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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 January 2020 (16.01.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/014523 A1

(51) International Patent Classification:
A61K 48/00 (2006.01) *C12N 15/63* (2006.01)
C07K 14/47 (2006.01) *C12N 15/79* (2006.01)
C12N 15/09 (2006.01)

(21) International Application Number:
PCT/US2019/041465

(22) International Filing Date:
11 July 2019 (11.07.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/697,302 12 July 2018 (12.07.2018) US

(71) Applicant: **ROCKET PHARMACEUTICALS, LTD.** [US/US]; The Empire State Building, 350 Fifth Avenue, Suite 7530, New York, New York 10118 (US).

(72) Inventors: **KERAVALA, Annahita**; c/o Rocket Pharmaceuticals, Ltd., The Empire State Building, 350 Fifth Avenue. Suite 7530, New York, New York 10118 (US). **MOORE, Simon**; c/o Rocket Pharmaceuticals, Ltd., The Empire State Building, 350 Fifth Avenue, Suite 7530, New York, New York 10118 (US). **RICKS, David**; c/o Rocket Pharmaceuticals, Ltd., The Empire State Building, 350 Fifth Avenue. Suite 7530, New York, New York 10118 (US).

(74) Agent: **CHRISTIANSEN, William** et al.; Cooley LLP, 1299 Pennsylvania Aveenue, NW, Suite 700, Washington, District of Columbia 20004 (US).

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

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

(54) Title: GENE THERAPY VECTORS FOR TREATMENT OF DANON DISEASE

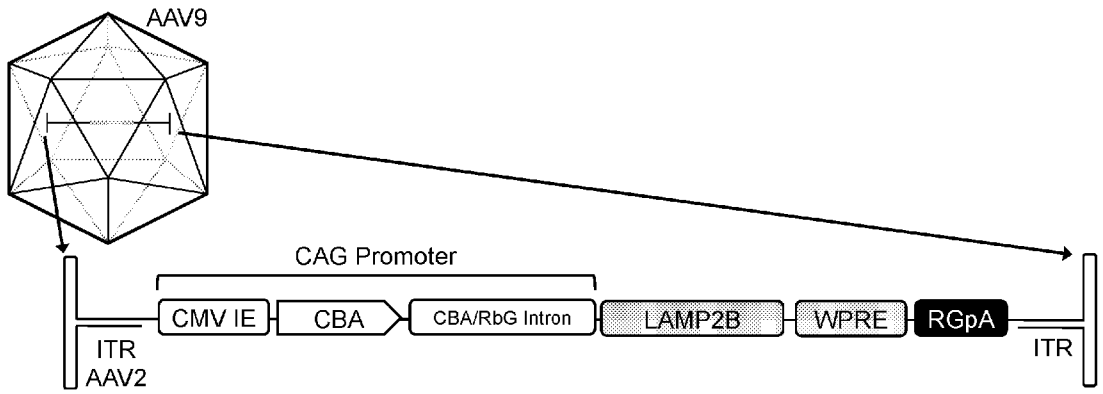

FIG. 1A

(57) Abstract: The invention relates generally to gene therapy for diseases associated with mutations in lysosome-associated membrane protein 2 (LAMP-2, also known as CD107b). In one aspect, the disclosure provides a gene therapy vector comprising an expression cassette comprising a transgene encoding an isoform of LAMP-2 or a functional variant thereof, wherein the transgene is codon-optimized for expression in a human host cell. In another aspect, the disclosure provides methods of preventing, mitigating, ameliorating, reducing, inhibiting, eliminating and/or reversing one or more symptoms of Danon disease or another autophagy disorder in a subject in need thereof, comprising administering to the subject any gene therapy vector of the disclosure.

WO 2020/014523 A1

[Continued on next page]

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

— *with international search report (Art. 21(3))*

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

GENE THERAPY VECTORS FOR TREATMENT OF DANON DISEASE

RELATED APPLICATIONS

This application claims priority to U.S. Provisional Patent Appl. No. 62/697,302, filed July 12, 2018, which is incorporated herein by reference in its entirety.

SEQUENCE LISTING

This application is being filed electronically via EFS-Web and includes an electronically submitted sequence listing in .txt format. The .txt file contains a sequence listing entitled "ROPA_011_01WO_SeqList_ST25.txt" created on July 11, 2019 and having a size of ~62 kilobytes. The sequence listing contained in this .txt file is part of the specification and is incorporated herein by reference in its entirety.

FIELD OF INVENTION

The invention relates generally to gene therapy for diseases associated with mutations in lysosome-associated membrane protein 2 (LAMP-2, also known as CD107b).

BACKGROUND

Lysosome-associated membrane protein 2 (LAMP-2, also known as CD107b) is a gene that encodes a lysosome-associated membrane glycoprotein. Alternative splicing of the gene produces three isoforms: LAMP-2A, LAMP-2B, and LAMP-2C. Loss-of-function mutations in LAMP-2 are associated with human diseases, including Danon disease, a familial cardiomyopathy associated with impaired autophagy.

International Patent Application Publication No. WO2017127565A1 discloses that overexpression of LAMP-2 in human induced pluripotent stem cells (hiPSCs) derived from patients with LAMP-2 mutations, as described in Hashem, et al., Stem Cells. 2015 Jul;33(7):2343-50, results in reduced oxidative stress levels and apoptotic cell death, confirming the importance of LAMP-2B in disease pathophysiology.

There remains a need in the art for gene therapy vectors for LAMP-2. The present disclosure provides such gene therapy vectors, methods of use thereof, pharmaceutical compositions, and more.

SUMMARY OF THE INVENTION

The present disclosure provides improved gene therapy vectors comprising a polynucleotide sequence encoding a LAMP-2 polypeptide, methods of use thereof, pharmaceutical compositions, and more.

In one aspect, the disclosure provides a gene therapy vector comprising an expression cassette comprising a transgene encoding an isoform of lysosome-associated membrane protein 2 (LAMP-2) or a functional variant thereof, wherein the transgene is codon-optimized for expression in a human host cell.

In an embodiment, the expression cassette contains fewer CpG sites than SEQ ID: 2.

In an embodiment, the expression cassette contains fewer cryptic splice sites than SEQ ID: 2.

In an embodiment, the expression cassette encodes fewer alternative open reading frames than SEQ ID: 2.

In an embodiment, the transgene shares at least 95% identity to a sequence selected from SEQ ID NOs: 3-5.

In an embodiment, the transgene shares at least 99% identity to a sequence selected from SEQ ID NOs: 3-5.

In an embodiment, the transgene comprises a sequence selected from SEQ ID NOs: 3-5.

In an embodiment, the transgene shares at least 95% identity to SEQ ID NO: 3.

In an embodiment, the transgene shares at least 99% identity to SEQ ID NO: 3.

In an embodiment, the transgene comprises a sequence identical to SEQ ID NO: 3.

In an embodiment, the expression cassette comprises a consensus optimal Kozak sequence operatively linked to the transgene, wherein optionally the consensus optimal Kozak sequence comprises SEQ ID NO: 6.

In an embodiment, the expression cassette comprises a full-length polyA sequence operatively linked to the transgene, wherein optionally the full-length polyA sequence comprises SEQ ID NO: 7.

In an embodiment, the expression cassette comprises no start site 5′ to the transgene capable of generating alternative mRNAs.

In an embodiment, the expression cassette comprises operatively linked, in the 5′ to 3′ direction, a first inverse terminal repeat, an enhancer/promoter region, a consensus optimal Kozak sequence, the transgene, a 3′ untranslated region including a full-length polyA sequence, and a second inverse terminal repeat.

In an embodiment, the enhancer/promoter region comprises in the 5′ to 3′ direction a CMV IE enhancer and a chicken beta-actin promoter.

In an embodiment, the expression cassette shares at least 95% identity to a sequence selected from SEQ ID NOs: 8-10.

In an embodiment, the expression cassette shares complete identity to a sequence selected from SEQ ID NOs: 8-10.

In a second aspect, the disclosure provides a method of preventing, mitigating, ameliorating, reducing, inhibiting, eliminating and/or reversing one or more symptoms of Danon disease or another autophagy disorder in a subject in need thereof, comprising administering to the subject any gene therapy vector of the disclosure.

In an embodiment, the vector is administered via a route selected from the group consisting of intravenous, intra-arterial, intracardiac, intracoronary, intramyocardial, intrarenal, intraurethral, epidural, and intramuscular.

In an embodiment, the autophagy disorder is selected from the group consisting of end-stage heart failure, myocardial infarction, drug toxicities, diabetes, end-stage renal failure, and aging.

In an embodiment, the subject is a human.

In an embodiment, the subject is exhibiting symptoms of Danon disease or another autophagy disorder.

In an embodiment, the subject has been identified as having reduced or non-detectable LAMP-2 expression.

In an embodiment, the subject has been identified as having a mutated LAMP-2 gene.

In a third aspect, the disclosure provides a pharmaceutical composition for use in preventing, mitigating, ameliorating, reducing, inhibiting, eliminating and/or reversing one or more symptoms of Danon disease or another autophagy disorder, comprising any gene therapy vector of the disclosure.

Other features and advantages of the invention will be apparent from and encompassed by the following detailed description and claims.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1A provides a diagram of an illustrative embodiment of a viral vector of the disclosure.

FIG. 1B provides a diagram of an illustrative embodiment of an expression cassette of an adeno-associated virus (AAV) gene therapy vector.

FIG. 2 shows the expression cassette of plasmid-based green fluorescence protein (GFP) reporter system used to test and compare wildtype verses codon-optimized LAMP-2B constructs.

FIG. 3 is a graph showing transfection expression efficiency of LAMP-2B constructs tested using transfection of a plasmid-based GFP reporter system and measured as GFP+ cells per well. A wild-type LAMP-2B construct (WT) is compared to three codon-optimized ("CO") constructs, CO 1, CO 2, and CO 3, and a no-vector control (labeled "ViaFect Only").

FIG. 4 is a graph showing the gene expression level in cells transfected with plasmid-based GFP reporting systems used to test LAMP-2B constructs, which is measured as mean GFP intensity of GFP+ cells in absorbance units (A.U.). GFP expression by cells transfected with a wild-type LAMP-2B construct (WT) is compared to GFP expression by three codon-optimized ("CO") constructs, CO 1, CO 2, and CO 3, or a no-vector control (labeled "ViaFect Only").

FIG. 5 shows immunofluorescence images of induced pluripotent stem cell (iPSC)-derived cardiomyocytes two days after transfection with a plasmid-based GFP reporter system. Cells were transfected with no DNA, or LAMP-2B constructs expressing wild-type LAMP-2B, the CO 1 variant, the CO 2 variant, or the CO 3 variant.

FIG. 6 shows immunofluorescence images of iPSC-derived cardiomyocytes seven days after transfection with a plasmid-based GFP reporter system. Cells were transfected with

no DNA, or LAMP-2B constructs expressing wild-type LAMP-2B, the CO 1 variant, the CO 2 variant, or the CO 3 variant.

FIG. 7A shows an immunoblot of human LAMP-2B protein in CHO-Lec2 cells transduced with AAV9 viral vectors comprising the wild-type LAMP-2B v1.0 transgene (AAV9 1.0), the optimized variant LAMP-2B v1.2 transgene (AAV9 1.2) or a GFP transgene (AAV9 GFP). Molecular weight markers (MW) and a control LAMP-2B recombinant protein sample (LAMP2B (+ve control)) were also included.

FIG. 7B shows quantification of LAMP-2B protein by ELISA in CHO-Lec2 cells transduced with AAV9-wild-type LAMP-2B (v1.0), AAV9-optimized LAMP-2B (v1.2) or AAV9-GFP (GFP) vectors.

FIG. 8A shows immunofluorescence images of Danon patient iPSC-derived cardiomyocytes transduced with the indicated amounts of AAV9-Luc (Luc), AAV9-wild-type LAMP-2B (LAMP2B v1.0) or AAV9-optimized LAMP-2B (LAMP2B v1.2) vectors.

FIG. 8B shows quantification of immunofluorescence of human LAMP-2B protein in Danon patient iPSC-derived cardiomyocytes transduced with AAV9-Luc, AAV9-wild-type LAMP-2B (v1.0) or AAV9-optimized LAMP-2B (v1.2) vectors.

FIG. 8C shows an immunoblot of human LAMP-2B protein in Danon patient iPSC-derived cardiomyocytes transduced with AAV9-Luc, AAV9-wild-type LAMP-2B (v1.0) or AAV9-optimized LAMP-2B (v1.2) vectors.

FIG. 9A shows PCR quantification of viral vector DNA in cardiac tissue isolated from LAMP-2-deficient mice treated with AAV9-wild-type LAMP-2B (v1.0), AAV9-optimized LAMP-2B (v1.2) or an AAV9 vehicle control. Vector copy number was quantified as VCN/Diploid Nucleus in the cardiac tissue. Control wild-type mice not injected with vector were included as controls (WT).

FIG. 9B shows quantitative RT-PCR analyses of transgene mRNA, measured by RT-PCR using probes specific for the WPRE element, in cardiac tissue isolated from LAMP-2-deficient mice treated with AAV9-wild-type LAMP-2B (v1.0), AAV9-optimized LAMP-2B (v1.2) or an AAV9 vehicle control (Vehicle). Expression of mRNA was quantified as vector genomes (vg) per μg total cellular RNA using a standard curve to convert copy number to vector genomes.

FIG. 9C shows an immunoblot of LAMP-2B protein in cardiac tissue isolated from LAMP-2-deficient mice treated with AA9-wild-type LAMP-2B (v1.0), AAV9-optimized LAMP-2B (v1.2) or the AAV9 vehicle control (Vehicle) compared to untreated wild-type mice (Untreated).

FIG. 9D shows immunofluorescence images of human LAMP-2B protein in cardiac tissue isolated from LAMP-2-deficient mice treated with AAV9-wild-type LAMP-2B (v1.0), AAV9-optimized LAMP-2B (v1.2) or the AAV9 vehicle control.

FIG. 10A shows PCR quantification of viral vector DNA in heart, muscle, liver and brain tissue isolated from primates treated with the AAV9-optimized human LAMP-2B (treated) vector or no vector vehicle control (untreated). Individuals are denoted as black or white squares.

FIG. 10B shows PCR quantification of viral vector DNA in cardiac chambers isolated from primates treated with AAV9-optimized human LAMP-2B vector (treated) or no vector vehicle control (untreated). Individuals are denoted as B059 (male, M), A991 (female, F), and A602 (male, M).

FIG. 10C shows quantitative RT-PCR analyses of transgene mRNA, measured by RT-PCR using probes specific for the WPRE element, in heart, muscle, liver and brain tissue isolated from primates treated with the AAV9-optimized human LAMP-2B vector (treated) or no vector vehicle control (untreated).

FIG. 10D shows quantitative RT-PCR analyses of transgene mRNA in cardiac chambers isolated from primates injected with the AAV9-optimized human LAMP-2B vector (treated) or no vector vehicle control (untreated).

FIG. 10E shows percentage of cells expressing transgene mRNA *in situ* in heart, muscle and liver tissue isolated from primates injected with AAV9-optimized human LAMP-2B vector (treated) or no vector vehicle control (untreated). Individuals are denoted as B059 (male, M), A991 (female, F), and A602 (male, M).

FIG. 10F shows transgene mRNA staining *in situ* in heart tissue isolated from primates injectioned with the AAV9-optimized human LAMP-2B vector or no vector vehicle control (untreated). Individuals are denoted as B059 (male, M), A991 (female, F), and A602 (male, M).

FIG. 10G shows fold change of LAMP-2B protein assessed by western blot in heart, muscle and liver tissue isolated from primates treated with the AAV9-optimized human LAMP-2B vector relative to no vector (untreated).

FIG. 10H shows fold change of LAMP-2B protein assessed by western blot in cardiac chambers isolated from primates treated with the AAV9-optimized human LAMP-2B vector relative to no vector (untreated).

FIG. 10I shows quantification of LAMP-2B protein by ELISA in heart, muscle and liver tissue isolated from primates treated with the AAV9-optimized human LAMP-2B vector relative to no vector (untreated).

FIG. 10J shows quantification of LAMP-2B protein by ELISA in cardiac chambers isolated from primates treated with the AAV9-optimized human LAMP-2B vector relative to no vector (untreated).

DETAILED DESCRIPTION OF THE INVENTION

The present disclosure provides improved polynucleotide sequences, expression cassettes, and vectors encoding an isoform of LAMP-2 (e.g., LAMP-2B), as well as related pharmaceutical compositions, and their use to treat diseases and disorders associated with LAMP-2 deficiency or mutation. The present inventors have discovered that modifications to the gene sequence of LAMP-2B result in increased transgene expression. In addition, the presence of specific sequence elements in the expression cassettes of gene therapy vectors encoding LAMP-2B result in an improvement in transgene expression. Accordingly, the LAMP-2 polynucleotide sequences, expression cassettes, and vectors disclosed herein offer advantages for gene therapy as compared to previous gene therapy vectors, including the ability to achieve higher levels of LAMP-2 expression in therapeutically relevant tissues.

The wild-type polypeptide sequence of human LAMP-2B (SEQ ID NO: 1) and the wild-type polynucleotide sequence encoding human LAMP-2B (SEQ ID NO: 2) are, respectively:

```
  1 MVCFRLFPVP GSGLVLVCLV LGAVRSYALE LNLTDSENAT CLYAKWQMNF TVRYETTNKT
 61 YKTVTISDHG TVTYNGSICG DDQNGPKIAV QFGPGFSWIA NFTKAASTYS IDSVSFSYNT
121 GDNTTFPDAE DKGILTVDEL LAIRIPLNDL FRCNSLSTLE KNDVVQHYWD VLVQAFVQNG
181 TVSTNEFLCD KDKTSTVAPT IHTTVPSPTT TPTPKEKPEA GTYSVNNGND TCLLATMGLQ
241 LNITQDKVAS VININPNTTH STGSCRSHTA LLRLNSSTIK YLDFVFAVKN ENRFYLKEVN
301 ISMYLVNGSV FSIANNNLSY WDAPLGSSYM CNKEQTVSVS GAFQINTFDL RVQPFNVTQG
361 KYSTAQECSL DDDTILIPII VGAGLSGLII VIVIAYVIGR RKSYAGYQT
```

(SEQ ID NO: 1); and

```
   1 ATGGTGTGCT TCCGCCTCTT CCCGGTTCCG GGCTCAGGGC TCGTTCTGGT CTGCCTAGTC
  61 CTGGGAGCTG TGCGGTCTTA TGCATTGGAA CTTAATTTGA CAGATTCAGA AAATGCCACT
 121 TGCCTTTATG CAAAATGGCA GATGAATTTC ACAGTTCGCT ATGAAACTAC AAATAAAACT
 181 TATAAAACTG TAACCATTTC AGACCATGGC ACTGTGACAT ATAATGGAAG CATTTGTGGG
 241 GATGATCAGA ATGGTCCCAA AATAGCAGTG CAGTTCGGAC CTGGCTTTTC CTGGATTGCG
 301 AATTTTACCA AGGCAGCATC TACTTATTCA ATTGACAGCG TCTCATTTTC CTACAACACT
 361 GGTGATAACA CAACATTTCC TGATGCTGAA GATAAAGGAA TTCTTACTGT TGATGAACTT
 421 TTGGCCATCA GAATTCCATT GAATGACCTT TTTAGATGCA ATAGTTTATC AACTTTGGAA
 481 AAGAATGATG TTGTCCAACA CTACTGGGAT GTTCTTGTAC AAGCTTTTGT CCAAAATGGC
 541 ACAGTGAGCA CAAATGAGTT CCTGTGTGAT AAAGACAAAA CTTCAACAGT GGCACCCACC
 601 ATACACACCA CTGTGCCATC TCCTACTACA ACACCTACTC CAAAGGAAAA ACCAGAAGCT
 661 GGAACCTATT CAGTTAATAA TGGCAATGAT ACTTGTCTGC TGGCTACCAT GGGGCTGCAG
 721 CTGAACATCA CTCAGGATAA GGTTGCTTCA GTTATTAACA TCAACCCCAA TACAACTCAC
 781 TCCACAGGCA GCTGCCGTTC TCACACTGCT CTACTTAGAC TCAATAGCAG CACCATTAAG
 841 TATCTAGACT TTGTCTTTGC TGTGAAAAAT GAAAACCGAT TTTATCTGAA GGAAGTGAAC
 901 ATCAGCATGT ATTTGGTTAA TGGCTCCGTT TTCAGCATTG CAAATAACAA TCTCAGCTAC
 961 TGGGATGCCC CCCTGGGAAG TTCTTATATG TGCAACAAAG AGCAGACTGT TTCAGTGTCT
1021 GGAGCATTTC AGATAAATAC CTTTGATCTA AGGGTTCAGC CTTTCAATGT GACACAAGGA
1081 AAGTATTCTA CAGCCCAAGA GTGTTCGCTG GATGATGACA CCATTCTAAT CCCAATTATA
1141 GTTGGTGCTG GTCTTTCAGG CTTGATTATC GTTATAGTGA TTGCTTACGT AATTGGCAGA
1201 AGAAAAAGTT ATGCTGGATA TCAGACTCTG TAA
```

(SEQ ID NO: 2).

Disclosed herein are modified polynucleotide sequences encoding an isoform of lysosome-associated membrane protein 2 (LAMP-2) or a functional variant thereof. In certain embodiments, the modified polynucleotide sequences comprise one or more of the following modifications as compared to the wild-type polynucleotide encoding the isoform of LAMP-2: codon-optimization, CpG depletion, removal of cryptic splice sites, or a reduced number of alternative open-reading frames (ORFs). In some embodiments, the modified polynucleotide encodes LAMP-2A, LAMP-2B, LAMP-2C or a functional variant of any of these isoforms. In embodiments, the disclosure provides a polynucleotide sequence or transgene encoding LAMP-2B or a functional variant thereof comprising one or more nucleotide substitutions as compared to SEQ ID NO:2. In embodiments, the transgene shares at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or complete identity to a sequence selected from SEQ ID NOs: 3-5. The disclosure provides at least three illustrative variant transgene sequences encoding LAMP-2B (SEQ ID NOs: 3-5):

```
   1 ATGGTCTGCT TCAGACTGTT CCCTGTCCCT GGATCTGGTC TGGTGCTTGT GTGCTTGGTG
  61 CTGGGTGCTG TGAGATCCTA TGCCCTTGAG CTGAACCTGA CTGACTCAGA AAATGCCACT
 121 TGCCTGTATG CCAAGTGGCA GATGAACTTC ACTGTGAGAT ATGAGACTAC CAACAAGACC
 181 TACAAGACTG TGACCATCTC AGACCATGGC ACTGTCACCT ACAATGGATC AATCTGTGGT
 241 GATGATCAGA ATGGCCCAAA GATAGCAGTG CAGTTTGGGC CCGGTTTTTC CTGGATTGCT
 301 AACTTCACCA AGGCAGCCTC CACCTACAGC ATTGACTCAG TCAGCTTCAG CTACAACACT
 361 GGGGATAACA CCACCTTCCC TGACGCAGAG GACAAGGGAA TCCTTACTGT GGACGAACTC
 421 CTGGCAATCA GAATCCCCCT TAACGACCTG TTCAGATGCA ACTCCCTTTC AACCCTTGAA
```

```
 481 AAGAATGATG TGGTGCAACA CTATTGGGAC GTCCTGGTGC AAGCCTTTGT GCAGAATGGG
 541 ACAGTGAGTA CCAACGAGTT CCTCTGTGAC AAGGACAAGA CCAGCACTGT GGCCCCCACT
 601 ATCCACACCA CTGTGCCCAG CCCTACCACT ACCCCCACCC CTAAAGAGAA GCCAGAAGCT
 661 GGAACCTACT CAGTCAACAA TGGAAATGAC ACATGCCTCC TTGCCACCAT GGGACTGCAG
 721 CTGAACATCA CTCAGGACAA GGTGGCCTCA GTGATTAACA TCAACCCTAA CACCACTCAT
 781 AGCACTGGGA GCTGCAGATC ACATACAGCT CTGCTGAGGC TCAACTCCTC CACCATCAAG
 841 TACCTGGACT TTGTGTTTGC TGTGAAGAAT GAGAACAGGT TCTACCTCAA GGAAGTGAAC
 901 ATTTCCATGT ACCTGGTCAA TGGTTCAGTG TTCTCTATTG CCAACAACAA TCTGAGCTAC
 961 TGGGATGCAC CCCTGGGATC CTCCTACATG TGCAACAAGG AGCAGACTGT GAGTGTGTCA
1021 GGTGCTTTTC AGATCAACAC TTTTGACCTG AGGGTGCAGC CCTTCAATGT GACTCAGGGA
1081 AAGTACTCCA CTGCACAAGA GTGTTCCTTG GATGATGACA CTATCCTCAT CCCCATTATT
1141 GTGGGAGCTG GACTGTCAGG ATTGATTATA GTGATTGTGA TTGCTTATGT GATTGGAAGG
1201 AGAAAGAGCT ATGCTGGCTA CCAGACCCTG TAA
```

(SEQ ID NO: 3);

```
   1 ATGGTGTGCT TTAGACTGTT TCCTGTGCCT GGTTCAGGGC TGGTCCTGGT CTGTCTGGTG
  61 CTGGGGGCTG TCAGAAGCTA TGCCTTGGAG CTGAACCTCA CTGATAGTGA AAATGCCACT
 121 TGTCTGTATG CTAAGTGGCA GATGAACTTC ACTGTGAGAT ATGAAACCAC CAACAAGACT
 181 TACAAAACAG TGACCATCTC AGATCATGGA ACTGTGACCT ACAACGGCAG CATTTGTGGA
 241 GACGACCAGA ACGGACCAAA AATCGCTGTC CAATTTGGGC CTGGATTCTC CTGGATTGCC
 301 AATTTCACTA AAGCTGCCTC CACATATTCA ATTGACTCAG TGTCCTTCTC CTACAACACT
 361 GGGGACAACA CTACTTTCCC TGATGCTGAA GATAAGGGAA TCTTGACAGT GGATGAGCTG
 421 CTGGCTATCA GGATCCCTTT GAATGACCTG TTTAGGTGTA ATTCACTGAG CACTCTGGAG
 481 AAGAACGACG TGGTGCAGCA CTACTGGGAC GTGCTGGTGC AGGCCTTTGT GCAGAACGGC
 541 ACTGTGTCCA CCAACGAATT CCTGTGTGAT AAGGACAAAA CTTCCACTGT GGCACCTACA
 601 ATTCACACTA CTGTGCCTTC ACCTACCACC ACTCCAACTC CAAAGGAAAA GCCTGAAGCA
 661 GGAACCTACT CTGTGAACAA TGGCAATGAT ACCTGTCTGT TGGCCACCAT GGGCCTCCAA
 721 CTGAACATTA CTCAGGACAA GGTGGCCTCA GTGATTAACA TTAACCCCAA CACTACCCAC
 781 TCCACTGGCA GCTGTAGATC ACACACAGCC TTGCTCAGAC TGAATAGCAG CACCATCAAG
 841 TATTTGGATT TTGTGTTTGC AGTGAAGAAT GAAAACAGGT TCTACCTGAA GGAAGTCAAC
 901 ATCTCAATGT ACCTGGTGAA CGGCTCAGTG TTCAGCATTG CCAACAACAA CCTCTCCTAT
 961 TGGGACGCTC CACTGGGGAG CAGCTACATG TGTAACAAGG AACAGACTGT GTCAGTGTCA
1021 GGAGCCTTCC AGATTAACAC CTTTGATCTG AGGGTCCAAC CCTTTAATGT CACTCAAGGA
1081 AAGTATAGCA CTGCCCAGGA GTGCTCCCTG GATGATGACA CCATTCTGAT TCCAATCATT
1141 GTGGGTGCAG GACTTTCTGG GCTTATTATT GTGATTGTGA TTGCCTATGT GATTGGCAGA
1201 AGGAAATCCT ATGCAGGGTA CCAAACTCTG TAA
```

(SEQ ID NO: 4); and

```
   1 ATGGTCTGTT TTAGGCTGTT CCCTGTCCCT GGTTCAGGAC TGGTCTTAGT GTGTCTGGTG
  61 CTTGGAGCTG TCAGAAGCTA TGCCCTGGAG CTGAACCTGA CTGACTCAGA AAATGCCACT
 121 TGCCTGTATG CCAAGTGGCA GATGAACTTC ACTGTCAGAT ATGAAACCAC CAACAAGACC
 181 TATAAGACTG TGACCATCTC AGACCATGGC ACTGTGACTT ACAATGGGTC AATTTGTGGA
 241 GATGACCAGA ATGGCCCTAA GATAGCTGTC CAGTTTGGTC CAGGATTCAG CTGGATTGCC
 301 AACTTCACCA AGGCAGCCAG CACCTACAGC ATTGACTCTG TGTCCTTCTC CTACAACACA
 361 GGAGACAACA CCACTTTCCC TGATGCAGAG GACAAAGGTA TCCTGACTGT GGATGAGTTG
 421 CTGGCAATCA GGATCCCACT GAACGATCTG TTCAGGTGCA ACTCACTGTC CACTCTGGAA
 481 AAGAATGATG TGGTGCAGCA CTATTGGGAT GTGCTAGTCC AGGCCTTTGT CCAGAATGGG
 541 ACTGTGTCAA CTAATGAGTT CCTGTGTGAC AAGGACAAGA CAAGCACTGT AGCCCCCACT
 601 ATCCATACCA CAGTACCTAG CCCCACCACT ACTCCAACCC CCAAGGAGAA GCCTGAGGCT
 661 GGCACCTACT CAGTGAACAA TGGGAATGAC ACCTGTTTGC TGGCCACTAT GGGACTCCAA
 721 CTGAACATCA CCCAGGACAA AGTGGCCTCT GTGATCAATA TCAATCCCAA CACCACCCAC
 781 AGCACTGGGT CCTGCAGAAG CCACACTGCC CTCCTGAGGC TCAACTCATC AACTATCAAG
 841 TACTTGGATT TTGTGTTTGC AGTGAAGAAT GAGAACAGAT TCTACCTCAA AGAGGTCAAC
 901 ATTTCAATGT ACCTGGTGAA TGGGAGTGTG TTCTCCATTG CTAACAACAA CCTGAGCTAC
 961 TGGGATGCCC CTCTGGGCTC CTCATACATG TGCAACAAGG AACAGACTGT GAGTGTGTCA
1021 GGGGCCTTCC AGATCAACAC TTTTGACCTG AGAGTGCAGC CCTTTAATGT GACACAGGGA
```

```
1081 AAGTACAGCA CTGCTCAGGA GTGCAGCCTG GATGATGACA CTATCCTGAT CCCTATCATT
1141 GTGGGGGCAG GCCTGTCTGG ACTCATTATT GTGATTGTGA TTGCCTATGT GATAGGGAGA
1201 AGGAAGTCTT ATGCTGGATA CCAGACCCTG TAA
```

(SEQ ID NO: 5).

In an embodiment, the transgene shares at least 95% identity to a sequence selected from SEQ ID NOs: 3-5. In an embodiment, the transgene shares at least 99% identity to a sequence selected from SEQ ID NOs: 3-5. In an embodiment, the transgene comprises a sequence selected from SEQ ID NOs: 3-5. In an embodiment, the transgene shares at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or complete identity to SEQ ID NO: 3. In an embodiment, the transgene shares at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or complete identity to SEQ ID NO: 4. In an embodiment, the transgene shares at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or complete identity to SEQ ID NO: 5.

In some embodiments, the transgene is similar to or identical to a subsequence of any one of SEQ ID NOs: 3-5. In some embodiments, the transgene comprises a subsequence of any one of SEQ ID NOs: 3-5. In various embodiments, the subsequence may comprise any set of consecutive nucleotides (nt) in the full sequence having a length of at least about 50 nt, at least about 100 nt, at least about 150 nt, at least about 250 nt, at least about 200 nt, at least about 350 nt, at least about 450 nt, at least about 400 nt, at least about 450 nt, at least about 550 nt, at least about 600 nt, at least about 650 nt, at least about 600 nt, at least about 650 nt, at least about 700 nt, at least about 750 nt, at least about 800 nt, at least about 850 nt, at least about 900 nt, at least about 950 nt, or at least abou 1000 nt.

In some embodiments, the transgene shares at least 95% identity to a subsequence that comprises nucleotides 1-500, 250-750, 500-1000, or 750-1240 of any one of SEQ ID NO: 3-5. In an embodiment, the transgene shares at least 99% identity to a subsequence that comprises nucleotides 1-500, 250-750, 500-1000, or 750-1240 of any one of SEQ ID NO: 3-5. In an embodiment, the transgene comprises a sequence that comprises nucleotides 1-500, 250-750, 500-1000, or 750-1240 of any one of SEQ ID NOs: 3-5. In embodiment, the transgene shares at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or complete identity to a subsequence that comprises nucleotides 1-500, 250-750, 500-1000, or 750-1240 of any one of SEQ ID NOs: 3-5. In embodiments, the

transgene shares at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or complete identity to a subsequence that comprises nucleotides 1-500, 250-750, 500-1000, or 750-1240 of SEQ ID NO: 3. In embodiment, the transgene shares at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or complete identity to a subsequence that comprises nucleotides 1-500, 250-750, 500-1000, or 750-1240 of SEQ ID NO: 3.

