLRP
ADLLALILLVQDLYPSESTAEDDIQETDSFSLPEEYFTAPSPGQSSGSTIWLEGGTPPMDALQIAEDTLQTLVPH
CPVPSGPRRIFLDANVKESYCPLVPHTMYCLPLWQGINLVLLTRSPSAPLALVLSQLMDGFSMLEKKLKEGPEPGAS
LRSQPLVGDRLRQMDKFVKNRGAQEIQSTWLEFKAKAFSKSEPGSSWELLQACGKLKRQLCAIYRLNFLTTPASRG
PHLPQHLQDQVQRLMREKLTWDKDFLLVKSRRNITMVSYLEDFPGLVHFIYVDRTTGQMVAPSLNCSQKTSSELGK
25 PLAAAFVKTKVWSLIQLARRYLQKGYTTLLFQEGDFYCSYFLWFEENDMGYKLQMI EVPVLSDDSVPIGMLGGDYRKL
LRYYSKNRPTEAVRCYELLALHLSVIPDILLVQQAGQLARRLWEASRIPLL

An exemplary mouse HPS1 protein amino acid sequence as set forth in SEQ ID NO: 6 (NP_001333632.1):

30 MKCVLVATEGAEVLFYWTDDEEFAESLRLKLQQSEDEEEELPVLEDQLSTLLAPVIISSMTMMEKLSDTYTCFSTEND
NHLYVLHLFGEYLFVAINGDHSESEGLRRKLCVLKYLFEVHFGLVTVDGQLIRKELRPPDLEERARVWKHFQRLG
TYSYLRDREQSFAVEAVERLIHPQLCEQSIETLERHVQAINASPERGGEEVLHAFLLVHCKLLAFYSGHGASTLRP
ADLLALILLVQDLQPSPGTTEEEEEEEEDSDSPQVPEHKGNRRRPKSSQNIPVQQARSQSTSVPTRSSRETDTDSISL
PEEYFTAPSPGQSSGSLVWLDGGTPPSDALQMAEDTPEGLASHSPELSPRRIFLDANIKENYCPLVPHTMYCLP
35 LWPGINMVLLTKSPSTPLALILYQLLDGFSLEKKLKEGQEGSALRSQPFVADLRQKMDKFIKNRVGQEIQNTWLE
FKSKAFSRSEPGSSWELLQVCGKLRQLCVIYRLSFLVTAPSRGGPHLPQHLQDRAQKLMKERLLDWKDFLLVKSRR
NVTMVSYLEDFPGLVHFIYVDRTTGQMVAPSLSPNEKMSSELGKGPLAAAFVKAKVWALVRLARRYLQKGCTTLFQE
GDFRCSYFLWFEENDMGYKLQMI EVPVLSDDSVPIGVLGGDYRKLRLRYYSKSHPSPEPVRCYELLTLHLSVIPDILLV
QQASQLARRLGEASRVTLF
40

An exemplary mouse HPS1 protein amino acid sequence as set forth in SEQ ID NO: 7 (NP_062297.1):

MKCVLVATEGAEVLFYWTDDEEFAESLRLKLQQSEDEEEELPVLEDQLSTLLAPVIISSMTMMEKLSDTYTCFSTEND
NHLYVLHLFGEYLFVAINGDHSESEGLRRKLCVLKYLFEVHFGLVTVDGQLIRKELRPPDLEERARVWKHFQRLG
45 TYSYLRDREQSFAVEAVERLIHPQLCEQSIETLERHVQAINASPERGGEEVLHAFLLVHCKLLAFYSGHGASTLRP
ADLLALILLVQDLQPSPGTTEEEEEEEEDSDSPQRRPKSSQNIPVQQARSQSTSVPTRSSRETDTDSISLPEEYFTPA
PSPGQSSGSLVWLDGGTPPSDALQMAEDTPEGLASHSPELSPRRIFLDANIKENYCPLVPHTMYCLPLWPGINMV
LLTKSPSTPLALILYQLLDGFSLEKKLKEGQEGSALRSQPFVADLRQKMDKFIKNRVGQEIQNTWLEFKSKAFSR
SEPGSSWELLQVCGKLRQLCVIYRLSFLVTAPSRGGPHLPQHLQDRAQKLMKERLLDWKDFLLVKSRRNVTMVSYL
50 EDFPGLVHFIYVDRTTGQMVAPSLSPNEKMSSELGKGPLAAAFVKAKVWALVRLARRYLQKGCTTLFQEGDFRCSYF

- 9 -

LWFENDMGYKLQMIIEVPVLSDDSVPIGVLGDDYRKLRLRYYSKSHPSSEPVRCYELLTLHLSV IPTDLLVQQASQLAR
RLGEASRVTL P

An exemplary mouse HPS1 protein amino acid sequence as set forth in SEQ ID NO: 8

5 (NP_001349339):

MKCVLVATEGAEVLFFYWTDEEFAESLRLKLQQSEDEEEELPVLEDQLSTLLAPVIISSMTMMEKLSDTYTCFSTEND
NHLYVLHLFGYELFVAINGDHSESEGLRRKLCVLKYLFEVHFGLVTVDGQLIRKELRPPDLEERARVWKHFQRLLG
TYSYLRDREQSFAVEAVERLIHPQLCEQSIETLERHVQAINASPERGGEEVLHAFLLVHCKLLAFYSGHGASTLRP
ADLLALILLVQDLQPSPTTEEEEEEEEDSDSPQRRPKSSQNIPVQQARSQSTSVPTRSSRETDTDSISLPEEYFTPA
10 PSPGDQSSGSLVWLDGGTPPSDALQMAEDTPEGLASHSPELPSPRRIFLDANIKENYCPLVPHTMYCLPLWPGINMV
LLTKSPSTPLALILYQLLDGFSLLLEKKLKEGQEAGSALRSQPFVADLRQKMDKFIKNRVGQEIQNTWLEFKSKAFSR
SEPGSSWELLQVCGKLKRQLCVIYRLSFLVTAPSRGGPHLPQHLQDRAQKLMKERLLDWKDFLLVKSRRNVMTMVSYL
EDFPGLVHFIIYVDRTTGQMVAPSLSPNEKMSELGKGPLAAAFVKAKVWSPRTRVLRQHSALSLSCSHTPAEWTVGS
TQEASFSALLI

15

An exemplary nucleotide sequence encoding HPS1 protein (NP_000186.2) is set forth in
SEQ ID NO: 9 (NM_000195.5):

AAGCGCGCCCGGGCTCCTGCAGGCGGGGCGCTGTGCGCGCCGCGATCCGGTACGTGGGCCTCCGGGCTGTCCCCTCT
GGGGGCGGCGATCCTCCCCTCCGGAGCCCCCTTCAACCCCTCCCGGAAGTGAGGACCAGGGATGCTGTGCTGCTCTCC
20 CATGAGCCAGTCACCGAGTCGGTCTGCTGCAGCCCTTTCTGAACCTCTGGCCGTCTGGATGCTCCACTGTGCTTGCC
AAGATGAAGTGCCTTTGGTGGCCACTGAGGGCGCAGAGGTCCTCTTCTACTGGACAGATCAGGAGTTTGAAGAGAG
TCTCCGGCTGAAGTTCCGGGCAGTCAGAGAATGAGGAAGAAGAGCTCCCTGCCCTGGAGGACCAGTCAGCACCCCTCC
TAGCCCCGGTCATCATCTCCTCCATGACGATGCTGGAGAAGCTCTCGGACACCTACACCTGCTTCTCCACGAAAAT
GGCAACTTCTGTATGTCTTACCTGTTTGGAGAATGCCTGTTCAATTGCCATCAATGGTGACCACACCGAGAGCGA
25 GGGGACCTGCGGCGGAAGCTGTATGTGCTCAAGTACCTGTTTGAAGTGCACTTTGGGCTGGTGACTGTGGACGGTC
ATCTTATCCGAAAGGAGCTGCGGCCCCCAGACCTGGCGCAGCGTGTCAGCTGTGGGAGCACTTCAGAGCCTGCTG
TGGACCTACAGCCGCTGCGGGAGCAGGAGCAGTGCTTCGCCGTGGAGGCCCTGGAGCGACTGATTCACCCCCAGCT
CTGTGAGCTGTGCATAGAGGCGCTGGAGCGGCACGTATCCAGGCTGTCAACACCAGCCCCGAGCGGGGAGGCGAGG
AGGCCCTGCATGCCTTCTGCTCGTGCACTCCAAGCTGCTGGCATTCTACTCTAGCCACAGTGCCAGCTCCCTGCGC
30 CCGGCCGACCTGCTTGCCCTCATCCTCCTGGTTTCAGGACCTCTACCCAGCGAGAGCACAGCAGAGGACGACATTCA
GCCTTCCCCGCGGAGGGCCCCGAGCAGCCAGAACATCCCCGTGCAGCAGGCCTGGAGCCCTCACTCCACGGGCCCAA
CTGGGGGGAGCTCTGCAGAGACGGAGACAGACAGCTTCTCCCTCCCTGAGGAGTACTTCACACCAGCTCCTTCCCCT
GGCGATCAGAGCTCAGGTAGCACCATCTGGCTGGAGGGGGGCACCCCCCATGGATGCCCTTCAGATAGCAGAGGA
CACCTTCAAACACTGGTTCCCCACTGCCCTGTGCCCTTCCGGCCCCAGAAGGATCTTCTGGATGCCAACGTGAAGG
35 AAAGCTACTGCCCCCTAGTGCCCCACACCATGTACTGCCTGCCCTGTGGCAGGGCATCAACCTGGTGCTCCTGACC
AGGAGCCCCAGCGCGCCCCCTGGCCCTGGTTCTGTGCCAGCTGATGGATGGCTTCTCCATGCTGGAGAAGAAGCTGAA
GGAAGGGCCGAGCCCCGGGGCCCTCCCTGCGCTCCAGCCCCCTCGTGGGAGACCTGCGCCAGAGGATGGACAAGTTTG
TCAAGAATCGAGGGGACAGGAGATTAGAGCACCTGGCTGGAGTTTAAGGCCAAGGCTTTCTCCAAAAGTGAGCCC
GGATCCTCCTGGGAGCTGCTCCAGGCATGTGGGAAGCTGAAGCGGCAGCTCTGCGCCATCTACCGCTGAACCTTCT
40 GACCACAGCCCCCAGCAGGGGAGGGCCACACCTGCCCCAGCACCTGCAGGACCAAGTGACAGAGGCTCATGCGGGAGA
AGCTGACGGAAGTGAAGGACTTCTTGGTGGTGAAGAGCAGGAGGAACATCACCATGGTGTCTTACCTAGAAGACTTC
CCAGGCTTGGTGCACTTCATCTATGTGGACCGCACCCTGGGCAGATGGTGGCGCCTTCCCTCAACTGCAGTCAAAA
GACCTCGTCGGAGTTGGGCAAGGGGCGCTGGCTGCCTTTGTCAAAAATAAGGTCTGGTCTCTGATCCAGCTGGCGC
GCAGATACCTGCAGAAGGGCTACACCACGCTGCTGTTCCAGGAGGGGGATTCTACTGCTCCTACTTCTGTGGTTT
45 GAGAATGACATGGGGTACAACTCCAGATGATCGAGGTGCCCTGCTCTCCGACGACTCAGTGCTTATCGGCATGCT
GGGAGGAGACTACTACAGGAAGCTCCTGCGCTACTACAGCAAGAACC GCCCAACCAGGCTGTGAGGTGCTACGAGC
TGCTGGCCCTGCACCTGTCTGTCTATCCCCACTGACCTGCTGGTGCAGCAGGCCGGCCAGCTGGCCCGGCGCCTCTGG
GAGGCCTCCCGTATCCCCCTGCTCTAGGCCAAGGTGGCCGAGCTTGCCCTTGCATCCTGTCTCCAGCCACCCTTG
CTTGCCACTGTTCCCCATGACGAGAGCCTCCTGTCTGCAGTGGCCATCCTGAGGATAGGGCAGAGTGCCAGGGTGG
50 CCCCAGGGCTTCTAAAACCCACCTAGACCACCCTCCATGTGTCAGGTACTGAGCAAGGCCCCAGATCCTTCTCTCTGG
AGGAAGAGGGAAGCCCAGGGGTCTGTTTGTAAAAACAACGGTGGCAACAGCTCCTCTTCCAGAGCTGCCTCTGCCTT
TATCCTGGGAGATGGGGAGGAAGCCCCATCTCTGCTGTTCCCTGCGTGGAGGAAGCCACCCAGCAAGCTCTCTCCT
ACCCAGGTAAAAGGTGCTCCTTTGCCTGGGTTTGAATTCAGCGCTGCCACTTCTCTCTGCACCTCCTGGCAAGT

- 10 -

TTCTTCTATTCCCCACGTTTAAAGCGATGGCACCTCCGTCCCAGGGTGGTGTGAGGATTACCCAGTGTGGTAGGTGC
TCAATAAATGTTGGTCATTGTTATCACTGAAGCCCAACATGCTAGTGCTTCTAGACCCCTTCTGTCAGTGTGATAAG
CCCTTGCTAAGTCCCAGCCCCCTTCATGCTTGGCTGGCGTCTGCCCTAGGGCTGGGGTTCTCAAGCCCCCTGGCCCTGG
5 CCCAGAGATTTGGATTCCCTTGGCGGCCGTGGAGCCCAGCCTTTGATGTCTTTCAAAGCTTCTGTGGTGCGCCCTGG
ATTGAGAACCACCACCCGAGGGGTACAGCCCCCTCTCTTCCAACCCGAGAAGTTCCTGTCCCAGAATGGACCCAGGGAC
AAGAGACCCTGAGAGCCCTGGGACTGGGAGTGTCTGCTCCTCTGAGGCCAGGAGGCCGGTGTGGGCCAGAGAGGAC
GGCGTGGCGAAAGTCAGCGTCCACTGCAGCACAGGATCAGATGGCCGTGTGCTGTGCATGCAGGAGCCTCGCCTTCT
GTGTCTTTAGTCTTGAGCCAAAATTTGCTCAAAGACTGATCTCTTCCCTGCAGGGAACAGCTTTGGGGCTGGGGGAA
CTAGAACCACATGTTGGTCTAAACCCTGAGAAGGTGGCAGTGAGGAAGTATCCCCCTCAGGTGACTGGATCTGTGTT
10 CCTCCTTAACATCATCTGATGGAATGGCAATGAAAAGCGTGGATTGTGGAAAAATACAGAAAAACATAAAGGAAAAAA
CTCCAATCCCCTGAGCCCCACCACTGTTTCAGGACCCCTGCTTTTGTACCTACTATTTCCCTTTAGTTTTTAGCAGCG
GCTGGATGTGATATGTCTAGTTTAACCAGTCCCCCTTGATCTTTCTATATAATAAATAACACAGGAGTGAACATCCTG
AATCAGA

15 An exemplary nucleotide sequence encoding HPS1 protein (NP_001298274.1) is set forth in SEQ ID NO: 10 (NM_001311345.2):

AAGCGCGCCCGGGCTCCTGCAGGCGGGGCGCTGTGCGCGCCGCGATCCGGTACGTGGGCCTCCGGGCTGTCCCCTCT
GGGGGCGGCGATCCTCCCTCCGGAGCCCCCTTCAACCCTCCCGGAAGTGAGCATTTCAGGACCAGGGATGCTGTG
CTGCTCTCCCATGAGCCAGTCACCGAGTCGGTCTGCTGCAGCCCTTTCTGAACCTCTGGCCGTCTGGATGCTCCACT
20 GTGCTTGCCAAGATGAAGTGCCTTGGTGGCCACTGAGGGCGCAGAGGTCCCTCTTCTACTGGACAGATCAGGAGTT
TGAAGAGAGTCTCCGGCTGAAGTTCGGGCAGTCAGAGAATGAGGAAGAAGAGCTCCCTGCCCTGGAGGACCAGCTCA
GCACCCTCCTAGCCCCGGTCATCATCTCCTCCATGACGATGCTGGAGAAGCTCTCGGACACCTACACCTGCTTCTCC
ACGGAAAATGGCAACTTCCTGTATGTCTTCACTGTTTGGAGAATGCCTGTTCAATTGCCATCAATGGTGACCACAC
CGAGAGCGAGGGGGACCTGCGGCGGAAGCTGTATGTCTCAAGTACCTGTTTGAAGTGCACCTTTGGGCTGGTGACTG
25 TGGACGGTCACTTATCCGAAAGGAGCTGCGGCCCCCAGACCTGGCGCAGCGTGTCCAGCTGTGGGAGCACTTCCAG
AGCCTGCTGTGGACCTACAGCCGCCTGCGGGAGCAGGAGCAGTGCTTCGCCGTGGAGGTGACTACTGGGCGCGGGAG
TCCTCAGGACTTGAAGCTGTGAGGCCCTGGAGCGACTGATTACCCCCAGCTCTGTGAGCTGTGCATAGAGGCGCTG
GAGCGGCACGTATCCAGGCTGTCAACACCAGCCCCGAGCGGGGAGGCGAGGAGGCCCTGCATGCCTTCCCTGCTCGT
GCACTCCAAGCTGCTGGCATTCTACTCTAGCCACAGTGCCAGCTCCCTGCGCCCGGCCGACCTGCTTGCCCTCATCC
30 TCCTGGTTTCAGGACCTCTACCCAGCGAGAGCACAGCAGAGGACGACATTCAGGAGACAGACAGCTTCTCCCTCCCT
GAGGAGTACTTCACACCAGCTCCTTCCCCTGGCGATCAGAGCTCAGGTAGCACCATCTGGCTGGAGGGGGGCACCCC
CCCCATGGATGCCCTTCAGATAGCAGAGGACACCCCTCCAAACACTGGTTCCCCACTGCCCTGTGCCTTCCGGCCCCA
GAAGGATCTTCTGGATGCCAACGTGAAGGAAAGCTACTGCCCCCTAGTGCCCCACACCATGTACTGCCCTGCCCTG
TGGCAGGGCATCAACCTGGTGTCTCTGACCAGGAGCCCCAGCGCGCCCCCTGGCCCTGGTTCTGTCCCAGCTGATGGA
35 TGGCTTCTCCATGCTGGAGAAGAAGCTGAAGGAAGGGCCGGAGCCCCGGGGCCTCCCTGCGCTCCCAGCCCCCTCGTGG
GAGACCTGCGCCAGAGGATGGACAAGTTTGTCAAGAATCGAGGGGCACAGGAGATTTCAGAGCACCTGGCTGGAGTTT
AAGGCCAAGGCTTTCTCCAAAAGTGAGCCCGGATCCTCCTGGGAGCTGCTCCAGGCATGTGGGAAGCTGAAGCGGCA
GCTCTGCGCCATCTACCGCTGAACCTTTCTGACCACAGCCCCCAGCAGGGGAGGCCACACCTGCCCCAGCACCTGC
AGGACCAAGTGCAGAGGCTCATGCGGGAGAAGCTGACGGACTGGAAGGACTTCTTGCTGGTGAAGAGCAGGAGGAAC
40 ATCACCATGGTGTCTACCTAGAAGACTTCCAGGCTTGGTGCACTTCATCTATGTGGACCGCACCCTGGGCAGAT
GGTGGCGCCTTCCCTCAACTGCAGTCAAAGACCTCGTCGGAGTTGGGCAAGGGGCCGCTGGCTGCCTTTGTCAAAA
CTAAGGTCTGGTCTCTGATCCAGCTGGCGCGCAGATACCTGCAGAAGGGCTACACCACGCTGCTGTTCCAGGAGGGG
GATTTCTACTGCTCCTACTTCTGTGGTTTCGAGAATGACATGGGGTACAACTCCAGATGATCAGGTGCCCGTCTCT
CTCCGACGACTCAGTGCTATCGGCATGCTGGGAGGAGACTACTACAGGAAGCTCCTGCGCTACTACAGCAAGAACC
45 GCCCAACCGAGGCTGTGAGGTGCTACGAGCTGCTGGCCCTGCACCTGTCTGTCTATCCCCACTGACCTGCTGGTGCAG
CAGGCCGGCCAGCTGGCCCGGCGCCTCTGGGAGGCCCTCCCGTATCCCCCTGCTCTAGGCCAAGGTGGCCGCAGTCTG
CCTTTGCATCCTGTCTCCAGCCACCCTTGCTTGCCACTGTTCCCCATGACGAGAGCCTCCTGTCTGCAGTGGCCAT
CCTGAGGATAGGGCAGAGTGCCAGGGTGGCCCCAGGGCTTCTAAAACCCACCTAGACCACCTCCATGTGAGGTA
CTGAGCAAGGCCCCAGATCCTTCTCTCTGGAGGAAGAGGGAAGCCCAGGGGTCTGTGTTGTAACAAACAGGTGGCAA
50 CAGCTCCTCTTCCAGAGCTGCCTCTGCCTTTATCCTGGGAGATGGGGAGGAAGCCCCATCTCTGCTGTTCCCTGCGT
GGAGGAAGCCACCCAGCAAGCTCTCTCTTACCCAGGTAAAGGTGCTCCTTTGCTGGTGTGAATTCCAGCGCT
GCCACTTCTCTCTGCACCTCCTGGCAAGTTTCTTCTATTCCCCACGTTTAAAGCGATGGCACCTCCGTCCCAGGGT
GGTGTGAGGATTACCCAGTGTGGTAGGTGCTCAATAAATGTTGGTCATTGTTATCACTGAAGCCCAACATGCTAGTG
CTTCTAGACCCTTCTGTGAGTGTGATAAGCCCTTGCTAAGTCCCAGCCCCCTTCATGCTTGGCTGGCGTCTGCCCTA
55 GGGCTGGGGTTCTCAAGCCCCCTGGCCCTGGCCAGAGATTTGGATTCCCTTGGCGGCCGTGGAGCCCAGCCTTTGAT
GTCTTTCAAAGCTTCTGTGGTGCGCCCTGGATTGAGAACCACCACCCGAGGGGTACAGCCCCCTCTCTTCCAACCGAG

