US 20240009247A1

(19) **United States**

(12) **Patent Application Publication**

Cherqui et al.

(10) **Pub. No.: US 2024/0009247 A1**

(43) **Pub. Date: Jan. 11, 2024**

(54) **METHODS OF TREATING LYSOSOMAL DISORDERS**

(71) Applicant: **The Regents of the University of California**, Oakland, CA (US)

(72) Inventors: **Stephanie Cherqui**, La Jolla, CA (US); **Eric Adler**, La Jolla, CA (US); **Sylvia Evans**, La Jolla, CA (US)

(21) Appl. No.: **18/342,642**

(22) Filed: **Jun. 27, 2023**

Related U.S. Application Data

(62) Division of application No. 16/493,573, filed on Sep. 12, 2019, now Pat. No. 11,806,367, filed as application No. PCT/US2018/022598 on Mar. 15, 2018.

(60) Provisional application No. 62/471,741, filed on Mar. 15, 2017, provisional application No. 62/507,713, filed on May 17, 2017.

Publication Classification

(51) **Int. Cl.**

A61K 35/28	(2006.01)
A61P 3/00	(2006.01)
A61K 35/14	(2006.01)
C07K 14/705	(2006.01)
C12N 15/86	(2006.01)

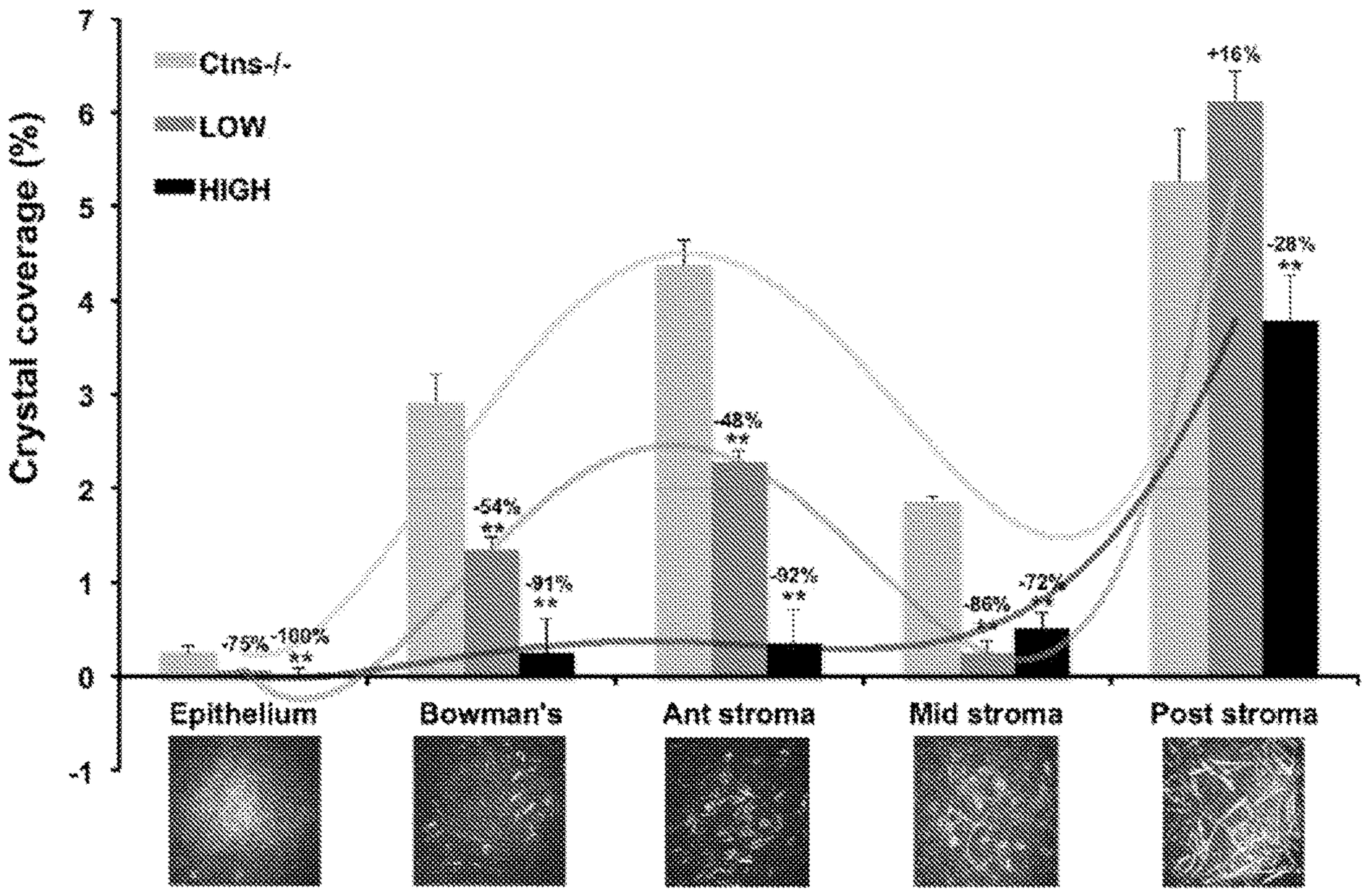
(52) **U.S. Cl.**

CPC ***A61K 35/28*** (2013.01); ***A61P 3/00*** (2018.01); ***A61K 35/14*** (2013.01); ***C07K 14/705*** (2013.01); ***C12N 15/86*** (2013.01); *A61K 38/00* (2013.01)

(57) **ABSTRACT**

Provided herein are methods for treating a lysosomal transmembrane protein disease or disorder through ex vivo introduction of a nucleic acid molecule into hematopoietic stem and progenitor cells (HSPCs) followed by transplantation of the HSPCs into a subject in need of treatment. Also provided are vectors containing the nucleic acid molecule.

Specification includes a Sequence Listing.



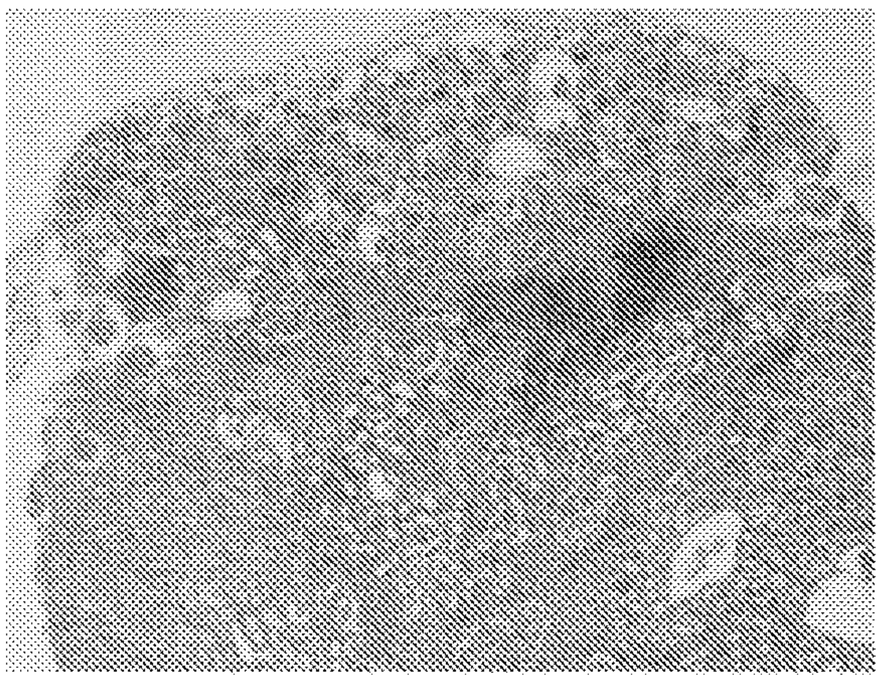


FIG. 1A

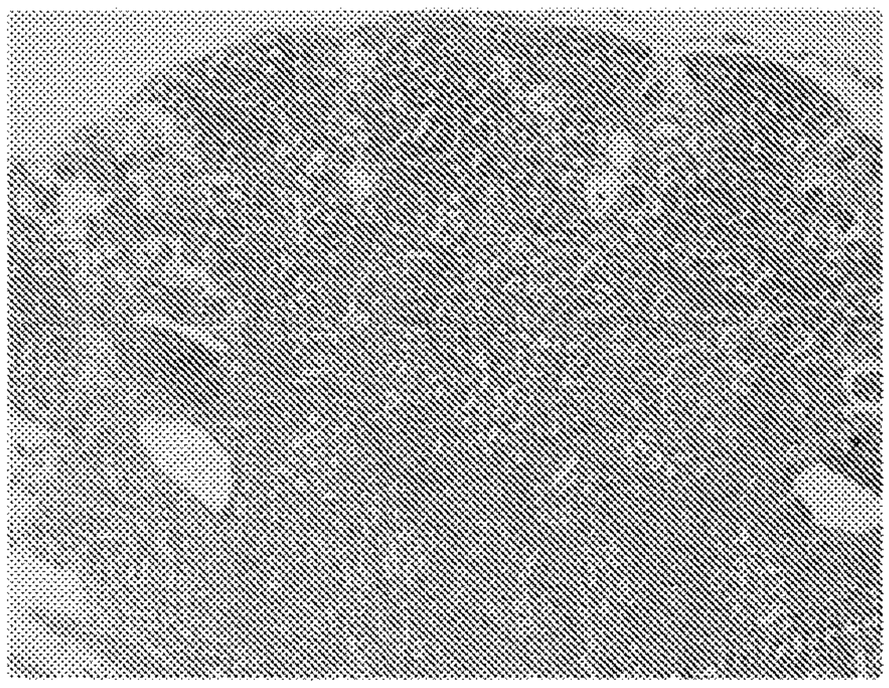


FIG. 1B

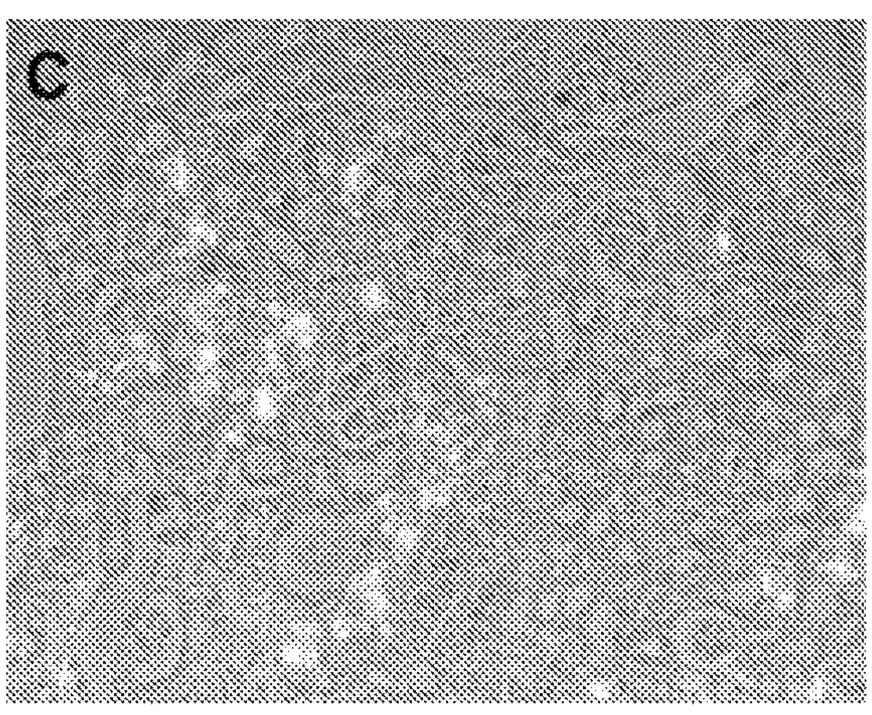


FIG. 1C

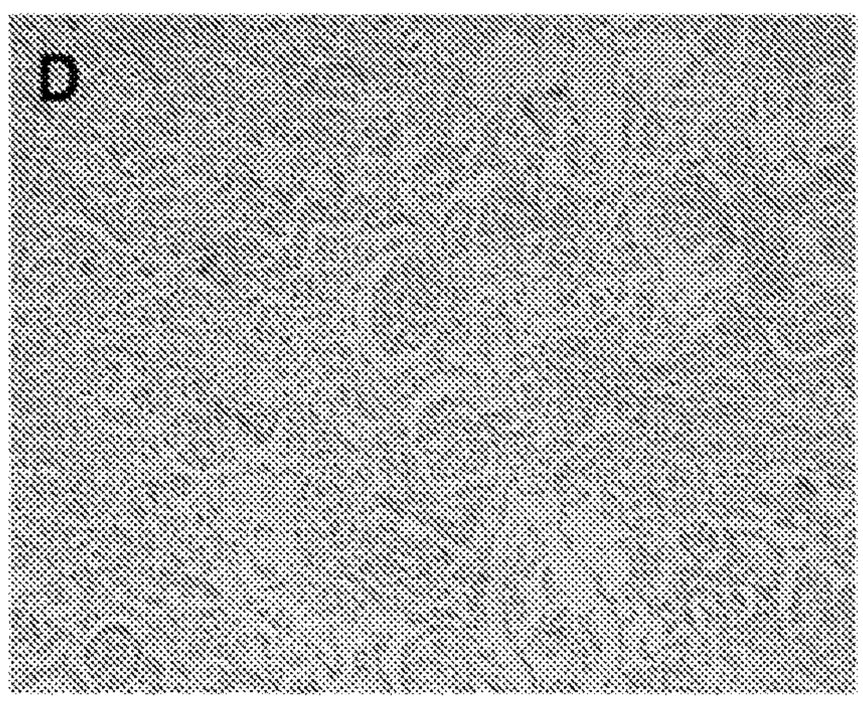


FIG. 1D

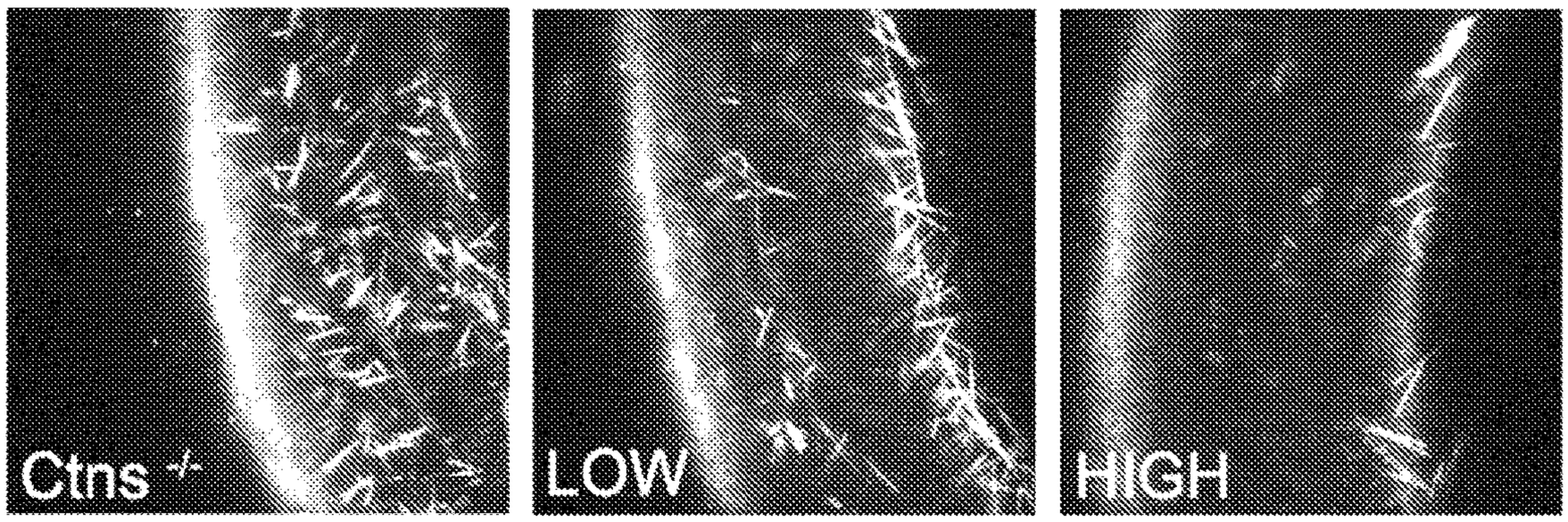

Ctns -/-
LOW
HIGH

FIG. 2A

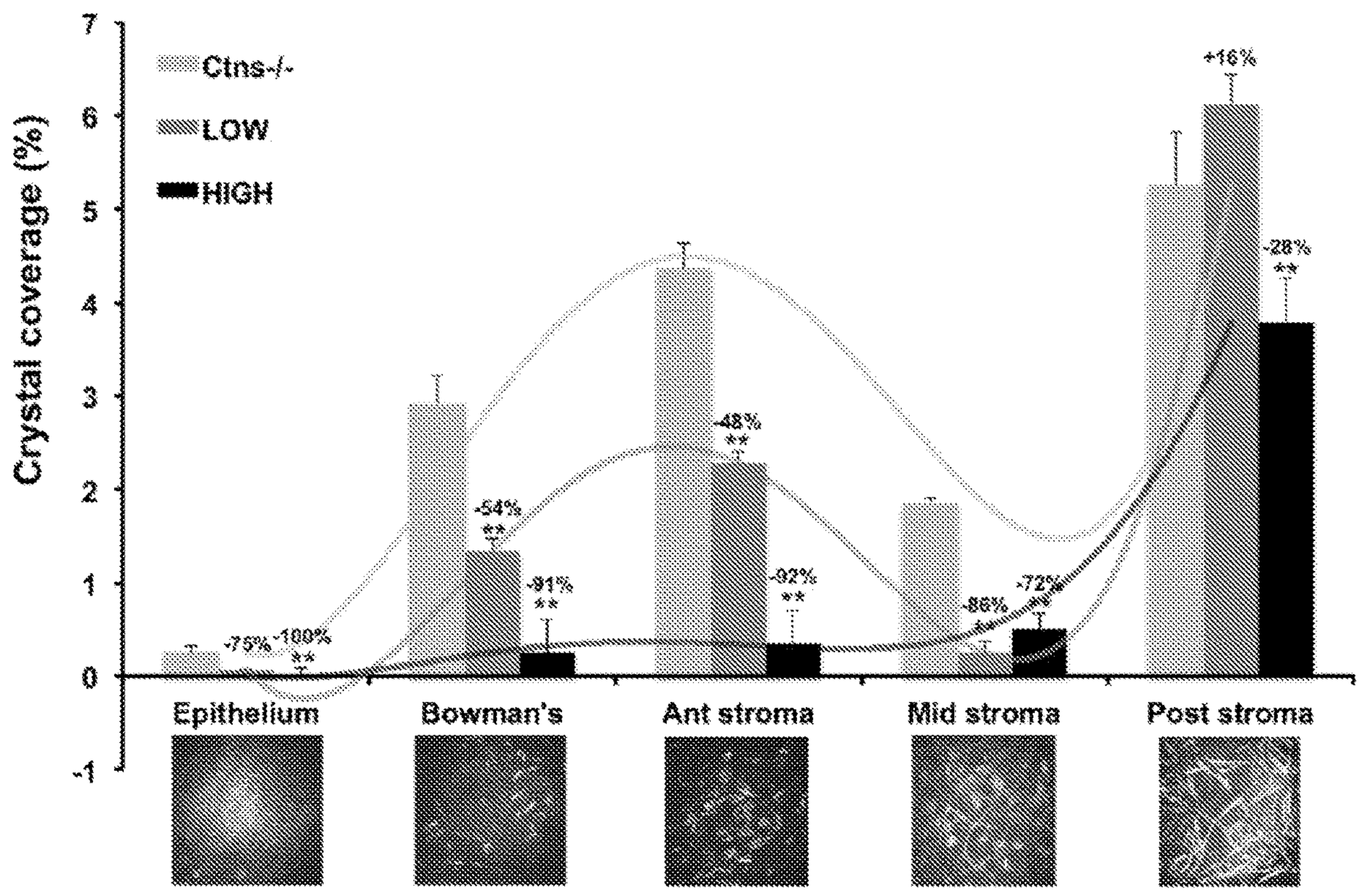

Crystal coverage (%)
7
6
5
4
3
2
1
0
-1
Ctns-/-
LOW
HIGH
-75%
-100%
**
-54%
**
-91%
**
-48%
**
-92%
**
-86%
-72%
+16%
-28%
**
Epithelium
Bowman's
Ant stroma
Mid stroma
Post stroma

