

US012385065B2

### (12) United States Patent

#### Wilson et al.

### (10) Patent No.: US 12,385,065 B2

#### (45) **Date of Patent:** \*Aug. 12, 2025

### (54) GENE THERAPY FOR TREATING CITRULLENEMIA

(71) Applicant: The Trustees of the University of Pennsylvania, Philadelphia, PA (US)

(72) Inventors: James M. Wilson, Philadelphia, PA

(US); Jenny Agnes Sidrane, Phoenixville, PA (US); Lili Wang,

Phoenixville, PA (US)

(73) Assignee: The Trustees of the University of

Pennsylvania, Philadelphia, PA (US)

(\*) Notice: Subject to any disclaimer, the term of this

patent is extended or adjusted under 35

U.S.C. 154(b) by 165 days.

This patent is subject to a terminal dis-

claimer.

(21) Appl. No.: 18/057,572

(22) Filed: Nov. 21, 2022

#### (65) Prior Publication Data

US 2023/0295660 A1 Sep. 21, 2023

#### Related U.S. Application Data

- (63) Continuation of application No. 16/478,971, filed as application No. PCT/US2018/016413 on Feb. 1, 2018, now Pat. No. 11,535,866.
- (60) Provisional application No. 62/453,424, filed on Feb. 1, 2017, provisional application No. 62/469,650, filed on Mar. 10, 2017.
- (51) Int. Cl. A61K 48/00 (2006.01) A61P 43/00 (2006.01) C12N 15/86 (2006.01)

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

(58) Field of Classification Search

CPC .... A61K 48/00; A61K 48/0058; C12N 15/86; C12N 2750/14143; C12N 2830/008; C12N 2830/42; C12Y 603/04005

See application file for complete search history.

#### (56) References Cited

#### U.S. PATENT DOCUMENTS

		Wilson
2003/0198620 A	1* 10/2003	435/235.1 Ozawa A61K 38/44
2004/0009151 A	1* 1/2004	435/235.1 Kay C12Y 304/21022

# 144 A1 12/2015 DeRosa et al.FOREIGN PATENT DOCUMENTS

WO	WO 2003/052051	6/2003
WO	WO 2005/033321	4/2005
WO	WO 2011/126808	10/2011
WO	WO 2013/049493	4/2013
WO	WO 2015/012924	1/2015
WO	WO 2017/015102	1/2017
WO	WO 2017/100676	6/2017

2015/0376144 A1

#### OTHER PUBLICATIONS

Calcedo et al., Worldwide Epidemiology of Neutralizing Antibodies to Adeno-Associated Viruses. J Infect Dis. Feb. 1, 2009;199(3):381-90

Castelhano-Carlos et al., Identification methods in newborn C57BL/6 mice: a developmental and behavioural evaluation. Lab Anim. Apr. 2010;44(2):88-103.

Chandler et al., Liver-directed adeno-associated virus serotype 8 gene transfer rescues a lethal murine model of citrullinemia type 1. Gene Ther. Dec. 2013;20(12):1188-91. Epub Oct. 17, 2013.

Engel et al., Mutations and polymorphisms in the human argininosuccinate synthetase (ASS1) gene. Hum Mutat. Mar. 2009;30(3):300-7. Gao et al., Adeno-associated viruses undergo substantial evolution in primates during natural infections. Proc Natl Acad Sci U S A. May 13, 2003;100(10):6081-6. Epub Apr. 25, 2003.

GenBank Accession: AF513852.1, Adeno-associated virus 8 nonstructural protein and capsid protein genes, complete cds, Sep. 5, 2002

Grimm et al., Titration of AAV-2 particles via a novel capsid ELISA: packaging of genomes can limit production of recombinant AAV-2. Gene Ther, Jul. 1999;6(7):1322-30.

Kelly et al., Splicing of many human genes involves sites embedded within introns. Nucleic Acids Res. May 19, 2015;43(9):4721-32. Epub Apr. 20, 2015.

Kobayashi et al. Messenger RNA coding for argininosuccinate synthetase in citrullinemia. Am J Hum Genet. May 1986;38(5):667-80

Kok et al., Adeno-associated virus-mediated rescue of neonatal lethality in argininosuccinate synthetase-deficient mice. Mol Ther. Oct. 2013;21(10):1823-31. doi: 10.1038/mt.2013.139. Epub Jul. 2, 2013.

#### (Continued)

Primary Examiner — Kevin K Hill (74) Attorney, Agent, or Firm — Howson & Howson LLP; Colleen M. Schaller

#### (57) ABSTRACT

Compositions and regimens useful in treating type I citrullenemia are provided. The compositions include recombinant adeno-associated virus (rAAV) with a transthyretin enhancer and promoter driving expression of a human Argininosuccinate Synthase 1 (ASS1).

#### 20 Claims, 18 Drawing Sheets

Specification includes a Sequence Listing.

#### (56) References Cited

#### OTHER PUBLICATIONS

Lock et al., Absolute determination of single-stranded and self-complementary adeno-associated viral vector genome titers by droplet digital PCR. Hum Gene Ther Methods. Apr. 2014;25(2):115-25. Epub Feb. 14, 2014.

Miyatake et al., Transcriptional targeting of herpes simplex virus for cell-specific replication. J Virol. Jul. 1997;71(7):5124-32.

NCBI Reference Sequence: YP\_077180.1, capsid protein [Adenoassociated virus—8], Mar. 11, 2010.

Saheki et al., Qualitative and quantitative abnormalities of argininosuccinate synthetase in citrullinemia. Clin Chim Acta. Feb. 5, 1981;109(3):325-35.

Sandig et al., HBV-derived promoters direct liver-specific expression of an adenovirally transduced LDL receptor gene. Gene Ther. Nov. 1996;3(11):1002-9.

Shepelev and Fedorov. Advances in the Exon-Intron Database. Brief Bioinform. Jun. 2006;7(2):178-85. Epub Mar. 9, 2006.

Sommer et al., Quantification of adeno-associated virus particles and empty capsids by optical density measurement. Mol Ther. Jan. 2003;7(1):122-8.

Thompson et al., A comprehensive comparison of multiple sequence alignments, Nucleic Acids Res. Jul. 1, 1999;27(13):2682-90.

Wobus et al., Monoclonal antibodies against the adeno-associated virus type 2 (AAV-2) capsid: epitope mapping and identification of capsid domains involved in AAV-2-cell interaction and neutralization of AAV-2 infection. J Virol. Oct. 2000;74(19):9281-93.

Wu et al., Effect of genome size on AAV vector packaging. Mol Ther. Jan. 2010;18(1):80-6. Epub Nov. 10, 2009.

Wu et al., Optimization of self-complementary AAV vectors for liver-directed expression results in sustained correction of hemophilia B at low vector dose. Mol Ther. Feb. 2008;16(2):280-9. Epub Dec. 4 2007

Search Report and Written Opinion in International Patent Application No. PCT/US18/16413, mailed Oct. 4, 2018.

Requirement for Election/Restriction in U.S. Appl. No. 16/478,971, issued Sep. 3, 2021.

Applicant's Response to Requirement for Election/Restriction in U.S. Appl. No. 16/478,971, filed Jan. 3, 2022.

Non-Final Rejection in U.S. Appl. No. 16/478,971, issued Mar. 4, 2022.

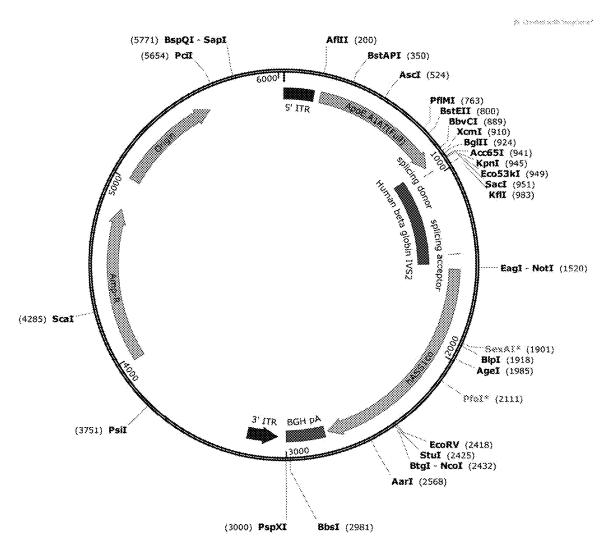
Applicant's Response to Non-Final Rejection in U.S. Appl. No. 16/478,971, filed Jul. 28, 2022.

Notice of Allowance in U.S. Appl. No. 16/478,971, issued Aug. 25, 2022.

\* cited by examiner

FIG. 1

MAP of pAAV8.ApoE.A1AT(full).IVS2.hASS1co.bGH.



pENN.AAV.ApoE.A1AT(full).IVS2.hASS1co.bGH (p4385)

FIG. 2A

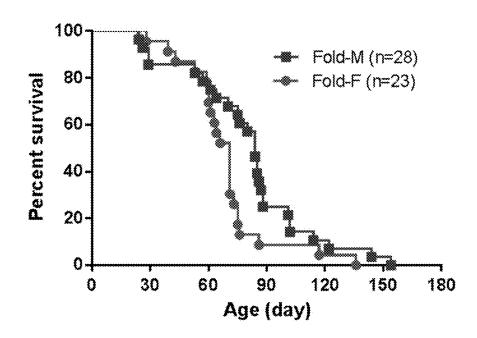


FIG. 2B

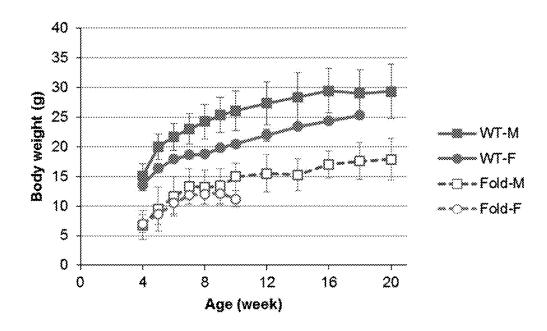


FIG. 2C

# Plasma NH<sub>3</sub>

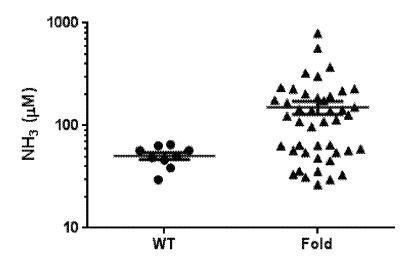


FİG. 2D

### Plasma citrulline

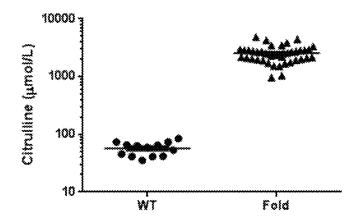


FIG. 2E

### Plasma arginine

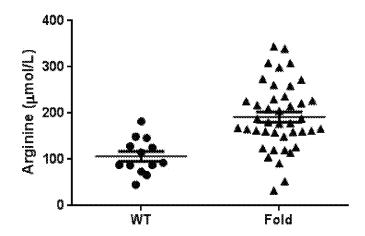


FIG. 2F

### Urine orotic acid

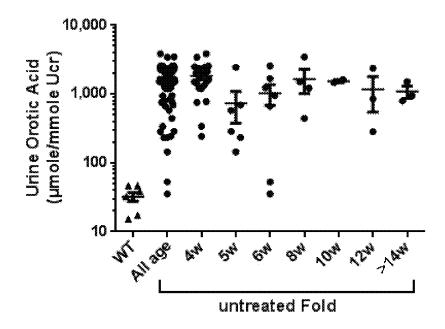


FIG. 3A

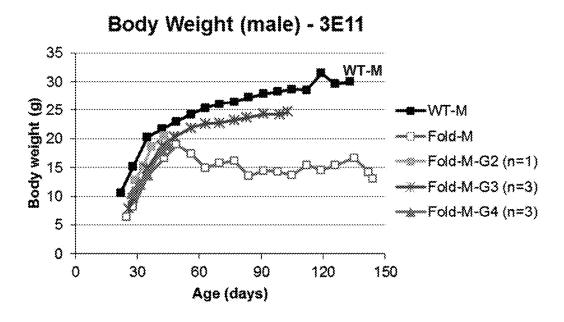


FIG. 3B

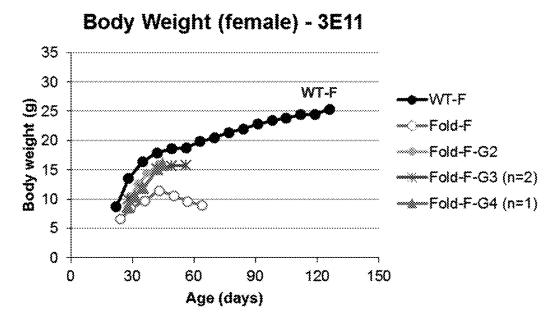


FIG. 3C

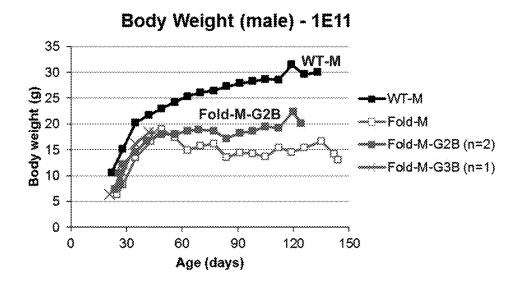


FIG. 3D

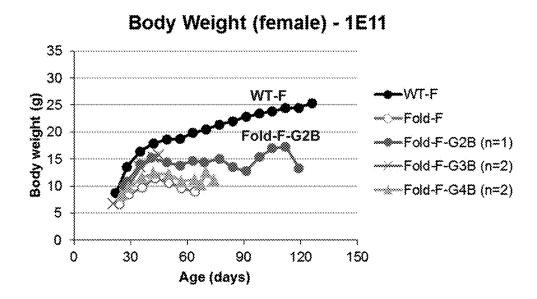


FIG. 3E

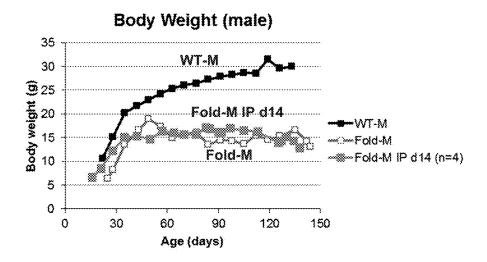
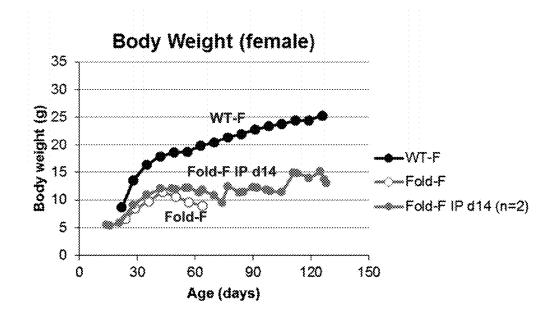
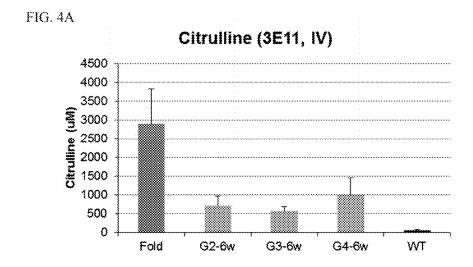


FIG. 3F





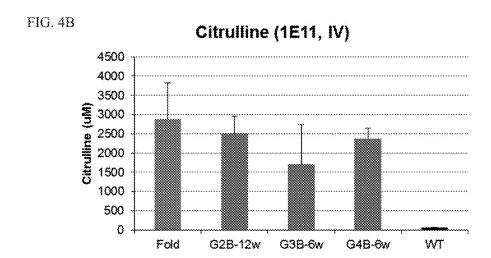


FIG. 5A

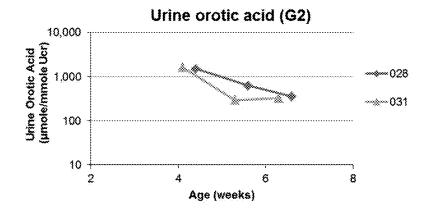


FIG. 5B

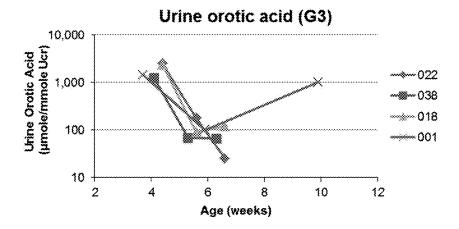


FIG. 5C

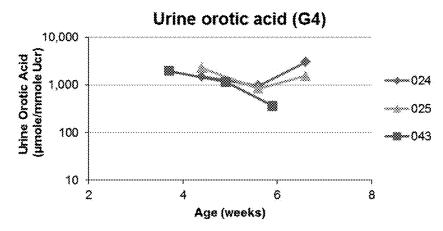


FIG. 5D

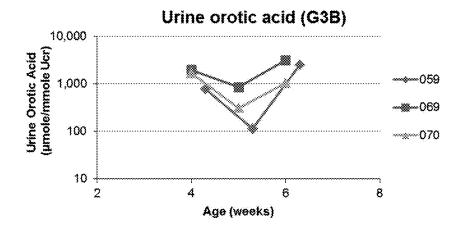


FIG. 6A

### AAV8.LSP.IVS2.hASS1co.bGH

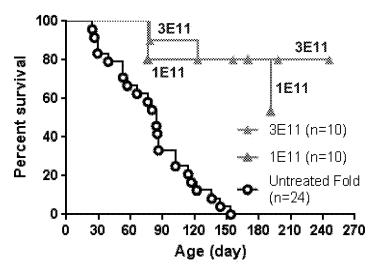


FIG. 6B

### AAV8.LSP.IVS2.hASS1co.bGH

### Body weight (male)

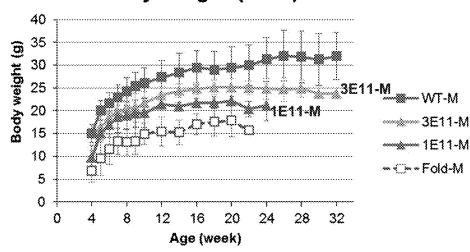


FIG. 6C

### AAV8.LSP.IVS2.hASS1co.bGH

#### **Body weight (female)** 30 Body weight (g) 25 20 WT-F 15 3E11-F 10 1E11-F --Fold-F 5 0 8 12 16 0 4 20 24 28 Age (week)

FIG. 7A

AAV8.ApoE.A1AT(full).IVS2.hASS1co.bGH

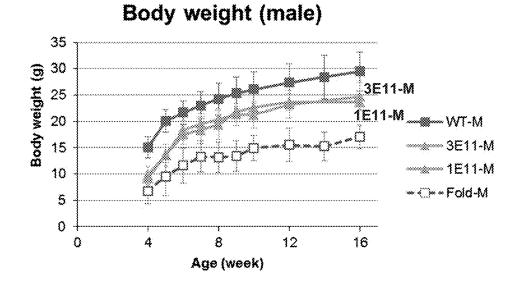
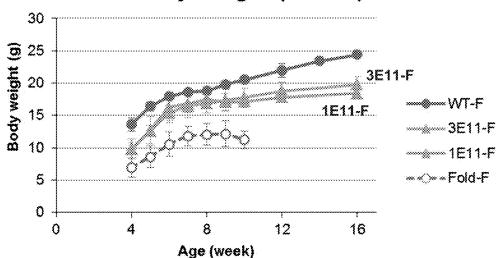
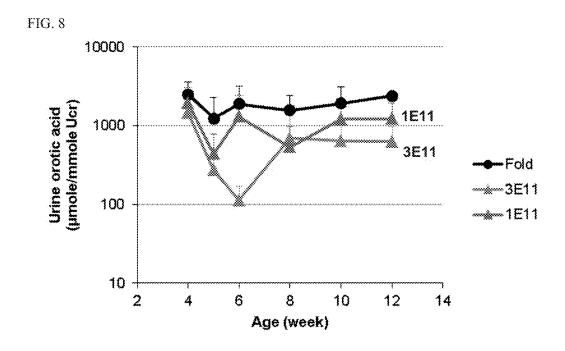


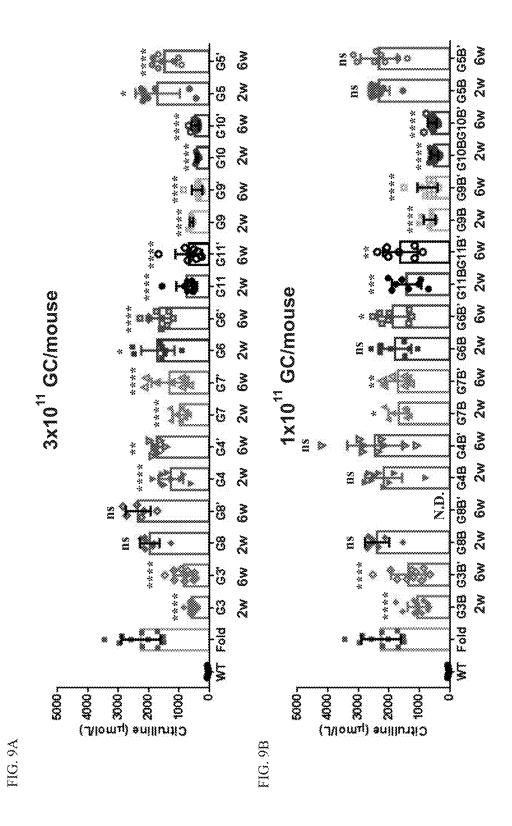
FIG. 7B

### AAV8.ApoE.A1AT(full).IVS2.hASS1co.bGH

# **Body weight (female)**







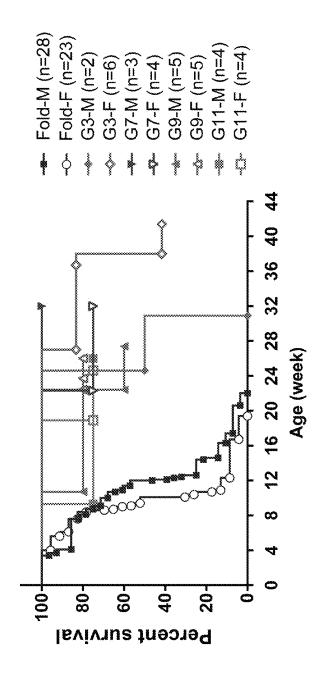


FIG. 10

FIG. 11A

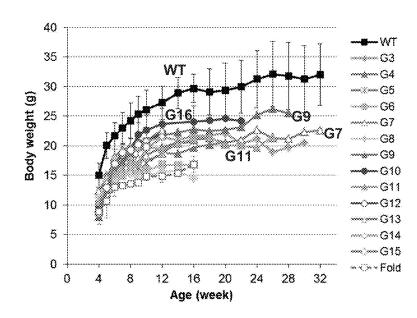
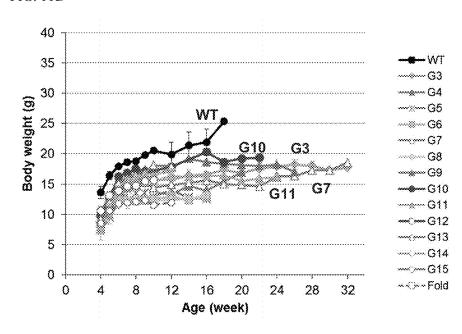
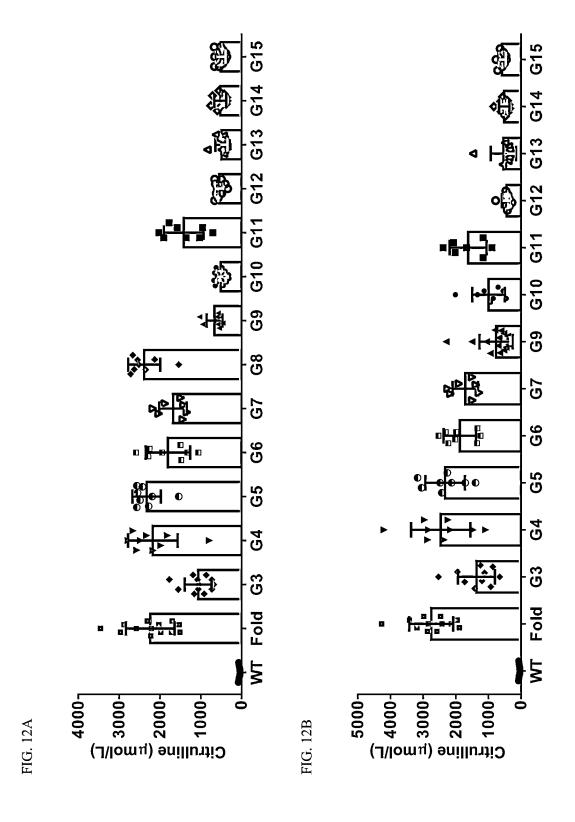


FIG. 11B





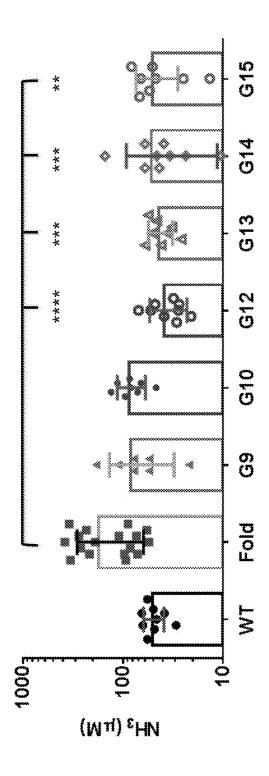


FIG. 13

#### GENE THERAPY FOR TREATING CITRULLENEMIA

#### CROSS-REFERENCE TO RELATED APPLICATIONS

This is a continuation of U.S. patent application Ser. No. 16/478,971, filed Jul. 18, 2019, which is a national stage application under 35 USC 371 of PCT/US2018/061413, filed Jan. 2, 2018, which claims the benefit under 35 USC 119 (e) of U.S. Provisional Patent Application No. 62/453, 424, filed Feb. 1, 2017, and U.S. Provisional Patent Application No. 62/469,650, filed Mar. 10, 2017. Each of these applications is hereby incorporated by reference in its entirety.

#### INCORPORATION-BY-REFERENCE OF MATERIAL SUBMITTED IN ELECTRONIC **FORM**

Applicant hereby incorporates by reference the Sequence 20 Listing material filed in electronic form herewith. This file is labeled "16-7938USC1.xml" (created Feb. 19, 2025 and 139,685 bytes in size).

#### 1. BACKGROUND

The application relates to embodiments useful for a gene therapy for treating type I citrullenemia. Type I citrullenemia is an autosomal recessive disease caused by mutations in argininosuccinate synthase 1 (ASS1) enzyme that catalyzes the synthesis of argininosuccinate from citrulline and aspartate, resulting in citrullinemia and buildup of ammonia.

The clinical spectrum of Type I citrullenemia (CTLN1) ranges from severe neonatal onset form to milder late-onset forms. Owing to its relatively recent addition to the newborn screening panel, patients with type I citrullinemia will be 35 identified early, allowing immediate implementation of treatment. However, despite this early identification of disease and treatment, some patients may progress. The untreated mortality rate in untreated classical CTLN1 is 100%, with most deaths occurring before 17 days of life.

Current treatment approaches for type I citrullenemia include dietary restriction (restriction of protein intake), medications (nitrogen scavenger therapy and carnitine), and arginine supplementation. Liver transplantation is curative for citrullinemia but transplant recipients are required to 45 maintain a constant immune suppression regimen to prevent rejection. Liver directed AAV treatments for type I citrullenemia have been shown. See, e.g., Chandler et al, Liverdirected adeno-associated virus serotype 8 gene transfer rescues a lethal murine model of citrullinemia type 1, Gene 50 Therapy (2013) 20, 1188-1191, which is incorporated herein by reference. However, liver-directed gene therapy did not fully correct the biochemical phenotype of systemic ASS1 deficiency; arginine levels plummeted in treated individuals of a murine model for CTLN1 (fold/fold) due to persistent 55 renal deficiency. See also, Kok et al, Adeno-associated Virus-mediated Rescue of Neonatal Lethality in Argininosuccinate Synthase-deficient Mice, Molecular Therapy vol. 21 no. 10, 1823-1831 Oct. 2013, which is incorporated herein by reference.

What are needed are more effective treatments for type 1 citrullinemia.

#### 2. SUMMARY

The embodiments described herein relate to an AAV gene therapy vector for delivering normal human Argininosucci-

nate Synthase 1 (ASS1) to a subject in need thereof, following intravenous administration of the vector resulting in long-term, perhaps 10 years or more, of clinically meaningful correction of Type I citrullenemia (CTLN1) (also sometimes called citrullinuria or ASS1 deficiency). The subject patient population is patients with moderate to severe Type I citrullenemia, including those with the acute neonatal form (the "classic" form), a milder late-onset form (the "nonclassic" form), or the form in which women have onset of severe symptoms during pregnancy or post-partum. The intended vector dose is, in one embodiment, intended to deliver ASS1 which results in near normal citrulline, glutamine and ammonia plasma levels. However, even nominal reductions in citrulline, glutamine and ammonia levels are desirable, and a desirable endpoint. As reported by Quinonez and Thoene, Pagon R A, Adam M P, Ardinger H H, et al., editors. Seattle (WA): University of Washington, Seattle; 1993-2016 (incorporated herein by reference), elevation of either citrulline or ammonia above acceptable levels (ammonia >100 μmol/L or plasma citrulline >~100 μmol/L) is sufficient evidence to initiate treatment for CTLN1. In another embodiment, a neonatal diagnosis based on genetic testing is sufficient to initiate treatment.

In one aspect, this application provides the use of a replication deficient adeno-associated virus (AAV) to deliver a human Argininosuccinate Synthase 1 (hASS1) gene to liver cells of patients (human subjects) diagnosed with CTLN1. The recombinant AAV vector (rAAV) used for delivering the hASS1 gene ("rAAV.hASS1") should have a tropism for the liver (e.g., a rAAV bearing an AAV8 capsid), and the hASS1 transgene should be controlled by liverspecific expression control elements. In one embodiment, the expression control elements include one or more of the following: an enhancer; a promoter; an intron; a WPRE; and a polyA signal. Such elements are further described herein.

In one embodiment, the ASS1 protein sequence is shown in SEQ ID NO: 1. In one embodiment, the hASS1 coding sequence is shown in SEQ ID NO: 3. The coding sequence for hASS1 is, in one embodiment, codon optimized for expression in humans. Such sequence may share at least 80% identity to the native hASS1 coding sequence (SEQ ID NO: 3). In another embodiment, the hASS1 coding sequence is that shown in SEQ ID NO: 3.

In another aspect, provided herein is an aqueous suspension suitable for administration to a CTLN1 patient which includes the rAAV described herein. In some embodiments, the suspension includes an aqueous suspending liquid and about  $1 \times 10^{12}$  to about  $1 \times 10^{14}$  genome copies (GC) of the rAAV/mL. The suspension is, in one embodiment, suitable for intravenous injection. In other embodiment, the suspension further includes a surfactant, preservative, and/or buffer dissolved in the aqueous suspending liquid.

In another embodiment, provided herein is a method of treating a patient having CTLN1 with an rAAV as described herein. In one embodiment, about  $1\times10^{11}$  to about  $3\times10^{13}$ genome copies (GC) of the rAAV/kg patient body weight are delivered the patient in an aqueous suspension.

#### 3. BRIEF DESCRIPTION OF THE DRAWINGS

60

FIG. 1 is a schematic representation of AAV.hASS1co cis plasmid.

FIG. 2A is a survival curve of ASS1fold/fold mice as described in Example 2. Squares are male mice and circles are female mice.

FIG. **2**B is a line graph of body weights in both genders (M, male; F, female) of ASS1<sup>fold/fold</sup> mice (Fold). Wild-type littermates (WT) were provided as controls.

FIG. 2C is a graph showing elevated plasma NH<sub>3</sub> levels of ASS1<sup>fold/fold</sup> mice. Wild-type (WT) mice were provided as controls. Each circle or triangle indicates one sample. Mean±SEM is also plotted. Mann Whitney test was performed and the p value between indicated groups is shown in the figure.

FIG. 2D is a graph showing elevated plasma citrulline 10 levels of ASS1<sup>fold/fold</sup> mice. Wild-type (WT) mice were provided as controls. Each circle or triangle indicates one sample. Mean±SEM is also plotted. Mann Whitney test was performed and the p value between indicated groups is shown in the figure.

FIG. **2**E is a graph showing elevated plasma arginine levels of ASS1<sup>fold/fold</sup> mice. Wild-type (WT) mice were provided as controls. Each circle or triangle indicates one sample. Mean±SEM is also plotted. Mann Whitney test was performed and the p value between indicated groups is 20 shown in the figure.

FIG. **2**F is a graph showing elevated urine orotic acid levels of ASS1<sup>fold/fold</sup> mice at all age, or at the age of 4 weeks, 5 weeks, 6 weeks, 8 weeks, 10 weeks, 12 weeks and 14 weeks and older. Wild-type (WT) mice were provided as 25 controls. Each dot indicates one sample. Mean±SEM is also plotted.

FIG. **3**A is a line graph of body weights of male ASS1<sup>fold/fold</sup> mice injected intravenously at birth with 3×10<sup>11</sup> GC/pup of AAV8.TBG.PI.hASS1co. WPRE.bGH 30 was acquired and plo (Fold-M-G2, solid gray square, n=1) or AAV8.LSP.IVS2.hASS1co.bGH (Fold-M-G3, asterisk, n=3) or AAV8.TBG.PI.hASS1co.bGH (Fold-M-G4, triangle, n=3). Male ASS1<sup>fold/fold</sup> (Fold-M, open square) and wild-type (WT-M, solid black square) mice without treatment 35 acquired and plotted. FIG. **5**D is a line

FIG. **3**B is a line graph of body weights of female ASS1<sup>fold/fold</sup> mice injected intravenously at birth with 3×10<sup>11</sup> GC/pup of AAV8.TBG.PI.hASS1co. WPRE.bGH (Fold-F-G2, solid grey circle, n=1) or AAV8.LSP.IVS2.hASS1co.bGH (Fold-F-G3, asterisk, n=2) or AAV8.TBG.PI.hASS1co.bGH (Fold-F-G4, triangle, n=1). Female ASS1<sup>fold/fold</sup> (Fold-F, open circle) and wild-type (WT-F, solid black circle) mice without treatment were provided as controls.

FIG. **3**C is a line graph of body weights of male ASS1<sup>fold/fold</sup> mice injected intravenously at birth with 1×10<sup>11</sup> GC/pup of AAV8.TBG.PI.hASS1co.WPRE.bGH (Fold-M-G2, red, n=2) or AAV8.LSP.IVS2.hASS1co.bGH (Fold-M-G3, purple, n=1). Male ASS1<sup>fold/fold</sup> (Fold-M, blue) 50 and wild-type (WT-M, black) mice without treatment were provided as controls.

FIG. **3**D is a line graph of body weights of female ASS1<sup>fold/fold</sup> mice injected intravenously at birth with 1×10<sup>11</sup> GC/pup of AAV8.TBG.PI.hASS1co.WPRE.bGH 55 (Fold-F-G2, red, n=1) or AAV8.LSP.IVS2.hASS1co.bGH (Fold-F-G3, purple, n=2) or AAV8.TBG.PI.hASS1co.bGH (Fold-F-G4, green, n=2). Female ASS1<sup>fold/fold</sup> (Fold-F, blue) and wild-type (WT-F, black) mice without treatment were provided as controls.

FIG. **3**E is a line graph of body weights of male ASS1<sup>fold/fold</sup> mice injected intraperitoneal on postnatal day 14 with 1×10<sup>11</sup> GC/mouse of AAV8.TBG.PI.hASS1co.WPRE.bGH (Fold-M-IP d14, light blue, n=4). Male ASS1<sup>fold/fold</sup> (Fold-M, blue) and wild-type 65 (WT-M, black) mice without treatment were provided as controls.

4

FIG. **3**F is a line graph of body weights of female ASS1<sup>fold/fold</sup> mice injected intraperitoneal on postnatal day 14 with 1×10<sup>11</sup> GC/mouse of AAV8.TBG.PI.hASS1co. WPRE.bGH (Fold-F-IP d14, magenta, n=2). Female ASS1<sup>fold/fold</sup> (Fold-F, blue) and wild-type (WT-F, black) mice without treatment were provided as controls.

FIG. 4A is a bar graph of citrulline levels in the blood of ASS1<sup>fold/fold</sup> mice injected intravenously at birth with 3×10<sup>11</sup> GC/pup of AAV8.TBG.PI.hASS1co.WPRE.bGH (G2-6w) or AAV8.LSP.IVS2.hASS1co.bGH (G3-6w) or AAV8.TBG.PI.hASS1co.bGH (G4-6w). ASS1<sup>fold/fold</sup> (Fold) and wild-type (WT-M) mice without treatment were provided as controls.

FIG. 4B is a bar graph of citrulline levels in the blood of ASS1<sup>fold/fold</sup> mice injected intravenously at birth with 1×10<sup>11</sup> GC/pup of AAV8.TBG.PI.hASS1co.WPRE.bGH (G2B-12w) or AAV8.LSP.IVS2.hASS1co.bGH (G3B-6w) or AAV8.TBG.PI.hASS1co.bGH (G4B-6w). ASS1<sup>fold/fold</sup> (Fold) and wild-type (WT-M) mice without treatment were provided as controls.

FIG. **5**A is a line graph of urine orotic acid levels in ASS1<sup>fold/fold</sup> mice injected intravenously at birth with 3×10<sup>11</sup> GC/pup of AAV8.TBG.PI.hASS1co. WPRE.bGH. Data from mice with identification number 028 and 031 was acquired and plotted.

FIG. **5**B is a line graph of urine orotic acid levels in ASS1<sup>fold/fold</sup> mice injected intravenously at birth with 3×10<sup>11</sup> GC/pup of AAV8.LSP.IVS2.hASS1co.bGH. Data from mice with identification number 022, 038, 018 and 001 was acquired and plotted.

FIG. **5**C is a line graph of urine orotic acid levels in ASS1<sup>fold/fold</sup> mice injected intravenously at birth with 3×10<sup>11</sup> GC/pup of AAV8.TBG.PI.hASS1co.bGH. Data from mice with identification number 024, 025 and 043 was acquired and plotted.

FIG. **5**D is a line graph of urine orotic acid levels in ASS1<sup>fold/fold</sup> mice injected intravenously at birth with 1×10<sup>11</sup> GC/pup of AAV8.LSP.IVS2.hASS1co.bGH. Data from mice with identification number 059, 069 and 070 was acquired and plotted.

FIG. **6**A is a survival curve of ASS1<sup>fold/fold</sup> mice demonstrating liver-directed gene therapy improves the survival of ASS1<sup>fold/fold</sup>. Four-week old ASS1<sup>fold/fold</sup> (both males and females) received a single retro-orbital injection of AAV8.LSP.IVS.hASS1co.bGH vector at the dose of 3×10<sup>11</sup> GC/mouse (n=10) or 1×10<sup>11</sup> GC/mouse (n=10). Survival was monitored. Untreated ASS1<sup>fold/fold</sup> mice were provided as controls (n=24).

FIG. **6**B is a line graph of body weights of male ASS1<sup>fold/fold</sup> mice demonstrating liver-directed gene therapy improves the body weight of ASS1<sup>fold/fold</sup> Four-week old ASS1<sup>fold/fold</sup> received a single retro-orbital injection of AAV8.LSP.IVS.hASS1co.bGH vector at the dose of 3×10<sup>11</sup> GC/mouse or 1×10<sup>11</sup> GC/mouse. Body weights were monitored. Gender-matched untreated ASS1<sup>fold/fold</sup> (Fold) and wild-type (WT) mice were provided as controls.

FIG. **6**C is a line graph of body weights of female ASS1<sup>fold/fold</sup> mice demonstrating liver-directed gene therapy improves the body weight of ASS1<sup>fold/fold</sup> Four-week old ASS1<sup>fold/fold</sup> received a single retro-orbital injection of AAV8.LSP.IVS.hASS1co.bGH vector at the dose of 3×10<sup>11</sup> GC/mouse or 1×10<sup>11</sup> GC/mouse. Body weights were monitored. Gender-matched untreated ASS1<sup>fold/fold</sup> (Fold) and wild-type (WT) mice were provided as controls.

FIG. 7A is a line graph of body weights of male ASS1<sup>fold/fold</sup> mice demonstrating liver-directed gene therapy improves the body weight of ASS1<sup>fold/fold</sup> Four-week old

ASS1fold/fold received a single retro-orbital injection of AAV8.ApoE.A1AT (full).IVS2.hASS1co.bGH vector at the dose of 3×10<sup>11</sup> GC/mouse or 1×10<sup>11</sup> GC/mouse. Body weights were monitored. Gender-matched untreated ASS1 fold/fold (Fold) and wild-type (WT) mice were provided 5

FIG. 7B is a line graph of body weights of female ASS1<sup>fold/fold</sup> mice demonstrating liver-directed gene therapy improves the body weight of ASS1fold/fold Four-week old ASS1<sup>fold/fold</sup> received a single retro-orbital injection of AAV8.ApoE.A1AT (full).IVS2.hASS1co.bGH vector at the dose of 3×10<sup>11</sup> GC/mouse or 1×10<sup>11</sup> GC/mouse. Body weights were monitored. Gender-matched untreated ASS1<sup>foldifold</sup> (Fold) and wild-type (WT) mice were provided 15

FIG. 8 is a line graph showing reduction of urine orotic acid in ASS1fold/fold following vector administration. Four-week old ASS1 fold fold received a single retro-orbital injection of AAV8.LSP.IVS2.hASS1co.bGH vector at the 20 dose of  $3\times10^{11}$  or  $1\times10^{11}$  GC/mouse. ASS1 fold/fold (Fold) mice were provided as controls.

FIGS. 9A to 9B provide graphs showing citrulline levels in ASS1 fold/fold mice at 2 weeks and 6 weeks post AAV8 vector administration. Four-week old ASS1fold/fold 25 received a single retro-orbital injection of the indicated vector at the dose of  $3\times10^{11}$  GC/mouse (FIG. 9A) or  $1\times10^{11}$ GC/mouse (FIG. 9B). Code for each vector is listed in Table 1. Untreated ASS1 fold/fold (Fold) and wild-type (WT) mice were provided as controls. Significant differences compared 30 to untreated fold mice were calculated using one-way ANOVA Dunnett's multiple comparisons test. \* P<0.05, \*\* P<0.01, \*\*\* P<0.001, \*\*\*\* P<0.0001.

single retro-orbital injection of the indicated vector at the dose of 1×10<sup>11</sup> GC/mouse. Code for each vector is listed in Table 1. Untreated ASS1<sup>fold/fold</sup> (Fold) and wild-type (WT) mice were provided as controls.