In certain embodiments, the transgene encodes any of the various isoforms of LAMP-2, including any of LAMP-2A, LAMP-2B, or LAMP-2C, or a functional fragment or variant of any of these isoforms. Thus, in particular embodiments, the expression cassette is an optimized polynucleotide sequence encoding any of LAMP-2A, LAMP-2B, or LAMP-2C, or a functional fragment or variant thereof, which comprises one or more modifications as compared to the corresponding wild-type polynucleotide sequence, including one or more modification selected from: codon-optimization of the transgene sequence encoding LAMP-2A, LAMP-2B, or LAMP-2C; the expression cassette or transgene sequence contains fewer CpG sites than its corresponding wild-type sequence; the expression cassette or transgene sequence contains fewer CpG sites than its corresponding wild-type sequence; the expression cassette or transgene sequence contains fewer cryptic splice sites than its corresponding wild-type sequence; and/or the expression cassette or transgene sequence contains fewer open reading frames than its corresponding wild-type sequence. In particular embodiments, the optimized sequence is optimized for increased expression in human cells. The wild-type human polynucleotide sequences encoding the LAMP-2A and LAMP-2C isoforms are set forth in SEQ ID NOs: 29 and 30, respectively. The wild-type sequences of human LAMP-2A and LAMP-2C proteins are set forth in SEQ ID NOs: 34 and 35, respectively. The sequences of the wild-type LAMP-2 isoforms and coding sequences are also publicly available. While the specification describes specific embodiments with respect to LAMP-2B, it is understood that LAMP-2A or LAMP-2C could alternatively be used in each embodiment.

The coding sequences of wild-type LAMP-2A (SEQ ID NO: 29) and wild-type LAMP-2C (SEQ ID NO: 30) are 100% identical to the coding sequence of wild-type LAMP-2B (SEQ ID NO: 2) across at least nucleotides 1 – 1080. Accordingly, it will be readily recognized by those in the art that that transgenes, expression cassettes, and vectors disclosed herein can be adapted for expression of these isoforms of LAMP-2 by substituting the 3′ end

(nucleotides 1081 – end) of either of LAMP-2A (SEQ ID NO: 29) or wild-type LAMP-2C (SEQ ID NO: 30) in place of nucleotides 1081 – 1233 of LAMP-2B (e.g., an optimized LAMP-2B of any of SEQ ID NO: 3-5). For example, embodiments of the invention utilize nucleotides 1-1080 of the optimized LAMP-2B gene sequences, SEQ ID NOs: 3-5, which are, respectively, SEQ ID NOs: 31-33.

In an embodiment, the transgene shares at least 95% identity to a sequence selected from SEQ ID NOs: 31-33. In an embodiment, the transgene shares at least 99% identity to a sequence selected from SEQ ID NOs: 31-33. In an embodiment, the transgene comprises a sequence selected from SEQ ID NOs: 31-33. In an embodiment, the transgene shares at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or complete identity to SEQ ID NO: 31. In an embodiment, the transgene shares at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or complete identity to SEQ ID NO: 32. In an embodiment, the transgene shares at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or complete identity to SEQ ID NO: 33.In some cases, the transgene has a polynucleotide sequence that is different from the polynucleotide sequence of a reference sequence, *e.g.*, a "native" or "wild-type" LAMP-2B sequence. In some embodiments, the transgene shares at most 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, or 95% identity with a reference sequence. In some embodiments, the reference sequence is SEQ ID NO: 2. For example, SEQ ID NO: 3 shares 78.5% identity to SEQ ID NO: 2.

In some cases, the transgene has a polynucleotide sequence that is different from the polynucleotide sequence of a reference sequence, *e.g.*, a "native" or "wild-type" LAMP-2A sequence. In some embodiments, the transgene shares at most 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, or 95% identity with a reference sequence. In some embodiments, the reference sequence is the wild-type human LAMP-2A sequence set forth in SEQ ID NO: 29.

In some cases, the transgene has a polynucleotide sequence that is different from the polynucleotide sequence of a reference sequence, *e.g.*, a "native" or "wild-

type" LAMP-2C sequence. In some embodiments, the transgene shares at most 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, or 95% identity with a reference sequence. In some embodiments, the reference sequence is the wild-type human LAMP-2A sequence set forth in SEQ ID NO: 30.

In an embodiment, the transgene is codon-optimized for expression in a human host cell. In an embodiment, the transgene coding sequence is modified, or "codon optimized" to enhance expression by replacing infrequently represented codons with more frequently represented codons. The coding sequence is the portion of the mRNA sequence that encodes the amino acids for translation. During translation, each of 61 trinucleotide codons are translated to one of 20 amino acids, leading to a degeneracy, or redundancy, in the genetic code. However, different cell types, and different animal species, utilize tRNAs (each bearing an anticodon) coding for the same amino acids at different frequencies. When a gene sequence contains codons that are infrequently represented by the corresponding tRNA, the ribosome translation machinery may slow, impeding efficient translation. Expression can be improved via "codon optimization" for a particular species, where the coding sequence is altered to encode the same protein sequence, but utilizing codons that are highly represented, and/or utilized by highly expressed human proteins (Cid-Arregui et al., 2003; J. Virol. 77: 4928).

In some embodiments, the coding sequence of the transgene is modified to replace codons infrequently expressed in mammal or in primates with codons frequently expressed in primates. For example, in some embodiments, the transgene encodes a polypeptide having at least 85% sequence identity to a reference polypeptide (*e.g.* wild-type LAMP-2B; SEQ ID NO: 3)—for example, at least 90% sequence identity, at least 95% sequence identity, at least 98% identity, or at least 99% identity to the reference polypeptide—wherein at least one codon of the coding sequence has a higher tRNA frequency in humans than the corresponding codon in the sequence disclosed above or herein.

In an embodiment, the transgene comprises fewer alternative open reading frames than SEQ ID: 2. In an embodiment, the transgene is modified to enhance expression by

termination or removal of open reading frames (ORFs) that do not encode the desired transgene. An open reading frame (ORF) is the nucleic acid sequence that follows a start codon and does not contain a stop codon. ORFs may be in the forward or reverse orientation, and may be "in frame" or "out of frame" compared with the gene of interest. Such open reading frames have the potential to be expressed in an expression cassette alongside the gene of interest, and could lead to undesired adverse effects. In some embodiments the transgene has been modified to remove open reading frames by further altering codon usage. This may be done by eliminating one or more start codons (ATG) and/or introducing one or more stop codons (TAG, TAA, or TGA) in reverse orientation or out-of-frame to the desired ORF, while preserving the encoded amino acid sequence and, optionally, maintaining highly utilized codons in the gene of interest (*i.e.*, avoiding codons with frequency < 20%).

In some embodiments, the expression cassette comprises at most one, at most two, at most three, at most four, or at most five start codons 5′ to the start codon of the transgene. In some embodiments, the expression cassette comprises no start codon 5′ to the start codon of the transgene. In some embodiments, one or more ATG codons in the 5′ UTR, the promoter, the enhance, the promoter/enhancer element, or other sequences 5′ to the start codon of the transgene remain after one or more cryptic start sites are removed. In some embodiments, the expression cassette comprises no cryptic starts sites upstream of transgene to generate erroneous mRNAs.

In variations of the present disclosure, the transgene coding sequence may be optimized by either codon optimization or removal of non-transgene ORFs or using both techniques. In some cases, one removes or minimizes non-transgene ORFs after codon optimization in order to remove ORFs introduced during codon optimization.

In an embodiment, the transgene contains fewer CpG sites than SEQ ID: 2. Without being bound by theory, it is believed that the presence of CpG sites in a polynucleotide sequence is associated with the undesirable immunological responses of the host against a viral vector comprising the polynucleotide sequence. In some embodiments, the transgene is designed to reduce the number of CpG sites. Exemplary methods are provides in U.S. Patent Application Publication No. US20020065236A1.

In an embodiment, the transgene contains fewer cryptic splice sites than SEQ ID: 2. For the optimization, GeneArt® software may be used, *e.g.*, to increase the GC content

and/or remove cryptic splice sites in order to avoid transcriptional silencing and, therefore, increase transgene expression. Alternatively, any optimization method known in the art may be used. Removal of cryptic splice sites is described, for example, in International Patent Application Publication No. WO2004015106A1.

Also disclosed herein are expression cassettes and gene therapy vectors encoding LAMP-2B. In certain embodiments, the expression cassettes and gene therapy vectors comprise a codon-optimized or variant LAMP-2B polynucleotide sequence or transgene sequence disclosed herein.

In particular embodiments, an expression cassette or gene therapy vector encoding LAMP-2B comprises: a consensus optimal Kozak sequence, a full-length polyadenylation (polyA) sequence (or substitution of full-length polyA by a truncated polyA), and minimal or no upstream (*i.e.* 5′) or cryptic start codons (*i.e.* ATG sites). In some embodiments, the expression cassette comprises no start site 5′ to the transgene capable of generating alternative mRNAs. In certain embodiments, the expression cassette or gene therapy vector comprises a sequence encoding LMAP-2B, *e.g.*, a codon-optimized or variant LAMP-2B polynucleotide sequence or transgene sequence disclosed herein.

In some cases, the expression cassette contains two or more of a first inverted terminal repeat, an enhancer/promoter region, a consensus optimal Kozak sequence, a transgene (*e.g.*, a transgene encoding a LAMP-2B disclosed herein), a 3′ untranslated region including a full-length polyA sequence, and a second inverted terminal repeat. In some embodiments, one or both of the inverted terminal repeats (ITRs) are AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, or AAV9 ITRs, or any one ITR known in the art. In some embodiments, the expression cassette comprises exactly two ITRs. In some embodiments, both ITRs are AAV2, AAV5, or AAV9 ITRs. In some embodiments, both ITRs are AAV2 ITRs.

In an embodiment, the expression cassette comprises a Kozak sequence operatively linked to the transgene. In an embodiment, the Kozak sequence is a consensus optimal Kozak sequence comprising or consisting of SEQ ID NO: 6:

GCCGCCACCATGG (SEQ ID NO: 6).

In various embodiments, the expression cassette comprises an alternative Kozak sequence operatively linked to the transgene. In an embodiment, the Kozak

sequence is an alternative Kozak sequence comprising or consisting of any one of SEQ ID NOs. 14-18:

(gcc)gccRccAUGG (SEQ ID NO: 14);

AGNNAUGN (SEQ ID NO: 15);

ANNAUGG (SEQ ID NO: 16);

ACCAUGG (SEQ ID NO: 17);

GACACCAUGG (SEQ ID NO: 18).

In some embodiments, the expression cassette comprises no Kozak sequence.

In SEQ ID NO: 14, a lower-case letter denotes the most common base at a position where the base can nevertheless vary; an upper-case letter indicate a highly conserved base; indicates adenine or guanine. In SEQ ID NO: 14, the sequence in parentheses (gcc) is optional. In SEQ ID NOs: 15-17, 'N' denotes any base.

A variety of sequences can be used in place of this consensus optimal Kozak sequence as the translation-initiation site and it is within the skill of those in the art to identify and test other sequences. *See* Kozak M. An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.* 115 (4): 887–903 (1991).

In an embodiment, the expression cassette comprises a full-length polyA sequence operatively linked to the transgene. In an embodiment, the full-length polyA sequence comprises SEQ ID NO: 7:

```
  1 TGGCTAATAA AGGAAATTTA TTTTCATTGC AATAGTGTGT TGGAATTTTT TGTGTCTCTC
 61 ACTCGGAAGG ACATATGGGA GGGCAAATCA TTTAAAACAT CAGAATGAGT ATTTGGTTTA
121 GAGTTTGGCA ACATATGCCC ATATGCTGGC TGCCATGAAC AAAGGTTGGC TATAAAGAGG
181 TCATCAGTAT ATGAAACAGC CCCCTGCTGT CCATTCCTTA TTCCATAGAA AAGCCTTGAC
241 TTGAGGTTAG ATTTTTTTTA TATTTTGTTT TGTGTTATTT TTTTCTTTAA CATCCCTAAA
301 ATTTTCCTTA CATGTTTTAC TAGCCAGATT TTTCCTCCTC TCCTGACTAC TCCCAGTCAT
361 AGCTGTCCCT CTTCTCTTAT GGAGATC
```

(SEQ ID NO: 7).

Various alternative polyA sequences may be used in expression cassettes of the present disclosure, including without limitation, bovine growth hormone polyadenylation signal (bGHpA) (SEQ ID NO: 19), the SV40 early/late polyadenylation signal (SEQ ID NO: 20), and human growth hormone (HGH) polyadenylation signal (SEQ ID NO: 21):

```
  1 TCGACTGTGC CTTCTAGTTG CCAGCCATCT GTTGTTTGCC CCTCCCCCGT GCCTTCCTTG
 61 ACCCTGGAAG GTGCCACTCC CACTGTCCTT TCCTAATAAA ATGAGGAAAT TGCATCGCAT
```

```
121 TGTCTGAGTA GGTGTCATTC TATTCTGGGG GGTGGGGTGG GGCAGGACAG CAAGGGGGAG
181 GATTGGGAGG ACAATAGCAG GCATGCTGGG GATGCGGTGG GCTCTATGGC TTCTG
```

(SEQ ID NO: 19);

```
  1 CAGACATGAT AAGATACATT GATGAGTTTG GACAAACCAC AACTAGAATG CAGTGAAAAA
 61 AATGCTTTAT TTGTGAAATT TGTGATGCTA TTGCTTTATT TGTAACCATT ATAAGCTGCA
121 ATAAACAAGT TAACAACAAC AATTGCATTC ATTTTATGTT TCAGGTTCAG GGGGAGATGT
181 GGGAGGTTTT TTAAAGCAAG TAAAACCTCT ACAAATGTGG TA
```

(SEQ ID NO: 20);

```
  1 CTGCCCGGGT GGCATCCCTG TGACCCCTCC CCAGTGCCTC TCCTGGCCCT GGAAGTTGCC
 61 ACTCCAGTGC CCACCAGCCT TGTCCTAATA AAATTAAGTT GCATCATTTT GTCTGACTAG
121 GTGTCCTTCT ATAATATTAT GGGGTGGAGG GGGGTGGTAT GGAGCAAGGG GCCCAAGTTG
181 GGAAGAAACC TGTAGGGCCT GC
```

(SEQ ID NO: 21).

In some embodiments, the expression cassette comprises an active fragment of a polyA sequence. In particular embodiments, the active fragment of the polyA sequence comprises or consists of less than 20 base pair (bp), less than 50 bp, less than 100 bp, or less than 150 bp, *e.g.*, of any of the polyA sequences disclosed herein.

In some cases, expression of the transgene is increased by ensuring that the expression cassette does not contain competing ORFs. In an embodiment, the expression cassette comprises no start codon within 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, or 500 base pairs 5′ of the start codon of the transgene. In some embodiment, the expression cassette comprises no start codon 5′ of the start codon of the transgene. In some embodiments, the expression cassette comprises no start site 5′ to the transgene capable of generating alternative mRNAs.

In an embodiment, the expression cassette comprises operatively linked, in the 5′ to 3′ direction, a first inverted terminal repeat, an enhancer/promoter region, introns, a consensus optimal Kozak sequence, the transgene, a 3′ untranslated region including a full-length polyA sequence, and a second inverted terminal repeat, wherein the expression cassette comprises no start site 5′ to the transgene capable of generating alternative mRNAs.

In some embodiments, the enhancer/promoter region comprises, in the 5′ to 3′ direction: a CMV IE enhancer and a chicken beta-actin promoter. In an embodiment, the enhancer/promoter region comprises a CAG promoter (SEQ ID NO: 22). As used herein "CAG promoter" refers to a polynucleotide sequence comprising a CMV early enhancer

element, a chicken beta-actin promoter, the first exon and first intron of the chicken beta-actin gene, and a splice acceptor from the rabbit beta-globin gene.

```
   1 CTAGTCGACA TTGATTATTG ACTAGTTATT AATAGTAATC AATTACGGGG TCATTAGTTC
  61 ATAGCCCATA TATGGAGTTC CGCGTTACAT AACTTACGGT AAATGGCCCG CCTGGCTGAC
 121 CGCCCAACGA CCCCCGCCCA TTGACGTCAA TAATGACGTA TGTTCCCATA GTAACGCCAA
 181 TAGGGACTTT CCATTGACGT CAATGGGTGG AGTATTTACG GTAAACTGCC CACTTGGCAG
 241 TACATCAAGT GTATCATATG CCAAGTACGC CCCCTATTGA CGTCAATGAC GGTAAATGGC
 301 CCGCCTGGCA TTATGCCCAG TACATGACCT TATGGGACTT TCCTACTTGG CAGTACATCT
 361 ACGTATTAGT CATCGCTATT ACCATGGTCG AGGTGAGCCC CACGTTCTGC TTCACTCTCC
 421 CCATCTCCCC CCCCTCCCCA CCCCCAATTT TGTATTTATT TATTTTTTAA TTATTTTGTG
 481 CAGCGATGGG GGCGGGGGGG GGGGGGGGGC GCGCGCCAGG CGGGGCGGGG CGGGGCGAGG
 541 GGCGGGGCGG GGCGAGGCGG AGAGGTGCGG CGGCAGCCAA TCAGAGCGGC GCGCTCCGAA
 601 AGTTTCCTTT TATGGCGAGG CGGCGGCGGC GGCGGCCCTA TAAAAAGCGA AGCGCGCGGC
 661 GGGCGGGAGT CGCTGCGCGC TGCCTTCGCC CCGTGCCCCG CTCCGCCGCC GCCTCGCGCC
 721 GCCCGCCCCG GCTCTGACTG ACCGCGTTAC TCCCACAGGT GAGCGGGCGG GACGGCCCTT
 781 CTCCTCCGGG CTGTAATTAG CGCTTGGTTT AATGACGGCT TGTTTCTTTT CTGTGGCTGC
 841 GTGAAAGCCT TGAGGGGCTC CGGGAGGGCC CTTTGTGCGG GGGGAGCGGC TCGGGGGGTG
 901 CGTGCGTGTG TGTGTGCGTG GGGAGCGCCG CGTGCGGCTC CGCGCTGCCC GGCGGCTGTG
 961 AGCGCTGCGG GCGCGGCGCG GGGCTTTGTG CGCTCCGCAG TGTGCGCGAG GGGAGCGCGG
1021 CCGGGGGCGG TGCCCCGCGG TGCGGGGGGG GCTGCGAGGG GAACAAAGGC TGCGTGCGGG
1081 GTGTGTGCGT GGGGGGGTGA GCAGGGGGTG TGGGCGCGTC GGTCGGGCTG CAACCCCCCC
1141 TGCACCCCCC TCCCCGAGTT GCTGAGCACG GCCCGGCTTC GGGTGCGGGG CTCCGTACGG
1201 GGCGTGGCGC GGGGCTCGCC GTGCCGGGCG GGGGGTGGCG GCAGGTGGGG GTGCCGGGCG
1261 GGGCGGGGCC GCCTCGGGCC GGGGAGGGCT CGGGGGAGGG GCGCGGCGGC CCCCGGAGCG
1321 CCGGCGGCTG TCGAGGCGCG GCGAGCCGCA GCCATTGCCT TTTATGGTAA TCGTGCGAGA
1381 GGGCGCAGGG ACTTCCTTTG TCCCAAATCT GTGCGGAGCC GAAATCTGGG AGGCGCCGCC
1441 GCACCCCCTC TAGCGGGCGC GGGGCGAAGC GGTGCGGCGC CGGCAGGAAG GAAATGGGCG
1501 GGGAGGGCCT TCGTGCGTCG CCGCGCCGCC GTCCCCTTCT CCCTCTCCAG CCTCGGGGCT
1561 GTCCGCGGGG GGACGGCTGC CTTCGGGGGG GACGGGGCAG GGCGGGGTTC GGCTTCTGGC
1621 GTGTGACCGG CGGCTCTAGA GCCTCTGCTA ACCATGTTCA TGCCTTCTTC TTTTTCCTAC
1681 AGCTCCTGGG CAACGTGCTG GTTATTGTGC TGTCTCATCA TTTTGGCAAA
```

(SEQ ID NO: 22).

In some embodiments, the enhancer/promoter region comprises a ubiquitous promoter. In some embodiments, the enhancer/promoter region comprises a CMV promoter (SEQ. ID NO: 23), an SV40 promoter (SEQ ID NO: 24), a PGK promoter (SEQ ID NO: 25), and/or a human beta-actin promoter (SEQ ID NO: 26). In some embodiments, the enhancer/promoter region comprises a polynucleotide that shares at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with any one of SEQ ID NOs: 23-26:

```
   1 GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC ACGGGGATTT
  61 CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTTTT GGCACCAAAA TCAACGGGAC
 121 TTTCCAAAAT GTCGTAACAA CTCCGCCCCA TTGACGCAAA TGGGCGGTAG GCGTGTACGG
 181 TGGGAGGTCT ATATAAGCAG AGCT
```

(SEQ ID NO: 23);

```
   1 GGTGTGGAAA GTCCCCAGGC TCCCCAGCAG GCAGAAGTAT GCAAAGCATG CATCTCAATT
  61 AGTCAGCAAC CAGGTGTGGA AAGTCCCCAG GCTCCCCAGC AGGCAGAAGT ATGCAAAGCA
 121 TGCATCTCAA TTAGTCAGCA ACCATAGTCC CGCCCCTAAC TCCGCCCATC CCGCCCCTAA
 181 CTCCGCCCAG TTCCGCCCAT TCTCCGCCCC ATGGCTGACT AATTTTTTTT ATTTATGCAG
 241 AGGCCGAGGC CGCCTCGGCC TCTGAGCTAT TCCAGAAGTA GTGAGGAGGC TTTTTTGGAG
 301 GCCTAGGCTT TTGCAAA
```

(SEQ ID NO: 24);

```
   1 GGGTAGGGGA GGCGCTTTTC CCAAGGCAGT CTGGAGCATG CGCTTTAGCA GCCCCGCTGG
  61 GCACTTGGCG CTACACAAGT GGCCTCTGGC CTCGCACACA TTCCACATCC ACCGGTAGGC
 121 GCCAACCGGC TCCGTTCTTT GGTGGCCCCT TCGCGCCACC TTCTACTCCT CCCCTAGTCA
 181 GGAAGTTCCC CCCCGCCCCG CAGCTCGCGT CGTGCAGGAC GTGACAAATG GAAGTAGCAC
 241 GTCTCACTAG TCTCGTGCAG ATGGACAGCA CCGCTGAGCA ATGGAAGCGG GTAGGCCTTT
 301 GGGGCAGCGG CCAATAGCAG CTTTGCTCCT TCGCTTTCTG GGCTCAGAGG CTGGGAAGGG
 361 GTGGGTCCGG GGGCGGGCTC AGGGGCGGGC TCAGGGGCGG GGCGGGCGCC CGAAGGTCCT
 421 CCGGAGGCCC GGCATTCTGC ACGCTTCAAA AGCGCACGTC TGCCGCGCTG TTCTCCTCTT
 481 CCTCATCTCC GGGCCTTTCG
```

(SEQ ID NO: 25);

```
   1 CCTGCAGGGC CCACTAGTTC CATGTCCTTA TATGGACTCA TCTTTGCCTA TTGCGACACA
  61 CACTCAATGA ACACCTACTA CGCGCTGCAA AGAGCCCCGC AGGCCTGAGG TGCCCCCACC
 121 TCACCACTCT TCCTATTTTT GTGTAAAAAT CCAGCTTCTT GTCACCACCT CCAAGGAGGG
 181 GGAGGAGGAG GAAGGCAGGT TCCTCTAGGC TGAGCCGAAT GCCCCTCTGT GGTCCCACGC
 241 CACTGATCGC TGCATGCCCA CCACCTGGGT ACACACAGTC TGTGATTCCC GGAGCAGAAC
 301 GGACCCTGCC CACCCGGTCT TGTGTGCTAC TCAGTGGACA GACCCAAGGC AAGAAAGGGT
 361 GACAAGGACA GGGTCTTCCC AGGCTGGCTT TGAGTTCCTA GCACCGCCCC GCCCCCAATC
 421 CTCTGTGGCA CATGGAGTCT TGGTCCCCAG AGTCCCCCAG CGGCCTCCAG ATGGTCTGGG
 481 AGGGCAGTTC AGCTGTGGCT GCGCATAGCA GACATACAAC GGACGGTGGG CCCAGACCCA
 541 GGCTGTGTAG ACCCAGCCCC CCCGCCCCGC AGTGCCTAGG TCACCCACTA ACGCCCCAGG
 601 CCTGGTCTTG GCTGGGCGTG ACTGTTACCC TCAAAAGCAG GCAGCTCCAG GGTAAAAGGT
 661 GCCCTGCCCT GTAGAGCCCA CCTTCCTTCC CAGGGCTGCG GCTGGGTAGG TTTGTAGCCT
 721 TCATCACGGG CCACCTCCAG CCACTGGACC GCTGGCCCCT GCCCTGTCCT GGGGAGTGTG
 781 GTCCTGCGAC TTCTAAGTGG CCGCAAGCCA CCTGACTCCC CCAACACCAC ACTCTACCTC
 841 TCAAGCCCAG GTCTCTCCCT AGTGACCCAC CCAGCACATT TAGCTAGCTG AGCCCCACAG
 901 CCAGAGGTCC TCAGGCCCTG CTTTCAGGGC AGTTGCTCTG AAGTCGGCAA GGGGGAGTGA
 961 CTGCCTGGCC ACTCCATGCC CTCCAAGAGC TCCTTCTGCA GGAGCGTACA GAACCCAGGG
1021 CCCTGGCACC CGTGCAGACC CTGGCCCACC CCACCTGGGC GCTCAGTGCC CAAGAGATGT
1081 CCACACCTAG GATGTCCCGC GGTGGGTGGG GGGCCCGAGA GACGGGCAGG CCGGGGGCAG
1141 GCCTGGCCAT GCGGGGCCGA ACCGGGCACT GCCCAGCGTG GGGCGCGGGG GCCACGGCGC
1201 GCGCCCCCAG CCCCCGGGCC CAGCACCCCA AGGCGGCCAA CGCCAAAACT CTCCCTCCTC
1261 CTCTTCCTCA ATCTCGCTCT CGCTCTTTTT TTTTTTCGCA AAAGGAGGGG AGAGGGGGTA
1321 AAAAAATGCT GCACTGTGCG GCGAAGCCGG TGAGTGAGCG GCGCGGGGCC AATCAGCGTG
1381 CGCCGTTCCG AAAGTTGCCT TTTATGGCTC GAGCGGCCGC GGCGGCGCCC TATAAAACCC
1441 AGCGGCGCGA CGCGCCACCA CCGCCGAGAC CGCGTCCGCC CCGCGAGCAC AGAGCCTCGC
1501 CTTTGCCGAT CCGCCGCCCG TCCACACCCG CCGCCAGGTA AGCCCGGCCA GCCGACCGGG
1561 GCAGGCGGCT CACGGCCCGG CCGCAGGCGG CCGCGGCCCC TTCGCCCGTG CAGAGCCGCC
1621 GTCTGGGCCG CAGCGGGGGG CGCATGGGGG GGGAACCGGA CCGCCGTGGG GGGCGCGGGA
1681 GAAGCCCCTG GGCCTCCGGA GATGGGGGAC ACCCCACGCC AGTTCGGAGG CGCGAGGCCG
1741 CGCTCGGGAG GCGCGCTCCG GGGGTGCCGC TCTCGGGGCG GGGGCAACCG GCGGGGTCTT
1801 TGTCTGAGCC GGGCTCTTGC CAATGGGGAT CGCAGGGTGG GCGCGGCGGA GCCCCCGCCA
1861 GGCCCGGTGG GGGCTGGGGC GCCATTGCGC GTGCGCGCTG GTCCTTTGGG CGCTAACTGC
1921 GTGCGCGCTG GGAATTGGCG CTAATTGCGC GTGCGCGCTG GGACTCAAGG CGCTAACTGC
1981 GCGTGCGTTC TGGGGCCCGG GGTGCCGCGG CCTGGGCTGG GGCGAAGGCG GGCTCGGCCG
2041 GAAGGGGTGG GGTCGCCGCG GCTCCCGGGC GCTTGCGCGC ACTTCCTGCC CGAGCCGCTG
```

```
2101 GCCGCCCGAG GGTGTGGCCG CTGCGTGCGC GCGCGCCGAC CCGGCGCTGT TTGAACCGGG
2161 CGGAGGCGGG GCTGGCGCCC GGTTGGGAGG GGGTTGGGGC CTGGCTTCCT GCCGCGCGCC
2221 GCGGGGACGC CTCCGACCAG TGTTTGCCTT TTATGGTAAT AACGCGGCCG GCCCGGCTTC
2281 CTTTGTCCCC AATCTGGGCG CGCGCCGGCG CCCCCTGGCG GCCTAAGGAC TCGGCGCGCC
2341 GGAAGTGGCC AGGGCGGGGG CGACCTCGGC TCACAGCGCG CCCGGCTATT CTCGCAGCTC
2401 ACC
```

(SEQ ID NO: 26).

Further exemplary promoters include, but are not limited to, human Elongation Factor 1 alpha promoter (EFS), SV40 early promoter, mouse mammary tumor virus long terminal repeat (LTR) promoter; adenovirus major late promoter (Ad MLP); a herpes simplex virus (HSV) promoter, an endogenous cellular promoter that is heterologous to the gene of interest, a cytomegalovirus (CMV) promoter such as the CMV immediate early promoter region (CMVIE), a Rous sarcoma virus (RSV) promoter, synthetic promoters, hybrid promoters, and the like

In some embodiments, the 3′ UTR comprises a polynucleotide (WPRE element) that shares at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with SEQ ID NO: 27:

```
   1 ATTCGAGCAT CTTACCGCCA TTTATTCCCA TATTTGTTCT GTTTTTCTTG ATTTGGGTAT
  61 ACATTTAAAT GTTAATAAAA CAAAATGGTG GGGCAATCAT TTACATTTTT AGGGATATGT
 121 AATTACTAGT TCAGGTGTAT TGCCACAAGA CAAACATGTT AAGAAACTTT CCCGTTATTT
 181 ACGCTCTGTT CCTGTTAATC AACCTCTGGA TTACAAAATT TGTGAAAGAT TGACTGATAT
 241 TCTTAACTAT GTTGCTCCTT TTACGCTGTG TGGATATGCT GCTTTAATGC CTCTGTATCA
 301 TGCTATTGCT TCCCGTACGG CTTTCGTTTT CTCCTCCTTG TATAAATCCT GGTTGCTGTC
 361 TCTTTATGAG GAGTTGTGGC CCGTTGTCCG TCAACGTGGC GTGGTGTGCT CTGTGTTTGC
 421 TGACGCAACC CCCACTGGCT GGGGCATTGC CACCACCTGT CAACTCCTTT CTGGGACTTT
 481 CGCTTTCCCC CTCCCGATCG CCACGGCAGA ACTCATCGCC GCCTGCCTTG CCCGCTGCTG
 541 GACAGGGGCT AGGTTGCTGG GCACTGATAA TTCCGTGGTG TTGTCGGGGA AGGGCC
```

(SEQ ID NO: 27).

