



US 20250263678A1

(19) **United States**(12) **Patent Application Publication** (10) **Pub. No.: US 2025/0263678 A1**
(43) **Pub. Date:** **Aug. 21, 2025**(54) **COMPOSITIONS AND METHODS FOR
DELIVERY OF RNA**(71) Applicant: **REJUVENATION TECHNOLOGIES,
INC.**, Mountain View, CA (US)(72) Inventors: **John RAMUNAS**, Mountain View, CA
(US); **Glenn Jeremy MARKOV**,
Mountain View, CA (US); **William
Gillis OLSEN**, Mountain View, CA
(US)(73) Assignee: **REJUVENATION TECHNOLOGIES,
INC.**, Mountain View, CA (US)(21) Appl. No.: **18/270,179**(22) PCT Filed: **Dec. 28, 2021**(86) PCT No.: **PCT/US2021/065386**

§ 371 (c)(1),

(2) Date: **Jun. 28, 2023****Related U.S. Application Data**(60) Provisional application No. 63/131,528, filed on Dec.
29, 2020.**Publication Classification**(51) **Int. Cl.**

C12N 9/12 (2006.01)
A61K 9/51 (2006.01)
A61K 38/00 (2006.01)
A61K 47/54 (2017.01)
A61K 47/69 (2017.01)
A61K 48/00 (2006.01)
A61P 1/16 (2006.01)

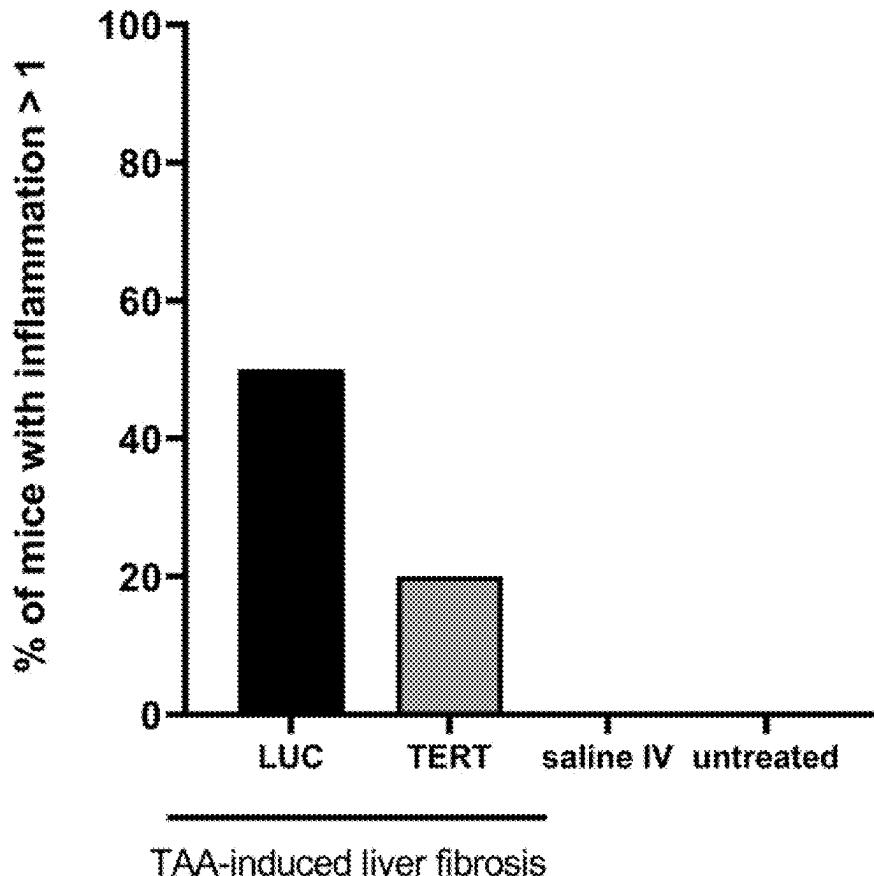
(52) **U.S. Cl.**

CPC *C12N 9/1276* (2013.01); *A61K 9/5123*
(2013.01); *A61K 47/543* (2017.08); *A61K
47/6929* (2017.08); *A61K 48/0033* (2013.01);
A61P 1/16 (2018.01); *A61K 38/00* (2013.01);
C12Y 207/07049 (2013.01)

ABSTRACT

The disclosure relates to compositions and methods for the treatment of fibrotic diseases and disorders and/or liver diseases and disorders, with one or more synthetic messenger ribonucleic acids (mRNAs) encoding telomerase reverse transcriptase (TERT).

Specification includes a Sequence Listing.

Mice with high inflammation in the liver

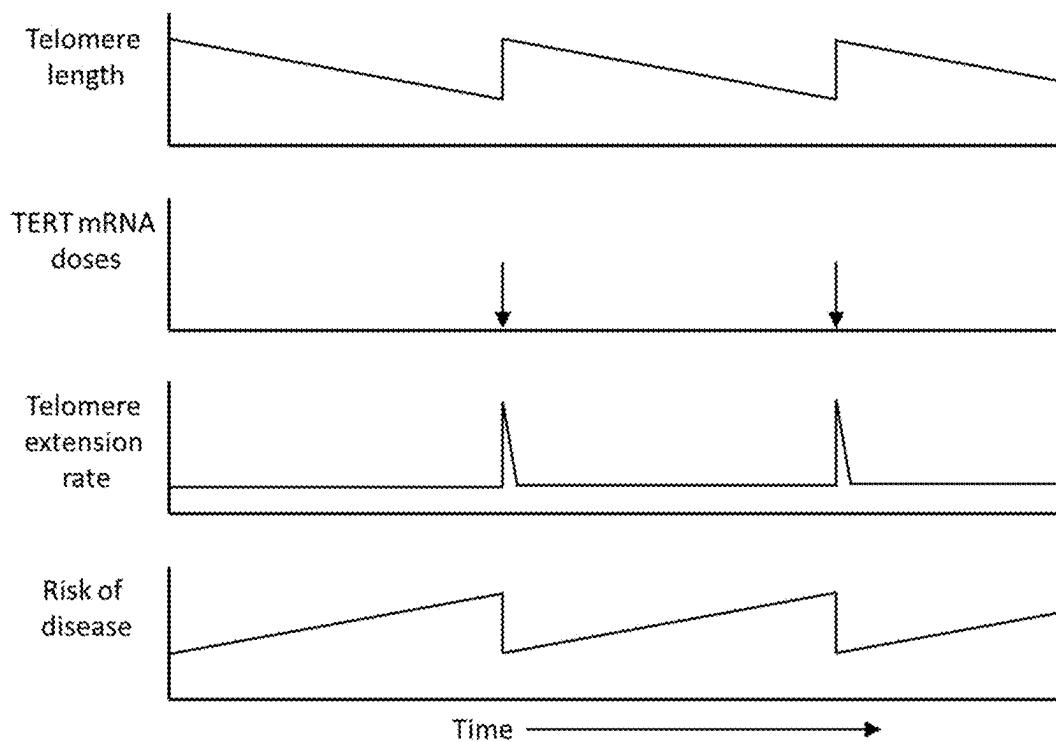


FIG. 1

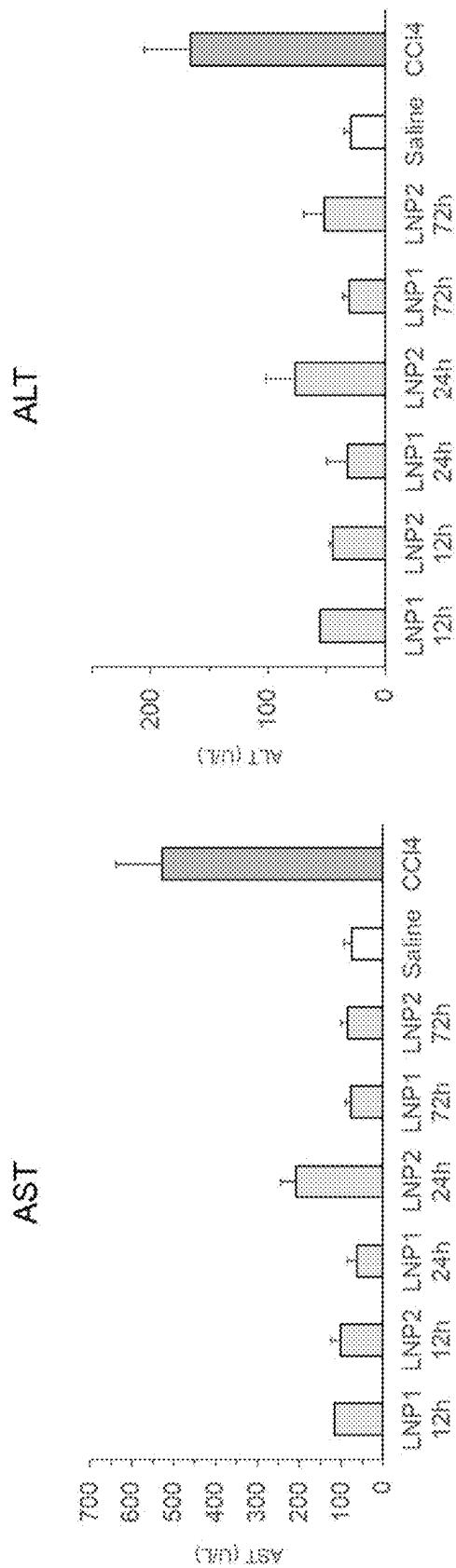


FIG. 2

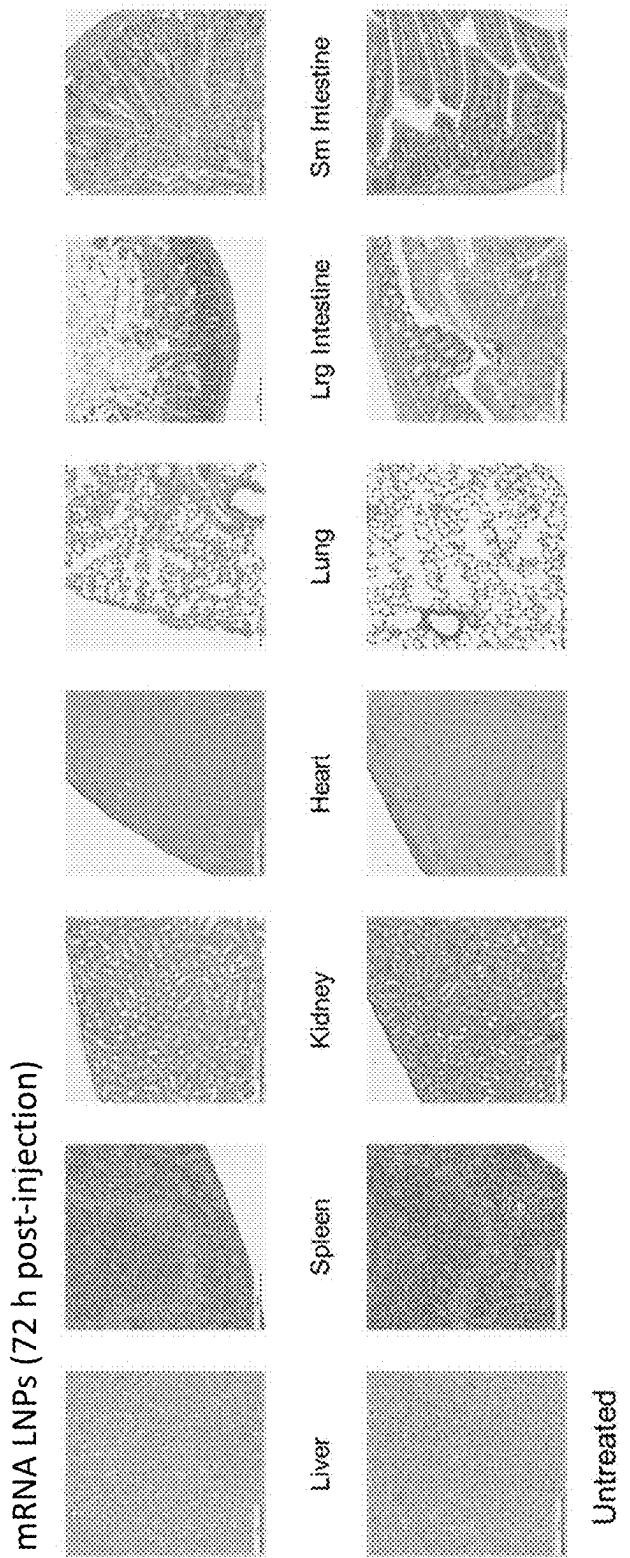


FIG. 3

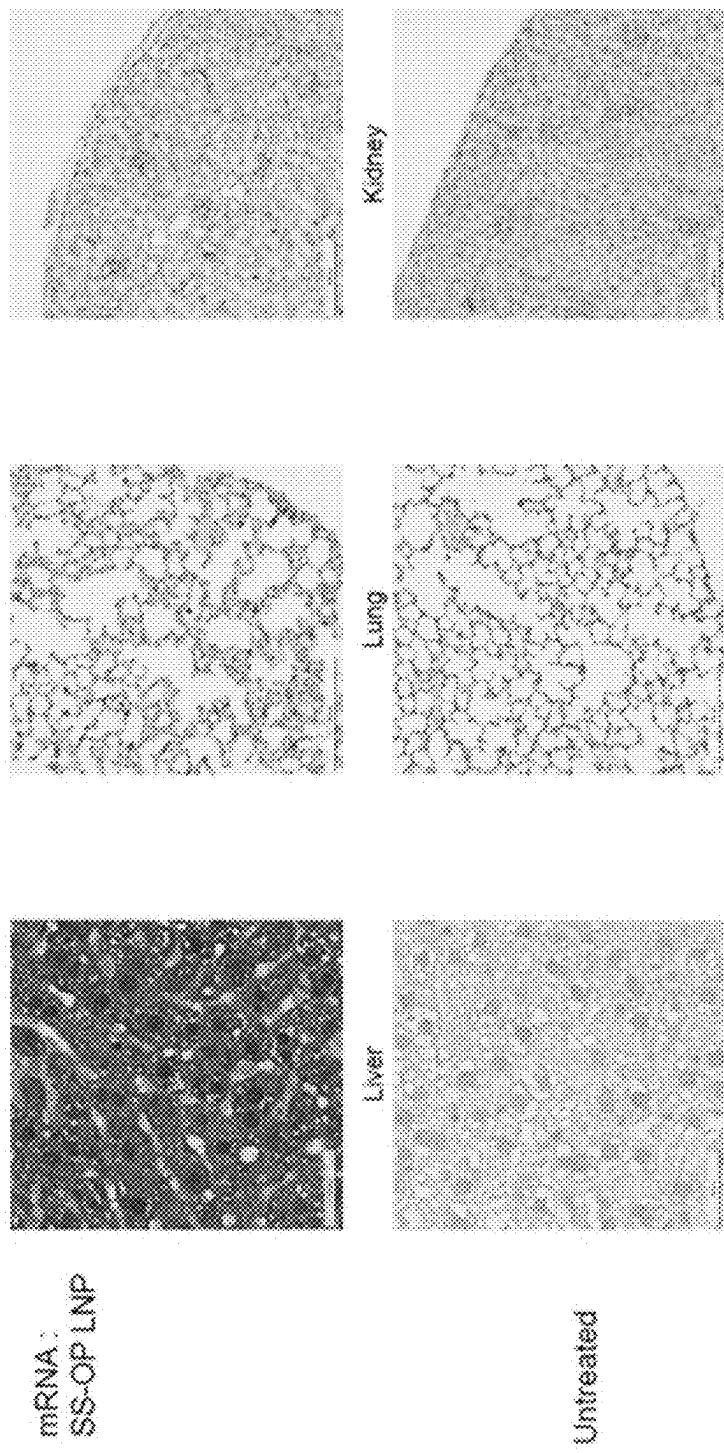


FIG. 4

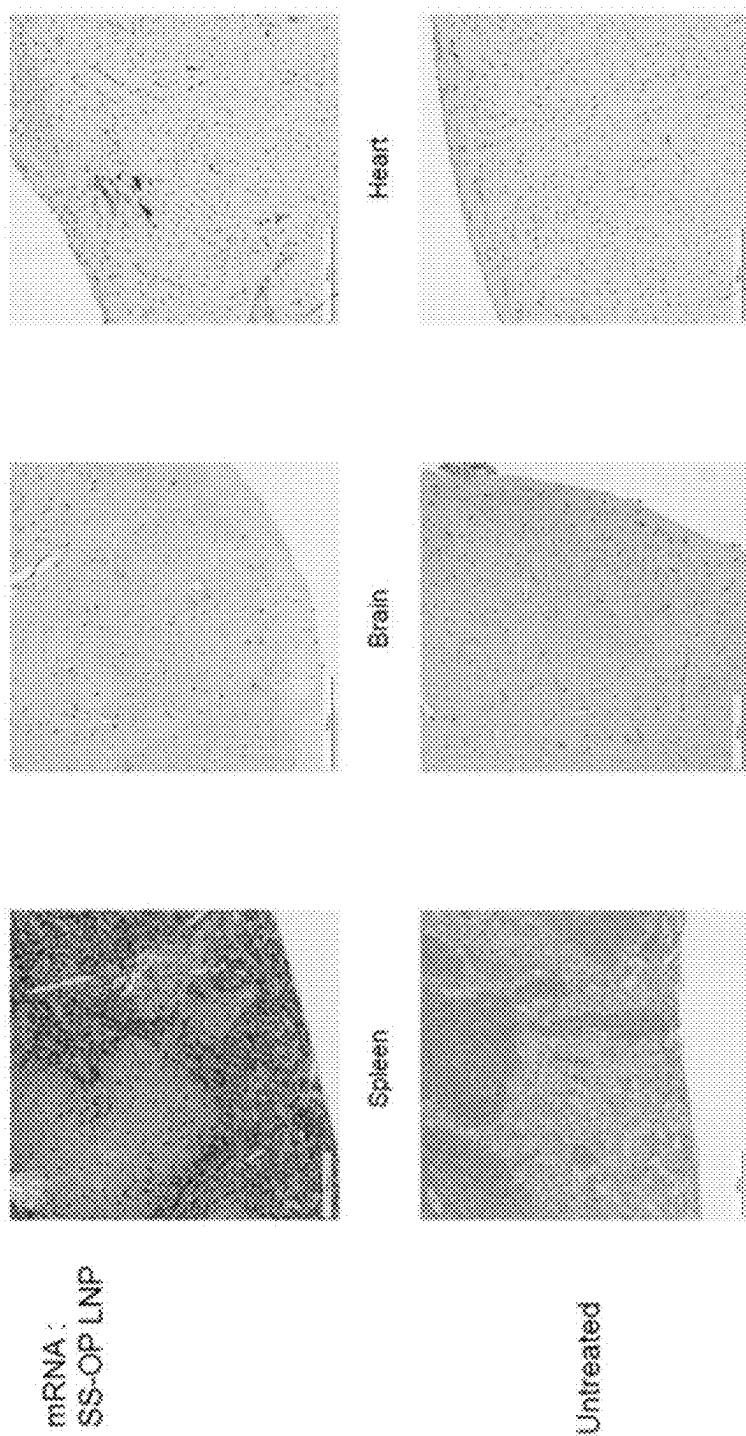


FIG. 5

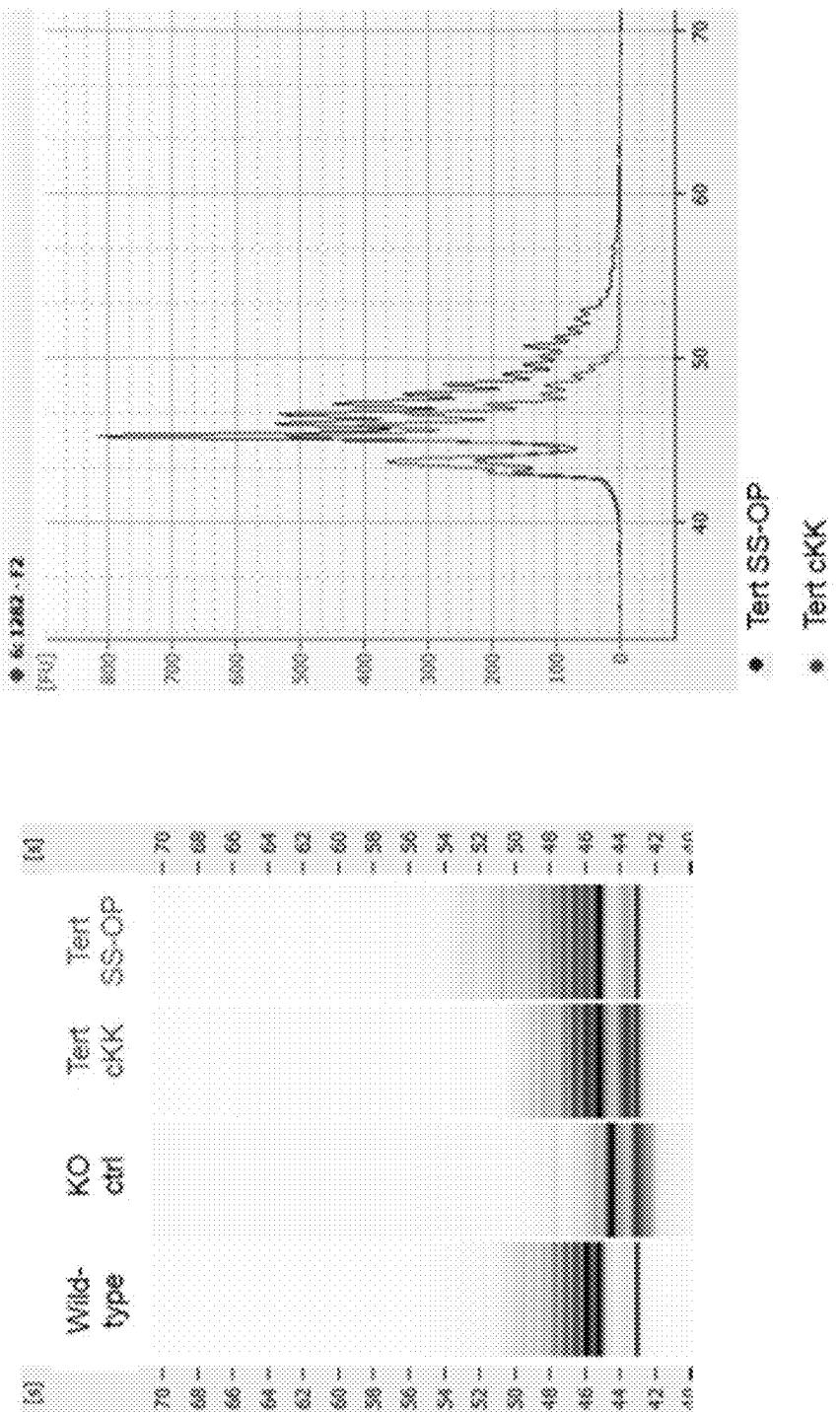


FIG. 6

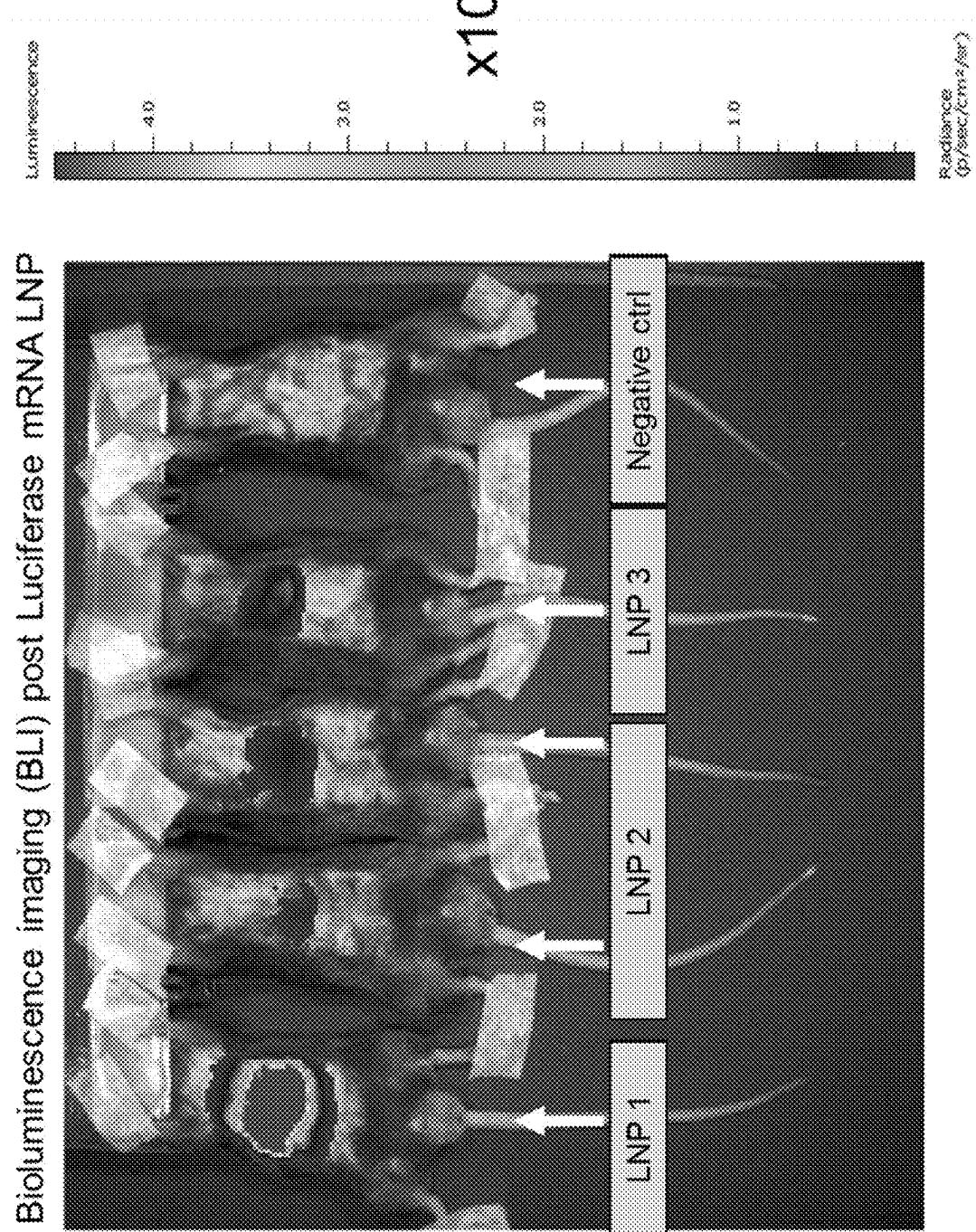


FIG. 7

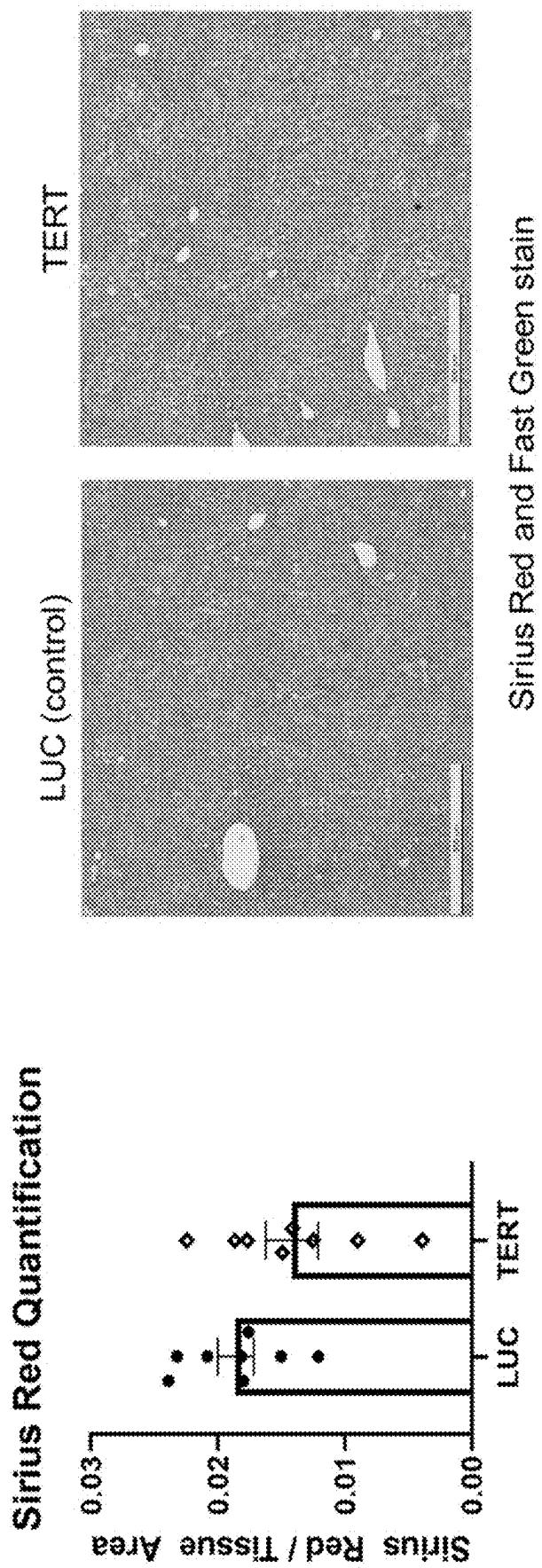


FIG. 8

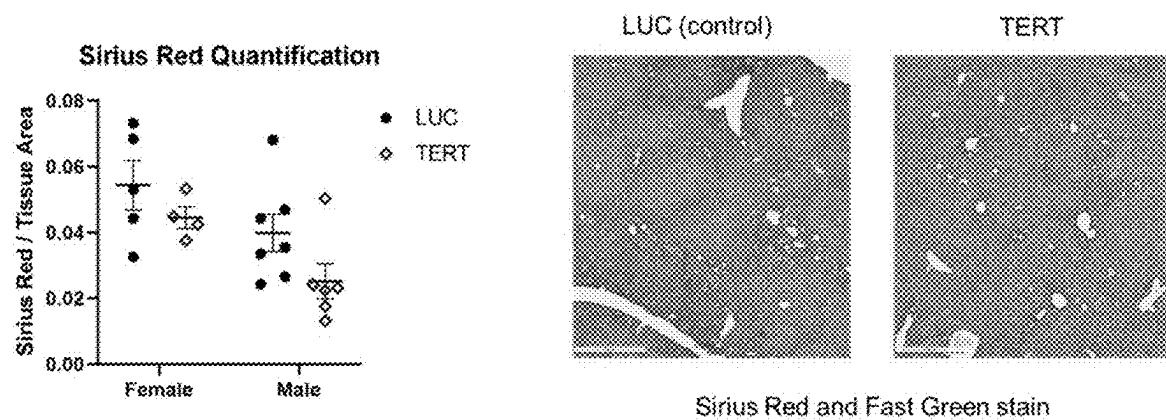


FIG. 9A

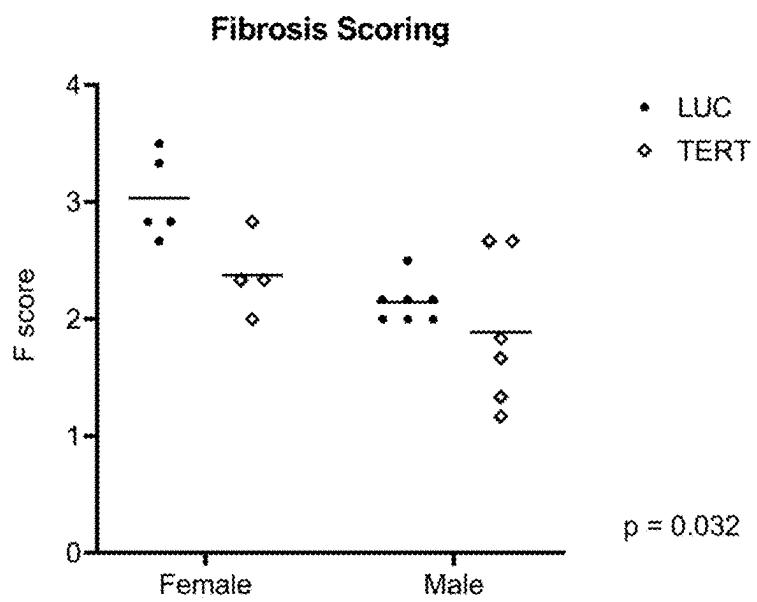


FIG. 9B

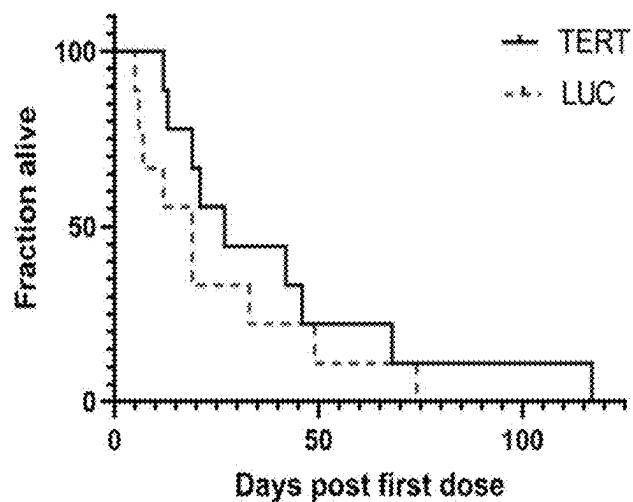


FIG. 10A

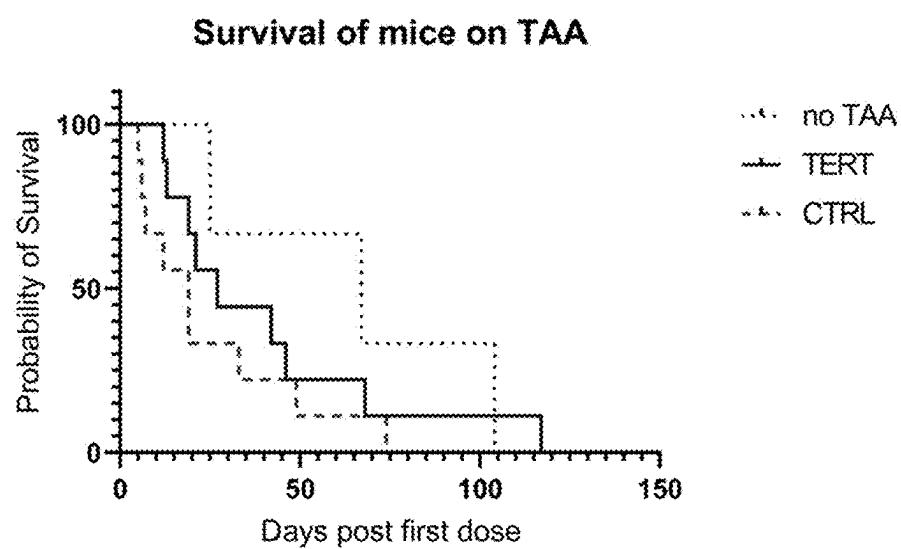


FIG. 10B

Inflammation score in fibrotic livers

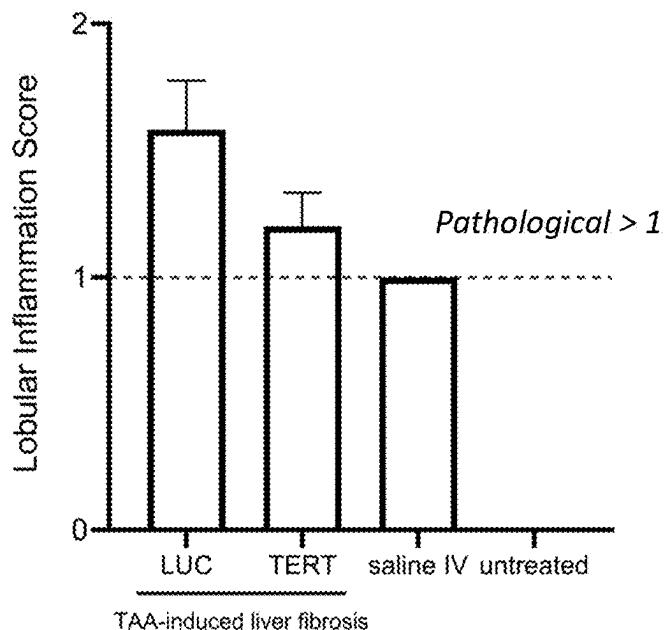


FIG. 11A

Mice with high inflammation in the liver

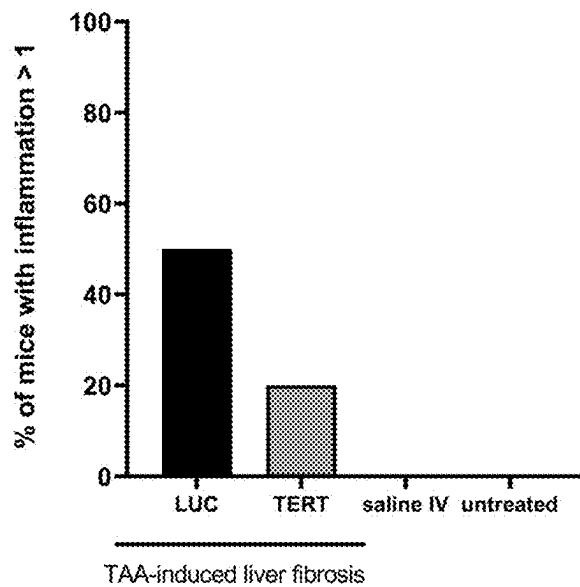


FIG. 11B

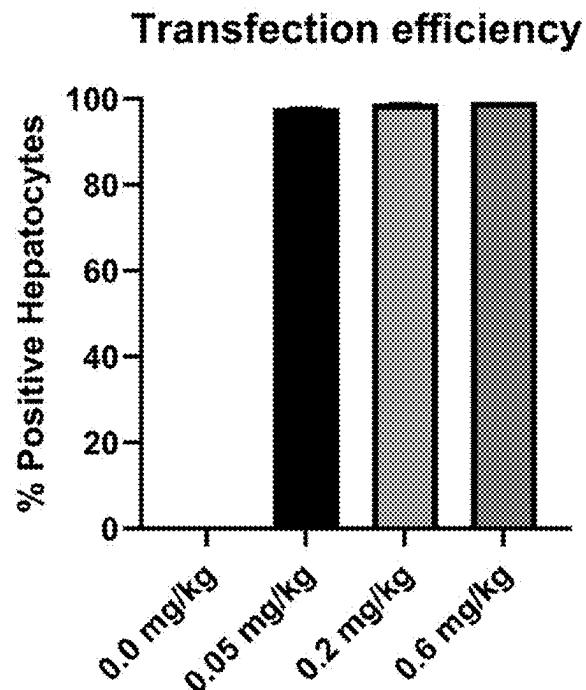
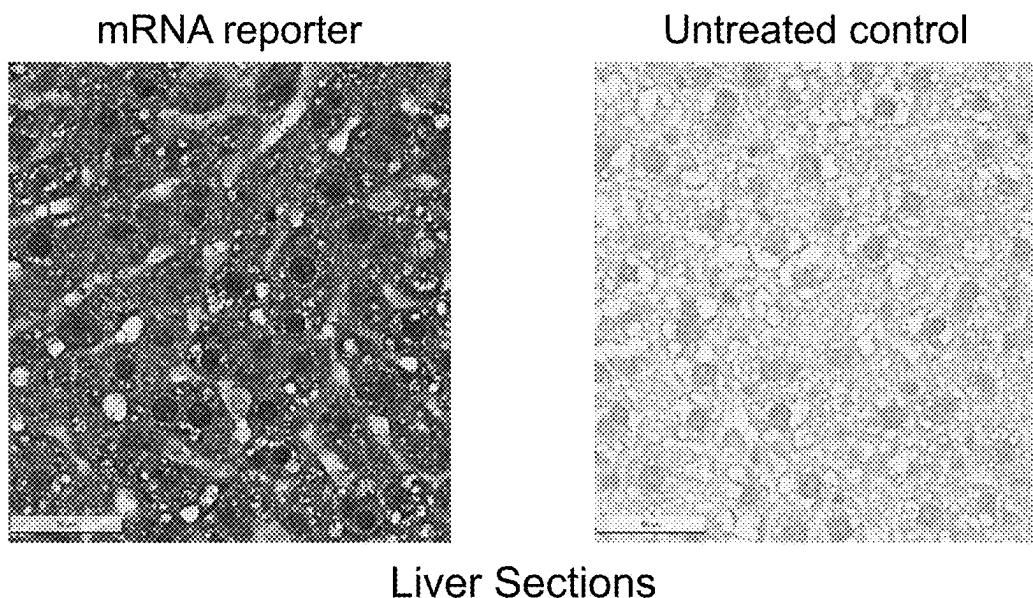


FIG. 12A



Liver Sections

FIG. 12B

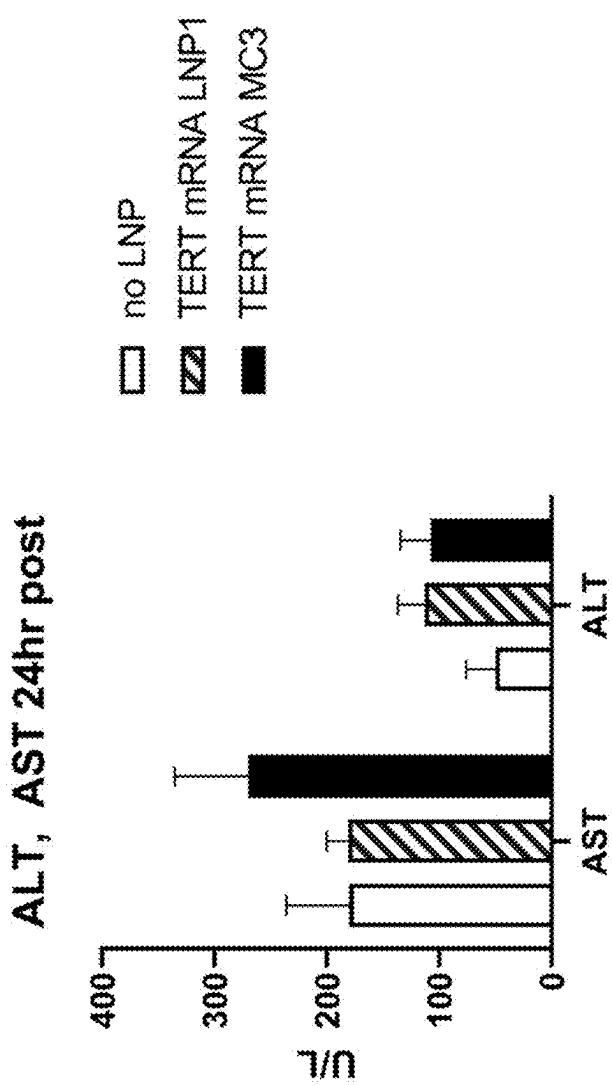


FIG. 13

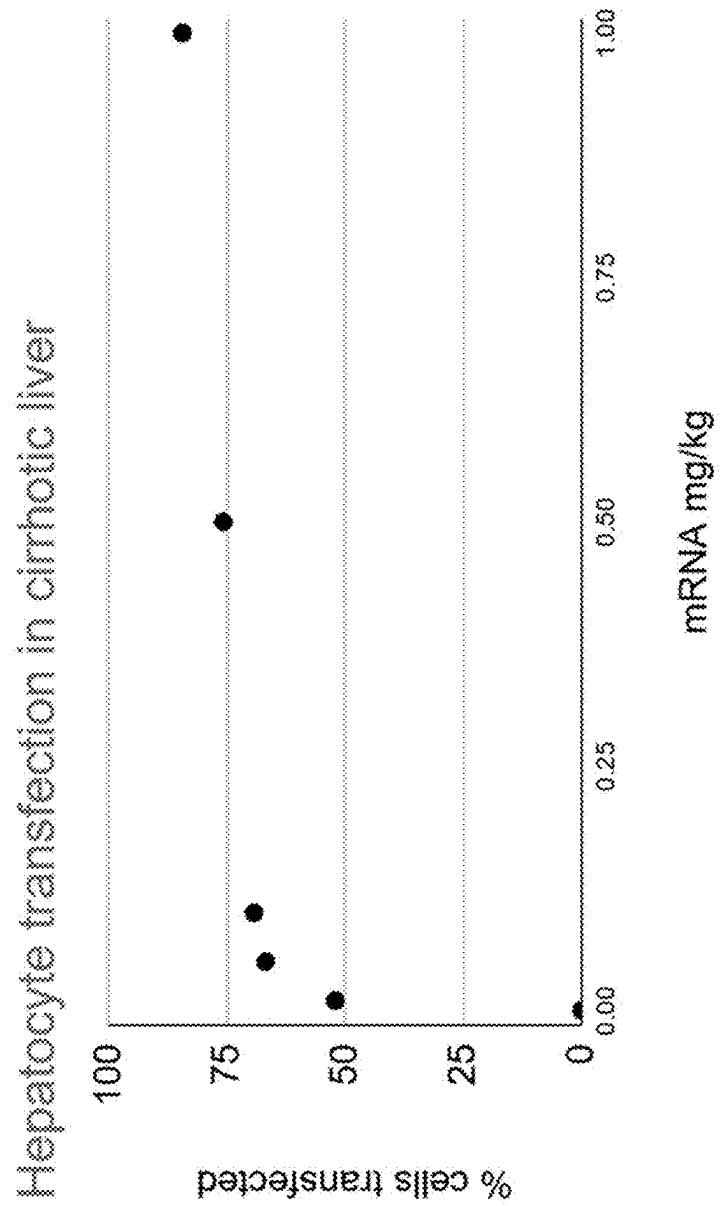


FIG. 14A

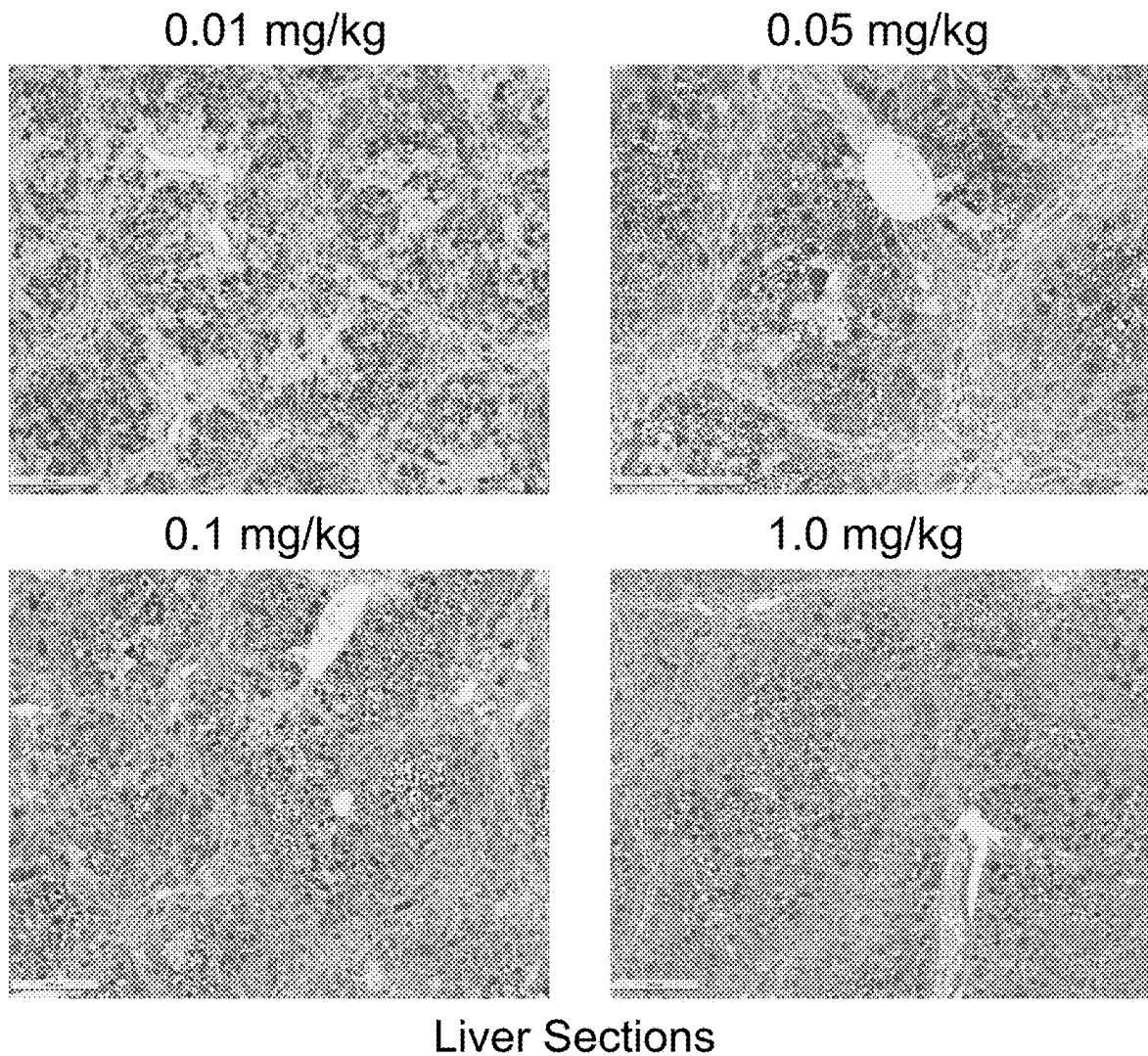


FIG. 14B

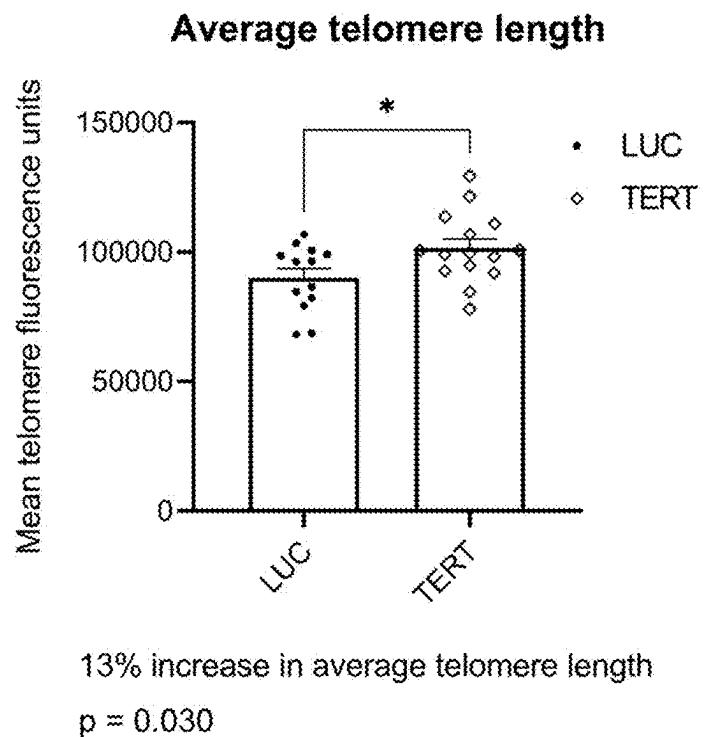
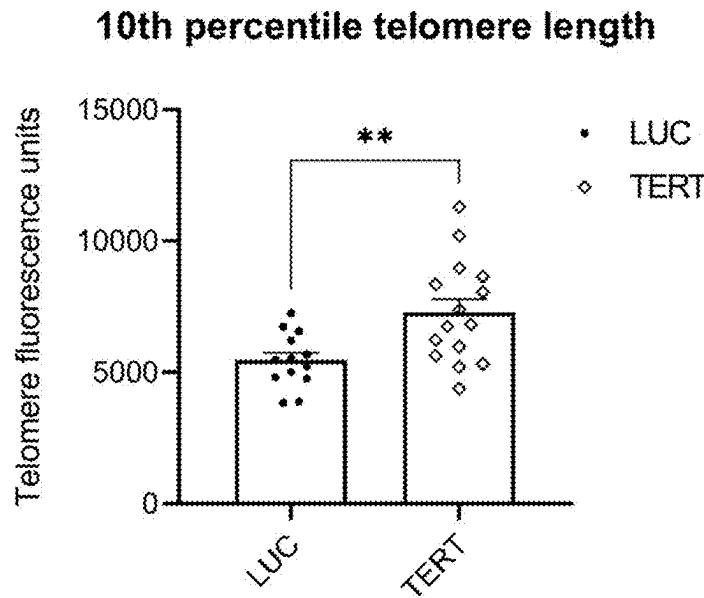


FIG. 15A



33% increase in 10th percentile telomere length
p = 0.006

FIG. 15B

TERT mRNA

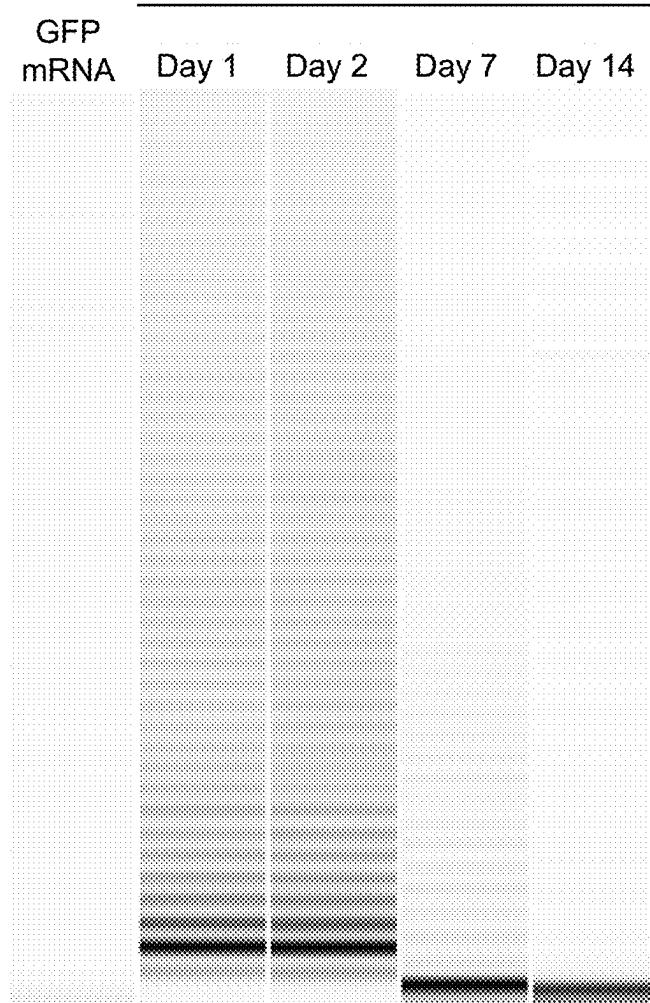
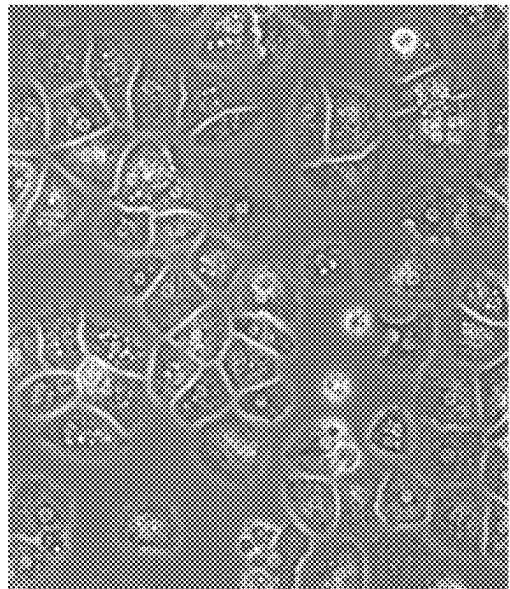


FIG. 16A



Human hepatocytes from 51-year-old donor

FIG. 16B

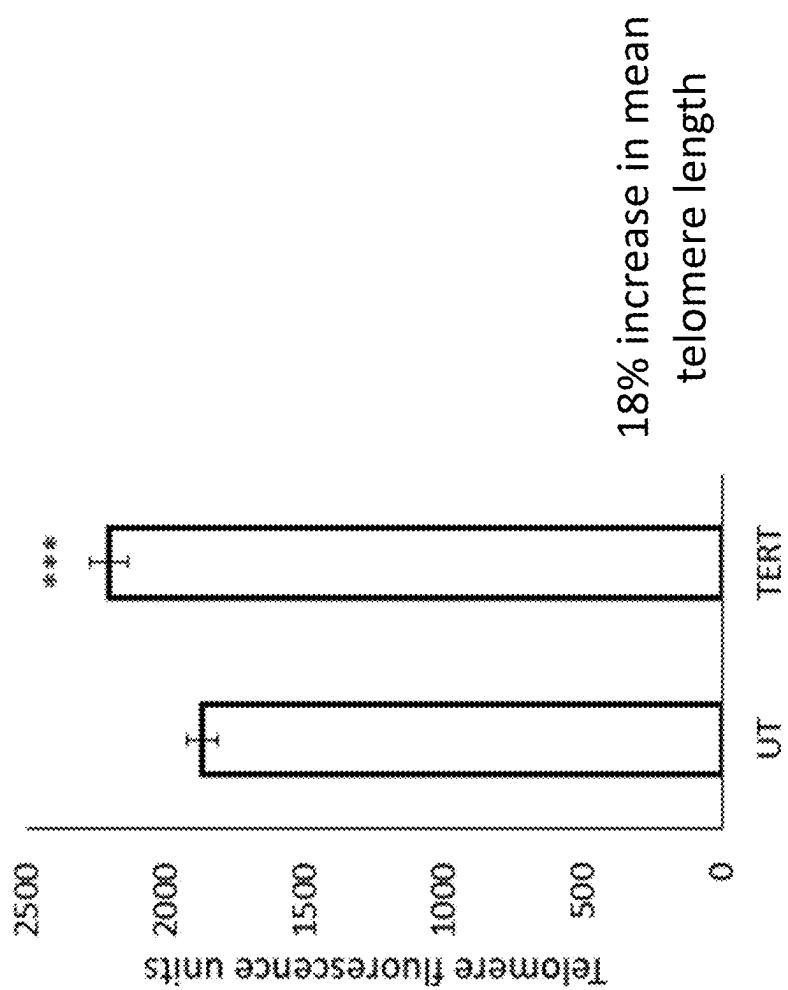


FIG. 17A

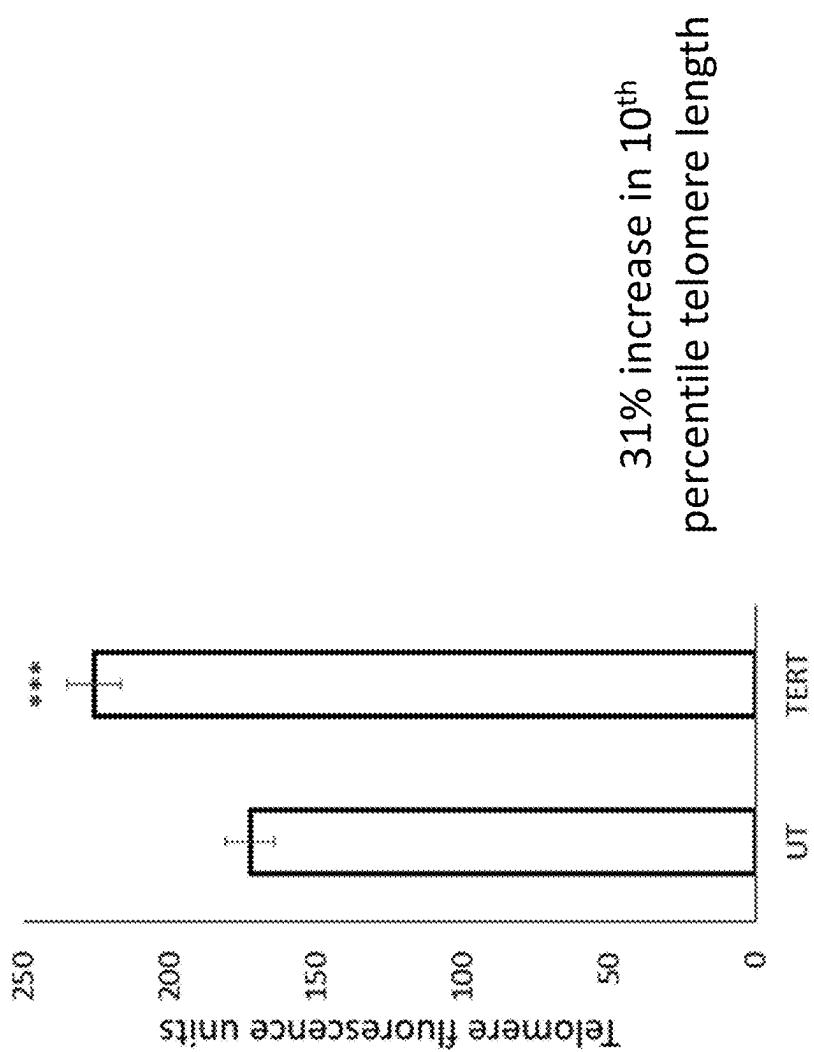


FIG. 17B

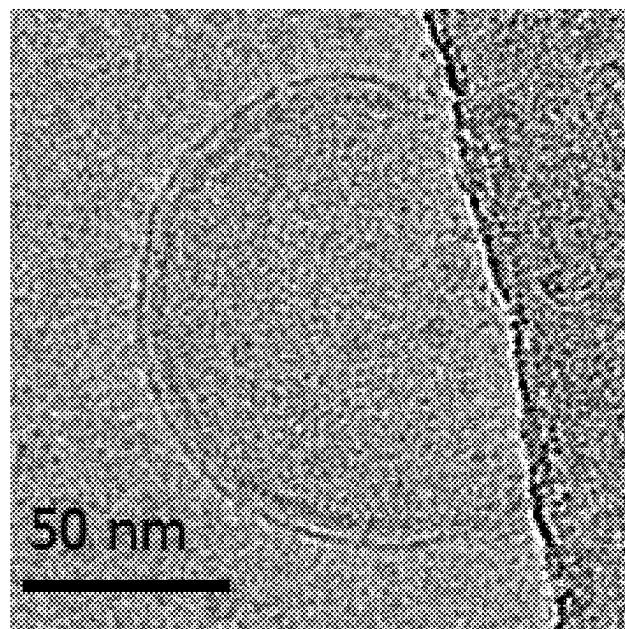


FIG. 18A

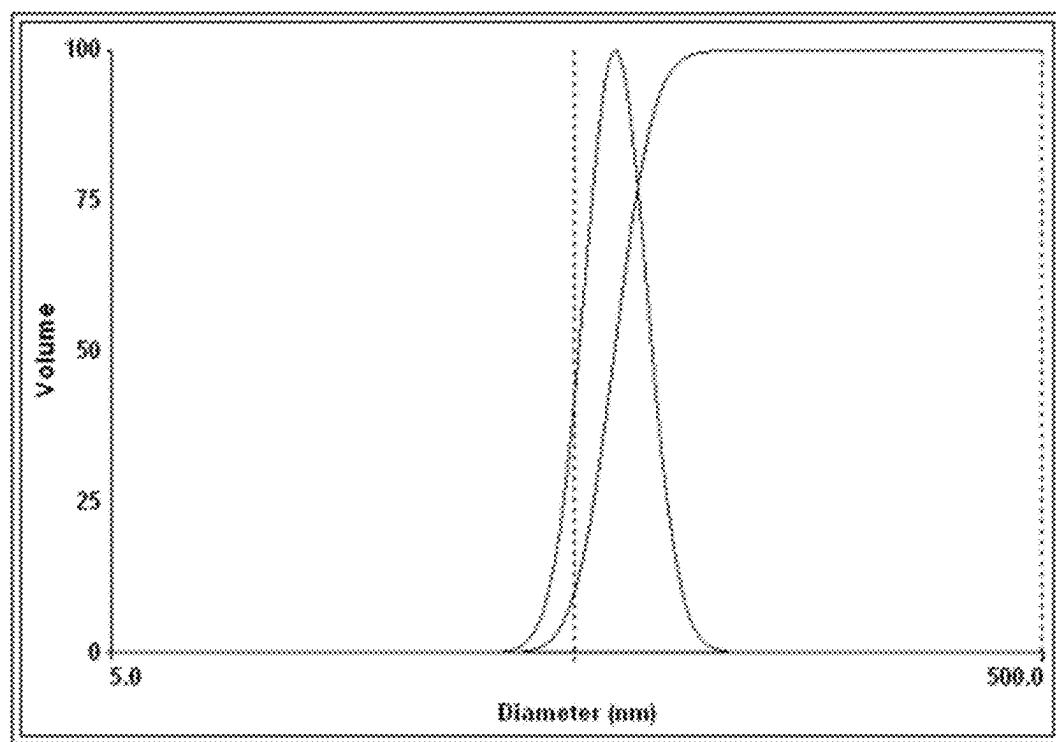


FIG. 18B

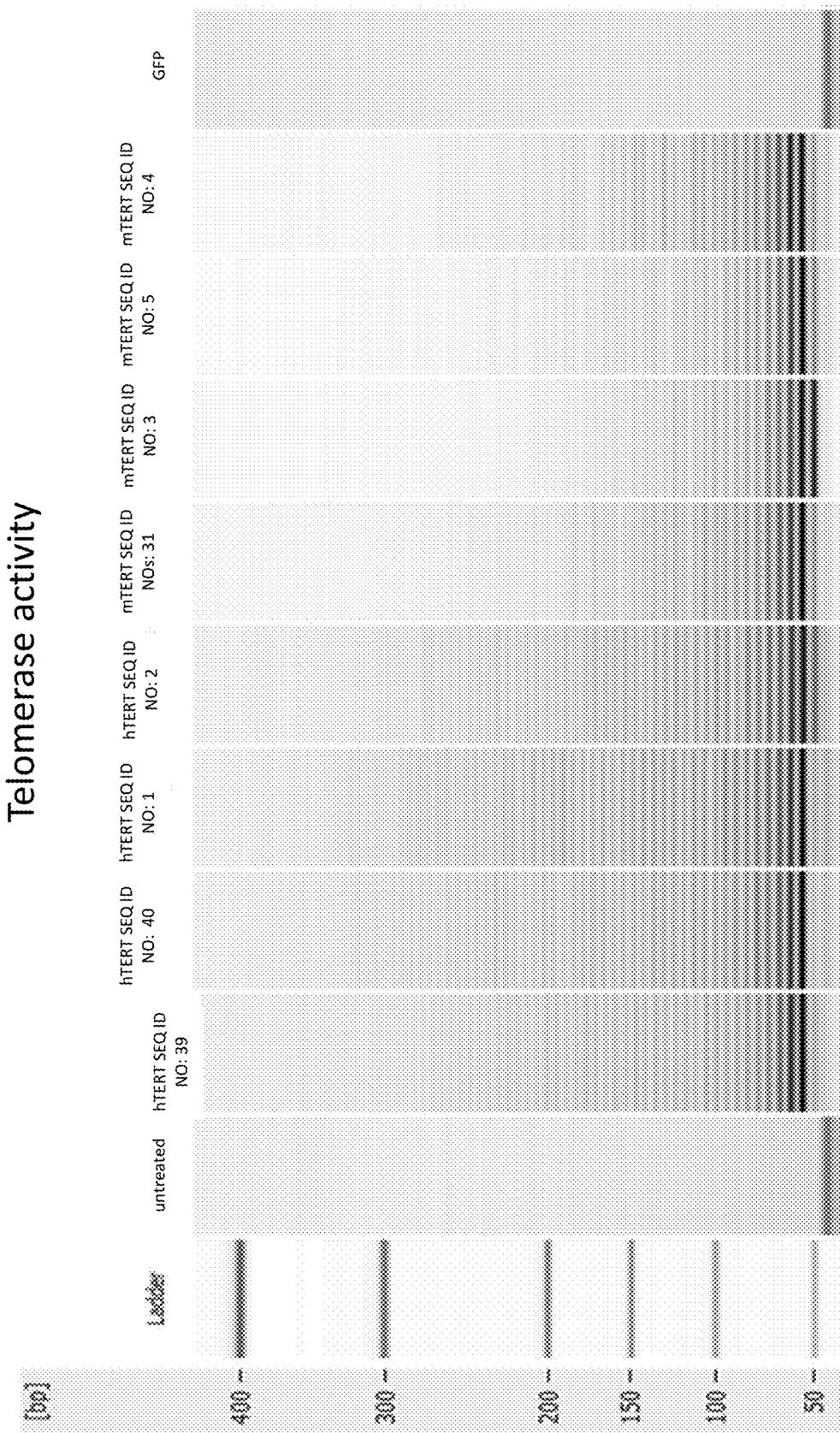


FIG. 19

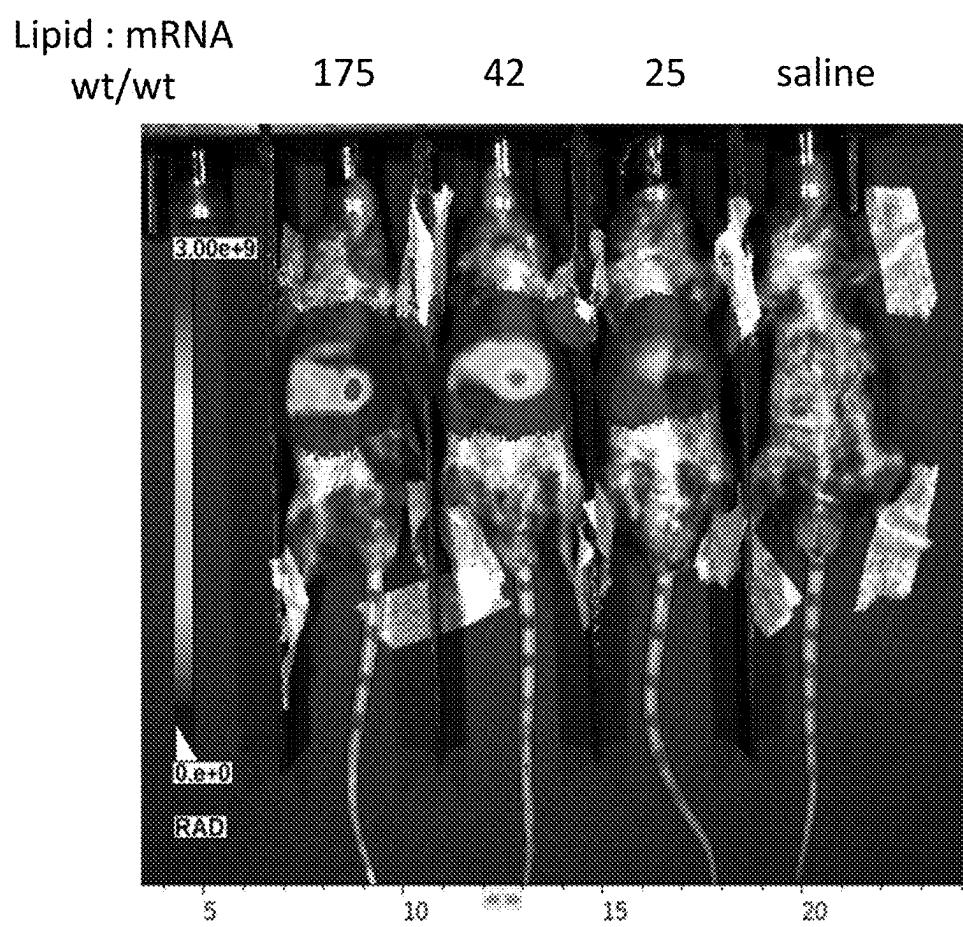


FIG. 20

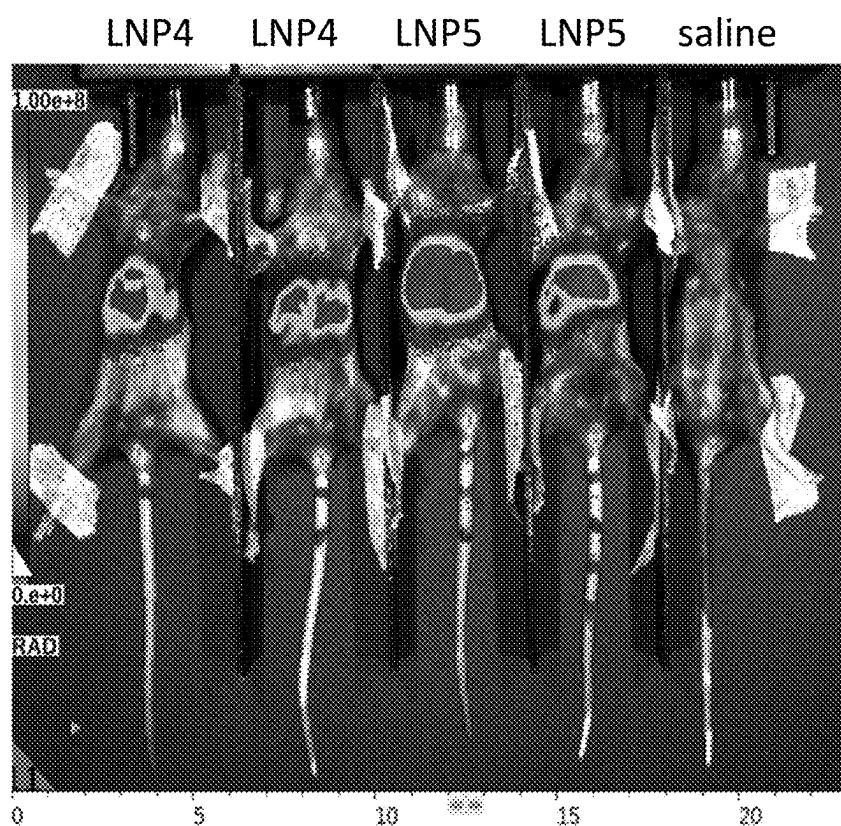


FIG. 21

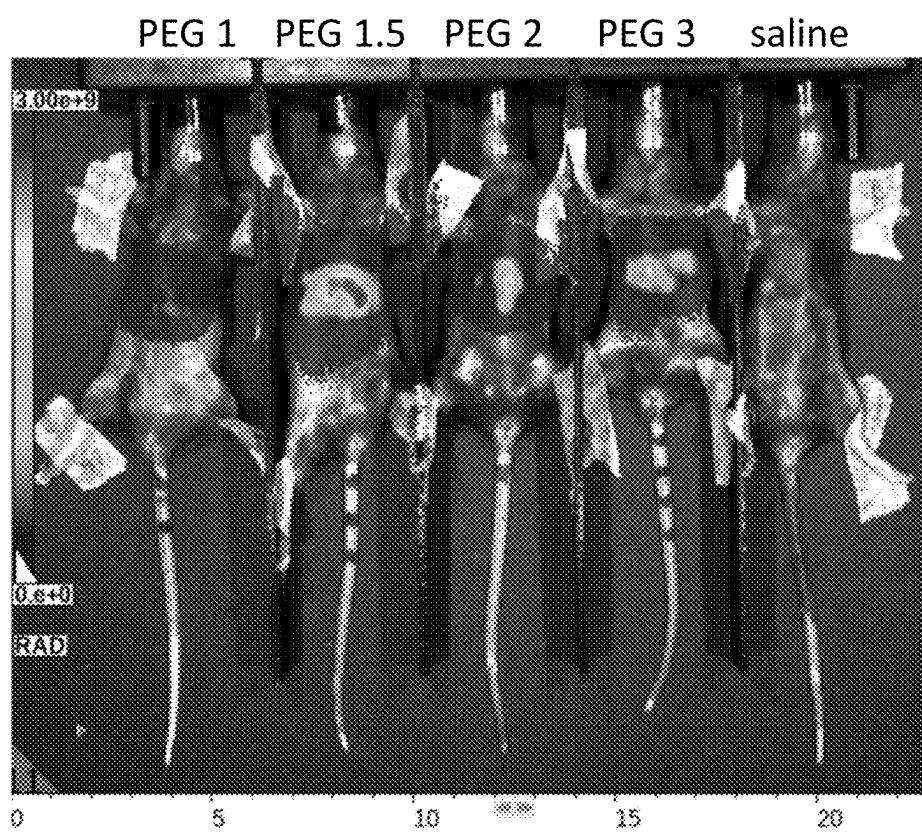


FIG. 22

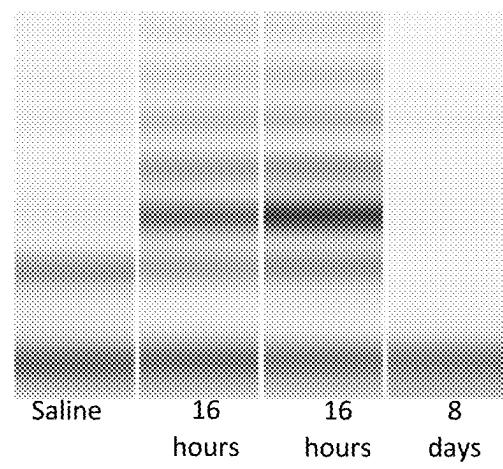


FIG. 23

COMPOSITIONS AND METHODS FOR DELIVERY OF RNA

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to and the benefit of U.S. Provisional Patent Application No. 63/131,528 entitled “COMPOSITIONS AND METHODS FOR DELIVERY OF RNA ENCODING TERT USING LIPIDS,” filed Dec. 29, 2020, the disclosure of which is hereby incorporated by reference in its entirety.

INCORPORATION OF THE SEQUENCE LISTING

[0002] The contents of the text file named “REJU-002_03WO_SeqList_ST25.txt,” which was created on Dec. 28, 2021 and is 205 KB in size, are hereby incorporated by reference in their entirety.

FIELD OF THE DISCLOSURE

[0003] The present disclosure relates generally to telomerase reverse transcriptase (TERT) messenger ribonucleic acid (mRNA) therapies for the treatment of fibrotic diseases and liver diseases.

BACKGROUND

[0004] Drug treatment of fibrotic diseases and liver diseases remains elusive as evidenced by the high mortality rates of these diseases. Currently, cessation of the damaging activity or disease is the primary method for treating fibrosis, e.g., of the liver, and liver disease. Therefore, a need exists for pharmaceutical therapies to treat fibrotic diseases and liver diseases.

SUMMARY

[0005] The disclosure relates to telomerase reverse transcriptase (TERT) messenger ribonucleic acid (mRNA) therapies for the treatment of fibrotic diseases and conditions, e.g. of the liver, and liver diseases and conditions. Treatment with compositions comprising TERT mRNA may prevent, reverse or treat fibrosis and other pathological features of fibrotic disease and/or liver disease leading to improvements in overall organ function and subject health. Accordingly, in some embodiments, provided herein are compositions comprising one or more synthetic messenger ribonucleic acids (mRNAs) encoding telomerase reverse transcriptase (TERT).

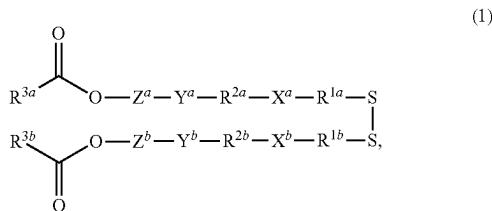
[0006] In some embodiments, the composition comprises: (i) a ribonucleic acid (RNA) encoding telomerase reverse transcriptase (TERT) and (ii) a delivery vehicle, wherein the RNA of (i) comprises one or more modified nucleotides and wherein the delivery vehicle of (ii) is operably-linked to the RNA of (i).

[0007] In some embodiments of the compositions of the disclosure, the delivery vehicle comprises one or more of a nanoparticle, a liposome, a cationic lipid, an exosome, an extracellular vesicle, a lipid nanoparticle, a natural lipoprotein particle, and an artificial lipoprotein particle.

[0008] In some embodiments of the compositions of the disclosure, the delivery vehicle comprises a lipid nanoparticle (LNP). In some embodiments, the delivery vehicle comprises an ionizable and/or cationic lipid.

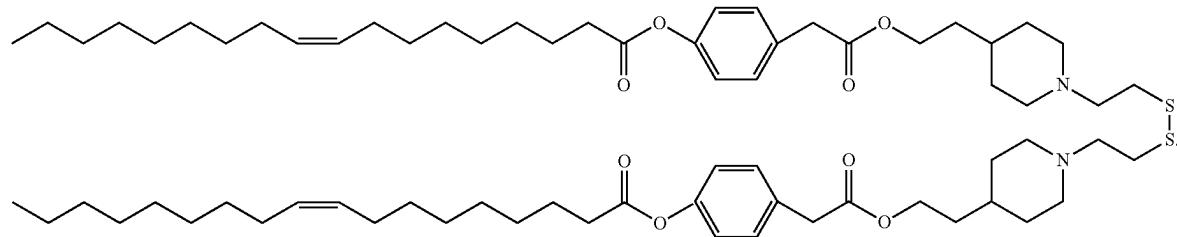
[0009] In some embodiments, the delivery vehicle comprises a targeting moiety. In some embodiments, the targeting moiety results in the delivery vehicle specifically or selectively interacting with a liver cell. In some embodiments, the targeting moiety comprises cholesterol. In some embodiments, the targeting moiety is a lipid, a peptide, and/or an antibody. In some embodiments, the LNP comprises an ionizable lipid, a phospholipid, a cholesterol, and/or a PEGylated lipid. In some embodiments, the LNP comprises a molar ratio of about 30-70 moles of an ionizable lipid, to about 0.1 to about 20 moles of a phospholipid, about 20 to about 60 moles of cholesterol, and about 0.1 to about 5.5 moles of PEGylated lipid.

[0010] In some embodiments of the compositions of the disclosure, including those in which the delivery vehicle is a lipid nanoparticle (LNP), the delivery vehicle comprises a compound of Formula I:

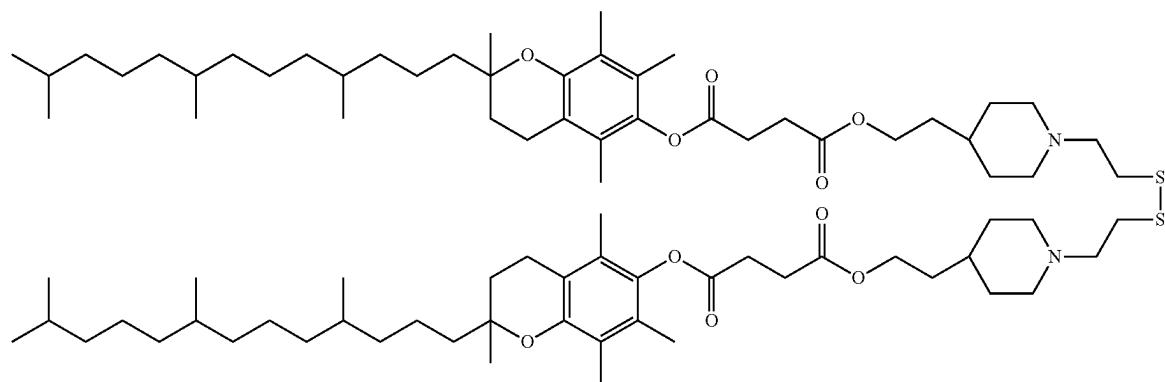


wherein R^{1a} and R^{1b} each independently represents an alkylene group having 1 to 6 carbon atoms, wherein X^a and X^b are each independently an acyclic alkyl tertiary amino group having 1 to 6 carbon atoms and 1 tertiary amino group, or 2 to 5 carbon atoms, and A cyclic alkylene tertiary amino group having 1 to 2 tertiary amino groups, wherein R^{2a} and R^{2b} each independently represent an alkylene group having 8 or less carbon atoms or an oxydialkylene group, wherein Y^a and Y^b each independently represent an ester bond, an amide bond, a carbamate bond, an ether bond or a urea bond; wherein Z^a and Z^b are each independently a divalent group derived from an aromatic compound having 3 to 16 carbon atoms, having at least one aromatic ring, and optionally having a hetero atom, and wherein R^{3a} and R^{3b} each independently represent a residue derived from a reaction product of a fat-soluble vitamin having a hydroxyl group and succinic anhydride or glutaric anhydride, or a sterol derivative having a hydroxyl group and succinic anhydride or a residue derived from a reaction product with glutaric anhydride or an aliphatic hydrocarbon group having 12 to 22 carbon atoms.

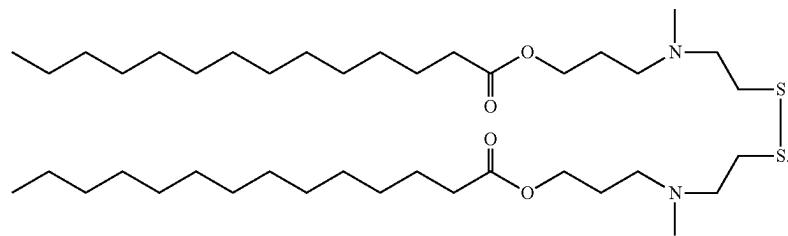
[0011] In some embodiments of the compositions of the disclosure, including those in which the delivery vehicle is a lipid nanoparticle (LNP), the compound of Formula I is:



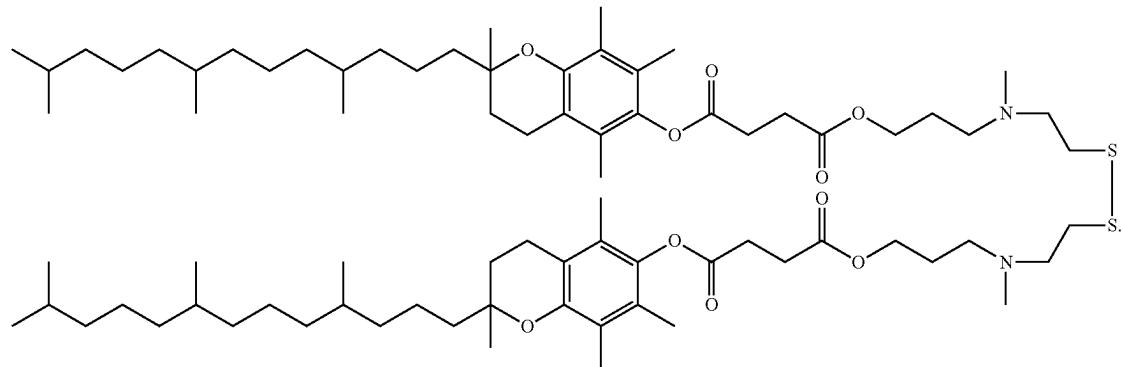
[0012] In some embodiments of the compositions of the disclosure, including those in which the delivery vehicle is a lipid nanoparticle (LNP), the compound of Formula I is:



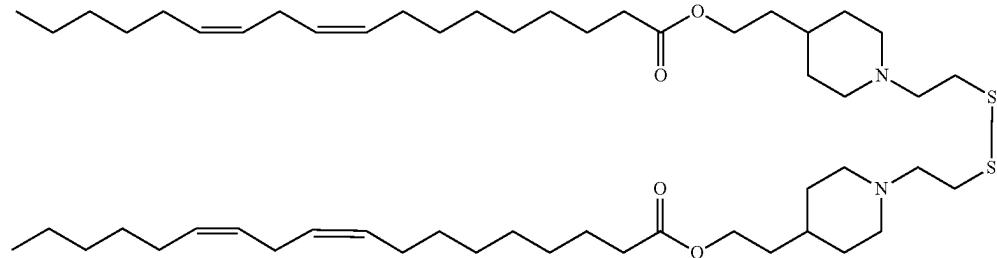
[0013] In some embodiments of the compositions of the disclosure, including those in which the delivery vehicle is a lipid nanoparticle (LNP), the compound of Formula I is:



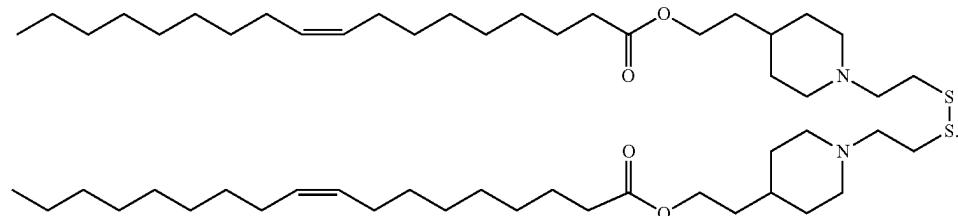
[0014] In some embodiments of the compositions of the disclosure, including those in which the delivery vehicle is a lipid nanoparticle (LNP), the compound of Formula I is:



[0015] In some embodiments of the compositions of the disclosure, including those in which the delivery vehicle is a lipid nanoparticle (LNP), the compound of Formula I is:



[0016] In some embodiments of the compositions of the disclosure, including those in which the delivery vehicle is a lipid nanoparticle (LNP), the compound of Formula I is:



[0017] In some embodiments, the RNA comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to any one of SEQ ID NOS: 1-5, 30-31, or 37-40.

[0018] In some embodiments of the compositions of the disclosure, including those in which the delivery vehicle is a lipid nanoparticle (LNP), the RNA comprises a 5' cap. In some embodiments, the 5'cap comprises an anti-reverse cap analog (ARCA). In some embodiments, the ARCA comprises an 3'-O-Me-m7G(5')ppp(5')G structure. In some embodiments, the 5' cap comprises m7G(5')ppp(5')(2'OMeA)pG. In some embodiments, the 5' cap comprises m7(3'OMeG)(5')ppp(5')(2'OMeA)pG.

[0019] In some embodiments of the compositions of the disclosure, including those in which the delivery vehicle is a lipid nanoparticle (LNP), the RNA further comprises at least one untranslated region (UTR). The UTR may comprise a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to any one of SEQ ID NOS: 32-36. In some embodiments, the at least one UTR is positioned 5' to (i) the RNA encoding TERT. In some embodiments, the at least one UTR is positioned 3' to the RNA of (i). In some embodiments, the UTR comprises a human sequence. In some embodiments, the UTR comprises a non-human or synthetic sequence. In some embodiments, the UTR comprises a chimeric sequence. In some embodiments, the UTR increases stability, increases half-life, increases a transcription rate or decreases a time until initiation of transcription of the RNA of (i). In some embodiments, the UTR comprises a sequence having at least 70% identity to a UTR sequence

isolated or derived from one or more of α -globin, β -globin, c-fos, and a tobacco etch virus.

[0020] In some embodiments of the compositions of the disclosure, including those in which the delivery vehicle is a lipid nanoparticle (LNP), the one or more modified nucleotides of the RNA of (i) comprise one or more of a modified adenine or analog thereof, a modified cytidine or analog thereof, a modified guanosine or analog thereof, and a modified uridine or analog thereof. In some embodiments, the one or more modified nucleotides of the RNA of (i) comprise one or more of 1-methylpseudouridine also known as N1-Methylpseudouridine, pseudouridine (N1m), 2-thiouridine, and 5-methylcytidine. In some embodiments, the one or more modified nucleotides of the RNA of (i) comprise 5-methoxyuridine (5-moU). In some embodiments, the one or more modified nucleotides of the RNA of (i) comprise one or more of m1A 1-methyladenosine, m6A N6-methyladenosine, Am 2'-O-methyladenosine, i6A N6-isopentenyladenosine, io6A N6-(cis-hydroxyisopentenyl)adenosine, ms2io6A 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine, g6A N6-glycylcarbamoyladenosine, t6A N6-threonylcarbamoyladenosine, ms2t6A 2-methylthio-N6-threonyl carbamoyladenosine, Ar(p) 2'-O-ribosyladenosine (phosphate), m6 2A N6,N6-dimethyladenosine, m6Am N6,2'-O-dimethyladenosine, m6 2Am N6,N6,2'-O-trimethyladenosine, m1Am 1,2'-O-dimethyladenosine, m3C 3-methylcytidine, m5C 5-methylcytidine, Cm 2'-O-methylcytidine, ac4C N4-acetylcytidine, f5C 5-formylcytidine, m4C N4-methylcytidine, hm5C 5-hydroxymethylcytidine, f5Cm 5-formyl-2'-O-methylcytidine, m1G 1-methylguanosine, m2G N2-methylguanosine, m7G 7-methylguanosine, Gm 2'-O-methylguanosine, m2 2G N2,N2-dimethylguanos-

ine, Gr(p) 2'-O-ribosylguanosine (phosphate), yW wybutsine, o2yW peroxywybutosine, OHyW hydroxywybutosine, OHyW* undermodified hydroxywybutosine, imG wyosine, m2,7G N2,7-dimethylguanosine, m2,2,7G N2,N2,7-trimethylguanosine I inosine, m1I 1-methylinosine, Im 2'-O-methylinosine, Q queuosine, galQ galactosyl-queuosine, manQ mannosyl-queuosine, Ψ pseudouridine, D dihydrouridine, m5U 5-methyluridine, Um 2'-O-methyluridine, m5Um 5,2'-O-dimethyluridine, m1Ψ 1-methylpseudouridine, Ψm 2'-O-methylpseudouridine, s2U 2-thiouridine, ho5U 5-hydroxyuridine, chm5U 5-(carboxyhydroxymethyl)uridine, mchm5U 5-(carboxyhydroxymethyl)uridine, methyl ester mcm5U 5-methoxycarbonylmethyluridine, mcm5Um 5-methoxycarbonylmethyl-2'-O-methyluridine, mcm5s2U 5-methoxycarbonylmethyl-2-thiouridine, nc5U 5-carbamoylmethyluridine, ncm5Um 5-carbamoylmethyl-2'-O-methyluridine, cmnm5U 5-carboxymethylaminomethyluridine, m3U 3-methyluridine, m1acp3Ψ 1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine, cm5U 5-carboxymethyluridine, m3Um 3,2'-O-dimethyluridine, m5D 5-methylidihydrouridine, TM5U 5-taurinomethyluridine, tm5s2U 5-taurinomethyl-2-thiouridine, 2-Aminoadenosine, 2-Amino-6-chloropurineriboside, 8-Azaadenosine, 6-Chloropurineriboside, 5-Iodocytidine, 5-Iodouridine, Inosine, 2'-O-Methylinosine, Xanthosine, 4-Thiouridine, 06-Methylguanosine, 5,6-Dihydrouridine, 2-Thiacytidine, 6-Azacytidine, 6-Azauridine, 2'-O-Methyl-2-aminoadenosine, 2'-O-Methylpseudouridine, N1-Methyladenosine, 2'-O-Methyl-5-methyluridine, 7-Deazaguanosine, 8-Azidoadenosine, 5-Bromocytidine, 5-Bromouridine, 7-Deazaadenosine, 5-Aminoallyluridine, 5-Aminoallylcytidine, 8-Oxoguanosine, 2-Aminopurine-riboside, Pseudoisocytidine, N1-Methylpseudouridine, 5,6-Dihydro-5-Methyluridine, N6-Methyl-2-Aminoadenosine, 5-Carboxycytidine, 5-Hydroxymethyluridine, Thienoguanosine, 5-Hydroxycytidine, 5-Formyluridine, 5-Carboxyuridine, 5-Methoxyuridine, 5-Methoxycytidine, Thienouridine, 5-Carboxymethyl-esteruridine, Thienocytidine, 8-Oxaadenosine, Isoguanosine, N1-Ethylpseudouridine, N1-Methyl-2'-O-Methylpseudouridine, N1-Methoxymethylpseudouridine, N1-Propylpseudouridine, 2'-O-Methyl-N6-Methyladenosine, 2-Amino-6-Cl-purine-2'-deoxyriboside, 2-Amino-2'-deoxyadenosine, 2-Aminopurine-2'-deoxyriboside, 5-Bromo-2'-deoxycytidine, 5-Bromo-2'-deoxyuridine, 6-Chloropurine-2'-deoxyriboside, 7-Deaza-2'-deoxyadenosine, 7-Deaza-2'-deoxyguanosine, 2'-Deoxyinosine, 5-Propynyl-2'-deoxycytidine, 5-Propynyl-2'-deoxyuridine, 5-Fluoro-2'-deoxyuridine, 5-Iodo-2'-deoxycytidine, 5-Iodo-2'-deoxyuridine, N6-Methyl-2'-deoxyadenosine, 5-Methyl-2'-deoxycytidine, 06-Methyl-2'-deoxyguanosine, N2-Methyl-2'-deoxyguanosine, 8-Oxo-2'-deoxyadenosine, 8-Oxo-2'-deoxyguanosine, 2-Thiothymidine, 2'-Deoxy-P-nucleoside, 5-Hydroxy-2'-deoxycytidine, 4-Thiothymidine, 2-Thio-2'-deoxycytidine, 6-Aza-2'-deoxyuridine, 6-Thio-2'-deoxyguanosine, 8-Chloro-2'-deoxyadenosine, 5-Aminoallyl-2'-deoxycytidine, 5-Aminoallyl-2'-deoxyuridine, N4-Methyl-2'-deoxycytidine, 2'-Deoxyzebularine, 5-Hydroxymethyl-2'-deoxyuridine, 5-Hydroxymethyl-2'-deoxycytidine, 5-Propargylamino-2'-deoxycytidine, 5-Propargylamino-2'-deoxyuridine, 5-Carboxy-2'-deoxycytidine, 5-Formyl-2'-deoxycytidine, 5-[3-Indolyl]propionamide-N-allyl]-2'-deoxyuridine, 5-Carboxy-2'-deoxyuridine, 5-Formyl-2'-deoxyuridine, 7-Deaza-7-Propargylamino-2'-deoxyadenosine, 7-Deaza-7-Propargylamino-2'-deox-

guanosine, Biotin-16-Aminoallyl-2'-dUTP, Biotin-16-Aminoallyl-2'-dCTP, Biotin-16-Aminoallylcytidine, N4-Biotin-OBEA-2'-deoxycytidine, Biotin-16-Aminoallyluridine, DabcyL-5-3-Aminoallyl-2'-dUTP, Desthiobiotin-6-Aminoallyl-2'-deoxycytidine, Desthiobiotin-16-Aminoallyl-Uridine, Biotin-16-7-Deaza-7-Propargylamino-2'-deoxyguanosine, Cyanine 3-5-Propargylamino-2'-deoxycytidine, Cyanine 3-6-Propargylamino-2'-deoxyuridine, Cyanine 5-6-Propargylamino-2'-deoxycytidine, Cyanine 5-6-Propargylamino-2'-deoxyuridine, Cyanine 3-Aminoallylcytidine, Cyanine 3-Aminoallyluridine, Cyanine 5-Aminoallylcytidine, Cyanine 5-Aminoallyluridine, Cyanine 7-Aminoallyluridine, 2'-Fluoro-2'-deoxyadenosine, 2'-Fluoro-2'-deoxycytidine, 2'-Fluoro-2'-deoxyguanosine, 2'-Fluoro-2'-deoxyuridine, 2'-O-Methyladenosine, 2'-O-Methylcytidine, 2'-O-Methylguanosine, 2'-O-Methyluridine, Puromycin, 2'-Amino-2'-deoxycytidine, 2'-Amino-2'-deoxyuridine, 2'-Azido-2'-deoxycytidine, 2'-Azido-2'-deoxyuridine, Araacytidine, Arauridine, 2'-Azido-2'-deoxyadenosine, 2'-Amino-2'-deoxyadenosine, Araadenosine, 2'-Fluoro-thymidine, 3'-O-Methyladenosine, 3'-O-Methylcytidine, 3'-O-Methylguanosine, 3'-O-Methyluridine, 2'-Azido-2'-deoxyguanosine, Araguanosine, 2'-Deoxyuridine, 3'-O-(2-nitrobenzyl)-2'-Deoxyadenosine, 3'-O-(2-nitrobenzyl)-2'-Deoxyinosine, 3'-Deoxyadenosine, 3'-Deoxyguanosine, 3'-Deoxycytidine, 3'-Deoxy-5-Methyluridine, 3'-Deoxyuridine, 2',3'-Dideoxyadenosine, 2',3'-Dideoxyguanosine, 2',3'-Dideoxyuridine, 2',3'-Dideoxythymidine, 2',3'-Dideoxycytidine, 3'-Azido-2',3'-dideoxyadenosine, 3'-Azido-2',3'-dideoxythymidine, 3'-Amino-2',3'-dideoxyadenosine, 3'-Amino-2',3'-dideoxycytidine, 3'-Amino-2',3'-dideoxyuridine, 3'-Azido-2',3'-dideoxyuridine, 5-Bromo-2',3'-dideoxyuridine, 2',3'-Dideoxyinosine, 2'-Deoxyadenosine-5'-O-(1-Thiophosphate), 2'-Deoxycytidine-5'-O-(1-Thiophosphate), 2'-Deoxythymidine-5'-O-(1-Thiophosphate), Adenosine-5'-O-(1-Thiophosphate), Cytidine-5'-O-(1-Thiophosphate), Guanosine-5'-O-(1-Thiophosphate), Uridine-5'-O-(1-Thiophosphate), 2',3'-Dideoxyadenosine-5'-O-(1-Thiophosphate), 2',3'-Dideoxycytidine-5'-O-(1-Thiophosphate), 2',3'-Dideoxyguanosine-5'-O-(1-Thiophosphate), 3'-Deoxythymidine-5'-O-(1-Thiophosphate), 3'-Azido-2',3'-dideoxythymidine-5'-O-(1-Thiophosphate), 2',3'-Dideoxyuridine-5'-O-(1-Thiophosphate), 2'-Deoxyadenosine-5'-O-(1-Boranophosphate), 2'-Deoxycytidine-5'-O-(1-Boranophosphate), 2'-Deoxyguanosine-5'-O-(1-Boranophosphate), and 2'-Deoxythymidine-5'-O-(1-Boranophosphate).

