

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2021337530 B2**

(54) Title
Codon Optimized RPGRorf 15 Genes And Uses Thereof

(51) International Patent Classification(s)
A61K 9/127 (2006.01) **A61K 38/00** (2006.01)
A61K 9/14 (2006.01) **C07K 14/47** (2006.01)
A61K 9/50 (2006.01) **C12N 15/88** (2006.01)

(21) Application No: **2021337530** (22) Date of Filing: **2021.08.30**

(87) WIPO No: **WO22/051232**

(30) Priority Data

(31) Number	(32) Date	(33) Country
63/073,843	2020.09.02	US

(43) Publication Date: **2022.03.10**

(44) Accepted Journal Date: **2023.08.24**

(71) Applicant(s)
4D Molecular Therapeutics Inc.

(72) Inventor(s)
KIRN, David H.;KOTTERMAN, Melissa A.;SCHAFFER, David;FRANCIS, Peter

(74) Agent / Attorney
WRAYS PTY LTD, L7 863 Hay St, Perth, WA, 6000, AU

(56) Related Art
WO 2015/160893 A1



(51) International Patent Classification:

A61K 9/127 (2006.01) A61K 38/00 (2006.01)
A61K 9/14 (2006.01) C07K 14/47 (2006.01)
A61K 9/50 (2006.01) C12N 15/88 (2006.01)

(21) International Application Number:

PCT/US2021/048267

(22) International Filing Date:

30 August 2021 (30.08.2021)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/073,843 02 September 2020 (02.09.2020) US

(71) Applicant: **4D MOLECULAR THERAPEUTICS INC.**

[US/US]; 5858 Horton Street, Suite 455, Emeryville, CA 94608 (US).

(72) Inventors: **KIRN, David, H.**; c/o 4D Molecular Therapeutics Inc., 5858 Horton Street, Suite 455, Emeryville, CA 94608 (US). **KOTTERMAN, Melissa, A.**; c/o 4D Molecular Therapeutics Inc., 5858 Horton Street, Suite 455, Emeryville, CA 94608 (US). **SCHAFFER, David**; c/o 4D Molecular Therapeutics Inc., 5858 Horton Street, Suite 455, Emeryville, CA 94608 (US). **FRANCIS, Peter**; c/o 4D

Molecular Therapeutics Inc., 5858 Horton Street, Suite 455, Emeryville, CA 94608 (US).

(74) Agent: **CABRAL, Christopher, M.** et al.; Much She-
list PC, 191 N. Wacker Drive, Suite 1800, Chicago, IL
60606-1615 (US).

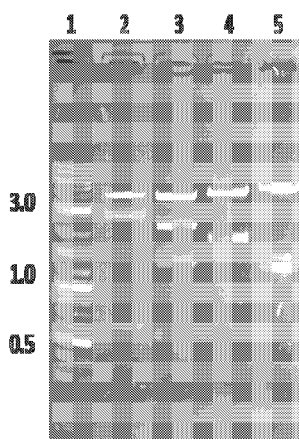
(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,
HR, HU, ID, IL, IN, IR, IS,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

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ,
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

(54) Title: CODON OPTIMIZED RPGRORF 15 GENES AND USES THEREOF

(57) Abstract: The present disclosure provides codon optimized
RPGRorf15 sequences, vectors, and host cells comprising codon
optimized RPGRorf15 sequences, and methods of treating retinal
disorders such as XLRP comprising administering to the subject a
codon optimized RPGRorf15 sequence.

FIGURE 1



TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *of inventorship (Rule 4.17(iv))*

Published:

- *with international search report (Art. 21(3))*
- *with sequence listing part of description (Rule 5.2(a))*

CODON OPTIMIZED RPGRORF15 GENES AND USES THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of United States Provisional Patent Application Serial No. 63/073,843, filed September 2, 2020, the full disclosure of which is incorporated herein by reference.

SEQUENCE LISTING SUBMISSION VIA EFS-WEB

[0002] A computer readable text file, entitled “090400-5012-WO-Sequence-Listing” created on or about August 11, 2021, with a file size of about 37 KB contains the sequence listing for this application and is hereby incorporated by reference in its entirety.

BACKGROUND OF THE INVENTION

[0003] X-linked retinitis pigmentosa (XLRP) is a relatively severe and genetically heterogenous inherited retinal degeneration. Approximately 70% of XLRP cases are caused by mutations in the Retinitis Pigmentosa GTPase Regulator (*RPGR*) gene. The *RPGR* gene encodes several distinct alternatively –spliced transcripts that are widely expressed. The function of the encoded protein is not well understood, but studies suggest that it plays an important role in cell structures called cilia.

[0004] One *RPGR* isoform contains a unique 3' region called ORF15, a Gly- and Glu-rich carboxyl terminal domain of 567 amino acids. This version of the *RPGR* protein, containing exons 1-13 of the *RPGR* gene and the ORF15 region, is expressed predominantly in photoreceptors in the retina. Mutations in the ORF15 region of *RPGR* account for about 60% of all XLRP cases.

[0005] Several preclinical studies support the use of wild type cDNA of *RPGRorf15* to rescue the XLRP disease phenotype. However, poor sequence stability of the wild type sequence poses challenges to maintaining sequence integrity during vector production and suboptimal expression level of the wild type sequence in human photoreceptors are challenges to gene therapy approaches to treat XLRP.

SUMMARY OF THE INVENTION

[0006] Disclosed are codon optimized nucleic acid molecules encoding a human retinitis pigmentosa GTPase regulator (RPGR) protein. In one aspect, the disclosure provides a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 or a nucleic acid comprising a nucleotide sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to the nucleotide sequence of SEQ ID NO:1 and which encodes a human RPGR polypeptide having the amino acid sequence of SEQ ID NO:2. In some embodiments, a nucleic acid comprising or consisting of the nucleotide sequence of SEQ ID NO:1 is provided. In related embodiments, the nucleic acid is expressed at a higher level compared with the level of expression of a wild type RPGR nucleic acid sequence (e.g. SEQ ID NO:3) in an otherwise identical cell.

[0006a] In one embodiment of the present invention, there is provided a nucleic acid encoding human retinitis pigmentosa GTPase regulator (RPGR) protein of SEQ ID NO:2 and codon optimized for expression in humans, the nucleic acid comprising the nucleotide sequence set forth as SEQ ID NO: 1 or comprising a nucleotide sequence at least 95% identical thereto.

[0007] In some aspects, a codon optimized nucleic acid molecule as herein described has a human codon adaptation index that is increased relative to that of the wild type RPGR cDNA (GenBank Accession No. NM_001034853; SEQ ID NO:3). In some embodiments, the codon optimized nucleic acid molecule has a human codon adaptation index of at least about 0.85, at least about 0.88, or at least about 0.89.

[0008] In certain embodiments, the nucleic acid contains a higher percentage of G/C nucleotides compared to the percentage of G/C nucleotides in SEQ ID NO:3. In other embodiments, the nucleic acid contains a percentage of G/C nucleotides that is at most about 59%, at most about 58%, or at most about 57%. In some aspects, the average G/C content of the nucleic acid is from about 55% to about 59%, from about 56% to about 58%. In some preferred embodiments, the average G/C content is about 57%.

[0009] In other embodiments, the nucleic acid comprises one or more optimized parameters relative to SEQ ID NO:3 selected from removal of negative cis-acting

sites including without limitation TATA-boxes and splice sites, and increasing the frequency of optimal codons.

[0010] In another embodiment, the nucleic acid is operatively linked to at least one transcription control sequence, preferably a transcription control sequence that is heterologous to the nucleic acid. In some aspects, the transcription control sequence is a cell- or tissue-specific promoter that results in cell-specific expression of the nucleic acid e.g. in

[TEXT CONTINUED ON PAGE 3]

photoreceptor cells such as human rod photoreceptor-specific human G-protein coupled receptor rhodopsin kinase 1 (hGRK) promoter or a human interphotoreceptor retinoid-binding protein (IRBP) promoter. In preferred embodiments, the transcription control sequence comprises a human rod photoreceptor-specific human G-protein coupled receptor rhodopsin kinase 1 (hGRK) promoter. In other aspects, the transcription control sequence is a constitutive promoter that results in similar expression level of the nucleic acid in many cell types (e.g. a CAG, CBA, CMV, or PGK promoter). In preferred embodiments, the transcription control sequence comprises a human G protein-coupled receptor kinase (hGRK, also known as Rhodopsin Kinase) promoter as described in Young *et al.*, Investigative Ophthalmology and Visual Science, 44(9):4076-4085 (2003). In a particularly preferred embodiment, the hGRK promoter comprises the sequence of SEQ ID NO:4 or comprises a sequence at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical thereto:

GGGCCCCAGAAGCCTGGTGGTTGTTTGTCCCTTCTCAGGGGAAAAGTGAGGCGGC
 CCCTTGGAGGAAGGGGCGGGCAGAATGATCTAATCGGATTCGAAGCAGCTCAG
 GGGATTGTCTTTTTCTAGCACCTTCTTGCCACTCCTAAGCGTCCTCCGTGACCCCG
 GCTGGGATTTAGCCTGGTGTGTGTCAGCCCCGGG (SEQ ID NO:4)

[0011] In related embodiments, provided herein is an expression cassette comprising a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, or a nucleotide sequence at least 90% identical thereto, operably linked to an expression control sequence.

[0012] In related embodiments, provided herein is a vector comprising a comprising a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, or a nucleotide sequence at least 90% identical thereto. In preferred embodiments, the vector is a recombinant adeno-associated (rAAV) expression vector. In some embodiments, the rAAV vector comprises a native capsid (e.g. a capsid of AAV serotype 2 or AAV serotype 5 or AAV serotype 8). In other embodiments, the rAAV vector comprises a capsid that is modified (e.g. comprises one or more peptide insertions and/or one or more amino acid substitutions (e.g. tyrosine to phenylalanine) and/or amino acid insertions or amino acid deletions) relative to a native AAV capsid (e.g. comprising one or more modifications relative to an AAV capsid of serotype 2, 5 or 8).

[0013] In another embodiment, provided herein is a host cell comprising a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, or a nucleotide sequence at least 90% identical thereto. In some aspects, the host cell is a mammalian cell, including without limitation, a CHO cell, an HEK293 cell, a HeLa cell, a BHK21 cell, a Vero cell or a V27 cell. In related aspects, the host cell is selected from a CHO cell, an HEK293 cell, an HEK293T cell, a HeLa cell, a BHK21 cell and a Vero cell. In other aspects, the host cell is a photoreceptor cell (e.g. rods; cones), a retinal ganglion cell (RGC), a glial cell (e.g. a Müller glial cell, a microglial cell), a bipolar cell, an amacrine cell, a horizontal cell, or a retinal pigmented epithelium (RPE) cell. In related embodiments, the disclosure provides a method of increasing expression of a polypeptide of SEQ ID NO: 2 comprising culturing the host cell under conditions whereby a polypeptide of SEQ ID NO: 2 is expressed by the nucleic acid molecule, wherein the expression of the polypeptide is increased relative to a host cell cultured under the same conditions comprising a reference nucleic acid comprising the nucleotide sequence of SEQ ID NO:3 (comparator sequence).

[0014] In another embodiment, the disclosure provides a method of increasing expression of a polypeptide of SEQ ID NO: 2 in a human subject comprising administering to the subject an isolated nucleic acid molecule comprising a nucleotide sequence at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to the nucleotide sequence of SEQ ID NO:1 and which encodes a polypeptide having the amino acid sequence of SEQ ID NO:2 or a vector comprising such a nucleotide sequence, wherein the expression of the polypeptide is increased relative to a reference nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:3 (comparator sequence) or a vector comprising the reference nucleic acid molecule.

[0015] In some embodiments, the disclosure provides a method of treating an ocular disorder associated with insufficient RGRP ORF15 activity in a human subject comprising administering to the subject a nucleic acid molecule or a vector disclosed herein. In some embodiments, the retinal disorder is X-linked retinitis pigmentosa.

[0015a] In another embodiment, there is provided the use of an infectious rAAV comprising (i) an AAV capsid and (ii) a nucleic acid comprising from 5' to 3': (a) an AAV2 terminal repeat (b) an hGRK promoter (c) a nucleotide sequence at least 95% identical to the nucleotide sequence as set forth in SEQ ID NO:1 (d) an SV40 polyadenylation sequence and (e) an AAV2 terminal repeat, in the manufacture of a medicament for the treatment of XLRP.

DESCRIPTION OF THE DRAWINGS

[0016] **Figure 1** illustrates gel electrophoresis of restriction digests of pAAV-GRK-cohRPGRorf15-SV40. Maxiprep DNA was digested with various enzymes and analyzed by

agarose gel electrophoresis: Lane 1 = 2-log ladder; Lane 2 = BsrGI-H+BglII; Lane 3 = Pml +Sph-HF; Lane 4 = HindIII-HF + Sph-HF; Lane 5 = Pst. The resulting restriction fragments matched the predicted fragments in all digests (Lane 2 fragments of 3.9, 2.5, 0.6 kb; Lane 3 fragments of 3.7, 2.1, 1.3 kb; Lane 4 fragments of 3.9, 1.7 and 1.5 kb; Lane 5 fragments of 4.6, 1.4 and 1.2 kb). The sizes of the prominent 2-log ladder bands in kilobase pairs are indicated to the left of the gel.

[0017] **Figure 2** is Western Blot of cell lysates from HEK293T cells transfected with pAAV-GRK-cohRPGRorf15-SV40. Expression of human RPGRorf15 protein in HEK293 cells was assessed with the indicated primary antibodies (Sigma; CT-15; Polyglut GT335) For each antibody, lane 1 = untransfected control; lane 2 = pAAV-GRK-cohRPGRorf15-SV40; lane 3 = pAAV-PGK-cohRPGRorf15-SV40. The arrows indicate hRPGRorf15 protein. Molecular weight marker (in kilodaltons) is shown on the left-hand side.

[0018] **Figure 3** Transduction with recombinant AAV (rAAV) virions comprising codon optimized RPGRorf15 of SEQ ID NO:1 under the control of an hGRK1 promoter leads to a robust increase of cohRPGRorf15 (SEQ ID NO:1) transcript levels in XLRP-iPSC-derived photoreceptor cells. Digital droplet PCR was performed on RNA extracted from XLRP-iPSC derived photoreceptor cultures following transduction with rAAV comprising pAAV-GRK-cohRPGRorf15-SV40 and capsid of SEQ ID NO:9 at MOI of 50,000, thirty days post transduction. hRPGR1-19 (internal control) and cohRPGRorf15 transcript levels were determined and quantified as copies/mL above a set threshold and plotted on a log scale. Following transduction, codon optimized hRPGRorf15 (SEQ ID NO:1) transcript level was statistically greater than hRPGR1-19. NT=non-transduced, MOI=multiplicity of infection, hRPGR1-19=human retinitis pigmentosa GTPase regulator exon 1-19, constitutive isoform, cohRPGRorf15 = codon optimized human retinitis pigmentosa GTPase regulator open reading frame 15, retinal specific isoform of SEQ ID NO:1. * $p \leq 0.05$ compared to MOI 50,000 hRPGR1-19, † $p \leq 0.05$ compared to NT cohRPGRorf15. Error bars \pm Standard Deviation. n=3 per Patient. Y-axis in log scale.

[0019] **Figure 4** Transduction with rAAV comprising codon optimized RPGRorf15 of SEQ ID NO:1 under the control of an hGRK promoter increases hRPGRorf15 protein levels in XLRP photoreceptor cultures. XLRP-iPSC derived photoreceptor cultures were transduced at MOI of 50,000 and protein lysates were harvested 30 days post transduction.

SDS-PAGE and Western blot showed an increase in hRPGRorf15, at 127 kDa, compared to non-transduced cells (NT) for both patients, normalized to the loading control α -tubulin. Band intensity was quantified and averaged between patients. Transduction with rAAV yielded a significant increase in hRPGRorf15 protein. * $p \leq 0.05$ compared to NT. Error bars \pm Standard Deviation. n=3 per Patient.

[0020] Figure 5 Glutamylation of hRPGRorf15 following transduction with rAAV comprising codon optimized RPGRorf15 of SEQ ID NO:1 under the control of an hGRK promoter in XLRP photoreceptor cultures. XLRP-iPSC derived photoreceptor cultures were transduced at a MOI of 50,000 and protein lysates were harvested 30 days post transduction. SDS-PAGE and Western blot analyses showed an increase in glutamylation of a 127kDa protein, hRPGRorf15, compared to nontransduced (NT) control for both patients, normalized to the loading control, α -tubulin band intensity was quantified and averaged between patients. Transduction with rAAV yielded a significant increase in glutamylation of hRPGRorf15 protein. GT335= anti-glutamylation antibody, NT= non-transduced, MOI= Multiplicity of Infection, hRPGRorf15= human Retinitis Pigmentosa GTPase Regulator Open Reading Frame 15, retinal specific isoform. * $p \leq 0.05$ compared to NT. Error bars \pm Standard Deviation. n=3 per Patient.

[0021] Figure 6 Constitutive promoter drives increase in hRPGRorf15 protein and glutamylation in XLRP photoreceptor cultures. XLRP-iPSC derived photoreceptor cultures were transduced with rAAV comprising codon optimized RPGRorf15 of SEQ ID NO:1 under the control of a PGK promoter at MOIs of 5,000, 10,000 and 20,000. Protein lysates were harvested 30 days post transduction. SDS-PAGE and Western blot showed an increase in hRPGRorf15, and glutamylation at 127kDa, compared to non-transduced (NT) control for Patient 78, normalized to the loading control, α -tubulin. Band intensity was quantified. Transduction yielded a significant increase in hRPGRorf15 protein. NT= non-transduced, MOI= Multiplicity of Infection, hRPGRorf15= human Retinitis Pigmentosa GTPase Regulator Open Reading Frame 15, retinal specific isoform, GT335= anti-glutamylation antibody. * $p \leq 0.05$ compared to NT. Error bars \pm Standard Deviation. n=3.

[0022] Figure 7 is the codon optimized sequence of SEQ ID NO:1 and the encoded amino acid sequence.

[0023] **Figure 8** is a schematic of the transgene cassette contained within the rAAV described in the Examples below. The transgene cassette comprises a 5' AAV2 ITR, a human rhodopsin kinase (aka hGRK) Promoter, a Codon Optimized Human RPGRorf15 cDNA of SEQ ID NO:1, a late SV40 Polyadenylation Signal, and a 3' AAV2 ITR and has the nucleotide sequence of SEQ ID NO:5.

[0024] **Figure 9** illustrates safety of 4D-125 (comprising the transgene cassette shown in Figure 8 and a capsid protein of SEQ ID NO:9) through quantification of ocular inflammation, as assessed by aqueous flare, aqueous cells, and vitreous cells. Ophthalmoscopic signs of transient mild ocular inflammation were observed at the high dose. These changes responded to an increase in the systemic steroid treatment. There were no adverse findings considered related to 4D-125. IOP values were within normal limits for all animals at the different examination intervals. ERG values and OCT images including macular morphology were also within normal limits.

[0025] **Figure 10** illustrates vector genome biodistribution in selected retinal, ocular, and non-ocular tissues, as measured by qPCR at 3 necropsy timepoints in NHPs intravitreally administered 4D-125. LOD = lower limit of detection; all samples "BLOD" graphed at LOD value for visualization purposes.

[0026] **Figure 11** illustrates RPGR transgene mRNA expression in selected retinal, ocular, and non-ocular tissues, as measured by RT-qPCR at 3 necropsy timepoints in NHPs intravitreally administered 4D-125. LOD = lower limit of detection; all samples "BLOD" graphed at LOD value for visualization purposes.

DETAILED DESCRIPTION OF THE INVENTION

[0027] Definitions

[0028] A "codon adaptation index," as used herein, refers to a measure of codon usage bias. A codon adaptation index (CAI) measures the deviation of a given protein coding gene sequence with respect to a reference set of genes (Sharp P M and Li W H, *Nucleic Acids Res.* 15(3):1281-95 (1987)). CAI is calculated by determining the geometric mean of the weight associated to each codon over the length of the gene sequence (measured in codons):

$$CAI = \exp\left(1/L \sum_{l=1}^L \ln(w_l(l))\right), \quad (I)$$

For each amino acid, the weight of each of its codons, in CAI, is computed as the ratio between the observed frequency of the codon (f_i) and the frequency of the synonymous codon (f_j) for that amino acid:

$$w_i = \frac{f_i}{\max(f_j)} \quad ij \in [\text{synonymous codons for amino acid}] \quad (II)$$

[0029] The term "isolated" designates a biological material (cell, nucleic acid or protein) that has been removed from its original environment (the environment in which it is naturally present). For example, a polynucleotide present in the natural state in a plant or an animal is not isolated, however the same polynucleotide separated from the adjacent nucleic acids in which it is naturally present, is considered "isolated."

[0030] The term "4D-125" refers to a recombinant AAV particle comprising (i) a capsid protein comprising the amino acid sequence of SEQ ID NO:9 and a heterologous nucleic acid comprising the nucleotide sequence of SEQ ID NO:5.

[0031] The term "R100" refers to a variant AAV capsid protein comprising the amino acid sequence of SEQ ID NO:9.

[0032] As used herein, a "coding region" or "coding sequence" is a portion of polynucleotide which consists of codons translatable into amino acids. Although a "stop codon" (TAG, TGA, or TAA) is typically not translated into an amino acid, it can be considered to be part of a coding region, but any flanking sequences, for example promoters, ribosome binding sites, transcriptional terminators, introns, and the like, are not part of a coding region. The boundaries of a coding region are typically determined by a start codon at the 5' terminus, encoding the amino terminus of the resultant polypeptide, and a translation stop codon at the 3' terminus, encoding the carboxyl terminus of the resulting polypeptide. Two or more coding regions can be present in a single polynucleotide construct, e.g., on a single vector, or in separate polynucleotide constructs, e.g., on separate (different) vectors. It

follows, then that a single vector can contain just a single coding region, or comprise two or more coding regions.