In some embodiment, the expression cassette shares at least 95% identity to a sequence selected from SEQ ID NOs: 8-10. In an embodiment, the expression cassette shares complete identity to a sequence selected from SEQ ID NOs: 8-10, or shares at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to a sequence selected from SEQ ID NOs: 8-10:

```
   1 CTGCGCGCTC GCTCGCTCAC TGAGGCCGCC CGGGCAAAGC CCGGGCGTCG GGCGACCTTT
  61 GGTCGCCCGG CCTCAGTGAG CGAGCGAGCG CGCAGAGAGG GAGTGGCCAA CTCCATCACT
 121 AGGGGTTCCT TGTAGTTAAT GATTAACCCG CCATGCTACT TATCTACCAG GGTAATGGGG
 181 ATCCTCTAGA ACTATAGCTA GTCGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT
 241 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC TTACGGTAAA
 301 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA TGACGTATGT
```

```
 361 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT ATTTACGGTA
 421 AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTACGCCCC CTATTGACGT
 481 CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAT GGGACTTTCC
 541 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTCGAGG TGAGCCCCAC
 601 GTTCTGCTTC ACTCTCCCCA TCTCCCCCCC CTCCCCACCC CCAATTTTGT ATTTATTTAT
 661 TTTTTAATTA TTTTGTGCAG CGATGGGGGC GGGGGGGGGG GGGGGGCGCG CGCCAGGCGG
 721 GGCGGGGCGG GGCGAGGGGC GGGGCGGGGC GAGGCGGAGA GGTGCGGCGG CAGCCAATCA
 781 GAGCGGCGCG CTCCGAAAGT TTCCTTTTAT GGCGAGGCGG CGGCGGCGGC GGCCCTATAA
 841 AAAGCGAAGC GCGCGGCGGG CGGGAGTCGC TGCGCGCTGC CTTCGCCCCG TGCCCCGCTC
 901 CGCCGCCGCC TCGCGCCGCC CGCCCCGGCT CTGACTGACC GCGTTACTCC CACAGGTGAG
 961 CGGGCGGGAC GGCCCTTCTC CTCCGGGCTG TAATTAGCGC TTGGTTTAAT GACGGCTTGT
1021 TTCTTTTCTG TGGCTGCGTG AAAGCCTTGA GGGGCTCCGG GAGGGCCCTT TGTGCGGGGG
1081 GAGCGGCTCG GGGGGTGCGT GCGTGTGTGT GTGCGTGGGG AGCGCCGCGT GCGGCTCCGC
1141 GCTGCCCGGC GGCTGTGAGC GCTGCGGGCG CGGCGCGGGG CTTTGTGCGC TCCGCAGTGT
1201 GCGCGAGGGG AGCGCGGCCG GGGGCGGTGC CCCGCGGTGC GGGGGGGGCT GCGAGGGGAA
1261 CAAAGGCTGC GTGCGGGGTG TGTGCGTGGG GGGGTGAGCA GGGGGTGTGG GCGCGTCGGT
1321 CGGGCTGCAA CCCCCCCTGC ACCCCCCTCC CCGAGTTGCT GAGCACGGCC CGGCTTCGGG
1381 TGCGGGGCTC CGTACGGGGC GTGGCGCGGG GCTCGCCGTG CCGGGCGGGG GGTGGCGGCA
1441 GGTGGGGGTG CCGGGCGGGG CGGGGCCGCC TCGGGCCGGG GAGGGCTCGG GGGAGGGGCG
1501 CGGCGGCCCC CGGAGCGCCG GCGGCTGTCG AGGCGCGGCG AGCCGCAGCC ATTGCCTTTT
1561 ATGGTAATCG TGCGAGAGGG CGCAGGGACT TCCTTTGTCC CAAATCTGTG CGGAGCCGAA
1621 ATCTGGGAGG CGCCGCCGCA CCCCCTCTAG CGGGCGCGGG GCGAAGCGGT GCGGCGCCGG
1681 CAGGAAGGAA ATGGGCGGGG AGGGCCTTCG TGCGTCGCCG CGCCGCCGTC CCCTTCTCCC
1741 TCTCCAGCCT CGGGGCTGTC CGCGGGGGGA CGGCTGCCTT CGGGGGGGAC GGGGCAGGGC
1801 GGGGTTCGGC TTCTGGCGTG TGACCGGCGG CTCTAGAGCC TCTGCTAACC ATGTTCATGC
1861 CTTCTTCTTT TTCCTACAGC TCCTGGGCAA CGTGCTGGTT ATTGTGCTGT CTCATCATTT
1921 TGGCAAAGAA TTCGAGCGGC CGCCAGCCGC CACCATGGTC TGCTTCAGAC TGTTCCCTGT
1981 CCCTGGATCT GGTCTGGTGC TTGTGTGCTT GGTGCTGGGT GCTGTGAGAT CCTATGCCCT
2041 TGAGCTGAAC CTGACTGACT CAGAAAATGC CACTTGCCTG TATGCCAAGT GGCAGATGAA
2101 CTTCACTGTG AGATATGAGA CTACCAACAA GACCTACAAG ACTGTGACCA TCTCAGACCA
2161 TGGCACTGTC ACCTACAATG GATCAATCTG TGGTGATGAT CAGAATGGCC CAAAGATAGC
2221 AGTGCAGTTT GGGCCCGGTT TTTCCTGGAT TGCTAACTTC ACCAAGGCAG CCTCCACCTA
2281 CAGCATTGAC TCAGTCAGCT TCAGCTACAA CACTGGGGAT AACACCACCT TCCCTGACGC
2341 AGAGGACAAG GGAATCCTTA CTGTGGACGA ACTCCTGGCA ATCAGAATCC CCCTTAACGA
2401 CCTGTTCAGA TGCAACTCCC TTTCAACCCT TGAAAAGAAT GATGTGGTGC AACACTATTG
2461 GGACGTCCTG GTGCAAGCCT TTGTGCAGAA TGGGACAGTG AGTACCAACG AGTTCCTCTG
2521 TGACAAGGAC AAGACCAGCA CTGTGGCCCC CACTATCCAC ACCACTGTGC CCAGCCCTAC
2581 CACTACCCCC ACCCCTAAAG AGAAGCCAGA AGCTGGAACC TACTCAGTCA ACAATGGAAA
2641 TGACACATGC CTCCTTGCCA CCATGGGACT GCAGCTGAAC ATCACTCAGG ACAAGGTGGC
2701 CTCAGTGATT AACATCAACC CTAACACCAC TCATAGCACT GGGAGCTGCA GATCACATAC
2761 AGCTCTGCTG AGGCTCAACT CCTCCACCAT CAAGTACCTG GACTTTGTGT TTGCTGTGAA
2821 GAATGAGAAC AGGTTCTACC TCAAGGAAGT GAACATTTCC ATGTACCTGG TCAATGGTTC
2881 AGTGTTCTCT ATTGCCAACA ACAATCTGAG CTACTGGGAT GCACCCCTGG GATCCTCCTA
2941 CATGTGCAAC AAGGAGCAGA CTGTGAGTGT GTCAGGTGCT TTTCAGATCA ACACTTTTGA
3001 CCTGAGGGTG CAGCCCTTCA ATGTGACTCA GGGAAAGTAC TCCACTGCAC AAGAGTGTTC
3061 CTTGGATGAT GACACTATCC TCATCCCCAT TATTGTGGGA GCTGGACTGT CAGGATTGAT
3121 TATAGTGATT GTGATTGCTT ATGTGATTGG AAGGAGAAAG AGCTATGCTG GCTACCAGAC
3181 CCTGTAAAAG GGCGAATTCC AGCACACGCG TCCTAGGAGC TCGAGTACTA CTGGCGGCCG
3241 TTACTAGTGG ATCCGCGGTA CAAGTAAGCA TGCAAGCTTC GAGGACGGGG TGAACTACGC
3301 CTGAATCAAG CTTATCGATA AATTCGAGCA TCTTACCGCC ATTTATTCCC ATATTTGTTC
3361 TGTTTTTCTT GATTTGGGTA TACATTTAAA TGTTAATAAA ACAAAATGGT GGGGCAATCA
3421 TTTACATTTT TAGGGATATG TAATTACTAG TTCAGGTGTA TTGCCACAAG ACAAACATGT
3481 TAAGAAACTT TCCCGTTATT TACGCTCTGT TCCTGTTAAT CAACCTCTGG ATTACAAAAT
3541 TTGTGAAAGA TTGACTGATA TTCTTAACTA TGTTGCTCCT TTTACGCTGT GTGGATATGC
3601 TGCTTTAATG CCTCTGTATC ATGCTATTGC TTCCCGTACG GCTTTCGTTT TCTCCTCCTT
3661 GTATAAATCC TGGTTGCTGT CTCTTTATGA GGAGTTGTGG CCCGTTGTCC GTCAACGTGG
3721 CGTGGTGTGC TCTGTGTTTG CTGACGCAAC CCCCACTGGC TGGGGCATTG CCACCACCTG
3781 TCAACTCCTT TCTGGGACTT TCGCTTTCCC CCTCCCGATC GCCACGGCAG AACTCATCGC
3841 CGCCTGCCTT GCCCGCTGCT GGACAGGGGC TAGGTTGCTG GGCACTGATA ATTCCGTGGT
```

```
3901 GTTGTCGGGG AAGGGCCTCG ATACCGTCGA TATCGATCCT GGCTAATAAA GGAAATTTAT
3961 TTTCATTGCA ATAGTGTGTT GGAATTTTTT GTGTCTCTCA CTCGGAAGGA CATATGGGAG
4021 GGCAAATCAT TTAAAACATC AGAATGAGTA TTTGGTTTAG AGTTTGGCAA CATATGCCCA
4081 TATGCTGGCT GCCATGAACA AAGGTTGGCT ATAAAGAGGT CATCAGTATA TGAAACAGCC
4141 CCCTGCTGTC CATTCCTTAT TCCATAGAAA AGCCTTGACT TGAGGTTAGA TTTTTTTTAT
4201 ATTTTGTTTT GTGTTATTTT TTTCTTTAAC ATCCCTAAAA TTTTCCTTAC ATGTTTTACT
4261 AGCCAGATTT TTCCTCCTCT CCTGACTACT CCCAGTCATA GCTGTCCCTC TTCTCTTATG
4321 GAGATCGAAG CAATTCGTTG ATCTGAATTT CGACCACCCA TAATAGATCT CCCATTACCC
4381 TGGTAGATAA GTAGCATGGC GGGTTAATCA TTAACTACAA GGAACCCCTA GTGATGGAGT
4441 TGGCCACTCC CTCTCTGCGC GCTCGCTCGC TCACTGAGGC CGGGCGACCA AAGGTCGCCC
4501 GACGCCCGGG CTTTGCCCGG GCGGCCTCAG TGAGCGAGCG AGCGCGCAG
```

(SEQ ID NO: 8);

```
   1 CTGCGCGCTC GCTCGCTCAC TGAGGCCGCC CGGGCAAAGC CCGGGCGTCG GGCGACCTTT
  61 GGTCGCCCGG CCTCAGTGAG CGAGCGAGCG CGCAGAGAGG GAGTGGCCAA CTCCATCACT
 121 AGGGGTTCCT TGTAGTTAAT GATTAACCCG CCATGCTACT TATCTACCAG GGTAATGGGG
 181 ATCCTCTAGA ACTATAGCTA GTCGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT
 241 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC TTACGGTAAA
 301 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA TGACGTATGT
 361 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT ATTTACGGTA
 421 AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTACGCCCC CTATTGACGT
 481 CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAT GGGACTTTCC
 541 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTCGAGG TGAGCCCCAC
 601 GTTCTGCTTC ACTCTCCCCA TCTCCCCCCC CTCCCCACCC CCAATTTTGT ATTTATTTAT
 661 TTTTTAATTA TTTTGTGCAG CGATGGGGGC GGGGGGGGGG GGGGGGCGCG CGCCAGGCGG
 721 GGCGGGGCGG GGCGAGGGGC GGGGCGGGGC GAGGCGGAGA GGTGCGGCGG CAGCCAATCA
 781 GAGCGGCGCG CTCCGAAAGT TTCCTTTTAT GGCGAGGCGG CGGCGGCGGC GGCCCTATAA
 841 AAAGCGAAGC GCGCGGCGGG CGGGAGTCGC TGCGCGCTGC CTTCGCCCCG TGCCCCGCTC
 901 CGCCGCCGCC TCGCGCCGCC CGCCCCGGCT CTGACTGACC GCGTTACTCC CACAGGTGAG
 961 CGGGCGGGAC GGCCCTTCTC CTCCGGGCTG TAATTAGCGC TTGGTTTAAT GACGGCTTGT
1021 TTCTTTTCTG TGGCTGCGTG AAAGCCTTGA GGGGCTCCGG GAGGGCCCTT TGTGCGGGGG
1081 GAGCGGCTCG GGGGGTGCGT GCGTGTGTGT GTGCGTGGGG AGCGCCGCGT GCGGCTCCGC
1141 GCTGCCCGGC GGCTGTGAGC GCTGCGGGCG CGGCGCGGGG CTTTGTGCGC TCCGCAGTGT
1201 GCGCGAGGGG AGCGCGGCCG GGGGCGGTGC CCCGCGGTGC GGGGGGGGCT GCGAGGGGAA
1261 CAAAGGCTGC GTGCGGGGTG TGTGCGTGGG GGGGTGAGCA GGGGGTGTGG GCGCGTCGGT
1321 CGGGCTGCAA CCCCCCCTGC ACCCCCCTCC CCGAGTTGCT GAGCACGGCC CGGCTTCGGG
1381 TGCGGGGCTC CGTACGGGGC GTGGCGCGGG GCTCGCCGTG CCGGGCGGGG GGTGGCGGCA
1441 GGTGGGGGTG CCGGGCGGGG CGGGGCCGCC TCGGGCCGGG GAGGGCTCGG GGGAGGGGCG
1501 CGGCGGCCCC CGGAGCGCCG GCGGCTGTCG AGGCGCGGCG AGCCGCAGCC ATTGCCTTTT
1561 ATGGTAATCG TGCGAGAGGG CGCAGGGACT TCCTTTGTCC CAAATCTGTG CGGAGCCGAA
1621 ATCTGGGAGG CGCCGCCGCA CCCCCTCTAG CGGGCGCGGG GCGAAGCGGT GCGGCGCCGG
1681 CAGGAAGGAA ATGGGCGGGG AGGGCCTTCG TGCGTCGCCG CGCCGCCGTC CCCTTCTCCC
1741 TCTCCAGCCT CGGGGCTGTC CGCGGGGGGA CGGCTGCCTT CGGGGGGGAC GGGGCAGGGC
1801 GGGGTTCGGC TTCTGGCGTG TGACCGGCGG CTCTAGAGCC TCTGCTAACC ATGTTCATGC
1861 CTTCTTCTTT TTCCTACAGC TCCTGGGCAA CGTGCTGGTT ATTGTGCTGT CTCATCATTT
1921 TGGCAAAGAA TTCGAGCGGC CGCCAGCCGC CACCATGGTG TGCTTTAGAC TGTTTCCTGT
1981 GCCTGGTTCA GGGCTGGTCC TGGTCTGTCT GGTGCTGGGG GCTGTCAGAA GCTATGCCTT
2041 GGAGCTGAAC CTCACTGATA GTGAAAATGC CACTTGTCTG TATGCTAAGT GGCAGATGAA
2101 CTTCACTGTG AGATATGAAA CCACCAACAA GACTTACAAA ACAGTGACCA TCTCAGATCA
2161 TGGAACTGTG ACCTACAACG GCAGCATTTG TGGAGACGAC CAGAACGGAC CAAAAATCGC
2221 TGTCCAATTT GGGCCTGGAT TCTCCTGGAT TGCCAATTTC ACTAAAGCTG CCTCCACATA
2281 TTCAATTGAC TCAGTGTCCT TCTCCTACAA CACTGGGGAC AACACTACTT TCCCTGATGC
2341 TGAAGATAAG GGAATCTTGA CAGTGGATGA GCTGCTGGCT ATCAGGATCC CTTTGAATGA
2401 CCTGTTTAGG TGTAATTCAC TGAGCACTCT GGAGAAGAAC GACGTGGTGC AGCACTACTG
2461 GGACGTGCTG GTGCAGGCCT TTGTGCAGAA CGGCACTGTG TCCACCAACG AATTCCTGTG
2521 TGATAAGGAC AAAACTTCCA CTGTGGCACC TACAATTCAC ACTACTGTGC CTTCACCTAC
2581 CACCACTCCA ACTCCAAAGG AAAAGCCTGA AGCAGGAACC TACTCTGTGA ACAATGGCAA
2641 TGATACCTGT CTGTTGGCCA CCATGGGCCT CCAACTGAAC ATTACTCAGG ACAAGGTGGC
2701 CTCAGTGATT AACATTAACC CCAACACTAC CCACTCCACT GGCAGCTGTA GATCACACAC
```

```
2761 AGCCTTGCTC AGACTGAATA GCAGCACCAT CAAGTATTTG GATTTTGTGT TTGCAGTGAA
2821 GAATGAAAAC AGGTTCTACC TGAAGGAAGT CAACATCTCA ATGTACCTGG TGAACGGCTC
2881 AGTGTTCAGC ATTGCCAACA ACAACCTCTC CTATTGGGAC GCTCCACTGG GGAGCAGCTA
2941 CATGTGTAAC AAGGAACAGA CTGTGTCAGT GTCAGGAGCC TTCCAGATTA ACACCTTTGA
3001 TCTGAGGGTC CAACCCTTTA ATGTCACTCA AGGAAAGTAT AGCACTGCCC AGGAGTGCTC
3061 CCTGGATGAT GACACCATTC TGATTCCAAT CATTGTGGGT GCAGGACTTT CTGGGCTTAT
3121 TATTGTGATT GTGATTGCCT ATGTGATTGG CAGAAGGAAA TCCTATGCAG GGTACCAAAC
3181 TCTGTAAAAG GGCGAATTCC AGCACACGCG TCCTAGGAGC TCGAGTACTA CTGGCGGCCG
3241 TTACTAGTGG ATCCGCGGTA CAAGTAAGCA TGCAAGCTTC GAGGACGGGG TGAACTACGC
3301 CTGAATCAAG CTTATCGATA AATTCGAGCA TCTTACCGCC ATTTATTCCC ATATTTGTTC
3361 TGTTTTTCTT GATTTGGGTA TACATTTAAA TGTTAATAAA ACAAAATGGT GGGGCAATCA
3421 TTTACATTTT TAGGGATATG TAATTACTAG TTCAGGTGTA TTGCCACAAG ACAAACATGT
3481 TAAGAAACTT TCCCGTTATT TACGCTCTGT TCCTGTTAAT CAACCTCTGG ATTACAAAAT
3541 TTGTGAAAGA TTGACTGATA TTCTTAACTA TGTTGCTCCT TTTACGCTGT GTGGATATGC
3601 TGCTTTAATG CCTCTGTATC ATGCTATTGC TTCCCGTACG GCTTTCGTTT TCTCCTCCTT
3661 GTATAAATCC TGGTTGCTGT CTCTTTATGA GGAGTTGTGG CCCGTTGTCC GTCAACGTGG
3721 CGTGGTGTGC TCTGTGTTTG CTGACGCAAC CCCCACTGGC TGGGGCATTG CCACCACCTG
3781 TCAACTCCTT TCTGGGACTT TCGCTTTCCC CCTCCCGATC GCCACGGCAG AACTCATCGC
3841 CGCCTGCCTT GCCCGCTGCT GGACAGGGGC TAGGTTGCTG GGCACTGATA ATTCCGTGGT
3901 GTTGTCGGGG AAGGGCCTCG ATACCGTCGA TATCGATCCT GGCTAATAAA GGAAATTTAT
3961 TTTCATTGCA ATAGTGTGTT GGAATTTTTT GTGTCTCTCA CTCGGAAGGA CATATGGGAG
4021 GGCAAATCAT TTAAAACATC AGAATGAGTA TTTGGTTTAG AGTTTGGCAA CATATGCCCA
4081 TATGCTGGCT GCCATGAACA AAGGTTGGCT ATAAAGAGGT CATCAGTATA TGAAACAGCC
4141 CCCTGCTGTC CATTCCTTAT TCCATAGAAA AGCCTTGACT TGAGGTTAGA TTTTTTTTAT
4201 ATTTTGTTTT GTGTTATTTT TTTCTTTAAC ATCCCTAAAA TTTTCCTTAC ATGTTTTACT
4261 AGCCAGATTT TTCCTCCTCT CCTGACTACT CCCAGTCATA GCTGTCCCTC TTCTCTTATG
4321 GAGATCGAAG CAATTCGTTG ATCTGAATTT CGACCACCCA TAATAGATCT CCCATTACCC
4381 TGGTAGATAA GTAGCATGGC GGGTTAATCA TTAACTACAA GGAACCCCTA GTGATGGAGT
4441 TGGCCACTCC CTCTCTGCGC GCTCGCTCGC TCACTGAGGC CGGGCGACCA AAGGTCGCCC
4501 GACGCCCGGG CTTTGCCCGG GCGGCCTCAG TGAGCGAGCG AGCGCGCAG
```

(SEQ ID NO: 9); and

```
   1 CTGCGCGCTC GCTCGCTCAC TGAGGCCGCC CGGGCAAAGC CCGGGCGTCG GGCGACCTTT
  61 GGTCGCCCGG CCTCAGTGAG CGAGCGAGCG CGCAGAGAGG GAGTGGCCAA CTCCATCACT
 121 AGGGGTTCCT TGTAGTTAAT GATTAACCCG CCATGCTACT TATCTACCAG GGTAATGGGG
 181 ATCCTCTAGA ACTATAGCTA GTCGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT
 241 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC TTACGGTAAA
 301 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA TGACGTATGT
 361 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT ATTTACGGTA
 421 AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTACGCCCC CTATTGACGT
 481 CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAT GGGACTTTCC
 541 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTCGAGG TGAGCCCCAC
 601 GTTCTGCTTC ACTCTCCCCA TCTCCCCCCC CTCCCCACCC CCAATTTTGT ATTTATTTAT
 661 TTTTTAATTA TTTTGTGCAG CGATGGGGGC GGGGGGGGGG GGGGGGCGCG CGCCAGGCGG
 721 GGCGGGGCGG GGCGAGGGGC GGGGCGGGGC GAGGCGGAGA GGTGCGGCGG CAGCCAATCA
 781 GAGCGGCGCG CTCCGAAAGT TTCCTTTTAT GGCGAGGCGG CGGCGGCGGC GGCCCTATAA
 841 AAAGCGAAGC GCGCGGCGGG CGGGAGTCGC TGCGCGCTGC CTTCGCCCCG TGCCCCGCTC
 901 CGCCGCCGCC TCGCGCCGCC CGCCCCGGCT CTGACTGACC GCGTTACTCC CACAGGTGAG
 961 CGGGCGGGAC GGCCCTTCTC CTCCGGGCTG TAATTAGCGC TTGGTTTAAT GACGGCTTGT
1021 TTCTTTTCTG TGGCTGCGTG AAAGCCTTGA GGGGCTCCGG GAGGGCCCTT TGTGCGGGGG
1081 GAGCGGCTCG GGGGGTGCGT GCGTGTGTGT GTGCGTGGGG AGCGCCGCGT GCGGCTCCGC
1141 GCTGCCCGGC GGCTGTGAGC GCTGCGGGCG CGGCGCGGGG CTTTGTGCGC TCCGCAGTGT
1201 GCGCGAGGGG AGCGCGGCCG GGGGCGGTGC CCCGCGGTGC GGGGGGGGCT GCGAGGGGAA
1261 CAAAGGCTGC GTGCGGGGTG TGTGCGTGGG GGGGTGAGCA GGGGGTGTGG GCGCGTCGGT
1321 CGGGCTGCAA CCCCCCCTGC ACCCCCCTCC CCGAGTTGCT GAGCACGGCC CGGCTTCGGG
1381 TGCGGGGCTC CGTACGGGGC GTGGCGCGGG GCTCGCCGTG CCGGGCGGGG GGTGGCGGCA
1441 GGTGGGGGTG CCGGGCGGGG CGGGGCCGCC TCGGGCCGGG GAGGGCTCGG GGGAGGGGCG
```

```
1501 CGGCGGCCCC CGGAGCGCCG GCGGCTGTCG AGGCGCGGCG AGCCGCAGCC ATTGCCTTTT
1561 ATGGTAATCG TGCGAGAGGG CGCAGGGACT TCCTTTGTCC CAAATCTGTG CGGAGCCGAA
1621 ATCTGGGAGG CGCCGCCGCA CCCCCTCTAG CGGGCGCGGG GCGAAGCGGT GCGGCGCCGG
1681 CAGGAAGGAA ATGGGCGGGG AGGGCCTTCG TGCGTCGCCG CGCCGCCGTC CCCTTCTCCC
1741 TCTCCAGCCT CGGGGCTGTC CGCGGGGGGA CGGCTGCCTT CGGGGGGGAC GGGGCAGGGC
1801 GGGGTTCGGC TTCTGGCGTG TGACCGGCGG CTCTAGAGCC TCTGCTAACC ATGTTCATGC
1861 CTTCTTCTTT TTCCTACAGC TCCTGGGCAA CGTGCTGGTT ATTGTGCTGT CTCATCATTT
1921 TGGCAAAGAA TTCGAGCGGC CGCCAGCCGC CACCATGGTC TGTTTTAGGC TGTTCCCTGT
1981 CCCTGGTTCA GGACTGGTCT TAGTGTGTCT GGTGCTTGGA GCTGTCAGAA GCTATGCCCT
2041 GGAGCTGAAC CTGACTGACT CAGAAAATGC CACTTGCCTG TATGCCAAGT GGCAGATGAA
2101 CTTCACTGTC AGATATGAAA CCACCAACAA GACCTATAAG ACTGTGACCA TCTCAGACCA
2161 TGGCACTGTG ACTTACAATG GGTCAATTTG TGGAGATGAC CAGAATGGCC CTAAGATAGC
2221 TGTCCAGTTT GGTCCAGGAT TCAGCTGGAT TGCCAACTTC ACCAAGGCAG CCAGCACCTA
2281 CAGCATTGAC TCTGTGTCCT TCTCCTACAA CACAGGAGAC AACACCACTT TCCCTGATGC
2341 AGAGGACAAA GGTATCCTGA CTGTGGATGA GTTGCTGGCA ATCAGGATCC CACTGAACGA
2401 TCTGTTCAGG TGCAACTCAC TGTCCACTCT GGAAAAGAAT GATGTGGTGC AGCACTATTG
2461 GGATGTGCTA GTCCAGGCCT TTGTCCAGAA TGGGACTGTG TCAACTAATG AGTTCCTGTG
2521 TGACAAGGAC AAGACAAGCA CTGTAGCCCC CACTATCCAT ACCACAGTAC CTAGCCCCAC
2581 CACTACTCCA ACCCCCAAGG AGAAGCCTGA GGCTGGCACC TACTCAGTGA ACAATGGGAA
2641 TGACACCTGT TTGCTGGCCA CTATGGGACT CCAACTGAAC ATCACCCAGG ACAAAGTGGC
2701 CTCTGTGATC AATATCAATC CCAACACCAC CCACAGCACT GGGTCCTGCA GAAGCCACAC
2761 TGCCCTCCTG AGGCTCAACT CATCAACTAT CAAGTACTTG GATTTTGTGT TTGCAGTGAA
2821 GAATGAGAAC AGATTCTACC TCAAAGAGGT CAACATTTCA ATGTACCTGG TGAATGGGAG
2881 TGTGTTCTCC ATTGCTAACA ACAACCTGAG CTACTGGGAT GCCCCTCTGG GCTCCTCATA
2941 CATGTGCAAC AAGGAACAGA CTGTGAGTGT GTCAGGGGCC TTCCAGATCA ACACTTTTGA
3001 CCTGAGAGTG CAGCCCTTTA ATGTGACACA GGGAAAGTAC AGCACTGCTC AGGAGTGCAG
3061 CCTGGATGAT GACACTATCC TGATCCCTAT CATTGTGGGG GCAGGCCTGT CTGGACTCAT
3121 TATTGTGATT GTGATTGCCT ATGTGATAGG GAGAAGGAAG TCTTATGCTG GATACCAGAC
3181 CCTGTAAAAG GGCGAATTCC AGCACACGCG TCCTAGGAGC TCGAGTACTA CTGGCGGCCG
3241 TTACTAGTGG ATCCGCGGTA CAAGTAAGCA TGCAAGCTTC GAGGACGGGG TGAACTACGC
3301 CTGAATCAAG CTTATCGATA AATTCGAGCA TCTTACCGCC ATTTATTCCC ATATTTGTTC
3361 TGTTTTTCTT GATTTGGGTA TACATTTAAA TGTTAATAAA ACAAAATGGT GGGGCAATCA
3421 TTTACATTTT TAGGGATATG TAATTACTAG TTCAGGTGTA TTGCCACAAG ACAAACATGT
3481 TAAGAAACTT TCCCGTTATT TACGCTCTGT TCCTGTTAAT CAACCTCTGG ATTACAAAAT
3541 TTGTGAAAGA TTGACTGATA TTCTTAACTA TGTTGCTCCT TTTACGCTGT GTGGATATGC
3601 TGCTTTAATG CCTCTGTATC ATGCTATTGC TTCCCGTACG GCTTTCGTTT TCTCCTCCTT
3661 GTATAAATCC TGGTTGCTGT CTCTTTATGA GGAGTTGTGG CCCGTTGTCC GTCAACGTGG
3721 CGTGGTGTGC TCTGTGTTTG CTGACGCAAC CCCCACTGGC TGGGGCATTG CCACCACCTG
3781 TCAACTCCTT TCTGGGACTT TCGCTTTCCC CCTCCCGATC GCCACGGCAG AACTCATCGC
3841 CGCCTGCCTT GCCCGCTGCT GGACAGGGGC TAGGTTGCTG GGCACTGATA ATTCCGTGGT
3901 GTTGTCGGGG AAGGGCCTCG ATACCGTCGA TATCGATCCT GGCTAATAAA GGAAATTTAT
3961 TTTCATTGCA ATAGTGTGTT GGAATTTTTT GTGTCTCTCA CTCGGAAGGA CATATGGGAG
4021 GGCAAATCAT TTAAAACATC AGAATGAGTA TTTGGTTTAG AGTTTGGCAA CATATGCCCA
4081 TATGCTGGCT GCCATGAACA AAGGTTGGCT ATAAAGAGGT CATCAGTATA TGAAACAGCC
4141 CCCTGCTGTC CATTCCTTAT TCCATAGAAA AGCCTTGACT TGAGGTTAGA TTTTTTTTAT
4201 ATTTTGTTTT GTGTTATTTT TTTCTTTAAC ATCCCTAAAA TTTTCCTTAC ATGTTTTACT
4261 AGCCAGATTT TTCCTCCTCT CCTGACTACT CCCAGTCATA GCTGTCCCTC TTCTCTTATG
4321 GAGATCGAAG CAATTCGTTG ATCTGAATTT CGACCACCCA TAATAGATCT CCCATTACCC
4381 TGGTAGATAA GTAGCATGGC GGGTTAATCA TTAACTACAA GGAACCCCTA GTGATGGAGT
4441 TGGCCACTCC CTCTCTGCGC GCTCGCTCGC TCACTGAGGC CGGGCGACCA AAGGTCGCCC
4501 GACGCCCGGG CTTTGCCCGG GCGGCCTCAG TGAGCGAGCG AGCGCGCAG
```

(SEQ ID NO: 10).

In certain embodiments, the expression cassette comprises one or more modifications as compared to a sequence selected from SEQ ID NOs: 8-10, including but not limited to any

of the modifications disclosed herein. In particular embodiments, the one or more modifications comprise one or more of: removal of one or more (*e.g.*, all) upstream ATG sequences, replacement of the Kozak sequence with an optimized consensus Kozak sequence or another Kozak sequence, including but not limited to any of those disclosed herein, and/or replacement of the polyadenylation sequence with a full-length polyadenylation sequence or another polyadenylation sequence, including but not limited to any of those disclosed herein. An illustrative configuration of genetic elements within these exemplary expression cassettes is depicted in **FIG. 1B**.

In an embodiment, the vector is an adeno-associated virus (AAV) vector. In an embodiment, the expression cassette comprises inverted terminal repeat (ITR) sequences selected from SEQ ID NOs: 11 and 12:

```
  1 CTGCGCGCTC GCTCGCTCAC TGAGGCCGCC CGGGCAAAGC CCGGGCGTCG GGCGACCTTT
 61 GGTCGCCCGG CCTCAGTGAG CGAGCGAGCG CGCAGAGAGG GAGTGGCCAA CTCCATCACT
121 AGGGGTTCCT
```

(SEQ ID NO: 11);

```
  1 AGGAACCCCT AGTGATGGAG TTGGCCACTC CCTCTCTGCG CGCTCGCTCG CTCACTGAGG
 61 CCGGGCGACC AAAGGTCGCC CGACGCCCGG GCTTTGCCCG GGCGGCCTCA GTGAGCGAGC
121 GAGCGCGCAG
```

(SEQ ID NO: 12).

In related embodiments, the disclosure provides gene therapy vectors comprising an expression cassette disclosed herein. Generally, the gene therapy vectors described herein comprise an expression cassette comprising a polynucleotide encoding one or more isoforms of lysosome-associated membrane protein 2 (LAMP-2), and allows for the expression of LAMP-2 to partially or wholly rectify deficient LAMP-2 protein expression levels and/or autophagic flux in a subject in need thereof (*e.g.*, a subject having Danon disease or another disorder characterized by deficient autophagic flux at least in part due to deficient LAMP-2 expression). In particular embodiments, the expression cassette comprises a polynucleotide sequence encoding LAMP-2 disclosed herein, *e.g.*, SEQ ID NOs: 3-5, or a functional variant thereof. In some embodiments, the variant sequence has at least 90%, at least 95%, at least 98%, or at least 99% identity to any of SEQ ID NOs: 3-5. In some embodiments, the variant is a fragment of any of SEQ ID NOs: 3-5, *e.g.*, a fragment having at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95% of the sequence of any of SEQ ID Nos: 3-5. The gene therapy vectors can be viral or non-viral vectors. Illustrative non-viral

vectors include, *e.g.*, naked DNA, cationic liposome complexes, cationic polymer complexes, cationic liposome-polymer complexes, and exosomes. Examples of viral vectors include, but are not limited to, adenoviral, retroviral, lentiviral, herpesvirus and adeno-associated virus (AAV) vectors.