[0021] In some embodiments of the compositions of the disclosure, including those in which the delivery vehicle is a lipid nanoparticle (LNP), the delivery vehicle comprises the RNA encoding TERT. In some embodiments, one or more of a surface, a layer or a volume of the delivery vehicle comprises the RNA encoding TERT. In some embodiments, the surface comprises an outer surface or an inner surface. In some embodiments, the layer comprises a lipid monolayer or lipid bi-layer. In some embodiments, the volume comprises an internal volume.

[0022] In some embodiments, the disclosure provides a composition comprising a (i) a ribonucleic acid (RNA) encoding telomerase reverse transcriptase (TERT) and (ii) a

delivery vehicle, wherein the RNA of (i) comprises one or more modified nucleotides and wherein the delivery vehicle of (ii) is operably-linked to the RNA of (i).

[0023] In some embodiments of the compositions of the disclosure, including those in which the delivery vehicle is a lipid nanoparticle (LNP), the composition further comprises a ribonucleic acid (RNA) encoding Telomerase RNA Component (TERC). In some embodiments, the delivery vehicle is operably-linked to a ribonucleic acid (RNA) encoding Telomerase RNA Component (TERC). In some embodiments, the delivery vehicle comprises the RNA encoding TERC. In some embodiments, one or more of a surface, a layer or a volume of the delivery vehicle comprises the RNA encoding TERC. In some embodiments, the surface comprises an outer surface or an inner surface. In some embodiments, the layer comprises a lipid monolayer or lipid bi-layer. In some embodiments, the volume comprises an internal volume.

[0024] In some embodiments the RNA encoding TERT comprises a sequence with at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% sequence identity to any one of SEQ ID NOS: 1-5, 7, 9, 14-17, 19, 21, 23, 25, 27, 29-31, 37-40. In some embodiments, the RNA encoding TERT comprises a UTR sequence, optionally a UTR sequence with at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% sequence identity to any one of SEQ ID NOS: 32-34, 35, and 36.

[0025] In some embodiments, the RNA comprises a self-replicating RNA. In some embodiments, the RNA comprises a circular RNA.

[0026] The disclosure provides a method of increasing telomerase activity in a cell, the method comprising contacting the cell and the composition of the disclosure. In some embodiments, the cell is *in vivo*, *ex vivo* or *in vitro*.

[0027] The disclosure provides a method of extending telomeres in a cell, the method comprising contacting the cell and the composition of the disclosure. In some embodiments, the cell is *in vivo*, *ex vivo* or *in vitro*.

[0028] The disclosure provides a cell comprising the composition of the disclosure.

[0029] The disclosure provides a formulation comprising the cell of the disclosure, which comprises a composition of the disclosure. In some embodiments of the formulation, a plurality of cells comprises a cell of the disclosure, which comprises a composition of the disclosure. In some embodiments of the formulation, each cell of the plurality is a cell of the disclosure, which comprises a composition of the disclosure.

[0030] The disclosure provides a method of treating a disease or disorder comprising administering to a subject an effective amount of a composition of the disclosure.

[0031] The disclosure provides a method of treating a disease or disorder comprising administering to a subject an effective amount of a cell of the disclosure, which comprises a composition of the disclosure.

[0032] The disclosure provides a method of treating a disease or disorder comprising administering to a subject an effective amount of a formulation of the disclosure.

[0033] The disclosure provides a method of delaying the onset of a disease comprising administering to a subject an effective amount of a composition of the disclosure.

[0034] The disclosure provides a method of delaying the onset of a disease comprising administering to a subject an

effective amount of a cell of the disclosure, which comprises a composition of the disclosure.

[0035] The disclosure provides a method of delaying the onset of a disease comprising administering to a subject an effective amount of a formulation of the disclosure.

[0036] In some embodiments, the disclosure provides a method of treating a fibrotic disease in a subject in need thereof, comprising: administering to the subject an effective amount of a composition comprising one or more synthetic messenger ribonucleic acids (mRNAs) encoding telomerase reverse transcriptase (TERT).

[0037] In some embodiments of the method, the composition comprises a delivery vehicle, optionally wherein the delivery vehicle is a nanoparticle, optionally a lipid nanoparticle (LNP). In some embodiments, the LNP comprises an ionizable lipid, a phospholipid, a cholesterol, and/or a PEGylated lipid.

[0038] In some embodiments, the LNP comprises a molar ratio of about 50 to about 60 moles of an ionizable lipid, to about 4 to about 6 moles of a phospholipid, about 35 to about 45 moles of cholesterol, and about 1.0 to about 2.0 moles of PEGylated lipid.

[0039] In some embodiments, the LNP comprises a molar ratio of about 30 to 40 moles of an ionizable lipid, to about 14 to about 18 moles of a phospholipid, about 40 to about 50 moles of a cholesterol, and about 2.0 to about 3.0 moles of a PEGylated lipid.

[0040] In some embodiments, the LNP comprises a molar ratio of about 55 moles of an ionizable lipid, to about 5 moles of a phospholipid, about 40 moles of a cholesterol, and about 1.5 moles of a PEGylated lipid.

[0041] In some embodiments, the LNP comprises a molar ratio of about 52.5 moles of an ionizable lipid, to about 7.5 moles of a phospholipid, about 40 moles of a cholesterol, and about 1.5 moles of a PEGylated lipid.

[0042] In some embodiments, the TERT synthetic mRNA comprises at least one modified nucleoside from the list in Table 1B. In some embodiments, the modified nucleoside is pseudouridine or a pseudouridine analog, optionally wherein the pseudouridine analog is N-1-methylpseudouridine. In some embodiments, the modified nucleoside is 5-methoxyuridine.

[0043] In some embodiments, the TERT synthetic mRNA comprises an untranslated region (UTR). In some embodiments the UTR comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to any one of SEQ ID NOS: 32-36.

[0044] In some embodiments, the TERT synthetic mRNA comprises a 5' cap structure, wherein the 5' cap structure is IRES, Cap0, Cap1, ARCA, inosine, N1-methyl-guanosine, 2'-fluoro-guanosine, 7-deaza-guanosine, CleanCap™, m7(3'OMeG)(5')ppp(5')(2'OMeA)pG, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, 2-azido-guanosine, Cap2, Cap4, CAP-003, or CAP-225.

[0045] In some embodiments, the TERT synthetic mRNA comprises a poly-adenosine (poly-A) nucleotide sequence 3' to the encoding region.

[0046] In some embodiments, the TERT synthetic mRNA comprises a chain terminating nucleotide, wherein the nucleotide is 3'-deoxyadenosine (cordycepin), 3'-deoxyuridine, 3'-deoxycytosine, 3'-deoxyguanosine, 3'-deoxythymine, 2',3'-dideoxynucleosides, 2',3'-dideoxyadenosine,

2',3'-dideoxyuridine, 2',3'-dideoxycytosine, 2',3'-dideoxyguanosine, 2',3'-dideoxythymine, a 2'-deoxynucleoside, or —O— methylnucleoside.

[0047] In some embodiments, the TERT synthetic mRNA is codon optimized. In some embodiments, the TERT synthetic mRNA comprises at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to any one of SEQ ID NOS: 1, 2, 7, 9, 30, 39, or 40.

[0048] In some embodiments, the delivery vehicle is a liposome, an ionizable lipid, an extracellular vesicle, or an exosome. In some embodiments the delivery vehicle is an extracellular vesicle of an exosome, optionally wherein the extracellular vesicle or exosome comprises a targeting moiety of one or more of a lipid, a peptide, or an antibody.

[0049] In some embodiments, the method reduces fibrosis.

[0050] In some embodiments, the subject is human.

[0051] In some embodiments, the disclosure describes a composition for use. In some embodiments the composition for use is a pharmaceutical composition comprising one or more pharmaceutically acceptable solvents or excipients.

[0052] In some embodiments, the disclosure provides a kit for treating a fibrotic disease in a subject, comprising a composition and instructions for use thereof.

[0053] In some embodiments, the disclosure provides a method of treating a liver disease in a subject in need thereof, comprising administering to the subject a composition comprising one or more synthetic messenger ribonucleic acids (mRNAs) encoding telomerase reverse transcriptase (TERT).

[0054] In some embodiments, the method reduces liver fibrosis. In some embodiments, the liver disease is non-alcoholic steatohepatitis (NASH) or non-alcoholic fatty liver disease (NAFLD).

[0055] In some embodiments the liver disease is alcoholic hepatitis, liver cirrhosis, liver fibrosis, compensated cirrhosis, decompensated cirrhosis, acute-on-chronic liver failure, fibrotic stage F4 Non-alcoholic steatohepatitis (NASH), biliary atresia, primary biliary cirrhosis, primary sclerosing cholangitis, and/or chronic liver disease, hemochromatosis, Wilson's disease, or ischemic hepatitis.

INCORPORATION BY REFERENCE

[0056] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0057] The novel features of the disclosure are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present disclosure will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the disclosure are utilized, and the accompanying drawings of which:

[0058] FIG. 1 is a schematic illustrating long-term therapeutic benefit from transient, rapid telomere extension via telomerase reverse transcriptase (TERT) mRNA. In particular, the speed of telomere extension made possible by TERT mRNA treatment enables telomere maintenance by very infrequent TERT mRNA dosing. The telomerase activity

resulting from TERT mRNA delivery rapidly extends telomeres in a brief period, before the mRNA is turned over, thus allowing the protective anti-cancer mechanism of telomere shortening to function most of the time. Between treatments, normal telomerase activity and telomere shortening is present, and therefore the anti-cancer safety mechanism of telomere shortening to prevent out-of-control proliferation remains intact, while the risk of short telomere-related disease remains low. In contrast, the best existing small molecule treatment for extending telomeres requires chronic delivery, and thus presents a chronic cancer risk, and even then has a small, inconsistent effect on telomere length, with no detectable effect on telomere length at all in about half of patients.

[0059] FIG. 2 is a series of graphs showing that TERT mRNA LNPs exhibit low toxicity by liver panel. Mice were dosed intravenously with reporter mRNA encapsulated in a lipid nanoparticle employing either LNP1 or LNP2, as shown in Table 5 (N=1-4 per condition). Mice were sacrificed and blood was collected at the time points indicated (12, 24, and 72 hours). Mice receiving saline (N=4) and carbon tetrachloride (CCl₄, N=4) served as negative and positive controls, respectively. Error bars display standard error of the mean.

[0060] FIG. 3 is a series of photographs showing that TERT mRNA LNPs cause normal histology. Cre mRNA was encapsulated into LNP 1 (Table 5) and delivered intravenously (i.v.) into tdTomato fl/fl mice. Organs were harvested 72 hours later, fixed, paraffin embedded, and sectioned. Organs from an untreated tdTomato fl/fl mouse are shown for reference.

[0061] FIG. 4 is a series of photographs showing that TERT mRNA LNPs transfect hepatocytes with high efficiency. Cre mRNA was encapsulated into LNP 1 (Table 5) and delivered intravenously (i.v.) into tdTomato fl/fl mice. Organs were harvested 72 hours later, fixed, paraffin embedded, and sectioned. Photographs depict immunohistochemistry (IHC) with anti-tdTomato. Organs from an untreated tdTomato fl/fl mouse are shown for reference.

[0062] FIG. 5 is a series of photographs showing that TERT mRNA LNPs also target some cells in spleen. Cre mRNA was encapsulated into LNP 1 (Table 5) and delivered intravenously (i.v.) into tdTomato fl/fl mice. Organs were harvested 72 hours later, fixed, paraffin embedded, and sectioned. Photographs depict immunohistochemistry (IHC) with anti-tdTomato. Organs from an untreated tdTomato fl/fl mouse are shown for reference.

[0063] FIG. 6 is a pair of graphs showing that TERT mRNA LNPs cause high telomerase activity in liver. Tert mRNA was formulated with LNP1 or LNP2 (Table 5) and delivered to i.v. into TERT KO mice. 20 hours later, the livers were harvested for telomerase repeat amplification protocol (TRAP). Wild-type C57B16/J and untreated TERT KO mouse livers were used as positive and negative controls, respectively.

[0064] FIG. 7 is a photograph depicting the results of an assay demonstrating that luciferase mRNA LNPs cause high bioluminescence signal in liver. Various LNPs designated as Lipid Nanoparticle 1 (LNP1), Lipid Nanoparticle 2 (LNP2), or Lipid Nanoparticle 3 (LNP3) (Table 5). Empty LNP shown as a negative control (ctrl). Luciferase mRNA was formulated with LNP 1, 2, and 3 and delivered via IV

injection into C57B16/J mice. 20 hours later, these mice were shaved and imaged after injection with luciferin using the IVIS BLI system.

[0065] FIG. 8 is a graph and a series of photographs of a first study demonstrating that TERT LNPs reduce fibrosis in thioacetamide (TAA) drinking water model. The addition of thioacetamide (TAA) to drinking water represents an art-recognized model for the induction of experimental liver fibrosis in rodents (Wallace M C, Hamesch K, Lunova M, et al. Standard operating procedures in experimental liver research: thioacetamide model in mice and rats. Lab Anim 2015; 49:21-9). In this experiment, TERT KO mice received 0.3 g/L TAA in their drinking water for 9.5 weeks. Mice were treated with LNPs carrying either Tert mRNA or Luciferase (LUC) mRNA once weekly. Liver sections were stained with Picosirius red (PSR), and a quantification of showed a 24% mean reduction in PSR stained tissue in mice treated with TERT LNPs compared to those treated with LUC LNPs. Scale bar on photographs equals 500 μ m.

[0066] FIGS. 9A AND 9B are graphs and photographs of a study demonstrating that TERT LNPs Reduce Fibrosis in a Thioacetamide (TAA) Drinking Water Model. TERT KO mice received 0.3 g/L TAA in their drinking water for 9.4 weeks and were treated with TERT or LUC mRNA-LNPs once weekly. By picrosirius red (PSR) staining, there was an 18% mean reduction in fibrosis in female mice and a 37% mean reduction was observed in males treated with TERT mRNA-LNPs, representing a significant ($p=0.041$) reduction in fibrosis. The scale bar on photographs is 500 μ m. Additionally, using the 0 through 4 scoring system developed by the Pathology Committee of the NASH Clinical Research Network (Kleiner et al Hepatology 2005), animals treated with TERT mRNA LNPs had a significant reduction in fibrosis compared to control animals treated with LUC (luciferase) mRNA LNPs ($p=0.032$) as seen in FIG. 9B. For all scoring, liver fibrosis was scored independently for each of 3 lobes per mouse (right, median, and left) in a blinded manner. The scores were averaged together to get a score per mouse, which were then plotted in the graphs (FIGS. 8, 9A, and 9B).

[0067] FIGS. 10A AND 10B are graphs demonstrating that TERT mRNA improves survival. Survival plotted as fraction of mice alive as a function of days post first dose of either TERT or a Luciferase (Luc) negative control. Same experimental procedure was followed as described in FIG. 9, except that the mice were 4th generation (G4) TERT KO mice aged to over 30 weeks at the start of the study. These mice were dosed once weekly with TERT or LUC, and survival was recorded after the first dose. TERT treated mice showed a 42% increase in median survival and a 58% increase in maximal survival. FIG. 10B shows the survival of a group of aged mice that did not receive thioacetamide (TAA) in drinking water (no TAA).

[0068] FIGS. 11A AND 11B show the effects of TERT mRNA on pathological liver fibrosis in mice. FIG. 11A shows the assessment of pathological fibrosis in liver sections from TERT knockout mice with TAA-induced liver fibrosis, as described in FIG. 9, by a certified pathologist based on the Non-alcoholic Fatty Liver Disease (NAFLD) Activity Score (NAS). TERT knockout mice without TAA-induced liver fibrosis (saline IV) exhibited a mild inflammation score of 1 (<2 foci per 200x field of view). Untreated control mice (C57Bl/6 strain) exhibited no detectable inflammation. FIG. 11B shows inflammation was signifi-

cantly reduced in TERT knockout mice treated with TERT mRNA compared to the control (LUC): TERT mRNA treatment resulted in a 60% reduction in the number of animals with a score of >1.

[0069] FIGS. 12A AND 12B show the transfection efficiency of mRNA in liver. FIG. 12A shows the quantification of percent (%) positive hepatocytes after different doses of Cre mRNA encapsulated in lipid nanoparticle using ionizable lipid 1 (LNP1) delivered intravenously to tdTomato fl/fl (HTT flox/flox) mice. Hepatocytes were identified in liver tissue sections using nuclear size and circularity by QuPath software. The experimental design was the same as for FIGS. 3-5. FIG. 12B is representative images of immunohistochemistry (IHC) using an anti-tdTomato antibody in liver sections.

[0070] FIG. 13 shows levels of liver damage makers following TERT mRNA delivery. TERT mRNA was formulated with LNP1 or D-Lin-MC3-DMA (MC3) and delivered intravenously into C57B16 mice at 0.6 mg/kg. 24 hours later, the plasma was taken for measurement of alanine aminotransferase (ALT) and aspartate aminotransferase (AST). TERT-LNP1 had equivalent or lower levels of ALT and AST compared to MC3. These liver enzymes serve as markers of acute toxicity.

[0071] FIGS. 14 AND 14B show the transfection efficiency in fibrotic liver. FIG. 14A shows the quantification of percent (%) positive hepatocytes following delivery of Cre mRNA encapsulated in lipid nanoparticle using LNP1 to tdTomato fl/fl mice after 16 weeks of treatment with thioacetamide (TAA) in drinking water at 0.3 g/L. Hepatocytes were identified using nuclear size and circularity by QuPath software. The experimental design was the same as for FIGS. 3-5. FIG. 12B: is representative immunohistochemical (IHC) images using an anti-tdTomato antibody in liver sections.

[0072] FIGS. 15A AND 15B show telomere extension in liver. The experimental design is as follows: 3 doses of either TERT mRNA (SEQ ID NO: 37) or Luciferase (LUC) mRNA formulated with LNP1 were delivered to TERT KO mice intravenously once weekly at 0.5 mg/kg. The mRNA-LNP dosing was preceded two days prior by a dose of thioacetamide intraperitoneally (i.p.) at 50 mg/kg. Mice were harvested 1 week after the final dose of mRNA-LNP. Telomere length was quantified in hepatocytes using Q-FISH. Liver tissue was fixed, sectioned, and stained with TelC fluorescent probe that labels the telomeres. Individual telomere fluorescence was quantified on a per cell basis and the average was taken for each mouse. The median fluorescence is shown in FIG. 15A and 10th percentile fluorescence is shown in FIG. 15B. Each point represents a single mouse. Hepatocytes in mice treated with TERT mRNA had significantly longer telomeres than LUC mRNA treated control animals. At least 300 cells were analyzed per mouse per treatment group.

[0073] FIGS. 16A AND 16B show telomerase activity in human hepatocytes. Human hepatocytes from a 51-year-old donor cultured were transfected with green fluorescent protein (GFP) mRNA or TERT mRNA using Messenger Max from Thermo Scientific at 1 μ g/ml. Cells were harvested for each time point indicated in FIG. 16A for the TRAP assay to measure telomerase activity. Telomerase activity produces a ladder pattern that was detected using an Agilent Bioanalyzer. Telomerase activity returned to baseline by day 14.

[0074] FIGS. 17A AND 17B show that telomere length was quantified from human hepatocytes from a 51-year-old donor. In the experimental design human hepatocytes were cultured on glass coverslips. Cells were transfected once with TERT mRNA at 1 μ g/ml using Messenger Max from Thermo Scientific or left untreated (UT). Cells were fixed and stained with TelC fluorescent probe using the Q-FISH protocol. Individual telomere fluorescence was quantified on a per cell level. The mean fluorescence is shown in FIG. 17A and the 10th percentile fluorescence is shown in FIG. 17B. At least 150 cells were analyzed per treatment group.

[0075] FIGS. 18A AND 18B show TERT mRNA (SEQ ID NO: 40) formulated with LNP1 and imaged at high resolution using the Thermo Scientific Talos Glacios Cryo transmission electron microscope (TEM) at 34,000 magnification and 200 kv voltage. A representative image is shown in FIG. 18A; the TEM copper grid is the dark region on the right. The particle size was characterized using dynamic light scattering (DLS) using a Brookhaven 90Plus Particle Analyzer as shown in FIG. 18B.

[0076] FIG. 19 shows results of the telomerase activity assay “telomerase repeat amplification protocol” (TRAP) in human fibroblasts treated for 24 hours with 1 μ g/ml TERT mRNAs of from left to right, untreated cells, SEQ ID NOS:39, 40, 1, 2, 31, 3, 5, and 4 respectively, and a GFP mRNA control. Telomerase activity is indicated by a characteristic ladder pattern as shown by the transfection of TERT mRNAs of SEQ ID NOS: 39, 40, 1, 2, 31, 3, 5, and 4 to varying degrees. The samples transfected with human TERT mRNA showed higher levels of telomerase activity than the samples transfected with mouse TERT mRNA. Untreated and GFP mRNA samples did exhibit telomerase activity.

[0077] FIG. 20 shows in vivo delivery of mRNA in a photograph depicting the results demonstrating that luciferase mRNA LNPs cause high bioluminescence signal in liver. Luciferase mRNA was formulated with SS-OP using the lipid ratios for LNP1, as shown in Table 5. The lipid:mRNA ratios (wt/wt) were varied. The formulated mRNA LNPs were delivered via IV injection into C57B16/J mice at 0.6 mg/kg. As a negative control, a mouse was injected with saline. 24 hours later, these mice were shaved and imaged after injection with luciferin using the Lago instrument from Spectral Instruments Imaging. Depicted is an BLI image from mice dosed with lipid:mRNA ratios of 175, 42, and 25. The signal was highest in the mice receiving LNPs with a lipid:mRNA ratio (wt/wt) of 175 and 42. The other data presented here using LNP1 uses a wt/wt ratio of 42.

[0078] FIG. 21 shows in vivo delivery of mRNA in a photograph depicting the results demonstrating that luciferase mRNA LNPs causing high bioluminescence signal in liver. LNPs designated as Lipid Nanoparticle 4 (LNP4) or Lipid Nanoparticle 5 (LNP5) were formulated using the recipe in Table 5 with luciferase mRNA. These LNPs were delivered via IV injection into C57B16/J mice at 0.6 mg/kg. As a negative control, a mouse was injected with saline. 20 hours later, these mice were shaved and imaged after injection with luciferin using the Lago instrument from Spectral Instruments Imaging. LNP4 consisted of the formula for LNP2, but with SS-OP substituted for cKK-E12. LNP5 consisted of the formula for LNP1, but with cKK-E12 substituted for SS-OP. Bioluminescent imaging indicates that both of these LNPs had successful delivery to the liver.

[0079] FIG. 22 shows in vivo delivery of mRNA in a photograph depicting the results demonstrating that luciferase mRNA LNPs causing high bioluminescence signal in liver. Luciferase mRNA was formulated with lipids per the recipe for LNP1 in Table 5. The ingredient that was varied was the molar ratio of DMG-PEG2000. As shown in FIG. 22, DMG-PEG2000 was added as either 1, 1.5, 2, or 3 parts relative to the molar sum of all lipids, while the molar ratio for the other 3 lipids is held constant. This corresponds to a molar percentage for DMG-PEG2000 of approximately 1.0%, 1.5%, 2.0%, and 2.9%. 20 hours after intravenous delivery at 0.6 mg/kg, the C57BL/6J mice were shaved and imaged following luciferin injection using the Lago instrument from Spectral Instruments Imaging. The signal was strong from all of the mice receiving active Luciferase mRNA LNPs, and the best signal was seen when DMG-PEG2000 was added in a molar ratio of 1.5:101.5 for a total of (~1.5%). The other data presented here uses LNP1 with this molar ratio of DMG-PEG2000.

[0080] FIG. 23 is a capillary electrophoresis gel image showing that TERT mRNA LNPs cause high telomerase activity in liver. Tert mRNA (mTert SEQ 37) was formulated with LNP3, a lipid nanoparticle containing DLin-MC3-DMA (Table 5) and delivered i.v into TERT KO mice at 0.6 mg/kg. 16 hours or 8 days later (as indicated in the image), the livers were harvested for telomerase repeat amplification protocol (TRAP). The negative control was a TRAP performed on a liver from a TERT KO mouse that was injected with saline. Livers from mice treated with TERT mRNA LNP3 exhibit elevated telomerase activity which returns to baseline levels, indicating the increase in telomerase activity was transient.

DETAILED DESCRIPTION

[0081] Provided herein are compositions and methods that may be used for preventing or treating fibrotic diseases and liver diseases or disorders. The present disclosure describes the surprising result that compositions comprising an mRNA encoding telomerase reverse transcriptase (TERT) reduce liver fibrosis. TERT mRNA therapies as described herein may be delivered in lipid nanoparticles (LNPs), or by other delivery vehicles. Diseases that may be treated include, without limitation, fibrotic diseases, e.g. of the liver, and other liver diseases.

[0082] Telomerase reverse transcriptase (TERT) is an enzyme known to maintain and extend chromosomal ends (telomeres). The TERT enzyme is a catalytic subunit of the ribonucleoprotein telomerase. TERT adds simple sequence repeats to telomeres by copying a template sequence 5'-GGTTAG-3' within the RNA component of telomerase. This addition of repetitive deoxyribonucleic acid (DNA) sequences helps slow telomere shortening, which occurs over time, e.g., due to incomplete DNA replication during mitosis.

[0083] TERT translocates between the nucleus and cytoplasm and has been shown to be a critical factor in a number of other biological processes, including cell proliferation and cancer metastasis. Thus, the level of TERT in the nucleus may be a critical step in regulating cell and organismal health.

[0084] Telomerase reverse transcriptase (TERT) is also known in the art as TRT, cutaneous malignant melanoma 9 (CMM9), dyskeratosis congenita autosomal dominant 2 (DKCA2), autosomal recessive dyskeratosis congenita-4

(DKCB4), human ever shorter telomeres 2 (HEST2), pulmonary fibrosis/bone marrow failure telomere related 1 (PFBMFT1), telomerase catalytic subunit (TCS1), and telomerase associated protein 2 (TP2).

[0085] In some embodiments, the treatments described herein may stop, slow, or reverse progression of a fibrotic disease, e.g., a liver disease, or other liver diseases.

[0086] TERT mRNA is transient and only requires a few hours to extend telomeres in human cells before being degraded. Therefore, TERT mRNA leaves the protective anti-cancer telomere shortening mechanism intact. The present disclosure provides compositions and methods for delivery of TERT mRNA and treatment of fibrotic diseases and liver diseases.

[0087] During normal aging, telomeres shorten by approximately 30-100 base pairs per year due to oxidation and incomplete DNA replication during S phase of the cell cycle (Kurenova E V, et al. Telomere functions. A review. Biochemistry (Mosc) 1997; 62:1242-53). Telomerase, consisting of the TERT protein and a polynucleotide template (TERC), extends telomeres, but in humans, it is inactive in most somatic cell types and is only active at low levels that are insufficient to prevent net telomere shortening in many progenitor cell types. The exception is the spermatogenic lineage, in which telomerase is active enough to maintain telomere length over the human lifespan (Takubo K, Aida J, Izumiya-Shimomura N, et al. Changes of telomere length with aging. Geriatric Gerontology Int 2010; 10 Suppl 1:S197-206). As the TERC component is present at high levels in all cell types, typically over 10,000 copies per cell, TERT is the limiting component. Because short telomeres limit the proliferative and regenerative capacities of cells, they are associated with aging, early death, and a vast number of diseases and conditions.

[0088] Telomeres comprise repetitive DNA sequences at the ends of linear chromosomes that, when sufficiently long, can allow each chromosome end to form a loop that protects the ends from acting as double-stranded or single-stranded DNA breaks. Telomeres can shorten over time, due in part to oxidative damage and incomplete DNA replication, eventually leading to critically short telomeres unable to form the protective loop, exposure of the chromosome ends, chromosome-chromosome fusions, DNA damage responses, and cellular senescence, apoptosis, or malignancy.

[0089] Telomere length maintenance can play a role in preventing cellular senescence and apoptosis and resulting cellular and organ dysfunction. In many diseases, the need for cell replication to replace cells damaged or killed by the underlying disease mechanism shortens telomeres more rapidly than normal, exhausting the replicative capacity of cells, and leading to tissue dysfunction, exacerbated or additional symptoms, disability, or death. Further, genetic mutations of telomerase enzyme (TERT) can be linked to fatal inherited diseases of inadequate telomere maintenance.

[0090] The prospect of preventing, delaying, or treating dysfunction, conditions, and diseases by telomere extension motivates a need for safe and effective treatments to extend telomeres in animal cells *in vivo* and/or *in vitro*, and safe and effective compositions and methods for delivering therapies to the animal cells to extend telomeres. Further, there is a need to safely and rapidly extend telomeres in cells for use in cell therapy, cell and tissue engineering, and regenerative medicine. At the same time, however, there can be a danger in the constitutive, as opposed to transient, activation of

telomerase activity. Indeed, for cell therapy applications, there is a need to avoid cell immortalization. To this end, transient, rather than constitutive, telomerase activity can be advantageous for safety, e.g., if the elevated telomerase activity is not only brief but extends telomeres rapidly enough that the treatment does not need to be repeated continuously.

[0091] Thus, there is need for therapies that safely extend telomeres to potentially prevent, delay, ameliorate, or treat these and other conditions and diseases, to do the same for the gradual decline in physical form and function and mental function that accompanies chronological aging, and to enable cell therapies and regenerative medicine. Furthermore, there is need for improved methods of delivering these therapies, e.g., nucleic acid molecules encoding telomerase, to cells.

[0092] Unless otherwise defined herein, scientific and technical terms used in this application shall have the meanings that are commonly understood by those of ordinary skill in the art. Generally, nomenclature used in connection with, and techniques of, chemistry, molecular biology, cell and cancer biology, immunology, microbiology, pharmacology, and protein and nucleic acid chemistry, described herein, are those well-known and commonly used in the art.

[0093] It must be noted that, as used herein and in the appended claims, the singular forms "a," "and," and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a drug candidate" refers to one or mixtures of such candidates, and reference to "the method" includes reference to equivalent steps and methods known to those skilled in the art, and so forth.

[0094] As used herein, the term "approximately" or "about," as applied to one or more values of interest, refers to a value that is similar in magnitude and/or within a similar range to a stated reference value. In certain embodiments, the term "approximately" or "about" may refer to a range of values that fall within 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, or less in either direction (greater than or less than) of the stated reference value unless otherwise stated or otherwise evident from the context (except where such number would exceed 100% of a possible value).

[0095] Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limit of that range and any other stated or intervening value in that stated range is encompassed within the disclosure. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges and also encompassed within the disclosure, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the disclosure.

[0096] "G," "C," "A," "T" and "U" generally stand for the bases, guanine, cytosine, adenine, thymidine and uracil, respectively. Nucleobases can form nucleosides by the addition of a five carbon sugar. If the sugar is ribose then the nucleoside is a ribonucleoside. Nucleosides can in turn form nucleotides by the addition of one or more linker groups such as phosphate groups. Nucleotides can in turn form polymers (polynucleotides) which include short polymers

(oligonucleotides). However, it will be understood that the terms "base", "nucleobase", "nucleoside", "ribonucleoside", "nucleotide", "ribonucleotide" can also refer to a modified base, nucleobase, nucleoside, ribonucleoside, nucleotide, or ribonucleotide, as further detailed below, or a surrogate replacement moiety (see, e.g., Table 1B and elsewhere herein). The skilled person is well aware that guanine, cytosine, adenine, thymidine, uracil can be replaced by other moieties without substantially impairing one or more of certain properties (such as base pairing properties, translatability, or protein binding properties) of an oligonucleotide or polynucleotide comprising a nucleotide bearing such replacement moiety. Sequences containing such replacement moieties are suitable for the compositions and methods featured in the disclosure. Similarly, the skilled person is well aware that ribose can be replaced with other moieties without impairing certain properties (such as base pairing properties, translatability, or protein binding properties) of an oligonucleotide or polynucleotide comprising a nucleotide bearing such replacement moiety. Sequences containing such replacement moieties are suitable for the compositions and methods featured in the disclosure. Similarly, the skilled person is well aware that phosphate can be replaced with other moieties without impairing certain properties (such as base pairing properties, translatability, or protein binding properties) of an oligonucleotide or polynucleotide comprising a nucleotide bearing such replacement moiety. Sequences containing such replacement moieties are suitable for the compositions and methods featured in the disclosure.

[0097] As used herein, the terms "polypeptide," "peptide," and "protein" refer to polymers of amino acids of any length. The terms also encompass an amino acid polymer that has been modified; for example, to include disulfide bond formation, glycosylation, lipidation, phosphorylation, or conjugation with a labeling component.

[0098] As used herein, the terms "identity" and "identical" refer, with respect to a polypeptide or polynucleotide sequence-of-interest, to the percentage of exact matching residues in an alignment of that the sequence-of-interest to a reference sequence, such as an alignment generated by the BLAST algorithm. Identity is calculated, unless specified otherwise, across the full length of the reference sequence. Thus a sequence-of-interest "shares at least x % identity to" a reference sequence if, when the reference sequence is aligned (as a query sequence) is aligned to the sequence-of-interest (as subject sequence), at least x % (rounded down) of the residues in the subject sequence are aligned as an exact match to a corresponding residue in the query sequence, the denominator being the full length of the reference sequence plus the lengths of any gaps inserted into the reference sequence by alignment of the reference sequence to the sequence-of-interest. Where the subject sequence has variable positions (e.g., residues denoted X), an alignment to any residue in the query sequence is counted as a match. Sequence alignments may be performed using the NCBI Blast service (BLAST+ version 2.12.0) or another program giving the same results.

[0099] The term "native" or "wild-type" as used herein refers to a nucleotide sequence, e.g. gene, or gene product, e.g. RNA or polypeptide, that is present in a wild-type cell, tissue, organ or organism. The term "variant" as used herein refers to a mutant of a reference polynucleotide or polypeptide sequence, for example a native polynucleotide or poly-

peptide sequence, i.e., having less than 100% sequence identity with the reference polynucleotide or polypeptide sequence. Put another way, a variant comprises at least one nucleotide difference (e.g., nucleotide substitution, nucleotide insertion, nucleotide deletion) or one amino acid difference (e.g., amino acid substitution, amino acid insertion, amino acid deletion) relative to a reference polynucleotide sequence, e.g. a native polynucleotide or polypeptide sequence. For example, a variant may be a polynucleotide having a sequence identity of 50% or more, 60% or more, or 70% or more with a full length native polynucleotide sequence, e.g. an identity of 75% or 80% or more, such as 85%, 90%, or 95% or more, for example, 98% or 99% identity with the full length native polynucleotide sequence. As another example, a variant may be a polypeptide having a sequence identity of 70% or more with a full length native polypeptide sequence, e.g. an identity of 75% or 80% or more, such as 85%, 90%, or 95% or more, for example, 98% or 99% identity with the full length native polypeptide sequence. Variants may also include variant fragments of a reference, e.g. native, sequence sharing a sequence identity of 70% or more with a fragment of the reference, e.g. native, sequence, e.g. an identity of 75% or 80% or more, such as 85%, 90%, or 95% or more, for example, 98% or 99% identity with the native sequence.

[0100] As used herein, the term "codon optimized" refers to any process used to improve gene expression and increase the translational efficiency of a gene of interest by accommodating the codon bias of the host organism, and/or to reduce the immunogenicity of the polynucleotide.

[0101] The terms "treating" or "treatment" are used herein to generally mean obtaining a desired pharmacologic and/or physiologic effect with a therapeutic agent. The effect may be prophylactic in terms of completely or partially preventing a disease or symptom thereof, e.g. reducing the likelihood that the disease or symptom thereof occurs in the subject, and/or may be therapeutic in terms of completely or partially reducing a symptom, or a partial or complete cure for a disease and/or adverse effect attributable to the disease. "Treatment" as used herein covers any treatment of a disease in a mammal, and includes: (a) preventing the disease from occurring in a subject which may be predisposed to the disease but has not yet been diagnosed as having it; (b) inhibiting or slowing the onset or development of the disease; or (c) relieving the disease, e.g., causing regression of the disease or symptoms associated with the disease. The therapeutic agent may be administered before, during or after the onset of disease. The treatment of ongoing disease, where the treatment stabilizes or reduces the undesirable clinical symptoms of the patient, may be of particular interest. In some embodiments, treatment is performed prior to complete loss of function in the affected tissues. In some embodiments, the subject therapy will be administered before the symptomatic stage of the disease; and, in some embodiments, during the symptomatic stage of the disease; and, in some embodiments, after the symptomatic stage of the disease.

[0102] In some embodiments, therapies as described herein treat fibrotic diseases or liver diseases, including but not limited to fibrotic liver diseases. The fibrotic or liver diseases may be associated with a TERT mutation, mutation in other genes, or non-genetic causes. Diseases that may be treated using the composition and methods of the present disclosure include non-alcoholic fatty liver disease

(NAFLD), non-alcoholic steatohepatitis (NASH), e.g., stage F4 NASH, alcoholic hepatitis, alcoholic liver disease, liver cirrhosis, e.g. compensated and non-compensated cirrhosis, liver fibrosis, hemochromatosis, biliary atresia, primary biliary cirrhosis, primary sclerosing cholangitis, chronic liver disease, acute-on-chronic liver failure (ACLF), Wilson's disease, or ischemic hepatitis.

[0103] The terms "individual," "subject," and "patient" are used interchangeably herein and refer to any subject for whom treatment or therapy is desired. The subject may be a mammalian subject. Mammalian subjects include, e.g., humans, non-human primates, rodents, (e.g., rats, mice), lagomorphs (e.g., rabbits), ungulates (e.g., cows, sheep, pigs, horses, goats, and the like), etc. In some embodiments, the subject is a human. In some embodiments, the subject is a non-human primate, for example a cynomolgus monkey. In some embodiments, the subject is a companion or service animal (e.g. cats or dogs).

[0104] A subject "in need thereof," as used herein, refers to any subject suffering from or identified to be at risk of developing a fibrotic disease or liver disease.

[0105] It is to be understood that this disclosure is not limited to the particular methodology, products, apparatus and factors described, as such methods, apparatus and formulations may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and it is not intended to limit the scope of the present disclosure which will be limited only by appended claims.

I. Synthetic mRNAs

[0106] A synthetic mRNA as used herein may refer to any sequence comprising a mutation (point or deletion) or additional nucleotides not found in the wild type sequence. For example, a synthetic TERT mRNA may refer to a wild type sequence encoding a human TERT sequence, flanked by the addition of 1, 2, 3, 10, 100 or more nucleotides. Similarly, the nucleotides themselves may encode amino acids distinct from the wild type, or be modified to reduce immunogenicity in the cell or tissue. An mRNA sequence in some embodiments may comprise any of the following modifications, including but not limited to an untranslated region (UTR), a 5' cap, and a poly-adenosine tail. In some embodiments, the RNA may be circular and/or self-replicating. Illustrative methods of making circular mRNAs are provided in Chen et al. *Science*. 1995 Apr. 21; 268(5209):415-7; Perriman R. (2002) Circular mRNA Encoding for Monomeric and Polymeric Green Fluorescent Protein. In: Hicks B. W. (eds) Green Fluorescent Protein. Methods in Molecular Biology, vol 183. Humana Press; Wang et al. *RNA*. 2015 Feb; 21(2):172-9. doi: 10.1261/ma.048272.114. Epub 2014 Dec 1; Wesselhoeft et al. *Nat Commun.* 2018 Jul. 6; 9(1): 2629; and Wesselhoeft et al. *Mol Cell*. 2019 May 2; 74(3): 508-520.e4. Illustrative methods of making self-replicating mRNAs are provided in Tews B. A., Meyers G. (2017) Self-Replicating RNA. In: Kramps T., Elbers K. (eds) RNA Vaccines. *Methods in Molecular Biology*, vol 1499. Humana Press; Leyman et al. *Mol Pharm.* 2018 Feb. 5; 15(2):377-384; and Huysmans et al. *Mol Ther Nucleic Acids*. 2019 Sep. 6; 17:388-395.

TERT mRNAs

[0107] In some embodiments, a composition may comprise a reverse transcriptase telomerase (TERT) mRNA sequence to treat one or more phenotypes or symptoms associated with a fibrotic disease or liver disease. In some

embodiments, a TERT mRNA refers to an mRNA encoding any full length, functional fragment or portion of a TERT protein, including wild type sequences or variants thereof.

[0108] In some embodiments, a TERT mRNA may comprise a codon-optimized sequence. In some embodiments, a TERT mRNA may comprise a uridine depleted human TERT sequence. In some embodiments, the codon-optimized sequence may comprise a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO 1:

```

ATGCCAGAGCCCCAGATGCAGAGCCGTGAGAAGCCTGCTGAGAACCCA
CTACAGAGGGTCTGCCCTGGCACCTTCGTGAGAACACTGGGCCCC
AGGGCTGGAGACTGGTGCAGAGAGGCACCCGCCCTCAGAGCCCTG
GTGGCCCAGTGCCTGGTGTGCGTGCCCTGGACGCCAGACCCCTCCGC
CGCCCCCAGCTTCAGACAGGTGAGCTGCCCTGAAGGAGCTGGTGGCCAGAG
TGCTGCAGAGACTGTGCAGAGAGGCCAAGAACGTGCTGGCCTTCGGC
TTGCCCTGCTGGACGCCCTGGGGCCAGAGGGGGCCCTCCAGGCCCTCACAC
CAGCGTGAGAACACTGCCAACACCGTGACCGACGCCCTGAGAGGCA
GCCGCCCTGGGGCTGCTGCTGAGAAGAGTGGCGACGACGTGCTGGTG
CACCTGCTGGCCAGATGCCCTGTTCTGCTGGTGGCCCCAGCTGCC
CTACCAGGTGTGCCCTCCCTGTACAGCTGGGCCGCCACCCAGG
CCAGACCCCCCTCCACGCCAGCGGCCAGAGAACACTGGCCTGCCAG
AGAGCCTGGAACCACAGCGTGAGAGAGGCCGGTGGCCCTGGCCTGCC
CGCCCCCGGCCAGAAGAACAGGCCAGCGCAGCAGAACGCTGCC
TGCCCAAGAGACCCAGAAGAGGCCGCCGCCAGGCCAGAGAACCCCC
GTGGGCCAGGGCAGCTGGGCCACCCCGCAGAACCCAGAGGCCAGCGA
CAGAGGCTTCTGCGTGGTGAGCCCCGCCAGACCCGCCAGGAGGCCACCA
GCCCTGGAGGGGCCCTGAGCGGCCACAGACAGCCACCCAGCGTGGC
AGACAGCACCACGCCGCCCTCCAGCACCAGCAGACCTCCAGACCTG
GGACACCCCCCTGCCCTCCCTGTACGCCAGGCCACAGCACTTCTGTACA
GCAGCGCGACAAGGAGCAGCTGAGACCCAGCTTCTGCTGAGCAGCTG
AGACCCAGCTGACCGCGCCAGAACAGACTGGTGGAGACCATCTTCTGG
CAGCAGACCCCTGGATGCCGGCACCCCAAGAACAGACTGCCAGACTGCC
AGAGATACTGGCAGATGAGACCCCTGTTCTGGAGCAGCTGGCAACAC
GCCCAGTCCCCCTACGCCGTGCTGAGAACCCACTGCCCTGAGAGC
CGCCGTACCCCCGCCGGCGTGTGCGCCAGAGAGAACAGGCCAGGGCA
GCGTGGCGCCCCCGAGGAGGAGCACCGACCCAGAACAGACTGGTGCAG
CTGCTGAGACAGCACAGCAGCCCTGGCAGGTGTACGGCTTCGTGAGAGC
CTGCTGAGAACACTGGTGCCTCCGGCTGTGGGGCAGCAGACACAACG
AGAGAAAGATTCTGAGAAAACCCAAGAAGTTCATCAGCCTGGCAAGCAC
GCCAAGCTGAGCTGCAAGGAGCTGACCTGAAAGATGAGCGTGAGAGACTG
CGCTGGCTGAGAAGAACAGGCCCGCGTGGCTGCGTGCCCGCCGAGC

```

- continued

```

ACAGACTGAGAGAGGAGATCTGGCCAAGTCTCTGCAC TGCGTGTGAGC
GTGTACGTGGTGGAGCTGCTGAGAAGCTTCTTCTACGTGACCGAGACCAC
CTTCCAGAAGAACAGACTGTTCTTCTACAGAAAGAGCGTGTGGAGCAAGC
TGCAGAGCATCGGCATCAGACAGCACCTGAAGAGAGTGAGCTGAGAGAG
CTGAGCGAGGCCAGGTGAGACAGCACAGAGAGGCCAGACCCGCCCTGCT
GACCAGCAGACTGAGATTCATCCCCAAGCCCGACGCCCTGAGACCCATCG
TGAACATGGACTACGTGGTGGCGCCAGAACCTTCAGAAGAGAGAAGAGA
GCCGAGAGACTGACCAGCAGAGTGAAAGCCCTGTTCA CGCGTGTGAACTA
CGAGAGGCCAGAACAGCCGCCCTGCTGGGCCAGCGTGTGGGCCCTGG
ACGACATCCACAGAGCCTGGAGAACCTTCGTGCTGAGAGTGAGAGCCAG
GATCCCCCTCCCGAGCTGTACTCGTAAGGTGGACGTGACCGGCCCTA
CGACACCATCCCCCAGGACAGACTGACCGAGGTGATGCCAGCATCATCA
AGCCCCAGAACACCTACTGCGTGAGAAGATA CGCCGTGGTGAGAACGCC
GCCCACGCCACGTGAGAACGCCCTTAAGAGCCACGTGAGCACCCCTGAC
CGACCTGCAGCCCTACATGAGACAGTTCGTGGCCACCTGCGAGGAGACCA
GCCCCCTGAGAGACGCCGTGGTGTACGAGCAGAGCAGCAGCCTGAACGAG
GCCAGCAGCGGCCCTGTCAGCTGAGATTCATGTGCCACCCACGC
CGTGAGAATCAGAGGCAAGAGCTACGTGCA GTGCCAGGGCATCCCCCAGG
GCAGCATCCTGAGCACCCCTGCTGTGACGCCCTGTGACCGACATGGAG
AACAGCTGTTGCCCGCATCAGAAGAGACGCCCTGCTGAGACTGGT
GGACGACTTCCTGCTGGTGACACCCACCTGACCCACGCCAACCTCC
TGAGAACCCCTGGTGAGAGCCGTGCCCGAGTACGGCTGCGTGGTGAACCTG
AGAAAGACCGTGGTGAACCTCCCGTGGAGGACGAGGCCCTGGCGGCAC
CGCCTTCGTGAGATGCCGCCAACGCCCTGTTCCCTGGTGGCGGCCCTGC
TGCTGGACACCGAACCTGGAGGTGCA GAGCGACTACAGCAGCCTACGCC
AGAACCCAGCATCAGAGCCAGCCTGACCTAACAGAGGCTTCAAGGCCGG
CAGAAACATGAGAACAGAACGCTGTCGGCGTGTGAGACTGAAGTGCACA
GCCTGTTCTGGACCTGCAGGTGAACGCCCTGCA GAGCCGTGTGACCAAC
ATCTACAAGATCCTGCTGCTGCA GGGCTACAGATTCCACGCCCTGCGTGT
GCAGCTGCCCTTCCACCAAGCAGGTGTGGAGAACCCCCACCTTCTTCTG
GAGTGATCAGCGACACGCCAGCCTGTGCTACAGCATCCTGAAGGCCAG
AACGCCGGCATGAGCCTGGCGCCAAGGGGCCGCCGGCCCCCTGCCAG
CGAGGCCGTGCA GTGGCTGTGCCACCCAGGCCCTGCTGAAGCTGACCA
GACACAGAGTGACCTACGTGCCCTGCTGGCAGCCTGAGAACGCCAG
ACCCAGCTGAGCAGAACAGCTGCCCGCACCCCTGACGCCCTGGAGGC
CGCCGCCAACCCGCCCTGCCAGCGACTGCTGAGCTCCCTTATGTCACGGAGAC

```

[1019] In some embodiments, a TERT mRNA may comprise a mutant human TERT sequence. In some embodiments, the mutant human TERT mRNA may encode a Y707F mutation in the resulting peptide sequence. In some embodiments a mutation in the TERT mRNA sequence

encodes a mutation in the nuclear export signal which may result in nuclear retention of the TERT peptide. In some embodiments, the mutant TERT mRNA sequence may comprise a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO 2:

```

ATGCCGCGCGCTCCCCCTGCCAGCCGTGCGCTCCCTGCTGCGCAGCCA
CTACCGCGAGGTGCTGCCGCTGGCACGTTCTGCGGCCCTGGGGCC
AGGGCTGGCGCTGGTGCA CGCGCGGGACCCGGCGCTTCCGCGCCTG
GTGGCCCAGTGCCCTGGTGTGCGTGCCCTGGGACGCACGCCGCCCGC
CGCCCCCTCCCTCCGCCAGGTGCTCCCTGAAGGAGCTGGTGGCCCGAG
TGCTGCAAGGGCTGTGCGAGCGCGGCCGAAGAACGTCGCTGGCCTTCGGC
TTCGCGCTGCTGGACGGGCCCGCGGGGCCCGAGGCGCTTACAC
CAGCGTGCAGCTACCTGCCAACACGGTGACCGACGC ACTGCGGGGGA
GCCGGCGTGGGGCTGCTGCGGCCGTGGCGACGACGTGCTGGTT
CACCTGCTGGCACGCTGCGCGCTTGTGCTGGTGTGCTCCAGCTGCGC
CTACCAAGGTGTGCGGGCGCCGCTGTACCGCTCGCGCTGCCACTCAGG
CCCGGCCCCGCCACACGCTAGTGACCCGAAGGCGCTGGGATGCGAA
CGGGCCTGGAACCATAGCGTCAGGGAGGCCGGGTCCCCCTGGGCTGCC
AGCCCCGGGTGCGAGGAGGCCGGCGAGTGCCAGCCGAAGTCTGCCGT
TGCCCAAGAGGCCAGGCCGTGGCGCTGCCCTGAGCCGGAGCGGCC
GTTGGCAGGGGTCTGGGCCACCCGGCAGGACGCGTGGACCGAGTGA
CCGTGGTTCTGTTGCTGGTGACCTGCCAGACCCGCCGAAGAACCCACT
CTTGGAGGGTGCCTCTGGCACGCCACTCCACCCATCCGTGGC
CGCCAGCACCACGCCGCCCTACACATGCCGCCACACGTCCCTG
GGACACGCCCTGTCCCCGGTGTACGCCAGACCAAGCACTTCCCTACT
CCTCAGCGACAAGGAGCAGCTGCCCTCCCTACTCAGCTCTCTG
AGGCCAGCCTGACTGGCGCTCGGAGGCTGTGGAGACCATCTTCTGGG
TTCCAGGCCCTGGATGCCAGGGACTCCCGCAGGTGCCCCGCTGCC
AGCGCTACTGGCAAATGCCGCCCTGTTCTGGAGCTGCTGGGAACAC
GCCAGTGCCCTACGGGTGCTCTCAAGACGCACTGCCGCTGCGAGC
TGGCGTACCCCAAGCAGCCGGTGTCTGCCCCGGAGAACCCCCAGGGCT
CTGTGGCGGCCAGGACAGAGGCCCTGCGCTGGAGGACACAGAC
CTGCTCCGCCAGCACAGCAGCCCTGGCAGGTGTACGGCTTGTGCGGGC
CTGCCCTGCCGCCGTGGTGGCCCTGAGGCCCTCTGGGCTCCAGGCACAACG
AACGCCGCTTCTCAGGAACACCAAGAAGTTCATCTCCCTGGGAAGCAT
GCCAAGCTCTCGCTGCA GGAGAGCTGACGTGGAGATGAGCGTGCAGGACTG
CGCTTGCTGCGCAGGAGCCAGGGTTGGCTGTGTTCCGGCCGAGAGC
ACCGTCTCGCTGAGGAGATCCTGGCCAAGTCTCTGCAC TGCGTGTAGT
GTGTACGTGCTGAGGCTGCTCAGGTCTTCTTATGTCACGGAGACAC
GTTTCAAAAAGAACAGGCTCTTTCTACCGGAAGAGTGTCTGGAGCAAGT

```

- continued

```

TGCCTAAAGCATTGGAATCAGACAGCAGCTTGAGAGGGTGCAGCTGCGGGAG
CTGTCTGGAAAGCAGAGGTCAAGGAGCATCGGAAAGGCCAGGCCCTGCT
GACGTCCAGACTCCGTTCATCCCCAAGCCTGACGGCTGCGGCCGATTG
TGAAACATGGACTACGTGTTGGAGGCCAGAACGTTCCGAGAGAAAAGAGG
GCCGAGCGTCTCACCTCGAGGGTGAAGGCAGTGTTCAGCGTGCCTAAGT
CGAGCGGCCGCGCCGCCGCGCCCTCTGGGCCCTCTGTGCTGGGCCAG
ACGATATCCACAGGGCTGGCCACCTTCGTGCTGCGTGTGCGGCCAG
GACCCGCCGCTGAGCTGTTGTCAAGGTGGATGTGACGGCGCGTA
CGACACCATCCCCAAGGACAGGCTACGGAGGTACGCCAGCATCATCA
AACCCCAGAACACGTAATGCGTGCCTGGTATGCCGTGGTCCAGAACGCC
GCCCATGGCACGTCCGAAGGCCCTCAAGAGCCACGTCCTACCTTGAC
AGACCTCCAGCCGTACATGCGACAGTTGCTGGCTCACCTGCAGGAGACCA
GCCCGCTGAGGGATGCCGTCGTATCGAGCAGAGCTCCCTGAATGAG
GCCAGCAGTGGCTCTCGACGTCTTCTACGCTTACGCTGCGACACACG
CGTGCATCAGGGCAAGTCTACGTCAGTCCAGGGATCCCGCAGG
GCTCCATCCTCTCACGCTGCTCTGCAGCCTGTGCTACGGCGACATGGAG
AACAGCTGTTGGGGATTCCGGGGGACGGGCTGCTCTGGTTGGT
GGATGATTCTTGTGGTACACCTCACCTCACCCACCGGAAACACCTCC
TCAGGACCTGGTCCAGGTGTCCTGAGTATGGCTGCGTGGTGAACCTG
CGGAAGACAGTGGTGAACCTCCCTGTTAGAAGACGAGGCCCTGGTGGCAC
GGCTTTGTCAGATGCCGCCACGGCTATTCCCTGGTGGCCCTG
TGCTGGATACCCGACCTGGAGGTGCAAGAGGACTACTCCAGCTATGCC
CGGACCTCCATCAGAGCAGTCACCTTCAACCGCGGCTCAAGGCTGG
GAGGAACATGCGTCGCAAACCTTTGGGTCTTGCCTGCTGAAGTGTACA
GCCCTGTTCTGGATTTGCAGGTGAACAGCCTCAGACGGTGTGACCAAC
ATCTACAAGATCCTCTGCTGCAAGGGTACAGGTTTACGCTGTGCT
GCAGCTCCATTTCATCAGCAAGTTGGAAAGAACCCACATTTCCTG
GCGTCATCTGACACGCCCTCCCTGCTACTCCATCCTGAAAGCCAAG
AACGCAGGGATGTCGCTGGGGCAAGGGGCCGCGCCGCCCTGCCCCTC
CGAGGCCGTGCAAGTGGCTGTGCCACCAAGCATTCTGCTCAAGCTGACTC
GACACCGTGTCACTACGTCACCTGGGTCACTCAGGACAGGCCAG
ACGCAGCTGAGTCGGAAGCTCCGGGACGACGCTGACTGCCCTGGAGGC
CGCAGCCAACCCGGACTGCCCTCAGACTCAAGACCATCCTGGACTGA

```

[0110] In some embodiments, a mouse TERT mRNA may comprise a codon-optimized sequence. In some embodiments, a TERT mRNA may comprise a uridine depleted mouse TERT sequence. In some embodiments, the codon-optimized sequence may comprise a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO 3:

```

ATGACCAGGGCCCTAGGTGCCCTGCCGTGAGGAGCTGCTGAGGAGCAG
GTACAGGGAGGTGTGCCCTGCCACCTCGTGAAGGAGCTGGCCCTG
AGGGCAGGAGGTGGTGCAGCCTGGCACCCTAACATCTACAGGACCTG
GTGGCCAGTGCCTGGTGTGCATGCAGTGGGCCAGCCAGCCTCCCTG
CGACCTGAGCTTCCACCCAGGTGAGCAGCCTGAAGGAGCTGGGCCAGG
TGGTGCAGAGGCTGTGCGAGAGGAACGAGAGGAACGTGCTGGCCCTGGC
TTCGAGCTGCTGAACGAGGCCAGGGGCCCTCCATGGCCTCACAG
CAGCGTGAAGGAGCTACCTGCCTAACACCGTATCGAGACCTGAGGGTGA
GCCGCCGCTGGATGCTGCTGAGCAGGGTGGGCCAGCAGCCTGCTGGT
TACCTGCTGCCACTGCCCTGTACCTGCTGGCCTCTAGCTGCGC
CTACCAAGGTGTGCCAGCCCTGCTGACAGATCTGCCACCCACCGACA
TCTGGCTAGCGTGAAGGCCAGCTACAGGCCCTACAGGCCCTGTGGCAGG
AACTTCACCAACCTGAGGTTCTGCAAGCAGATCAAGAGCAGCAGCAGGA
GGAGGCCCTAACCTCTGGCCCTGCCTAGCAGGGCACCAGAGGCCAC
TGAGCCTGACCAGCACAGCGTGCCTAGCGCCAAGAAGGCCAGGTGCTAC
CCTGTGCTAGGGTGGAGGAGGCCCTCACAGGCAGGTGCTGCTAACCCC
TAGCGGCAAGAGCTGGTGCCTAGCCCTGCCAGGAGGCCCTGAGGTGCTA
CCGCCGAGAAGGACCTGAGCAGCAAGGCAAGGTGAGCAGCTGAGCTG
AGCGGCAGCGTGTGCTCAAGCACAAGCCTAGCAGCAGCACCAGCTGCTGAG
CCCTCCTAGGCAGAACCCCTTCCAGCTGAGGCCCTTCATCGAGACCAAGGC
ACTTCCTGTACAGCAGGGCGACGCCAGGAGAGGCTAACCTAGCTTC
CTGCTGAGCAACCTGCAGCTAACCTGACCGCGCCAGGAGGCTGGTGG
GATCATCTCCTGGCAGCAGGCCCTAGGACAGCGGCCCTCTGTGCGAGGA
CCACAGGCTGAGCAGGAGGTACTGGCAGATGAGGCCCTGTTCCAGCAG
CTGCTGGTGAACACGCCAGGTGCAAGTACGTCAGGCTGCTGAGGAGCCA
CTGCAGGTTCAAGGCCAACAGCAGGTGACCGAGGCCCTGAACACCA
GCCCTCCTCACCTGATGGACCTGCTGAGGCTGCACAGCAGCCCTGGCAG
GTGTACGGCTTCTGAGGGCCTGCCCTGCAAGGTGGTGAAGGCCAGCCT
GTGGGGCACCAGGCACACAGAGAGGAGGTTCTCAAGAACCTGAAGAAGT
TCATCAGCCTGGCAAGTACGCAAGCTGAGCCTGAGGAGCTGATGTGG
AAGATGAAGGTGGAGGACTGCCACTGGCTGAGGAGCAGCCCTGGCAAGGA
CAGGGTGCCTGCCGCCAGCACAGGCTGAGGGAGAGGATCTGGCCACCT
TCCCTGTTCTGGCTGATGGACACCTACGTCAGGAGCTGCTGAGGAGCTTC
TTCTACATCACCGAGAGCACCTCCAGAAGAACAGGCTGTTCTACAG
GAAGAGCGTGTGGAGCAAGCTGCAGAGCATCGCGTGAAGGCAGCACCTGG
AGAGGGTGGAGGCTGAGGGAGGCTGAGGCCAGGAGGAGGACCCAG
GACACCTGGCTGCCATGCCCTATCTGCAAGGCTGAGGTCATCCCTAACGCC
TAACGGCCTGAGGCCTACGTCAGACATGAGCTACAGCATGGGCCAGAGGG
CCCTGGCAGGAGGAAGCAGGCCAGCAGCAACCCAGAGGCTGAAGACC

```

- continued

CTGTTCAGCATGCTGAAC TACGAGAGGACCAAGCACCCTCACCTGATGGG
 CAGCAGCGTGTGGCATGAACGACATCTACAGGACCTGGAGGGCCTCG
 TGCTGAGGGTGAGGGCCCTGGACCAGACCCCTAGGATGTACTTCGTAAG
 GCGCACGTGACCGCCCTACGACGCCATCCCTCAGGGCAAGCTGGTGA
 GGTGGTGCCAAACATGATCAGGCACAGCGAGAGCACCCTACTGCATCAGGC
 AGTACGCCGTGGTGAGGAGGGACAGCCAGGCCAGGTGCACAAGAGCTTC
 AGGGCAGGTGACCACCCCTGAGCAGCCTGAGCCTTACATGGGAGCTT
 CCTGAAGCACCTGCAGGACAGCGACGCCAGCGCCCTGAGGAACAGCGTGG
 TGATCGAGCAGAGCATCAGCATGAACGAGAGCAGCAGCAGCCGTGAC
 TTCTCCTGCACTCCTGAGGCACAGCGTGGTGAAGATCGGCCAGGGT
 CTACACCCAGTGCAGGGCATTCCCTCAGGGCAGCAGCCTGAGCACCCTGC
 TGTCAGCAGCTGTGCTTCGGCGACATGGAGAACAGCTGTTCCCGAGGTG
 CAGAGGGACGGCTGCTGCTGAGGTTCTGGACGACTTCCTGCTGGTGCAC
 CCCTCACCTGGACCAGGCCAAGACCTCCTGAGCACCCCTGGTGCACGGCG
 TGCTGAGTACGGCTGCATGATCAACCTGCAGAACAGCCGTGGTGAACCTC
 CCTGTGGAGCCTGGCACCCCTGGCGGCCGCCCCCTTACAGCTGCGCTGC
 CCACTGCCTGTCCTGGTGCCTGCTGCTGGACACCCAGACCCCTGG
 AGGTGTTCTGCGACTACAGCGCTACGCCAGACAGCATCAAGACCAGC
 CTGACCTCCAGAGCGTGTCAAGGCCAGAACCATGAGGAACAAAGCT
 GCTGAGCGTGTGAGGCTGAAGTGCCACGGCTGTTCTGGACCTGCGAG
 TGAACAGCCTGCAGACCGTGTGCATCAACATCACAAGATCTCCGCTG
 CAGGCCTACAGGTTCCACGCCTGCGTGTCCAGCTGCCTTGCACAGAG
 GGTGAGGAGAACCTGACCTTCTCCTGGCATCATCAGCAGCCAGGCCA
 GCTGCTGCTACGCCATCCTGAAGGTGAAGAACCCCTGGCATGACCCCTGA
 GCCAGCGCAGCTCCCTCCCTGAGGCCAACAGCGTGTGCTTACAGG
 CTTCTGCTGAAGCTGCCGCCAACAGCGTGTACAAAGTGCCTGCTGG
 GCCCTCTGAGGACGCCAACAGCTGCTGTGCAGGAAGCTGCCCTGAGGCC
 ACCATGACCATCTGAAGGCCGCCAACAGCTGCCGACCCCTGCCCTGAGGCC
 CCAGACCATCCTGGACTGA

[0111] In some embodiments, a mouse TERT mRNA may comprise a mutant mouse TERT sequence. In some embodiments, the mutant mouse TERT mRNA may encode a Y707F mutation in the resulting peptide sequence. In some embodiments a mutation in the TERT mRNA sequence encodes a mutation in the nuclear export signal which may result in nuclear retention of the TERT peptide. In some embodiments, the mutant mouse TERT mRNA sequence may comprise a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO 4:

ATGACCCGCGCTCTCGTTGCCCGGGTGCCTCTGCTGCCAGCCG
 ATACCGGGAGGTGTGGCCGCTGGCACCTTGTGCGGCCCTGGGGCCG

- continued

AGGGCAGGCCCTGTGCAACCCGGGACCGAAGATCTACCGCACTTTG
 GTTGCCCAATGCCTAGTGTGCATGCAGCCTGCACAGCCTCCACCTGC
 CGACCTTCCCTCCACCAAGGTGTCATCCCTGAAAGAGCTGGTGGCCAGGG
 TTGTGCAAGAGACTCTGGAGCGCAACAGAGAGAACAGTGTGGCTTTGGC
 TTTGAGCTGCTAACAGGGCCAGAGCGGGCCCTCCATGGCCTTCACTAG
 TAGCGTGCCTAGCTACTTGCCAACACTGTTATTGAGACCCCTGCGTGTCA
 GTGGTGCATGGATGCTACTGTTGAGGCCAGTGGGCGACGACCTGCGTGTCA
 TACCTGCTGGCACACTGTGCTTTATCTCTGGTGGCCCCAGCTGTGC
 CTACCAAGGTGTGGCTCCCTGTCACCAAATTGTGCCACACAGGATA
 TCTGGCCCTCTGTGTCGCTAGTTACAGGCCACCCGACCCGTGGCAGG
 AATTCTACTAACCTTAGGTTCTTACAACAGATCAAGAGCAGTAGTCGCCA
 GGAAGCAGGAAACCCCTGGCTTGCATCTCGAGGTACAAAGAGGCATC
 TGAGTCTCACCAGTACAAGTGTGCCTCAGCTAAGAAGGGCAGATGCTAT
 CCTGTCCCAGAGTGGAGGAGGGACCCACAGGCAGGTGCTACCAACCCC
 ATCAGGCAAATCATGGGTGCCAAGTCCTGCTCGTCCCCGAGGTGCTTA
 CTGCAGAGAAAGATTGTCTTCTAAAGGAAAGGTGTGACCTGAGTCTC
 TCTGGGTCGGTGTGCTGTAAACACAAGGCCAGCTCCACATCTGCTGTG
 ACCACCCGCCAAATGCCTTCAGCTCAGGCCATTATTGAGACCAAGAC
 ATTTCCTTACTCCAGGGAGATGGCCAAGAGCGTCTAAACCCCTCATTC
 CTACTCAGCAACCTCCAGCCTAACCTGACTGGGCCAGGAGACTGGTGG
 GATCATTTCTGGGCTCAAGGCCCTAGGACATCAGGACCACTCTGCGAGA
 CACACCGTCTATCGCGTGCATACTGGAGATGCCAGTGGCCCCCTGTTCCAACAG
 CTGCTGGTGAACCATGCAGAGTGCCTAGTCAGACTCCTCAGGTGACA
 TTGCAGGTTTCGAACAGCAAACCAACAGGTGACAGATGCCTGAACACCA
 GCCCACCGCACCTCATGGATTGCTCCGCCTGCACAGCAGTCCCTGGCAG
 GTATATGGTTCTCGGGCTGTCTGCAAGGTGGTGTGCTAGTCT
 CTGGGTACCGGCACAATGAGCGCCCTCTTAAGAAACTTAAAGAAGT
 TCATCTCGTGGGAAATACGGCAAGCTATCACTGCGAGACTGATGTGG
 AAGATGAAAAGTAGAGGATTGCCACTGGCTCCGCAGCAGCCGGGAGAGA
 CCGTGTCCCCGCTGCAGAGCAGCGTCTGAGGGAGGGATCTGGCTACGT
 TCCCTGTTCTGGCTGATGGACACATACGTGGTACAGCTGCTTAGGTCTTC
 TTTTACATCACAGAGAGCACATTCCAGAAGAACAGGCTCTCTTACCCG
 TAAGAGTGTGTGGAGCAAGCTGCAGAGCATTGGAGTCAGGCAACACCTG
 AGAGAGTGCAGCTACGGAGCTGTCAAGAGAGGAGGTAGGCATCACCAG
 GACACCTGGCTAGCCATGCCCATCTGCAGACTGCGCTCATCCCCAAGCC
 CAACGGCCTGCCGCCATTGTGAACATGAGTTATAGCATGGGTACCGAG
 CTTGGGAGAAGGAAGCAGGCCAGCATTTCACCCAGCGTCTCAAGACT
 CTCTTCAGCATGCTCAACTATGAGCGAACAAACATCCTCACCTTATGGG
 GTCTTCTGTAUTGGGTATGAATGACATCTACAGGACCTGGCGGGCTTG

- continued

TGCTGCGTGTGCGTGCCTGGACCAGACACCCAGGATGTACTTTGTTAAG
 GCAGATGTGACCGGGGCITGTGATGCCATCCCCCAGGTAAGCTGGTGA
 GGTGTTGCCAATATGTCAGGCACTCGGAGAGCAGCTACTGTATCCGCC
 AGTATGCACTGGTCCGGAGAGATAGCCAAGGCCAAGTCCACAAGTCC
 AGGAGACAGGTACCACCCCTCTGACCTCCAGGCCATACAGGGCAGTT
 CCTTAAGCATCTGCAAGATTCAAGATGCCAGTGCACTGAGGAACCTCGTTG
 TCATCGAGCAGACATCTCTATGAATGAGAGCAGCAGCAGCCGTTGAC
 TTCTCCTGCACCTCCTGCGTACAGTGTGTAAGATTGGTACAGGTG
 CTATAACGCACTGCCAGGGCATTCCCAGGGCTCCAGCCTATCCACCC
 TCTGCAGTCTGTGTTCGGAGACATGGAGAACAGCTGTTGCTGAGGTG
 CAGCGGGATGGGTTGCTTTACGTTTGTGATGACTTCTGTTGGTGC
 GCCTCACTGGACCAAGCAAACCTCCTCACGCACCCCTGGTCCATGGCG
 TTCTGAGATGGGTGCATGATAAACTGCAAGAACAGTGGTGAACCTC
 CCTGTGGAGCCTGGTACCCCTGGGTGGTGCAGCTCCATACCAGCTGC
 TCACTGCCTGTTCCCTGGTGTGGCTGCTGGACACTCAGACTTGG
 AGGTGTTCTGACTACTCAGGTTATGCCAGACCTCAATTAGACGAGC
 CTCACCCCTCAGAGTGTCTCAAAGCTGGGAAGACCATGCGGAACAGCT
 CCTGCGGTTGCGGTTGAAGTGTACGGCTATTCTAGACTTGCAGG
 TGAAACAGCCTCAGACAGTCTGCATCAATATACAAAGATCTCTGCTT
 CAGGCCTACAGGTTCATGCATGTGATTCACTCCCTTGACAGCG
 TGTTAGGAAACCTCACATTCTTCTGGCATCATCTCCAGCCAAGCAT
 CCTGCTGCTATGCTATCCTGAAGGTCAAGAATCCAGGAATGACACTAAAG
 GCCTCTGGCTCTTCTCCTGAAGCCGCACATTGGCTCTGCTACCC
 CCTCCTGCTCAAGCTGGCTGCTCATTCTGTATCTACAAATGCTCTGG
 GACCTCTGAGGACAGCCAAAAACTGCTGTGCGGGAGCTCCAGAGGCG
 ACAATGACCATCTTAAAGCTGCAGCTGACCCAGCCCTAACGACAGACTT
 TCAGACCATTTGGACTAA

[0112] In some embodiments, a mouse TERT mRNA may comprise a mutant mouse TERT sequence. In some embodiments, the mutant mouse TERT mRNA may encode a Y697F mutation in the resulting peptide sequence. In some embodiments a mutation in the TERT mRNA sequence encodes a mutation in the nuclear export signal which may result in nuclear retention of the TERT peptide. In some embodiments, the mutant mouse TERT mRNA sequence may comprise a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO 5:

ATGACCCCGCCTCTCGTTGCCCGCGTGCCTCTGCTGCCAGCCG
 ATACCGGGAGGTGTGGCGCTGGCAACCTTGTGCGGCCCTGGGCG
 AGGGCAGGCCCTGTGCAACCCGGGACCGAAGATCTACCGCATTG
 GTTGCCCAATGCCATGTGCACTGGGCTCACAGCCTCACCTGC

- continued

CGACCTTCCTCCACCAGGTGTCATCCCTGAAAGAGCTGGTGGCCAGGG
 TTGTGCAAGAGACTCTGCAGCGCAACGAGAGAAACGTGCTGGCTTTGGC
 TTTGAGCTGCTAACGAGGCCAGAGCGGGCCCTCCATGGCCTTCACTAG
 TAGCGTGGTAGCTACTGCCCACACTGTTATTGAGACCTCGCTGTCA
 GTGGTGCATGGATGCTACTGTTGAGGCCAGTGGCGACGACCTGCTGGC
 TACCTGCTGGCACACTGTGCTCTTATCTCTGGTCCCCCAGCTGTGC
 CTACCAAGGTGTTGGCTCCCTGTACCAAATTGTGCCACACGGATA
 TCTGGCCCTCTGTGTCAGCTAGTTACAGGCCACCCGACCGTGGCAGG
 AATTCACAACTACAGGTTCTTACACAGATCAAGAGCAGTAGTCGCCA
 GGAAGCACCAGAACCCCTGGCTTGCCATCTCGAGGTACAAAGAGGCATC
 TGAGTCTACCAAGTGTGCTTCAGCTAAAGAGGCCAGATGCTAT
 CCTGTCAGAGACTGGGAGGGAGGGACCCACAGGCAGGTGCTACCAACCC
 ATCAGGCAAATCATGGGTGCAAGTCTGCTCGGTCCCCGAGGTGCTA
 CTGCAGAGAAAGATTGTCTTAAAGGAAAGGTGCTGACCTGAGTCTC
 TCTGGGTCGGTGTGCTGTAACACAAGGCCAGCTCCACATCTGCTGTG
 ACCACCCCGCAAATGCTTTCAGCTCAGGCCATTATTGAGACCAAGAC
 ATTTCTTACTCCAGGGAGATGGCCAAGAGCGTCTAAACCCCTCATTC
 CTACTCAGCAACCTCCAGCCTAACCTGACTGGGCCAGGAGACTGGTGA
 GATCATTTCTGGGCTCAAGGCCAGGACATCAGGACACTCTGAGGA
 CACACCGCTATCGCGTGCAGACTGGCAGATGCGGCCCTGTTCAAACAG
 CTGCTGGTGAACCATGCAAGAGCTGCAAGTCTGCTCAGGTGAC
 TTGCAAGGTTGAAACAGCAAACCAACAGGTGACAGATGCTGAACACCA
 GCCCACCGCACCTCATGGATTGCTCCGCTGCACAGCAGTCCCTGGCAG
 GTATATGGTTTCTCGGGCTGTCTGCAAGGTGGTGTGCTAGTCT
 CTGGGGTACCGGCACAAATGAGCGCCCTCTTAAAGAACTTAAAGAAGT
 TCATCTCGTTGGGAAATACGGCAAGCTACTGCAGGAACGTGATGTGG
 AAAGATGAAAAGTAGAGGATTGCCACTGGCTCCGCAGCAGCCGGGAAAGGA
 CCGTGTCCCCGCGTGCAGAGCACCGTCTGAGGGAGGGATCTGGCTACGT
 TCCCTGTTCTGGCTGATGGACACATACGTTGCTACGCTGCTTAGGTCTTC
 TTTTACATCACAGAGAGCACATTCCAGAAGAACAGGCTCTTCTACCG
 TAAGAGTGTGAGGCAAGCTGCAGAGCATGGAGTCAGGCAACACCTTG
 AGAGAGTGGCTACGGAGCTGTCACAAGAGGAGGTGAGGCATCACCAG
 GACACCTGGCTGCCATGCCATTCTGAGACTGCGCTCATCCCCAAGCC
 CAACGGCTGCCATTGTGAACATGAGTTATAGCATGGTACCAAGAG
 CTTGGGCAGAAGGAAGCAGGCCAGCATTTCACCCAGCGTCTCAAGACT
 CTCTTCAGCATGCTCAACTATGAGCGAACAAACATCCTCACCTTATGGG
 GTCTTCTGACTGGGTATGAAATGACATCTACAGGACCTGGCGGGCTTTG
 TGCTGCGTGTGCGTCTGGACCAGACACCCAGGATGTTCTTGTAAAG
 GCAGATGTGACCGGGCCTATGATGCCATCCCCCAGGGTAAGCTGGTGA

- continued

```

GGTTGTTGCCAATATGATCAGGCACACTGGAGAGCACGTACTGTATCCGCC
AGTATGCAGTGGTCCGGAGAGATAGCCAAGGCCAAGTCCACAAGTCCTT
AGGAGACAGGTACCACCCCTCTGACCTCCAGGCCATACATGGGAGCTT
CTTAAGCATCTGCAGGATTCAAGATGCCAGTGCACGTGAGGAACCTCGGTG
TCATCGAGCAGACATCTCTATGAATGAGAGCAGCAGCAGCCGTGTTGAC
TTCTCCTGCACCTCCTGCGTACAGTGTGTAAGATTGGTGACAGGTG
CTATACGCAGTGCAGGGCATTCCCAGGGCTCCAGCCTATCCACCCCTGC
TCTGCAGTCTGTGTTCGGAGACATGGAGAACAGCTGTTGCTGAGGTG
CAGCGGGATGGGTTGCTTTACGTTTGTGATGACTTCTGTTGGTGCAC
GCCTCACTGGACCAAGCAAAACCTCCTCAGCACCCCTGGTCCATGGCG
TTCCTGAGTATGGGTGCATGATAAACTGCAAGAACAGCTGGTGAACCTC
CTGTGGAGCCTGGTACCCCTGGTGGTGCAGCTCCATACCAAGCTGCTGC
TCACTGCCTGTTCCCTGGTGTGGCTGCTGCGACACTCAGACTTGG
AGGTGTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAGACGAGC
CTCACCTCCAGAGTGTCTCAAAGCTGGAAAGACCATGCGGAACAGCT
CCTGTCGGTCTTGCAGTTGAAGTGTACGGTCTATTCTAGACTTGCAGG
TGAACAGCCTCCAGACAGTCTGCATCAATATACAAAGATCTCCCTGCTT
CAGGCCTACAGGTTCCATGCATGTGTGATTCAAGCTTCCCTTGACCAGCG
TGTTAGGAAGAACCTCACATTCTTCTGGCATCATCTCCAGCCAAGCAT
CCTGCTGCTATGCTATCTGAAGGTCAAGAACCCAGGAATGACACTAAAG
GCCTCTGGCTCTTCCCTGCAAGCCGCACATTGGCTCTGCTACCCAGGC
CTTCCTGCTCAAGCTGGCTGCTCATTCTGTATCTACAAATGCTCTGG
GACCTCTGAGGACAGCCCCAAACTGCTGTGCGGGAGCTCCAGAGGCG
ACAATGACCACCTTAAAGCTGCAGCTGACCCAGCCCTAACGACAGACTT
TCAGACCATTGGACTAA

```

[0113] The compositions comprise a ribonucleic acid, e.g., a synthetic ribonucleic acid coding for a telomerase reverse transcriptase (TERT), wherein telomeres are extended within a cell treated with the compound. The ribonucleic acids used in the transient expression of TERT can comprise a ribonucleic acid coding for a TERT protein. The ribonucleic acids can further comprise one or more sequences that affect the expression and/or stability of the ribonucleic acid in a cell. For example, the ribonucleic acids can contain a 5' cap and untranslated region (UTR) to the 5' and/or 3' side of the coding sequence. The ribonucleic acids may further contain a 3' tail, such as a poly-A tail. The poly-A tail can, for example, increase the stability of the ribonucleic acid. In some embodiments, the poly-A tail comprises at least 25 nucleotides, at least 50 nucleotides, at least 75 nucleotides, at least 100 nucleotides, at least 125 nucleotides, at least 150 nucleotides, at least 200 nucleotides, at least 225 nucleotides, at least 250 nucleotides. In some embodiments, the poly-A tail comprises between 1 and 25 nucleotides, between 25 and 50 nucleotides, between 50 and 75 nucleotides, between 75 and 100 nucleotides, between 100 and 125 nucleotides, between 125 and 150 nucleotides, between 150 and 175 nucleotides, between 175 and 200 nucleotides,

between 200 and 225 nucleotides, or between 225 and 250 nucleotides, inclusive of the endpoints for each range. In some embodiments, the poly-A tail comprises between 100 and 200 nucleotides, inclusive of the endpoints.

[0114] In some embodiments, the 5' cap of the ribonucleic acid is a non-immunogenic cap. In some embodiments, the 5' cap may increase the translation of the ribonucleic acid. In some embodiments, the 5' cap may be treated with phosphatase to modulate the innate immunogenicity of the ribonucleic acid. In some embodiments, the 5' cap is an anti-reverse cap analog ("ARCA"), such as a 3'-O-Me-m7G(5')ppp(5')G RNA cap structure analog. In some embodiments, the 5' cap is m7G(5')ppp(5')(2'OMeA)pG (also known as CleanCap[®] AG). In some embodiments, the 5' cap is m7(3'OMeG)(5')ppp(5')(2'OMeA)pG (also known as CleanCap[®] AG (3' Ome)).

[0115] The above features, or others, may increase translation of the TERT protein encoded by the ribonucleic acid, may increase or decrease the stability of the ribonucleic acid itself in a cell type-specific or cell type-independent manner, or may do both. In some embodiments, the 5' UTR and/or the 3' UTR are from a gene that has a very stable mRNA and/or an mRNA that is rapidly translated, for example, α -globin or β -globin, c-fos, or tobacco etch virus. In some embodiments, the 5' UTR and 3' UTR are from different genes or are from different species than the species into which the compositions are being delivered. The UTRs may also be assemblies of parts of UTRs from the mRNAs of different genes, where the parts are selected to achieve a certain combination of stability and efficiency of translation. The UTRs may also comprise designed sequences that confer properties to the RNA such as cell type-specific stability or cell type-independent stability.

[0116] The ribonucleic acids of the present disclosure may comprise one or more modified nucleosides, and/or comprise primary sequences of nucleosides, that modulate translation, stability, or immunogenicity of the RNA ("mRNA"). Most mature RNA molecules in eukaryotic cells contain nucleosides that are modified versions of the canonical unmodified RNA nucleosides, adenine, cytidine, guanosine, and uridine. For example, the 5' cap of mature RNA comprises a modified nucleoside, and other modified nucleosides often occur elsewhere in the RNA. Those modifications may prevent the RNA from being recognized as a foreign RNA. Synthetic RNA molecules made using certain nucleosides are much less immunogenic than unmodified RNA. The immunogenicity can be reduced even further by purifying the synthetic mRNA, for example by using high performance liquid chromatography (HPLC). The modified nucleosides may be, for example, chosen from the nucleosides listed below. The nucleosides are, in some embodiments, pseudouridine, 1-methylpseudouridine, 2-thiouridine, 5-methoxyuridine, or 5-methylcytidine. The primary sequence may be modified in ways that increase or decrease immunogenicity. Under some circumstances, it may be desirable for the modified RNA to retain some immunogenicity.

[0117] Accordingly, in some embodiments, the ribonucleic acids of the instant compositions comprise a 1-methylpseudouridine, pseudouridine, a 5-methoxyuridine (5-moU), a 2-thiouridine, a 5-methylcytidine, or another modified nucleoside. Modified nucleosides found in eukaryotic cells include m1A 1-methyladenosine, m6A N6-methyladenosine, Am 2'-O-methyladenosine, i6A N6-isopente-

nyladenosine, io6A N6-(cis-hydroxyisopentenyl)adenosine, ms2io6A 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine, g6A N6-glycylcarbamoyladenose, t6A N6-threonylcarbamoyladenose, ms2t6A 2-methylthio-N6-threonyl carbamoyladenose, Ar(p) 2'-O-ribosyladenose (phosphate), m6 2A N6,N6-dimethyladenose, m6Am N6,2'-O-dimethyladenose, m6 2Am N6,N6,2'-O-trimethyladenose, m1Am 1,2'-O-dimethyladenose, m3C 3-methylcytidine, m5C 5-methylcytidine, Cm 2'-O-methylcytidine, ac4C N4-acetylcytidine, f5C 5-formylcytidine, m4C N4-methylcytidine, hm5C 5-hydroxymethylcytidine, f5Cm 5-formyl-2'-O-methylcytidine, m1G 1-methylguanosine, m2G N2-methylguanosine, m7G 7-methylguanosine, Gm 2'-O-methylguanosine, m2 2G N2,N2-dimethylguanosine, Gr(p) 2'-O-ribosylguanosine (phosphate), yW wybutsine, o2yW peroxywybutosine, OhyW hydroxywybutosine, OhyW* undermodified hydroxywybutosine, imG wyosine, m2,7G N2,7-dimethylguanosine, m2,2,7G N2,N2,7-trimethylguanosine I inosine, m1I 1-methylinosine, Im 2'-O-methylinosine, Q queuosine, galQ galactosyl-queuosine, manQ mannosyl-queuosine, Ψ pseudouridine, D dihydrouridine, m5U 5-methyluridine, Um 2'-O-methyluridine, m5Um 5,2'-O-dimethyluridine, m1Ψ 1-methylpseudouridine, Ψm 2'-O-methylpseudouridine, s2U 2-thiouridine, ho5U 5-hydroxyuridine, chm5U 5-(carboxyhydroxymethyl)uridine, mchm5U 5-(carboxyhydroxymethyl)uridine, methyl ester mcm5U 5-methoxycarbonylmethyluridine, mcm5Um 5-methoxycarbonylmethyl-2'-O-methyluridine, mcm5s2U 5-methoxycarbonylmethyl-2-thiouridine, ncm5U 5-carbamoylmethyluridine, ncm5Um 5-carbamoylmethyl-2'-O-methyluridine, cmnm5U 5-carboxymethylaminomethyluridine, m3U 3-methyluridine, m1acp3Ψ 1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine, cm5U 5-carboxymethyluridine, m3Um 3,2'-O-dimethyluridine, m5D 5-methylidihydrouridine, tm5U 5-taurinomethyluridine, tm5s2U 5-taurinomethyl-2-thiouridine, 2-Aminoadenosine, 2-Amino-6-chloropurineriboside, 8-Azaadenosine, 6-Chloropurineriboside, 5-Iodocytidine, 5-Iodouridine, Inosine, 2'-O-Methylinosine, Xanthosine, 4-Thiouridine, 06-Methylguanosine, 5,6-Dihydrouridine, 2-Thiacytidine, 6-Azacytidine, 6-Azauridine, 2'-O-Methyl-2-aminoadenosine, 2'-O-Methylpseudouridine, N1-Methyladenosine, 2'-O-Methyl-5-methyluridine, 7-Deazaguanosine, 8-Azidoadenosine, 5-Bromocytidine, 5-Bromouridine, 7-Deazaadenosine, 5-Aminoallyluridine, 5-Aminoallylcystidine, 8-Oxoguanosine, 2-Aminopurine-riboside, Pseudoisocytidine, N1-Methylpseudouridine, 5,6-Dihydro-5-Methyluridine, N6-Methyl-2-Aminoadenosine, 5-Carboxycytidine, 5-Hydroxymethyluridine, Thienoguanosine, 5-Hydroxycytidine, 5-Formyluridine, 5-Carboxyuridine, 5-Methoxyuridine, 5-Methoxycytidine, Thienouridine, 5-Carboxymethylesteruridine, Thienocytidine, 8-Oxoadenosine, Isoguanosine, N1-Ethylpseudouridine, N1-Methyl-2'-O-Methylpseudouridine, N1-Methoxymethylpseudouridine, N1-Propylpseudouridine, 2'-O-Methyl-N6-Methyladenosine, 2-Amino-6-Cl-purine-2'-deoxyriboside, 2-Amino-2'-deoxyadenosine, 2-Aminopurine-2'-deoxyriboside, 5-Bromo-2'-deoxycytidine, 5-Bromo-2'-deoxyuridine, 6-Chloropurine-2'-deoxyriboside, 7-Deaza-2'-deoxyadenosine, 7-Deaza-2'-deoxyguanosine, 2'-Deoxyinosine, 5-Propynyl-2'-deoxycytidine, 5-Propynyl-2'-deoxyuridine, 5-Fluoro-2'-deoxyuridine, 5-Iodo-2'-deoxycytidine, 5-Iodo-2'-deoxyuridine, N6-Methyl-2'-deoxyadenosine, 5-Methyl-2'-deoxycytidine, 06-Methyl-2'

deoxyguanosine, N2-Methyl-2'-deoxyguanosine, 8-Oxo-2'-deoxyadenosine, 8-Oxo-2'-deoxyguanosine, 2-Thiothymidine, 2'-Deoxy-P-nucleoside, 5-Hydroxy-2'-deoxycytidine, 4-Thiothymidine, 2-Thio-2'-deoxycytidine, 6-Aza-2'-deoxyuridine, 6-Thio-2'-deoxyguanosine, 8-Chloro-2'-deoxyadenosine, 5-Aminoallyl-2'-deoxycytidine, 5-Aminoallyl-2'-deoxyuridine, N4-Methyl-2'-deoxycytidine, 2'-Deoxyzebularine, 5-Hydroxymethyl-2'-deoxyuridine, 5-Hydroxymethyl-2'-deoxycytidine, 5-Propargylamino-2'-deoxycytidine, 5-Propargylamino-2'-deoxyuridine, 5-Carboxy-2'-deoxycytidine, 5-Formyl-2'-deoxycytidine, 5-[(3-Indolyl)propionamide-N-allyl]-2'-deoxyuridine, 5-Carboxy-2'-deoxyuridine, 5-Formyl-2'-deoxyuridine, 7-Deaza-7-Propargylamino-2'-deoxyadenosine, 7-Deaza-7-Propargylamino-2'-deoxyguanosine, Biotin-16-Aminoallyl-2'-dUTP, Biotin-16-Aminoallyl-2'-dCTP, Biotin-16-Aminoallylcystidine, N4-Biotin-OBEA-2'-deoxycytidine, Biotin-16-Aminoallyl-Uridine, DabcyL-5-3-Aminoallyl-2'-dUTP, Desthiobiotin-6-Aminoallyl-2'-deoxycytidine, Desthiobiotin-16-Aminoallyl-Uridine, Biotin-16-7-Deaza-7-Propargylamino-2'-deoxyguanosine, Cyanine 3-5-Propargylamino-2'-deoxycytidine, Cyanine 3-6-Propargylamino-2'-deoxyuridine, Cyanine 5-6-Propargylamino-2'-deoxycytidine, Cyanine 5-6-Propargylamino-2'-deoxyuridine, Cyanine 3-Aminoallylcystidine, Cyanine 3-Aminoallyluridine, Cyanine 5-Aminoallylcystidine, Cyanine 5-Aminoallyluridine, Cyanine 7-Aminoallyluridine, 2'-Fluoro-2'-deoxyadenosine, 2'-Fluoro-2'-deoxycytidine, 2'-Fluoro-2'-deoxyguanosine, 2'-Fluoro-2'-deoxyuridine, 2'-O-Methyladenosine, 2'-O-Methylcytidine, 2'-O-Methylguanosine, 2'-O-Methyluridine, Puromycin, 2'-Amino-2'-deoxycytidine, 2'-Amino-2'-deoxyuridine, 2'-Azido-2'-deoxycytidine, 2'-Azido-2'-deoxyuridine, Aracytidine, Arauridine, 2'-Azido-2'-deoxyadenosine, 2'-Amino-2'-deoxyadenosine, Araadenosine, 2'-Fluoro-thymidine, 3'-O-Methyladenosine, 3'-O-Methylcytidine, 3'-O-Methylguanosine, 3'-O-Methyluridine, 2'-Azido-2'-deoxyguanosine, Araguanosine, 2'-Deoxyuridine, 3'-O-(2-nitrobenzyl)-2'-Deoxyadenosine, 3'-O-(2-nitrobenzyl)-2'-Deoxyinosine, 3'-Deoxyadenosine, 3'-Deoxyguanosine, 3'-Deoxycytidine, 3'-Deoxy-5-Methyluridine, 3'-Deoxyuridine, 2',3'-Dideoxyadenosine, 2',3'-Dideoxyguanosine, 2',3'-Dideoxyuridine, 2',3'-Dideoxythymidine, 2',3'-Dideoxycytidine, 3'-Azido-2',3'-dideoxyadenosine, 3'-Azido-2',3'-dideoxythymidine, 3'-Amino-2',3'-dideoxyadenosine, 3'-Amino-2',3'-dideoxyguanosine, 3'-Amino-2',3'-dideoxythymidine, 3'-Azido-2',3'-dideoxycytidine, 3'-Azido-2',3'-dideoxyuridine, 3'-Deoxyguanosine, 3'-Deoxyuridine, 3'-Deoxycytidine, 3'-Deoxythymidine, 3'-O-(1-Thiophosphate), 2'-Deoxycytidine-5'-O-(1-Thiophosphate), 2'-Deoxyguanosine-5'-O-(1-Thiophosphate), 2'-Deoxythymidine-5'-O-(1-Thiophosphate), Adenosine-5'-O-(1-Thiophosphate), Cytidine-5'-O-(1-Thiophosphate), Guanosine-5'-O-(1-Thiophosphate), Uridine-5'-O-(1-Thiophosphate), 2',3'-Dideoxyadenosine-5'-O-(1-Thiophosphate), 2',3'-Dideoxyguanosine-5'-O-(1-Thiophosphate), 2',3'-Dideoxyuridine-5'-O-(1-Thiophosphate), 3'-Deoxythymidine-5'-O-(1-Thiophosphate), 3'-Azido-2',3'-dideoxythymidine-5'-O-(1-Thiophosphate), 2',3'-Dideoxyuridine-5'-O-(1-Thiophosphate), 2'-Deoxyadenosine-5'-O-(1-Boranophosphate), 2'-Deoxycytidine-5'-O-(1-

Boranophosphate), 2'-Deoxyguanosine-5'-O-(1-Borano-phosphate), and 2'-Deoxythymidine-5'-O-(1-Boranophosphate).

[0118] Without intending to be bound by theory, the presence of the modified nucleosides, and/or sequences of nucleosides that alter secondary structure of the RNA and/or binding of RNA to RNA binding proteins or microRNA, may enable mRNA to avoid activation of an immune response mediated by various receptors, including the Toll-like receptors and RIG-1. Non-immunogenic mRNA has been used as a therapeutic agent in mice via topical delivery. Kormann et al. (2011) *Nature Biotechnology* 29:154-157. In some embodiments, the ribonucleic acids comprise more than one of the above nucleosides or combination of the above nucleosides. In some embodiments, the ribonucleic acids comprise 1-methylpseudouridine, 5-methoxyuridine, or pseudouridine and 5-methylcytidine.

[0119] In some embodiments, an immune response to the mRNA may be desired, and the RNA may be modified to induce an optimal level of innate immunity. In other embodiments, an immune response to the mRNA may not be desired, and the RNA may be modified in order to minimize such a reaction. The RNA can be modified for either situation.