- 11 -

AAGTTCCTGTCCCAGAATGGACCCAGGGACAAGAGACCCTGAGAGCCCTGGGACTGGGAGTGTCTGCTCCTCTGAGG
 CCAGGAGGCCCGGTGCTGGGCCAGAGAGGACGGCGTGCGAAAAGTCAGCGTCCACTGCAGCACAGGATCAGATGGCCG
 TGTGCTGTGCATGCAGGAGCCTCGCCTTCTGTGTCTTTAGTCTTGAGCCAAAATTTGCTCAAAGACTGATCTCTTCC
 TTGCAGGGAACAGCTTTGGGGCTGGGGGAACTAGAACCCACATGTTGGTCTAAACCCTGAGAAGGTGGCAGTGAGGA
 5 AGTATCCCCCTCAGGTGACTGGATCTGTGTTCCCTCCTTAACATCATCTGATGGAATGGCAATGAAAAGCGTGGATTGT
 GGAAAATACAGAAAAACATAAAGGAAAAAACTCCAATCCCCCTGAGCCCACCCTGTTTCAGGACCCCTGCTTTTGTCA
 CCTACTATTTCCCTTTAGTTTTTAGCAGCGGCTGGATGTGATATGTCTAGTTTAACCAGTCCCCCTTGATCTTTCTAT
 ATAATAAATAACACAGGAGTGAACATCCTGAATCAGA

10 An exemplary nucleotide sequence encoding HPS1 protein (NP_001309405.1) is set
 forth in SEQ ID NO: 11 (NM_001322476.2):

AAGCGCGCCCGGGCTCCTGCAGGCGGGGCGCTGTGCGCGCCGCGATCCGCATTTGCAGGACCAGGGATGCTGTGCTG
 CTCTCCCATGAGCCAGTCACCGAGTCGGTCTGCTGCAGCCCTTTCTGAACCTCTGGCCGTCTGGATGCTCCACTGTG
 CTTGCCAAGATGAAGTGCCTCTTGGTGGCCACTGAGGGCGCAGAGGTCCCTCTTCTACTGGACAGATCAGGAGTTTGA
 15 AGAGAGTCTCCGGCTGAAGTTCGGGCAGTCAGAGAATGAGGAAGAAGAGCTCCCTGCCCTGGAGACCAGCTCAGCA
 CCGCTCCTAGCCCCGGTCATCTCCTCCATGACGATGCTGGAGAAGCTCTCGGACACCTACACCTGCTTCTCCACG
 GAAAATGGCAACTTCCTGTATGTCTTACCTGTTTGGAGAATGCCGTGTTCAATTGCCATCAATGGTGACCACACCGA
 GAGCGAGGGGGACCTGCGGCGGAAGCTGTATGTGCTCAAGTACCTGTTTGAAGTGCACCTTTGGGCTGGTGAAGTGTG
 ACGGTCTCTTATCCGAAAGGAGCTGCGGCCCCCAGACCTGGCGCAGCGTGTCCAGCTGTGGGAGCACTTCCAGAGC
 20 CTGCTGTGGACCTACAGCCGCTGCGGGAGCAGGAGCAGTGTTCGCCGTGGAGGCCCTGGAGCGACTGATTACCCC
 CCAGCTCTGTGAGCTGTGCATAGAGGCGCTGGAGCGGCACGTCTCCAGGCTGTCAACACCAGCCCCGAGCGGGGAG
 GCGAGGAGGCCCTGCATGCCTTCTGCTCGTGCACCTCCAAGCTGCTGGCATTCTACTCTAGCCACAGTGCCAGCTCC
 CTGCGCCCCGGCGACCTGCTTGCCTCATCCTCTGGTTTCAGGACCTCTACCCAGCGAGAGCACAGCAGAGGACGA
 CATTTCAGCCTTCCCCGCGGAGGGCCCCGAGCAGCCAGAACATCCCCGTGCAGCAGGCCTGGAGCCCTCACTCCACGG
 25 GCCCAACTGGGGGGAGCTCTGCAGAGACGGAGACAGACAGCTTCTCCCTCCCTGAGGAGTACTTCACACCAGCTCCT
 TCCCCCTGGCGATCAGAGCTCAGGTAGCACCATCTGGCTGGAGGGGGGCACCCCCCATGGATGCCCTTCAGATAGC
 AGAGGACACCCCTCCAAACACTGGTTCCCCACTGCCCTGTGCCCTTCCGGCCCCAGAAGGATCTTCTGGATGCCAACG
 TGAAGGAAAGCTACTGCCCCCTAGTGCCCCACACCATGTACTGCCTGCCCCGTGGCAGGGCATCAACCTGGTGCTC
 CTGACCAGGAGCCCCAGCGCGCCCCCTGGCCCTGGTTCTGTCCAGCTGATGGATGGCTTCTCCATGCTGGAGAAGAA
 30 GCTGAAGGAAGGGCCGGAGCCCCGGGGCCTCCCTGCGCTCCAGCCCCCTCGTGGGAGACCTGCGCCAGAGGATGGACA
 AGTTTGTCAAGAATCGAGGGGCACAGGAGATTGAGAGCACCTGGCTGGAGTTTAAGGCCAAGGCTTCTCCAAAAGT
 GAGCCCCGATCCTCCTGGGAGCTGCTCCAGGCATGTGGGAAGCTGAAGCGGCAGCTCTGCGCCATCTACCGGCTGAA
 CTTTCTGACCACAGCCCCAGCAGGGGAGGCCACACCTGCCCCAGCACCTGCAGGACCAAGTGCAGAGGCTCATGC
 GGGAGAAGCTGACGGACTGGAAGGACTTCTTGTGTTGAAGAGCAGGAGGAACATCACCATGGTGTCTTACCTAGAA
 35 GACTTCCCAGGCTTGGTGCACCTTCATCTATGTGGACCGCACCACTGGGCAGATGGTGGCGCCTTCCCTCAACTGCAG
 TCAAAAGACCTCGTCGGAGTTGGGCAAGGGGCCGCTGGCTGCCTTTGTCAAAACTAAGGTCTGGTCTCTGATCCAGC
 TGGCGCGCAGATACCTGCAGAAGGGCTACACCACGCTGCTGTTCAGGAGGGGGATTTCTACTGCTCCTACTTCTCTG
 TGGTTTCGAGAATGACATGGGGTACAACTCCAGATGATCGAGGTGCCCGTCCCTCTCCGACGACTCAGTGCCTATCGG
 CATGCTGGGAGGAGACTACTACAGGAAGCTCCTGCGCTACTACAGCAAGAACC GCCCAACCGAGGCTGTGAGGTGCT
 40 ACGAGCTGCTGGCCCTGCACCTGTCTGTCTATCCCCACTGACCTGCTGGTGCAGCAGGCCGGCCAGCTGGCCCGGCGC
 CTCTGGGAGGCCCTCCCGTATCCCCCTGCTCTAGGCCAAGGTGGCCGCACTGTGCCTTTGCATCCTGTCTCCAGCCA
 CCCTTGCTTGCCACTGTTCCCCATGACGAGAGCCTCCTGTCTGCAGTGGCCATCCTGAGGATAGGGCAGAGTGCCCA
 GGGTGGCCCCAGGGCTTCTAAAACCCACCTAGACCACCTCCATGTCTAGGTACTGAGCAAGGCCCCAGATCCTTCT
 CTCTGGAGGAAGAGGGAAGCCCCAGGGGTCTGTTTGTAAAAACAACGGTGGCAACAGCTCCTCTTCCAGAGCTGCCTC
 45 TGCCTTTATCCTGGGAGATGGGGAGGAAGCCCCATCTCTGCTGTTCCCTGCGTGGAGGAAGCCCACCCAGCAAGCTC
 TCTCCTACCCAGGTAAAAGGTGCTCCTTTGCCTGGGTTTGAATTCACGCGTGCCACTTCTCTCTGCACCTCCTG
 GCAAGTTTCTTCTATTCCCCACGTTTAAAGCGATGGCACCTCCGTCCAGGGTGGTGTGAGGATTACCCAGTGTGGT
 AGGTGCTCAATAAATGTTGGTCATTGTTATCACTGAAGCCCAACATGCTAGTGCTTCTAGACCTTCTGTGAGTGTCT
 GATAAGCCCTTGCTAAGTCCCAGCCCCCTCATGCTTGGCTGGCGTCTGCCCTAGGGCTGGGGTTCTCAAGCCCCCTGG
 50 CCGTGGCCCCAGAGATTGGATTCCCTTGGCGGCGCTGGAGCCAGCCTTTGATGTCTTTCAAAGCTTCTGTGGTGGC
 CCCTGGATTGAGAACCACCCAGGGGTACAGCCCCCTCTCTTCCCAACCAGAAAGTTCTCTCCAGAAATGGAGCCC
 AGGGACAAGAGACCCTGAGAGCCCTGGGACTGGGAGTGTCTGCTCCTCTGAGGCCAGGAGGCCGGTGTGGGCCAGA
 GAGGACGGCGTGGCGAAAAGTCAGCGTCCACTGCAGCACAGGATCAGATGGCCGTGTGCTGTGCATGCAGGAGCCTCG
 CCTTCTGTGTCTTTAGTCTTGAGCCAAAATTTGCTCAAAGACTGATCTCTTCCCTTGCAGGGAACAGCTTTGGGGCTG
 55 GGGGAAGTAGAACCACACATGTTGGTCTAAACCCTGAGAAGGTGGCAGTGAGGAAGTATCCCCCTCAGGTGACTGGATC
 TGTGTTCTCCTTAACATCATCTGATGGAATGGCAATGAAAAGCGTGGATTGTGGAAAATACAGAAAAACATAAAGG

- 12 -

AAAAAACTCCAATCCCCTGAGCCCCACCACTGTTTCAGGACCCCTGCTTTTGTACCTACTATTTCCCTTTAGTTTTTA
GCAGCGGCTGGATGTGATATGTCTAGTTTAAACCAGTCCCCTTGATCTTTCTATATAATAAATAACACAGGAGTGAAC
ATCCTGAATCAGA

5 Exemplary nucleotide sequence encoding HPS1 protein (NP_001309406.1) is set forth in
SEQ ID NO: 12 (NM_001322477.2):

AAGCGCGCCCCGGGCTCCTGCAAGGCGGGGCGCTGTGCGCGCCGCGATCCGGTACGTGGGCCTCCGGGCTGTCCCCCTCT
GGGGGCGGCGATCCTCCCTCCGGAGCCCCCTTCAACCCTCCCGGAAGTGAGCATTTCAGGACCAGGGATGCTGTG
CTGCTCTCCCATGAGCCAGTCAACGAGTCGGTCTGCTGCAGCCCTTTCTGAACCTCTGGCCGTCTGGATGCTCCACT
10 GTGCTTGCCAAGATGAAGTGCCTTGGTGGCCACTGAGGGCGCAGAGGTCCCTCTTCTACTGGACAGATCAGGAGTT
TGAAGAGAGTCTCCGGCTGAAGTTCGGGCAGTCAGAGAATGAGGAAGAAGAGCTCCCTGCCCTGGAGGACCAGCTCA
GCACCTCCTAGCCCCGGTCATCATCTCCTCCATGACGATGCTGGAGAAGCTCTCGGACACCTACACCTGCTTCTCC
ACGGAAAATGGCAACTTCCTGTATGTCTTCACCTGTTTGGAGAATGCCTGTTTCATTGCCATCAATGGTGACCACAC
CGAGAGCGAGGGGACCTGCGGCGGAAGCTGTATGTGCTCAAGTACCTGTTTGAAGTGCACCTTTGGGCTGGTGAATG
15 TGGACGGTCACTTTATCCGAAAGGAGCTGCGGCCCCCAGACCTGGCGCAGCGTGTCCAGCTGTGGGAGCACTTCCAG
AGCCTGCTGTGGACCTACAGCCGCTGCGGGAGCAGGAGCAGTGTTCGCCGTGGAGGCCCTGGAGCGACTGATTCA
CCCCCAGCTCTGTGAGCTGTGCATAGAGGCGCTGGAGCGGCACGTATCCAGGCTGTCAACACCAGCCCCGAGCGGG
GAGGCGAGGAGGCCCTGCATGCCTTCCTGCTCGTGCCTTCCAAGCTGCTGGCATTCTACTCTAGCCACAGTGCCAGC
TCCCTGCGCCCGGCCGACCTGCTTGCCCTCATCTCCTGGTTCAGGACCTCTACCCCAGCGAGAGCACAGCAGAGGA
20 CGACATTACAGCCTTCCCCGCGGAGGGCCCGGAGCAGCCAGAACATCCCCGTGCAGCAGGCCCTGGAGCCCTCACTCCA
CGGGCCCAACTGGGGGAGCTCTGCAGAGACGGAGACAGACAGCTTCTCCCTCCCTGAGGAGTACTTCACACCAGCT
CCTTCCCTTGCGATCAGAGCTCAGGTAGCACCATCTGGCTGGAGGGGGCACCCCCCATGGATGCCCTTCAGAT
AGCAGAGGACACCTCCAAACACTGGTTCCTCCACTGCCCTGTGCCTTCCGGCCCCAGAAGGATCTTCCCTGGATGCCA
ACGTGAAGGAAAGCTACTGCCCCCTAGTGCCCCACACCATGTACTGCCTGCCCTGTGGCAGGGCATCAACCTGGTG
25 CTCCTGACCAGGAGCCCCAGCGCGCCCTGGCCCTGGTTCTGTCCCAGCTGATGGATGGCTTCTCCATGCTGGAGAA
GAAGCTGAAGGAAGGGCCGGAGCCCCGGGGCCTCCCTGCGCTCCCAGCCCCCTCGTGGGAGACCTGCGCCAGAGGATGG
ACAAGTTTGTCAAGAAATCGAGGGGCGACAGGAGATTTCAGAGCACCTGGCTGGAGTTTAAGGCCAAGGCTTTCTCCAAA
AGTGAGCCCGGATCCTCCTGGGAGCTGCTCCAGGCATGTGGGAAGCTGAAGCGGCAGCTCTGCGCCATCTACCGGCT
GAACTTTCTGACCACAGCCCCCAGCAGGGGAGGCCACACCTGCCCCAGCACCTGCAGGACCAAGTGCAGAGGCTCA
30 TGCGGGAGAAGCTGACGGACTGGAAGGACTTCTTGCTGGTGAAGAGCAGGAGGAACATCACCATGGTGTCTTACCTA
GAAGACTTCCCAGGCTTGGTGCCTTCTATGTGAGCCGACCACTGGGCAGATGGTGGCGCCTTCCCTCAACTG
CAGTCAAAAGACCTCGTCGGAGTTGGGCAAGGGGCGCTGGCTGCCCTTTGTCAAAACTAAGGTTCTGGTCTCTGATCC
AGCTGAGCGCGCAGATACCTGCAGAAGGGCTACACCACGCTGCTGTTCCAGGAGGGGATTTCTACTGCTCTCTACTTC
CTGTGGTTTCGAGAATGACATGGGGTACAAACTCCAGATGATCGAGGTGCCCTTCTCCGACGACTCAGTGCCTAT
35 CGGCATGCTGGGAGGAGACTACTACAGGAAGCTCCTGCGCTACTACAGCAAGAACCGCCCAACCGAGGCTGTCAAGT
GCTACGAGCTGCTGGCCCTGCACCTGTCTGTCTATCCCCACTGACCTGCTGGTGCAGCAGGCGGCCAGCTGGCCCGG
CGCCTCTGGGAGGCCTCCCGTATCCCCCTGCTCTAGGCCAAGGTGGCCGCGAGTCTGCCCTTTCATCTGTCTCTCCAG
CCACCTTGTCTTGCCACTGTTCCCCATGACGAGAGCCTCCTGTCTGCAGTGGCCATCCTGAGGATAGGGCAGAGTGC
CCAGGGTGGCCCCAGGGCTTCTAAACCCACCTAGACCACCTCCATGTCAGGTACTGAGCAAGGCCCCAGATCCT
40 TCTCTCTGGAGGAAGAGGGAAGCCAGGGGTCTGTTTGTAAACAACGGTGGCAACAGCTCCTCTTCCAGAGCTGC
CTCTGCCCTTATCCTGGGAGATGGGGAGGAAGCCCCATCTCTGCTGTTCCCTGCGTGGAGGAAGCCCCACCCAGCAAG
CTCTCTCCTACCCCAGGTAAAGGTGCTCCTTTGCCCTGGGTTTGAATTCCAGCGCTGCCACTTCTCTCTGCACCTC
CTGGCAAGTTTCTTCTATTCCCCACGTTTAAAGCGATGGCACCTCCGTCCCAGGGTGGTGTGAGGATTACCCAGTGT
GGTAGGTGCTCAATAAATGTTGGTCATTGTTATCACTGAAGCCCCAATGCTAGTGCTTCTAGACCCTTCTGTCAAGT
45 GCTGATAAGCCCTTGCTAAGTCCCAGCCCCCTTCATGCTTGCTGGCTGGCGTCTGCCCTAGGGCTGGGGTTCTCAAGCCCC
TGGCCCTGGCCCAGAGATTTGGATTCCCTTGGCGGCCGTGGAGCCCAGCCTTTGATGTCTTTCAAAGCTTCTGTGGT
GCGCCCTGGATTGAGAACCACCCAGGGGTACAGCCCCCTCTCTTCCAACCGAGAAGTTCCTGTCCCAGAATGGA
CCCAGGGACAAGAGACCCTGAGAGCCCTGGGACTGGGAGTGTCTGCTCCTCTGAGGCCAGGAGGCCGTGCTGGGCC
AGAGAGGACGGCGTGGCGAAAGTCAGCGTCCACTGCAGCACAGGATCAGATGGCCGTGTGCTGTGCATGCAGGAGCC
50 TCGCCTTCTGTGCTTTTACGTCTTGAGCCAAAATTTGCTCAAAGACTGATCTTCTCCTTGACAGGAACAGCTTTGGGG
CTGGGGGAACATAAGAACCATGTTGGTCTAAACCTGAGAAAGTGGCAGTGAGGAAGTATCCCCCTCAGGTGACTGG
ATCTGTGTTTCTCTTAACATCATCTGATGGAATGGCAATGAAAAGCGTGGATTGTGGAATAACAGAAAAACATAA
AGGAAAAAACTCCAATCCCCTGAGCCCCACCACTGTTTCAGGACCCCTGCTTTTGTACCTACTATTTCCCTTTAGTTT
TTAGCAGCGGCTGGATGTGATATGTCTAGTTTAAACCAGTCCCCTTGATCTTTCTATATAATAAATAACACAGGAGTG
55 AACATCCTGAATCAGA