In certain embodiments, the viral vector is an AAV vector. AAV is a 4.7 kb, single stranded DNA virus. Recombinant vectors based on AAV are associated with excellent clinical safety, since wild-type AAV is nonpathogenic and has no etiologic association with any known diseases. In addition, AAV offers the capability for highly efficient gene delivery and sustained transgene expression in numerous tissues. By an "AAV vector" is meant a vector derived from an adeno-associated virus serotype, including without limitation, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAVrh.10, AAVrh.74, *etc.* AAV vectors can have one or more of the AAV wild-type genes deleted in whole or part, *e.g.*, the rep and/or cap genes, but retain functional flanking inverted terminal repeat (ITR) sequences. Functional ITR sequences are necessary for the rescue, replication and packaging of the AAV virion. Thus, an AAV vector is defined herein to include at least those sequences required in cis for replication and packaging (*e.g.*, functional ITRs) of the virus. The ITRs need not be the wild-type nucleotide sequences, and may be altered, *e.g.* by the insertion, deletion or substitution of nucleotides, as long as the sequences provide for functional rescue, replication and packaging. AAV vectors may comprise other modifications, including but not limited to one or more modified capsid protein (*e.g.*, VP1, VP2 and/or VP3). For example, a capsid protein may be modified to alter tropism and/or reduce immunogenicity.

Recombinant vectors based on AAV are associated with excellent clinical safety, since wild-type AAV is nonpathogenic and has no etiologic association with any known diseases. In addition, AAV offers the capability for highly efficient gene delivery and sustained transgene expression in numerous tissues. Various serotypes of AAV are known, including, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAVrh.10, AAVrh.74, *etc.*. AAV vectors can have one or more of the AAV wild-type genes deleted in whole or part, *e.g.*, the *rep* and/or *cap* genes, but retain functional flanking inverted terminal repeat (ITR) sequences. The serotype of a recombinant AAV vector is determined by its capsid. International Patent Publication No. WO2003042397A2 discloses various capsid sequences including those of AAV1, AAV2, AAV3, AAV8, AAV9, and rh10.

International Patent Publication No. WO2013078316A1 discloses the polypeptide sequence of the VP1 from AAVrh74. Numerous diverse naturally occurring or genetically modified AAV capsid sequences are known in the art.

An exemplary, non-limiting capsid is an AAV9 capsid, having the sequence of SEQ ID NO: 28 (or the VP1, VP2, or VP3 fragments thereof). In some embodiments, the AAV vectors of the disclosure comprise capsid proteins that share at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 100% identity of the entire sequence of SEQ ID NO: 28, or over amino acids 138 to 736 of SEQ ID NO: 28, or over amino acids 203 to 736 of SEQ ID NO: 28.

```
  1 MAADGYLPDW LEDNLSEGIR EWWALKPGAP QPKANQQHQD NARGLVLPGY KYLGPGNGLD
 61 KGEPVNAADA AALEHDKAYD QQLKAGDNPY LKYNHADAEF QERLKEDTSF GGNLGRAVFQ
121 AKKRLLEPLG LVEEAAKTAP GKKRPVEQSP QEPDSSAGIG KSGAQPAKKR LNFGQTGDTE
181 SVPDPQPIGE PPAAPSGVGS LTMASGGGAP VADNNEGADG VGSSSGNWHC DSQWLGDRVI
241 TTSTRTWALP TYNNHLYKQI SNSTSGGSSN DNAYFGYSTP WGYFDFNRFH CHFSPRDWQR
301 LINNNWGFRP KRLNFKLFNI QVKEVTDNNG VKTIANNLTS TVQVFTDSDY QLPYVLGSAH
361 EGCLPPFPAD VFMIPQYGYL TLNDGSQAVG RSSFYCLEYF PSQMLRTGNN FQFSYEFENV
421 PFHSSYAHSQ SLDRLMNPLI DQYLYYLSKT INGSGQNQQT LKFSVAGPSN MAVQGRNYIP
481 GPSYRQQRVS TTVTQNNNSE FAWPGASSWA LNGRNSLMNP GPAMASHKEG EDRFFPLSGS
541 LIFGKQGTGR DNVDADKVMI TNEEEIKTTN PVATESYGQV ATNHQSAQAQ AQTGWVQNQG
601 ILPGMVWQDR DVYLQGPIWA KIPHTDGNFH PSPLMGGFGM KHPPPQILIK NTPVPADPPT
661 AFNKDKLNSF ITQYSTGQVS VEIEWELQKE NSKRWNPEIQ YTSNYYKSNN VEFAVNTEGV
721 YSEPRPIGTR YLTRNL
```

(SEQ ID NO: 28).

AAV expression vectors are constructed using known techniques to at least provide as operatively linked components in the direction of transcription, control elements including a transcriptional initiation region, the DNA of interest (*i.e.* the LAMP-2 gene) and a transcriptional termination region.

In some embodiments, the viral vector is an AAV9 vector. In some embodiments, the expression cassette of the viral vector is flanked by AAV2 inverted terminal repeats (ITRs). ITRs used in alternative embodiments of the disclosed vectors include, but are not limited to, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, and AAV9. In some embodiments, the viral vector is an AAV2/9 vector. The notation AAV2/9 refers to an AAV vector have the ITRs of AAV2 and the capsid of AAV9. Other embodiments of the disclosure include without limitation AAV2/9, AAV5/9, AAVrh74, AAV2/rh74, AAV5/9, and AAV5/rh74 vectors. Other ITRs known in the art may be used. Exemplary ITRs (and other AAV components) useful in the vectors of the present disclosure include, without limitation, those

described in US6936466B2, US9169494B2, US20050220766A1, US20190022249A1, and US7282199B2, which are each incorporated by reference herein in their entireties.

In some embodiments, the vector is a retroviral vector, or more specifically, a lentiviral vector. As used herein, the term "retrovirus" or "retroviral" refers an RNA virus that reverse transcribes its genomic RNA into a linear double-stranded DNA copy and subsequently covalently integrates its genomic DNA into a host genome. Retrovirus vectors are a common tool for gene delivery (Miller, 2000, *Nature*. 357: 455-460). Once the virus is integrated into the host genome, it is referred to as a "provirus." The provirus serves as a template for RNA polymerase II and directs the expression of RNA molecules encoded by the virus.

Illustrative retroviruses (family Retroviridae) include, but are not limited to: (1) genus gammaretrovirus, such as, Moloney murine leukemia virus (M-MuLV), Moloney murine sarcoma virus (MoMSV), murine mammary tumor virus (MuMTV), gibbon ape leukemia virus (GaLV), and feline leukemia virus (FLV), (2) genus spumavirus, such as, simian foamy virus, (3) genus lentivirus, such as, human immunodeficiency virus-1 and simian immunodeficiency virus.

As used herein, the term "lentiviral" or "lentivirus" refers to a group (or genus) of complex retroviruses. Illustrative lentiviruses include, but are not limited to: HIV (human immunodeficiency virus; including HIV type 1, and HIV type 2; visna-maedi virus (VMV) virus; the caprine arthritis-encephalitis virus (CAEV); equine infectious anemia virus (EIAV); feline immunodeficiency virus (FIV); bovine immune deficiency virus (BIV); and simian immunodeficiency virus (SIV). In one embodiment, HIV-based vector backbones (*i.e.*, HIV cis-acting sequence elements) are preferred.

Retroviral vectors, and more particularly, lentiviral vectors, may be used in practicing the present invention. Accordingly, the term "retroviral vector," as used herein is meant to include "lentiviral vector"; and the term "retrovirus" as used herein is meant to include "lentivirus."

The term viral vector may refer either to a vector or viral particle capable of transferring a nucleic acid into a cell or to the transferred nucleic acid itself. Viral vectors contain structural and/or functional genetic elements that are primarily derived from a virus. The term "retroviral vector" refers to a viral vector containing structural and functional

genetic elements, or portions thereof, that are primarily derived from a retrovirus. The term "lentiviral vector" refers to a viral vector containing structural and functional genetic elements, or portions thereof, including LTRs that are primarily derived from a lentivirus. The term "hybrid" refers to a vector, LTR or other nucleic acid containing both retroviral, *e.g.*, lentiviral, sequences and non-lentiviral viral sequences. In one embodiment, a hybrid vector refers to a vector or transfer plasmid comprising retroviral, *e.g.*, lentiviral, sequences for reverse transcription, replication, integration and/or packaging.

In particular embodiments, the terms "lentiviral vector" and "lentiviral expression vector" may be used to refer to lentiviral transfer plasmids and/or infectious lentiviral particles. Where reference is made herein to elements such as cloning sites, promoters, regulatory elements, heterologous nucleic acids, *etc.*, it is to be understood that the sequences of these elements are present in RNA form in the lentiviral particles of the invention and are present in DNA form in the DNA plasmids of the invention.

According to certain specific embodiments, most or all of the viral vector backbone sequences are derived from a lentivirus, *e.g.*, HIV-1. However, it is to be understood that many different sources of lentiviral sequences can be used, and numerous substitutions and alterations in certain of the lentiviral sequences may be accommodated without impairing the ability of a transfer vector to perform the functions described herein. Moreover, a variety of lentiviral vectors are known in the art, *see* Naldini *et al.*, (1996a, 1996b, and 1998); Zufferey *et al.*, (1997); Dull *et al.*, 1998, U.S. Pat. Nos. 6,013,516; and 5,994,136, many of which may be adapted to produce a viral vector or transfer plasmid of the present invention.

The LAMP-2B transgene sequences disclosed herein are, in various embodiments, used in any vector system known in the art or prospectively discovered. The invention is not limited to any particular viral vector described herein, as it is within the skill of those in the art to use a transgene sequence in other vector systems without undue experimentation and with a reasonable expectation of success.

Gene delivery viral vectors useful in the practice of the present invention can be constructed utilizing methodologies well known in the art of molecular biology. Typically, viral vectors carrying transgenes are assembled from polynucleotides encoding the transgene, suitable regulatory elements and elements necessary for production of viral proteins, which mediate cell transduction. Such recombinant viruses may be produced by techniques known

in the art, *e.g.*, by transfecting packaging cells or by transient transfection with helper plasmids or viruses. Typical examples of virus packaging cells include but are not limited to HeLa cells, SF9 cells (optionally with a baculovirus helper vector), 293 cells, etc. A Herpesvirus-based system can be used to produce AAV vectors, as described in US20170218395A1. Detailed protocols for producing such replication-defective recombinant viruses may be found for instance in W095/14785, W096/22378, U.S. Pat. No. 5,882,877, U.S. Pat. No. 6,013,516, U.S. Pat. No. 4,861,719, U.S. Pat. No. 5,278,056 and W094/19478, the complete contents of each of which is hereby incorporated by reference.

The present disclosure also provides pharmaceutical compositions comprising an expression cassette or vector (*e.g.*, gene therapy vector) disclosed herein and one or more pharmaceutically acceptable carriers, diluents or excipients. In particular embodiments, the pharmaceutical composition comprises an AAV vector comprising an expression cassette disclosed herein, *e.g.*, wherein the expression cassette comprises a codon-optimized transgene encoding LAMP-2B, *e.g.*, any of SEQ ID NOs: 3-5 and variants thereof. Provided are pharmaceutical compositions, *e.g.*, for use in preventing or treating a disorder characterized by deficient autophagic flux (*e.g.*, Danon disease) which comprises a therapeutically effective amount of an expression cassette or vector disclosed herein that comprises a nucleic acid sequence of a polynucleotide that encodes one or more isoforms of LAMP-2.

AAV vectors useful in the practice of the present invention can be packaged into AAV virions (viral particles) using various systems including adenovirus-based and helper-free systems. Standard methods in AAV biology include those described in Kwon and Schaffer. *Pharm Res.* (2008) 25(3):489-99; Wu et al. *Mol. Ther.* (2006) 14(3):316-27. Burger et al. *Mol. Ther.* (2004) 10(2):302-17; Grimm et al. *Curr Gene Ther.* (2003) 3(4):281-304; Deyle DR, Russell DW. *Curr Opin Mol Ther.* (2009) 11(4):442-447; McCarty et al. *Gene Ther.* (2001) 8(16):1248-54; and Duan et al. *Mol Ther.* (2001) 4(4):383-91. Helper-free systems included those described in US 6,004,797; US 7,588,772; and US 7,094,604;

The pharmaceutical compositions that contain the expression cassette or vector may be in any form that is suitable for the selected mode of administration, for example, for intraventricular, intramyocardial, intracoronary, intravenous, intra-arterial, intra-renal, intraurethral, epidural or intramuscular administration. The gene therapy vector comprising a polynucleotide encoding one or more LAMP-2 isoforms can be administered, as sole active

agent, or in combination with other active agents, in a unit administration form, as a mixture with conventional pharmaceutical supports, to animals and human beings. In some embodiments, the pharmaceutical composition comprises cells transduced *ex vivo* with any of the gene therapy vectors of the disclosure.

In some embodiments, the viral vector (*e.g.* AAV vector), or a pharmaceutical composition comprising that vector, is effective when administered systemically. For example, the viral vectors of the disclosure, in some cases, demonstrate efficacy when administered intravenously to subject (*e.g.*, a primate, such as a non-human primate or a human). In some embodiments, the viral vectors of the disclosure are capable of inducing expression of LAMP-2B in various tissues when administered systemically (*e.g.*, in heart, muscle, and/or lung). In particular embodiments, administration of an AAV9 vector comprising a transgene substantially identical to, or identical to, SEQ ID NO: 3 to a subject intravenously results in detectable expression of LAMP-2B in heart tissue. In some embodiments, expression of LAMP-2B is detectable in one or more, or all, of the left ventricle, the right ventricle, the left atrium, and the right atrium of the heart of the subject. In some embodiments, expression of LAMP-2B is detectable in sub-region 1 and/or sub-region 2 of the left ventricle of the heart of the subject.

"Detectable expression" typically refers to transgene expression at least 5%, 10%, 15%, 20% or more compared to a control subject or tissue not treated with the viral vector. In some embodiments, detectable expression means expression at least 1.5-fold, at least 2-fold, at least 2.5-fold, at least 3-fold, at least 5-fold, at least 10-fold, at least 20-fold, at least 50-fold, or at least 100-fold greater than a no-vector control. Transgene expression can be determined as the increase over expression of the wild-type or endogenous gene in the cell (accounting for the potential that expression of the transgene may influence expression of the endogenous gene). Transgene expression can also be determined by RT-PCR detection of sequences that are present on the transgene mRNA transcript but not on the mRNA transcript of the endogenous gene. For example, the 3′ UTR of the transcript may be used to determine the expression of the transgene independent of the expression of the endogenous gene (which may have a different 3′ UTR). Expression of the polypeptide encoded by the transgene can be assessed by western blot or enzyme-linked immunosorbent assay (ELISA), as described in the examples that follow, or other methods known in the art. Antibodies cross-reactive to the wild-type and exogenous copies of the protein may be used. In some cases, an antibody

specific to the exogenous protein can be identified and used to determine transgene expression. Those of skill in the art can design appropriate detection methodologies taking into account the target cell or tissue. In some cases, expression is measured quantitatively using a standard curve. Standard curves can be generated using purified LAMP-2 protein, by methods described in the examples or known in the art. Alternatively, expression of the transgene can be assessed by quantification of the corresponding mRNA.

In some embodiments, detectable expression of LAMP-2B in heart tissue occurs at doses, in vector genomes (vg) per kilogram weight of subject (kg), of $5x10^{14}$ vg/kg or less, 3×10^{14} vg/kg or less, 2×10^{14} vg/kg or less, 1×10^{14} vg/kg or less, 9×10^{13} vg/kg or less, 8×10^{13} vg/kg or less, 7×10^{13} vg/kg or less, 6×10^{13} vg/kg or less, 5×10^{13} vg/kg or less, 4×10^{13} vg/kg or less, 3×10^{13} vg/kg or less, 2×10^{13} vg/kg or less, or 1×10^{13} vg/kg or less.

In some embodiments, detectable expression of LAMP-2B in heart tissue occurs at doses, in vector genomes (vg) per kilogram weight of subject (kg), of 1×10^{13} vg/kg to 2×10^{13} vg/kg, 2×10^{13} vg/kg to 3×10^{13} vg/kg, 3×10^{13} vg/kg to 4×10^{13} vg/kg, 4×10^{13} vg/kg to 5×10^{13} vg/kg, 5×10^{13} vg/kg to 6×10^{13} vg/kg, 6×10^{13} vg/kg to 7×10^{13} vg/kg, 7×10^{13} vg/kg to 8×10^{13} vg/kg, 8×10^{13} vg/kg to 9×10^{13} vg/kg, 9×10^{13} vg/kg to 1×10^{14} vg/kg, 1×10^{14} vg/kg to 2×10^{14} vg/kg, 2×10^{14} vg/kg to 3×10^{14} vg/kg, or 3×10^{14} vg/kg to $5x10^{14}$ vg/kg.

In some embodiments, detectable expression of LAMP-2B in heart tissue occurs at doses, in vector genomes (vg) per kilogram weight of subject (kg), of 1×10^{13} vg/kg to 3×10^{13} vg/kg, 3×10^{13} vg/kg to 5×10^{13} vg/kg, 5×10^{13} vg/kg to 7×10^{13} vg/kg, 7×10^{13} vg/kg to 9×10^{13} vg/kg, 9×10^{13} vg/kg to 2×10^{14} vg/kg, or 2×10^{14} vg/kg to $5x10^{14}$ vg/kg. In some embodiments, detectable expression of LAMP-2B in heart tissue occurs at doses, in vector genomes (vg) per kilogram weight of subject (kg), of 1×10^{13} vg/kg to 5×10^{13} vg/kg, 5×10^{13} vg/kg to 9×10^{13} vg/kg, 9×10^{13} vg/kg or to $5x10^{14}$ vg/kg. In some embodiments, detectable expression of LAMP-2B in heart tissue occurs at doses, in vector genomes (vg) per kilogram weight of subject (kg), of 1×10^{13} vg/kg to 9×10^{13} vg/kg, or 9×10^{13} vg/kg or to $5x10^{14}$ vg/kg.

In some embodiments, detectable expression of LAMP-2B in heart tissue occurs at doses, in vector genomes (vg) per kilogram weight of subject (kg), of 1×10^{13} vg/kg to 5×10^{13} vg/kg, 5×10^{13} vg/kg to 1×10^{14} vg/kg, or 1×10^{14} vg/kg to $5x10^{14}$ vg/kg.

In some embodiments, detectable expression of LAMP-2B in heart tissue occurs at doses, in vector genomes (vg) per kilogram weight of subject (kg), of 1×10^{13} vg/kg to $5x10^{14}$

vg/kg. In some embodiments, detectable expression of LAMP-2B in heart tissue occurs at doses, in vector genomes (vg) per kilogram weight of subject (kg), of 1×10^{13} vg/kg to 1×10^{14}.

In various embodiments, the pharmaceutical compositions contain vehicles (*e.g.*, carriers, diluents and excipients) that are pharmaceutically acceptable for a formulation capable of being injected. Exemplary excipients include a poloxamer. Formulation buffers for viral vectors (including AAV) general contains salts to prevent aggregation and other excipients (*e.g.* poloxamer) to reduce stickiness of the vector. These may be in particular isotonic, sterile, saline solutions (monosodium or disodium phosphate, sodium, potassium, calcium or magnesium chloride and the like or mixtures of such salts), or dry, especially freeze-dried compositions which upon addition, depending on the case, of sterilized water or physiological saline, permit the constitution of injectable solutions. Advantageously, the formulation is stable for storage and use when frozen (*e.g.* at less than 0°C, about -60°C, or about -72°C).

Exemplary methods of treating lysosomal disorders and/or Danon disease are provided in WO 2018/170239 A1, which is incorporated herein in its entirety. The transgenes, expression cassettes, and vectors of the disclosure are useful for both *in vivo* (*e.g.* systemic, particularly intravenous use) and also *ex vivo* use. LAMP-2B transgene and a functional promoter can be used to *ex vivo* gene-correct patients' autologous hematopoietic stem and progenitor cells (HSPCs), which can then be re-transplanted in the patients to repopulate their bone marrow, which is a reservoir of "healthy" cells for the rest of the life of the patients. In some embodiments, lentiviral vectors are used for *ex vivo* gene corrected, but other non-viral or viral vectors may be used in place of a lentiviral vector. The disclosure are envisions allogeneic transplant using donor HSPCs. In some embodiments, the lentiviral vector is a self-inactivating (SIN) lentivirus vector. In some embodiments, the HSPCs are derived from peripheral blood mobilized using, *e.g.*, granulocyte-colony stimulating factor (G-CSF) and/or plerixafor.

In another aspect, the disclosure provides methods of preventing, mitigating, ameliorating, reducing, inhibiting, eliminating and/or reversing one or more symptoms of Danon disease or another autophagy disorder in a subject in need thereof, comprising administering to the subject a gene therapy vector of the disclosure. The term "Danon disease" refers to an X-linked dominant skeletal and cardiac muscle disorder with

multisystem clinical manifestations. Danon disease mutations lead to an absence of lysosome-associated membrane protein 2 (LAMP-2) protein expression. Major clinical features include skeletal and cardiac myopathy, cardiac conduction abnormalities, cognitive difficulties, and retinal disease. Men are typically affected earlier and more severely than women.

In an embodiment, the vector is administered via a route selected from the group consisting of parenteral, intravenous, intra-arterial, intracardiac, intracoronary, intramyocardial, intrarenal, intraurethral, epidural, and intramuscular. In an embodiment, the vector is administered multiple times. In an embodiment, the vector is administered by intramuscular injection of the vector. In an embodiment, the vector is administered by injection of the vector into skeletal muscle. In an embodiment, the expression cassette comprises a muscle-specific promoter, optionally a muscle creatine kinase (MCK) promoter or a MCK/SV40 hybrid promoter as described in Takeshita et al. Muscle creatine kinase/SV40 hybrid promoter for muscle-targeted long-term transgene expression. *Int J Mol Med.* 2007 Feb;19(2):309-15. In an embodiment, the vector is administered by intracardiac injection.

In an embodiment, the vector, e.g., AAV vector, is administered systemically, and more particularly, intravenously. Advantageously, the vector is administered at a dose (in vg per mL, vg/kg body mass, or vg/min/kg) less than the dose required to observe the same response when an original or wild-type LAMP-2B sequence is used. In particular embodiments, the vector is an AAV2/9 vector comprising an expression cassette comprising a polynucleotide encoding LAMP-2B disclosed herein.

In some embodiments, the disclosure provides a method of expressing LAMP-2B in a subject, comprising systemically administering an adeno-associated viral (AAV) vector to the subject, wherein the AAV vector comprises an expression cassette comprising a transgene sharing at least 95% identity with SEQ ID NO: 3 or is identical to SEQ ID NO: 3, the transgene operatively linked to an enhancer/promoter region, wherein systemic administration of the AAV vector to the subject results in increased expression of LAMP-2B compared to expression of LAMP-2B prior to administration of the AAV vector or expression of LAMP-2B in an untreated control subject. In some embodiments, the AAV vector is an AAV2/9 vector. In particular embodiments, the expression cassette comprises

any of the elements disclosed herein. In some embodiment, systemic administration comprises intravenous administration. In some embodiments, the subject is exhibiting symptoms of Danon disease. In some embodiments, the subject suffers from, or is at risk for, Danon disease.

In some embodiments, the AAV vector is administered at a dose of between about 1×10^{12} and 5×10^{14} vector genomes (vg) of the AAV vector per kilogram (vg) of total body mass of the subject (vg/kg). In some embodiments, the AAV vector is administered at a dose of between about 1×10^{13} and 5×10^{14} vg/kg. In some embodiments, the AAV vector is administered at a dose of between about 5×10^{13} and 3×10^{14} vg/kg. In some embodiments, the AAV vector is administered at a dose of between about 5×10^{13} and 1×10^{14} vg/kg. In some embodiments, the AAV vector is administered at a dose of less than about 1×10^{12} vg/kg, less than about 3×10^{12} vg/kg, less than about 5×10^{12} vg/kg, less than about 7×10^{12} vg/kg, less than about 1×10^{13} vg/kg, less than about 3×10^{13} vg/kg, less than about 5×10^{13} vg/kg, less than about 7×10^{13} vg/kg, less than about 1×10^{14} vg/kg, less than about 3×10^{14} vg/kg, less than about 5×10^{14} vg/kg, less than about 7×10^{14} vg/kg, less than about 1×10^{15} vg/kg, less than about 3×10^{15} vg/kg, less than about 5×10^{15} vg/kg, or less than about 7×10^{15} vg/kg.

In some embodiments, the AAV vector is administered at a dose of about 1×10^{12} vg/kg, about 3×10^{12} vg/kg, about 5×10^{12} vg/kg, about 7×10^{12} vg/kg, about 1×10^{13} vg/kg, about 3×10^{13} vg/kg, about 5×10^{13} vg/kg, about 7×10^{13} vg/kg, about 1×10^{14} vg/kg, about 3×10^{14} vg/kg, about 5×10^{14} vg/kg, about 7×10^{14} vg/kg, about 1×10^{15} vg/kg, about 3×10^{15} vg/kg, about 5×10^{15} vg/kg, or about 7×10^{15} vg/kg.

In some embodiments, the AAV vector is administered at a dose of 1×10^{12} vg/kg, 3×10^{12} vg/kg, 5×10^{12} vg/kg, 7×10^{12} vg/kg, 1×10^{13} vg/kg, 3×10^{13} vg/kg, 5×10^{13} vg/kg, 7×10^{13} vg/kg, 1×10^{14} vg/kg, 3×10^{14} vg/kg, 5×10^{14} vg/kg, 7×10^{14} vg/kg, 1×10^{15} vg/kg, 3×10^{15} vg/kg, 5×10^{15} vg/kg, or 7×10^{15} vg/kg.

In some embodiments, the lentiviral vector is administered at a dose of between about 1×10^{12} and 5×10^{14} vector genomes (vg) of the lentiviral vector per kilogram (vg) of total body mass of the subject (vg/kg). In some embodiments, the lentiviral vector is administered at a dose of between about 1×10^{13} and 5×10^{14} vg/kg. In some embodimets, the lentiviral vector is administered at a dose of between about 5×10^{13} and 3×10^{14} vg/kg. In some embodiments, the lentiviral vector is administered at a dose of between about 5×10^{13} and

1×10^{14} vg/kg. In some embodiments, the lentiviral vector is administered at a dose of less than about 1×10^{12} vg/kg, less than about 3×10^{12} vg/kg, less than about 5×10^{12} vg/kg, less than about 7×10^{12} vg/kg, less than about 1×10^{13} vg/kg, less than about 3×10^{13} vg/kg, less than about 5×10^{13} vg/kg, less than about 7×10^{13} vg/kg, less than about 1×10^{14} vg/kg, less than about 3×10^{14} vg/kg, less than about 5×10^{14} vg/kg, less than about 7×10^{14} vg/kg, less than about 1×10^{15} vg/kg, less than about 3×10^{15} vg/kg, less than about 5×10^{15} vg/kg, or less than about 7×10^{15} vg/kg.

In some embodiments, the lentiviral vector is administered at a dose of about 1×10^{12} vg/kg, about 3×10^{12} vg/kg, about 5×10^{12} vg/kg, about 7×10^{12} vg/kg, about 1×10^{13} vg/kg, about 3×10^{13} vg/kg, about 5×10^{13} vg/kg, about 7×10^{13} vg/kg, about 1×10^{14} vg/kg, about 3×10^{14} vg/kg, about 5×10^{14} vg/kg, about 7×10^{14} vg/kg, about 1×10^{15} vg/kg, about 3×10^{15} vg/kg, about 5×10^{15} vg/kg, or about 7×10^{15} vg/kg.

In some embodiments, the lentiviral vector is administered at a dose of 1×10^{12} vg/kg, 3×10^{12} vg/kg, 5×10^{12} vg/kg, 7×10^{12} vg/kg, 1×10^{13} vg/kg, 3×10^{13} vg/kg, 5×10^{13} vg/kg, 7×10^{13} vg/kg, 1×10^{14} vg/kg, 3×10^{14} vg/kg, 5×10^{14} vg/kg, 7×10^{14} vg/kg, 1×10^{15} vg/kg, 3×10^{15} vg/kg, 5×10^{15} vg/kg, or 7×10^{15} vg/kg.

In some embodiments, the viral vector is administered at a dose of between about 1×10^{12} and 5×10^{14} vector genomes (vg) of the viral vector per kilogram (vg) of total body mass of the subject (vg/kg). In some embodiments, the viral vector is administered at a dose of between about 1×10^{13} and 5×10^{14} vg/kg. In some embodimets, the viral vector is administered at a dose of between about 5×10^{13} and 3×10^{14} vg/kg. In some embodiments, the viral vector is administered at a dose of between about 5×10^{13} and 1×10^{14} vg/kg. In some embodiments, the viral vector is administered at a dose of less than about 1×10^{12} vg/kg, less than about 3×10^{12} vg/kg, less than about 5×10^{12} vg/kg, less than about 7×10^{12} vg/kg, less than about 1×10^{13} vg/kg, less than about 3×10^{13} vg/kg, less than about 5×10^{13} vg/kg, less than about 7×10^{13} vg/kg, less than about 1×10^{14} vg/kg, less than about 3×10^{14} vg/kg, less than about 5×10^{14} vg/kg, less than about 7×10^{14} vg/kg, less than about 1×10^{15} vg/kg, less than about 3×10^{15} vg/kg, less than about 5×10^{15} vg/kg, or less than about 7×10^{15} vg/kg.

In some embodiments, the viral vector is administered at a dose of about 1×10^{12} vg/kg, about 3×10^{12} vg/kg, about 5×10^{12} vg/kg, about 7×10^{12} vg/kg, about 1×10^{13} vg/kg, about 3×10^{13} vg/kg, about 5×10^{13} vg/kg, about 7×10^{13} vg/kg, about 1×10^{14} vg/kg, about

3×10^{14} vg/kg, about 5×10^{14} vg/kg, about 7×10^{14} vg/kg, about 1×10^{15} vg/kg, about 3×10^{15} vg/kg, about 5×10^{15} vg/kg, or about 7×10^{15} vg/kg.

In some embodiments, the viral vector is administered at a dose of 1×10^{12} vg/kg, 3×10^{12} vg/kg, 5×10^{12} vg/kg, 7×10^{12} vg/kg, 1×10^{13} vg/kg, 3×10^{13} vg/kg, 5×10^{13} vg/kg, 7×10^{13} vg/kg, 1×10^{14} vg/kg, 3×10^{14} vg/kg, 5×10^{14} vg/kg, 7×10^{14} vg/kg, 1×10^{15} vg/kg, 3×10^{15} vg/kg, 5×10^{15} vg/kg, or 7×10^{15} vg/kg.

In some embodiments, the AAV vector is administered systemically at a dose of between about 1×10^{12} and 5×10^{14} vector genomes (vg) of the AAV vector per kilogram (vg) of total body mass of the subject (vg/kg). In some embodiments, the AAV vector is administered systemically at a dose of between about 1×10^{13} and 5×10^{14} vg/kg. In some embodiments, the AAV vector is administered systemically at a dose of between about 5×10^{13} and 3×10^{14} vg/kg. In some embodiments, the AAV vector is administered systemically at a dose of between about 5×10^{13} and 1×10^{14} vg/kg. In some embodiments, the AAV vector is administered systemically at a dose of less than about 1×10^{12} vg/kg, less than about 3×10^{12} vg/kg, less than about 5×10^{12} vg/kg, less than about 7×10^{12} vg/kg, less than about 1×10^{13} vg/kg, less than about 3×10^{13} vg/kg, less than about 5×10^{13} vg/kg, less than about 7×10^{13} vg/kg, less than about 1×10^{14} vg/kg, less than about 3×10^{14} vg/kg, less than about 5×10^{14} vg/kg, less than about 7×10^{14} vg/kg, less than about 1×10^{15} vg/kg, less than about 3×10^{15} vg/kg, less than about 5×10^{15} vg/kg, or less than about 7×10^{15} vg/kg.

In some embodiments, the AAV vector is administered systemically at a dose of about 1×10^{12} vg/kg, about 3×10^{12} vg/kg, about 5×10^{12} vg/kg, about 7×10^{12} vg/kg, about 1×10^{13} vg/kg, about 3×10^{13} vg/kg, about 5×10^{13} vg/kg, about 7×10^{13} vg/kg, about 1×10^{14} vg/kg, about 3×10^{14} vg/kg, about 5×10^{14} vg/kg, about 7×10^{14} vg/kg, about 1×10^{15} vg/kg, about 3×10^{15} vg/kg, about 5×10^{15} vg/kg, or about 7×10^{15} vg/kg.

In some embodiments, the AAV vector is administered systemically at a dose of 1×10^{12} vg/kg, 3×10^{12} vg/kg, 5×10^{12} vg/kg, 7×10^{12} vg/kg, 1×10^{13} vg/kg, 3×10^{13} vg/kg, 5×10^{13} vg/kg, 7×10^{13} vg/kg, 1×10^{14} vg/kg, 3×10^{14} vg/kg, 5×10^{14} vg/kg, 7×10^{14} vg/kg, 1×10^{15} vg/kg, 3×10^{15} vg/kg, 5×10^{15} vg/kg, or 7×10^{15} vg/kg.