[0120] The ribonucleic acid molecules can be synthetic ribonucleic acids. The term "synthetic", as used herein, can mean that the ribonucleic acids are in some embodiments prepared using the tools of molecular biology under the direction of a human, for example as described below. The synthetic ribonucleic acids may, for example, be prepared by *in vitro* synthesis using cellular extracts or purified enzymes and nucleic acid templates. The synthetic ribonucleic acids may in some embodiments be prepared by chemical synthesis, either partially or completely. Alternatively, or in addition, the synthetic ribonucleic acids may in some embodiments be prepared by engineered expression in a cell, followed by disruption of the cell and at least partial purification of the ribonucleic acid.

[0121] The ribonucleic acids of the present disclosure may be prepared using a variety of techniques, as would be

understood by one of ordinary skill in the art. In some embodiments, the ribonucleic acids may be prepared by *in vitro* synthesis. In some embodiments, the ribonucleic acids may be prepared by chemical synthesis. In some embodiments, the ribonucleic acids may be prepared by a combination of *in vitro* synthesis and chemical synthesis. As described above, the term "synthetic" should be understood to include ribonucleic acids that are prepared either by chemical synthesis, by *in vitro* synthesis, by expression *in vivo* and at least partial purification, or by a combination of such, or other, chemical or molecular biological methods.

[0122] The ribonucleic acids may, in some embodiments, be purified. As noted above, purification may reduce immunogenicity of the ribonucleic acids and may be advantageous in some circumstances. In some embodiments, the ribonucleic acids are purified by one or more of HPLC, DNase treatment, protease treatment, or by affinity capture and elution.

[0123] The protein structure of TERT can include at least three distinct domains: a long extension at the amino-terminus (the N-terminal extension, NTE) that contains conserved domains and an unstructured linker region; a catalytic reverse-transcriptase domain in the middle of the primary sequence that includes seven conserved reverse transcriptase (RT) motifs; and a short extension at the carboxyl-terminus. In some embodiments, the ribonucleic acid codes for a full-length TERT. In some embodiments, the ribonucleic acid codes for a catalytic reverse transcriptase domain of TERT. In some embodiments, the ribonucleic acid codes for a polypeptide having TERT activity. TERT activity may be measured using known methods including the telomerase repeat amplification protocol (TRAP).

[0124] The TERT encoded by the ribonucleic acids of the instant disclosure may be a mammalian, avian, reptilian, or fish TERT. In some embodiments, the TERT is a mammalian TERT, such as human TERT. Meyerson et al. (1997) *Cell* 90:785-795; Nakamura et al. (1997) *Science* 277:955-959; Wick et al. (1999) *Gene* 232:97-106.

[0125] The amino acid sequence of two human TERT isoforms are available as NCBI Reference Sequences: NP_937983.2 and NP_001180305.1.

The amino acid sequence of human TERT isoform 1 may comprise or consist of the sequence of SEQ ID NO: 6 (also described at GenBank Accession No. NP_937983.2):
 1 mprparcrav rsllrshyre vplatfvrr lgpqgwrlvq rgdpaafra lvaqlvcvpw
 61 darpppaaps frqvsclkel varvlqrlnce rgaknvlafg falldgargg ppeafsttsvr
 121 sylpntvtada lrgsgawgll lrrvgddvlv hllarcalfv lvapscayqv cgpplyqlga
 181 atqarpppha sgprrrlgce rawnhsvrea gvplglpapg arrrggsasr slplpkrrpr
 241 gaapepertp vggswahpg rtrgpsdrgf cvvsparpae eatslegals gtrhshpsvg
 301 rqhhagppst srpprpwdtp cppvyaethk flyssgdkeq lrpsflssl rpsltgarl
 361 vetiflgsrp wmpgtprrlp rlpqrywqmr plflellgnh aqcpygvl1k thcplraavt
 421 paagvcarek pgqsvaapee edtdprrlvq llrqhsspwq vygfvrac1r rlvpqglwgs
 481 rhnerrflrn tkkfislgkh akls1qeltw kmsvrdcawl rrspgvgcvp aaehrlreei
 541 lakflhwlm s vvvellrsf fyvtettfqk nrlffyrksv wsklqsigir qhlkrvqlre
 601 lseaevrqhr earpalltsr lrfipkpdgl rpivnmdyvv gartfrrekr aerltsrvka
 661 lfsvlnyera rrpqllgasv lglddihraw rtfvrlvraq dpppelyfvk vdvtgaydti

-continued

721 pqdrltevia siikpqntyc vrryavvqka ahghvrkafk shvstltdlq pymrqfvahl
 781 qetsplrdav vieqssslne assglfdvfl rfmchhvri rgksyvqcqg ipqgsilstl
 841 lcslcygdm e nklgagirrd gllrlvddf llvphltha ktflrtlvrq vpeygcvvnl
 901 rktvvnfpve dealggtavf qmpahglfpw cgllldtrtl evqsdysya rtsirasltf
 961 nrgfkagrnm rrklfgvrlr kchsifldlq vnslqtvctn iykillqay rfhacvlqlp
 1021 fhqqvwknpt fflrvvisdta slcysilkak nagmslgakg aagplpseav qwlchqafll
 1081 kltrhrvtyv pllgslrtaq tqlsrklpgt tltaleaaan palpsdfkti ld.

The nucleic acid sequence of human TERT isoform 1 may comprise or consist of the sequence SEQ ID NO: 7 (also described at GenBank Accession No. NM_198253.3):

1 ctctccatcgcc ggccggagtt tcaggcagcg ctgcgtctcg ctgcgcacgt gggaaaggccct
 61 ggccccggcc acccccgcga tgeccgcgcgc tccccgctgc cgageccgtgc gctccctgtct
 121 ggcgcagccac taccgcgagg tgcgcgcgtt ggcacacgttc gtgcggcgcc tggggcccca
 181 gggctggccg ctggtgccgc gcggggaccc ggccggcttc cgcgcgcgtgg tggcccagtgg
 241 cctggtgtgc gtgcgcctggg acgcacggcc gccccccggcc gcccccttcct tccggccaggt
 301 gtcctgcctg aaggagctgg tggcccgagt gtcgcagagg ctgtgcgcgc gcccgcgcgaa
 361 gaacgtgtgc gccttcggct tcgcgcgtgc ggacggggcc cggggggggcc ccccccggaggc
 421 cttcaccacc accgtgcgcga gctacactgcc caacacgggt accgcacgcac tgccggggagg
 481 cggggcgtgg gggctgcgtgc tgcgcgcgtt gggcgacgc gtgcgtggcc acctgtgtggc
 541 acgctgcgcg ctctttgtgc tgggtggctcc cagctgcgc taccagggtt gccccccgc
 601 gctgttaccag ctccggcgctg ccactcaggc cccggggggcc ccacacgccta gtggaccccg
 661 aaggcgtctg ggtatgcgaac gggcttgaa ccatacggtc agggaggccg gggccccccct
 721 gggcctgcca gccccgggtt cgaggaggcc cggggggcgtt gccacgcgaa gtctggccgtt
 781 gccaagagg cccaggcggt ggcgtgcctt tgagccggag cggacggccg ttggccagg
 841 gtcctggccc cacccggca ggacgcgtgg accgaggttgc cgtggtttct gtgtgggtgc
 901 acctggccaga cccggccgaag aagccaccc tttggagggt ggcgtctctg gcacgcgc
 961 ctccccaccca tccgtggccc gcaacgcacca cggggggccccc ccacccacat cggggccacc
 1021 acgtccctgg gacacgcctt gtcctgggggt gtacgcccag accaaggact tcctctactc
 1081 ctcaggcgac aaggaggcgc tgcggccctc ctccctactc agctctctga ggcccaaggct
 1141 gactggcgctt cggaggcgctt tggagaccat ctttctgggt tccaggccctt ggtatggccagg
 1201 gactccccgc aggttggccc gcctggccca ggcgtactgg caaatgcggc ccctgtttct
 1261 ggagctgctt gggaaaccacg cgcagtgccc ctacgggggtt ctccctcaaga cgcactgccc
 1321 gctgcgtactt cgggtcaccc cagcagccgg tgcgtgtgc cgggagaagc cccagggttc
 1381 tgtggccggcc cccggaggagg aggacacaga ccccccgtgc tgggtgcgc tgctccgc
 1441 gcacacgcgc cccctggcagg tgcgtggctt cgtgcggggcc tgcctgcgc ggcgtggcc
 1501 cccaggccctc tggggctcca ggcacaacacg acgcgcgttc ctccaggaaaca ccaagaagtt
 1561 catctccctg gggaaacatg ccaagctctc gtcgcaggag ctgcacgtgg agatgagcgt
 1621 gggggactgc gcttggctgc gcaggagccc aggggtggc tgcgttccgg cgcacaggca
 1681 ccgtctgcgtt gaggagatcc tggcaagtt cctgcactgg ctgcacgtgg tgatgagtg
 1741 cgagctgcgc aggttttctt ttatgtca gggacaccacg tttcaaaaaga acaggcttt
 1801 tttctaccgg aagagtgtctt ggagcaagttt gcaaaacattt ggaatcagac agcacttgaa

-continued

1861 gaggggtgcag ctgcgggagc tgcggaaagc agaggtcagg cagcatcgaa aagccaggcc
1921 cgccctgctg acgtccagac tccgcttcat ccccaagcct gacgggctgc ggccgattgt
1981 gaacatggac tacgtcggtt gagccagaac gttccgcaga gaaaagaggg ccgagcgtct
2041 cacctcgagg gtgaaggcac tgttcagcgt gctcaactac gagcgggcgc ggcccccgg
2101 cctcctggc gcctctgtgc tgggccttggc cgatataccac agggcctggc gcaccttcgt
2161 gctgcgtgtc cgccccagg accggccgc tgagctgtac ttgtcaagg tggatgtgac
2221 gggcgctac gacaccatcc cccaggacag gtcacggag gtcatcgcca gcatcatcaa
2281 accccagaac acgtactcg tgcgtcgta tgccgtggc cagaaggccg cccatggca
2341 cgtcccaag gccttcaaga gccacgtctc taccttgaca gacccatcgac cgtacatcg
2401 acagttcggtg gtcacactgc aggagaccag cccgctgagg gatgcgtcg tcacgagca
2461 gagctctcc ctgaatgagg ccagcgtgg cctctcgac gtcctcttac gtcctatgtg
2521 ccaccacgac gtgcgcataa gggcaagtc ctacgtccag tgccaggggg tcccgaggg
2581 ctccatccctc tccacgctgc tctgcgcctt gtgctacggc gacatggaga acaagctgtt
2641 tgcggggatt cggcgccggc ggctgcgttgc ggcgtttggc gatgtttct tggatgtgac
2701 acctcacctc accccacgcga aaaccttcctt caggaccctg gtccgagggtg tccctgagta
2761 tggctcgctg gtgaacttgc ggaagacagt ggtgaacttc cctgtagaag acggggccct
2821 gggtgtgcacg gctttgttc agatgcggc ccacggctta ttccctggc gggcgtgt
2881 gctggatacc cggaccctgg aggtgcagag cgactactcc agctatgccc ggacctccat
2941 cagaggcagt ctcacccatca accggcgctt caagggtggg aggaacatgc gtcgcaaact
3001 ctttgggtc ttgcggctga agtgcacag cctgtttctg gatggcagg tgaacagct
3061 ccagacggtg tgcaccaaca tctacaagat cctctcgac caggggtaca ggttcaecgc
3121 atgtgtgctg cagctccat ttcatcagca agtttggaaag aacccacat tttccctgct
3181 cgtcatctc gacacggcct ccctctgcta ctccatccctg aaagccaaga acgcaggat
3241 gtcgctgggg gccaaggccg cccggccccc tctgcctcc gagggcgtgc agtggctgt
3301 ccaccaagca ttccctgctca agtgcactcg acaccgtgtc acctacgtgc cactctgg
3361 gtcactcagg acagccaga cgcgcgttag tcggaaacttc cccggggacga cgctgactgc
3421 cctggaggcc gcagccaaacc cggcactgcc ctcagacttc aagaccatcc togactgatg
3481 gccacccggc cacagccagg cccggagcag acaccggcag ccctgtcactc cccggctcta
3541 cgtcccaagg agggggggc ggccacacc caggccgcac ccgcgtggag tctgaggcc
3601 gagtgagtgt ttggccgagg cctgcgtgtc cggctgaagg ctgagtgtcc ggctgaggcc
3661 tgagcgtgtc tccagccaaag ggctgagtgtt ccagcaca tgcgtcttc acttcccccac
3721 aggctggcgc tcggctccac cccaggccca gctttccctc accaggagcc cggcttccac
3781 tccccacata ggaatagtcc atccccagat tgcgcattgt tcacccctcg ccctggccctc
3841 ctttgccttc caccggccacc atccagggtt agaccctgtt aaggaccctg ggagctctgg
3901 gaattttggag tgaccaaagg tgcgtccatgtt acacaggcga ggaccctgca cctggatggg
3961 ggtccctgtg ggtcaaaattt gggggagggtt ctgtggagttt aaaaatactga atatatgatg
4021 ttttcagttt tgaaaaaaa.

-continued

The amino acid sequence of human TERT isoform 2 may comprise or consist of the sequence of SEQ ID NO: 8 (also described at GenBank Accession No. NP_001180305.1):

```

1 mprapprcrav rsllrshyre vlplatfvrr lgpqgwr1vq rgdpaafral vaqclvcvpw
61 darpppaaps frqvsclkel varvlqlrve rgaknvlafg falldgargg ppeafttsvr
121 sylpnvttda lrgsgawgll lrrvgddv1v hllarcalfv lvapscayqv cgpplyqlga
181 atqarpppha sgprrrlgce rawnhsvrea gvplglpapg arrrggsasr slplpkrr
241 gaapepertp vgggswahpg rtrgpsdrgf cvvsparpae eatslegals gtrhshpsvg
301 rqhhagppst srpprpwdtp cppvyaetkh flyssgdkeq lrpsflssl rpsltgarrl
361 vetiflgsrp wmpgtprrlp rlpqrywqmr plflellgnh aqcpygvllk thcplraavt
421 paagvcarek pqgsvaapee edtdprrlvq llrqhssp1wq vygfvraclr rlvpplwg
481 rhnerrflrn tkkfislgkh aklslqeltw kmsvrdcawl rrspgvgcvp aaeahlreei
541 lakflhwlm1 vvvvellarf fyvtettf1qk nrlffyrksv wsklqsigir qhlkrvqlre
601 lseaevrqhr earpalltsr lrfipkp1dgl rpivnmdyvv gartfrrekr aerltsrvka
661 lfsvlnyera rrp1llgasv lglddi1raw rtfv1rvraq dpppelyfvk vdvtgaydti
721 pqdr1tevia siikpqntyc vrryavvqka ahghvrkafk shvst1tdlq pymrqfvahl
781 qetsplrdav vieqsslne assglfdvfl rfmchhavri rgksyvqcqg ipqgsil1stl
841 lcs1cygdme nk1fagirrd g1llrlvddf llvtp1hla ktflsyarts irasltfnrg
901 fkagrnmrrk lfgv1rlkch slf1dlqvns lqtvctniyk ill1qayrfh acvlqlpfhq
961 qvwknptffl rvisdtas1c ysilkaknag mslgakgaag plpseavqwl chqaf11klt
1021 rhrvtyvpl1 gslrtaqtql srklpg1tt1 aleaaanpal psdfktild.

```

The amino acid sequence of human TERT isoform 2 may comprise or consist of the sequence of SEQ ID NO: 9 (also described at GenBank Accession No. NP_001193376.3):

```

1 ctctcc1tcgc ggcgcgagtt tcaggcagcg ctgcgtctg ctgcgcacgt gggaa1ccct
61 ggccccggcc acccccgcga tgcggcgagc tccccgtgc cgagccgtgc gtc1ccctgtc
121 ggcgcagccac taccgcgagg tgctgcgcgt ggccacgttc gtgcggcgcc tggggccccca
181 gggctggcg1 ctgg1tgcage gcggggaccc ggccggcttgc cgcgcgctgg tggcccagtgc
241 cctgg1tgcgc tgcgcctggg acgcacggcc gc1ccccccgc gc1cccctect tccgc1cagg1t
301 gtcctgcctg aaggagctgg tggcccaggt gctgcagagg ctgtgcgcagc gcggcgccgaa
361 gaacgtgtgc gccttcggct tcgcgcgtgc ggacggggcc cgcggggggcc cccccggagc
421 cttcaccaacc agcgtgcgca gctacactgca caacacggtg accgcacgcac tgcggggggag
481 cggggcgtgg gggctgcgtgc tgcgcctggcgt gggcgacgc gtgcgtgg1tc acctgtgcgc
541 acgctgcgcg ctctttgtgc tgg1tgcgc tgcgtgcgc taccaggtgt gcggggccgc
601 gctgtaccag ctgcggcgctc ccactcaggc cccggcccccg ccacacgcta gtggaccccg
661 aaggcgtctg ggatgcgaac gggctggaa ccatacgctc agggaggccg gggccccct
721 gggcctgcca gccccgggtg cgaggaggcg cggggggcgt gccagccgaa gtc1tgcgcgtt
781 gccaagagg cccaggcg1tgc ggcgtgc1ccc tgagccggag cggacggcccg ttgggcagg
841 gtcctggggcc caccggggca ggacgcgtgc accgagtgac cgtgg1ttct gtgtgg1tgc
901 acctgcccaga cccggcgaag aagccaccc1t tttggagggt ggcgtctctg gca1cgccca
961 ctcccaccca tccgtggggcc gca1gcacca cgcggggcccc cc1atccacat cgcggccacc
1021 acgtccctgg gacacgcctt gtccccccgt gta1gcgcag acca1agcact tcctctactc
1081 ctcaggcgac aaggaggcgc tgcggccctc ct1ccctactc agctctctga ggcccagcc

```

-continued

1141 gactggcgct cgaggagctcg tggagaccat ctttctgggt tccaggccct ggatgccagg
1201 gactccccgc aggttccccc gcctgccccca gcgctactgg caaatgcggc ccctgttct
1261 ggagctgctt gggAACCAcG cgcaGtGCC ctacGGGGTG ctccCTcaAGA cgcaCTGCC
1321 gctgCGAGCT gCGGTcACCC cAGCAGCGG TGtCTGTGCC CGGGAGAAGC CCCAGGGCTC
1381 tGTGGCGGCC CCGAGGAGG AGGACACAGA CCCCCGTcGC CTGGTGCAGC TGtCCGCCA
1441 GCAcAGCAGC CCCTGGCAGG TGTACGGTT CGTGCAGGCC TGCCtGCAGC GGtGGtGCC
1501 CCCAGGcCTC TGGGGCTCCA GGCACAACGA ACGCCGCTC CTCAGGAACA CCAAGAAgTT
1561 CATCTCCCTG GGGAAAGCATG CCAAGCTCTC GETGCAGGAG CTGACGTGGA AGATGAGCgT
1621 GCGGGACTG CCGTGGCTG GCAGGAGGCC AGGGGTTGGC TGtGtCCGG CGCAGAGCA
1681 CGGTCTGCGT GAGGAGATCC TGGCCAAGT CCTGCACTGG CTGATGAGTG TGtACGTGt
1741 CGAGCTGCTC AGGTCTTCT TTtATGTcAC GGAGACCACG TTtCAAAGA ACAGGCTT
1801 TTtCTACCGG AAGAGTGTCT GGAGCAAGTT GCAAAGCATT GGAATCAGAC AGCACTTGA
1861 GAGGGTGCAG CTGCGGGAGC TGtCGGAAGC AGAGGTcAGG CAGCATCGGG AAGCCAGGCC
1921 CGCCCTGCTG ACGTCCAGAC TCCGCTTcAT CCCAAGCCT GACGGGCTGC GGCCATTG
1981 GAACATGGAC TACGTcGTGG GAGCCAGAAc GTtCCGCAGA GAAAAGAGGG CGAGCGTCT
2041 CACCTCGAGG GTGAAGGCAC TGtTCAGCGT GCTCAACTAC GAGCAGGCC GGCAGCCCG
2101 CCTCCTGGC GCTCTGTGc TGGGCTGGA CGATATCCAC AGGGCCTGGC GCACTTCTG
2161 GCTGCCTGTG CGGGCCcAGG ACCCGCCGCC TGAGCTGTAC TTtGTCAAGG TGGATGTGAC
2221 GGGCGCTAC GACACCATCC CCCAGGAGC GCTCACGGAG GTCATGCCA GCACTCATCAA
2281 ACCCCAGAAC ACGTACTGEG TGCGTGGTA TGCCGTGGC CAGAAGGCCG CCCATGGGCA
2341 CGTCCGCAAG GCCTTCAAGA GCCACGTCTC TACCTTGACA GACCTCCAGC CGTACATCG
2401 ACAGTTCTG GCTACACTGC AGGAGACCAg CCGCTGAGG GATGCGTcG TCACTGAGCA
2461 GAGCTCTCC CTGAATGAGG CGAGCAGTGG CCTCTTCGAC GCTTCTCTAC GCTTCTGTG
2521 CCACCAcGCC GTGCGCATCA GGGCAAGTC CTACGTCCAG TOCCAGGGGA TCCCGCAGGG
2581 CTCCATCTC TCCACGCTG TCTGCAGCT GtGCTACGGC GACATGGAGA ACAAGCTGTT
2641 TGCGGGGATT CGGGGGAGC GGCTGCTCT GCGTTGGTG GATGATTCT TGtGGTGA
2701 ACCTCACCTC ACCACAGCGA AAACCTTCTC CAGCTATGCC CGGACCTCCA TCAAGGCCAG
2761 TCTCACCTC AACCGCGGCT TCAAGGCTGG GAGGAACATG CGTCGAAAC TCTTGGGTT
2821 CTTGCGGTG AAGTGTcACA GCTGTTCTC GGATTTGAG GtGAACAGCC TCCAGACCGT
2881 GTGCACCAAC ATCTACAAGA TCTCTCTGCT GCAAGGCGTAC AGGTTCAcG CAtGTGTG
2941 GCAAGCTCCCA TTtCATCAGe AAGTTGGAA GAAcCCcACA TTtTCTGc GCGTCACTC
3001 TGACACGGCC TCCCTCTGCT ACTCCACCTC GAAAGCCAAG AACGCAAGGA TGtCGCTGG
3061 GCGCAAGGGC GCGCGGGCC CTCGCCCCC CGAGGGCGTG CAGTGGCTGT GCGACCAAGC
3121 ATTCTGCTC AAGCTGACTC GACACCGTGT CACCTACGTG CCACtCTGG GGTCACTCAG
3181 GACAGCCAG ACGCAAGCTGA GTCGGAAGCT CCGGGGGAGC ACGCTGACTG CCCTGGAGGC
3241 CGCAGCCAAC CGGGCACTGC CCTCAGACTT CAAGACCATC CTGGACTGAT GCGCACCCG
3301 CCACAGCCAG GCGAGAGCA GACACCGAGA GCGCTGCAC GCGGGCTCT ACGTCCCAGG
3361 GAGGGAGGGG CGGCCACAC CCAGGCCCCG ACCGCTGGGA GTCTGAGGCC TGAGTGA
3421 TTtGGCCGAG GCGCTGATGT CGGGCTGAAG GCTGAGTGTc CGGCTGAGGC CTGAGCGAGT

- continued

```

3481 gtccagccaa gggtgagtg tccagcacac ctggcgctt cactccccca caggctggcg
3541 ctcggctcca ccccgaggcc agttttctt caccaggagc ccggcttcca ctccccat
3601 aggaatagtc catccccaga ttgcgcattt ttcacccctc gcacctgcctt cttgcctt
3661 ccaccccccac catccaggtg gagaccctga gaaggaccct gggagctctg gaaatttgg
3721 gtgaccaaag gtgtgcctg tacacaggcg aggaccctgc acctggatgg gggccctgt
3781 gggtaaaatt gggggaggt gctgtgggaa taataactg aatatatgat ttttcagtt
3841 ttgaaaaaaaaa.

```

[0126] In some embodiments, a human TERT mRNA may comprise a wild type TERT sequence. In some embodiments, the wild type TERT sequence may comprise a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO 30:

```

ATGCCGCGCCTCCCCGTGCCGAGCCGTGCGCTCCCTGTCGCGC
AGCCACTACCGCGAGGTGCTGCCGCTGGCACGTCGTGCGCGC
CTGGGGCCCCAGGGCTGGCGCTGGTGCAGCGCGGGGACCCGGG
GCTTCCGCGCCTGGTGGCCAGTGCTGGTGTGCGTGCCTGG
GACGCACGGCGCCCCCGCGCCCCCTCTTCCGCCAGGTGTCC
TGCCTGAAGGAGCTGGTGGCCGAGTGCTGCAGAGGCTGTGCGAG
CGCGCGCGAAGAACGTGCTGGCCTTCGCGTTCGCGCTGGC
GGGGCCCCGGGGCCCCCGAGGCCTTACACCACCGCGTGC
AGCTACCTGCCAACACGGTAGCCGACGCACTGCGGGGAGCGG
GCGTGGGGCTGCTGCTGCCGCGTGGCGACGACGTGCTGGTT
CACCTGCTGGCACGCTGCCGCTTTGTGCTGGTGGCTCCAGC
TGCCTACAGGTGTGGGGCGCCGCTGTACAGCTGCCGCT
GCCACTCAGGCCGGCCCCGCCACACGCTAGTGGACCCGAAGG
CGTCTGGGATGCGAACGGCCTGGAACCATAGCGTCAGGGAGGCC
GGGTCCCCCTGGGCCCTGCCAGCCCCGGGCGAGGAGGCCGGG
GGCAGTGCCAGCCGAAGTCTGCCGTTGCCAAGAGGCCAGGGT
GGCCTGCCCTGAGCGGAGCGGACGCCGTTGGCAGGGTCC
TGGGCCACCCGGGAGGACGCGTGGACCGAGTGACCGTGGTTTC
TGTGTGGTGTACCTGCCAGACCCGCCGAAGAACGCCACCTTTG
GAGGGTGCCTCTGGCACGCCACTCCCACCCATCCGTGGG
GCCAGCACGCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
CCCTGGGACACGCCCTGTCCCCGGTGTACGCCAGACCAAGCAC
TCCCTACTCCTCAGGGACAAGGAGCAGCTGCCGCTCCCTTC
CTACTCAGCTCTGAGGCCAGCCTGACTGGCCTGGAGGCC
GTGGAGACCATTTCTGGGTTCCAGGCCCTGGATGCCAGGGACT
CCCCGCAGGTTGCCCGCCTGCCAGCGCTACTGGCAAATGCCG

```

- continued

```

CCCCCTGTTCTGGAGCTGCTGGAAACCACGCGCAGTGCCCCCTAC
GGGGTGCCTCAAGACGCACTGCCGCTGCGAGCTGCGGTCA
CCAGCAGCCGGTGTCTGTGCCGGGAGAAGCCCCAGGGCTCTGTG
GCGGCCCCCGAGGAGGAGACACAGACCCCCGTGCGCTGGTGCAG
CTGCTCGGCGACAGCAGCCCTGGCAGGTGTACGGCTTCG
CGGGCCTGCCCTGCCGGCTGGTGCCTGGGAGGCTCTGGGGCTCC
AGGCACAACGAACGCCGCTTCCTCAGGAACACCAAGAAGTTCATC
TCCCTGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGACGTGG
AAGATGAGCGTGCAGGACTGCGCTGGCTGCGCAGGAGCCAGGG
GTTGGCTGTGTTCCGGCCGAGACGACCCGCTGCGTGAGGAGATC
CTGGCCAAGTTCCTGCACTGGCTGATGAGTGTACGTCGTGAG
CTGCTCAGGTCTTCTTTATGTACGGAGACCAAGTTCAAAAG
AACAGGCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAA
AGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCC
CTGCGGAAGCAGAGGTCAAGCAGCATCGGAAGCCAGGCCGCC
CTGCTGACGTCCAGACTCCGCTTCATCCCCAAGCCTGACGGGCTG
CGGCGATTGTGAACATGGACTACGTCGTGGAGGCCAGACGTT
CGCAGAGAAAAGAGGCCGAGCGTCTCACCTCGAGGGTAAGGCA
CTGTTCAGGTGTCAACTACGAGCGGGCGCGGGCCCCGGGCTC
CTGGGCGCTCTGTGCTGGCCTGGACGATATCCACAGGGCTGG
CGCACCTCGTGTGCGTGTGCGGGCCAGGACCCGCCCTGAG
CTGTAACCTGTCAAGGTGGATGTGACGGCGCGTACGACACCAC
CCCCAGGACAGGCTACGGAGGTCACTGCCAGCATCATCAAACCC
CAGAACACGTACTGCGTGCCTGGTATGCCGTGGCTCACCTG
GCCCATGGCACGTCCGAAGGCCCTCAAGAGCCACGTCTTAC
TTGACAGACCTCCAGCGTACATGCCAGCTGACAGTGGCTCACCTG
CAGGAGACCAAGCCGCTGAGGGATGCCGTGTCATCGAGCAGAGC
TCCCTCCCTGAATGAGGCCAGCAGTGGCCTCTCGACGTCTTCTA
CGCTTCATGTGCCACCACGCCGTGCGCATCAGGGCAAGTCC
GTCCAGTGCCAGGGATCCCGCAGGGCTCCATCCTCTCCACGCTG

```

- continued

CTCTGCAGCCTGTGCTACGGCGACATGGAGAACAGCTTTGCG
 GGGATTGGCGGGACGGCTGCTCCGTTGGTGGATGATTG
 TTGTTGGTGAACACCTCACCTCACCCACGCGAAAACCTTCCCTCAGG
 ACCCTGGTCCGAGGTGTCCTGAGATGGCTGGCTGGTGAACTTG
 CGGAAGACAGTGGTAACCTCCCTGAGAAGACGAGGCCCTGGT
 GGCACGGCTTGTTCAGATGCCGGCCACGGCTATTCCCCTGG
 TGCAGGCTGCTGGATAACCGGACCCCTGGAGGTGCAAGGCAC
 TACTCCAGCTATGCCGGACCTCCATCAGGCCAGTCTCACCTC
 AACCGCGGCTTAAGGCTGGAGGAACATGCGTCGCAAACCTTT
 GGGTCTTGGGCTGAAGTGTACAGCCTGTTCTGGATTGCA
 GTGAACAGCCTCCAGACGGTGTGCAACACATCTACAAGATCCTC
 CTGCTGCAGGCCAGGTTCAAGGTTCACGCATGTGCTGCAGCTCCA
 TTCATCAGCAAGTTGGAGAACCCCCACATTTCTGGCGCTC
 ATCTCTGACACGGCTCCCTCTGCTACTCCATCCTGAAAGCCAAG
 AACGCAGGGATGTCGCTGGGGCCAAGGGCGCCGCCGGCTCTG
 CCCTCCGAGGCCGTGCAGTGGCTGTGCCACCAAGCATTCTGCTC
 AAGCTGACTCGACACCCTGTCACCTACGTGCCACTCCTGGGTCA
 CTCAGGACAGCCAGACGAGCTGAGTCGGAAGCTCCCCGGACG
 ACGCTGACTGCCCTGGAGGCCGAGCCAACCCGGACTGCCCTCA
 GACTTCAAGACCATCCTGGACTGA

[0127] In some embodiments, a mouse TERT mRNA may comprise a wild type TERT sequence. In some embodiments, the wild type TERT sequence may comprise a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO 31:

ATGACCCGCGCTCCTCGTTGCCCCCGGGTGCCTCTGCTGCGC
 AGCCGATACCGGGAGGTGTGGCGCTGGCAACCTTGTGCGCGC
 CTGGGGCCGAGGGCAGGGCGCTTGTGCAACCCGGGACCGAAG
 ATCTACCGCCTTGGTGCACATGCCATGCTAGTGTGCACTGG
 GGCTCACAGCCTCCACCTGCCACCTTCTTCCACCAAGGTGTCA
 TCCCTGAAAGAGCTGGTGGCCAGGGTTGTGAGAGACTCTGCGAG
 CGCAACAGAGAGAACGTGCTGGCTTTGGCTTGAGCTGCTTAAC
 GAGGCCAGAGGCCCTCCATGGCCTCACTAGTAGCGTGCCT
 AGCTACTTGCCCAACACTGTTATTGAGACCCGTGCTGAGTGGT
 GCATGGATGCTACTGTTGAGCCAGTGGCGACGACCTGCTGGTC
 TACCTGCTGGCACACTGTGCTTTATCTCTGGTGGCCCCCAGC
 TGTGCCTACCAAGGTGTGGCTCCCTGACCAAATTGTGCC
 ACCACGGATATCTGGCCCTGTGTCGGTAGTTACAGGCCAC
 CGACCCGTGGCAGGAATTCACTAACCTTAGGTTCTAACACAG

- continued

ATCAAGAGCAGTAGTCGCCAGGAAGCACCAGAACCCCTGGCCTG
 CCATCTCGAGGTACAAGAGGCATCTGAGTCTCACAGTACAAGT
 GTGCCCTCAGCTAAGAAGGCCAGTGTATCCTGTCAGGAGAGT
 GAGGAGGGACCCACAGGCCAGTGCTACCAACCCATCAGGCAA
 TCATGGGTGCCAAGTCCTGCTCGTCCCCGAGGTGCTACTGCA
 GAGAAAGATTGTCTTCAAAGGAAAGGTGTCTGACCTGAGTCTC
 TCTGGTCGGTGTGCTGTAACACAAGGCCAGCTCCACATCTCTG
 CTGTCACCACCCGCCAAATGCCTTCAGCTCAGGCCATTATT
 GAGACCAGACATTCTTTACTCCAGGGAGATGGCCAAGAGCGT
 CTAAACCCCTCATCCTACTCAGCAACCTCCAGCCTAACTTGACT
 GGGGCCAGGAGACTGGTGGAGATCATCTTCTGGCTCAAGGCCT
 AGGACATCAGGACCACTCTGCAGGACACACCGCTATCGCGTCA
 TACTGGCAGATGCCGCCCCCTGTTCAACAGCTGCTGGTGAACCAT
 GCAGAGTGCCAATATGTCAGACTCCTCAGGTACATTGAGGTT
 CGAACAGCAAACCAACAGGTGACAGATGCCTGAACACCAGCC
 CCGCACCTCATGGATTGCTCCGCCGTGACAGCAGTCCCTGGCAG
 GTATATGGTTTCTCGGGCCGTCTCTGCAAGGTGGTGTCTG
 AGTCTCTGGGTACCAGGCACAAATGAGCGCCGCTCTTAAGAAC
 TTAAAGAAGTTCATCTCGTTGGGAAATACGGCAAGCTATCACTG
 CAGGAACGACTGTTGGAGATGAAAGTAGAGGATTGCCACTGGCT
 CGCAGCAGCCGGGAGGACCGTGTCCCCGCTGCAGAGCACCGT
 CTGAGGGAGAGGATCCTGGCTACGTTCTGCTGGCTGATGGAC
 ACATACGTGGTACAGCTGCTTAGGTCTTACATCACAGAG
 AGCACATTCCAGAAGAACAGGCTTCTTCTACCGTAAGAGTGTG
 TGGAGCAAGCTGCAAGCATTGGAGTCAGGCAACACCTTGAGAGA
 GTGCGGCTACGGAGCTGTCACAAGAGGAGGTAGGCATCACCAG
 GACACCTGGCTAGCCATGCCATCTGCAGACTGCGCTTCATCCCC
 AAGCCCAACGGCTGGCCATTGTGAACATGAGTTATAGCATG
 GGTACAGAGCTTGGCAGAAGGAAGCAGGCCAGCATTCACC
 CAGCGTCTCAAGACTCTCTCAGCATGCTCAACTATGAGCGGACA
 AAACATCCTCACCTTATGGGTCTTCTGACTGGGTATGAATGAC
 ATCTACAGGACCTGGCGGGCTTGTGCTGCGTGTGCGTCTG
 GACAGACACCCAGGATGTTAGTGTGAGGAGATGTGACCGGG
 GCCTATGATGCCATCCCCAGGGTAAGCTGGTGGAGGTTGTTG
 AATATGATCAGGCACTGGAGAGCACGTACTGTATCCGCCAGTAT
 GCAGTGGTCCGGAGAGATGCCAAGCCAAGTCCACAAGTCC
 AGGAGACAGGTACCCACCTCTGACCTCCAGGCCATACATGGG
 CAGTTCTTAAGCAGTGCAGGATTCAAGATGCCAGTGCAG
 AACCTCGTTGTCATCGAGCAGAGCATCTATGAATGAGAGCAGC

- continued

```

AGCAGCCTGTTGACTTCTCCCTGCACCTCCGTGCGTCACAGTGTC
GTAAAGATTGGTGACAGGTGCTATACGCAGTGCCAGGGCATCCCC
CAGGGCTCCAGCCTATCCACCCCTGCTCTGCAGTCTGTGTTCGGA
GACATGGAGAACAGCTGTTGCTGAGGTGCAGCGGGATGGGTTG
CTTTACGTTTGTGATGACTTCTGTTGGTGACGCCCTCACTTG
GACCAAGCAAAACCTCCTCAGCACCCCTGGTCCATGGCGTCCCT
GAGTATGGGTGATGATAAACTTGAGAAGACAGTGGTGAACATTC
CCTGTGGAGCCTGGTACCCCTGGGTGGTGAGCTCCATACCAGCTG
CTGCTCACTGCCTGTTCCCTGGTGTGGCTTGCTGCTGGACACT
CAGACTTGGAGGTGTTCTGTGACTACTCAGGTTATGCCAGACC
TCAATTAAGACGAGCCTCACCTTCAGAGTGTCTCAAAGCTGGG
AAGACCATGCGGAACAAGCTCCTGTCGGCTTGCGGTTGAAGTGT
CACGGTCTATTCTAGACTTGAGGTGAACAGCCTCCAGACAGTC
TGCATCAATATACAAGATCTCCTGCTTCAGGCCTACAGGTTC
CATGCGATGTTGATTAGCTCAGGTTCCCTTGACCGCGTGTAGGA
AACCTCACATTCTTCTGGCATCATCTCAGCCAAGCATCCTGC
TGCTATGCTATCCTGAAGGTCAAGAATCCAGGAATGACACTAAAG
GCCTCTGGCTCTTCCTCTGAAGCCGACATTGGCTTGCTAC
CAGGCCCTCCTGCTCAAGCTGGCTGCTCATCTGTCACTACAAA
TGTCTCTGGGACCTCTGAGGACAGCCCCAAAATGCTTGCGG
AAGCTCCAGGGCGACAATGACCATCCTAAAGCTGCGACTGAC
CCAGCCCTAACGACAGACTTCAGACCATTGGACTAA

```

[0128] In some embodiments, a TERT mRNA may comprise a nucleic acid sequence at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% identical to any one of SEQ ID NOS: 1-5, 7, 9 or 30.

[0129] In some embodiments, a TERT mRNA may encode a modified TERT protein containing one or more amino acid substitutions, deletions, and/or insertions as compared to SEQ ID NOS: 6 or 8, while retaining substantial TERT activity. In some embodiments, a TERT mRNA may encode an amino acid sequence at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% identical to SEQ ID NO: 6 or SEQ ID NO: 8.

[0130] In other embodiments, a TERT mRNA may encode an amino acid sequence with a mutation of L55Q, P65A,

V70M, A202T, A279T, V299M, H412Y, a deletion of residue 441, R522K, K570N, R631Q, G682D, V694M, Y697F, P704S, Y707F, A716T, P721R, T726M, Y772C, P785L, V791I, R811C, L841F, R865H, V867M, R901W, K902N, P923L, S948R, R979W, V1025F, A1062T, V1090M, T1110M, and/or F1127L relative to the amino acid sequences of SEQ ID NO: 6. In some embodiments, the TERT mRNA may encode a TERT isoform in which the translated protein lacks amino acid residues 711-722, 764-807, 808-1132, or 885-947 relative to the amino acid sequences of SEQ ID NO: 6. In some embodiments about 1, about 5, about 10, about 20, or about 100 amino acids preceding or following the domain are also deleted from the amino acid sequence of SEQ ID NO: 6.

[0131] In some embodiments, the TERT mRNA may encode an amino acid sequence in which one or more of the protein regions are deleted or repeated relative to the amino acid sequences of SEQ ID NO: 6: residues 1-230 corresponding to the RNA-interacting domain 1, residues 58-197 corresponding to a “GQ” residue motif, residues 137-141 associated with the specificity of telomeric DNA and primer elongation, residues 210-320 corresponding to a disordered region, residues 231-324 associated with a linker sequence, residues 301-538 associated with oligomerization, residues 325-550 or 460-594 corresponding to an RNA-interacting domain, residues 376-521 corresponding to a “QFP” residue motif, residues 397-417 corresponding to a “CP” residue motif, residues 825-884 corresponding to a DNA repeat template, residues 618-729 corresponding to a reverse transcriptase like element, residues 914-928 associated with oligomerization, residues 930-934 associated with a primer grip sequence, and/or residues 936-1132 corresponding to the C-terminus. In some embodiments about 1, about 5, about 10, about 20, or about 100 amino acids preceding or following the domain are also deleted or repeated.

[0132] In some embodiments, a TERT mRNA may comprise or consist of a nucleotide sequence at least at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or more identical to any of disclosed nucleic acid sequences, or to any subsequence of a disclosed nucleic acid sequence, e.g., any 100 base pair (bp), 200 bp, 300 bp, 400 bp, 500 bp, or more of a disclosed nucleic acid sequence. In some embodiments, a TERT mRNA may encode an amino acid sequence at least at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or more identical to any of one of the disclosed polypeptide sequences, or to any subsequence of a disclosed polypeptide sequence, e.g., any 50 amino acid (aa), 100 aa, 200 aa, 300 aa, 400 aa, 500 aa, or more of a disclosed polypeptide sequence.

[0133] Non-limiting TERT sequences of the disclosure, include TERT nucleic acid and amino acid sequences listed in Table 1A.

TABLE 1A

TERT Species	Amino Acid SEQ ID NO: Sequence	Example Nucleic Acid SEQ ID NO: Sequence	Example Nucleic Acid SEQ ID NO: Sequence
Cat	ASO67359.1	KX620456.1	
Dog	NP_001026800.1	NM_001031630.1	
Mouse	AAI27069.1	BC127068.1	

TABLE 1A-continued

TERT Species	Amino Acid SEQ ID NO:	Amino Acid Sequence	Example Nucleic Acid SEQ ID NO:	Example Nucleic Acid Sequence
Mouse, isoform 1	10	NP_033380.1	14	NM_009354.2
Mouse, isoform 2	11	NP_001349316.1	15	NM_001362387.1
Mouse, isoform 3	12	NP_001349317.1	16	NM_001362388.1
Mouse		EDL37055.1		Machine reverse translation of EDL37055.1
Cow		NP_001039707.1		NM_001046242.1
Sheep, isoform 1		XP_027835754.1		XM_027979953.1
Sheep, isoform 2		XP_027835755.1		XM_027979954.1
Pig		NP_001231229.1		NM_001244300.1
African Elephant		XP_023401395.1		XM_023545627.1
Chicken		NP_001026178.1		NM_001031007.1
Rat	13	NP_445875.1	17	NM_053423.1
Zebrafish		NP_001077335.1		NM_001083866.1
Japanese medaka		NP_001098286.1		NM_001104816.1
Horse, isoform 1		XP_023481649.1		XM_023625881.1
Horse, isoform 2		XP_023481650.1		XM_023625882.1
Horse, isoform 3		XP_023481651.1		XM_023625883.1

[0134] In some embodiments of the compositions and methods of the disclosure, an amino acid sequence of TERT may comprise or consist of a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99%, 100% or any percentage in between of identity to one or more of SEQ ID Nos: 6-8 or 10-13. In some embodiments of the compositions and methods of the disclosure, an amino acid sequence of a portion of TERT, functional or non-functional, may comprise or consist of a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99%, 100% or any percentage in between of identity to one or more of SEQ ID Nos: 6-8 or 10-13.

[0135] In some embodiments of the compositions and methods of the disclosure, a nucleic acid sequence of TERT may comprise or consist of a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99%, 100% or any percentage in between of identity to one or more of SEQ ID Nos: 1-5, 7, 9, 14-17, 30 or 31. In some embodiments of the compositions and methods of the disclosure, a nucleic acid sequence of a portion of non-human TERT, functional or non-functional, may comprise or consist of a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99%, 100% or any percentage in between of identity to one or more of SEQ ID Nos: 1-5, 7, 9, 14-17, 30 or 31.

The amino acid sequence of non-human primate TERT isoform 1 may comprise or consist of the sequence of SEQ ID NO: 18 (also described at GenBank Accession No. XP_016808391.2):

```

1 mpraprrcrav rslllrshyre vlplatfvrr lgpqgwrlvq rgdpaafral vaqlcvcpw
61 darpppaaps frqvsclkel varvlqrlnce rgaknvlafg falldgargg ppeafttsvr
121 sylpntvtada lrgsgawgll lrrvgddvly hllarcalfv lvapseyqv cgpplyqlga
181 atqarpppha sgprrrlgce rawnhsvrea gvplglpapg arrrggsasr slplpkrr
241 gaapepertp vgggswahpg rtrgpsdrgf cvvsparpae eatslegals gtrhshpsvg
301 rqhhagppst srpprpwdtp cppvytaetkh flyssgdkeq lrpsflssl rpsltgarrr
361 vetiflgsrp wmpgtprrlp rlpqrywqmr plflellgnh aqcpygvllk thcplraavt
421 paagvcarek pgqsvaapee edtdprrlvq llrqhsspwmq vygfvraclr rlvpplwg
481 rhnerrflrn tkkfislgkh aklslqeltw kmsvrdcawl rrspvgvgsvp aaehrlreei
541 lakflhwlmvs vyvvellrsf fyvtettfqk nrlffyrksv wsklqsigir qhlkrvqlre
601 lseaevrqhq earpalltsr lrfipkpfdgl rpivnmddyv gartfrrekr aerltsrvka

```

-continued

661 lfsvlnyera rrpgilgasv lglddihrav rtfvlrvraq dpppelyfvk vdvtgaydti
 721 pqdrltevia siikpqntyc vrryavvqka ahghvrkafk shvstltdlq pymrqlfvahl
 781 qetsplrdav iieqssslne assglfdvfl rfvcvhavri rgksyvqcqg ipqgsilstl
 841 lcslcyygdme nklfagirrd gllrlvddf llvtpchltha kafirltlvrg vpeygcvnvl
 901 rktvvnfpve dealgtafv qlpahglfpw cglldtrtl evqsdysya rtsirasltf
 961 nrgfkagrnm rrklfqvrl kchslfldlq vnslqtvctn iykillqay rfhacvlqlp
 1021 fhqqvwknpt fflriisdtal slcysilkak nagmslgakg aaglpseam qwlchqafll
 1081 kltrhrvtyv pllgslrtaq tqlsrklpgt tsaleaaan palpsdfkti ld.

The nucleic acid sequence of non-human primate TERT isoform 1 may comprise or consist of the sequence of SEQ ID NO: 19 (also described at GenBank Accession No. XM_016952902.2):

1 ctcgcggcgc gagtttcagg cagcgctgcg tcctgctgcg cacgtggaa gccctggccc
 61 cggccacccc cgcgatgccg cgcgctcccc gctgccgagc cgtgcgcgtcc ctgctgcgcga
 121 gccactaccg cgaggtgctg ccgctggcca cgttcgtgcg ggcgcctgggg ccccaggggct
 181 ggccggcttgt gcagcgcggg gacccggcgg ctttccgcgc gctgggtggcc cagtgcctgg
 241 tgtgcgtgcc ctgggacgcga cggccggccc cgcgcggccc ctccctccgc caggtgtcct
 301 gcctgaagga gctggggcc cgagtgcgtc agaggctgtg cgagcgcggc gccaagaacg
 361 tgctggcctt cggcttcgcg ctgctggacg gggcccgccg gggccccccc gaggcattca
 421 ccaccagcgt ggcgcagctac ctgcaccaaca cggtgaccga cgcactgcgg gggagcgggg
 481 cgtggggcgt gctgcgtgcg cgcgtggcgc acgacgtgt gggtcacctg ctggcacgc
 541 ggcgcgttct tgtgcgtgt gtcggcgt ggccttacca ggtgtgcggg cgcgcgtgt
 601 accagctcg cgcgtgcact caggccggc cccgcacaca cgctagtggc cccgcaggc
 661 gtcttagatg cgaacgggcc ttggaaaccata ggcgtcaggga ggcggggc cccctggccc
 721 tgccagcccc gggtgccagg aggccgggg gcaactgcgtc cggaaatgtc cccgttgc
 781 agaggcccaag ggcgtggcgct gcccctgagc cggagcggac gcccgtggg caggggtcct
 841 gggccacccc gggcaggacg cgtggaccga gtgaccgtgg tttctgtgt gtgtcacctg
 901 ccagaccgc cgaagaagcc acctttttgg agggctgcgtc ctctggcactc cgcactcccc
 961 acccatccgt gggccgcacg caccacgcgg gccccccatc cacatgcgg ccaccacgtc
 1021 cctggacac gccttgccttcc cccgtgtacgc cccgacccaa gcaacttcctc tacttcgt
 1081 ggcacaagga gcaactgcgg cccttgccttcc tactcagtc totgaggccc agcctgactg
 1141 ggcgtggcgt gctcgtggag accatcttc tgggttccatc ggcctggatc ccaggactc
 1201 cccgcagggtt gccccgcgtt ccccgccgt actggcaat gggccctgt tttctggagc
 1261 tgcttggaa ccacgcgcag tgcccttacg ggggtgttccatc caagaegcact tgccgcgtgc
 1321 gagctgcggcgtt caccacgcac gcccgtgttccatc gtgcccggga gaagccccag ggcttgc
 1381 cggcccccgcg gggaggaggac acacccccc gtcgcgtgttccatc gcaactgcgtc cggccgcaca
 1441 gcaactgcgttccatc gcaactgcgtc gggccgtgttccatc ggcgcggcgt gtcggcccg
 1501 ggccttggggccatc ctcgcggcacttccatc aacgcacccccc gtcgcgtgttccatc
 1561 ccctggggaa gcaactgcgtc gcaactgcgtc gggccgtgttccatc ggcgcggcgt
 1621 actgcgtgttccatc gcaactgcgtc gggccgtgttccatc ggcgcggcgt
 1681 tgcttggaa gatcctggcc aacttgcgttccatc gcaactgcgtc gggccgtgttccatc
 1741 tgcttggaa gatcctggcc aacttgcgttccatc gcaactgcgtc gggccgtgttccatc

-continued

1801 accggaaagatgtctggaggc aagttgcacaa gcaattggaaat cagacacgacat ttgaagagg
1861 tgcagctgcg ggagctgtcg gaagcagagg tcaggcagca tcaggaagcc agggccccc
1921 tgctgacgtc cagactccgc ttcatcccc agcctgacgg gctcgccggc atttgtgaaca
1981 tggactacgt cgtgggagcc agaacgttcc gcaagaaaaa gagggcccgag cgtctcaect
2041 cgagggtgaa ggcactgttc aegctgtca actacgagcg ggcggggcgc cccggctcc
2101 tggggccctc tggctgggc ctggacgata tccacaggc ctggcgcacc ttctgtctgc
2161 gtgtgcgggc ccaggaccccg ccgcctgagc tgtaactttgt caaggtggat gtgacggggc
2221 cgtacgacac catccccca gacaggtctca cggaggtcat cggccagcatc atcaaacc
2281 agaacacgta ctgcgtgcgt cggtatgtcg tggccagaa ggccgcggcat gggcacgtcc
2341 gcaaggccctt caagagccac gtctctaccc tgcacagacct ccagccgtac atgcgcagat
2401 tcgtggctca cctgcaggag accagccac tgagggtatgc cgtcatcatc gacgcagat
2461 cctccctgaa tgaggccage agtggccctt tgcacgttcc cctacgcttc ttgtgcggcc
2521 acgcgtgcg catcaggggc aagtcttacg tccagtgcca gggatcccg cagggtccca
2581 tcctgtccac gctgtctgc aegctgtct acggcgacat ggagaacaag ctgttttgcgg
2641 ggattcggcg ggacgggctg ctctgcgtt tggggatga tttcttgcgg tgcacaccc
2701 acctcaccca cgcgaaagcc ttcttcagga cctctggccg aggtgtccct gatgtatgg
2761 gcgtggtaa cttgcggaaag acagtagtga achtccctgt agaagatgag gcccgggt
2821 gcacggcttt tggtaagctg cggcccccacg goctattccc ctggtgccgc ctgtgtctgg
2881 acacccggac cctggagggtg cagagcgact achtccagcta tgccggacc tccatca
2941 ccagtcacat cttaaccgc ggcttcaagg ctggggaggaa catcgctcgc aaactcttt
3001 gggcttgcg gctgaagtgt cacagctgt ttctggattt gcaggtgaac agcctccaga
3061 cggtgtgcac caacatctac aagatctcc tgcgtcgaggc gtacagggtt cacgcatgt
3121 tgctgcagct cccatttcat cagcaagttt ggaagaaccc cacattttc ctgcgcatca
3181 tctctgacac ggctccctc tgtaacttca tectgaaagc caagaacgca gggatgtcg
3241 tggggccaa gggtgcggcc gggccctctgc cctccggcgc catcgagtgg ctgtgcggcc
3301 aagcattccct gctcaagctg actcgacacc ggtcaccta cgtgccacte ctggggtcac
3361 tcaggacage ccagacgcag ctgagtcgga agctccggg gacgacgctg agtgccttgc
3421 aggccgcagc caacccggca ctgcctctacg acttcaagac catcctggac tggatggcc
3481 ccgccccacag ccggggccgag agcagacacc agcagccctg tcaacggggg ctctacgtcc
3541 cagggaggaa gggggccccc acaccccgac cgcacccgt gggagtctga ggcctgagtg
3601 agtgtctggc caaggccctgc atgtccggct gaaggctgag tgcgtcgactg aggccctgag
3661 gagtgtccag ccaagggtcg agtgtccagc acacctgccc tcttcacttc cccacaggct
3721 ggcgtcgcc tccacccag ggcagcttt tccctgcggc gagcccggt tccctactcc
3781 cacatggaa tagtccatcc ccagattcgc cattgtccac ccctcgccct gcccctcc
3841 gcctccacg cccaccatcc agatggagac cctgagaagg accctgggag ctctggaaat
3901 ttggagtgc caaagggtgtc ccctgtacac aggtgaggac cctgcacactg gatgggggt
3961 cctgtgggtc aaattggggg ggggggtgtg tggagtaaa atactgaata tatgagttt
4021 tcagttttaaaa.

-continued

The amino acid sequence of non-human primate TERT isoform 2 may comprise or consist of the sequence of SEQ ID NO: 20, GenBank Accession No. PNI27662.1:
 1 mpraprrcav rslllrshyre vplatfvrr lgpqgwrlvq rgdpaafraal vaqclvcvpw

```

 61 darpppaaps frqvsclkel varvlqrice rgaknvlafg falldgargg ppeaf ttsvr
121 sylpntvtda lrgsgawgll lrrvgddvlv hllarcalfv lvapscayqv cgpplqlga
181 atqarpppha sgprrrlgce rawnhsvrea gvplglpapg arrrggsasr slplpkprrr
241 gaapepertp vgqgswahpg rtrgpsdrgf cvvsparpae eatslegals gtrhshpsvg
301 rqhhagppst srpprpdtp cppvyaetkh flyssgdkeq lrpsfllssl rpsltgarl
361 vetiflgsrp wmpgtprrlp rlpqrywqmr plflellgnh aqcpqygvllk thcplraavt
421 paagvcarek pqgsavaapee edtdprrlvq llrqhsspwlq vygfvraclr rlvppglwgs
481 rhnerrflrn tkkfislgkh aklslqeltw kmsvrdcawl rrspgvgsvp aaehrlreei
541 lakflhwlm vvvvelrsf fyvtettfiqk nrlffyriksv wsklqsigir qhlkrvqlre
601 lseaevrqhq earpalltsr lrfipkpdgl rpivnmdyvv gartfrrekr aerltsrvka
661 lfsvlnyera rrppllgasv lglddihrw rtfvrlvraq dpppelyfvk vdvtgaydti
721 pqdrltevia siikpqntyc vrriavvqka ahghvrkafk shvstltdlq pymrqfvahl
781 qetsplrdav iieqssslne assglfdvfl rfvcvhavri rgksyvqcqg ipqgsilstl
841 lcslcygdm e nklfagirrd glllrlvddf lltvphltha kaflrtlvrg vpeygcvvnl
901 rktvvnfve dealggtafv qlpahglfpw cgllldtrtl evqsdysya rtsirasltf
961 nrgfkagrn m rrklnfgvrlr kchslfllq vnsllqtvctn iykillqay rfhacvlqlp
1021 fhqqvwknpt fflriisdt a slcysilkak nagmislakg aagplpseam qwlchqafll
1081 kltrhrvtyv pllgslrtaq tqlsrklpgt tlaaleaaa palpsdfkti ld.
```

The nucleic acid sequence of non-human primate TERT isoform 2 may comprise or consist of the sequence of SEQ ID NO: 21 (reverse machine translation of GenBank Accession No. PNI27662.1):

```

 1 atgccgcgcg cggccgcgtc cgcgcgcgtg cgcagccgtc tgccgcggca ttatcgccaa
 61 gtgctgcccgc tggcgacatt tggcgccgc ctggggccgc agggctggcg cctgggtcag
121 cgccggcgtc cggccggcggt tcgcgcgtc gtggcgca gctgggtgtg cgtgcgtgg
181 gatgcgcgcg cggccgcggc ggcgcgcgagc ttccgcagg tgagctgcct gaaagaactg
241 gtggcgccgc tgctgcagcg cctgtgcgaa cgcggcgccgaa aaaaactgtc ggccgtttggc
301 ttgcgcgtgc tggatggcgc gcggggcgcc cccgcggaaag cgtttaccac cagcgtgcgc
361 agctatgtc cgaacaccgt gaccgatgcg ctgcgcggca gggcgccgtg gggcctgtcg
421 ctgcgcgcgc tggcgatgtc tggatggcgt catctgtgg cgcgcgtgc gctgtttgtg
481 ctggcgccgc cggatgcgc gtatcagggt tgcggccgc cgcgttatca gctggcgccg
541 gggccggccggc gggccggccggc gggccggccggc gggccggccggc gggccggccggc
601 cggccgtggc accatagcgt gcggcgccggc ggccggccggc tggccgtggc gggccggccggc
661 gggccggccggc gggccggccggc cggccggccggc gggccggccggc gggccggccggc
721 gggccggccggc gggccggccggc gggccggccggc gggccggccggc gggccggccggc
781 cggccggccggc gggccggccggc gggccggccggc gggccggccggc gggccggccggc
841 gggccggccggc gggccggccggc gggccggccggc gggccggccggc gggccggccggc
901 cggccggccggc gggccggccggc gggccggccggc gggccggccggc gggccggccggc
961 tggccggccggc tggatgcgcga aaccaaacat tttctgtata gcaagccggccga taaagaacag
```

-continued

1021 ctgcgccgaa gctttctgtc gagcagccctg cggccgagcc tgaccggcgcc gcgcgcgc
1081 gtggaaacca tttttctggg cagccggccg tggatgcccgc gcaccccgcc ccgcctgccc
1141 cgccctgcccgc agcgcttatttgc gcagatgcgc cccgtgtttc tggactgtctt gggcaaccat
1201 gcccaggccgc cgtatggcgt gctgtgtaaa acccattggcc cgctgcgcgc ggcgggtgacc
1261 cccggccggg gcggtgtgcgc gcggaaaaaa cccgaggccgc gcggtggccgc gccggaaagaa
1321 gaagataccg atccggccgc cctggtgcag ctgtgcgcgc agcatagcag cccgtggcag
1381 gtgtatggct ttgtgcgcgc gtgcctgcgc cccctggcgc cgcggggccctt gtggggcagc
1441 cgccataaacg aacggccgtt tctgcgcac accaaaaaat ttattagccctt gggcaaacat
1501 gcgaaactga gcctgcagga actgcacctgg aaaatgagcg tgcgcgatttgc cgcgtggcgt
1561 cgccgcagcc cgggcgtggg cagcgtgcgc gcggcggaaatcgcctgcgc cgaagaaatt
1621 ctggcgaat ttctgcatttgc gctgtatgtgg tggactgtctt gcgcaagctt
1681 ttttatgtga cccggaaaccac ctttcagaaa aaccgcctgt tttttatcg caaaagcgtg
1741 tggagcaaaccatgc ctcgcatttcgc cagcatctga aacgcgtgcgc gctgcgcgaa
1801 ctgagcgaag cggaaatgcgc ccagcatcag gaagcgcgcgc cggcgctgtt gaccagccgc
1861 ctgcgcctta ttccggaaacc cggatggcctgc cccgcatttgc tgaacatggaa ttatgtggtg
1921 ggccgcgcgc ccttcgcgc cggaaaaacgc gcggaaacgcgc tgaccagccgc cgtgaaagcg
1981 ctgttttagcg tgctgaacta tgaacgcgcgc cggccggccgc gcctgtggg cgcgcgcgt
2041 ctgggcctgg atgatattca tcgcgcgtgg cgcaccccttgc tgctgcgcgtt gcgccgcgc
2101 gatccgcgcgc cggaaactgttgc ttttgtaaa gtggatgtga cccggcgcgttgc tgataccatt
2161 cccgcaggatc gcctgaccga agtgcattgcgc agcattattaa aaccgcagaa cacctattgc
2221 gtgcgcgcgttgc atgcgggtgttgc cccggaaaccgcgc atgtgcgcgcgc agcgtttaaa
2281 agccatgtga gcaccctgcgc cgcgcgtgc cccgcattatgc gcacgtttgtt ggcgcgcgc
2341 caggaaacca gcccgcgcgc cgcgcgtgcgc attattgcgc agagcgcgcgc cccgcgcgc
2401 gcgcgcgcgc gcctgtttgc tggtttctgc cccgcattatgc gcgcgcgc
2461 cccggccaaa gctatgtgcgc tgccaggccgc attccgcagg gcacgcatttgc gaggccctgc
2521 ctgtgcgcgc tttgtgtatgg cccatggaa aacaaactgttgc tgccggccatgc tccgcgcgc
2581 ggccctgtgc tgccgcgtgg ggtgtttttgc tgccgcgtgg cccgcattatgc gaccatgcgc
2641 aaagcggttgc tccgcgcgcgc cccgcggccgc gtcgcggaaatgcgcgtt ggtgcgcgc
2701 cggaaaaaccgc tggtaacttgc tccggccgc gatgcgcgcgc tggccggccac cccgcgttgc
2761 cccgcgcgc cccatggccgc gttccgtgg tgccgcgcgc tgctggataccgc cccgcgcgc
2821 gaagtgcaga cccatggccgc cccatggccgc tccgcgcgcgc cccgcgcgc
2881 aaccgcgcgc taaaacgcggccgc cccgcgcgcgc cccgcgcgcgc tgccgcgcgc
2941 aaatgcgcgc tccgcgcgcgc cccgcggccgc gtcgcggccgc gtcgcgcgc
3001 atttataaaa ttctgcgttgc cccatggccgc cccatggccgc tgccgcgcgc
3061 tttcatcaggc aggtgtggaa aaacccgcacc tttttctgc gcattattatgc cccatggccgc
3121 agccatgtgttgc atgcatttgc cccatggccgc cccatggccgc tgccgcgcgc
3181 cccgcggccgc cccatggccgc cccatggccgc cccatggccgc tgccgcgcgc
3241 aaactgcaccgc cccatggccgc cccatggccgc cccatggccgc cccatggccgc

-continued

3301 acccagctga gcccggaaact gccggggcacc accctgagcg cgcttggaaagc ggcggggcaac

3361 ccggcgctgc cgagcgattt taaaaccatt ctggat,

The amino acid sequence of non-human primate TERT isoform 3 may comprise or consist of the sequence of SEQ ID NO: 22 (also described at GenBank Accession No. PNI27663.1).

1 mpraprcrav rsllrrshyre vlplatfvrr lgpqgwrllvq rgdpaafraal vaqclvcvpm
61 darpppaaps frqvsclkel varvlqlrlce rgaknvlafg falldgargg ppeafftsvr
21 sylplntvtda lrgsgawgll lrrvgddvlv hllarcalfv lvapscayqv cgpplyqlga
81 atqarpppha sgprrrlgce rawnhsvrea gvplglpapg arrrggsasr slplpkrrpr
41 gaapepertp vggswahpg rtrgpsdrgf cvvsparpae eatslegals gtrhshpsvg
01 rqhhagppst srpprpwdtp cppvyaetkh flyssgdkeq lrpsflssl rpstlgarrl
61 vetiflgsrp wmpgtprrlp rlpqrywqmr plflellgnh aqcpygvlkk thcplraavt
21 paagvcarek pqgsvaapee edtdprrlvq llrqhsspwy vygfvraclr r1vppglwgs
81 rhnerrflrn tkkfislgkh aklslqeltw kmsvrdcawl rrspvgvsvp aaehrlreei
41 lakflhwlmvs vyvvellrsf fyvtettfqk nrlffyrksv wsklqsigir qhlkrvqlre
01 lseaevrqhq earpalltsr lrfipkpdgl rpivnmnyvv gartfrrekr aerltsrvka
61 lfsvlnyera rrpqllgasv lglddihrav rtfvrlvraq dpppelyfvk vdvtgaydi
21 pqdrltevia siikpqntyc vrriyavvqka ahghvrkafk shvlpvpgd paglhvpaha
81 lqpvlrrhge qavcgdsagr aapafogg.

The nucleic acid sequence of non-human primate TERT isoform 3 may comprise or consist of the sequence of SEQ ID NO: 23 (reverse machine translation of GenBank Accession No. PNI27663.1):

1 atgcgcgcgc cgccgcgtc ccgagcgggtc cgcagectgc tgcgacggca ttatcgcgaa
61 gtgcgtccgc tggcacctt tgcgcgcgc ctggggccgc agggctggcg cctgggtcgac
121 cgcggcgatc cggggcggtt tcgcgcgtc gtggcgact gcctgggtgtc cgtggcggtgg
181 gatgcgcgcc cgccgcggc ggcccggcggc ttgcgcagg tgagctgcct gaaaactgt
241 gtggcgcgcg tgctgcageg cctgtggaa cgcggcgca aaaactgtc ggcgtttggc
301 ttgcgcgtc tggatggcgc gcgcggcgcc cgccggaaag cgtttaccac cagcgtcgcc
361 agctatctgc cgaacaccgt gaccgtatgc ctgcgcggca gcggcgcggt gggcctgtcg
421 ctgcgcggcg tggcgatga tgcgtgggtg catctgttgc cgcgtgcgc gtcgtttgtg
481 ctggggcgcc cgagctgcgc gtatcagggtg tgccggccgc cgctgtatca gctggggcgcc
541 ggcacccagg cgccgcggcc gcccgcattgcg agccggccgcg ggcggccgcet gggctgcgaa
601 cgcgcgtggaa accatagcgt gcgcgaagcg ggctgcgcgc tgggcgtgcg ggcgcggggc
661 ggcgcggccgccc gggcgccgcg cgcgagccgc agcctgcgcgc tgccgaaacg cccgcggccgc
721 ggcgcggccgc cggAACCGGA acgcaccccg tggggccagg gcagctggc gcatccggggc
781 cgcacccgcg gcccgcgcg tgcgggttt tgcgtgtga gccccggcgcccggggaa
841 gaagcgacca gcctggagg cgcgcgtgac ggcacccgcg atagccatcc gagcgtgggc
901 cgcacccatc atgcggggccc gcccgcgcacc agccggccgc cgcgcgggtg ggataccccc
961 tgccgcggg tgcgtgcgg aaccaaactt ttctgtata gcagcggcgta taaagaacacg
1021 ctgcgcggcga gcttctgtc gagcagcgtc cgcggcgcc tgacggcgcc ggcggccgtc
1081 gtggaaacca tttttctggg cggccggcccg tggatgcggc gcaaccccgcg cggcgtgcgg
1141 cgcctgcgcg acgcgtattt gcagatgcgc cgcgtgttc tggaaactgtc gggcaacat

-continued

```

1201 ggcagtgcc cgtatggcgt gctgctgaaa acccattgcc cgctgegcgc ggcgggtgacc
1261 ccggcgccgg gcgtgtgcgc ggcggaaaaa ccgcaggcga gcgtggccgc gcccggaa
1321 gaagataaccg atcccgcccg cctgggtgcag ctgtctgcgc agcatagcag cccgtggcag
1381 gtgtatggct ttgtgcgcgc gtgcctgcgc cgccctggc gcgcggccct gtggggcage
1441 cgccataacg aacgcgcgtt tctgogcaac accaaaaaat ttattagcct gggcaaacat
1501 ggcggaaactga gcctgeagga actgacacctgg aaaatgagc tgccgcatttgc cgccgtggctg
1561 cgccgcagcc cggggcggtggg cagcggtgcgc gcccggaaac atcgccctgcgc cgaagaaatt
1621 ctggcgaataat ttctgcatttgc gctgtatggc gtgtatgtgg tggaaactgtc ggcgcagctt
1681 ttttatgtga ccgaaaccac ctttcagaaa aaccgcctgt ttttttatcg caaaagcgtg
1741 tggagcaaac tgcagagcat tggcattcgc cagcatctga aacgcgtgca gctgcgcgaa
1801 ctgagcgaag cggaaagtgcgc ccagcatcag gaagcgcgcgc cggcgctgtc gaccagccgc
1861 ctgcgcctta ttccgaaacc ggtatggcctgc cgcccgatttgc tgaacatggta ttatgtggc
1921 ggccgcgcga ccttcgcgc cggaaacgcg cggaaacgcgc tgaccagccg cgtaaagcg
1981 ctgttttagcgt tgctgaacta tgaacgcgcgc cggccggccgc ggcgtgtgg cgcgcgcgtg
2041 ctgggcctgg atgatattca tgcgcgcgtgg cgcacccatttgc tgctgcgcgt ggcgcgcag
2101 gatccgcgcgc cggaaactgttca ttttgaaa gttggatgtga cggcgcgcgt tgataccatt
2161 cccgaggatc gcctgaccga agtattttgcg agcattatta aaccgcagaa cacctattgc
2221 gtgcgcgcgt atgcgggttgttgc agcggaaatgcg ggcgcgcgcgc atgtgcgcgcgcgc
2281 agccatgtgc tgcgcgcgcgttgc gccggggcgttgc cggcgcgcgcgcgcgcgcgcgc
2341 ctgcagccggc tgctgcgcgc ccatggcga caggcgggtgt gggcgcatacg cggcggccgc
2401 gggcgcgcgc cgtttggcggc c.

```

The amino acid sequence of non-human primate TERT isoform 4 may comprise or consist of the sequence of SEQ ID NO: 24 (also described at GenBank Accession No. PNI27664.1):

```

1 mpraprcrav rsllrlshyre vplplatfvrr lgpqgwrlvq rgdpaafral vaqclvcvpw
61 darpppaaps frqvsc1kel varvlqr1ce rgaknvlafg falldgargg ppeaf1tsvr
121 sylpn1tvtda lrgsgawg1l lrrvgddv1v hllarcalfv Ivaps1cayqv cgppl1yqlga
181 atqarpppha sgprrrlgce rawnhsvrea gvplglpapg arrrggsasr slplpkprrr
241 gaapepertp vggswahpg rtrgpsdrgf cvvsparpae eatslegals gtrhshpsvg
301 rqhhagppst srpprwdtp cppvyaetkh flyssgdkeq lrpsfl1ssl rpsltgarrl
361 vetiflgsrp wmpgtprrlp rlpqrywqmr plflellgnh aqcpygvllk thcp1raavt
421 paagvcarek pqgsavaapee edtdprrrlvq llrqhssp1wq vygfvrac1r r1vppglwg1
481 rhnerrflrn tkkfislgkh ak1sl1qel1w k1msvrdcawl rrspgvgs1vp aae1hrlreei
541 lakflhw1ms vyyvvel1rsf fyyt1tfqk nrlffyrks1w sk1qlqsig1r qhlkrvqlre
601 lseaevrqhq earpalltsr lrfipkpdgl rpi1vnmdyvv gartfrrekr aerltsrvka
661 lfs1vlnyera rrp1glgasv lg1ddihraw rtfv1rvraq dpppelyfvk vdvtgaydti
721 pqdr1tevia siikpqntyc vrryavvqka ahghvrkafk shvst1td1q pymrqfvahl
781 qetsplrdav iieqss1ne assglfdvfl rfvc1rhavri rgksyvqcqg ipqgs1st1
841 lcs1c1ygdme nk1fagirrd g111rlvddf llvtp1hltha kaflsyarts iras1tfnrg
901 fkagrnmrrk lfgv1rlkch slf1ld1qvns lqtvctniyk ill1qayrfh acv1qlpfh1

```

-continued

961 qvwknptffl riisdtaslc ysilkaknag mslgakgaag plpseamqwl chqafllklt

1021 rhrvtyvp11 gslrtaqtql srklpgttls aleaaanpal psdfktild.

The nucleic acid sequence of non-human primate TERT isoform 4 may comprise or consist of the sequence of SEQ ID NO: 25 (reverse machine translation of GenBank Accession No. PNI27664.1):

```

1 atgccgcgcg cgccgcgcgtc ccgcgcgggtc cgcacgcgtc tgccgcggca ttatcgccaa
61 gtgtcgccgc tggcgacattt tgtgcccgc ctggggccgc agggctggcg cctgggtgcag
121 cgccgcgcgtc cggcgccgtt tcgcgcgcgtc gtggcgca ggcgcgttgc cgtgcgcgtgg
181 gatgcgcgcgc cgccgcgcggc ggccgcgcgc tttcgccagg tgagctgcct gaaaagaactg
241 gtggcgccgc tgctgcgcgc cctgtgcgc aaacgcgtgc ggcgttggc
301 tttgcgtgc tggatggcgc gcggggggc cgcggggcaag cgtttaccac cagcggtgcgc
361 agctatgc cgaacaccgt gacccatgcg ctgcgcggca gcggcgccgtc gggcctgcgt
421 ctgcgcgcgc tggggcatga tgtgtggtg catctgtgg cgcgcgtgcgc gctgtttgt
481 ctggcgccgc cgagctgcgc gtatcagggt tgcggccgcgc cgctgtatca gctggggcg
541 ggcggccagg cgccgcgcgc gcccgcgtcg aacggggccgc ggcggccgc gggctgcgaa
601 cgccgcgtgg accatagcgt gcgcgaaageg ggccgcgtgg tgggcctggc ggcgcggggc
661 ggcgcgcgcgc gccggggcag cgccgcgcgc aacccatgcg tggccaaacg cccgcgcgc
721 ggcgcgcgcgc cggaaccggc acgcaccccg gtggggccagg gcaagctggc gcatccgggc
781 cgccaccgcg gccccagcga tcggggctt tgcgtggta gcccggcgcc cccggggaa
841 gaagcgacca gcctggaaagg cgccgcgtcg ggcacccgcg atagccatcc gagegtggc
901 cgccagcatc atgcggggcc gcccgcgtcg agccgcgcgc cgcgcgcgtg ggataccccc
961 tggccgcgcgc tggatgcgg aaccaaacat ttctgtata gcaagccgcgaa taaaacat
1021 ctgcgcgcgc gctttctgtc gagcagcctg cgcgcgcgc tgaccggcgcc ggcgcgcgt
1081 gtggaaacca ttttttggg cagccgcgcg tggatgcgg gcaccccgcc cccgcgtgg
1141 cgccgcgcgc aacccatgcg cccgcgttc tggaaactgcg gggcaaccat
1201 ggcgcgtgc cgtatggcgt gctgtgaaa acccattgcg cgcgcgtgc ggcgggtgacc
1261 cccggggccgg ggcgtgtgcgc ggcggaaaaaa cccgcaggcc ggcgtggccgc gccggaaagaa
1321 gaagataaccg atccgcgcgc cctgggtgcag ctgcgcgcgc agcatagcag cccgtggcag
1381 gtgtatggc ttgtgcgcgc gtcgcgtgc cgcgcgtgc cgcgcgcgc gtcggggcagc
1441 cgccataacg aacccgcgtt tctgcgcac accaaaaaat ttattagccct gggcaacat
1501 gcaactgaa ggcgcgtcg actgcacccgtt aaaatgagcg tgcgcgtcg cgcgcgtcg
1561 cgccgcgcgc cggggcgccg cagcgccgcg gcccgcgcgc atccgcgtcg cgcgcgtcg
1621 ctggcgaaat ttctgcattt gctgtatgcg tggatgtgg tggaaactgcg ggcgcgttt
1681 ttttatgtga ccgaaaccac ctttcagaaa aacccgcgtt tttttatgc caaaacgcgt
1741 tggagcaac tgcagacat tggcattgcg cagcatctga aacgcgtgcg cgcgcgtcgaa
1801 ctgagcgaag cggaagtgcg ccagcatcg gaagcgccgc cggcgccgtc gaccagccgc
1861 ctgcgcgttta ttccgaaacc ggcgcgtcg cgcgcgtcg tggatgtgg tggatgtgg
1921 ggcgcgcgc ccttcgcgcg cggaaacgcg gccggacgcg tggccgcgcg cgtgaaacgc
1981 ctgttttagcg tgctgaacta tgaacgcgcg cgcgcgcgc ggcgcgtgg cgcgcgtcg
2041 ctgggcgtgg atgatattca tcgcgcgtgg cgcacccgttgc tggatgtgg cgcgcgcgc

```

-continued

2101 gatccgcgcg cggaactgta ttttgtaaa gtggatgtga ccggcgctgta tgataccatt
 2161 ccgcaggatc gcctgaccga agtGattgcg agcattatta aaccgcagaa cacctattgc
 2221 gtgcgcgcgt atgcgggtgt gcagaaagcg ggcgcattggcc atgtgcgcaaa agcgtttaaa
 2281 agccatgtga gcacccctgac cgatctgeag cgcgttatgc ggcgcattgtt ggcgcattctg
 2341 caggaaacca gcccgcgtcg cgtgcgggtt attattgaac agagcagcag cctgaacgaa
 2401 gcgagcagcg gcctgtttga tgtgtttctg cgcttgcgtt ggcgcattgc ggtgcgcatt
 2461 cgccggccaaa gctatgtca gtgcgcaggc attccgcaggc gcaaggatctt gggccacccctg
 2521 ctgtgcagcc tgcgcgtatgg cgatatggaa aacaaactgt ttgcggcat tcgcgcgcgt
 2581 ggcctgcgtgc tgcgcgtgtt ggtatgtttt ctgcgtgttgc cccgcattctt gacccatgc
 2641 aaagcgtttc tgagctatgc gcgcaccaggc attcgccgcga ggcctgacccctt taaccgcggc
 2701 tttaaagcg gccgcaacat gcgcgcgaaa ctgtttggc tgctgcgcctt gaaatgcgc
 2761 agcctgtttc tggatctgca ggtgaacagc ctgcagaccg tgcgttgcacaa catttataaa
 2821 attctgtgc tgcaggcgttgc tgcgttgc tgcgtgc tgcgtgc tgcgtgc
 2881 caggtgttgcgaaa aacccgcac ctgttttgcg cgcattatta ggcgcgcgc ggcgcgc
 2941 tatagcattc tgaaagcgaa aaacgcgggc atgagcctgg ggcgcgaaagg cgcggcgcc
 3001 cgcgtgcgcg gcaagcgtat gcaaggcgttgc tgcgtgc gacccgcgc gacccagctg
 3061 cgcgcgcgcg tgcgttgcg tgcgttgcg tgcgtgc gacccgcgc gacccagctg
 3121 agccgcggcac tgcgcggcac caccctgago ggcgcgttgcgaa cccggcgctg
 3181 ccgagcgttgc ttaaaaaccat tctggat.

The amino acid sequence of non-human primate TERT isoform 5 may comprise or consist of the sequence of SEQ ID NO: 26 (also described at GenBank Accession No. PNI27665.1):

1 mpraprrcrav rslllrshyre vplplatfvrr lgpqgwrlvq rgdpaafral vaqclvcvpw
 61 darpppaaps frqvsclkel varvlqrle rgaknvlafg falldgargg ppeafhtsvr
 121 sylpntvtada lrgsgawgll lrrvgddvlv hllarcalfv Iavpscayqv cggpplqlga
 181 atqarpppha sgprrrlgce rawnhsvrea gvplglpapg arrrggsasr slplpkprrr
 241 gaapepertp vggswahpg rtrgpsdrgf cvvsparpae eatslegals gtrhshpsvg
 301 rqhhagppst srpprpdtp cppvyetaekh flyssgdkeq lrpsflssl rpsltgarrr
 361 vetiflgsrp wmpgtprrlp rlpqrywqmr plflellghn aqcpqygvllk thcplraavt
 421 paagvcarek pggsavaapee edtdprrlvq llrqhsspawq vygfvraclr rlvppglwgs
 481 rhnerrflrn tkkfislgkh aklslqeltw kmsvrdcawl rrspvgvgsvp aaehrlreei
 541 lakflhwlm vvvvellarf fyvtettfqk nrlffyrrksv wsklqsigir qhlkrvqlre
 601 lseaevrqhq earpalltsr lrfipkpdgl rpivnmddyv gartfrrekr aerltsrvka
 661 lfsvlnyera rrpqllgasv lglddihrw rtfvrlrvraq dpppelyfvk drlteviashi
 721 ikpqntycvr ryavvvqkaah ghvrkafksh vlrpvpgdpa glhpvhaalq pvlrrhgeqa
 781 vcgdsagraa pafgg.

The nucleic acid sequence of non-human primate TERT isoform 5 may comprise or consist of the sequence of SEQ ID NO: 27 (reverse machine translation of GenBank Accession No. PNI27665.1):

1 atgcgcgcgcg cgcgcgcgtt cgcgcgcgtt ggcgcgcgc ggcgcgcgc ttatcgccaa
 61 gtgcgtgcgc tggcgcaccc ttgtgcgcgc cttggccgcg aggctgggg cttgggtgcag
 121 cgcggcgatc cggcggcggtt tgcgcgcgtt gtggcgactt ggcctgggtgtt cgtgcgcgtgg

-continued

181 gatgcggecc cgccggcgcc ggccgcgac tttcgccagg tgagctgcct gaaagaactg
 241 gtggcgcgcg tgctgcagcg cctgtgcgaa cgccggcgca aaaacgtgct ggcttggc
 301 tttgcgtgc tggatggcgcc gcggggggcc ccgcggaaag cgtttaccac cagcgtgcgc
 361 agctatgtc cgaacaccgt gaccgatgcg ctgcgcggca gcggcgcggt gggcctgctg
 421 ctgcgcgcg tggcgatga tgtgtggtg catctgtgg cgccgtgcgc gctgtttgtg
 481 ctgggtggcgcc cgagctgcgc gtatcagggtg tgccggccgc cgctgtatca gctggggcg
 541 gegaccagg cgccggccgc gcccgcattcg ageggggccgc gcccggccct gggctgcgaa
 601 cgccgcgtgaa accatagcgt gcgcgaaagcg ggctgcgcg tggcctgccc ggccggggc
 661 gcccgcgcgc gccggcgccag cgccgcgcgc agccgtgcgc tgccgaaacg cccgcgcgc
 721 ggccgcggcgcc cggaaccggaa acgcaccccg gtggggccagg gcagctggcgc gcatccggc
 781 cgccaccccgcc gccccggcgca tgccgggttt tgccgtggta gcccggcgcc cccggggaa
 841 gaagcgacca gccttggagg cgccgtgagc ggcacccgc atagccatcc gagegtggc
 901 cgccagcatac atgcggggccc gccgagcacc agccggccgc cgccggccgtg ggataaccccg
 961 tgcccgccgg tggatcgca aaccaaacat ttctgtata gcagcgccga taaagaacag
 1021 ctgcgcgcga gctttctgtc gagcagcctg cgcccgagcc tgaccggcgcc gcccgcctg
 1081 gtggaaaccca ttttttgggg cagccggcccg tggatgggg gcacccggcg cccgcctgcg
 1141 cgccctgcgcg acgcgtattt gcagatgcgc ccgcgttttc tggactgtct gggcaaccat
 1201 gcccgcgtgcc cgtatggcggt gctgtgaaa acccattgcc cgctgcgcgc ggccgggtgacc
 1261 cccggggggcc gccgtgtgcgc gcccggaaaa ccgcaggccca gcgtggggcc gcccggaa
 1321 gaagataacccg atcccgcccg cctgggtgcag ctgcgcgcgc agcatagcag cccgtggcag
 1381 gtgtatggct ttgtgcgcgc gtccgtgcgc cgcctgggtc cgccggccct gtggggcagc
 1441 cgccataacg aacgcgcgtt tctgcgcac accaaaaaat ttattagcct gggcaacat
 1501 gcgaaactga gcctgcagga actgacctgg aaaaatgagcg tgccgcattt cgccgtggctg
 1561 cgccgcgcgc cggggcgccgg cagccgtgcgc gcccggaaatc atccgcgcgc cggaaattt
 1621 ctggcgaaat ttctgcattt gctgtatgcgc gtgtatgtgg tggactgtct gcccgcgtt
 1681 ttttatgtca ccgaaaccac ctttcagaaaa aaccgcgtt tttttatgc caaaacgtg
 1741 tggagcaaac tgcagagcat tggcatttcgc cagcatctga aacgcgtgca gctgcgcga
 1801 ctgagcgaag cggaagtgcg ccagcatcg gaagcgccgc cgccgcgtct gaccagccgc
 1861 ctgcgcgttta ttccgaaacc ggtatggcctg cgcccgattt tgaacatggaa ttatgtgg
 1921 ggccgcgcgc ccttcgcgc cggaaacgc gcccggccgc tgaccaggcc cgtggaaagecg
 1981 ctgttttagc tgctgaacta tgaacgcgcg cgcgcgttgc gctgtgggg cgccgcgtg
 2041 ctgggcctgg atgatattca tgcgcgtgg cgccaccccttgc tgcgcgtgc gcccgcgcag
 2101 gatccgcgcgc cggaactgtt ttttgtgaaa gatgcgcgttgc cggaaactgtat tgcgcgttgc
 2161 attaaacccg agaacaacctt ttgcgtgcgc cgcgtatgcgc tggtgcagaa agccggcgcat
 2221 ggccatgtgc gcaaaacgc taaaagccat gtgcgcgcgc cgggtggggcc cgcgcgcgc
 2281 ggccatgtgc cgggtgcgc ggcgcgtgc cgggtgcgc gcccgcgttgc cggaaacaggcg
 2341 gtgtgcggcg atagcgccgg ccgcgcggcg cccggcggttgc gggc.

-continued

The amino acid sequence of non-human primate TERT isoform 6 may comprise or consist of the sequence of SEQ ID NO: 28 (also described at GenBank Accession No. PNI27666.1):

```

1 mpraprrcav rslllrshyre vplatfvrr lgpqgwrlvq rgdpaafral vaqclvcvpw
61 darpppaaps frqvscmekl varvlqlrce rgaknvlafg falldgargg ppeafittsvr
121 sylpntvtda lrgsgawgll lrrvgddvlv hllarcalfv lvapscayqv cgpplqlga
181 atqarpppha sgprrrlgce rawnhsvrea gvplglpapg arrrggsasr slplpkprrr
241 gaapepertp vgqgswahpg rtrgpsdrgf cvvsparpae eatslegals gtrhshpsvg
301 rqhhagppst srpprpwdtp cppvyaetkh flyssgdkeq lrpsfllssl rpsltgarl
361 vetiflgsrp wmpgtprrlp rlpqrwywqmr plflellgnh aqcpqygvllk thcplraavt
421 paagvcarek pqgsvaapee edtdprrlvq llrqhsspwm vygfvraclr rlvppglwgs
481 rhnerrflrn tkkfislgkh aklslqeltw kmsvrdcawl rrspgvgsvp aaehrlreei
541 lakflhwlm vvvvelrsf fyvtettfqk nrlffyriksv wsklqsigir qhlkrvqlre
601 lseaevrqhq earpalltsr lrfipkpdgl rpivnmdyvv gartfrrekr aerltsrvka
661 lfsvlnyera rrppllgasv lglddihrw rtfvrlrvraq dpppelyfvk vdvtgaydti
721 pqdrltevia siikpqntyc vrriyavvqka ahghvrkafk shvstltdlq pymrqfvahl
781 qetsplrdav iieqssslne assglfdvfl rfvcvhavri rgksyvqcqg ipqgsilstl
841 lcslcygdm e nklfagirrd glllrlvddf llvphltha kaflrtlvrq vpeygcvvnl
901 rktvvnfve dealggtafv qlpahglfpw cgllldtrtl evqsdysya rtsirasltf
961 nrgfkagrn m rrklnfgvrlr kchslfllq vnsllqtvcn iykillqay rfhacvlqlp
1021 fhqvwknpt fflriisdt slcysilkak naaqtqlsrk lpgttlsale aanpalpsd
1081 fktild.

```

The nucleic acid sequence of non-human primate TERT isoform 6 may comprise or consist of the sequence of SEQ ID NO: 29 (reverse machine translation of GenBank Accession No. PNI27666.1):

```

1 atgccgcgcg cgccgcgcgtg ccgcgcgcgtg cgcagccgtc tgccgcgcgca ttatcgcgaa
61 gtgctgccgc tggcgacctt tgtgcgcgcg ctggggccgc agggctggcg cctgggtgcag
121 cgccgcgcgtc cggcgccgtt tcgcgcgcgtc gtggcgcaagt gcctgggtgtg cgtgcgcgtgg
181 gatgcgcgcg cgcgcgcgc ggcgcgcgagc tttcgccagg tgagctgcct gaaagaactg
241 gtggcgccgc tgctgcagcg cctgtgcgaa cgcggcgccgaa aaaaactgtc ggcgtttggc
301 tttgcgcgtgc tggatggcgc gcgcggcgcc cccgcggaaag cgtttaccac cagcgtgcgc
361 agctatgtc cgaacaccgt gaccgatgcg ctgcgcgcgc gggcgccgtg gggcctgcgt
421 ctgcgcgcgc tggcgatgtc tctgtggcgt catctgtgg cgcgcgcgcgtc gctgtttgtg
481 ctggcgccgc cgagctgcgc gtatcagggtc tgccgcgcgc cgcgttatca gctggcgccgc
541 gegaccagg cgcgcgcgc gcccgcgtcg agcggccgcg cccgcgcgcct gggctgcgaa
601 cgcgcgtggc accatagcgt gcgcgaaagcg ggctgcgcgc tggccgtgc ggcgcggggc
661 ggcgcgcgc gcccgcgcgc gcccgcgtcg agcggccgcg cccgcgcgc
721 ggcgcgcgcgc cggAACCGGA acgcaccccg gtggggccagg gcagctggcgc gcatccgggc
781 cgcaccaggcgcg gcccgcgcgc tgcggcgtt tgcgtggcgt gcccggcgcc cccggcgaa
841 gaagcgacca gcctggcagg cgcgcgtcg ggcacccgcg atagccatcc gagegtggc
901 cgcacccgcg atgcggggccc gcccgcgtcg agccgcgcgc cgcgcgcgtc ggataccccg
961 tgcccgccgc tgcgtgcgcgaa aaccaaacat tttctgtata gcagcggcga taaagaacag

```

-continued

1021 ctgcggcccg gctttctgtc gagcagccgt cgccccgagcc tgacccggcgc gcgcgcgc
1081 gtggaaacca ttttctggg cagccgcgg tggatgccgg gcaccccgcc cccgcctgc
1141 cgccctgcgc agcgctattt gcagatgcgc cccgtgttc tggactgtc gggcaaccat
1201 ggcgcgtgcc cgtatggcgt gctgctgaaa acccattgcc cgcgcgcgc ggccgggtacc
1261 ccggcggccg gctgtgcgc gcgcgaaaaa cccgcaggca gcgtggccgc gcccggaaagaa
1321 gaagataccg atccgcgcgc cctgggtcag ctgcgtgcgc a诶catagcag cccgtggc
1381 gtgtatggct ttgtgcgcgc tgccctgcgc cgcctggcgc cgcggggcct gtggggcagc
1441 cgccataacg aacgcgcgtt tctgcgcac accaaaaaat ttattagcct gggcaaacat
1501 gcgaaactga gcctgcagga actgcacctgg aaaaatgagcg tgcgcgattt cgcgtggc
1561 cggccgcgc cgggcgtggg cagcgtgcgc gcggcggaaac atgcgcgcgc cgaagaaa
1621 ctggcgaaat ttctgcattt gctgatgagc gtgtatgtgg tggactgtc ggcgcgc
1681 ttttatgtga cccggaaaccac ctttcagaaa aaccgcgtt tttttatgc caaaaggcgt
1741 tggagcaaac tgcagagcat tggcattcgc cagcatctga aacgcgtgc gctgcgc
1801 ctgagcgaag cggaaagtgcgc ccgcgcatttgc gaaagcgcgc cggcgcgtgc
1861 ctgcgcgttta ttccgaaacc cggatggccttgc cgcgcgcatttgc tgaacatgg
1921 ggcgcgcgc ccttcgcgc cggaaaacgc cggggacgcgc tgaccagcgc cgtgaaagc
1981 ctgttagcg tgctgaacta tgaacgcgcgc cgcgcgcgcgc gcctgcgtgg
2041 ctgggcctgg atgatattca tcgcgcgtgg cgcacatttgc tgctgcgcgc
2101 gatccgcgc cggaaactgttgc ttttgtaaa gtggatgtga cggggcgatgc
2161 cccgcaggatc gcctgaccgc agtggatgcgc agcattatta aaccgcagaa
2221 gtgcgcgcgc atgcgggtggt gcagaaaagcg cgcgcgcgc
2281 agccatgtga gcacccgtac cgcgcgcgc
2341 caggaaacca gcccgcgc cgcgcgcgc
2401 ggcgcgcgc gcctgtttga tggatggatgc
2461 cgcgcgcgc
2521 ctgtgcgcgc
2581 ggcgcgcgc
2641 aaagcggttgc
2701 cgcgcgcgc
2761 cgcgcgcgc
2821 gaagtgcaga
2881 aaccgcggc
2941 aaatgcgc
3001 atttataaaa
3061 tttcatc
3121 agccgtgc
3181 ctgcggcc
3241 tttaaaaacca

[0136] In some embodiments of the compositions and methods of the disclosure, an amino acid sequence of TERT may comprise or consist of a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99%, 100% or any percentage in between of identity to one or more of SEQ ID NOS: 6, 8, 10-13, 18, 20, 22, 24, 26, or 28. In some embodiments of the compositions and methods of the disclosure, an amino acid sequence of a portion of TERT, functional or non-functional, may comprise or consist of a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99%, 100% or any percentage in between of identity to one or more of SEQ ID NOS: 6, 8, 10-13, 18, 20, 22, 24, 26, or 28.

[0137] In some embodiments of the compositions and methods of the disclosure, a nucleic acid sequence of TERT may comprise or consist of a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99%, 100% or any percentage in between of identity to one or more of SEQ ID Nos: 1-5, 7, 9, 14-17, 19, 21, 23, 25, 27, 29, 30, or 31. In some embodiments of the compositions and methods of the disclosure, a nucleic acid sequence of a portion of non-human primate TERT, functional or non-functional, may comprise or consist of a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99%, 100% or any percentage in between of identity to one or more of SEQ ID Nos: 1-5, 7, 9, 14-17, 19, 21, 23, 25, 27, 29, 30, or 31. In some embodiments, the instant ribonucleic acids may correspond to the native gene sequences coding for the above-listed TERT proteins or may correspond to variants that are made possible due to the redundancy of the genetic code, as would be understood by one of ordinary skill in the art. In some embodiments, the codon selection may be optimized to optimize protein expression and/or reduced or increased immunogenicity using algorithms and methods known by those of ordinary skill in the art.

[0138] In some embodiments, an mRNA sequence may be synthesized as an unmodified or modified mRNA. An mRNA may be modified to enhance stability and/or evade immune detection and degradation. A modified mRNA may include, for example, one or more of a nucleotide modification, a nucleoside modification, a backbone modification, a sugar modification, and/or a base modification. In some embodiments, the modified nucleoside is pseudouridine or a pseudouridine analog. In some embodiments, the pseudouridine analog is N-1-methylpseudouridine. In some embodiments, the modified nucleoside is 5-methoxyuridine. In some embodiments a modified nucleotide as used herein may comprise any of the moieties listed in Table 1B.

