

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7645929号
(P7645929)

(45)発行日 令和7年3月14日(2025.3.14)

(24)登録日 令和7年3月6日(2025.3.6)

(51)国際特許分類		F I	
A 6 1 K	9/51 (2006.01)	A 6 1 K	9/51 Z N A
A 6 1 K	31/7088(2006.01)	A 6 1 K	31/7088
A 6 1 P	43/00 (2006.01)	A 6 1 P	43/00 1 1 1
A 6 1 K	48/00 (2006.01)	A 6 1 K	48/00
A 6 1 K	47/18 (2017.01)	A 6 1 K	47/18
請求項の数 35 (全414頁) 最終頁に続く			
(21)出願番号	特願2023-82312(P2023-82312)	(73)特許権者	520107630
(22)出願日	令和5年5月18日(2023.5.18)		インテリア セラピューティクス、イン
(62)分割の表示	特願2020-540244(P2020-540244)の分割		コーポレーテッド
原出願日	平成30年9月28日(2018.9.28)		アメリカ合衆国 0 2 1 3 9 マサチュー
(65)公開番号	特開2023-103421(P2023-103421 A)		セッツ州、ケンブリッジ、エリー スト
(43)公開日	令和5年7月26日(2023.7.26)	(74)代理人	110002572
審査請求日	令和5年6月16日(2023.6.16)		弁理士法人平木国際特許事務所
(31)優先権主張番号	62/566,240	(72)発明者	ウッド、クリスティー、エム.
(32)優先日	平成29年9月29日(2017.9.29)		アメリカ合衆国 0 2 4 8 1 マサチュー
(33)優先権主張国・地域又は機関	米国(US)		セッツ州、ウェルスレイ、アール ロード
		(72)発明者	3 4
			ガードナー、ノア、ポール
			アメリカ合衆国 0 1 5 9 0 マサチュー
			セッツ州、サットン、マククレラン 口
			最終頁に続く

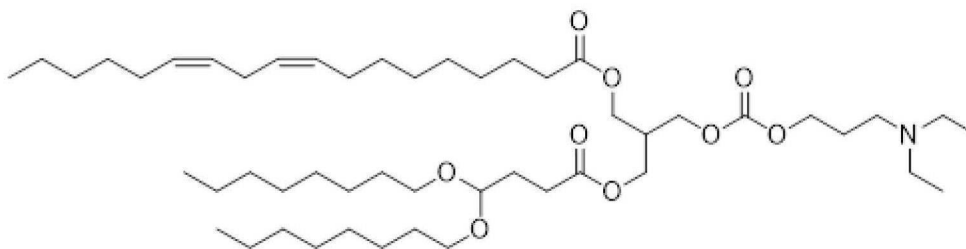
(54)【発明の名称】 製剤

(57)【特許請求の範囲】

【請求項1】

脂質ナノ粒子（「LNP」）組成物であって、
RNA成分であって、（i）RNA誘導型DNA結合因子をコードするmRNA、及び（
ii）gRNA核酸を含む、前記RNA成分と、
脂質成分と
を含み、
ここで、
前記脂質成分は、
40～60モル%のアミン脂質と、
0～10モル%の中性脂質と、
1.5～10モル%のPEG脂質と
を含み、
ここで、前記脂質成分の残部はヘルパー脂質であり、
前記LNP組成物のN／P比は4.5～8であり、
前記アミン脂質は、リピドA又はリピドAのアセタール類似体であり、リピドAは以下の
構造式：

【化 1】



で表される、
前記組成物。

【請求項 2】

前記 RNA 誘導型 DNA 結合因子は Cas ヌクレアーゼである、請求項 1 に記載の組成物。

【請求項 3】

前記 RNA 誘導型 DNA 結合因子はクラス 2 Cas ヌクレアーゼである、請求項 1 または 2 に記載の組成物。

【請求項 4】

前記クラス 2 Cas ヌクレアーゼは Cas 9 ヌクレアーゼである、請求項 3 に記載の組成物。

【請求項 5】

前記 mRNA は、配列番号 1、4、7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または 66 のいずれか 1 つに対する同一性が少なくとも 90% である配列を含み、ここで、前記 mRNA は、前記 RNA 誘導型 DNA 結合因子をコードするオープンリーディングフレームを含む、請求項 1～4 のいずれか一項に記載の組成物。

【請求項 6】

前記 gRNA 核酸は gRNA である、請求項 1～5 のいずれか一項に記載の組成物。

【請求項 7】

前記 gRNA は修飾されている、請求項 6 に記載の組成物。

【請求項 8】

前記 gRNA は、2'-O-メチル (2'-O-Me) 修飾ヌクレオチド、ヌクレオチド間のホスホロチオアート (PS) 結合、及び 2'-フルオロ (2'-F) 修飾ヌクレオチドから選択される修飾を含む、請求項 7 に記載の組成物。

【請求項 9】

前記 RNA 誘導型 DNA 結合因子はクラス 2 Cas ヌクレアーゼであり、前記 gRNA 及びクラス 2 Cas ヌクレアーゼ mRNA は、重量基準で 10:1～1:10 の範囲の比で存在する、請求項 1～8 のいずれか一項に記載の組成物。

【請求項 10】

前記 gRNA 及びクラス 2 Cas ヌクレアーゼ mRNA は、重量基準で 5:1～1:5、3:1～1:1 または 2:1～1:1 の範囲の比で存在する、請求項 9 に記載の組成物。

【請求項 11】

前記 gRNA 及びクラス 2 Cas ヌクレアーゼ mRNA は、重量基準で 2:1、1:1 または 1:2 の比で存在する、請求項 9 または 10 に記載の組成物。

【請求項 12】

少なくとも 1 つの鑄型核酸をさらに含む、請求項 1～11 のいずれか一項に記載の組成物。

【請求項 13】

前記 LNP 組成物の N/P 比は 6 ± 1 である、請求項 1～12 のいずれか一項に記載の組成物。

10

20

30

40

50

【請求項 14】

前記 N / P 比は 6 ± 0.5 である、請求項 13 に記載の組成物。

【請求項 15】

前記 PEG 脂質のモル%は、脂質成分の 2 モル% ~ 10 モル% または脂質成分の 2 モル% ~ 8 モル% または脂質成分の 2 モル% ~ 4 モル% である、請求項 1 ~ 14 のいずれか一項に記載の組成物。

【請求項 16】

前記 PEG 脂質のモル%は、脂質成分の 2.5 モル% である、請求項 15 に記載の組成物。

【請求項 17】

前記 PEG 脂質のモル%は、脂質成分の 1.5 モル% である、請求項 1 ~ 14 のいずれか一項に記載の組成物。

10

【請求項 18】

前記アミン脂質のモル%が、脂質成分の 45 モル% ~ 60 モル%、または 50 モル% ~ 55 モル%、47 ~ 53 モル%、または 48 ~ 53 モル% である、請求項 1 ~ 17 のいずれか一項に記載の組成物。

【請求項 19】

前記アミン脂質のモル%が、脂質成分の 50 モル%、45 モル%、47 モル%、48 モル%、53 モル%、または 55 モル% である、請求項 18 に記載の組成物。

【請求項 20】

前記中性脂質のモル%が、脂質成分の 5 モル% ~ 10 モル%、または 0 モル% ~ 5 モル% である、請求項 1 ~ 19 のいずれか一項に記載の組成物。

20

【請求項 21】

前記中性脂質のモル%が、脂質成分の 9 モル%、5 モル%、6 モル%、7 モル%、8 モル%、または 10 モル% である、請求項 20 に記載の組成物。

【請求項 22】

前記ヘルパー脂質のモル%が、脂質成分の 30 モル% ~ 50 モル%、または 30 モル% ~ 40 モル% である、請求項 1 ~ 21 のいずれか一項に記載の組成物。

【請求項 23】

前記アミン脂質はリピド A である、請求項 1 ~ 22 のいずれか一項に記載の組成物。

【請求項 24】

前記アミン脂質はリピド A の C4 - C12 アセタール類似体である、請求項 1 ~ 22 のいずれか一項に記載の組成物。

30

【請求項 25】

前記ヘルパー脂質はステロイド、ステロールまたはアルキルレソルシノールである、請求項 1 ~ 24 のいずれか一項に記載の組成物。

【請求項 26】

前記ヘルパー脂質はコレステロールである、請求項 25 に記載の組成物。

【請求項 27】

前記中性脂質は中性リン脂質である、請求項 1 ~ 26 のいずれか一項に記載の組成物。

【請求項 28】

前記中性脂質は DSPC または DPPC である、請求項 27 に記載の組成物。

40

【請求項 29】

前記 PEG 脂質はジミリスチルグリセロール (DMG)、PEG-2k、または PEG-DMG を含む、請求項 1 ~ 28 のいずれか一項に記載の組成物。

【請求項 30】

前記 PEG 脂質は PEG-DMG である、請求項 29 に記載の組成物。

【請求項 31】

前記 PEG 脂質は PEG 2k-DMG である、請求項 30 に記載の組成物。

【請求項 32】

細胞を、(a) インビトロでの遺伝子編集または (b) インビトロでの遺伝子操作された

50

細胞の作製における、請求項 1 ～ 3 1 のいずれか一項に記載の組成物の使用。

【請求項 3 3】

治療における使用のための、請求項 1 ～ 3 1 のいずれか一項に記載の組成物。

【請求項 3 4】

前記治療が以下：

(a) 対象における遺伝子編集または

(b) 対象における遺伝子操作された細胞の作製

を含む、請求項 3 3 に記載の組成物。

【請求項 3 5】

前記遺伝子操作された細胞の作製が、少なくとも 1 つの鋳型核酸を細胞に導入することをさらに含む、請求項 3 2 に記載の使用又は請求項 3 4 に記載の組成物。

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、2017年9月29日に出願された米国仮特許出願第62／566，240号の優先権の利益を主張し、その記載内容は、参照によりその全体が本明細書に組み込まれる。

【0002】

本明細書では、生物学的に活性な薬剤、特にRNA、mRNA、及びガイドRNAを送達するための特性が改善された脂質ナノ粒子（「LNP」）組成物を提供する。LNP組成物は、細胞膜を通過するRNA剤の送達を容易にし、特定の実施形態では、かかる組成物により、遺伝子編集用の成分及び組成物を生細胞内に導入する。

【背景技術】

【0003】

細胞への送達が特に困難な生物学的活性薬剤としては、タンパク質、核酸系薬物、及びその誘導体が挙げられる。有望な遺伝子編集技術を細胞内に送達するための組成物、例えば、CRISPR／Cas9システム成分を送達するための組成物は特に関心が持たれる。

【0004】

現在、細胞の遺伝子をインビボで編集するための成分及びシステムは多数存在し、疾患の治療に非常に大きな可能性を提供している。CRISPR／Cas9遺伝子編集システムは細胞内ではリボ核タンパク質複合体として活性である。RNA依存性ヌクレアーゼは細胞内のDNA配列に結合し、切断を導く。この部位特異的ヌクレアーゼ活性は、細胞自身の自然の過程を介した遺伝子編集を促進する。例えば、細胞は、二本鎖DNA切断（DSB）に応答して、非相同末端結合（「NHEJ」）として知られる誤りがちな修復過程を用いる。NHEJの際、細胞によりDNA末端でヌクレオチドの付加または除去が行われ、切断された配列から変化した配列が生じる場合がある。他の状況では、細胞は、相同組換え修復（「HDR」）または相同組換え（「HR」）という機構によりDSBを修復し、その際、内在性または外来性の鋳型を使用して切断の修復を導くことができる。これらの編集技術のいくつかでは、一本鎖切断（SSB）またはDSBを修復する細胞機構が利用される。

【0005】

CRISPR／Casのタンパク質及び核酸成分を細胞、例えば、患者の細胞などに送達するための組成物が必要とされている。特に、CRISPRタンパク質成分をコードするmRNAを送達するための組成物、及びCRISPRガイドRNAを送達するための組成物は特に関心が持たれる。RNA成分を安定化させ送達することができる、インビトロ及びインビボでの送達に有用な特性を有する組成物も特に関心が持たれる。

【発明の概要】

【0006】

本明細書では、有用な特性、特に、CRISPR／Cas9遺伝子編集成分の送達に有用な特性を有する脂質ナノ粒子を用いる組成物を提供する。

【0007】

特定の実施形態では、LNP組成物はRNA成分及び脂質成分を含み、ここで、かかる脂質成分は、(1) 約50～60モル%のアミン脂質、(2) 約8～10モル%の中性脂質、及び(3) 約2.5～4モル%のPEG脂質を含み、脂質成分の残部はヘルパー脂質であり、LNP組成物のN/P比は約6である。さらなる実施形態では、LNP組成物は、(1) RNA成分、(2) 約50～60モル%のアミン脂質、(3) 約27～39.5モル%のヘルパー脂質、(4) 約8～10モル%の中性脂質、及び(5) 約2.5～4モル%のPEG脂質を含み、ここで、LNP組成物のN/P比は約5～7である。

【0008】

他の実施形態では、LNP組成物はRNA成分及び脂質成分を含み、ここで、かかる脂質成分は、(1) 約50～60モル%のアミン脂質、(2) 約5～15モル%の中性脂質、及び(3) 約2.5～4モル%のPEG脂質を含み、脂質成分の残部はヘルパー脂質であり、LNP組成物のN/P比は約3～10である。さらなる実施形態では、LNP組成物は、(1) 約40～60モル%のアミン脂質、(2) 約5～15モル%の中性脂質、及び(3) 約2.5～4モル%のPEG脂質が含まれる脂質成分を含み、ここで、かかる脂質成分の残部はヘルパー脂質であり、LNP組成物のN/P比は約6である。別の実施形態では、LNP組成物は、(1) 約50～60モル%のアミン脂質、(2) 約5～15モル%の中性脂質、及び(3) 約1.5～10モル%のPEG脂質が含まれる脂質成分を含み、ここで、かかる脂質成分の残部はヘルパー脂質であり、LNP組成物のN/P比は約6である。

【0009】

いくつかの実施形態では、LNP組成物はRNA成分及び脂質成分を含み、ここで、かかる脂質成分は、(1) 約40～60モル%のアミン脂質、(2) 約0～5モル%の中性脂質、例えば、リン脂質、及び(3) 約1.5～10モル%のPEG脂質を含み、脂質成分の残部はヘルパー脂質であり、LNP組成物のN/P比は約3～10である。いくつかの実施形態では、LNP組成物はRNA成分及び脂質成分を含み、ここで、かかる脂質成分は、(1) 約40～60モル%のアミン脂質、(2) 約1モル%未満の中性脂質、例えば、リン脂質、及び(3) 約1.5～10モル%のPEG脂質を含み、脂質成分の残部はヘルパー脂質であり、LNP組成物のN/P比は約3～10である。特定の実施形態では、LNP組成物は中性脂質を本質的に含まない。いくつかの実施形態では、LNP組成物はRNA成分及び脂質成分を含み、ここで、脂質成分は、(1) 約40～60モル%のアミン脂質、及び(2) 約1.5～10モル%のPEG脂質を含み、脂質成分の残部はヘルパー脂質であり、LNP組成物のN/P比は約3～10であり、かつ、LNP組成物は中性脂質、例えば、リン脂質を含まない。特定の実施形態では、LNP組成物は中性リン脂質を本質的に含まないかまたは含まない。特定の実施形態では、LNP組成物は中性脂質、例えば、リン脂質を本質的に含まないかまたは含まない。

【0010】

特定の実施形態では、RNA成分は、RNA誘導型DNA結合因子(例えば、Casヌクレアーゼまたはクラス2Casヌクレアーゼ)などのmRNAを含む。特定の実施形態では、RNA成分はgRNAを含む。

【図面の簡単な説明】

【0011】

【図1】CRISPR/Cas遺伝子編集成分であるCas9 mRNA及びgRNAをLNP組成物に入れて送達した後のマウス肝臓において達成されたTTT遺伝子編集の割合を示し、1mpk(図1A)または0.5mpk(図1B)の単回投与の場合を示す。

【図2】Cas9 mRNA及びgRNAを含むLNP組成物の粒子分布データを示す。

【図3】LNP組成物の物理化学的性質を示し、対数で表す組成物のモル質量差(図3A)及び組成物の分子量測定値の平均(図3B)を比較する。

【図4】図3のLNP組成物を分析する、多分散度計算を図4Aに、またBurchard-Stockmeyer分析を図4Bに示す。

10

20

30

40

50

【図5】PEG脂質濃度を高めたLNP組成物がラットにおける単回投与後に及ぼす血清TTRノックダウン、肝臓での遺伝子編集、及びサイトカインMCP-1レベルに対する効果を評価する実験結果を記載する。図5Aは血清TTRレベルをグラフで表し、図5Bは肝臓試料での編集率をグラフで表し、図5CはMCP-1レベルをpg/mL単位で示している。

【図6-1】LNP組成物では各種PEG脂質で遺伝子編集効力が維持されることを示す（血清TTRレベル（図6A及び図6B）及び編集率（図6C）により測定）。

【図6-2】図6-1の続きである。

【図7-1】マウスにおいて、リピドA類似体は、単回投与後の肝臓での編集率（%）で測定した場合に、LNP組成物中の遺伝子編集カーゴを効果的に送達することを示す。

【図7-2】図7-1の続きである。

【図7-3】図7-2の続きである。

【図7-4】図7-3の続きである。

【図8-1】カニクイザル初代肝細胞における各種LNP組成物での編集率の用量反応曲線を示す。

【図8-2】図8-1の続きである。

【図9-1】図9A及び図9Bは、gRNAとmRNAの比をさまざまに変えた場合のマウスにおける単回投与後の血清TTR及び編集率についての結果を示し、図9C及び図9Dは、Cas9 mRNAの量は一定のままgRNAをさまざまに変えた場合のマウスにおける単回投与後の肝臓での血清TTR及び編集率についての結果を示す。

【図9-2】図9-1の続きである。

【図10】図10A及び図10Bは、中性脂質を用いたLNP組成物及び中性脂質を用いないLNP組成物を投与した後の血清TTR及び肝臓編集の結果を示す。

【発明を実施するための形態】

【0012】

本開示は、細胞に送達するためのCRISPR/Cas成分RNA（「カーゴ」）などのRNAの脂質ナノ粒子（LNP）組成物の実施形態及びそれらの使用方法を提供する。LNP組成物は、従来の送達技術と比較して改善された特性を示し得る。LNP組成物は、本明細書で定義されるようなRNA成分及び脂質成分を含有してよい。特定の実施形態では、RNA成分には、クラス2CasヌクレアーゼなどのCasヌクレアーゼが含まれる。特定の実施形態では、カーゴまたはRNA成分には、クラス2CasヌクレアーゼをコードするmRNA、及びガイドRNAまたはガイドRNAをコードする核酸が含まれる。遺伝子編集方法及び工学的に操作された細胞の作製方法も提供される。

【0013】

CRISPR/Casカーゴ

LNP製剤を介して送達されるCRISPR/Casカーゴには目的タンパク質をコードするmRNA分子が含まれてよい。例えば、緑色蛍光タンパク質（GFP）のようなタンパク質とRNA誘導型DNA結合因子とを発現させるためのmRNA、またはCasヌクレアーゼが含まれる。Cas9タンパク質の細胞内発現を可能にするCasヌクレアーゼmRNA、例えば、クラス2CasヌクレアーゼmRNAなどを含むLNP組成物を提供する。さらに、カーゴは、ガイドRNAまたはガイドRNAをコードする核酸を1つ以上含有してよい。組成物中に例えば、修復または組換えなどのための鋳型核酸も含まれてよく、または鋳型核酸を本明細書に記載する方法において使用してよい。

【0014】

「mRNA」とは、ポリペプチドに翻訳され得る（すなわち、リボソーム及びアミノアシル化tRNAによる翻訳のための基質として使用され得る）オープンリーディングフレームを含むポリヌクレオチドを指す。mRNAは、リボース残基またはその類似体、例えば、2'-メトキシリボース残基など、リン酸塩-糖骨格を含むことができる。いくつかの実施形態では、mRNAリン酸塩-糖骨格の糖は本質的にリボース残基、2'-メトキシリボース残基、またはその組み合わせからなる。一般に、mRNAは実質的な量のチミジン

10

20

30

40

50

残基を含有しない（例えば、チミジン残基が0であるかまたは30、20、10、5、4、3、もしくは2より少ない、あるいはチミジン含量が10%未満、9%未満、8%未満、7%未満、6%未満、5%未満、4%未満、4%未満、3%未満、2%未満、1%未満、0.5%未満、0.2%未満、または0.1%未満である）。mRNAは、そのウリジンの位置のうち一部または全部において修飾ウリジンを含有することができる。

【0015】

CRISPR/Casヌクレアーゼ系

開示される製剤の成分の一つは、RNA誘導型DNA結合因子、例えば、CasヌクレアーゼなどをコードするmRNAである。

【0016】

本明細書で使用する場合、「RNA誘導型DNA結合因子」とは、RNA及びDNA結合活性を有するポリペプチドもしくはポリペプチドの複合体、またはそのような複合体のDNA結合サブユニットを意味し、ここで、DNA結合活性は配列特異的であり、RNAの配列に依存する。例示的なRNA誘導型DNA結合因子としては、Casクリベース/ニッカーゼ及びその不活性型（「dCas DNA結合物質」）が挙げられる。本明細書で使用する場合、「Casヌクレアーゼ」には、Casクリベース、Casニッカーゼ、及びdCas DNA結合物質が包含される。Casクリベース/ニッカーゼ及びdCas DNA結合物質としては、III型CRISPR系のCsm複合体またはCmr複合体、そのサブユニットのCas10、Csm1、またはCmr2、I型CRISPR系のCascade複合体、そのサブユニットのCas3、及びクラス2Casヌクレアーゼが挙げられる。本明細書で使用する場合、「クラス2Casヌクレアーゼ」は、RNA誘導型のDNA結合活性のある一本鎖ポリペプチドである。クラス2Casヌクレアーゼには、RNA誘導によるDNAクリベース活性またはニッカーゼ活性をさらに有するクラス2Casクリベース/ニッカーゼ（例えば、H840A、D10A、またはN863Aという変異型）、及びクリベース/ニッカーゼ活性が不活性化されているクラス2 dCas DNA結合物質が含まれる。クラス2Casヌクレアーゼとしては、例えば、Cas9、Cpf1、C2c1、C2c2、C2c3、HF Cas9というタンパク質（例えば、N497A、R661A、Q695A、Q926Aという変異型）、HypaCas9タンパク質（例えば、N692A、M694A、Q695A、H698Aという変異型）、eSPCas9（1.0）タンパク質（例えば、K810A、K1003A、R1060Aという変異型）、及びeSPCas9（1.1）タンパク質（例えば、K848A、K1003A、R1060Aという変異型）、及びその改変体が挙げられる。Cpf1タンパク質（Zetsche et al., Cell, 163:1-13 (2015)）はCas9に相同であり、RuvC様ヌクレアーゼドメインを含有する。ZetscheのCpf1配列は、参照によりその全体が組み込まれる。例えば、Zetscheの表S1及び表S3を参照のこと。例えば、Makarova et al., Nat Rev Microbiol, 13(11):722-36 (2015)、Shmakov et al., Molecular Cell, 60:385-397 (2015)を参照のこと。

【0017】

いくつかの実施形態では、RNA誘導型DNA結合因子はクラス2Casヌクレアーゼである。いくつかの実施形態では、RNA誘導型DNA結合因子はクリベース活性を有し、これは二本鎖エンドヌクレアーゼ活性とも呼ばれ得る。いくつかの実施形態では、RNA誘導型DNA結合因子は、クラス2Casヌクレアーゼ（これは、例えば、II型、V型、またはVI型のCasヌクレアーゼであってよい）などのCasヌクレアーゼを含む。クラス2Casヌクレアーゼとしては、例えば、Cas9、Cpf1、C2c1、C2c2、及びC2c3タンパク質ならびにその改変体が挙げられる。Cas9ヌクレアーゼの例としては、S. pyogenes、S. aureus、及び他の原核生物（例えば、次段落の一覧を参照のこと）のII型CRISPR系のもの、ならびにその改変型（例えば、工学的に操作されたもの、または変異体）が挙げられる。例えば、U.S. 2016/0312198 A1、U.S. 2016/0312199 A1を参照のこと。Cas

10

20

30

40

50

ヌクレアーゼの他の例としては、III型CRISPR系のCsm複合体もしくはCmr複合体またはそのサブユニットのCas10、Csm1、もしくはCmr2、及びI型CRISPR系のCascade複合体、またはそのサブユニットのCas3が挙げられる。いくつかの実施形態では、Casヌクレアーゼは、IIA型、IIB型、またはIIC型の系のものであってよい。さまざまなCRISPR系及びCasヌクレアーゼの考察については、例えば、Makarova et al., Nat. Rev. Microbiol. 9:467-477 (2011)、Makarova et al., Nat. Rev. Microbiol. 13:722-36 (2015)、Shmakov et al., Molecular Cell, 60:385-397 (2015)を参照のこと。

[0018]

Casヌクレアーゼが由来し得る非限定的な例示的種としては、*Streptococcus pyogenes*、*Streptococcus thermophilus*、*Streptococcus sp.*、*Staphylococcus aureus*、*Listeria innocua*、*Lactobacillus gasseri*、*Francisella novicida*、*Wolinella succinogenes*、*Sutterella wadsworthensis*、*Gammaproteobacterium*、*Neisseria meningitidis*、*Campylobacter jejuni*、*Pasteurella multocida*、*Fibrobacter succinogenes*、*Rhodospirillum rubrum*、*Nocardiaopsis dassonvillei*、*Streptomyces pristinaespiralis*、*Streptomyces viridochromogenes*、*Streptomyces viridochromogenes*、*Streptoporum roseum*、*Streptoporum roseum*、*Alicyclobacillus acidocaldarius*、*Bacillus pseudomycoides*、*Bacillus selenitireducens*、*Exiguobacterium sibiricum*、*Lactobacillus delbrueckii*、*Lactobacillus salivarius*、*Lactobacillus buchneri*、*Treponema denticola*、*Microscilla marina*、*Burkholderiales bacterium*、*Polaromonas naphthalenivorans*、*Polaromonas sp.*、*Crocospaera watsonii*、*Cyanothece sp.*、*Microcystis aeruginosa*、*Synechococcus sp.*、*Acetohalobium arabaticum*、*Ammonifex degensii*、*Caldicelulosiruptor beccarii*、*Candidatus Desulforudis*、*Clostridium botulinum*、*Clostridium difficile*、*Finnegoldia magna*、*Natronaerobius thermophilus*、*Pelotomaculum thermopropionicum*、*Acidithiobacillus caldus*、*Acidithiobacillus ferrooxidans*、*Allochrochromatium vinosum*、*Marinobacter sp.*、*Nitrosococcus halophilus*、*Nitrosococcus watsoni*、*Pseudoalteromonas haloplanktis*、*Ktedonobacter racemifer*、*Methanohalobium evestigatum*、*Anabaena variabilis*、*Nodularia spumigena*、*Nostoc sp.*、*Arthrospira maxima*、*Arthrospira platensis*、*Arthrospira sp.*、*Lyngbya sp.*、*Microcoleus chthonoplastes*、*Oscillatoria sp.*、*Petrogobacter mobilis*、*Thermosiphon africanus*、*Streptococcus pasteurianus*、*Neisseria cinerea*、*Campylobacter lari*、*Parvibaculum lav*

10

20

30

40

50

amentivorans、Corynebacterium diphtheria、Acidaminococcus sp.、Lachnospiraceae bacterium ND2006、及びAcaryochloris marinaが挙げられる。

【0019】

いくつかの実施形態では、CasヌクレアーゼはStreptococcus pyogenes由来のCas9ヌクレアーゼである。いくつかの実施形態では、CasヌクレアーゼはStreptococcus thermophilus由来のCas9ヌクレアーゼである。いくつかの実施形態では、CasヌクレアーゼはNeisseria meningitidis由来のCas9ヌクレアーゼである。いくつかの実施形態では、CasヌクレアーゼはStaphylococcus aureus由来のCas9ヌクレアーゼである。いくつかの実施形態では、CasヌクレアーゼはFrancisella novicida由来のCpf1ヌクレアーゼである。いくつかの実施形態では、CasヌクレアーゼはAcidaminococcus sp. 由来のCpf1ヌクレアーゼである。いくつかの実施形態では、CasヌクレアーゼはLachnospiraceae bacterium ND2006由来のCpf1ヌクレアーゼである。さらなる実施形態では、CasヌクレアーゼはFrancisella tularensis、Lachnospiraceae bacterium、Butyrivibrio proteoclasticus、Peregrinibacteria bacterium、Parcubacteria bacterium、Smithella、Acidaminococcus、Candidatus Methanoplasma termittum、Eubacterium eligens、Moraxella bovoculi、Leptospira inadai、Porphyromonas crevioricanis、Prevotella disiens、またはPorphyromonas macacae由来のCpf1ヌクレアーゼである。特定の実施形態では、CasヌクレアーゼはAcidaminococcusまたはLachnospiraceaeに由来するCpf1ヌクレアーゼである。

【0020】

野生型Cas9はRuvC及びHNHという2つのヌクレアーゼドメインを有する。RuvCドメインは非標的DNA鎖を切断し、HNHドメインは標的DNA鎖を切断する。いくつかの実施形態では、Cas9ヌクレアーゼは、2つ以上のRuvCドメイン及び／または2つ以上のHNHドメインを含む。いくつかの実施形態では、Cas9ヌクレアーゼは野生型Cas9である。いくつかの実施形態では、Cas9は、標的DNAに二本鎖切断を誘導することができる。特定の実施形態では、Casヌクレアーゼは、dsDNAを切断するか、dsDNAの1本鎖を切断するか、またはDNAのクリベース活性もニックーゼ活性も持たない場合がある。例示的なCas9アミノ酸配列を配列番号3として記載する。開始コドン及び終止コドンを含む、例示的なCas9 mRNA ORF配列を配列番号4として記載する。融合タンパク質に含めるのに好適な例示的Cas9 mRNAコード配列を配列番号10として記載する。

【0021】

いくつかの実施形態では、キメラCasヌクレアーゼを使用し、その場合、かかるタンパク質の1つのドメインまたは領域は、異なるタンパク質の一部で置換される。いくつかの実施形態では、Casヌクレアーゼドメインを、FokIなどの異なるヌクレアーゼに由来するドメインで置き換えてよい。いくつかの実施形態では、Casヌクレアーゼは修飾ヌクレアーゼであってよい。

【0022】

他の実施形態では、CasヌクレアーゼはI型CRISPR/Cas系に由来してよい。いくつかの実施形態では、Casヌクレアーゼは、I型CRISPR/Cas系のCascade複合体の成分であってよい。いくつかの実施形態では、CasヌクレアーゼはCas3タンパク質であってよい。いくつかの実施形態では、CasヌクレアーゼはII型CRISPR/Cas系に由来してよい。いくつかの実施形態では、Casヌクレア

ーゼはRNA切断活性を有してよい。

【0023】

いくつかの実施形態では、RNA誘導型DNA結合因子は一本鎖ニッカーゼ活性を有し、すなわち、一本のDNA鎖を切断して、「ニック」としても知られる一本鎖の切れ目を生じさせることができる。いくつかの実施形態では、RNA誘導型DNA結合因子はCasニッカーゼを含む。ニッカーゼは、dsDNAにニックをつくる、すなわち、DNA二重らせんの1本の鎖を切断するが、もう1本の鎖は切断しない酵素である。いくつかの実施形態では、CasニッカーゼはCasヌクレアーゼ（例えば、上述のCasヌクレアーゼ）の変形の1種であり、例えば、触媒ドメイン内の1つ以上の変化（例えば、点変異）によって、エンドヌクレアーゼによる分解活性部位が不活性化されている。Casニッカーゼ及び例示的な触媒ドメイン改変に関する考察については、例えば、米国特許第8,889,356号を参照のこと。いくつかの実施形態では、Cas9ニッカーゼなどのCasニッカーゼは不活性化されたRuvCドメインまたはHNHドメインを有する。例示的なCas9ニッカーゼのアミノ酸配列を配列番号6として記載する。開始コドン及び終止コドンが含まれる例示的なCas9ニッカーゼmRNAのORF配列を配列番号7として記載する。融合タンパク質に含めるのに好適な例示的なCas9ニッカーゼmRNAのコード配列を配列番号11として記載する。

10

【0024】

いくつかの実施形態では、RNA誘導型DNA結合因子は、機能するヌクレアーゼドメインを1つのみ含有するよう修飾される。例えば、因子のタンパク質は、ヌクレアーゼドメインの1つを変異させるかまたは完全もしくは部分的に欠失させてその核酸切断活性を低下させるよう修飾されてよい。いくつかの実施形態では、活性を低下させたRuvCドメインを有するニッカーゼを使用する。いくつかの実施形態では、不活性なRuvCドメインを有するニッカーゼを使用する。いくつかの実施形態では、活性を低下させたHNHドメインを有するニッカーゼを使用する。いくつかの実施形態では、不活性なHNHドメインを有するニッカーゼを使用する。

20

【0025】

いくつかの実施形態では、Casタンパク質ヌクレアーゼドメイン内の保存されたアミノ酸を置換してヌクレアーゼ活性を低下または変化させる。いくつかの実施形態では、Casヌクレアーゼは、RuvCまたはRuvC様ヌクレアーゼドメインにアミノ酸置換を含んでよい。RuvCまたはRuvC様ヌクレアーゼドメインにおける例示的なアミノ酸置換としては、D10A (*S. pyogenes* Cas9タンパク質に基づく)が挙げられる。例えば、Zetsche et al. (2015) Cell Oct 22:163(3):759-771を参照のこと。いくつかの実施形態では、Casヌクレアーゼは、HNHまたはHNH様ヌクレアーゼドメインにアミノ酸置換を含んでよい。HNHまたはHNH様ヌクレアーゼドメインにおける例示的なアミノ酸置換としては、E762A、H840A、N863A、H983A、及びD986A (*S. pyogenes* Cas9タンパク質に基づく)が挙げられる。例えば、Zetsche et al. (2015)を参照のこと。さらなる例示的なアミノ酸置換としては、D917A、E1006A、及びD1255A (*Francisella novicida* U112 Cpf1 (FnCpf1)配列 (UniProtKB-A0Q7Q2 (CPF1_FRNATN)に基づく)が挙げられる。

30

40

【0026】

いくつかの実施形態では、ニッカーゼをコードするmRNAは、標的配列のセンス鎖及びアンチセンス鎖のそれぞれに相補的な一対のガイドRNAと組み合わせて提供される。この実施形態では、ガイドRNAはニッカーゼを標的配列へと導き、標的配列の反対鎖にニックを生じさせることによりDSBを導入する（すなわち、ダブルニックング）。いくつかの実施形態では、ダブルニックングの使用により特異性が改善され、オフターゲット作用が低減され得る。いくつかの実施形態では、ニッカーゼを、対抗するDNA鎖を標的にする2つの別々のガイドRNAと共に使用して、標的DNAにダブルニックを作製する

50

。いくつかの実施形態では、ニッカーゼを、近接するよう選択された2つの別々のガイドRNAと共に使用して、標的DNAにダブルニックを作製する。

【0027】

いくつかの実施形態では、RNA誘導型DNA結合因子にはクリベース活性及びニッカーゼ活性がない。いくつかの実施形態では、RNA誘導型DNA結合因子はdCas DNA結合ポリペプチドを含む。dCasポリペプチドはDNA結合活性を有するが、本質的に触媒（クリベース／ニッカーゼ）活性がない。いくつかの実施形態では、dCasポリペプチドはdCas9ポリペプチドである。いくつかの実施形態では、クリベース活性及びニッカーゼ活性がないRNA誘導型DNA結合因子またはdCas DNA結合ポリペプチドは、Casヌクレアーゼ（例えば、上述のCasヌクレアーゼ）の変形の1種であり、例えば、その触媒ドメイン内の1つ以上の変化（例えば、点変異）によってそのエンドヌクレアーゼによる分解活性部位が不活性化されている。例えば、U.S. 2014/0186958 A1、U.S. 2015/0166980 A1を参照のこと。例示的なdCas9アミノ酸配列を配列番号8として記載する。開始コドン及び終止コドンが含まれる例示的なCas9 mRNA ORF配列を配列番号9として記載する。融合タンパク質に含めるのに好適な例示的Cas9 mRNAコード配列を配列番号12として記載する。

10

【0028】

いくつかの実施形態では、RNA誘導型DNA結合因子は、1つ以上の異種機能ドメイン（例えば、融合ポリペプチドであるかまたはそれを含む）を含む。

20

【0029】

いくつかの実施形態では、異種機能ドメインは、細胞の核内へのRNA誘導型DNA結合因子の輸送を容易にし得る。例えば、異種機能ドメインは核局在化シグナル（NLS）であってよい。いくつかの実施形態では、RNA誘導型DNA結合因子を1～10のNLS（複数可）と融合させてよい。いくつかの実施形態では、RNA誘導型DNA結合因子を1～5のNLS（複数可）と融合させてよい。いくつかの実施形態では、RNA誘導型DNA結合因子を1つのNLSと融合させてよい。1つのNLSを使用する場合、かかるNLSを、RNA誘導型DNA結合因子配列のN末端またはC末端にて連結させてよい。かかる1つのNLSを、RNA誘導型DNA結合因子配列の内部に挿入してもよい。他の実施形態では、RNA誘導型DNA結合因子を2つ以上のNLSと融合させてよい。いくつかの実施形態では、RNA誘導型DNA結合因子を2つ、3つ、4つ、または5つのNLSと融合させてよい。いくつかの実施形態では、RNA誘導型DNA結合因子を2つのNLSと融合させてよい。特定の状況では、2つのNLSは同一であっても（例えば、SV40 NLSが2つ）異なってもよい。いくつかの実施形態では、RNA誘導型DNA結合因子を、2つのSV40 NLS配列とカルボキシ末端で連結して融合させる。いくつかの実施形態では、RNA誘導型DNA結合因子を2つのNLSと融合させて、その際、1つはN末端で連結し、1つはC末端で連結してよい。いくつかの実施形態では、RNA誘導型DNA結合因子を3つのNLSと融合させてよい。いくつかの実施形態では、RNA誘導型DNA結合因子をNLSと融合させなくてよい。いくつかの実施形態では、NLSは、例えば、SV40 NLSであるPKKKRKVまたはPKKKRRVのような一分（monopartite）配列であってよい。いくつかの実施形態では、NLSは、ヌクレオプラスミンのNLSであるKRPAATKKAGQAKKKKのような二分（bipartite）配列であってよい。特定の実施形態では、単一のPKKKRKV NLSをRNA誘導型DNA結合因子のC末端で連結させてよい。任意選択で1つ以上のリンカーが融合部位において含まれる。

30

40

【0030】

いくつかの実施形態では、異種機能ドメインは、RNA誘導型DNA結合因子の細胞内半減期を改変する能力があってよい。いくつかの実施形態では、RNA誘導型DNA結合因子の半減期を延長させる能力があってよい。いくつかの実施形態では、RNA誘導型DNA結合因子の半減期は短縮されてよい。いくつかの実施形態では、異種機能ドメインは

50

、RNA誘導型DNA結合因子の安定性を増大させる能力があってよい。いくつかの実施形態では、異種機能ドメインは、RNA誘導型DNA結合因子の安定性を低下させる能力があってよい。いくつかの実施形態では、異種機能ドメインは、タンパク質分解のシグナルペプチドとして作用してよい。いくつかの実施形態では、タンパク質分解は、例えば、プロテアソーム、リソソームのプロテアーゼ、またはカルパインプロテアーゼなどのタンパク質分解酵素によって媒介されてよい。いくつかの実施形態では、異種機能ドメインはPEST配列を含んでよい。いくつかの実施形態では、RNA誘導型DNA結合因子は、ユビキチンまたはポリユビキチン鎖の付加によって修飾されてよい。いくつかの実施形態では、ユビキチンはユビキチン様タンパク質(UBL)であってよい。ユビキチン様タンパク質の非限定的な例としては、低分子ユビキチン様修飾因子(SUMO)、ユビキチン交差反応性タンパク質(UCRP; インターフェロン誘導性遺伝子15(ISG15)としても知られる)、ユビキチン関連修飾因子1(URM1)、神経前駆細胞発現し発生段階で下方制御されたタンパク質8(NEDD8; *S. cerevisiae*ではRub1とも呼ばれる)、ヒト白血球抗原F関連(FAT10)、オートファジー8(ATG8)及びオートファジー12(ATG12)、Fauユビキチン様タンパク質(FUB1)、膜アンカー型UBL(MUB)、ユビキチンフォールド修飾因子1(UFM1)、及びユビキチン様タンパク質5(UBL5)が挙げられる。

【0031】

いくつかの実施形態では、異種機能ドメインはマーカードメインであってよい。マーカードメインの非限定的な例としては、蛍光タンパク質、精製タグ、エピトープタグ、及びレポーター遺伝子配列が挙げられる。いくつかの実施形態では、マーカードメインは蛍光タンパク質であってよい。適切な蛍光タンパク質の非限定的な例としては、緑色蛍光タンパク質(例えば、GFP、GFP-2、タグGFP、turboGFP、sfGFP、EGFP、Emerald、Azami Green、単量体Azami Green、CopGFP、AceGFP、ZsGreen1)、黄色蛍光タンパク質(例えば、YFP、EYFP、Citrine、Venus、YPet、PhiYFP、ZsYellow1)、青色蛍光タンパク質(例えば、EBFP、EBFP2、Azurite、mKalamal、GFPuv、Sapphire、T-sapphire)、シアン蛍光タンパク質(例えば、ECFP、Cerulean、CyPet、AmCyan1、Midorishii-Cyan)、赤色蛍光タンパク質(例えば、mKate、mKate2、mPlum、DsRed単量体、mCherry、mRFP1、DsRed-Express、DsRed2、DsRed-Monomer、HcRed-Tandem、HcRed1、AsRed2、eqFP611、mRaspberry、mStrawberry、Jred)、及びオレンジ色蛍光タンパク質(mOrange、mKO、Kusabira-Orange、単量体Kusabira-Orange、mTangerine、tdTomato)または他の適切な任意の蛍光タンパク質が挙げられる。他の実施形態では、マーカードメインは精製タグ及び／またはエピトープタグであってよい。非限定的な例示的タグとしては、グルタチオン-S-トランスフェラーゼ(GST)、キチン結合タンパク質(CBP)、マルトース結合タンパク質(MBP)、チオレドキシン(TRX)、ポリ(NANP)、タンデムアフィニティー精製(TAP)タグ、myc、AcV5、AU1、AU5、E、ECS、E2、FLAG、HA、nus、Softag1、Softag3、Strep、SBP、Glu-Glu、HSV、KT3、S、S1、T7、V5、VSV-G、6xHis、8xHis、ビオチンカルボキシルキャリアタンパク質(BCCP)、ポリ-His、及びカルモジュリンが挙げられる。非限定的な例示的レポーター遺伝子としては、グルタチオン-S-トランスフェラーゼ(GST)、西洋わさびペルオキシダーゼ(HRP)、クロラムフェニコールアセチルトランスフェラーゼ(CAT)、ベータガラクトシダーゼ、ベータグルクロニダーゼ、ルシフェラーゼ、または蛍光タンパク質が挙げられる。

【0032】

さらなる実施形態では、異種機能ドメインは、RNA誘導型DNA結合因子を、特定の

細胞小器官、細胞型、組織、または臓器に指向させてよい。いくつかの実施形態では、異種機能ドメインは、RNA誘導型DNA結合因子をミトコンドリアに指向させてよい。

【0033】

さらなる実施形態では、異種機能ドメインはエフェクタードメインであってよい。RNA誘導型DNA結合因子をその標的配列へ配向させる場合、例えば、CasヌクレアーゼをgRNAによって標的配列へ配向させる場合、エフェクタードメインは標的配列を修飾するかまたはそれに影響を与えてよい。いくつかの実施形態では、エフェクタードメインは、核酸結合ドメイン、ヌクレアーゼドメイン（例えば、非Casヌクレアーゼドメイン）、エピジェネティック修飾ドメイン、転写活性化ドメイン、または転写抑制因子ドメインから選んでよい。いくつかの実施形態では、異種機能ドメインはFokIヌクレアーゼなどのヌクレアーゼである。例えば、米国特許第9,023,649号を参照のこと。いくつかの実施形態では、異種機能ドメインは転写活性化因子または転写抑制因子である。例えば、Qi et al., "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression," Cell 152:1173-83 (2013)、Perez-Pinera et al., "RNA-guided gene activation by CRISPR-Cas9-based transcription factors," Nat. Methods 10:973-6 (2013)、Mali et al., "CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering," Nat. Biotechnol. 31:833-8 (2013)、Gilbert et al., "CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes," Cell 154:442-51 (2013)を参照のこと。したがって、RNA誘導型DNA結合因子は本質的に、ガイドRNAを使用して所望の標的配列と結合させるために配向可能な転写因子になる。特定の実施形態では、DNA修飾ドメインは脱メチル化ドメインまたはメチルトランスフェラーゼドメインなどのメチル化ドメインである。特定の実施形態では、エフェクタードメインは塩基編集ドメインなどのDNA修飾ドメインである。特定の実施形態では、DNA修飾ドメインは、デアミナーゼドメインなど、DNA内に特定の修飾を導入する核酸編集ドメインである。例えば、WO2015/089406、U.S. 2016/0304846を参照のこと。WO2015/089406及びU.S. 2016/0304846に記載の核酸編集ドメイン、デアミナーゼドメイン、及びCas9変異型は参照により本明細書に組み込まれる。

【0034】

ヌクレアーゼは、ガイドRNA（「gRNA」）と相互作用する少なくとも1つのドメインを含んでよい。さらに、ヌクレアーゼをgRNAによって標的配列に配向させてよい。クラス2Casヌクレアーゼ系では、gRNAはヌクレアーゼ及び標的配列と相互作用し、それにより標的配列への結合が導かれるようにする。いくつかの実施形態では、gRNAは目標とする切断に対する特異性を提供する一方、ヌクレアーゼは汎用性がある、異なるgRNAと対形成して異なる標的配列を切断し得る。クラス2Casヌクレアーゼは、上掲の型、オルソログ、及び例示的種のgRNA足場構造と対形成してよい。

【0035】

ガイドRNA (gRNA)

本開示のいくつかの実施形態では、LNP製剤のカーゴには少なくとも1つのgRNAが含まれる。gRNAは、Casヌクレアーゼまたはクラス2Casヌクレアーゼを標的核酸分子上の標的配列に誘導してよい。いくつかの実施形態では、gRNAはクラス2Casヌクレアーゼと結合して、クラス2Casヌクレアーゼによる切断の特異性を提供する。いくつかの実施形態では、gRNAとCasヌクレアーゼは、LNP組成物により送達され得るリボ核タンパク質（RNP）、例えば、CRISPR/Cas9複合体のよう

なCRISPR/Cas複合体などを形成してよい。いくつかの実施形態では、CRISPR/Cas複合体はII型CRISPR/Cas9複合体であってよい。いくつかの実施形態では、CRISPR/Cas複合体はCpf1/ガイドRNA複合体などのV型CRISPR/Cas複合体であってよい。Casヌクレアーゼと特異的gRNAとを対形成させてよい。各クラス2Casヌクレアーゼと対形成するgRNA足場構造は個々のCRISPR/Cas系によって異なる。

【0036】

「ガイドRNA」、「gRNA」、及び単に「ガイド」は本明細書では同じ意味で使用され、crRNA (CRISPR RNAとしても知られる)、またはcrRNAとtrRNAの組み合わせ (tracrRNAとしても知られる) のいずれも指す。crRNAとtrRNAは、単一RNA分子として会合していても (単一ガイドRNA、sgRNA)、または2つの別々のRNA分子 (デュアルガイドRNA、dgRNA) であってもよい。「ガイドRNA」または「gRNA」とはそれぞれの型を指す。trRNAは、天然に生じる配列、または天然に生じる配列と比較して修飾または変形を有するtrRNA配列のいずれであってもよい。

【0037】

本明細書で使用する場合、「ガイド配列」とは、ガイドRNA内にある配列であって、標的配列に相補的であり、かつ、RNA誘導型DNA結合因子が結合または修飾 (例えば、切断) を行うためにガイドRNAを標的配列へと導くよう機能する配列を指す。「ガイド配列」は、「指向性配列」または「スペーサー配列」とも呼ばれ得る。ガイド配列は、例えば、*Streptococcus pyogenes* (すなわち、Spy Cas9) 及び関連Cas9相同体/オルソログの場合、長さが20塩基対であり得る。それより短いまたは長い配列、例えば、15ヌクレオチド長、16ヌクレオチド長、17ヌクレオチド長、18ヌクレオチド長、19ヌクレオチド長、21ヌクレオチド長、22ヌクレオチド長、23ヌクレオチド長、24ヌクレオチド長、または25ヌクレオチド長の配列をガイドとして使用することができる。いくつかの実施形態では、標的配列は、例えば、遺伝子内または染色体上にあり、ガイド配列に相補的である。いくつかの実施形態では、ガイド配列とその対応する標的配列の間の相補性または同一性の程度は約75%、80%、85%、90%、95%、96%、97%、98%、99%、または100%であってよい。いくつかの実施形態では、ガイド配列と標的領域は100%相補的であっても同一であってもよい。他の実施形態では、ガイド配列と標的領域には少なくとも1つのミスマッチが含有されてよい。例えば、ガイド配列と標的配列には、1つ、2つ、3つ、または4つのミスマッチが含有されてよく、その場合、標的配列の全長は少なくとも17塩基対、18塩基対、19塩基対、20塩基対、またはそれ以上である。いくつかの実施形態では、ガイド配列と標的領域には1~4つのミスマッチが含有されてよく、その場合、ガイド配列は少なくとも17個、18個、19個、20個、またはそれ以上のヌクレオチドを含む。いくつかの実施形態では、ガイド配列と標的領域には1つ、2つ、3つ、または4つのミスマッチが含有されてよく、その場合、ガイド配列は20個のヌクレオチドを含む。

【0038】

Casタンパク質のための基質となる核酸は二本鎖核酸であるため、Casタンパク質の標的配列にはゲノムDNAのプラス鎖とマイナス鎖の両方 (すなわち、所与の配列と、かかる配列の逆相補鎖) が含まれる。したがって、ガイド配列が「標的配列に相補的である」という場合は、かかるガイド配列がガイドRNAを導いて標的配列の逆相補鎖に結合させ得ることを理解されるべきである。したがって、いくつかの実施形態では、ガイド配列が標的配列の逆相補鎖と結合する場合、かかるガイド配列は、ガイド配列のTがUで置換される以外は標的配列 (例えば、PAMを含まない標的配列) の特定のヌクレオチドと同一である。

【0039】

指向性配列の長さは使用するCRISPR/Cas系及び成分によって異なってよい。例えば、異なる細菌種に由来する異なるクラス2Casヌクレアーゼでは最適な指向性配

10

20

30

40

50

列長がさまざまである。したがって、指向性配列は、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、または50超のヌクレオチド長を含んでよい。いくつかの実施形態では、指向性配列長は、天然に生じるCRISPR/Cas系のガイド配列よりも0ヌクレオチド、1ヌクレオチド、2ヌクレオチド、3ヌクレオチド、4ヌクレオチド、または5ヌクレオチド長いかまたは短い。特定の実施形態では、Casヌクレアーゼ及びgRNA足場は同じCRISPR/Cas系に由来する。いくつかの実施形態では、指向性配列は18~24個のヌクレオチドを含むかまたはそれからなっておりよい。いくつかの実施形態では、指向性配列は19~21個のヌクレオチドを含むかまたはそれからなっておりよい。いくつかの実施形態では、指向性配列は20個のヌクレオチドを含むかまたはそれからなっておりよい。

10

【0040】

いくつかの実施形態では、sgRNAは、Cas9タンパク質によるRNA誘導性DNA切断を媒介することができる「Cas9 sgRNA」である。いくつかの実施形態では、sgRNAは、Cpf1タンパク質によるRNA誘導性DNA切断を媒介することができる「Cpf1 sgRNA」である。特定の実施形態では、gRNAは、Cas9タンパク質との活性複合体を形成してRNA誘導性DNA切断を媒介するのに十分な、crRNA及びtracrRNAを含む。特定の実施形態では、gRNAは、Cpf1タンパク質との活性複合体を形成してRNA誘導性DNA切断を媒介するのに十分なcrRNAを含む。Zetsche 2015を参照のこと。

20

【0041】

本発明の特定の実施形態はまた、本明細書に記載のgRNAをコードする核酸、例えば、発現カセットも提供する。「ガイドRNA核酸」は、本明細書ではガイドRNA（例えば、sgRNAまたはdgRNA）及びガイドRNA発現カセットを指すために使用される、1つ以上のガイドRNAをコードする核酸である。

【0042】

いくつかの実施形態では、核酸はDNA分子であってよい。いくつかの実施形態では、核酸は、crRNAをコードするヌクレオチド配列を含んでよい。いくつかの実施形態では、crRNAをコードするヌクレオチド配列は、天然に生じるCRISPR/Cas系由来の繰り返し配列の全部または一部が隣接する指向性配列を含む。いくつかの実施形態では、核酸は、tracrRNAをコードするヌクレオチド配列を含んでよい。いくつかの実施形態では、crRNA及びtracrRNAは、2つの別々の核酸によってコードされてよい。他の実施形態では、crRNA及びtracrRNAは単一核酸によってコードされてよい。いくつかの実施形態では、crRNA及びtracrRNAは単一核酸の反対鎖によってコードされてよい。他の実施形態では、crRNA及びtracrRNAは単一核酸の同一鎖によってコードされてよい。いくつかの実施形態では、gRNA核酸はsgRNAをコードする。いくつかの実施形態では、gRNA核酸はCas9ヌクレアーゼsgRNAをコードする。いくつかの実施形態では、gRNA核酸はCpf1ヌクレアーゼsgRNAをコードする。

30

【0043】

ガイドRNAをコードするヌクレオチド配列は、少なくとも1つの転写性または調節性の制御配列、例えば、プロモーター、3'UTR、または5'UTRなどに機能的に連結されてよい。一例では、プロモーターは、tRNAプロモーター、例えば、tRNA^{lys}3、またはtRNAキメラであってよい。Mefferd et al., RNA. 2015 21:1683-9, Scherer et al., Nucleic Acids Res. 2007 35:2620-2628を参照のこと。特定の実施形態では、プロモーターは、RNAポリメラーゼIII（Pol III）によって認識されてよい。Pol IIIプロモーターの非限定的な例としてはU6プロモーター及びH1プロモーターも挙げられる。いくつかの実施形態では、ガイドRNAをコードするヌクレオチド配列は、マウスまたはヒトのU6プロモーターに機能的に連結されてよい。いくつかの実施形態では

40

50

、gRNA核酸は修飾核酸である。特定の実施形態では、gRNA核酸には、修飾されたヌクレオシドまたはヌクレオチドが含まれる。いくつかの実施形態では、gRNA核酸には、5'末端修飾、例えば、核酸を安定させ、かつ、核酸の組み込みを妨げる修飾されたヌクレオシドまたはヌクレオチドなどが含まれる。いくつかの実施形態では、gRNA核酸は、各鎖に5'末端修飾を有する二本鎖DNAを含む。特定の実施形態では、gRNA核酸には、逆位ジデオキシ-Tまたは逆位脱塩基ヌクレオシドもしくはヌクレオチドが5'末端修飾として含まれる。いくつかの実施形態では、gRNA核酸には、ビオチン、デスチオビオテン-TEG、ジゴキシゲニン、ならびに蛍光マーカー、例えば、FAM、ROX、TAMRA、及びAlexaFluorなどといった標識が含まれる。

【0044】

特定の実施形態では、2つ以上のgRNA核酸、例えば、gRNAなどをCRISPR/Casヌクレアーゼ系と共に使用することができる。CRISPR/Cas系により2つ以上の標的配列が切断されるよう、gRNA核酸ごとに異なる指向性配列を含有してよい。いくつかの実施形態では、1つ以上のgRNAは、CRISPR/Cas複合体内で特性、例えば、活性または安定性などが同一であっても異なってもよい。2つ以上のgRNAを使用する場合、それぞれのgRNAは同一gRNA核酸上または異なるgRNA核酸上のいずれでもコードされ得る。2つ以上のgRNAの発現誘導に使用するプロモーターは同一であっても異なってもよい。

【0045】

修飾RNA

特定の実施形態では、LNP組成物は修飾RNAを含む。

【0046】

修飾されたヌクレオシドまたはヌクレオチドは、RNA、例えば、gRNAまたはmRNAなどに存在し得る。例えば、1つ以上の修飾されたヌクレオシドまたはヌクレオチドを含むgRNAまたはmRNAは「修飾」RNAと呼ばれ、標準残基A、G、C、及びUの代わりまたは追加で使用される、天然に生じない成分もしくは構成及び／または天然に生じる成分もしくは構成が1つ以上存在することを表す。いくつかの実施形態では、修飾RNAは、本明細書で「修飾」と呼ばれる非標準ヌクレオシドまたはヌクレオチドを用いて合成される。

【0047】

修飾ヌクレオシド及び修飾ヌクレオチドには、(i)ホスホジエステル骨格結合におけるリン酸基の非連結酸素のうち一方もしくは両方及び／またはリン酸基の連結酸素のうち1つ以上の変化、例えば、置換(例示的な骨格修飾)、(ii)リボース糖の構成要素、例えば、リボース糖の2'位のヒドロキシルなどの変化、例えば、置換(例示的な糖修飾)、(iii)「デホスホ」リンカーでのリン酸部分の大幅な置換(例示的な骨格修飾)、(iv)非標準核酸塩基を用いるなど、天然に生じる核酸塩基の修飾または置換(例示的な塩基修飾)、(v)リボース-リン酸骨格の置換または修飾(例示的な骨格修飾)、(vi)オリゴヌクレオチドの3'末端または5'末端の修飾、例えば、末端リン酸基を除去、修飾もしくは置換するか、または部分、キャップもしくはリンカーを結合させる修飾(そのような3'または5'のキャップ修飾は糖修飾及び／または骨格修飾を含んでよい)、及び(vii)糖の修飾または置換(例示的な糖修飾)のうち1つ以上が含まれ得る。特定の実施形態は、mRNA、gRNA、または核酸に対する5'末端修飾を含む。特定の実施形態は、mRNA、gRNA、または核酸に対する3'末端修飾を含む。修飾RNAは、5'末端及び3'末端の修飾を含有し得る。修飾RNAは、末端ではない位置に1つ以上の修飾残基を含有し得る。特定の実施形態では、gRNAには少なくとも1つの修飾残基が含まれる。特定の実施形態では、mRNAには少なくとも1つの修飾残基が含まれる。

【0048】

本明細書で使用する場合、第1の配列と第2の配列のアラインメントにより、第1の配列が、第2の配列全体の位置のX%以上と一致することが示される場合、第1の配列は、第2の配列「に対する同一性が少なくともX%である配列を含む」とみなされる。例えば

10

20

30

40

50

、配列 A A G A は、配列 A A G に対する同一性が 100% の配列を含むが、これは、アラインメントにより、第 2 の配列の 3 つの位置すべてに対して一致が見られる 100% という同一性が示されると考えられるためである。関連するヌクレオチド（チミジン、ウリジン、または修飾ウリジンなど）が同一の補体を有する限り（例えば、チミジン、ウリジン、または修飾ウリジンのいずれについてもアデノシンであり、別の例では、シトシン及び 5-メチルシトシンはいずれもグアノシンまたは修飾グアノシンを補体として有する）、RNA と DNA の違い（一般に、ウリジンをチミジンに交換、またはその逆）及び修飾ウリジンなどのヌクレオシド類似体の存在は、ポリヌクレオチド間の同一性または相補性の違いに寄与しない。したがって、例えば、X が任意の修飾ウリジン、例えば、プソイドウリジン、N1-メチルプソイドウリジン、または 5-メトキシウリジンなどである 5'-A X G という配列は A U G と 100% 同一であるとみなされ、両配列とも同じ配列（5'-C A U）に対して完全に相補的である。例示的なアラインメントアルゴリズムは、S m i t h-W a t e r m a n 及び N e e d l e m a n-W u n s c h アルゴリズムであり、これらは当該技術分野で周知である。当業者は、所与の整列させるべき配列対にどのアルゴリズム及びパラメータ設定を選択するのが適切かを理解するが、概して長さが同様で、期待されるアミノ酸同一性が 50% 超またはヌクレオチド同一性が 75% 超である配列の場合、N e e d l e m a n-W u n s c h アルゴリズムを、www.ebi.ac.uk ウェブサーバーにて E B I より提供されている N e e d l e m a n-W u n s c h アルゴリズムインターフェースのデフォルト設定で用いるのが一般に適切である。

【0049】

mRNA

いくつかの実施形態では、本明細書に開示する組成物または製剤は、本明細書に記載するような C a s ヌクレアーゼ、またはクラス 2 C a s ヌクレアーゼなどの RNA 誘導型 DNA 結合因子をコードするオープンリーディングフレーム（ORF）を含む mRNA を含む。いくつかの実施形態では、C a s ヌクレアーゼまたはクラス 2 C a s ヌクレアーゼなどの RNA 誘導型 DNA 結合因子をコードする ORF を含む mRNA を提供するか、使用するか、または投与する。いくつかの実施形態では、RNA 誘導型 DNA 結合因子をコードする ORF は、「修飾 RNA 誘導型 DNA 結合因子 ORF」または単に「修飾 ORF」であり、ORF が、以下の方法のうち 1 つ以上の方法で修飾されることを示す簡略表現として使用される：（1）修飾 ORF はウリジン含量が、その最小ウリジン含量から、かかる最小ウリジン含量の 150% までの範囲である、（2）修飾 ORF はウリジンジヌクレオチド含量が、その最小ウリジンジヌクレオチド含量から、かかる最小ウリジンジヌクレオチド含量の 150% までの範囲である、（3）修飾 ORF は、配列番号 1、4、7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または 66 のいずれか 1 つに対する同一性が少なくとも 90% である、（4）修飾 ORF は、コドンの少なくとも 75% が所与のアミノ酸についての最小ウリジンコドン（複数可）、例えば、ウリジンが最も少ないコドン（複数可）（最小ウリジンコドンは、ウリジンを 2 つ有するフェニルアラニンのコドンを除いては、通常 0 または 1 つである）である、1 セットのコドンからなる、または（5）修飾 ORF は少なくとも 1 つの修飾ウリジンを含む。いくつかの実施形態では、修飾 ORF は、上述の方法のうち少なくとも 2 つ、3 つ、または 4 つの方法で修飾される。いくつかの実施形態では、修飾 ORF は少なくとも 1 つの修飾ウリジンを含み、上記（1）～（4）のうち少なくとも 1 つ、2 つ、3 つ、またはすべての方法で修飾される。

【0050】

本明細書では「修飾ウリジン」は、水素結合受容体がウリジンの場合と同じであり、かつ、ウリジンとの構造上の違いが 1 つ以上ある、チミジン以外のヌクレオシドを指すために使用される。いくつかの実施形態では、修飾ウリジンは、置換ウリジン、すなわち、1 つ以上の非プロトン置換基（例えば、メトキシなどのアルコキシ）がプロトンに取って代わるウリジンである。いくつかの実施形態では、修飾ウリジンはプソイドウリジンである。いくつかの実施形態では、修飾ウリジンは、置換プソイドウリジン、すなわち、1 つ以

上の非プロトン置換基（例えば、メチルなどのアルキル）がプロトンに取って代わるプソイドウリジンである。いくつかの実施形態では、修飾ウリジンは、置換ウリジン、プソイドウリジン、または置換プソイドウリジンのいずれかである。

【0051】

本明細書で使用する「ウリジン位置」とは、ポリヌクレオチドにおいてウリジンまたは修飾ウリジンが占めている位置を指す。したがって、例えば、「ウリジン位置の100%が修飾ウリジン」であるポリヌクレオチドは、それと同じ配列の従来のRNA（この場合、塩基はいずれも標準塩基のA、U、C、またはGである）ではウリジンであるはずのどの位置においても修飾ウリジンを含有する。特に明記しない限り、本開示内または本開示に添付の配列表（sequence table）もしくは配列表（sequence listing）に記載のポリヌクレオチド配列におけるUはウリジンまたは修飾ウリジンであり得る。

【表1】

表1. 最小ウリジンコドン

	アミノ酸	最小ウリジンコドン
A	アラニン	GCAまたはGCCまたはGCG
G	グリシン	GGAまたはGGCまたはGGG
V	バリン	GUCまたはGUAまたはGUG
D	アスパラギン酸	GAC
E	グルタミン酸	GAAまたはGAG
I	イソロイシン	AUCまたはAUAまたはAUG
T	スレオニン	ACAまたはACCまたはACG
N	アスパラギン	AAC
K	リジン	AGGまたはAAA
S	セリン	AGC
R	アルギニン	AGAまたはAGG
L	ロイシン	CUGまたはCUAまたはCUC
P	プロリン	CCGまたはCCAまたはCCC
H	ヒスチジン	CACまたはCAAまたはCAG
Q	グルタミン	CAGまたはCAA
F	フェニルアラニン	UUC
Y	チロシン	UAC
C	システイン	UGC
W	トリプトファン	UGG
M	メチオニン	AUG

【0052】

上述の実施形態のいずれにおいても、修飾ORFは、コドンのうち少なくとも75%、80%、85%、90%、95%、98%、99%、または100%が最小ウリジンコドンの表に記載のコドンである、1セットのコドンで構成されてよい。上述の実施形態のいずれにおいても、修飾ORFは、配列番号1、4、7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または66のいずれか1つに対する同一性が少なくとも90%、95%、98%、99%、または100%である配列を含んでよい。

【0053】

上述の実施形態のいずれにおいても、修飾ORFは、配列番号1、4、7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または66のいずれか1つに対する同一性が少なくとも90%、95%、98%、99%、または100%である配列を含んでよい。

【0054】

上述の実施形態のいずれにおいても、修飾ORFは、ウリジン含量が、その最小ウリジン含量から、かかる最小ウリジン含量の150%、145%、140%、135%、130%、125%、120%、115%、110%、105%、104%、103%、102%、または101%までの範囲であってよい。

10

【0055】

上述の実施形態のいずれにおいても、修飾ORFは、ウリジンジヌクレオチド含量が、その最小ウリジンジヌクレオチド含量から、かかる最小ウリジンジヌクレオチド含量の150%、145%、140%、135%、130%、125%、120%、115%、110%、105%、104%、103%、102%、または101%までの範囲であってよい。

【0056】

上述の実施形態のいずれにおいても、修飾ORFはウリジン位置の少なくとも1つ、複数、またはすべてにおいて修飾ウリジンを含んでよい。いくつかの実施形態では、修飾ウリジンは、5位において、例えば、ハロゲン、メチル、またはエチルなどで修飾されたウリジンである。いくつかの実施形態では、修飾ウリジンは、1位において、例えば、ハロゲン、メチル、またはエチルなどで修飾されたプソイドウリジンである。修飾ウリジンは、例えば、プソイドウリジン、N1-メチルプソイドウリジン、5-メトキシウリジン、5-ヨードウリジン、またはその組み合わせであり得る。いくつかの実施形態では、修飾ウリジンは5-メトキシウリジンである。いくつかの実施形態では、修飾ウリジンは5-ヨードウリジンである。いくつかの実施形態では、修飾ウリジンはプソイドウリジンである。いくつかの実施形態では、修飾ウリジンはN1-メチルプソイドウリジンである。いくつかの実施形態では、修飾ウリジンはプソイドウリジンとN1-メチルプソイドウリジンの組み合わせである。いくつかの実施形態では、修飾ウリジンはプソイドウリジンと5-メトキシウリジンの組み合わせである。いくつかの実施形態では、修飾ウリジンはN1-メチルプソイドウリジンと5-メトキシウリジンの組み合わせである。いくつかの実施形態では、修飾ウリジンは5-ヨードウリジンとN1-メチルプソイドウリジンの組み合わせである。いくつかの実施形態では、修飾ウリジンはプソイドウリジンと5-ヨードウリジンの組み合わせである。いくつかの実施形態では、修飾ウリジンは5-ヨードウリジンと5-メトキシウリジンの組み合わせである。

20

30

【0057】

いくつかの実施形態では、本開示によるmRNAのウリジン位置のうち少なくとも10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、98%、99%、または100%は修飾ウリジンである。いくつかの実施形態では、本開示によるmRNAのウリジン位置のうち10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または90~100%は修飾ウリジン、例えば、5-メトキシウリジン、5-ヨードウリジン、N1-メチルプソイドウリジン、プソイドウリジン、またはその組み合わせである。いくつかの実施形態では、本開示によるmRNAのウリジン位置のうち10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または90~100%は5-メトキシウリジンである。いくつかの実施形態では、本開示によるmRNAのウリジン位置のうち10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または90~100%はプソイドウリジンである。いくつかの実施

40

50

形態では、本開示によるmRNAのウリジン位置のうち10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または90~100%はN1-メチルプソイドウリジンである。いくつかの実施形態では、本開示によるmRNAのウリジン位置のうち10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または90~100%は5-ヨードウリジンである。いくつかの実施形態では、本開示によるmRNAのウリジン位置のうち10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または90~100%は5-メトキシウリジンであり、残りのウリジン位置はN1-メチルプソイドウリジンである。いくつかの実施形態では、本開示によるmRNAのウリジン位置のうち10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または90~100%は5-ヨードウリジンであり、残りのウリジン位置はN1-メチルプソイドウリジンである。

【0058】

上述の実施形態のいずれにおいても、修飾ORFは、可能な限り低いウリジンジヌクレオチド(UU)含量など、低いウリジンジヌクレオチド含量を含んでよく、例えば、(a)どの位置にも最小ウリジンコドン(上述)を使用し、(b)所与のORFと同じアミノ酸配列をコードするORFである。ウリジンジヌクレオチド(UU)含量は、ORF内のUUジヌクレオチドの計数として絶対値で表すか、またはウリジンジヌクレオチドのウリジンが占めている位置の割合として比率で表すことができる(例えば、AUUAUの場合であれば、5つの位置のうち2つがウリジンジヌクレオチドのウリジンで占められているのでウリジンジヌクレオチド含量は40%となる)。最小ウリジンジヌクレオチド含量を評価するために、修飾ウリジン残基はウリジンと同等であるとみなす。

【0059】

いくつかの実施形態では、mRNAは、発現された哺乳類mRNA、例えば、構成的に発現しているmRNAなどに由来するUTRを少なくとも1つ含む。mRNAは、健常成体哺乳類の少なくとも1つの組織で継続的に転写されている場合、哺乳類で構成的に発現されているとみなされる。いくつかの実施形態では、mRNAは、発現された哺乳類のRNA、例えば、構成的に発現された哺乳類のmRNAなどに由来する5'UTR、3'UTR、または5'と3'の両UTRを含む。アクチンmRNAは構成的に発現しているmRNAの一例である。

【0060】

いくつかの実施形態では、mRNAは、17-ベータヒドロキシステロイドデヒドロゲナーゼ4(HSD17B4またはHSD)由来の少なくとも1つのUTR、例えば、HSD由来の5'UTRなどを含む。いくつかの実施形態では、mRNAは、グロビンmRNA、例えば、ヒトアルファグロビン(HBA)mRNA、ヒトベータグロビン(HBB)mRNA、またはアフリカツメガエル(Xenopus laevis)ベータグロビン(XBG)mRNAなどに由来する少なくとも1つのUTRを含む。いくつかの実施形態では、mRNAは、グロビンmRNA、例えば、HBA、HBB、またはXBGなどに由来する5'UTR、3'UTR、または5'と3'の両UTRを含む。いくつかの実施形態では、mRNAは、ウシ成長ホルモン、サイトメガロウイルス(CMV)、マウスHba-a1、HSD、アルブミン遺伝子、HBA、HBB、またはXBGに由来する5'UTRを含む。いくつかの実施形態では、mRNAは、ウシ成長ホルモン、サイトメガロウイルス、マウスHba-a1、HSD、アルブミン遺伝子、HBA、HBB、またはXBGに由来する3'UTRを含む。いくつかの実施形態では、mRNAは、ウシ成長ホルモン、サイトメガロウイルス、マウスHba-a1、HSD、アルブミン遺伝子、HBA、HBB、XBG、熱ショックタンパク質90(Hsp90)、グリセルアルデヒド3-リン酸デヒドロゲナーゼ(GAPDH)、ベーターアクチン、アルファチューブリン、腫瘍タンパク質(p53)、または上皮成長因子受容体(EGFR)に由来する5'と3'の両UTRを含

む。

【0061】

いくつかの実施形態では、mRNAは、同じ供給源に由来する5'と3'の両UTR、例えば、アクチン、アルブミン、またはグロビン、例えば、HBA、HBB、もしくはXBGなどのような構成的に発現しているmRNAなどに由来する5'と3'の両UTRを含む。

【0062】

いくつかの実施形態では、mRNAは5'UTRを含まず、例えば、5'キャップと開始コドンの間にはさらなるヌクレオチドはない。いくつかの実施形態では、mRNAは5'キャップと開始コドンの間にコザック配列（以下に記載）を含むが、さらなる5'UTRは何もない。いくつかの実施形態では、mRNAは3'UTRを含まず、例えば、終止コドンとポリA尾部の間にはさらなるヌクレオチドはない。

10

【0063】

いくつかの実施形態では、mRNAはコザック配列を含む。コザック配列は、翻訳開始及びmRNAから翻訳されたポリペプチドの全体としての収率に影響を及ぼし得る。コザック配列には、開始コドンとして機能し得るメチオニンコドンが含まれる。最小コザック配列はNNNRUGNであり、以下のうち少なくとも1つが真である、すなわち、1番目のNはAまたはGであり、かつ2番目のNはGである。ヌクレオチド配列においては、Rはプリン（AまたはG）を意味する。いくつかの実施形態では、コザック配列はRNNRUGN、NNNRUGG、RNNRUGG、RNN AUGN、NNNAUGG、またはRNN AUGGである。いくつかの実施形態では、コザック配列はrccRUGgであり、ミスマッチがゼロであるか、または小文字の位置に対するミスマッチを最高で1つもしくは2つ有する。いくつかの実施形態では、コザック配列はrccAUGgであり、ミスマッチがゼロであるか、または小文字の位置に対するミスマッチを最高で1つもしくは2つ有する。いくつかの実施形態では、コザック配列はgccRccAUGGであり、ミスマッチがゼロであるか、または小文字の位置に対するミスマッチを最高で1つ、2つ、もしくは3つ有する。いくつかの実施形態では、コザック配列はgccAccAUGであり、ミスマッチがゼロであるか、または小文字の位置に対するミスマッチを最高で1つ、2つ、3つ、もしくは4つ有する。いくつかの実施形態では、コザック配列はGCCACCAUGである。いくつかの実施形態では、コザック配列はgccgccRccAUGGであり、ミスマッチがゼロであるか、または小文字の位置に対するミスマッチを最高で1つ、2つ、3つ、もしくは4つ有する。

20

30

【0064】

いくつかの実施形態では、RNA誘導型DNA結合因子をコードするORFを含むmRNAは、配列番号43に対する同一性が少なくとも90%である配列を含み、ここで、任意選択で、配列番号43のORF（すなわち、配列番号4）は、配列番号7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または66のいずれか1つの代替的ORFで置換される。

【0065】

いくつかの実施形態では、RNA誘導型DNA結合因子をコードするORFを含むmRNAは、配列番号44に対する同一性が少なくとも90%である配列を含み、ここで、任意選択で、配列番号44のORF（すなわち、配列番号4）は、配列番号7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または66のいずれか1つの代替的ORFで置換される。

40

【0066】

いくつかの実施形態では、RNA誘導型DNA結合因子をコードするORFを含むmRNAは、配列番号56に対する同一性が少なくとも90%である配列を含み、ここで、任意選択で、配列番号56のORF（すなわち、配列番号4）は、配列番号7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または66のいずれか1つの代替的ORFで置換される。

【0067】

50

いくつかの実施形態では、RNA誘導型DNA結合因子をコードするORFを含むmRNAは、配列番号57に対する同一性が少なくとも90%である配列を含み、ここで、任意選択で、配列番号57のORF（すなわち、配列番号4）は、配列番号7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または66のいずれか1つの代替的ORFで置換される。
【0068】

いくつかの実施形態では、RNA誘導型DNA結合因子をコードするORFを含むmRNAは、配列番号58に対する同一性が少なくとも90%である配列を含み、ここで、任意選択で、配列番号58のORF（すなわち、配列番号4）は、配列番号7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または66のいずれか1つの代替的ORFで置換される。
【0069】

いくつかの実施形態では、RNA誘導型DNA結合因子をコードするORFを含むmRNAは、配列番号59に対する同一性が少なくとも90%である配列を含み、ここで、任意選択で、配列番号59のORF（すなわち、配列番号4）は、配列番号7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または66のいずれか1つの代替的ORFで置換される。
【0070】

いくつかの実施形態では、RNA誘導型DNA結合因子をコードするORFを含むmRNAは、配列番号60に対する同一性が少なくとも90%である配列を含み、ここで、任意選択で、配列番号60のORF（すなわち、配列番号4）は、配列番号7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または66のいずれか1つの代替的ORFで置換される。
【0071】

いくつかの実施形態では、RNA誘導型DNA結合因子をコードするORFを含むmRNAは、配列番号61に対する同一性が少なくとも90%である配列を含み、ここで、任意選択で、配列番号61のORF（すなわち、配列番号4）は、配列番号7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または66のいずれか1つの代替的ORFで置換される。
【0072】

いくつかの実施形態では、mRNAは、配列番号7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または66のいずれか1つの代替的ORFを含む。

【0073】

いくつかの実施形態では、任意選択で置換された、配列番号43、44、または56～61の配列に対する同一性の程度は95%である。いくつかの実施形態では、任意選択で置換された、配列番号43、44、または56～61の配列に対する同一性の程度は98%である。いくつかの実施形態では、任意選択で置換された、配列番号43、44、または56～61の配列に対する同一性の程度は99%である。いくつかの実施形態では、任意選択で置換された、配列番号43、44、または56～61の配列に対する同一性の程度は100%である。

【0074】

いくつかの実施形態では、本明細書に開示するmRNAは、5'キャップ、例えば、Cap0、Cap1、またはCap2などを含む。5'キャップは一般に、mRNAの5'から3'方向への鎖の1番目のヌクレオチドの5'位に5'-三リン酸を介して連結されている7-メチルグアニンリボヌクレオチド（これを、以下に考察するように、例えば、ARCAに関してさらに修飾してよい）であり、すなわち、第1のキャップ近傍ヌクレオチドである。Cap0では、mRNAの第1及び第2のキャップ近傍ヌクレオチドのリボースはいずれも2'-ヒドロキシルを含む。Cap1では、mRNAの第1及び第2の転写されたヌクレオチドのリボースはそれぞれ、2'-メトキシ及び2'-ヒドロキシルを含む。Cap

10

20

30

40

50

2では、mRNAの第1及び第2のキャップ近傍ヌクレオチドのリボースはいずれも2'-メトキシを含む。例えば、Katibah et al. (2014) Proc Natl Acad Sci USA 111 (33):12025-30、Abbas et al. (2017) Proc Natl Acad Sci USA 114 (11):E2106-E2115を参照のこと。ヒトmRNAなど哺乳類mRNAを含め、高等真核生物のほとんどの内在性mRNAはCap1またはCap2を含む。Cap0、ならびにCap1及びCap2とは異なる他のキャップ構造は、ヒトなどの哺乳類ではIFIT-1及びIFIT-5などの自然免疫系の成分によって「非自己」として認識されるために免疫原性である場合があり、これによりI型インターフェロンを含めサイトカインレベルが上昇し得る。IFIT-1及びIFIT-5などの自然免疫系の成分はまた、Cap1またはCap2以外のキャップを有するmRNAの結合についてeIF4Eと競合する場合もあり、mRNAの翻訳を阻害する可能性がある。

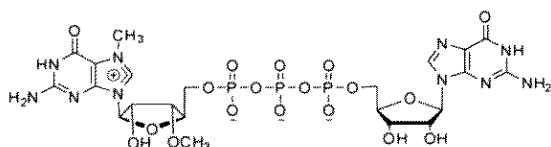
10

【0075】

キャップは共転写によって含めることができる。例えば、ARCA（アンチリバースキャップアナログ；Thermo Fisher Scientificカタログ番号AM8045）は、グアニンリボヌクレオチドの5'位に連結された7-メチルグアニン3'-メトキシ-5'-三リン酸を含む、インビトロでの転写開始時に転写産物内に組み込まれ得るキャップアナログである。ARCAでは、第1キャップ近傍ヌクレオチドの2'位がヒドロキシルであるCap0のキャップになる。例えば、Stepinski et al., (2001) "Synthesis and properties of mRNAs containing the novel 'anti-reverse' cap analogs 7-methyl (3'-O-methyl) GpppG and 7-methyl (3'-deoxy) GpppG," RNA 7:1486-1495を参照のこと。ARCAの構造を以下に示す。

20

【化1】



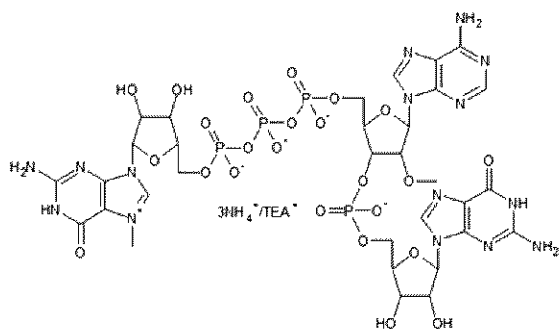
30

【0076】

CleanCap（商標）AG (m7G (5') ppp (5') (2' OMeA) pG; Trilink Biotechnologiesカタログ番号N-7113) またはCleanCap（商標）GG (m7G (5') ppp (5') (2' OMeG) pG; Trilink Biotechnologiesカタログ番号N-7133) を使用すると、共転写によってCap1構造を得ることができる。CleanCap（商標）AG及びCleanCap（商標）GGの3'-O-メチル化型もそれぞれTrilink Biotechnologiesよりカタログ番号N-7413及びN-7433として入手可能である。CleanCap（商標）AG構造を以下に示す。

【化2】

40



50

【0077】

別法として、転写後にRNAにキャップを付加することができる。例えば、ワクシニアのキャッピング酵素が市販されているが（New England Biolabsカタログ番号M2080S）、これは、そのD1サブユニットにより与えられるRNAトリホスファターゼ活性及びグアニリルトランスフェラーゼ活性、ならびにそのD12サブユニットにより与えられるグアニンメチルトランスフェラーゼを有する。したがって、この酵素は、S-アデノシルメチオニン及びGTPの存在下、Cap0を与えるようRNAに7-メチルグアニンを付加することができる。例えば、Guo, P. and Moss, B. (1990) Proc. Natl. Acad. Sci. USA 87, 4023-4027、Mao, X. and Shuman, S. (1994) J. Biol. Chem. 269, 24472-24479を参照のこと。

10

【0078】

いくつかの実施形態では、mRNAはさらにポリアデニル化（ポリA）尾部を含む。いくつかの実施形態では、ポリA尾部は少なくとも20個、30個、40個、50個、60個、70個、80個、90個、または100個のアデニン、任意選択で最高300個のアデニンを含む。いくつかの実施形態では、ポリA尾部は、95個、96個、97個、98個、99個、または100個のアデニンヌクレオチドを含む。いくつかの例では、ポリA尾部は、ポリA尾部内の1つ以上の位置において1個以上の非アデニンヌクレオチド「アンカー」で「中断」される。ポリA尾部は少なくとも8個の連続するアデニンヌクレオチドを含んでよいが、1個以上の非アデニンヌクレオチドも含む。本明細書で使用する場合、「非アデニンヌクレオチド」とは、アデニンを含まない任意の天然または非天然のヌクレオチドを指す。グアニン、チミン、及びシトシンヌクレオチドは例示的な非アデニンヌクレオチドである。したがって、本明細書に記載するmRNAのポリA尾部は、RNA誘導型DNA結合因子または目的配列をコードするヌクレオチドの3'に位置する連続するアデニンヌクレオチドを含んでよい。いくつかの例では、mRNAのポリA尾部は、RNA誘導型DNA結合因子または目的配列をコードするヌクレオチドの3'に位置する非連続的なアデニンヌクレオチドを含み、ここで、非アデニンヌクレオチドによりアデニンヌクレオチドが規則的または不規則に空いた間隔で中断される。

20

【0079】

いくつかの実施形態では、mRNAはさらにポリアデニル化（ポリA）尾部を含む。いくつかの実施形態では、ポリA尾部は少なくとも20個、30個、40個、50個、60個、70個、80個、90個、または100個のアデニン、任意選択で最高300個のアデニンを含む。いくつかの実施形態では、ポリA尾部は、95個、96個、97個、98個、99個、または100個のアデニンヌクレオチドを含む。いくつかの例では、ポリA尾部は、ポリA尾部内の1つ以上の位置において1個以上の非アデニンヌクレオチド「アンカー」で「中断」される。ポリA尾部は少なくとも8個の連続するアデニンヌクレオチドを含んでよいが、1個以上の非アデニンヌクレオチドも含む。本明細書で使用する場合、「非アデニンヌクレオチド」とは、アデニンを含まない任意の天然または非天然のヌクレオチドを指す。グアニン、チミン、及びシトシンヌクレオチドは例示的な非アデニンヌクレオチドである。したがって、本明細書に記載するmRNAのポリA尾部は、RNA誘導型DNA結合因子または目的配列をコードするヌクレオチドの3'に位置する連続するアデニンヌクレオチドを含んでよい。いくつかの例では、mRNAのポリA尾部は、RNA誘導型DNA結合因子または目的配列をコードするヌクレオチドの3'に位置する非連続的なアデニンヌクレオチドを含み、ここで、非アデニンヌクレオチドによりアデニンヌクレオチドが規則的または不規則に空いた間隔で中断される。

30

40

【0080】

いくつかの実施形態では、連続するアデニンヌクレオチドの長く伸びた領域にポリ（A）結合タンパク質が結合できるよう、連続するアデニンヌクレオチドを中断するため1個以上の非アデニンヌクレオチドを配する。いくつかの実施形態では、少なくとも8個、9個、10個、11個、または12個連続するアデニンヌクレオチドの後に1個以上の非ア

50

デニンヌクレオチド（複数可）を位置させる。いくつかの実施形態では、少なくとも8～50個連続するアデニンヌクレオチドの後に1個以上の非アデニンヌクレオチドを位置させる。いくつかの実施形態では、少なくとも8～100個連続するアデニンヌクレオチドの後に1個以上の非アデニンヌクレオチドを位置させる。いくつかの実施形態では、非アデニンヌクレオチドは1個、2個、3個、4個、5個、6個、または7個のアデニンヌクレオチドの後にあり、その後に少なくとも8個連続するアデニンヌクレオチドが続く。

【0081】

ポリA尾部は、連続するアデニンヌクレオチドの配列を1つ、その後に1個以上の非アデニンヌクレオチド、任意選択でその後にさらなるアデニンヌクレオチドを含んでよい。

【0082】

いくつかの実施形態では、ポリA尾部は、非アデニンヌクレオチドを1個、または非アデニンヌクレオチドが2～10個連続する長い領域を1つ含むかまたは含有する。いくつかの実施形態では、少なくとも8個、9個、10個、11個、または12個連続するアデニンヌクレオチドの後に非アデニンヌクレオチド（複数可）を位置させる。いくつかの例では、少なくとも8～50個連続するアデニンヌクレオチドの後に1個以上の非アデニンヌクレオチドを位置させる。いくつかの実施形態では、少なくとも8個、9個、10個、11個、12個、13個、14個、15個、16個、17個、18個、19個、20個、21個、22個、23個、24個、25個、26個、27個、28個、29個、30個、31個、32個、33個、34個、35個、36個、37個、38個、39個、40個、41個、42個、43個、44個、45個、46個、47個、48個、49個、または50個連続するアデニンヌクレオチドの後に1個以上の非アデニンヌクレオチドを位置させる。

【0083】

いくつかの実施形態では、非アデニンヌクレオチドはグアニン、シトシン、またはチミンである。いくつかの例では、非アデニンヌクレオチドはグアニンヌクレオチドである。いくつかの実施形態では、非アデニンヌクレオチドはシトシンヌクレオチドである。いくつかの実施形態では、非アデニンヌクレオチドはチミンヌクレオチドである。2個以上の非アデニンヌクレオチドが存在するいくつかの例では、非アデニンヌクレオチドは、a) グアニンヌクレオチド及びチミンヌクレオチド、b) グアニンヌクレオチド及びシトシンヌクレオチド、c) チミンヌクレオチド及びシトシンヌクレオチド、またはd) グアニンヌクレオチド、チミンヌクレオチド及びシトシンヌクレオチドから選択してよい。非アデニンヌクレオチドを含む例示的なポリA尾部を配列番号62として記載する。

【0084】

いくつかの実施形態では、mRNAを精製する。いくつかの実施形態では、沈殿方法（例えば、LiCl沈殿、アルコール沈殿など、または同等の方法、例えば、本明細書に記載の方法など）を使用してmRNAを精製する。いくつかの実施形態では、HPLCを用いる方法など、クロマトグラフィーを用いる方法または同等の方法（例えば、本明細書に記載の方法）を使用してmRNAを精製する。いくつかの実施形態では、沈殿方法（例えば、LiCl沈殿）及びHPLCを用いる方法の両方を使用してmRNAを精製する。

【0085】

いくつかの実施形態では、本明細書に開示のmRNAと組み合わせて少なくとも1つのgRNAを提供する。いくつかの実施形態では、gRNAは、mRNAとは別の分子として提供される。いくつかの実施形態では、gRNAは、本明細書に開示のmRNAの一部として、例えば、UTRの一部として提供される。

【0086】

化学修飾gRNA

いくつかの実施形態では、gRNAを化学的に修飾する。1つ以上の修飾されたヌクレオチドまたはヌクレオチドを含むgRNAは「修飾」gRNAまたは「化学修飾」gRNAと呼ばれ、標準残基A、G、C、及びUの代わりに使用されるかまたは追加で使用される天然には生じない成分もしくは構成及び／または天然に生じる成分もしくは構成が1つ

10

20

30

40

50

以上存在することを表す。いくつかの実施形態では、修飾 gRNA は非標準のヌクレオシドまたはヌクレオチドを用いて合成され、これを本明細書では「修飾」と呼ぶ。修飾ヌクレオシド及び修飾ヌクレオチドには、(i) ホスホジエステル骨格結合におけるリン酸基の非連結酸素のうち的一方もしくは両方及び／またはリン酸基の連結酸素のうち1つ以上の変化、例えば、置換（例示的な骨格修飾）、(ii) リボース糖の構成要素、例えば、リボース糖の2'位のヒドロキシルなどの変化、例えば、置換（例示的な糖修飾）、(iii) 「デホスホ」リンカーでのリン酸部分の大幅な置換（例示的な骨格修飾）、(iv) 非標準核酸塩基を用いるなど、天然に生じる核酸塩基の修飾または置換（例示的な塩基修飾）、(v) リボース-リン酸骨格の置換または修飾（例示的な骨格修飾）、(vi) オリゴヌクレオチドの3'末端または5'末端の修飾、例えば、末端リン酸基を除去、修飾もしくは置換するか、または部分、キャップもしくはリンカーを結合させる修飾（そのような3'または5'のキャップ修飾は糖修飾及び／または骨格修飾を含んでよい）、及び(vii) 糖の修飾または置換（例示的な糖修飾）のうち1つ以上が含まれ得る。

【0087】

いくつかの実施形態では、gRNAは、一部または全部のウリジン位置において修飾ウリジンを含む。いくつかの実施形態では、修飾ウリジンは、5位において、例えば、ハロゲンまたはC1-C6アルコキシなどで修飾されたウリジンである。いくつかの実施形態では、修飾ウリジンは、1位において、例えば、C1-C6アルキルなどで修飾されたプソイドウリジンである。修飾ウリジンは、例えば、プソイドウリジン、N1-メチルプソイドウリジン、5-メトキシウリジン、5-ヨードウリジン、またはその組み合わせであり得る。いくつかの実施形態では、修飾ウリジンは5-メトキシウリジンである。いくつかの実施形態では、修飾ウリジンは5-ヨードウリジンである。いくつかの実施形態では、修飾ウリジンはプソイドウリジンである。いくつかの実施形態では、修飾ウリジンはN1-メチルプソイドウリジンである。いくつかの実施形態では、修飾ウリジンはプソイドウリジンとN1-メチルプソイドウリジンの組み合わせである。いくつかの実施形態では、修飾ウリジンはプソイドウリジンと5-メトキシウリジンの組み合わせである。いくつかの実施形態では、修飾ウリジンはN1-メチルプソイドウリジンと5-メトキシウリジンの組み合わせである。いくつかの実施形態では、修飾ウリジンは5-ヨードウリジンとN1-メチルプソイドウリジンの組み合わせである。いくつかの実施形態では、修飾ウリジンはプソイドウリジンと5-ヨードウリジンの組み合わせである。いくつかの実施形態では、修飾ウリジンは5-ヨードウリジンと5-メトキシウリジンの組み合わせである。

【0088】

いくつかの実施形態では、本開示によるgRNAのウリジン位置のうち少なくとも10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、98%、99%、または100%は修飾ウリジンである。いくつかの実施形態では、本開示によるgRNAのウリジン位置のうち10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または90~100%は修飾ウリジン、例えば、5-メトキシウリジン、5-ヨードウリジン、N1-メチルプソイドウリジン、プソイドウリジン、またはその組み合わせである。いくつかの実施形態では、本開示によるgRNAのウリジン位置のうち10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または90~100%は5-メトキシウリジンである。いくつかの実施形態では、本開示によるgRNAのウリジン位置のうち10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または90~100%はプソイドウリジンである。いくつかの実施形態では、本開示によるgRNAのウリジン位置のうち10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または90~100%はN1-メチルプソイドウリジンである。いく

つかの実施形態では、本開示による gRNA のウリジン位置のうち 10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または 90~100% は 5-ヨードウリジンである。いくつかの実施形態では、本開示による gRNA のウリジン位置のうち 10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または 90~100% は 5-メトキシウリジンであり、残りのウリジン位置は N1-メチルプソイドウリジンである。いくつかの実施形態では、本開示による gRNA のウリジン位置のうち 10%~25%、15~25%、25~35%、35~45%、45~55%、55~65%、65~75%、75~85%、85~95%、または 90~100% は 5-ヨードウリジンであり、残りのウリジン位置は N1-メチルプソイドウリジンである。

10

【0089】

上掲のような化学修飾を組み合わせ、2つ、3つ、4つ、またはそれ以上の修飾を有し得るヌクレオシド及びヌクレオチド（総称して「残基」と呼ぶ）を含む修飾 gRNA を得ることができる。例えば、修飾残基は修飾糖及び修飾核酸塩基を有し得る。いくつかの実施形態では、gRNA のどの塩基も修飾する、例えば、塩基すべてがホスホロチオアート基などの修飾リン酸基を有する。特定の実施形態では、gRNA 分子のリン酸基の全部または実質的に全部をホスホロチオアート基で置き換える。いくつかの実施形態では、修飾 gRNA は、RNA の 5' 末端またはその近傍に少なくとも 1 つの修飾残基を含む。いくつかの実施形態では、修飾 gRNA は、RNA の 3' 末端またはその近傍に少なくとも 1 つの修飾残基を含む。

20

【0090】

いくつかの実施形態では、gRNA は修飾残基を 1 つ、2 つ、3 つ、またはそれ以上含む。いくつかの実施形態では、修飾 gRNA の位置のうち少なくとも 5%（例えば、少なくとも 5%、少なくとも 10%、少なくとも 15%、少なくとも 20%、少なくとも 25%、少なくとも 30%、少なくとも 35%、少なくとも 40%、少なくとも 45%、少なくとも 50%、少なくとも 55%、少なくとも 60%、少なくとも 65%、少なくとも 70%、少なくとも 75%、少なくとも 80%、少なくとも 85%、少なくとも 90%、少なくとも 95%、または 100%）は修飾されたヌクレオシドまたはヌクレオチドである。

【0091】

非修飾核酸は、例えば、細胞内または血清中に見られるヌクレアーゼなどによって分解される傾向があり得る。例えば、ヌクレアーゼは核酸のホスホジエステル結合を加水分解することができる。したがって、一態様では、本明細書に記載する gRNA は、1 つ以上の修飾されたヌクレオシドまたはヌクレオチドを含有して、例えば、細胞内または血清中のヌクレアーゼに対する安定性を導入することができる。いくつかの実施形態では、本明細書に記載の修飾 gRNA 分子は、インピボでもエキソピボでも細胞集団に導入した場合に低い自然免疫応答を示し得る。用語「自然免疫応答」には、一本鎖核酸などの外来性核酸に対する細胞応答が含まれ、サイトカインの発現及び放出、特にインターフェロンの放出、ならびに細胞死の誘導を行う。

30

【0092】

骨格修飾のいくつかの実施形態では、修飾残基のリン酸基は、酸素の 1 つ以上を異なる置換基で置き換えることによって修飾することができる。さらに、修飾残基、例えば、修飾核酸に存在する修飾残基には、本明細書に記載する修飾リン酸基での非修飾リン酸部分の大幅な置換が含まれ得る。いくつかの実施形態では、リン酸骨格の骨格修飾には、非荷電リンカーまたは電荷分布が非対称的な荷電リンカーのいずれかをもたらす変化が含まれ得る。

40

【0093】

修飾リン酸基の例としては、ホスホロチオアート、ホスホロセレナート、ボラノリン酸、ボラノリン酸エステル、ホスホン酸水素、ホスホロアミダート、ホスホン酸アルキルまたはホスホン酸アリール及びホストリエステルが挙げられる。非修飾リン酸基のリン原

50

子はアキラルである。ただし、非架橋酸素のうちの1個を上記の原子または原子団の1つで置き換えることにより、そのリン原子をキラルにすることができる。不斉リン原子は、「R」立体配置（本明細書ではR p）または「S」立体配置（本明細書ではS p）のいずれも有することができる。骨格はまた、架橋酸素（すなわち、リン酸基とヌクレオシドを連結する酸素）を窒素（架橋ホスホロアミダート）、硫黄（架橋ホスホロチオアート）及び炭素（架橋メチレンホスホナート）で置き換えることによって修飾することもできる。置き換えは、連結酸素のいずれか、または連結酸素の両方において出現してよい。

【0094】

特定の骨格修飾では、リン酸基をリン不含のコネクタで置き換えることができる。いくつかの実施形態では、荷電リン酸基を中性部分で置き換えることができる。リン酸基に換えることができる部分の例としては、限定することなく、例えば、メチルホスホナート、ヒドロキシルアミノ、シロキサン、カルボナート、カルボキシメチル、カルバマート、アミド、チオエーテル、エチレンオキシドリinker、スルホナート、スルホンアミド、チオホルムアセタール、ホルムアセタール、オキシム、メチレンイミノ、メチレンメチルイミノ、メチレンヒドラゾ、メチレンジメチルヒドラゾ及びメチレンオキシメチルイミノを挙げることができる。