In some embodiments, the lentiviral vector is administered systemically at a dose of between about 1×10^{12} and 5×10^{14} vector genomes (vg) of the lentiviral vector per kilogram (vg) of total body mass of the subject (vg/kg). In some embodiments, the lentiviral vector is

administered systemically at a dose of between about 1×10^{13} and 5×10^{14} vg/kg. In some embodimets, the lentiviral vector is administered systemically at a dose of between about 5×10^{13} and 3×10^{14} vg/kg. In some embodiments, the lentiviral vector is administered systemically at a dose of between about 5×10^{13} and 1×10^{14} vg/kg. In some embodiments, the lentiviral vector is administered systemically at a dose of less than about 1×10^{12} vg/kg, less than about 3×10^{12} vg/kg, less than about 5×10^{12} vg/kg, less than about 7×10^{12} vg/kg, less than about 1×10^{13} vg/kg, less than about 3×10^{13} vg/kg, less than about 5×10^{13} vg/kg, less than about 7×10^{13} vg/kg, less than about 1×10^{14} vg/kg, less than about 3×10^{14} vg/kg, less than about 5×10^{14} vg/kg, less than about 7×10^{14} vg/kg, less than about 1×10^{15} vg/kg, less than about 3×10^{15} vg/kg, less than about 5×10^{15} vg/kg, or less than about 7×10^{15} vg/kg.

In some embodiments, the lentiviral vector is administered systemically at a dose of about 1×10^{12} vg/kg, about 3×10^{12} vg/kg, about 5×10^{12} vg/kg, about 7×10^{12} vg/kg, about 1×10^{13} vg/kg, about 3×10^{13} vg/kg, about 5×10^{13} vg/kg, about 7×10^{13} vg/kg, about 1×10^{14} vg/kg, about 3×10^{14} vg/kg, about 5×10^{14} vg/kg, about 7×10^{14} vg/kg, about 1×10^{15} vg/kg, about 3×10^{15} vg/kg, about 5×10^{15} vg/kg, or about 7×10^{15} vg/kg.

In some embodiments, the lentiviral vector is administered systemically at a dose of 1×10^{12} vg/kg, 3×10^{12} vg/kg, 5×10^{12} vg/kg, 7×10^{12} vg/kg, 1×10^{13} vg/kg, 3×10^{13} vg/kg, 5×10^{13} vg/kg, 7×10^{13} vg/kg, 1×10^{14} vg/kg, 3×10^{14} vg/kg, 5×10^{14} vg/kg, 7×10^{14} vg/kg, 1×10^{15} vg/kg, 3×10^{15} vg/kg, 5×10^{15} vg/kg, or 7×10^{15} vg/kg.

In some embodiments, the viral vector is administered systemically at a dose of between about 1×10^{12} and 5×10^{14} vector genomes (vg) of the viral vector per kilogram (vg) of total body mass of the subject (vg/kg). In some embodiments, the viral vector is administered systemically at a dose of between about 1×10^{13} and 5×10^{14} vg/kg. In some embodimets, the viral vector is administered systemically at a dose of between about 5×10^{13} and 3×10^{14} vg/kg. In some embodiments, the viral vector is administered systemically at a dose of between about 5×10^{13} and 1×10^{14} vg/kg. In some embodiments, the viral vector is administered systemically at a dose of less than about 1×10^{12} vg/kg, less than about 3×10^{12} vg/kg, less than about 5×10^{12} vg/kg, less than about 7×10^{12} vg/kg, less than about 1×10^{13} vg/kg, less than about 3×10^{13} vg/kg, less than about 5×10^{13} vg/kg, less than about 7×10^{13} vg/kg, less than about 1×10^{14} vg/kg, less than about 3×10^{14} vg/kg, less than about 5×10^{14}

vg/kg, less than about 7×10^{14} vg/kg, less than about 1×10^{15} vg/kg, less than about 3×10^{15} vg/kg, less than about 5×10^{15} vg/kg, or less than about 7×10^{15} vg/kg.

In some embodiments, the viral vector is administered systemically at a dose of about 1×10^{12} vg/kg, about 3×10^{12} vg/kg, about 5×10^{12} vg/kg, about 7×10^{12} vg/kg, about 1×10^{13} vg/kg, about 3×10^{13} vg/kg, about 5×10^{13} vg/kg, about 7×10^{13} vg/kg, about 1×10^{14} vg/kg, about 3×10^{14} vg/kg, about 5×10^{14} vg/kg, about 7×10^{14} vg/kg, about 1×10^{15} vg/kg, about 3×10^{15} vg/kg, about 5×10^{15} vg/kg, or about 7×10^{15} vg/kg.

In some embodiments, the viral vector is administered systemically at a dose of 1×10^{12} vg/kg, 3×10^{12} vg/kg, 5×10^{12} vg/kg, 7×10^{12} vg/kg, 1×10^{13} vg/kg, 3×10^{13} vg/kg, 5×10^{13} vg/kg, 7×10^{13} vg/kg, 1×10^{14} vg/kg, 3×10^{14} vg/kg, 5×10^{14} vg/kg, 7×10^{14} vg/kg, 1×10^{15} vg/kg, 3×10^{15} vg/kg, 5×10^{15} vg/kg, or 7×10^{15} vg/kg.

In some embodiments, the AAV vector is administered intravenously at a dose of between about 1×10^{12} and 5×10^{14} vector genomes (vg) of the AAV vector per kilogram (vg) of total body mass of the subject (vg/kg). In some embodiments, the AAV vector is administered intravenously at a dose of between about 1×10^{13} and 5×10^{14} vg/kg. In some embodiments, the AAV vector is administered intravenously at a dose of between about 5×10^{13} and 3×10^{14} vg/kg. In some embodiments, the AAV vector is administered intravenously at a dose of between about 5×10^{13} and 1×10^{14} vg/kg. In some embodiments, the AAV vector is administered intravenously at a dose of less than about 1×10^{12} vg/kg, less than about 3×10^{12} vg/kg, less than about 5×10^{12} vg/kg, less than about 7×10^{12} vg/kg, less than about 1×10^{13} vg/kg, less than about 3×10^{13} vg/kg, less than about 5×10^{13} vg/kg, less than about 7×10^{13} vg/kg, less than about 1×10^{14} vg/kg, less than about 3×10^{14} vg/kg, less than about 5×10^{14} vg/kg, less than about 7×10^{14} vg/kg, less than about 1×10^{15} vg/kg, less than about 3×10^{15} vg/kg, less than about 5×10^{15} vg/kg, or less than about 7×10^{15} vg/kg.

In some embodiments, the AAV vector is administered intravenously at a dose of about 1×10^{12} vg/kg, about 3×10^{12} vg/kg, about 5×10^{12} vg/kg, about 7×10^{12} vg/kg, about 1×10^{13} vg/kg, about 3×10^{13} vg/kg, about 5×10^{13} vg/kg, about 7×10^{13} vg/kg, about 1×10^{14} vg/kg, about 3×10^{14} vg/kg, about 5×10^{14} vg/kg, about 7×10^{14} vg/kg, about 1×10^{15} vg/kg, about 3×10^{15} vg/kg, about 5×10^{15} vg/kg, or about 7×10^{15} vg/kg.

In some embodiments, the AAV vector is administered intravenously at a dose of 1×10^{12} vg/kg, 3×10^{12} vg/kg, 5×10^{12} vg/kg, 7×10^{12} vg/kg, 1×10^{13} vg/kg, 3×10^{13} vg/kg, 5×10^{13}

vg/kg, 7×10^{13} vg/kg, 1×10^{14} vg/kg, 3×10^{14} vg/kg, 5×10^{14} vg/kg, 7×10^{14} vg/kg, 1×10^{15} vg/kg, 3×10^{15} vg/kg, 5×10^{15} vg/kg, or 7×10^{15} vg/kg.

In some embodiments, the lentiviral vector is administered intravenously at a dose of between about 1×10^{12} and 5×10^{14} vector genomes (vg) of the lentiviral vector per kilogram (vg) of total body mass of the subject (vg/kg). In some embodiments, the lentiviral vector is administered intravenously at a dose of between about 1×10^{13} and 5×10^{14} vg/kg. In some embodimets, the lentiviral vector is administered intravenously at a dose of between about 5×10^{13} and 3×10^{14} vg/kg. In some embodiments, the lentiviral vector is administered intravenously at a dose of between about 5×10^{13} and 1×10^{14} vg/kg. In some embodiments, the lentiviral vector is administered intravenously at a dose of less than about 1×10^{12} vg/kg, less than about 3×10^{12} vg/kg, less than about 5×10^{12} vg/kg, less than about 7×10^{12} vg/kg, less than about 1×10^{13} vg/kg, less than about 3×10^{13} vg/kg, less than about 5×10^{13} vg/kg, less than about 7×10^{13} vg/kg, less than about 1×10^{14} vg/kg, less than about 3×10^{14} vg/kg, less than about 5×10^{14} vg/kg, less than about 7×10^{14} vg/kg, less than about 1×10^{15} vg/kg, less than about 3×10^{15} vg/kg, less than about 5×10^{15} vg/kg, or less than about 7×10^{15} vg/kg.

In some embodiments, the lentiviral vector is administered intravenously at a dose of about 1×10^{12} vg/kg, about 3×10^{12} vg/kg, about 5×10^{12} vg/kg, about 7×10^{12} vg/kg, about 1×10^{13} vg/kg, about 3×10^{13} vg/kg, about 5×10^{13} vg/kg, about 7×10^{13} vg/kg, about 1×10^{14} vg/kg, about 3×10^{14} vg/kg, about 5×10^{14} vg/kg, about 7×10^{14} vg/kg, about 1×10^{15} vg/kg, about 3×10^{15} vg/kg, about 5×10^{15} vg/kg, or about 7×10^{15} vg/kg.

In some embodiments, the lentiviral vector is administered intravenously at a dose of 1×10^{12} vg/kg, 3×10^{12} vg/kg, 5×10^{12} vg/kg, 7×10^{12} vg/kg, 1×10^{13} vg/kg, 3×10^{13} vg/kg, 5×10^{13} vg/kg, 7×10^{13} vg/kg, 1×10^{14} vg/kg, 3×10^{14} vg/kg, 5×10^{14} vg/kg, 7×10^{14} vg/kg, 1×10^{15} vg/kg, 3×10^{15} vg/kg, 5×10^{15} vg/kg, or 7×10^{15} vg/kg.

In some embodiments, the viral vector is administered intravenously at a dose of between about 1×10^{12} and 5×10^{14} vector genomes (vg) of the viral vector per kilogram (vg) of total body mass of the subject (vg/kg). In some embodiments, the viral vector is administered intravenously at a dose of between about 1×10^{13} and 5×10^{14} vg/kg. In some embodimets, the viral vector is administered intravenously at a dose of between about 5×10^{13} and 3×10^{14} vg/kg. In some embodiments, the viral vector is administered intravenously at a dose of between about 5×10^{13} and 1×10^{14} vg/kg. In some embodiments, the viral vector is

administered intravenously at a dose of less than about 1×10^{12} vg/kg, less than about 3×10^{12} vg/kg, less than about 5×10^{12} vg/kg, less than about 7×10^{12} vg/kg, less than about 1×10^{13} vg/kg, less than about 3×10^{13} vg/kg, less than about 5×10^{13} vg/kg, less than about 7×10^{13} vg/kg, less than about 1×10^{14} vg/kg, less than about 3×10^{14} vg/kg, less than about 5×10^{14} vg/kg, less than about 7×10^{14} vg/kg, less than about 1×10^{15} vg/kg, less than about 3×10^{15} vg/kg, less than about 5×10^{15} vg/kg, or less than about 7×10^{15} vg/kg.

In some embodiments, the viral vector is administered intravenously at a dose of about 1×10^{12} vg/kg, about 3×10^{12} vg/kg, about 5×10^{12} vg/kg, about 7×10^{12} vg/kg, about 1×10^{13} vg/kg, about 3×10^{13} vg/kg, about 5×10^{13} vg/kg, about 7×10^{13} vg/kg, about 1×10^{14} vg/kg, about 3×10^{14} vg/kg, about 5×10^{14} vg/kg, about 7×10^{14} vg/kg, about 1×10^{15} vg/kg, about 3×10^{15} vg/kg, about 5×10^{15} vg/kg, or about 7×10^{15} vg/kg.

In some embodiments, the viral vector is administered intravenously at a dose of 1×10^{12} vg/kg, 3×10^{12} vg/kg, 5×10^{12} vg/kg, 7×10^{12} vg/kg, 1×10^{13} vg/kg, 3×10^{13} vg/kg, 5×10^{13} vg/kg, 7×10^{13} vg/kg, 1×10^{14} vg/kg, 3×10^{14} vg/kg, 5×10^{14} vg/kg, 7×10^{14} vg/kg, 1×10^{15} vg/kg, 3×10^{15} vg/kg, 5×10^{15} vg/kg, or 7×10^{15} vg/kg.

Systemic (or more particularly intravenous) administration in some embodiments results in expression of LAMP-2B polynucleotide as mRNA, in the form of an mRNA expressed from the transgene, in one or more tissues (*e.g.* heart, muscle, and/or liver) of the subject. In some embodiments, expression of the LAMP-2B polynucleotide as mRNA is increased at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 1.6-fold, at least about 1.7-fold, at least about 1.8-fold, at least about 1.9-fold, at least about 2.0-fold, at least about 2.2-fold, at least about 2.3-fold, at least about 2.4-fold, at least about 2.5-fold, at least about 3-fold, or at least about 4-fold in the heart compared to expression in an untreated subject or a subject treated with a control vector. In some embodiments, expression of LAMP-2B polynucleotide as mRNA is increased at least 1.2-fold, at least 1.3-fold, at least 1.4-fold, at least 1.5-fold, at least 1.6-fold, at least 1.7-fold, at least 1.8-fold, at least 1.9-fold, at least 2.0-fold, at least 2.2-fold, at least 2.3-fold, at least 2.4-fold, at least 2.5-fold, at least 3-fold, or at least 4-fold in the heart compared to expression in an untreated subject or a subject treated with a control vector. In some embodiments, expression of LAMP-2B polynucleotide as mRNA is increased 1.2-fold, 1.3-fold, 1.4-fold, 1.5-fold, 1.6-fold, 1.7-fold, 1.8-fold, 1.9-fold, 2.0-fold, 2.2-fold, 2.3-fold, 2.4-fold, 2.5-fold,

3-fold, or 4-fold in the heart compared to expression in an untreated subject or a subject treated with a control vector.

In some embodiments, expression of LAMP-2B polynucleotide as mRNA is increased at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 1.6-fold, at least about 1.7-fold, at least about 1.8-fold, at least about 1.9-fold, at least about 2.0-fold, at least about 2.2-fold, at least about 2.3-fold, at least about 2.4-fold, at least about 2.5-fold, at least about 3-fold, or at least about 4-fold in the muscle compared to expression in an untreated subject or a subject treated with a control vector. In some embodiments, expression of LAMP-2B polynucleotide as mRNA is increased at least 1.2-fold, at least 1.3-fold, at least 1.4-fold, at least 1.5-fold, at least 1.6-fold, at least 1.7-fold, at least 1.8-fold, at least 1.9-fold, at least 2.0-fold, at least 2.2-fold, at least 2.3-fold, at least 2.4-fold, at least 2.5-fold, at least 3-fold, or at least 4-fold in the muscle compared to expression in an untreated subject or a subject treated with a control vector. In some embodiments, expression of LAMP-2B polynucleotide as mRNA is increased 1.2-fold, 1.3-fold, 1.4-fold, 1.5-fold, 1.6-fold, 1.7-fold, 1.8-fold, 1.9-fold, 2.0-fold, 2.2-fold, 2.3-fold, 2.4-fold, 2.5-fold, 3-fold, or 4-fold in the muscle compared to expression in an untreated subject or a subject treated with a control vector.

In some embodiments, the LAMP-2B transgene is expressed in the heart and not expressed in the liver of the subject. In some embodiments, expression of LAMP-2B polynucleotide as mRNA is observed to be at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 1.6-fold, at least about 1.7-fold, at least about 1.8-fold, at least about 1.9-fold, at least about 2.0-fold, at least about 2.2-fold, at least about 2.3-fold, at least about 2.4-fold, at least about 2.5-fold, at least about 3-fold, or at least about 4-fold in the heart compared to the liver. In some embodiments, expression of LAMP-2B polynucleotide as mRNA is observed to be at least 1.2-fold, at least 1.3-fold, at least 1.4-fold, at least 1.5-fold, at least 1.6-fold, at least 1.7-fold, at least 1.8-fold, at least 1.9-fold, at least 2.0-fold, at least 2.2-fold, at least 2.3-fold, at least 2.4-fold, at least 2.5-fold, at least 3-fold, or at least 4-fold in the heart compared to the liver. In some embodiments, expression of LAMP-2B polynucleotide as mRNA is observed to be 1.2-fold, 1.3-fold, 1.4-fold, 1.5-fold, 1.6-fold, 1.7-fold, 1.8-fold, 1.9-fold, 2.0-fold, 2.2-fold, 2.3-fold, 2.4-fold, 2.5-fold, 3-fold, or 4-fold in the heart compared to the liver.

In some embodiments, expression of wild-type or functional LAMP-2B protein is increased at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 1.6-fold, at least about 1.7-fold, at least about 1.8-fold, at least about 1.9-fold, at least about 2.0-fold, at least about 2.2-fold, at least about 2.3-fold, at least about 2.4-fold, at least about 2.5-fold, at least about 3-fold, or at least about 4-fold in the heart compared to expression in an untreated subject or a subject treated with a control vector. In some embodiments, expression of wild-type or functional LAMP-2B protein is increased at least 1.2-fold, at least 1.3-fold, at least 1.4-fold, at least 1.5-fold, at least 1.6-fold, at least 1.7-fold, at least 1.8-fold, at least 1.9-fold, at least 2.0-fold, at least 2.2-fold, at least 2.3-fold, at least 2.4-fold, at least 2.5-fold, at least 3-fold, or at least 4-fold in the heart compared to expression in an untreated subject or a subject treated with a control vector. In some embodiments, expression of wild-type or functional LAMP-2B protein is increased 1.2-fold, 1.3-fold, 1.4-fold, 1.5-fold, 1.6-fold, 1.7-fold, 1.8-fold, 1.9-fold, 2.0-fold, 2.2-fold, 2.3-fold, 2.4-fold, 2.5-fold, 3-fold, or 4-fold in the heart compared to expression in an untreated subject or a subject treated with a control vector.

In some embodiments, expression of wild-type or functional LAMP-2B protein is observed to be at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 1.6-fold, at least about 1.7-fold, at least about 1.8-fold, at least about 1.9-fold, at least about 2.0-fold, at least about 2.2-fold, at least about 2.3-fold, or at least 5-fold, in the heart compared to the liver. In some embodiments, expression of wild-type or functional LAMP-2B protein is observed to be at least 1.2-fold, at least 1.3-fold, at least 1.4-fold, at least 1.5-fold, at least 1.6-fold, at least 1.7-fold, at least 1.8-fold, at least 1.9-fold, at least 2.0-fold, at least 2.2-fold, at least 2.3-fold, at least 2.4-fold, at least 2.5-fold, at least 3-fold, or at least 4-fold in the heart compared to the liver. In some embodiments, expression of wild-type or functional LAMP-2B protein is observed to be 1.2-fold, 1.3-fold, 1.4-fold, 1.5-fold, 1.6-fold, 1.7-fold, 1.8-fold, 1.9-fold, 2.0-fold, 2.2-fold, 2.3-fold, 2.4-fold, 2.5-fold, 3-fold, or 4-fold in the heart compared to the liver.

In some embodiments, administration of the gene therapy vector results in expression of wild-type or functional LAMP-2B protein in the liver of at most about 1.1-fold, at most about 1.2-fold, at most about 1.3-fold, at most about 1.4-fold, at most about 1.5-fold, at most about 1.6-fold, at most about 1.7-fold, at most about 1.8-fold, at most about 1.9-fold, or at most about 2-fold increased compared to expression in the liver of an untreated subject. In

some embodiments, administration of the gene therapy vector results in expression of wild-type or functional LAMP-2B protein in the liver of at most 1.1-fold, at most 1.2-fold, at most 1.3-fold, at most 1.4-fold, at most 1.5-fold, at most 1.6-fold, at most 1.7-fold, at most 1.8-fold, at most 1.9-fold, or at most 2-fold increased compared to expression in the liver of an untreated subject. In some embodiments, administration of the gene therapy vector results in expression of wild-type or functional LAMP-2B protein in the liver of 1.1-fold, 1.2-fold, 1.3-fold, 1.4-fold, 1.5-fold, 1.6-fold, 1.7-fold, 1.8-fold, 1.9-fold, or 2-fold increased compared to expression in the liver of an untreated subject.

In an embodiment, the disclosure provides a method of treating a disease or disorder, optionally Danon disease, in a subject in need thereof, comprising contacting cells with a gene therapy vector according to the present disclosure and administering the cells to the subject. In an embodiment, the cells are stem cells, optionally pluripotent stem cells. In an embodiment, the stem cells are capable of differentiation into cardiac tissue. In an embodiment, the stem cells are capable of differentiation into muscle tissue, *e.g.*, cardiac muscle tissue and/or skeletal muscle tissue. In an embodiment, the stem cells are autologous. In an embodiment, the stem cells are induced pluripotent stem cells (iPSCs).

In an embodiment, the disease or disorder is an autophagy disorder. In some embodiments, the autophagy disorder is selected from the group consisting of, but not limited to, end-stage heart failure, myocardial infarction, drug toxicities, diabetes, end-stage renal failure, and aging. In an embodiment, the subject is a mammal, *e.g.*, a human. In an embodiment, the subject is exhibiting symptoms of Danon disease or another autophagy disorder. In an embodiment, the subject has been identified as having reduced or non-detectable LAMP-2 expression. In an embodiment, the subject has been identified as having a mutated LAMP-2 gene.

Subjects/patients amenable to treatment using the methods described herein include, but are not limited to, individuals at risk of a disease or disorder characterized by insufficient autophagic flux (*e.g.*, Danon disease as well as other known disorders of autophagy including, but not limited to, systolic and diastolic heart failure, myocardial infarction, drug toxicities (for example, anthracyclines chloroquine and its derivatives), diabetes, end-stage renal disease, and aging) but not showing symptoms, as well as subjects presently showing

symptoms. Such subject may have been identified as having a mutated LAMP-2 gene or as having reduced or non-detectable levels of LAMP-2 expression.

In some embodiments, the patient is a human. In some embodiments, the patient is a pediatric, adolescent, or adult human. In some embodiments, the patient is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 years old, or more than 20 years old. In some embodiments, the patient is 20 to 50 years old. In some embodiments, the patient is 50 to 65 years old. In some embodiments, the patient is 1 to 5, 2 to 6, 3 to 7, 4 to 8, 5 to 9, 6 to 10, 7 to 11, 8 to 12, 9 to 13, 10 to 14, 11 to 15, 12 to 16, 13 to 17, 14 to 18, 15 to 19, or 16 to 20 years old. In some embodiments, the patient is 5 to 6, 6 to 7, 7 to 8, 8 to 9, 9 to 10, 10 to 11, 11 to 12, 12 to 13, 13 to 14, 14 to 15, 15 to 16, 16 to 17, 17 to 18, 18 to 19, 19 to 20, or 20 to 21 years old. In a particular embodiment, the patient is 15 to 16 years old.

In some embodiments, the patient is a human male. In some embodiments, the patient is a pediatric, adolescent, or adult human male. In some embodiments, the patient is a 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 years old male, or a more than 20 years old male. In some embodiments, the patient is a 20 to 50 years old male. In some embodiments, the patient is a 50 to 65 years old male. In some embodiments, the patient is a 1 to 5, 2 to 6, 3 to 7, 4 to 8, 5 to 9, 6 to 10, 7 to 11, 8 to 12, 9 to 13, 10 to 14, 11 to 15, 12 to 16, 13 to 17, 14 to 18, 15 to 19, or 16 to 20 years old male. In some embodiments, the patient is a 5 to 6, 6 to 7, 7 to 8, 8 to 9, 9 to 10, 10 to 11, 11 to 12, 12 to 13, 13 to 14, 14 to 15, 15 to 16, 16 to 17, 17 to 18, 18 to 19, 19 to 20, or 20 to 21 year old male. In a particular embodiment, the patient is 15 to 16 years old.

In some embodiments, the patient is a human female. In some embodiments, the patient is a pediatric, adolescent, or adult human female. In some embodiments, the patient is a 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 years old female, or a more than 20 years old female. In some embodiments, the patient is a 20 to 50 years old female. In some embodiments, the patient is a 50 to 65 years old female.

In some embodiments, the subject is exhibiting symptoms of a disease or disorder characterized by insufficient autophagic flux (*e.g.*, Danon disease as well as other known disorders of autophagy including, but not limited to, systolic and diastolic heart failure, myocardial infarction, drug toxicities, diabetes, end-stage renal disease, and aging). The symptoms may be actively manifesting, or may be suppressed or controlled (*e.g.*, by

medication) or in remission. The subject may or may not have been diagnosed with the disorder, *e.g.*, by a qualified physician.

Definitions

The terms "lysosome-associated membrane protein 2" and "LAMP-2" interchangeably refer to nucleic acids and polypeptide polymorphic variants, alleles, mutants, and interspecies homologs that: (1) have an amino acid sequence that has greater than about 90% amino acid sequence identity, for example, 91 %, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater amino acid sequence identity, preferably over a region of at least about 25, 50, 100, 200, 300, 400, or more amino acids, or over the full-length, to an amino acid sequence encoded by a LAMP-2 nucleic acid (*see, e.g.*, GenBank Accession Nos. NM_002294.2 (isoform A). NM_013995.2 (isoform B), NM_001122606.1 (isoform C)) or to an amino acid sequence of a LAMP-2 polypeptide (*see e.g.*, GenBank Accession Nos. NP_002285.1 (isoform A), NP_054701.1 (isoform B), NP_001116078.1 (isoform C)); (2) bind to antibodies, *e.g.*, polyclonal antibodies, raised against an immunogen comprising an amino acid sequence of a LAMP-2 polypeptide (*e.g.*, LAMP-2 polypeptides described herein); or an amino acid sequence encoded by a LAMP-2 nucleic acid (*e.g.*, LAMP-2 polynucleotides described herein), and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to an anti-sense strand corresponding to a nucleic acid sequence encoding a LAMP-2 protein, and conservatively modified variants thereof; (4) have a nucleic acid sequence that has greater than about 90%, preferably greater than about 91 %, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher nucleotide sequence identity, preferably over a region of at least about 25, 50, 100, 200, 500, 1000, 2000 or more nucleotides, or over the full-length, to a LAMP-2 nucleic acid (*e.g.*, LAMP-2 polynucleotides, as described herein, and LAMP-2 polynucleotides that encode LAMP-2 polypeptides, as described herein).

The terms "lysosome-associated membrane protein 2B" and "LAMP-2B" interchangeably refer to nucleic acids and polypeptide polymorphic variants, alleles, mutants, and interspecies homologs that: (1) have an amino acid sequence that has greater than about 90% amino acid sequence identity, for example, 91 %, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater amino acid sequence identity, preferably over a region of at least about 25, 50, 100, 200, 300, 400, or more amino acids, or over the full-length, to an amino

acid sequence encoded by a LAMP-2B nucleic acid (*see e.g.,* NM_013995.2) or to an amino acid sequence of a LAMP-2B polypeptide (*see e.g.*, NP_054701.1); (2) bind to antibodies, *e.g.*, polyclonal antibodies, raised against an immunogen comprising an amino acid sequence of a LAMP-2B polypeptide (*e.g.*, LAMP-2B polypeptides described herein); or an amino acid sequence encoded by a LAMP-2B nucleic acid (*e.g.*, LAMP-2B polynucleotides described herein), and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to an anti-sense strand corresponding to a nucleic acid sequence encoding a LAMP-2B protein, and conservatively modified variants thereof; (4) have a nucleic acid sequence that has greater than about 90%, preferably greater than about 91 %, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher nucleotide sequence identity, preferably over a region of at least about 25, 50, 100, 200, 500, 1000, 2000 or more nucleotides, or over the full-length, to a LAMP-2B nucleic acid (*e.g.*, LAMP-2B polynucleotides, as described herein, and LAMP-2B polynucleotides that encode LAMP-2B polypeptides, as described herein).

The term "functional variant" in respect to a protein (*e.g.* a LAMP-2B) refers to a polypeptide sequence, or a fragment of a polypeptide sequence having at least about 30, at least about 40, at least about 50, at least about 60, at least about 70, or at least about 80 amino acid resides, that retains one or more functional attributes of the protein. For example, a functional variant of LAMP-2B is a LAMP-2B (as defined herein) that retains one or more functions such as: (1) regulating human cardiomyocyte function (Chi et al. (2019) *PNAS USA* 116 (2) 556-565); (2) improving metabolic and physiological function in Danon disease (Adler et al. (2019) *J. Am. College Cardiology* S0735-1097(19)31295-1); and/or (3) autophagy (Rowland et al. (2016) *J. Cell Sci.* (2016) 129, 2135-2143).

LAMP-2B has a lumenal domain (residues 29-375), a transmembrane domain (residues 376-399), and a cytoplasmic domain (residues 400-410), *see* UniProt Accession No. P13473. LAMP-2B functions in include chaperone-mediated autophagy, a process that mediates lysosomal degradation of proteins in response to various stresses and as part of the normal turnover of proteins with a long biological half-live (Cuervo et al. *Science* 273:501-503(1996), Cuervo et al. *J. Cell Sci.* 113:4441-4450(2000), Bandyopadhyay et al. *Mol. Cell. Biol.* 28:5747-5763(2008), Li et al. *Exp. Cell Res.* 327:48-56(2014), Hubert et al. *Biol. Open* 5:1516-1529(2016)). LAMP-2B may target GAPDH and MLLT11 for lysosomal degradation. LAMP-2B may be required for the fusion of autophagosomes with lysosomes

during autophagy. It has been suggested that cells that lack LAMP2 express normal levels of VAMP8, but fail to accumulate STX17 on autophagosomes, which is the most likely explanation for the lack of fusion between autophagosomes and lysosomes. LAMP-2B may be required for normal degradation of the contents of autophagosomes. LAMP-2B may be required for efficient MHCII-mediated presentation of exogenous antigens via its function in lysosomal protein degradation; antigenic peptides generated by proteases in the endosomal/lysosomal compartment are captured by nascent MHCII subunits (Crotzer et al. *Immunology* 131:318-330(2010)).

Functional variants of LAMP-2B therefore include fragments of LAMP-2B that are capable of mediating any of the foregoing functions. In some embodiments, the function fragment of LAMP-2B includes one or more of the lumenal, transmembrane, and cytoplasmic domains. In some embodiments, the functional variant of LAMP-2B comprises one or more C-terminal or N-terminal deletions with respect to native LAMP-2B. In some embodiments, the functional variant of LAMP-2B comprises one or more internal deletions with respect to native LAMP-2B.

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (*i.e.* , share at least about 80% identity, for example, at least about 85%, 90%, 91 %, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identity over a specified region to a reference sequence, *e.g.*, LAMP-2 polynucleotide or polypeptide sequence as described herein, when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Such sequences are then said to be "substantially identical." This definition also refers to the compliment of a test sequence. Preferably, the identity exists over a region that is at least about 25 amino acids or nucleotides in length, for example, over a region that is 50, 100, 200, 300, 400 amino acids or nucleotides in length, or over the full-length of a reference sequence.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if

necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters. For sequence comparison of nucleic acids and proteins to LAMP-2 nucleic acids and proteins, the BLAST and BLAST 2.0 algorithms and the default parameters are used.

A "comparison window," as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith & Waterman, Adv. Appl. Math. 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol. 48:443 (1970), by the search for similarity method of Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wl), or by manual alignment and visual inspection (see, *e.g.*, Ausubel et al., eds., Current Protocols in Molecular Biology (1995 supplement)). Examples of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al., J. Mol. Biol. 215:403-410 (1990) and Altschul et al., Nucleic Acids Res. 25:3389-3402 (1977), respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (on the worldwide web at ncbi.nlm.nih.gov/).

An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions. Yet

another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequence.

As used herein, "administering" refers to local and systemic administration, *e.g.*, including enteral, parenteral, pulmonary, and topical/transdermal administration. Routes of administration for compounds (*e.g.*, polynucleotide encoding one or more LAMP- 2 isoforms) that find use in the methods described herein include, *e.g.*, oral (per os (P.O.)) administration, nasal or inhalation administration, administration as a suppository, topical contact, transdermal delivery (*e.g.*, via a transdermal patch), intrathecal (IT) administration, intravenous ("iv") administration, intraperitoneal ("ip") administration, intramuscular ("im") administration, intralesional administration, or subcutaneous ("sc") administration, or the implantation of a slow-release device *e.g.*, a mini-osmotic pump, a depot formulation, etc. , to a subject. Administration can be by any route including parenteral and transmucosal (*e.g.*, oral, nasal, vaginal, rectal, or transdermal). Parenteral administration includes, *e.g.*, intravenous, intramuscular, intraarterial, intrarenal, intraurethral, intracardiac, intracoronary, intramyocardial, intradermal, epidural, subcutaneous, intraperitoneal, intraventricular, ionophoretic and intracranial. Other modes of delivery include, but are not limited to, the use of liposomal formulations, intravenous infusion, transdermal patches, etc.