FIG. 2B

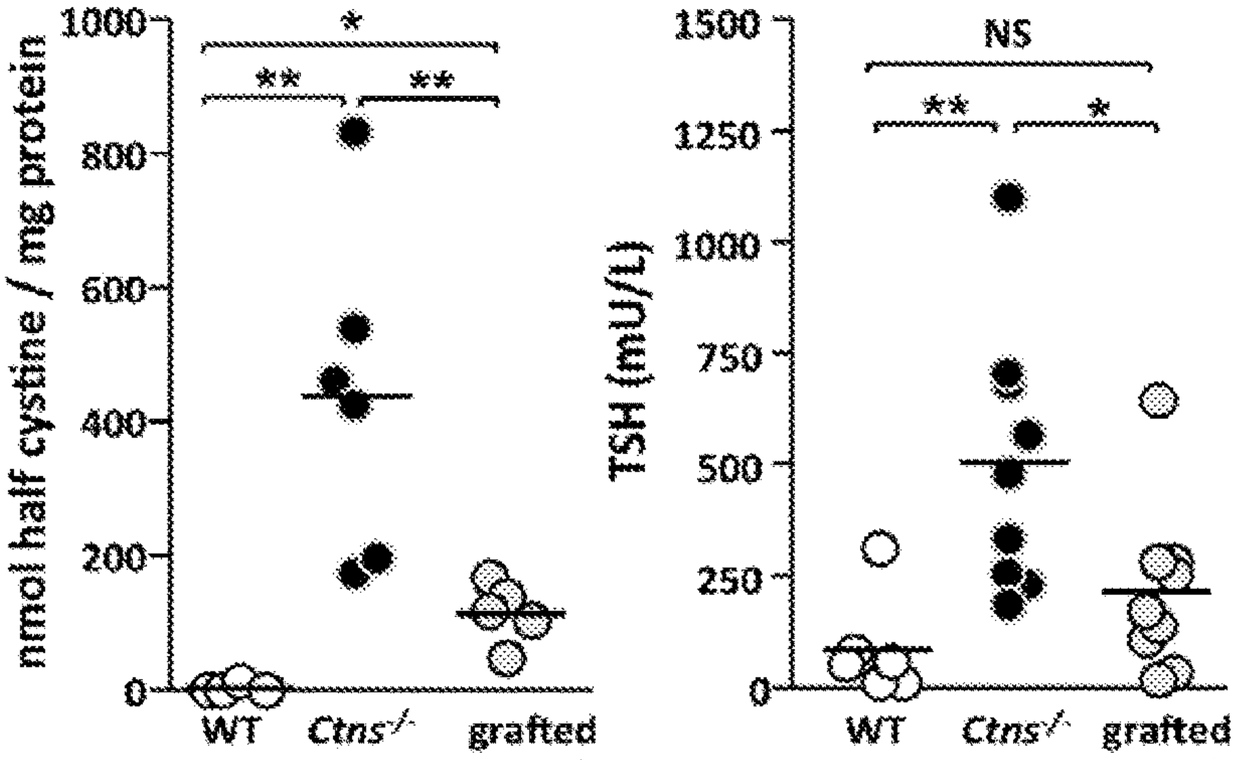


FIG. 3

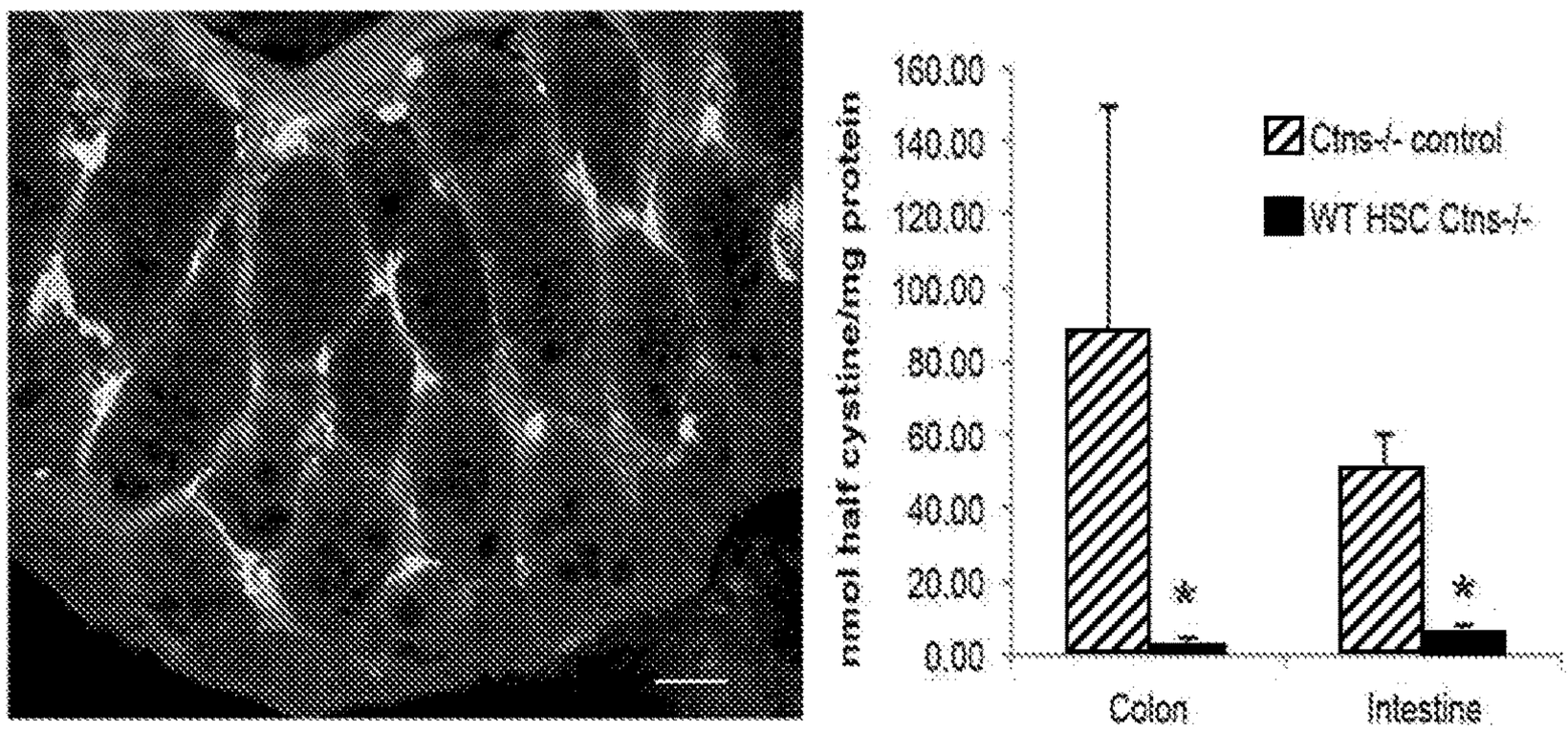


FIG. 4

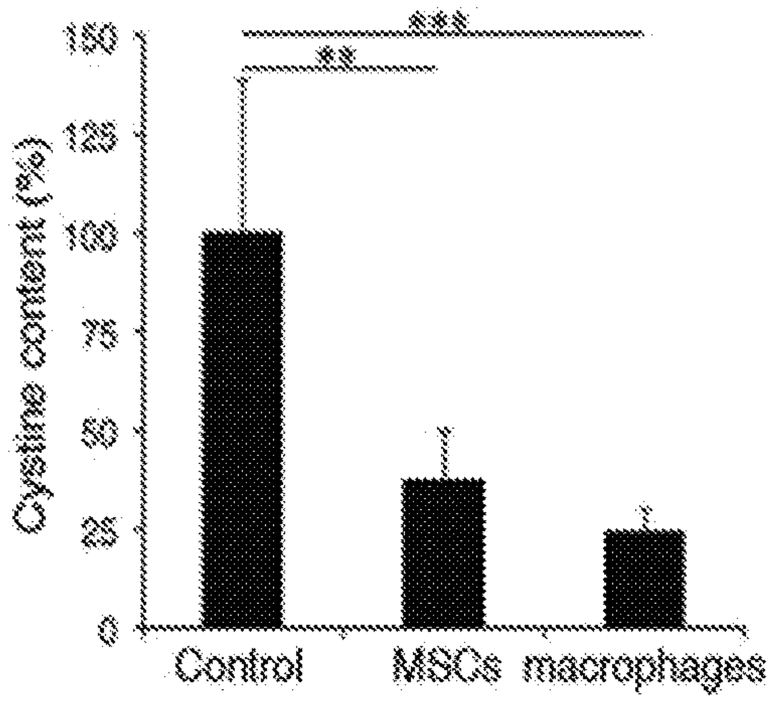


FIG. 5A

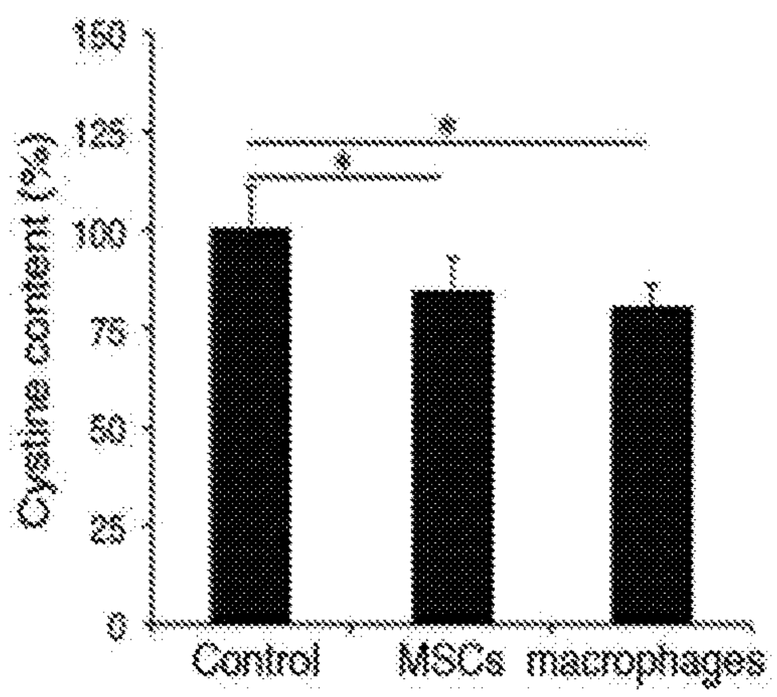


FIG. 5B

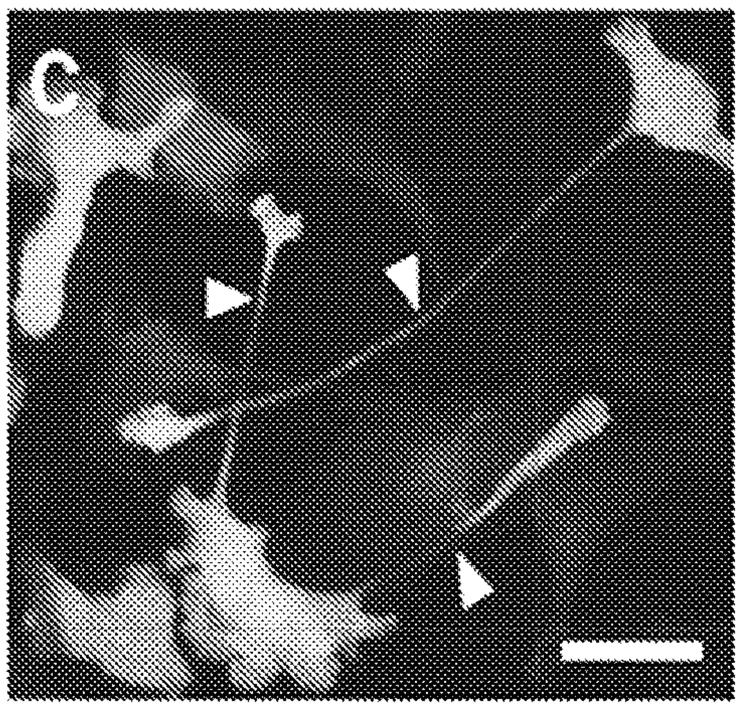


FIG. 5C

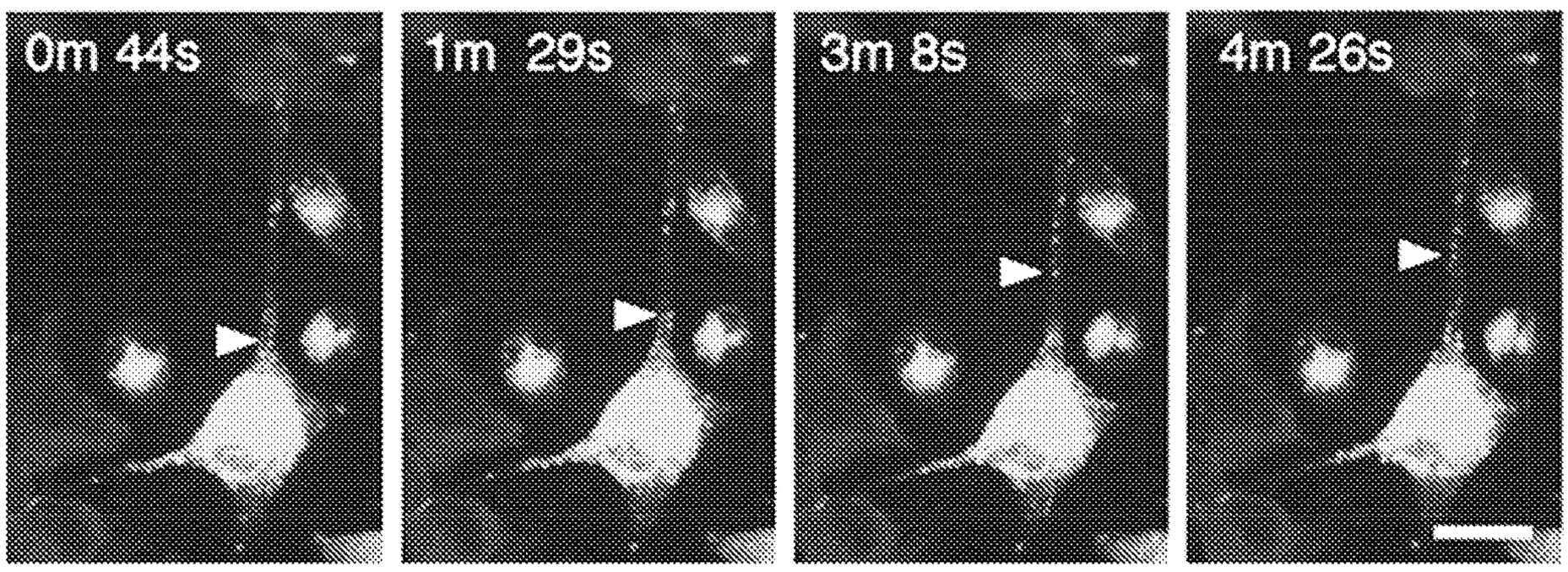


FIG. 5D

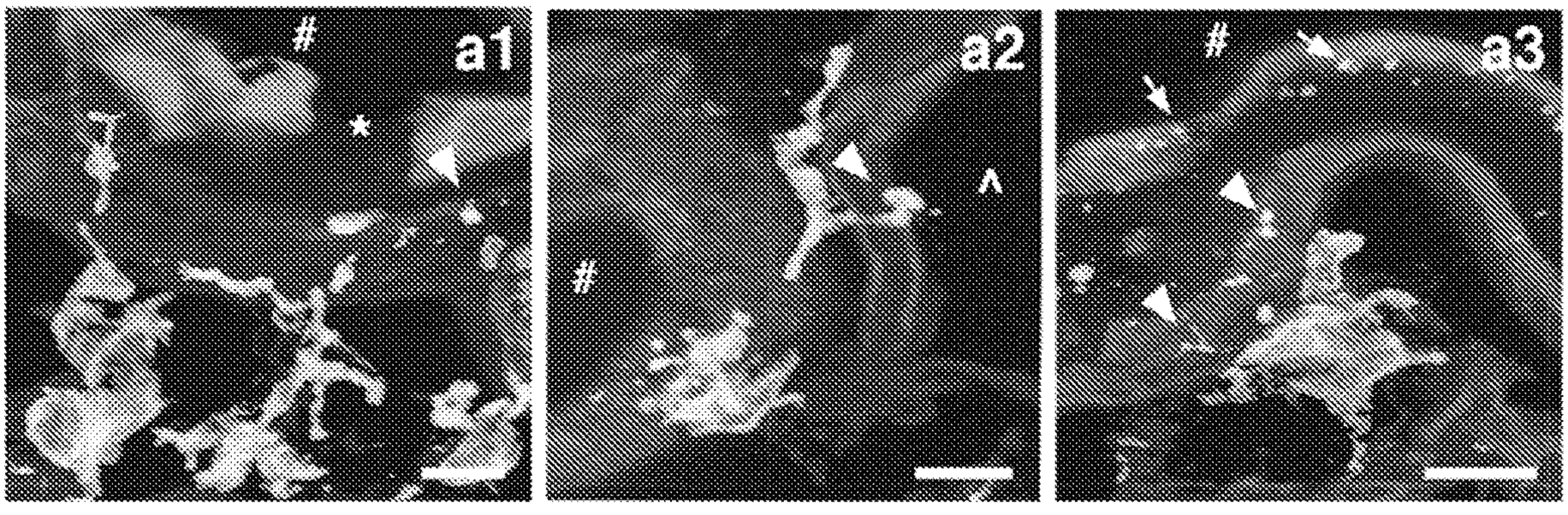


FIG. 6A

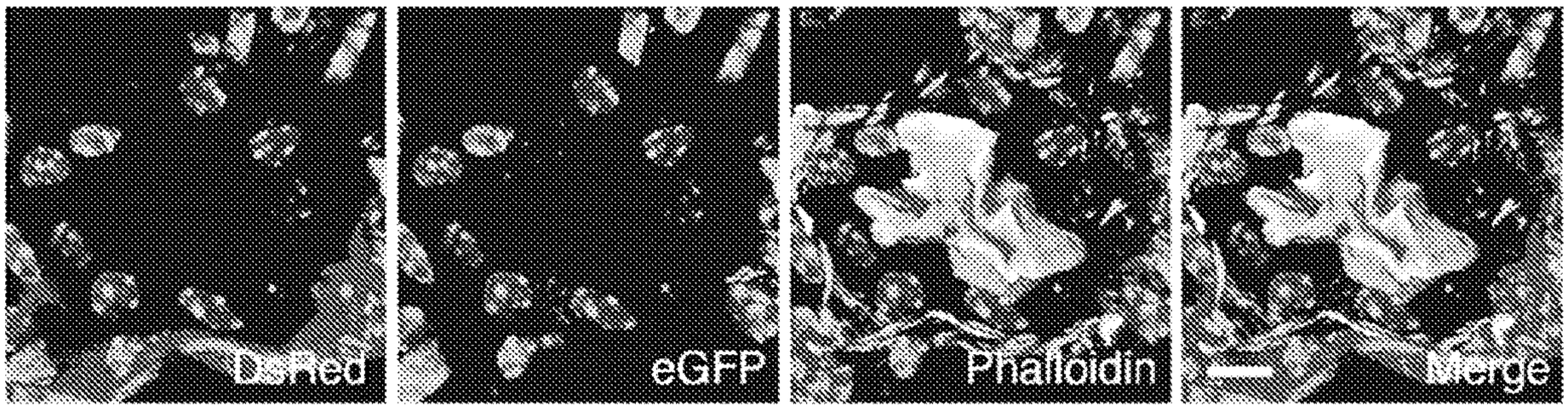


FIG. 6B

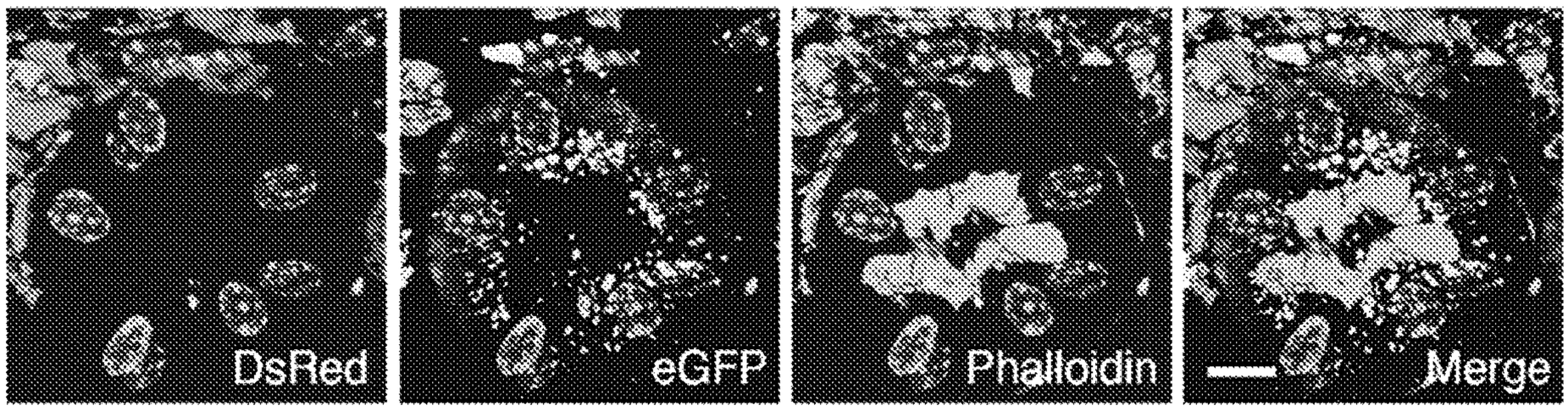


FIG. 6C

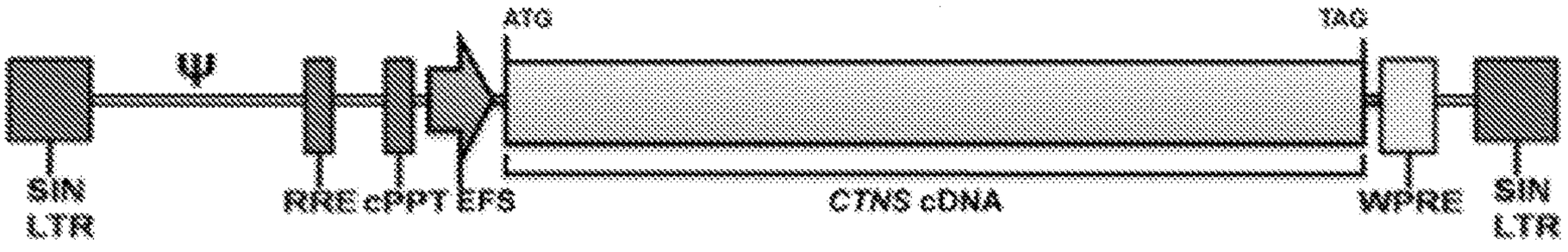


FIG. 7

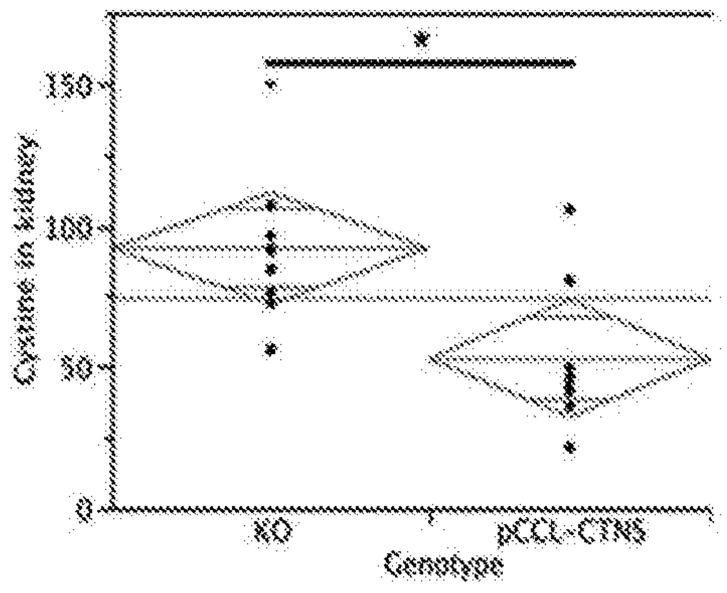


FIG. 8A

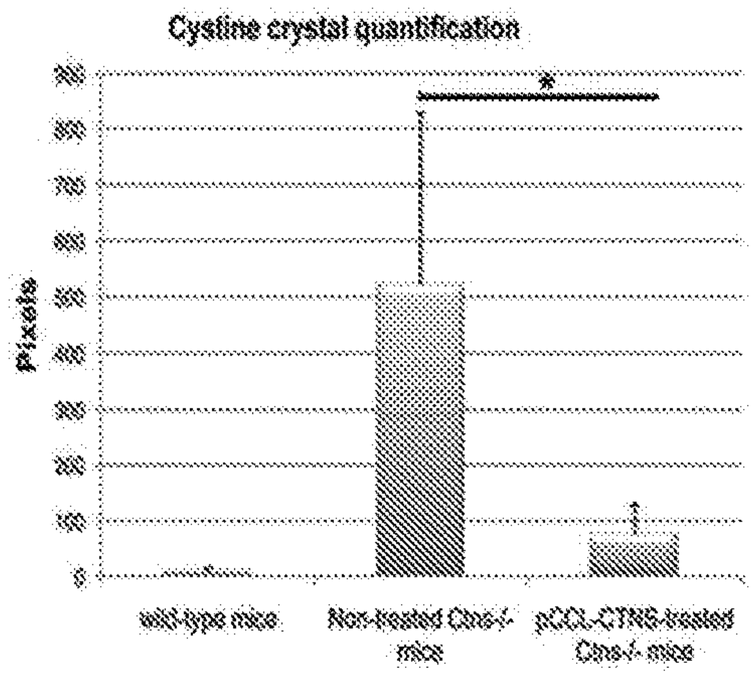


FIG. 8B

Non-Treated Ctns$^{-/-}$ mice

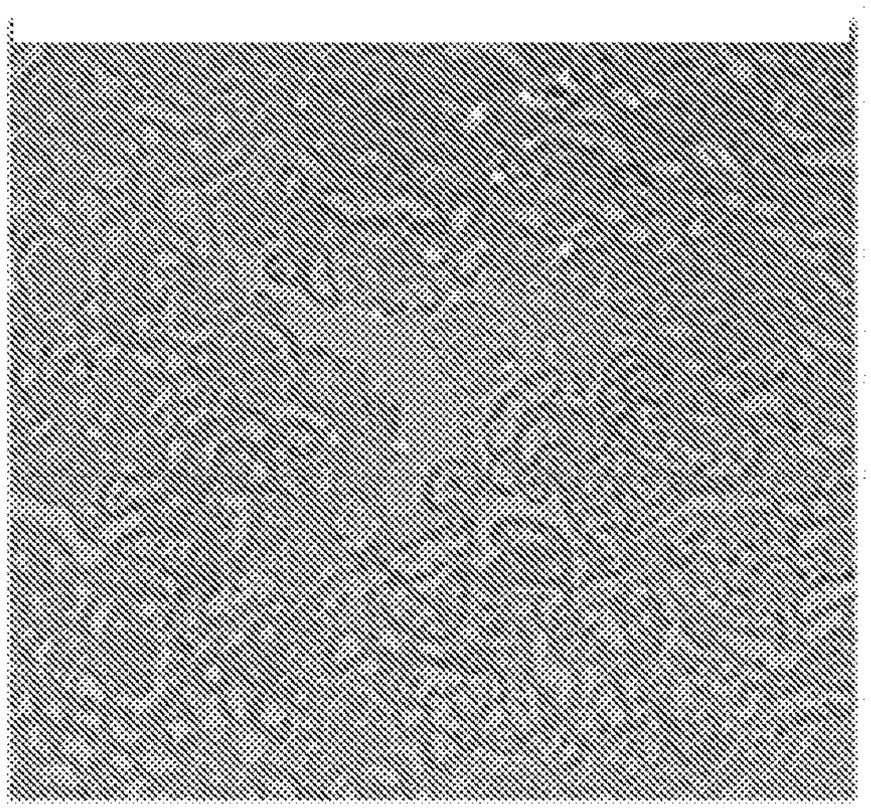


FIG. 8C

pCCL-CTNS-treated mice

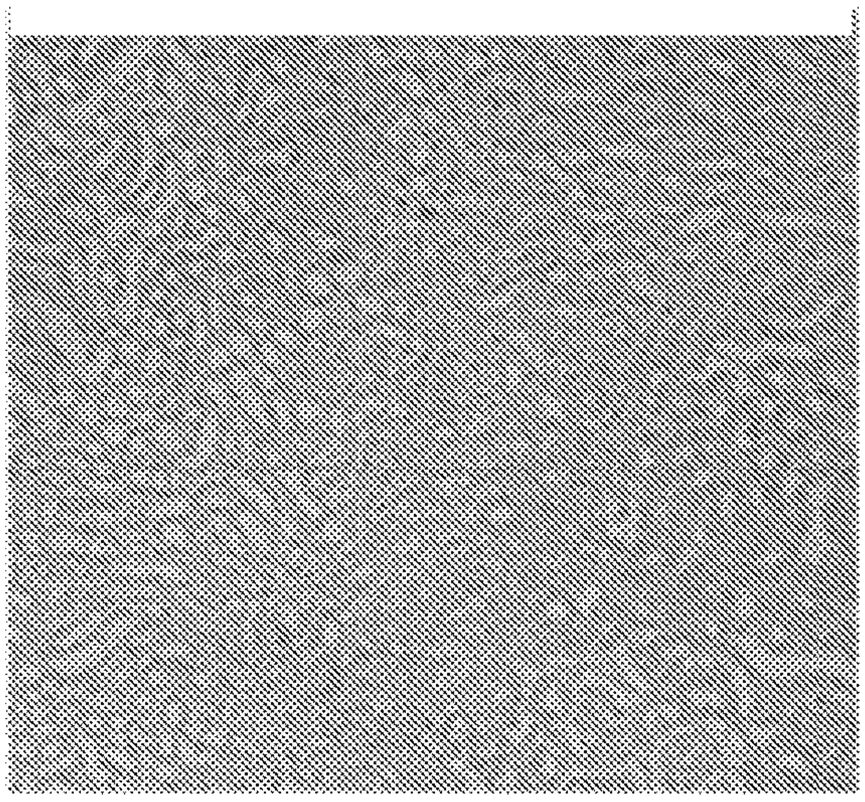


FIG. 8D

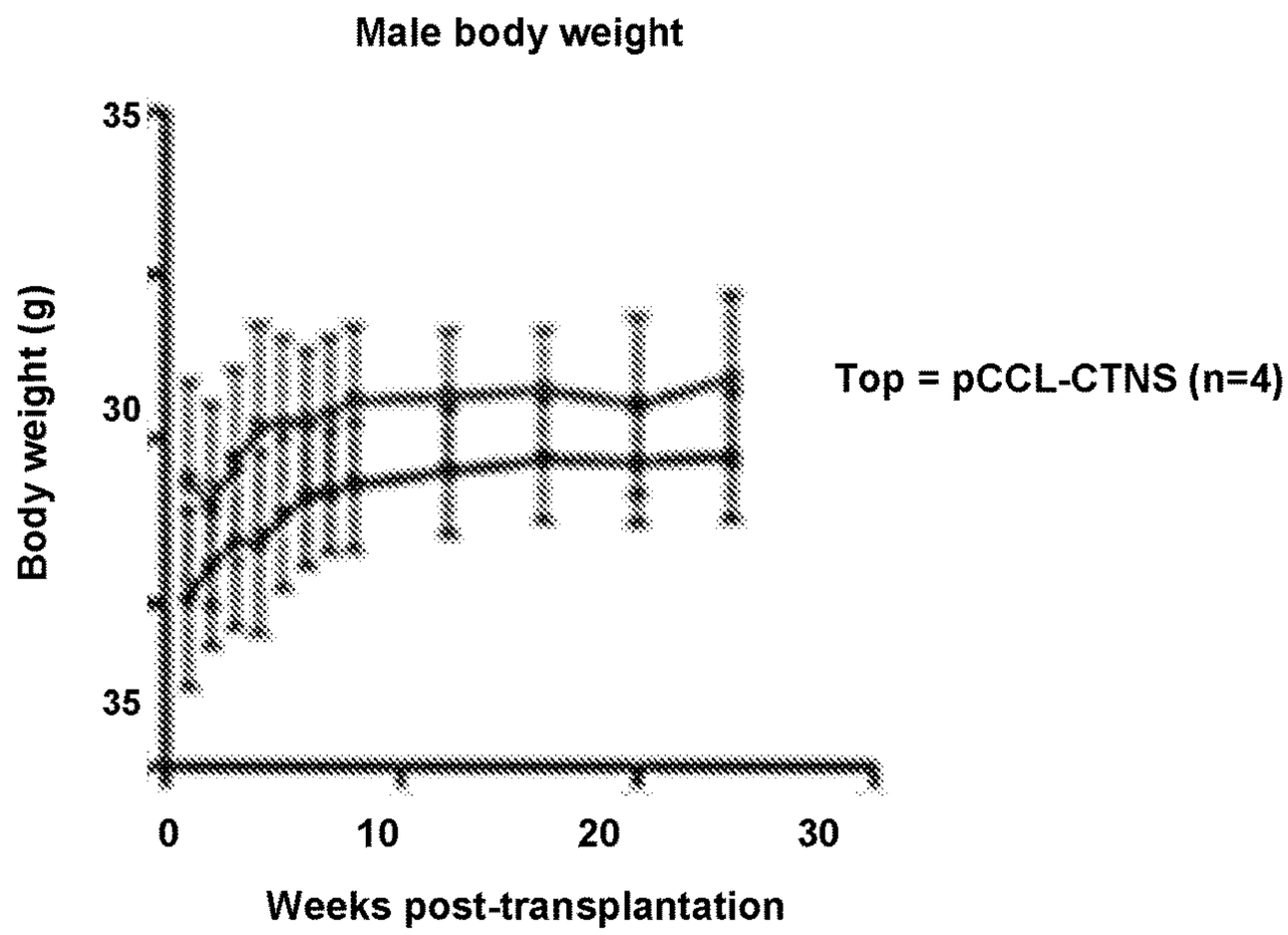


FIG. 9A-1

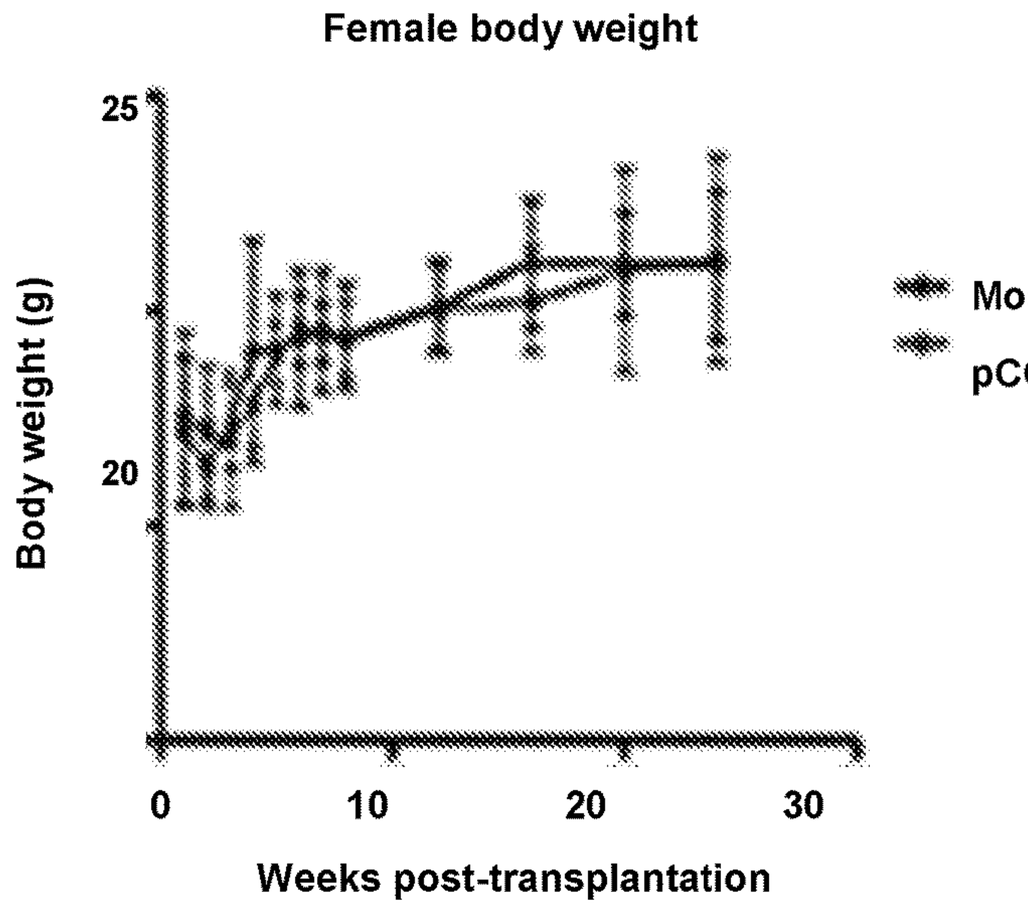


FIG. 9A-2

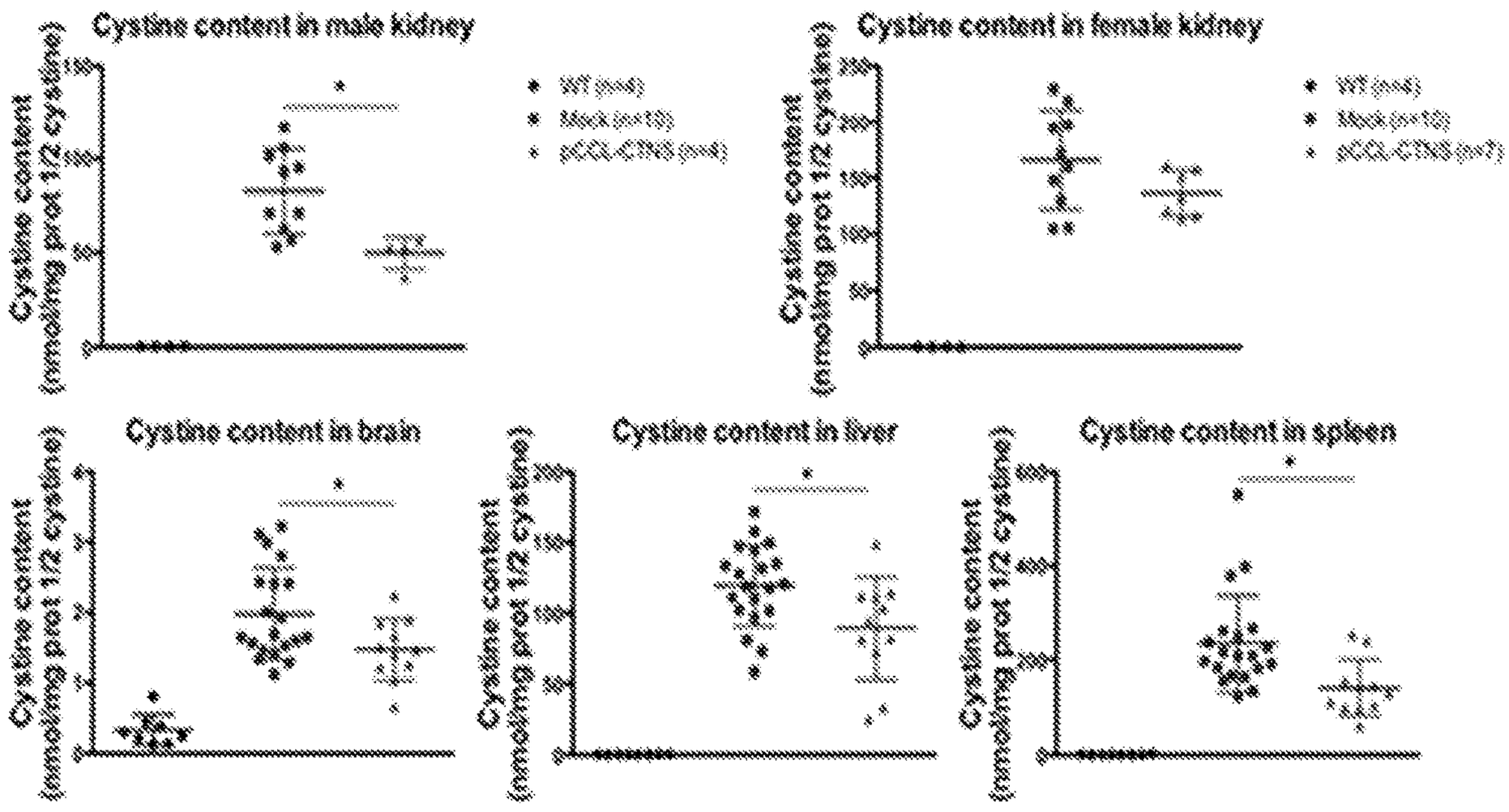


FIG. 9B

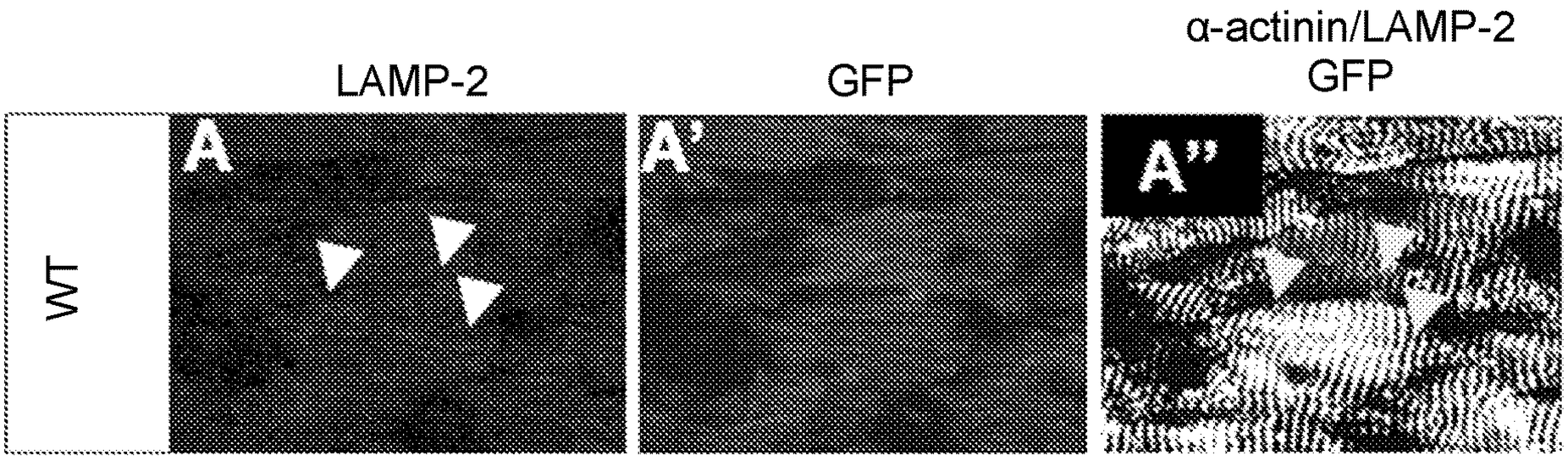


FIG. 10A

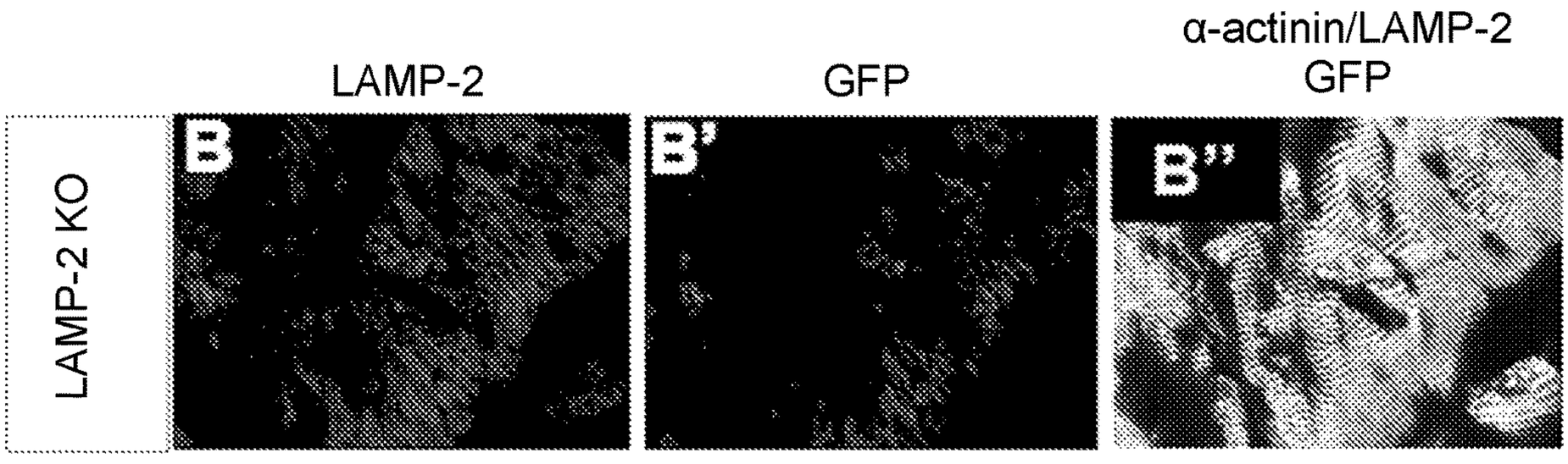


FIG. 10B

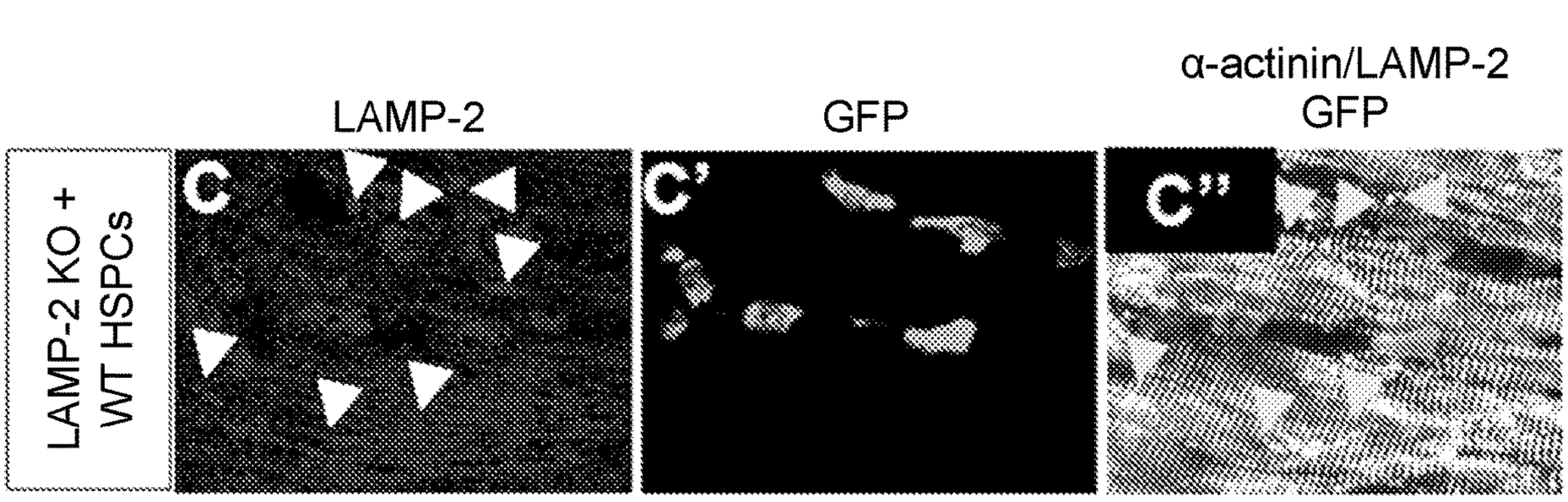


FIG. 10C

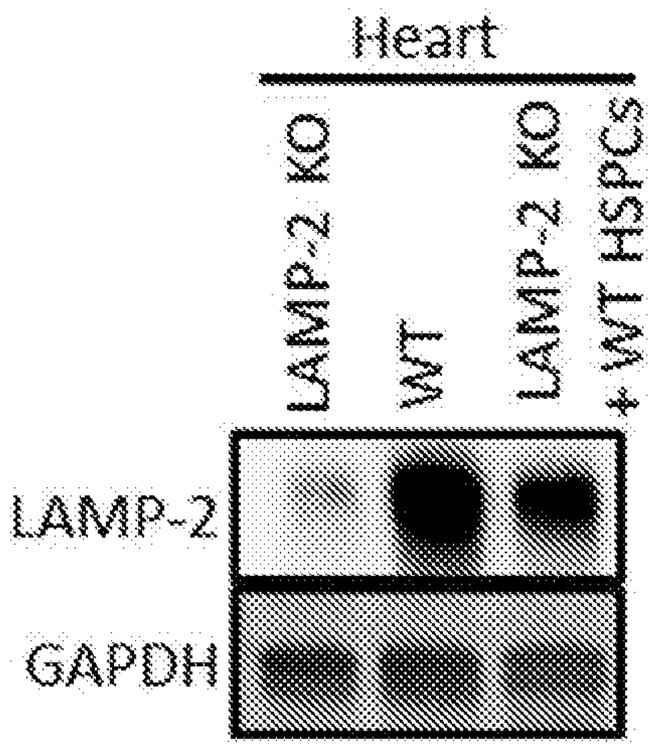


FIG. 10D

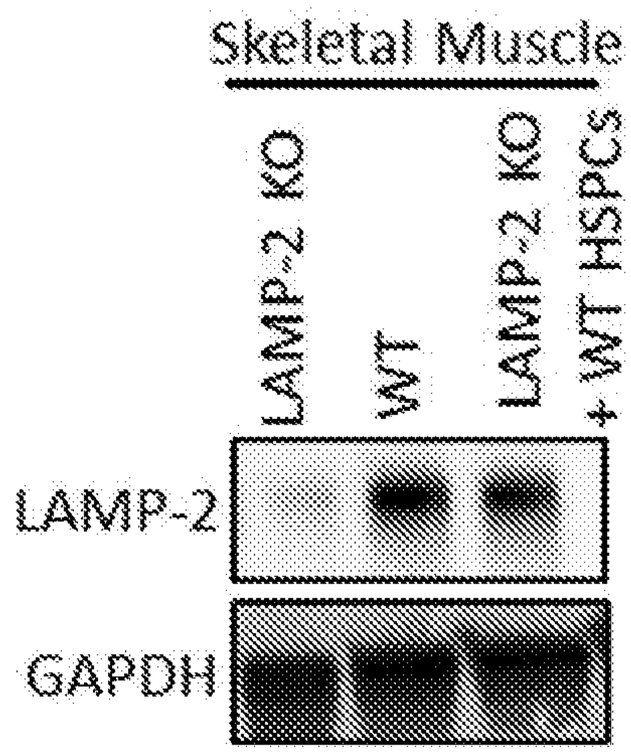


FIG. 10E

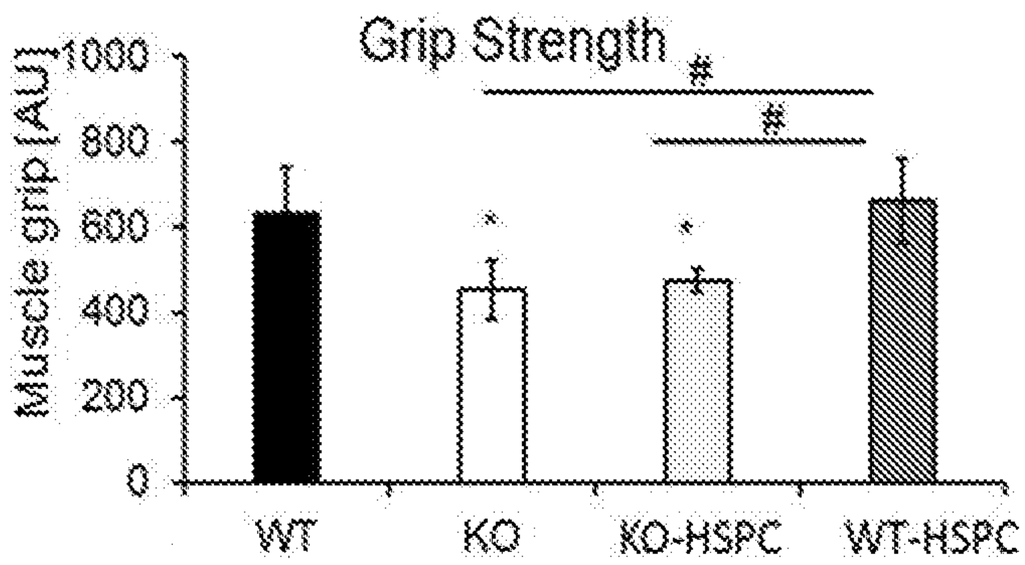


FIG. 11

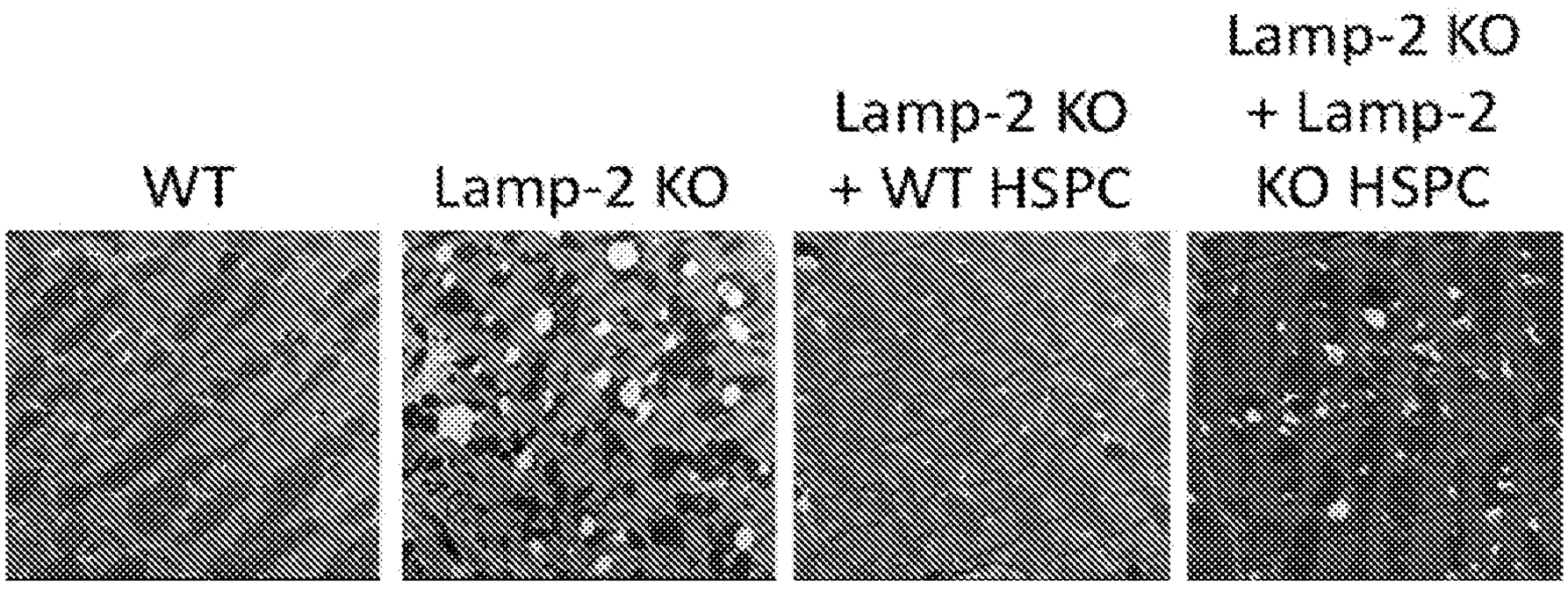


FIG. 12A

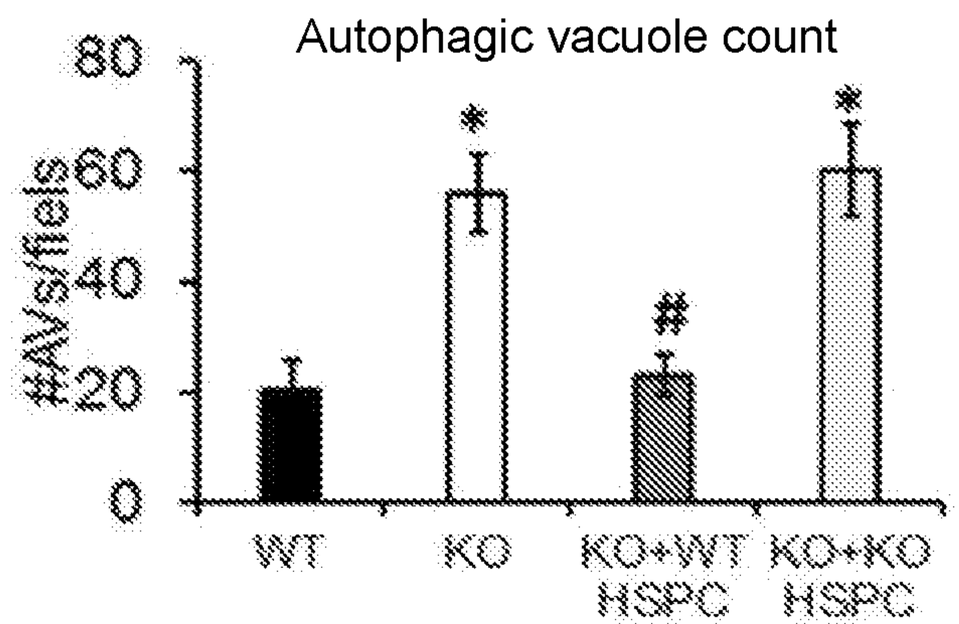


FIG. 12B

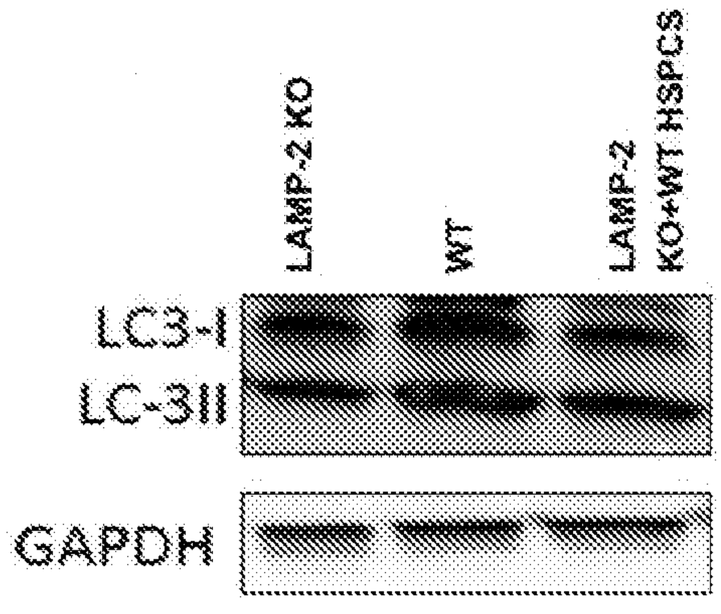


FIG. 12C

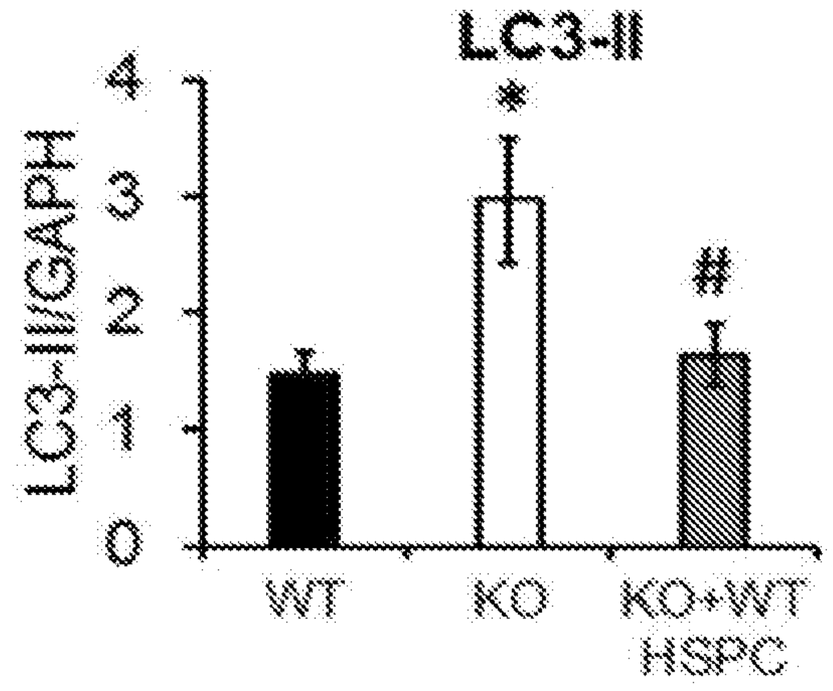


FIG. 12D

METHODS OF TREATING LYSOSOMAL DISORDERS

CROSS-REFERENCE TO RELATED PATENT APPLICATIONS

[0001] This application is a divisional application of U.S. Ser. No. 16/493,573, filed Sep. 12, 2019, which is a US national phase application under 35 U.S.C. § 371 of International Patent Application No. PCT/US2018/022598, filed Mar. 15, 2018, which claims the benefit of priority under 35 U.S.C. § 119(e) of U.S. Ser. No. 62/471,741, filed Mar. 15, 2017, and of U.S. Ser. No. 62/507,713, filed May 17, 2017, the entire content of each of which is incorporated herein by reference.

GRANT INFORMATION

[0002] This invention was made with government support under Grant Nos. DK090058 and HL107755 awarded by the National Institutes of Health. The United States government has certain rights in the invention.

SEQUENCE LISTING

[0003] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on Mar. 15, 2018, is named 20378-201753_SL.txt and is 109 kilobytes in size.

BACKGROUND OF THE INVENTION

Field of the Invention

[0004] The invention relates generally to lysosomal diseases associated with dysfunctional transmembrane lysosomal proteins and more specifically to treatment of such diseases with hematopoietic stem and progenitor cell (HSPC) gene therapy.

Background Information

[0005] Lysosomal membrane proteins act at several crucial steps of the lysosome life cycle, including lumen acidification, metabolite export, molecular motor recruitment and fusion with other organelles. Lysosomal storage diseases are a group of inherited metabolic disorders that result from defects in lysosomal function. Lysosomes are sacs of enzymes within cells that digest large molecules and pass the fragments on to other parts of the cell for recycling. This process requires several critical enzymes. If one of these enzymes is defective (for example, because of a mutation), the large molecules accumulate within the cell, eventually killing it.

[0006] Among the ~50 known lysosomal storage diseases, several are caused by lysosomal membrane protein dysfunction. One such lysosomal membrane protein disease is cystinosis, which is characterized by the abnormal accumulation of the amino acid cystine in all cells of the body leading to multi-organ failure. Cystinosis is caused by mutations in the CTNS gene that codes for cystinosin, the lysosomal membrane-specific transporter for cystine. Intracellular metabolism of cystine, as it happens with all amino acids, requires its transport across the cell membrane. After degradation of endocytosed protein to cystine within lysosomes, it is normally transported to the cytosol. But if there is a defect in the carrier protein, cystine is accumulated in lysosomes. As cystine is highly insoluble, when its concentration in tissue lysosomes increase, its solubility is immediately exceeded and crystalline precipitates are formed in almost all organs and tissues. Another example is Danon disease, which is caused by mutations in the LAMP-2 gene, a lysosomal transmembrane protein critical for autophagic flux.

[0007] To date, there are no known cures or preventative measures for such lysosomal diseases, with current therapies being directed to treating the associated symptoms. Thus, there is a need in the art for alternative or improved methods for treating lysosomal diseases/disorders.

SUMMARY OF THE INVENTION

[0008] Accordingly, in one aspect, the invention provides a method of treating a lysosomal transmembrane protein disease or disorder in a subject. The method includes introducing a corresponding functional human lysosomal transmembrane protein into hematopoietic stem and progenitor cells (HSPCs) of the subject, and transplanting the HSPCs into the subject, thereby treating the lysosomal transmembrane protein disease or disorder. Thus, when the lysosomal transmembrane protein disease or disorder is cystinosis, the corresponding functional human lysosomal transmembrane protein is cystinosin (CTNS); the lysosomal transmembrane protein disease or disorder is Salla disease or infantile sialic acid storage disorder, the corresponding functional human lysosomal transmembrane protein is sialin (SLC17A5); the lysosomal transmembrane protein disease or disorder is Cobalamin F-type disease, the corresponding functional human lysosomal transmembrane protein is LMBD1; the lysosomal transmembrane protein disease or disorder is late infantile neuronal ceroid lipofuscinosis, the corresponding functional human lysosomal transmembrane protein is CLN7; the lysosomal transmembrane protein disease or disorder is juvenile neuronal ceroid lipofuscinosis, the corresponding functional human lysosomal transmembrane protein is Battenin (CLN3); the lysosomal transmembrane protein disease or disorder is malignant infantile osteopetrosis, the corresponding functional human lysosomal transmembrane protein is C1C-7 or OSTM1; the lysosomal transmembrane protein disease or disorder is mucolipidosis IV, the corresponding functional human lysosomal transmembrane protein is TRPML-1; the lysosomal transmembrane protein disease or disorder is mucopolysaccharidosis type IIIC, the corresponding functional human lysosomal transmembrane protein is HGSNAT; the lysosomal transmembrane protein disease or disorder is Neiman-Pick Type C, the corresponding functional human lysosomal transmembrane protein is NPC-1; and the lysosomal transmembrane protein disease or disorder is Danon disease, the corresponding functional human lysosomal transmembrane protein is LAMP2.

[0009] In various embodiments, the step of introducing may include contacting a vector comprising a polynucleotide encoding functional human lysosomal transmembrane protein and a functional promoter with the HSPCs and allowing expression of the functional human lysosomal transmembrane protein. In various embodiments, the lysosomal transmembrane protein disease or disorder is cystinosis and the functional human lysosomal transmembrane protein is CTNS. In various embodiments, the lysosomal transmembrane protein disease or disorder is Danon disease

and the functional human lysosomal transmembrane protein is LAMP2. The LAMP2 may be an isoform selected from the group consisting of LAMP-2A, LAMP-2B, LAMP-2C. The subject may be a mammal, such as a human. In various embodiments, the vector is a viral vector selected from the group consisting of a lentiviral, adenoviral, and AAV vector. In various embodiments, the vector is a lentiviral vector. In various embodiments, the vector is an adenoviral vector. In various embodiments, the vector is an AAV vector. In various embodiments, the vector is a self-inactivating (SIN)-lentivirus vector, such as pCCL-CTNS or pCCL-LAMP2. In various embodiments, the step of introducing is performed ex vivo. In various embodiments, the HSPCs are isolated from the bone marrow of the subject.