【0095】

鋳型核酸

本明細書に開示する組成物及び方法には鋳型核酸が含まれ得る。鋳型を使用して、Casヌクレアーゼの標的部位またはその近傍にて核酸配列を変化させるかまたは挿入してよい。いくつかの実施形態では、方法は、細胞に鋳型を導入することを含む。いくつかの実施形態では、単一の鋳型が提供され得る。他の実施形態では、2つ以上の標的部位にて編集が生じるよう2つ以上の鋳型が提供され得る。例えば、ある細胞の単一遺伝子、またはある細胞の2つの異なる遺伝子を編集するため、異なる鋳型が提供され得る。

【0096】

いくつかの実施形態では、鋳型を相同組換えで使用してよい。いくつかの実施形態では、相同組換えにより、標的核酸分子内への鋳型配列または鋳型配列の一部の組み込みがもたらされ得る。他の実施形態では、鋳型を、核酸の切断部位でDNA鎖の侵入が起こる相同組換え修復で使用してよい。いくつかの実施形態では、相同組換え修復により、鋳型配列が、編集された標的核酸分子内に含まれるということがもたらされ得る。さらに別の実施形態では、非同源末端結合によって媒介される遺伝子編集で鋳型を使用してよい。いくつかの実施形態では、鋳型配列には、切断部位近傍の核酸配列に対する類似性は全くない。いくつかの実施形態では、鋳型または鋳型配列の一部を組み込む。いくつかの実施形態では、鋳型には、隣接する末端逆位反復（ITR）配列が含まれる。

【0097】

いくつかの実施形態では、鋳型は、切断部位のそれぞれ上流及び下流に位置する配列に相補的な、第1のホモロジーアーム及び第2のホモロジーアーム（第1のヌクレオチド配列及び第2のヌクレオチド配列とも呼ばれる）を含んでよい。鋳型が2つのホモロジーアームを含有する場合、各アームは長さが同じであっても異なってもよく、ホモロジーアームに挟まれた配列は、ホモロジーアームに挟まれた標的配列と実質的に類似であるかもしくは同一であり得る、または全く無関係の配列であり得る。いくつかの実施形態では、鋳型上の第1のヌクレオチド配列と切断部位の上流の配列との相補性または同一性パーセントの程度、及び鋳型上の第2のヌクレオチド配列と切断部位の下流の配列との相補性または同一性パーセントの程度により、鋳型と標的核酸分子の間の相同組換え、例えば、高忠実度相同組換えなどが可能になり得る。いくつかの実施形態では、相補性の程度は約50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、97%、98%、99%、または100%であってよい。いくつかの実施形態では、相補性の程度は約95%、97%、98%、99%、または100%であってよい。いくつかの実施形態では、相補性の程度は少なくとも98%、99%、または100%であってよい。いくつかの実施形態では、相補性の程度は100%であってよい。いくつかの実施形

10

20

30

40

50

態では、同一性パーセントは約 50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、97%、98%、99%、または 100% であってよい。いくつかの実施形態では、同一性パーセントは約 95%、97%、98%、99%、または 100% であってよい。いくつかの実施形態では、同一性パーセントは少なくとも 98%、99%、または 100% であってよい。いくつかの実施形態では、同一性パーセントは 100% であってよい。

【0098】

いくつかの実施形態では、鋳型配列は、標的細胞の内在性配列に対応するか、それを含むか、またはそれからなっている。さらに、または別法として、鋳型配列は、標的細胞の外来配列に対応するか、それを含むか、またはそれからなっている。本明細書で使用する場合、用語「内在性配列」とは、その細胞にとって天然である配列を指す。用語「外来配列」とは、その細胞にとって天然ではない配列、またはその細胞のゲノムにおける天然での位置が別の異なる位置にある配列を指す。いくつかの実施形態では、内在性配列は細胞のゲノム配列であってよい。いくつかの実施形態では、内在性配列は染色体の配列または染色体外の配列であってよい。いくつかの実施形態では、内在性配列は細胞のプラスミド配列であってよい。いくつかの実施形態では、鋳型配列は、切断部位またはその近傍において細胞の内在性配列の一部と実質的に同一であってよいが、少なくとも 1 つのヌクレオチドの変化を含む。いくつかの実施形態では、切断された標的核酸分子を鋳型で編集することにより、標的核酸分子の 1 つ以上のヌクレオチドの挿入、欠失、または置換を含む変異がもたらされ得る。いくつかの実施形態では、変異により、標的配列を含む遺伝子により発現されたタンパク質において 1 つ以上のアミノ酸の変化がもたらされ得る。いくつかの実施形態では、変異は、標的遺伝子により発現された RNA に 1 つ以上のヌクレオチドの変化をもたらし得る。いくつかの実施形態では、変異は標的遺伝子の発現レベルを変化させ得る。いくつかの実施形態では、変異は標的遺伝子の発現の増大または低下をもたらし得る。いくつかの実施形態では、変異は遺伝子ノックダウンをもたらし得る。いくつかの実施形態では、変異は遺伝子ノックアウトをもたらし得る。いくつかの実施形態では、変異は遺伝子機能の回復をもたらし得る。いくつかの実施形態では、切断された標的核酸分子を鋳型で編集することにより、DNA などの標的核酸分子エキソン配列、イントロン配列、制御配列、転写制御配列、翻訳制御配列、スプライシング部位、または非コード配列の変化がもたらされ得る。

【0099】

他の実施形態では、鋳型配列は外来配列を含んでよい。いくつかの実施形態では、外来配列は、外来性プロモーター配列に機能的に連結された、タンパク質または RNA をコードする配列を含んでよく、外来配列が標的核酸分子に組み込まれた際に、かかる組み込み配列でコードされたタンパク質または RNA を細胞が発現することができるようにしてよい。他の実施形態では、外来配列が標的核酸分子内に組み込まれた際、かかる組み込み配列の発現は内在性プロモーター配列によって制御されてよい。いくつかの実施形態では、外来配列は、タンパク質または該タンパク質の一部をコードする cDNA 配列を提供してよい。さらに別の実施形態では、外来配列は、エキソン配列、イントロン配列、制御配列、転写制御配列、翻訳制御配列、スプライシング部位、または非コード配列を含むかまたはそれからなっている。いくつかの実施形態では、外来配列の組み込みにより遺伝子機能の回復がもたらされ得る。いくつかの実施形態では、外来配列の組み込みにより遺伝子ノックインがもたらされ得る。いくつかの実施形態では、外来配列の組み込みにより遺伝子ノックアウトがもたらされ得る。

【0100】

鋳型は、適切な任意の長さであってよい。いくつかの実施形態では、鋳型は、10 ヌクレオチド長、15 ヌクレオチド長、20 ヌクレオチド長、25 ヌクレオチド長、50 ヌクレオチド長、75 ヌクレオチド長、100 ヌクレオチド長、150 ヌクレオチド長、200 ヌクレオチド長、500 ヌクレオチド長、1000 ヌクレオチド長、1500 ヌクレオチド長、2000 ヌクレオチド長、2500 ヌクレオチド長、3000 ヌクレオチド長、

3500ヌクレオチド長、4000ヌクレオチド長、4500ヌクレオチド長、5000ヌクレオチド長、5500ヌクレオチド長、6000ヌクレオチド長、またはそれ以上のヌクレオチド長を含んでよい。鋳型は一本鎖核酸であってよい。鋳型は二本鎖または部分的二本鎖の核酸であり得る。特定の実施形態では、一本鎖鋳型は、20ヌクレオチド長、30ヌクレオチド長、40ヌクレオチド長、50ヌクレオチド長、75ヌクレオチド長、100ヌクレオチド長、125ヌクレオチド長、150ヌクレオチド長、175ヌクレオチド長、または200ヌクレオチド長である。いくつかの実施形態では、鋳型は、標的配列（すなわち、「ホモロジーアーム」）を含む標的核酸分子の一部に相補的なヌクレオチド配列を含んでよい。いくつかの実施形態では、鋳型は、標的核酸分子上の切断部位の上流または下流に位置する配列に相補的なホモロジーアームを含んでよい。

10

【0101】

いくつかの実施形態では、鋳型は、末端逆位反復（ITR）配列が隣接して含有されるssDNAまたはdsDNAを含有する。いくつかの実施形態では、鋳型は、ベクター、プラスミド、ミニサークル、ナノサークル、またはPCR産物として提供される。

【0102】

核酸の精製

いくつかの実施形態では、核酸を精製する。いくつかの実施形態では、沈殿方法（例えば、LiCl沈殿、アルコール沈殿、または同等の方法、例えば、本明細書に記載の方法など）を使用して核酸を精製する。いくつかの実施形態では、HPLCを用いる方法など、クロマトグラフィーを用いる方法または同等の方法（例えば、本明細書に記載の方法）を使用して核酸を精製する。いくつかの実施形態では、沈殿方法（例えば、LiCl沈殿）及びHPLCを用いる方法の両方を使用して核酸を精製する。

20

【0103】

標的配列

いくつかの実施形態では、本開示のCRISPR/Cas系を標的核酸分子上の標的配列に配向させ、切断してよい。例えば、標的配列は、Casヌクレアーゼによって認識され、切断されてよい。特定の実施形態では、Casヌクレアーゼの標的配列は、かかるヌクレアーゼの特異的PAM配列の近傍に位置する。いくつかの実施形態では、gRNAによって標的核酸分子の標的配列にクラス2Casヌクレアーゼを配向させ、そこで、gRNAは標的配列とハイブリダイズし、クラス2Casタンパク質が標的配列を切断してよい。いくつかの実施形態では、ガイドRNAは、クラス2Casヌクレアーゼの特異的PAMに隣接するかまたはそれを含む標的配列とハイブリダイズし、クラス2Casヌクレアーゼはかかる標的配列を切断する。いくつかの実施形態では、標的配列は、ガイドRNAの指向性配列に相補的であってよい。いくつかの実施形態では、ガイドRNAの指向性配列と、それに対応する、ガイドRNAがハイブリダイズする標的配列の部分との相補性の程度は、約50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、97%、98%、99%、または100%であってよい。いくつかの実施形態では、ガイドRNAの指向性配列と、それに対応する、ガイドRNAがハイブリダイズする標的配列の部分との同一性パーセントは、約50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、97%、98%、99%、または100%であってよい。いくつかの実施形態では、標的の相同性領域は特異的PAM配列に隣接している。いくつかの実施形態では、標的配列は、ガイドRNAの指向性配列と100%相補的な配列を含んでよい。他の実施形態では、標的配列は、ガイドRNAの指向性配列と比べて少なくとも1つのミスマッチ、欠失、または挿入を含んでよい。

30

40

【0104】

標的配列の長さは使用するヌクレアーゼ系によって異なってよい。例えば、CRISPR/Cas系用のガイドRNAの指向性配列は、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、または50超のヌクレオチド長を含んでよく、標的配列は対応する長さであり、任意選択でPAM配列に隣接する。い

50

いくつかの実施形態では、標的配列は15～24ヌクレオチド長を含んでよい。いくつかの実施形態では、標的配列は17～21ヌクレオチド長を含んでよい。いくつかの実施形態では、標的配列は20ヌクレオチド長を含んでよい。ニッカーゼを使用する場合、標的配列は、DNA分子の反対鎖を切断する一対のニッカーゼによって認識される一対の標的配列を含んでよい。いくつかの実施形態では、標的配列は、DNA分子の同じ鎖を切断する一対のニッカーゼによって認識される一対の標的配列を含んでよい。いくつかの実施形態では、標的配列は、1つ以上のCasヌクレアーゼによって認識される標的配列の一部を含んでよい。

【0105】

標的核酸分子は、細胞にとって内在性または外来性であるどのDNA分子またはRNA分子であってもよい。いくつかの実施形態では、標的核酸分子は、細胞由来または細胞内のエピソームDNA、プラスミド、ゲノムDNA、ウイルスゲノム、ミトコンドリアDNA、または染色体DNAであってよい。いくつかの実施形態では、標的核酸分子の標的配列は、ヒト細胞などの細胞由来または細胞内のゲノム配列であってよい。

【0106】

さらなる実施形態では、標的配列はウイルス配列であってよい。さらなる実施形態では、標的配列は病原体配列であってよい。さらに別の実施形態では、標的配列は合成配列であってよい。さらなる実施形態では、標的配列は染色体配列であってよい。特定の実施形態では、標的配列は転座接合部、例えば、がんに関連した転座を含んでよい。いくつかの実施形態では、標的配列はヒト染色体など真核生物の染色体上にあってよい。特定の実施形態では、標的配列は肝臓特異的配列であり、その場合、その配列は肝細胞に発現する。

【0107】

いくつかの実施形態では、標的配列は遺伝子のコード配列内、遺伝子のイントロン配列内、制御配列内、遺伝子の転写制御配列内、遺伝子の翻訳制御配列内、スプライシング部位または遺伝子間の非コード配列内に位置してよい。いくつかの実施形態では、遺伝子は、タンパク質をコードする遺伝子であってよい。他の実施形態では、遺伝子はノンコーディングRNA遺伝子であってよい。いくつかの実施形態では、標的配列は疾患関連遺伝子の全部または一部を含んでよい。いくつかの実施形態では、標的配列は、ゲノムの非遺伝子機能部位、例えば、足場部位または遺伝子座制御領域のようなクロマチン組織の側面を制御する部位に位置してよい。

【0108】

クラス2CasヌクレアーゼなどのCasヌクレアーゼが関係する実施形態では、標的配列は、プロトスペーサー隣接モチーフ（「PAM」）に隣接してよい。いくつかの実施形態では、PAMは、標的配列の3'末端に隣接するかまたはその1ヌクレオチド、2ヌクレオチド、3ヌクレオチド、または4ヌクレオチド以内にあってよい。PAMの長さ及び配列は、使用するCasタンパク質によって異なってよい。例えば、PAMは、特定のCas9タンパク質またはCas9オルソログについてのコンセンサス配列または特定のPAM配列から選択してよく、これには、関連開示それぞれが参照により本明細書に組み込まれる、Ran et al., Nature, 520:186-191 (2015) の図1、及びZetsche 2015の図S5に開示のものが含まれる。いくつかの実施形態では、PAMは、2ヌクレオチド長、3ヌクレオチド長、4ヌクレオチド長、5ヌクレオチド長、6ヌクレオチド長、7ヌクレオチド長、8ヌクレオチド長、9ヌクレオチド長、または10ヌクレオチド長であってよい。非限定的な例示的PAM配列としては、NGG、NGGNG、NG、NAAAAN、NNAAA AW、NNNNACA、GNNNCNNA、TTN、及びNNNNGATTが挙げられる（ここで、Nは任意のヌクレオチドとして定義され、WはAまたはTのいずれかとして定義される）。いくつかの実施形態では、PAM配列はNGGであってよい。いくつかの実施形態では、PAM配列はNGGNGであってよい。いくつかの実施形態では、PAM配列はTTNであってよい。いくつかの実施形態では、PAM配列はNNAAA AWであってよい。

【0109】

脂質製剤

本明細書では、C R I S P R / C a s カーゴなど、RNAのL N P製剤のさまざまな実施形態を開示する。そのようなL N P製剤には、ヘルパー脂質、中性脂質、及びP E G脂質とともに「アミン脂質」が含まれる。いくつかの実施形態では、そのようなL N P製剤には、ヘルパー脂質及びP E G脂質とともに「アミン脂質」が含まれる。いくつかの実施形態では、L N P製剤には、1パーセント未満の中性リン脂質が含まれる。いくつかの実施形態では、L N P製剤には、0.5パーセント未満の中性リン脂質が含まれる。「脂質ナノ粒子」は、互いに分子間力によって物理的に会合している複数（すなわち2つ以上）の脂質分子を含む粒子を意味する。

【0110】

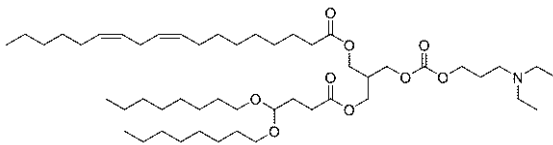
アミン脂質

生物学的に活性な薬剤を送達するためのL N P組成物は、リピドAのアセタール類似体など、リピドAまたはその同等物として定義される「アミン脂質」を含む。

【0111】

いくつかの実施形態では、アミン脂質はリピドAであり、これは(9Z, 12Z) - 3 - ((4, 4-ビス(オクチルオキシ)ブタノイル)オキシ) - 2 - (((3-(ジエチルアミノ)プロポキシ)カルボニル)オキシ)メチル)プロピルオクタデカ-9, 12-ジエノアートであり、3 - ((4, 4-ビス(オクチルオキシ)ブタノイル)オキシ) - 2 - (((3-(ジエチルアミノ)プロポキシ)カルボニル)オキシ)メチル)プロピル(9Z, 12Z) - オクタデカ-9, 12-ジエノアートとも呼ばれる。リピドAは、

【化3】



のように表され得る。

【0112】

リピドAはW O 2 0 1 5 / 0 9 5 3 4 0（例えば、84～86頁）に従って合成してよい。特定の実施形態では、アミン脂質はリピドAと同等物である。

【0113】

特定の実施形態では、アミン脂質はリピドAの類似体である。特定の実施形態では、リピドA類似体はリピドAのアセタール類似体である。特定のL N P組成物では、アセタール類似体はC 4 - C 1 2アセタール類似体である。いくつかの実施形態では、アセタール類似体はC 5 - C 1 2アセタール類似体である。さらなる実施形態では、アセタール類似体はC 5 - C 1 0アセタール類似体である。さらなる実施形態では、アセタール類似体は、C 4、C 5、C 6、C 7、C 9、C 1 0、C 1 1、及びC 1 2のアセタール類似体から選ばれる。

【0114】

本明細書に記載するL N Pでの使用に好適なアミン脂質はインビボで生物分解性であり、RNAなどの生物学的に活性な物質を細胞に送達するのに好適である。アミン脂質は毒性が低い（例えば、10mg/kg以上のRNAカーゴ量で動物モデルでの忍容性があり、有害作用がない）。特定の実施形態では、アミン脂質を含むL N Pには、アミン脂質の少なくとも75%が8時間以内、10時間以内、12時間以内、24時間以内、もしくは48時間以内、または3日以内、4日以内、5日以内、6日以内、7日以内、もしくは10日以内に血漿から消失するものが含まれる。特定の実施形態では、アミン脂質を含むL N Pには、mRNAまたはgRNAの少なくとも50%が8時間以内、10時間以内、12時間以内、24時間以内、もしくは48時間以内、または3日以内、4日以内、5日以内、6日以内、7日以内、もしくは10日以内に血漿から消失するものが含まれる。特定の実施形態では、アミン脂質を含むL N Pには、例えば、脂質（例えば、アミン脂質）、

10

20

30

40

50

RNA（例えば、mRNA）、または別の成分を測定して、LNPの少なくとも50%が8時間以内、10時間以内、12時間以内、24時間以内、もしくは48時間以内、または3日以内、4日以内、5日以内、6日以内、7日以内、もしくは10日以内に血漿から消失するものが含まれる。特定の実施形態では、LNPの脂質、RNA、または核酸成分を脂質に封入した場合と遊離の場合とで測定する。

【0115】

脂質クリアランスは、文献に記載のとおり測定してよい。Maier, M. A., et al. Biodegradable Lipids Enabling Rapidly Eliminated Lipid Nanoparticles for Systemic Delivery of RNAi Therapeutics. Mol. Ther. 2013, 21 (8), 1570-78 (「Maier」)を参照のこと。例えば、Maierでは、ルシフェラーゼ指向性siRNAを含有するLNP-siRNA系を0.3mg/kgにて6~8週齢雄C57BL/6マウスに側尾静脈を介して静脈内ボラス注射により投与した。投与から0.083時間、0.25時間、0.5時間、1時間、2時間、4時間、8時間、24時間、48時間、96時間、及び168時間の後、血液、肝臓、及び脾臓の各試料を採取した。マウスを生理食塩水で灌流してから組織を採取し、血液試料を処理して血漿を得た。全試料を処理し、LC-MSで分析した。さらに、Maierは、LNP-siRNA製剤投与後の毒性評価の手法を記載している。例えば、雄Sprague-Dawleyラットに対しルシフェラーゼ指向性siRNAを0mg/kg、1mg/kg、3mg/kg、5mg/kg、及び10mg/kg（動物5匹/群）にて5mL/kgの投与容量で単一静脈内ボラス注射により投与した。24時間後、約1mLの血液を覚醒動物の頸静脈から得、血清を単離した。投与から72時間後、全動物を剖検用に安楽死させた。臨床徴候、体重、血清化学検査、臓器重量及び組織病理学の評価を実施した。Maierは、siRNA-LNP製剤を評価するための方法を記載しており、これらの方法を本開示のLNP組成物の投与に関するクリアランス、薬物動態、及び毒性の評価に適用してよい。

【0116】

アミン脂質はクリアランス速度の増大を引き起こし得る。いくつかの実施形態では、クリアランス速度は、脂質クリアランス速度、例えば、血液、血清、または血漿から脂質が消失する速度である。いくつかの実施形態では、クリアランス速度は、RNAクリアランス速度、例えば、血液、血清、または血漿からmRNAまたはgRNAが消失する速度である。いくつかの実施形態では、クリアランス速度は、血液、血清、または血漿からLNPが消失する速度である。いくつかの実施形態では、クリアランス速度は、肝組織または脾臓組織などの組織からLNPが消失する速度である。特定の実施形態では、高クリアランス速度は実質的に有害作用のない安全性プロファイルをもたらす。アミン脂質は、循環中及び組織内でのLNPの蓄積を低下させ得る。いくつかの実施形態では、循環中及び組織内でのLNPの蓄積低下は、実質的に有害作用のない安全性プロファイルをもたらす。

【0117】

本開示のアミン脂質は、それが含まれている培地のpHに応じてイオン化され得る（例えば、塩を形成してよい）。例えば、わずかに酸性の培地では、アミン脂質はプロトン化されてよく、したがって正電荷を帯びてよい。逆に、例えば、血液などのようにpHが約7.35であるわずかに塩基性の培地ではアミン脂質はプロトン化されない場合があり、したがって電荷を帯びない場合がある。いくつかの実施形態では、本開示のアミン脂質は、pHが少なくとも約9でプロトン化され得る。いくつかの実施形態では、本開示のアミン脂質は、pHが少なくとも約9でプロトン化され得る。いくつかの実施形態では、本開示のアミン脂質は、pHが少なくとも約10でプロトン化され得る。

【0118】

アミン脂質が優勢にプロトン化されるpHは、その内因性pKaに関連している。いくつかの実施形態では、本開示のアミン脂質は各々独立して、pKaが約5.1~約7.4の範囲にあってよい。いくつかの実施形態では、本開示のアミン脂質は各々独立して、p

Kaが約5.5～約6.6の範囲にあってよい。いくつかの実施形態では、本開示のアミン脂質は各々独立して、pKaが約5.6～約6.4の範囲にあってよい。いくつかの実施形態では、本開示のアミン脂質は各々独立して、pKaが約5.8～約6.2の範囲にあってよい。例えば、本開示のアミン脂質は各々独立して、pKaが約5.8～約6.5の範囲にあってよい。pKaが約5.1～約7.4の範囲のカチオン性脂質はカーゴをインビボで、例えば、肝臓などに送達する際に有効であることが見出されていることから、アミン脂質のpKaはLNPの製剤化において考慮すべき重要な事項であり得る。さらに、pKaが約5.3～約6.4の範囲のカチオン性脂質はインビボで、例えば、腫瘍などに送達する際に有効であることが見出されている。例えば、WO2014/136086を参照のこと。

【0119】

さらなる脂質

本開示の脂質組成物での使用に好適な「中性脂質」としては、例えば、多種多様な中性脂質、非荷電脂質または両性イオン型脂質などが挙げられる。本開示での使用に好適な中性リン脂質の例としては、5-ヘプタデシルベンゼン-1,3-ジオール（レソルシン）、ジパルミトイルホスファチジルコリン（DPPC）、ジステアロイルホスファチジルコリン（DSPC）、ホスホコリン（DOPC）、ジミリストイルホスファチジルコリン（DMPC）、ホスファチジルコリン（PLPC）、1,2-ジステアロイル-sn-グリセロ-3-ホスホコリン（DAPC）、ホスファチジルエタノールアミン（PE）、卵ホスファチジルコリン（EPC）、ジラウリロイルホスファチジルコリン（dilaurylphosphatidylcholine）（DLPC）、ジミリストイルホスファチジルコリン（DMPC）、1-ミリストイル-2-パルミトイルホスファチジルコリン（MPPC）、1-パルミトイル-2-ミリストイルホスファチジルコリン（PMP C）、1-パルミトイル-2-ステアロイルホスファチジルコリン（PSPC）、1,2-ジアラキドイル-sn-グリセロ-3-ホスホコリン（DBPC）、1-ステアロイル-2-パルミトイルホスファチジルコリン（SPPC）、1,2-ジエイコセノイル-sn-グリセロ-3-ホスホコリン（DEPC）、パルミトイルオレオイルホスファチジルコリン（POPC）、リゾホスファチジルコリン、ジオレオイルホスファチジルエタノールアミン（DOPE）、ジリノレオイルホスファチジルコリン、ジステアロイルホスファチジルエタノールアミン（DSPE）、ジミリストイルホスファチジルエタノールアミン（DMPE）、ジパルミトイルホスファチジルエタノールアミン（DPPE）、パルミトイルオレオイルホスファチジルエタノールアミン（POPE）、リゾホスファチジルエタノールアミン及びその組み合わせが挙げられるが、これに限定されるものではない。一実施形態では、中性リン脂質は、ジステアロイルホスファチジルコリン（DSPC）及びジミリストイルホスファチジルエタノールアミン（DMPE）からなる群から選択されてよい。別の実施形態では、中性リン脂質はジステアロイルホスファチジルコリン（DSPC）であってよい。別の実施形態では、中性リン脂質はジパルミトイルホスファチジルコリン（DPPC）であってよい。

【0120】

「ヘルパー脂質」としては、ステロイド、ステロール、及びアルキルレソルシノールが挙げられる。本開示での使用に好適なヘルパー脂質としては、コレステロール、5-ヘプタデシルレソルシノール、及びコレステロールヘミスクシナートが挙げられるが、これに限定されるものではない。一実施形態では、ヘルパー脂質はコレステロールであってよい。一実施形態では、ヘルパー脂質はコレステロールヘミスクシナートであってよい。

【0121】

PEG脂質は、ナノ粒子がインビボ（例えば、血液中）で存在できる時間の長さを変化させるステルス脂質である。PEG脂質は、例えば、粒子の凝集抑制及び粒径制御などにより製剤工程の補助となり得る。本明細書で使用するPEG脂質により、LNPの薬物動態特性が調節され得る。典型的に、PEG脂質は、脂質部分と、PEGに基づくポリマー部分とを含む。

10

20

30

40

50

【0122】

いくつかの実施形態では、脂質部分はジアシルグリセロールまたはジアシルグリカミドから得ることができ、これには、独立して、アルキル鎖長に約C4から約C40の飽和もしくは不飽和の炭素原子を含むジアルキルグリセロール基またはジアルキルグリカミド基を含むものが含まれ、ここで、かかる鎖は、例えば、アミドまたはエステルなど、1つ以上の官能基を含んでよい。いくつかの実施形態では、アルキル鎖長は約C10からC20を含む。ジアルキルグリセロール基またはジアルキルグリカミド基はさらに1つ以上の置換アルキル基を含むことができる。鎖長は対称でも非対称でもよい。

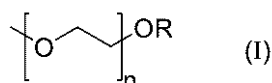
【0123】

特に明記しない限り、本明細書で使用する「PEG」という用語は、任意のポリエチレングリコールまたは他のポリアルキレンエーテルポリマーを意味する。一実施形態では、PEG部分は、任意選択で置換された、エチレングリコールまたはエチレンオキシドの直鎖ポリマーまたは分岐ポリマーである。別法として、PEG部分を、例えば、1つ以上のアルキル基、アルコキシ基、アシル基、ヒドロキシ基、またはアリール基などによって置換してよい。一実施形態では、PEG部分には、PEG-ポリウレタンまたはPEG-ポリプロピレンなどのPEG共重合体（例えば、J. Milton Harris, Poly(ethylene glycol) chemistry: biotechnical and biomedical applications (1992)を参照のこと）が含まれ、別法として、PEG部分はPEG共重合体を含まず、例えば、PEGモノポリマーであってよい。一実施形態では、PEGは分子量が約130～約50,000であり、下位実施形態では約150～約30,000であり、下位実施形態では約150～約20,000であり、下位実施形態では約150～約15,000であり、下位実施形態では約150～約10,000であり、下位実施形態では約150～約6,000であり、下位実施形態では約150～約5,000であり、下位実施形態では約150～約4,000であり、下位実施形態では約150～約3,000であり、下位実施形態では約300～約3,000であり、下位実施形態では約1,000～約3,000であり、下位実施形態では約1,500～約2,500である。

【0124】

特定の実施形態では、PEG（例えば、ステルス脂質などの脂質部分または脂質に結合させたもの）は「PEG-2K」であり、「PEG 2000」とも呼ばれ、平均分子量は約2,000ダルトンである。PEG-2Kは本明細書では以下の式（I）

【化4】



で表され、ここで、nは45であり、数平均重合度は約45のサブユニットを含むことを意味する。しかしながら、当該技術分野で公知の他のPEG実施形態を使用してよく、これには、例えば、数平均重合度が約23のサブユニット（n=23）、及び／または68のサブユニット（n=68）を含む実施形態が含まれる。いくつかの実施形態では、nは約30から約60であってよい。いくつかの実施形態では、nは約35から約55であってよい。いくつかの実施形態では、nは約40から約50であってよい。いくつかの実施形態では、nは約42から約48であってよい。いくつかの実施形態では、nは45であってよい。いくつかの実施形態では、Rは、H、置換アルキル、及び非置換アルキルから選択してよい。いくつかの実施形態では、Rは非置換アルキルであってよい。いくつかの実施形態では、Rはメチルであってよい。

【0125】

本明細書に記載のいずれの実施形態においても、PEG脂質は、PEG-ジラウロイルグリセロール、PEG-ジミリストイルグリセロール（PEG-DMG）（カタログ番号GM-020、NOF（Tokyo, Japan）製）、PEG-ジパルミトイルグリセ

ロール、PEG-ジステアロイルグリセロール (PEG-DSPE) (カタログ番号 DSPE-020CN、NOF (Tokyo, Japan) 製)、PEG-ジラウリルグリカミド、PEG-ジミリストイルグリカミド、PEG-ジパルミトイルグリカミド、及び PEG-ジステアロイルグリカミド、PEG-コレステロール (1-[8'-(コレスター5-エン-3[ベータ]-オキシ)カルボキサミド-3', 6'-ジオキサオクタニル]カルバモイル-[オメガ]-メチル-ポリ(エチレングリコール)、PEG-DMB (3, 4-ジテトラデコキシルベンジル-[オメガ]-メチル-ポリ(エチレングリコール)エーテル)、1, 2-ジミリストイル-sn-グリセロ-3-ホスホエタノールアミン-N-[メトキシ(ポリエチレングリコール)-2000] (PEG2k-DMG) (カタログ番号 880150P、Avanti Polar Lipids (Alabaster, Alabama, USA) 製)、1, 2-ジステアロイル-sn-グリセロ-3-ホスホエタノールアミン-N-[メトキシ(ポリエチレングリコール)-2000] (PEG2k-DSPE) (カタログ番号 880120C、Avanti Polar Lipids (Alabaster, Alabama, USA) 製)、1, 2-ジステアロイル-sn-グリセロール、メトキシポリエチレングリコール (PEG2k-DSG; GS-020、NOF (Tokyo, Japan) 製)、ポリ(エチレングリコール)-2000-ジメタクリレート (PEG2k-DMA)、及び 1, 2-ジステアリルオキシプロピル-3-アミン-N-[メトキシ(ポリエチレングリコール)-2000] (PEG2k-DSA) から選択してよい。一実施形態では、PEG 脂質は PEG2k-DMG であってよい。いくつかの実施形態では、PEG 脂質は PEG2k-DSG であってよい。一実施形態では、PEG 脂質は PEG2k-DSPE であってよい。一実施形態では、PEG 脂質は PEG2k-DMA であってよい。一実施形態では、PEG 脂質は PEG2k-C-DMA であってよい。一実施形態では、PEG 脂質は、WO2016/010840 で段落 [00240] から段落 [00244] までに開示されている化合物 S027 であってよい。一実施形態では、PEG 脂質は PEG2k-DSA であってよい。一実施形態では、PEG 脂質は PEG2k-C11 であってよい。いくつかの実施形態では、PEG 脂質は PEG2k-C14 であってよい。いくつかの実施形態では、PEG 脂質は PEG2k-C16 であってよい。いくつかの実施形態では、PEG 脂質は PEG2k-C18 であってよい。

【0126】

LNP 製剤

本開示の実施形態は、製剤中の成分脂質のそれぞれのモル比に従って表される脂質組成物を提供する。一実施形態では、アミン脂質のモル%は約 30 モル%~約 60 モル%であってよい。一実施形態では、アミン脂質のモル%は約 40 モル%~約 60 モル%であってよい。一実施形態では、アミン脂質のモル%は約 45 モル%~約 60 モル%であってよい。一実施形態では、アミン脂質のモル%は約 50 モル%~約 60 モル%であってよい。一実施形態では、アミン脂質のモル%は約 55 モル%~約 60 モル%であってよい。一実施形態では、アミン脂質のモル%は約 50 モル%~約 55 モル%であってよい。一実施形態では、アミン脂質のモル%は約 50 モル%であってよい。一実施形態では、アミン脂質のモル%は約 55 モル%であってよい。いくつかの実施形態では、LNP バッチのアミン脂質のモル%は、標的モル%の $\pm 30\%$ 、 $\pm 25\%$ 、 $\pm 20\%$ 、 $\pm 15\%$ 、 $\pm 10\%$ 、 $\pm 5\%$ 、または $\pm 2.5\%$ になる。いくつかの実施形態では、LNP バッチのアミン脂質のモル%は、標的モル%の ± 4 モル%、 ± 3 モル%、 ± 2 モル%、 ± 1.5 モル%、 ± 1 モル%、 ± 0.5 モル%、または ± 0.25 モル%になる。モル%数値はいずれも LNP 組成物脂質成分に対する割合として与えられる。特定の実施形態では、アミン脂質のモル%の LNP ロット間変動は 15%未満、10%未満または 5%未満になる。

【0127】

一実施形態では、中性脂質、例えば、中性リン脂質などのモル%は約 5 モル%~約 15 モル%であってよい。一実施形態では、中性脂質、例えば、中性リン脂質などのモル%は約 7 モル%~約 12 モル%であってよい。一実施形態では、中性脂質、例えば、中性リン

脂質などのモル％は約 0 モル％～約 5 モル％であってよい。一実施形態では、中性脂質、例えば、中性リン脂質などのモル％は約 0 モル％～約 10 モル％であってよい。一実施形態では、中性脂質、例えば、中性リン脂質などのモル％は約 5 モル％～約 10 モル％であってよい。一実施形態では、中性脂質、例えば、中性リン脂質などのモル％は約 8 モル％～約 10 モル％であってよい。

【0128】

一実施形態では、中性脂質、例えば、中性リン脂質などのモル％は約 5 モル％、約 6 モル％、約 7 モル％、約 8 モル％、約 9 モル％、約 10 モル％、約 11 モル％、約 12 モル％、約 13 モル％、約 14 モル％、または約 15 モル％であってよい。一実施形態では、中性脂質、例えば、中性リン脂質などのモル％は約 9 モル％であってよい。

10

【0129】

一実施形態では、中性脂質、例えば、中性リン脂質などのモル％は約 1 モル％～約 5 モル％であってよい。一実施形態では、中性脂質のモル％は約 0.1 モル％～約 1 モル％であってよい。一実施形態では、中性リン脂質などの中性脂質のモル％は約 0.1 モル％、約 0.2 モル％、約 0.5 モル％、1 モル％、約 1.5 モル％、約 2 モル％、約 2.5 モル％、約 3 モル％、約 3.5 モル％、約 4 モル％、約 4.5 モル％、または約 5 モル％であってよい。

【0130】

一実施形態では、中性脂質、例えば、中性リン脂質などのモル％は約 1 モル％未満であってよい。一実施形態では、中性脂質、例えば、中性リン脂質などのモル％は約 0.5 モル％未満であってよい。一実施形態では、中性脂質、例えば、中性リン脂質などのモル％は約 0 モル％、約 0.1 モル％、約 0.2 モル％、約 0.3 モル％、約 0.4 モル％、約 0.5 モル％、約 0.6 モル％、約 0.7 モル％、約 0.8 モル％、約 0.9 モル％、または約 1 モル％であってよい。いくつかの実施形態では、本明細書に開示の製剤は中性脂質を含まない（すなわち、中性脂質は 0 モル％）。いくつかの実施形態では、本明細書に開示の製剤は中性脂質を本質的に含まない（すなわち、中性脂質は約 0 モル％）。いくつかの実施形態では、本明細書に開示の製剤は中性リン脂質を含まない（すなわち、中性リン脂質は 0 モル％）。いくつかの実施形態では、本明細書に開示の製剤は中性リン脂質を本質的に含まない（すなわち、中性リン脂質は約 0 モル％）。

20

【0131】

いくつかの実施形態では、LNP バッチの中性脂質のモル％は、標的中性脂質のモル％の $\pm 30\%$ 、 $\pm 25\%$ 、 $\pm 20\%$ 、 $\pm 15\%$ 、 $\pm 10\%$ 、 $\pm 5\%$ 、または $\pm 2.5\%$ になる。特定の実施形態では、LNP ロット間変動は 15% 未満、10% 未満または 5% 未満になる。

30

【0132】

一実施形態では、ヘルパー脂質のモル％は約 20 モル％～約 60 モル％であってよい。一実施形態では、ヘルパー脂質のモル％は約 25 モル％～約 55 モル％であってよい。一実施形態では、ヘルパー脂質のモル％は約 25 モル％～約 50 モル％であってよい。一実施形態では、ヘルパー脂質のモル％は約 25 モル％～約 40 モル％であってよい。一実施形態では、ヘルパー脂質のモル％は約 30 モル％～約 50 モル％であってよい。一実施形態では、ヘルパー脂質のモル％は約 30 モル％～約 40 モル％であってよい。一実施形態では、脂質成分が 100 モル％となるようアミン脂質、中性脂質、及び PEG 脂質の濃度に基づいてヘルパー脂質のモル％を調整する。一実施形態では、脂質成分が 100 モル％となるようアミン脂質及び PEG 脂質の濃度に基づいてヘルパー脂質のモル％を調整する。一実施形態では、脂質成分が少なくとも 99 モル％となるようアミン脂質及び PEG 脂質の濃度に基づいてヘルパー脂質のモル％を調整する。いくつかの実施形態では、LNP バッチのヘルパーのモル％は、標的のモル％の $\pm 30\%$ 、 $\pm 25\%$ 、 $\pm 20\%$ 、 $\pm 15\%$ 、 $\pm 10\%$ 、 $\pm 5\%$ 、または $\pm 2.5\%$ になる。特定の実施形態では、LNP ロット間変動は 15% 未満、10% 未満または 5% 未満になる。

40

【0133】

50

一実施形態では、PEG脂質のモル%は約1モル%~約10モル%であってよい。一実施形態では、PEG脂質のモル%は約2モル%~約10モル%であってよい。一実施形態では、PEG脂質のモル%は約2モル%~約8モル%であってよい。一実施形態では、PEG脂質のモル%は約2モル%~約4モル%であってよい。一実施形態では、PEG脂質のモル%は約2.5モル%~約4モル%であってよい。一実施形態では、PEG脂質のモル%は約3モル%であってよい。一実施形態では、PEG脂質のモル%は約2.5モル%であってよい。いくつかの実施形態では、LNPバッチのPEG脂質のモル%は、標的PEG脂質のモル%の $\pm 30\%$ 、 $\pm 25\%$ 、 $\pm 20\%$ 、 $\pm 15\%$ 、 $\pm 10\%$ 、 $\pm 5\%$ 、または $\pm 2.5\%$ になる。特定の実施形態では、LNPロット間変動は15%未満、10%未満または5%未満になる。

【0134】

特定の実施形態では、カーゴには、RNA誘導型DNA結合因子（例えば、Casヌクレアーゼ、クラス2Casヌクレアーゼ、またはCas9）をコードするmRNA、及びgRNAもしくはgRNAをコードする核酸、またはmRNAとgRNAの組み合わせが含まれる。一実施形態では、LNP組成物はリピドAまたはその同等物を含んでよい。いくつかの態様では、アミン脂質はリピドAである。いくつかの態様では、アミン脂質はリピドA同等物、例えば、リピドAの類似体である。特定の態様では、アミン脂質はリピドAのアセタール類似体である。さまざまな実施形態では、LNP組成物は、アミン脂質、中性脂質、ヘルパー脂質、及びPEG脂質を含む。特定の実施形態では、ヘルパー脂質はコレステロールである。特定の実施形態では、中性脂質はDSPCである。具体的な実施形態では、PEG脂質はPEG2k-DMGである。いくつかの実施形態では、LNP組成物は、リピドA、ヘルパー脂質、中性脂質、及びPEG脂質を含んでよい。いくつかの実施形態では、LNP組成物は、アミン脂質、DSPC、コレステロール、及びPEG脂質を含む。いくつかの実施形態では、LNP組成物は、DMGを含むPEG脂質を含む。特定の実施形態では、アミン脂質は、リピドA、及びリピドAのアセタール類似体などリピドA同等物から選択される。さらなる実施形態では、LNP組成物は、リピドA、コレステロール、DSPC、及びPEG2k-DMGを含む。

【0135】

さまざまな実施形態では、LNP組成物は、アミン脂質、ヘルパー脂質、中性脂質、及びPEG脂質を含む。さまざまな実施形態では、LNP組成物は、アミン脂質、ヘルパー脂質、中性リン脂質、及びPEG脂質を含む。さまざまな実施形態では、LNP組成物は、アミン脂質、ヘルパー脂質、中性脂質、及びPEG脂質からなる脂質成分を含む。さまざまな実施形態では、LNP組成物は、アミン脂質、ヘルパー脂質、及びPEG脂質を含む。特定の実施形態では、LNP組成物は中性リン脂質などの中性脂質を含まない。さまざまな実施形態では、LNP組成物は、アミン脂質、ヘルパー脂質、及びPEG脂質からなる脂質成分を含む。特定の実施形態では、中性脂質は、DSPC、DPPC、DAPC、DMPC、DOPC、DOPE、及びDSPEのうち1つ以上から選ばれる。特定の実施形態では、中性脂質はDSPCである。特定の実施形態では、中性脂質はDPPCである。特定の実施形態では、中性脂質はDAPCである。特定の実施形態では、中性脂質はDMPCである。特定の実施形態では、中性脂質はDOPCである。特定の実施形態では、中性脂質はDOPEである。特定の実施形態では、中性脂質はDSPEである。特定の実施形態では、ヘルパー脂質はコレステロールである。具体的な実施形態では、PEG脂質はPEG2k-DMGである。いくつかの実施形態では、LNP組成物は、リピドA、ヘルパー脂質、及びPEG脂質を含んでよい。いくつかの実施形態では、LNP組成物は、リピドA、ヘルパー脂質、及びPEG脂質からなる脂質成分を含んでよい。いくつかの実施形態では、LNP組成物は、アミン脂質、コレステロール、及びPEG脂質を含む。いくつかの実施形態では、LNP組成物は、アミン脂質、コレステロール、及びPEG脂質からなる脂質成分を含む。いくつかの実施形態では、LNP組成物は、DMGを含むPEG脂質を含む。特定の実施形態では、アミン脂質は、リピドA及びリピドAのアセタール類似体などリピドA同等物から選択される。特定の実施形態では、アミン脂質は、C5

10

20

30

40

50

－C 1 2 または C 4－C 1 2 の、リピド A のアセタール類似体である。さらなる実施形態では、L N P 組成物は、リピド A、コレステロール、及び P E G 2 k－D M G を含む。

【0 1 3 6】

本開示の実施形態はまた、アミン脂質の正電荷のアミン基（N）と封入されるべき核酸の負電荷のリン酸基（P）とのモル比に従って表される脂質組成物も提供する。これは、式 N／P で数学的に表され得る。いくつかの実施形態では、L N P 組成物は、アミン脂質、ヘルパー脂質、中性脂質、及び P E G 脂質を含む脂質成分と、核酸成分とを含んでよく、ここで、N／P 比は約 3～1 0 である。いくつかの実施形態では、L N P 組成物は、アミン脂質、ヘルパー脂質、及び P E G 脂質を含む脂質成分と、核酸成分とを含んでよく、ここで、N／P 比は約 3～1 0 である。いくつかの実施形態では、L N P 組成物は、アミン脂質、ヘルパー脂質、中性脂質、及びヘルパー脂質を含む脂質成分と、RNA 成分とを含んでよく、ここで、N／P 比は約 3～1 0 である。いくつかの実施形態では、L N P 組成物は、アミン脂質、ヘルパー脂質、及び P E G 脂質を含む脂質成分と、RNA 成分とを含んでよく、ここで、N／P 比は約 3～1 0 である。一実施形態では、N／P 比は約 5～7 であってよい。一実施形態では、N／P 比は約 3～7 であってよい。一実施形態では、N／P 比は約 4．5～8 であってよい。一実施形態では、N／P 比は約 6 であってよい。一実施形態では、N／P 比は 6±1 であってよい。一実施形態では、N／P 比は 6±0．5 であってよい。いくつかの実施形態では、N／P 比は標的 N／P 比の±3 0 %、±2 5 %、±2 0 %、±1 5 %、±1 0 %、±5 %、または±2．5 % になる。特定の実施形態では、L N P ロット間変動は 1 5 % 未満、1 0 % 未満または 5 % 未満になる。

【0 1 3 7】

いくつかの実施形態では、核酸成分、例えば、RNA 成分は、mRNA、例えば、C a s ヌクレアーゼをコードする mRNA などを含んでよい。RNA 成分には、RNA が、任意選択でさらなる核酸及び／またはタンパク質、例えば、R N P カーゴなどと共に含まれる。一実施形態では、RNA は C a s 9 mRNA を含む。C a s ヌクレアーゼをコードする mRNA を含むいくつかの組成物では、L N P はさらに gRNA などの gRNA 核酸を含む。いくつかの実施形態では、RNA 成分は C a s ヌクレアーゼ mRNA 及び gRNA を含む。いくつかの実施形態では、RNA 成分はクラス 2 C a s ヌクレアーゼ mRNA 及び gRNA を含む。

【0 1 3 8】

特定の実施形態では、L N P 組成物は、クラス 2 C a s ヌクレアーゼなどの C a s ヌクレアーゼをコードする mRNA、アミン脂質、ヘルパー脂質、中性脂質、及び P E G 脂質を含んでよい。特定の実施形態では、L N P 組成物は、クラス 2 C a s ヌクレアーゼなどの C a s ヌクレアーゼをコードする mRNA、アミン脂質、ヘルパー脂質、及び P E G 脂質を含んでよい。クラス 2 C a s ヌクレアーゼなどの C a s ヌクレアーゼをコードする mRNA を含む特定の L N P 組成物では、ヘルパー脂質はコレステロールである。クラス 2 C a s ヌクレアーゼなどの C a s ヌクレアーゼをコードする mRNA を含む他の組成物では、中性脂質は D S P C である。クラス 2 C a s ヌクレアーゼなどの C a s ヌクレアーゼをコードする mRNA を含むさらなる実施形態では、P E G 脂質は P E G 2 k－D M G または P E G 2 k－C 1 1 である。クラス 2 C a s ヌクレアーゼなどの C a s ヌクレアーゼをコードする mRNA を含む特定の組成物では、アミン脂質は、リピド A 及びその同等物、例えば、リピド A のアセタール類似体などから選択される。

【0 1 3 9】

いくつかの実施形態では、L N P 組成物は gRNA を含んでよい。特定の実施形態では、L N P 組成物は、アミン脂質、gRNA、ヘルパー脂質、中性脂質、及び P E G 脂質を含んでよい。特定の実施形態では、L N P 組成物は、アミン脂質、gRNA、ヘルパー脂質、及び P E G 脂質を含んでよい。gRNA を含む特定の L N P 組成物では、ヘルパー脂質はコレステロールである。gRNA を含むいくつかの組成物では、中性脂質は D S P C である。gRNA を含むさらなる実施形態では、P E G 脂質は P E G 2 k－D M G または P E G 2 k－C 1 1 である。特定の実施形態では、アミン脂質は、リピド A 及びその同等

物、例えば、リピドAのアセタール類似体などから選択される。

【0140】

一実施形態では、LNP組成物はsgRNAを含んでよい。一実施形態では、LNP組成物はCas9 sgRNAを含んでよい。一実施形態では、LNP組成物はCpf1 sgRNAを含んでよい。sgRNAを含むいくつかの組成物では、LNPには、アミン脂質、ヘルパー脂質、中性脂質、及びPEG脂質が含まれる。sgRNAを含むいくつかの組成物では、LNPには、アミン脂質、ヘルパー脂質、及びPEG脂質が含まれる。sgRNAを含む特定の組成物では、ヘルパー脂質はコレステロールである。sgRNAを含む他の組成物では、中性脂質はDSPCである。sgRNAを含むさらなる実施形態では、PEG脂質はPEG2k-DMGまたはPEG2k-C11である。特定の実施形態では、アミン脂質は、リピドA及びその同等物、例えば、リピドAのアセタール類似体などから選択される。

10

【0141】

特定の実施形態では、LNP組成物は、CasヌクレアーゼをコードするmRNAと、sgRNAであってよいgRNAとを含む。一実施形態では、LNP組成物は、アミン脂質、CasヌクレアーゼをコードするmRNA、gRNA、ヘルパー脂質、中性脂質、及びPEG脂質を含んでよい。一実施形態では、LNP組成物は、アミン脂質、ヘルパー脂質、中性脂質、及びPEG脂質からなる脂質成分、ならびにCasヌクレアーゼをコードするmRNAとgRNAとからなる核酸成分を含んでよい。一実施形態では、LNP組成物は、アミン脂質、ヘルパー脂質、及びPEG脂質からなる脂質成分、ならびにCasヌクレアーゼをコードするmRNAとgRNAとからなる核酸成分を含んでよい。CasヌクレアーゼをコードするmRNA及びgRNAを含む特定の組成物では、ヘルパー脂質はコレステロールである。CasヌクレアーゼをコードするmRNA及びgRNAを含むいくつかの組成物では、中性脂質はDSPCである。CasヌクレアーゼをコードするmRNA及びgRNAを含む特定の組成物は、約1モル%未満の中性脂質、例えば、中性リン脂質を含む。CasヌクレアーゼをコードするmRNA及びgRNAを含む特定の組成物は、約0.5モル%未満の中性脂質、例えば、中性リン脂質を含む。特定の組成物では、LNPは中性脂質、例えば、中性リン脂質を含まない。CasヌクレアーゼをコードするmRNA及びgRNAを含むさらなる実施形態では、PEG脂質はPEG2k-DMGまたはPEG2k-C11である。特定の実施形態では、アミン脂質は、リピドA及びその同等物、例えば、リピドAのアセタール類似体などから選択される。

20

30

【0142】

特定の実施形態では、LNP組成物には、クラス2Cas mRNAなどのCasヌクレアーゼmRNA及び少なくとも1つのgRNAが含まれる。特定の実施形態では、LNP組成物には、gRNAと、クラス2CasヌクレアーゼmRNAなどのCasヌクレアーゼmRNAとが約25:1~約1:25の比で含まれる。特定の実施形態では、LNP製剤には、gRNAと、クラス2CasヌクレアーゼmRNAなどのCasヌクレアーゼmRNAとが約10:1~約1:10の比で含まれる。特定の実施形態では、LNP製剤には、gRNAと、クラス2CasヌクレアーゼmRNAなどのCasヌクレアーゼmRNAとが約8:1~約1:8の比で含まれる。本明細書で測定した比は重量基準である。いくつかの実施形態では、LNP製剤には、gRNAと、クラス2Cas mRNAなどのCasヌクレアーゼmRNAとが約5:1~約1:5の比で含まれる。いくつかの実施形態では、比の範囲は、約3:1~1:3、約2:1~1:2、約5:1~1:2、約5:1~1:1、約3:1~1:2、約3:1~1:1、約3:1、約2:1~1:1である。いくつかの実施形態では、gRNAとmRNAの比は約3:1または約2:1である。いくつかの実施形態では、gRNAとクラス2CasヌクレアーゼなどのCasヌクレアーゼmRNAとの比は約1:1である。比は、約25:1、10:1、5:1、3:1、1:1、1:3、1:5、1:10、または1:25であってよい。

40

【0143】

本明細書に開示するLNP組成物には鋳型核酸が含まれ得る。鋳型核酸を、Casヌク

50

レアーゼをコードするmRNA、例えば、クラス2 CasヌクレアーゼmRNAなどと共に合剤にしてよい。いくつかの実施形態では、鋳型核酸をガイドRNAと共に合剤にしてよい。いくつかの実施形態では、鋳型核酸を、CasヌクレアーゼをコードするmRNA及びガイドRNAの両方と共に合剤にしてよい。いくつかの実施形態では、鋳型核酸を、CasヌクレアーゼをコードするmRNAまたはガイドRNAとは別に製剤化してよい。鋳型核酸をLNP組成物と共に送達しても、LNP組成物とは別に送達してもよい。いくつかの実施形態では、鋳型核酸は、所望の修復機構に応じて一本鎖でも二本鎖でもよい。鋳型は、標的DNAまたは標的DNAに隣接する配列に対する相同性領域を有してよい。

【0144】

いくつかの実施形態では、水性RNA溶液と有機溶媒系脂質溶液、例えば、100%エタノールなどを混合することによってLNPを形成する。適切な溶液または溶媒には、水、PBS、Tris緩衝液、NaCl、クエン酸緩衝液、エタノール、クロロホルム、ジエチルエーテル、シクロヘキサン、テトラヒドロフラン、メタノール、イソプロパノールが含まれるか、またはそれらを含めてよい。薬理的に許容される緩衝液を、例えば、LNPをインビボで投与するためなどに使用してよい。特定の実施形態では、緩衝液を使用して、LNPを含む組成物のpHをpH6.5以上に維持する。特定の実施形態では、緩衝液を使用して、LNPを含む組成物のpHをpH7.0以上に維持する。特定の実施形態では、組成物はpHが約7.2~約7.7の範囲である。さらなる実施形態では、組成物はpHが約7.3~約7.7の範囲、または約7.4~約7.6の範囲である。さらなる実施形態では、組成物はpHが約7.2、7.3、7.4、7.5、7.6、または7.7である。マイクロpHプローブで組成物のpHを測定してよい。特定の実施形態では、組成物に凍結保護剤が含まれる。凍結保護剤の非限定的な例としては、ショ糖、トレハロース、グリセロール、DMSO、及びエチレングリコールが挙げられる。例示的な組成物には、例えば、ショ糖などの凍結保護剤が最高10%まで含まれてよい。特定の実施形態では、LNP組成物には約1%、2%、3%、4%、5%、6%、7%、8%、9%、または10%の凍結保護剤が含まれてよい。特定の実施形態では、LNP組成物には約1%、2%、3%、4%、5%、6%、7%、8%、9%、または10%のショ糖が含まれてよい。いくつかの実施形態では、LNP組成物には緩衝液が含まれてよい。いくつかの実施形態では、緩衝液は、リン酸緩衝液(PBS)、Tris緩衝液、クエン酸緩衝液、またはその混合物を含んでよい。特定の例示的な実施形態では、緩衝液はNaClを含む。特定の実施形態では、NaClを除外する。NaClの例示的な量は、約20mMから約45mMに及んでよい。NaClの例示的な量は、約40mMから約50mMに及んでよい。いくつかの実施形態では、NaClの量は約45mMである。いくつかの実施形態では、緩衝液はTris緩衝液である。Trisの例示的な量は、約20mMから約60mMに及んでよい。Trisの例示的な量は、約40mMから約60mMに及んでよい。いくつかの実施形態では、Trisの量は約50mMである。いくつかの実施形態では、緩衝液はNaCl及びTrisを含む。LNP組成物の特定の例示的な実施形態では、5%のショ糖と45mMのNaClとを溶解させたTris緩衝液が含有される。他の例示的な実施形態では、組成物は、約5w/v%の量のショ糖、約45mMのNaCl、及び約50mMのpH7.5のTrisを含有する。塩、緩衝液、及び凍結保護剤の量は、製剤全体としての浸透圧が維持されるよう変えてよい。例えば、最終浸透圧は450mOsm/L未満に維持してよい。さらなる実施形態では、浸透圧は350mOsm/Lと250mOsm/Lの間である。特定の実施形態では最終浸透圧は300+/-20mOsm/Lである。

【0145】

いくつかの実施形態では、マイクロ流体混合、T-混合、または交差混合を使用する。特定の態様では、流量、合流部サイズ、合流部の位置関係、合流部の形状、管径、溶液、及び/またはRNA濃度及び脂質濃度を変えてよい。LNPまたはLNP組成物を、例えば、透析、接線流ろ過、またはクロマトグラフィーなどにより濃縮または精製してよい。LNPは、例えば、懸濁液、エマルジョン、または凍結乾燥粉末などとして保存してよい

10

20

30

40

50

。いくつかの実施形態では、LNP組成物を2～8℃で保存し、特定の態様では、LNP組成物を室温で保存する。さらなる実施形態では、LNP組成物を、例えば、-20℃または-80℃などで凍結保存する。他の実施形態では、LNP組成物を約0℃～約-80℃の範囲の温度で保存する。凍結LNP組成物を、使用前に、例えば、氷上、室温、または25℃で解凍してよい。

【0146】

LNPは、例えば、ミクロ粒子（単層小胞及び多層小胞、例えば、「リポソーム」という、いくつかの実施形態では実質的に球状である、ラメラ相の脂質二重層などを含み、さらに特定の実施形態では水性核、例えば、RNA分子の実質的な部分を含むことができる）、エマルジョンの分散相、ミセル、または懸濁液の内相などであってよい。

10

【0147】

さらに、LNP組成物は生物分解性であり、治療有効量において細胞毒性濃度まで生体内に蓄積することがない。いくつかの実施形態では、LNP組成物は、治療用量において、実質的有害作用をもたらす自然免疫応答を引き起こさない。いくつかの実施形態では、本明細書で提供されるLNP組成物は、治療用量において毒性を生じない。

【0148】

いくつかの実施形態では、pdiは約0.005から約0.75に及んでよい。いくつかの実施形態では、pdiは約0.01から約0.5に及んでよい。いくつかの実施形態では、pdiは約ゼロから約0.4に及んでよい。いくつかの実施形態では、pdiは約ゼロから約0.35に及んでよい。いくつかの実施形態では、pdiは約ゼロから約0.35に及んでよい。いくつかの実施形態では、pdiは約ゼロから約0.3に及んでよい。いくつかの実施形態では、pdiは約ゼロから約0.25に及んでよい。いくつかの実施形態では、pdiは約ゼロから約0.2に及んでよい。いくつかの実施形態では、pdiは約0.08未満、0.1未満、0.15未満、0.2未満、または0.4未満であってよい。

20

【0149】

本明細書に開示するLNPは大きさ（例えば、Z-平均径）が約1～約250nmである。いくつかの実施形態では、LNPは大きさが約10～約200nmである。さらなる実施形態では、LNPは大きさが約20～約150nmである。いくつかの実施形態では、LNPは大きさが約50～約150nmである。いくつかの実施形態では、LNPは大きさが約50～約100nmである。いくつかの実施形態では、LNPは大きさが約50～約120nmである。いくつかの実施形態では、LNPは大きさが約60～約100nmである。いくつかの実施形態では、LNPは大きさが約75～約150nmである。いくつかの実施形態では、LNPは大きさが約75～約120nmである。いくつかの実施形態では、LNPは大きさが約75～約100nmである。特に指定しない限り、本明細書で言及する大きさはいずれも、Malvern Zetasizerで動的光散乱法により測定した完全形成粒子の平均サイズ（径）である。計数率が約200～400kcpsとなるよう、ナノ粒子試料をリン酸緩衝食塩水（PBS）で希釈する。データは強度測定値の加重平均として提示される（Z-平均径）。

30

【0150】

いくつかの実施形態では、LNPは、平均封入率が約50%～約100%の範囲で形成される。いくつかの実施形態では、LNPは、平均封入率が約50%～約70%の範囲で形成される。いくつかの実施形態では、LNPは、平均封入率が約70%～約90%の範囲で形成される。いくつかの実施形態では、LNPは、平均封入率が約90%～約100%の範囲で形成される。いくつかの実施形態では、LNPは、平均封入率が約75%～約95%の範囲で形成される。

40

【0151】

いくつかの実施形態では、LNPは、平均分子量が約1.00E+05g/mol～約1.00E+10g/molの範囲で形成される。いくつかの実施形態では、LNPは、平均分子量が約5.00E+05g/mol～約7.00E+07g/molの範囲で形

50

成される。いくつかの実施形態では、LNPは、平均分子量が約 $1.00E+06\text{ g/mol}$ ～約 $1.00E+10\text{ g/mol}$ の範囲で形成される。いくつかの実施形態では、LNPは、平均分子量が約 $1.00E+07\text{ g/mol}$ ～約 $1.00E+09\text{ g/mol}$ の範囲で形成される。いくつかの実施形態では、LNPは、平均分子量が約 $5.00E+06\text{ g/mol}$ ～約 $5.00E+09\text{ g/mol}$ の範囲で形成される。

【0152】

いくつかの実施形態では、多分散度 (M_w/M_n ；重量平均モル質量 (M_w) と数平均モル質量 (M_n) の比) は約1.000から約2.000に及んでよい。いくつかの実施形態では、 M_w/M_n は約1.00から約1.500に及んでよい。いくつかの実施形態では、 M_w/M_n は約1.020から約1.400に及んでよい。いくつかの実施形態では、 M_w/M_n は約1.010から約1.100に及んでよい。いくつかの実施形態では、 M_w/M_n は約1.100から約1.350に及んでよい。

【0153】

細胞を工学的に操作する方法；工学的に操作された細胞

本明細書に開示するLNP組成物は、インビボ及びインビトロの両方において、遺伝子編集による細胞の工学的操作方法で使用してよい。いくつかの実施形態では、方法には、細胞と本明細書に記載のLNP組成物との接触を伴う。

【0154】

いくつかの実施形態では、方法には、ヒトなどの哺乳類などの対象の細胞を接触させることを伴う。いくつかの実施形態では、細胞は臓器、例えば、肝臓、例えば、哺乳類の肝臓、例えば、ヒトの肝臓にある。いくつかの実施形態では、細胞は肝臓細胞、例えば、哺乳類の肝臓細胞、例えば、ヒト肝臓細胞である。いくつかの実施形態では、細胞は肝細胞、例えば、哺乳類の肝細胞、例えば、ヒト肝細胞である。いくつかの実施形態では、肝臓細胞は幹細胞である。いくつかの実施形態では、ヒト肝臓細胞は肝臓の類洞壁内皮細胞 (LSEC) であってよい。いくつかの実施形態では、ヒト肝臓細胞はクッパー細胞であってよい。いくつかの実施形態では、ヒト肝臓細胞は肝星細胞であってよい。いくつかの実施形態では、ヒト肝臓細胞は腫瘍細胞であってよい。さらなる実施形態では、細胞はApoe結合受容体を含む。いくつかの実施形態では、肝細胞などの肝臓細胞はインサイチュである。いくつかの実施形態では、肝細胞などの肝臓細胞は、例えば、初代培養のような培養などで単離されている。本明細書に開示する使用に対応する方法も提供し、これは、本明細書に開示するLNP組成物を対象に投与すること、または上記のような細胞と本明細書に開示するLNP組成物とを接触させることを含む。

【0155】

いくつかの実施形態では、工学的に操作された細胞、例えば、先行段落での細胞型のうちいずれか1つに由来する工学的に操作された細胞を提供する。そのような工学的に操作された細胞は本明細書に記載する方法に従って作製される。いくつかの実施形態では、工学的に操作された細胞は対象体内の組織または臓器、例えば、肝臓などの内部にある。

【0156】

本明細書に記載する方法及び細胞のいくつかでは、細胞は、標的配列のヌクレオチドの修飾、例えば、挿入もしくは欠失 (「インデル」) または置換を含む。いくつかの実施形態では、修飾は、標的配列の1個、2個、3個、4個または5個またはそれ以上のヌクレオチドの挿入を含む。いくつかの実施形態では、修飾は、標的配列の1個または2個のヌクレオチドの挿入を含む。他の実施形態では、修飾は、標的配列の1個、2個、3個、4個、5個、6個、7個、8個、9個、10個、15個、20個または25個またはそれ以上のヌクレオチドの欠失を含む。いくつかの実施形態では、修飾は、標的配列の1個または2個のヌクレオチドの欠失を含む。いくつかの実施形態では、修飾は、標的配列にフレームシフト突然変異をもたらすインデルを含む。いくつかの実施形態では、修飾は、標的配列の1個、2個、3個、4個、5個、6個、7個、8個、9個、10個、15個、20個または25個またはそれ以上のヌクレオチドの置換を含む。いくつかの実施形態では、

修飾は、標的配列の1個または2個のヌクレオチドの置換を含む。いくつかの実施形態では、修飾は、鋳型核酸、例えば、本明細書に記載する鋳型核酸のいずれかの組み込みにより生じたヌクレオチドの挿入、欠失、または置換のうち1つ以上を含む。

【0157】

いくつかの実施形態では、工学的に操作された細胞を含む細胞集団、例えば、本明細書に記載する方法に従って工学的に操作された細胞を含む細胞集団を提供する。いくつかの実施形態では、集団は、インビトロで培養された工学的操作をされた細胞を含む。いくつかの実施形態では、集団は、対象体内の組織または臓器、例えば、肝臓などの内部にある。いくつかの実施形態では、集団内の細胞のうち少なくとも5%、少なくとも10%、少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%または少なくとも95%またはそれ以上を工学的に操作する。特定の実施形態では、本明細書に開示する方法により、少なくとも5%、少なくとも10%、少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%または少なくとも95%の編集効率（または「編集率」）がもたらされ、これはインデルの検出により定義される。他の実施形態では、本明細書に開示する方法により、少なくとも5%、少なくとも10%、少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%または少なくとも95%のDNA修飾効率がもたらされ、これは、挿入、欠失、置換によるか、または他の方法によるかを問わず、配列変化の検出により定義される。特定の実施形態では、本明細書に開示する方法により、細胞集団における編集効率レベルまたはDNA修飾効率レベルは、約5%～約100%、約10%～約50%、約20～約100%、約20～約80%、約40～約100%、または約40～約80%になる。

【0158】

本明細書に記載する方法及び細胞のいくつかでは、集団内の細胞は、標的配列における修飾、例えば、インデルまたは置換などを含む。いくつかの実施形態では、修飾は、標的配列の1個、2個、3個、4個または5個またはそれ以上のヌクレオチドの挿入を含む。いくつかの実施形態では、修飾は、標的配列の1個または2個のヌクレオチドの挿入を含む。他の実施形態では、修飾は、標的配列の1個、2個、3個、4個、5個、6個、7個、8個、9個、10個、15個、20個または25個またはそれ以上のヌクレオチドの欠失を含む。いくつかの実施形態では、修飾は、標的配列の1個または2個のヌクレオチドの欠失を含む。いくつかの実施形態では、修飾により標的配列にフレームシフト突然変異がもたらされる。いくつかの実施形態では、修飾は、標的配列にフレームシフト突然変異をもたらしインデルを含む。いくつかの実施形態では、集団の工学的操作された細胞のうち少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、または少なくとも99%またはそれ以上はフレームシフト突然変異を含む。いくつかの実施形態では、修飾は、標的配列の1個、2個、3個、4個、5個、6個、7個、8個、9個、10個、15個、20個または25個またはそれ以上のヌクレオチドの置換を含む。いくつかの実施形態では、修飾は、標的配列の1個または2個のヌクレオチドの置換を含む。いくつかの実施形態では、修飾は、鋳型核酸、例えば、本明細書に記載する鋳型核酸のいずれかの組み込みにより生じたヌクレオチドの挿入、欠失、または置換のうち1つ以上を含む。

【0159】

遺伝子編集方法

インビボ及びインビトロで遺伝子編集を行うために本明細書に開示するLNP組成物を使用してよい。一実施形態では、本明細書に記載の1つ以上のLNP組成物を、それを必要とする対象に投与してよい。一実施形態では、本明細書に記載の1つ以上のLNP組成物は細胞と接触し得る。一実施形態では、治療的有效量の本明細書に記載する組成物は、それを必要とする対象の細胞と接触し得る。一実施形態では、遺伝子操作された細胞は、細胞と本明細書に記載のLNP組成物とを接触させることにより作製され得る。さまざまな実施形態では、方法は、上記のように鋳型核酸を細胞または対象に導入することを含む。

【0160】

いくつかの実施形態では、方法には、肝臓障害に関連した細胞に対するLNP組成物の投与を伴う。いくつかの実施形態では、方法には、肝臓障害の治療を伴う。特定の実施形態では、方法には、肝細胞とLNP組成物との接触を伴う。特定の実施形態では、方法には、肝細胞とLNP組成物との接触を伴う。いくつかの実施形態では、方法には、ApoE結合細胞とLNP組成物との接触を伴う。

【0161】

一実施形態では、クラス2CasヌクレアーゼをコードするmRNA及びgRNAを含むLNP組成物をApoE結合細胞などの細胞に投与してよい。さらなる実施形態では、鋳型核酸も細胞に導入される。特定の例では、クラス2Casヌクレアーゼ及びsgRNAを含むLNP組成物をApoE結合細胞などの細胞に投与してよい。一実施形態では、クラス2CasヌクレアーゼをコードするmRNA、gRNA、及び鋳型を含むLNP組成物を細胞に投与してよい。特定の例では、Casヌクレアーゼ及びsgRNAを含むLNP組成物を肝細胞に投与してよい。場合によっては、肝細胞は対象内にある。

【0162】

特定の実施形態では、対象はLNP組成物の単回投与を受けてよい。他の例では、対象はLNP組成物の多回投与を受けてよい。いくつかの実施形態では、LNP組成物を2～5回投与する。2用量以上を投与する場合、用量は、約1日、2日、3日、4日、5日、6日、7日、14日、21日、もしくは28日の間隔を置くか、約2か月、3か月、4か月、5か月、もしくは6か月の間隔を置くか、または約1年、2年、3年、4年、もしくは5年の間隔を置いて投与してよい。特定の実施形態では、LNP組成物の再投与時、編集が改善する。

【0163】

一実施形態では、クラス2CasヌクレアーゼなどのCasヌクレアーゼをコードするmRNAを含むLNP組成物は、gRNAを含む組成物の投与とは別に細胞に投与してよい。一実施形態では、クラス2CasヌクレアーゼなどのCasヌクレアーゼをコードするmRNA及びgRNAを含むLNP組成物は、細胞への鋳型核酸の投与とは別に細胞に投与してよい。一実施形態では、クラス2CasヌクレアーゼなどのCasヌクレアーゼをコードするmRNAを含むLNP組成物を細胞に投与し、その後、gRNAを含むLNP組成物を逐次投与してから、鋳型を細胞に投与してよい。CasヌクレアーゼをコードするmRNAを含むLNP組成物を投与してからgRNAを含むLNP組成物を投与する実施形態では、投与と投与の間を約4時間、6時間、8時間、12時間、もしくは24時間、または2日、3日、4日、5日、6日、もしくは7日空けてよい。

【0164】

一実施形態では、LNP組成物を使用して、遺伝子ノックアウトを生じさせる遺伝子を編集してよい。一実施形態では、LNP組成物を使用して、細胞集団に遺伝子ノックダウンを生じさせる遺伝子を編集してよい。別の実施形態では、LNP組成物を使用して、遺伝子修正を生じさせる遺伝子を編集してよい。さらなる実施形態では、LNP組成物を使用して、遺伝子挿入を生じさせる細胞を編集してよい。

【0165】

一実施形態では、LNP組成物の投与により、持続的応答を生じさせる遺伝子編集がもたらされ得る。例えば、投与により、反応の持続時間が1日、1か月、1年、またはそれ

10

20

30

40

50

以上になり得る。本明細書で使用する場合、「反応の持続時間」とは、本明細書に開示する L N P 組成物を使用して細胞が編集された後、それにより生じる修飾が、L N P 組成物投与後のある一定期間、依然として存在することを意味する。修飾は、標的タンパク質レベルの測定によって検出され得る。修飾は、標的 D N A の検出によって検出され得る。いくつかの実施形態では、反応の持続時間は少なくとも 1 週間であり得る。他の実施形態では、反応の持続時間は少なくとも 2 週間であり得る。一実施形態では、反応の持続時間は少なくとも 1 か月であり得る。いくつかの実施形態では、反応の持続時間は少なくとも 2 か月であり得る。一実施形態では、反応の持続時間は少なくとも 4 か月であり得る。一実施形態では、反応の持続時間は少なくとも 6 か月であり得る。特定の実施形態では、反応の持続時間は約 2 6 週間であり得る。いくつかの実施形態では、反応の持続時間は少なくとも 1 年であり得る。いくつかの実施形態では、反応の持続時間は少なくとも 5 年であり得る。いくつかの実施形態では、反応の持続時間は少なくとも 1 0 年であり得る。いくつかの実施形態では、持続的反応は少なくとも 0 . 5 か月、1 か月、2 か月、3 か月、4 か月、5 か月、6 か月、7 か月、8 か月、9 か月、1 0 か月、1 1 か月、1 2 か月、1 5 か月、1 8 か月、2 1 か月、または 2 4 か月の後、標的タンパク質レベルの測定によっても標的 D N A の検出によっても検出可能である。いくつかの実施形態では、持続的反応は少なくとも 1 年、2 年、3 年、4 年、5 年、6 年、7 年、8 年、9 年、1 0 年、1 2 年、1 4 年、1 6 年、1 8 年、または 2 0 年の後、標的タンパク質レベルの測定によっても標的 D N A の検出によっても検出可能である。

【0166】

L N P 組成物は、非経口投与が可能である。L N P 組成物を血流中、組織内、筋内、または内臓内へ直接投与してよい。投与は、例えば、注射または注入などで全身投与してよい。投与は局所投与であってよい。適切な投与手段としては、静脈内、動脈内、髄腔内、脳室内、尿道内、胸骨内、頭蓋内、網膜下、硝子体内、前房内、筋肉内、滑膜内、皮内、及び皮下が挙げられる。適切な投与デバイスとしては、針（including マイクロニードル）インジェクター、無針、浸透圧ポンプ、及び注入技術が挙げられる。

【0167】

L N P 組成物は必ずというわけではないが一般に、1 つ以上の薬理学的に許容される添加剤を含む製剤として投与される。用語「添加剤」には、本開示の化合物（複数可）以外の任意の成分、他の脂質成分（複数可）及び生物学的に活性な薬剤が含まれる。添加剤により、機能的（例えば、薬物放出速度制御）特徴及び／または非機能的（例えば、処理補助または希釈剤）特徴のいずれが製剤に付与されてもよい。添加剤の選択は、個々の投与方法、添加剤が溶解度及び安定性に及ぼす影響、及び剤形の性質などの要因によって大きく異なる。

【0168】

非経口製剤は典型的に水性もしくは油性の溶液または懸濁液である。製剤が水性の場合、添加剤は糖（グルコース、マンニトール、ソルビトール等が含まれるが、これに限定されるわけではない）、塩、炭水化物及び緩衝剤（好ましくは p H が 3 ~ 9）などであるが、一部の用途ではこれらは、滅菌非水溶液を用いるか、または滅菌パイロジェンフリー水（W F I）などの適切なビヒクルと併せて使用する乾燥形態としてさらに好適に製剤化されてよい。

【0169】

本発明は例示実施形態と併せて記載されるが、それらは、本発明を記載の実施形態に限定することを意図するものではないと理解される。むしろ、本発明は、代替案、変更、及び具体的な特徴の同等物を含む同等物のすべてを包含することを意図し、これらは添付の特許請求の範囲によって定義される本発明の範囲内に含まれ得る。

【0170】

上述された一般的記載及び詳細な説明の両方、ならびに以下の実施例は例示及び説明するのみであり、教示を限定するものではない。本明細書で使用する項目見出しは構成のみを目的としており、所望の主題を何ら限定するものではないと解釈されるべきである。参照

10

20

30

40

50

により組み込まれる文献が本明細書で定義した用語と矛盾する場合、本明細書が優先する。本出願に記載されるすべての範囲には、特に断らない限り、端点が包含される。

【0171】

本出願で使用する場合、単数形「a」、「an」及び「the」には、文脈で特に明確に指示されない限り、複数の指示対象が含まれることに留意すべきである。したがって、例えば、「組成物 (a composition)」への言及には複数の組成物が含まれ、「細胞 (a cell)」への言及には複数の細胞が含まれ、他も同様である。「または」の使用は包括的なものであり、特に断らない限り、「及び／または」を意味する。

【0172】

数字範囲には、その範囲を定義する数字が含まれる。測定値及び測定可能な値は概数であり、測定に関連する有効数字及び誤差を考慮すべきであると理解される。用語「約 (about)」または「約 (approximately)」とは、当業者により決定された特定の値について許容される誤差を意味し、値の測定法または決定法によって幾分異なる。範囲の前または列挙されている値の前に「約 (about)」などの修飾語の使用により、範囲の各端点または列挙にある各々の値が修飾を受ける。「約 (about)」には値または端点も含まれる。例えば、「約50～55」では「約50～約55」が包含される。また、「含む (comprise)」、「含む (comprises)」、「含む (comprising)」、「含有する (contain)」、「含有する (contains)」、「含有している (containing)」、「含む (include)」、「含む (include)」、及び「含む (including)」の使用は限定的なものではない。

【0173】

上記明細書で特に断りがない限り、明細書中、さまざまな成分を「含む (comprising)」という記述がある実施形態は、記述の成分「からなる (consisting of)」かまたはそれ「から本質的になる (consisting essentially of)」ことも意図され、明細書中、さまざまな成分「からなる (consisting of)」という記述がある実施形態は、記述の成分を「含む (consisting)」かまたはそれ「から本質的になる (consisting essentially of)」ことも意図され、明細書中、さまざまな成分「について (about)」という記述がある実施形態は、記述の成分「にて (at)」であることも意図され、また、明細書中、さまざまな成分「から本質的になる (consisting essentially of)」という記述がある実施形態は、記述の成分「からなる (consisting of)」かまたはそれを「含む (comprising)」ことも意図される（このような互換性は請求項内でこれらの用語を使用する際には適用されない）。

本発明は、例えば以下の実施形態を包含する：

〔実施形態1〕RNA成分と、

脂質成分とを含み、ここで、前記脂質成分は、

約50～60モル%のアミン脂質と、

約8～10モル%の中性脂質と、

約2.5～4モル%のPEG脂質と

を含み、ここで、前記脂質成分の残部はヘルパー脂質であり、

LNP組成物のN/P比は約6である、前記脂質ナノ粒子（「LNP」）組成物。

〔実施形態2〕RNA成分と、

約50～60モル%のアミン脂質と、

約27～39.5モル%のヘルパー脂質と、

約8～10モル%の中性脂質と、

約2.5～4モル%のPEG脂質と

を含み、ここで、LNP組成物のN/P比は約5～7である、前記LNP組成物。

〔実施形態3〕前記N/P比は約6である、実施形態2に記載のLNP組成物。

〔実施形態４〕RNA成分と、
脂質成分とを含み、ここで、前記脂質成分は、
約５０～６０モル％のアミン脂質と、
約５～１５モル％の中性脂質と、
約２．５～４モル％のPEG脂質と
を含み、ここで、前記脂質成分の残部はヘルパー脂質であり、
LNP組成物のN／P比は約３～１０である、前記LNP組成物。

〔実施形態５〕RNA成分と、
脂質成分とを含み、ここで、前記脂質成分は、
約４０～６０モル％のアミン脂質と、
約５～１５モル％の中性脂質と、
約２．５～４モル％のPEG脂質と
を含み、ここで、前記脂質成分の残部はヘルパー脂質であり、
LNP組成物のN／P比は約６である、前記LNP組成物。

〔実施形態６〕RNA成分と、
脂質成分とを含み、ここで、前記脂質成分は、
約５０～６０モル％のアミン脂質と、
約５～１５モル％の中性脂質と、
約１．５～１０モル％のPEG脂質と
を含み、ここで、前記脂質成分の残部はヘルパー脂質であり、
LNP組成物のN／P比は約６である、前記LNP組成物。

〔実施形態７〕RNA成分と、
脂質成分とを含み、ここで、前記脂質成分は、
約４０～６０モル％のアミン脂質と、
約０～１０モル％の中性脂質と、
約１．５～１０モル％のPEG脂質と
を含み、ここで、前記脂質成分の残部はヘルパー脂質であり、
LNP組成物のN／P比は約３～１０である、前記LNP組成物。

〔実施形態８〕RNA成分と、
脂質成分とを含み、ここで、前記脂質成分は、
約４０～６０モル％のアミン脂質と、
約１モル％未満の中性脂質と、
約１．５～１０モル％のPEG脂質と
を含み、ここで、前記脂質成分の残部はヘルパー脂質であり、
LNP組成物のN／P比は約３～１０である、前記LNP組成物。

〔実施形態９〕RNA成分と、
脂質成分とを含み、ここで、前記脂質成分は、
約４０～６０モル％のアミン脂質と、
約１．５～１０モル％のPEG脂質と
を含み、ここで、前記脂質成分の残部はヘルパー脂質であり、
LNP組成物のN／P比は約３～１０であり、かつ
前記LNP組成物は中性リン脂質を本質的に含まないかまたは含まない、前記LNP組成物。

〔実施形態１０〕RNA成分と、
脂質成分とを含み、ここで、前記脂質成分は、
約５０～６０モル％のアミン脂質と、
約８～１０モル％の中性脂質と、
約２．５～４モル％のPEG脂質と
を含み、ここで、前記脂質成分の残部はヘルパー脂質であり、
LNP組成物のN／P比は約３～７である、前記LNP組成物。

10

20

30

40

50

〔実施形態 1 1〕 前記 RNA 成分は mRNA を含む、先行実施形態いずれかに記載の組成物。

〔実施形態 1 2〕 前記 RNA 成分は、RNA 誘導型 DNA 結合因子、例えば、Cas ヌクレアーゼ mRNA などを含む、先行実施形態いずれかに記載の組成物。

〔実施形態 1 3〕 前記 RNA 成分はクラス 2 Cas ヌクレアーゼ mRNA を含む、先行実施形態いずれかに記載の組成物。

〔実施形態 1 4〕 前記 RNA 成分は Cas 9 ヌクレアーゼ mRNA を含む、先行実施形態いずれかに記載の組成物。

〔実施形態 1 5〕 前記 mRNA は修飾 mRNA である、実施形態 1 1 ～ 1 4 のいずれかに記載の組成物。

10

〔実施形態 1 6〕 前記 RNA 成分は、RNA 誘導型 DNA 結合因子をコードするオープンリーディングフレームを含む RNA を含み、ここで、前記オープンリーディングフレームはウリジン含量が、その最小ウリジン含量から、前記最小ウリジン含量の 1 5 0 % までの範囲である、先行実施形態いずれかに記載の組成物。

〔実施形態 1 7〕 前記 RNA 成分は、RNA 誘導型 DNA 結合因子をコードするオープンリーディングフレームを含む mRNA を含み、ここで、前記オープンリーディングフレームはウリジンジヌクレオチド含量が、その最小ウリジンジヌクレオチド含量から、前記最小ウリジンジヌクレオチド含量の 1 5 0 % までの範囲である、先行実施形態いずれかに記載の組成物。

〔実施形態 1 8〕 前記 RNA 成分は、配列番号 1、4、7、9、10、11、12、14、15、17、18、20、21、23、24、26、27、29、30、50、52、54、65、または 66 のいずれか 1 つに対する同一性が少なくとも 9 0 % である配列を含む mRNA を含み、ここで、前記 mRNA は、RNA 誘導型 DNA 結合因子をコードするオープンリーディングフレームを含む、先行実施形態いずれかに記載の組成物。

20

〔実施形態 1 9〕 前記 RNA 成分は g RNA 核酸を含む、先行実施形態のいずれかに記載の組成物。

〔実施形態 2 0〕 前記 g RNA 核酸は g RNA である、実施形態 1 9 に記載の組成物。

〔実施形態 2 1〕 前記 RNA 成分はクラス 2 Cas ヌクレアーゼ mRNA 及び g RNA を含む、先行実施形態いずれかに記載の組成物。

〔実施形態 2 2〕 前記 g RNA 核酸は、二重ガイド RNA (dg RNA) であるかまたはそれをコードする、実施形態 1 9 ～ 2 1 のいずれかに記載の組成物。

30

〔実施形態 2 3〕 前記 g RNA 核酸は、sg RNA であるかまたはそれをコードする、実施形態 1 9 ～ 2 1 のいずれかに記載の組成物。

〔実施形態 2 4〕 前記 g RNA は修飾されている、実施形態 1 9 ～ 2 3 のいずれかに記載の組成物。

〔実施形態 2 5〕 前記 g RNA は、2' -O -メチル (2' -O -Me) 修飾ヌクレオチド、ヌクレオチド間のホスホロチオアート (PS) 結合、及び 2' -フルオロ (2' -F) 修飾ヌクレオチドから選ばれる修飾を含む、実施形態 2 4 に記載の組成物。

〔実施形態 2 6〕 前記 g RNA は、5' 末端の最初の 5 ヌクレオチドの 1 つ以上における修飾を含む、実施形態 2 4 ～ 2 5 のいずれかに記載の組成物。

40

〔実施形態 2 7〕 前記 g RNA は、3' 末端の最後の 5 ヌクレオチドの 1 つ以上における修飾を含む、実施形態 2 4 ～ 2 6 のいずれかに記載の組成物。

〔実施形態 2 8〕 前記 g RNA は、最初の 4 ヌクレオチド間に PS 結合を含む、実施形態 2 4 ～ 2 7 のいずれかに記載の組成物。

〔実施形態 2 9〕 前記 g RNA は、最後の 4 ヌクレオチド間に PS 結合を含む、実施形態 2 4 ～ 2 8 のいずれかに記載の組成物。

〔実施形態 3 0〕 5' 末端の最初の 3 ヌクレオチドに 2' -O -Me 修飾ヌクレオチドをさらに含む、実施形態 2 4 ～ 2 9 のいずれかに記載の組成物。

〔実施形態 3 1〕 3' 末端の最後の 3 ヌクレオチドに 2' -O -Me 修飾ヌクレオチドをさらに含む、実施形態 2 4 ～ 3 0 のいずれかに記載の組成物。

50

〔実施形態 3 2〕前記 gRNA 及びクラス 2 Cas ヌクレアーゼ mRNA は、重量基準で約 10 : 1 ~ 約 1 : 10 の範囲の比で存在する、実施形態 19 ~ 31 のいずれかに記載の組成物。

〔実施形態 3 3〕前記 gRNA 及びクラス 2 Cas ヌクレアーゼ mRNA は、重量基準で約 5 : 1 ~ 約 1 : 5 の範囲の比で存在する、実施形態 19 ~ 31 のいずれかに記載の組成物。

〔実施形態 3 4〕前記 gRNA 及びクラス 2 Cas ヌクレアーゼ mRNA は、重量基準で約 3 : 1 ~ 約 1 : 1 の範囲の比で存在する、実施形態 19 ~ 33 のいずれかに記載の組成物。

〔実施形態 3 5〕前記 gRNA 及びクラス 2 Cas ヌクレアーゼ mRNA は、重量基準で約 2 : 1 ~ 約 1 : 1 の範囲の比で存在する、実施形態 19 ~ 34 のいずれかに記載の組成物。

10

〔実施形態 3 6〕前記 gRNA 及びクラス 2 Cas ヌクレアーゼ mRNA は、重量基準で約 2 : 1 の比で存在する、実施形態 19 ~ 35 のいずれかに記載の組成物。

〔実施形態 3 7〕前記 gRNA 及びクラス 2 Cas ヌクレアーゼ mRNA は、重量基準で約 1 : 1 の比で存在する、実施形態 19 ~ 35 のいずれかに記載の組成物。

〔実施形態 3 8〕少なくとも 1 つの鑄型をさらに含む、先行実施形態いずれかに記載の組成物。

〔実施形態 3 9〕前記 PEG 脂質のモル%は約 3 である、先行実施形態いずれかに記載の組成物。

20

〔実施形態 4 0〕前記アミン脂質のモル%は約 50 である、先行実施形態いずれかに記載の組成物。

〔実施形態 4 1〕前記アミン脂質のモル%は約 55 である、先行実施形態いずれかに記載の組成物。

〔実施形態 4 2〕前記アミン脂質のモル%は ± 3 モル%である、先行実施形態いずれかに記載の組成物。

〔実施形態 4 3〕前記アミン脂質のモル%は ± 2 モル%である、先行実施形態いずれかに記載の組成物。

〔実施形態 4 4〕前記アミン脂質のモル%は 47 ~ 53 モル%である、先行実施形態いずれかに記載の組成物。

30

〔実施形態 4 5〕前記アミン脂質のモル%は 48 ~ 53 モル%である、先行実施形態いずれかに記載の組成物。

〔実施形態 4 6〕前記アミン脂質のモル%は 53 ~ 57 モル%である、先行実施形態いずれかに記載の組成物。

〔実施形態 4 7〕前記 N / P 比は 6 ± 1 である、先行実施形態いずれかに記載の組成物。

〔実施形態 4 8〕前記 N / P 比は 6 ± 0.5 である、先行実施形態いずれかに記載の組成物。

〔実施形態 4 9〕前記アミン脂質はリピド A である、先行実施形態いずれかに記載の組成物。

〔実施形態 5 0〕前記アミン脂質はリピド A の類似体である、先行実施形態いずれかに記載の組成物。

40

〔実施形態 5 1〕前記類似体はアセタール類似体である、実施形態 5 0 に記載の組成物。

〔実施形態 5 2〕前記アセタール類似体は C4 - C12 アセタール類似体である、実施形態 5 1 に記載の組成物。

〔実施形態 5 3〕前記アセタール類似体は C5 - C12 アセタール類似体である、実施形態 5 0 に記載の組成物。

〔実施形態 5 4〕前記アセタール類似体は C5 - C10 アセタール類似体である、実施形態 5 0 に記載の組成物。

〔実施形態 5 5〕前記アセタール類似体は、C4 類似体、C5 類似体、C6 類似体、C7 類似体、C9 類似体、C10 類似体、C11 類似体、及び C12 類似体から選ばれる、実

50

施形態 50 に記載の組成物。

〔実施形態 56〕 前記ヘルパー脂質はコレステロールである、先行実施形態いずれかに記載の組成物。

〔実施形態 57〕 前記中性脂質は DSPC である、先行実施形態いずれかに記載の組成物。

〔実施形態 58〕 前記中性脂質は DPPC である、先行実施形態いずれかに記載の組成物。

〔実施形態 59〕 前記 PEG 脂質はジミリストイルグリセロール (DMG) を含む、先行実施形態いずれかに記載の組成物。

〔実施形態 60〕 前記 PEG 脂質は PEG-2k を含む、先行実施形態いずれかに記載の組成物。

〔実施形態 61〕 前記 PEG 脂質は PEG-DMG である、先行実施形態いずれかに記載の組成物。

〔実施形態 62〕 前記 PEG-DMG は PEG 2k-DMG である、実施形態 61 に記載の組成物。

〔実施形態 63〕 前記 LNP 組成物は中性脂質を本質的に含まない、実施形態 9 に記載の組成物。

〔実施形態 64〕 前記中性脂質はリン脂質である、実施形態 63 に記載の組成物。

〔実施形態 65〕 細胞を、実施形態 12～64 のいずれかに記載の LNP 組成物と接触させることを含む、遺伝子編集方法。

〔実施形態 66〕 クラス 2 Cas ヌクレアーゼ mRNA 及びガイド RNA 核酸を細胞に送達することを含む遺伝子編集方法であって、前記クラス 2 Cas mRNA 及び前記ガイド RNA 核酸は、実施形態 13～64 のいずれかに記載の少なくとも 1 つの LNP 組成物として製剤化される、前記方法。

〔実施形態 67〕 細胞を、実施形態 12～64 のいずれかに記載の少なくとも 1 つの LNP 組成物と接触させることを含む、遺伝子操作された細胞を作製する方法。

〔実施形態 68〕 前記 LNP 組成物を少なくとも 2 回投与する、実施形態 65～67 のいずれかに記載の方法。

〔実施形態 69〕 前記 LNP 組成物を 2～5 回投与する、実施形態 68 に記載の方法。

〔実施形態 70〕 再投与時に編集が改善する、実施形態 68 または 69 に記載の方法。

〔実施形態 71〕 少なくとも 1 つの鋳型核酸を前記細胞に導入することをさらに含む、実施形態 65～70 のいずれかに記載の方法。

〔実施形態 72〕 前記 mRNA を第 1 の LNP 組成物に製剤化し、前記ガイド RNA 核酸を第 2 の LNP 組成物に製剤化する、実施形態 65～71 のいずれかに記載の方法。

〔実施形態 73〕 前記第 1 及び第 2 の LNP 組成物を同時に投与する、実施形態 72 に記載の方法。

〔実施形態 74〕 前記第 1 及び第 2 の LNP 組成物を順次投与する、実施形態 72 に記載の方法。

〔実施形態 75〕 前記 mRNA 及び前記ガイド RNA 核酸を単一の LNP 組成物に製剤化する、実施形態 65～73 のいずれかに記載の方法。

【実施例】

【0174】

実施例 1—マウスにおけるインビボでの編集用 LNP 組成物

さまざまな LNP 組成物の小規模調製物を調製し、それらの特性を調べた。マウスにおける肝臓での編集割合についてのアッセイでは、Cas9 mRNA と、化学修飾したマウス TTR 配列指向性 sgRNA とを LNP 中に製剤化し、その際、PEG のモル%、リピド A のモル%、及び N:P 比を下記表 2 に記載のようにさまざまに変えた。

10

20

30

40

50

【表 2】

表 2. LNP 組成物.