FIG. 11A is a graph showing body weight in male 40 ASS1 fold/fold mice. Four-week old ASS1 fold/fold received a single retro-orbital injection of the indicated vector at the dose of 1×10<sup>11</sup> GC/mouse. Code for each vector is listed in Table 1. Untreated ASS1 fold/fold (Fold) and wild-type (WT) mice were provided as controls. Liver-directed gene therapy 45 improves the body weight of  ${\rm ASS1}^{fold/fold}$ .

FIG. 11B is a graph showing body weight in female ASS1<sup>fold/fold</sup> mice. Four-week old ASS1<sup>fold/fold</sup> received a single retro-orbital injection of the indicated vector at the dose of 1×10<sup>11</sup> GC/mouse. Code for each vector is listed in 50 Table 1. Untreated ASS1<sup>fold/fold</sup> (Fold) and wild-type (WT) mice were provided as controls. Liver-directed gene therapy improves the body weight of ASS1fold/fold

FIGS. 12A and 12B are graphs showing a reduction in plasma citrulline levels by AAV. Four-week old ASS1 fold/fold 55 (both males and females) received a single retro-orbital injection of AAV8-hASS1 vectors at the dose of  $1\times10^{11}$ GC/mouse. Code for each vector is listed in Table 1. Plasma citrulline levels at 2 weeks (A) and 6 weeks (B) post vector administration. \*\* P<0.01; \*\*\*\* P<0.0001, one-way 60 ANOVA, Kruskal-Wallis test, compared to untreated fold

FIG. 13 is a graph showing a reduction in plasma citrulline levels by AAV. Four-week-old fold mice received a single retro orbital injection of AAV8-hASS1 vectors at the 65 dose of 1.0×10<sup>11</sup> GC. Plasma NH<sub>3</sub> levels at 6 weeks post vector injection are shown here. \*\* P<0.01; \*\*\*

P<0.001; \*\*\*\* P<0.0001, one-way ANOVA, Kruskal-Wallis test, compared to untreated fold mice.

#### 4. DETAILED DESCRIPTION

The embodiments described in the application relate to the use of a replication deficient adeno-associated virus (AAV) to deliver a human Argininosuccinate Synthase 1 (hASS1) gene to liver cells of patients (human subjects) diagnosed with type 1 citrullenemia (CTLN1). The recombinant AAV vector (rAAV) used for delivering the hASS gene ("rAAV.hASS1") should have a tropism for the liver (e.g., an rAAV bearing an AAV8 capsid), and the hASS1 transgene should be controlled by liver-specific expression control elements. In one embodiment, the expression control elements include one or more of the following: an enhancer; a promoter; an intron; a WPRE; and a polyA signal. Such elements are further described herein.

As used herein, "AAV8 capsid" refers to the AAV8 capsid having the amino acid sequence of GenBank, accession: YP\_077180.1, SEQ ID NO: 19, and/or an AAV8 capsid encoded by the nucleic acid sequence of GenBank: AF513852.1, nt 2121-4337, SEQ ID NO: 36, which sequences are incorporated by reference herein. Some variation from this encoded sequence is permitted, which may include sequences having about 99% identity to the referenced amino acid sequence in YP\_077180.1 and WO 2003/ 052051 (which is incorporated herein by reference) (i.e., less than about 1% variation from the referenced sequence). Methods of generating the capsid, coding sequences therefore, and methods for production of rAAV viral vectors have been described. See, e.g., Gao, et al, Proc. Natl. Acad. Sci. U.S.A. 100 (10), 6081-6086 (2003) and US 2015/0315612.

FIG. 10 is a graph showing percent survival in As used herein, the term "NAb titer" a measurement of ASS1<sup>fold/fold</sup> mice. Four-week old ASS1<sup>fold/fold</sup> received a 35 how much neutralizing antibody (e.g., anti-AAV Nab) is As used herein, the term "NAb titer" a measurement of produced which neutralizes the physiologic effect of its targeted epitope (e.g., an AAV). Anti-AAV NAb titers may be measured as described in, e.g., Calcedo, R., et al., Worldwide Epidemiology of Neutralizing Antibodies to Adeno-Associated Viruses. Journal of Infectious Diseases, 2009. 199 (3): p. 381-390, which is incorporated by reference herein.

> The terms "percent (%) identity", "sequence identity", "percent sequence identity", or "percent identical" in the context of amino acid sequences refers to the residues in the two sequences which are the same when aligned for correspondence. Percent identity may be readily determined for amino acid sequences over the full-length of a protein, polypeptide, about 32 amino acids, about 330 amino acids, or a peptide fragment thereof or the corresponding nucleic acid sequence coding sequencers. A suitable amino acid fragment may be at least about 8 amino acids in length, and may be up to about 700 amino acids. Generally, when referring to "identity", "homology", or "similarity" between two different sequences, "identity", "homology" or "similarity" is determined in reference to "aligned" sequences. "Aligned" sequences or "alignments" refer to multiple nucleic acid sequences or protein (amino acids) sequences, often containing corrections for missing or additional bases or amino acids as compared to a reference sequence. Alignments are performed using any of a variety of publicly or commercially available Multiple Sequence Alignment Programs. Sequence alignment programs are available for amino acid sequences, e.g., the "Clustal X", "MAP", "PIMA", "MSA", "BLOCKMAKER", "MEME", and "Match-Box" programs. Generally, any of these programs are used at default settings, although one of skill in the art

can alter these settings as needed. Alternatively, one of skill in the art can utilize another algorithm or computer program which provides at least the level of identity or alignment as that provided by the referenced algorithms and programs. See, e.g., J. D. Thomson et al, Nucl. Acids. Res., "A 5 comprehensive comparison of multiple sequence alignments", 27 (13): 2682-2690 (1999).

As used herein, the term "operably linked" refers to both expression control sequences that are contiguous with the gene of interest and expression control sequences that act in trans or at a distance to control the gene of interest.

A "replication-defective virus" or "viral vector" refers to a synthetic or artificial viral particle in which an expression cassette containing a gene of interest is packaged in a viral 15 capsid or envelope, where any viral genomic sequences also packaged within the viral capsid or envelope are replicationdeficient; i.e., they cannot generate progeny virions but retain the ability to infect target cells. In one embodiment, the genome of the viral vector does not include genes 20 encoding the enzymes required to replicate (the genome can be engineered to be "gutless"-containing only the transgene of interest flanked by the signals required for amplification and packaging of the artificial genome), but these genes may be supplied during production. Therefore, it is deemed safe 25 for use in gene therapy since replication and infection by progeny virions cannot occur except in the presence of the viral enzyme required for replication.

It is to be noted that the term "a" or "an" refers to one or more. As such, the terms "a" (or "an"), "one or more," and 30 "at least one" are used interchangeably herein. The words "comprise", "comprises", and "comprising" are to be interpreted inclusively rather than exclusively. The words "consist", "consisting", and its variants, are to be interpreted exclusively, rather than inclusively. While various embodiments in the specification are presented using "comprising" language, under other circumstances, a related embodiment is also intended to be interpreted and described using "consisting of" or "consisting essentially of" language.

As used herein, the term "about" means a variability of 40 10% from the reference given, unless otherwise specified.

Unless defined otherwise in this specification, technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art and by reference to published texts, which provide one skilled in 45 the art with a general guide to many of the terms used in the present application.

#### 4.1 Gene Therapy Vectors

In one aspect, a recombinant adeno-associated virus (rAAV) vector carrying the human ASS1 gene is provided 50 for use in gene therapy. The rAAV.hASS1 vector should have a tropism for the liver (e.g., a rAAV bearing an AAV8 capsid) and the hASS1 transgene should be controlled by liver-specific expression control elements. In another embodiment, the rAAV.hASS1 vector has a tropism for 55 kidney. The vector is formulated in a buffer/carrier suitable for infusion in human subjects. The buffer/carrier should include a component that prevents the rAAV from sticking to the infusion tubing but does not interfere with the rAAV binding activity in vivo.

#### 4.1.1. The rAAV.hASS Vector

#### 4.1.1.1. The hASS1 Sequence

Citrullinemia type I (CTLN1) (also called "classic citrullinemia") results from deficiency of the enzyme argininosuccinate synthase 1 (ASS1), the third step in the urea cycle, 65 in which citrulline is condensed with aspartate to form arginosuccinic acid.

8

Type I citrullinemia shows kinetically abnormal ASS1 in the liver, kidney, and cultured fibroblasts. In quantitative-type citrullinemia, low ASS1 is found in the liver but not in kidney or cultured skin fibroblasts. Residual enzyme in the liver has normal kinetic properties (Saheki et al., 1981). In a study of mRNA coding for ASS1, Kobayashi et al. (Am J. Hum. Genet., 38:667-80, 1986, which is incorporated herein by reference) found that patients with the quantitative type of citrullinemia had, as demonstrated in previous studies, about 10% of the control value of the enzyme in the liver but a normal level of mRNA. They concluded that in quantitative-type citrullinemia, the decrease in the enzyme protein is due either to increased degradation of the enzyme or to decreased or inhibited translation in the liver.

Although certain pathogenic variants are identified with some phenotypes, the phenotype cannot be predicted in all instances. Severe, classic citrullinemia type I typically results from 22 defined pathogenic variants (Engel et al, Human Mutation, 2009 March; 30 (3): 300-7, which is incorporated herein by reference). The pathogenic variant in exon 15, p.Gly390Arg, remains the most prevalent associated with the classic phenotype. Mild (i.e., late-onset) citrullinemia type I is associated with 12 pathogenic variants.

One goal of therapies described herein would provide functional ASS1 enzyme resulting in citrulline, glutamine, and/or ammonia levels less than 100 µmol/L. In another embodiment, any reduction in citrulline, glutamine, and/or ammonia levels is desirable. Other suitable clinical outcomes may include reduction in the use of scavenger, less restrictive diet or no need for liver transplant.

In one embodiment, the "subject" or "patient" is a mammalian subject having CTLN1 as described above. It is intended that a patient having CTLN1 of any severity is the intended subject. In addition, it is intended that a patient having any mutation in their native ASS1 gene is the intended subject.

In one embodiment, the hASS1 gene encodes the hASS1 protein shown in SEQ ID NO: 1. Thus, in one embodiment, the hASS1 transgene can include, but is not limited to, the sequence provided by SEQ ID NO:2 or SEQ ID NO: 3 which are provided in the attached Sequence Listing, which is incorporated by reference herein. SEQ ID NO: 3 provides the cDNA for native human ASS1. SEQ ID NO: 2 provides an engineered cDNA for human ASS1, which has been codon optimized for expression in humans (sometimes referred to herein as hASS1co). It is to be understood that reference to hASS1 herein may, in some embodiments, refer to the hASS1 native or codon optimized sequence. Alternatively or additionally, web-based or commercially available computer programs, as well as service based companies may be used to back translate the amino acid sequences to nucleic acid coding sequences, including both RNA and/or cDNA. See, e.g., backtranseq by EMBOSS,/www.ebi.ac.uk/Tools/ st/; Gene Infinity (www.geneinfinity.org/sms-/sms\_backtranslation.html); ExPasy (www.expasy.org/tools/). It is intended that all nucleic acids encoding the described hASS1 polypeptide sequences are encompassed, including nucleic acid sequences which have been optimized for expression in the desired target subject (e.g., by codon 60 optimization).

In one embodiment, the nucleic acid sequence encoding hASS1 shares at least 95% identity with the native hASS1 coding sequence of SEQ ID NO: 3. In another embodiment, the nucleic acid sequence encoding hASS1 shares at least 90, 85, 80, 75, 70, or 65% identity with the native hASS1 coding sequence of SEQ ID NO: 3. In one embodiment, the nucleic acid sequence encoding hASS1 shares about 84%

identity with the native hASS1 coding sequence of SEQ ID NO: 3. In one embodiment, the nucleic acid sequence encoding hASS1 is SEQ ID NO: 2.

In one embodiment, the hASS1 coding sequence is codon optimized for expression in the desirable subject species, 5 e.g., humans. Codon-optimized coding regions can be designed by various different methods. This optimization may be performed using methods which are available online (e.g., GeneArt,), published methods, or a company which provides codon optimizing services, e.g., as DNA2.0 (Menlo Park, CA). One codon optimizing approach is described, e.g., in International Patent Publication No. WO 2015/012924, which is incorporated by reference herein. See also, e.g., US Patent Publication No. 2014/0032186 and US Patent Publication No. 2006/0136184. Suitably, the entire 15 length of the open reading frame (ORF) for the product is modified. However, in some embodiments, only a fragment of the ORF may be altered (e.g., one or more of the individual immunoglobulin domains). By using one of these methods, one can apply the frequencies to any given poly-20 peptide sequence, and produce a nucleic acid fragment of a codon-optimized coding region which encodes the polypep-

A number of options are available for performing the actual changes to the codons or for synthesizing the codon- 25 optimized coding regions designed as described herein. Such modifications or synthesis can be performed using standard and routine molecular biological manipulations well known to those of ordinary skill in the art. In one approach, a series of complementary oligonucleotide pairs 30 of 80-90 nucleotides each in length and spanning the length of the desired sequence are synthesized by standard methods. These oligonucleotide pairs are synthesized such that upon annealing, they form double stranded fragments of 80-90 base pairs, containing cohesive ends, e.g., each oli-35 gonucleotide in the pair is synthesized to extend 3, 4, 5, 6, 7, 8, 9, 10, or more bases beyond the region that is complementary to the other oligonucleotide in the pair. The single-stranded ends of each pair of oligonucleotides are designed to anneal with the single-stranded end of another 40 pair of oligonucleotides. The oligonucleotide pairs are allowed to anneal, and approximately five to six of these double-stranded fragments are then allowed to anneal together via the cohesive single stranded ends, and then they ligated together and cloned into a standard bacterial cloning 45 vector, for example, a TOPO® vector available from Thermo Fisher Scientific Inc. The construct is then sequenced by standard methods. Several of these constructs consisting of 5 to 6 fragments of 80 to 90 base pair fragments ligated together, i.e., fragments of about 500 base pairs, are 50 prepared, such that the entire desired sequence is represented in a series of plasmid constructs. The inserts of these plasmids are then cut with appropriate restriction enzymes and ligated together to form the final construct. The final construct is then cloned into a standard bacterial cloning 55 vector, and sequenced. Additional methods would be immediately apparent to the skilled artisan. In addition, gene synthesis is readily available commercially.

#### 4.1.1.2. The rAAV Vector

Because ASS1 is natively expressed in the liver, it is 60 desirable to use an AAV which shows tropism for liver. In one embodiment, the AAV supplying the capsid is AAV8. In another embodiment, the AAV supplying the capsid is AAVrh.10. In yet another embodiment, the AAV supplying the capsid is a Clade E AAV. Such AAV include rh.2; rh.10; 65 rh. 25; bb.1, bb.2, pi.1, pi.2, pi.3, rh.38, rh.40, rh.43, rh.49, rh.50, rh.51, rh.52, rh.53, rh.57, rh.58, rh.61, rh.64, hu.6, hu.

10

17, hu.37, hu.39, hu.40, hu.41, hu.42, hu.66, and hu.67. This clade further includes modified rh. 2; modified rh. 58; and modified rh.64. See, WO 2005/033321, which is incorporated herein by reference. However, any of a number of rAAV vectors with liver tropism can be used. In another embodiment, the rAAV vector has a tropism for kidney.

In a specific embodiment described in the Examples, infra, the gene therapy vector is an AAV8 vector expressing an hASS1 transgene under control of a thyroxine binding globulin (TBG) promoter referred AAV8.TBG.PI.hASS1co. WPRE.bGH. In another embodiment, the WPRE is removed. In another embodiment, the gene therapy vector is an AAV8 vector expressing an hASS1 transgene under control of a A1AT promoter, with an ApoE1 enhancer referred to as AAV8.ApoE.A1AT (full) .IVS2.hASS1co.bGH. The external AAV vector component is a serotype 8, T=1 icosahedral capsid consisting of 60 copies of three AAV viral proteins, VP1, VP2, and VP3, at a ratio of 1:1:10. The capsid contains a single-stranded DNA rAAV vector genome.

In one embodiment, the rAAV.hASS1 genome contains an hASS1 transgene flanked by two AAV inverted terminal repeats (ITRs). In one embodiment, the hASS1 transgene includes one or more of an enhancer, promoter, an intron, a Woodchuck Hepatitis Virus (WHP) Posttranscriptional Regulatory Element (WPRE), an hASS1 coding sequence, and polyadenylation (polyA) signal. These control sequences are "operably linked" to the hASS1 gene sequences. The expression cassette containing these sequences may be engineered onto a plasmid which is used for production of a viral vector.

The ITRs are the genetic elements responsible for the replication and packaging of the genome during vector production and are the only viral cis elements required to generate rAAV. The minimal sequences required to package the expression cassette into an AAV viral particle are the AAV 5' and 3' ITRs, which may be of the same AAV origin as the capsid, or which of a different AAV origin (to produce an AAV pseudotype). In one embodiment, the ITR sequences from AAV2, or the deleted version thereof  $(\Delta ITR)$ , are used. However, ITRs from other AAV sources may be selected. Where the source of the ITRs is from AAV2 and the AAV capsid is from another AAV source, the resulting vector may be termed pseudotyped. Typically, an expression cassette for an AAV vector comprises an AAV 5' ITR, the hASS1 coding sequences and any regulatory sequences, and an AAV 3' ITR. However, other configurations of these elements may be suitable. A shortened version of the 5' ITR, termed  $\Delta$ ITR, has been described in which the D-sequence and terminal resolution site (trs) are deleted. In other embodiments, the full-length AAV 5' and 3' ITRs are used. In one embodiment, the 5' ITR is that shown in SEQ ID NO: 16. In one embodiment, the 3' ITR is that shown in SEQ ID NO: 17.

Exemplary production plasmids to generate rAAVs are shown in SEQ ID NOs: 22 to 35. In one embodiment, provided herein is the plasmid of SEQ ID NO: 22. In another embodiment, provided herein is the plasmid of SEQ ID NO: 23. In another embodiment, provided herein is the plasmid of SEQ ID NO: 24. In another embodiment, provided herein is the plasmid of SEQ ID NO: 25. In another embodiment, provided herein is the plasmid of SEQ ID NO: 26. In another embodiment, provided herein is the plasmid of SEQ ID NO: 27. In another embodiment, provided herein is the plasmid of SEQ ID NO: 28. In another embodiment, provided herein is the plasmid of SEQ ID NO: 29. In another embodiment, provided herein is the plasmid of SEQ ID NO: 30. In another

embodiment, provided herein is the plasmid of SEQ ID NO: 31. In another embodiment, provided herein is the plasmid of SEQ ID NO: 32. In another embodiment, provided herein is the plasmid of SEQ ID NO: 33. In another embodiment, provided herein is the plasmid of SEQ ID NO: 34. In another 5 embodiment, provided herein is the plasmid of SEQ ID NO:

Expression of the hASS1 coding sequence is driven from a liver-specific promoter. An illustrative plasmid and vector described herein uses the thyroxine binding globulin (TBG) 10 promoter (SEQ ID NO: 9), or a modified version thereof. One modified version of the TBG promoter is a shortened version, termed TBG-S1. A modified thyroxine binding globulin (TBG-S1) promoter sequence is shown in SEQ ID NO: 8. Alternatively, other liver-specific promoters may be 15 used such as the transthyretin promoter (TTR) (SEQ ID NO: 11). Another suitable promoter is the alpha 1 anti-trypsin (A1AT), or a modified version thereof (which sequence is shown in SEQ ID NO: 10). In one embodiment, the promoter is an A1AT promoter combined with an ApoE 20 enhancer, sometimes referred to as ApoE.A1AT (full). In one embodiment, the sequence is shown in SEQ ID NO: 20. Another suitable promoter is the Liver specific promoter LSP (TH-binding globulin promoter/alpha1-microglobulin/ bikunin enhancer) (SEQ ID NO: 21). Other suitable pro- 25 moters include human albumin (Miyatake et al., J. Virol., 71:5124 32 (1997)), humAlb; and hepatitis B virus core promoter, (Sandig et al., Gene Ther., 3:1002-9 (1996). See, e.g., The Liver Specific Gene Promoter Database, Cold Spring Harbor, rulai.schl.edu/LSPD, which is incorporated 30 by reference. Although less desired, other promoters, such as viral promoters, constitutive promoters, regulatable promoters [see, e.g., WO 2011/126808 and WO 2013/04943], or a promoter responsive to physiologic cues may be used may be utilized in the vectors described herein.

In one embodiment, the expression control sequences include one or more enhancer. In one embodiment, the En34 enhancer is included (34 bp core enhancer from the human apolipoprotein hepatic control region), which is shown in enhancer sequence from transthyretin) is included. Such sequence is shown in SEQ ID NO: 5. See, Wu et al, Molecular Therapy, 16 (2): 280-289, February 2008, which is incorporated herein by reference. In yet another embodiment, the α1-microglogulin/bikunin precursor enhancer is 45 included. In yet another embodiment, the ABPS (shortened version of the 100 bp distal enhancer from the al-microglogulin/bikunin precursor [ABP] to 42 bp) enhancer is included. Such sequence is shown in SEQ ID NO: 6. In yet another embodiment, the ApoE enhancer is included. Such 50 sequence is shown in SEQ ID NO: 7. In another embodiment, more than one enhancer is present. Such combination may include more than one copy of any of the enhancers described herein, and/or more than one type of enhancer.

In addition to a promoter, an expression cassette and/or a 55 vector may contain other appropriate transcription initiation, termination, enhancer sequences, and efficient RNA processing signals. Such sequences include splicing and polyadenylation (polyA) signals; regulatory elements that enhance expression (i.e., WPRE (SEQ ID NO: 15); sequences that 60 stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (i.e., Kozak consensus sequence); sequences that enhance protein stability; and when desired, sequences that enhance secretion of the encoded product. In one embodiment, a polyadenylation (polyA) signal is 65 included to mediate termination of hASS1 mRNA transcripts. Examples of other suitable polyA sequences include,

12 e.g., bovine growth hormone (SEQ ID NO: 12), SV40, rabbit beta globin, and TK polyA, amongst others.

In one embodiment, the regulatory sequences are selected such that the total rAAV vector genome is about 2.0 to about 5.5 kilobases in size. In one embodiment, the regulatory sequences are selected such that the total rAAV vector genome is about 2.1, 2.3, 2.8, 3.1, 3.2, 3.3 or 4.0 kb in size. In one embodiment, it is desirable that the rAAV vector genome approximate the size of the native AAV genome. Thus, in one embodiment, the regulatory sequences are selected such that the total rAAV vector genome is about 4.7 kb in size. In another embodiment, the total rAAV vector genome is less about 5.2 kb in size. The size of the vector genome may be manipulated based on the size of the regulatory sequences including the promoter, enhancer, intron, poly A, etc. See, Wu et al, Mol Ther, January 2010 18 (1): 80-6, which is incorporated herein by reference.

Thus, in one embodiment, an intron is included in the vector. Suitable introns include the human beta globin IVS2 (SEO ID NO: 13). See, Kelly et al, Nucleic Acids Research, 43 (9): 4721-32 (2015), which is incorporated herein by reference. Another suitable promoter includes the Promega chimeric intron (SEQ ID NO: 14). See, Almond, B. and Schenborn, E. T. A Comparison of pCI-neo Vector and pcDNA4/HisMax Vector. [Internet] 2000, which is incorporated herein by reference. Available from: www.promega-.com/resources/pubhub/enotes/a-comparison-of-pcineovector-and-pcdna4hismax-vector/). Another suitable intron includes the hFIX intron (SEQ ID NO: 18). Various introns suitable herein are known in the art and include, without limitation, those found at bpg.utoledo.edu/~afedorov/lab/ eid.html, which is incorporated herein by reference. See also, Shepelev V., Fedorov A. Advances in the Exon-Intron Database. Briefings in Bioinformatics 2006, 7:178-185, 35 which is incorporated herein by reference.

In one embodiment, the rAAV vector genome comprises a sequence selected from nt 1 to nt 3216 of SEQ ID NO: 22, nt 1 to nt 2331 of SEQ ID NO: 23, nt 1 to nt 3261 of SEQ ID NO: 24, nt 1 to nt 3325 of SEQ ID NO: 25, nt 1 to nt 2777 SEQ ID NO: 4. In another embodiment, the EnTTR (100 bp 40 of SEQ ID NO: 26, nt 1 to nt 2777 of SEQ ID NO: 27, nt 1 to nt 3216 of SEQ ID NO: 28, nt 1 to nt 3066 of SEQ ID NO: 29, nt 1 to nt 2083 of SEQ ID NO: 30, nt 1 to nt 2121 of SEQ ID NO: 31, nt 1 to nt 3221 of SEQ ID NO: 32, nt 1 to nt 4040 of SEQ ID NO: 33, nt 1 to nt 2798 of SEQ ID NO: 34, or nt 1 to nt 3066 of SEQ ID NO: 35.

#### 4.1.2. Compositions

In one embodiment, the rAAV.hASS1 virus is provided in a pharmaceutical composition which comprises an aqueous carrier, excipient, diluent or buffer. In one embodiment, the buffer is PBS. In a specific embodiment, the rAAV.hASS1 formulation is a suspension containing an effective amount of rAAV.hASS1 vector suspended in an aqueous solution containing 0.001% Pluronic F-68 in TMN200 (200 mM sodium chloride, 1 mM magnesium chloride, 20 mM Tris, pH 8.0). However, various suitable solutions are known including those which include one or more of: buffering saline, a surfactant, and a physiologically compatible salt or mixture of salts adjusted to an ionic strength equivalent to about 100 mM sodium chloride (NaCl) to about 250 mM sodium chloride, or a physiologically compatible salt adjusted to an equivalent ionic concentration.

For example, a suspension as provided herein may contain both NaCl and KCl. The pH may be in the range of 6.5 to 8.5, or 7 to 8.5, or 7.5 to 8. A suitable surfactant, or combination of surfactants, may be selected from among Poloxamers, i.e., nonionic triblock copolymers composed of a central hydrophobic chain of polyoxypropylene (poly

(propylene oxide)) flanked by two hydrophilic chains of polyoxyethylene (poly(ethylene oxide)), SOLUTOL® HS 15 (Macrogol-15 Hydroxystearate), LABRASOL® (Polyoxy capryllic glyceride), polyoxy 10 oleyl ether, TWEEN® (polyoxyethylene sorbitan fatty acid esters), ethanol and 5 polyethylene glycol. In one embodiment, the formulation contains a poloxamer. These copolymers are commonly named with the letter "P" (for poloxamer) followed by three digits: the first two digits x 100 give the approximate molecular mass of the polyoxypropylene core, and the last digit x 10 gives the percentage polyoxyethylene content. In one embodiment Poloxamer 188 is selected. The surfactant may be present in an amount up to about 0.0005% to about 0.001% of the suspension. In another embodiment, the vector is suspended in an aqueous solution containing 180 mM sodium chloride, 10 mM sodium phosphate, 0.001% Poloxamer 188, pH 7.3.

In one embodiment, the formulation is suitable for use in human subjects and is administered intravenously. In one embodiment, the formulation is delivered via a peripheral 20 vein by bolus injection. In one embodiment, the formulation is delivered via a peripheral vein by infusion over about 10 minutes (+5 minutes). In one embodiment, the formulation is delivered via a peripheral vein by infusion over about 20 minutes (+5 minutes). In one embodiment, the formulation 25 is delivered via a peripheral vein by infusion over about 30 minutes (+5 minutes). In one embodiment, the formulation is delivered via a peripheral vein by infusion over about 60 minutes (+5 minutes). In one embodiment, the formulation is delivered via a peripheral vein by infusion over about 90 30 minutes (+10 minutes). However, this time may be adjusted as needed or desired. Any suitable method or route can be used to administer an AAV-containing composition as described herein, and optionally, to co-administer other active drugs or therapies in conjunction with the AAV- 35 mediated delivery of hASS1 described herein. Routes of administration include, for example, systemic, oral, inhalation, intranasal, intratracheal, intraarterial, intraocular, intravenous, intramuscular, subcutaneous, intradermal, and other parental routes of administration.

In one embodiment, the formulation may contain, e.g., about  $1.0 \times 10^{11}$  genome copies per kilogram of patient body weight (GC/kg) to about  $1 \times 10^{14}$  GC/kg, about  $5 \times 10^{11}$  genome copies per kilogram of patient body weight (GC/kg) to about  $3 \times 10^{13}$  GC/kg, or about  $1 \times 10^{12}$  to about  $1 \times 10^{14}$  45 GC/kg, as measured by oqPCR or digital droplet PCR (ddPCR) as described in, e.g., M. Lock et al, Hum Gene Ther Methods. 2014 April; 25 (2): 115-25. doi: 10.1089/hgtb.2013.131. Epub 2014 Feb. 14, which is incorporated herein by reference. In one embodiment, the rAAV.hASS 50 formulation is a suspension containing at least  $1 \times 10^{13}$  genome copies (GC)/mL, or greater, as measured by oqPCR or digital droplet PCR (ddPCR) as described in, e.g., M. Lock et al, supra.

In order to ensure that empty capsids are removed from 55 the dose of AAV.hASS1 that is administered to patients, empty capsids are separated from vector particles during the vector purification process, e.g., using the method discussed herein. In one embodiment, the vector particles containing packaged genomes are purified from empty capsids using 60 the process described in U.S. Patent Appln No. 62/322,098, filed on Apr. 13, 2016, and entitled "Scalable Purification Method for AAV8", which is incorporated by reference herein. Briefly, a two-step purification scheme is described which selectively captures and isolates the genome-containing rAAV vector particles from the clarified, concentrated supernatant of a rAAV production cell culture. The process

14

utilizes an affinity capture method performed at a high salt concentration followed by an anion exchange resin method performed at high pH to provide rAAV vector particles which are substantially free of rAAV intermediates. Similar purification methods can be used for vectors having other capsids.

While any conventional manufacturing process can be utilized, the process described herein (and in U.S. Patent Appln No. 62/322,098) yields vector preparations wherein between 50 and 70% of the particles have a vector genome, i.e., 50 to 70% full particles. Thus for an exemplary dose of  $1.6\times10^{12}$  GC/kg, and the total particle dose will be between  $2.3\times10^{12}$  and  $3\times10^{12}$  particles. In another embodiment, the proposed dose is one half log higher, or  $5\times10^{12}$  GC/kg, and the total particle dose will be between  $7.6\times10^{12}$  and  $1.1\times10^{13}$  particles. In one embodiment, the formulation is be characterized by an rAAV stock having a ratio of "empty" to "full" of 1 or less, preferably less than 0.75, more preferably, 0.5, preferably less than 0.3.

A stock or preparation of rAAV8 particles (packaged genomes) is "substantially free" of AAV empty capsids (and other intermediates) when the rAAV8 particles in the stock are at least about 75% to about 100%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least 99% of the rAAV8 in the stock and "empty capsids" are less than about 1%, less than about 5%, less than about 10%, less than about 15% of the rAAV8 in the stock or preparation.

Generally, methods for assaying for empty capsids and AAV vector particles with packaged genomes have been known in the art. See, e.g., Grimm et al., Gene Therapy (1999) 6:1322-1330; Sommer et al., Molec. Ther. (2003) 7:122-128. To test for denatured capsid, the methods include subjecting the treated AAV stock to SDS-polyacrylamide gel electrophoresis, consisting of any gel capable of separating the three capsid proteins, for example, a gradient gel containing 3-8% Tris-acetate in the buffer, then running the gel until sample material is separated, and blotting the gel onto nylon or nitrocellulose membranes, preferably nylon. Anti-40 AAV capsid antibodies are then used as the primary antibodies that bind to denatured capsid proteins, preferably an anti-AAV capsid monoclonal antibody, most preferably the B1 anti-AAV-2 monoclonal antibody (Wobus et al., J. Virol. (2000) 74:9281-9293). A secondary antibody is then used, one that binds to the primary antibody and contains a means for detecting binding with the primary antibody, more preferably an anti-IgG antibody containing a detection molecule covalently bound to it, most preferably a sheep anti-mouse IgG antibody covalently linked to horseradish peroxidase. A method for detecting binding is used to semi-quantitatively determine binding between the primary and secondary antibodies, preferably a detection method capable of detecting radioactive isotope emissions, electromagnetic radiation, or colorimetric changes, most preferably a chemiluminescence detection kit. For example, for SDS-PAGE, samples from column fractions can be taken and heated in SDS-PAGE loading buffer containing reducing agent (e.g., DTT), and capsid proteins were resolved on pre-cast gradient polyacylamide gels (e.g., Novex). Silver staining may be performed using SilverXpress (Invitrogen, CA) according to the manufacturer's instructions. In one embodiment, the concentration of AAV vector genomes (vg) in column fractions can be measured by quantitative real time PCR (Q-PCR). Samples are diluted and digested with DNase I (or another suitable nuclease) to remove exogenous DNA. After inactivation of the nuclease, the samples are further diluted and amplified using primers and a TaqMan<sup>TM</sup> fluorogenic probe specific

for the DNA sequence between the primers. The number of cycles required to reach a defined level of fluorescence (threshold cycle, Ct) is measured for each sample on an Applied Biosystems Prism 7700 Sequence Detection System. Plasmid DNA containing identical sequences to that contained in the AAV vector is employed to generate a standard curve in the Q-PCR reaction. The cycle threshold (Ct) values obtained from the samples are used to determine vector genome titer by normalizing it to the Ct value of the plasmid standard curve. End-point assays based on the digital PCR can also be used.

In one aspect, an optimized q-PCR method is provided herein which utilizes a broad spectrum serine protease, e.g., proteinase K (such as is commercially available from Qiagen). More particularly, the optimized qPCR genome titer assay is similar to a standard assay, except that after the DNase I digestion, samples are diluted with proteinase K buffer and treated with proteinase K followed by heat inactivation. Suitably samples are diluted with proteinase K 20 buffer in an amount equal to the sample size. The proteinase K buffer may be concentrated to 2 fold or higher. Typically, proteinase K treatment is about 0.2 mg/mL, but may be varied from 0.1 mg/mL to about 1 mg/mL. The treatment step is generally conducted at about 55° C. for about 15 25 minutes, but may be performed at a lower temperature (e.g., about 37° C. to about 50° C.) over a longer time period (e.g., about 20 minutes to about 30 minutes), or a higher temperature (e.g., up to about 60° C.) for a shorter time period (e.g., about 5 to 10 minutes). Similarly, heat inactivation is generally at about 95° C. for about 15 minutes, but the temperature may be lowered (e.g., about 70 to about 90° C.) and the time extended (e.g., about 20 minutes to about 30 minutes). Samples are then diluted (e.g., 1000 fold) and subjected to TaqMan analysis as described in the standard assav.

Additionally, or alternatively, droplet digital PCR (ddPCR) may be used. For example, methods for determining single-stranded and self-complementary AAV vector 40 genome titers by ddPCR have been described. See, e.g., M. Lock et al, Hu Gene Therapy Methods, Hum Gene Ther Methods. 2014 April; 25 (2): 115-25. doi: 10.1089/hgtb.2013.131. Epub 2014 Feb. 14.

#### 4.2 Patient Population

As discussed above, a subject having CTLN1 of any severity is the intended recipient of the compositions and methods described herein.

Subjects may be permitted to continue their standard of care treatment(s) (e.g., protein restricted diet, and/or medications (including nitrogen scavenger therapy and carnitine)) prior to and concurrently with the gene therapy treatment at the discretion of their caring physician. In the alternative, the physician may prefer to stop standard of care therapies prior to administering the gene therapy treatment 55 and, optionally, resume standard of care treatments as a co-therapy after administration of the gene therapy.

Desirable endpoints of the gene therapy regimen are an increase in ASS activity resulting in citrulline levels below about 100 µmol/L and/or ammonia levels below about 100 µmol/L. In another embodiment, any reduction in citrulline, glutamine, and/or ammonia levels is desirable. Other suitable clinical outcomes may include reduction in the use of scavenger, less restrictive diet or no need for liver transplant. In one embodiment, patients achieve reduced circulating 65 ASS1 levels after treatment with rAAV.hASS1, alone and/or combined with the use of adjunctive treatments.

16

#### 4.3. Dosing & Route of Administration

In one embodiment, the rAAV.hASS1 vector is delivered as a single dose per patient. In one embodiment, the subject is delivered the minimal effective dose (MED) (as determined by preclinical study described in the Examples herein). As used herein, MED refers to the rAAV.hASS1 dose required to achieve ASS1 activity resulting in citrulline levels below about 100  $\mu$ mol/L and/or ammonia levels below about 100  $\mu$ mol/L.

As is conventional, the vector titer is determined on the basis of the DNA content of the vector preparation. In one embodiment, quantitative PCR or optimized quantitative PCR as described in the Examples is used to determine the DNA content of the rAAV.hASS1 vector preparations. In one embodiment, digital droplet PCR as described in the Examples is used to determine the DNA content of the rAAV.hASS1 vector preparations. In one embodiment, the dosage is about  $1\times10^{11}$  genome copies (GC)/kg body weight to about 1×10<sup>13</sup> GC/kg, inclusive of endpoints. In one embodiment, the dosage is 5×10<sup>11</sup> GC/kg. In another embodiment, the dosage is 5×10<sup>12</sup> GC/kg. In specific embodiments, the dose of rAAV.hASS1 administered to a patient is at least  $5\times10^{11}$  GC/kg,  $1\times10^{12}$  GC/kg,  $1.5\times10^{12}$  GC/kg,  $2.0\times10^{12}$  GC/kg,  $2.5\times10^{12}$  GC/kg,  $3.0\times10^{12}$  GC/kg,  $3.5 \times 10^{12}$  GC/kg,  $4.0 \times 10^{12}$  GC/kg,  $4.5 \times 10^{12}$  GC/kg, 5.0 $\times 10^{12}$  GC/kg,  $5.5 \times 10^{12}$  GC/kg,  $6.0 \times 10^{12}$  GC/kg,  $6.5 \times 10^{12}$ GC/kg,  $7.0 \times 10^{12}$  GC/kg, or  $7.5 \times 10^{12}$  GC/kg. Also, the replication-defective virus compositions can be formulated in dosage units to contain an amount of replication-defective virus that is in the range of about 1.0×10° GC to about  $1.0 \times 10^{15}$  GC. As used herein, the term "dosage" can refer to the total dosage delivered to the subject in the course of treatment, or the amount delivered in a single (of multiple) administration.

In some embodiments, rAAV.hASS1 is administered in combination with one or more therapies for the treatment of CTLN1, such as a low protein diet or administration nitrogen scavenger therapy or dialysis.

#### 4.4. Measuring Clinical Objectives

Measurements of efficacy of treatment can be measured by transgene expression and activity as determined by ammonia or citrulline levels and/or ASS1 activity. Further assessment of efficacy can be determined by clinical assessment of dietary citrulline tolerance.

As used herein, the rAAV.hASS1 vector herein "functionally replaces" or "functionally supplements" the patients defective ASS1 with active ASS1 when the patient expresses a sufficient level of ASS1 to achieve ASS1 activity resulting in citrulline and/or ammonia levels less than about 100 µmol/L. In another embodiment, the rAAV.hASS1 vector functionally replaces or functionally supplements the patient's defective ASS1 when partial rescue is provided. This allows for treatment with a combination of scavengers and dietary control. In one embodiment, the treatment provides sufficient rescue such that liver transplantation is not required. In another embodiment, a reduction in the rate of complications such as viral diseases (e.g., flu) is desired.

#### 5. EXAMPLES

The following examples are illustrative only and are not intended to limit the present invention.

#### EXAMPLE 1: AAV Vectors Containing hASS1

An exemplary gene therapy vector AAV8.TBG.PI.hASS1co. WPRE.bGH was constructed by

an AAV8 vector bearing a codon-optimized human ASS1 cDNA (hASS1co) under the control of TBG, a hybrid promoter based on the human thyroid hormone-binding globulin promoter and microglobin/bikunin enhancer. The ASS1 expression cassette was flanked by AAV2 derived 5 inverted terminal repeats (ITRs) and the expression was driven by a hybrid of the TBG enhancer/promoter and the Woodchuck Hepatitis Virus (WHP) posttranscriptional regulatory element (WPRE) as an enhancer. The transgene also included the Promega SV40 misc intron (PI) and a bovine 10 growth hormone polyadenylation signal (bGH). Another Exemplary Gene Therapy Vector

AAV8.ApoE.A1AT (full).IVS2.hASS1co.bGH was constructed by an AAV8 vector bearing a codon-optimized human ASS1 cDNA (hASS1co) under the control of

A1AT promoter and a ApoE enhancer (FIG. 1). The ASS1 expression cassette was flanked by AAV2 derived inverted terminal repeats (ITRs) and the expression was driven by the ApoE.A1AT enhancer/promoter. The transgene also included the human beta globin IVS2 as an intron and a 20 healthy littermates. bovine growth hormone polyadenylation signal (bGH).

The vector AAV8.TBG.PI.hASS1co.bGH was constructed as described above without WPRE as an enhancer.

The AAV8.LSP.IVS2.hASS1co.bGH vector encodes a codon-optimized human ASS1 cDNA (hASS1co) under the <sup>25</sup> control of a liver-specific promoter (LSP), with intervening sequence 2 (IVS2) and a bovine growth hormone polyadenylation signal (bGH).

The vector was prepared using conventional triple transfection techniques in 293 cells as described e.g., by Mizukami, Hiroaki, et al. *A Protocol for AAV vector production and purification*. Diss. Di-vision of Genetic Therapeutics, Center for MolecularMedicine, 1998., which is incorporated herein by reference. All vectors were produced by the Vector Core at the University of Pennsylvania as previously 35 described [Lock, M., et al, Hum Gene Ther, 21:1259-1271 (2010)].

#### EXAMPLE 2: Natural History Study

All animal procedures were performed in accordance with protocols approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Pennsylvania.