- 13 -

An exemplary nucleotide sequence encoding HPS1 protein (NP_001309407.1) is set forth in SEQ ID NO: 13 (NM_001322478.2):

```

AAGCGCGCCCGGGCTCCTGCGAGGCGGGGCGCTGTGCGCGCCGCGATCCGCATTTGCAGGACCAGGGATGCTGTGCTG
CTCTCCCATGAGCCAGTCACCGAGTCGGTCTGCTGCGAGCCCTTTCTGAACCTCTGGCCGTCTGGATGCTCCACTGTG
5 CTTGCCAAGATGAAGTTCGCTCTTGGTGGCCACTGAGGGCGCAGAGGTCCTCTTCTACTGGACAGATCAGGAGTTTGA
AGAGAGTCTCCGGCTGAAGTTCGGGCAGTCAGAGAATGAGGAAGAAGAGCTCCCTGCCCTGGAGGACCAGCTCAGCA
CCCTCCTAGCCCCGGTCATCATCTCCTCCATGACGATGCTGGAGAAGCTCTCGGACACCTACACCTGCTTCTCCACG
GAAAATGGCAACTTCTGTATGTCCTTCACCTGTTTGGAGAATGCCTGTTTCAATTGCCATCAATGGTGACCACACCGA
GAGCGAGGGGGACCTGCGGCGGAAGCTGTATGTGCTCAAGTACCTGTTTGAAGTGCACCTTTGGGCTGGTGAAGTGTGG
10 ACGGTCATCTTATCCGAAAGGAGCTGCGGCCCCCAGACCTGGCGCAGCGTGTCCAGCTGTGGGAGCACTTCCAGAGC
CTGCTGTGGACCTACAGCCGCTGCGGGAGCAGGAGCAGTGTTCGCCGTGGAGGCCCTGGAGCGACTGATTACCCC
CCAGCTCTGTGAGCTGTGCATAGAGGCGCTGGAGCGGCACGTCAATCCAGGCTGTCAACACCAGCCCCGAGCGGGGAG
GCGAGGAGGCCCTGCATGCCTTCTGCTCGTGCACCTCCAAGCTGCTGGCATTCTACTCTAGCCACAGTGCCAGCTCC
CTGCGCCCCGGCCGACCTGCTTGCCCTCATCCTCCTGGTTTCAGGACCTCTACCCCAGCGAGAGCACAGCAGAGGACGA
15 CATTTCAGGAGACAGACAGCTTCTCCCTCCCTGAGGAGTACTTCACACCAGCTCCTTCCCCCTGGCGATCAGAGCTCAG
GTAGCACCATCTGGCTGGAGGGGGGACCCCCCATGGATGCCCTTCAGATAGCAGAGGACACCCCTCCAAACACTG
GTTCCCCACTGCCCTGTGCCTTCCGGCCCCAGAAGGATCTTCTGGATGCCAACGTGAAGGAAAGCTACTGCCCCCT
AGTGCCCCACACCATGTACTGCCTGCCCTGTGGCAGGGCATCAACCTGGTGTCTCTGACAGGAGCCCCAGCGCGC
CCCTGGCCCTGGTTCTGTCTCCAGCTGATGGATGGCTTCTCCATGCTGGAGAAGAAGCTGAAGGAAGGGCCGGAGCCC
20 GGGGCTCCCTGCGCTGCCCTCCAGCCCCCTCGTGGGAGACCTGCGCCAGAGGATGGACAAGTTTGTCAAGAATCGAGGGGC
ACAGGAGATTTCAGAGCACCTGGCTGGAGTTTAAGGCCAAGGCTTCTCCAAAAGTGAGCCCGGATCCTCCTGGGAGC
TGCTCCAGGCATGTGGGAAGCTGAAGCGGCAGCTCTGCGCCATCTACCGGCTGAACTTTCTGACCACAGCCCCCAGC
AGGGGAGGCCCCACACCTGCCCCAGCACCTGCAGGACCAAGTGCAGAGGCTCATGCGGGAGAAGCTGACGGACTGGAA
GGACTTCTTGCTGGTGAAGAGCAGGAGGAACATCACCATGGTGTCTTACCTAGAAAGACTTCCCAGGCTTGGTGCAC
25 TCATCTATGTGGACCGCACCACTGGGCAGATGGTGGCGCCTTCCCTCAACTGCAGTCAAAAGACCTCGTCGGAGTTG
GGCAAGGGGCGCTGGCTGCCTTTGTCAAACTAAGGTCTGGTCTCTGATCCAGCTGGCGCGCAGATACCTGCAGAA
GGGCTACACCACGCTGCTGTTCCAGGAGGGGATTTCTACTGCTCCTACTTCCCTGTGGTTCGAGAATGACATGGGGT
ACAAACTCCAGATGATCGAGGTGCCCGTCTCTCCGACGACTCAGTGCCTATCGGCATGCTGGGAGGAGACTACTAC
AGGAAGCTCCTGCGCTACTACAGCAAGAACCGCCCAACCGAGGCTGTCTAGGTGCTACGAGCTGCTGGCCCTGCACCT
30 GTCTGTCTATCCCCACTGACCTGCTGGTGCAGCAGGCGCGCCAGCTGGCCCGGCGCCTCTGGGAGGCCTCCCGTATCC
CCCTGCTCTAGGCCAAGGTGGCCGCGAGTCTGCCTTTGCATCCTGTCTCCAGCCACCCCTTGCTTGCCACTGTTCCCC
ATGACGAGAGCCTCCTGTCTGCAGTGGCCATCCTGAGGATAGGGCAGAGTGGCCAGGGTGGCCCCAGGGCTTCTAAA
ACCCACCTAGACCACCCCTCCATGTCAGGTACTGAGCAAGGCCCCAGATCCTTCTCTCTGGAGGAAGAGGGAAGCCC
AGGGGTCTGTTTGTAAACAACGGTGGCAACAGCTCCTCTTCCAGAGCTGCCTCTGCCTTTATCCTGGGAGATGGG
35 GAGGAAGCCCCATCTCTGCTGTTCCCTGCGTGGAGGAAGCCCACCCAGCAAGCTCTCTCCTACCCAGGTAAAAGGT
GCTCCTTTTGGCTGGGTTTGAATTCAGCGCTGCCACTTCTCTCTGACCTCCTGGCAAGTTTCTTCTATTCCCCAC
GTTTAAAGCATGGCACCTCCGTCCAGGGTGGTGTGAGGATTACCCAGTGTGGTAGGTGCTCAATAAATGTTGGTC
ATTGTTATCACTGAAGCCCAACATGCTAGTCTTCTAGACCTTCTGTCTAGTGTGATAAGCCCTTGCTAAGTCCCA
GCCCCCTTCATGCTTGGCTGGCGTCTGCCCTAGGGCTGGGGTTCTCAAGCCCCCTGGCCCTGGCCAGAGATTTGGATT
40 CCCTTGCGGCGCGTGGAGCCCAGCCTTTGATGTCTTTCAAAGCTTCTGTGGTGCGCCCTGGATTGAGAACCACCACC
CGAGGGGTACAGCCCTCTCTTCCAACCGAGAAGTCTGTGCCAGAAATGGACCCAGGGACAAGAGACCCCTGAGAGC
CCTGGGACTGGGAGTGTCTGCTCCTCTGAGGCCAGGAGGCCGGTGTGGGCCAGAGAGGACGGCGTGGCGAAAGTCA
GCGTCCACTGCAGCACAGGATCAGATGGCCGTGTGCTGTGCATGCAGGAGCCTCGCCTTCTGTGTCTTTAGTCTTGA
GCCAAAATTTGCTCAAAGACTGATCTCTTCTTGCAGGGAACAGCTTTGGGGCTGGGGGAACTAGAACCCACATGTT
45 GGTCTAAACCCTGAGAAGGTGGCAGTGAGGAAGTATCCCTCAGGTGACTGGATCTGTGTTCTCTTAAACATCATC
TGATGGAATGGCAATGAAAAGCGTGGATTGTGAAAAATACAGAAAAACATAAAGGAAAAAACTCCAATCCCCTGAGC
CCACCACTGTTTCAGGACCCCTGCTTTTGTACCTACTATTTCCCTTTAGTTTTTAGCAGCGGCTGGATGTGATATGT
CTAGTTTAAACAGTCCCCCTTGATCTTTCTATATAATAAATAACACAGGAGTGAACATCCTGAATCAGA

```

50 An exemplary nucleotide sequence encoding mouse HPS1 protein (NP_001333632.1) is set forth in SEQ ID NO: 14 (NM_001346703.2):

```

ACGTGTCTACCCGGAAGCGCGCCGCGCAGGCTGGCTACACGCTGGGCGGTGCGAGCCGCGGATCCGGTACGTGGGCAT
CGTGGGTGACCAAGCGCAGTCTGGGTGCGCCCTCGGTGGGCACCTCCGGCAAGCTCTCCAGTTTGGAGAGGCCCCG
GGCACCCGGGGAAACCCGCATCAGTCTCATGCCTTTTGGAGCGGGCGTCAACAAGCCAGGCCTTCGTCCCCGGGAGGA

```

- 14 -

GAGGGGGAGGCATTTGCAGGCCCCAAGGATGCTGTGCTGCTGTTGTCTGCTGACCTGGTCAGAGTATCTAGCTCCTCA
 GAAGCCCTGCTGTGTGCTTTGCCAAGGCTGAACTCACGGATGTGACCTTTGCCCTCGCTTTTCCCTGGAAGATGAAG
 TGCCTGTTGGTGGCCACCGAGGGCGCAGAGGTCTCTTCTACTGGACAGATGAGGAGTTTGTCTGAGAGCCTCCGGCT
 5 GAAGCTCCAGCAGTCAGAGGATGAGGAAGAGGAGCTCCCTGTGCTGGAGGACCAGCTCAGCACCCTCCTGGCCCCAG
 TCATCATCTCCTCCATGACGATGATGGAGAAGCTCTCGGACACATATACCTGCTTCTCAACAGAAAATGACAACCAC
 CTGTATGTCTCTGCATCTGTTTGGAGAGTACCTGTTCTGTCGCCATCAACGGAGACCACAGTGAGAGCGAGGGGGACCT
 GCGGAGGAAACTGTGTGTGCTCAAGTATCTATTTCGAGGTGCACTTCGGGCTGGTGACTGTGGACGGCCAGCTCATCC
 GGAAGGAGCTTCGTCCACCCGACCTGGAAGAGCGTGCCCGGTGTGGAAGCACTTTCAGAGACTGCTGGGGACCTAC
 AGCTACCTGCGGGACCGGGAGCAGAGCTTTGCTGTGGAGGAGTGGAGCGGCTCATTACCCCCAGCTCTGCGAACA
 10 GAGCATCGAGACGCTAGAGCGGCACGTGGTTCAAGCCATCAACGCCAGCCCCGAGCGGGCGGTGAAGAGGTCTCTGC
 ACGCCTTCCTGCTGGTGCACGTGCAAGCTGTTGGCTTTCTACTCCGGCCATGGTGCAGACCCCTGCGCCCCGCAGAC
 CTCCTGGCGCTCATCCTTCTGGTGCAGGACCTCCAGCCCTCCCCAGGAACCACAGAGGAGGAGGAGGAGGAGGAGGA
 TAGCGACAGCCCTCAGGTCCCGGAGCACAAGGGGAACAGGAGGAGACCCAAAGAGCAGCCAAAACATCCCAGTGCAGC
 AGGCAAGGAGCCAGAGCACCTCAGTCCCCACCAGGAGTTCAGGGAGACAGACACAGACAGCATCTCCCTCCCTGAA
 15 GAGTACTTCACTCCTGCTCCTTCCCCAGGCGATCAGAGTTCAGGTAGCCTCGTCTGGCTGGATGGTGGCACCCCGCC
 CAGCGATGCTCTTCAGATGGCTGAGGACACCCCGAGGGGCTGGCGTCTCACTCCCCAGAACTTCCCAGCCCCAGAA
 GGATCTTCCCTGGATGCCAACATAAAAGAAAACACTACTGTCCCTTAGTGCCCCACACCATGTACTGCCTGCCCTGTGG
 CCAGGCATCAACATGGTGTGCTGCTGACCAAGAGCCCCAGCACCCTGAGCCCTGATCCTGTACCAACTGCTGGATGG
 CTTCTCCCTCCTGGAGAAAAAGCTGAAGGAAGGTGAGGAGGCCGGAAGTGCCCTGCGATCCCAGCCCTTTGTTGCAG
 20 ACCTGCGCCAGAAGATGGACAAGTTTATCAAGAATCGAGTTGGGCGAGGAGATTCAGAACACCTGGCTGGAGTTTAAG
 AGCAAAGCCTTCTCTAGAAGTGAGCCAGGATCGTCTGGGAGCTGCTCCAGGTCTGTGGAAGCTGAAGCGACAGCT
 GTGTGTCTATCTACCGGCTTAGCTTCTGGTCACTGCACCCAGCAGAGGAGGACCGCACCTGCCCCAGCACCTGCAGG
 ACCGAGCCCAGAAACTCATGAAGGAGAGGCTTCTGGACTGGAAGGACTTCTGTGGTGAAGAGCCGGAGAAATGTC
 ACCATGGTGTCTCTACCTGGAGGATTTCCAGGCCTGGTCCACTTCATCTATGTGGACCGTACAACCGGGCAGATGGT
 25 GGCCCCCTCTCTCAGCCCCAACGAAAAGATGTCTTCCGAGTTGGGCAAGGGGGCCCTGGCTGCCTTTGTCAAAGCCA
 AGGTCTGGGCTCTGCTCCGACTGGCGCGCAGGTACCTGCAGAAGGGCTGCACCACACTGCTGTTCCAGGAAGGGGAC
 TTCCGCTGCTCCTACTTCTGTGGTTCGAGAATGACATGGGATACAACTGCAGATGATTGAGGTGCCTGTCTCTC
 GGATGACTCCGTCCCCATTGGCGTGCTGGGAGGTGACTACTACAGGAAACTTCTGCGCTACTACAGCAAGAGCCACC
 CCTCGGAGCCCGTCAGGTGCTATGAGCTGCTCACGCTGCATCTGTCCGTCTATCCCCACGGACCTGCTGGTGCAGCAG
 30 GCCAGCCAGCTGGCCCCGGCGCTGGGGGAGGCCTCCCGGGTGACCTGCCCTAGGCCGTCTTGCTTGCAAGGCTTCTT
 CCAGTGCCCTTCTCTCATACCCCACTGTGGACTCCTCCACAAAAGGCTGAGGAGCAAGCCTTCTCTGAAGGGCAGGGAT
 GCCATGATTCTGTGGCCCCCAGAACTGTGTGGTCAGAGGTGGCAGCAGCTGCTCTGGGAAGCTGCCCTTCTGGGTGC
 CCGGGAGAGCAGAAGAAGCTCTTAGTGCTCAGGCTGTCTGAAGCCACCCACTTCTGCCATTCTGCATAGGGTAGCT
 GCTCCCTACCCCAATCAGAAGGTACTCCCTCTGGCTGGCGCCATTCCAGCTTGGCTCCTGCTCCCTTGCAACTGAGT
 35 GGATCCAGTTTCTCTACAAATTGAAGTACAAACACCTACCTTTCAGGGTGGTGCAGGAGGATTGCTGGATGTGCTCGG
 GTGCTCAATAAATGTTGATTCTTGTCAATTATTCAAG

An exemplary nucleotide sequence encoding mouse HPS1 protein (NP_062297.1) is set forth in SEQ ID NO: 15 (NM_019424.3):

40 ACGTGTCTTACCCGGAAGCGCGCCGAGGCTGGCTACACGCTGGGCGGTGCGAGCCGCGGATCCGCATTTGCAGGGC
 CAAGGATGCTGTGCTGCTGTTGTCTGCTGACCTGGTCAGAGTATCTAGCTCCTCAGAAGCCCTGCTGTGTGCTTTGC
 CAAGGCTGAACTCACGGATGTGACCTTTGCCCTCGCTTTTCCCTGGAAGATGAAGTGCCTGTTGGTGGCCACCGAGG
 GCGCAGAGGTCTCTTCTACTGGACAGATGAGGAGTTTGTCTGAGAGCCTCCGGCTGAAGCTCCAGCAGTCAGAGGAT
 GAGGAAGAGGAGCTCCCTGTGCTGGAGGACCAGCTCAGCACCCTCCTGGCCCCAGTCATCATCTCCTCCATGACGAT
 45 GATGGAGAAGCTCTCGGACACATATACCTGCTTCTCAACAGAAAATGACAACCACCTGTATGTCTCTGCATCTGTTT
 GAGAGTACCTGTTCTGTCGCCATCAACGGAGACCACAGTGAGAGCGAGGGGGACCTGCGGAGGAAACTGTGTGTGCTC
 AAGTATCTATTTCGAGGTGCACTTCGGGCTGGTGACTGTGGACGGCCAGCTCATCCGGAAGGAGCTTCGTCCACCCGA
 CCTGGAAGAGCGTGCCCGGTGTGGAAGCACTTTCAGAGACTGCTGGGGACCTACAGCTACCTGCGGGACCGGGAGC
 AGAGCTTTGCTGTGGAGGCAGTGGAGCGGCTCATTACCCCCAGCTCTGCGAACAGAGCATCGAGACGCTAGAGCGG
 50 CAGCTGGTTCAAGCCATCAACGCCAGCCCCGAGCGGGGCGGTGAAGAGGTCTGACGCTTCTGCTGGTGTGCTG
 CAAGCTGTTGGCTTTCTACTCCGGCCATGGTGCAGACCCCTGCGCCCCGAGACCTCCTGGCGCTCATCCTTCTGG
 TGCAGGACCTCCAGCCCTCCCCAGGAACCACAGAGGAGGAGGAGGAGGAGGATAGCGACAGCCCTCAGAGGAGA
 CCAAAGAGCAGCCAAAACATCCCAGTGCAGCAGGCAAGGAGCCAGAGCACCTCAGTCCCCACCAGGAGTTCCAGGGA
 GACAGACACAGACAGCATCTCCCTCCCTGAAGAGTACTTCACTCCTGCTCCTTCCCCAGGCGATCAGAGTTCAGGTA
 55 GCCTCGTCTGGCTGGATGGTGGCACCCCGCCAGCGATGCTCTTCAGATGGCTGAGGACACCCCGAGGGGCTGGCG
 TCTCACTCCCCAGAACTTCCCAGCCCCAGAAGGATCTTCTGGATGCCAACATAAAAGAAAACACTACTGTCCCTTAGT