LNP 番号	リピド A モル%	PEG-DMG モル%	N : P 比
(各種)	4.5	2, 2.5, 3, 4, 5	4.5
(各種)	4.5	2, 2.5, 3, 4, 5	6
(各種)	5.0	2, 2.5, 3, 4, 5	4.5
(各種)	5.0	2, 2.5, 3, 4, 5	6
(各種)	5.5	2, 2.5, 3, 4, 5	4.5
(各種)	5.5	2, 2.5, 3, 4, 5	6

10

【0175】

図 1 では、LNP 製剤は、「%CL ; N : P」で表わされる、それらのリピド A モル% 及び N : P 比に基づいて X 軸上で識別されている。図 1 の凡例で示されるように、2 モル%、2.5 モル%、3 モル%、4 モル%、または 5 モル%の濃度の PEG-2k-DMG は、(1) 4.5 モル%のリピド A ; N : P 比 4.5 (「4.5 ; 4.5」)、(2) 4.5 モル%のリピド A ; N : P 比 6 (「4.5 ; 6」)、(3) 5.0 モル%のリピド A ; N : P 比 4.5 (「5.0 ; 4.5」)、(4) 5.0 モル%のリピド A ; N : P 比 6 (「5.0 ; 6」)、(5) 5.5 モル%のリピド A ; N : P 比 4.5 (「5.5 ; 4.5」)、及び (6) 5.5 モル%のリピド A ; N : P 比 6 (「5.5 ; 6」) で製剤化した。DSPC のモル%は、9 モル%で一定に維持され、各製剤の脂質成分の残部が 100 モル%になるようコレステロール (モル%) を加えた。30 種の製剤の各々を以下に記載のように製剤化し、1 kg 当たり 1 mg または 1 kg 当たり 0.5 mg のトータル RNA 用量 (それぞれ、図 1 A 及び図 1 B) にて単回投与で投与した。

20

【0176】

LNP 製剤-NanoAssembler

脂質ナノ粒子成分を、脂質成分を上記のモル比で 100 % エタノールに溶解させた。RNA カーゴを 25 mM のクエン酸塩、100 mM の NaCl、pH 5.0 に溶解させ、RNA カーゴの濃度を約 0.45 mg/mL とした。LNP は、脂質アミンと RNA リン酸塩 (N : P) のモル比を約 4.5 または約 6、mRNA と gRNA の比を重量基準で 1 : 1 にして製剤化された。

30

【0177】

Precision Nanosystems NanoAssembler (商標) Benchtop Instrument を製造者の操作手順に従って使用し、脂質と RNA 溶液のマイクロ流体混合によって LNP を生成した。混合中は水性溶媒と有機溶媒の比を 2 : 1 に維持しつつ、異なる流量を使用した。混合した後、LNP を回収し、水に希釈 (約 1 : 1 v/v) して、室温で 1 時間保持し、さらに水で希釈 (約 1 : 1 v/v) してから最終の緩衝液交換を行った。50 mM の Tris、45 mM の NaCl、5 % (w/v) のショ糖、pH 7.5 (TSS) への最終緩衝液交換は PD-10 脱塩カラム (GE) で完了させた。必要に応じ、Amicon 100kDa 遠心式フィルター (Millipore) で遠心分離にかけて製剤を濃縮した。その後、得られる混合物を、0.2 µm 滅菌フィルターを使用してろ過した。最終 LNP は、以降の使用時まで -80 °C にて保存した。

40

【0178】

製剤分析法

本開示の LNP の多分散指数 (「pdi」) 及び大きさの特徴付けには、動的光散乱法 (「DLS」) を使用する。DLS は、試料を光源に当てて得られる光の散乱を測定する。DLS 測定から決定される PDI は、集団における粒径 (およその平均粒径) 分布を表

50

し、完全に均一な集団の場合、PDIはゼロである。

【0179】

特定pHにおけるLNPの表面電荷の特徴付けには電気泳動光散乱法を使用する。表面電荷、またはゼータ電位は、LNP懸濁液中の粒子間の静電的反発力／静電的引力の大きさの尺度である。

【0180】

非対称流フィールド・フロー・フラクシオネーション-多角度光散乱法（AF4-MALS）を使用して、流体力学的半径により製剤中の粒子を分離し、その後、分画粒子の分子量、流体力学的半径及び根平均二乗半径を測定する。これにより、分子量及び粒度分布ならびに二次的特徴、例えば、Burcharde-Stockmeyerプロット（粒子の内核密度を示唆する、根平均二乗（「rms」）半径と流体力学的半径の経時的な比）及びrmsコンホメーションプロット（得られる線形フィットの傾きがコンパクトなまとも（compactness）と伸長（elongation）の程度を与える、rms半径の対数と分子量の対数の対比）などを評価することができるようになる。

【0181】

製剤の粒度分布及び粒子濃度の決定にはナノ粒子トラッキング解析法（NTA、Malvern Nanosight）を使用することができる。LNP試料を適切に希釈し、顕微鏡スライド上に注入する。粒子がゆっくりと視野を通して注がれると、カメラが散乱光を記録する。ムービーが取得されるとナノ粒子トラッキング解析は画素を追跡して拡散係数を算出することによりムービーを処理する。この拡散係数は、粒子の流体力学的半径に翻訳され得る。この装置は、解析時に計数された個々の粒子数を計数して、粒子濃度も与える。

【0182】

LNPの粒径、形態、及び構造的特徴の決定には低温電子顕微鏡法（「cryo-EM」）を使用することができる。

【0183】

LNPの脂質組成分析は、液体クロマトグラフィー、その後の荷電化粒子検出（LC-CAD）により決定することができる。この分析では、実際の脂質含量と理論上の脂質含量の対比が得られる。

【0184】

LNP製剤を、平均粒径、多分散指数（pdi）、トータルRNA含量、RNAの封入率、及びゼータ電位について分析する。脂質分析、AF4-MALS、NTA、及び／またはcryo-EMによってLNP製剤をさらに特徴付けしてよい。平均粒径及び多分散度は、Malvern Zetasizer DLS装置を使用して動的光散乱法（DLS）により測定する。DLSで測定する前に、LNP試料をPBSに30X希釈した。平均粒径の強度基準の測定値であるZ-平均径が数平均径及びpdiと共に報告された。Malvern Zetasizer装置は、LNPのゼータ電位の測定にも使用される。測定に先立ち、試料を0.1X PBS（pH7.4）に1:17（50µLを800µLに）で希釈する。

【0185】

トータルRNA濃度及び遊離のRNAの決定には蛍光を基にしたアッセイ（Ribogreen（登録商標）、ThermoFisher Scientific）を使用する。封入率は、（トータルRNA-遊離のRNA）／トータルRNAで算出する。トータルRNAを決定するため0.2% Triton-X 100を含有している1X TE緩衝液を用いるか、または遊離のRNAを決定するため1X TE緩衝液を用いてLNP試料を適切に希釈する。製剤を作製するために使用され、その後1X TE緩衝液+／-0.2% Triton-X 100に希釈された、開始時のRNA溶液を使用して標準曲線を作成する。その後、希釈したRibogreen（登録商標）色素（製造者の取扱説明書に従う）を標準液及び試料のそれぞれに加え、光のない状態で室温にて約10分間インキュベートした。SpectraMax M5プレートリーダー（Molecular Dev

10

20

30

40

50

ices) を使用し、励起、オートカットオフ及び発光波長をそれぞれ488 nm、515 nm、及び525 nmに設定して試料の読み取りを行った。適切な標準曲線によりトータルRNA及び遊離RNAを決定する。

【0186】

封入率は、(トータルRNA－遊離のRNA) / トータルRNAで算出する。同様の手法を使用して、DNAに基づくかまたは核酸を含有するカーゴ成分の封入率を決定してよい。一本鎖DNAにはOligreen Dyeを使用してよく、二本鎖DNAにはPicogreen Dyeを使用してよい。

【0187】

AF4-MALSを使用して、分子量及び粒度分布、ならびにそれらの計算値による二次統計を調べる。LNPを適宜希釈し、それらが集められているHPLCオートサンプラーを使用してAF4分離チャンネルに注入し、その後、チャンネルを横切る交差流中に指数勾配で溶出させる。流体はすべてHPLCポンプ及びWyatt Eclipse装置で駆動される。粒子はAF4のチャンネル流から溶出し、UV検出器、多角度光散乱検出器、準弾性光散乱検出器及び示差屈折率検出器を通る。Debyeモデルを使用して生データを処理し、検出器信号から分子量及びrms半径を決定する。

【0188】

荷電化粒子検出器(CAD)に連結したHPLCにより、LNP中の脂質成分を定量的に分析する。クロマトグラフィーによる4種の脂質成分の分離は逆相HPLCにより達成された。CADは非揮発性化合物を漏れなく検出する質量基準の検出器であり、分析対象物の構造にかかわらずシグナルは一定である。

【0189】

Cas9 mRNA及びgRNAのカーゴ

Cas9 mRNAカーゴをインビトロ転写により調製した。直鎖状にしたプラスミドDNA鋳型及びT7 RNAポリメラーゼを使用してインビトロ転写により、1X-NLS(配列番号48)を含むキャップ化しポリアデニル化したCas9 mRNAを作製した。XbaIを以下の200 ng/μLのプラスミド、2 U/μLのXbaI(NEB)、及び1×反応緩衝液という条件で用いて37℃で2時間インキュベートすることにより、T7プロモーター及び100 ntのポリ(A/T)領域を含有しているプラスミドDNAを直鎖状にした。反応物を65℃で20分間加熱してXbaIを不活性化した。直鎖状にしたプラスミドをシリカマキシスピンカラム(Epoch Life Sciences)を使用して酵素及び緩衝塩により精製し、アガロースゲルで分析し、直鎖化を確認した。Cas9修飾mRNAを作製するためのIVT反応物を以下の条件で37℃にて4時間インキュベートした：50 ng/μLの直鎖状プラスミド；GTP、ATP、CTP、及びN1-メチルpseudo-UTP(Trilink)を各2 mMずつ；10 mMのARCA(Trilink)；5 U/μLのT7 RNAポリメラーゼ(NEB)；1 U/μLのマウスRNAse阻害剤(NEB)；0.004 U/μLのE.coli無機ピロホスファターゼ(NEB)；及び1×反応緩衝液。4時間インキュベーションした後、TURBO DNase(ThermoFisher)を加えて最終濃度0.01 U/μLとし、反応物をさらに30分間インキュベートしてDNA鋳型を除去した。Megaclear Transcription Cleanupキットを製造者(ThermoFisher)の操作手順に従って使用し、Cas9 mRNAを酵素及びヌクレオチドにより精製した。別法として、LiCl沈殿法でCas9 mRNAを精製した。

【0190】

本実施例でのsgRNAは化学合成したものであり、市販業者により供給されたものであった。sg282配列を以下に記載し、2'-O-メチル修飾及びホスホロチオアート結合は以下の通り表される(m=2'-OMe；*=ホスホロチオアート)：

mU*mU*mA*CA GCCACGUCUACAGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmC

m C m G m A m G m U m C m G m G m U m G m C m U * m U * m U * m U。 (配列番号 42)。

【0191】

L N P

最終 L N P の特徴付けを行い、封入率、多分散指数、及び平均粒径を上記の分析方法に従って決定した。

【0192】

マウスに L N P を投与し (1 m g / k g または 0 . 5 m g / k g での単回投与)、ゲノム DNA を単離して下記の N G S 分析に供した。

【0193】

インビボでの L N P 送達

各試験では 6 週齢から 10 週齢までの雌 C D - 1 マウスを使用した。動物の体重を測定し、体重に従って群に分け、群の平均体重に基づいて投与溶液を調製した。L N P を、動物 1 匹当たり 0 . 2 m L の容量 (体重 1 キログラム当たり約 10 m L) で側尾静脈を介して投与した。投与の約 6 時間後、動物を有害作用について観察した。投与の 24 時間後に体重を測定し、さまざまな時点において動物をイソフルラン麻酔下、心臓穿刺を介した放血により安楽死させた。血液を血清分離剤入り管または本明細書に記載のような血漿用の緩衝クエン酸ナトリウム含有チューブに採取した。インビボでの編集を行う試験用に、各動物の中葉または 3 つの独立した葉 (例えば、右中葉、左中葉、及び左側葉) から肝組織を採取し、DNA の抽出及び分析に供した。

【0194】

マウスのコホートを、次世代シーケンシング (N G S) 及び血清 T T R レベルによって肝臓での編集について測定した (データ図示せず)。

【0195】

トランスサイレチン (T T R) E L I S A 分析

血液を採取し、示されているように血清を単離した。マウスプレアルブミン (トランスサイレチン) E L I S キット (A v i v a S y s t e m s B i o l o g y、カタログ O K I A 00111) を使用してマウスの血清総 T T R レベルを決定した。ラット特異的 E L I S キット (A v i v a S y s t e m s B i o l o g y、カタログ番号 O K I A 00159) を製造者の操作手順に従って使用し、ラット血清 T T R レベルを測定した。簡単に言うと、血清をキットのサンプル希釈剤で段階的に希釈し、最終希釈度 10,000 倍とした。その後、この希釈サンプルを E L I S A プレートに加えた後、使用説明書に従ってアッセイを行った。

【0196】

N G S シークエンシング

手短に言えば、ゲノムにおける標的位置での編集効率を定量的に決定するため、ゲノム DNA を単離し、ディープシーケンシングを利用して、遺伝子編集により導入された挿入及び欠失の存在を同定した。

【0197】

標的部位 (例えば、T T R) 周辺に P C R プライマーを設計し、目的のゲノム領域を増幅させた。プライマー配列を以下に提供する。製造者の操作手順 (I l l u m i n a) に従ってさらなる P C R を実施し、シーケンシングに必要な化学を追加した。アンプリコンの配列決定を I l l u m i n a M i S e q 装置で行った。品質スコアの低いものを除去した後、読み取ったものをヒト基準ゲノム (例えば、h g 38) に対して整列させた。読み取りを含有している結果ファイルを基準ゲノム (B A M ファイル) に対してマッピングし、ここで、目的の標的領域と重複した読み取りを選択し、野生型の読み取り数と挿入、置換、または欠失含有する読み取り数との比を計算した。

【0198】

編集割合 (例えば、「編集効率」または「編集率」) は、野生型などの配列読み取り総数に対する、挿入または欠失を含む配列読み取り総数と定義される。

【0199】

図1は、NGSにより測定したマウス肝臓における編集割合を示す。図1Aに示すように、1kg当たり1mgのRNAを投与した場合、インビボでの編集割合は肝臓での編集が約20%～60%超である。1kg当たり0.5mgを投与した図1Bでは、約10%～60%の肝臓での編集が観察された。このマウスにおけるインビボ試験では、全組成物ともCas9 mRNA及びgRNAを肝細胞に効果的に送達し、各LNP組成物のNGSにより測定された、標的部位における活性なCRISPR/Casヌクレアーゼ活性が認められた。5%のPEG脂質を含有するLNPは封入率が低く（データ図示せず）、効力が幾分低かった。

【0200】

実施例2－LNP組成物分析法

LNPの特性解析は、リポドA及びPEG脂質を増量して製剤化したLNPにおいて、改善された物理化学的パラメータを示している。PEG脂質を2モル%または3モル%含む組成物（PEG2k-DMG）を下記表3に記載する。

【表3】

表3.

	LNP898	LNP897	LNP966	LNP969
CL／コレステロール／DSPC／PEG (理論上モル%)	45／44 ／9／2	45／43 ／9／3	50／38／ 9／3	55／33／ 9／3
mRNA	Cas9 配列番号48	Cas9 配列番号48	Cas9 U -dep 配列番号43	Cas9 U -dep 配列番号43
gRNA	G502 配列番号70	G502 配列番号70	G534 配列番号72	G534 配列番号72
N／P	4.5	4.5	6.0	6.0

【0201】

LNP製剤－交差流

脂質含有エタノールを、2容量のRNA溶液及び1容積の水と衝突噴流混合してLNPを形成した。脂質含有エタノールを、混合交差部を介して2容量のRNA溶液と混合する。第4流の水を、インラインT字部を介して交差部の出口流と混合する。（WO2016010840の図2を参照のこと）。LNPを室温で1時間維持し、その後、水でさらに希釈した（約1：1v/v）。希釈したLNPを、フラットシートカートリッジ（Sartorius、100kD MWCO）で接線流ろ過を使用して濃縮し、その後、緩衝液を、血液透析濾過によって50mMのTris、45mMのNaCl、5%（w/v）のショ糖、pH7.5（TSS）に交換した。別法として、TSSへの最終緩衝液交換をPD-10脱塩カラム（GE）で完了させた。必要に応じ、Amicon 100kDa遠心式フィルター（Millipore）で遠心分離にかけて製剤を濃縮した。その後、得られる混合物を、0.2μm滅菌フィルターを使用してろ過した。最終LNPを以降の使用時まで4℃または－80℃で保存した。

【0202】

Cas9 mRNA及びsgRNAを実施例1に記載のように調製したが、キャップ化及びポリアデニル化したCas9 U除去（Cas9 Udep）mRNAが配列番号43を含む点が異なっている。sg282は実施例1に記載され、sg534（「G534」）の配列を以下に記載する：

m A * m C * m G * C A A A U A U C A G U C C A G C G G U U U U A G A m G m C m U
m A m G m A m A m A m U m A m G m C A A G U U A A A A U A A G G C U A G U C C G
U U A U C A m A m C m U m U m G m A m A m A m A m G m U m G m G m C m A m C
m C m G m A m G m U m C m G m G m U m G m C m U * m U * m U * m U (配列番号 7 2)
【0203】

LNP製剤を、平均粒径、多分散度 (p d i)、トータルRNA含量及びRNAの封入率について実施例1に記載のように分析した。

【0204】

平均粒径、多分散度 (P D I)、トータルRNA含量及びRNAの封入率の分析を表4に示す。LNP組成物の理論上の脂質濃度に加え、脂質分析により、下記表5に示すような実際の脂質のモル%レベルが示された。

【表4】

表4.

L N P 識別 番号	N / P	Z - 平均 (n m)	P D I	数 平 均 (n m)	R N A 濃 度 (m g / m L)	封 入 率 (%)
L N P 8 9 8	4 . 5	8 7 . 9 1	0 . 0 3 0	7 1 . 3 3	1 . 5 3	9 8
L N P 8 9 7	4 . 5	7 4 . 0 5	0 . 0 3 6	5 8 . 5 5	1 . 4 3	9 8
L N P 9 6 6	6 . 0	8 2 . 7 8	0 . 0 1 0	6 7 . 8 6	2 . 1 2	9 8
L N P 9 6 9	6 . 0	9 2 . 9 7	0 . 0 4 2	7 5 . 5 2	2 . 0 9	9 7

【表5】

表5.

L N P 識別 番 号	脂質比 (リポドA / コレステロー ル / D S P C / P E G) (理論値及び 実測値)	リポドA (m g / m L) (理論値 及び実測 値)	コレステ ロール (m g / m L) (理論値 及び実測 値)	D S P C (m g / m L) (理論値 及び実測 値)	P E G (m g / m L) (理論値 及び実測 値)
L N P 8 9 8	4 5 / 4 4 / 9 / 2 4 6 . 1 / 4 2 . 6 / 9 . 2 / 2	1 8 . 0 1 8 . 3	8 . 0 7 . 7	3 . 3 3 . 4	2 . 3 2 . 4

10

20

30

40

50

L N P 8 9 7	4 5 / 4 3 / 9 / 3 4 4 . 8 / 4 2 . 9 / 9 . 2 / 3 . 1	1 8 . 0 1 7 . 8	7 . 8 7 . 7	3 . 3 3 . 4	3 . 5 3 . 6
L N P 9 6 6	5 0 / 3 8 / 9 / 3 5 0 . 0 / 3 8 . 0 / 8 . 8 / 3 . 1	3 3 . 4 3 5 . 6	1 1 . 5 1 2 . 3	5 . 6 5 . 8	5 . 8 6 . 5
L N P 9 6 9	5 5 / 3 3 / 9 / 3 5 4 . 8 / 3 3 . 2 / 8 . 8 / 3 . 2	3 3 . 4 3 1 . 6	9 . 1 8 . 7	5 . 1 4 . 7	5 . 3 5 . 4

10

20

【0205】

物理化学的性質をさらに分析するため、LNP897、LNP898、LNP966、及びLNP969を非対称流フィールド・フロー・フラクショネーション-多角度光散乱法（AF4-MALS）分析に供した。AF4-MALS装置は粒径及び分子量分布を測定し、粒子のコンホメーション及び密度についての情報を提供する。

【0206】

LNPは、それらが集められているHPLCオートサンプラーを使用してAF4分離チャンネルに注入し、その後、チャンネルを横切る交差流中に指数勾配で溶出させる。流体はすべてHPLCポンプ及びWyatt Eclipse装置で駆動される。粒子はAF4のチャンネル流から溶出し、UV検出器、Wyatt Heleos II多角度光散乱検出器、準弾性光散乱検出器及びWyatt Optilab T-rEX示差屈折率検出器を通る。Debyeモデルを使用してWyatt Astra 7ソフトウェアで生データを処理し、検出器信号から分子量及びrms半径を決定する。

30

【0207】

LNPについての対数で表すモル質量差プロットを図2Aとして示す。手短に言えば、X軸はモル質量（g/mol）を示し、Y軸は微分（differential number fraction）を示す。対数で表すモル質量差プロットにより、具体的な製剤について測定した異なる分子量の分布が示される。これにより、製剤中の分子量ならびに分子量の全体的分布の様子に関するデータが与えられ、粒子の異質性に関し平均分子量よりも良好な像が得られる。

40

【0208】

異なるモル質量モーメントを測定し、重量平均モル質量（Mw）と数平均モル質量（Mn）の比を計算してMw/Mnの多分散度を得ることによって、各種LNP製剤の異質性を決定する。これらの各種製剤の多分散度グラフを図2Bに記載する。

【0209】

図2Aに示すように、データは、PEGが3モル%、及びリピドAが50モル%及び55モル%の場合にN/P6.0で粒子分布が密になっていることを示す。これは、図2Bに示すように多分散度が密であることに反映されている。

【0210】

50

実施例 3-AF4 MALS データ-さらなる製剤

LNP の特性解析は、リピド A を増量して製剤化した LNP において、改善された物理化学的パラメータを示している。45 モル%、50 モル%、または 55 モル%いずれかのリピド A を 2 つの異なる gRNA と共に含む組成物を下記表 6 に記載する。

【表 6】

表 6.

	LNP 1021	LNP 1022	LNP 1023	LNP 1024	LNP 1025
CL/コレステロール/DSPC/PEG (理論上モル%)	50/3 8/9/3	55/3 3/9/3	45/4 3/9/3	50/3 8/9/3	55/3 3/9/3
mRNA	Cas9 Udep 配列番号 43	Cas9 Udep 配列番号 43	Cas9 Udep 配列番号 43	Cas9 Udep 配列番号 43	Cas9 Udep 配列番号 43
gRNA	G502 配列番号 70	G502 配列番号 70	G502 配列番号 70	G509 配列番号 71	G509 配列番号 71
N/P	6.0	6.0	4.5	6.0	6.0

【0211】

LNP を実施例 2 に記載のように形成した。

【0212】

Cas9 mRNA 及び sgRNA を上記のように調製した。

【0213】

LNP 組成物の特徴付けを行い、封入率、多分散指数、及び平均粒径を実施例 1 に記載のように決定した。

【0214】

平均粒径、多分散度 (PDI)、トータル RNA 含量及び RNA の封入率の分析を表 7 に示す。LNP 組成物の理論上の脂質濃度に加え、脂質分析により、下記表 8 に示すような実際の脂質のモル%レベルが示された。

【表 7】

表 7.

LNP 識別 番号	N/P	Z-平均 (nm)	PDI	数平均 (nm)	RNA 濃度 (mg/mL)	封入率 (%)
LNP 1021	6.0	83.18	0.027	67.15	1.63	98
LNP 1022	6.0	94.08	0.005	78.28	1.60	97

10

20

30

40

50

L N P 1 0 2 3	4. 5	7 4. 0 1	0. 0 1 7	6 1. 1 1	1. 6 1	9 7
L N P 1 0 2 4	6. 0	8 5. 3 7	0. 0 0 2	7 0. 4 2	1. 5 9	9 7
L N P 1 0 2 5	6. 0	9 4. 4 7	0. 0 1 8	7 7. 7 1	1. 6 0	9 8

10

【表 8】

表 8.

L N P 識 別 番 号	脂質比 (リピド A / コレステロー ル / D S P C / P E G) (理論値及び 実測値)	リピド A (m g / m L) (理論値 及び実測 値)	コレステ ロール m g / m L) (理論値 及び実測 値)	D S P C (m g / m L) (理論値 及び実測 値)	P E G (m g / m L) (理論値 及び実測 値)
L N P 1 0 2 1	5 0 / 3 8 / 9 / 3 5 0. 9 / 3 7. 4 / 8. 6 / 3. 1	2 3. 6 2 1. 6	8. 1 7. 2	3. 9 3. 4	4. 1 3. 8
L N P 1 0 2 2	5 5 / 3 3 / 9 / 3 5 5. 2 / 3 3. 0 / 8. 7 / 3. 1	2 3. 6 2 0. 4	6. 4 5. 5	3. 6 3. 0	3. 7 3. 4
L N P 1 0 2 3	4 5 / 4 3 / 9 / 3 4 5. 9 / 4 2. 4 / 8. 6 / 3. 1	1 7. 7 1 5. 3	7. 7 6. 4	3. 3 2. 7	3. 4 3. 0
L N P 1 0 2 4	5 0 / 3 8 / 9 / 3 5 0. 5 / 3 7. 9 / 8. 5 / 3. 0	2 3. 6 2 2. 4	8. 1 7. 6	3. 9 3. 5	4. 1 3. 9

20

30

40

50

L N P	5 5 / 3 3 /				
1 0 2	9 / 3	2 3 . 6	6 . 4	3 . 6	3 . 7
4	5 5 . 5 / 3	2 1 . 3	5 . 8	3 . 0	3 . 4
	3 . 1 / 8 . 5				
	/ 3 . 0				

【0215】

物理化学的性質をさらに分析するため、LNP1021、LNP1022、LNP1023、LNP1024及びLNP1025を非対称流フィールド・フロー・フラクシオネーション-多角度光散乱法（AF4-MALS）分析に供した。AF4-MALS装置は粒径及び分子量分布を測定し、粒子のコンホメーション及び密度についての情報を提供する。

【0216】

LNPに、実施例1に記載のようにAF4-MALSを実施した。

【0217】

LNPについての対数で表すモル質量差プロットを図3Aとして示す。手短に言えば、X軸はモル質量（g/mol）を示し、Y軸は微分数（differential number fraction）を示す。対数で表すモル質量差プロットにより、具体的な製剤について算出した異なる分子量の分布が示される。これにより、製剤中の分子量ならびに分子量の全体的分布の様子に関するデータが与えられ、粒子の異質性に関し平均分子量よりも良好な像が得られる。

【0218】

図3Bでは平均分子量をプロットした。平均分子量は分布全体の平均であるが、その分布の形状に関する情報は得られない。LNP1022及びLNP1025は平均分子量が同じであるが、分布はLNP1022の方がわずかに広い。

【0219】

異なるモル質量モーメントを調べ、重量平均モル質量（Mw）と数平均モル質量（Mn）の比を計算してMw/Mnの多分散度を得ることによって、各種LNP製剤の異質性を算出する。これらの各種製剤の多分散度グラフを図4Aに記載する。

【0220】

さらに、LNP製剤のBurchard-Stockmeyerプロットを図4Bとして示す。Burchard-Stockmeyerプロットでは、AF4チャンネルからの製剤の溶出液全体における、rms半径と流体力学的半径との比が示される。これにより、脂質ナノ粒子の内部密度に関する情報が得られる。図4Bは、この測定において、LNP1021、LNP1022及びLNP1023は異なるプロファイルを有することを示している。

【0221】

実施例4-PEG脂質増量では効力が維持されサイトカイン応答が低下する

別の試験では、PEG脂質を2モル%または3モル%含むLNP製剤で、PEG DMG脂質を比較した。PEG DMGを2モル%、または3モル%いずれかで含む組成物を下記表9に記載する。

【表9】

表9.

	LNP809	LNP810

10

20

30

40

50

C L / コレステロール / D S P C / P E G (理論上モル%)	4 5 / 4 4 / 9 / 2	4 5 / 4 3 / 9 / 3
m R N A	C a s 9 配列番号 4 8	C a s 9 配列番号 4 8
g R N A	G 3 9 0 配列番号 6 9	G 3 9 0 配列番号 6 9
N / P	4 . 5	4 . 5

10

【0222】

実施例2に記載の方法でLNPを形成した。

【0223】

C a s 9 mRNA及びs g R N Aを実施例1に記載のように調製し、s g 3 9 0 (「G 3 9 0」)の配列を以下に記載する：

m G * m C * m C * G A G U C U G G A G A G C U G C A G U U U U A G A m G m C m U
m A m G m A m A m A m U m A m G m C A A G U U A A A A U A A G G C U A G U C C G
U U A U C A m A m C m U m U m G m A m A m A m A m G m U m G m G m C m A m C
m C m G m A m G m U m C m G m G m U m G m C m U * m U * m U * m U (配列番号69)。

20

【0224】

L N P 製剤を、平均粒径、多分散度 (p d i)、トータルRNA含量及びRNAの封入率について実施例1に記載のように分析した。

【0225】

平均粒径、多分散度 (P D I)、トータルRNA含量及びRNAの封入率の分析を表10に示す。L N P 組成物の理論上の脂質濃度に加え、脂質分析により、下記表11に示すような実際の脂質のモル%レベルが示された。

【表10】

表10.

30

L N P 識別番号	N / P	Z - 平均 (n m)	P D I	数 平 均 (n m)	R N A 濃度 (m g / m L)	封入率 (%)
L N P 809	4 . 5	8 9 . 8 5	0 . 0 6 0	7 2 . 1 0	2 . 4 5	9 7
L N P 810	4 . 5	7 5 . 2 6	0 . 0 2 5	6 1 . 1 7	2 . 1 4	9 7

40

【表11】

表11.

50

L N P 識 別 番 号	脂質比 (リピド A / コレステロール / D S P C / P E G) (理論値及び 実測値)	リピド A (m g / m L) (理論値 及び実測 値)	コレステ ロール (m g / m L) (理論値 及び実測 値)	D S P C (m g / m L) (理論値 及び実測 値)	P E G (m g / m L) (理論値 及び実測 値)
L N P 8 0 9	4 5 / 4 4 / 9 / 2 4 5 . 7 / 4 3 . 3 / 9 . 0 / 2 . 1	2 8 . 6 3 0 . 5	1 2 . 7 1 3 . 1	5 . 3 5 . 6	3 . 7 4 . 0
L N P 8 1 0	4 5 / 4 3 / 9 / 3 4 5 . 0 / 4 2 . 3 / 9 . 7 / 3 . 0	2 5 . 2 2 4 . 7	1 0 . 9 1 0 . 5	4 . 7 4 . 9	4 . 9 4 . 7

10

20

【0226】

ラット血清サイトカインを、Luminex磁気ビーズマルチプレックスアッセイ (Milliplex MAP磁気ビーズアッセイ、Millipore Sigma製、カタログ番号RECYTMAG-65K) を使用して、MCP-1、IL-6、TNF-α及びIFN-γを分析して評価した。アッセイビーズをBioRad BioPlex-200で読み取り、BioPlex Managerソフトウェアバージョン6.1で4パラメータロジスティックフィットを使用して、標準曲線からサイトカイン濃度を計算した。データは図5にグラフ化されている。図5A (血清TTR)、図5B (肝臓での編集)、及び図5C (サイトカインp MCP1) を参照のこと。

30

【0227】

ラット特異的ELISAキット (Aviva Systems Biology、カタログ番号OKIA00159) を製造者の操作手順に従って使用し、ラット血清TTRレベルを測定した。簡単に言えば、血清をキットのサンプル希釈剤で段階的に希釈し、最終希釈度10,000倍とした。その後、この希釈サンプルをELISAプレートに加えた後、使用説明書に従ってアッセイを行った。

【0228】

約10mgの肝組織からゲノムDNAを単離し、上記のようにNGSを使用して解析した。増幅用PCRプライマー配列を以下に記載する。

40

【0229】

図5A及び図5Bは、PEGが2モル%及び3モル%の各製剤において、血清TTRノックダウン及び肝臓での編集が十分であったことを示す。図5Cは、PEGが3モル%の製剤使用時、MCP-1応答が低下していることを示す。

【0230】

実施例5-非ヒト霊長類へのLNP送達

実施例1に記載のように調製したLNP製剤を用いて3つの試験を行った。特定のモル量及びカーゴを表12～表26に記載する。Cas9 mRNA及びガイドRNA (gR

50

NA) を含有している各製剤はmRNA : gRNA比が重量基準で1 : 1であった。LNP投与量 (単位 : mg / kg、トータルRNA含量)、投与経路、及び動物へのデキサメタゾンでの前処置の有無を表に示す。デキサメタゾン (Dex) の前処置を受けた動物の場合、LNPまたはビヒクルの投与の1時間前に、Dexを2mg / kgでIVボラス注射により投与した。

【0231】

血液化学分析では、測定した各因子についての下記表で示されている時間に動物から血液を採取した。処理の前後NHPでサイトカイン誘導を測定した。拘束した覚醒動物の末梢静脈から最低の0.5mLの全血を4mL血清分離剤入り管に採取した。血液を室温で最低限30分間凝血させ、その後、2000xgで15分の遠心分離にかけた。血清を2本の120μLポリプロピレン製マイクロチューブに分注し、分析まで-60~-86℃で保存した。Meso Scale Discovery (MSD) 製非ヒト霊長類U-Plex Cytokineカスタムキットを使用して分析した。IL-6及びMCP-1に重点を置いて、以下のパラメータ、すなわち、INF-g、IL-1b、IL-2、IL-4、IL-6、IL-8、IL-10、IL-12p40、MCP-1及びTNF-aを分析に含めた。キット試薬及び標準液を製造者の操作手順で指示されているように調製した。NHP血清はそのまま使用した。MSD Sector Imager 6000でプレートの読み取りを実施し、MSD Discoveryワークベンチソフトウェアバージョン4012で分析を実施した。

【0232】

処置前後の動物で酵素免疫測定法により補体レベルを測定した。拘束した覚醒動物の末梢静脈から全血 (0.5mL) を0.5mLのk₂EDTAを含有する管に採取した。血液を2000xgで15分の遠心分離にかけた。血漿を2本の120μLポリプロピレン製マイクロチューブに分注し、分析まで-60~-86℃で保存した。Quidel MicroVue Complement Plus EIAキット (C3a-カタログ番号A031) または (Bb-カタログ番号A027) を使用して分析した。キット試薬及び標準液を製造者の操作手順で指示されているように調製した。MSD Sector Imager 6000で光学濃度を450nmにしてプレートの読み取りを実施した。4パラメータ曲線フィッティングを使用して結果を分析した。

【0233】

サイトカイン誘導及び補体活性化についてのデータを下記表に記載する。「BLQ」は、定量限界未満であることを意味する。

【表12】

表12. 試験1.

処置群	モル比 (それぞれ、 リピドA、コレステロール、DSPC、及びPEG2k-DMG	N:P	カーゴ	試料 サイズ (n)	経路	投与量、 トータル RNA含 量 (mg / kg)	D e x

(1) T SS (ビ ヒクル)	N/A	N / A	N/A	3	I V 注 入	N/A	なし
(2) L NP 6 9 9 G 5 0 2	4 5 / 4 4 / 9 / 2	4 . 5	C a s 9 mRNA (配列番 号 2) ;G 0 0 0 5 0 2	3	I V 注 入	3	なし
(3) L NP 6 8 8 G 5 0 6	4 5 / 4 4 / 9 / 2	4 . 5	C a s 9 mRNA (配列番 号 2) ;G 0 0 0 5 0 6	3	I V 注 入	3	なし
(4) L NP 6 8 9 G 5 0 9	4 5 / 4 4 / 9 / 2	4 . 5	C a s 9 mRNA (配列番 号 2) ;G 0 0 0 5 0 9	3	I V 注 入	3	なし
(5) L NP 6 9 0 G 5 1 0	4 5 / 4 4 / 9 / 2	4 . 5	C a s 9 mRNA (配列番 号 2) ;G 0 0 0 5 1 0	3	I V 注 入	3	なし

10

20

30

40

50

【表 1 3】

表 1 3 . 試験 2 .

処置群	モル比（それぞれ、リピドA、コレステロール、DSPC、及びPEG2k-DMG	N：P	カーゴ	試料サイズ（n）	経路	投与量、 トータルRNA含量（mg／kg）	D e x
（1）TSS（ビヒクル）	N／A	N／A		1	I V ボ ー ラ ス	N／A	有 り
（2）TSS（ビヒクル）	N／A	N／A		1	I V ボ ー ラ ス	N／A	な し
（3）LNP898 G5	4 5 ／ 4 4 ／ 9 ／ 2	4 . 5	C a s 9 mRNA （配列番	1	I V 注 入	3	有 り

10

20

30

40

50

0 2			号 2) ;G 0 0 0 5 0 2				
(4) L N P 8 9 8 G 5 0 2	4 5 / 4 4 / 9 / 2	4 . 5	C a s 9 m R N A (配 列 番 号 2) ;G 0 0 0 5 0 2	1	I V 注 入	3	な し
(5) L N P 8 9 7 G 5 0 2	4 5 / 4 3 / 9 / 3	4 . 5	C a s 9 m R N A (配 列 番 号 2) ;G 0 0 0 5 0 2	1	I V ボ ー ラ ス	3	有 り
(6) L N P 8 9 7 G 5 0 2	4 5 / 4 3 / 9 / 3	4 . 5	C a s 9 m R N A (配 列 番 号 2) ;G 0 0 0 5 0 2	1	I V ボ ー ラ ス	3	な し
(7) L N P 8 9 7 G 5 0 2	4 5 / 4 3 / 9 / 3	4 . 5	C a s 9 m R N A (配 列 番 号 2) ;G 0 0 0 5 0 2	1	I V 注 入	3	有 り
(8) L N P 8 9 7 G 5 0 2	4 5 / 4 3 / 9 / 3	4 . 5	C a s 9 m R N A (配 列 番 号 2) ;G 0 0 0 5 0 2	1	I V 注 入	3	な し
(9) L N P 9 1 6 G F P	4 5 / 4 3 / 9 / 3	4 . 5	e G F P m R N A (配 列 番 号)	1	I V 注 入	6	有 り
(1 0) L N P 9 1 6 G F P	4 5 / 4 3 / 9 / 3	4 . 5	e G F P m R N A (配 列 番 号)	1	I V 注 入	6	な し

【表 1 4】

表 1 4 . 試 験 3 .

10

20

30

40

50

処置群	モル比（それぞれ、 リピドA、コレステ ロール、DSPC、 及びPEG 2k- DMG	N : P	カーゴ	試料サ イ ズ (n)	経路	投与量、 トータル RNA含 量 (mg /kg)	D e x
(1) T SS	N/A	N / A	N/A	3	I V ボ ー ラス	N/A	なし
(2) L NP10 21 G 502	50 / 38 / 9 / 3	6	C a s 9 m RNA (配列 番号 1); G 000 502	3	I V ボ ー ラス	1	なし
(3) L NP10 21 G 502	50 / 38 / 9 / 4	6	C a s 9 m RNA (配列 番号 1); G 000 502	1	I V ボ ー ラス	1	有り
(4) L NP10 22 G 502	55 / 33 / 9 / 3	6	C a s 9 m RNA (配列 番号 1); G 000 502	3	I V ボ ー ラス	1	なし
(5) L NP10 23 G 502	45 / 43 / 9 / 3	4 . 5	C a s 9 m RNA (配列 番号 1); G 000 502	3	I V ボ ー ラス	3	なし
(6) L NP10	50 / 38 / 9 / 3	6	C a s 9 m	3	I V ボ ー	1	なし

10

20

30

40

50

2 4 G 5 0 9			R N A (配 列 番 号 1) ; G 0 0 0 5 0 9		ラス		
(7) L N P 1 0 2 4 G 5 0 9	5 0 / 3 8 / 9 / 4	6	C a s 9 m R N A (配 列 番 号 1) ; G 0 0 0 5 0 9	1	I V ボ ー ラス	1	有り
(8) L N P 1 0 2 5 G 5 0 9	5 5 / 3 3 / 9 / 3	6	C a s 9 m R N A (配 列 番 号 1) ; G 0 0 0 5 0 9	3	I V ボ ー ラス	1	なし
(9) L N P 1 0 2 1 G 5 0 2	5 0 / 3 8 / 9 / 3	6	C a s 9 m R N A (配 列 番 号 1) ; G 0 0 0 5 0 2	1	I V ボ ー ラス	3	なし
(1 0) L N P 1 0 2 2 G 5 0 2	5 0 / 3 8 / 9 / 3	6	C a s 9 m R N A (配 列 番 号 1) ; G 0 0 0 5 0 2	1	I V ボ ー ラス	3	なし

10

20

30

40

【表 1 5】

表 1 5 . 試験 1 による I L - 6 測定値.

処置群	事前採血	6 時間	2 4 時間
(1) T S S (ビ ヒクル)	5 . 7 1 ± 2 . 7 0	2 9 . 1 ± 2 0 . 3 7	7 . 0 5 ± 3 . 4 9

(2) LNP 6 9 9 G 5 0 2	9. 7 3 ± 8. 3 4	1 2 9 6. 4 1 ± 6 6 4. 7 1	5. 4 3 ± 7. 6 8
(3) LNP 6 8 8 G 5 0 6	1 6. 8 3 ± 4. 0 8	1 7 4 9. 4 7 ± 1 7 2 7. 2 2	3 8. 5 7 ± 3 9. 3 9
(4) LNP 6 8 9 G 5 0 9	1 8. 1 1 ± 1 1. 5 1	1 3 5 3. 4 9 ± 7 6 6. 6 6	3 2. 4 2 ± 1 8. 4 0
(5) LNP 6 9 0 G 5 1 0	1 3. 9 5 ± 1. 8 5	1 1 8 3 8 ± 1 7 1 6 1. 7 4	9 0. 0 7 ± 9 6. 0 2

10

【表 1 6】

表 1 6. 試験 1 による MCP-1 測定値.

処置群	事前採血	6 時間	2 4 時間
(1) TSS (ビ ヒクル)	8 1 0. 4 9 ± 1 7 8. 2 7	1 3 5 1. 1 6 ± 3 9 7. 3 1	7 4 5. 2 5 ± 5 6. 4 9
(2) LNP 6 9 9 G 5 0 2	8 4 2. 3 1 ± 3 5 0. 6 5	1 9 2 9 8. 4 9 ± 1 1 9 8 1. 1 4	2 0 9 2. 8 9 ± 1 7 1. 2 1
(3) LNP 6 8 8 G 5 0 6	1 1 9 0. 7 9 ± 3 8 3. 6 4	1 3 5 0 0. 1 7 ± 1 2 6 9 1. 6 0	1 4 1 4. 7 1 ± 4 2 2. 4 3
(4) LNP 6 8 9 G 5 0 9	8 3 8. 6 3 ± 2 8 4. 4 2	1 4 4 2 7. 7 ± 8 7 1 5. 4 8	1 5 9 0 ± 8 1 3. 2 3
(5) LNP 6 9 0 G 5 1 0	7 8 5. 3 2 ± 1 0 8. 9 7	5 2 5 5 7. 2 4 ± 4 8 0 3 4. 6 8	6 3 1 9. 7 7 ± 9 8 3. 3 7

20

30

【表 1 7】

表 1 7. 試験 1 による補体 C 3 a 測定値.

処置群	事前採血	6 時間	7 日目
(1) TSS (ビ ヒクル)	2 3. 9 ± 1 1. 9 5	2 5. 5 1 ± 1 4. 7 9	3 0. 6 7 ± 1 8. 3 6
(2) LNP 6 9 9 G 5 0 2	3 2. 3 6 ± 1 1. 2 9	9 4. 3 3 ± 5 8. 4 5	3 8. 5 0 ± 1 2. 6 9
(3) LNP 6 8 8 G 5 0 6	2 2. 3 0 ± 1. 7 3	1 2 7. 0 0 ± 2 2. 3 4	3 7. 8 0 ± 6. 8 6
(4) LNP 6 8 9 G 5 0 9	3 5. 8 3 ± 2 1. 9 4	1 7 4. 0 0 ± 4 4. 5 1	5 0. 8 3 ± 2 1. 9 2
(5) LNP 6 9 0 G 5 1 0	3 6. 3 0 ± 8. 2 1	1 6 3. 0 0 ± 4 0. 6 0	4 2. 5 0 ± 1 2. 4 4

40

50

【表 18】

表 18. 試験 1 による補体 b b 測定値.

処置群	0 4 - b b	事前採血	6 時間	7 日目
(1) T S S (ビ ヒ ク ル)	対 照	1 . 5 3 ± 0 . 1 9	3 . 3 7 ± 2 . 1 3	1 . 4 3 ± 0 . 7 1
(2) L N P 6 9 9 G 5 0	G 5 0 2	1 . 4 5 ± 0 . 3 9	9 . 0 1 ± 5 . 2 8	1 . 5 7 ± 0 . 5 4
2				
(3) L N P 6 8 8 G 5 0 6	G 5 0 6	1 . 4 5 ± 0 . 7 8	1 1 . 7 8 ± 2 . 3 3	1 . 7 8 ± 0 . 8 4
(4) L N P 6 8 9 G 5 0 9	G 5 0 9	1 . 9 5 ± 0 . 9 9	1 5 . 7 3 ± 2 . 2 3	2 . 8 3 ± 0 . 8 8
(5) L N P 6 9 0 G 5 1 0	G 5 1 0	2 . 1 2 ± 0 . 4 4	1 3 . 5 7 ± 1 . 2 3	2 . 2 1 ± 0 . 7 2

10

20

30

40

50

【表 19】

表 19. 試験 2 による IL-6 測定値.

処置群	事前採血	90分	6時間	24時間	7日目
(1) TSS (ビヒクル)	1. 77	11. 46	4. 2	2. 76	3. 01
(2) TSS (ビヒクル)	5. 23	18. 11	20. 36	13. 2	6. 36
(3) LNP8 98 G502	2. 02	1305. 75	1138. 22	383. 32	16. 02
(4) LNP8 98 G502	2. 34	37. 19	91. 59	14. 11	3. 07
(5) LNP8 97 G502	2. 1	55. 79	6. 89	2. 26	2. 01
(6) LNP8 97 G502	6. 8	10. 1	44. 72	5. 4	2. 01
(7) LNP8 97 G502	1. 97	44. 87	32. 61	2. 97	1. 11
(8) LNP8 97 G502	3. 14	37. 68	73. 41	8. 58	2. 22
(9) LNP9 16 GFP	1. 6	BLQ	95. 32	27. 58	BLQ
(10) LNP 916 GFP	2. 43	BLQ	883. 0 1	66. 71	BLQ

【表 20】

表 20. 試験 2 による MCP-1 測定値.

処置群	事前採血	90分	6時間	24時間	7日目
(1) TSS (ビヒクル)	312. 12	197. 24	145. 36	177. 02	403. 82
(2) TSS (ビヒクル)	232. 44	175. 08	187. 72	136. 64	325. 69
(3) L	249.	2183. 5	1814. 64	1887. 41	372.

10

20

30

40

50

N P 8 9 8 G 5 0 2	1				3 8
(4) L N P 8 9 8 G 5 0 2	3 4 9 . 5 1	4 3 0 . 4 9	5 6 3 5 . 5 5	9 5 3 . 0 5	2 3 6 . 6
(5) L N P 8 9 7 G 5 0 2	4 9 2 . 3	9 8 9 . 9 8	4 0 9 . 0 8	3 0 2 . 9 7	5 0 6 . 8 2
(6) L N P 8 9 7 G 5 0 2	2 8 3 . 7 9	2 2 5 . 1	1 1 4 1 . 0 8	4 8 4 . 5 9	2 5 9 . 4 6
(7) L N P 8 9 7 G 5 0 2	2 2 3 . 1 6	3 4 9 . 7 9	3 9 8 . 5 7	1 7 2 . 6 7	2 8 7 . 0 9
(8) L N P 8 9 7 G 5 0 2	5 8 4 . 4 2	8 5 3 . 5 1	3 8 8 0 . 8 1	1 5 8 8 . 4 6	6 9 2 . 9 9
(9) L N P 9 1 6 G F P	3 2 5 . 8 4	B L Q	1 1 8 9 . 9 7	2 2 7 9 . 8 2	B L Q
(1 0) L N P 9 1 6 G F P	1 7 5 . 4 7	B L Q	3 2 8 4 . 1 6	2 0 2 3 . 5 3	B L Q

10

20

30

40

50

【表 2 1】

表 2 1. 試験 2 による補体 C 3 a 測定値.

処置群	事前採血	9 0 分	6 時間	2 4 時間	7 日目
(1) T S S (ビヒクル)	0. 0 8 7	0. 0 9 6	0. 0 4 8	0. 0 3 3	0. 0 3 8
(2) T S S (ビヒクル)	0. 3 6 9	0. 3 1 1	0. 1 4 6	0. 1	0. 1 0 6
(3) L N P 8 9 8 G 5 0 2	0. 0 8 7	0. 9 5 3	0. 6 4 7	0. 2 7 7	0. 0 6 5
(4) L N P 8 9 8 G 5 0 2	0. 0 9 9	0. 2 6 2	0. 1 2 3	0. 0 4 9	0. 0 4 4
(5) L N P 8 9 7	0. 0 6 7	0. 4 7 9	0. 2 0 9	0. 0 3 6	0. 0 3 6

10

G 5 0 2					
(6) L N P 8 9 7 G 5 0 2	0. 1 4 1	0. 4 3 3	0. 3 4	0. 1 1	0. 0 7 4
(7) L N P 8 9 7 G 5 0 2	0. 1	0. 3 4 5	0. 3 9 6	0. 0 9 6	0. 1 2 7
(8) L N P 8 9 7 G 5 0 2	0. 2 6 1	0. 4 5 8	0. 4 0 9	0. 2 4 4	0. 3 1 3
(9) L N P 9 1 6 G F P	0. 1 4 9	B L Q	0. 7 1 4	0. 3 8 2	B L Q
(10) L N P 9 1 6 G F P	0. 1 1 7	B L Q	0. 7 5 2	0. 7 2 3	B L Q

20

30

40

50

【表 2 2】

表 2 2. 試験 2 による補体 b b 測定値.

処置群	事前採血	9 0 分	6 時間	2 4 時間	7 日目
(1) T S S (ビヒクル)	0. 0 8 7	0. 0 9 6	0. 0 4 8	0. 0 3 3	0. 0 3 8
(2) T S S (ビヒクル)	0. 3 6 9	0. 3 1 1	0. 1 4 6	0. 1	0. 1 0 6
(3) L N P 8 9 8 G 5 0 2	0. 0 8 7	0. 9 5 3	0. 6 4 7	0. 2 7 7	0. 0 6 5
(4) L N P 8 9 8 G 5 0 2	0. 0 9 9	0. 2 6 2	0. 1 2 3	0. 0 4 9	0. 0 4 4
(5) L N P 8 9 7 G 5 0 2	0. 0 6 7	0. 4 7 9	0. 2 0 9	0. 0 3 6	0. 0 3 6
(6) L N P 8 9 7 G 5 0 2	0. 1 4 1	0. 4 3 3	0. 3 4	0. 1 1	0. 0 7 4
(7) L N P 8 9 7 G 5 0 2	0. 1	0. 3 4 5	0. 3 9 6	0. 0 9 6	0. 1 2 7
(8) L N P 8 9 7 G 5 0 2	0. 2 6 1	0. 4 5 8	0. 4 0 9	0. 2 4 4	0. 3 1 3
(9) L N P 9 1 6 G F P	0. 1 4 9	B L Q	0. 7 1 4	0. 3 8 2	B L Q
(1 0) L N P 9 1 6 G F P	0. 1 1 7	B L Q	0. 7 5 2	0. 7 2 3	B L Q

【表 2 3】

表 2 3. 試験 3 による I L - 6 測定値.

処置群	事前採血	9 0 分	6 時間	2 4 時間	7 日目
(1) T S S	1. 8 9 ± 0. 9 7	2. 5 6 ± 1. 4 1	0. 9 0 ± 0. 7 1	B L Q	0. 0 8
(2) L N P 1 0 2 1 G 5 0 2	2 1 0 ± 0. 3 5	7. 4 4 ± 5. 1 6	6. 9 4 ± 8. 4 5	1. 0 7 ± 1. 1 1	1. 7 6 ± 0. 9 8
(3) L N P 1	0. 7 9	2. 9 6	4. 2 5	0. 6 7	0. 2 7

10

20

30

40

50

0 2 1 G 5 0 2					
(4) LNP 1 0 2 2 G 5 0 2	1. 5 4 $\pm 1. 3$ 2	2 0. 4 2 ± 3 1. 6 0	1 3. 9 4 ± 1 0. 1 0	0. 9 8 $\pm 0. 4$ 1	2. 0 4 $\pm$ 0. 6 5
(5) LNP 1 0 2 3 G 5 0 2	2. 9 2 $\pm 1. 6$ 8	6. 2 8 $\pm 7. 1$ 8	6. 0 6 $\pm 2. 3$ 1	3. 6 2 $\pm 4. 6$ 8	2. 0 0 $\pm$ 1. 2 1
(6) LNP 1 0 2 4 G 5 0 9	1. 4 3 $\pm 0. 6$ 2	2. 6 4 $\pm 1. 9$ 2	7. 7 2 $\pm 1 1.$ 9 6	0. 4 5 $\pm 0. 1$ 9	0. 8 8 $\pm$ 0. 7 9
(7) LNP 1 0 2 4 G 5 0 9	1. 3 5 $\pm 0. 7$ 4	2. 6 4 $\pm 2. 3$ 5	1. 7 1 $\pm 0. 4$ 1	0. 3 6 $\pm 0. 5$ 8	0. 5 1 $\pm$ 0. 3 2
(8) LNP 1 0 2 5 G 5 0 9	1. 6 4	2. 6 8	2 5. 6 5	0. 5 8	2. 0 0
(9) LNP 1 0 2 1 G 5 0 2	0. 5 6	6. 1 5	2 8. 8 0	0. 8 5	0. 6 1
(10) LNP 1 0 2 2 G 5 0 2	1. 7 6	8. 6 6	2 9 0 7. 8 6	1 1. 2 6	1. 7 2

10

20

30

40

50

【表 2 4】

表 2 4. 試験 2 による MCP-1 測定値.

処置群	事前採血	90分	6時間	24時間	7日目
(1) TS S	204. 01 ± 4 6.39	197.6 2 ± 19. 54	310.8 4 ± 45. 87	179.0 7 ± 20. 77	234.61 ± 71.79
(2) LN P1021 G502	303. 67 ± 3 6.37	337.6 3 ± 19 5.18	755.2 0 ± 58 1.45	339.7 5 ± 20 6.20	214.82 ± 40.81
(3) LN P1021 G502	229. 30	358.1 0	3182. 00	413.5 6	178.30
(4) LN P1022 G502	393. 63 ± 1 87.8 1	467.7 2 ± 22 1.61	1852. 94 ± 21 99.66	497.1 2 ± 41 2.30	382.19 ± 67.27
(5) LN P1023 G502	213. 72 ± 8.85	196.1 8 ± 62. 81	1722. 18 ± 14 13.90	197.8 3 ± 74. 01	156.16 ± 18.87
(6) LN	237.	210.3	468.5	22.32	141.20

P1024 G509	76 ± 9 6.36	7 ± 95. 17	3 ± 25 0.42	± 69.0 6	± 71.90
(7) LN P1024 G509	207. 36	183.0 7	1885. 66	235.7 0	163.11
(8) LN P1025 G509	259. 57 ± 1 12.9 8	299.2 1 ± 30 4.89	1193. 10 ± 97 4.04	258.8 2 ± 88. 53	219.86 ± 219.8 6
(9) LN P1021 G502	199. 29	286.0 4	2001. 23	197.5 7	196.44
(10) L NP102 2 G50 2	305. 81	970.6 5	7039. 06	8379. 05	203.47

10

20

30

40

【表 2 5】

表 2 5. 試験 3 による補体 C 3 a 測定値.

処置群	事前採血	9 0 分	6 時間	2 4 時間	7 日 目
(1) T S S	4 2 . 4 7 ± 1 0 . 3 0	5 5 . 4 0 ± 1 3 . 5 8	2 9 . 3 0 ± 1 4 . 4 6	4 1 . 7 0 ± 2 3 . 6 5	2 7 . 4 3 ± 1 2 . 4 3
(2) L N P 1 0 2 1 G 5 0 2	3 4 . 3 7 ± 0 . 5 0	8 6 . 5 0 ± 3 . 6 6	9 0 . 0 7 ± 4 . 8 5	5 6 . 6 0 ± 2 . 2 5	3 2 . 5 3 ± 0 . 9 3
(3) L N P 1 0 2 1 G 5 0 2	3 4 . 3 0	1 2 8 . 0 0 0	9 3 . 3 0	3 3 . 4 0	2 8 . 2 0
(4) L N P 1 0 2 2 G 5 0 2	4 1 . 5 5 ± 1 3 . 5 1	1 5 1 . 3 7 ± 1 0 9 . 9 8	8 2 . 0 0 ± 3 1 . 8 2	4 5 . 5 7 ± 1 8 . 5 8	3 2 . 7 7 ± 6 . 4 5
(5) L N P 1 0 2 3 G 5 0 2	3 1 . 6 7 ± 3 . 1 9	7 4 . 4 0 ± 2 2 . 0 8	7 4 . 1 3 ± 4 8 . 6 1	3 3 . 8 3 ± 9 . 7 5	2 7 . 7 0 ± 8 . 0 5
(6) L N P 1 0 2 4 G 5 0 9	5 6 . 6 0 ± 2 5 . 6 1	1 0 0 . 3 7 ± 7 7 . 9 5	7 4 . 7 3 ± 7 0 . 1 5	5 5 . 2 0 ± 4 8 . 3 4	4 9 . 9 7 ± 3 9 . 9 4
(7) L N P 1 0 2 4 G 5 0 9	3 3 . 8 0	3 3 . 9 0	3 3 . 7 0	2 6 . 1 0	2 0 . 9 0
(8) L N P 1 0 2 5 G 5 0 9	3 9 . 9 0 ± 1 3 . 0 1	7 5 . 7 3 ± 1 . 3 8	4 6 . 1 3 ± 3 0 . 5 6	2 5 . 0 0 ± 3 . 8 0	2 3 . 9 0 ± 7 . 1 8

(9) L N P 1 0 2 1 G 5 0 2	3 4	8 5 . 7 0	1 3 3 . 0 0	6 2 . 0 0	2 5 . 5 0
(10) L N P 1 0 2 2 G 5 0 2	2 9 . 8	6 8 . 1 0	1 1 3 . 0 0	7 1 . 7 0	2 3 . 3 0

10

20

30

40

50

【表 2 6】

表 2 6．試験 3 による補体 b b 測定値.

処置群	事前採血	9 0 分	6 時間	2 4 時間	7 日目
(1) T S S	1 . 4 6 ± 0 . 7 0	2 . 1 8 ± 0 . 7 8	1 . 9 6 ± 0 . 6 4	0 . 9 4 5 ± 0 . 1 5	1 . 3 4 ± 0 . 5 0
(2) L N P 1 0 2 1 G 5 0 2	1 . 7 7 ± 0 . 6 0	6 . 5 1 ± 3 . 6 6	1 1 . 0 0 ± 4 . 8 5	3 . 5 9 ± 2 . 2 5	2 . 0 7 ± 0 . 9 3
(3) L N P 1 0 2 1 G 5 0 2	1 . 2 4	2 . 9 0	1 1 . 5 0	2 . 9 7	1 . 2 4
(4) L N P 1 0 2 2 G 5 0 2	1 . 5 2 ± 0 . 3 4	5 . 6 7 ± 2 . 2 8	1 0 . 2 ± 3 . 3 6	3 . 6 6 ± 1 . 6 8	1 . 8 4 ± 0 . 2 4
(5) L N P 1 0 2 3 G 5 0 2	1 . 6 5 ± 0 . 9 4	4 . 4 ± 1	7 . 6 8 ± 4 . 6 7	2 . 6 4 ± 1 . 1 8	2 . 0 8 ± 1 . 3 2
(6) L N P 1 0 2 4 G 5 0 9	1 . 6 1 ± 0 . 1 3	4 . 5 2 ± 1 . 8 1	4 . 5 0 ± 3 . 2 2	1 . 6 3 ± 0 . 8 4	1 . 6 3 ± 0 . 3 2
(7) L N P 1 0 2 4 G 5 0 9	0 . 9 6	2 . 9 9	2 . 6 4	1 . 1 3	1 . 0 7
(8) L N P 1 0 2 5 G 5 0 9	1 . 3 7 ± 0 . 1 7	4 . 9 ± 4 . 5 1	3 . 7 9 ± 3 . 8 4	1 . 6 6 ± 1 . 4 3	1 . 3 5 ± 0 . 4 4
(9) L N P 1 0 2 1 G	1 . 4 1	5 . 6 7	1 1 . 5 0	4 . 6 4	1 . 3 8

5 0 2					
(1 0) L N P 1 0 2 2 G 5 0 2	1 . 2 8	5 . 2 2	1 4 . 1 0	5 . 6 4	1 . 8 7

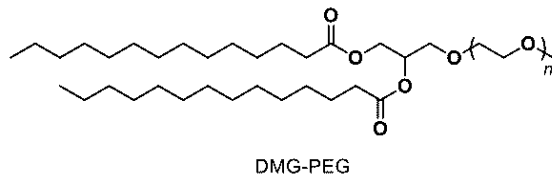
実施例 6－PEG 脂質スクリーニング

別の試験では、PEG 脂質を 2 モル% または 3 モル% 含む LNP 製剤で代替的 PEG 脂質を比較した。

【0235】

試験では 3 種の PEG 脂質を使用し、脂質 1 (DMG-PEG 2k; No f) は、以下のように表される。

【化 5】

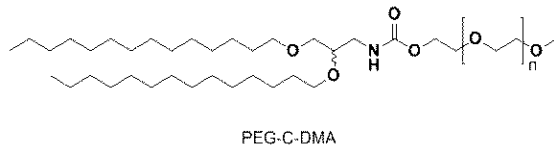


10

【0236】

Heyes, et al., J. Controlled Release, 107 (2005), pp. 278-279 (「Synthesis of PEG2000-C-DM A」を参照のこと) に記載のように合成される脂質 2 は、以下のように表すことができ、

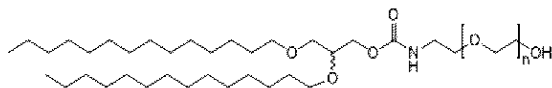
【化 6】



20

WO 2016 / 010840 (化合物 S 027、段落 [00240] から段落 [00244] を参照のこと) 及び WO 2011 / 076807 に開示される脂質 3 は、以下のように表すことができる。

【化 7】



30

【0237】

各 PEG 脂質を 2 モル% 及び 3 モル% 用いてリピド A を製剤化した。脂質ナノ粒子成分を、脂質成分を上記のモル比で 100% エタノールに溶解させた。手短に言えば、RNA カーゴを 25 mM のクエン酸塩、100 mM の NaCl、pH 5.0 に調製し、RNA カーゴの濃度を約 0.45 mg/mL とした。LNP は、脂質アミンと RNA リン酸塩 (N : P) のモル比を約 4.5、mRNA と gRNA の比を重量基準で 1 : 1 にして製剤化された。

【表 27】

表 27.

40

LNP 番号	LNP 784	LNP 785	LNP 786	LNP 787	LNP 788	LNP 789

50

CL／コレステロール／DSPC／PEG (理論上モル%)	4 5 / 4 4 / 9 / 2	4 5 / 4 3 / 9 / 3	4 5 / 4 4 / 9 / 2	4 5 / 4 3 / 9 / 3	4 5 / 4 4 / 9 / 2	4 5 / 4 3 / 9 / 3
mRNA	C a s 9 配列番号 4 8	C a s 9 配列番号 4 8	C a s 9 配列番号 4 8	C a s 9 配列番号 4 8	C a s 9 配列番号 4 8	C a s 9 配列番号 4 8
gRNA	G 2 8 2 配列番号 4 2	G 2 8 2 配列番号 4 2	G 2 8 2 配列番号 4 2	G 2 8 2、 配列番号 4 2	G 2 8 2、 配列番号 4 2	G 2 8 2 配列番号 4 2
PEGの種類	脂質 1	脂質 1	脂質 2	脂質 2	脂質 3	脂質 3
N／P	4 . 5	4 . 5	4 . 5	4 . 5	4 . 5	4 . 5

10

【0238】

20

C a s 9 mRNA、s g 2 8 2、及びLNPを実施例1に記載のように調製した。

【0239】

脂質1、脂質2、または脂質3を含むLNP組成物を雌CD-1マウスに投与し、1mg/kg体重及び0.5mg/kg体重で実施例1に記載のように評価した。マウスのコホートを、実施例1の方法に従って次世代シーケンシング(NGS)及び血清TTRレベルによって肝臓での編集について測定した。

【0240】

図6A及び図6Bでは、PEG脂質製剤間の血清TTRレベル比較している。図6Aは血清TTRをμg/mL単位で示し、図6Bはノックダウンの割合(TSS%)としてのデータを示す。図6Cは肝臓で達成された編集率を示す。データは、被験PEG脂質の各々を含むLNP組成物は2モル%及び3モル%において効力を示し、脂質1は一貫して脂質2及び脂質3よりも性能が良好であることを示す。

30

【0241】

実施例7-リポドA類似体

リポドAの構造類似体を多数合成し、本明細書に記載のLNP組成物で試験した。

【0242】

合成：リポドAは、4,4-ビス(オクチルオキシ)ブタン酸(WO2015/095340の実施例13の「中間体13b」と、(9Z,12Z)-3-ヒドロキシ-2-(ヒドロキシメチル)プロピルオクタデカ-9,12-ジエノアート(「中間体13c」とを反応させてから、中間体13bと中間体13cの生成物を、3-ジエチルアミノ-1-プロパノールと反応させて頭部基を付加することにより作られる。(WO2015/095340の84~86頁を参照のこと)。

40

【0243】

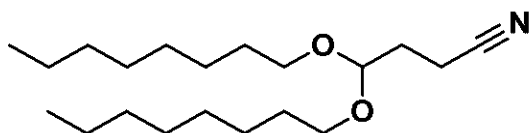
WO2015/095340の中間体13b(4,4-ビス(オクチルオキシ)ブタン酸)を4,4-ビス(オクチルオキシ)ブタンニトリルにより以下のとおり合成した。

【0244】

中間体13a: 4,4-ビス(オクチルオキシ)ブタンニトリル

50

【化 8】



【0245】

4, 4-ジエトキシブタンニトリル (9.4 g、60 mmol) 及びオクタン-1-オール (23.1 g、178 mmol) の混合物にピリジニウム *p*-トルエンスルホン酸 (748 mg、3.0 mmol) を室温で加えた。混合物を 105℃ に温め、反応槽を開にし、空気に触れる状態で、還流冷却器を取り付けずに 18 時間攪拌した。その後、反応混合物を室温まで冷却し、シリカゲル (酢酸エチル含有ヘキサンの勾配 0~5%) で精製して 10.1 g (31.0 mmol) の中間体 13a を澄明な油状物として得た。¹H NMR (400 MHz, CDCl₃) δ 4.55 (t, J = 5.3 Hz, 1H), 3.60 (dt, J = 9.2, 6.6 Hz, 2H), 3.43 (dt, J = 9.2, 6.6 Hz, 2H), 2.42 (t, J = 7.4 Hz, 2H), 1.94 (td, J = 7.4, 5.3 Hz, 2H), 1.63 – 1.50 (m, 4H), 1.38 – 1.19 (m, 20H), 0.93 – 0.82 (m, 6H) ppm.

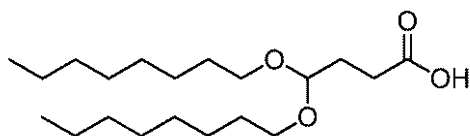
【0246】

次に、中間体 13a (8.42 g、31 mmol) のエタノール溶液 (30 mL) に、31 mL の水性水酸化カリウム (2.5 M、30.9 mL、77.3 mmol) を室温で加えた。反応槽に還流冷却器を取り付けた時点で、混合物を 110℃ まで加熱し、24 時間攪拌した。その後、混合物を室温まで冷却し、水性塩化水素酸 (1 N) で酸性化して pH 5 とし、ヘキサン中に 3 回抽出した。合わせた有機抽出物を水 (2 回) 及び食塩水で洗浄して無水硫酸マグネシウムで乾燥させた後、減圧濃縮し、8.15 g (23.6 mmol) の中間体 13b を澄明な油状物として得、これをそれ以上精製せずに使用した。¹H NMR (400 MHz, CDCl₃) δ 4.50 (t, J = 5.5 Hz, 1H), 3.57 (dt, J = 9.4, 6.7 Hz, 2H), 3.41 (dt, J = 9.3, 6.7 Hz, 2H), 2.40 (t, J = 7.4 Hz, 2H), 1.92 (td, J = 7.4, 5.3 Hz, 2H), 1.56 (m, 4H), 1.37 – 1.21 (m, 20H), 0.92 – 0.83 (m, 6H) ppm (下記の構造)。

【0247】

中間体 13b

【化 9】



【0248】

上記方法を使用して、C (炭素数 5、6、7、9、及び 10) -アセタール酸性中間体 (中間体 B3~F3 と呼び、以下に表す) を、適切なアルカン-1-オール試薬を使用して調製した。

【0249】

中間体 B3、4, 4-ビス (ペンチルオキシ) ブタン酸

10

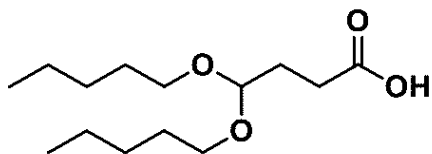
20

30

40

50

【化10】



【0250】

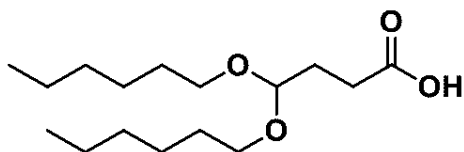
$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 4.52 (t, $J = 5.5$ Hz, 1 H), 3.58 (dt, $J = 9.3, 6.6$ Hz, 2 H), 3.41 (dt, $J = 9.3, 6.7$ Hz, 2 H), 2.45 (t, $J = 7.4$ Hz, 2 H), 1.94 (m, 2 H), 1.57 (m, 4 H), 1.32 (m, $J = 3.7$ Hz, 8 H), 0.95 – 0.83 (m, 6 H) ppm.

10

【0251】

中間体C3：4,4-ビス(ヘキシルオキシ)ブタン酸

【化11】



20

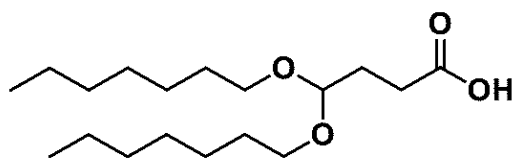
【0252】

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 4.44 (t, $J = 5.6$ Hz, 1 H), 3.49 (dt, $J = 9.3, 6.9$ Hz, 2 H), 3.39 (dt, $J = 9.3, 6.8$ Hz, 2 H), 2.12 (t, $J = 7.6$ Hz, 2 H), 1.79 (q, $J = 7.0$ Hz, 2 H), 1.54 (m, 4 H), 1.29 (m, 12 H), 0.94 – 0.82 (m, 6 H) ppm.

【0253】

中間体D3：4,4-ビス(ヘプチルオキシ)ブタン酸

【化12】



30

【0254】

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 8.85 (br s, 1 H), 4.46 (t, $J = 5.6$ Hz, 1 H), 3.52 (dt, $J = 9.4, 6.8$ Hz, 2 H), 3.39 (dt, $J = 9.3, 6.8$ Hz, 2 H), 2.26 (t, $J = 7.6$ Hz, 2 H), 1.85 (q, $J = 7.0$ Hz, 2 H), 1.53 (m, 4 H), 1.29 (m, 16 H), 0.94 – 0.80 (m, 6 H) ppm.

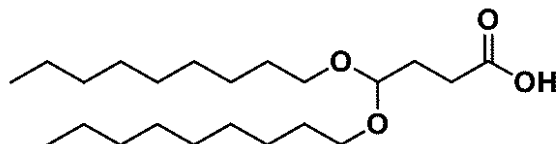
40

【0255】

中間体E3：4,4-ビス(ノニルオキシ)ブタン酸

50

【化13】



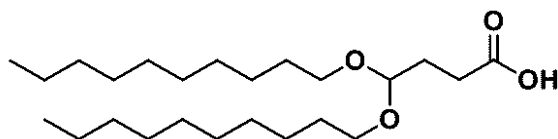
【0256】

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 5.32 (br s, 1H), 4.44 (t, $J = 5.6$ Hz, 1H), 3.49 (dt, $J = 9.3, 6.9$ Hz, 2H), 3.38 (dt, $J = 9.4, 6.9$ Hz, 2H), 2.10 (t, $J = 7.6$ Hz, 2H), 1.78 (q, $J = 7.0$ Hz, 2H), 1.53 (m, 4H), 1.27 (m, 24H), 0.88 (t, $J = 6.6$ Hz, 6H) ppm.

【0257】

中間体 F3: 4, 4-ビス(デシルオキシ)ブタン酸:

【化14】



【0258】

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 4.48 (t, $J = 5.5$ Hz, 1H), 3.55 (m, 2H), 3.42 (m, 2H), 2.29 (dd, $J = 10.8, 7.5$ Hz, 2H), 1.90 – 1.82 (m, 2H), 1.55 (m, 4H), 1.27 (m, 28H), 0.88 (t, $J = 6.7$ Hz, 6H) ppm.

【0259】

リピドAのアセタール類似体(C(8))は、C(炭素数5、6、7、9、または10)ーアセタール酸性中間体(B3～F3)と、中間体13cとを反応させてから、そのステップの生成物を3-ジエチルアミノ-1-プロパノールと反応させることによって合成した。(WO2015/095340の84～86頁を参照のこと)。各類似体を合成し、 $^1\text{H NMR}$ で特徴付けを行った(データ図示せず)。

【0260】

リピドA類似体を45モル%、DMG-PEG2kを2モル%、DSPCを9モル%、及びコレステロールを44モル%にし、N:P比4.5にてC7、C9、及びC10の類似体を製剤化した。それぞれの類似体は、リピドA類似体を55モル%、DMG-PEG2kを2.5モル%、DSPCを9モル%、及びコレステロールを38.5モル%にし、N:P比6にても製剤化された。脂質ナノ粒子成分を、脂質成分を上記のモル比で100%エタノールに溶解させた。RNAカーゴを25mMのクエン酸塩、100mMのNaCl、pH5.0に調製し、RNAカーゴの濃度を約0.45mg/mLとした。

【0261】

RNAカーゴには、配列番号43を含むCas9 mRNA及びsg282が含まれ、上記のように調製した。LNPを実施例1に記載のように形成した。

【0262】

C(5)及びC(6)リピドA類似体を含むLNP組成物を含めた拡張パネルのアセタール類似体を、先のパネルと共に試験した。2つの新たな類似体は、リピドA類似体を55モル%、DMG-PEG2kを2.5モル%、DSPCを9モル%、及びコレステロールを33.5モル%にし、N/P比6にて、上記のように製剤化した。分析により、LNPの大きさはいずれも120nmを下回り、PDIは0.2未満、及び封入されたRNAの割合(%)は80%より高いことが示された。製剤の分析結果を下記表28に記載する。

【表 28】

表 28.

L N P 識別番号	リピド A 類似体	類似体 (モ ル%)	P E G %	N / P	R N A 濃度 (m g / m L)	E E %	Z - 平 均 (n m)	P D I	数平均 (n m)
L N P 1 1 2 2	C 5 類似 体 (L P O 0 0 0 3 0 - 0 0 1)	5 5	2 . 5	6	0 . 0 6 3	8 8	1 1 8 . 8	0 . 1 0 3	8 8 . 1 7
L N P 1 1 2 3	C 6 類似 体 (L P O 0 0 0 3 1 - 0 0 1)	5 5	2 . 5	6	0 . 0 6 7	9 5	1 0 7 . 6	0 . 0 3 8	8 8 . 1
L N P 1 0 0 4	C 7 類似 体 (L P O O 0 0 2 0 - 0 0 1)	5 5	2 . 5	6	0 . 0 6 8	9 8	1 0 0	0 . 0 1 2	8 1 . 5 5
L N P 1 0 0 2	L P O O 0 0 0 1 - 0 1 1	5 5	2 . 5	6	0 . 0 6 7	9 8	9 5 . 0 6	0 . 0 1 5	7 8 . 9 5
L N P 1 0 0 6	C 9 類似 体 (L P O 0 0 0 2 1 - 0 0 1)	5 5	2 . 5	6	0 . 0 6 7	9 7	9 5 . 4 3	0 . 0 2 2	8 0 . 3 5
L N P 1 0 0 8	C 1 0 類 似 体 (L P O 0 0 0 2 2 - 0 0 1)	5 5	2 . 5	6	0 . 0 6 9	9 5	1 0 3 . 9	0 . 0 0 8	8 6 . 7 9

10

20

30

40

【0263】

水に溶解させた6-(p-トリルイジノ)-6-ナフタレンスルホン酸(「TNS」)を使用し、類似体をpKaについて評価した。本アッセイでは、0.1Mのリン酸緩衝液を4.5~10.5の範囲の異なるpH値で調製した。各類似体を100%エタノール中に個々に調製した。その後、脂質及びTNSを個々のpH緩衝液に加えてプレートに移し、SpectraMaxマイクロプレートリーダーで波長321~488nmにて分析した。値をプロットしてpKaを生成した。対数IC₅₀をpKaとして使用する。

【0264】

雌CD-1マウスに、実施例1に記載のように、0.3mg/kg(図7A-図7E)

50

、または1kg当たり0.1mg（図7F～図7G）を投与した。手短に言えば、Charles River Laboratoriesから得た雌CD-1マウス（n=5/群）にLNP組成物をさまざまな用量で投与した。剖検時（投与後7日）、TTR分析用に血清を採取し、編集分析用に肝臓を採取した。血清TTRアッセイ及び編集率アッセイを実施例1に記載のように実施した。図7Aから図7Eの血清TTRレベル及び肝臓での編集では、すべての類似体が、体重1キログラム当たり0.3ミリグラムにおいてリピドAに匹敵する性能であることが示されている。図7F～図7Gでは、リピドAが最も効力が高かったが、新たに合成した類似体はいずれもTTRノックダウン及び肝臓での編集が適切であることが示されている。

【0265】

実施例8－用量反応曲線－初代Cyno肝細胞

肝臓の初代肝細胞。初代カニクイザル肝臓肝細胞（PCH）（Gibco）を解凍し、添加物を含有する肝細胞融解培地（Gibco、カタログCM7000）に再懸濁させ、その後、80gで4分間遠心分離にかけた。上清を廃棄し、ペレット状にした細胞を添加物入り肝細胞播種培地パック（Invitrogen、カタログA1217601及びCM3000）に再懸濁させた。細胞を計数し、Bio-coatコーラゲンコート96ウェルプレート（ThermoFisher、カタログ877272）に50,000細胞/ウェルの密度で播種した。播種細胞を沈殿させ、組織培養インキュベーター（37℃、5%CO₂雰囲気）に入れて24時間接着させてから、LNPを投与した。インキュベーション後、細胞を単層形成について確認し、肝細胞培地及び無血清の添加物パック（Invitrogen、カタログA1217601及びCM4000）で培地を交換した。

【0266】

本試験用のLNP製剤（LNP1021、LNP1022、LNP1023、LNP1024、LNP1025、及びLNP897）を上記のように調製した。

【0267】

カニクイザル初代肝細胞で修飾sgRNAを含有する脂質ナノ粒子製剤をさまざまな用量で試験し、用量反応曲線を作成した。播種して24時間培養した後、6%のカニクイザル血清を含有する肝細胞維持培地でLNPを37℃で5分間インキュベートした。インキュベーション後、LNPをカニクイザル初代肝細胞に、8.2倍の用量反応曲線で100ngのmRNAで開始して加えた。細胞を、処理後72時間溶解させ、実施例1に記載のようにNGS分析を行った。各種LNP組成物について編集率を決定したデータを図8Aにグラフ化している。Cas9 mRNA（配列番号48）及びU-除去Cas9 mRNA（配列番号43）での編集率（%）を図8Bに記載する。LNP組成物は表2（LNP897）及び表5（LNP1021、LNP1022、LNP1023、LNP1024、及びLNP1025）に記載されている。

【0268】

結果は比較効力評価の定量的評価を示しており、mRNA及びLNP組成物の両方が効力に影響を与えることを示している。

【0269】

実施例9－RNAカーゴ：mRNA及びgRNA合剤

この試験では、gRNAとmRNAの比が異なる場合のマウスにおけるインビボでの有効性を評価した。ORFの配列番号4、HSD 5'UTR、ヒトアルブミン3'UTR、コザック配列、及びポリA尾部を有するCleanCap（商標）キャップ化Cas9 mRNAを、実施例1に示すウリジン三リン酸の代わりにN1-メチルプソイドウリジン三リン酸を用いてIVT合成により作製した。

【0270】

LNP製剤を、リピドA、コレステロール、DSPC、及びPEG2k-DMGのモル比を50:38:9:3にし、N:P比6にて記載mRNA及びsg282（配列番号42; G282）から実施例2に記載のように調製した。製剤のgRNA:Cas9 mRNA重量比は表29に示すとおりであった。

【表 29】

表 29.

L N P 識 別番号	ガイド：C a s 9 m R NA比 (w/ w)	R N A 濃度 (m g / m L)	E E (%)	Z - 平均 サ イ ズ (n m)	粒 子 P D I	数 平 均 (n m)
1 1 1 0	8 : 1	0 . 9 2	9 9	6 9 . 5 2	0 . 0 2 2	5 6 . 4 7
1 1 1 1	4 : 1	0 . 8 6	9 7	7 6 . 6 5	0 . 0 6 5	5 7 . 3 6
1 1 1 2	2 : 1	0 . 9 0	9 9	7 6 . 5 8	0 . 0 3 6	6 3 . 1 1
1 1 1 3	1 : 1	0 . 9 7	9 9	7 6 . 6 0	0 . 0 7 1	5 8 . 9 2
1 1 1 4	1 : 2	1 . 0 5	9 9	7 6 . 3 4	0 . 0 1 8	6 2 . 8 2
1 1 1 5	1 : 4	0 . 6 5	9 9	8 2 . 6 4	0 . 0 1 8	6 6 . 6 3
1 1 1 6	1 : 8	0 . 7 5	1 0 0	8 2 . 0 1	0 . 0 3 9	6 5 . 0 5

【0271】

インビボでの特性解析をするため、上記のLNPを、1kg当たり0.1mgのトータルRNA（ガイドRNA（mg）+mRNA（mg））にてマウスに投与した（n=5／群）。投与後7～9日目、動物を殺して血液及び肝臓を採取し、血清TTR及び肝臓での編集を実施例1に記載のように測定した。血清TTR及び肝臓での編集の結果を図9A及び図9Bに示す。陰性対照マウスにはTSSビヒクルを投与した。

【0272】

さらに、上記LNPをマウスに投与し、その際、mRNAを1kg当たり0.05mgのmRNA（n=5／群）という一定用量にし、gRNAを1kg当たり0.06mg～1kg当たり0.4mgで用量を変えた。投与後7～9日目、動物を殺して血液及び肝臓を採取し、血清TTR及び肝臓での編集を測定した。血清TTR及び肝臓での編集の結果を図9C及び図9Dに示す。陰性対照マウスにはTSSビヒクルを投与した。

【0273】

実施例10－中性脂質

LNPのインビボでの有効性を評価するため、実施例2のmRNA、及びsg534（配列番号72；G534）を用いて実施例2に記載のようにLNP製剤を調製した。脂質ナノ粒子成分を下記の脂質成分モル比で100％エタノールに溶解させた。手短に言えば、25mMのクエン酸塩及び100mMのNaClの緩衝液（pH5.0）にRNAカーゴを調製し、RNAカーゴの濃度を約0.45mg/mLとした。LNPは、脂質アミンとRNAリン酸塩（N：P）のモル比を約6、mRNAとgRNAの比を重量基準で1：2にして製剤化された。

【0274】

LNP製剤を、平均粒径、多分散度（pdi）、トータルRNA含量及びRNAの封入率について実施例1に記載のように分析した。平均粒径、多分散度（PDI）、トータルRNA含量及びRNAの封入率の分析を表30に示す。脂質のモル比は、アミン脂質（リ

ピドA)／中性脂質／ヘルパー脂質(コレステロール)／PEG脂質(PEG2k-DMG)として記載される。中性脂質は、示されているように、DSP、DPPCであるかまたは不含であった。

【表30】

表30. LNP組成物及びデータ。(脂質のモル比は、アミン脂質(リピドA)／中性脂質／ヘルパー脂質(コレステロール)／PEG脂質(PEG2k-DMG)として記載される)。

試料識別番号	中性脂質	モル比	RNA濃度 (mg/mL)	EE%	Z-平均 (nm)	PDI	数平均 (nm)	肝臓での編集率(%)	血清TR(μg/mL)	TTTR(KD)
TSS 対照								0.0	1248.9	
CO2 41	—	50.0 /0.0 /47.0 /3.0	1.46	94	75.64	0.090	54.21	1.8	1070.2	14.3
CO2 42	—	59.0 /0.0 /38.0 /3.0	1.51	94	92.25	0.019	75.56	12.06	819.4	34.4

10

20

30

40

50

CO ₂ 43	—	54.5 /0.0 /42. 5/3. 0	1.6 2	94	78. 90	0.0 52	61.49	3.3	126 0.5	— 0.9
CO ₂ 44	DSP C	50.0 /9.0 /39. 0/2. 0	1.5 0	93	10 1.3	0.0 44	80.73	27.4	741. 0	4 0.7
CO ₀ 34	DSPC	50.0 /9.0 /38. 0/3. 0	1.4 8	97	84. 23	0.0 40	66.96	34.2	630. 1	4 9.6
CO ₂ 45	DSP C	52.5 /4.0 /42. 5/3. 0	1.5 5	95	81. 88	0.0 54	64.54	5.8	846. 6	3 2.2
CO ₂ 46	DPP C	50.0 /9.0 /38. 0/3. 0	1.5 2	96	87. 11	0.0 40	70.04	35.9	528. 6	5 7.7
CO ₂ 47	DPP C	52.5 /4.0 /42. 5/3. 0	1.5 4	97	83. 67	0.0 50	66.43	18.3	726. 6	4 1.8

10

20

30

【0275】

インビボでの特性解析をするため、雌Sprague Dawleyラットに上記のLNPを体重1kg当たり0.3mgのトータルRNA（ガイドRNA及びmRNA）にて静脈内投与した。ラットは1群当たり5匹であった。投与後7日目、動物を殺して血液及び肝臓を採取し、血清TTR及び肝臓での編集を実施例1に記載のように測定した。陰性対照動物にはTSSビヒクルを投与した。血清TTR及び肝臓での編集の結果を図10A及び図10B、及び表30（上記）に示す。

40

50

【表 3 1】

開示配列の簡単な説明

配列番号	詳細
1	開始コドン及び終止コドンを有し、表 3 に記載の最小ウリジンコドンのチミジン類似体を使用する、C a s 9 の DNA コード配列
2	一般にヒトで高発現するコドンを使用する C a s 9 の DNA コード配

	列	
3	C 末端の 7 個のアミノ酸として 1 つの核局在化シグナル (1 x NLS) を有する C a s 9 のアミノ酸配列	10
4	開始コドン及び終止コドンを有し、表 3 に記載の最小ウリジンコドンを使用する C a s 9 mRNA ORF	
5	開始コドン及び終止コドンを有し、一般にヒトで高発現するコドンを使用する C a s 9 mRNA ORF	
6	C 末端の 7 個のアミノ酸として 1 x NLS を有する C a s 9 ニッカーゼのアミノ酸配列	
7	開始コドン及び終止コドンを有し、表 3 に記載の最小ウリジンコドンを使用する、配列番号 6 をコードする C a s 9 ニッカーゼ mRNA ORF	
8	C 末端の 7 個のアミノ酸として 1 x NLS を有する d C a s 9 のアミノ酸配列	20
9	開始コドン及び終止コドンを有し、表 3 に記載の最小ウリジンコドンを使用する、配列番号 8 をコードする d C a s 9 mRNA ORF	
10	表 3 に記載の最小ウリジンコドンを使用する C a s 9 mRNA コード配列 (開始コドンまたは終止コドンはない; 融合タンパク質コード配列に含める際に好適)	
11	表 3 に記載の最小ウリジンコドンを使用する C a s 9 ニッカーゼ mRNA コード配列 (開始コドンまたは終止コドンはない; 融合タンパク質コード配列に含める際に好適)	
12	表 3 に記載の最小ウリジンコドンを使用する d C a s 9 mRNA コード配列 (開始コドンまたは終止コドンはない; 融合タンパク質コード配列に含める際に好適)	30
13	C a s 9 のアミノ酸配列 (NLS なし)	
14	開始コドン及び終止コドンを有し、表 3 に記載の最小ウリジンコドンを使用する、配列番号 13 をコードする C a s 9 mRNA ORF	
15	表 3 に記載の最小ウリジンコドンを使用する、配列番号 13 をコードする C a s 9 コード配列 (開始コドンまたは終止コドンはない; 融合タンパク質コード配列に含める際に好適)	
16	C a s 9 ニッカーゼのアミノ酸配列 (NLS なし)	
17	開始コドン及び終止コドンを有し、表 3 に記載の最小ウリジンコドンを使用する、配列番号 16 をコードする C a s 9 ニッカーゼ mRNA ORF	40
18	表 3 に記載の最小ウリジンコドンを使用する、配列番号 16 をコードする C a s 9 ニッカーゼコード配列 (開始コドンまたは終止コドンはない; 融合タンパク質コード配列に含める際に好適)	
19	d C a s 9 のアミノ酸配列 (NLS なし)	
20	開始コドン及び終止コドンを有し、表 3 に記載の最小ウリジンコドンを使用する、配列番号 13 をコードする d C a s 9 mRNA ORF	

2 1	表 3 に記載の最小ウリジンコドンを使用する、配列番号 1 3 をコードする d C a s 9 コード配列 (開始コドンまたは終止コドンはない; 融合タンパク質コード配列に含める際に好適)
2 2	C 末端のアミノ酸として 2 つの核局在化シグナル (2 x N L S) を有する、C a s 9 のアミノ酸配列
2 3	開始コドン及び終止コドン有し、表 3 に記載の最小ウリジンコドンを使用する、配列番号 1 3 をコードする C a s 9 m R N A O R F
2 4	表 3 に記載の最小ウリジンコドンを使用する、配列番号 1 3 をコードする C a s 9 コード配列 (開始コドンまたは終止コドンはない; 融合タンパク質コード配列に含める際に好適)
2 5	C 末端のアミノ酸として 2 つの核局在化シグナルを有する、C a s 9 ニッカーゼのアミノ酸配列
2 6	開始コドン及び終止コドン有し、表 3 に記載の最小ウリジンコドンを使用する、配列番号 1 6 をコードする C a s 9 ニッカーゼ m R N A O R F
2 7	表 3 に記載の最小ウリジンコドンを使用する、配列番号 1 6 をコードする C a s 9 ニッカーゼコード配列 (開始コドンまたは終止コドンはない; 融合タンパク質コード配列に含める際に好適)
2 8	C 末端のアミノ酸として 2 つの核局在化シグナルを有する、d C a s 9 のアミノ酸配列
2 9	開始コドン及び終止コドン有し、表 3 に記載の最小ウリジンコドンを使用する、配列番号 1 3 をコードする d C a s 9 m R N A O R F
3 0	表 3 に記載の最小ウリジンコドンを使用する、配列番号 1 3 をコードする d C a s 9 コード配列 (開始コドンまたは終止コドンはない; 融合タンパク質コード配列に含める際に好適)
3 1	T 7 プロモーター
3 2	ヒトベータグロビン 5' UTR
3 3	ヒトベータグロビン 3' UTR
3 4	ヒトアルファグロビン 5' UTR
3 5	ヒトアルファグロビン 3' UTR
3 6	X e n o p u s l a e v i s ベータグロビン 5' UTR
3 7	X e n o p u s l a e v i s ベータグロビン 3' UTR
3 8	ウシ成長ホルモン 5' UTR
3 9	ウシ成長ホルモン 3' UTR
4 0	M u s m u s c u l u s ヘモグロビンアルファ、成体鎖 1 (H b a - a 1)、3' UTR
4 1	H S D 1 7 B 4 5' UTR
4 2	マウス T T R 遺伝子指向性 G 2 8 2 単一ガイド RNA
4 3	H S D の 5' UTR、配列番号 4 に対応する O R F、コザック配列、及び A L B の 3' UTR を有する C a s 9 転写産物

10

20

30

40

50

44	HSDの5' UTR、配列番号4に対応するORF、及びALBの3' UTRを有するCas9転写産物
45	U含量が19.36%である代替的Cas9 ORF
46	HSDの5' UTR、配列番号45に対応するORF、コザック配列、及びALBの3' UTRを有するCas9転写産物
47	HSDの5' UTR、配列番号45に対応するORF、及びALBの3' UTRを有するCas9転写産物
48	一般にヒトで高発現するコドンを使用するCas9 ORFを含むCas9転写産物
49	一般にヒトで高発現するコドンを使用するCas9 ORFを有する、コザック配列を含むCas9転写産物
50	スプライス部位が除去されたCas9 ORF; U含量12.75%
51	HSDの5' UTR、配列番号50に対応するORF、コザック配列、及びALBの3' UTRを有するCas9転写産物
52	一般にヒトで頻用される最小ウリジンコドンを含むCas9 ORF; U含量12.75%
53	HSDの5' UTR、配列番号52に対応するORF、コザック配列、及びALBの3' UTRを有するCas9転写産物
54	一般にヒトでまれに使用される最小ウリジンコドンを含むCas9 ORF; U含量12.75%
55	HSDの5' UTR、配列番号54に対応するORF、コザック配列、及びALBの3' UTRを有するCas9転写産物
56	CleanCap (商標)と共に使用するための最初の3ヌクレオチドとしてのAGG、HSDの5' UTR、配列番号4に対応するORF、コザック配列、及びALBの3' UTRを有するCas9転写産物
57	CMV由来の5' UTR、配列番号4に対応するORF、コザック配列、及びALBの3' UTRを有するCas9転写産物
58	HBB由来の5' UTR、配列番号4に対応するORF、コザック配列、及びHBBの3' UTRを有するCas9転写産物
59	XBG由来の5' UTR、配列番号4に対応するORF、コザック配列、及びXBGの3' UTRを有するCas9転写産物
60	CleanCap (商標)と共に使用するための最初の3ヌクレオチドとしてのAGG、XBG由来の5' UTR、配列番号4に対応するORF、コザック配列、及びXBGの3' UTRを有するCas9転写産物
61	CleanCap (商標)と共に使用するための最初の3ヌクレオチドとしてのAGG、HSD由来の5' UTR、配列番号4に対応するORF、コザック配列、及びALBの3' UTRを有するCas9転写産物
62	30/30/39ポリA配列
63	ポリA100配列
64	マウスTTR遺伝子指向性G209単一ガイドRNA
65	開始コドン及び終止コドンを含み、表3に記載の最小ウリジンコドンを使用する、Neisseria meningitidis Cas9

10

20

30

40

50

	s 9 をコードする O R F
6 6	表 3 に記載の最小ウリジンコドンを使用する、Neisseria meningitidis Cas 9 をコードする O R F（開始コドンまたは終止コドンはない；融合タンパク質コード配列に含める際に好適）
6 7	配列番号 6 5 を含む転写産物（Neisseria meningitidis Cas 9 をコードする）
6 8	Neisseria meningitidis Cas 9 のアミノ酸配列
6 9	ラット T T R 遺伝子指向性 G 3 9 0 単一ガイド RNA
7 0	カニクイザル T T R 遺伝子指向性 G 5 0 2 単一ガイド RNA
7 1	カニクイザル T T R 遺伝子指向性 G 5 0 9 単一ガイド RNA
7 2	ラット T T R 遺伝子指向性 G 5 3 4 単一ガイド RNA

10

【 0 2 7 6 】

配列そのものについては下記配列表を参照のこと。転写産物配列には一般に、A R C A と共に使用するための最初の 3 ヌクレオチドとして G G G が含まれるか、または C l e a n C a p（商標）と共に使用するための最初の 3 ヌクレオチドとして A G G が含まれる。したがって、最初の 3 ヌクレオチドは、ワクシニアのキャッピング酵素などの、他のキャッピング手法と共に使用するために修飾され得る。プロモーター及びポリ A 配列は転写産物配列には含まれない。T 7 プロモーター（配列番号 3 1）などのプロモーター及び配列番号 6 2 または 6 3 などのポリ A 配列は、開示の転写産物配列に 5' 末端及び 3' 末端でそれぞれ付加することができる。ほとんどのヌクレオチド配列は DNA として提供されるが、それらは T を U に変えることにより容易に RNA に変換可能である。

20

【 0 2 7 7 】

配列表

以下の配列表は、本明細書に開示する配列の一覧を提供する。DNA 配列（T を含む）が RNA に関して参照される場合は、T を U に置き換えなければならない（状況に応じて修飾の場合と非修飾の場合がある）、その逆の場合は、逆の置き換えをしなければならないと理解される。

30

【表 3 2】

詳細	配列	配 列 番 号
C a s 9 D N A コード配 列 2	A T G G A C A A G A A G T A C A G C A T C G G A C T G G A C A T C G G A A C A A A C A G C G T C G G A T G G G C A G T C A T C A C A G A C G A A T A C A A G G T C C C G A G C A A G A A G T T C A A G G T C C T G G G A A A C A C A G A C A G A C A C A G C A T C A A G A A G A A C C T G A T C G G A G C A C T G C T G T T C G A C A G C G G A G A A A C A G C A G A A G C A A C A A G A C T G A A G A G A A C A G C A A G A A G A A G A T A C A C A A G A A G A A A G A A C A G A A T C T G C T A C C T G C A G G A A A T C T T C A G C A A C G A A A T G G C A A A G G T C G A C G A C A	1

40

50

	GCTTCTTCCACAGACTGGAAGAAAGC TTCCTGGTCGAAGAAGACAAGAAGCA CGAAAGACACCCGATCTTCGGAAACA TCGTGACGAAGTCGCATACCACGAA AAGTACCCGACAATCTACCACCTGAG AAAGAAGCTGGTCGACAGCACAGACA AGGCAGACCTGAGACTGATCTACCTG GCACTGGCACACATGATCAAGTTCAG AGGACACTTCCTGATCGAAGGAGACC TGAACCCGGACAACAGCGACGTGAC AAGCTGTTTCATCCAGCTGGTCCAGAC ATACAACCAGCTGTTTCGAAGAAAACC CGATCAACGCAAGCGGAGTCGACGCA AAGGCAATCCTGAGCGCAAGACTGAG CAAGAGCAGAAGACTGGAAAACCTGA TCGCACAGCTGCCGGGAGAAAAGAAG AACGGACTGTTTCGGAAACCTGATCGC ACTGAGCCTGGGACTGACACCGAACT TCAAGAGCAACTTCGACCTGGCAGAA GACGCAAAGCTGCAGCTGAGCAAGGA CACATACGACGACGACCTGGACAACC TGCTGGCACAGATCGGAGACCAGTAC GCAGACCTGTTCTGCTGGCAGCAAAGAA CCTGAGCGACGCAATCCTGCTGAGCG ACATCCTGAGAGTCAACACAGAAATC ACAAAGGCACCGCTGAGCGCAAGCAT GATCAAGAGATACGACGAACACCACC AGGACCTGACACTGCTGAAGGCACTG GTCAGACAGCAGCTGCCGGAAAAGTA CAAGGAAATCTTCTTCGACCAGAGCA AGAACGGATACGCAGGATACATCGAC GGAGGAGCAAGCCAGGAAGAATTCTA CAAGTTTCATCAAGCCGATCCTGGAAA AGATGGACGGAACAGAAGAAGTCTG GTCAAGCTGAACAGAGAAGACCTGCT GAGAAAGCAGAGAACATTCGACAACG GAAGCATCCCGCACCAAGATCCACCTG GGAGAAGTGCACGCAATCCTGAGAAG ACAGGAAGACTTCTACCCGTTCTCTGA AGGACAACAGAGAAAAGATCGAAAAG ATCCTGACATTCAGAATCCCGTACTA CGTCGGACCGCTGGCAAGAGGAAACA	

10

20

30

40

50

	GCAGATTTCGCATGGATGACAAGAAAG AGCGAAGAAACAATCACACCGTGGAA CTTCGAAGAAGTCGTGACAAGGGAG CAAGCGCACAGAGCTTCATCGAAAGA ATGACAAACTTCGACAAGAACCTGCC GAACGAAAAGGTCCTGCCGAAGCACA GCCTGCTGTACGAATACTTCACAGTC TACAACGAACTGACAAAGGTCAAGTA CGTCACAGAAGGAATGAGAAAGCCGG CATTCCTGAGCGGAGAACAGAAGAAG GCAATCGTCGACCTGCTGTTCAAGAC AAACAGAAAGGTCACAGTCAAGCAGC TGAAGGAAGACTACTTCAAGAAGATC GAATGCTTCGACAGCGTCGAAATCAG CGGAGTCGAAGACAGATTCAACGCAA GCCTGGGAACATACCACGACCTGCTG AAGATCATCAAGGACAAGGACTTCCT GGACAACGAAGAAAACGAAGACATCC TGGAAGACATCGTCCTGACACTGACA CTGTTTCGAAGACAGAGAAATGATCGA AGAAAGACTGAAGACATACGCACACC TGTTTCGACGACAAGGTCATGAAGCAG CTGAAGAGAAGAAGATACACAGGATG GGGAAGACTGAGCAGAAAGCTGATCA ACGGAATCAGAGACAAGCAGAGCGGA AAGACAATCCTGGACTTCCTGAAGAG CGACGGATTTCGCAAACAGAAACTTCA TGCAGCTGATCCACGACGACAGCCTG ACATTCAAGGAAGACATCCAGAAGGC ACAGGTCAGCGGACAGGGAGACAGCC TGCACGAACACATCGCAAACCTGGCA GGAAGCCCGGCAATCAAGAAGGGAAT CCTGCAGACAGTCAAGGTCGTGACG AACTGGTCAAGGTCATGGGAAGACAC AAGCCGGAAAACATCGTCATCGAAAT GGCAAGAGAAAACCAGACAACACAGA AGGGACAGAAGAACAGCAGAGAAAGA ATGAAGAGAATCGAAGAAGGAATCAA GGAACCTGGGAAGCCAGATCCTGAAGG AACACCCGGTCGAAAACACACAGCTG CAGAACGAAAAGCTGTACCTGTACTA CCTGCAGAACGGAAGAGACATGTACG	
		10
		20
		30
		40