Citrullinemia is an autosomal recessive disease caused by mutations in argininosuccinate synthase (ASS1) enzyme 45 that catalyzes the synthesis of argininosuccinate from citrulline and aspartate, results in citrullinemia and buildup of ammonia. ASS1<sup>fold/fold</sup> mouse express deficient argininosuccinate synthase 1 (ASS1) [Harris BS, et al., Follicular dystrophy: a new skin and hair mutation on mouse Chro- 50 mosome 2. MGI Direct Data Submission. 2007]. FOLD allele mice have a T389I substitution in exon 15 leading to an unstable protein structure with normal ASS1 mRNA and protein levels. Homozygotes survive up to 3 weeks or longer, have 5-10% enzyme activity and display clinical and 55 biochemical parameters similar to CTLN1. At 1 week of age fold homozygotes lack hair such that they can be distinguished from their control littermates, and at 2 weeks of age fold homozygotes have wrinkled skin, a sparse coat does grow in. By P14 mice show a 10- to 40-fold increase in the 60 levels of citrulline, and a 1.5- to threefold increase in the plasma levels of many amino acids, including glutamine, cystine, methionine, and lysine and arginine, glutamic acid, leucine, and ornithine levels are decreased.

ASS1<sup>fold/fold</sup> mice thus served as a mouse model for 65 Citrullinemia. Survival curve was generated based on the observation of 28 male and 23 female ASS1<sup>fold/fold</sup> mice

18

(FIG. 2A). The result demonstrated that functionally deficient ASS1 reduced the lifespan of ASS1 fold/fold mice significantly.

As a second parameter for normal development and growth, the body weights of ASS1<sup>fold/fold</sup> mice and healthy littermates after weaning were closely monitored and recorded. The results showed that female ASS1<sup>fold/fold</sup> mice weight remained relatively constant and the body weights of male ASS1<sup>fold/fold</sup> mice reached plateau at about 8 weeks old while both genders of wild-type littermates exhibited a steady growth over the observation period (FIG. 2B). It further demonstrated that functionally deficient ASS1 compromised the development and growth of the mice.

Additional testing was done on the untreated ASS1<sup>fold/fold</sup> mice and healthy littermates, including measurement of plasma ammonia (FIG. 2C), plasma citrulline (FIG. 2D). plasma arginine (FIG. 2E), urine orotic acid (FIG. 2E). Ammonia, citrulline, arginine and orotic acid levels were all elevated in untreated ASS1<sup>fold/fold</sup> mice as compared to healthy littermates.

### EXAMPLE 3: AAV8.hASS1 vectors in the Model of Citrullenemia

To evaluate the efficacy and determine the dose-dependent effects of AAV8.hASS1co vectors, ASS1<sup>fold/fold</sup> mice were injected with 1×10<sup>11</sup> GC/mouse or 3×10<sup>11</sup> GC/mouse of the gene therapy vectors intravenously at birth, as shown in the Table 1 below. Wild-type and heterozygous littermates served as controls.

3×10<sup>11</sup> GC/pup of AAV8.LSP.IVS2.hASS1co.bGH vector successfully increased the rate of weight gain in male ASS1<sup>fold/fold</sup> mice (FIG. 3A). In females, 3×10<sup>11</sup> GC/pup of all three tested vectors rescued the reduction in body weight upon growth (FIG. 3B). Meanwhile, mice injected intravenously at birth with 1×10<sup>11</sup> GC/pup of the vectors demonstrated a slight increase in weights (FIG. 3C and FIG. 3D). Mice which received an intraperitoneal injection of the AAV8.TBG.PI.hASS1co.WPRE.bGH vector at postnatal day 14 did not exhibit any increase in body weight (FIG. 3E and FIG. 3F).

A further experiment was performed to assess the survival of citrullenemia mice treated with the AAV8.hAASco vectors. To assess citrulline accumulation in the blood, concentration of citrulline in ASS1<sup>fold/fold</sup> mice with intravenous injections of AAV8.TBG.PI.hASS1co.WPRE.bGH, AAV8.LSP.IVS2.hASS1co.bGH or AAV8.TBG.PI.hASS1co.bGH or 3×10<sup>11</sup> GC/pup were examined (FIG. 4B and FIG. 4A, respectively). Injection of 1×10<sup>11</sup> GC/pup of the vectors resulted in a minor decrease in citrulline compared to the untreated ASS1<sup>fold/fold</sup> mice (FIG. 4B), while 3×10<sup>11</sup> GC/pup of the vectors successfully brought the citrulline level down (FIG. 4A).

A further study of expression and enzyme activity of ASS1 in the injected ASS1 fold/fold mice is performed. Livers from the tested mice injected with AAV8.hASSco vectors and the healthy littermate controls are collected and lysates are prepared. The mRNA is extracted and the expression of ASS1 is evaluated via RT-PCR. The protein expression of ASS1 is determined by western blot and immunohistochemistry. Experiments are also performed to assess the ASS1 activity in the ASS1 fold/fold mice treated with the vector as well as controls.

An additional experiment was performed, where weight (FIG. **5**A), ALT (FIG. **5**B), and alkaline phosphatase (FIG. **5**C) were measured on ASS1<sup>fold/fold</sup> mice treated with 1×10<sup>11</sup>

of the vector AAV8.TBG.PI.hASS1co. WPRE.bGH. Wild-type and heterozygous littermates serves as controls.

Urine orotic acid levels were measured in ASS1fold/fold mice injected intravenously at birth with 3×10<sup>11</sup> GC/pup of 5 AAV8.TBG.PI.hASS1co. WPRE.bGH (FIG. 6A). AAV8.LSP.IVS2.hASS1co.bGH (FIG. AAV8.TBG.PI.hASS1co.bGH (FIG. 6C), or with  $1\times10^{11}$ GC/pup of AAV8.LSP.IVS2.hASS1co.bGH (FIG. 6D).

#### **EXAMPLE 4: Additional Vectors**

Additional AAV vectors as shown in Table below were produced.  ${\rm ASS1}^{fold/fold}$  mice were injected on postnatal day 0 at  $1 \times 10^{11}$  or  $3 \times 10^{11}$  GC/pup with the noted vectors. Citrulline levels at two and six week post injection were 15 measured (FIGS. 9A and 9B).

FIG. 10 shows a survival curve of male and female mice injected with G3, G7, G9 and G11. All vectors tested provided significant increase in survival of ASS1fold/fold mice. Likewise, all vectors tested provided an increase in 20 body weight in both male (FIG. 11A) and female (FIG. 11B) mice, a decrease in citrulline levels at 2 weeks (FIG. 12A) and 6 weeks (FIG. 12B) post injection, and a decrease in ammonia levels (FIG. 13).

In conclusion, a single injection of AAV8.hASSco vectors 25 resulted in substantial blood citrulline reduction and con20

comitant functional correction when administered intravenously in ASS1-deficient mice.

> EXAMPLE 5: Long-Term Rescue of a Hypomorphic Lethal Murine Model of Citrullinemia Type I by Liver-Directed, AAV8-Mediated Gene Therapy

Citrullinemia type I (CTLN1) is an autosomal, recessive disorder of the urea cycle caused by a deficiency of argininosuccinate synthase 1 (ASS1). The clinical spectrum of CTLN1 ranges from a severe neonatal onset form to a milder form with later onset. Affected patients have persistent elevated plasma citrulline levels and are at risk of lifethreatening elevation of ammonia that can lead to irreversible cognitive impairment, coma, and death. Current treatment for CTLN1 patients, which includes a low protein diet, supplementation of arginine and administration of nitrogen scavengers, is often unable to prevent ongoing hyperammonemic crises. Liver transplantation has shown successful reduction of plasma ammonia and citrulline levels, but donor liver is limiting, the procedure itself carries significant morbidity, and immunosuppressive drugs are necessary for the duration of the graft. Therefore, there is a need for other approaches to therapy for CTLN1.

			-	TABLE 1				
	Sumi	nary of Vecto	rs for liver di	ected therapy	for treatment o	f citrullinem	nia	
					Key elements in vector			
Grp	Vector				Enhancer- Promoter	Intron	Transgene	
G2	AAV8.TBG.PI.hASS1co.WPRE.bGH (p3795)			ABPx2 - TBG	PI	hASS1co		
G3	AAV8.TBG.IVS2.hASS1co.bGH (p4169)			1169)	ABPx2 - TBG	IVS2	hASS1co	
G4	AAV8.TBG.PI.hASS1co.bGH (p4157)			7)	ABPx2 - TBG	PI	hASS1co	
G5	AAV8.EnTTR.TTR.hASS1co.bGH (p4319)			p4319)	EnTTR - TTR	_	hASS1co	
G6	AAV8.E	n34.A1AT.hA	SS1co.bGH (p	4320)	En34 - A1AT	_	hASS1co	
G7	AAV8.T	BG.PI.hASS1	-native.bGH (1	54339)	ABPx2 -	PI	hASS1-	
					TBG		native	
G8	AAV8.T	BG.hFIXintro	n.hASS1co.bC	H (p4382)	ABPx2 -	hFIX	hASS1co	
				• ,	TBG	intron		
G9	AAV8.ApoE.A1AT.IVS2.hASS1co.bGH (p4383)			GH (p4383)	ApoE - A1AT	IVS2	hASS1co	
<b>G</b> 10	AAV8.ApoE.A1AT(full).IVS2.hASS1co.bGH (p4385)			1co.bGH	ApoE - A1AT (full)	IVS2	hASS1co	
G11	AAV8.En34.A1AT.PI.hASS1co.bGH (p4340)			(p4340)	En34 - A1AT	PΙ	hASS1co	
G12	AAV8.ApoE.A1AT(full).IVS2.hASS1-				ApoE -	IVS2	hASS1-	
	native.bGH (p4456)				A1AT (full)		native	
G13			S2.hASS1-nat	ive.bGH	ApoE -	IVS2	hASS1-	
	(p4457)	r			A1AT		native	
G14	AAV8.TBG.IVS2.hASS1-native.bGH (p4458)		H (p4458)	ABPx2 -	IVS2	hASS1-		
01.	161.0.160.1.02.11.6501 (hadve.0011 (p++30)			- (p ·)	TBG		native	
G15	AAV8.E	nTTR.TTR.IV	S2.hASS1-nat	ive.bGH	EnTTR -	IVS2	hASS1-	
	(p4459)				TTR		native	
					lline (µM)		ulline (%	
				Mean	± STD	redu	ction)	
		Size (bp)	Yield (GC)/	3E11	1E11	3E11	1E11	
	Grp	(ITR-ITR)	Cellstack	GC/mouse	GC/mouse		GC/mouse	
	Gip	(11K-11K)	Censtack	GC/Illouse	GC/Illouse	GC/IIIouse	GC/IIIouse	
	G2	3325	1.21E+14	714 ± 161 (n = 5)****	N.D.	68.4%	N.D.	
	G3	3261	8.60E+13	564 ± 122 (n = 13)****	$1065 \pm 329$ (n = 12)	75	52	
	G4	2777	1.26E+14	$1272 \pm 406$ (n = 9)	(n - 12) 2176 ± 607 (n = 9)	43	3	

21 TABLE 1-continued

Sumn	nary of Vecto	ors for liver di	rected therapy	for treatment of	citrullinemi	a
G5	2121	9.90E+13	$1707 \pm 732$ (n = 8)	$2328 \pm 348$ (n = 8)	24	-4
G6	2083	6.60E+13	$1702 \pm 549$ (n = 8)	$1808 \pm 542$ (n = 8)	24	19
G7	2777	1.17E+14	$980 \pm 231$ (n = 8)	1681 ± 337	56	25
G8	4040	1.38E+14	1966 ± 319 (n = 8)		12	-7
G9	3066	6.75E+13		664 ± 196	74	70
<b>G</b> 10	3216	1.01E+14	$404 \pm 45$ (n = 9)****	509 ± 118	82	77
G11	2331	1.17E+14	746 ± 348 (n = 8)**		67	37
G12	3216	7.27E+13	N.D.	548 ± 111 (n = 10)****	N.D.	76
G13	3066	1.10E+14	N.D.	494 ± 148 (n = 10)****	N.D.	78
G14	3221	7.12E+13	N.D.	$523 \pm 147$ (n = 9)****	N.D.	77
G15	2798	1.37E+14	N.D.	$513 \pm 119$ (n = 10)****	N.D.	77
		Fold	2239 ± 59	96 (n = 13)		

AAV vector-based gene therapy provides an alternative to 25 vector. Reduction of plasma citrulline levels was chosen as current treatment options as long as the vector delivers sufficient and sustained transgene expression in the liver without substantial toxicity. Several candidate AAV8 vectors for CTLN1 were generated with different liver-specific promoters, introns, and cDNA sequences (native or codonoptimized hASS1 cDNA). In vivo evaluation of vectors was performed in a murine model of CTLN1 (ASS1fold/fold). Homozygous ASS1 fold/fold (fold) mice carried a hypomorphic mutation and display lethality after weaning. Half of the 35 untreated fold mice perished before the age of 12 weeks old, while a few (5%) lived up to 5 months. In addition to significantly elevated plasma citrulline levels, untreated fold mice had significantly reduced body weight, variable elevated plasma ammonia levels and urine orotic acid levels, 40 and they were not fertile.

Four-week-old fold mice were dosed via retro-orbital or intraperitoneal injection with 3×10<sup>11</sup> GC or 1×10<sup>11</sup> GC of the main criteria to differentiate the performance of different vectors. A lead vector containing the ApoE enhancer-alpha 1 antitrypsin promoter and the beta globulin intervening sequence 2 achieved 77% reduction of citrulline levels two weeks post vector administration at the dose of  $1 \times 10^{11}$  GC. Intron played an important role in expression of ASS1 and vectors carrying the same promoter with other introns, or no intron, showed significantly perturbed efficiency in reducing citrulline levels. Vector with native cDNA sequences performed slightly better than a vector with codon-optimized cDNA sequences. Fold mice treated with the top candidate vectors gained weight, became fertile, and survived more than 9 months (still on-going).

#### Sequence Listing Free Text

The following information is provided for sequences containing free text under numeric identifier <223>.

SEQ ID NO	Amino Acid or Nucleic Acid Sequence	Free text under <223>	Description
1	Amino Acid		human ASS1
2	Nucleic Acid	<223> constructed sequence	engineered cDNA for human ASS1
3	Nucleic Acid		cDNA for native human ASS1
4	Nucleic Acid	<223> constructed sequence	En34 enhancer
5	Nucleic Acid	<223> constructed sequence	EnTTR enhancer
6	Nucleic Acid	<223> constructed sequence	ABPS enhancer
7	Nucleic Acid	<223> constructed sequence	ApoE enhancer
8	Nucleic Acid	<223> constructed sequence	TBG-S1
9	Nucleic Acid	<223> constructed sequence	TBG promoter
10	Nucleic Acid	<223> constructed sequence	A1AT promoter
11	Nucleic Acid	<223> constructed sequence	TTR promoter
12	Nucleic Acid	<223> constructed sequence	bGH polyA
13	Nucleic Acid	<223> constructed sequence	human beta globin IVS2
14	Nucleic Acid	<223> constructed sequence	Promega ® chimeric intron
15	Nucleic Acid	<223> constructed sequence	WPRE
16	Nucleic Acid	<223> constructed sequence	5' ITR
17	Nucleic Acid	<223> constructed sequence	3' ITR
18	Nucleic Acid	<223> constructed sequence	hFIX intron
19	Amino Acid	<223> AAV8	AAV8 capsid

#### -continued

SEQ ID NO	Amino Acid or Nucleic Acid Sequence	Free text under <223>	Description
20	Nucleic Acid	<223> constructed sequence	A1AT promoter combined with an ApoE enhancer
21	Nucleic Acid	<223> constructed sequence	liver-specific promoter LSP
22	Nucleic Acid	<223> constructed sequence	ApoE.A1AT(full).IVS2.hASS1co.bGH cis plasmid
23	Nucleic Acid	<223> constructed sequence	En34.A1AT.PI.hASS1co.bGH cis plasmid
24	Nucleic Acid	<223> constructed sequence	TBG.IVS2.hASS1co.bGH cis plasmid
25	Nucleic Acid	<223> constructed sequence	TBG.PI.hASS1co.WPRE.bGH cis plasmid
26	Nucleic Acid	<223> constructed sequence	TBG.PI.hASS1co.bGH cis plasmid
27	Nucleic Acid	<223> constructed sequence	TBG.PI.hASS1-native.bGH cis plasmid
28	Nucleic Acid	<223> constructed sequence	ApoE.A1AT(full).IVS2.hASS1-native.bGH cis plasmid
29	Nucleic Acid	<223> constructed sequence	ApoE.A1AT.IVS2.hASS1co.bGH cis plasmid
30	Nucleic Acid	<223> constructed sequence	En34.A1AT.hASS1co.bGH cis plasmid
31	Nucleic Acid	<223> constructed sequence	EnTTR.TTR.hASS1co.bGH cis plasmid
32	Nucleic Acid	<223> constructed sequence	TBG.IVS2.hASS1-native.bGH cis plasmid
33	Nucleic Acid	<223> constructed sequence	TBG.hFIXintron.hASS1co.bGH cis plasmid
34	Nucleic Acid	<223> constructed sequence	EnTTR.TTR.IVS2.hASS1-native.bGH cis plasmid
35	Nucleic Acid	<223> constructed sequence	ApoE.A1AT.IVS2.hASS1-native.bGH cis
36	Nucleic Acid		AAV8 capsid

All publications cited in this specification are incorporated herein by reference in their entireties, as are U.S. Provisional Patent Application No. 62/453,424, filed Feb. 1, 2017, and U.S. provisional Patent Application No. 62/469, 650, filed Mar. 10, 2017. Similarly, the SEQ ID NOs which are referenced herein and which appear in the appended

Sequence Listing are incorporated by reference. While the invention has been described with reference to particular embodiments, it will be appreciated that modifications can be made without departing from the spirit of the invention. Such modifications are intended to fall within the scope of the appended claims.

24

SEQUENCE LISTING

```
Sequence total quantity: 36
SEQ ID NO: 1
                       moltype = AA length = 412
FEATURE
                       Location/Qualifiers
source
                       1..412
                       mol_type = protein
                       organism = Homo sapiens
SEQUENCE: 1
MSSKGSVVLA YSGGLDTSCI LVWLKEQGYD VIAYLANIGQ KEDFEEARKK ALKLGAKKVF
IEDVSREFVE EFIWPAIQSS ALYEDRYLLG TSLARPCIAR KQVEIAQREG AKYVSHGATG
KGNDQVRFEL SCYSLAPQIK VIAPWRMPEF YNRFKGRNDL MEYAKQHGIP IPVTPKNPWS
MDENLMHISY EAGILENPKN QAPPGLYTKT QDPAKAPNTP DILEIEFKKG VPVKVTNVKD
GTTHQTSLEL FMYLNEVAGK HGVGRIDIVE NRFIGMKSRG IYETPAGTIL YHAHLDIEAF
TMDREVRKIK QGLGLKFAEL VYTGFWHSPE CEFVRHCIAK SQERVEGKVQ VSVLKGQVYI
LGRESPLSLY NEELVSMNVQ GDYEPTDATG FININSLRLK EYHRLQSKVT AK
SEQ ID NO: 2
                       moltype = DNA length = 1236
FEATURE
                       Location/Qualifiers
misc_feature
                       1..1236
                       note = constructed sequence
                       1..1236
source
                       mol type = other DNA
                       organism = synthetic construct
SEQUENCE: 2
atqaqcaqca aqqqctctqt qqtqctqqcc tactctqqcq qcctqqacac caqctqtatc
                                                                   120
ctegtgtggc tgaaagaaca gggctacgac gtgategeet acetggeeaa categgeeag
aaaqaqqact tcqaqqaaqc ccqqaaqaaq qccctqaaqc tqqqcqccaa qaaqqtqttc
                                                                   180
                                                                   240
ategaggaeg tgteeegega gttegtggaa gagtteatet ggeeegeeat eeagageage
gccctgtacg aggacagata cctgctgggc accagcctgg ccagaccctg tatcgcccgg
                                                                   300
aaacaggtgg aaatcgccca gegegaggge gecaaataeg tgteteaegg egecaeegge
                                                                   360
aagggcaacg accaggtgcg ctttgagctg agctgctact ccctggcccc ccagatcaaa
                                                                   420
gtgatcgccc cttggcggat gcccgagttc tacaaccggt tcaagggccg gaacgacctg
                                                                   480
atggaatacg ccaagcagca cggcatcccc atccccgtga cccccaagaa cccttggagc
                                                                    540
atggacgaga acctgatgca catcagctac gaggccggca tcctggaaaa ccccaagaat
caggeeette ceggeetgta cacaaagace caggaeeetg ecaaggeeee caacaceeee
                                                                    660
gacattetgg aaategagtt caagaaagge gtgeeegtga aagtgaceaa egtgaaggae
                                                                   720
ggcaccaccc accagacctc cctggaactg ttcatgtacc tgaacgaggt ggccggcaag
                                                                    780
cacggcgtgg gcagaatcga catcgtggaa aacagattca tcggcatgaa gtcccggggc
                                                                    840
atctacgaga caccagoogg caccatootg taccaogooc acctggatat cgaggootto
```

```
accatggacc gggaagtgcg gaagatcaag cagggcctgg gcctgaagtt cgccgagctg
gtgtacacag gcttttggca cagccccgag tgcgagtttg tgcggcactg cattgccaag
                                                                     1020
agccaggaac gggtggaagg caaggtgcag gtgtccgtgc tgaagggcca ggtgtacatt
                                                                     1080
ctgggcagag agagcccct gagcctgtac aacgaggaac tggtgtctat gaacgtgcag
                                                                     1140
ggcgactacg agcccaccga cgccaccggc ttcatcaaca tcaacagcct gagactgaaa
                                                                     1200
gagtaccacc ggctgcagtc caaagtgacc gccaag
                                                                     1236
SEQ ID NO: 3
                       moltype = DNA length = 1236
FEATURE
                       Location/Qualifiers
source
                       1..1236
                       mol_type = genomic DNA
                       organism = Homo sapiens
SEQUENCE: 3
atgtccagca aaggctccgt ggttctggcc tacagtggcg gcctggacac ctcgtgcatc
ctcgtgtggc tgaaggaaca aggctatgac gtcattgcct atctggccaa cattggccag
aaggaagact tcgaggaagc caggaagaag gcactgaagc ttggggccaa aaaggtgttc
attgaggatg tcagcaggga gtttgtggag gagttcatct ggccggccat ccagtccagc
geactgtatg aggacegeta ceteetggge acetetettg ecaggeeetg categeeege
aaacaagtgg aaatcgccca gcgggagggg gccaagtatg tgtcccacgg cgccacagga
aaggggaacg atcaggtccg gtttgagctc agctgctact cactggcccc ccagataaag
gtcattgctc cctggaggat gcctgaattc tacaaccggt tcaagggccg caatgacctg
atggagtacg caaagcaaca cgggattccc atcccggtca ctcccaagaa cccgtggagc
atggatgaga acctcatgca catcagctac gaggctggaa tcctggagaa ccccaagaac
                                                                     600
caagcgcctc caggtctcta cacgaagacc caggacccag ccaaagcccc caacacccct
                                                                     660
gacatteteg agategagtt caaaaaaggg gteeetgtga aggtgaccaa egteaaggat
                                                                     720
ggcaccaccc accagacctc cttggagctc ttcatgtacc tgaacgaagt cgcgggcaag
                                                                     780
catggcgtgg gccgtattga catcgtggag aaccgcttca ttggaatgaa gtcccgaggt
                                                                     840
atctacgaga ccccagcagg caccatcctt taccacgctc atttagacat cgaggccttc
                                                                     900
accatggacc gggaagtgcg caaaatcaaa caaggcctgg gcttgaaatt tgctgagctg
                                                                     960
gtgtataccg gtttctggca cagccctgag tgtgaatttg tccgccactg catcgccaag tcccaggagc gagtggaagg gaaagtgcag gtgtccgtcc tcaagggcca ggtgtacatc
                                                                     1020
                                                                     1080
cteggeeggg agteeceact gtetetetae aatgaggage tggtgageat gaacgtgeag
                                                                     1140
ggtgattatg agccaactga tgccaccggg ttcatcaaca tcaattccct caggctgaag
                                                                     1200
gaatatcatc gtctccagag caaggtcact gccaaa
                                                                     1236
SEQ ID NO: 4
                       moltype = DNA length = 34
FEATURE
                       Location/Qualifiers
misc_feature
                       1..34
                       note = constructed sequence
source
                       1..34
                       mol_type = other DNA
                       organism = synthetic construct
SEOUENCE: 4
tgtttgctgc ttgcaatgtt tgcccatttt aggg
                                                                     34
                       moltype = DNA length = 100
SEO ID NO: 5
FEATURE
                       Location/Qualifiers
misc feature
                       1..100
                       note = constructed sequence
                       1..100
source
                       mol_type = other DNA
                       organism = synthetic construct
SEOUENCE: 5
ctacctcgtg atcgcccggc ccctgttcaa acatgtccta atactctgtc tctgcaaggg
tcatcagtag ttttccatct tactcaacat cctcccagtg
                                                                     100
SEQ ID NO: 6
                       moltype = DNA length = 42
                       Location/Qualifiers
FEATURE
misc feature
                       1..42
                       note = constructed sequence
source
                       1..42
                       mol type = other DNA
                       organism = synthetic construct
SEOUENCE: 6
aggttaattt ttaaactgtt tgctctggtt aataatctca gg
                                                                     42
SEO ID NO: 7
                       moltype = DNA length = 322
FEATURE
                       Location/Qualifiers
misc feature
                       1..322
                       note = constructed sequence
source
                       1..322
                       mol_type = other DNA
                       organism = synthetic construct
aaggeteaga ggeacacagg agtttetggg eteaccetge eccetteeaa ecceteagtt
cccatcctcc agcagctgtt tgtgtgctgc ctctgaagtc cacactgaac aaacttcagc
ctactcatgt ccctaaaatg ggcaaacatt gcaagcagca aacagcaaac acacagccct
ccctgcctgc tgaccttgga gctggggcag aggtcagaga cctctctggg cccatgccac
ctccaacatc cactcgaccc cttggaattt cggtggagag gagcagaggt tgtcctggcg
```

```
tggtttaggt agtgtgagag gg
                                                                   322
                       moltype = DNA length = 176
SEO ID NO: 8
FEATURE
                       Location/Qualifiers
misc feature
                       1..176
                       note = constructed sequence
source
                       1..176
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 8
actcaaagtt caaaccttat cattttttgc tttgttcctc ttggccttgg ttttgtacat
cagetttgaa aataceatee cagggttaat getggggtta atttataaet aagagtgete
tagttttgca atacaggaca tgctataaaa atggaaagat gttgctttct gagaga
SEQ ID NO: 9
                       moltype = DNA length = 477
                       Location/Qualifiers
FEATURE
                       1..477
misc feature
                       note = constructed sequence
                       1..477
source
                       mol type = other DNA
                       organism = synthetic construct
SEQUENCE: 9
agggctggaa gctacctttg acatcatttc ctctgcgaat gcatgtataa tttctacaga
acctattaga aaggatcacc cagcctctgc ttttgtacaa ctttccctta aaaaactgcc
                                                                   120
aattccactg ctgtttggcc caatagtgag aactttttcc tgctgcctct tggtgctttt
geetatggee cetattetge etgetgaaga cactettgee ageatggaet taaacceete
                                                                   240
cagetetgae aatcetettt etetttegtt ttacatgaag ggtetggeag ccaaageaat
                                                                   300
cactcaaagt tcaaacctta tcattttttg ctttgttcct cttggccttg gttttgtaca
                                                                   360
tcagctttga aaataccatc ccagggttaa tgctggggtt aatttataac taagagtgct
                                                                   420
ctagttttgc aatacaggac atgctataaa aatggaaaga tgttgctttc tgagaga
                       moltype = DNA length = 218
SEO ID NO: 10
                       Location/Qualifiers
FEATURE
                       1..218
misc_feature
                       note = constructed sequence
                       1..218
source
                       mol_type = other DNA
                       organism = synthetic construct
SEOUENCE: 10
tggacacagg acgctgtggt ttctgagcca gggggcgact cagatcccag ccagtggact
tageceetgt ttgeteetee gataactggg gtgaeettgg ttaatattea eeageageet
                                                                   120
ccccgttgc ccctctggat ccactgctta aatacggacg aggacagggc cctgtctcct
                                                                   180
cagetteagg caccaccact gaeetgggae agtgaata
                                                                   218
SEQ ID NO: 11
                       moltype = DNA length = 190
                       Location/Qualifiers
FEATURE
misc_feature
                       1..190
                       note = constructed sequence
source
                       1..190
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 11
atttcataga acgaatgttc cgatgctcta atctctctag acaaggttca tatttgtatg
ggttacttat tctctctttg ttgactaagt caataatcag aatcagcagg tttgcagtca
gattggcagg gataagcagc ctagctcagg agaagtgagt ataaaagccc caggctggga
gcagccatca
SEQ ID NO: 12
                       moltype = DNA length = 215
FEATURE
                       Location/Qualifiers
misc_feature
                       1..215
                       note = constructed sequence
                       1..215
source
                       mol type = other DNA
                       organism = synthetic construct
SEQUENCE: 12
gcctcgactg tgccttctag ttgccagcca tctgttgttt gcccctcccc cgtgccttcc
ttgaccctgg aaggtgccac tcccactgtc ctttcctaat aaaatgagga aattgcatcg 120
cattgtctga gtaggtgtca ttctattctg gggggtgggg tggggcagga cagcaagggg
                                                                   180
gaggattggg aagacaatag caggcatgct gggga
SEO ID NO: 13
                       moltype = DNA length = 572
FEATURE
                       Location/Qualifiers
misc_feature
                       1..572
                       note = constructed sequence
                       1..572
source
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 13
agettacttq tqqtaccqaq ctcqqatcct qaqaacttca qqqtqaqtct atqqqaccct 60
```

```
tgatgttttc tttccccttc ttttctatgg ttaagttcat gtcataggaa ggggagaagt
aacagggtac acatattgac caaatcaggg taattttgca tttgtaattt taaaaaatgc
                                                                   180
tttcttcttt taatactt ttttgtttat cttatttcta atactttccc taatctcttt
                                                                   240
ctttcagggc aataatgata caatgtatca tgcctctttg caccattcta aagaataaca
                                                                   300
gtgataattt ctgggttaag gcaatagcaa tatttctgca tataaatatt tctgcatata
                                                                   360
aattgtaact gatgtaagag gtttcatatt gctaatagca gctacaatcc agctaccatt
                                                                   420
ctgcttttat tttatggttg ggataaggct ggattattct gagtccaagc taggcccttt
                                                                   480
tgctaatcat gttcatacct cttatcttcc tcccacagct cctgggcaac gtgctggtct
                                                                   540
gtgtgctggc ccatcacttt ggcaaagaat tg
SEQ ID NO: 14
                       moltype = DNA length = 133
FEATURE
                       Location/Qualifiers
misc_feature
                       1..133
                       note = constructed sequence
                       1..133
source
                       mol type = other DNA
                       organism = synthetic construct
SEOUENCE: 14
gtaagtatca aggttacaag acaggtttaa ggagaccaat agaaactggg cttgtcgaga
cagagaagac tottgogttt otgataggca cotattggto ttactgacat coactttgco
                                                                   120
tttctctcca cag
SEQ ID NO: 15
                       moltype = DNA length = 542
FEATURE
                       Location/Qualifiers
misc feature
                       1..542
                       note = constructed sequence
                       1..542
source
                       mol_type = other DNA
                       organism = synthetic construct
SEOUENCE: 15
aatcaacctc tggattacaa aatttgtgaa agattgactg gtattcttaa ctatgttgct
cettttacgc tatgtggata egetgettta atgeetttgt ateatgetat tgetteeegt
                                                                   120
atggctttca ttttctcctc cttgtataaa tcctggttgc tgtctcttta tgaggagttg
tggcccgttg tcaggcaacg tggcgtggtg tgcactgtgt ttgctgacgc aacccccact
                                                                   240
ggttggggca ttgccaccac ctgtcagctc ctttccggga ctttcgcttt ccccctccct
                                                                   300
attgccacgg cggaactcat cgccgcctgc cttgcccgct gctggacagg ggctcggctg
                                                                   360
ttgggcactg acaattccgt ggtgttgtcg gggaaatcat cgtcctttcc ttggctgctc
                                                                   420
gcctgtgttg ccacctggat tctgcgcggg acgtccttct gctacgtccc ttcggccctc
                                                                   480
aatccagegg acctteette eegeggeetg etgeeggete tgeggeetet teegegtett
                                                                   540
cg
                                                                   542
SEQ ID NO: 16
                       moltype = DNA length = 168
FEATURE
                       Location/Qualifiers
misc_feature
                       1..168
                       note = constructed sequence
                       1..168
source
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 16
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact
                                                                   120
aggggttcct tgtagttaat gattaacccg ccatgctact tatctacg
                                                                   168
SEQ ID NO: 17
                       moltype = DNA length = 168
FEATURE
                       Location/Qualifiers
misc feature
                       1..168
                       note = constructed sequence
                       1..168
source
                       mol_type = other DNA
                       organism = synthetic construct
SEOUENCE: 17
cgtagataag tagcatggcg ggttaatcat taactacaag gaacccctag tgatggagtt
ggccactccc tctctgcgcg ctcgctcgct cactgaggcc gggcgaccaa aggtcgcccg
                                                                   120
acgcccgggc tttgcccggg cggcctcagt gagcgagcga gcgcgcag
                                                                   168
SEQ ID NO: 18
                       moltype = DNA length = 1421
FEATURE
                       Location/Qualifiers
misc_feature
                       1..1421
                       note = constructed sequence
                       1..1421
source
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 18
gtttgtttcc ttttttaaaa tacattgagt atgcttgcct tttagatata gaaatatctg 60
atgctgtctt cttcactaaa ttttgattac atgatttgac agcaatattg aagagtctaa
                                                                   120
cagecageae geaggttggt aagtaetggt tetttgttag etaggtttte ttettettea
tttttaaaac taaatagatc gacaatgctt atgatgcatt tatgtttaat aaacactgtt
cagttcatga tttggtcatg taattcctgt tagaaaacat tcatctcctt ggtttaaaaa
aattaaaagt gggaaaacaa agaaatagca gaatatagtg aaaaaaaata accacattat 360
```

```
ttttgtttgg acttaccact ttgaaatcaa aatgggaaac aaaagcacaa acaatggcct
tatttacaca aaaagtctga ttttaagata tatgacattt caaggtttca gaagtatgta
atqaqqtqtq tctctaattt tttaaattat atatcttcaa tttaaaqttt taqttaaaac
                                                                   540
ataaagatta acctttcatt agcaagctgt tagttatcac caaagctttt catggattag
                                                                   600
gaaaaaatca ttttgtctct atgtcaaaca tcttggagtt gatatttggg gaaacacaat
                                                                   660
actcagttga gttccctagg ggagaaaagc aagcttaaga attgacataa agagtaggaa
                                                                   720
gttagctaat gcaacatata tcactttgtt ttttcacaac tacagtgact ttatgtattt
                                                                   780
cccagaggaa ggcatacagg gaagaaatta tcccatttgg acaaacagca tgttctcaca
                                                                   840
ggaagcattt atcacactta cttgtcaact ttctagaatc aaatctagta gctgacagta
                                                                   900
ccaggatcag gggtgccaac cctaagcacc cccagaaagc tgactggccc tgtggttccc
                                                                   960
actocagaca tgatgtcago tggaccataa ttaggottot gttottcagg agacatttgt
                                                                   1020
tcaaagtcat ttgggcaacc atattctgaa aacagcccag ccagggtgat ggatcacttt
                                                                   1080
gcaaagatcc tcaatgagct attttcaagt gatgacaaag tgtgaagtta accgctcatt
tgagaacttt ctttttcatc caaagtaaat tcaaatatga ttagaaatct gaccttttat
                                                                   1200
tactggaatt ctcttgacta aaagtaaaat tgaattttaa ttcctaaatc tccatgtgta
tacagtactg tgggaacatc acagattttg gctccatgcc ctaaagagaa attggctttc
                                                                   1320
agattatttg gattaaaaac aaagactttc ttaagagatg taaaattttc atgatgtttt
cttttttgct aaaactaaag aattattctt ttacatttca g
SEQ ID NO: 19
                      moltype = AA length = 738
FEATURE
                      Location/Qualifiers
REGION
                       1..738
                      note = AAV8
source
                       1..738
                       mol type = protein
                       organism = unidentified
SEQUENCE: 19
MAADGYLPDW LEDNLSEGIR EWWALKPGAP KPKANOOKOD DGRGLVLPGY KYLGPFNGLD
KGEPVNAADA AALEHDKAYD QQLQAGDNPY LRYNHADAEF QERLQEDTSF GGNLGRAVFQ
AKKRVLEPLG LVEEGAKTAP GKKRPVEPSP QRSPDSSTGI GKKGQQPARK RLNFGQTGDS
                                                                   180
ESVPDPOPLG EPPAAPSGVG PNTMAAGGGA PMADNNEGAD GVGSSSGNWH CDSTWLGDRV
                                                                   240
ITTSTRTWAL PTYNNHLYKQ ISNGTSGGAT NDNTYFGYST PWGYFDFNRF HCHFSPRDWQ
                                                                   300
RLINNNWGFR PKRLSFKLFN IOVKEVTONE GTKTIANNLT STIOVFTDSE YOLPYVLGSA
                                                                   360
HQGCLPPFPA DVFMIPQYGY LTLNNGSQAV GRSSFYCLEY FPSQMLRTGN NFQFTYTFED
                                                                   420
VPFHSSYAHS QSLDRLMNPL IDQYLYYLSR TQTTGGTANT QTLGFSQGGP NTMANQAKNW
                                                                   480
LPGPCYRQQR VSTTTGQNNN SNFAWTAGTK YHLNGRNSLA NPGIAMATHK DDEERFFPSN
                                                                   540
GILIFGKONA ARDNADYSDV MLTSEEEIKT TNPVATEEYG IVADNLOOON TAPOIGTVNS
                                                                   600
QGALPGMVWQ NRDVYLQGPI WAKIPHTDGN FHPSPLMGGF GLKHPPPQIL IKNTPVPADP
                                                                   660
PTTFNQSKLN SFITQYSTGQ VSVEIEWELQ KENSKRWNPE IQYTSNYYKS TSVDFAVNTE
                                                                   720
GVYSEPRPIG TRYLTRNL
                                                                   738
SEO ID NO: 20
                      moltype = DNA length = 735
                      Location/Qualifiers
FEATURE
misc_feature
                      1..735
                      note = constructed sequence
                       1..735
source
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 20
gctcagaggc acacaggagt ttctgggctc accctgcccc cttccaaccc ctcagttccc
atcctccagc agctgtttgt gtgctgcctc tgaagtccac actgaacaaa cttcagccta
                                                                   120
ctcatgtccc taaaatgggc aaacattgca agcagcaaac agcaaacaca cagccctccc
tgcctgctga ccttggagct ggggcagagg tcagagacct ctctgggccc atgccacctc
                                                                   240
caacatccac tcgacccctt ggaatttcgg tggagaggag cagaggttgt cctggcgtgg
tttaggtagt gtgagagggc gcgccgatct tgctaccagt ggaacagcca ctaaggattc
                                                                   360
tgcagtgaga gcagagggcc agctaagtgg tactctccca gagactgtct gactcacgcc
acceceteca cettggacae aggacgetgt ggtttetgag ceaggtacaa tgaeteettt
                                                                   480
cggtaagtgc agtggaagct gtacactgcc caggcaaagc gtccgggcag cgtaggcggg
cgactcagat cccagccagt ggacttagcc cctgtttgct cctccgataa ctggggtgac
cttggttaat attcaccagc agcctccccc gttgcccctc tggatccact gcttaaatac
ggacgaggac agggccctgt ctcctcagct tcaggcacca ccactgacct gggacagtag
atctagctta cttgt
SEQ ID NO: 21
                       moltype = DNA length = 727
FEATURE
                       Location/Qualifiers
                      1..727
misc_feature
                       note = constructed sequence
source
                       1..727
                       mol type = other DNA
                      organism = synthetic construct
SEOUENCE: 21
agttaatttt taaaaagcag tcaaaagtcc aagtgccctt gcgagcattt actctctctg
tttgctctgg ttaataatct caggagcaca aacattcctt actagttcta ggagttaatt
tttaaaaagc agtcaaaagt ccaagtgccc ttgcgagcat ttactctctc tgtttgctct
ggttaataat ctcaggagca caaacattcc ttactagttc tagagcggcc gccagtgtgc
tggaattcgg cttttttagg gctggaagct acctttgaca tcatttcctc tgcgaatgca
tgtataattt ctacagaacc tattagaaag gatcacccag cctctgcttt tgtacaactt
tcccttaaaa aactgccaat cccactgctg tttggcccaa tagtgagaac tttttcctgc
tgcctcttgg tgcttttgcc tatggcccct attctgcctg ctgaagacac tcttgccagc
```

```
atggacttaa acccctccag ctctgacaat cctctttctc ttttgtttta catgaagggt
                                                                     540
ctggcagcca aagcaatcac tcaaagttca aaccttatca ttttttgctt tgttcctctt
                                                                     600
ggccttggtt ttgtacatca gctttgaaaa taccatccca gggttaatgc tggggttaat
                                                                     660
ttataactga gagtgctcta gttctgcaat acaggacatg ctataaaaaat ggaaagatgt
                                                                     720
                                                                     727
SEQ ID NO: 22
                       moltype = DNA length = 6026
FEATURE
                       Location/Qualifiers
misc_feature
                       1..6026
                       note = constructed sequence
                       1..6026
source
                       mol type = other DNA
                       organism = synthetic construct
SEQUENCE: 22
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact
aggggtteet tgtagttaat gattaaceeg ceatgetaet tatetaegta geeatgetet
aggaagateg gaattegeee ttaaggetea gaggeacaca ggagtttetg ggeteaceet
geocettee aacceteag tteccateet ceageagetg tttgtgtget geetetgaag
tocacactga acaaacttca gootactcat gtocotaaaa tgggcaaaca ttgcaagcag
caaacagcaa acacacagcc ctccctgcct gctgaccttg gagctggggc agaggtcaga
gacetetetg ggeceatgee acetecaaca tecactegae ecettggaat tteggtggag
aggagcagag gttqtcctqq cqtqqtttaq qtaqtqtqaq agqqcqccc qatcttqcta
ccagtggaac agccactaag gattctgcag tgagagcaga gggccagcta agtggtactc tcccagagac tgtctgactc acgccaccc ctccaccttg gacacaggac gctgtggttt
                                                                     600
                                                                     660
ctgagccagg tacaatgact cctttcggta agtgcagtgg aagctgtaca ctgcccaggc
                                                                     720
aaagcgtccg ggcagcgtag gcgggcgact cagatcccag ccagtggact tagcccctgt
                                                                     780
ttgctcctcc gataactggg gtgaccttgg ttaatattca ccagcagcct cccccgttgc
                                                                     840
ccctctggat ccactgctta aatacggacg aggacagggc cctgtctcct cagcttcagg
                                                                     900
caccaccact gacctgggac agtagatcta gcttacttgt ggtaccgagc tcggatcctg
                                                                     960
agaacttcag ggtgagtcta tgggaccctt gatgttttct ttccccttct tttctatggt
                                                                     1020
taagtteatg teataggaag gggagaagta acagggtaca catattgace aaatcagggt aattttgcat ttgtaatttt aaaaaatget ttettettt aatataettt tttgtttate
                                                                     1080
                                                                     1140
ttatttctaa tactttccct aatctctttc tttcaqqqca ataatqatac aatqtatcat
                                                                     1200
gcctctttgc accattctaa agaataacag tgataatttc tgggttaagg caatagcaat
                                                                     1260
atttctgcat ataaatattt ctgcatataa attgtaactg atgtaagagg tttcatattg
                                                                     1320
ctaatagcag ctacaatcca gctaccattc tgcttttatt ttatggttgg gataaggctg
                                                                     1380
gattattctg agtccaagct aggccctttt gctaatcatg ttcatacctc ttatcttcct
                                                                     1440
cccacagete etgggeaacg tgetggtetg tgtgetggee cateaetttg gcaaagaatt
                                                                     1500
gatetegagt aactgaagge ggeegeeace atgageagea agggetetgt ggtgetggee
                                                                     1560
tactctggcg gcctggacac cagctgtatc ctcgtgtggc tgaaagaaca gggctacgac
                                                                     1620
gtgatcgcct acctggccaa catcggccag aaagaggact tcgaggaagc ccggaagaag
                                                                     1680
gccctgaagc tgggcgccaa gaaggtgttc atcgaggacg tgtcccgcga gttcgtggaa
                                                                     1740
gagttcatct ggcccgccat ccagagcagc gccctgtacg aggacagata cctgctgggc
                                                                     1800
accagectgg ecagacectg tategeeegg aaacaggtgg aaategeeea gegegaggge
                                                                     1860
gccaaatacg tgtctcacgg cgccaccggc aagggcaacg accaggtgcg ctttgagctg
                                                                     1920
agetgetact ccctggcccc ccagatcaaa gtgatcgccc cttggcggat gcccgagttc
                                                                     1980
tacaaccggt tcaagggccg gaacgacctg atggaatacg ccaagcagca cggcatcccc
                                                                     2040
atccccgtga cccccaagaa cccttggagc atggacgaga acctgatgca catcagctac
                                                                     2100
gaggeeggea teetggaaaa eeccaagaat eaggeeette eeggeetgta eacaaagace
                                                                     2160
caggaccetg ccaaggeeee caacaceeee gacattetgg aaategagtt caagaaagge
                                                                     2220
gtgcccgtga aagtgaccaa cgtgaaggac ggcaccaccc accagacctc cctggaactg
                                                                     2280
ttcatgtacc tgaacgaggt ggccggcaag cacggcgtgg gcagaatcga catcgtggaa
                                                                     2340
aacagattca tcggcatgaa gtcccggggc atctacgaga caccagccgg caccatcctg
taccacgccc acctggatat cgaggccttc accatggacc gggaagtgcg gaagatcaag
                                                                     2460
cagggcctgg gcctgaagtt cgccgagctg gtgtacacag gcttttggca cagccccgag
                                                                     2520
tgcgagtttg tgcggcactg cattgccaag agccaggaac gggtggaagg caaggtgcag
                                                                     2580
gtgtccgtgc tgaagggcca ggtgtacatt ctgggcagag agagccccct gagcctgtac
aacgaggaac tggtgtctat gaacgtgcag ggcgactacg agcccaccga cgccaccggc
ttcatcaaca tcaacagcct gagactgaaa gagtaccacc ggctgcagtc caaagtgacc
gccaagtgat aagcatgcgg atctgcctcg actgtgcctt ctagttgcca gccatctgtt
gtttgcccct cccccgtgcc ttccttgacc ctggaaggtg ccactcccac tgtcctttcc
taataaaatg aggaaattgc atcgcattgt ctgagtaggt gtcattctat tctggggggt
                                                                     2940
ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca tgctggggac
togaqttaaq qqoqaattoo oqataaqqat ottootaqaq catqqotacq taqataaqta
gcatggcggg ttaatcatta actacaagga acccctagtg atggagttgg ccactccctc
                                                                     3120
tetgegeget egetegetea etgaggeegg gegaceaaag gtegeeegae geeegggett
                                                                     3180
tgcccgggcg gcctcagtga gcgagcgagc gcgcagcctt aattaaccta attcactggc
                                                                     3240
cgtcgtttta caacgtcgtg actgggaaaa ccctggcgtt acccaactta atcgccttgc
                                                                     3300
agcacatece cetttegeea getggegtaa tagegaagag geeegeaceg ategeeette
                                                                     3360
ccaacagttg cgcagcctga atggcgaatg ggacgcgccc tgtagcggcg cattaagcgc
                                                                     3420
ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgccc tagcgcccgc
                                                                     3480
teettteget ttetteeett eetttetege eacgttegee ggettteeee gteaagetet
                                                                     3540
aaatcggggg ctccctttag ggttccgatt tagtgcttta cggcacctcg accccaaaaa
acttgattag ggtgatggtt cacgtagtgg gccatcgccc tgatagacgg tttttcgccc
tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact
caaccctatc tcggtctatt cttttgattt ataagggatt ttgccgattt cggcctattg
                                                                     3780
gttaaaaaat gagctgattt aacaaaaatt taacgcgaat tttaacaaaa tattaacgct
tacaatttag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc
```