[0033] As used herein, the term "regulatory region" refers to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding region, and which influence the transcription, RNA processing, stability, or translation of the associated coding region. Regulatory regions can include promoters, translation leader sequences, introns, polyadenylation recognition sequences, RNA processing sites, effector binding sites and stem-loop structures. If a coding region is intended for expression in a eukaryotic cell, a polyadenylation signal and transcription termination sequence will usually be located 3' to the coding sequence.

[0034] As used herein, the term "nucleic acid" is interchangeable with "polynucleotide" or "nucleic acid molecule" and a polymer of nucleotides is intended.

[0035] A polynucleotide which encodes a gene product, e.g., a polypeptide, can include a promoter and/or other transcription or translation control elements operably associated with one or more coding regions. In an operable association a coding region for a gene product, e.g., a polypeptide, is associated with one or more regulatory regions in such a way as to place expression of the gene product under the influence or control of the regulatory region(s). For example, a coding region and a promoter are "operably associated" if induction of promoter function results in the transcription of mRNA encoding the gene product encoded by the coding region, and if the nature of the linkage between the promoter and the coding region does not interfere with the ability of the promoter to direct the expression of the gene product or interfere with the ability of the DNA template to be transcribed. Other transcription control elements, besides a promoter, for example enhancers, operators, repressors, and transcription termination signals, can also be operably associated with a coding region to direct gene product expression.

[0036] "Transcriptional control sequences" refer to DNA regulatory sequences, such as promoters, enhancers, terminators, and the like, that provide for the expression of a coding sequence in a host cell. A variety of transcription control regions are known to those skilled in the art. These include, without limitation, transcription control regions which function in vertebrate cells, such as, but not limited to, promoter and enhancer segments from

cytomegaloviruses (the immediate early promoter, in conjunction with intron-A), simian virus 40 (the early promoter), and retroviruses (such as Rous sarcoma virus). Other transcription control regions include those derived from vertebrate genes such as actin, heat shock protein, bovine growth hormone and rabbit beta-globin, as well as other sequences capable of controlling gene expression in eukaryotic cells. Additional suitable transcription control regions include tissue-specific promoters and enhancers as well as lymphokine-inducible promoters (e.g., promoters inducible by interferons or interleukins).

[0037] Similarly, a variety of translation control elements are known to those of ordinary skill in the art. These include, but are not limited to ribosome binding sites, translation initiation and termination codons, and elements derived from picornaviruses (particularly an internal ribosome entry site, or IRES, also referred to as a CITE sequence).

[0038] The term "expression" as used herein refers to a process by which a polynucleotide produces a gene product, for example, an RNA or a polypeptide. It includes without limitation transcription of the polynucleotide into messenger RNA (mRNA), transfer RNA (tRNA), small hairpin RNA (shRNA), small interfering RNA (siRNA) or any other RNA product, and the translation of an mRNA into a polypeptide. Expression produces a "gene product." As used herein, a gene product can be either a nucleic acid, e.g., a messenger RNA produced by transcription of a gene, or a polypeptide which is translated from a transcript. Gene products described herein further include nucleic acids with post transcriptional modifications, e.g., polyadenylation or splicing, or polypeptides with post translational modifications, e.g., methylation, glycosylation, the addition of lipids, association with other protein subunits, or proteolytic cleavage.

[0039] A "vector" refers to any vehicle for the cloning of and/or transfer of a nucleic acid into a host cell. A vector can be a replicon to which another nucleic acid segment can be attached so as to bring about the replication of the attached segment. The term "vector" includes both viral and nonviral vehicles for introducing the nucleic acid into a cell in vitro, ex vivo or in vivo. A large number of vectors are known and used in the art including, for example, plasmids, modified eukaryotic viruses, or modified bacterial viruses. Insertion, of a polynucleotide into a suitable vector can be accomplished by ligating the appropriate polynucleotide fragments into a chosen vector that has complementary cohesive termini.

[0040] Vectors can be engineered to encode selectable markers or reporters that provide for the selection or identification of cells that have incorporated the vector. Expression of selectable markers or reporters allows identification and/or selection of host cells that incorporate and express other coding regions contained on the vector. Examples of selectable marker genes known and used in the art include: genes providing resistance to ampicillin, streptomycin, gentamycin, kanamycin, hygromycin, bialaphos herbicide, sulfonamide, and the like; and genes that are used as phenotypic markers, i.e., anthocyanin regulatory genes, isopentenyl transferase gene, and the like. Examples of reporters known and used in the art include: luciferase (Luc), green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), -galactosidase (LacZ), -glucuronidase (Gus), and the like. Selectable markers can also be considered to be reporters.

[0041] Eukaryotic viral vectors that can be used include, but are not limited to, adenovirus vectors, retrovirus vectors, adeno-associated virus vectors, poxvirus, e.g., vaccinia virus vectors, baculovirus vectors, or herpesvirus vectors. Non-viral vectors include plasmids, liposomes, electrically charged lipids (cytofectins), DNA-protein complexes, and biopolymers.

[0042] "Promoter" and "promoter sequence" are used interchangeably and refer to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. Promoters can be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters can direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental or physiological conditions. Promoters that cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters." Promoters that cause a gene to be expressed in a specific cell type are commonly referred to as "cell-specific promoters" or "tissue-specific promoters." Promoters that cause a gene to be expressed at a specific stage of development or cell differentiation are commonly referred to as "developmentally-specific promoters" or "cell differentiation-specific promoters." Promoters that are induced and cause a gene to be expressed following exposure or treatment of the cell with an agent, biological molecule, chemical, ligand, light, or the like that induces the promoter are commonly referred to as "inducible promoters" or "regulatable promoters."

It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths can have identical promoter activity.

[0043] The term "plasmid" refers to an extra-chromosomal element often carrying a gene that is not part of the central metabolism of the cell, and usually in the form of circular double-stranded DNA molecules. Such elements can be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear, circular, or supercoiled, of a single- or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing a promoter fragment and DNA sequence for a selected gene product along with appropriate 3' untranslated sequence into a cell.

[0044] A polynucleotide or polypeptide has a certain percent "sequence identity" to another polynucleotide or polypeptide, meaning that, when aligned, that percentage of bases or amino acids are the same when comparing the two sequences. Sequence similarity can be determined in a number of different manners. To determine sequence identity, sequences can be aligned using the methods and computer programs, including BLAST, available over the world wide web at ncbi.nlm.nih.gov/BLAST/. Another alignment algorithm is FASTA, available in the Genetics Computing Group (GCG) package, from Madison, Wis., USA. Other techniques for alignment are described in *Methods in Enzymology*, vol. 266: *Computer Methods for Macromolecular Sequence Analysis* (1996), ed. Doolittle, Academic Press, Inc. Of particular interest are alignment programs that permit gaps in the sequence. The Smith-Waterman is one type of algorithm that permits gaps in sequence alignments. See *Meth. Mol. Biol.* 70: 173-187 (1997). Also, the GAP program using the Needleman and Wunsch alignment method can be utilized to align sequences. See *J. Mol. Biol.* 48: 443-453 (1970).

[0045] In one embodiment, the present invention provides a modified nucleic acid molecule comprising a nucleotide sequence that encodes a polypeptide of SEQ ID NO:2 (human RGPGR ORF15), wherein the nucleic acid sequence has been codon optimized. In another embodiment, the starting nucleic acid sequence that encodes a polypeptide of SEQ ID NO:2 and that is subject to codon optimization has the nucleotide sequence set forth as SEQ ID NO:3. In preferred embodiments, the sequence that encodes a polypeptide of SEQ ID

NO:2 is codon optimized for human expression. SEQ ID NO:1 is a codon optimized version of SEQ ID NO:3, optimized for human expression:

ATGAGAGAACCCGAGGAACTGATGCCCGACTCTGGCGCCGTGTTTACCTTCGGC
AAGAGCAAGTTCGCCGAGAACAACCCCGGCAAGTTCTGGTTCAAGAACGACGTG
CCAGTGCACCTGAGCTGCGGAGATGAACACTCTGCCGTGGTCACCGGCAACAAC
AAGCTGTACATGTTTCGGCAGCAACAACCTGGGGCCAGCTCGGCCTGGGATCTAAG
TCTGCCATCAGCAAGCCTACCTGCGTGAAGGCCCTGAAGCCTGAGAAAGTGAAA
CTGGCCGCCTGCGGCAGAAATCACACCCTGGTTTCTACCGAAGGCGGCAATGTGT
ATGCCACCGGCGGAAACAATGAGGGACAGCTTGGACTGGGCGACACCGAGGAA
AGAAACACCTTCCACGTGATCAGCTTTTTTACCAGCGAGCACAAGATCAAGCAG
CTGAGCGCCGGCTCTAATACCTCTGCCGCTCTGACAGAGGACGGCAGACTGTTTA
TGTGGGGCGACAATTCTGAGGGCCAGATCGGACTGAAGAACGTGTCCAATGTGT
GCGTGCCCCAGCAAGTGACAATCGGCAAGCCTGTGTCTTGGATCAGCTGCGGCT
ACTACCACAGCGCCTTTGTGACAACCGATGGCGAGCTGTATGTGTTTCGGCGAGCC
AGAGAATGGCAAGCTGGGACTGCCTAACCAGCTGCTGGGCAATCACAGAACCCC
TCAGCTGGTGTCTGAGATCCCCGAAAAAGTGATCCAGGTGGCCTGTGGCGGAGA
GCACACAGTGGTGCTGACAGAGAATGCCGTGTACACCTTTGGCCTGGGCCAGTTT
GGACAACCTCGGACTGGGAACCTTCTGTTTCGAGACAAGCGAGCCC AAAGTGATC
GAGAACATCCGGGACCAGACCATCAGCTACATCAGCTGTGGCGAGAACCACACA
GCCCTGATCACAGACATCGGCCTGATGTACACATTTCGGCGACGGAAGGCATGGA
AAGCTCGGACTTGGCCTGGAAAACCTTACCAACCACTTCATCCCTACGCTGTGCA
GCAACTTCTGCGGTTTCATTGTGAAGCTGGTGGCCTGCGGAGGATGCCACATGGT
GGTTTTTGTGCTGCCCTCACAGAGGCGTGGCCAAAGAGATTGAGTTCGACGAGATC
AACGATACCTGCCTGAGCGTGGCCACCTTCTGCCTTACAGCAGCCTGACATCTG
GCAACGTGCTGCAGAGGACACTGAGCGCCAGAATGCGCAGACGGGAAAGAGAG
AGAAGCCCCGACAGCTTCAGCATGAGAAGAACCCTGCCTCCAATCGAGGGCACA
CTGGGCCTGTCTGCCTGCTTTCTGCCTAACAGCGTGTTCCTCCAGATGCAGCGAGA
GAAACCTGCAAGAGAGCGTGCTGAGCGAGCAGGATCTGATGCAGCCTGAGGAAC
CCGACTACCTGCTGGACGAGATGACCAAAGAGGCCGAGATCGACAACAGCAGCA
CAGTGGAAAGCCTGGGCGAGACAACCGACATCCTGAACATGACCCACATCATGA
GCCTGAACAGCAACGAGAAGTCTCTGAAGCTGAGCCCCGTGCAGAAGCAGAAGA
AGCAGCAGACCATCGGCGAGCTGACACAGGATACTGCCCTGACCGAGAACGACG
ACAGCGACGAGTACGAAGAGATGAGCGAGATGAAGGAAGGCAAGGCCTGCAAG

CAGCACGTGTCCCAGGGCATCTTTATGACCCAGCCTGCCACCACCATCGAGGCCT
TTTCCGACGAGGAAGTGGAAATCCCCGAGGAAAAAGAGGGCGCCGAGGACAGC
AAAGGCAACGGCATTGAGGAACAAGAGGTGGAAGCCAACGAAGAGAACGTGAA
GGTGCACGGCGGACGGAAAGAAAAGACCGAGATCCTGAGCGACGACCTGACCG
ATAAGGCCGAGGTTTCCGAGGGCAAAGCCAAGTCTGTGGGAGAAGCCGAGGATG
GACCTGAAGGCCGCGGAGATGGAACCTGTGAAGAAGGATCTAGCGGAGCCGAG
CACTGGCAGGATGAGGAACGCGAGAAGGGCGAGAAAGACAAAGGCAGAGGCGA
GATGGAAAGACCCGGCGAGGGCGAAAAAGAGCTGGCCGAGAAAGAGGAATGGA
AGAAACGCGACGGCGAAGAACAAGAGCAGAAAGAAAGAGAGCAGGGCCACCA
GAAAGAACGGAATCAAGAGATGGAAGAAGGCGGCGAGGAAGAACACGGCGAA
GGGGAAGAAGAGGAAGGCGACCGAGAGGAAGAAGAAGAGAAAGAAGGCGAAG
GCAAAGAAGAAGGCGAGGGCGAAGAGGTGGAAGGCGAGCGTGAAAAAGAAGA
GGGCGAACGCAAGAAAGAAGAACGCGCCGGAAAAGAGGAAAAGGCGAGGAA
GAGGGCGACCAAGGCGAAGGCGAGGAAGAAGAACTGAAGGCAGAGGGGAAG
AGAAAGAGGAAGGCGGCGAAGTCGAAGGCGGAGAGGTTGAAGAAGGCAAAGG
CGAGCGAGAAGAGGAAGAAGAAGAAGGCGAAGGCGAGGAAGAGGAAGGCGAA
GGCGAAGAGGAAGAAGGCGAAGGGGAAGAAGAAGAAGGCGAAGGCAAGGGCG
AAGAGGAGGGCGAAGAAGGCGAGGGCGAAGAGGAGGGCGAAGAAGGCGAAGG
CGAGGGCGAAGAAGAAGAAGGCGAAGGCGAAGGCGAGGAAGAAGGCGAAGGC
GAAGGGGAAGAAGAGGAAGGCGAAGGCGAAGGCGAAGAAGAAGGCGAAGGCG
AGGGCGAAGAGGAAGAAGGCGAAGGCAAAGGGGAAGAAGAAGGCGAGGAAGG
CGAAGGCGAAGGCGAGGAAGAAGAAGGCGAAGGCGAGGGCGAAGATGGCGAA
GGCGAAGGCGAAGAGGAAGAGGGCGAGTGGGAGGGCGAAGAAGAGGAAGGCG
AAGGCGAGGGCGAAGAGGAAGGCGAAGGCGAGGGCGAAGAAGGCGAAGGCGA
AGGCGAGGAAGAGGAAGGCGAAGGCGAAGGGGAAGAAGAAGAGGGCGAAGAA
GAAGGCGAAGAGGAAGGCGAAGGGGAAGAAGAAGGCGAAGGCGAAGGCGAAG
AAGAGGAAGAGGGCGAAGTTGAAGGCGAGGTTGAGGGCGAAGAAGGCGAAGGC
GAAGGGGAAGAAGAAGAAGGCGAGGAAGAAGGGGAAGAGAGAGAAAAAGAAG
GCGAGGGCGAAGAAAACCGCCGGAACCGCGAAGAGGAAGAGGAAGAAGAGGG
CAAGTACCAAGAGACTGGCGAGGAAGAGAACGAGCGGCAGGATGGCGAAGAGT
ACAAGAAGGTGTCCAAGATCAAGGGCAGCGTGAAGTACGGCAAGCACAAGACC
TACCAGAAGAAGTCCGTCACCAACACGCAAGGCAATGGAAAAGAACAGCGGAG
CAAGATGCCCGTGCAGTCCAAGAGGCTGCTGAAGAATGGCCCTAGCGGCAGCAA

GAAATTCTGGAACAATGTGCTGCCCCACTACCTCGAGCTGAAGTGA (SEQ ID NO:1)

[0046] In some embodiments, a codon-optimized sequence encoding human RPGR ORF15 is provided lacking the TGA stop codon of SEQ ID NO:1 (i.e. consisting of nucleotides 1-3456 of SEQ ID NO:1).

[0047] In one aspect, the disclosure provides a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 or polynucleotide comprising a nucleotide sequence at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to the nucleotide sequence of SEQ ID NO:1 and which encodes a human RPGR polypeptide having the amino acid sequence of SEQ ID NO:2:

MREPEELMPDSGAVFTFGKSKFAENNP GKFWFKNDVPVHLS CGDEHS AVVTGNNK
LYMFGSNNWGQLGLGSKSAISKPTCVKALKPEKVKLAACGRNHTLVSTEGGNVYAT
GGNNEGQLGLGDTEERNTFHVISFFTSEHKIKQLSAGSNTSAALTEDGRLFMWGDNS
EGQIGLKNVSNVCVPQQVTIGKPVSWISCGYYHSAFVTTDGELYVFGEPENGLGLP
NQLLGNHRTPQLVSEIPEKVIQVACGGEHTVVL TENAVYTFGLGQFGQLGLGTFLFE
TSEPKVIENIRDQTISYISCGENHTALITDIGLMYTFGDGRHGKLGLENFTNHFIPTL
CSNFLRFIVKLVACGGCHMVVFAAPHRGVAKEIEFDEINDTCLSVATFLPYSSLTSGN
VLQRTLSARMRRRERERSPDSFSMRRTLPPIEGTLGLSACFLPNSVFPRCSERNLQESV
LSEQDLMQPEEPDYLLDEMTKEAEIDNSSTVESLGETTDILNMTHIMSLNSNEKSLKL
SPVQKQKKQQTIGELTQDTALTENDDSD EYEEMSEMKEGKACKQHVSQGIFMTQPA
TTIEAFSDEEVEIPEEKEGAEDSKNGIEEQEVEANEENVKVHGGRKEKTEILSDDL T
DKAEVSEGKAKSVGEAEDGPEGRGDGTCEEGSSGAEHWQDEEREKGEKDKGRGEM
ERPGEGEKELAEKEEWKKRDGEEQE QKEREQGHQKERNQEMEEGGEEEHGEGEEE
EGDREEEEEKEGEGKEEGEGEEVEGEREKEEGERKKEERAGKEEKGEEEGDQGEGE
EEETEGRGEEKKEEGGEVEGGEVEEGKGEREEEEEGEGEEEEEGEGEEEEEGEGEEE
EGKGEEEGEEGEGEEEGEGEGEEEEEGEGEGEEEEEGEGEGEEEEEGEGEGEEEEGE
GEEEEGEGKGEEEGEEGEGEGEEEEEGEGEGEDGEGEGEEEEGEWEGEEEEGEGEGEE
EGEGEGEEGEGEGEEEEEGEGEGEEEEEGEGEGEEEEEGEGEGEEEEEGEVEGEV
EGEEGEGEGEEEEEGEEEREKEGEGEENRRNREEEEEEGKYQETGEEENERQD G
EEYKKVSKIKGSVKYGKHKTYQKKSVTNTQGNNGKEQRSKMPVQSKRLLKNGPSGS
KKFWNNVLPHYLELK (SEQ ID NO:2)

[0048] The term "codon-optimized" as it refers to genes or coding regions of nucleic acid molecules for transformation of various hosts, refers to the alteration of codons in the gene or coding regions of the nucleic acid molecules to reflect the typical codon usage of the host organism without altering the polypeptide encoded by the DNA. Such optimization includes replacing at least one, or more than one, or a significant number, of codons with one or more codons that are more frequently used in the genes of that organism.

[0049] Deviations in the nucleotide sequence that comprises the codons encoding the amino acids of, any polypeptide chain allow for variations in the sequence coding for the gene. Since each codon consists of three nucleotides, and the nucleotides comprising DNA are restricted to four specific bases, there are 64 possible combinations of nucleotides, 61 of which encode amino acids (the remaining three codons encode signals ending translation). The "genetic code" which shows which codons encode which amino acids is reproduced herein as Table 1. As a result, many amino acids are designated by more than one codon. For example, the amino acids alanine and proline are coded for by four triplets, serine and arginine by six, whereas tryptophan and methionine are coded by just one triplet. This degeneracy allows for DNA base composition to vary over a wide range without altering the amino acid sequence of the proteins encoded by the DNA.

TABLE-US-00001 **TABLE 1** The Standard Genetic Code T C A G T TTT Phe (F) TCT Ser (S) TAT Tyr (Y) TGT Cys (C) TTC Phe (F) TCC Ser (S) TAC Tyr (Y) TGC TTA Leu (L) TCA Ser (S) TAA Stop TGA Stop TTG Leu (L) TCG Ser (S) TAG Stop TGG Trp (W) C CTT Leu (L) CCT Pro (P) CAT His (H) CGT Arg (R) CTC Leu (L) CCC Pro (P) CAC His (H) CGC Arg (R) CTA Leu (L) CCA Pro (P) CAA Gln (Q) CGA Arg (R) CTG Leu (L) CCG Pro (P) CAG Gln (Q) CGG Arg (R) A ATT Ile (I) ACT Thr (T) AAT Asn (N) AGT Ser (S) ATC Ile (I) ACC Thr (T) AAC Asn (N) AGC Ser (S) ATA Ile (I) ACA Thr (T) AAA Lys (K) AGA Arg (R) ATG Met (M) ACG Thr (T) AAG Lys (K) AGG Arg (R) G GTT Val (V) GCT Ala (A) GAT Asp (D) GGT Gly (G) GTC Val (V) GCC Ala (A) GAC Asp (D) GGC Gly (G) GTA Val (V) GCA Ala (A) GAA Glu (E) GGA Gly (G) GTG Val (V) GCG Ala (A) GAG Glu (E) GGG Gly (G)

[0050] Many organisms display a bias for use of particular codons to code for insertion of a particular amino acid in a growing peptide chain. Codon preference, or codon bias, differences in codon usage between organisms, is afforded by degeneracy of the genetic code,

and is well documented among many organisms. Codon bias often correlates with the efficiency of translation of messenger RNA (mRNA), which is in turn believed to be dependent on, inter alia, the properties of the codons being translated and the availability of particular transfer RNA (tRNA) molecules. The predominance of selected tRNAs in a cell is generally a reflection of the codons used most frequently in peptide synthesis. Accordingly, genes can be tailored for optimal gene expression in a given organism based on codon optimization.

[0051] Given the large number of gene sequences available for a wide variety of animal, plant and microbial species, the relative frequencies of codon usage have been calculated. Codon usage tables are available, for example, at the "Codon Usage Database" available at www.kazusa.or.jp/codon/ (visited Jun. 18, 2012). See Nakamura, Y., et al. Nucl. Acids Res. 28:292 (2000).