[0010] In another aspect, the present invention provides an expression cassette comprising a promoter functionally linked to a transgene encoding a functional human lysosomal transmembrane protein selected from the group consisting of CTNS, SLC17A5, LMBRD1, CLN7, CLN3, CLC-7, OSTM1, TRPML1, HGSNAT, NPC1, and LAMP2. Also provided are a vector, such as a self-inactivating (SIN)-lentivirus vector, that includes a promoter functionally linked to a polynucleotide encoding a functional human lysosomal transmembrane protein selected from the group consisting of CTNS, SLC17A5, LMBRD1, CLN7, CLN3, CLC-7, OSTM1, TRPML1, HGSNAT, NPC1, and LAMP2. In various embodiments, the functional human lysosomal transmembrane protein is CTNS. In various embodiments, the functional human lysosomal transmembrane protein is LAMP2.

[0011] In another aspect, the present invention provides a method of treating or ameliorating a lysosomal protein disease or disorder in a subject. The method includes isolating hematopoietic stem and HSPCs cells from a subject's bone marrow, introducing a functional human lysosomal transmembrane gene into the HSPCs, wherein the gene encodes a protein corresponding to the lysosomal protein disease or disorder, and transplanting the HSPCs back into the subject, thereby treating or ameliorating the lysosomal protein disease or disorder. Thus, when the lysosomal transmembrane protein disease or disorder is cystinosis, the functional human lysosomal transmembrane gene is CTNS; the lysosomal transmembrane protein disease or disorder is Salla disease or infantile sialic acid storage disorder, the functional human lysosomal transmembrane gene is SLC17A5; the lysosomal transmembrane protein disease or disorder is Cobalamin F-type disease, the functional human lysosomal transmembrane gene is LMBRD1; the lysosomal transmembrane protein disease or disorder is late infantile neuronal ceroid lipofuscinosis, the functional human lysosomal transmembrane gene is MFSD8; the lysosomal transmembrane protein disease or disorder is juvenile neuronal ceroid lipofuscinosis, the functional human lysosomal transmembrane gene is CLN3; the lysosomal transmembrane protein disease or disorder is malignant infantile osteopetrosis, the functional human lysosomal transmembrane gene is CLCN7 or OSTM1; the lysosomal transmembrane protein disease or disorder is mucolipidosis IV, the functional human lysosomal transmembrane gene is MCOLN1; the lysosomal transmembrane protein disease or disorder is mucopolysaccharidosis type IIIC, the functional human lysosomal transmembrane gene is HGSNAT; the lysosomal transmembrane protein disease or disorder is Neiman-Pick Type C, the functional human lysosomal transmembrane gene is NPC1; and the lysosomal transmembrane protein disease or disorder is Danon disease, the functional human lysosomal transmembrane gene is LAMP2.

[0012] In various embodiments, the HSPCs are CD34+ cells. In various embodiments the lysosomal protein disease or disorder is cystinosis and the functional human lysosomal transmembrane gene is CTNS. In various embodiments, the lysosomal protein disease or disorder is Danon disease and the functional human lysosomal transmembrane gene is LAMP2. In various embodiments, the step of introducing the functional human CTNS gene into the HSPCs includes using a vector, such as a viral vector. In various embodiments, the vector is a viral vector selected from the group consisting of a lentiviral, adenoviral, and AAV vector. In various embodiments, the level of cystine in the eye, skin, leukocytes, parenchymal tissue or gastrointestinal tract of the subject is reduced following treatment. In various embodiments, the dosage is about 1.0×10^6 to 5.0×10^6 cells/kg, such as 2.5×10^6 cells/kg, administered as a single dose.

[0013] The subject may be on cysteamine therapy, such as oral cysteamine therapy, prior to treatment. The dose administration may be intravenous. In various embodiments, cystine or cystine crystals are measure in the eye, skin, leukocytes, parenchymal tissue and/or gastrointestinal tract prior to and/or following treatment. In various embodiments, cystine or cystine crystals are measured in the eye prior to and/or following treatment. In various embodiments, cystine crystals are measured using in vivo confocal microscopy. In various embodiments, cystine levels may be measured prior to, during and/or following treatment. In various embodiments, cystine levels are measured using biological samples, such as blood, rectal biopsies, or buccal mucosa. In various embodiments, cystine levels are measured from rectal biopsies.

[0014] In another aspect, the present invention provides a method of treating or ameliorating a lysosomal protein disease or disorder in a subject. The method includes producing a functional human lysosomal transmembrane gene in the subject using gene editing. Thus, when the lysosomal transmembrane protein disease or disorder is cystinosis, the functional human lysosomal transmembrane gene is CTNS; the lysosomal transmembrane protein disease or disorder is Salla disease or infantile sialic acid storage disorder, the functional human lysosomal transmembrane gene is SLC17A5; the lysosomal transmembrane protein disease or disorder is Cobalamin F-type disease, the functional human lysosomal transmembrane gene is LMBRD1; the lysosomal transmembrane protein disease or disorder is late infantile neuronal ceroid lipofuscinosis, the functional human lysosomal transmembrane gene is MFSD8; the lysosomal transmembrane protein disease or disorder is juvenile neuronal ceroid lipofuscinosis, the functional human lysosomal transmembrane gene is CLN3; the lysosomal transmembrane protein disease or disorder is malignant infantile osteopetrosis, the functional human lysosomal transmembrane gene is CLCN7 or OSTM1; the lysosomal transmembrane protein disease or disorder is mucolipidosis IV, the functional human lysosomal transmembrane gene is MCOLN1; the lysosomal transmembrane protein disease or disorder is mucopolysaccharidosis type IIIC, the functional human lysosomal transmembrane gene is HGSNAT; the lysosomal transmembrane protein disease or disorder is Neiman-Pick Type C, the functional human lysosomal transmembrane gene is NPC1; and the lysosomal transmembrane protein

disease or disorder is Danon disease, the functional human lysosomal transmembrane gene is LAMP2.

[0015] In another aspect, the present invention provides a method of treating or ameliorating a lysosomal protein disease or disorder in a subject. The method includes contacting cells expressing a defective lysosomal transmembrane protein from the subject with a vector encoding a gene editing system that, when transfected into the cells, removes a trinucleotide extension mutation of the gene encoding the endogenous lysosomal transmembrane protein, thereby treating the lysosomal protein disease or disorder. Thus, when the lysosomal transmembrane protein disease or disorder is cystinosis, the lysosomal transmembrane protein is cystinosin (CTNS); the lysosomal transmembrane protein disease or disorder is Salla disease or infantile sialic acid storage disorder, the lysosomal transmembrane protein is sialin (SLC17A5); the lysosomal transmembrane protein disease or disorder is Cobalamin F-type disease, the lysosomal transmembrane protein is LMBD1; the lysosomal transmembrane protein disease or disorder is late infantile neuronal ceroid lipofuscinosis, the lysosomal transmembrane protein is CLN7; the lysosomal transmembrane protein disease or disorder is juvenile neuronal ceroid lipofuscinosis, the lysosomal transmembrane protein is Battenin (CLN3); the lysosomal transmembrane protein disease or disorder is malignant infantile osteopetrosis, the lysosomal transmembrane protein is C1C-7 or OSTM1; the lysosomal transmembrane protein disease or disorder is mucolipidosis IV, the lysosomal transmembrane protein is TRPML-1; the lysosomal transmembrane protein disease or disorder is mucopolysaccharidosis type IIIC, the lysosomal transmembrane protein is HGSNAT; the lysosomal transmembrane protein disease or disorder is Neiman-Pick Type C, the lysosomal transmembrane protein is NPC-1; and the lysosomal transmembrane protein disease or disorder is Danon disease, the lysosomal transmembrane protein is LAMP2.

[0016] In various embodiments, the gene editing system is selected from the group consisting of CRISPR/Cas, zinc finger nucleases, engineered meganucleases, ARCUS, and transcription activator-life effector nucleases. In various embodiments, the step of contacting comprises administering to the subject an effective amount of the vector. In various embodiments, the step of contacting comprises obtaining a sample of cells from the subject, transfecting the gene editing system into the sample of cells, and thereafter, transplanting the transfected cells into the subject. In various embodiments, the sample of cells is selected from the group consisting of blood cells and HSPCs.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] FIGS. **1**A-**1**D are pictorial diagrams showing histological analyses of kidney sections of 15 months old $Ctns^{-/-}$ mice. FIGS. **1**A and **1**B show the results from hematoxylin & eosin staining revealing severe anomalies in $Ctns^{-/-}$ mice (FIG. **1**A) whereas HSC-transplanted $Ctns^{-/-}$ mice exhibited only focal anomalies (FIG. **1**B). FIGS. **1**C and **1**D show the results of methylene blue staining revealed the presence of abundant cystine crystals in the kidney of the $Ctns^{-/-}$ mice (FIG. **1**C) and very few in the treated $Ctns^{-/-}$ mice (FIG. **1**D).

[0018] FIGS. **2**A and **2**B are pictorial and graphical diagrams showing cystine crystals in the cornea. FIG. **2**A shows lateral cornea IVCM representations of $Ctns^{-/-}$ controls and LOW and HIGH HSC-transplanted mice. FIG. **2**B shows surface crystal quantification within each layer of the full IVCM cornea scans from both eyes of $Ctns^{-/-}$ controls and transplanted (LOW and HIGH) mice. Error bars represent SEM (* $p<0.05$, ** $p<0.005$).

[0019] FIG. **3** is a graphical diagram showing the results of a Thyroid study. Measure of cystine content (Left panel) and TSH level (Right panel) in $Ctns^{-/-}$ mice compared to wild-type mice (WT) and $Ctns^{-/-}$ mice transplanted with Ctns-expressing HSCs (grafted $Ctns^{-/-}$).

[0020] FIG. **4** is a pictorial and graphical diagram showing the impact of HSC transplantation on gastrointestinal tract in $Ctns^{-/-}$ mice. Left panel: Representative confocal picture of the colon: abundant GFP-expressing HSC-derived cells can be seen. Right panel: Cystine content in colon and intestine in HSC-transplanted $Ctns^{-/-}$ mice compared to controls. *$p<0.05$.

[0021] FIGS. **5**A-**5**D are graphical and pictorial diagrams showing TNT-mediated transfer of cystinosin is the preferred mode of cross-correction. FIGS. **5**A and **5**B show histograms representing percent decrease in cystine content in DsRed-$Ctns^{-/-}$ fibroblasts (recipient cells) when plated together with contact co-culture assays (FIG. **5**A) or separated by 1-μm port transwell filters from transwell assays (FIG. **5**B) either GFP-MSCs or GFP-macrophages (donor cells) (N=4 replicates for each). Values are means±standard deviations. *$p<0.05$; **$p<0.01$; *** $p<0.005$. FIG. **5**C shows a confocal image of TNTs (arrowheads) extended from GFP-macrophages to DsRed-$Ctns^{-/-}$ fibroblasts. FIG. **5**D shows representative frames from a confocal movie showing migration of cystinosin-GFP-contacting vesicles via TNTs from a CTNS-GFP-expressing macrophage towards $Ctns^{-/-}$ fibroblasts (arrowheads). Bars: (FIG. **5**C) 30 μm; (FIG. **5**D) 20 μm.

[0022] FIGS. **6**A-**6**C are pictorial diagrams showing TNT-mediated transfer in vivo, study of the kidney. FIG. **6**A shows confocal images of kidney from 8-month-old $Ctns^{-/-}$ mice at 6 months post-transplantation with GFPWT HSPCs. GFP is in green and laminin in red. PTCs (lumen, #) was labeled by Lotus Tetragonobus-lectin (LT) (blue). FIGS. **6**A-a**1**, **6**A-a**2**, and **6**A a**3** show that eGFP-expressing HSC-derived cells display numerous extensions. Arrowheads indicate TBL crossing. Apoptotic PTC (*). FIG. **6**A-a**3** shows GFP-expressing green structures are located within PTCs. FIGS. **6**B-**6**D show Z-stack confocal images of kidneys obtained from $Ctns^{-/-}$ mice transplanted with DsRed-$Ctns^{-/-}$ HSPCs (control, FIG. **6**B) or DsRed-$Ctns^{-/-}$ HPCs lentivirally transduced to express cystinosin-GFP and stained for phalloidin (FIG. **6**C). Cystinosin-GFP-containing vesicles are abundant in the cytoplasm of PTCs (FIG. **6**C). FIGS. **6**B and **6**C show nuclei that are stained in blue (DAPI). Scale bars: μm (FIG. **6**A), 10 μm (FIGS. **6**B and **6**C).

[0023] FIG. **7** is a pictorial diagram showing pCCL-CTNS lentivirus vector structure. SIN-LTR=Self-inactivating long terminal repeat; Ψ=Psi sequence; RRE=rev responsive element; cPPT=central polypurine tract; EFS=elongation factor 1α short; CTNS cDNA=human CTNS cDNA; WPRE=woodchuck hepatitis post-transcriptional regulatory element.

[0024] FIGS. **8**A-**8**D are graphical and pictorial diagrams showing cystine and cystine crystal quantification in male kidney. FIG. **8**A shows cystine content in non-treated $Ctns^{-/-}$ mice (KO) compared with treated with pCCL-CTNS-HSCs. FIG. **8**B shows quantification of cystine crys-

tals on kidney sections stained with methylene blue. Abundant cystine crystals were observed in kidney sections from nontreated $Ctns^{-/-}$ mice (FIG. **8**C) in contrast to pCCLCTNS-treated mice (FIG. **8**D). Error bars are defined as Mean+SD, *P<0.05.

[0025] FIGS. **9**A-**9**B are graphical diagrams showing the results from in vivo toxicology studies. FIGS. **9**A-**1** and **9**A-**2** show body weight of $Ctns^{-/-}$ males (FIG. **9**A-**1**) and females treated (FIG. **9**A-**2**) with pCCL-CTNS-transduced HSCs and mock treated. FIG. **9**B shows cystine content in tissues of $Ctns^{-/-}$ mice treated with pCCL-CTNS-transduced HSCs and mock treated.

[0026] FIGS. **10**A-**10**E are pictorial diagrams showing LAMP2 expression in heart and skeletal muscle of WT-HSPC-transplant recipients. FIGS. **10**A-**10**C are images showing LAMP2 expression in the hearts of WT (FIG. **10**A), KO (FIG. **10**B), and WT-HSPC transplanted showing LAMP2 expressing vesicles in cardiomyocytes adjacent to WT-GFP+ macrophages (FIG. **10**C). Arrows demonstrate RFP+ vacuoles. Western blots of heart (Figure and skeletal muscle (FIG. **10**E) lysates show near WT-level restoration of LAMP2 expression in mice recipient of WT-HSPC transplant.

[0027] FIG. **11** is a graphical diagram showing the results of physiological assessment. Grip strength is rescued in mice recipients of WT-HSPC compared to KO (untreated) and KO-HSPC recipient mice. *p<0.05 vs. WT; #p<0.05 vs. WT-HSPC group.

[0028] FIGS. **12**A-**12**D are pictorial and graphical diagrams showing rescue of increased autophagic flux following WT BMT. FIG. **12**A shows representative EM images of the hearts of WT, KO, WT-HSPC transplanted, and KO-HSPC transplanted mice. FIG. **12**B shows quantification of EM images demonstrating rescue of the accumulation of AVs in WT-HSPC mice to near WT levels. FIGS. **12**C and **12**D show Western blots and results demonstrating decreased LC-II/GAPDH levels in WT-HSPC transplanted mice vs. KO mice. *p<0.05 vs. WT; #p<0.05 vs. KO.

DETAILED DESCRIPTION OF THE INVENTION

[0029] The present invention is based, in part, on the finding that a self-inactivating (SIN)-lentivirus vector containing the encoding human cystinosin (CTNS) or LAMP-2 cDNA and a functional promoter can be used to ex vivo gene-corrected 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. These cells mobilize and integrate into the disease tissues, brain, muscle, heart, leading to their rescue. While autologous HSPCs are used in the illustrative examples herein, one of skill in the art would recognize that other HSPCs would be useful as well (e.g., allogeneic).

[0030] Before the present compositions and methods are described, it is to be understood that this invention is not limited to particular compositions, methods, and experimental conditions described, as such compositions, methods, and conditions may vary. It is also to be understood that the terminology used herein is for purposes of describing particular embodiments only, and is not intended to be limiting, since the scope of the present invention will be limited only in the appended claims.

[0031] As used in this specification and the appended claims, the singular forms "a", "an", and "the" include plural references unless the context clearly dictates otherwise. Thus, for example, references to "the method" includes one or more methods, and/or steps of the type described herein which will become apparent to those persons skilled in the art upon reading this disclosure and so forth.

[0032] The term "comprising," which is used interchangeably with "including," "containing," or "characterized by," is inclusive or open-ended language and does not exclude additional, unrecited elements or method steps. The phrase "consisting of" excludes any element, step, or ingredient not specified in the claim. The phrase "consisting essentially of" limits the scope of a claim to the specified materials or steps and those that do not materially affect the basic and novel characteristics of the claimed invention. The present disclosure contemplates embodiments of the invention compositions and methods corresponding to the scope of each of these phrases. Thus, a composition or method comprising recited elements or steps contemplates particular embodiments in which the composition or method consists essentially of or consists of those elements or steps.

[0033] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the invention, the preferred methods and materials are now described.

[0034] The term "subject" or "host organism," as used herein, refers to any individual or patient to which the subject methods are performed. Generally the subject is human, although as will be appreciated by those in the art, the subject may be an animal. Thus other animals, including mammals such as rodents (including mice, rats, hamsters and guinea pigs), cats, dogs, rabbits, farm animals including cows, horses, goats, sheep, pigs, etc., and primates (including monkeys, chimpanzees, orangutans and gorillas) are included within the definition of subject.

[0035] The term "biological sample," refers to any sample taken from a participant, including but not limited to cells, blood, tissue, skin, urine, etc., or hair.

[0036] The term "buccal mucosa," refers to the inside lining of the cheeks and floor of the mouth.

[0037] The term "therapeutically effective amount" or "effective amount" means the amount of a compound or pharmaceutical composition that elicits the biological or medical response of a tissue, system, animal or human that is being sought by the researcher, veterinarian, medical doctor or other clinician. Thus, the term "therapeutically effective amount" is used herein to denote any amount of a formulation that causes a substantial improvement in a disease condition when applied to the affected areas repeatedly over a period of time. The amount varies with the condition being treated, the stage of advancement of the condition, and the type and concentration of formulation applied. Appropriate amounts in any given instance will be readily apparent to those skilled in the art or capable of determination by routine experimentation. In the context of cystinosin, an example of a therapeutically effective amount of an agent, such as a population of hematopoietic stem cells transduced, gene-edited, or otherwise modified to express a human cystinosin transgene, is an amount sufficient to reduce the quantity of cystine (e.g., crystalline cystine) in the

lysosomes of a cell in the patient, such as a cell in the kidney, liver, lung, spleen, muscle, brain, and/or heart.

[0038] A "dosage" or "dose" are defined to include a specified size, frequency, or exposure level are included within the definition.

[0039] A "therapeutic effect," as used herein, encompasses a therapeutic benefit and/or a prophylactic benefit as described herein.

[0040] The terms "administration" or "administering" are defined to include an act of providing a compound or pharmaceutical composition of the invention to a subject in need of treatment. The phrases "parenteral administration" and "administered parenterally" as used herein means modes of administration other than enteral and topical administration, usually orally or by injection, and includes, without limitation, intravenous, intramuscular, intraarterial, intrathecal, intracapsular, intraorbital, intracardiac, intradermal, intraperitoneal, transtracheal, subcutaneous, subcuticular, intra-articulare, subcapsular, subarachnoid, intraspinal and infrasternal injection and infusion. The phrases "systemic administration," "administered systemically," "peripheral administration" and "administered peripherally" as used herein mean the administration of a compound, drug or other material other than directly into the central nervous system, such that it enters the subject's system and, thus, is subject to metabolism and other like processes, for example, subcutaneous administration.

[0041] If a viral vector specific for the cell type is not available, the vector can be modified to express a receptor (or ligand) specific for a ligand (or receptor) expressed on the target cell, or can be encapsulated within a liposome, which also can be modified to include such a ligand (or receptor). A peptide agent can be introduced into a cell by various methods, including, for example, by engineering the peptide to contain a protein transduction domain such as the human immunodeficiency virus TAT protein transduction domain, which can facilitate translocation of the peptide into the cell. In addition, there are a variety of biomaterial-based technologies such as nano-cages and pharmacological delivery wafers (such as used in brain cancer chemotherapeutics) which may also be modified to accommodate this technology.

[0042] The viral vectors most commonly assessed for gene transfer are based on DNA-based adenoviruses (Ads) and adeno-associated viruses (AAVs) and RNA-based retroviruses and lentiviruses. Lentivirus vectors have been most commonly used to achieve chromosomal integration.

[0043] The term "parenchymal," refers to the functional parts of an organ, which sometimes includes structural parts of the same and/or adjacent organ.

[0044] As used herein, the terms "reduce" and "inhibit" are used together because it is recognized that, in some cases, a decrease can be reduced below the level of detection of a particular assay. As such, it may not always be clear whether the expression level or activity is "reduced" below a level of detection of an assay, or is completely "inhibited." Nevertheless, it will be clearly determinable, following a treatment according to the present methods.

[0045] As used herein, "treatment" or "treating" means to administer a composition to a subject or a system with an undesired condition. The condition can include a disease or disorder. "Prevention" or "preventing" means to administer a composition to a subject or a system at risk for the condition. The condition can include a predisposition to a disease or disorder. The effect of the administration of the composition to the subject (either treating and/or preventing) can be, but is not limited to, the cessation of one or more symptoms of the condition, a reduction or prevention of one or more symptoms of the condition, a reduction in the severity of the condition, the complete ablation of the condition, a stabilization or delay of the development or progression of a particular event or characteristic, or minimization of the chances that a particular event or characteristic will occur.

[0046] As used herein, the term "genetic modification" is used to refer to any manipulation of an organism's genetic material in a way that does not occur under natural conditions. Methods of performing such manipulations are known to those of ordinary skill in the art and include, but are not limited to, techniques that make use of vectors for transforming cells with a nucleic acid sequence of interest. Included in the definition are various forms of gene editing in which DNA is inserted, deleted or replaced in the genome of a living organism using engineered nucleases, or "molecular scissors." These nucleases create site-specific double-strand breaks (DSBs) at desired locations in the genome. The induced double-strand breaks are repaired through nonhomologous end-joining (NHEJ) or homologous recombination (HR), resulting in targeted mutations (i.e., edits).

[0047] There are several families of engineered nucleases used in gene editing, for example, but not limited to, meganucleases, zinc finger nucleases (ZFNs), transcription activator-like effector-based nucleases (TALEN), the CRISPR-Cas system, and ARCUS. However, it should be understood that any known gene editing system utilizing engineered nucleases may be used in the methods described herein.

[0048] CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) is an acronym for DNA loci that contain multiple, short, direct repetitions of base sequences. The prokaryotic CRISPR/Cas system has been adapted for use as gene editing (silencing, enhancing or changing specific genes) for use in eukaryotes (see, for example, Cong, Science, 15:339(6121):819-823 (2013) and Jinek, et al., Science, 337(6096):816-21 (2012)). By transfecting a cell with elements including a Cas gene and specifically designed CRISPRs, nucleic acid sequences can be cut and modified at any desired location. Methods of preparing compositions for use in genome editing using the CRISPR/Cas systems are described in detail in US Pub. No. 2016/0340661, US Pub. No. 20160340662, US Pub. No. 2016/0354487, US Pub. No. 2016/0355796, US Pub. No. 20160355797, and WO 2014/018423, which are specifically incorporated by reference herein in their entireties.

[0049] Thus, as used herein, "CRISPR system" refers collectively to transcripts and other elements involved in the expression of or directing the activity of CRISPR-associated ("Cas") genes, including sequences encoding a Cas gene, a tracr (trans-activating CRISPR) sequence (e.g., tracrRNA or an active partial tracrRNA), a tracr-mate sequence (encompassing a "direct repeat" and a tracrRNA-processed partial direct repeat in the context of an endogenous CRISPR system), a guide sequence (also referred to as a "spacer", "guide RNA" or "gRNA" in the context of an endogenous CRISPR system), or other sequences and transcripts from a CRISPR locus. One or more tracr mate sequences operably linked to a guide sequence (e.g., direct repeat-spacer-direct

repeat) can also be referred to as "pre-crRNA" (pre-CRISPR RNA) before processing or crRNA after processing by a nuclease.

[0050] In some embodiments, a tracrRNA and crRNA are linked and form a chimeric crRNA-tracrRNA hybrid where a mature crRNA is fused to a partial tracrRNA via a synthetic stem loop to mimic the natural crRNA:tracrRNA duplex as described in Cong, Science, (2013) and Jinek, et al., Science, 337(6096):816-21 (2012)). A single fused crRNA-tracrRNA construct can also be referred to as a guide RNA or gRNA (or single-guide RNA (sgRNA)). Within an sgRNA, the crRNA portion can be identified as the 'target sequence' and the tracrRNA is often referred to as the 'scaffold'.

[0051] There are many resources available for helping practitioners determine suitable target sites once a desired DNA target sequence is identified. For example, numerous public resources, including a bioinformatically generated list of about 190,000 potential sgRNAs, targeting more than 40% of human exons, are available to aid practitioners in selecting target sites and designing the associate sgRNA to affect a nick or double strand break at the site. See also, crispr.u-psud.fr, a tool designed to help scientists find CRISPR targeting sites in a wide range of species and generate the appropriate crRNA sequences.

[0052] In some embodiments, one or more vectors driving expression of one or more elements of a CRISPR system are introduced into a target cell such that expression of the elements of the CRISPR system direct formation of a CRISPR complex at one or more target sites. While the specifics can be varied in different engineered CRISPR systems, the overall methodology is similar. A practitioner interested in using CRISPR technology to target a DNA sequence can insert a short DNA fragment containing the target sequence into a guide RNA expression plasmid. The sgRNA expression plasmid contains the target sequence (about 20 nucleotides), a form of the tracrRNA sequence (the scaffold) as well as a suitable promoter and necessary elements for proper processing in eukaryotic cells. Such vectors are commercially available (see, for example, Addgene). Many of the systems rely on custom, complementary oligos that are annealed to form a double stranded DNA and then cloned into the sgRNA expression plasmid. Co-expression of the sgRNA and the appropriate Cas enzyme from the same or separate plasmids in transfected cells results in a single or double strand break (depending on the activity of the Cas enzyme) at the desired target site.

[0053] Zinc-finger nucleases (ZFNs) are artificial restriction enzymes generated by fusing a zinc finger DNA-binding domain to a DNA-cleavage domain. Zinc finger domains can be engineered to target specific desired DNA sequences and this enables zinc-finger nucleases to target unique sequences within complex genomes. By taking advantage of endogenous DNA repair machinery, these reagents can be used to precisely alter the genomes of higher organisms. The most common cleavage domain is the Type IIS enzyme Fok1. Fok1 catalyzes double-stranded cleavage of DNA, at 9 nucleotides from its recognition site on one strand and 13 nucleotides from its recognition site on the other. See, for example, U.S. Pat. Nos. 5,356,802; and 5,487,994; as well as Li et al. *Proc., Natl. Acad. Sci. USA* 89 (1992):4275-4279; Li et al. *Proc. Natl. Acad. Sci. USA,* 90:2764-2768 (1993); Kim et al. *Proc. Natl. Acad. Sci. USA.* 91:883-887 (1994a); Kim et al. *J Biol. Chem.* 269:31, 978-31,982 (1994b), all of which are incorporated herein by reference. One or more of these enzymes (or enzymatically functional fragments thereof) can be used as a source of cleavage domains.

[0054] Transcription activator-like effector nucleases (TALENs) have an overall architecture similar to that of ZFNs, with the main difference being that the DNA-binding domain comes from TAL effector proteins, transcription factors from plant pathogenic bacteria. The DNA-binding domain of a TALEN is a tandem array of amino acid repeats, each about 34 residues long. The repeats are very similar to each other; typically they differ principally at two positions (amino acids 12 and 13, called the repeat variable diresidue, or RVD). Each RVD specifies preferential binding to one of the four possible nucleotides, meaning that each TALEN repeat binds to a single base pair, though the NN RVD is known to bind adenines in addition to guanine. TAL effector DNA binding is mechanistically less well understood than that of zinc-finger proteins, but their seemingly simpler code could prove very beneficial for engineered-nuclease design. TALENs also cleave as dimers, have relatively long target sequences (the shortest reported so far binds 13 nucleotides per monomer) and appear to have less stringent requirements than ZFNs for the length of the spacer between binding sites. Monomeric and dimeric TALENs can include more than 10, more than 14, more than 20, or more than 24 repeats. Methods of engineering TAL to bind to specific nucleic acids are described in Cermak, et al, *Nucl. Acids Res.* 1-11 (2011); US Published Application No. 2011/0145940, which discloses TAL effectors and methods of using them to modify DNA; Miller et al. *Nature Biotechnol* 29: 143 (2011) reported making TALENs for site-specific nuclease architecture by linking TAL truncation variants to the catalytic domain of Fok1 nuclease. The resulting TALENs were shown to induce gene modification in immortalized human cells. General design principles for TALE binding domains can be found in, for example, WO 2011/072246. Each of the foregoing references are incorporated herein by reference in their entireties.

[0055] The nuclease activity of the genome editing systems described herein cleave target DNA to produce single or double strand breaks in the target DNA. Double strand breaks can be repaired by the cell in one of two ways: non-homologous end joining, and homology-directed repair. In non-homologous end joining (NHEJ), the double-strand breaks are repaired by direct ligation of the break ends to one another. As such, no new nucleic acid material is inserted into the site, although some nucleic acid material may be lost, resulting in a deletion. In homology-directed repair, a donor polynucleotide with homology to the cleaved target DNA sequence is used as a template for repair of the cleaved target DNA sequence, resulting in the transfer of genetic information from a donor polynucleotide to the target DNA. As such, new nucleic acid material can be inserted/copied into the site. Therefore, in some embodiments, the genome editing vector or composition optionally includes a donor polynucleotide. The modifications of the target DNA due to NHEJ and/or homology-directed repair can be used to induce gene correction, gene replacement, gene tagging, transgene insertion, nucleotide deletion, gene disruption, gene mutation, etc.

[0056] Accordingly, cleavage of DNA by the genome editing vector or composition can be used to delete nucleic acid material from a target DNA sequence by cleaving the

target DNA sequence and allowing the cell to repair the sequence in the absence of an exogenously provided donor polynucleotide. Alternatively, if the genome editing composition includes a donor polynucleotide sequence that includes at least a segment with homology to the target DNA sequence, the methods can be used to add, i.e., insert or replace, nucleic acid material to a target DNA sequence (e.g., to "knock in" a nucleic acid that encodes for a protein, an siRNA, an miRNA, etc.), to add a tag (e.g., 6×His (SEQ ID NO: 27), a fluorescent protein (e.g., a green fluorescent protein; a yellow fluorescent protein, etc.), hemagglutinin (HA), FLAG, etc.), to add a regulatory sequence to a gene (e.g., promoter, polyadenylation signal, internal ribosome entry sequence (IRES), 2A peptide, start codon, stop codon, splice signal, localization signal, etc.), to modify a nucleic acid sequence (e.g., introduce a mutation), and the like. As such, the compositions can be used to modify DNA in a site-specific, i.e., "targeted" way, for example gene knock-out, gene knock-in, gene editing, gene tagging, etc., as used in, for example, gene therapy.

[0057] ARCUS is a genome editing platform derived from a natural genome editing enzyme referred to as a "homing endonuclease." Homing endonucleases are site-specific DNA-cutting enzymes encoded in the genomes of many eukaryotic species that are able to precisely recognize long DNA sequences (12-40 base pairs). These non-destructive enzymes trigger gene conversion events that modify the genome in a very precise way, most frequently by the insertion of a new DNA sequence. Thus, the ARCUS genome editing platform relies upon engineered ARC nucleases, which are fully synthetic enzymes similar to a homing endonuclease, but with improved specificity to recognize a DNA sequence within any target gene.

[0058] The terms "polypeptide," "peptide," and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer.

[0059] The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, a-carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., a carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

[0060] Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

[0061] As used herein, a "regulatory gene" or "regulatory sequence" is a nucleic acid sequence that encodes products (e.g., transcription factors) that control the expression of other genes.

[0062] As used herein, a "protein coding sequence" or a sequence that encodes a particular protein or polypeptide, is a nucleic acid sequence that is transcribed into mRNA (in the case of DNA) and is translated (in the case of mRNA) into a polypeptide in vitro or in vivo when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' terminus (N-terminus) and a translation stop nonsense codon at the 3' terminus (C-terminus). A coding sequence can include, but is not limited to, cDNA from eukaryotic mRNA, genomic DNA sequences from eukaryotic DNA, and synthetic nucleic acids. A transcription termination sequence will usually be located 3' to the coding sequence.

[0063] As used herein, a "promoter" is defined as a regulatory DNA sequence generally located upstream of a gene that mediates the initiation of transcription by directing RNA polymerase to bind to DNA and initiating RNA synthesis. A promoter can be a constitutively active promoter (i.e., a promoter that is constitutively in an active/"ON" state), it may be an inducible promoter (i.e., a promoter whose state, active/"ON" or inactive/"OFF", is controlled by an external stimulus, e.g., the presence of a particular compound or protein), it may be a spatially restricted promoter (i.e., transcriptional control element, enhancer, etc.)(e.g., tissue specific promoter, cell type specific promoter, etc.), and it may be a temporally restricted promoter (i.e., the promoter is in the "ON" state or "OFF" state during specific stages of embryonic development or during specific stages of a biological process. Thus, in various embodiments, the promoter may be a stem cell-specific promoter that drives transgene expression. For example, constitutive promoters of different strengths can be used. Expression vectors and plasmids in accordance with the present invention may include one or more constitutive promoters, such as viral promoters or promoters from mammalian genes that are generally active in promoting transcription. 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

[0064] As used herein, the term "gene" means the deoxyribonucleotide sequences comprising the coding region of a structural gene. A "gene" may also include non-translated sequences located adjacent to the coding region on both the 5' and 3' ends such that the gene corresponds to the length of the full-length mRNA. The sequences which are located 5' of the coding region and which are present on the mRNA are referred to as 5' non-translated sequences. The sequences which are located 3' or downstream of the coding region and which are present on the mRNA are referred to as 3' non-translated sequences. The term "gene" encompasses

both cDNA and genomic forms of a gene. A genomic form or clone of a gene contains the coding region interrupted with non-coding sequences termed "introns" or "intervening regions" or "intervening sequences." Introns are segments of a gene which are transcribed into heterogenous nuclear RNA (hnRNA); introns may contain regulatory elements such as enhancers. Introns are removed or "spliced out" from the nuclear or primary transcript; introns therefore are absent in the messenger RNA (mRNA) transcript. The mRNA functions during translation to specify the sequence or order of amino acids in a nascent polypeptide.

[0065] As used herein, the terms "functionally linked" and "operably linked" are used interchangeably and refer to a functional relationship between two or more DNA segments, in particular gene sequences to be expressed and those sequences controlling their expression. For example, a promoter/enhancer sequence, including any combination of cis-acting transcriptional control elements is operably linked to a coding sequence if it stimulates or modulates the transcription of the coding sequence in an appropriate host cell or other expression system. Promoter regulatory sequences that are operably linked to the transcribed gene sequence are physically contiguous to the transcribed sequence.

[0066] "Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refer to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence.

[0067] As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

[0068] The term "antibody" as used herein refers to polyclonal and monoclonal antibodies and fragments thereof, and immunologic binding equivalents thereof. The term "antibody" refers to a homogeneous molecular entity, or a mixture such as a polyclonal serum product made up of a plurality of different molecular entities, and broadly encompasses naturally occurring forms of antibodies (for example, IgG, IgA, IgM, IgE) and recombinant antibodies such as single-chain antibodies, chimeric and humanized antibodies and multi-specific antibodies. The term "antibody" also refers to fragments and derivatives of all of the foregoing, and may further comprise any modified or derivatized variants thereof that retains the ability to specifically bind an epitope. Antibody derivatives may comprise a protein or chemical moiety conjugated to an antibody. A monoclonal antibody is capable of selectively binding to a target antigen or epitope. Antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, camelized antibodies, single chain antibodies (scFvs), Fab fragments, F(ab')2 fragments, disulfide-linked Fvs (sdFv) fragments, for example, as produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, intrabodies, nanobodies, synthetic antibodies, and epitope-binding fragments of any of the above.

[0069] As used herein, the term "humanized mouse" (Humouse) is a mouse developed to carry functioning human genes, cells, tissues, and/or organs. Humanized mice are commonly used as small animal models in biological and medical research for human therapeutics. Immunodeficient mice are often used as recipients for human cells or tissues, because they can relatively easily accept heterologous cells due to lack of host immunity.

[0070] HSCs possess the ability of multipotency (i.e., one HSC can differentiate into all functional blood cells) and self-renewal (i.e., HSCs can divide and give rise to an identical daughter cell, without differentiation). Through a series of lineage commitment steps, HSCs give rise to progeny that progressively lose self-renewal potential and successively become more and more restricted in their differentiation capacity, generating multi-potential and lineage-committed progenitor cells, and ultimately mature functional circulating blood cells.

[0071] The ability of hematopoietic stem and progenitor cells (HSPCs) to self-renew and differentiate is fundamental for the formation and maintenance of life-long hematopoiesis and deregulation of these processes may lead to severe clinical consequences. HSPCs are also highly valuable for their ability to reconstitute the hematopoietic system when transplanted and this has enabled their use in the clinic to treat a variety of disorders including bone marrow failure, myeloproliferative disorders and other acquired or genetic disorders that affect blood cells.

[0072] As used herein, a "pluripotent cell" refers to a cell derived from an embryo produced by activation of a cell containing DNA of all female or male origin that can be maintained in vitro for prolonged, theoretically indefinite period of time in an undifferentiated state that can give rise to different differentiated tissue types, i.e., ectoderm, mesoderm, and endoderm. "Embryonic stem cells" (ES cells) are pluripotent stem cells derived from the inner cell mass of a blastocyst, an early-stage preimplantation embryo.

[0073] As used herein, an "autologous transplant" refers to a transplant that uses a subject's own stem cells. These cells are collected in advance and returned at a later stage. Thus, an "allogeneic transplant" refers to a transplant where the donor and the recipient of the stem cells are different people.

Exemplary allogeneic cells include, but are not limited to, syngeneic cells, WIC-matched cells, etc.

[0074] As used herein "pharmaceutically acceptable carrier" encompasses any of the standard pharmaceutical carriers, such as a phosphate buffered saline solution, water and emulsions such as an oil/water or water/oil emulsion, and various types of wetting agents.

[0075] As used herein, a "lysosomal protein disorder" or "lysosomal protein disease" refers to any metabolic disorders that result from defects in lysosomal function. Also referred to as "lysosomal storage disorders", such diseases/disorders are typically caused by lysosomal dysfunction usually as a consequence of deficiency of a single enzyme required for the metabolism of lipids, glycoproteins (sugar containing proteins) or so-called mucopolysaccharides. Exemplary lysosomal storage disorders include, but are not limited to, cystinosis, Salla disease, infantile sialic acid storage disorder, Cobalamin F-type disease, neuronal ceroid lipofuscinosis (both late infantile and juvenile forms), malignant infantile osteopetrosis, mucolipidosis IV, mucopolysaccharidosis type IIIC (Sanfilippo syndrome C), Niemann-Pick Type C, and Danon disease (Ruivo, et al. *Biochimica et Biophysica Acta* 1793 (2009) 636-649, incorporated herein by reference).

[0076] For example, Cystinosis is an autosomal metabolic disease that belongs to the family of the lysosomal storage disorders. Cystinosis has a devastating impact on the affected individuals, primarily children and young adults, even with cysteamine treatment. The prevalence of cystinosis is 1:100,000 to 1:200,000. The gene involved in cystinosis is the gene CTNS that encodes for the 7-transmembrane lysosomal cystine transporter, cystinosin. The most severe and the most frequent form of cystinosis is the infantile form, also called nephropathic cystinosis. Children develop a renal Fanconi syndrome at 6-8 months of age, characterized by severe fluid and electrolyte disturbance, growth retardation and rickets. Progressive loss of glomerular function leads to renal failure; according to NAPRTCS' (North American Pediatric Renal Trials and Collaborative Studies), 1.4% of children on dialysis (2011 Annual Dialysis Report) and 2.1% with kidney transplants (2010 Annual Transplant Report) have cystinosis. Cystinosis as a clinical entity is also a progressive dysfunction of multiple organs caused by the accumulation of cystine in the lysosomes of all the cells in the body; affected patients store 50-100 times the normal amounts of cystine in their cells.

[0077] Cystine storage leads to the formation of cystine crystals in all tissues. The main clinical complications in cystinosis include to diabetes, hypothyroidism, myopathy and central nervous system deterioration. Corneal cystine crystals appear from the first decade of life resulting in photophobia and visual impairment. Swallowing difficulties, directly correlated with muscle atrophy, is a major cause of death in cystinosis. In addition to cystine builds up, cellular dysfunctions such as abnormal vesicular trafficking, autophagy and TFEB (Transcription Factor EB) signaling have also been described as responsible for the pathogenesis of cystinosis.

[0078] The current treatment for cystinosis is the drug cysteamine (mercaptoethylamine), which reduces the intracellular cystine content. However, this therapy only delays disease progression and has no effect on renal Fanconi syndrome nor does it prevent end stage renal failure in affected patients. Cysteamine has also been shown to be inefficient to improve cellular dysfunctions in CTNS-deficient cells, proving that cellular defects in cystinosis are not only due to cystine accumulation but also due to the lack of the cystinosin itself that interacts directly with key cellular components.

[0079] In addition, cysteamine has to be taken every 6 hours including at night, and results in bad body odor as well as severe gastrointestinal side effects such as vomiting and diarrhea that render treatment compliance difficult. In 2013, a delayed-release formulation of cysteamine (PROCYSBI®) was FDA-approved, which requires dosing every 12 hours. While PROCYSBI® reduces the number of doses improving the patients' quality of life, the impact on the disease is similar than immediate release cysteamine and patients still experience gastric side effects. Moreover, the cost of this medication is very high, $300,000-$600,000 per year per patient.

[0080] The ocular pathology in cystinosis requires topical administration of cysteamine eye drops every hour, which causes irritation and burning so compliance is very challenging. The cost of eye drops is about $50,000 per year per patient. Cysteamine and the supportive treatment for all the complications associated with cystinosis requires patients to take up to 60 pills per day; the kids often require placement of a gastric tube to be able to tolerate the medications and get essential caloric intake. Medical complications increase in severity and number with age resulting in new and ever-increasing symptoms and treatments. There are unending doctor appointments, G-tube feedings, frequent blood draws, growth hormones shots, bone pain, daily vomiting, eye pain and severe gastrointestinal side effects. As the disease progress, their bodies deteriorate. The most severe complications for adults are myopathy, pulmonary issues and progression of corneal cystinosis. Patients with renal failure require dialysis or transplantation, both of which have significant negative health effects and due to the severe shortage of donor organs, patients may wait three to six years for transplantation. Thus, the current standard of care does not prevent the progression of the disease and significantly impacts the quality of life for patients with cystinosis who still die in early adulthood.