### -continued

```
taaatacatt caaatatgta toogotoatg agacaataac ootgataaat gottoaataa
                                                                    3960
tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt
                                                                    4020
geggeatttt geetteetgt tittgeteae eeagaaaege tggtgaaagt aaaagatget
                                                                    4080
gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc
                                                                    4140
                                                                    4200
cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta
tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac
                                                                    4260
tattctcaga atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc
                                                                    4320
atgacagtaa gagaattatg cagtgctgcc ataaccatga gtgataacac tgcggccaac
                                                                    4380
ttacttctga caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg
                                                                    4440
gatcatgtaa ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac
                                                                    4500
gagegtgaca ceaegatgee tgtageaatg geaacaaegt tgegeaaact attaaetgge
                                                                    4560
gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt
                                                                    4620
gcaggaccac ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga
gccggtgagc gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc
                                                                    4740
cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag
                                                                    4800
atcgctgaga taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca
                                                                    4860
tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc
ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca
gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc
tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta
ccaactettt tteegaaggt aaetggette ageagagege agataceaaa taetgttett
                                                                    5160
ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc
                                                                    5220
getetgetaa teetgetaee agtggetget gecagtggeg ataagtegtg tettaeeggg
ttggactcaa gacgatagtt accggataag gcgcagcggt cgggctgaac ggggggttcg
tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag
                                                                    5340
                                                                    5400
ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc
                                                                    5460
agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat
                                                                    5520
agtectgteg ggtttegeea cetetgaett gagegtegat ttttgtgatg etegteaggg
                                                                    5580
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc
                                                                    5640
tggccttttg ctcacatgtt ctttcctgcg ttatcccctg attctgtgga taaccgtatt
                                                                    5700
accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca
                                                                    5760
gtgagcgagg aagcggaaga gcgcccaata cgcaaaccgc ctctccccgc gcgttggccg
                                                                    5820
atteattaat geagetggea egacaggttt eeegactgga aagegggeag tgagegeaac
                                                                    5880
gcaattaatg tgagttaget cactcattag gcaccccagg ctttacactt tatgetteeg
                                                                    5940
gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac
                                                                    6000
catgattacg ccagatttaa ttaagg
                                                                    6026
SEQ ID NO: 23
                       moltype = DNA length = 5141
                       Location/Qualifiers
FEATURE
misc_feature
                       1..5141
                       note = constructed sequence
                       1..5141
source
                       mol_type = other DNA
organism = synthetic construct
SEOUENCE: 23
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact
                                                                    120
aggggtteet tgtagttaat gattaacceg ceatgetaet tatetaegta gecatgetet
                                                                    180
aggaagateg gaattegeee ttaagetage tgtttgetge ttgcaatgtt tgcccatttt
                                                                    240
agggtggaca caggacgctg tggtttctga gccagggggc gactcagatc ccagccagtg
                                                                    300
gacttagece etgtttgete eteegataac tggggtgace ttggttaata tteaceagea
                                                                    360
gcctcccccg ttgcccctct ggatccactg cttaaatacg gacgaggaca gggccctgtc
                                                                    420
tootoagott caggoaccac cactgacetg ggacagtgaa tetgcagaag ttggtegtga
                                                                    480
ggcactgggc aggtaagtat caaggttaca agacaggttt aaggagacca atagaaactg
                                                                    540
ggcttgtcga gacagagaag actcttgcgt ttctgatagg cacctattgg tcttactgac
atccactttg cctttctctc cacaggtgtc caggcggccg ccaccatgag cagcaagggc
                                                                    660
tetgtggtge tggeetacte tggeggeetg gacaccaget gtateetegt gtggetgaaa
                                                                    720
gaacagggct acgacgtgat cgcctacctg gccaacatcg gccagaaaga ggacttcgag
gaagcccgga agaaggccct gaagctgggc gccaagaagg tgttcatcga ggacgtgtcc
                                                                    840
cgcgagttcg tggaagagtt catctggccc gccatccaga gcagcgccct gtacgaggac
agatacetge tgggcaceag cetggceaga ceetgtateg eeeggaaaca ggtggaaate
gcccagcgcg agggcgccaa atacgtgtct cacggcgcca ccggcaaggg caacgaccag
gtgcgctttg agctgagctg ctactccctg gccccccaga tcaaagtgat cgccccttgg
cggatgcccg agttctacaa ccggttcaag ggccggaacg acctgatgga atacgccaag
                                                                    1140
caqcacqqca tocccatoco oqtqacococ aaqaaccett qqaqcatqqa cqaqaacctq
                                                                    1200
atgcacatca gctacgaggc cggcatcctg gaaaacccca agaatcaggc ccctcccggc
                                                                    1260
ctgtacacaa agacccagga ccctgccaag gcccccaaca cccccgacat tctggaaatc
                                                                    1320
gagttcaaga aaggcgtgcc cgtgaaagtg accaacgtga aggacggcac cacccaccag
                                                                    1380
acctccctgg aactgttcat gtacctgaac gaggtggccg gcaagcacgg cgtgggcaga
                                                                    1440
atcgacatcg tggaaaacag attcatcggc atgaagtccc ggggcatcta cgagacacca
                                                                    1500
qccqqcacca tcctgtacca cgcccacctg gatatcgagg ccttcaccat ggaccgggaa
                                                                    1560
gtgcggaaga tcaagcaggg cctgggcctg aagttcgccg agctggtgta cacaggcttt
                                                                    1620
tggcacagcc ccgagtgcga gtttgtgcgg cactgcattg ccaagagcca ggaacgggtg
                                                                    1680
gaaggcaagg tgcaggtgtc cgtgctgaag ggccaggtgt acattctggg cagagagagc
cccctgagcc tgtacaacga ggaactggtg tctatgaacg tgcagggcga ctacgagccc
                                                                    1800
accgacgcca ccggcttcat caacatcaac agcctgagac tgaaagagta ccaccggctg
                                                                    1860
cagtccaaag tgaccgccaa gtgataagca tgcggatctg cctcgactgt gccttctagt
                                                                    1920
tgccagccat ctgttgtttg cccctccccc gtgccttcct tgaccctgga aggtgccact
                                                                    1980
```

cccactgtcc tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat

## -continued

```
tctattctgg ggggtggggt ggggcaggac agcaaggggg aggattggga agacaatagc
aggeatgetg gggaetegag ttaagggega attecegatt aggatettee tagageatgg
                                                                   2160
ctacgtagat aagtagcatg gcgggttaat cattaactac aaggaacccc tagtgatgga
                                                                   2220
gttggccact ccctctctgc gcgctcgctc gctcactgag gccgggcgac caaaggtcgc
                                                                   2280
ccgacgcccg ggctttgccc gggcggcctc agtgagcgag cgagcgcgca gccttaatta
                                                                   2340
acctaattca ctggccgtcg ttttacaacg tcgtgactgg gaaaaccctg gcgttaccca
                                                                   2400
acttaatcgc cttgcagcac atcccccttt cgccagctgg cgtaatagcg aagaggcccg
                                                                   2460
caccgatege cetteceaac agttgegeag cetgaatgge gaatgggaeg egecetgtag
                                                                   2520
cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag
                                                                   2580
cgccctagcg cccgctcctt tcgctttctt cccttccttt ctcgccacgt tcgccggctt
                                                                   2640
tccccgtcaa gctctaaatc gggggctccc tttagggttc cgatttagtg ctttacggca
                                                                   2700
cctcgacccc aaaaaacttg attagggtga tggttcacgt agtgggccat cgccctgata
                                                                   2760
gacggttttt cgccctttga cgttggagtc cacgttcttt aatagtggac tcttgttcca
aactggaaca acactcaacc ctatctcggt ctattctttt gatttataag ggattttgcc
                                                                   2880
gatttcggcc tattggttaa aaaatgagct gatttaacaa aaatttaacg cgaattttaa
                                                                   2940
caaaatatta acgcttacaa tttaggtggc acttttcggg gaaatgtgcg cggaacccct
                                                                   3000
atttgtttat ttttctaaat acattcaaat atgtatccgc tcatgagaca ataaccctga
taaatgcttc aataatattg aaaaaggaag agtatgagta ttcaacattt ccgtgtcgcc
cttattccct tttttgcggc attttgcctt cctgtttttg ctcacccaga aacgctggtg
aaagtaaaag atgctgaaga tcagttgggt gcacgagtgg gttacatcga actggatctc
aacagcggta agatccttga gagttttcgc cccgaagaac gttttccaat gatgagcact
tttaaagttc tgctatgtgg cgcggtatta tcccgtattg acgccgggca agagcaactc
                                                                   3360
ggtcgccgca tacactattc tcagaatgac ttggttgagt actcaccagt cacagaaaag
catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat
                                                                   3480
aacactgcgg ccaacttact tctgacaacg atcggaggac cgaaggagct aaccgctttt
                                                                   3540
ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt gggaaccgga gctgaatgaa
                                                                   3600
gccataccaa acgacgagcg tgacaccacg atgcctgtag caatggcaac aacgttgcgc
                                                                   3660
aaactattaa ctggcgaact acttactcta gcttcccggc aacaattaat agactggatg
                                                                   3720
gaggeggata aagttgeagg accaettetg egeteggeee tteeggetgg etggtttatt
                                                                   3780
getgataaat etggageegg tgagegtggg tetegeggta teattgeage aetggggeea
                                                                   3840
gatggtaagc cctcccgtat cgtagttatc tacacgacgg ggagtcaggc aactatggat
                                                                   3900
gaacgaaata gacagatcgc tgagataggt gcctcactga ttaagcattg gtaactgtca
                                                                   3960
gaccaagttt actcatatat actttagatt gatttaaaac ttcattttta atttaaaagg
                                                                   4020
atctaggtga agatcctttt tgataatctc atgaccaaaa tcccttaacg tgagttttcg
                                                                   4080
ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tcctttttt
                                                                   4140
ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggt ggtttgtttg
                                                                   4200
ccggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata
                                                                   4260
ccaaatactg ttcttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca
                                                                   4320
cogcotacat acctogotot gotaatootg ttaccagtgg otgotgocag tggogataag
                                                                   4380
tegtgtetta eegggttgga eteaagaega tagttaeegg ataaggegea geggteggge
                                                                   4440
tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga
                                                                   4500
tacctacage gtgagetatg agaaagegee aegetteeeg aagggagaaa ggeggaeagg
                                                                   4560
tatccggtaa gcggcagggt cggaacagga gagcgcacga gggagcttcc agggggaaac
                                                                   4620
gcctggtatc tttatagtcc tgtcgggttt cgccacctct gacttgagcg tcgatttttg
                                                                   4680
tgatgctcgt caggggggcg gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg
                                                                   4740
ttcctggcct tttgctggcc ttttgctcac atgttctttc ctgcgttatc ccctgattct
                                                                   4800
gtggataacc gtattaccgc ctttgagtga gctgataccg ctcgccgcag ccgaacgacc
                                                                   4860
gagogoagog agtoagtgag ogaggaagog gaagagogoo caataogoaa acogoototo
                                                                   4920
cccgcgcgtt ggccgattca ttaatgcagc tggcacgaca ggtttcccga ctggaaagcg
                                                                   4980
ggcagtgagc gcaacgcaat taatgtgagt tagctcactc attaggcacc ccaggcttta
                                                                   5040
cactttatgc ttccggctcg tatgttgtgt ggaattgtga gcggataaca atttcacaca
                                                                   5100
ggaaacagct atgaccatga ttacgccaga tttaattaag g
                                                                   5141
SEQ ID NO: 24
                       moltype = DNA length = 6071
FEATURE
                       Location/Qualifiers
misc_feature
                       1..6071
                       note = constructed sequence
                       1..6071
source
                       mol_type = other DNA
organism = synthetic construct
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact
aggggtteet tgtagttaat gattaacceg ceatgetaet tatetaegta geeatgetet
aggaaqatcq qaattcqccc ttaaqctaqq qqqqatccac taqtactcqa qacctaqqaq
ttaattttta aaaagcagtc aaaagtccaa gtgcccttgc gagcatttac tctctctgtt
tgctctggtt aataatctca ggagcacaaa cattccttac tagttctagg agttaatttt
                                                                   360
taaaaagcag tcaaaagtcc aagtgccctt gcgagcattt actctctctg tttgctctgg
                                                                   420
ttaataatct caggagcaca aacattcctt actagttcta gagcggccgc cagtgtgctg
gaattegget tttttaggge tggaagetae etttgacate attteetetg egaatgeatg
                                                                   540
tataatttct acagaaccta ttagaaagga tcacccagcc tctgcttttg tacaactttc
ccttaaaaaa ctgccaatcc cactgctgtt tggcccaata gtgagaactt tttcctgctg
                                                                   660
cctcttggtg cttttgccta tggcccctat tctgcctgct gaagacactc ttgccagcat
                                                                   720
ggacttaaac ccctccagct ctgacaatcc tctttctctt ttgttttaca tgaagggtct
qqcaqccaaa qcaatcactc aaaqttcaaa ccttatcatt ttttqctttq ttcctcttqq
                                                                   840
ccttggtttt gtacatcagc tttgaaaata ccatcccagg gttaatgctg gggttaattt
                                                                   900
ataactgaga gtgctctagt tctgcaatac aggacatgct ataaaaatgg aaagatgttg
                                                                   960
ctttctgaga gatcagctta cttgtggtac cgagctcgga tcctgagaac ttcagggtga
```

gtctatggga cccttgatgt tttctttccc cttctttct atggttaagt tcatgtcata

```
ggaaggggag aagtaacagg gtacacatat tgaccaaatc agggtaattt tgcatttgta
attttaaaaa atgctttctt cttttaatat acttttttgt ttatcttatt tctaatactt
                                                                  1140
                                                                  1200
tecetaatet etttettea gggeaataat gatacaatgt ateatgeete tttgeaceat
                                                                  1260
tctaaagaat aacagtgata atttctgggt taaggcaata gcaatatttc tgcatataaa
                                                                  1320
tatttctgca tataaattgt aactgatgta agaggtttca tattgctaat agcagctaca
                                                                  1380
atccagctac cattctgctt ttattttatg gttgggataa ggctggatta ttctgagtcc
                                                                  1440
aagctaggcc cttttgctaa tcatgttcat acctcttatc ttcctcccac agctcctggg
                                                                  1500
caacgtgctg gtctgtgtgc tggcccatca ctttggcaaa gaattgatct cgagtaactg
                                                                  1560
aaggeggeeg ceaceatgag eageaaggge tetgtggtge tggeetacte tggeggeetg
                                                                  1620
gacaccagct gtatcctcgt gtggctgaaa gaacagggct acgacgtgat cgcctacctg
                                                                  1680
gccaacatcg gccagaaaga ggacttcgag gaagcccgga agaaggccct gaagctgggc
                                                                  1740
gccaagaagg tgttcatcga ggacgtgtcc cgcgagttcg tggaagagtt catctggccc
                                                                  1800
gccatccaga gcagcgccct gtacgaggac agatacctgc tgggcaccag cctggccaga
ccctgtatcg cccggaaaca ggtggaaatc gcccagcgcg agggcgccaa atacgtgtct
                                                                  1920
1980
gcccccaga tcaaagtgat cgcccttgg cggatgcccg agttctacaa ccggttcaag
                                                                  2040
ggccggaacg acctgatgga atacgccaag cagcacggca tccccatccc cgtgaccccc
aagaaccett ggagcatgga cgagaacctg atgcacatca gctacgagge cggcatcetg
                                                                  2160
gaaaacccca agaatcaggc ccctcccggc ctgtacacaa agacccagga ccctgccaag
gececcaaca ecceegacat tetggaaate gagtteaaga aaggegtgee egtgaaagtg
accaacgtga aggacggcac caccaaccag acctccctgg aactgttcat gtacctgaac
gaggtggccg gcaagcacgg cgtgggcaga atcgacatcg tggaaaacag attcatcggc
atgaaqtccc qqqqcatcta cqaqacacca qccqqcacca tcctqtacca cqcccacctq
                                                                  2460
gatatcgagg ccttcaccat ggaccgggaa gtgcggaaga tcaagcaggg cctgggcctg
                                                                  2520
aagttegeeg agetggtgta cacaggettt tggcacagee eegagtgega gtttgtgegg
                                                                  2580
cactgcattg ccaagagcca ggaacgggtg gaaggcaagg tgcaggtgtc cgtgctgaag
                                                                  2640
ggccaggtgt acattctggg cagagagagc cccctgagcc tgtacaacga ggaactggtg
                                                                  2700
tctatgaacg tgcagggcga ctacgagccc accgacgcca ccggcttcat caacatcaac
                                                                  2760
agcetgagac tgaaagagta ccaceggetg cagtecaaag tgacegecaa gtgataagca
                                                                  2820
tgcggatctg cctcgactgt gccttctagt tgccagccat ctgttgtttg cccctcccc
                                                                  2880
gtgccttcct tgaccctgga aggtgccact cccactgtcc tttcctaata aaatgaggaa
                                                                  2940
3000
agcaaggggg aggattggga agacaatagc aggcatgctg gggactcgag ttaagggcga
                                                                  3060
attcccgatt aggatcttcc tagagcatgg ctacgtagat aagtagcatg gcgggttaat
                                                                  3120
cattaactac aaggaacccc tagtgatgga gttggccact ccctctctgc gcgctcgctc
                                                                  3180
geteactgag geegggegae caaaggtege eegaegeeeg ggetttgeee gggeggeete
                                                                  3240
agtgagcgag cgagcgcgca gccttaatta acctaattca ctggccgtcg ttttacaacg
                                                                  3300
togtgactgg gaaaaccotg gogttaccca acttaatcgc cttgcagcac atcccccttt
                                                                  3360
cgccagctgg cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgcgcag
                                                                  3420
cctgaatggc gaatgggacg cgccctgtag cggcgcatta agcgcggcgg gtgtggttggt
                                                                  3480
tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt tcgctttctt
                                                                  3540
cccttccttt ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc gggggctccc
                                                                  3600
tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg attagggtga
                                                                  3660
tggttcacgt agtgggccat cgccctgata gacggttttt cgccctttga cgttggagtc
                                                                  3720
cacgttettt aatagtggae tettgtteea aactggaaca acacteaace etateteggt
                                                                  3780
ctattctttt gatttataag ggattttgcc gatttcggcc tattggttaa aaaatgagct
                                                                  3840
gatttaacaa aaatttaacg cgaattttaa caaaatatta acgcttacaa tttaggtggc
                                                                  3900
acttttcggg gaaatgtgcg cggaacccct atttgtttat ttttctaaat acattcaaat
                                                                  3960
atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggaag
                                                                  4020
agtatgagta ttcaacattt ccgtgtcgcc cttattccct tttttgcggc attttgcctt
                                                                  4080
cctgtttttg ctcacccaga aacgctggtg aaagtaaaag atgctgaaga tcagttgggt
                                                                  4140
gcacgagtgg gttacatcga actggatctc aacagcggta agatccttga gagttttcgc
                                                                  4200
cccgaagaac gttttccaat gatgagcact tttaaagttc tgctatgtgg cgcggtatta
                                                                  4260
tecegtattg acgeegggea agageaacte ggtegeegea tacactatte teagaatgae
                                                                  4320
ttggttgagt actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa
                                                                  4380
ttatgcagtg ctgccataac catgagtgat aacactgcgg ccaacttact tctgacaacg
                                                                  4440
atcggaggac cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc
                                                                  4500
cttgatcgtt gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg
atgcctgtag caatggcaac aacgttgcgc aaactattaa ctggcgaact acttactcta
getteeegge aacaattaat agactggatg gaggeggata aagttgeagg accaettetg
cgctcggccc ttccggctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg
tetegeggta teattgeage actggggeea gatggtaage cetecegtat egtagttate
tacacgacgg ggagtcaggc aactatggat gaacgaaata gacagatcgc tgagataggt
                                                                  4860
gcctcactga ttaagcattg gtaactgtca gaccaagttt actcatatat actttagatt
gatttaaaac ttcatttta atttaaaaqq atctaqqtqa aqatcctttt tqataatctc
atgaccaaaa tcccttaacg tgagttttcg ttccactgag cgtcagaccc cgtagaaaag
                                                                  5040
atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa
                                                                  5100
aaaccaccgc taccagcggt ggtttgtttg ccggatcaag agctaccaac tctttttccg
                                                                  5160
aaggtaactg gcttcagcag agegcagata ccaaatactg ttcttctagt gtagecgtag
ttaggccacc acttcaagaa ctctgtagca ccgcctacat acctcgctct gctaatcctg
                                                                  5280
ttaccagtgg ctgctgccag tggcgataag tcgtgtctta ccgggttgga ctcaagacga
                                                                  5340
tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc
                                                                  5400
ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagctatg agaaagcgcc
                                                                  5460
acgetteccg aagggagaaa ggeggacagg tateeggtaa geggeagggt eggaacagga
gagcgcacga gggagcttcc agggggaaac gcctggtatc tttatagtcc tgtcgggttt
                                                                  5580
egecacetet gaettgageg tegatttttg tgatgetegt eaggggggeg gageetatgg
                                                                  5640
aaaaacgcca gcaacgcggc ctttttacgg ttcctggcct tttgctggcc ttttgctcac
                                                                  5700
atgttctttc ctgcgttatc ccctgattct gtggataacc gtattaccgc ctttgagtga
                                                                  5760
gctgataccg ctcgccgcag ccgaacgacc gagcgcagcg agtcagtgag cgaggaagcg
                                                                  5820
```

```
gaagagegee caataegeaa acegeetete eeegegegtt ggeegattea ttaatgeage
                                                                   5880
tggcacgaca ggtttcccga ctggaaagcg ggcagtgagc gcaacgcaat taatgtgagt
                                                                   5940
tageteacte attaggeace ecaggettta caetttatge tteeggeteg tatgttgtgt
                                                                   6000
ggaattgtga gcggataaca atttcacaca ggaaacagct atgaccatga ttacgccaga
                                                                   6060
tttaattaag g
                                                                   6071
SEQ ID NO: 25
                      moltype = DNA length = 6135
FEATURE
                       Location/Qualifiers
                      1..6135
misc_feature
                      note = constructed sequence
source
                       1..6135
                       mol_type = other DNA
                      organism = synthetic construct
SEQUENCE: 25
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact
aggggtteet tgtagttaat gattaaceeg ceatgetaet tatetaegta geeatgetet
aggaagatcg gaattcgccc ttaagctagc aggttaattt ttaaaaaagca gtcaaaagtc
caagtggccc ttggcagcat ttactctctc tgtttgctct ggttaataat ctcaggagca
caaacattcc agatccaggt taatttttaa aaagcagtca aaagtccaag tggcccttgg
cagcatttac tetetetgtt tgetetggtt aataatetea ggagcacaaa cattecagat
ccggcgcgcc agggctggaa gctacctttg acatcatttc ctctgcgaat gcatgtataa
tttctacaga acctattaga aaggatcacc cagcctctgc ttttgtacaa ctttccctta
aaaaactgcc aattccactg ctgtttggcc caatagtgag aactttttcc tgctgcctct
                                                                   600
tggtgctttt gcctatggcc cctattctgc ctgctgaaga cactcttgcc agcatggact
                                                                   660
taaacccctc cagctctgac aatcctcttt ctcttttgtt ttacatgaag ggtctggcag
                                                                   720
ccaaagcaat cactcaaagt tcaaacctta tcattttttg ctttgttcct cttggccttg
                                                                   780
gttttgtaca tcagctttga aaataccatc ccagggttaa tgctggggtt aatttataac
                                                                   840
taagagtgct ctagttttgc aatacaggac atgctataaa aatggaaaga tgttgctttc
                                                                   900
tgagagactg cagaagttgg tcgtgaggca ctgggcaggt aagtatcaag gttacaagac
                                                                   960
aggtttaagg agaccaatag aaactgggct tgtcgagaca gagaagactc ttgcgtttct
                                                                   1020
gataggcacc tattggtctt actgacatcc actttgcctt tctctccaca ggtgtccagg
                                                                   1080
cggccgccac catgagcagc aagggctctg tggtgctggc ctactctggc ggcctggaca
                                                                   1140
ccagctgtat cctcgtgtgg ctgaaagaac agggctacga cgtgatcgcc tacctggcca
                                                                   1200
acateggeca gaaagaggae ttegaggaag ceeggaagaa ggeeetgaag etgggegeea
                                                                   1260
agaaggtgtt catcgaggac gtgtcccgcg agttcgtgga agagttcatc tggcccgcca
                                                                   1320
tccagagcag cgccctgtac gaggacagat acctgctggg caccagcctg gccagaccct
                                                                   1380
gtatcgcccg gaaacaggtg gaaatcgccc agcgcgaggg cgccaaatac gtgtctcacg
                                                                   1440
gcgccaccgg caagggcaac gaccaggtgc gctttgagct gagctgctac tccctggccc
                                                                   1500
cccagatcaa agtgatcgcc ccttggcgga tgcccgagtt ctacaaccgg ttcaagggcc
                                                                   1560
ggaacgacct gatggaatac gccaagcagc acggcatccc catccccgtg acccccaaga
                                                                   1620
accettggag catggacgag aacctgatge acateageta egaggeegge atcetggaaa
                                                                   1680
accccaagaa tcaggcccct cccggcctgt acacaaagac ccaggaccct gccaaggccc
                                                                   1740
ccaacacccc cgacattctg gaaatcgagt tcaagaaagg cgtgcccgtg aaagtgacca
                                                                   1800
acgtgaagga cggcaccacc caccagacct ccctggaact gttcatgtac ctgaacgagg
                                                                   1860
tggccggcaa gcacggcgtg ggcagaatcg acatcgtgga aaacagattc atcggcatga
                                                                   1920
agtcccgggg catctacgag acaccagccg gcaccatcct gtaccacgcc cacctggata
                                                                   1980
togaggoott caccatggac ogggaagtgo ggaagatcaa gcagggootg ggootgaagt
                                                                   2040
tegeegaget ggtgtacaca ggettttgge acageecega gtgegagttt gtgeggeact
                                                                   2100
gcattgccaa gagccaggaa cgggtggaag gcaaggtgca ggtgtccgtg ctgaagggcc
                                                                   2160
aggtgtacat tctgggcaga gagagccccc tgagcctgta caacgaggaa ctggtgtcta
                                                                   2220
tgaacgtgca gggcgactac gagcccaccg acgccaccgg cttcatcaac atcaacagcc
                                                                   2280
tgagactgaa agagtaccac cggctgcagt ccaaagtgac cgccaagtga taagcatgcg
                                                                   2340
gatccaatca acctctggat tacaaaattt gtgaaagatt gactggtatt cttaactatg
                                                                   2400
ttgctccttt tacgctatgt ggatacgctg ctttaatgcc tttgtatcat gctattgctt
                                                                   2460
cccgtatggc tttcattttc tcctccttgt ataaatcctg gttgctgtct ctttatgagg
                                                                   2520
agttgtggcc cgttgtcagg caacgtggcg tggtgtgcac tgtgtttgct gacgcaaccc
                                                                   2580
ccactggttg gggcattgcc accacctgtc agctcctttc cgggactttc gctttccccc
tecetattge caeggeggaa eteategeeg eetgeettge eegetgetgg acaggggete
ggctgttggg cactgacaat tccgtggtgt tgtcggggaa gctgacgtcc tttccatggc
tgctcgcctg tgttgccacc tggattctgc gcgggacgtc cttctgctac gtcccttcgg
ccctcaatcc agcggacctt ccttcccgcg gcctgctgcc ggctctgcgg cctcttccgc
gtettegaga tetgeetega etgtgeette tagttgeeag ceatetgttg tttgeeeete
ccccgtgcct tccttgaccc tggaaggtgc cactcccact gtcctttcct aataaaatga
qqaaattqca tcqcattqtc tqaqtaqqtq tcattctatt ctqqqqqqtq qqqtqqqca
ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctggggact cgagttaagg
                                                                   3120
gcgaattccc gattaggatc ttcctagagc atggctacgt agataagtag catggcgggt
                                                                   3180
taatcattaa ctacaaggaa cccctagtga tggagttggc cactccctct ctgcgcgctc
                                                                   3240
getegeteac tgaggeeggg egaceaaagg tegeeegaeg eeegggettt geeegggegg
                                                                   3300
cctcagtgag cgagcgagcg cgcagcctta attaacctaa ttcactggcc gtcgttttac
                                                                   3360
aacgtcgtga ctgggaaaac cctggcgtta cccaacttaa tcgccttgca gcacatcccc
                                                                   3420
ctttcgccag ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc caacagttgc
                                                                   3480
gcagcctgaa tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg
                                                                   3540
tggttacgcg cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt
tettecette etttetegee aegttegeeg gettteeeeg teaageteta aateggggge
                                                                   3660
tecetttagg gtteegattt agtgetttae ggeacetega eeceaaaaaa ettgattagg
                                                                   3720
gtgatggttc acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg
                                                                   3780
agtocacgtt ctttaatagt ggactottgt tocaaactgg aacaacactc aaccotatct
cggtctattc ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg
```

### -continued

```
agctgattta acaaaaattt aacgcgaatt ttaacaaaat attaacgctt acaatttagg
tggcactttt cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc
                                                                    4020
aaatatgtat cogotoatga gacaataaco otgataaatg ottoaataat attgaaaaag
                                                                    4080
gaagagtatg agtattcaac atttccgtgt cgcccttatt cccttttttg cggcattttg
                                                                    4140
                                                                    4200
ccttcctgtt tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt
gggtgcacga gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt
                                                                    4260
tcgccccgaa gaacgttttc caatgatgag cacttttaaa gttctgctat gtggcgcggt
                                                                    4320
attatcccgt attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa
                                                                    4380
tgacttggtt gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag
                                                                    4440
agaattatgc agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac
                                                                    4500
aacgatcgga ggaccgaagg agctaaccgc ttttttgcac aacatggggg atcatgtaac
                                                                    4560
togocttgat ogttgggaac oggagotgaa tgaagocata ocaaacgaog agogtgacac
                                                                    4620
cacgatgcct gtagcaatgg caacaacgtt gcgcaaacta ttaactggcg aactacttac
totagettee eggeaacaat taatagactg gatggaggeg gataaagttg caggaceact
                                                                    4740
tetgegeteg geeetteegg etggetggtt tattgetgat aaatetggag eeggtgageg
                                                                    4800
tgggtctcgc ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt
                                                                    4860
tatctacacg acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat
aggtgcctca ctgattaagc attggtaact gtcagaccaa gtttactcat atatacttta
                                                                    4980
gattgattta aaacttcatt tttaatttaa aaggatctag gtgaagatcc tttttgataa
                                                                    5040
totoatgaco aaaatooott aacgtgagtt ttogttocac tgagogtoag accoogtaga
aaagatcaaa ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac
                                                                    5160
aaaaaaacca ccgctaccag cggtggtttg tttgccggat caagagctac caactctttt
                                                                    5220
tecqaaqqta actqqettea qeaqaqeqea qataccaaat actqttette taqtqtaqee
gtagttaggc caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaat
                                                                    5340
cctgttacca gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag
                                                                    5400
acgatagtta ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc
                                                                    5460
cagcttggag cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag
                                                                    5520
cgccacgctt cccgaaggga gaaaggcgga caggtatccg gtaagcggca gggtcggaac
                                                                    5580
aggagagege acgagggage ttccaggggg aaacgcetgg tatetttata gteetgtegg
                                                                    5640
gtttcgccac ctctgacttg agcgtcgatt tttgtgatge tcgtcagggg ggcggagcct atggaaaaac gccagcaacg cggccttttt acggttcctg gccttttgc ggcctttttc
                                                                    5700
                                                                    5760
tcacatgttc tttcctgcgt tatcccctga ttctgtggat aaccgtatta ccgcctttga
                                                                    5820
gtgagctgat accgctcgcc gcagccgaac gaccgagcgc agcgagtcag tgagcgagga
                                                                    5880
agoggaagag ogcocaatac gcaaacogco totococgcg cgttggccga ttcattaatg
                                                                    5940
cagetggeac gaeaggttte eegaetggaa agegggeagt gagegeaacg caattaatgt
                                                                    6000
gagttagctc actcattagg caccccaggc tttacacttt atgcttccgg ctcgtatgtt
                                                                    6060
gtgtggaatt gtgagcggat aacaatttca cacaggaaac agctatgacc atgattacgc
                                                                    6120
cagatttaat taagg
                                                                    6135
                       moltype = DNA length = 5587
Location/Qualifiers
SEQ ID NO: 26
FEATURE
                       1 5587
misc_feature
                       note = constructed sequence
                       1..5587
source
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 26
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact
                                                                    120
aggggtteet tgtagttaat gattaaceeg ecatgetaet tatetaegta gecatgetet
                                                                    180
aggaagatcg gaattcgccc ttaagctagc aggttaattt ttaaaaagca gtcaaaagtc
                                                                    240
caagtggccc ttggcagcat ttactctctc tgtttgctct ggttaataat ctcaggagca
                                                                    300
caaacattcc agatccaggt taatttttaa aaagcagtca aaagtccaag tggcccttgg
                                                                    360
cagcatttac tctctctgtt tgctctggtt aataatctca ggagcacaaa cattccagat
ccggcgcgcc agggctggaa gctacctttg acatcatttc ctctgcgaat gcatgtataa
                                                                    480
tttctacaga acctattaga aaggatcacc cagcctctgc ttttgtacaa ctttccctta
                                                                    540
aaaaactgcc aattccactg ctgtttggcc caatagtgag aactttttcc tgctgcctct
                                                                    600
tggtgctttt gcctatggcc cctattctgc ctgctgaaga cactcttgcc agcatggact
taaacccctc cagctctgac aatcctcttt ctcttttgtt ttacatgaag ggtctggcag
                                                                    720
ccaaagcaat cactcaaagt tcaaacctta tcattttttg ctttgttcct cttggccttg
gttttgtaca tcagctttga aaataccatc ccagggttaa tgctggggtt aatttataac
                                                                    840
taagagtgct ctagttttgc aatacaggac atgctataaa aatggaaaga tgttgctttc
tgaqaqactq cagaaqttqq tcqtqaqqca ctqqqcaqqt aaqtatcaaq qttacaaqac
aggtttaagg agaccaatag aaactgggct tgtcgagaca gagaagactc ttgcgtttct
gataggcacc tattggtctt actgacatcc actttgcctt tctctccaca ggtgtccagg
cggccgccac catgagcagc aagggctctg tggtgctggc ctactctggc ggcctggaca
                                                                    1140
ccagctgtat cctcgtgtgg ctgaaagaac agggctacga cgtgatcgcc tacctggcca
                                                                    1200
acateggeca gaaagaggae ttegaggaag ceeggaagaa ggeeetgaag etgggegeca
                                                                    1260
agaaggtgtt catcgaggac gtgtcccgcg agttcgtgga agagttcatc tggcccgcca
tccagagcag cgccctgtac gaggacagat acctgctggg caccagcctg gccagaccct
                                                                    1380
gtatcgcccg gaaacaggtg gaaatcgccc agcgcgaggg cgccaaatac gtgtctcacg
                                                                    1440
gegecacegg caagggcaac gaccaggtge getttgaget gagetgetae teeetggeee
                                                                    1500
cccagatcaa agtgatcgcc ccttggcgga tgcccgagtt ctacaaccgg ttcaagggcc
                                                                    1560
ggaacgacct gatggaatac gccaagcagc acggcatccc catccccgtg acccccaaga
accettggag catggacgag aacetgatge acateageta egaggeegge ateetggaaa
accccaagaa tcaggccct cccggcctgt acacaaagac ccaggaccct gccaaggccc
                                                                    1740
ccaacacccc cgacattctg gaaatcgagt tcaagaaagg cgtgcccgtg aaagtgacca
                                                                    1800
acgtgaagga cggcaccacc caccagacct ccctggaact gttcatgtac ctgaacgagg
```

tggccggcaa gcacggcgtg ggcagaatcg acatcgtgga aaacagattc atcggcatga

```
-continued
```

agtcccgggg catctacgag acaccagccg gcaccatcct gtaccacgcc cacctggata

```
tcgaggcctt caccatggac cgggaagtgc ggaagatcaa gcagggcctg ggcctgaagt
tegeegaget ggtgtacaca ggettttgge acageeeega gtgegagttt gtgeggeaet
                                                                   2100
gcattgccaa gagccaggaa cgggtggaag gcaaggtgca ggtgtccgtg ctgaagggcc
                                                                   2160
                                                                   2220
aggtgtacat tctgggcaga gagagccccc tgagcctgta caacgaggaa ctggtgtcta
tgaacgtgca gggcgactac gagcccaccg acgccaccgg cttcatcaac atcaacagcc
                                                                   2280
tgagactgaa agagtaccac cggctgcagt ccaaagtgac cgccaagtga taagcatgcg
                                                                   2340
gatetgeete gaetgtgeet tetagttgee agecatetgt tgtttgeece teeceegtge
                                                                   2400
cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat gaggaaattg
                                                                   2460
2520
agggggagga ttgggaagac aatagcaggc atgctgggga ctcgagttaa gggcgaattc
                                                                   2580
ccgattagga tcttcctaga gcatggctac gtagataagt agcatggcgg gttaatcatt
                                                                   2640
aactacaagg aaccectagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc
actgaggeeg ggegaccaaa ggtegeeega egeeeggget ttgeeeggge ggeeteagtg
                                                                   2760
agcgagcgag cgcgcagcct taattaacct aattcactgg ccgtcgtttt acaacgtcgt
                                                                   2820
gactgggaaa accetggcgt tacceaactt aatcgcettg cagcacatce ccetttegce
                                                                   2880
agetggegta atagegaaga ggeeegeace gategeeett eecaacagtt gegeageetg
aatggcgaat gggacgcgcc ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg
cgcagcgtga ccgctacact tgccagcgcc ctagcgcccg ctcctttcgc tttcttccct
teettteteg ceaegttege eggettteee egteaagete taaategggg geteeettta
gggttccgat ttagtgcttt acggcacctc gaccccaaaa aacttgatta gggtgatggt
tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg
                                                                   3240
ttotttaata gtggactott gttocaaact ggaacaacac toaaccotat ctoggtotat
                                                                   3300
tettttgatt tataagggat tttgeegatt teggeetatt ggttaaaaaa tgagetgatt
                                                                   3360
taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc ttacaattta ggtggcactt
                                                                   3420
ttcggggaaa tgtgcgcgga acccctattt gtttattttt ctaaatacat tcaaatatgt
                                                                   3480
atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta
                                                                   3540
tgagtattca acatttccgt gtcgccctta ttcccttttt tgcggcattt tgccttcctg
                                                                   3600
tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac
                                                                   3660
gagtgggtta catcgaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg
                                                                   3720
aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc
                                                                   3780
gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg
ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat
                                                                   3840
                                                                   3900
gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg
                                                                   3960
gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg
                                                                   4020
atcgttggga accggagetg aatgaageea taccaaacga cgagegtgae accaegatge
                                                                   4080
ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt
                                                                   4140
cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct
                                                                   4200
cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc
                                                                   4260
geggtateat tgeageactg gggeeagatg gtaageeete eegtategta gttatetaea
                                                                   4320
cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct
                                                                   4380
cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt
                                                                   4440
taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat aatctcatga
                                                                   4500
ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca
                                                                   4560
aaggatette ttgagateet ttttttetge gegtaatetg etgettgeaa acaaaaaaa
                                                                   4620
caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg
                                                                   4680
taactggctt cagcagagcg cagataccaa atactgttct tctagtgtag ccgtagttag
                                                                   4740
gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac
                                                                   4800
cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt
                                                                   4860
taccggataa ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag cccagcttgg
                                                                   4920
agegaacgae ctacacegaa etgagatace tacagegtga getatgagaa agegecaege
                                                                   4980
ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc
                                                                   5040
gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc
                                                                   5100
acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa
                                                                   5160
acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt
                                                                   5220
tettteetge gttateeect gattetgtgg ataacegtat tacegeettt gagtgagetg
                                                                   5280
ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag
                                                                   5340
agegeceaat aegeaaaceg ceteteceeg egegttggee gatteattaa tgeagetgge
acgacaggtt tcccgactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc
                                                                   5460
tcactcatta ggcacccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa
                                                                   5520
ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgattac gccagattta
                                                                   5580
attaagg
SEQ ID NO: 27
                       moltype = DNA length = 5587
FEATURE
                       Location/Qualifiers
                       1..5587
misc_feature
                       note = constructed sequence
source
                       1..5587
                       mol type = other DNA
                       organism = synthetic construct
SEQUENCE: 27
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact
aggggtteet tgtagttaat gattaaeceg ceatgetaet tatetaegta geeatgetet
aggaagatcg gaattcgccc ttaagctagc aggttaattt ttaaaaagca gtcaaaagtc
                                                                   240
caagtggccc ttggcagcat ttactctctc tgtttgctct ggttaataat ctcaggagca
                                                                   300
caaacattcc agatccaggt taatttttaa aaagcagtca aaagtccaag tggcccttgg
                                                                   360
cagcatttac tctctctgtt tgctctggtt aataatctca ggagcacaaa cattccagat
ccggcgcgcc agggctggaa gctacctttg acatcatttc ctctgcgaat gcatgtataa
```