TCGACCAGGAACTGGACATCAACAGA CTGAGCGACTACGACGTCGACCACAT CGTCCCGCAGAGCTTCCTGAAGGACG ACAGCATCGACAACAAGGTCCTGACA AGAAGCGACAAGAACAGAGGAAAGAG CGACAACGTCCCGAGCGAAGAAGTCG TCAAGAAGATGAAGAACTACTGGAGA CAGCTGCTGAACGCAAAGCTGATCAC ACAGAGAAAGTTTCGACAACCTGACAA AGGCAGAGAGAGGAGGACTGAGCGAA CTGGACAAGGCAGGATTCATCAAGAG ACAGCTGGTTCGAAACAAGACAGATCA CAAAGCACGTTCGCACAGATCCTGGAC AGCAGAATGAACACAAAGTACGACGA AAACGACAAGCTGATCAGAGAAGTCA AGGTCATCACACTGAAGAGCAAGCTG GTCAGCGACTTTCAGAAAGGACTTCCA GTTCTACAAGGTCAGAGAAATCAACA ACTACCACCACGCACACGACGCATAC CTGAACGCAGTCGTCGGAACAGCACT GATCAAGAAGTACCCGAAGCTGGAAA GCGAATTTCGTCTACGGAGACTACAAG GTCTACGACGTCAGAAAGATGATCGC AAAGAGCGAACAGGAAATCGGAAAGG CAACAGCAAAGTACTTCTTCTACAGC AACATCATGAAC TTCTTCAAGACAGA AATCACACTGGCAAACGGAGAAATCA GAAAGAGACCGCTGATCGAAACAAAC GGAGAAACAGGAGAAATCGTCTGGGA CAAGGGAAGAGACTTCGCAACAGTCA GAAAGGTCCTGAGCATGCCGCAGGTC AACATCGTCAAGAAGACAGAAGTCCA GACAGGAGGATTCAGCAAGGAAAGCA TCCTGCCGAAGAGAAACAGCGACAAG CTGATCGCAAGAAAGAAGGACTGGGA CCCGAAGAAGTACGGAGGATTCGACA GCCCGACAGTCGCATACAGCGTCCTG GTCGTCGCAAAGGTCGAAAAGGGAAA GAGCAAGAAGCTGAAGAGCGTCAAGG AACTGCTGGGAATCACCAATCATGGAA AGAAGCAGCTTTCGAAAAGAACCCGAT CGACTTCCTGGAAGCAAAGGGATACA	

10

20

30

40

50

	AGGAAGTCAAGAAGGACCTGATCATC AAGCTGCCGAAGTACAGCCTGTTCGA ACTGGAAAACGGAAGAAAGAGAATGC TGGCAAGCGCAGGAGAACTGCAGAAAG GGAAACGAACTGGCACTGCCGAGCAA GTACGTCAAACCTTCCTGTACCTGGCAA GCCACTACGAAAAGCTGAAGGGGAAGC CCGGAAGACAACGAACAGAAGCAGCT GTTTCGTGGAACAGCACAAAGCACTACC TGGACGAAATCATCGAACAGATCAGC GAATTTCAGCAAGAGAGTTCATCCTGGC AGACGCAAACCTGGACAAGGTCCTGA GCGCATACAACAAGCACAGAGACAAG CCGATCAGAGAACAGGCAGAAAACAT CATCCACCTGTTTCACACTGACAAACC TGGGAGCACCGGCAGCATTCAAGTAC TTCGACACAACAATCGACAGAAAGAG ATACACAAGCACAAAGGAAGTCCTGG ACGCAACACTGATCCACCAGAGCATC ACAGGACTGTACGAAACAAGAATCGA CCTGAGCCAGCTGGGAGGAGACGGAG GAGGAAGCCCGAAGAAGAAGAGAAAG GTCTAG	
C a s 9 D NAコード配 列 1	ATGGATAAGAAGTACTCAATCGGGCT GGATATCGGAACTAATTCCGTGGGTT GGGCAGTGATCACGGATGAATACAAA GTGCCGTCCAAGAAGTTCAAGGTCCT GGGGAACACCGATAGACACAGCATCA AGAAAAATCTCATCGGAGCCCTGCTG TTTGACTCCGGCGAAACCGCAGAAGC GACCCGGCTCAAACGTACCGCGAGGC GACGCTACACCCGGCGGAAGAATCGC ATCTGCTATCTGCAAGAGATCTTTTC GAACGAAATGGCAAAGGTCGACGACA GCTTCTTCCACCGCCTGGAAGAATCT TTCCTGGTGGAGGAGGACAAGAAGCA TGAACGGCATCCTATCTTTGGAAACA TCGTGACGAAGTGGCGTACCACGAA AAGTACCCGACCATCTACCATCTGCG GAAGAAGTTGGTTGACTCAACTGACA AGGCCGACCTCAGATTGATCTACTTG GCCCTCGCCCATATGATCAAATTCCG	2

10

20

30

40

50

50

	CGTTACTGAAGGAATGAGGAAGCCGG CCTTTCTGTCCGGAGAACAGAAGAAA GCAATTGTCGATCTGCTGTTCAAGAC CAACCGCAAGGTGACCGTCAAGCAGC TTAAAGAGGACTACTTCAAGAAGATC GAGTGTTTCGACTCAGTGGAATCAG CGGGGTGGAGGACAGATTCAACGCTT CGCTGGGAACCTATCATGATCTCCTG AAGATCATCAAGGACAAGGACTTCC TGACAACGAGGAGAACGAGGACATCC TGGAAGATATCGTCCTGACCTTGACC CTTTTCGAGGATCGCGAGATGATCGA GGAGAGGCTTAAGACCTACGCTCATC TCTTCGACGATAAGGTCATGAAACAA CTCAAGCGCCGCCGGTACACTGGTTG GGGCCGCCTCTCCCGCAAGCTGATCA ACGGTATTTCGCGATAAACAGAGCGGT AAAACCTATCCTGGATTTCTCAAATC GGATGGCTTCGCTAATCGTAACTTCA TGCAATTGATCCACGACGACAGCCTG ACCTTTAAGGAGGACATCCAAAAAGC ACAAGTGTCCGGACAGGGAGACTCAC TCCATGAACACATCGCGAATCTGGCC GGTTTCGCCGGCGATTAAAGAAGGGAAT TCTGCAAACTGTGAAGGTGGTCGACG AGCTGGTGAAGGTGATGGGACGGCAC AAACCGGAGAATATCGTGATTGAAAT GGCCCGAGAAAACAGACTACCCAGA AGGGCCAGAAAACTCCCGCGAAAGG ATGAAGCGGATCGAAGAAGGAATCAA GGAGCTGGGCAGCCAGATCCTGAAAG AGCACCCGGTGGAAAACACGCAGCTG CAGAACGAGAAGCTCTACCTGTACTA TTTGCAAAATGGACGGGACATGTACG TGGACCAAGAGCTGGACATCAATCGG TTGTCTGATTACGACGTGGACCACAT CGTTCCACAGTCCTTTCTGAAGGATG ACTCGATCGATAACAAGGTGTTGACT CGCAGCGACAAGAACAGAGGGAAGTC AGATAATGTGCCATCGGAGGAGGTGCG TGAAGAAGATGAAGAATTACTGGCGG CAGCTCCTGAATGCGAAGCTGATTAC	

10

20

30

40

50

	CCAGAGAAAGTTTGACAATCTCACTA AAGCCGAGCGCGGCGGACTCTCAGAG CTGGATAAGGCTGGATTTCATCAAACG GCAGCTGGTCGAGACTCGGCAGATTA CCAAGCACGTGGCGCAGATCTTGGAC TCCCGCATGAACACTAAATACGACGA GAACGATAAGCTCATCCGGGAAGTGA AGGTGATTACCCTGAAAAGCAAACCTT GTGTCGGACTTTTCGGAAGGACTTTCA GTTTTTACAAAGTGAGAGAAATCAACA ACTACCATCACGCGCATGACGCATAC CTCAACGCTGTGGTCGGTACCGCCCT GATCAAAAAGTACCCTAAACTTGAAT CGGAGTTTGTGTACGGAGACTACAAG GTCTACGACGTGAGGAAGATGATAGC CAAGTCCGAACAGGAAATCGGGAAAG CAACTGCGAAATACTTCTTTTACTCA AACATCATGAACCTTTTTTCAAGACTGA AATTACGCTGGCCAATGGAGAAATCA GGAAGAGGCCACTGATCGAAACTAAC GGAGAAACGGGCGAAATCGTGTGGGA CAAGGGCAGGGACTTCGCAACTGTTC GCAAAGTGCTCTCTATGCCGCAAGTC AATATTGTGAAGAAAACCGAAGTGCA AACCGGCGGATTTTTCAAAGGAATCGA TCCTCCCAAAGAGAAATAGCGACAAG CTCATTGCACGCAAGAAAGACTGGGA CCCGAAGAAGTACGGAGGATTTCGATT CGCCGACTGTTCGCATACTCCGTCCTC GTGGTGGCCAAGGTGGAGAAGGGAAA GAGCAAAAAGCTCAAATCCGTCAAAG AGCTGCTGGGGATTACCATCATGGAA CGATCCTCGTTTCGAGAAGAACCCGAT TGATTTCTCTCGAGGCGAAGGGTTACA AGGAGGTGAAGAAGGATCTGATCATC AAACTCCCCAAGTACTCACTGTTTCGA ACTGGAAAATGGTCGGAAGCGCATGC TGGCTTCGGCCGGAGAACTCCAAAAA GGAAATGAGCTGGCCTTGCCTAGCAA GTACGTCAACTTCCTCTATCTTGCTT CGCACTACGAAAAACTCAAAGGGTCA CCGGAAGATAACGAACAGAAGCAGCT	

10

20

30

40

50

	TTTCGTGGAGCAGCACAAAGCATTATC TGGATGAAATCATCGAACAAATCTCC GAGTTTTCAAAGCGCGTGATCCTCGC CGACGCCAACCTCGACAAAGTCCTGT CGGCCTACAATAAGCATAGAGATAAG CCGATCAGAGAACAGGCCGAGAACAT TATCCACTTGTTTACCCCTGACTAACC TGGGAGCCCCAGCCGCCTTCAAGTAC TTCGATACTACTATCGATCGCAAAAG ATACACGTCCACCAAGGAAGTTCTGG ACGCGACCCTGATCCACCAAAGCATC ACTGGACTCTACGAAACTAGGATCGA TCTGTTCGCAGCTGGGTGGCGATGGCG GTGGATCTCCGAAAAAGAAGAGAAAG GTGTAATGA	
C a s 9 アミ ノ酸配列	MDKKYSIGLDIGTNSVGWAVITDEYK VPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFHRL EES FLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAE DAKLQLSKD TYDDDLNLLAQIGDQY ADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKAL VRQQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHL GELHA ILRRQEDFY PFLKDNREKIEK ILTFRIPYYVGPLARGNSRFAWMTRK SEETITPWNFEEVVDKGASAQSFIER MTNFDKNLPNEKVLPHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKI ECFDSVEISGVEDRFNASLGTYHDL L KIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKYAHLFDDKVMKQ LKRRRYTGWGRLSRKLINGIRDKQSG	3

10

20

30

40

50

	K T I L D F L K S D G F A N R N F M Q L I H D D S L T F K E D I Q K A Q V S G Q G D S L H E H I A N L A G S P A I K K G I L Q T V K V V D E L V K V M G R H K P E N I V I E M A R E N Q T T Q K G Q K N S R E R M K R I E E G I K E L G S Q I L K E H P V E N T Q L Q N E K L Y L Y Y L Q N G R D M Y V D Q E L D I N R L S D Y D V D H I V P Q S F L K D D S I D N K V L T R S D K N R G K S D N V P S E E V V K K M K N Y W R Q L L N A K L I T Q R K F D N L T K A E R G G L S E L D K A G F I K R Q L V E T R Q I T K H V A Q I L D S R M N T K Y D E N D K L I R E V K V I T L K S K L V S D F R K D F Q F Y K V R E I N N Y H H A H D A Y L N A V V G T A L I K K Y P K L E S E F V Y G D Y K V Y D V R K M I A K S E Q E I G K A T A K Y F F Y S N I M N F F K T E I T L A N G E I R K R P L I E T N G E T G E I V W D K G R D F A T V R K V L S M P Q V N I V K K T E V Q T G G F S K E S I L P K R N S D K L I A R K K D W D P K K Y G G F D S P T V A Y S V L V V A K V E K G K S K K L K S V K E L L G I T I M E R S S F E K N P I D F L E A K G Y K E V K K D L I I K L P K Y S L F E L E N G R K R M L A S A G E L Q K G N E L A L P S K Y V N F L Y L A S H Y E K L K G S P E D N E Q K Q L F V E Q H K H Y L D E I I E Q I S E F S K R V I L A D A N L D K V L S A Y N K H R D K P I R E Q A E N I I H L F T L T N L G A P A A F K Y F D T T I D R K R Y T S T K E V L D A T L I H Q S I T G L Y E T R I D L S Q L G G D G G G S P K K K R K V	
C a s 9 m RNAオープン ンリーディン グフレーム (ORF) 2	A U G G A C A A G A A G U A C A G C A U C G G A C U G G A C A U C G G A A C A A A C A G C G U C G G A U G G G C A G U C A U C A C A G A C G A A U A C A A G G U C C C G A G C A A G A A G U U C A A G G U C C U G G G A A A C A C A G A C A G A C A C A G C A U C A A G A A G A A C C U G A U C G G A G C A C U G C U G U U C G A C A G C G G A G A A A C A G C A G A A G C A A C A A G A C U G A A G A G A A C A G C A A G A A G A A G A U A C A C A A G A A G A A A G A A C A G A A U C U G C U A C C U G C A G G A A A U C U U C A G C A A C G A A A U G G C A A A G G U C G A C G A C A G C U U C U U C C A C A G A C U G G A A G A A A G C U U C C U G G U C G A A G A A G A C A A G A A G C A	4

10

20

30

40

50

	CGAAAGACACCCGAUCUUCGGAAACA UCGU CGACGAAGUCGCAUACCACGAA AAGUACCCGACAAUCUACCACCUGAG AAAGAAGCUGGUCGACAGCACAGACA AGGCAGACCUGAGACUGAUCUACCUG GCACUGGCACACAUGAUCAAGUUCAG AGGACACUUCCUGAUCGAAGGAGACC UGAACC CGGACAACAGCGACGUCGAC AAGCUGUUCAUCCAGCUGGUCCAGAC AUACAACCAGCUGUUCGAAGAAAACC CGAUCAACGCAAGCGGAGUCGACGCA AAGGCAAUCCUGAGCGCAAGACUGAG CAAGAGCAGAAGACUGGAAAACCUGA UCGCACAGCUGCCGGGAGAAAAGAAG AACGGACUGUUCGGAAACCUGAUCGC ACUGAGCCUGGGACUGACACCGAACU UCAAGAGCAACUUCGACCUGGCAGAA GACGCAAAGCUGCAGCUGAGCAAGGA CACAUACGACGACGACCUGGACAACC UGCUGGCACAGAUCCGAGACCAGUAC GCAGACCUGUUCUGGCAGCAAAGAA CCUGAGCGACGCAAUCCUGCUGAGCG ACAUCCUGAGAGUCAACACAGAAAUC ACAAAGGCACCGCUGAGCGCAAGCAU GAUCAAGAGAUACGACGAACACCACC AGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUA CAAGGAAAUCUUCUUCGACCAGAGCA AGAACGGAUACGCAGGAUACAUCGAC GGAGGAGCAAGCCAGGAAGAAUUCUA CAAGUUCAUCAAGCCGAUCCUGGAAA AGAUGGACGGAACAGAAGAACUGCUG GUCAAGCUGAACAGAGAAGACCUGCU GAGAAAGCAGAGAACAUCGACAACG GAAGCAUCCCGCACCCAGAUCCACCUG GGAGAACUGCACGCAAUCCUGAGAAG ACAGGAAGACUUCUACCCGUUCCUGA AGGACAACAGAGAAAAGAUCGAAAAG AUCCUGACAUCAGAAUCCCGUACUA CGUCGGACCGCUGGCAAGAGGAAACA GCAGAUUCGCAUGGAUGACAAGAAAG AGCGAAGAAACAUAUCACACCCUGGAA	
		10
		20
		30
		40

	CUUCGAAGAAGUCGUCGACAAGGGAG CAAGCGCACAGAGCUUCAUCGAAAGA AUGACAAACUUCGACAAGAACCUGCC GAACGAAAAGGUCCUGCCGAAGCACA GCCUGCUGUACGAAUACUUCACAGUC UACAACGAACUGACAAAGGUCAAGUA CGUCACAGAAGGAAUGAGAAAGCCGG CAUUCCUGAGCGGAGAACAGAAGAAG GCAAUCGUCGACCUGCUGUUCAAGAC AAACAGAAAGGUCACAGUCAAGCAGC UGAAGGAAGACUACUUCAAGAAGAUC GAAUGCUUCGACAGCGUCGAAAUCAG CGGAGUCGAAGACAGAUUCAACGCAA GCCUGGGAACAUAACACGACCUGCUG AAGAUCAUCAAGGACAAGGACUUCCU GGACAACGAAGAAAACGAAGACAUCC UGGAAGACAUCGUCCUGACACUGACA CUGUUCGAAGACAGAGAAAUGAUCGA AGAAAGACUGAAGACAUACGCACACC UGUUCGACGACAAGGUCAUGAAGCAG CUGAAGAGAAGAAGAUACACAGGAUG GGGAAGACUGAGCAGAAAGCUGAUCA ACGGAAUCAGAGACAAGCAGAGCGGA AAGACAAUCCUGGACUUCCUGAAGAG CGACGGAUUCGCAAACAGAAACUUCA UGCAGCUGAUCCACGACGACAGCCUG ACAUUCAAGGAAGACAUCCAGAAGGC ACAGGUCAGCGGACAGGGAGACAGCC UGCACGAACACAUCGCAAACCUGGCA GGAAGCCCGGCAAUCAAGAAGGGAAU CCUGCAGACAGUCAAGGUUCGUCGACG AACUGGUCAAGGUCAUGGGAAGACAC AAGCCGGAAAACAUCGUCAUCGAAAU GGCAAGAGAAAACCAGACAACACAGA AGGGACAGAAGAACAGCAGAGAAAGA AUGAAGAGAAUCGAAGAAGGAAUCAA GGAACUGGGAAGCCAGAUC CUGAAGG AACACCCGGUCGAAAACACACAGCUG CAGAACGAAAAGCUGUACCUGUACUA CCUGCAGAACGGAAGAGACAUGUACG UCGACCAGGAACUGGACAUCAACAGA CUGAGCGACUACGACGUCGACCACAU	
		10
		20
		30
		40

	CGUCCCGCAGAGCUUCCUGAAGGACG ACAGCAUCGACAACAAGGUCCUGACA AGAAGCGACAAGAACAGAGGAAAGAG CGACAACGUCCCGAGCGAAGAAGUCG UCAAGAAGAUGAAGAACUACUGGAGA CAGCUGCUGAACGCAAAGCUGAUCAC ACAGAGAAAGUUCGACAACCUGACAA AGGCAGAGAGAGGAGGACUGAGCGAA CUGGACAAGGCAGGAUUCAUAAGAG ACAGCUGGUCGAAACAAGACAGAUCA CAAAGCACGUCGCACAGAUCCUGGAC AGCAGAAUGAACACAAAGUACGACGA AAACGACAAGCUGAUCAGAGAAGUCA AGGUCAUACACACUGAAGAGCAAGCUG GUCAGCGACUUCAGAAAGGACUUCCA GUUCUACAAGGUCAGAGAAAUCAACA ACUACCACACGACACGACGCAUAC CUGAACGCAGUCGUCGGAACAGCACU GAUCAAGAAGUACCCGAAGCUGGAAA GCGAAUUCGUCUACGGAGACUACAAG GUCUACGACGUCAGAAAGAUGAUCGC AAAGAGCGAACAGGAAAUCGGAAAGG CAACAGCAAAGUACUUCUUCUACAGC AACAUCAUGAACUUCUUCAAGACAGA AAUCACACUGGCAAACGGAGAAAUCA GAAAGAGACCGCUGAUCGAAACAAAC GGAGAAACAGGAGAAAUCGUCUGGGA CAAGGGAAGAGACUUCGCAACAGUCA GAAAGGUCCUGAGCAUGCCGCAGGUC AACAUUCGUCAAGAAGACAGAAGUCCA GACAGGAGGAUUCAGCAAGGAAAGCA UCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGA CCCGAAGAAGUACGGAGGAUUCGACA GCCCGACAGUCGCAUACAGCGUCCUG GUCGUCGCAAAGGUUCGAAAAGGGAAA GAGCAAGAAGCUGAAGAGCGUCAAGG AACUGCUGGGAAUCACAAUCAUGGAA AGAAGCAGCUUCGAAAAGAACCCGAU CGACUUCUGGAAGCAAAGGGAUACA AGGAAGUCAAGAAGGACCUGAUCAUC AAGCUGCCGAAGUACAGCCUGUUCGA	
		10
		20
		30
		40

40

50

	CCUUUCUGUCCGGAGAACAGAAGAAA GCAAUUGUCGAUCUGCUGUUCAAGAC CAACCGCAAGGUGACCGUCAAGCAGC UUAAGAGGACUACUUCAAGAAGAUC GAGUGUUUCGACUCAGUGGAAAUCAG CGGGGUGGAGGACAGAUUCAACGCUU CGCUGGGAACCUAUC AUGAUCUCCUG AAGAUCAUCAAGGACAAGGACUUCCU UGACAACGAGGAGAACGAGGACAUCC UGGAAGAUAU CGUCCUGACCUUGACC CUUUUCGAGGAUCGCGAGAUGAUCGA GGAGAGGCUUAAGACCUACGCUCAUC UCUUCGACGAUAAGGUCAUGAAACAA CUCAAGCGCCGCGGUACACUGGUUG GGGCCGCCUCUCCCGCAAGCUGAUCA ACGGUAUUCGCGAUAAACAGAGCGGU AAAACUAUCCUGGAUUUCCUCAAAUC GGAUGGCUUCGCUAAUCGUAACUUCA UGCAAUUGAUCCACGACGACAGCCUG ACCUUUAAAGGAGGACAUCCAAAAAGC ACAAGUGUCCGGACAGGGAGACUCAC UCCAUGAACACAUCGCGAAUCUGGCC GGUUCGCCGCGGAUUAAGAAGGGAAU UCUGCAAACUGUGAAGGUGGUCGACG AGCUGGUGAAGGUCAUGGGACGGCAC AAACCGGAGAAUAUCGUGAUUGAAAU GGCCCGAGAAAACCAGACUACCCAGA AGGGCCAGAAAAACUCCCGCGAAAGG AUGAAGCGGAUCGAAGAAGGAAUCAA GGAGCUGGGCAGCCAGAUCCUGAAAG AGCACCCGGUGGAAAACACGCAGCUG CAGAACGAGAAGCUCUACCUGUACUA UUUGCAAAAUGGACGGGACAUGUACG UGGACCAAGAGCUGGACAUCAAU CGG UUGUCUGAUUACGACGUGGACCACAU CGUUCCACAGUCCUUUCUGAAGGAUG ACUCGAUCGAUAACAAGGUGUUGACU CGCAGCGACAAGAACAGAGGGAAGUC AGAUAAUGUGCCAUCGGAGGAGGUCG UGAAGAAGAUGAAGAAUUACUGGCGG CAGCUCCUGAAUGCGAAGCUGAUUAC CCAGAGAAAGUUUGACAAUCUCACUA	
		10
		20
		30
		40

	AAGCCGAGCGCGGGCGGACUCUCAGAG CUGGAUAAGGCUGGAUUCAAAACG GCAGCUGGUCGAGACUCGGCAGAUUA CCAAGCACGUGGCGCAGAUUCUUGGAC UCCCGCAUGAACACUAAAUACGACGA GAACGAUAAGCUCAUCCGGGAAGUGA AGGUGAUUACCCUGAAAAGCAAACUU GUGUCGGACUUUCGGAAGGACUUUCA GUUUUACAAAGUGAGAGAAAUCAACA ACUACCAUCACGCGCAUGACGCAUAC CUCAACGCUGUGGUCGGUACCGCCCU GAUCAAAAAGUACCCUAAACUUGAAU CGGAGUUUGUGUACGGAGACUACAAG GUCUACGACGUGAGGAAGAUGAUAGC CAAGUCCGAACAGGAAAUCGGGAAAG CAACUGCGAAAUACUUCUUUUACUCA AACAUCAUGAACUUUUUCAAGACUGA AAUUACGCUGGCCAAUGGAGAAAUCA GGAAGAGGCCACUGAUCGAAACUAAAC GGAGAAACGGGCGAAAUCGUGUGGGA CAAGGGCAGGGACUUCGCAACUGUUC GCAAAGUGCUCUCUAUGCCGCAAGUC AAUAUUGUGAAGAAAACCGAAGUGCA AACCGGCGGAUUUUCAAAGGAAUCGA UCCUCCCAAAGAGAAAUAGCGACAAG CUCAUUGCACGCAAGAAAGACUGGGA CCCGAAGAAGUACGGAGGAUUCGAUU CGCCGACUGUCGCAUACUCCGUCCUC GUGGUGGCCAAGGUGGAGAAGGGAAA GAGCAAAAAGCUCAAAUCCGUCAAAG AGCUGCUGGGGAUUACCAUCAUGGAA CGAUCCUCGUUCGAGAAGAACCCGAU UGAUUUCCUCGAGGCGAAGGGUUACA AGGAGGUGAAGAAGGAUCUGAUCAUC AAACUCCCCAAGUACUCACUGUUCGA ACUGGAAAAAUGGUCGGAAGCGCAUGC UGGCUUCGGCCGGAGAACUCCAAAAA GGAAAUGAGCUGGCCUUGCCUAGCAA GUACGUCAACUUCUUAUCUUGCUU CGCACUACGAAAAACUCAAGGGUCA CCGGAAGAUAACGAACAGAAGCAGCU UUUCGUGGAGCAGCACAAAGCAUUAUC	
		10
		20
		30
		40

	UGGAUGAAAUCAUCGAACAAAUCUCC GAGUUUUCAAAGCGCGUGAUCCUCGC CGACGCCAACCU CGACAAAGUCCUGU CGGCCUACAAUAAGCAUAGAGAUAAAG CCGAUCAGAGAACAGGCCGAGAACAU UAUCCACUUGUUCACCCUGACUAACC UGGGAGCCCCAGCCGCCUUCAAGUAC UUCGAUACUACUAUCGAUCGCAAAAG AUACACGUCCACCAAGGAAGUUCUGG ACGCGACCCUGAUCCACCAAAGCAUC ACUGGACUCUACGAAACUAGGAUCGA UCUGUCGCAGCUGGGUGGGCGAUGGCG GUGGAUCUCCGAAAAAGAAGAGAAAG GUGUAAUGA	
Cas9 ニッ カーゼ (D1 0A) アミノ 酸配列	MDKKYSIGLAIGTNSVGWAVITDEYK VPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFHRLEES FLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEEENPINASGVDA KAILSARLSKSRRENLI AQLPGEKK NGLFGNLI ALSLGLTPNFKSNFDLAE DAKLQLSKDITYDDDDLNDLLAQIGDQY ADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKAL VRQQLPKEYKEIFFDQSKNGYAGYID GGASQEEFYKFIPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHL GELHAILRRQEDFYFPFLKDNREKIEK ILTFRIPYYVGPLARGNSRFAWMTRK SEETITPWNFEEVVDKGASAQSFIER MTNFDKNLPNEKVL PKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKI ECFDSVEISGVEDRFNASLGTYHDLL KIIKDKDFLDNEENEDI LEDIVLTLT LFEDREMIEERLKYAHLFDDKVMKQ LKRRRYTGWGRLSRKLINGIRDKQSG KTI LDFLKS DGFANRNF MQLIHDDSL	6

	TFKEDIQKAQVSGQGDSLHEHIANLA GSPA I K K G I L Q T V K V V D E L V K V M G R H K P E N I V I E M A R E N Q T T Q K G Q K N S R E R M K R I E E G I K E L G S Q I L K E H P V E N T Q L Q N E K L Y L Y Y L Q N G R D M Y V D Q E L D I N R L S D Y D V D H I V P Q S F L K D D S I D N K V L T R S D K N R G K S D N V P S E E V V K K M K N Y W R Q L L N A K L I T Q R K F D N L T K A E R G G L S E L D K A G F I K R Q L V E T R Q I T K H V A Q I L D S R M N T K Y D E N D K L I R E V K V I T L K S K L V S D F R K D F Q F Y K V R E I N N Y H H A H D A Y L N A V V G T A L I K K Y P K L E S E F V Y G D Y K V Y D V R K M I A K S E Q E I G K A T A K Y F F Y S N I M N F F K T E I T L A N G E I R K R P L I E T N G E T G E I V W D K G R D F A T V R K V L S M P Q V N I V K K T E V Q T G G F S K E S I L P K R N S D K L I A R K K D W D P K K Y G G F D S P T V A Y S V L V V A K V E K G K S K K L K S V K E L L G I T I M E R S S F E K N P I D F L E A K G Y K E V K K D L I I K L P K Y S L F E L E N G R K R M L A S A G E L Q K G N E L A L P S K Y V N F L Y L A S H Y E K L K G S P E D N E Q K Q L F V E Q H K H Y L D E I I E Q I S E F S K R V I L A D A N L D K V L S A Y N K H R D K P I R E Q A E N I I H L F T L T N L G A P A A F K Y F D T T I D R K R Y T S T K E V L D A T L I H Q S I T G L Y E T R I D L S Q L G G D G G G S P K K K R K V	
C a s 9 ニッ カーゼ (D1 0 A) m R N A O R F	AUGGACAAGAAGUACAGCAUCGGACU GGCAAUCGGAACAAACAGCGUCGGAU GGGCAGUCAUCACAGACGAAUACAAG GUCCCGAGCAAGAAGUUCAAGGUCCU GGGAAACACAGACAGACACAGCAUCA AGAAGAACCUGAUCGGAGCACUGCUG UUCGACAGCGGAGAAACAGCAGAAGC AACAAGACUGAAGAGAACAGCAAGAA GAAGAUACACAAGAAGAAAGAACAGA AUCUGCUACCUGCAGGAAAUCUUCAG CAACGAAAUGGCAAAGGUCGACGACA GCUUCUUCACAGACUGGAAGAAAGC UUCUGGUCGAAGAAGACAAGAAGCA CGAAAGACACCCGAUCUUCGGAAACA	7

10

20

30

40

50

50

	CAAGCGCACAGAGCUUCAUCGAAAGA AUGACAAACUUCGACAAGAACCUGCC GAACGAAAAGGUCCUGCCGAAGCACA GCCUGCUGUACGAAUACUUCACAGUC UACAACGAACUGACAAAGGUCAAGUA CGUCACAGAAGGAAUGAGAAAGCCGG CAUUCCUGAGCGGAGAACAGAAGAAG GCAAUCGUCGACCUGCUGUUCAAGAC AAACAGAAAGGUACAGUCAAGCAGC UGAAGGAAGACUACUUCAAGAAGAUC GAAUGCUUCGACAGCGUCGAAAUCAG CGGAGUCGAAGACAGAUUCAACGCAA GCCUGGGAACAUAACACGACCUGCUG AAGAUCAUCAAGGACAAGGACUUCCU GGACAACGAAGAAAACGAAGACAUCC UGGAAGACAUCGUCCUGACACUGACA CUGUUCGAAGACAGAGAAAUGAUCGA AGAAAGACUGAAGACAUACGCACACC UGUUCGACGACAAGGUCAUGAAGCAG CUGAAGAGAAGAAGAUACACAGGAUG GGGAAGACUGAGCAGAAAGCUGAUCA ACGGAAUCAGAGACAAGCAGAGCGGA AAGACAAUCCUGGACUUCCUGAAGAG CGACGGAUUCGCAAACAGAAACUUCA UGCAGCUGAUCCACGACGACAGCCUG ACAUUCAAGGAAGACAUCCAGAAGGC ACAGGUCAGCGGACAGGGAGACAGCC UGCACGAACACAUCGCAAACCUGGCA GGAAGCCCGGCAAUCAAGAAGGGAAU CCUGCAGACAGUCAAGGUUCGUCGACG AACUGGUCAAGGUCAUGGGGAAGACAC AAGCCGGAAAACAUCGUCAUCGAAAU GGCAAGAGAAAACCAGACAACACAGA AGGGACAGAAGAACAGCAGAGAAAGA AUGAAGAGAAUUCGAAGAAGGAAUCAA GGAACUGGGAAGCCAGAUCCUGAAGG AACACCCGGUCGAAAACACACAGCUG CAGAACGAAAAGCUGUACCUGUACUA CCUGCAGAACGGAAGAGACAUGUACG UCGACCAGGAACUGGACAUCAACAGA CUGAGCGACUACGACGUCGACCACAU CGUCCCGCAGAGCUUCCUGAAGGACG	

10

20

30

40

50

50

	UGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAA GUACGUCAACUUCCUGUACCUGGCAA GCCACUACGAAAAGCUGAAGGGAAGC CCGGAAGACAACGAACAGAAGCAGCU GUUCGUCGAACAGCACAAAGCACUACC UGGACGAAAUCAUCGAACAGAUACAGC GAAUUCAGCAAGAGAGUCAUCCUGGC AGACGCAAACCUGGACAAGGUCCUGA GCGCAUACAACAAGCACAGAGACAAG CCGAUCAGAGAACAGGCAGAAAACA CAUCCACCUGUUCACACUGACAAACC UGGGAGCACCGGCAGCAUUCAAGUAC UUCGACACAACAUAUCGACAGAAAGAG AUACACAAGCACAAAGGAAGUCCUGG ACGCAACACUGAUCCACCAGAGCAUC ACAGGACUGUACGAAACAAGAAUCGA CCUGAGCCAGCUGGGAGGAGACGGAG GAGGAAGCCCGAAGAAGAAGAGAAAG GUCUAG	
d C a s 9 (D 1 0 A H 8 4 0 A) アミノ酸配列	MDKKYSIGLAIGTNSVGWAVITDEYK VPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFHRLEES FLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEEENPINASGVDA KAILSARLSKSRRLLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAE DAKLQLSKDITYDDDLNLLAQIGDQY ADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKAL VRQQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHL GELHAILRRQEDFYPPFLKDNREKIEK ILTFRIPYYVGPLARGNSRFAWMTRK SEETITPWNFEENVVDKGASAQSFIER MTNFDKNLPNEKVLPHKSLLEYEFTV YNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKI	8

10

20

30

40

50

40

	AUCUGCUACCUUGCAGGAAAUCUUCAG CAACGAAAUGGCAAAGGUCGACGACA GCUUCUUCACACAGACUGGAAGAAAGC UUCCUGGUCGAAGAAGACAAGAAGCA CGAAAGACACCCGAUCUUCGGAAACA UCGUCGACGAAGUCGCAUACCACGAA AAGUACCCGACAAUCUACCACCUGAG AAAGAAGCUGGUCGACAGCACAGACA AGGCAGACCUGAGACUGAUCUACCUG GCACUGGCACACAUGAUCAAGUUCAG AGGACACUUCUUGAUCGAAGGAGACC UGAACCCGGACAACAGCGACGUCGAC AAGCUGUUCAUCCAGCUGGUCCAGAC AUACAACCAGCUGUUCGAAGAAAACC CGAUCAACGCAAGCGGAGUCGACGCA AAGGCAAUCCUGAGCGCAAGACUGAG CAAGAGCAGAAGACUGGAAAACCUGA UCGCACAGCUGCCGGGAGAAAAGAAG AACGGACUGUUCGGAAACCUGAUCGC ACUGAGCCUGGGACUGACACCGAACU UCAAGAGCAACUUCGACCUGGCAGAA GACGCAAAGCUGCAGCUGAGCAAGGA CACAUACGACGACGACCUGGACAACC UGCUGGCACAGAUUCGGAGACCAGUAC GCAGACCUGUUCUUGGCAGCAAAGAA CCUGAGCGACGCAAUCCUGCUGAGCG ACAUCCUGAGAGUCAACACAGAAAUC ACAAAGGCACCGCUGAGCGCAAGCAU GAUCAAGAGAUACGACGAACACCACC AGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUA CAAGGAAAUCUUCUUCGACCAGAGCA AGAACGGAUACGCAGGAUACAUCGAC GGAGGAGCAAGCCAGGAAGAAUUCUA CAAGUUCAUCAAGCCGAUCCUGGAAA AGAUGGACGGAACAGAAGAACUGCUG GUCAAGCUGAACAGAGAAGACCUGCU GAGAAAGCAGAGAACAUUCGACAACG GAAGCAUCCCGCACCAAGAUCCACCUG GGAGAACUGCACGCAAUCCUGAGAAG ACAGGAAGACUUCUACCCGUUCCUGA AGGACAACAGAGAAAAGAUCGAAAAG	
		10
		20
		30
		40

50

	CAGAACGAAAAGCUGUACCUGUACUA CCUGCAGAACGGAAGAGACAUGUACG UCGACCAGGAACUGGACAUCAACAGA CUGAGCGACUACGACGUCGACGCAAU CGUCCCGCAGAGCUUCCUGAAGGACG ACAGCAUCGACAACAAGGUCCUGACA AGAAGCGACAAGAACAGAGGAAAGAG CGACAACGUCCCGAGCGAAGAAGUCG UCAAGAAGAUGAAGAACUACUGGAGA CAGCUGCUGAACGCAAAGCUGAUCAC ACAGAGAAAGUUCGACAACCUGACAA AGGCAGAGAGAGGAGGACUGAGCGAA CUGGACAAGGCAGGAUUCAUCAAGAG ACAGCUGGUCGAAACAAGACAGAUCA CAAAGCACGUCGCACAGAUCCUGGAC AGCAGAAUGAACACAAAGUACGACGA AAACGACAAGCUGAUCAGAGAAGUCA AGGUCAUCACACUGAAGAGCAAGCUG GUCAGCGACUUCAGAAAGGACUUCCA GUUCUACAAGGUCAGAGAAAUCAACA ACUACCACCACGCACACGACGCAUAC CUGAACGCAGUCGUCGGAACAGCACU GAUCAAGAAGUACCCGAAGCUGGAAA GCGAAUUCGUCUACGGAGACUACAAG GUCUACGACGUCAGAAAGAUGAUCGC AAAGAGCGAACAGGAAAUCGGAAAGG CAACAGCAAAGUACUUCUUCUACAGC AACAUCAUGAACUUCUUCAAGACAGA AAUCACACUGGCAAACGGAGAAAUCA GAAAGAGACCGCUGAUCGAAACAAAC GGAGAAACAGGAGAAAUCGUCUGGGA CAAGGGAAGAGACUUCGCAACAGUCA GAAAGGUCCUGAGCAUGCCGCAGGUC AACAUCGUCAAGAAGACAGAAGUCCA GACAGGAGGAUUCAGCAAGGAAAGCA UCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGA CCCGAAGAAGUACGGAGGAUUCGACA GCCCGACAGUCGCAUACAGCGUCCUG GUCGUCGCAAAGGUCGAAAAGGGAAA GAGCAAGAAGCUGAAGAGCGUCAAGG AACUGCUGGGAAUCACAAUCAUGGAA	
		10
		20
		30
		40

	AGAAGCAGCUUCGAAAAGAACCCGAU CGACUUCCUGGAAGCAAAGGGAUACA AGGAAGUCAAGAAGGACCUGAUCauc AAGCUGCCGAAGUACAGCCUGUUCGA ACUGGAAAACGGAAGAAAGAGAAUGC UGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAA GUACGUCAACUUCCUGUACCUGGCAA GCCACUACGAAAAGCUGAAGGGGAAGC CCGGAAGACAACGAACAGAAGCAGCU GUUCGUUGAACAGCACAAAGCACUACC UGGACGAAAUCAUCGAACAGAUcAGC GAAUUCAGCAAGAGAGUCAUCCUGGC AGACGCAAACCUGGACAAGGUCCUGA GCGCAUACAACAAGCACAGAGACAAG CCGAUCAGAGAACAGGCAGAAAACAu CAUCCACCUGUUCACACUGACAAACC UGGGAGCACCCGGCAGCAUUCAAGUAC UUCGACACAACAUAUCGACAGAAAGAG AUACACAAGCACAAAGGAAGUCCUGG ACGCAACACUGAUCCACCAGAGCAUC ACAGGACUGUACGAAACAAGAAUCGA CCUGAGCCAGCUGGGAGGAGACGGAG GAGGAAGCCCGAAGAAGAAGAGAAAG GUCUAG	
C a s 9 の裸 のコード配列	GACAAGAAGUACAGCAUCGGACUGGA CAUCGGAACAAACAGCGUCGGAUGGG CAGUCAUCACAGACGAAUACAAGGUC CCGAGCAAGAAGUUCAAGGUCCUGGG AAACACAGACAGACACAGCAUCAAGA AGAACCUGAUUCGGAGCACUGCUGUUC GACAGCGGAGAAACAGCAGAAGCAAC AAGACUGAAGAGAACAGCAAGAAGAA GAUACACAAGAAGAAAGAACAGAAUC UGCUACCUGCAGGAAAUCUUCAGCAA CGAAAUGGCAAAGGUcGACGACAGCU UCUUCCACAGACUGGAAGAAAGCUUC CUGGUcGAAGAAGACAAGAAGCACGA AAGACACCCGAUCUUCGGAACAUCG UCGACGAAGUCGCAUACCACGAAAAG UACCCGACAAUCUACCACCUGAGAAA GAAGCUGGUcGACAGCACAGACAAGG	1 0

10

20

30

40

50

	CAGACCUGAGACUGAUCUACCUGGCA CUGGCACACAUGAUCAAGUUCAGAGG ACACUUCCUGAUCGAAGGAGACCUGA ACCCGGACAACAGCGACGUCGACAAG CUGUUCAUCCAGCUGGUCCAGACAU CAACCAGCUGUUCGAAGAAAACCCGA UCAACGCAAGCGGAGUCGACGCAAAG GCAAUCCUGAGCGCAAGACUGAGCAA GAGCAGAAGACUGGAAAACCUGAUCG CACAGCUGCCGGGAGAAAAGAAGAAC GGACUGUUCGGAAACCUGAUCGCACU GAGCCUGGGACUGACACCGAACUUCA AGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACAC AUACGACGACGACCUGGACAACCUGC UGGCACAGAUUCGGAGACCAGUACGCA GACCUGUUCUGGCAGCAAAGAACCU GAGCGACGCAAUCCUGCUGAGCGACA UCCUGAGAGUCAACACAGAAAUCACA AAGGCACCGCUGAGCGCAAGCAUGAU CAAGAGAUACGACGAACACCACCAGG ACCUGACACUGCUGAAGGCACUGGUC AGACAGCAGCUGCCGGAAAAGUACAA GGAAAUCUUCUUCGACCAGAGCAAGA ACGGAUACGCAGGAUACAUCGACGGA GGAGCAAGCCAGGAAGAAUUCUACAA GUUCAUCAAGCCGAUCCUGGAAAAGA UGGACGGAACAGAAGAACUGCUGGUC AAGCUGAACAGAGAAGACCUGCUGAG AAAGCAGAGAACAUUCGACAACGGAA GCAUCCCGCACCCAGAUCCACCUGGGA GAACUGCACGCAAUCCUGAGAAGACA GGAAGACUUCUACCCGUUCCUGAAGG ACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUUCAGAAUCCCGUACUACGU CGGACCGCUGGCAAGAGGAAACAGCA GAUUCGCAUGGAUGACAAGAAAGAGC GAAGAAACAAUACACCCUGGAACUU CGAAGAAGUCGUCGACAAGGGAGCAA GCGCACAGAGCUUCAUCGAAAGAAUG ACAAACUUCGACAAGAACCUGCCGAA CGAAAAGGUCCUGCCGAAGCACAGCC	
		10
		20
		30
		40

	UGCUGUACGAAUACUUCACAGUCUAC AACGAACUGACAAAGGUCAAGUACGU CACAGAAGGAAUGAGAAAGCCGGCAU UCCUGAGCGGAGAACAGAAGAAGGCA AUCGUCGACCUGCUGUUCAAGACAAA CAGAAAGGUCACAGUCAAGCAGCUGA AGGAAGACUACUUCAAGAAGAUCGAA UGC UUCGACAGCGUCGAAAUCAGCGG AGUCGAAGACAGAUUCAACGCAAGCC UGGGAACAUAACACGACCUGCUGAAG AUCAUCAAGGACAAGGACUUCUGGA CAACGAAGAAAACGAAGACAUCCUGG AAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAUCGAAGA AAGACUGAAGACAUAACGCACACCUGU UCGACGACAAGGUCAUGAAGCAGCUG AAGAGAAGAAGAUACACAGGAUGGGG AAGACUGAGCAGAAAGCUGAUCAACG GAAUCAGAGACAAGCAGAGCGGAAAG ACAAUCCUGGACUUCUGAAGAGCGA CGGAUUCGCAAACAGAAACUUCAUGC AGCUGAUCCACGACGACAGCCUGACA UUCAAGGAAGACAUCCAGAAGGCACA GGUCAGCGGACAGGGAGACAGCCUGC ACGAACACAUCGCAAACCUGGCAGGA AGCCCGGCAAUCAAGAAGGGAAUCCU GCAGACAGUCAAGGUUCGUCGACGAAC UGGUCAAGGUCAUGGGAAGACACAAG CCGGAAAACAUCGUCAUCGAAAUGGC AAGAGAAAACCAGACAACACAGAAGG GACAGAAGAACAGCAGAGAAAGAAUG AAGAGAAUCGAAGAAGGAAUCAAGGA ACUGGGAAGCCAGAUCCUGAAGGAAC ACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCU GCAGAACGGAAGAGACAUGUACGUCG ACCAGGAACUGGACAUCAACAGACUG AGCGACUACGACGUCGACCACAUCGU CCCGCAGAGCUUCCUGAAGGACGACA GCAUCGACAACAAGGUCCUGACAAGA AGCGACAAGAACAGAGGAAAGAGCGA CAACGUCCCGAGCGAAGAAGUCGUCA	

10

20

30

40

50

50

	ACUACGAAAAGCUGAAGGGAAGCCCG GAAGACAACGAACAGAAGCAGCUGUU CGUCGAACAGCACAAAGCACUACCUGG ACGAAAUCAUCGAACAGAUAGCGAA UUCAGCAAGAGAGUCAUCCUGGCAGA CGCAAACCUGGACAAGGUCCUGAGCG CAUACAACAAGCACAGAGACAAGCCG AUCAGAGAACAGGCAGAAAACAUCAU CCACCUGUUCACACUGACAAACCUGG GAGCACCGGCAGCAUUCAAGUACUUC GACACAACAUAUCGACAGAAAGAGUA CACAAGCACAAAGGAAGUCCUGGACG CAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGACCU GAGCCAGCUGGGAGGAGACGGAGGAG GAAGCCCCGAAGAAGAAGAGAAAGGUC	
C a s 9 ニッ カーゼの裸の コード配列	GACAAGAAGUACAGCAUCGGACUGGC AAUCGGAACAAACAGCGUCGGAUGGG CAGUCAUCACAGACGAAUACAAGGUC CCGAGCAAGAAGUUCAAGGUCCUGGG AAACACAGACAGACACAGCAUCAAGA AGAACCUGAUUCGGAGCACUGCUGUUC GACAGCGGAGAAACAGCAGAAGCAAC AAGACUGAAGAGAAACAGCAAGAAGAA GAUACACAAGAAGAAAGAACAGAAUC UGC UACCUGCAGGAAAUCUUCAGCAA CGAAAUGGCAAAGGUUCGACGACAGCU UCUUCCACAGACUGGAAGAAAGCUUC CUGGUUCGAAGAAGACAAGAAGCACGA AAGACACCCGAUCUUCGGAAACAUCG UCGACGAAGUCGCAUACCACGAAAAG UACCCGACAAUCUACCACCUGAGAAA GAAGCUGGUUCGACAGCACAGACAAGG CAGACCUGAGACUGAUUCUACCUGGCA CUGGCACACAUGAUCAAGUUCAGAGG ACACUUCUGAUUCGAAGGAGACCUGA ACCCGGACAACAGCGACGUCGACAAG CUGUUCAUCCAGCUGGUCCAGACAU CAACCAGCUGUUCGAAGAAAACCCGA UCAACGCAAGCGGAGUCGACGCAAAG GCAAUCCUGAGCGCAAGACUGAGCAA	1 1

10

20

30

40

50

	GAGCAGAAGACUGGAAAACCUGAUCG CACAGCUGCCGGGAGAAAAGAAGAAC GGACUGUUCGGAAACCUGAUCGCACU GAGCCUGGGACUGACACCGAACUUCA AGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACAC AUACGACGACGACCUGGACAACCUGC UGGCACAGAUUCGGAGACCAGUACGCA GACCUGUUCUUGGCAGCAAAGAACCU GAGCGACGCAAUCCUGCUGAGCGACA UCCUGAGAGUCAACACAGAAAUCACA AAGGCACCGCUGAGCGCAAGCAUGAU CAAGAGAUACGACGAACACCACCAGG ACCUGACACUGCUGAAGGCACUGGUC AGACAGCAGCUGCCGGAAAAGUACAA GGAAAUCUUCUUCGACCAGAGCAAGA ACGGAUACGCAGGAUACAUCGACGGA GGAGCAAGCCAGGAAGAAUUCUACAA GUUCAUCAAGCCGAUCCUGGAAAAGA UGGACGGAACAGAAGAACUGCUGGUC AAGCUGAACAGAGAAGACCUGCUGAG AAAGCAGAGAACAUAUCGACAACGGAA GCAUCCCGCACCCAGAUCCACCUGGGA GAACUGCACGCAAUCCUGAGAAGACA GGAAGACUUCUACCCGUUCCUGAAGG ACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUAUCAGAAUCCCGUACUACGU CGGACCGCUGGCAAGAGGAAACAGCA GAUUCGCAUGGAUGACAAGAAAGAGC GAAGAAACAAUCACACCGUGGAACUU CGAAGAAGUCGUCGACAAGGGAGCAA GCGCACAGAGCUUCAUCGAAAGAAUG ACAAACUUCGACAAGAACCUGCCGAA CGAAAAGGUCCUGCCGAAGCACAGCC UGCUGUACGAAUACUUCACAGUCUAC AACGAACUGACAAAGGUCAAGUACGU CACAGAAGGAAUGAGAAAGCCGGCAU UCCUGAGCGGAGAACAGAAGAAGGCA AUCGUCGACCUGCUGUUCAAGACAAA CAGAAAGGUCACAGUCAAGCAGCUGA AGGAAGACUACUUCAAGAAGAUCGAA UGCUUCGACAGCGUCGAAAUCAGCGG	
		10
		20
		30
		40

50

	CGACAAGCUGAUCAGAGAAGUCAAGG UCAUCACACUGAAGAGCAAGCUGGUC AGCGACUUCAGAAAGGACUUCCAGUU CUACAAGGUCAGAGAAAUCAACAACU ACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAU CAAGAAGUACCCGAAGCUGGAAAGCG AAUUCGUCUACGGAGACUACAAGGUC UACGACGUCAGAAAGAUGAUCGCAA GAGCGAACAGGAAAUCGGAAAGGCAA CAGCAAAGUACUUCUUCUACAGCAAC AUCAUGAACUUCUUCAAGACAGAAAU CACACUGGCAAACGGAGAAAUCAGAA AGAGACCGCUGAUCGAAACAAACGGA GAAACAGGAGAAAUCGUCUGGGACAA GGGAAGAGACUUCGCAACAGUCAGAA AGGUCCUGAGCAUGCCGCAGGUCAAC AUCGUCAAGAAGACAGAAGUCCAGAC AGGAGGAUUCAGCAAGGAAAGCAUCC UGCCGAAGAGAAAACAGCGACAAGCUG AUCGCAAGAAAGAAGGACUGGGACCC GAAGAAGUACGGAGGAUUCGACAGCC CGACAGUCGCAUACAGCGUCCUGGUC GUCGCAAAGGUCGAAAAGGGAAAGAG CAAGAAGCUGAAGAGCGUCAAGGAAC UGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGA CUUCCUGGAAGCAAAGGGAUACAAGG AAGUCAAGAAGGACCUGAUCAUCAAG CUGCCGAAGUACAGCCUGUUCGAACU GGAAAACGGAAGAAAGAGAAUGCUGG CAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUA CGUCAACUUCUGUACCUGGCAAGCC ACUACGAAAAGCUGAAGGGAAGCCCG GAAGACAACGAACAGAAGCAGCUGUU CGUCGAACAGCACAAAGCACUACCUGG ACGAAAUCAUCGAACAGAUCAGCGAA UUCAGCAAGAGAGUCAUCCUGGCAGA CGCAAACCUGGACAAGGUCCUGAGCG CAUACAACAAGCACAGAGACAAGCCG AUCAGAGAACAGGCAGAAAACAUCAU	

10

20

30

40

50

	CCACCUGUUCACACUGACAAACCUGG GAGCACCGGCAGCAUUCAAGUACUUC GACACAACAAUCGACAGAAAGAGAU CACAAGCACAAAGGAAGUCCUGGACG CAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGACCU GAGCCAGCUGGGAGGAGACGGAGGAG GAAGCCCCGAAGAAGAAGAGAAAGGUC	
d C a s 9 の 裸のコード配 列	GACAAGAAGUACAGCAUCGGACUGGC AAUCGGAACAAACAGCGUCGGAUGGG CAGUCAUCACAGACGAAUACAAGGUC CCGAGCAAGAAGUUCAAGGUCCUGGG AAACACAGACAGACACAGCAUCAAGA AGAACCUGAUCCGGAGCACUGCUGUUC GACAGCGGAGAAACAGCAGAAGCAAC AAGACUGAAGAGAACAGCAAGAAGAA GAUACACAAGAAGAAAGAACAGAAUC UGC UACCUGCAGGAAAUCUUCAGCAA CGAAAUGGCAAAGGUCGACGACAGCU UCUUCACAGACUGGAAGAAAGCUUC CUGGUCGAAGAAGACAAGAAGCACGA AAGACACCCGAUCUUCGGAAACAUCG UCGACGAAGUCGCAUACCACGAAAAG UACCCGACAAUCUACCACCUGAGAAA GAAGCUGGUCGACAGCACAGACAAGG CAGACCUGAGACUGAUUCUACCUGGCA CUGGCACACAUGAUCAAGUUCAGAGG ACACUUCUGAUUCGAAGGAGACCUGA ACCCGGACAAACAGCGACGUCGACAAG CUGUUCAUCCAGCUGGUCCAGACAU CAACCAGCUGUUCGAAGAAAACCCGA UCAACGCAAGCGGAGUCGACGCAAAG GCAAUCCUGAGCGCAAGACUGAGCAA GAGCAGAAGACUGGAAAACCUGAUCG CACAGCUGCCGGGAGAAAAGAAGAAC GGACUGUUCGGAACCUGAUCGCACU GAGCCUGGGACUGACACCGAACUUCA AGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACAC AUACGACGACGACCUGGACAACCUGC UGGCACAGAUCCGGAGACCAGUACGCA GACCUGUUCUGGCAGCAAAGAACCU	1 2

10

20

30

40

50

	GAGCGACGCAAUCCUGCUGAGCGACA UCCUGAGAGUCAACACAGAAAUCACA AAGGCACCGCUGAGCGCAAGCAUGAU CAAGAGAUACGACGAACACCACCAGG ACCUGACACUGCUGAAGGCACUGGUC AGACAGCAGCUGCCGGAAAAGUACAA GGAAAUCUUCUUCGACCAGAGCAAGA ACGGAUACGCAGGAUACAUCGACGGA GGAGCAAGCCAGGAAGAAUUCUACAA GUUCAUCAAGCCGAUCCUGGAAAAGA UGGACGGAACAGAAAGAACUGCUGGUC AAGCUGAACAGAGAAGACCUGCUGAG AAAGCAGAGAACAUCGACAACGGAA GCAUCCCGCACCCAGAUCCACCUGGGA GAACUGCACGCAAUCCUGAGAAGACA GGAAGACUUCUACCCGUUCCUGAAGG ACAACAGAGAAAAGAUCGAAAAGAU CUGACAUCAGAAUCCCGUACUACGU CGGACCGCUGGCAAGAGGAAACAGCA GAUUCGCAUGGAUGACAAGAAAGAGC GAAGAAACAUCACACCGUGGAACUU CGAAGAAGUCGUCGACAAGGGAGCAA GCGCACAGAGCUUCAUCGAAAGAAUG ACAAACUUCGACAAGAACCUGCCGAA CGAAAAGGUCCUGCCGAAGCACAGCC UGCUGUACGAAUACUUCACAGUCUAC AACGAACUGACAAAGGUCAAGUACGU CACAGAAGGAAUGAGAAAGCCGGCAU UCCUGAGCGGAGAACAGAAAGAAGGCA AUCGUUCGACCUGCUGUUCAAGACAAA CAGAAAGGUCCAGUCAAGCAGCUGA AGGAAGACUACUUCAAGAAGAUCGAA UGCUCGACAGCGUCGAAAUCAGCGG AGUCGAAGACAGAUUCAACGCAAGCC UGGGAACAUAACACGACCUGCUGAAG AUCAUCAAGGACAAGGACUUCCUGGA CAACGAAGAAAACGAAGACAUCCUGG AAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAUCGAAGA AAGACUGAAGACAUAACGCACACCUGU UCGACGACAAGGUCAUGAAGCAGCUG AAGAGAAGAAGAUACACAGGAUGGGG	
		10
		20
		30
		40

	AAGACUGAGCAGAAAGCUGAUC AACG GAAUCAGAGACAAGCAGAGCGGAAAG ACAAUCCUGGACU UCCUGAAGAGCGA CGGAUUCGCAAACAGAAACUUCAUGC AGCUGAUCCACGACGACAGCCUGACA UUCAAGGAAGACA UCCAGAAGGCACA GGUCAGCGGACAGGGAGACAGCCUGC ACGAACACAUCGCAAACCUGGCAGGA AGCCC GGCAAUCAAGAAGGGAAUCCU GCAGACAGUCAAGGUCGUCGACGAAC UGGUCAAGGUCAUGGGAAGACACAAG CCGGAAAACAUCGUCAUCGAAAUGGC AAGAGAAAACCAGACAACACAGAAGG GACAGAAGAACAGCAGAGAAAGAAUG AAGAGAAUCGAAGAAGGAAUCAAGGA ACUGGGAAGCCAGA UCCUGAAGGAAC ACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCU GCAGAACGGAAGAGACAUGUACGUCG ACCAGGAACUGGACAUCAACAGACUG AGCGACUACGACGUCGACGCAAUCGU CCCGCAGAGCUUCCUGAAGGACGACA GCAUCGACAACAAGGUCCUGACAAGA AGCGACAAGAACAGAGGAAAGAGCGA CAACGUCCCGAGCGAAGAAGUCGUCA AGAAGAUGAAGAACUACUGGAGACAG CUGCUGAACGCAAAGCUGAUCACACA GAGAAAGUUCGACAACCUGACAAAGG CAGAGAGAGGAGGACUGAGCGAACUG GACAAGGCAGGAUUCAUCAAGAGACA GCUGGUUCGAAACAAGACAGAUCACAA AGCACGUCGCACAGA UCCUGGACAGC AGAAUGAACACAAAGUACGACGAAAA CGACAAGCUGAUCAGAGAAGUCAAGG UCAUCACACUGAAGAGCAAGCUGGUC AGCGACUUCAGAAAGGACU UCCAGUU CUACAAGGUCAGAGAAAUCAACAACU ACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAU CAAGAAGUACCCGAAGCUGGAAAGCG AAUUCGUCUACGGAGACUACAAGGUC UACGACGUCAGAAAGAUGAUCGCAA	

10

20

30

40

50

	GAGCGAACAGGAAAUCGGAAAGGCAA CAGCAAAGUACUUCUUCUACAGCAAC AUCAUGAACUUCUUCAAGACAGAAAU CACACUGGCAAACGGAGAAAUCAGAA AGAGACCGCUGAUUCGAAACAAACGGA GAAACAGGAGAAAUCGUCUGGGACAA GGGAAGAGACUUCGCAACAGUCAGAA AGGUCCUGAGCAUGCCGCAGGUCAAC AUCGUCAAGAAGACAGAAGUCCAGAC AGGAGGAUUCAGCAAGGAAAGCAUCC UGCCGAAGAGAAAACAGCGACAAGCUG AUCGCAAGAAAGAAGGACUGGGACCC GAAGAAGUACGGAGGAUUCGACAGCC CGACAGUCGCAUACAGCGUCCUGGUC GUCGCAAAGGUCGAAAAGGGAAAGAG CAAGAAGCUGAAGAGCGUCAAGGAAC UGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGA CUUCCUGGAAGCAAAGGGAUACAAGG AAGUCAAGAAGGACCUGAUCAUCAAG CUGCCGAAGUACAGCCUGUUCGAACU GGAAAACGGAAGAAAGAGAAUGCUGG CAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUA CGUCAACUUCUGUACCUGGCAAGCC ACUACGAAAAGCUGAAGGGGAAGCCCG GAAGACAACGAACAGAAGCAGCUGUU CGUCGAACAGCACAAAGCACUACCUGG ACGAAAUCAUCGAACAGAUCAAGCGAA UUCAGCAAGAGAGUCAUCCUGGCAGA CGCAAACCUGGACAAGGUCCUGAGCG CAUACAACAAGCACAGAGACAAGCCG AUCAGAGAACAGGCAGAAAACAUCAU CCACCUGUUCACACUGACAAACCUGG GAGCACCGGCAGCAUUCAAGUACUUC GACACAACAUAUCGACAGAAAGAGUA CACAAGCACAAAGGAAGUCCUGGACG CAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGACCU GAGCCAGCUGGGAGGAGACGGAGGAG GAAGCCCGAAGAAGAAGAGAAAGGUC	
C a s 9 のア	MDKKYSIGLDIGTNSVGWAVITDEYK	1 3

10

20

30

40

50

ミノ酸配列 (NLSなし)	VPSKKFKVLGNTDRHSIKKNLIGALL	
	FDSGETAEATRLKRTARRRYTRRKNR	
	ICYLQEIFSNEMAKVDDSFHRL EES	
	FLVEEDKKHERHPIFGNIVDEVAYHE	
	KYPTIYHLRKKLVDSTDKADLRRLIYL	
	ALAHMIKFRGHFLIEGDLNPDNSDVD	
	KLFIQLVQTYNQLFEE NPINASGVDA	
	KAILSARLSKSRRL ENLIAQLPG EKK	
	NGLFGNLIALSLGLTPNFKSNFDLAE	10
	DAKLQLSKD TYDDDDL DNLLAQIGDQY	
	ADLFLAAKNLSDAILLSDILRVNTEI	
	TKAPLSASMIKRYDEHHQDLTLLKAL	
	VRQQLP EKYKEIFFDQSKNGYAGYID	
	GGASQEEFYKF IKPILEKMDGTEELL	
	VKLNREDLLRKQRTFDNGSIPHQIHL	
	GELHA ILRRQEDFY PFLKDNREKIEK	
	ILTFRI PYYVG PLARGNSRFAWMTRK	
	SEETITPWNFEEVVDKGASAQS FIER	
	MTNFDKNLPNEKVL PKHSLLYEYFTV	
	YNELTKVKYVTEGMRKPAFLSGEQKK	20
	AIVDLLFKTNRKVTVKQLKEDYFKKI	
	ECFDSVEISGVEDRFNASLGTYHDLL	
	KIIKDKDFLDNEENEDI LEDIVLTLT	
	LFEDREMIEERLKTYAHLFDDKVMKQ	
	LKRRRYTGWGRLSRK LINGIRDKQSG	
	KTILDFLKSDGFANRNFMQLIHDDSL	
	TFKEDIQKAQVSGQGDSLHEHIANLA	
	GSPA IKKGILQTVKVVD ELVKVMGRH	
	KPENIVIEMARENQTTQKGQKNSRER	
	MKRIEEGIKELGSQILKEHPVENTQL	
	QNEKLYLYYLQNGRDMYVDQELDINR	30
	LSDYDVDHIVPQSFLKDDSIDNKVLT	
	RSDKNRGKSDNVPSEEVVKMKMKNYWR	
	QLLNAKLITQRKF DNLTKAERGGLSE	
	LDKAGFIKRQLVETRQITKHVAQILD	
	SRMNTKYDENDKLIREVKVITLKSKL	
	VSDFRKDFQFYKVREINNYHHAHDAY	
	LNAVVGTA LIKKY PKLESEFVYGDYK	
	VYDVRKMI AKSEQEIGKATAKYFFYS	
	NIMNFFKTEITLANGEIRKRPLIETN	
	GETGEIVWDKGRDFATVRKVLSMPQV	
	NIVKKTEVQTGGFSKESILPKRNSDK	40

	LIARAKKWDPKKYGGFDSPTVAYSVL VVAKVEKKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLII KLPKYSLFELENGKRKRLASAGELQK GNELALPSKYVNFLLYLA SHYEKLKGS PEDNEQKQLFVEQHKHYLDEII EQIS EFSKRVI LADANLDKVLSAYNKH RDK PIREQAENIIHLFTLTNLGAPAAFKY FDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGD	
開始コドン及 び終止コドン を有し、表3 に記載の最小 ウリジンコドン を使用する、 配列番号 13をコード するCas9 mRNA ORF	AUGGACAAGAAGUACAGCAUCGGACU GGACAUCGGAACAAACAGCGUCGGAU GGGCAGUCAUCACAGACGAAUACAAG GUCCCGAGCAAGAAGUUCAAGGUCCU GGGAAACACAGACAGACACAGCAUCA AGAAGAACCUGAUCGGAGCACUGCUG UUCGACAGCGGAGAAACAGCAGAAGC AACAAAGACUGAAGAGAACAGCAAGAA GAAGAUACACAAGAAGAAAGAACAGA AUCUGCUACCUGCAGGAAAUCUUCAG CAACGAAAUGGCAAAGGUCGACGACA GCUUCUUCACACAGACUGGAAGAAAGC UUCCUGGUCGAAGAAGACAAGAAGCA CGAAAGACACCCGAUCUUCGGAAACA UCGUCGACGAAGUCGCAUACCACGAA AAGUACCCGACAAUCUACCACCUGAG AAAGAAGCUGGUCGACAGCACAGACA AGGCAGACCUGAGACUGAUCUACCUG GCACUGGCACACAUGAUC AAGUUCAG AGGACACUUCUGAUCGAAGGAGACC UGAACCCGGACAACAGCGACGUCGAC AAGCUGUUCAUCCAGCUGGUCCAGAC AUACAACCAGCUGUUCGAAGAAAACC CGAUC AACGCAAGCGGAGUCGACGCA AAGGCAAUCCUGAGCGCAAGACUGAG CAAGAGCAGAAGACUGGAAAACCUGA UCGCACAGCUGCCGGGAGAAAAGAAAG AACGGACUGUUCGGAAACCUGAUCGC ACUGAGCCUGGGACUGACACCGAACU UCAAGAGCAACUUCGACCUGGCAGAA GACGCAAAGCUGCAGCUGAGCAAGGA	14

50

AGAAAGACUGAAGACAUACGCACACC UGUUCGACGACAAGGUCAUGAAGCAG CUGAAGAGAAGAAGAUACACAGGAUG GGGAAGACUGAGCAGAAAGCUGAUCA ACGGAAUCAGAGACAAGCAGAGCGGA AAGACAAUCCUGGACUUCCUGAAGAG CGACGGAUUCGCAAACAGAAACUUCA UGCAGCUGAUCCACGACGACAGCCUG ACAUUCAAGGAAGACAUCCAGAAGGC ACAGGUCAGCGGACAGGGAGACAGCC UGCACGAACACAUCGCAAACCUGGCA GGAAGCCCGGCAAUCAAGAAGGGAAU CCUGCAGACAGUCAAGGUUCGUCGACG AACUGGUCAAGGUCAUGGGAAGACAC AAGCCGGAAAACAUCGUCAUCGAAAU GGCAAGAGAAAACCAGACAACACAGA AGGGACAGAAGAACAGCAGAGAAAGA AUGAAGAGAAUCGAAGAAGGAAUCAA GGAACUGGGAAGCCAGAUCCUGAAGG AACACCCGGUCGAAAACACACAGCUG CAGAACGAAAAGCUGUACCUGUACUA CCUGCAGAACGGAAGAGACAUGUACG UCGACCAGGAACUGGACAUCAACAGA CUGAGCGACUACGACGUCGACCACAU CGUCCCGCAGAGCUUCCUGAAGGACG ACAGCAUCGACAACAAGGUCCUGACA AGAAGCGACAAGAACAGAGGAAAGAG CGACAACGUCCCGAGCGAAGAAGUCG UCAAGAAGAUGAAGAACUACUGGAGA CAGCUGCUGAACGCAAAGCUGAUCAC ACAGAGAAAGUUCGACAACCUGACAA AGGCAGAGAGAGGAGGACUGAGCGAA CUGGACAAGGCAGGAUUCAUCAAGAG ACAGCUGGUUCGAAACAAGACAGAUCA CAAAGCACGUCGCACAGAUCCUGGAC AGCAGAAUGAACACAAAGUACGACGA AAACGACAAGCUGAUCAGAGAAGUCA AGGUCAUCACACUGAAGAGCAAGCUG GUCAGCGACUUCAGAAAGGACUUCCA GUUCUACAAGGUCAGAGAAAUCAACA ACUACCACCACGCACACGACGCAUAC CUGAACGCAGUCGUCGGAACAGCACU	

10

20

30

40

50

	GAUCAAGAAGUACCCGAAGCUGGAAA GCGAAUUCGUCUACGGAGACUACAAG GUCUACGACGUCAGAAAGAUGAUCGC AAAGAGCGAACAGGAAAUCGGAAAGG CAACAGCAAAGUACUUCUUCUACAGC AACAUCAUGAACUUCUUCAAGACAGA AAUCACACUGGCCAAACGGAGAAAUCA GAAAGAGACCGCUGAUCGAAACAAAC GGAGAAACAGGAGAAAUCGUCUGGGA CAAGGGAAGAGACUUCGCAACAGUCA GAAAGGUCCUGAGCAUGCCGCAGGUC AACAUUCGUCAAGAAGACAGAAGUCCA GACAGGAGGAUUCAGCAAGGAAAGCA UCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGA CCCGAAGAAGUACGGAGGAUUCGACA GCCCGACAGUCGCAUACAGCGUCCUG GUCGUCGCAAAGGUCGAAAAGGGAAA GAGCAAGAAGCUGAAGAGCGUCAAGG AACUGCUGGGAAUCACAAUCAUGGAA AGAAGCAGCUUCGAAAAGAACCCGAU CGACUUCCUGGAAGCAAAGGGAUACA AGGAAGUCAAGAAGGACCUGAUCAUC AAGCUGCCGAAGUACAGCCUGUUCGA ACUGGAAAACGGAAGAAAGAGAAUGC UGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAA GUACGUCAACUUCUGUACCUGGCAA GCCACUACGAAAAGCUGAAGGGAAGC CCGGAAGACAACGAACAGAAGCAGCU GUUCGUCGAACAGCACAAAGCACUACC UGGACGAAAUCAUCGAACAGAUCAAGC GAAUUCAGCAAGAGAGUCAUCCUGGC AGACGCAAACCUGGACAAGGUCCUGA GCGCAUACAACAAGCACAGAGACAAG CCGAUCAGAGAACAGGCAGAAAACAU CAUCCACCUGUUCACACUGACAAACC UGGGAGCACCGGCAGCAUUCAAGUAC UUCGACACAACAUCGACAGAAAGAG AUACACAAGCACAAAGGAAGUCCUGG ACGCAACACUGAUCCACCAGAGCAUC ACAGGACUGUACGAAACAAGAAUCGA	
		10
		20
		30
		40

	CCUGAGCCAGCUGGGAGGAGACUAG	
表3に記載の 最小ウリジン コドンを使用 する、配列番 号13をコー ドするCas 9コード配列 (開始コドン または終止コ ドンはない； 融合タンパク 質コード配列 に含める際に 好適)	GACAAGAAGUACAGCAUCGGACUGGA CAUCGGAACAAACAGCGUCGGAUGGG CAGUCAUCACAGACGAAUACAAGGUC CCGAGCAAGAAGUUCAAGGUCCUGGG AAACACAGACAGACACAGCAUCAAGA AGAACCUGAUCGGAGCACUGCUGUUC GACAGCGGAGAAACAGCAGAAGCAAC AAGACUGAAGAGAAACAGCAAGAAGAA GAUACACAAGAAGAAAGAACAGAAUC UGC UACCUGCAGGAAAUCUUCAGCAA CGAAAUGGCAAAGGUUCGACGACAGCU UCU UCCACAGACUGGAAGAAAGCUUC CUGGU CGAAGAAGACAAGAAGCACGA AAGACACCCGAUCUUCGGAAACAUCG UCGACGAAGUCGCAUACCACGAAAAG UACCCGACAAUCUACCACCUGAGAAA GAAGCUGGUCGACAGCACAGACAAGG CAGACCUGAGACUGAUCUACCUGGCA CUGGCACACAUGAUC AAGUUCAGAGG ACACU UCCUGAUCGAAGGAGACCUGA ACCCGGACAAACAGCGACGUCGACAAG CUGUUC AUCCAGCUGGUCCAGACAUA CAACCAGCUGUUCGAAGAAAACCCGA UCAACGCAAGCGGAGUCGACGCAAAG GCAAUCCUGAGCGCAAGACUGAGCAA GAGCAGAAGACUGGAAAACCUGAUCG CACAGCUGCCGGGAGAAAAGAAGAAC GGACUGUUCGGAAACCUGAUCGCACU GAGCCUGGGACUGACACCGAACUUCA AGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACAC AUACGACGACGACCUGGACAACCUGC UGGCACAGAU CGGAGACCAGUACGCA GACCUGU UCCUGGCAGCAAAGAACC GAGCGACGCAAUCCUGCUGAGCGACA UCCUGAGAGUCAACACAGAAAUCACA AAGGCACCGCUGAGCGCAAGCAUGAU CAAGAGAUACGACGAACACCACCAGG ACCUGACACUGCUGAAGGCACUGGUC AGACAGCAGCUGCCGGAAAAGUACAA	1 5

10

20

30

40

50

	GGAAAUCUUCUUCGACCAGAGCAAGA ACGGAUACGCAGGAUACAUCGACGGA GGAGCAAGCCAGGAAGAAUUCUACAA GUUCAUCAAGCCGAUCCUGGAAAAGA UGGACGGAACAGAAGAACUGCUGGUC AAGCUGAACAGAGAAGACCUGCUGAG AAAGCAGAGAAACAUUCGACAACGGAA GCAUCCCGCACCCAGAUCCACCUGGGA GAACUGCACGCAAUCCUGAGAAGACA GGAAGACUUCUACCCGUUCCUGAAGG ACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUUCAGAAUCCCGUACUACGU CGGACCGCUGGCAAGAGGAAACAGCA GAUUCGCAUGGAUGACAAGAAAGAGC GAAGAAACAAUCACACCGUGGAACUU CGAAGAAGUCGUCGACAAGGGAGCAA GCGCACAGAGCUUCAUCGAAAGAAUG ACAAACUUCGACAAGAACCUGCCGAA CGAAAAGGUCCUGCCGAAGCACAGCC UGCUGUACGAAUACUUCACAGUCUAC AACGAACUGACAAAGGUCAAGUACGU CACAGAAGGAAUGAGAAAGCCGGCAU UCCUGAGCGGAGAACAGAAGAAGGCA AUCGUCGACCUGCUGUUCAAGACAAA CAGAAAGGUCACAGUCAAGCAGCUGA AGGAAGACUACUUCAAGAAGAUCGAA UGC UUCGACAGCGUCGAAAUCAGCGG AGUCGAAGACAGAUUCAACGCAAGCC UGGGAACAUAACACGACCUGCUGAAG AUCAUCAAGGACAAGGACUUCCUGGA CAACGAAGAAAACGAAGACAUCCUGG AAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAUCGAAGA AAGACUGAAGACAUAACGCACACCUGU UCGACGACAAGGUCAUGAAGCAGCUG AAGAGAAGAAGAUACACAGGAUGGGG AAGACUGAGCAGAAAGCUGAUCAACG GAAUCAGAGACAAGCAGAGCGGAAAG ACAAUCCUGGACUUCCUGAAGAGCGA CGGAUUCGCAAACAGAAACUUCAUGC AGCUGAUCCACGACGACAGCCUGACA UUCAAGGAAGACAUCCAGAAGGCACA	
		10
		20
		30
		40

	GGUCAGCGGACAGGGAGACAGCCUGC ACGAACACAUCGCAAACCUGGCAGGA AGCCCGGCAAUCAAGAAGGGAAUCCU GCAGACAGUCAAGGUCGUCGACGAAC UGGUCAAGGUCAUGGGAAGACACAAG CCGGAAAACAUCGUCAUCGAAAUGGC AAGAGAAAACCAGACAACACAGAAGG GACAGAAGAACAGCAGAGAAAGAAUG AAGAGAAUCGAAGAAGGAUCAAAGGA ACUGGGGAAGCCAGAUCCUGAAGGAAC ACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCU GCAGAACGGAAGAGACAUGUACGUCG ACCAGGAACUGGACAUCAACAGACUG AGCGACUACGACGUCGACCACAUCGU CCCGCAGAGCUUCCUGAAGGACGACA GCAUCGACAACAAGGUCCUGACAAGA AGCGACAAGAACAGAGGAAAGAGCGA CAACGUCCCGAGCGAAGAAGUCGUCA AGAAGAUGAAGAACUACUGGAGACAG CUGCUGAACGCAAAGCUGAUCACACA GAGAAAGUUCGACAACCUGACAAAGG CAGAGAGAGGAGGACUGAGCGAACUG GACAAGGCAGGAUUCAUCAAGAGACA GCUGGUUCGAAACAAGACAGAUCACAA AGCACGUCGCACAGAUAUCCUGGACAGC AGAAUGAACAACAAGUACGACGAAAA CGACAAGCUGAUCAGAGAAGUCAAGG UCAUCACACUGAAGAGCAAGCUGGUC AGCGACUUCAGAAAGGACUUCAGUU CUACAAGGUCAGAGAAAUCAACAACU ACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAU CAAGAAGUACCCGAAGCUGGAAAGCG AAUUCGUCUACGGAGACUACAAGGUC UACGACGUCAGAAAGAUGAUCGCAA GAGCGAACAGGAAAUCGGAAAGGCAA CAGCAAAGUACUUCUUCUACAGCAAC AUCAUGAACUUCUUCAAGACAGAAAU CACACUGGCAAACGGAGAAAUCAGAA AGAGACCGCUGAUCGAAACAAACGGA GAAACAGGAGAAAUCGUCUGGGACAA	
		10
		20
		30
		40

	GGGAAGAGACUUCGCAACAGUCAGAA AGGUCCUGAGCAUGCCGCAGGUCAAC AUCGUCAAGAAGACAGAAGUCCAGAC AGGAGGAUUCAGCAAGGAAAGCAUCC UGCCGAAGAGAAACAGCGACAAGCUG AUCGCAAGAAAGAAGGACUGGGACCC GAAGAAGUACGGAGGAUUCGACAGCC CGACAGUCGCAUACAGCGUCCUGGUC GUCGCAAAGGUCGAAAAGGGAAAGAG CAAGAAGCUGAAGAGCGUCAAGGAAC UGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGA CUUCCUGGAAGCAAAGGGAUACAAGG AAGUCAAGAAGGACCUGAUCAUCAAG CUGCCGAAGUACAGCCUGUUCGAACU GGAAAACGGAAGAAAGAGAAUGCUGG CAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUA CGUCAACUUCCUGUACCUGGCAAGCC ACUACGAAAAGCUGAAGGGAAAGCCCG GAAGACAACGAACAGAAGCAGCUGUU CGUCGAACAGCACAAAGCACUACCUGG ACGAAAUCAUCGAACAGAUCAAGCGAA UUCAGCAAGAGAGUCAUCCUGGCAGA CGCAAACCUGGACAAGGUCCUGAGCG CAUACAACAAGCACAGAGACAAGCCG AUCAGAGAACAGGCAGAAAACAUCAU CCACCUGUUCACACUGACAAACCUGG GAGCACCGGCAGCAUUCAAGUACUUC GACACAACAUAUCGACAGAAAGAGAU CACAAGCACAAAGGAAGUCCUGGACG CAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGACCU GAGCCAGCUGGGAGGAGAC	
C a s 9 ニッ カーゼのアミ ノ酸配列 (N LSなし)	MDKKYSIGLAIGTNSVGWAVITDEYK VPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFHRLEES FLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVD	16

10

20

30

40

50

	KLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAE DAKLQLSKDITYDDDLNLLAQIGDQY ADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKAL VRQQLPKEYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHL GELHAILLRRQEDFYPLKDNREKIEK ILTFRIPYYVGPLARGNSRFAWMTRK SEETITPWNFEEVVDKGASAQSFIER MTNFDKNLPNEKVLPKHSLLEYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKI ECFDSVEISGVEDRFNASLGTYHDLL KIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKTYAHLFDDKVMKQ LKRRRYTGWGRLSRKLINGIRDKQSG KTI LDFLKS DGFANRNF MQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLA GSPAIIKKGILQTVKVVDDELVKVMGRH KPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQL QNEKLYLYYLQNGRDMYVDQELDINR LSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKMKMKNYWR QLLNAKLITQRKFDNLTKAERGGLSE LDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKS KL VSDFRKDFQFYKVREINNYHHAHDAY LNAVVG TALIKKYPKLESEFVYG DYK VYDVRKMI AKSEQEIGKATAKYFFYS NIMNFFKTEITLANGEIRKRPLIETN GETGEIVWDKGRDFATVRKVL SMPQV NIVKKTEVQTGGFSKESILPKRNSDK LIARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLII KL PKYSLFELENGRKRMLASAGELQK GNE LALPSKYVNFLYLASHYEK LKGS PEDNEQKQLFVEQHKHYLDEII EQIS	

10

20

30

40

50

	EFSKRVI LADANLDKVL SAYNKHRDK PIREQAEN I IHLFTLTNLGAPAAF KY FDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDLSQLGGD	
開始コドン及 び終止コドン を有し、表3 に記載の最小 ウリジンコド ンを使用す る、配列番号 16をコード するCas9 ニッカーゼm RNA OR F	AUGGACAAGAAGUACAGCAUCGGACU GGCAAUCGGAACAAACAGCGUCGGAU GGGCAGUCAUCACAGACGAAUACAAG GUCCCGAGCAAGAAGUUCAAGGUCCU GGGAAACACAGACAGACACAGCAUCA AGAAGAACCUGAUUCGGAGCACUGCUG UUCGACAGCGGAGAAACAGCAGAAGC AACAAAGACUGAAGAGAACAGCAAGAA GAAGAUACACAAGAAGAAAGAACAGA AUCUGCUACCUGCAGGAAAUCUUCAG CAACGAAAUGGCAAAGGUCGACGACA GCUUCUUCACAGACUGGAAGAAAGC UUCUGGUCGAAGAAGACAAGAAGCA CGAAAGACACCCGAUCUUCGGAAACA UCGUCGACGAAGUCGCAUACCACGAA AAGUACCCGACAAUCUACCACCUGAG AAAGAAGCUGGUCGACAGCACAGACA AGGCAGACCUGAGACUGAUCUACCUG GCACUGGCACACAUGAUCAAGUUCAG AGGACACUUCUGAUCGAAGGAGACC UGAACCCGGACAACAGCGACGUCGAC AAGCUGUUCAUCCAGCUGGUCCAGAC AUACAACCAGCUGUUCGAAGAAAACC CGAUCAAACGCAAGCGGAGUCGACGCA AAGGCAAUCCUGAGCGCAAGACUGAG CAAGAGCAGAAGACUGGAAAACCUGA UCGCACAGCUGCCGGGAGAAAAGAAG AACGGACUGUUCGGAAACCUGAUCGC ACUGAGCCUGGGACUGACACCGAACU UCAAGAGCAACUUCGACCUGGCAGAA GACGCAAAGCUGCAGCUGAGCAAGGA CACAUACGACGACGACCUGGACAACC UGCUGGCACAGAUCGGAGACCAGUAC GCAGACCUGUUCUGGCAGCAAAGAA CCUGAGCGACGCAAUCCUGCUGAGCG ACAUCCUGAGAGUCAACACAGAAAUC ACAAAGGCACCGCUGAGCGCAAGCAU	17

10

20

30

40

50

	GAUCAAGAGAUACGACGAACACCACC AGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUA CAAGGAAAUCUUCUUCGACCAGAGCA AGAACGGAUACGCAGGAUACAUCGAC GGAGGAGCAAGCCAGGAAGAAUUCUA CAAGUUCAUCAAGCCGAUCCUGGAAA AGAUGGACGGAACAGAAGAACUGCUG GUCAAGCUGAACAGAGAAGACCUGCU GAGAAAGCAGAGAAACAUUCGACAACG GAAGCAUCCCGCACCCAGAUCCACCUG GGAGAACUGCACGCAAUCCUGAGAAG ACAGGAAGACUUCUACCCGUUCCUGA AGGACAACAGAGAAAAGAUCGAAAAG AUCCUGACAUUCAGAAUCCCGUACUA CGUCGGACCGCUGGCAAGAGGAAACA GCAGAUUCGCAUGGAUGACAAGAAAG AGCGAAGAAACAUAUCACACCGUGGAA CUUCGAAGAAGUCGUCGACAAGGGAG CAAGCGCACAGAGCUUCAUCGAAAGA AUGACAAACUUCGACAAGAACCUGCC GAACGAAAAGGUCCUGCCGAAGCACA GCCUGCUGUACGAAUACUUCACAGUC UACAACGAACUGACAAAGGUCAAGUA CGUCACAGAAGGAAUGAGAAAGCCGG CAUUCCUGAGCGGAGAACAGAAGAAG GCAAUCGUCGACCUGCUGUUCAAGAC AAACAGAAAGGUACAGUCAAGCAGC UGAAGGAAGACUACUUCAAGAAGAUC GAAUGCUUCGACAGCGUCGAAAUCAG CGGAGUCGAAGACAGAUUCAACGCAA GCCUGGGAACAUAACACGACCUGCUG AAGAUCAUCAAGGACAAGGACUUCCU GGACAACGAAGAAAACGAAGACAUCC UGGAAGACAUCGUCCUGACACUGACA CUGUUCGAAGACAGAGAAAUGAUCGA AGAAAGACUGAAGACAUAACGCACACC UGUUCGACGACAAGGUCAUGAAGCAG CUGAAGAGAAGAAGAUACACAGGAUG GGGAAGACUGAGCAGAAAGCUGAUCA ACGGAAUCAGAGACAAGCAGAGCGGA AAGACAAUCCUGGACUUCCUGAAGAG	
		10
		20
		30
		40

	CGACGGAUUCGGCAAACAGAAACUUUA UGCAGCUGAUCCACGACGACAGCCUG ACAUUCAAGGAAGACAUCCAGAAGGC ACAGGUCAGCGGACAGGGAGACAGCC UGCACGAACACAUCGCAAACCUGGCA GGAAGCCCGGCAAUCAAGAAGGGAAU CCUGCAGACAGUCAAGGUUCGUCGACG AACUGGUCAAGGUCAUGGGAAGACAC AAGCCGGAAAACAUCGUCAUCGAAAU GGCAAGAGAAAACCAGACAACACAGA AGGGACAGAAGAACAGCAGAGAAAGA AUGAAGAGAAUCGAAGAAGGAAUCAA GGAACUGGGAAGCCAGAUC CUGAAGG AACACCCGGUCGAAAACACACAGCUG CAGAACGAAAAGCUGUACCUGUACUA CCUGCAGAACGGAAGAGACAUGUACG UCGACCAGGAACUGGACAUCAACAGA CUGAGCGACUACGACGUCGACCACAU CGUCCCGCAGAGCUUCCUGAAGGACG ACAGCAUCGACAACAAGGUCCUGACA AGAAGCGACAAGAACAGAGGAAAGAG CGACAACGUCCCGAGCGAAGAAGUCG UCAAGAAGAUGAAGAACUACUGGAGA CAGCUGCUGAACGCAAAGCUGAUCAC ACAGAGAAAGUUCGACAACCUGACAA AGGCAGAGAGAGGAGGACUGAGCGAA CUGGACAAGGCAGGAUUCAUCAAGAG ACAGCUGGUUCGAAACAAGACAGAUCA CAAAGCACGUCGCACAGAUCCUGGAC AGCAGAAUGAACACAAAGUACGACGA AAACGACAAGCUGAUCAGAGAAGUCA AGGUCAUCACACUGAAGAGCAAGCUG GUCAGCGACUUCAGAAAGGACUUCCA GUUCUACAAGGUCAGAGAAAUCAACA ACUACCACCACGCACACGACGCAUAC CUGAACGCAGUCGUCGGAACAGCACU GAUCAAGAAGUACCCGAAGCUGGAAA GCGAAUUCGUCUACGGAGACUACAAG GUCUACGACGUCAGAAAGAUGAUCGC AAAGAGCGAACAGGAAAUCGGAAAGG CAACAGCAAAGUACUUCUUCUACAGC AACAUCAUGAACUUCUUCAAGACAGA	
		10
		20
		30
		40

	AAUCACACUGGC AAAACGGAGAAAUCA GAAAGAGACCGCUGAUCGAAACAAAC GGAGAAACAGGAGAAAUCGUCUGGGA CAAGGGAAGAGACUUCGCAACAGUCA GAAAGGUCCUGAGCAUGCCGCAGGUC AACAUUCGUCAAGAAGACAGAAGUCCA GACAGGAGGAUUCAGCAAGGAAAGCA UCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGA CCCGAAGAAGUACGGAGGAUUCGACA GCCCGACAGUCGCAUACAGCGUCCUG GUCGUCGCAAAGGUCGAAAAGGGAAA GAGCAAGAAGCUGAAGAGCGUCAAGG AACUGCUGGGAAUCACAAUCAUGGAA AGAAGCAGCUUCGAAAAGAACCCGAU CGACUUCCUGGAAGCAAAGGGAUACA AGGAAGUCAAGAAGGACCUGAUCAUC AAGCUGCCGAAGUACAGCCUGUUCGA ACUGGAAAACGGAAGAAAGAGAAUGC UGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAA GUACGUCAACUUCCUGUACCUGGCAA GCCACUACGAAAAGCUGAAGGGAAGC CCGGAAGACAACGAACAGAAGCAGCU GUUCGUCGAACAGCACAAAGCACUACC UGGACGAAAUCAUCGAACAGAUCAAGC GAAUUCAGCAAGAGAGUCAUCCUGGC AGACGCAAACCUGGACAAGGUCCUGA GCGCAUACAACAAGCACAGAGACAAG CCGAUCAGAGAACAGGCAGAAAACAU CAUCCACCUGUUCACACUGACAAACC UGGGAGCACCCGGCAGCAUUCAAGUAC UUCGACACAACAUAUCGACAGAAAGAG AUACACAAGCACAAAGGAAGUCCUGG ACGCAACACUGAUCCACCAGAGCAUC ACAGGACUGUACGAAACAAGAAUCGA CCUGAGCCAGCUGGGAGGAGACUAG	
表 3 に記載の 最小ウリジン コドンを使用 する、配列番 号 16 をコー	GACAAGAAGUACAGCAUCGGACUGGC AAUCGGAACAAACAGCGUCGGAUGGG CAGUCAUCACAGACGAAUACAAGGUC CCGAGCAAGAAGUUCAAGGUCCUGGG AAACACAGACAGACACAGCAUCAAGA	18

10

20

30

40

50

40

	GCAUCCCCGCACCAGAUCCACCUGGGA GAACUGCACGCAAUCCUGAGAAGACA GGAAGACUUCUACCCGUUCCUGAAGG ACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUUCAGAAUCCCGUACUACGU CGGACCGCUGGCAAGAGGAAACAGCA GAUUCGCAUGGAUGACAAGAAAGAGC GAAGAAACAAUCACACCGUGGAACUU CGAAGAAGUCGUCGACAAGGGAGCAA GCGCACAGAGCUUCAUCGAAAGAAUG ACAAACUUCGACAAGAACCUGCCGAA CGAAAAGGUCCUGCCGAAGCACAGCC UGCUGUACGAAUACUUCACAGUCUAC AACGAACUGACAAAGGUCAAGUACGU CACAGAAGGAAUGAGAAAGCCGGCAU UCCUGAGCGGAGAACAGAAGAAGGCA AUCGUCGACCUGCUGUUCAAGACAAA CAGAAAGGUCACAGUCAAGCAGCUGA AGGAAGACUACUUCAAGAAGAUCGAA UGC UUCGACAGCGUCGAAAUCAGCGG AGUCGAAGACAGAUUCAACGCAAGCC UGGGAACAUAACCACGACCUGCUGAAG AUCAUCAAGGACAAGGACUUCCUGGA CAACGAAGAAAACGAAGACAUCCUGG AAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAUCGAAGA AAGACUGAAGACAUAACGCACACCUGU UCGACGACAAGGUCAUGAAGCAGCUG AAGAGAAGAAGAUACACAGGAUGGGG AAGACUGAGCAGAAAGCUGAUCAACG GAAUCAGAGACAAGCAGAGCGGAAAG ACAAUCCUGGACUUCCUGAAGAGCGA CGGAUUCGCAAACAGAAACUUCAUGC AGCUGAUCCACGACGACAGCCUGACA UUCAAGGAAGACAUCCAGAAGGCACA GGUCAGCGGACAGGGAGACAGCCUGC ACGAACACAUCGCAAACCUGGCAGGA AGCCCGGCAAUCAAGAAGGGAAUCCU GCAGACAGUCAAGGUUCGUCGACGAAC UGGUCAAGGUCAUGGGAAGACACAAG CCGGAAAACAUCGUCAUCGAAAUGGC AAGAGAAAACCAGACAACACAGAAGG	
		10
		20
		30
		40

	GACAGAAGAACAGCAGAGAAAGAAUG AAGAGAAUCGAAGAAGGAAUCAAGGA ACUGGGAAGCCAGAUCUGAAGGAAC ACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCU GCAGAACGGAAGAGACAUGUACGUUCG ACCAGGAACUGGACAUCAACAGACUG AGCGACUACGACGUCGACCACAUCGU CCCGCAGAGCUUCCUGAAGGACGACA GCAUCGACAACAAGGUCCUGACAAGA AGCGACAAGAACAGAGGAAAGAGCGA CAACGUCCCGAGCGAAGAAGUCGUCA AGAAGAUGAAGAACUACUGGAGACAG CUGCUGAACGCAAAGCUGAUCACACA GAGAAAGUUCGACAACCUGACAAAGG CAGAGAGAGGAGGACUGAGCGAACUG GACAAGGCAGGAUUCAUCAAGAGACA GCUGGUCGAAACAAGACAGAUCACAA AGCACGUCGCACAGAUCUGGACAGC AGAAUGAACACAAAGUACGACGAAAA CGACAAGCUGAUCAGAGAAGUCAAGG UCAUCACACUGAAGAGCAAGCUGGUC AGCGACUUCAGAAAGGACUUCCAGUU CUACAAGGUCAGAGAAAUCAACAACU ACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAU CAAGAAGUACCCGAAGCUGGAAAGCG AAUUCGUCUACGGAGACUACAAGGUC UACGACGUCAGAAAGAUGAUCGCAA GAGCGAACAGGAAAUCGGAAAGGCAA CAGCAAAGUACUUCUUCUACAGCAAC AUCAUGAACUUCUUCAAGACAGAAAU CACACUGGCAAACGGAGAAAUCAGAA AGAGACCGCUGAUCGAAACAAACGGA GAAACAGGAGAAAUCGUCUGGGACAA GGGAAGAGACUUCGCAACAGUCAGAA AGGUCCUGAGCAUGCCGCAGGUCAAC AUCGUCAAGAAGACAGAAGUCCAGAC AGGAGGAUUCAGCAAGGAAAGCAUCC UGCCGAAGAGAAACAGCGACAAGCUG AUCGCAAGAAAGAAGGACUGGGACCC GAAGAAGUACGGAGGAUUCGACAGCC	
		10
		20
		30
		40

	CGACAGUCGCAUACAGCGUCCUGGUC GUCGCAAAGGUCGAAAAGGGAAAGAG CAAGAAGCUGAAGAGCGUCAAGGAAC UGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGA CUUCCUGGAAGCAAAGGGGAUACAAGG AAGUCAAGAAGGACCUGAUCAUCAAG CUGCCGAAGUACAGCCUGUUCGAACU GGAAAACGGAAGAAAGAGAAUGCUGG CAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUA CGUCAACUUCUGUACCUGGCAAGCC ACUACGAAAAGCUGAAGGGAAGCCCG GAAGACAACGAACAGAAGCAGCUGUU CGUCGAACAGCACAAAGCACUACCUGG ACGAAAUCAUCGAACAGAUACAGCGAA UUCAGCAAGAGAGUCAUCCUGGCAGA CGCAAACCUGGACAAGGUCCUGAGCG CAUACAACAAGCACAGAGACAAGCCG AUCAGAGAACAGGCAGAAAACAUCAU CCACCUGUUCACACUGACAAACCUGG GAGCACCGGCAGCAUUCAAGUACUUC GACACAACAUAUCGACAGAAAGAGAU CACAAGCACAAAGGAAGUCCUGGACG CAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGACCU GAGCCAGCUGGGAGGAGAC	10
d C a s 9 の アミノ酸配列 (N L S な し)	MDKKYSIGLAIGTNSVGWAVITDEYK VPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFHRLEES FLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEEENPINASGVDA KAILSARLSKSRRLLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAE DAKLQLSKDITYDDDLNLLAQIGDQY ADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKAL VRQQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELL	19
		30
		40