- 15 -

GCCCCACACCATGTACTGCCTGCCCCCTGTGGCCAGGCATCAACATGGTGCTGCTGACCAAGAGCCCCAGCACCCCCAC
TGGCCCTGATCCTGTACCAACTGCTGGATGGCTTCTCCCTCCTGGAGAAAAAGCTGAAGGAAGGTCAGGAGGCCGGA
AGTGCCCTGCGATCCCAGCCCTTTGTTGCAGACCTGCGCCAGAAGATGGACAAGTTTATCAAGAATCGAGTTGGGCA
GGAGATTGAGAACACCTGGCTGGAGTTTAAAGAGCAAAGCCTTCTCTAGAAGTGAGCCAGGATCGTCTGGGAGCTGC
5 TCCAGGTCTGTGGAAAGCTGAAGCGACAGCTGTGTGTCATCTACCGGCTTAGCTTCTGGTCACTGCACCCAGCAGA
GGAGGACCGCACCTGCCCCAGCACCTGCAGGACCGAGCCCAGAACTCATGAAGGAGAGGCTTCTGGACTGGAAGGA
CTTCTGTGTTGGTGAAGAGCCGGAGAAATGTCACCATGGTGTCTTACCTGGAGGATTTCCAGGCCCTGGTCCACTTCA
TCTATGTGGACCGTACAACCGGGCAGATGGTGGCCCCCTCTCTCAGCCCCAACGAAAAGATGTCTTCCGAGTTGGGC
AAGGGGCCCCCTGGCTGCCCTTTGTCAAAGCCAAGGTCTGGGCTCTGGTCCGACTGGCGCGCAGGTACCTGCAGAAGGG
10 CTGCACCACACTGCTGTGTTCCAGGAAGGGGACTTCCGCTGCTCCTACTTCTCTGTGGTTCGAGAATGACATGGGATACA
AACTGCAGATGATTGAGGTGCCTGTCTCGGATGACTCCGTCCCCATTGGCGTGCTGGGAGGTGACTACTACAGG
AAACTTCTGCGCTACTACAGCAAGAGCCACCCCTCGGAGCCCCGTCAGGTGCTATGAGCTGCTCAGCTGCATCTGTC
CGTCATCCCCACGGACCTGCTGGTGCAGCAGGCCAGCCAGCTGGCCCCGGCGCTGGGGGAGGCCCTCCCGGTGACCC
TGCCCTAGGCCGTCTTGCTTGCAGGCTTCTTCCAGTGCCCTTCTCTCATACCCCACTGTGGACTCCTCCACAAAGGCT
15 GAGGAGCAAGCCTTCTCTGAAGGGCAGGGATGCCATGATTCTGTGGCCCCCAGAACTGTGTGGTGCAGAGGTGGCAGC
AGCTGCTCTGGGAAGCTGCCCTTCTGGGTGCCCGGGAGAGCAGAAGAACGCTCTTAGTGCTCAGGCTGTCTGAAGCCA
CCCACCTTCTGCCATTCTGCATAGGGTAGCTGCTCCCTACCCCAATCAGAAGGTACTCCCTCTGGCTGGCGCCATTCT
CAGCTTGGCTCCTGCTCCCTTGCAACTGAGTGATCCAGTTTCTCTACAAATTGAAGTACAAACACCTACCTTTTCAG
GGTGGTCAAGGAGATTGCTGGATGTGCTCGGGTGCTCAATAAATGTTGATTCTTGTATTATTCAAG

An exemplary nucleotide sequence encoding mouse HPS1 protein (NP_001349339.1) is
set forth in SEQ ID NO: 16 (NM_001362410.1):

CATTTGCAGGCCCCAAGGATGCTGTGCTGCTGTTGTCTGCTGACCTGGTCAAGTATCTAGCTCCTCAGAAGCCCTGC
TGTGTGCTTTGCCAAGGCTGAACCTACGGATGTGACCTTTGCCCTCGCTTTTCCCTGGAAGATGAAGTGCCTGTTGG
25 TGGCCACCGAGGGCGCAGAGGTCTCTTCTACTGGACAGATGAGGAGTTTGCTGAGAGCCTCCGGCTGAAGCTCCAG
CAGTCAGAGGATGAGGAAGAGGAGCTCCCTGTGCTGGAGGACCAGCTCAGCACCCCTCCTGGCCCCAGTCATCATCTC
CTCCATGACGATGATGGAGAAGCTCTCGGACACATATACCTGCTTCTCAACAGAAAATGACAACCACCTGTATGTCC
TGCATCTGTTTGGAGAGTACCTGTTTCGTGCGCCATCAACGGAGACCACAGTGAGAGCGAGGGGGACCTGCGGAGGAAA
CTGTGTGTGCTCAAGTATCTATTTCGAGGTGCACTTCGGGCTGGTGACTGTGGACGGCCAGCTCATCCGGAAGGAGCT
30 TCGTCCACCCGACCTGGAAGAGCGTGCCCGGTGTGGAAGCACTTTCAGAGACTGCTGGGGACCTACAGCTACCTGC
GGGACCGGGAGCAGAGCTTTGCTGTGGAGGCAGTGAGCGGCTCATTACCCCCAGCTCTGCGAACAGAGCATCGAG
ACGCTAGAGCGGCACGTGGTTCAAGCCATCAACGCCAGCCCCGAGCGGGGCGGTGAAGAGGTCTGCACGCCCTTCCT
GCTGGTGACTGCAAGCTGTTGGCTTTCTACTCCGGCCATGGTGCGAGCACCTGCGCCCCGAGACCTTCTGGCGC
TCATCCTTCTGGTGCAGGACCTCCAGCCCTCCCCAGGAACACAGAGGAGGAGGAGGAGGATAGCGACAGC
35 CCTCAGAGGAGACCCAAGAGCAGCCAAAACATCCAGTGACAGCAGGCAAGGAGCCAGAGCACCTCAGTCCCCACCAG
GAGTTCCAGGGAGACAGACACAGACAGCATCTCCCTCCCTGAAGAGTACTTCACTCCTGCTCCTTCCCCAGGCGATC
AGAGTTCAGGTAGCCTCGTCTGGCTGGATGGTGGCACCCCCGCCAGCGATGCTCTTCAGATGGCTGAGGACACCCCC
GAGGGGCTGGCGTCTCACTCCCCAGAACTTCCAGCCCCAGAAGGATCTTCTGGATGCCAACATAAAAGAAAATA
CTGTCCCTTAGTGCCCCACACCATGTACTGCCTGCCCTGTGGCCAGGCATCAACATGGTGTGCTGACCAAGAGCC
40 CCAGCACCCCCACTGGCCCTGATCCTGTACCAACTGCTGGATGGCTTCTCCCTCCTGGAGAAAAAGCTGAAGGAAGGT
CAGGAGGCCGGAAGTGCCCTGCGATCCCAGCCCTTTGTTGCAGACCTGCGCCAGAAGATGGACAAGTTTATCAAGAA
TCGAGTTGGGCAGGAGATTGAGAACACCTGGCTGGAGTTTAAAGAGCAAAGCCTTCTCTAGAAGTGAGCCAGGATCGT
CCTGGGAGCTGCTCCAGGTCTGTGGAAAGCTGAAGCGACAGCTGTGTGTCATCTACCGGCTTAGCTTCTGGTCACT
GCACCCAGCAGAGGAGGACCGCACCTGCCCCAGCACCTGCAGGACCGAGCCCAGAACTCATGAAGGAGAGGCTTCT
45 GGACTGGAAGGACTTCTGTTGGTGAAGAGCCGGAGAAAATGTCACCATGGTGTCTTACCTGGAGGATTTCCAGGCC
TGGTCCACTTCATCTATGTGGACCGTACAACCGGGCAGATGGTGGCCCCCTCTCTCAGCCCCAACGAAAAGATGTCT
TCCGAGTTGGGCAAGGGGCCCCCTGGCTGCCCTTTGTCAAAGCCAAGGTATGGAGCCCAAGGACTCGTGTACTGAGACA
GCACTCTGCCCTCTCTCTGAGCTGCTCCACACCCCCAGCCGAGTGACAGGGGTGGTTTCCACTCAGGAAGCATCGT
TCTCTGCTCTTCTGATTTAATGACGCTTCTTGACAGGCTGGTAGTTTTTAAAGGAAAGACTCCGCCATTTCTAAAA
50 GTTTTATTTTTATTTTATCTATATAAAATGTATCAGCATAGATAAGCTAGGTGGTGAGAGCCAGAGCTGGAAGCC
TGGTCAGTTCTGGTTGACTGTGTGGCCCTGACGTGCTTTGGTTATATTTTCAAGTTCTGGTTTCAAAATATCCATG
GCAGAGATAATGCCGTTGTAAGGGTATAGGTCATAAACACAGATGAATTCACAGAACAGGGGCAGAGCTGAGCCA
AGCATGCCTCACACAGCGCGGCCAGCTTCTCACCTACCACATTACCCCTGCATCCTGCTCTGGGGCAGAGGGGACT
CTAGGCCATTGCCCTTCATTGTCTTCATCTCAGTGCTGCCCACTCTGCATGCACAGGCCTGTGTGACAGATTCTCAT
55 GTTGTGTGAGGTGCCTCTCTTCTGCTGTTAACTCTTTCATGCCTCGCCTGTAGCAACCTATGATGCAGGCGTAG
CAGAATTCACACAAGCCGTTCCCCGCTGAGTCCCTGGTTCAGATGGGAGGGCTGTTATTGAAGTCATCTATTTCTGTG

- 16 -

GTCAGGGGTACATTCCCTGGAGCTGGCAACTGATTCCCGTGTCTATTGCAGATATAATTAGATAGTATATTGTGTCAAG
 TGCAAAGAAAAGACCACATCCCTTCACTGAGGCTTCTGACAGAGTGGGTTATTTCCCAGCCCTGGCACACAGTAGGC
 AAGCCTAGTAGAGCAGATCTTCCCTCAGGCTGCTCTTAGGCAGCCTAGGTGCCAACACTCGCTTTTGGCAGAACATT
 CCCAGGGCGCGAATGTCTTCAGAGCCTCTGTTTCCAGCCACCATATAGTGAGCCTACGTTTGGGCCTGACACCTAA
 5 GCAGGTGATGTCAGAGCTCAGGGCTCACTGTACATTTAGGTACCCCCAAGGCCCTCCAGCACAGAGCCTGCACCAATT
 CTTCCCAGCCCTGCAGGCCTGGGCAGTTGGATTTATTACAGTCTTCTTGGCTGTTTTTGGTCTCAACGAGCAGTTT
 ACTAATCCATTGGGTTGTGTGCAGTAAAGTTACATGAAGC

In some embodiments, the HPS1 is any one of the HPS1 described herein. In some
 10 embodiments, the HPS1 comprises an amino acid sequence at least 70%, at least 75%, at least
 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least
 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to any one of SEQ
 ID NOs: 1-8. In some embodiments, the expression cassette comprises a nucleotide sequence
 encoding a HPS1 comprising an amino acid sequence at least 70%, at least 75%, at least 80%, at
 15 least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at
 least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to any one of SEQ ID
 NOs: 1-8. In some embodiments, a HPS1 is encoded by a nucleotide sequence at least 70%, at
 least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at
 least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical
 20 to any one of SEQ ID NOs: 9-16.

In some aspects, the disclosure relates to isolated nucleic acids comprising or lacking
 certain regulatory sequences. In some embodiments, isolated nucleic acids and rAAVs
 described herein comprise (or lack) one or more of the following structural features (*e.g.*, control
 or regulatory sequences): a 5' untranslated region (5'UTR), a promoter, an intron, a Kozak
 25 sequence, one or more miRNA binding sites, a SV40 poly A sequence, and a 3' untranslated
 region (3'UTR). In some embodiments, one or more of the foregoing control sequences is
 operably linked to a nucleic acid sequence encoding the HPS1 protein.

In some embodiments, the expression cassette of the isolated nucleic acid further
 comprises a promoter operably linked to nucleotide sequence encoding the HPS1 protein. As
 30 used herein, a nucleic acid sequence (*e.g.*, coding sequence) and regulatory sequences are said to
 be "operably linked" when they are covalently linked in such a way as to place the expression or
 transcription of the nucleic acid sequence under the influence or control of the regulatory
 sequences. If it is desired that the nucleic acid sequences be translated into a functional protein,
 two DNA sequences are said to be operably linked if induction of a promoter in the 5' regulatory
 35 sequences results in the transcription of the coding sequence and if the nature of the linkage

- 17 -

between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably linked to a nucleic acid sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide. Similarly two or more coding regions are operably linked when they are linked in such a way that their transcription from a common promoter results in the expression of two or more proteins having been translated in frame.

A "promoter" refers to a DNA sequence recognized by the synthetic machinery of the cell, or introduced synthetic machinery, required to initiate the specific transcription of a gene. The phrases "operatively linked," "operatively positioned," "under control" or "under transcriptional control" means that the promoter is in the correct location and orientation in relation to the nucleic acid to control RNA polymerase initiation and expression of the gene. In some embodiments, a transgene comprises a nucleic acid sequence encoding a HPS1 protein operably linked to a first promoter and a multi-gRNA expression cassette operably linked to a second promoter.

Generally, a promoter can be a constitutive promoter, inducible promoter, or a tissue-specific promoter.

Examples of constitutive promoters include, without limitation, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer) [see, *e.g.*, Boshart et al., Cell, 41:521-530 (1985)], the SV40 promoter, the dihydrofolate reductase promoter, the β -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1 α promoter [Invitrogen]. In some embodiments, the promoter is a chicken beta-actin (CB) promoter or a small synthetic promoter. In some embodiments, the promoter is a hybrid promoter. In some embodiments, the promoter is a cytomegalovirus immediate – early enhancer/chicken β -actin hybrid promoter. In some embodiments, a promoter is an RNA pol II promoter. In some embodiments, a promoter is an RNA pol III promoter, such as U6 or H1. In some embodiments, a promoter is an RNA pol II promoter. In some embodiments, a nucleic acid encoding a RGN is operably linked to a CB6 promoter. In some embodiments, a nucleic acid sequence encoding a multi-RNA expression

- 18 -

cassette is operably linked to a RNA pol III promoter. In some embodiments, the RNA pol III promoter is a U6 promoter. In some embodiments, the promoter is a U1a promoter.

Examples of inducible promoters regulated by exogenously supplied promoters include the zinc-inducible sheep metallothionein (MT) promoter, the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter, the T7 polymerase promoter system (WO 98/10088); the ecdysone insect promoter (No et al., Proc. Natl. Acad. Sci. USA, 93:3346-3351 (1996)), the tetracycline-repressible system (Gossen et al., Proc. Natl. Acad. Sci. USA, 89:5547-5551 (1992)), the tetracycline-inducible system (Gossen et al., Science, 268:1766-1769 (1995), see also Harvey et al., Curr. Opin. Chem. Biol., 2:512-518 (1998)), the RU486-inducible system (Wang et al., Nat. Biotech., 15:239-243 (1997) and Wang et al., Gene Ther., 4:432-441 (1997)) and the rapamycin-inducible system (Magari et al., J. Clin. Invest., 100:2865-2872 (1997)). Still other types of inducible promoters which may be useful in this context are those which are regulated by a specific physiological state, *e.g.*, temperature, acute phase, a particular differentiation state of the cell, or in replicating cells only.

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

In some embodiments, the regulatory sequences impart tissue-specific gene expression capabilities. In some cases, the tissue-specific regulatory sequences bind tissue-specific transcription factors that induce transcription in a tissue specific manner. Such tissue-specific regulatory sequences (*e.g.*, promoters, enhancers, etc.) are well known in the art. Exemplary tissue-specific regulatory sequences include, but are not limited to the following tissue specific promoters: retinoschisin proximal promoter, interphotoreceptor retinoid-binding protein enhancer (RS/IRBPa), rhodopsin kinase (RK), liver-specific thyroxin binding globulin (TBG) promoter, an insulin promoter, a glucagon promoter, a somatostatin promoter, a pancreatic polypeptide (PPY) promoter, a synapsin-1 (Syn) promoter, a creatine kinase (MCK) promoter, a mammalian desmin (DES) promoter, a α -myosin heavy chain (α -MHC) promoter, or a cardiac Troponin T (cTnT) promoter. Other exemplary promoters include Beta-actin promoter, hepatitis

- 19 -

B virus core promoter, Sandig et al., Gene Ther., 3:1002-9 (1996); alpha-fetoprotein (AFP) promoter, Arbuthnot et al., Hum. Gene Ther., 7:1503-14 (1996)), bone osteocalcin promoter (Stein et al., Mol. Biol. Rep., 24:185-96 (1997)); bone sialoprotein promoter (Chen et al., J. Bone Miner. Res., 11:654-64 (1996)), CD2 promoter (Hansal et al., J. Immunol., 161:1063-8 (1998); immunoglobulin heavy chain promoter; T cell receptor α -chain promoter, neuronal such as neuron-specific enolase (NSE) promoter (Andersen et al., Cell. Mol. Neurobiol., 13:503-15 (1993)), neurofilament light-chain gene promoter (Piccioli et al., Proc. Natl. Acad. Sci. USA, 88:5611-5 (1991)), and the neuron-specific vgf gene promoter (Piccioli et al., Neuron, 15:373-84 (1995)), among others which will be apparent to the skilled artisan. In some embodiments, the tissue-specific promoter is a lung cell specific promoter. In some embodiments, the tissue-specific promoter is a lung epithelial cell promoter. In some embodiments, the tissue-specific promoter is an alveolar type II cell promoter. Non-limiting examples of alveolar type II cell specific promoter include a surfactant protein C (SP-C) promoter, a Surfactant Protein B (SP-B) promoter, or a FoxM1 promoter.

In some aspects, the disclosure relates to isolated nucleic acids comprising an expression cassette that comprises one or more miRNA binding sites. Without wishing to be bound by any particular theory, incorporation of miRNA binding sites into gene expression constructs allows for regulation of transgene expression (*e.g.*, inhibition of transgene expression) in cells and tissues where the corresponding miRNA is expressed. In some embodiments, incorporation of one or more miRNA binding sites into a transgene allows for de-targeting of transgene expression in a cell-type specific manner. In some embodiments, one or more miRNA binding sites are positioned in a 3' untranslated region (3' UTR) of a transgene, for example between the last codon of a nucleic acid sequence encoding a HPS1 protein, and a poly A sequence.

In some embodiments, a transgene comprises one or more (*e.g.*, 1, 2, 3, 4, 5, or more) miRNA binding sites that de-target expression of the HPS1 protein from central nervous system (CNS) cells. In some embodiments, a transgene comprises one or more (*e.g.*, 1, 2, 3, 4, 5, or more) miRNA binding sites that de-target expression of the HPS1 from immune cells. In some embodiments, an expression cassette comprises one or more (*e.g.*, 1, 2, 3, 4, 5, or more) miRNA binding sites that de-target expression of the HPS1 protein from immune cells (*e.g.*, antigen presenting cells (APCs), such as macrophages, dendrites, *etc.*). Incorporation of miRNA binding sites for immune-associated miRNAs may de-target transgene expression from antigen presenting cells and thus reduce or eliminate immune responses (cellular and/or humoral)

- 20 -

produced in the subject against products of the transgene, for example as described in US 2018/0066279, the entire contents of which are incorporated herein by reference.

As used herein an “immune-associated miRNA” is an miRNA preferentially expressed in a cell of the immune system, such as an antigen presenting cell (APC). In some embodiments, an immune-associated miRNA is an miRNA expressed in immune cells that exhibits at least a 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold higher level of expression in an immune cell compared with a non-immune cell (*e.g.*, a control cell, such as a HeLa cell, HEK293 cell, mesenchymal cell, *etc.*). In some embodiments, the cell of the immune system (immune cell) in which the immune-associated miRNA is expressed is a B cell, T cell, Killer T cell, Helper T cell, $\gamma\delta$ T cell, dendritic cell, macrophage, monocyte, vascular endothelial cell, or other immune cell. In some embodiments, the cell of the immune system is a B cell expressing one or more of the following markers: B220, BLAST-2 (EBVCS), Bu-1, CD19, CD20 (L26), CD22, CD24, CD27, CD57, CD72, CD79a, CD79b, CD86, chB6, D8/17, FMC7, L26, M17, MUM-1, Pax-5 (BSAP), and PC47H. In some embodiments, the cell of the immune system is a T cell expressing one or more of the following markers: ART2, CD1a, CD1d, CD11b (Mac-1), CD134 (OX40), CD150, CD2, CD25 (interleukin 2 receptor alpha), CD3, CD38, CD4, CD45RO, CD5, CD7, CD72, CD8, CRTAM, FOXP3, FT2, GPCA, HLA-DR, HML-1, HT23A, Leu-22, Ly-2, Ly-m22, MICG, MRC OX 8, MRC OX-22, OX40, PD-1 (Programmed death-1), RT6, TCR (T cell receptor), Thy-1 (CD90), and TSA-2 (Thymic shared Ag-2). In some embodiments, the immune-associated miRNA is selected from: miR-15a, miR-16-1, miR-17, miR-18a, miR-19a, miR-19b-1, miR-20a, miR-21, miR-29a/b/c, miR-30b, miR-31, miR-34a, miR-92a-1, miR-106a, miR-125a/b, miR-142-3p, miR-146a, miR-150, miR-155, miR-181a, miR-223 and miR-424, miR-221, miR-222, let-7i, miR-148, and miR-152.