46

1980

2040

540

## -continued

tttctacaga acctattaga aaggatcacc cagcctctgc ttttgtacaa ctttccctta aaaaactgcc aattccactg ctgtttggcc caatagtgag aactttttcc tgctgcctct 600 tggtgctttt gcctatggcc cctattctgc ctgctgaaga cactcttgcc agcatggact 660 taaacccctc cagctctgac aatcctcttt ctcttttgtt ttacatgaag ggtctggcag 720 ccaaagcaat cactcaaagt tcaaacctta tcattttttg ctttgttcct cttggccttg 780 gttttgtaca tcagctttga aaataccatc ccagggttaa tgctggggtt aatttataac 840 taagagtgct ctagttttgc aatacaggac atgctataaa aatggaaaga tgttgctttc 900 tgagagactg cagaagttgg tcgtgaggca ctgggcaggt aagtatcaag gttacaagac 960 aggtttaagg agaccaatag aaactgggct tgtcgagaca gagaagactc ttgcgtttct 1020 gataggeace tattggtett actgacatee actttgeett teteteeaca ggtgteeagg 1080 eggeegeeac catgiceage aaaggeteeg tggttetgge etacagtgge ggeetggaca 1140 cctcgtgcat cctcgtgtgg ctgaaggaac aaggctatga cgtcattgcc tatctggcca 1200 acattggcca gaaggaagac ttcgaggaag ccaggaagaa ggcactgaag cttggggcca aaaaggtgtt cattgaggat gtcagcaggg agtttgtgga ggagttcatc tggccggcca 1320 tccagtccag cgcactgtat gaggaccgct acctcctggg cacctctctt gccaggccct 1380 gcatcgcccg caaacaagtg gaaatcgccc agcgggaggg ggccaagtat gtgtcccacg 1440 gcgccacagg aaaggggaac gatcaggtcc ggtttgagct cagctgctac tcactggccc cccagataaa ggtcattgct ccctggagga tgcctgaatt ctacaaccgg ttcaagggcc gcaatgacct gatggagtac gcaaagcaac acgggattcc catcccggtc actcccaaga accogtggag catggatgag aacctcatgc acatcagcta cgaggctgga atcctggaga accccaagaa ccaagcgcct ccaggtctct acacgaagac ccaggaccca gccaaagcc ccaacacccc tgacattctc gagatcgagt tcaaaaaaagg ggtccctgtg aaggtgacca acgtcaagga tggcaccacc caccagacct ccttggagct cttcatgtac ctgaacgaag 1860 tcgcgggcaa gcatggcgtg ggccgtattg acatcgtgga gaaccgcttc attggaatga 1920 agtcccgagg tatctacgag accccagcag gcaccatcct ttaccacgct catttagaca 1980 tcgaggcctt caccatggac cgggaagtgc gcaaaatcaa acaaggcctg ggcttgaaat 2040 ttgctgagct ggtgtatacc ggtttctggc acagccctga gtgtgaattt gtccgccact 2100 gcatcgccaa gtcccaggag cgagtggaag ggaaagtgca ggtgtccgtc ctcaagggcc 2160 aggtgtacat cctcggccgg gagtccccac tgtctctcta caatgaggag ctggtgagca 2220 tgaacgtgca gggtgattat gagccaactg atgccaccg gttcatcaac atcaattccc tcaggctgaa ggaatatcat cgtctccaga gcaaggtcac tgccaaatga taagcatgcg 2280 2340 gatetgeete gaetgtgeet tetagttgee agecatetgt tgtttgeece teeceegtge 2400 cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat gaggaaattg 2460 2520 agggggagga ttgggaagac aatagcaggc atgctgggga ctcgagttaa gggcgaattc 2580 ccgattagga tcttcctaga gcatggctac gtagataagt agcatggcgg gttaatcatt 2640 aactacaagg aaccectagt gatggagttg gecacteeet etetgegege tegetegete 2700 actgaggeeg ggegaceaaa ggtegeeega egeeeggget ttgeeeggge ggeeteagtg 2760 agegagegag egegeageet taattaaeet aatteaetgg eegtegttit acaaegtegt 2820 gactgggaaa accetggegt tacceaactt aategeettg cageacatee eeetttegee 2880 agetggegta atagegaaga ggecegeace gategeeett cecaacagtt gegeageetg 2940 aatggcgaat gggacgcgcc ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg 3000 egeagegtga eegetacaet tgeeagegee etagegeeeg eteetttege tttetteeet 3060 teettteteg ceaegttege eggettteee egteaagete taaategggg geteeettta 3120 gggttccgat ttagtgcttt acggcacctc gaccccaaaa aacttgatta gggtgatggt 3180 tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg 3240 ttctttaata gtggactctt gttccaaact ggaacaacac tcaaccctat ctcggtctat 3300 tettttgatt tataagggat tttgccgatt teggeetatt ggttaaaaaa tgagetgatt 3360 taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc ttacaattta ggtggcactt 3420 ttcggggaaa tgtgcgcgga acccctattt gtttattttt ctaaatacat tcaaatatgt 3480 atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta 3540 tgagtattca acatttccgt gtcgccctta ttcccttttt tgcggcattt tgccttcctg 3600 tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac 3660 gagtgggtta catcgaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg 3720 aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc 3780 gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg 3840 ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat 3900 gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg 4020 atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt 4140 cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct eggeeettee ggetggetgg tttattgetg ataaatetgg ageeggtgag egtgggtete 4260 gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca cgacqqqqaq tcaqqcaact atqqatqaac qaaataqaca qatcqctqaq ataqqtqcct cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt 4440 taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat aatctcatga 4500 ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca 4560 aaggatette ttgagateet tttttetge gegtaatetg etgettgeaa acaaaaaac caccgctacc agoggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg 4680 taactggctt cagcagagcg cagataccaa atactgttct tctagtgtag ccgtagttag 4740 gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac 4800 cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt 4860 taccggataa ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc 4980 ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc 5040 gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc 5100 acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa 5160 acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt

```
tettteetge gttateeeet gattetgtgg ataacegtat tacegeettt gagtgagetg
                                                                     5280
ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag
                                                                     5340
agegeceaat aegeaaaceg ceteteeceg egegttggee gatteattaa tgeagetgge
                                                                     5400
acgacaggtt tcccgactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc
                                                                     5460
tcactcatta ggcaccccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa
                                                                     5520
ttgtgagegg ataacaattt cacacaggaa acagctatga ccatgattac gccagattta
                                                                     5580
                                                                     5587
SEQ ID NO: 28
                       moltype = DNA length = 6026
FEATURE
                       Location/Qualifiers
misc feature
                       1..6026
                       note = constructed sequence
                       1..6026
source
                       mol_type = other DNA
organism = synthetic construct
SEQUENCE: 28
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt
ggtcgccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact
aggggtteet tgtagttaat gattaacceg ceatgetaet tatetaegta geeatgetet
aggaagatcg gaattcgccc ttaaggctca gaggcacaca ggagtttctg ggctcaccct
geoceettee aacceeteag tteccateet ceageagetg tttgtgtget geetetgaag
tccacactga acaaacttca gcctactcat gtccctaaaa tgggcaaaca ttgcaagcag
caaacagcaa acacacagcc ctccctgcct gctgaccttg gagctggggc agaggtcaga
gacctctctg ggcccatgcc acctccaaca tccactcgac cccttggaat ttcggtggag
                                                                     480
aggagcagag gttgtcctgg cgtggtttag gtagtgtgag agggcgcgcc gatcttgcta
ccagtggaac agccactaag gattotgcag tgagagcaga gggccagcta agtggtactc tcccagagac tgtctgactc acgccacccc ctccaccttg gacacaggac gctgtggttt
                                                                     600
                                                                     660
ctgagccagg tacaatgact cctttcggta agtgcagtgg aagctgtaca ctgcccaggc
                                                                     720
aaagcgtccg ggcagcgtag gcgggcgact cagatcccag ccagtggact tagccctgt
                                                                     780
ttgctcctcc gataactggg gtgaccttgg ttaatattca ccagcagcct cccccgttgc
                                                                     840
ccctctggat ccactgctta aatacggacg aggacagggc cctgtctcct cagcttcagg
                                                                     900
caccaccact gacctgggac agtagateta gettaettgt ggtaccgage teggatectg
                                                                     960
agaacttcag ggtgagtcta tgggaccctt gatgttttct ttccccttct tttctatggt
                                                                     1020
taagttcatg tcataggaag gggagaagta acagggtaca catattgacc aaatcagggt
aattttgcat ttgtaattt aaaaaatgct ttcttcttt aatatactt tttgttatc
                                                                     1080
                                                                     1140
ttatttctaa tactttccct aatctctttc tttcagggca ataatgatac aatgtatcat
                                                                     1200
gcctctttgc accattctaa agaataacag tgataatttc tgggttaagg caatagcaat
                                                                     1260
atttctgcat ataaatattt ctgcatataa attgtaactg atgtaagagg tttcatattg
                                                                     1320
ctaatagcag ctacaatcca gctaccattc tgcttttatt ttatggttgg gataaggctg
                                                                     1380
gattattctg agtccaaget aggccctttt gctaatcatg ttcatacete ttatetteet
                                                                     1440
cccacagete etgggeaacg tgetggtetg tgtgetggee cateaetttg gcaaagaatt
                                                                     1500
gatetegagt aactgaagge ggeegeeace atgteeagea aaggeteegt ggttetggee
                                                                     1560
tacagtggcg gcctggacac ctcgtgcatc ctcgtgtggc tgaaggaaca aggctatgac
                                                                     1620
gtcattgcct atctggccaa cattggccag aaggaagact tcgaggaagc caggaagaag
                                                                     1680
gcactgaagc ttggggccaa aaaggtgttc attgaggatg tcagcaggga gtttgtggag
                                                                     1740
gagttcatct ggccggccat ccagtccagc gcactgtatg aggaccgcta cctcctgggc
                                                                     1800
acctetettg ccaggecetg categeeege aaacaagtgg aaategeeea gegggagggg
                                                                     1860
gccaagtatg tgtcccacgg cgccacagga aaggggaacg atcaggtccg gtttgagctc
                                                                     1920
agetgetact caetggeece ceagataaag gteattgete eetggaggat geetgaatte
                                                                     1980
tacaaccggt tcaagggccg caatgacctg atggagtacg caaagcaaca cgggattccc
                                                                     2040
atcccggtca ctcccaagaa cccgtggagc atggatgaga acctcatgca catcagctac
                                                                     2100
gaggetggaa teetggagaa eeccaagaac caagegeete caggteteta caegaagace
                                                                     2160
caggacccag ccaaagcccc caacacccct gacattctcg agatcgagtt caaaaaaggg
                                                                     2220
gtccctgtga aggtgaccaa cgtcaaggat ggcaccaccc accagacctc cttggagctc
ttcatgtacc tgaacgaagt cgcgggcaag catggcgtgg gccgtattga catcgtggag
                                                                     2340
aaccgcttca ttggaatgaa gtcccgaggt atctacgaga ccccagcagg caccatcctt
                                                                     2400
taccacgete atttagacat egaggeette accatggace gggaagtgeg caaaatcaaa
                                                                     2460
caaggeetgg gettgaaatt tgetgagetg gtgtataceg gtttetggea cageeetgag
tgtgaatttg tccgccactg catcgccaag tcccaggagc gagtggaagg gaaagtgcag
gtgtccgtcc tcaagggcca ggtgtacatc ctcggccggg agtccccact gtctctctac
aatgaggagc tggtgagcat gaacgtgcag ggtgattatg agccaactga tgccaccggg
ttcatcaaca tcaattccct caggctgaag gaatatcatc gtctccagag caaggtcact
gccaaatgat aagcatgegg atetgeeteg actgtgeett etagttgeea gccatetgtt
                                                                     2820
gtttgcccct cccccgtgcc ttccttgacc ctggaaggtg ccactcccac tgtcctttcc
taataaaatq aqqaaattqc atcqcattqt ctqaqtaqqt qtcattctat tctqqqqqqt
ggggtgggc aggacagcaa gggggaggat tgggaagaca atagcaggca tgctggggac
                                                                     3000
tcgagttaag ggcgaattcc cgataaggat cttcctagag catggctacg tagataagta
                                                                     3060
gcatggcggg ttaatcatta actacaagga acccctagtg atggagttgg ccactccctc
                                                                     3120
tetgegeget egetegetea etgaggeegg gegaceaaag gtegeeegae geeegggett
                                                                     3180
tgcccgggcg gcctcagtga gcgagcgagc gcgcagcctt aattaaccta attcactggc
                                                                     3240
cgtcgtttta caacgtcgtg actgggaaaa ccctggcgtt acccaactta atcgccttgc
                                                                     3300
agcacatece cetttegeca getggegtaa tagegaagag geeegeaceg ategeeette
                                                                     3360
ccaacagttg cgcagcctga atggcgaatg ggacgcgccc tgtagcggcg cattaagcgc
                                                                     3420
ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgccc tagcgcccgc
teettteget ttetteeett eetttetege eacgttegee ggettteeee gteaagetet
                                                                     3540
aaatcggggg ctccctttag ggttccgatt tagtgcttta cggcacctcg accccaaaaa
                                                                     3600
acttgattag ggtgatggtt cacgtagtgg gccatcgccc tgatagacgg tttttcgccc
                                                                     3660
tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact
caaccctatc toggtotatt cttttgattt ataagggatt ttgccgattt cggcctattg 3780
```

```
-continued
```

```
gttaaaaaat gagctgattt aacaaaaatt taacgcgaat tttaacaaaa tattaacgct
tacaatttag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc
                                                                   3900
taaatacatt caaatatgta teegeteatg agacaataac eetgataaat getteaataa
                                                                   3960
tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt
                                                                   4020
geggeatttt geetteetgt ttttgeteac ceagaaaege tggtgaaagt aaaagatget
                                                                   4080
gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc
                                                                   4140
cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta
                                                                   4200
tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac
                                                                   4260
tatteteaga atgaetteget tgagtaetea eeagteacag aaaageatet taeggatgge
                                                                   4320
atgacagtaa gagaattatg cagtgctgcc ataaccatga gtgataacac tgcggccaac
                                                                   4380
ttacttctga caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg
                                                                   4440
gatcatgtaa ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac
                                                                   4500
gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc
                                                                   4560
gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt
                                                                   4620
gcaggaccac ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga
                                                                   4680
geoggtgage gtgggteteg eggtateatt geageaetgg ggeoagatgg taageeetee
                                                                   4740
cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag
atcgctgaga taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca
                                                                   4860
tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc
ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca
gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc
tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta
                                                                   5100
ccaactettt tteegaaggt aactggette ageagagege agataccaaa taetgttett
                                                                   5160
ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc
                                                                   5220
getetgetaa teetgttaee agtggetget geeagtggeg ataagtegtg tettaeeggg
                                                                   5280
ttggactcaa gacgatagtt accggataag gcgcagcggt cgggctgaac ggggggttcg
                                                                   5340
tgcacacage ccagettgga gegaacgace tacacegaac tgagatacet acagegtgag
                                                                   5400
ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc
                                                                   5460
agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat
                                                                   5520
agtectgteg ggtttegeea cetetgaett gagegtegat ttttgtgatg etegteaggg
                                                                   5580
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc
                                                                   5640
tggccttttg ctcacatgtt ctttcctgcg ttatcccctg attctgtgga taaccgtatt
                                                                   5700
accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca
                                                                   5760
gtgagcgagg aagcggaaga gcgcccaata cgcaaaccgc ctctccccgc gcgttggccg
                                                                   5820
attcattaat gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac
                                                                   5880
gcaattaatg tgagttaget cactcattag gcaccccagg etttacaett tatgetteeg
                                                                   5940
gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac
                                                                   6000
catgattacg ccagatttaa ttaagg
                                                                   6026
                       moltype = DNA length = 5876
Location/Qualifiers
SEO ID NO: 29
FEATURE
                       1..5876
misc_feature
                       note = constructed sequence
                       1..5876
source
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 29
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact
                                                                   120
aggggtteet tgtagttaat gattaacceg ceatgetaet tatetaegta gecatgetet
                                                                   180
aggaagatcg gaattcgccc ttaagctagc cgctgacaag gctcagaggc acacaggagt
                                                                   240
ttctgggctc accetgcccc cttccaaccc ctcagttccc atcctccage agctgtttgt
                                                                   300
gtgctgcctc tgaagtccac actgaacaaa cttcagccta ctcatgtccc taaaatgggc
                                                                   360
aaacattgca agcagcaaac agcaaacaca cagccctccc tgcctgctga ccttggagct
ggggcagagg tcagagacct ctctgggccc atgccacctc caacatccac tcgacccctt
                                                                   480
ggaatttcgg tggagaggag cagaggttgt cctggcgtgg tttaggtagt gtgagaggga
                                                                   540
gateeggege geetggacae aggaegetgt ggtttetgag eeagggggeg acteagatee
                                                                   600
cagccagtgg acttagcccc tgtttgctcc tccgataact ggggtgacct tggttaatat
tcaccagcag cctcccccgt tgcccctctg gatccactgc ttaaatacgg acgaggacag
                                                                   720
ggccctgtct cctcagcttc aggcaccacc actgacctgg gacagtgaat agcgagatct
agettacttg tggtaccage teggateetg agaactteag ggtgagteta tgggaceett
                                                                   840
gatgttttct ttccccttct tttctatggt taagttcatg tcataggaag gggagaagta
acagggtaca catattgacc aaatcagggt aattttgcat ttgtaatttt aaaaaatgct
                                                                   960
ttettettt aatataettt tttgtttate ttatttetaa taettteeet aatetette
tttcaqqqca ataatqatac aatqtatcat qcctctttqc accattctaa aqaataacaq
tgataatttc tgggttaagg caatagcaat atttctgcat ataaatattt ctgcatataa
                                                                   1140
attgtaactg atgtaagagg tttcatattg ctaatagcag ctacaatcca gctaccattc
                                                                   1200
tgcttttatt ttatggttgg gataaggctg gattattctg agtccaagct aggccctttt
                                                                   1260
gctaatcatg ttcatacctc ttatcttcct cccacagctc ctgggcaacg tgctggtctg
tgtgctggcc catcactttg gcaaagaatt gatctcgagt aactgaaggc ggccgccacc
                                                                   1380
atgagcagca agggetetgt ggtgetggee tactetggeg geetggacae cagetgtate
                                                                   1440
ctcgtgtggc tgaaagaaca gggctacgac gtgatcgcct acctggccaa catcggccag
                                                                   1500
aaagaggact tcgaggaagc ccggaagaag gccctgaagc tgggcgccaa gaaggtgttc
                                                                   1560
atcgaggacg tgtcccgcga gttcgtggaa gagttcatct ggcccgccat ccagagcagc
gccctgtacg aggacagata cctgctgggc accagcctgg ccagaccctg tatcgcccgg
                                                                   1680
aaacaggtgg aaatcgccca gcgcgagggc gccaaatacg tgtctcacgg cgccaccggc
                                                                   1740
aagggcaacg accaggtgcg ctttgagctg agctgctact ccctggcccc ccagatcaaa
                                                                   1800
gtgatcgccc cttggcggat gcccgagttc tacaaccggt tcaagggccg gaacgacctg
```

atggaatacg ccaagcagca cggcatcccc atccccgtga cccccaagaa cccttggagc

```
atggacgaga acctgatgca catcagctac gaggccggca tcctggaaaa ccccaagaat
caggecete eeggeetgta cacaaagace eaggaceetg eeaaggeeee eaacaceee
                                                                      2040
gacattetgg aaategagtt caagaaagge gtgeeegtga aagtgaeeaa egtgaaggae
                                                                      2100
ggcaccaccc accagacctc cctggaactg ttcatgtacc tgaacgaggt ggccggcaag
                                                                      2160
cacggcgtgg gcagaatcga catcgtggaa aacagattca tcggcatgaa gtcccggggc
                                                                      2220
atctacgaga caccagccgg caccatcctg taccacgccc acctggatat cgaggccttc
                                                                      2280
accatggacc gggaagtgcg gaagatcaag cagggcctgg gcctgaagtt cgccgagctg
                                                                      2340
gtgtacacag gcttttggca cagccccgag tgcgagtttg tgcggcactg cattgccaag
                                                                      2400
agccaggaac gggtggaagg caaggtgcag gtgtccgtgc tgaagggcca ggtgtacatt
                                                                      2460
ctgggcagag agagccccct gagcctgtac aacgaggaac tggtgtctat gaacgtgcag
                                                                      2520
ggcgactacg agcccaccga cgccaccggc ttcatcaaca tcaacagcct gagactgaaa
                                                                      2580
gagtaccacc ggctgcagtc caaagtgacc gccaagtgat aagcatgcgg atctgcctcg
                                                                      2640
actgtgcctt ctagttgcca gccatctgtt gtttgcccct ccccgtgcc ttccttgacc
ctggaaggtg ccactcccac tgtcctttcc taataaaatg aggaaattgc atcgcattgt
                                                                      2760
ctgagtaggt gtcattctat tctggggggt ggggtggggc aggacagcaa gggggaggat
                                                                      2820
tgggaagaca atagcaggca tgctggggac tcgagttaag ggcgaattcc cgataaggat
                                                                      2880
cttcctagag catggctacg tagataagta gcatggcggg ttaatcatta actacaagga
acceptagtg atggagttgg coactcocte tetgegeget egetegetea etgaggeegg
gcgaccaaag gtcgcccgac gcccgggctt tgcccgggcg gcctcagtga gcgagcgagc
gcgcagcctt aattaaccta attcactggc cgtcgtttta caacgtcgtg actgggaaaa
ccctggcgtt acccaactta atcgccttgc agcacatccc cctttcgcca gctggcgtaa
tagogaagag gooogcacog atogoootto coaacagttg ogcagootga atggogaatg
                                                                      3240
ggacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac
                                                                      3300
cgctacactt gccagcgccc tagcgcccgc tcctttcgct ttcttccctt cctttctcgc
                                                                      3360
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt
                                                                      3420
tagtgettta eggeaceteg acceeaaaaa acttgattag ggtgatggtt eacgtagtgg
                                                                      3480
gccatcgccc tgatagacgg tttttcgccc tttgacgttg gagtccacgt tctttaatag
                                                                      3540
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt
                                                                      3600
                                                                      3660
taacgcgaat titaacaaaa tattaacgct tacaatitag gtggcacttt tcgggggaaat gtgcgcggaa cccctatttg titattitic taaatacatt caaatatgta tccgctcatg
                                                                      3720
                                                                      3780
agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt gccttcctgt ttttgctcac
                                                                      3840
                                                                      3900
ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac
                                                                      3960
atcgaactgg atctcaacag cggtaagatc cttgagagtt ttcgccccga agaacgtttt
                                                                      4020
ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccg tattgacgcc
                                                                      4080
gggcaagagc aactcggtcg ccgcatacac tattctcaga atgacttggt tgagtactca
                                                                      4140
ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgctgcc
                                                                      4200
ataaccatga gtgataacac tgcggccaac ttacttctga caacgatcgg aggaccgaag
                                                                      4260
gagetaaccg cttttttgca caacatgggg gatcatgtaa ctcgccttga tcgttgggaa
                                                                      4320
ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg
                                                                      4380
gcaacaacgt tgcgcaaact attaactggc gaactactta ctctagcttc ccggcaacaa
                                                                      4440
ttaatagact ggatggagge ggataaagtt gcaggaccac ttctgcgctc ggcccttccg
                                                                      4500
gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt
                                                                      4560
gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt
                                                                      4620
caggcaacta tggatgaacg aaatagacag atcgctgaga taggtgcctc actgattaag
                                                                      4680
cattggtaac tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat
                                                                      4740
ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct
                                                                      4800
taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct
                                                                      4860
tgagateett tttttetgeg egtaatetge tgettgeaaa caaaaaaace acegetacea
                                                                      4920
geggtggttt gtttgeegga teaagageta eeaactettt tteegaaggt aactggette
                                                                      4980
agcagagege agataceaaa tactgttett etagtgtage egtagttagg ceaceaette
                                                                      5040
aagaactetg tagcacegee tacatacete getetgetaa teetgttace agtggetget
                                                                      5100
gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag
                                                                      5160
gcgcagcggt cgggctgaac ggggggttcg tgcacacagc ccagcttgga gcgaacgacc
                                                                      5220
tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg
                                                                      5280
agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa caggagagcg cacgagggag
                                                                      5340
cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt
gagegtegat ttttgtgatg etegteaggg gggeggagee tatggaaaaa egeeageaac
geggeetttt taeggtteet ggeettttge tggeettttg eteacatgtt ettteetgeg
ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga taccgctcgc
cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga gcgcccaata
cgcaaaccgc ctctcccgc gcgttggccg attcattaat gcagctggca cgacaggttt
cccgactgga aagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag
qcaccccaqq ctttacactt tatqcttccq qctcqtatqt tqtqtqqaat tqtqaqcqqa
                                                                      5820
taacaatttc acacaggaaa cagctatgac catgattacg ccagatttaa ttaagg
                                                                      5876
SEQ ID NO: 30
                        moltype = DNA length = 4893
                       Location/Qualifiers
FEATURE
misc_feature
                        1..4893
                        note = constructed sequence
                        1..4893
source
                        mol_type = other DNA
                        organism = synthetic construct
SEQUENCE: 30
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt 60
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact 120
aggggtteet tgtagttaat gattaacceg ceatgetaet tatetaetta agtgtttget
gettgeaatg tttgeecatt ttaggggaat tetggacaca ggacgetgtg gtttetgage 240
```

cagggggcga	ctcagatccc	agccagtgga	cttagcccct	gtttgctcct	ccgataactg	300
agataacctt	ggttaatatt	caccagcagc	ctcccccatt	gcccctctgg	atccactact	360
						420
				ggcaccacca		
acagtgaata	gcggccgcca	ccatgagcag	caagggctct	gtggtgctgg	cctactctgg	480
caacctaaac	accadetata	tecteatata	actaaaaaaa	cagggctacg	acqtqatcqc	540
						600
				gcccggaaga		
getgggegee	aagaaggtgt	tcatcgagga	cgtgtcccgc	gagttcgtgg	aagagttcat	660
ctqqcccqcc	atccaqaqca	qcqccctqta	cqaqqacaqa	tacctgctgg	qcaccaqcct	720
				cagcgcgagg		780
				cgctttgagc		840
ctccctggcc	ccccagatca	aagtgatcgc	cccttggcgg	atgcccgagt	tctacaaccg	900
gttcaagggc	cggaacgacc	tgatggaata	caccaaacaa	cacggcatcc	ccatccccqt	960
				cacatcagct		1020
catcctggaa	aaccccaaga	atcaggcccc	teeeggeetg	tacacaaaga	cccaggaccc	1080
tgccaaggcc	cccaacaccc	ccgacattct	ggaaatcgag	ttcaagaaag	gcgtgcccgt	1140
				tccctggaac		1200
						1260
				gacatcgtgg		
catcggcatg	aagtcccggg	gcatctacga	gacaccagcc	ggcaccatcc	tgtaccacgc	1320
ccacctqqat	atcqaqqcct	tcaccatqqa	ccqqqaaqtq	cggaagatca	aqcaqqqcct	1380
				cacageeeeg		1440
				ggcaaggtgc		1500
gctgaagggc	caggtgtaca	ttctgggcag	agagagcccc	ctgagcctgt	acaacgagga	1560
actoototct	atgaacgtgc	aggggagta	cdadcccacc	gacgccaccg	gcttcatcaa	1620
						1680
				tccaaagtga		
ataagcatgc	ggatctgcct	cgactgtgcc	ttctagttgc	cagccatctg	ttgtttgccc	1740
ctcccccqtq	ccttccttqa	ccctqqaaqq	tqccactccc	actgtccttt	cctaataaaa	1800
				attctggggg		1860
gcaggacagc	aagggggagg	attgggaaga	caatagcagg	catgctgggg	actcgagtag	1920
ataagtagca	tggcgggtta	atcattaact	acaaggaacc	cctagtgatg	gagttggcca	1980
				accaaaggtc		2040
						2100
				cagccttaat		
cactggccgt	cgttttacaa	cgtcgtgact	gggaaaaccc	tggcgttacc	caacttaatc	2160
gccttgcagc	acatccccct	ttcgccagct	ggcgtaatag	cgaagaggcc	cgcaccgatc	2220
				cgcgccctgt		2280
-						2340
				tacacttgcc		
cgcccgctcc	tttcgctttc	ttcccttcct	ttctcgccac	gttcgccggc	tttccccgtc	2400
aaqctctaaa	tcqqqqqctc	cctttagggt	tccgatttag	tgctttacgg	cacctcgacc	2460
				atcgccctga		2520
				actcttgttc		2580
caacactcaa	ccctatctcg	gtctattctt	ttgatttata	agggattttg	ccgatttcgg	2640
cctattqqtt	aaaaaatqaq	ctqatttaac	aaaaatttaa	cgcgaatttt	aacaaaatat	2700
				cgcggaaccc		2760
						2820
				caataaccct		
tcaataatat	tgaaaaagga	agagtatgag	tattcaacat	ttccgtgtcg	cccttattcc	2880
cttttttgcg	gcattttgcc	ttcctgtttt	tgctcaccca	gaaacgctgg	tgaaagtaaa	2940
				gaactggatc		3000
						3060
				atgatgagca		
tctgctatgt	ggcgcggtat	tatcccgtat	tgacgccggg	caagagcaac	teggtegeeg	3120
catacactat	tctcagaatg	acttggttga	gtactcacca	gtcacagaaa	agcatcttac	3180
				accatgagtg		3240
						3300
				ctaaccgctt		
catgggggat	catgtaactc	gccttgatcg	ttgggaaccg	gagctgaatg	aagccatacc	3360
aaacgacgag	cgtgacacca	cgatgcctgt	agcaatggca	acaacgttgc	gcaaactatt	3420
aactggcgaa	ctacttactc	tagetteecg	gcaacaatta	atagactgga	t.ggaggggga	3480
						3540
				ggctggttta		
			_	gcactggggc	0 00	3600
gccctcccgt	atcgtagtta	tctacacgac	ggggagtcag	gcaactatgg	atgaacgaaa	3660
				tggtaactgt		3720
				taatttaaaa		3780
				cgtgagtttt		3840
agcgtcagac	cccgtagaaa	agatcaaagg	atcttcttga	gatccttttt	ttctgcgcgt	3900
aatctgctgc	ttqcaaacaa	aaaaaccacc	actaccaaca	gtggtttgtt	taccadatca	3960
				agagcgcaga		4020
tgttcttcta	gtgtagccgt	agttaggcca	ccacttcaag	aactctgtag	caccgcctac	4080
atacctcgct	ctgctaatcc	tgttaccagt	ggctgctgcc	agtggcgata	agtcgtgtct	4140
taccaaatta	gactcaagac	gatagttacc	agataagaca	cagcggtcgg	actaaacaaa	4200
	_		_	accgaactga	_	4260
gcgtgagcta	tgagaaagcg	ccacgcttcc	cgaagggaga	aaggcggaca	ggtatccggt	4320
				ccagggggaa		4380
						4440
_		_		cgtcgatttt		
gtcagggggg	cggagcctat	ggaaaaacgc	cagcaacgcg	gcctttttac	ggttcctggc	4500
				tcccctgatt		4560
				agccgaacga		4620
cgagtcagtg	agcgaggaag	cggaagagcg	cccaatacgc	aaaccgcctc	tccccgcgcg	4680
				gactggaaag		4740
						4800
				ccccaggctt		
gcttccggct	cgtatgttgt	gtggaattgt	gagcggataa	caatttcaca	caggaaacag	4860
ctatgaccat	gattacgcca	gatttaatta	agg			4893
J	5 - 5 - 7					

-continued

moltype = DNA length = 4931 SEQ ID NO: 31 FEATURE Location/Qualifiers misc\_feature 1..4931 note = constructed sequence 1..4931 source