	VKLNREDLLRKQRTFDNGS I PHQ I HL GELHAILRRQEDFY P FLKDNREKIEK I L T F R I P Y Y V G P L A R G N S R F A W M T R K S E E T I T P W N F E E V V D K G A S A Q S F I E R M T N F D K N L P N E K V L P K H S L L Y E Y F T V Y N E L T K V K Y V T E G M R K P A F L S G E Q K K A I V D L L F K T N R K V T V K Q L K E D Y F K K I E C F D S V E I S G V E D R F N A S L G T Y H D L L K I I K D K D F L D N E E N E D I L E D I V L T L T L F E D R E M I E E R L K T Y A H L F D D K V M K Q L K R R R Y T G W G R L S R K L I N G I R D K Q S G K T I L D F L K S D G F A N R N F M Q L I H D D S L T F K E D I Q K A Q V S G Q G D S L H E H I A N L A G S P A I K K G I L Q T V K V V D E L V K V M G R H K P E N I V I E M A R E N Q T T Q K G Q K N S R E R M K R I E E G I K E L G S Q I L K E H P V E N T Q L Q N E K L Y L Y Y L Q N G R D M Y V D Q E L D I N R L S D Y D V D A I V P Q S F L K D D S I D N K V L T R S D K N R G K S D N V P S E E V V K K M K N Y W R Q L L N A K L I T Q R K F D N L T K A E R G G L S E L D K A G F I K R Q L V E T R Q I T K H V A Q I L D S R M N T K Y D E N D K L I R E V K V I T L K S K L V S D F R K D F Q F Y K V R E I N N Y H H A H D A Y L N A V V G T A L I K K Y P K L E S E F V Y G D Y K V Y D V R K M I A K S E Q E I G K A T A K Y F F Y S N I M N F F K T E I T L A N G E I R K R P L I E T N G E T G E I V W D K G R D F A T V R K V L S M P Q V N I V K K T E V Q T G G F S K E S I L P K R N S D K L I A R K K D W D P K K Y G G F D S P T V A Y S V L V V A K V E K G K S K K L K S V K E L L G I T I M E R S S F E K N P I D F L E A K G Y K E V K K D L I I K L P K Y S L F E L E N G R K R M L A S A G E L Q K G N E L A L P S K Y V N F L Y L A S H Y E K L K G S P E D N E Q K Q L F V E Q H K H Y L D E I I E Q I S E F S K R V I L A D A N L D K V L S A Y N K H R D K P I R E Q A E N I I H L F T L T N L G A P A A F K Y F D T T I D R K R Y T S T K E V L D A T L I H Q S I T G L Y E T R I D L S Q L G G D	
開始コドン及 び終止コドン を有し、表 3	AUGGACAAGAAGUACAGCAUCGGACU GGCAAUCGGAACAAACAGCGUCGGAU GGGCAGUCAUCACAGACGAAUACAAG	20

10

20

30

40

50

に記載の最小 ウリジンコド ンを使用す る、配列番号 19をコード する d C a s 9 m R N A O R F	GUCCCGAGCAAGAAGUUCAAGGUCCU GGGAAACACAGACAGACACAGCAUCA AGAAGAACCUGAUCGGAGCACUGCUG UUCGACAGCGGAGAAACAGCAGAAGC AACAAGACUGAAGAGAACAGCAAGAA GAAGAUACACAAGAAGAAAGAACAGA AUCUGCUACCUGCAGGAAAUCUUCAG CAACGAAAUGGC AAAAGGUCGACGACA GCUUCUUCACAGACUGGAAGAAAGC UUCUGGUCGAAGAAGACAAGAAGCA CGAAAGACACCCGAUCUUCGGAAACA UCGUCGACGAAGUCGCAUACCACGAA AAGUACCCGACAAUCUACCACCUGAG AAAGAAGCUGGUCGACAGCACAGACA AGGCAGACCUGAGACUGAUCUACCUG GCACUGGCACACAUGAUCAAGUUCAG AGGACACUUCUGAUCGAAGGAGACC UGAACCCGGACAACAGCGACGUCGAC AAGCUGUUCAUCCAGCUGGUCCAGAC AUACAACCAGCUGUUCGAAGAAAACC CGAUCAACGCAAGCGGAGUCGACGCA AAGGCAAUCCUGAGCGCAAGACUGAG CAAGAGCAGAAGACUGGAAAACCUGA UCGCACAGCUGCCGGGAGAAAAGAAG AACGGACUGUUCGGAAACCUGAUCGC ACUGAGCCUGGGACUGACACCGAACU UCAAGAGCAACUUCGACCUGGCAGAA GACGCAAAGCUGCAGCUGAGCAAGGA CACAUACGACGACGACCUGGACAACC UGCUGGCACAGAUCCGGAGACCAGUAC GCAGACCUGUUCUGGCAGCAAAGAA CCUGAGCGACGCAAUCCUGCUGAGCG ACAUCCUGAGAGUCAACACAGAAAUC ACAAAGGCACCGCUGAGCGCAAGCAU GAUCAAGAGAUACGACGAACACCACC AGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGA AAAAGUA CAAGGAAAUCUUCUUCGACCAGAGCA AGAACGGAUACGCAGGAUACAUCGAC GGAGGAGCAAGCCAGGAAGAAUUCUA CAAGUUCAUCAAGCCGAUCCUGGAAA AGAUGGACGGAACAGAAGAACUGCUG	10 20 30 40

	GUCAAGCUGAACAGAGAAGACCUGCU GAGAAAGCAGAGAACAUUCGACAACG GAAGCAUCCCGCACCAAGAUCCACCUG GGAGAACUGCACGCAAUCCUGAGAAG ACAGGAAGACUUCUACCCGUUCCUGA AGGACAACAGAGAAAAGAUCGAAAAG AUCCUGACAUUCAGAAUCCCGUACUA CGUCGGACCGCUGGCAAGAGGAAACA GCAGAUUCGCAUGGAUGACAAGAAAG AGCGAAGAAACAAUCACACCGUGGAA CUUCGAAGAAGUCGUCGACAAGGGAG CAAGCGCACAGAGCUUCAUCGAAAGA AUGACAAACUUCGACAAGAACCUGCC GAACGAAAAGGUCCUGCCGAAGCACA GCCUGCUGUACGAAUACUUCACAGUC UACAACGAACUGACAAAGGUCAAGUA CGUCACAGAAGGAAUGAGAAAGCCGG CAUUCCUGAGCGGAGAACAGAAGAAG GCAAUCGUCGACCUGCUGUUCAAGAC AAACAGAAAGGUCACAGUCAAGCAGC UGAAGGAAGACUACUUCAAGAAGAUC GAAUGCUUCGACAGCGUCGAAAUCAG CGGAGUCGAAGACAGAUUCAACGCAA GCCUGGGAACAUAACACGACCUGCUG AAGAUCAUCAAGGACAAGGACUUCCU GGACAACGAAGAAAACGAAGACAUCC UGGAAGACAUCGUCCUGACACUGACA CUGUUCGAAGACAGAGAAAUGAUCGA AGAAAGACUGAAGACAUAACGCACACC UGUUCGACGACAAGGUCAUGAAGCAG CUGAAGAGAAGAAGAUACACAGGAUG GGGAAGACUGAGCAGAAAGCUGAUCA ACGGAAUCAGAGACAAGCAGAGCGGA AAGACAAUCCUGGACUUCCUGAAGAG CGACGGAUUCGCAAACAGAAACUUCA UGCAGCUGAUCCACGACGACAGCCUG ACAUUCAAGGAAGACAUCCAGAAGGC ACAGGUCAGCGGACAGGGAGACAGCC UGCACGAACACAUCGCAAACCUGGCA GGAAGCCCGGCAAUCAAGAAGGGAAU CCUGCAGACAGUCAAGGUCGUCGACG AACUGGUCAAGGUCAUGGGAAGACAC	

10

20

30

40

50

	AAGCCGGAAAACAUCGUCAUCGAAAU GGCAAGAGAAAACCAGACAACACAGA AGGGACAGAAAGAACAGCAGAGAAAGA AUGAAGAGAAUCGAAGAAGGAAUCAA GGAACUGGGAAGCCAGA UCCUGAAGG AACACCCGGUCGAAAACACACAGCUG CAGAACGAAAAGCUGUACCUGUACUA CCUGCAGAACGGAAGAGACAUGUACG UCGACCAGGAACUGGACAUCAACAGA CUGAGCGACUACGACGUCGACGCAAU CGUCCCGCAGAGCUUCCUGAAGGACG ACAGCAUCGACAACAAGGUCCUGACA AGAAGCGACAAGAACAGAGGAAAGAG CGACAACGUCCCGAGCGAAGAAGUCG UCAAGAAGAUGAAGAACUACUGGAGA CAGCUGCUGAACGCAAAGCUGAUCAC ACAGAGAAAGUUCGACAACCUGACAA AGGCAGAGAGAGGAGGACUGAGCGAA CUGGACAAGGCAGGAUUCAUCAAGAG ACAGCUGGUCGAAACAAGACAGAUCA CAAAGCACGUCGCGACAGAUCCUGGAC AGCAGAAUGAACACAAAGUACGACGA AAACGACAAGCUGAUCAGAGAAGUCA AGGUCAUCACACUGAAGAGCAAGCUG GUCAGCGACUUCAGAAAGGACUUCCA GUUCUACAAGGUCAGAGAAAUCAACA ACUACCACCACGCGACACGACGCAUAC CUGAACGCGAGUCGUCGGAACAGCACU GAUCAAGAAGUACCCGAAGCUGGAAA GCGAAUUCGUCUACGGAGACUACAAG GUCUACGACGUCAGAAAGAUGAUCGC AAAGAGCGAACAGGAAAUCGGAAAGG CAACAGCAAAGUACUUCUUCUACAGC AACAUCAUGAACUUCUUCAAGACAGA AAUCACACUGGCAAACGGAGAAAUCA GAAAGAGACCGCUGAUCGAAACAAAC GGAGAAACAGGAGAAAUCGUCUGGGA CAAGGGAAGAGACUUCGCAACAGUCA GAAAGGUCCUGAGCAUGCCGCAGGUC AACAUCGUCAAGAAGACAGAAGUCCA GACAGGAGGAUUCAGCAAGGAAAGCA UCCUGCCGAAGAGAAACAGCGACAAG	

10

20

30

40

50

	CUGAUCGCAAGAAAGAAGGACUGGGA CCCGAAGAAGUACGGAGGAUUCGACA GCCCGACAGUCGCAUACAGCGUCCUG GUCGUCGCAAAGGUCGAAAAGGGAAA GAGCAAGAAGCUGAAGAGCGUCAAGG AACUGCUGGGAAUCACAAUCAUGGAA AGAAGCAGCUUCGAAAAGAACCCGAU CGACUUCCUGGAAGCAAAGGGAUACA AGGAAGUCAAGAAGGACCUGAUCAUC AAGCUGCCGAAGUACAGCCUGUUCGA ACUGGAAAACGGAAGAAAGAGAAUGC UGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAA GUACGUCAACUUCUGUACCUGGCAA GCCACUACGAAAAGCUGAAGGGAAGC CCGGAAGACAACGAACAGAAGCAGCU GUUCGUCGAACAGCACAAAGCACUACC UGGACGAAAUCAUCGAACAGAUCAAGC GAAUUCAGCAAGAGAGUCAUCCUGGC AGACGCAAACCUGGACAAGGUCCUGA GCGCAUACAACAAGCACAGAGACAAG CCGAUCAGAGAACAGGCAGAAAACA CAUCCACCUGUUCACACUGACAAACC UGGGAGCACCCGGCAGCAUUCAAGUAC UUCGACACAACAUAUCGACAGAAAGAG AUACACAAGCACAAAGGAAGUCCUGG ACGCAACACUGAUCCACCAGAGCAUC ACAGGACUGUACGAAACAAGAAUCGA CCUGAGCCAGCUGGGAGGAGACUAG		10
表3に記載の 最小ウリジン コドンを使用 する、配列番 号19をコード するdCas9コード配 列（開始コド ンまたは終止 コドンはない；融合タン パク質コード 配列に含める	GACAAGAAGUACAGCAUCGGACUGGC AAUCGGAACAAACAGCGUCGGAUGGG CAGUCAUCACAGACGAAUACAAGGUC CCGAGCAAGAAGUUCAAGGUCCUGGG AAACACAGACAGACACAGCAUCAAGA AGAACCUGAUUCGGAGCACUGCUGUUC GACAGCGGAGAAACAGCAGAAGCAAC AAGACUGAAGAGAACAGCAAGAAGAA GAUACACAAGAAGAAAGAACAGAAUC UGC UACCUGCAGGAAAUCUUCAGCAA CGAAAUGGCAAAGGUCGACGACAGCU UCUUCACACAGACUGGAAGAAAGCUUC CUGGUCGAAGAAGACAAGAAGCACGA	21	30
			40

際に好適)	AAGACACCCGAUCUUCGGAAACAUCG	
	UCGACGAAGUCGCAUACCACGAAAAG	
	UACCCGACAAUCUACCACCUGAGAAA	
	GAAGCUGGUCGACAGCACAGACAAGG	
	CAGACCUGAGACUGAUCUACCUGGCA	
	CUGGCACACAUGAUCAAGUUCAGAGG	
	ACACUUCCUGAUCGAAGGAGACCUGA	
	ACCCGGACAACAGCGACGUCGACAAG	
	CUGUUCAUCCAGCUGGUCCAGACAUA	10
	CAACCAGCUGUUCGAAGAAAACCCGA	
	UCAACGCAAGCGGAGUCGACGCAAAG	
	GCAAUCCUGAGCGCAAGACUGAGCAA	
	GAGCAGAAGACUGGAAAACCUGAUCG	
	CACAGCUGCCGGGAGAAAAGAAGAAC	
	GGACUGUUCGGAAACCUGAUCGCACU	
	GAGCCUGGGACUGACACCGAACUUCA	
	AGAGCAACUUCGACCUGGCAGAAAGAC	
	GCAAAGCUGCAGCUGAGCAAGGACAC	
	AUACGACGACGACCUGGACAACCUGC	
	UGGCACAGAUUCGGAGACCAGUACGCA	20
	GACCUGUUCUGGCAGCAAAGAACCU	
	GAGCGACGCAAUCCUGCUGAGCGACA	
	UCCUGAGAGUCAACACAGAAAUCACA	
	AAGGCACCGCUGAGCGCAAGCAUGAU	
	CAAGAGAUACGACGAACACCACCAGG	
	ACCUGACACUGCUGAAGGCACUGGUC	
	AGACAGCAGCUGCCGGAAAAGUACAA	
	GGAAAUCUUCUUCGACCAGAGCAAGA	
	ACGGAUACGCAGGAUACAUCGACGGA	
	GGAGCAAGCCAGGAAGAAUUCUACAA	
	GUUCAUCAAGCCGAUCCUGGAAAAGA	30
	UGGACGGAACAGAAGAACUGCUGGUC	
	AAGCUGAACAGAGAAGACCUGCUGAG	
	AAAGCAGAGAACAUUCGACAACGGAA	
	GCAUCCCGCACCCAGAUCCACCUGGGA	
	GAACUGCACGCAAUCCUGAGAAGACA	
	GGAAGACUUCUACCCGUUCCUGAAGG	
	ACAACAGAGAAAAGAUCGAAAAGAUC	
	CUGACAUUCAGAAUCCCGUACUACGU	
	CGGACCGCUGGCAAGAGGAAACAGCA	
	GAUUCGCAUGGAUGACAAGAAAGAGC	
	GAAGAAACAAUCACACCGUGGAACUU	40

	CGAAGAAGUCGUCGACAAGGGAGCAA GCGCACAGAGCUUCAUCGAAAGAAUG ACAAACUUCGACAAGAACCUGCCGAA CGAAAAGGUCCUGCCGAAGCACAGCC UGCUGUACGAAUACUUCACAGUCUAC AACGAACUGACAAAGGUCAAGUACGU CACAGAAGGAAUGAGAAAGCCGGCAU UCCUGAGCGGAGAACAGAAGAAGGCA AUCGUCGACCUGCUGUUCAAGACAAA CAGAAAGGUCACAGUCAAGCAGCUGA AGGAAGACUACUUCAAGAAGAU CGAA UGC UUCGACAGCGUCGAAAUCAGCGG AGUCGAAGACAGAUUCAACGCAAGCC UGGGAACAUAACCACGACCUGCUGAAG AUCAUCAAGGACAAGGACU UCCUGGA CAACGAAGAAAACGAAGACA UCCUGG AAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAU CGAAGA AAGACUGAAGACAUAACGCACACCUGU UCGACGACAAGGUCAUGAAGCAGCUG AAGAGAAGAAGAUACACAGGAUGGGG AAGACUGAGCAGAAAGCUGAUCAACG GAAUCAGAGACAAGCAGAGCGGAAAG ACAAUCCUGGACU UCCUGAAGAGCGA CGGAUUCGCAAACAGAAACUUCAUGC AGCUGAUCCACGACGACAGCCUGACA UUCAAGGAAGACA UCCAGAAGGCACA GGUCAGCGGACAGGGAGACAGCCUGC ACGAACACAUCGCAAACCUGGCAGGA AGCCC GGCAAUCAAGAAGGGAAUCCU GCAGACAGUCAAGGUUCGUCGACGAAC UGGUCAAGGUCAUGGGAAGACACAAG CCGGAAAACAUCGUCAUCGAAAUGGC AAGAGAAAACCAGACAACACAGAAGG GACAGAAGAACAGCAGAGAAAGAAUG AAGAGAAUCGAAGAAGGAAUCAAGGA ACUGGGGAAGCCAGA UCCUGAAGGAAC ACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCU GCAGAACGGAAGAGACAUGUACGUUCG ACCAGGAACUGGACA UCAACAGACUG AGCGACUACGACGUCGACGCAAUCGU	
		10
		20
		30
		40

	CCCGCAGAGCUUCCUGAAGGACGACA GCAUCGACAACAAGGUCCUGACAAGA AGCGACAAGAACAGAGGAAAGAGCGA CAACGUCCCGAGCGAAGAAGUCGUCA AGAAGAUGAAGAACUACUGGAGACAG CUGCUGAACGCAAAGCUGAUCACACA GAGAAAGUUCGACAACCUGACAAGG CAGAGAGAGGAGGACUGAGCGAACUG GACAAGGCAGGAUUCAUCAAGAGACA GCUGGUCGAAACAAGACAGAUCACAA AGCACGUUCGCACAGAUCUGGACAGC AGAAUGAACACAAAGUACGACGAAAA CGACAAGCUGAUCAGAGAAGUCAAGG UCAUCACACUGAAGAGCAAGCUGGUC AGCGACUUCAGAAAGGACUUCCAGUU CUACAAGGUCAGAGAAAUCAACAACU ACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAU CAAGAAGUACCCGAAGCUGGAAAGCG AAUUCGUCUACGGAGACUACAAGGUC UACGACGUCAGAAAGAUGAUCGCAAA GAGCGAACAGGAAAUCGGAAAGGCAA CAGCAAAGUACUUCUUCUACAGCAAC AUCAUGAACUUCUUCAAGACAGAAAU CACACUGGCAAACGGAGAAAUCAGAA AGAGACCGCUGAUCGAAACAAACGGA GAAACAGGAGAAAUCGUCUGGGACAA GGGAAGAGACUUCGCAACAGUCAGAA AGGUCCUGAGCAUGCCGCAGGUCAAC AUCGUCAAGAAGACAGAAGUCCAGAC AGGAGGAUUCAGCAAGGAAAGCAUCC UGCCGAAGAGAAAACAGCGACAAGCUG AUCGCAAGAAAGAAGGACUGGGACCC GAAGAAGUACGGAGGAUUCGACAGCC CGACAGUCGCAUACAGCGUCCUGGUC GUCGCAAAGGUCGAAAAGGGAAAGAG CAAGAAGCUGAAGAGCGUCAAGGAAC UGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGA CUUCCUGGAAGCAAAGGGAUACAAGG AAGUCAAGAAGGACCUGAUCAUCAAG CUGCCGAAGUACAGCCUGUUCGAACU	

10

20

30

40

50

	GGAAAACGGAAGAAAGAGAAUGCUGG CAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUA CGUCAACUUCUGUACCUUGGCAAGCC ACUACGAAAAGCUGAAGGGAAGCCCG GAAGACAACGAACAGAAGCAGCUGUU CGUCGAACAGCACAAAGCACUACCUUGG ACGAAAUCAUCGAACAGAUCAGCGAA UUCAGCAAGAGAGUCAUCCUGGCAGA CGCAAACCUUGGACAAGGUCCUGAGCG CAUACAACAAGCACAGAGACAAGCCG AUCAGAGAAACAGGCAGAAAACAUCAU CCACCUUGUUCACACUGACAAACCUUGG GAGCACCGGCAGCAUUCAAGUACUUC GACACAACAUAUCGACAGAAAGAGAU CACAAGCACAAAGGAAGUCCUGGACG CAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGACCU GAGCCAGCUGGGAGGAGACGGAGGAG GAAGC	
C末端のアミノ酸として2つの核局在化シグナルを有する、Cas9のアミノ酸配列	MDKKYSIGLDIGTNSVGWAVITDEYK VPSKKFKVLGNTRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFHRLEES FLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKS NFDLAE DAKLQLSKD TYDDDLNLLAQIGDQY ADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKAL VRQQLPKEYKEIFFDQSKNGYAGYID GGASQEEFYKFIPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHL GELHAILRRQEDFYFPLKDNREKIEK ILTFRIPYYVGPLARGNSRFAWMTRK SEETITPWNFEVVDKGASAQSFIER MTNFDKNLPNEKVL PKHSLLEYEFTV YNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKI	22

10

20

30

40

50

	ECFDSVEISGVEDRNFNASLGTYHDL LKIIKDKDFLDNEENEDILEDIVLT LTFEDREMIEERLKTIAHLFDDKVMKQ LKRRRYTGWGRLSRKLINGIRDKQSG KTI LDFLKS DGFANRNF MQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLA GSPAIIKKGILQTVKVVDDELVKVMGRH KPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQL QNEKLYLYYLQNGRDMYVDQELDINR LSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKMKMKNYWR QLLNAKLITQRKFDNLTKAERGGLSE LDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKSKL VSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTAIIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEI GKATAKYFFYS NIMNFFKTEITLANGEIRKRPLIETN GETGEIVWDKGRDFATVRKVL SMPQV NIVKKTEVQTGGFSKESILPKRNSDK LIARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLII KLPKYSLFELENGRKRMLASAGELQK GNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEII EQIS EFSKRVILADANLDKVLSAYNKHHRDK PIREQAENIIHLFTLTNLGAPAAFKY FDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGD GSGSPKKKR KVDGSPKKKRKVD SG		10
			20
			30
開始コドン及び終止コドン を有し、表3 に記載の最小 ウリジンコ ンを使用す る、配列番号 22をコード するCas9	AUGGACAAGAAGUACAGCAUCGGACU GGACAUCGGAACAAACAGCGUCGGAU GGGCAGUCAUCACAGACGAAUACAAG GUCCCGAGCAAGAAGUUCAAGGUCCU GGGAAACACAGACAGACACAGCAUCA AGAAGAACCUGAUCGGAGCACUGCUG UUCGACAGCGGAGAAACAGCAGAAGC AACAAGACUGAAGAGAACAGCAAGAA GAAGAUACACAAGAAGAAAGAACAGA	23	40

mRNA RF	AUCUGCUACCUUGCAGGAAAUCUUCAG CAACGAAAUGGCAAAGGU CGACGACA GCUUCUUCACAGACUGGAAGAAAGC UUCUGGUCGAAGAAGACAAGAAGCA CGAAAGACACCCGAUCUUCGGAAACA UCGUCGACGAAGUCGCAUACCAAGAA AAGUACCCGACAAUCUACCAACUGAG AAAGAAGCUGGUCGACAGCACAGACA AGGCAGACCUGAGACUGAUCUACCU GCACUGGCACACAUGAUC AAGUUCAG AGGACACUUCUGAUCGAAGGAGACC UGAACCCGGACAACAGCGACGUCGAC AAGCUGUUCAUCCAGCUGGUCCAGAC AUACAACCAGCUGUUCGAAGAAAACC CGAUC AACGCAAGCGGAGUCGACGCA AAGGCAAUCCUGAGCGCAAGACUGAG CAAGAGCAGAAGACUGGAAAACCU UGGCACAGCUGCCGGGAGAAAAGAAG AACGGACUGUUCGGAAACCU GAUCGC ACUGAGCCUGGGACUGACACCGAACU UCAAGAGCAAACUUCGACCUGGCAGAA GACGCAAAGCUGCAGCUGAGCAAGGA CACAUACGACGACGACCUGGACAACC UGCUGGCACAGAU CGGAGACCAGUAC GCAGACCUGUUCUGGCAGCAAAGAA CCUGAGCGACGCAAUCCUGCUGAGCG ACAUCCUGAGAGUCAACACAGAAAU ACAAAGGCACCGCUGAGCGCAAGCAU GAUCAAGAGAUACGACGAACACCACC AGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUA CAAGGAAAUCUUCUUCGACCAGAGCA AGAACGGAUACGCAGGAUACAUCGAC GGAGGAGCAAGCCAGGAAGAAUUCUA CAAGUUCAUCAAGCCGAUCCUGGAAA AGAUGGACGGAACAGAAGAACUGCUG GUCAAGCUGAACAGAGAAGACCUGCU GAGAAAGCAGAGAACAUUCGACAACG GAAGCAUCCCGCACCAAGAUCCACCUG GGAGAACUGCACGCAAUCCUGAGAAG ACAGGAAGACUUCUACCCGUUCCUGA AGGACAACAGAGAAAAGAU CGAAAAG	

10

20

30

40

50

50

50

	AGAAAGCAGCUUCGAAAAGAACC CGAU CGACUUCCUGGAAGCAAAGGGAUACA AGGAAGUCAAGAAGGACCUGAUCAUC AAGCUGCCGAAGUACAGCCUGUUCGA ACUGGAAAACGGAAGAAAGAGAAUGC UGGCAAGCGCAGGAGAACUGCAGAAAG GGAAACGAACUGGCACUGCCGAGCAA GUACGUCAACUUCCUGUACCUGGCAA GCCACUACGAAAAGCUGAAGGGGAAGC CCGGAAGACAACGAACAGAAGCAGCU GUUCGUCGAACAGCACAAAGCACUACC UGGACGAAAUCAUCGAACAGAUCAAGC GAAUUCAGCAAGAGAGUCAUCCUGGC AGACGCAAACCUGGACAAGGUCCUGA GCGCAUACAACAAGCACAGAGACAAG CCGAUCAGAGAACAGGCAGAAAACAUC CAUCCACCUGUUCACACUGACAAACC UGGGAGCACCCGGCAGCAUUCAAGUAC UUCGACACAACAACUUCGACAGAAAGAG AUACACAAGCACAAAGGAAGUCCUGG ACGCAACACUGAUCCACCAGAGCAUC ACAGGACUGUACGAAACAAGAAUCGA CCUGAGCCAGCUGGGAGGAGACGGAA GCGGAAGCCC GAAGAAGAAGAGAAAG GUCGACGGAAGCCC GAAGAAGAAGAG AAAGGU CGACAGCGGAUAG	
表 3 に記載の 最小ウリジン コドンを使用 する、配列番 号 2 3 をコー ドする C a s 9 コード配列 (開始コドン または終止コ ドンはない； 融合タンパク 質コード配列 に含める際に 好適)	GACAAGAAGUACAGCAUCGGACUGGA CAUCGGAAACAAACAGCGUCGGAUGGG CAGUCAUCACAGACGAAUACAAGGUC CCGAGCAAGAAGUUCAAGGUCCUGGG AAACACAGACAGACACAGCAUCAAGA AGAACCUGAUUCGGAGCACUGCUGUUC GACAGCGGAGAAACAGCAGAAGCAAC AAGACUGAAGAGAACAGCAAGAAGAA GAUACACAAGAAGAAAGAACAGAAUC UGC UACCUGCAGGAAAUCUUCAGCAA CGAAAUGGCAAAGGU CGACGACAGCU UCU UCCACAGACUGGAAGAAAGCUUC CUGGU CGAAGAAGACAAGAAGCACGA AAGACACCCGAUCUUCGGAAACAUCG UCGACGAAGUCGCAUACCACGAAAAG	2 4

10

20

30

40

50

	UACCCGACAAUCUACCACCUGAGAAA GAAGCUGGUCGACAGCACAGACAAGG CAGACCUGAGACUGAUCUACCUGGCA CUGGCACACAUGAUCAAGUUCAGAGG ACACUUCCUGAUCGAAGGAGACCUGA ACCCGGACAACAGCGACGUCGACAAG CUGUUCAUCCAGCUGGUCCAGACAU CAACCAGCUGUUCGAAGAAAACCCGA UCAACGCAAGCGGAGUCGACGCAAAG GCAAUCCUGAGCGCAAGACUGAGCAA GAGCAGAAGACUGGAAAACCCUGAUCG CACAGCUGCCGGGAGAAAAGAAGAAC GGACUGUUCGGAAACCUGAUCGCACU GAGCCUGGGACUGACACCGAACUUCA AGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACAC AUACGACGACGACCUGGACAACCUGC UGGCACAGAUCCGGAGACCAGUACGCA GACCUGUUCUGGCAGCAAAGAACCU GAGCGACGCAAUCCUGCUGAGCGACA UCCUGAGAGUCAACACAGAAAUCACA AAGGCACCGCUGAGCGCAAGCAUGAU CAAGAGAUACGACGAACACCACCAGG ACCUGACACUGCUGAAGGCACUGGUC AGACAGCAGCUGCCGGAAAAGUACAA GGAAAUCUUCUUCGACCAGAGCAAGA ACGGAUACGCAGGAUACAUCGACGGA GGAGCAAGCCAGGAAGAAUUCUACAA GUUCAUCAAGCCGAUCCUGGAAAAGA UGGACGGAAACAGAAGAACUGCUGGUC AAGCUGAACAGAGAAGACCUGCUGAG AAAGCAGAGAACAUUCGACAACGGAA GCAUCCCGCACCCAGAUCCACCUGGGA GAACUGCACGCAAUCCUGAGAAGACA GGAAGACUUCUACCCGUUCCUGAAGG ACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUUCAGAAUCCCGUACUACGU CGGACCGCUGGCAAGAGGAAACAGCA GAUUCGCAUGGAUGACAAGAAAGAGC GAAGAAACAAUCACACCGUGGAACUU CGAAGAAGUCGUCGACAAGGGAGCAA GCGCACAGAGCUUCAUCGAAAGAAUG	

10

20

30

40

50

50

50

	AACGAACUGGCACUGCCGAGCAAGUA CGUCAACUUCUGUACCUGGCAAGCC ACUACGAAAAGCUGAAGGGAAGCCCG GAAGACAACGAACAGAAGCAGCUGUU CGUCGAACAGCACAAAGCACUACCUGG ACGAAAUCAUCGAACAGAUAGCGAA UUCAGCAAGAGAGUCAUCCUGGCAGA CGCAAACCUGGACAAGGUCCUGAGCG CAUACAACAAGCACAGAGACAAGCCG AUCAGAGAACAGGCAGAAAACAUCAU CCACCUGUUCACACUGACAAACCUGG GAGCACCGGCAGCAUUCAAGUACUUC GACACAACAUAUCGACAGAAAGAGAU CACAAGCACAAAGGAAGUCCUGGACG CAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGACCU GAGCCAGCUGGGAGGAGACGGAAGCG GAAGCCCGAAGAAGAAGAGAAAGGUC GACGGAAGCCCGAAGAAGAAGAGAAA GGUCGACAGCGGA		10
			20
C末端のアミノ酸として2つの核局在化シグナルを有する、Cas9ニッカーゼのアミノ酸配列	MDKKYSIGLAIGTNSVGWAVITDEYK VPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFHRLEES FLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAE DAKLQLSKDTYDDDLNLLAQIGDQY ADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKAL VRQQLPKEYKEIFFDQSKNGYAGYID GGASQEEFYKFIPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHL GELHAILLRRQEDFYFPFLKDNREKIEK ILTFRIPYYVGPLARGNSRFAWMTRK SEETITPWNFEEVVDKGASAQSFIER MTNFDKNLPNEKVLPHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKK	25	30
			40

	A I V D L L F K T N R K V T V K Q L K E D Y F K K I E C F D S V E I S G V E D R F N A S L G T Y H D L L K I I K D K D F L D N E E N E D I L E D I V L T L T L F E D R E M I E E R L K T Y A H L F D D K V M K Q L K R R R Y T G W G R L S R K L I N G I R D K Q S G K T I L D F L K S D G F A N R N F M Q L I H D D S L T F K E D I Q K A Q V S G Q G D S L H E H I A N L A G S P A I K K G I L Q T V K V V D E L V K V M G R H K P E N I V I E M A R E N Q T T Q K G Q K N S R E R M K R I E E G I K E L G S Q I L K E H P V E N T Q L Q N E K L Y L Y Y L Q N G R D M Y V D Q E L D I N R L S D Y D V D H I V P Q S F L K D D S I D N K V L T R S D K N R G K S D N V P S E E V V K K M K N Y W R Q L L N A K L I T Q R K F D N L T K A E R G G L S E L D K A G F I K R Q L V E T R Q I T K H V A Q I L D S R M N T K Y D E N D K L I R E V K V I T L K S K L V S D F R K D F Q F Y K V R E I N N Y H H A H D A Y L N A V V G T A L I K K Y P K L E S E F V Y G D Y K V Y D V R K M I A K S E Q E I G K A T A K Y F F Y S N I M N F F K T E I T L A N G E I R K R P L I E T N G E T G E I V W D K G R D F A T V R K V L S M P Q V N I V K K T E V Q T G G F S K E S I L P K R N S D K L I A R K K D W D P K K Y G G F D S P T V A Y S V L V V A K V E K G K S K K L K S V K E L L G I T I M E R S S F E K N P I D F L E A K G Y K E V K K D L I I K L P K Y S L F E L E N G R K R M L A S A G E L Q K G N E L A L P S K Y V N F L Y L A S H Y E K L K G S P E D N E Q K Q L F V E Q H K H Y L D E I I E Q I S E F S K R V I L A D A N L D K V L S A Y N K H R D K P I R E Q A E N I I H L F T L T N L G A P A A F K Y F D T T I D R K R Y T S T K E V L D A T L I H Q S I T G L Y E T R I D L S Q L G G D G S G S P K K K R K V D G S P K K K R K V D S G	
開始コドン及び終止コドン を有し、表3 に記載の最小 ウリジンコド ンを使用す る、配列番号 25をコード	A U G G A C A A G A A G U A C A G C A U C G G A C U G G C A A U C G G A A C A A A C A G C G U C G G A U G G G C A G U C A U C A C A G A C G A A U A C A A G G U C C C G A G C A A G A A G U U C A A G G U C C U G G G A A A C A C A G A C A G A C A C A G C A U C A A G A A G A A C C U G A U C G G A G C A C U G C U G U U C G A C A G C G G A G A A A C A G C A G A A G C A A C A A G A C U G A A G A G A A C A G C A A G A A	26

10

20

30

40

50

するC a s 9 ニッカーゼm RNA OR F	GAAGAUACACAAGAAGAAAGAACAGA	
	AUCUGCUACCUUGCAGGAAAUCUUCAG	
	CAACGAAAUGGCAAAGGUCGACGACA	
	GCUUCUUCCACAGACUGGAAGAAAGC	
	UUCCUGGUCGAAGAAGACAAGAAGCA	
	CGAAAGACACCCGAUCUUCGGAAACA	
	UCGUCGACGAAGUCGCAUACCACGAA	
	AAGUACCCGACAAUCUACCACCUAGAG	
	AAAGAAGCUGGUCGACAGCACAGACA	10
	AGGCAGACCUAGAGACUGAUCUACCUG	
	GCACUGGCACACAUGAUCAAGUUCAG	
	AGGACACUUCUUGAUCGAAGGAGACC	
	UGAACCCGGACAACAGCGACGUCGAC	
	AAGCUGUUCAUCCAGCUGGUCCAGAC	
	AUACAACCAGCUGUUCGAAGAAAACC	
	CGAUCAACGCAAGCGGAGUCGACGCA	
	AAGGCAAUCCUGAGCGCAAGACUGAG	
	CAAGAGCAGAAGACUGGAAAACCUGA	
	UCGCACAGCUGCCGGGAGAAAAGAAG	
	AACGGACUGUUCGGAAACCUGAUCGC	20
	ACUGAGCCUGGGACUGACACCGAACU	
	UCAAGAGCAACUUCGACCUGGCAGAA	
	GACGCAAAGCUGCAGCUGAGCAAGGA	
	CACAUACGACGACGACCUGGACAACC	
	UGCUGGCACAGAUCGGAGACCAGUAC	
	GCAGACCUGUUCUUGGCAGCAAAGAA	
	CCUGAGCGACGCAAUCCUGCUGAGCG	
	ACAUCCUGAGAGUCAACACAGAAAUC	
	ACAAAGGCACCGCUGAGCGCAAGCAU	
	GAUCAAGAGAUACGACGAACACCACC	
	AGGACCUGACACUGCUGAAGGCACUG	30
	GUCAGACAGCAGCUGCCGGAAAAGUA	
	CAAGGAAAUCUUCUUCGACCAGAGCA	
	AGAACGGAUACGCAGGAUACAUCGAC	
	GGAGGAGCAAGCCAGGAAGAAUUCUA	
	CAAGUUCAUCAAGCCGAUCCUGGAAA	
	AGAUGGACGGAACAGAAGAACUGCUG	
	GUCAAGCUGAACAGAGAAGACCUGCU	
	GAGAAAGCAGAGAACAUUCGACAACG	
	GAAGCAUCCCGCACCCAGAUCCACCUG	
	GGAGAACUGCACGCAAUCCUGAGAAG	
	ACAGGAAGACUUCUACCCGUUCCUGA	40

AGGACAACAGAGAAAAGAUUCGAAAAG AUCCUGACAUUCAGAAUCCCGUACUA CGUCGGACCGCUGGCAAGAGGAAACA GCAGAUUCGCAUGGAUGACAAGAAAG AGCGAAGAAACAUAUCACACCGUGGAA CUUCGAAGAAGUCGUCGACAAGGGAG CAAGCGCACAGAGCUUCAUCGAAAGA AUGACAAACUUCGACAAGAACCUGCC GAACGAAAAGGUCCUGCCGAAGCACA GCCUGCUGUACGAAUACUUCACAGUC UACAACGAACUGACAAAGGUCAAGUA CGUCACAGAAGGAAUGAGAAAGCCGG CAUUCCUGAGCGGAGAACAGAAGAAG GCAAUCGUCGACCUGCUGUUCAAGAC AAACAGAAAGGUACAGUCAAGCAGC UGAAGGAAGACUACUUCAAGAAGAUC GAAUGCUUCGACAGCGUCGAAAUCAG CGGAGUCGAAGACAGAUUCAACGCAA GCCUGGGAACAUAACACGACCUGCUG AAGAUCAUCAAGGACAAGGACUUCU GGACAACGAAGAAAACGAAGACAUC UGGAAGACAUCGUCCUGACACUGACA CUGUUCGAAGACAGAGAAAUGAUCGA AGAAAGACUGAAGACAUAACGCACACC UGUUCGACGACAAGGUCAUGAAGCAG CUGAAGAGAAGAAGAUACACAGGAUG GGGAAGACUGAGCAGAAAGCUGAUCA ACGGAAUCAGAGACAAGCAGAGCGGA AAGACAAUCCUGGACUUCUGAAGAG CGACGGAUUCGCAAACAGAAACUUCA UGCAGCUGAUCCACGACGACAGCCUG ACAUUCAAGGAAGACAUCCAGAAGGC ACAGGUCAGCGGACAGGGAGACAGCC UGCACGAACACAUCGCAAACCUGGCA GGAAGCCCGGCAAUCAAGAAGGGAAU CCUGCAGACAGUCAAGGUUCGUCGACG AACUGGUCAAGGUCAUGGGAAGACAC AAGCCGGAAAACAUCGUCAUCGAAAU GGCAAGAGAAAACCAGACAACACAGA AGGGACAGAAGAACAGCAGAGAAAGA AUGAAGAGAAUCGAAGAAGGAAUCAA GGAACUGGGAAGCCAGAUCUGAAGG	

10

20

30

40

50

	AACACCCGGUCGAAAACACACAGCUG CAGAACGAAAAGCUGUACCUGUACUA CCUGCAGAACGGAAGAGACAUGUACG UCGACCAGGAACUGGACAUAACAGA CUGAGCGACUACGACGUCGACCACAU CGUCCCGCAGAGCUUCCUGAAGGACG ACAGCAUCGACAACAAGGUCCUGACA AGAAGCGACAAGAACAGAGGAAAGAG CGACAACGUCCCGAGCGAAGAAGUCG UCAAGAAGAUGAAGAACUACUGGAGA CAGCUGCUGAACGCAAAGCUGAUCAC ACAGAGAAAGUUCGACAACCUGACAA AGGCAGAGAGAGGAGGACUGAGCGAA CUGGACAAGGCAGGAUUCAUCAAGAG ACAGCUGGUCGAAAACAAGACAGAUCA CAAAGCACGUCGACACAGAUCCUGGAC AGCAGAAUGAACACAAAGUACGACGA AAACGACAAGCUGAUCAGAGAAGUCA AGGUCAUCACACUGAAGAGCAAGCUG GUCAGCGACUUCAGAAAGGACUUCCA GUUCUACAAGGUCAGAGAAAUCAACA ACUACCACCACGCACACGACGCAUAC CUGAACGCAGUCGUCGGAACAGCACU GAUCAAGAAGUACCCGAAGCUGGAAA GCGAAUUCGUCUACGGAGACUACAAG GUCUACGACGUCAGAAAGAUGAUCGC AAAGAGCGAACAGGAAAUCGGAAAGG CAACAGCAAAGUACUUCUUCUACAGC AACAUCAUGAACUUCUUCAAGACAGA AAUCACACUGGCAAACGGAGAAAUCA GAAAGAGACCGCUGAUCGAAACAAAC GGAGAAACAGGAGAAAUCGUCUGGGA CAAGGGAAGAGACUUCGCAACAGUCA GAAAGGUCCUGAGCAUGCCGCAGGUC AACAUCGUCAAGAAGACAGAAGUCCA GACAGGAGGAUUCAGCAAGGAAAGCA UCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGA CCCGAAGAAGUACGGAGGAUUCGACA GCCCGACAGUCGCAUACAGCGUCCUG GUCGUCGCAAAGGUCGAAAAGGGAAA GAGCAAGAAGCUGAAGAGCGUCAAGG	

10

20

30

40

50

	AACUGCUGGGAAUCACAAUCAUGGAA AGAAGCAGCUUCGAAAAGAACCCGAU CGACUUCCUGGAAGCAAAGGGAUACA AGGAAGUCAAGAAGGACCUGAUCAUC AAGCUGCCGAAGUACAGCCUGUUCGA ACUGGAAAACGGAAGAAAGAGAAUGC UGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAA GUACGUCAACUUCUGUACCUGGCAA GCCACUACGAAAAGCUGAAGGGGAAGC CCGGAAGACAACGAACAGAAGCAGCU GUUCGUCGAACAGCACAAAGCACUACC UGGACGAAAUCAUCGAACAGAUCAGC GAAUUCAGCAAGAGAGUCAUCCUGGC AGACGCAAACCUGGACAAGGUCCUGA GCGCAUACAACAAGCACAGAGACAAG CCGAUCAGAGAACAGGCAGAAAACAU CAUCCACCUGUUCACACUGACAAACC UGGGAGCACCCGGCAGCAUUCAAGUAC UUCGACACAACAUAUCGACAGAAAGAG AUACACAAGCACAAAGGAAGUCCUGG ACGCAACACUGAUCCACCAGAGCAUC ACAGGACUGUACGAAACAAGAAUCGA CCUGAGCCAGCUGGGAGGAGACGGAA GCGGAAGCCCGAAGAAGAAGAGAAAG GUCGACGGAAGCCCGAAGAAGAAGAG AAAGGUUCGACAGCGGAUAG		10
			20
表 3 に記載の 最小ウリジン コドンを使用 する、配列番 号 25 をコー ドする Cas 9 ニッカーゼ コード配列 (開始コドン または終止コ ドンはない； 融合タンパク 質コード配列 に含める際に 好適)	GACAAGAAGUACAGCAUCGGACUGGC AAUCGGAACAAACAGCGUCGGAUGGG CAGUCAUCACAGACGAAUACAAGGUC CCGAGCAAGAAGUUCAAGGUCCUGGG AAACACAGACAGACACAGCAUCAAGA AGAACCUGAUUCGGAGCACUGCUGUUC GACAGCGGAGAAACAGCAGAAGCAAC AAGACUGAAGAGAAACAGCAAGAAGAA GAUACACAAGAAGAAAGAACAGAAUC UGC UACCUGCAGGAAAUCUUCAGCAA CGAAAUGGCAAAGGUUCGACGACAGCU UCUUCACAGACUGGAAGAAAGCUUC CUGGUUCGAAGAAGACAAGAAGCACGA AAGACACCCGAUCUUCGGAAACAUCG UCGACGAAGUCGCAUACCACGAAAAG	27	30
			40

UACCCGACAAUCUACCACCUGAGAAA GAAGCUGGUCGACAGCACAGACAAGG CAGACCUGAGACUGAUCUACCUGGCA CUGGCACACAUGAUCAAGUUCAGAGG ACACUUCCUGAUCGAAGGAGACCUGA ACCCGGACAACAGCGACGUCGACAAG CUGUUCAUCCAGCUGGUCCAGACAUA CAACCAGCUGUUCGAAGAAAACCCGA UCAACGCAAGCGGAGUCGACGCAAAG GCAAUCCUGAGCGCAAGACUGAGCAA GAGCAGAAGACUGGAAAACCUGAUCG CACAGCUGCCGGGAGAAAAGAAGAAC GGACUGUUCGGAAACCUGAUCGCACU GAGCCUGGGACUGACACCGAACUUCA AGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACAC AUACGACGACGACCUGGACAACCUGC UGGCACAGAUUCGGAGACCAGUACGCA GACCUGUUCUGGCAGCAAAGAACCU GAGCGACGCAAUCCUGCUGAGCGACA UCCUGAGAGUCAACACAGAAAUCACA AAGGCACCGCUGAGCGCAAGCAUGAU CAAGAGAUACGACGAACACCACCAGG ACCUGACACUGCUGAAGGCACUGGUC AGACAGCAGCUGCCGGAAAAGUACAA GGAAAUCUUCUUCGACCAGAGCAAGA ACGGAUACGCAGGAUACAUCGACGGA GGAGCAAGCCAGGAAGAAUUCUACAA GUUCAUCAAGCCGAUCCUGGAAAAGA UGGACGGAAACAGAAGAACUGCUGGUC AAGCUGAACAGAGAAGACCUGCUGAG AAAGCAGAGAACAUUCGACAACGGAA GCAUCCCGCACCCAGAUCCACCUGGGA GAACUGCACGCAAUCCUGAGAAGACA GGAAGACUUCUACCCGUUCCUGAAGG ACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUUCAGAAUCCCGUACUACGU CGGACCGCUGGCAAGAGGAAACAGCA GAUUCGCAUGGAUGACAAGAAAGAGC GAAGAAACAAUCACACCGUGGAACUU CGAAGAAGUCGUCGACAAGGGAGCAA GCGCACAGAGCUUCAUCGAAAGAAUG	

10

20

30

40

50

	ACAAACUUCGACAAGAACCUGCCGAA CGAAAAGGUCCUGCCGAAGCACAGCC UGCUGUACGAAUACUUCACAGUCUAC AACGAACUGACAAAGGUCAAGUACGU CACAGAAGGAAUGAGAAAGCCGGCAU UCCUGAGCGGAGAACAGAAGAAGGCA AUCGUCGACCUGCUGUUCAAGACAAA CAGAAAGGUACAGUCAAGCAGCUGA AGGAAGACUACUUCAAGAAGAU CGAA UGC UUCGACAGCGUCGAAAUCAGCGG AGUCGAAGACAGAUUCAACGCAAGCC UGGGAACAUAACCAACGACCUGCUGAAG AUCAUCAAGGACAAGGACU UCCUGGA CAACGAAGAAAACGAAGACA UCCUGG AAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAU CGAAGA AAGACUGAAGACAUAACGCACACCUGU UCGACGACAAGGUCAUGAAGCAGCUG AAGAGAAGAAGAUACACAGGAUGGGG AAGACUGAGCAGAAAGCUGAUCAACG GAAUCAGAGACAAGCAGAGCGGAAAG ACAAUCCUGGACU UCCUGAAGAGCGA CGGAUUCGCAAACAGAAACUUCAUGC AGCUGAUCCACGACGACAGCCUGACA UUCAAGGAAGACA UCCAGAAGGCACA GGUCAGCGGACAGGGAGACAGCCUGC ACGAACACAUCGCAAACCUGGCAGGA AGCCCGGCAAUCAAGAAGGGA AUCCU GCAGACAGUCAAGGU CGUCGACGAAC UGGUCAAGGUCAUGGGAAGACACAAG CCGGAAAACAUCGUCAUCGAAAUGGC AAGAGAAAACCAGACAACACAGAAGG GACAGAAGAACAGCAGAGAAAGAAUG AAGAGAAUCGAAGAAGGAAUCAAGGA ACUGGGAAGCCAGAUCCUGAAGGAAC ACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCU GCAGAACGGAAGAGACAUGUACGUCG ACCAGGAACUGGACAUCAACAGACUG AGCGACUACGACGUCGACCACAUCGU CCCGCAGAGCU UCCUGAAGGACGACA GCAUCGACAACAAGGUCCUGACAAGA	
		10
		20
		30
		40

10

20

30

40

50

AGCGACAAGAACAGAGGAAAGAGCGA CAACGUCCCGAGCGAAGAAGUCGUCA AGAAGAUGAAGAACUACUGGAGACAG CUGCUGAACGC AAAGCUGAUCACACA GAGAAAGUUCGACAACCUGACAAAGG CAGAGAGAGGAGGACUGAGCGAACUG GACAAGGCAGGAUUCAUCAAGAGACA GCUGGUCGAAACAAGACAGAUCACAA AGCACGU CGCACAGAUCCUGGACAGC AGAAUGAACACAAAGUACGACGAAAA CGACAAGCUGAUCAGAGAAGUCAAGG UCAUCACACUGAAGAGCAAGCUGGUC AGCGACUUCAGAAAGGACUUCCAGUU CUACAAGGUCAGAGAAAUCAACAACU ACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAU CAAGAAGUACCCGAAGCUGGAAAGCG AAUUCGUCUACGGAGACUACAAGGUC UACGACGUCAGAAAGAUGAUCGCAAA GAGCGAACAGGAAAUCGGAAAGGCAA CAGCAAAGUACUUCUUCUACAGCAAC AUCAUGAACUUCUUCAAGACAGAAAU CACACUGGCAAACGGAGAAAUCAGAA AGAGACCGCUGAUCGAAACAAACGGA GAAACAGGAGAAAUCGUCUGGGACAA GGGAAGAGACUUCGCAACAGUCAGAA AGGUCCUGAGCAUGCCGCAGGUCAAC AUCGUCAAGAAGACAGAAGUCCAGAC AGGAGGAUUCAGCAAGGAAAGCAUCC UGCCGAAGAGAAAACAGCGACAAGCUG AUCGCAAGAAAGAAGGACUGGGACCC GAAGAAGUACGGAGGAUUCGACAGCC CGACAGUCGCAUACAGCGUCCUGGUC GUCGCAAAGGUCGAAAAGGGAAAGAG CAAGAAGCUGAAGAGCGUCAAGGAAC UGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGA CUUCCUGGAAGCAAAGGGAUACAAGG AAGUCAAGAAGGACCUGAUCAUCAAG CUGCCGAAGUACAGCCUGUUCGAACU GGAAAACGGAAGAAAGAGAAUUGCUGG CAAGCGCAGGAGAACUGCAGAAGGGA	

10

20

30

40

50

	AACGAACUGGCACUGCCGAGCAAGUA CGUCAACUCCUGUACCUGGCAAGCC ACUACGAAAAGCUGAAGGGAAGCCCG GAAGACAACGAACAGAAGCAGCUGUU CGUCGAACAGCACAAAGCACUACCUGG ACGAAAUCAUCGAACAGAUAGCGAA UUCAGCAAGAGAGUCAUCCUGGCAGA CGCAAACCUGGACAAGGUCCUGAGCG CAUACAACAAGCACAGAGACAAGCCG AUCAGAGAACAGGCAGAAAACAUCAU CCACCUGUUCACACUGACAAACCUGG GAGCACCGGCAGCAUUCAAGUACUUC GACACAACAUAUCGACAGAAAGAGAU CACAAGCACAAAGGAAGUCCUGGACG CAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGACCU GAGCCAGCUGGGAGGAGAC GGAAGC GGAAGCCCGAAGAAGAAGAGAAAGGU CGACGGAAGCCCGAAGAAGAAGAGAA AGGUUCGACAGCGGA	
C末端のアミノ酸として2つの核局在化シグナルを有する、dCas9のアミノ酸配列	MDKKYSIGLAIGTNSVGWAVITDEYK VPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFHRLEES FLVEEDKKHERHPIFGNIVDEVAYHE KYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAE DAKLQLSKDTYDDDLNLLAQIGDQY ADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKAL VRQQLPKEYKEIFFDQSKNGYAGYID GGASQEEFYKFIPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHL GELHAILRRQEDFYFPLKDNREKIEK ILTFRIPIYYVGPLARGNSRFAWMTRK SEETITPWNFEEVVDKGASAQSFIER MTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKI	28

10

20

30

40

50

40

9 mRNA ORF	AUCUGCUACCUUGCAGGAAAUCUUCAG	
	CAACGAAAUGGCAAAGGUCGACGACA	
	GCUUCUUCACACAGACUGGAAGAAAGC	
	UUCUGGUCGAAGAAGACAAGAAGCA	
	CGAAAGACACCCGAUCUUCGGAAACA	
	UCGUCGACGAAGUCGCAUACCACGAA	
	AAGUACCCGACAAUCUACCACCUGAG	
	AAAGAAGCUGGUCGACAGCACAGACA	
	AGGCAGACCUGAGACUGAUCUACCUG	
	GCACUGGGCACACAUGAUCAAGUUCAG	
	AGGACACUUCUGAUCGAAGGAGACC	
	UGAACCCGGACAACAGCGACGUCGAC	
	AAGCUGUUCAUCCAGCUGGUCCAGAC	
	AUACAACCAGCUGUUCGAAGAAAACC	
	CGAUCAACGCAAGCGGAGUCGACGCA	
	AAGGCAAUCCUGAGCGCAAGACUGAG	
	CAAGAGCAGAAGACUGGAAAACCUGA	
	UCGCACAGCUGCCGGGAGAAAAGAAG	
	AACGGACUGUUCGGAAACCUGAUCGC	
	ACUGAGCCUGGGACUGACACCGAACU	
	UCAAGAGCAACUUCGACCUGGCAGAA	
	GACGCAAAGCUGCAGCUGAGCAAGGA	
	CACAUACGACGACGACCUGGACAACC	
	UGCUGGCACAGAUCCGAGACCAGUAC	
	GCAGACCUGUUCUGGCAGCAAAGAA	
	CCUGAGCGACGCAAUCCUGCUGAGCG	
	ACAUCCUGAGAGUCAACACAGAAAUC	
	ACAAAGGCACCGCUGAGCGCAAGCAU	
	GAUCAAGAGAUACGACGAACACCACC	
	AGGACCUGACACUGCUGAAGGCACUG	
	GUCAGACAGCAGCUGCCGGAAAAGUA	
	CAAGGAAAUCUUCUUCGACCAGAGCA	
	AGAACGGAUACGCAGGAUACAUCGAC	
	GGAGGAGCAAGCCAGGAAGAAUUCUA	
	CAAGUUCAUCAAGCCGAUCCUGGAAA	
	AGAUGGACGGAACAGAAGAACUGCUG	
	GUCAAGCUGAACAGAGAAGACCUGCU	
	GAGAAAGCAGAGAAACAUCGACAACG	
	GAAGCAUCCCGCACCAAGAUCCACCUG	
	GGAGAACUGCACGCAAUCCUGAGAAG	
	ACAGGAAGACUUCUACCCGUUCCUGA	
	AGGACAACAGAGAAAAGAUCGAAAAG	

10

20

30

40

50

	A U C C U G A C A U U C A G A A U C C C G U A C U A C G U C G G A C C G C U G G C A A G A G G A A A C A G C A G A U U C G C A U G G A U G A C A A G A A A G A G C G A A G A A A C A A U C A C A C C G U G G A A C U U C G A A G A A G U C G U C G A C A A G G G A G C A A G C G C A C A G A G C U U C A U C G A A A G A A U G A C A A A C U U C G A C A A G A A C C U G C C G A A C G A A A A G G U C C U G C C G A A G C A C A G C C U G C U G U A C G A A U A C U U C A C A G U C U A C A A C G A A C U G A C A A A G G U C A A G U A C G U C A C A G A A G G A A U G A G A A A G C C G G C A U U C C U G A G C G G A G A A C A G A A G A A G G C A A U C G U C G A C C U G C U G U U C A A G A C A A A C A G A A A G G U C A C A G U C A A G C A G C U G A A G G A A G A C U A C U U C A A G A A G A U C G A A U G C U U C G A C A G C G U C G A A A U C A G C G G A G U C G A A G A C A G A U U C A A C G C A A G C C U G G G A A C A U A C C A C G A C C U G C U G A A G A U C A U C A A G G A C A A G G A C U U C C U G G A C A A C G A A G A A A A C G A A G A C A U C C U G G A A G A C A U C G U C C U G A C A C U G A C A C U G U U C G A A G A C A G A G A A A U G A U C G A A G A A A G A C U G A A G A C A U A C G C A C A C C U G U U C G A C G A C A A G G U C A U G A A G C A G C U G A A G A G A A G A A G A U A C A C A G G A U G G G G A A G A C U G A G C A G A A A G C U G A U C A A C G G A A U C A G A G A C A A G C A G A G C G G A A A G A C A A U C C U G G A C U U C C U G A A G A G C G A C G G A U U C G C A A A C A G A A A C U U C A U G C A G C U G A U C C A C G A C G A C A G C C U G A C A U U C A A G G A A G A C A U C C A G A A G G C A C A G G U C A G C G G A C A G G G A G A C A G C C U G C A C G A A C A C A U C G C A A A C C U G G C A G G A A G C C C G G C A A U C A A G A A G G G A A U C C U G C A G A C A G U C A A G G U C G U C G A C G A A C U G G U C A A G G U C A U G G G A A G A C A C A A G C C G G A A A A C A U C G U C A U C G A A A U G G C A A G A G A A A A C C A G A C A A C A C A G A A G G G A C A G A A G A A C A G C A G A G A A A G A A U G A A G A G A A U C G A A G A A G G A A U C A A G G A A C U G G G A A G C C A G A U C C U G A A A G A A C A C C C G G U C G A A A A C A C A C A G C U G	

50

	AGAAAGCAGCUUUCGAAAAGAACCCGAU CGACUUCCUGGAAGCAAAGGGGAUACA AGGAAGUCAAGAAGGACCUGAUCAUC AAGCUGCCGAAGUACAGCCUGUUCGA ACUGGAAAACGGAAGAAAGAGAAUGC UGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAA GUACGUCAACUUCCUGUACCUGGCAA GCCACUACGAAAAGCUGAAGGGGAAGC CCGGAAGACAACGAACAGAAGCAGCU GUUCGUCGAACAGCACAAAGCACUACC UGGACGAAAUCAUCGAACAGAUCAGC GAAUUCAGCAAGAGAGUCAUCCUGGC AGACGCAAACCUGGACAAGGUCCUGA GCGCAUACAACAAGCACAGAGACAAG CCGAUCAGAGAACAGGCAGAAAACA CAUCCACCUGUUCACACUGACAAACC UGGGAGCACCCGGCAGCAUUCAAGUAC UUCGACACAACAUAUCGACAGAAAGAG AUACACAAGCACAAAGGAAGUCCUGG ACGCAACACUGAUCCACCAGAGCAUC ACAGGACUGUACGAAACAAGAAUCGA CCUGAGCCAGCUGGGAGGAGAC GGA AGCGGAAGCCCGAAGAAGAAGAGAAA GGUCGACGGAAGCCCGAAGAAGAAGA GAAAGGUCGACAGCGGAUAG	10
表3に記載の 最小ウリジン コドンを使用 する、配列番 号28をコー ドするdCa s9コード配 列（開始コ ドンまたは終 止コドン はない；融 合タンパク 質コード配 列に含める 際に好適）	GACAAGAAGUACAGCAUCGGACUGGC AAUCGGAAACAAACAGCGUCGGAUGGG CAGUCAUCACAGACGAAUACAAGGUC CCGAGCAAGAAGUUCAAGGUCCUGGG AAACACAGACAGACACAGCAUCAAGA AGAACCCUGAUUCGGAGCACUGCUGUUC GACAGCGGAGAAACAGCAGAAGCAAC AAGACUGAAGAGAAACAGCAAGAAGAA GAUACACAAGAAGAAAGAACAGAAUC UGCUACCUGCAGGAAAUCUUCAGCAA CGAAAUGGCAAAGGUUCGACGACAGCU UCUUCCACAGACUGGAAGAAAGCUUC CUGGUCGAAGAAGACAAGAAGCACGA AAGACACCCGAUCUUCGGAAACAUCG UCGACGAAGUCGCAUACCACGAAAAG UACCCGACAAUCUACCACCUGAGAAA	30
		40

	GAAGCUGGUCGACAGCACAGACAAGG CAGACCUGAGACUGAUCUACCUGGCA CUGGCACACAUGAUCAAGUUCAGAGG ACACUUCCUGAUCGAAGGAGACCUGA ACCCGGACAACAGCGACGUCGACAAG CUGUUCAUCCAGCUGGUCCAGACAU CAACCAGCUGUUCGAAGAAAACCCGA UCAACGCAAGCGGAGUCGACGCAAAG GCAAUCCUGAGCGCAAGACUGAGCAA GAGCAGAAGACUGGAAAACUGAUCG CACAGCUGCCGGGAGAAAAGAAGAAC GGACUGUUCGGAAACCUGAUCGCACU GAGCCUGGGACUGACACCGAACUUCA AGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACAC AUACGACGACGACCUGGACAACCUGC UGGCACAGAUUCGGAGACCAGUACGCA GACCUGUUCUGGCAGCAAAGAACCU GAGCGACGCAAUCCUGCUGAGCGACA UCCUGAGAGUCAACACAGAAAUCACA AAGGCACCGCUGAGCGCAAGCAUGAU CAAGAGAUACGACGAACACCACCAGG ACCUGACACUGCUGAAGGCACUGGUC AGACAGCAGCUGCCGGAAAAGUACAA GGAAAUCUUCUUCGACCAGAGCAAGA ACGGAUACGCAGGAUACAUCGACGGA GGAGCAAGCCAGGAAGAAUUCUACAA GUUCAUCAAGCCGAUCCUGGAAAAGA UGGACGGAACAGAAGAACUGCUGGUC AAGCUGAACAGAGAAGACCUGCUGAG AAAGCAGAGAAACAUUCGACAACGGAA GCAUCCCCGCACCAGAUCCACCUGGGA GAACUGCACGCAAUCCUGAGAAGACA GGAAGACUUCUACCCGUUCCUGAAGG ACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUUCAGAAUCCCGUACUACGU CGGACCGCUGGCAAGAGGAAACAGCA GAUUCGCAUGGAUGACAAGAAAGAGC GAAGAAACAAUCACACCGUGGAACUU CGAAGAAGUCGUCGACAAGGGAGCAA GCGCACAGAGCUUCAUCGAAAGAAUG ACAAACUUCGACAAGAACCUGCCGAA	

10

20

30

40

50

	CGAAAAGGUCCUGCCGAAGCACAGCC UGCUGUACGAAUACUUCACAGUCUAC AACGAACUGACAAAGGUCAAGUACGU CACAGAAGGAAUGAGAAAGCCGGCAU UCCUGAGCGGAGAACAGAAGAAGGCA AUCGUCGACCUGCUGUUCAAGACAAA CAGAAAGGUCACAGUCAAGCAGCUGA AGGAAGACUACUUCAAGAAGAUCGAA UGC UUCGACAGCGUCGAAAUCAGCGG AGUCGAAGACAGAUUCAACGCAAGCC UGGGAACAUAACCACGACCUGCUGAAG AUCAUCAAGGACAAGGACUUCCUGGA CAACGAAGAAAACGAAGACA UCCUGG AAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAUCGAAGA AAGACUGAAGACAUAACGCACACCUGU UCGACGACAAGGUCAUGAAGCAGCUG AAGAGAAGAAGAUACACAGGAUGGGG AAGACUGAGCAGAAAGCUGAUCAACG GAAUCAGAGACAAGCAGAGCGGAAAG ACAAUCCUGGACU UCCUGAAGAGCGA CGGAUUCGCAAACAGAAACUUCAUGC AGCUGAUCCACGACGACAGCCUGACA UUCAAGGAAGACA UCCAGAAGGCACA GGUCAGCGGACAGGGAGACAGCCUGC ACGAACACAUCGCAAACCUGGCAGGA AGCCCGGCAAUCAAGAAGGGAAUCCU GCAGACAGUCAAGGUUCGUCGACGAAC UGGUCAAGGUCAUGGGAAGACACAAG CCGGAAAACAUCGUCAUCGAAAUGGC AAGAGAAAACCAGACAACACAGAAGG GACAGAAGAACAGCAGAGAAAGAAUG AAGAGAAUCGAAGAAGGAAUCAAGGA ACUGGGGAAGCCAGA UCCUGAAGGAAC ACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCU GCAGAACGGAAGAGACAUGUACGUCG ACCAGGAACUGGACAUCAACAGACUG AGCGACUACGACGUCGACGCAAUCGU CCCGCAGAGCUUCCUGAAGGACGACA GCAUCGACAACAAGGUCCUGACAAGA AGCGACAAGAACAGAGGAAAGAGCGA	

10

20

30

40

50

	CAACGUCCCCGAGCGAAGAAGUCGUCA AGAAGAUGAAGAACUACUGGAGACAG CUGCUGAACGCAAAGCUGAUCACACA GAGAAAGUUCGACAACCUGACAAAGG CAGAGAGAGGAGGACUGAGCGAACUG GACAAGGCAGGAUUCAUCAAGAGACA GCUGGUUCGAAACAAGACAGAUCACAA AGCACGUCGACACAGAUCUGGACAGC AGAAUGAACACAAAGUACGACGAAAA CGACAAGCUGAUCAGAGAAGUCAAGG UCAUCACACUGAAGAGCAAGCUGGUC AGCGACUUCAGAAAGGACUUCCAGUU CUACAAGGUCAGAGAAAUCAACAACU ACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAU CAAGAAGUACCCGAAGCUGGAAAGCG AAUUCGUCUACGGAGACUACAAGGUC UACGACGUCAGAAAGAUGAUCGCAAA GAGCGAACAGGAAAUCGGAAAGGCAA CAGCAAAGUACUUCUUCUACAGCAAC AUCAUGAACUUCUUCAAGACAGAAAU CACACUGGCAAACGGAGAAAUCAGAA AGAGACCGCUGAUCGAAACAAACGGA GAAACAGGAGAAAUCGUCUGGGACAA GGGAAGAGACUUCGCAACAGUCAGAA AGGUCCUGAGCAUGCCGCAGGUCAAC AUCGUCAAGAAGACAGAAGUCCAGAC AGGAGGAUUCAGCAAGGAAAGCAUCC UGCCGAAGAGAAACAGCGACAAGCUG AUCGCAAGAAAGAAGGACUGGGACCC GAAGAAGUACGGAGGAUUCGACAGCC CGACAGUCGCAUACAGCGUCCUGGUC GUCGCAAAGGUUCGAAAAGGGAAAGAG CAAGAAGCUGAAGAGCGUCAAGGAAC UGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGA CUUCCUGGAAGCAAAGGGAUACAAGG AAGUCAAGAAGGACCUGAUCAUCAAG CUGCCGAAGUACAGCCUGUUCGAACU GGAAAACGGAAGAAAGAGAAUGCUGG CAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUA	

10

20

30

40

50

	CGUCAACUUCCUGUACCUGGCAAGCC ACUACGAAAAGCUGAAGGGAAGCCCG GAAGACAACGAACAGAAGCAGCUGUU CGUCGAACAGCACAAAGCACUACCUGG ACGAAAUCAUCGAACAGAUACAGCGAA UUCAGCAAGAGAGUCAUCCUGGCAGA CGCAAACCUGGACAAGGUCCUGAGCG CAUACAACAAGCACAGAGACAAGCCG AUCAGAGAACAGGCAGAAAACAUCAU CCACCUGUUCACACUGACAAACCUGG GAGCACCGGCAGCAUUCAAGUACUUC GACACAACAUAUCGACAGAAAGAGAU CACAAGCACAAAGGAAGUCCUGGACG CAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGACCU GAGCCAGCUGGGAGGAGAC GGAAGC GGAAGCCCGAAGAAGAAGAGAAAGGU CGACGGAAGCCCGAAGAAGAAGAGAA AGGUCGACAGCGGA	
T7プロモーター	TAATACGACTCACTATA	3 1
ヒトベータグロビン 5' UTR	ACATTTTGCTTCTGACACAACCTGTGTT CACTAGCAACCTCAAACAGACACC	3 2
ヒトベータグロビン 3' UTR	GCTCGCTTTTCTTGCTGTCCAATTTCT ATTAAAGGTTCTTTGTTCCCTAAGT CCAACTACTAAACTGGGGGATATTAT GAAGGGCCTTGAGCATCTGGATTCTG CCTAATAAAAAACATTTATTTTCATT GC	3 3
ヒトアルファグロビン 5' UTR	CATAAACCCCTGGCGCGCTCGCGGCC GGCACTCTTCTGGTCCCCACAGACTC AGAGAGAACCCACC	3 4
ヒトアルファグロビン 3' UTR	GCTGGAGCCTCGGTGGCCATGCTTCT TGCCCCCTTGGGCCTCCCCCAGCCCC TCCTCCCCCTTCCTGCACCCGTACCCC CGTGGTCTTTGAATAAAGTCTGAGTG GGCGGC	3 5
Xenopus laevis ベータグロビン 5'	AAGCTCAGAATAAACGCTCAACTTTG GCC	3 6

10

20

30

40

50

UTR		
Xenopus laevis i s ベータグロビン 3' UTR	ACCAGCCTCAAGAACACCCGAATGGA GTCTCTAAGCTACATAATACCAACTT ACACTTTTACAAAAATGTTGTCCCCCAA AATGTAGCCATTTCGTATCTGCTCCTA ATAAAAAGAAAGTTTCTTCACATTCT 37	
ウシ成長ホルモン 5' UTR	CAGGGTCCTGTGGACAGCTCACCAGC T	38
ウシ成長ホルモン 3' UTR	TTGCCAGCCATCTGTTGTTTGCCCCCT CCCCCGTGCCCTTCCTTGACCCTGGAA GGTGCCACTCCCCTGTCCTTTCTCTA ATAAAATGAGGAAATTGCATCGCA	39
Mus musculus ヘモグロビン アルファ、成 体鎖1 (Hb a - a 1)、 3' UTR	GCTGCCTTCTGCGGGGCTTGCCCTTCT GGCCATGCCCTTCTTCTCTCCCTTGCT ACCTGTACCTCTTGGTCTTTGAATAA AGCCTGAGTAGGAAG	40
HSD17B4 5' UTR	TCCCGCAGTCGGCGTCCAGCGGCTCT GCTTGTTTCGTGTGTGTGTCGTTGCAG GCCTTATTC	41
TTR指向性 G282ガイドRNA	mU*mU*mA*CAGCCACGUCUACAGCA GUUUUAGAmGmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAUAAGGCUAGUC CGUUUAUCAmAmCmUmUmGmAmAmAmA mAmGmUmGmGmCmAmCmCmGmAmGmU mCmGmGmUmGmCmU*mU*mU*mU	42
HSDの5' UTR、配列 番号4に対応 するORF、 コザック配 列、及びALB の3' UTR を有するCas9 転写産物	GGGTCCCGCAGTCGGCGTCCAGCGGC TCTGCTTGTTTCGTGTGTGTGTCGTTG CAGGCCTTATTCGGATCCGCCACCAT GGACAAGAAGTACAGCATCGGACTGG ACATCGGAACAAACAGCGTCGGATGG GCAGTCATCACAGACGAATACAAGGT CCCGAGCAAGAAGTTCAAGGTCCTGG GAAACACAGACAGACACAGCATCAAG AAGAACCTGATCGGAGCACTGCTGTT CGACAGCGGAGAAACAGCAGAAGCAA CAAGACTGAAGAGAACAGCAAGAAGA	43

10

20

30

40

50

	AGATACACAAGAAGAAGAACAGAAAT CTGCTACCTGCAGGAAATCTTTCAGCA ACGAAATGGCAAAGGTCGACGACAGC TTCTTCCACAGACTGGAAGAAAGCTT CCTGGTTCGAAGAAGACAAGAAGCACG AAAGACACCCGATCTTTCGGAAACATC GTCGACGAAGTCGCATACCACGAAAA GTACCCGACAATCTACCACCTGAGAA AGAAGCTGGTCGACAGCACAGACAAG GCAGACCTGAGACTGATCTACCTGGC ACTGGCACACATGATCAAGTTCAGAG GACACTTCTGATCGAAGGAGACCTG AACCCGGACAACAGCGACGTCGACAA GCTGTTCATCCAGCTGGTCCAGACAT ACAACCAGCTGTTCGAAGAAAACCCG ATCAACGCAAGCGGAGTCGACGCAAA GGCAATCCTGAGCGCAAGACTGAGCA AGAGCAGAAGACTGGAAAACCTGATC GCACAGCTGCCGGGAGAAAAGAAGAA CGGACTGTTCGGAAACCTGATCGCAC TGAGCCTGGGACTGACACCGAACTTC AAGAGCAACTTCGACCTGGCAGAAGA CGCAAAGCTGCAGCTGAGCAAGGACA CATACGACGACGACCTGGACAACCTG CTGGCACAGATCGGAGACCAGTACGC AGACCTGTTCCTGGCAGCAAAGAACC TGAGCGACGCAATCCTGCTGAGCGAC ATCCTGAGAGTCAACACAGAAATCAC AAAGGCACCGCTGAGCGCAAGCATGA TCAAGAGATACGACGAACACCACCAG GACCTGACACTGCTGAAGGCACTGGT CAGACAGCAGCTGCCGGAAAAGTACA AGGAAATCTTCTTCGACCAGAGCAAG AACGGATACGCAGGATACATCGACGG AGGAGCAAGCCAGGAAGAATTCTACA AGTTTCATCAAGCCGATCCTGGAAAAG ATGGACGGAACAGAAGAACTGCTGGT CAAGCTGAACAGAGAAGACCTGCTGA GAAAGCAGAGAACATTTCGACAACGGA AGCATCCCGCACCCAGATCCACCTGGG AGAACTGCACGCAATCCTGAGAAGAC AGGAAGACTTCTACCCGTTCTGAAAG	

	GACAACAGAGAAAAGATCGAAAAGAT CCTGACATTTCAGAATCCCGTACTACG TCGGACCGCTGGCAAGAGGAAACAGC AGATTTCGCATGGATGACAAGAAAGAG CGAAGAAACAATCACACCGTGGAACCT TCGAAGAAGTCGTTCGACAAGGGAGCA AGCGCACAGAGCTTCATCGAAAGAAT GACAAACTTCGACAAGAACCTGCCGA ACGAAAAGGTCCTGCCGAAGCACAGC CTGCTGTACGAATACTTCACAGTCTA CAACGAACCTGACAAAGGTCAAGTACG TCACAGAAGGAATGAGAAAGCCGGCA TTCCTGAGCGGAGAACAGAAGAAGGC AATCGTCGACCTGCTGTTCAAGACAA ACAGAAAGGTCACAGTCAAGCAGCTG AAGGAAGACTACTTCAAGAAGATCGA ATGCTTCGACAGCGTCGAAATCAGCG GAGTCGAAGACAGATTCAACGCAAGC CTGGGAACATACCACGACCTGCTGAA GATCATCAAGGACAAGGACTTCCTGG ACAACGAAGAAAACGAAGACATCCTG GAAGACATCGTCCTGACACTGACACT GTTTCGAAGACAGAGAAATGATCGAAG AAAGACTGAAGACATACGCACACCTG TTCGACGACAAGGTCATGAAGCAGCT GAAGAGAAGAAGATACACAGGATGGG GAAGACTGAGCAGAAAGCTGATCAAC GGAATCAGAGACAAGCAGAGCGGAAA GACAATCCTGGACTTCCTGAAGAGCG ACGGATTTCGCAAACAGAAACTTCATG CAGCTGATCCACGACGACAGCCTGAC ATTCAAGGAAGACATCCAGAAGGCAC AGGTCAGCGGACAGGGAGACAGCCTG CACGAACACATCGCAAACCTGGCAGG AAGCCCGGCAATCAAGAAGGGAATCC TGCAGACAGTCAAGGTCGTTCGACGAA CTGGTCAAGGTCATGGGAAGACACAA GCCGGAAAACATCGTCATCGAAATGG CAAGAGAAAACCAGACAACACAGAAG GGACAGAAGAACAGCAGAGAAAGAAT GAAGAGAATCGAAGAAGGAATCAAGG AACTGGGAAGCCAGATCCTGAAGGAA	
		10
		20
		30
		40

	CACCCGGTCGAAAACACACAGCTGCA GAACGAAAAGCTGTACCTGTACTACC TGCAGAACGGAAGAGACATGTACGTC GACCAGGAACCTGGACATCAACAGACT GAGCGACTACGACGTCGACCACATCG TCCCGCAGAGCTTCCTGAAGGACGAC AGCATCGACAACAAGGTCCTGACAAG AAGCGACAAGAACAGAGGAAAGAGCG ACAACGTCCCGAGCGAAGAAGTCGTC AAGAAGATGAAGAACTACTGGAGACA GCTGCTGAACGCAAAGCTGATCACAC AGAGAAAGTTTCGACAACCTGACAAAG GCAGAGAGAGGAGGACTGAGCGAACT GGACAAGGCAGGATTTCATCAAGAGAC AGCTGGTCGAAAACAAGACAGATCACA AAGCACGTTCGCACAGATCCTGGACAG CAGAATGAACACAAAGTACGACGAAA ACGACAAGCTGATCAGAGAAGTCAAG GTCATCACACTGAAGAGCAAGCTGGT CAGCGACTTCAGAAAGGACTTCCAGT TCTACAAGGTCAGAGAAATCAACAAC TACCACCACGCACACGACGCATACCT GAACGCAGTCGTTCGGAACAGCACTGA TCAAGAAGTACCCGAAGCTGGAAAGC GAATTTCGTCTACGGAGACTACAAGGT CTACGACGTCAGAAAGATGATCGCAA AGAGCGAACAGGAAATCGGAAAGGCA ACAGCAAAGTACTTCTTCTACAGCAA CATCATGAACTTCTTCAAGACAGAAA TCACACTGGCAAACGGAGAAATCAGA AAGAGACCGCTGATCGAAACAAACGG AGAAACAGGAGAAATCGTCTGGGACA AGGGAAGAGACTTCGCAACAGTCAGA AAGGTCCTGAGCATGCCGCAGGTCAA CATCGTCAAGAAGACAGAAGTCCAGA CAGGAGGATTTCAGCAAGGAAAGCATC CTGCCGAAGAGAAACAGCGACAAGCT GATCGCAAGAAAGAAGGACTGGGACC CGAAGAAGTACGGAGGATTTCGACAGC CCGACAGTCGCATACAGCGTCCTGGT CGTCGCAAAGGTCGAAAAGGGAAAGA GCAAGAAGCTGAAGAGCGTCAAGGAA	

10

20

30

40

50

	CTGCTGGGAATCACAATCATGGAAAG AAGCAGCTTCGAAAAGAACCCGATCG ACTTCCTGGAAGCAAAGGGATACAAG GAAGTCAAGAAGGACCTGATCATCAA GCTGCCGAAGTACAGCCTGTTTGAAC TGGAACACGGAAGAAAGAGAATGCTG GCAAGCGCAGGAGAAGTGCAGAAGGG AAACGAACTGGCACTGCCGAGCAAGT ACGTCAAACCTTCCTGTACCTGGCAAGC CACTACGAAAAGCTGAAGGGGAAGCCC GGAAGACAACGAACAGAAGCAGCTGT TCGTGGAACAGCACAAAGCACTACCTG GACGAAATCATCGAACAGATCAGCGA ATTCAGCAAGAGAGTTCATCCTGGCAG ACGCAAACCTGGACAAGGTCCTGAGC GCATACAACAAGCACAGAGACAAGCC GATCAGAGAACAGGCAGAAAACATCA TCCACCTGTTTCACTGACAAACCTG GGAGCACCGGCAGCATTCAAGTACTT CGACACAACAATCGACAGAAAGAGAT ACACAAGCACAAAGGAAGTCCTGGAC GCAACACTGATCCACCAGAGCATCAC AGGACTGTACGAAACAAGAATCGACC TGAGCCAGCTGGGAGGAGACGGAGGA GGAAGCCCCGAAGAAGAAGAGAAAGGT CTAGCTAGCCATCACATTTAAAAGCA TCTCAGCCTACCATGAGAATAAGAGA AAGAAAATGAAGATCAATAGCTTATT CATCTCTTTTTCTTTTTTCGTTGGTGT AAAGCCAACACCCTGTCTAAAAAACA TAAATTTCTTTAATCATTTTGCCTCT TTTCTCTGTGCTTCAATTAATAAAAA ATGGAAAGAACCTCGAG	
HSDの5' UTR、配列番号4に対応するORF、及びALBの3' UTRを有するCas9転写産物	GGGTCCCGCAGTCGGCGTCCAGCGGC TCTGCTTGTTTCGTGTGTGTGTCGTTG CAGGCCTTATTTCGGATCCATGGACAA GAAGTACAGCATCGGACTGGACATCG GAACAAACAGCGTCGGATGGGCAGTC ATCACAGACGAATACAAGGTCCCGAG CAAGAAGTTCAAGGTCCTGGGAAACA CAGACAGACACAGCATCAAGAAGAAC	4 4

10

20

30

40

50

CTGATCGGAGCACTGCTGTTTCGACAG CGGAGAAACAGCAGAAGCAACAAGAC TGAAGAGAACAGCAAGAAGAAGATAC ACAAGAAGAAAGAACAGAATCTGCTA CCTGCAGGAAATCTTCAGCAACGAAA TGGCAAAGGTCGACGACAGCTTCTTC CACAGACTGGAAGAAAGCTTCCTGGT CGAAGAAGACAAGAAGCACGAAAGAC ACCCGATCTTCGGAAACATCGTCGAC GAAGTCGCATACCACGAAAAGTACCC GACAATCTACCACCTGAGAAAGAAGC TGGTCGACAGCACAGACAAGGCAGAC CTGAGACTGATCTACCTGGCACTGGC ACACATGATCAAGTTCAGAGGACACT TCCTGATCGAAGGAGACCTGAACCCG GACAACAGCGACGTCGACAAGCTGTT CATCCAGCTGGTCCAGACATACAACC AGCTGTTCTGAAGAAAACCCGATCAAC GCAAGCGGAGTCGACGCAAAGGCAAT CCTGAGCGCAAGACTGAGCAAGAGCA GAAGACTGGAAAACCTGATCGCACAG CTGCCGGGAGAAAAGAAGAACGGACT GTTCTGGAAACCTGATCGCACTGAGCC TGGGACTGACACCGAACTTCAAGAGC AACTTCGACCTGGCAGAAGACGCAAA GCTGCAGCTGAGCAAGGACACATACG ACGACGACCTGGACAACCTGCTGGCA CAGATCGGAGACCAGTACGCAGACCT GTTCTGTCAGCAAAGAACCTGAGCG ACGCAATCCTGCTGAGCGACATCCTG AGAGTCAACACAGAAATCACAAAGGC ACCGCTGAGCGCAAGCATGATCAAGA GATACGACGAACACCACCAGGACCTG ACACTGCTGAAGGCACTGGTCAGACA GCAGCTGCCGGAAAAGTACAAGGAAA TCTTCTTCGACCAAGAGCAAGAACGGA TACGCAGGATACATCGACGGAGGAGC AAGCCAGGAAGAATTCTACAAGTTCA TCAAGCCGATCCTGGAAAAGATGGAC GGAACAGAGAAGAACTGCTGGTCAAGCT GAACAGAGAAGACCTGCTGAGAAAGC AGAGAACATTTCGACAACGGAAGCATC	

10

20

30

40

50

50

	AAGAACAGCAGAGAAAGAATGAAGAG AATCGAAGAAGGAATCAAGGAACTGG GAAGCCAGATCCTGAAGGAACACCCG GTCGAAAACACACAGCTGCAGAACGA AAAGCTGTACCTGTACTACCTGCAGA ACGGAAGAGACATGTACGTCTGACCAG GAACTGGACATCAACAGACTGAGCGA CTACGACGTCTGACCACATCGTCCCGC AGAGCTTCCTGAAGGACGACAGCATC GACAACAAGGTCCTGACAAGAAGCGA CAAGAACAGAGGAAAGAGCGACAACG TCCCGAGCGAAGAAGTCGTCAAGAAG ATGAAGAACTACTGGAGACAGCTGCT GAACGCAAAGCTGATCACACAGAGAA AGTTTCGACAACCTGACAAAGGCAGAG AGAGGAGGACTGAGCGAACTGGACAA GGCAGGATTTCATCAAGAGACAGCTGG TCGAAACAAGACAGATCACAAAGCAC GTCGCACAGATCCTGGACAGCAGAAT GAACACAAAGTACGACGAAAACGACA AGCTGATCAGAGAAGTCAAGGTCATC ACACTGAAGAGCAAGCTGGTCAGCGA CTTCAGAAAGGACTTCCAGTTCTACA AGGTCAGAGAAATCAACAACCTACCAC CACGCACACGACGCATACCTGAACGC AGTCGTCTGGAACAGCACTGATCAAGA AGTACCCGAAGCTGGAAAGCGAATTCT GTCTACGGAGACTACAAGGTCTACGA CGTCAGAAAGATGATCGCAAAGAGCG AACAGGAAATCGGAAAGGCAACAGCA AAGTACTTCTTCTACAGCAACATCAT GAACTTCTTCAAGACAGAAATCACAC TGGCAAACGGAGAAATCAGAAAGAGA CCGCTGATCGAAACAAACGGAGAAAC AGGAGAAATCGTCTGGGACAAGGGAA GAGACTTCGCAACAGTCAGAAAGGTC CTGAGCATGCCGCAGGTCAACATCGT CAAGAAGACAGAAGTCCAGACAGGAG GATTCAGCAAGGAAAGCATCCTGCCG AAGAGAAACAGCGACAAGCTGATCGC AAGAAAGAAGGACTGGGACCCGAAGA AGTACGGAGGATTTCGACAGCCCGACA	

10

20

30

40

50

	GTCGCATACAGCGTCCTGGTCGTCGC AAAGGTCGAAAAGGGAAAGAGCAAGA AGCTGAAGAGCGTCAAGGAACTGCTG GGAATCACAAATCATGGAAAGAAGCAG CTTCGAAAAGAACCCGATCGACTTCC TGGAAGCAAAGGGATACAAGGAAGTC AAGAAGGACCTGATCATCAAGCTGCC GAAGTACAGCCTGTTTCGAACTGGAAA ACGGAAGAAAGAGAATGCTGGCAAGC GCAGGAGAACTGCAGAAGGGAAACGA ACTGGCACTGCCGAGCAAGTACGTCA ACTTCCTGTACCTGGCAAGCCACTAC GAAAAGCTGAAGGGAAGCCCGGAAGA CAACGAACAGAAGCAGCTGTTTCGTCG AACAGCACAAAGCACTACCTGGACGAA ATCATCGAACAGATCAGCGAATTTCAG CAAGAGAGTCATCCTGGCAGACGCAA ACCTGGACAAGGTCCTGAGCGCATAC AACAAAGCACAGAGACAAGCCGATCAG AGAACAGGCAGAAAACATCATCCACC TGTTTCACACTGACAAACCTGGGAGCA CCGGCAGCATTCAAGTACTTCGACAC AACAAATCGACAGAAAGAGATACACAA GCACAAAGGAAGTCCTGGACGCAACA CTGATCCACCAGAGCATCACAGGACT GTACGAAACAAGAATCGACCTGAGCC AGCTGGGAGGAGACGGAGGAGGAAGC CCGAAGAAGAAGAGAAAGGTCTAGCT AGCCATCACATTTAAAAGCATCTCAG CCTACCATGAGAATAAGAGAAAGAAA ATGAAGATCAATAGCTTATTCATCTC TTTTTCTTTTTTCGTTGGTGTAAGCC AACACCCTGTCTAAAAAACATAAATT TCTTTAATCATTTTTGCCTCTTTTCTC TGTGCTTCAATTAATAAAAAAATGGAA AGAACCCTCGAG	
U 含量が 1 9.36%で ある代替的C as9 OR F	ATGGATAAGAAGTACTCGATCGGGCT GGATATCGGAACTAATTCCGTGGGTT GGGCAGTGATCACGGATGAATACAAA GTGCCGTCCAAGAAGTTCAAGGTCCT GGGGAACACCGATAGACACAGCATCA	4 5

10

20

30

40

50

	AGAAGAATCTCATCGGAGCCCTGCTG TTTGA CTCCGGCGAAACCGCAGAAGC GACCCGGCTCAAACGTACCGCGAGGC GACGCTACACCCGGCGGAAGAATCGC ATCTGCTATCTGCAAGAAATCTTTTC GAACGAAATGGCAAAGGTGGACGACA GCTTCTTCCACCGCCTGGAAGAATCT TTCCTGGTGGAGGAGGACAAGAAGCA TGAACGGCATCCTATCTTTGGAAACA TCGTGGACGAAGTGGCGTACCACGAA AAGTACCCGACCATCTACCATCTGCG GAAGAAGTTGGTTGACTCAACTGACA AGGCCGACCTCAGATTGATCTACTTG GCCCTCGCCCATATGATCAAATTCCG CGGACACTTCCTGATCGAAGGCGATC TGAACCCTGATAACTCCGACGTGGAT AAGCTGTTCAATTCAACTGGTGCAGAC CTACAACCAACTGTTCGAAGAAAACC CAATCAATGCCAGCGGCGTCGATGCC AAGGCCATCCTGTCCGCCCGGCTGTC GAAGTCGCGGCGCCTCGAAAACCTGA TCGCACAGCTGCCGGGAGAGAAGAAG AACGGACTTTTTCGGCAACTTGATCGC TCTCTCACTGGGACTCACTCCCAATT TCAAGTCCAATTTTGACCTGGCCGAG GACGCGAAGCTGCAACTCTCAAAGGA CACCTACGACGACGACTTGGAACAATT TGCTGGCACAAATTGGCGATCAGTAC GCGGATCTGTTCTTGCCGCTAAGAA CCTTTCGGACGCAATCTTGCTGTCCG ATATCCTGCGCGTGAACACCGAAATA ACCAAAGCGCCGCTTAGCGCCTCGAT GATTAAGCGGTACGACGAGCATCACC AGGATCTCACGCTGCTCAAAGCGCTC GTGAGACAGCAACTGCCTGAAAAGTA CAAGGAGATTTTCTTCGACCAGTCCA AGAATGGGTACGCAGGGTACATCGAT GGAGGCGCCAGCCAGGAAGAGTTCTA TAAGTTCATCAAGCCAATCCTGGAAA AGATGGACGGAACCGAAGAACTGCTG GTCAAGCTGAACAGGGAGGATCTGCT CCGCAAACAGAGAACCTTTGACAACG	

10

20

30

40

50

	GAAGCATTCCACACCAGATCCATCTG GGTGAGCTGCACGCCATCTTGCGGCG CCAGGAGGACTTTTACCCATTCTCA AGGACAACCGGGAAAAGATCGAGAAA ATTCTGACGTTCCGCATCCCGTATTA CGTGGGCCCCACTGGCGCGCGGCAATT CGCGCTTCGCGTGGATGACTAGAAAA TCAGAGGAAACCATCACTCCTTGGAA TTTCGAGGAAGTTGTGGATAAAGGGAG CTTCGGCACAAATCCTTCATCGAACGA ATGACCAACTTCGACAAGAATCTCCC AAACGAGAAGGTGCTTCCTAAGCACA GCCTCCTTTACGAATACTTCACTGTC TACAACGAACTGACTAAAGTGAAATA CGTTACTGAAGGAATGAGGAAGCCGG CCTTTCTGAGCGGAGAACAGAAGAAA GCGATTGTCGATCTGCTGTTCAAGAC CAACCGCAAGGTGACCGTCAAGCAGC TTAAAGAGGACTACTTCAAGAAGATC GAGTGTTTCGACTCAGTGGAAATCAG CGGAGTGGAGGACAGATTCAACGCTT CGCTGGGAACCTATCATGATCTCCTG AAGATCATCAAGGACAAGGACTTCCT TGACAACGAGGAGAACGAGGACATCC TGGAAGATATCGTCCTGACCTTGACC CTTTTTCGAGGATCGCGAGATGATCGA GGAGAGGCTTAAGACCTACGCTCATC TCTTCGACGATAAGGTCATGAAACAA CTCAAGCGCCGCGGTACACTGGTTG GGGCCGCTCTCCCGCAAGCTGATCA ACGGTATTTCGCGATAAACAGAGCGGT AAAACATATCCTGGATTTCTCTCAAATC GGATGGCTTCGCTAATCGTAACTTCA TGCAGTTGATCCACGACGACAGCCTG ACCTTTAAGGAGGACATCCAGAAAGC ACAAGTGAGCGGACAGGGAGACTCAC TCCATGAACACATCGCGAATCTGGCC GGTTCGCCGGCGATTAAAGAAGGGAAT CCTGCAAACCTGTGAAGGTGGTGGACG AGCTGGTGAAGGTCATGGGACGGCAC AAACCGGAGAATATCGTGATTGAAAT GGCCCGAGAAAACCAGACTACCCAGA	

10

20

30

40

50

AGGGCCAGAAGAACTCCCGCGAAAGG	
ATGAAGCGGATCGAAGAAGGAATCAA	
GGAGCTGGGCAGCCAGATCCTGAAAG	
AGCACCCGGTGGAAAACACGCAGCTG	
CAGAACGAGAAGCTCTACCTGTACTA	
TTTGCAAAATGGACGGGACATGTACG	
TGGACCAAGAGCTGGACATCAATCGG	
TTGTCTGATTACGACGTGGACCACAT	
CGTTCCACAGTCCTTTCTGAAGGATG	
ACTCCATCGATAACAAGGTGTTGACT	
CGCAGCGACAAGAACAGAGGGGAAGTC	
AGATAATGTGCCATCGGAGGAGGTCG	
TGAAGAAGATGAAGAATTACTGGCGG	
CAGCTCCTGAATGCGAAGCTGATTAC	
CCAGAGAAAGTTTGACAATCTCACTA	
AAGCCGAGCGCGGGCGGACTCTCAGAG	
CTGGATAAAGGCTGGATTTCATCAAACG	
GCAGCTGGTCGAGACTCGGCAGATTA	
CCAAGCACGTGGCGCAGATCCTGGAC	
TCCCGCATGAACACTAAATACGACGA	
GAACGATAAGCTCATCCGGGAAGTGA	
AGGTGATTACCCTGAAAAGCAAACCTT	
GTGTCGGACTTTTCGGAAGGACTTTCA	
GTTTTTACAAAGTGAGAGAAATCAACA	
ACTACCATCACGCGCATGACGCATAC	
CTCAACGCTGTGGTCGGCACCGCCCT	
GATCAAGAAGTACCCTAAACTTGAAT	
CGGAGTTTGTGTACGGAGACTACAAG	
GTCTACGACGTGAGGAAGATGATAGC	
CAAGTCCGAACAGGAAATCGGGAAAG	
CAACTGCGAAATACTTCTTTTACTCA	
AACATCATGAACTTCTTCAAGACTGA	
AATTACGCTGGCCAATGGAGAAATCA	
GGAAGAGGCCACTGATCGAAACTAAC	
GGAGAAACGGGCGAAATCGTGTGGGA	
CAAGGGCAGGGACTTCGCAACTGTTC	
GCAAAGTGCTCTCTATGCCGCAAGTC	
AATATTGTGAAGAAAACCGAAGTGCA	
AACCGGCGGATTTTCAAAGGAATCGA	
TCCTCCCAAAGAGAAATAGCGACAAG	
CTCATTGCACGCAAGAAAGACTGGGA	
CCCGAAGAAGTACGGAGGATTTCGATT	

10

20

30

40

50

	CGCCGACTGTGCGCATACTCCGTCCTC GTGGTGGCCAAAGGTGGAGAAGGGAAA GAGCAAGAAGCTCAAATCCGTCAAAG AGCTGCTGGGGATTACCATCATGGAA CGATCCTCGTTTCGAGAAGAACCCGAT TGATTTTCCTGGAGGCGAAGGGTTACA AGGAGGTGAAGAAGGATCTGATCATC AAACTGCCCAAGTACTCACTGTTCGA ACTGGAAAATGGTTCGGAAGCGCATGC TGGCTTCGGCCGGAGAAGTCCAGAAA GGAAATGAGCTGGCCTTGCTTAGCAA GTACGTCAACTTCCTCTATCTTGCTT CGCACTACGAGAACTCAAAGGGTCA CCGGAAGATAACGAACAGAAGCAGCT TTTCGTGGAGCAGCACAAAGCATTATC TGGATGAAATCATCGAACAAATCTCC GAGTTTTTCAAAGCGCGTGATCCTCGC CGACGCCAACCTCGACAAAGTCCTGT CGGCCTACAATAAGCATAGAGATAAG CCGATCAGAGAACAGGCCGAGAACAT TATCCACTTGTTTACCCCTGACTAACC TGGGAGCTCCAGCCGCTTCAAGTAC TTCGATACTACTATCGACCGCAAAAG ATACACGTCCACCAAGGAAGTTCTGG ACGCGACCCCTGATCCACCAAAGCATC ACTGGACTCTACGAACTAGGATCGA TCTGTGCGAGCTGGGTGGCGATGGTG GCGGTGGATCCTACCCATACGACGTG CCTGACTACGCCTCCGGAGGTGGTGG CCCCAAGAAGAAACGGAAGGTGTGAT AG	
HSDの5' UTR、配列番号45に対応するORF、コザック配列、及びALBの3' UTRを有するCas9転写産物	GGGTCCCGCAGTCGGCGTCCAGCGGC TCTGCTTGTTTCGTGTGTGTGTCGTTG CAGGCCTTATTTCGGATCTGCCACCAT GGATAAGAAGTACTCGATCGGGCTGG ATATCGGAACATAATTCGTTGGGTGG GCAGTGATCACGGATGAATACAAAGT GCCGTCCAAGAAGTTCAAGGTCCTGG GGAACACCGATAGACACAGCATCAAG AAGAATCTCATCGGAGCCCTGCTGTT TGA CTCCGGCGAAACCGCAGAAGCGA	46