In some embodiments, an expression cassette comprises one or more (*e.g.*, 1, 2, 3, 4, 5, or more) miRNA binding sites that de-target expression of the HPS1 proteins from liver cells. For example, in some embodiments, an expression cassette comprises one or more miR-122 binding sites.

An isolated nucleic acid described by the disclosure may comprise an expression cassette that further comprises a polyadenylation (poly A) sequence. In some embodiments, an expression cassette comprises a poly A sequence that is SV40 poly A sequence, a rabbit beta-globin (RBG) poly A sequence, or a bovine growth hormone poly A sequence. In some embodiments, an expression cassette comprises a SV40 poly A sequence.

- 21 -

The isolated nucleic acids of the disclosure may be recombinant adeno-associated virus (AAV) vectors (rAAV vectors). In some embodiments, an isolated nucleic acid as described by the disclosure comprises adeno-associated virus (AAV) inverted terminal repeats (ITRs), or a variant thereof. The isolated nucleic acid (*e.g.*, the recombinant AAV vector) may be packaged
5 into a capsid protein and administered to a subject and/or delivered to a selected target cell. "Recombinant AAV (rAAV) vectors" are typically composed of, at a minimum, an expression cassette and its regulatory sequences, and 5' and 3' AAV inverted terminal repeats (ITRs). The isolated nucleic acid may comprise a region encoding, for example, a protein and/or an expression control sequence (*e.g.*, a poly-A tail), as described elsewhere in the disclosure.

10 Generally, ITR sequences are about 145 bp in length. Preferably, substantially the entire sequences encoding the ITRs are used in the molecule, although some degree of minor modification of these sequences is permissible. The ability to modify these ITR sequences is within the skill of the art. (See, *e.g.*, texts such as Sambrook et al., "Molecular Cloning. A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory, New York (1989); and K. Fisher et al., J Virol., 70:520 532 (1996)). An example of such a molecule employed in the disclosure is a
15 "cis-acting" plasmid containing the transgene, in which the selected transgene sequence and associated regulatory elements are flanked by the 5' and 3' AAV ITR sequences. The AAV ITR sequences may be obtained from any known AAV, including presently identified mammalian AAV types. In some embodiments, the isolated nucleic acid further comprises a region (*e.g.*, a
20 second region, a third region, a fourth region, *etc.*) comprising a second AAV ITR. In some embodiments, an isolated nucleic acid encoding a transgene is flanked by AAV ITRs (*e.g.*, in the orientation 5'-ITR-transgene-ITR-3'). In some embodiments, the AAV ITRs are AAV2 ITRs. In some embodiments, at least one of the AAV ITRs is a Δ ITR, which lacks a terminal resolution site and induces formation of a self-complementary AAV (scAAV) vector. In some
25 embodiments, the AAV ITRs are selected from the group consisting of AAV1 ITR, AAV2 ITR, AAV3 ITR, AAV4 ITR, AAV5 ITR, and AAV6 ITR. In some embodiments, the AAV ITRs are AAV2 ITRs or a variant thereof. In some embodiments, the AAV2 ITR is a mutated AAV2 ITR. Without wishing to be bound by any particular theory, a rAAV vector comprising an ITR lacking a functional TRS produces a self-complementary rAAV vector, for example as described
30 by McCarthy (2008) Molecular Therapy 16(10):1648-1656. In some embodiments, vectors described herein comprise one or more AAV ITRs, and at least one ITR is an ITR variant of a known AAV serotype ITR. In some embodiments, the AAV ITR variant is a synthetic AAV ITR

- 22 -

(e.g., AAV ITRs that do not occur naturally). In some embodiments, the AAV ITR variant is a hybrid ITR (e.g., a hybrid ITR comprises sequences derived from ITRs of two or more different AAV serotypes).

In some embodiments, an isolated nucleic acid (e.g., a rAAV vector) as described herein comprises, in 5' to 3' order: (a) a 5' AAV2 mutant ITR; (b) a cytomegalovirus immediate – early enhancer/chicken b-actin hybrid promoter; (c) a nucleotide sequence encoding an Hermansky–Pudlak syndrome 1 (HPS1) protein; (d) a SV40 polyA signal; and (e) a 3' AAV2 mutant ITR.

As used herein, the term "vector" includes any genetic element, such as a plasmid, phage, transposon, cosmid, chromosome, artificial chromosome, virus, virion, etc., which is capable of replication when associated with the proper control elements and which can transfer gene sequences between cells. In some embodiments, a vector is a viral vector, such as an rAAV vector, a lentiviral vector, an adenoviral vector, a retroviral vector, an anellovirus vector (e.g., Anellovirus vector as described in US20200188456A1), etc. Thus, the term includes cloning and expression vehicles, as well as viral vectors. In some embodiments, useful vectors are contemplated to be those vectors in which the nucleic acid segment to be transcribed is positioned under the transcriptional control of a promoter. In some embodiments, HPS1 protein can be delivered to a subject via a non-viral platform. In some embodiments, the HPS1 protein can be delivered to a subject via closed-ended linear duplex DNA (ceDNA). Delivery of a transgene (e.g., HPS1 protein) has been described previously, see e.g., WO2017152149, the entire contents of which are incorporated herein by reference. In some embodiments, the nucleic acids having asymmetric terminal sequences (e.g., asymmetric interrupted self-complementary sequences) form closed-ended linear duplex DNA structures (e.g., ceDNA) that, in some embodiments, exhibit reduced immunogenicity compared to currently available gene delivery vectors. In some embodiments, ceDNA behaves as linear duplex DNA under native conditions and transforms into single-stranded circular DNA under denaturing conditions. Without wishing to be bound by any particular theory, ceDNA are useful, in some embodiments, for the delivery of a transgene (e.g., HPS1 protein) to a subject.

Recombinant adeno-associated viruses (rAAVs)

Aspects of the disclosure relate to isolated adeno-associated viruses (AAVs) comprising an isolated nucleic acid comprising an expression cassette comprising a nucleotide sequence

- 23 -

encoding a Hermansky–Pudlak syndrome 1 (HPS1) protein. As used herein with respect to AAVs, the term “isolated” refers to an AAV that has been artificially produced or obtained. Isolated AAVs may be produced using recombinant methods. Such AAVs are referred to herein as “recombinant AAVs”. Recombinant AAVs (rAAVs) preferably have tissue-specific targeting capabilities, such that a transgene of the rAAV will be delivered specifically to one or more predetermined tissue(s) (*e.g.*, Alveolar type II cells). The AAV capsid is an important element in determining these tissue-specific targeting capabilities (*e.g.*, tissue tropism). Thus, an rAAV having a capsid appropriate for the tissue being targeted can be selected.

Methods for obtaining recombinant AAVs having a desired capsid protein are well known in the art. (See, for example, US 2003/0138772), the contents of which are incorporated herein by reference in their entirety). Typically the methods involve culturing a host cell which contains a nucleic acid sequence encoding an AAV capsid protein; a functional rep gene; a recombinant AAV vector composed of AAV inverted terminal repeats (ITRs) and a transgene; and sufficient helper functions to permit packaging of the recombinant AAV vector into the AAV capsid proteins. In some embodiments, capsid proteins are structural proteins encoded by the cap gene of an AAV. AAVs comprise three capsid proteins, virion proteins 1 to 3 (named VP1, VP2 and VP3), all of which are transcribed from a single cap gene via alternative splicing. In some embodiments, the molecular weights of VP1, VP2 and VP3 are respectively about 87 kDa, about 72 kDa and about 62 kDa. In some embodiments, upon translation, capsid proteins form a spherical 60-mer protein shell around the viral genome. In some embodiments, the functions of the capsid proteins are to protect the viral genome, deliver the genome and interact with the host. In some aspects, capsid proteins deliver the viral genome to a host in a tissue specific manner.

In some embodiments, an AAV capsid protein has a tropism for lung tissue (*e.g.*, ATII cells, *etc.*).

In some embodiments, an AAV capsid protein is of an AAV serotype selected from the group consisting of AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV9.hr, AAVrh8, AAVrh10, AAVrh39, AAVrh43, AAV.PHP.B, AAV.PHP.eB, and variants of any of the foregoing. In some embodiments, an AAV capsid protein is of a serotype derived from a non-human primate, for example AAVrh8 serotype. In some embodiments, the AAV capsid protein is an AAV8 capsid protein. In some embodiments, the AAV capsid is AAV5

- 24 -

capsid, or a variant thereof. In some embodiments, the AAV capsid is AAV8 capsid, or a variant thereof.

In some embodiments, an rAAV vector or rAAV particle comprises a mutant ITR that lacks a functional terminal resolution site (TRS). The term “lacking a terminal resolution site” can refer to an AAV ITR that comprises a mutation (*e.g.*, a sense mutation such as a non-synonymous mutation, or missense mutation) that abrogates the function of the terminal resolution site (TRS) of the ITR, or to a truncated AAV ITR that lacks a nucleic acid sequence encoding a functional TRS (*e.g.*, a Δ TRS ITR). Without wishing to be bound by any particular theory, a rAAV vector comprising an ITR lacking a functional TRS produces a self-complementary rAAV vector, for example as described by McCarthy (2008) *Molecular Therapy* 16(10):1648-1656.

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

In some embodiments, the disclosure relates to a host cell containing a nucleic acid that comprises an expression cassette comprising a nucleotide sequence encoding a Hermansky–Pudlak syndrome 1 (HPS1) protein. A “host cell” refers to any cell that harbors, or is capable of harboring, a substance of interest. Often a host cell is a mammalian cell. In some embodiments, a host cell is a neuron. In some embodiments, a host cell is a photoreceptor cell. A host cell may be used as a recipient of an AAV helper construct, an AAV minigene plasmid, an accessory

- 25 -

function vector, or other transfer DNA associated with the production of recombinant AAVs. The term includes the progeny of the original cell which has been transfected. Thus, a "host cell" as used herein may refer to a cell which has been transfected with an exogenous DNA sequence. It is understood that the progeny of a single parental cell may not necessarily be
5 completely identical in morphology or in genomic or total DNA complement as the original parent, due to natural, accidental, or deliberate mutation. In some embodiments, the host cell is a mammalian cell, a yeast cell, a bacterial cell, an insect cell, a plant cell, or a fungal cell. In some embodiments, the host cell is a hepatocyte.

The recombinant AAV vector, *rep* sequences, *cap* sequences, and helper functions
10 required for producing the rAAV of the disclosure may be delivered to the packaging host cell using any appropriate genetic element (vector). The selected genetic element may be delivered by any suitable method, including those described herein. The methods used to construct any embodiment of this disclosure are known to those with skill in nucleic acid manipulation and include genetic engineering, recombinant engineering, and synthetic techniques. See, *e.g.*,
15 Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. Similarly, methods of generating rAAV virions are well known and the selection of a suitable method is not a limitation on the disclosure. See, *e.g.*, K. Fisher et al., *J. Virol.*, 70:520-532 (1993) and U.S. Pat. No. 5,478,745.

In some embodiments, recombinant AAVs may be produced using the triple transfection
20 method (described in detail in U.S. Pat. No. 6,001,650). Typically, the recombinant AAVs are produced by transfecting a host cell with an AAV vector (comprising a transgene flanked by ITR elements) to be packaged into AAV particles, an AAV helper function vector, and an accessory function vector. An AAV helper function vector encodes the "AAV helper function" sequences (*e.g.*, *rep* and *cap*), which function in trans for productive AAV replication and
25 encapsidation. Preferably, the AAV helper function vector supports efficient AAV vector production without generating any detectable wild-type AAV virions (*e.g.*, AAV virions containing functional *rep* and *cap* genes). Non-limiting examples of vectors suitable for use with the disclosure include pHLP19, described in U.S. Pat. No. 6,001,650 and pRep6cap6 vector, described in U.S. Pat. No. 6,156,303, the entirety of both incorporated by reference
30 herein. The accessory function vector encodes nucleotide sequences for non-AAV derived viral and/or cellular functions upon which AAV is dependent for replication (*e.g.*, "accessory functions"). The accessory functions include those functions required for AAV replication,

- 26 -

including, without limitation, those moieties involved in activation of AAV gene transcription, stage specific AAV mRNA splicing, AAV DNA replication, synthesis of cap expression products, and AAV capsid assembly. Viral-based accessory functions can be derived from any of the known helper viruses such as adenovirus, herpes virus (other than herpes simplex virus type-1), and vaccinia virus.

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

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

Methods

Methods for delivering a transgene (*e.g.*, an isolated nucleic acid described herein) to a subject are provided by the disclosure. The methods typically involve administering to a subject an effective amount of an isolated nucleic acid encoding the transgene(s). In some embodiments, expression constructs described by the disclosure are useful for treating Hermansky–Pudlak syndrome-associated pulmonary fibrosis in a subject. Pulmonary fibrosis is a lung disease that occurs when lung tissue becomes damaged and scarred. This thickened, stiff tissue makes it more difficult for the lungs to work properly. As pulmonary fibrosis worsens, the patient becomes progressively more short of breath.

In some embodiments, the method comprising administering to a subject in need thereof an effective amount of an isolated nucleic acid or an rAAV as described herein. A subject may be any mammalian organism, for example a human, non-human primate, horse, pig, dog, cat rodent, *etc.* In some embodiments a subject is a human.

- 27 -

An “effective amount” of a substance is an amount sufficient to produce a desired effect. In some embodiments, an effective amount of an isolated nucleic acid is an amount sufficient to transfect (or infect in the context of rAAV mediated delivery) a sufficient number of target cells of a target tissue of a subject. In some embodiments, a target tissue is lung tissue (*e.g.*, alveolar type II cells, *etc.*). In some embodiments, an effective amount of an isolated nucleic acid (*e.g.*, which may be delivered via an rAAV) may be an amount sufficient to have a therapeutic benefit in a subject, *e.g.*, to improve lung function of the subject, to extend the lifespan of a subject, to improve in the subject one or more symptoms of disease (*e.g.*, shortness of breath), *etc.* The effective amount will depend on a variety of factors such as, for example, the species, age, weight, health of the subject, and the tissue to be targeted, and may thus vary among subject and tissue as described elsewhere in the disclosure.

As used herein, the term “treating” refers to the application or administration of a composition encoding a transgene(s) to a subject, who has a mutation in the HPS1 gene, with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve, or affect the disorder, or the symptom of the disease. In some embodiments, the administration of a composition described herein increases a responsible gene (*e.g.*, HPS1 gene) expression level and/or activity by 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% compared to a reference value. Methods of measuring gene (*e.g.*, HPS1 gene) expression level and/or activity are known in the art. Non-limiting exemplary reference value can be gene (*e.g.*, HPS1 gene) expression and/or activity of the same subject prior to receiving the treatment.

Alleviating a disease includes delaying the development or progression of the disease, or reducing disease severity. Alleviating the disease does not necessarily require curative results. As used therein, “delaying” the development of a disease (such as Hermansky–Pudlak syndrome-associated pulmonary fibrosis) means to defer, hinder, slow, retard, stabilize, and/or postpone progression of the disease. This delay can be of varying lengths of time, depending on the history of the disease and/or individuals being treated. A method that “delays” or alleviates the development of a disease, or delays the onset of the disease, is a method that reduces probability of developing one or more symptoms of the disease in a given time frame and/or reduces extent of the symptoms in a given time frame, when compared to not using the method. Such comparisons are typically based on clinical studies, using a number of subjects sufficient to give a statistically significant result.

- 28 -

"Development" or "progression" of a disease means initial manifestations and/or ensuing progression of the disease. Development of the disease can be detectable and assessed using standard clinical techniques as well known in the art. However, development also refers to progression that may be undetectable. For purpose of this disclosure, development or
5 progression refers to the biological course of the symptoms. "Development" includes occurrence, recurrence, and onset. As used herein "onset" or "occurrence" of a disease includes initial onset and/or recurrence.

The isolated nucleic acids and rAAVs of the disclosure may be delivered to a subject in compositions according to any appropriate methods known in the art. For example, an rAAV,
10 preferably suspended in a physiologically compatible carrier (*i.e.*, in a composition), may be administered to a subject, *i.e.* host animal, such as a human, mouse, rat, cat, dog, sheep, rabbit, horse, cow, goat, pig, guinea pig, hamster, chicken, turkey, or a non-human primate (*e.g.*, Macaque). In some embodiments a host animal does not include a human.

Delivery of the rAAVs to a mammalian subject may be by, for example, intramuscular
15 injection or by administration into the bloodstream of the mammalian subject. Administration into the bloodstream may be by injection into a vein, an artery, or any other vascular conduit. In some embodiments, the rAAVs are administered into the bloodstream by way of isolated limb perfusion, a technique well known in the surgical arts, the method essentially enabling the artisan to isolate a limb from the systemic circulation prior to administration of the rAAV
20 virions. A variant of the isolated limb perfusion technique, described in U.S. Pat. No. 6,177,403, can also be employed by the skilled artisan to administer the virions into the vasculature of an isolated limb to potentially enhance transduction into muscle cells or tissue.

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

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

- 29 -

Optionally, the compositions of the disclosure may contain, in addition to the rAAV and carrier(s), other conventional pharmaceutical ingredients, such as preservatives, or chemical stabilizers. Suitable exemplary preservatives include chlorobutanol, potassium sorbate, sorbic acid, sulfur dioxide, propyl gallate, the parabens, ethyl vanillin, glycerin, phenol, and
5 parachlorophenol. Suitable chemical stabilizers include gelatin and albumin.

The rAAVs are administered in sufficient amounts to transfect the cells of a desired tissue and to provide sufficient levels of gene transfer and expression without undue adverse effects. Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, direct delivery to the selected organ (*e.g.*, lungs), oral, inhalation (including
10 intranasal and intratracheal delivery), intravenous, intramuscular, subcutaneous, intradermal, , and other parental routes of administration. Routes of administration may be combined, if desired. In some embodiments, the rAAVs or composition thereof is delivered to the lungs by intratracheal instillation, oro-tracheal Injection, intranasal administration, oropharyngeal aspiration, or nebulization administration.

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

An effective amount of an rAAV is an amount sufficient to target infect an animal, target a desired tissue (*e.g.*, lungs). In some embodiments, an effective amount of an rAAV is an amount sufficient to produce a stable somatic transgenic animal model. The effective amount
25 will depend primarily on factors such as the species, age, weight, health of the subject, and the tissue to be targeted, and may thus vary among animal and tissue. For example, an effective amount of the rAAV is generally in the range of from about 1 ml to about 100 ml of solution containing from about 10^9 to 10^{16} genome copies. In some cases, a dosage between about 10^{11} to 10^{13} rAAV genome copies is appropriate. In certain embodiments, 10^{12} or 10^{13} rAAV
30 genome copies is effective to target lung tissue (*e.g.*, ATII cells). In some cases, stable transgenic animals are produced by multiple doses of an rAAV.