mol\_type = other DNA organism = synthetic construct SEQUENCE: 31 ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact 120 aggggtteet tgtagttaat gattaaeceg ceatgetaet tatetaetta agetaeeteg tgatcgcccg gcccctgttc aaacatgtcc taatactctg tctctgcaag ggtcatcagt 240 agttttccat cttactcaac atcctcccag tggaattcat ttcatagaac gaatgttccg atgetetaat etetetagae aaggtteata titgtatggg ttaettatte tetetitgti 360 gactaagtca ataatcagaa tcagcaggtt tgcagtcaga ttggcaggga taagcagcct ageteaggag aagtgagtat aaaageeeca ggetgggage ageeateage ggeegeeace atgagcagca agggctctgt ggtgctggcc tactctggcg gcctggacac cagctgtatc ctcgtgtggc tgaaagaaca gggctacgac gtgatcgcct acctggccaa catcggccag aaagaggact tcgaggaagc ccggaagaag gccctgaagc tgggcgccaa gaaggtgttc atcgaggacg tgtcccgcga gttcgtggaa gagttcatct ggcccgccat ccagagcagc geoctgtacg aggacagata cetgetggge accageetgg ceagaceetg tategeeegg aaacaggtgg aaatcgccca gcgcgagggc gccaaatacg tgtctcacgg cgccaccggc 840 aaqqqcaacq accaqqtqcq ctttqaqctq aqctqctact ccctqqcccc ccaqatcaaa gtgatcgccc cttggcggat gcccgagttc tacaaccggt tcaagggccg gaacgacctg 960 atggaatacg ccaagcagca cggcatcccc atccccgtga cccccaagaa cccttggagc 1020 atggacgaga acctgatgca catcagctac gaggccggca tcctggaaaa ccccaagaat 1080 caggoccete coggoctgta cacaaagace caggaceetg ccaaggocce caacaceeee 1140 gacattctgg aaatcgagtt caagaaaggc gtgcccgtga aagtgaccaa cgtgaaggac 1200 ggcaccaccc accagacctc cctggaactg ttcatgtacc tgaacgaggt ggccggcaag 1260 cacggcgtgg gcagaatcga catcgtggaa aacagattca tcggcatgaa gtcccggggc 1320 atctacgaga caccageegg caccateetg taccaegeee acetggatat egaggeette 1380 accatggace gggaagtgeg gaagateaag cagggeetgg geetgaagtt egeegagetg gtgtacacag gettttggea cageecegag tgegagtttg tgeggeactg cattgecaag 1440 1500 agccaggaac gggtggaagg caaggtgcag gtgtccgtgc tgaagggcca ggtgtacatt 1560 ctgggcagag agagcccct gagcctgtac aacgaggaac tggtgtctat gaacgtgcag 1620 ggcgactacg agcccaccga cgccaccggc ttcatcaaca tcaacagcct gagactgaaa 1680 gagtaccacc ggctgcagtc caaagtgacc gccaagtgat aagcatgcgg atctgcctcg 1740 actgtgcctt ctagttgcca gccatctgtt gtttgcccct cccccgtgcc ttccttgacc 1800 ctggaaggtg ccactcccac tgtcctttcc taataaaatg aggaaattgc atcgcattgt 1860 ctgagtaggt gtcattctat tctggggggt ggggtggggc aggacagcaa gggggaggat 1920 tgggaagaca atagcaggca tgctggggac tcgagtagat aagtagcatg gcgggttaat 1980 cattaactac aaggaacccc tagtgatgga gttggccact ccctctctgc gcgctcgctc 2040 geteactgag geegggegae caaaggtege eegaegeeeg ggetttgeee gggeggeete 2100 agtgagcgag cgagcgcgca gccttaatta acctaattca ctggccgtcg ttttacaacg 2160 tegtgactgg gaaaaccetg gegttaceca acttaatege ettgeageae atceceettt 2220 cgccagctgg cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgcgcag 2280 cctgaatggc gaatgggacg cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt 2340 tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt tcgctttctt 2400 cccttccttt ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc gggggctccc 2460 tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg attagggtga 2520 tggttcacgt agtgggccat cgccctgata gacggttttt cgccctttga cgttggagtc 2580 cacgttcttt aatagtggac tcttgttcca aactggaaca acactcaacc ctatctcggt 2640 ctattctttt gatttataag ggattttgcc gatttcggcc tattggttaa aaaatgagct 2700 gatttaacaa aaatttaacg cgaattttaa caaaatatta acgcttacaa tttaggtggc 2760 acttttcggg gaaatgtgcg cggaacccct atttgtttat ttttctaaat acattcaaat 2820 atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggaag 2880 agtatgagta ttcaacattt ccgtgtcgcc cttattccct tttttgcggc attttgcctt 2940 cctgtttttg ctcacccaga aacgctggtg aaagtaaaag atgctgaaga tcagttgggt gcacgagtgg gttacatcga actggatctc aacagcggta agatccttga gagttttcgc 3060 cccgaagaac gttttccaat gatgagcact tttaaagttc tgctatgtgg cgcggtatta 3120 tcccgtattg acgccgggca agagcaactc ggtcgccgca tacactattc tcagaatgac ttggttgagt actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat aacactgcgg ccaacttact tctgacaacg atcggaggac cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttqatcqtt qqqaaccqqa qctqaatqaa qccataccaa acqacqaqcq tqacaccacq atgcctgtag caatggcaac aacgttgcgc aaactattaa ctggcgaact acttactcta 3480 gcttcccggc aacaattaat agactggatg gaggcggata aagttgcagg accacttctg 3540 cgctcggccc ttccggctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg 3600 tctcgcggta tcattgcagc actggggcca gatggtaagc cctcccgtat cgtagttatc tacacgacgg ggagtcaggc aactatggat gaacgaaata gacagatcgc tgagataggt 3720 gcctcactga ttaagcattg gtaactgtca gaccaagttt actcatatat actttaqatt 3780 gatttaaaac ttcatttta atttaaaagg atctaggtga agatcctttt tgataatctc 3840 atgaccaaaa tcccttaacg tgagttttcg ttccactgag cgtcagaccc cgtagaaaag 3900 atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggt ggtttgtttg ccggatcaag agctaccaac tctttttccg 4020 aaggtaactg gcttcagcag agcgcagata ccaaatactg ttcttctagt gtagccgtag 4080 ttaggccacc acttcaagaa ctctgtagca ccgcctacat acctcgctct gctaatcctg 4140 ttaccagtgg ctgctgccag tggcgataag tcgtgtctta ccgggttgga ctcaagacga 4200 tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc

```
ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagctatg agaaagcgcc
acgetteccg aagggagaaa ggeggacagg tateeggtaa geggeagggt eggaacagga
                                                                  4380
gagegeaega gggagettee agggggaaae geetggtate tttatagtee tgtegggttt
                                                                  4440
cgccacctct gacttgagcg tcgatttttg tgatgctcgt caggggggcg gagcctatgg
                                                                  4500
aaaaacgcca gcaacgcggc ctttttacgg ttcctggcct tttgctggcc ttttgctcac
                                                                  4560
atgttctttc ctgcgttatc ccctgattct gtggataacc gtattaccgc ctttgagtga
                                                                  4620
gctgataccg ctcgccgcag ccgaacgacc gagcgcagcg agtcagtgag cgaggaagcg
                                                                  4680
gaagagegee caataegeaa acegeetete eeegegegtt ggeegattea ttaatgeage
                                                                  4740
tggcacgaca ggtttcccga ctggaaagcg ggcagtgagc gcaacgcaat taatgtgagt
                                                                  4800
tageteacte attaggeace ecaggettta caetttatge tteeggeteg tatgttgtgt
                                                                  4860
ggaattgtga gcggataaca atttcacaca ggaaacagct atgaccatga ttacgccaga
                                                                  4920
tttaattaaq q
                                                                  4931
SEQ ID NO: 32
                      moltype = DNA length = 6031
FEATURE
                      Location/Qualifiers
misc_feature
                      1..6031
                      note = constructed sequence
                      1..6031
source
                      mol type = other DNA
                      organism = synthetic construct
SEOUENCE: 32
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt
qqtcqcccqq cctcaqtqaq cqaqcqaqcq cqcaqaqaqq qaqtqqccaa ctccatcact
aggggtteet tgtagttaat gattaaceeg ceatgeteet tatetacgta gecatgetet aggaagateg gaattegeee ttaagetagg agttaatttt taaaaageag teaaaagtee
aagtgeeett gegageattt actetetetg tttgetetgg ttaataatet caggageaca
                                                                  300
aacatteett aetagtteta ggagttaatt tttaaaaaage agteaaaagt eeaagtgeee
                                                                  360
ttgcgagcat ttactctctc tgtttgctct ggttaataat ctcaggagca caaacattcc
                                                                  420
ttactagttc tagagcggcc gccagtgtgc tggaattcgg cttttttagg gctggaagct
                                                                  480
acctttgaca tcatttcctc tgcgaatgca tgtataattt ctacagaacc tattagaaag
                                                                  540
gatcacccag cctctgcttt tgtacaactt tcccttaaaa aactgccaat cccactgctg
                                                                  600
tttqqcccaa taqtqaqaac tttttcctqc tqcctcttqq tqcttttqcc tatqqcccct
                                                                  660
attetgeetg etgaagacae tettgeeage atggaettaa acceeteeag etetgaeaat
                                                                  720
                                                                  780
cctctttctc ttttgtttta catgaagggt ctggcagcca aagcaatcac tcaaagttca
aacettatca ttttttgctt tgttcctctt ggccttggtt ttgtacatca gctttgaaaa
                                                                  840
taccatecea gggttaatge tggggttaat ttataactga gagtgeteta gttetgeaat
                                                                  900
acaggacatg ctataaaaat ggaaagatgt tgctttctga gagatcagct tacttgtggt
                                                                  960
accgageteg gateetgaga actteagggt gagtetatgg gaccettgat gttttettte
                                                                  1020
cccttctttt ctatggttaa gttcatgtca taggaagggg agaagtaaca gggtacacat
                                                                  1080
attgaccaaa tcagggtaat tttgcatttg taattttaaa aaatgctttc ttcttttaat
                                                                  1140
1200
atgatacaat gtatcatgcc tctttgcacc attctaaaga ataacagtga taatttctgg
                                                                  1260
gttaaggcaa tagcaatatt tctgcatata aatatttctg catataaatt gtaactgatg
                                                                  1320
1380
tggttgggat aaggetggat tattetgagt ceaagetagg ceettttget aateatgtte
                                                                  1440
atacetetta tetteeteee acageteetg ggcaaegtge tggtetgtgt getggeeeat
                                                                  1500
cactttggca aagaattgat ctcgaggccg ccaccatgtc cagcaaaggc tccgtggttc
                                                                  1560
tggcctacag tggcggcctg gacacctcgt gcatcctcgt gtggctgaag gaacaaggct
                                                                  1620
atgacgtcat tgcctatctg gccaacattg gccagaagga agacttcgag gaagccagga
                                                                  1680
agaaggcact gaagettggg gccaaaaagg tgttcattga ggatgtcagc agggagtttg
                                                                  1740
tggaggagtt catctggccg gccatccagt ccagcgcact gtatgaggac cgctacctcc
                                                                  1800
tgggcacctc tcttgccagg ccctgcatcg cccgcaaaca agtggaaatc gcccagcggg
                                                                  1860
agggggccaa gtatgtgtcc cacggcgcca caggaaaggg gaacgatcag gtccggtttg
                                                                  1920
ageteagetg ctacteactg geoccecaga taaaggteat tgeteeetgg aggatgeetg
                                                                  1980
aattotacaa coggttcaag ggoogcaatg acotgatgga gtacgcaaag caacacggga
                                                                  2040
ttcccatccc ggtcactccc aagaacccgt ggagcatgga tgagaacctc atgcacatca
                                                                  2100
gctacgaggc tggaatcctg gagaacccca agaaccaagc gcctccaggt ctctacacga
                                                                  2160
agacccagga cccagccaaa gcccccaaca cccctgacat tctcgagatc gagttcaaaa
aaggggtccc tgtgaaggtg accaacgtca aggatggcac cacccaccag acctccttgg
agetetteat gtacetgaac gaagtegegg geaageatgg egtgggeegt attgacateg
tggagaaccg cttcattgga atgaagtccc gaggtatcta cgagacccca gcaggcacca
tcctttacca cgctcattta gacatcgagg ccttcaccat ggaccgggaa gtgcgcaaaa
tcaaacaagg cctgggcttg aaatttgctg agctggtgta taccggtttc tggcacagcc
                                                                  2520
ctgagtgtga atttgtccgc cactgcatcg ccaagtccca ggagcgagtg gaagggaaag
tgcaqqtqtc cqtcctcaaq qqccaqqtqt acatcctcqq ccqqqaqtcc ccactqtctc
                                                                  2640
tctacaatga ggagctggtg agcatgaacg tgcagggtga ttatgagcca actgatgcca
                                                                  2700
ccgggttcat caacatcaat tccctcaggc tgaaggaata tcatcgtctc cagagcaagg
                                                                  2760
tcactgccaa atgataagca tgcggatctg cctcgactgt gccttctagt tgccagccat
                                                                  2820
ctgttgtttg cccctcccc gtgccttcct tgaccctgga aggtgccact cccactgtcc
tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat tctattctgg
                                                                  2940
ggggtggggt ggggcaggac agcaaggggg aggattggga agacaatagc aggcatgctg
                                                                  3000
gggactcgag ttaagggcga attcccgata aggatcttcc tagagcatgg ctacgtagat
                                                                  3060
aagtagcatg gcgggttaat cattaactac aaggaacccc tagtgatgga gttggccact
                                                                  3120
ccctctctgc gcgctcgctc gctcactgag gccgggcgac caaaggtcgc ccgacgcccg
ggctttgccc gggcggcctc agtgagcgag cgagcgcgca gccttaatta acctaattca
ctggccgtcg ttttacaacg tcgtgactgg gaaaaccctg gcgttaccca acttaatcgc
cttgcagcac atcccccttt cgccagctgg cgtaatagcg aagaggcccg caccgatcgc
                                                                  3360
ccttcccaac agttgcgcag cctgaatggc gaatgggacg cgccctgtag cggcgcatta
agegeggegg gtgtggtggt taegegeage gtgacegeta caettgeeag egeeetageg 3480
```

## -continued

```
cocgctcctt togctttctt cocttccttt ctcgccacgt tcgccggctt tccccgtcaa
getetaaate gggggeteee tttagggtte egatttagtg etttaeggea eetegaeeee
                                                                    3600
aaaaaacttg attagggtga tggttcacgt agtgggccat cgccctgata gacggttttt
                                                                    3660
cgccctttga cgttggagtc cacgttcttt aatagtggac tcttgttcca aactggaaca
                                                                    3720
acactcaacc ctatctcggt ctattctttt gatttataag ggattttgcc gatttcggcc
                                                                    3780
tattggttaa aaaatgagct gatttaacaa aaatttaacg cgaattttaa caaaatatta
                                                                    3840
acgettacaa tttaggtgge actttteggg gaaatgtgeg eggaaceeet atttgtttat
                                                                    3900
ttttctaaat acattcaaat atgtatccgc tcatgagaca ataaccctga taaatgcttc
                                                                    3960
aataatattg aaaaaggaag agtatgagta ttcaacattt ccgtgtcgcc cttattccct
                                                                    4020
tttttgcggc attttgcctt cctgtttttg ctcacccaga aacgctggtg aaagtaaaag
                                                                    4080
atgctgaaga tcagttgggt gcacgagtgg gttacatcga actggatctc aacagcggta
                                                                    4140
agateettga gagttttege eeegaagaac gtttteeaat gatgageact tttaaagtte
                                                                    4200
tgctatgtgg cgcggtatta tcccgtattg acgccgggca agagcaactc ggtcgccgca
tacactattc tcagaatgac ttggttgagt actcaccagt cacagaaaag catcttacgg
                                                                    4320
atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat aacactgcgg
                                                                    4380
ccaacttact tctgacaacg atcggaggac cgaaggagct aaccgctttt ttgcacaaca
tgggggatca tgtaactcgc cttgatcgtt gggaaccgga gctgaatgaa gccataccaa
acgacgageg tgacaccacg atgcctgtag caatggcaac aacgttgege aaactattaa
ctggcgaact acttactcta gcttcccggc aacaattaat agactggatg gaggcggata
aagttgcagg accacttctg cgctcggccc ttccggctgg ctggtttatt gctgataaat
ctggagccgg tgagcgtggg tctcgcggta tcattgcagc actggggcca gatggtaagc
cctcccgtat cgtagttatc tacacgacgg ggagtcaggc aactatggat gaacgaaata
gacagatogo tgagataggt gootcactga ttaagcattg gtaactgtca gaccaagttt
                                                                    4860
acticatatat actitagatt gatitaaaac ticattitta atitaaaagg atctaggtga agateettit tgataatete atgaccaaaa teeettaagg tgagtitteg ticeaetgag
                                                                    4920
                                                                    4980
cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tcctttttt ctgcgcgtaa
                                                                    5040
tctgctgctt gcaaacaaaa aaaccaccgc taccagcggt ggtttgtttg ccggatcaag
                                                                    5100
agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata ccaaatactg
                                                                    5160
ttettetagt gtageegtag ttaggecace actteaagaa etetgtagea eegeetacat
                                                                    5220
acctegetet getaateetg ttaccagtgg etgetgecag tggegataag tegtgtetta
                                                                    5280
ccgggttgga ctcaagacga tagttaccgg ataaggcgca gcggtcgggc tgaacggggg
                                                                    5340
gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc
                                                                    5400
gtgagctatg agaaagcgcc acgcttcccg aagggagaaa ggcggacagg tatccggtaa
                                                                    5460
geggeagggt eggaacagga gagegeacga gggagettee agggggaaac geetggtate
                                                                    5520
tttatagtcc tgtcgggttt cgccacctct gacttgagcg tcgatttttg tgatgctcgt
                                                                    5580
caggggggg gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg ttcctggcct
                                                                    5640
tttgctggcc ttttgctcac atgttctttc ctgcgttatc ccctgattct gtggataacc
                                                                    5700
gtattaccgc ctttgagtga gctgataccg ctcgccgcag ccgaacgacc gagcgcagcg
                                                                    5760
agtcagtgag cgaggaagcg gaagagcgcc caatacgcaa accgcctctc cccgcgcgtt
                                                                    5820
ggccgattca ttaatgcagc tggcacgaca ggtttcccga ctggaaagcg ggcagtgagc
                                                                    5880
qcaacqcaat taatqtqaqt taqctcactc attaqqcacc ccaqqcttta cactttatqc
                                                                    5940
ttccggctcg tatgttgtgt ggaattgtga gcggataaca atttcacaca ggaaacagct
                                                                    6000
atgaccatga ttacgccaga tttaattaag g
                                                                    6031
                       moltype = DNA length = 6850
Location/Qualifiers
SEQ ID NO: 33
FEATURE
misc feature
                       1..6850
                       note = constructed sequence
source
                       1..6850
                       mol_type = other DNA
                       organism = synthetic construct
SEQUENCE: 33
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact
aggggttcct tgtagttaat gattaacccg ccatgctact tatctacgta gccatgctct
aggaagatcg gaattcgccc ttaagctagc ggatccaggt taatttttaa aaagcagtca
aaagtccaag tggcccttgg cagcatttac tctctctgtt tgctctggtt aataatctca
ggagcacaaa cattccagat ccaggttaat ttttaaaaaag cagtcaaaaag tccaagtggc
ccttggcagc atttactctc tctgtttgct ctggttaata atctcaggag cacaaacatt
ccagatccgg cgcgccaggg ctggaagcta cctttgacat catttcctct gcgaatgcat
gtataatttc tacagaacct attagaaagg atcacccagc ctctgctttt gtacaacttt
cccttaaaaa actgccaatt ccactgctgt ttggcccaat agtgagaact ttttcctgct
geetettggt gettttgeet atggeeceta ttetgeetge tgaagacaet ettgeeagea
                                                                    660
tggacttaaa cccctccagc tctgacaatc ctctttctct tttgttttac atgaagggtc
tggcagccaa agcaatcact caaagttcaa accttatcat tttttgcttt gttcctcttg
gccttggttt tgtacatcag ctttgaaaat accatcccag ggttaatgct ggggttaatt
                                                                    840
tataactaag agtgctctag ttttgcaata caggacatgc tataaaaatg gaaagatgtt
                                                                    900
gctttctgag agagcggccg cgtttgtttc cttttttaaa atacattgag tatgcttgcc
                                                                    960
ttttagatat agaaatatct gatgctgtct tcttcactaa attttgatta catgatttga
cagcaatatt gaagagteta acagccagca cgcaggttgg taagtactgg ttetttgtta
                                                                    1080
qctaqqtttt cttcttcttc atttttaaaa ctaaatagat cgacaatgct tatgatgcat
                                                                    1140
ttatgtttaa taaacactgt tcagttcatg atttggtcat gtaattcctg ttagaaaaca
                                                                    1200
ttcatctcct tggtttaaaa aaattaaaag tgggaaaaca aagaaatagc agaatatagt
                                                                    1260
gaaaaaaaat aaccacatta tttttgtttg gacttaccac tttgaaatca aaatgggaaa
caaaagcaca aacaatggcc ttatttacac aaaaagtctg attttaagat atatgacatt
                                                                    1380
tcaaggtttc agaagtatgt aatgaggtgt gtctctaatt ttttaaatta tatatcttca
                                                                    1440
atttaaagtt ttagttaaaa cataaagatt aacctttcat tagcaagctg ttagttatca
                                                                    1500
ccaaagcttt tcatggatta ggaaaaaatc attttgtctc tatgtcaaac atcttggagt
```

tgatatttgg ggaaacacaa tactcagttg agttccctag gggagaaaag caagcttaag 1620

63

```
aattgacata aagagtagga agttagctaa tgcaacatat atcactttgt tttttcacaa
ctacagtgac tttatgtatt tcccagagga aggcatacag ggaagaaatt atcccatttg
gacaaacagc atgttctcac aggaagcatt tatcacactt acttgtcaac tttctagaat
                                                                   1800
caaatctagt agctgacagt accaggatca ggggtgccaa ccctaagcac ccccagaaag
                                                                   1860
ctgactggcc ctgtggttcc cactccagac atgatgtcag ctggaccata attaggcttc
                                                                   1920
tgttcttcag gagacatttg ttcaaagtca tttgggcaac catattctga aaacagccca
                                                                   1980
gccagggtga tggatcactt tgcaaagatc ctcaatgagc tattttcaag tgatgacaaa
                                                                   2040
gtgtgaagtt aaccgctcat ttgagaactt tctttttcat ccaaagtaaa ttcaaatatg
                                                                   2100
attagaaatc tgacctttta ttactggaat tctcttgact aaaagtaaaa ttgaatttta
                                                                   2160
attectaaat eteeatgtgt atacagtaet gtgggaacat eacagatttt ggeteeatge
                                                                   2220
cctaaagaga aattggcttt cagattattt ggattaaaaa caaagacttt cttaagagat
                                                                   2280
gtaaaatttt catgatgttt tcttttttgc taaaactaaa gaattattct tttacatttc
                                                                   2340
agatggccgc caccatgagc agcaagggct ctgtggtgct ggcctactct ggcggcctgg
acaccagetg tateetegtg tggetgaaag aacagggeta egaegtgate geetaeetgg
                                                                   2460
ccaacatcgg ccagaaagag gacttcgagg aagcccggaa gaaggccctg aagctgggcg
                                                                   2520
ccaagaaggt gttcatcgag gacgtgtccc gcgagttcgt ggaagagttc atctggcccg
ccatccagag cagcgccctg tacgaggaca gatacctgct gggcaccagc ctggccagac
cctgtatcgc ccggaaacag gtggaaatcg cccagcgcga gggcgccaaa tacgtgtctc
acggcgccac cggcaagggc aacgaccagg tgcgctttga gctgagctgc tactccctgg
cccccagat caaagtgatc gccccttggc ggatgcccga gttctacaac cggttcaagg
geoggaacga cetgatggaa tacgecaage ageacggeat ceccateece gtgaccecea
agaaccettg gagcatggac gagaacctga tgcacatcag ctacgaggcc ggcatcctgg
aaaaccccaa qaatcaqqcc cctccqqcc tqtacacaaa qacccaqqac cctqccaaqq
                                                                   3000
ccccaacac ccccgacatt ctggaaatcg agttcaagaa aggcgtgccc gtgaaagtga
                                                                   3060
ccaacgtgaa ggacggcacc acccaccaga cctccctgga actgttcatg tacctgaacg
                                                                   3120
aggtggccgg caagcacggc gtgggcagaa tcgacatcgt ggaaaacaga ttcatcggca
                                                                   3180
tgaagtcccg gggcatctac gagacaccag ccggcaccat cctgtaccac gcccacctgg
                                                                   3240
agttcgccga gctggtgtac acaggctttt ggcacagccc cgagtgcgag tttgtgcggc
                                                                   3300
                                                                   3360
actgoattge caagagecag gaacgggtgg aaggcaaggt geaggtgtee gtgetgaagg
gecaggtgta cattetggge agagagagee eeetgageet gtacaacgag gaactggtgt
                                                                   3420
                                                                   3480
ctatqaacqt qcaqqqqac tacqaqcca ccqacqccac cqqcttcatc aacatcaaca
                                                                   3540
gcctgagact gaaagagtac caccggctgc agtccaaagt gaccgccaag tgataagcat
                                                                   3600
geggatetge etegactgtg cettetagtt gecagecate tgttgtttge eeeteceeeg
                                                                   3660
tgccttcctt gaccctggaa ggtgccactc ccactgtcct ttcctaataa aatgaggaaa
                                                                   3720
3780
gcaaggggga ggattgggaa gacaatagca ggcatgctgg ggactcgagt taagggcgaa
                                                                   3840
ttcccgataa ggatcttcct agagcatggc tacgtagata agtagcatgg cgggttaatc
                                                                   3900
attaactaca aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg
                                                                   3960
ctcactgagg ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca
                                                                   4020
gtgagcgagc gagcgcgcag ccttaattaa cctaattcac tggccgtcgt tttacaacgt
                                                                   4080
cgtgactggg aaaaccctgg cgttacccaa cttaatcgcc ttgcagcaca tccccctttc
                                                                   4140
gecagetgge gtaatagega agaggeeege acegategee etteecaaca gttgegeage
                                                                   4200
ctgaatggcg aatgggacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtgtt
                                                                   4260
acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc
                                                                   4320
cetteettte tegecaegtt egeeggettt eeeegteaag etetaaateg ggggeteeet
                                                                   4380
ttagggttcc gatttagtgc tttacggcac ctcgacccca aaaaacttga ttagggtgat
                                                                   4440
ggttcacgta gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc
                                                                   4500
acgttettta atagtggaet ettgtteeaa aetggaacaa eaeteaaece tateteggte
                                                                   4560
tattettttg atttataagg gattttgeeg attteggeet attggttaaa aaatgagetg
                                                                   4620
atttaacaaa aatttaacgc gaattttaac aaaatattaa cgcttacaat ttaggtggca
                                                                   4680
cttttcgggg aaatgtgcgc ggaaccccta tttgtttatt tttctaaata cattcaaata
                                                                   4740
tgtatccgct catgagacaa taaccctgat aaatgcttca ataatattga aaaaggaaga
                                                                   4800
gtatgagtat tcaacatttc cgtgtcgccc ttattccctt ttttgcggca ttttgccttc
                                                                   4860
ctgtttttgc tcacccagaa acgctggtga aagtaaaaga tgctgaagat cagttgggtg
                                                                   4920
cacgagtggg ttacatcgaa ctggatctca acagcggtaa gatccttgag agttttcgcc
                                                                   4980
ccgaagaacg ttttccaatg atgagcactt ttaaagttct gctatgtggc gcggtattat
                                                                   5040
cccgtattga cgccgggcaa gagcaactcg gtcgccgcat acactattct cagaatgact
tggttgagta ctcaccagtc acagaaaagc atcttacgga tggcatgaca gtaagagaat
tatgcagtgc tgccataacc atgagtgata acactgcggc caacttactt ctgacaacga
toggaggaco gaaggagota acogottttt tgcacaacat gggggatcat gtaactogco
ttgatcgttg ggaaccggag ctgaatgaag ccataccaaa cgacgagcgt gacaccacga
tgcctgtagc aatggcaaca acgttgcgca aactattaac tggcgaacta cttactctag
cttcccggca acaattaata gactggatgg aggcggataa agttgcagga ccacttctgc
gctcqqccct tccqqctqqc tqqtttattq ctqataaatc tqqaqccqqt qaqcqtqqqt
                                                                   5520
ctcgcggtat cattgcagca ctggggccag atggtaagcc ctcccgtatc gtagttatct
                                                                   5580
acacgacggg gagtcaggca actatggatg aacgaaatag acagatcgct gagataggtg
                                                                   5640
cctcactgat taagcattgg taactgtcag accaagttta ctcatatata ctttagattg
                                                                   5700
atttaaaact tcatttttaa tttaaaagga tctaggtgaa gatccttttt gataatctca
tgaccaaaat cccttaacgt gagttttcgt tccactgagc gtcagacccc gtagaaaaga
                                                                   5820
tcaaaggatc ttcttgagat ccttttttc tgcgcgtaat ctgctgcttg caaacaaaaa
                                                                   5880
aaccaccgct accagcggtg gtttgtttgc cggatcaaga gctaccaact ctttttccga
                                                                   5940
aggtaactgg cttcagcaga gcgcagatac caaatactgt tcttctagtg tagccgtagt
                                                                   6000
taggccacca cttcaagaac tctgtagcac cgcctacata cctcgctctg ctaatcctgt
taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttggac tcaagacgat
                                                                   6120
agttacegga taaggegeag eggteggget gaaegggggg ttegtgeaca eageeeaget
                                                                   6180
tggagcgaac gacctacacc gaactgagat acctacagcg tgagctatga gaaagcgcca
                                                                   6240
cgcttcccga agggagaaag gcggacaggt atccggtaag cggcagggtc ggaacaggag
                                                                   6300
agegeaegag ggagetteea gggggaaaeg cetggtatet ttatagteet gtegggttte
```

## -continued

```
gccacctctg acttgagcgt cgatttttgt gatgctcgtc aggggggggg agcctatgga
                                                                    6420
aaaacgccag caacgcggcc tttttacggt tcctggcctt ttgctggcct tttgctcaca
                                                                    6480
tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagtgag
                                                                    6540
ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg
                                                                    6600
aagagegeee aataegeaaa eegeetetee eegegegttg geegatteat taatgeaget
                                                                    6660
ggcacgacag gtttcccgac tggaaagcgg gcagtgagcg caacgcaatt aatgtgagtt
                                                                    6720
ageteactea ttaggeacce caggetttae aetttatget teeggetegt atgttgtgtg
                                                                    6780
gaattgtgag cggataacaa tttcacacag gaaacagcta tgaccatgat tacgccagat
                                                                    6840
ttaattaagg
                                                                    6850
SEQ ID NO: 34
                       moltype = DNA length = 5608
FEATURE
                       Location/Qualifiers
                       1..5608
misc_feature
                       note = constructed sequence
                       1..5608
source
                       mol type = other DNA
                       organism = synthetic construct
SEOUENCE: 34
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact
aggggttcct tgtagttaat gattaacccg ccatgctact tatctacgta gccatgctct
aggaagateg gaattegeee ttaagetace tegtgatege eeggeeeetg tteaaacatg
tectaatact etgtetetge aagggteate agtagtttte catettacte aacateetee
cagtggaatt catttcatag aacgaatgtt ccgatgctct aatctctcta gacaaggttc
                                                                    360
atatttgtat gggttactta ttctctcttt gttgactaag tcaataatca gaatcagcag
                                                                    420
gtttgcagtc agattggcag ggataagcag cctagctcag gagaagtgag tataaaagcc ccaggctggg agcagccatc agcggccgat ctagcttact tgtggtacca gctcggatcc
                                                                    480
                                                                    540
tgagaacttc agggtgagtc tatgggaccc ttgatgtttt ctttcccctt cttttctatg
                                                                    600
gttaagttca tgtcatagga aggggagaag taacagggta cacatattga ccaaatcagg
                                                                    660
gtaattttgc atttgtaatt ttaaaaaaatg ctttcttctt ttaatatact tttttgttta
                                                                    720
tettattet aataetttee etaatetett tettteaggg caataatgat acaatgtate
                                                                    780
atgeetettt geaccattet aaagaataac agtgataatt tetgggttaa ggeaatagea
                                                                    840
atatttctgc atataaatat ttctgcatat aaattgtaac tgatgtaaga ggtttcatat
                                                                    900
tgctaatagc agctacaatc cagctaccat tctgctttta ttttatggtt gggataaggc
                                                                    960
tggattattc tgagtccaag ctaggccctt ttgctaatca tgttcatacc tcttatcttc
                                                                    1020
ctcccacagc tcctgggcaa cgtgctggtc tgtgtgctgg cccatcactt tggcaaagaa
                                                                    1080
ttgatctcga qtaactgaag qcqqccqcca ccatqtccag caaaqqctcc qtqqttctqq
                                                                    1140
cctacagtgg cggcctggac acctcgtgca tcctcgtgtg gctgaaggaa caaggctatg
                                                                    1200
acgtcattgc ctatctggcc aacattggcc agaaggaaga cttcgaggaa gccaggaaga
                                                                    1260
aggcactgaa gcttggggcc aaaaaggtgt tcattgagga tgtcagcagg gagtttgtgg
                                                                    1320
aggagttcat ctggccggcc atccagtcca gcgcactgta tgaggaccgc tacctcctgg
                                                                    1380
geacetetet tgecaggece tgeategece geaaacaagt ggaaategee cagegggagg
                                                                    1440
gggccaagta tgtgtcccac ggcgccacag gaaaggggaa cgatcaggtc cggtttgagc
                                                                    1500
tcagctgcta ctcactggcc ccccagataa aggtcattgc tccctggagg atgcctgaat
                                                                    1560
tctacaaccg gttcaagggc cgcaatgacc tgatggagta cgcaaagcaa cacgggattc
                                                                    1620
ccatcccggt cactcccaag aacccgtgga gcatggatga gaacctcatg cacatcagct
                                                                    1680
acgaggetgg aateetggag aaccecaaga accaagegee tecaggtete tacaegaaga
                                                                    1740
cccaggaccc agccaaagcc cccaacaccc ctgacattct cgagatcgag ttcaaaaaag
                                                                    1800
gggtccctgt gaaggtgacc aacgtcaagg atggcaccac ccaccagacc tccttggagc
                                                                    1860
tetteatgta eetgaacgaa gtegegggea ageatggegt gggeegtatt gacategtgg
                                                                    1920
agaaccgctt cattggaatg aagtcccgag gtatctacga gaccccagca ggcaccatcc
                                                                    1980
tttaccacgc tcatttagac atcgaggcct tcaccatgga ccgggaagtg cgcaaaatca
                                                                    2040
aacaaggeet gggettgaaa tttgetgage tggtgtatae eggtttetgg cacageeetg
                                                                    2100
agtgtgaatt tgtccgccac tgcatcgcca agtcccagga gcgagtggaa gggaaagtgc
aggtgtccgt cctcaagggc caggtgtaca tcctcggccg ggagtcccca ctgtctctct
                                                                    2220
acaatgagga gctggtgagc atgaacgtgc agggtgatta tgagccaact gatgccaccg
                                                                    2280
ggttcatcaa catcaattcc ctcaggctga aggaatatca tcgtctccag agcaaggtca
                                                                    2340
ctgccaaatg ataagcatgc ggatctgcct cgactgtgcc ttctagttgc cagccatctg
ttgtttgccc ctcccccgtg ccttccttga ccctggaagg tgccactccc actgtccttt
cctaataaaa tgaggaaatt gcatcgcatt gtctgagtag gtgtcattct attctggggg
gtggggtggg gcaggacagc aagggggagg attgggaaga caatagcagg catgctgggg
actogagtta agggcgaatt cocgataagg atottoctag agcatggcta cgtagataag
tagcatggcg ggttaatcat taactacaag gaaccectag tgatggagtt ggccactece
tetetgegeg etegeteget caetgaggee gggegaceaa aggtegeeeg acgeeeggge
                                                                    2760
tttqcccqqq cqqcctcaqt qaqcqaqcqa qcqcqcaqcc ttaattaacc taattcactq
                                                                    2820
gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg ttacccaact taatcgcctt
                                                                    2880
gcagcacatc cccctttcgc cagctggcgt aatagcgaag aggcccgcac cgatcgccct
                                                                    2940
tcccaacagt tgcgcagcct gaatggcgaa tgggacgcgc cctgtagcgg cgcattaagc
                                                                    3000
geggegggtg tggtggttac gegeagegtg acceptacac ttgccagege cetagegeec
                                                                    3060
geteettteg etttetteee tteetttete gecaegtteg eeggetttee eegteaaget
                                                                    3120
ctaaatcggg ggctcccttt agggttccga tttagtgctt tacggcacct cgaccccaaa
                                                                    3180
aaacttgatt agggtgatgg ttcacgtagt gggccatcgc cctgatagac ggtttttcgc
                                                                    3240
cctttgacgt tggagtccac gttctttaat agtggactct tgttccaaac tggaacaaca
                                                                    3300
ctcaacccta tctcggtcta ttcttttgat ttataaggga ttttgccgat ttcggcctat
tggttaaaaa atgagctgat ttaacaaaaa tttaacgcga attttaacaa aatattaacg
                                                                    3420
cttacaattt aggtggcact tttcggggaa atgtgcgcgg aaccectatt tgtttatttt
                                                                    3480
tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat
                                                                    3540
aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt
                                                                    3600
```

ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg 3660

### -continued

```
ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga
teettgagag ttttegeece gaagaaegtt tteeaatgat gageaetttt aaagttetge
tatgtggege ggtattatee egtattgaeg eegggeaaga geaacteggt egeegeatae
                                                                   3840
actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg
                                                                   3900
gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca
                                                                   3960
acttacttct gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg
                                                                   4020
gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg
                                                                   4080
acgagegtga caccacgatg cetgtageaa tggcaacaac gttgegeaaa etattaactg
                                                                   4140
gcgaactact tactctagct tcccggcaac aattaataga ctggatggag gcggataaag
                                                                   4200
ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg
                                                                   4260
gagccggtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct
                                                                   4320
cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac
                                                                   4380
agatcgctga gataggtgcc tcactgatta agcattggta actgtcagac caagtttact
catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga
                                                                   4500
tcctttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt
                                                                   4560
cagaccccgt agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaatct
gctgcttgca aacaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc
taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc
ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc
tegetetget aatcetgtta ceagtggetg etgecagtgg egataagteg tgtettaceg
ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtt
cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg
agctatqaqa aaqcqccacq cttcccqaaq qqaqaaaqqc qqacaqqtat ccqqtaaqcq
                                                                   5040
gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt
                                                                   5100
atagteetgt egggtttege cacetetgae ttgagegteg atttttgtga tgetegteag
                                                                   5160
gggggggag cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt
                                                                   5220
gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta
                                                                   5280
ttaccgcctt tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt
                                                                   5340
cagtgagega ggaageggaa gagegeecaa taegeaaace geeteteece gegegttgge
                                                                   5400
cgattcatta atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca
                                                                   5460
acgcaattaa tgtgagttag ctcactcatt aggcacccca ggctttacac tttatgcttc
                                                                   5520
cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg
                                                                   5580
accatgatta cgccagattt aattaagg
                                                                   5608
SEQ ID NO: 35
                       moltype = DNA length = 5876
                       Location/Qualifiers
FEATURE
                      1..5876
misc_feature
                       note = constructed sequence
source
                      1..5876
                       mol_type = other DNA
                      organism = synthetic construct
SECUENCE: 35
ctgegegete getegeteae tgaggeegee egggeaaage eegggegteg ggegaeettt
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact
                                                                   120
aggggtteet tgtagttaat gattaaceeg ceatgetaet tatetaegta gecatgetet
                                                                   180
aggaagateg gaattegeee ttaagetage egetgacaag geteagagge acacaggagt
                                                                   240
ttctgggctc accetgeccc ettccaacce etcagttece atectccage agetgtttgt
                                                                   300
gtgctgcctc tgaagtccac actgaacaaa cttcagccta ctcatgtccc taaaatgggc
                                                                   360
aaacattgca agcagcaaac agcaaacaca cagccctccc tgcctgctga ccttggagct
                                                                   420
ggggcagagg tcagagacct ctctgggccc atgccacctc caacatccac tcgacccctt
                                                                   480
ggaatttcgg tggagaggag cagaggttgt cctggcgtgg tttaggtagt gtgagaggga
                                                                   540
gatccggcgc gcctggacac aggacgctgt ggtttctgag ccagggggcg actcagatcc
                                                                   600
cagocagtgg acttagcccc tgtttgctcc tccgataact ggggtgacct tggttaatat
                                                                   660
traccagrag cotroccogt tgcccctctg gatccactgc ttaaatacgg acgaggacag
ggccctgtct cctcagcttc aggcaccacc actgacctgg gacagtgaat agcgagatct
                                                                   780
agettacttg tggtaccage teggateetg agaactteag ggtgagteta tgggaecett
gatgttttct ttccccttct tttctatggt taagttcatg tcataggaag gggagaagta
                                                                   900
acagggtaca catattgacc aaatcagggt aattttgcat ttgtaatttt aaaaaatgct
ttettettt aatataettt tttgtttate ttatttetaa taettteeet aatetette
tttcagggca ataatgatac aatgtatcat gcctctttgc accattctaa agaataacag
tgataatttc tgggttaagg caatagcaat atttctgcat ataaatattt ctgcatataa
attgtaactg atgtaagagg tttcatattg ctaatagcag ctacaatcca gctaccattc
tgcttttatt ttatggttgg gataaggctg gattattctg agtccaagct aggccctttt
                                                                   1260
gctaatcatg ttcatacctc ttatcttcct cccacagctc ctgggcaacg tgctggtctg
tgtqctqqcc catcactttq qcaaaqaatt qatctcqaqt aactqaaqqc qqccqccacc
                                                                   1380
atgtccagca aaggctccgt ggttctggcc tacagtggcg gcctggacac ctcgtgcatc
                                                                   1440
ctcgtgtggc tgaaggaaca aggctatgac gtcattgcct atctggccaa cattggccag
                                                                   1500
aaggaagact tcgaggaagc caggaagaag gcactgaagc ttggggccaa aaaggtgttc
                                                                   1560
attgaggatg tcagcaggga gtttgtggag gagttcatct ggccggccat ccagtccagc
geactgtatg aggacegeta ceteetggge acetetettg ecaggeeetg categeeege
                                                                   1680
aaacaagtgg aaatcgccca gcgggagggg gccaagtatg tgtcccacgg cgccacagga
aaggggaacg atcaggtccg gtttgagctc agctgctact cactggcccc ccagataaag
                                                                   1800
gtcattgctc cctggaggat gcctgaattc tacaaccggt tcaagggccg caatgacctg
                                                                   1860
atggagtacg caaagcaaca cgggattccc atcccggtca ctcccaagaa cccgtggagc
atggatgaga acctcatgca catcagctac gaggctggaa tcctggagaa ccccaagaac
caagegeete caggteteta caegaagace caggacecag ccaaageeee caacacecet
gacattctcg agatcgagtt caaaaaaggg gtccctgtga aggtgaccaa cgtcaaggat
ggcaccaccc accagacctc cttggagctc ttcatgtacc tgaacgaagt cgcgggcaag
```

catggcgtgg gccgtattga catcgtggag aaccgcttca ttggaatgaa gtcccgaggt

```
atctacgaga ccccagcagg caccatcett taccacgete atttagacat cgaggeette
accatggacc gggaagtgcg caaaatcaaa caaggcctgg gcttgaaatt tgctgagctg
                                                                     2340
gtgtataccg gtttctggca cagccctgag tgtgaatttg tccgccactg catcgccaag
                                                                     2400
tcccaggagc gagtggaagg gaaagtgcag gtgtccgtcc tcaagggcca ggtgtacatc
                                                                     2460
ctcggccggg agtccccact gtctctctac aatgaggagc tggtgagcat gaacgtgcag
                                                                     2520
ggtgattatg agccaactga tgccaccggg ttcatcaaca tcaattccct caggctgaag
                                                                     2580
gaatatcatc gtctccagag caaggtcact gccaaatgat aagcatgcgg atctgcctcg
                                                                     2640
actgtgcctt ctagttgcca gccatctgtt gtttgcccct ccccgtgcc ttccttgacc
                                                                     2700
                                                                     2760
ctggaaggtg ccactcccac tgtcctttcc taataaaatg aggaaattgc atcgcattgt
ctgagtaggt gtcattctat tctggggggt ggggtggggc aggacagcaa gggggaggat
                                                                     2820
tgggaagaca atagcaggca tgctggggac tcgagttaag ggcgaattcc cgataaggat
                                                                     2880
cttcctagag catggctacg tagataagta gcatggcggg ttaatcatta actacaagga
                                                                     2940
acceptagtg atggagttgg coactcocte tetgegeget egetegetea etgaggeegg
                                                                     3000
gcgaccaaag gtcgcccgac gcccgggctt tgcccgggcg gcctcagtga gcgagcgagc
                                                                     3060
gcgcagcctt aattaaccta attcactggc cgtcgtttta caacgtcgtg actgggaaaa
                                                                     3120
ccctggcgtt acccaactta atcgccttgc agcacatccc cctttcgcca gctggcgtaa
                                                                     3180
tagogaagag gooogcacog atogocotto coaacagttg ogcagootga atggogaatg
ggacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac
cgctacactt gccagcgccc tagcgcccgc tcctttcgct ttcttccctt cctttctcgc
                                                                     3360
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt
                                                                     3420
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatggtt cacgtagtgg
gccatcgccc tgatagacgg tttttcgccc tttgacgttg gagtccacgt tctttaatag
                                                                     3540
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt
                                                                     3600
ataagggatt ttgccgattt cggcctattg gttaaaaaaat gagctgattt aacaaaaatt
                                                                     3660
taacgcgaat tttaacaaaa tattaacgct tacaatttag gtggcacttt tcggggaaat
                                                                     3720
gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg
                                                                     3780
agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa
                                                                     3840
cattleegtg tegecettat teeetttttt geggeatttt geetteetgt ttttgeteac
                                                                     3900
ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac
                                                                     3960
atcgaactgg atctcaacag cggtaagatc cttgagagtt ttcgccccga agaacgtttt
                                                                     4020
ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccg tattgacgcc
                                                                     4080
gggcaagage aactoggtog ocgoatacae tattotoaga atgacttggt tgagtactoa coagtoacag aaaagcatot tacggatgge atgacagtaa gagaattatg cagtgctgce
                                                                     4140
                                                                     4200
ataaccatga gtgataacac tgcggccaac ttacttctga caacgatcgg aggaccgaag
                                                                     4260
gagctaaccg cttttttgca caacatgggg gatcatgtaa ctcgccttga tcgttgggaa
                                                                     4320
ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta ctctagcttc ccggcaacaa
                                                                     4380
                                                                     4440
ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg
                                                                     4500
gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt
                                                                     4560
gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt
                                                                     4620
caggcaacta tggatgaacg aaatagacag atcgctgaga taggtgcctc actgattaag
                                                                     4680
cattggtaac tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat
                                                                     4740
ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct
                                                                     4800
taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct
                                                                     4860
tgagateett tttttetgeg egtaatetge tgettgeaaa caaaaaaace acegetacea
                                                                     4920
gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt aactggcttc
                                                                     4980
agcagagege agataceaaa tactgttett etagtgtage egtagttagg ceaceaette
                                                                     5040
aagaactetg tageacegee tacatacete getetgetaa teetgttace agtggetget
                                                                     5100
gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag
                                                                     5160
gegeageggt egggetgaac ggggggtteg tgeacacage ceagettgga gegaacgace
                                                                     5220
tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg
                                                                     5280
agaaaggegg acaggtatee ggtaagegge agggteggaa caggagageg caegagggag
                                                                     5340
cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt
                                                                     5400
gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac
                                                                     5460
geggeetttt taeggtteet ggeettttge tggeettttg eteacatgtt ettteetgeg
                                                                     5520
ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga taccgctcgc
                                                                     5580
cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga gcgcccaata
                                                                     5640
cgcaaaccgc ctctccccgc gcgttggccg attcattaat gcagctggca cgacaggttt
cccgactgga aagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag
                                                                     5760
gcaccccagg ctttacactt tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgga
taacaatttc acacaggaaa cagctatgac catgattacg ccagatttaa ttaagg
SEQ ID NO: 36
                       moltype = DNA length = 4393
                       Location/Qualifiers
FEATURE
source
                       1..4393
                       mol_type = genomic DNA
                       organism = adeno-associated virus 8
SEOUENCE: 36
cagagaggga gtggccaact ccatcactag gggtagcgcg aagcgcctcc cacgctgccg
cgtcagcgct gacgtaaatt acgtcatagg ggagtggtcc tgtattagct gtcacgtgag
                                                                     120
tgettttgeg geattttgeg acaccaegtg geeatttgag gtatatatgg eegagtgage
gagcaggatc tccattttga ccgcgaaatt tgaacgagca gcagccatgc cgggcttcta
                                                                     240
cgagatcgtg atcaaggtgc cgagcgacct ggacgagcac ctgccgggca tttctgactc
                                                                     300
gtttgtgaac tgggtggccg agaaggaatg ggagctgccc ccggattctg acatggatcg
gaatctgatc gagcaggcac ccctgaccgt ggccgagaag ctgcagcgcg acttcctggt
                                                                     420
ccaatggege egegtgagta aggeeeegga ggeeetette tttgtteagt tegagaaggg
                                                                     480
cgagagctac tttcacctgc acgttctggt cgagaccacg ggggtcaagt ccatggtgct
                                                                     540
aggeogette etgagteaga ttegggaaaa gettggteea gaccatetae eegeggggte
                                                                     600
gagececace ttgeceaact ggttegeggt gaccaaagae geggtaatgg egeeggeggg
```