40

	CCCGGCTCAAACGTACCGCGAGGCGA CGCTACACCCGGCGGAAGAATCGCAT CTGCTATCTGCAAGAAATCTTTTCGA ACGAAATGGCAAAGGTGGACGACAGC TTCTTCCACCGCCTGGAAGAATCTTT CCTGGTGGAGGAGGACAAGAAGCATG AACGGCATCCTATCTTTGGAAACATC GTGGACGAAGTGGCGTACCACGAAAA GTACCCGACCATCTACCATCTGCGGA AGAAGTTGGTTGACTCAACTGACAAG GCCGACCTCAGATTGATCTACTTGGC CCTCGCCCATATGATCAAATTCGCG GACACTTCCTGATCGAAGGCGATCTG AACCTGATAACTCCGACGTGGATAA GCTGTTTCATTCAACTGGTGCAGACCT ACAACCAACTGTTCGAAGAAAACCCA ATCAATGCCAGCGGCGTCGATGCCAA GGCCATCCTGTCCGCCCGGCTGTCTGA AGTCGCGGCGCCTCGAAAACCTGATC GCACAGCTGCCGGGAGAGAAGAAGAA CGGACTTTTTCGGCAACTTGATCGCTC TCTCACTGGGACTCACTCCCAATTTTC AAGTCCAATTTTGTACCTGGCCGAGGA CGCGAAGCTGCAACTCTCAAAGGACA CCTACGACGACGACTTGGACAATTTG CTGGCACAAATTGGCGATCAGTACGC GGATCTGTTCCTTGCCGCTAAGAACC TTTCGGACGCAATCTTGCTGTCCGAT ATCCTGCGCGTGAACACCGAAATAAC CAAAGCGCCGCTTAGCGCCTCGATGA TTAAGCGGTACGACGAGCATCACCAG GATCTCACGCTGCTCAAAGCGCTCGT GAGACAGCAACTGCCTGAAAAGTACA AGGAGATTTTCTTCGACCAGTCCAAG AATGGGTACGCAGGGTACATCGATGG AGGCGCCAGCCAGGAAGAGTTCTATA AGTTCATCAAGCCAATCCTGGAAAAG ATGGACGGAACCGAAGAAGTGTGGT CAAGCTGAACAGGGAGGATCTGCTCC GCAAACAGAGAACCTTTGACAACGGA AGCATTCACACACAGATCCATCTGGG TGAGCTGCACGCCATCTTGCGGCGCC	

10

20

30

40

50

50

	AGCTGGGCGAGCCAGATCCTGAAAGAG CACCCGGTGGAAAACACGCAGCTGCA GAACGAGAAGCTCTACCTGTACTATT TGCAAAATGGACGGGACATGTACGTG GACCAAGAGCTGGACATCAATCGGTT GTCTGATTACGACGTGGACCACATCG TTCCACAGTCCCTTTCTGAAGGATGAC TCCATCGATAACAAGGTGTTGACTCG CAGCGACAAGAACAGAGGGAAGTCAG ATAATGTGCCATCGGAGGAGGTCGTG AAGAAGATGAAGAATTACTGGCGGCA GCTCCTGAATGCGAAGCTGATTACCC AGAGAAAGTTTGGACAATCTCACTAAA GCCGAGCGCGGCGGACTCTCAGAGCT GGATAAGGCTGGATTTCATCAAACGGC AGCTGGTCGAGACTCGGCAGATTACC AAGCACGTGGCGCAGATCCTGGACTC CCGCATGAACACTAAATACGACGAGA ACGATAAGCTCATCCGGGAAGTGAAG GTGATTACCCTGAAAAGCAAACCTTGT GTCGGACTTTTCGGAAGGACTTTTCAGT TTTACAAAGTGAGAGAAATCAACAAC TACCATCACGCGCATGACGCATACCT CAACGCTGTGGTCGGCACCGCCCTGA TCAAGAAGTACCCTAAACTTTGAATCG GAGTTTGTGTACGGAGACTACAAGGT CTACGACGTGAGGAAGATGATAGCCA AGTCCGAACAGGAAATCGGGAAAGCA ACTGCGAAATACTTCTTTTACTCAAA CATCATGAACTTCTTCAAGACTGAAA TTACGCTGGCCAATGGAGAAATCAGG AAGAGGCCACTGATCGAAACTAACGG AGAAACGGGCGAAATCGTGTGGGACA AGGGCAGGGACTTCGCAACTGTTTCGC AAAGTGCTCTCTATGCCGCAAGTCAA TATTGTGAAGAAAACCGAAGTGCAAA CCGGCGGATTTTCAAAGGAATCGATC CTCCCAAAGAGAAATAGCGACAAGCT CATTGCACGCAAGAAAGACTGGGACC CGAAGAAGTACGGAGGATTTCGATTCTG CCGACTGTCTGCATACTCCGTCCTCGT GGTGGCCAAGGTGGAGAAGGGAAAGA	

10

20

30

40

50

	GCAAGAAGCTCAAATCCGTCAAAGAG CTGCTGGGGATTACCATCATGGAACG ATCCTCGTTTCGAGAAGAACCCGATTG ATTTCTCTGGAGGCGAAGGGTTACAAG GAGGTGAAGAAGGATCTGATCATCAA ACTGCCCAAGTACTCACTGTTTCGAAC TGGAAAATGGTCGGAAGCGCATGCTG GCTTCGGCCCGGAGAAGTCCAGAAAGG AAATGAGCTGGCCTTGCCTAGCAAGT ACGTCAACTTCCTCTATCTTGCTTCG CACTACGAGAACTCAAAGGGTCCACC GGAAGATAACGAACAGAAGCAGCTTT TCGTGGAGCAGCACAAAGCATTTATCTG GATGAAATCATCGAACAAATCTCCGA GTTTTCAAAGCGCGTGATCCTCGCCG ACGCCAACCTCGACAAAGTCCTGTGCG GCCTACAATAAGCATAGAGATAAGCC GATCAGAGAACAGGCCGAGAACATTA TCCACTTGTTCCACCCTGACTAACCTG GGAGCTCCAGCCGCCTTCAAGTACTT CGATACTACTATCGACCGCAAAAGAT ACACGTCCACCAAGGAAGTTCTGGAC GCGACCCTGATCCACCAAAGCATCAC TGGACTCTACGAAACTAGGATCGATC TGTCGCAGCTGGGTGGCGATGGTGGC GGTGGATCCTACCCATACGACGTGCC TGA CTACGCCTCCGGAGGTGGTGGCC CCAAGAAGAAACGGAAGGTGTGATAG CTAGCCATCACATTTAAAAGCATCTC AGCCTACCATGAGAATAAGAGAAAGA AAATGAAGATCAATAGCTTATTCATC TCTTTTTTCTTTTTTCGTTGGTGTAAG CCAACACCCTGTCTAAAAAACATAAA TTTCTTTTAATCATTTTGCCTCTTTTC TCTGTGCTTCAATTAATAAAAAATGG AAAGAACCCTCGAG	
HSDの5' UTR、配列番号45に対応するORF、及びAL	GGGTCCCGCAGTCGGCGTCCAGCGGC TCTGCTTGTTTCGTGTGTGTGTCGTTG CAGGCCTTATTCGGATCTATGGATAA GAAGTACTCGATCGGGCTGGATATCG GAACTAATTCCGTGGGTGGGCAGTG	47

10

20

30

40

50

Bの3' UTRを有するCas9転写産物	ATCACGGATGAATACAAAGTGCCGTC	
	CAAGAAGTTCAAGGTCCTGGGGAACA	
	CCGATAGACACAGCATCAAGAAGAAT	
	CTCATCGGAGCCCTGCTGTTTGACTC	
	CGGCGAAACCGCAGAAGCGACCCGGC	
	TCAAACGTACCGCGAGGCGACGCTAC	
	ACCCGGCGGAAGAATCGCATCTGCTA	
	TCTGCAAGAAATCTTTTCGAACGAAA	
	TGGCAAAGGTGGACGACAGCTTCTTC	10
	CACCGCCTGGAAGAATCTTTCCTGGT	
	GGAGGAGGACAAGAAGCATGAACGGC	
	ATCCTATCTTTGGAAACATCGTGGAC	
	GAAGTGGCGTACCACGAAAAGTACCC	
	GACCATCTACCATCTGCGGAAGAAGT	
	TGGTTGACTCAACTGACAAGGCCGAC	
	CTCAGATTGATCTACTTGGCCCTCGC	
	CCATATGATCAAATTCCGCGGACACT	
	TCCTGATCGAAGGCGATCTGAACCCCT	
	GATAACTCCGACGTGGATAAGCTGTT	
	CATTCAACTGGGTGCAGACCTACAACC	20
	AACTGTTTCGAAGAAAACCCAATCAAT	
	GCCAGCGGCGTCGATGCCAAGGCCAT	
	CCTGTCCGCCCGGCTGTCTGAAGTCGC	
	GGCGCCTCGAAAACCTGATCGCACAG	
	CTGCCGGGAGAGAAGAAGAACGGACT	
	TTTCGGCAACTTGATCGCTCTCTCAC	
	TGGGACTCACTCCCAATTTCAAGTCC	
	AATTTTGTACCTGGCCGAGGACGCGAA	
	GCTGCAACTCTCAAAGGACACCTACG	
	ACGACGACTTGGACAATTTGCTGGCA	
	CAAATTGGCGATCAGTACGCGGATCT	30
	GTTTCCTTGCCGCTAAGAACCTTTTCGG	
	ACGCAATCTTGCTGTCCGATATCCTG	
	CGCGTGAACACCGAAATAACCAAAGC	
	GCCGCTTAGCGCCTCGATGATTAAGC	
	GGTACGACGAGCATCACCAAGGATCTC	
	ACGCTGCTCAAAGCGCTCGTGAGACA	
	GCAACTGCCTGAAAAGTACAAGGAGA	
	TTTTCTTCGACCAGTCCAAGAATGGG	
	TACGCAGGGTACATCGATGGAGGCGC	
	CAGCCAGGAAGAGTTCTATAAGTTCA	
	TCAAGCCAATCCTGGAAAAGATGGAC	40

	GGAACCGAAGAAGCTGCTGGTCAAGCT GAACAGGGAGGATCTGCTCCGCAAAC AGAGAACCTTTGACAACGGAAGCATT CCACACCAGATCCATCTGGGTGAGCT GCACGCCATCTTGCGGCGCCAGGAGG ACTTTTACCCATTCTCAAGGACAAC CGGGAAAAGATCGAGAAAATTCTGAC GTTCCGCATCCCGTATTACGTGGGCC CACTGGCGCGCGGGCAATTCGCGCTTC GCGTGGATGACTAGAAAATCAGAGGA AACCATCACTCCTTGGAATTTTCGAGG AAGTTGTGGATAAGGGAGCTTCGGCA CAATCCTTCATCGAACGAATGACCAA CTTCGACAAGAATCTCCCAAACGAGA AGGTGCTTCCTAAGCACAGCCTCCTT TACGAATACTTCACTGTCTACAACGA ACTGACTAAAGTGAAATACGTTACTG AAGGAATGAGGAAGCCGGCCTTTCTG AGCGGAGAACAGAAGAAAGCGATTGT CGATCTGCTGTTCAAGACCAACCGCA AGGTGACCGTCAAGCAGCTTAAAGAG GACTACTTCAAGAAGATCGAGTGTTT CGACTCAGTGGAAATCAGCGGAGTGG AGGACAGATTCAACGCTTCGCTGGGA ACCTATCATGATCTCCTGAAGATCAT CAAGGACAAGGACTTCCTTGACAACG AGGAGAACGAGGACATCCTGGAAGAT ATCGTCCTGACCTTGACCCTTTTCGA GGATCGCGAGATGATCGAGGAGAGGC TTAAGACCTACGCTCATCTCTTCGAC GATAAGGTCATGAAACAACCTCAAGCG CCGCCGGGTACACTGGTTGGGGCCGCC TCTCCCGCAAGCTGATCAACGGTATT CGCGATAAACAGAGCGGTAAAACCTAT CCTGGATTTCTCTCAAATCGGATGGCT TCGCTAATCGTAACTTCATGCAGTTG ATCCACGACGACAGCCTGACCTTTAA GGAGGACATCCAGAAAGCACAAAGTGA GCGGACAGGGAGACTCACTCCATGAA CACATCGCGAATCTGGCCGGTTTCGCC GGCGATTAAAGAAGGGAATCCTGCAAA CTGTGAAGGTGGTGGACGAGCTGGTG	

10

20

30

40

50

	AAGGTCATGGGACGGCACAAACCGGA GAATATCGTGATTGAAATGGCCCGAG AAAACCAGACTACCCAGAAGGGCCAG AAGAACTCCCGCGAAAGGATGAAGCG GATCGAAGAAGGAATCAAGGAGCTGG GCAGCCAGATCCTGAAAGAGCACCCG GTGGAAAACACGCAGCTGCAGAACGA GAAGCTCTACCTGTACTATTTGCAAA ATGGACGGGACATGTACGTGGACCAA GAGCTGGACATCAATCGGTTGTCTGA TTACGACGTGGACCACATCGTTCCAC AGTCCTTTTCTGAAGGATGACTCCATC GATAACAAGGTGTTGACTCGCAGCGA CAAGAACAGAGGGAAGTCAGATAATG TGCCATCGGAGGAGGTCGTGAAGAAG ATGAAGAATTACTGGCGGCAGCTCCT GAATGCGAAGCTGATTACCCAGAGAA AGTTTGACAATCTCACTAAAGCCGAG CGCGGCGGACTCTCAGAGCTGGATAA GGCTGGATTTCATCAAACGGCAGCTGG TCGAGACTCGGCAGATTACCAAGCAC GTGGCGCAGATCCTGGACTCCCGCAT GAACACTAAATACGACGAGAACGATA AGCTCATCCGGGAAGTGAAGGTGATT ACCCTGAAAAGCAAACCTTGTTGTCGGA CTTTCGGAAGGACTTTTCAGTTTTTACA AAGTGAGAGAAATCAACAACCTACCAT CACGCGCATGACGCATACCTCAACGC TGTGGTCGGCACCGCCCTGATCAAGA AGTACCCTAAACTTTGAATCGGAGTTT GTGTACGGAGACTACAAGGTCTACGA CGTGAGGAAGATGATAGCCAAGTCCG AACAGGAAATCGGGAAAGCAAACCTGCG AAATACTTCTTTTACTCAAACATCAT GAACTTCTTCAAGACTGAAATTACGC TGGCCAATGGAGAAATCAGGAAGAGG CCACTGATCGAAACTAACGGAGAAAC GGGCGAAATCGTGTGGGACAAGGGCA GGGACTTCGCAACTGTTTCGCAAAGTG CTCTCTATGCCGCAAGTCAATATTGT GAAGAAAACCGAAGTGCAAACCGGCG GATTTTCAAAGGAATCGATCCTCCCA	
		10
		20
		30
		40

	AAGAGAAATAGCGACAAGCTCATTGC ACGCAAGAAAGACTGGGACCCGAAGA AGTACGGAGGATTTCGATTTCGCCGACT GTCGCATACTCCGTCCTCGTGGTGGC CAAGGTGGAGAAGGGAAAGAGCAAGA AGCTCAAATCCGTCAAAGAGCTGCTG GGGATTACCATCATGGAACGATCCTC GTTCGAGAAGAACC CGATTGATTTCC TGGAGGCGAAGGGTTACAAGGAGGTG AAGAAGGATCTGATCATCAAAC TGCC CAAGTACTCACTGTTTCGAACTGGAAA ATGGTTCGGAAGCGCATGCTGGCTTCG GCCGGAGA ACTCCAGAAAGGAAATGA GCTGGCCTTGCC TAGCAAGTACGTCA ACTTCCTCTATCTTGCTTCGCACTAC GAGAAACTCAAAGGGTCACCGGAAGA TAACGAACAGAAAGCAGCTTTTTCGTGG AGCAGCACAAAGCATTATCTGGATGAA ATCATCGAACAAATCTCCGAGTTTTTC AAAGCGCGTGATCCTCGCCGACGCCA ACCTCGACAAAGTCCTGTCTGGCCTAC AATAAGCATAGAGATAAGCCGATCAG AGAACAGGCCGAGAACATTATCCACT TGTTTCAACCCTGACTAACCTGGGAGCT CCAGCCGCCTTCAAGTACTTCGATAC TACTATCGACCGCAAAAGATACACGT CCACCAAGGAAGTTCTGGACGCGACC CTGATCCACCAAAGCATCACTGGACT CTACGAAACTAGGATCGATCTGTCTGC AGCTGGGTGGCGATGGTGGCGGTGGA TCCTACCCATACGACGTGCCTGACTA CGCCTCCGGAGGTGGTGGCCCCAAGA AGAAACGGAAGGTGTGATAGCTAGCC ATCACATTTTAAAAGCATCTCAGCCTA CCATGAGAATAAGAGAAAGAAAATGA AGATCAATAGCTTATTTCATCTCTTTT TCTTTTTTCGTTGGTGTAAGCCAACA CCCTGTCTAAAAAACATAAATTTCTT TAATCATTTTTGCCTCTTTTCTCTGTG CTTCAATTAATAAAAAATGGAAAGAA CCTCGAG	

10

20

30

40

50

一般にヒトで 高発現するコ ドンを使用す る C a s 9 ORFを含む C a s 9 転写 産物	GGGTCCCGCAGTCGGCGTCCAGCGGC	4 8
	TCTGCTTGTTCGTGTGTGTGTCGTTG	
	CAGGCCTTATTCGGATCCATGCCTAA	
	GAAAAAGCGGAAGGTCGACGGGGATA	
	AGAAGTACTCAATCGGGCTGGATATC	
	GGAAC TAATTCCGTGGGTTGGGCAGT	
	GATCACGGATGAATACAAAGTGCCGT	
	CCAAGAAGTTCAAGGTCCTGGGGAAC	
	ACCGATAGACACAGCATCAAGAAAAA	
	TCTCATCGGAGCCCTGCTGTTTGA	
	CCGGCGAAACCGCAGAAGCGACCCGG	
	CTCAAACGTACCGCGAGGCGACGCTA	
	CACCCGGCGGAAGAATCGCATCTGCT	
	ATCTGCAAGAGATCTTTTTCGAACGAA	
	ATGGCAAAGGTCGACGACAGCTTCTT	
	CCACCGCCTGGAAGAATCTTTCTG	
	TGGAGGAGGACAAGAAGCATGAACGG	
	CATCCTATCTTTGGAAACATCGTCGA	
	CGAAGTGGCGTACCACGAAAAGTACC	
	CGACCATCTACCATCTGCGGAAGAAG	
	TTGGTTGACTCAACTGACAAGGCCGA	
	CCTCAGATTGATCTACTTGGCCCTCG	
	CCCATATGATCAAATTCCGCGGACAC	
	TTCTGATCGAAGGCGATCTGAACCC	
	TGATAACTCCGACGTGGATAAGCTTT	
	TCATTCAACTGGTGCAGACCTACAAC	
	CAACTGTTTCGAAGAAAACCCAATCAA	
	TGCTAGCGGCGTCGATGCCAAGGCCA	
	TCCTGTCCGCCCGGCTGTGGAAGTCG	
	CGGCGCCTCGAAAACCTGATCGCACA	
	GCTGCCGGGAGAGAAAAAGAACGGAC	
	TTTTTCGGCAACTTGATCGCTCTCTCA	
	CTGGGACTCACTCCCAATTTCAAGTC	
	CAATTTTGACCTGGCCGAGGACGCGA	
	AGCTGCAACTCTCAAAGGACACCTAC	
	GACGACGACTTGGAACAATTTGCTGGC	
	ACAAATTGGCGATCAGTACGCGGATC	
	TGTTCCCTTGCCGCTAAGAACCTTTTCG	
	GACGCAATCTTGCTGTCCGATATCCT	
	GCGCGTGAACACCGAAATAACCAAAG	
	CGCCGCTTAGCGCCTCGATGATTAAG	
	CGGTACGACGAGCATCACCAGGATCT	

10

20

30

40

50

CACGCTGCTCAAAGCGCTCGTGAGAC	
AGCAACTGCCTGAAAAGTACAAGGAG	
ATCTTCTTCGACCAGTCCAAGAATGG	
GTACGCAGGGTACATCGATGGAGGCG	
CTAGCCAGGAAGAGTTCTATAAGTTC	
ATCAAGCCAATCCTGGAAAAGATGGA	
CGGAACCGAAGAACTGCTGGTCAAGC	
TGAACAGGGAGGATCTGCTCCGGAAA	
CAGAGAACCCTTTGACAACGGATCCAT	
TCCCCACCAGATCCATCTGGGTGAGC	10
TGCACGCCATCTTGCGGCGCCAGGAG	
GACTTTTACCCATTCCCTCAAGGACAA	
CCGGGAAAAGATCGAGAAAATTCTGA	
CGTTCCGCATCCCGTATTACGTGGGC	
CCACTGGCGCGCGGCAATTGCGCCTT	
CGCGTGGATGACTAGAAAATCAGAGG	
AAACCATCACTCCTTGGAATTTCGAG	
GAAGTTGTGGATAAGGGAGCTTCGGC	
ACAAAGCTTCATCGAACGAATGACCA	
ACTTCGACAAGAATCTCCCAAACGAG	20
AAGGTGCTTCCTAAGCACAGCCTCCT	
TTACGAATACTTCACTGTCTACAACG	
AACTGACTAAAGTGAAATACGTTACT	
GAAGGAATGAGGAAGCCGGCCTTTCT	
GTCCGGGAGAACAGAAGAAAGCAATTG	
TCGATCTGCTGTTCAAGACCAACCGC	
AAGGTGACCGTCAAGCAGCTTAAAGA	
GGACTACTTCAAGAAGATCGAGTGTT	
TCGACTCAGTGGAAATCAGCGGGGTG	
GAGGACAGATTCAACGCTTCGCTGGG	
AACCTATCATGATCTCCTGAAGATCA	30
TCAAGGACAAGGACTTCCTTGACAAC	
GAGGAGAACGAGGACATCCTGGAAGA	
TATCGTCCTGACCTTGACCCTTTTCG	
AGGATCGCGAGATGATCGAGGAGAGG	
CTTAAGACCTACGCTCATCTCTTCGA	
CGATAAGGTCATGAAACAACCTCAAGC	
GCCGCCGGTACACTGGTTGGGGCCGC	
CTCTCCCGCAAGCTGATCAACGGTAT	
TCGCGATAAACAGAGCGGTAAAATA	
TCCTGGATTTCTCTCAAATCGGATGGC	
TTCGCTAATCGTAACTTCATGCAATT	40

10

20

30

40

50

	GATCCACGACGACAGCCTGACCTTTA AGGAGGACATCCAAAAAGCACAAAGTG TCCGGACAGGGAGACTCACTCCATGA ACACATCGCGAATCTGGCCGGTTTCGC CGGCGATTAAAGAAGGGAATTCTGCAA ACTGTGAAGGTGGTCGACGAGCTGGT GAAGGTCATGGGACGGCACAAACCGG AGAATATCGTGATTGAAATGGCCCGA GAAAACCAGACTACCCAGAAGGGCCA GAAAAACTCCCGCGAAAGGATGAAGC GGATCGAAGAAGGAATCAAGGAGCTG GGCAGCCAGATCCTGAAAGAGCACCC GGTGGAAAACACGCAGCTGCAGAACG AGAAGCTCTACCTGTACTATTTGCAA AATGGACGGGACATGTACGTGGACCA AGAGCTGGACATCAATCGGTTGTCTG ATTACGACGTGGACCACATCGTTCCA CAGTCCTTTCTGAAGGATGACTCGAT CGATAACAAGGTGTTGACTCGCAGCG ACAAGAACAGAGGGAAGTCAGATAAT GTGCCATCGGAGGAGGTCGTGAAGAA GATGAAGAATTACTGGCGGCAGCTCC TGAATGCGAAGCTGATTACCCAGAGA AAGTTTGACAATCTCACTAAAGCCGA GCGCGGCGGACTCTCAGAGCTGGATA AGGCTGGATTTCATCAAACGGCAGCTG GTCGAGACTCGGCAGATTACCAAGCA CGTGGCGCAGATCTTGGACTCCCGCA TGAACACTAAATACGACGAGAACGAT AAGCTCATCCGGGAAGTGAAGGTGAT TACCCTGAAAAGCAAACCTTGTGTCGG ACTTTCGGAAGGACTTTCAGTTTTAC AAAGTGAGAGAAATCAACAACCTACCA TCACGCGCATGACGCATACCTCAACG CTGTGGTTCGGTACCGCCCTGATCAAA AAGTACCCTAAACTTGAATCGGAGTT TGTGTACGGAGACTACAAGGTCTACG ACGTGAGGAAGATGATAGCCAAGTCC GAACAGGAAATCGGGAAAGCAACTGC GAAATACTTCTTTTACTCAAACATCA TGAACTTTTTCAAGACTGAAATTACG CTGGCCAATGGAGAAATCAGGAAGAG	

10

20

30

40

50

	G C C A C T G A T C G A A A C T A A C G G A G A A A C G G G C G A A A T C G T G T G G G A C A A G G G C A G G G A C T T C G C A A C T G T T C G C A A A G T G C T C T C T A T G C C G C A A G T C A A T A T T G T G A A G A A A A C C G A A G T G C A A A C C G G C G G A T T T T C A A A G G A A T C G A T C C T C C C A A A G A G A A A T A G C G A C A A G C T C A T T G C A C G C A A G A A A G A C T G G G A C C C G A A G A A G T A C G G A G G A T T C G A T T C G C C G A C T G T C G C A T A C T C C G T C C T C G T G G T G G C C A A G G T G G A G A A G G G A A A G A G C A A A A A G C T C A A A T C C G T C A A A G A G C T G C T G G G G A T T A C C A T C A T G G A A C G A T C C T C G T T C G A G A A G A A C C C G A T T G A T T T C C T C G A G G C G A A G G G T T A C A A G G A G G T G A A G A A G G A T C T G A T C A T C A A A C T C C C C A A G T A C T C A C T G T T C G A A C T G G A A A A T G G T C G G A A G C G C A T G C T G G C T T C G G C C G G A G A A C T C C A A A A A G G A A A T G A G C T G G C C T T G C C T A G C A A G T A C G T C A A C T T C C T C T A T C T T G C T T C G C A C T A C G A A A A A C T C A A A G G G T C A C C G G A A G A T A A C G A A C A G A A G C A G C T T T T C G T G G A G C A G C A C A A G C A T T A T C T G G A T G A A A T C A T C G A A C A A A T C T C C G A G T T T T C A A A G C G C G T G A T C C T C G C C G A C G C C A A C C T C G A C A A A G T C C T G T C G G C C T A C A A T A A G C A T A G A G A T A A G C C G A T C A G A G A A C A G G C C G A G A A C A T T A T C C A C T T G T T C A C C C T G A C T A A C C T G G G A G C C C C A G C C G C C T T C A A G T A C T T C G A T A C T A C T A T C G A T C G C A A A A G A T A C A C G T C C A C C A A G G A A G T T C T G G A C G C G A C C C T G A T C C A C C A A A G C A T C A C T G G A C T C T A C G A A A C T A G G A T C G A T C T G T C G C A G C T G G G T G G C G A T T G A T A G T C T A G C C A T C A C A T T T A A A A G C A T C T C A G C C T A C C A T G A G A A T A A G A G A A A G A A A A T G A A G A T C A A T A G C T T A T T C A T C T C T T T T T C T T T T T C G T T G G T G T A A A G C C A A C A C C C T G T C T A A A A A A C A T A A A T T T C T T T A A T C A T T T T G C C T C T T T T C T C T G	
		10
		20
		30
		40

	TGCTTCAATTAAATAAAAAATGGAAAG AACCTCGAG	
一般にヒトで 高発現するコ ドンを使用す る C a s 9 ORFを有す る、コザック 配列を含むC a s 9 転写産 物	GGGTCCCCGCAGTCGGCGTCCAGCGGC TCTGCTTGTTTCGTGTGTGTGTCGTTG CAGGCCTTATTTCGGATCCGCCACCAT GCCTAAGAAAAAGCGGAAGGTTCGACG GGGATAAGAAGTACTCAATCGGGCTG GATATCGGAACATAATTCCGTGGGTG GGCAGTGATCACGGATGAATACAAAG TGCCGTCCAAGAAGTTCAAGGTCCTG GGGAACACCGATAGACACAGCATCAA GAAAAATCTCATTCGGAGCCCTGCTGT TTGACTCCGGCGAAACCGCGAGAAGCG ACCCGGCTCAAACGTACCGCGAGGCG ACGCTACACCCGGCGGAAGAATCGCA TCTGCTATCTGCAAGAGATCTTTTTCG AACGAAATGGCAAAGGTTCGACGACAG CTTCTTCCACCGCCTGGAAGAATCTT TCCTGGTGAGAGGAGGACAAGAAGCAT GAACGGCATCCTATCTTTGGAAACAT CGTCGACGAAGTGCGGTACCAACGAAA AGTACCCGACCATCTACCATCTGCGG AAGAAGTTGGTTGACTCAACTGACAA GGCCGACCTCAGATTGATCTACTTGG CCCTCGCCCATATGATCAAATTCCGC GGACACTTCCTGATCGAAGGCGATCT GAACCCTGATAACTCCGACGTGGATA AGCTTTTTCATTCAACTGGTGACAGACC TACAACCAACTGTTTCGAAGAAAACCC AATCAATGCTAGCGGCGTCGATGCCA AGGCCATCCTGTCCGCCCGGCTGTGCG AAGTCGCGGCGCCTCGAAAACCTGAT CGCACAGCTGCCGGGAGAGAAAAAGA ACGGACTTTTTCGGCAACTTGATCGCT CTCTCACTGGGACTCACTCCCAATTT CAAGTCCAATTTTGACCTGGCCGAGG ACGCGAAGCTGCAACTCTCAAAGGAC ACCTACGACGACGACTTGGACAATTT GCTGGCACAAATTGGCGATCAGTACG CGGATCTGTTCTTGCCGCTAAGAAC CTTTTCGGACGCAATCTTGCTGTCCGA	49

50

	CGGTATTTCGCGATAAACAGAGCGGTA AAACTATCCTGGATTTTCCTCAAATCG GATGGCTTCGCTAATCGTAACTTCAT GCAATTGATCCACGACGACAGCCTGA CCTTTAAGGAGGACATCCAAAAAGCA CAAGTGTCCGGACAGGGAGACTCACT CCATGAACACATCGCGAATCTGGCCG GTTCGCCGGCGATTAAAGAAGGGAATT CTGCAAACCTGTGAAGGTGGTTCGACGA GCTGGTGAAGGTCATGGGACGGCACA AACCGGAGAATATCGTGATTGAAATG GCCCGAGAAAACCAGACTACCCAGAA GGGCCAGAAAAACTCCCGCGAAAGGA TGAAGCGGATCGAAGAAGGAATCAAG GAGCTGGGCAGCCAGATCCTGAAAGA GCACCCGGTGGAAAACACGCAGCTGC AGAACGAGAAGCTCTACCTGTACTAT TTGCAAAATGGACGGGACATGTACGT GGACCAAGAGCTGGACATCAATCGGT TGTCTGATTACGACGTGGACCAACATC GTTCCACAGTCCTTTCTGAAGGATGA CTCGATCGATAACAAGGTGTTGACTC GCAGCGACAAGAACAGAGGGAAGTCA GATAATGTGCCATCGGAGGAGGTCGT GAAGAAGATGAAGAATTACTGGCGGC AGCTCCTGAATGCGAAGCTGATTACC CAGAGAAAGTTTGACAATCTCACTAA AGCCGAGCGCGGGCGGACTCTCAGAGC TGGATAAAGGCTGGATTTCATCAAACGG CAGCTGGTCGAGACTCGGCAGATTAC CAAGCACGTGGCGCAGATCTTGGACT CCCGCATGAACACTAAATACGACGAG AACGATAAGCTCATCCGGGAAGTGAA GGTGATTACCCTGAAAAGCAAACCTTG TGTCGGACTTTTCGGAAGGACTTTTCAG TTTTACAAAGTGAGAGAAATCAACAA CTACCATCACGCGCATGACGCATACC TCAACGCTGTGGTCGGTACCGCCCTG ATCAAAAAGTACCCTAAACTTGAATC GGAGTTTGTGTACGGAGACTACAAGG TCTACGACGTGAGGAAGATGATAGCC AAGTCCGAACAGGAAATCGGGAAAGC	

10

20

30

40

50

	A A C T G C G A A A T A C T T C T T T T A C T C A A A C A T C A T G A A C T T T T T C A A G A C T G A A A T T A C G C T G G C C A A T G G A G A A A T C A G G A A G A G G C C A C T G A T C G A A A C T A A C G G A G A A A C G G G C G A A A T C G T G T G G G A C A A G G G C A G G G A C T T C G C A A C T G T T C G C A A A G T G C T C T C T A T G C C G C A A G T C A A T A T T G T G A A G A A A A C C G A A G T G C A A A C C G G C G G A T T T T C A A A G G A A T C G A T C C T C C C A A A G A G A A A T A G C G A C A A G C T C A T T G C A C G C A A G A A A G A C T G G G A C C C G A A G A A G T A C G G A G G A T T C G A T T C G C C G A C T G T C G C A T A C T C C G T C C T C G T G G T G G C C A A G G T G G A G A A G G G A A A G A G C A A A A A G C T C A A A T C C G T C A A A G A G C T G C T G G G G A T T A C C A T C A T G G A A C G A T C C T C G T T C G A G A A G A A C C C G A T T G A T T T C C T C G A G G C G A A G G G T T A C A A G G A G G T G A A G A A G G A T C T G A T C A T C A A A C T C C C C A A G T A C T C A C T G T T C G A A C T G G A A A A T G G T C G G A A G C G C A T G C T G G C T T C G G C C G G A G A A C T C C A A A A A G G A A A T G A G C T G G C C T T G C C T A G C A A G T A C G T C A A C T T C C T C T A T C T T G C T T C G C A C T A C G A A A A A C T C A A A G G G T C A C C G G A A G A T A A C G A A C A G A A G C A G C T T T T C G T G G A G C A G C A C A A G C A T T A T C T G G A T G A A A T C A T C G A A C A A A T C T C C G A G T T T T C A A A G C G C G T G A T C C T C G C C G A C G C C A A C C T C G A C A A A G T C C T G T C G G C C T A C A A T A A G C A T A G A G A T A A G C C G A T C A G A G A A C A G G C C G A G A A C A T T A T C C A C T T G T T C A C C C T G A C T A A C C T G G G A G C C C C A G C C G C C T T C A A G T A C T T C G A T A C T A C T A T C G A T C G C A A A A G A T A C A C G T C C A C C A A G G A A G T T C T G G A C G C G A C C C T G A T C C A C C A A A G C A T C A C T G G A C T C T A C G A A A C T A G G A T C G A T C T G T C G C A G C T G G G T G G C G A T T G A T A G T C T A G C C A T C A C A T T T A A A A G C A T C T C A G C C T A C C A T G A G A A T A A G A G A A A G A A A A T G A A G A T C A A T A G C T T A T T C A	
		10
		20
		30
		40

	TCTCTTTTTCTTTTTTCGTTGGTGTAAGCCAACACCCCTGTCTAAAAAACATAAATTTCTTTAATCATTTTGCCTCTTTTCTCTGTGTGCTTCAATTAATAAAAAATGGAAAGAACCTCGAG	
スプライス部位が除去された Cas 9 ORF ; U 含量 12.75 %	ATGGACAAGAAGTACAGCATCGGACTGGACATCGGAACAAACAGCGTCGGATGGGCAGTCATCACAGACGAATACAAGGTCCCAGAGCAAGAAGTTCAAGGTCCTGGGAAACACAGACAGACACAGCATCAAGAAGAACCTGATCGGAGCACTGCTGTTCGACAGCGGAGAAACAGCAGAAGCAACAAGACTGAAGAGAACAGCAAGAAAGAGATACACAAGAAGAAAGAACAGAACTGTGCTACCTGCAGGAAATCTTTCAGCAACGAAATGGCAAAGGTCGACGACAGCTTCTTCCAC c g g CTGGAAGAAAGCTTCCTGGTCGAAGAAGACAAGAAGCACGAAAGACACCCGATCTTCGGAAACATCGTCGACGAAGTCGCATACCAACGAAAGTACCCGACAATCTACCACCTGAGAAAGAAGCTGGTTCGACAGCACAGACAAGGCAGACCTGAGACTGATCTACCTGGCACTGGCACACATGATCAAGTTCAGAGGACACTTCCTGATCGAAGGAGACCTGAACCCGGACAACAGCGACGTCGACAAGCTGTTTCATCCAGCTGGTCCAGACATACAACCAGCTGTTTCGAAGAAAACCCGATCAACGCAAGCGGAGTCGACGCAAGGCAATCCTGAGCGCAAGACTGAGCAAGAGCAGAAGACTGGAAAACCTGATCGCACAGCTGCCGGGAGAAAAGAAGAACGGACTGTTTCGGAAACCTGATCGCACTGAGCCTGGGACTGACACCGAACTTCAAGAGCAACTTCGACCTGGCAGAAGACGCAAAGCTGCAGCTGAGCAAGGACACATACGACGACGACCTGGACAACCTGCTGGCACAGATCGGAGACCAGTACGCAGACCTGTTTCTGGCAGCAAAGAACCTGAGCGACGCAATCCTGCTGAGCGACATCCTGAGAGTCAACACAGAAATC	50

10

20

30

40

50

	ACAAAGGCACCGCTGAGCGCAAGCAT GATCAAGAGATACGACGAACACCACC AGGACCTGACACTGCTGAAGGCACTG GTCAGACAGCAGCTGCCGGAAAAGTA CAAGGAAATCTTCTTCGACCAGAGCA AGAACGGATACGCAGGATACATCGAC GGAGGAGCAAGCCAGGAAGAATTCTA CAAGTTTCATCAAGCCGATCCTGGAAA AGATGGACGGAACAGAAGAACTGCTG GTCAAGCTGAACAGAGAAGACCTGCT GAGAAAGCAGAGAACATTTCGACAACG GAAGCATCCCGCACCCAGATCCACCTG GGAGAAGCTGCACGCAATCCTGAGAAG ACAGGAAGACTTCTACCCGTTCTCTGA AGGACAACAGAGAAAAGATCGAAAAG ATCCTGACATTTCAGAATCCCGTACTA CGTCGGACCGCTGGCAAGAGGAAACA GCAGATTTCGCATGGATGACAAGAAAG AGCGAAGAAACAATCACACCGTGGAA CTTCGAAGAAGTCGTTCGACAAGGGAG CAAGCGCACAGAGCTTCATCGAAAGA ATGACAAACTTCGACAAGAACCTGCC GAACGAAAAGGTCCTGCCGAAGCACA GCCTGCTGTACGAATACTTCACAGTC TACAACGAACTGACAAAGGTCAAGTA CGTCACAGAAGGAATGAGAAAGCCGG CATTCCTGAGCGGAGAACAGAAGAAG GCAATCGTCGACCTGCTGTTCAAGAC AAACAGAAAGGTCACAGTCAAGCAGC TGAAGGAAGACTACTTCAAGAAGATC GAATGCTTCGACAGCGTCGAAATCAG CGGAGTCGAAGACAGATTCAACGCAA GCCTGGGAACATACCACGACCTGCTG AAGATCATCAAGGACAAGGACTTCCT GGACAACGAAGAAAACGAAGACATCC TGGAAGACATCGTCCTGACACTGACA CTGTTTCGAAGACAGAGAAATGATCGA AGAAAGACTGAAGACATACGCACACC TGTTTCGACGACAAGGTCATGAAGCAG CTGAAGAGAAGAAGATACACAGGATG GGGAAGACTGAGCAGAAAGCTGATCA ACGGAATCAGAGACAAGCAGAGCGGA	

10

20

30

40

50

	AAGACAATCCTGGACTTCCTGAAGAG CGACGGATTTCGCAAACAGAAACTTCA TGCAGCTGATCCACGACGACAGCCTG ACATTCAAGGAAGACATCCAGAAGGC ACAGGTCAGCGGACAGGGAGACAGCC TGCACGAACACATCGCAAACCTGGCA GGAAGCCCGGCAATCAAGAAGGGAAT CCTGCAGACAGTCAAGGTCGTTCGACG AACTGGTCAAGGTCATGGGAAGACAC AAGCCGGAAAACATCGTCATCGAAAT GGCAAGAGAAAACAGACAACACAGA AGGGACAGAAGAACAGCAGAGAAAGA ATGAAGAGAATCGAAGAAGGAATCAA GGAACCTGGGAAGCCAGATCCTGAAGG AACACCCCGGTTCGAAAACACACAGCTG CAGAACGAAAAGCTGTACCTGTACTA CCTGCAaAACGGAAGAGACATGTACG TCGACCAGGAACCTGGACATCAACAGA CTGAGCGACTACGACGTCGACCACAT CGTCCCAGCAGAGCTTCCTGAAGGACG ACAGCATCGACAACAAGGTCCTGACA AGAAGCGACAAGAACAGAGGAAAGAG CGACAACGTCCCGAGCGAAGAAGTCG TCAAGAAGATGAAGAACTACTGGAGA CAGCTGCTGAACGCAAAGCTGATCAC ACAGAGAAAGTTTCGACAACCTGACAA AGGCAGAGAGAGGAGGACTGAGCGAA CTGGACAAGGCAGGATTCATCAAGAG ACAGCTGGTCGAAACAAGACAGATCA CAAAGCACGTTCGCACAGATCCTGGAC AGCAGAATGAACACAAAGTACGACGA AAACGACAAGCTGATCAGAGAAGTCA AGGTCATCACACTGAAGAGCAAGCTG GTCAGCGACTTCAGAAAGGACTTCCA GTTCTACAAGGTCAGAGAAATCAACA ACTACCACCACGCACACGACGCATAC CTGAACGCAGTCGTTCGGAACAGCACT GATCAAGAAGTACCCGAAGCTGGAAA GCGAATTTCGTCTACGGAGACTACAAG GTCTACGACGTCAGAAAGATGATCGC AAAGAGCGAACAGGAAATCGGAAAGG CAACAGCAAAGTACTTCTTCTACAGC	

10

20

30

40

50

	AACATCATGAACTTCTTCAAGACAGA AATCACACTGGCAAACGGAGAAATCA GAAAGAGACCGCTGATCGAAACAAAC GGAGAAACAGGAGAAATCGTCTGGGA CAAGGGAAGAGACTTTCGCAACAGTCA GAAAGGTCCTGAGCATGCCGCAGGTC AACATCGTCAAGAAGACAGAAGTCCA GACAGGAGGATTTCAGCAAGGAAAGCA TCCTGCCGAAGAGAAACAGCGACAAG CTGATCGCAAGAAAGAAGGACTGGGA CCCGAAGAAGTACGGAGGATTTCGACA GCCCGACAGTCGCATACAGCGTCCTG GTCGTCGCAAAGGTCGAAAAGGGAAA GAGCAAGAAGCTGAAGAGCGTCAAGG AACTGCTGGGAATCACCAATCATGGAA AGAAGCAGCTTTCGAAAAGAACCCGAT CGACTTCCTGGAAGCAAAGGGATACA AGGAAGTCAAGAAGGACCTGATCATC AAGCTGCCGAAGTACAGCCTGTTCTGA ACTGGAAAACCGGAAGAAAGAGAATGC TGGCAAGCGCAGGAGAACTGCAGAAG GGAAACGAACTGGCACTGCCGAGCAA GTACGTCAACTTCCTGTACCTGGCAA GCCACTACGAAAAGCTGAAGGGAAGC CCGGAAGACAACGAACAGAAGCAGCT GTTCTGTCGAACAGCACAAAGCACTACC TGGACGAAATCATCGAACAGATCAGC GAATTTCAGCAAGAGAGTCATCCTGGC AGACGCAAACCTGGACAAGGTCCTGA GCGCATACAACAAGCACAGAGACAAG CCGATCAGAGAACAGGCAGAAAACAT CATCCACCTGTTTCACACTGACAAACC TGGGAGCACCGGCAGCATTC AAGTAC TTCGACACAACAATCGACAGAAAGAG ATACACAAGCACAAAGGAAGTCCTGG ACGCAACACTGATCCACCAGAGCATC ACAGGACTGTACGAAACAAGAATCGA CCTGAGCCAGCTGGGAGGAGACGGAG GAGGAAGCCC GAAGAAGAAGAGAAAG GTCTAG	
HSDの5'	GGGTCCCGCAGTCGGCGTCCAGCGGC	5 1

10

20

30

40

50

[illegible]

AGGAAATCTTCTTCGACCAGAGCAAG AACGGATACGCAGGATACATCGACGG AGGAGCAAGCCAGGAAGAATTCTACA AGTTCATCAAGCCGATCCTGGAAAAG ATGGACGGAACAGAAGAACTGCTGGT CAAGCTGAACAGAGAAGACCTGCTGA GAAAGCAGAGAACATTCGACAACGGA AGCATCCCGCACCCAGATCCACCTGGG AGAACTGCACGCAATCCTGAGAAGAC AGGAAGACTTCTACCCGTTCTGGAAG GACAACAGAGAAAAGATCGAAAAGAT CCTGACATTTCAGAATCCCGTACTACG TCGGACCGCTGGCAAGAGGAAACAGC AGATTTCGCATGGATGACAAGAAAGAG CGAAGAAACAATCACACCGTGGAAC TCGAAGAAGTCGTCGACAAGGGAGCA AGCGCACAGAGCTTCATCGAAAGAAT GACAAACTTCGACAAGAACCTGCCGA ACGAAAAGGTCCTGCCGAAGCACAGC CTGCTGTACGAATACTTCACAGTCTA CAACGAACTGACAAAGGTCAAGTACG TCACAGAAGGAATGAGAAAGCCGGCA TTCCTGAGCGGAGAACAGAAGAAGGC AATCGTCGACCTGCTGTTCAAGACAA ACAGAAAGGTCACAGTCAAGCAGCTG AAGGAAGACTACTTCAAGAAGATCGA ATGCTTCGACAGCGTCGAAATCAGCG GAGTCGAAGACAGATTCAACGCAAGC CTGGGAACATAACACGACCTGCTGAA GATCATCAAGGACAAGGACTTCCTGG ACAACGAAGAAAACGAAGACATCCTG GAAGACATCGTCCTGACACTGACACT GTTCGAAGACAGAGAAATGATCGAAG AAAGACTGAAGACATACGCACACCTG TTCGACGACAAGGTCATGAAGCAGCT GAAGAGAAGAAGATACACAGGATGGG GAAGACTGAGCAGAAAGCTGATCAAC GGAATCAGAGACAAGCAGAGCGGAAA GACAATCCTGGACTTCCTGAAGAGCG ACGGATTTCGCAAACAGAACTTCATG CAGCTGATCCACGACGACAGCCTGAC ATTCAAGGAAGACATCCAGAAGGCAC	

10

20

30

40

50

	AGGTCAGCGGACAGGGAGACAGCCTG CACGAACACATCGCAAACCTGGCAGG AAGCCCGGCAATCAAGAAGGGAATCC TGCAGACAGTCAAGGTCGTGACGAA CTGGTCAAGGTCATGGGAAGACACAA GCCGGAAAACATCGTCATCGAAATGG CAAGAGAAAACCAGACAACACAGAAG GGACAGAAGAACAGCAGAGAAAGAAT GAAGAGAATCGAAGAAGGAATCAAGG AACTGGGAAGCCAGATCCTGAAGGAA CACCCGGTCGAAAACACACAGCTGCA GAACGAAAAGCTGTACCTGTACTACC TGCA _a AACGGAAGAGACATGTACGTC GACCAGGAACTGGACATCAACAGACT GAGCGACTACGACGTCGACCACATCG TCCCGCAGAGCTTCCTGAAGGACGAC AGCATCGACAACAAGGTCCTGACAAG AAGCGACAAGAACAGAGGAAAGAGCG ACAACGTCCCGAGCGAAGAAGTCGTC AAGAAGATGAAGAACTACTGGAGACA GCTGCTGAACGCAAAGCTGATCACAC AGAGAAAGTTTCGACAACCTGACAAAG GCAGAGAGAGGAGGACTGAGCGAACT GGACAAGGCAGGATTTCATCAAGAGAC AGCTGGTCGAAACAAGACAGATCACA AAGCACGTGCGCACAGATCCTGGACAG CAGAATGAACACAAAGTACGACGAAA ACGACAAGCTGATCAGAGAAGTCAAG GTCATCACACTGAAGAGCAAGCTGGT CAGCGACTTCAGAAAGGACTTCCAGT TCTACAAGGTCAGAGAAATCAACAAC TACCACCACGCGACACGACGCATACCT GAACGCAGTCGTGGAACAGCACTGA TCAAGAAGTACCCGAAGCTGGAAAGC GAATTTCGTCTACGGAGACTACAAGGT CTACGACGTCAGAAAGATGATCGCAA AGAGCGAACAGGAAATCGGAAAGGCA ACAGCAAAGTACTTCTTCTACAGCAA CATCATGAACTTCTTCAAGACAGAAA TCACACTGGCAAACGGAGAAATCAGA AAGAGACCGCTGATCGAAACAAACGG AGAAACAGGAGAAATCGTCTGGGACA	

10

20

30

40

50

AGGGAAGAGACTTCGCAACAGTCAGA AAGGTCCTGAGCATGCCGCAGGTCAA CATCGTCAAGAAGACAGAAGTCCAGA CAGGAGGATTTCAGCAAGGAAAGCATC CTGCCGAAGAGAAACAGCGACAAGCT GATCGCAAGAAAGAAGGACTGGGACC CGAAGAAGTACGGAGGATTTCGACAGC CCGACAGTCGCATACAGCGTCCTGGT CGTCGCAAAGGTTCGAAAAGGGAAAGA GCAAGAAGCTGAAGAGCGTCAAGGAA CTGCTGGGAATCACAAATCATGGAAAG AAGCAGCTTCGAAAAGAACCCGATCG ACTTCCTGGAAGCAAAGGGATACAAG GAAGTCAAGAAGGACCTGATCATCAA GCTGCCGAAGTACAGCCTGTTTCGAAC TGGAAAACGGAAGAAAGAGAATGCTG GCAAGCGCAGGAGAACTGCAGAAGGG AAACGAACTGGCACTGCCGAGCAAGT ACGTCAACTTCCTGTACCTGGCAAGC CACTACGAAAAGCTGAAGGGAAAGCCC GGAAGACAACGAACAGAAGCAGCTGT TCGTGGAACAGCACAAAGCACTACCTG GACGAAATCATCGAACAGATCAGCGA ATTCAGCAAGAGAGTCAATCCTGGCAG ACGCAAACCTGGACAAGGTCCTGAGC GCATACAACAAGCACAGAGACAAGCC GATCAGAGAACAGGCAGAAAACATCA TCCACCTGTTTACACTGACAAACCTG GGAGCACCGGCAGCATTCAAGTACTT CGACACAACAATCGACAGAAAGAGAT ACACAAGCACAAAGGAAGTCCTGGAC GCAACACTGATCCACCAGAGCATCAC AGGACTGTACGAAACAAGAATCGACC TGAGCCAGCTGGGAGGAGACGGAGGA GGAAGCCCGAAGAAGAAGAGAAAGGT CTAGCTAGCCATCACATTTAAAAGCA TCTCAGCCTACCATGAGAATAAGAGA AAGAAAATGAAGATCAATAGCTTATT CATCTCTTTTTCTTTTTTCGTTGGTGT AAAGCCAACACCCTGTCTAAAAACA TAAATTTCTTTAATCATTTTGCCTCT TTTCTCTGTGCTTCAATTAATAAAAA	

10

20

30

40

50

40

CAAGGAGATCTTCTTCGACCAGAGCA AGAACGGCTACGCCGGCTACATCGAC GGCGGCGCCAGCCAGGAGGAGTTCTA CAAGTTTCATCAAGCCCATCCTGGAGA AGATGGACGGCACCAGGAGCTGCTG GTGAAGCTGAACAGAGAGGACCTGCT GAGAAAGCAGAGAACCTTCGACAACG GCAGCATCCCCCACCAGATCCACCTG GGCGAGCTGCACGCCATCCTGAGAAG ACAGGAGGACTTCTACCCCTTCCTGA AGGACAACAGAGAGAAGATCGAGAAG ATCCTGACCTTCAGAATCCCCTACTA CGTGGGCCCCCTGGCCAGAGGCAACA GCAGATTTCGCCTGGATGACCAGAAAG AGCGAGGAGACCATCACCCCTGGAA CTTCGAGGAGGTGGTGGACAAGGGCG CCAGCGCCCAGAGCTTCATCGAGAGA ATGACCAACTTCGACAAGAACCTGCC CAACGAGAAGGTGCTGCCCAAGCACA GCCTGCTGTACGAGTACTTCACCGTG TACAACGAGCTGACCAAGGTGAAGTA CGTGACCGAGGGCATGAGAAAGCCCG CCTTCCTGAGCGGCGAGCAGAAGAAG GCCATCGTGGACCTGCTGTTCAAGAC CAACAGAAAGGTGACCGTGAAGCAGC TGAAGGAGGACTACTTCAAGAAGATC GAGTGCTTCGACAGCGTGGAGATCAG CGGCGTGGAGGACAGATTCAACGCCA GCCTGGGGCACCTACCACGACCTGCTG AAGATCATCAAGGACAAGGACTTCCT GGACAACGAGGAGAACGAGGACATCC TGGAGGACATCGTGCTGACCCTGACC CTGTTCGAGGACAGAGAGATGATCGA GGAGAGACTGAAGACCTACGCCCACC TGTTTCGACGACAAGGTGATGAAGCAG CTGAAGAGAAGAAGATACACCGGCTG GGGCAGACTGAGCAGAAAGCTGATCA ACGGCATCAGAGACAAGCAGAGCGGC AAGACCATCCTGGACTTCCTGAAGAG CGACGGCTTCGCCAACAGAACTTCA TGCAGCTGATCCACGACGACAGCCTG ACCTTCAAGGAGGACATCCAGAAGGC	

10

20

30

40

50

CCAGGTGAGCGGCCAGGGCGACAGCC	
TGCACGAGCACATCGCCAACCTGGCC	
GGCAGCCCCGCCATCAAGAAGGGCAT	
CCTGCAGACCGTGAAGGTGGTGGACG	
AGCTGGTGAAGGTGATGGGCAGACAC	
AAGCCCGAGAACATCGTGATCGAGAT	
GGCCAGAGAGAACCAGACCACCCAGA	
AGGGCCAGAAGAACAGCAGAGAGAGA	
ATGAAGAGAATCGAGGAGGGCATCAA	
GGAGCTGGGCAGCCAGATCCTGAAGG	
AGCACCCCGTGGAGAACACCCAGCTG	
CAGAACGAGAAGCTGTACCTGTACTA	
CCTGCAGAACGGCAGAGACATGTACG	
TGGACCAGGAGCTGGACATCAACAGA	
CTGAGCGACTACGACGTGGACCACAT	
CGTGCCCCAGAGCTTCCTGAAGGACG	
ACAGCATCGACAACAAGGTGCTGACC	
AGAAGCGACAAGAACAGAGGCAAGAG	
CGACAACGTGCCCAGCGAGGAGGTGG	
TGAAGAAGATGAAGAACTACTGGAGA	
CAGCTGCTGAACGCCAAGCTGATCAC	
CCAGAGAAAGTTTCGACAACCTGACCA	
AGGCCGAGAGAGGCGGCCTGAGCGAG	
CTGGACAAGGCCGGCTTCATCAAGAG	
ACAGCTGGTGGAGACCAGACAGATCA	
CCAAGCACGTGGCCCAGATCCTGGAC	
AGCAGAATGAACACCAAGTACGACGA	
GAACGACAAGCTGATCAGAGAGGTGA	
AGGTGATCACCTGAAGAGCAAGCTG	
GTGAGCGACTTCAGAAAGGACTTCCA	
GTTCTACAAGGTGAGAGAGATCAACA	
ACTACCACCACGCCCACGACGCCTAC	
CTGAACGCCGTGGTGGGCACCGCCCT	
GATCAAGAAGTACCCCAAGCTGGAGA	
GCGAGTTCGTGTACGGCGACTACAAG	
GTGTACGACGTGAGAAAGATGATCGC	
CAAGAGCGAGCAGGAGATCGGCAAGG	
CCACCGCCAAGTACTTCTTCTACAGC	
AACATCATGAACTTCTTCAAGACCGA	
GATCACCTGGCCAACGGCGAGATCA	
GAAAGAGACCCCTGATCGAGACCAAC	
GGCGAGACCGGCGAGATCGTGTGGGA	

10

20

30

40

50

	CAAGGGCAGAGACTTCGCCACCGTGAGAAAGGTGCTGAGCATGCCCCAGGTG AACATCGTGAAGAAGACCGAGGTGCA GACCGGCGGCTTTCAGCAAGGAGAGCATCCTGCCCAAGAGAAACAGCGACAAG CTGATCGCCAGAAAGAAGGACTGGGACCCCAAGAAGTACGGCGGCTTTCGACAGCCCCACCGTGGCCTACAGCGTGCTG GTGGTGGCCAAAGGTGGAGAAGGGCAA GAGCAAGAAGCTGAAGAGCGTGAAGGAGCTGCTGGGCATCACCATCATGGAG AGAAGCAGCTTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACA AGGAGGTGAAGAAGGACCTGATCATCAAGCTGCCCAAGTACAGCCTGTTCGAGCTGGAGAACGGCAGAAAGAGAATGCTGGCCAGCGCCGGCGAGCTGCAGAAG GGCAACGAGCTGGCCCTGCCAGCAAGTACGTGAACCTTCCTGTACCTGGCCAGCCACTACGAGAAGCTGAAGGGCAGC CCCGAGGACAACGAGCAGAAGCAGCTGTTTCGTGGAGCAGCACAAAGCACTACCTGGACGAGATCATCGAGCAGATCAGC GAGTTCAGCAAGAGAGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGAGCGCCTACAACAAGCACAGAGACAAG CCCATCAGAGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGGCGCCCCCGCCGCCTTCAAGTAC TTCGACACCACCATCGACAGAAAGAGATACACCAGCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATC ACCGGCCTGTACGAGACCAGAATCGACCTGAGCCAGCTGGGGCGGCGACGGCGGCGGCAGCCCCAAGAAGAAGAGAAAG GTGTGA	
HSDの5' UTR、配列番号52に対応するORF、コザック	GGGTCCCGCAGTCGGCGTCCAGCGGC TCTGCTTGTTTCGTGTGTGTGTCGTTG CAGGCCTTATTCGGATCCGCCACCAT GGACAAGAAGTACAGCATCGGCCTGG ACATCGGCACCAACAGCGTGGGCTGG	53

10

20

30

40

50

配列、及びA LBの3' U TRを有する C a s 9 転写 産物	GCCGTGATCACCACGAGTACAAGGT	
	GCCCAGCAAGAAGTTCAAGGTGCTGG	
	GCAACACCGACAGACACAGCATCAAG	
	AAGAACCTGATCGGCGCCCTGCTGTT	
	CGACAGCGGCGAGACCGCCGAGGCCA	
	CCAGACTGAAGAGAAACCGCCAGAAGA	
	AGATACACCAGAAGAAAGAACAGAAT	
	CTGCTACCTGCAGGAGATCTTCAGCA	
	ACGAGATGGCCAAGGTGGACGACAGC	10
	TTCTTCCACAGACTGGAGGAGAGCTT	
	CCTGGTGGAGGAGGACAAGAAGCACG	
	AGAGACACCCCATCTTCGGCAACATC	
	GTGGACGAGGTGGCCTACCACGAGAA	
	GTACCCACCATCTACCACCTGAGAA	
	AGAAGCTGGTGGACAGCACCGACAAG	
	GCCGACCTGAGACTGATCTACCTGGC	
	CCTGGCCCCACATGATCAAGTTCAGAG	
	GCCACTTCCTGATCGAGGGCGACCTG	
	AACCCCGACAACAGCGACGTGGACAA	
	GCTGTTCATCCAGCTGGTGCAGACCT	20
	ACAACCAGCTGTTTCGAGGAGAACCCC	
	ATCAACGCCAGCGGCGTGACGCCAA	
	GGCCATCCTGAGCGCCAGACTGAGCA	
	AGAGCAGAAGACTGGAGAACCTGATC	
	GCCCAGCTGCCCGGCGAGAAGAAGAA	
	CGGCCTGTTTCGGCAACCTGATCGCCC	
	TGAGCCTGGGCCTGACCCCCAACTTC	
	AAGAGCAACTTCGACCTGGCCGAGGA	
	CGCCAAGCTGCAGCTGAGCAAGGACA	
	CCTACGACGACGACCTGGACAACCTG	
	CTGGCCCAGATCGGCGACCAAGTACGC	30
	CGACCTGTTTCCTGGCCGCCAAGAACC	
	TGAGCGACGCCATCCTGCTGAGCGAC	
	ATCCTGAGAGTGAACACCGAGATCAC	
	CAAGGCCCCCCCTGAGCGCCAGCATGA	
	TCAAGAGATACGACGAGCACCAACAG	
	GACCTGACCTGCTGAAGGCCCTGGT	
	GAGACAGCAGCTGCCCGAGAAGTACA	
	AGGAGATCTTCTTCGACCAGAGCAAG	
	AACGGCTACGCCGGCTACATCGACGG	
	CGGCGCCAGCCAGGAGGAGTTCTACA	
	AGTTCATCAAGCCCATCCTGGAGAAG	40

ATGGACGGCACCCGAGGAGCTGCTGGT	
GAAGCTGAACAGAGAGGACCTGCTGA	
GAAAGCAGAGAACCTTCGACAACGGC	
AGCATCCCCCACCAGATCCACCTGGG	
CGAGCTGCACGCCATCCTGAGAAGAC	
AGGAGGACTTCTACCCCTTCCTGAAG	
GACAACAGAGAGAAGATCGAGAAGAT	
CCTGACCTTCAGAATCCCCTACTACG	
TGGGCCCCCTGGCCAGAGGCAACAGC	
AGATTTCGCCTGGATGACCAGAAAGAG	
CGAGGAGACCATCACCCCTGGAACT	
TCGAGGAGGTGGTGGACAAGGGCGCC	
AGCGCCCAGAGCTTCATCGAGAGAAT	
GACCAACTTCGACAAGAACCTGCCCA	
ACGAGAAGGTGCTGCCCAAGCACAGC	
CTGCTGTACGAGTACTTCACCGTGTA	
CAACGAGCTGACCAAGGTGAAGTACG	
TGACCGAGGGCATGAGAAAGCCCGCC	
TTCCTGAGCGGCGAGCAGAAGAAGGC	
CATCGTGGACCTGCTGTTCAAGACCA	
ACAGAAAGGTGACCGTGAAGCAGCTG	
AAGGAGGACTACTTCAAGAAGATCGA	
GTGCTTCGACAGCGTGGAGATCAGCG	
GCGTGGAGGACAGATTCAACGCCAGC	
CTGGGCACCTACCACGACCTGCTGAA	
GATCATCAAGGACAAGGACTTCCTGG	
ACAACGAGGAGAACGAGGACATCCTG	
GAGGACATCGTGCTGACCCTGACCCT	
GTTTCGAGGACAGAGAGATGATCGAGG	
AGAGACTGAAGACCTACGCCACCTG	
TTCGACGACAAGGTGATGAAGCAGCT	
GAAGAGAAGAAGATACACCGGCTGGG	
GCAGACTGAGCAGAAAGCTGATCAAC	
GGCATCAGAGACAAGCAGAGCGGCAA	
GACCATCCTGGACTTCCTGAAGAGCG	
ACGGCTTCGCCAACAGAAACTTCATG	
CAGCTGATCCACGACGACAGCCTGAC	
CTTCAAGGAGGACATCCAGAAGGCC	
AGGTGAGCGGCCAGGGCGACAGCCTG	
CACGAGCACATCGCCAACCTGGCCGG	
CAGCCCCGCCATCAAGAAGGGCATCC	
TGCAGACCGTGAAGGTGGTGGACGAG	

10

20

30

40

50

	CTGGTGAAGGTGATGGGCAGACACAA GCCCCGAGAACATCGTGATCGAGATGG CCAGAGAGAACCAGACCACCCAGAAG GGCCAGAAGAACAGCAGAGAGAGAAT GAAGAGAATCGAGGAGGGCATCAAGG AGCTGGGCAGCCAGATCCTGAAGGAG CACCCCGTGGAGAACACCCAGCTGCA GAACGAGAAGCTGTACCTGTACTACC TGCAGAACGGCAGAGACATGTACGTG GACCAGGAGCTGGACATCAACAGACT GAGCGACTACGACGTGGACCACATCG TGCCCCAGAGCTTCCTGAAGGACGAC AGCATCGACAACAAGGTGCTGACCAG AAGCGACAAGAACAGAGGCAAGAGCG ACAACGTGCCCAGCGAGGAGGTGGTG AAGAAGATGAAGAACTACTGGAGACA GCTGCTGAACGCCAAGCTGATCACCC AGAGAAAGTTTCGACAACCTGACCAAG GCCGAGAGAGGCGGCCTGAGCGAGCT GGACAAGGCCGGCTTCATCAAGAGAC AGCTGGTGGAGACCAGACAGATCACCC AAGCACGTGGCCCAGATCCTGGACAG CAGAATGAACACCAAGTACGACGAGA ACGACAAGCTGATCAGAGAGGTGAAG GTGATCACCCCTGAAGAGCAAGCTGGT GAGCGACTTCAGAAAGGACTTCCAGT TCTACAAGGTGAGAGAGATCAACAAC TACCACCACGCCCACGACGCCTACCT GAACGCCGTGGTGGGCACCGCCCTGA TCAAGAAGTACCCCAAGCTGGAGAGC GAGTTCGTGTACGGCGACTACAAGGT GTACGACGTGAGAAAGATGATCGCCA AGAGCGAGCAGGAGATCGGCAAGGCC ACCGCCAAGTACTTCTTCTACAGCAA CATCATGAACTTCTTCAAGACCGAGA TCACCCTGGCCAACGGCGAGATCAGA AAGAGACCCCTGATCGAGACCAACGG CGAGACCGGCGAGATCGTGTGGGACA AGGGCAGAGACTTCGCCACCGTGAGA AAGGTGCTGAGCATGCCCCAGGTGAA CATCGTGAAGAAGACCGAGGTGCAGA CCGGCGGCCTTCAGCAAGGAGAGCATC	
		10
		20
		30
		40

	CTGCCCAAGAGAAACAGCGACAAGCT GATCGCCAGAAAGAAGGACTGGGACC CCAAGAAGTACGGCGGCTTCGACAGC CCCACCGTGGCCTACAGCGTGCTGGT GGTGGCCAAGGTGGAGAAGGGCAAGA GCAAGAAGCTGAAGAGCGTGAAGGAG CTGCTGGGCGATCACCATCATGGAGAG AAGCAGCTTCGAGAAGAACCCCATCG ACTTCCTGGAGGCCAAGGGCTACAAG GAGGTGAAGAAGGACCTGATCATCAA GCTGCCCAAGTACAGCCTGTTCGAGC TGGAGAACGGCAGAAAGAGAATGCTG GCCAGCGCCGGCGAGCTGCAGAAGGG CAACGAGCTGGCCCTGCCCAGCAAGT ACGTGAACTTCCTGTACCTGGCCAGC CACTACGAGAAGCTGAAGGGCAGCCC CGAGGACAACGAGCAGAAGCAGCTGT TCGTGGAGCAGCACAAAGCACTACCTG GACGAGATCATCGAGCAGATCAGCGA GTTTCAGCAAGAGAGTGATCCTGGCCG ACGCCAACCTGGACAAGGTGCTGAGC GCCTACAACAAGCACAGAGACAAGCC CATCAGAGAGCAGGCCGAGAACATCA TCCACCTGTTCACCCCTGACCAACCTG GGCGCCCCCGCCGCTTCAAGTACTT CGACACCACCATCGACAGAAAGAGAT ACACCAGCACCAAGGAGGTGCTGGAC GCCACCCTGATCCACCAGAGCATCAC CGGCCTGTACGAGACCAGAATCGACC TGAGCCAGCTGGGCGGGCGACGGCGGC GGCAGCCCCAAGAAGAAGAGAAAGGT GTGACTAGCCATCACATTTAAAAGCA TCTCAGCCTACCATGAGAATAAGAGA AAGAAAATGAAGATCAATAGCTTATT CATCTCTTTTTTCTTTTTTCGTTGGTGT AAAGCCAACACCCTGTCTAAAAAACA TAAATTTCTTTAATCATTTTGCCTCT TTTCTCTGTGCTTCAATTAATAAAAA ATGGAAAGAACCTCGAG	
一般にヒトで まれに使用さ	ATGGACAAAAAATACAGCATAGGGCT AGACATAGGGACGAACAGCGTAGGGT	5 4

10

20

30

40

50

[illegible]

AAATGGACGGGACGGAAGAATACTACTA GTAAAACATAAACCGAGAAGACCTACT ACGAAAACAACGAACGTTTCGACAACG GGAGCATACCGCACCAATACACCTA GGGGAATACTACACGCGATACTACGACG ACAAGAAGACTTCTACCCGTTCTCTAA AAGACAACCGAGAAAAAATAGAAAAA ATACTAACGTTCCGAATACCGTACTA CGTAGGGCCGCTAGCGCGAGGGGAACA GCCGATTTCGCGTGGATGACGCGAAAA AGCGAAGAAACGATAACGCCGTGGAA CTTCGAAGAAGTAGTAGACAAAGGGG CGAGCGCGCAAAGCTTCATAGAACGA ATGACGAACCTTCGACAAAAACCTACC GAACGAAAAAGTACTACCGAAACACA GCCTACTATACGAATACTTCACGGTA TACAACGAACATAACGAAAGTAAAATA CGTAACGGGAAGGGATGCGAAAACCGG CGTTCCTAAGCGGGGAACAAAAAAA GCGATAGTAGACCTACTATTCAAAAC GAACCGAAAAGTAACGGTAAAACAAC TAAAAGAAGACTACTTCAAAAAAATA GAATGCTTCGACAGCGTAGAAATAAG CGGGGTAGAAGACCGATTCAACGCGA GCCTAGGGACGTACCACGACCTACTA AAAATAATAAAAGACAAAGACTTCCT AGACAACGAAGAAAACGAAGACATAC TAGAAGACATAGTACTAACGCTAACG CTATTTCGAAGACCGAGAAATGATAGA AGAACGACTAAAAACGTACGCGCACCC TATTTCGACGACAAAGTAATGAAACAA CTAAAACGACGACGATACACGGGGTG GGGGCGACTAAGCCGAAAATAATAA ACGGGATACGAGACAAACAAAGCGGG AAAACGATACTAGACTTCCTAAAAAG CGACGGGTTCGCGAACCGAAACTTCA TGCAACTAATACACGACGACAGCCTA ACGTTCAAAGAAGACATACAAAAAGC GCAAGTAAGCGGGCAAGGGGACAGCC TACACGAACACATAGCGAACCTAGCG GGGAGCCCGGCGATAAAAAAAGGGAT ACTACAAACGGTAAAAGTAGTAGACG	

10

20

30

40

50

	AACTAGTAAAAGTAATGGGGCGACAC AAACCGGAAAACATAGTAATAGAAAT GGCGCGAGAAAACCAAACGACGCAAA AAGGGCAAAAAAACAGCCGAGAACGA ATGAAACGAATAGAAGAAGGGATAAA AGAACTAGGGAGGCCAAATACTAAAAG AACACCCGGTAGAAAACACGCAACTA CAAAACGAAAAACTATACCTATACTA CCTACAAAACGGGGCGAGACATGTACG TAGACCAAGAACTAGACATAAACCGA CTAAGCGACTACGACGTAGACCCACAT AGTACCGCAAAGCTTCCTAAAAGACG ACAGCATAGACAACAAAGTACTAACG CGAAGCGACAAAAAACCGAGGGGAAAAG CGACAACGTACCGAGCGAAGAAGTAG TAAAAAAAATGAAAAACTACTGGCGA CAACTACTAAACGCGAAACTAATAAC GCAACGAAAATTTCGACAACCTAACGA AAGCGGAACGAGGGGGGGCTAAGCGAA CTAGACAAAGCGGGGGTTCATAAACG ACAACTAGTAGAAACGCGACAAATAA CGAAACACGTAGCGCAAATACTAGAC AGCCGAATGAACACGAAATACGACGA AAACGACAAACTAATACGAGAAGTAA AAGTAATAACGCTAAAAAGCAAACCTA GTAAGCGACTTCCGAAAAGACTTCCA ATTCTACAAAGTACGAGAAATAAACA ACTACCACCACGCGCACGACGCGGTAC CTAAACGCGGTAGTAGGGACGGCGCT AATAAAAAAATACCCGAAACTAGAAA GCGAATTTCGTATACGGGGACTACAAA GTATACGACGTACGAAAAAATGATAGC GAAAAGCGAACAAGAAATAGGGAAAG CGACGGCGAAATACTTCTTCTACAGC AACATAATGAACTTCTTCAAAACGGA AATAACGCTAGCGAACGGGGAAATAC GAAAACGACCGCTAATAGAAACGAAC GGGGAAACGGGGGAAATAGTATGGGA CAAAGGGCGAGACTTCGCGACGGTAC GAAAAGTACTAAGCATGCCGCAAGTA AACATAGTAAAAAAAACGGAAGTACA AACGGGGGGGGTTCAGCAAAGAAAGCA	

10

20

30

40

50

50

産物	CGACAGCGGGGAAACGGCGGAAGCGA	
	CGCGACTAAAACGAACGGCGCGACGA	
	CGATACACGCGACGAAAAAACCGAAT	
	ATGCTACCTACAAGAAATATTCAGCA	
	ACGAAATGGCGAAAGTAGACGACAGC	
	TTCTTCCACCGACTAGAAGAAAGCTT	
	CCTAGTAGAAGAAGACAAAAAACACG	
	AACGACACCCGATATTCGGGAACATA	
	GTAGACGAAGTAGCGTACCACGAAAA	10
	ATACCCGACGATATACCACCTACGAA	
	AAAAACTAGTAGACAGCACGGACAAA	
	GCGGACCTACGACTAATATACCTAGC	
	GCTAGCGCACATGATAAAATTCCGAG	
	GGCACTTCCTAATAGAAGGGGACCTA	
	AACCCGGACAACAGCGACGTAGACAA	
	ACTATTTCATACAACCTAGTACAAACGT	
	ACAACCAACTATTCGAAGAAAACCCG	
	ATAAACGCGAGCGGGGTAGACGCGAA	
	AGCGATACTAAGCGCGCGACTAAGCA	
	AAAGCCGACGACTAGAAAACCTAATA	20
	GCGCAACTACCGGGGGGAAAAAAAAAA	
	CGGGCTATTCGGGAACCTAATAGCGC	
	TAAGCCTAGGGCTAACGCCGAACCTC	
	AAAAGCAACTTCGACCTAGCGGAAGA	
	CGCGAAACTACAACCTAAGCAAAGACA	
	CGTACGACGACGACCTAGACAACCTA	
	CTAGCGCAAATAGGGGACCAATACGC	
	GGACCTATTCCTAGCGGCGAAAAACC	
	TAAGCGACGCGATACTACTAAGCGAC	
	ATACTACGAGTAAACACGGAAATAAC	
	GAAAGCGCCGCTAAGCGCGAGCATGA	30
	TAAAACGATACGACGAACACCACCAA	
	GACCTAACGCTACTAAAAGCGCTAGT	
	ACGACAACAACCTACCGGAAAAATACA	
	AAGAAATATTCTTCGACCAAAGCAAA	
	AACGGGTACGCGGGGTACATAGACGG	
	GGGGGCGAGCCAAGAAGAATTCTACA	
	AATTCATAAAAACCGATACTAGAAAAA	
	ATGGACGGGACGGAAGAAGCTACTAGT	
	AAAACATAAACCGAGAAGACCTACTAC	
	GAAAACAACGAACGTTTCGACAACGGG	
	AGCATACCGCACCAAATACACCTAGG	40

	GGA ACTACACGCGATACTACGACGAC AAGAAGACTTCTACCCGTTCTAAAA GACAACCGAGAAAAAATAGAAAAAAT ACTAACGTTCCGAATACCGTACTACG TAGGGCCGCTAGCGCGAGGGAACAGC CGATTTCGCGTGGATGACGCGAAAAAG CGAAGAAACGATAACGCCGTGGA ACT TCGAAGAAGTAGTAGACAAAGGGGCG AGCGCGCAAAGCTTCATAGAACGAAT GACGAACTTCGACAAAAACCTACCGA ACGAAAAAGTACTACCGAAACACAGC CTACTATACGAATACTTCACGGTATA CAACGAACTAACGAAAGTAAAATACG TAACGGAAGGGATGCGAAAACCGGCG TTCCTAAGCGGGGAACAAAAAAAAGC GATAGTAGACCTACTATTCAAAACGA ACCGAAAAAGTAACGGTAAAACAACTA AAAGAAGACTACTTCAAAAAAATAGA ATGCTTCGACAGCGTAGAAATAAGCG GGGTAGAAGACCGATTCAACGCGAGC CTAGGGACGTACCACGACCTACTAAA AATAATAAAAGACAAAGACTTCCTAG ACAACGAAGAAAAACGAAGACATACTA GAAGACATAGTACTAACGCTAACGCT ATTCGAAGACCGAGAAATGATAGAAG AACGACTAAAAACGTACGCGCACCTA TTCGACGACAAAGTAATGAAACA ACT AAAACGACGACGATACACGGGGTGGG GGCGACTAAGCCGAAA ACTAATAAAC GGGATACGAGACAAACAAAGCGGGAA AACGATACTAGACTTCCTAAAAAGCG ACGGGTTCGCGAACCGAA ACTTCATG CAACTAATACACGACGACAGCCTAAC GTTCAAAGAAGACATACAAAAAGCGC AAGTAAGCGGGCAAGGGGACAGCCTA CACGAACACATAGCGAACCTAGCGGG GAGCCCGGCGATAAAAAAAGGGATAC TACAAACGGTAAAAGTAGTAGACGAA CTAGTAAAAGTAATGGGGCGACACAA ACCGGAAAACATAGTAATAGAAATGG CGCGAGAAAACCAAACGACGCAAAAA GGGCAAAAAAACAGCCGAGAACGAAT	
		10
		20
		30
		40

	GAAACGAATAGAAAGGAGGGATAAAAG AACTAGGGAGCCAAATACTAAAAGAA CACCCGGTAGAAAACACGCAACTACA AAACGAAAAAACTATACCTATACTACC TACAAAACGGGCGAGACATGTACGTA GACCAAGAACTAGACATAAACCGACT AAGCGACTACGACGTAGACCACATAG TACCGCAAAGCTTCCTAAAAGACGAC AGCATAGACAACAAAGTACTAACGCG AAGCGACAAAAAACCGAGGGGAAAAGCG ACAACGTACCGAGCGAAGAAGTAGTA AAAAAAAATGAAAAACTACTGGCGACA ACTACTAAACGCGAAACTAATAACGC AACGAAAAATTTCGACAACCTAACGAAA GCGGAACGAGGGGGGCTAAGCGAACT AGACAAAGCGGGGTTCATAAAACGAC AACTAGTAGAAACGCGACAAATAACG AAACACGTAGCGCAAATACTAGACAG CCGAATGAACACGAAATACGACGAAA ACGACAAACTAATACGAGAAGTAAAA GTAATAACGCTAAAAAGCAAAGTAGT AAGCGACTTCCGAAAAGACTTCCAAT TCTACAAAGTACGAGAAATAAACAAC TACCACCACGCGCACGACGCGTACCT AAACGCGGTAGTAGGGACGGCGCTAA TAAAAAAAATACCCGAAACTAGAAAGC GAATTTCGTATACGGGGACTACAAAGT ATACGACGTACGAAAAAATGATAGCGA AAAGCGAACAAGAAATAGGGAAAGCG ACGGCGAAATACTTCTTCTACAGCAA CATAATGAACTTCTTCAAAACGGAAA TAACGCTAGCGAACGGGGAAATACGA AAACGACCGCTAATAGAAACGAACGG GGAAACGGGGGAAATAGTATGGGACA AAGGGCGAGACTTCGCGACGGTACGA AAAGTACTAAGCATGCCGCAAGTAAA CATAGTAAAAAAAACGGAAGTACAAA CGGGGGGGTTCAGCAAAGAAAGCATA CTACCGAAACGAAACAGCGACAAACT AATAGCGCGAAAAAAAAGACTGGGACC CGAAAAAATACGGGGGGTTCGACAGC CCGACGGTAGCGTACAGCGTACTAGT	
		10
		20
		30
		40

	AGTAGCGAAAGTAGAAAAAGGGAAAA GCAAAAAAATAAAAAAGCGTAAAAGAA CTACTAGGGATAACGATAATGGAACG AAGCAGCTTCGAAAAAAACCCGATAG ACTTCCTAGAAGCGAAAGGGTACAAA GAAGTAAAAAAAGACCTAATAATAAA ACTACCGAAATACAGCCTATTTCGAAC TAGAAAACGGGCGAAAAACGAATGCTA GCGAGCGCGGGGGAACTACAAAAAGG GAACGAAGTAGCGCTACCGAGCAAAT ACGTAAACTTCCTATACCTAGCGAGC CACTACGAAAAAATAAAAAGGGAGCCC GGAAGACAACGAACAAAAACAACATAT TCGTAGAACAACACAAACACTACCTA GACGAAATAATAGAACAAATAAGCGA ATTCAGCAAACGAGTAATACTAGCGG ACGCGAACCTAGACAAAGTACTAAGC GCGTACAACAACACCGAGACAAACC GATACGAGAACAAGCGGAAAACATAA TACACCTATTACAGCTAACGAACCTA GGGGCGCCGGCGGCGTTCAAATACTT CGACACGACGATAGACCGAAAACGAT ACACGAGCACGAAAGAAGTACTAGAC GCGACGCTAATACACCAAAGCATAAC GGGGCTATACGAAACGCGAATAGACC TAAGCCAACTAGGGGGGGGACGGGGGG GGGAGCCCGAAAAAAAACGAAAAGT ATGACTAGCCATCACATTTAAAAGCA TCTCAGCCTACCATGAGAATAAGAGA AAGAAAATGAAGATCAATAGCTTATT CATCTCTTTTTTCTTTTTTCGTTGGTGT AAAGCCAACACCCTGTCTAAAAAACA TAAATTTCTTTAATCATTTTGCCTCT TTTCTCTGTGCTTCAATTAATAAAAA ATGGAAAGAACCTCGAG	
CleanC ap (商標)と 共に使用する ための最初の 3ヌクレオチ ドとしてのA	AGGTCCCGCAGTCGGCGTCCAGCGGC TCTGCTTGTTTCGTGTGTGTGTCGTTG CAGGCCTTATTTCGGATCCGCCACCAT GGACAAGAAGTACAGCATCGGACTGG ACATCGGAACAAACAGCGTCGGATGG GCAGTCATCACAGACGAATACAAGGT	56

10

20

30

40

50

GG、HSD の5' UTR、 配列番号4に 対応するOR F、コザック 配列、及びA LBの3' U TRを有する Cas9転写 産物	CCCCGAGCAAGAAGTTCAAGGTCCTGG GAAACACAGACAGACACAGCATCAAG AAGAACCTGATCGGAGCACTGCTGTT CGACAGCGGAGAAACAGCAGAAGCAA CAAGACTGAAGAGAACAGCAAGAAGA AGATACACAAGAAGAAAGAACAGAAT CTGCTACCTGCAGGAAATCTTCAGCA ACGAAATGGCAAAGGTCGACGACAGC TTCTTCCACAGACTGGAAGAAAGCTT CCTGGTTCGAAGAAGACAAGAAGCACG AAAGACACCCGATCTTCGGAAACATC GTCGACGAAGTCGCATACCACGAAAA GTACCCGACAATCTACCACCTGAGAA AGAAGCTGGTCGACAGCACAGACAAG GCAGACCTGAGACTGATCTACCTGGC ACTGGCACACATGATCAAGTTCAGAG GACACTTCCTGATCGAAGGAGACCTG AACCCGGACAACAGCGACGTTCGACAA GCTGTTTCATCCAGCTGGTCCAGACAT ACAACCAGCTGTTTCGAAGAAAACCCG ATCAACGCAAGCGGAGTCGACGCAAA GGCAATCCTGAGCGCAAGACTGAGCA AGAGCAGAAGACTGGAAAACCTGATC GCACAGCTGCCGGGAGAAAAGAAGAA CGGACTGTTTCGGAAACCTGATCGCAC TGAGCCTGGGACTGACACCGAACTTC AAGAGCAACTTCGACCTGGCAGAAGA CGCAAAGCTGCAGCTGAGCAAGGACA CATACGACGACGACCTGGACAACCTG CTGGCACAGATCGGAGACCAAGTACGC AGACCTGTTTCCTGGCAGCAAAGAACC TGAGCGACGCAATCCTGCTGAGCGAC ATCCTGAGAGTCAACACAGAAATCAC AAAGGCACCGCTGAGCGCAAGCATGA TCAAGAGATACGACGAACACCACCAG GACCTGACACTGCTGAAGGCACTGGT CAGACAGCAGCTGCCGGAAAAGTACA AGGAAATCTTCTTCGACCAGAGCAAG AACGGATACGCAGGATACATCGACGG AGGAGCAAGCCAGGAAGAATTCTACA AGTTTCATCAAGCCGATCCTGGAAAAG ATGGACGGAACAGAAGAACTGCTGGT	10 20 30 40

	CAAGCTGAACAGAGAAGACCTGCTGA GAAAGCAGAGAACATTCGACAACGGA AGCATCCCGCACCCAGATCCACCTGGG AGAACTGCACGCAATCCTGAGAAGAC AGGAAGACTTCTACCCGTTCTGAAAG GACAACAGAGAAAAGATCGAAAAGAT CCTGACATTCAGAATCCCGTACTACG TCGGACCGCTGGCAAGAGGAAACAGC AGATTTCGCATGGATGACAAGAAAGAG CGAAGAAACAATCACACCGTGGAAC TCGAAGAAGTCGTGACAAGGGAGCA AGCGCACAGAGCTTCATCGAAAGAAT GACAAACTTCGACAAGAACCTGCCGA ACGAAAAGGTCCTGCCGAAGCACAGC CTGCTGTACGAATACTTCACAGTCTA CAACGAACTGACAAAGGTCAAGTACG TCACAGAAGGAATGAGAAAGCCGGCA TTCCTGAGCGGAGAACAGAAGAAGGC AATCGTCGACCTGCTGTTCAAGACAA ACAGAAAGGTCACAGTCAAGCAGCTG AAGGAAGACTACTTCAAGAAGATCGA ATGCTTCGACAGCGTCGAAATCAGCG GAGTCGAAGACAGATTCAACGCAAGC CTGGGAACATAACACGACCTGCTGAA GATCATCAAGGACAAGGACTTCCTGG ACAACGAAGAAAACGAAGACATCCTG GAAGACATCGTCCTGACACTGACACT GTTCTGAAGACAGAGAAATGATCGAAG AAAGACTGAAGACATACGCACACCTG TTCGACGACAAGGTCATGAAGCAGCT GAAGAGAAGAAGATACACAGGATGGG GAAGACTGAGCAGAAAGCTGATCAAC GGAATCAGAGACAAGCAGAGCGGAAA GACAATCCTGGACTTCCTGAAGAGCG ACGGATTTCGCAAACAGAACTTCATG CAGCTGATCCACGACGACAGCCTGAC ATTCAAGGAAGACATCCAGAAGGCAC AGGTCAGCGGACAGGGAGACAGCCTG CACGAACACATCGCAAACCTGGCAGG AAGCCCGGCAATCAAGAAGGGAATCC TGCAGACAGTCAAGGTCGTGACGAA CTGGTCAAGGTCATGGGAAGACACAA	
		10
		20
		30
		40

	GCCGGAAAACATCGTCATCGAAATGG CAAGAGAAAACCAGACAACACAGAAG GGACAGAAGAACAGCAGAGAAAGAAT GAAGAGAATCGAAGAAGGAATCAAGG AACTGGGAAGCCAGATCCTGAAGGAA CACCCGGTCGAAAACACACAGCTGCA GAACGAAAAGCTGTACCTGTACTACC TGCAGAACGGAAGAGACATGTACGTC GACCAGGAACCTGGACATCAACAGACT GAGCGACTACGACGTCGACCACATCG TCCCGCAGAGCTTCCTGAAGGACGAC AGCATCGACAACAAGGTCCTGACAAG AAGCGACAAGAACAGAGGAAAGAGCG ACAACGTCCCGAGCGAAGAAGTCGTC AAGAAGATGAAGAAGCTACTGGAGACA GCTGCTGAACGCAAAGCTGATCACAC AGAGAAAGTTTCGACAACCTGACAAAG GCAGAGAGAGGAGGACTGAGCGAACT GGACAAGGCAGGATTTCATCAAGAGAC AGCTGGTCGAAACAAGACAGATCACA AAGCACGTCGCACAGATCCTGGACAG CAGAATGAACACAAAGTACGACGAAA ACGACAAGCTGATCAGAGAAGTCAAG GTCATCACACTGAAGAGCAAGCTGGT CAGCGACTTCAGAAAGGACTTCCAGT TCTACAAGGTCAGAGAAATCAACAAC TACCACCACGCACACGACGCATACCT GAACGCAGTCGTCGGAACAGCACTGA TCAAGAAGTACCCGAAGCTGGAAAGC GAATTTCGTCTACGGAGACTACAAGGT CTACGACGTCAGAAAGATGATCGCAA AGAGCGAACAGGAAATCGGAAAGGCA ACAGCAAAGTACTTCTTCTACAGCAA CATCATGAACTTCTTCAAGACAGAAA TCACACTGGCAAACGGAGAAATCAGA AAGAGACCGCTGATCGAAACAAACGG AGAAACAGGAGAAATCGTCTGGGACA AGGGAAGAGACTTCGCAACAGTCAGA AAGGTCCTGAGCATGCCGCAGGTCAA CATCGTCAAGAAGACAGAAGTCCAGA CAGGAGGATTTCAGCAAGGAAAGCATC CTGCCGAAGAGAAACAGCGACAAGCT	
		10
		20
		30
		40

	GATCGCAAGAAAGAAGGACTGGGACC CGAAGAAGTACGGAGGATTTCGACAGC CCGACAGTCGCATACAGCGTCCTGGT CGTCGCAAAGGTCGAAAAGGGAAAGA GCAAGAAGCTGAAGAGCGTCAAGGAA CTGCTGGGAATCACAATCATGGAAAG AAGCAGCTTCGAAAAGAACCCGATCG ACTTCCTGGAAGCAAAGGGATACAAG GAAGTCAAGAAGGACCTGATCATCAA GCTGCCGAAGTACAGCCTGTTTCGAAC TGGAAAACGGAAGAAAGAGAATGCTG GCAAGCGCAGGAGAACTGCAGAAGGG AAACGAACTGGCACTGCCGAGCAAGT ACGTCAACTTCCTGTACCTGGCAAGC CACTACGAAAAGCTGAAGGGAAGCCC GGAAGACAACGAACAGAAGCAGCTGT TCGTGGAACAGCACAAAGCACTACCTG GACGAAATCATCGAACAGATCAGCGA ATTCAGCAAGAGAGTCATCCTGGCAG ACGCAAACCTGGACAAGGTCCTGAGC GCATACAACAAGCACAGAGACAAGCC GATCAGAGAACAGGCAGAAAACATCA TCCACCTGTTTCACACTGACAAACCTG GGAGCACCGGCAGCATTCAAGTACTT CGACACAACAATCGACAGAAAGAGAT ACACAAGCACAAAGGAAGTCCTGGAC GCAACACTGATCCACCAGAGCATCAC AGGACTGTACGAAACAAGAATCGACC TGAGCCAGCTGGGAGGAGACGGAGGA GGAAGCCCGAAGAAGAAGAGAAAGGT CTAGCTAGCCATCACATTTAAAAGCA TCTCAGCCTACCATGAGAATAAGAGA AAGAAAATGAAGATCAATAGCTTATT CATCTCTTTTTTCTTTTTTCGTTGGTGT AAAGCCAACACCCTGTCTAAAAACA TAAATTTCTTTAATCATTTTGCCTCT TTTCTCTGTGCTTCAATTAATAAAAA ATGGAAAGAACCTCGAG	
CMV由来の 5'UTR、配 列番号4に対	GGGCAGATCGCCTGGAGACGCCATCC ACGCTGTTTTTGACCTCCATAGAAGAC ACCGGGACCGATCCAGCCTCCGCGGC	5 7

10

20

30

40

50

応するOR F、コザック 配列、及びA LBの3' U TRを有する Cas 9 転写 産物	CGGGAACGGTGCATTGGAACGCGGAT TCCCCGTGCCAAGAGTGACTCACCGT CCTTGACACGGCCACCATGGACAAGA AGTACAGCATCGGACTGGACATCGGA ACAAACAGCGTCGGATGGGCAGTCAT CACAGACGAATACAAGGTCCCGAGCA AGAAGTTCAAGGTCCTGGGAAACACA GACAGACACAGCATCAAGAAGAACCT GATCGGAGCACTGCTGTTTCGACAGCG GAGAAACAGCAGAAGCAACAAGACTG AAGAGAACAGCAAGAAGAAGATACAC AAGAAGAAAGAACAGAATCTGCTACC TGCAGGAAATCTTCAGCAACGAAATG GCAAAGGTGACGACAGCTTCTTCCA CAGACTGGAAGAAAGCTTCCTGGTCG AAGAAGACAAGAAGCACGAAAGACAC CCGATCTTCGGAAACATCGTCGACGA AGTCGCATACCACGAAAAGTACCCGA CAATCTACCACCTGAGAAAGAAGCTG GTCGACAGCACAGACAAGGCAGACCT GAGACTGATCTACCTGGCACTGGCAC ACATGATCAAGTTTCAGAGGACACTTC CTGATCGAAGGAGACCTGAACCCGGA CAACAGCGACGTGACAAAGCTGTTCA TCCAGCTGGTCCAGACATACAACCAG CTGTTTCGAAGAAAACCCGATCAACGC AAGCGGAGTCGACGCAAAGGCAATCC TGAGCGCAAGACTGAGCAAGAGCAGA AGACTGGAAAACCTGATCGCACAGCT GCCGGGAGAAAAGAAGAACGGACTGT TCGGAAACCTGATCGCACTGAGCCTG GGACTGACACCGAACTTCAAGAGCAA CTTCGACCTGGCAGAAGACGCAAAGC TGCAGCTGAGCAAGGACACATACGAC GACGACCTGGACAACCTGCTGGCACA GATCGGAGACCAGTACGCAGACCTGT TCCTGGCAGCAAAGAACCTGAGCGAC GCAATCCTGCTGAGCGACATCCTGAG AGTCAACACAGAAATCACAAAGGCAC CGCTGAGCGCAAGCATGATCAAGAGA TACGACGAACACCACCAGGACCTGAC ACTGCTGAAGGCACTGGTCAGACAGC	

10

20

30

40

50

AGCTGCCGGAAAAGTACAAGGAAATC	
TTCTTTCGACCAGAGCAAGAACGGATA	
CGCAGGATACATCGACGGAGGAGCAA	
GCCAGGAAGAATTCTACAAGTTTCATC	
AAGCCGATCCTGGAAAAGATGGACGG	
AACAGAAGAAGCTGCTGGTCAAGCTGA	
ACAGAGAAGACCTGCTGAGAAAGCAG	
AGAACATTTCGACAACGGAAGCATCCC	
GCACCAGATCCACCTGGGAGAACTGC	
ACGCAATCCTGAGAAGACAGGAAGAC	
TTCTACCCGTTTCTGAAAGGACAACAG	
AGAAAAGATCGAAAAGATCCTGACAT	
TCAGAATCCCGTACTACGTCCGACCG	
CTGGCAAGAGGAAACAGCAGATTTCGC	
ATGGATGACAAGAAAGAGCGAAGAAA	
CAATCACACCGTGGAAGCTTCGAAGAA	
GTCGTCGACAAGGGAGCAAGCGCACA	
GAGCTTCATCGAAAGAATGACAAACT	
TCGACAAGAACCTGCCGAACGAAAAG	
GTCCTGCCGAAGCACAGCCTGCTGTA	
CGAATACTTCACAGTCTACAACGAAC	
TGACAAAGGTCAAGTACGTACAGAA	
GGAATGAGAAAGCCGGCATTCCTGAG	
CGGAGAACAGAAGAAGGCAATCGTCG	
ACCTGCTGTTCAAGACAAACAGAAAG	
GTCACAGTCAAGCAGCTGAAGGAAGA	
CTACTTCAAGAAGATCGAATGCTTCG	
ACAGCGTCGAAATCAGCGGAGTCGAA	
GACAGATTCAACGCAAGCCTGGGAAC	
ATACCACGACCTGCTGAAGATCATCA	
AGGACAAGGACTTCCTGGACAACGAA	
GAAAACGAAGACATCCTGGAAGACAT	
CGTCCTGACACTGACACTGTTTCGAAG	
ACAGAGAAATGATCGAAGAAAGACTG	
AAGACATACGCACACCTGTTTCGACGA	
CAAGGTCATGAAGCAGCTGAAGAGAA	
GAAGATACACAGGATGGGGAAGACTG	
AGCAGAAAGCTGATCAACGGAATCAG	
AGACAAGCAGAGCGGAAAGACAATCC	
TGGACTTCCTGAAGAGCGACGGATTTC	
GCAAACAGAACTTCATGCAGCTGAT	
CCACGACGACAGCCTGACATTCAAGG	

10

20

30

40

50

	AAGACATCCAGAAGGCACAGGTCAGC GGACAGGGAGACAGCCTGCACGAACA CATCGCAAACCTGGCAGGAAGCCCGG CAATCAAGAAGGGAATCCTGCAGACA GTCAAGGTCGTTCGACGAACTGGTCAA GGTCATGGGAAGACACAAGCCGGAAA ACATCGTCATCGAAATGGCAAGAGAA AACCAGACAACACAGAAGGGACAGAA GAACAGCAGAGAAAGAATGAAGAGAA TCGAAGAAGGAATCAAGGAACTGGGA AGCCAGATCCTGAAGGAACACCCGGT CGAAAACACACAGCTGCAGAACGAAA AGCTGTACCTGTACTACCTGCAGAAC GGAAGAGACATGTACGTTCGACCAGGA ACTGGACATCAACAGACTGAGCGACT ACGACGTTCGACCACATCGTCCCGCAG AGCTTCCTGAAGGACGACAGCATCGA CAACAAGGTCCTGACAAGAAGCGACA AGAACAGAGGAAAGAGCGACAACGTC CCGAGCGAAGAAGTCGTCAAGAAGAT GAAGAAGTACTGGAGACAGCTGCTGA ACGCAAAGCTGATCACACAGAGAAAG TTCGACAACCTGACAAGGCAGAGAG AGGAGGACTGAGCGAACTGGACAAGG CAGGATTTCATCAAGAGACAGCTGGTC GAAACAAGACAGATCACAAAGCACGT CGCACAGATCCTGGACAGCAGAATGA ACACAAAGTACGACGAAAACGACAAG CTGATCAGAGAAGTCAAGGTCATCAC ACTGAAGAGCAAGCTGGTCAGCGACT TCAGAAAGGACTTCCAGTTCTACAAG GTCAGAGAAATCAACAACCTACCACCA CGCACACGACGCATACCTGAACGCAG TCGTTCGGAACAGCACTGATCAAGAAG TACCCGAAGCTGGAAAGCGAATTTCGT CTACGGAGACTACAAGGTCTACGACG TCAGAAAGATGATCGCAAAGAGCGAA CAGGAAATCGGAAAGGCAACAGCAAA GTACTTCTTCTACAGCAACATCATGA ACTTCTTCAAGACAGAAATCACACTG GCAAACGGAGAAATCAGAAAGAGACC GCTGATCGAAACAAACGGAGAAACAG	