[0052] Randomly assigning codons at an optimized frequency to encode a given polypeptide sequence can be done manually by calculating codon frequencies for each amino acid, and then assigning the codons to the polypeptide sequence randomly. Additionally, various algorithms and computer software programs can be used to calculate an optimal sequence.

[0053] Non-Viral Vectors

[0054] In some embodiments, a non-viral vector (e.g. an expression plasmid) comprising a modified nucleic acid as herein described is provided. Preferably, the non-viral vector is a plasmid comprising a nucleic acid sequence of SEQ ID NO: 1, or a sequence at least 90% identical thereto.

[0055] Viral Vectors

[0056] In preferred embodiments, a viral vector comprising a modified (codon optimized) nucleic acid as herein described is provided. Preferably, the viral vector comprises a nucleic acid sequence of SEQ ID NO: 1, or a sequence at least 90% identical thereto, operably linked to an expression control sequence. Examples of suitable viral vectors include but are not

limited to adenoviral, retroviral, lentiviral, herpesvirus and adeno-associated virus (AAV) vectors.

[0057] In a preferred embodiment, the viral vector includes a portion of a parvovirus genome, such as an AAV genome with the rep and cap genes deleted and/or replaced by the modified RPGRorf15 gene sequence and its associated expression control sequences. The modified human RPGRorf15 gene sequence is typically inserted adjacent to one or two (i.e., is flanked by) AAV TRs or TR elements adequate for viral replication (Xiao et al., 1997, J. Virol. 71(2): 941-948), in place of the nucleic acid encoding viral rep and cap proteins. Other regulatory sequences suitable for use in facilitating tissue-specific expression of the modified RPGRorf15 gene sequence in the target cell may also be included.

[0058] In some preferred embodiments, the AAV viral vector comprises a nucleic acid comprising from 5' to 3': (a) an AAV2 terminal repeat (b) an hGRK promoter (c) a codon optimized RPGRorf15 gene as herein described (d) a polyadenylation sequence and (e) an AAV2 terminal repeat. In a particularly preferred embodiment, the AAV viral vector comprises a nucleic acid (transgene cassette) comprising the sequence of SEQ ID NO:5 or a sequence at least 90%, at least 95%, at least 98% or at least 99% identical thereto:

```

TTGGCCACTC CCTCTCTGCG CGCTCGCTCG CTCACTGAGG CCGGGCGACC AAAGGTCGCC 60
CGACGCCCGG GCTTTGCCCG GGCGCCTCA GTGAGCGAGC GAGCGCGCAG AGAGGGAGTG 120
GCCAACTCCA TCACTAGGGG TTCCTATCGA TTGAATTCCC CGGGGATCCG GGCCCCAGAA 180
GCCTGGTGGT TGTTTGTCCCT TCTCAGGGGA AAAGTGAGGC GGCCCCTTG AGGAAGGGGC 240
CGGGCAGAAT GATCTAATCG GATTCCAAGC AGCTCAGGGG ATTGTCTTTT TCTAGCACCT 300
TCTTGCCACT CCTAAGCGTC CTCCGTGACC CCGGCTGGGA TTAGCCTGG TGCTGTGTCA 360
GCCCCGGGTC TAGAGTCGAC CTGCAGAAAGC TTCCACCATG AGAGAACCCG AGGAAGTATG 420
GCCCGACTCT GGCGCCGTGT TTACCTTCGG CAAGAGCAAG TTCGCCGAGA ACAACCCCGG 480
CAAGTTCTGG TTCAAGAACG ACGTGCCAGT GCACCTGAGC TCGCGAGATG AACACTCTGC 540
CGTGGTCACC GGCAACAACA AGCTGTACAT GTTCGGCAGC AACAACTGGG GCCAGCTCGG 600
CCTGGGATCT AAGTCTGCCA TCAGCAAGCC TACCTGCGTG AAGGCCCTGA AGCCTGAGAA 660
AGTGAAGACTG GCCGCCTGCG GCAGAAATCA CACCCTGGTT TCTACCGAAG GCGGCAATGT 720
GTATGCCACC GGCGGAAACA ATGAGGGACA GCTTGGACTG GCGGACACCG AGGAAAGAAA 780
CACCTTCACG GTGATCAGCT TTTTCACCAG CGAGCACAAG ATCAAGCAGC TGAGCGCCGG 840
CTCTAATACC TCTGCCGCTC TGACAGAGGA CGGCAGACTG TTTATGTGGG GCGACAATTC 900
TGAGGGCCAG ATCGGACTGA AGAACGTGTC CAATGTGTGC GTGCCCCAGC AAGTGACAAT 960
CGGCAAGCCT GTGTCTTGGG TCAGCTGCGG CTAATACCAC AGCGCCTTTG TGACAACCGA 1020
TGGCGAGCTG TATGTGTTCC GCGAGCCAGA GAATGGCAAG CTGGGACTGC CTAACCAGCT 1080
GCTGGGCAAT CACAGAACCC CTCAGCTGGT GTCTGAGATC CCCGAAAAAG TGATCCAGGT 1140
GGCCTGTGGC GGAGAGCACA CAGTGGTGCT GACAGAGAAT GCCGTGTACA CCTTTGGCCT 1200
GGGCCAGTTT GGACAACTCG GACTGGGAAC CTTCTGTTC GAGACAAGCG AGCCCAAAGT 1260
GATCGAGAAC ATCCGGGACC AGACCATCAG CTACATCAGC TGTGGCGAGA ACCACACAGC 1320
CCTGATCACA GACATCGGCC TGATGTACAC ATTCGGCGAC GGAAGGCATG GAAAGCTCGG 1380
ACTTGGCCTG GAAAACCTCA CCACCCTT CATCCCTACG CTGTGCAGCA ACTTCTGCG 1440
GTTCAATTGTG AAGCTGGTGG CCTGCGGAGG ATGCCACATG GTGGTTTTTTG CTGCCCTCA 1500
CAGAGGCGTG GCCAAAGAGA TTGAGTTCGA CGAGATCAAC GATACCTGCC TGAGCGTGGC 1560
CACCTTCCTG CCTTACAGCA GCCTGACATC TGGCAACGTG CTGCAGAGGA CACTGAGCGC 1620
CAGAAATGCGC AGACGGGAAA GAGAGAGAAG CCCCAGACAGC TTCAGCATGA GAAGAACCCT 1680
    
```

GCCTCCAATC GAGGGCACAC TGGGCTGTC TGCCTGCTTT CTGCCTAACA GCGTGTTCCT 1740
 CAGATGCAGC GAGAGAAACC TGCAAGAGAG CGTGCTGAGC GAGCAGGATC TGATGCAGCC 1800
 TGAGGAACCC GACTACCTGC TGGACGAGAT GACCAAAGAG GCCGAGATCG ACAACAGCAG 1860
 CACAGTGGAAG AGCCTGGGCG AGACAACCGA CATCCTGAAC ATGACCCACA TCATGAGCCT 1920
 GAACAGCAAC GAGAAGTCTC TGAAGCTGAG CCCCCTGCAG AAGCAGAAGA AGCAGCAGAC 1980
 CATCGGCGAG CTGACACAGG ATACTGCCCT GACCGAGAAC GACGACAGCG ACGAGTACGA 2040
 AGAGATGAGC GAGATGAAGG AAGGCAAGGC CTGCAAGCAG CACGTGTCCC AGGGCATCTT 2100
 TATGACCCAG CCTGCCACCA CCATCGAGGC CTTTTCCGAC GAGGAAGTGG AAATCCCCGA 2160
 GGAAAAGAG GGCGCCGAGG ACAGCAAAGG CAACGGCATT GAGGAACAAG AGGTGGAAGC 2220
 CAACGAAGAG AACGTGAAGG TGCACGGCGG ACGGAAAGAA AAGACCGAGA TCCTGAGCGA 2280
 CGACCTGACC GATAAGGCCG AGGTTTCCGA GGGCAAAGCC AAGTCTGTGG GAGAAGCCGA 2340
 GGATGGACCT GAAGGCCGCG GAGATGGAAC CTGTGAAGAA GGATCTAGCG GAGCCGAGCA 2400
 CTGGCAGGAT GAGGAACGCG AGAAGGGCGA GAAAGACAAA GGCAGAGGCG AGATGGAAAG 2460
 ACCCGGCGAG GGCGAAAAG AGCTGGCCGA GAAAGAGGAA TGGAAAGAAC GCGACGGCGA 2520
 AGAACAAAGAG CAGAAAGAAA GAGAGCAGGG CCACCAGAAA GAACGGAATC AAGAGATGGA 2580
 AGAAGGCGGC GAGGAAGAAC ACGGCGAAGG GGAAGAAGAG GAAGGCGACC GAGAGGAAGA 2640
 AGAAGAGAAA GAAGGCGAAG GCAAAGAAGA AGGCGAGGGC GAAGAGGTGG AAGGCGAGCG 2700
 TGAAAAGAA GAGGGCGAAC GCAAGAAAAG AGAACGCGCC GGAAAAGAGG AAAAAGGCCA 2760
 GGAAGAGGGC GACCAAGGCG AAGGCGAGGA AGAAGAACT GAAGGCAGAG GGAAGAGAA 2820
 AGAGGAAGGC GGCGAAGTCC AAGGCGGAGA GGTGGAAGAA GGCAAAGGCG AGCGAGAAGA 2880
 GGAAGAAGAA GAAGGCCAAG GCGAGGAAGA GGAAGGCCGA GCGGAAGAGG AAGAAGCCGA 2940
 AGGGGAAGAA GAAGAAGGCG AAGGCAAGGG CGAAGAGGAG GCGGAAGAAG GCGAGGGCGA 3000
 AGAGGAGGGC GAAGAAGGCG AAGGCGAGGG CGAAGAAGAA GAAGGCGAAG GCGAAGGCCA 3060
 GGAAGAAGGC GAAGGCCAAG GGAAGAAGA GGAAGGCCAA GCGGAAGGCG AAGAAGAAGG 3120
 CGAAGGCGAG GCGGAAGAGG AAGAAGGCCA AGGCAAAGGG GAAGAAGAAG GCGAGGAAGG 3180
 CGAAGGCCAA GGCAGAGGAG AAGAAGGCCA AGGCGAGGGC GAAGATGGCG AAGGCGAAGG 3240
 CGAAGAGGAA GAGGGCGAGT GGGAGGGCGA AGAAGAGGAA GCGGAAGGCG AGGGCGAAGA 3300
 GGAAGGCCAA GGCAGGGGCG AAGAAGGCCA AGGCGAAGGC GAGGAAGAGG AAGGCGAAGG 3360
 CGAAGGGGAA GAAGAAGAGG GCGAAGAAGA AGGCGAAGAG GAAGGCGAAG GGAAGAAGA 3420
 AAGCGAAGGC GAAGGCCAAG AAGAGGAAGA GGGCGAAGTT GAAGGCCAGG TTGAGGGCGA 3480
 AGAAGGCCAA GGCGAAGGGG AAGAAGAAGA AGGCGAGGAA GAAGGGGAAG AGAGAGAAAA 3540
 AGAAGGCCAG GGCGAAGAAA ACCGCCGGA CCGCGAAGAG GAAGAGGAAG AAGAGGGCAA 3600
 GTACCAAGAG ACTGGCGAGG AAGAGAACGA GCGGCAGGAT GCGGAAGAGT ACAAGAAGGT 3660
 GTCCAAGATC AAGGGCAGCG TGAAGTACGG CAAGCACAAG ACCTACCAGA AGAAGTCCGT 3720
 CACCAACACG CAAGGCAATG GAAAAGAACA GCGGAGCAAG ATGCCCGTGC AGTCCAAGAG 3780
 GCTGTGAAG AATGGCCCTA GCGGCAGCAA GAAATTCTGG AACAAATGTG TGCCCCACTA 3840
 CCTCGAGCTG AAGTGAGCCT CGAGCAGCGC TGCTCGAGAG ATCTGCGGCC GCGAGCTCGG 3900
 GGATCCAGAG ATGATAAGAT ACATTGATGA GTTTGGACAA ACCACAATA GAATGCAGTG 3960
 AAAAAAATGC TTTATTTGTG AAATTTGTGA TGCTATTGCT TTATTTGTAA CCATTATAAG 4020
 CTGCAATAAA CAAGTTAACA ACAACAATTG CATTCAATTT ATGTTTCAGG TTCAGGGGGA 4080
 GGTGTGGGAG GTTTTTTAAA GCAAGTAAAA CCTCTACAAA TGTGGTATGG CTGATTATGA 4140
 TCAATGCATC CTAGCCGGAG GAACCCCTAG TGATGGAGTT GGCCACTCCC TCTCTGCGCG 4200
 CTCGCTCGCT CACTGAGGCC GCCCGGGCAA AGCCCGGGCG TCGGGCGACC TTTGGTCTGCC 4260
 CGGCTCAGT GAGCGAGCGA GCGCGCAGAG AGGGAGTGGC CAA 4303 (SEQ ID NO:5)

[0059] The components of the transgene cassette of SEQ ID NO:5 and their respective locations are identified in Table 2 below:

Table 2

Location (bp)	Component	Length (bp)
1-145	5' ITR	145
170-368	GRK promoter	199
398-3856	RPGRorf15 cDNA	3459
3899-4143	SV40 Poly A	245

Location (bp)	Component	Length (bp)
4159-4304	3' ITR	145

[0060] The 5' ITR has the following sequence:

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGG
 TCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA
 GAGAGGGAGTGGCCAACTCCATCACTAGGGGTTTCCT (SEQ ID NO:6)

[0061] The 3' ITR has the following sequence:

AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAC
 TGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTC
 AGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA (SEQ ID NO:7)

[0062] The SV40 polyadenylation sequence has the following sequence:

GGGGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACA ACTAGA
 ATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGT
 AACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATG
 TTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACA
 AATGTGGTATGGCTGATTATGATCA (SEQ ID NO:8)

[0063] Those skilled in the art will appreciate that an AAV vector comprising a transgene and lacking virus proteins needed for viral replication (e.g., cap and rep), cannot replicate since such proteins are necessary for virus replication and packaging. Helper viruses include, typically, adenovirus or herpes simplex virus. Alternatively, as discussed below, the helper functions (E1a, E1b, E2a, E4, and VA RNA) can be provided to a packaging cell including by transfecting the cell with one or more nucleic acids encoding the various helper elements and/or the cell can comprise the nucleic acid encoding the helper protein. For instance, HEK 293 were generated by transforming human cells with adenovirus 5 DNA and now express a number of adenoviral genes, including, but not limited to E1 and E3 (see, e.g., Graham et al., 1977, J. Gen. Virol. 36:59-72). Thus, those helper functions can be provided by the HEK 293

packaging cell without the need of supplying them to the cell by, e.g., a plasmid encoding them.

[0064] The viral vector may be any suitable nucleic acid construct, such as a DNA or RNA construct and may be single stranded, double stranded, or duplexed (i.e., self complementary as described in WO 2001/92551).

[0065] The viral capsid component of the packaged viral vectors may be a parvovirus capsid. AAV Cap and chimeric capsids are preferred. For example, the viral capsid may be an AAV capsid (e.g., AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7 AAV8, AAV9, AAV10, AAV11, AAV12, AAV1.1, AAV2.5, AAV6.1, AAV6.3.1, AAV9.45, AAVrh10, AAVrh74, RHM4-1, AAV2-TT, AAV2-TT-S312N, AAV3B-S312N, AAV-LK03, snake AAV, avian AAV, bovine AAV, canine AAV, equine AAV, ovine AAV, goat AAV, shrimp AAV, and any other AAV now known or later discovered. see, e.g., Fields et al., VIROLOGY, volume 2, chapter 69 (4.sup.th ed., Lippincott-Raven Publishers).

[0066] In some embodiments, the viral capsid component of the packaged viral vector is a variant of a native AAV capsid (i.e. comprises one or more modifications relative to a native AAV capsid). In some embodiments, the capsid is a variant of an AAV2, AAV5 or AAV8 capsid. In preferred embodiments, the capsid is a variant of an AAV2 capsid, such as those described in U.S. Patent Application Publication Number 2019/0255192A1 (e.g. comprising the amino acid sequence of any of SEQ ID NOs: 42-59). In a particularly preferred embodiment, the capsid comprises a VP1 capsid protein having the following amino acid sequence:

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKAAERHKDDSRGLVLPGYKYLGP
 FNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSF
 GGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQP
 ARKRLNFGQTGDADSVDPQPLGQPPAAPSGLGTNTMATGSGAPMADNNEGADGV
 GNSSGNWHCDSTWMGDRVITSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYST
 PWGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTQNDGTTTIA
 NNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNNGSQAVGR
 SSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQLDRLMNPLIDQYLYLSR
 TNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTADNNNSEYSWT

GATKYHLNGRDSL VNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMIT
DEEEIRTTNPVATEQYGSVSTNLQRGNLAISDQTKHARQAATADVNTQGVLPGMVW
QDRDVYLQGPWIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSA
KFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSNKNKSVNVDFTVDTNGVYS
EPRPIGTRYLTRNL (SEQ ID NO:9)

[0067] The variant AAV capsid protein of SEQ ID NO:9 contains the following modifications relative to native AAV2 capsid: (i) a proline (P) to alanine (A) mutation at amino acid position 34, which is located inside the assembled capsid (VP1 protein only), and (ii) an insertion of 10 amino acids (leucine-alanine-isoleucine-serine-aspartic acid-glutamine-threonine-lysine-histidine-alanine/LAISDQTKHA) at amino acid position 588, which is present in VP1, VP2, and VP3.

[0068] A full complement of AAV Cap proteins includes VP1, VP2, and VP3. The ORF comprising nucleotide sequences encoding AAV VP capsid proteins may comprise less than a full complement AAV Cap proteins or the full complement of AAV Cap proteins may be provided.

[0069] In yet another embodiment the present invention provides for the use of ancestral AAV vectors for use in therapeutic in vivo gene therapy. Specifically, in silico-derived sequences were synthesized de novo and characterized for biological activities. This effort led to the generation of nine functional putative ancestral AAVs and the identification of Anc80, the predicted ancestor of AAV serotypes 1, 2, 8 and 9 (Zinn et al., 2015, Cell Reports 12:1056-1068). Predicting and synthesis of such ancestral sequences in addition to assembling into a virus particle may be accomplished by using the methods described in WO 2015/054653, the contents of which are incorporated by reference herein. Notably, the use of the virus particles assembled from ancestral viral sequences may exhibit reduced susceptibility to pre-existing immunity in current day human population than do contemporary viruses or portions thereof.

[0070] The invention includes packaging cells, which are encompassed by "host cells," which may be cultured to produce packaged viral vectors of the invention. The packaging cells of the invention generally include cells with heterologous (1) viral vector function(s),

(2) packaging function(s), and (3) helper function(s). Each of these component functions is discussed in the ensuing sections.

[0071] Initially, the vectors can be made by several methods known to skilled artisans (see, e.g., WO 2013/063379). A preferred method is described in Grieger, et al. 2015, *Molecular Therapy* 24(2):287-297, the contents of which are incorporated by reference herein for all purposes. Briefly, efficient transfection of HEK293 cells is used as a starting point, wherein an adherent HEK293 cell line from a qualified clinical master cell bank is used to grow in animal component-free suspension conditions in shaker flasks and WAVE bioreactors that allow for rapid and scalable rAAV production. Using the triple transfection method (e.g., WO 96/40240), the suspension HEK293 cell line generates greater than 10^5 vector genome containing particles (vg)/cell or greater than 10^{14} vg/L of cell culture when harvested 48 hours post-transfection. More specifically, triple transfection refers to the fact that the packaging cell is transfected with three plasmids: one plasmid encodes the AAV rep and cap genes, another plasmid encodes various helper functions (e.g., adenovirus or HSV proteins such as E1a, E1b, E2a, E4, and VA RNA, and another plasmid encodes the transgene and its various control elements (e.g., modified RPGRorf15 gene and hGRK promoter).

[0072] To achieve the desired yields, a number of variables are optimized such as selection of a compatible serum-free suspension media that supports both growth and transfection, selection of a transfection reagent, transfection conditions and cell density. A universal purification strategy, based on ion exchange chromatography methods, was also developed that resulted in high purity vector preps of AAV serotypes 1-6, 8, 9 and various chimeric capsids. This user-friendly process can be completed within one week, results in high full to empty particle ratios (>90% full particles), provides post-purification yields (>1.times.10.sup.13 vg/L) and purity suitable for clinical applications and is universal with respect to all serotypes and chimeric particles. This scalable manufacturing technology has been utilized to manufacture GMP Phase I clinical AAV vectors for retinal neovascularization (AAV2), Hemophilia B (scAAV8), Giant Axonal Neuropathy (scAAV9) and Retinitis Pigmentosa (AAV2), which have been administered into patients. In addition, a minimum of a 5-fold increase in overall vector production by implementing a perfusion method that entails harvesting rAAV from the culture media at numerous time-points post-transfection.

[0073] The packaging cells include viral vector functions, along with packaging and vector functions. The viral vector functions typically include a portion of a parvovirus genome, such as an AAV genome, with rep and cap deleted and replaced by the modified RPGRorf15 sequence and its associated expression control sequences. The viral vector functions include sufficient expression control sequences to result in replication of the viral vector for packaging. Typically, the viral vector includes a portion of a parvovirus genome, such as an AAV genome with rep and cap deleted and replaced by the transgene and its associated expression control sequences. The transgene is typically flanked by two AAV TRs, in place of the deleted viral rep and cap ORFs. Appropriate expression control sequences are included, such as a tissue-specific promoter and other regulatory sequences suitable for use in facilitating tissue-specific expression of the transgene in the target cell. The transgene is typically a nucleic acid sequence that can be expressed to produce a therapeutic polypeptide or a marker polypeptide.