TABLE 1B

Common name
pseudouridine
N-1-methylpseudouridine
5-methoxyuridine
1,2'-O-dimethyladenosine
1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine
1-methyladenosine
1-methylguanosine
1-methylinosine
1-methylpseudouridine
2,2-dimethyl-guanosine
2',3'-dideoxyadenosine
2',3'-Dideoxycytidine
2',3'-Dideoxyguanosine
2',3'-Dideoxyinosine
2',3'-Dideoxynucleosides
2',3'-Dideoxythymidine
2',3'-dideoxythymine
2',3'-Dideoxyuridine
2,6-diaminopurine
2'-O-ribosyladenosine (phosphate)
2'-Amino-2'-deoxyadenosine
2'-Amino-2'-deoxyadenosine
2'-Amino-2'-deoxyuridine
2-Amino-6-chloropurineriboside
2-Amino-6-Cl-purine-2'-deoxyriboside
2-aminoadenosine
2-Aminoadenosine
2-Aminopurine-2'-deoxyriboside
2-Aminopurine-riboside
2'-Azido-2'-deoxyadenosine
2'-Azido-2'-deoxycytidine
2'-Azido-2'-deoxyguanosine
2'-Azido-2'-deoxyuridine
2'-Deoxyinosine
2'-Deoxy-P-nucleoside
2'-Deoxyuridine
2'-Deoxyzebularaine
2'-Fluoro-2'-deoxyadenosine
2'-Fluoro-2'-deoxycytidine
2'-Fluoro-2'-deoxyguanosine
2'-Fluoro-2'-deoxyuridine
2'-Fluoro-thymidine
2-methyl-adenosine
2-methyl-guanosine
2-methylthio-N ⁶ -(cis-hydroxyisopentenyl) adenosine
2-methylthio-N-6-isopentenyl-adenosine
2-methylthio-N ⁶ -threonylcarbamoyladenosine
2'-O-Methyl-2-aminoadenosine
2'-O-Methyl-5-methyluridine
2'-O-methyladenosine
2'-O-methylcytidine
2'-O-methylguanosine
2'-O-methylinosine
2'-O-Methyl-N ⁶ -Methyladenosine
2'-O-methylpseudo尿idine
2'-O-methyluridine
2'-O-ribosylguanosine (phosphate)
2-Thio-2'-deoxycytidine
2-Thiocytidine
2-Thiothymidine
2-thiouridine
3,2'-O-dimethyluridine
3'-Amino-2',3'-dideoxyadenosine
3'-Amino-2',3'-dideoxycytidine
3'-Amino-2',3'-dideoxyguanosine
3'-Amino-2',3'-dideoxythymidine
3'-Azido-2',3'-dideoxyadenosine
3'-Azido-2',3'-dideoxycytidine
3'-Azido-2',3'-dideoxythymidine
3'-Azido-2',3'-dideoxyuridine
3'-Deoxy-5-Methyluridine
3'-Deoxyadenosine
3'-deoxyadenosine (cordycepin)
3'-Deoxycytidine
3'-Deoxyguanosine
3'-deoxythymine
3'-Deoxyuridine
3-methylcytidine
3-methyluridine
3'-O-(2-nitrobenzyl)-2'-Deoxyadenosine
3'-O-(2-nitrobenzyl)-2'-Deoxyinosine
3'-O-Methyladenosine
3'-O-Methylcytidine
3'-O-Methylguanosine
3'-O-Methyluridine
4-acetyl-cytidine
4-Thiothymidine
4-Thiouridine
5-(carboxyhydroxymethyl) uridine methyl ester

TABLE 1B-continued

TABLE 1B-continued

Common name
5-(carboxyhydroxymethyl)uridine
5,2'-O-dimethyluridine
5,6-Dihydro-5-Methyluridine
5,6-Dihydrouridine
5-[(3-Indolyl)propionamide-N-allyl]-2'-deoxyuridine
5-Aminoallyl-2'-deoxycytidine
5-Aminoallyl-2'-deoxyuridine
5-Aminoallylcytidine
5-Aminoallyluridine
5-Bromo-2',3'-dideoxyuridine
5-Bromo-2'-deoxycytidine
5-Bromo-2'-deoxyuridine
5-Bromocytidine
5-Bromouridine
5-carbamoylmethyl-2'-O-methyluridine
5-carbamoylmethyluridine
5-Carboxy-2'-deoxycytidine
5-Carboxy-2'-deoxycytidine
5-carboxymethylaminomethyl-2-thio-uridine
5-carboxymethylaminomethyluridine
5-Carboxymethylesteruridine
5-carboxymethyluridine
5-Carboxyuridine
5-Fluoro-2'-deoxyuridine
5-fluoro-uridine
5-Formyl-2'-deoxycytidine
5-Formyl-2'-deoxyuridine
5-formyl-2'-O-methylcytidine
5-formylcytidine
5-Formyluridine
5-Hydroxy-2'-deoxycytidine
5-Hydroxycytidine
5-Hydroxymethyl-2'-deoxycytidine
5-Hydroxymethyl-2'-deoxyuridine
5-hydroxymethylcytidine
5-Hydroxymethyluridine
5-hydroxyuridine
5-Iodo-2'-deoxycytidine
5-Iodo-2'-deoxyuridine
5-Iodocytidine
5-Iodouridine
5-methoxyaminomethyl-2-thio-uridine
5-methoxycarbonylmethyl-2'-O-methyluridine
5-methoxycarbonylmethyl-2-thiouridine
5'-methoxycarbonylmethyl-uridine
5-methoxycarbonylmethyluridine
5-Methoxycytidine
5-Methoxyuridine
5-methoxy-uridine
5-Methyl-2'-deoxycytidine
5-methyl-2-thio-uridine
5-methyldiaminomethyl-uridine
5-methylcytidine
5-methylhydrouridine
5-methyluridine
5-Propargylamino-2'-deoxycytidine
5-Propargylamino-2'-deoxyuridine
5-Propynyl-2'-deoxycytidine
5-taurinomethyl-2-thiouridine
5-taurinomethyluridine
6-Aza-2'-deoxyuridine
6-Azacytidine
6-Azauridine
6-chloropurine riboside
6-Chloropurine-2'-deoxyriboside
6-O-methylguanosine
6-Thio-2'-deoxyguanosine
7-Deaza-2'-deoxyadenosine
7-Deaza-2'-deoxyguanosine
7-Deaza-7-Propargylamino-2'-deoxyadenosine
7-Deaza-7-Propargylamino-2'-deoxyguanosine
7-Deazadenosine
7-Deazaguanosine
7-methylguanosine
7-methyl-guanosine

TABLE 1B-continued

Common name
8-Azaadenosine
8-Azidoadenosine
8-Chloro-2'-deoxyadenosine
8-Oxo-2'-deoxyadenosine
8-Oxo-2'-deoxyguanosine
8-Oxoadenosine
8-Oxoguanosine
a 2'-deoxynucleoside
ac4C N4-acetylcytidine
Am 2'-O-methyladenosine
an —O-methylnucleoside
Ar(p) 2' —O-ribosyladenosine (phosphate)
Araadenosine
Aracytidine
Araguanosine
Arauridin
benzimidazole riboside
beta-D-mannosyl-queosine
Biotin-16-7-Deaza-7-Propargylamino-2'-deoxyguanosine
Biotin-16-Aminoallyl-2'-dCTP
Biotin-16-Aminoallyl-2'-dTTP
Biotin-16-Aminoallylcytidine
Biotin-16-Aminoallyluridine
chm5U 5-(carboxyhydroxymethyl)uridine
2'-O-methylcytidine
5-carboxymethyluridine
5-carboxymethylaminomethyluridine
Cyanine 3-5-Propargylamino-2'-deoxycytidine
Cyanine 3-6-Propargylamino-2'-deoxyuridine
Cyanine 3-Aminoallylcytidine
Cyanine 3-Aminoallyluridine
Cyanine 5-6-Propargylamino-2'-deoxycytidine
Cyanine 5-6-Propargylamino-2'-deoxyuridine
Cyanine 5-Aminoallylcytidine
Cyanine 5-Aminoallyluridine
Cyanine 7-Aminoallyluridine
dihydrouridine
DabcyL-5-3-Aminoallyl-2'-dUTP
Desthiobiotin-16-Aminoallyl-Uridine
Desthiobiotin-6-Aminoallyl-2'-deoxycytidine
dihydrouridine
5-formylcytidine
5-formyl-2'-O-methylcytidine
N6-glycinylcarbamoyladenosine
galactosyl-queosine
2'-O-methylguanosine
2'-O-ribosylguanosine (phosphate)
5-hydroxymethylcytidine
5-hydroxyuridine
hydroxywybutosine
N6-isopentenyladenosine
2'-O-methylinosine
wyosine
inosine
N6-(cis-hydroxisopentenyl)adenosine
Isoguanosine
l-methylguanosine
1-methyladenosine
1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine
1,2'-O-dimethyladenosine
1-methylguanosine
1-methylinosine
1-methylpseudoouridine
N2,N2-dimethylguanosine
N2,N2,7-trimethylguanosine I inosine
N2,7-dimethylguanosine
N2-methylguanosine
3-methylcytidine
3-methyluridine
3,2'-O-dimethyluridine
N4-methylcytidine
5-methylcytidine
5-methylhydrouridine
5-methyluridine
5,2'-O-dimethyluridine

TABLE 1B-continued

Common name
N6,N6-dimethyladenosine
N6,N6,2'-O-trimethyladenosine
N6-methyladenosine
N6,2'-O-dimethyladenosine
7-methylguanosine
mannosyl-queuosine
5-(carboxyhydroxymethyl)uridine
5-methoxycarbonylmethyl-2-thiouridine
5-methoxycarbonylmethyl-2'-O-methyluridine
5-methoxycarbonylmethyluridine
2-methylthio-N6-(cis-hydroxyisopentenyl) adenosine
2-methylthio-N6-threonyl carbamoyladenosine
N1-Ethylpseudouridine
N1-Methoxymethylpseudouridine
N1-Methyl-2'-O-Methylpseudouridine
N1-Methyladenosine
N1-Propylpseudouridine
N ² ,7-dimethylguanosine
N ² ,N ² ,7-trimethylguanosine
N ² ,N ² -dimethylguanosine
N2-Methyl-2'-deoxyguanosine
N ² -methylguanosine
N ⁴ -acetylcytidine
N4-Biotin-OBEA-2'-deoxycytidine
N4-Methyl-2'-deoxycytidine
N ⁴ -methylcytidine
N ⁶ -(cis-hydroxyisopentenyl)-adenosine
N ⁶ ,2'-O-dimethyladenosine
N ⁶ ,N ⁶ ,2'-O-trimethyladenosine
N ⁶ ,N ⁶ -dimethyladenosine
N ⁶ -glycimylcarbamoyladenosine
N ⁶ -isopentenyladenosine
N6-isopentenyl-adenosine
N6-Methyl-2-Aminoadenosine
N6-Methyl-2'-deoxyadenosine
N ⁶ -methyladenosine
N6-methyladenosine
N ⁶ -threonylcarbamoyladenosine
ncm5U 5-carbamoylmethyluridine
ncm5Um 5-carbamoylmethyl-2'-O-methyluridine
N1-methyladenosine
N-uridine-5-oxyacetidmethylester
peroxywybutosine
O6-Methyl-2'-deoxyguanosine
O6-Methylguanosine
hydroxywybutosine
undermodified hydroxywybutosine
O-Methylpseudouridine
peroxywybutosine
Pseudoisocytidine
Puromycin
queosine
2-thiouridine
N6-threonylcarbamoyladenosine
Thienocytidine
Thienoguanosine
Thienouridine
2'-O-methyluridine
undermodified hydroxywybutosine
uridine-5-oxyaceticacid(v)
uridine-5-oxyaceticacidmethylester
wybutosine
wybutoxosine
wynosine
Xanthosine
5-taurinomethyl-2-thiouridine
5-taurinomethyluridine
2'-O-methylpseudouridine

[0139] In some embodiments, an mRNA may be synthesized from naturally occurring bases and/or base analogs (modified bases) including, but not limited to, purines (adenine (A), guanine (G)) or pyrimidines (thymine (T), cytosine (C), uracil (U)), and analogues and derivatives thereof,

e.g. 1-methyl-adenine, 2-methyl-adenine, 2-methylthio-N-6-isopentenyl-adenine, N6-methyl-adenine, N6-isopentenyl-adenine, 2-thio-cytosine, 3-methyl-cytosine, 4-acetyl-cytosine, 5-methyl-cytosine, 2,6-diaminopurine, 1-methyl-guanine, 2-methyl-guanine, 2,2-dimethyl-guanine, 7-methyl-guanine, inosine, 1-methyl-inosine, pseudouracil (5-uracil), pseudouridine, N-1-methyl-pseudouridine, dihydro-uracil, 2-thio-uracil, 4-thio-uracil, 5-carboxymethylaminomethyl-2-thio-uracil, 5-(carboxyhydroxymethyl)-uracil, 5-fluoro-uracil, 5-bromo-uracil, 5-carboxymethylaminomethyl-uracil, 5-methyl-2-thio-uracil, 5-methyl-uracil, N-uracil-5-oxyacetic acid methyl ester, 5-methylaminomethyl-uracil, 5-methoxyaminomethyl-2-thio-uracil, 5'-methoxycarbonylmethyl-uracil, 5-methoxy-uracil, uracil-5-oxyacetic acid methyl ester, uracil-5-oxyacetic acid (v), 1-methyl-pseudouracil, queosine, beta-D-mannosyl-queuosine, wybutoxosine, and phosphoramides, phosphorothioates, peptide nucleotides, methylphosphonates, 7-deazaguanosine, 5-methylcytosine and inosine.

[0140] In some embodiments, an mRNA may be synthesized from naturally occurring nucleosides and/or nucleoside analogs (modified nucleosides) including, but not limited to, nucleosides comprising adenosine (A), guanosine (G)) or pyrimidines (thymine (T), cytidine (C), uridine (U)), and nucleoside comprising analogues and derivatives thereof, e.g., 3'-deoxyadenosine (cordycepin), 3'-deoxyuridine, 3'-deoxycytosine, 3'-deoxyguanosine, 3'-deoxythymine, 2',3'-dideoxynucleosides, 2',3'-dideoxyadenosine, 2',3'-dideoxyuridine, 2',3'-dideoxycytosine, 2',3'-dideoxyguanosine, 2',3'-dideoxythymine, a 2'-deoxynucleoside, —O— methylnucleoside, 1-methyl-adenine, 2-methyl-adenine, 2-methylthio-N-6-isopentenyl-adenine, N6-methyl-adenine, N6-isopentenyl-adenine, 2-thio-cytosine, 3-methyl-cytosine, 4-acetyl-cytosine, 5-methyl-cytosine, 2,6-diaminopurine, 1-methyl-guanine, 2-methyl-guanine, 2,2-dimethyl-guanine, 7-methyl-guanine, inosine, 1-methyl-inosine, pseudouridine, N-1-methyl-pseudouridine, dihydro-uracil, 2-thio-uracil, 4-thio-uridine, 5-carboxymethylaminomethyl-2-thio-uridine, 5-(carboxyhydroxymethyl)-uridine, 5-fluoro-uridine, 5-bromo-uridine, 5-carboxymethylaminomethyl-uridine, 5-methyl-2-thio-uridine, 5-methyl-uridine, N-uridine-5-oxyacetic acid methyl ester, 5-methylaminomethyl-uridine, 5-methoxyaminomethyl-2-thio-uridine, 5'-methoxycarbonylmethyl-uracil, 5-methoxy-uracil, uracil-5-oxyacetic acid methyl ester, uracil-5-oxyacetic acid (v), 1-methyl-pseudouracil, queosine, beta-D-mannosyl-queuosine, wybutoxosine, 7-deazaguanosine, 5-methylcytosine, and inosine.

[0141] The preparation of such base, nucleoside, nucleotide, and backbone analogues, modifications, and derivatives is known to a person skilled in the art e.g. from the U.S. Pat. Nos. 4,373,071, 4,401,796, 4,415,732, 4,458,066, 4,500,707, 4,668,777, 4,973,679, 5,047,524, 5,132,418, 5,153,319, 5,262,530 and 5,700,642, all of which are incorporated by reference in their entirety.

[0142] In some embodiments, uracil nucleosides of the mRNA are about 80%, about 90%, 95%, 99%, or 100% depleted and replaced with a uracil nucleoside analog, e.g., pseudouridine, 5-methoxyuridine, or N-1-methyl-pseudouridine.

[0143] In some embodiments, an mRNA may contain an RNA backbone modification. Typically, a backbone modification is a modification in which the phosphates of the backbone of the nucleotides contained in the RNA are

chemically modified. Exemplary backbone modifications may include, but are not limited to, modifications in which the phosphodiester linkage is replaced with a member from the group consisting of peptides, methylphosphonates, methylphosphoramidates, phosphoramidates, phosphorothioates (e.g., cytidine 5'-O-(1-thiophosphate)), boranophosphates, and/or positively charged guanidinium groups, or other means of replacing the phosphodiester linkage.

[0144] In some embodiments, an mRNA may contain sugar modifications. A sugar modification may include but is not limited to, 2' O-methyl sugar modifications, 2' fluoro sugar modifications (e.g. 2'-fluororibose), 3' amino sugar modifications, 2' thio sugar modifications, 2'-O-alkyl sugar modifications, 5-methylthioribose, and sugar modifications of 2'-deoxy-2'-fluoro-ribonucleotide (2'-fluoro-2'-deoxycytidine, 2'-fluoro-2'-deoxyuridine), 2'-deoxy-2'-deamino-ribonucleotide (2'-amino-2'-deoxycytidine, 2'-amino-2'-deoxyuridine), 2'-O-alkylribonucleotide, 2'-deoxy-2'-C-alkylribonucleotide (2'-O-methylcytidine, 2'-methyluridine), 2'-C-alkylribonucleotide, and isomers thereof (2'-aracytidine, 2'-arauridine), or azidophosphates (2'-azido-2'-deoxycytidine, 2'-azido-2'-deoxyuridine).

[0145] In some embodiments, an mRNA may be synthesized from one or more of the nucleotide triphosphates comprising any of the nucleosides and nucleotides disclosed herein, or any of the following nucleoside triphosphates: 2'-Deoxyadenosine-5'-O-(1-Thiotriphosphate), 2'-Deoxycytidine-5'-O-(1-Thiotriphosphate), 2'-Deoxyguanosine-5'-O-(1-Thiotriphosphate), 2'-Deoxythymidine-5'-O-(1-Thiotriphosphate), Adenosine-5'-O-(1-Thiotriphosphate), Cytidine-5'-O-(1-Thiotriphosphate), Guanosine-5'-O-(1-Thiotriphosphate), Uridine-5'-O-(1-Thiotriphosphate), 2',3'-Dideoxyadenosine-5'-O-(1-Thiotriphosphate), 2',3'-Dideoxycytidine-5'-O-(1-Thiotriphosphate), 2',3'-Dideoxyguanosine-5'-O-(1-Thiotriphosphate), 3'-Deoxythymidine-5'-O-(1-Thiotriphosphate), 3'-Azido-2',3'-dideoxythymidine-5'-O-(1-Thiotriphosphate), 2',3'-Dideoxyuridine-5'-O-(1-Thiotriphosphate), 2'-Deoxyadenosine-5'-O-(1-Boranotriphosphate), 2'-Deoxycytidine-5'-O-(1-Boranotriphosphate), 2'-Deoxyguanosine-5'-O-(1-Boranotriphosphate), and 2'-Deoxythymidine-5'-O-(1-Boranotriphosphate).

[0146] In some embodiments, an mRNA may include the addition of a “cap” on the N-terminal (5') end, and a “tail” on the C-terminal (3') end. The presence of the cap may provide resistance to nucleases found in eukaryotic cells. The presence of a “tail” may protect the mRNA from exonuclease degradation.

Cap Structure

[0147] In some embodiments, an mRNA may include a 5' cap structure. A 5' cap may comprise for example, a triphosphate linkage and a guanine nucleotide in which the 7-nitrogen is methylated. Examples of cap structures include, but are not limited to, m7G(5')ppp(5')A, G(5')ppp(5')A, and G(5')ppp(5')G. Naturally occurring cap structures comprise a 7-methyl guanosine that is linked via a triphosphate bridge to the 5'-end of the first transcribed nucleotide, resulting in a dinucleotide cap of m7G(5')ppp(5')N, where N is any nucleoside. In vivo, the cap is added in the nucleus by the enzyme guanylyl transferase immediately after initiation of transcription.

[0148] In some embodiments, a 5' cap may comprise an m7(3'OMeG)(5')ppp(5')(2'OMeA)pG or (CleanCapTM 3'

Ome) structure. In some embodiments, a 5' cap may comprise a m7G(5')ppp(5')G. In some embodiments, the Anti-Reverse Cap Analog (“ARCA”) or modified ARCA, is a 5' cap in which the 2' or 3' OH group is replaced with —OCH₃. In some embodiments, the ARCA comprises an 3'-O-Me-m7G(5')ppp(5')G structure. In some embodiments, the 5' cap comprises m7G(5')ppp(5')(2'OMeA)pG. Additional mRNA caps may include, but are not limited to, a chemical structures selected from the group consisting of m7GpppG, m7GpppA, m7GpppC; unmethylated caps (e.g., GpppG); a 71emethylated cap (e.g., m2'7GpppG), a trimethylated cap analog, or anti reverse cap analogs (e.g., ARCA; m7,2'0meGpppG, m72'dGpppG, m7'3'0meGpppG, m7,3 dGpppG and their tetraphosphate derivatives) (see, e.g., Jemielity, J. et al, ‘‘Wove anti-reverse cap analogs with superior translational properties’’, RNA, 9: 1108-1122 (2003)).

[0149] In some embodiments, a suitable cap is a 7-methyl guanylate (“m7G”) linked via a triphosphate bridge to the 5'-end of the first transcribed nucleotide, resulting in m7G (5')ppp(5')N, where N is any nucleoside. A embodiment of a m7G cap utilized in embodiments of the disclosure is m7G(5')ppp(5')G. In some embodiments, the cap is a Cap0 structure. Cap0 structures lack a 2'-O-methyl residue of the ribose attached to bases 1 and 2. In some embodiments, the cap is a Cap1 structure. Cap1 structures have a 2'-O-methyl residue at base 2. In some embodiments, the cap is a Cap2 structure. Cap2 structures have a 2'-O-methyl residue attached to both bases 2 and 3.

[0150] A variety of m7G cap analogs are known in the art, many of which are commercially available. These include the m7 GpppG described above, as well as the ARCA 3'-OCH₃ and 2'-OCH₃ cap analogs (Jemielity, J. et al., RNA, 9: 1108-1122 (2003)). Additional cap analogs for use in embodiments of the disclosure include N7-benzylated dinucleoside tetraphosphate analogs (described in Grudzien, E. et at, RNA, 10: 1479-1487 (2004)), phosphorothioate cap analogs (described in Grudzien-Nogalska, E., et al, RNA, 13: 1745-1755 (2007)), and cap analogs (including biotinylated cap analogs) described in U.S. Pat. Nos. 8,093,367 and 8,304,529, incorporated by reference herein.

[0151] In some embodiments, the 5' cap is inosine, N1-methyl-guanosine, 2'fluoro-guanosine, 7-deaza-guanosine, m7(3'OMeG)(5')ppp(5')(2'OMeA)pG, CleanCapTM m7(3'OMeG)(5')ppp(5')(2'OMeA)pG, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, 2-azido-guanosine, Cap2, Cap4, CAP-003, or CAP-225.

[0152] In some embodiments, the 5' cap comprises or consists of an internal ribosome entry site (IRES). In some embodiments the IRES is within the 5' UTR. In some embodiments, the 5' cap comprises or consists of a 2A self-cleavage peptide, e.g, one or more of P2A, T2A, E2A and F2A.

Tail Structure

[0153] The presence of a “tail” may serve to protect an mRNA from exonuclease degradation. The poly-A tail is thought to stabilize natural messengers and synthetic sense RNA. Therefore, in certain embodiments a long poly-A tail can be added to an mRNA molecule thus rendering the RNA more stable. Poly-A tails can be added using a variety of art-recognized techniques. For example, long poly-A tails can be added to synthetic or in vitro transcribed RNA using poly-A polymerase (Yokoe, et al. Nature Biotechnology).

1996; 14: 1252-1256). A transcription vector can also encode long poly-A tails. In addition, poly-A tails can be added by transcription directly from PCR products. Poly-A may also be ligated to the 3' end of a sense RNA with RNA ligase (see, e.g., Molecular Cloning A Laboratory Manual, 2nd Ed., ed. By Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1991 edition)).

[0154] In some embodiments, an mRNA may include a 3' poly(A) tail structure. The length of the poly-A tail may be at least about 10, 50, 100, 200, 300, 400 or at least about 500 nucleotides. In some embodiments, a poly-A tail on the 3' terminus of an mRNA may include about 10 to 300 adenosine nucleotides (e.g., about 10 to 200 adenosine nucleotides, about 10 to 150 adenosine nucleotides, about 10 to 100 adenosine nucleotides, about 20 to 70 adenosine nucleotides, or about 20 to 60 adenosine nucleotides). In some embodiments, the poly A tail is 120 adenosine nucleotides.

[0155] In some embodiments, an mRNA may include a 3' polyI tail structure. A poly-C tail on the 3' terminus of mRNA may include about 10 to 200 cytosine nucleotides (e.g., about 10 to 150 cytosine nucleotides, about 10 to 100 cytosine nucleotides, about 20 to 70 cytosine nucleotides, about 20 to 60 cytosine nucleotides, or about 10 to 40 cytosine nucleotides). The poly-C tail may be added to the poly-A tail or may substitute the poly-A tail. In some embodiments, the length of the poly-A or poly C tail is associated with the stability of a modified sense mRNA and, therefore, the transcription of the protein. For example, because the length of the poly-A tail may influence the half-life of a sense mRNA molecule, the length of the poly-A tail may be adjusted to modify the level of resistance of the mRNA to nucleases, thereby providing more control over the time course of polynucleotide expression and/or polypeptide production.'

5' an' 3' Untranslated Regions (UTRs)

[0156] In some embodiments, an mRNA may include 5' untranslated region (UTR) and/or a 3' UTR. In some embodiments, a 5' UTR may include one or more elements that affect the stability or translation of an mRNA. In some embodiments, the 5'UTR for example, may include an iron responsive element. In some embodiments, 5' UTR may be between about 50 to about 100, or from about 50 to about 500 nucleotides in length. In some embodiments, 3' UTR includes one or more of a poly-A signal, a binding site for proteins that may affect mRNA stability or localization, or one or more binding sites for miRNAs. In some embodiments, 3' UTR may be between about 0 and about 50 nucleotides, or about 50 to about 100 nucleotides in length.

[0157] Example 3' an' 5' UTR sequences may be derived from mRNAs with relatively long half-lives (e.g., globin, actin, GAPDH, tubulin, histone, or citric acid cycle enzymes) to increase the stability of the sense mRNA molecule. For example, 5' UTR sequence may include a partial sequence of a cytomegalovirus (CMV) immediate-early 1 (IE1) gene, or a fragment thereof to improve the nuclease resistance and/or improve the half-life of the polynucleotide. Generally, these modifications improve the stability and/or pharmacokinetic properties (e.g., half-life) of the polynucleotide relative to their unmodified counterparts, and include, for example modifications made to improve such polynucleotide's resistance to in vivo nuclease digestion.

[0158] In some embodiments, a UTR may improve tissue specific expression, e.g., in the liver.

[0159] In some embodiments, the 3' UTR is a mouse alpha-globin 3' UTR. In some embodiments, the 3' UTR comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO: 32

```
TTAAGCTGCCTCTGGGGCTTGCTCTGGCATGCCCTCTT  
CTCTCCCTGACCTGTACCTCTGGTCTTGAATAAGCCTGAG  
TAGGAAGTCTAG
```

[0160] In some embodiments, the 3' UTR is a wild type human beta-globin 3' UTR. In some embodiments, the 3' UTR comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO: 33

```
CAATTGGCTCGCTTCTTGCTGTCCAATTCTATTAAGGTTCCCT  
TTGTTCCCTAACGTCACACTAAACTGGGGATATTATGAAGGG  
CCTTGAGCATCTGGATTCTGCCTAATAAAAACATTTATTTCAT  
TGGCAATTTC
```

[0161] In some embodiments, the 3' UTR is a variant human beta-globin 3' UTR. In some embodiments, the 3' UTR comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO: 34

```
CAATTGGCTCGCTTCTTGCTGTCCAATTCTATTAAGGTTCCCT  
TTGTTCCCTAACGTCACACTAAACTGGGGATATTATGAAGGG  
GCCTTGAGCATCTGGATTCTGCCTAATAAAAACATTTATTTCAT  
TGGCAATTTC
```

[0162] In some embodiments, the 5' UTR is a synthetic 5' UTR. In some embodiments, the 5' UTR comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO: 35

```
AGGAAAAATAAGAGAGAAAAAGAAGAGTAAGAAGAAATAAGAGCCA  
CC
```

[0163] In some embodiments, the 5' UTR is a human beta-globin 5' UTR. In some embodiments, the 5' UTR comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO: 36

```
GGACATTGCTCTGACACAACGTGTTCACTAGCAACCTCAAAC  
AACTAGTACACC
```

[0164] In some embodiments, the UTR may be any of, or functional variants of, those described in any of PCT Application No. WO2017053297A1 and U.S. patent Ser. No. 10/519,189B2, both of which are incorporated herein in their entirety.

Exemplary Therapeutic TERT mRNA Sequences

[0165] In some embodiments, a TERT mRNA may refer to the full length mRNA sequence, ie. coding and non-coding, delivered to the tissue, e.g. the liver. Example sequences include the sequences comprising mouse TERT of SEQ ID NOS: 37 and 38, and the sequences comprising human TERT of SEQ ID NOS: 39 and 40.

[0166] In some embodiments, the mouse TERT mRNA comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO: 37

```

CAACCGGTACACCAGACCCGCGCTCCCTCGTTGCCCGCGGTGC
GCTCTCTGCGCAGCGATAACCGGGAGGTGTGGCCCGTGGCAA
CCTTTGTGCGCGCCTGGGGCCGAGGGCAGGCCGCGCTTGCAAC
CCGGGGACCGAAGATCTACCGCATTGGTTGCCAATGCCAG
TGTGCATGCACTGGGCTCACAGCCTCACCTGCCGACCTTCCT
TCCACCAGGTGTCATCCCTGAAAGAGCTGGTGGCCAGGGTTGTG
AGAGACTCTGCCAGGCCAACGAGAGAAACGTGCTGGCTTGGCT
TTGAGCTGCTTAACGAGGGCAGAGGCCGCGCTCCCATGGCTTC
CTAGTAGCGTGCCTAGCTACTGCCAACACTGTTATTGAGACCC
TGCCTGTCAGTGGTGCATGGATGCTACTGTTGAGCCGAGTGGCG
ACGACCTGCTGGTCTACCTGCTGGCACACTGTGCTTTATCTC
TGGTGCCCCCAGCTGTGCCCTACCAAGGTGTGGTCTCCCTGT
ACCAAATTGTGCCACCACGGATATCTGCCCTCTGTGTCGCTA
GTTACAGGCCACCCGACCGTGGCAGGAATTCACTAACCTTA
GGTCTTACAACAGATCAAGAGCAGTAGTCGCCAGGAAGCACCAG
AACCCCTGCCCTGCCATCTGAGGTACAAAGAGGCATGAGCT
TCACCAAGTACAAGTGTGCCCTCAGCTAACAGGCCAGATGCTATC
CTGTCCCGAGAGTGGAGGAGGGACCCCACAGGCAGGTGCTACCA
CCCCATCAGGCAAATCATGGTGCCAAGTCCTGCTCGTCCCCG
AGGTGCCCTACTGCAGAGAAAGATTGTCTCTAAAGGAAAGGTG
CTGACCTGAGTCTCTGGTCGGTGTGCTGTAAACACAAGCCA
GCTCACATCTGCTGTCAACCCCCGCCAAAATGCCCTTCAGC
TCAGGCCATTATTGAGACAGACATTCTTACTCCAGGGAG
ATGGCCAAGAGCGCTCTAACCCCTCATTCCTACTCAGCAACCTCC
AGCCTAACCTGACTGGGCCAGGAGACTGGTGGAGATCATCTTC
TGGGCTCAAGGCCCTAGGACATCAGGACACTCTGCAGGACACCC
GTCTATCGCGTCGATACTGGCAGATGCGGCCCTGTTCCAACAGC
TGCTGGTGAACCATGCAGAGTGCCAATATGTCAGACTCCTCAGG
CACATTGCAGGTTCGAACAGCAAACCAACAGGTGACAGATGCC
TGAACACCAGGCCACCGCACCTCATGGATTGCTCCGCCGTGACA
GCAGTCCCTGGCAGGTATATGGTTCTCGGGCTGTCTGCA
AGGTGGTGTCTGCTAGTCTCTGGGTACCAGGACAATGAGCGCC

```

-continued

```

GCTCTTTAAAGAACCTAAAGAACAGTCATCTCGTTGGGAAATACG
GCAAGCTATCACTGCAGGAACATGATGTTGGAAAGATGAAAGTAGAGG
ATTGCCACTGGCTCCGCAGCAGCCGGGGAGGGACCGTGTCCCCG
CTGCAGAGCACCGCTCTGGAGAGGATCCTGGCTACGTTCTGT
TCTGGCTGATGGACACATACGTGGTACAGCTGCTTAGGTCTTCT
TTTACATCACAGAGAGCACATTCCAGAAGAACAGGCTCTTCTTCT
ACCGTAAGAGTGTGGAGCAAGCTGCAGAGCATTGGAGTCAGGC
AACACCTTGAGAGAGTGCAGGCTACGGAGCTGTACAAGAGGAGG
TCAGGCATCACAGGACACCTGGCTAGCCATGCCATCTGCAGAC
TGCCTTCATCCCCAGCCAAACGGCTGCGGCCATTGTGAACA
TGAGTTATAGCATGGTACAGAGCTTGGCAGAAGGAAGCAGG
CCCAGCATTTCACCCAGCGCTCAAGACTCTTCTAGCATGCTCA
ACTATGAGCGAACAAACATCTCACCTTATGGGTCTTCTGTAC
TGGGTATGAATGACATCTACAGGACTGGCGGGCTTGTGCTG
GTGTGCGTCTGGACCAGACACCCAGGATGTACTTGTAAAGG
CAGATGTGACCGGGCTATGATGCCATCCCCAGGTAAGCTGG
TGGAGGTTGTGCCAATATGATCAGGACTCGGAGAGCACGTA
GTATCCGCCAGTATGCACTGGTCCGGAGAGATGCCAAGGCCAAG
TCCACAAGTCCTTAGGAGACAGGTACCCCTCTGACCTCC
AGCCATACATGGGCCAGTCCTTAAGCATCTGCAGGATTCAAGATG
CCAGTCAGTGGAAACTCCGTTGTCATCGAGCAGACATCTCTA
TGAATGAGAGCACAGCAGCCGCTGTTGACTTCTCTGCACTTCC
TGCCTCACAGTGTGCTAAAGATTGGTACAGGTGCTATCGCAGT
GCCAGGGCATCCCCAGGGCTCCAGCCTATCCACCCCTGCTCTGCA
GTCTGTGTTGGAGACATGGAGAACAGCTGTTGCTGAGGTG
AGCGGGATGGTTGCTTTACGTTGATGACTTCTGTTGG
TGACGCCCTACTGGGCCAAGCAGGAAACCTCCCTCAGCACCCCTG
TCCATGGCTTCCTGAGTATGGTGCATGATAAAACTGAGAACAGA
CAGTGGTAACCTCCCTGTTGGAGGCTGGTACCCCTGGTGGTGCA
CTCCATACAGCTGCCCTGCTACTGCCCTGTTCCCTGGTGTGGCT
TGCTGCTGGACACTCAGACTTGGAGGTGTTGTGACTACTCAG
GTTATGCCAGACCTCAATTAAGACAGGCCCTCACCTCCAGAGTG
TCTCTAACAGCTGGAGAACCATGCCAACAGCTCTGTCGGTCT
TGCGGTTGAAGTGTACGGCTATTTCTAGACTTGCAGGTGAACA
GCCCTCAGACAGTGTGCTCAATATACAAGATCTTCTGCTTC
AGGCCTACAGGTTCCATGCATGTTGATTGAGCTTCCCTTGACCC
AGCGTGTAGGAAAGAACCTCACATTCTTCTGGCTCATCTCCA
GCCAAGCATCTGCTGCTATGCTATCCTGAAGGTCAAGAACATCCAG
GAATGACACTAAAGGCCCTCTGGCTCTTCTCCTGAAGCCGCAC

```

- continued

ATTGGCTCTGCTACCAGGCCCTCGCTAACAGCTGGCTCTCATT
 CTGTATCATACAAATGTCCTCGTGGACCTCTGAGGACAGCCCCAA
 AACTGCTGTGCCGGAGCTCCAGAGGGACAATGACCATCCTTA
 AAGCTGCAGCTGACCCAGGCCCTAACGCACAGACTTCAGACCATT
 TGGACTAACAAATTGGCTCGTTCTTGCTGCTCAATTCTATTAA
 AGGTTCTTTGTTCCCTAACGACTAACTAACTGGGGATAT
 TATGAAGGGCTTGAGCATCTGGATTCTGCTTAATAAAAACATT
 TCTTTCATTCGAATTCAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAA

[0167] In some embodiments, the mouse TERT mRNA comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO: 38

AGGAAATAAGAGAGAAAAAGAGTAAGAAGAAATAAGAGCCA
 CCATGACCGCGCTCTCGTTGCCCGCGGTGCGCTCTGCTC
 GCAGCCGATACCGGGAGGTGTGGCGCTGGCACACCTTGTGCGC
 GCCTGGGGCCGAGGGCAGGGCCTTGTCAACCCGGGACCGA
 AGATCTACCGCACTTGGTGCCAAATGCCATGTGCTGACT
 GGGGCTCACAGCCTCCACCTGCCACCTTCCACCAAGGT
 CATCCCTGAAAGAGCTGGTGGCCAGGGTTGTGAGAGACTCTGCG
 AGCGCAACGAGAGAACGTCGCTGGCTTGGCTTGAGCTGCTTA
 ACGAGGCCAGAGGCAGGGCCTCCATGCCCTCACTAGTAGCGTGC
 GTAGCTACTGCCAACACTGTTATTGAGACCCCTGCGTCAGTG
 GTGCAATGGATGCTACTGTTGAGCCGAGTGGCGACGACCTGCTGG
 TCTACCTGCTGGCACACTGTGCTTTATCTCTGGTGGCCCCCA
 GCTGTGCCCTACCAAGGTGTGGTCTCCCTGTACCAAATTG
 CCACCA CGGATATCTGGCCCTGTGTCGGCTAGTTACAGGCCA
 CCCGACCCGTGGCAGGAATTCAACTAACCTTAGGTTTACAAAC
 AGATCAAGAGCAGTAGTCGCCAGGAAGCACCAGAACCCCTGGC
 TGCCATCTCGAGGTACAAAGAGGCATCTGAGTCTCACCAAGTACAA
 GTGTGCCCTCAGCTAACAGGCCAGATGCTATCCTGTCGGAGAG
 TGGAGGAGGGACCCACAGGCAGGTGCTACCAACCCATCAGGCA
 AATCATGGGTGCCAACGTCCTGCTCGGTCCCCGAGGTGCTACTG
 CAGAGAAAAGATTGCTTCTAAAGGAAAGGTGTGACCTGAGTC
 TCTCTGGTCGGTGTGCTGAAACACAAGCCCAGCTCCACATCTC
 TGCTGTCACCACCCGCCAAATGCCCTTCAGCTCAGGCCATTAA
 TTGAGACCAGACATTCCCTTACTCCAGGGAGATGGCCAAGAGC

- continued

GTCCTAAACCCCTCATTCCTACTCAGCAACCTCCAGCCTAACCTGA
 CTGGGGCCAGGAGACTGGTGGAGATCATCTTCTGGCTCAAGGC
 CTAGGACATCAGGACCACTCTGCAAGGACACACCGCTATCGCGTC
 GATACTGGCAGATGCGGCCCTGTCCAACAGCTGCTGGTGAACC
 ATGCAGAGTGCCAATATGTCAGACTCCTCAGGTACATTGCAAGGT
 TTCGAACAGCAACCCAACAGGTGACAGATGCCTGAACACCAGGC
 CACCGCACCTCATGGATTGCTCCGCTGCACAGCAGTCCTGGC
 AGGTATATGGTTTCTCGGGCTGTCTGCAAGGTGGTGTCTG
 CTAGTCTCTGGGTACAGGACAAATGAGCGCCGTTCTTAAGA
 ACTTAAAGAAGTTCATCTCGTGGGAAATACGGCAAGCTATCAC
 TGCAGGAACGTGATGTGAAGATGAAAGTAGAGGATTGCCACTGGC
 TCCGCAGCAGCCGGGAAGGACCGTGTCCCGCTGCAGAGCACC
 GTCTGAGGGAGAGGATCCTGGCTACGTTCTGTTCTGGCTGATGG
 ACACATACGTGGTACAGCTGCTTAGGTATTCTTACATCACAG
 AGAGCACATTCCAGAAGAACAGGCTCTTCTACCGTAAGAGTG
 TGTGGAGCAAGCTGCAAGCAGCATTGGAGTCAGGAAACACCTTGAGA
 GAGTGCAGCTACGGGAGCTGTCAACAGAGGGTCAAGGCATCACC
 AGGACACCTGGCTAGCCATGCCATGCAAGACTGCGCTTCATCC
 CCAAGCCAAACGCCCTGCCGCCATTGTGAACATGAGTTATAGCA
 TGGGTACAGAGCTTGGCAGAAGGAAGCAGGCCAGCATTCA
 CCCAGCGTCTCAAGACTCTCTCAGCATGCTCAACTATGAGCGGA
 CAAACATCCTCACCTTATGGGTCTCTGTACTGGTATGAATG
 ACATCTACAGGACCTGGCGGGCTTGTGCTGCGTGTGCGTGT
 TGGACCAGACACCCAGGATGTTGTTAAGGCAGATGTGACCG
 GGGCCTATGATGCCATCCCCAGGTAAGCTGGTGGAGGTTGTTG
 CCAATATGATCAGGCACTCGGAGAGCACGTACTGTATCCGCCAGT
 ATGCAGTGGTCCGGAGAGATGCCAAGGCCAAGTCCACAAGTCCT
 TAGGAGACAGGTCAACCAACCTCTGACCTCCAGCCATACATGG
 GCCAGTCCTAACGATCTGCAAGGATTCAAGTGCAGTCAGTGC
 GGAACCTCGTTGTCATCGAGCAGAGCATCTATGAATGAGAGCA
 GCAGCAGCCTGTTGACTTCTCCTGCACTTCCTGCGTACAGTG
 TCGTAAAGATTGGTACAGGTGCTACGCAGTGCCAGGGCATCC
 CCCAGGCCAGCCATACCCCTGCTGCACTGCTGTGTTTC
 GAGACATGGAGAACAGCTGTTGCTGAGGTGCAAGGGATGGGT
 TGCTTTACGTTTGTGATGACTTCTGTTGGTACGCGCTCACT
 TGGACCAAGCAAAACCTCCTCAGCACCCCTGGCCATGGCGTTC
 CTGAGTATGGGTGCACTGATAAACTGCAAGACAGTGGTGAAC
 TCCCTGTTGGAGCCTGGTACCCCTGGTGGTGCAGCTCCATACCAGC
 TGCCTGCTCACTGCCCTGTTCCCTGGTGTGGCTTGCTGCTGGACA

- continued

```

CTCAGACTTGGAGGTGTTCTGTGACTACTCAGTTATGCCAGA
CCTCAATTAAGACGAGCCTACCTCCAGAGTGTCTCAAAGCTG
GGAAGACCATGCGAACAAAGCTCCTGTCGGTCTTGCAGGTGAAGT
GTCACGGTCTATTCTAGACTTGCAAGTGAAACAGCCTCAGACAG
TCTGCATCAATATATAAGATCTCCTGCTTCAGGCCTACAGGT
TCCATGCATGTGATTCAAGCTTCCCTTGGACCGCTGTTAGGA
AGAACCTCACATTCTTCTGGCATCATCCAGCCAAGCATTCT
GCTGCTATGCTATCCTGAAGGTCAAGAATCCAGGAATGACACTAA
AGGCTCTGGCTCTTCTCTGAAGGCCACATTGGCTCTGCT
ACCAGGCCCTCCTGCTCAAGCTGGCTGCTATTGTCTACATACA
AATGTCTCTGGACCTCTGAGGACAGCCAAAAACTGCTGTGCC
GGAAGCTCCCAGAGGCACATGACCATTCTAAAGCTGAGCTG
ACCCAGCCCTAACGACAGACTTCAGACCATTGGACTAAATAT
TAAGCTGCTTCTGGGGCTTGCTTCTGGCATGCCCTCTTC
TCTCCCTTGACCTGTACCTCTTGGTCTTGAATAAAGCCTGAGT
AGGAAGTCTAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
A

```

[0168] In some embodiments, the human TERT mRNA comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO: 39

```

GGGACATTGCTTCTGACACAACGTGTTCACTAGCAACCTCAA
CAACTAGTACACCATGCCGCGCTCCCCGCTGCCGAGCCGTGCG
CTCCCTGCGCGCAGCCACTACCGCGAGGTGCTGCCGCTGGCC
GTTCGTGGCGCCCTGGGCCCCAGGGCTGGCGCTGGTGCAGCG
CGGGGACCCGGCGCTTCCGCGCGCTGGTGGCCAGTGCCTGGT
GTGCGTGCCTGGACGCACGGCCGCCCCCGCCGCCCCCTCCTT
CCGCCAGGTGCTCTGCCCTGAAGGAGCTGGTGGCCGAGTGTGCA
GAGGCTGTGCGAGCGCGCGGAAGAACGTGCTGGCCTCGGCTT
CGCGCTGCGGGACGGGCCCCCGGGGGCCCCCGAGGCGCTCAC
CACCAAGCGTGCAGCTACCTGCCAACACGGTGACCGACGCACT
GCCGGGGAGCGGGGGCGTGGGGCTGCTGCGCGCGTGGCGA
CGACGTGCTGGTTACCTGCTGGCACGCTGCGCTCTTGTGCT
GGTGGCTCCAGCTGCGCTACCAGGTGTCGGGCCGCGCTGTA
CCAGCTCGCGCTGCCACTCAGGCCGGCCCCGCCACACGCTAG
TGGACCCCGAAGGCCTGGGATGCGAACGGCCTGGAACCATAG
CGTCAGGGAGGCCGGGGTCCCCCTGGGCCAGCCCCGGGTGCA
GAGGAGGCCGGGGCAGTGCCAGCCGAAGTCTGCCGTTGCCAA

```

- continued

```

GAGGCCAGGCCTGGCGCTGCCCTGAGCCGGAGCGGACGCCGCT
TGGCAGGGCTCTGGGCCACCCGGCAGGACCGTGGACCGAG
TGACCGTGGTTCTGTGTTGTGACCTGCCAGACGCCGCTGGAC
AGCCACCTTCTGGAGGGTGCCTCTGGCACCGGCCACTCCCA
CCCATCCGTGGCCGCCAGCACGCCGGCCCCCATCCACATC
GCCGCCACACGCCCTGGACACGCCCTGTCCCCGGTGTACCC
CGAGACCAAGCATTCTACTCCTCAGGCACAAGGAGCAGCT
GCCGCCCTCTTCTACTCAGCTCTGAGGCCAGCCTGACTGG
CGCTCGAGGCTCGTGGAGACCATCTTCTGGGTTCCAGGCCCTG
GATGCCAGGGACTCCCCGAGGTTGCCCGCTGCCAGCGCTA
CTGGCAAATGCCGCCCTGTTCTGGAGCTGCTGGAAACCACG
GCAGTGCCCTACGGGGTGCCTCTCAAGACGCACTGCCGCTGCG
AGCTGCGGTACCCAGCAGCGGTGCTGTGCCGGAGAACGCGTA
CCAGGGCTCTGTGGCGCCCCCGAGGAGGAGGACACAGCCCCG
TCGCGCTGGTGCAGCTGCTCCGCCAGCACGAGCCCTGGCAGG
GTACGGCTTGTGCGGGCTGCCCTGCCGCGCCGGCTGGTGCC
CCTCTGGGCTCAGGCCAACGAACGCCGCTTCTCAGGAACAC
CAAGAAGTTCATCTCCCTGGGAAGCATGCCAACGCTCTGCTGCA
GGAGCTGACGTGAAAGATGAGCGTGCAGGACTGCCCTGGCTGCG
CAGGAGCCAGGGTTGGCTGTGTTCCGGCCGAGAGCACCGTCT
GCCGTGAGGAGATCTGGCAAGTCTGCACTGGCTGATGAGTGT
GTACGTGCTGAGCTGCTCAGCTTTCTTATGTACCGAGAC
CACGTTCAAAGAACAGGCTTTTCTACCGAAGAGTGTCTG
GAGCAAGTTGCAAAGCATTGAAATCAGACAGCAGCTGAAGAGGGT
GCAGCTGCGGGAGCTGCGGAAGCAGAGGTGAGCAGCATCGGG
AGCCAGGCCGCCCCCTGCTGACGTTCCAGACTCCGTTCATCCCCA
GCCGTGAGGGCTGCCGCGATTGTGAACATGGACTACGTCGTTGG
AGCCAGAACGTTCCGAGAGAAAAGAGGGCCGAGCGTCTCACCTC
GAGGGTGAAGGACTGTTGAGCGTGTCAACTACGAGCGGGCGCG
GCCCGCCGGCTCTGGGCCCTCTGTGCTGGGCTGGACGATAT
CCACAGGGCTGCCGACCTCTGCTGCGTGTGCGTGGGCCAGGA
CCCGCCGCTGAGCTGACTTGTCAAGGTGGATGTGACGGCGC
GTACGACACCATCCCCAGGACAGGCTACGGAGGTGATGCCAG
CATCATCAAACCCAGAACACGTACTGCGTGCCTGGTATGCCGT
GGTCCAGAAGGCCGCCATGGCAGCTGCCAGGCCCTCAAGAG
CCACGTCTCTACCTGACAGACCTCCAGCGTACATGCGACAGTT
CGTGGCTCACCTGCAAGGAGACCGAGCCGCTGAGGGATGCCGTGCT
CATCGAGCAGAGCTCTCCCTGAATGAGGCCAGCAGTGGCCTCTT
CGACGTCTTACGCTTACGCTGCGACCACGCCGCTGCGCATCAG

```

- continued

GGGCAAGTCCCTACGTCCAGTGCCAGGGATCCCGCAGGGCTCCAT
 CCTCTCCACCGCTGCTCTGCCAGCCTGTGCTACGGCAGATGGAGAA
 CAAGCTGTTGCGGGATTGGCGGGACGGGCTGCTCCTGCCTT
 GGTGGATGATTCTTGTGGTGAACACCTCACCTCACCCACCGCAA
 AACCTCCCTCAGGACCTGGTCCGAGGTGTCCTGAGTATGGCTG
 CGTGGTGAACTTGCGGAAGACAGTGGTGAACCTTCCTGTAGAAGA
 CGAGGCCCTGGGTGGCACGGCTTTGTTAGATGCCGGCCACGG
 CCTATTCCCCCTGGTGCAGGCTGCTGGATAACCGGACCCCTGGA
 GGTGCAGAGCGACTACTCCAGCATGCCGGACCTCCATCAGAGC
 CAGTCTCACCTCAACCGCGCTTCAAGGCTGGAGGAACATGCG
 TCGCAAACCTTTGGGTCTGCGGCTGAAGTGTACAGCCTGTT
 TCTGGATTGCAAGGTGAAACAGCCTCCAGACGGTGTGCACCAACAT
 CTACAAGATCCTCCTGCTGCAGGCGTACAGGTTACCGATGTGT
 GCTGCAGCTCCATTTCATCAGCAAGTTGAAAGAACCCACATT
 TTTCCTGCGCGTCATCTGACACGGCCTCCCTGCTACTCCAT
 CCTGAAAGCCAAGAACGCAGGGATGTCGCTGGGGCCAAGGGCGC
 CGCCGGCCCTCTGCCCTCCGAGGGCGTGAGTGGCTGCCACCA
 AGCATTCTGCTCAAGCTGACTCGACACCGTGTACCTACGTGCC
 ACTCCTGGGTCACTCAGGACAGCCCAGCAGCTGAGTCGGAA
 GCTCCGGGGACGACGCTGACTGCCCTGGAGGCCAGCCAACCC
 GGCACTGCCCTCAGACTTAAGACCATCTGGACTGACAATTGGC
 TCGTTTCTGCTGTCATTCTATTAAAGTTCTTTGTTCC
 CTAAGTCCAACACTAAACTGGGGATATTATGAAGGGCCTTGAG
 CATCTGGATTCTGCTTAATAAAAACATTCTTTCATTCGAAT
 TCAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAA

[0169] In some embodiments, the human TERT mRNA comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to SEQ ID NO: 40

AGGAAAATAAGAGAGAAAAGAGATAAGAAGAAATAAGAGCCA
 CCATGCCGCGCCTCCCGCTGCCAGGCCGTGCGCTCCCTGCTGC
 GCAGCCACTACCGCGAGGTGCTGCCGCTGGCACGTTCTGCC
 GCCTGGGGCCCCAGGGCTGGCGCTGGTCAGCGCGGGACCCGG
 CGGCTTCCGCGCCTGGTGGCCAGTGCCTGGTGTGCGTGCCT
 GGGACGCACGGCCGCCCGGCCGCCCCCTCCCTGCCAGGTGT
 CCTGCCTGAAGGAGCTGGTGGCCGAGTGCCTGAGAGGCTGTGCG

- continued

AGCGCGCGCGAAGAACGTGCTGCCCTCGCCTCGCGCTGCTGG
 ACGGGGCCGCGGGGGCCCCCGAGGCCCTCACACCAGCGTC
 GCAGCTACCTGCCAACACGGTGACCGACGCAGTGGGGAGCG
 GGGCGTGGGGCTGCTGCTGCCGCGTGGCGACGACGTGCTGG
 TTCACCTGCTGGCACGCTGCCGCTCTTGTGCTGGCTGGCTCC
 GCTGCCCTACCAAGGTGTCGCCGCCGCGCTGACAGCTGGCG
 CTGCCACTCAGGCCGGCCCCGCCACACGCTAGTGGACCCCGAA
 GCGCTCTGGATGCGAACGGCTGGAACCATAGCGTCAGGGAGG
 CGGGGCTCCCCCTGGGCTGCCAGGCCGGGTGCGAGGAGCGCG
 GGGCAGTGCAGCCGAAGTGTGCCGTTGCCAAGGCCAGGCC
 GTGGCGCTGCCCTGAGCCGGAGCGGCCGTTGGCAGGG
 CCTGGGCCACCCGGCAGGACGCGTGGACCGAGTGACCGTGGTT
 TCTGTGTTGTCACCTGCCAGACCCGCCGAAGAAGCCACCTCTT
 TGGAGGGTGCCTCTGGCACGCCACTCCCACCCATCCG
 GCCGCCAGCACGCCGGCCCCATCCACATGCCGGCCACCAC
 GTCCCTGGACACGCCCTGTCCTGGGAGACGCCAGGCC
 ACTCCTCTACTCCCTGCCAGCGACAAGGAGCAGCTGCCCTCC
 TCCTACTCAGCTCTGAGGCCAGCCTGACTGCCGCTGGAGGC
 TCGTGGAGACCATTTCTGGGTTCCAGGCCCTGGATGCCAGGG
 CTCCCGCAGGTTGCCCGCCTGCCAGCGCTACTGCCAATGC
 GGCCCTGTTCTGGAGCTGCTGGAAACACGCCAGCGCAGTGG
 ACGGGGTGCTCCTCAAGACCGACTGCCGCTGCCAGCTGCC
 CCCAGCAGCCGGTGTCTGCCGGAGAAGGCCAGGGCTCTG
 TGGCGGCCGGAGGGAGGACACAGACCCCGTCGCTGGTGC
 AGCTGCTCCGCCAGCACAGCAGCCCTGGCAGGTGTACGCC
 TGCGGCCCTGCCCTGCCGGCTGGTGCCTGGAGGCCCTGG
 CCAGGCACACGAACGCCCTGCCAGAACACCAAGAAGTTCA
 TCTCCCTGGGAAGCATGCCAGCTCGCTGCAGGAGCTGACGT
 GGAAGATGAGCGTGCAGGACTGCCCTGGCTGCCAGGCCAG
 GGGTTGGCTGTGTCAGGCCAGGCCAGGCCAGCGTGCAGGAGA
 TCCCTGGCCAAGTCCCTGCACTGGCTGATGAGTGTACGCGT
 AGCTGCTCAGGTCTTCTTTATGTCACGGAGACACGTTCAA
 AGAACAGGGCTTTCTACCGGAAGAGTGTCTGGAGCAAGTT
 AAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCC
 AGCTGTCGGAGACAGGGTCAGGCCAGCATGCCAGGCCAGGCC
 CCCCTGCTGACGTCAGACTGCCCTCACCCCAAGCCTGACGCC
 TGCGGCCGATTGTAACATGGACTACGTCGTGGAGGCCAGAC
 TCCGCAGAGAAAAGAGGGCCGAGCGTCTACCTCGAGGGTGAAG
 CACTGTTCAGCGTGCTCAACTACGAGCGGGCGGGGCC
 CG

- continued

```

TCCTGGGCCTCTGTGCTGGCCTGGACGATATCCACAGGGCCT
GGCGCACCTTGTGCTGCGTGTGCGGGCCCAAGGACCCGGCGCTG
AGCTGTACTTGTCAGGTGGATGTGACGGGCGCGTACGACACCA
TCCCCCAGGACAGGCTCACGGAGGTACATGCCAGCATCATCAAAC
CCCAGAACACGTACTGCGTGCCTGGTATGCCGTGGTCCAGAAGG
CCGCCCATTGGCACGTCCGAAGGCCTTCAAGAGGCCACGCTCTCA
CCTTGACAGACCTCCAGCGTACATGCGACAGTCGTGGCTCACC
TGCAGGAGACCAGCCGCTGAGGGATGCCGTGTCATCGAGCAGA
GCTCCCTCTGAATGAGGGCAGCAGTGGCTTCTCGACGCTCTCC
TACGCTTCATGTGCCACACGCCGTGCGCATCAGGGCAAGTCCT
ACGTCCAGTGCCAGGGATCCCGCAGGGCTCCATCTCCACGC
TGCTCTGCAGCCTGTGCTACGGCAGATGGAGAACAGCTGTTG
CGGGGATTGCGCGGACGGGCTGCTCCCTGCGTTGGATGATT
TCTTGTGGTACACCTCACCTCACCCACGCCAAACCTCTCA
GGACCCCTGGTCCGAGGTGTCCTGAGTATGGCTGCGTGGTGAAC
TGCAGAACAGTGGTGAACCTCCCTGAGAACAGCAGGCCCTGG
GTGGCACGGCTTTGTCAGATGCCGCCACGGCTATTCCCT
GGTGCAGGCCCTGCTGGATACCCGGACCCCTGGAGGTGAGAGCG
ACTACTCCAGCTATGCCGCCACCTCATCAGAGCAGTCACCT
TCAACCGGGCTTCAAGGCTGGGAGGAACATGCGTCGAAACTCT
TTGGGGTCTTGCAGGCTGAAGTGTACAGCTGTTCTGGATTG
AGGTGAACAGCCTCCAGACGGTGTGACCAACATCTAACAGATCC
TCCCTGTCAGGCGTACAGGTTTACCGCATGTTGCTGAGCTCC
CATTCATCAGCAAGTTGAGAACCCCCACATTTCTGCGCG
TCATCTCTGACACGCCCTCCCTGCTACTCCATCTGAAAGCCA
AGAACGCAGGGATGTCGCTGGGGCCAAGGGCGCCGCCCTC
TGCCCTCCAGGCCGTGCACTGGCTGTGCCACCAAGCATTCTGC
TCAAGCTGACTCGACACCGTGTACCTACGTGCCACTCTGGGT
CACTCAGGACAGGCCAGACGCAGCTGAGTCGGAAAGCTCCGGGA
CGACGCTGACTGCCCTGGAGGCCAGCCAACCCGGACTGCCCT
CAGACTCAAGACCATCTGGACTGATAATTAAGCTGCCCTCTG
GGGGCTTGCCTCTGGCATGCCCTCTCTCCCTTGACCTG
TACCTCTGGTCTTGAATAAGCCTGAGTAGGAAGAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAA

```

[0170] In some embodiments, a TERT mRNA may comprise a nucleic acid sequence at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%,

at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% identical to any one of SEQ ID NOS: 38-40.

[0171] The disclosure provides compositions for the extension of telomeres in a cell, the compositions comprising a compound of the present disclosure, as described above, and a further component. In some embodiments, the further component comprises a telomerase RNA component (TERC). In some embodiments, the compositions further comprise a telomerase RNA component (TERC). In some embodiments, the compositions further comprise one or more additional components that may facilitate delivery of the RNA to cells *in vitro* and/or *in vivo*. In some embodiments, the one or more additional components comprise a nanoparticle. In some embodiments, the nanoparticle comprises a lipid. In some embodiments, the nanoparticle or the lipid comprise a coatsome-like lipid or a compound of the disclosure. In some embodiments, the nanoparticle or the lipid comprise a compound of the disclosure according to Formula I.

II. Delivery Vehicles

[0172] In some embodiments, one or more mRNAs may be delivered to a cell or tissue via delivery vehicles. In some embodiments a delivery vehicle may be a nanoparticle. In some embodiments, the delivery vehicle is a lipid nanoparticle (LNP) including but not limited to a nanoparticle comprising lipids and/or polymers, a liposome, a liposomal nanoparticle, a cationic lipid, or an exosome. As used herein, liposomal nanoparticles may be characterized as microscopic vesicles having an interior aqueous space sequestered from an outer medium by a membrane of one or more bilayers.

[0173] In some embodiments, the nanoparticle is a polymeric nanoparticle. In some embodiments, the nanoparticle is a metal nanoparticle. In other embodiments, the delivery vehicle comprises or consists of a recombinant virus or virus-like particle, e.g., an adenovirus, adeno-associated virus (AAV), herpesvirus, or retrovirus, e.g., lentivirus. In some embodiments, the delivery vehicle comprises or consists of a modified viral vector, e.g., an adenovirus dodecahedron or recombinant adenovirus conglomerate. In other embodiments, the delivery vehicle may comprise or consist of calcium phosphate nucleotides, aptamers, cell-penetrating peptides or other vectorial tags.

A. Liposomal Delivery Vehicles

[0174] In some embodiments, a suitable delivery vehicle is a lipid nanoparticle (LNP). Exemplary LNPs may comprise one or more different lipids and/or polymers. In some embodiments, an LNP comprises one or more of ionizable lipids, neutral lipids, cholesterol, and/or stabilizing lipids (e.g., PEGylated lipids).

Ionizable Lipids

[0175] In some embodiments, an LNP may comprise an ionizable lipid. An ionizable lipid may refer to any of a number of lipid species that have a net positive charge at a selected pH, such as a physiological pH. An ionizable lipid may also, for example, refer to a lipid in an ionized state, e.g., a cationic lipid. In some embodiments, an LNP may comprise an ionizable lipid as disclosed in either of WO

2010/053572 or WO 2012/170930, or variations thereof, both of which are incorporated herein by reference in their entirety.

[0176] In some embodiments, an LNP for liver delivery of a TERT mRNA may comprise one or more of MC3 (((6Z, 9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate), 1,2-dilinoleyl-3-dimethylammonium-propane (DLinDAP), DLin-MC3-DMA 4-(dimethylamino)-butanoic acid, (10Z,13Z)-1-(9Z,12Z)-9,12-octadecadien-1-yl-10,13-nonadecadien-1-yl ester and/or cKK-E12 3,6-Bis(4-(bis(2-hydroxydodecyl)amino)butyl) piperazine-2,5-dione. In some embodiments the LNP comprises 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (Dlin-KC2-DMA, 1) and/or (6Z,9Z,28Z,31 Z)-Heptatriaconta-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate. In some embodiments, the ionizable lipid may have a pKa range of 6.1-6.7, optionally a pKa range of 6.2-6.5.

[0177] In some embodiments, an LNP comprises 1,2-dioleoyl-3-trimethylammonium-propane (DOTAP), N,N-distearyl-N,N-dimethylammonium bromide (DABB), or 1,2-dimyristoyl-sn-glycero-3-ethylphosphocholine (EPC). In some embodiments, an LNP comprises a ionizable lipid wherein the ionizable lipid is one or more of N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethylammonium chloride (DOTMA), 5-carboxyspermylglycinedioctadecylamide (DOGS), 2,3-dioleyloxy-N-[2(spermine-carboxamido)ethyl]-N,N-dimethyl-1-propanaminium (DOSPA), 1,2-Dioleoyl-3-Dimethylammonium-Propane (DODAP), and/or 1,2-Dioleoyl-3-Trimethylammonium-Propane (DOTAP), or variations thereof. An LNP may also comprise one or more of 1,2-distearyoxy-N,N-dimethyl-3-aminopropane (DSDMA), 1,2-dioleyloxy-N,N-dimethyl-3-aminopropane (DODMA), 1,2-dilinoleyoxy-N,N-dimethyl-3-aminopropane (DLinDMA), 1,2-dilinoleyoxy-N,N-dimethyl-3-aminopropane or (DLenDMA), 4-(dimethylamino)-butanoic acid, (10Z,13Z)-1-(9Z,12Z)-9,12-octadecadien-1-yl-10,13-nonadecadien-1-yl ester (DLin-MC3-DMA), N-dioleyl-N,N-dimethylammonium chloride (DODAC), or variations thereof. In other embodiments, an LNP may comprise a ionizable lipid of XTC (2,2-Dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane), MC3 (((6Z,9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate), ALNY-100 ((3aR,5s,6s)-N,N-dimethyl-2,2-di((9Z,12Z)-octadeca-9,12-dienyl)tetrahydro-3aH-cyclopenta[d][1,3]dioxol-5-amine)), NC98-5 (4,7,13-tris(3-oxo-3-(undecylamino)propyl)-N1,N16-diundecyl-4,7,10,13-tetraazahexadecane-1,16-diamide), or variations thereof.

[0178] In some embodiments, an LNP may comprise an ionizable lipid, e.g., one or more of (15Z,18Z)-N,N-dimethyl-6-(9Z,12Z)-octadeca-9,12-dien-1-yl) tetracosa-15,18-dien-1-amine, (15Z,18Z)-N,N-dimethyl-6-((9Z,12Z)-octadeca-9,12-dien-1-yl) tetracosa-4,15,18-trien-1-amine, and (15Z,18Z)-N,N-dimethyl-6-((9Z,12Z)-octadeca-9, and 12-dien-1-yl) tetracosa-5, 15,18-trien-1-amine (HGT5002).

[0179] In some embodiments, an LNP may comprise a cleavable ionizable lipid comprising a disulfide bond, e.g., COATSOME™ SS-OP, i.e. SS-OP™, COATSOME™ SS-M, COATSOME™ SS-E, COATSOME™ SS-EC, COATSOME™ SS-LC, COATSOME™ SS-OC and variations thereof. In some embodiments, an LNP may comprise about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, or about 90% ionizable lipids relative to the other lipids.

Other “Helper” Phospholipids

[0180] In some embodiments, an LNP may comprise additional lipids selected from one more of: distearoylphosphatidylcholine (DSPC), dioleoylphosphatidylcholine (DOPC), dipalmitoylphosphatidylcholine (DPPC), dioleoylphosphatidylglycerol (DOPG), dipalmitoylphosphatidylglycerol (DPPG), dioleoylphosphatidylethanolamine (DOPE), palmitoyloleoylphosphatidylcholine (POPC), palmitoyloleoyl-phosphatidylethanolamine (POPE), dioleoyl-phosphatidylethanolamine 4-(N-maleimidomethyl)-cyclohexane-1-carboxylate (DOPE-mal), dipalmitoyl phosphatidyl ethanolamine (DPPE), dimyristoylphosphoethanolamine (DMPE), distearoyl-phosphatidyl-ethanolamine (DSPE), 16-O-monomethyl PE, 16-O-dimethyl PE, 18-1-trans PE, 1-stearoyl-2-oleoyl-phosphatidylethanolamine (SOPE), or variants thereof. In some embodiments, an LNP may include one or more phosphatidyl lipids, for example, the phosphatidyl compounds (e.g., phosphatidylglycerol, phosphatidylcholine, phosphatidylserine and phosphatidylethanolamine). In some embodiments, an LNP may comprise sphingolipids, for example but not limited to, sphingosine, ceramide, sphingomyelin, cerebroside and ganglioside. In some embodiments, the aforementioned “helper” lipids contribute to the stability and/or specificity of the LNP composition.

Cholesterol-Based Lipids

[0181] In some embodiments, an LNP may comprise one or more cholesterol-based lipids. A cholesterol-based lipid may include but is not limited to: PEGylated cholesterol, DC-Choi (N,N-dimethyl-N-ethylcarboxamidocholesterol), 1,4-bis(3-N-oleylamino-propyl)piperazine. In some embodiments, an LNP may comprise about 2% to about 30%, or about 5% to about 20% of cholesterol relative to the total lipid present.

PEGylated Lipids

[0182] Without wishing to be bound by theory, it is contemplated that the addition of a lipid modified with an insulating molecule such as a protein or other polymer such as polyethylene-glycol (PEG), also known as a PEGylated lipid, may prevent complex aggregation and increase circulation lifetime to facilitate the delivery of the liposome encapsulated mRNA to the target cell. In some embodiments, the addition of a PEGylated lipid protects the LNP from immune targeting. In some embodiments, the PEGylated lipid forms a hydrophilic barrier around the hydrophobic LNP, preventing opsonization of plasma proteins and bypassing macrophage uptake. In some embodiments, lipids modified with other hydrophilic molecules may be substituted for PEGylated lipids in an LNP delivery vehicle.

[0183] In some embodiments of the disclosure, an LNP may comprise one or more PEGylated lipids. For example, the use of polyethylene glycol (PEG)-modified phospholipids and derivatized lipids such as derivatized ceramides (PEG-CER), including N-Octanoyl-Sphingosine-1-{Succinyl}(Methoxy Polyethylene Glycol)-2000](C8 PEG-2000 ceramide) is contemplated by the present disclosure in combination with one or more of the ionizable and/or other lipids. In some embodiments, PEGylated lipids comprise PEG-ceramides having shorter acyl chains (e.g., C14 or C18). In some embodiments, the PEGylated lipid DSPE-PEG-Maleimide-Lectin may be used. Other contemplated

PEG-modified lipids include, but are not limited to, a polyethylene glycol chain of up to 5 kDa in length covalently attached to a lipid with alkyl chain(s) of C6-C20 length. Without wishing to be bound by a particular theory, it is contemplated that the addition of PEGylated lipids may prevent complex aggregation and increase circulation lifetime to facilitate the delivery of the liposome encapsulated mRNA to the target cell.

[0184] In some embodiments, PEGylated lipids may comprise about 0% to about 20%, about 0% to about 15%, about 0% to about 10%, about 1% to about 10%, about 1% to about 8%, 1% to about 6%, 1% to about 5%, about 2% to about 10%, about 4% to about 10%, of the total lipids present in the liposome by molar ratio. In some embodiments, the percentage of PEGylated lipids may be less than about 20%, about 15%, about 10%, about 9%, about 8%, about 7%, about 6%, about 5%, about 4%, about 3%, about 2%, or about 1% of the total lipids present in the liposome by molar ratio. In some embodiments, the percentage of PEGylated lipids may be greater than about 1%, about 2%, about 3%, about 4%, about 5%, about 6%, about 7%, about 8%, about 9%, about 10%, about 15%, or about 20% of the total lipids present in the liposome by molar ratio.

[0185] In some embodiments, a lipid nanoparticle formulation may comprise, consist essentially of or consist of any of those described in U.S. Pat. Nos. 11,185,595; 9,868,693; 10,195,156; 9,877,919; 9,738,593; 10,399,937; 10,106,490; 9,738,593; 10,821,186; or 8,058,069, each of which is incorporated by reference herein in its entirety; or described in U.S. Patent Application Publication Nos. US20180085474A1, US20210259980A1, US20200206362A1, US20210267895A1, US20200283372A1, or US20200163878A1, each of which is incorporated by reference herein in its entirety.

Lipid Nanoparticle (LNP) Compositions

[0186] The following example LNP formulations are not intended to be limiting.

[0187] In some embodiments, an LNP may comprise a molar ratio of about 35, about 40, about 45, about 50, about 55, about 60, about 65, about 70, or about 75 moles of an ionizable lipid. In some embodiments, an LNP may comprise a molar ratio of about 0.1, about 1.0, about 2.0, about 3.0, about 4.0, about 5.0, about 6.0, about 7.0, about 8.0, about 10, about 12, about 14, about 16, about 18, about 20, about 25, about 30, about 40, or about 50 moles of another phospholipid. In some embodiments, an LNP may comprise a molar ratio of about 10, about 15, about 20, about 25, about 30, about 35, about 40, about 45, about 50, about 55, about 60, about 65, or about 70 moles of cholesterol. In some embodiments, an LNP may comprise a molar ratio of about 0.1, about 0.25, about 0.5, about 0.75, about 1.0, about 1.25, about 1.5, about 1.75, about 2.0, about 2.5, about 3.0, about 3.5, about 4.0, about 4.5 or about 5.0 moles of a PEGylated lipid.

[0188] In some embodiments, an LNP comprises a molar ratio of about 40-70 moles of an ionizable lipid to about 0.1 to about 20 moles of another phospholipid, about 20 to about 60 moles of cholesterol, and about 0.1 to about 5 moles of PEGylated lipid. In some embodiments, the LNP delivery vehicle comprises a molar ratio of about 50-60 moles of an ionizable lipid to about 4-18 moles of another phospholipid, about 35-50 moles of cholesterol, and about 1-3 moles of PEGylated lipid.

[0189] In some embodiments, an LNP may comprise a molar ratio of about 50 to about 60 moles of an ionizable lipid, about 4 to about 6 moles of a phospholipid, about 35 to about 45 moles of cholesterol, and about 1 to about 2 moles of PEGylated lipid.

[0190] In some embodiments, an LNP may comprise a molar ratio of about 30 to 40 moles of an ionizable lipid, about 14 to about 18 moles of a phospholipid, about 40 to about 50 moles of a cholesterol, and about 2.0 to about 3.0 moles of a PEGylated lipid.

[0191] In some embodiments, an LNP may comprise the ionizable lipid SS-OPTTM, the phospholipid DOPC, a cholesterol lipid, and the PEGylated lipid DMG-PEG2000. In some embodiments, an LNP may comprise a molar ratio of 55 moles of SS-OPTTM, to 5 moles of DOPC, 40 moles of a cholesterol lipid, and 1.5 moles of PEGylated lipid DMG-PEG2000. In some embodiments, an LNP may comprise a molar ratio of 52.5 moles of SS-OPTTM, to 7.5 moles of DOPC, 40 moles of a cholesterol lipid, and 1.5 moles of PEGylated lipid DMG-PEG2000.

[0192] In some embodiments, an LNP may comprise the ionizable lipid cKK-E12, the phospholipid DOPE, a cholesterol lipid, and the PEGylated lipid 14:0 PEG2000 PE. In some embodiments, an LNP may comprise a molar ratio of about 35 moles of cKK-E12, to about 16 moles of DOPE, about 46.5 moles of a cholesterol lipid, and about 2.5 moles of PEGylated lipid 14:0 PEG2000 PE.

[0193] In some embodiments, an LNP may comprise the ionizable lipid DLin-MC3-DMA, the phospholipid DSPC, a cholesterol lipid, and the PEGylated lipid DMG-PEG2000. In some embodiments, an LNP may comprise a molar ratio of about 50 moles of DLin-MC3-DMA, about 10 moles of the phospholipid DSPC, about 40 moles of a cholesterol lipid, and about 1.5 moles of the PEGylated lipid DMG-PEG2000.

[0194] In some embodiments, an LNP may comprise the ionizable lipid SS-OPTTM, the phospholipid DOPE, a cholesterol lipid, and the PEGylated lipid 14:0 PEG2000 PE. In some embodiments, an LNP may comprise a molar ratio of about 35 moles of SS-OPTTM, to about 16 moles of DOPE, about 46.5 moles of a cholesterol lipid, and about 2.5 moles of PEGylated lipid 14:0 PEG2000 PE.

[0195] In some embodiments, an LNP may comprise the ionizable lipid cKK-E12, the phospholipid DOPC, a cholesterol lipid, and the PEGylated lipid DMG-PEG2000. In some embodiments, an LNP may comprise a molar ratio of about 55 moles of cKK-E12, to about 5 moles of DOPC, about 40 moles of a cholesterol lipid, and about 1.5 moles of PEGylated lipid DMG-PEG2000.

B. Polymer Nanoparticles

[0196] In some embodiments, a suitable delivery vehicle is formulated using a polymer as a carrier, alone or in combination with other carriers including various lipids described herein. Thus, in some embodiments, liposomal delivery vehicles, as used herein, also encompass polymer containing nanoparticles. Suitable polymers may include, for example, polyacrylates, polyalkylenoacrylates, polylactide, polylactide-polyglycolide copolymers, polycaprolactones, dextran, albumin, gelatin, alginate, collagen, chitosan, cyclodextrins, protamine, polyethylene glycol (PEG)-modified (PEGylated) protamine, poly-D-lysine (PLL), PEGylated PLL and polyethylenimine (PEI). When PEI is present, it may be linear or branched PEI of a molecular

weight ranging from 10 to 40 kDa, e.g., 25 kDa branched PEI (Sigma #408727). In some embodiments the PEGylated lipid is 14:0 PEG2000 PE and/or DMG-PEG2000.

C. Delivery Vehicles Targeting Liver

[0197] In some embodiments, delivery vehicles disclosed herein preferentially target specific organs, e.g., the liver. In various embodiments, the delivery vehicles may delivery mRNA to liver cells 10, 10², 10³, 10⁴, 10⁵, 10⁶, 10⁷, 10⁸, 10⁹, or 10¹⁰-fold more effectively compared a reference cell type (e.g., lung cells). However, it will be understood that some level of delivery to non-target cells/organs may be tolerated without decreasing the effectiveness in the target organ/cell. In some embodiments, the lipid composition of a delivery vehicle enhances delivery to the liver relative to other lipid compositions known in the art. In other embodiments, the lipid composition of a delivery vehicle enhances delivery to the liver relative to other lipid compositions. In some embodiments, the presence or level of cholesterol enhances delivery of a delivery vehicle, e.g. an LNP or extracellular vesicle to the liver.

[0198] In some embodiments, a delivery vehicle, e.g. an LNP, targeting the liver may comprise a molar ratio of about 35, about 40, about 45, about 50, about 55, about 60, about 65, about 70, or about 75 moles of an ionizable lipid. In some embodiments, an LNP may comprise a molar ratio of about 0.1, about 1.0, about 2.0, about 3.0, about 4.0, about 5.0, about 6.0, about 7.0, about 8.0, about 10, about 12, about 14, about 16, about 18, about 20, about 25, about 30, about 40, or about 50 moles of another phospholipid. In some embodiments, an LNP may comprise a molar ratio of about 10, about 15, about 20, about 25, about 30, about 35, about 40, about 45, about 50, about 55, about 60, about 65, or about 70 moles of cholesterol. In some embodiments, an LNP may comprise a molar ratio of about 0.1, about 0.25, about 0.5, about 0.75, about 1.0, about 1.25, about 1.5, about 1.75, about 2.0, about 2.5, about 3.0, about 3.5, about 4.0, about 4.5 or about 5.0 moles of a PEGylated lipid.

[0199] In some embodiments, a delivery vehicle, e.g. an LNP, targeting the liver comprises a molar ratio of about 40-70 moles of an ionizable lipid to about 0.1 to about 20 moles of another phospholipid, about 20 to about 60 moles of cholesterol, and about 0.1 to about 5 moles of PEGylated lipid. In some embodiments, the LNP delivery vehicle comprises a molar ratio of about 50-60 moles of an ionizable lipid to about 4-18 moles of another phospholipid, about 35-50 moles of cholesterol, and about 1-3 moles of PEGylated lipid.

[0200] In some embodiments, a delivery vehicle, e.g. an LNP, targeting the liver may comprise a molar ratio of about 50 to about 60 moles of an ionizable lipid, about 4 to about 6 moles of a phospholipid, about 35 to about 45 moles of cholesterol, and about 1 to about 2 moles of PEGylated lipid.

[0201] In some embodiments, a delivery vehicle, e.g. an LNP, targeting the liver may comprise a molar ratio of about 30 to 40 moles of an ionizable lipid, about 14 to about 18 moles of a phospholipid, about 40 to about 50 moles of a cholesterol, and about 2.0 to about 3.0 moles of a PEGylated lipid.

[0202] In some embodiments, a delivery vehicle, e.g. an LNP, targeting the liver may comprise the ionizable lipid SS-OPT™, the phospholipid DOPC, a cholesterol lipid, and the PEGylated lipid DMG-PEG2000. In some embodiments, an LNP may comprise a molar ratio of 55 moles of SS-OPT™,

to 5 moles of DOPC, 40 moles of a cholesterol lipid, and 1.5 moles of PEGylated lipid DMG-PEG2000.

[0203] In some embodiments, a delivery vehicle, e.g. an LNP, targeting the liver may comprise the ionizable lipid cKK-E12, the phospholipid DOPE, a cholesterol lipid, and the PEGylated lipid 14:0 PEG2000 PE. In some embodiments, an LNP may comprise a molar ratio of about 35 moles of cKK-E12, to about 16 moles of DOPE, about 46.5 moles of a cholesterol lipid, and about 2.5 moles of PEGylated lipid 14:0 PEG2000 PE.

[0204] In some embodiments, a delivery vehicle, e.g. an LNP, targeting the liver may comprise the ionizable lipid DLin-MC3-DMA, the phospholipid DSPC, a cholesterol lipid, and the PEGylated lipid DMG-PEG2000. In some embodiments, an LNP may comprise a molar ratio of about 50 moles of DLin-MC3-DMA, about 10 moles of the phospholipid DSPC, about 40 moles of a cholesterol lipid, and about 1.5 moles of the PEGylated lipid DMG-PEG2000.

[0205] In some embodiments, a delivery vehicle, e.g. an LNP, targeting the liver may comprise the ionizable lipid SS-OPT™, the phospholipid DOPE, a cholesterol lipid, and the PEGylated lipid 14:0 PEG2000 PE. In some embodiments, an LNP may comprise a molar ratio of about 35 moles of SS-OPT™, to about 16 moles of DOPE, about 46.5 moles of a cholesterol lipid, and about 2.5 moles of PEGylated lipid 14:0 PEG2000 PE.

[0206] In some embodiments, a delivery vehicle, e.g. an LNP, targeting the liver may comprise the ionizable lipid cKK-E12, the phospholipid DOPC, a cholesterol lipid, and the PEGylated lipid DMG-PEG2000. In some embodiments, an LNP may comprise a molar ratio of about 55 moles of cKK-E12, to about 5 moles of DOPC, about 40 moles of a cholesterol lipid, and about 1.5 moles of PEGylated lipid DMG-PEG2000.

[0207] In some embodiments, a delivery vehicle comprises an organ-specific targeting ligand to enhance delivery to a particular organ, e.g. the liver. Ligands may include but are not limited to proteins (e.g., human serum albumin HSA), low-density lipoprotein (LDL), or globulin); carbohydrates (e.g., a dextran, pullulan, chitin, chitosan, inulin, cyclodextrin, N-acetylgalactosamine, or hyaluronic acid); or lipids. The ligand may also be a recombinant or synthetic molecule, such as a synthetic polymer, e.g., a synthetic polyamino acid. Examples of polyamino acids include polyamino acid is a polylysine (PLL), poly L-aspartic acid, poly L-glutamic acid, styrene maleic acid anhydride copolymer, poly(L-lactide-co-glycolide) copolymer, divinyl ether maleic anhydride copolymer, N-(2-hydroxypropyl) methacrylamide copolymer (HMPA), polyethylene glycol (PEG), polyvinyl alcohol (PVA), polyurethane, poly(2-ethylacrylic acid), N-isopropylacrylamide polymers, or polyphosphazine. Examples of polyamines include: polyethyl-enimine, polylysine (PLL), spermine, spermidine, polyamine, pseudopeptide-polyamine, peptidomimetic polyamine, dendrimer polyamine, arginine, amidine, protamine, cationic lipid, cationic porphyrin, quaternary salt of a polyamine, or an alpha helical peptide. Ligands can also include targeting groups, e.g., a cell or tissue targeting agent, e.g., a lectin, glycoprotein, lipid or protein, e.g., an antibody, that binds to a specified cell type such as a kidney cell. A targeting group can be a thyrotropin, melanotropin, lectin, glycoprotein, surfactant protein A, Mucin carbohydrate, multivalent lactose, multivalent galactose, N acetyl-galactosamine, N-acetyl-glucosamine multivalent mannose, mul-

tivalent fucose, glycosylated polyaminoacids, multivalent galactose, transferrin, bisphosphonate, polyglutamate, polyaspartate, a lipid, cholesterol, a steroid, bile acid, folate, vitamin B12, vitamin A, biotin, or an RGD peptide or RGD peptide mimetic. Other examples of ligands include dyes, intercalating agents (e.g. acridines), cross linkers (e.g. psoralene, mitomycin C), porphyrins (TPPC4, texaphyrin, Sapphyrin), polycyclic aromatic hydrocarbons (e.g., phenazine, dihydrophenazine), artificial endonucleases (e.g. EDTA), lipophilic molecules, e.g., cholesterol, cholic acid, adamantane acetic acid, 1-pyrene butyric acid, dihydrotestosterone, 1,3-Bis-O(hexadecyl)glycerol, geranyloxyhexyl group, hexadecylglycerol, bomeol, menthol, 1,3-propanediol, heptadecyl group, palmitic acid, myristic acid, 03-(oleoyl)lithocholic acid, 03-(oleoyl)cholenic acid, dimethoxytrityl, or phenoxyazine) and peptide conjugates (e.g., antennapedia peptide, Tat peptide), alkylating agents, phosphate, amino, mercapto, PEG (e.g., PEG-40K), MPEG, [MPEG]2, polyamino, alkyl, substituted alkyl, radiolabeled markers, enzymes, haptens (e.g. biotin), transport/absorption facilitators (e.g., aspirin, vitamin E, folic acid), synthetic ribonucleases (e.g., imidazole, bisimidazole, histamine, imidazole clusters, acridine-imidazole conjugates, Eu3+ complexes of tetraazamacrocycles), dinitrophenyl, HRP, or AP.

[0208] In some embodiments, the organ targeting ligands are proteins, e.g., glycoproteins, or peptides, e.g., molecules having a specific affinity for a co-ligand, or antibodies e.g., an antibody, that binds to a specified cell type, e.g., a liver cell. In some embodiments, the ligands may be hormones or hormone receptors. Ligands may also be non-peptidic species, such as lipids, lectins, carbohydrates, vitamins, cofactors, multivalent lactose, multivalent galactose, N-acetyl galactosamine, N-acetyl glucosamine multivalent mannose, or multivalent fructose.

[0209] In some embodiments, a delivery vehicle to target the liver, e.g. an LNP, may comprise an apoE ligand and/or a ligand comprising a multivalent N-acetylgalactosamine (GalNAc)-cluster, which binds with high affinity to the asialoglycoprotein receptor (ASGPR) expressed on hepatocytes. In some embodiments, an LNP may comprise Retinol Binding protein (RBP) for targeting hepatic cells, which express the RBP receptor.

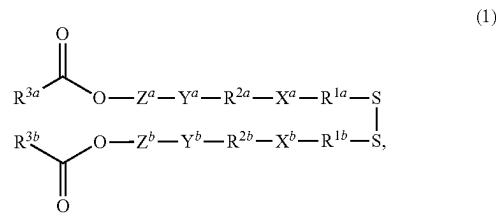
[0210] In some embodiments, a delivery vehicle may comprise an extracellular vesicle, e.g. an exosome, to target the liver. In some embodiments, an extracellular vesicle comprises one or more tissue targeting moieties, including but not limited to lipids, peptides or antibodies.

[0211] Compositions of the disclosure may comprise one or more components that may facilitate delivery of the RNA to cells. Collectively or in part, components of the composition may comprise a delivery vehicle. In some embodiments, the delivery vehicle facilitates targeting and uptake of the ribonucleic acid of a composition of the disclosure to a target cell. Exemplary delivery vehicles include, but are not limited to, nanoparticles, lipid nanoparticles (LNPs), liposomes, micelles, exosomes, cationic lipids and a natural or artificial lipoprotein particle.

[0212] In some embodiments, a delivery vehicle comprises an ionizable lipid. An ionizable lipid may refer to any of a number of lipid species that have a net positive charge at a selected pH, such as a physiological pH. An ionizable lipid may also, for example, refer to a lipid in an ionized state, e.g., a cationic lipid.

[0213] In some embodiments, a cationic lipid formulation comprises a cationic lipid and a structural or matrix lipid. Cationic lipids may be composed of a cationic amine moiety and a lipid moiety, and the cationic amine moiety and a polyanion nucleic acid may interact to form a positively charged liposome or lipid membrane structure. In some embodiments, reference to a lipid "moiety" and a "lipid" may be equivalent. Thus, uptake into cells may be promoted and nucleic acids delivered into cells.

[0214] In some embodiments, the ionizable lipid may be a compound of Formula (1):



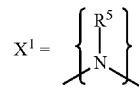
[0215] In the formula (1): R^{1a} and R^{1b} each independently represents an alkylene group having 1 to 6 carbon atoms, and may be linear or branched. The alkylene group may have 1 to 4 carbon atoms, or may have 1 to 2. Specific examples of the alkylene group having 1 to 6 carbon atoms include a methylene group, an ethylene group, a trimethylene group, an isopropylene group, a tetramethylene group, an isobutylene group, a pentamethylene group, and a neopentylene group. R^{1a} and R^{1b} may be each independently a methylene group, an ethylene group, a trimethylene group, an isopropylene group, or a tetramethylene group, and may be an ethylene group.

[0216] R^{1a} may be different or be the same as R^{1b} .

[0217] X^a and X^b are each independently an acyclic alkyl tertiary amino group having 1 to 6 carbon atoms and 1 tertiary amino group, or 2 to 5 carbon atoms, and a cyclic alkylene tertiary amino group having 1 to 2 tertiary amino groups, and/or each independently a cyclic alkylene having 2 to 5 carbon atoms and 1 to 2 tertiary amino groups and an alkylene tertiary amino group.

[0218] The alkyl group having 1 to 6 carbon atoms in the acyclic alkyl tertiary amino group having 1 to 6 carbon atoms and 1 tertiary amino group is branched even if it is linear. The alkyl group may be annular. The alkyl group may have 1 to 3 carbon atoms. Specific examples of the alkyl group having 1 to 6 carbon atoms include methyl group, ethyl group, propyl group, isopropyl group, n-butyl group, sec-butyl group, isobutyl group, tert-butyl group, pentyl group, and isopentyl group. Neopentyl group, t-pentyl group, 1,2-dimethylpropyl group, 2-methylbutyl group, 2-methylpentyl group, 3-methylpentyl group, 2,2-dimethylbutyl group, 2,3-dimethylbutyl group, A cyclohexyl group etc. can be mentioned.

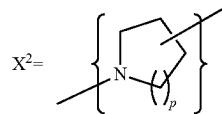
[0219] A specific structure of an acyclic alkyl tertiary amino group having 1 to 6 carbon atoms and 1 tertiary amino group is represented by X^1 .



[0220] R^5 of X^1 represents an alkyl group having 1 to 6 carbon atoms and may be linear, branched or cyclic. The alkyl group may have 1 to 3 carbon atoms. Specific examples of the alkyl group having 1 to 6 carbon atoms include methyl group, ethyl group, propyl group, isopropyl group, n-butyl group, sec-butyl group, isobutyl group, tert-butyl group, pentyl group, and isopentyl group. Neopentyl group, t-pentyl group, 1,2-dimethylpropyl group, 2-methylbutyl group, 2-methylpentyl group, 3-methylpentyl group, 2,2-dimethylbutyl group, 2,3-dimethylbutyl group, A cyclohexyl group etc. can be mentioned.

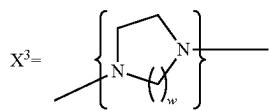
[0221] The number of carbon atoms in the cyclic alkylene tertiary amino group having 2 to 5 carbon atoms and 1 to 2 tertiary amino groups may be 4 to 5. Specific examples of the cyclic alkylene tertiary amino group having 2 to 5 carbon atoms and 1 to 2 tertiary amino groups include aziridylene group, azetidylene group, pyrrolidylene group, piperidylene group, imidazolidylene group, a piperazylene group, optionally a pyrrolidylene group, a piperidylene group or a piperazylene group.

[0222] Number is 2 to 5 carbon atoms, and specific structure of alkylene tertiary amino groups containing 1 annular tertiary amino group represented by X^2 .



[0223] P of X^2 is 1 or 2. When p is 1, X^2 is a pyrrolidylene group, and when p is 2, X^2 is a piperidylene group.

[0224] A specific structure of a cyclic alkylene tertiary amino group having 2 to 5 carbon atoms and 2 tertiary amino groups is represented by X^3 .



[0225] W of X^3 is 1 or 2. When w is 1, X^3 is an imidazolidylene group, and when w is 2, X^3 is a piperazylene group.

[0226] X^a may be different be identical to X^b .

[0227] R^{2a} and R^{2b} each independently represent an alkylene group or an oxydialkylene group having 8 or less carbon atoms, optionally each independently an alkylene group having 8 or less carbon atoms.

[0228] The alkylene group having 8 or less carbon atoms may be linear or branched but is optionally linear. The number of carbon atoms contained in the alkylene group is optionally 6 or less, and optionally 4 or less. Specific examples of the alkylene group having 8 or less carbon atoms include methylene group, ethylene group, propylene group, isopropylene group, tetramethylene group, isobutylene group, pentamethylene group, hexamethylene group, heptamethylene group, octamethylene group, and the like. In some embodiments included are a methylene group, an ethylene group, a propylene group, and a tetramethylene group.

[0229] The oxydialkylene group having 8 or less carbon atoms refers to an alkylene group (alkylene-O-alkylene) via an ether bond, and the total number of carbon atoms of two alkylene groups is 8 or less. Here, the two alkynes may be the same or different, but are optionally the same. Specific examples of the oxydialkylene group having 8 or less carbon atoms include an oxydimethylene group, an oxydiethylene group, an oxydipropylene group, and an oxydibutylene group.

[0230] R^{2a} may be same or different and R^{2b} .

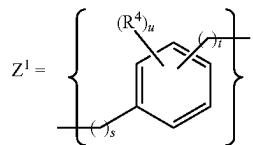
[0231] Y^a and Y^b are each independently an ester bond, an amide bond, a carbamate bond, an ether bond or a urea bond, optionally each independently an ester bond, an amide bond or a carbamate bond. While Y binding orientation of Y^a and Y^b are not limited, if Y^a and Y^b is an ester bond, optionally, $-Z^a-CO-R^{2a}$ and $-Z^b-CO-O-R^{2b}$ -Structure.

[0232] Y^a may be different or identical to Y^b .

[0233] Z^a and Z^b are each independently a divalent group derived from an aromatic compound having 3 to 16 carbon atoms, having at least one aromatic ring, and optionally having a hetero atom. Represents. The number of carbon atoms contained in the aromatic compound is optionally 6 to 12, or 6 to 7. Moreover, the number of aromatic rings contained in the aromatic compound is optionally one.

[0234] As the types of aromatic rings contained in the aromatic compound having 3 to 16 carbon atoms, as for aromatic hydrocarbon rings, benzene ring, naphthalene ring, anthracene ring, and aromatic heterocycles as imidazole ring, pyrazole ring, oxazole ring, Isoxazole ring, thiazole ring, isothiazole ring, triazine ring, pyrrole ring, furanthiophene ring, pyrimidine ring, pyridazine ring, pyrazine ring, pyridine ring, purine ring, pteridine ring, benzimidazole ring, indole ring, benzofuran ring, quinazoline ring, phthalazine ring, quinoline ring, isoquinoline ring, coumarin ring, chromone ring, benzodiazepine ring, phenoxazine ring, phenothiazine ring, acridine ring, etc., optionally benzene ring, naphthalene ring, anthracene ring. The aromatic ring may have a substituent. Examples of the substituent include an acyl group having 2 to 4 carbon atoms, an alkoxy carbonyl group having 2 to 4 carbon atoms, a carbamoyl group having 2 to 4 carbon atoms, and 2 to 2 carbon atoms. 4 acyloxy groups, acylamino groups having 2 to 4 carbon atoms, alkoxy carbonyl amino groups having 2 to 4 carbon atoms, fluorine atoms, chlorine atoms, bromine atoms, iodine atoms, alkylsulfanyl groups having 1 to 4 carbon atoms, 1 carbon atom Alkylsulfonyl group having 4 to 4, arylsulfonyl group having 6 to 10 carbon atoms, nitro group, trifluoromethyl group, cyano group, alkyl group having 1 to 4 carbon atoms, ureido group having 1 to 4 carbon atoms, 1 to carbon atoms 4 alkoxy groups, aryl groups having 6 to 10 carbon atoms, aryloxy groups having 6 to 10 carbon atoms, and the like. Some examples include acetyl groups, methoxycarbonyl groups, methyl carbonate groups, and the like, moyl group, acetoxy group, acetamide group, methoxycarbonyl amino group, fluorine atom, chlorine atom, bromine atom, iodine atom, methylsulfanyl group, phenylsulfonyl group, nitro group, trifluoromethyl group, cyano group, methyl group, ethyl group Propyl group, isopropyl group, t-butyl group, ureido group, methoxy group, ethoxy group, propoxy group, isopropoxy group, t-butoxy group, phenyl group and phenoxy group.

[0235] A specific structure of Z^a and Z^b includes Z^1 .



[0236] Wherein, s represents an integer of 0 to 3, t represents an integer of 0 to 3, u represents an integer of 0 to 4, represents a u-number of R 4 is independently a substituent.

[0237] S in Z^1 is optionally an integer of 0 to 1.

[0238] T in Z^1 is optionally an integer of 0 to 2.

[0239] U in Z^1 is optionally an integer of 0 to 2.

[0240] R 4 in Z^1 is a substituent of an aromatic ring (benzene ring) contained in an aromatic compound having 3 to 16 carbon atoms that does not inhibit the reaction in the process of synthesizing the ionizable lipid. Examples of the substituent include an acyl group having 2 to 4 carbon atoms, an alkoxy carbonyl group having 2 to 4 carbon atoms, a carbamoyl group having 2 to 4 carbon atoms, an acyloxy group having 2 to 4 carbon atoms, and an acylamino group having 2 to 4 carbon atoms, an alkoxy carbonylamino group having 2 to 4 carbon atoms, fluorine atom, chlorine atom, bromine atom, iodine atom, alkylsulfanyl group having 1 to 4 carbon atoms, alkylsulfonyl group having 1 to 4 carbon atoms, 6 to 10 carbon atoms Arylsulfonyl group, nitro group, trifluoromethyl group, cyano group, alkyl group having 1 to 4 carbon atoms, ureido group having 1 to 4 carbon atoms, alkoxy group having 1 to 4 carbon atoms, aryl group having 6 to 10 carbon atoms And aryloxy groups having 6 to 10 carbon atoms, and examples include acetyl, methoxycarbonyl, methylcarbamoyl, acetoxy, Mido group, methoxycarbonylamino group, fluorine atom, chlorine atom, bromine atom, iodine atom, methylsulfanyl group, phenylsulfonyl group, nitro group, trifluoromethyl group, cyano group, methyl group, ethyl group, propyl group, isopropyl group, T-butyl group, ureido group, methoxy group, ethoxy group, propoxy group, isopropoxy group, t-butoxy group, phenyl group and phenoxy group. When a plurality of R^4 are present, each R^4 may be the same or different.