The terms "systemic administration" and "systemically administered" refer to a method of administering a compound or composition to a mammal so that the compound or composition is delivered to sites in the body, including the targeted site of pharmaceutical action, via the circulatory system. Systemic administration includes, but is not limited to, oral, intranasal, rectal and parenteral (*e.g.*, other than through the alimentary tract, such as intramuscular, intravenous, intra-arterial, transdermal and subcutaneous) administration.

The term "co-administering" or "concurrent administration", when used, for example with respect to the compounds (*e.g.*, LAMP-2 polynucleotides) and/or analogs thereof and another active agent, refers to administration of the compound and/or analogs and the active agent such that both can simultaneously achieve a physiological effect. The two agents, however, need not be administered together. In certain embodiments, administration of one agent can precede administration of the other. Simultaneous physiological effect need not necessarily require presence of both agents in the circulation at the same time. However, in certain embodiments, co-administering typically results in both agents being simultaneously

present in the body (*e.g.*, in the plasma) at a significant fraction (*e.g.*, 20% or greater, *e.g.*, 30% or 40% or greater, *e.g.*, 50% or 60% or greater, *e.g.*, 70% or 80% or 90% or greater) of their maximum serum concentration for any given dose.

The term "effective amount" or "pharmaceutically effective amount" refer to the amount and/or dosage, and/or dosage regime of one or more compounds (*e.g.*, gene therapy vectors) necessary to bring about the desired result *e.g.*, increased expression of one or more LAMP-2 isoforms in an amount sufficient to reduce the ultimate severity of a disease characterized by impaired or deficient autophagy (*e.g.*, Danon disease).

The phrase "cause to be administered" refers to the actions taken by a medical professional (*e.g.*, a physician), or a person controlling medical care of a subject, that control and/or permit the administration of the agent(s)/compound(s) at issue to the subject. Causing to be administered can involve diagnosis and/or determination of an appropriate therapeutic or prophylactic regimen, and/or prescribing particular agent(s)/compounds for a subject. Such prescribing can include, for example, drafting a prescription form, annotating a medical record, and the like.

As used herein, the terms "treating" and "treatment" refer to delaying the onset of, retarding or reversing the progress of, reducing the severity of, or alleviating or preventing either the disease or condition to which the term applies, or one or more symptoms of such disease or condition. The terms "treating" and "treatment" also include preventing, mitigating, ameliorating, reducing, inhibiting, eliminating and/or reversing one or more symptoms of the disease or condition.

The term "mitigating" refers to reduction or elimination of one or more symptoms of that pathology or disease, and/or a reduction in the rate or delay of onset or severity of one or more symptoms of that pathology or disease, and/or the prevention of that pathology or disease. In certain embodiments, the reduction or elimination of one or more symptoms of pathology or disease can include, *e.g.*, measurable and sustained increase in the expression levels of one or more isoforms of LAMP-2.

As used herein, the phrase "consisting essentially of refers to the genera or species of active pharmaceutical agents recited in a method or composition, and further can include other agents that, on their own do not have substantial activity for the recited indication or purpose.

The terms "subject," "individual," and "patient" interchangeably refer to a mammal, preferably a human or a non-human primate, but also domesticated mammals (*e.g.*, canine or feline), laboratory mammals (*e.g.*, mouse, rat, rabbit, hamster, guinea pig) and agricultural mammals (*e.g.*, equine, bovine, porcine, ovine). In various embodiments, the subject can be a human (*e.g.*, adult male, adult female, adolescent male, adolescent female, male child, female child).

The terms "gene transfer" or "gene delivery" refer to methods or systems for reliably inserting foreign DNA into host cells. Such methods can result in transient expression of non-integrated transferred DNA, extrachromosomal replication and expression of transferred replicons (*e.g.* episomes), or integration of transferred genetic material into the genomic DNA of host cells.

The term "vector" is used herein to refer to a nucleic acid molecule capable transferring or transporting another nucleic acid molecule. The transferred nucleic acid is generally linked to, *e.g.*, inserted into, the vector nucleic acid molecule. A vector may include sequences that direct autonomous replication or reverse transcription in a cell, or may include sequences sufficient to allow integration into host cell DNA. "vectors" include gene therapy vectors. As used herein, the term "gene therapy vector" refers to a vector capable of use in performing gene therapy, *e.g.*, delivering a polynucleotide sequence encoding a therapeutic polypeptide to a subject. Gene therapy vectors may comprise a nucleic acid molecule ("transgene") encoding a therapeutically active polypeptide, *e.g.*, a LAMP-2B or other gene useful for replacement gene therapy when introduced into a subject. Useful vectors include, but are not limited to, viral vectors.

As used herein, the term "expression cassette" refers to a DNA segment that is capable in an appropriate setting of driving the expression of a polynucleotide (a "transgene") encoding a therapeutically active polypeptide (*e.g.*, LAMP-2B) that is incorporated in said expression cassette. When introduced into a host cell, an expression cassette *inter alia* is capable of directing the cell's machinery to transcribe the transgene into RNA, which is then usually further processed and finally translated into the therapeutically active polypeptide. The expression cassette can be comprised in a gene therapy vector. Generally, the term expression cassette excludes polynucleotide sequences 5′ to the 5′ ITR and 3′ to the 3′ ITR.

All patents, patent publications, and other publications referenced and identified in the present specification are individually and expressly incorporated herein by reference in their entirety for all purposes.

EXAMPLES

EXAMPLE 1: Enhanced Gene Expression using Lamp-2B Transgene Variants

A gene expression cassette depicted in **FIG. 2** was constructed in a plasmid-based green fluorescence protein (GFP) reporter system for LAMP-2B transgene expression. The plasmid included a single open reading frame encoding the LAMP-2B transgene, a 2A peptide, and an enhanced green fluorescence protein (eGFP). Post-translational self-cleavage of the 2A peptide resulted in co-expression of LAMP-2B and eGFP in equimolar amounts. A wild-type LAMP-2B coding sequence (SEQ ID NO: 2) and three codon variants of the LAMP-2B coding sequence (codon variants 1, 2 and 3; SEQ ID NOs: 3-5, respectively) were tested as the transgenes. The three codon variants contained a reduced number of CpGs, removal of cryptic sites, and a reduced number of open reading frames as compared to the wild-type LAMP-2B coding sequence.

Forty wells of a CellBIND 96-well plate (NUNC #3300) were coated with 0.1% gelatin in water (Millipore ES-006-B) for 1 hour at room temperature. Approximately 88,000 induced pluripotent stem cell (iPSC)-derived cardiomyocytes (VWR MSPP-CMC10001) were plated into each well in plating media (VWR #M1001) at 37°C and 5% carbon dioxide (CO_2). After 4 hours, the media was changed to maintenance media (VWR #M1003) that was pre-equilibrated to 37°C and 5% CO_2. A transfection mixture was prepared by adding 6 µL of transfection reagent (ViaFect Promega #E4982) to 128 µL of 0.015 µg/µL plasmid (wildtype or codon variants 1, 2, or 3) in OptiMEM or OptiMEM + ViaFect only (negative control) and incubated for 10-20 min. 100 µL of this transfection mixture was added to 1 mL of maintenance media that was pre-equilibrated to 37°C and 5% CO_2.

Approximately 28 hours after initial plating, 100 µL of this transfection mixture in maintenance media was added to each well. Approximately 48 hours after adding media with transfection mixture, the cells were imaged and analyzed on an automated confocal microscope (Perkin Elmer Operetta CLS, Harmony version 4.5 software) for GFP positive

cells **(FIG. 3)** and their average fluorescent intensity **(FIG. 4)**. Immunofluorescence images of the cells two or seven days after transfection are shown in **FIG. 5** and **FIG. 6**, respectively.

FIG. 3 shows that the use of codon variants 1 ("CO 1") or 2 ("CO 2") resulted in expression of GFP in a significantly higher number of cells (~ 9-fold higher) as compared to the wild-type transgene. Similarly, **FIG. 4** showed mean GFP intensity in cells transfected with codon variants 1 ("CO 1") and 2 ("CO 2") to be 1.5-fold higher than wild-type.

EXAMPLE 2: Optimized LAMP-2B Gene Therapy Vectors

Optimized AAV gene therapy vectors are produced by inserting the LAMP-2B optimized variant, CO1 sequence described in Example 1 into the expression cassette of a recombinant AAV vector. The AAV regulatory cassette is modified by removal of upstream cryptic ATG sequence, use of an optimized consensus Kozak sequence, and/or a full-length polyadenylation sequence. The vectors are tested in comparison to control recombinant AAV vectors containing one or more additional ATG sites upstream of the transgene, a non-optimal Kozak sequence, and/or a non-full-length polyadenylation sequence. Vectors are tested *in vitro* in Danon patient iPSC-derived cardiomyocytes and in a LAMP-2$^{-/-}$ knockout mouse model of Danon disease. The optimized AAV gene therapy cassettes and vectors are expected to result in a higher level of expression and/or expression in a higher percentage of cells as compared to the control recombinant AAV vectors.

EXAMPLE 3: In Vitro Evaluation of AAV9-LAMP-2B.v1.2

AAV gene therapy cassette and vector were produced by inserting the LAMP-2B variant sequence CO1 (SEQ ID NO: 3) into a recombinant AAV plasmid vector having no cryptic start sites upstream of the transgene, an optimized consensus Kozak sequence, and a full-length polyadenylation (polyA) sequence from rabbit globin ("LAMP-2B.v.1.2"; expression cassette: SEQ ID NO: 8). LAMP-2B.v1.2 was compared to LAMP-2B v1.0, which is the regulatory cassette having a wild-type LAMP-2B transgene (transgene sequence: SEQ ID NO: 2) without an optimal Kozak sequence and a mini-polyA.

HEK293 cells were used to generate viral particles with three-plasmid, helper virus-free system was used to generate recombinant AAV particles containing serotype 9 capsid

proteins and viral genomes that have AAV2 ITRs flanking the LAMP-2B expression cassette. The expression cassette contains the human codon-optimized LAMP-2B coding sequence (v1.2 or v1.0) driven by an upstream chimeric "CAG" promoter containing the CMV IE enhancer (CMV IE), the chicken β-actin (CBA) promoter, and a CBA intron splice donor (**FIG. 1A**). The expression cassette also includes a downstream WPRE element and is terminated by the rabbit beta-globin polyadenylation signal (RGpA). The HEK293 cells were transiently transfected with the LAMP-2B.v1.2 plasmid vector or the LAMP-2B.v1.0 plasmid vector, a pAAV2/9 packaging plasmid, and pAd-Helper adenovirus helper plasmid.

CHO-Lec2 cells were seeded in a 24 well plate at 1.2×10^5 cells/mL in MEM-α containing 10% FBS and 1% Normocin. The following day, CHO-Lec2 cells were transduced in serum-free MEMα medium with either AAV9-LAMP-2B.v1.0, AAV9-LAMP-2B.v1.2, or the same vector having GFP in place of the LAMP-2B transgene (at MOI of 3×10^5). Seven days post-transduction, lysates were harvested using the Mammalian Cell Lysis kit (Sigma) and total protein was quantified using the MicroBCA kit per manufacturer's instructions. Proteins were separated by SDS-PAGE, transferred onto nitrocellulose membranes and immunoblotted for LAMP-2B (1:500) and GAPDH (1:10,000). CHO-Lec2 cells transduced with AAV9-optimized LAMP-2B.v1.2 showed increased expression compared to CHO-Lec2 cells transduced with the original AAV9-wild-type LAMP-2B.v1.0 (**FIG. 7A**). LAMP-2B was not detected in cells transduced with AAV9-GFP vector alone (**FIG. 7A**).

LAMP-2B expression was also quantitated in cell lysates by ELISA. Briefly, a 96 well plate was coated with anti-LAMP-2B antibody (clone: H4B4), lysates were added to the wells, and detection was performed using anti-LAMP-2B polyclonal antibody (1:500, Thermo Fisher PA; 5-24575) followed by incubation with HRP-conjugated anti-rabbit antibody (1:3000, Sigma). Transduction with AAV9-optimized LAMP-2B.v1.2 vector resulted in an approximately 7-fold increase in LAMP-2B expression compared to cells transduced with the AAV9-wild-type LAMP-2B.v1.0 (**FIG. 7B**).

Cardiomyocytes were derived from iPSCs generated from individuals with Danon disease. Following rhythmic contraction and selection for purity, Danon disease cardiomyocytes were transduced with various viral genome copies (vg) of AAV9-Luc (negative control), AAV9-wild-type LAMP-2B.v1.0 or AAV9-optimized LAMP-2B.v1.2. Ten days post-transduction, transduced cardiomyocytes were fixed with 4%

paraformaldehyde, permeabilized, blocked for 30 min in 5% IgG-free BSA and incubated for 1 hour with either mouse anti-human LAMP-2B antibody (1:25, clone: H4B4) or rabbit anti-α-actinin antibody (1:200, #A7811, Sigma). Cells were washed with 1X PBS to remove residual unbound primary antibody and then subjected to the appropriate anti-mouse AlexaFluor tagged secondary antibody and 200 ng/mL DAPI for 60 minutes at room temperature. The wells were then washed with PBS prior to imaging. Human LAMP-2B expression was expressed at a higher level in cardiomyocytes transduced with low titer (1.56×10^8 vg/well) AAV9-optimized LAMP-2B.v1.2 vector compared to cardiomyocytes transduced with the highest titer (8.45×10^{10} vg/well) of AAV9-wild-type LAMP-2B.v1.0 (**FIG. 8A** and **FIG. 8B**).

Western blot analyses were performed on the transduced Danon disease cardiomyocytes. AAV9-optimized LAMP-2B.v1.2 at 0.983×10^9 vg/well showed significant expression of LAMP-2B protein compared to no detection of LAMP-2B protein in cells transduced with either AAV9-wild-type LAM2B.v1.0 (1.347×10^9 vg/well) or AAV9-Luc (1.167×10^9 vg/well) vectors (**FIG. 8C**). Collectively, these results demonstrate that the optimized AAV9-LAMP-2B.v1.2 vector mediates human LAMP-2B expression in Danon disease cardiomyocytes at a significantly higher level than the original AAV9-wild-type LAMP-2B.v1.0 vector.

EXAMPLE 4: In Vivo Evaluation of AAV9-LAMP-2B.v1.2 in Mouse Model of Danon Disease

LAMP-2-deficient mice were intravenously injected with 1×10^{13} vg/kg of AAV9 viral vectors containing original human LAMP-2B (AAV9-LAMP-2B.v1.0), optimized human LAMP-2B (AAV9-LAMP-2B.v1.2, codon variant 1 – SEQ ID NO: 3) or vehicle alone. Six weeks post-treatment, mice were sacrificed and heart tissue was collected for analysis of LAMP-2B expression.

Methods

For quantitative analyses of vector copy number, total DNA was isolated from frozen tissues using the DNeasy Blood and Tissue kit according to manufacturer's guidelines. DNA concentration and integrity was assessed spectrophotometrically. qPCR was performed to calculate viral genome copies per μg of DNA using TaqPath ProAmp Master Mix (Applied

Biosystems) with forward (5′-ATCATGCTATTGCTTCCCGTA-3′; SEQ ID NO: 36) and reverse (5′-GGGCCACAACTCCTCATAAA-3′; SEQ ID NO: 37) primers and a probe (5′-CCTCCTTGTATAAATCCTGGTTGCTGTCT-3′; SEQ ID NO: 38) for the WPRE gene. RNase P was used as an endogenous control (Thermofisher, #4403328). A standard curve was generated using a linearized plasmid that contained the vector genome (WPRE) used for virus production. Quantification of DNA per sample was calculated using TaqMan copy number reference assay and was represented as vector copy number per diploid nucleus (VCN/Diploid Nucleus).

RNA was extracted and purified from heart using RNeasy Fibrous Tissue Mini kit according to the manufacturer's protocol. RNA concentration and integrity were assessed spectrophotometrically. RNA was reverse-transcribed using iScript cDNA Synthesis kit and cDNA was used as a template for quantitative real-time (qRT)-PCR. qRT-PCR was performed on cDNA using TaqPath ProAmp Master Mix with forward (5′-ATCATGCTATTGCTTCCCGTA-3′; SEQ ID NO: 36) and reverse (5′-GGGCCACAACTCCTCATAAA-3′; SEQ ID NO: 37) primers and a probe (5′-CCTCCTTGTATAAATCCTGGTTGCTGTCT-3′; SEQ ID NO: 38) for the WPRE gene.

For protein extraction, tissues were flash-frozen and pulverized, and the subsequent tissue powder was digested in protein lysis buffer (100 mM Tris, 300 mM NaCl, 20 mM EDTA, 2% NP-40, 0.2% SDS) containing protease and phosphatase inhibitor cocktails. Partial protein lysates were passed through a glass tissue grinder and sonicated with 3 bursts of 5 second on ice, with 10 seconds intervals in between at 30 amplitude microns power. Samples were centrifuged for 15 min at 12000 rpm and then the supernatant was collected. Concentration of protein in samples was determined by Lowry assay. Proteins (20 µg/sample) were separated using 10-20% SDS-PAGE and transferred to PVDF membranes by rapid dry transfer technique. Membranes were then blocked in 5% milk (non-fat dry milk solubilized in PBS containing 0.1% Tween-20) for 1h, and incubated with anti-human LAMP-2B (1:100, H4B4), anti-mouse LAMP-2B (1:100) or anti-GAPDH (1:1000, #32233, Santa Cruz) antibodies overnight at 4°C. Membranes were washed and then incubated with the appropriate HRP-conjugated secondary antibodies (1:10,000) for 1 hour at room temperature. The blots were developed using WesternBright™ Sirius substrates followed by imaging on a BioRad gel imager.

For immunofluorescence analyses, tissues were cryoprotected in 30% sucrose/PBS at 4°C, embedded in optimal cutting temperature (OCT) mounting media and then tissue was cut to 8-10 μm thickness on a standard cryotome. Cryosections were then fixed with 4% PFA for 5 min, permeabilized with 0.2% Triton-X for 5 min and blocked with 1% BSA, 3% serum, 1% cold water fish gelatin in PBS for 30 minutes. The sections were incubated with mouse anti-human LAMP-2B antibody (1:50, H4B4) directly conjugated to Alexa Fluor 647 and rabbit anti-dystrophin antibody overnight at 4° C. The slides were then incubated with anti-rabbit Alexa Fluor 488 secondary antibody and DAPI (1:2000, #D9542, Sigma) for 30 min at room temperature. Slides were then imaged using an Olympus FluoView FV1000 confocal microscope. Scan speed, off set, voltage, and gain were kept constant during the acquisition of all images on a given day.

Results

Quantitative PCR was performed on cardiac tissue of AAV9-treated LAMP-2-deficient mice. Although similar viral copy numbers were observed in cardiac tissue of mice treated with wild-type and optimized LAMP-2B containing vector (**FIG. 9A**), transcription of AAV9-optimized LAMP-2B.v1.2 was increased nearly 7-fold compared to the AAV9-wild-type LAMP-2B.v1.0 (**FIG. 9B**). Despite similar transduction of v1.0 and v1.2 viral vectors in cardiac tissue, induction of human LAMP-2B mRNA expression was significantly enhanced using v1.2.

LAMP-2-deficient mice intravenously injected with AAV9-optimized LAMP-2B.v1.2 vector also showed significantly higher levels of human LAMP-2B protein in cardiac tissue compared to LAMP2-deficient mice treated with AAV9-wild-type LAMP-2B.v1.0 or the vehicle control (**FIG. 9C**). Similar results were achieved with immunofluorescence staining: human LAMP-2B was highly induced in cardiac tissue of LAMP-2-deficient mice treated with AAV9-optimized LAMP-2B.v1.2 (**FIG.9D**). Collectively, these data show that viral transduction using the AAV9-optimized LAMP-2B.v1.2 vector leads to increased expression of human LAMP-2B protein in cardiac tissue *in vivo* compared to AAV9-wild-type LAMP-2B.v1.0 at the same dose.

EXAMPLE 4: In Vivo Evaluation of AAV9-LAMP-2B.v1.2 in Non-Human Primates

Non-human primates were intravenously injected with 1×10^{13} vg/kg of either the AAV9 viral vector containing codon variant LAMP-2B (v1.2, codon variant 1 – SEQ ID NO: 3) described in Example 2, or vehicle control. Eight weeks post-treatment, the non-human primates were humanely sacrificed, and heart, muscle, liver and brain tissue was collected for analysis of LAMP-2B expression.

Methods

For quantitative analyses of vector copy number, total DNA was isolated from frozen tissues using the Qiagen DNeasy kit according to manufacturer's guidelines. DNA concentration and integrity were assessed spectrophotometrically. Quantitative PCR on isolated DNA was performed using TaqMan Universal Master Mix II (Applied Biosystems) with forward (5′-ATCATGCTATTGCTTCCCGTA-3′; SEQ ID NO: 36) and reverse (5′-GGGCCACAACTCCTCATAAA-3′; SEQ ID NO: 37) primers and a probe (5′-CCTCCTTGTATAAATCCTGGTTGCTGTCT-3′; SEQ ID NO: 38) for the WPRE gene. RNaseP was used as an endogenous control (#4403328, ThermoFisher). A standard curve was generated using a linearized plasmid that contained the vector genome used for virus production. Quantification of DNA per sample was calculated using the TaqMan copy number reference assay and was represented as vector copy number per diploid nucleus (VCN/Diploid Nucleus).

RNA was extracted and purified from heart and skeletal muscle using the RNeasy Fibrous Tissue Mini kit (Qiagen) and from liver and brain using the RNeasy Lipid Tissue kit (Qiagen) according to manufacturer's protocol. RNA concentration and integrity was assessed using the NanoDrop spectrophotometer. RNA was reverse-transcribed using SuperScript IV VILO master mix (ThermoFisher) and cDNA was used as a template for quantitative real-time (qRT)-PCR. qRT-PCR was performed on cDNA using TaqMan Universal Master Mix II with forward (5′-ATCATGCTATTGCTTCCCGTA-3′; SEQ ID NO: 36) and reverse (5′-GGGCCACAACTCCTCATAAA-3′; SEQ ID NO: 37) and a probe (5′-CCTCCTTGTATAAATCCTGGTTGCTGTCT-3′; SEQ ID NO: 38) of the WPRE gene. Human HPRT-1 was used as an endogenous control. A standard curve was generated using a linearized plasmid that contained the vector genome used for virus production.

For semi-quantitative analysis of mRNA using RNAScope technology, cardiac tissue was fixed in 10% neutral buffered formalin, embedded in paraffin and sectioned. Transgene mRNA was detected using WPRE-O3 probe (#518628, ACD) with RNAscope 2.5 LS RED. Cells with greater than 1 dot were considered positive and the percentage of positive cells were binned into five categories: 0%, 1-25%, 26-50%, 51-75% or 100%.

For western blot analyses, 125 mg of cardiac tissue was homogenized in 500 µL of lysis buffer using the Next Advance Bullet System. Protein concentration was determined using the BCA kit (ThermoFisher) and proteins (50 µg/sample) were separated using SDS-PAGE and then transferred to nitrocellulose membranes. Membranes were then probed with mouse anti-human LAMP2 (1:100), washed and then incubated with HRP-conjugated anti-mouse antibody. The blots were developed using ECL substrate and the BioRad ChemiDoc MP system.

For the LAMP-2B ELISA, protein extraction was performed as described above. A plate was coated with mouse anti-LAMP2 antibody (clone: H4B4, #NBP2-22217, Novus Biologicals), 100 µg of tissue lysate was added to each well, and detection was performed using anti-LAMP2 polyclonal antibody (#AF6228, R&D Systems) followed by incubation with HRP-conjugated donkey anti-goat antibody (#AP180P, Millipore).

Results

Quantitative PCR was performed on various tissues of AAV9-treated primates. Viral copy numbers were increased in heart, muscle and liver tissue of primates injected with AAV9-LAMP-2B.v1.2 vector at 1×10^{13} vg/kg compared to vehicle control (**FIG. 10A**). Vector genomes were detected in all cardiac chambers examined, including the left and right ventricles and left and right atriums **(FIG. 10B)**. Vector mRNA was detected at significant levels in the heart, skeletal muscle and liver tissue of treated primates compared to the untreated vehicle control (**FIG. 10C** and **FIG. 10D**). *In situ* RNA analysis showed approximately 25-75% of heart and liver tissue expressing vector mRNA (**FIG. 10E** and **FIG. 10F**). These data demonstrate that systemic administration of 1×10^{13} vg/kg AAV9-optimized LAMP-2B.v1.2 to a primate results in efficient transduction and expression in heart tissue *in vivo*.

Western blot analyses showed that primates systemically treated with LAMP-2B.v1.2 at 1×10^{13} vg/kg showed increased human LAMP-2B protein in the left and right ventricles

and left atrium of the heart compared to an untreated control (**FIG. 10G** and **FIG. 10H**). ELISA also showed that human LAMP-2B protein was increased in the left ventricle and atrium of the heart, as well as skeletal muscle tissue of primates treated with AAV9-LAMP-2B.v1.2 vector compared to an untreated control (**FIG. 10I** and **FIG. 10J**). Vector transduction using AAV9.LAMP-2B.v1.2 leads to expression of human LAMP-2B protein in cardiac tissue of primates *in vivo*.

CLAIMS

What is claimed is:

1. A gene therapy vector comprising an expression cassette comprising a transgene encoding an isoform of lysosome-associated membrane protein 2 (LAMP-2) or a functional variant thereof, wherein the transgene is optimized for expression in a human host cell.

2. The gene therapy vector of claim 1, wherein the transgene is codon-optimized for expression in a human host cell.

3. The gene therapy vector of claim 1 or claim 2, wherein the expression cassette contains fewer CpG sites than SEQ ID: 2.

4. The gene therapy vector of any one of claims 1-3, wherein the expression cassette contains fewer cryptic splice sites than SEQ ID: 2.

5. The gene therapy vector of any one of claims 1-4, wherein the expression cassette encodes fewer alternative open reading frames than SEQ ID: 2.

6. The gene therapy vector of any one of claims 1-4, wherein the transgene shares at least 95% identity or at least 99% identity to a sequence selected from SEQ ID NOs: 3-5.

7. The gene therapy vector of claim 6, wherein the transgene comprises a sequence selected from SEQ ID NOs: 3-5.

8. The gene therapy vector of claim 7, wherein the transgene shares at least 95% identity to SEQ ID NO: 3 or is identical to SEQ ID NO: 3.

9. The gene therapy vector of any one of claims 1-8, wherein the expression cassette comprises a consensus optimal Kozak sequence, wherein optionally the consensus optimal Kozak sequence comprises SEQ ID NO: 6.

10. The gene therapy vector of any one of claims 1-9, wherein the expression cassette comprises a full-length polyA sequence, wherein optionally the full-length polyA sequence comprises SEQ ID NO: 7.

11. The gene therapy vector of any one of claims 1-10, wherein the expression cassette comprises no start site 5′ to the transgene capable of generating alternative mRNAs.

12. The gene therapy vector of any one of claims 1-11, wherein the expression cassette comprises, in the 5′ to 3′ direction, a first inverted terminal repeat, an enhancer/promoter region, introns, a consensus optimal Kozak sequence, the transgene, a 3′ untranslated region including a full-length polyA sequence, and a second inverted terminal repeat.

13. The gene therapy vector of claim 12, wherein the enhancer/promoter region comprises in the 5′ to 3′ direction a CMV IE enhancer and a chicken beta-actin promoter, and optionally wherein the enhancer/promoter region further comprises a first exon and first intron of a chicken beta-actin gene and a splice acceptor of a rabbit beta-globin gene.

14. The gene therapy vector of any one of claims 1-13, wherein the expression cassette shares at least 95% identity to a sequence selected from SEQ ID NOs: 8-10.

15. The gene therapy vector of claim 14, wherein the expression cassette shares complete identity to a sequence selected from SEQ ID NOs: 8-10.

16. The gene therapy vector of claim 14, wherein the expression cassette shares at least 95% identity to SEQ ID NO: 3 or is identical to SEQ ID NO: 8.

17. The gene therapy vector of any one of claims 1-16, wherein the vector is an adeno-associated virus (AAV) vector.

18. A pharmaceutical composition comprising the gene therapy vector of any one of claims 1 to 17.

19. A method of treating or preventing Danon disease or another autophagy disorder in a subject in need thereof, comprising administering to the subject the gene therapy vector of any one of claims 1 to 17 or the pharmaceutical composition of claim 18.

20. The method of claim 19, wherein the vector or pharmaceutical composition is administered via a route selected from the group consisting of intravenous, intra-arterial, intracardiac, intracoronary, intramyocardial, intrarenal, intraurethral, epidural, and intramuscular.

21. The method of claim 19 or claim 20, wherein the autophagy disorder is selected from the group consisting of end-stage heart failure, myocardial infarction, drug toxicities, diabetes, end-stage renal failure, and aging.

22. The method of any one of claims 19-21, wherein the subject is a human.

23. The method of any one of claims 19-22, wherein the subject is exhibiting symptoms of Danon disease or another autophagy disorder.

24. The method of any one of claims 19-23, wherein the subject has been identified as having reduced or non-detectable LAMP-2 expression.

25. The method of any one of claims 19-24, wherein the subject has been identified as having a mutated LAMP-2 gene.

26. The method of any one of claims 19-25, wherein administration of the gene therapy vector results in increased expression of LAMP-2B polynucleotide and/or LAMP-2B protein in the heart compared to an untreated subject.

27. The method of any one of claims 19-26, wherein administration of the gene therapy vector results in at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 1.6-fold, at least about 1.7-fold, at least about 1.8-fold, at least about 1.9-fold, at least about 2.0-fold, at least about 2.2-fold, at least about 2.3-fold, at least about 2.4-fold, at least about 2.5-fold, at least about 3-fold, or at least about 4-fold increased expression of LAMP-2B polynucleotide and/or LAMP-2B protein in the heart compared to an untreated subject.

28. The method of any one of claims 19-27, wherein administration of the gene therapy vector results in expression of LAMP-2B polynucleotide and/or LAMP-2B protein in the liver of at most about 1.1-fold, at most about 1.2-fold, at most about 1.3-fold, at most about 1.4-fold, or at most about 1.5-fold increased compared to expression in the liver of an untreated subject.

29. A method of expressing LAMP-2B in a subject, comprising systemically administering an adeno-associated viral (AAV) vector to the subject, wherein the AAV vector comprises an expression cassette comprising a transgene sharing at least 95% identity with SEQ ID NO: 3 or is identical to SEQ ID NO: 3, the transgene operatively linked to an enhancer/promoter region, wherein systemic administration of the AAV vector to the subject results in increased expression of LAMP-2B compared to expression of LAMP-2B prior to administration of the AAV vector or expression of LAMP-2B in an untreated control subject.

30. The method of claim 29, wherein the systemic administration comprises intravenous administration.

31. The method of claim 29 or claim 30, wherein the AAV vector is administered at a dose of between about 1×10^{12} and 5×10^{14} vector genomes (vg) of the AAV vector per kilogram (vg) of total body mass of the subject (vg/kg).

32. The method of claim 31, wherein the AAV vector is administered at a dose of between about 1×10^{13} and 5×10^{14} vg/kg.

33. The method of claim 31, wherein the AAV vector is administered at a dose of between about 5×10^{13} and 3×10^{14} vg/kg.

34. The method of claim 31, wherein the AAV vector is administered at a dose of between about 5×10^{13} and 1×10^{14} vg/kg.

35 The method of any one of claim 29-34, wherein the enhancer/promoter region comprises in the 5′ to 3′ direction a CMV IE Enhancer and a Chicken Beta-Actin Promoter, and optionally wherein the enhancer/promoter region further comprises a first exon and first intron of a chicken beta-actin gene and a splice acceptor of a rabbit beta-globin gene.

36. The method of any one of claims 29-35, wherein the expression cassette comprises a consensus optimal Kozak sequence operatively linked to the transgene, wherein optionally the consensus optimal Kozak sequence comprises SEQ ID NO: 6.

37. The method of any one of claims 29-36, wherein the expression cassette comprises a full-length polyA sequence operatively linked to the transgene, wherein optionally the full-length polyA sequence comprises SEQ ID NO: 7.

38. The method of any one of claims 29-37, wherein the expression cassette comprises no start site 5′ to the transgene capable of generating alternative mRNAs.

39. The method of any one of claims 29-38, wherein the expression cassette comprises operatively linked, in the 5′ to 3′ direction, a first inverted terminal repeat, an enhancer/promoter region, introns, a consensus optimal Kozak sequence, the transgene, a 3′ untranslated region

including a full-length polyA sequence, and a second inverted terminal repeat, wherein the expression cassette comprises no start codon 5′ to the start codon of the transgene.

40. The method of any one of claims 29-39, wherein the subject is a primate.

41. The method of any one of claims 29-40, wherein the subject has been identified as having, or is suspected of having, reduced or non-detectable LAMP-2 expression.