- 30 -

In some embodiments, a dose of rAAV is administered to a subject no more than once per calendar day (*e.g.*, a 24-hour period). In some embodiments, a dose of rAAV is administered to a subject no more than once per 2, 3, 4, 5, 6, or 7 calendar days. In some embodiments, a dose of rAAV is administered to a subject no more than once per calendar week (*e.g.*, 7 calendar days). In some embodiments, a dose of rAAV is administered to a subject no more than bi-weekly (*e.g.*, once in a two-calendar week period). In some embodiments, a dose of rAAV is administered to a subject no more than once per calendar month (*e.g.*, once in 30 calendar days). In some embodiments, a dose of rAAV is administered to a subject no more than once per six calendar months. In some embodiments, a dose of rAAV is administered to a subject no more than once per calendar year (*e.g.*, 365 days or 366 days in a leap year).

In some embodiments, rAAV compositions are formulated to reduce aggregation of AAV particles in the composition, particularly where high rAAV concentrations are present (*e.g.*, $\sim 10^{13}$ GC/ml or more). Methods for reducing aggregation of rAAVs are well known in the art and include, for example, addition of surfactants, pH adjustment, salt concentration adjustment, *etc.* (See, *e.g.*, Wright FR, et al., Molecular Therapy (2005) 12, 171–178, the contents of which are incorporated herein by reference.)

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

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

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations

- 31 -

contain a preservative to prevent the growth of microorganisms. In many cases the form is sterile and fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

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

Sterile injectable solutions are prepared by incorporating the active rAAV in the required amount in the appropriate solvent with various of the other ingredients enumerated herein, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation

- 32 -

are vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

The rAAV compositions disclosed herein may also be formulated in a neutral or salt form. Pharmaceutically-acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, drug-release capsules, and the like.

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

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

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

Liposomes have been used successfully with a number of cell types that are normally resistant to transfection by other procedures. In addition, liposomes are free of the DNA length

- 33 -

constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, drugs, radiotherapeutic agents, viruses, transcription factors and allosteric effectors into a variety of cultured cell lines and animals. In addition, several successful clinical trials examining the effectiveness of liposome-mediated drug delivery have
5 been completed.

Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs)). MLVs generally have diameters of from 25 nm to 4 μ m. Sonication of MLVs results in the formation of small unilamellar vesicles (SUVs) with diameters in the range of 200 to 500 Å,
10 containing an aqueous solution in the core.

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

In addition to the methods of delivery described above, the following techniques are also contemplated as alternative methods of delivering the rAAV compositions to a host.

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

25 VI. General techniques

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, biochemistry, and immunology, which are within the skill of the art. Molecular Cloning: A Laboratory Manual, second edition (Sambrook, et al., 1989) Cold Spring
30 Harbor Press; Oligonucleotide Synthesis (M. J. Gait, ed., 1984); Methods in Molecular Biology, Humana Press; Cell Biology: A Laboratory Notebook (J. E. Cellis, ed., 1998) Academic Press; Animal Cell Culture (R. I. Freshney, ed., 1987); Introduction to Cell and Tissue Culture (J. P.

- 34 -

Mather and P. E. Roberts, 1998) Plenum Press; Cell and Tissue Culture: Laboratory Procedures (A. Doyle, J. B. Griffiths, and D. G. Newell, eds., 1993-8) J. Wiley and Sons; Methods in Enzymology (Academic Press, Inc.); Handbook of Experimental Immunology (D. M. Weir and C. C. Blackwell, eds.); Gene Transfer Vectors for Mammalian Cells (J. M. Miller and M. P. Calos, eds., 1987); Current Protocols in Molecular Biology (F. M. Ausubel, et al., eds., 1987); PCR: The Polymerase Chain Reaction, (Mullis, et al., eds., 1994); Current Protocols in Immunology (J. E. Coligan et al., eds., 1991); Short Protocols in Molecular Biology (Wiley and Sons, 1999); Immunobiology (C. A. Janeway and P. Travers, 1997); Antibodies (P. Finch, 1997); Antibodies: a practical approach (D. Catty., ed., IRL Press, 1988-1989); Monoclonal antibodies: a practical approach (P. Shepherd and C. Dean, eds., Oxford University Press, 2000); Using antibodies: a laboratory manual (E. Harlow and D. Lane (Cold Spring Harbor Laboratory Press, 1999); The Antibodies (M. Zanetti and J. D. Capra, eds., Harwood Academic Publishers, 1995)

Without further elaboration, it is believed that one skilled in the art can, based on the present disclosure, utilize the present invention to its fullest extent. The following specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever. All publications cited herein are incorporated by reference for the purposes or subject matter referenced herein.

Exemplary embodiments of the invention will be described in more detail by the following examples. These embodiments are exemplary of the invention, which one skilled in the art will recognize is not limited to the exemplary embodiments.

EXAMPLES

Example 1

Hermansky-Pudlak syndrome (HPS) is a group of 10 autosomal recessive multisystemic disorders, each defined by the deficiency of a specific gene. HPS is an orphan disease occurring in 1 in 500,000 to 1,000,000 individuals worldwide, and more than 50% of cases are diagnosed on the Caribbean island of Puerto Rico. HPS-associated genes encode components of four ubiquitously expressed protein complexes: Adaptor protein-3 (AP-3) and biogenesis of lysosome-related organelles complex-1 (BLOC-1) through -3 that are critical for intracellular protein trafficking. All individuals with HPS exhibit albinism, and bleeding diathesis; additional

- 35 -

phenotypes occur depending on the defective protein complex. HPS-associated pulmonary fibrosis (HPS-PF) remains the most serious complication for patients with HPS-1, -2 and -4 types of the disease; most patients with HPS type 1 develop lethal form of pulmonary fibrosis that manifests as earlier as at 30 – 40 years of age. The progression of HPS1- PF is characterized by the development of dyspnea and increasingly debilitating hypoxemia. No curable therapies are currently approved for the HPS1-PF. The prognosis of HPS-PF is extremely poor, and currently, lung transplantation remains the only potentially life-prolonging treatment.

It was suggested that alveolar type II pneumocytes (ATII) play a significant role in the susceptibility of HPS patients to PF. Thus, correcting BLOC-3 function in ATII cells by HPS1 gene transfer or gene editing might be the most effective therapy for treating HPS1-PF. The first in vitro attempt to transfect isolated mouse HPS deficient ATII cells by electroporation with plasmid DNA encoding a specialized Rb38 protein (part of the BLOC-3) has demonstrated largely improved regulation of the surfactant production by ATII cells. In vivo delivery, however, requires a vector mediated gene transfer system such as a recombinant adeno-associated virus (rAAV). Advantages of rAAV as a respiratory gene delivery vector include: efficient transduction, lack of pathogenicity, low cellular immunogenicity, lack of integration into host genome, ability to infect dividing and non-dividing cells, and persistent expression of therapeutic gene of interest.

In this study, an adeno-associated virus mediated gene augmentation therapy to treat pulmonary fibrosis in Hermansky-Pudlak Syndrome type 1 was developed. The results of this study showed that AAV-mediated delivery of the (HPS1 Biogenesis Of Lysosomal Organelles Complex 3 Subunit 1 (*HPS1*) gene provides clinically relevant therapeutic effect for treating pulmonary fibrosis in Hermansky-Pudlak Syndrome type 1.

The AAV vectors encode either murine *Hps1* gene (FIGs 1A-1B), or human *HPS1* gene (FIGs. 1C-1D). Both constructs contain a single-stranded DNA molecule with mutated AAV serotype 2 inverted terminal repeats (ITRs) flanking a gene cassette, which includes a cytomegalovirus immediate – early enhancer/chicken b-actin hybrid promoter, complementary DNA encoding wild-type murine HPS1 protein (AAV.CB-mHps1) or human HPS1 protein (AAV.CB-hHPS1), and an SV40 polyadenylation signal (FIGs. 1A-1D).

The AAV vectors were packaged either in rAAV serotype 5 capsid, or rAAV serotype 8 capsid, and purified. The purified AAV5.CB-mHps1 had a titer of 3.0×10^{12} genome copies/ml, and AAV8.CB-mHps1 had a titer 2.0×10^{12} genome copies/ml.

- 36 -

Initially, the efficiency and stability of the rAAV-eGFP reporter gene transfer via intratracheal instillation into adult wild-type (C57Bl6/J) mouse lungs was studied. The results based on in situ GFP protein RNA hybridization showed that AAV serotype 5 capsid effectively transduced the targeted lung epithelial cells in adult mice (FIGs. 2A-2I).

Further, the rAAV-mediated *Hps1* gene transduction, protein expression and its functionality were assessed in *Hps1* (Hermansky-Pudlak Syndrome type 1) deficient mice B6-GE(Δ HPS1) ^{Δ HPS1Ex5} and B6-GE(Δ HPS1) ^{Δ HPS1Prom} (*Hps1* KO mice).

Hps1 KO mice largely recapitulate human phenotypes, in particular, hypopigmentation, bleeding diathesis and alveolar type 2 cell dysplasia associated with increased susceptibility to development of pulmonary fibrosis. Lung cell morphopathology (FIGs. 3A-3D) and pulmonary phenotypes were characterized in these mice. Aged naïve *Hps1* KO mice develop spontaneous fibrotic-like pulmonary phenotype. Such fibrotic development that can be accelerated by pro-inflammatory agents like lipopolysaccharide (LPS) or bleomycin in dose-dependent manner triggering injury to the lung epithelial lining in younger *Hps1* KO mice (Figs. 4A-4G).

AAV.CB-mHps1 was delivered to the lungs of adult *Hps1* KO mice two weeks prior LPS challenge. LPS challenge was chosen over bleomycin because it closely mimics the innate immune response to the pathogenic bacteria. Animals received single oro-tracheal vector instillation at 10^{11} genome copies in a total volume of 40 μ l (or PBS for control) into the lungs via trachea. The results show that *Hps1* KO mice treated with AAV.CB-mHps1 were significantly protected from the pulmonary exacerbation as compared to the LPS challenged untreated mice (FIGs. 5A-5C).

The hypothesis that AAV-Hps1 intervention before the onset of pulmonary fibrosis can prevent the alveolar type 2 cell dysplasia in *Hps1* KO mice. Two cohorts of mice were tested either with AAV5-eGFP, or AAV8-eGFP. eGFP reporter expression was assessed 1 month later, and eGFP were highly expressed in the lungs of the mice treated with both AAV5-eGFP and AAV8-eGFP. Another group of neonatal mice (one-day old) was treated with a single systemic dose of AAV.CB-mHps1 at 10^{12} genome copies per mouse in a total volume of 50 μ l via facial vein. Mice are aged for six months, and lung phenotype are assessed.

Importantly, AAV-mediated *Hps1* gene augmentation in *Hps1* KO mice increased *Hps1* gene expression in the lungs, as measured by qPCR. In addition, it was also shown that AAV-mediated *Hps1* gene augmentation led to the reduction in the number of foamy type II alveolar epithelial cells by histology.

- 37 -

The results indicate that the Hps1 gene expression resulted in stable over time transduction of the target lung cells and in production of the functional Hps1 protein, leading to phenotypic correction of the alveolar type 2 cells and improved pulmonary function in a mouse model of Hermansky-Pudlak Syndrome type 1. Moreover, the rAAV proved to be safe for
5 targeted single-dose delivery through either the respiratory system (oro-tracheal instillation), intrapleural injection, or infused systemically (via vein injection).

EQUIVALENTS

While several inventive embodiments have been described and illustrated herein, those
10 of ordinary skill in the art will readily envision a variety of other means and/or structures for performing the function and/or obtaining the results and/or one or more of the advantages described herein, and each of such variations and/or modifications is deemed to be within the scope of the inventive embodiments described herein. More generally, those skilled in the art will readily appreciate that all parameters, dimensions, materials, and configurations described
15 herein are meant to be exemplary and that the actual parameters, dimensions, materials, and/or configurations will depend upon the specific application or applications for which the inventive teachings is/are used. Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific inventive embodiments described herein. It is, therefore, to be understood that the foregoing embodiments are presented
20 by way of example only and that, within the scope of the appended claims and equivalents thereto, inventive embodiments may be practiced otherwise than as specifically described and claimed. Inventive embodiments of the present disclosure are directed to each individual feature, system, article, material, kit, and/or method described herein. In addition, any combination of two or more such features, systems, articles, materials, kits, and/or methods, if
25 such features, systems, articles, materials, kits, and/or methods are not mutually inconsistent, is included within the inventive scope of the present disclosure.

All definitions, as defined and used herein, should be understood to control over dictionary definitions, definitions in documents incorporated by reference, and/or ordinary meanings of the defined terms.

30 All references, patents and patent applications disclosed herein are incorporated by reference with respect to the subject matter for which each is cited, which in some cases may encompass the entirety of the document.

- 38 -

The indefinite articles “a” and “an,” as used herein in the specification and in the claims, unless clearly indicated to the contrary, should be understood to mean “at least one.”

The phrase “and/or,” as used herein in the specification and in the claims, should be understood to mean “either or both” of the elements so conjoined, i.e., elements that are

5 conjunctively present in some cases and disjunctively present in other cases. Multiple elements listed with “and/or” should be construed in the same fashion, i.e., “one or more” of the elements so conjoined. Other elements may optionally be present other than the elements specifically identified by the “and/or” clause, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, a reference to “A and/or B”, when used in
10 conjunction with open-ended language such as “comprising” can refer, in one embodiment, to A only (optionally including elements other than B); in another embodiment, to B only (optionally including elements other than A); in yet another embodiment, to both A and B (optionally including other elements); etc.

As used herein in the specification and in the claims, “or” should be understood to have
15 the same meaning as “and/or” as defined above. For example, when separating items in a list, “or” or “and/or” shall be interpreted as being inclusive, i.e., the inclusion of at least one, but also including more than one, of a number or list of elements, and, optionally, additional unlisted items. Only terms clearly indicated to the contrary, such as “only one of” or “exactly one of,” or, when used in the claims, “consisting of,” will refer to the inclusion of exactly one element of
20 a number or list of elements. In general, the term “or” as used herein shall only be interpreted as indicating exclusive alternatives (i.e. “one or the other but not both”) when preceded by terms of exclusivity, such as “either,” “one of,” “only one of,” or “exactly one of.” “Consisting essentially of,” when used in the claims, shall have its ordinary meaning as used in the field of patent law.

25 As used herein in the specification and in the claims, the phrase “at least one,” in reference to a list of one or more elements, should be understood to mean at least one element selected from any one or more of the elements in the list of elements, but not necessarily including at least one of each and every element specifically listed within the list of elements and not excluding any combinations of elements in the list of elements. This definition also
30 allows that elements may optionally be present other than the elements specifically identified within the list of elements to which the phrase “at least one” refers, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, “at least one of A and

- 39 -

B” (or, equivalently, “at least one of A or B,” or, equivalently “at least one of A and/or B”) can refer, in one embodiment, to at least one, optionally including more than one, A, with no B present (and optionally including elements other than B); in another embodiment, to at least one, optionally including more than one, B, with no A present (and optionally including elements other than A); in yet another embodiment, to at least one, optionally including more than one, A, and at least one, optionally including more than one, B (and optionally including other elements); etc.

In the claims, as well as in the specification above, all transitional phrases such as “comprising,” “including,” “carrying,” “having,” “containing,” “involving,” “holding,” and the like are to be understood to be open-ended, i.e., to mean including but not limited to. Only the transitional phrases “consisting of” and “consisting essentially of” shall be closed or semi-closed transitional phrases, respectively, as set forth in the United States Patent Office Manual of Patent Examining Procedures, Section 2111.03.

Use of ordinal terms such as “first,” “second,” “third,” etc., in the claims to modify a claim element does not by itself connote any priority, precedence, or order of one claim element over another or the temporal order in which acts of a method are performed, but are used merely as labels to distinguish one claim element having a certain name from another element having a same name (but for use of the ordinal term) to distinguish the claim elements.

It should also be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one step or act, the order of the steps or acts of the method is not necessarily limited to the order in which the steps or acts of the method are recited.

SEQUENCES

All NCBI Gene and Accession Number Sequences are incorporated herein by reference in their entireties.

- 40 -

CLAIMS

What is claimed is:

1. An isolated nucleic acid comprising an expression cassette comprising a nucleotide sequence encoding a Hermansky–Pudlak syndrome 1 (HPS1) protein, wherein the expression cassette is flanked by adeno-associated virus (AAV) inverted terminal repeats (ITRs).

2. The isolated nucleic acid of claim 1, wherein the Hermansky–Pudlak syndrome 1 (HPS1) protein comprises an amino acid sequence at least 80% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 1-8.

3. The isolated nucleic acid of claims 1 or 2, wherein the nucleotide sequence encoding the HPS1 protein comprises a sequence at least 80% identical to the nucleic acid sequence set forth in any one of SEQ ID NO: 9-16.

4. The isolated nucleic acid of any one of claims 1-3, wherein the expression cassette further comprises a promoter operably linked to the nucleotide sequence encoding the HPS1 protein.

5. The isolated nucleic acid of claim 4, wherein the promoter is a cytomegalovirus immediate – early enhancer/chicken β -actin hybrid promoter.

6. The isolated nucleic acid of any one of claims 1 to 5, further comprising a polyadenylation (polyA) signal.

7. The isolated nucleic acid of claim 6, wherein the polyA signal is SV40 polyadenylation signal.

8. The isolated nucleic acid of any one of claims 1-7, wherein the ITRs are adeno-associated virus ITRs of a serotype selected from the group consisting of AAV1 ITR, AAV2 ITR, AAV3 ITR, AAV4 ITR, AAV5 ITR, and AAV6 ITR.

- 41 -

9. The isolated nucleic acid of claim 8, wherein the ITRs are AAV2 ITRs, or a variant thereof.

10. The isolated nucleic acid of claim 9, wherein the AAV2 ITR variant is a mutated AAV2 ITR.

11. A vector comprising the isolated nucleic acid of any one of claims 1-10.

12. The vector of claim 11, wherein the vector is a rAAV vector.

13. A recombinant adeno-associated virus (rAAV) vector comprising a nucleic acid comprising, in 5' to 3' order:

(a) a 5' AAV2 mutant ITR;

(b) a cytomegalovirus immediate – early enhancer/chicken b-actin hybrid promoter;

(c) a nucleotide sequence encoding an Hermansky–Pudlak syndrome 1 (HPS1) protein;

(d) a SV40 polyA signal; and

(e) a 3' AAV2 mutant ITR.

14. A recombinant adeno-associated virus (rAAV) comprising:

(i) an AAV capsid protein; and

(ii) the isolated nucleic acid of any one of claims 1-10, or the vector of any one of claims 11-13.

15. The rAAV of claim 14, wherein the capsid protein is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9 capsid protein, or a variant thereof.

16. The rAAV of claims 14 or 15, wherein the capsid protein is an AAV5 capsid protein, or a variant thereof.

17. The rAAV of claims 14 or 15, wherein the capsid protein is an AAV8 capsid protein, or a variant thereof.

- 42 -

18. The rAAV of any one of claims 14-17, wherein the capsid protein is capable of delivering the HPS1 protein into lung cells.

19. The rAAV of claim 18, wherein the lung cells are alveolar type II epithelial (ATII) cells.

20. A pharmaceutical composition comprising the isolated nucleic acid of any one of claims 1-10, the vector of any one of claims 11-14, or the rAAV of any one of claims 14-19.

21. The pharmaceutical composition of claim 20, further comprising a pharmaceutically acceptable carrier.

22. A host cell comprising the isolated nucleic acid of any one of claims 1-10, the vector of any one of claims 11-14, the rAAV of any one of claims 14-19, or the pharmaceutical composition of claims 20 or 21.

23. A method for delivering a Hermansky–Pudlak syndrome 1 (HPS1) protein to a cell or a subject, the method comprising delivering to the cell or the subject the isolated nucleic acid of any one of claims 1-10, the vector of any one of claims 11-14, the rAAV of any one of claims 14-19, the pharmaceutical composition of claims 20 or 21, or the host cell of claim 22.