[0241] Z^a may be different even identical to the Z^b .

[0242] R^{3a} and R^{3b} are each independently a residue derived from a reaction product of a fat-soluble vitamin having a hydroxyl group and succinic anhydride or glutaric anhydride, or a sterol derivative having a hydroxyl group and succinic anhydride or glutaric acid. Represents a residue derived from a reaction product with an anhydride, or an aliphatic hydrocarbon group having 12 to 22 carbon atoms, and optionally each independently a fat-soluble vitamin having a hydroxyl group and succinic anhydride or glutaric anhydride. Or a C 12-22 aliphatic hydrocarbon group, and optionally each independently an aliphatic hydrocarbon group having 12-22 carbon atoms.

[0243] Examples of the fat-soluble vitamin having a hydroxyl group include retinol, ergosterol, 7-dehydrocholesterol, calciferol, corcalciferol, dihydroergocalciferol, dihydrotaxolol, tocopherol, and tocotrienol. The fat-soluble vitamin having a hydroxyl group is optionally tocopherol.

[0244] Examples of the sterol derivative having a hydroxyl group include cholesterol, cholestanol, stigmas-

terol, P-sitosterol, lanosterol, ergosterol and the like, optionally cholesterol or cholestanol.

[0245] The aliphatic hydrocarbon group having 12 to 22 carbon atoms may be linear or branched. The aliphatic hydrocarbon group may be saturated or unsaturated. In the case of an unsaturated aliphatic hydrocarbon group, the number of unsaturated bonds contained in the aliphatic hydrocarbon group is usually 1 to 6, optionally 1 to 3, or 1 to 2. Unsaturated bonds include carbon-carbon double bonds and carbon-carbon triple bonds. The number of carbon atoms contained in the aliphatic hydrocarbon group is optionally 13 to 19, or 13 to 17. The aliphatic hydrocarbon group includes an alkyl group, an alkenyl group, an alkynyl group and the like, and optionally includes an alkyl group or an alkenyl group. Specific examples of the aliphatic hydrocarbon group having 12 to 22 carbon atoms include dodecyl, tridecyl, tetradecyl, pentadecyl, hexadecyl, heptadecyl, octadecyl, nonadecyl, icosyl, heicosyl, docosyl, Dodecenyl group, tridecenyl group, tetradecenyl group, pentadecenyl group, hexadecenyl group, heptadecenyl group, octadecenyl group, nonadecenyl group, icocenyl group, henicocenyl group, dococenyl group, dodecadienyl group, tridecadienyl group, tetradecadienyl group, pentadecadienyl group Group, hexadecadienyl group, heptadecadienyl group, octadecadienyl group, nonadecadienyl group, icosadienyl group, heneicosadienyl group, docosadienyl group, octadecatrienyl group, icosatrienyl group, Cosatetraenyl group, icosapentaenyl group, docosahexaenyl group, isostearyl group, 1-hexylheptyl group, 1-hexylnonyl group, 1-octylnonyl group, 1-octylundecyl group, 1-decylundecyl group, etc. be able to. The aliphatic hydrocarbon group having 12 to 22 carbon atoms is optionally a tridecyl group, a pentadecyl group, a heptadecyl group, a nonadecyl group, a heptadecenyl group, a heptadecadienyl group, or a 1-hexylnonyl group, or a tridecyl group, A heptadecyl group, a heptadecenyl group, and a heptadecadienyl group.

[0246] In one embodiment of the present disclosure, the aliphatic hydrocarbon group having 12 to 22 carbon atoms represented by R^{3a} and R^{3b} is derived from a fatty acid. In this case, the carbonyl carbon derived from the fatty acid is contained in CO—O—in the formula (1). Specific examples of the aliphatic hydrocarbon group include a heptadecenyl group when linoleic acid is used as the fatty acid, and a heptadecenyl group when oleic acid is used as the fatty acid.

[0247] R^{3a} may be different be the same as R^{3b} .

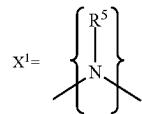
[0248] In one embodiment of the present disclosure, R^{1a} is the same as R^{1b} , X^a is the same as X^b , R^{2a} is the same as R^{2b} , Y^a is the same as Y^b , and Z^a is identical to the Z^b , R^{3a} is the same as R^{3b} .

[0249] Preferable examples of the ionizable lipid represented by the formula (1) include the following ionizable lipids: Ionizable lipid (1-1); R^{1a} and R^{1b} are each independently an alkylene group having 1 to 6 carbon atoms (eg, methylene group, ethylene group); X a and X b are each independently an acyclic alkyl tertiary amino group having 1 to 6 carbon atoms and 1 tertiary amino group (eg, N(CH₃)₃), Or a cyclic alkylene tertiary amino group having 2 to 5 carbon atoms and 1 to 2 tertiary amino groups (eg, piperidylene group); R^{2a} and R^{2b} are each independently an alkylene group having 8 or less carbon atoms (eg, methylene group, ethylene group, propylene group); Y^a and Y^b are each independently an ester bond or an amide bond; Z^a and Z^b are each independently a divalent group derived from an aromatic compound having 3 to 16 carbon atoms, having at

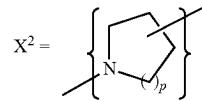
least one aromatic ring, and optionally having a hetero atom. (Eg, C 6 H 4 CH 2, CH 2 C 6 H 4 CH 2); R^{3a} and R^{3b} are each independently a residue derived from a reaction product of a fat-soluble vitamin having a hydroxyl group (eg, tocopherol) and succinic anhydride or glutaric anhydride, or an aliphatic group having 12 to 22 carbon atoms A hydrocarbon group (eg, heptadecenyl group, heptadecadienyl group, 1-hexylnonyl group);

[0250] Ionizable lipid (1-2); Ria and R^{1b} are each independently an alkylene group having 1 to 4 carbon atoms (eg, methylene group, ethylene group); X a and X b are each independently an acyclic alkyl tertiary amino group having 1 to 3 carbon atoms and 1 tertiary amino group (eg, —N(CH₃)). Or a cyclic alkylene tertiary amino group having 2 to 5 carbon atoms and 1 tertiary amino group (eg, piperidylene group); R^{2a} and R^{2b} are each independently an alkylene group having 6 or less carbon atoms (eg, methylene group, ethylene group, propylene group); Y^a and Y^b are each independently an ester bond or an amide bond; Z^a and Z^b are each independently a divalent group derived from an aromatic compound having 6 to 12 carbon atoms, one aromatic ring, and optionally having a hetero atom (Eg, C 6 H 4 CH 2, CH 2 C 6 H 4 CH 2); R^{3a} and R^{3b} are each independently a residue derived from a reaction product of a fat-soluble vitamin having a hydroxyl group (eg, tocopherol) and succinic anhydride, or an aliphatic hydrocarbon group having 13 to 19 carbon atoms (eg., Heptadecenyl group, heptadecadienyl group, 1-hexylnonyl group).

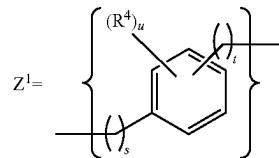
[0251] Ionizable lipid (1-3); R^{1a} and R^{1b} are each independently an alkylene group having 1 to 2 carbon atoms (eg, methylene group, ethylene group); X^a and X^b are each independently X¹:



wherein R⁵ is an alkyl group having 1 to 3 carbon atoms (eg, a methyl group)), or X²:



wherein p is 1 or 2), R^{2a} and R^{2b} are each independently an alkylene group having 4 or less carbon atoms (eg, methylene group, ethylene group, propylene group); Y^a and Y^b are each independently an ester bond or an amide bond; Z^a and Z^b are each independently Z¹:



wherein s is an integer from 0 to 1, t is an integer from 0 to 2, u is an integer from 0 to 2 (optionally 0), and (R⁴)_u are each independently represents a substituent. R^{3a} and R^{3b} are each independently a residue derived from a reaction product of a fat-soluble vitamin having a hydroxyl group (eg, tocopherol) and succinic anhydride, or an aliphatic hydrocarbon group having 13 to 17 carbon atoms (eg, Heptadecenyl group, heptadecadienyl group, 1-hexylnonyl group); Ionizable lipid (1).

[0252] Specific examples of the ionizable lipid (1) of the present disclosure include the following O-Ph-P3C1, O-Ph-P4C1, O-Ph-P4C2, O-Bn-P4C2, E-Ph-P4C2, L-Ph-P4C2, HD-Ph-P4C2, O-Ph-amide-P4C2, and O-Ph-C3M as seen in Tables 2, 3, and 4.

TABLE 2

Ionizable lipids

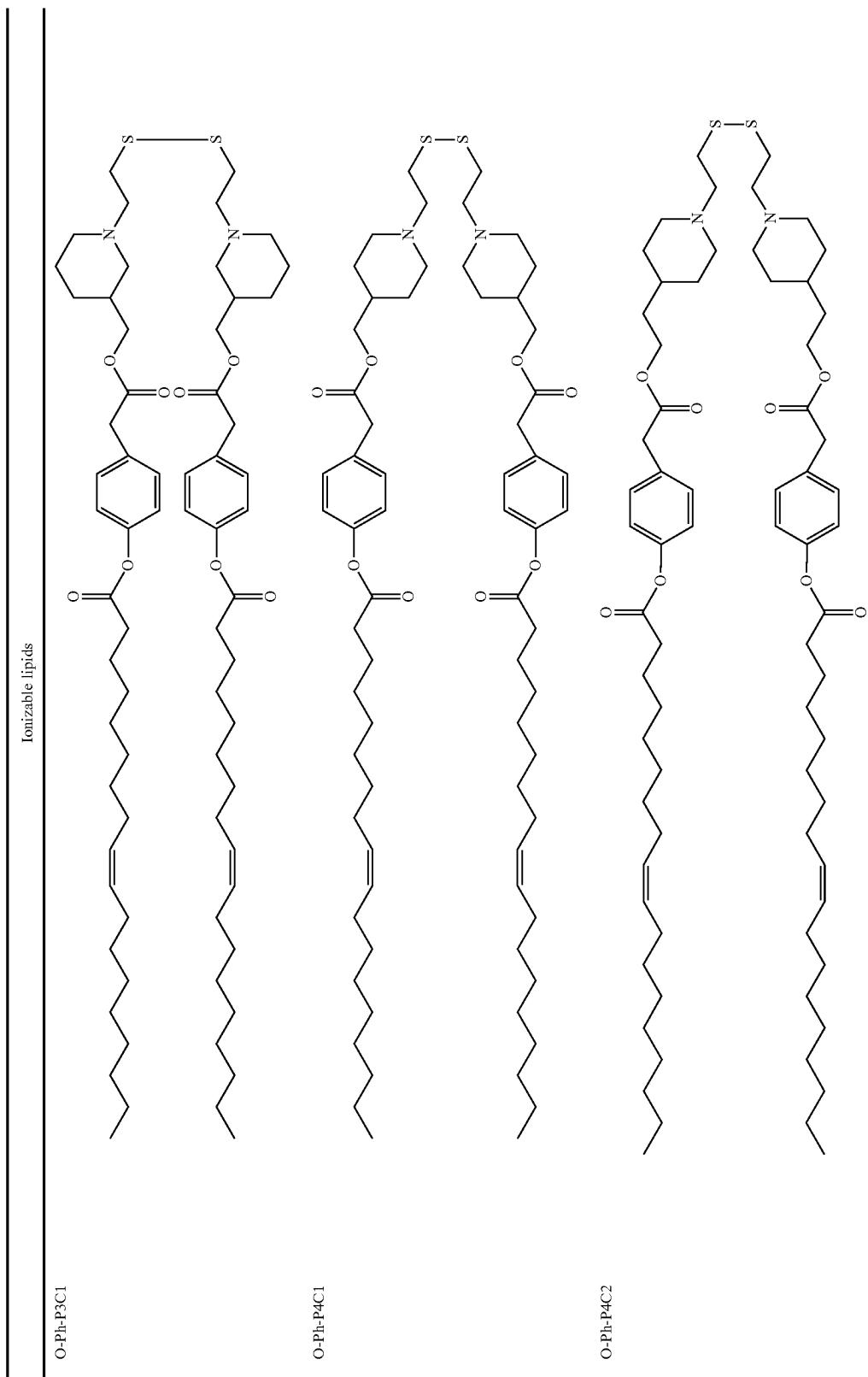


TABLE 2-continued

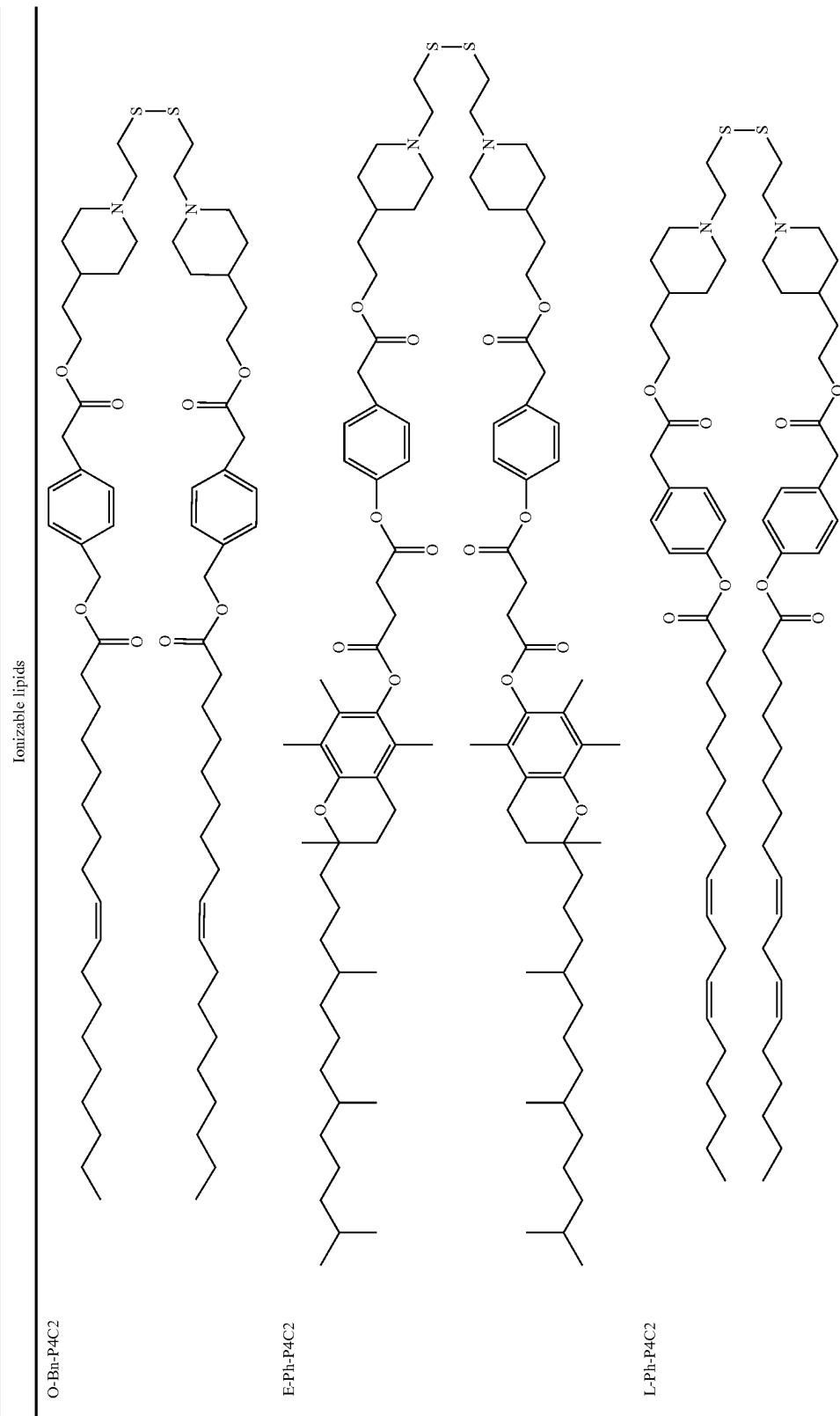


TABLE 2-continued

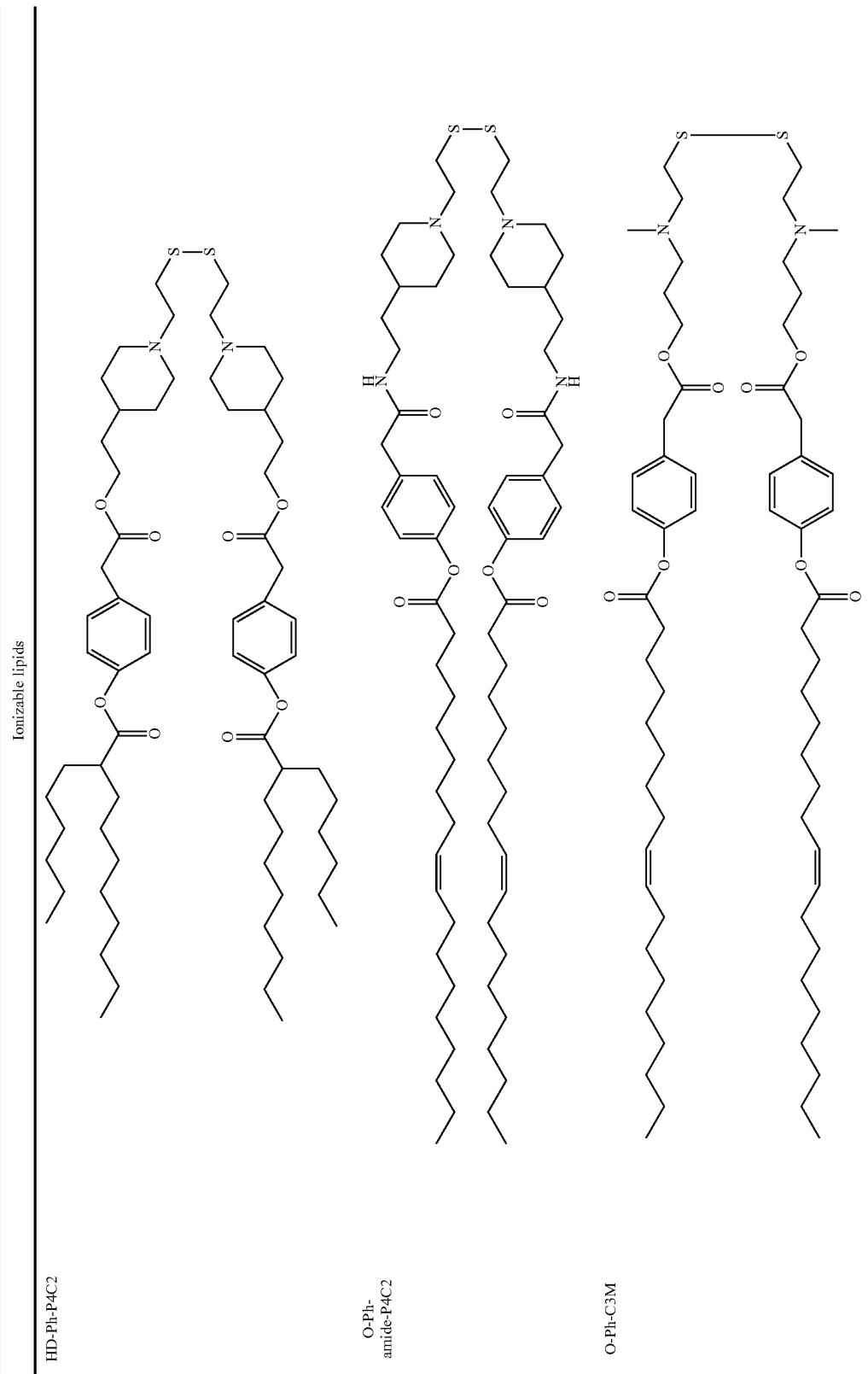
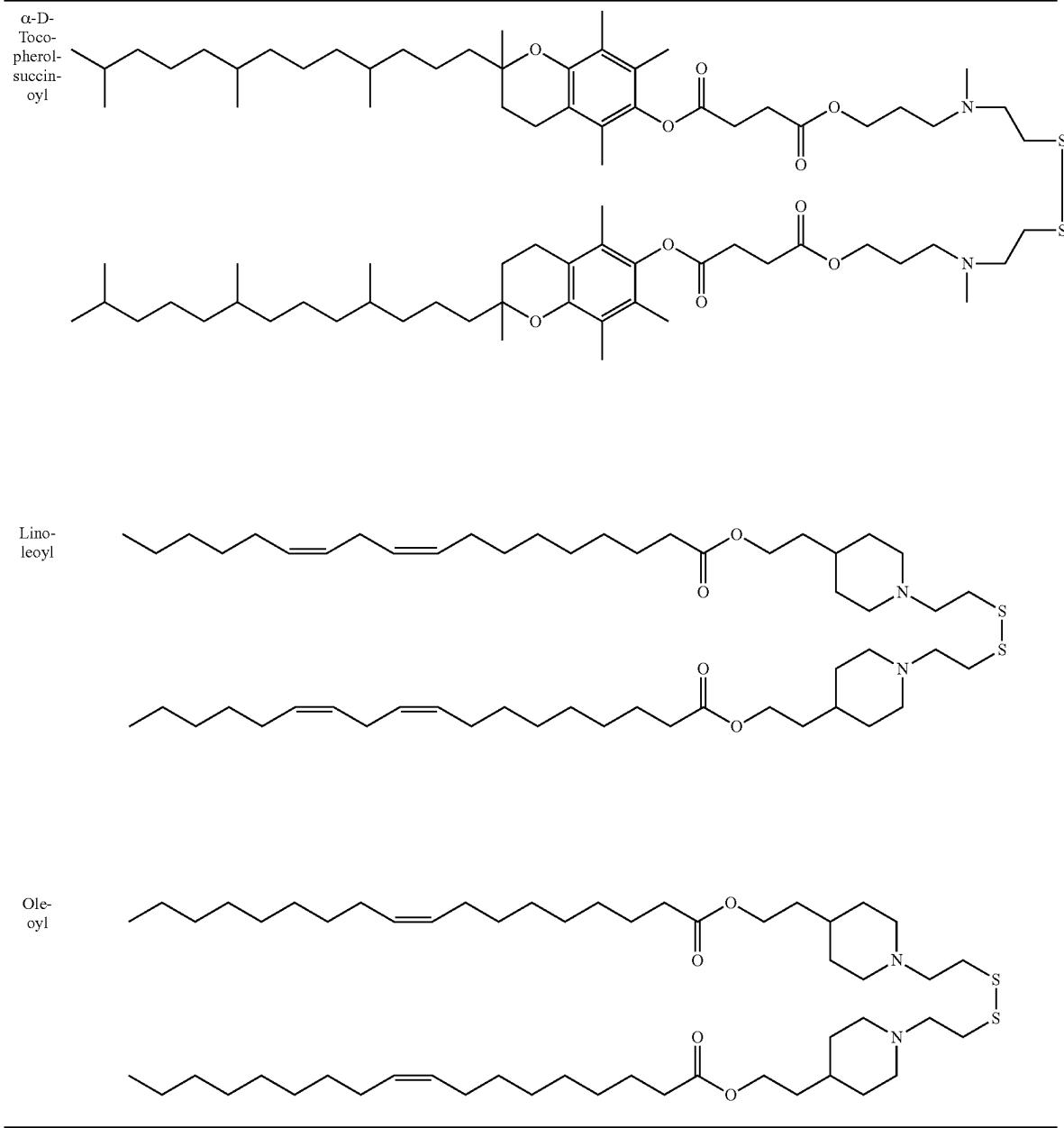


TABLE 3

Ionizable lipids



[0253] In some embodiments, the delivery vehicle is an LNP capable of transfecting at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95% of a population of liver cells wherein the ionizable lipid is at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, or at least 80% of the molar percentage of the LNP.

[0254] In some embodiments, the LNP comprises an ionizable lipid. In some embodiments, the ionizable lipid is no more than 20%, no more than 30%, no more than 40%, no more than 50%, no more than 60%, no more than 70%, or no more than 90% of the molar percentage of the LNP.

Exemplary ionizable lipids include, but are not limited to: imidazole cholesterol ester (ICE), (15Z,18Z)-N,N-dimethyl-6-(9Z,12Z)-octadeca-9,12-dien-1-yl)tetracosa-15,18-dien-1-amine (HGT5000), (15Z,18Z)-N,N-dimethyl-6-(9Z, 12Z)-octadeca-9,12-dien-1-yl)tetracosa-4,15,18-trien-1-amine (HGT5001), and (15Z,18Z)-N,N-dimethyl-6-(9Z, 12Z)-octadeca-9,12-dien-1-yl)tetracosa-5,15,18-trien-1-amine (HGT5002).

[0255] Lipids having the structure of Formula I are shown in Table 4 below. For example, SS-OP is also named 0-Ph-P4C2. The term "SS-OP analog" as used herein refers to a compound of Formula I.

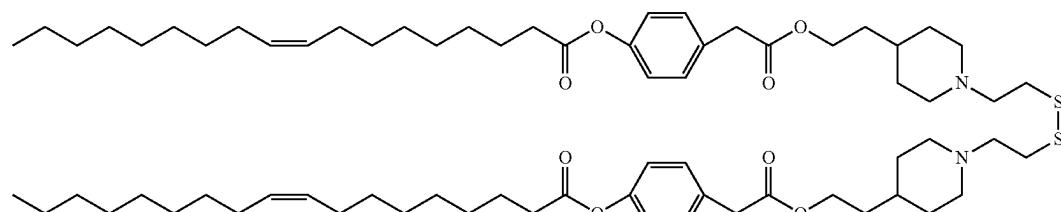
TABLE 4

Nomenclature of Lipids

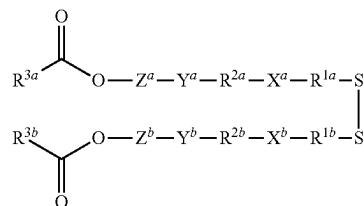
Name	Structure
SS-M	
SS-E	
SS-EC	
SS-LC	
SS-OC	

TABLE 4-continued

Nomenclature of Lipids

Name	Structure
SS-OP	

[0256] Provided herein are compositions comprising i) a ribonucleic acid (RNA) coding for telomerase reverse transcriptase (TERT) and ii) a compound of Formula (I):



[0257] In the formula (I): R^{1a} and R^{1b} can each independently represent an alkylene group having 1 to 6 carbon atoms, and may be linear or branched, but is optionally linear. The alkylene group optionally has 1 to 4 carbon atoms, or 1 to 2. Specific examples of the alkylene group having 1 to 6 carbon atoms include a methylene group, an ethylene group, a trimethylene group, an isopropylene group, a tetramethylene group, an isobutylene group, a pentamethylene group, and a neopentylene group. R^{1a} and R^{1b} are optionally each independently a methylene group, an ethylene group, a trimethylene group, an isopropylene group, or a tetramethylene group, or an ethylene group.

[0258] R^{1a} may be different or be the same as R^{1b}.

[0259] X^a and X^b can each independently be an acyclic alkyl tertiary amino group having 1 to 6 carbon atoms and 1 tertiary amino group, or 2 to 5 carbon atoms, and a cyclic alkylene tertiary amino group having 1 to 2 tertiary amino groups, optionally each independently a cyclic alkylene having 2 to 5 carbon atoms and 1 to 2 tertiary amino groups and an alkylene tertiary amino group.

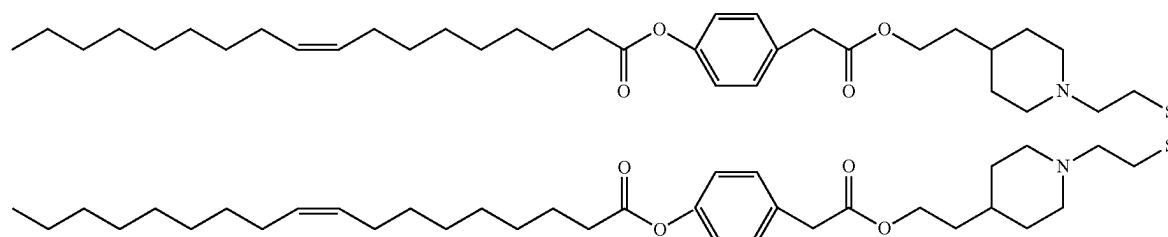
[0260] In some embodiments, the compound of Formula II is

[0261] The RNA can be a synthetic RNA. The RNA can comprise at least one modified nucleoside. Also provided herein are methods for delivery of the compositions to a cell. The compound of Formula I can be used to aid in delivery of the RNA to a cell *in vitro* or *in vivo*. Once delivered to a cell, the synthetic RNA can transiently express exogenous telomerase in the cell, and telomeres within the cell treated with the synthetic RNA can be extended. Thus the compositions can be used to extend telomeres within a cell.

III. Formulation of mRNA and Nanoparticle Delivery Vehicle Compositions

[0262] The methods of synthesis of mRNA and lipid nanoparticles (LNPs) are well established. Synthetic mRNAs, e.g., comprising a 5' cap, 5' and 3' UTRs coding sequence, and a poly-A tail, may be synthesized from modified and unmodified nucleotides by *in vitro* transcription of a DNA template using an RNA polymerase, for example T7 RNA polymerase. The DNA template may be generated, for example, by PCR or plasmid amplification and restriction digest, followed by purification.

[0263] Lipid nanoparticles (LNPs), liposomes, or polymer nanoparticle delivery vehicles carrying mRNA may be produced, for example, by mixing the lipids or polymers in an organic solvent, e.g., ethanol, with one or more mRNAs in an aqueous buffer, and then subject to buffer exchange and concentration. In some embodiments, the LNP, liposome, or polymer nanoparticle delivery vehicle may be produced using a microfluidic device to rapidly mix reagents and form monodisperse particles of controlled size. For example, the microfluidic mixer could be a staggered herringbone mixer (SHM). For example, the microfluidic mixer could be produce by the NanoAssemblr made by Precision Nanosystems (PNI). In other embodiments, the LNP, liposome, or polymer nanoparticle delivery vehicle may be produced by a T-mixer. In some embodiments, the LNP, liposome, or polymer nanoparticle may encapsulate an mRNA and/or



associate with one or more mRNAs through electrostatic interactions. The buffer exchange and concentration of the LNP, liposome, or polymer nanoparticle may be performed by tangential flow filtration. In other embodiments, the buffer exchange and concentration of the LNP, liposome, or polymer nanoparticle may be performed by centrifugal ultrafiltration using a membrane with a nominal molecule weight cutoff of <=500,000 Da, for example 100,000 Da.

[0264] In some embodiments, the lipid nanoparticle particles (LNP) formulations provided herein are capable of transfecting at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95% of a population of liver cells.

[0265] The form of the lipid membrane structure of the present disclosure is not particularly limited. For example, as a form 1 which the ionizable lipid of the present disclosure is dispersed in an aqueous solvent, liposomes (for example, monolayer liposomes, multilamellar liposomes, etc.), spherical micelles, string micelles, lipid nanoparticles (LNPs) or unspecified layered structures.

[0266] The lipid membrane structure of the present disclosure may further contain other components in addition to the ionizable lipid of the present disclosure. Examples of the other components include lipids (phospholipids (such as phosphatidylinositol, phosphatidylethanolamine, phosphatidylserine, phosphatidic acid, phosphatidylglycerol, phosphatidylcholine), glycolipids, peptide lipids, cholesterol, ionizable lipids other than cationic lipids, PEG lipids, etc.), surfactants (eg 3-[(3-cholamidopropyl) dimethylammonio] propane sulfonate, cholic acid sodium salt, octyl glycoside, ND-gluco-N-methylalkanamides), polyethylene glycol, proteins and the like. The content of the other constituents in the lipid membrane structure of the present disclosure is usually 5 to 95 mol %, optionally 10 to 90 mol %, or 30 to 80 mol %.

[0267] The content of the ionizable lipid of the present disclosure contained in the lipid membrane structure of the present disclosure is not particularly limited.

[0268] The lipid membrane structure of the present disclosure is prepared by dispersing the ionizable lipid of the present disclosure and other components (lipids, etc.) in a suitable solvent or dispersion medium, for example, an aqueous solvent or an alcoholic solvent, and if necessary, tissue. It can be prepared by performing an operation that induces crystallization.

[0269] Examples of the "operation for inducing organization" include an ethanol dilution method using a microchannel or a vortex, a simple hydration method, an ultrasonic treatment, a heating, a vortex, an ether injection method, a French press method, and a cholic acid method. Examples thereof include, but are not limited to, methods known per se such as Ca 2+ fusion method, freeze-thaw method, and reverse phase evaporation method.

[0270] The nucleic acid can be introduced into the cell in vivo and/or in vitro by encapsulating the nucleic acid in the lipid membrane structure containing the ionizable lipid of the present disclosure and bringing it into contact with the cell. Therefore, the present disclosure provides a nucleic acid introduction agent comprising the ionizable lipid or lipid membrane structure of the present disclosure.

[0271] The nucleic acid introduction agent of the present disclosure can introduce any nucleic acid into cells. Examples of the nucleic acid include, but are not limited to, DNA, RNA, RNA chimeric nucleic acid, DNA/RNA hybrid,

and the like. The nucleic acid can be any one of 1 to 3 strands, but is optionally single strand or double strand. Nucleic acids may be other types of nucleotides that are N-glycosides of purine or pyrimidine bases, or other oligomers having a non-nucleotide backbone (e.g., commercially available peptide nucleic acids (PNA), etc.) or other oligomers with special linkages. The oligomer may contain nucleotides having a configuration that allows base pairing or base attachment as found in DNA or RNA. In addition, the nucleic acid may be substituted with, for example, a known modified nucleic acid, a labeled nucleic acid, a capped nucleic acid, a methylated nucleic acid, or one or more natural nucleotides known in the art, intramolecular nucleotide modified nucleic acids, nucleic acids with uncharged bonds (e.g., methyl sulfonate, phosphotriester, phosphoramidate, carbamate, etc.), charged bonds or sulfur containing bonds (eg phosphorothioate), side chain groups such as proteins (e.g., nucleases, nuclease inhibitors, toxins, antibodies, signal peptides, poly-L-lysine, etc.) and sugars (eg, monosaccharides), nucleic acids and nucleic acids with intercurrent compounds (eg, acridine, psoralen, etc.), nucleic acids containing chelate compounds (eg, metals, radioactive metals, boron, oxidizing metals, etc.), nucleic acids containing alkylating agents, and nucleic acids with modified bonds (eg, alpha anomeric nucleic acids, etc.).

[0272] The type of DNA that can be used in the present disclosure is not particularly limited, and can be appropriately selected depending on the purpose of use. Examples include plasmid DNA, cDNA, antisense DNA, chromosomal DNA, PAC, BAC, and CpG oligo, optionally plasmid DNA, cDNA, and antisense DNA, or plasmid DNA. Circular DNA such as plasmid DNA can be appropriately digested with a restriction enzyme or the like and used as linear DNA.

[0273] The type of RNA that can be used in the present disclosure is not particularly limited, and can be appropriately selected depending on the purpose of use. For example, siRNA, miRNA, shRNA, antisense RNA, messenger RNA (mRNA), single-stranded RNA genome, double-stranded RNA genome, RNA replicon, transfer RNA, ribosomal RNA, etc., optionally siRNA, miRNA, shRNA, miRNA, antisense RNA, RNA replicon.

[0274] The nucleic acid used in the present disclosure is optionally purified by a method commonly used by those skilled in the art.

[0275] The nucleic acid-introducing agent of the present disclosure encapsulating nucleic acid can be administered in vivo for the purpose of, for example, prevention and/or treatment of diseases. Accordingly, the nucleic acid used in the present disclosure is optionally a nucleic acid having preventive and/or therapeutic activity against a given disease (prophylactic/therapeutic nucleic acid). Examples of such nucleic acids include nucleic acids used for so-called gene therapy.

[0276] In order to introduce a nucleic acid into a cell using the nucleic acid introduction agent of the present disclosure, the nucleic acid was encapsulated by coexisting the target nucleic acid when forming the lipid membrane structure of the present disclosure. The lipid membrane structure of the present disclosure is formed. For example, when liposomes are formed by the ethanol dilution method, the aqueous solution of nucleic acid and the ethanol solution of the components of the lipid membrane structure of the present disclosure (lipids, etc.) are vigorously mixed by vortex or microchannel, etc, is diluted with an appropriate buffer.

When liposomes are formed by the simple hydration method, the components (lipids, etc.) of the lipid membrane structure of the present disclosure are dissolved in an appropriate organic solvent, the solution is placed in a glass container, and the solvent is retained by drying under reduced pressure and left to obtain a lipid film. Here, an aqueous solution of nucleic acid is added and hydrated, followed by sonication with a sonicator. The present disclosure also provides the above lipid membrane structure in which such a nucleic acid is encapsulated.

[0277] An example of a lipid membrane structure in which a nucleic acid is encapsulated is LNP encapsulated in a nucleic acid by forming an electrostatic complex between the nucleic acid and a ionizable lipid. This LNP can be used as a drug delivery system for selectively delivering a nucleic acid or the like into a specific cell. For example, a DNA vaccine by introducing an antigen gene into a dendritic cell, a gene therapy drug for a tumor, RNA It is useful for nucleic acid drugs that suppress the expression of target genes using interference.

[0278] The particle diameter of the lipid membrane structure of the present disclosure encapsulating nucleic acid is not particularly limited, but is optionally 10 nm to 500 nm, or 30 nm to 300 nm. The particle diameter can be measured using a particle size distribution measuring apparatus such as Zetasizer Nano (Malvern). The particle diameter of the lipid membrane structure can be appropriately adjusted according to the method for preparing the lipid membrane structure.

[0279] The surface potential (zeta potential) of the lipid membrane structure of the present disclosure encapsulating nucleic acid is not particularly limited, but may be -60 to +60 mV, -45 to 45 mV, -30 to +30 mV, -15 to +15 mV, or -10 to -10 mV. In conventional gene transfer, particles having a positive surface potential have been mainly used. While this is useful as a method to positive electrostatic interaction with negatively charged cell surface heparin sulfate and promote cellular uptake, positive surface charge is delivered intracellularly. There is a possibility that the nucleic acid release from the carrier due to the interaction with the nucleic acid is suppressed, and the protein synthesis due to the interaction between the mRNA and the delivery nucleic acid is suppressed. By adjusting the surface charge within the above range, this problem can be solved. The surface charge can be measured by using a zeta potential measuring device such as Zetasizer Nano. The surface charge of the lipid membrane structure can be adjusted by the composition of the components of the lipid membrane structure containing the ionizable lipid of the present disclosure.

[0280] The lipid membrane surface pKa (hereinafter referred to as Liposomal pKa) of the lipid membrane structure of the present disclosure is not particularly limited, but may have a pKa of 0.5 to 72, or a pKa of 6.0, to 6.8. Liposomal pKa is used as an index indicating that the lipid membrane structure taken up by endocytosis is susceptible to protonation of the lipid membrane structure in a weakly acidic environment within the endosome. Liposomal pKa can be adjusted by the composition of the components of the lipid membrane structure containing the ionizable lipid of any of the above embodiments.

[0281] The hemolysis activity (membrane fusion ability) of a lipid membrane structure of the present disclosure is not particularly limited, but may have no hemolysis activity

(less than 5%) at physiological pH (pH 7.4), and may be endosomal. The higher the hemolysis activity, the more efficiently the nucleic acid can be delivered into the cytoplasm. However, if the hemolysis activity is present at physiological pH, the nucleic acid will be delivered to unintended cells during residence in the blood, resulting in decreased target-directedness and toxicity. Therefore, it is preferable to have hemolysis activity only in the endosomal environment as described above. The hemolysis activity can be adjusted by the composition of the components of the lipid membrane structure containing the ionizable lipid of the present disclosure.

[0282] By bringing the lipid membrane structure of the present disclosure in which nucleic acid is encapsulated into contact with the cell, the encapsulated nucleic acid can be introduced into the cell. The cell may be a cultured cell line containing cancer cells, a cell isolated from an individual or tissue, or a tissue or tissue piece of cell. Further, the cells may be adherent cells or nonadherent cells.

[0283] The step of bringing the lipid membrane structure of the present disclosure encapsulating nucleic acid into contact with cells in vitro will be specifically described below.

[0284] Cells are suspended in an appropriate medium several days before contact with the lipid membrane structure and cultured under appropriate conditions. Upon contact with the lipid membrane structure, the cell may or may not be in the growth phase.

[0285] The culture medium at the time of the contact may be a serum-containing medium or a serum-free medium, but the serum concentration in the medium may be 30% by weight or less, more may be 20% by weight or less. If the medium contains excessive protein such as serum, the contact between the lipid membrane structure and the cell may be inhibited.

[0286] The cell density at the time of the contact is not particularly limited and can be appropriately set in consideration of the cell type, but is usually in the range of 1×10^4 to 1×10^7 cells/mL.

[0287] For example, a suspension of the lipid membrane structure of the present disclosure in which the above-described nucleic acid is encapsulated is added to the cells thus prepared. The addition amount of the suspension is not particularly limited, and can be appropriately set in consideration of the number of cells and the like. The concentration of the lipid membrane structure at the time of contacting the cell is not particularly limited as long as the introduction of the target nucleic acid into the cell can be achieved, but the lipid concentration is usually 1 to 100 nmol/mL, and may be 0.1 to 10 µg/mL.

[0288] After adding the above suspension to the cells, the cells are cultured. The culture temperature, humidity, CO₂ concentration, etc. are appropriately set in consideration of the cell type. When the cells are mammalian cells, the temperature is usually about 37° C., the humidity is about 95%, and the CO₂ concentration is about 5%. In addition, the culture time can be appropriately set in consideration of conditions such as the type of cells used, but may be in the range of 0.1 to 76 hours, or in the range of 0.2 to 24 hours, and may be 0.5-12 hours. If the culture time is too short, the nucleic acid is not sufficiently introduced into the cells, and if the culture time is too long, the cells may be weakened.

[0289] The nucleic acid is introduced into the cells by the above-described culture. The medium may be replaced with

a fresh medium, or the fresh medium is added to the medium and the cultivation is further continued. If the cells are mammalian cells, the fresh medium may contain serum or nutrient factors.

[0290] The lipid membrane structure of the present disclosure may further contain other components in addition to the ionizable lipid of the present disclosure. Examples of the other components include lipids (phospholipids (such as phosphatidylinositol, phosphatidylethanolamine, phosphatidylserine, phosphatidic acid, phosphatidylglycerol, phosphatidylcholine), glycolipids, peptide lipids, cholesterol, ionizable lipids other than cationic lipids, PEG lipids, etc.), surfactants (eg 3-[(3-cholamidopropyl) dimethylammonio] propane sulfonate, cholic acid sodium salt, octyl glycoside, ND-gluco-N-methylalkanamides), polyethylene glycol, proteins and the like.

[0291] The lipid membrane structure of the present disclosure is prepared by dispersing the ionizable lipid of the present disclosure and other components (lipids, etc.) in a suitable solvent or dispersion medium, for example, an aqueous solvent or an alcoholic solvent, and if necessary, tissue. It can be prepared by performing an operation that induces crystallization.

[0292] Examples of the "operation for inducing organization" include an ethanol dilution method using a microchannel or a vortex, a simple hydration method, an ultrasonic treatment, a heating, a vortex, an ether injection method, a French press method, and a cholic acid method. Examples thereof include, but are not limited to, methods known per se such as Ca 2+ fusion method, freeze-thaw method, and reverse phase evaporation method.

[0293] The nucleic acid can be introduced into the cell in vivo and/or in vitro by encapsulating the nucleic acid in the lipid membrane structure containing the ionizable lipid of the present disclosure and bringing it into contact with the cell. Therefore, the present disclosure provides a nucleic acid introduction agent comprising the ionizable lipid or lipid membrane structure of the present disclosure.

[0294] The nucleic acid introduction agent of the present disclosure can introduce any nucleic acid into cells, Examples of the nucleic acid include, but are not limited to, DNA, RNA, RNA chimeric nucleic acid, DNA/RNA hybrid, and the like. The nucleic acid can be any one of 1 to 3 strands, but may be single strand or double strand. Nucleic acids may be other types of nucleotides that are N-glycosides of purine or pyrimidine bases, or other oligomers having a non-nucleotide backbone (eg, commercially available peptide nucleic acids (PNA), etc.) or other oligomers with special linkages. The oligomer may contain nucleotides having a configuration that allows base pairing or base attachment as found in DNA or RNA.

[0295] The type of RNA that can be used in the present disclosure is not particularly limited, and can be appropriately selected depending on the purpose of use. For example, siRNA, miRNA, shRNA, antisense RNA, messenger RNA (mRNA), single-stranded RNA genome, double-stranded RNA genome, RNA replicon, transfer RNA, ribosomal RNA, etc., or siRNA, miRNA, shRNA, mRNA, antisense RNA, or an RN A replicon.

[0296] The nucleic acid used in the present disclosure may be purified by a method commonly used by those skilled in the art.

[0297] The nucleic acid-introducing agent of the present disclosure encapsulating nucleic acid can be administered in

vivo for the purpose of, for example, prevention and/or treatment of diseases. Accordingly, the nucleic acid used in the present disclosure may be a nucleic acid having preventive and at/or therapeutic activity against a given disease (prophylactic/therapeutic nucleic acid). Examples of such nucleic acids include nucleic acids used for so-called gene therapy.

IV. Methods of Treatment

[0298] Methods of treatment as described herein refer to the treatment of fibrotic disease and/or liver disease in a subject in need thereof by administration of a composition comprising one or more TERT mRNA sequences. Compositions and methods of the disclosure may be used for the treatment of fibrotic conditions, including fibrosis. In some embodiments, compositions and/or methods of use of compositions of the disclosure intended for treatment of fibrotic conditions, including fibrosis, induce TERT expression or increase TERT activity in a liver cell. In some embodiments, compositions and/or methods of use of compositions of the disclosure intended for treatment of fibrotic conditions, including fibrosis, do not induce cellular, tissue or systemic toxicity. In some embodiments, compositions and/or methods of use of compositions of the disclosure intended for treatment of fibrotic conditions, including fibrosis, induce TERT expression or increase TERT activity in a spleen cell. Compositions may be administered systemically, e.g., intravenously.

A. Dosage and Timing of Telomerase Reverse Transcriptase (TERT) mRNA

[0299] In the compositions and methods described herein, in some embodiments, a TERT mRNA is administered in a dose of about 0.001 mg/kg per the subject's body weight to about 2.0 mg/kg per the subject's body weight to a subject in need thereof. In some embodiments, a TERT mRNA is administered to a subject in need thereof in a dose of about 0.01 mg/kg; in some embodiments in a dose of about 0.025 mg/kg; in some embodiments in a dose of about 0.05 mg/kg; in some embodiments in a dose of about 0.075 mg/kg; in some embodiments in a dose of about 0.1 mg/kg; in some embodiments in a dose of about 0.125 mg/kg; in some embodiments in a dose of about 0.150 mg/kg; in some embodiments in a dose of about 0.175 mg/kg; in some embodiments in a dose of about 0.2 mg/kg; in some embodiments in a dose of about 0.5 mg/kg; in some embodiments in a dose of about 0.75 mg/kg; in some embodiments in a dose of about 1.0 mg/kg; in some embodiments, in a dose of about 1.25 mg/kg; in some embodiment in a dose of about 1.5 mg/kg; or in some embodiment in a dose of about 2.0 mg/kg. In some embodiments the TERT mRNA is administered to a subject in need thereof in a dose of 0.1 mg/kg. In some embodiments the TERT mRNA is administered to a subject in need thereof in a dose of 0.125 mg/kg.

[0300] In some embodiments the TERT mRNA is administered to a subject in need thereof in a single dose. In some embodiments the TERT mRNA is administered to a subject in need thereof two, three, four, or five or more times. In some embodiments, the TERT mRNA is administered twice a week, every week, every two weeks, every four weeks, every six weeks, every twelve weeks, or every fifteen weeks. In some embodiments, the TERT mRNA is administered every month, every two months, every six months, once a year, on an ongoing basis, or as determined by their physician.

B. TERT mRNA Co-Therapies

[0301] In some embodiments, co-administration of a TERT mRNA may be combined with other anti-fibrotic drugs used in the treatment of fibrotic diseases and/or liver diseases. Drugs that may be used include, but are not limited to nintedanib, pirfenidone, prednisone, azathioprine, cyclophosphamide, mycophenolate mofetil, Pamrevlumab, and N-acetylcysteine.

C. Routes of Administration

[0302] In some embodiments, a TERT mRNA may be delivered orally, subcutaneously, intravenously, intranasally, intradermally, transdermally, intraperitoneally, intramuscularly, intrapulmonarily, vaginally, rectally, or intraocularly. In example embodiments a TERT mRNA may be administered intravenously or through inhalation.

D. Subjects and Treatment

[0303] The methods of treatment described herein are useful for the treatment of fibrotic diseases, conditions and disorders, and liver diseases, conditions, and disorders in a subject in need thereof. Fibrotic diseases and conditions of the disclosure include, but are not limited to, non-alcoholic hepatitis, hepatitis A, hepatitis B, hepatitis C, alcoholic hepatitis, liver cirrhosis, hemochromatosis, Wilson's disease, nonalcoholic steatohepatitis (NASH), NASH with fibrosis stage F4 according to the METAVIR scoring system, compensated liver cirrhosis, decompensated liver cirrhosis, acute-on-chronic liver cirrhosis, biliary atresia, primary biliary cirrhosis, primary sclerosing cholangitis, auto-immune hepatitis, cryptic cirrhosis, and ischemic hepatitis.

[0304] In some embodiments, a subject in need of treatments described herein is a subject with a genetic disorder or mutation in telomerase reverse transcriptase (TERT). In some embodiments the subject has no symptoms of fibrosis or liver disease. In other embodiments, the subject has symptoms and the treatment completely or partially ameliorates the symptoms. In other embodiments, the treatment slows progression of the symptoms.

[0305] In some embodiments, the subject is human.

[0306] In some embodiments, administration of a TERT mRNA reduces fibrotic tissue relative to a subject without treatment. In some embodiments, fibrotic tissue levels are measured by the METAVIR scoring system. In some embodiments, a TERT mRNA reduces the fibrotic stage of the tissue (e.g., from F4 to F3, F3 to F2, F2 to F1, or F1 to F0, or variations thereof) according to the METAVIR scoring system. In some embodiments, administration of a TERT mRNA reduces collagen levels.

[0307] In some embodiments, administration of a TERT mRNA reduces fibrotic tissue in a subject by at least 5%, at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 99%, or at least 100% over the treatment period and/or after the treatment period.

[0308] In some embodiments, administration of a TERT mRNA stops or slows the increase in fibrotic tissue over time relative to a subject without treatment. In some embodiments, the administration of a TERT mRNA slows the increase in amount of fibrotic tissue in a subject by at least 5%, at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least

90%, at least 95%, at least 99%, or at least 100% over the treatment period and/or after the treatment period.

[0309] In some embodiments, administration of a TERT mRNA increases liver function relative to a subject without treatment. In some embodiments, the administration of a TERT mRNA increases liver function in a subject by at least 5%, at least 10%, at least 20%, at least 30%, at least 40%, or at least 50% over the treatment period and/or after the treatment period.

[0310] In some embodiments, administration of a TERT mRNA extends survival relative to a subject without treatment. In some embodiments, administration of a TERT mRNA extends liver transplant-free survival relative to a subject without treatment. In some embodiments, the administration of a TERT mRNA extends survival of a subject by at least 5%, at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 99%, at least 100%, at least 200%, at least 300%, at least 400%, at least 500%, at least 1000%, over the treatment period and/or after the treatment period. In some embodiments, administration of a TERT mRNA reduces hospitalization time and/or number of hospitalization visits to treat the fibrotic disease or liver disease. In some embodiments, administration of a TERT mRNA delays time to liver transplant.

[0311] Liver function may be measured by methods including but not limited to the Hepatic Quantification test (HepQuant SHUNT), the Child-Pugh Score, the Model for End stage Liver Disease (MELD) score, the Lillie Model, the Acute on Chronic Liver Failure (CLIF-C ACLF) score, the Glasgow Alcoholic Hepatitis Score (GAHS), the International Normalized Ratio (INR) score, the "Prothrombin Time" and other measures of coagulation enzymes, the presence or development of ascites, the presence or development of encephalopathy, platelet count, white blood cell count, mean arterial pressure, blood urea nitrogen (BUN) level, total bilirubin level, indirect bilirubin level, albumin level, alanine aminotransferase (ALT) level, aspartate aminotransferase (AST) level, alkaline phosphatase (ALP) level, and/or sodium creatinine level.

V. Pharmaceutical Combinations

[0312] In some embodiments, a composition comprising a TERT mRNA includes an excipient, or carrier, e.g., an aqueous carrier. A variety of aqueous carriers can be used, e.g., buffered saline. The compositions may contain pharmaceutically acceptable auxiliary substances as those required to approximate physiological conditions such as pH and buffering agents, toxicity countering agents, e.g., sodium acetate, sodium chloride, sodium citrate, potassium chloride, calcium chloride, and sodium lactate. In some embodiments, the pharmaceutical composition comprises 10 mM sodium citrate buffered to pH 6.4. The composition may contain a cryoprotectant, e.g., glycerol, ethylene glycol, propylene glycol, or dimethylsulfoxide (DMSO). The concentration of active agent in these formulations can vary and are selected based on fluid volumes, viscosities, and body weight in accordance with the particular mode of administration selected and the patient's needs (e.g., Remington's Pharmaceutical Science (15th ed., 1980) and Goodman & Gillman, The Pharmacological Basis of Therapeutics (Hardman et al., eds., 1996)).

VI. Methods of Extending Telomeres

[0313] In another aspect, the instant disclosure provides methods of extending telomeres, comprising the step of administering any of the above-described compounds or compositions to a cell with shortened telomeres, wherein telomeres are extended within the cell. The instant disclosure also provides methods of treatment, comprising the step of administering any of the above-described compounds or compositions to an animal subject in need of, or that may benefit from, telomere extension.

[0314] In some embodiments, the compounds or compositions are administered to a cell, wherein the cell is an isolated cell or is part of a cell culture, an isolated tissue culture, an isolated organ, or the like (i.e., the administration is *in vitro*).

[0315] In other embodiments, the compounds or compositions are administered without isolating the cell or cells, the tissue, or the organ from the subject (i.e., the administration is *in vivo*). In some of these embodiments, the compound or composition is delivered to all, or almost all, cells in the subject's body. In some embodiments, the compound or composition is delivered to a specific cell, cell type, tissue, or organ in the subject's body.

[0316] Administration of the compounds or compositions of the instant disclosure may result in the transient expression of a telomerase activity in the cell. The increased activity may be measured by various assays, such as, for example, the telomerase repeat amplification protocol (TRAP) assay. Commercial versions of the TRAP assay are available, for example the Trapeze® telomerase detection kit (Millipore), which provides a sensitive detection and quantitation of telomerase activity, although other measurement techniques are also possible.

[0317] As previously noted, one of the advantages of the instant techniques is that the expression of telomerase activity is transient in the treated cells. In particular, such transient expression is in contrast to previous techniques where a telomerase reverse transcriptase gene persists in an episomal DNA moiety, or is inserted into the genomic sequence of the cell or otherwise permanently modifies the genetic make-up of the targeted cell and results in constitutive activity of the nucleic acid sequence.

[0318] FIG. 1 graphically illustrates some of the advantages of the compounds, compositions, and methods disclosed herein. In particular, the speed of telomere extension made possible with these compounds, compositions, and methods enables telomere maintenance by very infrequent delivery of TERT mRNA. The expressed telomerase activity rapidly extends telomeres in a brief period, before being turned over, thus allowing the protective anti-cancer mechanism of telomere-shortening to function most of the time. Between treatments, normal telomerase activity and telomere shortening is present, and therefore the anti-cancer safety mechanism of telomere shortening to prevent out-of-control proliferation remains intact, while the risk of short telomere-related disease remains low. In contrast, small molecule treatments for extending telomeres may require chronic delivery, and thus present a chronic cancer risk, with minimal therapeutic benefit.

[0319] In some embodiments of the instant methods, the transient expression is independent of cell cycle.

[0320] As noted above, the transient expression of telomerase reverse transcriptase results in the extension of shortened telomeres in treated cells. Telomere length can be

measured using techniques such as terminal restriction fragment (TRF) length analysis, qPCR, MMqPCR, TeSLA, flow FISH, and Q-FISH, as would be understood by one of ordinary skill in the art. In some embodiments, the instant methods increase average telomere length in treated cells by at least 0.1 kb, at least 0.2 kb, at least 0.3 kb, at least 0.4 kb, at least 0.5 kb, at least 1 kb, at least 2 kb, at least 3 kb, at least 4 kb, at least 5 kb, or even more. In some embodiments, the instant methods reduce the percentage of telomeres with lengths below a certain length, for example 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, or more.

[0321] One of the advantages of the instant compounds, compositions, and methods, is the rapidity of extension of telomeres achieved by these techniques. The techniques allow treatments to be brief, and thus the interval between treatments can be long, and thus the treatments can be safe because the normal protective telomere shortening mechanism remains intact for most of the time i.e. between treatments.

[0322] The transient expression of telomerase reverse transcriptase also results in an increased replicative capacity in treated cells. Increased replicative capacity is readily monitored in cells that are approaching replicative senescence by measuring additional population doublings in such cells. Senescent cells do not divide in response to many conditions that cause normal cells to divide, for example passage in culture or treatment with serum. Senescent cells are further often characterized by the expression of pH-dependent P-galactosidase activity, expression of cell cycle inhibitors p53 and p19, and other altered patterns of gene expression, and an enlarged cell size. It is known in the art that, absent treatment with TERT mRNA, certain types of cells (e.g., human lung fibroblast cells) typically double 50-60 times after birth before senescing; with TERT mRNA treatments, however, these cells achieve an additional 16-28 population doublings. If treated again several weeks later, additional proliferative capacity is conferred again. This process of intermittent treatments to periodically re-extend telomeres may be applied additional times, with the interval between treatments depending on factors such as the rate of telomere shortening, the rate of cell divisions, and the amount of telomere extension provided by the treatment. Likewise, human microvascular dermal endothelial cells from an aged individual, absent treatment with the instant compositions, may achieve only 1-2 population doublings, whereas treated cells may achieve 3, 4, or even more population doublings.

[0323] Accordingly, in some embodiments, the instant treatment methods increase the number of population doublings of treated cells.

VII. Therapeutic Kits

[0324] Therapeutic kits comprising a pharmaceutical composition of a TERT mRNA, or sequences thereof (including complementary sequences), and instructions for use are also contemplated herein. In some embodiments, the therapeutic kit comprises devices for administration, including but not limited to syringes, inhalers, nebulizers, and vials or containers.

[0325] In another aspect, the instant disclosure provides ready-to-use kits for use in extending telomeres in a mammalian cell. The kits comprise any of the above-described compounds or compositions, together with instructions for their use. In some embodiments, the kits further comprise

packaging materials. In some embodiments, the packaging materials are air-tight. In these embodiments, the packaging materials may optionally be filled with an inert gas, such as, for example, nitrogen, argon, or the like. In some embodiments, the packaging materials comprise a metal foil container, such as, for example, a sealed aluminum pouch or the like. Such packaging materials are well known by those of ordinary skill in the art. The kit may also comprise a delivery vehicle, such as a lipid as described herein. In some embodiments, one or more components of the formulation are provided frozen with a cryoprotectant, or lyophilized.

[0326] In some embodiments, the kit may further comprise a desiccant, a culture medium, an RNase inhibitor, or other such components. In some embodiments, the kit may further comprise a combination of more than one of these additional components. In some kit embodiments, the composition of the kit is sterile.

ENUMERATED EMBODIMENTS

[0327] The disclosure may be defined by reference to the following enumerated, illustrative embodiments.

[0328] Embodiment 1. A composition comprising a (i) a ribonucleic acid (RNA) encoding telomerase reverse transcriptase (TERT) and (ii) a delivery vehicle, wherein the RNA of (i) comprises one or more modified nucleotides and wherein the delivery vehicle of (ii) is operably-linked to the RNA of (i).

[0329] Embodiment 2. The composition of embodiment 1, wherein the delivery vehicle comprises one or more of a nanoparticle, a liposome, a cationic lipid, an exosome, an extracellular vesicle, a lipid nanoparticle (LNP), a natural lipoprotein particle and an artificial lipoprotein particle.

[0330] Embodiment 3. The composition of embodiment 1, wherein the delivery vehicle comprises a lipid nanoparticle (LNP).

[0331] Embodiment 4. The composition of embodiment 1, wherein the delivery vehicle comprises an ionizable lipid nanoparticle.

[0332] Embodiment 5. The composition of any one of embodiments 1-4, wherein the delivery vehicle comprises a targeting moiety.

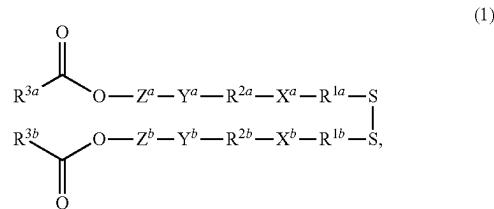
[0333] Embodiment 6. The composition of embodiment 5, wherein the delivery vehicle specifically or selectively interacts with a liver cell.

[0334] Embodiment 7. The composition of embodiment 5, wherein the targeting moiety is a lipid, a peptide, and/or an antibody.

[0335] Embodiment 8. The composition of embodiment 3, wherein the LNP comprises an ionizable lipid, a phospholipid, a cholesterol, and/or a PEGylated lipid.

[0336] Embodiment 9. The composition of embodiment 8, wherein the LNP comprises a molar ratio of about 50 to about 60 moles of an ionizable lipid, about 4 to about 6 moles of a phospholipid, about 35 to about 45 moles of cholesterol, and about 1 to about 2 moles of PEGylated lipid.

[0337] Embodiment 10. The composition of any one of embodiments 1-9, wherein the delivery vehicle comprises a compound of Formula I:

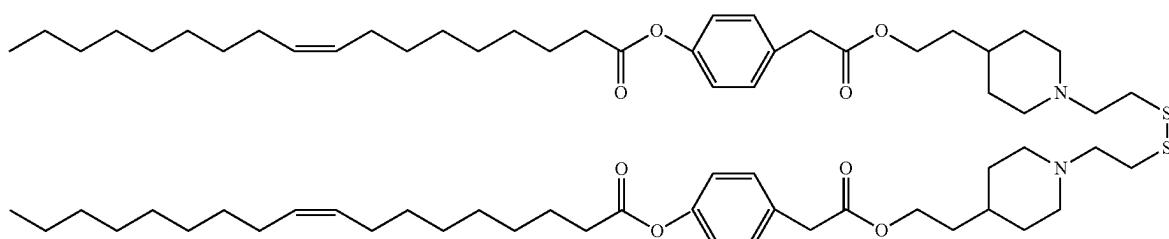


[0338] wherein R^{1a} and R^{1b} each independently represents an alkylene group having 1 to 6 carbon atoms, wherein X^a and X^b are each independently an acyclic alkyl tertiary amino group having 1 to 6 carbon atoms and 1 tertiary amino group, or 2 to 5 carbon atoms, and A cyclic alkylene tertiary amino group having 1 to 2 tertiary amino groups, wherein R^{2a} and R^{2b} each independently represent an alkylene group having 8 or less carbon atoms or an oxydialkylene group, wherein Y^a and Y^b each independently represent an ester bond, an amide bond, a carbamate bond, an ether bond or a urea bond;

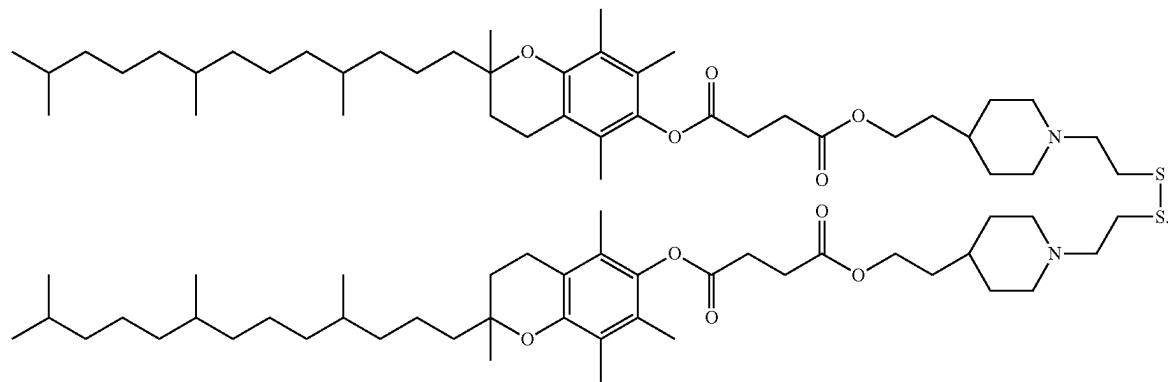
[0339] wherein Z^a and Z^b are each independently a divalent group derived from an aromatic compound having 3 to 16 carbon atoms, having at least one aromatic ring, and optionally having a hetero atom, and

[0340] wherein R^{3a} and R^{3b} each independently represent a residue derived from a reaction product of a fat-soluble vitamin having a hydroxyl group and succinic anhydride or glutaric anhydride, or a sterol derivative having a hydroxyl group and succinic anhydride or a residue derived from a reaction product with glutaric anhydride or an aliphatic hydrocarbon group having 12 to 22 carbon atoms.

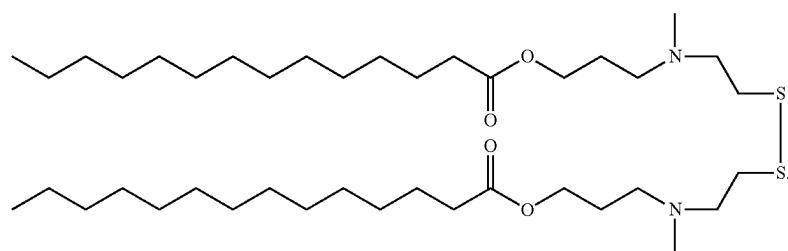
[0341] Embodiment 11. The composition of embodiment 10, wherein the compound of Formula I is:



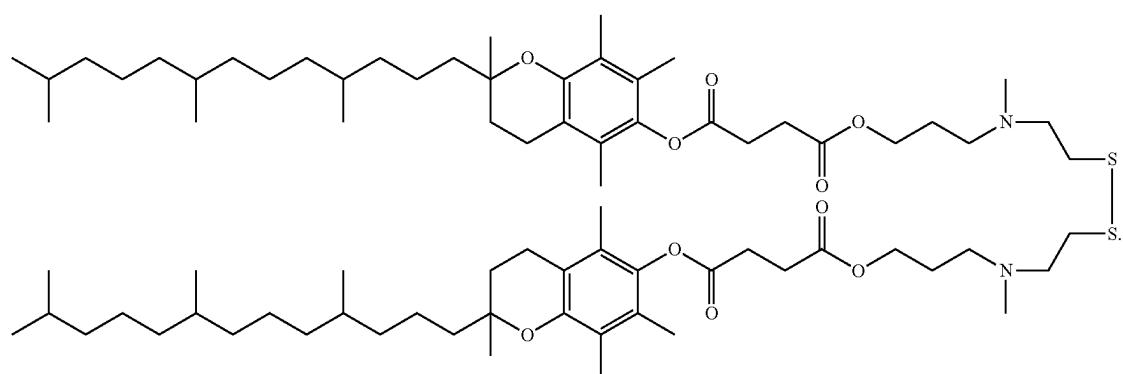
[0342] Embodiment 12. The composition of embodiment 10, wherein the compound of Formula I is:



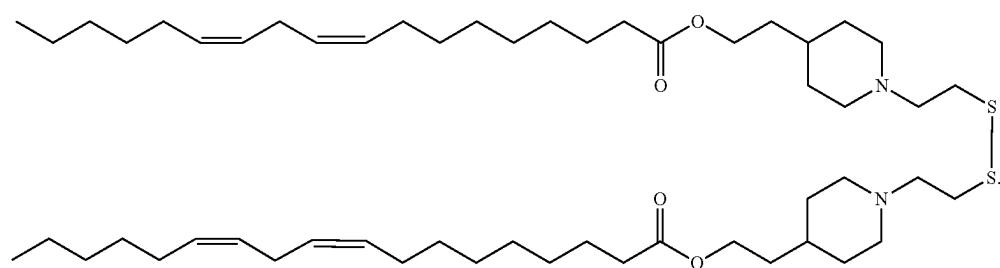
[0343] Embodiment 13. The composition of embodiment 10, wherein the compound of Formula I is:



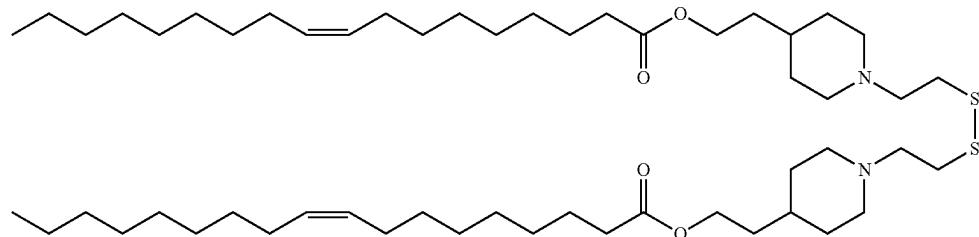
[0344] Embodiment 14. The composition of embodiment 10, wherein the compound of Formula I is:



[0345] Embodiment 15. The composition of embodiment 10, wherein the compound of Formula I is:



[0346] Embodiment 16. The composition of embodiment 10, wherein the compound of Formula I is:



[0347] Embodiment 17. The composition of any one of embodiments 1-16, wherein the RNA comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to any one of SEQ ID NOS: 1-5, 30-31, or 37-40.

[0348] Embodiment 18. The composition of embodiment 17, wherein the RNA comprises a 5' cap.

[0349] Embodiment 19. The composition of embodiment 18, wherein the 5'cap comprises an anti-reverse cap analog (ARCA).

[0350] Embodiment 20. The composition of embodiment 19, wherein the ARCA comprises a 3'-O-Me-m7G(5')ppp(5')G structure.

[0351] Embodiment 21. The composition of embodiment 18, wherein the 5' cap comprise' m7(3'O'eG)(5'ppp'5')(2'OMeA)pG.

[0352] Embodiment 22. The composition of any one of embodiments 1-21, wherein the RNA further comprises at least one untranslated region (UTR).

[0353] Embodiment 23. The composition of embodiment 22, wherein the at least one UTR is positioned 5' to the RNA of (i).

[0354] Embodiment 24. The composition of embodiment 22, wherein the at least one UTR is positioned 3' to the RNA of (i).

[0355] Embodiment 25. The composition of any one of embodiments 22-24, wherein the UTR comprises a human sequence.

[0356] Embodiment 26. The composition of any one of embodiments 22-24, wherein the UTR comprises a non-human sequence.

[0357] Embodiment 27. The composition of any one of embodiments 22-26, wherein the UTR comprises a chimeric sequence.

[0358] Embodiment 28. The composition of embodiment 27, wherein the chimeric sequence increases stability, increases a transcription rate or decreases a time until initiation of transcription of the RNA of (i).

[0359] Embodiment 29. The composition of any one of embodiments 22-28, wherein the UTR comprises a sequence having at least 70% identity to a UTR sequence isolated or derived from one or more of α -globin, β -globin, c-fos, and a tobacco etch virus.

[0360] Embodiment 30. The composition of any one of embodiments 1-29, wherein the one or more modified nucleotides of the RNA of (i) comprise one or more of a modified adenine or analog thereof, a modified cytidine or analog thereof, a modified guanosine or analog thereof, and a modified uridine or analog thereof.

[0361] Embodiment 31. The composition of any one of embodiments 1-30, wherein the one or more modified nucleotides of the RNA of (i) comprise one or more of 1-methylpseudo尿idylate, pseudo尿idylate, 2-thiouridylate, and 5-methylcytidylate.

[0362] Embodiment 32. The composition of any one of embodiments 1-31, wherein the one or more modified nucleotides of the RNA of (i) comprise 5-methoxyuridylate (5-moU).

[0363] Embodiment 33. The composition of any one of embodiments 1-32, wherein the one or more modified nucleotides of the RNA of (i) comprise one or more of m1A 1-methyladenosine, m6A N6-methyladenosine, Am 2'-O-methyladenosine, i6A N6-isopentenyladenosine, io6A N6-(cis-hydroxyisopentenyl)adenosine, ms2io6A 2-methylthio-N6-(cis-hydroxyisopentenyl) adenosine, g6A N6-glycylcarbamoyladenine, t6A N6-threonylcarbamoyladenine, ms2t6A 2-methylthio-N6-threonyl carbamoyladenine, Ar(p) 2'-O-ribosyladenosine (phosphate), m6 2A N6,N6-dimethyladenosine, m6Am N6,2'-O-dimethyladenosine, m6 2Am N6,N6,2'-O-trimethyladenosine, m1Am 1,2'-O-dimethyladenosine, m3C 3-methylcytidine, m5C 5-methylcytidine, Cm 2'-O-methylcytidine, ac4C N4-acetylcytidine, f5C 5-formylcytidine, m4C N4-methylcytidine, hm5C 5-hydroxymethylcytidine, f5Cm 5-formyl-2'-O-methylcytidine, m1G 1-methylguanosine, m2G N2-methylguanosine, m7G 7-methylguanosine, Gm 2'-O-methylguanosine, m2 2G N2,N2-dimethylguanosine, Gr(p) 2'-O-ribosylguanosine (phosphate), yW wybutoxine, o2yW peroxywybutoxine, OHyW hydroxywybutoxine, OHyW* undermodified hydroxywybutoxine, imG wyosine, m2,7G N2,7-dimethylguanosine, m2,2,7G N2,N2,7-trimethylguanosine I inosine, m1I 1-methylinosine, Im 2'-O-methylinosine, Q queuosine, galQ galactosyl-queuosine, manQmannosyl-queuosine, Ψ pseudouridine, D dihydrouridine, m5U 5-methyluridine, Um 2'-O-methyluridine, m5Um 5,2'-O-dimethyluridine, m1 Ψ 1-methylpseudo尿idylate, Ψ m 2'-O-methylpseudo尿idylate, s2U 2-thiouridine, ho5U 5-hydroxyuridine, chm5U 5-(carboxyhydroxymethyl)uridine, mchm5U 5-(carboxyhydroxymethyl)uridine, methyl ester mcm5U 5-methoxycarbonylmethyluridine, mcm5Um 5-methoxycarbonylmethyl-2'-O-methyluridine, mcm5s2U 5-methoxycarbonylmethyl-2-thiouridine, ncm5U 5-carbamoylmethyl-2'-O-methyluridine, cmnm5U 5-carboxymethylaminomethyluridine, m3U 3-methyluridine, m1acp3 Ψ 1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine, cm5U 5-carboxymethyluridine, m3Um 3,2'-O-dimethyluridine, m5D 5-methyldihydrouridine, tm5U 5-taurinomethyluridine, tm5s2U 5-taurinom-

ethyl-2-thiouridine, 2-Aminoadenosine, 2-Amino-6-chloropurineriboside, 8-Azaadenosine, 6-Chloropurineriboside, 5-Iodocytidine, 5-Iodouridine, Inosine, 2'-O-Methylinosine, Xanthosine, 4-Thiouridine, 06-Methylguanosine, 5,6-Dihydouridine, 2-Thiocytidine, 6-Azacytidine, 6-Azauridine, 2'-O-Methyl-2-aminoadenosine, 2'-O-Methylpseudouridine, N1-Methyladenosine, 2'-O-Methyl-5-methyluridine, 7-Deazaguanosine, 8-Azidoadenosine, 5-Bromocytidine, 5-Bromouridine, 7-Deazaadenosine, 5-Aminoallyluridine, 5-Aminoallylcytidine, 8-Oxoguanosine, 2-Aminopurine-riboside, Pseudoisocytidine, N1-Methylpseudouridine, 5,6-Dihydro-5-Methyluridine, N6-Methyl-2-Aminoadenosine, 5-Carboxycytidine, 5-Hydroxymethyluridine, Thienoguanosine, 5-Hydroxycytidine, 5-Formyluridine, 5-Carboxyuridine, 5-Methoxyuridine, 5-Methoxycytidine, Thienouridine, 5-Carboxymethylesteruridine, Thienocytidine, 8-Oxaadenosine, Isoguanosine, N1-Ethylpseudouridine, N1-Methyl-2'-O-Methylpseudouridine, N1-Methoxymethylpseudouridine, N1-Propylpseudouridine, 2'-O-Methyl-N6-Methyladenosine, 2-Amino-6-Cl-purine-2'-deoxyriboside, 2-Amino-2'-deoxyadenosine, 2-Aminopurine-2'-deoxyriboside, 5-Bromo-2'-deoxycytidine, 5-Bromo-2'-deoxyuridine, 6-Chloropurine-2'-deoxyriboside, 7-Deaza-2'-deoxyadenosine, 7-Deaza-2'-deoxymanosine, 2'-Deoxyinosine, 5-Propynyl-2'-deoxycytidine, 5-Propynyl-2'-deoxyuridine, 5-Fluoro-2'-deoxycytidine, 5-Iodo-2'-deoxycytidine, 5-Iodo-2'-deoxyuridine, N6-Methyl-2'-deoxyadenosine, 5-Methyl-2'-deoxycytidine, 06-Methyl-2'-deoxymanosine, N2-Methyl-2'-deoxymanosine, 8-Oxo-2'-deoxyadenosine, 8-Oxo-2'-deoxymanosine, 2-Thiothymidine, 2'-Deoxy-P-nucleoside, 5-Hydroxy-2'-deoxycytidine, 4-Thiothymidine, 2-Thio-2'-deoxycytidine, 6-Aza-2'-deoxyuridine, 6-Thio-2'-deoxymanosine, 8-Chloro-2'-deoxyadenosine, 5-Aminoallyl-2'-deoxycytidine, 5-Aminoallyl-2'-deoxyuridine, N4-Methyl-2'-deoxycytidine, 2'-Deoxyzebaraline, 5-Hydroxymethyl-2'-deoxyuridine, 5-Hydroxymethyl-2'-deoxycytidine, 5-Propargylamino-2'-deoxycytidine, 5-Propargylamino-2'-deoxyuridine, 5-Carboxy-2'-deoxycytidine, 5-Formyl-2'-deoxycytidine, 5-[3-Indolyl]propionamide-N-allyl]-2'-deoxyuridine, 5-Carboxy-2'-deoxyuridine, 5-Formyl-2'-deoxyuridine, 7-Deaza-7-Propargylamino-2'-deoxyadenosine, 7-Deaza-7-Propargylamino-2'-deoxymanosine, Biotin-16-Aminoallyl-2'-dUTP, Biotin-16-Aminoallyl-2'-dCTP, Biotin-16-Aminoallylcytidine, N4-Biotin-OBEA-2'-deoxycytidine, Biotin-16-Aminoallyluridine, Dabcyll-5-3-Aminoallyl-2'-dUTP, Desthiobiotin-6-Aminoallyl-2'-deoxycytidine, Desthiobiotin-16-Aminoallyl-Uridine, Biotin-16-7-Deaza-7-Propargylamino-2'-deoxymanosine, Cyanine 3-5-Propargylamino-2'-deoxycytidine, Cyanine 3-6-Propargylamino-2'-deoxyuridine, Cyanine 5-6-Propargylamino-2'-deoxycytidine, Cyanine 5-6-Propargylamino-2'-deoxyuridine, Cyanine 3-Aminoallylcytidine, Cyanine 3-Aminoallyluridine, Cyanine 5-Aminoallylcytidine, Cyanine 5-Aminoallyluridine, Cyanine 7-Aminoallyluridine, 2'-Fluoro-2'-deoxyadenosine, 2'-Fluoro-2'-deoxycytidine, 2'-Fluoro-2'-deoxymanosine, 2'-Fluoro-2'-deoxyuridine, 2'-O-Methyladenosine, 2'-O-Methylcytidine, 2'-O-Methylguanosine, 2'-O-Methyluridine, Puromycin, 2'-Amino-2'-deoxycytidine, 2'-Amino-2'-deoxyuridine, 2'-Azido-2'-deoxycytidine, 2'-Azido-2'-deoxyuridine, Aracytidine, Arauridine, 2'-Azido-2'-deoxyadenosine, 2'-Amino-2'-deoxyadenosine, Araadenosine, 2'-Fluoro-thymidine, 3'-O-

Methyladenosine, 3'-O-Methylcytidine, 3'-O-Methylguanosine, 3'-O-Methyluridine, 2'-Azido-2'-deoxymanosine, Araguanosine, 2'-Deoxyuridine, 3'-O-(2-nitrobenzyl)-2'-Deoxyadenosine, 3'-O-(2-nitrobenzyl)-2'-Deoxynosine, 3'-Deoxyadenosine, 3'-Deoxymanosine, 3'-Deoxycytidine, 3'-Deoxy-5-Methyluridine, 3'-Deoxuryidine, 2',3'-Dideoxyadenosine, 2',3'-Dideoxymanosine, 2',3'-Dideoxyuridine, 2',3'-Dideoxythymidine, 2',3'-Dideoxycytidine, 3'-Azido-2',3'-dideoxyadenosine, 3'-Azido-2',3'-dideoxythymidine, 3'-Amino-2',3'-dideoxyadenosine, 3'-Amino-2',3'-dideoxycytidine, 3'-Amino-2',3'-dideoxymanosine, 3'-Azido-2',3'-dideoxyuridine, 3'-Azido-2',3'-dideoxythymidine, 3'-Azido-2',3'-dideoxycytidine, 3'-Azido-2',3'-dideoxyuridine, 3'-Azido-2',3'-dideoxymanosine, 2'-Deoxyadenosine-5'-O-(1-Thiophosphate), 2'-Deoxycytidine-5'-O-(1-Thiophosphate), 2'-Deoxythymidine-5'-O-(1-Thiophosphate), Adenosine-5'-O-(1-Thiophosphate), Cytidine-5'-O-(1-Thiophosphate), Guanosine-5'-O-(1-Thiophosphate), Uridine-5'-O-(1-Thiophosphate), 2',3'-Dideoxyadenosine-5'-O-(1-Thiophosphate), 2',3'-Dideoxycytidine-5'-O-(1-Thiophosphate), 2',3'-Dideoxymanosine-5'-O-(1-Thiophosphate), 3'-Deoxythymidine-5'-O-(1-Thiophosphate), 3'-Azido-2',3'-dideoxythymidine-5'-O-(1-Thiophosphate), 2',3'-Dideoxyuridine-5'-O-(1-Thiophosphate), 2'-Deoxyadenosine-5'-O-(1-Boranophosphate), 2'-Deoxycytidine-5'-O-(1-Boranophosphate), 2'-Deoxyguanosine-5'-O-(1-Boranophosphate), and 2'-Deoxythymidine-5'-O-(1-Boranophosphate).

[0364] Embodiment 34. The composition of any one of embodiments 1-33, wherein the composition further comprises a ribonucleic acid (RNA) encoding Telomerase RNA Component (TERC).

[0365] Embodiment 35. The composition of any one of embodiments 1-34, wherein the delivery vehicle comprises the RNA encoding TERT.

[0366] Embodiment 36. The composition of embodiment 35, wherein the RNA encoding TERT comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to any one of SEQ ID NOS: 11-5, 7, 9, 14-17, 19, 21, 23, 25, 27, 29-31, 37-40.

[0367] Embodiment 37. The composition of embodiment 35, wherein the RNA encoding TERT comprises a full length or part thereof, of a UTR of one of SEQ ID NOS: 32-34, 35, and 36.

[0368] Embodiment 38. The composition of any one of embodiments 1-36, wherein the RNA comprises a self-replicating RNA.

[0369] Embodiment 39. The composition of any one of embodiments 1-38, wherein the RNA comprises a circular RNA.

[0370] Embodiment 40. The composition of embodiment 8, wherein the layer comprises a lipid monolayer or lipid bi-layer.

[0371] Embodiment 41. The composition of embodiment 41, wherein the delivery vehicle comprises an internal volume.

[0372] The composition of any one of embodiments 1-39, wherein the delivery vehicle is operably-linked to a ribonucleic acid (RNA) encoding Telomerase RNA Component (TERC).

- [0373] Embodiment 41. The composition of embodiment 40, wherein the delivery vehicle comprises the RNA encoding TERC.
- [0374] Embodiment 42. The composition of embodiment 35, wherein one or more of a surface, a layer or a volume of the delivery vehicle comprises the RNA encoding TERC.
- [0375] Embodiment 43. The composition of embodiment 42, wherein the surface comprises an outer surface or an inner surface.
- [0376] Embodiment 44. The composition of embodiment 42, wherein the layer comprises a lipid monolayer or lipid bi-layer.
- [0377] Embodiment 45. The composition of embodiment 42, wherein the volume comprises an internal volume.
- [0378] Embodiment 46. A method of increasing telomerase activity in a cell, the method comprising contacting the cell and the composition of any one of embodiments 1-45.
- [0379] Embodiment 47. A method of extending telomeres in a cell, the method comprising contacting the cell and the composition of any one of embodiments 1-45.
- [0380] Embodiment 48. The method of embodiment 46 or 47, wherein the cell is *in vivo*, *ex vivo* or *in vitro*.
- [0381] Embodiment 49. A cell comprising the composition of any one of embodiments 1-45.
- [0382] Embodiment 50. A formulation comprising the cell of embodiment 49.
- [0383] Embodiment 51. The formulation of embodiment 50, wherein a plurality of cells comprises the cell of embodiment 29.
- [0384] Embodiment 52. The formulation of embodiment 51, wherein each cell of the plurality is a cell according to embodiment 49.
- [0385] Embodiment 53. A method of treating a disease or disorder comprising administering to a subject an effective amount of a composition according to any one of embodiments 1-45.
- [0386] Embodiment 54. A method of treating a disease or disorder comprising administering to a subject an effective amount of a cell according to embodiment 49.
- [0387] Embodiment 55. A method of treating a disease or disorder comprising administering to a subject an effective amount of a formulation according to any one of embodiments 50-52.
- [0388] Embodiment 56. A method of delaying the onset of a disease comprising administering to a subject an effective amount of a composition according to any one of embodiments 1-45.
- [0389] Embodiment 57. A method of delaying the onset of a disease comprising administering to a subject an effective amount of a cell according to embodiment 49.
- [0390] Embodiment 58. A method of delaying the onset of a disease comprising administering to a subject an effective amount of a formulation according to any one of embodiments 50-52.
- [0391] Embodiment 59. A method of treating a fibrotic disease in a subject in need thereof, comprising: administering to the subject an effective amount of a composition comprising one or more synthetic messenger ribonucleic acids (mRNAs) encoding telomerase reverse transcriptase (TERT).
- [0392] Embodiment 60. The method of embodiment 59, wherein the composition comprises a delivery vehicle.
- [0393] Embodiment 61. The method of embodiment 60, wherein the delivery vehicle is a nanoparticle.
- [0394] Embodiment 62. The method of embodiment 61, wherein the nanoparticle is a lipid nanoparticle (LNP).
- [0395] Embodiment 63. The method of embodiment 62, wherein the LNP comprises an ionizable lipid, a phospholipid, a cholesterol, and/or a PEGylated lipid.
- [0396] Embodiment 64. The method of embodiment 63, wherein the LNP comprises a molar ratio of about 50 to about 60 moles of an ionizable lipid, to about 4 to about 6 moles of a phospholipid, about 35 to about 45 moles of cholesterol, and about 1.0 to about 2.0 moles of PEGylated lipid.
- [0397] Embodiment 65. The method of embodiment 63, wherein the LNP comprises a molar ratio of about 30 to 40 moles of an ionizable lipid, to about 14 to about 18 moles of a phospholipid, about 40 to about 50 moles of a cholesterol, and about 2.0 to about 3.0 moles of a PEGylated lipid.
- [0398] Embodiment 66. The method of any one of embodiments 59-65, wherein the TERT synthetic mRNA comprises at least one modified nucleoside from the list in Table 1B.
- [0399] Embodiment 67. The method of embodiment 66, wherein the modified nucleoside is pseudouridine or a pseudouridine analog.
- [0400] Embodiment 68. The method of embodiment 67, wherein the pseudouridine analog is N-1-methylpseudouridine.
- [0401] Embodiment 69. The method of embodiment 66, wherein the modified nucleoside is 5-methoxyuridine.
- [0402] Embodiment 70. The method of any one of embodiments 59-69, wherein the TERT synthetic mRNA comprises an untranslated region (UTR).
- [0403] Embodiment 71. The method of embodiment 70, wherein the UTR comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to any one of SEQ ID NOS: 32-36.
- [0404] Embodiment 72. The method of any one of embodiments 59-71, wherein the TERT synthetic mRNA comprises a 5' cap structure, wherein the 5' cap structure is IRES, Cap0, Cap1, ARCA, inosine, N1-methylguanosine, 2'fluoro-guanosine, 7-deaza-guanosine, CleanCap™, m7(3'O'eG)(5'“ppp’5’)(2'OMeA)pG, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, 2-azido-guanosine, Cap2, Cap4, CAP-003, or CAP-225.
- [0405] Embodiment 73. The method of any one of embodiments 59-72, wherein the TERT synthetic mRNA comprises a poly-adenosine (poly-A) nucleotide sequence 3' to the encoding region.
- [0406] Embodiment 74. The method of any one of embodiments 59-73, wherein the TERT synthetic mRNA comprises a chain terminating nucleotide, wherein the nucleotide' is 3'-deoxyadenosine (cordycepin), 3'-deoxyurid'ne, 3'-deoxycytos'ne, 3'-deoxyguanos'ne, 3'-deoxythym'ne' 2',3'-dideoxynucleosi'es' 2',3'-dideoxyadenos'ne' 2',3'-dideoxyurid'ne' 2',3'-dideoxycytos'ne' 2',3'-deoxyguanos'ne' 2',3'-dideoxythym'ne', a 2'-deoxynucleoside, or —O— methylnucleoside.
- [0407] Embodiment 75. The method of any one of embodiments 59-74, wherein the TERT synthetic mRNA is codon optimized.
- [0408] Embodiment 76. The method of any one of embodiments 59-73, wherein the TERT synthetic mRNA comprises a sequence of a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%,

at least 97%, at least 99%, or 100% identical to any one of SEQ ID NOS: 1, 2, 7, 9, 30, 39, or 40.

[0409] Embodiment 77. The method of embodiment 60, wherein the delivery vehicle is a liposome, an ionizable lipid, or an exosome.

[0410] Embodiment 78. The method of embodiment 77, wherein the delivery vehicle is an exosome, and wherein the exosome comprises a targeting moiety of one or more of a lipid, a peptide, or an antibody.

[0411] Embodiment 79. The method of any one of embodiments 59-78, wherein the method reduces fibrosis.

[0412] Embodiment 80. The method of any one of embodiments 59-79, wherein the subject is human.

[0413] Embodiment 81. A composition according to any one of embodiments 1-58 for use in a method according to any one of embodiments 59-80.

[0414] Embodiment 82. The composition for use of embodiment 81, wherein the composition is a pharmaceutical composition comprising one or more pharmaceutically acceptable solvents or excipients.

[0415] Embodiment 83. A kit for treating a fibrotic disease in a subject, the kit comprising a composition according to any one of embodiment 1-58, and instructions for use thereof.

[0416] Embodiment 84. A method of treating a liver disease in a subject in need thereof, comprising: administering to the subject a composition comprising one or more synthetic messenger ribonucleic acids (mRNAs) encoding telomerase reverse transcriptase (TERT).

[0417] Embodiment 85. The method of embodiment 84, wherein the composition comprises a delivery vehicle.

[0418] Embodiment 86. The method of embodiment 85, wherein the delivery vehicle is a nanoparticle.

[0419] Embodiment 87. The method of embodiment 86, wherein the nanoparticle is a lipid nanoparticle (LNP).

[0420] Embodiment 88. The method of embodiment 87, wherein the LNP comprises an ionizable lipid, a phospholipid, a cholesterol, and/or a PEGylated lipid.

[0421] Embodiment 89. The method of embodiment 88, wherein the LNP comprises a molar ratio of about 50 to about 60 moles of an ionizable lipid, to about 4 to about 6 moles of a phospholipid, about 35 to about 45 moles of cholesterol, and about 1 to about 2 moles of PEGylated lipid.

[0422] Embodiment 90. The method of embodiment 88, wherein the LNP comprises a molar ratio of about 55 moles of an ionizable lipid, to about 5 moles of a phospholipid, about 40 moles of a cholesterol, and about 1.5 moles of a PEGylated lipid.

[0423] Embodiment 91. The method of any one of embodiments 84-90, wherein the TERT synthetic mRNA comprises at least one modified nucleoside from the list in Table 1B.

[0424] Embodiment 92. The method of embodiment 91, wherein the modified nucleoside is pseudouridine or a pseudouridine analog.

[0425] Embodiment 93. The method of embodiment 92, wherein the pseudouridine analog is N-1-methylpseudouridine.

[0426] Embodiment 94. The method of embodiment 91, wherein the modified nucleoside is 5-methoxyuridine.

[0427] Embodiment 95. The method of any one of embodiments 84-94, wherein the TERT synthetic mRNA comprises an untranslated region (UTR).

[0428] Embodiment 96. The method of embodiment 95, wherein the UTR comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to any one of SEQ ID NOS: 32-36.

[0429] Embodiment 97. The method of any one of embodiments 84-96, wherein the TERT synthetic mRNA comprises a 5' cap structure, wherein the 5' cap structure is IRES, Cap0, Cap1, ARCA, inosine, N1-methylguanosine, 2'fluoro-guanosine, 7-deaza-guanosine, Clean-Cap™, m7(3'O'eG)(5'ppp'5')(2'OMeA)pG, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, 2-azido-guanosine, Cap2, Cap4, CAP-003, or CAP-225.

[0430] Embodiment 98. The method of any one of embodiments 84-97, wherein the TERT synthetic mRNA comprises a poly-adenosine (poly-A) nucleotide sequence 3' to the encoding region.

[0431] Embodiment 99. The method of any one of embodiments 84-98, wherein the TERT synthetic mRNA comprises a chain terminating nucleotide, wherein the nucleotide' is 3'-deoxyadenosine (cordycepin), 3'-deoxyurid'ne, 3'-deoxycytos'ne, 3'-deoxyguanos'ne, 3'-deoxythym'ne' 2',3'-dideoxynucleosi'es' 2',3'-dideoxyadenos'ne' 2',3'-dideoxyurid'ne' 2',3'-dideoxycytos'ne' 2',3'-deoxyguanosine' 2',3'-dideoxythym'ne, a 2'-deoxynucleoside, or —O— methylnucleoside.

[0432] Embodiment 100. The method of any one of embodiments 84-99, wherein the TERT synthetic mRNA is codon optimized.

[0433] Embodiment 101. The method of any one of embodiments 84-99, wherein the TERT synthetic mRNA comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to any one of SEQ ID NOS: 1, 2, 7, 9, 30, 39, or 40.

[0434] Embodiment 102. The method of embodiment 85, wherein the delivery vehicle is a liposome, a cationic lipid, or an exosome.

[0435] Embodiment 103. The method of any one of embodiments 84-102, wherein the method reduces liver fibrosis.

[0436] Embodiment 104. The method of any one of embodiments 84-103, wherein the liver disease is non-alcoholic steatohepatitis (NASH) or non-alcoholic fatty liver disease (NAFLD).

[0437] Embodiment 105. The method of any one of embodiments 84-103, wherein the liver disease is alcoholic hepatitis.

[0438] Embodiment 106. The method any one of embodiments 84-103, wherein the liver disease is liver cirrhosis or liver fibrosis.

[0439] Embodiment 107. The method of any one of embodiments 84-103, wherein the liver disease is compensated cirrhosis, decompensated cirrhosis, or acute-on-chronic liver failure.

[0440] Embodiment 108. The method of any one of embodiments 84-103, wherein the liver disease is fibrotic stage F4 Non-alcoholic steatohepatitis (NASH).

[0441] Embodiment 109. The method of any one of embodiments 84-103, wherein the liver disease is biliary atresia, primary biliary cirrhosis, primary sclerosing cholangitis, and/or chronic liver disease.

[0442] Embodiment 110. The method of any one of embodiments 84-103, wherein the liver disease is hemochromatosis, Wilson's disease, or ischemic hepatitis.

[0443] Embodiment 111. The method of any one of embodiments 84-110, wherein the subject is human.

[0444] Embodiment 112. A composition according to any one of embodiments 1-58 for use in a method according to any one of embodiments 84-111.

[0445] Embodiment 113. The composition for use of embodiment 112, wherein the composition is a pharmaceutical composition comprising one or more pharmaceutically acceptable solvents or excipients.

[0446] Embodiment 114. A kit for treating a liver disease in a subject, the kit comprising the composition according to any one of embodiments 1-58, and instructions for use thereof.

EXAMPLES

[0447] The following examples are included for illustrative purposes and are not intended to limit the scope of the disclosure.

Example 1: LNP Formulations for mRNA Expression in Liver

[0448] This Example demonstrates that three diverse lipid nanoparticle (LNP) formulations may be used to deliver an mRNA encoding a heterologous protein to the liver of a subject animal. An LNP formulation that includes an SS-OP lipid resulted in the highest levels of expression and/or activity of the heterologous protein. In this Example, intravenous administration was used. Furthermore, the SS-OP-based formulation (LNP1) caused less tissue toxicity than a cKK-E12-based formulation (LNP2). These experiments demonstrate expression of the report Luciferase or the therapeutic gene TERT in the liver of subject animals. In particular, an SS-OP-based formulation is here shown to cause high TERT expression and/or activity with low toxicity.

[0449] Table 5 shows illustrative lipid nanoparticle (LNP) formulations targeting the liver in total lipid/mRNA ratios by weight/weight (wt/wt).