[0074] The terminal repeats (TR(s)) (resolvable and non-resolvable) selected for use in the viral vectors are preferably AAV sequences, with serotypes 1, 2, 3, 4, 5 and 6 being preferred. Resolvable AAV TRs need not have a wild-type TR sequence (e.g., a wild-type sequence may be altered by insertion, deletion, truncation or missense mutations), as long as the TR mediates the desired functions, e.g., virus packaging, integration, and/or provirus rescue, and the like. The TRs may be synthetic sequences that function as AAV inverted terminal repeats, such as the "double-D sequence" as described in U.S. Pat. No. 5,478,745 to Samulski et al., the entire disclosure of which is incorporated in its entirety herein by reference. Typically, but not necessarily, the TRs are from the same parvovirus, e.g., both TR sequences are from AAV2.

[0075] The packaging functions include capsid components. The capsid components are preferably from a parvoviral capsid, such as an AAV capsid or a chimeric AAV capsid function. Examples of suitable parvovirus viral capsid components are capsid components from the family Parvoviridae, such as an autonomous parvovirus or a Dependovirus. For example, the capsid components may be selected from AAV capsids, e.g., AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAVrh10, AAVrh74, RHM4-1, RHM15-1, RHM15-2, RHM15-3/RHM15-5, RHM15-4, RHM15-6, AAV Hu.26, AAV1.1, AAV2.5, AAV6.1, AAV6.3.1, AAV9.45, AAV2i8, AAV2G9, AAV2i8G9, AAV2-TT, AAV2-TT-S312N, AAV3B-S312N, and AAV-LK03, and other

novel capsids as yet unidentified or from non-human primate sources. Capsid components may include components from two or more AAV capsids.

[0076] The packaged viral vector generally includes the modified RPGRorf15 gene sequence and expression control sequences flanked by TR elements, referred to herein as the "transgene" or "transgene expression cassette," sufficient to result in packaging of the vector DNA and subsequent expression of the modified RPGRorf15 gene sequence in the transduced cell. The viral vector functions may, for example, be supplied to the cell as a component of a plasmid or an amplicon. The viral vector functions may exist extrachromosomally within the cell line and/or may be integrated into the cell's chromosomal DNA.

[0077] Any method of introducing the nucleotide sequence carrying the viral vector functions into a cellular host for replication and packaging may be employed, including but not limited to, electroporation, calcium phosphate precipitation, microinjection, cationic or anionic liposomes, and liposomes in combination with a nuclear localization signal. In embodiments wherein the viral vector functions are provided by transfection using a virus vector; standard methods for producing viral infection may be used.

[0078] The packaging functions include genes for viral vector replication and packaging. Thus, for example, the packaging functions may include, as needed, functions necessary for viral gene expression, viral vector replication, rescue of the viral vector from the integrated state, viral gene expression, and packaging of the viral vector into a viral particle. The packaging functions may be supplied together or separately to the packaging cell using a genetic construct such as a plasmid or an amplicon, a Baculovirus, or HSV helper construct. The packaging functions may exist extrachromosomally within the packaging cell, but are preferably integrated into the cell's chromosomal DNA. Examples include genes encoding AAV Rep and Cap proteins.

[0079] The helper functions include helper virus elements needed for establishing active infection of the packaging cell, which is required to initiate packaging of the viral vector. Examples include functions derived from adenovirus, baculovirus and/or herpes virus sufficient to result in packaging of the viral vector. For example, adenovirus helper functions will typically include adenovirus components E1a, E1b, E2a, E4, and VA RNA. The

packaging functions may be supplied by infection of the packaging cell with the required virus. The packaging functions may be supplied together or separately to the packaging cell using a genetic construct such as a plasmid or an amplicon. See, e.g., pXR helper plasmids as described in Rabinowitz et al., 2002, *J. Virol.* 76:791, and pDG plasmids described in Grimm et al., 1998, *Human Gene Therapy* 9:2745-2760. The packaging functions may exist extrachromosomally within the packaging cell, but are preferably integrated into the cell's chromosomal DNA (e.g., E1 or E3 in HEK 293 cells).

[0080] Any suitable helper virus functions may be employed. For example, where the packaging cells are insect cells, baculovirus may serve as a helper virus. Herpes virus may also be used as a helper virus in AAV packaging methods. Hybrid herpes viruses encoding the AAV Rep protein(s) may advantageously facilitate for more scalable AAV vector production schemes.

[0081] Any method of introducing the nucleotide sequence carrying the helper functions into a cellular host for replication and packaging may be employed, including but not limited to, electroporation, calcium phosphate precipitation, microinjection, cationic or anionic liposomes, and liposomes in combination with a nuclear localization signal. In embodiments wherein the helper functions are provided by transfection using a virus vector or infection using a helper virus; standard methods for producing viral infection may be used.

[0082] Any suitable permissive or packaging cell known in the art may be employed in the production of the packaged viral vector. Mammalian cells or insect cells are preferred. Examples of cells useful for the production of packaging cells in the practice of the invention include, for example, human cell lines, such as VERO, WI38, MRC5, A549, HEK 293 cells (which express functional adenoviral E1 under the control of a constitutive promoter), B-50 or any other HeLa cells, HepG2, Saos-2, HuH7, and HT1080 cell lines. In one aspect, the packaging cell is capable of growing in suspension culture, more preferably, the cell is capable of growing in serum-free culture. In one embodiment, the packaging cell is a HEK293 that grows in suspension in serum free medium. In another embodiment, the packaging cell is the HEK293 cell described in U.S. Pat. No. 9,441,206 and deposited as ATCC No. PTA 13274. Numerous rAAV packaging cell lines are known in the art, including, but not limited to, those disclosed in WO 2002/46359. In another aspect, the

packaging cell is cultured in the form of a cell stack (e.g. 10-layer cell stack seeded with HEK293 cells).

[0083] Cell lines for use as packaging cells include insect cell lines. Any insect cell which allows for replication of AAV and which can be maintained in culture can be used in accordance with the present invention. Examples include *Spodoptera frugiperda*, such as the Sf9 or Sf21 cell lines, *Drosophila* spp. cell lines, or mosquito cell lines, e.g., *Aedes albopictus* derived cell lines. A preferred cell line is the *Spodoptera frugiperda* Sf9 cell line. The following references are incorporated herein for their teachings concerning use of insect cells for expression of heterologous polypeptides, methods of introducing nucleic acids into such cells, and methods of maintaining such cells in culture: *Methods in Molecular Biology*, ed. Richard, Humana Press, N J (1995); O'Reilly et al., *Baculovirus Expression Vectors: A Laboratory Manual*, Oxford Univ. Press (1994); Samulski et al., 1989, *J. Virol.* 63:3822-3828; Kajigaya et al., 1991, *Proc. Nat'l. Acad. Sci. USA* 88: 4646-4650; Ruffing et al., 1992, *J. Virol.* 66:6922-6930; Kimbauer et al., 1996, *Virol.* 219:37-44; Zhao et al., 2000, *Virol.* 272:382-393; and Samulski et al., U.S. Pat. No. 6,204,059.

[0084] Virus capsids according to the invention can be produced using any method known in the art, e.g., by expression from a baculovirus (Brown et al., (1994) *Virology* 198:477-488). As a further alternative, the virus vectors of the invention can be produced in insect cells using baculovirus vectors to deliver the rep/cap genes and rAAV template as described, for example, by Urabe et al., 2002, *Human Gene Therapy* 13:1935-1943.

[0085] In another aspect, the present invention provide for a method of rAAV production in insect cells wherein a baculovirus packaging system or vectors may be constructed to carry the AAV Rep and Cap coding region by engineering these genes into the polyhedrin coding region of a baculovirus vector and producing viral recombinants by transfection into a host cell. Notably when using Baculavirus production for AAV, preferably the AAV DNA vector product is a self-complementary AAV like molecule without using mutation to the AAV ITR. This appears to be a by-product of inefficient AAV rep nicking in insect cells which results in a self-complementary DNA molecule by virtue of lack of functional Rep enzyme activity. The host cell is a baculovirus-infected cell or has introduced therein additional nucleic acid encoding baculovirus helper functions or includes these baculovirus helper functions therein.

These baculovirus viruses can express the AAV components and subsequently facilitate the production of the capsids.

[0086] During production, the packaging cells generally include one or more viral vector functions along with helper functions and packaging functions sufficient to result in replication and packaging of the viral vector. These various functions may be supplied together or separately to the packaging cell using a genetic construct such as a plasmid or an amplicon, and they may exist extrachromosomally within the cell line or integrated into the cell's chromosomes.

[0087] The cells may be supplied with any one or more of the stated functions already incorporated, e.g., a cell line with one or more vector functions incorporated extrachromosomally or integrated into the cell's chromosomal DNA, a cell line with one or more packaging functions incorporated extrachromosomally or integrated into the cell's chromosomal DNA, or a cell line with helper functions incorporated extrachromosomally or integrated into the cell's chromosomal DNA

[0088] The rAAV vector may be purified by methods standard in the art such as by column chromatography or cesium chloride gradients. Methods for purifying rAAV vectors are known in the art and include methods described in Clark et al., 1999, Human Gene Therapy 10(6):1031-1039; Schenpp and Clark, 2002, Methods Mol. Med. 69:427-443; U.S. Pat. No. 6,566,118 and WO 98/09657.

[0089] Treatment methods

[0090] In certain embodiments, a method is provided for the treatment of XLRP in a subject in need of such treatment by administering to the subject a therapeutically effective amount of a nucleic acid having a nucleotide sequence at least 90%, at least 95%, at least 98% identical, or 100% identical to the nucleotide sequence of SEQ ID NO:1 or a pharmaceutical composition comprising such a nucleic acid and at least one pharmaceutically acceptable excipient.

[0091] In related aspects, a nucleic acid comprising a nucleotide sequence at least 90%, at least 95%, at least 98% identical or 100% identical to the nucleotide sequence of SEQ ID NO:1 for use in the treatment of XLRP is provided.

[0092] In other related aspects, the use of a nucleic acid comprising a nucleotide sequence at least 90%, at least 95%, at least 98% identical or 100% identical to the nucleotide sequence of SEQ ID NO:1 for the manufacture of a medicament is provided.

[0093] In other related aspects, the use of a nucleic acid comprising a nucleotide sequence at least 90%, at least 95%, at least 98% identical or 100% identical to the nucleotide sequence of SEQ ID NO:1 for the manufacture of a medicament for the treatment of XLRP is provided.

[0094] In some aspects, the nucleotide sequence at least 90%, at least 95%, at least 98% identical or 100% identical to the nucleotide sequence of SEQ ID NO:1 is operably linked to an expression control sequence. In some embodiments, the nucleotide sequence of SEQ ID NO:1 is operably linked to a human G protein-coupled receptor rhodopsin kinase 1 (hGRK) promoter. In some preferred embodiments, the hGRK promoter has the sequence of SEQ ID NO:4.

[0095] In some embodiments, the nucleotide sequence at least 90%, at least 95%, at least 98% identical or 100% identical to the nucleotide sequence of SEQ ID NO:1 forms part of an expression cassette. In some aspects, the expression cassette comprises from 5' to 3': (a) an AAV2 terminal repeat (b) an hGRK promoter (c) codon optimized RPGRorf15 gene of SEQ ID NO:1 (d) an SV40 polyadenylation sequence and (e) an AAV2 terminal repeat. In preferred embodiments, the 5' AAV2 terminal repeat has the nucleotide sequence set forth as SEQ ID NO:6 and/or the hGRK promoter has the nucleotide sequence set forth as SEQ ID NO:4 and/or the SV40 polyadenylation sequence has the nucleotide sequence set forth as SEQ ID NO:8 and/or the 3' AAV2 terminal repeat has the nucleotide sequence set forth as SEQ ID NO:7. In a particularly preferred embodiment, the expression cassette comprises a nucleic acid comprising the nucleotide sequence of SEQ ID NO:5 or a sequence at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95% , at least 96%, at least 97%, at least 98% or at least 99% identical thereto.

[0096] In further embodiments, a method is provided for the treatment of XLRP in a subject in need of such treatment by administering to the subject a therapeutically effective amount of a recombinant AAV (rAAV) virion, or a pharmaceutical composition comprising same, the rAAV virion comprising (i) a nucleic acid having a nucleotide sequence at least 90%, at least 95%, at least 98% identical or 100% identical to the nucleotide sequence of SEQ ID NO:1 operably linked to an expression control sequence and (ii) an AAV capsid.

[0097] In related embodiments, provided is the use of a recombinant AAV (rAAV) virion comprising (i) a nucleic acid having a nucleotide sequence at least 90%, at least 95%, at least 98% identical or 100% identical to the nucleotide sequence of SEQ ID NO:1 operably linked to an expression control sequence and (ii) an AAV capsid for the treatment of XLRP.

[0098] In other related embodiments, provided is the use of a recombinant AAV (rAAV) virion comprising (i) a nucleic acid having a nucleotide sequence at least 90%, at least 95%, at least 98% identical or 100% identical to the nucleotide sequence of SEQ ID NO:1 operably linked to an expression control sequence and (ii) an AAV capsid for the manufacture of a medicament for the treatment of XLRP.

[0099] In some embodiments, the rAAV virion comprises a native AAV2, AAV4, AAV5 or AAV8 capsid. In other embodiments, the rAAV virion comprises a variant AAV capsid that comprises one or more modifications relative to AAV2, AAV4, AAV5 or AAV8. In a preferred embodiment, the AAV capsid comprises a capsid protein comprising the sequence of SEQ ID NO:9.

[00100] In some embodiments, the rAAV virion comprises (i) a native AAV2 capsid or variant thereof and (ii) an expression cassette comprising from 5' to 3': (a) an AAV2 terminal repeat (b) an hGRK promoter (c) codon optimized RPGRorf15 gene of SEQ ID NO:1 (d) an SV40 polyadenylation sequence and (e) an AAV2 terminal repeat. In preferred embodiments, the rAAV comprises (i) a capsid comprising a capsid protein of SEQ ID NO:9 and (ii) a nucleic acid comprising a 5' AAV2 terminal repeat of SEQ ID NO:6, an hGRK promoter of SEQ ID NO:4, an SV40 polyadenylation sequence of SEQ ID NO:8 and a 3' AAV2 terminal repeat of SEQ ID NO:7. In a particularly preferred embodiment, the rAAV comprises (i) a capsid comprising a capsid protein of SEQ ID NO:9 and (ii) an expression cassette comprising the nucleotide sequence of SEQ ID NO:5.

[00101] In particularly preferred embodiments, the use of an rAAV in the treatment of XLRP or for the manufacture of a medicament for the treatment of XLRP is provided, wherein the rAAV comprises (i) a nucleic acid comprising a nucleotide sequence of SEQ ID NO:5 and (ii) a capsid comprising a capsid protein having the amino acid sequence of SEQ ID NO:9. In some aspects, the rAAV is administered by intravitreal injection.

[00102] In other particularly preferred embodiments, a method for the treatment of XLRP is provided comprising administering to the subject an effective amount of an rAAV comprising (i) a nucleic acid comprising a nucleotide sequence of SEQ ID NO:5 and (ii) a capsid comprising a capsid protein having the amino acid sequence of SEQ ID NO:9. In some aspects, the rAAV is administered to the subject by intravitreal injection.

[00103] In other aspects, a pharmaceutical composition is provided comprising a nucleic acid having a nucleotide sequence at least 90%, at least 95% at least 98% identical or 100% identical to the nucleotide sequence of SEQ ID NO:1, optionally operably linked to an expression control sequence, and at least one pharmaceutically acceptable excipient.

[00104] In some embodiments, the pharmaceutical composition comprises a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 operably linked to a constitutive promoter, preferably an hGRK promoter having a sequence at least 90%, at least 95% at least 98% identical or 100% identical to the nucleotide sequence of SEQ ID NO:4.

[00105] In other aspects, a pharmaceutical composition is provided comprising at least one pharmaceutically acceptable excipient and an infectious rAAV comprising (i) an AAV capsid and (ii) a nucleic acid comprising from 5' to 3': (a) an AAV2 terminal repeat (b) an hGRK promoter (c) codon optimized RPGRorf15 gene of SEQ ID NO:1 (d) an SV40 polyadenylation sequence and (e) an AAV2 terminal repeat. In related embodiments, the pharmaceutical composition comprises between 10^9 and 10^{14} vg, preferably between 10^{10} and 10^{13} vg of the rAAV, more preferably comprises 3×10^{11} vg or 1×10^{12} vg of the rAAV.

[00106] In preferred embodiments, the pharmaceutical composition comprises an rAAV comprising (i) a capsid comprising a capsid protein comprising or consisting of the sequence of SEQ ID NO:9 and (ii) a nucleic acid comprising a 5' AAV2 terminal repeat of SEQ ID NO:6 and/or an hGRK promoter of SEQ ID NO:4 and/or an SV40 polyadenylation sequence

of SEQ ID NO:8 and/or an AAV2 terminal repeat of SEQ ID NO:7. In related embodiments, the pharmaceutical composition comprises between 10^9 vg and 10^{14} vg, preferably between 10^{10} vg and 10^{13} vg of the rAAV, more preferably comprises about 3×10^{11} vg or about 1×10^{12} vg of the rAAV.

[00107] In some embodiments, a method for expressing RPGR in one or more photoreceptor cells of a human subject is provided comprising administering to the human subject an effective amount of an infectious rAAV as herein described, wherein the RPGR is expressed in the one or more photoreceptor cells. In some preferred embodiments, the effective amount of infectious rAAV is 10^9 to 10^{14} vg/eye and/or a single dose of the rAAV is intravitreally administered (bilaterally or unilaterally) to the human subject and/or the rAAV comprises a capsid of SEQ ID NO:9 and/or the rAAV comprises a heterologous nucleic acid comprising the nucleotide sequence of SEQ ID NO:5.

[00108] In a particularly preferred embodiment, a pharmaceutical composition is provided comprising at least one pharmaceutically acceptable excipient and an infectious rAAV comprising (i) a capsid comprising a capsid protein comprising or consisting of the sequence of SEQ ID NO:9 and (ii) a nucleic acid comprising or consisting of the nucleotide sequence of SEQ ID NO:5. In related embodiments, the pharmaceutical composition comprises between 10^{10} and 10^{13} vg of the rAAV, preferably comprises about 3×10^{11} vg or about 1×10^{12} vg of the rAAV.

[00109] In some embodiments, a nucleic acid or infectious rAAV as herein described is administered by periocular or intraocular (intravitreal, suprachoroidal or subretinal) injection to a human with XLRP, whereby the XLRP is treated in the subject. In other embodiments, a nucleic acid or infectious rAAV as herein described is administered subretinally or intravitreally to a human with XLRP, whereby the XLRP is treated in the subject. In preferred embodiments, a human subject with XLRP is administered a single intravitreal injection (bilateral or unilateral) of an rAAV as herein described.

[00110] In related aspects, treatment of XLRP in a treated subject comprises (i) an improvement (i.e. gain) in visual function or functional vision relative to a control (e.g. relative to a baseline measurement in the treated patient prior to treatment, relative to the untreated eye if the nucleic acid or rAAV is administered unilaterally, or relative to an

untreated concurrent or historical control group of XLRP patients) and/or (ii) a decrease in loss of visual function and/or retinal degeneration in a treated eye compared to a control (e.g. untreated eye in same patient or untreated control group) at e.g. 6 months, 12 months or 24 months after treatment. These improvements can be assessed by an appropriate ophthalmological test, including but not limited to visual acuity testing, microperimetry and other visual field testing, anatomical testing, such as optical coherence tomography scans and fundus autofluorescence imaging, retinal electrophysiology, and/or quality of life (QoL) assessments.

[00111] In some aspects, an effective amount of a nucleic acid or rAAV (or pharmaceutical composition comprising same) as herein described is an amount effective to treat XLRP in a human patient. In related aspects, an effective amount of an rAAV as herein described is between 10^9 and 10^{14} rAAV particles (or vector genomes (vg))/eye, preferably between 10^{10} and 10^{13} vg/eye or between 1×10^{11} vg/eye and 5×10^{12} vg/eye, more preferably is about 3×10^{11} vg/eye or about 1×10^{12} vg/eye. In some preferred embodiments, a single dose of about 3×10^{11} vg/eye or about 1×10^{12} vg/eye is intravitreally administered to a human patient with XLRP, whereby the XLRP is treated.

EXAMPLES

[00112] The following examples illustrate preferred embodiments of the present invention and are not intended to limit the scope of the invention in any way. While this invention has been described in relation to its preferred embodiments, various modifications thereof will be apparent to one skilled in the art from reading this application.