24. A method for treating Hermansky–Pudlak syndrome-associated pulmonary fibrosis in a subject, the method comprising administering to the subject an effective amount of the isolated nucleic acid of any one of claims 1-10, the vector of any one of claims 11-14, the rAAV of any one of claims 14-19, the pharmaceutical composition of claims 20 or 21, or the host cell of claim 22.

25. The method of claims 23 or 24, wherein the subject is a human or a non-human mammal.

26. The method of any one of claims 22-25, wherein the administration results in delivery of the HPS1 protein to the lungs.

- 43 -

27. The method of claim 26, wherein the administration results in delivery of the HPS1 protein to alveolar type II epithelial (ATII) cells.

5 28. The method of claim 27, wherein the administration results in at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or 100% increase in HPS1 expression level or activity in the ATII cells of the subject.

29. The method of any one of claims 22-28, wherein the administration comprises
10 systemic administration.

30. The method of claim 29, wherein the systemic administration is intravenous administration, intramuscular administration, subcutaneous administration, or intraperitoneal administration.

15

31. The method of any one of claims 22-28, wherein the administration comprises local administration to the lungs.

32. The method of claim 31, wherein the local administration to the lungs is
20 Intratracheal instillation, oro-tracheal Injection, intranasal administration, oropharyngeal aspiration, or nebulization administration.

25

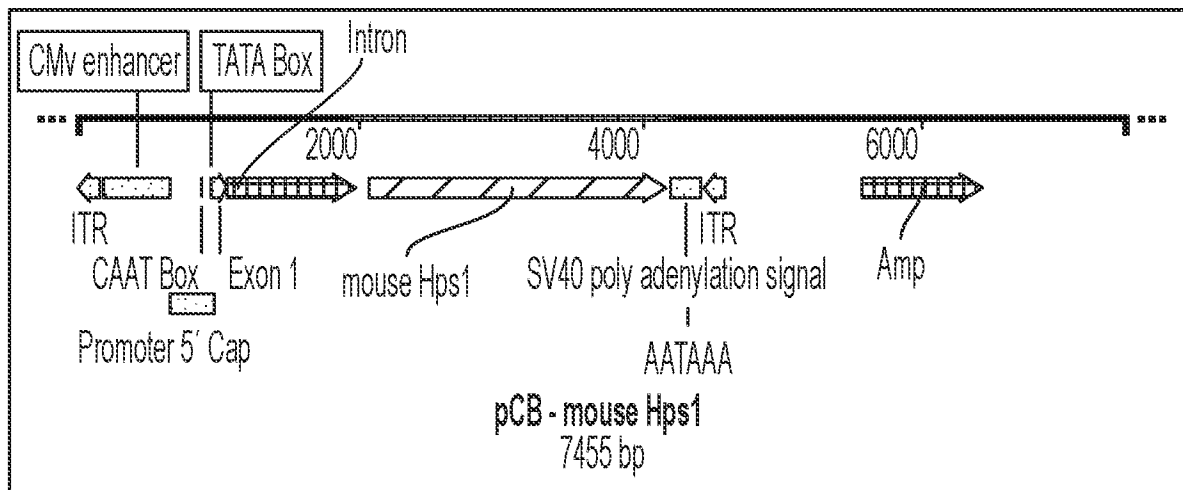


FIG. 1A

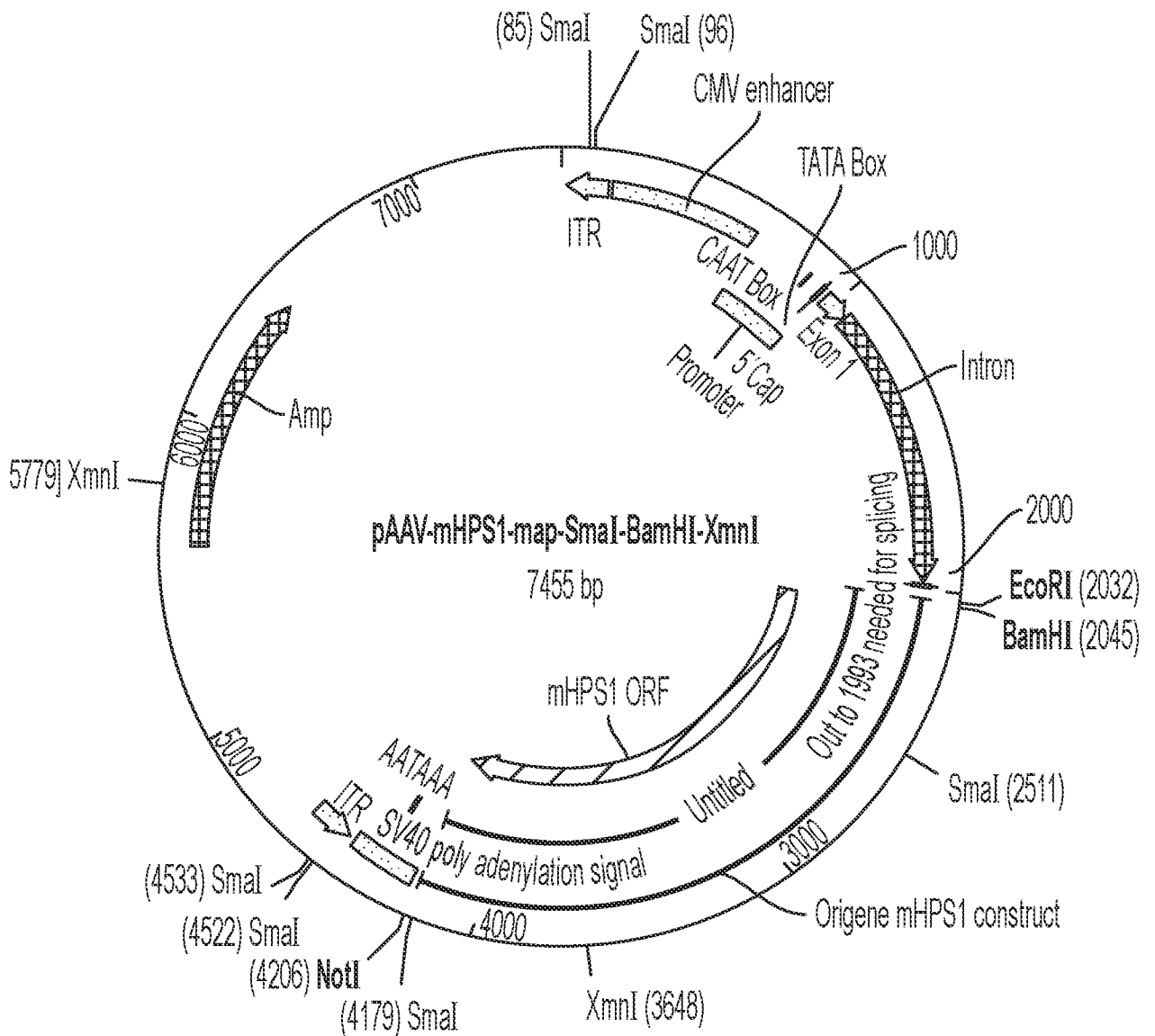


FIG. 1B

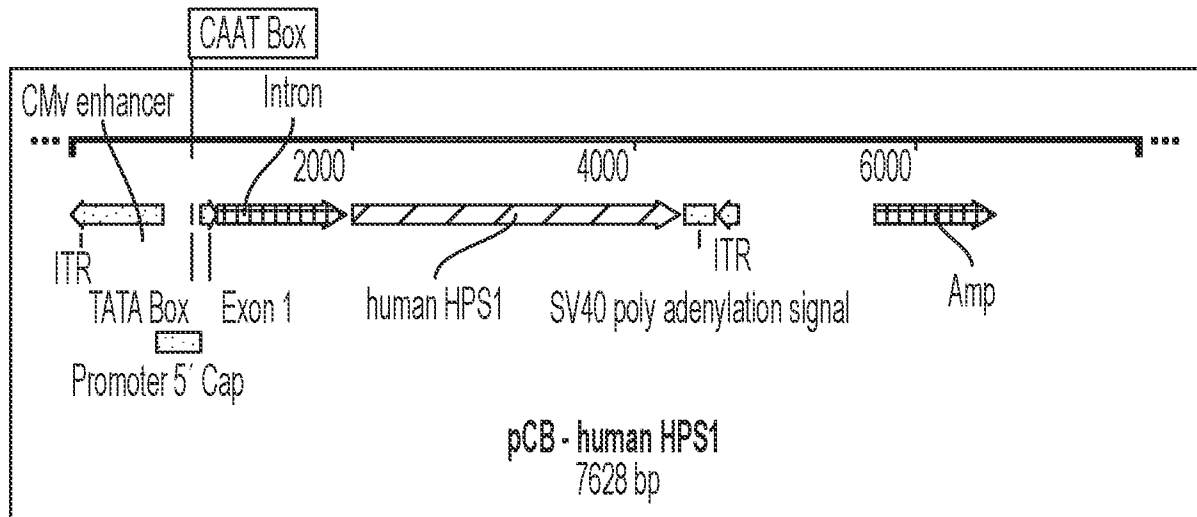


FIG. 1C

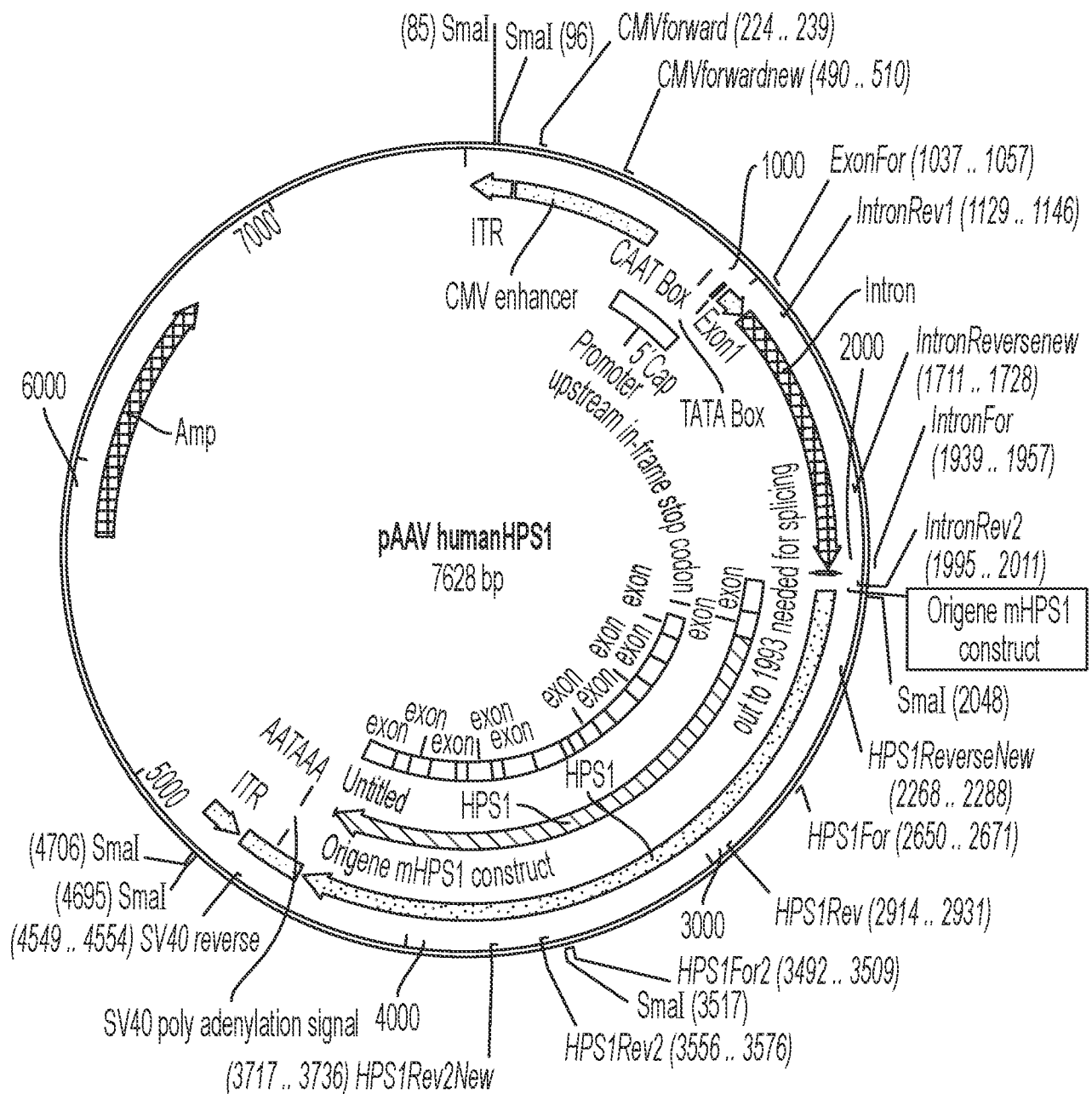
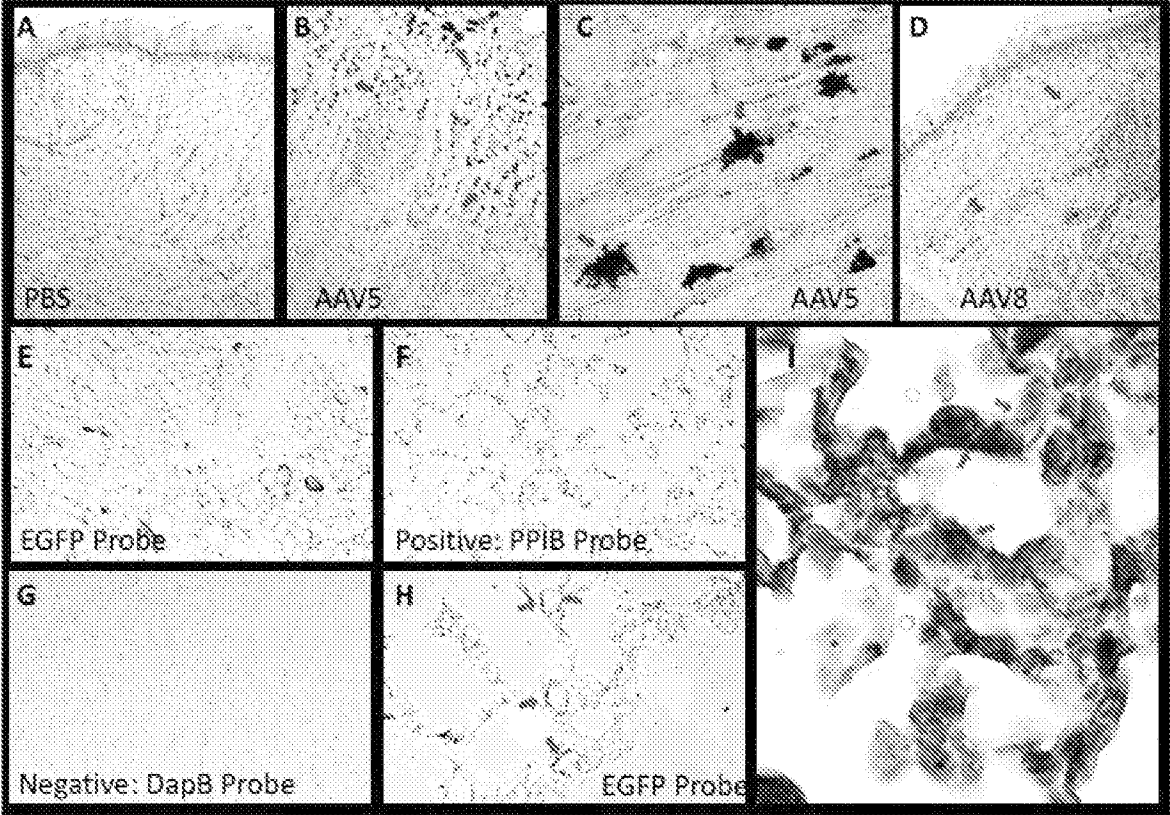
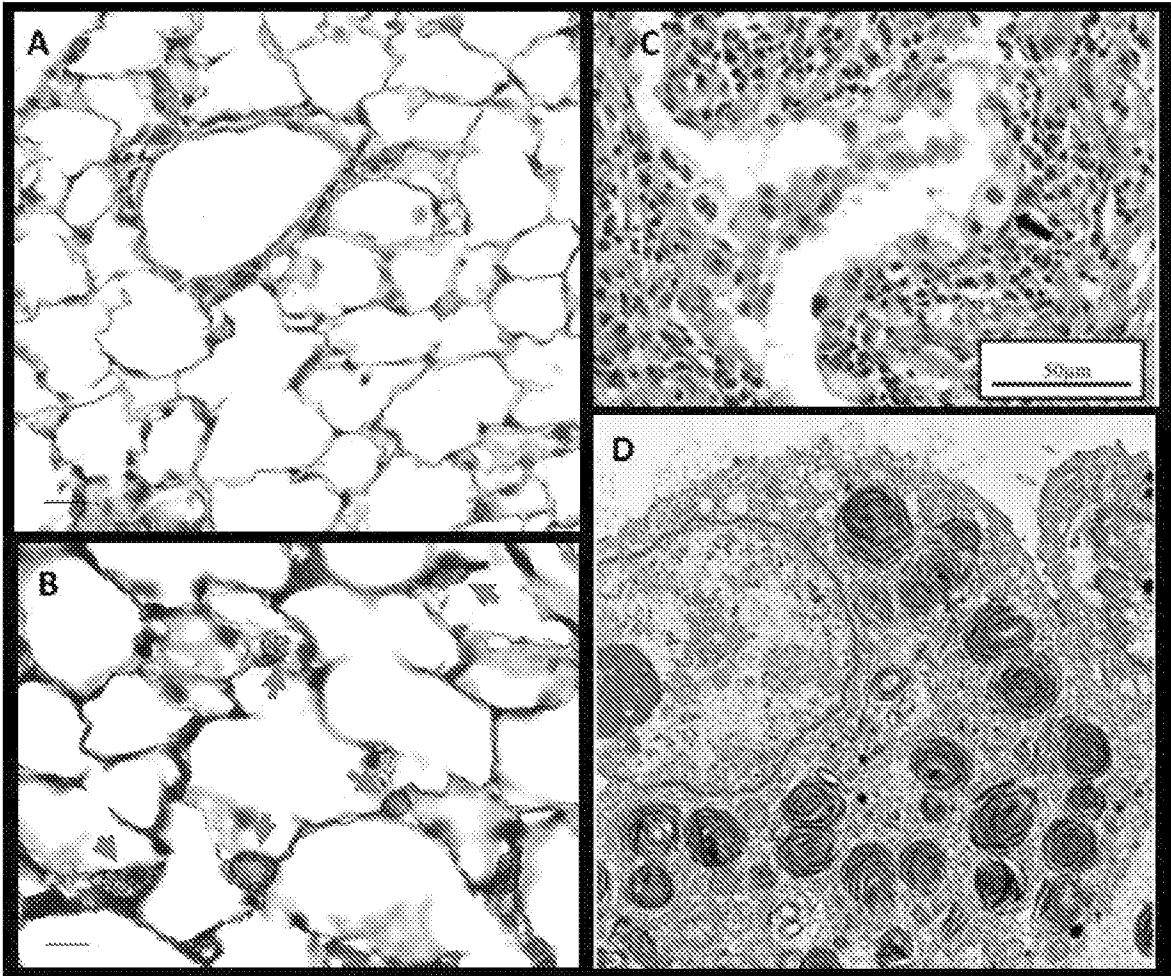