[0081] Danon disease has many similarities to other lysosomal membrane protein diseases and is characterized as a disorder of autophagy that affects the degradation of many cellular components and thus does not result in the accumulation of a single substrate. Danon disease has been more recently described as an autophagic vacuolar myopathy. Danon disease is caused by mutations in the gene encoding lysosomal associated membrane protein 2 (LAMP-2), resulting in decreased expression of the LAMP-2 protein. Loss of LAMP-2 expression disrupts autophagic flux, impairing the ability of cells to respond to stress and remove damaged cellular components.

[0082] Thus, the present disclosure demonstrates that one-time hematopoietic stem and progenitor cell (HSPC) transplantation holds the potential to become a life-long curative therapy for a disease or disorder associated with a defective lysosomal transmembrane protein. The therapy may further prevent kidney transplantation and long-term complications associated with cystinosis including unexpectedly the clearance of the corneal cystine crystals. This should also allow patients to withdraw from oral cysteamine, cysteamine eye drops and any other medications used for treating symptoms

associated with the disease. As such, the quality of life of the patients is greatly improved and the cost of treatment highly decreased.

[0083] Due to the multi-systemic nature of cystinosis and all the drugs necessary to compensate for the absence of the protein, cystinosin, in every tissues, a gene therapy approach was investigated. Gene therapy has the potential to become an important new approach for the third millennium to treat both rare and common severe diseases because its reach extends well beyond that of conventional drugs and offers the prospect of a curative stem cell-based therapy with limited risks as compared to allogeneic HSC transplantation. Hematopoietic stem and progenitor cells (HSPCs) are therefore ideal candidates for use in regenerative medicine and cell replacement therapies because of their ease of isolation, self-renewal capacity, and safety. Moreover, gene therapy can address unmet medical need such as in the case of cystinosis, especially this strategy overcomes the unavailability of matched HSC donor and makes the treatment potentially available to all patients.

[0084] Using a rodent model of cystinosis, $Cins^{-/-}$ mice, it has been shown that transplantation of HSCs expressing a functional Ctns gene resulted in abundant tissue integration of bone marrow-derived cells, significant decrease of cystine accumulation (up to 97% clearance), and long-term kidney preservation. Indeed, while non-treated $Cins^{-/-}$ mice progressed to end-stage renal failure, age-matched $Cins^{-/-}$ mice transplanted with wild-type HSCs maintained normal renal function after more than a year post-transplant. Few to no cystine crystals were observed in the kidneys of treated mice, in contrast to non-treated $Cins^{-/-}$ mice, in which abundant cystine crystals were consistently observed in the kidney. It has also recently been demonstrated that HSC transplantation rescues eye defects in the $Cins^{-/-}$ mice. Treated $Cins^{-/-}$ mice exhibited almost complete resolution of cystine crystals from the epithelial layer to the middle stroma (100% to 72% reduction, respectively), and normal corneal thickness and intraocular pressure. The impact of transplanted HSCs on the thyroid gland has also been studied. $Cins^{-/-}$ mice present with sustained TSH activation combined with thyrocyte hypertrophy, hyperplasia and vascular proliferation. In contrast, $Cins^{-/-}$ mice treated with transplanted HSCs exhibited normalization of cystine and TSH values and normal histology. These studies are the first proof of concept that one single HSC transplantation could prevent the multi-organ degeneration associated with cystinosis

[0085] As such, the present disclosure evaluates the impact of HSPC transplantation in a mouse model for cystinosis ($Ctns^{-/-}$ mice). Using a mouse model of cystinosis ($Ctns^{-/-}$ mice), the present disclosure demonstrates that transplantation of wildtype (WT) murine hematopoietic stem cells (mHSCs) led to abundant tissue integration of bone marrow-derived cells, significant decrease of tissue cystine accumulation (up to 97% reduction) and long-term kidney, eye and thyroid preservation. Given the risks of mortality and morbidity associated with allogeneic HSC transplantation, such as graft-versus-host diseases (GVHD), an autologous transplantation protocol of HSCs was developed for ex vivo modification. Using a self-inactivated-lentiviral vector (SIN-LV) to introduce a functional version of the CTNS cDNA, pCCL-CTNS (backbone pCCL-EFS-X-WPRE), efficacy in $Ctns^{-/-}$ mice has been shown.

[0086] In vitro studies using human CD34+ HSPCs isolated from peripheral blood of healthy donors and cystinosis patients have now completed, and the serial transplantation in the Ctns−/− mice has been significantly advanced. Thus, the data provided herein demonstrates efficacy of transplantation of CD34+ HSCs from G-CSF mobilized peripheral blood stem cells (PB SC) of patients with cystinosis, modified by ex vivo transduction using the pCCL-CTNS LV.

[0087] Cystinosis and Danon disease both arise from loss of function mutations in transmembrane lysosomal proteins, Cystinosin and LAMP-2, respectively. In fact, Cystinosin is localized to LAMP-2 positive vesicles that are transferred during cross-correction. Thus, the present disclosure also demonstrates that bone marrow was harvested from patients with Danon disease and sorted for CD34+ hematopoeitic stems cells (HPSCs). After harvest, patient HPSCs are genetically modified using viral transduction vectors including, but not limited to, lentiviruses and other retroviruses carrying any normal variant of the LAMP-2 gene and/or any of the LAMP-2 splice isoforms (e.g., LAMP-2A, LAMP-2B, LAMP-2C), referred to hereafter collectively as "wild-type LAMP-2" or "WT LAMP-2", is inserted into the genome of the harvested HPSCs. After infection, the viral vector inserts the wild-type LAMP-2 transgene into the host cell genome at specific sites that limit genome disruption. This insertion allows the wild-type LAMP-2 transgene to then be stably expressed by the host cell. Following translation, the wild-type LAMP-2 protein is trafficked to the lysosomal membrane where it embeds and assumes its normal intracellular position. Introduction of the wild-type LAMP-2 protein into the lysosomal membrane restores autophagic flux, allowing the cell to function normally.

[0088] After the wild-type LAMP-2 gene has been introduced, the HPSCs are transplanted back into the patient from which they were harvested. These cells then re-engraft in the patient's bone marrow and begin to produce progenitor cells. Some of these progenitor cells differentiate into monocytes carrying the wild-type LAMP-2 gene. Monocytes with the wild-type LAMP-2 gene enter the circulation and subsequently invade the peripheral tissues where they transform into tissue resident macrophages. These macrophages, through a variety of mechanisms including, but not limited to, the formation of tunneling nanotubes, vesicular release, and direct cell-cell adhesion they transfer their lysosomes, which carry membrane-bound wild-type LAMP-2 protein, to diseased peripheral cells. Wild-type LAMP-2 protein may also be transferred between macrophages and diseased peripheral cells in additional forms including, but not limited to, as free protein or bound to other proteins, membranes or organelles. The transfer of wild-type LAMP-2 containing lysosomes or wild-type LAMP-2 in other forms restores normal autophagic flux in the diseased cells resulting in partial or complete amelioration of the Danon phenotype.

[0089] Accordingly, in one aspect, the invention provides a method of treating a lysosomal transmembrane protein disease or disorder in a subject. The method includes introducing ex vivo a functional human transmembrane protein corresponding to the disorder to be treated into HSPCs of the subject, and thereafter transplanting the HSPCs into the subject, thereby treating the lysosomal transmembrane protein disease or disorder. Thus, for example, when the disease or disorder to be treated is cystinosis, the functional human transmembrane protein to be introduced is CTNS. In various embodiments, the vector is a self-inactivating (SIN)-lenti-

virus vector, such as, for example, pCCL-CTNS (in the case of CTNS). Likewise, when the disease or disorder to be treated is Danon disease, the functional human transmembrane protein to be introduced is LAMP-2. In various embodiments, the step of introducing may include contacting a vector comprising a polynucleotide encoding the functional protein (e.g., CTNS or LAMP-2) and a functional promoter (e.g., an ubiquitous or endogenous promoter of the functional protein) with the HSPCs and allowing expression of the functional protein. As such, the present disclosure provides a method for autologous transplantation of ex vivo gene modified HSPCs to introduce a functional protein associated with a specific lysosomal transmembrane protein disease or disorder.

[0090] In various embodiments, the lysosomal transmembrane protein diseases or disorders include, but are not limited to, cystinosis, Salla disease, infantile sialic acid storage disorder, Cobalamin F-type disease, neuronal ceroid lipofuscinosis (both late infantile and juvenile forms), malignant infantile osteopetrosis, mucolipidosis IV, mucopolysaccharidosis type IIIC (Sanfilippo syndrome C), Niemann-Pick Type C, and Danon disease. Without being bound by theory, in cystinosis and free sialic acid storage diseases, transporters for cystine and acidic monosaccharides, respectively, are blocked or retarded. A putative cobalamin transporter and a hybrid transporter/transferase of acetyl groups are defective in cobalamin F type disease and mucopolysaccharidosis type IIIC, respectively. In neurodegenerative forms of osteopetrosis, mutations of a proton/chloride exchanger impair the charge balance required for sustained proton pumping by the V-type ATPase, thus resulting in bone-resorption lacuna neutralization. However, the mechanism leading to lysosomal storage and neurodegeneration remains unclear. Mucolipidosis type IV is caused by mutations of a lysosomal cation channel named TRPML1; its gating properties are still poorly understood and the ion species linking this channel to lipid storage and membrane traffic defects is debated. Finally, the autophagy defect of Danon disease apparently arises from a role of LAMP2 in lysosome/autophagosome fusion, possibly secondary to a role in dynein-based centripetal motility. (Ruivo, et al. *Biochimica et Biophysica Acta* 1793 (2009) 636-649, incorporated herein by reference).

[0091] Table 1 sets forth the exemplary lysosomal transmembrane protein diseases or disorders to be treated with ex vivo introduction of corresponding functional human transmembrane proteins.

TABLE 1

Human disease/disorder	Causative gene, locus	Protein name (aliases)	Protein size, # of transmembrane helices (TM)
Cystinosis	CTNS, 17p13	Cystinosin	367 aa; 7 TM
Salla disease, infantile sialic acid storage disorder	SLC17A5, 6q14-q15	Sialin	495 aa; 12 TM
Cobalamin F-type disease	LMBRD1, 6q13	LMBD1 (probable lysosomal cobalamin transporter)	540 aa; 9 TM
Neuronal ceroid lipofuscinosis, late infantile variant	MFSD8, 4q28.1-q28.2	CLN7 (major facilitator superfamily domain-containing protein 8)	518 aa; 12 TM
Neuronal ceroid lipofuscinosis, juvenile form	CLN3, 16p12.1	CLN3 (Battenin)	438 aa; 6 TM
Malignant infantile osteopetrosis	CLCN7, 16p13 OSTM1, 6921	C1C-7 OSTM1	805 aa; 18 TM 338 aa; 1 TM
Mucolipidosis IV	MCOLNJ, 19p13.3-p13.2	TRPML1 (mucolipin-1, MLN1)	580 aa; 6 TM
Mucopolysaccharidosis type IIIC (Sanfilippo syndrome C)	HGSNAT, 8p11.1	HGSNAT (TMEM76)	663 aa; 11 TM
Niemann-Pick Type C	NPC1, 18q11-q12	NPC1	1278 aa; 11 TM
Danon disease	LAMP2, Xq24	LAMP2 (LAMPB, LGP110)	410 aa; 1 TM

[0092] Vectors derived from lentiviruses have supplanted y-retroviral vector for gene therapy due to their superior gene transfer efficiency and better biosafety profile. Indeed, all cases of leukemogenic complications observed to date in clinical trials or animal models involved the use of retroviral vectors with LTR containing strong enhancer/promoters that can trigger distant enhancer activation. In contrast, the third generation of lentivirus vectors, SIN-LV, with the deletions in their LTR, contains only one internal enhancer/promoter, which reduces the incidence of interactions with nearby cellular genes, and thus, decreases the risk of oncogenic integration. SIN-LV are also designed to prevent the possibility of developing replication competent lentivirus (RCL) during production of viral supernatants with three packaging plasmids necessary for production. Lentivirus vectors efficiently transduce HSPCs and do not alter their repopulation properties, which make this type of vector an attractive vehicle for stem cell gene therapy.

[0093] Clinical trials using SIN-LV to gene-correct human HSPCs are being undertaken in the U.S. and Europe for several conditions including HIV-1, β-thalassemia, immune deficiencies, metabolic diseases and cancers. For immune deficiency disorders, 35 patients have been transplanted with SIN-LV-modified HSPCs so far. A clinical trial in patients with Adrenoleukodystrophy (ALD) has achieved stable gene

correction in ~20% of hematopoietic cells in two patients. Cerebral demyelination was arrested without further progression over three years of follow-up, which represents a clinical outcome comparable to that observed after allogeneic transplantation; there was no evidence of clonal dominance. Recently, a clinical trial for Wilskott-Aldrich syndrome was reported in three patients 32 months post-transplantation. Stable and long-term engraftment of the gene modified HSPCs (25-50%) resulted in improved platelet counts, protection from bleeding and infections, and resolution of eczema. Another clinical success was recently reported in three pre-symptomatic patients with Metachromatic Leukodystrophy. Transduced cell-derived blood cell engraftment achieved 45 to 80%, and up to 24 months later, protein activity was reconstituted to above normal values in cerebrospinal fluid associated with a clear therapeutic benefit.

[0094] The recent gene therapy successes using AAV vectors in the MCK mice not only prevented heart failure when given to pre-symptomatic animals, but also reversed the cardiomyopathy when given after the onset. While encouraging, this approach presents potential safety and logistic concerns: i) localized delivery by direct viral injection to affected sites poses certain challenges in accessing sites such as heart and brain and leads only to tissue-specific rescue, ii) systemic AAV delivery remains difficult in humans due to the high levels of vector necessary, leading to vector synthesis and safety concerns. In contrast, HSPC gene therapy approach has the key advantages: i) it treats all the complications by a single infusion of stem cells, ii) gene-correction occurs ex vivo in a controlled environment allowing cell characterization prior to transplantation, iii) gene-corrected HSPCs reside in the bone marrow niche after transplantation where they self-renew and become a reservoir of healthy cells for the lifespan of the patients, iv) it avoids immune reaction as compared to allogeneic transplantation. Thus, autologous HSPC gene therapy could provide a cure for lysosomal transmembrane protein diseases or disorders.

[0095] Amino acid and nucleic acid sequences for the human proteins set forth in Table 1 are known in the art. See, for example, GenBank Accession No.: Y15924.1, human CTNS gene, exon 3, flanking intronic regions and joined CDS, which provides the amino acid sequence (SEQ ID NO: 1):

MIRNWLTIFILFPLKLVEKCESSVSLTVPPVVKLENGSSTNVSLTLRPP

LNATLVITFEITFRSKNITILELPDEVVVPPGVTNSSFQVTSQNVGQLT

VYLHGNHSNQTGPRIRFLVIRSSAISIINQVIGWIYFVAWSISFYPQVI

MNWRRKSVIGLSFDFVALNLTGFVAYSVFNIGLLWVPYIKEQFLLKYPN

GVNPVNSNDVFFSLHAVVLTLIIIVQCCLYERGGQRVSWPAIGFLVLAW

LFAFVTMIVAAVGVITWLQFLFCFSYIKLAVTLVKYFPQAYMNFYYKST

EGWSIGNVLLDFTGGSFSLLQMFLQSYNNDQWTLIFGDPTKFGLGVFSI

VFDVVFFIQHFCLYRKRPGYDQLN

GenBank Accession No.: AJ222967.1, human CTNS mRNA, which provides the nucleic acid sequence (SEQ ID NO: 2):

cgcctctccc aaagtctagc cgggcagggg aacgcggtgc attcctgacc ggcacctggc gaggctcatg cgtcccgtga

gggcggttcc tcgagcctgg gggcgctcag attgctttgg agacgctgag agaacctttg cgagagcgcc ggttgacgtg

cggagtgcgg ggctccgggg gactgagcag cacgagaccc catcctcccc tccgggtttt cacactgggc gaagggagga

ctcctgagct ctgcctcttc cagtaacatt gaggattact gtgttttgtg agagctcgct aggcgcccta agcaacagag ttctgagaaa

tcgagaaaca tgataaggaa ttggctgact atttttatcc tttttccccct gaagctcgta gagaaatgtg agtcaagcgt cagcctcact

gttcctcctg tcgtaaagct ggagaacggc agctcgacca acgtcagcct caccctgcgg ccaccattaa atgcaaccct

ggtgatcact tttgaaatca catttcgttc caaaaatatt actatccttg agctccccga tgaagttgtg gtgcctcctg gagtgacaaa

ctcctctttt caagtgacat ctcaaaatgt tggacaactt actgtttatc tacatggaaa tcactccaat cagaccggcc cgaggatacg

ctttcttgtg atccgcagca gcgccattag catcataaac caggtgattg gctggatcta ctttgtggcc tggtccatct ccttctaccc

tcaggtgatc atgaattgga ggcggaaaag tgtcattggt ctgagcttcg acttcgtggc tctgaacctg acaggcttcg

tggcctacag tgtattcaac atcggcctcc tctgggtgcc ctacatcaag gagcagtttc tcctcaaata ccccaacgga

gtgaaccccg tgaacagcaa cgacgtcttc ttcagcctgc acgcggttgt cctcacgctg atcatcatcg tgcagtgctg

cctgtatgag cgcggtggcc agcgcgtgtc ctggcctgcc atcggcttcc tggtgctcgc gtggctcttc gcatttgtca ccatgatcgt

ggctgcagtg ggagtgatca cgtggctgca gtttctcttc tgcttctcct acatcaagct cgcagtcacg ctggtcaagt attttccaca

ggcctacatg aacttttact acaaaagcac tgagggctgg agcattggca acgtgctcct ggacttcacc gggggcagct

tcagcctcct gcagatgttc ctccagtcct acaacaacga ccagtggacg ctgatcttcg gagacccaac caagtttgga

ctcggggtct tctccatcgt cttcgacgtc gtcttcttca tccagcactt ctgtttgtac agaaagagac cggggtatga ccagctgaac

tagcacccag ggacccagtg tacccagcct ctggcctcgt gccctgctgg ggaaggcctc acccagcgaa ggccggagaa

gcggttgggc cctggcacac agggctggct cagtgtgcgg acagaggaga ccactctgct cctggggcca gaggccattc

-continued

aatagcctgc cttcgtccgg gcccctcctg ggcctccccg gccaggcacg tggcaccgtc gccttgacac cgccatctct

tttctttaag gcttcaggca gcgcgcacag gctctggcag ccgtctcagg caggactggg caccaagctt gcagccgaag

gccttgcccc aaactaccag cgtttctgca agcagcttga agggctgacc ttgcagccgg gtgagccaag ggcactttgc

tgccaccgct gcattcccag agatcaagca gcccggtgcc gtggccagtg aactcagagg tgctggtgga cgggctagga

ctttggggtt aggccatggg gctctttctc tgaaggccac tttcctgacg tactctctgt acataactca gcgtccgtga ctgcagtaac

agccagccct acccagagta tttctgagcc atgaggggcc caccagattg gttctgaatt ggattcatgc ccagcgcatt

agcatagtaa ctcctttcag atttttgga gggacgtttg gaagtggctt actctcttct gccctctctc ctacctccac cttctcagat

gagccccatc tgagcacatc cagctgctcc ttacccagca tctggagtac aggacatagc tctctcctgc taccagtctg

tgccttagag gtcgttaggc ctgccaaacg gcgaccagct cccctggagc gagggcaggc cccttccctc tctttcccca

gacacctact tgagactcac caatttctgg cctgttcagg agcctcagat aagtatttgt acttgagacc acctcacaca atctgtatgg

gcccaaccct gatctcaaac ctccttccct ctgcccaaag ctgtccttcc tatggcagga gggggggggg tcccaggacg

tgcctcatac atgacttgag cttgtcagtc cactgagttt ccttctacga gatcaacgcg aggggcctgt atcttgaatt aaagcctact

cgcttccttt c

GenBank Accession No.: CAB62540.1, human sialin, which provides the amino acid sequence (SEQ ID NO: 3):

MRSPVRDLAR NDGEESTDRT PLLPGAPRAE AAPVCCSARY NLAILAFFGF

FIVYALRVNL SVALVDMVDS NTTLEDNRTS KACPEHSAPI KVHHNQTGKK

YQWDAETQGW ILGSFFYGYI ITQIPGGYVA SKIGGKMLLG FGILGTAVLT

LFTPIAADLG VGPLIVLRAL EGLGEGVTFP AMHAMWSSWA PPLERSKLLS

ISYAGAQLGT VISLPLSGII CYYMNWTYVF YFFGTIGIFW FLLWIWLVSD TPQKHKRISH

YEKEYILSSL RNQLSSQKSV PWVPILKSLP LWAIWAHFS YNWTFYTLLT LLPTYMKEIL

RFNVQENGFL SSLPYLGSWL CMILSGQAAD NLRAKWNFST LCVRRIFSLI

GMIGPAVELV AAGFIGCDYS LAVAFLTIST TLGGFCSSGF SINHLDIAPS YAGILLGITN

TFATIPGMVG PVIAKSLTPD NTVGEWQTVF YIAAAINVFG AIFFTLFAKG

EVQNWALNDH HGHRH

GenBank Accession No .: AJ387747.1, human sialin mRNA, which provides the nucleic acid sequence (SEQ ID NO: 4):

cggctacttt gcgccaatcc tacgagaact cccagaactc cgcttcccta gtccaaccca agccagagtt gcccacacct

aagatggcgg cgggggggcgg agtcggcgcg gccgcctctg ggcgggaccg cggggactag acgtggccgc ggggcggtgt

catcgccccc gccccgcccg gtccagccag ctcggcccgg gggcttcggg ctgtcgggcc ggcgctccct tctctgccag

gtggcgagta cacctgctca cgtaggcgtc atgaggtctc cggttcgaga cctggcccgg aacgatggcg aggagagcac

ggaccgcacg cctcttctac cgggcgcccc acgggccgaa gccgctccag tgtgctgctc tgctcgttac aacttagcaa

ttttggcctt ttttggtttc ttcattgtgt tagaagataa gggtaagaag atcacacaga ttggcactgc cagagcacta ccccctcttg

ttcctctttc aatattttgg tatgaaaagg ccattttaaa tttattgaca tcttcattgc caaaatggaa attcctggta acactgggag

tcctcctggg gacccctgat gccattttct gacactgaag gccttctgta atcatgaaat ggacttctac ttttagggcc ataattctca

cacttagacc ctgaggggag tatcagaggc tcaatgtgta aaactttccc atgcattacg tagaacttcc taccaatggg ttcctggagg

tgtcctcacc gaaggactag aaagaagcaa tggaataatt tttcttttgt aatacattct atccctgcca ttattgccta cttatttagg

tttttcaact gctgctggct gcttttgctc catcacaaat aacactgttg ttacactatt gaaccaataa ttgtgtaagc atcactagtt

tttaggttca atactaaagg ggtactgata tcaaacttta tgtgcccagg agcatccaag gataaagccc ttcccagcac tgtgaatctg

aaggcgtgtc atgcagaaac atatgttgcc ctgttcactc gagagggtgt acttcttagc tgctactata ggatctggtt ttcatcatta

ctttgggcta cttatatgaa ctcttggtta ttatgtgttc tcattggctg ttctggattt acatttgcca gagaatggca cgccaaaggt

-continued

ataatcctgc attctatgtc gccagaataa catacctgcc gaatgagctg aacacctgtt gcatctctgt cagctgccaa cccagagcca

tgagtaggca agaggaatat agtgttgcgt cagagcattc tcaaggatgg agcaaaatag ccattgctgc tacatttcca atttcatatg

tgaattggac agttagtgac agaaatcagc tcgtagttgc ggagatccta tgtatgatcc gcagaatttt tgattattct agcatcaacc

ctattccagg aaccgtgttc gaagtacaaa ctctattaat ttttttaat taaaatgaac tgctagtcgg aaacagacct gttgttcact

ggagctgcca gcactccctc gtgtcgactc agagcagtga tggctggcat tagtggatat tgctcccata attctcggtt

gggggaaaat agatttagga gccatgcatg caggagcaca ttatgtcttc acaccacaaa tttcttcaca acacttttct aggttcaatg

tgtctggtca tagccttata ttggccgttg atctggatat aatggttggg tatattgctg actgggctct gtatttttat tgtacttgta

tgtgtttaat gcaacatgaa cctgatacct ttcctcataa tccactgtat cctggcttca ttcggctggt gatccactgc gcaacctgca

ggtagattca aatacaactt aaagttcatc ataatcaaac ccttttttta tggctacatc gctgctagga tttgggatcc gttggaccac

tcattgtact ccatgtggtc ttcttgggct gcttgggaca gtaatttctc tacttttttg gtactattgg aacacaagag aatttcccat

gaagtcagtg ccgtgggtac tacaactgga ctttttatac ttcaagagaa tgggtttta agctgctgac aatttaaggg ggaatgattg

gacctgcagt ctttcctaac tatatcaaca tgctccttcg tatgctggta cccgtcattg ctaaaagtct ctgctattaa tgtttttggt

caatgatcac catggacaca ttatcatgta acctcaaagt ttagattttt aaggcctata tatgaataat atgtaagcta gtaggacagt

tctgttgatt ttgcttaatt aaactagatg aaattgtcag ctctctctga aatttcgcct ggcaactgga gggtcagagt gcccagcgtt

gcctttcctc tgaggggcta tatggtcttg atacatcctc aaagaaaaat gc

GenBank Accession No.: CCP79466.1, human LMBD1, which provides the amino acid sequence (SEQ ID NO: 5):

GAASAELVIGWCIFGPLLLAIFAFCWIYVRKYQSQRESEVVSTITAIFSLAIALITSALLPV

DIFLVSYMKNQNGTFKDWANANVSRQIEDTVLYGYYTLYSVILFCVFFWIPFVYFYYEE

KDDDDTSKCTQIKTAFKYTLGFAVICALLLLVGAFVPLNVPNNKNSTEWEKVKFLFEEL

GSSHGLAALSFSISSLTLIGMLAAITYTAYGMSALPLNLIKGTRSAAYERLENTEDIEEVE

QHIQTIKSKSKDGRPLPARDKRALKQFEERLRTLRKRERHLEYIENSWWTKFCGALRPL

KIIWGIFFILVALLFIISLFLSNLDKALHSAGIDSGFIIFGANLSNPLNMLLPVLQTVFPLDYI

LITIIIMYFIFTSMAGIRNIGIWFFWVRLYKIRRGRTRPQALLFLCMILLLIVLHTSYMIYSL

APQYVMYGSQNYLIESNITYDDHKNNSAFPVPKRCDADAPEDQCTVTRTYLFLHKFWF

FSAAYYFGNWAFLVVFLIGLIVSCCKGKKSVIEGVDEDDSDISDDEPSVYSV

GenBank Accession No.: HAAF01007642.1, human LMBD1 transcribed RNA, which provides the nucleic acid sequence (SEQ ID NO: 6):

ggcgcggctt tctgttggat ttttttctctg aaaaatcaaa tatatggtta ctactatgaa ttgggatttg acaaaaattc tgcattgtca

ggcatgtctg ctgaagacat gccagcaagg cacttagaat ggggaatatt cggcggagct atatgttcgt gcgattgcac

atggtacatt ctacacctta gaaaaggatg ctgtaatttg tacagagtgg ttttctatta cattaccttt tgaagaagtg gataaacgcg

acattgaaaa tttcatctta ggtgatcggc aaataccaaa ttatcacatc taaggactgg tattctgtta atgatgatac tgcacttctt

gaaaaagtga gttctctgac aaatctaata gagcaacaca ccttaaaaca cagctggtgg gttgcattgc tggtgcatat

gtcagcggga agcacttctt gccaatgcta tattattctg tagtaaatgt cttttagttg agttcctgtt cttgattgga aaaggcacta

ttcaaacgat atttgaagaa acaaaatttt tgtttataat ttggcccctt aagtgaagtt ccagtggata atgtcagcag tgtgtttttc

actcaaatta gagcttttgt tgaagaactt atgttggcag gaagcgctgc taaatcaaaa aggttaagaa gtggtgctct ttctctcttc

actactggct gtctccacca tatttttggt acagatcgag tggatccctt aaactgcatt tcctctaaat ggaagtagtc ctataactta

ttacgaacgt agcaaagatg cacttaggaa gcgtcccctg ctgtcaaatt atttttgcat taacggcaat ttcttacatg gacactgtgt

ttgtctactt caagtatact gttcctaata atggtttagc cacagcctat ttagaaaaca gtcggccttt aagagagagg aagatcattt

tggataaagc

GenBank Accession No.: AAH295036.1, human CLN7, which provides the amino acid sequence (SEQ ID NO: 7):

-continued

MAGLRNESEQ EPLLGDTPGS REWDILETEE HYKSRWRSIR ILYLTMFLSS

VGFSVVMMSI WPYLQKIDPT ADTSFLGWVI ASYSLGQMVA SPIFGLWSNY

RPRKEPLIVS ILISVAANCL YAYLHIPASH NKYYMLVARG LLGIGAGNVA

VVRSYTAGAT SLQERTSSMA NISMCQALGF ILGPVFQTCF TFLGEKGVTW

DVIKLQINMY TTPVLLSAFL GILNIILILA ILREHRVDDS GRQCKSINFE EASTDEAQVP

QGNIDQVAVV AINVLFFVTL FIFALFETII TPLTMDMYAW TQEQAVLYNG

IILAALGVEA VVIFLGVKLL SKKIGERAIL LGGLIVVWVG FFILLPWGNQ

FPKIQWEDLH NNSIPNTTFG EIIIGLWKSP MEDDNERPTG CSIEQAWCLY TPVIHLAQFL

TSAVLIGLGY PVCNLMSYTL YSKILGPKPQ GVYMGWLTAS GSGARILGPM

FISQVYAHWG PRWAFSLVCG IIVLTITLLG VVYKRLIALS VRYGRIQE

GenBank Accession No.: BC029503.1, human CLN7 mRNA, which provides the nucleic acid sequence (SEQ ID NO: 8):

aggttacaag cagcagatcc caccttcagt cctggctctg acaagccctc cagcttcacg ccacccggga tgggagaaag

caggtgtcgc gagagttggg cgcaagacgc cttgtaggga gtgtaactat ggccggcctg cggaacgaaa gtgaacagga

gccgctctta ggcgacacac ctggaagcag agaatgggac attttagaga ctgaagagca ttataagagc cgatggagat

ctattaggat tttatatctt actatgtttc tcagcagtgt agggttttct gtagtgatga tgtccatatg gccatatctc caaaagattg

atccgacagc tgatacaagt tttttgggct gggttattgc ttcatatagt cttggccaaa tggtagcttc acctatattt ggtttatggt

ctaattatag accaagaaaa gagcctctta ttgtctccat cttgatttcc gtggcagcca actgcctcta tgcatatctc cacatcccag

cttctcataa taaatactac atgctggttg ctcgtggatt gttgggaatt ggagcaggaa atgtagcagt tgttagatca tatactgctg

gtgctacttc ccttcaggaa agaacaagtt ccatggcaaa cataagcatg tgtcaagcat taggttttat tctaggtcca gtttttcaga

cttgttttac attccttgga gaaaaaggtg tgacatggga tgtgattaaa ctgcagataa acatgtatac aacaccagtt ttacttagcg

ccttcctggg aattttaaat attattctga tccttgccat actaagagaa catcgtgtgg atgactcagg aagacagtgt aaaagtatta

attttgaaga agcaagtaca gatgaagctc aggttcccca aggaaatatt gaccaggttg ctgttgtggc catcaatgtt ctgttttttg

tgactctatt tatctttgcc ctttttgaaa ccatcattac tccattaaca atggatatgt atgcctggac tcaagaacaa gctgtgttat

ataatggcat aatacttgct gctcttgggg ttgaagccgt tgttattttc ttaggagtta agttgctttc caaaaagatt ggcgagcgtg

ctattctact gggaggactc atcgttgtat gggttggctt ctttatcttg ttaccttggg gaaatcaatt tcccaaaata cagtgggaag

atttgcacaa taattcaatc cctaatacca catttgggga aattattatt ggtctttgga agtctccaat ggaagatgac aatgaaagac

caactggttg ctcgattgaa caagcctggt gcctctacac cccggtgatt catctggccc agttccttac atcagctgtg ctaataggat

taggctatcc agtctgcaat cttatgtcct atactctata ttcaaaaatt ctaggaccaa aacctcaggg tgtatacatg ggctggttaa

cagcatctgg aagtggagcc cggattcttg ggcctatgtt catcagccaa gtgtatgctc actggggacc acgatgggca

ttcagcctgg tgtgtggaat aatagtgctc accatcaccc tcctgggagt ggtttacaaa agactcattg ctctttctgt aagatatggg

aggattcagg aataaactag ctaagactgt gatggaaact acttgctgtg tggcacttcc tggtctaaag ctctgctaga caattgcggt

gagccagtct ccaagaatca gactacagat attgcagatt ttgaagaaca agaacatatg ttgaataaca gagagaattc tacatgtcat

tgtgaatagt aggttatata aaaacatact agatgataatt tcaaaaaaa aaaaaaaaa

GenBank Accession No.: AAB51075.1, human CLN3, which provides the amino acid sequence (SEQ ID NO: 9):

MGGCAGSRRRFSDSEGEETVPEPRLPLLDHQGAHWKNAVGFWLLGLCNNFSYVVMLS

AAHDILSHKRTSGNQSHVDPGPTPIPHNSSSRFDCNSVSTAAVLLADILPTLVIKLLAPLG

LHLLPYSPRVLVSGICAAGSFVLVAFSHSVGTSLCGVVFASISSGLGEVTFLSLTAFYPRA

VISWWSSGTGGAGLLGALSYLGLTQAGLSPQQTLLSMLGIPALLLASYFLLLTSPEAQDP

GGEEEAESAARQPLIRTEAPESKPGSSSSLSLRERWTVFKGLLWYIVPLVVVYFAEYFIN

-continued

QGLFELLFFWNTSLSHAQQYRWYQMLYQAGVFASRSSLRCCRIRFTWALALLQCLNLV

FLLADVWFGFLPSIYLVFLIILYEGLLGGAAYVNTFHNIALETSDEHREFAMAATCISDTL

GISLSGLLALPLHDFLCQLS

GenBank Accession No.: U32680.1, human CLN3 mRNA complete cds, which provides the nucleic acid sequence (SEQ ID NO: 10):

ccccctagaca cctcccgtgg ggaggctgtg tccctctgtt caacttctct cagagccatg tctctacggc ccttcacctg gttgcctttt

gggaggtcac tgggggagct accctgctgt aggcccagga cccggagtcg ctgtggtaca tcctcttttt tggcgtcttt

cagtgcctca tcctgatcat ggagaccagt ctgtcggggc gacgcaggtc tgaactgtgc ggggccagcc tgtccatgg

agccggagct gagccccctt caggctcgcg ggaccatcag tatgtggtga tggacccagg tgctgtgctc ctgccctaca

ctcattctgt cttcctctcc gggctgctgg ccatgctggg ccctggaggg aagccaggct ttgttccctt ctggaacact gcctcccgct

acctggtgtt tctgtatgag gatgagcacc tcctggcttt acattcacct tcccagcctt attgtctccc gggaccggca tggacactct

gcggcgcttt ggcgcgcatt tgctgagtgc cccaacgccg ctggggaca gcccccgggt ggggaccagc ctcactgcct

gggccctgtc tatccctgcc gaagaagaag ccagctccag ggtcgtagtt tccctgagtc cttctctccg cctgctggca

gggctcctgg gggagtttgc gcctctgcat gtgggcagag cccggcaggt acttggggag atcgggcgtt atgaccctgg

tcggattccg ggaagaacgc cgcccacgac atcccccaca tcctccccac tctcgtcagt ctgtgtggtg tctaccccag

ctacctgggc ctgctgctgg cagagagcgc cctctccctt tactttgccg acgctcagca ctgctgtcgc gacgtgtggt

gaggcgcagc aatggcggcc gacttcctct ggacaggtca ctgggagtag tttcttcctg gatccttgtc accctcgggg

aggggggagga ggtgggcttc atccttagcc acagctcatc actcgtcatc gggatttgtg tggtcttcgc ggccgtgatc

ctcacccagg ccagctattt agcccggcag cgggaaaggt agtatttcat ataccgctgg atccgtttca tcggctttct ctacgtgaac

acctgcatct gccagctctc gacacccagg ggaagggctg gcatcatgcc acctgtcgca gacctgaact gaccgtcccg

tggctgctgg acaagaggac acgatttgac aaattgttgg ctgctggaag tagcatctca tcctggtggt ccggcctctc cttgttgctc

cccctcataa ggacagtatt taaccaggga taccagatgc cctgggccct gccaagcatc accttccaca ctgacacact

ctgatactcg cccaccccag aagccttgtt ttctgaataa gaccctcatc tgatgcgatg gagccccggc gcctttgcaa

atcgggaaac tgcaactctg ctcctcttgg cttcgtcctg taggccttg cctcagggac ccctcagcag acatctcctg

gaaccgaggc caagggtctg ctttttgaac tgtaccaggc ggccctgctg tacctcgtct acatcgccct ggggatctcc

ggatcctcag agaccctcca tccttgcagg atgccgattt

GenBank Accession No.: AAF34711.1, human CLCN7, which provides the amino acid sequence (SEQ ID NO: 11):

MANVSKKVSWSGRDRDDEEAAPLLRRTARPGGGTPLLNGAGPGAARQSPRSALFRVG

HMSSVELDDELLDPDMDPPHPFPKEIPHNEKLLSLKYESLDYDNSENQLFLEEERRINHT

AFRTVEIKRWVICALIGILTGLVACFIDIVVENLAGLKYRVIKGNIDKFTEKGGLSFSLLL

WATLNAAFVLVGSVIVAFIEPVAAGSGIPQIKCFLNGVKIPHVVRLKTLVIKVSGVILSVV

GGLAVGKEGPMIHSGSVIAAGISQGRSTSLKRDFKIFEYFRRDTEKRDFVSAGAAAGVSA

AFGAPVGGVLFSLEEGASFWNQFLTWRIFFASMISTFTLNFVLSIYHGNMWDLSSPGLIN

FGRFDSEKMAYTIHEIPVFIAMGVVGGVLGAVFNALNYWLTMFRIRYIHRPCLQVIEAV

LVAAVTATVAFVLIYSSRDCQPLQGGSMSYPLQLFCADGEYNSMAAAFFNTPEKSVVSL

FHDPPGSYNPLTLGLFTLVYFFLACWTYGLTVSAGVFIPSLLIGAAWGRLFGISLSYLTGA

AIWADPGKYALMGAAAQLGGIVRMTLSLTVIMMEATSNVTYGFPIMLVLMTAKIVGDV

FIEGLYDMHIQLQSVPFLHWEAPVTSHSLTAREVMSTPVTCLRRREKVGVIVDVLSDTAS

NHNGFPVVEHADDTQPARLQGLILRSQLIVLLKHKVFVERSNLGLVQRRLRLKDFRDAY

PRFPPIQSIHVSQDERECTMDLSEFMNPSPYTVPQEASLPRVFKLFRALGLRHLVVVDNR

-continued

NQVVGLVTRKDLARYRLGKRGLEELSLAQT

GenBank Accession No.: AF224741.1, human CLCN7 mRNA complete cds, which provides the nucleic acid sequence (SEQ ID NO: 12):

gccggcgctt cccggccggt gtcgctccgc ggcgggccat ggccaacgtc tctaagaagg tgtcctggtc cggccgggac

cgggacgacg aggaggcggc gccgctgctg cggaggacgg cgcggcccgg cgggggggacg ccgctgctga acggggctgg

gcccggggct gcgcgccagt caccacgttc tgcgcttttc cgagtcggac atatgagcag cgtggagctg gatgatgaac

ttttggaccc ggatatggac cctccacatc ccttccccaa ggagatccca cacaacgaga agctcctgtc cctcaagtat

gagagcttgg actatgacaa cagtgagaac cagctgttcc tggaggagga gcggcggatc aatcacacgg ccttccggac

ggtggagatc aagcgctggg tcatctgcgc cctcattggg atcctcacgg gcctcgtggc ctgcttcatt gacatcgtgg

tggaaaacct ggctggcctc aagtacaggg tcatcaaggg caatatcgac aagttcacag agaagggcgg actgtccttc

tccctgttgc tgtgggccac gctgaacgcc gccttcgtgc tcgtgggctc tgtgattgtg gctttcatag agccggtggc

tgctggcagc ggaatccccc agatcaagtg cttcctcaac ggggtgaaga tcccccacgt ggtgcggctc aagacgttgg

tgatcaaagt gtccggtgtg atcctgtccg tggtcggggg cctggccgtg ggaaaggaag ggccgatgat ccactcaggt

tcagtgattg ccgccgggat ctctcaggga aggtcaacgt cactgaaacg agatttcaag atcttcgagt acttccgcag

agacacagag aagcgggact tcgtctccgc aggggctgcg gccggagtgt cagcggcgtt tggagccccc gtgggtgggg

tcctgttcag cttggaggag ggtgcgtcct tctggaacca gttcctgacc tggaggatct tctttgcttc catgatctcc acgttcaccc

tgaattttgt tctgagcatt taccacggga acatgtggga cctgtccagc ccaggcctca tcaacttcgg aaggtttgac tcggagaaaa

tggcctacac gatccacgag atcccggtct tcatcgccat gggcgtggtg ggcggtgtgc ttggagctgt gttcaatgcc

ttgaactact ggctgaccat gtttcgaatc aggtacatcc accggccctg cctgcaggtg attgaggccg tgctggtggc

cgccgtcacg gccacagttg ccttcgtgct gatctactcg tcgcgggatt gccagcccct gcaggggggc tccatgtcct

acccgctgca gctcttttgt gcagatggcg agtacaactc catggctgcg gccttcttca acaccccgga gaagagcgtg

gtgagcctct tccacgaccc gccaggctcc tacaaccccc tgaccctcgg cctgttcacg ctggtctact tcttcctggc

ctgctggacc tacgggctca cggtgtctgc cggggtcttc atcccgtccc tgctcatcgg ggctgcctgg ggccggctct

ttgggatctc cctgtcctac ctcacggggg cggcgatctg ggcggacccc ggcaaatacg ccctgatggg agctgctgcc

cagctgggcg ggattgtgcg gatgacactg agcctgaccg tcatcatgat ggaggccacc agcaacgtga cctacggctt

ccccatcatg ctggtgctca tgaccgccaa gatcgtgggc gacgtcttca ttgagggcct gtacgacatg cacattcagc

tgcagagtgt gcccttcctg cactgggagg ccccggtcac ctcacactca ctcactgcca gggaggtgat gagcacacca

gtgacctgcc tgaggcggcg tgagaaggtc ggcgtcattg tggacgtgct gagcgacacg gcgtccaatc acaacggctt

ccccgtggtg gagcatgccg atgacaccca gcctgcccgg ctccagggcc tgatcctgcg ctcccagctc atcgttctcc

taaagcacaa ggtgtttgtg gagcggtcca acctgggcct ggtacagcgg cgcctgaggc tgaaggactt ccgagacgcc

tacccgcgct tcccacccat ccagtccatc cacgtgtccc aggacgagcg ggagtgcacc atggacctct ccgagttcat

gaacccctcc ccctacacgg tgccccagga ggcgtcgctc ccacgggtgt tcaagctgtt ccgggccctg ggcctgcggc

acctggtggt ggtggacaac cgcaatcagg ttgtcgggtt ggtgaccagg aaggacctcg ccaggtaccg cctgggaaag

agaggcttgg aggagctctc gctggcccag acgtgaggcc cagccctgcc cataatgggc actggcgctg gcaccccggc

ccttctgcat ttcctcccgg agtcactggt ttctcggccc aaaccatgct ccccagcagt ggcaatggcg agcaccctgc

agctgggcgg gcaggcggca ggcgcggaac tgaccctctc gcgggactga ccctgttgtg ggcagtggtc tccccccttg

gcgcctcctt gcgcaggccc agcctccact ctcctcgtct aggtttcttt acctccaggg atcagctgtg tgtgtgtgac ctccctaccg

ggctatcggc ctcttgggag ccagcggcag ggccggcacc tgcgtgcctg tgcccgtgtg cgtgagacag agcccttgcc

cctgctgctg ccccgagggc tgccctgccc tggaagggcc cctctgcctc cacaccagtg gagtcttcga gacttgggag

ctgcttggcc tcattttcag ccatgagcag acggcctgtg gtccctgggc ctgaggcacg gactcgtagc accagggttt