TABLE 5

Compound	Molar ratio	Total lipid:mRNA ratio wt/wt
<u>LNP1</u>		
SS-OP	55	42
DOPC	5	
Cholesterol	40	
DMG-PEG2000	1.5	
<u>LNP2</u>		
cKK-E12	35	20.5
DOPE	16	
Cholesterol	46.5	
14:0 PEG2000	2.5	
PE		
<u>LNP3</u>		
DLin-MC3-DMA	50	35
DSPC	10	
Cholesterol	40	
DMG-PEG2000	1.5	

TABLE 5-continued

Compound	Molar ratio	Total lipid:mRNA ratio wt/wt
<u>LNP4</u>		
SS-OP	35	38.3
DOPE	16	
Cholesterol	46.5	
14:0 PEG2000	2.5	
PE		
<u>LNP5</u>		
cKK-E12	55	14.5
DOPC	5	
Cholesterol	40	
DMG-PEG2000	1.5	

[0450] Compositions and methods of the disclosure may be used for the treatment of cirrhosis. In some embodiments, compositions and/or methods of use of compositions of the disclosure intended for treatment of cirrhosis induce TERT expression or increase TERT activity in a liver cell. In some embodiments, compositions and/or methods of use of compositions of the disclosure intended for treatment of cirrhosis do not induce cellular, tissue or systemic toxicity. Compositions may be administered systemically, e.g., intravenously.

[0451] FIG. 2 is a series of graphs showing that mRNA LNPs exhibit low toxicity by liver panel. Mice were dosed intravenously with GRP or CRE mRNA encapsulated in a lipid nanoparticle employing either LNP1 (comprising SS-OP™) or LNP2 (comprising cKK-E12) (N=1-4 per condition). Mice were sacrificed and blood was collected at the time points indicated (12, 24, and 72 hours). Mice receiving saline (N=4) and carbon tetrachloride (CCl₄, N=4) served as negative and positive controls, respectively. Error bars display standard error of the mean.

[0452] FIG. 3 is a series of photographs showing that intravenous delivery of TERT mRNA LNPs does not result in abnormal histology. 11 µg Cre mRNA was encapsulated into LNP1 and delivered intravenously (i.v.) into tdTomato fl/fl mice. Organs were harvested 72 hours later, fixed, paraffin embedded, and sectioned. Organs from an untreated tdTomato fl/fl mouse are shown for reference.

[0453] FIG. 4 is a series of photographs showing that TERT mRNA LNPs transfect hepatocytes with high efficiency. 11 µg Cre mRNA was encapsulated into LNP1 and delivered intravenously (i.v.) into tdTomato fl/fl mice. Organs were harvested 72 hours later, fixed, paraffin embedded, and sectioned. Photographs depict immunohistochemistry (IHC) with anti-tdTomato. Organs from an untreated tdTomato fl/fl mouse are shown for reference.

[0454] FIG. 5 is a series of photographs showing that TERT mRNA LNPs also target some cells in spleen, particularly in the red pulp area. 11 µg Cre mRNA was encapsulated into LNP1 and delivered intravenously (i.v.) into tdTomato fl/fl mice. Organs were harvested 72 hours later, fixed, paraffin embedded, and sectioned. Photographs depict immunohistochemistry (IHC) with anti-tdTomato. Organs from an untreated tdTomato fl/fl mouse are shown for reference.

[0455] FIG. 6 is a pair of graphs showing that TERT mRNA LNPs cause high telomerase activity in liver. Tert mRNA (SEQ ID NO: 37) was formulated with LNP1 or LNP2 and delivered intravenously in a concentration of 0.6 mg/kg into TERT KO mice. 20 hours later, the livers were

harvested for TRAP. Wild-type C57B16/J and untreated TERT KO mouse livers were used as positive and negative controls, respectively.

[0456] The TRAP assay uses lysate from cells or tissues incubated with an artificial telomere (DNA oligonucleotide) to detect telomerase. If active telomerase is present, it extends the artificial telomere 6 base pairs (bp) at a time, producing a ladder pattern. This extension reaction is amplified by PCR and run on a gel (in this case Agilent bioanalyzer, a microfluidic agarose gel). The presence of a ladder in 6 bp increments indicates telomerase activity.

[0457] FIG. 7 is a photograph demonstrating that exemplary LNP formulations deliver luciferase (LUC) mRNA to the liver, as demonstrated by the high bioluminescence signals. Shown are LNP1 (comprising SS-OPT™), LNP2 (comprising cKK), and LNP3 (comprising DLin-MC3-DMA). An empty LNP formulation is also shown as a negative control (ctrl). Luciferase mRNA was formulated with the aforementioned LNPs 1, 2, and 3 and delivered intravenously into C57B16/J mice. 20 hours later, these mice were shaved and imaged after luciferin injection using an IVIST™ Bioluminescence imaging system.

[0458] FIG. 20 shows in vivo delivery of mRNA in a photograph depicting the results demonstrating that luciferase mRNA LNPs causing high bioluminescence signal in liver. Luciferase mRNA was formulated with SS-OP using the lipid ratios for LNP1, as shown in Table 5. The lipid:mRNA ratios (wt/wt) were varied. The formulated mRNA LNPs were delivered via IV injection into C57B16/J mice at 0.6 mg of total mRNA/kg of body weight. As a negative control, a mouse was injected with saline. 24 hours later, these mice were shaved and imaged after injection with luciferin using a Lago instrument from Spectral Instruments Imaging. Depicted is an BLI image from mice dosed with a lipid:mRNA ratio of 175, 42, and 25. The signal was highest in the mice receiving LNPs with a lipid:mRNA ratio (wt/wt) of 175 and 42. The other data presented here using LNP1 uses a wt/wt ratio of 42.

[0459] FIG. 21 shows in vivo delivery of mRNA in a photograph depicting the results demonstrating that luciferase mRNA LNPs causing high bioluminescence signal in liver. LNPs designated as Lipid Nanoparticle 4 (LNP4) or Lipid Nanoparticle 5 (LNP5) were formulated using the recipe in Table 5 with luciferase mRNA. These LNPs were delivered via IV injection into C57B16/J mice at 0.6 mg/kg. As a negative control, a mouse was injected with saline. 20 hours later, these mice were shaved and imaged after injection with luciferin using the Lago instrument from Spectral Instruments Imaging. LNP4 consisted of the formula for LNP2, but with SS-OP substituted for cKK-E12. LNP5 consisted of the formula for LNP1, but with cKK-E12 substituted for SS-OP. Bioluminescent imaging indicates that both of these LNPs had successful delivery to the liver.

[0460] FIG. 22 shows in vivo delivery of mRNA in a photograph depicting the results demonstrating that luciferase mRNA LNPs causing high bioluminescence signal in liver. Luciferase mRNA was formulated with lipids per the recipe for LNP1 in Table 5. The ingredient that was varied was the molar ratio of DMG-PEG2000. As shown in FIG. 22, DMG-PEG2000 was added as either 1, 1.5, 2, or 3 parts relative to the molar sum of all lipids, while the molar ratio for the other 3 lipids is held constant. This corresponds to a molar percentage for DMG-PEG2000 of approximately 1.0%, 1.5%, 2.0%, and 2.9%, 20 hours after intravenous

delivery at 0.6 mg/kg, the C57B16/J mice were shaved and imaged following luciferin injection (75 mg/kg) using the Lago instrument from Spectral Instruments Imaging. The signal was strong from all of the mice receiving active Luciferase mRNA LNPs, and the best signal was seen when DMG-PEG2000 was added in a molar ratio of 1.5:101.5 of total (~1.5%). The other data presented here use LNP1 with this molar ratio of DMG-PEG2000.

[0461] FIG. 23 is a capillary electrophoresis gel image showing that TERT mRNA LNPs cause high telomerase activity in liver. Tert mRNA (mTert SEQ 37) was formulated with LNP3, a lipid nanoparticle containing DLin-MC3-DMA (Table 5) and delivered i.v into TERT KO mice at 0.6 mg/kg. 16 hours or 8 days later (as indicated in the image), the livers were harvested for telomerase repeat amplification protocol (TRAP). The negative control was a TRAP performed on a liver from a TERT KO mouse that was injected with saline. Livers from mice treated with TERT mRNA LNP3 exhibit elevated telomerase activity which returns to baseline levels, indicating the increase in telomerase activity was transient.

Example 2: Treatment of Fibrosis in a TAA Mouse Model with TERT mRNA

[0462] This Example demonstrates that an LNP formulation with SS-OP (LNP1 in Table 5), administered intravenously, effectively delivered an mRNA encoding TERT to the liver in an amount effective to treat liver fibrosis. Treatment was demonstrated by reduced liver scarring in both female and male animals (FIG. 9A, graph on left and FIG. 9B).

[0463] FIG. 8 is a graph and a series of photographs of a first study demonstrating that TERT LNPs reduce fibrosis in Thioacetamide (TAA) drinking water model. The addition of thioacetamide (TAA) to drinking water represents an art-recognized model for the induction of experimental liver fibrosis in rodents (Wallace et al. Standard operating procedures in experimental liver research: thioacetamide model in mice and rats. Lab Anim. 49:21-9 (2015)). In this experiment, TERT KO mice received 0.3 g/L TAA in their drinking water for 9.5 weeks. Mice were treated once weekly with LNP1 carrying 0.6 mg/kg of TERT mRNA (SEQ ID NO: 37) or Luciferase (LUC) mRNA. Liver sections were stained with Picosirius red (PSR), and a quantification showed a 24% mean reduction in PSR stained tissue in mice treated with TERT LNPs compared to those treated with LUC LNPs. Scale bar on photographs equals 500 μ m.

[0464] FIGS. 9A and 9B are graphs and photographs of a second study demonstrating that TERT LNPs Reduce Fibrosis in Thioacetamide (TAA) Drinking Water Model. TERT KO mice received 0.3 g/L TAA in their drinking water for 9.4 weeks and were treated with TERT or LUC LNPs once weekly. By picrosirius red (PSR) staining, there was an 18% mean reduction in fibrosis in female mice and a 37% mean reduction was observed in males treated with TERT LNPs, representing a significant ($p=0.041$) reduction in fibrosis. Scale bar on photographs equals 500 μ m. Additionally, using the 0 through 4 scoring system developed by the Pathology Committee of the NASH Clinical Research Network (Kleiner et al. Hepatology 2005), animals treated with TERT mRNA LNPs had a significant reduction in fibrosis compared to control animals treated with LUC (luciferase) mRNA LNPs ($p=0.032$) as seen in FIG. 9B. For all scoring, liver fibrosis was scored independently for each of 3 lobes

per mouse (right, median, and left) in a blinded manner. The scores were averaged together to get a score per mouse, which is then plotted in the graph (FIG. 8, FIG. 9A, and FIG. 9B).

Example 3: Improved Survival in TAA Mouse Model after Treatment with TERT mRNA

[0465] This Example demonstrates that an LNP formulation with SS-OP (LNP1 in Table 5), administered intravenously, effectively delivered an mRNA encoding TERT to the liver in an amount effective to treat liver fibrosis. Treatment was demonstrated by increased survival in the treatment group (TERT) compared to the control group (LUC) in FIG. 10. TERT mRNA treated mice showed a 42% increase in median survival and a 58% increase in maximal survival. Moreover, the maximal lifespan for the TERT mRNA treated mice on the TAA liver toxin was 12.5% longer (117 vs 104 days) than the control mice that did not receive TAA.

[0466] FIGS. 10A and 10B are graphs demonstrating that TERT mRNA improves survival. Survival plotted as fraction of mice alive as a function of days post first dose of either TERT (SEQ ID NO: 37) or a Luciferase (LUC) negative control. Same experimental procedure was followed as described in FIG. 9, but mice were 4th generation (G4) TERT KO's aged to over 30 weeks at the start of the study. These mice were dosed once weekly for 8 weeks with 0.6 mg/kg TERT or LUC mRNA, and survival was recorded after the first dose.

Example 4: Decreased Inflammation and Increase Telomere Length after TERT mRNA Treatment

[0467] This Example demonstrates that TERT mRNA treatment decreased inflammation and increased telomere length in the livers of treatment subject animals with liver fibrosis. Furthermore, it confirms that mRNA was delivered and translated at all dose levels tested (0.05 mg/kg to 0.6 mg/kg) in nearly all hepatocytes with an SS-OP-based LNP (LNP1). Lastly, it shows that both an SS-OP-based LNP and an MC3-based LNP are tolerated, with the SS-OP-based LNP having less toxicity (lower AST) than the MC3-based LNP. Further, the number of mice with pathological inflammation was significantly reduced in mice treated with TERT mRNA.

Transfection Efficiency with Reporter mRNA in Healthy and Fibrotic Subject Animals

[0468] The high in vivo transfection efficiency of reporter mRNA in the liver with LNP1 is shown in FIG. 12A. Different doses of Cre mRNA encapsulated in lipid nanoparticles with ionizable LNP1 delivered intravenously to *tdTomato* fl/fl mice were quantified. *tdTomato* flox/flox (fl/fl) mice refers to knock-in of the *tdTomato* gene in which portions of the gene are flanked by two Cre recombinase recognition sites. FIG. 12B shows representative images of immunohistochemistry (IHC) using an anti-*tdTomato* antibody in liver sections from the knock-in mice. Hepatocyte cells were identified from mouse liver tissue sections using nuclear size and circularity with QuPath software.

[0469] Low levels of liver damage markers were observed with successful TERT mRNA delivery. TERT mRNA was formulated with LNP1 or D-Lin-MC3-DMA (MC3) (LNP3) and delivered intravenously into C57B16 mice at 0.6 mg/kg. 24 hours later, the liver toxicity markers alanine aminotrans-

ferase (ALT) and aspartate aminotransferase (AST) were measured. LNP1 delivery of TERT mRNA had equivalent or lower levels of ALT and AST compared to MC3 delivery of TERT mRNA (FIG. 13).

[0470] In vivo transfection efficiency of reporter mRNA was also high in fibrotic liver. Hepatocytes were identified using nuclear size and circularity by QuPath software, as described above. FIG. 14 B shows representative IHC images using an anti-*tdTomato* antibody in liver sections. Telomere Extension and Reduction in Inflammation in Liver of Subject Animals Treated with Fibrosis-Inducing Liver Toxin

[0471] FIG. 11 are two graphs demonstrating that TERT LNPs reduce lobular inflammation in the livers of mice on the thioacetamide (TAA) drinking water model. The addition of thioacetamide (TAA) to drinking water represents an art-recognized model for the induction of experimental liver fibrosis in rodents. In this experiment, TERT KO mice received 0.3 g/L TAA in their drinking water for 9.5 weeks. Mice were treated once weekly with LNP1 carrying 0.6 mg/kg of TERT mRNA (SEQ ID NO: 37) or Luciferase (LUC) mRNA. TERT mRNA (SEQ ID NO: 37) in vivo delivery with the LNP1 formulation resulted in a 60% reduction in the number of animals with a score of >1 (FIG. 11B). Lobular inflammation was performed by a certified pathologist based on the non-alcoholic fatty liver disease NAFLD Activity Score (NAS) (Kleiner et al Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* Jun 41(6); 1313-21; 2005). The measurement was performed on hematoxylin and eosin (H&E) stained liver sections from the thioacetamide (TAA) water experiment described in Example 2. Saline-treated animals had a mild inflammation score of 1 (<2 foci per 200× field of view).

[0472] Three doses of either TERT mRNA (SEQ ID NO: 37) or LUC mRNA formulated with LNP1 were delivered to TERT knock out (KO) mice once weekly intravenously at 0.5 mg/kg. The mRNA-LNP1 dosing was preceded two days prior by a dose of thioacetamide (TAA) intraperitoneally (i.p.) at 50 mg/kg. Mice were harvested 1 week after the final dose of mRNA LNP1. Telomere length was quantified in hepatocytes using Q-FISH. Liver tissues were fixed, sectioned, and stained with a TelC fluorescent probe that labels the telomeres. Individual telomere fluorescence was quantified on a per cell basis (the median is shown in FIG. 15A and the 10th percentile is shown in FIG. 15B), and the average was taken for each mouse. Each point represents a single mouse. Hepatocytes in mice treated with TERT mRNA had significantly longer telomeres than luciferase mRNA treated control animals. At least 300 cells were analyzed per mouse per treatment group.

Telomere Extension in Human Hepatocytes

[0473] To measure telomerase activity in ex vivo human samples, the telomerase repeat amplification protocol (TRAP) assay was used on lysates from human hepatocytes incubated with an artificial telomere (DNA oligonucleotide). As described above in Example 1, when active telomerase is present, it extends the artificial telomere 6 base pairs bp at a time, producing a ladder pattern. This extension reaction is amplified by PCR and run on a gel (in this case Agilent bioanalyzer, a microfluidic agarose gel). The presence of a ladder in 6 bp increments indicates telomerase activity.

[0474] Human hepatocytes from a 51-year-old donor were cultured and transfected with GFP mRNA or TERT mRNA (SEQ ID NO: 39) using Messenger Max™ from Thermo Scientific at 1 µg/ml. Cells were harvested at each time point indicated in FIG. 16A for the TRAP assay to measure telomerase activity. An Agilent Bioanalyzer was used to detect the characteristic telomerase activity a ladder pattern as shown in FIG. 16A. Telomerase activity was detected strongly on day 1 and day 2 post-transfection, and weakly on day 7 post-transfection. It was not detected on day 14 post-transfection or in hepatocytes treated with GFP mRNA.

[0475] Telomere length was quantified using a fluorescent probe to label the telomeres. Individual telomere fluorescence was quantified on a per cell level as the mean for FIG. 17A and the 10th percentile for FIG. 17B. At least 150 cells were analyzed per treatment group. It was observed that telomerase activity returned to baseline by day 14.

[0476] FIG. 19 shows results of the telomerase activity assay “telomerase repeat amplification protocol” (TRAP) in human fibroblasts treated for 24 hours with 1 µg/ml TERT mRNAs of from left to right, untreated cells, SEQ ID NOS: 39, 40, 1, 2, 31, 3, 5, and 4 respectively, and a GFP mRNA control. Telomerase activity is indicated by a characteristic ladder pattern as shown by the transfection of TERT mRNAs of SEQ ID NOS: 39, 40, 1, 2, 31, 3, 5, and 4 to varying degrees. Untreated and GFP mRNA samples did not exhibit telomerase activity.

Imaging of LNP1-TERT mRNA Formulation

[0477] The LNP1-TERT mRNA (SEQ ID NO: 40) formulation was imaged at high resolution using the Thermo

Scientific Talos Glacios Cryo transmission electron microscope (TEM) at 34,000× magnification and 200 kv voltage. A representative image is show in FIG. 18A; the TEM copper grid is the dark region on the right. The particle size was characterized using dynamic light scattering (DLS) using a Brookhaven 90Plus Particle Analyzer (FIG. 18B).

[0478] LNP1 nanoparticles comprising TERT mRNA were observed to have the following exemplary characteristics, shown in Table 6.

TABLE 6

Characteristic	Broad Range	Narrow Range
Particle Size	50-150 nm	70-100 nm
Zeta Potential	5-30 mV	18-20 mV
Encapsulation	70-100%	85-98%
Polydispersity index (PDI)	<0.2	<0.1
Ratio of total lipid to mRNA	30-300 nmol/µg	40-120 nmol/µg

[0479] While embodiments of the instant disclosure have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the disclosure. It should be understood that various alternatives to the embodiments of the disclosure described herein may be employed in practicing the disclosure. It is intended that the following claims define the scope of the disclosure and that methods and structures within the scope of these claims and their equivalents be covered thereby.

SEQUENCE LISTING

```

<160> NUMBER OF SEQ ID NOS: 40

<210> SEQ ID NO 1
<211> LENGTH: 3399
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 1

atgcccagatg ccccccaggatg cagagccgtg agaaggctgc tgagaagccca ctacagagag 60
gtgctgcccc tggccacatt cgtgagaaga ctggggccccc agggctggag actggtgca 120
agaggcgacc ccgcgcgcctt cagagccctg gtggcccaagt gcctgggtgtg cgtgcctgg 180
gacgcccagac cccatccccgc cgcccccaagg ttcagacagg tgagctgcct gaaggagctg 240
gtggccagag tgctgcagag actgtgcgag agaggcgcca agaacctgtgc ggccttcggc 300
ttcgccctgc tggacggcgcc cagaggcgcc cctcccgagg ctttcaccac cagcgtgaga 360
agctacactgc ccaacacccgt gaccgacgcc ctgagaggca gggggccctg gggccctgtg 420
ctgagaagag tggggcgacga cgtgctggtg cacctgtgtt ccagatgcgc cctgttgtg 480
ctgggtggccc ccagatgcgc ctaccaggatg tgccggccctc ccctgttacca gctggggcc 540
gccacccagg ccagacccccc tccccacgcc agcgcccccga gaagaagact gggctgcgag 600
agagcctggaa accacagcgat gagagaggcc ggcgtgcgcc tggggctgcc cgcggccggc 660
gccagaagaa gaggccggcag cgccagcaga agcctgcctc tgcccaagag acccagaaga 720
ggcgccggccc ccgagcccgaa gagaacccccc gtggggccagg gcagctggc ccacccggc 780

```

-continued

agaaccagag	gccccagcga	cagaggcttc	tgegtggta	gccccccag	acccgcccag	840
gaggccacca	gcctggaggg	cgcctgago	ggcaccagac	acagccaccc	cagcgtggc	900
agacagcacc	acgcgggccc	tcccagcaco	agcagacaccc	ccagaccctg	ggacaccccc	960
tgcctcccg	tgtacgcca	gaccaagcac	ttctctgtaca	gcagcggcga	caaggagcag	1020
ctgagaccca	gttctctgtc	gagcagcctg	agacccagcc	tgaccggcgc	cagaagactg	1080
gtggagacca	tcttcctggg	cagcagaccc	tggatgccc	gcaccccccag	aagactgcc	1140
agactgcccc	agagatactg	gcagatgaga	ccctgttcc	tggagctgct	ggcaaccac	1200
gcccagtgc	cctacggcgt	gctgtgaag	acccactgc	ccctgagac	cgccgtgacc	1260
cccgccgocg	gcgtgtgcgc	cagagagaag	ccccaggcga	gcgtggccgc	ccccgaggag	1320
gaggacacccg	accccagaag	actggtgtcag	ctgctgagac	agcacagcag	ccctggcag	1380
gtgtacggct	tcgtgagagc	ctgcctgaga	agactggtc	ctccggcct	gtggggcagc	1440
agacacaacg	agagaagatt	cctgagaaac	accaagaagt	tcatcagcct	ggcaagcac	1500
gccaagctga	gcctgcagga	gctgacctgg	aagatgagcg	tgagagactg	cgccctggct	1560
agaagaagcc	ccggcgtggg	ctgcgtgccc	gccgcccagc	acagactgag	agaggagatc	1620
ctggccaagt	tcctgcactg	gctgtgago	gtgtacgtgg	tggagctgct	gagaagcttc	1680
ttctacgtga	ccgagaccac	cttccagaag	aacagactgt	tcttctacag	aaagagcgtg	1740
tggagcaagc	tgcagagcat	cgccatcaga	cagcacctga	agagagtgc	gctgagagag	1800
ctgagcgagg	ccgaggtgag	acagcacaga	gaggccagac	ccgcctgct	gaccagcaga	1860
ctgagattca	tccccaagcc	cgacggcctg	agacccatcg	tgaacatgga	ctacgtgg	1920
gggcccaagaa	ctttcagaag	agagaagaga	gccgagagac	tgaccagcag	agtgaaggcc	1980
ctgttcagcg	tgctgaacta	cgagagagcc	agaagacccg	gcctgctggg	cgccagcgtg	2040
ctgggcctgg	acgacatcca	cagacccctgg	agaacccctcg	tgctgagagt	gagagccag	2100
gatccccctc	ccgagctgta	cttcgtgaag	gtggacgtga	ccggcgccta	cgacaccatc	2160
ccccaggaca	gactgaccga	ggtgatgcgc	agcatcatca	agccccagaa	cacctactgc	2220
gtgagaagat	acgcggcgtgt	gcagaaggcc	gcccacggcc	acgtgagaaa	ggccctcaag	2280
agccacgtga	gcaccctgac	cgacccctgc	ccctacatga	gacagttcg	ggcccacctg	2340
caggagacca	gccccctgag	agacccctgc	gtgatcgac	agagcagcag	cctgaacccg	2400
gccagcagcg	ccctgttgc	cgtgttcc	agattcatgt	gccaccacgc	cgtgagaatc	2460
agaggcaaga	gctacgtgca	gtgccaggc	atccccagg	gcagcatcct	gaccaccc	2520
ctgtgcagcc	tgtgctacgg	cgacatggag	aacaagctgt	tgcggccat	cagaagagac	2580
ggccctgtgc	tgagactggt	ggacacttc	ctgctggta	cacccaccc	gaccacgccc	2640
aagacccctcc	tgagaaccct	ggtgagaggg	gtgcccaggt	acggctgcgt	ggtgaacctg	2700
agaaagacccg	tggtaactt	ccccgtggag	gacgaggccc	tggggggcac	cgccctcg	2760
cagatgcccgg	cccacggccct	gttccccctgg	tgcggccctgc	tgctggacac	cagaaccc	2820
gaggtgcaga	gcgactacag	cagctacgc	agaaccagca	tcaagccag	cctgacc	2880
aacacaggct	tcaaggccgg	cagaacacatg	agaagaaagc	tgttccggcgt	gctgagactg	2940
aagtgccaca	gcctgttcc	ggacccctgc	gtgaacagcc	tgcagaccgt	gtgcaccaac	3000
atctacaaga	tcctgtgtc	gcaggccctac	agattccacg	cctgcgtgct	gcagctgccc	3060

-continued

ttccaccaggc	aggtgtggaa	gaaccccacc	ttcttcttgc	gagtgtatcg	cgacaccggc	3120
agcctgtgt	acagcatcct	gaaggccaag	aacgcggca	tgagcctgg	cgccaagggc	3180
gccccggcc	ccctgeccag	cgaggccgt	cagtggctgt	gccaccaggc	cttctgtgt	3240
aagctgacca	gacacagagt	gacctacgt	ccccctgtgg	gcagcctgag	aaccgcccag	3300
acccagctga	gcagaaaagct	gcceggcacc	accctgaccg	ccctggaggc	cgccgccaac	3360
cccgccctgc	ccagcgactt	caagaccatc	ctggactga			3399

<210> SEQ ID NO 2
<211> LENGTH: 3399
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 2						
atgcgcgcgc	ctcccccgt	ccgagccgt	cgtccctgc	tgcgcagcca	ctaccgcgag	60
gtgctgcgc	tggccacgtt	cgtgcggcgc	ctggggccccc	agggctggcg	gctgggtcag	120
cgcggggacc	cggcggttt	ccgcgcgt	gtggcccagt	gcctgggtgt	cgtgcctgg	180
gacgcacggc	cgccccccgc	cgccccctcc	ttccgcccagg	tgtcctgcct	gaaggagctg	240
gtggcccgag	tgctgcagag	gctgtgcgag	cgcggcgca	agaacgtgt	ggccttcggc	300
ttcgcgcgtc	tggacggggc	ccgcgggggc	ccccccgagg	ccttcaccac	cagcgtgcgc	360
agctacctgc	ccaacacggt	gaccgacgca	ctgcggggga	gcggggcggt	ggggctgtcg	420
ctgcgcgcgc	tggggcgtca	cgtgtggtt	cacctgtgt	cacgcgtgcgc	gtctttgtg	480
ctgggtggctc	ccagctgcgc	ctaccagggt	tgcggccgc	cgtgtacca	gtcggcgct	540
gccactcagg	ccccggccccc	gccacacgt	agtggacccc	gaaggcgct	gggatgcgaa	600
cgggccttga	accatagcgt	cagggaggcc	ggggtcccc	tgggcctgcc	agccccgggt	660
gcgaggaggc	cgggggggcag	tgccagccga	agtctgcgt	tgcggcaagag	gccaggcggt	720
ggcgctgcgc	ctgagccgga	gcggacgccc	gttggcagg	gttctgggc	ccacccgggc	780
aggacgcgtg	gaccgagtga	ccgtggtttc	tgtgtgggt	cacgtccag	acccggccaa	840
gaagccacct	ctttggaggg	tgcgtctct	ggcacgcgc	actcccaccc	atccgtggc	900
cgcacgcacc	acgcggggcc	cccatccaca	tgcggccac	cacgtccctg	ggacacgcct	960
tgtccccccgg	tgtacgcga	gaccaagcac	ttctctact	cctcaggcga	caaggagcag	1020
ctgcggccct	ccttctact	cagctctgt	aggcccagcc	tgactggcgc	tcggaggctc	1080
gtggagacca	tctttctggg	ttccaggccc	tggatgccag	ggactcccc	caggttgcc	1140
cgcctgcggcc	agcgctactg	gcaaattgcgg	ccccctgtttc	tggagctgt	tggaaaccac	1200
gcgcagtgcc	cctacgggg	gctctcaag	acgcactgca	cgctgcgagc	tgcggtcacc	1260
ccagcagccg	gtgtctgtgc	ccgggagaag	ccccagggt	ctgtggcgc	ccccaggag	1320
gaggacacag	accccccgtcg	cctgggtgcag	ctgtccgc	agcacagcag	ccccctggcag	1380
gtgtacggct	tcgtgcgggc	ctgcctgcgc	cggctgggtc	ccccaggct	ctggggctcc	1440
aggcacaacg	aacgcggctt	cctcaggaaac	accaagaagt	tcatctccct	gggaaagcat	1500
gcgaagctct	cgctgcagga	gctgacgtgg	aaagatgagcg	tgcgggactg	cgcttggct	1560
cgcaggagcc	caggggttgg	ctgtgttccg	gccgcagagc	accgtctgcg	tgaggagatc	1620

-continued

ctggccaagt	tcctgcactg	gctgatgagt	gtgtacgtcg	tcgagctgtc	caggctttc	1680
tttatgtca	cgagaccac	gttcaaaggaa	aacaggctc	tttctaccc	gaagagtgtc	1740
tggagcaagt	tgcaaaggcat	tggaatcaga	cagcactga	agagggtgca	gctgcggag	1800
ctgtcgaaag	cagaggctcg	gcagcatcg	gaagccaggc	ccgcctgtc	gacgtccaga	1860
ctccgcttca	tcccaagcc	tgacgggctg	cggccgattt	tgaacatgga	ctacgtcg	1920
ggagccagaa	cgtccgcag	agaaaagagg	gccgagcg	tcacctcgag	ggtgaaggca	1980
ctgttcagcg	tgctcaacta	cgagcggcgc	cggcccccg	gcctctggg	cgccctgt	2040
ctgggcctgg	acgatatacc	cagggcctgg	cgcacccctcg	tgctgegtgt	gcgggcccag	2100
gaccgcgcgc	ctgagctgtt	ctttgtcaag	gtggatgtga	cgggcgcgt	cgacaccatc	2160
ccccaggaca	ggctcacgg	ggtcatcgcc	agcatcatca	aacccagaa	cacgtactgc	2220
gtgcgttgtt	atgcgttgtt	ccagaaggcc	gcccatggc	acgtccgca	ggcctcaag	2280
agccacgtct	ctaccttgac	agacctccag	ccgtacatgc	gacagttcg	ggctcacctg	2340
caggagacca	gcccgtgag	ggatgcgcgt	gtcatcgac	agagctcc	cctgaatgag	2400
gccagcagt	gcctcttgc	cgtcttcata	cgcttcatgt	gccaccacgc	cgtgcgcac	2460
aggggcaagt	cctacgttca	gtgccagggg	atcccgagg	gctccatc	ctccacgt	2520
ctctgcagcc	tgtgtacgg	cgacatggag	aacaagctgt	ttgggggat	tcgggggac	2580
gggctgtcc	tgcgttttgtt	ggatgattt	ttgttgtga	cacccatc	cacccacgc	2640
aaaaacttcc	tcaggaccct	ggtccgaggt	gtccctgagt	atggctgcgt	ggtaactt	2700
cggaagacag	ttgtgaactt	ccctgttagaa	gacgaggccc	ttgggtggc	ggcttttgtt	2760
cagatgcgg	cccacggcct	atccccctgg	tgcggcctgc	tgtggatac	ccggaccctg	2820
gagggtcaga	gcgactactc	cagctatgcc	cggacccctca	tcaagccag	tctcacctc	2880
aaccgcggct	tcaaggctgg	gaggaacatg	cgtcgaaac	tctttgggt	cttgcggct	2940
aagtgtcaca	gcctgtttct	ggatttgcag	gtgaacagcc	tccagacgg	tgccaccaac	3000
atctacaaga	tcctctgt	gcaggcgtac	aggtttcag	catgtgtgt	gcagtc	3060
tttcatcagc	aatgttgaa	gaacccaca	ttttctgc	gcgtcatc	tgacacggcc	3120
tccctctgt	actccatcct	gaaagccaa	aacgcagg	tgtcgctgg	ggccaagg	3180
gccgcggccc	ctctgcctc	cgaggccgt	cagtggctgt	gccacca	attctgtc	3240
aagctgactc	gacaccgtgt	cacctacgt	ccactctgg	ggtcactc	gacagcc	3300
acgcagtg	gtcggaaagct	cccggggac	acgctgact	ccctggagg	cgcagccaa	3360
ccggcactgc	cctcagactt	caagaccatc	ctggactga			3399

```

<210> SEQ ID NO 3
<211> LENGTH: 3369
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

<400> SEQUENCE: 3

atgaccagg	cccctagg	ccctgcgt	aggaggctgc	tgaggagg	gtacagg	60
gtgtggc	tggccac	ctgtgagg	ctggccct	agggcagg	gctgg	120
cctggc	ctaagatct	caggacc	gtggcc	actgtgt	catgcact	180

-continued

ggcagccaggc	ctccctcctgc	cgcacctgago	ttccaccagg	tgagcagcct	gaaggagctg	240
gtggccaggg	tgggtcagag	gctgtcgag	aggaacgaga	ggaacgtgct	ggccttcggc	300
ttcgagctgc	tgaacgaggc	caggggcggo	cctcctatgg	ccttcaccag	cagcgtgagg	360
agctacacgtc	ctaaccaccgt	gategagacc	ctgaggggtga	gcccgcgcctg	gatgtgtcg	420
ctgagcaggg	tggggegacga	cctgctggtg	tacctgctgg	cccactgcgc	cctgtacactg	480
ctgggtcctc	ctagctgcgc	ctaccagggt	tgccggcagcc	ctctgtacca	gatctgcgcc	540
accaccgaca	tctggcctag	cgtgagcgc	agctacaggc	ctaccaggcc	tgtgggcagg	600
aacttcacca	acctgagggtt	cctgcagcag	atcaagagca	gcagcaggca	ggaggccct	660
aaggcctctgg	ccctgcctag	caggggcacc	aagaggcacc	tgagcctgac	cagcaccagc	720
gtgcctagcg	ccaagaaggc	caggtgctac	cctgtgccta	gggtggagga	gggcctcacc	780
aggcagggtgc	tcgcctacccc	tagccggcaag	agctgggtgc	ctagccctgc	caggagccct	840
gagggtgccta	ccggccgagaa	ggacactgago	agcaaggcc	aggtgagcga	cctgagcctg	900
agcggcagcg	tgtgctgcaa	gcacaaggct	agcagcacca	gcctgctgag	ccctccttagg	960
cagaacgcct	tccagctgag	gcctttcato	gagaccaggc	acttcctgta	cagcaggggc	1020
gacggccagg	agaggctgaa	ccctagcttc	ctgctgagca	acctgcagcc	taacctgacc	1080
ggcgccaggag	ggctgggtgg	gatcatcttc	ctgggcagca	ggcctaggac	cagcggccct	1140
ctgtgcagga	cccacaggct	gagcaggagg	tactggcaga	tgaggcctct	gttccagcag	1200
ctgctgggtga	accacgcccga	gtgccagtag	gtgaggctgc	tgaggagcc	ctgcaggttc	1260
aggaccgc	accagcagg	gaccgacgco	ctgaacacca	gcctcctca	cctgtatggac	1320
ctgctgaggc	tgcacagcag	cccttggcag	gtgtacggct	tcctgagggc	ctgcctgtgc	1380
aagggtgggtga	gcgcgcgcct	gtggggcacc	aggcacaac	agaggagg	tttcaagaac	1440
ctgaagaagt	tcatcagcct	ggcaagctac	ggcaagctga	gcctgcagga	gctgtatgtgg	1500
aagatgaagg	tggaggactg	ccactggctg	aggagcagcc	ctggcaagga	cagggtgcct	1560
ggccgcgc	acaggctgag	ggagaggatc	ctggccacct	tcctgttctg	gctgtatggac	1620
acctacgtgg	tgcagctgct	gaggagcttc	ttctacatca	ccgagacac	cttccagaag	1680
aacaggctgt	tcttctacag	gaagagcgtg	tggagcaagc	tgcagagcat	cggcgtgagg	1740
cagcacccctgg	agagggtgag	gctgagggag	ctgagccagg	aggaggtgag	gcaccaccag	1800
gacacccctggc	tggccatgcc	tatctgcagg	ctgagggttca	tccttaagcc	taacggcctg	1860
aggcctatcg	tgaacatgag	ctacagcatg	ggcaccagg	ccctggcag	gaggaagcag	1920
gcccagca	tcacccagag	gctgaagacc	ctgttcagca	tgctgaacta	cgagaggacc	1980
aagcaccctc	acctgtatgg	cagcagcgtg	ctgggcatt	acgacatcta	caggacccctgg	2040
aggcccttcg	tgctgagggt	gaggccctg	gaccagacc	ctaggatgta	cttcgtgaag	2100
ggccacgtga	ccggccgcata	cgacgcctac	cctcagggg	agctgggtgg	ggtgtggcc	2160
aacatgtatca	ggcacagcga	gagcacctac	tgcacatggc	agtacgcctg	ggtgaggagg	2220
gacagccagg	gccaggtgca	caagagcttc	aggaggcagg	tgaccaccct	gagcgcac	2280
cagccttaca	tggggcagtt	cctgaagcac	ctgcaggaca	gagcgcocag	cgccctgagg	2340
aacagcgtgg	tgatcgagca	gagcatcago	atgaacgaga	gcagcagcag	cctgttcgac	2400
tttttctgc	acttcctgag	gcacagcgtg	gtgaagatcg	gagcagagg	ctacaccag	2460

-continued

tgcacaggca	tccctcaggg	cagcagcctg	agcacccctgc	tgtgcagcct	gtgcttcggc	2520
gacatggaga	acaagctgtt	cgcgcagggtg	cagagggacg	gcctgtgtc	gaggttcggt	2580
gacgacttcc	tgctgggtac	ccctcacctg	gaccaggcca	agacacctct	gagcacctg	2640
gtgcacggcg	tgcctgagta	cggtgcatt	atcaacctgc	agaagaccgt	ggtaacttc	2700
cctgtggagc	ctggcaccct	ggggggggcc	gccccttacc	agctgcctgc	ccactgcctg	2760
ttcccttgg	gccccctgct	gctggacacc	cagacccctgg	aggtgttctg	cgactacagc	2820
ggctacgccc	agaccagcat	caagaccago	ctgacacctcc	agagcgtgtt	caaggccggc	2880
aagaccatga	ggaacaagct	gctgagcgtg	ctgaggctga	agtgcacagg	cctgttctg	2940
gacctgcagg	tgaacagcct	gcagaccgtg	tgcataaca	tctacaagat	cttctctgt	3000
caggcctaca	ggttccacgc	ctgcgtgatc	cagtcgcctt	tgcaccagag	ggtgaggaag	3060
aacctgaccc	ttttcttggg	catcatcago	agccaggcca	gctgctgcata	cgccatctg	3120
aagggtgaaga	accctggcat	gaccctgaag	gccagggca	gcttccctcc	tgaggccggc	3180
caactggctgt	gttaccaggc	cttctgtctg	aagctggccg	cccacagcgt	gatctacaag	3240
tgccctgctgg	gcccctgtag	gaccggccag	aagctgctgt	gcaggaagct	gcctgaggcc	3300
accatgacca	tcctgaaggc	cgccgcccac	cctgcctgta	gcacccgactt	ccagaccatc	3360
ctggactga						3369

<210> SEQ ID NO 4

<211> LENGTH: 3369

<212> TYPE: DNA

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 4

atgaccccg	ctccctcggt	ccccgggtg	cgctctctgc	tgcgacccg	ataccggag	60
gtgtggccgc	tggcaacctt	tgtggggcgc	ctggggcccg	agggcaggcg	gcttgtgcaa	120
ccgggggacc	cgaagatcta	ccgcactttg	gttgccaaat	gcctagtgtg	catgcactgg	180
ggctcacagc	ctccacactgc	cgacccctcc	ttccaccagg	tgtcataccct	gaaagagctg	240
gtggccaggg	ttgtgcagag	actctgcgag	cgcaacgaga	gaaacgtgt	ggctttggc	300
tttgagctgc	ttaacgaggc	cagaggccgg	cctccatgg	ccttcactag	tagcgtgcgt	360
agctacttgc	ccaacactgt	tattgagacc	ctgcgtgtca	gtggtgcatg	gatgtactg	420
ttgagccgag	tggggcagca	cctgtgtgtc	tacctgtctg	cacactgtgc	tctttatctt	480
ctggtgcccc	ccagctgtgc	ctaccagggt	tgtgggtctc	ccctgtacca	aatttgc	540
accacggata	tctggccctc	tgtgtccgt	agttacaggc	ccacccgacc	cgtggccagg	600
aatttcacta	accttaggtt	cttacaacag	atcaagagca	gtagtcgc	ggaagcaccc	660
aaacccctgg	ccttgcacatc	tcgaggatca	aaggaggatc	ttagtctcac	cagtacaagt	720
gtgccttcag	ctaagaaggc	cagatgctat	cctgtcccga	gagtgaggg	gggacccac	780
aggcaggtgc	taccaacccc	atcaggcaaa	tcatgggtgc	caagtccctgc	tcggcccc	840
gagggtgcata	ctgcagagaa	agatttgtct	tctaaaggaa	aggtgtctga	cctgagtc	900
tctgggtcgg	tgtgtgtaa	acacaagccc	agctccacat	ctctgtgtc	accacccgc	960
caaaaatgcct	ttcagctcag	gccatttatt	gagaccagac	atttccttta	ctccagggg	1020

-continued

gatggccaag	agcgtctaaa	cccctcatc	ctactcagca	acctccaggc	taacttgact	1080
ggggccagga	gactgggtgga	gatcatctt	ctgggctcaa	ggcctaggac	atcaggagcca	1140
ctctgcagga	cacaccgtct	atcgctcgaa	tactggcaga	tgccggccct	gttccaacag	1200
ctgctgggtga	accatgcaga	gtgcacaatat	gtcagactcc	ttaggtcaca	ttgcagggtt	1260
cgaacagcaa	accaacaggt	gacagatgcc	ttgaacacca	gccccccgca	cctcatggat	1320
ttgctccggc	tgcacagcag	tccctggcag	gtataatgggt	ttcttcgggc	ctgtctctgc	1380
aagggtggtgt	ctgcttagtct	ctgggggtacc	aggcacaatg	agcgcggctt	ctttaagaac	1440
ttaaagaaga	tcatctcggt	ggggaaatac	ggcaagctat	cactgcagga	actgtatgg	1500
aagatgaaag	tagaggattg	ccactggctc	cgcagcagcc	cggggaaagga	ccgtgtcccc	1560
gctgcagagc	accgtctgag	ggagaggatc	ctggctacgt	tcctgttctg	gctgtatggac	1620
acatacgtgg	tacagctgct	taggtcatc	ttttacatca	cagagagcac	attccagaag	1680
aacaggtct	ttttctaccg	taagagtgtg	tggagcaagc	tgcagagcat	tggagtcaagg	1740
caacacccctg	agagagtgcg	gctacggggag	ctgtcacaag	aggaggtcag	gcatcaccag	1800
gacacctggc	tagccatgcc	catctgcaga	ctgcgcctca	tccccaaagcc	caacggcctg	1860
cggcccatcg	tgaacatgag	ttatagcatg	ggtaccagag	ctttgggcag	aaggaagcag	1920
gcccagcatt	tcacccagcg	tctcaagact	ctcttcagca	tgctcaacta	tgagcggaca	1980
aaacatccctc	accttatggg	gtcttctgt	ctgggtatga	atgacatcta	caggacctgg	2040
cgggccttgc	tgctgcgtgt	gcgtgcctcg	gaccagacac	ccaggatgta	ctttgttaag	2100
gcagatgtga	cggggccctt	ttagtgcata	ccccagggtt	agctgggtga	gggttgtgcc	2160
aatatgtatca	ggcactcgga	gagcacgtac	tgtatccgcc	agtatgcagt	ggtccggaga	2220
gatagccaa	gccaagttca	caagtccctt	aggagacagg	tcaccaccct	ctctgacctc	2280
cagccataca	tggggccagtt	ccttaagcat	ctgcaggatt	cagatgccag	tgcactgagg	2340
aactccgttg	tcatcgagca	gagcatctct	atgaatgaga	gcagcagcag	cctgtttgac	2400
ttcttcctgc	acttcctgcg	tcacagtgtc	gtaaagattt	gtgacaggtt	ctatacgcag	2460
tgccaggcgc	tccccccagg	ctccagccta	tccaccctgc	tctgcagttt	gtgtttcgga	2520
gacatggaga	acaagctgtt	tgctgagggt	cagcggatgt	gggtgttttt	acgttttgtt	2580
gatgactttc	tgttggtgac	gcctcaactt	gaccaagcaa	aaacccctt	cagcaccctg	2640
gtccatggcg	ttccctgagta	tgggtgcatt	ataaacttgc	agaagacagt	ggtgaacttc	2700
cctgtggcgc	ctgggtaccct	ggggtgtca	gctccatacc	agctgcctgc	tcactgcctg	2760
tttccctgg	gtggcttgct	gctggacact	cagactttgg	aggtgttctg	tgactactca	2820
ggttatgcct	agacccatca	taagacgago	ctcacccctt	agagtgttctt	caaagctggg	2880
aagaccatgc	ggaacaagct	cctgtcggtc	ttgcgggtga	agtgcacgg	tctatttcta	2940
gacttgagg	tgaacacgct	ccagacagtc	tgcataata	tataacaagat	cttccctgtt	3000
caggcctaca	gggtccatgc	atgtgtgatt	cagctccctt	ttgaccagcg	tgtttaggaag	3060
aacctcacat	tctttctggg	catcatctcc	agccaagcat	cctgctgcta	tgctatcctg	3120
aagggtcaaga	atccaggaat	gacactaaag	gcctctggct	ctttctctcc	tgaagccgca	3180
cattggctct	gtcaccaggc	ctccctgctc	aagctggctg	ctcattctgt	catctacaaa	3240
tgtctctgg	gacccctgag	gacagcccaa	aaactgctgt	gccgaaagct	cccagagcgc	3300

-continued

acaatgacca tccttaaagc tgcagctgac ccagccctaa gcacagactt tcagaccatt	3360
ttggactaa	3369
<210> SEQ ID NO 5	
<211> LENGTH: 3369	
<212> TYPE: DNA	
<213> ORGANISM: Artificial	
<220> FEATURE:	
<223> OTHER INFORMATION: Synthetic	
<400> SEQUENCE: 5	
atgaccgcg ctccctcggtg cccgcggtg cgctctctgc tgcgcagccg ataccggag	60
gtgtggccgc tggcaacctt tgtgcggcgc ctggggcccg agggcaggcg gcttgtcaa	120
cccgccccacc cgaagatcta ccgcactttg gttgccaat gccttagtgc catgcactgg	180
ggctcacagc ctccacactgc cgaccttcc ttccaccagg tgcacccct gaaagagctg	240
gtggccaggg ttgtgcagag actctgcgag cgcaacgaga gaaacgtgct ggctttggc	300
tttgagctgc ttaacgaggc cagaggcggg cttccatgg ctttcactag tagcgtgcgt	360
agctacttgc ccaacactgt tattgagacc ctgcgtgtca gtgggtgcata gatgctactg	420
ttgagccgag tggggcagca cctgctggc tacctgctgg cacactgtgc tctttatctt	480
ctgggtgcccc ccagctgtgc ctaccagggt tggtggctc ccctgtacca aatttgc	540
accacggata tctggccctc tgtgtccgct agttacaggc ccacccgacc cgtggccagg	600
aatttcaacta accttaggtt cttacaacag atcaagagca gtagtcgcca ggaagcacccg	660
aaacccctgg cttggccatc tcgaggtaca aagaggcatc tgagtctcac cagtacaagt	720
gtgccttcag ctaagaaggc cagatgctat cctgtcccgag gatggggaggc gggaccccac	780
aggcagggtgc taccacaaaa atcaggcacaat tcatgggtgc caagtcctgc tcggcccc	840
gaggtgccta ctgcagagaa agatttgtct tctaaaggaa aggtgtctga cctgagtctc	900
tctgggtcg tgcgtgtaa acacaagccc agtccacat ctctgtgtc accacccgc	960
caaatgcct ttcagctcag gccattttt gagaccagac atttcctta ctccaggggaa	1020
gatggccaaag agcgtctaaa cccctcatc ctactcaga acctccagcc taacttgact	1080
ggggccagga gactgggtgaa gatcatctt ctgggtcaa ggcctaggac atcaggacca	1140
ctctgcaggaa cacaccgtct atcgcgtcga tactggcaga tgccggccct gttccaacag	1200
ctgctgggtga accatgcaga gtgcataat gtcagactcc tcaaggcacaat ttgcaggtt	1260
cgaacacagaa accaacacagg gacagatgcc ttgaacacca gcccacccca cctcatggat	1320
ttgtccggcc tgcacagcag tccctggcag gtatatggtt ttcttcgggc ctgtctctgc	1380
aagggtgggtgt ctgcttagtct ctggggtaacc aggcacaatg agcgcggctt cttaagaac	1440
ttaaagaagt tcatctcggtt gggaaatac ggcaagctat cactgcaggaa actgtatgtgg	1500
aaatgtaaag tagaggattt ccactggcctc cgcagcagcc cggggaaaggc ccgtgtcccc	1560
gctgcagagc accgtctgag ggagaggatc ctggctacgt tccctgttctg gctgtatggac	1620
acatacgtgg tacagctgct taggtcattt ttttacatca cagagagcac attccagaag	1680
aacaggctct ttttctaccg taagatgtgt tggagcaagc tgcagagcat tggagtcagg	1740
caacacccctg agagagtgcg gctacgggag ctgtcacaag aggaggctag gcatcaccag	1800
gacacccctggc tagccatgcc catctgcaga ctgcgttca tcccaagcc caacggctg	1860

- continued

cggccccattt	tgaacatgag	ttatagcatg	ggtaccagag	ctttgggcag	aaggaaagcag	1920
geccagccattt	tcaccccgacg	tctcaagact	ctttcagca	tgctcaacta	tgaggggaca	1980
aaacatcctt	accttatggg	gtcttctgtt	ctgggtatga	atgacatctt	caggacctgg	2040
cgggcctttt	tgctgctgtt	gcgtgctctg	gaccagacac	ccaggatgtt	ctttgttaag	2100
gcagatgttt	ccggggccctt	tgtgccatc	ccccagggtt	agctggtgga	ggtttgttgc	2160
aatatgtatc	ggcactcgat	gagcacgtac	tgtatccgc	agtatgcagt	ggtccggaga	2220
gatagccaagg	gcctaaatccca	caagtcctt	aggagacagg	tcaccaccc	ctctgaccc	2280
cagccatacata	tggggccagg	ccttaagcat	ctgcaggatt	cagatgccag	tgcaactgagg	2340
aactccgttg	tcatcgagca	gagcatctt	atgaatgaga	gcagcagcag	cctgtttgac	2400
ttcttcctgc	acttcctgcg	tcacagtgtt	gtaaagattt	gtgacagggt	ctatacgac	2460
tgccaggggca	tccccccagg	ctccagccta	tccaccctgc	tctgcagtct	gtgtttcgga	2520
gacatggaga	acaagctgtt	tgctgagggt	cagcgggatg	ggttgttttt	acgttttgtt	2580
gatgactttc	tgttggtgac	gcctcactt	gaccaagcaa	aaacccctt	cagcaccct	2640
gtccatggcg	ttccctgagta	tgggtgcatt	ataaaacttgc	agaagacagt	ggtaacttc	2700
cctgtggagc	ctggtaacct	gggtggtgca	gctccatacc	agtcgcctgc	tcactgcct	2760
ttttccctgg	gtggcttgct	gctggacact	cagactttgg	aggtgttctg	tgactactca	2820
ggttatgccc	agacctaata	taagacgagc	ctcacccccc	agagtgtctt	caaagctggg	2880
aagccatgc	ggaacaagct	cctgtcggtc	ttgcgggttga	agtgtcacgg	tctatttctt	2940
gacttgcagg	tgaacagct	ccagacagtc	tgcataata	tatacaagat	cttcctgttt	3000
caggccataca	ggttccatgc	atgtgtgatt	cagttccctt	ttgaccagcg	tgttaggaag	3060
aacccatcacat	tctttctggg	catcatctcc	agccaaagcat	cctgcgtcta	tgctatcc	3120
aaaggtaaga	atcccgaaat	gacactaaag	gcctctgggt	cctttccccc	tgaagccgc	3180
cattggctct	gttaccaggc	cttccctgtt	aaagttggct	ctcattctgt	catctacaaa	3240
tgtctctgg	gacccctgtt	gacagcccaa	aaactgtgtt	gcccggaaat	cccagaggcg	3300
acaatgacca	tccttaaagc	tgcagctgac	ccagccctaa	gcacagactt	tcagaccatt	3360
ttggactaa						3369

<210> SEQ ID NO 6
<211> LENGTH: 1132
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 6

Met Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu Leu Arg Ser
1 5 10 15

His Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg Arg Leu Gly
20 25 30

Pro Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro Ala Ala Phe Arg
35 40 45

Ala Leu Val Ala Gln Cys Leu Val Cys Val Pro Trp Asp Ala Arg Pro
50 55 60

-continued

Val Ala Arg Val Leu Gln Arg Leu Cys Glu Arg Gly Ala Lys Asn Val			
85	90	95	
Leu Ala Phe Gly Phe Ala Leu Leu Asp Gly Ala Arg Gly Pro Pro			
100	105	110	
Glu Ala Phe Thr Thr Ser Val Arg Ser Tyr Leu Pro Asn Thr Val Thr			
115	120	125	
Asp Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu Leu Leu Arg Arg Val			
130	135	140	
Gly Asp Asp Val Leu Val His Leu Leu Ala Arg Cys Ala Leu Phe Val			
145	150	155	160
Leu Val Ala Pro Ser Cys Ala Tyr Gln Val Cys Gly Pro Pro Leu Tyr			
165	170	175	
Gln Leu Gly Ala Ala Thr Gln Ala Arg Pro Pro Pro His Ala Ser Gly			
180	185	190	
Pro Arg Arg Arg Leu Gly Cys Glu Arg Ala Trp Asn His Ser Val Arg			
195	200	205	
Glu Ala Gly Val Pro Leu Gly Leu Pro Ala Pro Gly Ala Arg Arg Arg			
210	215	220	
Gly Gly Ser Ala Ser Arg Ser Leu Pro Leu Pro Lys Arg Pro Arg Arg			
225	230	235	240
Gly Ala Ala Pro Glu Pro Glu Arg Thr Pro Val Gly Gln Gly Ser Trp			
245	250	255	
Ala His Pro Gly Arg Thr Arg Gly Pro Ser Asp Arg Gly Phe Cys Val			
260	265	270	
Val Ser Pro Ala Arg Pro Ala Glu Glu Ala Thr Ser Leu Glu Gly Ala			
275	280	285	
Leu Ser Gly Thr Arg His Ser His Pro Ser Val Gly Arg Gln His His			
290	295	300	
Ala Gly Pro Pro Ser Thr Ser Arg Pro Pro Arg Pro Trp Asp Thr Pro			
305	310	315	320
Cys Pro Pro Val Tyr Ala Glu Thr Lys His Phe Leu Tyr Ser Ser Gly			
325	330	335	
Asp Lys Glu Gln Leu Arg Pro Ser Phe Leu Leu Ser Ser Leu Arg Pro			
340	345	350	
Ser Leu Thr Gly Ala Arg Arg Leu Val Glu Thr Ile Phe Leu Gly Ser			
355	360	365	
Arg Pro Trp Met Pro Gly Thr Pro Arg Arg Leu Pro Arg Leu Pro Gln			
370	375	380	
Arg Tyr Trp Gln Met Arg Pro Leu Phe Leu Glu Leu Leu Gly Asn His			
385	390	395	400
Ala Gln Cys Pro Tyr Gly Val Leu Leu Lys Thr His Cys Pro Leu Arg			
405	410	415	
Ala Ala Val Thr Pro Ala Ala Gly Val Cys Ala Arg Glu Lys Pro Gln			
420	425	430	
Gly Ser Val Ala Ala Pro Glu Glu Asp Thr Asp Pro Arg Arg Leu			
435	440	445	
Val Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln Val Tyr Gly Phe			
450	455	460	
Val Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu Trp Gly Ser			
465	470	475	480
Arg His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser			

-continued

485	490	495
Leu Gly Lys His Ala Lys Leu Ser	Leu Gln Glu Leu Thr Trp Lys Met	
500	505	510
Ser Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro	Gly Val Gly Cys	
515	520	525
Val Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile	Leu Ala Lys Phe	
530	535	540
Leu His Trp Leu Met Ser Val Tyr Val Val Glu Leu	Leu Arg Ser Phe	
545	550	555
Phe Tyr Val Thr Glu Thr Phe Gln Lys Asn Arg Leu	Phe Phe Tyr	
565	570	575
Arg Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly	Ile Arg Gln His	
580	585	590
Leu Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala	Glu Val Arg Gln	
595	600	605
His Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg	Leu Arg Phe Ile	
610	615	620
Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met	Asp Tyr Val Val	
625	630	635
Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu	Arg Leu Thr Ser	
645	650	655
Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu	Arg Ala Arg Arg	
660	665	670
Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp	Asp Ile His Arg	
675	680	685
Ala Trp Arg Thr Phe Val Leu Arg Val Ala Gln Asp	Pro Pro Pro	
690	695	700
Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala	Tyr Asp Thr Ile	
705	710	715
Pro Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile	Ile Lys Pro Gln	
725	730	735
Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln	Lys Ala Ala His	
740	745	750
Gly His Val Arg Lys Ala Phe Lys Ser His Val Ser	Thr Leu Thr Asp	
755	760	765
Leu Gln Pro Tyr Met Arg Gln Phe Val Ala His	Leu Gln Glu Thr Ser	
770	775	780
Pro Leu Arg Asp Ala Val Val Ile Glu Gln Ser Ser	Ser Leu Asn Glu	
785	790	795
Ala Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe	Met Cys His His	
805	810	815
Ala Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys	Gln Gly Ile Pro	
820	825	830
Gln Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu	Cys Tyr Gly Asp	
835	840	845
Met Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp	Gly Leu Leu Leu	
850	855	860
Arg Leu Val Asp Asp Phe Leu Leu Val Thr Pro His	Leu Thr His Ala	
865	870	875
Lys Thr Phe Leu Arg Thr Leu Val Arg Gly Val Pro	Glu Tyr Gly Cys	
885	890	895

-continued

Val Val Asn Leu Arg Lys Thr Val Val Asn Phe Pro Val Glu Asp Glu
900 905 910

Ala Leu Gly Gly Thr Ala Phe Val Gln Met Pro Ala His Gly Leu Phe
915 920 925

Pro Trp Cys Gly Leu Leu Leu Asp Thr Arg Thr Leu Glu Val Gln Ser
930 935 940

Asp Tyr Ser Ser Tyr Ala Arg Thr Ser Ile Arg Ala Ser Leu Thr Phe
945 950 955 960

Asn Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys Leu Phe Gly
965 970 975

Val Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu Gln Val Asn
980 985 990

Ser Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu Leu Leu Gln
995 1000 1005

Ala Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe His Gln
1010 1015 1020

Gln Val Trp Lys Asn Pro Thr Phe Phe Leu Arg Val Ile Ser Asp
1025 1030 1035

Thr Ala Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly
1040 1045 1050

Met Ser Leu Gly Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu
1055 1060 1065

Ala Val Gln Trp Leu Cys His Gln Ala Phe Leu Leu Lys Leu Thr
1070 1075 1080

Arg His Arg Val Thr Tyr Val Pro Leu Leu Gly Ser Leu Arg Thr
1085 1090 1095

Ala Gln Thr Gln Leu Ser Arg Lys Leu Pro Gly Thr Thr Leu Thr
1100 1105 1110

Ala Leu Glu Ala Ala Ala Asn Pro Ala Leu Pro Ser Asp Phe Lys
1115 1120 1125

Thr Ile Leu Asp
1130

<210> SEQ ID NO 7

<211> LENGTH: 4039

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 7

ctctcctcgc ggcgcgagtt tcaggcagcg ctgcgtctg ctgcgcacgt gggaaagccct	60
ggccccggcc acccccgccgat tgccgcgcgc tccccgctgc cgagccgtgc gtcctctgt	120
gcccgcggccat taccgcgagg tgctgccgtt ggccacgttc gtgcggcgcc tggggcccca	180
gggctggcggtt ctgggtgcagg gcggggaccc ggccggcttc cgccgcgtgg tggcccagt	240
cctgggtgtgc gtgcctggg acgcacggcc gccccccgcc gccccctcttccgcaggat	300
gtcctgcctg aaggagctgg tggcccgagt gctgcagagg ctgtgcgagc gccggcgccaa	360
gaacgtgtgc gccttcggct tcgcgtgtgtt ggacggggcc cggggggggcc cccccggggc	420
cttcaccacc acgcgtgcgcac gtcacacttc caacacgggtt accgacgcac tgccggggag	480
cggggcggtgg gggctgtgtgc tgccgcgcgtt gggcgacgac gtgctgggttc acctgtggc	540
acgcgtgcgcg ctctttgtgc tggtggttcc cagctgcgcctt acccagggtgt gccggccggcc	600

-continued

gctgtaccag	ctcggegctg	ccactcaggo	ccggcccccg	ccacacgcta	gtggaccgg	660
aaggcgtctg	ggatgegaac	gggcctggaa	ccatagegta	agggaggccg	gggtccccct	720
gggcctgcca	gccccgggtg	cgaggaggcg	cggggccagt	gccagccaa	gtctggcgtt	780
gccccaggagg	cccaggcgta	gcgtgcggcc	tgagccggag	eggacgccc	ttgggcaggg	840
gtcctggcc	cacccgggca	ggacgcgtgg	accgagtgac	cgtggtttct	gtgtggtgtc	900
acctgccaga	cccggcgaag	aagccaccc	tttggaggg	gcgtctctg	gcacgcgcca	960
ctccccaccca	tccgtggcc	gccagcacca	cgcgggcccc	ccatccacat	cgcggccacc	1020
acgtccccgg	gacacgcctt	gtccccgggt	gtacgcccag	accaagca	tccctctactc	1080
ctcaggcgac	aaggagcagc	tgccggccctc	cttcctactc	agctctctga	ggcccgccct	1140
gactggcgct	cggaggctcg	tggagaccat	ctttctgggt	tccaggccct	ggatgcagg	1200
gactccccgc	agggttgc	gcctgcccc	gcgtactgg	caaatacgcc	ccctgtttct	1260
ggagactgttt	gggaaccacg	cgcagtgc	ctacgggggt	tccctcaaga	cgcactgccc	1320
gctgcgagct	gogggtcaccc	cagcagccgg	tgtctgtgcc	cgggagaagc	cccagggtct	1380
tgtggccggcc	cccgaggagg	aggacacaga	cccccgctcg	ctgggtgc	tgctccgcca	1440
gcacagcagc	ccctggcagg	tgtacggctt	cgtgcgggccc	tgcctgcg	ggctggtgcc	1500
ccaggccctc	tggggctcca	ggcacaacga	acggccgttc	ctcaggaaaca	ccaagaagtt	1560
catctccctg	gggaagcatg	ccaagctctc	gctgcaggag	ctgacgtgga	agatgagcgt	1620
cggggactgc	gcttggctgc	gcaggagccc	aggggttggc	tgtgttccgg	ccgcagagca	1680
ccgtctcggt	gaggagatcc	tggcaagtt	cctgcactgg	ctgatgagtg	tgtacgtcgt	1740
cgagctgctc	aggctttct	tttatgtcad	ggagaccacg	tttcaaaaga	acaggctctt	1800
tttctaccgg	aagagtgtct	ggagcaagtt	gcaaaggatt	ggaatcagac	agcacttgaa	1860
gagggtgcag	ctgcgggagc	tgtcggaaagc	agaggtcagg	cagcatcg	aagccaggcc	1920
cgcctctgtc	acgtccagac	tccgtttcat	ccccaaagcc	gacgggctgc	ggccgattgt	1980
gaacatggac	tacgtcgtgg	gagccagaac	gttccgcaga	aaaaagaggg	ccgagcgct	2040
cacctcgagg	gtgaaggcac	tgttcagcgt	gctcaactac	gagcgggccc	ggcgc	2100
cctcctggcc	gcctctgtgc	tgggcctgg	cgatatccac	agggcctggc	gcaccttcgt	2160
gctgcgtgt	cgggcccgagg	acccggccgc	tgagctgtac	tttgtcaagg	tggatgtgac	2220
gggcgcgtac	gacaccatcc	cccaggacag	gctcacggag	gtcatgc	ccatcatcaa	2280
accccaagaac	acgtactgcg	tgccgtcggt	tgcctgtggc	cagaaggccg	cccatgggca	2340
cgtccgc	gacccataga	gccacgtctc	taccttgaca	gacccatcg	cgtacatgcg	2400
acagttcgtg	gctcacctgc	aggagaccag	cccgtcgagg	gtgcgcgtc	tcatcgagca	2460
gagctctcc	ctgaatgagg	ccagcagtgg	cctcttcgac	gtttctac	gtttcatgt	2520
ccaccacgc	gtgcgcatca	ggggcaagtc	ctacgtccag	tgc	ccgcagg	2580
ctccatcttc	tccacgtgc	tctgcagcct	gtgctacggc	gacatggaga	acaagctgtt	2640
tgcggggatt	cgccgggacg	ggctgctcct	gcgttgggt	gatgatttct	tgttggtgac	2700
acotcacctc	acccacgcga	aaacccct	caggaccctg	gtccgcgg	tccctgagta	2760
tggctgcgtg	gtgaacttgc	ggaagacagt	ggtgaacttc	cctgtagaag	acgaggccct	2820
gggtggcaca	gctttgttc	agatgcggc	ccacggccta	ttccctgg	gcggcctgct	2880

-continued

gctggatacc	cgaccctgg	agggtcagag	cgactactcc	agctatgcc	ggacccat	2940
cagagccagt	ctcacccca	accgcggctt	caaggctggg	aggaacatgc	gtcgaaact	3000
ctttggggtc	ttgcggctga	agtgtcacag	cctgtttctg	gatttgagg	tgaacaggct	3060
ccagacggtg	tgcaccaaca	tctacaagat	cctcctgtcg	caggcgtaca	ggttcacgc	3120
atgtgtgtcg	cagctcccat	ttcatcagca	agtttggaa	aaccccacat	tttcctgcg	3180
cgtcatctct	gacacggcct	ccctctgcta	ctccatctg	aaagccaaga	acgcaggat	3240
gtcgctgggg	gccaaggggcg	ccgcggcccc	tctgcctcc	gaggccgtgc	agtggctgt	3300
ccaccaagca	ttcctgctca	agctgactcg	acaccgtgtc	acctacgtgc	cactcctgg	3360
gtcaactcagg	acagcccaga	cgcagctgag	tccgaagctc	ccggggacga	cgctgactgc	3420
cctggaggcc	gcagccaaacc	cggcactgco	ctcagacttc	aagaccatcc	tggactgtat	3480
gccacccggcc	cacagccagg	ccgagagcag	acaccagcag	ccctgtacg	ccgggctcta	3540
cgtcccagg	agggagggggc	ggcccacacc	caggccccga	ccgctgggag	tctgaggcct	3600
gagtgagtgt	ttggccgagg	cctgcatgtc	cggctgaagg	ctgagtgtcc	ggctgaggcc	3660
tgagcgagt	tccagccaag	ggctgagtgt	ccagcacacc	tgcgtcttc	acttccccac	3720
aggctggcgc	tccggctccac	cccaggggca	gctttcctc	accaggagcc	cggttccac	3780
tcccccacata	ggaatagtcc	atccccagat	tccgcattgt	tcacccctcg	ccctgcctc	3840
ctttgccttc	caccccccacc	atccaggtgg	agaccctgag	aaggaccctg	ggagctctgg	3900
gaatttggag	tgaccaaagg	tgtgccctgt	acacaggcga	ggaccctgca	cctggatggg	3960
ggtccctgtg	ggtcaaattt	gggggaggtg	ctgtgggagt	aaaatactga	atatatgagt	4020
tttcagttt	tgaaaaaaaa					4039

<210> SEQ ID NO 8
<211> LENGTH: 1069
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 8

Met	Pro	Arg	Ala	Pro	Arg	Cys	Arg	Ala	Val	Arg	Ser	Leu	Leu	Arg	Ser
1									5		10			15	

His	Tyr	Arg	Glu	Val	Leu	Pro	Leu	Ala	Thr	Phe	Val	Arg	Arg	Leu	Gly	
														20	25	30

Pro	Gln	Gly	Trp	Arg	Leu	Val	Gln	Arg	Gly	Asp	Pro	Ala	Ala	Phe	Arg	
														35	40	45

Ala	Leu	Val	Ala	Gln	Cys	Leu	Val	Cys	Val	Pro	Trp	Asp	Ala	Arg	Pro	
														50	55	60

Pro	Pro	Ala	Ala	Pro	Ser	Phe	Arg	Gln	Val	Ser	Cys	Leu	Lys	Glu	Leu		
														65	70	75	80

Val	Ala	Arg	Val	Leu	Gln	Arg	Leu	Cys	Glu	Arg	Gly	Ala	Lys	Asn	Val	
														85	90	95

Leu	Ala	Phe	Gly	Phe	Ala	Leu	Leu	Asp	Gly	Ala	Arg	Gly	Pro	Pro		
														100	105	110

Glu	Ala	Phe	Thr	Thr	Ser	Val	Arg	Ser	Tyr	Leu	Pro	Asn	Thr	Val	Thr	
														115	120	125

Asp	Ala	Leu	Arg	Gly	Ser	Gly	Ala	Trp	Gly	Leu	Leu	Leu	Arg	Arg	Val	
														130	135	140

-continued

Gly	Asp	Asp	Val	Leu	Val	His	Leu	Leu	Ala	Arg	Cys	Ala	Leu	Phe	Val
145			150			155			160						
Leu	Val	Ala	Pro	Ser	Cys	Ala	Tyr	Gln	Val	Cys	Gly	Pro	Pro	Leu	Tyr
	165				170			175							
Gln	Leu	Gly	Ala	Ala	Thr	Gln	Ala	Arg	Pro	Pro	Pro	His	Ala	Ser	Gly
	180				185			190							
Pro	Arg	Arg	Arg	Leu	Gly	Cys	Glu	Arg	Ala	Trp	Asn	His	Ser	Val	Arg
	195				200			205							
Glu	Ala	Gly	Val	Pro	Leu	Gly	Leu	Pro	Ala	Pro	Gly	Ala	Arg	Arg	Arg
	210				215			220							
Gly	Gly	Ser	Ala	Ser	Arg	Ser	Leu	Pro	Leu	Pro	Lys	Arg	Pro	Arg	Arg
	225				230			235			240				
Gly	Ala	Ala	Pro	Glu	Pro	Glu	Arg	Thr	Pro	Val	Gly	Gln	Gly	Ser	Trp
	245				250			255							
Ala	His	Pro	Gly	Arg	Thr	Arg	Gly	Pro	Ser	Asp	Arg	Gly	Phe	Cys	Val
	260				265			270							
Val	Ser	Pro	Ala	Arg	Pro	Ala	Glu	Glu	Ala	Thr	Ser	Leu	Glu	Gly	Ala
	275				280			285							
Leu	Ser	Gly	Thr	Arg	His	Ser	His	Pro	Ser	Val	Gly	Arg	Gln	His	His
	290				295			300							
Ala	Gly	Pro	Pro	Ser	Thr	Ser	Arg	Pro	Pro	Arg	Pro	Trp	Asp	Thr	Pro
	305				310			315			320				
Cys	Pro	Pro	Val	Tyr	Ala	Glu	Thr	Lys	His	Phe	Leu	Tyr	Ser	Ser	Gly
	325				330			335							
Asp	Lys	Glu	Gln	Leu	Arg	Pro	Ser	Phe	Leu	Leu	Ser	Ser	Leu	Arg	Pro
	340				345			350							
Ser	Leu	Thr	Gly	Ala	Arg	Arg	Leu	Val	Glu	Thr	Ile	Phe	Leu	Gly	Ser
	355				360			365							
Arg	Pro	Trp	Met	Pro	Gly	Thr	Pro	Arg	Arg	Leu	Pro	Arg	Leu	Pro	Gln
	370				375			380							
Arg	Tyr	Trp	Gln	Met	Arg	Pro	Leu	Phe	Leu	Glu	Leu	Gly	Asn	His	
	385				390			395			400				
Ala	Gln	Cys	Pro	Tyr	Gly	Val	Leu	Leu	Lys	Thr	His	Cys	Pro	Leu	Arg
	405				410			415							
Ala	Ala	Val	Thr	Pro	Ala	Ala	Gly	Val	Cys	Ala	Arg	Glu	Lys	Pro	Gln
	420				425			430							
Gly	Ser	Val	Ala	Ala	Pro	Glu	Glu	Asp	Thr	Asp	Pro	Arg	Arg	Leu	
	435				440			445							
Val	Gln	Leu	Leu	Arg	Gln	His	Ser	Ser	Pro	Trp	Gln	Val	Tyr	Gly	Phe
	450				455			460							
Val	Arg	Ala	Cys	Leu	Arg	Arg	Leu	Val	Pro	Pro	Gly	Leu	Trp	Gly	Ser
	465				470			475			480				
Arg	His	Asn	Glu	Arg	Arg	Phe	Leu	Arg	Asn	Thr	Lys	Lys	Phe	Ile	Ser
	485				490			495							
Leu	Gly	Lys	His	Ala	Lys	Leu	Ser	Leu	Gln	Glu	Leu	Thr	Trp	Lys	Met
	500				505			510							
Ser	Val	Arg	Asp	Cys	Ala	Trp	Leu	Arg	Arg	Ser	Pro	Gly	Val	Gly	Cys
	515				520			525							
Val	Pro	Ala	Ala	Glu	His	Arg	Leu	Arg	Glu	Glu	Ile	Leu	Ala	Lys	Phe
	530				535			540							
Leu	His	Trp	Leu	Met	Ser	Val	Tyr	Val	Val	Glu	Leu	Leu	Arg	Ser	Phe

-continued

545	550	555	560												
Phe	Tyr	Val	Thr	Glu	Thr	Thr	Phe	Gln	Lys	Asn	Arg	Leu	Phe	Phe	Tyr
565				570				575							
Arg	Lys	Ser	Val	Trp	Ser	Lys	Leu	Gln	Ser	Ile	Gly	Ile	Arg	Gln	His
580					585				590						
Leu	Lys	Arg	Val	Gln	Leu	Arg	Glu	Leu	Ser	Glu	Ala	Glu	Val	Arg	Gln
595				600			605								
His	Arg	Glu	Ala	Arg	Pro	Ala	Leu	Leu	Thr	Ser	Arg	Leu	Arg	Phe	Ile
610				615			620								
Pro	Lys	Pro	Asp	Gly	Leu	Arg	Pro	Ile	Val	Asn	Met	Asp	Tyr	Val	Val
625				630			635			640					
Gly	Ala	Arg	Thr	Phe	Arg	Arg	Glu	Lys	Arg	Ala	Glu	Arg	Leu	Thr	Ser
645					650			655							
Arg	Val	Lys	Ala	Leu	Phe	Ser	Val	Leu	Asn	Tyr	Glu	Arg	Ala	Arg	Arg
660				665			670								
Pro	Gly	Leu	Leu	Gly	Ala	Ser	Val	Leu	Gly	Leu	Asp	Asp	Ile	His	Arg
675				680			685								
Ala	Trp	Arg	Thr	Phe	Val	Leu	Arg	Val	Arg	Ala	Gln	Asp	Pro	Pro	Pro
690				695			700								
Glu	Leu	Tyr	Phe	Val	Lys	Val	Asp	Val	Thr	Gly	Ala	Tyr	Asp	Thr	Ile
705				710			715			720					
Pro	Gln	Asp	Arg	Leu	Thr	Glu	Val	Ile	Ala	Ser	Ile	Ile	Lys	Pro	Gln
725					730			735							
Asn	Thr	Tyr	Cys	Val	Arg	Arg	Tyr	Ala	Val	Val	Gln	Lys	Ala	Ala	His
740					745			750							
Gly	His	Val	Arg	Lys	Ala	Phe	Lys	Ser	His	Val	Ser	Thr	Leu	Thr	Asp
755					760			765							
Leu	Gln	Pro	Tyr	Met	Arg	Gln	Phe	Val	Ala	His	Leu	Gln	Glu	Thr	Ser
770				775			780								
Pro	Leu	Arg	Asp	Ala	Val	Val	Ile	Glu	Gln	Ser	Ser	Ser	Leu	Asn	Glu
785				790			795			800					
Ala	Ser	Ser	Gly	Leu	Phe	Asp	Val	Phe	Leu	Arg	Phe	Met	Cys	His	His
805					810			815							
Ala	Val	Arg	Ile	Arg	Gly	Lys	Ser	Tyr	Val	Gln	Cys	Gln	Gly	Ile	Pro
820				825			830								
Gln	Gly	Ser	Ile	Leu	Ser	Thr	Leu	Leu	Cys	Ser	Leu	Cys	Tyr	Gly	Asp
835				840			845								
Met	Glu	Asn	Lys	Leu	Phe	Ala	Gly	Ile	Arg	Arg	Asp	Gly	Leu	Leu	Leu
850				855			860								
Arg	Leu	Val	Asp	Asp	Phe	Leu	Leu	Val	Thr	Pro	His	Leu	Thr	His	Ala
865				870			875			880					
Lys	Thr	Phe	Leu	Ser	Tyr	Ala	Arg	Thr	Ser	Ile	Arg	Ala	Ser	Leu	Thr
885				890			895								
Phe	Asn	Arg	Gly	Phe	Lys	Ala	Gly	Arg	Asn	Met	Arg	Arg	Lys	Leu	Phe
900				905			910								
Gly	Val	Leu	Arg	Leu	Lys	Cys	His	Ser	Leu	Phe	Leu	Asp	Leu	Gln	Val
915				920			925								
Asn	Ser	Leu	Gln	Thr	Val	Cys	Thr	Asn	Ile	Tyr	Lys	Ile	Leu	Leu	Leu
930				935			940								
Gln	Ala	Tyr	Arg	Phe	His	Ala	Cys	Val	Leu	Gln	Leu	Pro	Phe	His	Gln
945				950			955			960					

-continued

Gln Val Trp Lys Asn Pro Thr Phe Phe Leu Arg Val Ile Ser Asp Thr
965 970 975

Ala Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly Met Ser
980 985 990

Leu Gly Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu Ala Val Gln
995 1000 1005

Trp Leu Cys His Gln Ala Phe Leu Leu Lys Leu Thr Arg His Arg
1010 1015 1020

Val Thr Tyr Val Pro Leu Leu Gly Ser Leu Arg Thr Ala Gln Thr
1025 1030 1035

Gln Leu Ser Arg Lys Leu Pro Gly Thr Thr Leu Thr Ala Leu Glu
1040 1045 1050

Ala Ala Ala Asn Pro Ala Leu Pro Ser Asp Phe Lys Thr Ile Leu
1055 1060 1065

Asp

<210> SEQ ID NO 9
<211> LENGTH: 3850
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

ctctccctcg	ggcgcgagtt	tcaggcageg	ctgcgtcctg	ctgcgcacgt	ggaaagccct	60
ggccccggcc	accccccgcg	tgccgcgcgc	tccccgtgc	cgagcgcgtgc	gtccctgct	120
gcccagccac	taccgcgagg	tgctgccgt	ggccacgttc	gtgcggcgcc	tggggcccca	180
gggctggcg	ctgggtcagc	gccccggacc	ggggggatcc	cgcgcgatgg	tggcccaagt	240
cctgggtgc	gtgccttggg	acgcacggcc	gccccccgc	gccccctct	tccgcacagg	300
gtcctgcctg	aaggagctgg	tggcccgagt	gtgcagagg	ctgtgcgagc	gccccggcga	360
gaacgtgtc	gccttcggct	tcgcgcgtgt	ggacggggcc	cgccccggcc	cccccgaggc	420
cttcaccacc	agcgtgcgca	gctacctgc	caacacggtg	accgaegcac	tgcggggag	480
ccccggcgtgg	gggctgtgc	tgccgcgcgt	gggcgacgac	gtgctgggtc	acctgtggc	540
acgctgegca	ctctttgtgc	tggtggctcc	cagctgcgcc	taccaggtgt	gccccggcc	600
gctgtaccag	ctcgccgtg	ccactcaggg	ccggcccccc	ccacacgcta	gtggaccgg	660
aaggcgtctg	ggatgcgaac	gggcctggaa	ccatagcgta	agggaggccg	gggtccccct	720
gggcctgcca	gccccgggtg	cgaggaggcg	cgggggcagt	gccagccgaa	gtctgcgtt	780
cccccaaggagg	cccaggcggt	gctgcgtcccc	ttagccggag	cgacgcggcc	ttggggcagg	840
gtcctgggcc	cacccgggca	ggacgcgtgg	accgagtgc	cgtgggttct	gtgtgggtgc	900
acctgccaga	cccgccgaag	aagccacctc	tttggagggt	gcgcgtctctg	gcacgcgcca	960
ctccccaccca	tccgtgggcc	gccagcacca	cgcccccccc	ccatccacat	cgccggccacc	1020
acgtccctgg	gacacgcctt	gtcccccgtt	gtacgcgcag	accaagoact	tccctactc	1080
ctcaggcgac	aaggagcagc	tgccggccctc	cttcctactc	agctctctga	ggcccagcct	1140
gactggcgct	cggaggctcg	tggagaccat	cttctgggt	tccaggccct	ggatgccagg	1200
gactccccgc	agggttgcggc	gcctgcccc	gcgcgtactgg	caaatgcggc	ccctgtttct	1260
ggagctgctt	ggaaaccacg	cgcagtgc	ctacggggtg	ctcctcaaga	cgcaactgccc	1320

-continued

gctgcgact gcggtcaccc cagcagccgg tgcgtgtgcc	cgggagaagc cccagggttc	1380
tgtggcgccc cccgaggagg aggacacaga cccccgtcgc	ctgggtcagc tgctccgcca	1440
gcacagcagc ccctggcagg tgtacggctt	cgtgcgggcc tgcctgcgcc ggctgggtgcc	1500
cccaaggttc tggggctcca ggcacaacga acgcccgttc	ctcaggaaca ccaagaagtt	1560
catctccctg gggaaagcatg ccaagcttc	gctgcaggag ctgacgtgga agatgagcgt	1620
gccccactgc gcttggctgc gcaggagccc	aggggttggc tgcgttccgg ccgcagagca	1680
cctgtctcgat	tgccaaagtt cctgcactgg ctgatgagtg tgcgtcgt	1740
cgagctgctc aggttttct tttatgtcac	ggagaccacg ttcaaaaaga acaggcttt	1800
tttctacogg aagagtgtct	ggagcaagtt gcaaaggatt ggaatcagac agcacttgaa	1860
gagggtgcag ctgcgggagc tgcggaaagc	agaggtcagg cagcatcgaa aagccaggcc	1920
cgcctcgctc acgtccagac tccggttcat	ccccaaagctt gacgggtgtc ggccgattgt	1980
gaacatggac tacgttgtgg gagccagaac	gttccgcaga gaaaagaggg ccgagcgtct	2040
cacctcgagg gtgaaggcac tggtaagcgt	gctcaactac gagcggggcgc ggcccccgg	2100
cctcctggcc	gcctctgtgc tgggctcgaa cgatatccac agggctggc gcacccgt	2160
gctgcgtgtc	cgggcccccagg acceggccgc ttagctgtac ttgtcaagg tggatgtgac	2220
gggcgcgtac	gacaccatcc cccaggacag gctcacggag gtcategcac gcatcatcaa	2280
accccaaaac acgtactgcg	tgccgtggta tgcgtggc cagaaggccg cccatggca	2340
cgtccgcac	gcctcaaga gccacgtctc tacctgaca gacccctcagc cgtacatgcg	2400
acagttcg	gctcacctgc aggagaccag cccgctgagg gatgcgtcg tcacatgca	2460
gagctctcc	ctgaatgagg ccagcgtgg cctcttcgac gtttcttac gttcatgt	2520
ccaccacgc	gtgcgtatca gggcaagtc ctacgtccag tgccagggg tcccgaggg	2580
ctccatctc	tccacgctgc tctgcagctt gtgtacggc gacatggaga acaagctgtt	2640
tgcgggatt	cgccgggacg ggctgtcttgc gctttgggt gatgtttct tggatgtgac	2700
acctcacctc	acccacgcga aaacccctt cagctatgcc cggacccca tcagagccag	2760
tctcaccc	aaccggggct tcaaggctgg gaggaacatg gtcgtcaaac tctttgggt	2820
cttgcggctg	aagtgtcaca gcctgtttct ggatttgcag gtgaacagcc tccagacgg	2880
gtgcaccaac	atctacaaga tccctctgtc gcaggcgtac aggtttcacc catgtgtgt	2940
gcagctccca	tttcatcagc aagtttggaa gaaccccaaa ttttctgc gctcatctc	3000
tgacacggcc	tccctctgtc actccatctt gaaagccaag aacgcaggaa tgccgttgg	3060
ggccaagggc	gcccggggcc ctctgcccgc cgaggccgtg cagtggtgtt gcccaccaac	3120
attcctgctc	aagctgactc gacaccgtgt cacctacgtt ccactctgg ggtcactcag	3180
gacagcccaag	acgcagctga gtcggaaagct cccggggacg acgctgactg ccctggaggc	3240
cgcagccaaac	ccggcactgc cctcagactt caagaccatc ctggactgtat gcccacccgc	3300
ccacagccag	gcccggggcc gacaccatc gcccgttac gcccgggtt acgtcccaagg	3360
gagggagggg	cgccccacac ccaggccccgc accgctggaa gtctgaggcc tgagtgtgt	3420
tttggccgg	gcctgcgtgtt ccggctgtaa gctgagtgtc cggctgaggc ctgagcgt	3480
gtccagccaa	gggctgtgtt tccagcacac ctgcgttccat cacttccca caggctggcg	3540
ctcggttcca	ccccaggggcc agttttctt caccaggacg ccggcttcca ctccccacat	3600

-continued

aggaatagtc	catccccaga	ttcgccattg	ttcacccctc	gcctgcct	ccttgcctt	3660
ccaccccccac	catccagggtg	gagaccctga	gaaggaccct	gggagctctg	ggaatttggaa	3720
gtgaccaaaag	gtgtgcctg	tacacaggcg	aggaccctgc	acctggatgg	gggtccctgt	3780
gggtcaaatt	ggggggaggt	gctgtggag	taaaatactg	aatatatgag	tttttcagtt	3840
ttgaaaaaaaaa						3850

<210> SEQ ID NO 10
<211> LENGTH: 1122
<212> TYPE: PRT
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 10

Met	Thr	Arg	Ala	Pro	Arg	Cys	Pro	Ala	Val	Arg	Ser	Leu	Leu	Arg	Ser
1									5	10				15	
Arg	Tyr	Arg	Glu	Val	Trp	Pro	Leu	Ala	Thr	Phe	Val	Arg	Arg	Leu	Gly
				20					25					30	
Pro	Glu	Gly	Arg	Arg	Leu	Val	Gln	Pro	Gly	Asp	Pro	Lys	Ile	Tyr	Arg
					35		40					45			
Thr	Leu	Val	Ala	Gln	Cys	Leu	Val	Cys	Met	His	Trp	Gly	Ser	Gln	Pro
					50		55				60				
Pro	Pro	Ala	Asp	Leu	Ser	Phe	His	Gln	Val	Ser	Ser	Leu	Lys	Glu	Leu
				65		70			75			80			
Val	Ala	Arg	Val	Val	Gln	Arg	Leu	Cys	Glu	Arg	Asn	Glu	Arg	Asn	Val
					85		90				95				
Leu	Ala	Phe	Gly	Phe	Glu	Leu	Leu	Asn	Glu	Ala	Arg	Gly	Pro	Pro	
				100		105					110				
Met	Ala	Phe	Thr	Ser	Ser	Val	Arg	Ser	Tyr	Leu	Pro	Asn	Thr	Val	Ile
				115		120			125						
Glu	Thr	Leu	Arg	Val	Ser	Gly	Ala	Trp	Met	Leu	Leu	Ser	Arg	Val	
				130		135			140						
Gly	Asp	Asp	Leu	Leu	Val	Tyr	Leu	Leu	Ala	His	Cys	Ala	Leu	Tyr	Leu
				145		150			155			160			
Leu	Val	Pro	Pro	Ser	Cys	Ala	Tyr	Gln	Val	Cys	Gly	Ser	Pro	Leu	Tyr
				165		170			175						
Gln	Ile	Cys	Ala	Thr	Thr	Asp	Ile	Trp	Pro	Ser	Val	Ser	Ala	Ser	Tyr
				180		185			190						
Arg	Pro	Thr	Arg	Pro	Val	Gly	Arg	Asn	Phe	Thr	Asn	Leu	Arg	Phe	Leu
				195		200			205						
Gln	Gln	Ile	Lys	Ser	Ser	Arg	Gln	Glu	Ala	Pro	Lys	Pro	Leu	Ala	
				210		215			220						
Leu	Pro	Ser	Arg	Gly	Thr	Lys	Arg	His	Leu	Ser	Leu	Thr	Ser	Thr	Ser
				225		230			235			240			
Val	Pro	Ser	Ala	Lys	Lys	Ala	Arg	Cys	Tyr	Pro	Val	Pro	Arg	Val	Glu
				245		250			255			255			
Glu	Gly	Pro	His	Arg	Gln	Val	Leu	Pro	Thr	Pro	Ser	Gly	Lys	Ser	Trp
				260		265			270			270			
Val	Pro	Ser	Pro	Ala	Arg	Ser	Pro	Glu	Val	Pro	Thr	Ala	Glu	Lys	Asp
				275		280			285			285			
Leu	Ser	Ser	Lys	Gly	Lys	Val	Ser	Asp	Leu	Ser	Leu	Ser	Gly	Ser	Val
				290		295			300						
Cys	Cys	Lys	His	Lys	Pro	Ser	Ser	Thr	Ser	Leu	Leu	Ser	Pro	Pro	Arg

-continued

305	310	315	320
Gln Asn Ala Phe Gln Leu Arg Pro Phe Ile Glu Thr Arg His Phe Leu			
325	330	335	
Tyr Ser Arg Gly Asp Gly Gln Glu Arg Leu Asn Pro Ser Phe Leu Leu			
340	345	350	
Ser Asn Leu Gln Pro Asn Leu Thr Gly Ala Arg Arg Leu Val Glu Ile			
355	360	365	
Ile Phe Leu Gly Ser Arg Pro Arg Thr Ser Gly Pro Leu Cys Arg Thr			
370	375	380	
His Arg Leu Ser Arg Arg Tyr Trp Gln Met Arg Pro Leu Phe Gln Gln			
385	390	395	400
Leu Leu Val Asn His Ala Glu Cys Gln Tyr Val Arg Leu Leu Arg Ser			
405	410	415	
His Cys Arg Phe Arg Thr Ala Asn Gln Gln Val Thr Asp Ala Leu Asn			
420	425	430	
Thr Ser Pro Pro His Leu Met Asp Leu Leu Arg Leu His Ser Ser Pro			
435	440	445	
Trp Gln Val Tyr Gly Phe Leu Arg Ala Cys Leu Cys Lys Val Val Ser			
450	455	460	
Ala Ser Leu Trp Gly Thr Arg His Asn Glu Arg Arg Phe Phe Lys Asn			
465	470	475	480
Leu Lys Lys Phe Ile Ser Leu Gly Lys Tyr Gly Lys Leu Ser Leu Gln			
485	490	495	
Glu Leu Met Trp Lys Met Lys Val Glu Asp Cys His Trp Leu Arg Ser			
500	505	510	
Ser Pro Gly Lys Asp Arg Val Pro Ala Ala Glu His Arg Leu Arg Glu			
515	520	525	
Arg Ile Leu Ala Thr Phe Leu Phe Trp Leu Met Asp Thr Tyr Val Val			
530	535	540	
Gln Leu Leu Arg Ser Phe Phe Tyr Ile Thr Glu Ser Thr Phe Gln Lys			
545	550	555	560
Asn Arg Leu Phe Phe Tyr Arg Lys Ser Val Trp Ser Lys Leu Gln Ser			
565	570	575	
Ile Gly Val Arg Gln His Leu Glu Arg Val Arg Leu Arg Glu Leu Ser			
580	585	590	
Gln Glu Glu Val Arg His His Gln Asp Thr Trp Leu Ala Met Pro Ile			
595	600	605	
Cys Arg Leu Arg Phe Ile Pro Lys Pro Asn Gly Leu Arg Pro Ile Val			
610	615	620	
Asn Met Ser Tyr Ser Met Gly Thr Arg Ala Leu Gly Arg Arg Lys Gln			
625	630	635	640
Ala Gln His Phe Thr Gln Arg Leu Lys Thr Leu Phe Ser Met Leu Asn			
645	650	655	
Tyr Glu Arg Thr Lys His Pro His Leu Met Gly Ser Ser Val Leu Gly			
660	665	670	
Met Asn Asp Ile Tyr Arg Thr Trp Arg Ala Phe Val Leu Arg Val Arg			
675	680	685	
Ala Leu Asp Gln Thr Pro Arg Met Tyr Phe Val Lys Ala Asp Val Thr			
690	695	700	
Gly Ala Tyr Asp Ala Ile Pro Gln Gly Lys Leu Val Glu Val Val Ala			
705	710	715	720

-continued

Asn Met Ile Arg His Ser Glu Ser Thr Tyr Cys Ile Arg Gln Tyr Ala
 725 730 735
 Val Val Arg Arg Asp Ser Gln Gly Gln Val His Lys Ser Phe Arg Arg
 740 745 750
 Gln Val Thr Thr Leu Ser Asp Leu Gln Pro Tyr Met Gly Gln Phe Leu
 755 760 765
 Lys His Leu Gln Asp Ser Asp Ala Ser Ala Leu Arg Asn Ser Val Val
 770 775 780
 Ile Glu Gln Ser Ile Ser Met Asn Glu Ser Ser Ser Ser Leu Phe Asp
 785 790 795 800
 Phe Phe Leu His Phe Leu Arg His Ser Val Val Lys Ile Gly Asp Arg
 805 810 815
 Cys Tyr Thr Gln Cys Gln Gly Ile Pro Gln Gly Ser Ser Leu Ser Thr
 820 825 830
 Leu Leu Cys Ser Leu Cys Phe Gly Asp Met Glu Asn Lys Leu Phe Ala
 835 840 845
 Glu Val Gln Arg Asp Gly Leu Leu Leu Arg Phe Val Asp Asp Phe Leu
 850 855 860
 Leu Val Thr Pro His Leu Asp Gln Ala Lys Thr Phe Leu Ser Thr Leu
 865 870 875 880
 Val His Gly Val Pro Glu Tyr Gly Cys Met Ile Asn Leu Gln Lys Thr
 885 890 895
 Val Val Asn Phe Pro Val Glu Pro Gly Thr Leu Gly Ala Ala Pro
 900 905 910
 Tyr Gln Leu Pro Ala His Cys Leu Phe Pro Trp Cys Gly Leu Leu Leu
 915 920 925
 Asp Thr Gln Thr Leu Glu Val Phe Cys Asp Tyr Ser Gly Tyr Ala Gln
 930 935 940
 Thr Ser Ile Lys Thr Ser Leu Thr Phe Gln Ser Val Phe Lys Ala Gly
 945 950 955 960
 Lys Thr Met Arg Asn Lys Leu Leu Ser Val Leu Arg Leu Lys Cys His
 965 970 975
 Gly Leu Phe Leu Asp Leu Gln Val Asn Ser Leu Gln Thr Val Cys Ile
 980 985 990
 Asn Ile Tyr Lys Ile Phe Leu Leu Gln Ala Tyr Arg Phe His Ala Cys
 995 1000 1005
 Val Ile Gln Leu Pro Phe Asp Gln Arg Val Arg Lys Asn Leu Thr
 1010 1015 1020
 Phe Phe Leu Gly Ile Ile Ser Ser Gln Ala Ser Cys Cys Tyr Ala
 1025 1030 1035
 Ile Leu Lys Val Lys Asn Pro Gly Met Thr Leu Lys Ala Ser Gly
 1040 1045 1050
 Ser Phe Pro Pro Glu Ala Ala His Trp Leu Cys Tyr Gln Ala Phe
 1055 1060 1065
 Leu Leu Lys Leu Ala Ala His Ser Val Ile Tyr Lys Cys Leu Leu
 1070 1075 1080
 Gly Pro Leu Arg Thr Ala Gln Lys Leu Leu Cys Arg Lys Leu Pro
 1085 1090 1095
 Glu Ala Thr Met Thr Ile Leu Lys Ala Ala Asp Pro Ala Leu
 1100 1105 1110

-continued

Ser Thr Asp Phe Gln Thr Ile Leu Asp
1115 1120

<210> SEQ ID NO 11
<211> LENGTH: 729
<212> TYPE: PRT
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 11

Met Arg Pro Leu Phe Gln Gln Leu Leu Val Asn His Ala Glu Cys Gln
1 5 10 15

Tyr Val Arg Leu Leu Arg Ser His Cys Arg Phe Arg Thr Ala Asn Gln
20 25 30

Gln Val Thr Asp Ala Leu Asn Thr Ser Pro Pro His Leu Met Asp Leu
35 40 45

Leu Arg Leu His Ser Ser Pro Trp Gln Val Tyr Gly Phe Leu Arg Ala
50 55 60

Cys Leu Cys Lys Val Val Ser Ala Ser Leu Trp Gly Thr Arg His Asn
65 70 75 80

Glu Arg Arg Phe Phe Lys Asn Leu Lys Lys Phe Ile Ser Leu Gly Lys
85 90 95

Tyr Gly Lys Leu Ser Leu Gln Glu Leu Met Trp Lys Met Lys Val Glu
100 105 110

Asp Cys His Trp Leu Arg Ser Ser Pro Gly Lys Asp Arg Val Pro Ala
115 120 125

Ala Glu His Arg Leu Arg Glu Arg Ile Leu Ala Thr Phe Leu Phe Trp
130 135 140

Leu Met Asp Thr Tyr Val Val Gln Leu Leu Arg Ser Phe Phe Tyr Ile
145 150 155 160

Thr Glu Ser Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr Arg Lys Ser
165 170 175

Val Trp Ser Lys Leu Gln Ser Ile Gly Val Arg Gln His Leu Glu Arg
180 185 190

Val Arg Leu Arg Glu Leu Ser Gln Glu Glu Val Arg His His Gln Asp
195 200 205

Thr Trp Leu Ala Met Pro Ile Cys Arg Leu Arg Phe Ile Pro Lys Pro
210 215 220

Asn Gly Leu Arg Pro Ile Val Asn Met Ser Tyr Ser Met Gly Thr Arg
225 230 235 240

Ala Leu Gly Arg Arg Lys Gln Ala Gln His Phe Thr Gln Arg Leu Lys
245 250 255

Thr Leu Phe Ser Met Leu Asn Tyr Glu Arg Thr Lys His Pro His Leu
260 265 270

Met Gly Ser Ser Val Leu Gly Met Asn Asp Ile Tyr Arg Thr Trp Arg
275 280 285

Ala Phe Val Leu Arg Val Arg Ala Leu Asp Gln Thr Pro Arg Met Tyr
290 295 300

Phe Val Lys Ala Asp Val Thr Gly Ala Tyr Asp Ala Ile Pro Gln Gly
305 310 315 320

Lys Leu Val Glu Val Val Ala Asn Met Ile Arg His Ser Glu Ser Thr
325 330 335

Tyr Cys Ile Arg Gln Tyr Ala Val Val Arg Arg Asp Ser Gln Gly Gln
340 345 350

-continued

Val His Lys Ser Phe Arg Arg Gln Val Thr Thr Leu Ser Asp Leu Gln
355 360 365

Pro Tyr Met Gly Gln Phe Leu Lys His Leu Gln Asp Ser Asp Ala Ser
370 375 380

Ala Leu Arg Asn Ser Val Val Ile Glu Gln Ser Ile Ser Met Asn Glu
385 390 395 400

Ser Ser Ser Ser Leu Phe Asp Phe Phe Leu His Phe Leu Arg His Ser
405 410 415

Val Val Lys Ile Gly Asp Arg Cys Tyr Thr Gln Cys Gln Gly Ile Pro
420 425 430

Gln Gly Ser Ser Leu Ser Thr Leu Leu Cys Ser Leu Cys Phe Gly Asp
435 440 445

Met Glu Asn Lys Leu Phe Ala Glu Val Gln Arg Asp Gly Leu Leu Leu
450 455 460

Arg Phe Val Asp Asp Phe Leu Leu Val Thr Pro His Leu Asp Gln Ala
465 470 475 480

Lys Thr Phe Leu Ser Thr Leu Val His Gly Val Pro Glu Tyr Gly Cys
485 490 495

Met Ile Asn Leu Gln Lys Thr Val Val Asn Phe Pro Val Glu Pro Gly
500 505 510

Thr Leu Gly Gly Ala Ala Pro Tyr Gln Leu Pro Ala His Cys Leu Phe
515 520 525

Pro Trp Cys Gly Leu Leu Leu Asp Thr Gln Thr Leu Glu Val Phe Cys
530 535 540

Asp Tyr Ser Gly Tyr Ala Gln Thr Ser Ile Lys Thr Ser Leu Thr Phe
545 550 555 560

Gln Ser Val Phe Lys Ala Gly Lys Thr Met Arg Asn Lys Leu Leu Ser
565 570 575

Val Leu Arg Leu Lys Cys His Gly Leu Phe Leu Asp Leu Gln Val Asn
580 585 590

Ser Leu Gln Thr Val Cys Ile Asn Ile Tyr Lys Ile Phe Leu Leu Gln
595 600 605

Ala Tyr Arg Phe His Ala Cys Val Ile Gln Leu Pro Phe Asp Gln Arg
610 615 620

Val Arg Lys Asn Leu Thr Phe Phe Leu Gly Ile Ile Ser Ser Gln Ala
625 630 635 640

Ser Cys Cys Tyr Ala Ile Leu Lys Val Lys Asn Pro Gly Met Thr Leu
645 650 655

Lys Ala Ser Gly Ser Phe Pro Pro Glu Ala Ala His Trp Leu Cys Tyr
660 665 670

Gln Ala Phe Leu Leu Lys Leu Ala Ala His Ser Val Ile Tyr Lys Cys
675 680 685

Leu Leu Gly Pro Leu Arg Thr Ala Gln Lys Leu Leu Cys Arg Lys Leu
690 695 700

Pro Glu Ala Thr Met Thr Ile Leu Lys Ala Ala Ala Asp Pro Ala Leu
705 710 715 720

Ser Thr Asp Phe Gln Thr Ile Leu Asp
725

-continued

<212> TYPE: PRT

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 12