Example 1- Codon Optimization of RPGRorf15 cDNA Sequence with Improved Stability

[00113] The human Retinitis Pigmentosa GTPase Regulator open reading frame 15 (hRPGRorf15) sequence contains a highly repetitive, purine-rich region that leads to sequence instability during transgene cassette cloning and plasmid amplification. The hRPGRorf15 cDNA sequence (NCBI Reference Sequence NM_001034853.1) was codon optimized to generate an RPGRorf15 cDNA sequence with increased expression in human cells and improved sequence stability

[00114] The codon optimized nucleotide sequence is set forth below:

ATGAGAGAGCCTGAAGAGCTGATGCCTGATAGCGGAGCAGTGTTTACCTTTGGG
AAGAGCAAGTTCGCAGAGAATAACCCTGGGAAATTCTGGTTTAAGAACGACGTG
CCCGTGCACCTGAGCTGTGGCGATGAGCACTCCGCCGTGGTGACAGGCAACAAT
AAGCTGTACATGTTCCGGCTCTAACAAATTGGGGACAGCTGGGCCTGGGAAGCAAG
TCCGCCATCAGCAAGCCAACCTGCGTGAAGGCCCTGAAGCCCGAGAAGGTGAAG
CTGGCCGCCTGTGGCAGAAACCACACACTGGTGAGCACCGAGGGAGGAAACGTG
TACGCAACAGGAGGCAACAATGAAGGCCAGCTGGGCCTGGGCGACACAGAGGA
GAGGAATACCTTTCACGTGATCAGCTTCTTTACCTCCGAGCACAAGATCAAGCAG
CTGTCCGCCGGCTCTAACACAAGCGCCGCCCTGACCGAGGACGGCCGCCTGTTCA
TGTGGGGCGATAATAGCGAGGGCCAGATCGGCCTGAAGAACGTGTCCAACGTGT
GCGTGCCTCAGCAGGTGACCATCGGCAAGCCAGTGTCTGATCTCTTGTGGCTA
CTATCACAGCGCCTTCGTGACCACAGATGGCGAGCTGTACGTGTTTGGAGAGCCA
GAGAACGGCAAGCTGGGCCTGCCTAACAGCTGCTGGGCAATCACCGGACACCC
CAGCTGGTGTCCGAGATCCCTGAGAAAGTGATCCAGGTGGCATGCGGAGGAGAG
CACACAGTGGTGTGACCGAGAATGCCGTGTATACCTTCGGCCTGGGACAGTTTG
GACAGCTGGGCCTGGGCACATTCCTGTTTGAGACAAGCGAGCCAAAAGTGATCG
AGAACATCCGCGACCAGACAATCAGCTACATCTCCTGCGGGGAGAATCACACAG
CCCTGATCACCGACATCGGCCTGATGTATACCTTTGGCGATGGCCGGCACGGCAA
GCTGGGCCTGGGCCTGGAGAACTTCACAAATCACTTTATCCCCACCCTGTGCTCT
AACTTCCTGCGGTTTCATCGTGAAGCTGGTGGCCTGCGGCGGCTGTCACATGGTGG
TGTTTCGACACCTCACAGGGGAGTGGCCAAGGAGATCGAGTTTGACGAGATCA
ACGATACATGCCTGTCCGTGGCCACCTTCCTGCCATACAGCTCCCTGACATCCGG
CAATGTGCTGCAGCGCACCTGTCTGCCAGGATGCGGAGAAGGGAGAGGGAGCG
GTCCCCTGACTCTTTCAGCATGAGGCGGACACTGCCACCTATCGAGGGCACCTG
GGCCTGTCTGCCTGCTTCTGCCTAACAGCGTGTTCCCAAGATGTAGCGAGAGGA
ATCTGCAGGAGTCTGTGCTGAGCGAGCAGGATCTGATGCAGCCAGAGGAGCCCG
ACTACCTGCTGGATGAGATGACAAAGGAGGCCGAGATCGACAACCTTAGCACCG
TGGAGAGCCTGGGCGAGACAACAGATATCCTGAATATGACACACATCATGTCCC
TGAACCTAATGAGAAGTCTCTGAAGCTGAGCCCAGTGCAGAAGCAGAAGAAGC
AGCAGACCATCGGCGAGCTGACCCAGGACACAGCCCTGACCGAGAACGACGATT
CTGATGAGTATGAGGAGATGAGCGAGATGAAGGAGGGCAAGGCCTGTAAGCAG
CACGTGTCCCAGGGCATCTTCATGACCCAGCCAGCCACCACAATCGAGGCCTTTT

CTGACGAAGAGGTGGAGATCCCCGAGGAGAAGGAGGGCGCCGAGGATAGCAAG
GGCAATGGCATCGAGGAGCAGGAGGTGGAGGCCAACGAGGAGAATGTGAAGGT
GCACGGCGGCAGAAAGGAGAAGACAGAGATCCTGTCCGACGATCTGACCGACA
AGGCCGAGGTGTCCGAGGGCAAGGCCAAGTCTGTGGGAGAGGCAGAGGACGGA
CCAGAGGGACGCGGCGATGGAACCTGCGAGGAGGGATCCTCTGGAGCAGAGCA
CTGGCAGGACGAAGAAAGAGAGAAGGGCGAGAAGGATAAAGGCAGAGGAGAG
ATGGAGAGGCCTGGAGAGGGAGAGAAGGAGCTGGCAGAGAAGGAGGAGTGA
AGAAGAGGGACGGCGAGGAGCAGGAGCAGAAGGAGAGAGAGCAGGGCCACCA
GAAGGAGAGGAACCAGGAGATGGAGGAGGGAGGAGAGGAGGAGCACGGCGAG
GGAGAGGAGGAGGAGGGCGATAGAGAGGAAGAAGAGGAGAAGGAGGGAGAGG
GCAAGGAGGAAGGCAGGGAGAGGAGGTGGAGGGAGAAAGGGAGAAGGAGGA
GGGAGAGCGCAAGAAGGAAGAAAGAGCAGGCAAGGAAGAGAAGGGAGAGGAG
GAGGGCGATCAGGGCGAAGGAGAGGAGGAGGAGACAGAGGGAAGGGGAGAGG
AGAAGGAGGAGGGAGGAGAGGTCTGAAGGAGGAGAAGTGGAGGAGGGCAAGGG
CGAAAGAGAAGAGGAGGAGGAGGAAGGCGAGGGCGAAGAAGAGGAGGGCGAG
GGCGAGGAAGAAGAGGGCGAGGGCGAAGAGGAAGAAGGCAGGGCAAGGGCG
AGGAGGAGGGCGAAGAAGGCGAAGGGGAGGAGGAGGGCGAAGAGGGAGAGGG
CGAGGGCGAGGAGGAAGAAGGCGAAGGCGAAGGCGAAGAAGAAGGAGAAGGA
GAGGGCGAAGAGGAGGAAGGCGAAGGAGAAGGAGAGGAGGAAGGAGAAGGGG
AGGGCGAAGAGGAGGAGGGAGAAGGCAAGGGAGAAGAAGAAGGCGAAGAAGG
CGAGGGAGAAGGCGAGGAAGAAGAAGGCGAGGGAGAGGGAGAGGACGGCGAA
GGCGAGGGCGAGGAAGAGGAAGGAGAGTGGGAGGGCGAGGAAGAGGAGGGAG
AAGGAGAAGGCGAAGAAGAAGGGGAAGGAGAGGGCGAGGAAGGAGAAGGCGA
AGGCGAAGAGGAGGAGGGGGAAGGGGAGGGCGAGGAGGAAGAGGGAGAAGAG
GAAGGCGAAGAAGAGGGAGAAGGCGAAGAGGAAGGAGAAGGCGAGGGAGAAG
AAGAGGAGGAGGGCGAGGTCTGAAGGCGAGGTGGAGGGCGAAGAGGGGGAAGG
CGAAGGCGAGGAGGAGGAAGGGGAAGAAGAAGGCGAGGAGAGAGAGAAAGAA
GGCGAGGGCGAGGAGAACAGAAGGAATCGCGAAGAAGAAGAGGAAGAAGAGG
GCAAGTACCAGGAGACAGGCGAGGAGGAGAACGAGCGGCAGGATGGCGAGGAG
TATAAGAAGGTGTCCAAGATCAAGGGCTCTGTGAAGTACGGCAAGCACAAAGACC
TATCAGAAGAAGAGCGTGACCAACACACAGGGCAATGGCAAGGAGCAGCGCAG
CAAGATGCCTGTGCAGTCCAAGCGGCTGCTGAAGAATGGCCCCTCTGGGAGCAA
GAAGTTTTTGAATAATGTCTGCCACACTACCTGGAGCTGAAATGA (SEQ ID
NO:10)

[00115] AAV plasmids containing the codon optimized hRPGRorf15 gene (SEQ ID NO:10) under the control of either the control of human G protein-coupled receptor kinase 1 promoter, also known as the human rhodopsin kinase promoter (hGRK) or the ubiquitous 3-phosphoglycerate kinase (PGK) promoter were constructed by GenScript.

[00116] 20 ng of AAV plasmid DNA was used to transform competent *E. coli* (Cat. #C3040H, New England BioLabs, Ipswich, MA) and the cells were spread on Kanamycin 50 µg/ml plates (#L1025, Teknova, Hollister, CA). Miniprep cultures were grown from the resulting colonies, DNA was prepared with the GeneJET Plasmid Miniprep kit (Cat. #0503, ThermoFisher, Waltham, MA) and restriction digested to identify positive clones.

[00117] Despite codon optimization, sequence instability of the codon optimized hRPGRorf15 (SEQ ID NO:10) during plasmid production was detected following restriction digestion.

[00118] A second codon optimized hRPGRorf15 sequence was developed using a different optimization algorithm that included parameters including, but not limited to, codon usage bias, GC content, AT-rich or GC-rich regions, mRNA secondary structure, RNA instability motifs, cryptic splicing sites, internal chi sites and ribosomal binding sites, and repeat sequences. The codon usage bias in humans was changed by upgrading the codon adaptation index (CAI) to 0.89. The average GC content was optimized from 59.16 in the native sequence to 57 in the optimized sequence to prolong the half-life of the mRNA. The resulting codon optimized nucleotide sequence, set forth herein as SEQ ID NO:1, contains improved codon usage, altered GC content, better mRNA stability, and modification of negative cis acting elements.

[00119] An AAV plasmid (pAAV-GRK promoter-cohRPGRorf15-SV40) was constructed comprising the nucleotide sequence of SEQ ID NO:5 (SEQ ID NO:5 comprises (i) 5' AAV2 ITR (SEQ ID NO:6); (ii) codon optimized hRPGRorf15 cDNA (SEQ ID NO:1) under the control of hGRK promoter (SEQ ID NO:4); (iii) SV40 late polyA element (SEQ ID NO:8) and (iv) 3' AAV2 ITR (SEQ ID NO:7)).

[00120] pAAV-GRK promoter-cohRPGRorf15-SV40 DNA was prepared as follows. Plasmid DNA from GenScript (20 ng) was used to transform competent *E. coli* (Cat.

#C3040H, New England BioLabs, Ipswich, MA) and the cells were spread on Kanamycin 50 µg/ml plates (#L1025, Teknova, Hollister, CA). Miniprep cultures were grown from the resulting colonies, DNA was prepared with the GeneJET Plasmid Miniprep kit (Cat. #0503, ThermoFisher, Waltham, MA) and restriction digested to identify positive clones. A 50 ml culture in Terrific Broth was grown from one positive clone and DNA was prepared with the Qiagen EndoFree Plasmid Maxi Kit (Cat. #12362, Qiagen, Hilden, Germany). The maxiprep of pAAV-GRK-cohRPGRorf15-SV40 was digested with multiple restriction enzymes to verify the identity of the plasmid. Gel electrophoresis of the restriction digests and the expected fragments are shown in Figure 1. All actual fragments matched the expected fragments. The sequence of the expression cassette was verified by Sanger DNA sequencing.

[00121] Conclusion: The maxiprep of pAAV-GRK-cohRPGRorf15-SV40 mapped correctly by restriction digest and its integrity was verified by Sanger DNA sequencing. Thus, the codon optimized hRPGRorf15 sequence set forth as SEQ ID NO:1 exhibits superior stability relative to both the native sequence of SEQ ID NO:3 and the codon optimized sequence of SEQ ID NO:10.

Example 2 – Expression and Activity of human RPGRorf15 protein expressed from Codon Optimized hRPGRorf15 of SEQ ID NO:1

[00122] Expression and activity of human RPGRorf15 protein expressed from pAAV-GRK-cohRPGRorf15-SV40 was assessed in transfected HEK293T cells.

[00123] Briefly, HEK293T cells were seeded in 12-well plates at 2.0×10^5 cells/well in 1.0 ml DMEM/10% FBS media. HEK293T cells were used due to their high transfectability and protein expression. The next day, 1.0 µg AAV plasmid DNA complexed with 3.0 µl FuGene6 (Cat.# E2691, Promega, Madison, WI) was added to the cells in duplicate wells. Two days after transfection, the cells were washed with PBS and lysed in 0.25 ml 1x Passive Lysis Buffer (Promega) containing 1x Halt Protease Inhibitor (ThermoFisher), rocking for 15 minutes at room temperature. Cell debris was pelleted by centrifugation in a microcentrifuge at 12,000 g for 10 minutes at 4°C. The supernatant was collected and stored at -20°C. No-plasmid and pAAV-PGK promoter-cohRPGRorf15-SV40 samples were included in the transfection as negative and positive controls, respectively. pAAV-PGK promoter-cohRPGRorf15-SV40 is identical to the aforementioned AAV vector except that codon

optimized hRPGRorf15 is operably linked to a ubiquitous promoter 3-phosphoglycerate kinase (PGK) promoter rather than an hGRK promoter.

[00124] Cell lysate (20 μ l) was mixed with 10 μ l 4x LDS, 4 μ l 10x Reducing Agent, and 6 μ l water (final volume = 40 μ l) and denatured at 70°C for 10 minutes. Samples were loaded on a 12-well Bolt 4-12% Bis-Tris Plus polyacrylamide gel (Invitrogen, NW04122BOX) and ran in 1x MOPS buffer at 200 V for 32 minutes. Separated proteins were transferred to a nitrocellulose filter with the iBlot 2 device (ThermoFisher) for 10 minutes and probed with primary anti-RPGR (Sigma HPA001593 1:2000 and GenScript CT-15 U1729DC260_16 1:500), and anti-polyglutamylation GT335 (AG-20B-0020 1:500, Adipogen, San Diego, CA) antibodies using the iBind Flex device (ThermoFisher). Secondary antibodies were HRP-conjugated goat anti-rabbit (ThermoFisher 31460) for the anti-RPGR primary antibodies and HRP-conjugated goat anti-mouse (ThermoFisher 31430) for the anti-polyglutamylation primary antibody. Proteins were visualized with SuperSignal West Dura Chemiluminescent Substrate (ThermoFisher 34076) and imaged on a ChemiDoc MP (BioRad, Hercules, CA). All antibodies used are listed below in Table 3.

[00125] **Table 3:** Western Blot Antibodies

Antibody	Host species	Vendor	Catalog #	Dilution
Anti-RPGR polyclonal	Rabbit	Sigma	HPA001593	1:2,000
Anti-CT-15	Rabbit	GenScript	U1729DC260_16	1:500
Anti-Polyglutamylation GT335	Mouse	Adipogen	AG-20B-0020	1:500
HRP anti-Rabbit IgG (H+L)	Goat	Thermo	31460	1:5,000
HRP anti-Mouse IgG (H+L)	Goat	Thermo	31430	1:5,000

[00126] Figure 2 shows an image of a representative Western blot of lysates from transfected HEK293T cells. The CT-15 and Sigma antibodies detect the same 135-140 kD species that appears to be RPGRorf15, as it is present in RPGR-transfected but not untransfected lysates, is the correct size and is recognized by the polyglutamylation-detecting antibody GT335. Expression is higher when driven by the ubiquitous PGK promoter, which is not preferentially active in photoreceptor cells.

[00127] Conclusion - Western blot analysis of lysates from transfected HEK293T cells demonstrates expression and poly-glutamylation of the correct size hRPGRorf15 protein expressed from the codon optimized hRPGRorf15 of SEQ ID NO:1.

Example 3 – Functional Expression of hRPGRorf15 in an *in vitro* model of human XLRP

[00128] A human *in vitro* model system was generated to evaluate correction of the X-linked Retinitis Pigmentosa (XLRP) disease phenotype with the codon optimized human RPGRorf15 nucleic acid having the nucleotide sequence of SEQ ID NO:1. To that end, an AAV vector was constructed comprising the nucleotide sequence of SEQ ID NO:1 driven by the human G-protein coupled receptor rhodopsin kinase 1 (hGRK) promoter (i.e. the AAV vector backbone described in Examples 1 and 2, having the sequence of SEQ ID NO:5) and a variant capsid protein having the amino acid sequence of SEQ ID NO:9. The hGRK promoter was chosen to limit expression of RPGRorf15 to photoreceptors.

[00129] Peripheral blood mononuclear cells (PBMCs) were isolated from whole blood drawn from individuals with XLRP and reprogrammed into induced pluripotent stem cells (iPSCs) using the CytoTune iPS 2.0 Sendai Reprogramming Kit (Thermo Fisher Scientific, Waltham, MA). Pluripotency of the pluripotent stem cells was confirmed by immunocytochemistry examining iPSC markers including Sox2, Oct4 and Nanog. The induced pluripotent stem cells were then differentiated into photoreceptors by the methods described in Gonzalez-Cordero *et al.*, Stem Cell Report, 9, 820:837 (2017); Gonzalez Cordero *et al.*, Human Gene Therapy, 29(1) (2018); and Meyer *et al.*, Stem Cells, 29(8):1206-1218 (2011). Photoreceptor differentiation was confirmed by immunocytochemistry examining specific markers, Recoverin and Rhodopsin. The photoreceptors were confirmed to lack hRPGRorf15 protein expression and glutamylation of the hRPGorf15 protein, which is known to confer functionality.

[00130] Immunocytochemistry was as follows: Cells were fixed with 4% paraformaldehyde (PFA) (Santa Cruz Biotechnologies, Dallas, TX) for 15 minutes at 4°C. All antibody staining was done in a blocking solution of PBS with 0.2% Triton-X100 (Sigma-Aldrich), 2% bovine serum albumin (Millipore Sigma, Burlington, MA), and 5% goat serum (Thermo Fisher Scientific). Primary antibody incubations were done overnight at 4°C. Cells were then incubated with secondary antibodies for one hour at room temperature and then

counterstained with DAPI (Sigma Aldrich) in PBS for five minutes at room temperature. Cells were imaged using a Zeiss Axio Observer.D1 Fluorescent Microscope. Image processing was performed using Zeiss Zen 2 software (Carl Zeiss Microscopy LLC, White Plains, NY). A list of primary and secondary antibodies is provided at Table 4:

Table 4

Antibody	Host	Company-Catalog No.	Dilution
<i>Primary Antibodies</i>			
OCT4	Mouse	Millipore- MAB4401	1:50
Nanog	Rabbit	Abcam- ab21624	1:50
SOX2	Rabbit	Abcam- ab92494	1:50
Beta-Tubulin III	Mouse	Sigma- T8578	1:200
HNF4- α	Rabbit	Santa Cruz-SC-8987	1:100
A-SMA	Mouse	Sigma Aldrich- A2547	1:500
Recoverin	Rabbit	EMD Millipore- AB5585	1:100
Rhodopsin	Mouse	Abcam- AB98887	1:100
RPGR	Rabbit	Sigma- HPA001593	1:2000
GT335	Mouse	Fisher Adipogen- 50-463-394 \ddagger	1:4000
Alpha Tubulin	Rabbit	Abcam- ab52866	1:4000
<i>Secondary Antibodies</i>			
Alexa Fluor488 anti-rabbit	Goat	Invitrogen-A11078	1: 500
Alexa Fluor555 anti-rabbit	Goat	Invitrogen-A21428	1:500
Alexa Fluor680 anti-rabbit	Goat	Invitrogen-A21109	1:500
Alexa Fluor488 anti-mouse	Goat	Invitrogen-A11029	1:500
Alexa Fluor555 anti-mouse	Goat	Invitrogen-A21422	1:500
Alexa Fluor680 anti-mouse	Goat	Invitrogen-35518	1:500
Horseradish Peroxidase anti-Rabbit IgG (H+L)	Goat	Thermo- 31460	1:5000
Horseradish Peroxidase anti-Mouse IgG (H+L)	Goat	Thermo- 31430	1:5000

[00131] To assess transcript levels of codon optimized RPGRorf15 transgene following transduction into the XLRP-iPSC derived diseased photoreceptors, XLRP photoreceptors (PR) were transduced with the above-described AAV vectors at a multiplicity of infection (MOI, viral genomes per cell) of 50,000 to ensure levels above the limit of detection of the

assays. RNA was isolated 30 days post transduction and cDNA was synthesized. Digital droplet PCR was run on the prepared samples and transcript levels per droplet were analyzed as a copies/mL value. Quantification of the number of droplets, above the set threshold, containing the transcript of the primer/probe set was examined. Two primer/probe sets were created to specifically differentiate the codon optimized human RPGRorf15 transgene from the endogenous human RPGR1-19 constitutive isoform (hRPGR1-19).

[00132] Non-transduced XLRP diseased cells expressed low, background levels of cohRPGRorf15 transcript, as expected. Following transduction with AAV vector, cells showed over a 400-fold increase of cohRPGRorf15 transcript levels compared to hRPGR1-19. Transduced cells displayed over a 1000-fold increase in cohRPGRorf15 transcript compared to non-transduced cell cohRPGRorf15 levels. Non-transduced cells had a higher level of hRPGR1-19 than cohRPGRorf15. See Figure 3. Analysis was done in triplicate and levels were averaged. Transduction with AAV vector comprising codon optimized hRPGRorf15 of SEQ ID NO:1 significantly increased transcript levels of cohRPGRorf15 in photoreceptor cultures.

[00133] To assess protein levels of codon optimized human RPGRorf15 transgene produced by transduction of XLRP-iPSC derived photoreceptor cells with the AAV vectors, XLRP-iPSC derived diseased photoreceptors were transduced at a MOI of 50,000 vg/cell. Cell lysates were collected 30 days post transduction and SDS-PAGE and Western blot analysis were carried out to evaluate hRPGRorf15 protein levels. Band intensity was quantified and is depicted as a histogram in Figure 4. Transduction with AAV vector elicited a significant increase in expression of human RPGRorf15 protein, compared to non-transduced cells.

[00134] In order to determine whether the cohRPGRorf15 protein exogenously introduced into photoreceptors was functional, glutamylation, a surrogate of function, was examined. Glutamylation of hRPGRorf15 and protein function are strongly correlated according to published work. (Fischer et al., 2017; Rao et al., 2016; Sun et al., 2016). XLRP-iPSC-derived diseased PR were transduced at a MOI of 50,000 vg/cell. Cell lysates were collected 30 days post transduction and SDS-PAGE and Western blot analysis was carried out to evaluate glutamylation of the expressed hRPGRorf15 protein. Glutamylation was determined by probing the membrane with a glutamylation specific antibody, GT335, and examining

positive banding patterns at the hRPGRorf15 size, 127kDa. Band intensity was quantified and depicted as a histogram at Figure 5. Transduction of PR cells with AAV vector comprising codon optimized hRPGRorf15 nucleotide sequence led to a significant increase in glutamylation of human RPGRorf15 protein, compared to non-transduced cells in both XLRP patient-derived diseased photoreceptors.