## -continued

ggggacaaa glegotggac aleacataga ggaatataa agocgac georgeaaa cggetegtg georgacaac ggetegtgg egaagacac gaccacaca egacagagaaa gagacagaa gagagagaa gagagagaa gagagaga							
	ggggaacaag	gtggtggacg	agtgctacat	ccccaactac	ctcctgccca	agactcagcc	720
ciasgigadat         cigarceca attetgace         gercettata         agtecasaa         ectogegogo         900           ciatatagaga         cagacectot         cactatected         cacacecagoa         gercetagat         1020           casagagoaga         cagacetot         cactatected         cacacecagoa         tecacacecago         cacacecagoa         1080           cetagagagaa         cacacacecagoa         cetagagaat         cacacecagoa         tecacacecago         1140           cetagagaaa         cacacecagoa         cacacecagoa         tecacacegoa         1140           cetterente         cacacecagoa         tecacecagoa         tecacecagoa         1200           cetterente         tetterente         tegacaceagoa         tecacecagoa         1200           cetterente         tegacaceagoa         tegacaceagoa         tecacecagoa         1200           cettere         tegacaceagoa         tegacaceagoa         cacaceagoa         1200           cettere         tegacaceagoa         tegacaceagoa         1200         1440           cette         tegacaceagoa         terte <tegacaceagoa< td="">         1200         1440           cette         tegacaceagoa         terte<tegaceagoa< td="">         1400         1440           cetagagaacea</tegaceagoa<></tegacaceagoa<>							780
catatagoag   cligotogot   gotogotogot   cacoccogaa   accaccoga   agocatogot   catatoccogaa   accaccoga   accaccoga   cacogagact   catacogagac   catacogagagagagag   catacogagagagagagagagagagagagagagagagagagag	cgagcgcaaa	cggctcgtgg	cgcagcacct	gacccacgtc	agccagacgc	aggagcagaa	840
catataggag   clagacaggt   gacatecet   acatecet   acat	caaggagaat	ctgaacccca	attctgacgc	gcccgtgatc	aggtcaaaaa	cctccgcgcg	900
caggaggag c aggectopt acatetoctt casegoegy totogaagat casegoegy of totogaagat of totogaagat casegoegy							960
1980							1020
							1080
totoànogic tacgacocti cotaciored chocytett otogotogi otacqacacaca 1200 gittoggaaaa gocaacacaca totgotocaca cotgotocact chocytette gaacacacgaa agaccacaca totgotocaca cotgotocact chocytette otogotogic gocaacqacacacacacacaa cottocacacacacaa cottocacacacacacacacacacacacacacacacacac							
Strongsaka cgoaacaca totgocqut togaccogc accaccggca agaccaaca totgocacqut cyclogucatt chacagging chacactgga cacatggaa totgocquad cyclogucatgac cyclogucatgac gadagacga gagacgaga gagacgacga gagacgaga gagacgacga gagacgaga gagacgacga gagacgaga gagacgacga gagacgaga aga							
tseggaagec ategeceaeg cegtgecett chaeggetge genaetugge caatagaaga 1320 ctttecette ategataget tegacaagat giggatetge geggaagagag getgggaacag 1440 aaaagtgeaag tegtgeege caatgeeae caccecegtg ategacege geggaacaacacacaa 1500 catgtgeege tgattgaag ggaacagac caccecegtga ategatagate gegtgaacaa 1500 catgtgeege tgattgaag ggaacagac caccecegag accacagagag ctetcaagag cagagatta aaagagtett teegetggg cagtgateae gtgaacagag tagtgacaaa 1620 cagaagteta aaagaagtet teegetggg cagtgateae gtgaacagag tagtgacaaa 1740 cagaagteta titgeegaaa ggacagaaga aaagacegee cecagtagag gggataaaaag cagaggeta titgeegaaa ggacagaag aaaagtete gegtaagagag ggataaaag attteagagag gggacagaa aaaagatete cecggatgag cagtaacaag attgatgttt cotgeaaaa gtgacagaga aaaagttete gegtaacagag ggataaaag attgatgttt cotgeaaaa gtgtgegaag aatgaateaa gattteaaca tttgettaac cagagggeta aaagaagtgg atcagagaga aattgaatea gattteaaca tttgettaac cagaagaga agaategta agagtgttgg gatteaacag gattgaacag gattgatgtta 1980 caacaagagg agatagag aactgaggag atcagagagagagagagagagagagagagagagagaacaagacaagaa agaceteteaa gagatgtgtgg gattgaacaagagagagaagaa gacacaagaagaga gacgacagaaga gacgacagaaga gacgacagaaga gacgacagaaga gacgacagaa gacgacagaaga gacgacagaaga gacgacagaaga gacgacagaaga gacgacagaaga gacgaagaagaga gacgaagaagagagaagaagaagaagaagaagaagaagaa							
etticoctic asigatique ceasagea gigatecteg tyggagage geagatigae geocaagee tyggagage geagatigae gegeocaagee caageeee tecteagage etatique geagageee caceteegig acceptace ceaseaceaa 1800 catatiguege gigatique gigaacegee caceteegig acceptace etateagega etetecagga 1860 cagaagegee gigatecateg gigatecage gigacegae gigategeegae titigeegae gigacegae gigatecateg gigategeegae 1800 cagaagegae gigategeegae acceptacege gigategeegae 1800 cagaagegeegae acceptacege gigategeegae acceptacege gigategeegae acceptacege gigategeegae acceptacege acceptacege acceptacegeegae acceptacegae acceptacegeegae acceptacegeegae acceptacegeegae acceptacegae acceptacegeegae acceptacegee							
general grigagator or casagaceat tectograge ageaaggic or acasagacaa aaaagatgaan togtecaceo acacecetya ateguatore casacaceaa casacaceaa acaceteroga cacecaceaaga etcacaceaga etcacaceaga garattagaa gagtacaaaa agetaceaa aagatacaaa gaattagaa gagtacaaaaa le20 gaagaagata aaagaatta tecacaceaga gacetagaaa gattitagaa aggtacaaaaa le20 gaagaagata aaagaaceaga cocquataaaaga gagtacaaaa le20 gaagaagaaga gaceagaaa aagaaceaga cocquataaaaga gadacaaaa le20 gaagaagaaga tettagacaaga gaceagaaaagaaga gacaacaaaa aaagaaceaga cocquataaaaga gadacagaa le30 gaagaagaaga tettagaaagaagaaga gacaacaaaa aagaaceaga cocquataaaaga gaagaagaaga gacaacaaaa aaacagaagaaga aaagaagaaa gagaagaaga aacagaagaa gacagaagaa gacagaagaa aacagaagaa gacagaagaa gacagaagaa gacagaagaa gacagaagaa gacagaagaa gacagaagaa gacagaagaaga cocquataaaaga gacagaagaaga gacagaagaa gacagaagaa aagacagaaa gacagaagaa gacagaagaa aacagaagaa gacagaagaaga cocquataaaagaagaagaa gacagaagaaga gacagaagaaga gacagaagaaga gacagaagaagaagaagaagaagaagaagaagaagaagaa							
asagtogaag togtocogoo agatogacoo cacctocagu atogtocaco cacatogaco cactogacogo cactogacogo cactogacogo cactogacogo cactogacogo caccacogaco cottocagogo gacgagaagaa cacatogaco cactogacogo gacgagaagaaga gacgagaagaaga gacgagaagaaga gacgagaagaagaagacogoo cagagaagaagaagaagaagaagaagaagaagaagaagaa							
catifactogo gradattaca gagaacaa cacettedaga cacetagaa etcagaaga tecagaataata aagtagaacaa cacegaagatta aagtagacaaa aagtagacaaa aagtagacaaaa gagaagaaga tecagaagaaga tecagaagaagaaga gacegagaa aagaaccaga cocagaagaga gagaagagaag							
coggafith         aagtittagaa         tocqcoggo         cattagaca         agtataaaa         1620           goagagaata         aagagaccaga         aagaccaga         agaacagaag         1740           goagaccaag         aagaccaga         aagaccaga         1740           goagaccaag         gogaccaag         aagacagag         1740           goagaccaag         gogaccaag         aaccagg         1740           goagaccaga         gogaccaga         aaccagg         1740           goagaccaga         gogaccaga         aactagg         1860           gatgottit         coctoctotaga         aactatoga         aactatoga           gattagtotto         cogactyptog         gottogacatag         2190           aataatagact         taaacagag         attiggacatag         gottogaaca         2160           cocaccagca         aagacgattag         gogtogacagag         gottogaagat         2220           cocaccagca         gocacctroa         cagacgate         cagacgate         cagacgate         2220           cocaccagca         goctocagat         cagacgate         cagacgate         2220         2240           cocacagaga         gottocaga         goctocaga         2220         2280           goagaccaag							
grangsagtc asagasttct tecgettagge captateae grancesag tgegeagas 1680 grittlacets agasagtoge gageacagea asagacetge eccaptage gageacagas 1680 tetegragaea (gageacaga) 2000 tetegragaea (gageacaga) 1740 tetegragaea (tracegaea) gageacagaea (acategaea) captageagaea ettegragaea captageaeaa (captageaea) categaeaeaa (gageacagaeaeaeaeaeaeaeaeaeaeaeaeaeaeaeae							
guertacida agalaggeg gagecageaa algaccegec ecgateaseg eggataaaaag 1740   cegagecaaa eggagectgee octeagtege eggateatega eggateagaag 1800   cegagecaaaa eggagectgee ettegegaaaa etaatetga etteaaea tttgettea 1820   cacagaggete tegectgea actetgtgea gatacaaaa eaaatgitt eggateaaaagaag acgatacegaa actetgtge gatteatea etteaaea tttgettea 1920   cacagagagaa teggateatega actetgtgea eggateatea etteaaea tttgettea 2040   gattgettge teggectgea actetgtgea eggateatea etteaaea tttgetaea 1920   cacacaagaa acgatacegaa actetgtgaa eggatatea etgagacega gattgetgat ttteaaeaa ttgagacega 2160   cacacagea taaacagagaa gacgacegag eggatatea tegagacega eacetetega gageacetga eggacetae eggacegag eggacetae eggacegag eggacetae eggacegag eggacetae eggacegag eggacetae eggacegag eggacetae eggacegaa eggacetae eggacegaa eggacetae eggagagaa ectetagag ecgaaagaaga gacgagagaa ectetagaga egcaaagaaga eggagetga eacetegga eggacetae eggagagaa ectetagaga egcaaagaaga eggagetga eacetegge eggagagaa ectetagagaaga egcaaagaga eggacetae eggagagaaga egcaaagaga egcaaagaga eggagetga eacetegga eggagagaa ectetagagaaga egcaaagaga egcaaagaga eggagetga eacetegga eggagagaaga egcaaagaga eggagagaagaaga eggagagaga eacetegga eggagagaga eacetagaga eggagagaagaaga egcaaagaga egcaaagaga eggagagaagaga	ccggatgttt	aagttcgaac	tcacccgccg	tctggagcac	gactttggca	aggtgacaaa	1620
cgagcocaag cgggectyce ctcaqtege ggatecateg acgtacqaacg cggaaggacg tteogggget tteoggtggat tteogctggag adaptacqaa gatecatega ggatecatega gatecatega gatecatega gatecatega acateggggg catgytte ceggacggg catgytte cagaactec acacggggt gaggtette agacacgaagagga gatgatggg actgatega atteated ctggtgggg gggtecega 2040 gatgytty teggectgga actgytag cggatetatet ctggtgggg gggtecega 2160 acctetctga gggcattega gaggggggg gtggtggtgt teetgggtag ctggagaaga agaccaaga gaccaacaaga gaccaacaaga gaccacgaaa ggcatgaga agaggagggg gtgtggtgt teetggata aagaccaacaag gaccacgaaa ggcacgaaga gagggggggg	gcaggaagtc	aaagagttct	tccgctgggc	cagtgatcac	gtgaccgagg	tggcgcatga	1680
teogrogaga titigeogaea ggtaceaaaa caaatgttet egfeaegegg geatgettte cotsgeaaaa egtagetgttt cotsgeaaaa egtagetgtt coeggegtg teagaatete aaceggegtg 1980 gatagetgttt cotsgeaaaa egtacegga gateateteat eggetgegggggggggggggggg	gttttacgtc	agaaagggcg	gagccagcaa	aagacccgcc	cccgatgacg	cggataaaag	1740
teogtiggaa titigeogaea gigtaceaaa caastitti cigtacegeg geatgette 1920 acaceggigte agagatiget coceggegit teagaatte accegtiget 1980 cagaaagag aegtategga actegateggit teagaatte acceggite 1980 cagaaagag aegtategga acteggite teagaatte acceggite 1980 cagaaagag aegtategga acteggite 2980 gattyetty teggeotteg acteggite 2980 gattyetty teagacegga atteggite 2980 cateacagaga aagaagaga atgaateg 2980 cacaceagaa aagaagaga atgagtegga getyaaca 2980 cacacagaa aagaagaga gacggeeggg getyaaca 2980 cacacagaa agaacagaga gacggeegg getyaaca 2980 cacacagaa agaacagaga caagatgg agaggaga cogtaaaga 2980 cacaacagaa ggoctacaga caagatgg agaggaga cogtaaaga 2980 cacaacagaa agactico agaagaaga gacggagaga accectaa acceacaga 2280 cacaacagaa agactico agaagaaga gacggagaa accetaagaa 2880 cacagegaa agacggatea ggaagaaga gacggagaa accetaagaa 2880 cagagactaa gacggatea ggaagaaga gacagagaga gacaacacac 2980 caagactect tacgggaata gacggatga agaccetaa accetagga gacagaacacacacacacacacacacacacacaca	cgagcccaag	cgggcctgcc	cctcagtcgc	ggatccatcg	acgtcagacg	cggaaggagc	1800
aatgeggto agaactgg actested gegagaga atgateag attteaaca tttgetted cagagagtg cagaactgg gegetecega 2040 acacagagga aegatetega actetgte gatteateat etgetggge gggetecega 2100 acatetetga gggettega tegagagagagagagagagagagagagagagagagagaga							1860
acacagogate agajactget cajagatgiti coccagogit teagaatete aacegitegt caganagaga acetaterge gatteateate etgetgage aggetecega gattgettge teagacetge atetggteaa etgegtacet ataaatgacet taaaceaggi ateggtegeg atgegtateate teagatgitege cacaceagca aaagaagga cacatetca gagettgeg gattgytygg getgaaace tegageceg aaggecaaag agacectaa cagaactgac aaggaggagg etgtgytet teetggetae aagtaceteg agacagaaa ggoetacgaa aaggaggagg etgtgytet teetggetae aagtaceteg agacagaagaa gacacaagaaga cacaagaaga aaggaggaga aggagcaca gagaggagaagaaga aaggaggaga aggagaagaag							
cagalagaga acitatega alcitatege galteateat etgetgaget gagetecega 2040 gategategt teggectege ateggracet gagagetegt trettgagea 2100 acctetetga gagetatege ateggraceteg gatgatetge tretgagea 2100 acctetetga gagetatege geggeagea etgegageae 2220 accaecagea aaageaggae gaeggeegg gettgatet tretgageae 2220 agaecetteaa gageategaa gaeggeegga gatgategt tretgageae 2220 agaecetteaa gageategaa gaeggeegga gatgagetgate tretgagegga gaegeegea gaeggeeteeg 2240 accaecagea gaegageae agaeggagea accaecagea geggeagea geggeeteeg 2240 accaecagea geggeagea gatetateat gagagegga accaecagea gaeggaga accaecagea gaeggagaa accaecagea gaeggagaa accaecagea gaeggagaa accaecagea gaeggagaa accaecaga gaeggagaa accaecaga gaeggagaa gaeggagaa accaecaga gaeggagaa accaecaga gaeggagaa gaeggagaa accaecaa accaecaga aagaetacate treggeagaaaga gaeggagaa gaeggagaa gaeggagaa aacaacaaa agaetaaata treggagaa gaeggagaa accaecaga gagagagaa gaeggagaa gaeggagaa gaeggagaa accaeaaa agaetaaata gaactaaata tagaacaacaa agaecacaaa accaeaaaa agaecaaaa  agaecaaaaa agaecaaaaa agaecaaaaa agaecaaaaaaaaaa							
gattgetty to teggecteg alteggteen eggstacteg gatgatetty the the teggen and atanatyact thanacagy attgytygg getganace tegangeeg teggsgated getganace tegangeeg and gatgatety the tecangating to tegangeed and accepted and accepted and general getting the tector of the teggetgan and gatgatety and gatgatety the tetting and and gatgatety and gatgatgate and gatgatgatgate and gatgatgatgate and gatgatgatgate and gatgatgatgatgatgatgatgatgatgatgatgatgatg					_		
ataatgact taaacaggt atggttggg gegtgaacc tggagccog aagoccaaag 2220 ccaaccagca aagocgaagae gacggceggg gtctgatet tectggtac aagtaceteg 2280 gaccetteaa cggactegac aagoggagge cetecaacge ggeggacgac ageggeceteg 2340 accaccagcag agoctacgac aagoggagge cetecaacge ggeggacga geggeceteg 2400 accaccagcag agoctacga caagoaggte tecaagaaga tacgtettt gggggacacc 2400 accaccagcag agoctace ggcaagaaag ggctetega agocggtag acctecage tegggtaga 2520 aaggggada agocgac ggaagaaag ggcaacagaa acctecteggt tegggagacc 2400 acgaccatca tacgggate ggaagaaag gcaacagac cgccagaaaa agactcaat 2580 aagaccacta tacgggate ggaagaaga gcaacagac cgccagaaaa agactcaat 2580 aagactecte tacgggate ggaagaaga gcaacagac cgccagaaaa agactcaac 2580 aagacgccta tggtgtggg cetaatacaa tggctgaag cgdtgggac acaataggaa 2700 aacaataacga aggcgcaac ggagtgggt agteetegg cggtggega caataggaa 2700 aacaataacga ggacgaac atcacacac ggaacctaa acctectaga gaacatcaca 2700 aacaataacga ggacgaac atcacacac ggaacctaa actetegga agaggcacc acaataggaag 2700 aacaatagga gaacagate atcacacac ggaaccagaa ctgggcect ccaactaca 2820 acaacaacact ctacaagaa atctccaacg ggaactagga 2930 actactactog tacaagaac atctccaac ggaaccagaa 2930 accatacacct ggactggac cacacacc agagtgttac ggaatagaa ggaaccaca aagacaaca 2940 ttcacacac ggactataa accacaca agagagtca ggaatagaa ggaatagaa ggaatagaa 2940 ccatactogg tacaagaac ttcacaca acaacacat ggaattacag 2930 ccatactaga ggctttcaac accagggca 2930 cttactacta cggctttac agaccatca agagagtcac gaaatagaa ggaaccaaga 3120 ccatactagga gactagaac ttcactaca acaacacat ggaattacgg 293 cgaataaca 2940 ccatacaggac 2940 ccatacagga 2940 ccatacagga 2940 ccatacagga 2940 ccatacagga 2940 ccatacagga 2940 ccatacaga 2940 ccatac							
acactectega gagcatege gagtggtggg getgaaace tegagecega aagcecaaag 2220 ceaaccagea aaagcagaa gacggeeggg gtetggtgtet teetggetac aagtaceteg 2280 gaccetteaa eggactegaa aagggggggg eegtcaacge ggeggagea geggeetetg 2340 agcacgacaa ggectacgac cagcagette agggeggtgt aaatcegtac etgeggtata 2400 accacgeega egcagattt caggagegte tgeaagaag tacgtettt gggggacaac 2460 teggggaga agtettecag gecaagaag gggttetega accteteggt etggttgagg 2520 aaggegeta agcggetet gggaaagaa gecaacagee eggetace etggttgagg 2520 cagacteete tacgggcate ggcaagaag gecaacagee eggetagggggaacac 2580 cagacteete tggttggga ectaatacaa tggetgageta ecceteggg aaactecace 2700 cagacteete tggttggga ectaatacaa tggetgagg eggtggggacac 2760 acaatacgaa aggegeega ggagtgggta gtteetegga gagtgggega acatgggag 2760 acaatacga aggegegae ggagtgggta gtteeteggg aaattgggat tggattee 2880 acaatacga aggegegae ggatgggta gtteeteggg aaggggeed ecaactaca 2880 acaatacga ggacagagat atcacacaca gacceagaac etgggeeteg ecaactaca 2880 acaataccac etacagacac ecetggggg attttgag gggatece acaceacaca 2880 acaatacga gtactggaa gactaataa acaacaacag gggattee etacacaca 320 acaatecace tgactgggaa ggacatactaa acaacaacag gggattee etacacaca 3300 actactecteg etacagacac ecetggggga tttttgact tacagacaca aaggacaaca 2940 cetacttegg etacagacac acaacagaca eggattete acacacaga 3000 actactgecaa tacactacac acaacacaca aggattetac gacacagaa 3120 cetactecac gtactggea etacacaca acaacagata ecaacagaa 3120 cetactecaca gacacteca acaacacaca aggattetac gacacagaa 3120 cetactecaca egacacata ettectege agategtga acceggetg ggacgtgtte 3240 cetactacaca etaacctaca acaacacaca agacegaca acatecacaga 3120 cetacacacacacacacacacacacacacacacacacaca							
ccaacagca aaagcaggac gacggoggg gtctggtgct toctggctac aagtacotog 2280 gacccttcac ggactcgac aagggggggg cogtcacqcc ggoggacgac goggcctcq 2340 agcacgacaa ggoctacgac aaggagggtc toctgagatc ctgoggtac 2400 accacgocga gtcttccag gocaagaagac gggttctcga acctcctggt ctggtgagg 2520 aagggctaa gacgtcct gggaaagaag ggctactcga gcgacgacc 2460 togggcggac agtcttcag gocaagaaga ggcaaccagc cggcggacac 2520 aagggctaa gacggctca gagtcagtc ggaaagaag ggcaaccagc cgcagaaaa agactcaat 2520 aaggcgcta taggggatc ggaaagaag ggcaaccagc cgcagaaaa agactcaa 2700 cacgacgctc tggttggga ccatatcaaa taggtgaag cggtgggac caataggcag 2700 cacgacgctc tggttggga ctaatcacac ggacccaa acctctogg agaacctcag 2700 cacgacgctggg cgacagagtc atcaccaca ggacccgaa ccatatggcat tggtgatca 2820 cataggttggg cgacagagtc atcacacca gacccgaac ctgggcctg ccaactagac 2700 cactactcag taaccacac ctacaagaa atctccaaga gaactggga gttctcgg aggacgacac 2820 cataggttgg cfacacac accaggag attttpactt taacagattc cactgcaact 2820 catagctca taacctacc aggactaca agagaggacac cgaaatgga 2940 cttcactctcg tgactggca cacacacac aggactacac aggacacaca 2940 ccatactcgcaa taacctcac agacactca aggattcag ggaattacg 2940 ccatacgctaa taacctcac agacactca aggattcag 2940 ccatacgcaa taacctcac agacacacac 2940 ccatacgcaa taacctcac agacacacac 2940 ccatacgcaa taacctcac agacacacac 2940 ccatacgcaa taacctcac 2940 ccatacgcaa taacctcac 2940 ccatacgcaa taacctcac 2940 ccatacgcaa taacctcac 2940 ccatacgaga 2940 ccatacaga 2940	_						
gaccttcaa ggactacgac aggaggaga cegtaaacge ggggacga geggcatect 2400 accacgocga ggctdacgac cagaagacgt aggaggtga aatcegtac ctggggtga 2400 accacgocga agcttcacga gecaagaaga gggttetega acctetegt ctgggtgagg 2520 aagaggagac agcaggatct cagaagaaga gggttetega acctetegt ctggttgagg 2520 aagaggactactc tacgggcatc ggcaagaaga gccaacaccc cacgettete 2580 cagactoctc tacgggcatc ggcaagaaga gccaacacgc cgcaagaaa agactcaatt 2640 ttggtcaagac tggcgactca ggcaggtgat gccaacacgc cagcgtcatcac 2700 caagcgcoctc tggttgtggga cctaatacaa tggctgaga gcgactacaccc cacatggcag 2760 acaataacga agggcgcac ggagtgggta gttcetegg aaattggcat tgggttcac 2820 acaataacga agggcgcac gagtgggta gttcetegg aaattggcat tgggttcac 2820 acaataacga agggcgcac cactaggaga ggacaccaa accacaa gcgaccaca 2940 acaataacga gcgacagagat atcacacaa ggacactaca cacacacaa 2940 acaatacttgg ctacagaac cettgggggt attttgactt taacagattc cactgcact 3000 tttacaccact tagctggac ggacatcaca aggagtcac aggattcag aggattcag 2940 actactttagg ctacagaac accetaca agaagstcac aggattcag agcaccaaga 3120 actacgctac taggtcac agaaccaca aggagttcac acaacatg ggattcag accaagaag 3120 actacgcaa taacctcaa accacaca aggatttaca acaacatg ggattcag accaagaag 3120 accatcgcaa taacgcaa tttecttca acaacacatg aggattcag accagaaga 3120 accatcacac gctetteaa gttectcac acaacacaca aggattcac acaacacag 3120 accatcacac gctetteaa gttectcac acaacacaca acaacagtag gagattcag 3120 accatcacac gctetteac gttectegaa acaacacaca acaacagtag 3240 accatcacac gctetteac gttectegaa acaacacaca acaacagtag 3240 accagaagaac ggaaatacac attgacacaca acaacacaca acaacacaca 3340 accacacacacacacacacacacacacacacacacaca							
agacagacaa agcetaagac cagaagctgc aggaggtga caatcagtac etgoggtata 2460 tegggeggac agtetteeag gecaagaaga tagaagat tagtetttt gggggcaac 2460 aagacgetaa gacggteet ggaaagaag gagtetega aceteteggt etggttgagg 2520 aaggegetaa gacggteet ggaaagaag gacggtagag gecateace cagegttete 2580 cagacteete tacggggate ggaaagaag gacagtagag gecateace 2700 teggteagac tggegactaa gagtagtte cagaceetea acetetegga gaacetecag 2700 cagegeete tggttgggga cetaatacaa tggetgeag gggtgggga caattggeat 2760 acaataacga aggeggacgac ggagtgggta gtteeteggg agattgggat cagaatggaag 2700 categgttggg cgacagaagte ateaceacea geaceegaac etgggecet etgegatteea 2820 categgttgg cgacagaagte ateaceacea geaceegaac etgggecet eccacetaca 2880 cactactegg ctacagaca atetecaaceg ggacategga aggageacea cacagacaaca 2940 cetacttegg ctacaggac cectaggggt attittgact taacagate cacagacaaca 2940 cetacttegg ctacaggac cectaggggt attittgact taacagatec cacageagaa 2940 cetacttegg ctacaggac gacteatea acaacaacteg ggatteegg cecaagagac 3000 cettedgataa getetteaa atecaggtca aggaggtea ggaatgagggace cacaggagac 3120 ceategeaa taaceteace agacacatea aggattetta ggacteggg gacactacagg 3120 ceateggeaa tagacetes ctacacacaca acaacageacaca 3120 ceategeaa taaceteace acaagaget geetgeetee gtteeggag taccagateg 3240 teatettacac ctgggatac ctacacaca acaacagga accaggaag accagaaga 3120 ceatetactg cetggatac ctacacaca acaacagga accaggaaga accagaagac 3240 teatetacac ctteggaatac ctacacaca acaacaggaa accagacaca accagtagac 3340 cettetactg cetggates ctacacacaca agattegga accagacaca agagtegga 3340 cettetactg cetggates gaatecteg ctacacacaca agaggacac acaacacaa 340 ceaacaggac agaaatacg cagactete ggetgecte gtteegga accacacaa 340 ceaacaggaca ggaaatacg cagactete getteac ctggacac accacacaa 340 cacaggaggaca ggaaatacg cagactetgg ctgcacagata cetgtacta ctgtetegg accacacaca 3340 caacaggagaca ggaaatacg cagactetga ggaggategta gaacacacaa 3340 caacaggagaca acaacacaca agaacacaca agaagagac 3720 aaggagttttt tecaacag ggaatectga ttttggca acaacacaca agaagagaca 3720 aaggagtttttt tecaacag aggaagacac acagagaa atcaacacaca 3720 aaggagtagac ggaatacga tagg							
aceacgccga egcegagttt caggagctt tgaagaggt tgaagaaga tacgtctttt gggggcaacc 2460 tcgggggaga agtcttccag gcaagaaga gcagtagag ggettctega acctctogt ctggttgagg 250 aagagctcat tacgggcatc ggcaagaaag gcaacaagc cgcagaaaa agactcaatt 2640 ttggtcagac tggggactca ggcagaaga gcaacagc cgcagaaaa agactcaatt 2640 ttggtcagac tgggggctca gagtagttc cagaccctca acctctcgg gaactcaccc 2700 caagcgcctc tggttggga cctaatacaa tggctgcagg cggtggggca ccaataggaa 2760 acaataacga agggccgac ggagtgggt atcccaca acctctcgg gaactcacc 2880 acaatacacct ctacaagcaa atctccaacg ggacatcgg aggggcacc acaatggcag 2760 acatggctggg cgacagagtc atcaccacca gcaccgaac ctgggccctg ccaactacac 2880 acaaccacct ctacaagcaa atctccaacg ggacatcgg aggagcacc acagaccaca 2940 actacttcgg ctacagcac accetggggct atttggatt tacagacaca 2940 cctacttcgg ctacagcac accetggggct atttggatt tacaggacca 2940 actacttcaa gctctcaac accaagcac aggagttcg ggagttcgg cccaagagaac 3060 tcagcttcaa gctctcaac accaaggtca aggagtcac gcagaaatga ggcaccaaga 3120 ccatcgcaa taacctcacc agaccatcc aggagttcg ggaactggg ccaagagaac 3180 acgtacgttca cggctaga ctacacaca acaacggtag cgctgccc gttcccggg ggacgttcta 3240 acgtacgttc cggattac cttcctcgc acaagagct gtcagggag accggtgg ggacgtgtca 3300 accttctacac cttcgagaga gtgcctttc acagcagct accggcag ggagtgttca 3420 accggtgta gaaaccctgg gtgctttc acagcagct accggcag gaaccgtgg 3420 accggtgta gaaaccacaa accacaca accaggcta accggcaca acattcaga 3300 accggtgta ggaaaccac ggaaatacg accattgg accgggtag accggagaga accggagaga 3420 accaggagag aaaacacaaa accacaca accaggact accggacaa accggaacaa 3420 accaggagaga gaaaccacaa accacaca accaggacaca accggacaca accggacaca 3420 accaggagaga gaaaccacaa accacacaca accagaccacacaca	-						
tegggegaac agtettecag gecaagaag gggttetega accteteggt etggttagag geoggtea gaeggeteet ggaaagaagag gaeggtagag gecateacee eagegttete 2500 cagegeteet tegggegate ggeaagaaag gecagtagag gecateacee eagegttete 2610 cagegeete tegggagate ggeaagaaag gecagtagag gecateacee eagegttete 2610 cagegeete tegggagate ggaagaaag gecagtagag gaaceteegag 2700 cagegeete teggttgegga etataceaa gecaetea teggtgegga ecaatagaga 2760 accaatacaa tetecaacea ggaceetea etacaceae teggggegga aattgggag agaaceteaa 2820 catggetggg egaaagaagte atcaceacea geaceegaac etggggeete ecaaetacaa 2820 catggetggg egaacaacaa attecaace ggacateggg agageecae acaetaceae 2940 cetacteegg egaceaeae atteegggt attettggatt taacagatte eaceaceae geaceeaeae 2940 cetacteegg ggatteegg agggeteete egaaagaace 2940 cetacteegg egaceeaeae accagggag agggeeaeae acaegaagaa 3120 ceatgeeae taaceeaeae accaegagate ggatteegg gggatteegg ecaagagaa 3120 ceatgeeaa ataceeteae ageaceaeae aggtgtttae gggatteegg gaegateeag 23180 ceatgetee eggtetee egtegetee gtteeeggg gaegtgttea 3180 ceatteaeae acaeggaga eacaegagag taceagaete 3180 ceatteaeae atcaeggaga eacaeggaga accaggeagg gaegtetee 3180 cettetaete eggetetee egtegeeaeae eacaeggaga accaeggaga accaegaga 3120 ceatteaeae atcaeggaa gaegetee ggaaceteeae 3180 cettetaete eggetetee egtegeeaeae eacaeggag gaegeteea 3180 cettetaeae atcaeggaa gaegetee ggaaceteeae 3180 cettetaeae acaeggaga accaeggaae acaeaeaeae 3180 cettetaeae eacaeggaa accaeaeaeae 3180 cettetaeae eacaeggaae accaeaeaeaeaeaeaeaeaeaeaeaeaeaeaeaea	agcacgacaa	ggcctacgac	cagcagctgc	aggcgggtga	caatccgtac	ctgcggtata	
aaggogtaa gacggctcct ggaaagaag gaccggtag gccatcacc cagcgttctc cagactcctc tacgggcatc ggaagaagaag gccaacacc cgccagaaaa agactcaatt ttggtcagac tggcggcacca gagacgagtag gccaccagaca cettergag gaactccad 2700 cagcgcctct tggtgtggga cctaataccaa tggctgcagg cggtggcga ccaataggcag 2760 acaatgactggg cgacagagtc atcaccaca gcacccgaac cggtggcgac ccaatggcag 2760 acaatgactggg cgacagagtc atcaccaca gcacccgaac ctggggccctg ccaacctaca acaaccacct ctaccaacga atcaccaca gacccgaac ctggggccctg ccaacctaca acaaccacct ctacaagcaa atcaccaca gacaccgaac acagacacac acagacaaca cctacttcgg tacaagaaca atcacaaca ggactcaca acaacaactg ggattcaca acaacaactg ggattcaca acaacaactg ggattcacaa atacctacaca agaccatca agaggtcac gaattcagg cacactaca aggtttaca cactgcacat tacactacac	accacgccga	cgccgagttt	caggagcgtc	tgcaagaaga	tacgtctttt	gggggcaacc	2460
cagactecte tacgggeate ggeaagaaag gccaacagee egcagaaaa agacteaatt tggegactea gagteggtet cagacectea acctetegga gaaceteeag 2700 cagegeetee tggetgggga ectaatacaa tggetgeaga eggtgggga eggtgggga gtteetegga aaattggeat tgegatteea 2820 categgetggg ggaeggggga atteeteaag gagaateegga 2880 cacaceacace eccacetace acaaceacee eccaceacea 2920 cetaetteegg etacagaagaa atteecaace gacaceagaa aggagageace cacaceaca 2940 cetaetteegg etacagaagaa atteecaage gagaateegg aggagtgeete eccaagaagaa 2940 cetaetteegg etacagaagaa eccaageaa aggagageace eccaagaagaa 2940 cetaeteaa gecaceace aggagatteegg eccaagaagaa 2940 cetaeteaa gecaceace aggaggteete eacaagagaa 2940 cetaeteeaa agacacace aggaggteete eacaagagaagaa 2940 cetaeteaa agacaceace aggaggteete eacaagagaagaagaagaagaagaagaagaagaagaagaag	tcgggcgagc	agtcttccag	gccaagaagc	gggttctcga	acctctcggt	ctggttgagg	2520
ttggtcagac tggcgactca gagtcagttc cagacccca acctetegga gaacctccag 2700 cagcgccctc tggtgtggga cctaatacaa tggctgagg gggtggggac ccaatggcag 2760 caacgacaga aggcgccgac gagatggta gttceteggg aaattggact gcaactacaa 2820 catggctggg cgacagagtc atcaccacca gcacccgaac ctgggccctg cccactaca 2880 acaaccacct ctacaagcaa atctccaacg ggacatcggg aggagccacc aacgacaaca 2940 cctacttcgg ctacagcacc coctgggggt attttgactt taacagattc cactgccact 3000 ttttcaccacg tgactggcag cgactcatca acaacaactg gggattccgg cccaagaagac 3000 ttttacacacg tgactggcag cgactcatca acaacaactg gggattccgg cccaagaagac 3000 tctagcttcaa gctctcaac accaaggtca aggagtcac ggaattcagg cccaagaaga 3120 ccatcgctcaa faacctcacc agcaccatcc aggtgtttac ggactcggag taccaggag 3180 ccatcgccaa taacctcacc agcaccatcc aggtgtttac ggactcgggag taccaggtg 3180 ccttctactg cggattac cttcetegc gctgcctcc gttcccggcg gacgtgttca 3240 tgattcocca gtacggtta cttctctcgc agatgctgg gatccggg gaggtctct 3300 ccttctactg cctggaatac tttccttcgc agatgctggg gacgctgg gacggtctc 3300 ccttctactg cttggaatac tttccttcgc agatgctggg gacgctgg gacggttca 3300 ccttctactg cttggaatac tttccttcgc acacaggag tcagggcac aacttccag 3300 cttactacac cttcgaggac gtgcctttcc acagagagta ccggcaaa aacttccagt 3300 ccaatcaagg ggcaaatacg ctggctttc acagagagta ccggcaaa aacttccag 3300 ccaatcaagg ggcaaatacg ctggctttcc acagagagta ctgggaca aacttccag 3300 ccaaccggtgat gaatccttg attgaccagt acctftacta cttftctcgg aactcaaaca 3480 ccaatcaggc aaacaaca agaacctgg gtctcaggac aggtggcct aatacaatga 3480 ccaatcaggc aaacaacaa agcaaccttg gcttcagca aggtgggcct aatacaatga 3540 ccaatcagga aacaacaat agcaactttg cctggaactg tgggaccaaa taccatctg 3600 caaccggaaaaaa aggaatacga atctggaa atcggaaga aatcaacaa agcagagaga 3780 aatggagatt caggagaga atcatgg atgaccaaa acaacaacaa acaacacaaa 3780 aatggagatt caggagaga atcggagaga atcacaaca acaacacaa 3780 aatggagatt caggagaga atcggagaga atcacaacc acaacacaca 3780 aatggagat ggaaccag atggagaga aattgaac accaacac acaacacaca 3900 aatacaacac aggagagaca ttccacacacacacacacacacacacacacacacac	aaggcgctaa	gacggctcct	ggaaagaaga	gaccggtaga	gccatcaccc	cagcgttctc	2580
teggtagaac tggtgagaa cagatcagatc cagaccatca accategag gaacctcag 2700 caagcgcctct tggtgtggga cctaatacaa tggctgagg cggtgggga caataggatca 2760 caacgataacga agcgccgac gagatgggta gttcctcggg aaattggattca 2820 catggctggg cgacagagtc atcaccacca gcaccgaac ctgggcctg cccactaca 2880 acaaccacct ctaccagcaa atctccaacg ggacatcggg agaggcacc aacgacacac 2940 cctacttcgg ctacagacac cctgggggt attttgatt taaccagatc cacategcag 3000 tttcaccacg tgactggcag cgactcatca acaacactg gggattccgg cccaagagac tcagcttcaa gctcttcaac accacgacac agaggtgtac ggagttccgg cccaagagac 3000 tcagcttcaa gctcttcaac accacgacactc agagtgttta ggactcggt ggcaccaaga 3120 ccatcgccaa taacctcacc accacaggcd gcctgccc gtcccgggg gacgtttca 3240 sqattcocca gtacggtac ctaacactca acaacactg ggttccggg gacgtgttca 3240 tgattcocca gtacggtac ctaacactca acaacacggtag tcaaggcgtg ggacgtctc 3300 ccttctactg cctggaatac tttccttcgc agatgctgag acacggcaac aacttccagt 3300 ttacttcactg cttggaatac tttccttcgc agatgctggagaa aactgggaac aacttcaga 3360 ttacttactac cttcgagaac cttctctcgc acacacacgagac cacacacaca 3360 ttacttactac cttcgagaac cttctctcgc acacacacac acacacacacacacacacacacaca	cagactcctc	tacgggcatc	ggcaagaaag	gccaacagcc	cgccagaaaa	agactcaatt	2640
caagegcoctt tggtgtggga cctaatacaa tggctgcagg eggtggca ccaatggcag 2760 acaataacaga aggegcogaa ggagtgggta gtctceggg aattggcat tggcatca 2820 acaataacaga aggegcogaa atcaccacag gcaccagaac ctgggacctg ccaactaca 2880 acaaccacct ctacaagcaa atctccaacg ggactacggg aggagcacc aacgacaaca 2940 accatcatcgg ctacagagtc atcaccacag ggactacggg aggagcacc aacgacaaca 2940 accatcatcgg tgactggcag cgactaatca acaacaactg gggattccgg ccaagagac 3060 ttagcatca agctctcaac accacca aggagtgta tttgactt taacagattc cactggcaac 3060 tcagcacca taacctacc agcaccatca aggagtgtac gcagaatgaa ggcaccaaga 3120 accatcagcaa taacctacca agcaccatca aggagtgttac ggactggga gacggtgtta 3240 aggagtcacc atcacacca gtacggctac ctaacaccac gcagagtgttac 2920 ttacttacac ctcagagac gtcattca acaacgagtag tcaggcggg gacggtgtta 3240 accaggagga ggaacacca tttecttcgc agattggtgag aaccggcaac aacttccaag 3300 accatcagcaac tttecttcgc agattgtgag aaccggcaac acattccag 3300 accatcagcaac ggcaaatacg cagacttctg gttcagcaca accaggctgt 3320 accaggaggaca ggaaacacg gacgattct accaggtga accagagacg 3420 accggctgat gaatcctctg attgaccag acctgtacac aggtgggcct actaacaca 3480 acaggaggaga aaacaacaa cagactttg gcttbacc acaggaggca acaacacag 3480 acaaccgggaa aaacaacaac gagactttg acctgcaca aggtgggcct aatacaatgg 3420 acaaccgggaa aaacaacaa dacacac actgttacac aggaggagaa aacacacacg gtctaaacaa 3600 acaaccgggaa aaacaacaa dacacac actgttacac aggaggagaa aacacacaca 3480 acaacgggaaa aacaacaac accagaac acctgttacac aggaggagaa aacacacaca 3480 acaaccgggaa aaacaacaa dacacacacacacacacacacacacac							2700
acaataacga aggogcogac ggagtgggta gttcctogg aattggcat tgcgattcca 2820 catggttggg ogacagatc atcaccacca gcaccogaac ctgggccttg ccacctacca 2880 acaaccacct ctacaagcaa atctccaacg ggacatcggg aggagcacc aaggacaaca 2940 cctacttcgg ctacagcac ccctgggggt attttgactt taacagattc cactgccact 3000 tttcaccacg tgactggcag ogactcatca acaacaactg gggattccgg cccaagaggac 3120 ccatcgccaa taacctcacc agcaccatcc aggatgttac ggactcgggg taccaggggt accatggcg gacgtgttca ggactcagtac accatgccac atcaccagtca cacaggggt gactcgggg gacgtgttca ggactcagtcc gttcaccag gtacggttccag gacggtttca ggactcagtac cttcatg cctggaatac tttccttcgc agatgctgctg tttcacggcg gacggtttca ggacggtgtcac accagggag taccaggag accacagag ggacggttca gcttcatga cttcatga cttcatg cttggaatac tttccttcgc agatgctgag accacagac cagaggttgg 3300 ccttctactg cctggaatac tttccttcgc agatgctgag accacagac cagaggttgg 3300 ccttctactg cctggaatac tttccttcgc agatgctgag accacagac cagaggttgg 3300 ccttctactg cattggaatac tttccttcgc agatgctgag accacagac cagaggttgg 3420 accaggaggaga ggacaatacag gagactctgg gtttcacgg cccacagc cagagttgg 3420 acaccggtgat gaatccctg attgaccagt acctgatca cttgttctcgg actcaaacaa 3540 ccaaccaggca aaacaaat agcaacttg gcttgacga attggggct aatacaatgg 3540 ccaaccaggca aaacaaat agcaacttg ctggactgc tgggaccaaa accacaaa gacgagagg 3720 agggttttt tcccagtaac gggatcctga tttttggaa accaaaaagcg gtcaagagaa 3720 agggttttt tcccagtaac gggatcctga tttttggaa accaaaaaagcg acaggagag 3720 aaattggaac tgtcaacag ggatcctga ttccacca ggaggagaa accaaaaa agcagcacca 3900 aaattggaac tgtcaacag gggggct taccagga ggacggaac ttccaccaca 3900 aaattggaac tgtcaacag ggggggct taccacag ggacggaac ttccaacaca 3900 aaattggaac tgtcaacag ggcggctt taccacggt ggcggaac ttccaacaca 3900 aaattggaac tggcggctt gggcggaac atcctcacaca aaacgacgaac acggcctcc 3900 aaattggaac tggagac cagagaac atccacaca aaactgca aaacgaacacacacacacacacacacacacacacac	caqcqccctc	tqqtqtqqqa	cctaatacaa	tggctgcagg	cqqtqqcqca	ccaatqqcaq	2760
cataggetggg egacagagte atcaccaca geaceegaac etgggeetg eccacetaca 2880 acaaccacct ctacaageaa atctecaacg ggacateggg agagecace aacgacaaca 2940 cctacttegg ctacageace ecctgggggt attttgactt taacagatte eactgeeact 3000 ttteaccacg tgactggag egacteatea acaaacatg gggatteeg eccaagagac 3060 teagettea getetteaac atccaggea aggaggteac ggattgag ggactggga gacgtgtta 2940 ccategteat taaceccace aggetetge eacaagaggt gettgag gacgtgtta 2940 cgtacgttet eggetetge eacaaggget gettgetee gtteeggag taccageggg 3120 cgtacgttet eggetetge eacaaggget gettgetee gtteeggag gacgtgtta 3240 ttactteaca ettegaggac etteegea acaacggtag teaggeetg ggacgteet 3300 accggetgat gaatecteg atteettee acaagggtag teaggeetg ggacgteet 3300 accggetgat gaatectetg atteettee acaagggtag teaggeetg ggacgteet 3300 accggetgat gaatectetg atteettee acaagggat ecceacage eagagettgg 3420 accggggggeet ggaatacg egeaaatacg etgeettee acaaggagat eccaacage acceacage acaacaggag 3420 accggggggagggggggggggggggggggggggggggg							2820
acaaccact ctacaagcaa atctccaacg ggacateggg aggagcacc aacgacaca 2940 cctacttegg ctacagcacc ccctgggggt attttgactt taacagattc catgccact 3000 ttttcaccacg tgactggcag cgactcatca acaacaactg gggattcgg cccaagagac 3060 tcagcttcaa gctcttcaac atccaggtca aggaggtcac ggagattcgg cccaagagac 3120 ccategccaa taacctcacc agaccatcc aggtgtttac ggactcggag taccaggag taccaggtg 3240 cgtacgttct cggctctgcc caccagggct gcctgcctcc gttcccggcg gacgtgttca 3240 tgattccca gtacggctac ctaacactca acaacgtag tcaggcgtg gacgtgttca 3240 ccttctactg cctggaatac tttccttcgc agaatgctag tcaggcagt ggacgctcct 3300 ccttctactg cttcgaggac gtgccttcc acagaggtg tcaggcaga aaccggcaga accggcaga accggcaga accggcaga ggacgctgg 3420 accaggaggac ggcaaatacg ctgccattcg acagagcta cctgtacac aggtggtca accggtagta gactctaga 3480 ccaactgggca aaacaacaa agcaactttg cttgaacga aggtgggcct aatacaatagg 3540 ccaactgggca aaacaacaa agcaactttg cttggactgc caacaacaca gacgagaaca 3600 acaaccgggca aaacaacaa agcaactttg cttggactgc caacaacaac gccaacaacaa 3600 acaaccgggca aacacaacaa agcaactttg cttggactgc aacacacaaa gacgagagaa 3720 agggtttttt tcccagtaac gggatcctga tttttggcaa acaacacaaa gacgacgaga 3720 aggggtttttt tcccagtaac gggatcctga ataacttgca acaacaacaa gacgacgaga 3780 actgggaaca tgccaacaa tacgtggaa ataacttgca ggacgaaaa accaacacaa 3840 tggaacacag ggatccaaca tagggaggcc taccaaca acaacacaa agcagcaggag 3960 aaattggaac tgtcaacag caggaggcc taccacaa taccaacaa accaggacg 3960 aaattggaac tgtcaacag caggaggca taccaggta aaatggaaca atcctccgc tcagaaca accgggacg 3960 tcacctgtaat ggcggcttt ggcccaaca atcctccgc tcagaacaa tctcaacccgt 3900 aaattggaac tgcggaac tcggaaca atcctccgc tcagaacaa tctcaacccgt 3900 aaattggaac tgcggaac tcggaacaa tcctccgcc tcagaacaa tcttcaacc 3960 accacggaacaacaa accggaaca 3960 accacggaacaacaacaa accggaacaa atcctccgc tcagaacaa accggaacaa 4000 accacacacaacaa accggaacaa acccggaacaacaacaa accacacaacaa 3960 accacacacacaacaa accggaaca acccggaacaacaacaa accacacaca	_						2880
cetacttegg etacagace ceetggggf attttgact taacagatte cactgecact tttcaccacg tgactgacag cgactcatca acaacactg gggatteegg eccaagagac 3000 accategocaa taacetcace agcaccate aggagtee gagatteegg eccaagagac 3120 accategocaa taacetcace agcaccatec aggttttac ggacteeggag taccagatga 3120 accategocaa taacetcace agcaccatec aggttttac ggacteeggag taccagetge 3180 accatecace gtacegotac etacactca acaacggtag tecegogg gacgeteet 3300 acctteatecg etacgatac ettecetoge geetgeetee gteeggag acceptace 3300 acctteatecg etaggaatac ettecegoga gaagetggg gacgeteet 3300 acctteatecg etaggaatac ettecetoga agateetggg gacgeteete 3300 acctgetatea ettecateg etaggagata etegeggaac aaceteeagt 3360 accaggagagaca etegegaga accegogaac acctgacacac acaacacaca 3480 accaggagagaca ggcaaatacg etgecaagac etgetacag etaggagaca acctgacacacacacacacacacacacacacacacacaca							2940
tttcaccacg tgactgcag cgactcatca acaacactg gggattccgg cccaagagac 3060 tcagcttcaa gctcttcaac acaccatcc aggagtcac gcagaatgaa ggcaccaaga 3120 ccatcgccaa taacctcacc agcacatcc aggagtttac ggactcggag taccaggag 3180 cgtacgttct cggctctgcc caccagggct gcctgcctc gttcccggag taccaggactg 3240 tgattcccca gtacggctac ctaacactca acaacggtag tcaggcgtg gacgttcta 3240 ccttctactg cctggaatac tttccttcgc agatgctgag accggcaac aacttccagt 3360 ttacttacac cttcgaggac gtgcctttcc acaggagcta ccgccacaagc cagagcttgg 3420 accggctgat gaatcctctg attgaccagt accggcagac acctgacacaa 3480 caggaggaca ggcaaatacg ctgccaggac cctgttaccg ccaacaagc aggtgggca accaacaa 3480 caaggaggaca ggcaaatacg ctgccaggac cctgttaccg ccaacaagc gtctcaacga 3600 caaccggcga aaacaacaat agcaactttg ctggactcc tgggaccaaa taccatctga 3600 caaccgggca aacaacaaa agcaactttg ctggactgc tgggaccaaa taccatctga 3600 caaccgggaa aacaacaat agcaactttg ctggaactgc tgggaccaaa taccatctga 3600 caaccgggaagaa ttcattggct atcctgga ttttttggcaa acaacaacaa gacgaggagg 3720 agcgtttttt tcccagtaac gggatcctga ttttttggcaa acaacacaaa accaggagag 3780 atggcggatta cagcgatgt atggcaga atacttgaa gcagcaaaaa accaggaca 3780 atggcggatta cagcgatgt atggcaga atacctgaa gcagcaaaaa accaggaca 3780 atggctacaaga ggatccatc tgggcaaga ttcctcaac ggaggaaca ttccaacccg 3840 tggactctga gggtccatc tgggccaaga ttcctcaac ggaggaaaa accagagaag 3780 aaattggaac tgcaacag caggggcct taccacac ggaggaaaa accagaaca accggaca 3990 aaaattggaac tgcaacac ttcaccaca ggacggaaa ttcctcaac ggacggaaa accaggaaca 3960 ttgtacctgaa gggtccatc tgggccaaga ttcctcaac ggacggaaa ttccaacccg 4020 ctccgctgaac tgcggatct cggaccacc tcaacacac acaacacac accaggaaca 4080 cgcaatacag caccggaag gtcagcgtga aaccacaca accaggaaca 4200 gcaaatacag caccggaaa tccgttaatc ctgaacccg cccaatacc acaaactc  acaaactcgaccgt acccgtaac accggaca acccgacc accgacca accaccac ac		_	_			-	3000
teagetteaa getetteaae atecaggtea aggagteae geagaatgaa ggeaceaaga 3120 ceateegeeaa taaceteaee ageaceatee aggtgtttaee ggacteegag taccaagea 3180 cettedactite eggeteee eacacaggget geetgeee gteecegge gaegtgttea 3240 teateteeea gtaeggetae etaacaatea acaacggtag teacggegtg gaegteete 3300 cettetacteg eetggaatae titeettege agatgetgag aaceggeaae aactteeagt 3360 titaettacae ettegaggae gteettee acageageta eggeacaeae acatteeagt 3420 aceggetgat gaateetetg attgaceagt acetgtaeta ettgetegg aceggetigg 3420 aceggetgat gaateetetg attgaceagt acetgtaeta ettgetegg aceacaaca 3480 ceaacagggea gaaacaacaat ageacettig getteageea aggtgggeet aatacaacaa 3540 ceaacagggea aaacaacaat ageacettig ettgetagea aceacaacae getetaacaa 3540 ceaacagggea aaacaacaat ageacettig ettgetagea aceacaacaa 3600 caacegggea aaacaacaat ageacettig eetggactge tgggaceaaa taccaatega 3540 ceaacagggea aaacaacaat ageacettig eetggactge tgggaceaaa taccaatega 3600 aatggaagaaa treattgget aateetggaa tettitiggeaa acaacaacaa agacgaagag 3720 aagggttitt t teccagtaac gggateetga tittitiggeaa aacaacaaaa acaaacacag 3780 atggegatta eagegatget atgeteacaca ggaaggaaga aatacaaaaca acggeteete 3900 aaaattggaac tgteaacage cagggggeet taccacag ggacgaaaac acegggaag 3960 teggetacaga gggeteete tgggecaaaa teeteegee teagaacete aaaggeteg 3900 aaaattggaac tgeggateet tgggecaaaa ateeteegee teagaacete 4020 ceccatacaacaa gggegggetet tgggecaaaa aceteegee teagaacete 4020 ceccatacaacaa ggegggggett tgggaaggaagaacaa teetteeace 4080 acegeaatacaa teeggaacaa gteeggaagaa ateeteegee teagaacete 4020 acegeaatacaa teeggaacaa ggeggggggeteac eetecaaca caaaacaa aaaggegaaaca 4200 acegeaatacaa teeggaaa teetegaace eetecaaca caaaacea 4200 acegeaatacea teeggaacaa teetegaace eetecaaca caaaacaa 4200 acegeaatacea teeggaacaa teetegaacea eetecaacaa aaaggetgaacea 4200 acegeaatacea teegtaate eetecaacaa aaaacaacaa 4200 acegeaatacea teeggaacaa teetegaacaa teetecaacaa aaaacaacaa 4200 acegeaaacaacaa tacaacaa aaaacaacaa aaaacaacaa aaaacaacaa		_		_	_	_	
coategecaa taacetcace ageaceatec aggitettac ggaetegaag taceagetge 3180 cgtaegttet eggetetgee caceaggget geetgeetee gtteeeggeg gaegtgttea 3240 tgatteeea gtaeggetae etaacaetea acaacggtag teageggetg ggaeggeteet 3300 cettetaetg cetggaatae ttteettege agatgetgag teageggaac acteeagga 3360 ttaettaeea ettegaggae gtgeetttee acageaggeta eggeeagaegeggetgg 3420 aceggetgat gaateetetg attgaeeag acetgtaeta ettgtetegg acteaacaa 3480 caggaggea ggeaaataeg eggeacaeteg ettgeeagae ettgtaeeg agggegeet aatacaatagg 3540 ceaateagge aaagaactgg etgeeaggae ettgtaeeg eggeacaaa aceaategg 3660 atggaagaaa tteattgget aateetggea tegttaege acaacaeaeg gteteaaega 3660 atggaagaaa teatttgget aateetggea tegttaege acaacaeaea ggaegetgt 3420 acggattitt teeeagtae gggateetga tegttaege acaacaeaeg gteteaaega 3660 atggaggatae eageagtte atgeteaea gggagaaga acaacaeaaa gaegaegagg 3720 ageggttitt teeeagtae gggateetga tittitggea acaacaeaea geeagagaaga 3720 atggggattae eggeagtge atgetegag ataacttgea geageaaaa aceggeteet 3900 acaateagg ggaataeggt atgetegag ataacttgea ggaegaaaa aceggeteet 3900 acaateagg ggateeeate tgggeeaaga teeetggaa ggaeggaac acaacaeaea 4020 eteegetgat gggeggett tgggeeaaa teeteegee teagateetg aceggaaga 4020 eteegetgatee tgeggateet eggaecaeet teaacaeae eteaaaeae 4020 geaaataeag gaaceeegag gteageegg aaattgaatg ggaegtgeaa teetteaae 4200 geaaataeag gaaceeegaa gteageaga aceeegaae cacaaceae 4200 geaaataeag gaaceeegaa gteageaga aceeegaae cacaaceae 4200 geaaataeag gaaceeegaa gteageaga aceeeaaea eeceaatee acaaatee acaagtgaga 4260 actttgetgt taatacaagaa ggegtgtaet etgaaceeeg eeceattgge aceeettaee 4320 teaceegtaa teetgtaate eeggttaate etgaaceeeg eeceattgge aceegttaee 4320 teaceegtaa teetgtaate eeggttaate etgaaceeeg eeceaatee 4320							
tgategttet eggetetgee caceaggget geetgeetee gtteeeggeg gaegtgttea 3240 tgatteeea gtaeggetae etteetgee aaaaaggstag teaggeegte ggaeggeteet 3300 cettetaetg cetggaatae ttteettege agatgetgag teaggeaggeae aactteeagg 3360 ttaettacae ettegaggae gtgeetttee acageageta eggeeaaaaa acctaeaaaaa 3420 aceggetgat gaateetetg attgaeeag acetgtaeta ettgtetegg acteaaaaa 3480 caggaggeae ggeaaataeg eggeeaaaaa ageaactteg getteagea aggtgggeet aatacaategg 3540 ceaateagge aaagaactgg etgeeaggae eetgttaeeg eeaaaaaaaa teataeategg acetggaae eetgttaeeg eegaaaaaa teataetggea aacetggae etgetaatge aacaaaaaaa gaegaetga acetggaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa							
tgatteccca gtacggetac ctaacactca acaacggtag teaggecett ggacgetect 3300 cettetactg cetggaatac ttteettege agatgetgag aaceggeac aacttecagt 3360 ttacttacac ettegaggac gtgeetttee acageageta egeceacage cagagettgg 3420 aceggetgat gaatectetg attgaccagt acetgtacta ettgtetegg acteaaacaa 3480 caggaggeac ggeaaatacg cagactetgg getteagea aggtgggeet aatacaatgg 3540 ccaatcagge aaacacaat ageaacttg cetgtaceg ccaacaacg getetaacaga 3600 caacegggaa aacaacaat ageaacttg cetggactge tgggaccaaa taccatetga 3600 caacegggaa aacaacaat ageaacttg cetggactge tgggaccaaa taccatetga 3600 acaggaggaagaa tetattgget aatectggea tegetatgge acaacacaaa gacgacgagg 3720 agggttttt teccagtaac gggatectga tttttggaaa acaacacaaa acagacgagag 3780 atggggatta cagegatgte atgeteaca gegaggaaga aateaaaacc actaaceetg 3840 tggactacaga ggaatacggt atggtggaag ataacttga geagaaaac acggeteete 3900 aaaattggaac tgteaacage caggaggeet taccggaat ggetetggaag accgggaaga 4020 ctccgetgaa gggegettt gggeetgaaac ateeteegee teagateetg 4020 ctccgetgat ggggggetet tgggeagaa ateeteegee teagateetg 4080 cgcaatacag caceggaaag gtcagedgg aaattgaatg ggaggtaca teetteacac 4080 cgcaatacag gaaceeegg gtcageaga ateeteegee teagateetg acagaacaa 4080 cgcaatacag gaaceeegg gtcageaga ateeteegee teagateetg acagaacaa 4080 actttgetgt taatacagaa ggegtgtact etgaaceeeg ecceattgg accegttace 4200 actttgetgt taatacagaa ggegtgtact etgaaceeeg ecceattgg accegttace 4200 actttgetgt taatacagaa ggegtgtact etgaaceeeg ecceattgg accegttace 4320 acaaceegtaa teetgtaatt eetgtaatte etgaaceeeg ecceattgg accegttace 4320 acaceegtaa teetgtaatt eetgtaatt eetgtaatee etgaaceeeg ecceattgg accegttace 4320							
cettetactg cetggaatac titectiege agatgetgag aaceggeaac aacticeagt 3360 titactiacac ettegaggac gtgeetitee acageageta egeceacage cagagetigg 3420 aceggetgat gaatectetg attgaceagt acetgtacta ettgetegg acteaacaa 3480 ccaacagggeac ggeaaatacg eagacetigg gettecagea aggtgggeet aatacaacag 3540 ccaacagggea aaacaacaat ageaactitg eetggactge egagaceaaa taceatetga 3600 aatggaagaaa ticattgget aatectggae tegetatgge aacaacaacag getetaacag 3600 atggaagaaa ticattgget atgeteacag tegetatgge aacaacaaaa gaegacgagg 3720 atggegatta eagegatge atgeteaca gggateetga tittiggeaa aacaacaaa acaaaatgee geeagagaca 3780 atggegatta eagegatge atgeteaca gegaggaaga aateaatage geagagaaca aceggeteete 3900 aaaattggaac tgteaacage eagggggeet taceeggtat ggtetggeag aaceeggaacg 3960 tegetetgaac gggeceate tgggecaaga teetecaca ggaeggeaac teetecacae ggaeggeaace teetecacae ggaeggeaga ateeteetga 4020 etcegetgat gggeggett ggeetgaaac ateeteegee teagateetg ateaagaaca 4080 egecaatacag eaceggaacag gteagedgg aaattgaatg ggagetgeag aagaggaaca 4200 gecaatacag gaaceceaga ateeagteac ectecaact etacaaace acaaatge 4260 acttigetgt taatacagaa ggegtgtact etgaaceeg ecceattgge accepttace 4320 teaceegtaa tetgtaattg eetgttaate aataaaceg ttgateegt teagttgaac 4380							
ttacttacac cttcgaggac gtgcctttcc acagaggta cgcccacagc cagagettgg 3420 accggctgat gaatcctctg attgaccagt acctgtacta cttgtctcgg actcaaacaa 3480 caggaggagcac ggcaaatacg cagactctgg gcttcagcca aggtgggcct aatcaatagg 3540 ccaatcagggc aaagaactgg ctgccaggac cctgttaccg ccaacaacgc gtctcaacga 3600 caaccgggca aaacacaat agcaactttg cctggactgc tgggaccaaa taccattgg 3660 attggaagaaa ttcattggct acceptaca 4ggactgca 4ggaccaaa 4acaatgg 3720 agcgtttttt tcccagtaac gggatcctga 4ttttggcaa acacacaaa gacgacgagg 3720 agcggttttt tcccagtac 4ggatcctga 4ttttggcaa acacacaaa gacgacgagg 3780 attggctacaga ggaatacgg attggcag ataccttga 2gcagcaaca 4acactgc 3840 ttggctacaga ggaatacgg attggcag atacctgca 2gcagcaaca 2gcgctcct 3900 aaattggaac tgtcaacag cagggggcct tacccggtat ggtctggcag acccgggag 3960 ttgacctgca gggcgcttt gggccaaga ttcccacac 2gacggaaca 4ccaccgg 3960 ttgacctgca 2gggggctt 2ggccaaca 4ccccgcc 4cagatcctg 4cagaaccac 4cacacagt 4caagaaca 4cacacaga 4ccacagaaca 4ccaccaga 4ccacagaaca 4ccaccaga 4ccacagaaca 4ccaccaga 4ccacagaaca 4ccacaacaa 4ccaacaga 4ccacagaaca 4ccac							
accggotgat gaatectetg attgaccagt acctgtacta ettgtetegg acteaacaa 3480 caggaggcae ggeaaataeg cagactetgg getteagea aggtgggeet aatacaatagg 3540 ccaateagge aaagaactgg etgeeaggae eetgttaeeg eeaacaaege gteteaaega 3600 caategggea aaacaacaat aggeaactttg eetggactge tgggaccaaa taccatetga 3600 atggaagaaa teattgget aateetggea tegttaegge aacaacaaaa gaegaeggagg 3720 agggttttt teecagtaae gggateetga tttttggeaa acaaacaaa gaegaegagg 3720 atggggatta eagegatgte atgeteacaa gegaggaaga aateaaaaee actaaeeetg 3840 ttggetacaaga ggaataeggt ategtggeag ataacttgea geageaaaaa acggeteete 3900 aaaattggaae tgteaacage eaggggeet taceeggtat ggtetggeag aaccgggaeg 3960 ttgtaeetgaa gggeteete tggggeeaaae ateeteege teagateetg 4020 eteegetgat gggeggettt ggeetgaaae ateeteegee teagateetg ateaagaaea 4080 egeaataeeag eaceeggaaag gteageggga aattgaatg ggagetgeaa acaagtgag 4260 actttgetgt taatacagaa ggegtgtaet eegaaceeg eeceattgge accegttaee 4320 teaceegtaa teegtaattg eetgttaate aataaaeeg ttgateegt teagttgaae 4320 teaceegtaa teegtaattg eetgttaate aataaaeeg ttgategt teagttgaae 4380							
cagagaggaca ggcaaatacg cagactctgg gcttcagcca aggtgggcct aatacaatgg 3540 ccaatcaggc aaagaactgg ctgccaggac cctgttaccg ccaacaacagc gtctcaacga 3600 caaccgggca aaacaacaat agcaactttg cctggactgc tgggaccaaa taccatctga 3660 atggaagaaa ttcattggct aatcctggca tcgctatggc aacaacaacag gccaagaaga 3720 agcgttttt tcccagtaac gggatcctga tttttggcaa acaaaatgct gccagagaca 3780 atggggatta cagcgatgtc atgtcacca gcgaggaaga aatcaaaacc actacctg 3840 tgggtacaaga ggaatacggt atgttggcag ataacttgca gcagcaaaca acggtcctc 3900 aaaattggaac tgtcaacag cagggggct tacccggtat gggccgaaa atcctacaaca ggaaggagaa acccggaac 3960 tgtacctgca gggtcccatc tgggccaaga tcctcaaca ggaaggaaca tcctcaaca ggacggcaaca tccaccggt 4020 ctccgctgat gggcggcttt ggcctgaaac accctccacc tcaaccagtc acaagacca 4080 cgcaatacag caccggacag gtcagcagg aattgaatg ggagctgcaa acaagtggaca 4200 gccaatacag gaaccccgaa gtcagcaga accccaacta ctacaaatc acaagtgtgg 4260 actttgctgt taatacagaa gggtgtact ctgaacccg tcagtcgt tcagttgac 320 tcacccgtaa tctgtaattg cctgttaatc aataaaccg ttgattcgt tcagttgac 4380							
ccaatcaggc aaagaactgg ctgccaggac cctgttaccg ccaacaacgc gtctcaacga 3600 caaccgggca aaacacaat agcaactttg cctggactgc tgggaccaaa taccatctga 3660 atggaagaaa ttcattggct aatcctggca tcgctatggc accacacaaa gacgacgagg 3720 atggcgatta cagcgatgtc atgctcacca gcgaggaaga aatcaaaaacc actaaccctg 3780 atggcgatta cagcgatgtc atgctcacca gcgaggaaga aatcaaaaacc actaaccctg 3840 tggctacaga ggaatacggt atcgtggcag ataacttgca gcagcaaaac acggctcctc 3900 aaattggaac tgtcaacagc cagggggcct tacccggtat ggtctggcag aaccgggaagg 3960 tgtacctgca gggtcccatc tgggccaaga ttcctcacac ggacggcaac ttccacccg 4020 ctccgctgat gggcggttt ggcctgaaac atcctccgcc tcaagatcctg atcaagaaca 4080 cgccatfacc tgcggatcct ccgaccacct tcaaccagtc aaagctgaac tctttcatca 4140 cgcaataccag gaaccccgag atccagtaca cctccaacta ctacaaatct acaagtgtgg 4260 actttgctgt taatacagaa ggcgtgtact ctgaaccccg cccattggc acccgttaac 4320 tcacccgtaa tctgtaattg cctgttaatc aataaaccgg ttgattcgtt tcagttgaac 4380	accggctgat	gaatcctctg	attgaccagt	acctgtacta	cttgtctcgg	actcaaacaa	
caaccgggca aaacaacaat agcaactttg cctggactgc tgggaccaaa taccatctga 3660 atggaagaaa ttcattggct aatcctggca tcgctatggc aacacacaaa gacgacgagg 3720 agcgtttttt tcccagtaac gggatcctga tttttggcaa acaaaatgct gccagagaca 3780 atggcggatta cagcgatgtc atgctcacca gcgaggaaga aatcaaaacc actaaccctg 3840 tggctacaga ggaatacggt atcgtgcag ataacttgca gcagcaaaac acggctcctc 3900 aaattggaac tgtcaacagc cagggggcct tacccggtat ggtctggcaga acacgggacg 3960 tgtacctgca gggtcccatc tgggccaaga ttcctcacac ggacggaaca ttccacccgt 4020 ctccgctgat ggggggcttt ggcctgaaac atcctccgcc tcagatcctg atcaagaaca 4080 cgccatfacc tgcggatcct ccgaccacct tcaaccagt aaagctgaac tctttatca 4200 gcaagcgctg gaaccccgag gtcagcagtac cctccaacta ctacaaatct acaagttgg actttgctgt taatacagaa ggcgtgtact ctgaaccccg cccattggc acccgttacc 4320 tcacccgtaa tctgtaattg cctgttaatc aataaaccgg ttgattcgtt tcagttgaac 4380	caggaggcac	ggcaaatacg	cagactctgg	gcttcagcca	aggtgggcct	aatacaatgg	
atggaagaaa ttcattggct aatcctggca tcgctatggc aacacaaaa gacgacgagg 3720 agcgttttt tcccagtaac gggatcctga tttttggcaa acaaaatgct gccagagaca 3780 atgcggatta cagcgatgtc atgctcacca gcgaggaaga aatcaaaacc actaaccctg 3840 tggctacaga ggaatacggt atcgtggcag ataacttgca gcagcaaaac acggctcctc 3900 aaattggaac tgtcaacagc cagggggcct tacccggtat ggtctggcag acccgggacg 3960 tgtacctgca gggtcccatc tgggccaaga ttcctcacac ggacggacac ttccacccgt 4020 ctccgctgat gggcggcttt ggcctgaaac atcctccgcc tcagatcctg atcaagaaca 4080 egcctgtacc tggggatcct ccgaccacct tcaaccagt aaagctgaac tctttcatca 4200 gcaagcgctg gaaccccgag gtcagcaggac actccaacta ctacaaatct acaagttgg 4260 actttgctgt taatacagaa ggcgtgtact ctgaaccccg cccattggc acccgttaac tcgtaattg actgtaact aataaccgg ttgattcgtt tcagttgaac 4380	ccaatcaggc	aaagaactgg	ctgccaggac	cctgttaccg	ccaacaacgc	gtctcaacga	3600
atggaagaaa ttcattggct aatcctggca tcgctatggc aacacaaaa gacgacgagg 3720 agcgttttt tcccagtaac gggatcctga tttttggcaa acaaaatgct gccagagaca 3780 atgcggatta cagcgatgtc atgctcacca gcgaggaaga aatcaaaacc actaaccctg 3840 tggctacaga ggaatacggt atcgtggcag ataacttgca gcagcaaaac acggctcctc 3900 aaattggaac tgtcaacagc cagggggcct tacccggtat ggtctggcag acccgggacg 3960 tgtacctgca gggtcccatc tgggccaaga ttcctcacac ggacggacac ttccacccgt 4020 ctccgctgat gggcggcttt ggcctgaaac atcctccgcc tcagatcctg atcaagaaca 4080 egcctgtacc tggggatcct ccgaccacct tcaaccagt aaagctgaac tctttcatca 4200 gcaagcgctg gaaccccgag gtcagcaggac actccaacta ctacaaatct acaagttgg 4260 actttgctgt taatacagaa ggcgtgtact ctgaaccccg cccattggc acccgttaac tcgtaattg actgtaact aataaccgg ttgattcgtt tcagttgaac 4380	caaccgggca	aaacaacaat	agcaactttg	cctggactgc	tgggaccaaa	taccatctga	3660
agggtttttt teccagtaac gggatectga tttttggeaa acaaaatget geeagagaca 3780 atgeggatta eagegatgte atgeteacea gegaggaaga aatecaaaace actaaceetg 3840 tggetacaaga ggaataeggt ateetggeag ataacttgea geageaaaac acggeteete 3900 aaattggaac tgteaacaage eaggggeet taceeggtat ggtetggeag aaceeggacg 3960 tgtacetgea gggteceate tgggecaaga tteeteacac ggaeggaac tteeaceegt 4020 etecgetgat gggeggettt ggeetgaac ateeteegee teagateetg ateaagaaca 4080 egeeatgtace tgeggateet eegaceacet teaaceagte aaagetgaac tetteatea 4140 egeaaatacaag eaceeggacag gteagegtgg aaattgaatg ggagetgeag agagaaaca 4200 geaagegetg gaaceeegag ateeagtaca eeteeaacta etacaaatet acaagtgtgg 4260 actttgetgt taatacagaa ggegtgtaet etgaaceeg tegategt teagttgaac 4320 teaceegtaa tetgtaattg eetgttaate aataaaaceg ttgattegt teagttgaac 4380							3720
atgeggatta cagegatgte atgeteacea gegaggaaga aateaaaace actaaceetg 3840 tggetacaga ggaataeggt ategtggeag ataacttgea geageaaaace acggeteete 3900 aaattggaac tgteaacage cagggaggeet taceeggtat ggtetggeag aacegggaeg 3960 tgtacetgea gggteecate tgggeeaaga tteeteacae ggaeggeaac tteeaceeegt 4020 eteegetgat gggeggettt ggeetgaaac ateeteegee teagateetg ateaagaaca 4080 egeetgtace tgeggateet eegaceacet teaaceagte aaagetgaac tettteatea 4140 egeaatacag caceggacag gteagegtgg aaattgaatt ggagetgeag aaggaaaca 4200 gecagetgt gaaceeegga ateeagtaca eeteeaacta etacaaaatet acaagtgtgg 4260 actttgetgt taatacagaa ggegtgtaet etgaaceeeg eeceattgge accegttace 4320 teaceegtaa tetgtaattg eetgttaate aataaacegg ttgattegtt teagttgaac 4380							
tggctacaga ggaatacggt atcgtggcag ataacttgca gcagcaaaac acggctcctc 3900 aaattggaac tgtcaacagc cagggggcct tacccggtat ggtctggcag aaccgggacg 3960 tgtacctgca gggtcccatc tgggccaaga ttcctcacac ggacggcaac ttccacccgt 4020 ctccgctgat gggcggtctt ggcctgaaac atcctccgcc tcaagcaccat tctttcatca 4080 cgccattacag caccggacag gtcagcgtgg aaattgaatg ggagctgaac tctttcatca 4140 cgcaatacag caccggacag gtcagcgtgg aaattgaatg ggagctgcag aaggaaaaca 4200 gcaagcgctg gaacccgga atccagtaca cctccaacta ctaccaaatac acaagtgtgg 4260 actttgctgt taatacagaa ggcgtgtact ctgaacccg ccccattggc acccgttacc 4320 tcacccgtaa tctgtaattg cctgttaatc aataaaccgg ttgattcgtt tcagttgaac 4380							
aaattggaac tgtcaacagc cagggggcct tacceggtat ggtctggcag aaccgggacg 3960 tgtacctgca gggtcccatc tgggccaaga ttcctcacac ggacggcaac ttccacccgt 4020 ctccgctgat gggcggcttt ggcctgaaac atcctccgcc tcagatcctg atcaagaaca 4080 cgccatgtacc tgcggactcct ccgaccacct tcaaccagtc aaagcgtgaac tcttcatca 4140 cgcaatacag caccggacag gtcagcgtgg aaattgaatg ggagctgcag aaggaaaaca 4200 gcaagcgctg gaaccccgag atccagtaca cctccaacta ctaccaaatct acaagttgg 4260 actttgctgt taatacagaa ggcgtgtact ctgaacccg ccccattggc acccgttacc 4320 tcacccgtaa tctgtaattg cctgttaatc aataaaccgg ttgattcgtt tcagttgaac 4380							
tgtacetgca gggteceate tgggecaaga ttecteacae ggaeggeaae ttecaecegt 4020 eteegetgat gggeggettt ggeetgaaae ateeteegee teagateetg ateaagaaea 4080 egeetgtaee tgeggateet egaceacet teaaecagte aaagetgaae tetteatea 4140 egeaataeag caeeggaeag gteagegtgg aaattgaatg ggagetgeag aagaaaaea 4200 geaagegetg gaaeceegag ateeagtaea eeteeaaeta etaeaaatet aeaagttgg 4260 actttgetgt taatacagaa ggegtgtaet etgaaeceeg eeeeattgge aeeegttaee 4320 teaecegtaa tetgtaattg eetgttaate aataaaeegg ttgattegtt teagttgaae 4380							
ctccgctgat gggcggcttt ggcctgaaac atcctccgcc tcagatcctg atcaagaaca 4080 cgcctgtacc tgcggatcct cggaccacct tcaaccagtc aaagctgaac tctttcatca 4140 cgcaatacag caccggacag gtcagcgtgg aaattgaatg ggagctgcag aaggaaaaca 4200 gcaagcgctg gaaccccgag atccagtaca cctccaacta ctacaaatct acaagtgtgg 4260 actttgctgt taatacagaa ggcgtgtact ctgaaccccg ccccattggc acccgttacc 4320 tcacccgtaa tctgtaattg cctgttaatc aataaaccgg ttgattcgtt tcagttgaac 4380							
egectgtace tgeggateet eegaceacet teaaceagte aaagetgaae tettteatea 4140 egeaatacag caceggacag gteagegtgg aaattgaatg ggagetgeag aaggaaaaca 4200 geaagegetg gaacecegag atecagtaca ecteeaacta etacaaatet acaagtgtgg 4260 actttgetgt taatacagaa ggegtgtact etgaaceeg eeceattgge accegttace 4320 teaceegtaa tetgtaattg eetgttaate aataaacegg ttgattegtt teagttgaac 4380							
ogcaatacag cacoggacag gtcagogtgg aaattgaatg ggagotgcag aaggaaaaca 4200 gcaagogotg gaacooogag atocagtaca ootocaaota otacaaatot acaagtgtgg 4260 actttgotgt taatacagaa ggogtgtact otgaacooog coccattggo accogttaco 4320 toacoogtaa totgtaattg ootgttaato aataaacogg ttgattogtt toagttgaac 4380				_	-	_	
gcaagogotg gaacooogag atooagtaca ootooaacta otacaaatot acaagtgtgg 4260 actttgotgt taatacagaa ggogtgtact otgaacooog coccattggo accogttaco 4320 toacoogtaa totgtaattg ootgttaato aataaacogg ttgattogtt toagttgaac 4380							
actttgetgt taatacagaa ggegtgtaet etgaaceeeg ceecattgge accegttace 4320 teaceegtaa tetgtaattg eetgttaate aataaacegg ttgattegtt teagttgaac 4380							
tcaccegtaa tetgtaattg cetgttaate aataaacegg ttgattegtt teagttgaac 4380							
3 3 3 3 3 3 3 3 3							
tttggtgtgt aca 4393	tcacccgtaa	tctgtaattg	cctgttaatc	aataaaccgg	ttgattcgtt	tcagttgaac	4380
	tttaatetet	gcg					4393