```

Met Asp Thr Tyr Val Val Gln Leu Leu Arg Ser Phe Phe Tyr Ile Thr
1           5          10          15

Glu Ser Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr Arg Lys Ser Val
20          25          30

Trp Ser Lys Leu Gln Ser Ile Gly Val Arg Gln His Leu Glu Arg Val
35          40          45

Arg Leu Arg Glu Leu Ser Gln Glu Glu Val Arg His His Gln Asp Thr
50          55          60

Trp Leu Ala Met Pro Ile Cys Arg Leu Arg Phe Ile Pro Lys Pro Asn
65          70          75          80

Gly Leu Arg Pro Ile Val Asn Met Ser Tyr Ser Met Gly Thr Arg Ala
85          90          95

Leu Gly Arg Arg Lys Gln Ala Gln His Phe Thr Gln Arg Leu Lys Thr
100         105         110

Leu Phe Ser Met Leu Asn Tyr Glu Arg Thr Lys His Pro His Leu Met
115         120         125

Gly Ser Ser Val Leu Gly Met Asn Asp Ile Tyr Arg Thr Trp Arg Ala
130         135         140

Phe Val Leu Arg Val Arg Ala Leu Asp Gln Thr Pro Arg Met Tyr Phe
145         150         155         160

Val Lys Ala Asp Val Thr Gly Ala Tyr Asp Ala Ile Pro Gln Gly Lys
165         170         175

Leu Val Glu Val Val Ala Asn Met Ile Arg His Ser Glu Ser Thr Tyr
180         185         190

Cys Ile Arg Gln Tyr Ala Val Val Arg Arg Asp Ser Gln Gly Gln Val
195         200         205

His Lys Ser Phe Arg Arg Gln Val Thr Thr Leu Ser Asp Leu Gln Pro
210         215         220

Tyr Met Gly Gln Phe Leu Lys His Leu Gln Asp Ser Asp Ala Ser Ala
225         230         235         240

Leu Arg Asn Ser Val Val Ile Glu Gln Ser Ile Ser Met Asn Glu Ser
245         250         255

Ser Ser Ser Leu Phe Asp Phe Phe Leu His Phe Leu Arg His Ser Val
260         265         270

Val Lys Ile Gly Asp Arg Cys Tyr Thr Gln Cys Gln Gly Ile Pro Gln
275         280         285

Gly Ser Ser Leu Ser Thr Leu Leu Cys Ser Leu Cys Phe Gly Asp Met
290         295         300

Glu Asn Lys Leu Phe Ala Glu Val Gln Arg Asp Gly Leu Leu Leu Arg
305         310         315         320

Phe Val Asp Asp Phe Leu Leu Val Thr Pro His Leu Asp Gln Ala Lys
325         330         335

Thr Phe Leu Ser Thr Leu Val His Gly Val Pro Glu Tyr Gly Cys Met
340         345         350

Ile Asn Leu Gln Lys Thr Val Val Asn Phe Pro Val Glu Pro Gly Thr
355         360         365

Leu Gly Gly Ala Ala Pro Tyr Gln Leu Pro Ala His Cys Leu Phe Pro
370         375         380

```

-continued

Trp Cys Gly Leu Leu Leu Asp Thr Gln Thr Leu Glu Val Phe Cys Asp
385 390 395 400

Tyr Ser Gly Tyr Ala Gln Thr Ser Ile Lys Thr Ser Leu Thr Phe Gln
405 410 415

Ser Val Phe Lys Ala Gly Lys Thr Met Arg Asn Lys Leu Leu Ser Val
420 425 430

Leu Arg Leu Lys Cys His Gly Leu Phe Leu Asp Leu Gln Val Asn Ser
435 440 445

Leu Gln Thr Val Cys Ile Asn Ile Tyr Lys Ile Phe Leu Leu Gln Ala
450 455 460

Tyr Arg Phe His Ala Cys Val Ile Gln Leu Pro Phe Asp Gln Arg Val
465 470 475 480

Arg Lys Asn Leu Thr Phe Phe Leu Gly Ile Ile Ser Ser Gln Ala Ser
485 490 495

Cys Cys Tyr Ala Ile Leu Lys Val Lys Asn Pro Gly Met Thr Leu Lys
500 505 510

Ala Ser Gly Ser Phe Pro Pro Glu Ala Ala His Trp Leu Cys Tyr Gln
515 520 525

Ala Phe Leu Leu Lys Leu Ala Ala His Ser Val Ile Tyr Lys Cys Leu
530 535 540

Leu Gly Pro Leu Arg Thr Ala Gln Lys Leu Leu Cys Arg Lys Leu Pro
545 550 555 560

Glu Ala Thr Met Thr Ile Leu Lys Ala Ala Ala Asp Pro Ala Leu Ser
565 570 575

Thr Asp Phe Gln Thr Ile Leu Asp
580

<210> SEQ ID NO 13

<211> LENGTH: 1125

<212> TYPE: PRT

<213> ORGANISM: Rattus norvegicus

<400> SEQUENCE: 13

Met Pro Arg Ala Pro Arg Cys Pro Ala Val Arg Ser Leu Leu Arg Ser
1 5 10 15

Arg Tyr Arg Glu Val Trp Pro Leu Ala Thr Phe Val Arg Arg Leu Gly
20 25 30

Leu Glu Gly Ser Arg Leu Val Gln Pro Gly Asp Pro Lys Val Phe Arg
35 40 45

Thr Leu Val Ala Gln Cys Leu Val Cys Val Pro Trp Gly Ser Gln Pro
50 55 60

Pro Pro Ala Asp Leu Ser Phe His Gln Val Ser Ser Leu Lys Glu Leu
65 70 75 80

Val Ser Arg Val Val Gln Lys Leu Cys Glu Arg Gly Glu Arg Asn Val
85 90 95

Leu Ala Phe Gly Phe Ala Leu Leu Asn Gly Ala Arg Gly Pro Pro
100 105 110

Met Ala Phe Thr Thr Ser Val His Ser Tyr Leu Pro Asn Ser Val Thr
115 120 125

Glu Ser Leu Cys Val Ser Gly Ala Trp Met Leu Leu Leu Ser Arg Val
130 135 140

Gly Asp Asp Leu Leu Val Tyr Leu Leu Ser His Cys Ala Leu Tyr Leu

-continued

145	150	155	160
Leu Val Pro Pro Ser Cys Ala Tyr Gln Val Cys Gly Ser Pro Leu Tyr			
165	170	175	
Gln Ile Cys Ala Thr Thr Asp Thr Trp Ser Ser Val Pro Ala Gly Tyr			
180	185	190	
Arg Pro Thr Arg Pro Val Gly Gly Asn Phe Thr Asn Leu Gly Ser Ala			
195	200	205	
His Gln Ile Lys Asn Ser Gly His Gln Glu Ala Pro Lys Pro Gln Ala			
210	215	220	
Leu Pro Ser Arg Gly Thr Lys Arg Leu Leu Ser Leu Thr Ser Thr Asn			
225	230	235	240
Val Pro Ser Ala Lys Lys Ala Arg Phe Glu Pro Ala Leu Arg Val Asp			
245	250	255	
Lys Gly Pro His Arg Gln Val Val Pro Thr Pro Ser Gly Lys Thr Trp			
260	265	270	
Ala Pro Ser Pro Ala Ala Ser Pro Lys Val Pro Pro Ala Ala Lys Asn			
275	280	285	
Leu Ser Leu Lys Gly Lys Ala Ser Asp Pro Ser Leu Ser Gly Ser Val			
290	295	300	
Cys Cys Lys His Lys Pro Ser Ser Ser Ser Leu Leu Ser Ser Pro Pro			
305	310	315	320
Gln Asp Ala Glu Lys Leu Arg Pro Phe Thr Glu Thr Arg His Phe Leu			
325	330	335	
Tyr Ser Arg Gly Gly Gln Glu Glu Leu Asn Pro Ser Phe Leu Leu			
340	345	350	
Asn Ser Leu Pro Pro Ser Leu Thr Gly Ala Arg Arg Leu Val Glu Ile			
355	360	365	
Ile Phe Leu Gly Ser Arg Pro Arg Thr Ser Gly Pro Phe Cys Arg Thr			
370	375	380	
Arg Arg Leu Pro Arg Arg Tyr Trp Gln Met Arg Pro Leu Phe Gln Gln			
385	390	395	400
Leu Leu Met Asn His Ala Lys Cys Gln Tyr Val Arg Phe Leu Arg Ser			
405	410	415	
His Cys Arg Phe Arg Thr Ala Asn Gln Arg Val Pro Asp Ala Met Asp			
420	425	430	
Thr Ser Pro Ser His Leu Thr Ser Leu Leu Arg Leu His Ser Ser Pro			
435	440	445	
Trp Gln Val Tyr Gly Phe Leu Arg Ala Cys Leu Arg Glu Leu Val Pro			
450	455	460	
Ala Gly Leu Trp Gly Thr Arg His Asn Glu Arg Arg Phe Leu Lys Asn			
465	470	475	480
Val Lys Lys Phe Ile Ser Leu Gly Lys Tyr Ala Lys Leu Ser Leu Gln			
485	490	495	
Glu Leu Met Trp Arg Val Lys Val Glu Asp Cys His Trp Leu Arg Ser			
500	505	510	
Ser Pro Glu Lys Asp Thr Val Pro Ala Ala Glu His Arg Leu Arg Glu			
515	520	525	
Arg Ile Leu Ala Met Phe Leu Phe Trp Leu Met Asp Thr Tyr Val Val			
530	535	540	
Gln Leu Leu Arg Ser Phe Phe Tyr Ile Thr Glu Thr Thr Phe Gln Lys			
545	550	555	560

-continued

Asn	Arg	Leu	Phe	Phe	Tyr	Arg	Lys	Ser	Val	Trp	Ser	Lys	Leu	Gln	Ser
565									570						575
Ile	Gly	Ile	Arg	Gln	Gln	Leu	Glu	Arg	Val	Gln	Leu	Arg	Glu	Leu	Ser
580									585						590
Gln	Glu	Glu	Val	Lys	His	His	Gln	Asp	Thr	Trp	Leu	Ala	Met	Pro	Ile
595									600						605
Cys	Arg	Leu	Arg	Phe	Ile	Pro	Lys	Leu	Asn	Gly	Leu	Arg	Pro	Ile	Val
610									615						620
Asn	Met	Ser	Tyr	Gly	Met	Asp	Thr	Arg	Ala	Phe	Gly	Lys	Lys	Lys	Gln
625									630						640
Thr	Gln	Cys	Phe	Thr	Gln	Ser	Leu	Lys	Thr	Leu	Phe	Ser	Val	Leu	Asn
645									650						655
Tyr	Glu	Arg	Thr	Lys	His	Pro	Asn	Leu	Met	Gly	Ala	Ser	Val	Leu	Gly
660									665						670
Thr	Ser	Asp	Ser	Tyr	Arg	Ile	Trp	Arg	Thr	Phe	Val	Leu	Arg	Val	Arg
675									680						685
Ala	Leu	Asp	Gln	Thr	Pro	Arg	Met	Tyr	Phe	Val	Lys	Ala	Asp	Val	Thr
690									695						700
Gly	Ala	Tyr	Asp	Ala	Ile	Pro	Gln	Asp	Lys	Leu	Val	Glu	Ile	Val	Ala
705									710						720
Asn	Ile	Ile	Arg	Arg	Ser	Glu	Ser	Met	Tyr	Cys	Ile	Arg	Gln	Tyr	Ala
725									730						735
Val	Val	Gln	Lys	Asp	Ser	Gln	Gly	Gln	Val	His	Lys	Ser	Phe	Arg	Arg
740									745						750
Gln	Val	Ser	Thr	Leu	Ser	Asp	Leu	Gln	Pro	Tyr	Met	Gly	Gln	Phe	Thr
755									760						765
Lys	His	Leu	Gln	Asp	Ser	Asp	Ala	Ser	Ala	Leu	Arg	Asn	Ser	Val	Val
770									775						780
Ile	Glu	Gln	Ser	Ile	Ser	Met	Asn	Glu	Thr	Gly	Ser	Ser	Leu	Leu	His
785									790						800
Phe	Phe	Leu	Arg	Phe	Val	Arg	His	Ser	Val	Val	Lys	Ile	Asp	Gly	Arg
805									810						815
Phe	Tyr	Val	Gln	Cys	Gln	Gly	Ile	Pro	Gln	Gly	Ser	Ser	Leu	Ser	Thr
820									825						830
Leu	Leu	Cys	Ser	Leu	Cys	Phe	Gly	Asp	Met	Glu	Asn	Lys	Leu	Phe	Ala
835									840						845
Glu	Val	Gln	Gln	Asp	Gly	Leu	Leu	Leu	Arg	Phe	Val	Asp	Asp	Phe	Leu
850									855						860
Leu	Val	Thr	Pro	His	Leu	Ala	His	Ala	Lys	Ala	Phe	Leu	Ser	Thr	Leu
865									870						880
Val	His	Gly	Val	Pro	Glu	Tyr	Gly	Cys	Met	Ile	Asn	Leu	Gln	Lys	Thr
885									890						895
Val	Val	Asn	Phe	Pro	Val	Glu	Thr	Gly	Ala	Leu	Gly	Gly	Ala	Ala	Pro
900									905						910
His	Gln	Leu	Pro	Ala	His	Cys	Leu	Phe	Pro	Trp	Cys	Gly	Leu	Leu	Leu
915									920						925
Asp	Thr	Arg	Thr	Leu	Glu	Val	Phe	Cys	Asp	Tyr	Ser	Gly	Tyr	Gly	Arg
930									935						940
Thr	Ser	Ile	Lys	Met	Ser	Leu	Thr	Phe	Gln	Gly	Val	Ser	Arg	Ala	Gly
945									950						960

-continued

Lys	Thr	Met	Arg	Tyr	Lys	Leu	Leu	Ser	Val	Leu	Arg	Leu	Lys	Cys	His
965					970					975					
Gly	Leu	Phe	Leu	Asp	Leu	Gln	Val	Asn	Ser	Leu	Gln	Thr	Val	Cys	Ile
980					985					990					
Asn	Ile	Tyr	Lys	Ile	Phe	Leu	Leu	Gln	Ala	Tyr	Arg	Phe	His	Ala	Cys
995					1000					1005					
Val	Ile	Arg	Leu	Pro	Phe	Gly	Gln	His	Val	Arg	Lys	Asn	His	Ala	
1010					1015					1020					
Phe	Phe	Leu	Gly	Ile	Ile	Ser	Asn	Leu	Ala	Ser	Cys	Cys	Tyr	Ala	
1025					1030					1035					
Ile	Leu	Lys	Val	Lys	Asn	Pro	Gly	Val	Ser	Leu	Arg	Ala	Lys	Gly	
1040					1045					1050					
Ala	Pro	Gly	Ser	Phe	Pro	Pro	Glu	Ala	Thr	Arg	Trp	Leu	Cys	Tyr	
1055					1060					1065					
Gln	Ala	Phe	Leu	Leu	Lys	Leu	Ala	Ala	His	Ser	Val	Thr	Tyr	Lys	
1070					1075					1080					
Cys	Leu	Leu	Gly	Pro	Leu	Arg	Thr	Ala	Gln	Lys	Gln	Leu	Cys	Arg	
1085					1090					1095					
Lys	Leu	Pro	Glu	Ala	Thr	Met	Thr	Leu	Leu	Lys	Thr	Ala	Ala	Asp	
1100					1105					1110					
Pro	Ala	Leu	Ser	Thr	Asp	Phe	Gln	Thr	Ile	Leu	Asp				
1115					1120					1125					

<210> SEQ ID NO 14

<211> LENGTH: 4335

<212> TYPE: DNA

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 14

gttcccagcc	tcatctttt	cgtcggtggac	tctcagtggc	ctgggtccctg	gctgtttct	60
aaggcacaccc	ttgcatcttg	gttcccgac	gtgggaggcc	catccggcc	ttgagcacaa	120
tgaccgcgc	tcctcggtgc	cccgcggtgc	gctctctgtc	gcgcagccga	taccgggagg	180
tgtggccgt	ggcaaccttt	gtgcgggcgc	tggggcccg	ggcaggccgg	cttgtcaac	240
ccggggaccc	aaagatctac	cgcactttgg	ttgccaatg	cctagtgtgc	atgcactggg	300
gctcacagcc	tccacctgcc	gaccttcct	tccaccagg	gtcatccctg	aaagagctgg	360
tggccagggt	tgtgcagaga	ctctgcgago	gcaacgagag	aaacgtgtcg	gttttggtct	420
tttgagctgt	taacgaggcc	agaggcgccc	ctccccatggc	cttcactagt	agcgtgcgt	480
gtctacttgcc	caacactgtt	atttagaccc	tgcgtgtca	tggtgtcatgg	atgctactgt	540
tgagccgagt	gggcgacgac	ctgtcggtct	acctgctggc	acactgtgt	ctttatcttc	600
ttgtgcccc	cagctgtgcc	taccagggt	gtgggtctcc	cctgtaccaa	atttgtgcca	660
ccacggatat	ctggccctct	gtgtccgcta	gttacaggcc	cacccgaccc	gtgggcagga	720
atttcaactaa	ccttaggttc	ttacaacaga	tcaagagcg	tagtcgccc	gaagcaccga	780
aaaaacctggc	cttgcacatct	cgaggatcaa	agaggcatct	gagtctcacc	agtacaagt	840
tgccttcagc	taagaaggcc	agatgctatc	ctgtcccggag	agtggaggag	ggaccccaca	900
ggcagggtgt	accaacccca	tcaggcaat	catgggtgc	aagtccctgt	cggtcccccg	960
agggtgcctac	tgcagagaaa	gatttgtctt	ctaaaggaaa	ggtgtctgac	ctgagtctct	1020
ctgggtcggt	gtgctgtaaa	cacaagccca	gtcccacatc	tctgctgtca	ccaccccgcc	1080

-continued

aaaatgcctt	tca	gctcagg	ccatttattt	agaccagaca	tttccttac	tccagggag	1140
atggccaaga	gcgtctaaac	ccctcatcc	tactcagcaa	cctccagcct	aacttgactg		1200
gggccaggag	actgggtggag	atcatcttcc	tgggctcaag	gccttaggaca	tcaggacac		1260
tctgcaggac	acacccgtcta	tcgegtcgat	actggcagat	gcccgcctg	ttccaacagc		1320
tgcgttgtaa	ccatgcagag	tgccaatatg	tcagactcc	caggtcacat	tgcaggttcc		1380
gaacagcaaa	ccaacagggt	acagatgcct	tgaacaccag	cccacccgcac	ctcatggatt		1440
tgcgttgttt	gcacagcagt	ccctggcagg	tatatggttt	tcttcgggccc	tgtctctgca		1500
agggtgggtgc	tgcgttgtc	tgggtgttca	ggcacaatga	gcccgcctc	tttaagaact		1560
taaagaagtt	catctcggtt	gggaaatacg	gcaagctatc	actgcaggaa	ctgtatgttgc		1620
agatgaaagt	agaggattgc	cactggctcc	gcagcagccc	ggggaaaggac	cgtgtccccg		1680
ctgcagagca	ccgtctgagg	gagaggatcc	tggctacgtt	cctgttctgg	ctgtatggaca		1740
catacgttgtt	acagctgctt	aggtcatttc	tttacatcac	agagagcaca	ttccagaaga		1800
acaggctctt	cttctaccgt	aagagtgtgt	ggagcaagct	gcagagcatt	ggagtccaggc		1860
aacaccttga	gagagtgcgg	ctacgggago	tgtcacaaga	ggaggtcagg	catcaccagg		1920
acacacctggct	agccatgccc	atctgcagac	tgcgcttcat	ccccaaagccc	aacggcctgc		1980
ggcccatgtt	gaacatgagt	tatagcatgg	gtaccagac	tttgggcaga	aggaagcagg		2040
cccgacattt	cacccagcgt	ctcaagactc	tcttcagcat	gctcaactat	gagcggacaa		2100
aacatcctca	ccttatgggg	tcttcgttac	tgggtatgaa	tgacatctac	aggacctggc		2160
gggcctttgt	gctgcgtgt	cgtgtctgg	accagacacc	caggatgtac	tttggtaagg		2220
cagatgttgc	cggggcctat	gatgccatcc	cccaggtaa	gctggggag	gttggccca		2280
atatgatcag	gcactcggag	agcacgtact	gtatccgcac	gtatgcagtg	gtccggagag		2340
atagccaagg	ccaagtccac	aagtccctta	ggagacaggt	caccaccctc	tctgacccctc		2400
agccatacat	ggggccatcc	cttaagacatc	tgcaggattc	agatgcagg	gcactgagga		2460
actccgttgt	catcgaggcag	agcatctcta	tgaatgagag	cagcaggcagc	ctgtttgact		2520
tcttcctgc	cttcctgcgt	cacagtgtcg	taaaatgg	tgacaggtgc	tatacgcagt		2580
gccaggccat	cccccagggc	tccagctat	ccaccctgct	ctgcagtcgt	tgtttcgag		2640
acatggagaa	caagctgttt	gctgggtgc	agcgggatgg	gttgccttta	cgttttttg		2700
atgactttct	gttgggtgacg	cctcaacttgg	accaagcaaa	aaccccttc	agcacccctgg		2760
tccatggcgt	tcctgtgtat	gggtgtcatga	taaaatgc	gaagacagtg	gtgaacttcc		2820
ctgtggagcc	tggtaaccctg	gggtgggtgc	ctccatacc	gtgcctgct	ca		2880
tccctgggt	ttggcttgcgt	ctggacactc	agactttgg	gttggttgt	gactactcag		2940
gttatgccc	gacctaatt	aagacgagcc	tcaccccttca	gagtgtctc	aaagctggga		3000
agaccatgcg	gaacaagctc	ctgtcggtct	tgcgggtgaa	gtgtcacgg	ctatttctag		3060
acttgcaggt	gaacagccctc	cagacagtct	gcatcaat	atacaagatc	ttcctgcctc		3120
aggcctacag	gttccatgca	tgtgtgattc	agcttccctt	tgaccagcgt	gttaggaaga		3180
acotcacatt	ctttctgggc	atcatctca	gccaagcatc	ctgtgtat	gtatccctga		3240
aggtaaagaa	tccaggaatg	acactaaagg	cctctggctc	ctttccct	gaagccgcac		3300
attggctctg	ctaccaggcc	ttccctgc	agctggctc	tcatttgc	atctacaaat		3360

-continued

gtctcctggg acctctgagg acagccaaa aactgctgtg ccggaagctc ccagaggcga	3420
caatgaccat ccttaaagct gcagctgacc cagccctaag cacagactt cagaccattt	3480
tggactaacc ctgtctcctt ccgctagatg aacatggca ttgttagctc agcactctg	3540
gatccacgta acaagaggga ctggtcagtt gtgaggctag gtcatectcc aaacctctgt	3600
gtcatgggtg gtatgggaga ttgtcccagt gccttgttcc ctgtaacagg cttgattct	3660
ttcttgatgc cctcaggag gcagatccta tcccttttag tggcaggat ccactagcac	3720
cagcacatga ggagtgcacc cagtgcacat gggcaetggg acagtggaca ggtgtgagat	3780
tcctgggcc tggagtctt tcacacctaa ccatggagcc tgtcccagta catcagatg	3840
cctcgagat gaaaaaggac atcgagccag tgacctaata tacagcctga atatactctg	3900
aattcatgtg actgccttag ctacttctct actgctgtg agtaaaacac caagccaact	3960
tataaaaagca ggatttcct actggagcag cagctgagag ttacatctt gatccataag	4020
cacaaaagca caagacagag agagagagag agagagagag agagagagag agagagagag	4080
agagagagag agagagtca gtcagtcagta taacaaataa ctaagaaagg tgaagggtga	4140
tgaagtccac agggatcacg cttagggatgt tccatgcctt ctctgaagct aagattcctt	4200
ggcagcgctt gacagtaacc atagtggta cctactgaga tcactataaa gataaaatag	4260
ggggaaagcgt atttgtactg aactggaaaa acatacaaataa aaagagtaaa tcatggaaaa	4320
aaaaaaaaaaaa aaaaa	4335

<210> SEQ ID NO 15
<211> LENGTH: 3826
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 15

gttcccagcc tcatttttt cgtcgtggac tctcagtggc ctgggtctg gctgtttct	60
aagcacaccc ttgcattttt gttcccgac gttggaggcc cttccggcc ttgagcacaa	120
tgacccgccc tcctcggtgc cccgggtgc gctctctgt gtcggccga taccgggagg	180
tgtggccgtt ggcaaccttt gtgcggccgc tggggcccgaa gggcaggccg cttgtcaac	240
ccggggaccc gaagatctac cgcactttgg ttgcccataatg cctagtgtgc atgcactgg	300
gctcacagcc tccacctgcc gaccccttct tccaccaggat gtcatccctg aaagagctgg	360
tggccagggt tgcgcagaga ctctgcgcgc gcaacgcagaa aaacgtgtgc gctttggct	420
ttgagctgtc taacggggcc agaggccggc cttccatggc cttcactagt agcgtgcgt	480
gctacttgcc caacactgtt attgagaccc tgcgtgtca gttgtcatgg atgcactgt	540
tgagccagggt gggcgacgc ctgcgtgtctt acctgcgtggc acactgtgtc ctttatctt	600
tggtgcccc cagctgtgcc taccaggggaa gatggccaaag agcgtctaaa cccctcatc	660
ctactcagca acctccagcc taacttgact gggccaggaa gactgggtggaa gatcatctt	720
ctgggctcaa ggcctaggac atcaggacca ctctgcaggaa cacaccgtct atcgcgtcga	780
tactggcaga tgcggccctt gttccaaacag ctgcgtgtca accatgcaga gtgccaat	840
gtcagactcc tcaggcaca ttgcagggtt cgaacagcaa accaacaggt gacagatgcc	900
ttgaacacca gcccacccgca cctcatggat ttgcacccgc tgcacacgcag tccctggcag	960
gtatatggtt ttcttcgggc ctgtctctgc aagggtgggt ctgcgtgtct ctggggtaacc	1020

-continued

aggcacaatg	agcgccgctt	ctttaagaac	ttaaagaagt	tcatctcggtt	ggggaaatac	1080
ggcaagctat	cactgcagga	actgatgtgg	aagatgaaag	tagaggatg	ccactggctc	1140
cgcagcagcc	cggggaagga	ccgtgtcccc	gctgcagagc	accgtctgag	ggagaggatc	1200
ctggctacgt	tcctgttctg	gctgtatggac	acatacgttg	tacagetgct	taggtcattc	1260
ttttacatca	cagagagcac	attccagaag	aacaggtct	tcttctaccg	taagagtgtg	1320
tggagcaagc	tgcagagcat	tggagtcagg	caacacctt	agagagtgcg	gctacggag	1380
ctgtcacaag	aggaggtcag	gcatcaccag	gacacctggc	tagccatgcc	catctgcaga	1440
ctgcgcctca	tccccaagcc	caacggcctg	cggcccattt	tgaacatgag	ttatagcatg	1500
ggtaccagag	ctttggcag	aaggaagcag	gcccagcatt	tcacccagcg	tctcaagact	1560
ctttcagca	tgctcaacta	tgagcggaca	aaacatccctc	accttatggg	gtcttctgta	1620
ctgggtatga	atgacatcta	caggacctgg	cgggcctt	tgctgcgtgt	gcgtgctctg	1680
gaccagacac	ccaggatgta	ctttgttaag	gcagatgtga	ccggggccct	tgtgccatc	1740
ccccagggtta	agctggtgg	ggttgttgc	aatatgtca	ggcactcgga	gagcacgtac	1800
tgtatccgcc	agtatgcagt	ggtccggaga	gatagccaag	gccaagtcca	caagtcctt	1860
aggagacagg	tcaccaccc	ctctgaccc	cagccataca	tgggcccattt	ccttaagcat	1920
ctgcaggatt	cagatgccag	tgcactgagg	aactccgtt	tcatcgagca	gagcatctct	1980
atgaatgaga	gcagcagcag	cctgtttgac	ttcttcctgc	acttcctgcg	tcacagtgtc	2040
gtaaagattt	gtgacaggtg	ctatacgcag	tgccaggggca	tcccccagg	ctccagccta	2100
tccaccctgc	tctgcagttt	gtgtttcgg	gacatggaga	acaagcttt	tgctgaggtg	2160
cagcgggatg	ggttgctttt	acgtttttgtt	gatgactttc	tgttgttgc	gcctcactt	2220
gaccaagcaa	aaacccctt	cagccccctg	gtccatggcg	tccctgagta	tgggtgcatt	2280
ataaaacttgc	agaagacagt	ggtgaactt	cctgtggagc	ctggtaacct	gggtggtgca	2340
gctccatacc	agctgcctgc	tcactgcctg	ttttccctgtt	gtggcttgc	gctggacact	2400
cagactttgg	aggtgttctg	tgactacta	ggttatgcc	agacactaat	taagacgagc	2460
ctcaccccttcc	agagtgtctt	caaagctgg	aagaccatgc	ggaacaagct	cctgtcggtc	2520
ttgcgggttga	agtgtcacgg	tctatttcta	gacttgcagg	tgaacagcct	ccagacagtc	2580
tgcataata	tatacaagat	cttccctgtt	caggcctaca	ggttccatgc	atgtgtgatt	2640
cagcttccctt	ttgaccagcg	tgttaggaag	aacctcacat	tcttctggg	catcatctcc	2700
agccaaagcat	cctgctgctt	tgctatccctg	aaggtaaga	atccaggaat	gacactaaag	2760
gcctctggct	ccttccctcc	tgaagccgca	cattggctct	gttaccaggc	cttcctgctc	2820
aagctggctt	ctcattctgt	catctacaaa	tgtctctgg	gacactctgag	gacagccaa	2880
aaactgctgt	gccggaagct	cccaagggcg	acaatgacca	tccttaaagc	tgcagctgac	2940
ccagccctaa	gcacagactt	tcagaccatt	ttggactaaac	cctgtctctt	tccgcttagat	3000
gaacatgggc	atttgcgcct	cagcactctt	ggatccacgt	cacaagaggg	actggctagt	3060
tgtgaggctt	ggtcaccc	caaaccctctg	tgtcatgggt	ggtatggag	attgtccct	3120
tgccttgc	cctgttaacag	gctgtattt	tttccctgtat	ccctcaggga	ggcagatctt	3180
atccctttta	gtggcaggga	tccactagca	ccagcacatg	aggagtgcac	ccagtgacaca	3240
tgggcactgg	gacagtggac	agggtgtgaga	ttcctgggccc	ctggagtc	ttcacacacta	3300

-continued

accatggagc ctgtcccagt acatcagagt gcctcgaggaa tgaaaaaggaa catcgaggca	3360
gtgacctaaa ttacagcctg aataactctt gaattcatgt gactgectta gctacttctc	3420
tactgctgtg tagtaaaaca ccaagccaaac ttataaaaaggc aggattttcc tactggagca	3480
gcagctgaga gtttacatct tggatccataa gcacaaaaggc acaagacaga gagagagaga	3540
gagagagaga gagagagaga gagagagaga gagagagaga gagagagtca gtcagtca	3600
ctaacaataa actaagaaag gtgaagggtt atgaagtcca caggatcac gctaggatg	3660
ttccatgcct tctctgaagc taagattctt tggcagcggt tgacagtaac catagtgggt	3720
acactacttagt atcaactataa agataaaaata gggggaaaggc tattttactt gaactggaaa	3780
aacatacaaaa taaagagtaa atcatggaaa aaaaaaaaaaaa aaaaaaa	3826

<210> SEQ ID NO 16

<211> LENGTH: 3633

<212> TYPE: DNA

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 16

gttccccagcc tcatactttt cgtcgtggac tctcagtggc ctgggtccctg gctgtttct	60
aaggcacaccc ttgcatacttg gttccccgac gtggggaggcc catcccccggcc ttgagcacaa	120
tgcacccgcgc tcctcggtgc cccgcgggtgc gctctctgtc ggcgcagccgc taccgggagg	180
tgtggccgcgt ggcaacacctt gtgcggcgcc tggggcccgaa gggcaggccgg cttgtgcac	240
ccggggacccc gaagatctac cgcactttgg ttggccaaatg cctagtggtc atgcactgg	300
gctcacagcc tccacatgcc gacctttctt tccaccagggt gtcatccctg aaagagctgg	360
tggccagggt tgcgcagaga ctctgcgcgc gcaacgcgagaa aacatgtgtc gctttggct	420
tttagctgtct taacgaggcc agaggcgggc ctcccatggc cttcaactgt agcgtgcgt	480
gtacttgtcc caacactgtt attgagaccc tgcgtgtcaag tgggtgcattt atgcactgt	540
tgcgcgcgtt gggcgacgc acgtgtgtt acactgtgtc ctttatcttc	600
tggtgcccccc cagctgtgcc taccaggggaa gatggccaaag agcgtctaaa ccccttattc	660
ctactcagca acctccagcc taacttgact ggggcaggaa gactgggtggaa gatcatctt	720
ctgggctcaa ggcctaggac atcaggacca ctctgcaggaa cacaccgtct atcgcgtcga	780
tactggcaga tgcggccctt gttccaaacag ctgctgggtga accatgcaga gtgcacat	840
gtcagactcc tcaggtcaca ttgcagggtt cgaacagcaa accaacaggt gacagatgcc	900
ttgaacacca gcccacccgc cctcatggat ttgcacccgc tgcacagcag tccctggcag	960
ggaaggaccc tgcccccgct gcagagcacc gtctgaggaa gaggatccctg gctacgttcc	1020
tgtttctggct gatggacaca tacgtgttac agctgttttgc tgcattttt tacatcacag	1080
agagcacatt ccagaagaac aggctttctt tctaccgtaa gagttgtgttgg agcaagctgc	1140
agagcattgg agtcaggccaa caccttggaa gagttgtgttgg acggggagctg tcacaagagg	1200
aggctcaggccaa tcaccaggac acctggcttag ccatggccat ctgcagactg cgcttcatcc	1260
ccaaaggccaa cggcctgcgg cccattgtga acatgagtttta tagcatgggtt accagagctt	1320
tggcgcaggaa gaaaggccaa cagcatgttca cccagcgctt caagactctc ttctgtgtc	1380
tcaactatgtt gcccacaaaa catcctcacc ttatgggttc ttctgtgttgg ggtatgtat	1440
acatctacag gacctggccgg gcctttgtgc tgcgtgtgc tgctctggac cagacacccaa	1500

-continued

ggatgtactt	tgttaaggca	gatgtgacgg	gggcctatga	tgccatcccc	caggtaagc	1560
tggtgtgggt	tgttgccaat	atgatcaggo	actcgagag	cacgtactgt	atccgccagt	1620
atgcagtgg	ccggagagat	agccaaggcc	aagtccaca	gtccctttagg	agacaggctca	1680
ccaccctctc	tgacctccag	ccatacatgg	gccagttct	taagcatctg	caggatttag	1740
atgccagtgc	actgaggaac	tccgttgtca	tcgagcagag	catctctatg	aatgagagca	1800
gcagcagcct	gtttgacttc	ttctctgact	tcctgcgtca	cagtgtcgta	aagattggtg	1860
acaggtgcta	tacgcagtgc	cagggcattcc	cccagggtctc	cagcctatcc	accctgctct	1920
cgagtctgt	tttcggagac	atggagaaca	agctgtttgc	tgaggtgcag	cgggatgggt	1980
tgttttacg	ttttgttgc	gactttctgt	tggtgacgccc	tcacttggac	caagcaaaaa	2040
ccttcctcag	caccctggtc	catggcgttc	ctgagtatgg	gtgcataata	aacttgcaga	2100
agacagtgg	gaacttccct	gtggggcctg	gtaccctggg	tggtgacgct	ccataccagc	2160
tgcctgtca	ctgcctgttt	ccctgggtgt	gcttgctgtc	ggacactcag	actttggagg	2220
tgttctgtga	ctactcaggt	tatgcccaga	cctcaattaa	gacgagccctc	accttccaga	2280
gtgtcttcaa	agctggaaag	accatgcgga	acaagctct	gtcggctctg	cggttgaagt	2340
gtcacggct	atttctagac	ttgcaggtga	acagcctcca	gacagtctgc	atcaatata	2400
acaagatctt	cctgcattcag	gcctacaggt	tccatgcatg	tgtgattcag	cttcccttg	2460
accagcgtgt	taggaagaac	ctcacattct	ttctgggcat	catctccagc	caagcatcct	2520
gctgctatgc	tatcctgaag	gtcaagaatc	caggaatgac	actaaaggcc	tctggctcct	2580
ttccctcctga	agccgcacat	tggctctgt	accaggcctt	cctgctcaag	ctggctgctc	2640
attctgtcat	ctacaaatgt	ctcctggac	ctctgaggac	agcccaaaaa	ctgctgtgcc	2700
ggaagctccc	agaggcgaca	atgaccatcc	ttaaagctgc	agctgaccca	gccctaagca	2760
cagactttca	gaccattttg	gactaaccct	gtctccttcc	gttagatgaa	catggcatt	2820
gtagccttag	cactcctgga	tccacgtcac	aagagggact	ggtcagttgt	gaggctaggt	2880
catcctccaa	acctctgtgt	catgggtgg	atgggagatt	gtcccagtg	cttggcttct	2940
gtacacaggct	tgatttcttt	cctgtatgcc	tcagggaggc	agatcctatc	ccttttagtg	3000
gcagggatcc	actagcacca	gcacatgagg	agtgcaccca	gtgcacatgg	gcactggac	3060
agtggacagg	tgtgagattc	ctggggcctg	gagtctttc	acacctaacc	atggagcctg	3120
tcccagtaca	tcaagtgcc	tcggagatga	aaaaggacat	cgagccagtg	acctaaatta	3180
cagcctgaat	atactctgaa	ttcatgtgac	tgccttagt	acttctctac	tgcgtgttag	3240
taaaacacca	agccaaactta	taaaagcagg	atttcctac	tggagcagca	gctgagagtt	3300
tacatcttga	tccataagca	caaaaacaca	agacagag	agagagagag	agagagagag	3360
agagagagag	agagagagag	agagagagag	agagtca	agtca	acaaataact	3420
aagaaagggt	aagggtgt	aagtccacag	ggatcacgt	agggatgtc	catgcctct	3480
ctgaagctaa	gattcctgg	cagcgttga	cagtaaccat	agtgggtacc	tactgagatc	3540
actataaaaga	taaaataggg	ggaagcgtat	ttgtactgaa	ctggaaaaac	atacaataaa	3600
agagtaaatc	atggaaaaaa	aaaaaaaaaa	aaa			3633

-continued

```
<212> TYPE: DNA
<213> ORGANISM: Rattus norvegicus
<400> SEQUENCE: 17
atgcggcccg ctcctcgtt cccggccgtg cgctctctac tgcgcaagccg atatcgagg 60
gtgtggccgc tggcgacctt tggcgccgc ctggggcttg agggcagtcg gcttgtcaa 120
cccgccccacc cgaaggcttt ccgcacgtt gttgcccagt gcctagtgtg cgtccctgg 180
ggctcacagc cgccacactgc tgaccttcc ttccaccagg tgcataccct gaaagagctg 240
gtgtccaggg ttgtcagaa actttgcag cgcgggtgaga ggaatgtgt ggctttggc 300
tttgcaactgc ttaacggggc cagaggtggg cttccatgg cttcacgac cagcgtgcat 360
agctacttgc ccaactcggt tactgagtc ctgtgtgtca gtgggtcatg gatgtactg 420
tttagccgag tggcgacgaa cctgtggc tacctgtgt cgcactgtgc gctctacctg 480
ctgggtgcccc ccagctgtgc ctaccaggta tgccgggtcac ccctgtacca aatttgc 540
accacggata cctggccctc tggcccgctt gggtacaggc ccactcgacc cgtgggggg 600
aatttcaacta accttgggtc cgcacaccag ataaaaaca gtggtaccca ggaagcacca 660
aaaccccccagg cttggccatc acggaggta cggaggcttc tggatctcac cagtacaaac 720
gtgccttcag ctaagaaggc cagggttcaa cttggccatc gagtggataa gggacccac 780
aggcagggtgg taccaccccc atcaggcaaa acatgggcgc caagtccgtc tgcgtcccc 840
aagggtgcctc ctgcagcgaa aaacttgtct ttgaaaggaa aggcatctga cccgagtctc 900
tctgggtcg tggctgtaa acacaagccc agtcctcgat ccctgtgtc atcaccaccc 960
caagatgtcg aaaagtcag gccattcaact gagaccacat atttcctta ctccagggg 1020
ggtgtggcaag aggagctaaa tccctcatc ctactcaaca gcctccggcc tagttgacc 1080
ggggccagga gactgggtgg gatcatctt ctgggctcaa ggcttaggac atcaggacca 1140
ttctgcagga cccggccctt gccccgtcg tactggcaga tggacccctt attccagcag 1200
ctgctcatga accacgcaaa gtgcaatat gtcagattcc tccggcgca ctgcagattt 1260
cgaacagcaa accagegggt gcccgtcg atggacacca gcccacccca cctcagcagt 1320
ttgctccgggt tacacacgag cccctggcag gtatacggt ttcttcgggc ctgcctccgc 1380
gagctgggtgc ctggcggtct ctggggcacc aggcacaatg agcggcgctt cttaaagaac 1440
gtgaagaagt tcatactcggtt ggggaagtac gccaagctat ccctgcagga actgtatgtgg 1500
agggtgaaag tggaggactg ccactggctc cgcagcagcc cagagaagga cactgtccct 1560
gccgcagacgc accgtctgag ggagaggatc ttgcgcatgt tctgttctg gctaattggac 1620
acatatgtgg tacagctgtc gaggtcattt ttctacatca cagagaccac gttccagaag 1680
aacccgcattt ttttctaccg taagatgtgt tggagcaagc tgcagacat tggaaatcagg 1740
caacagcttgc agagagttca gctacgggaa ctgtcacaag aggaggctaa gcatcaccc 1800
gacacttggc tggccatgcc tatctgcaga ttgcgcattca tcccaagct caatggctc 1860
cggcccatgg tgaacatgag ttatggcatg gacaccagag cttttggcaa aaagaagcag 1920
acccagtgtt tcactcagag tctcaagact ttgttcagcg tgctcaacta cgagcggacc 1980
aaacatctca accttatggg tgcttcagta ctgggtacga gtgacagctt caggatctgg 2040
cgaccccttcg tgctcgtgt gctgtctcg gaccagacac ccaggatgtt ctttggtaag 2100
cgacatgtga cagggcccta tgcgtccatc cccaggaca agctcggtgg aatttgc 2160
```

-continued

aatataatca	ggcgctcaga	gagcatgtac	tgtatccgcc	agtatgcagt	ggttcagaaa	2220
gatagccaag	gcacaaggcca	caagtccttc	aggagacagg	tctccaccct	ctctgacactc	2280
cagccatacata	tggggcagtt	caccaagcat	ctgcaggact	cagatgccag	tgcactgagg	2340
aactctgttgc	tcatecgagca	gagcatctcc	atgaatgaga	ctggcagtag	cctgetccac	2400
ttcttcctgc	gttttgtccg	tcacagtgtc	gtgaagatcg	atggcaggtt	ctatgtgcaa	2460
tgccaggggca	tccccccaggg	ctccagccctg	tccaccctgc	tctgcagttct	gtgtttcgga	2520
gacatggaga	acaagctgtt	tgccgagggtg	cagcaggacg	gcttgcgtttt	acgttttgtc	2580
gtgacttttc	tgttgtgtac	acctcacctg	gcccattgcaa	aaggcccttct	cagcacccctg	2640
gtccatggcg	tgcccgagta	tggctgcattg	ataaaactgc	agaagacagt	ggtgaacctc	2700
cctgtggaga	ccggcccccct	gggaggtgca	gccccgcacc	agctgcctgc	tcactgcctg	2760
tttccctgg	gtggcttact	gctggacact	cggaacttttg	aagtattctg	tgactactca	2820
ggttacggac	ggacctcaat	taagatgago	ctcaccttcc	agggtgtctc	cagggccggg	2880
aagaccatgc	ggtacaagct	cttgtcagtc	ttggcgttga	agtgtcatgg	tctgtttcta	2940
gacttgcagg	tgaacagcct	gcagacagtc	tgcattcaata	tatacaagat	cttcctgctt	3000
caggcctaca	ggttccatgc	atgtgtgatt	cggtttccct	ttggccagca	tgttaggaag	3060
aaccatgcat	tctttctggg	catcatctcc	aacctagcat	cctgctgcta	cggccatctg	3120
aaggtaaga	atccaggagt	gtcactaagg	gccaagggtg	cccctggctc	ctttccgccc	3180
gaggccacac	gttggctctg	ctaccaagcc	ttcctgctca	agctggctgc	tcattctgtc	3240
acctacaagt	gtctccctggg	acctcttagg	acagccaa	aacagctgtg	ccggaagctc	3300
ccagaggcaa	caatgaccct	ccttaagact	gcagctgacc	cagccctaag	cacagattt	3360
cagaccattt	tggactaa					3378

<210> SEQ ID NO 18

<211> LENGTH: 1132

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 18

Met	Pro	Arg	Ala	Pro	Arg	Cys	Arg	Ala	Val	Arg	Ser	Leu	Leu	Arg	Ser
1									5		10			15	

His	Tyr	Arg	Glu	Val	Leu	Pro	Leu	Ala	Thr	Phe	Val	Arg	Arg	Leu	Gly
									20		25			30	

Pro	Gln	Gly	Trp	Arg	Leu	Val	Gln	Arg	Gly	Asp	Pro	Ala	Ala	Phe	Arg
									35		40			45	

Ala	Leu	Val	Ala	Gln	Cys	Leu	Val	Cys	Val	Pro	Trp	Asp	Ala	Arg	Pro
									50		55			60	

Pro	Pro	Ala	Ala	Pro	Ser	Phe	Arg	Gln	Val	Ser	Cys	Leu	Lys	Glu	Leu
								65		70		75		80	

Val	Ala	Arg	Val	Leu	Gln	Arg	Leu	Cys	Glu	Arg	Gly	Ala	Lys	Asn	Val
								85		90			95		

Leu	Ala	Phe	Gly	Phe	Ala	Leu	Leu	Asp	Gly	Ala	Arg	Gly	Pro	Pro	
								100		105			110		

Glu	Ala	Phe	Thr	Thr	Ser	Val	Arg	Ser	Tyr	Leu	Pro	Asn	Thr	Val	Thr
								115		120			125		

-continued

Asp Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu	Leu Leu Arg Arg Val		
130	135	140	
Gly Asp Asp Val Leu Val His Leu	Leu Ala Arg Cys Ala	Leu Phe Val	
145	150	155	160
Leu Val Ala Pro Ser Cys Ala Tyr Gln Val	Cys Gly Pro Pro	Leu Tyr	
165	170	175	
Gln Leu Gly Ala Ala Thr Gln Ala Arg Pro	Pro Pro His Ala Ser Gly		
180	185	190	
Pro Arg Arg Leu Gly Cys Glu Arg Ala Trp Asn	His Ser Val Arg		
195	200	205	
Glu Ala Gly Val Pro Leu Gly Leu Pro Ala Pro	Gly Ala Arg Arg Arg		
210	215	220	
Gly Gly Ser Ala Ser Arg Ser Leu Pro Leu Pro	Lys Arg Pro Arg Arg		
225	230	235	240
Gly Ala Ala Pro Glu Pro Glu Arg Thr Pro	Val Gly Gln Gly Ser Trp		
245	250	255	
Ala His Pro Gly Arg Thr Arg Gly Pro Ser Asp	Arg Gly Phe Cys Val		
260	265	270	
Val Ser Pro Ala Arg Pro Ala Glu Glu Ala Thr	Ser Leu Glu Gly Ala		
275	280	285	
Leu Ser Gly Thr Arg His Ser His Pro Ser Val	Gly Arg Gln His His		
290	295	300	
Ala Gly Pro Pro Ser Thr Ser Arg Pro Pro Arg	Pro Trp Asp Thr Pro		
305	310	315	320
Cys Pro Pro Val Tyr Ala Glu Thr Lys His	Phe Leu Tyr Ser Ser Gly		
325	330	335	
Asp Lys Glu Gln Leu Arg Pro Ser Phe Leu	Leu Ser Ser Leu Arg Pro		
340	345	350	
Ser Leu Thr Gly Ala Arg Arg Leu Val Glu Thr	Ile Phe Leu Gly Ser		
355	360	365	
Arg Pro Trp Met Pro Gly Thr Pro Arg Arg	Leu Pro Arg Leu Pro Gln		
370	375	380	
Arg Tyr Trp Gln Met Arg Pro Leu Phe Leu	Glu Leu Leu Gly Asn His		
385	390	395	400
Ala Gln Cys Pro Tyr Gly Val Leu Leu Lys	Thr His Cys Pro Leu Arg		
405	410	415	
Ala Ala Val Thr Pro Ala Ala Gly Val Cys	Ala Arg Glu Lys Pro Gln		
420	425	430	
Gly Ser Val Ala Ala Pro Glu Glu Asp Thr	Asp Pro Arg Arg Leu		
435	440	445	
Val Gln Leu Leu Arg Gln His Ser Ser Pro	Trp Gln Val Tyr Gly Phe		
450	455	460	
Val Arg Ala Cys Leu Arg Arg Leu Val Pro	Pro Gly Leu Trp Gly Ser		
465	470	475	480
Arg His Asn Glu Arg Arg Phe Leu Arg Asn	Thr Lys Lys Phe Ile Ser		
485	490	495	
Leu Gly Lys His Ala Lys Leu Ser Leu Gln	Glu Leu Thr Trp Lys Met		
500	505	510	
Ser Val Arg Asp Cys Ala Trp Leu Arg Arg	Ser Pro Gly Val Gly Ser		
515	520	525	

-continued

Val	Pro	Ala	Ala	Glu	His	Arg	Leu	Arg	Glu	Glu	Ile	Leu	Ala	Lys	Phe
530				535			540								
Leu	His	Trp	Leu	Met	Ser	Val	Tyr	Val	Val	Glu	Leu	Leu	Arg	Ser	Phe
545				550			555			560					
Phe	Tyr	Val	Thr	Glu	Thr	Thr	Phe	Gln	Lys	Asn	Arg	Leu	Phe	Phe	Tyr
				565			570		585		590		575		
Arg	Lys	Ser	Val	Trp	Ser	Lys	Leu	Gln	Ser	Ile	Gly	Ile	Arg	Gln	His
							595		590						
Leu	Lys	Arg	Val	Gln	Leu	Arg	Glu	Leu	Ser	Glu	Ala	Glu	Val	Arg	Gln
				595		600		605							
His	Gln	Glu	Ala	Arg	Pro	Ala	Leu	Leu	Thr	Ser	Arg	Leu	Arg	Phe	Ile
				610		615		620							
Pro	Lys	Pro	Asp	Gly	Leu	Arg	Pro	Ile	Val	Asn	Met	Asp	Tyr	Val	Val
				625		630		635		640					
Gly	Ala	Arg	Thr	Phe	Arg	Arg	Glu	Lys	Arg	Ala	Glu	Arg	Leu	Thr	Ser
				645		650		655		660		665		670	
Arg	Val	Lys	Ala	Leu	Phe	Ser	Val	Leu	Asn	Tyr	Glu	Arg	Ala	Arg	Arg
				660		665		670							
Pro	Gly	Leu	Leu	Gly	Ala	Ser	Val	Leu	Gly	Leu	Asp	Asp	Ile	His	Arg
				675		680		685							
Ala	Trp	Arg	Thr	Phe	Val	Leu	Arg	Val	Arg	Ala	Gln	Asp	Pro	Pro	Pro
				690		695		700							
Glu	Leu	Tyr	Phe	Val	Lys	Val	Asp	Val	Thr	Gly	Ala	Tyr	Asp	Thr	Ile
				705		710		715		720					
Pro	Gln	Asp	Arg	Leu	Thr	Glu	Val	Ile	Ala	Ser	Ile	Ile	Lys	Pro	Gln
				725		730		735							
Asn	Thr	Tyr	Cys	Val	Arg	Arg	Tyr	Ala	Val	Val	Gln	Lys	Ala	Ala	His
				740		745		750							
Gly	His	Val	Arg	Lys	Ala	Phe	Lys	Ser	His	Val	Ser	Thr	Leu	Thr	Asp
				755		760		765							
Leu	Gln	Pro	Tyr	Met	Arg	Gln	Phe	Val	Ala	His	Leu	Gln	Glu	Thr	Ser
				770		775		780							
Pro	Leu	Arg	Asp	Ala	Val	Ile	Ile	Glu	Gln	Ser	Ser	Ser	Leu	Asn	Glu
				785		790		795		800					
Ala	Ser	Ser	Gly	Leu	Phe	Asp	Val	Phe	Leu	Arg	Phe	Val	Cys	Arg	His
				805		810		815							
Ala	Val	Arg	Ile	Arg	Gly	Lys	Ser	Tyr	Val	Gln	Cys	Gln	Gly	Ile	Pro
				820		825		830							
Gln	Gly	Ser	Ile	Leu	Ser	Thr	Leu	Leu	Cys	Ser	Leu	Cys	Tyr	Gly	Asp
				835		840		845							
Met	Glu	Asn	Lys	Leu	Phe	Ala	Gly	Ile	Arg	Arg	Asp	Gly	Leu	Leu	
				850		855		860							
Arg	Leu	Val	Asp	Asp	Phe	Leu	Leu	Val	Thr	Pro	His	Leu	Thr	His	Ala
				865		870		875		880					
Lys	Ala	Phe	Leu	Arg	Thr	Leu	Val	Arg	Gly	Val	Pro	Glu	Tyr	Gly	Cys
				885		890		895							
Val	Val	Asn	Leu	Arg	Lys	Thr	Val	Val	Asn	Phe	Pro	Val	Glu	Asp	Glu
				900		905		910							
Ala	Leu	Gly	Gly	Thr	Ala	Phe	Val	Gln	Leu	Pro	Ala	His	Gly	Leu	Phe
				915		920		925							
Pro	Trp	Cys	Gly	Leu	Leu	Leu	Asp	Thr	Arg	Thr	Leu	Glu	Val	Gln	Ser

-continued

930	935	940
Asp Tyr Ser Ser Tyr Ala Arg Thr Ser Ile Arg Ala Ser Leu Thr Phe		
945	950	955
		960
Asn Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys Leu Phe Gly		
965	970	975
Val Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu Gln Val Asn		
980	985	990
Ser Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu Leu Leu Gln		
995	1000	1005
Ala Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe His Gln		
1010	1015	1020
Gln Val Trp Lys Asn Pro Thr Phe Phe Leu Arg Ile Ile Ser Asp		
1025	1030	1035
Thr Ala Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly		
1040	1045	1050
Met Ser Leu Gly Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu		
1055	1060	1065
Ala Met Gln Trp Leu Cys His Gln Ala Phe Leu Leu Lys Leu Thr		
1070	1075	1080
Arg His Arg Val Thr Tyr Val Pro Leu Leu Gly Ser Leu Arg Thr		
1085	1090	1095
Ala Gln Thr Gln Leu Ser Arg Lys Leu Pro Gly Thr Thr Leu Ser		
1100	1105	1110
Ala Leu Glu Ala Ala Ala Asn Pro Ala Leu Pro Ser Asp Phe Lys		
1115	1120	1125
Thr Ile Leu Asp		
1130		

```

<210> SEQ_ID NO 19
<211> LENGTH: 4036
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 19

ctcgccggcgc gagtttcagg cagcgctgctg tccgtctgcg cacgtggaa gcccggccc 60
cggccaccccc cgccatggccg cgccgtccccc gctgccgagc cgtgcgcctcc ctgtctgcaca 120
gccactaccg cgagggtgtcg ccgctggcca cgttcgtgcg ggcctgggg ccccaggct 180
gggggctgtgt gcagcgccgg gacccggcg cttcccgccg cgtggtgccc cagtgcctgg 240
tgtgcgtgcc ctgggacgca cggccggccc cccggccccc ctccctccgc caggtgtcct 300
ccctgtgaagg gctggtgccc cgagtgtcg agaggctgtg cgagcgccggc gcgagaacg 360
tgtggccctt cggcttcgcg ctgtggacg gggcccgccg gggccccccc gaggcctca 420
ccaccagcggt ggcgcgactac ctgccccaca cggtgaccga cgcactgcgg gggagcgccc 480
cgtgggggct gctgctgcgc cgcgtggcg acgacgtgtc ggttcacctg ctggcacgct 540
gcgcgcgttt tgtgctggcg gctcccgatc ggcgcctacca ggtgtgcggg cccggcgctgt 600
accagcgccgg cgcgtgcact caggccggc cccggccaca cgcttagtggc ccccgaaaggc 660
gtcttaggatc cgaacggggc tggaccata gcgtcaggaa ggccggggc cccctggcc 720
tgccagcccc gggtgccagg aggccggggc gcagtgcctg ccgaagtctg ccgttgcaca 780

```

-continued

agaggcccag	gcgtggcgct	gccccctgago	cggagcggac	gcccgttggg	caggggtctt	840
ggcccccaccc	gggcaggacg	cgtggaccga	gtgaccgtgg	tttctgtgtg	gtgtcacctg	900
ccagacccgc	cgaagaagcc	acctctttgg	agggtgcgc	ctctggcacg	cgccactccc	960
acccatccgt	ggccggccag	caccacgcgg	gccccccatc	cacategcgg	ccaccacgtc	1020
cctgggacac	gccttgtccc	ccggtgtaeg	ccgagaccaa	gcacttcctc	tacttcctag	1080
gcgacaagga	gcagctcgcc	ccctccttcc	tactcagtc	tctgaggccc	agcctgactg	1140
gcgctcggag	gctcgtggag	accatcttc	tgggttccag	gcccctggatg	ccagggactc	1200
cccgcaagggt	gccccgcctg	ccccagcgc	actggcaaat	gccccccctg	tttctggagc	1260
tgcttggaa	ccacgcgcag	tgcccctacg	gggtgcttcc	caagacgcac	tgcccgcgtc	1320
gagctgcgg	cacccagca	gcccgtgtct	gtgcccggga	gaagccccag	ggctctgtgg	1380
cgccccccga	ggaggaggac	acagacccccc	gtcgccctgg	gcagctgc	cgccagcaca	1440
gcagccccctg	gcaggtgtac	ggcttcgtgc	gggcctgc	gcgcggctg	gtgccccag	1500
gcctctgggg	ctccaggcac	aacgaacgc	gcttcctc	gaacaccaag	aagttcatct	1560
ccctggggaa	gcatgccaag	ctctcgctgc	aggagctgac	gtggaagatg	agcgtgcggg	1620
actgcgcgtt	gctgcgcagg	agcccagggg	tggctctgt	tccggccgca	gagcaccgtc	1680
tgcgtgagga	gatcttgcc	aagttcttgc	actggctgt	gagtgtgtac	gttgcgagc	1740
tgctcaggtc	tttcttttat	gtcacggaga	ccacgtttca	gaagaacagg	ctcttttct	1800
accggaaagag	tgtctggagc	aagttgcaaa	gcatttgaat	cagacagcac	ttgaagaggg	1860
tgcagctgcg	ggagctgtcg	gaagcagagg	tcaggcagca	tcaggaagcc	aggccccccc	1920
tgctgacgtc	cagactccgc	ttcatcccc	agcctgacgg	gctgcggccg	attgtgaaca	1980
tggactacgt	cgtggagcc	agaacgttcc	gcagagaaaa	gaggggccgag	cgtctcacct	2040
cgaggggtgaa	ggcactgttc	agcgtgtca	actacgagcg	ggcgccgc	cccgccctcc	2100
tggcgccctc	tgtgtgggc	ctggacgata	tccacaggc	ctggcgcacc	ttcgtgtgc	2160
gtgtgcgggc	ccaggacccg	ccgcctgago	tgtactttgt	caaggtggat	gtgacggcgc	2220
cgtacgacac	catccccag	gacaggctca	cgagggtcat	cgccagcatc	atcaaacccc	2280
agaacacgt	ctgcgtgcgt	cggtatgt	tggtccagaa	ggccgc	ccat ggcacgtcc	2340
gcaaggcctt	caagagccac	gtctctacct	tgacagac	ccagccgtac	atgcgacagt	2400
tctgtggctca	cctgcaggag	accagccac	tgagggatgc	cgtcatcatc	gagcagagct	2460
cctccctgaa	tgagggccac	agtggctct	tgcacgttct	cctacgttcc	gtgtgcgc	2520
acggccgtgc	catcaggggc	aagtcttacg	tccagtgc	ggggatccc	cagggctcca	2580
tctgtccac	gtgtctgc	agcctgtgt	acggcgacat	ggagaacaag	ctgtttgcgg	2640
ggattcggcg	ggacgggctg	ctctgcgtt	tggtggatg	tttctgtgt	gtgacac	2700
acctcacc	cgcgaaagcc	ttcctcagga	ccctggctcc	aggtgtccct	gagtatggct	2760
gcgtgggtgaa	cttgcggaaag	acagtagtga	acttccctgt	agaagatgag	gccctgggtg	2820
gcacggctt	tgttcagctg	ccggccac	gcctattccc	ctggtgccgc	ctgctgtgg	2880
acacccggac	cctggagg	tgagcgtact	actccagcta	tgcccgacc	tccatcagag	2940
ccagtcac	cttcaaccgc	ggcttcaagg	ctgggaggaa	catgcgtc	gc aaactcttg	3000
gggtcttgcg	gctgaagtgt	cacagctgt	ttctggattt	gcaggtgaac	agcctccaga	3060

-continued

cggtgtgcac	caacatctac	aagatcctcc	tgcgtcaggc	gtacaggttt	cacgcatttg	3120
tgcgtcagct	cccatttc	catacgat	ttt ggaagaacc	cacatttttc	ctgcgcac	3180
tctctgacac	ggcctccctc	tgctactcca	tcctgaaagc	caagaacgc	ca gggatgtcgc	3240
tgggggceaa	gggtgcggcc	ggcctctgc	cctccgaggc	catgcagtgg	ctgtgccacc	3300
aagcattct	gctcaagctg	actcgacacc	gcgtcaccta	cgtgccactc	ctggggtac	3360
tcaggacagc	ccagacgcag	ctgagtcgga	agctccccgg	gacgacgcgt	agtgcctgg	3420
aggccgcagc	caacccggca	ctgcctcag	acttcaagac	catcctggac	tgtggccac	3480
ccgcccacag	ccggccggag	agcagacacc	agcagccctg	tcacgcccgg	ctctacgtcc	3540
cagggaggg	ggggggggccc	acacccagac	ccgcacccgt	gggagtctga	ggcctgagtg	3600
agtgtctggc	caaggcctgc	atgtccggct	gaaggctgag	tgtccagctg	aggcctgagc	3660
gagtgtccag	ccaaggcgt	agtgtccago	acacctgcgc	tcttcactc	cccacaggct	3720
ggcgctcggc	tccacccca	ggccagctt	tcctcgccag	gagcccgct	tcccactccc	3780
cacatggaa	tagtccatcc	ccagattcgc	cattgtccac	ccctcgccct	gcccctctt	3840
gccttccacg	cccaccatcc	agatggagac	cctgagaagg	accctgggag	ctctggaaat	3900
ttggagtgc	caaagggtgt	ccctgtacac	aggtgaggac	cctgcacctg	gatgggggtc	3960
cctgtgggtc	aaattgggg	gggggtgctg	tgggagtaaa	atactgaata	tatgagttt	4020
tcagtttga	aaaaaa					4036

<210> SEQ ID NO 20
<211> LENGTH: 1132
<212> TYPE: PRT
<213> ORGANISM: Pan troglodytes

<400> SEQUENCE: 20

Met	Pro	Arg	Ala	Pro	Arg	Cys	Arg	Ala	Val	Arg	Ser	Leu	Leu	Arg	Ser
1									5		10			15	

His	Tyr	Arg	Glu	Val	Leu	Pro	Leu	Ala	Thr	Phe	Val	Arg	Arg	Leu	Gly
									20		25			30	

Pro	Gln	Gly	Trp	Arg	Leu	Val	Gln	Arg	Gly	Asp	Pro	Ala	Ala	Phe	Arg
									35		40			45	

Ala	Leu	Val	Ala	Gln	Cys	Leu	Val	Cys	Val	Pro	Trp	Asp	Ala	Arg	Pro
									50		55			60	

Pro	Pro	Ala	Ala	Pro	Ser	Phe	Arg	Gln	Val	Ser	Cys	Leu	Lys	Glu	Leu
65									70		75			80	

Val	Ala	Arg	Val	Leu	Gln	Arg	Leu	Cys	Glu	Arg	Gly	Ala	Lys	Asn	Val
									85		90			95	

Leu	Ala	Phe	Gly	Phe	Ala	Leu	Leu	Asp	Gly	Ala	Arg	Gly	Pro	Pro	
									100		105			110	

Glu	Ala	Phe	Thr	Thr	Ser	Val	Arg	Ser	Tyr	Leu	Pro	Asn	Thr	Val	Thr
									115		120			125	

Asp	Ala	Leu	Arg	Gly	Ser	Gly	Ala	Trp	Gly	Leu	Leu	Leu	Arg	Arg	Val
									130		135			140	

Gly	Asp	Asp	Val	Leu	Val	His	Leu	Leu	Ala	Arg	Cys	Ala	Leu	Phe	Val
145									145		150			155	

Leu	Val	Ala	Pro	Ser	Cys	Ala	Tyr	Gln	Val	Cys	Gly	Pro	Pro	Leu	Tyr
									165		170			175	

-continued

Gln	Leu	Gly	Ala	Ala	Thr	Gln	Ala	Arg	Pro	Pro	Pro	His	Ala	Ser	Gly
180						185						190			
Pro	Arg	Arg	Arg	Leu	Gly	Cys	Glu	Arg	Ala	Trp	Asn	His	Ser	Val	Arg
195						200					205				
Glu	Ala	Gly	Val	Pro	Leu	Gly	Leu	Pro	Ala	Pro	Gly	Ala	Arg	Arg	Arg
210						215					220				
Gly	Gly	Ser	Ala	Ser	Arg	Ser	Leu	Pro	Leu	Pro	Lys	Arg	Pro	Arg	Arg
225						230				235		240			
Gly	Ala	Ala	Pro	Glu	Pro	Glu	Arg	Thr	Pro	Val	Gly	Gln	Gly	Ser	Trp
245						250				255					
Ala	His	Pro	Gly	Arg	Thr	Arg	Gly	Pro	Ser	Asp	Arg	Gly	Phe	Cys	Val
	260					265					270				
Val	Ser	Pro	Ala	Arg	Pro	Ala	Glu	Glu	Ala	Thr	Ser	Leu	Glu	Gly	Ala
275						280					285				
Leu	Ser	Gly	Thr	Arg	His	Ser	His	Pro	Ser	Val	Gly	Arg	Gln	His	His
290						295					300				
Ala	Gly	Pro	Pro	Ser	Thr	Ser	Arg	Pro	Pro	Arg	Pro	Trp	Asp	Thr	Pro
305						310				315		320			
Cys	Pro	Pro	Val	Tyr	Ala	Glu	Thr	Lys	His	Phe	Leu	Tyr	Ser	Ser	Gly
	325					330				335					
Asp	Lys	Glu	Gln	Leu	Arg	Pro	Ser	Phe	Leu	Leu	Ser	Ser	Leu	Arg	Pro
	340					345				350					
Ser	Leu	Thr	Gly	Ala	Arg	Arg	Leu	Val	Glu	Thr	Ile	Phe	Leu	Gly	Ser
	355					360				365					
Arg	Pro	Trp	Met	Pro	Gly	Thr	Pro	Arg	Arg	Leu	Pro	Arg	Leu	Pro	Gln
	370					375				380					
Arg	Tyr	Trp	Gln	Met	Arg	Pro	Leu	Phe	Leu	Glu	Leu	Gly	Asn	His	
385						390				395		400			
Ala	Gln	Cys	Pro	Tyr	Gly	Val	Leu	Leu	Lys	Thr	His	Cys	Pro	Leu	Arg
	405					410				415					
Ala	Ala	Val	Thr	Pro	Ala	Ala	Gly	Val	Cys	Ala	Arg	Glu	Lys	Pro	Gln
	420					425				430					
Gly	Ser	Val	Ala	Ala	Pro	Glu	Glu	Asp	Thr	Asp	Pro	Arg	Arg	Leu	
	435					440				445					
Val	Gln	Leu	Leu	Arg	Gln	His	Ser	Ser	Pro	Trp	Gln	Val	Tyr	Gly	Phe
	450					455				460					
Val	Arg	Ala	Cys	Leu	Arg	Arg	Leu	Val	Pro	Pro	Gly	Leu	Trp	Gly	Ser
465						470				475		480			
Arg	His	Asn	Glu	Arg	Arg	Phe	Leu	Arg	Asn	Thr	Lys	Lys	Phe	Ile	Ser
	485					490				495					
Leu	Gly	Lys	His	Ala	Lys	Leu	Ser	Leu	Gln	Glu	Leu	Thr	Trp	Lys	Met
	500					505				510					
Ser	Val	Arg	Asp	Cys	Ala	Trp	Leu	Arg	Arg	Ser	Pro	Gly	Val	Gly	Ser
	515					520				525					
Val	Pro	Ala	Ala	Glu	His	Arg	Leu	Arg	Glu	Glu	Ile	Leu	Ala	Lys	Phe
	530					535				540					
Leu	His	Trp	Leu	Met	Ser	Val	Tyr	Val	Val	Glu	Leu	Leu	Arg	Ser	Phe
	545					550				555		560			
Phe	Tyr	Val	Thr	Glu	Thr	Thr	Phe	Gln	Lys	Asn	Leu	Phe	Phe	Tyr	
	565					570				575					
Arg	Lys	Ser	Val	Trp	Ser	Lys	Leu	Gln	Ser	Ile	Gly	Ile	Arg	Gln	His

-continued

580	585	590
Leu Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln		
595	600	605
His Gln Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile		
610	615	620
Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val		
625	630	635
Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser		
645	650	655
Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg		
660	665	670
Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg		
675	680	685
Ala Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro		
690	695	700
Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile		
705	710	715
Pro Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln		
725	730	735
Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala His		
740	745	750
Gly His Val Arg Lys Ala Phe Lys Ser His Val Ser Thr Leu Thr Asp		
755	760	765
Leu Gln Pro Tyr Met Arg Gln Phe Val Ala His Leu Gln Glu Thr Ser		
770	775	780
Pro Leu Arg Asp Ala Val Ile Ile Glu Gln Ser Ser Ser Leu Asn Glu		
785	790	795
Ala Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe Val Cys Arg His		
805	810	815
Ala Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys Gln Gly Ile Pro		
820	825	830
Gln Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu Cys Tyr Gly Asp		
835	840	845
Met Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp Gly Leu Leu Leu		
850	855	860
Arg Leu Val Asp Asp Phe Leu Leu Val Thr Pro His Leu Thr His Ala		
865	870	875
Lys Ala Phe Leu Arg Thr Leu Val Arg Gly Val Pro Glu Tyr Gly Cys		
885	890	895
Val Val Asn Leu Arg Lys Thr Val Val Asn Phe Pro Val Glu Asp Glu		
900	905	910
Ala Leu Gly Gly Thr Ala Phe Val Gln Leu Pro Ala His Gly Leu Phe		
915	920	925
Pro Trp Cys Gly Leu Leu Asp Thr Arg Thr Leu Glu Val Gln Ser		
930	935	940
Asp Tyr Ser Ser Tyr Ala Arg Thr Ser Ile Arg Ala Ser Leu Thr Phe		
945	950	955
Asn Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys Leu Phe Gly		
965	970	975
Val Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu Gln Val Asn		
980	985	990

-continued

```

Ser Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu Leu Leu Gln
995          1000          1005

Ala Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe His Gln
1010          1015          1020

Gln Val Trp Lys Asn Pro Thr Phe Phe Leu Arg Ile Ile Ser Asp
1025          1030          1035

Thr Ala Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly
1040          1045          1050

Met Ser Leu Gly Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu
1055          1060          1065

Ala Met Gln Trp Leu Cys His Gln Ala Phe Leu Leu Lys Leu Thr
1070          1075          1080

Arg His Arg Val Thr Tyr Val Pro Leu Leu Gly Ser Leu Arg Thr
1085          1090          1095

Ala Gln Thr Gln Leu Ser Arg Lys Leu Pro Gly Thr Thr Leu Ser
1100          1105          1110

Ala Leu Glu Ala Ala Ala Asn Pro Ala Leu Pro Ser Asp Phe Lys
1115          1120          1125

Thr Ile Leu Asp
1130

```

```

<210> SEQ_ID NO 21
<211> LENGTH: 3396
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 21

atgccgcgcg cgccgcgcgtc ccgcgcggcg cgcagcctgc tgccgcagcca ttatcgcaaa 60
gtgtgcgcgc tggcgacattt tgtgcgcgcg ctggggccgc agggctggcg cctgggtcag 120
cgccggcgatc cggcgccgtt tcgcgcgtc gtggcgca gctgggtgtg cgtccgtgg 180
gtgcgcgcgc ccgcgcgggc ggcgcggcgcg tttcgccagg tgagctgcct gaaagaactg 240
gtggcgcgcg tgctgcagcg cctgtgcgaa cgccgcgcga aaaacgtgt ggcgtttggc 300
tttgcgtgc tggatggcgcc gcgcggcgcc ccgcgggaag cgtttaccac cagcgtgcgc 360
agctatctgc cgaacaccgt gaccgatgcg ctgcgcggca gcccgcgcgtg gggcctgcgt 420
ctgcgcgcgc tggcgatga tgtgtgggtg catctgtgg cgcgtgcgc gctgtttgtg 480
ctgggtggcgcc cgagctgcgc gtatcagggtg tgccgcgcgc cgctgtatca gctggcgcc 540
gcgcacccagg cgccgcgcgc gcccgcgtgc agccgcgcgc gcccgcgcgt gggctgcgaa 600
cgccgcgtggaa accatagcggt gcgcgcga ggcgtgcgcg tggccgtgc ggcgcgcggc 660
gcgcgcgcgc gccgcgcgcg cgcgcgcgcg agccgcgcgc tgccgaaacg cccgcgcgc 720
ggcgccgcgc cggaaaccggc acgcaccccg gtggggccagg gcagctggc gcatccggc 780
cgccaccccgcc gccccggcga tcgcggcgtt tgctgtgtga gccccggcgcg cccggcgaa 840
gaaggcgacca gcctggaaagg cgccgtgcgc ggcacccgc atagccatcc gagcgtggc 900
cgccacccgcg atgcgggccc gcccggcacc acgcgcgcgc cgcgcgcgtg ggatacccg 960
tgcccgccgg tggatgcggaa aaccaaacat tttctgtata gcagcggcga taaaacag 1020
ctgcgcgcgcg gctttctgt gaggcgcctg cggccggcgc tgaccggcgc ggcgcgcctg 1080

```

-continued

gtggaaacca	ttttctggg	cagecgcccg	tggatgcgg	gcaccccgcg	ccgcctgccg	1140
cgccctgcgc	agcgctattt	gcagatgcgc	ccgctgtttc	tggaactgt	ggcaaccat	1200
gcccgcgtcc	cgtatggcgt	gctgtgaaa	accattgc	cgctgcgcgc	ggcggtgacc	1260
ccggcgccgg	gcgtgtgcgc	gcgcgaaaaa	ccgcaggca	gcgtggcgcc	gccggaagaa	1320
gaagataaccg	atccgcgcg	cctggtgca	ctgctgcgc	agcatagcag	cccggtgcag	1380
gtgtatggct	ttgtgcgcgc	gtgcgcgc	cgcctggtc	cgccgggcct	gtggggcagc	1440
cgcataacg	aacgcgcctt	tctgcgcac	accaaaaat	ttattagcct	ggcaaacat	1500
gcgaaactga	gcctgcagga	actgacctgg	aaaatgagc	tgcgcgatg	cgcgtggctg	1560
cgcgcagcc	cgggcgtggg	cagegtgcgc	gcggcgaaac	atgcgcgc	cgaagaaatt	1620
ctggcgaaat	ttctgcattt	gctgtgago	gtgtatgtgg	tggaactgt	gcccgcgttt	1680
tttatgtga	ccgaaaccac	cttcagaaa	aaccgcctgt	tttttatcg	caaaagcgtg	1740
tggagcaaac	tgcagagcat	tggcattcgc	cagcatctga	aacgcgtgca	gctgcgcgaa	1800
ctgagcgaag	cggaaagtgcgc	ccagcatcag	gaagcgcgc	cggcgcgtgt	gaccagccgc	1860
ctgcgcgtta	ttccgaaacc	ggatggcctg	cgcgcatttgc	tgaacatgg	ttatgtgg	1920
ggcgcgca	ccttcgccc	cgaaaaacgc	gcggaaacgc	tgaccagcc	cgtgaaagcg	1980
ctgttagcg	tgctgaacta	tgaacgcgc	cgcgcggcgg	gcctgtgg	cgcgagcgtg	2040
ctgggcctgg	atgatattca	tgcgcgtgg	cgcaccttgc	tgctgcgcgt	gcccgcgcag	2100
gatccgcgc	cggaaactgt	ttttgtgaaa	gtggatgtga	ccggcgcgt	tgataccatt	2160
ccgcaggatc	gcctgaccga	agtgttgcgc	agcattatta	aaccgcagaa	cacctattgc	2220
gtgcgcgc	atgcgggttgt	gcagaaagcg	gcgcattggcc	atgtgcgc	agcgtttaaa	2280
agccatgtga	gcacccgtac	cgtatgcgc	ccgtatgtc	gccagttgt	ggcgcacatgc	2340
caggaaacca	gcgcgcgtgc	cgcgcgtgc	attattgt	agagcagcag	cctgaacgaa	2400
gcgcgcgc	gcctgtttga	tgtgtttctg	cgcgttgcgt	gcccgcattc	ggcgcgcatt	2460
cgcggcaaaa	gtatgtgc	gtgcgcgggc	atccgcagg	gcagcatttgc	gagcaccc	2520
ctgtgcgc	tgtgtatgg	cgtatggaa	aaacaactgt	ttgcgggc	tgcgcgcgt	2580
ggcgtgtgc	tgcgcctgg	ggatgattt	ctgctggta	ccccgcattc	gaccatgcgc	2640
aaagcgttgc	tgcgcaccc	ggtgcgcggc	gtgcgcggaa	atggcgtgcgt	ggcgcaccc	2700
cgcacccgc	tggtaactt	tccggtgaa	gatgaagcgc	tggcggcac	cgcgtttgt	2760
cagctgcgc	cgcattggc	gtttcgtgg	tgcggcgtgc	tgctggatac	ccgcaccc	2820
gaagtgcaga	gcgattatag	cagctatgcgc	cgcaccagca	tgcgcgcgc	cctgaccc	2880
aaccgcgc	ttaaagcggg	ccgcaacatgc	cgcgcacaa	tgtttggcgt	gtgcgcgt	2940
aaatgcgcata	gcctgtttct	ggatctgcgc	gtgaacagc	tgcagaccgt	gtgcaccaac	3000
atttataaaa	ttctgtgt	gcaggcgtat	cgcatttc	cgatgcgtgt	gcagctgcgc	3060
tttcatcgc	aggtgtggaa	aaacccgcacc	tttttctgc	gcattattag	cgataccgcgc	3120
agcctgtgt	atagcatttgc	gaaagcgaaa	aacgcggca	tgagcctgg	cgcgaaaggc	3180
gcggcgccgc	cgcgcgcgc	cgcgcgcgc	cgcgcgcgc	gttcgtgt	3240	
aaactgaccc	gcacatgcgc	gacctatgt	ccgcgtgt	gcagcctgcgc	caccgcgcag	3300
acccagctga	gcgcacaa	actgcggcacc	accctgagc	cgctggaa	ggcggcgac	3360

-continued

ccggcgctgc cgagcgattt taaaaccatt ctggat	3396
<210> SEQ ID NO 22	
<211> LENGTH: 807	
<212> TYPE: PRT	
<213> ORGANISM: Pan troglodytes	
<400> SEQUENCE: 22	
Met Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu Leu Arg Ser	
1 5 10 15	
His Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg Arg Leu Gly	
20 25 30	
Pro Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro Ala Ala Phe Arg	
35 40 45	
Ala Leu Val Ala Gln Cys Leu Val Cys Val Pro Trp Asp Ala Arg Pro	
50 55 60	
Pro Pro Ala Ala Pro Ser Phe Arg Gln Val Ser Cys Leu Lys Glu Leu	
65 70 75 80	
Val Ala Arg Val Leu Gln Arg Leu Cys Glu Arg Gly Ala Lys Asn Val	
85 90 95	
Leu Ala Phe Gly Phe Ala Leu Leu Asp Gly Ala Arg Gly Pro Pro	
100 105 110	
Glu Ala Phe Thr Thr Ser Val Arg Ser Tyr Leu Pro Asn Thr Val Thr	
115 120 125	
Asp Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu Leu Arg Arg Val	
130 135 140	
Gly Asp Asp Val Leu Val His Leu Leu Ala Arg Cys Ala Leu Phe Val	
145 150 155 160	
Leu Val Ala Pro Ser Cys Ala Tyr Gln Val Cys Gly Pro Pro Leu Tyr	
165 170 175	
Gln Leu Gly Ala Ala Thr Gln Ala Arg Pro Pro Pro His Ala Ser Gly	
180 185 190	
Pro Arg Arg Arg Leu Gly Cys Glu Arg Ala Trp Asn His Ser Val Arg	
195 200 205	
Glu Ala Gly Val Pro Leu Gly Leu Pro Ala Pro Gly Ala Arg Arg Arg	
210 215 220	
Gly Gly Ser Ala Ser Arg Ser Leu Pro Leu Pro Lys Arg Pro Arg Arg	
225 230 235 240	
Gly Ala Ala Pro Glu Pro Glu Arg Thr Pro Val Gly Gln Gly Ser Trp	
245 250 255	
Ala His Pro Gly Arg Thr Arg Gly Pro Ser Asp Arg Gly Phe Cys Val	
260 265 270	
Val Ser Pro Ala Arg Pro Ala Glu Glu Ala Thr Ser Leu Glu Gly Ala	
275 280 285	
Leu Ser Gly Thr Arg His Ser His Pro Ser Val Gly Arg Gln His His	
290 295 300	
Ala Gly Pro Pro Ser Thr Ser Arg Pro Pro Arg Pro Trp Asp Thr Pro	
305 310 315 320	
Cys Pro Pro Val Tyr Ala Glu Thr Lys His Phe Leu Tyr Ser Ser Gly	
325 330 335	
Asp Lys Glu Gln Leu Arg Pro Ser Phe Leu Leu Ser Ser Leu Arg Pro	
340 345 350	

-continued

Ser Leu Thr Gly Ala Arg Arg Leu Val Glu Thr Ile Phe Leu Gly Ser
 355 360 365
 Arg Pro Trp Met Pro Gly Thr Pro Arg Arg Leu Pro Arg Leu Pro Gln
 370 375 380
 Arg Tyr Trp Gln Met Arg Pro Leu Phe Leu Glu Leu Leu Gly Asn His
 385 390 395 400
 Ala Gln Cys Pro Tyr Gly Val Leu Leu Lys Thr His Cys Pro Leu Arg
 405 410 415
 Ala Ala Val Thr Pro Ala Ala Gly Val Cys Ala Arg Glu Lys Pro Gln
 420 425 430
 Gly Ser Val Ala Ala Pro Glu Glu Asp Thr Asp Pro Arg Arg Leu
 435 440 445
 Val Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln Val Tyr Gly Phe
 450 455 460
 Val Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu Trp Gly Ser
 465 470 475 480
 Arg His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser
 485 490 495
 Leu Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr Trp Lys Met
 500 505 510
 Ser Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro Gly Val Gly Ser
 515 520 525
 Val Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu Ala Lys Phe
 530 535 540
 Leu His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe
 545 550 555 560
 Phe Tyr Val Thr Glu Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr
 565 570 575
 Arg Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His
 580 585 590
 Leu Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln
 595 600 605
 His Gln Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile
 610 615 620
 Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val
 625 630 635 640
 Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser
 645 650 655
 Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg
 660 665 670
 Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg
 675 680 685
 Ala Trp Arg Thr Phe Val Leu Arg Val Ala Gln Asp Pro Pro Pro
 690 695 700
 Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile
 705 710 715 720
 Pro Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln
 725 730 735
 Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala His
 740 745 750

-continued

Gly	His	Val	Arg	Lys	Ala	Phe	Lys	Ser	His	Val	Leu	Arg	Pro	Val	Pro
755															

Gly	Asp	Pro	Ala	Gly	Leu	His	Pro	Val	His	Ala	Ala	Leu	Gln	Pro	Val
770															

Leu	Arg	Arg	His	Gly	Glu	Gln	Ala	Val	Cys	Gly	Asp	Ser	Ala	Gly	Arg
785															

Ala	Ala	Pro	Ala	Phe	Gly	Gly									
							805								

<210> SEQ ID NO 23

<211> LENGTH: 2421

<212> TYPE: DNA

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 23

atgcccgcgc	cggcgccgtg	ccggcgccgtg	cgcagcctgc	tgcgcagccaa	ttatcgcaaa	60
gtgtgtgcgc	tggcgacccctt	tgttgccgcgc	ctggggccgc	agggtggcg	cctgggtgcag	120
cggggcgatc	ggggcgccgtt	tcggcgccgtg	gtggcgccgt	gcttgggtgt	cgtgcgcgtgg	180
gtatgcgcgc	cggcgccggc	ggcgccgcgc	tttcgccagg	ttagctgcct	gaaaagaactg	240
gtggcgccgc	tgctgcagcg	cctgtgcgaa	cggggcgccga	aaaacgtgt	ggcggtttggc	300
tttgcgtgc	tggatggcgc	gcggggcgcc	ccggccggaa	cgtttaccac	cagcgtgcgc	360
agctatctgc	cgaacaccgt	gacccatgt	ctggcgccgc	gccccgtgt	gggcctgtgt	420
ctggccgcgc	tggggcgatga	tgtgtgtgt	catctgtgg	cgcgcgtgcgc	gtgtttgtg	480
ctgggtggcgc	cgagctgcgc	gtatcagggt	tggggccgc	cgctgtatca	gctggggcg	540
ggcacccagg	cgcgcggcgc	gccgcattgt	ageggggccgc	gccccgcct	gggcgtcgaa	600
cgcgcgtgg	accatagcgt	gcgcgaagcg	ggcggtgcgc	tgggcctg	ggcgccgggc	660
ggcgccgcgc	ggggggccgc	cgcgcggcgc	agectgcgc	tgcgcgaaacg	cccgccgcgc	720
ggcgccggcgc	cggAACCCGG	acgcaccccg	gtggggccagg	gcagctgggc	gcataccggc	780
cgcaccccg	ccccggcgca	tcggggctt	tgcgtgggt	gcggggcg	cccgccggaa	840
gaagcgacca	gccttggagg	cgcgtgtgc	ggcacccgc	atagccatcc	gagcgtggc	900
cgcacccatc	atgcggggcc	gccggcgacc	agccggccgc	cgcgcggcgt	ggatacccg	960
tgcggccgg	tgtatgcgga	aaccaaacat	tttctgtata	gcagcggcga	taaagaacag	1020
cgcggccgc	gttttctgt	gagcagccct	cgcggccgc	tgaccggcgc	gcccgcctg	1080
gtggaaacca	tttttctggg	cagccggcc	tggatgcgc	gcaccccg	ccgcctgcgc	1140
cgcctgcgc	agcgctattt	gcagatgcgc	ccgcgttttgc	tggaaactgt	ggcaaccat	1200
gcccggcgc	cgtatggcgt	gctgtgtaaa	accattgtcc	cgcgtgcgc	ggcggtgacc	1260
ccggcgccgg	gcgtgtgcgc	gcgcgaaaaa	ccgcaggggca	gcgtggccgc	gccggaaagaa	1320
gaagataccg	atccggccgc	cctgggtgcag	ctgtgtgcgc	acatagcgt	ccccgtggcag	1380
gtgtatggct	ttgtgcgcgc	gtgcctgcgc	cgccctgggt	cgccgggcct	gtggggcagc	1440
cgcataacg	aacggccgtt	tctgcgcac	accaaaaaat	ttattagcct	ggcaaccat	1500
gcgaaactga	gcctgcagga	actgacccgtt	aaaatgagcg	tgcgcgat	cgcggtgcgt	1560
cgcgcagcc	cgggcggtgg	cagegtgcgc	gcggcgaaac	atgcgcctgcgc	cgaagaaatt	1620

-continued

ctggcgaaat ttctgcattt gctgatgago gtgttatgtgg tggaaactgtc ggcagcgttt	1680
ttttatgtga ccgaaaccac ctttcagaaa aaccgcctgt tttttatcg caaaagcgtg	1740
tggagcaaac tgcagagcat tggcattcgc cagcatctga aacgcgtgca gctgcgcgaa	1800
ctgagcgaag cgaaagtgcg ccagcatcg gaagcgcgcg cggcgctgtc gaccagccgc	1860
ctgcgcctta ttccgaaacc ggatggcctg cggccgattt gtaacatgga ttatgtggtg	1920
ggcgccgca ctttcgccc cgaaaaacgc gggaaacgc tgaccagccg cgtaaaacgc	1980
ctgttttagcg tgctgaacta tgaacgcgcg cggcccccgg gcgtgtggg cgcgagcgtg	2040
ctgggcctgg atgatattca tcgcgcgtgg cgcaccccttg tgctgcccgt ggcgcgcag	2100
gatccgcgcg cggaaactgtt ttttgtaaa gtggatgtga cggcgcgota tgataccatt	2160
ccgcaggatc gcctgaccga agtggattgcg agcattatta aaccgcagaa cacctattgc	2220
gtgcgcgcgt atgcgggtggt gcagaagcgc ggcgcgcgcg atgtgcgcgaa agcgtttaaa	2280
agccatgtgc tgccgcgcgtt gcccggcgat cccggccggcc tgcatccggt gcatgcgcg	2340
ctgcagccgg tgctgcgcgcg ccatggcga caggccgtgt gcccgcgatag cccggccgc	2400
ggggcgccgg cggttggcg c	2421

<210> SEQ ID NO 24

<211> LENGTH: 1069

<212> TYPE: PRT

<213> ORGANISM: Pan troglodytes

<400> SEQUENCE: 24

Met Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu Leu Arg Ser			
1	5	10	15

His Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg Arg Leu Gly			
20	25	30	

Pro Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro Ala Ala Phe Arg			
35	40	45	

Ala Leu Val Ala Gln Cys Leu Val Cys Val Pro Trp Asp Ala Arg Pro			
50	55	60	

Pro Pro Ala Ala Pro Ser Phe Arg Gln Val Ser Cys Leu Lys Glu Leu			
65	70	75	80

Val Ala Arg Val Leu Gln Arg Leu Cys Glu Arg Gly Ala Lys Asn Val			
85	90	95	

Leu Ala Phe Gly Phe Ala Leu Leu Asp Gly Ala Arg Gly Pro Pro			
100	105	110	

Glu Ala Phe Thr Thr Ser Val Arg Ser Tyr Leu Pro Asn Thr Val Thr			
115	120	125	

Asp Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu Leu Arg Arg Val			
130	135	140	

Gly Asp Asp Val Leu Val His Leu Leu Ala Arg Cys Ala Leu Phe Val			
145	150	155	160

Leu Val Ala Pro Ser Cys Ala Tyr Gln Val Cys Gly Pro Pro Leu Tyr			
165	170	175	

Gln Leu Gly Ala Ala Thr Gln Ala Arg Pro Pro Pro His Ala Ser Gly			
180	185	190	

Pro Arg Arg Arg Leu Gly Cys Glu Arg Ala Trp Asn His Ser Val Arg			
195	200	205	

Glu Ala Gly Val Pro Leu Gly Leu Pro Ala Pro Gly Ala Arg Arg Arg	
---	--

-continued

210	215	220														
Gly	Gly	Ser														
Ala	Ala	Ser														
Arg	Arg	Ser														
Leu	Leu	Pro														
Pro	Leu	Pro														
Lys	Arg	Pro														
Arg	Arg	Arg														
225	230	235	240													
Gly	Gly	Ala	Ala	Pro	Glu	Pro	Glu	Arg	Thr	Pro	Val	Gly	Gln	Gly	Ser	Trp
245	250	255														
Ala	His	Pro	Gly	Arg	Thr	Arg	Gly	Pro	Ser	Asp	Arg	Gly	Phe	Cys	Val	
260	265	270														
Val	Ser	Pro	Ala	Arg	Pro	Ala	Glu	Glu	Ala	Thr	Ser	Leu	Glu	Gly	Ala	
275	280	285														
Leu	Ser	Gly	Thr	Arg	His	Ser	His	Pro	Ser	Val	Gly	Arg	Gln	His	His	
290	295	300														
Ala	Gly	Pro	Pro	Ser	Thr	Ser	Arg	Pro	Pro	Arg	Pro	Trp	Asp	Thr	Pro	
305	310	315	320													
Cys	Pro	Pro	Val	Tyr	Ala	Glu	Thr	Lys	His	Phe	Leu	Tyr	Ser	Ser	Gly	
325	330	335														
Asp	Lys	Glu	Gln	Leu	Arg	Pro	Ser	Phe	Leu	Leu	Ser	Ser	Leu	Arg	Pro	
340	345	350														
Ser	Leu	Thr	Gly	Ala	Arg	Arg	Leu	Val	Glu	Thr	Ile	Phe	Leu	Gly	Ser	
355	360	365														
Arg	Pro	Trp	Met	Pro	Gly	Thr	Pro	Arg	Arg	Leu	Pro	Arg	Leu	Pro	Gln	
370	375	380														
Arg	Tyr	Trp	Gln	Met	Arg	Pro	Leu	Phe	Leu	Glu	Leu	Leu	Gly	Asn	His	
385	390	395	400													
Ala	Gln	Cys	Pro	Tyr	Gly	Val	Leu	Leu	Lys	Thr	His	Cys	Pro	Leu	Arg	
405	410	415														
Ala	Ala	Val	Thr	Pro	Ala	Ala	Gly	Val	Cys	Ala	Arg	Glu	Lys	Pro	Gln	
420	425	430														
Gly	Ser	Val	Ala	Ala	Pro	Glu	Glu	Glu	Asp	Thr	Asp	Pro	Arg	Arg	Leu	
435	440	445														
Val	Gln	Leu	Leu	Arg	Gln	His	Ser	Ser	Pro	Trp	Gln	Val	Tyr	Gly	Phe	
450	455	460														
Val	Arg	Ala	Cys	Leu	Arg	Arg	Leu	Val	Pro	Pro	Gly	Leu	Trp	Gly	Ser	
465	470	475	480													
Arg	His	Asn	Glu	Arg	Arg	Phe	Leu	Arg	Asn	Thr	Lys	Lys	Phe	Ile	Ser	
485	490	495														
Leu	Gly	Lys	His	Ala	Lys	Leu	Ser	Leu	Gln	Glu	Leu	Thr	Trp	Lys	Met	
500	505	510														
Ser	Val	Arg	Asp	Cys	Ala	Trp	Leu	Arg	Arg	Ser	Pro	Gly	Val	Gly	Ser	
515	520	525														
Val	Pro	Ala	Ala	Glu	His	Arg	Leu	Arg	Glu	Glu	Ile	Leu	Ala	Lys	Phe	
530	535	540														
Leu	His	Trp	Leu	Met	Ser	Val	Tyr	Val	Val	Glu	Leu	Leu	Arg	Ser	Phe	
545	550	555	560													
Phe	Tyr	Val	Thr	Glu	Thr	Thr	Phe	Gln	Lys	Asn	Arg	Leu	Phe	Phe	Tyr	
565	570	575														
Arg	Lys	Ser	Val	Trp	Ser	Lys	Leu	Gln	Ser	Ile	Gly	Ile	Arg	Gln	His	
580	585	590														
Leu	Lys	Arg	Val	Gln	Leu	Arg	Glu	Leu	Ser	Glu	Ala	Glu	Val	Arg	Gln	
595	600	605														
His	Gln	Glu	Ala	Arg	Pro	Ala	Leu	Leu	Thr	Ser	Arg	Leu	Arg	Phe	Ile	
610	615	620														

-continued

Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val
 625 630 635 640
 Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser
 645 650 655
 Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg
 660 665 670
 Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg
 675 680 685
 Ala Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro
 690 695 700
 Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile
 705 710 715 720
 Pro Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln
 725 730 735
 Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala His
 740 745 750
 Gly His Val Arg Lys Ala Phe Lys Ser His Val Ser Thr Leu Thr Asp
 755 760 765
 Leu Gln Pro Tyr Met Arg Gln Phe Val Ala His Leu Gln Glu Thr Ser
 770 775 780
 Pro Leu Arg Asp Ala Val Ile Ile Glu Gln Ser Ser Ser Leu Asn Glu
 785 790 795 800
 Ala Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe Val Cys Arg His
 805 810 815
 Ala Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys Gln Gly Ile Pro
 820 825 830
 Gln Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu Cys Tyr Gly Asp
 835 840 845
 Met Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp Gly Leu Leu Leu
 850 855 860
 Arg Leu Val Asp Asp Phe Leu Leu Val Thr Pro His Leu Thr His Ala
 865 870 875 880
 Lys Ala Phe Leu Ser Tyr Ala Arg Thr Ser Ile Arg Ala Ser Leu Thr
 885 890 895
 Phe Asn Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys Leu Phe
 900 905 910
 Gly Val Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu Gln Val
 915 920 925
 Asn Ser Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu Leu Leu
 930 935 940
 Gln Ala Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe His Gln
 945 950 955 960
 Gln Val Trp Lys Asn Pro Thr Phe Phe Leu Arg Ile Ile Ser Asp Thr
 965 970 975
 Ala Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly Met Ser
 980 985 990
 Leu Gly Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu Ala Met Gln
 995 1000 1005
 Trp Leu Cys His Gln Ala Phe Leu Leu Lys Leu Thr Arg His Arg
 1010 1015 1020

-continued

Val	Thr	Tyr	Val	Pro	Leu	Leu	Gly	Ser	Leu	Arg	Thr	Ala	Gln	Thr
1025					1030					1035				

Gln	Leu	Ser	Arg	Lys	Leu	Pro	Gly	Thr	Thr	Leu	Ser	Ala	Leu	Glu
1040					1045					1050				

Ala	Ala	Ala	Asn	Pro	Ala	Leu	Pro	Ser	Asp	Phe	Lys	Thr	Ile	Leu
1055					1060					1065				

Asp