-continued

ggaggctgcg accgccccgg agagcagctt cacactggcg ccacagagga gccccacgtg cactccccgg cctgcatccg

gcttgggtac acaggcccag aggactgggg tgactcacgg gccctgtgct gtgatgttga gagctgagaa aaacctccaa

ggccctgagc cccatgccca gccctgcctt ggtcccccaa tccccagagc ttggagtctg ggccccacac ccagccctgc

cttggtccct gagcctcaaa gcgtggaatt gctgccctgt ggacact

GenBank Accession No.: AAH68581.1, human OSTM1, which provides the amino acid sequence (SEQ ID NO: 13):
MEPGPTAAQR RCSLPPWLPL GLLLWSGLAL GALPFGSSPH RVFHDLLSEQ

QLLEVEDLSL SLLQGGGLGP LSLPPDLPDL DPECRELLLD FANSSAELTG

CLVRSARPVR LCQTCYPLFQ QWSKMDNIS RAAGNTSESQ SCARSLLMAD

RMQIWILSE FFNTTWQEAN CANCLTNNSE ELSNSTVYFL NLFNHTLTCF

EHNLQGNAHS LLQTKNYSEV CKNCREAYKT LSSLYSEMQK MNELENKAEP

GTHLCIDVED AMNITRKLWS RTFNCSVPCS DTVPVIAVSV FILFLPWFY LSSFLHSEQK

KRKLILPKRL KSSTSFANIQ ENSN

GenBank Accession No.: BC068581. 1, human OSTM1 mRNA, which provides the nucleic acid sequence (SEQ ID NO: 14):
ggctgtccgc ggtgccggct gggggcggag aggcggcggt gggctccctg gggtgtgtga gcccggtgat ggagccgggc

ccgacagccg cgcagcggag gtgttcgttg ccgccgtggc tgccgctggg gctgctgctg tggtcggggc tggccctggg

cgcgctcccc ttcggcagca gtccgcacag ggtcttccac gacctcctgt cggagcagca gttgctggag gtggaggact

tgtccctgtc cctcctgcag ggtggagggc tggggcctct gtcgctgccc ccggacctgc cggatctgga tcctgagtgc

cgggagctcc tgctggactt cgccaacagc agcgcagagc tgacagggtg tctggtgcgc agcgcccggc ccgtgcgcct

ctgtcagacc tgctaccccc tcttccaaca ggtcgtcagc aagatggaca acatcagccg agccgcgggg aatacttcag

agagtcagag ttgtgccaga agtctcttaa tggcagatag aatgcaaata gttgtgattc tctcagaatt ttttaatacc acatggcagg

aggcaaattg tgcaaattgt ttaacaaaca acagtgaaga attatcaaac agcacagtat atttccttaa tctatttaat cacaccctga

cctgctttga acataacctt caggggaatg cacatagtct tttacagaca aaaaattatt cagaagtatg caaaaactgc cgtgaagcat

acaaaactct gagtagtctg tacagtgaaa tgcaaaaaat gaatgaactt gagaataagg ctgaacctgg aacacattta tgcattgatg

tggaagatgc aatgaacatc actcgaaaac tatggagtcg aactttcaac tgttcagtcc cttgcagtga cacagtgcct gtaattgctg

tttctgtgtt cattctcttt ctacctgttg tcttctacct tagtagcttt cttcactcag agcaaaagaa acgcaaactc attctgccca

aacgtctcaa gtccagtacc agttttgcaa atattcagga aaattcaaac tgagacctac aaaatggaga attgacatat cacgtgaatg

aatggtggaa gacacaactt ggtttcagaa agaagataaa ctgtgatttg acaagtcaag ctcttaagaa atacaaggac

ttcagatcca tttttaaata agaattttcg atttttcttt ccttttccac ttctttctaa cagatttgga tatttttaat ttccaggcat

agcagtgtta

tctattttaa tgtgtatttg tcacaataac agaacatgca agaacaatca ttattttatt ttataggcat ttgattacta ttctagactt

ctggtatctt cttactaaca taagtatctc aagtagaaaa gtttttgaaa actaacattt aaaaattaat cagttacagt aaagactttg

aaaaagaaat gtacttgtta ggaagtagct taattacccc ccattgcagt attattgtta tatatatagt taatatgttg tacatcacaa

taatatataa ttcagtctct agtttcccta gagtcatttt tgaaaccact gattgcaaac ctccctgaca atttttaaaa gtagtaagcc

acattacatt tatctttgta aaaagattta tggtaactgg tttcttactt gacttttata aatagtattt tacatcttat ttttgccttt

atttcataag

taatttaaaa atcactggat tgctttatta tattcagggc aatatggatt atttttatac caaggatttg catcgtgaat taaattaagt

tatttggcaa tttataattt attactactt taaatcaaat gtagcattat cacactgtat ttaaattgtc atttttaaa ggaatatttt

cttcttaaga

tatatagagg attttggaga agagagacag gaggggtaaa accagcttaa ggttcagcga gcagaaaggg acctgagagg

-continued

atgctcactg taagactgtt ggacagtggt gtgtattgag gggatgaatc ggaacgatag tctcatgcag aaaatagtga gattaagatc

atccttattg tttctaaatt atttcaatca gatgaaagtg atacgattga aatgaaatca catagttcgt gctcagaaat tctattttgg

tatgtttgta ttagccttta gaaaaaacac tccgtttcag aattgttcac agttttattt cttaggtttt tagagttcag gatttcattt

attaatttct tcttgctttt ttggtggaaa taggctttgt tgtaaacatt aagaatataa aatctcctct atatagaaac aagaattttg

ttaaaaagag aatttgaatc ccttcctata ctataaaatg ctctataggg agacaaagtg tttctttttt cttttatgtt tactgtttat

gtggagtgaa atataaggct cttggatgta taacatactc aaaagctgtt acactttctc tgatctgctg tgatccactg aaaatgtgct

ggggtttgtt ctgctgtcac tgtttatgct gctggaactt agcactgtct tgatttgaag catatgattg agagccattt gaagcaatct

tcattaatgc agataaaaca agtttacatg tgcagagtta gaaaatgaca tgttcaattc tgtaagtggt gacttttga gcacctttca

gtattatgta tttgtaaaaa ccattgtttt tggatataaa gctaataagc actttaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa

aaaaaaaaaa aaaaaaaaaa aaagaaaaaa aaaaaaaaa

GenBank Accession No.: AA000797.1, human MCOLNI, which provides the amino acid sequence (SEQ ID NO: 15):
MTAPAGPRGSETERLLTPNPGYGTQAGPSPAPPTPPEEEDLRRRLKYFFMSPCDKFRAKG

RKPCKLMLQVVKILVVTVQLILFGLSNQLAVTFREENTIAFRHLFLLGYSDGADDTFAAY

TREQLYQAIFHAVDQYLALPDVSLGRYAYVRGGGDPWINGSGLALCQRYYHRGHVDP

ANDTFDIDPMVVTDCIQVDPPERPPPPPSDDLTLLESSSSYKNLTLKFHKLVNVTIHFRLK

TINLQSLINNEIPDCYTFSVLITFDNKAHSGRIPISLETQAHIQECKHPSVFQHGDNSFRLLF

DVVVILTCSLSFLLCARSLLRGFLLQNEFVGFMWRQRGRVISLWERLEFVNGWYILLVT

SDVLTISGTIMKIGIEAKNLASYDVCSILLGTSTLLVWVGVIRYLTFFHNYNILIATLRVAL

PSVMRFCCCVAVIYLGYCFCGWIVLGPYHVKFRSLSMVSECLFSLINGDDMFVTFAAMQ

AQQGRSSLVWLFSQLYLYSFISLFIYMVLSLFIALITGAYDTIKHPGGAGAEESELQAYIA

QCQDSPTSGKFRRGSGSACSLLCCCGRDPSEEHSLLVN

GenBank Accession No.: AF287269.1, human MCOLN1 mRNA, complete cds, which provides the nucleic acid sequence (SEQ ID NO: 16):
agatcagctg atgccggagg gtttgaagcc gcgccgcgag ggagcgaggt cgcagtgaca gcggcgggcg atcggaccca

ggctgccccg ccgtacccgc ctgcgtcccg cgctcccgcc ccagcatgac agccccggcg ggtccgcgcg gctcagagac

cgagcggctt ctgaccccca accccgggta tgggacccag gcggggcctt caccggcccc tccgacaccc ccagaagagg

aagaccttcg ccgtcgtctc aaatactttt tcatgagtcc ctgcgacaag tttcgagcca agggccgcaa gccctgcaag

ctgatgctgc aagtggtcaa gatcctggtg gtcacggtgc agctcatcct gtttgggctc agtaatcagc tggctgtgac

attccgggaa gagaacacca tcgccttccg acacctcttc ctgctgggct actcggacgg agcggatgac accttcgcag

cctacacgcg ggagcagctg taccaggcca tcttccatgc tgtggaccag tacctggcgt tgcctgacgt gtcactgggc

cggtatgcgt atgtccgtgg tgggggtgac ccttggacca atggctcagg gcttgctctc tgccagcggt actaccaccg

aggccacgtg gacccggcca acgacacatt tgacattgat ccgatggtgg ttactgactg catccaggtg gatccccccg

agcggccccc tccgcccccc agcgacgatc tcaccctctt ggaaagcagc tccagttaca agaacctcac gctcaaattc

cacaagctgg tcaatgtcac catccacttc cggctgaaga ccattaacct ccagagcctc atcaataatg agatcccgga

ctgctatacc ttcagcgtcc tgatcacgtt tgacaacaaa gcacacagtg ggcggatccc catcagcctg gagacccagg

cccacatcca ggagtgtaag caccccagtg tcttccagca cggagacaac agcttccggc tcctgtttga cgtggtggtc

atcctcacct gctccctgtc cttcctcctc tgcgcccgct cactccttcg aggcttcctg ctgcagaacg agtttgtggg gttcatgtgg

cggcagcggg gacgggtcat cagcctgtgg gagcggctgg aatttgtcaa tggctggtac atcctgctcg tcaccagcga

tgtgctcacc atctcgggca ccatcatgaa gatcggcatc gaggccaaga acttggcgag ctacgacgtc tgcagcatcc

tcctgggcac ctcgacgctg ctggtgtggg tgggcgtgat ccgctacctg accttcttcc acaactacaa tatcctcatc gccacactgc

-continued

gggtggccct gcccagcgtc atgcgcttct gctgctgcgt ggctgtcatc tacctgggct actgcttctg tggctggatc

gtgctggggc cctatcatgt gaagttccgc tcactctcca tggtgtctga gtgcctgttc tcgctcatca atggggacga catgtttgtg

acgttcgccg ccatgcaggc gcagcagggc cgcagcagcc tggtgtggct cttctcccag ctctaccttt actccttcat cagcctcttc

atctacatgg tgctcagcct cttcatcgcg ctcatcaccg gcgcctacga caccatcaag catcccggcg gcgcaggcgc

agaggagagc gagctgcagg cctacatcgc acagtgccag gacagcccca cctccggcaa gttccgccgc gggagcggct

cggcctgcag ccttctctgc tgctgcggaa gggacccctc ggaggagcat tcgctgctgg tgaattgatt cgacctgact

gccgttggac cgtaggccct ggactgcaga gacccccgcc cccgaccccg cttatttatt tgtagggttt gcttttaagg atcggctccc

tgtcgcgccc gaggagggcc tggacctttc gtgtcggacc cttgggggcg gggagactgg gtggggaggg tgttgaataa a

GenBank Accession No .: Q68CP4.2, human HGSNAT, which provides the amino acid sequence (SEQ ID NO: 17):

MTGARASAAE QRRAGRSGQA RAAERAAGMS GAGRALAALL LAASVLSAAL

LAPGGSSGRD AQAAPPRDLD KKRHAELKMD QALLLIHNEL LWTNLTVYWK

SECCYHCLFQ VLVNVPQSPK AGKPSAAAAS VSTQHGSILQ LNDTLEEKEV

CRLEYRFGEF GNYSLLVKNI HNGVSEIACD LAVNEDPVDS NLPVSIAFLI GLAVIIVISF

LRLLLSLDDF NNWISKAISS RETDRLINSE LGSPSRTDPL DGDVQPATWR LSALPPRLRS

VDTFRGIALI LMVFVNYGGG KYWYFKHASW NGLTVADLVF PWFVFIMGSS

IFLSMTSILQ RGCSKFRLLG KIAWRSFLLI CIGIIIVNPN YCLGPLSWDK VRIPGVLQRL

GVTYFVVAVL ELLFAKPVPE HCASERSCLS LRDITSSWPQ WLLILVLEGL

WLGLTFLLPV PGCPTGYLGP GGIGDFGKYP NCTGGAAGYI DRLLLGDDHL

YQHPSSAVLY HTEVAYDPEG ILGTINSIVM AFLGVQAGKI LLYYKARTKD

ILIRFTAWCC ILGLISVALT KVSENEGFIP VNKNLWSLSY VTTLSSFAFF ILLVLYPVVD

VKGLWTGTPF FYPGMNSILV YVGHEVFENY FPFQWKLKDN QSHKEHLTQN

IVATALWVLI AYILYRKKIF WKI

GenBank Accession No.: NM 152419, human HGSNAT mRNA, which provides the nucleic acid sequence (SEQ ID NO: 18):

agggcggggc gcagcgggca ggcaagggcg gccgagcggg cggcgggcat gagcggggcg ggcagggcgc tggccgcgct

gctgctggcc gcgtccgtgc tgagcgccgc gctgctggcc cccggcggct cttcggggcg cgatgcccag gccgcgccgc

cacgagactt agacaaaaaa agacatgcag agctgaagat ggatcaggct ttgctactca tccataatga acttctctgg

accaacttga ccgtctactg gaaatctgaa tgctgttatc actgcttgtt tcaggttctg gtaaacgttc ctcagagtcc aaaagcaggg

aagcctagtg ctgcagctgc ctctgtcagc acccagcacg gatctatcct gcagctgaac gacaccttgg aagagaaaga

agtttgtagg ttggaataca gatttggaga atttggaaac tattctctct tggtaaagaa catccataat ggagttagtg aaattgcctg

tgacctggct gtgaacgagg atccagttga tagtaacctt cctgtgagca ttgcattcct tattggtctt gctgtcatca ttgtgatatc

ctttctgagg ctcttgttga gtttggatga ctttaacaat tggatttcta aagccataag ttctcgagaa actgatcgcc tcatcaattc

tgagctggga tctcccagca ggacagaccc tctcgatggt gatgttcagc cagcaacgtg gcgtctatct gccctgccgc

cccgcctccg cagcgtggac accttcaggg ggattgctct tatactcatg gtctttgtca attatggagg aggaaaatat tggtacttca

aacatgcaag ttggaatggg ctgacagtgg ctgacctcgt gttcccgtgg tttgtattta ttatgggatc ttccattttt ctatcgatga

cttctatact gcaacggggg tgttcaaaat tcagattgct ggggaagatt gcatggagga gtttcctgtt aatctgcata ggaattatca

ttgtgaatcc caattattgc cttggtccat tgtcttggga caaggtgcgc attcctggtg tgctgcagcg attgggagtg acatactttg

tggttgctgt gttggagctc ctctttgcta aacctgtgcc tgaacattgt gcctcggaga ggagctgcct ttctcttcga gacatcacgt

ccagctggcc ccagtggctg ctcatcctgg tgctggaagg cctgtggctg ggcttgacat tcctcctgcc agtccctggg

tgccctactg gttatcttgg tcctgggggc attggagatt ttggcaagta tccaaattgc actggaggag ctgcaggcta catcgaccgc

-continued

ctgctgctgg gagacgatca cctttaccag cacccatctt ctgctgtact ttaccacacc gaggtggcct atgaccccga

gggcatcctg ggcaccatca actccatcgt gatggccttt ttaggagttc aggcaggaaa aatactattg tattacaagg

ctcggaccaa agacatcctg attcgattca ctgcttggtg ttgtattctt gggctcattt ctgttgctct gacgaaggtt tctgaaaatg

aaggctttat tccagtaaac aaaaatctct ggtccctttc gtatgtcact acgctcagtt cttttgcctt cttcatcctg ctggtcctgt

acccagttgt ggatgtgaag gggctgtgga caggaacccc attcttttat ccaggaatga attccattct ggtatatgtc ggccacgagg

tgtttgagaa ctacttcccc tttcagtgga agctgaagga caaccagtcc cacaaggagc acctgactca gaacatcgtc

gccactgccc tctgggtgct cattgcctac atcctctata gaaagaagat tttttggaaa atctgatggc tcccactgag atgtgctgct

ggaagactct agtaggcctg cagggaggac tgaagcagcc tttgttaaag ggaagcattc attaggaaat tgactggctg

cgtgtttaca gactctgggg gaagacactg atgtcctcaa actggttaac tgtgacacgg ctcgccagaa ctctgcctgt ctatttgtga

cttacagatt tgaaatgtaa ttgtcttttt tcctccatct tctgtggaaa tggatgtctt tggaacttca ttccgaggag ataagcttta

actttccaaa agggaattgc catgggtgtt tttcttctgt ggtgagtgaa acaatctgag gtctggttct tgctgacctt gttgccctgc

aaacttcctt tccacgtgta cgcgcacacc aacacgaaat gccatcactc ctactgcggc tgctatgaag cttactggtt gtgatgtgtt

ataatttagt ctgtttttt gattgaatgc agtttaatgt ttccagaaag ccaaagtaat tttcttttca gatatgcaag gctttggtgg

gtccaaaaaa tgtctatcac aagccatttt ttccttttcc tctctcgaaa agttaaaata tctatgtgtt attcccaaac cctcttacct

atgtatctgc ctgtctgtcc atcatcttcc ttcctcccta tctctgtgta tctggatggc agccgctgcc caggggagtg gctgtgggga

gggcaggtac tgtctttgcc tgtgggtcca gctgagccat ccctgctggg tgatgctggg caagaccctt ggcccgtctg

ggccttggct tcctcacttg tgaaatgagc gggaagatga ctctcagttc cttccacctc ttagacatgg tgaggtaaca gacatcaaaa

gcttttctga aatcttcaga agaaatagtt ccattacaga aaactcttca aaataaatag tagtgaaaac ttttaaaaac tctcattgga

gtaagtcttt tcaagatgat cctccacaat ggaggcagcg ttcctacttg tcatcacaca gctgaagaca ttgtttctta ggtgtgaaat

cggggacaaa ggacaaacag agacacacgg cattgttcat gggaggcatc gtcaccctcc tgggtgttct gtgggaattt

cctgtgtgag gaaaacgtgg ccacagggtt gtgctgtacc caccttccc cggcgagatg gccctcggcc tgtgccgctg

cttccaccct cgccactcca tggcagcttt tggtctgttt ccggctctgc cctctgccct gaactctcat ccggcttgta cctgcctgct

ggacccctcc acctggaggc cagcccatgt ctcaggccca gccctagcct cttctcctca aattctaagt gttttctctt taggtttccc

tggctttgtg aatggatcat gtgtctctag gtataaacct gacatcatct ttccacccgg cttacctcca ccagatctcc ccagttctgt

ctccatcttc tacctgcagc tgctctgttc tcatggtcac tgctgcatca ctgagtctgg acccttgtta tcattttcaa actggcctcc

ttccctcgtt ccccacttct taaagtcacc tgtccattgc caccagatta agctttctcc agccagatca cctctctctg agaaacctcc

attgacatgg aaacaccatt gtctggcaca catactcaca tactcacctt cccgtcttga tccccacaca tctttccagc ctcccctccc

actccactcc ctgctctctc ctccacctcc ccatcctctt gtctcccctc ccctctgaat ccagcccagc ggggcttctc ctgcctccat

cacatcacag aagtacctcc tgcttctggt tttaattaga gccttccccg attacatttt cctctgaatt ttttcctatc tacatttgat

ctgtcatgtt taaacccccc acttctaagg gaacttctct aatctcttat cctcatcccc aaatagtgtt ttcttcctct gggttcttat

aatgttggta tcaatctcac agcatttagt gcttcctgcc tggtgtgaca gttacctgtg tgcatgtgca atttctaatt tcccacgcta

gactgtgagc ttcctaaggc aagaatcatg ccttgttggt ttctgtattc ctcatggtgc caaacacagt gccttctaca ttgcaggcgc

tgaataaaca tttttaaagc aaaatgatgt ggatttttaa aataaatatt taagtgctgg taagatgagc atgtatccgg ggtgcccatg

aaatgttctt ggggccgtgt ggggacagtc gtcattcctc ctcctgccac ccttttcttt cagtgagtca ctgtggatgg tcccagctgt

gtcatcccaa agttcagcag ggaaagctga gctgggcctc tccaggtgag ttttctagaa gcatttctca aactgtgggt tacatcaact

tgggtgtctt gagctgtaag gaaggaactc cggagtcagc tgggctacag gggagcttct ctaagtcctg cgggaggcca

gacccagcct gagcttgctg ttagctagcg gaggcagctg ctggtggccc aggtgctcga caccaggcat cccctctcct

cccacgaagg gtgtgccata atccccttca acaggaaatg cttcccagaa gcctctcagc agcctcccct cctgtcctat

cagctagaag cgcctcgctt gtcccaagac cagcagggac agggaactgt ccgagcccgt ggctgtgtgg aggaaggcga

-continued

cccccagcac aagattggtt tcctttggga agggaagagg gagtgtgttg gggtaagggg tagagcagag gaatggtcag

ggggcaacaa ccgctgacag ctgcaacagg tgcatggcat ctcacaggga ggcagggagg tgcgagctcc taagtaatgg

agcaaaaaaa ttctattctg tagaatgggg agagaaaatg tgacatttta attttttttt gcatttatat tcctaattcc tacttaaagt

gaatatactg ccgctgtaga tcataaaatg tatcttttcc atggccaaca aggggcatct tttataaatg cataataacc cagtttgtat

caaagggtat cgacttaagt gaaatttcaa catgctgtta ctttttcctt ttaatgtaat tctgttttcc aaataaatgg gggagacaaa

tggaaaaaaa aaaaaaaa

GenBank Accession No.: AAB63982.1, human NPC1, which provides the amino acid sequence (SEQ ID NO: 19):
MTARGLALGLLLLLLCPAQVFSQSCVWYGECGIAYGDKRYNCEYSGPPKPLPKDGYDL

VQELCPGFFFGNVSLCCDVRQLQTLKDNLQLPLQFLSRCPSCFYNLLNLFCELTCSPRQS

QFLNVTATEDYVDPVTNQTKTNVKELQYYVGQSFANAMYNACRDVEAPSSNDKALGL

LCGKDADACNATNWIEYMFNKDNGQAPFTITPVFSDFPVHGMEPMNNATKGCDESVD

EVTAPCSCQDCSIVCGPKPQPPPPPAPWTILGLDAMYVIMWITYMAFLLVFFGAFFAVW

CYRKRYFVSEYTPIDSNIAFSVNASDKGEASCCDPVSAAFEGCLRRLFTRWGSFCVRNPG

CVIFFSLVFITACSSGLVFVRVTTNPVDLWSAPSSQARLEKEYFDQHFGPFFRTEQLIIRAP

LTDKHIYQPYPSGADVPFGPPLDIQILHQVLDLQIAIENITASYDNETVTLQDICLAPLSPY

NTNCTILSVLNYFQNSHSVLDHKKGDDFFVYADYHTHFLYCVRAPASLNDTSLLHDPCL

GTFGGPVFPWLVLGGYDDQNYNNATALVITFPVNNYYNDTEKLQRAQAWEKEFINFVK

NYKNPNLTISFTAERSIEDELNRESDSDVFTVVISYAIMFLYISLALGHIKSCRRLLVDSKV

SLGIAGILIVLSSVACSLGVFSYIGLPLTLIVIEVIPFLVLAVGVDNIFILVQAYQRDERLQG

ETLDQQLGRVLGEVAPSMFLSSFSETVAFFLGALSVMPAVHTFSLFAGLAVFIDFLLQITC

FVSLLGLDIKRQEKNRLDIFCCVRGAEDGTSVQASESCLFRFFKNSYSPLLLKDWMRPIVI

AIFVGVLSFSIAVLNKVDIGLDQSLSMPDDSYMVDYFKSISQYLHAGPPVYFVLEEGHDY

TSSKGQNMVCGGMGCNNDSLVQQIFNAAQLDNYTRIGFAPSSWIDDYFDWVKPQSSCC

RVDNITDQFCNASVVDPACVRCRPLTPEGKQRPQGGDFMRFLPMFLSDNPNPKCGKGG

HAAYSSAVNILLGHGTRVGATYFMTYHTVLQTSADFIDALKKARLIASNVTETMGINGS

AYRVFPYSVFYVFYEQYLTIIDDTIFNLGVSLGAIFLVTMVLLGCELWSAVIMCATIAMV

LVNMFGVMWLWGISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSRVERAEEALA

HMGSSVFSGITLTKFGGIVVLAFAKSQIFQIFYFRMYLAMVLLGATHGLIFLPVLLSYIGP

SVNKAKSCATEERYKGTERERLLNF

GenBank Accession No.: AF002020J, human NPC1 mRNA, complete cds, which provides the nucleic acid sequence (SEQ ID NO: 20):
tttgctcctg ctcctccgct cctcctgcgc ggggtgctga aacagcccgg ggaagtagag ccgcctccgg ggagcccaac

cagccgaacg ccgccggcgt cagcagcctt gcgcggccac agcatgaccg ctcgcggcct ggcccttggc ctcctcctgc

tgctactgtg tccagcgcag gtgttttcac agtcctgtgt ttggtatgga gagtgtggaa ttgcatatgg ggacaagagg tacaattgcg

aatattctgg cccaccaaaa ccattgccaa aggatggata tgacttagtg caggaactct gtccaggatt cttctttggc aatgtcagtc

tctgttgtga tgttcggcag cttcagacac taaaagacaa cctgcagctg cctctacagt ttctgtccag atgtccatcc tgtttttata

acctactgaa cctgttttgt gagctgacat gtagccctcg acagagtcag tttttgaatg ttacagctac tgaagattat gttgatcctg

ttacaaacca gacgaaaaca aatgtgaaag agttacaata ctacgtcgga cagagttttg ccaatgcaat gtacaatgcc

tgccgggatg tggaggcccc ctcaagtaat gacaaggccc tgggactcct gtgtgggaag gacgctgacg cctgtaatgc

caccaactgg attgaataca tgttcaataa ggacaatgga caggcacctt ttaccatcac tcctgtgttt tcagattttc cagtccatgg

-continued

gatggagccc atgaacaatg ccaccaaagg ctgtgacgag tctgtggatg aggtcacagc accatgtagc tgccaagact

gctctattgt ctgtggcccc aagccccagc ccccacctcc tcctgctccc tggacgatcc ttggcttgga cgccatgtat gtcatcatgt

ggatcaccta catggcgttt ttgcttgtgt tttttggagc atttttgca gtgtggtgct acagaaaacg gtattttgtc tccgagtaca

ctcccatcga tagcaatata gctttttctg ttaatgcaag tgacaaagga gaggcgtcct gctgtgaccc tgtcagcgca gcatttgagg

gctgcttgag gcggctgttc acacgctggg ggtctttctg cgtccgaaac cctggctgtg tcattttctt ctcgctggtc ttcattactg

cgtgttcgtc aggcctggtg tttgtccggg tcacaaccaa tccagttgac ctctggtcag cccccagcag ccaggctcgc

ctggaaaaag agtactttga ccagcacttt gggcctttct tccggacgga gcagctcatc atccgggccc ctctcactga

caaacacatt taccagccat acccttcggg agctgatgta ccctttggac ctccgcttga catacagata ctgcaccagg ttcttgactt

acaaatagcc atcgaaaaca ttactgcctc ttatgacaat gagactgtga cacttcaaga catctgcttg gcccctcttt caccgtataa

cacgaactgc accattttga gtgtgttaaa ttacttccag aacagccatt ccgtgctgga ccacaagaaa ggggacgact tctttgtgta

tgccgattac cacacgcact ttctgtactg cgtacgggct cctgcctctc tgaatgatac aagtttgctc catgacccct gtctgggtac

gtttggtgga ccagtgttcc cgtggcttgt gttgggaggc tatgatgatc aaaactacaa taacgccact gcccttgtga ttaccttccc

tgtcaataat tactataatg atacagagaa gctccagagg gcccaggcct gggaaaaaga gtttattaat tttgtgaaaa actacaagaa

tcccaatctg accatttcct tcactgctga acgaagtatt gaagatgaac taaatcgtga aagtgacagt gatgtcttca ccgttgtaat

tagctatgcc atcatgtttc tatatatttc cctagccttg gggcacatca aaagctgtcg caggcttctg gtggattcga aggtctcact

aggcatcgcg ggcatcttga tcgtgctgag ctcggtggct tgctccttgg gtgtcttcag ctacattggg ttgcccttga ccctcattgt

gattgaagtc atcccgttcc tggtgctggc tgttggagtg gacaacatct tcattctggt gcaggcctac cagagagatg aacgtcttca

aggggaaacc ctggatcagc agctgggcag ggtcctagga gaagtggctc ccagtatgtt cctgtcatcc ttttctgaga

ctgtagcatt tttcttagga gcattgtccg tgatgccagc cgtgcacacc ttctctctct ttgcgggatt ggcagtcttc attgactttc

ttctgcagat tacctgtttc gtgagtctct tggggttaga cattaaacgt caagagaaaa atcggctaga catcttttgc tgtgtcagag

gtgctgaaga tggaacaagc gtccaggcct cagagagctg tttgtttcgc ttcttcaaaa actcctattc tccacttctg ctaaaggact

ggatgagacc aattgtgata gcaatatttg tgggtgttct gtcattcagc atcgcagtcc tgaacaaagt agatattgga ttggatcagt

ctctttcgat gccagatgac tcctacatgg tggattattt caaatccatc agtcagtacc tgcatgcggg tccgcctgtg tactttgtcc

tggaggaagg gcacgactac acttcttcca aggggcagaa catggtgtgc ggcggcatgg gctgcaacaa tgattccctg

gtgcagcaga tatttaacgc ggcgcagctg gacaactata cccgaatagg cttcgccccc tcgtcctgga tcgacgatta

tttcgactgg gtgaagccac agtcgtcttg ctgtcgagtg gacaatatca ctgaccagtt ctgcaatgct tcagtggttg accctgcctg

cgttcgctgc aggcctctga ctccggaagg caaacagagg cctcaggggg gagacttcat gagattcctg cccatgttcc

tttcggataa ccctaacccc aagtgtggca aaggggggaca tgctgcctat agttctgcag ttaacatcct ccttggccat

ggcaccaggg tcggagccac gtacttcatg acctaccaca ccgtgctgca gacctctgct gactttattg acgctctgaa

gaaagcccga cttatagcca gtaatgtcac cgaaaccatg ggcattaacg gcagtgccta ccgagtattt ccttacagtg tgttttatgt

cttctacgaa cagtacctga ccatcattga cgacactatc ttcaacctcg gtgtgtccct gggcgcgata tttctggtga ccatggtcct

cctgggctgt gagctctggt ctgcagtcat catgtgtgcc accatcgcca tggtcttggt caacatgttt ggagttatgt ggctctgggg

catcagtctg aacgctgtat ccttggtcaa cctggtgatg agctgtggca tctccgtgga gttctgcagc cacataacca gagcgttcac

ggtgagcatg aaaggcagcc gcgtggagcg cgcggaagag gcacttgccc acatgggcag ctccgtgttc agtggaatca

cacttacaaa atttggaggg attgtggtgt tggcttttgc caaatctcaa attttccaga tattctactt caggatgtat ttggccatgg

tcttactggg agccactcac ggattaatat ttctccctgt cttactcagt tacatagggc catcagtaaa taaagccaaa agttgtgcca

ctgaagagcg atacaaagga acagagcgcg aacggcttct aaatttctag ccctctcgca gggcatcctg actgaactgt

gtctaagggt cggtcggttt accactggac gggtgctgca tcggcaaggc caagttgaac accggatggt gccaaccatc

ggttgtttgg cagcagcttt gaacgtagcg cctgtgaact caggaatgca cagttgactt gggaagcagt attactagat

-continued

ctggaggcaa ccacaggaca ctaaacttct cccagcctct tcaggaaaga aacctcattc tttggcaagc aggaggtgac

actagatggc tgtgaatgtg atccgctcac tgacactctg taaaggccaa tcaatgcact gtctgtcctc tccttttag gagtaagcca

tcccacaagt tctataccat atttttagtg acagttgagg ttgtagatac actttataac attttatagt ttaaagagct ttattaatgc

aataaattaa ctttgtacac atttttatat aaaaaaacag caagtgattt cagaatgttg taggcctcat tagagcttgg tctccaaaaa

tctgtttgaa aaaagcaaca tgttcttcac agtgttcccc tagaaaggaa gagatttaat tgccagttag atgtggcatg aaatgaggga

caaagaaagc atctcgtagg tgtgtctact gggttttaac ttatttttct ttaataaaat acattgtttt cctaaaaaaa aaa

GenBank Accession No.: CAA54416.1, human LAMP-2A, which provides the amino acid sequence (SEQ ID NO: 21):
MVCFRLFPVP GSGLVLVCLV LGAVRSYALE LNLTDSENAT CLYAKWQMNF

TVRYETTNKT YKTVTISDHG TVTYNGSICG DDQNGPKIAV QFGPGFSWIA

NFTKAASTYS IDSVSFSYNT GDNTTFPDAE DKGILTVDEL LAIRIPLNDL FRCNSLSTLE

KNDWQHYWD VLVQAFVQNG TVSTNEFLCD KDKTSTVAPT IHTTVPSPTT

TPTPKEKPEA GTYSVNNGND TCLLATMGLQ LNITQDKVAS VININPNTTH

STGSCRSHTA LLRLNSSTIK YLDFVFAVKN ENRFYLKEVN ISMYLVNGSV

FSIANNNLSY WDAPLGSSYM CNKEQTVSVS GAFQINTFDL RVQPFNVTQG

KYSTAQDCSA DDDNFLVPIA VGAALAGVLI LVLLAYFIGL KHHHAGYEQF

GenBank Accession No.: X77196. 1, human LAMP2 mRNA, which provides the nucleic acid sequence (SEQ ID NO: 22):
ccgattcctg gcttttgcaa ggctgtggtc ggtggtcatc agtgctcttg acccaggtcc agcgagcctt ttccctggtg ttgcagctgt

tgttgtaccg ccgccgtcgc cgccgtcgcc gcctgctctg cggggtcatg gtgtgcttcc gcctcttccc ggttccgggc

tcagggctcg ttctggtctg cctagtcctg ggagctgtgc ggtcttatgc attggaactt aatttgacag attcagaaaa tgccacttgc

ctttatgcaa aatggcagat gaatttcaca gtacgctatg aaactacaaa taaaacttat aaaactgtaa ccatttcaga ccatggcact

gtgacatata atggaagcat ttgtggggat gatcagaatg gtcccaaaat agcagtgcag ttcggacctg gcttttcctg gattgcgaat

tttaccaagg cagcatctac ttattcaatt gacagcgtct cattttccta caacactggt gataacacaa catttcctga tgctgaagat

aaaggaattc ttactgttga tgaacttttg gccatcagaa ttccattgaa tgacctttt agatgcaata gtttatcaac tttggaaaag

aatgatgttg tccaacacta ctgggatgtt cttgtacaag cttttgtcca aaatggcaca gtgagcacaa atgagttcct gtgtgataaa

gacaaaactt caacagtggc acccaccata cacaccactg tgccatctcc tactacaaca cctactccaa aggaaaaacc

agaagctgga acctattcag ttaataatgg caatgatact tgtctgctgg ctaccatggg gctgcagctg aacatcactc aggataaggt

tgcttcagtt attaacatca accccaatac aactcactcc acaggcagct gccgttctca cactgctcta cttagactca atagcagcac

cattaagtat ctagactttg tctttgctgt gaaaaatgaa aaccgatttt atctgaagga agtgaacatc agcatgtatt tggttaatgg

ctccgttttc agcattgcaa ataacaatct cagctactgg gatgcccccc tgggaagttc ttatatgtgc aacaaagagc agactgtttc

agtgtctgga gcatttcaga taaatacctt tgatctaagg gttcagcctt tcaatgtgac acaaggaaag tattctacag ctcaagactg

cagtgcagat gacgacaact tccttgtgcc catagcggtg ggagctgcct tggcaggagt acttattcta gtgttgctgg cttattttat

tggtctcaag caccatcatg ctggatatga gcaattttag aatctgcaac ctgattgatt atataaaaat acatgcaaat aacaagattt

tcttacctct cagttgttga aacactttgc ttcttaaaat tgatatgttg aaactttaat tcttttatca atcccagcat tttgagatca

gtctttatta ataaaacctg ttctctttaa tcagcttaaa atccaaagtg tcatatttac tggtcctgga gacaaacttg ttcaaaagaa

catcaacgtg caatgtttta aggtctatct taagaagccc tggccaaatt ttgatcctaa ccttgaagta tgccttgaac ttattaacat

ggccattata agaataaaat atgtagttgt gtcttaatgg aattaataaa tgtcatttca ctactggtgt tctgttttca atgtataagg

actatagtga tttaaactca tcaatgtgcc tttgcataaa gttgattaaa taaatattga tgtggtataa atgcccatca gatatgct

GenBank Accession No.: AAA91149.1, human LAMP-2B, which provides the amino acid sequence (SEQ ID NO: 23):
MVCFRLFPVPGSGLVLVCLVLGAVRSYALELNLTDSENATCLYAKWQMNFTVRYETTN

-continued

KTYKTVTISDHGTVTYNGSICGDDQNGPKIAVQFGPGFSWIANFTKAASTYSIDSVSFSY

NTGDNTTFPDAEDKGILTVDELLAIRIPLNDLFRCNSLSTLEKNDVVQHYWDVLVQAFV

QNGTVSTNEFLCDKDKTSTVAPTIHTTVPSPTTTPTPKEKPEAGTYSVNNGNDTCLLATM

GLQLNITQDKVASVININPNTTHSTGSCRSHTALLRLNSSTIKYLDFVFAVKNENRFYLKE

VNISMYLVNGSVFSIANNNLSYWDAPLGSSYMCNKEQTVSVSGAFQINTFDLRVQPFNV

TQGKYSTAQECSLDDDTILIPIIVGAGLSGLIIVIVIAYVIGRRKSYAGYQTL

GenBank Accession No.: U36336.1, human LAMP-2B mRNA, complete cds, which provides the nucleic acid sequence (SEQ ID NO: 24):

ccgattcctg gcttttgcaa ggctgtggtc ggtggtcatc agtgctcttg acccaggtcc agcgagcctt ttccctggtg ttgcagctgt

tgttgtaccg ccgccgtcgc cgccgtcgcc gcctgctctg cggggtcatg gtgtgcttcc gcctcttccc ggttccgggc

tcagggctcg ttctggtctg cctagtcctg ggagctgtgc ggtcttatgc attggaactt aatttgacag attcagaaaa tgccacttgc

ctttatgcaa aatggcagat gaatttcaca gtacgctatg aaactacaaa taaaacttat aaaactgtaa ccatttcaga ccatggcact

gtgacatata atggaagcat ttgtggggat gatcagaatg gtcccaaaat agcagtgcag ttcggacctg gcttttcctg gattgcgaat

tttaccaagg cagcatctac ttattcaatt gacagcgtct cattttccta caacactggt gataacacaa catttcctga tgctgaagat

aaaggaattc ttactgttga tgaacttttg gccatcagaa ttccattgaa tgacctttt agatgcaata gtttatcaac tttggaaaag

aatgatgttg tccaacacta ctgggatgtt cttgtacaag cttttgtcca aaatggcaca gtgagcacaa atgagttcct gtgtgataaa

gacaaaactt caacagtggc acccaccata cacaccactg tgccatctcc tactacaaca cctactccaa aggaaaaacc

agaagctgga acctattcag ttaataatgg caatgatact tgtctgctgg ctaccatggg gctgcagctg aacatcactc aggataaggt

tgcttcagtt attaacatca accccaatac aactcactcc acaggcagct gccgttctca cactgctcta cttagactca atagcagcac

cattaagtat ctagactttg tctttgctgt gaaaaatgaa aaccgatttt atctgaagga agtgaacatc agcatgtatt tggttaatgg

ctccgttttc agcattgcaa ataacaatct cagctactgg gatgcccccc tgggaagttc ttatatgtgc aacaaagagc agactgtttc

agtgtctgga gcatttcaga taaatacctt tgatctaagg gttcagcctt tcaatgtgac acaaggaaag tattctacag cccaagagtg

ttcgctggat gatgacacca ttctaatccc aattatagtt ggtgctggtc tttcaggctt gattatcgtt atagtgattg cttacgtaat

tggcagaaga aaaagttatg ctggatatca gactctgtaa cactaatcaa tacgtgatct ctgttacaaa agaaaaaagc aagtacaagt

tccaacatgc aatactggtc aacttaaggt atatttagtt gcagtccagc tctttagaat gggtggtatg ggggatttca aacttaaaca

aaaaactatc aactacaaat tagttgcctg actttggttt ttccaaccaa ggaatttaaa actgttattt ttacagcaaa agatgtgcaa

aatcactgga ttataagttc tattttactg tcttgaatta gtatttcagt gttttcattt tagacattca gactaaaaat acaccgttta

gaaaaaacaa tttttgaaaa agagattttt tttccctgca ggtagttgag ttgaacaaca tgttctaccg tggatttgta cttgctcctt

ttgctctttt tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgattttt gtttgcaggt taacttagct actttggcat tgctgcatat

ttgacctttg agagatataa tagtagattt gaacaggggc tggtattatt atgttcttag caataaatgc ttttctaatg ccttttgaat

acatttgtat ttatgtggct gtaatgacaa aagatacaaa agctttttaa aatttagagt aggtattaat cttattgttt aatctttttt

ttaaaaaaac tggatatttc aatcttttaa attgcaatat ataagactat tccaactggg catttcaatc catttttag gtgctttaga

gataattgct tgccagtgcc aattgagggc attagtactt tgtgctcata aattggcctc tgtatgcagt actaaaatta atgcagattt

ctctttagcc ttccaacatt tcttgttgat agtgatgtat tttattattt tctttttctt aagaaatgcc agtgtgtcct agaacctaga

taacgaagtg cacttacact tataaaataa cttgcatcta ggctgggcgt ggcggctcac gcctgtaatc ccagcacttt

gggaggccga agtgggtgga tcacttgagg ccaggagttt gagaccagcc tggccaacat ggtgaaaccc catctctatc

agaaatacaa aaaattagct gggcatggtg gtgggcgcct gtaatcccag ttactcggga ggctgaggca ggagaatcac

ttgaacccgg gaggcagagg ttgcggtgag ccaagagcgc accattgcac tccagccttg ggcgacaaaa acgaaactcc

atcttcaaaa caaaacaaaa caaaacaaac aaacaaacaa aacttgcatc ttaaccaaaa gtcttggttt tatcttaatc cattaaaagt

tggtcttgtt tccagcttgc attgattgct acaacatcac taatttggct ttcacattta aatggttctg tgctaatcaa aactttcgtt

-continued

gttattattc gttatggtag aatcattttt aattcacgtg ctttgtgttc agttttgtgg tctgagagat gtaccaattg tcaaattacc

gtgtaccacc taatgtttat aggagaaagc aaaatacatc agcttggtag ttaacacatc aaatatttct tgctgcttct aggagaactt

ttttggtgtg tgttggaatg gctgagcaaa tattaaaatt gttaatatgc agccatatat ggaaggttcc tgtggggttg ttttttcgtg

ttttttttt ttgtggtggg attatgtgcc tcccattcac tagaaaatga gaaaattgtc tgggttccaa aatattgaca ttgaatggat

caatacacac acacagacat atatatatat atatgcacac atatataggc agttgcatgc ctagcatggg tattttataa ccatataact

gagttatatt ggaattataa atattttccg tcacttaaat ttgttctttg tttagcctga aaacctttat ggctcaagat cagattcctg

actaacccct ctcttagagc tacagcgagc tgcattacca gcttaaaaca cttcttaggg attaaatata gatgtaattt ttcaaaatcg

tttttaattt aaactgtgtt ttagtgtaaa attgttaacc ttgtaagatg gataatgtgt ataagaatgt aggccttaac tatttcacat

gagtcaaaac aaagcagctt taaaaaaata attggaagca caatgcatgg cactgactga atgctgttaa tatttctaaa agtttctaca

ttcagattat atgcctgatt catagtaaaa tacctctaat aaacactgtt ttatagaaaa cctgacttca gtgaatattt ttgtatttta

catgggccag tttatatact gctatttaca ctattatttc ctatagctac atgttctttg tacctttttgt agttttattt gtattactag

attcatacct

tgatggtaac gctctatctg gttttgggtg tttttcatgt tttagcattt gtataaagaa actggtccat gtaaatactt tccatgtttt

ttcttcaaat gtttaaacca ctagttgatg tatggtatct ttagatattt gcctgtctgt ttgctcaaaa ttgcttctaa aacaataaag

attctt

GenBank Accession No.: AAS67876.1, human LAMP-2C, which provides the amino acid sequence (SEQ ID NO: 25):
MVCFRLFPVPGSGLVLVCLVLGAVRSYALELNLTDSENATCLYAKWQMNFTVRYETTN

KTYKTVTISDHGTVTYNGSICGDDQNGPKIAVQFGPGFSWIANFTKAASTYSIDSVSFSY

NTGDNTTFPDAEDKGILTVDELLAIRIPLNDLFRCNSLSTLEKNDVVQHYWDVLVQAFV

QNGTVSTNEFLCDKDKTSTVAPTIHTTVPSPTTTPTPKEKPEAGTYSVNNGNDTCLLATM

GLQLNITQDKVASVININPNTTHSTGSCRSHTALLRLNSSTIKYLDFVFAVKNENRFYLKE

VNISMYLVNGSVFSIANNNLSYWDAPLGSSYMCNKEQTVSVSGAFQINTFDLRVQPFNV

TQGKYSTAEECSADSDLNFLIPVAVGVALGFLIIVVFISYMIGRRKSRTGYQSV

GenBank Accession No.: AY561849.1, human LAMP-2C mRNA, complete cds, which provides the nucleic acid sequence (SEQ ID NO: 26):
atggtgtgct tccgcctctt cccggttccg ggctcagggc tcgttctggt ctgcctagtc ctgggagctg tgcggtctta tgcattggaa

cttaatttga cagattcaga aaatgccact tgcctttatg caaaatggca gatgaatttc acagtacgct atgaaactac aaataaaact

tataaaactg taaccatttc agaccatggc actgtgacat ataatggaag catttgtggg gatgatcaga atggtcccaa aatagcagtg

cagttcggac ctggcttttc ctggattgcg aattttacca aggcagcatc tacttattca attgacagcg tctcattttc ctacaacact

ggtgataaca caacatttcc tgatgctgaa gataaaggaa ttcttactgt tgatgaactt ttggccatca gaattccatt gaatgacctt

tttagatgca atagtttatc aactttggaa aagaatgatg ttgtccaaca ctactgggat gttcttgtac aagcttttgt ccaaaatggc

acagtgagca caaatgagtt cctgtgtgat aaagacaaaa cttcaacagt ggcacccacc atacacacca ctgtgccatc

tcctactaca acacctactc caaaggaaaa accagaagct ggaacctatt cagttaataa tggcaatgat acttgtctgc tggctaccat

ggggctgcag ctgaacatca ctcaggataa ggttgcttca gttattaaca tcaaccccaa tacaactcac tccacaggca

gctgccgttc tcacactgct ctacttagac tcaatagcag caccattaag tatctagact ttgtctttgc tgtgaaaaat gaaaaccgat

tttatctgaa ggaagtgaac atcagcatgt atttggttaa tggctccgtt ttcagcattg caaataacaa tctcagctac tgggatgccc

ccctgggaag ttcttatatg tgcaacaaag agcagactgt ttcagtgtct ggagcatttc agataaatac ctttgatcta agggttcagc

ctttcaatgt gacacaagga aagtattcta cagctgaaga atgttctgct gactctgacc tcaactttct tattcctgtt gcagtgggtg

tggccttggg cttccttata attgttgtct ttatctctta tatgattgga agaaggaaaa gtcgtactgg ttatcagtct gtgtaa

[0096] In another aspect, the method of treating lysosomal transmembrane protein disease or disorder in a subject includes contacting cells expressing a protein associated with the particular disease or disorder (see Table 1) from the

subject with a vector encoding a gene editing system that when transfected into the cells removes a mutation (e.g., a trinucleotide repeat expansion mutation) of the endogenous protein, thereby treating the lysosomal transmembrane protein disease or disorder. In various embodiments, the gene editing system is selected from the group consisting of CRISPR/Cas, zinc finger nucleases, and transcription activator-life effector nucleases. The step of contacting may be performed ex vivo by first obtaining a sample of cells from the subject, transfecting the gene editing system into the sample of cells, and thereafter transplanting the transfected cells into the subject, thereby treating the lysosomal transmembrane protein disease or disorder. The sample of cells may be any cells expressing the protein associated with the lysosomal transmembrane protein disease or disorder, such as, for example, blood cells or HSPCs of the subject.