The invention claimed is:

- 1. A recombinant adeno-associated virus (AAV) comprising an AAV capsid, and a vector genome packaged therein, the vector genome comprising an expression cassette comprising
  - a promoter operably linked to
  - a nucleotide sequence encoding a human argininosuccinate synthase 1 (ASS1);
  - wherein the nucleotide sequence encoding the human ASS1 comprises SEQ ID NO: 2.
- **2**. The recombinant AAV according to claim **1**, wherein the AAV capsid is an AAV8 capsid or variant thereof.

- 3. The recombinant AAV of claim 1, wherein the promoter is a liver-specific promoter.
- **4**. The recombinant AAV of claim **1**, wherein the promoter is a thyroxine binding globulin (TBG) promoter, an alpha 1 anti-trypsin (A1AT) promoter, or a transthyretin promoter (TTR) promoter.
- **5**. The recombinant AAV of claim **1**, wherein the promoter is a constitutive promoter.
- 6. The recombinant AAV of claim 1, wherein the vector genome further comprises an AAV 5' ITR sequence from AAV2 and an AAV 3' ITR sequence from AAV2.

72

- 7. A pharmaceutical composition comprising the recombinant AAV of claim 1 and an aqueous carrier, excipient, diluent, or buffer.
- **8**. A method of treating citrullinemia in a patient in need thereof, the method comprising administering the pharmaceutical composition of claim **7** to the patient.
- **9.** A recombinant adeno-associated virus (AAV) comprising an AAV capsid, and a vector genome packaged therein, the vector genome comprising:
  - (a) an AAV 5' ITR sequence;
  - (b) a nucleotide sequence encoding a human argininosuccinate synthase 1(ASS1); and
  - (c) an AAV 3' ITR sequence,
  - wherein the nucleotide sequence encoding the human ASS1 comprises SEQ ID NO: 2.
- 10. The recombinant AAV of claim 9, wherein the AAV capsid is an AAV8 capsid or variant thereof.
- 11. The recombinant AAV of claim 9, wherein the AAV 5' ITR sequence from AAV2 and the AAV 3' ITR sequence are from AAV2.
- 12. The recombinant AAV of claim 9, wherein the AAV 5' ITR sequence comprises SEQ ID NO: 16 and the AAV 3' ITR sequence comprises SEQ ID NO: 17.

74

- 13. A plasmid comprising an expression cassette, the expression cassette comprising a promoter operably linked to a nucleotide sequence encoding a human argininosuccinate synthase 1 (ASS1),
- wherein the nucleotide sequence encoding the human ASS1 comprises SEQ ID NO: 2.
- **14**. The plasmid of claim **13**, wherein the promoter is a thyroxine binding globulin (TBG) promoter, an alpha 1 anti-trypsin (A1AT) promoter, or a transthyretin promoter (TTR) promoter.
- 15. The plasmid of claim 13, wherein the promoter is a constitutive promoter.
- 16. The plasmid of claim 13, further comprising an AAV 5' ITR sequence from AAV2 and an AAV 3' ITR sequence 5 from AAV2, wherein the AAV 5'ITR sequence and the AAV 3' ITR sequence flank the expression cassette.
  - 17. The plasmid of claim 16, wherein the AAV 5' ITR sequence comprises SEQ ID NO: 16 and the AAV 3' ITR sequence comprises SEQ ID NO: 17.
    - 18. A 293 cell comprising the plasmid of claim 13.
    - 19. A 293 cell comprising the plasmid of claim 15.
    - 20. A 293 cell comprising the plasmid of claim 17.

\* \* \* \* \*