42. The method of any one of claims 29-41, wherein the subject has a mutated LAMP-2 gene.

43. The method of any one of claims 29-42, wherein the subject is exhibiting symptoms of Danon disease.

44. The method of any one of claims 29-43, wherein the subject suffers from, or is at risk for, Danon disease.

45. The method of any one of claims 29-44, wherein administration of the gene therapy vector results in increased expression of LAMP-2B polynucleotide and/or LAMP-2B protein in the heart compared to an untreated subject.

46. The method of any one of claims 29-45, wherein administration of the gene therapy vector results in at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 1.6-fold, at least about 1.7-fold, at least about 1.8-fold, at least about 1.9-fold, at least about 2.0-fold, at least about 2.2-fold, at least about 2.3-fold, at least about 2.4-fold, at least about 2.5-fold, at least about 3-fold, or at least about 4-fold increased expression of LAMP-2B polynucleotide and/or LAMP-2B protein in the heart compared to an untreated subject.

47. The method of any one of claims 29-47, wherein administration of the gene therapy vector results in expression of LAMP-2B polynucleotide and/or LAMP-2B protein in the liver of at most about 1.1-fold, at most about 1.2-fold, at most about 1.3-fold, at most about 1.4-fold, or at most about 1.5-fold increased compared to expression in the liver of an untreated subject.

48. A polynucleotide comprising a polynucleotide sequence that shares 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 at least 100% identity to any one of SEQ ID NOs: 3-5 or a functional variant thereof.

49. The polynucleotide of claim 48, wherein the polynucleotide sequence that shares 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 at least 100% identity to SEQ ID NO: 3 or a functional variant thereof.

50 The polynucleotide of claim 49, wherein the polynucleotide sequence that shares 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 at least 100% identity to SEQ ID NO: 3.

51. The polynucleotide of claim 50, wherein the polynucleotide sequence consists of SEQ ID NO: 3.

52. The polynucleotide of any one of claims 48-51, wherein the polynucleotide is a transgene encoding LAMP-2B or a functional variant thereof.

53. An expression cassette comprising the polynucleotide of any one of claims 48-52.

54. The expression cassette of claim 53, wherein the expression cassette comprises an enhancer/promoter region that shares at least 80% identity, 90% identity, 95% identity, or 100% identity to a CAG promoter (SEQ. ID NO: 22).

55. The expression cassette of claim 53, wherein the expression cassette comprises an enhancer/promoter region that shares at least 80% identity, 90% identity, 95% identity, or 100% identity to a sequence selected from a CMV promoter (SEQ. ID NO: 23), an SV40 promoter (SEQ ID NO: 24), a PGK promoter (SEQ ID NO: 25), and/or a human beta-actin promoter (SEQ ID NO: 26).

56. The expression cassette of any one of claims 53-55, wherein the expression cassette comprises a 3′ UTR that shares at least 80% identity, 90% identity, 95% identity, or 100% identity to SEQ ID NO: 27.

57. The expression cassette of any one of claims 53-56, wherein the expression cassette comprises a poly-adenylation sequence that shares at least 80% identity, 90% identity, 95% identity, or 100% identity to a CAG promoter (SEQ ID NO: 7).

58. An expression cassette comprising, in 5′ to 3′ order:

(a) an enhancer/promoter region that shares at least 80% identity, 90% identity, 95% identity, or 100% identity to a CAG promoter (SEQ. ID NO: 22);

(b) a polynucleotide sequence encoding a LAMP-2B protein that shares 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 at least 100% identity to SEQ ID NO: 3;

(c) a 3′ UTR that shares at least 80% identity, 90% identity, 95% identity, or 100% identity to SEQ ID NO: 27; and

(d) a poly-adenylation sequence that shares at least 80% identity, 90% identity, 95% identity, or 100% identity to a CAG promoter (SEQ ID NO: 7).

59. The expression cassette of claim 58, wherein the functional LAMP-2B shares at least 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 at least 100% identity to SEQ ID NO: 1.

60. The expression cassette of claim 58 or claim 59, wherein the polynucleotide sequence encoding a LAMP-2B protein that shares at most 90%, at most 91%, at most 92%, at most 93%, at most 94%, at most 95%, at most 96%, at most 97%, at most 98%, at most 99%, or at most 100% identity to SEQ ID NO: 2.

61. An AAV vector comprising the expression cassette of any one of claim 53-69 flanked by (i) a 5′ ITR that shares at least 80% identity, 90% identity, 95% identity, or 100% identity to SEQ ID NO: 11 and (ii) a 5′ ITR that shares at least 80% identity, 90% identity, 95% identity, or 100% identity to SEQ ID NO: 12.

62. The AAV vector of claim 61, wherein the AAV vector comprises an AAV9 capsid.

63. The AAV vector of claim 62, wherein the AAV9 capsid shares at least 95% identity to amino acids 1 to 736 of SEQ ID NO: 27, to amino acids 138 to 736 of SEQ ID NO: 27, and/or to amino acids 203 to 736 of SEQ ID NO: 27.

64. A polynucleotide, comprising a polynucleotide sequence that shares 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 at least 100% identity to any one of SEQ ID NOs: 8-10.

65. The polynucleotide of claim 64, wherein the polynucleotide sequence that shares 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 at least 100% identity to SEQ ID NO:8.

66. The polynucleotide of claim 65, wherein the polynucleotide sequence consists of SEQ ID NO: 8.

67. The polynucleotide of any one of claims 65-66, wherein the polynucleotide comprises an expression cassette encoding LAMP-2B or a functional variant thereof.

68. The polynucleotide of any one of claim 65-67, wherein the polynucleotide comprises either or both of (i) a 5′ ITR that shares at least 80% identity, 90% identity, 95% identity, or 100% identity to SEQ ID NO: 11 and (ii) a 5′ ITR that shares at least 80% identity, 90% identity, 95% identity, or 100% identity to SEQ ID NO: 12.

69. The AAV vector comprising the polynucleotide of any one of claims 65-68.

70. The AAV vector of claim 69, wherein the AAV vector comprises an AAV9 capsid.

71. The AAV vector of claim 71, wherein the AAV9 capsid shares at least 95% identity to amino acids 1 to 736 of SEQ ID NO: 27, to amino acids 138 to 736 of SEQ ID NO: 27, and/or to amino acids 203 to 736 of SEQ ID NO: 27.

72. A method of treating Danon disease in a subject in need thereof, comprising intravenously administering the AAV vector of any one of claims 61-63 or claims 69-71 to the subject.

73. The method of claim 72, wherein administration of the gene therapy vector results in at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least

about 1.6-fold, at least about 1.7-fold, at least about 1.8-fold, at least about 1.9-fold, at least about 2.0-fold, at least about 2.2-fold, at least about 2.3-fold, at least about 2.4-fold, at least about 2.5-fold, at least about 3-fold, or at least about 4-fold increased expression of LAMP-2B polynucleotide and/or LAMP-2B protein in the heart compared to an untreated subject.

74. The method of claim 72 or claim 73, wherein administration of the gene therapy vector results in expression of LAMP-2B polynucleotide and/or LAMP-2B protein in the liver of at most about 1.1-fold, at most about 1.2-fold, at most about 1.3-fold, at most about 1.4-fold, or at most about 1.5-fold increased compared to expression in the liver of an untreated subject.

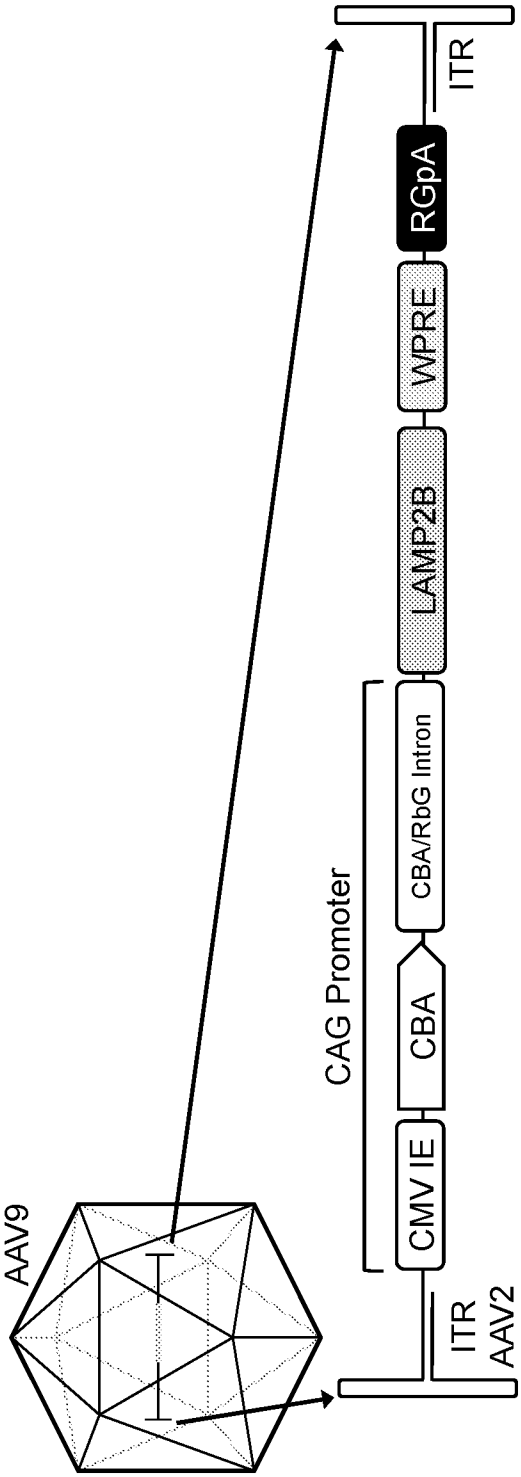
AAV9
ITR
AAV2
CMV IE
CBA
CBA/RbG Intron
CAG Promoter
LAMP2B
WPRE
RGpA
ITR

FIG. 1A

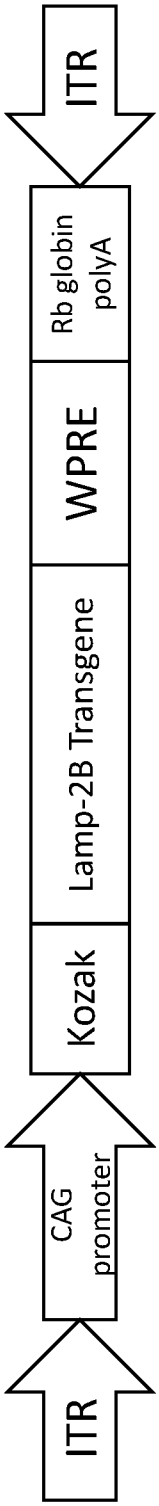
ITR
CAG promoter
Kozak
Lamp-2B Transgene
WPRE
Rb globin polyA
ITR

FIG. 1B

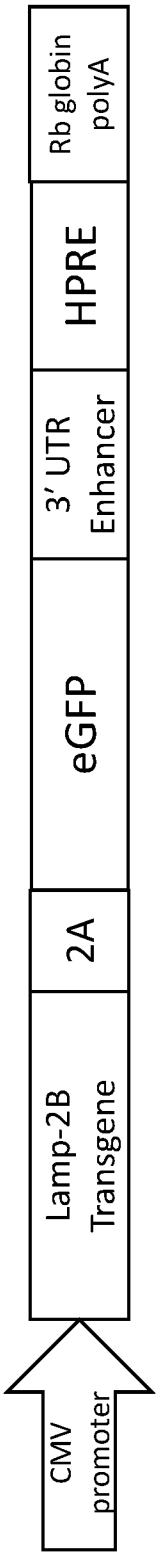
CMV promoter
Lamp-2B Transgene
2A
eGFP
3' UTR Enhancer
HPRE
Rb globin polyA

FIG. 2

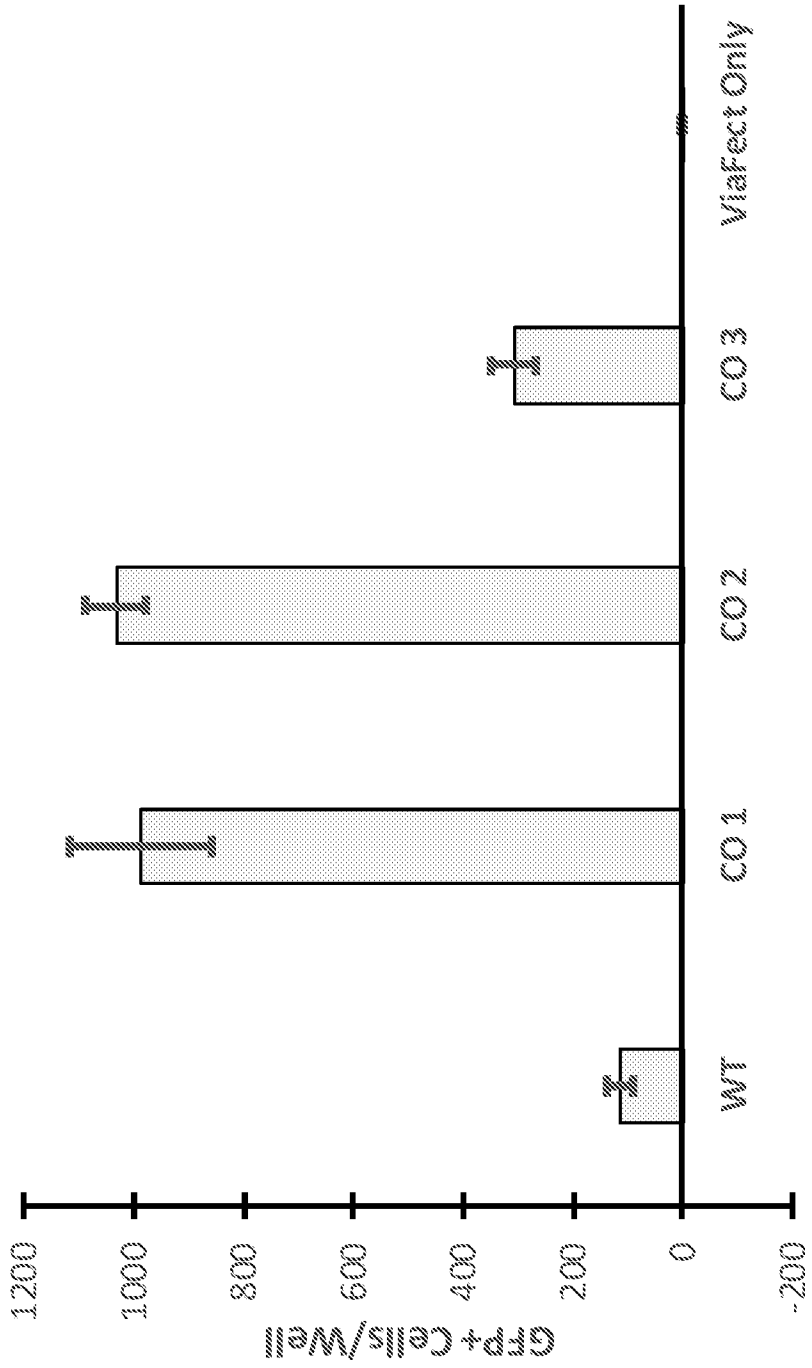
GFP+ Cells/Well
1200
1000
800
600
400
200
0
-200
WT
CO 1
CO 2
CO 3
ViaFect Only

FIG. 3

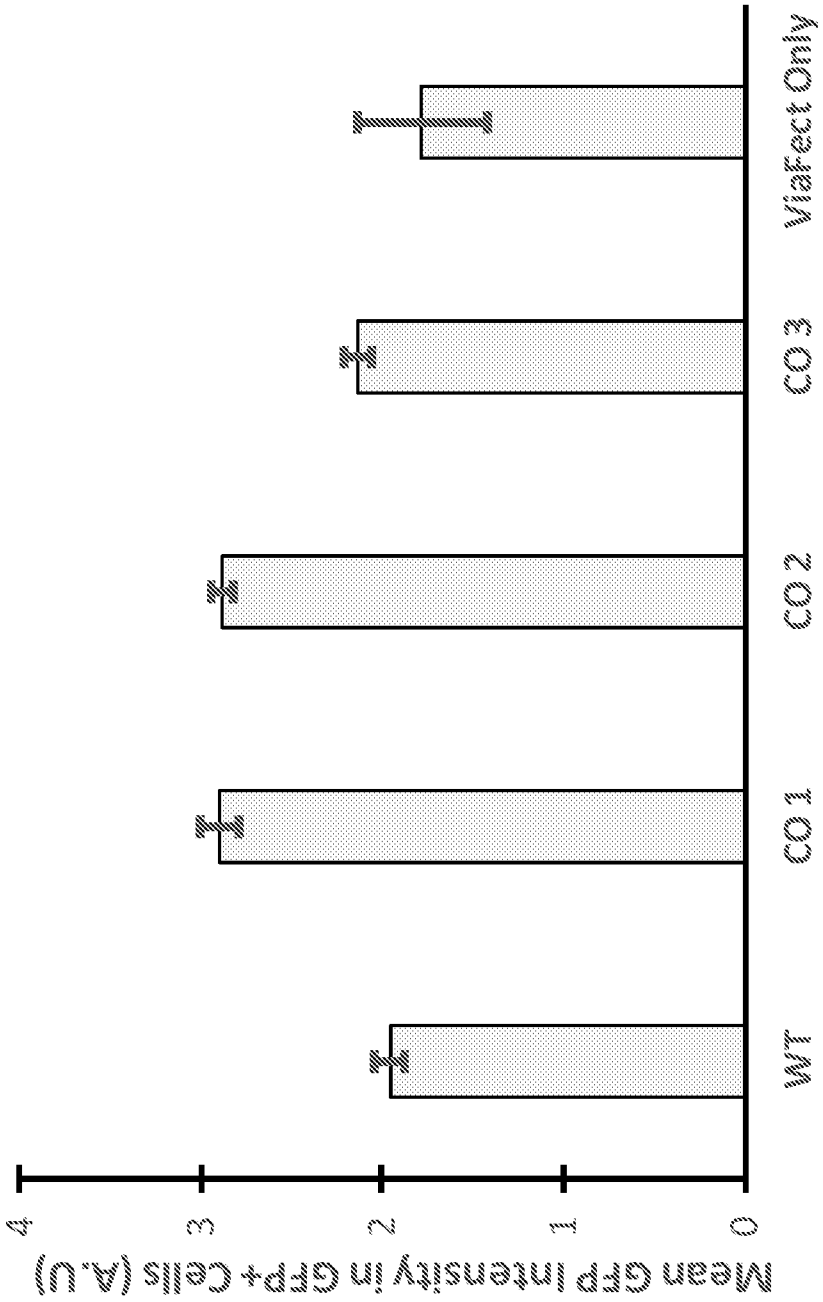
Mean GFP Intensity in GFP+ Cells (A.U)
0
1
2
3
4
WT
CO 1
CO 2
CO 3
ViaFect Only

FIG. 4

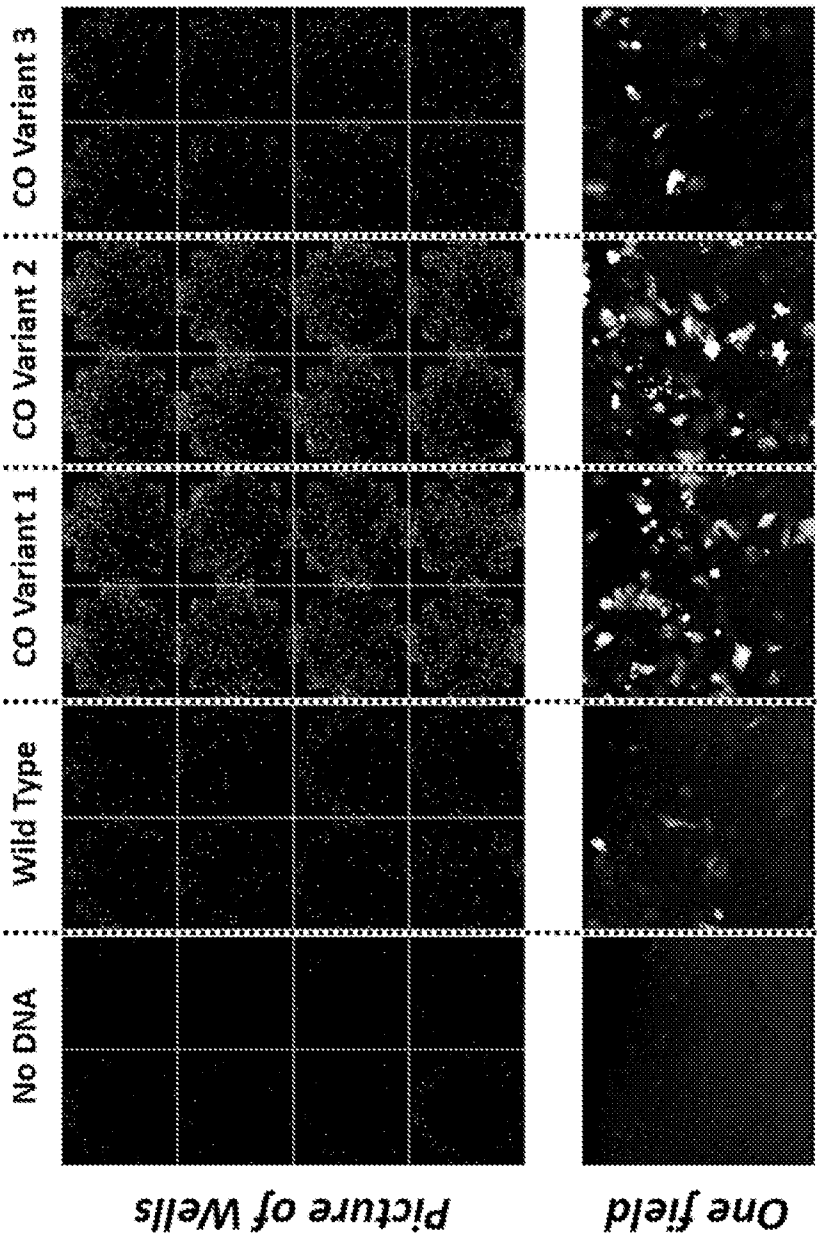
No DNA
Wild Type
CO Variant 1
CO Variant 2
CO Variant 3
Picture of Wells
One field

FIG. 5

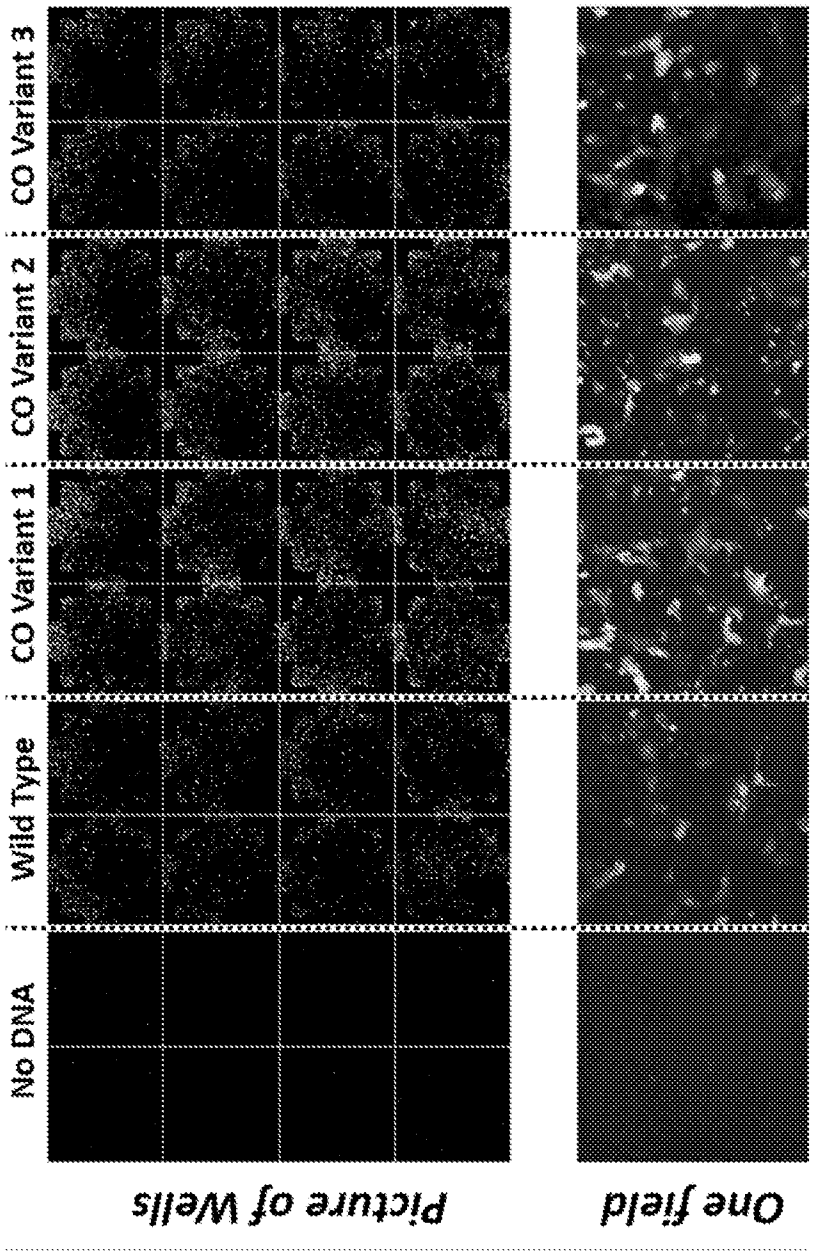
No DNA
Wild Type
CO Variant 1
CO Variant 2
CO Variant 3
Picture of Wells
One field

FIG. 6

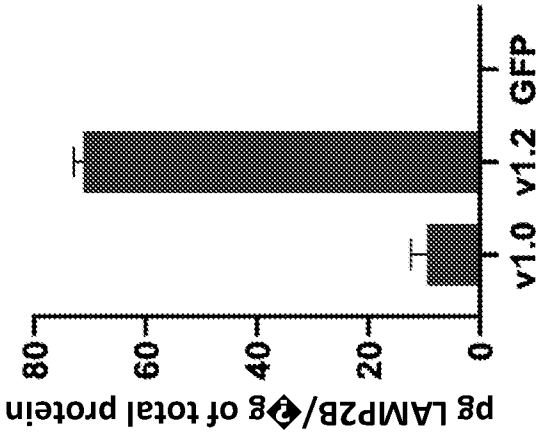
pg LAMP2B/pg of total protein
80
60
40
20
0
v1.0
v1.2
GFP

FIG. 7B

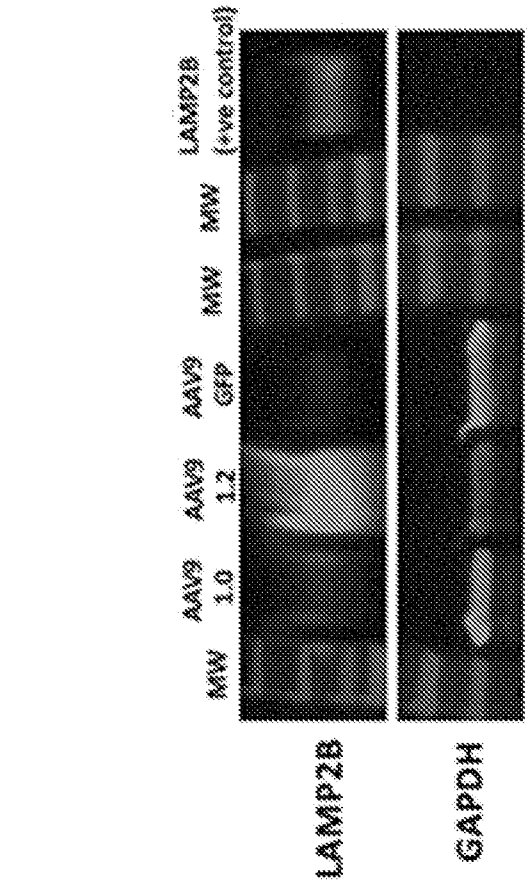
LAMP2B
GAPDH
MW
AAV9 1.0
AAV9 1.2
AAV9 GFP
MW
MW
LAMP2B (+ve control)

FIG. 7A

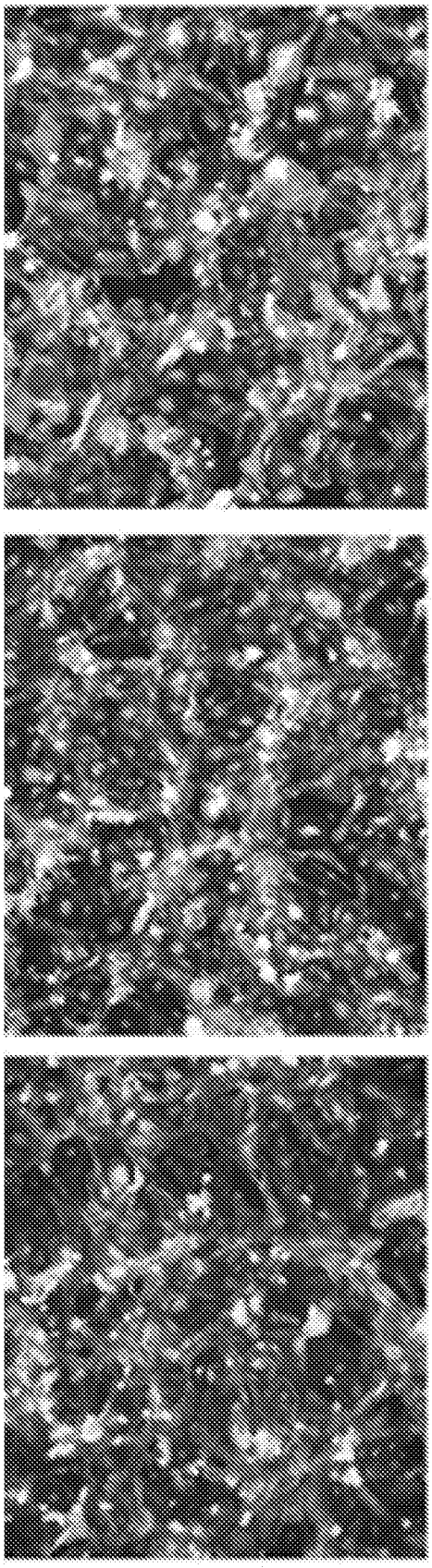
Luc 8.45 E+10
(highest titer tested)
LAMP28 v1.0 8.45 E+10
(highest titer tested)
LAMP28 v1.2 1.56 E+08
(lowest titer tested)

FIG. 8A

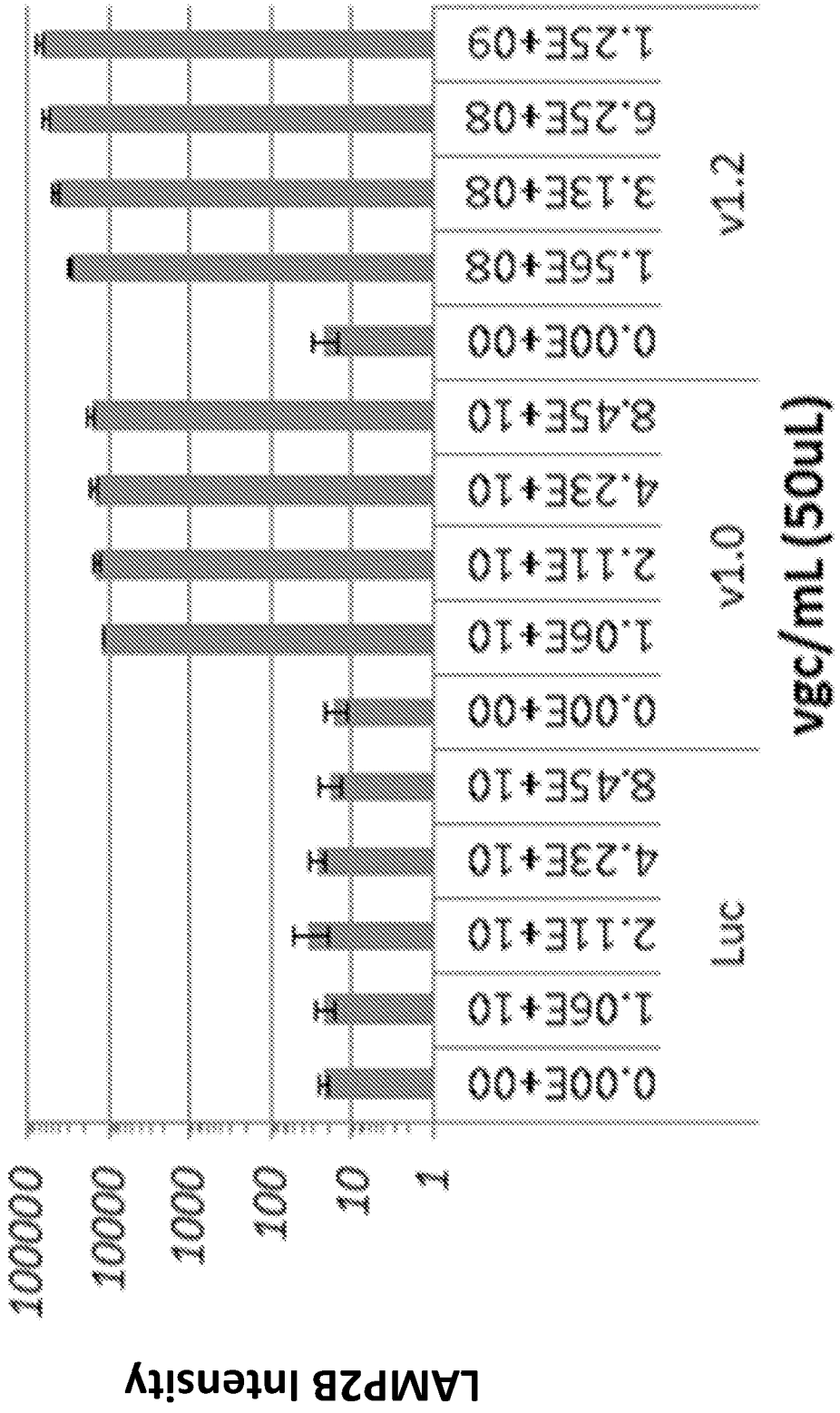
LAMP2B Intensity
100000
10000
1000
100
10
1
vgc/mL (50uL)
Luc
v1.0
v1.2
0.00E+00
1.06E+10
2.11E+10
4.23E+10
8.45E+10
0.00E+00
1.06E+10
2.11E+10
4.23E+10
8.45E+10
0.00E+00
1.56E+08
3.13E+08
6.25E+08
1.25E+09