[00135] Due to the low hRPGRorf15 protein levels detected in the Western blot with use of a high MOI, a dose response of the hRPGRorf15 codon optimized transgene (cohRPGRorf15) was verified. To this end, an AAV vector was constructed comprising the codon optimized RPGRorf15 sequence of SEQ ID NO:1 operably linked to a ubiquitous promoter 3-phosphoglycerate kinase (PGK) and a capsid of SEQ ID NO:9 (this AAV vector was identical to the AAV vector described above aside from the promoter). Diseased photoreceptors were transduced at three MOIs, 5,000, 10,000 and 20,000. Cell lysates were collected 30 days post transduction and SDS-PAGE and Western blot analyses were carried out to evaluate hRPGRorf15 protein levels and glutamylation (GT335 = anti-glutamylation antibody). Band intensity was quantified and depicted as a histogram (Figure 6). Although there was high variability, due to the heterogeneity of the cultures, hRPGRorf15 protein and glutamylation of hRPGRorf15 were observed at lower MOIs using a constitutive promoter to drive cohRPGRorf15 expression.

[00136] Conclusion – the *in vitro* studies with iPSC-derived photoreceptors have demonstrated that AAV-mediated delivery of codon optimized hRPGRorf15 of SEQ ID NO:1 restores human RPGRorf15 transcript and transgene expression in human XLRP diseased photoreceptors. Furthermore, the RPGRorf15 protein, expressed following transduction of 4D-125, was post-translationally glutamylated. Based on published literature, glutamylation confers functionality of RPGRorf15.

Example 4 – Assessment of Safety and Biodistribution of Codon Optimized R100 via Intravitreal Administration in Non-Human Primates

[00137] Materials and Methods

[00138] *GLP Toxicology and Biodistribution Studies*

[00139] Male cynomolgus macaques (*macaca fascicularis*) aged 2-14 years were dosed via two 50 μ L intravitreal injections into each eye through the sclera for a total dose volume of 100 μ L/eye. Doses of 1×10^{11} vg/eye and 1×10^{12} vg/eye were evaluated. The animals were anesthetized with Ketamine IM and given topical ophthalmic solutions to eliminate pain. 20-80 mg of methylprednisolone was administered by IM injection weekly post-injection. Euthanasia was performed by trained veterinary staff at Week 3, Week 13, and Week 26 post-administration.

[00140] 4D-125 (rAAV comprising a capsid protein of SEQ ID NO:9 and a heterologous nucleic acid comprising the nucleotide sequence of SEQ ID NO:5) genome biodistribution was assessed in all major ocular compartments (retina, optic nerve, ciliary body, iris, trabecular meshwork), and major systemic organs (including the testes) using validated, GLP-compliant qPCR assay. In tissues where genomes were detected, transgene expression was assessed by a qualified, GLP-compliant RT-qPCR assay.

[00141] Serial toxicology assessments performed in the study were: clinical ocular evaluations (complete ophthalmic examinations, including SD-OCT imaging and ERG), systemic evaluations, clinical pathology, gross pathology and microscopic pathology. Assays were validated to determine the anti-capsid and anti-transgene antibody responses. ELISpot assays were validated to detect cellular responses to the R100 (comprising a variant capsid protein of SEQ ID NO:9) capsid and expressed proteins.

[00142] *Neutralizing Antibody Assay*

[00143] 2v6.11 cells were plated at a density of 3×10^4 cells/well 24 hours prior to infection. rAAV vectors encoding firefly luciferase driven by the CAG promoter were incubated at

37°C for 1 hour with individual serum samples prior to infection, and cells were then infected at a genomic MOI of 1,000. Luciferase activity was assessed 48 hours post infection using the Luc-Screen Extended-Glow Luciferase Reporter Gene Assay System (Invitrogen) or the ONE-Glo Luciferase Assay System (Promega) and quantified using the BioTek Cytation 3 Cell Imaging Multi-Mode Reader and Gen5 software.

[00144] Prior to enrollment in studies, non-human primates (NHP) serum was screened for the presence of neutralizing antibodies against R100. NHPs were enrolled in studies when samples resulted in less than 50% neutralization of AAV transduction at a 1:10 serum dilution.

[00145] *AAV Manufacturing*

[00146] Recombinant R100 viral vectors were produced by transient transfection in HEK293 cells. Cells were cultured in DMEM supplemented with FBS and were maintained at 37°C in a 5% CO₂ environment. Cells were triply transfected (payload, capsid, and helper plasmids) using polyethylenimine (PEI). 48-96 hours post-transfection, viral particles were harvested from cells and/or supernatant and cells lysed via microfluidization. Cell lysate and/or supernatant was enzymatically treated to degrade plasmid and host-cell DNA, then clarified and concentrated by tangential flow filtration (TFF). The TFF retentate was then loaded onto an affinity resin column for purification. Following pH-gradient elution, post-affinity material was buffer exchanged, then further purified (if needed) by anion-exchange chromatography. Purified rAAV was then formulated into DPBS with 0.001% polysorbate-20, sterile filtered, and filled to yield rAAV Drug Product.

[00147] **Results**

[00148] *4D-125 delivery is safe and results in expression of therapeutic transgene in NHP*

[00149] 4D-125 (R100.GRK-cohRPGRorf15) has been advanced into a Phase 1-2 clinical trial. Investigational New Drug (IND)-enabling data for this product includes evaluation in a 6-month Good Laboratory Practices (GLP) toxicology and biodistribution study (Table 5). A total of 30 eyes of 30 NHPs were injected by intravitreal injection with a single eye administration.

Table 5: Good Laboratory Practices (GLP) Toxicology and Biodistribution Studies

4DMT Study Number	Lot Number	Number	Gender	Eye(s)	Dose	In-Life
4D18-08	N/A	1	Male	OD	vehicle	3 weeks
	4DEP000008.01	5	Male	OD	1E+11 vg/eye	
		5	Male	OD	1E+12 vg/eye	
	N/A	1	Male	OD	vehicle	13 weeks
	4DEP000008.01	5	Male	OD	1E+11 vg/eye	
		5	Male	OD	1E+12 vg/eye	
N/A	1	Male	OD	vehicle	26 weeks	
4DEP000008.01	5	Male	OD	1E+11 vg/eye		
	5	Male	OD	1E+12 vg/eye		

[00150] No significant toxicities were observed with 4D-125 at either dose level, as determined by clinical observations, histopathology, OCT, or ERG. Administration of 4D-125 into a single eye resulted in only minimal to mild anterior uveitis that was restricted to the immediate post-administration period and resolved by Week 3 (**Figure 9**); in some cases systemic steroid doses were transiently increased.

[00151] Very high levels of vector genomes were present in the retina of the treated eye at all timepoints (week 3, left panel; week 13, middle panel; week 26, right panel), indicating persistence of the vector in ocular tissue (**Figure 10**). In addition to the retina, vector genomes were detected in the treated eye within samples from the aqueous humor, vitreous humor, iris/ciliary body, and the optic nerve at all timepoints. Non-ocular tissues generally had no detectable vector genomes with the exception of low levels in liver, spleen, and the lymph nodes (**Figure 10**). R100 vector-derived transgene expression was detected in the treated retina and iris/ciliary body from both low and high dose groups (**Figure 11**). Gene expression was dose-dependent and increased from Week 3 to Week 13 and remained stable at Week 26 (**Figure 11, left, middle and right panel respectively**). No non-ocular vector expression was detected at Week 26 (**Figure 11**).

[00152] Using an ELISpot assay to evaluate cellular immune responses, no animals developed significant responses to R100 capsid peptides or transgene peptides (data not

shown). A majority of animals dosed with 4D-125 generated an anti-capsid antibody response post-administration (data not shown).

[00153] Summary

[00154] 4D-125 (R100.GRK-cohRPGRorf15) has recently been translated into a clinical trial for the inherited retinal disease x-linked retinitis pigmentosa (NCT04517149). This therapeutic product has been evaluated in a GLP toxicology and biodistribution studies (Table 5). A total of 30 NHPs were injected with a single eye administration; a total of 30 NHP eyes were injected. No significant test-article-related adverse events or T-cell responses were reported. Mild to moderate, transient corticosteroid-responsive anterior uveitis was observed. Transgene expression was localized to the retina, and expression was not detected in any of the systemic organs evaluated. Human clinical trials are underway in order to determine the safety, pharmacodynamics, and efficacy (including through serial visual field testing and optical coherence tomography scans) of this product by intravitreal injection.

Example 5 – Assessment of Safety of Codon Optimized RPGRorf15 cDNA Sequence Delivered by R100 via Intravitreal Administration in Human X-Linked Retinitis Pigmentosa Patients

[00155] Initial Phase 1 Dose Escalation Safety and Tolerability Data Summary

[00156] Clinical trial designs and enrollment

[00157] The clinical trial employed a standard “3+3” dose-escalation designed to assess the safety, tolerability and biologic activity of a single intravitreal injection of 4D-125 at two dose levels (3E11 or 1E12 vg/eye). A total of six patients were enrolled across dose escalation cohorts, with three at each dose level. Patients received a standard immunosuppression regimen with taper; adjustments were determined by investigators. The results described are based on cut-offs between 4-9 months post-administration.

[00158] Initial Tolerability and Adverse Event Profile

[00159] 4D-125 was well-tolerated throughout the assessment period as outlined in the treatment-emergent adverse event (AE) summary table (Table 6):

[00160] Table 6: Adverse Event Summary

Patient # enrolled	6
Doses	3E11 or 1E12 vg/eye
Follow-up at data cut-off (months)	4-9 months
Dose-Limiting Toxicities (DLTs)	0 (0%)
Serious AE	0 (0%)
Any CTCAE Grade \geq 3	0 (0%)
Retinal AE (Any Grade)	0 (0%)
Uveitis CTCAE Grade 2 (moderate)	1/6 (17%)
Uveitis CTCAE Grade 1 (mild)	2/6 (33%)

[00161] Clinical Assessments

[00162] Preliminary biological activity was assessed using microperimetry (MP) to measure retinal sensitivity and SD-OCT to measure ellipsoid zone area (EZA). Seven subjects (median age 42.5 years; range 27-56 years) received 4D-125 (3×10^{11} vg/eye (n=3) and 1×10^{12} vg/eye (n=4)) with follow-up of 4.2-12.5 months. Intraocular inflammation (4/7 subjects) was mild or moderate, transient (duration 0.9-1.6 months) and steroid-responsive. Most of the subjects had advanced disease, with only 2 having both measurable EZA and mean MP retinal sensitivity (mMPRS) at baseline (BL) in both eyes and follow-up of at least 4 months. Both subjects had a greater increase from BL in mMPRS in the treated vs. untreated eye (+1.65 dB vs. +0.25 dB at 9 months and +0.50 dB vs. +0.10 dB at 4 months; BL values 1.5-3.2 dB) and number of loci gaining \geq 7 dB sensitivity (6 vs. 1 at 9 months and 3 vs. 0 at 4 months). Relative decreases from BL EZA were less in the treated vs. untreated eye for both subjects (-12.4% vs. -16.2% at 9 months and -20.2% vs. -28.7% at 6 months).

[00163] During the Phase 1/2 study, patients’ ocular and systemic status is closely monitored including detailed ophthalmic evaluations and retinal imaging together with blood testing and systemic examinations, as necessary. A variety of visual function and anatomical assessments are performed to detect any preliminary efficacy signal. These assessments include, but are not limited to, measurements of ellipsoid zone (EZ) area, fundus autofluorescence, microperimetry, static automated perimetry, and best corrected visual acuity (BCVA).

[00164] Conclusion

[00165] Intravitreally administered 4D-125 was well-tolerated with mild or moderate, transient, and steroid-responsive intraocular inflammation. Preliminary signs of biologic activity were observed in 2 evaluable dose escalation subjects based on microperimetry and SD-OCT. These findings support dose expansion with the 1×10^{12} vg/eye dose in XLRP subjects with less advanced disease in the ongoing Phase 1/2 study.

[00166] While the materials and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the method described herein without departing from the concept, spirit and scope of the invention. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention.

[00167] Throughout the specification and claims, unless the context requires otherwise, the word “comprise” or variations such as “comprises” or “comprising”, will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers.

CLAIMS

1. A nucleic acid encoding human retinitis pigmentosa GTPase regulator (RPGR) protein of SEQ ID NO:2 and codon optimized for expression in humans, the nucleic acid comprising the nucleotide sequence set forth as SEQ ID NO: 1 or comprising a nucleotide sequence at least 95% identical thereto.
2. The nucleic acid according to claim 1, wherein the nucleotide sequence has a codon adaptation index of at least 0.89.
3. The nucleic acid according to claim 1, comprising the nucleotide sequence set forth as SEQ ID NO: 1.
4. An expression cassette comprising the nucleic acid according to any one of claims 1 to 3, wherein the nucleotide sequence is operably linked to an expression control sequence.
5. The expression cassette of claim 4, wherein the expression control sequence comprises a constitutive promoter.
6. The expression cassette of claim 4, wherein the expression control sequence comprises a promoter that directs preferential expression of the nucleic acid in rods and cones.
7. The expression cassette of claim 6, comprising from 5' to 3': (a) an AAV2 terminal repeat (b) an hGRK promoter (c) a nucleotide sequence at least 95% identical to the nucleotide sequence set forth in SEQ ID NO:1 (d) an SV40 polyadenylation sequence and (e) an AAV2 terminal repeat.
8. The expression cassette of claim 7, wherein the 5' AAV2 terminal repeat has the nucleotide sequence set forth as SEQ ID NO:6 and/or wherein the hGRK promoter has the nucleotide sequence set forth as SEQ ID NO:4 and/or wherein the SV40 polyadenylation sequence has the nucleotide sequence set forth as SEQ ID NO:8 and/or wherein the 3' AAV2 terminal repeat has the nucleotide sequence set forth as SEQ ID NO:7.
9. The expression cassette of claim 8, comprising or consisting of the nucleotide sequence of SEQ ID NO:5 or a sequence at least 95%, or at least 98% identical thereto.
10. A vector comprising the nucleic acid according to any one of claims 1 to 3 or an expression cassette according to any one of claims 4 to 9.
11. The vector of claim 10, wherein the vector is a recombinant adeno-associated (rAAV) vector.
12. The vector of claim 11, wherein the rAAV vector comprises an AAV capsid of serotype 2, 5 or 8 or a variant thereof.
13. The vector of claim 12, wherein the rAAV vector comprises an AAV2 capsid variant comprising a capsid protein comprising or consisting of the sequence of SEQ ID NO:9.
14. The vector of any one of claims 11-13, wherein the rAAV vector comprises a nucleic acid comprising from 5' to 3': (a) an AAV2 terminal repeat (b) an hGRK promoter (c) a nucleotide

sequence at least 95% identical to the nucleotide sequence set forth in SEQ ID NO:1 and (d) an AAV2 terminal repeat.

15. The vector of claim 14, wherein the 5' AAV2 terminal repeat has the nucleotide sequence set forth as SEQ ID NO:6 and/or wherein the hGRK promoter has the nucleotide sequence set forth as SEQ ID NO:4 and/or wherein the SV40 polyadenylation sequence has the nucleotide sequence set forth as SEQ ID NO:8 and/or wherein the 3' AAV2 terminal repeat has the nucleotide sequence set forth as SEQ ID NO:7.

16. The vector of claim 15, wherein the rAAV vector comprises a nucleic acid comprising the nucleotide sequence of SEQ ID NO:5 or a sequence, at least 95% or at least 98% identical thereto.

17. The vector of claim 16, wherein the rAAV vector comprises (i) a capsid comprising a capsid protein comprising or consisting of the sequence of SEQ ID NO:9 and (ii) a nucleic acid comprising or consisting of the nucleotide sequence of SEQ ID NO:5.

18. A host cell comprising the nucleic acid according to any one of claims 1 to 3 or an expression cassette according to any one of claims 4 to 9.

19. The host cell according to claim 18, wherein the host cell is a mammalian cell.

20. The host cell of claim 18 or 19, wherein the host cell is a CHO cell, an HEK293 cell, an HEK293T cell, a HeLa cell, a BHK21 cell or a Vero cell and/or wherein the host cell is grown in a suspension or cell stack culture and/or wherein the host cell is a photoreceptor cell, a retinal ganglion cell, a glial cell, a bipolar cell, an amacrine cell, a horizontal cell, or a retinal pigmented epithelium cell.

21. A method of treating XLRP in a subject in need thereof, comprising administering to the subject a therapeutically effective amount of a nucleic acid according to any one of claims 1-3, an expression cassette according to any one of claims 4-9 or a vector according to any one of claims 10-17.

22. A method of treating XLRP in a subject in need thereof, comprising administering to the subject an infectious rAAV comprising (i) an AAV capsid and (ii) a nucleic acid comprising from 5' to 3': (a) an AAV2 terminal repeat (b) an hGRK promoter (c) a nucleotide sequence at least 95% identical to the nucleotide sequence set forth in SEQ ID NO:1 (d) an SV40 polyadenylation sequence and (e) an AAV2 terminal repeat.

23. The method according to claim 22, wherein the 5' AAV2 terminal repeat has the nucleotide sequence set forth as SEQ ID NO:6 and/or wherein the hGRK promoter has the nucleotide sequence set forth as SEQ ID NO:4 and/or wherein the SV40 polyadenylation sequence has the nucleotide sequence set forth as SEQ ID NO:8 and/or wherein the 3' AAV2 terminal repeat has the nucleotide sequence set forth as SEQ ID NO:7.

24. The method according to claim 22 or 23, wherein the rAAV comprises (i) a capsid comprising a capsid protein comprising or consisting of the sequence of SEQ ID NO:9 and (ii) a nucleic acid comprising or consisting of the nucleotide sequence of SEQ ID NO:5.

25. The method according to any one of claims 21-24, wherein the nucleic acid or vector is administered to the subject by periocular, intravitreal, suprachoroidal or subretinal injection and/or wherein the vector is administered to the subject at a dosage from about 10^{10} vector genomes (vg)/eye to about 10^{13} vg/eye.

26. The method according to claim 25, wherein the vector is an rAAV vector and is administered to the subject at a dosage of about 3×10^{11} vg/eye or at a dosage of about 1×10^{12} vg/eye.

27. Use of nucleic acid according to any one of claims 1-3, an expression cassette according to any one of claims 4-9 or a vector according to any one of claims 10-17 in the manufacture of a medicament for the treatment of XLRP.

28. The method according to claim 26 or the use according to claim 27, wherein the nucleic acid or vector is administered by periocular, intravitreal, suprachoroidal or subretinal injection and/or wherein the vector is for administration at a dosage from about 10^{10} vector genomes (vg)/eye to about 10^{13} vg/eye.

29. Use of an infectious rAAV comprising (i) an AAV capsid and (ii) a nucleic acid comprising from 5' to 3': (a) an AAV2 terminal repeat (b) an hGRK promoter (c) a nucleotide sequence at least 95% identical to the nucleotide sequence as set forth in SEQ ID NO:1 (d) an SV40 polyadenylation sequence and (e) an AAV2 terminal repeat, in the manufacture of a medicament for the treatment of XLRP.

30 The use according to claim 29, wherein the 5' AAV2 terminal repeat has the nucleotide sequence set forth as SEQ ID NO:6 and/or wherein the hGRK promoter has the nucleotide sequence set forth as SEQ ID NO:4 and/or wherein the SV40 polyadenylation sequence has the nucleotide sequence set forth as SEQ ID NO:8 and/or wherein the 3' AAV2 terminal repeat has the nucleotide sequence set forth as SEQ ID NO:7.

31. The use according to claim 30, wherein the rAAV comprises (i) a capsid comprising a capsid protein comprising or consisting of the sequence of SEQ ID NO:9 and (ii) a nucleic acid comprising or consisting of the nucleotide sequence of SEQ ID NO:5.

32. The use according to any one of claims 29-31, wherein the rAAV is administered by intravitreal injection and/or wherein the vector is administered at a dosage from about 10^{10} vector genomes (vg)/eye to about 10^{13} vg/eye.

33. A pharmaceutical composition comprising a nucleic acid according to any one of claims 1-3, an expression cassette according to any one of claims 4-9 or a vector according to any one of claims 10-17, and at least one pharmaceutically acceptable excipient.

34. A pharmaceutical composition comprising an infectious rAAV comprising (i) an AAV capsid and (ii) a nucleic acid comprising from 5' to 3': (a) an AAV2 terminal repeat (b) an hGRK promoter (c) a nucleotide sequence at least 95% identical to the nucleotide sequence set forth in SEQ ID NO:1 (d) an SV40 polyadenylation sequence and (e) an AAV2 terminal repeat.

35. The pharmaceutical composition according to claim 34, wherein the 5' AAV2 terminal repeat has the nucleotide sequence set forth as SEQ ID NO:6 and/or wherein the hGRK promoter has the nucleotide sequence set forth as SEQ ID NO:4 and/or wherein the SV40 polyadenylation sequence has the nucleotide sequence set forth as SEQ ID NO:8 and/or wherein the 3' AAV2 terminal repeat has the nucleotide sequence set forth as SEQ ID NO:7.

36. The pharmaceutical composition according to claim 35, wherein the rAAV comprises (i) a capsid comprising a capsid protein comprising or consisting of the sequence of SEQ ID NO:9 and (ii) a nucleic acid comprising or consisting of the nucleotide sequence of SEQ ID NO:5.

37. The pharmaceutical composition according to any one of claims 34-36, wherein the pharmaceutical composition comprises between 10^9 vg and 10^{14} vg of the rAAV.

38. The pharmaceutical composition of claim 37, wherein the pharmaceutical composition comprises between 10^{10} vg and 10^{13} vg of the rAAV.

39. The pharmaceutical composition of claim 37, wherein the pharmaceutical composition comprises about 3×10^{11} vg or about 1×10^{12} vg of the rAAV.