FIG. 1D



FIGs. 2A-2I



FIGs. 3A-3D

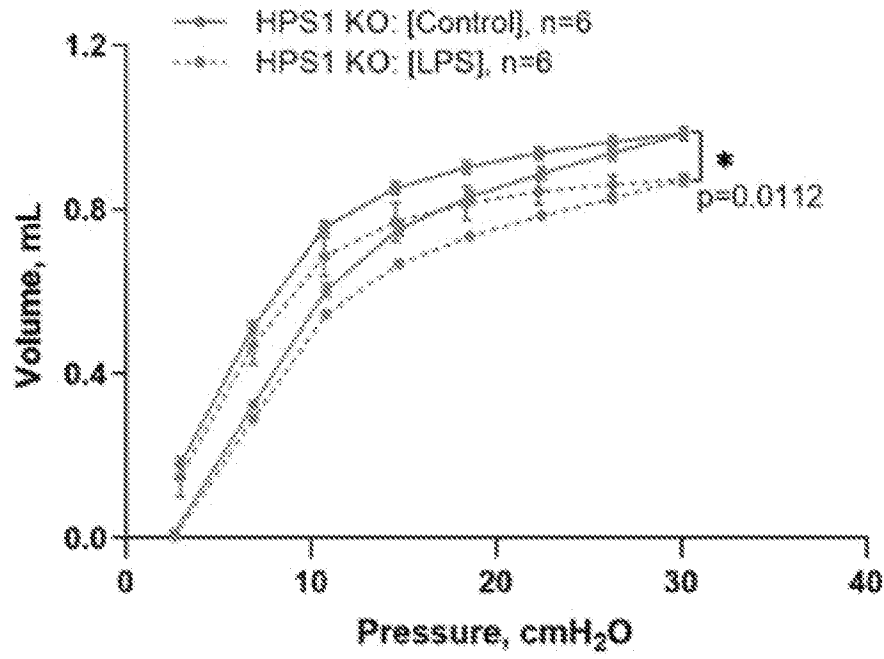


FIG. 4A

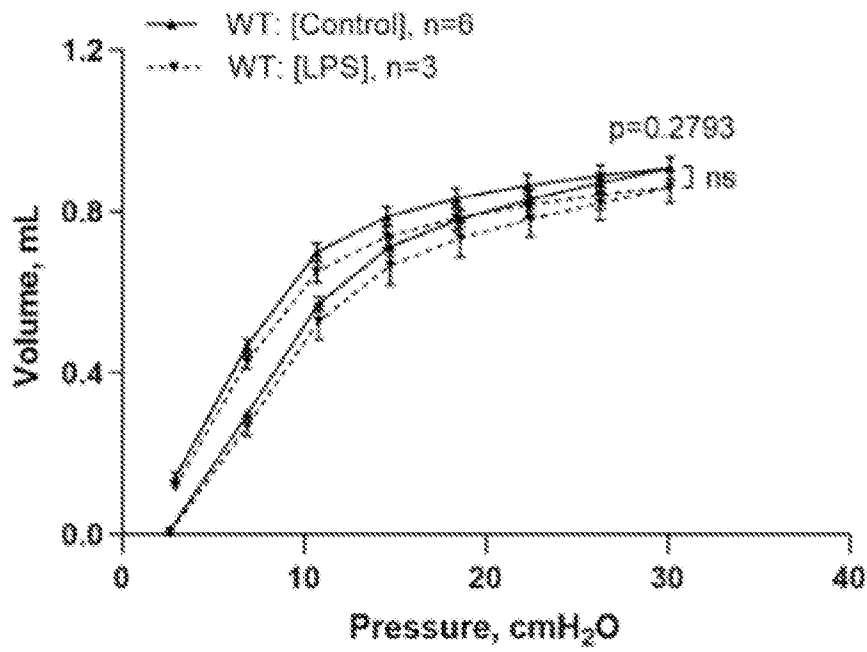


FIG. 4B

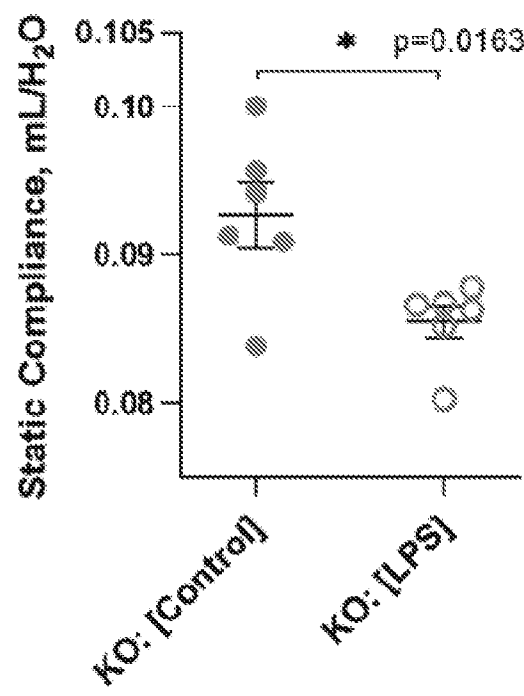


FIG. 4C

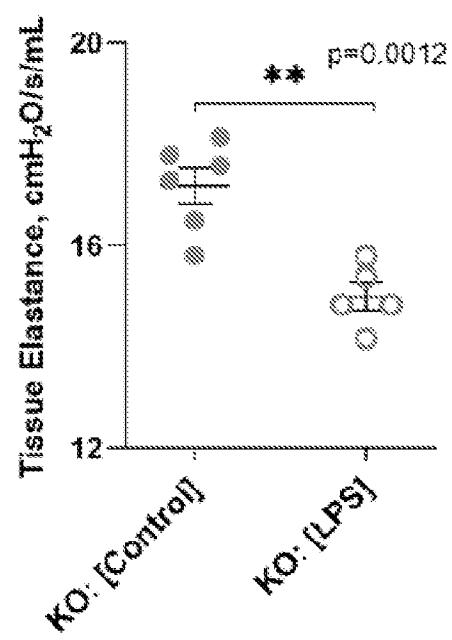


FIG. 4D

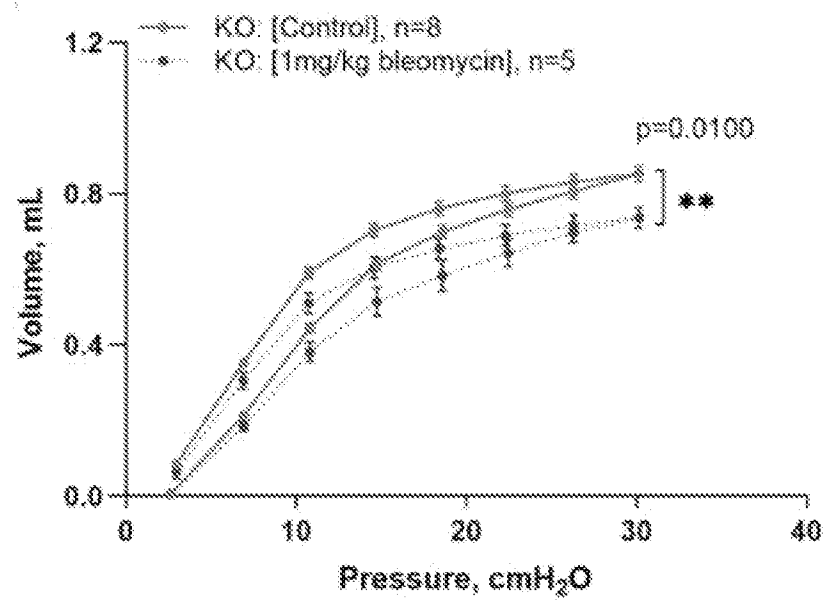


FIG. 4E

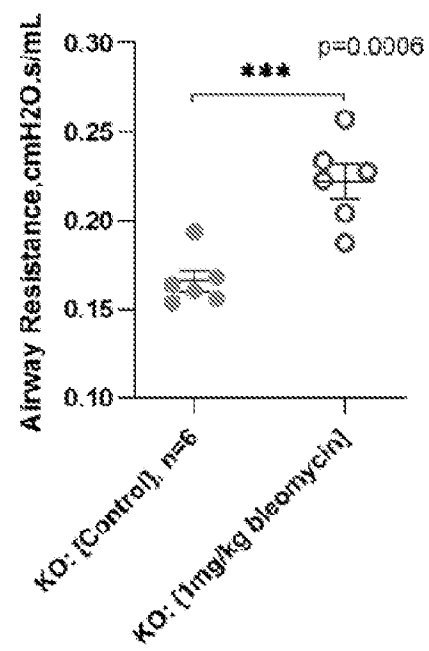


FIG. 4F

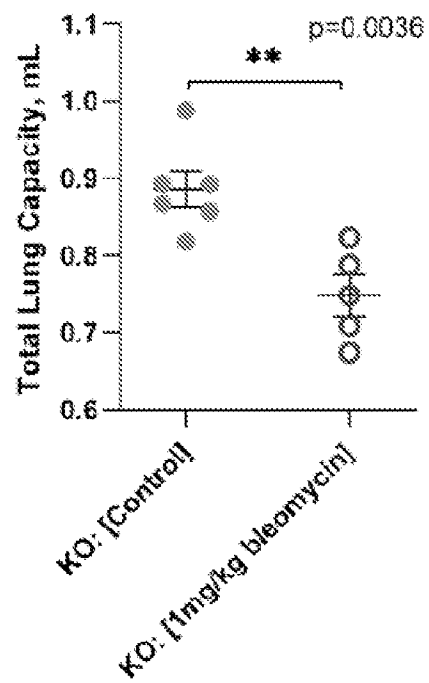


FIG. 4G

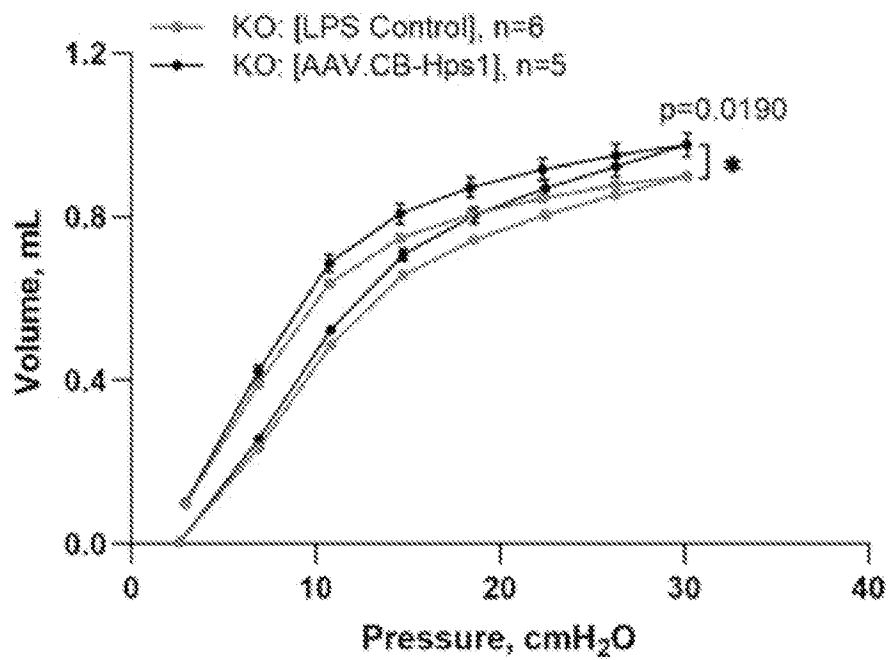


FIG. 5A

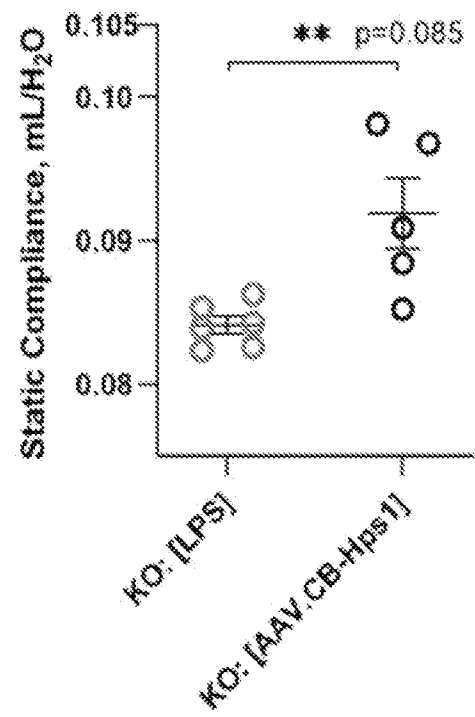


FIG. 5B

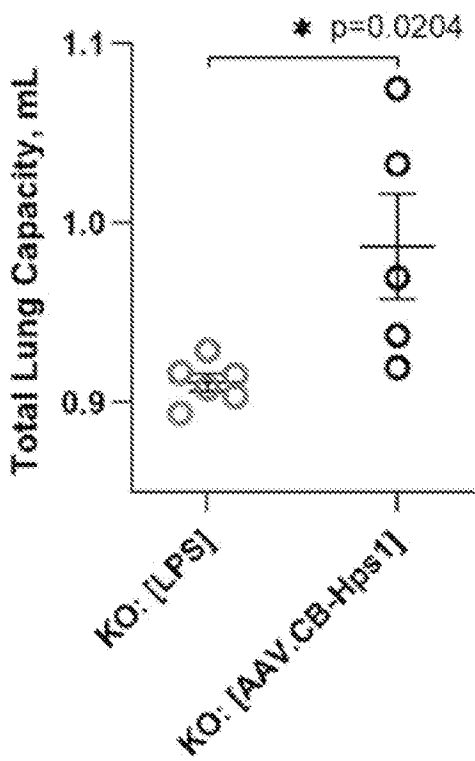


FIG. 5C

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2022/026101

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - A61K 38/16; A61K 35/76; A61K 48/00; A61P 43/00; C12N 15/63; C12N 15/86 (2022.01)

CPC - A61K 38/1709; A61K 35/76; A61K 48/00; C07K 14/005; C12N 7/00; C12N 15/86; C12N 2750/14143 (2022.08)

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

see Search History document

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

see Search History document

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

see Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2020/0316221 A1 (UNIVERSITY OF MASSACHUSETTS) 08 October 2020 (08.10.2020) entire document	1-3
Y	US 2013/0303558 A1 (CELERA CORPORATION) 14 November 2013 (14.11.2013) entire document	1-3
Y	US 2018/0110788 A1 (FRED HUTCHINSON CANCER RESEARCH CENTER) 26 April 2018 (26.04.2018) entire document	1-3
A	US 2013/0131145 A1 (CARTHEW et al) 23 May 2013 (23.05.2013) entire document	1-3

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

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

26 August 2022

Date of mailing of the international search report

SEP 21 2022

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Taina Matos

Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2022/026101

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

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

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

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

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

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

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

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

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

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

3. Additional comments:

SEQ ID NOs: 1 and 9 were searched.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2022/026101

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

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

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

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

3. ☒ Claims Nos.: 4-12, 14-32
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

This International Searching Authority found multiple inventions in this international application, as follows:

See extra sheet(s).

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-3

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

Continued from Box No. III Observations where unity of invention is lacking

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

Group I+: claims 1-3, and 13 are drawn to isolated nucleic acids and rAAVs encoding a Hermansky-Pudlak syndrome 1 (HPS 1) protein.

The first invention of Group I+ is restricted to an expression cassette comprising a Hermansky-Pudlak syndrome 1 (HPS1) protein selected to be SEQ ID NO: 1, encoded by SEQ ID NO: 9, where the HPS1 nucleic acid is flanked by adeno-associated virus (AAV) inverted terminal repeats (ITRs). It is believed that claims 1-3 read on this first named invention and thus these claims will be searched without fee to the extent that they read on SEQ ID NOs: 1 and 9.

Applicant is invited to elect additional Hermansky-Pudlak syndrome 1 (HPS1) proteins, encoding nucleic acids, and their respective, corresponding, SEQ ID NOs to be searched in a specific combination by paying additional fee for each set of election. An exemplary election would be an expression cassette comprising a Hermansky-Pudlak syndrome 1 (HPS 1) protein selected to be SEQ ID NO: 2, encoded by SEQ ID NO: 10, where the HPS1 nucleic acid is flanked by adeno-associated virus (AAV) inverted terminal repeats (ITRs). Additional Hermansky-Pudlak syndrome 1 (HPS1) proteins, encoding nucleic acids, and their respective, corresponding, SEQ ID NOs will be searched upon the payment of additional fees. Applicants must specify the claims that read on any additional elected inventions. Applicants must further indicate, if applicable, the claims which read on the first named invention if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined.

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

The Groups I+ formulas do not share a significant structural element responsible for treating Hermansky-Pudlak syndrome-associated pulmonary fibrosis, requiring the selection of alternative Hermansky-Pudlak syndrome 1 (HPS1) proteins and encoding nucleic acids where "the Hermansky-Pudlak syndrome 1 (HPS1) protein comprises an amino acid sequence at least 80% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 1-8," and "the nucleotide sequence encoding the HPS 1 protein comprises a sequence at least 80% identical to the nucleic acid sequence set forth in any one of SEQ ID NO: 9-16."

Additionally, even if Groups I+ were considered to share the technical features of an isolated nucleic acid comprising an expression cassette comprising a nucleotide sequence encoding a Hermansky-Pudlak syndrome 1 (HPS1) protein, wherein the expression cassette is flanked by adeno-associated virus (AAV) inverted terminal repeats (ITRs); a recombinant adeno-associated virus (rAAV) vector comprising a nucleic acid comprising, in 5' to 3' order: (a) a 5' AAV2 mutant ITR; (b) a cytomegalovirus immediate - early enhancer/chicken b-actin hybrid promoter; (c) a nucleotide sequence encoding an Hermansky-Pudlak syndrome 1 (HPS 1) protein; (d) a SV40 polyA signal; and e) a 3' AAV2 mutant ITR. However, these shared technical features do not represent a contribution over the prior art.

Specifically, US 2020/0316221 A1 to University Of Massachusetts discloses an isolated nucleic acid (provides an isolated nucleic acid, Para. [0005]) comprising an expression cassette comprising a nucleotide sequence encoding a Hermansky-Pudlak syndrome 1 (HPS1) protein (the disclosure relates to an AAV comprising a nucleic acid encoding a protein or functional RNA useful for the treatment of a condition, disease or disorder ...HPS1 ...recombinant AAVs comprising nucleic acids that express one or more of the foregoing genes or fragments thereof, Para. [0152]), wherein the expression cassette is flanked by adeno-associated virus (AAV) inverted terminal repeats (ITRs) (an isolated nucleic acid comprises AAV inverted terminal repeats (ITRs) flanking the transgene, Para. [0051]); a recombinant adeno-associated virus (rAAV) vector comprising a nucleic acid (recombinant AAVs (rAAVs), Para. [0052]) comprising, in 5' to 3' order: (a) a 5' AAV2 mutant ITR (the selected transgene sequence and associated regulatory elements are flanked by the 5' and 3' AAV ITR sequences, Para. [0132]; AAV inverted terminal repeats (ITRs) ...AAV is enhanced or changed by mutations occurring in the capsid proteins, Para. [0105]); (b) a cytomegalovirus immediate - early enhancer/chicken b-actin hybrid promoter (the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer) ...the SV40 promoter, Para. [0140]); (c) a nucleotide sequence encoding an Hermansky-Pudlak syndrome 1 (HPS1) protein (the disclosure relates to an AAV comprising a nucleic acid encoding a protein or functional RNA useful for the treatment of a condition, disease or disorder ...HPS1 ...recombinant AAVs comprising nucleic acids that express one or more of the foregoing genes or fragments thereof, Para. [0152]); (d) a SV40 polyA signal (the SV40 promoter, Para. [0140]; other native expression control elements, such as enhancer elements, polyadenylation sites, Para. [0142]); and e) a 3' AAV2 mutant ITR (the selected transgene sequence and associated regulatory elements are flanked by the 5' and 3' AAV ITR sequences, Para. [0132]; AAV inverted terminal repeats (ITRs) ...AAV is enhanced or changed by mutations occurring in the capsid proteins, Para. [0105]).

The inventions listed in Groups I+ therefore lack unity under Rule 13 because they do not share a same or corresponding special technical features.