FIG. 8B

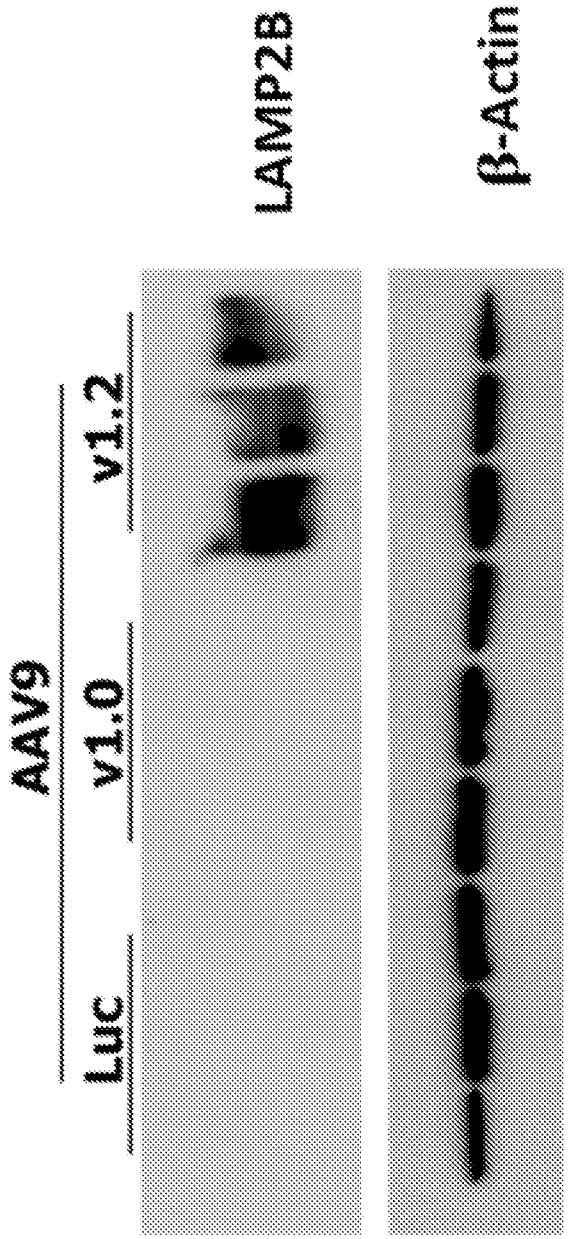

FIG. 8C

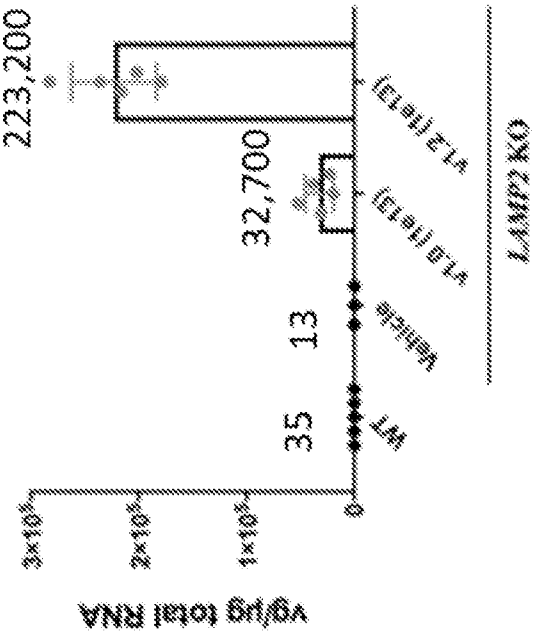
223,200
32,700
13
35
3×10⁵
2×10⁵
1×10⁵
0
vg/μg total RNA
WT
Vehicle
VT3 (1e+13)
VT3 (1e+11)
LAMP2 KO

FIG. 9B

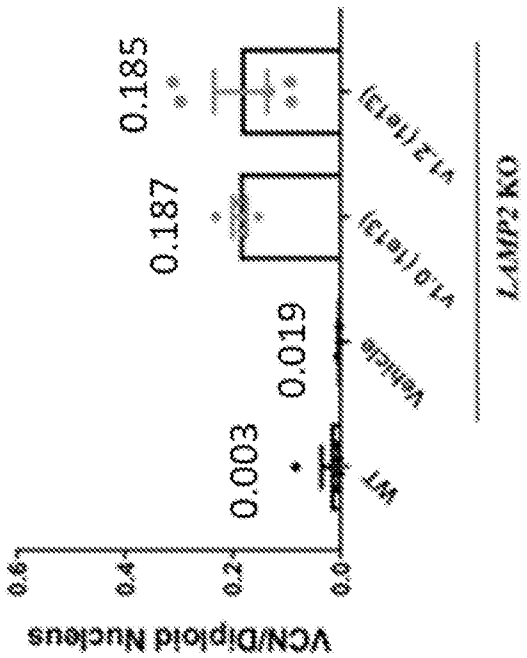
0.185
0.187
0.019
0.003
0.6
0.4
0.2
0.0
VCN/Diploid Nucleus
WT
Vehicle
VT3 (1e+13)
VT3 (1e+11)
LAMP2 KO

FIG. 9A

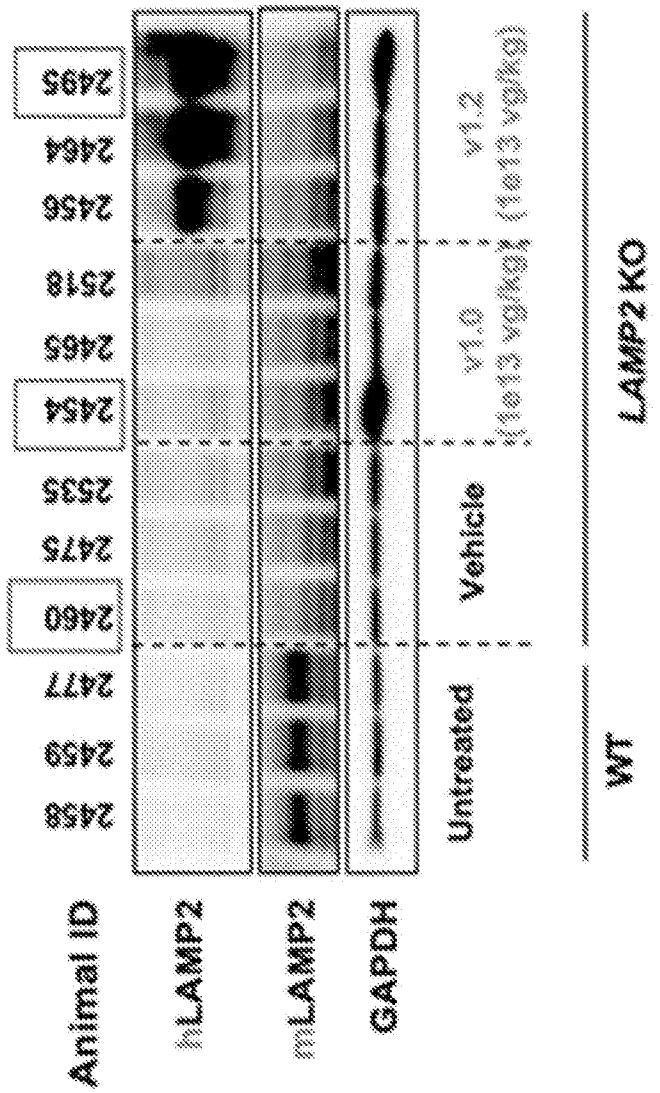
Animal ID
hLAMP2
mLAMP2
GAPDH
2458
2459
2477
2460
2475
2535
2454
2465
2518
2456
2464
2495
Untreated
Vehicle
v1.0 (1e13 vg/kg)
v1.2 (1e13 vg/kg)
WT
LAMP2 KO

FIG. 9C

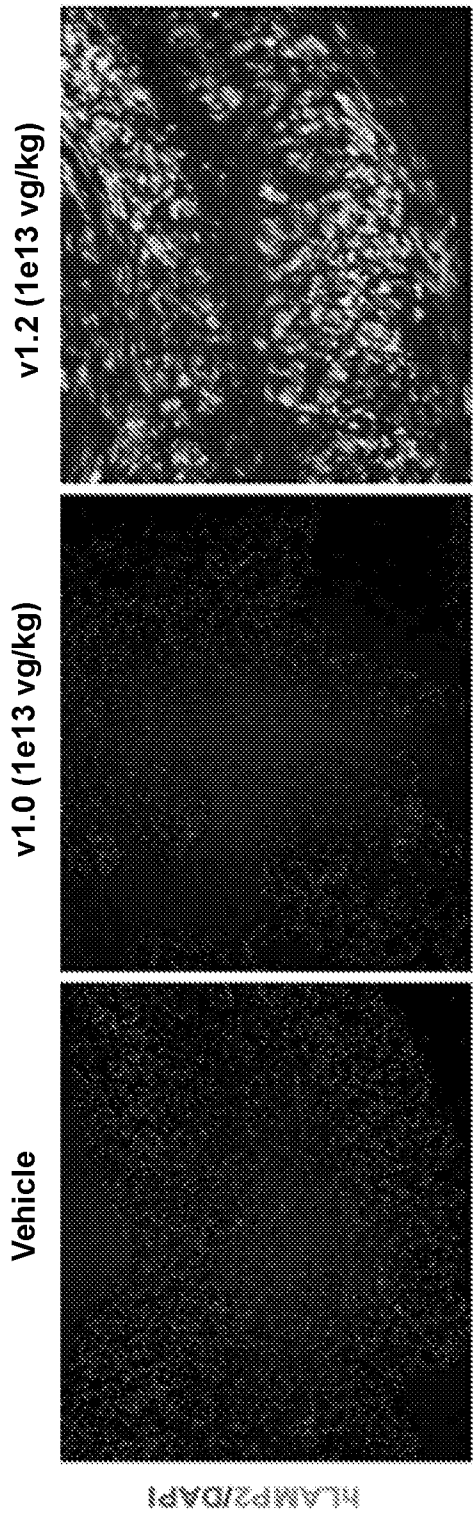

FIG. 9D

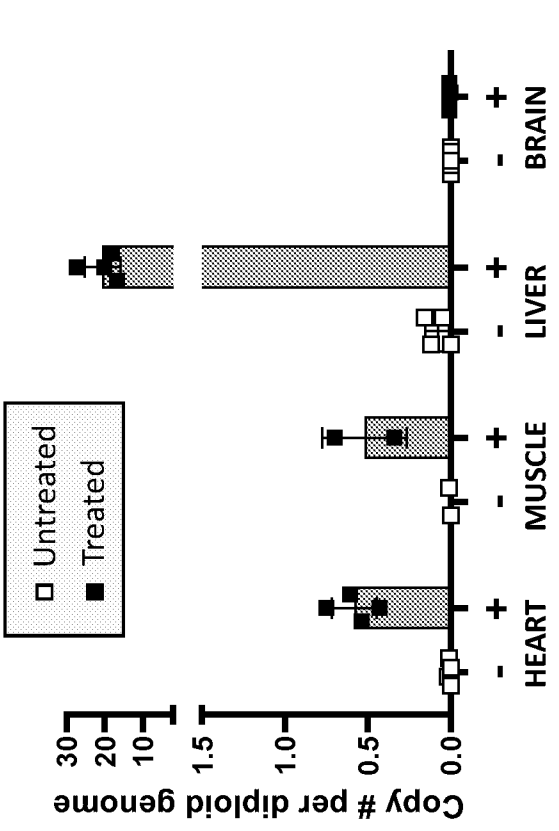

FIG. 10A

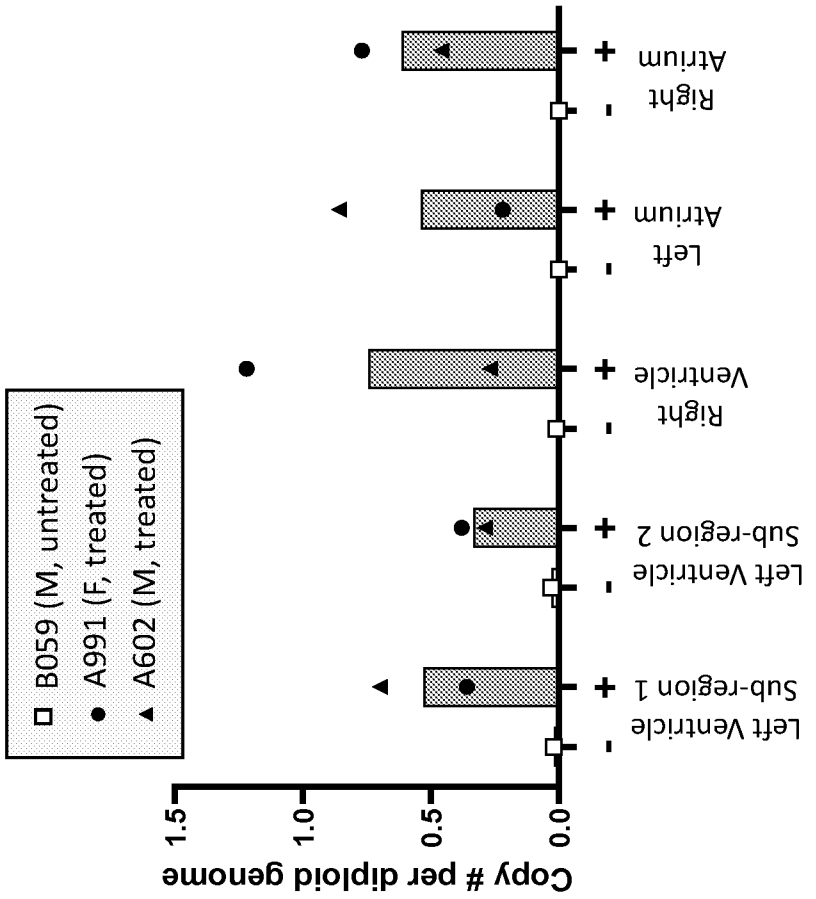

FIG. 10B

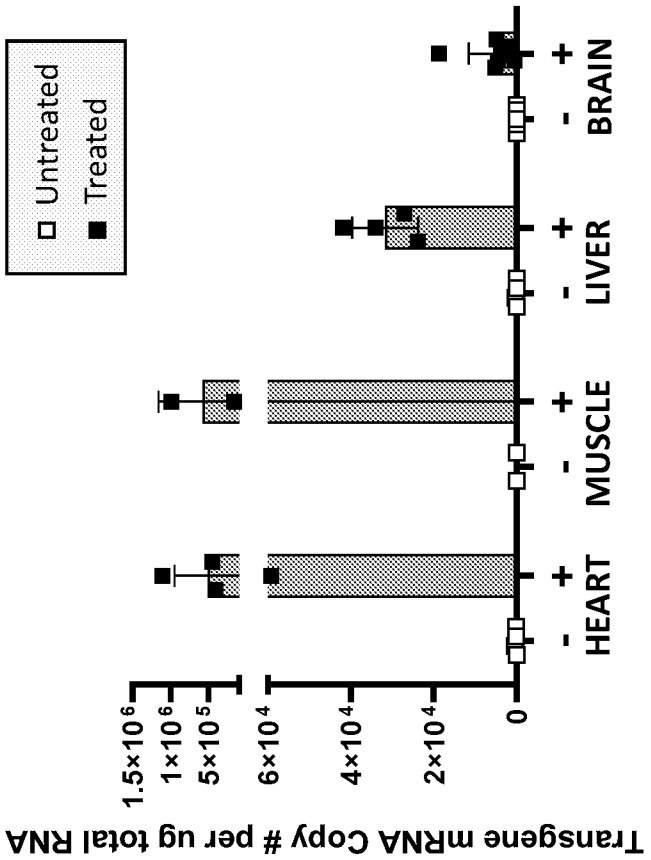
Untreated
Treated
Transgene mRNA Copy # per ug total RNA
1.5×10^6
1×10^6
5×10^5
6×10^4
4×10^4
2×10^4
0
- + - + - + - +
HEART
MUSCLE
LIVER
BRAIN

FIG. 10C

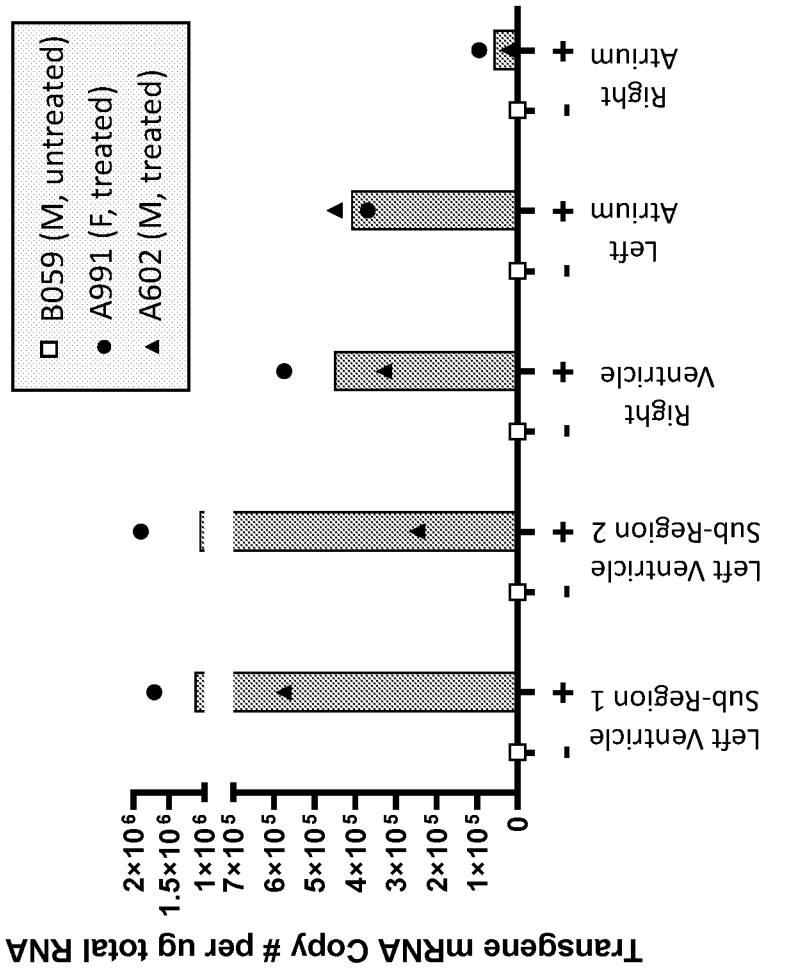
Transgene mRNA Copy # per ug total RNA
2×10^6
1.5×10^6
1×10^6
7×10^5
6×10^5
5×10^5
4×10^5
3×10^5
2×10^5
1×10^5
0
Left Ventricle Sub-Region 1
Left Ventricle Sub-Region 2
Right Ventricle
Left Atrium
Right Atrium
+
-
B059 (M, untreated)
A991 (F, treated)
A602 (M, treated)

FIG. 10D

Animal	Tissue	Location	Treatment	% Expressing Cells
B059	Heart	left ventricle	Untreated	0%
B059	Muscle	quadricep	Untreated	0%
B059	Liver	left lobe	Untreated	0%
A991	Heart	left ventricle	Treated	26-50%
A991	Heart	right ventricle	Treated	26-50%
A991	Heart	left atrium	Treated	1-25%
A991	Heart	right atrium	Treated	26-50%
A991	Muscle	quadricep	Treated	0%
A991	Muscle	gastrocnemius	Treated	0%
A991	Liver	left lobe	Treated	26-50%
A991	Liver	right lobe	Treated	51-75%
A602	Heart	left ventricle	Treated	1-25%
A602	Heart	right ventricle	Treated	1-25%
A602	Heart	left atrium	Treated	26-50%
A602	Heart	right atrium	Treated	1-25%
A602	Muscle	quadricep	Treated	1-25%
A602	Muscle	gastrocnemius	Treated	1-25%
A602	Liver	left lobe	Treated	51-75%
A602	Liver	right lobe	Treated	51-75%

FIG. 10E

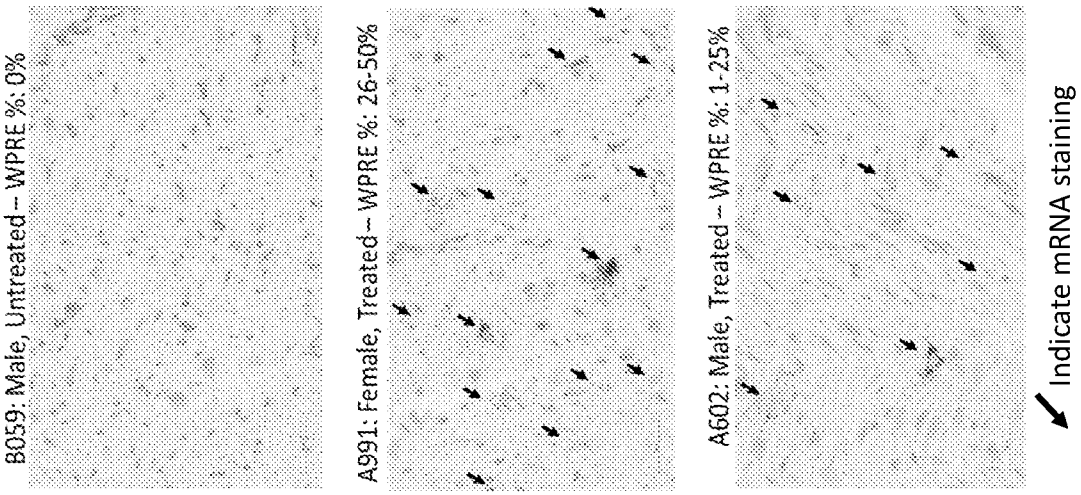

FIG. 10F

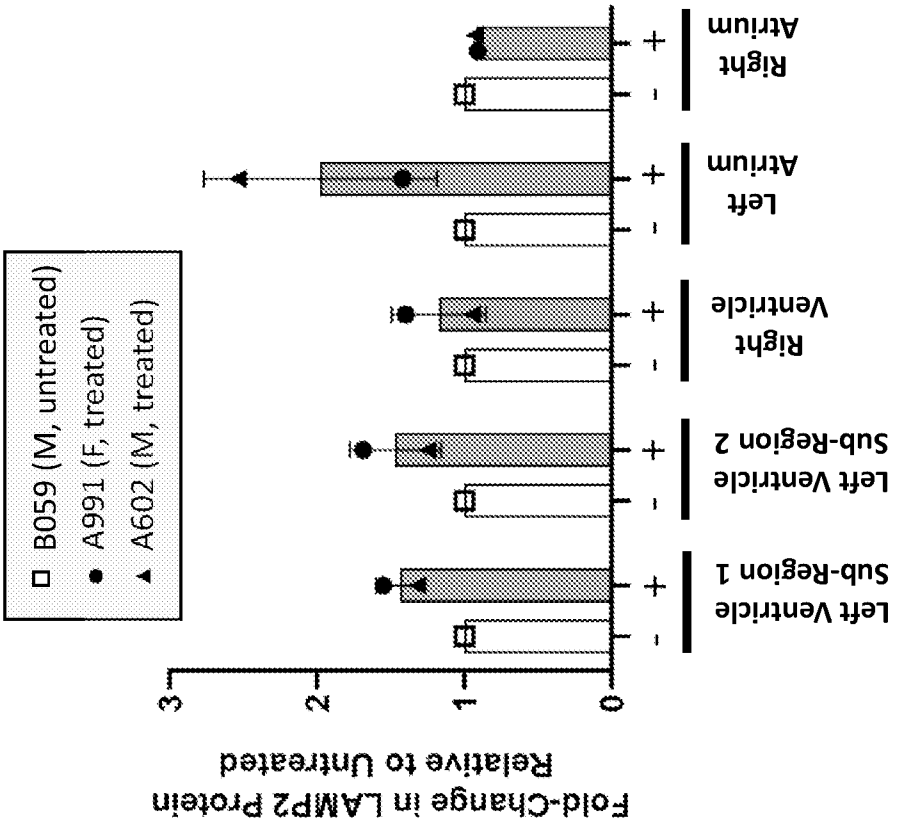

FIG. 10H

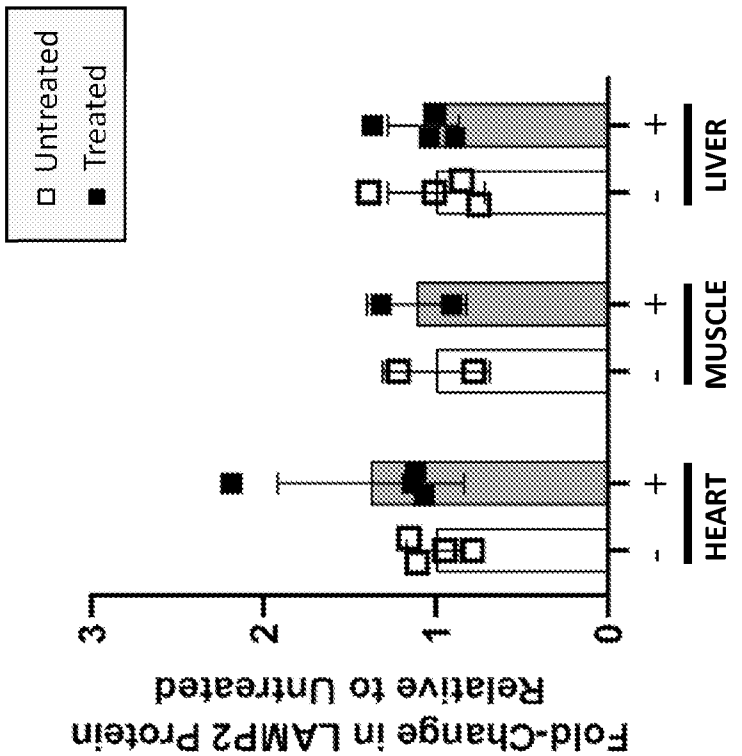

FIG. 10G

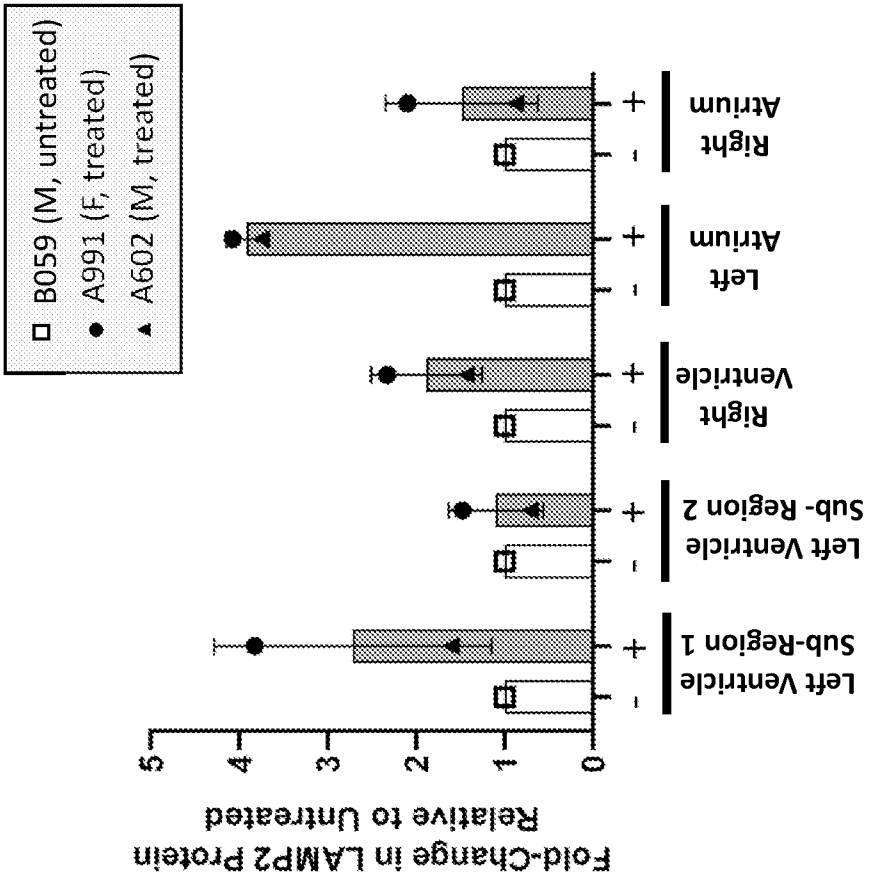

FIG. 10J

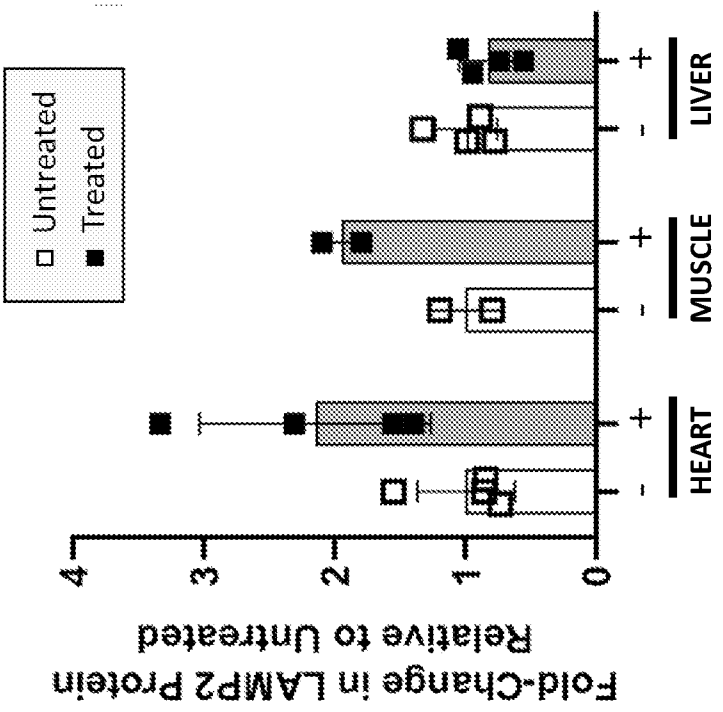

FIG. 10I

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2019/041465

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - A61K 48/00; C07K 14/47; C12N 15/09; C12N 15/63; C12N 15/79 (2019.01)
CPC - A61K 48/00; C07K 14/70596; C12N 15/09; C12N 15/11; C12N 15/79; C12N 15/86 (2019.08)

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
USPC - 435/455; 435/320.1; 514/44R; 536/23.5 (keyword delimited)

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	WO 2017/127565 A1 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 27 July 2017 (27.07.2017) entire document	1-3
Y	BROWN et al. "Target-Cell-Directed Bioengineering Approaches for Gene Therapy of Hemophilia A," Mol Ther Methods Clin Dev, 15 June 2018 (15.06.2018), Vol. 9, Pgs. 57-69. entire document	1-3
A	WO 2004/048537 A2 (EXELIXIS, INC. et al) 10 June 2004 (10.06.2004) entire document	1-3, 29-34, 48-52, 58-60, 64-67
A	US 2013/0184223 A1 (LAND et al) 18 July 2013 (18.07.2013) entire document	1-3, 29-34, 48-52, 58-60, 64-67
A	US 2016/0060656 A1 (SANGAMO BIOSCIENCES, INC.) 03 March 2016 (03.03.2016) entire document	1-3, 29-34, 48-52, 58-60, 64-67
A	US 2004/0053870 A1 (YEW et al) 18 March 2004 (18.03.2004) entire document	1-3, 29-34, 48-52, 58-60, 64-67

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

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

Date of the actual completion of the international search	Date of mailing of the international search report
03 October 2019	28 OCT 2019
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 Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (January 2015)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2019/041465

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.: 4-28, 35-47, 53-57, 61-63, 68-74
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:

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.:

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.