FIGURE 1

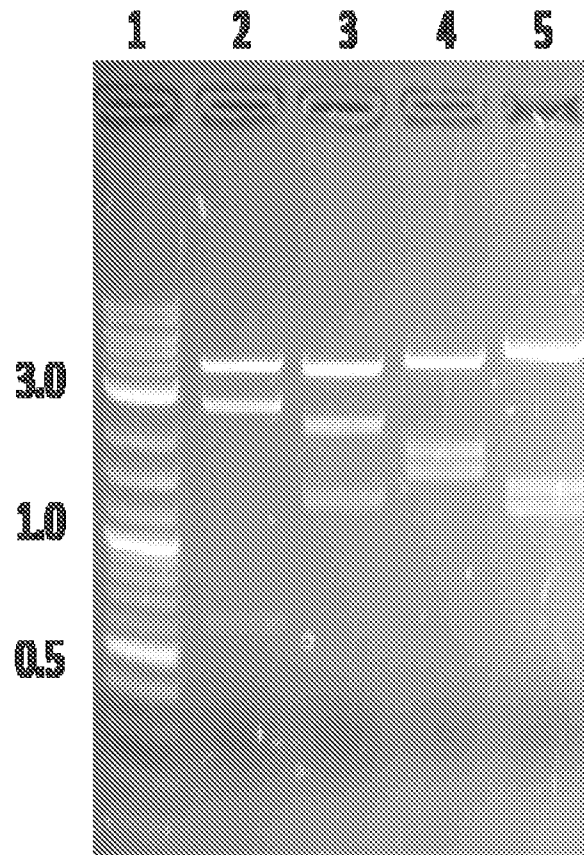


FIGURE 2

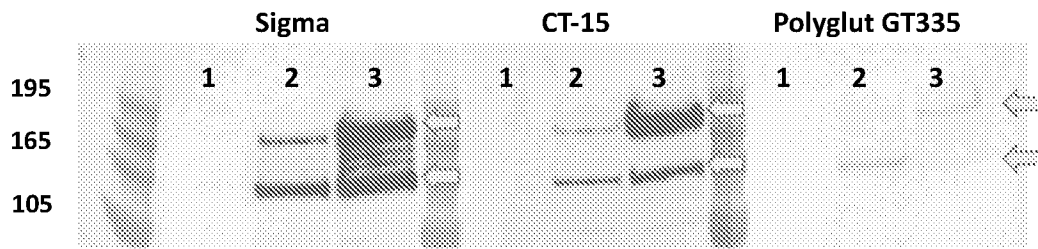


FIGURE 3

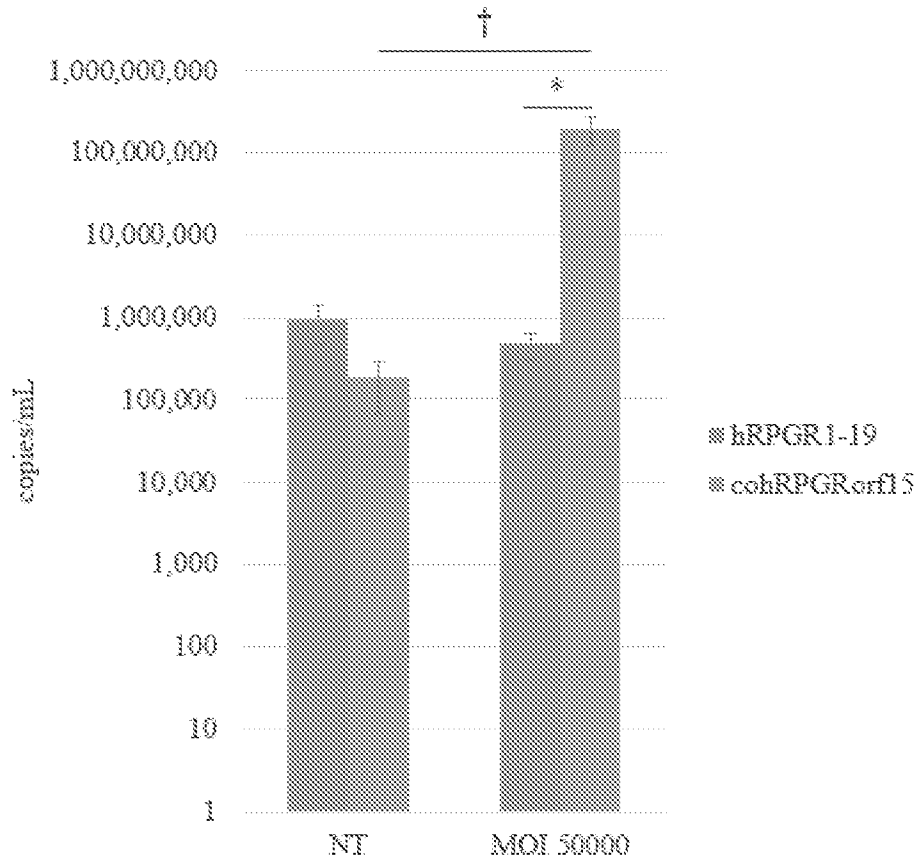


FIGURE 4

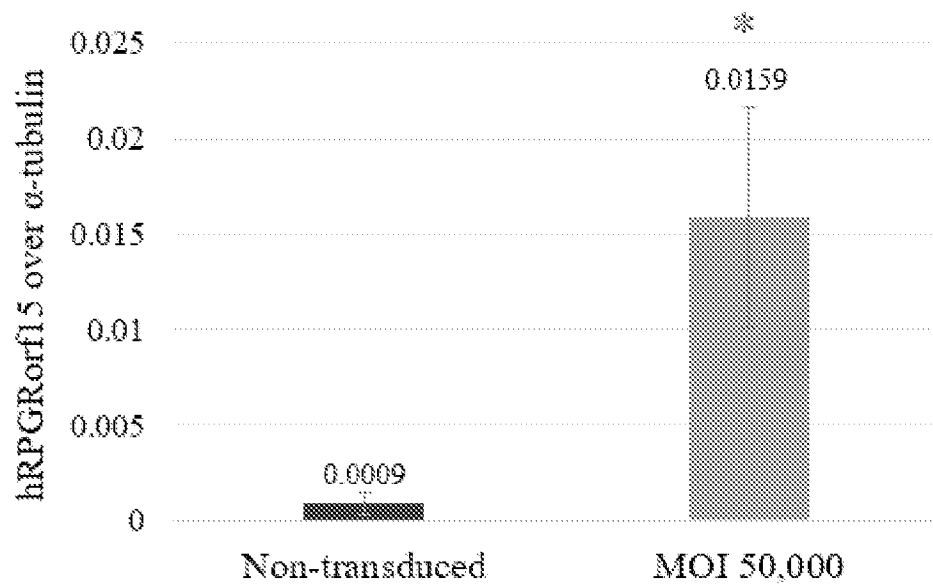


FIGURE 5

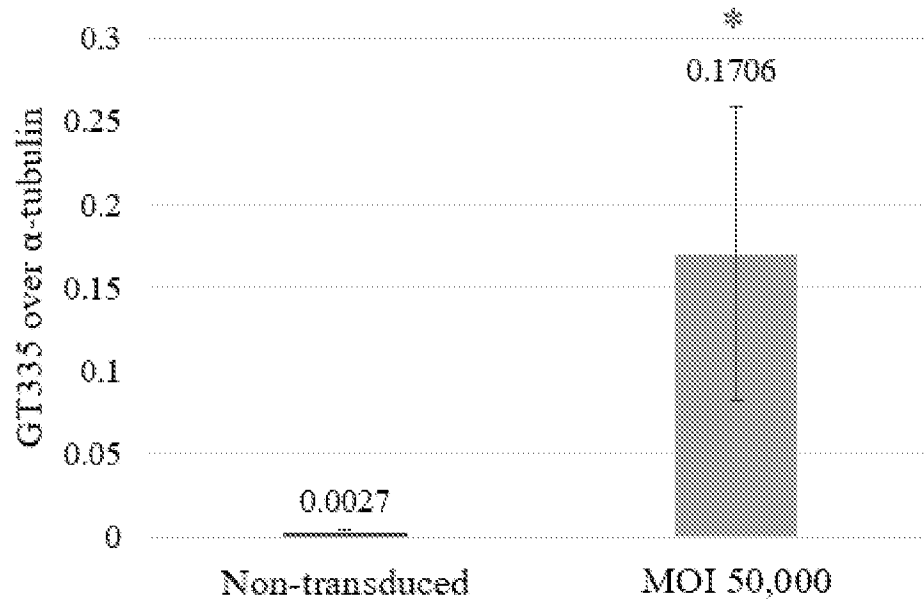


FIGURE 6

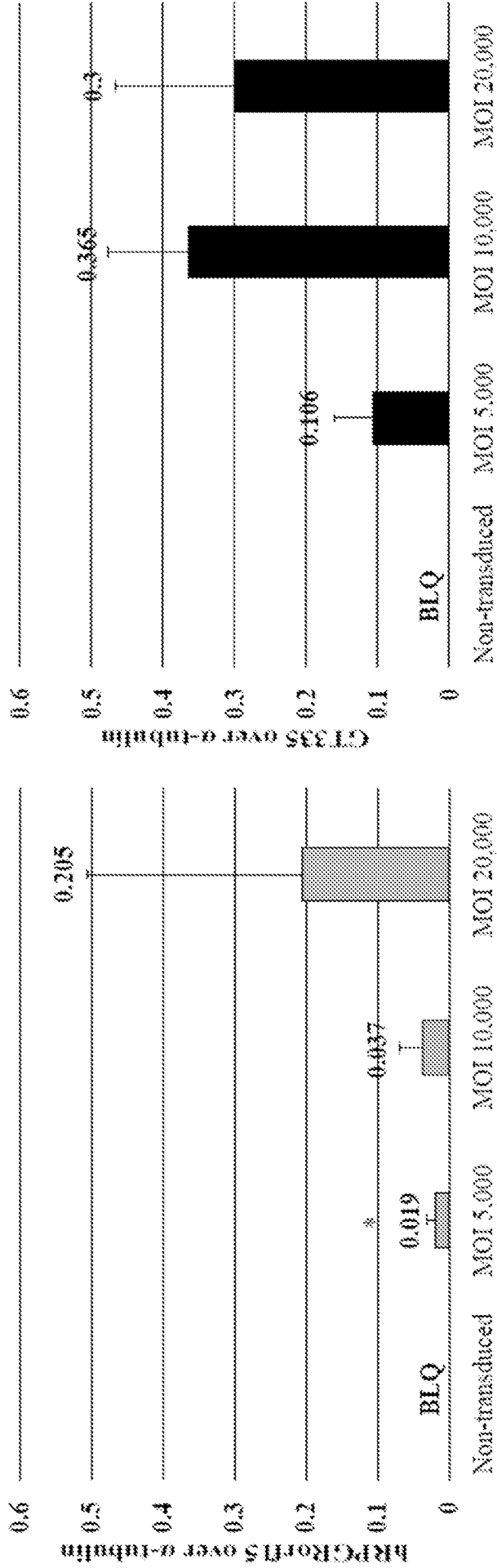


FIGURE 7

NarI

KasI

ATGAGAGAACCCGAGGAAGTCTGATGCCCCACTCTGGCCCGTGTTTACCTTCGGCAAGAGC
 1 -----+-----+-----+-----+-----+-----+-----+
 TACTCTCTTGGGCTCCTTGACTACGGGCTGAGACCGCGGCACAAATGGAAGCCGTTCTCG
M R E P E E L M P D S G A V F T F G K S
1 3 5 7 9 11 13 15 17 19

AAGTTCGCCGAGAACACCCCGGCAAGTTCTGGTTCAAGAACGACGTGCCAGTGCACCTG 61 --
 -----+-----+-----+-----+-----+-----+-----+
 TTCAAGCGGCTCTTGTGGGGCCGTTCAAGACCAAGTTCTTGCTGCACGGTACAGTGGAC
K F A E N N P G K F W F K N D V P V H L
21 23 25 27 29 31 33 35 37 39

BstEII

AGCTGCGGAGATGAACACTCTGCCGTGGTCAACCGGCAACAACAAGCTGTACATGTTCCGGC
 121 -----+-----+-----+-----+-----+-----+-----+
 TCGACGCCTCTACTTGTGAGACGGCACCAGTGGCCGTTGTTGTTTCGACATGTACAAGCCG
S C G D E H S A V V T G N N K L Y M F G
41 43 45 47 49 51 53 55 57 59

BspMI

AGCAACAACCTGGGGCCAGCTCGGCCCTGGGATCTAAGTCTGCCATCAGCAAGCCTACCTGC
 181 -----+-----+-----+-----+-----+-----+-----+
 TCGTTGTTGACCCCGGTCGAGCCGACCCCTAGATTTCAGACGGTAGTCGTTCCGATGGACG
S N N W G Q L G L G S K S A I S K P T C
61 63 65 67 69 71 73 75 77 79

GTGAAGGCCCTGAAGCCTGAGAAAGTGAACCTGGCCGCTGCGGCAGAAATCACACCCTG 241 -
 -----+-----+-----+-----+-----+-----+-----+
 CACTTCCGGGACTTCCGACTCTTTTCACTTTGACCGGCGGACGCCGTCTTTAGTGTGGGAC
V K A L K P E K V K L A A C G R N H T L
81 83 85 87 89 91 93 95 97 99

GTTTCTACCGAAGGCGCAATGTGTATGCCACCGCGGAAACAATGAGGGACAGCTTGA 301 -
 -----+-----+-----+-----+-----+-----+-----+
 CAAAGATGGCTTCCGCGTTACACATACGGTGGCCGCTTTGTTACTCCCTGTCTGAACCT
V S T E G G N V Y A T G G N N E G Q L G
101 103 105 107 109 111 113 115 117 119

BclI

CTGGGCGACACCGAGGAAAGAAACACCTTCCACGTGATCAGCTTTTTCCACGCGGAGCAC
 361 -----+-----+-----+-----+-----+-----+-----+
 GACCCGCTGTGGCTCCTTTCTTTGTGGAAGGTGCACTAGTCGAAAAAGTGGTCCGCTCGTG
L G D T E E R N T F H V I S F F T S E H
121 123 125 127 129 131 133 135 137 139

PvuII

AAGATCAAGCAGCTGAGCGCCGCTCTAATACCTCTGCCGCTCTGACAGAGGACGGCAGA
 421 -----+-----+-----+-----+-----+-----+-----+
 TTCTAGTTCGTCGACTCGCGCCGAGATTATGGAGACGGCGAGACTGTCTCCTGCCGTCT
K I K Q L S A G S N T S A A L T E D G R
141 143 145 147 149 151 153 155 157 159

CTGTTTATGTGGGGCGACAATTCTGAGGGCCAGATCGGACTGAAGAACGTGTCCAATGTG 481 -
 -----+-----+-----+-----+-----+-----+-----+
 GACAAATACACCCCGCTGTTAAGACTCCCGGTCTAGCCTGACTTCTTGCACAGGTTACAC
L F M W G D N S E G Q I G L K N V S N V
161 163 165 167 169 171 173 175 177 179

FIGURE 7 (cont'd)

PvuII

TGCGTGCCCCAGCAAGTGACAATCGGCAAGCCTGTGTCTTGGATCAGCTGCGGCTACTAC
 541 -----+-----+-----+-----+-----+-----+-----+
 ACGCACGGGGTCGTTCACTGTTAGCCGTTCCGGACACAGAACCTAGTCGACGCCGATGATG
 C V P Q Q V T I G K P V S W I S C G Y Y
 181 183 185 187 189 191 193 195 197 199

Pf1MI

CACAGCGCCTTTGTGACAACCGATGGCGAGCTGTATGTGTTCCGGCGAGCCAGAGAATGGC
 601 -----+-----+-----+-----+-----+-----+-----+
 GTGTCCGGAAACACTGTTGGCTACCGCTCGACATACACAAGCCGCTCGGTCTCTTACCG
 H S A F V T T D G E L Y V F G E P E N G
 201 203 205 207 209 211 213 215 217 219

PvuII
Pf1MI *PvuII*

AAGCTGGGACTGCCTAACCAGCTGCTGGGCAATCACAGAACCCTCAGCTGGTGTCTGAG
 661 -----+-----+-----+-----+-----+-----+-----+
 TTCGACCCTGACGGATTGGTCGACGCCGTTAGTGTCTTGGGGAGTCGACCACAGACTC
 K L G L P N Q L L G N H R T P Q L V S E
 221 223 225 227 229 231 233 235 237 239

ATCCCCGAAAAGTGATCCAGGTGGCCTGTGGCGGAGAGCACACAGTGGTGTGACAGAG 721 -
 -----+-----+-----+-----+-----+-----+-----+
 TAGGGGCTTTTTCACTAGGTCCACCGGACACCGCCTCTCGTGTGTACCACGACTGTCTC
 I P E K V I Q V A C G G E H T V V L T E
 241 243 245 247 249 251 253 255 257 259

AATGCCGTGTACACCTTTGGCCTGGGCCAGTTTGGACAACCTCGGACTGGGAACCTCCTG 781 -
 -----+-----+-----+-----+-----+-----+-----+
 TTACGGCACATGTGGAAACCGGACCCGGTCAAACCTGTTGAGCCTGACCCTTGGGAAGGAC
 N A V Y T F G L G Q F G Q L G L G T F L
 261 263 265 267 269 271 273 275 277 279

PvuII

TTCGAGACAAGCGAGCCCAAAGTGATCGAGAACATCCGGGACCAGACCATCAGCTACATC
 841 -----+-----+-----+-----+-----+-----+-----+
 AAGCTCTGTTCCGCTCGGTTTCACTAGCTCTTGTAGGCCCTGGTCTGGTAGTCGATGTAG
 F E T S E P K V I E N I R D Q T I S Y I
 281 283 285 287 289 291 293 295 297 299

BclI

AGCTGTGGCGAGAACCACACAGCCCTGATCACAGACATCGGCCTGATGTACACATTCGGC
 901 -----+-----+-----+-----+-----+-----+-----+
 TCGACACCGCTCTTGGTGTGTCCGGACTAGTGTCTGTAGCCGGACTACATGTGTAAGCCG
 S C G E N H T A L I T D I G L M Y T F G
 301 303 305 307 309 311 313 315 317 319

GACGGAAGGCATGGAAAGCTCGGACTTGGCCTGGAAAACCTTACCAACCACTTCATCCCT 961
 -----+-----+-----+-----+-----+-----+-----+
 CTGCCTTCCGTACCTTCGAGCCTGAACCGGACCTTTGAAGTGGTTGGTGAAGTAGGGA
 D G R H G K L G L G L E N F T N H F I P
 321 323 325 327 329 331 333 335 337 339

Pf1MI

ACGCTGTGCAGCAACTTCCTGCGGTTCAATTGTGAAGCTGGTGGCCTGCGGAGGATGCCAC
 1021 -----+-----+-----+-----+-----+-----+-----+
 TCGCACCGTCGTTGAAGGACGCCAAGTAACACTTCGACCACCGGACGCCTCCTACGGTG
 T L C S N F L R F I V K L V A C G G C H
 341 343 345 347 349 351 353 355 357 359

FIGURE 7 (cont'd)

ATGGTGGTTTTTGGCTGCCCTCACAGAGGCGTGGCCAAAGAGATTGAGTTCGACGAGATC
 1081 -----+-----+-----+-----+-----+-----+-----+
 TACCACCAAAAACGACGGGGAGTGTCTCCGCACCGGTTTCTCTAACTCAAGCTGCTCTAG
 M V V F A A P H R G V A K E I E F D E I
 361 363 365 367 369 371 373 375 377 379

BspMI
 AACGATACCTGCCTGAGCGTGGCCACCTTCCTGCCTTACAGCAGCCTGACATCTGGCAAC
 1141 -----+-----+-----+-----+-----+-----+-----+
 TTGCTATGGACGGACTCGCACCGGTGGAAGGACGGAATGTCGTCGGACTGTAGACCGTTG
 N D T C L S V A T F L P Y S S L T S G N
 381 383 385 387 389 391 393 395 397 399

PstI
 GTGCTGCAGAGGACACTGAGCGCCAGAATGCGCAGACGGGAAAAGAGAGAAGCCCCGAC
 1201 -----+-----+-----+-----+-----+-----+-----+
 CACGACGTCTCCTGTGACTCGCGGTCTACGCGTCTGCCCTTCTCTCTCTCCGGGGTG
 V L Q R T L S A R M R R R E R E R S P D
 401 403 405 407 409 411 413 415 417 419

AGCTTCAGCATGAGAAGAACCCTGCCTCCAATCGAGGGCACACTGGGCCTGTCTGCCTGC
 1261 -----+-----+-----+-----+-----+-----+-----+
 TCGAAGTCGTACTCTTCTGGGACGGAGTTAGCTCCCGTGTGACCCGGACAGACGGACG
 S F S M R R T L P P I E G T L G L S A C
 421 423 425 427 429 431 433 435 437 439

BspMI
 TTTCTGCCTAACAGCGTGTTCCTCCAGATGCAGCGAGAGAAACCTGCAAGAGAGCGTGCTG
 1321 -----+-----+-----+-----+-----+-----+-----+
 AAAGACGGATTGTTCGACCAAGGGGTCTACGTGCGTCTCTTTGGACGTTCTCTCGCACGAC
 F L P N S V F P R C S E R N L Q E S V L
 441 443 445 447 449 451 453 455 457 459

BspMI
 AGCGAGCAGGATCTGATGCAGCCTGAGGAACCCGACTACCTGCTGGACGAGATGACCAAA
 1381 -----+-----+-----+-----+-----+-----+-----+
 TCGCTCGTCCTAGACTACGTCCGACTCCTTGGGCTGATGGACGACCTGCTCTACTGTTTT
 S E Q D L M Q P E E P D Y L L D E M T K
 461 463 465 467 469 471 473 475 477 479

GAGGCCGAGATCGACAACAGCAGCACAGTGGAAAGCCTGGGCGAGACAACCGACATCCTG
 1441 -----+-----+-----+-----+-----+-----+-----+
 CTCCGGCTCTAGCTGTTGTCGTGTCACCTTTCGGACCCGCTCTGTTGGCTGTAGGAC
 E A E I D N S S T V E S L G E T T D I L
 481 483 485 487 489 491 493 495 497 499

AACATGACCCACATCATGAGCCTGAACAGCAACGAGAAGTCTCTGAAGCTGAGCCCCGTG
 1501 -----+-----+-----+-----+-----+-----+-----+
 TTGTAAGGGTGTAGTACTCGGACTTGTGCTTTCAGAGACTTCGACTCGGGGCAC
 N M T H I M S L N S N E K S L K L S P V
 501 503 505 507 509 511 513 515 517 519