```

<210> SEQ ID NO 25
<211> LENGTH: 3207
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

<400> SEQUENCE: 25

atgcgcgcgc	cgccgcgctg	ccgcgcggcgt	cgcagcctgc	tgcgcagccca	ttatcgcaaa	60
gtgtgcgcgc	tggcgacctt	tgtgcgcgcgc	ctggggccgc	agggctggcg	cctgggtgcag	120
cgcggcgatc	cggcgccggtt	tcgcgcgctg	gtggcgcaagt	gcctgggtgtg	cgtgcgcgtgg	180
gatgcgcgcgc	cgccgcgcggc	ggcgccgcgc	tttcgcagg	tgagctgcct	gaaagaactg	240
gtggcgcgcg	tgctgcagcg	cctgtgcgaa	cgcggcgccga	aaaacgtgct	ggcgtttggc	300
tttgcgtgc	tggatggcgc	gchgccccggc	ccgcggggaa	cgtttaccac	cagcgtgcgc	360
agctatctgc	cgaacaccgt	gaccgatgcg	ctgcgcggca	gcggcgccgtg	gggcctgctg	420
ctgcgcgcgc	tggcgatga	tgtgcgttgt	catctgctgg	cgcgcgtgcgc	gctgttttgt	480
ctgggtggcgc	c gagactgcgc	gtatcagggt	tgccggccgc	cgctgtatca	gctggggcg	540
gcgacccagg	cgcgcggccgc	gccgcatgcg	agcgccccgc	gccgcggccct	gggctgcgaa	600
cgcgcgtggaa	accatagcgt	g c g c g a a g c g	ggcgtgcgcgc	tgggcctgca	ggcgcggggc	660
gcgcgcgcgc	gcggcgccgc	cgcgagccgc	agccgtccgc	tgcgcggaaac	cccgcgcgc	720
ggcgccggcgc	cggAACCGGA	acgcaccccg	gtggggccagg	gcagctgggc	gcattccggc	780
cgcaccccg	gcggcgccgc	tcgcggctt	tgcgttgtga	gcccggcg	cccggcgaa	840
gaagcgcacca	gcctggaaagg	cgcgcgtgcgc	ggcaccccg	atagccatcc	gagcgtggc	900
cgcgcgcgc	atgcggggccc	gccgagcacc	agccgcggc	cgcgcggcgt	ggatacccg	960
tgcccgccgg	tgtatgcgga	aaccaaacat	tttctgtata	gcagcggcga	taaagaacag	1020
ctgcgcgcgc	gttttctgtct	gagcagcctg	cgcggcgcgc	tgaccggcgc	gcccgcgc	1080
gtggaaaccca	tttttctggg	cagccggccgc	tggatgcggg	gcaccccg	ccgcctgcgc	1140
cgcctgcgcgc	agcgctattg	gcagatgcgc	ccgcgtttt	tggactgtct	ggcaaccat	1200
gcgcagtgc	cgtatggcgt	gctgtgtaaa	accattgc	cgctgcgcgc	ggcggtgacc	1260
ccggcgccgg	gcgtgtgcgc	gcgcgaaaaaa	ccgcaggc	gcgtggcg	gccgaaagaa	1320
gaagataccg	atccgcgcgc	cctgggtgcag	ctgcgcgc	agcatagcag	cccgtggcag	1380
gtgtatggct	ttgtgcgcgc	gtgcctgcgc	cgcctgggtgc	cgcggggcct	gtggggcagc	1440
cgcataacg	aacgcgcctt	tctgcgcac	accaaaaaat	ttattagcct	ggcaacat	1500
gcgaaactga	gcctgcagga	actgacctgg	aaaatgagc	tgcgcgattg	cgcgtggct	1560
cgcgcgcgc	cggcgctggg	cagcgtgcgc	gcggcgaaac	atgcgcgc	cgaagaatt	1620
ctggcgaaat	ttctgcattt	gctgtatgago	gtgtatgtgg	tggactgtct	gcgcagctt	1680

-continued

tttatgtga	ccgaaaccac	ctttcagaaa	aaccgcctgt	tttttatcg	caaaagcgtg	1740
tggagcaaac	tgcagagcat	tggcattcgc	cagcatctga	aacgcgtgca	gctgcccga	1800
ctgagcgaag	cggaagtgcg	ccgcatcg	gaagcgcgc	ggcgctgt	gaccagccgc	1860
ctgcgcctta	ttccgaaacc	ggatggcctg	cgeccgattt	tgaacatgga	ttatgtggtg	1920
ggcgcgcgca	ccttcgccc	cgaaaaacgc	gcggaacgcc	tgaccagccg	cgtgaaagcg	1980
ctgttttagcg	tgctgaacta	tgaacgcgc	cgccgcgg	gcctgtggg	cgcgagcgt	2040
ctgggcctgg	atgatattca	tcgcgcgtgg	cgcaccttg	tgctgcgcgt	gcgcgcgcag	2100
gatccgcgc	cggaactgta	ttttgtgaaa	gtggatgtga	ccggcgcgt	tgataccatt	2160
ccgcaggatc	gcctgaccga	agtgattgcg	agcattatta	aaccgcagaa	cacctattgc	2220
gtgcgcgcgt	atgcgggtgg	gcagaaagcg	gcgcatggcc	atgtgcgc	agcgtttaaa	2280
agccatgtga	gcacctgac	cgatctgcg	ccgtatatgc	gccagttgt	ggcgcacat	2340
caggaaacca	ccccgcgtcg	cgatgcgg	attattgaac	agagcagcag	cctgaacgaa	2400
gcgagcagcg	gcctgttga	tgtgttctg	cgcttgcgt	gccgcctatc	ggtgcgcatt	2460
cgcggcaaaa	gctatgtca	gtgccaggc	attccgcagg	gcagcattct	gagcaccc	2520
ctgtgcagcc	tgtgctatgg	cgatattggaa	aacaactgt	ttgcgggc	atcgccgc	2580
ggcctgtgc	tgcgcctgg	ggatgat	tttctgtgt	ccccgcatct	gaccatgcg	2640
aaagcgtttc	ttagctatgc	gcgcaccago	attcgcgcga	gcctgac	taaccgcggc	2700
tttaaagcgg	gccgcaacat	gcgccc	ctgttggcg	tgctgcgc	gaaatgc	2760
agcctgttgc	tggatctgc	ggtaaacago	ctgcagacc	tgtgcac	aaatataaa	2820
attctgtgc	tgcaggcgta	tgcgttcat	gcgtgcgtgc	tgcagctgc	gttcatcag	2880
caggtgtgga	aaaacccgac	cttttctg	cgcattatta	gcgataccgc	gagcgtgc	2940
tatagcattc	tgaaagcgaa	aaacgcggc	atgagcctgg	gcgcgaaagg	cgcggcggc	3000
ccgctgcga	gcgaagcgat	gcagtggctg	tgccatcagg	cgtttctgt	gaaactgacc	3060
cgccatcg	tgacctatgt	gcccgtgt	ggcagcctgc	gcacccgc	gaccagctg	3120
agccgcaaac	tgccgggcac	caccctgago	gcgcgtggaa	ggcggcgaa	cccgccgt	3180
ccgagcgt	ttaaaaccat	tctggat				3207

<210> SEQ ID NO 26

<211> LENGTH: 795

<212> TYPE: PRT

<213> ORGANISM: Pan troglodytes

<400> SEQUENCE: 26

Met	Pro	Arg	Ala	Pro	Arg	Cys	Arg	Ala	Val	Arg	Ser	Leu	Leu	Arg	Ser
1				5				10				15			

His	Tyr	Arg	Glu	Val	Leu	Pro	Leu	Ala	Thr	Phe	Val	Arg	Arg	Ley	Gly
				20			25				30				

Pro	Gln	Gly	Trp	Arg	Leu	Val	Gln	Arg	Gly	Asp	Pro	Ala	Ala	Phe	Arg
				35			40				45				

Ala	Leu	Val	Ala	Gln	Cys	Leu	Val	Cys	Val	Pro	Trp	Asp	Ala	Arg	Pro
				50			55			60					

Pro	Pro	Ala	Ala	Pro	Ser	Phe	Arg	Gln	Val	Ser	Cys	Leu	Lys	Glu	Ley
				65			70			75			80		

-continued

Val Ala Arg Val Leu Gln Arg Leu Cys Glu Arg Gly Ala Lys Asn Val			
85	90	95	
Leu Ala Phe Gly Phe Ala Leu Leu Asp Gly Ala Arg Gly Pro Pro			
100	105	110	
Glu Ala Phe Thr Thr Ser Val Arg Ser Tyr Leu Pro Asn Thr Val Thr			
115	120	125	
Asp Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu Leu Leu Arg Arg Val			
130	135	140	
Gly Asp Asp Val Leu Val His Leu Leu Ala Arg Cys Ala Leu Phe Val			
145	150	155	160
Leu Val Ala Pro Ser Cys Ala Tyr Gln Val Cys Gly Pro Pro Leu Tyr			
165	170	175	
Gln Leu Gly Ala Ala Thr Gln Ala Arg Pro Pro Pro His Ala Ser Gly			
180	185	190	
Pro Arg Arg Arg Leu Gly Cys Glu Arg Ala Trp Asn His Ser Val Arg			
195	200	205	
Glu Ala Gly Val Pro Leu Gly Leu Pro Ala Pro Gly Ala Arg Arg Arg			
210	215	220	
Gly Gly Ser Ala Ser Arg Ser Leu Pro Leu Pro Lys Arg Pro Arg Arg			
225	230	235	240
Gly Ala Ala Pro Glu Pro Glu Arg Thr Pro Val Gly Gln Gly Ser Trp			
245	250	255	
Ala His Pro Gly Arg Thr Arg Gly Pro Ser Asp Arg Gly Phe Cys Val			
260	265	270	
Val Ser Pro Ala Arg Pro Ala Glu Glu Ala Thr Ser Leu Glu Gly Ala			
275	280	285	
Leu Ser Gly Thr Arg His Ser His Pro Ser Val Gly Arg Gln His His			
290	295	300	
Ala Gly Pro Pro Ser Thr Ser Arg Pro Pro Arg Pro Trp Asp Thr Pro			
305	310	315	320
Cys Pro Pro Val Tyr Ala Glu Thr Lys His Phe Leu Tyr Ser Ser Gly			
325	330	335	
Asp Lys Glu Gln Leu Arg Pro Ser Phe Leu Leu Ser Ser Leu Arg Pro			
340	345	350	
Ser Leu Thr Gly Ala Arg Arg Leu Val Glu Thr Ile Phe Leu Gly Ser			
355	360	365	
Arg Pro Trp Met Pro Gly Thr Pro Arg Arg Leu Pro Arg Leu Pro Gln			
370	375	380	
Arg Tyr Trp Gln Met Arg Pro Leu Phe Leu Glu Leu Leu Gly Asn His			
385	390	395	400
Ala Gln Cys Pro Tyr Gly Val Leu Leu Lys Thr His Cys Pro Leu Arg			
405	410	415	
Ala Ala Val Thr Pro Ala Ala Gly Val Cys Ala Arg Glu Lys Pro Gln			
420	425	430	
Gly Ser Val Ala Ala Pro Glu Glu Asp Thr Asp Pro Arg Arg Leu			
435	440	445	
Val Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln Val Tyr Gly Phe			
450	455	460	
Val Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu Trp Gly Ser			
465	470	475	480
Arg His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser			

-continued

485	490	495
Leu Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr Trp Lys Met		
500	505	510
Ser Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro Gly Val Gly Ser		
515	520	525
Val Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu Ala Lys Phe		
530	535	540
Leu His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe		
545	550	555
Phe Tyr Val Thr Glu Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr		
565	570	575
Arg Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His		
580	585	590
Leu Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln		
595	600	605
His Gln Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile		
610	615	620
Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val		
625	630	635
Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser		
645	650	655
Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg		
660	665	670
Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg		
675	680	685
Ala Trp Arg Thr Phe Val Leu Arg Val Ala Gln Asp Pro Pro Pro		
690	695	700
Glu Leu Tyr Phe Val Lys Asp Arg Leu Thr Glu Val Ile Ala Ser Ile		
705	710	715
Ile Lys Pro Gln Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln		
725	730	735
Lys Ala Ala His Gly His Val Arg Lys Ala Phe Lys Ser His Val Leu		
740	745	750
Arg Pro Val Pro Gly Asp Pro Ala Gly Leu His Pro Val His Ala Ala		
755	760	765
Leu Gln Pro Val Leu Arg Arg His Gly Glu Gln Ala Val Cys Gly Asp		
770	775	780
Ser Ala Gly Arg Ala Ala Pro Ala Phe Gly Gly		
785	790	795

<210> SEQ ID NO 27
<211> LENGTH: 2385
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 27

atgccgcg	cggcgccgtg	ccgcgcggcgtc	cgcagcctgc	tgcgcagccaa	60
gtgtgcgcgc	tggcgaccctt	tgtgcgcgcgc	ctggggccgc	agggctggcg	120
cgcggcgatc	cggcgccgtt	tgcgcgcgtg	gtggcgca	gcctgggtgt	180
gatgcgcgccc	cggcgccggc	ggcgccgago	tttcgcccagg	tgagctgcct	240

-continued

gtggcgcg	tgctgcagcg	cctgtgcgaa	cggcgcgaga	aaaacgtgt	ggcggttggc	300
tttcgctgc	tggatggcg	gcgccccggc	ccggccggaa	cgtttaccac	cagcgtgcgc	360
agctatctgc	cgAACACCGT	gaccgatcg	ctggcgccgc	gcccgcgcgt	gggcctgtg	420
ctgcgcgcg	tggggatga	tgtgtgtgt	catctgtgt	cgcgctgcgc	gcttttgt	480
ctgggtggcg	cgagctgcgc	gtatcagggt	tgcggccccgc	cgctgtatca	gctgggcgcg	540
gcgcacccagg	cgcgccccgc	gccgcgtgcgc	agcgccccgc	gcccgcgcct	gggcgtcgaa	600
cgcgcggtgg	accatagcgt	gcgcaageg	ggcggtgcgc	tgggcctgca	ggcgccgggc	660
gcgcgcgcgc	cgggcgccag	cgcgagccgc	agcctgcgc	tgcggaaacgc	cccgcgcgc	720
ggcgcgccgc	cggaaccgg	acgcaccccg	gtggggccagg	gcagctggc	gcattccggc	780
cgcaccccg	gcggagcga	tgcgggttt	tgcgtgtgt	gcccggcg	cccgccggaa	840
gaagcgtacca	gcctggagg	cgcgctgago	ggcacccgc	atagccatcc	gagcgtggc	900
cgccagcatc	atgcggggcc	gccgagcacc	agccgcgcgc	cgcgccccgt	ggataaccccg	960
tgcgcgcgc	tgtatgcgg	aaccaaacat	tttctgtata	gcagcggcga	taaagaacag	1020
ctgcgcgcga	gttttctgt	gagcagcctg	cgccccgagcc	tgaccggcgc	gcccgcgcct	1080
gtggaaacca	tttttctgg	cagccgcgc	tggatgcgg	gcacccgcgc	ccgcctgcgc	1140
cgccgcgcgc	agcgctattt	gcagatgcgc	ccgctgtttc	tggactgt	ggcaacccat	1200
gcmcagtgc	cgtatggcgt	gctgtgtaaa	accattgc	cgctgcgcgc	ggcggtgacc	1260
ccggcgccgg	gcgtgtgcgc	gcgcgaaaaa	ccgcaggcga	gcgtggcg	gccggaagaa	1320
gaagataccg	atccgcgcg	cctgggtcag	ctgctgcgc	agcatagcag	cccggtgcag	1380
gtgtatggct	ttgtgcgcgc	gtgcctgcgc	cgccctgg	cgccgggcct	gtggggcagc	1440
cgccataacg	aacgcgcctt	tctgcgcac	accaaaaat	ttattagcct	ggcaacccat	1500
gcgaaactga	gcctgcagga	actgacctgg	aaaatgagcg	tgcgcgattt	cgcggtgc	1560
cgccgcgcgc	cgggcggtgg	cagcgccgc	gcggcgaa	atgcgcgc	cgaagaaatt	1620
ctggcgaaat	ttctgcattt	gctgtatgc	gtgtatgtgg	tggactgt	gcgcagctt	1680
tttttatgtga	ccgaaaccac	cttcagaaa	aaccgcctgt	tttttatgc	caaaagcgt	1740
tggagcaaac	tgcagagcat	tggcattcgc	cagcatctga	aacgcgtgc	gctgcgcga	1800
ctgagcgaag	cggaagtgcg	ccagcatcg	gaagcgcgc	cgccgcgt	gaccagccgc	1860
ctgcgcgttta	ttccgaaacc	ggatggcctg	cgcccgattt	tgaacatgg	ttatgtgg	1920
ggcgcgccgc	ccttcgcgc	cgaaaaacgc	gcggaaacgc	tgaccaggc	cgtgaaacgc	1980
ctgttttagcg	tgctgaacta	tgaacgcgc	cgccgcgcgg	gcctgtgtgg	cgcgagcgt	2040
ctgggcctgg	atgatattca	tcgcgcgtgg	cgccaccc	tgctgcgcgt	gcgcgcgc	2100
gatccgcgcgc	cggaactgt	ttttgtgaaa	gatgcgcgt	ccgaagtgt	tgcgagcatt	2160
attaaaccgc	agaacaccta	ttgcgtgcgc	cgctatgcgg	tggtgcagaa	agcggcgcat	2220
ggccatgtgc	gaaaagccat	gtgcgtgcgc	cggtgcgg	cgatccggcg	2280	
ggcctgcatac	cggtgcatac	ggcgctgcag	ccggcgtgc	gcccgcattt	cgaacaggcg	2340
gtgtgcggcg	atagcgcggg	ccgegcgg	ccggcggtt	ggggc		2385

-continued

<212> TYPE: PRT

<213> ORGANISM: Pan troglodytes

<400> SEQUENCE: 28

Met Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu Leu Arg Ser
1 5 10 15

His Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg Arg Leu Gly
20 25 30

Pro Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro Ala Ala Phe Arg
35 40 45

Ala Leu Val Ala Gln Cys Leu Val Cys Val Pro Trp Asp Ala Arg Pro
50 55 60

Pro Pro Ala Ala Pro Ser Phe Arg Gln Val Ser Cys Leu Lys Glu Leu
65 70 75 80

Val Ala Arg Val Leu Gln Arg Leu Cys Glu Arg Gly Ala Lys Asn Val
85 90 95

Leu Ala Phe Gly Phe Ala Leu Leu Asp Gly Ala Arg Gly Pro Pro
100 105 110

Glu Ala Phe Thr Thr Ser Val Arg Ser Tyr Leu Pro Asn Thr Val Thr
115 120 125

Asp Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu Leu Leu Arg Arg Val
130 135 140

Gly Asp Asp Val Leu Val His Leu Leu Ala Arg Cys Ala Leu Phe Val
145 150 155 160

Leu Val Ala Pro Ser Cys Ala Tyr Gln Val Cys Gly Pro Pro Leu Tyr
165 170 175

Gln Leu Gly Ala Ala Thr Gln Ala Arg Pro Pro Pro His Ala Ser Gly
180 185 190

Pro Arg Arg Arg Leu Gly Cys Glu Arg Ala Trp Asn His Ser Val Arg
195 200 205

Glu Ala Gly Val Pro Leu Gly Leu Pro Ala Pro Gly Ala Arg Arg Arg
210 215 220

Gly Gly Ser Ala Ser Arg Ser Leu Pro Leu Pro Lys Arg Pro Arg Arg
225 230 235 240

Gly Ala Ala Pro Glu Pro Glu Arg Thr Pro Val Gly Gln Gly Ser Trp
245 250 255

Ala His Pro Gly Arg Thr Arg Gly Pro Ser Asp Arg Gly Phe Cys Val
260 265 270

Val Ser Pro Ala Arg Pro Ala Glu Glu Ala Thr Ser Leu Glu Gly Ala
275 280 285

Leu Ser Gly Thr Arg His Ser His Pro Ser Val Gly Arg Gln His His
290 295 300

Ala Gly Pro Pro Ser Thr Ser Arg Pro Pro Arg Pro Trp Asp Thr Pro
305 310 315 320

Cys Pro Pro Val Tyr Ala Glu Thr Lys His Phe Leu Tyr Ser Ser Gly
325 330 335

Asp Lys Glu Gln Leu Arg Pro Ser Phe Leu Leu Ser Ser Leu Arg Pro
340 345 350

Ser Leu Thr Gly Ala Arg Arg Leu Val Glu Thr Ile Phe Leu Gly Ser
355 360 365

Arg Pro Trp Met Pro Gly Thr Pro Arg Arg Leu Pro Arg Leu Pro Gln
370 375 380

-continued

Arg Tyr Trp Gln Met Arg Pro Leu Phe Leu Glu Leu Leu Gly Asn His			
385	390	395	400
Ala Gln Cys Pro Tyr Gly Val Leu Leu Lys Thr His Cys Pro Leu Arg			
405	410	415	
Ala Ala Val Thr Pro Ala Ala Gly Val Cys Ala Arg Glu Lys Pro Gln			
420	425	430	
Gly Ser Val Ala Ala Pro Glu Glu Asp Thr Asp Pro Arg Arg Leu			
435	440	445	
Val Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln Val Tyr Gly Phe			
450	455	460	
Val Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu Trp Gly Ser			
465	470	475	480
Arg His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser			
485	490	495	
Leu Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr Trp Lys Met			
500	505	510	
Ser Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro Gly Val Gly Ser			
515	520	525	
Val Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu Ala Lys Phe			
530	535	540	
Leu His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe			
545	550	555	560
Phe Tyr Val Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr			
565	570	575	
Arg Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His			
580	585	590	
Leu Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln			
595	600	605	
His Gln Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile			
610	615	620	
Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val			
625	630	635	640
Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser			
645	650	655	
Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg			
660	665	670	
Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg			
675	680	685	
Ala Trp Arg Thr Phe Val Leu Arg Val Ala Gln Asp Pro Pro Pro			
690	695	700	
Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile			
705	710	715	720
Pro Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln			
725	730	735	
Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala His			
740	745	750	
Gly His Val Arg Lys Ala Phe Lys Ser His Val Ser Thr Leu Thr Asp			
755	760	765	
Leu Gln Pro Tyr Met Arg Gln Phe Val Ala His Leu Gln Glu Thr Ser			
770	775	780	

-continued

Pro	Leu	Arg	Asp	Ala	Val	Ile	Ile	Glu	Gln	Ser	Ser	Ser	Leu	Asn	Glu
785					790			795							800
Ala Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe Val Cys Arg His															
805						810			815						
Ala Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys Gln Gly Ile Pro															
820						825			830						
Gln Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu Cys Tyr Gly Asp															
835						840			845						
Met Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp Gly Leu Leu Leu															
850						855			860						
Arg Leu Val Asp Asp Phe Leu Leu Val Thr Pro His Leu Thr His Ala															
865						870			875						880
Lys Ala Phe Leu Arg Thr Leu Val Arg Gly Val Pro Glu Tyr Gly Cys															
885						890			895						
Val Val Asn Leu Arg Lys Thr Val Val Asn Phe Pro Val Glu Asp Glu															
900						905			910						
Ala Leu Gly Gly Thr Ala Phe Val Gln Leu Pro Ala His Gly Leu Phe															
915						920			925						
Pro Trp Cys Gly Leu Leu Asp Thr Arg Thr Leu Glu Val Gln Ser															
930						935			940						
Asp Tyr Ser Ser Tyr Ala Arg Thr Ser Ile Arg Ala Ser Leu Thr Phe															
945						950			955						960
Asn Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys Leu Phe Gly															
965						970			975						
Val Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu Gln Val Asn															
980						985			990						
Ser Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu Leu Leu Gln															
995						1000			1005						
Ala Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe His Gln															
1010						1015			1020						
Gln Val Trp Lys Asn Pro Thr Phe Phe Leu Arg Ile Ile Ser Asp															
1025						1030			1035						
Thr Ala Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Ala															
1040						1045			1050						
Gln Thr Gln Leu Ser Arg Lys Leu Pro Gly Thr Thr Leu Ser Ala															
1055						1060			1065						
Leu Glu Ala Ala Ala Asn Pro Ala Leu Pro Ser Asp Phe Lys Thr															
1070						1075			1080						
Ile Leu Asp															
1085															

```

<210> SEQ ID NO 29
<211> LENGTH: 3258
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 29

atgcgcgcg cggcgccgtc cggcgccgtc cgccgcgcgc tgcgcagccaa ttatcgca 60
gtgctgccgc tggcgacctt tgtgcgccgc ctggggccgc agggctggcg cctggtgca 120
cgccggcgatc cggcgccgtt tcgcgcgcgt gtcggcgact gcctggtgca cgtgcgcgtt 180

```

-continued

gatgcgcggcc	cgccgcggc	ggcgcggago	tttcgccagg	tgagctgcct	gaaagaactg	240
gtggcgcgcg	tgctgcagcg	cctgtgcgaa	cgccgcgcga	aaaacgtgt	ggcgttggc	300
tttgcgtgc	tggatggcgc	gcccggggc	ccgcggaa	cgtttaccac	cagcgtgcgc	360
agctatctgc	cgaacaccgt	gaccgtatgc	ctgcgcggca	gcccgcgtg	gggcctgctg	420
ctgcgcgcgc	tgggcgatga	tgtgtgggt	catctgtgg	cgcgctgcgc	gctgtttgt	480
ctgggtggcgc	cgagctgcgc	gtatcagggt	tgcggccccg	cgctgtata	gctggggcgcg	540
gcccgcagg	cgccgcgcgc	gccgcgtgcgc	agcgccccgc	gcccgcgcct	gggcgtgcga	600
cgcgcgtgga	accatagcgt	gcgcgaageg	ggcggtgcgc	tgggcctgca	gggcgcgggc	660
gcccgcgcgc	cgccggggcag	cgcgagccgc	agectgcgcgc	tgccgcggaa	cccgcgcgc	720
ggcgccggcgc	cggaaccggc	acgcaccccg	gtggggcagg	cgagctgggc	gcataccggc	780
cgcaccccg	cccccgagcga	tcgcggcttt	tgcgtggta	gcccggcgcg	ccccggcgaa	840
gaagcgcacca	gccttggaaagg	cgcgctgago	ggcacccgc	atagccatcc	gagcgtggc	900
cgcacgcac	atgcggggcc	gccgagcacc	agccgcgcgc	cgccgcgcgt	ggatacccg	960
tgcccgccgg	tgtatgcgga	aaccaaacat	ttttgtata	cgagcggcga	taaagaacag	1020
ctgcgcgcga	gttttgcgt	gagcagccctg	cgcccgagcg	tgaccggcgc	gcccgcgcct	1080
gtggaaacca	tttttctggg	cagecgcccg	tggatgcgg	gcaccccgcg	ccgcctgccc	1140
cgcctgcgcgc	acgcgtattt	gcagatgcgc	ccgcgtttt	tggactgt	ggcaaccat	1200
gcccgcgtcc	cgtatggcgt	gctgtgaaa	accattgccc	cgctgegcgc	ggcggtgacc	1260
ccggcgccgg	gcgtgtgcgc	gcccggaaaa	ccgcaggcgc	gcgtggcgcc	gcccggaa	1320
gaagataacc	atcccgcccg	cctggtgcag	ctgcgtgcgc	agcatagcag	cccggtgcag	1380
gtgtatggct	ttgtgcgcgc	gtgcgtgcgc	cgcctggtgc	cgccgggcct	gtggggcagc	1440
cgcacataac	aacgcgcgtt	tctgcgcac	accaaaaaaaaat	ttattagcct	ggcaaccat	1500
gcgaaactga	gcctgcagga	actgacctgg	aaaatgagcg	tgcgcgatgt	cgcggtgct	1560
cgcgcgcagcc	cggcgctggg	cagegtgcgc	gcggcgaaac	atgcgcgcgc	cgaagaatt	1620
ctggcgaaat	ttctgcattt	gctgtatgago	gtgtatgtt	tggactgt	gcccgcgcct	1680
tttatgtga	ccgaaaccac	cttcagaaa	aaccgcctgt	tttttatcg	caaaacgtg	1740
tggagcaaac	tgcagagcat	tggcattgc	cagcatctga	aacgcgtgca	gctgcgcgaa	1800
ctgagcgaag	cggaagtgcg	ccagcatcg	gaagcgcgc	cggcgctgt	gaccagccgc	1860
ctgcgcctta	ttccgaaacc	ggatggcctg	cgcccgattt	tgaacatgga	ttatgtgt	1920
ggcgccgcga	cctttccgc	cgaaaaacgc	gccaacgcgc	tgaccagccg	cgtaaagcg	1980
ctgtttagcg	tgctgaacta	tgaacgcgc	cgccggccgg	ccctgtggg	cgcgacgt	2040
ctgggcctgg	atgatattca	tcgcgcgtgg	cgcaccttgc	tgctgcgcgt	gcccgcgcag	2100
gatccgcgc	cggaactgt	ttttgtgaaa	gtggatgtga	ccggcgccgt	tgataccatt	2160
cgcaggatc	gcctgaccga	agtgtattgc	agcattatta	aaccgcagaa	cacctattgc	2220
gtgcgcgcct	atgcgggtgt	gcagaaagcg	gcccgcgc	atgtgcgc	agcgttaaa	2280
agccatgtga	gcacccctgac	cgatctgcgc	ccgtatgtc	gccagttgt	ggcgcatctg	2340
caggaaacca	gcccgcgc	cgatgcgggt	attattgaac	agagcagcag	cctgaacgaa	2400
gcgagcagcg	gcctgtttga	tgtgtttctg	cgcttgcgt	ggcccatgc	ggtgcgcatt	2460

-continued

cgcggcaaaa	gctatgtcga	tgccaggc	attccgcagg	gcagcattct	gagcacccctg	2520
ctgtcgaccc	tgtgctatgg	cgatatggaa	aacaactgt	ttgcgggcat	tcgcccgcat	2580
ggcctgtgc	tgcgccctgg	ggatgatttt	ctgctggta	ccccgcac	gaccatgcg	2640
aaagcgtttc	tgcgccacct	ggtgcgcggc	gtgcggaaat	atggctgcgt	ggtaaacctg	2700
cgcaaaaaccg	tggtaaactt	tccggtgaa	gatgaagcgc	tggggggcac	cgcgttgtg	2760
cagctgcggg	cgcattggct	gtttccgtgg	tgccgcctgc	tgctggatac	ccgcacccctg	2820
gaagtgcaga	gcgattatag	cagctatgcg	cgcaccagca	ttcgcgcgag	cctgacccttt	2880
aaccgcggct	ttaaagcggg	ccgcaacatg	cgcgcacaaac	tgtttggcgt	gctgcgcctg	2940
aaatgccata	gctgtttct	ggatctgcag	gtgaacagcc	tgcagaccgt	gtgcaccaac	3000
atttataaaa	ttctgtgtct	gcaggcgtat	cgcttcatg	cgtgctgtct	gcagctgcgg	3060
tttcateagc	aggtgtggaa	aaaccgcacc	tttttctgc	gcattattag	cgataccgcg	3120
agcctgtgtct	atagcattct	gaaagcgaaa	aacgcgcggc	agaccaggct	gagccgc当地	3180
ctgccgggca	ccaccctgag	cgcgtggaa	gcggcggcga	acccggcgct	gccgagcgat	3240
tttaaaaacca	ttctggat					3258

```
<210> SEQ ID NO 30
<211> LENGTH: 3399
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic
```

<400> SEQUENCE: 30
atgcccggcg ctcggccgtg ccggagccgtg cgctccctgc tgccgcagcca ctaccggcag 60
gtgtgtgcgc tggccacgtt cgtggggcgc ctggggcccc agggctggcg gctgtgcag 120
cgcgccccacc cggcggtttt ccggcgcgctg gtggcccaagt gcttgggttg cgtccccctgg 180
gacgcacggc cgccccccgc cgccccctcc ttccggcagg tgcctcgct gaaggagctg 240
gtggccccag tgcgtcagag gctgtgcagag cgccggcgaga agaacgtgtc ggccttcggc 300
ttccggcgctgc tggacggggc ccggcgggggc ccccccgggg ctttcaccac cagcgctgcgc 360
agtcacactgc ccaacacggt gaccgacgca ctggggggga gggggcggtg ggggctgtcg 420
ctggcccgcg tggggcgacga cgtgtggttt cacatgtgtt caccgtgcgc gcttttgtg 480
ctgggtggctc ccagctgcgc ctaccagggtg tgccggccgc cgctgtacca gctggcgct 540
ggccactcagg cccggcccccc gccacacgct agtggaccccc ggaggcgctc gggatcgaa 600
cgggccctggg accatagcgt cagggaggcc ggggtcccccc tggggctgcc agcccggggt 660
ggcgaggaggc gggggggcag tgccggccga agtctggcgat tgcccaagag gcccaggcg 720
ggcgctgcgc ctgagccggc gccggacggcc gttgggcagg ggtccctggc ccaccgggc 780
aggacggcggt gaccggagtga ccgtgggttc tgggtgggtt cacatgtgtt acggccggaa 840
gaagccaccc tttggagggt tgccgtctt ggcacggccg actccacccc atccgtggc 900
cgccaggcacc acggggggccccc cccatccaca tcggggccac cacatgtgtt ggacacggct 960
tgtccccccgg tgcgtcccgat gaccaaggac ttccgttact cctcaggcgat caaggaggcg 1020
ctggggccct ctttcttact cagatctctg agggcccgatc tgactggcgatc tggggaggctc 1080
gtggagacca ttttttggg ttccaggccc tggatggccg gggactcccg cagggtggcc 1140

-continued

cgccctgcccc	agcgctactg	gcaaatgcgg	cccctgttcc	tggagctgt	tggaaaccac	1200
gcccagtgcc	cctacgggg	gtcttcaga	acgcactgcc	cgctgcgagc	tgcggtcacc	1260
ccaggcagccg	gtgtctgtgc	ccgggagaag	ccccagggg	ctgtggcgcc	ccccgaggag	1320
gaggacacag	accccccgtcg	cctgggtcgag	ctgtccgccc	agcacagcag	cccctggcag	1380
gtgtacggct	tcgtggggc	ctgcgtgcgc	cggctggtg	ccccaggcct	ctggggctcc	1440
aggcacaacg	aacgcgcgtt	cctcaggaac	accaagaagt	tcatctccct	gggaaagcat	1500
gccaagctct	cgctgcagga	gctgacgtgg	aaatgagcgc	tgccggactg	cgcttggctg	1560
cgcaggagcc	caggggttgg	ctgtgttccg	gccgcagagc	accgtctgcg	tgaggagatc	1620
ctggccaa	gttgcactg	gctgtatgagt	gtgtacgtcg	tgcgttgt	caggctttc	1680
tttatgtca	cggagaccac	gtttcaaaag	aacaggctct	tttctaccg	gaagagtgtc	1740
tggagcaagt	tgcaaaagcat	tggatcaga	cagcacttgc	agagggtgca	gtcgccggag	1800
ctgtcggaag	cagaggtcag	gcagoatcg	gaagccaggc	cgcgcctgt	gacgtccaga	1860
ctccgcgttca	tcccaagcc	tgacgggtcg	cggccgattt	tgaacatgg	ctacgtcg	1920
ggagccagaa	cgttccgcag	agaaaagagg	gccgagcg	tacactcg	ggtaaggca	1980
ctgttcagcg	tgctcaacta	cgaggggg	cggccgc	gcctctgg	cgcctctgt	2040
ctgggcctgg	acgatatcca	cagggcctgg	cgeaccttc	tgctgegtgt	gcgggcccag	2100
gaccgcgc	ctgagctgt	ctttgtcaag	gtggatgt	cgggcgcgt	cgacaccatc	2160
ccccaggaca	ggctcacg	ggcatcgcc	ageatcatca	aaaaaggaa	cacgtactgc	2220
gtgcgtcggt	atgcccgtgt	ccagaaggcc	gcccattgg	acgtccgc	ggccttcaag	2280
agccacgtct	ctaccttgc	agacctccag	ccgtacatgc	gacagttcg	ggctcacctg	2340
caggagacca	gcccgtgag	ggatgcgc	gtcatcg	agagctcc	cctgaatgag	2400
gccagcgt	gccttcc	cgttccatg	gccaccac	cgtgcgc	2460	
agggcaagt	cctacgtcc	gtgcaggg	atcccgc	gtccatc	ctccacgt	2520
ctctgcagcc	tgtgtacgg	cgacatgg	aacaagct	ttgcgggat	tcgggggac	2580
gggctgtcc	tgcgtttgg	ggatgattt	ttgttgg	cacctcac	cacccac	2640
aaaacccctcc	tcaggacc	ggtccgaggt	gtccctg	atggctgc	ggtaact	2700
ccgaagacag	tggtaactt	ccctgttagaa	gacgagcc	tgggtgg	ggctttgtt	2760
cagatgcgc	cccacggc	atccccctgg	tgccgc	tgctggata	ccggacc	2820
gaggtgcaga	gcgactactc	cagctatgc	cgacacttca	tca	gagccag	2880
aaccgcgc	tcaaggctgg	gaggaacat	cgtcg	caac	tctcac	2940
aagtgtcaca	gcctgttct	ggatgtcg	gtgaa	acagc	gtgcacca	3000
atctacaaga	tcctcctgt	gcaggcgt	agg	ttc	gcac	3060
tttcatcagc	aagttggaa	gaacccaca	ttttcctgc	gcgtcat	tgcac	3120
tcctctgt	actccatc	gaaagcc	aacgc	agg	ggccaagg	3180
gccggccgc	ctctgc	cgaggccgt	cag	ttgc	gtgc	3240
aagctgactc	gacaccgt	cac	cttgc	gg	tcgtactc	3300
acgcagctga	gtcggaa	cccggg	acg	ctgt	gacagcc	3360
ccggcactgc	cctcagactt	caagaccatc	ctggactg			3399

-continued

<210> SEQ ID NO 31
<211> LENGTH: 3369
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 31

atgaccgcgc	tcctcggtt	ccccgcggtg	cgtctctgc	tgcgcagccg	ataccgggag	60
gtgtggccgc	tggcaacctt	tgtggggcgc	ctggggcccg	agggcaggcg	gcttgtcaa	120
cccggggacc	cgaagatcta	ccgacttttgc	gttgccaaat	gcctagtgtg	catgcactgg	180
ggctcacage	ctccacactgc	cgacccccc	ttccaccagg	tgtcatccct	gaaagagctg	240
gtggccaggg	ttgtgcagag	actctgcgag	cgcaacgaga	gaaacgtgct	ggctttggc	300
tttgagctgc	ttaacgaggc	cagaggccgg	cctccatgg	ccttcaactag	tagcgtgcgt	360
agctacttgc	ccaacactgt	tattgagacc	ctgcgtgtca	gtggtgcatg	gatgctactg	420
tttagccgag	tgggcacga	cctgtggtc	tacctgtgg	cacactgtgc	tctttatctt	480
ctgggtcccc	ccagctgtgc	ctaccagggt	tgtgggtctc	ccctgtacca	aatttgtgcc	540
accacggata	tctggccctc	tgtgtccgt	agttacaggc	ccacccgacc	cgtgggcagg	600
aatttcacta	accttaggtt	cttacaacag	atcaagagca	gtagtegc当地	ggaagcacccg	660
aaacccctgg	ccttgccttc	tcgaggtaca	aaaggccatc	tgagtctcac	cagtacaagt	720
gtgccttcag	ctaagaaggc	cagatgttat	cctgtcccg	gagtggagga	gggaccacac	780
aggcagggtgc	taccaacccc	atcaggccaa	tcatgggtgc	caagtcttc	tcgggtcccc	840
gaggtgccta	ctgcagagaa	agatttgtct	tctaaaggaa	aggtgtctga	cctgagtc	900
tctgggtcgg	tgtgtgtaa	acacaagccc	agtcacat	ctctgtgtc	accacccgc	960
caaatgcct	ttagtcag	gccattttt	gagaccagac	atttcctta	ctccagggg	1020
gatggccaag	agcgtctaaa	cccttcattt	ctactcaga	acctccagcc	taacttgact	1080
ggggccagga	gactgggtgg	gatcatctt	ctgggtctaa	ggccttaggac	atcaggacca	1140
ctctgcagga	cacacccgt	atcgcgtcga	tactggcaga	tgcggeccct	gttccaacag	1200
ctgctgggt	accatgcaga	gtgccaat	gtcagactcc	tcaggtcaca	ttgcagg	1260
cgaacacgaa	accaacaggt	gacagatgcc	ttgaacacca	gcccacccgca	cctcatggat	1320
ttgctccgccc	tgcacagcag	tccctggcag	gtatatgggt	ttcttgggc	ctgtctctgc	1380
aagggtgggt	ctgcttagtct	ctggggtacc	aggcacaatg	agcgccgctt	ctttaagaac	1440
ttaaagaagt	tcatctcg	ggggaaatac	ggcaagctat	cactgcagga	actgtatgtgg	1500
aaagatgaaag	tagaggattt	ccactggctc	cgcagcagcc	cggggaaagga	ccgtgtcccc	1560
gctgcagagc	accgtctgag	ggagaggatc	ctggctacgt	tcctgttctg	gctgtatggac	1620
acatacgtgg	tacagctgt	taggtcattt	ttttacatca	cagagagcac	attccagaag	1680
aacaggctct	tcttctaccg	taagatgtg	tggagcaagc	tgcagagcat	tggagtcaagg	1740
caacacccctt	agagagtgcg	gctacgggag	ctgtcacaag	aggaggtcag	gcatcacca	1800
gacacccctggc	tagccatgcc	catctcaga	ctgcgttca	tcccaagcc	caacggctcg	1860
cggcccatgg	tgaacatgag	ttatagcatg	ggtaccagag	cttggggcag	aaggaagcag	1920
gcccagcatt	tcacccagcg	tctcaagact	ctcttcagca	tgctcaacta	tgagcggaca	1980

-continued

aacacatcctc accttatggg gtcttctgt a ctgggttatga atgacatcta caggacctgg	2040
cgggcctttg tgctgcgtgt gcggtgtctg gaccagacac ccaggatgt a ctttgttaag	2100
gcagatgtga cccggggctta tgatgccatc ccccagggtt a gctgggttga gggtgttgcc	2160
aatatgtatca ggcactcgga gaggcacgtac t gatccgccc agtatgcagt ggtccggaga	2220
gatagccaag gccaaggcca caagtccctt aggagacagg tcaccaccc ctctgaccc	2280
cagccataca tgggcagtt ccttaagcat ctgcaggatt cagatgccag tgcactgagg	2340
aactccgttg tcatcgagca gagcatctct atgaatgaga gcagcagcag cctgtttgac	2400
tttttcctgc acttcctgcg tcacagtgtc gtaaagattt gtgacagggtt ctataccgc	2460
tgcaggcga tccccccagggtt ctccagccata tccaccctgc tctgcgtct gtgtttcgga	2520
gacatggaga acaagctgtt tgctgagggtt cagcgggatgt gggtgtttt acgtttgtt	2580
gatgacttgc tggtgtgtac gcctcaacttgc gaccaagcaaa aaacccctt cagcaccctg	2640
gtccatggcg ttccctgagta tgggtgcattt ataaacttgc agaagacagt ggtgaacttc	2700
cctgtggagc ctggtagccctt ggggtgtca gctccatacc agtgcctgc tcactgcctg	2760
tttccctgtt gtggcttgc gctggacactt cagactttgg aggtgttctg tgactactca	2820
ggttatgccca agacccatcaat taagacgago ctcaccccttcc agagtgttctt caaagctggg	2880
aagaccatgc ggaacaagct cctgtcggtt ttgcgggttga agtgcacccg tctatttcta	2940
gacttgcagg tgaacacgcctt ccagacagtc tgcataataa tatacaagat cttccctgtt	3000
caggcctaca ggttccatgc atgtgtgattt cagttccctt ttgaccaggcg tgtaggaag	3060
aacccatccat tctttctggg catcatctcc agccaagcat cctgctgtca tgctatccctg	3120
aaggtaaga atccaggaat gacactaaag gcctctggctt ctttccctcc tgaagccgca	3180
cattggctct gtcaccaggc ctccctgctt aagctggctt ctcattctgt catctacaaa	3240
tgtctctgg gacctctgag gacagcccaa aaactgctgtt gccggaaagctt cccagaggcg	3300
acaatgacca tccttaaagc tgcagctgac ccagccctaa gcacagactt tcagaccattt	3360
ttggactaa	3369

<210> SEQ ID NO 32
<211> LENGTH: 102
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 32
ttaagctgcc ttctgcgggg ctgccttctt ggccatgccc ttcttccttc ccttgcacctt 60
gtacccctttt gtcattttttt aatggcttgcg taggtttttt ag 102

<210> SEQ ID NO 33
<211> LENGTH: 144
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 33
caattggctt gttttttttt gtcattttttt tttttttttt ccctaaatgtcc 60
aactactaaa ctgggggata ttatgtttttt ctttgcgtttt ctttgcgtttt 120

-continued

aacattttattt ttcattgcga attc	144
<210> SEQ ID NO 34	
<211> LENGTH: 145	
<212> TYPE: DNA	
<213> ORGANISM: Artificial	
<220> FEATURE:	
<223> OTHER INFORMATION: Synthetic	
<400> SEQUENCE: 34	
caattggctc gctttttgc tgtccaattt ctattaaagg ttccctttgt tccctaagtc	60
caactactaa actgggggat attatgaagg gccttgagca tctggattct gcctaataaa	120
aaacatttct ttcattgcg aattc	145
<210> SEQ ID NO 35	
<211> LENGTH: 47	
<212> TYPE: DNA	
<213> ORGANISM: Artificial	
<220> FEATURE:	
<223> OTHER INFORMATION: Synthetic	
<400> SEQUENCE: 35	
agggaaaataag agagaaaaga agagtaagaa gaaaatataag agccacc	47
<210> SEQ ID NO 36	
<211> LENGTH: 57	
<212> TYPE: DNA	
<213> ORGANISM: Artificial	
<220> FEATURE:	
<223> OTHER INFORMATION: Synthetic	
<400> SEQUENCE: 36	
ggacattgc ttctgacaca actgtgttca ctagcaacct caaacaacta gtacacc	57
<210> SEQ ID NO 37	
<211> LENGTH: 3723	
<212> TYPE: DNA	
<213> ORGANISM: Artificial	
<220> FEATURE:	
<223> OTHER INFORMATION: Synthetic	
<400> SEQUENCE: 37	
ggcacatttgc tttctgacac aactgtgttca actagcaacc tcaaacaacc ggtacaccat	60
gaccgcgcgt cctcggtgcc ccgcgggtgcg ctctctgtcg cgccagccat accggggaggt	120
gtggccgcgtg gcaaccttttgc tgccgcgcctt gggccccag ggcaggcgcc ttgtgcaacc	180
cggggaccccg aagatctacc gcactttgtt tgcccaatgc ctatgtgcg tgcactgggg	240
ctcacagccctt ccacccgtcc acctttccctt ccaccagggt tcatccctga aagagctgg	300
ggccaggggtt gtgcagagac tctgcgagcg caacgagaga aacgtgttgc cttttggctt	360
tgagctgtt aacgaggccca gagggggccc tcccatggcc ttcaactagta gcgtgcgttag	420
ctacttgcctt aacactgttta ttgagaccctt gcgtgtcagt ggtgcatttgc tgctactgtt	480
gagccgagtg ggcgacgacc tgctggctca cctgctggca cactgtgcctc tttatctct	540
gggtccccccccc agctgtgcctt accagggtgtg tgggtctccc ctgttacaaaa tttgtgcac	600
cacggatatac tggccctctg tgcgtgcctt ttacaggccc acccgaccgg tgggcaggaa	660
tttcaactaac cttaggttct tacaacagat caagagcgtt agtcgcagg aagcaccgaa	720

-continued

accctggcc ttgcatctc gaggtacaaa gaggcatcg agtctcacca gtacaagtgt	780
gccttcagct aagaaggcca gatgttatcc tgtcccaga gtggaggagg gacccacag	840
gcaggtgcta ccaacccat caggcaaatac atgggtgcca agtctgtctc ggtccccga	900
ggtgtctact gcagagaaag atttgtttc taaaggaaag gtgtctgacc tgagtctctc	960
tgggtcggtg tgctgtaaac acaagcccag ctccacatct ctgtgtcac caccggcca	1020
aatgccttt cagctcaggc catttattga gaccagacat ttcccttaact ccaggggaga	1080
tggccaagag cgtctaaacc cctcattcct actcagcaac ctccagccata ctttgactgg	1140
ggccaggaga ctggtgaga tcatcttct gggctcaagg cctaggacat caggaccact	1200
ctgcaggaca caccgtctat cgccgtcgata ctggcagatg cggccctgtt tccaaagct	1260
gctggtaac catgcagagt gccaatatgt cagactcctc aggtcacatt gcaggttcg	1320
aacagcaaac caacaggtga cagatgcctt gaacaccaggc ccaccgcacc tcattggattt	1380
gctccgcctg cacagcagtc cctggcaggt atatggttt cttccggcct gtctctgcaa	1440
ggtgtgtct gtagtctct ggggtaccag gcacaatgag cggccgttct ttaagaactt	1500
aaagaagttc atctcggtgg gggaaatacgg caagctatca ctgcaggaac tgatgtggaa	1560
gatgaaagta gaggattgcc actggctccg cagcagccgg gggaaaggacc gtgtccccgc	1620
tgcagagcac cgtctgaggg agaggatctt ggtacgttc ctgttctggc tgatggacac	1680
atacgtggta cagctgctta ggtattttt ttacatcaca gagagcacat tccagaagaa	1740
caggcttcc ttctaccgtt agagtgtgtg gagcaagctg cagagcattt ggtcaggca	1800
acacctttag agagtgccgc tacggagct gtcacaagag gaggtcaggc atcaccaggaa	1860
cacctggcta gccatgccc tctcagact gcttcatc cccaaaggcca acggcgtcg	1920
gcccatgtt aacatgagtt atacatggg taccagact ttggcagaa ggaaggcaggc	1980
ccagcatttc acccagcgtc tcaagactctt cttcagcatg ctcaactatg agcggacaaa	2040
acatcctcac cttatgggtt cttctgtact gggtatgaat gacatctaca ggacctggcg	2100
ggccttggta ctgcgtgtgc gtgtctggta ccagacaccc aggatgtact ttgttaaggc	2160
agatgtgacc gggccatgt atgecatccc ccaggtaag ctgggtggagg ttgttgccaa	2220
tatgatcagg cactcggtt gcaacgtactg tatccggccatg tatgcgttgg tccggagaga	2280
tagccaaggc caagtccaca agtcttttag gagacaggctt accaccctt ctgacctcca	2340
gccatacatg ggccagttcc ttaagcatct gcaggattca gatgccagtg cactgaggaa	2400
ctccgttgc atcgagcaga gcatctctat gaatgagacg agcagcagcc tgtttactt	2460
ctccctgcac ttcctgcgtc acatgtcgta aaagattggt gacagggtct atacgcgtg	2520
ccagggcatc ccccaaggctt ccagcctatc caccctgtctc tgcaagtctgtt gtttggaga	2580
catggagaac aagctgtttt ctgagggtgca gggggatgggg ttgttttac gttttgttga	2640
tgactttctg ttggtgacgc ctcacttggaa ccaagcaaaa accttcctca gcaccctgg	2700
ccatggcggtt ctttgcgtatg ggtgcgtatg aaacttgcag aagacagtgg tgaacttccc	2760
tgtggagctt ggtaccctgg gtggtgccgc tccataccag ctgcctgtctc actgcctgtt	2820
tccctgggtt ggcttgcgtc tggacactca gactttggag gtgttctgtg actactcagg	2880
ttatgcccac acctcaatta agacgacgc tacccttccat agtgcgttca aagctggaa	2940
gaccatgcgg aacaagctcc tgcgttgcgtt gcggttgaag tgcacggc tatttctaga	3000

-continued

cttgcagggt aacagccccc agacagtctg catcaataata tacaagatct tcctgcttca	3060
ggcctacagg ttccatgcat gtgtgattca gttccctt gaccagcgtg ttagaaagaa	3120
cctcacattt tttctggcca tcatctccag ccaagacatcc tgctgtatcg ctatctggaa	3180
ggtaagaat ccaggaatga cactaaaggc ctctggctcc tttccctctg aagccgcaca	3240
ttggctctgc taccaggcct tcctgctcaa gctggctgtcattctgtca tctacaaaatg	3300
tctcctggga cctctgagga cagccccaaa actgctgtc cgaaagctcc cagaggcgcac	3360
aatgaccatc cttaaagctg cagctgaccc agccoctaagc acagactttc agaccattt	3420
ggactaacaa ttggctcgct ttcttgctgt ccaatttcta ttaaaggttc cttttgttcc	3480
ctaagtccaa ctactaaact gggggatatt atgaagggcc ttgagcatct ggattctgcc	3540
taataaaaaaa catttctttt cattgcgaat tcaaaaaaaaaaaaaaaa aaaaaaaaaaaa	3600
aaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa	3660
aaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa	3720
aaa	3723

```
<210> SEQ ID NO 38
<211> LENGTH: 3601
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic
```

<400> SEQUENCE: 38

aggaaataag agagaaaaga agagtaagaa gaaatataag agccaccatg acccgcgctc	60
ctcggtgccg cgccggcgcgc tctctgtgc gcagccgata cggggagggtg tggccgcgtgg	120
caaccttgc gggcccgagg gcaggccgt tgcgtcaaccc ggggaccgaa	180
agatctaccg cactttgggt gcccaatgcc tagtgcgtcat gcactggggc tcacagcctc	240
cacccgtccg ccttccttc caccagggtgt catccctgaa agagctggtg gccagggttg	300
tgcagagact ctgcgagcgc aacgagagaa acgtgcgtgc ttttggctt gagctgccta	360
acgaggccag aggccggcgt cccatggct tcacttagtag cgtgcgtagc tacttgcaca	420
acactgttat tgagaccctg cgtgtcagtgt gtgcattgtt gctactgttg agccaggtgg	480
gcgacgacccgt gctggctcac ctgcgtggcac actgtgcgtct ttatcttgcgttgc	540
gtgtgtgcata ccagggtgtgt gggctcccc tgcgtccaaat ttgtgccacc acggatatct	600
ggccctctgt gtccgcgttgt tacaggccca cccgaccctgt gggcaggaat ttcactaacc	660
ttaggttctt acaacagatc aagagcagta gtcgcaggaa agcaccgaaa cccctggccct	720
tgccatctcg aggtacaaag aggcatctga gtctcaccatc tacaagtgtg ctttcgtacta	780
agaaggccag atgctatcct gtcccgagag tggaggaggacccacacagg caggtgcata	840
caacccatc aggcaaatca tgggtgcaca gtctgtctcg gtcccccggatgtcgtactcg	900
cagagaaaga tttgtcttct aaaggaaagg tgcgtacact gactctcttgcgttgcgtgt	960
gtgttaaaca caagcccgac tccacatctc tgcgtgcacc accccgcacaa aatgccttc	1020
agctcaggcc atttatttgag accagacatt tccttactc cagggaggat ggccaaagac	1080
gtctaaaccc ctcattccca ctcagcaacc tccagcctaa cttgactggg gccaggagac	1140
tgggtggagat catctttctg ggctcaaggc cttaggacatc aggaccactc tgcaggacac	1200

-continued

accgtctatc	gcgtcgatac	tggcagatgc	ggccctgtt	ccaacagctg	ctggtaacc	1260
atgcagagtgc	ccaatatgtc	agactcctca	ggtcacattg	caggttcga	acagcaaacc	1320
aacagggtgac	agatgccttg	aacaccagcc	caccgcacct	catggatttgc	ctccgcctgc	1380
acagcagtc	ctggcaggta	tatggtttc	ttcgggcctg	tctctgcaag	gtggtgtctg	1440
ctagtctctg	gggttaccagg	cacaatgago	gcccgttctt	taagaactta	aagaagttca	1500
tctcggtggg	gaaatacggc	aagctatcac	tgcaggaact	gatgtggaag	atgaaaatgt	1560
aggattgc	ctggctccgc	agcagccccg	ggaaggaccc	tgtccccgt	gcagagcacc	1620
gtctgaggga	gaggatcctg	gctacgttcc	tgttctggct	gatggacaca	tacgtggtag	1680
agctgcttag	gtcattcttt	tacatcacag	agagcacatt	ccagaagaac	aggcttctt	1740
tctaccgtaa	gagtgtgtgg	agcaagctgc	agagcattgg	agtcaggccaa	cacccgtgaga	1800
gagtgccgc	acggggactg	tcacaagagg	aggtcaggcc	tcaccaggac	acctggctag	1860
ccatgccat	ctgcagactg	cgcttcatcc	ccaagccaa	cggcctgcgg	cccattgtga	1920
acatgagtt	tagcatgggt	accagagtt	tggcagaag	gaagcaggcc	cagcatttca	1980
cccagcgtct	caagactctc	ttcagcatgc	tcaactatga	gcccacaaaa	catcctcacc	2040
ttatggggct	ttctgtactg	ggttatgaatg	acatctacag	gacccggcc	gcctttgtgc	2100
tgcgtgtgc	tgctctggac	cagacaccca	ggatgtactt	tgttaaggca	gatgtgaccg	2160
ggccctatga	tgccatcccc	cagggtaago	tggtgagggt	tgttgcera	atgatcaggc	2220
actcggagag	cacgtactgt	atccggccat	atgcagtggt	ccggagagat	agccaaggcc	2280
aagtccacaa	gtcctttagg	agacaggctc	ccaccctctc	tgacccctccag	ccatacatgg	2340
gccagttcct	taagcatctg	caggattcag	atgccagtc	actgaggaac	tccgttgtca	2400
tcgagcagag	catctctatg	aatgagagca	gcagcagect	gtttgacttc	ttctgtcact	2460
tcctgcgtca	cagtgtcgta	aagattgggt	acaggtgcta	tacgcagtc	cagggcatcc	2520
cccagggtct	cagcctatcc	accctgtct	gcagtcgtgt	tttcggagac	atggagaaca	2580
agctgtttgc	tgaggtgcag	cggatgggt	tgcctttacg	ttttgttgc	gactttctgt	2640
tggtaegcc	tcacttggac	caagaaaaaa	ccttcctcag	caccctggtc	catggcggtc	2700
ctgagtatgg	gtgcata	aacttgcaga	agacagtgg	gaacttccct	gtggagctg	2760
gtaccctggg	tggtgcagct	ccataccago	tgcctgtca	ctgcctgtt	ccctgggtgt	2820
gcttgctgct	ggacactcag	acttggagg	tgttctgt	ctactcagg	tatgcccaga	2880
cctcaattaa	gacgagcc	acccctcaga	gtgtcttca	agctggaaag	accatcgcca	2940
acaagctct	gtcggcttc	cggtgaat	gtcacggct	atttctagac	ttgcaggtga	3000
acagcctcca	gacagtctgc	atcaatata	acaagatctt	cctgcttcag	gcctacaggt	3060
tccatgcatg	tgtgattcag	cttcccttgc	accagcgtgt	taggaagaac	ctcacattct	3120
ttctggcat	catctccagc	caagcatct	gctgctatgc	tatcctgaag	gtcaagaatc	3180
caggaatgac	actaaaggcc	tctggctct	ttctctctga	agccgcacat	tggctctgt	3240
accaggcctt	cctgctcaag	ctggctgtc	attctgtcat	ctacaaatgt	ctccctggac	3300
ctctgaggac	agccccaaaa	ctgctgtgc	ggaagctccc	agaggcgaca	atgaccatcc	3360
ttaaagctgc	agctgaccca	gccctaagca	cagacttca	gaccatttg	gactaataat	3420
taagctgc	tctgcggggc	ttgccttc	gccatgc	tcttcctcc	cttgcaccc	3480

-continued

tacctcttgg	tctttgaata	aagcctgagt	aggaagtcta	aaaaaaaaaaa	aaaaaaaaaaa	3540
aaaaaaaaaaa	aaaaaaaaaaa	aaaaaaaaaaa	aaaaaaaaaaa	aaaaaaaaaaa	aaaaaaaaaaa	3600
a						3601
<210> SEQ_ID NO 39						
<211> LENGTH: 3753						
<212> TYPE: DNA						
<213> ORGANISM: Artificial						
<220> FEATURE:						
<223> OTHER INFORMATION: Synthetic						
<400> SEQUENCE: 39						
gggacatttg	cttctgacac	aactgtgttc	actagcaacc	tcaaacaact	agtacaccat	60
gccgcgegct	ccccgctgcc	gagccgtgctg	cgcagccact	accgcgaggt		120
gctgcccgtc	gcaacgttgc	tgccggcgct	ggggcccaag	ggctggcgcc	tggtgtcagcg	180
cggggacccg	gccccgttcc	gcgcgcgtgg	ggcccaagtgc	ctgggtgtgc	tgccctggga	240
cgcacggccg	ccccccgcgg	ccccctccct	ccgcagggtg	tccgtcgtga	aggagctgg	300
ggcccgagtg	ctgcagaggc	tgtgcgagcg	cgccgcgaaag	aacgtgtgg	ccttcggctt	360
cgcgcgtctg	gacggggccc	gcggggggcc	ccccaggcc	ttcaccacca	gcgtgcgcag	420
ctacctgccc	aacacgggtg	ccgacgcact	gcggggggagc	ggggcgtggg	ggctgtctgt	480
gcgcgcgtgc	ggcgcacgcg	tgctgggtca	cctgctggc	cgctgcgcgc	tctttgtgt	540
gggtggctccc	agctgcgcct	accagggtgt	cgggccgcgc	ctgtaccagc	tcggcgctgc	600
cactcaggcc	cggccccccgc	cacacgctag	tggaccccg	aggcgtctgg	gtgtgtgt	660
ggccttggAAC	catagegtca	gggaggccgg	ggtccccctg	ggcctgcacag	ccccgggtgc	720
gaggaggccgc	ggggcgttg	ccagccgaag	tctgcgttgc	cccaagaggc	ccaggcgtgg	780
cgtgcgcct	gagccggagc	ggacgcgcgt	tgggcagggg	tccctggccc	acccggcag	840
gacgcgtggA	ccaggtgacc	gtgggttctg	tgtgggtca	cctgccagac	ccgcgaaga	900
agccacccct	ttggagggtg	cgctctctgg	cacgcgcac	tcccaaccat	ccgtggcccg	960
ccagcaacc	cgggcccccc	catecacato	gcggccacca	cgccctggg	acacgcctt	1020
tccccgggt	tacggcgaga	ccaagcactt	cctctactcc	ttaggggaca	aggaggagct	1080
cgccccctcc	tccctactca	gctctctgag	gcccagcctg	actggcgctc	ggaggctcgt	1140
ggagaccatc	tttctgggtt	ccaggccctg	gatgccaggg	actccccgca	ggttggcccg	1200
cctgccccag	cgctactggc	aatgcggcc	cctgtttctg	gagctgtctg	ggaaccacgc	1260
gcagtgcggcc	tacgggggtgc	tcctcaagac	gcactgcgcg	ctgcgagctg	cggtcacccc	1320
agcagccgggt	gtctgtgccc	gggagaagcc	ccagggtct	gtggcgccccc	ccgaggaggaa	1380
ggacacagac	ccccgtcgcc	tggtgcagct	gctccgcac	cacaggcagcc	cctggcaggt	1440
gtacggcttc	gtgcggccct	gcctgcgcgc	gctggtgcctc	ccaggcctct	ggggctccag	1500
gcacaacgaa	cgccgcgtcc	tcaggaacac	caagaagttc	atctccctgg	ggaagcatgc	1560
caagctctcg	ctgcaggagc	tgacgtggaa	gatgagcgtg	cgggactgcg	cttggctgcg	1620
caggagccca	ggggttggct	gtgttccggc	cgcagagcac	cgtctgcgtg	aggagatcct	1680
ggccaagttc	ctgcactggc	tgtatgagtgt	gtacgtcgtc	gagctgtca	ggtctttctt	1740
ttatgtcactg	gagaccacgt	ttcaaaaagaa	caggctttt	ttctaccgga	agagtgtctg	1800

-continued

gagcaaggttg caaaggcattt gaatcagaca gcacttgaag agggtgcagc tgcggggagct	1860
gtcggaagca gaggtcaggc agcatcgaa agccaggccc gcccgtgtga cgtccagact	1920
ccgcttcatac cccaaggctg acgggctgcg gccgattgtg aacatggact acgtcggtgg	1980
agccagaacg ttccgcagag aaaagagggc cgagcgtctc acctcgaggg tgaaggcact	2040
gttcagegtg ctcaactacg agcggggcgcg ggcggccggc ctccctggcg cctctgtgct	2100
gggcctggac gataccaca gggcctggcg cacccgtgtg ctgcgtgtgc gggcccaagga	2160
cccgccgcct gagctgtact ttgtcaaggt ggatgtgacg ggccgcgtacg acaccatccc	2220
ccaggacagg ctcacggagg tcatacgccag catcatcaa cccagaaca cgtactgcgt	2280
gctcggttat ggcgtggcc agaaggcccgc ccatgggcac gtccgcgaagg ccttcaagag	2340
ccacgtctct accttgacag acctccagcc gtacatgcga cagttcggtt ctcacctgca	2400
ggagaccaggc cccgctgaggg atgcgtcgat catcgacag agctccccc tgaatgaggg	2460
cagcagtggc ctcttcgacg tcttccatcg ctccatgtgc caccacgcgg tgccgtatcg	2520
ggcaagttcc tacgttcaggc ggcaggggat cccgcaggggc tccatctct ccacgtgtct	2580
ctgcagccctg tgctacggcg acatggagaa caagctgtt gggggatc ggccggacgg	2640
gctgctctcg cggttgggtgg atgatttctt gttggtaaca cctcacctca cccacgcgaa	2700
aaccttccatc aggaccctgg tccgagggtt ccctgagttt ggctgtgttgg tgaacttgcg	2760
gaagacatgtg tgaaacttcc ctgtagaaga cgaggccctg ggtggacagg ctttgttca	2820
gatgccggcc cacggeatat tccccctggc cggcctgtgtg ctggataccc ggaccctgg	2880
ggtgccggcgc gactactcca gctatggccg gacccatccat agagccatgc tcaccccaa	2940
cccgccgttc aaggctgggaa ggaacatgcg tcgcaaaactc tttgggtct tgccgtgaa	3000
gtgtcacatgc ctgtttctgg atttgcaggtaa gaaacgcctc cagacgggtt gcaacacat	3060
ctacaagatc ctccctgtgc aggccgtacag gtttccacgc tttgtgtgtc agctccatt	3120
tcatcagcaaa gtttggaaaga accccacatt tttccctgcgc gtcatctctg acacggctc	3180
cctctgttac tccatccatc aagccaaagaa cgcaggatgtt tggctgggg ccaaggccgc	3240
cgccggccctt ctggccctccg aggcgtgtca gtggctgtgc caccaaggcat tcctgtcaa	3300
gctgacttgcg caccgtgtca cctacgtgc actccctgggg tcaactcaggaa cagcccagac	3360
cgagctgttgtt cggaaaggctcc cggggacgcac gctgacttgcg ctggaggccgc cagccaaaccc	3420
ggcactgccc tcagacttca agaccatctt ggactgacaa ttggctcgct ttcttgcgtt	3480
ccaaatttcta taaaagggttc ctttgttcc ctaagtccaa ctactaaact gggggatatt	3540
atgaaggggcc ttgagcatct ggattctgc taataaaaaaa cattttttt cattgcgtat	3600
taaaaaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa	3660
aaaaaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa	3720
aaaaaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa	3753

<210> SEQ ID NO 40
<211> LENGTH: 3665
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 40

-continued

agggaaaataag agagaaaaga agagtaagaa gaaatataag agccaccatg ccgcgcgctc	60
cccgctgcg agccgtgcgc tccctgtgc gcagccacta ccgcgagggt ctggcgctgg	120
ccacgttctgt gggggcgcctg gggccccagg gctggcgctg ggtgcagcgc ggggaccgg	180
cggctttccg cgccgtggtg gcccagtgtcc tggtgtgcgt gcccctggac gcacggcgc	240
ccccccgcgc ccccttccttc cgccagggtgt cctgcctgaa ggagctggtg gcccaggatgc	300
tgcagaggct gtgcgagcgc ggcgcgaaga aegtgtggc ttccggatc gcgtgtgtgg	360
acggggccgc cggggggccccc cccggggct tcaccaccatc cgtgcgcagc tacctgcca	420
acacgggtgac cgacgcactg cggggggageg gggcggtgggg gctgctgtcg cgccgcgtgg	480
gcgcgcgcgt gctgggtcac ctgcgtggcac gctgcgcgcgt ctttgtgtcg gtggctccca	540
gctgcgccta ccagggtgtgc gggccgcgcg tgtaccagct cggcgctgcc actcaggccc	600
ggccccccgc acacgttagt ggaccccgaa ggctgtctggg atgcgaacgg gcttggaaacc	660
atagcgtcag ggaggccggg gtcccccctgg gcctgccagc cccgggtgcg aggaggcgcg	720
ggggcagtgc cagccgaagt ctgcgtgtgc ccaagaggcc caggcggtggc gctgcctctg	780
agccggagcg gacgcgggtt gggcagggggt cctgggcaca cccggggagg acgcgtggac	840
cgagtgaccg tggtttctgt gtgggtcac ctgcagacc cggcgaagaa gccaccttt	900
tggagggtgc gctctctggc acgcgcact cccacccatc cgtgggcgc cagcaccacg	960
cgggcccccc atccacatcg cggccaccac gtccctggg cacgccttgt ccccccgtgt	1020
acgcccagac caagcacttc ctctactctt caggcgacaa ggagcagctg cggccctctt	1080
tcctacttag ctctctgagg cccagctgtc ctggcgctcg gaggctcgag gagaccatct	1140
ttctgggttc caggccctgg atgcaggaa ctccccgeag gttggccgc ctgccccagc	1200
gtactggca aatgcggccc ctgtttctgg agctgtttgg gaaccacgcg cagtgcctt	1260
acgggggtct cctcaagacg cactgcccgc tgccgcgtgc ggtcaccacca gcagccgcgt	1320
tctgtgecccg ggagaagcccc cagggtctgt tggggcccccc cgaggaggag gacacagacc	1380
cccggtcgct ggtcagctg ctccgcagc acagcagccc ctggcagggt tacggctcg	1440
tgcgggectg cctgcgcgg ctgggtcccc caggcctctg gggctecagg cacaacgaa	1500
gccgcttctt caggaacacc aagaagttca tctccctggg gaagcatgcc aagctctcg	1560
tgcaggagct gacgttggaaat atgcgtgc gggactgcgc ttggctgcgc aggagccag	1620
gggtttggctg tggccggcc gcagagcacc gtctgcgtga ggagatcctg gccaagtcc	1680
tgcactggct gatgatgtt tacgtcgatc agctgttcag gtctttcttt tatgtcacgg	1740
agaccacgtt taaaaaaac aggtctttt tctaccggaa gagttgtctgg agcaagtgc	1800
aaagcattgg aatcagacag cacttgaaga ggggtcgatc gcggggagctg tcggaagcag	1860
aggtcaggca gcatcggaa gccaggcccc ccctgcgtac gtccagactc cgcttcatcc	1920
ccaagcctga cgggcgtgcgg ccgattgtga acatggacta cgtcggtggaa gccagaacgt	1980
tccgcagaga aaagaggccc gagegtctca cctcgagggtt gaaggcactg ttccagcgtgc	2040
tcaactacga gccccggcgg ccgcggcc tccctggccgc ctctgtgtcg ggcctggacg	2100
atatccacag ggcctggcgc accttcgtgc tgcgtgtgc gggccaggac ccggccctcg	2160
agctgtactt tgtcaagggtt gatgtgacgg ggcgtacga caccatcccc caggacaggc	2220
tcacggaggt catgcgcagc atcatcaaacc cccagaacac gtactgcgtgc cgtcggtatg	2280

-continued

ccgtggtcca gaaggccgccc catgggcacg tccgcaaggc cttcaagagc cacgtctcta	2340
ccttgacaga cctccagccg tacatgcgac agttcggtgc tcacactgcag gagaccagcc	2400
cgctgaggga tgccgtcgtc atcgagcaga gctccctccct gaatgaggcc agcagtggcc	2460
tcttcgacgt cttccctacgc ttcatgtgcc accacgcccgt ggcgcattcagg ggcaagtccct	2520
acgtccagtgc ccaagggtatc ccgcagggtatc ccatcccttc caccgtgttc tgcagccgt	2580
gtacggcga catggagaac aagctgtttt cggggattcg cggggacggg ctgtccctgc	2640
gtttgggtggaa tgatttcttg ttgggtgacac ctcacccctac ccacgcgaaa accttcctca	2700
ggaccctgggt ccgaggtgtc cctgagttatg gctgcgtgggt gaacttgcgg aagacagtgg	2760
tgaacttcccc ttgttagaagac gaggccctgg gtggcacggc ttttggttcaag atgcggggccc	2820
acggccctatt cccctgggtgc ggcctgctgc tggatacccg gaccctggag gtgcagagcg	2880
actactccatcg ctatgcggcc acctccatca gagccagtctc cacccttcaac cgcggcttca	2940
aggctgggag gaacatgcgt cgcaaactct ttgggggttgc ggggctgaag tgtcacagcc	3000
tgtttctggaa ttgcagggtg aacagccctcc agacgggtgtg caccaacatc tacaagatcc	3060
tccctgctgca ggcgtacagg tttcacgcatt gtgtgctgca gtccttcatc catcagaag	3120
tttggaaagaa ccccacattt ttccctgcggc tcatctctgc cacggcctcc ctctgctact	3180
ccatcctgaa agccaagaac gcagggatgt cgctgggggc caagggcggcc gcccggccctc	3240
tgcctcctgaa ggcgtgcag tggctgtgc accaaggcatt cctgctcaag ctgactcgac	3300
accgtgtcac ctacgtgcca ctccctgggt cactcaggac agcccagacg cagctgagtc	3360
ggaaagctccc ggggacgacg ctgactgccc tggaggccgc agccaaacccg gcactgcct	3420
cagacttcaa gaccatcctg gactgataat taagctgcct tctgcggggc ttgccttctg	3480
ccatgcctc ttttctctcc cttgcacctg tacctcttgg tctttgaata aagctgagt	3540
aggaagaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	3600
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	3660
aaaaaa	3665

1.114. (canceled)

115. A composition comprising (i) a ribonucleic acid (RNA) encoding telomerase reverse transcriptase (TERT) and (ii) a lipid nanoparticle (LNP), wherein the LNP comprises:

- (a) a phospholipid at a molar ratio of between at or about 1 to 20 moles versus total moles of lipid in the LNP;
- (b) a PEGylated lipid at a molar ratio of between at or about 0.1 to 3 moles versus total moles of lipid in the LNP;
- (c) a cholesterol lipid at a molar ratio of between at or about 20 to 60 moles versus total moles of lipid in the LNP; and
- (d) an ionizable lipid comprising SS-OP or an SS-OP analog at a molar ratio of between at or about 30 to 70 moles versus total moles of lipid in the LNP;
wherein the liposomal pKa of the LNP is between at or about 5.5 to 7.2; and
wherein the RNA comprises one or more modified nucleotides and is encapsulated in the LNP.

116. The composition of claim 115, wherein the phospholipid, the PEGylated lipid, the cholesterol, and the ionizable lipid comprise a total lipid weight of the composition, and wherein the ratio of the total lipid weight to total RNA weight in the composition is between at or about 14 to 42.

izable lipid comprise a total lipid weight of the composition, and wherein the ratio of the total lipid weight to total RNA weight in the composition is between at or about 14 to 42.

117. The composition of claim 115, wherein the phospholipid, the PEGylated lipid, the cholesterol, and the ionizable lipid comprise a total lipid weight of the composition, and wherein the ratio of the total lipid weight to total RNA weight in the composition is between at or about 175 to 25.

118. The composition of claim 115, wherein the phospholipid, the PEGylated lipid, the cholesterol, and the ionizable lipid comprise a total lipid weight of the composition, and wherein the ratio of the total lipid weight to total RNA weight in the composition is at or about 42.

119. The composition of claim 115, wherein the phospholipid is included in the LNP at a molar ratio of between at or about 4 to 6 moles versus total moles of lipid in the LNP; the PEGylated lipid is included in the LNP at a molar ratio of between at or about 1 to 2 moles versus total moles of lipid in the LNP; the cholesterol lipid is included in the LNP at a molar ratio of between at or about 35 to 45 moles versus total moles of lipid in the LNP; and the ionizable lipid

is included in the LNP at a molar ratio of between at or about 50 to 60 moles versus total moles of lipid in the LNP.

120. The composition of claim 115, wherein the phospholipid is included in the LNP at a molar ratio of at or about 7.5 moles versus total moles of lipid in the LNP; the PEGylated lipid is included in the LNP at a molar ratio of at or about 1.5 moles versus total moles of lipid in the LNP; the cholesterol lipid is included in the LNP at a molar ratio of at or about 40 moles versus total moles of lipid in the LNP; and the ionizable lipid is included in the LNP at a molar ratio of at or about 52.5 moles versus total moles of lipid in the LNP.

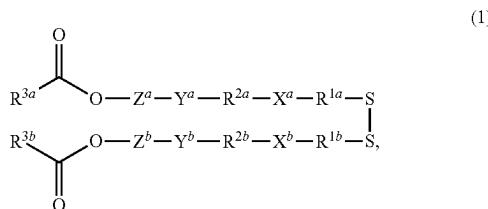
121. The composition of claim 115, wherein the phospholipid is included in the LNP at a molar ratio of between at or about 14 to 18 moles versus total moles of lipid in the LNP; the PEGylated lipid is included in the LNP at a molar ratio of between at or about 1 to 3 moles versus total moles of lipid in the LNP; the cholesterol lipid is included in the LNP at a molar ratio of between at or about 40 to 50 moles versus total moles of lipid in the LNP; and the ionizable lipid is included in the LNP at a molar ratio of between at or about 30 to 40 moles versus total moles of lipid in the LNP.

122. The composition of claim 115, wherein the pKa of the LNP is between 6.0 to 6.8.

123. The composition of claim 115, further comprising a targeting moiety operably integrated in or attached to the LNP.

124. The composition of claim 123, wherein the composition is adapted to specifically or selectively interact with a liver cell.

125. The composition of claim 115, wherein the SS-OP analog comprises a compound of Formula I:



wherein R^{1a} and R^{1b} each independently represents an alkylene group having 1 to 6 carbon atoms, wherein X^a and X^b are each independently an acyclic alkyl tertiary amino group having 1 to 6 carbon atoms and 1 tertiary amino group, or 2 to 5 carbon atoms, and A cyclic alkylene tertiary amino group having 1 to 2 tertiary amino groups, wherein R^{2a} and R^{2b} each independently represent an alkylene group having 8 or less carbon atoms or an oxydialkylene group, wherein Y^a and Y^b each independently represent an ester bond, an amide bond, a carbamate bond, an ether bond or a urea bond; wherein Z^a and Z^b are each independently a divalent group derived from an aromatic compound having 3 to 16 carbon atoms, having at least one aromatic ring, and optionally having a hetero atom, and wherein R^{3a} and R^{3b} each independently represent a residue derived from a reaction product of a fat-soluble vitamin having a hydroxyl group and succinic anhydride or glutaric anhydride, or a sterol derivative hav-

ing a hydroxyl group and succinic anhydride or a residue derived from a reaction product with glutaric anhydride or an aliphatic hydrocarbon group having 12 to 22 carbon atoms.

126. The composition of claim 115, wherein the RNA comprises a sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99%, or 100% identical to any one of SEQ ID NOS: 1-5, 7, 9, 14-17, 19, 21, 23, 25, 27, 29-31, 37-40.

127. The composition of claim 115, wherein the one or more modified nucleotides of the RNA of (i) comprise one or more of a modified adenine or analog thereof, a modified cytidine or analog thereof, a modified guanosine or analog thereof, and a modified uridine or analog thereof, one or more of 1-methylpseudouridine, pseudouridine, 2-thiouridine, and 5-methylcytidine, 5-methoxyuridine (5-moU), and/or one or more of m1A 1-methyladenosine, m6A N6-methyladenosine, Am 2'-O-methyladenosine, i6A N6-isopentenyladenosine, i6A N6-(cis-hydroxyisopentenyl)adenosine, ms2i6A 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine, g6A N6-glycinylcarbamoyladenosine, t6A N6-threonylcarbamoyladenosine, ms2t6A 2-methylthio-N6-threonyl carbamoyladenosine, Ar(p) 2'-O-ribosyladenosine (phosphate), m6 2A N6,N6-dimethyladenosine, m6Am N6,2'-O-dimethyladenosine, m6 2Am N6,N6,2'-O-trimethyladenosine, m1Am 1,2'-O-dimethyladenosine, m3C 3-methylcytidine, m5C 5-methylcytidine, Cm 2'-O-methylcytidine, ac4C N4-acetylcytidine, f5C 5-formylcytidine, m4C N4-methylcytidine, hm5C 5-hydroxymethylcytidine, f5Cm 5-formyl-2'-O-methylcytidine, m1G 1-methylguanosine, m2G N2-methylguanosine, m7G 7-methylguanosine, Gm 2'-O-methylguanosine, m2 2G N2,N2-dimethylguanosine, Gr(p) 2'-O-ribosylguanosine (phosphate), yW wybutsine, o2yW peroxywybutosine, OHyW hydroxywybutosine, OHyW* undermodified hydroxywybutosine, imG wyosine, m2,7G N2,7-dimethylguanosine, m2,2,7G N2,N2,7-trimethylguanosine I inosine, m11 1-methylinosine, Im 2'-O-methylinosine, Q queuosine, galQ galactosyl-queuosine, manQmannosyl-queuosine, W pseudouridine, D dihydrouridine, m5U 5-methyluridine, Um 2'-O-methyluridine, m5Um 5,2'-O-dimethyluridine, m1Ψ 1-methylpseudouridine, Ψm 2'-O-methylpseudouridine, s2U 2-thiouridine, ho5U 5-hydroxyuridine, chm5U 5-(carboxyhydroxymethyl)uridine, mchm5U 5-(carboxyhydroxymethyl)uridine, methyl ester mcm5U 5-methoxycarbonylmethyluridine, mcm5Um 5-methoxycarbonylmethyl-2'-O-methyluridine, mcm5s2U 5-methoxycarbonylmethyl-2-thiouridine, ncm5U 5-carbamoylmethyluridine, ncm5Um 5-carbamoylmethyl-2'-O-methyluridine, cmnm5U 5-carboxymethylaminomethyluridine, m3U 3-methyluridine, m1acp3Ψ 1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine, cm5U 5-carboxymethyluridine, m3Um 3,2'-O-dimethyluridine, m5D 5-methylidihydrouridine, cm5U 5-taurinomethyluridine, tm5s2U 5-taurinomethyl-2-thiouridine, 2-Aminoadenosine, 2-Amino-6-chloropurineriboside, 8-Azaadenosine, 6-Chloropurineriboside, 5-lodocytidine, 5-lodouridine, Inosine, 2'-O-Methylinosine, Xanthosine, 4-Thiouridine, 06-Methylguanosine, 5,6-Dihydrouridine, 2-Thiocytidine, 6-Azacytidine, 6-Azauridine, 2'-O-Methyl-2-aminoadenosine, 2'-O-Methylpseudouridine, N1-Methyladenosine, 2'-O-Methyl-5-methyluridine, 7-Deazaguanosine, 8-Azidoadenosine, 5-Bromocytidine, 5-Bromouridine, 7-Deazaadenosine, 5-Aminoallyluridine, 5-Aminoallylcytidine, 8-Oxoguanosine, 2-Aminopurine-riboside, Pseudoiso-

cytidine, N1-Methylpseudouridine, 5,6-Dihydro-5-Methyluridine, N6-Methyl-2-Aminoadenosine, 5-Carboxycytidine, 5-Hydroxymethyluridine, Thienoguanosine, 5-Hydroxycytidine, 5-Formyluridine, 5-Carboxyuridine, 5-Methoxyuridine, 5-Methoxycytidine, Thienouridine, 5-Carboxymethylesteruridine, Thienocytidine, 8-Oxaadenosine, Isoguanosine, N1-Ethylpseudouridine, N1-Methyl-2'-O-Methylpseudouridine, N1-Methoxymethylpseudouridine, N1-Propylpseudouridine, 2'-O-Methyl-N6-Methyladenosine, 2-Amino-6-Cl-purine-2'-deoxyriboside, 2-Amino-2'-deoxyadenosine, 2-Aminopurine-2'-deoxyriboside, 5-Bromo-2'-deoxycytidine, 5-Bromo-2'-deoxyuridine, 6-Chloropurine-2'-deoxyriboside, 7-Deaza-2'-deoxyadenosine, 7-Deaza-2'-deoxyguanosine, 2'-Deoxyinosine, 5-Propynyl-2'-deoxycytidine, 5-Propynyl-2'-deoxyuridine, 5-Fluoro-2'-deoxyuridine, 5-Iodo-2'-deoxycytidine, 5-Iodo-2'-deoxyuridine, N6-Methyl-2'-deoxyadenosine, 5-Methyl-2'-deoxycytidine, 06-Methyl-2'-deoxyguanosine, N2-Methyl-2'-deoxyguanosine, 8-Oxo-2'-deoxyadenosine, 8-Oxo-2'-deoxyguanosine, 2-Thiothymidine, 2'-Deoxy-P-nucleoside, 5-Hydroxy-2'-deoxycytidine, 4-Thiothymidine, 2-Thio-2'-deoxycytidine, 6-Aza-2'-deoxyuridine, 6-Thio-2'-deoxyguanosine, 8-Chloro-2'-deoxyadenosine, 5-Aminoallyl-2'-deoxycytidine, 5-Aminoallyl-2'-deoxyuridine, N4-Methyl-2'-deoxycytidine, 2'-Deoxyzebaraline, 5-Hydroxymethyl-2'-deoxyuridine, 5-Hydroxymethyl-2'-deoxycytidine, 5-Propargylamino-2'-deoxyuridine, 5-Carboxy-2'-deoxycytidine, 5-Formyl-2'-deoxycytidine, 5-[3-Indolyl]propionamide-N-allyl]-2'-deoxyuridine, 5-Carboxy-2'-deoxyuridine, 5-Formyl-2'-deoxyuridine, 7-Deaza-7-Propargylamino-2'-deoxyadenosine, 7-Deaza-7-Propargylamino-2'-deoxyguanosine, Biotin-16-Aminoallyl-2'-dUTP, Biotin-16-Aminoallyl-2'-dCTP, Biotin-16-Aminoallylcytidine, N4-Biotin-OBEA-2'-deoxycytidine, Biotin-16-Aminoallyluridine, Dabcyd-5-3-Aminoallyl-2'-dUTP, Desthiobiotin-6-Aminoallyl-2'-deoxycytidine, Desthiobiotin-16-Aminoallyl-Uridine, Biotin-16-7-Deaza-7-Propargylamino-2'-deoxyguanosine, Cyanine 3-5-Propargylamino-2'-deoxycytidine, Cyanine 3-6-Propargylamino-2'-deoxyuridine, Cyanine 5-6-Propargylamino-2'-deoxycytidine, Cyanine 5-6-Propargylamino-2'-deoxyuridine, Cyanine 3-Aminoallylcytidine, Cyanine 3-Aminoallyluridine, Cyanine 5-Aminoallylcytidine, Cyanine 5-Aminoallyluridine, Cyanine 7-Aminoallyluridine, 2'-Fluoro-2'-deoxyadenosine, 2'-Fluoro-2'-deoxycytidine, 2'-Fluoro-2'-deoxyguanosine, 2'-Fluoro-2'-deoxyuridine, 2'-O-Methyladenosine, 2'-O-Methyleytidine, 2'-O-Methylguanosine, 2'-O-Methyluridine, Puromycin, 2'-Amino-2'-deoxycytidine, 2'-Amino-2'-deoxyuridine, 2'-Azido-2'-deoxycytidine, 2'-Azido-2'-deoxyuridine, Aracytidine, Arauridine, 2'-Azido-2'-deoxyadenosine, 2'-Amino-2'-deoxyadenosine, Araadenosine, 2'-Fluoro-thymidine, 3'-O-Methyladenosine, 3'-O-Methylcytidine, 3'-O-Methylguanosine, 3'-O-Methyluridine, 2'-Azido-2'-deoxyguanosine, Araguanosine, 2'-Deoxyuridine, 3'-O-(2-nitrobenzyl)-2'-Deoxyadenosine, 3'-O-(2-nitrobenzyl)-2'-Deoxyinosine, 3'-Deoxyadenosine, 3'-Deoxyguanosine, 3'-Deoxycytidine, 3'-Deoxy-5-Methyluridine, 3'-Deoxyuridine, 2',3'-Dideoxyadenosine, 2',3'-Dideoxyguanosine,

Dideoxyuridine, 2',3'-Dideoxythymidine, 2',3'-Dideoxycytidine, 3'-Azido-2',3'-dideoxyadenosine, 3'-Azido-2',3'-dideoxythymidine, 3'-Amino-2',3'-dideoxycytidine, 3'-Amino-2',3'-dideoxyguanosine, 3'-Amino-2',3'-dideoxythymidine, 3'-Azido-2',3'-dideoxycytidine, 3'-Azido-2',3'-dideoxyuridine, 5-Bromo-2',3'-dideoxyuridine, 2',3'-Dideoxyinosine, 2'-Deoxyadenosine-5'-O-(1-Thiophosphate), 2'-Deoxyguanosine-5'-O-(1-Thiophosphate), 2'-Deoxythymidine-5'-O-(1-Thiophosphate), Adenosine-5'-O-(1-Thiophosphate), Cytidine-5'-O-(1-Thiophosphate), Guanosine-5'-O-(1-Thiophosphate), Uridine-5'-O-(1-Thiophosphate), 2',3'-Dideoxyadenosine-5'-O-(1-Thiophosphate), 2',3'-Dideoxycytidine-5'-O-(1-Thiophosphate), 2',3'-Dideoxyguanosine-5'-O-(1-Thiophosphate), 3'-Deoxythymidine-5'-O-(1-Thiophosphate), 3'-Azido-2',3'-dideoxythymidine-5'-O-(1-Thiophosphate), 2',3'-Dideoxyuridine-5'-O-(1-Thiophosphate), 2'-Deoxyadenosine-5'-O-(1-Boranophosphate), 2'-Deoxycytidine-5'-O-(1-Boranophosphate), 2'-Deoxyguanosine-5'-O-(1-Boranophosphate), and 2'-Deoxythymidine-5'-O-(1-Boranophosphate).

128. The composition of claim 115, wherein the composition further comprises a ribonucleic acid (RNA) encoding Telomerase RNA Component (TERC).

129. The composition of claim 115, wherein the RNA comprises a self-replicating RNA or a circular RNA.

130. A cell comprising the composition of claim 126.

131. A method of increasing telomerase activity in a target cell or extending telomeres in the target cell, the method comprising contacting the target cell and the composition of claim 115 and permitting the target cell to uptake the composition.

132. A method of treating a disease or disorder, or delaying the onset of a disease, comprising administering to a subject an effective amount of a composition according to claim 126.

133. A method of treating a disease or disorder, or delaying the onset of a disease, comprising administering to a subject an effective amount of a cell according to claim 130.

134. The method of claim 132, wherein the disease is fibrotic disease or liver disease, and wherein the liver disease is non-alcoholic steatohepatitis (NASH), non-alcoholic fatty liver disease (NAFLD), alcoholic hepatitis, liver cirrhosis, liver fibrosis, compensated cirrhosis, decompensated cirrhosis, acute-on-chronic liver failure, biliary atresia, primary biliary cirrhosis, primary sclerosing cholangitis, chronic liver disease hemochromatosis, Wilson's disease, and/or ischemic hepatitis.

135. The method of claim 133, wherein the disease is fibrotic disease or liver disease, and wherein the liver disease is non-alcoholic steatohepatitis (NASH), non-alcoholic fatty liver disease (NAFLD), alcoholic hepatitis, liver cirrhosis, liver fibrosis, compensated cirrhosis, decompensated cirrhosis, acute-on-chronic liver failure, biliary atresia, primary biliary cirrhosis, primary sclerosing cholangitis, chronic liver disease hemochromatosis, Wilson's disease, and/or ischemic hepatitis.