[0097] In another aspect, the present invention provides a method of treating or ameliorating a lysosomal protein disease or disorder in a subject. The method includes transplanting a population of HSPCs into the subject, wherein the HSPCs have been genetically modified by introduction of a transgene encoding a corresponding functional human lysosomal transmembrane protein, thereby treating the lysosomal transmembrane protein disease or disorder. Thus, when the lysosomal transmembrane protein disease or disorder is cystinosis, the functional human lysosomal transmembrane gene is CTNS; the lysosomal transmembrane protein disease or disorder is Salla disease or infantile sialic acid storage disorder, the functional human lysosomal transmembrane gene is SLC17A5; the lysosomal transmembrane protein disease or disorder is Cobalamin F-type disease, the functional human lysosomal transmembrane gene is LMBRD1; the lysosomal transmembrane protein disease or disorder is late infantile neuronal ceroid lipofuscinosis, the functional human lysosomal transmembrane gene is MFSD8; the lysosomal transmembrane protein disease or disorder is juvenile neuronal ceroid lipofuscinosis, the functional human lysosomal transmembrane gene is CLN3; the lysosomal transmembrane protein disease or disorder is malignant infantile osteopetrosis, the functional human lysosomal transmembrane gene is CLCN7 or OSTM1; the lysosomal transmembrane protein disease or disorder is mucolipidosis IV, the functional human lysosomal transmembrane gene is MCOLN1; the lysosomal transmembrane protein disease or disorder is mucopolysaccharidosis type IIIC, the functional human lysosomal transmembrane gene is HGSNAT; the lysosomal transmembrane protein disease or disorder is Neiman-Pick Type C, the functional human lysosomal transmembrane gene is NPC1; and the lysosomal transmembrane protein disease or disorder is Danon disease, the functional human lysosomal transmembrane gene is LAMP2. In various embodiments, the HSPCs are isolated from the subject, such as from the bone marrow of the subject.

[0098] While the present invention has been demonstrated with regard to cystinosis and Danon disease, it should be understood that the methods are applicable to any of the diseases or disorders set forth in Table 1. Thus, this strategy turns HSPCs into intelligent and widespread delivery vehicles to obtain stable and sustained cross-correction after their differentiation into monocytes that enter the circulation and subsequently invade the peripheral tissues where they transform into tissue resident macrophages. These macrophages, through a variety of mechanisms including, but not limited to, the formation of tunneling nanotubes, vesicular release, and direct cell-cell adhesion, transfer their lysosomes, which carry the respective protein to diseased peripheral cells. As such, this work demonstrates the development of a HSPC gene therapy strategy for treating lysosomal transmembrane protein diseases or disorders.

[0099] The following examples are intended to illustrate but not limit the invention.

Example 1

Preclinical Model of Cystinosis for Testing Therapeutic Approaches

[0100] Stem cell therapeutic approaches have been tested on the mouse model of cystinosis, the $Ctns^{-/-}$ mice. This murine model was engineered to produce defective cystinosin, and is thus unable to properly transport cystine out of the lysosomes. The defect results in accumulation of cystine and formation of cystine crystals, pathognomonic of cystinosis. Cystine accumulation is present from birth and increases with age. The original $Ctns^{-/-}$ mice have been backcrossed to generate a pure strain of C57BL/6 $Ctns^{-/-}$ mice, which develop renal dysfunction from 6-months of age, as observed biochemically (elevated serum urea and creatinine) and histologically, and these mice are in end stage renal failure by 18 months. Renal Fanconi syndrome also starts around 6 months old (polyuria, phosphaturia and proteinuria), proximal tubular cells appear de-differentiated, and exhibit the typical "swan-neck" deformity found in mice and humans with cystinosis resulting in atubular glomeruli. Finally, heavy infiltration of inflammatory cells can be observed in the kidney of the $Ctns^{-/-}$ mice. $Ctns^{-/-}$ mice also develop ocular defects with corneal cystine crystal depositions and thyroid dysfunction similar to those observed in affected patients.

Example 2

Impact of BMC, HSC and MSC Transplantation on Cystinosis

[0101] To determine the appropriate cell population for transplantation in the context of cystinosis, syngeneic bone marrow cell (BMC), Scar hematopoietic stem cell (HSC) and mesenchymal stem cell (MSC) transplantations were performed in 2-month-old irradiated $Cins^{-/-}$ mice. The cells were isolated from either green fluorescent protein (GFP)-transgenic wild-type (WT) mice or from $Cins^{-/-}$ mice as controls. Analyses of disease parameters were performed 4 months post-transplantation. MSCs had only a short-term limited beneficial impact on the disease. In contrast, tissue cystine content was significantly reduced in all organs tested in the WT BMC and HSC-treated mice (from 57% to 94% decrease depending on tissues). Abundant GFP^{+} bone marrow-derived cells were present in all organs and kidney function was improved. This was the first proof of concept that HSCs could rescue cystinosis even if cystinosin is a transmembrane lysosomal protein as opposed to a secreted enzyme.

Example 3

Long-Term Effect of HSC Transplantation in $Ctns^{-/-}$ Mice

[0102] It was then determined if this treatment was stable for the life of the mice and could result in multi-organ preservation.

[0103] Kidney analysis: Transplantation of WT HSCs was able to provide long-term protection of the kidney function and structure and prevented the progression of the renal disease up to 15 months post-transplantation (last time point tested; FIG. **1**). However, effective therapy depends on achieving a relatively high level of donor-derived blood cell engraftment of Ctns-expressing cells (>50%), which is directly linked to the quantity of Ctns-expressing cells found within the kidney. In contrast, kidney preservation was not dependent on the age of the mice at the time of transplant. Indeed, up to 10-month-old mice could exhibit normal kidney function after stem cell treatment, suggesting that if tissue injury is not consolidated, kidney could be rescued. It was also shown that cystine content was significantly decreased in all tissues (from 54% in the kidney to 96.5% in the liver) proving that the treatment, consisting in a one-time HSC transplantation, led to long-term and stably low levels of tissue cystine for the life span of the mice. Moreover, few to no cystine crystals were observed in all kidneys from treated mice whereas abundant cystine crystals were consistently observed in kidneys from non-treated $Ctns^{-/-}$ mice.

[0104] Eye analysis: GFP+ WT HSC transplantation led to the long-term preservation of the eyes in $Ctns^{-/-}$ mice. Abundant GFP+ bone marrow-derived cells were detected within the cornea but also in the sclera, ciliary body, retina, choroid, and lens in the treated mice. To quantify cystine crystals within the cornea, in vivo confocal microscopy (IVCM) in live mice was performed. Effective therapy was dependent on the level of donor-derived blood cell engraftment as previously demonstrated for the kidney. While $Ctns^{-/-}$ mice with low level of engraftment (<50%; LOW; n=5) presented a partial reduction of crystal counts, the mice with high engraftment levels (>50%; HIGH; n=5) exhibited almost a complete resolution of crystals from the epithelial layer to the middle stroma (100% to 72% clearance, respectively; FIG. **2**). One-year post-transplantation, HSC-treated $Ctns^{-/-}$ mice exhibited normal corneal thickness and structure and normal intraocular pressure. This work was the first demonstration that transplanted HSCs could rescue corneal defects and brings new perspectives for ocular regenerative medicine.

[0105] Thyroid analysis: Since the thyroid gland is also affected in cystinosis, thyroid function and structure from $Cins^{-/-}$ mice and HSC-transplanted mice were analyzed. Sustained thyroid stimulating hormone (TSH) activation combined with morphological evidence for increased thyroglobulin synthesis was shown in $Cins^{-/-}$ mice. Follicular changes included thyrocyte hypertrophy, hyperplasia, colloid exhaustion and vascular proliferation. In contrast, $Cins^{-/-}$ mice treated by HSC transplantation presented virtually normal histology and normalization of cystine and TSH values (FIG. **3**).

[0106] Gastro-intestinal tract analysis: Gastrointestinal mucosal biopsies can be used to measure gene-modified stem cell tissue engraftment and their impact on cystine and cystine crystal levels in subjects enrolled in the HSC gene therapy clinical trial for cystinosis. It has been previously described that a histologic technique for evaluating tissue cystine crystal levels on intestinal mucosal biopsies. It was shown that cystine crystal counts could be correlated with renal function and could help evaluating the response to cysteamine treatment. Thus, a rectal biopsy is planned before and every 6 months after gene-corrected HSC transplantation in the subjects with cystinosis. Up to 9 biopsies can be obtained at a time so it is possible to measure Vector Copy number (VCN), CTNS expression, cystine content and cystine crystals in this tissue at each time point. To establish if this tissue is representative of the efficacy of the treatment, the impact of GFP^+ WT HSC transplantation on the gastro-intestinal tract in $Ctns^{-/-}$ mice was studied. Six months posttransplant, abundant GFP^+ HSC-derived cells were observed in both intestine and colon tissues and cystine content was significantly decreased in treated mice compared to controls in these compartments (FIG. **4**).

[0107] Skin analysis: In vivo confocal microscopy is used on the skin as a noninvasive imaging technology for the visualization and quantification of tissue cystine crystals before and after HSC transplantation in the subjects with cystinosis enrolled in the clinical trial. Chiaverini et al. (Journal of the American Academy of Dermatology 68, el 11 (2013)) showed that this technology was able to detect dermal cystine deposition in patients with cystinosis. For this purpose, a reflectance confocal imager (Caliber VIVASCOPE® 3000) adapted for skin imaging was used to test patients with cystinosis. It was also shown in the HSC-transplanted $Ctns^{-/-}$ mice that abundant GFP+ bone marrow-derived cells engrafted within the skin leading to significant cystine decrease in this tissue (79±0.87 in HSC-treated $Ctns^{-/-}$ vs 193±78 in controls, $p<0.05$).

Example 4

Myeloablative Conditioning Regimen: Efficacy and Toxicity in $Ctns^{-/-}$ Mice

[0108] $Ctns^{-/-}$ mice were exposed to myeloablative drugs currently used in clinic for HSC transplantation, Busulfan (Bu) and Cyclophosphamide (Cy) to test if drug-mediated myeloablation allow efficient engraftment of Ctns-expressing HSC in the preclinical model, decreased tissue cystine and to determine if any unexpected toxicity occurs because of cystinosis. Drugs were injected intraperitoneally (IP) in $Ctns^{-/-}$ mice and WT mice as controls. The mice were analyzed at 4 months post-transplant, demonstrating that: i) $Ctns^{-/-}$ mice did not present any toxicity to Bu or Cy compared to WT controls; ii) Renal function was similar to non-treated age-matched WT controls; iii) Myeloablation was successful in both cases and reached a donor cell engraftment measured in the peripheral blood of 94.2±1.6% for Bu/Cy and 94.0±0.8% for Bu alone; and iv) Treated $Ctns^{-/-}$ mice had a significant decrease of cystine in all tissues tested compared to non-treated. Thus, Bu and Cy are not toxic in the mouse model for cystinosis and drug-mediated myeloablation and HSC transplantation in $Ctns^{-/-}$ mice led to significant decrease of cystine in all tissues.

[0109] The dosing for myeloablation can be done using busulfan alone without cyclophosphamide. Cyclophosphamide does not ablate the hematopoietic stem cells (i.e., does not make engraftment space), and is immune suppressive and anti-leukemic. Since the HSC transplant is autologous, and not for leukemia, cyclophosphamide is not needed, as it adds unnecessary toxicity to the conditioning regimen. Moreover, it is unusual for severe nephrotoxicity to arise directly due to commonly used-conditioning regimen agents such as busulfan. Note that Dr. Donald Kohn's sickle cell trial (ClinicalTrials.gov Identifier: NCT02247843) is only using busulfan, and the sickle cell and thalassemia trials of

bluebird bio, Inc. (ClinicalTrials.gov Identifier: NCT02151526) are also using busulfan without cyclophosphamide.

Example 5

Mechanism of Therapeutic Action

[0110] The extent of efficacy of HSCs to rescue cystinosis was surprising especially considering that the ability of HSC transplantation to rescue non-hematopoietic tissue remains contentious and that cystinosin is a transmembrane lysosomal protein. To elucidate the mechanism of HSC-mediated tissue repair, a novel mouse model was developed, in which $Ctns^{-/-}$ mice back-crossed on a DsRed background so as to ubiquitously express the DsRed reporter gene (Harrison et al., Mol Ther 21, 433 (2013)). When transplanted with GFP-expressing HSCs derived from GFP-transgenic mice, this generated a biofluorescent mouse mode that not only allowed us to track the fate of the transplanted HSCs in an in vivo setting, but also enabled sensitive identification and unequivocal discrimination of events such as fusion, differentiation, and transdifferentiation.

[0111] Using this model, it was first shown that HSCs differentiated into macrophages within tissues (Naphade et al., Stem Cells 33, 301 (2015)). In vitro co-culture experiments were then performed using WT GFP-macrophages and DsRed-Ctns4-fibroblasts. When WT macrophages were co-cultured with $Ctns^{-/-}$ fibroblasts, cystine levels decreased by '75% in FACS-sorted fibroblasts (FIG. **5**A). In contrast, when the two populations were physically separated using a transwell porous to microvesicles, cystine levels decreased only by '20% (FIG. **5**B). These findings showed that cross-correction occurs even if cystinosin is a lysosomal transmembrane protein and that direct cell:cell contact is the main pathway for cross-correction. Using confocal microscopy (FIG. **5**C), it was observed that macrophages extended long membrane protrusions called tunneling nanotubes (TNTs) ('40 μm) that established contact with the fibroblasts. To determine whether TNTs could mediate the physical transfer of cystinosin-bearing vesicles, DsRed-$Ctns^{-/-}$ fibroblasts were co-cultured with macrophages stably transduced with a lentivirus vector expressing cystinosin-GFP fusion protein (CTNS-GFP-macrophages). Live confocal microscopy revealed that vesicles containing cystinosin-GFP could migrate along TNTs towards DsRed-$Ctns^{-/-}$ fibroblasts (FIG. **5**D). LysoTracker staining identified these vesicles as lysosomes (Naphade et al., Stem Cells 33, 301 (2015)).

[0112] Very little is known about TNTs in vivo. It was thus examined whether intercellular vesicular exchange involving nanotubules could be detected in vivo, so as to account for the long-term kidney preservation in $Ctns^{-/-}$ mice. The initial focus was on the kidney not only because of the early occurrence of cystinosis in proximal tubular cells (PTCs) but also because of their physical isolation by the dense tubular basement lamina (TBL). In the two-color grafted mice, GFP^{+} bone marrow-derived cells were observed surrounding but never within the proximal tubules and numerous tubular extensions emanated from the HSC-derived macrophages and crossed the TBL (FIGS. **6**A-a**1** to **6**A-a**3**). GFP-containing structures were observed within PTCs, indicating physical transfer of cytoplasm from the macrophages into the epithelia (FIG. **6**A-a**3**). To test this hypothesis, $Ctns^{-/-}$ mice were transplanted with DsRed-$Ctns^{-/-}$ HSPCs stably expressing cystinosin-GFP fusion protein (FIG. **4**C) or with DsRed-$Ctns^{-/-}$ HSPCs (FIG. **4**B). Many cystinosin-GFP-vesicles were observed in PTCs (FIG. **4**C) (Naphade et al., Stem Cells 33, 301 (2015)). This is the first evidence of direct transfer of proteins from interstitial macrophages to epithelial cells via TNTs penetrating the TBL, so as to correct a genetic defect leading to PTC degeneration. Similar data were obtained for the mechanism of HSC-mediated therapeutic action for the ocular defects (Rocca et al., Investigative ophthalmology & visual science 56, 7214 (2015)) and for the thyroid rescue (Chevronnay et al., Endocrinology In press, (2016)) in the $Ctns^{-/-}$ mice. These findings on HSC-mediated tissue repair bring new perspectives to regenerative medicine, as they should be applicable to other multi-compartment disorders involving deficient intracellular organelles.

Example 6

Clinical Study for Hematopoietic Stem Cell Transplantation

[0113] The work described above represents the first proof-of-concept for using HSC transplantation as a therapy for cystinosis. To minimize the risks of Graft-versus-host (GVHD), subjects are required to have a sibling bone marrow donor who is HLA-matched on 10 of 10 alleles. This study was designed to include six subjects who are either adults ages 18 years and older with significant signs of disease progression or adolescents ages 13-17 years who do not tolerate cysteamine. However, given the rarity of the disease and strict donor requirements, the candidates so far were not complete matches with their sibling. In addition, the risk-benefit ratio for allogeneic HSC transplantation may not be ideal for young patients where the introduction of regular use of the drug cysteamine has permitted patients to live to adulthood, albeit with significant medical problems (Cherqui, Kidney Int 81, 127 (2012)). Indeed, there are significant risks of morbidity and mortality associated with allogeneic transplantation. GVHD is a major complication; in recent studies, acute GVHD grade II-IV occurred in 20% to 32% of patients and chronic GVHD in 16% to 59%, both significantly impacting survival of the recipients (Cutler et al., Blood 109, 3108 (2007); Geyer et al., Br J Haematol 155, 218 (2011); and Schleuning et al., Bone Marrow Transplant 43, 717 (2009)). Thus, a preferred candidate therapy would utilize the patient's own stem cells for an autologous HSC transplantation, thereby mitigating the risks of graft rejection and GVHD.

Example 7

Viral Vector Selection

[0114] Given the risks associated with allogeneic HSC transplantation and considering the preclinical data for HSC gene therapy, transplantation of autologous HSC modified to express function cystinosin represents a safer approach.

[0115] With regard to gene therapy, vectors derived from lentiviruses have supplanted y-retroviral vectors due to their superior gene transfer efficiency and improved biosafety profile (Case et al., Proc Natl Acad Sci USA 96, 2988 (1999); Miyoshi, et al. Science 283, 682 (1999); Naldini et al., Science 272, 263 (1996); and Varma et al., Nature 389, 239 (1997)). Specifically:

[0116] 1. All cases of leukemogenic complications observed to date in clinical trials or animal models of gene therapy involved the use of y-retroviral vectors such as Moloney Leukemia Virus (MLV) retrovirus with long terminal repeats (LTR) containing strong enhancer/promoters that can trigger distant enhancer activation (Hacein-Bey-Abina et al., J Clin Invest 118, 3132 (2008); Li et al., Science 296, 497 (2002)).

[0117] 2. In contrast, the third generation of lentivirus vectors, Self-inactivated (SIN)-lentivirus vectors (LV), with the deletions in their LTR, contain only one internal enhancer/promoter, which reduces the incidence of interactions with nearby cellular genes and thus decreases the risk of oncogenic integration (Modlich et al., Blood 108, 2545 (2006); Montini et al., J Clin Invest 119, 964 (2009)). Moreover, in contrast to the MLV, lentiviruses are not associated with oncogenesis. Importantly, leukemia is not a recognized side effect of HIV patients even though memory T cells are known to carry integrated virus for years.

[0118] 3. SIN-LTR are also designed to prevent the possibility of developing replication competent lentivirus (RCL) during production of the viral supernatants. Indeed, transient transfection systems with three packaging plasmids are usually employed for vector production—gag, pol, and rev (Dull et al., J Virol 72, 8463 (1998)). A fourth plasmid containing the gene coding for the envelope and vesicular stomatitis virus glycoprotein (VSV-G) is frequently used as the choice of envelope. So far RCL has never been reported with this commonly used viral production system in patients after infusion of the vector transduced cell products (Sastry et al., Mol Ther 8, 830 (2003)).

[0119] 4. LV efficiently transduce HSCs and do not alter their repopulation properties (Montini et al., J Clin Invest 119, 964 (2009); Gonzalez-Murillo et al., Blood 112, 3138 (2008)).

[0120] 5. Clinical trials using SIN-LV to transduce human HSCs are being undertaken in the U.S. and Europe for several conditions including HIV-1, 0-thalassemia, immune deficiencies and cancers (DiGiusto et al., Viruses 5, 2898 (2013); Drakopoulou et al., Current molecular medicine 13, 1314 (2013); Porter et al., N Engl J Med 365, 725 (2011); and Zhang et al., Gene Ther 20, 963 (2013)). For immune deficiency disorders, 35 patients have been transplanted with SIN-LV-modified HSCs so far (Bigger et al., Discovery medicine 17, 207 (2014)). A clinical trial using a SIN-LV to correct ex vivo HSCs in patients with X-adrenoleukodystrophy showed that cerebral demyelination was arrested in the two enrolled patients without further progression over 3 years of follow-up; and there was no evidence of clonal dominance (Cartier et al., Methods Enzymol 507, 187 (2012); Cartier et al., Science 326, 818 (2009)). Recently, a clinical trial for Wilskott-Aldrich was reported in three patients 32 months post-transplantation. Stable and long-term engraftment of the gene-modified HSCs (25-50%) resulted in improved platelet counts, protection from bleeding and infections, and resolution of eczema (Aiuti et al., Science 341, 1233151 (2013)). Another clinical success was recently reported in three pre-symptomatic patients with Metachromatic Leukodystrophy. Donor-derived blood cell engraftment of transduced cells achieved 45 to 80% and up to 24 months later the protein activity was reconstituted to above normal values in cerebrospinal fluid associated with a clear therapeutic benefit (Biffi et al., Science 341, 1233158 (2013)).

[0121] pCCL-CTNS lentiviral vector—a third-generation SIN-lentiviral vector in which human CTNS cDNA has been subcloned, pCCL-CTNS (FIG. **7**), was prepared for use. The vector backbone pCCL-EFS-X-WPRE, described by Zufferey et al. (J Virol 72, 9873 (1998)), was provided by Dr. Donald Kohn (UCLA). A central polypurine tract (cPPT) fragment that increases the nuclear import of viral DNA was added to the CCL vector backbone (Demaison et al., Hum Gene Ther 13, 803 (2002)). A Woodchuck hepatitis virus Posttranslational Regulatory Element (WPRE) is present to boost titer and gene expression. However, its open-reading frame was eliminated (Zanta-Boussif et al., Gene Ther 16, 605 (2009)) because it overlapped with the woodchuck hepatitis virus X protein, a transcriptional activator involved in the development of liver tumors (Kingsman et al., Gene Ther 12, 3 (2005)). The transgene expression is driven by the ubiquitously expressed short intron-less human Elongation Factor 1 alpha promoter (EFS, 242 bp) (Wakabayashi-Ito, S. Nagata, J Biol Chem 269, 29831 (1994)). The EFS promoter, which lacks the intron and enhancers of the larger element used in many expression plasmids, has been shown to direct high-level transcription of reporter genes in murine HSCs and to have significantly reduced trans-activation potential compared to y-retroviral LTR (Zychlinski et al., Mol Ther, (2008)).

[0122] Vectors with this backbone are used in clinical trials conducted by Dr. Kohn: i) Autologous Transplantation of Bone Marrow CD34+ Stem/Progenitor Cells after Addition of a Normal Human ADA cDNA by the EFS-ADA Lentiviral Vector for Adenosine Deaminase (ADA)-Deficient Severe Combined Immunodeficiency (SCID) (BB IND 15440; NCT01852071); ii) Autologous Bone Marrow Stem Cells (CD34+) Cultured W/Cytokines; Transduced W/Self-inactivating (SIN) Lentiviral Vector Expressing Human 0-globin (LENTI/BetaAS3-FB); following Busulfan (BB IND 16028; NCT02247843).

Example 8

Preclinical Studies for Transplantation of pCCL-CTN-Transduced HSCs

[0123] $Sca1^{+}$ HSCs isolated from $Ctns^{-/-}$ mice were transduced ex vivo with pCCL-CTNS using our optimized protocol for mHSCs and transplanted into 1- to 4-month-old $Ctns^{-/-}$ mice. Cystine content in brain, eye, heart, kidney, liver, muscle, and spleen were analyzed after 4 (group 1; n=8) and 8 (group 2; n=12) months post-transplantation. As controls, age matched non-treated Ctns−/− mice (n=7 and n=12) were used or $Ctns^{-/-}$ mice transplanted with WT HSCs (n=4 and n=4). Decreases in cystine content were statistically significant in all the tissues tested in mice treated with pCCL-CTNS-transduced HSCs compared to $Ctns^{-/-}$ controls (FIG. **8**A). The impact of Ctns−/−HSCs transduced with control vector, pCCL-GFP, was also tested on tissue cystine levels to exclude the possibility that the presence of any transgene results in cystine decreases. No decrease in any tissue was observed in mice transplanted with pCCL-GFP-$Ctns^{-/-}$ HSCs compared to non-treated $Ctns^{-/-}$ mice (Harrison et al., Mol Ther 21, 433 (2013)).

[0124] Renal glomerular and tubular function was assessed by measuring creatinine, urea, and phosphate levels in the serum, and creatinine clearance in 24-hour urine in males at 8 months post-transplant and compared to age-matched WT males (n=6). All the parameters were increased and the creatinine clearance decreased in non-treated $Ctns^{-/-}$ mice compared to WT mice. In the pCCL-CTNS-treated $Ctns^{-/-}$ mice, serum creatinine, urine phosphate and urine volume were significantly decreased compared to controls, showing a beneficial effect of the genetically modified HSC on kidney function in the $Ctns^{-/-}$ mice. Significant reduction of cystine crystals present in kidney sections was demonstrated in the treated $Ctns^{-/-}$ mice compared controls (FIGS. **8**B and **8**C). Note that we showed that cystine content in female kidneys was five times more elevated then in male kidney in $Ctns^{-/-}$ mice, thus studies on kidney have to be performed on males and females separately (Harrison et al., Mol Ther 21, 433 (2013)).

[0125] Quantitative PCR (qPCR) was performed on genomic DNA isolated from blood collected from pCCL-CTNS-transplanted $Ctns^{-/-}$ mice using lentiviral-specific primers to determine the Vector Copy Numbers per cell (VCN). Average VCN was 1.573±1.868, which fell in the targeted range of VCN 1-3. To determine if lentivirus levels could be predicted in tissues, linear regression analyses was performed between pCCL levels in the different tissues as a function of blood VCN levels. Direct correlation between the levels of lentivirus present in the blood and the levels present in tissues was demonstrated (Harrison et al., Mol Ther 21, 433 (2013)), which is useful to follow the future subjects enrolled in the clinical trial.

Example 9

Pre-Clinical Pharmacology and Toxicology

[0126] The pharmacology/toxicology studies for HSCs ex-vivo gene-modified with pCCL-CTNS are performed using a batch of pCCL-CTNS lentiviral vector preparation produced under comparable-Good Manufacturing Practice (GMPc) obtained from the Indiana University Vector Production Facility (IUVPF), directed by Dr. Kenneth Cornetta. The targeted VCN range that was proposed to the FDA for safety is included between 1 and 3.

[0127] The In Vitro Immortalization (IVIM) assays, a genotoxicity test, was performed by the Translational Trials Development and Support Laboratory at the Cincinnati Children's Hospital Medical Center. This assay consists in mass culture expansion of transduced murine Lin-BMC for 2 weeks followed by culture in 96-well plates at a density of 100 or 10 cells/well for up to 7 weeks (Arumugam et al., Mol Ther 17, 1929 (2009); Modlich et al., Mol Ther 17, 1919 (2009)). The positive wells are counted and the frequency of replating cells calculated and compared to a negative (mock transduced) and positive control (MLV vector). The IVIM assays were performed in triplicates using GMPc pCCL-CTNS preparation with a VCN ranged between 1-3. No immortalized clone was produced with the construct, thus demonstrating an excellent safety profile.

[0128] In vivo pharmacology/toxicology studies are currently being conducted with the analogous cell therapy product consisting in $Sca1^{+}$ mHSCs in the $Ctns^{-/-}$ mice involving serial transplantation. 15-20 $Cins^{-/-}$ mice (10 males and 10 females) were transplanted with pCCL-CTNS-transduced mHSCs (with a VCN 1-3) and 20 with mock-transduced $Ctns^{-/-}$ mHSCs as Primary Recipients. Subsequently, bone marrow cells derived from each of these mice are transplanted into Secondary $Ctns^{-/-}$ mice. The Primary and Secondary mice have to be fully analyzed 6 months post-transplantation by comprehensive molecular, clinical and histological analyses. So far, we have 32 Primary Recipients that reached the 6-month time points: 11 $Ctns^{-/-}$ mice transplanted with pCCL-CTNS-transduced $Ctns^{-/-}$ mHSCs (with VCN included in 1-3) and 21 mock-treated mice, and 18 Secondary mice. No adverse event has been detected so far, the data show efficacy of the product as the weight of the mice treated with pCCL-CTNS-HSCs is higher and the cystine content in the tissues tested is significantly lower than the mock-treated controls (FIGS. **9**A-**9**B). Thus, we have to reach the 6 month-time point for up to 9 additional $Ctns^{-/-}$ Primary Recipients treated with pCCL-CTNS-transduced HSCs and for 15 Secondary Recipients.

Example 10

Manufacturing: Process Development

[0129] Using the GMPc pCCL-CTNS preparation, a protocol was optimized to transduce human CD34+ HSCs from healthy donors to obtain a VCN included between 1 and 3. This protocol involved a one-hit vector transduction at a MOI 20 for 20 hours. Colony Forming Unit (CFU) assays were then performed using human CD34+ peripheral blood stem cell (PBSC) isolated from five healthy donors and four cystinosis patients and neither showed aberrant proliferation or differentiation potential with pCCL-CTNS LV compared to mock-transduced controls. Moreover, Vector Integration Site (VIS) analyses in the patient's cells showed no enrichment of the integration sites near proto-oncogene 5' ends. However, while this protocol led to an average VCN of 2 in healthy CD34+ cells, the average VCN in cystinosis patients was Therefore, the protocol was further optimized with the cystinosis patients' cells to achieve a higher level of transduction as we have demonstrated that a higher level of cells expressing CTNS leads to a better therapeutic response (Yeagy et al., Kidney Int 79, 1198 (2011); Rocca et al., Investigative ophthalmology & visual science 56, 7214 (2015); Harrison et al., Mol Ther 21, 433 (2013)). This protocol involves a two-hit vector transduction at a MOI 20 each for 24-hour total and an average VCN of 1.9 with patients' cells was obtained. CFU assays and VIS have now to be repeated with this new protocol.

[0130] For the clinical trial, the transduction protocol is performed according to the GMP facility's standard operating procedures and uses the optimal protocol for cystinosis patients' CD34+ cells. Note that prior to enrolling the first patient, optimal conditions for large-scale transductions using the GMP-grade pCCL-CTNS vector preparation and optimal protocol are validated in small scales and Proficiency Runs using human CD34+ cells from healthy donors at the GMP facility. The clinical trial will include six patients affected with cystinosis, four adults and two adolescents. This will be a first-in-human clinical trial for an autologous stem cell and gene therapy treatment strategy for cystinosis. If successful, this treatment could be a life-long therapy that may eliminate or reduce renal deterioration and the need for kidney transplantation, as well as, the long-term complications associated with cystinosis. Additionally, the mechanism by which transplantation of pCCL-CTNS-modified

CD34+ HSCs provide beneficial and protective effects may be applicable to other inherited multi-organ degenerative disorders.

Example 11

HSPC Transplantation for Danon Disease

[0131] The objective of this experiment is to determine whether Danon disease can be rescued by HSPC transplantation and to determine whether lysosomal cross-correction occurs. Using the mouse model described herein, it has been demonstrated that the hearts of Lamp2 KO mice exhibit increased numbers of abnormal mitochondria, and impairments in mitophagy and mitochondrial respiration, which is consistent with prior studies in induced pluripotent stem cell (hiPSC)-derived cardiac myocytes from Danon patients (Cherqui, *Kidney Int* 81, 127 (2012)), confirming similarities between the mouse model and human disease.

[0132] To evaluate the ability of WT HSPCs to rescue Danon disease, two-month-old lethally irradiated Lamp2 KO mice were transplanted with $Sca1^+$ HSPCs isolated from congenic C57BL/6 WT that ubiquitously expressed cytoplasmic eGFP (Tg(ACTB-EGFP)1Osb/J from Jackson Laboratory) (WT-HSPCs) using transplantation previously described protocols (Yeagy et al., Kidney Int 79, 1198 (2011); Naphade et al., Stem Cells 33, 301 (2015); Case et al., Proc Natl Acad Sci USA 96, 2988 (1999)). As negative controls, Lamp2 KO mice were also transplanted with $Sca1^+$ HSPCs from Lamp2 KO mice that constitutively expressed eGFP (KO-HSPCs). Skeletal muscle strength was assayed using previously described techniques and demonstrated significantly decreased grip strength in Lamp2 KO mice in comparison to both WT and Lamp2 mice that had received WT HSPCs (FIG. **11**). LAMP2 protein expression, as assessed by Western blot analyses, was restored to near WT levels in hearts and skeletal muscle of Lamp2 KO mice transplanted with WT HSPCs (FIGS. **10**D-**10**E).

[0133] To demonstrate that LAMP2 was expressed within cardiac myocytes of recipient Lamp2 KO mice, and not just within donor macrophages residing in those hearts, immunofluorescence studies were performed that demonstrated LAMP2+ vesicles in cardiomyocytes (α-actinin: white) located adjacent to donor macrophages (FIGS. **10**A-**10**C). EM analyses showed decreased vacuoles in Lamp2 KO mice that received WT HSPCs in comparison to Lamp2 KO Mice (FIGS. **11**A and **11**B), exhibiting an appearance similar to that of WT mice. Improved autophagic flux following WT HSPC transplant of Lamp2 KO mice was confirmed by assessing LC3-II/GAPDH levels (FIGS. **11**C-**11**D). In summary, these studies demonstrate the restoration of physiologic and metabolic function in Lamp2 KO mice treated with WT HSPC transplantation.

[0134] Although the invention has been described with reference to the above examples, it will be understood that modifications and variations are encompassed within the spirit and scope of the invention. Accordingly, the invention is limited only by the following claims.