CAGAAGCAGAAGAAGCAGCAGACCATCGGCGAGCTGACACAGGATACTGCCCTGACCGAG
 1561 -----+-----+-----+-----+-----+-----+-----+
 GTCTTCGTCTTCTTCGTCGTCTGGTAGCCGCTCGACTGTGTCCTATGACGGGACTGGCTC
 Q K Q K K Q Q T I G E L T Q D T A L T E
 521 523 525 527 529 531 533 535 537 539

FIGURE 7 (cont'd)

StuI

AACGACGACAGCGACGAGTACGAAGAGATGAGCGAGATGAAGGAAGGCAAGGCCTGCAAG
 1621 -----+-----+-----+-----+-----+-----+-----+
 TTGCTGCTGTCGCTGCTCATGCTTCTCTACTCGCTCTACTTCCCTCCGGTCCGGACGTTC
 N D D S D E Y E E M S E M K E G K A C K
 541 543 545 547 549 551 553 555 557 559

StuI

CAGCACGTGTCCCAGGGCATCTTTATGACCCAGCCTGCCACCACCATCGAGGCCTTTTCC
 1681 -----+-----+-----+-----+-----+-----+-----+
 GTCGTGCACAGGGTCCCGTAGAAATACTGGGTCCGACGGTGGTGGTAGCTCCGGAAAAAGG
 Q H V S Q G I F M T Q P A T T I E A F S
 561 563 565 567 569 571 573 575 577 579

NarI
KasI

GACGAGGAAGTGGAAATCCCCGAGGAAAAAGAGGGCGCCGAGGACAGCAAAGGCAACGGC
 1741 -----+-----+-----+-----+-----+-----+-----+
 CTGCTCCTTCACCTTTAGGGGCTCCTTTTCTCCCGCGGCTCCTGTCTTCCGTTGCCG
 D E E V E I P E E K E G A E D S K G N G
 581 583 585 587 589 591 593 595 597 599

ATTGAGGAACAAGAGGTGGAAGCCAACGAAGAGAACGTGAAGGTGCACGGCGGACGGAAA
 1801 -----+-----+-----+-----+-----+-----+-----+
 TAACTCCTTGTCTCCACCTTCGGTTGCTTCTCTTGCACTTCCACGTGCCGCTGCCTTT
 I E E Q E V E A N E E N V K V H G G R K
 601 603 605 607 609 611 613 615 617 619

GAAAAGACCGAGATCCTGAGCGACGACCTGACCGATAAGGCCGAGGTTTCCGAGGGCAAA
 1861 -----+-----+-----+-----+-----+-----+-----+
 CTTTTCTGGCTCTAGGACTCGCTGCTGGACTGGCTATTCCGGCTCCAAAGGCTCCCGTTT
 E K T E I L S D D L T D K A E V S E G K
 621 623 625 627 629 631 633 635 637 639

SacII

GCCAAGTCTGTGGGAGAAGCCGAGGATGGACCTGAAGGCCGCGGAGATGGAACCTGTGAA
 1921 -----+-----+-----+-----+-----+-----+-----+
 CGGTTTCAGACACCCTCTTCGGCTCCTACCTGGACTTCCGGCGCCTCTACCTTGACACTT
 A K S V G E A E D G P E G R G D G T C E
 641 643 645 647 649 651 653 655 657 659

GAAGGATCTAGCGGAGCCGAGCACTGGCAGGATGAGGAACGCGAGAAGGGCGAGAAAGAC
 1981 -----+-----+-----+-----+-----+-----+-----+
 CTTCTAGATCGCCTCGGCTCGTGACCGTCTACTCCTTGCCTCTTCCCGCTCTTTCTG
 E G S S G A E H W Q D E E R E K G E K D
 661 663 665 667 669 671 673 675 677 679

AAAGGCAGAGGCGAGATGGAAAGACCCGGCGAGGGCGAAAAAGAGCTGGCCGAGAAAGAG
 2041 -----+-----+-----+-----+-----+-----+-----+
 TTTCCGTCTCCGCTCTACCTTTCTGGGCCGCTCCCGCTTTTCTCGACCGGCTCTTTCTC
 K G R G E M E R P G E G E K E L A E K E
 681 683 685 687 689 691 693 695 697 699

GAATGGAAGAAACGCGACGGCGAAGAACAAGAGCAGAAAGAAAGAGAGCAGGGCCACCAG
 2101 -----+-----+-----+-----+-----+-----+-----+
 CTTACCTTCTTTGCGTCCGCTTCTGTTCTCGTCTTTCTTTCTCTCGTCCCGGTGGTC
 E W K K R D G E E Q E Q K E R E Q G H Q
 701 703 705 707 709 711 713 715 717 719

FIGURE 7 (cont'd)

AAAGAACGGAATCAAGAGATGGAAGAAGGCGGCCGAGGAAGAACACGGCGAAGGGGAAGAA
2161 -----+-----+-----+-----+-----+-----+-----+
TTTCTTGCCTTAGTTCTCTACCTTCTTCCGCGCTCCTTCTTGTGCCGCTTCCCCTTCTT
K E R N Q E M E E G G E E E H G E G E E
721 723 725 727 729 731 733 735 737 739

GAGGAAGGCGACCGAGAGGAAGAAGAAGAAAGAAGGCGAAGGCAAAGAAGAAGGCGAG
2221 -----+-----+-----+-----+-----+-----+-----+
CTCCTTCCGCTGGCTCTCCTTCTTCTTCTTTCTTCCGCTTCCGTTCTTCTTCCGCTC
E E G D R E E E E E K E G E G K E E G E
741 743 745 747 749 751 753 755 757 759

GGCGAAGAGGTGGAAGGCGAGCGTGAAAAAGAAGAGGGCGAACGCAAGAAAGAAGAACGC
2281 -----+-----+-----+-----+-----+-----+-----+
CCGCTTCTCCACCTTCCGCTCGACTTTTTCTTCTCCGCTTGCCTTCTTTCTTCTTGGC
G E E V E G E R E K E E G E R K K E E R
761 763 765 767 769 771 773 775 777 779

GCCGAAAAGAGGAAAAGGCGAGGAAGAGGGCGACCAAGGCGAAGGCGAGGAAGAAGAA
2341 -----+-----+-----+-----+-----+-----+-----+
CGGCCTTTTCTCCTTTTCCGCTCCTTCTCCGCTGGTTCCGCTTCCGCTCCTTCTTCTT
A G K E E K G E E E G D Q G E G E E E E
781 783 785 787 789 791 793 795 797 799

ACTGAAGGCAGAGGGGAAGAGAAGAGGAAGGCGGCGAAGTCGAAGGCGGAGAGGTTGAA
2401 -----+-----+-----+-----+-----+-----+-----+
TGACTCCGTCTCCCCTTCTCTTTCTCCTTCCGCGCTTCCGCTTCCGCTTCCCACTT
T E G R G E E K E E G G E V E G G E V E
801 803 805 807 809 811 813 815 817 819

GAAGGCAAAGGCGAGCGAGAAGAGGAAGAAGAAGGCGAAGGCGAGGAAGAGGAAGGC
2461 -----+-----+-----+-----+-----+-----+-----+
CTTCCGTTTCCGCTCGCTCTTCTCCTTCTTCTTCTTCCGCTTCCGCTCCTTCTCCTCCG
E G K G E R E E E E E E G E G E E E E G
821 823 825 827 829 831 833 835 837 839

GAAGGCGAAGAGGAAGAAGGCGAAGGGGAAGAAGAAGAAGGCGAAGGCAAGGGCGAAGAG
2521 -----+-----+-----+-----+-----+-----+-----+
CTTCCGCTTCTCCTTCTTCCGCTTCCGCTTCTTCTTCTTCCGCTTCCGTTCCGCTTCTC
E G E E E E G E G E E E E G E G K G E E
841 843 845 847 849 851 853 855 857 859

GAGGGCGAAGAAGGCGAGGGCGAAGAGGGCGAAGAAGGCGAAGGCGAGGGCGAAGAA
2581 -----+-----+-----+-----+-----+-----+-----+
CTCCGCTTCTTCCGCTCCGCTTCTCCTCCGCTTCTTCCGCTTCCGCTTCCGCTTCTT
E G E E G E G E E E G E E G E G E G E E
861 863 865 867 869 871 873 875 877 879

GAAGAAGGCGAAGGCGAAGGCGAGGAAGAAGGCGAAGGCGAAGGGGAAGAAGAGGAAGGC
2641 -----+-----+-----+-----+-----+-----+-----+
CTTCTTCCGCTTCCGCTTCCGCTCCTTCTTCCGCTTCCGCTTCCGCTTCCGCTTCTTCTCCTTCCG
E E G E G E G E E E G E G E G E E E E G
881 883 885 887 889 891 893 895 897 899

GAAGGCGAAGGCGAAGAAGAAGGCGAAGGCGAGGGCGAAGAGGAAGAAGGCGAAGGCGAAA
2701 -----+-----+-----+-----+-----+-----+-----+
CTTCCGCTTCCGCTTCTTCTTCCGCTTCCGCTTCCGCTTCCGCTTCTCCTTCTTCCGCTTCCGTTT
E G E G E E E G E G E G E E E E G E G K
901 903 905 907 909 911 913 915 917 919

FIGURE 7 (cont'd)

2761 -----+-----+-----+-----+-----+-----+-----+
 GGGGAAGAAGAAGGCCGAGGAAGGCCAAGGCCAAGGCCGAGGAAGAAGAAGGCCAAGGCCGAG
 CCCCTTCTTCTCCGCTCCTTCCGCTTCCGCTTCCGCTCCTTCTTCTCCGCTTCCGCTC
 G E E E G E E G E E G E E G E E E E G E E
 921 923 925 927 929 931 933 935 937 939

2821 -----+-----+-----+-----+-----+-----+-----+
 GGCGAAGATGGCGAAGGCCAAGGCCAAGAGGAAGAGGGCGAGTGGGAGGGCGAAGAAGAG
 CCGCTTCTACCGCTTCCGCTTCCGCTTCTCCTTCTCCCGCTCACCTCCCGCTTCTTCTC
 G E D G E G E G E E E E G E W E G E E E
 941 943 945 947 949 951 953 955 957 959

2881 -----+-----+-----+-----+-----+-----+-----+
 GAAGGCCAAGGCCGAGGGCGAAGAGGAAGGCCAAGGCCGAGGGCGAAGAAGGCCAAGGCCGAA
 CTTCCGCTTCCGCTCCGCTTCTCCTTCCGCTTCCGCTCCGCTTCTCCGCTTCCGCTT
 E G E G E G E E E G E G E G E E G E G E
 961 963 965 967 969 971 973 975 977 979

2941 -----+-----+-----+-----+-----+-----+-----+
 GGCGAGGAAGAGGAAGGCCAAGGCCAAGGGGAAGAAGAAGAGGGCGAAGAAGAAGGCCGAA
 CCGCTCCTTCTCCTTCCGCTTCCGCTTCCCTTCTTCTTCTCCCGCTTCTTCTCCGCTT
 G E E E E G E G E G E E E E G E E E G E
 981 983 985 987 989 991 993 995 997 999

3001 -----+-----+-----+-----+-----+-----+-----+
 GAGGAAGGCCAAGGGGAAGAAGAAGGCCAAGGCCAAGGCCAAGAAGAGGAAGAGGGCGAA
 CTCCTTCCGCTTCCCCTTCTTCTTCCGCTTCCGCTTCCGCTTCTTCTCCTTCTCCCGCTT
 E E G E G E E E G E G E G E E E E E G E
 1001 1003 1005 1007 1009 1011 1013 1015 1017 1019

3061 -----+-----+-----+-----+-----+-----+-----+
 GTTGAAGGCCGAGGTTGAGGGCGAAGAAGGCCAAGGCCAAGGGGAAGAAGAAGAAGGCCGAG
 CAACTCCGCTCCAACTCCCGCTTCTTCCGCTTCCGCTTCCCTTCTTCTTCTTCCGCTC
 V E G E V E G E E G E G E G E E E E G E
 1021 1023 1025 1027 1029 1031 1033 1035 1037 1039

3121 -----+-----+-----+-----+-----+-----+-----+
 GAAGAAGGGGAAGAGAGAGAAAAAGAAGGCCGAGGGCGAAGAAAACCGCCGAACCGCGAA
 CTTCTTCCCCTTCTCTCTTTTTCTTCCGCTCCCGCTTCTTTTGGGGCCTTGGCGCTT
 E E G E E R E K E G E G E E N R R N R E
 1041 1043 1045 1047 1049 1051 1053 1055 1057 1059

3181 -----+-----+-----+-----+-----+-----+-----+
 GAGGAAGAGGAAGAAGAGGGCAAGTACCAAGAGACTGGCGAGGAAGAGAACGAGCGGCAG
 CTCCTTCTCCTTCTTCCCGTTTCTGTTCTCTGACCGCTCCTTCTTCTGCTCGCCGTC
 E E E E E E G K Y Q E T G E E E N E R Q
 1061 1063 1065 1067 1069 1071 1073 1075 1077 1079

3241 -----+-----+-----+-----+-----+-----+-----+
 GATGGCGAAGAGTACAAGAAGGTGTCCAAGATCAAGGGCAGCGTGAAGTACGGCAAGCAC
 CTACCGCTTCTCATGTTCTTCCACAGGTTCTAGTTCCCGTCCGACTTCATGCCGTTCCGTG
 D G E E Y K K V S K I K G S V K Y G K H
 1081 1083 1085 1087 1089 1091 1093 1095 1097 1099

3301 -----+-----+-----+-----+-----+-----+-----+
 AAGACCTACCAGAAGAAGTCCGTCACCAACACGCAAGGCAATGGAAAAGAACAGCGGAGC
 TTCTGGATGGTCTTCTCAGGCAGTGGTTGTGCGTTCCGTTACCTTTTCTTGTGCGCTCG
 K T Y Q K K S V T N T Q G N G K E Q R S
 1101 1103 1105 1107 1109 1111 1113 1115 1117 1119

FIGURE 7 (concluded)

```

AAGATGCCCGTGCAGTCCAAGAGGCTGCTGAAGAATGGCCCTAGCGGCAGCAAGAAATTC
3361 -----+-----+-----+-----+-----+-----+-----+
TTCTACGGGCACGTCAGGTTCTCCGACGACTTCTTACCGGGATCGCCGTCGTTCTTTAAG
K M P V Q S K R L L K N G P S G S K K F
1121 1123 1125 1127 1129 1131 1133 1135 1137 1139

```

XhoI

```

TGGAACAATGTGCTGCCCCACTACCTCGAGCTGAAGTGA
3421 -----+-----+-----+-----+-----+-----+-----+
ACCTTGTTACACGACGGGGTGATGGAGCTCGACTTCACT
W N N V L P H Y L E L K *
1141 1143 1145 1147 1149 1151 1153

```

FIGURE 8

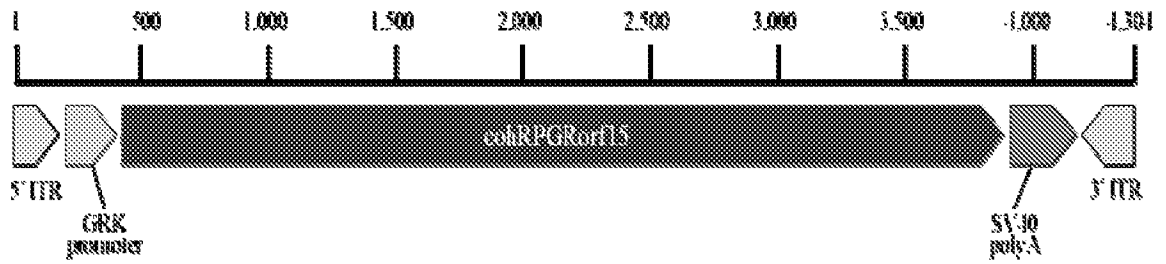


FIGURE 9

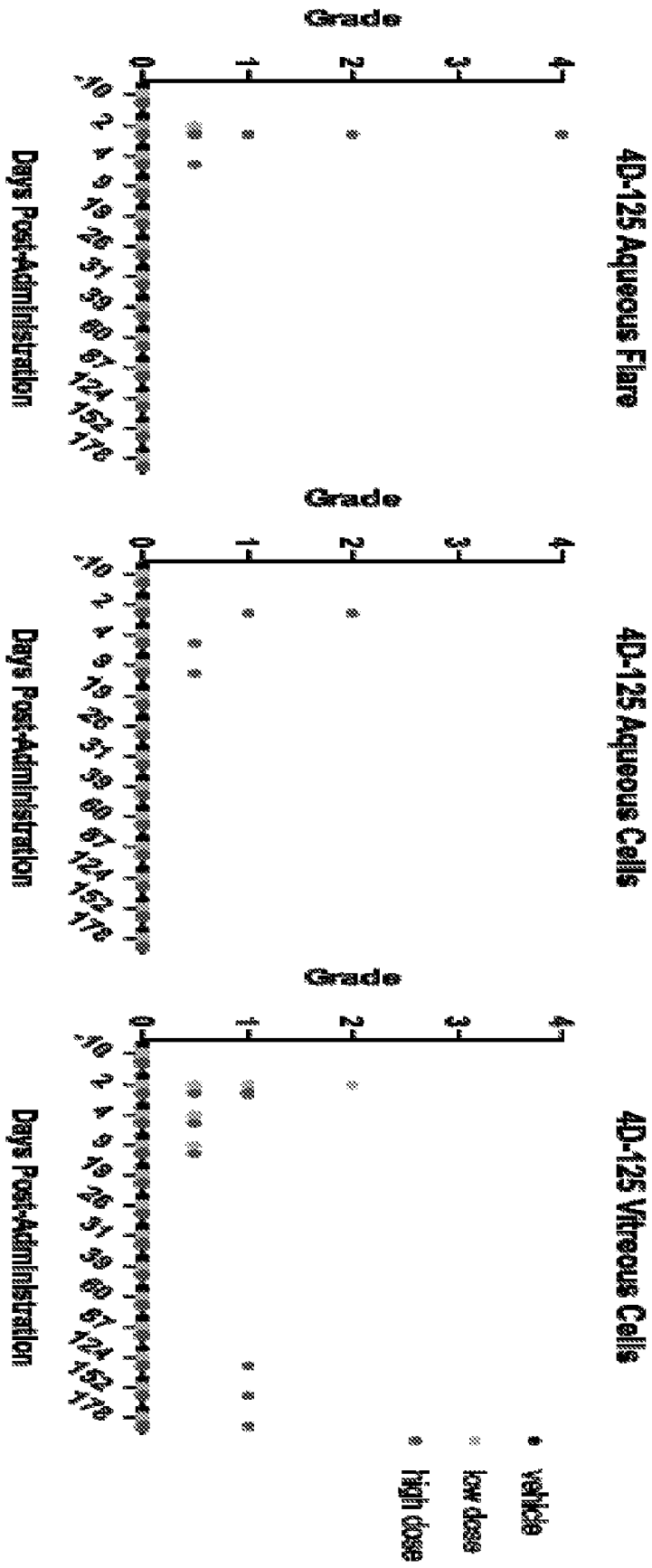


FIGURE 10

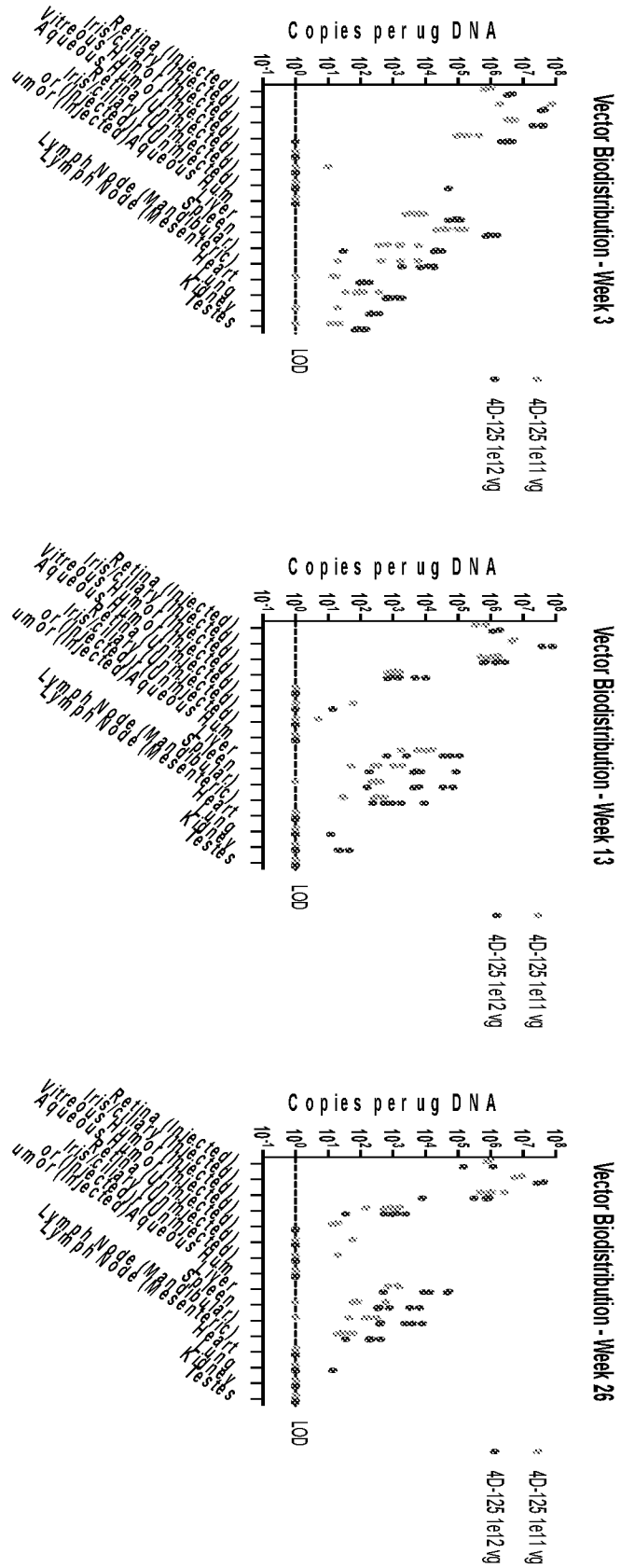


FIGURE 11