SEQUENCE LISTING

Sequence total quantity: 27
SEQ ID NO: 1 moltype = AA length = 367
FEATURE Location/Qualifiers
source 1..367
mol_type = protein
organism = Homo sapiens
SEQUENCE: 1
MIRNWLTIFI LFPLKLVEKC ESSVSLTVPP VVKLENGSST NVSLTLRPPL NATLVITFEI 60
TFRSKNITIL ELPDEVVVPP GVTNSSFQVT SQNVGQLTVY LHGNHSNQTG PRIRFLVIRS 120
SAISIINQVI GWIYFVAWSI SFYPQVIMNW RRKSVIGLSF DFVALNLTGF VAYSVFNIGL 180
LWVPYIKEQF LLKYPNGVNP VNSNDVFFSL HAVVLTLIII VQCCLYERGG QRVSWPAIGF 240
LVLAWLFAFV TMIVAAVGVI TWLQFLFCFS YIKLAVTLVK YFPQAYMNFY YKSTEGWSIG 300
NVLLDFTGGS FSLLQMFLQS YNNDQWTLIF GDPTKFGLGV FSIVFDVVFF IQHFCLYRKR 360
PGYDQLN 367

SEQ ID NO: 2 moltype = DNA length = 2611
FEATURE Location/Qualifiers
source 1..2611
mol_type = other DNA
organism = Homo sapiens
SEQUENCE: 2
cgcctctccc aaagtctagc cgggcagggg aacgcggtgc attcctgacc ggcacctggc 60
gaggctcatg cgtcccgtga gggcggttcc tcgagcctgg gggcgctcag attgctttgg 120
agacgctgag agaacctttg cgagagcgcc ggttgacgtg cggagtgcgg ggctccgggg 180
gactgagcag cacgagaccc catcctcccc tccgggtttt cacactgggc gaagggagga 240
ctcctgagct ctgcctcttc cagtaacatt gaggattact gtgttttgtg agagctcgct 300
aggcgcccta agcaacagag ttctgagaaa tcgagaaaca tgataaggaa ttggctgact 360
atttttatcc ttttttcccct gaagctcgta gagaaatgtg agtcaagcgt cagcctcact 420
gttcctcctg tcgtaaagct ggagaacggc agctcgacca acgtcagcct caccctgcgg 480
ccaccattaa atgcaaccct ggtgatcact tttgaaatca catttcgttc caaaaatatt 540
actatccttg agctccccga tgaagttgtg gtgcctcctg gagtgacaaa ctcctctttt 600
caagtgacat ctcaaaatgt tggacaactt actgtttatc tacatggaaa tcactccaat 660
cagaccggcc cgaggatacg ctttcttgtg atccgcagca gcgccattag catcataaac 720
caggtgattg gctggatcta ctttgtggcc tggtccatct ccttctaccc tcaggtgatc 780
atgaattgga ggcggaaaag tgtcattggt ctgagcttcg acttcgtggc tctgaacctg 840
acaggcttcg tggcctacag tgtattcaac atcggcctcc tctgggtgcc ctacatcaag 900
gagcagtttc tcctcaaata ccccaacgga gtgaaccccg tgaacagcaa cgacgtcttc 960
ttcagcctgc acgcggttgt cctcacgctg atcatcatcg tgcagtgctg cctgtatgag 1020

-continued

```
cgcggtggcc agcgcgtgtc ctggcctgcc atcggcttcc tggtgctcgc gtggctcttc  1080
gcatttgtca ccatgatcgt ggctgcagtg ggagtgatca cgtggctgca gtttctcttc  1140
tgcttctcct acatcaagct cgcagtcacg ctggtcaagt attttccaca ggcctacatg  1200
aacttttact acaaaagcac tgagggctgg agcattggca acgtgctcct ggacttcacc  1260
gggggcagct tcagcctcct gcagatgttc ctccagtcct acaacaacga ccagtggacg  1320
ctgatcttcg gagacccaac caagtttgga ctcggggtct tctccatcgt cttcgacgtc  1380
gtcttcttca tccagcactt ctgtttgtac agaaagagac cggggtatga ccagctgaac  1440
tagcacccag ggacccagtg tacccagcct ctggcctcgt gccctgctgg ggaaggcctc  1500
acccagcgaa ggccggagaa gcggttgggc cctggcacac agggctggct cagtgtgcgg  1560
acagaggaga ccactctgct cctggggcca gaggccattc aatagcctgc cttcgtccgg  1620
gcccctcctg ggcctccccg gccaggcacg tggcaccgtc gccttgacac cgccatctct  1680
tttctttaag gcttcaggca gcgcgcacag gctctggcag ccgtctcagg caggactggg  1740
caccaagctt gcagccgaag gccttgcccc aaactaccag cgtttctgca agcagcttga  1800
agggctgacc ttgcagccgg gtgagccaag ggcactttgc tgccaccgct gcattcccag  1860
agatcaagca gcccggtgcc gtggccagtg aactcagagg tgctggtgga cgggctagga  1920
ctttggggtt aggccatggg gctctttctc tgaaggccac tttcctgacg tactctctgt  1980
acataactca gcgtccgtga ctgcagtaac agccagccct acccagagta tttctgagcc  2040
atgaggggcc caccagattg gttctgaatt ggattcatgc ccagcgcatt agcatagtaa  2100
ctcctttcag atttttgga gggacgtttg gaagtggctt actctcttct gccctctctc  2160
ctacctccac cttctcagat gagccccatc tgagcacatc cagctgctcc ttacccagca  2220
tctggagtac aggacatagc tctctcctgc taccagtctg tgccttagag gtcgttaggc  2280
ctgccaaacg gcgaccagct cccctggagc gagggcaggc cccttccctc tctttcccca  2340
gacacctact tgagactcac caatttctgg cctgttcagg agcctcagat aagtatttgt  2400
acttgagacc acctcacaca atctgtatgg gcccaaccct gatctcaaac ctccttccct  2460
ctgcccaaag ctgtccttcc tatggcagga ggggtggggg tcccaggacg tgcctcatac  2520
atgacttgag cttgtcagtc cactgagttt ccttctacga gatcaacgcg aggggcctgt  2580
atcttgaatt aaagcctact cgcttccttt c                                  2611

SEQ ID NO: 3           moltype = AA  length = 495
FEATURE                Location/Qualifiers
source                 1..495
                       mol_type = protein
                       organism = Homo sapiens
SEQUENCE: 3
MRSPVRDLAR NDGEESTDRT PLLPGAPRAE AAPVCCSARY NLAILAFFGF FIVYALRVNL  60
SVALVDMVDS NTTLEDNRTS KACPEHSAPI KVHHNQTGKK YQWDAETQGW ILGSFFYGYI  120
ITQIPGGYVA SKIGGKMLLG FGILGTAVLT LFTPIAADLG VGPLIVLRAL EGLGEGVTFP  180
AMHAMWSSWA PPLERSKLLS ISYAGAQLGT VISLPLSGII CYYMNWTYVF YFFGTIGIFW  240
FLLWIWLVSD TPQKHKRISH YEKEYILSSL RNQLSSQKSV PWVPILKSLP LWAIVVAHFS  300
YNWTFYTLLT LLPTYMKEIL RFNVQENGFL SSLPYLGSWL CMILSGQAAD NLRAKWNFST  360
LCVRRIFSLI GMIGPAVFLV AAGFIGCDYS LAVAFLTIST TLGGFCSSGF SINHLDIAPS  420
YAGILLGITN TFATIPGMVG PVIAKSLTPD NTVGEWQTVF YIAAAINVFG AIFFTLFAKG  480
EVQNWALNDH HGHRH                                                   495

SEQ ID NO: 4           moltype = DNA  length = 2512
FEATURE                Location/Qualifiers
source                 1..2512
                       mol_type = other DNA
                       organism = Homo sapiens
SEQUENCE: 4
cggctacttt gcgccaatcc tacgagaact cccagaactc cgcttcccta gtccaaccca  60
agccagagtt gcccacacct aagatggcgg cgggggggcgg agtcggcgcg gccgcctctg  120
ggcgggaccg cggggactag acgtggccgc ggggcggtgt catcgccccc gccccgcccg  180
gtccagccag ctcggcccgg gggcttcggg ctgtcgggcc ggcgctccct tctctgccag  240
gtggcgagta cacctgctca cgtaggcgtc atgaggtctc cggttcgaga cctggcccgg  300
aacgatggcg aggagagcac ggaccgcacg cctcttctac cgggcgcccc acgggccgaa  360
gccgctccag tgtgctgctc tgctcgttac aacttagcaa ttttggcctt ttttggtttc  420
ttcattgtgt atgcattacg tgtgaatctg agtgttgcgt tagtggatat ggtagattca  480
aatacaactt tagaagataa tagaacttcc aaggcgtgtc cagagcattc tgctcccata  540
aaagttcatc ataatcaaac gggtaagaag taccaatggg atgcagaaac tcaaggatgg  600
attctcggtt ccttttttta tggctacatc atcacacaga ttcctggagg atatgttgcc  660
agcaaaatag gggggaaaat gctgctagga tttgggatcc ttggcactgc tgtcctcacc  720
ctgttcactc ccattgctgc agatttagga gttggaccac tcattgtact cagagcacta  780
gaaggactag gagagggtgt tacatttcca gccatgcatg ccatgtggtc ttcttgggct  840
ccccctcttg aaagaagcaa acttcttagc atttcatatg caggagcaca gcttgggaca  900
gtaatttctc ttcctctttc tggaataatt tgctactata tgaattggac ttatgtcttc  960
tacttttttg gtactattgg aatattttgg tttcttttgt ggatctggtt agttagtgac  1020
acaccacaaa aacacaagag aatttcccat tatgaaaagg aatacattct ttcatcatta  1080
agaaatcagc tttcttcaca gaagtcagtg ccgtgggtac ccattttaaa atccctgcca  1140
ctttgggcta tcgtagttgc acacttttct tacaactgga ctttttatac tttattgaca  1200
ttattgccta cttatatgaa ggagatccta aggttcaatg ttcaagagaa tgggttttta  1260
tcttcattgc cttatttagg ctcttggtta tgtatgatcc tgtctggtca agctgctgac  1320
aatttaaggg caaaatggaa tttttcaact ttatgtgttc gcagaatttt tagccttata  1380
ggaatgattg gacctgcagt attcctggta gctgctggct tcattggctg tgattattct  1440
ttggccgttg ctttcctaac tatatcaaca acactgggag gcttttgctc ttctggattt  1500
agcatcaacc atctggatat tgctccttcg tatgctggta tcctcctggg catcacaaat  1560
```

-continued

acatttgcca ctattccagg aatggttggg cccgtcattg ctaaaagtct gacccctgat 1620
aacactgttg gagaatggca aaccgtgttc tatattgctg ctgctattaa tgtttttggt 1680
gccatttttct ttacactatt cgccaaaggt gaagtacaaa actgggctct caatgatcac 1740
catggacaca gacactgaag gaaccaataa ataatcctgc ctctattaat gtatttttat 1800
ttatcatgta acctcaaagt gccttctgta ttgtgtaagc attctatgtc ttttttaat 1860
tgtacttgta ttagattttt aaggcctata atcatgaaat atcactagtt gccagaataa 1920
taaaatgaac tgtgtttaat tatgaataat atgtaagcta ggacttctac tttaggttca 1980
catacctgcc tgctagtcgg gcaacatgaa gtaggacagt tctgttgatt ttttagggcc 2040
atactaaagg gaatgagctg aaacagacct cctgatacct ttgcttaatt aaactagatg 2100
ataattctca ggtactgata aacacctgtt gttgttcact ttcctcataa aaattgtcag 2160
ctctctctga cacttagacc tcaaacttta gcatctctgt ggagctgcca tccactgtat 2220
aatttcgcct ggcaactgga ctgaggggag tgtgcccagg cagctgccaa gcactccctc 2280
cctggcttca gggtcagagt gcccagcgtt tatcagaggc agcatccaag cccagagcca 2340
gtgtcgactc ttcggctggt gcctttcctc tgaggggcta tcaatgtgta gataaagccc 2400
tgagtaggca agagcagtga gatccactgc tatggtcttg atacatcctc aaactttccc 2460
ttcccagcac agaggaatat tggctggcat gcaacctgca aaagaaaaat gc 2512

SEQ ID NO: 5 moltype = AA length = 537
FEATURE Location/Qualifiers
source 1..537
mol_type = protein
organism = Homo sapiens
SEQUENCE: 5
GAASAELVIG WCIFGPLLLA IFAFCWIYVR KYQSQRESEV VSTITAIFSL AIALITSALL 60
PVDIFLVSYM KNQNGTFKDW ANANVSRQIE DTVLYGYYTL YSVILFCVFF WIPFVYFYYE 120
EKDDDDTSKC TQIKTAFKYT LGFAVICALL LLVGAFVPLN VPNNKNSTEW EKVKFLFEEL 180
GSSHGLAALS FSISSLTLIG MLAAITYTAY GMSALPLNLI KGTRSAAYER LENTEDIEEV 240
EQHIQTIKSK SKDGRPLPAR DKRALKQFEE RLRTLRKRER HLEYIENSWW TKFCGALRPL 300
KIIWGIFFIL VALLFIISLF LSNLDKALHS AGIDSGFIIF GANLSNPLNM LLPVLQTVFP 360
LDYILITIII MYFIFTSMAG IRNIGIWFFW VRLYKIRRGR TRPQALLFLC MILLLIVLHT 420
SYMIYSLAPQ YVMYGSQNYL IESNITYDDH KNNSAFPVPK RCDADAPEDQ CTVTRTYLFL 480
HKFWFFSAAY YFGNWAFLVV FLIGLIVSCC KGKKSVIEGV DEDDSDISDD EPSVYSV 537

SEQ ID NO: 6 moltype = DNA length = 1979
FEATURE Location/Qualifiers
source 1..1979
mol_type = other DNA
organism = Homo sapiens
SEQUENCE: 6
ggcgcggctt cggcggagct ggtgatcggc tggtgcatat ttggcccctt actactggct 60
atttttgcat tctgttggat atatgttcgt aaataccaaa gtcagcggga aagtgaagtt 120
gtctccacca taacggcaat ttttttctctg gcgattgcac ttatcacatc agcacttctt 180
ccagtggata tatttttggt ttcttacatg aaaaatcaaa atggtacatt taaggactgg 240
gccaatgcta atgtcagcag acagatcgag gacactgtgt tatatggtta ctacacctta 300
tattctgtta tattattctg tgtgtttttc tggatccctt ttgtctactt ctactatgaa 360
gaaaaggatg atgatgatac tagtaaatgt actcaaatta aaactgcatt caagtatact 420
ttgggatttg ctgtaatttg tgcacttctt cttttagttg gagcttttgt tcctctaaat 480
gttcctaata acaaaaattc tacagagtgg gaaaaagtga agttcctgtt tgaagaactt 540
ggaagtagtc atggtttagc tgcattgtca ttttctatta gttctctgac cttgattgga 600
atgttggcag ctataactta cacagcctat ggcatgtctg cattaccttt aaatctaata 660
aaaggcacta gaagcgctgc ttacgaacgt ttagaaaaca ctgaagacat tgaagaagtg 720
gagcaacaca ttcaaacgat taaatcaaaa agcaaagatg gtcggccttt gccagcaagg 780
gataaacgcg ccttaaaaca atttgaagaa aggttaagaa cacttaggaa aagagagagg 840
cacttagaat acattgaaaa cagctggtgg acaaaatttt gtggtgctct gcgtcccctg 900
aagatcattt ggggaatatt tttcatctta gttgcattgc tgtttataat ttctctcttc 960
ctgtcaaatt tggataaagc ccttcattca gctggaatag attctggttt tataattttt 1020
ggagctaact tgagtaatcc actgaatatg cttttgcctg tactacaaac agtgtttcct 1080
cttgattata ttcttataac aattattatt atgtacttta tttttacttc aatggcggga 1140
attcgaaata tcggcatatg gttcttttgg gttagactat ataaaattag aagaggtaga 1200
accaggcccc aggccctctt atttctttgc atgatacttc tgcttattgt ccttcacact 1260
agctacatga tttatagtct tgctccccaa tatgtcatgt atggaagcca aaattactta 1320
atagagagca atataactta tgatgaccat aaaaacaatt cagccttccc tgtgccaaag 1380
agatgtgatg ctgatgcccc tgaagaccaa tgtactgtta cgcggacata cctgttcctt 1440
cacaagttct ggttctttag tgctgcatac tattttggta actgggcttt tcttgtggta 1500
ttcttgattg gattaattgt atcctgttgt aaagggaaga aatcagtcat tgaaggagta 1560
gatgaagatg attcagacat aagtgatgat gagccctctg tctattctgt ttgagagcct 1620
ctgtcttagg ggttttataa tgctgactga atgtctatta tgcatttttt aaagtgttaa 1680
actaacatta ggatgaactg actagcttca tcaaaaatgg gagcatggct attaaaaaaa 1740
ctatattttt tatgttatct gaagtaacat tattgtatca tagattaaca tttaaaattg 1800
ctgtaataat tctatgtaaa tataaaacta tggactttgt gagggaatgt ttgtggaaat 1860
ctttttctc tagtgtataa tagtgttgaa ttgattaaaa gtcttccaga attaatattc 1920
cctcttgtca cttcttaaaa acataataaa tcacttctac ctgtgcaaaa aaaaaaaaa 1979

SEQ ID NO: 7 moltype = AA length = 518
FEATURE Location/Qualifiers
source 1..518

-continued

```
                     mol_type = protein
                     organism = Homo sapiens
SEQUENCE: 7
MAGLRNESEQ EPLLGDTPGS REWDILETEE HYKSRWRSIR ILYLTMFLSS VGFSVVMMSI  60
WPYLQKIDPT ADTSFLGWVI ASYSLGQMVA SPIFGLWSNY RPRKEPLIVS ILISVAANCL  120
YAYLHIPASH NKYYMLVARG LLGIGAGNVA VVRSYTAGAT SLQERTSSMA NISMCQALGF  180
ILGPVFQTCF TFLGEKGVTW DVIKLQINMY TTPVLLSAFL GILNIILILA ILREHRVDDS  240
GRQCKSINFE EASTDEAQVP QGNIDQVAVV AINVLFFVTL FIFALFETII TPLTMDMYAW  300
TQEQAVLYNG IILAALGVEA VVIFLGVKLL SKKIGERAIL LGGLIVVWVG FFILLPWGNQ  360
FPKIQWEDLH NNSIPNTTFG EIIIGLWKSP MEDDNERPTG CSIEQAWCLY TPVIHLAQFL  420
TSAVLIGLGY PVCNLMSYTL YSKILGPKPQ GVYMGWLTAS GSGARILGPM FISQVYAHWG  480
PRWAFSLVCG IIVLTITLLG VVYKRLIALS VRYGRIQE                          518

SEQ ID NO: 8          moltype = DNA  length = 1909
FEATURE               Location/Qualifiers
source                1..1909
                      mol_type = other DNA
                      organism = Homo sapiens
SEQUENCE: 8
aggttacaag cagcagatcc caccttcagt cctggctctg acaagccctc cagcttcacg  60
ccacccggga tgggagaaag caggtgtcgc gagagttggg cgcaagacgc cttgtaggga  120
gtgtaactat ggccggcctg cggaacgaaa gtgaacagga gccgctctta ggcgacacac  180
ctggaagcag agaatgggac attttagaga ctgaagagca ttataagagc cgatggagat  240
ctattaggat tttatatctt actatgtttc tcagcagtgt agggttttct gtagtgatga  300
tgtccatatg gccatatctc caaaagattg atccgacagc tgatacaagt tttttgggct  360
gggttattgc ttcatatagt cttggccaaa tggtagcttc acctatattt ggtttatggt  420
ctaattatag accaagaaaa gagcctctta ttgtctccat cttgatttcc gtggcagcca  480
actgcctcta tgcatatctc cacatcccag cttctcataa taaatactac atgctggttg  540
ctcgtggatt gttgggaatt ggagcaggaa atgtagcagt tgttagatca tatactgctg  600
gtgctacttc ccttcaggaa agaacaagtt ccatggcaaa cataagcatg tgtcaagcat  660
taggttttat tctaggtcca gtttttcaga cttgttttac attccttgga gaaaaaggtg  720
tgacatggga tgtgattaaa ctgcagataa acatgtatac aacaccagtt ttacttagcg  780
ccttcctggg aattttaaat attattctga tccttgccat actaagagaa catcgtgtgg  840
atgactcagg aagacagtgt aaaagtatta attttgaaga agcaagtaca gatgaagctc  900
aggttcccca aggaaatatt gaccaggttg ctgttgtggc catcaatgtt ctgtttttg  960
tgactctatt tatctttgcc ctttttgaaa ccatcattac tccattaaca atggatatgt  1020
atgcctggac tcaagaacaa gctgtgttat ataatggcat aatacttgct gctcttgggg  1080
ttgaagccgt tgttattttc ttaggagtta agttgctttc caaaaagatt ggcgagcgtg  1140
ctattctact gggaggactc atcgttgtat gggttggctt ctttatcttg ttaccttggg  1200
gaaatcaatt tcccaaaata cagtgggaag atttgcacaa taattcaatc cctaatacca  1260
catttgggga aattattatt ggtctttgga agtctccaat ggaagatgac aatgaaagac  1320
caactggttg ctcgattgaa caagcctggt gcctctacac cccggtgatt catctggccc  1380
agttccttac atcagctgtg ctaataggat taggctatcc agtctgcaat cttatgtcct  1440
atactctata ttcaaaaatt ctaggaccaa aacctcaggg tgtatacatg ggctggttaa  1500
cagcatctgg aagtggagcc cggattcttg ggcctatgtt catcagccaa gtgtatgctc  1560
actggggacc acgatgggca ttcagcctgg tgtgtggaat aatagtgctc accatcaccc  1620
tcctgggagt ggtttacaaa agactcattg ctctttctgt aagatatggg aggattcagg  1680
aataaactag ctaagactgt gatggaaact acttgctgtg tggcacttcc tggtctaaag  1740
ctctgctaga caattgcggt gagccagtct ccaagaatca gactacagat attgcagatt  1800
ttgaagaaca agaacatatg ttgaataaca gagagaattc tacatgtcat tgtgaatagt  1860
aggttatata aaaacatact agatgataat ttcaaaaaaa aaaaaaaaa              1909

SEQ ID NO: 9          moltype = AA  length = 438
FEATURE               Location/Qualifiers
source                1..438
                      mol_type = protein
                      organism = Homo sapiens
SEQUENCE: 9
MGGCAGSRRR FSDSEGEETV PEPRLPLLDH QGAHWKNAVG FWLLGLCNNF SYVVMLSAAH  60
DILSHKRTSG NQSHVDPGPT PIPHNSSSRF DCNSVSTAAV LLADILPTLV IKLLAPLGLH  120
LLPYSPRVLV SGICAAGSFV LVAFSHSVGT SLCGVVFASI SSGLGEVTFL SLTAFYPRAV  180
ISWWSSGTGG AGLLGALSYL GLTQAGLSPQ QTLLSMLGIP ALLLASYFLL LTSPEAQDPG  240
GEEEAESAAR QPLIRTEAPE SKPGSSSSLS LRERWTVFKG LLWYIVPLVV VYFAEYFINQ  300
GLFELLFFWN TSLSHAQQYR WYQMLYQAGV FASRSSLRCC RIRFTWALAL LQCLNLVFLL  360
ADVWFGFLPS IYLVFLIILY EGLLGGAAYV NTFHNIALET SDEHREFAMA ATCISDTLGI  420
SLSGLLALPL HDFLCQLS                                                438

SEQ ID NO: 10         moltype = DNA  length = 1689
FEATURE               Location/Qualifiers
source                1..1689
                      mol_type = other DNA
                      organism = Homo sapiens
SEQUENCE: 10
cccctagaca agccggagct gggaccggca atcgggcgtt gatccttgtc acctgtcgca  60
gaccctcatc cctcccgtgg gagccccctt tggacactct atgaccctgg accctcgggg  120
gacctgaact tgatgcgatg ggaggctgtg caggctcgcg gcggcgcttt tcggattccg  180
```

-continued

```
aggggagga gaccgtcccg gagccccggc tccctctgtt ggaccatcag ggcgcgcatt  240
ggaagaacgc ggtgggcttc tggctgctgg gcctttgcaa caacttctct tatgtggtga  300
tgctgagtgc cgcccacgac atccttagcc acaagaggac atcgggaaac cagagccatg  360
tggacccagg cccaacgccg atcccccaca acagctcatc acgatttgac tgcaactctg  420
tctctacggc tgctgtgctc ctggcggaca tcctccccac actcgtcatc aaattgttgg  480
ctcctcttgg ccttcacctg ctgccctaca gcccccgggt tctcgtcagt gggatttgtg  540
ctgctggaag cttcgtcctg gttgcctttt ctcattctgt ggggaccagc ctgtgtggtg  600
tggtcttcgc tagcatctca tcaggccttg gggaggtcac cttcctctcc ctcactgcct  660
tctaccccag ggccgtgatc tcctggtggt cctcagggac tgggggagct gggctgctgg  720
gggccctgtc ctacctgggc ctcacccagg ccggcctctc ccctcagcag accctgctgt  780
ccatgctggg tatccctgcc ctgctgctgg ccagctattt cttgttgctc acatctcctg  840
aggcccagga ccctggaggg gaagaagaag cagagagcgc agcccggcag cccctcataa  900
gaaccgaggc cccggagtcg aagccaggct ccagctccag cctctccctt cgggaaaggt  960
ggacagtatt caagggtctg ctgtggtaca ttgttccctt ggtcgtagtt tactttgccg  1020
agtatttcat taaccaggga ctttttgaac tcctcttttt ctggaacact tccctgagtc  1080
acgctcagca ataccgctgg taccagatgc tgtaccaggc tggcgtcttt gcctcccgct  1140
cttctctccg ctgctgtcgc atccgtttca cctgggccct ggccctgctg cagtgcctca  1200
acctggtgtt cctgctggca gacgtgtggt tcggctttct gccaagcatc tacctcgtct  1260
tcctgatcat tctgtatgag gggctcctgg gaggcgcagc ctacgtgaac accttccaca  1320
acatcgccct ggagaccagt gatgagcacc gggagtttgc aatggcggcc acctgcatct  1380
ctgacacact ggggatctcc ctgtcggggc tcctggcttt gcctctgcat gacttcctct  1440
gccagctctc ctgatactcg ggatcctcag gacgcaggtc acattcacct gtgggcagag  1500
ggacaggtca gacacccagg cccaccccag agaccctcca tgaactgtgc tcccagcctt  1560
cccggcaggt ctgggagtag ggaagggctg aagccttgtt tccttgcagg ggggccagcc  1620
attgtctccc acttggggag tttcttcctg gcatcatgcc ttctgaataa atgccgattt  1680
tgtccatgg                                                          1689

SEQ ID NO: 11          moltype = AA  length = 805
FEATURE                Location/Qualifiers
source                 1..805
                       mol_type = protein
                       organism = Homo sapiens
SEQUENCE: 11
MANVSKKVSW SGRDRDDEEA APLLRRTARP GGGTPLLNGA GPGAARQSPR SALFRVGHMS  60
SVELDDELLD PDMDPPHPFP KEIPHNEKLL SLKYESLDYD NSENQLFLEE ERRINHTAFR  120
TVEIKRWVIC ALIGILTGLV ACFIDIVVEN LAGLKYRVIK GNIDKFTEKG GLSFSLLLWA  180
TLNAAFVLVG SVIVAFIEPV AAGSGIPQIK CFLNGVKIPH VVRLKTLVIK VSGVILSVVG  240
GLAVGKEGPM IHSGSVIAAG ISQGRSTSLK RDFKIFEYFR RDTEKRDFVS AGAAAGVSAA  300
FGAPVGGVLF SLEEGASFWN QFLTWRIFFA SMISTFTLNF VLSIYHGNMW DLSSPGLINF  360
GRFDSEKMAY TIHEIPVFIA MGVVGGVLGA VFNALNYWLT MFRIRYIHRP CLQVIEAVLV  420
AAVTATVAFV LIYSSRDCQP LQGGSMSYPL QLFCADGEYN SMAAAFFNTP EKSVVSLFHD  480
PPGSYNPLTL GLFTLVYFFL ACWTYGLTVS AGVFIPSLLI GAAWGRLFGI SLSYLTGAAI  540
WADPGKYALM GAAAQLGGIV RMTLSLTVIM MEATSNVTYG FPIMLVLMTA KIVGDVFIEG  600
LYDMHIQLQS VPFLHWEAPV TSHSLTAREV MSTPVTCLRR REKVGVIVDV LSDTASNHNG  660
FPVVEHADDT QPARLQGLIL RSQLIVLLKH KVFVERSNLG LVQRRLRLKD FRDAYPRFPP  720
IQSIHVSQDE RECTMDLSEF MNPSPYTVPQ EASLPRVFKL FRALGLRHLV VVDNRNQVVG  780
LVTRKDLARY RLGKRGLEEL SLAQT                                        805

SEQ ID NO: 12          moltype = DNA  length = 3277
FEATURE                Location/Qualifiers
source                 1..3277
                       mol_type = other DNA
                       organism = Homo sapiens
SEQUENCE: 12
gccggcgctt cccggccggt gtcgctccgc ggcgggccat ggccaacgtc tctaagaagg  60
tgtcctggtc cggccgggac cgggacgacg aggaggcggc gccgctgctg cggaggacgg  120
cgcggcccgg cgggggggacg ccgctgctga acggggctgg gcccggggct gcgcgccagt  180
caccacgttc tgcgcttttc cgagtcggac atatgagcag cgtggagctg gatgatgaac  240
ttttggaccc ggatatggac cctccacatc ccttccccaa ggagatccca cacaacgaga  300
agctcctgtc cctcaagtat gagagcttgg actatgacaa cagtgagaac cagctgttcc  360
tggaggagga gcggcggatc aatcacacgg ccttccggac ggtggagatc aagcgctggg  420
tcatctgcgc cctcattggg atcctcacgg gcctcgtggc ctgcttcatt gacatcgtgg  480
tggaaaacct ggctggcctc aagtacaggg tcatcaaggg caatatcgac aagttcacag  540
agaagggcgg actgtccttc tccctgttgc tgtgggccac gctgaacgcc gccttcgtgc  600
tcgtgggctc tgtgattgtg gctttcatag agccggtggc tgctggcagc ggaatccccc  660
agatcaagtg cttcctcaac ggggtgaaga tcccccacgt ggtgcggctc aagacgttgg  720
tgatcaaagt gtccggtgtg atcctgtccg tggtcggggg cctggccgtg ggaaaggaag  780
ggccgatgat ccactcaggt tcagtgattg ccgccgggat ctctcaggga aggtcaacgt  840
cactgaaacg agatttcaag atcttcgagt acttccgcag agacacagag aagcgggact  900
tcgtctccgc aggggctgcg gccggagtgt cagcggcgtt tggagccccc gtgggtgggg  960
tcctgttcag cttggaggag ggtgcgtcct tctggaacca gttcctgacc tggaggatct  1020
tctttgcttc catgatctcc acgttcaccc tgaattttgt tctgagcatt taccacggga  1080
acatgtggga cctgtccagc ccaggcctca tcaacttcgg aaggtttgac tcggagaaaa  1140
tggcctacac gatccacgag atcccggtct tcatcgccat gggcgtggtg ggcggtgtgc  1200
ttggagctgt gttcaatgcc ttgaactact ggctgaccat gtttcgaatc aggtacatcc  1260
accggccctg cctgcaggtg attgaggccg tgctggtggc cgccgtcacg gccacagttg  1320
```

-continued

```
ccttcgtgct gatctactcg tcgcgggatt gccagcccct gcaggggggc tccatgtcct  1380
acccgctgca gctcttttgt gcagatggcg agtacaactc catggctgcg gccttcttca  1440
acaccccgga gaagagcgtg gtgagcctct tccacgaccc gccaggctcc tacaaccccc  1500
tgaccctcgg cctgttcacg ctggtctact tcttcctggc ctgctggacc tacgggctca  1560
cggtgtctgc cggggtcttc atcccgtccc tgctcatcgg ggctgcctgg ggccggctct  1620
ttgggatctc cctgtcctac ctcacggggg cggcgatctg ggcggacccc ggcaaatacg  1680
ccctgatggg agctgctgcc cagctgggcg ggattgtgcg gatgacactg agcctgaccg  1740
tcatcatgat ggaggccacc agcaacgtga cctacggctt ccccatcatg ctggtgctca  1800
tgaccgccaa gatcgtgggc gacgtcttca ttgagggcct gtacgacatg cacattcagc  1860
tgcagagtgt gcccttcctg cactgggagg ccccggtcac ctcacactca ctcactgcca  1920
gggaggtgat gagcacacca gtgacctgcc tgaggcggcg tgagaaggtc ggcgtcattg  1980
tggacgtgct gagcgacacg gcgtccaatc acaacggctt ccccgtggtg gagcatgccg  2040
atgacaccca gcctgcccgg ctccagggcc tgatcctgcg ctcccagctc atcgttctcc  2100
taaagcacaa ggtgtttgtg gagcggtcca acctgggcct ggtacagcgg cgcctgaggc  2160
tgaaggactt ccgagacgcc tacccgcgct tcccacccat ccagtccatc cacgtgtccc  2220
aggacgagcg ggagtgcacc atggacctct ccgagttcat gaacccctcc ccctacacgg  2280
tgccccagga ggcgtcgctc ccacgggtgt tcaagctgtt ccgggccctg ggcctgcggc  2340
acctggtggt ggtggacaac cgcaatcagg ttgtcgggtt ggtgaccagg aaggacctcg  2400
ccaggtaccg cctgggaaag agaggcttgg aggagctctc gctggcccag acgtgaggcc  2460
cagccctgcc cataatgggc actggcgctg gcaccccggc ccttctgcat ttcctcccgg  2520
agtcactggt ttctcggccc aaaccatgct ccccagcagt ggcaatggcg agcaccctgc  2580
agctgggcgg gcaggcggca ggcgcggaac tgaccctctc gcgggactga ccctgttgtg  2640
ggcagtggtc tccccccttg gcgcctcctt gcgcaggccc agcctccact ctcctcgtct  2700
aggtttcttt acctccaggg atcagctgtg tgtgtgtgac ctccctaccg ggctatcggc  2760
ctcttgggag ccagcggcag ggccggcacc tgcgtgcctg tgcccgtgtg cgtgagacag  2820
agcccttgcc cctgctgctg ccccgagggc tgccctgccc tggaagggcc cctctgcctc  2880
cacaccagtg gagtcttcga gacttgggag ctgcttggcc tcattttcag ccatgagcag  2940
acggcctgtg gtccctgggc ctgaggcacg gactcgtagc accagggttt ggaggctgcg  3000
accgccccgg agagcagctt cacactggcg ccacagagga gccccacgtg cactccccgg  3060
cctgcatccg gcttgggtac acaggcccag aggactgggg tgactcacgg gccctgtgct  3120
gtgatgttga gagctgagaa aaacctccaa ggccctgagc cccatgccca gccctgcctt  3180
ggtcccccaa tccccagagc ttggagtctg ggccccacac ccagccctgc cttggtccct  3240
gagcctcaaa gcgtggaatt gctgccctgt ggacact                           3277
```

SEQ ID NO: 13 moltype = AA length = 334
FEATURE Location/Qualifiers
source 1..334
mol_type = protein
organism = Homo sapiens

SEQUENCE: 13

```
MEPGPTAAQR RCSLPPWLPL GLLLWSGLAL GALPFGSSPH RVFHDLLSEQ QLLEVEDLSL  60
SLLQGGGLGP LSLPPDLPDL DPECRELLLD FANSSAELTG CLVRSARPVR LCQTCYPLFQ  120
QVVSKMDNIS RAAGNTSESQ SCARSLLMAD RMQIVVILSE FFNTTWQEAN CANCLTNNSE  180
ELSNSTVYFL NLFNHTLTCF EHNLQGNAHS LLQTKNYSEV CKNCREAYKT LSSLYSEMQK  240
MNELENKAEP GTHLCIDVED AMNITRKLWS RTFNCSVPCS DTVPVIAVSV FILFLPVVFY  300
LSSFLHSEQK KRKLILPKRL KSSTSFANIQ ENSN                              334
```

SEQ ID NO: 14 moltype = DNA length = 2869
FEATURE Location/Qualifiers
source 1..2869
mol_type = other DNA
organism = Homo sapiens

SEQUENCE: 14

```
ggctgtccgc ggtgccggct gggggcggag aggcggcggt gggctccctg gggtgtgtga  60
gcccggtgat ggagccgggc ccgacagccg cgcagcggag gtgttcgttg ccgccgtggc  120
tgccgctggg gctgctgctg tggtcggggc tggccctggg cgcgctcccc ttcggcagca  180
gtccgcacag ggtcttccac gacctcctgt cggagcagca gttgctggag gtggaggact  240
tgtccctgtc cctcctgcag ggtggagggc tggggcctct gtcgctgccc ccggacctgc  300
cggatctgga tcctgagtgc cgggagctcc tgctggactt cgccaacagc agcgcagagc  360
tgacagggtg tctggtgcgc agcgcccggc ccgtgcgcct ctgtcagacc tgctaccccc  420
tcttccaaca ggtcgtcagc aagatggaca acatcagccg agccgcgggg aatacttcag  480
agagtcagag ttgtgccaga agtctcttaa tggcagatag aatgcaaata gttgtgattc  540
tctcagaatt ttttaatacc acatggcagg aggcaaattg tgcaaattgt ttaacaaaca  600
acagtgaaga attatcaaac agcacagtat atttccttaa tctatttaat cacaccctga  660
cctgctttga acataacctt caggggaatg cacatagtct tttacagaca aaaaattatt  720
cagaagtatg caaaaactgc cgtgaagcat acaaaactct gagtagtctg tacagtgaaa  780
tgcaaaaaat gaatgaactt gagaataagg ctgaacctgg aacacattta tgcattgatg  840
tggaagatgc aatgaacatc actcgaaaac tatggagtcg aactttcaac tgttcagtcc  900
cttgcagtga cacagtgcct gtaattgctg tttctgtgtt cattctcttt ctacctgttg  960
tcttctacct tagtagcttt cttcactcag agcaaaagaa acgcaaactc attctgccca  1020
aacgtctcaa gtccagtacc agttttgcaa atattcagga aaattcaaac tgagacctac  1080
aaaatggaga attgacatat cacgtgaatg aatggtggaa gacacaactt ggtttcagaa  1140
agaagataaa ctgtgatttg acaagtcaag ctcttaagaa atacaaggac ttcagatcca  1200
tttttaaata agaattttcg atttttcttt ccttttccac ttctttctaa cagatttgga  1260
tatttttaat ttccaggcat agcagtgtta tctattttaa tgtgtatttg tcacaataac  1320
agaacatgca agaacaatca ttatttttatt ttataggcat ttgattacta ttctagactt  1380
```

-continued

ctggtatctt cttactaaca taagtatctc aagtagaaaa gtttttgaaa actaacattt 1440
aaaaattaat cagttacagt aaagactttg aaaaagaaat gtacttgtta ggaagtagct 1500
taattacccc ccattgcagt attattgtta tatatatagt taatatgttg tacatcacaa 1560
taatatataa ttcagtctct agtttcccta gagtcatttt tgaaaccact gattgcaaac 1620
ctccctgaca atttttaaaa gtagtaagcc acattacatt tatctttgta aaaagattta 1680
tggtaactgg tttcttactt gacttttata aatagtattt tacatcttat ttttgccttt 1740
atttcataag taatttaaaa atcactggat tgctttatta tattcagggc aatatggatt 1800
atttttatac caaggatttg catcgtgaat taaattaagt tatttggcaa tttataattt 1860
attactactt taaatcaaat gtagcattat cacactgtat ttaaattgtc atttttaaa 1920
ggaatatttt cttcttaaga tatatagagg attttggaga agagagacag gaggggtaaa 1980
accagcttaa ggttcagcga gcagaaaggg acctgagagg atgctcactg taagactgtt 2040
ggacagtggt gtgtattgag gggatgaatc ggaacgatag tctcatgcag aaaatagtga 2100
gattaagatc atccttattg tttctaaatt atttcaatca gatgaaagtg atacgattga 2160
aatgaaatca catagttcgt gctcagaaat tctattttgg tatgtttgta ttagccttta 2220
gaaaaaacac tccgtttcag aattgttcac agttttattt cttaggtttt tagagttcag 2280
gatttcattt attaatttct tcttgctttt ttggtggaaa taggctttgt tgtaaacatt 2340
aagaatataa aatctcctct atatagaaac aagaattttg ttaaaaagag aatttgaatc 2400
ccttcctata ctataaaatg ctctataggg agacaaagtg tttctttttt cttttatgtt 2460
tactgtttat gtggagtgaa atataaggct cttggatgta taacatactc aaaagctgtt 2520
acactttctc tgatctgctg tgatccactg aaaatgtgct ggggtttgtt ctgctgtcac 2580
tgtttatgct gctggaactt agcactgtct tgatttgaag catatgattg agagccattt 2640
gaagcaatct tcattaatgc agataaaaca agtttacatg tgcagagtta gaaaatgaca 2700
tgttcaattc tgtaagtggt gacttttga gcacctttca gtattatgta tttgtaaaaa 2760
ccattgtttt tggatataaa gctaataagc actttaaaaa aaaaaaaaaa aaaaaaaaaa 2820
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaagaaaaaa aaaaaaaaa 2869

SEQ ID NO: 15 moltype = AA length = 580
FEATURE Location/Qualifiers
source 1..580
mol_type = protein
organism = Homo sapiens
SEQUENCE: 15
MTAPAGPRGS ETERLLTPNP GYGTQAGPSP APPTPPEEED LRRRLKYFFM SPCDKFRAKG 60
RKPCKLMLQV VKILVVTVQL ILFGLSNQLA VTFREENTIA FRHLFLLGYS DGADDTFAAY 120
TREQLYQAIF HAVDQYLALP DVSLGRYAYV RGGGDPWTNG SGLALCQRYY HRGHVDPAND 180
TFDIDPMVVT DCIQVDPPER PPPPPSDDLT LLESSSSYKN LTLKFHKLVN VTIHFRLKTI 240
NLQSLINNEI PDCYTFSVLI TFDNKAHSGR IPISLETQAH IQECKHPSVF QHGDNSFRLL 300
FDVVVILTCS LSFLLCARSL LRGFLLQNEF VGFMWRQRGR VISLWERLEF VNGWYILLVT 360
SDVLTISGTI MKIGIEAKNL ASYDVCSILL GTSTLLVWVG VIRYLTFFHN YNILIATLRV 420
ALPSVMRFCC CVAVIYLGYC FCGWIVLGPY HVKFRSLSMV SECLFSLING DDMFVTFAAM 480
QAQQGRSSLV WLFSQLYLYS FISLFIYMVL SLFIALITGA YDTIKHPGGA GAEESELQAY 540
IAQCQDSPTS GKFRRGSGSA CSLLCCCGRD PSEEHSLLVN 580

SEQ ID NO: 16 moltype = DNA length = 2051
FEATURE Location/Qualifiers
source 1..2051
mol_type = other DNA
organism = Homo sapiens
SEQUENCE: 16
agatcagctg atgccggagg gtttgaagcc gcgccgcgag ggagcgaggt cgcagtgaca 60
gcggcgggcg atcggaccca ggctgccccg ccgtacccgc ctgcgtcccg cgctcccgcc 120
ccagcatgac agccccggcg ggtccgcgcg gctcagagac cgagcggctt ctgaccccca 180
accccgggta tgggacccag gcggggcctt caccggcccc tccgacaccc ccagaagagg 240
aagaccttcg ccgtcgtctc aaatactttt tcatgagtcc ctgcgacaag tttcgagcca 300
agggccgcaa gccctgcaag ctgatgctgc aagtggtcaa gatcctggtg gtcacggtgc 360
agctcatcct gtttgggctc agtaatcagc tggctgtgac attccgggaa gagaacacca 420
tcgccttccg acacctcttc ctgctgggct actcggacgg agcggatgac accttcgcag 480
cctacacgcg ggagcagctg taccaggcca tcttccatgc tgtggaccag tacctggcgt 540
tgcctgacgt gtcactgggc cggtatgcgt atgtccgtgg tgggggtgac ccttggacca 600
atggctcagg gcttgctctc tgccagcggt actaccaccg aggccacgtg gacccggcca 660
acgacacatt tgacattgat ccgatggtgg ttactgactg catccaggtg gatccccccg 720
agcggccccc tccgcccccc agcgacgatc tcaccctctt ggaaagcagc tccagttaca 780
agaacctcac gctcaaattc cacaagctgg tcaatgtcac catccacttc cggctgaaga 840
ccattaacct ccagagcctc atcaataatg agatcccgga ctgctatacc ttcagcgtcc 900
tgatcacgtt tgacaacaaa gcacacagtg ggcggatccc catcagcctg gagacccagg 960
cccacatcca ggagtgtaag caccccagtg tcttccagca cggagacaac agcttccggc 1020
tcctgtttga cgtggtggtc atcctcacct gctccctgtc cttcctcctc tgcgcccgct 1080
cactccttcg aggcttcctg ctgcagaacg agtttgtggg gttcatgtgg cggcagcggg 1140
gacgggtcat cagcctgtgg gagcggctgg aatttgtcaa tggctggtac atcctgctcg 1200
tcaccagcga tgtgctcacc atctcgggca ccatcatgaa gatcggcatc gaggccaaga 1260
acttggcgag ctacgacgtc tgcagcatcc tcctgggcac ctcgacgctg ctggtgtggg 1320
tgggcgtgat ccgctacctg accttcttcc acaactacaa tatcctcatc gccacactgc 1380
gggtggccct gcccagcgtc atgcgcttct gctgctgcgt ggctgtcatc tacctgggct 1440
actgcttctg tggctggatc gtgctggggc cctatcatgt gaagttccgc tcactctcca 1500
tggtgtctga gtgcctgttc tcgctcatca atggggacga catgtttgtg acgttcgccg 1560
ccatgcaggc gcagcagggc cgcagcagcc tggtgtggct cttctcccag ctctaccttt 1620

-continued

actccttcat cagcctcttc atctacatgg tgctcagcct cttcatcgcg ctcatcaccg 1680
gcgcctacga caccatcaag catcccggcg gcgcaggcgc agaggagagc gagctgcagg 1740
cctacatcgc acagtgccag gacagcccca cctccggcaa gttccgccgc gggagcggct 1800
cggcctgcag ccttctctgc tgctgcggaa gggacccctc ggaggagcat tcgctgctgg 1860
tgaattgatt cgacctgact gccgttggac cgtaggccct ggactgcaga gacccccgcc 1920
cccgaccccg cttatttatt tgtagggttt gcttttaagg atcggctccc tgtcgcgccc 1980
gaggagggcc tggacctttc gtgtcggacc cttgggggcg gggagactgg gtggggaggg 2040
tgttgaataa a 2051

SEQ ID NO: 17 moltype = AA length = 663
FEATURE Location/Qualifiers
source 1..663
mol_type = protein
organism = Homo sapiens
SEQUENCE: 17
MTGARASAAE QRRAGRSGQA RAAERAAGMS GAGRALAALL LAASVLSAAL LAPGGSSGRD 60
AQAAPPRDLD KKRHAELKMD QALLLIHNEL LWTNLTVYWK SECCYHCLFQ VLVNVPQSPK 120
AGKPSAAAAS VSTQHGSILQ LNDTLEEKEV CRLEYRFGEF GNYSLLVKNI HNGVSEIACD 180
LAVNEDPVDS NLPVSIAFLI GLAVIIVISF LRLLLSLDDF NNWISKAISS RETDRLINSE 240
LGSPSRTDPL DGDVQPATWR LSALPPRLRS VDTFRGIALI LMVFVNYGGG KYWYFKHASW 300
NGLTVADLVF PWFVFIMGSS IFLSMTSILQ RGCSKFRLLG KIAWRSFLLI CIGIIIVNPN 360
YCLGPLSWDK VRIPGVLQRL GVTYFVVAVL ELLFAKPVPE HCASERSCLS LRDITSSWPQ 420
WLLILVLEGL WLGLTFLLPV PGCPTGYLGP GGIGDFGKYP NCTGGAAGYI DRLLLGDDHL 480
YQHPSSAVLY HTEVAYDPEG ILGTINSIVM AFLGVQAGKI LLYYKARTKD ILIRFTAWCC 540
ILGLISVALT KVSENEGFIP VNKNLWSLSY VTTLSSFAFF ILLVLYPVVD VKGLWTGTPF 600
FYPGMNSILV YVGHEVFENY FPFQWKLKDN QSHKEHLTQN IVATALWVLI AYILYRKKIF 660
WKI 663

SEQ ID NO: 18 moltype = DNA length = 5228
FEATURE Location/Qualifiers
source 1..5228
mol_type = other DNA
organism = Homo sapiens
SEQUENCE: 18
agggcggggc gcagcgggca ggcaagggcg gccgagcggg cggcgggcat gagcggggcg 60
ggcagggcgc tggccgcgct gctgctggcc gcgtccgtgc tgagcgccgc gctgctggcc 120
cccggcggct cttcggggcg cgatgcccag gccgcgccgc cacgagactt agacaaaaaa 180
agacatgcag agctgaagat ggatcaggct ttgctactca tccataatga acttctctgg 240
accaacttga ccgtctactg gaaatctgaa tgctgttatc actgcttgtt tcaggttctg 300
gtaaacgttc ctcagagtcc aaaagcaggg aagcctagtg ctgcagctgc ctctgtcagc 360
acccagcacg gatctatcct gcagctgaac gacaccttgg aagagaaaga agtttgtagg 420
ttggaataca gatttggaga atttggaaac tattctctct tggtaaagaa catccataat 480
ggagttagtg aaattgcctg tgacctggct gtgaacgagg atccagttga tagtaacctt 540
cctgtgagca ttgcattcct tattggtctt gctgtcatca ttgtgatatc ctttctgagg 600
ctcttgttga gtttggatga ctttaacaat tggatttcta aagccataag ttctcgagaa 660
actgatcgcc tcatcaattc tgagctggga tctcccagca ggacagaccc tctcgatggt 720
gatgttcagc cagcaacgtg gcgtctatct gccctgccgc cccgcctccg cagcgtggac 780
accttcaggg ggattgctct tatactcatg gtctttgtca attatggagg aggaaaatat 840
tggtacttca aacatgcaag ttggaatggg ctgacagtgg ctgacctcgt gttcccgtgg 900
tttgtattta ttatgggatc ttccattttt ctatcgatga cttctatact gcaacggggg 960
tgttcaaaat tcagattgct ggggaagatt gcatggagga gtttcctgtt aatctgcata 1020
ggaattatca ttgtgaatcc caattattgc cttggtccat tgtcttggga caaggtgcgc 1080
attcctggtg tgctgcagcg attgggagtg acatactttg tggttgctgt gttggagctc 1140
ctctttgcta aacctgtgcc tgaacattgt gcctcggaga ggagctgcct ttctcttcga 1200
gacatcacgt ccagctggcc ccagtggctg ctcatcctgg tgctggaagg cctgtggctg 1260
ggcttgacat tcctcctgcc agtccctggg tgccctactg gttatcttgg tcctgggggc 1320
attggagatt ttggcaagta tccaaattgc actggaggag ctgcaggcta catcgaccgc 1380
ctgctgctgg gagacgatca cctttaccag cacccatctt ctgctgtact ttaccacacc 1440
gaggtggcct atgaccccga gggcatcctg ggcaccatca actccatcgt gatggccttt 1500
ttaggagttc aggcaggaaa aatactattg tattacaagg ctcggaccaa agacatcctg 1560
attcgattca ctgcttggtg ttgtattctt gggctcattt ctgttgctct gacgaaggtt 1620
tctgaaaatg aaggctttat tccagtaaac aaaaatctct ggtccctttc gtatgtcact 1680
acgctcagtt cttttgcctt cttcatcctg ctggtcctgt acccagttgt ggatgtgaag 1740
gggctgtgga caggaacccc attcttttat ccaggaatga attccattct ggtatatgtc 1800
ggccacgagg tgtttgagaa ctacttcccc tttcagtgga agctgaagga caaccagtcc 1860
cacaaggagc acctgactca gaacatcgtc gccactgccc tctgggtgct cattgcctac 1920
atcctctata gaaagaagat tttttggaaa atctgatggc tcccactgag atgtgctgct 1980
ggaagactct agtaggcctg cagggaggac tgaagcagcc tttgttaaag ggaagcattc 2040
attaggaaat tgactggctg cgtgtttaca gactctgggg gaagacactg atgtcctcaa 2100
actggttaac tgtgacacgg ctcgccagaa ctctgcctgt ctatttgtga cttacagatt 2160
tgaaatgtaa ttgtcttttt tcctccatct tctgtggaaa tggatgtctt tggaacttca 2220
ttccgaggag ataagcttta actttccaaa agggaattgc catgggtgtt tttcttctgt 2280
ggtgagtgaa acaatctgag gtctggttct tgctgacctt gttgccctgc aaacttcctt 2340
tccacgtgta cgcgcacacc aacacgaaat gccatcactc ctactgcggc tgctatgaag 2400
cttactggtt gtgatgtgtt ataatttagt ctgtttttt gattgaatgc agtttaatgt 2460
ttccagaaag ccaaagtaat tttcttttca gatatgcaag gctttggtgg gtccaaaaaa 2520

-continued

```
tgtctatcac aagccatttt ttccttttcc tctctcgaaa agttaaaata tctatgtgtt  2580
attcccaaac cctcttacct atgtatctgc ctgtctgtcc atcatcttcc ttcctcccta  2640
tctctgtgta tctggatggc agccgctgcc caggggagtg gctgtgggga gggcaggtac  2700
tgtctttgcc tgtgggtcca gctgagccat ccctgctggg tgatgctggg caagaccctt  2760
ggcccgtctg ggccttggct tcctcacttg tgaaatgagc gggaagatga ctctcagttc  2820
cttccacctc ttagacatgg tgaggtaaca gacatcaaaa gcttttctga aatcttcaga  2880
agaaatagtt ccattacaga aaactcttca aaataaatag tagtgaaaac ttttaaaaac  2940
tctcattgga gtaagtcttt tcaagatgat cctccacaat ggaggcagcg ttcctacttg  3000
tcatcacaca gctgaagaca ttgtttctta ggtgtgaaat cggggacaaa ggacaaacag  3060
agacacacgg cattgttcat gggaggcatc gtcaccctcc tgggtgttct gtgggaattt  3120
cctgtgtgag gaaaacgtgg ccacagggtt gtgctgtacc caccttccc cggcgagatg  3180
gccctcggcc tgtgccgctg cttccaccct cgccactcca tggcagcttt tggtctgttt  3240
ccggctctgc cctctgccct gaactctcat ccggcttgta cctgcctgct ggacccctcc  3300
acctggaggc cagcccatgt ctcaggccca gccctagcct cttctcctca aattctaagt  3360
gttttctctt taggtttccc tggctttgtg aatggatcat gtgtctctag gtataaacct  3420
gacatcatct ttccacccgg cttacctcca ccagatctcc ccagttctgt ctccatcttc  3480
tacctgcagc tgctctgttc tcatggtcac tgctgcatca ctgagtctgg acccttgtta  3540
tcattttcaa actggcctcc ttccctcgtt ccccacttct taaagtcacc tgtccattgc  3600
caccagatta agctttctcc agccagatca cctctctctg agaaacctcc attgacatgg  3660
aaacaccatt gtctggcaca catactcaca tactcacctt cccgtcttga tccccacaca  3720
tctttccagc ctccctctcc actccactcc ctgctctctc ctccacctcc ccatcctctt  3780
gtctcccctc ccctctgaat ccagcccagc ggggcttctc ctgcctccat cacatcacag  3840
aagtacctcc tgcttctggt tttaattaga gccttccccg attacatttt cctctgaatt  3900
ttttcctatc tacatttgat ctgtcatgtt taaacccccc acttctaagg gaacttctct  3960
aatctcttat cctcatcccc aaatagtgtt ttcttcctct gggttcttat aatgttggta  4020
tcaatctcac agcatttagt gcttcctgcc tggtgtgaca gttacctgtg tgcatgtgca  4080
atttctaatt tcccacgcta gactgtgagc ttcctaaggc aagaatcatg ccttgttggt  4140
ttctgtattc ctcatggtgc caaacacagt gccttctaca ttgcaggcgc tgaataaaca  4200
tttttaaagc aaaatgatgt ggatttttaa aataaatatt taagtgctgg taagatgagc  4260
atgtatccgg ggtgcccatg aaatgttctt ggggccgtgt ggggacagtc gtcattcctc  4320
ctcctgccac ccttttcttt cagtgagtca ctgtggatgg tcccagctgt gtcatcccaa  4380
agttcagcag ggaaagctga gctgggcctc tccaggtgag ttttctagaa gcatttctca  4440
aactgtgggt tacatcaact tgggtgtctt gagctgtaag gaaggaactc cggagtcagc  4500
tgggctacag gggagcttct ctaagtcctg cgggaggcca gacccagcct gagcttgctg  4560
ttagctagcg gaggcagctg ctggtggccc aggtgctcga caccaggcat cccctctcct  4620
cccacgaagg gtgtgccata atccccttca acaggaaatg cttcccagaa gcctctcagc  4680
agcctccccc cctgtcctat cagctagaag cgcctcgctt gtcccaagac cagcagggac  4740
agggaactgt ccgagcccgt ggctgtgtgg aggaaggcga cccccagcac aagattggtt  4800
tcctttggga agggaagagg gagtgtgttg gggtaagggg tagagcagag gaatggtcag  4860
ggggcaacaa ccgctgacag ctgcaacagg tgcatggcat ctcacaggga ggcagggagg  4920
tgcgagctcc taagtaatgg agcaaaaaaa ttctattctg tagaatgggg agagaaaatg  4980
tgacatttta attttttttt gcatttatat tcctaattcc tacttaaagt gaatatactg  5040
ccgctgtaga tcataaaatg tatcttttcc atggccaaca aggggcatct tttataaatg  5100
cataataacc cagtttgtat caaagggtat cgacttaagt gaaatttcaa catgctgtta  5160
ctttttcctt ttaatgtaat tctgttttcc aaataaatgg gggagacaaa tggaaaaaaa  5220
aaaaaaaa                                                          5228

SEQ ID NO: 19          moltype = AA  length = 1278
FEATURE                Location/Qualifiers
source                 1..1278
                       mol_type = protein
                       organism = Homo sapiens
SEQUENCE: 19
MTARGLALGL LLLLLCPAQV FSQSCVWYGE CGIAYGDKRY NCEYSGPPKP LPKDGYDLVQ  60
ELCPGFFFGN VSLCCDVRQL QTLKDNLQLP LQFLSRCPSC FYNLLNLFCE LTCSPRQSQF  120
LNVTATEDYV DPVTNQTKTN VKELQYYVGQ SFANAMYNAC RDVEAPSSND KALGLLCGKD  180
ADACNATNWI EYMFNKDNGQ APFTITPVFS DFPVHGMEPM NNATKGCDES VDEVTAPCSC  240
QDCSIVCGPK PQPPPPPAPW TILGLDAMYV IMWITYMAFL LVFFGAFFAV WCYRKRYFVS  300
EYTPIDSNIA FSVNASDKGE ASCCDPVSAA FEGCLRRLFT RWGSFCVRNP GCVIFFSLVF  360
ITACSSGLVF VRVTTNPVDL WSAPSSQARL EKEYFDQHFG PFFRTEQLII RAPLTDKHIY  420
QPYPSGADVP FGPPLDIQIL HQVLDLQIAI ENITASYDNE TVTLQDICLA PLSPYNTNCT  480
ILSVLNYFQN SHSVLDHKKG DDFFVYADYH THFLYCVRAP ASLNDTSLLH DPCLGTFGGP  540
VFPWLVLGGY DDQNYNNATA LVITFPVNNY YNDTEKLQRA QAWEKEFINF VKNYKNPNLT  600
ISFTAERSIE DELNRESDSD VFTVVISYAI MFLYISLALG HIKSCRRLLV DSKVSLGIAG  660
ILIVLSSVAC SLGVFSYIGL PLTLIVIEVI PFLVLAVGVD NIFILVQAYQ RDERLQGETL  720
DQQLGRVLGE VAPSMFLSSF SETVAFFLGA LSVMPAVHTF SLFAGLAVFI DFLLQITCFV  780
SLLGLDIKRQ EKNRLDIFCC VRGAEDGTSV QASESCLFRF FKNSYSPLLL KDWMRPIVIA  840
IFVGVLSFSI AVLNKVDIGL DQSLSMPDDS YMVDYFKSIS QYLHAGPPVY FVLEEGHDYT  900
SSKGQNMVCG GMGCNNDSLV QQIFNAAQLD NYTRIGFAPS SWIDDYFDWV KPQSSCCRVD  960
NITDQFCNAS VVDPACVRCR PLTPEGKQRP QGGDFMRFLP MFLSDNPNPK CGKGGHAAYS  1020
SAVNILLGHG TRVGATYFMT YHTVLQTSAD FIDALKKARL IASNVTETMG INGSAYRVFP  1080
YSVFYVFYEQ YLTIIDDTIF NLGVSLGAIF LVTMVLLGCE LWSAVIMCAT IAMVLVNMFG  1140
VMWLWGISLN AVSLVNLVMS CGISVEFCSH ITRAFTVSMK GSRVERAEEA LAHMGSSVFS  1200
GITLTKFGGI VVLAFAKSQI FQIFYFRMYL AMVLLGATHG LIFLPVLLSY IGPSVNKAKS  1260
CATEERYKGT ERERLLNF                                               1278
```

-continued

SEQ ID NO: 20 moltype = DNA length = 4673
FEATURE Location/Qualifiers
source 1..4673
mol_type = other DNA
organism = Homo sapiens
SEQUENCE: 20

```
tttgctcctg ctcctccgct cctcctgcgc ggggtgctga aacagcccgg ggaagtagag  60
ccgcctccgg ggagcccaac cagccgaacg ccgccggcgt cagcagcctt gcgcggccac  120
agcatgaccg ctcgcggcct ggcccttggc ctcctcctgc tgctactgtg tccagcgcag  180
gtgttttcac agtcctgtgt ttggtatgga gagtgtggaa ttgcatatgg ggacaagagg  240
tacaattgcg aatattctgg cccaccaaaa ccattgccaa aggatggata tgacttagtg  300
caggaactct gtccaggatt cttctttggc aatgtcagtc tctgttgtga tgttcggcag  360
cttcagacac taaaagacaa cctgcagctg cctctacagt ttctgtccag atgtccatcc  420
tgtttttata acctactgaa cctgttttgt gagctgacat gtagccctcg acagagtcag  480
tttttgaatg ttacagctac tgaagattat gttgatcctg ttacaaacca gacgaaaaca  540
aatgtgaaag agttacaata ctacgtcgga cagagttttg ccaatgcaat gtacaatgcc  600
tgccgggatg tggaggcccc ctcaagtaat gacaaggccc tgggactcct gtgtgggaag  660
gacgctgacg cctgtaatgc caccaactgg attgaataca tgttcaataa ggacaatgga  720
caggcacctt ttaccatcac tcctgtgttt tcagattttc cagtccatgg gatggagccc  780
atgaacaatg ccaccaaagg ctgtgacgag tctgtggatg aggtcacagc accatgtagc  840
tgccaagact gctctattgt ctgtggcccc aagccccagc ccccacctcc tcctgctccc  900
tggacgatcc ttggcttgga cgccatgtat gtcatcatgt ggatcaccta catggcgttt  960
ttgcttgtgt tttttggagc attttttgca gtgtggtgct acagaaaacg gtattttgtc  1020
tccgagtaca ctcccatcga tagcaatata gctttttctg ttaatgcaag tgacaaagga  1080
gaggcgtcct gctgtgaccc tgtcagcgca gcatttgagg gctgcttgag gcggctgttc  1140
acacgctggg ggtctttctg cgtccgaaac cctggctgtg tcattttctt ctcgctggtc  1200
ttcattactg cgtgttcgtc aggcctggtg tttgtccggg tcacaaccaa tccagttgac  1260
ctctggtcag cccccagcag ccaggctcgc ctggaaaaag agtactttga ccagcacttt  1320
gggcctttct tccggacgga gcagctcatc atccgggccc ctctcactga caaacacatt  1380
taccagccat acccttcggg agctgatgta ccctttggac ctccgcttga catacagata  1440
ctgcaccagg ttcttgactt acaaatagcc atcgaaaaca ttactgcctc ttatgacaat  1500
gagactgtga cacttcaaga catctgcttg gcccctcttt caccgtataa cacgaactgc  1560
accattttga gtgtgttaaa ttacttccag aacagccatt ccgtgctgga ccacaagaaa  1620
ggggacgact tctttgtgta tgccgattac cacacgcact ttctgtactg cgtacgggct  1680
cctgcctctc tgaatgatac aagtttgctc catgaccctt gtctgggtac gtttggtgga  1740
ccagtgttcc cgtggcttgt gttgggaggc tatgatgatc aaaactacaa taacgccact  1800
gcccttgtga ttaccttccc tgtcaataat tactataatg atacagagaa gctccagagg  1860
gcccaggcct gggaaaaaga gtttattaat tttgtgaaaa actacaagaa tcccaatctg  1920
accatttcct tcactgctga acgaagtatt gaagatgaac taaatcgtga aagtgacagt  1980
gatgtcttca ccgttgtaat tagctatgcc atcatgtttc tatatatttc cctagccttg  2040
gggcacatca aaagctgtcg caggcttctg gtggattcga aggtctcact aggcatcgcg  2100
ggcatcttga tcgtgctgag ctcggtggct tgctccttgg gtgtcttcag ctacattggg  2160
ttgcccttga ccctcattgt gattgaagtc atcccgttcc tggtgctggc tgttggagtg  2220
gacaacatct tcattctggt gcaggcctac cagagagatg aacgtcttca aggggaaacc  2280
ctggatcagc agctgggcag ggtcctagga gaagtggctc ccagtatgtt cctgtcatcc  2340
ttttctgaga ctgtagcatt tttcttagga gcattgtccg tgatgccagc cgtgcacacc  2400
ttctctctct ttgcgggatt ggcagtcttc attgactttc ttctgcagat tacctgtttc  2460
gtgagtctct tggggttaga cattaaacgt caagagaaaa atcggctaga catcttttgc  2520
tgtgtcagag gtgctgaaga tggaacaagc gtccaggcct cagagagctg tttgtttcgc  2580
ttcttcaaaa actcctattc tccacttctg ctaaaggact ggatgagacc aattgtgata  2640
gcaatatttg tgggtgttct gtcattcagc atcgcagtcc tgaacaaagt agatattgga  2700
ttggatcagt ctctttcgat gccagatgac tcctacatgg tggattattt caaatccatc  2760
agtcagtacc tgcatgcggg tccgcctgtg tactttgtcc tggaggaagg gcacgactac  2820
acttcttcca aggggcagaa catggtgtgc ggcggcatgg gctgcaacaa tgattccctg  2880
gtgcagcaga tatttaacgc ggcgcagctg gacaactata cccgaatagg cttcgccccc  2940
tcgtcctgga tcgacgatta tttcgactgg gtgaagccac agtcgtcttg ctgtcgagtg  3000
gacaatatca ctgaccagtt ctgcaatgct tcagtggttg accctgcctg cgttcgctgc  3060
aggcctctga ctccggaagg caaacagagg cctcaggggg gagacttcat gagattcctg  3120
cccatgttcc tttcggataa ccctaacccc aagtgtggca aaggggggaca tgctgcctat  3180
agttctgcag ttaacatcct ccttggccat ggcaccaggg tcggagccac gtacttcatg  3240
acctaccaca ccgtgctgca gacctctgct gactttattg acgctctgaa gaaagcccga  3300
cttatagcca gtaatgtcac cgaaaccatg ggcattaacg gcagtgccta ccgagtattt  3360
ccttacagtg tgttttatgt cttctacgaa cagtacctga ccatcattga cgacactatc  3420
ttcaacctcg gtgtgtccct gggcgcgata tttctggtga ccatggtcct cctgggctgt  3480
gagctctggt ctgcagtcat catgtgtgcc accatcgcca tggtcttggt caacatgttt  3540
ggagttatgt ggctctgggg catcagtctg aacgctgtat ccttggtcaa cctggtgatg  3600
agctgtggca tctccgtgga gttctgcagc cacataacca gagcgttcac ggtgagcatg  3660
aaaggcagcc gcgtggagcg cgcggaagag gcacttgccc acatgggcag ctccgtgttc  3720
agtggaatca cacttacaaa atttggaggg attgtggtgt tggcttttgc caaatctcaa  3780
attttccaga tattctactt caggatgtat ttggccatgg tcttactggg agccactcac  3840
ggattaatat ttctccctgt cttactcagt tacatagggc catcagtaaa taaagccaaa  3900
agttgtgcca ctgaagagcg atacaaagga acagagcgcg aacggcttct aaatttctag  3960
ccctctcgca gggcatcctg actgaactgt gtctaagggt cggtcggttt accactggac  4020
gggtgctgca tcggcaaggc caagttgaac accggatggt gccaaccatc ggttgtttgg  4080
cagcagcttt gaacgtagcg cctgtgaact caggaatgca cagttgactt gggaagcagt  4140
attactagat ctggaggcaa ccacaggaca ctaaacttct cccagcctct tcaggaaaga  4200
```

-continued

aacctcattc tttggcaagc aggaggtgac actagatggc tgtgaatgtg atccgctcac 4260
tgacactctg taaaggccaa tcaatgcact gtctgtcctc tccttttag gagtaagcca 4320
tcccacaagt tctataccat atttttagtg acagttgagg ttgtagatac actttataac 4380
attttatagt ttaaagagct ttattaatgc aataaattaa ctttgtacac atttttatat 4440
aaaaaaacag caagtgattt cagaatgttg taggcctcat tagagcttgg tctccaaaaa 4500
tctgtttgaa aaaagcaaca tgttcttcac agtgttcccc tagaaaggaa gagatttaat 4560
tgccagttag atgtggcatg aaatgaggga caaagaaagc atctcgtagg tgtgtctact 4620
gggttttaac ttatttttct ttaataaaat acattgtttt cctaaaaaaa aaa 4673

SEQ ID NO: 21 moltype = AA length = 410
FEATURE Location/Qualifiers
source 1..410
mol_type = protein
organism = Homo sapiens
SEQUENCE: 21
MVCFRLFPVP GSGLVLVCLV LGAVRSYALE LNLTDSENAT CLYAKWQMNF TVRYETTNKT 60
YKTVTISDHG TVTYNGSICG DDQNGPKIAV QFGPGFSWIA NFTKAASTYS IDSVSFSYNT 120
GDNTTFPDAE DKGILTVDEL LAIRIPLNDL FRCNSLSTLE KNDVVQHYWD VLVQAFVQNG 180
TVSTNEFLCD KDKTSTVAPT IHTTVPSPTT TPTPKEKPEA GTYSVNNGND TCLLATMGLQ 240
LNITQDKVAS VININPNTTH STGSCRSHTA LLRLNSSTIK YLDFVFAVKN ENRFYLKEVN 300
ISMYLVNGSV FSIANNNLSY WDAPLGSSYM CNKEQTVSVS GAFQINTFDL RVQPFNVTQG 360
KYSTAQDCSA DDDNFLVPIA VGAALAGVLI LVLLAYFIGL KHHHAGYEQF 410

SEQ ID NO: 22 moltype = DNA length = 1868
FEATURE Location/Qualifiers
source 1..1868
mol_type = other DNA
organism = Homo sapiens
SEQUENCE: 22
ccgattcctg gcttttgcaa ggctgtggtc ggtggtcatc agtgctcttg acccaggtcc 60
agcgagcctt ttccctggtg ttgcagctgt tgttgtaccg ccgccgtcgc cgccgtcgcc 120
gcctgctctg cggggtcatg gtgtgcttcc gcctcttccc ggttccgggc tcagggctcg 180
ttctggtctg cctagtcctg ggagctgtgc ggtcttatgc attggaactt aatttgacag 240
attcagaaaa tgccacttgc ctttatgcaa aatggcagat gaatttcaca gtacgctatg 300
aaactacaaa taaaacttat aaaactgtaa ccatttcaga ccatggcact gtgacatata 360
atggaagcat ttgtggggat gatcagaatg gtcccaaaat agcagtgcag ttcggacctg 420
gcttttcctg gattgcgaat tttaccaagg cagcatctac ttattcaatt gacagcgtct 480
cattttccta caacactggt gataacacaa catttcctga tgctgaagat aaaggaattc 540
ttactgttga tgaacttttg gccatcagaa ttccattgaa tgacctttt agatgcaata 600
gtttatcaac tttggaaaag aatgatgttg tccaacacta ctgggatgtt cttgtacaag 660
cttttgtcca aaatggcaca gtgagcacaa atgagttcct gtgtgataaa gacaaaactt 720
caacagtggc acccaccata cacaccactg tgccatctcc tactacaaca cctactccaa 780
aggaaaaacc agaagctgga acctattcag ttaataatgg caatgatact tgtctgctgg 840
ctaccatggg gctgcagctg aacatcactc aggataaggt tgcttcagtt attaacatca 900
accccaatac aactcactcc acaggcagct gccgttctca cactgctcta cttagactca 960
atagcagcac cattaagtat ctagactttg tctttgctgt gaaaaatgaa aaccgatttt 1020
atctgaagga agtgaacatc agcatgtatt tggttaatgg ctccgttttc agcattgcaa 1080
ataacaatct cagctactgg gatgcccccc tgggaagttc ttatatgtgc aacaaagagc 1140
agactgtttc agtgtctgga gcatttcaga taaatacctt tgatctaagg gttcagcctt 1200
tcaatgtgac acaaggaaag tattctacag ctcaagactg cagtgcagat gacgacaact 1260
tccttgtgcc catagcggtg ggagctgcct tggcaggagt acttattcta gtgttgctgg 1320
cttattttat tggtctcaag caccatcatg ctggatatga gcaattttag aatctgcaac 1380
ctgattgatt atataaaaat acatgcaaat aacaagattt tcttacctct cagttgttga 1440
aacactttgc ttcttaaaat tgatatgttg aaactttaat tcttttatca atcccagcat 1500
tttgagatca gtctttatta ataaaacctg ttctctttaa tcagcttaaa atccaaagtg 1560
tcatatttac tggtcctgga gacaaacttg ttcaaaagaa catcaacgtg caatgtttta 1620
aggtctatct taagaagccc tggccaaatt ttgatcctaa ccttgaagta tgccttgaac 1680
ttattaacat ggccattata agaataaaat atgtagttgt gtcttaatgg aattaataaa 1740
tgtcatttca ctactggtgt tctgttttca atgtataagg actatagtga tttaaactca 1800
tcaatgtgcc tttgcataaa gttgattaaa taaatattga tgtggtataa atgcccatca 1860
gatatgct 1868

SEQ ID NO: 23 moltype = AA length = 410
FEATURE Location/Qualifiers
source 1..410
mol_type = protein
organism = Homo sapiens
SEQUENCE: 23
MVCFRLFPVP GSGLVLVCLV LGAVRSYALE LNLTDSENAT CLYAKWQMNF TVRYETTNKT 60
YKTVTISDHG TVTYNGSICG DDQNGPKIAV QFGPGFSWIA NFTKAASTYS IDSVSFSYNT 120
GDNTTFPDAE DKGILTVDEL LAIRIPLNDL FRCNSLSTLE KNDVVQHYWD VLVQAFVQNG 180
TVSTNEFLCD KDKTSTVAPT IHTTVPSPTT TPTPKEKPEA GTYSVNNGND TCLLATMGLQ 240
LNITQDKVAS VININPNTTH STGSCRSHTA LLRLNSSTIK YLDFVFAVKN ENRFYLKEVN 300
ISMYLVNGSV FSIANNNLSY WDAPLGSSYM CNKEQTVSVS GAFQINTFDL RVQPFNVTQG 360
KYSTAQECSL DDDTILIPII VGAGLSGLII VIVIAYVIGR RKSYAGYQTL 410

-continued

SEQ ID NO: 24 moltype = DNA length = 4006
FEATURE Location/Qualifiers
source 1..4006
mol_type = other DNA
organism = Homo sapiens
SEQUENCE: 24
ccgattcctg gcttttgcaa ggctgtggtc ggtggtcatc agtgctcttg acccaggtcc 60
agcgagcctt ttccctggtg ttgcagctgt tgttgtaccg ccgccgtcgc cgccgtcgcc 120
gcctgctctg cggggtcatg gtgtgcttcc gcctcttccc ggttccgggc tcagggctcg 180
ttctggtctg cctagtcctg ggagctgtgc ggtcttatgc attggaactt aatttgacag 240
attcagaaaa tgccacttgc ctttatgcaa aatggcagat gaatttcaca gtacgctatg 300
aaactacaaa taaaacttat aaaactgtaa ccatttcaga ccatggcact gtgacatata 360
atggaagcat ttgtggggat gatcagaatg gtcccaaaat agcagtgcag ttcggacctg 420
gcttttcctg gattgcgaat tttaccaagg cagcatctac ttattcaatt gacagcgtct 480
cattttccta caacactggt gataacacaa catttcctga tgctgaagat aaaggaattc 540
ttactgttga tgaacttttg gccatcagaa ttccattgaa tgacctttt agatgcaata 600
gtttatcaac tttggaaaag aatgatgttg tccaacacta ctgggatgtt cttgtacaag 660
cttttgtcca aaatggcaca gtgagcacaa atgagttcct gtgtgataaa gacaaaactt 720
caacagtggc acccaccata cacaccactg tgccatctcc tactacaaca cctactccaa 780
aggaaaaacc agaagctgga acctattcag ttaataatgg caatgatact tgtctgctgg 840
ctaccatggg gctgcagctg aacatcactc aggataaggt tgcttcagtt attaacatca 900
accccaatac aactcactcc acaggcagct gccgttctca cactgctcta cttagactca 960
atagcagcac cattaagtat ctagactttg tctttgctgt gaaaaatgaa aaccgatttt 1020
atctgaagga agtgaacatc agcatgtatt tggttaatgg ctccgttttc agcattgcaa 1080
ataacaatct cagctactgg gatgcccccc tgggaagttc ttatatgtgc aacaaagagc 1140
agactgtttc agtgtctgga gcatttcaga taaatacctt tgatctaagg gttcagcctt 1200
tcaatgtgac acaaggaaag tattctacag cccaagagtg ttcgctggat gatgacacca 1260
ttctaatccc aattatagtt ggtgctggtc tttcaggctt gattatcgtt atagtgattg 1320
cttacgtaat tggcagaaga aaaagttatg ctggatatca gactctgtaa cactaatcaa 1380
tacgtgatct ctgttacaaa agaaaaaagc aagtacaagt tccaacatgc aatactggtc 1440
aacttaaggt atatttagtt gcagtccagc tctttagaat gggtggtatg ggggatttca 1500
aacttaaaca aaaaactatc aactacaaat tagttgcctg actttggttt ttccaaccaa 1560
ggaatttaaa actgttattt ttacagcaaa agatgtgcaa aatcactgga ttataagttc 1620
tattttactg tcttgaatta gtatttcagt gttttcattt tagacattca gactaaaaat 1680
acaccgttta gaaaaaacaa tttttgaaaa agagattttt tttccctgca ggtagttgag 1740
ttgaacaaca tgttctaccg tggatttgta cttgctcctt ttgctctttt tgtgtgtgtg 1800
tgtgtgtgtg tgtgtgtgtg tgtgattttt gtttgcaggt taacttagct actttggcat 1860
tgctgcatat ttgacctttg agagatataa tagtagattt gaacaggggc tggtattatt 1920
atgttcttag caataaatgc ttttctaatg ccttttgaat acatttgtat ttatgtggct 1980
gtaatgacaa aagatacaaa agctttttaa aatttagagt aggtattaat cttattgttt 2040
aatctttttt ttaaaaaaac tggatatttc aatcttttaa attgcaatat ataagactat 2100
tccaactggg catttcaatc catttttag gtgctttaga gataattgct tgccagtgcc 2160
aattgagggc attagtactt tgtgctcata aattggcctc tgtatgcagt actaaaatta 2220
atgcagattt ctctttagcc ttccaacatt tcttgttgat agtgatgtat tttattattt 2280
tctttttctt aagaaatgcc agtgtgtcct agaacctaga taacgaagtg cacttacact 2340
tataaaataa cttgcatcta ggctgggcgt ggcggctcac gcctgtaatc ccagcacttt 2400
gggaggccga agtgggtgga tcacttgagg ccaggagttt gagaccagcc tggccaacat 2460
ggtgaaaccc catctctatc agaaatacaa aaaattagct gggcatggtg gtgggcgcct 2520
gtaatcccag ttactcggga ggctgaggca ggagaatcac ttgaacccgg gaggcagagg 2580
ttgcggtgag ccaagagcgc accattgcac tccagccttg ggcgacaaaa acgaaactcc 2640
atcttcaaaa caaaacaaaa caaaacaaac aaacaaacaa aacttgcatc ttaaccaaaa 2700
gtcttggttt tatcttaatc cattaaaagt tggtcttgtt tccagcttgc attgattgct 2760
acaacatcac taatttggct ttcacattta aatggttctg tgctaatcaa aactttcgtt 2820
gttattattc gttatggtag aatcattttt aattcacgtg ctttgtgttc agttttgtgg 2880
tctgagagat gtaccaattg tcaaattacc gtgtaccacc taatgtttat aggagaaagc 2940
aaaatacatc agcttggtag ttaacacatc aaatatttct tgctgcttct aggagaactt 3000
ttttggtgtg tgttggaatg gctgagcaaa tattaaaatt gttaatatgc agccatatat 3060
ggaaggttcc tgtggggttg ttttttcgtg ttttttttttt ttgtggtggg attatgtgcc 3120
tcccattcac tagaaaatga gaaaattgtc tgggttccaa aatattgaca ttgaatggat 3180
caatacacac acacagacat atatatatat atatgcacac atatataggc agttgcatgc 3240
ctagcatggg tattttataa ccatataact gagttatatt ggaattataa atattttccg 3300
tcacttaaat ttgttctttg tttagcctga aaacctttat ggctcaagat cagattcctg 3360
actaacccct ctcttagagc tacagcgagc tgcattacca gcttaaaaca cttcttaggg 3420
attaaatata gatgtaattt ttcaaaatcg tttttaattt aaactgtgtt ttagtgtaaa 3480
attgttaacc ttgtaagatg gataatgtgt ataagaatgt aggccttaac tatttcacat 3540
gagtcaaaac aaagcagctt taaaaaaata attggaagca caatgcatgg cactgactga 3600
atgctgttaa tatttctaaa agtttctaca ttcagattat atgcctgatt catagtaaaa 3660
tacctctaat aaacactgtt ttatagaaaa cctgacttca gtgaatattt ttgtatttta 3720
catgggccag tttatatact gctatttaca ctattatttc ctatagctac atgttctttg 3780
taccttttgt agttttattt gtattactag attcatacct tgatggtaac gctctatctg 3840
gttttgggtg tttttcatgt tttagcattt gtataaagaa actggtccat gtaaatactt 3900
tccatgtttt ttcttcaaat gtttaaacca ctagttgatg tatggtatct ttagatattt 3960
gcctgtctgt ttgctcaaaa ttgcttctaa aacaataaag attctt 4006

SEQ ID NO: 25 moltype = AA length = 411
FEATURE Location/Qualifiers

-continued

source 1..411
mol_type = protein
organism = Homo sapiens
SEQUENCE: 25
MVCFRLFPVP GSGLVLVCLV LGAVRSYALE LNLTDSENAT CLYAKWQMNF TVRYETTNKT 60
YKTVTISDHG TVTYNGSICG DDQNGPKIAV QFGPGFSWIA NFTKAASTYS IDSVSFSYNT 120
GDNTTFPDAE DKGILTVDEL LAIRIPLNDL FRCNSLSTLE KNDVVQHYWD VLVQAFVQNG 180
TVSTNEFLCD KDKTSTVAPT IHTTVPSPTT TPTPKEKPEA GTYSVNNGND TCLLATMGLQ 240
LNITQDKVAS VININPNTTH STGSCRSHTA LLRLNSSTIK YLDFVFAVKN ENRFYLKEVN 300
ISMYLVNGSV FSIANNNLSY WDAPLGSSYM CNKEQTVSVS GAFQINTFDL RVQPFNVTQG 360
KYSTAEECSA DSDLNFLIPV AVGVALGFLI IVVFISYMIG RRKSRTGYQS V 411

SEQ ID NO: 26 moltype = DNA length = 1236
FEATURE Location/Qualifiers
source 1..1236
mol_type = other DNA
organism = Homo sapiens
SEQUENCE: 26
atggtgtgct tccgcctctt cccggttccg ggctcagggc tcgttctggt ctgcctagtc 60
ctgggagctg tgcggtctta tgcattggaa cttaatttga cagattcaga aaatgccact 120
tgcctttatg caaaatggca gatgaatttc acagtacgct atgaaactac aaataaaact 180
tataaaactg taaccatttc agaccatggc actgtgacat ataatggaag catttgtggg 240
gatgatcaga atggtcccaa aatagcagtg cagttcggac ctggcttttc ctggattgcg 300
aattttacca aggcagcatc tacttattca attgacagcg tctcattttc ctacaacact 360
ggtgataaca caacatttcc tgatgctgaa gataaaggaa ttcttactgt tgatgaactt 420
ttggccatca gaattccatt gaatgacctt tttagatgca atagtttatc aactttggaa 480
aagaatgatg ttgtccaaca ctactgggat gttcttgtac aagcttttgt ccaaaatggc 540
acagtgagca caaatgagtt cctgtgtgat aaagacaaaa cttcaacagt ggcacccacc 600
atacacacca ctgtgccatc tcctactaca acacctactc caaaggaaaa accagaagct 660
ggaacctatt cagttaataa tggcaatgat acttgtctgc tggctaccat ggggctgcag 720
ctgaacatca ctcaggataa ggttgcttca gttattaaca tcaaccccaa tacaactcac 780
tccacaggca gctgccgttc tcacactgct ctacttagac tcaatagcag caccattaag 840
tatctagact ttgtctttgc tgtgaaaaat gaaaaccgat tttatctgaa ggaagtgaac 900
atcagcatgt atttggttaa tggctccgtt ttcagcattg caaataacaa tctcagctac 960
tgggatgccc ccctgggaag ttcttatatg tgcaacaaag agcagactgt ttcagtgtct 1020
ggagcatttc agataaatac ctttgatcta agggttcagc ctttcaatgt gacacaagga 1080
aagtattcta cagctgaaga atgttctgct gactctgacc tcaactttct tattcctgtt 1140
gcagtgggtg tggccttggg cttccttata attgttgtct ttatctctta tatgattgga 1200
agaaggaaaa gtcgtactgg ttatcagtct gtgtaa 1236

SEQ ID NO: 27 moltype = AA length = 6
FEATURE Location/Qualifiers
REGION 1..6
note = Synthetic construct
source 1..6
mol_type = protein
organism = synthetic construct
SEQUENCE: 27
HHHHHH 6

What is claimed is:

1. A method of treating a lysosomal transmembrane protein disease or disorder in a subject comprising:

introducing a corresponding functional human lysosomal transmembrane protein into hematopoietic stem and progenitor cells (HSPCs) of the subject; and

transplanting the HSPCs into the subject,

thereby treating the lysosomal transmembrane protein disease or disorder.

2. The method of claim **1**, wherein when:

(a) the lysosomal transmembrane protein disease or disorder is mucopolysaccharidosis type IIIC, the corresponding functional human lysosomal transmembrane protein is HGSNAT; and

(b) the lysosomal transmembrane protein disease or disorder is Neiman-Pick Type C, the corresponding functional human lysosomal transmembrane protein is NPC-1.

3. The method of claim **1**, wherein the step of introducing comprises contacting a vector comprising a polynucleotide encoding functional human lysosomal transmembrane protein and a functional promoter with the HSPCs and allowing expression of the functional human lysosomal transmembrane protein.

4. The method of claim **1**, wherein the subject is human.

5. The method of claim **1**, wherein the vector is a viral vector selected from the group consisting of a lentiviral, adenoviral, and AAV vector.

6. The method of claim **1**, wherein the step of introducing is performed ex vivo.

7. The method of claim **1**, wherein the HSPCs are isolated from the bone marrow of the subject.

8. A method of treating or ameliorating a lysosomal protein disease or disorder in a subject comprising:

isolating hematopoietic stem and progenitor cells (HSPCs) from bone marrow from the subject;

introducing a functional human lysosomal transmembrane gene into the HSPCs, wherein the gene encodes a protein corresponding to the lysosomal protein disease or disorder; and

transplanting the HSPCs back into the subject,

thereby treating or ameliorating the lysosomal protein disease or disorder.

9. The method of claim **8**, wherein when:

(a) the lysosomal transmembrane protein disease or disorder is mucopolysaccharidosis type IIIC, the functional human lysosomal transmembrane gene is HGSNAT; and

(b) the lysosomal transmembrane protein disease or disorder is Neiman-Pick Type C, the functional human lysosomal transmembrane gene is NPC1.

10. The method of claim **8**, wherein the HSPCs are CD34+ cells.

11. The method of claim **8**, wherein administration is intravenous.

12. A method of treating or ameliorating a lysosomal protein disease or disorder in a subject comprising: producing a functional human lysosomal transmembrane gene in the subject using gene editing.

13. The method of claim **12**, wherein when:

(a) the lysosomal transmembrane protein disease or disorder is mucopolysaccharidosis type IIIC, the functional human lysosomal transmembrane gene is HGSNAT; and

(b) the lysosomal transmembrane protein disease or disorder is Neiman-Pick Type C, the functional human lysosomal transmembrane gene is NPC1.

14. The method of claim **12**, wherein producing a functional human lysosomal gene un the subject comprises contacting cells expressing a defective lysosomal transmembrane protein from the subject with a vector encoding a gene editing system that, when transfected into the cells, removes a trinucleotide extension mutation of an endogenous gene encoding the lysosomal transmembrane protein, thereby treating the lysosomal protein disease or disorder.

15. The method of claim **14**, wherein the gene editing system is selected from the group consisting of CRISPR/Cas, zinc finger nucleases, and transcription activator-life effector nucleases.

16. The method of claim **14**, wherein the step of contacting comprises administering to the subject an effective amount of the vector.

17. The method of claim **14**, wherein the step of contacting comprises obtaining a sample of cells from the subject, transfecting the gene editing system into the sample of cells, and thereafter, transplanting the transfected cells into the subject.

18. The method of claim **17**, wherein the sample of cells is selected from the group consisting of blood cells and HSPCs.

* * * * *