10

20

30

40

50

50

40

	TGAGCGCAAGCATGATCAAGAGATAC GACGAACACCACCAGGACCTGACACT GCTGAAGGCACTGGTCAGACAGCAGC TGCCGGAAAAGTACAAGGAAATCTTC TTCGACCAGAGCAAGAACGGATACGC AGGATACATCGACGGAGGAGCAAGCC AGGAAGAATTCTACAAGTTCATCAAG CCGATCCTGGAAAAGATGGACGGAAC AGAAGAAGCTGCTGGTCAAGCTGAACA GAGAAGACCTGCTGAGAAAGCAGAGA ACATTTCGACAACGGAAGCATCCCGCA CCAGATCCACCTGGGAGAACTGCACG CAATCCTGAGAAGACAGGAAGACTTC TACCCGTTCTCTGAAGGACAACAGAGA AAAGATCGAAAAGATCCTGACATTCA GAATCCCGTACTACGTCCGACCGCTG GCAAGAGGAAACAGCAGATTTCGCATG GATGACAAGAAAGAGCGAAGAAACAA TCACACCGTGGAACCTTCGAAGAAGTC GTCGACAAGGGAGCAAGCGCACAGAG CTTCATCGAAAGAATGACAAACTTCG ACAAGAACCTGCCGAACGAAAAGGTC CTGCCGAAGCACAGCCTGCTGTACGA ATACTTTCACAGTCTACAACGAACTGA CAAAGGTCAAGTACGTCACAGAAGGA ATGAGAAAGCCGGCATTCTGAGCGG AGAACAGAAAGAAGGCAATCGTCCGAC TGCTGTTCAAGACAAACAGAAAGGTC ACAGTCAAGCAGCTGAAGGAAGACTA CTTCAAGAAGATCGAATGCTTTCGACA GCGTCGAAATCAGCGGAGTCGAAGAC AGATTCAACGCAAGCCTGGGAACATA CCACGACCTGCTGAAGATCATCAAGG ACAAGGACTTCTCTGGACAACGAAGAA AACGAAGACATCCTGGAAGACATCGT CCTGACACTGACACTGTTTCGAAGACA GAGAAATGATCGAAGAAAGACTGAAG ACATACGCACACCTGTTTCGACGACAA GGTCATGAAGCAGCTGAAGAGAAGAA GATACACAGGATGGGGAAGACTGAGC AGAAAGCTGATCAACGGAATCAGAGA CAAGCAGAGCGGAAAGACAATCCTGG	
		10
		20
		30
		40

	ACTTCCTGAAGAGCGACGGATTTCGCA AACAGAAACTTTCATGCAGCTGATCCA CGACGACAGCCTGACATTCAAGGAAG ACATCCAGAAGGCACAGGTCAGCGGA CAGGGAGACAGCCTGCACGAACACAT CGCAAACCTGGCAGGAAGCCCGGCAA TCAAGAAGGGAATCCTGCAGACAGTC AAGGTCGTGACGAACCTGGTCAAGGT CATGGGAAGACACAAGCCGGAAAACA TCGTTCATCGAAATGGCAAGAGAAAAC CAGACAACACAGAAGGGACAGAAGAA CAGCAGAGAAAGAATGAAGAGAATCG AAGAAGGAATCAAGGAACTGGGAAGC CAGATCCTGAAGGAACACCCGGTCGA AAACACACAGCTGCAGAACGAAAAGC TGTACCTGTACTACCTGCAGAACGGA AGAGACATGTACGTCGACCAGGAACT GGACATCAACAGACTGAGCGACTACG ACGTGCGACCACATCGTCCCGCAGAGC TTCCTGAAGGACGACAGCATCGACAA CAAGGTCCTGACAAGAAGCGACAAGA ACAGAGGAAAGAGCGACAACGTCCCG AGCGAAGAAGTCGTCAAGAAGATGAA GAACTACTGGAGACAGCTGCTGAACG CAAAGCTGATCACACAGAGAAAGTTC GACAACCTGACAAAGGCAGAGAGAGG AGGACTGAGCGAACTGGACAAGGCAG GATTCATCAAGAGACAGCTGGTCGAA ACAAGACAGATCACAAAGCACGTCGC ACAGATCCTGGACAGCAGAATGAACA CAAAGTACGACGAAAACGACAAGCTG ATCAGAGAAGTCAAGGTCATCACACT GAAGAGCAAGCTGGTCAGCGACTTCA GAAAGGACTTCCAGTTCTACAAGGTC AGAGAAATCAACAACCTACCACCACGC ACACGACGCATACCTGAACGCAGTCG TCGGAACAGCACTGATCAAGAAGTAC CCGAAGCTGGAAAGCGAATTCTGTCTA CGGAGACTACAAGGTCTACGACGTCA GAAAGATGATCGCAAAGAGCGAACAG GAAATCGGAAAGGCAACAGCAAAGTA CTTCTTCTACAGCAACATCATGAACT	

10

20

30

40

50

50

	a g g g c c t t g a g c a t c t g g a t t c t g c c t a a t a a a a a c a t t t a t t t t c a t t g c c t c g a g	
XBG由来の 5' UTR、配 列番号4に対 応するORF、 コザック配 列、及びXBG の3' UTRを有 するCas9転写 産物	GGGAagctcagaaataaacgctcaact ttgggccgggatctgcccacCATGGACAA GAAGTACAGCATCGGACTGGACATCG GAACAAACAGCGTCGGATGGGCAGTC ATCACAGACGAATACAAGGTCCCGAG CAAGAAGTTCAAGGTCCTGGGAACA CAGACAGACACAGCATCAAGAAGAAC CTGATCGGAGCACTGCTGTTTCGACAG CGGAGAAACAGCAGAAGCAACAAGAC TGAAGAGAACAGCAAGAAGAAGATAC ACAAGAAGAAAGAACAGAATCTGCTA CCTGCAGGAAATCTTTCAGCAACGAAA TGGCAAAGGTCGACGACAGCTTCTTC CACAGACTTGAAGAAAGCTTCTCTGGT CGAAGAAGACAAGAAGCACGAAAGAC ACCCGATCTTTCGGAAACATCGTCGAC GAAGTCGCATACCACGAAAAGTACCC GACAATCTACCACCTGAGAAAGAAGC TGGTCGACAGCACAGACAAGGCAGAC CTGAGACTGATCTACCTGGCACTGGC ACACATGATCAAGTTTCAGAGGACACT TCCTGATCGAAGGAGACCTGAACCCG GACAACAGCGACGTCGACAAGCTGTT CATCCAGCTGGTCCAGACATACAACC AGCTGTTTCGAAGAAAACCCGATCAAC GCAAGCGGAGTCGACGCAAAGGCAAT CCTGAGCGCAAGACTGAGCAAGAGCA GAAGACTGGAAAACCTGATCGCACAG CTGCCGGGAGAAAAGAAGAACGGACT GTTCGGAAACCTGATCGCACTGAGCC TGGGACTGACACCGAACTTCAAGAGC AACTTCGACCTGGCAGAAGACGCAAA GCTGCAGCTGAGCAAGGACACATACG ACGACGACCTGGACAACCTGCTGGCA CAGATCGGAGACCAGTACGCAGACCT GTTCCTGGCAGCAAAGAACCTGAGCG ACGCAATCCTGCTGAGCGACATCCTG AGAGTCAACACAGAAATCACAAAGGC	59

	ACCGCTGAGCGCAAGCATGATCAAGA GATACGACGAACACCACCAGGACCTG ACACTGCTGAAGGCACTGGTCAGACA GCAGCTGCCGGAAGTACAAGGAAA TCTTCTTCGACCAGAGCAAGAACGGA TACGCAGGATACATCGACGGAGGAGC AAGCCAGGAAGAATTCTACAAGTTCA TCAAGCCGATCCTGGAAAAGATGGAC GGAACAGAAAGAACTGCTGGTCAAGCT GAACAGAGAAGACCTGCTGAGAAAGC AGAGAACATTTCGACAACGGAAGCATC CCGCACCAGATCCACCTGGGAGAACT GCACGCAATCCTGAGAAGACAGGAAG ACTTCTACCCGTTCTCTGAAGGACAAC AGAGAAAAGATCGAAAAGATCCTGAC ATTCAGAATCCCGTACTACGTCGGAC CGCTGGCAAGAGGAAACAGCAGATTCT GCATGGATGACAAGAAAGAGCGAAGA AACAAATCACACCGTGGAACCTTCGAAG AAGTCGTGACAAAGGGAGCAAGCGCA CAGAGCTTCATCGAAAGAATGACAAA CTTCGACAAGAACCTGCCGAACGAAA AGGTCCTGCCGAAGCACAGCCTGCTG TACGAATACTTTCACAGTCTACAACGA ACTGACAAAGGTCAAGTACGTCACAG AAGGAATGAGAAAGCCGGCATTCTCTG AGCGGAGAACAGAAGAAGGCAATCGT CGACCTGCTGTTCAAGACAAACAGAA AGGTCACAGTCAAGCAGCTGAAGGAA GACTACTTCAAGAAGATCGAATGCTT CGACAGCGTCGAAATCAGCGGAGTCG AAGACAGATTCAACGCAAGCCTGGGA ACATACCACGACCTGCTGAAGATCAT CAAGGACAAGGACTTCTCTGGACAACG AAGAAAACGAAGACATCCTGGAAGAC ATCGTCTGACACTGACACTGTTCGA AGACAGAGAAATGATCGAAGAAAGAC TGAAGACATACGCACACCTGTTCGAC GACAAGGTCATGAAGCAGCTGAAGAG AAGAAGATACACAGGATGGGGAAGAC TGAGCAGAAAGCTGATCAACGGAATC AGAGACAAGCAGAGCGGAAAGACAAT	

10

20

30

40

50

	CCTGGACTTCCTGAAGAGCGACGGAT TCGCAAACAGAACTTCATGCAGCTG ATCCACGACGACAGCCTGACATTCAA GGAAGACATCCAGAAGGCACAGGTCA GCGGACAGGGAGACAGCCTGCACGAA CACATCGCAAACCTGGCAGGAAGCCC GGCAATCAAGAAGGGAATCCTGCAGA CAGTCAAGGTCGTCGACGAAC TGGTC AAGGTCATGGGAAGACACAAGCCGGA AAACATCGTCATCGAAATGGCAAGAG AAAACCAGACAACACAGAAGGGACAG AAGAACAGCAGAGAAAGAATGAAGAG AATCGAAGAAGGAATCAAGGAACTGG GAAGCCAGATCCTGAAGGAACACCCG GTCGAAAACACACAGCTGCAGAACGA AAAGCTGTACCTGTACTACCTGCAGA ACGGAAGAGACATGTACGTCGACCAG GAACTGGACATCAACAGACTGAGCGA CTACGACGTCGACCACATCGTCCCGC AGAGCTTCCTGAAGGACGACAGCATC GACAACAAGGTCCTGACAAGAAGCGA CAAGAACAGAGGAAAGAGCGACAACG TCCCGAGCGAAGAAGTCGTCAAGAAG ATGAAGAACTACTGGAGACAGCTGCT GAACGCAAAGCTGATCACACAGAGAA AGTTCGACAAACCTGACAAAGGCAGAG AGAGGAGGACTGAGCGAACTGGACAA GGCAGGATTTCATCAAGAGACAGCTGG TCGAAACAAGACAGATCACAAAGCAC GTCGCACAGATCCTGGACAGCAGAAT GAACACAAAGTACGACGAAAACGACA AGCTGATCAGAGAAGTCAAGGTCATC ACACTGAAGAGCAAGCTGGTCAGCGA CTTCAGAAAGGACTTCCAGTTCTACA AGGTCAGAGAAATCAACAAC TACCAC CACGCACACGACGCATACCTGAACGC AGTCGTCGGAACAGCACTGATCAAGA AGTACCCGAAGCTGGAAAGCGAATTC GTCTACGGAGACTACAAGGTCTACGA CGTCAGAAAGATGATCGCAAAGAGCG AACAGGAAATCGGAAAGGCAACAGCA AAGTACTTCTTCTACAGCAACATCAT	

10

20

30

40

50

	GAAC TTCTTCAAGACAGAAATCACAC TGGCAAACGGAGAAATCAGAAAGAGA CCGCTGATCGAAACAAACGGAGAAAC AGGAGAAATCGTCTGGGACAAGGGAA GAGACTTCGCAACAGTCAGAAAGGTC CTGAGCATGCCGCAGGTCAACATCGT CAAGAAGACAGAAGTCCAGACAGGAG GATTCAGCAAGGAAAGCATCCTGCCG AAGAGAAACAGCGACAAGCTGATCGC AAGAAAGAAGGACTGGGACCCGAAGA AGTACGGAGGATTTCGACAGCCCGACA GTCGCATACAGCGTCCTGGTCGTTCG AAAGGTTCGAAAAGGGAAAGAGCAAGA AGCTGAAGAGCGTCAAGGAACTGCTG GGAATCACAATCATGGAAAGAAGCAG CTTCGAAAAGAACCCGATCGACTTCC TGGAAGCAAAGGGATACAAGGAAGTC AAGAAGGACCTGATCATCAAGCTGCC GAAGTACAGCCTGTTTCGAACTGGAAA ACGGAAGAAAGAGAAATGCTGGCAAGC GCAGGAGAACTGCAGAAGGGAAACGA ACTGGCACTGCCGAGCAAGTACGTCA ACTTCCTGTACCTGGCAAGCCACTAC GAAAAGCTGAAGGGAAGCCCGGAAGA CAACGAACAGAAGCAGCTGTTTCGTTCG AACAGCACAAAGCACTACCTGGACGAA ATCATCGAACAGATCAGCGAATTTCAG CAAGAGAGTCATCCTGGCAGACGCAA ACCTGGACAAGGTCCTGAGCGCATAC AACAAAGCACAGAGACAAGCCGATCAG AGAACAGGCAGAAAACATCATCCACC TGTTTCACACTGACAAACCTGGGAGCA CCGGCAGCATTCAAGTACTTCGACAC AACAAATCGACAGAAAGAGATACACAA GCACAAAGGAAGTCCTGGACGCAACA CTGATCCACCAGAGCATCACAGGACT GTACGAAACAAGAATCGACCTGAGCC AGCTGGGAGGAGACGGAGGAGGAAGC CCGAAGAAGAAGAGAAAGGTCTAGct agcaccagcctcaagaaacccgaat ggagtctctaaagctacataataccaa cttacactttacaaaatgttgtcccc	
		10
		20
		30
		40

	c a a a a t g t a g c c a t t c g t a t c t g c t c c t a a t a a a a a g a a a g t t t c t t c a c a t t c t c t c g a g	
CleanCap (商標) と共に使用するための最初の3ヌクレオチドとしてのAGG、XBG由来の5' UTR、配列番号4に対応するORF、コザック配列、及びXBGの3' UTRを有するCas9転写産物	AGGaagctcagaaataaacgctcaact ttgggccggatctgcccacCATGGACAA GAAGTACAGCATCGGACTGGACATCG GAACAAACAGCGTCGGATGGGCAGTC ATCACAGACGAATACAAGGTCCCGAG CAAGAAGTTCAAGGTCCTGGGAACA CAGACAGACACAGCATCAAGAAGAAC CTGATCGGAGCACTGCTGTTTCGACAG CGGAGAAACAGCAGAAGCAACAAGAC TGAAGAGAACAGCAAGAAGAAGATAC ACAAGAAGAAAGAACAAGAATCTGCTA CCTGCAGGAAATCTTTCAGCAACGAAA TGGCAAAGGTCGACGACAGCTTCTTC CACAGACTTGAAGAAAGCTTCTCTGGT CGAAGAAGACAAGAAGCACGAAAGAC ACCCGATCTTTCGGAAACATCGTCGAC GAAGTCGCATACCACGAAAAGTACCC GACAATCTACCACCTGAGAAAGAAGC TGGTCGACAGCACAGACAAGGCAGAC CTGAGACTGATCTACCTGGCACTGGC ACACATGATCAAGTTTCAGAGGACACT TCCTGATCGAAGGAGACCTGAACCCG GACAACAGCGACGTCGACAAGCTGTT CATCCAGCTGGTCCAGACATACAACC AGCTGTTTCGAAGAAAACCCGATCAAC GCAAGCGGAGTCGACGCAAAGGCAAT CCTGAGCGCAAGACTGAGCAAGAGCA GAAGACTGGAAAACCTGATCGCACAG CTGCCGGGAGAAAAGAAGAACGGACT GTTTCGGAAACCTGATCGCACTGAGCC TGGGACTGACACCGAACTTCAAGAGC AACTTCGACCTGGCAGAAGACGCAAA GCTGCAGCTGAGCAAGGACACATACG ACGACGACCTGGACAACCTGCTGGCA CAGATCGGAGACCAGTACGCAGACCT GTTCTCTGGCAGCAAAGAACCTGAGCG ACGCAATCCTGCTGAGCGACATCCTG AGAGTCAACACAGAAATCACAAAGGC	60

	ACCGCTGAGCGCAAGCATGATCAAGA GATACGACGAACACCACCAGGACCTG ACACTGCTGAAGGCACTGGTCAGACA GCAGCTGCCGGAAGTACAAGGAAA TCTTCTTCGACCAGAGCAAGAACGGA TACGCAGGATACATCGACGGAGGAGC AAGCCAGGAAGAATTCTACAAGTTCA TCAAGCCGATCCTGGAAAAGATGGAC GGAACAGAAAGAACTGCTGGTCAAGCT GAACAGAGAAGACCTGCTGAGAAAGC AGAGAACATTTCGACAACGGAAGCATC CCGCACCAGATCCACCTGGGAGAACT GCACGCAATCCTGAGAAGACAGGAAG ACTTCTACCCGTTCTCTGAAGGACAAC AGAGAAAAGATCGAAAAGATCCTGAC ATTCAGAATCCCGTACTACGTCGGAC CGCTGGCAAGAGGAAACAGCAGATTCT GCATGGATGACAAGAAAGAGCGAAGA AACAAATCACACCGTGGAACCTTCGAAG AAGTCGTTCGACAAGGGAGCAAGCGCA CAGAGCTTCATCGAAAGAATGACAAA CTTCGACAAGAACCTGCCGAACGAAA AGGTCCTGCCGAAGCACAGCCTGCTG TACGAATACTTTCACAGTCTACAACGA ACTGACAAAGGTCAAGTACGTCACAG AAGGAATGAGAAAGCCGGCATTCTCTG AGCGGAGAACAGAAGAAGGCAATCGT CGACCTGCTGTTCAAGACAAACAGAA AGGTCACAGTCAAGCAGCTGAAGGAA GACTACTTCAAGAAGATCGAATGCTT CGACAGCGTCGAAATCAGCGGAGTCG AAGACAGATTCAACGCAAGCCTGGGA ACATACCACGACCTGCTGAAGATCAT CAAGGACAAGGACTTCTCTGGACAACG AAGAAAACGAAGACATCCTGGAAGAC ATCGTCCTGACACTGACACTGTTCGA AGACAGAGAAATGATCGAAGAAAGAC TGAAGACATACGCACACCTGTTCGAC GACAAGGTCATGAAGCAGCTGAAGAG AAGAAGATACACAGGATGGGGAAGAC TGAGCAGAAAGCTGATCAACGGAATC AGAGACAAGCAGAGCGGAAAGACAAT	

10

20

30

40

50

	CCTGGACTTCCTGAAGAGCGACGGAT TCGCAAACAGAACTTCATGCAGCTG ATCCACGACGACAGCCTGACATTCAA GGAAGACATCCAGAAGGCACAGGTCA GCGGACAGGGAGACAGCCTGCACGAA CACATCGCAAACCTGGCAGGAAGCCC GGCAATCAAGAAGGGAATCCTGCAGA CAGTCAAGGTCGTCGACGAACTGGTC AAGGTCATGGGAAGACACAAGCCGGA AAACATCGTCATCGAAATGGCAAGAG AAAACCAGACAACACAGAAGGGACAG AAGAACAGCAGAGAAAGAATGAAGAG AATCGAAGAAGGAATCAAGGAACTGG GAAGCCAGATCCTGAAGGAACACCCG GTCGAAAACACACAGCTGCAGAACGA AAAGCTGTACCTGTACTACCTGCAGA ACGGAAGAGACATGTACGTCGACCAG GAACTGGACATCAACAGACTGAGCGA CTACGACGTCGACCACATCGTCCCGC AGAGCTTCCTGAAGGACGACAGCATC GACAACAAGGTCCTGACAAGAAGCGA CAAGAACAGAGGAAAGAGCGACAACG TCCCGAGCGAAGAAGTCGTCAAGAAG ATGAAGAACTACTGGAGACAGCTGCT GAACGCAAAGCTGATCACACAGAGAA AGTTCGACAACCTGACAAGGCAGAG AGAGGAGGACTGAGCGAACTGGACAA GGCAGGATTTCATCAAGAGACAGCTGG TCGAAACAAGACAGATCACAAAGCAC GTCGCACAGATCCTGGACAGCAGAAT GAACACAAAGTACGACGAAAACGACA AGCTGATCAGAGAAGTCAAGGTCATC ACACTGAAGAGCAAGCTGGTCAGCGA CTTCAGAAAGGACTTCCAGTTCTACA AGGTCAGAGAAATCAACAACCTACCAC CACGCACACGACGCATACCTGAACGC AGTCGTCGGAACAGCACTGATCAAGA AGTACCCGAAGCTGGAAAGCGAATT GTCTACGGAGACTACAAGGTCTACGA CGTCAGAAAGATGATCGCAAAGAGCG AACAGGAAATCGGAAAGGCAACAGCA AAGTACTTCTTCTACAGCAACATCAT	
		10
		20
		30
		40

	GAAC TTCTTCAAGACAGAAATCACAC TGGCAAACGGAGAAATCAGAAAGAGA CCGCTGATCGAAACAAACGGAGAAAC AGGAGAAATCGTCTGGGACAAGGGAA GAGACTTCGCAACAGTCAGAAAGGTC CTGAGCATGCCGCAGGTCAACATCGT CAAGAAGACAGAAGTCCAGACAGGAG GATTCAGCAAGGAAAGCATCCTGCCG AAGAGAAACAGCGACAAGCTGATCGC AAGAAAGAAGGACTGGGACCCGAAGA AGTACGGAGGATTTCGACAGCCCGACA GTCGCATACAGCGTCCTGGTTCGTGCG AAAGGTTCGAAAAGGGAAAGAGCAAGA AGCTGAAGAGCGTCAAGGAACTGCTG GGAATCACAAATCATGGAAAGAAGCAG CTTCGAAAAGAACC CGATCGACTTCC TGGAAGCAAAGGGATACAAGGAAGTC AAGAAGGACCTGATCATCAAGCTGCC GAAGTACAGCCTGTTTCGAACTGGAAA ACGGAAGAAAGAGAATGCTGGCAAGC GCAGGAGAACTGCAGAAGGGAAACGA ACTGGCACTGCCGAGCAAGTACGTCA ACTTCCTGTACCTGGCAAGCCACTAC GAAAAGCTGAAGGGAAGCCCGGAAGA CAACGAACAGAAGCAGCTGTTTCGTG AACAGCACAAAGCACTACCTGGACGAA ATCATCGAACAGATCAGCGAATTCAG CAAGAGAGTCATCCTGGCAGACGCAA ACCTGGACAAGGTCCTGAGCGCATAC AACAAGCACAGAGACAAGCCGATCAG AGAACAGGCAGAAAACATCATCCACC TGTTTCACACTGACAAACCTGGGAGCA CCGGCAGCATTCAAGTACTTCGACAC AACAATCGACAGAAAGAGATACACAA GCACAAAGGAAGTCCTGGACGCAACA CTGATCCACCAGAGCATCACAGGACT GTACGAAACAAGAATCGACCTGAGCC AGCTGGGAGGAGACGGAGGAGGAAGC CCGAAGAAGAAGAGAAAGGTCTAGc t a g c a c c a g c c t c a a g a a c a c c c g a a t g g a g t c t c t a a g c t a c a t a a t a c c a a c t t a c a c t t t a c a a a a t g t t g t c c c c	
		10
		20
		30
		40

	c a a a a t g t a g c c a t t c g t a t c t g c t c c t a a t a a a a g a a a g t t t c t t c a c a t t c t c t c g a g	
C l e a n C a p (商標)と 共に使用する ための最初の 3ヌクレオチ ドとしてのA G G、H S D 由来の5' U T R、配列番 号4に対応す るO R F、コ ザック配列、 及びA L Bの 3' U T Rを 有するC a s 9転写産物	A G G T C C C G C A G T C G G C G T C C A G C G G C T C T G C T T G T T C G T G T G T G T C G T T G C A G G C C T T A T T C G G A T C C G C C A C C A T G G A C A A G A A G T A C A G C A T C G G A C T G G A C A T C G G A A C A A A C A G C G T C G G A T G G G C A G T C A T C A C A G A C G A A T A C A A G G T C C C G A G C A A G A A G T T C A A G G T C C T G G G A A A C A C A G A C A G A C A C A G C A T C A A G A A G A A C C T G A T C G G A G C A C T G C T G T T C G A C A G C G G A G A A A C A G C A G A A G C A A C A A G A C T G A A G A G A A C A G C A A G A A G A A G A T A C A C A A G A A G A A A G A A C A G A A T C T G C T A C C T G C A G G A A A T C T T C A G C A A C G A A A T G G C A A A G G T C G A C G A C A G C T T C T T C C A C A G A C T G G A A G A A A G C T T C C T G G T C G A A G A A G A C A A G A A G C A C G A A A G A C A C C C G A T C T T C G G A A A C A T C G T C G A C G A A G T C G C A T A C C A C G A A A A G T A C C C G A C A A T C T A C C A C C T G A G A A A G A A G C T G G T C G A C A G C A C A G A C A A G G C A G A C C T G A G A C T G A T C T A C C T G G C A C T G G C A C A C A T G A T C A A G T T C A G A G G A C A C T T C C T G A T C G A A G G A G A C C T G A A C C C G G A C A A C A G C G A C G T C G A C A A G C T G T T C A T C C A G C T G G T C C A G A C A T A C A A C C A G C T G T T C G A A G A A A A C C C G A T C A A C G C A A G C G G A G T C G A C G C A A A G G C A A T C C T G A G C G C A A G A C T G A G C A A G A G C A G A A G A C T G G A A A A C C T G A T C G C A C A G C T G C C G G G A G A A A A G A A G A A C G G A C T G T T C G G A A A C C T G A T C G C A C T G A G C C T G G G A C T G A C A C C G A A C T T C A A G A G C A A C T T C G A C C T G G C A G A A G A C G C A A A G C T G C A G C T G A G C A A G G A C A C A T A C G A C G A C G A C C T G G A C A A C C T G C T G G C A C A G A T C G G A G A C C A G T A C G C A G A C C T G T T C C T G G C A G C A A A G A A C C T G A G C G A C G C A A T C C T G C T G A G C G A C	61

10

20

30

40

50

ATCCTGAGAGTCAACACAGAAATCAC AAAGGCACCGCTGAGCGCAAGCATGA TCAAGAGATACGACGAACACCACCAG GACCTGACACTGCTGAAGGCACTGGT CAGACAGCAGCTGCCGGAAAAGTACA AGGAAATCTTCTTCGACCAGAGCAAG AACGGATACGCAGGATACATCGACGG AGGAGCAAGCCAGGAAGAATTCTACA AGTTCATCAAGCCGATCCTGGAAAAG ATGGACGGAACAGAGAAGACCTGCTGA CAAGCTGAACAGAGAAGACCTGCTGA GAAAGCAGAGAACATTTCGACAACGGA AGCATCCCCGCACCAGATCCACCTGGG AGAACTGCACGCAATCCTGAGAAGAC AGGAAGACTTCTACCCGTTCTCTGAAG GACAACAGAGAAAAGATCGAAAAGAT CCTGACATTTCAGAATCCCGTACTACG TCGGACCGCTGGCAAGAGGAAACAGC AGATTCGCATGGATGACAAGAAAGAG CGAAGAAACAATCACACCGTGGAAC TTCGAAGAAGTCGTCGACAAGGGAGCA AGCGCACAGAGCTTCATCGAAAGAAT GACAAACTTCGACAAGAACCTGCCGA ACGAAAAGGTCCTGCCGAAGCACAGC CTGCTGTACGAATACTTCACAGTCTA CAACGAACTGACAAAGGTCAAGTACG TCACAGAAGGAATGAGAAAGCCGGCA TTCCTGAGCGGAGAACAGAAGAAGGC AATCGTCGACCTGCTGTTCAAGACAA ACAGAAAGGTCACAGTCAAGCAGCTG AAGGAAGACTACTTCAAGAAGATCGA ATGCTTTCGACAGCGTCGAAATCAGCG GAGTCGAAGACAGATTCAACGCAAGC CTGGGAACATAACACGACCTGCTGAA GATCATCAAGGACAAGGACTTCCTGG ACAACGAAGAAAACGAAGACATCCTG GAAGACATCGTCCTGACACTGACACT GTTTCGAAGACAGAGAAATGATCGAAG AAAGACTGAAGACATACGCACACCTG TTCGACGACAAGGTCATGAAGCAGCT GAAGAGAAGAAGATACACAGGATGGG GAAGACTGAGCAGAAAGCTGATCAAC	

10

20

30

40

50

	GGAATCAGAGACAAGCAGAGCGGAAA GACAATCCTGGACTTCCTGAAGAGCG ACGGATTTCGCAAACAGAAACTTCATG CAGCTGATCCACGACGACAGCCTGAC ATTCAAGGAAGACATCCAGAAGGCAC AGGTCAGCGGACAGGGAGACAGCCTG CACGAACACATCGCAAACCTGGCAGG AAGCCCGGCAATCAAGAAGGGAATCC TGCAGACAGTCAAGGTCGTCGACGAA CTGGTCAAGGTCATGGGAAGACACAA GCCGGAAAACATCGTCATCGAAATGG CAAGAGAAAACCAGACAACACAGAAG GGACAGAAGAACAGCAGAGAAAGAAT GAAGAGAATCGAAGAAGGAATCAAGG AACTGGGAAGCCAGATCCTGAAGGAA CACCCGGTCGAAAACACACAGCTGCA GAACGAAAAGCTGTACCTGTACTACC TGCAGAACGGGAAGAGACATGTACGTC GACCAGGAACTGGACATCAACAGACT GAGCGACTACGACGTCGACCACATCG TCCCGCAGAGCTTCCTGAAGGACGAC AGCATCGACAACAAGGTCCTGACAAG AAGCGACAAGAACAGAGGAAAGAGCG ACAACGTCCCGAGCGAAGAAGTCGTC AAGAAGATGAAGAACTACTGGAGACA GCTGCTGAACGCAAAGCTGATCACAC AGAGAAAGTTTCGACAACCTGACAAAG GCAGAGAGAGGAGGACTGAGCGAACT GGACAAGGCAGGATTTCATCAAGAGAC AGCTGGTTCGAAACAAGACAGATCACA AAGCACGTCGCACAGATCCTGGACAG CAGAATGAACACAAAGTACGACGAAA ACGACAAGCTGATCAGAGAAGTCAAG GTCATCACACTGAAGAGCAAGCTGGT CAGCGACTTCAGAAAGGACTTCCAGT TCTACAAGGTCAGAGAAATCAACAAC TACCACCACGCACACGACGCATACCT GAACGCAGTCGTCGGAACAGCACTGA TCAAGAAGTACCCGAAGCTGGAAAGC GAATTTCGTCTACGGAGACTACAAGGT CTACGACGTCAGAAAGATGATCGCAA AGAGCGAACAGGAAATCGGAAAGGCA	

10

20

30

40

50

50

	AAGAAAATGAAGATCAATAGCTTATT CATCTCTTTTTTCTTTTTTCGTTGGTGT AAAGCCAACACCCTGTCTAAAAAACA TAAATTTCTTTAATCATTTTGCCTCT TTTCTCTGTGCTTCAATTAATAAAAA ATGGAAAGAACCTCGAG	
30/30/ 39ポリA配 列	AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAGCGAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAACCGAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA A	62
ポリA100 配列	AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAA	63
G209ガイ ドRNA	mC*mC*mA*GUCCAGCGAGGCAAAGG GUUUUAGAGCUAGAAAUAGCAAGUUA AAAUAAAGGCUAGUCCGUUAUCAACUU GAAAAAGUGGCACCGAGUCGGUGCmU *mU*mU*U	64
Neisseria meningitidis Cas9をコー ドするORF	ATGGCAGCATTCAAGCCGAACCTCGAT CAACTACATCCTGGGACTGGACATCG GAATCGCATCGGTCGGATGGGCAATG GTCGAAATCGACGAAGAAGAAAACCC GATCAGACTGATCGACCTGGGAGTCA GAGTCTTCGAAAGAGCAGAAGTCCCG AAGACAGGAGACTCGCTGGCAATGGC AAGAAGACTGGCAAGATCGGTCAGAA GACTGACAAGAAGAAGAGCACACAGA CTGCTGAGAACAAGAAGACTGCTGAA GAGAGAAGGAGTCTCTGCAGGCAGCAA ACTTCGACGAAAACGGACTGATCAAG TCGCTGCCGAACACACCGTGGCAGCT GAGAGCAGCAGCACTGGACAGAAAGC TGACACCGCTGGAATGGTCGGCAGTC CTGCTGCACCTGATCAAGCACAGAGG ATACCTGTTCGCAGAGAAAGAACGAAG GAGAAACAGCAGACAAGGAACTGGGA GCACTGCTGAAGGGAGTCGCAGGAAA CGCACACGCACTGCAGACAGGAGACT TCAGAACACCGGCAGAACTGGCACTG	65

10

20

30

40

50

	AACAAGTTTCGAAAAGGAATCGGGACA CATCAGAAACCAGAGATCGGACTACT CGCACACATTCTCGAGAAAGGACCTG CAGGCAGAACTGATCCTGCTGTTCGA AAAGCAGAAGGAATTCGGAAACCCGC ACGTCTCGGGAGGACTGAAGGAAGGA ATCGAAACACTGCTGATGACACAGAG ACCGGCACTGTCTGGGAGACGCAGTCC AGAAGATGCTGGGACACTGACATTC GAACCCGGCAGAACCGAAGGCAGCAAA GAACACATACACAGCAGAAAGATTCA TCTGGCTGACAAAGCTGAACAACCTG AGAATCCTGGAACAGGGATCGGAAAG ACCGCTGACAGACACAGAAAGAGCAA CACTGATGGACGAACCGTACAGAAAG TCGAAGCTGACATACGCACAGGCAAG AAAGCTGCTGGGACTGGAAGACACAG CATTCTTCAAGGGACTGAGATACGGA AAGGACAACGCAGAAGCATCGACACT GATGGAAATGAAGGCATACCACGCAA TCTCGAGAGCACTGGAAAAGGAAGGA CTGAAGGACAAGAAGTCGCCGCTGAA CCTGTCTGCCGGAACCTGCAGGACGAAA TCGGAACAGCATTCTCGCTGTTCAAG ACAGACGAAGACATCACAGGAAGACT GAAGGACAGAATCCAGCCGGAAATCC TGGAAGCACTGCTGAAGCACATCTCG TTCGACAAGTTCGTCCAGATCTCGCT GAAGGCACTGAGAAGAATCGTCCCGC TGATGGAACAGGGAAAGAGATACGAC GAAGCATGCGCAGAAATCTACGGAGA CCACTACGGAAAGAAGAACACAGAAG AAAAGATCTACCTGCCGCCGATCCCG GCAGACGAAATCAGAAACCCGGTCGT CCTGAGAGCACTGTCTGCAGGCAAGAA AGGTCATCAACGGAGTCGTCAAGAA TACGGATCGCCGGCAAGAATCCACAT CGAAACAGCAAGAGAAGTCGGAAAGT CGTTCAAGGACAGAAAGGAAATCGAA AAGAGACAGGAAGAAAACAGAAAGGA CAGAGAAAAGGCAGCAGCAAAGTTCA GAGAATACTTCCCGAACTTCGTCTGGA	
		10
		20
		30
		40

	GAACCGAAGTCGAAGGACATCCTGAA GCTGAGACTGTACGAACAGCAGCACG GAAAGTGCCCTGTACTCGGGAAAGGAA ATCAACCTGGGAAGACTGAACGAAAA GGGATACGTCGAAATCGACCACGCAC TGCCGTTCTCGAGAACATGGGACGAC TCGTTCAACAACAAGGTCCTGGTCCT GGGATCGGAAAACCAGAACAAGGGAA ACCAGACACCGTACGAATACTTCAAC GGAAAGGACAACCTCGAGAGAATGGCA GGAATTCAAGGCAAGAGTCGAAACAT CGAGATTCCCGAGATCGAAGAAGCAG AGAATCCTGCTGCGAGAAGTTCGACGA AGACGGATTCAAGGAAAGAAACCTGA ACGACACAAGATACGTCAACAGATTTC CTGTGCCAGTTTCGTGCGACAGACAAT GAGACTGACAGGAAAGGGAAAGAAGA GAGTCTTCGCATCGAACGGACAGATC ACAAACCTGCTGAGAGGATTCTGGGG ACTGAGAAAGGTCAGAGCAGAAAACG ACAGACACCACGCACTGGACGCAGTC GTCGTGCGCATGCTCGACAGTCGCAAT GCAGCAGAAGATCACAAAGATTCTGTCA GATACAAGGAAATGAACGCATTTCGAC GGAAAGACAATCGACAAGGAAACAGG AGAAGTCCTGCACCAGAAGACACACT TCCCGCAGCCGTGGGAATTCTTCGCA CAGGAAGTCATGATCAGAGTCTTCGG AAAGCCGGACGGAAAGCCGGAATTCTG AAGAAGCAGACACACTGGAAAAGCTG AGAACACTGCTGGCAGAAAAGCTGTC GTCGAGACCGGAAGCAGTCCACGAAT ACGTCACACCGCTGTTCTGTCTCGAGA GCACCGAACAGAAAGATGTCGGGACA GGGACACATGGAAACAGTCAAGTCGG CAAAGAGACTGGACGAAGGAGTCTCG GTCCTGAGAGTCCCGCTGACACAGCT GAAGCTGAAGGACCTGGAAAAGATGG TCAACAGAGAAAGAGAACCGAAGCTG TACGAAGCACTGAAGGCAAGACTGGA AGCACACAAGGACGACCCGGCAAAGG CATTCTGCAGAACC GTTCTACAAGTAC	

10

20

30

40

50

	GACAAGGCAGGAAACAGAACACAGCA GGTCAAGGCAGTCAGAGTCGAACAGG TCCAGAAGACAGGAGTCTGGGTCAGA AACCACAACGGAATCGCAGACAACGC AACAAATGGTCAGAGTAGACGTCTTCG AAAAGGGAGACAAGTACTACCTGGTC CCGATCTACTCGTGGCAGGTTCGCAA GGGAATCCTGCCGGACAGAGCAGTCG TCCAGGGAAAGGACGAAGAAGACTGG CAGCTGATCGACGACTCGTTCAACTT CAAGTTCTCGCTGCACCCGAACGACC TGGTCGAAGTCATCACAAAGAAGGCA AGAATGTTTCGGATACTTCGCATCGTG CCACAGAGGAAACAGGAAACATCAACA TCAGAATCCACGACCTGGACCACAAG ATCGGAAAGAACGGAATCCTGGAAGG AATCGGAGTCAAGACAGCACTGTCTGT TCCAGAAGTACCAGATCGACGAACTG GGAAAGGAAATCAGACCGTGCAGACT GAAGAAGAGACCGCCGGTCAGATCCG GAAAGAGAACAGCAGACGGATCGGAA TTCGAATCGCCGAAGAAGAAGAGAAA GGTCGAATGA		10
Neisseria meningitidis Cas9をコードするORF (開始コドンまたは終止コドンはない; 融合タンパク質コード配列に含める際に好適)	GCAGCATTCAAGCCGAACTCGATCAA CTACATCCTGGGACTGGACATCGGAA TCGCATCGGTTCGGATGGGCAATGGTC GAAATCGACGAAGAAGAAAACCCGAT CAGACTGATCGACCTGGGAGTCAGAG TCTTCGAAAGAGCAGAAGTCCCGAAG ACAGGAGACTCGCTGGCAATGGCAAG AAGACTGGCAAGATCGGTTCAGAAGAC TGACAAGAAGAAGAGCACACAGACTG CTGAGAACAAGAAGACTGCTGAAGAG AGAAGGAGTCCCTGCAGGCAGCAAAC TCGACGAAAACGGACTGATCAAGTCG CTGCCGAACACACCGTGGCAGCTGAG AGCAGCAGCACTGGACAGAAAGCTGA CACCGCTGGAATGGTTCGGCAGTCCCTG CTGCACCTGATCAAGCACAGAGGATA CCTGTTCGCAGAGAAAGAACGAAGGAG AAACAGCAGACAAGGAACTGGGAGCA CTGCTGAAGGGAGTCGCAGGAAACGC	66	30
			40

50

AGAAAAGGCAGCAGCAAAGTTCAGAG	
AATACTTCCCGAACTTCGTCCGAGAA	
CCGAAGTCGAAGGACATCCTGAAGCT	
GAGACTGTACGAACAGCAGCACGGAA	
AGTGCCCTGTACTCGGGAAAGGAAATC	
AACCTGGGAAGACTGAACGAAAAGGG	
ATACGTTCGAAATCGACCACGCACTGC	
CGTTCTCGAGAACATGGGACGACTCG	
TTCAACAACAAGGTCCTGGTCCTGGG	
ATCGGAAAACCAGAACAAAGGGAAACC	
AGACACCGTACGAATACTTCAACGGA	
AAGGACAACCTCGAGAGAATGGCAGGA	
ATTCAAGGCAAGAGTCGAAACATCGA	
GATTCCCGAGATCGAAGAAGCAGAGA	
ATCCTGCTGCAGAAGTTCGACGAAGA	
CGGATTCAAGGAAAGAAACCTGAACG	
ACACAAGATACGTCAACAGATTCTCTG	
TGCCAGTTCGTTCGCAGACAGAATGAG	
ACTGACAGGAAAGGGAAAGAAGAGAG	
TCTTCGCATCGAACGGACAGATCACA	
AACCTGCTGAGAGGATTCTGGGGACT	
GAGAAAGGTCAGAGCAGAAAACGACA	
GACACCACGCACTGGACGCAGTCGTC	
GTCGCATGCTCGACAGTCGCAATGCA	
GCAGAAGATCACAAAGATTTCGTCAGAT	
ACAAGGAAATGAACGCATTTCGACGGA	
AAGACAATCGACAAGGAAACAGGAGA	
AGTCCTGCACCAGAAGACACACTTCC	
CGCAGCCGTGGGAATTCTTCGCACAG	
GAAGTCATGATCAGAGTCTTCGGAAA	
GCCGGACGGAAAGCCGGAATTTCGAAG	
AAGCAGACACACTGGAAAAGCTGAGA	
ACACTGCTGGCAGAAAAGCTGTTCGTC	
GAGACCGGAAGCAGTCCACGAATACG	
TCACACCGCTGTTCGTCTCGAGAGCA	
CCGAACAGAAAGATGTCGGGACAGGG	
ACACATGGAAACAGTCAAGTCGGCAA	
AGAGACTGGACGAAGGAGTCTCGGTC	
CTGAGAGTCCCGCTGACACAGCTGAA	
GCTGAAGGACCTGGAAAAGATGGTCA	
ACAGAGAAAGAGAAACCGAAGCTGTAC	
GAAGCACTGAAGGCAAGACTGGAAGC	

10

20

30

40

50

	ACACAAGGACGACCCGGCAAAGGCAT TCGCAGAACCGTTCTACAAGTACGAC AAGGCAGGAAACAGAACACAGCAGGT CAAGGCAGTCAGAGTCGAACAGGTCC AGAAGACAGGAGTCTGGGTCAGAAAC CACAACGGAATCGCAGACAACGCAAC AATGGTCAGAGTAGACGTCTTCGAAA AGGGAGACAAAGTACTACCTGGTCCCG ATCTACTCGTGGCAGGTGCAAAGGG AATCCTGCCGGACAGAGCAGTCGTCC AGGGAAAGGACGAAGAAGACTGGCAG CTGATCGACGACTCGTTCAACTTCAA GTTCTCGCTGCACCCGAACGACCTGG TCGAAGTCATCACAAAGAAGGCAAGA ATGTTTCGGATACTTCGCATCGTGCCA CAGAGGAACAGGAAACATCAACATCA GAATCCACGACCTGGACCACAAGATC GGAAAGAACGGAATCCTGGAAGGAAT CGGAGTCAAGACAGCACTGTCGTTCC AGAAGTACCAGATCGACGAACTGGGA AAGGAAATCAGACCGTGCAGACTGAA GAAGAGACCGCCGGTCAGATCCGGAA AGAGAACAGCAGACGGATCGGAATTC GAATCGCCGAAGAAGAAGAGAAAGGT CGAA	
配列番号 65 を含む転写産 物 (Neis s e r i a m e n i n g i t i d i s C a s 9 をコ ードする)	GGGAGACCCAAGCTGGCTAGCGTTTA AACTTAAGCTTGGATCCGCCACCATG GCAGCATTCAGCCGAACCTCGATCAA CTACATCCTGGGACTGGACATCGGAA TCGCATCGGTTCGGATGGGCAATGGTC GAAATCGACGAAGAAGAAAACCCGAT CAGACTGATCGACCTGGGAGTCAGAG TCTTCGAAAGAGCAGAAGTCCCGAAG ACAGGAGACTCGCTGGCAATGGCAAG AAGACTGGCAAGATCGGTCAGAAGAC TGACAAGAAGAAGAGCACACAGACTG CTGAGAACAAGAAGACTGCTGAAGAG AGAAGGAGTCCTGCAGGCAGCAAACCT TCGACGAAAACGGACTGATCAAGTCG CTGCCGAACACACCGTGGCAGCTGAG AGCAGCAGCACTGGACAGAAAGCTGA CACCGCTGGAATGGTCGGCAGTCCTG	67

10

20

30

40

50

CTGCACCTGATCAAGCACAGAGGATA CCTGTCTGCAGAGAAAGAACGAAGGAG AAACAGCAGACAAGGAACTGGGAGCA CTGCTGAAGGGAGTCGCAGGAAACGC ACACGCACTGCAGACAGGAGACTTCA GAACACCGGCAGAACTGGCACTGAAC AAGTTTCGAAAAGGAATCGGGACACAT CAGAAACCAGAGATCGGACTACTCGC ACACATTCTCGAGAAAGGACCTGCAG GCAGAACTGATCCTGCTGTTTCGAAAA GCAGAAAGGAATTTCGGAAACCCGCACG TCTCGGGAGGACTGAAGGAAGGAATC GAAACACTGCTGATGACACAGAGACC GGCACTGTCTGGGAGACGCAGTCCAGA AGATGCTGGGACACTGCACATTCGAA CCGGCAGAAACCGAAGGCAGCAAAGAA CACATACACAGCAGAAAGATTTCATCT GGCTGACAAAGCTGAACAACCTGAGA ATCCTGGAACAGGGATCGGAAAGACC GCTGACAGACACAGAAAGAGCAACAC TGATGGACGAACCGTACAGAAAGTCG AAGCTGACATACGCACAGGCAAGAAA GCTGCTGGGACTGGAAGACACAGCAT TCTTCAAGGGACTGAGATACGGAAAG GACAACGCAGAAGCATCGACACTGAT GGAAATGAAGGCATACCACGCAATCT CGAGAGCACTGGAAAAGGAAGGACTG AAGGACAAGAAGTCGCCGCTGAACCT GTCGCCGGAACCTGCAGGACGAAATCG GAACAGCATTTCTCGCTGTTCAAGACA GACGAAGACATCACAGGAAGACTGAA GGACAGAATCCAGCCGGAAATCCTGG AAGCACTGCTGAAGCACATCTCGTTC GACAAGTTTCGTCCAGATCTCGCTGAA GGCACTGAGAAGAATCGTCCCGCTGA TGGAACAGGGAAAGAGATACGACGAA GCATGCGCAGAAATCTACGGAGACCA CTACGGAAAGAAGAACACAGAAGAAA AGATCTACCTGCCGCCGATCCCGGCA GACGAAATCAGAAACCCGGTTCGTCT GAGAGCACTGTCTGCAGGCAAGAAAGG TCATCAACGGAGTCGTCTAGAAGATAC	

10

20

30

40

50

	GGATCGCCGGCAAGAATCCACATCGA AACAGCAAGAGAAGTCGGAAAGTCGT TCAAGGACAGAAAGGAAATCGAAAAG AGACAGGAAGAAAACAGAAAGGACAG AGAAAAGGCAGCAGCAAAGTTCAGAG AATACTTCCCGAACTTCGTTCGGAGAA CCGAAGTCGAAGGACATCCTGAAGCT GAGACTGTACGAACAGCAGCACGGAA AGTGCCCTGTACTCGGGAAAGGAAATC AACCTGGGAAGACTGAACGAAAAGGG ATACGTCGAAATCGACCAACGCACTGC CGTTCTCGAGAACATGGGACGACTCG TTCAACAACAAGGTCCTGGTCCTGGG ATCGGAAAACCAGAACAAGGGAAACC AGACACCGTACGAATACTTCAACGGA AAGGACAACCTCGAGAGAATGGCAGGA ATTCAAGGCAAGAGTCGAAACATCGA GATTCCCGAGATCGAAGAAGCAGAGA ATCCTGCTGCAGAAGTTCGACGAAGA CGGATTCAAGGAAAGAAACCTGAACG ACACAAGATACGTCAACAGATTCTCTG TGCCAGTTCGTTCGCAGACAGAATGAG ACTGACAGGAAAGGGAAAGAAGAGAG TCTTCGCATCGAACGGACAGATCACA AACCTGCTGAGAGGATTCTGGGGACT GAGAAAGGTCAGAGCAGAAAACGACA GACACCACGCACTGGACGCAGTCGTC GTCGCATGCTCGACAGTCGCAATGCA GCAGAAGATCACAAGATTCTGTCAGAT ACAAGGAAATGAACGCATTTCGACGGA AAGACAATCGACAAGGAAACAGGAGA AGTCCTGCACCAGAAGACACACTTCC CGCAGCCGTGGGAATTCTTCGCACAG GAAGTCATGATCAGAGTCTTCGGAAA GCCGGACGGAAAGCCGGAATTTCGAAG AAGCAGACACACTGGAAAAGCTGAGA ACACTGCTGGCAGAAAAGCTGTCGTC GAGACCGGAAGCAGTCCACGAATACG TCACACCGCTGTTCGTCTCGAGAGCA CCGAACAGAAAGATGTTCGGGACAGGG ACACATGGAAACAGTCAAGTCGGCAA AGAGACTGGACGAAGGAGTCTCGGTC	

10

20

30

40

50

	CTGAGAGTCCCGCTGACACAGCTGAA GCTGAAGGACCTGGAAAAGATGGTCA ACAGAGAAAGAGAAACCGAAGCTGTAC GAAGCACTGAAGGCAAGACTGGAAGC ACACAAGGACGACCCGGCAAAGGCAT TCGCAGAACCGTTCTACAAGTACGAC AAGGCAGGAAACAGAACACAGCAGGT CAAGGCAGTCAGAGTCGAACAGGTCC AGAAGACAGGAGTCTGGGTCAGAAAC CACAACGGAATCGCAGACAACGCAAC AATGGTTCAGAGTAGACGTCTTCGAAA AGGGAGACAAGTACTACCTGGTCCCG ATCTACTCGTGGCAGGTGCGAAAGGG AATCCTGCCGGACAGAGCAGTCGTCC AGGGAAAGGACGAAGAAGACTGGCAG CTGATCGACGACTCGTTCAACTTCAA GTTCTCGCTGCACCCGAACGACCTGG TCGAAGTCATCACAAAGAAGGCAAGA ATGTTTCGGATACTTCGCATCGTGCCA CAGAGGAACAGGAAACATCAACATCA GAATCCACGACCTGGACCACAAGATC GGAAAGAACGGAATCCTGGAAGGAAT CGGAGTCAAGACAGCACTGTTCGTCC AGAAGTACCAGATCGACGAACTGGGA AAGGAAATCAGACCGTGCAGACTGAA GAAGAGACCGCCGGTCAGATCCGGAA AGAGAACAGCAGACGGATCGGAATT GAATCGCCGAAGAAGAAGAGAAAGGT CGAATGATAGCTAGCTCGAGTCTAGA GGGCCCGTTTAAACCCGCTGATCAGC CTCGACTGTGCCTTCTAGTTGCCAGC CATCTGTTGTTTGCCCCCTCCCCCGTG CCTTCCTTGACCCCTGGAAGGTGCCAC TCCCACCTGTCCTTTCCTAATAAAATG AGGAAATTGCATCGCATTGTCTGAGT AGGTGTCAATTCTATTCTGGGGGGTGG GGTGGGGCAGGACAGCAAGGGGGAGG ATTGGGAAGACAATAGCAGGCATGCT GGGGATGCGGTGGGCTCTATGG	
Neisse ria me ningit	MAAFKPNSINYILGLDIGIASVGWAM VEIDEEENPIRLIDLGVRFERAEVP KTGDSLAMARRLARSVRRLTRRRHR	68

10

20

30

40

50

i d i s C a s 9 の ア ミ ノ 酸 配 列	LLRTRRLLKREGVLQAANFDENGLIK SLPNTPWQLRAAALDRKLTPLEWSAV LLHLIKHRGYLSQRKNEGETADKELG ALLKGVAGNAHALQTGDFRTPAELAL NKFEKESGHIRNQSRSDYSHTFSRKDL QAELILLFEKQKEFGNPHVSGGLKEG IETLLMTQRPALSGDAVQKMLGHCTF EPAEPKAAKNTYTAERFIWLTKLNNL RILEQQSERPLTDTERATLMDEPYRK SKLTYAQARKLLGLEDTAFFKGLRYG KDNAEASTLMEMKAYHAISRALKKEG LKDKKSPNLSPELQDEIGTAFSLFK TDEDITGRLKDRIQPEILEALLKHIS FDKFBVQISLKALRRIVPLMEQGKRYD EACAEIYGDHYGKKNTEEKIYLPPI ADEIRNPVVLRALSQARKVINGVRR YGSPARIHIETAREVGKSFKDRKEIE KRQEENRKDREKAAAKFREYFPNFVG EPKSKDILKLRLYEQQHGKCLYSGKE INLGRNLNEKGYVEIDHALPFSRTWDD SFNNKVLVLGSENQNKGNQTPYEYFN GKDNSREWQEFKARVETSRFPRSKKQ RILLQKFDEDDGFKERNLNDTRYVNR LCQFVADRMRLTGKGKKRVFASNGQI TNLLRGFWGLRKVRAENDRHHALDAV VVACSTVAMQQKITRFVRYKEMNAFD GKTIDKETGEVLHQKTHFPQPWEFFA QEVMI R VFGKPDGKPEFEEADTLEKL RTL LAEKLSSRPEAVHEYVTPLFVSR APNRKMSGQGHEMETVKSARKRLDEGVS VLRVPLTQLKLKDLEKMVNREREP KL YEALKARLEAHKDDPAKAFAPFYKY DKAGNRTQQVKAVRVEQVQKTGVWVR NHNGIADNATMVRVDVFEKGDYKYYLV PIYSWQVAKGILPDRAVVQKDEEDW QLIDDSFNFKFSLHPNDLVEVITKKA RMFGYFASCHRG TG N I N I R I H D L D H K I G K N G I L E G I G V K T A L S F Q K Y Q I D E L G K E I R P C R L K K R P P V R S G K R T A D G S E F E S P K K K R K V E	
G 3 9 0 ガ イ ド RNA	mG*mC*mC*GAGUCUGGAGAGCUGCA GUUUUAGAmGmCmUmAmGmAmAmAmU	69

10

20

30

40

50

	mAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmA mAmGmUmGmGmCmAmCmCmGmAmGmU mCmGmGmUmGmCmU*mU*mU*mU	
G 5 0 2 ガイ ドRNA	mA*mC*mA*CAAAUACCAGUCCAGCG GUUUUAGAmGmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmA mAmGmUmGmGmCmAmCmCmGmAmGmU mCmGmGmUmGmCmU*mU*mU*mU	7 0
G 5 0 9 ガイ ドRNA	mA*mA*mA*GUUCUAGAUGCCGUCCG GUUUUAGAmGmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmA mAmGmUmGmGmCmAmCmCmGmAmGmU mCmGmGmUmGmCmU*mU*mU*mU	7 1
G 5 3 4 ガイ ドRNA	mA*mC*mG*CAAAUAUCAGUCCAGCG GUUUUAGAmGmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmA mAmGmUmGmGmCmAmCmCmGmAmGmU mCmGmGmUmGmCmU*mU*mU*mU	7 2

* = P S 結合 ; 「m」 = 2' - O - Meヌクレオチド

マウスの G 0 0 0 2 8 2 NGSプライマー配列

フォワードプライマー :

CACTCTTTCCCTACACGACGCTCTTCCGATCTGTTTTTGT
CCAGAGTCTATCACCG

リバースプライマー :

GGAGTTCAGACGTGTGCTCTTCCGATCTACACGAATAAGA
GCAAATGGGAAC

ラット G 0 0 0 3 9 0 NGSプライマー配列

フォワードプライマー :

CACTCTTTCCCTACACGACGCTCTTCCGATCTTGCAATTT
ATGAGACCGAAAACA

リバースプライマー :

GGAGTTCAGACGTGTGCTCTTCCGATCTGCTACAGTAGAG
CTGTACATAAACTT

GFP配列 :

TCGCGCGTTTTCGGTGATGACGGTGAAAACCTCTGACACAT
GCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGAT
GCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTG
TTGGCGGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGA
GCAGATTGTACTGAGAGTGCAACCATATGCGGTGTGAAATA
CCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCC
ATTCGCCATTTCAGGCTGCGCAACTGTTGGGAAGGGCGATC

10

20

30

40

50

GGTGC GGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGG
GGATGTGCTGCAAGGCGATTAAAGTTGGGTAAACGCCAGGGT
TTTCCCAGTCACGACGTTGTAAACGACGGCCAGTGAATT
CTAATACGACTCACTATAGGGTCCCCGCAGTCGGCGTCCAG
CGGCTCTGCTTGTTCGTGTGTGTGTCTGTTGCAGGCCCTTAT
TCGGATCCATGGTGAGCAAGGGCGAGGAGCTGTTTACCCGG
GGTGGTGCCCATCCTGGTTCGAGCTGGACGGCGACGTAAAC
GGCCACAAGTTCAGCGTGTCCGGCGAGGGGCGAGGGGCGATG
CCACCTACGGCAAGCTGACCCCTGAAGTTCATCTGCACCAC
CGGCAAGCTGCCCGTGCCCTGGCCCCACCCTCGTGACCACC
CTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACC
ACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGA
AGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGAC
GGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCG
ACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTT
CAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTAC
AACTACAACAGCCACAACGTCCTATATCATGGCCGACAAGC
AGAAGAACGGCATCAAGGTGAACCTTCAAGATCCGCCACAA
CATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAG
CAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCG
ACAACCCTACTCCTGAGCACCCAGTCCGCCCTGAGCAAAGA
CCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC
GTGACCGCCGCGCGGGATCACTCTCGGCATGGACGAGCTGT
ACAAGTAATAGGAATTATGCAGTCTAGCCATCACATTTAA
AAGCATCTCAGCCTACCATGAGAATAAGAGAAAGAAAATG
AAGATCAATAGCTTATTTCATCTCTTTTCTTTTCTGTTGG
TGTAAGCCAAACACCCTGTCTAAAAAACATAAATTTCTTT
AATCATTTTGGCCTCTTTTCTCTGTGCTTCAATTAATAAAA
AATGGAAAGAACCTCGAGAAAAAATAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATC
TAGACTTAAGCTTGATGAGCTCTAGCTTGGCGTAATCATG
GTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACA
ATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAG
CCTGGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGC
GTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCTG
TGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAG
GCGGTTTGCGTATTGGGCGCTCTTCCGCTTCTCTGCTCAC
TGACTCGCTGCGCTCGGTCTGTTCTGGCTGCGGCGAGCGGTA
TCAGCTCACTCAAAGGCGGGTAATACGGTTATCCACAGAAT
CAGGGGATAACGCAAGGAAGGAACATGTGAGCAAAAGGCCA
GCAAAAGGCCAGGAACCGTA AAAAGGCCGCGTTGCTGGCG
TTTTTCCATAGGCTCCGCCCCCCCTGACGAGCATCACAAAA
ATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT
ATAAAGATAACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG
CGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGT
CCGCCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAG
CTCACGCTGTAGGTATCTCAGTTCTGGTGTAGGTCTGTTCTGC
TCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTTCAGCCCG
ACCGCTGCGCCTTATCCGGTAACATCGTCTTGAGTCCAA

10

20

30

40

50

CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACT
GGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTA
CAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAG
AAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT
ACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGGCAAAC
AAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCA
GCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCT
TTGATCTTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAA
ACTCACGTTAAGGGATTTTGGTTCATGAGATTATCAAAAAG
GATCTTTCACCTAGATCCTTTTTTAAATTAATAAATGAAGTTTT
AAATCAATCTAAAGTATATATGAGTAAACTTTGGTCTGACA
GTTACCAATGCTTAATCAGTGAGGGCACCTATCTCAGCGAT
CTGTCTATTTTCGTTTCATCCATAGTTGCCCTGACTCCCCGTC
GTGTAGATAACTACGATACGGGGAGGGCTTACCATCTGGCC
CCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGC
TCCAGATTTTATCAGCAATAAACCAGCCAGCCGGAAGGGCC
GAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCC
AGTCTATTAAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTC
GCCAGTTAATAGTTTTCGCGCAACGTTGTTGCCATTGCTACA
GGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCAT
TCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATC
CCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCT
CCGATCGTTGTGCAAGTAAGTTGGCCGCGAGTGTATCAC
TCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCAT
GCCATCCGTAAGATGCTTTTTCTGTGACTGGTGAGTACTCA
ACCAAGTCATTCTGAGAATAGTGATGCGGCGACCGAGTT
GCTCTTGCCCGGCGTCAATACGGGATAATAACCGCGCCACA
TAGCAGAACTTTAAAAGTGCTCATCATTTGGAAAACGTTCT
TCGGGGGCGAAAACTCTCAAGGATCTTACCCTGTTGAGAT
CCAGTTCGATGTAACCCACTCGTGCAACCCAACTGATCTTC
AGCATCTTTTACTTTTACCAGCGTTTTCTGGGTGAGCAAAA
ACAGGAAGGC AAAATGCCGCAAAAAAGGGAATAAGGGCGA
CACGGAAATGTTGAATACTCATACTCTTCCTTTTTTCAATA
TTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGA
TACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGG
TTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTA
AGAAACCATTATTATCATGACATTAACCTATAAAAAATAGG
CGTATCACGAGGCCCTTTCTGTCG

【0278】

配列表

SEQUENCE LISTING

<110> INTELLIA THERAPEUTICS, INC.

<120> FORMULATIONS

<130> PA23-236

<150> US 62/566,240

<151> 2017-09-29

<160> 84

<170> PatentIn version 3.5

<210> 1

<211> 4140

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 1

atggacaaga agtacagcat cggactggac atcggaacaa acagcgtcgg atgggcagtc	60	
atcacagacg aatacaaggt cccgagcaag aagttcaagg tcctgggaaa cacagacaga	120	
cacagcatca agaagaacct gatcggagca ctgctgttcg acagcggaga aacagcagaa	180	10
gcaacaagac tgaagagaac agcaagaaga agatacacia gaagaaagaa cagaatctgc	240	
tacctgcagg aaatcttcag caacgaaatg gcaaaggctg acgacagctt cttccacaga	300	
ctggaagaaa gcttcctggt cgaagaagac aagaagcacg aaagacaccc gatcttcgga	360	
aacatcgtcg acgaagtgcg ataccacgaa aagtaccgca caatctacca cctgagaaag	420	
aagctggctg acagcacaga caaggcagac ctgagactga tctacctggc actggcacac	480	
atgatcaagt tcagaggaca cttcctgata gaaggagacc tgaaccgga caacagcgac	540	
gtcgacaagc tgttcatcca gctgggtccag acatacaacc agctgttcga agaaaacccg	600	
atcaacgcaa gcggagtcga cgcaaaggca atcctgagcg caagactgag caagagcaga	660	
agactggaaa acctgatcgc acagctgccg ggagaaaaga agaacggact gttcggaac	720	
ctgatcgcac tgagcctggg actgacaccg aacttcaaga gcaacttcga cctggcagaa	780	20
gacgcaaagc tgcagctgag caaggacaca tacgacgacg acctggacaa cctgctggca	840	
cagatcggag accagtacgc agacctgttc ctggcagcaa agaacctgag cgacgcaatc	900	
ctgctgagcg acatcctgag agtcaacaca gaaatcacia aggcaccgct gagcgcaagc	960	
atgatcaaga gatacgacga acaccaccag gacctgacac tgctgaaggc actggtcaga	1020	
cagcagctgc cggaaaagta caaggaaatc ttcttcgacc agagcaagaa cggatacgca	1080	
ggatacatcg acggaggagc aagccaggaa gaattctaca agttcatcaa gccgatcctg	1140	
gaaaagatgg acggaacaga agaactgctg gtcaagctga acagagaaga cctgctgaga	1200	
aagcagagaa cattcgacaa cggaagcatc ccgcaccaga tccacctggg agaactgcac	1260	
gcaatcctga gaagacagga agacttctac ccgttctga aggacaacag agaaaagatc	1320	
gaaaagatcc tgacattcag aatcccgtac tacgtcggac cgctggcaag aggaaacagc	1380	30
agattcgcat ggatgacaag aaagagcgaa gaaacaatca caccgtggaa cttcgaagaa	1440	
gtcgtcgaca agggagcaag cgcacagagc ttcacgaaa gaatgacaaa cttcgacaag	1500	
aacctgccga acgaaaaggt cctgccgaag cacagcctgc tgtacgaata cttcacagtc	1560	
tacaacgaac tgacaaaggt caagtacgtc acagaaggaa tgagaaagcc ggcattcctg	1620	
agcggagaaac agaagaaggc aatcgtcgac ctgctgttca agacaaacag aaaggtcaca	1680	
gtcaagcagc tgaaggaaga ctacttcaag aagatcgaat gttcgcagag cgtcgaaatc	1740	
agcggagtcg aagacagatt caacgaagc ctgggaacat accacgacct gctgaagatc	1800	
atcaaggaca aggacttcct ggacaacgaa gaaaacgaag acatcctgga agacatcgtc	1860	
ctgacactga cactgttcga agacagagaa atgatcgaag aaagactgaa gacatacgca	1920	
cacctgttcg acgacaaggt catgaagcag ctgaagagaa gaagatacac aggatgggga	1980	40
agactgagca gaaagctgat caacggaatc agagacaagc agagcggaaa gacaatcctg	2040	
gacttcctga agagcgacgg attcgcaaac agaaacttca tgcagctgat ccacgacgac	2100	
agcctgacat tcaaggaaga catccagaag gcacagggtc gcggacaggg agacagcctg	2160	
cacgaacaca tcgcaaacct ggcaggaagc ccggcaatca agaagggaat cctgcagaca	2220	
gtcaaggctg tcgacgaact ggtcaaggct atgggaagac acaagccgga aaacatcgtc	2280	
atcgaaatgg caagagaaaa ccagacaaca cagaaggggac agaagaacag cagagaaaga	2340	
atgaagagaa tcgaagaagg aatcaaggaa ctgggaagcc agatcctgaa ggaacacccg	2400	
gtcgaaaaca cacagctgca gaacgaaaag ctgtacctgt actacctgca gaacggaaga	2460	
gacatgtacg tcgaccagga actggacatc aacagactga gcgactacga cgtcgaccac	2520	
atcgtcccgc agagcttcct gaaggacgac agcatcgaca acaaggctcct gacaagaagc	2580	50

gacaagaaca gaggaagag cgacaacgtc cggagcgaag aagtcgtcaa gaagatgaag 2640
 aactactgga gacagctgct gaacgcaaag ctgatcacac agagaaagt cgacaacctg 2700
 acaaaggcag agagaggagg actgagcgaa ctggacaagg caggattcat caagagacag 2760
 ctggtcgaaa caagacagat cacaagcac gtcgcacaga tcctggacag cagaatgaac 2820
 acaaagtacg acgaaaacga caagctgac agagaagtca aggtcatcac actgaagagc 2880
 aagctggtca gcgacttcag aaaggacttc cagttctaca aggtcagaga aatcaacaac 2940
 taccaccacg cacacgacgc atacctgaac gcagtcgtcg gaacagcact gatcaagaag 3000
 taccgaagc tggaaagcga attcgtctac ggagactaca aggtctacga cgtcagaaaag 3060
 atgatcgcaa agagcgaaca ggaaatcgga aaggcaacag caaagtactt cttctacagc 3120
 aacatcatga acttcttcaa gacagaaatc acactggcaa acggagaaat cagaaagaga 3180
 ccgctgatcg aaacaaacgg agaaacagga gaaatcgtct gggacaaggg aagagacttc 3240
 gcaacagtca gaaaggctct gagcatgccg caggtaaca tcgtcaagaa gacagaagtc 3300
 cagacaggag gattcagcaa ggaaagcatc ctgccgaaga gaaacagcga caagctgatc 3360
 gcaagaaaga aggactggga cccgaagaag tacggaggat tcgacagccc gacagtcgca 3420
 tacagcgtcc tggtcgtcg aaaggtcgaa aagggaaga gcaagaagct gaagagcgtc 3480
 aaggaaactgc tgggaatcac aatcatggaa agaagcagct tcgaaaagaa cccgatcgac 3540
 ttcttggaag caaagggata caaggaagtc aagaaggacc tgatcatcaa gctgccgaag 3600
 tacagcctgt tcgaactgga aaacggaaga aagagaatgc tggcaagcgc aggagaactg 3660
 cagaagggaa acgaactggc actgccgagc aagtacgtca acttcctgta cctggcaagc 3720
 cactacgaaa agctgaaggg aagcccggaa gacaacgaac agaagcagct gttcgtcgaa 3780
 cagcacaagc actacctgga cgaaatcatc gaacagatca gcgaattcag caagagagtc 3840
 atcctggcag acgcaaact ggacaaggtc ctgagcgcat acaacaagca cagagacaag 3900
 ccgatcagag aacaggcaga aaacatcatc cacctgttca cactgacaaa cctgggagca 3960
 ccggcagcat tcaagtactt cgacacaaca atcgacagaa agagatacac aagcacaag 4020
 gaagtcctgg acgcaacact gatccaccag agcatcacag gactgtacga aacaagaatc 4080
 gacctgagcc agctgggagg agacggagga ggaagcccga agaagaagag aaaggcttag 4140

10

20

<210> 2

<211> 4143

<212> DNA

<213> Artificial Sequence

<220>

<221> source

 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 2

atggataaga agtactcaat cgggctggat atcggaacta attccgtggg ttgggcagtg 60
 atcacggatg aatacaaagt gccgtccaag aagttcaagg tcctggggaa caccgataga 120
 cacagcatca agaaaaatct catcgagacc ctgctgttg actccggcga aaccgcagaa 180
 gcgacccggc tcaaacttac cgcgaggcga cgctacaccc ggcggaagaa tcgcatctgc 240
 tatctgcaag agatctttc gaacgaaatg gcaaaggctc acgacagctt cttccaccgc 300
 ctggaagaat ctttctggt ggaggaggac aagaagcatg aacggcatcc tatctttgga 360
 aacatcgctc acgaagtggc gtaccacgaa aagtaccgca ccatctacca tctgcggaag 420
 aagttggttg actcaactga caaggccgac ctgagattga tctacttggc cctcgcccat 480
 atgatcaaat tccgcggaca cttcctgatc gaaggcgatc tgaaccctga taactccgac 540
 gtggataagc ttttcattca actggtgcag acctacaacc aactgttcga agaaaacca 600
 atcaatgcta gcggcgtcga tgccaaggcc atcctgtccg cccggctgtc gaagtcgcgg 660
 cgctcgaaa acctgatcgc acagctgccg ggagagaaaa agaacggact ttcggcaac 720
 ttgatcgctc tctcactggg actcactccc aatttcaagt ccaattttga cctggccgag 780
 gacgcgaagc tgcaactctc aaaggacacc tacgacgacg acttgacaaa ttgtctggca 840

30

40

50

caaattggcg atcagtacgc ggatctgttc cttgccgcta agaacccttc ggacgcaatc 900
 ttgctgtccg atatcctgcg cgtgaacacc gaaataacca aagcgccgct tagcgccctcg 960
 atgattaagc ggtacgacga gcatcaccag gatctcacgc tgctcaaagc gtcctgtgaga 1020
 cagcaactgc ctgaaaagta caaggagatc ttcttcgacc agtccaagaa tgggtacgca 1080
 gggtagatcg atggaggcgc tagccaggaa gagttctata agttcatcaa gccaatcctg 1140
 gaaaagatgg acggaaccga agaactgctg gtcaagctga acaggaggga tctgctccgg 1200
 aaacagagaa cctttgacaa cggatccatt cccaccaga tccatctggg tgagctgcac 1260
 gccatcttgc ggcgccagga ggacttttac ccattcctca aggacaaccg ggaaaagatc 1320
 gagaaaattc tgacgttccg catcccgtat tacgtgggccc cactggcgcg cggcaattcg 1380
 cgcttcgcgt ggatgactag aaaatcagag gaaaccatca ctcttgga tttcgaggaa 1440
 gttgtggata agggagcttc ggcacaaagc ttcacgaac gaatgacaa cttcgacaag 1500
 aatctcccaa acgagaaggt gcttcctaag cacagcctcc ttacgaata cttactgtc 1560
 tacaacgaac tgactaaagt gaaatacgtt actgaaggaa tgaggaagcc ggcctttctg 1620
 tccggagaac agaagaaagc aattgtcgat ctgctgttca agaccaaccg caaggtgacc 1680
 gtcaagcagc ttaaagagga ctacttcaag aagatcgagt gtttcgactc agtggaatc 1740
 agcggggtgg aggacagatt caacgcttcg ctgggaacct atcatgatct cctgaagatc 1800
 atcaaggaca aggacttctt tgacaacgag gagaacgagg acatcctgga agatatcgtc 1860
 ctgaccttga cctttttcga ggatcgcgag atgatcgagg agaggcttaa gacctacgt 1920
 catctcttcg acgataaggt catgaaacaa ctcaagcgcc gccggtagac tggttggggc 1980
 cgcctctccc gcaagctgat caacggtatt cgcgataaac agagcggtaa aactatcctg 2040
 gatttctca aatcggatgg cttcgcta atcgtaacttca tgcaattgat ccacgacgac 2100
 agcctgacct ttaaggagga catccaaaaa gcacaagtgt ccggacaggg agactcactc 2160
 catgaacaca tcgcgaatct ggccggttcg ccggcgatta agaagggaat tctgcaaat 2220
 gtgaagggtg tcgacgagct ggtgaaggct atgggacggc acaaaccgga gaatatcgtg 2280
 attgaaatgg cccgagaaaa ccagactacc cagaagggcc agaaaaactc ccgcgaaagg 2340
 atgaagcggg tcgaagaagg aatcaaggag ctgggcagcc agatcctgaa agagcaccgg 2400
 gtggaaaaca cgcagctgca gaacgagaag ctctacctgt actatttgca aaatggacgg 2460
 gacatgtacg tggaccaaga gctggacatc aatcggttgt ctgattacga cgtggaccac 2520
 atcgttccac agtcctttct gaaggatgac tcgatcgata acaagggtgt gactcgacg 2580
 gacaagaaca gaggggaagt agataatgt ccacggagg aggtcgtgaa gaagatgaag 2640
 aattactggc ggcagctcct gaatgcgaag ctgattaccc agagaaagtt tgacaatctc 2700
 actaaagccg agcgcggcgg actctcagag ctggataagg ctggattcat caaacggcag 2760
 ctggctgaga ctgcgcagat taccaagcac gtggcgaga tcttgactc ccgcatgaac 2820
 actaaatacg acgagaacga taagctcatc cggaagtga aggtgattac cctgaaaagc 2880
 aaacttgtgt cggactttcg gaaggacttt cagttttaca aagtgagaga aatcaacaac 2940
 taccatcacg cgcatgacgc atacctcaac gctgtggtcg gtaccgccct gatcaaaaag 3000
 taccctaaac ttgaatcgga gtttgtgtac ggagactaca aggtctacga cgtgaggaag 3060
 atgatagcca agtccgaaca ggaaatcggg aaagcaactg cgaaatactt cttttactca 3120
 aacatcatga actttttcaa gactgaaatt acgctggcca atggagaaat caggaagagg 3180
 ccactgatcg aaactaacgg agaaacgggc gaaatcgtgt gggacaaggg cagggacttc 3240
 gcaactgttc gcaaagtgt ctctatgccg caagtcaata ttgtgaagaa aaccgaagt 3300
 caaacggcg gattttcaaa ggaatcgatc ctcccaaaga gaaatagcga caagctcatt 3360
 gcacgcaaga aagactggga cccgaagaag tacggaggat tcgattcgcc gactgtcgca 3420
 tactccgtcc tcgtggtggc caagggtggg aagggaagaa gcaaaaagct caaatccgtc 3480
 aaagagctgc tggggattac catcatggaa cgatcctcgt tcgagaagaa cccgattgat 3540
 ttctcagag cgaagggtta caaggaggtg aagaaggatc tgatcatcaa actcccaaag 3600
 tactactgt tcgaactgga aaatggtcgg aagcgcatgc tggcttcggc cggagaactc 3660
 caaaaaggaa atgagctggc cttgcctagc aagtacgtca acttctcta tcttcttcg 3720
 cactacgaaa aactcaaagg gtcaccggaa gataacgaac agaagcagct tttcgtggag 3780
 cagcacaagc attatctgga tgaaatcatc gaacaaatct ccgagttttc aaagcgctg 3840

10

20

30

40

50

atcctcgccg acgccaacct cgacaaagtc ctgtcggcct acaataagca tagagataag 3900
ccgatcagag aacaggccga gaacattatc cactgttca ccctgactaa cctgggagcc 3960
ccagccgcct tcaagtactt cgatactact atcgatcgca aaagatacac gtccaccaag 4020
gaagttctgg acgcgaccct gatccaccaa agcatcactg gactctacga aactaggatc 4080
gatctgtcgc agctgggtgg cgatggcggg ggatctccga aaaagaagag aaaggtgtaa 4140
tga 4143

<210> 3

<211> 1379

<212> PRT

<213> Unknown

<220>

<221> source

<223> /note="Description of Unknown:

Cas9 sequence"

<400> 3

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile

35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu

50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys

65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser

85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys

100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr

115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp

130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His

145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro

165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr

10

20

30

40

50

180 185 190
Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
195 200 205
Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
210 215 220
Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225 230 235 240
Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
245 250 255
Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
260 265 270
Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
275 280 285
Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
290 295 300
Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305 310 315 320
Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
325 330 335
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
340 345 350
Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
355 360 365
Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
370 375 380
Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385 390 395 400
Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415
Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430
Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
435 440 445

10

20

30

40

50

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510

10

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560

20

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605

30

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670

40

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720

50

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
885 890 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
900 905 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
915 920 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
930 935 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val

10

20

30

40

50

980	985	990
Val Gly Thr Ala Leu Ile Lys Lys	Tyr Pro Lys Leu Glu	Ser Glu Phe
995	1000	1005
Val Tyr Gly Asp Tyr Lys Val	Tyr Asp Val Arg Lys	Met Ile Ala
1010	1015	1020
Lys Ser Glu Gln Glu Ile Gly	Lys Ala Thr Ala Lys	Tyr Phe Phe
1025	1030	1035
Tyr Ser Asn Ile Met Asn Phe	Phe Lys Thr Glu Ile	Thr Leu Ala
1040	1045	1050
Asn Gly Glu Ile Arg Lys Arg	Pro Leu Ile Glu Thr	Asn Gly Glu
1055	1060	1065
Thr Gly Glu Ile Val Trp Asp	Lys Gly Arg Asp Phe	Ala Thr Val
1070	1075	1080
Arg Lys Val Leu Ser Met Pro	Gln Val Asn Ile Val	Lys Lys Thr
1085	1090	1095
Glu Val Gln Thr Gly Gly Phe	Ser Lys Glu Ser Ile	Leu Pro Lys
1100	1105	1110
Arg Asn Ser Asp Lys Leu Ile	Ala Arg Lys Lys Asp	Trp Asp Pro
1115	1120	1125
Lys Lys Tyr Gly Gly Phe Asp	Ser Pro Thr Val Ala	Tyr Ser Val
1130	1135	1140
Leu Val Val Ala Lys Val Glu	Lys Gly Lys Ser Lys	Lys Leu Lys
1145	1150	1155
Ser Val Lys Glu Leu Leu Gly	Ile Thr Ile Met Glu	Arg Ser Ser
1160	1165	1170
Phe Glu Lys Asn Pro Ile Asp	Phe Leu Glu Ala Lys	Gly Tyr Lys
1175	1180	1185
Glu Val Lys Lys Asp Leu Ile	Ile Lys Leu Pro Lys	Tyr Ser Leu
1190	1195	1200
Phe Glu Leu Glu Asn Gly Arg	Lys Arg Met Leu Ala	Ser Ala Gly
1205	1210	1215
Glu Leu Gln Lys Gly Asn Glu	Leu Ala Leu Pro Ser	Lys Tyr Val
1220	1225	1230

10

20

30

40

50

Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser
 1235 1240 1245

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys
 1250 1255 1260

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys
 1265 1270 1275

Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala
 1280 1285 1290

Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn
 1295 1300 1305

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala
 1310 1315 1320

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
 1325 1330 1335

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
 1340 1345 1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
 1355 1360 1365

Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys Val
 1370 1375

<210> 4

<211> 4140

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 4

auggacaaga aguacagcau cggacuggac aucggaacaa acagcgucgg augggcaguc 60
 aucacagacg aaauacaaggu cccgagcaag aaguucaagg uccugggaaa cacagacaga 120
 cacagcauca agaagaaccu gaucggagca cugcuguucg acagcggaga aacagcagaa 180
 gcaacaagac ugaagagaac agcaagaaga agauacacaa gaagaaagaa cagaauucgc 240
 uaccugcagg aaauucuucag caacgaaaug gcaaaggucg acgacagcuu cuuccacaga 300
 cuggaagaaa gcuuccuggu cgaagaagac aagaagcacg aaagacaccc gaucuucgga 360
 aaucugucg acgaagucgc auaccacgaa aaguaccgca caaucuacca ccugagaaag 420
 aagcuggucg acagcacaga caaggcagac cugagacuga ucuaccuggc acuggcacac 480
 augaucaagu ucagaggaca cuuccugauc gaaggagacc ugaacccgga caacagcgac 540
 gucgacaagc uguucaucca gcugguccag acauacaacc agcuguucga agaaaacccg 600
 aucaacgcaa gcggagucga cgcaaaggca auccugagcg caagacugag caagagcaga 660

10

20

30

40

50

agacuggaaa accugaucgc acagcugccg ggagaaaaga agaacggacu guucggaaac	720	
cugaucgcac ugagccuggg acugacaccg aacuuaaga gcaacuucga ccuggcagaa	780	
gacgcaaagc ugcagcugag caaggacaca uacgacgacg accuggacaa ccugcuggca	840	
cagaucggag accaguacgc agaccuguuc cuggcagcaa agaaccugag cgacgcaauc	900	
cugcugagcg acauccugag agucaacaca gaaaucacaa aggcaccgcu gagcgcaagc	960	
augaucaaga gauacgacga acaccaccag gaccugacac ugcugaaggc acuggucaga	1020	
cagcagcugc cggaaaagua caaggaaauc uucuucgacc agagcaagaa cggauacgca	1080	
ggauacaucg acggaggagc aagccaggaa gaaauucuaca aguucaucaa gccgauccug	1140	
gaaaagaugg acggaacaga agaacugcug gucaagcuga acagagaaga ccugcugaga	1200	
aagcagagaa cauucgacaa cggaagcauc ccgcaccaga uccaccuggg agaacugcac	1260	10
gcaauccuga gaagacagga agacuucuac ccguuccuga aggacaacag agaaaagauc	1320	
gaaaagaucc ugacauucag aaucccgua uacgucggac cgcuggcaag aggaaacagc	1380	
agauucgcau ggauagacaag aaagagcgaa gaaacaauca caccguggaa cuucgaagaa	1440	
gucgucgaca agggagcaag cgcacagagc uucaucgaaa gaaugacaaa cuucgacaag	1500	
aaccugccga acgaaaaggu ccugccgaag cacagccugc uguacgaaua cuucacaguc	1560	
uacaacgaac ugacaaaaggu caaguacguc acagaaggaa ugagaaagcc ggcauuccug	1620	
agcggagaac agaagaaggc aaucgucgac cugcuguuca agacaaacag aaaggucaca	1680	
gucaagcagc ugaaggaaga cuacuuaag aagaucgaau gcuucgacag cgucgaauc	1740	
agcggagucg aagacagauu caacgcaagc cugggaacau accacgaccu gcugaagauc	1800	
aucaaggaca aggacuuccu ggacaacgaa gaaaacgaag acauccugga agacaucguc	1860	20
cugacacuga cacuguucga agacagagaa augaucgaag aaagacugaa gacauacgca	1920	
caccuguucg acgacaaggu caugaagcag cugaagagaa gaagauacac aggaugggga	1980	
agacugagca gaaagcugau caacggaauc agagacaagc agagcggaaa gacaauccug	2040	
gacuuccuga agagcgacgg auucgcaaac agaaacuua ugcagcugau ccacgacgac	2100	
agccugacau ucaaggaaga cauccagaag gcacagguca gcggacaggg agacagccug	2160	
cacgaacaca ucgcaaaccu ggcaggaagc ccggcaauca agaagggaau ccugcagaca	2220	
gucaaggucg ucgacgaacu ggucaagguc augggaagac acaagccgga aaacaucguc	2280	
aucgaaaugg caagagaaaa ccagacaaca cagaagggac agaagaacag cagagaaaga	2340	
augaagagaa ucgaagaagg aaucaggaa cugggaagcc agauccugaa ggaacacccg	2400	
gucgaaaaca cacagcugca gaacgaaaag cuguaccugu acuaccugca gaacggaaga	2460	30
gacauguacg ucgaccagga acuggacauc aacagacuga gcgacuacga cgucgaccac	2520	
aucguccgc agagcuuccu gaaggacgac agcaucgaca acaagguccu gacaagaagc	2580	
gacaagaaca gaggaaagag cgacaacguc ccgagcgaag aagucgucaa gaagaugaag	2640	
aacuacugga gacagcugcu gaacgcaaag cugaucacac agagaaaguu cgacaaccug	2700	
acaaaggcag agagaggagg acugagcgaa cuggacaagg caggauucau caagagacag	2760	
cuggucgaaa caagacagau cacaagcac gucgcacaga uccuggacag cagaaugaac	2820	
acaaaguacg acgaaaacga caagcugauc agagaaguca aggucaucac acugaagagc	2880	
aagcugguca gcgacuucag aaaggacuuc caguucuaca agguacagaga aaucacaac	2940	
uaccaccacg cacacgacgc auaccugaac gcagucgucg gaacagcacu gaucaagaag	3000	
uacccgaagc uggaaaagcga auucgucuac ggagacuaca aggucuacga cgucagaaag	3060	40
augaucgcaa agagcgaaca ggaaaucgga aaggcaacag caaaguacuu cuucuacagc	3120	
aacauauga acuuuuucaa gacagaaauc acacuggcaa acggagaaaau cagaaagaga	3180	
ccgcugaucg aaacaaacgg agaaacagga gaaaucgucu gggacaaggg aagagacuuc	3240	
gcaacaguca gaaagguccu gagcaugccg caggucaaca ucgucaagaa gacagaaguc	3300	
cagacaggag gauucagcaa ggaaagcauc cugccgaaga gaaacagcga caagcugauc	3360	
gcaagaaaga aggacuggga cccgaagaag uacggaggau ucgacagccc gacagucgca	3420	
uacagcgucc ugguucgucg aaaggucgaa aagggaagaa gcaagaagcu gaagagcguc	3480	
aaggaaucgc uggaauacac aaucaggaa agaagcagcu ucgaaaagaa cccgaucgac	3540	
uuccuggaag caaagggaau caaggaaguc aagaaggacc ugaucuucaa gcugccgaag	3600	
uacagccugu ucgaacugga aaacggaaga aagagaaguc uggcaagcgc aggagaacug	3660	50

cagaagggaa acgaacuggc acugccgagc aaguacguca acuuccugua ccuggcaagc 3720
cacuacgaaa agcugaaggg aagcccggaa gacaacgaac agaagcagcu guucgucgaa 3780
cagcacaagc acuaccugga cgaaaaucauc gaacagauca gcgaauucag caagagaguc 3840
auccuggcgag acgcaaaccu ggacaagguc cugagcgcau acaacaagca cagagacaag 3900
ccgaucagag aacaggcgaga aaacaucauc caccuguuua cacugacaaa ccuggggagca 3960
ccggcagcau ucaaguacuu cgacacaaca aucgacagaa agagauacac aagcacaag 4020
gaaguccugg acgcaacacu gauccaccag agcaucacag gacuguaacga aacaagaau 4080
gaccugagcc agcuggggagg agacggagga ggaagcccga agaagaagag aaaggucuag 4140

<210> 5

<211> 4143

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 5

auggauaaga aguacucaau cgggcuggau aucggaacua auuccguggg uugggcagug 60
aucacggaug aaucacaaagu gccguccaag aaguucaagg uccuggggaa caccgauaga 120
cacagcauca agaaaaauuc caucggagcc cugcuguuug acuccggcga aaccgcagaa 180
gcgacccggc ucaaacguac cgcgaggcga cgcuaacccc ggcggaagaa ucgcaucugc 240
uauugcaag agaucuuuuc gaacgaaaug gcaaaggucg acgacagcuu cuuccaccgc 300
cuggaagaau cuuuccuggu ggaggaggac aagaagcaug aacggcaucc uaucuuugga 360
aauaucgucg acgaaguggc guaccacgaa aaguaccgga ccaucuacca ucugcggaag 420
aaguugguug acucaacuga caaggccgac cucagauuga ucuacuuggc ccucgcccua 480
augaucaau uccgcggaca cuuccugauc gaaggcgau ugaaccuga uaacuccgac 540
guggauaagc uuuucauua acuggugcag accuacaacc aacuguuca agaaaaccca 600
aucaaugcua gcggcgucga ugccaaggcc auccuguccg cccggcuguc gaagucgcgg 660
cgccucgaaa accugaucgc acagcugccg ggagagaaaa agaaccgacu uuucggcaac 720
uugaucgcuc ucucacuggg acucacuccc aaauucaagu ccaauuuuga ccuggccgag 780
gacgcgaagc ugcaacucuc aaaggacacc uacgacgacg acuuggacaa uuugcuggca 840
caauuuggcg aucaguacgc ggaucuguuc cuugccgcua agaaccuuuc ggacgcauuc 900
uugcuguccg auauccugcg cgugaacacc gaaauaacca aagcgccgcu uagcgccucg 960
augauuaagc gguacgacga gcaucaccag gaucucacgc ugcuaaaagc gcucgugaga 1020
cagcaacugc cugaaaagua caaggagau uucucgacc aguccaagaa ugguuacgca 1080
ggguacaucg auggagggcg uagccaggaa gaguucuaa aguucaucaa gccaauccug 1140
gaaaagaugg acggaaccga agaacugcug gucaagcuga acaggaggga ucugcuccgg 1200
aaacagagaa ccuuugacaa cggauccauu cccaccaga uccaucuggg ugagcugcac 1260
gccaucuuugc ggcgccagga ggacuuuuac ccauuccuca aggacaaccg ggaaaagau 1320
gagaaaauuc ugacguuccg caucccgua uacgugggcc cacuggcgcg cggcaauucg 1380
cgcuucgcu ggaugacuag aaaauacagag gaaaccauca cuccuuggaa uuucgaggaa 1440
guuguggaua agggagcuuc ggcacaaagc uucaucgaac gaaugaccaa cuucgacaag 1500
aaucucccaa acgagaaggu gcuuccuaag cacagccucc uuuaacgaau cuucacuguc 1560
uacaacgaac ugacuaaagu gaaauacguu acugaaggaa ugaggaagcc ggccuuucug 1620
uccggagaac agaagaaagc aaauugcgau cugcuguuua agaccaaccg caaggugacc 1680
gucaagcagc uaaaagagga cuacuuaag aagaucgagu guuucgacuc aguggaaauc 1740
agcggggugg aggacagau caacgcuucg cugggaaccu aucaugaucu ccugaagauc 1800
aucaaggaca aggacuuccu ugacaacgag gagaacgagg acauccugga agauaucguc 1860
cugaccuuga cccuuuucga ggaucgcgag augaucgagg agaggcuua gaccuacgu 1920

10

20

30

40

50

caucucuucg acgauaaggu caugaaacaa cucaagcgcc gccgguacac ugguuggggc 1980
 cgccucuccc gcaagcugau caacgguaau cgcgauaaac agagcgguuaa aacuauccug 2040
 gauuuccuca aaucggauug cuucgcuaau cguaacuua ugcaauugau ccacgacgac 2100
 agccugaccu uuaaggagga cauccaaaaa gcacaagugu ccggacaggg agacucacuc 2160
 caugaacaca ucgcgaauuc ggccggauuc cgggcgauua agaagggaau ucugcaaacu 2220
 gugaaggugg ucgacgagcu ggugaagguc augggacggc acaaaccgga gaauaucgug 2280
 auugaaaugg cccgagaaaa ccagacuacc cagaaggggc agaaaaacuc ccgcgaaaagg 2340
 augaagcggg ucgaagaagg aaucaggag cugggcagcc agauccugaa agagcaccgc 2400
 guggaaaaca cgacgagcga gaacgagaag cucuaccugu acuaauugca aaauaggacgg 2460
 gacauguacg uggaccaaga gcuggacauc aaucgguguu cugauuacga cguggaccac 2520
 aucguccac aguccuuuc gaaggauag ucgaucgaa acaagguguu gacucgcagc 2580
 gacaagaaca gagggaguc agauaaugug ccaucggagg aggucgugaa gaagaugaag 2640
 aaauacuggc ggcagcuccu gaaugcgaag cugauuacc agagaaaguu ugacaauuc 2700
 acuaaaagcc agcgcgggcg acucucagag cuggauaagg cuggauucau caaacggcag 2760
 cuggucgaga cucggcagau uaccaagcac guggcgaga ucuuggacuc ccgcaugaac 2820
 acuaaaucg acgagaacga uaagcuauc cgggaaguga aggugauuac ccugaaaagc 2880
 aaacuugugu cggacuucg gaaggacuuc caguuuaca aagugagaga aaucacaac 2940
 uaccaucac cgcaugacgc auaccucaac gcuguggucg guaccgccc gaucaaaaag 3000
 uaccuuaac uugaauagg guuuguguac ggagacuaca aggucuacga cgugaggaag 3060
 augauagcca aguccgaaca ggaaucggg aaagcaacug cgaaauacuu cuuuuacuca 3120
 aacauauga acuuuuucaa gacugaaaau acgcuggcca auggagaaau caggaagagg 3180
 ccacugauc aaacuaacgg agaaacggc gaaucgugu gggacaaggg cagggacuuc 3240
 gcaacuguc gcaaagugcu cucuaugccg caagucaaua uugugaagaa aaccgaagug 3300
 caaacggcg gauuuucaa ggaucgauc cucccaaaga gaaauagcga caagcuauc 3360
 gcacgcaaga aagacuggga cccgaagaag uacggaggau ucgaucgcc gacugucga 3420
 uacuccgucc ucgugguggc caagguggag aagggaaga gcaaaaagcu caaauccguc 3480
 aaagagcugc uggggauuac caucaggaa cgauccucg ucgagaagaa cccgauugau 3540
 uuccucgagg cgaaggguua caaggaggug aagaaggauc ugaucauca acuccccaag 3600
 uacucacugu ucgaacugga aaauaggucg aagcgcaugc uggcuucggc cggagaacuc 3660
 caaaaaggaa augagcuggc cuugccuagc aaguacguc acuccucua ucuugcuucg 3720
 cacuacgaaa aacucaaaagg gucaccggaa gaaacgaac agaagcagcu uuucguggag 3780
 cagcacaagc auuauucgga ugaaaucauc gaacaaauuc ccgaguuuuc aaagcgcgug 3840
 auccucgccg acgccaaccu cgacaaaguc cugucggccu acaauaagca uagagauaag 3900
 ccgaucagag aacaggccga gaacauuac cacuuguuca ccugacuua ccugggagcc 3960
 ccagccgccu ucaaguacu cgauacuac aucgaucgca aaagauacac guccaccaag 4020
 gaaguucgg acgcgaccu gaucaccaa agcaucacug gacucuaag aacuaggauc 4080
 gaucugucgc agcugggugg cgauggcggu ggaucuccga aaaagaagag aaagguguaa 4140
 uga 4143

<210> 6

<211> 1379

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 6

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val

1

5

10

15

10

20

30

40

50

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
 20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
 35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
 50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
 65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
 85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
 100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
 115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
 130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
 145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
 165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
 180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
 195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
 210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
 225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
 245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
 260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp

10

20

30

40

50

275 280 285
Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
290 295 300
Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305 310 315 320
Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
325 330 335
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
340 345 350
Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
355 360 365
Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
370 375 380
Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385 390 395 400
Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415
Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430
Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
435 440 445
Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460
Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480
Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495
Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510
Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525
Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540

10

20

30

40

50

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605

10

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655

20

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700

30

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
755 760 765

40

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815

50

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
885 890 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
900 905 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
915 920 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
930 935 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val

10

20

30

40

50

1070	1075	1080
Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr		
1085	1090	1095
Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys		
1100	1105	1110
Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro		
1115	1120	1125
Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val		
1130	1135	1140
Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys		
1145	1150	1155
Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser		
1160	1165	1170
Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys		
1175	1180	1185
Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu		
1190	1195	1200
Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly		
1205	1210	1215
Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val		
1220	1225	1230
Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser		
1235	1240	1245
Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys		
1250	1255	1260
His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys		
1265	1270	1275
Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala		
1280	1285	1290
Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn		
1295	1300	1305
Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala		
1310	1315	1320

10

20

30

40

50

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
 1325 1330 1335

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
 1340 1345 1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
 1355 1360 1365

Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys Val
 1370 1375

10

<210> 7

<211> 4140

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

20

<400> 7

auggacaaga aguacagcau cggacuggca aucggaacaa acagcgucgg augggcaguc 60
 aucacagacg aaucacaaggu cccgagcaag aaguucaagg uccugggaaa cacagacaga 120
 cacagcauca agaagaaccu gaucggagca cugcuguucg acagcggaga aacagcagaa 180
 gcaacaagac ugaagagaac agcaagaaga agauacacaa gaagaaagaa cagaaucugc 240
 uaccugcagg aaaucuucag caacgaaaug gcaaaggucg acgacagcuu cuuccacaga 300
 cuggaagaaa gcuuccuggu cgaagaagac aagaagcacg aaagacaccc gaucuucgga 360
 aaucugucg acgaagucgc auaccacgaa aaguaccgga caaucuacca ccugagaaag 420
 aagcuggucg acagcacaga caaggcagac cugagacuga ucuaccuggc acuggcacac 480
 augaucaagu ucagaggaca cuuccugauc gaaggagacc ugaacccgga caacagcgac 540
 gucgacaagc uguucaucca gcugguccag acauacaacc agcuguucga agaaaacccg 600
 aucaacgcaa gcggagucga cgcaaaggca auccugagcg caagacugag caagagcaga 660
 agacuggaaa accugaucgc acagcugccg ggagaaaaga agaacggacu guucggaaac 720
 cugaucgcac ugagccuggg acugacaccg aacuucaaga gcaacuucga ccuggcagaa 780
 gacgcaaagc ugcagcugag caaggacaca uacgacgacg accuggacaa ccugcuggca 840
 cagaucggag accaguacgc agaccuguuc cuggcagcaa agaaccugag cgacgcaauc 900
 cugcugagcg acauccugag agucaacaca gaaucacaa aggcaccgcu gagcgcaagc 960
 augaucaaga gauacgacga acaccaccag gaccugacac ugcugaaggc acuggucaga 1020
 cagcagcugc cggaaaagua caaggaaauc uucuucgacc agagcaagaa cggauacgca 1080
 ggauacaucg acggaggagc aagccaggaa gaauucuaca aguucauca gccgauccug 1140
 gaaaagaugg acggaacaga agaacugcug gucaagcuga acagagaaga ccugcugaga 1200
 aagcagagaa cauucgacaa cggaagcauc ccgcaccaga uccaccuggg agaacugcac 1260
 gcaauccuga gaagacagga agacuucuac ccguuccuga aggacaacag agaaaagauc 1320
 gaaaagaucc ugacauucag aaucccgua uacgucggac cgcuggcaag aggaaacagc 1380
 agauucgcau ggauagacaag aaagagcgaa gaaacaauca caccguggaa cuucgaagaa 1440
 gucgucgaca agggagcaag cgcacagagc uucaucgaaa gaaugacaaa cuucgacaag 1500
 aaccugccga acgaaaaggu ccugccgaag cacagccugc uguacgaaua cuucacaguc 1560
 uacaacgaac ugacaaaaggu caaguacguc acagaaggaa ugagaaagcc ggcauuccug 1620
 agcggagaaac agaagaaggc aaucgucgac cugcuguuca agacaaacag aaaggucaca 1680
 gucaagcagc ugaaggaaga cuacuucaag aagaucgaau gcuucgacag cgucgaaauc 1740

30

40

50

agcggagucg aagacagauu caacgcaagc cugggaacau accacgaccu gcugaagauc 1800
 aucaaggaca aggacuuccu ggacaacgaa gaaaacgaag acauccugga agacaucguc 1860
 cugacacuga cacuguucga agacagagaa augaucgaag aaagacugaa gacauacgca 1920
 caccuguucg acgacaaggu caugaagcag cugaagagaa gaagauacac aggaugggga 1980
 agacugagca gaaagcugau caacggaauc agagacaagc agagcggaaa gacaauccug 2040
 gacuuccuga agagcgacgg auucgcaaac agaaacuua ugcagcugau ccacgacgac 2100
 agccugacau ucaaggaaga cauccagaag gcacagguca gcggacaggg agacagccug 2160
 cacgaacaca ucgcaaaccu ggacggaaagc ccggcaauca agaagggaau ccugcagaca 2220
 gucaaggucg ucgacgaacu ggucagguc augggaagac acaagccgga aaacaucguc 2280
 aucgaaaugg caagagaaaa ccagacaaca cagaagggac agaagaacag cagagaaaga 2340
 augaagagaa ucgaagaagg aaucaggaa cugggaagcc agauccugaa ggaacacccg 2400
 gucgaaaaca cacagcugca gaacgaaaag cuguaccugu acuaccugca gaacggaaga 2460
 gacauguacg ucgaccagga acuggacauc aacagacuga gcgacuacga cgucgaccac 2520
 aucgucccg agagcuuccu gaaggacgac agcaucgaca acaagguccu gacaagaagc 2580
 gacaagaaca gaggaagag cgacaacguc ccgagcgaag aagucgucaa gaagaugaag 2640
 aacuacugga gacagcugcu gaacgcaaag cugaucacac agagaaaguu cgacaaccug 2700
 acaaaggcag agagaggagg acugagcga cuggacaagg caggauucau caagagacag 2760
 cuggucgaaa caagacagau cacaagcac gucgcacaga uccuggacag cagaauaagc 2820
 acaaaguacg acgaaaacga caagcugauc agagaaguca aggucaucac acugaagagc 2880
 aagcugguca gcgacuucag aaaggacuuc caguucuaa agguacagaga aaucacaac 2940
 uaccaccacg cacacgacgc auaccugaac gcagucgucg gaacagcacu gaucaagaag 3000
 uacccgaagc uggaaagcga auucgucuac ggagacuaca agguacuacga cgucagaaag 3060
 augaucgcaa agagcgaaca ggaaucgga aaggcaacag caaaguacuu cuucuaacagc 3120
 acaucauga acuucuucaa gacagaaauc acacuggcaa acggagaaaau cagaaagaga 3180
 ccgugaucg aaacaaacgg agaaacagga gaaucgucu gggacaaggg aagagacuuc 3240
 gcaacaguca gaaagguccu gagcaugccg caggucaaca ucgucaagaa gacagaaguc 3300
 cagacaggag gauucagcaa ggaaagcauc cugccgaaga gaaacagcga caagcugauc 3360
 gcaagaaaga aggacuggga cccgaagaag uacggaggau ucgacagccc gacagucgca 3420
 uacagcgucc uggucgucgc aaaggucgaa aagggaaga gcaagaagcu gaagagcguc 3480
 aaggaacugc ugggaauac aaucaggaa agaagcagcu ucgaaaagaa cccgaucgac 3540
 uuccuggaag caaagggauc caaggaaguc aagaaggacc ugaucauca gcugccgaag 3600
 uacagccugu ucgaacugga aaacggaaga aagagaaguc uggcaagcgc aggagaacug 3660
 cagaaggga acgaacuggc acugccgagc aaguacguca acuuccugua ccuggcaagc 3720
 cacuacgaaa agcugaaggg aagcccggaa gacaacgaac agaagcagcu guucgucgaa 3780
 cagcacaagc acuaccugga cgaaaucauc gaacagauca gcgaauucag caagagaguc 3840
 auccuggcag acgcaaaccu ggacaagguc cugagcgcau acaacaagca cagagacaag 3900
 ccgaucagag aacaggcaga aaacaucauc caccuguuca cacugacaaa ccugggagca 3960
 ccggcagcau ucaaguacuu cgacacaaca aucgacagaa agagauacac aagcacaag 4020
 gaaguccug acgcaacacu gauccaccag agcaucacag gacuguacga aacaagaau 4080
 gaccugagcc agcugggagg agacggagga ggaagcccga agaagaagag aaaggucuag 4140

10

20

30

40

<210> 8

<211> 1379

<212> PRT

<213> Artificial Sequence

<220>

<221> source

 <223> /note="Description of Artificial Sequence: Synthetic
 polypeptide"

<400> 8

50

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val
1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
50 55 60

10

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
100 105 110

20

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145 150 155 160

30

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
210 215 220

40

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
260 265 270

50

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
290 295 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305 310 315 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
325 330 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
340 345 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
355 360 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
370 375 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385 390 395 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
435 440 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln

10

20

30

40

50

530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 795 800

10

20

30

40

50

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830

Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys
835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
850 855 860

10

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
885 890 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
900 905 910

20

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
915 920 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
930 935 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945 950 955 960

30

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
1010 1015 1020

40

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
1055 1060 1065

50

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
1085 1090 1095

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys
1100 1105 1110

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro
1115 1120 1125

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val
1130 1135 1140

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys
1145 1150 1155

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser
1160 1165 1170

Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys
1175 1180 1185

Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu
1190 1195 1200

Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly
1205 1210 1215

Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val
1220 1225 1230

Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser
1235 1240 1245

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys
1250 1255 1260

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys
1265 1270 1275

Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala
1280 1285 1290

Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn
1295 1300 1305

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala

10

20

30

40

50

1310	1315	1320
Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser		
1325	1330	1335
Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr		
1340	1345	1350
Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp		
1355	1360	1365
Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys Val		
1370	1375	

10

<210> 9
 <211> 4140
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"
 <400> 9

20

auggacaaga aguacagcau cggacuggca aucggaacaa acagcgucgg augggcaguc	60
aucacagacg aaaucaaggu cccgagcaag aaguucaagg uccugggaaa cacagacaga	120
cacagcauca agaagaaccu gaucggagca cugcuguucg acagcggaga aacagcagaa	180
gcaacaagac ugaagagaac agcaagaaga agauacacaa gaagaaagaa cagaaucugc	240
uaccugcagg aaaucuucag caacgaaaug gcaaaggucg acgacagcuu cuuccacaga	300
cuggaagaaa gcuuccuggu cgaagaagac aagaagcacg aaagacaccc gaucuucgga	360
aacaucgucg acgaagucgc auaccacgaa aaguaccgca caaucuacca ccugagaaag	420
aagcuggucg acagcacaga caaggcagac cugagacuga ucuaccuggc acuggcacac	480
augaucaagu ucagaggaca cuuccugauc gaaggagacc ugaaccgga caacagcgac	540
gucgacaagc uguucaucca gcugguccag acauacaacc agcuguucga agaaaacccg	600
aucaacgcaa gcgagagucga cgcaaaggca auccugagcg caagacugag caagagcaga	660
agacuggaaa accugaucgc acagcugccg ggagaaaaga agaacggacu guucggaaac	720
cugaucgcac ugagccuggg acugacaccg aacuuaaga gcaacuucga ccuggcagaa	780
gacgcaaagc ugcagcugag caaggacaca uacgacgacg accuggacaa ccugcuggca	840
cagaucggag accaguacgc agaccuguuc cuggcagcaa agaaccugag cgacgcaauc	900
cugcugagcg acauccugag agucaacaca gaaauacaaa aggcaccgcu gagcgcaagc	960
augaucaaga gauacgacga acaccaccag gaccugacac ugcugaaggc acuggucaga	1020
cagcagcugc cggaagaagaa caaggaaauc uucuucgacc agagcaagaa cggauacgca	1080
ggauacaucg acggaggagc aagccaggaa gaauucuaca aguucaucaa gccgauccug	1140
gaaaagaugg acggaacaga agaacugcug gucaagcuga acagagaaga ccugcugaga	1200
aagcagagaa cauucgacaa cggaagcauc ccgcaccaga uccaccuggg agaacugcac	1260
gcaauccuga gaagacagga agacuucuac ccguuccuga aggacaacag agaaaagauc	1320
gaaaagaucc ugacauucag aaucccgua uacgucggac cgcuggcaag aggaaacagc	1380
agauucgcau ggaugacaag aaagagcgaa gaaacaauca caccguggaa cuucgaagaa	1440
gucgucgaca agggagcaag cgcacagagc uucaucgaaa gaaugacaaa cuucgacaag	1500
aaccugccga acgaaaaggu ccugccgaag cacagccugc uguacgaaua cuucacaguc	1560
uacaacgaac ugacaaaaggu caaguacguc acagaaggaa ugagaaagcc ggcauuccug	1620

30

40

50

agcggagaac agaagaaggc aaucgucgac cugcuguuca agacaaacag aaaggucaca 1680
 gucaagcagc ugaaggaaga cuacuucaag aagaucgaau gcuucgacag cgucgaaauc 1740
 agcggagucg aagacagauu caacgcaagc cugggaacau accacgaccu gcugaagauc 1800
 aucaaggaca aggacuuccu ggacaacgaa gaaaacgaag acauccugga agacaucguc 1860
 cugacacuga cacuguucga agacagagaa augaucgaag aaagacugaa gacauacgca 1920
 caccuguucg acgacaaggu caugaagcag cugaagagaa gaagauacac aggaugggga 1980
 agacugagca gaaagcugau caacggaauc agagacaagc agagcggaaa gacaauccug 2040
 gacuuccuga agagcgcgacgg auucgcaaac agaaacuuca ugcagcugau ccacgacgac 2100
 agccugacau ucaaggaaga cauccagaag gcacagguca gcggacaggg agacagccug 2160
 cacgaacaca ucgcaaaccu ggcaggaagc ccggcaauca agaagggaau ccugcagaca 2220
 gucaaggucg ucgacgaacu ggucaagguc augggaagac acaagccgga aaacaucguc 2280
 aucgaaaugg caagagaaaa ccagacaaca cagaagggac agaagaacag cagagaaaga 2340
 augaagagaa ucgaagaagg aaucagaaga cugggaagcc agauccugaa ggaacacccg 2400
 gucgaaaaca cacagcugca gaacgaaaag cuguaccugu acuaccugca gaacggaaaga 2460
 gacauguacg ucgaccagga acuggacauc aacagacuga gcgacuacga cgucgacgca 2520
 aucgucccg agagcuuccu gaaggacgac agcaucgaca acaagguccu gacaagaagc 2580
 gacaagaaca gaggaagag cgacaacguc ccgagcgaag aagucgucaa gaagaugaag 2640
 aacucugga gacagcugcu gaacgcaaag cugaucacac agagaaaguu cgacaaccug 2700
 acaaaggcag agagaggagg acugagcga cuggacaagg caggauucau caagagacag 2760
 cuggucgaaa caagacagau cacaagcac gucgcacaga uccuggacag cagaaugaac 2820
 acaaaguacg acgaaaacga caagcugauc agagaaguca aggucaucac acugaagagc 2880
 aagcugguca gcgacuucag aaaggacuuc caguucuaa agguccagaga aaucacaac 2940
 uaccaccag cacacgacgc auaccugaac gcagucgucg gaacagcacu gaucaagaag 3000
 uaccgaagc uggaagcga auucgucua ggagacuaca agguccuacga cgucagaaag 3060
 augaucgcaa agagcgaaca ggaaucgga aaggcaacag caaaguacuu cuucacagc 3120
 aacauauga acuucuuca gacagaaauc acacuggcaa acggagaaau cagaaagaga 3180
 ccgcucaucg aaacaaacgg agaaacagga gaaucgucu gggacaaggg aagagacuuc 3240
 gcaacaguca gaaagguccu gagcaugccg cagguaaca ucgucaagaa gacagaaguc 3300
 cagacaggag gauucagcaa ggaaagcauc cugccgaaga gaaacagcga caagcugauc 3360
 gcaagaaaga aggacuggga cccgaagaag uacggaggau ucgacagccc gacagucgca 3420
 uacagcgucc ugguccgucg aaaggucgaa aagggaaga gcaagaagcu gaagagcguc 3480
 aaggaaucg uggaauacac aaucaggaaga agaagcagcu ucgaaaagaa cccgaucgac 3540
 uuccuggaag caaagggaau caaggaaguc aagaaggacc ugaucauca gcugccgaag 3600
 uacagccugu ucgaacugga aaacggaaaga aagagaaguc uggcaagcgc aggagaacug 3660
 cagaaggga acgaacuggc acugccgagc aaguacguc acuuccugua ccuggcaagc 3720
 cacuacgaaa agcugaaggg aagcccggaa gacaacgaac agaagcagcu guucgucgaa 3780
 cagcacaagc acuaccugga cgaaaucac gaacagauca gcgaauucag caagagaguc 3840
 auccuggcag acgcaaaccu ggacaagguc cugagcgau acaacaagca cagagacaag 3900
 ccgaucagag aacaggcaga aaacaucauc caccuguuca cacugacaaa ccugggagca 3960
 ccggcagcau ucaaguacuu cgacacaaca aucgacagaa agagauacac aagcacaag 4020
 gaaguccugg acgcaacacu gauccaccag agcaucacag gacugucgaa aacaagauc 4080
 gaccugagcc agcugggagg agacggagga ggaagcccga agaagaagag aaaggucuag 4140

10

20

30

40

<210> 10

<211> 4134

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic

50

polynucleotide"

<400> 10

gacaagaagu acagcaucgg acuggacauc ggaacaaaca gcgucggaug ggcagucauc	60	
acagacgaau acaagguccc gagcaagaag uucaaggucc ugggaaacac agacagacac	120	
agcaucaaga agaaccugau cggagcacug cuguucgaca gcggagaaac agcagaagca	180	
acaagacuga agagaacagc aagaagaaga uacacaagaa gaaagaacag aaucugcuac	240	
cugcaggaaa ucuucagcaa cgaauuggca aaggucgacg acagcuucuu ccacagacug	300	
gaagaaagcu uccuggucga agaagacaag aagcacgaaa gacacccgau cuucggaaac	360	
aucgucgacg aagucgcaua ccacgaaaag uacccgacaa ucuaccaccu gagaaagaag	420	
cuggucgaca gcacagacaa ggacagaccug agacugaucu accuggcacu ggcacacaug	480	10
aucaaguuca gaggacacuu ccugaucgaa ggagaccuga acccgacaa cagcgacguc	540	
gacaagcugu ucauccagcu gguccagaca uacaaccagc uguucgaaga aaacccgauc	600	
aacgcaagcg gagucgacgc aaaggcaauc cugagcgcaa gacugagcaa gagcagaaga	660	
cuggaaaacc ugaucgcaca gcugccggga gaaaagaaga acggacuguu cggaaccug	720	
aucgcacuga gccugggacu gacaccgaac uucaagagca acuucgaccu ggcagaagac	780	
gcaaagcugc agcugagcaa ggacacauac gacgacgacc uggacaaccu gcuggcacag	840	
aucggagacc aguacgcaga ccuguuccug gcagcaaaga accugagcga cgcaauccug	900	
cugagcgaca uccugagagu caacacagaa aucacaaagg caccgcugag cgcaagcaug	960	
aucaagagau acgacgaaca ccaccaggac cugacacugc ugaaggcacu ggucagacag	1020	
cagcugccgg aaaaguacaa ggaaaucuuc uucgaccaga gcaagaacgg auacgcagga	1080	20
uacaucgacg gaggagcaag ccaggaaaga uucuacaagu ucaucaagcc gauccuggaa	1140	
aagauggacg gaacagaaga acugcugguc aagcugaaca gagaagaccu gcugagaaag	1200	
cagagaacau ucgacaacgg aagcaucccg caccagaucc accugggaga acugcacgca	1260	
auccugagaa gacaggaaga cuucuacccg uuccugaagg acaacagaga aaagaucgaa	1320	
aagauccuga cauucagaau cccguacuac gucggaccgc uggcaagagg aaacagcaga	1380	
uucgcaugga ugacaagaaa gagcgaagaa acaaucacac cguggaacu u cgaagaaguc	1440	
gucgacaagg gagcaagcgc acagagcuuc aucgaaagaa ugacaaacuu cgacaagaac	1500	
cugccgaacg aaaagguccu gccgaagcac agccugcugu acgaauacu c acagucua	1560	
aacgaacuga caaaggucua guacgucaca gaaggaauga gaaagccggc auuccugagc	1620	
ggagaacaga agaaggcaau cgucgaccug cuguucaaga caaacagaaa ggucacaguc	1680	30
aagcagcuga aggaagacua cuucaagaag aucgaaugcu ucgacagcgu cgaaaucagc	1740	
ggagucgaag acagauucua cgcaagccug ggaacauacc acgaccugcu gaagaucac	1800	
aaggacaagg acuuccugga caacgaagaa aacgaagaca uccuggaaga caucguccug	1860	
acacugacac uguucgaaga cagagaaaug aucgaagaaa gacugaagac auacgcacac	1920	
cuguucgacg acaaggucua gaagcagcug aagagaagaa gauacacagg auggggaaga	1980	
cugagcagaa agcugaucaa cggaauacaga gacaagcaga gcggaagac aauccuggac	2040	
uuccugaaga gcgacggauu cgcaaacaga aacuucaugc agcugaucca cgacgacagc	2100	
cugacauuca aggaagacau ccagaaggca caggucagcg gacagggaga cagccugcac	2160	
gaacacaucg caaaccuggc aggaagcccg gcaaucaaga agggaauccu gcagacaguc	2220	
aaggucgucg acgaacuggu caaggucaug ggaagacaca agccggaaaa caucgucac	2280	40
gaaauggcaa gagaaaacca gacaacacag aagggacaga agaacagcag agaaagaag	2340	
aagagaauucg aagaaggaa u caaggaaucg ggaagccaga uccugaagga acacccgguc	2400	
gaaaacacac agcugcagaa cgaauagcug uaccuguacu accugcagaa cggaagagac	2460	
auguacgucg accaggaacu ggacaucaac agacugagcg acuacgacgu cgaccacauc	2520	
gucccgacga gcuuccugaa ggacgacagc aucgacaaca agguccugac aagaagcgac	2580	
aagaacagag gaaagagcga caacgucccg agcgaagaag ucgucaagaa gaugaagaac	2640	
uacuggagac agcugcugaa cgaaagcug aucacacaga gaaaguucga caaccugaca	2700	
aaggcagaga gaggaggacu gagcgaacug gacaaggcag gauucaucaa gagacagcug	2760	
gucgaaacaa gacagauac aaagcacguc gcacagauc uggacagcag aaugaacaca	2820	
aaguacgacg aaaacgacaa gcugaucaga gaagucaagg ucaucacacu gaagagcaag	2880	50

cuggucagcg acuucagaaa ggacuuccag uucuacaagg ucagagaaau caacaacuac 2940
 caccacgcac acgacgcaua ccugaacgca gucgucggaa cagcacugau caagaaguac 3000
 ccgaagcugg aaagcgaaau cgucuacgga gacuacaagg ucuacgacgu cagaaagaug 3060
 aucgcaaaga gcgaaacagga aaucggaaaag gcaacagcaa aguacuucuu cuacagcaac 3120
 aucaugaacu ucuucaagac agaaaucaca cuggcaaacg gagaaaucag aaagagaccg 3180
 cugaucgaaa caaacggaga aacaggagaa aucgucuggg acaagggaag agacuucgca 3240
 acagucagaa agguccugag caugccgcag gucaacaucg ucaagaagac agaaguccag 3300
 acaggaggau ucagcaagga aagcauccug ccgaagagaa acagcgacaa gcugaucgca 3360
 agaaagaagg acugggaccc gaagaaguac ggaggauucg acagcccgcac agucgcgauac 3420
 agcguccugg ucgucgcaaa ggucgaaaag ggaaagagca agaagcugaa gagcguaag 3480
 gaacugcugg gaaucacaau cauggaaaaga agcagcuucg aaaagaaccc gaucgacuuc 3540
 cuggaagcaa agggauacaa ggaagucaag aaggaccuga ucaucaagcu gccgaaguac 3600
 agccuguucg aacuggaaaa cggaagaaaag agaauvcugg caagcgcagg agaacugcag 3660
 aagggaacg aacuggcacu gccgagcaag uacgucaacu uccuguaccu ggcaagccac 3720
 uacgaaaagc ugaagggaag cccggaagac aacgaacaga agcagcuguu cgucgaacag 3780
 cacaagcacu accuggacga aaucaucgaa cagaucagcg aaucagcaa gagagucauc 3840
 cuggcagacg caaaccugga caagguccug agcgcauaca acaagcacag agacaagccg 3900
 aucagagaac aggcagaaaa caucauccac cuguucacac ugacaaaccu gggagcaccg 3960
 gcagcauua aguacuucga cacaacaauc gacagaaaga gauacacaag cacaaggaa 4020
 guccuggacg caacacugau ccaccagagc aucacaggac uguacgaaac aagaaucgac 4080
 cugagccagc ugggaggaga cggaggagga agcccgaaga agaagagaaa gguc 4134

10

20

<210> 11

<211> 4134

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

30

<400> 11

gacaagaagu acagcaucgg acuggcaauc ggaacaaaca gcgucggaug ggcagucauc 60
 acagacgaau acaagguccc gagcaagaag uucaaggucc ugggaaacac agacagacac 120
 agcaucaaga agaaccugau cggagcacug cuguucgaca gcgagaaaac agcagaagca 180
 acaagacuga agagaacagc aagaagaaga uacacaagaa gaaagaacag aaucugcuac 240
 cugcaggaaa ucuucagcaa cgaaauggca aaggucgacg acagcuucuu ccacagacug 300
 gaagaaagcu uccuggucga agaagacaag aagcacgaaa gacaccgga cuucggaaac 360
 aucgucgacg aagucgcaua ccacgaaaag uacccgacaa ucuaccaccu gagaaagaag 420
 cuggucgaca gcacagacaa ggcagaccug agacugaucu accuggcacu ggcacacaug 480
 aucaaguuca gaggacacuu ccugaucgaa ggagaccuga acccgacaa cagcgacguc 540
 gacaagcugu ucauccagcu gguccagaca uacaaccagc uguucgaaga aaaccggauc 600
 aacgcaagcg gagucgacgc aaaggcaauc cugagcgcaa gacugagcaa gagcagaaga 660
 cuggaaaacc ugaucgcaca gcugccggga gaaaagaaga acggacuguu cggaaccug 720
 aucgcacuga gccugggacu gacaccgaac uucaagagca acuucgaccu ggcagaagac 780
 gcaaagcugc agcugagcaa ggacacauac gacgacgacc uggacaaccu gcuggcacag 840
 aucggagacc aguacgcaga ccuguuccug gcagcaaaga accugagcga cgcaauccug 900
 cugagcgaca uccugagagu caacacagaa aucacaaagg caccgcugag cgcaagcaug 960
 aucaagagau acgacgaaca ccaccaggac cugacacugc ugaaggcacu ggucagacag 1020
 cagcugccgg aaaaguacaa ggaaaucuu uucgaccaga gcaagaacgg auacgcagga 1080
 ucaucgacg gaggagcaag ccaggaagaa uucuacaagu ucaucaagcc gauccuggaa 1140

40

50

aagauggacg gaacagaaga acugcugguc aagcugaaca gagaagaccu gcugagaaag	1200	
cagagaacau ucgacaacgg aagcaucccg caccagaucc accugggaga acugcacgca	1260	
auccugagaa gacaggaaga cuucuacccg uuccugaagg acaacagaga aaagaucgaa	1320	
aagauccuga cauucagaau cccguacuac gucggaccgc uggcaagagg aaacagcaga	1380	
uucgcaugga ugacaagaaa gagcgaagaa acaaucacac cguggaacu cgaagaaguc	1440	
gucgacaagg gagcaagcgc acagagcuuc aucgaaagaa ugacaaacu cgacaagaac	1500	
cugccgaacg aaaagguccu gccgaagcac agccugcugu acgaauacu cacagucuac	1560	
aacgaacuga caaaggucua guacgucaca gaaggaauga gaaagccggc auuccugagc	1620	
ggagaacaga agaaggcaau cgucgaccug cuguucaaga caaacagaaa ggucacaguc	1680	
aagcagcuga aggaagacua cuucaagaag aucgaaugcu ucgacagcgu cgaaaucagc	1740	10
ggagucgaag acagauucaa cgcaagccug ggaacauacc acgaccugcu gaagaucac	1800	
aaggacaagg acuuccugga caacgaagaa aacgaagaca uccuggaaga caucguccug	1860	
acacugacac uguucgaaga cagagaaaug aucgaagaaa gacugaagac auacgcacac	1920	
cuguucgacg acaaggucua gaagcagcug aagagaagaa gauacacagg auggggaaga	1980	
cugagcagaa agcugaucaa cggaaucaga gacaagcaga gcggaagac aauccuggac	2040	
uuccugaaga gcgacggauu cgcaaacaga aacuuaugc agcugaucca cgacgacagc	2100	
cugacauuca aggaagacau ccagaaggca caggucagcg gacagggaga cagccugcac	2160	
gaacacaucg caaaccuggc aggaagcccg gcaaucaaga agggaaucuu gcagacaguc	2220	
aaggucgucg acgaacuggu caaggucaug ggaagacaca agccggaaaa caucgucauc	2280	
gaaauggcaa gagaaaacca gacaacacag aaggagcaga agaacagcag agaaagaug	2340	20
aagagaaucg aagaaggaa caaggaacug ggaagccaga uccugaagga acaccgguc	2400	
gaaaacacac agcugcagaa cgaaaagcug uaccuguacu accugcagaa cggaagagac	2460	
auguacgucg accaggaacu ggacaucaac agacugagcg acuacgacgu cgaccacau	2520	
gucccgacga gcuuccugaa ggacgacagc aucgacaaca agguccugac aagaagcgac	2580	
aagaacagag gaaagagcga caacgucccg agcgaagaag ucgucaagaa gaugaagaac	2640	
uacuggagac agcugcugaa cgcaaacug aucacacaga gaaaguucga caaccugaca	2700	
aaggcagaga gaggaggacu gagcgaacug gacaaggcag gauucaucaa gagacagcug	2760	
gucgaaacaa gacagauca aaagcacguc gcacagaucc uggacagcag aaugaacaca	2820	
aaguacgacg aaaacgacaa gcugaucaga gaagucaagg ucaucacacu gaagagcaag	2880	
cuggucagcg acuucagaaa ggacuuccag uucuacaagg ucagagaaa caacaacuac	2940	30
caccacgcac acgacgcaua ccugaacgca gucgucggaa cagcacugau caagaaguac	3000	
ccgaagcugg aaagcgaauu cgucuacgga gacuacaagg ucuacgacgu cagaaagaug	3060	
aucgcaaaga gcgaacagga aaucggaaag gcaacagcaa aguacuucuu cuacagcaac	3120	
aucaugaacu ucuucaagac agaaaucaca cuggcaaacg gagaaaucag aaagagaccg	3180	
cugaucgaaa caaacggaga aacaggagaa aucgucuggg acaagggaag agacuucgca	3240	
acagucagaa agguccugag caugccgcag gucaacaucg ucaagaagac agaaguccag	3300	
acaggaggau ucagcaagga aagcauccug ccgaagagaa acagcgacaa gcugaucgca	3360	
agaaagaagg acugggaccc gaagaaguac ggaggauucg acagcccgcac agucgcacuac	3420	
agcguccugg ucgucgcaaa ggucgaaaag ggaagagca agaagcugaa gagcgucaag	3480	
gaacugcugg gaauacaaau cauggaaaga agcagcuucg aaaagaacc gaucgacuuc	3540	40
cuggaagcaa agggauacaa ggaagucaag aaggaccuga ucaucaagcu gccgaaguac	3600	
agccuguucg aacuggaaaa cggaagaaag agaauugcugg caagcgacagg agaacugcag	3660	
aagggaacg aacuggcacu gccgagcaag uacgucaacu uccuguaccu ggcaagccac	3720	
uacgaaaagc ugaagggaag cccggaagac aacgaacaga agcagcuguu cgucgaacag	3780	
cacaagcacu accuggacga aaucaucgaa cagauagcg aaucagcaa gagagucauc	3840	
cuggcagacg caaaccugga caagguccug agcgcauaca acaagcacag agacaagccg	3900	
aucagagaac aggcagaaaa caucauccac cuguucacac ugacaaaccu gggagcaccg	3960	
gcagcauua aguacuucga cacaacaauc gacagaaaga gauacacaag cacaaggaag	4020	
guccuggacg caacacugau ccaccagagc aucacaggac uguacgaaac aagaaucgac	4080	
cugagccagc ugggaggaga cggaggagga agcccgaaga agaagagaaa gguc	4134	50

<210> 12
 <211> 4134
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"
 <400> 12

gacaagaagu acagcaucgg acuggcaauc ggaacaaaca gcgucggaug ggcagucauc 60
 acagacgaau acaagguccc gagcaagaag uucaaggucc ugggaaacac agacagacac 120
 agcaucaaga agaaccugau cggagcacug cuguucgaca gcggagaaac agcagaagca 180
 acaagacuga agagaacagc aagaagaaga uacacaagaa gaaagaacag aaucugcuac 240
 cugcaggaaa ucuucagcaa cgaaauggca aaggucgacg acagcuucuu ccacagacug 300
 gaagaaagcu uccuggucga agaagacaag aagcacgaaa gacacccgau cuucggaaac 360
 aucgucgacg aagucgcaua ccacgaaaag uacccgacaa ucuaccaccu gagaaagaag 420
 cuggucgaca gcacagacaa ggacagaccug agacugaucu accuggcacu ggacacaug 480
 aucaaguuca gagggacacuu ccugaucgaa ggagaccuga acccgacaa cagcgacguc 540
 gacaagcugu ucauccagcu gguccagaca uacaaccagc uguucgaaga aaacccgauc 600
 aacgcaagcg gagucgacgc aaaggcaauc cugagcgcaa gacugagcaa gagcagaaga 660
 cuggaaaacc ugaucgcaca gcugccggga gaaaagaaga acggacuguu cggaaccug 720
 aucgcacuga gccugggacu gacaccgaac uucaagagca acuucgaccu ggacagaagac 780
 gcaaagcugc agcugagcaa ggacacauac gacgacgacc uggacaaccu gcuggcacag 840
 aucggagacc aguacgcaga ccuguuccug gcagcaaaga accugagcga cgcaauccug 900
 cugagcgaca uccugagagu caacacagaa aucacaaagg caccgcugag cgcaagcaug 960
 aucaagagau acgacgaaca ccaccaggac cugacacugc ugaaggcacu ggucagacag 1020
 cagcugccgg aaaaguacaa ggaaaucuuc uucgaccaga gcaagaacgg auacgcagga 1080
 ucaucgacg gagggagcaag ccaggaagaa uucuacaagu ucaucaagcc gauccuggaa 1140
 aagauggacg gaacagaaga acugcugguc aagcugaaca gagaagaccu gcugagaaag 1200
 cagagaacau ugcacaacgg aagcaucccg caccagauc accugggaga acugcacgca 1260
 auccugagaa gacaggaaga cuucuacccg uuccugaagg acaacagaga aaagaucgaa 1320
 aagauccuga cauucagaau cccguacuac gucggaccgc uggcaagagg aaacagcaga 1380
 uucgcaugga ugacaagaaa gagcgaagaa acaaucacac cguggaacuu cgaagaaguc 1440
 gucgacaagg gagcaagcgc acagagcuuc aucgaaagaa ugacaaacuu cgacaagaac 1500
 cugccgaacg aaaagguccu gccgaagcac agccugcugu acgaauacuu cacagucuac 1560
 aacgaacuga caaaggucua guacgucaca gaaggaauga gaaagccggc auuccugagc 1620
 ggagaacaga agaaggcaau cgucgaccug cuguucaaga caaacagaaa ggucacaguc 1680
 aagcagcuga aggaagacua cuucaagaag aucgaaugcu ucgacagcgu cgaaaucagc 1740
 ggagucgaag acagauucaa cgcaagccug ggaacauacc acgaccugcu gaagaucac 1800
 aaggacaagg acuuccugga caacgaagaa aacgaagaca uccuggaaga caucguccug 1860
 acacugacac uguucgaaga cagagaaaug aucgaagaaa gacugaagac auacgcacac 1920
 cuguucgacg acaaggucua gaagcagcug aagagaagaa gauacacagg auggggaaga 1980
 cugagcagaa agcugaucaa cggaaucaga gacaagcaga gcggaaagac aauccuggac 2040
 uuccugaaga gcgacggauu cgcaaacaga aacuucaugc agcugaucca cgacgacagc 2100
 cugacauuca aggaagacau ccagaaggca caggucagcg gacagggaga cagccugcac 2160
 gaacacaucg caaaccuggc aggaagcccg gcaaucaaga agggaaucuu gcagacaguc 2220
 aaggucgucg acgaacuggu caaggucaug ggaagacaca agccggaaaa caucgucauc 2280
 gaaauggcaa gagaaaacca gacaacacag aaggggacaga agaacagcag agaaagaag 2340
 aagagaauucg aagaaggaa caaggaacug ggaagccaga uccugaagga acacccgguc 2400

10

20

30

40

50

gaaaacacac agcugcagaa cgaaaagcug uaccuguacu accugcagaa cggaagagac 2460
 auguacgucg accaggaacu ggacaucaac agacugagcg acuacgacgu cgacgcaauc 2520
 gucccgcaga gcuuccugaa ggacgacagc aucgacaaca agguccugac aagaagcgac 2580
 aagaacagag gaaagagcga caacgucccg agcgaagaag ucgucaagaa gaugaagaac 2640
 uacuggagac agcugcugaa cgcaaagcug aucacacaga gaaaguucga caaccugaca 2700
 aaggcagaga gaggaggacu gagcgaacug gacaaggcag gauucauca gagacagcug 2760
 gucgaaacaa gacagauca aaagcacguc gcacagaucc uggacagcag aaugaacaca 2820
 aaguacgacg aaaacgacaa gcugaucaga gaagucaagg ucaucacacu gaagagcaag 2880
 cuggucagcg acuucagaaa ggacuuccag uucuacaagg ucagagaaaau caacaacuac 2940
 caccacgcac acgacgcaua ccugaacgca gucgucggaa cagcacugau caagaaguac 3000
 ccgaagcugg aaagcgaaau cgucuacgga gacuacaagg ucuacgacgu cagaaagaug 3060
 aucgcaaaga gcgaacagga aaucggaaag gcaacagcaa aguacuucuu cuacagcaac 3120
 aucaugaacu ucuucaagac agaaaucaca cuggcaaacg gagaaaucag aaagagaccg 3180
 cugaucgaaa caaacggaga aacaggagaa aucgucuggg acaagggaag agacuucgca 3240
 acagucagaa agguccugag caugccgcag gucaacaucg ucaagaagac agaaguccag 3300
 acaggaggau ucagcaagga aagcauccug ccgaagagaa acagcgacaa gcugaucgca 3360
 agaaagaagg acugggaccc gaagaaguac ggaggauucg acagcccgac agucgcauac 3420
 agcuccugg ucgucgcaaa ggucgaaaag ggaaagagca agaagcugaa gagcguaag 3480
 gaacugcugg gaaucacaa cauggaaaga agcagcuucg aaaagaaccc gaucgacuuc 3540
 cuggaagcaa agggauacaa ggaagucaag aaggaccuga ucaucaagcu gccgaaguac 3600
 agccuguucg aacuggaaaa cggaagaaag agaaugcugg caagcgagg agaacugcag 3660
 aagggaacg aacuggcacu gccgagcaag uacgucaacu uccuguaccu ggcaagccac 3720
 uacgaaaagc ugaagggaag cccggaagac aacgaacaga agcagcuguu cgucgaacag 3780
 cacaagcacu accuggacga aaucaucgaa cagaucagcg aaucagcaa gagagucauc 3840
 cuggcagacg caaacugga caagguccug agcgcauaca acaagcacag agacaagccg 3900
 aucagagaac aggcagaaaa caucauccac cuguucacac ugacaaaccu gggagcaccg 3960
 gcagcauua aguacuucga cacaacauc gacagaaaga gauacacaag cacaaggaa 4020
 guccuggacg caacacugau ccaccagagc aucacaggac uguacgaaac aagaaucgac 4080
 cugagccagc ugaggaggaga cggaggagga agcccgaaga agaagagaaa gguc 4134

10

20

30

<210> 13

<211> 1368

<212> PRT

<213> Artificial Sequence

<220>

<221> source

 <223> /note="Description of Artificial Sequence: Synthetic
 polypeptide"

<400> 13

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile

35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu

50 55 60

40

50

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
290 295 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305 310 315 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys

10

20

30

40

50

325 330 335
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
340 345 350
Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
355 360 365
Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
370 375 380
Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385 390 395 400
Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415
Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430
Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
435 440 445
Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460
Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480
Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495
Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510
Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525
Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540
Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560
Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575
Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590

10

20

30

40

50

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655

10

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700

20

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
740 745 750

30

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815

40

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
850 855 860

50

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
 865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
 885 890 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
 900 905 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
 915 920 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
 930 935 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
 945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
 965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
 980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
 995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
 1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
 1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
 1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
 1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
 1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
 1085 1090 1095

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys
 1100 1105 1110

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro

10

20

30

40

50

1115	1120	1125
Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val		
1130	1135	1140
Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys		
1145	1150	1155
Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser		
1160	1165	1170
Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys		
1175	1180	1185
Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu		
1190	1195	1200
Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly		
1205	1210	1215
Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val		
1220	1225	1230
Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser		
1235	1240	1245
Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys		
1250	1255	1260
His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys		
1265	1270	1275
Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala		
1280	1285	1290
Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn		
1295	1300	1305
Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala		
1310	1315	1320
Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser		
1325	1330	1335
Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr		
1340	1345	1350
Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp		
1355	1360	1365

10

20

30

40

50

<210> 14
 <211> 4107
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"
 <400> 14

auggacaaga aguacagcau cggacuggac aucggaacaa acagcgucgg augggcaguc	60	10
aucacagacg aauacaaggu cccgagcaag aaguucaagg uccugggaaa cacagacaga	120	
cacagcauca agaagaaccu gaucggagca cugcuguucg acagcggaga aacagcagaa	180	
gcaacaagac ugaagagaac agcaagaaga agauacacaa gaagaaagaa cagaauucgc	240	
uaccugcagg aaaucuucag caacgaaaug gcaaaaggucg acgacagcuu cuuccacaga	300	
cuggaagaaa gcuuccuggu cgaagaagac aagaagcacg aaagacaccc gaucuucgga	360	
aauaucgucg acgaagucgc auaccacgaa aaguacccga caaucuacca ccugagaaag	420	
aagcuggucg acagcacaga caaggcagac cugagacuga ucuaccuggc acuggcacac	480	
augaucaagu ucagaggaca cuuccugauc gaaggagacc ugaacccgga caacagcgac	540	
gucgacaagc uguucaucca gcugguccag acauacaacc agcuguucga agaaaacccg	600	
aucaacgcaa gcgagagucga cgcaaaggca auccugagcg caagacugag caagagcaga	660	20
agacuggaaa accugaucgc acagcugccg ggagaaaaga agaacggacu guucggaaac	720	
cugaucgcac ugagccuggg acugacaccg aacuuaaga gcaacuucga ccuggcagaa	780	
gacgcaaagc ugcagcugag caaggacaca uacgacgacg accuggacaa ccugcuggca	840	
cagaucggag accaguacgc agaccuguuc cuggcagcaa agaaccugag cgacgcaauc	900	
cugcugagcg acauccugag agucaacaca gaaauacaaa aggcaccgcu gagcgcaagc	960	
augaucaaga gauacgacga acaccaccag gaccugacac ugcugaaggc acuggucaga	1020	
cagcagcugc cggaagaaag caaggaaauu uucuucgacc agagcaagaa cggauacgca	1080	
ggauacaucg acggaggagc aagccaggaa gaaucuuaa aguucaucaa gccgauccug	1140	
gaaaagaugg acggaacaga agaacugcug gucaagcuga acagagaaga ccugcugaga	1200	
aagcagagaa cauucgacaa cggaagcauc ccgcaccaga uccaccuggg agaacugcac	1260	30
gcaauccuga gaagacagga agacuucuac ccguuccuga aggacaacag agaaaagauc	1320	
gaaaagaucc ugacauucag aaucccgua uacgucggac cgcuggcaag aggaaacagc	1380	
agauucgcau ggaugacaag aaagagcgaa gaaacaauca caccugggaa cuucgaagaa	1440	
gucgucgaca agggagcaag cgcacagagc uucaucgaaa gaaugacaaa cuucgacaag	1500	
aaccugccga acgaaaaggu ccugccgaag cacagccugc uguacgaaua cuucacaguc	1560	
uacaacgaac ugacaaaaggu caaguacguc acagaaggaa ugagaaagcc ggcauuccug	1620	
agcggagaac agaagaaggc aaucgucgac cugcuguuca agacaaacag aaaggucaca	1680	
gucaagcagc ugaaggaaga cuacuuaag aagaucgaau gcuucgacag cgucgaaauc	1740	
agcggagucg aagacagauu caacgcaagc cugggaacau accacgaccu gcugaagauc	1800	
aucaaggaca aggacuuccu ggacaacgaa gaaaacgaag acauccugga agacaucguc	1860	40
cugacacuga cacuguucga agacagagaa augaucgaag aaagacugaa gacauacgca	1920	
caccuguucg acgacaaggu caugaagcag cugaagagaa gaagauacac aggaugggga	1980	
agacugagca gaaagcugau caacggaauc agagacaagc agagcggaaa gacaauccug	2040	
gacuuccuga agagcgacgg auucgcaaac agaaacuua ugcagcugau ccacgacgac	2100	
agccugacau ucaaggaaga cauccagaag gcacagguca gcggacaggg agacagccug	2160	
cacgaacaca ucgcaaaccu ggcaggaaag ccggcaauca agaagggaau ccugcagaca	2220	
gucaaggucg ucgacgaacu ggucaagguc augggaagac acaagccgga aaacaucguc	2280	
aucgaaaugg caagagaaaa ccagacaaca cagaagggac agaagaacag cagagaaaga	2340	
augaagagaa ucgaagaagg aaucaggaa cugggaagcc agauccugaa ggaacacccg	2400	
gucgaaaaca cacagcugca gaacgaaaag cuguaccugu acuaccugca gaacggaaga	2460	50

gacauguacg ucgaccagga acuggacauc aacagacuga gcgacuacga cgucgaccac 2520
 aucgucccg agagcuuccu gaaggacgac agcaucgaca acaagguccu gacaagaagc 2580
 gacaagaaca gaggaagag cgacaacguc ccgagcgaag aagucgucaa gaagaugaag 2640
 aacuacugga gacagcugcu gaacgcaaag cugaucacac agagaaaguu cgacaaccug 2700
 acaaaggcag agagaggagg acugagcgaa cuggacaagg caggauucau caagagacag 2760
 cuggucgaaa caagacagau cacaagcac gucgcacaga uccuggacag cagaauaac 2820
 acaaaguacg acgaaaacga caagcugauc agagaaguca aggucaucac acugaagagc 2880
 aagcugguca gcgacuucag aaaggacuuc caguucuaa agguacagaga aaucacaac 2940
 uaccaccacg cacacgacgc auaccugaac gcagucgucg gaacagcacu gaucaagaag 3000
 uacccgaagc uggaaagcga auucgucuac ggagacuaca aggucuacga cgucagaaag 3060
 augaucgcaa agagcgaaca ggaaucgga aaggcaacag caaaguacuu cuucacagc 3120
 aacauauga acuuucaa gacagaaauc acacuggcaa acggagaaa cagaaagaga 3180
 ccgcuagucg aaacaaacgg agaaacagga gaaucgucu gggacaaggg aagagacuuc 3240
 gcaacaguca gaaagguccu gagcaugccg cagguaaca ucgucaaga gacagaaguc 3300
 cagacaggag gauucagcaa ggaaagcauc cugccgaaga gaaacagcga caagcugauc 3360
 gcaagaaaga aggacuggga cccgaagaag uacggaggau ucgacagccc gacagucgca 3420
 uacagcgucc uggucgucgc aaaggucgaa aagggaaga gcaagaagcu gaagagcguc 3480
 aaggaaucg ugggaauac aaucaugga agaagcagcu ucgaaaaga cccgaucgac 3540
 uuccuggaag caaagggauc caaggaaguc aagaaggacc ugaucacaa gcugccgaag 3600
 uacagccugu ucgaacugga aaacggaaga aagagaagc uggcaagcg aggagaacug 3660
 cagaaggga acgaacuggc acugccgagc aaguacguca acuuucgua ccuggcaagc 3720
 cacuacgaaa agcugaaggg aagcccggaa gacaacgaac agaagcagcu guucgucgaa 3780
 cagcacaagc acuaccugga cgaaucauc gaacagauca gcgaauucag caagagaguc 3840
 auccuggcag acgcaaaccu ggacaagguc cugagcgcau acaacaagca cagagacaag 3900
 ccgaucagag aacaggcaga aaacaucauc caccuguuca cacugacaaa ccugggagca 3960
 ccggcagcau ucaaguacuu cgacacaaca aucgacagaa agagauacac aagcacaag 4020
 gaaguccug acgcaacacu gauccaccag agcaucacag gacuguacga aacaagauc 4080
 gaccugagcc agcugggagg agacuag 4107

10

20

<210> 15

<211> 4101

<212> RNA

<213> Artificial Sequence

<220>

<221> source

 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 15

gacaagaagu acagcaucgg acuggacauc ggaacaaaca gcgucggaug ggcagucauc 60
 acagacgaau acaagguccc gagcaagaag uucaaggucc ugggaaacac agacagacac 120
 agcaucaaga agaaccugau cggagcacug cuguucgaca gcgagaaac agcagaagca 180
 acaagacuga agagaacagc aagaagaaga uacacaaga gaaagaacag aaucugcuac 240
 cugcaggaaa ucuucagcaa cgaauaggca aaggucgacg acagcuucuu ccacagacug 300
 gaagaaagcu uccuggucga agaagacaag aagcacgaaa gacaccgcu cuucggaaac 360
 aucgucgacg aagucgcau ccacgaaaag uacccgacaa ucuaccaccu gagaaagaag 420
 cuggucgaca gcacagacaa ggacagaccug agacugaucu accuggcacu ggacacaug 480
 aucaaguuca gaggacacuu ccugaucgaa ggagaccuga acccgacaa cagcagcugc 540
 gacaagcugu ucauccagcu gguccagaca uacaaccagc uguucgaaga aaaccgauc 600
 aacgcaagcg gagucgacgc aaaggcauc cugagcgcaa gacugagcaa gagcagaaga 660
 cuggaaaacc ugaucgcaca gcugccggga gaaaagaaga acggacuguu cggaaccug 720

30

40

50

aucgcacuga gccugggacu gacaccgaac uucaagagca acuucgaccu ggcagaagac 780
 gcaaagcugc agcugagcaa ggacacauac gacgacgacc uggacaaccu gcuggcacag 840
 aucggagacc aguacgcaga ccuguuccug gcagcaaaga accugagcga cgcaauccug 900
 cugagcgaca uccugagagu caacacagaa aucacaaagg caccgcugag cgcaagcaug 960
 aucaagagau acgacgaaca ccaccaggac cugacacugc ugaaggcacu ggucagacag 1020
 cagcugccgg aaaaguacaa ggaaauuc uucgaccaga gcaagaacgg auacgcagga 1080
 uacaucgacg gaggagcaag ccagggaaga uucuacaagu ucaucaagcc gauccuggaa 1140
 aagauggacg gaacagaaga acugcugguc aagcugaaca gagaagaccu gcugagaaag 1200
 cagagaacau ucgacaacgg aagcaucccg caccagaucc accugggaga acugcacgca 1260
 auccugagaa gacaggaaga cuucuacccg uuccugaagg acaacagaga aaagaucgaa 1320
 aagauccuga cauucagaau cccguacuac gucggaccgc uggcaagagg aaacagcaga 1380
 uucgcaugga ugacaagaaa gagcgaagaa acauacacac cguggaacuu cgaagaaguc 1440
 gucgacaagg gagcaagcgc acagagcuuc aucgaaagaa ugacaaacuu cgacaagaac 1500
 cugccgaacg aaaagguccu gccgaagcac agccugcugu acgaauacuu cacagucuac 1560
 aacgaacuga caaaggucua guacgucaca gaaggaauga gaaagccggc auuccugagc 1620
 ggagaacaga agaaggcaau cgucgaccug cuguucaaga caaacagaaa ggucacaguc 1680
 aagcagcuga aggaagacua cuucaagaag aucgaaugcu ucgacagcgu cgaaaucagc 1740
 ggagucgaag acagauucaa cgcaagccug ggaacauacc acgaccugcu gaagaucac 1800
 aaggacaagg acuuccugga caacgaagaa aacgaagaca uccuggaaga caucguccug 1860
 acacugacac uguucgaaga cagagaaaug aucgaagaaa gacugaagac auacgcacac 1920
 cuguucgacg acaaggucua gaagcagcug aagagaagaa gauacacagg auggggaaga 1980
 cugagcagaa agcugaucaa cggaaucaga gacaagcaga gcggaaagac aauccuggac 2040
 uuccugaaga gcgacggauu cgcaaacaga aacuucaugc agcugaucca cgacgacagc 2100
 cugacauuca aggaagacau ccagaaggca caggucagcg gacagggaga cagccugcac 2160
 gaacacaucg caaaccuggc aggaagcccg gcaaucaaga agggaauccu gcagacaguc 2220
 aaggucgucg acgaacuggu caaggucaug ggaagacaca agccggaaaa caucgucauc 2280
 gaaauggcaa gagaaaacca gacaacacag aaggagacaga agaacagcag agaaagaag 2340
 aagagaauucg aagaaggaa caaggaaucg ggaagccaga uccugaagga acaccgguc 2400
 gaaaacacac agcugcagaa cgaaaagcug uaccuguacu accugcagaa cggaagagac 2460
 auguacgucg accaggaacu ggacaucaac agacugagcg acuacgacgu cgaccacau 2520
 gucccgacaga gcuuccugaa ggacgacagc aucgacaaca agguccugac aagaagcagc 2580
 aagaacagag gaaagagcga caacgucccg agcgaagaag ucgucaagaa gaugaagaac 2640
 uacuggagac agcugcugaa cgcaaacgug aucacacaga gaaaguucga caaccugaca 2700
 aaggcagaga gaggaggacu gagcgaacug gacaaggcag gauucaucaa gagacagcug 2760
 gucgaaacaa gacagauca aaagcacguc gcacagaucc uggacagcag aaugaacaca 2820
 aaguacgacg aaaacgacaa gcugaucaga gaagucaagg ucaucacacu gaagagcaag 2880
 cuggucagcg acuucagaaa ggacuuccag uucuacaagg ucagagaaau caacaacuac 2940
 caccacgcac acgacgcaua ccugaacgca gucguccgaa cagcacugau caagaaguac 3000
 ccgaagcugg aaagcgaaau cgucuacgga gacuacaagg ucuacgacgu cagaaagaug 3060
 aucgcaaaga gcgaacagga aaucggaaag gcaacagcaa aguacuucuu cuacagcaac 3120
 aucaugaacu ucuucaagac agaaaucaca cuggcaaacg gagaaaucag aaagagaccg 3180
 cugaucgaaa caaacggaga aacaggagaa aucgucuggg acaagggaag agacuucgca 3240
 acagucagaa agguccugag caugccgcag gucaacaucg ucaagaagac agaaguccag 3300
 acaggaggau ucagcaagga aagcauccug ccgaagagaa acagcgacaa gcugaucgca 3360
 agaaagaagg acugggaccc gaagaaguac ggaggauucg acagcccgac agucgcauac 3420
 agcguccugg ucgucgcaaa ggucgaaaag ggaaagagca agaagcugaa gagcgucaag 3480
 gaacugcugg gaauacaa caugggaaaga agcagcuucg aaaagaaccc gaucgacuuc 3540
 cuggaagcaa agggauacaa ggaagucaag aaggaccuga ucaucaagcu gccgaaguac 3600
 agccguucg aacuggaaaa cggaagaaag agaauccugg caagcgcagg agaacugcag 3660
 aagggaacg aacuggcacu gccgagcaag uacgucaacu uccuguaccu ggcaagccac 3720

10

20

30

40

50

uacgaaaagc ugaagggaag cccggaagac aacgaacaga agcagcuguu cgucgaacag 3780
 cacaagcacu accuggacga aaucaucgaa cagaucagcg aaucagcaa gagagucauc 3840
 cuggcagacg caaaccugga caagguccug agcgcauaca acaagcacag agacaagccg 3900
 aucagagaac aggcagaaaa caucauccac cuguucacac ugacaaaccu gggagcaccg 3960
 gcagcauucg aguacuucga cacaacaauc gacagaaaga gauacacaag cacaagggaa 4020
 guccuggacg caacacugau ccaccagagc aucacaggac uguacgaaac aagaaucgac 4080
 cugagccagc ugggaggaga c 4101

<210> 16

<211> 1368

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 16

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val
 1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
 20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
 35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
 50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
 65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
 85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
 100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
 115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
 130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
 145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
 165 170 175

10

20

30

40

50

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225 230 235 240

10

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
275 280 285

20

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
290 295 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305 310 315 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
325 330 335

30

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
340 345 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
355 360 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
370 375 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385 390 395 400

40

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
435 440 445

50

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu

10

20

30

40

50

705 710 715 720
His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
 725 730 735
Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
 740 745 750
Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
 755 760 765
Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
 770 775 780
Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 795 800
Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
 805 810 815
Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
 820 825 830
Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
 835 840 845
Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
 850 855 860
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865 870 875 880
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
 885 890 895
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
 900 905 910
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
 915 920 925
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
 930 935 940
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945 950 955 960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
 965 970 975

10

20

30

40

50

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
1085 1090 1095

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys
1100 1105 1110

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro
1115 1120 1125

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val
1130 1135 1140

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys
1145 1150 1155

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser
1160 1165 1170

Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys
1175 1180 1185

Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu
1190 1195 1200

Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly
1205 1210 1215

Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val
1220 1225 1230

10

20

30

40

50

Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser
 1235 1240 1245

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys
 1250 1255 1260

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys
 1265 1270 1275

Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala
 1280 1285 1290

Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn
 1295 1300 1305

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala
 1310 1315 1320

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
 1325 1330 1335

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
 1340 1345 1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
 1355 1360 1365

<210> 17

<211> 4107

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 17

auggacaaga aguacagcau cggacuggca aucggaacaa acagcgucgg augggcaguc 60
 aucacagacg aaaucaaggu cccgagcaag aaguucaagg uccugggaaa cacagacaga 120
 cacagcauca agaagaaccu gaucggagca cugcuguucg acagcggaga aacagcagaa 180
 gcaacaagac ugaagagaac agcaagaaga agauacacaa gaagaaagaa cagaaucugc 240
 uaccugcagg aaaucuucag caacgaaaug gcaaaggucg acgacagcuu cuuccacaga 300
 cuggaagaaa gcuuccuggu cgaagaagac aagaagcacg aaagacaccc gaucuucgga 360
 aaucugucg acgaagucgc auaccacgaa aaguaccgca caaucuacca ccugagaaag 420
 aagcuggucg acagcacaga caaggcagac cugagacuga ucuaccuggc acuggcacac 480
 augaucaagu ucagaggaca cuuccugauc gaaggagacc ugaacccgga caacagcgac 540
 gucgacaagc uguucaucca gcugguccag acauacaacc agcuguucga agaaaacccg 600
 aucaacgcaa gcgagagucg cgcgaaggca auccugagcg caagacugag caagagcaga 660
 agacuggaaa accugaucgc acagcugccg ggagaaaaga agaacggacu guucggaaac 720
 cugaucgcac ugagccuggg acugacaccg aacuucaaga gcaacuucga ccuggcagaa 780

10

20

30

40

50

gacgcaaagc ugcagcugag caaggacaca uacgacgacg accuggacaa ccugcuggca 840
 cagaucggag accaguacgc agaccuguuc cuggcagcaa agaaccugag cgacgcaauc 900
 cugcugagcg acauccugag agucaacaca gaaaucacaa aggcaccgcu gagcgcaagc 960
 augaucaaga gauacgacga acaccaccag gaccugacac ugcugaaggc acuggucaga 1020
 cagcagcugc cggaagaauc caaggaaauc uucucgacc agagcaagaa cggauacgca 1080
 ggauacaucg acggaggagc aagccaggaa gaaucuaaca aguucacaa gccgaucug 1140
 gaaaagaugg acggaacaga agaacugcug gucaagcuga acagagaaga ccugcugaga 1200
 aagcagagaa cauucgacaa cggaagcauc ccgcaccaga uccaccuggg agaacugcac 1260
 gcaauccuga gaagacagga agacuucuaac ccguuccuga aggacaacag agaaaagauc 1320
 gaaaagauc ugcacauucag aaucuccuac uacgucggac cgcuggcaag aggaaacagc 1380
 agauucgcau ggaugacaag aaagagcgaa gaaacaauca caccguggaa cuucgaagaa 1440
 gucugcgaca agggagcaag cgcacagagc uucaucgaaa gaaugacaaa cuucgacaag 1500
 aaccugccga acgaaaaggu ccugccgaag cacagccugc uguacgaaua cuucacaguc 1560
 uacaacgaac ugacaaaaggu caaguacguc acagaaggaa ugagaaagcc ggcauuccug 1620
 agcggagaa acagaagggc aaucgucgac cugcuguuca agacaaacag aaaggucaca 1680
 gucaagcagc ugaaggaaga cuacuucuaag aagaucgaau gcuucgacag cgucgaaauc 1740
 agcggagucg aagacagauu caacgcaagc cugggaacau accacgaccu gcugaagauc 1800
 aucaaggaca aggacuuccu ggacaacgaa gaaaacgaag acauccugga agacaucguc 1860
 cugacacuga cacuguucga agacagagaa augaucgaag aaagacugaa gacauacgca 1920
 caccuguucg acgacaaggu caugaagcag cugaagagaa gaagauacac aggaugggga 1980
 agacugagca gaaagcugau caacggaauc agagacaagc agagcggaaa gacaauccug 2040
 gacuuccuga agagcgacgg auucgcaaac agaaacuua ugcagcugau ccacgacgac 2100
 agccugacau ucaaggaaga cauccagaag gcacagguca gcggacaggg agacagccug 2160
 cacgaacaca ucgcaaaccu ggcaggaagc ccggcaauca agaagggaau ccugcagaca 2220
 gucaaggucg ucgacgaacu ggucaagguc augggaagac acaagccgga aaacaucguc 2280
 aucgaaaugg caagagaaaa ccagacaaca cagaagggac agaagaacag cagagaaaga 2340
 augaagagaa ucgaagaagg aaucaggaa cugggaagcc agauccugaa ggaacaccg 2400
 gucgaaaaaca cacagcugca gaacgaaaag cuguaccugu acuaccugca gaacggaaga 2460
 gacauguacg ucgaccagga acuggacauc aacagacuga gcgacuacga cgucgaccac 2520
 aucguccgc agagcuuccu gaaggacgac agcaucgaca acaagguccu gacaagaagc 2580
 gacaagaaca gaggaaagag cgacaacguc ccgagcgaag aagucgucaa gaagaugaag 2640
 aacucugga gacagcugcu gaacgcaaag cugaucacac agagaaaguu cgacaaccug 2700
 acaaaggcag agagaggagg acugagcgaa cuggacaagg caggauucau caagagacag 2760
 cuggucgaaa caagacagau cacaagcac gucgcacaga uccuggacag cagaaugaac 2820
 acaaaguacg acgaaaacga caagcugauc agagaaguca aggucaucac acugaagagc 2880
 aagcugguca gcgacuucag aaaggacuuc caguucuaa agguacagaga aaucacaac 2940
 uaccaccagc cacacgacgc auaccugaac gcagucgucg gaacagcacu gaucaagaag 3000
 uaccgaagc uggaagcga auucgucuaac ggagacuaca aggucuaacga cgucagaaag 3060
 augaucgcaa agagcgaaca ggaaucgga aaggcaacag caaaguacuu cuucuaacagc 3120
 acaucauga acuucuaa gacagaaauc acacuggcaa acggagaaaau cagaaagaga 3180
 ccgcugaucg aaacaaacgg agaaacagga gaaucgucu gggacaaggg aagagacuuc 3240
 gcaacaguca gaaagguccu gagcaugccg cagguaaca ucgucaagaa gacagaaguc 3300
 cagacaggag gauucagcaa ggaaagcauc cugccgaaga gaaacagcga caagcugauc 3360
 gcaagaaaga aggacuggga cccgaagaag uacggaggau ucgacagccc gacagucgca 3420
 uacagcgucc ugguucgucg aaaggucgaa aagggaaga gcaagaagcu gaagagcguc 3480
 aaggaaucg uggaauacac aaucaggaa agaagcagcu ucgaaaagaa cccgaucgac 3540
 uuccuggaag caaagggaau caaggaaaguc aagaaggacc ugaucacaa gcugccgaag 3600
 uacagccugu ucgaacugga aaacggaaga aagagaaguc uggcaagcg aggagaacug 3660
 cagaaggga acgaacuggc acugccgagc aaguacguca acuuccugua ccuggcaagc 3720
 cacuacgaaa agcugaaggg aagcccggaa gacaacgaac agaagcagcu guucgucgaa 3780

10

20

30

40

50

cagcacaagc acuaccugga cgaaaaucauc gaacagauca gcgaauucag caagagaguc 3840
 auccuggcag acgcaaaccu ggacaagguc cugagcgcau acaacaagca cagagacaag 3900
 ccgaucagag aacaggcaga aaacaucauc caccuguuca cacugacaaa ccugggagca 3960
 ccggcagcau ucaaguacuu cgacacaaca aucgacagaa agagauacac aagcacaaag 4020
 gaaguccugg acgcaacacu gauccaccag agcaucacag gacuguacga aacaagaau 4080
 gaccugagcc agcugggagg agacuag 4107

<210> 18

<211> 4101

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 18

gacaagaagu acagcaucgg acuggcaauc ggaacaaaca gcgucggaug ggcagucauc 60
 acagacgaau acaagguccc gagcaagaag uucaaggucc ugggaaacac agacagacac 120
 agcaucaaga agaaccugau cggagcacug cuguucgaca gcggagaaac agcagaagca 180
 acaagacuga agagaacagc aagaagaaga uacacaagaa gaaagaacag aaucugcuac 240
 cugcaggaaa ucuucagcaa cgaaauggca aaggucgacg acagcuucuu ccacagacug 300
 gaagaaagcu uccuggucga agaagacaag aagcacgaaa gacacccgau cuucggaaac 360
 aucgucgacg aagucgcaua ccacgaaaag uacccgacaa ucuaccaccu gagaaagaag 420
 cuggucgaca gcacagacaa ggacagaccug agacugaucu accuggcacu ggacacaug 480
 aucaaguuca gaggacacuu ccugaucgaa ggagaccuga acccgacaa cagcgacguc 540
 gacaagcugu ucauccagcu gguccagaca uacaaccagc uguucgaaga aaacccgauc 600
 aacgcaagcg gagucgacgc aaaggcaauc cugagcgcaa gacugagcaa gagcagaaga 660
 cuggaaaacc ugaucgcaca gcugccggga gaaaagaaga acggacuguu cggaaccug 720
 aucgcacuga gccugggacu gacaccgaac uucaagagca acuucgaccu ggcagaagac 780
 gcaaagcugc agcugagcaa ggacacauac gacgacgacc uggacaaccu gcuggcacag 840
 aucggagacc aguacgcaga ccuguuccug gcagcaaaga accugagcga cgcaauccug 900
 cugagcgaca uccugagagu caacacagaa aucacaaagg caccgcugag cgcaagcaug 960
 aucaagagau acgacgaaca ccaccaggac cugacacugc ugaaggcacu ggucagacag 1020
 cagcugccgg aaaaguacaa ggaaaucuu uucgaccaga gcaagaacgg auacgcagga 1080
 ucaucgacg gaggagcaag ccaggaagaa uucaucaagu ucaucaagcc gaucgugaa 1140
 aagauggacg gaacagaaga acugcugguc aagcugaaca gagaagaccu gcugagaaag 1200
 cagagaacau ugcacaacgg aagcaucccg caccagaucc accugggaga acugcacgca 1260
 auccugagaa gacaggaaga cuucuacccg uuccugaagg acaacagaga aaagaucgaa 1320
 aagauccuga cauucagaau cccguacuac gucggaccgc uggcaagagg aaacagcaga 1380
 uucgcaugga ugacaagaaa gagcgaagaa acaaucacac cguggaacuu cgaagaaguc 1440
 gucgacaagg gagcaagcgc acagagcuuc aucgaaagaa ugacaaacuu cgacaagaac 1500
 cugccgaacg aaaagguccu gccgaagcac agccugcugu acgaauacuu cacagucuac 1560
 aacgaacuga caaaggucua guacgucaca gaaggaauga gaaagccggc auuccugagc 1620
 ggagaacaga agaaggcaau cgucgaccug cuguucaaga caaacagaaa ggucacaguc 1680
 aagcagcuga aggaagacua cuucaagaag aucgaaugcu ucgacagcgu cgaaaucagc 1740
 ggagucgaag acagauucaa cgcaagccug ggaacauacc acgaccugcu gaagaaucauc 1800
 aaggacaagg acuuccugga caacgaagaa aacgaagaca uccuggaaga caucguccug 1860
 acacugacac uguucgaaga cagagaaaug aucgaagaaa gacugaagac auacgcacac 1920
 cuguucgacg acaaggucuu gaagcagcug aagagaagaa gauacacagg auggggaaga 1980
 cugagcagaa agcugaucaa cggaaucaaga gacaagcaga gcggaaagac aauccuggac 2040

10

20

30

40

50

uuccugaaga gcgacggauu cgcaaacaga aacuucaugc agcugaucca cgacgacagc 2100
 cugacauuca aggaagacau ccagaaggca caggucagcg gacagggaga cagccugcac 2160
 gaacacaucg caaaccuggc aggaagcccg gcaaucaaga agggaauccu gcagacaguc 2220
 aaggucgucg acgaacuggu caaggucaug ggaagacaca agccggaaaa caucgucauc 2280
 gaaauggcaa gagaaaacca gacaacacag aagggacaga agaacagcag agaaagaug 2340
 aagagaaucg aagaaggauu caaggaacug ggaagccaga uccugaagga acacccgguc 2400
 gaaaacacac agcugcagaa cgaaaagcug uaccuguacu accugcagaa cggaagagac 2460
 auguacgucg accaggaacu ggacaucaac agacugagcg acuacgacgu cgaccacauc 2520
 gucccgagcaga gcuuccugaa ggacgacagc aucgacaaca agguccugac aagaagcgac 2580
 aagaacagag gaaagagcga caacgucccg agcgaagaag ucgucaagaa gaugaagaac 2640
 uacuggagac agcugcugaa cgaaaagcug aucacacaga gaaaguucga caaccugaca 2700
 aaggcagaga gaggaggacu gagcgaacug gacaaggcag gauucaucaa gagacagcug 2760
 gucgaaacaa gacagaucaac aaagcacguc gcacagaucc uggacagcag aaugaacaca 2820
 aaguacgacg aaaacgacaa gcugaucaga gaagucaagg ucaucacacu gaagagcaag 2880
 cuggucagcg acuucagaaa ggacuuccag uucuaacaagg ucagagaaaau caacaacuac 2940
 caccacgcac acgacgcuaa ccugaacgca gucgucggaa cagcacugau caagaaguac 3000
 ccgaagcugg aaagcgaaau cgucuacgga gacuacaagg ucuacgacgu cagaaagaug 3060
 aucgcaaaga gcgaacagga aaucggaaaag gcaacagcaa aguacuucuu cuacagcaac 3120
 aucaugaacu ucuucaagac agaaaucaca cuggcaaacg gagaaaucag aaagagaccg 3180
 cugaucgaaa caaacggaga aacaggagaa aucgucuggg acaagggaag agacuucgca 3240
 acagucagaa agguccugag caugccgcag gucaacaucg ucaagaagac agaaguccag 3300
 acaggaggau ucagcaagga aagcauccug ccgaagagaa acagcgacaa gcugaucgca 3360
 agaaagaagg acugggaccc gaagaaguac ggaggauucg acagcccgcac agucgcuauc 3420
 agcguccugg ucgucgcaaa ggucgaaaag ggaaagagca agaagcugaa gagcgucaag 3480
 gaacugcugg gaaucaacaau cauggaaaaga agcagcuucg aaaagaaccg gaucgacuuc 3540
 cuggaagcaa agggauacaa ggaagucaag aaggaccuga ucaucaagcu gccgaaguac 3600
 agccuguucg aacuggaaaa cggaagaaaag agaaugcugg caagcgacagg agaacugcag 3660
 aagggaacg aacuggcacu gccgagcaag uacgucaacu uccuguaccu ggcaagccac 3720
 uacgaaaagc ugaagggaag cccggaagac aacgaacaga agcagcuguu cgucgaacag 3780
 cacaagcacu accuggacga aaucaucgaa cagaucagcg aaucagcaa gagagucauc 3840
 cuggcagacg caaaccugga caagguccug agcgcauaca acaagcacag agacaagccg 3900
 aucagagaac aggcagaaaa caucauccac cuguucacac ugacaaaccu gggagcaccg 3960
 gcagcauuca aguacuucga cacaacaauc gacagaaaga gauacacaag cacaaggaa 4020
 guccuggacg caacacugau ccaccagagc aucacaggac uguacgaaac aagaaucgac 4080
 cugagccagc ugggaggaga c 4101

<210> 19

<211> 1368

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 19

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30

10

20

30

40

50

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp

10

20

30

40

50

290 295 300
 Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
 305 310 315 320
 Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
 325 330 335
 Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
 340 345 350
 Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
 355 360 365
 Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
 370 375 380
 Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
 385 390 395 400
 Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
 405 410 415
 Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
 420 425 430
 Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
 435 440 445
 Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
 450 455 460
 Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
 465 470 475 480
 Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
 485 490 495
 Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
 500 505 510
 Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
 515 520 525
 Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
 530 535 540
 Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
 545 550 555 560

10

20

30

40

50

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

10

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670

20

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720

30

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
770 775 780

40

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830

50

Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys
 835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
 850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
 865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
 885 890 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
 900 905 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
 915 920 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
 930 935 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
 945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
 965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
 980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
 995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
 1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
 1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
 1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
 1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
 1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr

10

20

30

40

50

1085	1090	1095
Glu Val Gln Thr Gly	Gly Phe Ser Lys	Glu Ser Ile Leu Pro Lys
1100	1105	1110
Arg Asn Ser Asp Lys	Leu Ile Ala Arg	Lys Lys Asp Trp Asp Pro
1115	1120	1125
Lys Lys Tyr Gly Gly	Phe Asp Ser Pro	Thr Val Ala Tyr Ser Val
1130	1135	1140
Leu Val Val Ala Lys	Val Glu Lys Gly	Lys Ser Lys Lys Leu Lys
1145	1150	1155
Ser Val Lys Glu Leu	Leu Gly Ile Thr	Ile Met Glu Arg Ser Ser
1160	1165	1170
Phe Glu Lys Asn Pro	Ile Asp Phe Leu	Glu Ala Lys Gly Tyr Lys
1175	1180	1185
Glu Val Lys Lys Asp	Leu Ile Ile Lys	Leu Pro Lys Tyr Ser Leu
1190	1195	1200
Phe Glu Leu Glu Asn	Gly Arg Lys Arg	Met Leu Ala Ser Ala Gly
1205	1210	1215
Glu Leu Gln Lys Gly	Asn Glu Leu Ala	Leu Pro Ser Lys Tyr Val
1220	1225	1230
Asn Phe Leu Tyr Leu	Ala Ser His Tyr	Glu Lys Leu Lys Gly Ser
1235	1240	1245
Pro Glu Asp Asn Glu	Gln Lys Gln Leu	Phe Val Glu Gln His Lys
1250	1255	1260
His Tyr Leu Asp Glu	Ile Ile Glu Gln	Ile Ser Glu Phe Ser Lys
1265	1270	1275
Arg Val Ile Leu Ala	Asp Ala Asn Leu	Asp Lys Val Leu Ser Ala
1280	1285	1290
Tyr Asn Lys His Arg	Asp Lys Pro Ile	Arg Glu Gln Ala Glu Asn
1295	1300	1305
Ile Ile His Leu Phe	Thr Leu Thr Asn	Leu Gly Ala Pro Ala Ala
1310	1315	1320
Phe Lys Tyr Phe Asp	Thr Thr Ile Asp	Arg Lys Arg Tyr Thr Ser
1325	1330	1335

10

20

30

40

50

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
 1340 1345 1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
 1355 1360 1365

<210> 20

<211> 4107

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 20

```

auggacaaga aguacagcau cggacuggca aucggaacaa acagcgucgg augggcaguc      60
aucacagacg aaaucaaggu cccgagcaag aaguucaagg uccugggaaa cacagacaga      120
cacagcauca agaagaaccu gaucggagca cugcuguucg acagcggaga aacagcagaa      180
gcaacaagac ugaagagaac agcaagaaga agauacacaa gaagaaagaa cagaaucugc      240
uaccugcagg aaaucuucag caacgaaaug gcaaaggucg acgacagcuu cuuccacaga      300
cuggaagaaa gcuuccuggu cgaagaagac aagaagcacg aaagacaccc gaucuucgga      360
aauaucgucg acgaagucgc auaccacgaa aaguacccga caaucuacca ccugagaaag      420
aagcuggucg acagcacaga caaggcagac cugagacuga ucuaccuggc acuggcacac      480
augaucaagu ucagaggaca cuuccugauc gaaggagacc ugaacccgga caacagcgac      540
gucgacaagc uguucaucca gcugguccag acauacaacc agcuguucga agaaaacccg      600
aucaacgcaa gcgagagucg cgcgaaggca auccugagcg caagacugag caagagcaga      660
agacuggaaa accugaucgc acagcugccg ggagaaaaga agaacggacu guucggaaac      720
cugaucgcac ugagccuggg acugacaccg aacuucaaga gcaacuucga ccuggcagaa      780
gacgcaaagc ugcagcugag caaggacaca uacgacgacg accuggacaa ccugcuggca      840
cagaucggag accaguacgc agaccuguuc cuggcagcaa agaaccugag cgacgcaauc      900
cugcugagcg acauccugag agucaacaca gaaauacaaa aggaccgcu gagcgcaagc      960
augaucaaga gauacgacga acaccaccag gaccugacac ugcugaaggc acuggucaga     1020
cagcagcugc cggaaaagua caaggaaauc uucuucgacc agagcaagaa cggauacgca     1080
ggauacaucg acggaggagc aagccaggaa gaaucuaca aguucauca gccgauccug     1140
gaaaagaugg acggaacaga agaacugcug gucaagcuga acagagaaga ccugcugaga     1200
aagcagagaa cauucgacaa cggaagcauc ccgcaccaga uccaccuggg agaacugcac     1260
gcaauccuga gaagacagga agacuucuac ccguuccuga aggacaacag agaaaagauc     1320
gaaaagaucc ugacauucag aaucccguac uacgucggac cgcuggcaag aggaaacagc     1380
agauucgcau ggaugacaag aaagagcgaa gaaacaauca caccugggaa cuucgaagaa     1440
gucgucgaca agggagcaag cgcacagagc uucaucgaaa gaaugacaaa cuucgacaag     1500
aaccugccga acgaaaaggu ccugccgaag cacagccugc uguacgaaua cuucacaguc     1560
uacaacgaac ugacaaaaggu caaguacguc acagaaggaa ugagaaagcc ggcauuccug     1620
agcggagaaac agaagaaggc aaucgucgac cugcuguuca agacaaacag aaaggucaca     1680
gucaagcagc ugaaggaaga cuacuucaag aagaucgaau gcuucgacag cgucgaaauc     1740
agcggagucg aagacagauu caacgcaagc cugggaacau accacgaccu gcugaagauc     1800
aucaaggaca aggacuuccu ggacaacgaa gaaaacgaag acauccugga agacaucguc     1860
cugacacuga cacuguucga agacagagaa augaucgaag aaagacugaa gacauacgca     1920
caccuguucg acgacaaggu caugaagcag cugaagagaa gaagauacac aggaugggga     1980
agacugagca gaaagcugau caacggaauc agagacaagc agagcggaaa gacaauccug     2040
gacuuccuga agagcgacgg auucgcaaac agaaacuuca ugcagcugau ccacgacgac     2100

```

10

20

30

40

50

agccugacau ucaaggaaga cauccagaag gcacagguca gcggacaggg agacagccug 2160
 cacgaacaca ucgcaaaccu ggcaggaagc ccggcaauca agaagggaau ccugcagaca 2220
 gucaaggucg ucgacgaacu ggucaagguc augggaagac acaagccgga aaacaucguc 2280
 aucgaaaugg caagagaaaa ccagacaaca cagaagggac agaagaacag cagagaaaaga 2340
 augaagagaa ucgaagaagg aaucaggaa cugggaagcc agauccugaa ggaacacccg 2400
 gucgaaaaca cacagcugca gaacgaaaag cuguaccugu acuaccugca gaacggaaga 2460
 gacauguacg ucgaccagga acuggacauc aacagacuga gcgacuacga cgucgacgca 2520
 aucgucccg agagcuuccu gaaggacgac agcaucgaca acaagguccu gacaagaagc 2580
 gacaagaaca gaggaagag cgacaacguc ccgagcgaag aagucgucaa gaagaugaag 2640
 aacuacugga gacagcugcu gaacgcaaag cugaucacac agagaaaguu cgacaaccug 2700
 acaaaggcag agagaggagg acugagcga cuggacaagg caggauucau caagagacag 2760
 cuggucgaaa caagacagau cacaagcac gucgcacaga uccuggacag cagaaugaac 2820
 acaaaguacg acgaaaacga caagcugauc agagaaguca aggucaucac acugaagagc 2880
 aagcugguca gcgacuucag aaaggacuuc caguucuaa agguacagaga aaucacaac 2940
 uaccaccag cacacgacgc auaccugaac gcagucgucg gaacagcacu gaucaagaag 3000
 uacccgaagc uggaagcga auucgucuac ggagacuaca aggucuacga cgucagaaag 3060
 augaucgcaa agagcgaaca ggaaucgga aaggcaacag caaaguacuu cuucuacagc 3120
 aacauauga acuucuuaa gacagaauc acacuggcaa acggagaaau cagaaagaga 3180
 ccgcugaucg aaacaaacgg agaaacagga gaaucgucu gggacaaggg aagagacuuc 3240
 gcaacaguca gaaagguccu gagcaugccg caggucaaca ucgucaaga gacagaaguc 3300
 cagacaggag gauucagcaa ggaaagcauc cugccgaaga gaaacagcga caagcugauc 3360
 gcaagaaaga aggacuggga cccgaagaag uacggaggau ucgacagccc gacagucgca 3420
 uacagcgucc ugguucgucg aaaggucgaa aagggaaga gcaagaagcu gaagagcguc 3480
 aaggaaucg uggaauac aaucaggaa agaagcagcu ucgaaaaga cccgaucgac 3540
 uuccuggaag caaagggaau caaggaaguc aagaaggacc ugaucauaa gcugccgaag 3600
 uacagccugu ucgaacugga aaacggaaga aagagaagc uggaagcgc aggagaacug 3660
 cagaaggga acgaacuggc acugccgagc aaguacguca acuuccugua ccuggcaagc 3720
 cacuacgaaa agcugaaggg aagcccggaa gacaacgaac agaagcagcu guucgucgaa 3780
 cagcacaagc acuaccugga cgaaaucauc gaacagauca gcgaauucag caagagaguc 3840
 auccuggcag acgcaaaccu ggacaagguc cugagcgcau acaacaagca cagagacaag 3900
 ccgaucagag aacaggcaga aaacaucauc caccuguuca cacugacaaa ccugggagca 3960
 ccggcagcau ucaaguacuu cgacacaaca aucgacagaa agagauacac aagcacaag 4020
 gaaguccugg acgcaacacu gaucaccag agcaucacag gacuguaacga aacaagauc 4080
 gaccugagcc agcugggagg agacuag 4107

<210> 21

<211> 4113

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 21

gacaagaagu acagcaucgg acuggcaauc ggaacaaaca gcgucggaug ggcagucauc 60
 acagacgaau acaaggucc gagcaagaag uucaaggucc uggaacac agacagacac 120
 agcaucaaga agaaccugau cggagcacug cuguucgaca gcggagaaac agcagaagca 180
 acaagacuga agagaacagc aagaagaaga uacacaaga gaaagaacag aaucugcuac 240
 cugcaggaaa ucuucagcaa cgaauggca aaggucgacg acagcuucuu ccacagacug 300
 gaagaaagcu uccuggucga agaagacaag aagcacgaaa gacacccgau cuucggaaac 360

aucgucgacg aagucgcaua ccacgaaaag uacccgacaa ucuaccaccu gagaaagaag 420
 cuggucgaca gcacagacaa ggacagaccug agacugaucu accuggcacu ggcacacaug 480
 aucaaguuca gaggacacuu ccugaucgaa ggagaccuga acccgacaa cagcgacguc 540
 gacaagcugu ucauccagcu gguccagaca uacaaccagc uguucgaaga aaacccgauc 600
 aacgcaagcg gagucgacgc aaaggcaauc cugagcgcaa gacugagcaa gagcagaaga 660
 cuggaaaacc ugaucgcaca gcugccggga gaaaagaaga acggacuguu cggaaccug 720
 aucgcacuga gccugggacu gacaccgaac uucaagagca acuucgaccu ggcagaagac 780
 gcaaagcugc agcugagcaa ggacacauac gacgacgacc uggacaaccu gcuggcacag 840
 aucggagacc aguacgcaga ccuguuccug gcagcaaaga accugagcga cgcaauccug 900
 cugagcgaca uccugagagu caacacagaa aucacaaagg caccgcugag cgcaagcaug 960 10
 aucaagagau acgacgaaca ccaccaggac cugacacugc ugaaggcacu ggucagacag 1020
 cagcugccgg aaaaguacaa ggaaaucuuc uucgaccaga gcaagaacgg auacgcagga 1080
 ucaucgacg gaggagcaag ccaggaagaa uucuacaagu ucaucaagcc gauccuggaa 1140
 aagauggacg gaacagaaga acugcugguc aagcugaaca gagaagaccu gcugagaaag 1200
 cagagaacau ucgacaacgg aagcaucccg caccagaucc accugggaga acugcacgca 1260
 auccugagaa gacaggaaga cuucuacccg uuccugaagg acaacagaga aaagaucgaa 1320
 aagauccuga cauucagaau cccguacuac gucggaccgc uggcaagagg aaacagcaga 1380
 uucgcaugga ugacaagaaa gagcgaagaa acaauacac cguggaacuu cgaagaaguc 1440
 gucgacaagg gagcaagcgc acagagcuuc aucgaaagaa ugacaaacuu cgacaagaac 1500
 cugccgaacg aaaagguccu gccgaagcac agccugcugu acgaauacuu cacagucuac 1560 20
 aacgaacuga caaagggucaa guacgucaca gaaggaauga gaaagccggc auuccugagc 1620
 ggagaacaga agaaggcaau cgucgaccug cuguucaaga caaacagaaa ggucacaguc 1680
 aagcagcuga aggaagacua cuucaagaag aucgaaugcu ucgacagcgu cgaaaucagc 1740
 ggagucgaag acagauucaa cgcaagccug ggaacauacc acgaccugcu gaagaucac 1800
 aaggacaagg acuuccugga caacgaagaa aacgaagaca uccuggaaga caucguccug 1860
 acacugacac uguucgaaga cagagaaaug aucgaagaaa gacugaagac auacgcacac 1920
 cuguucgacg acaaggucuu gaagcagcug aagagaagaa gauacacagg auggggaaga 1980
 cugagcagaa agcugaucaa cggaaucaaga gacaagcaga gcggaaagac aauccuggac 2040
 uuccugaaga gcgacggauu cgaaacaga aacuucagc agcugaucca cgacgacagc 2100
 cugacauuca aggaagacau ccagaaggca caggucagcg gacagggaga cagccugcac 2160 30
 gaacacaucg caaacccuggc aggaagcccg gcaaucaaga agggaauccu gcagacaguc 2220
 aaggucgucg acgaacuggu caaggucaug ggaagacaca agccggaaaa caucgucauc 2280
 gaaauggcaa gagaaaacca gacaacacag aagggacaga agaacagcag agaaagaau 2340
 aagagaau 2400
 aagaaggaau caaggaacug ggaagccaga uccugaagga acaccgguc 2460
 gaaaacacac agcugcagaa cgaaaagcug uaccuguacu accugcagaa cggaagagac 2520
 auguacgucg accaggaacu ggacaucaac agacugagcg acuacgacgu cgacgcauc 2580
 gucccgcaga gcuuccugaa ggacgacagc aucgacaaca agguccugac aagaagcgac 2640
 aagaacagag gaaagagcga caacgucccg agcgaagaag ucgucaagaa gaugaagaac 2700
 uacuggagac agcugcugaa cgaaagcug aucacacaga gaaaguucga caaccugaca 2760 40
 aaggcagaga gaggaggacu gagcgaacug gacaaggcag gauucaucaa gagacagcug 2820
 gucgaaacaa gacagaucaac aaagcacguc gcacagaucc uggacagcag aaugaacaca 2880
 aaguacgacg aaaacgacaa gcugaucaga gaagucaagg ucaucacacu gaagagcaag 2940
 cuggucagcg acuucagaaa ggacuuccag uucuacaagg ucagagaaau caacaacuac 3000
 caccacgcac acgacgcaua ccugaacgca gucguccgaa cagcacugau caagaaguac 3060
 ccgaagcugg aaagcgaaau cgucuacgga gacuacaagg ucuacgacgu cagaaagaug 3120
 aucgcaaaga gcgaacagga aaucggaaaag gcaacagcaa aguacuucuu cuacagcaac 3180
 aucaugaacu ucuucaagac agaaaucaac cuggcaaacg gagaaaucaag aaagagaccg 3240
 cugaucgaaa caaacggaga aacaggagaa aucgucuggg acaagggaag agacuucgca 3300
 acagucagaa agguccugag caugccgcag gucaacaucg ucaagaagac agaaguccag 3360 50
 acaggaggau ucagcaagga aagcauccug ccgaagagaa acagcgacaa gcugaucgca

agaaagaagg acugggaccc gaagaaguac ggaggauucg acagcccgac agucgcauac 3420
 agcguccugg ucgucgcaaa ggucgaaaag ggaaagagca agaagcugaa gagcguaag 3480
 gaacugcugg gaauacaaau cauggaaaga agcagcuucg aaaagaaccc gaucgacuuc 3540
 cuggaagcaa agggauacaa ggaagucaag aaggaccuga ucaucaagcu gccgaaguac 3600
 agccuguucg aacuggaaaa cggaagaaag agaauvcugg caagcgcagg agaacugcag 3660
 aagggaacg aacuggcacu gccgagcaag uacgucaacu uccuguaccu ggcaagccac 3720
 uacgaaaagc ugaagggaag cccggaagac aacgaacaga agcagcuguu cgucgaacag 3780
 cacaagcacu accuggacga aaucaucgaa cagaucagcg aaucagcaa gagagucauc 3840
 cuggcagacg caaacugga caagguccug agcgcauaca acaagcacag agacaagccg 3900
 aucagagaac aggcagaaaa caucauccac cuguucacac ugacaaaccu gggagcaccg 3960
 gcagcauua aguacuucga cacaacaauc gacagaaaga gauacacaag cacaagggaa 4020
 guccuggacg caacacugau ccaccagagc aucacaggac uguacgaaac aagaaucgac 4080
 cugagccagc ugaggaggaga cggaggagga agc 4113

10

<210> 22

<211> 1392

<212> PRT

<213> Artificial Sequence

<220>

<221> source

20

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 22

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile

35 40 45

30

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu

50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys

65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser

85 90 95

40

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys

100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr

115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp

130 135 140

50

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
195 200 205

10

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
245 250 255

20

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
290 295 300

30

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305 310 315 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
325 330 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
340 345 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
355 360 365

40

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
370 375 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385 390 395 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415

50

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
435 440 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe

10

20

30

40

50

675	680	685	
Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe			
690	695	700	
Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu			
705	710	715	720
His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly			
	725	730	735
Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly			
	740	745	750
Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln			
	755	760	765
Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile			
	770	775	780
Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro			
785	790	795	800
Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu			
	805	810	815
Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg			
	820	825	830
Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys			
	835	840	845
Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg			
	850	855	860
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys			
865	870	875	880
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys			
	885	890	895
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp			
	900	905	910
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr			
	915	920	925
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp			
	930	935	940

10

20

30

40

50

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
 945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
 965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
 980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
 995 1000 1005

10

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
 1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
 1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
 1040 1045 1050

20

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
 1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
 1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
 1085 1090 1095

30

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys
 1100 1105 1110

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro
 1115 1120 1125

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val
 1130 1135 1140

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys
 1145 1150 1155

40

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser
 1160 1165 1170

Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys
 1175 1180 1185

Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu
 1190 1195 1200

50

Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly
 1205 1210 1215

Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val
 1220 1225 1230

Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser
 1235 1240 1245

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys
 1250 1255 1260

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys
 1265 1270 1275

Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala
 1280 1285 1290

Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn
 1295 1300 1305

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala
 1310 1315 1320

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
 1325 1330 1335

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
 1340 1345 1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
 1355 1360 1365

Gly Ser Gly Ser Pro Lys Lys Lys Arg Lys Val Asp Gly Ser Pro
 1370 1375 1380

Lys Lys Lys Arg Lys Val Asp Ser Gly
 1385 1390

<210> 23

<211> 4179

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 23

auggacaaga aguacagcau cggacuggac aucggaacaa acagcgucgg augggcaguc

60

10

20

30

40

50

aucacagacg aaauacaaggu cccgagcaag aaguucaagg uccugggaaa cacagacaga 120
 cacagcauca agaagaaccu gaucggagca cugcuguucg acagcggaga aacagcagaa 180
 gcaacaagac ugaagagaac agcaagaaga agauacacaa gaagaaagaa cagaauucg 240
 uaccugcagg aaaucuucag caacgaaaug gcaaaaggucg acgacagcuu cuuccacaga 300
 cuggaagaaa gcuuccuggu cgaagaagac aagaagcacg aaagacaccc gaucuucgga 360
 acaucgucg acgaagucgc auaccacgaa aaguacccga caaucuacca ccugagaaag 420
 aagcuggucg acagcacaga caaggcagac cugagacuga ucuaccuggc acuggcacac 480
 augaucaagu ucagaggaca cuuccugauc gaaggagacc ugaacccgga caacagcgac 540
 gucgacaagc uguucaucca gcugguccag acauacaacc agcuguucga agaaaacccg 600
 aucaacgcaa gcgagagucga cgcaaaggca auccugagcg caagacugag caagagcaga 660
 agacuggaaa accugaucgc acagcugccg ggagaaaaga agaacggacu guucggaaac 720
 cugaucgcac ugagccuggg acugacaccg aacuuaaga gcaacuucga ccuggcagaa 780
 gacgcaaagc ugcagcugag caaggacaca uacgacgacg accuggacaa ccugcuggca 840
 cagaucggag accaguacgc agaccuguuc cuggcagcaa agaaccugag cgacgcaauc 900
 cugcugagcg acauccugag agucaacaca gaaauacaaa aggcaccgcu gagcgcaagc 960
 augaucaaga gauacgacga acaccaccag gaccugacac ugcugaaggc acuggucaga 1020
 cagcagcugc cgaaaaagua caaggaaauc uucuucgacc agagcaagaa cggaucgca 1080
 ggauacaucg acggaggagc aagccaggaa gaauucuaca aguucaucaa gccgauccug 1140
 gaaaagaugg acggaacaga agaacugcug gucaagcuga acagagaaga ccugcugaga 1200
 aagcagagaa cauucgacaa cggaagcauc ccgcaccaga uccaccuggg agaacugcac 1260
 gcaauccuga gaagacagga agacuucuac ccguuccuga aggacaacag agaaaagauc 1320
 gaaaagaucc ugacauucag aaucggcuac uacgucggac cgcuggcaag aggaaacagc 1380
 agauucgcau ggaugacaag aaagagcgaa gaaacaauca caccugggaa cuucgaagaa 1440
 gucgucgaca agggagcaag cgcacagagc uucaucgaaa gaaugacaaa cuucgacaag 1500
 aaccugccga acgaaaaggu ccugccgaag cacagccugc uguacgaaua cuucacaguc 1560
 uacaacgaac ugacaaaaggu caaguacguc acagaaggaa ugagaaagcc ggcauuccug 1620
 agcggagaac agaagaaggc aaucgucgac cugcuguuca agacaaacag aaaggucaca 1680
 gucaagcagc ugaaggaaga cuacuuaag aagaucgaau gcuucgacag cgucgaaauc 1740
 agcggagucg aagacagauu caacgcaagc cugggaacau accacgaccu gcugaagauc 1800
 aucaaggaca aggacuuccu ggacaacgaa gaaaacgaag acauccugga agacaucguc 1860
 cugacacuga cacuguucga agacagagaa augaucgaag aaagacugaa gacauacgca 1920
 caccuguucg acgacaaggu caugaagcag cugaagagaa gaagauacac aggaugggga 1980
 agacugagca gaaagcugau caacggaauc agagacaagc agagcggaaa gacaauccug 2040
 gacuuccuga agagcgacgg auucgcaaac agaaacuua ugcagcugau ccacgacgac 2100
 agccugacau ucaaggaaga cauccagaag gcacagguca gcggacaggg agacagccug 2160
 cacgaacaca ucgcaaaccu ggcaggaagc ccggcaauca agaagggaau ccugcagaca 2220
 gucaaggucg ucgacgaacu ggucaagguc augggaagac acaagccgga aaacaucguc 2280
 aucgaaaugg caagagaaaa ccagacaaca cagaagggac agaagaacag cagagaaaga 2340
 augaagagaa ucgaagaagg aaucaggaa cugggaagcc agauccugaa ggaacacccg 2400
 gucgaaaaca cacagcugca gaacgaaaag cuguaccugu acuaccugca gaacggaaga 2460
 gacauguacg ucgaccagga acuggacauc aacagacuga gcgacuacga cgucgaccac 2520
 aucgucccg agagcuuccu gaaggacgac agcaucgaca acaagguccu gacaagaagc 2580
 gacaagaaca gaggaagag cgacaacguc ccgagcgaag aagucgucaa gaagaugaag 2640
 aacuacugga gacagcugcu gaacgcaaag cugaucacac agagaaaguu cgacaaccug 2700
 acaaaggcag agagaggagg acugagcgaa cuggacaagg caggauucau caagagacag 2760
 cuggucgaaa caagacagau cacaagcac gucgcacaga uccuggacag cagaaugaac 2820
 acaaaguacg acgaaaacga caagcugauc agagaaguca aggucaucac acugaagagc 2880
 aagcugguca gcgacuucag aaaggacuuc caguucuaca agguacagaga aaucacaac 2940
 uaccaccacg cacacgacgc auaccugaac gcagucgucg gaacagcacu gaucaagaag 3000
 uaccgaagc uggaaagcga auucgucuac ggagacuaca agguacuacga cgucagaaag 3060

10

20

30

40

50

augaucgcaa agagcgaaca ggaaaucgga aaggcaacag caaaguacuu cuucucacagc 3120
 aacaucauga acuucuucaa gacagaaauc acacuggcaa acggagaaaau cagaaaagaga 3180
 ccgcugaucg aaacaaacgg agaaacagga gaaaucgucu gggacaaggg aagagacuuc 3240
 gcaacaguca gaaagguccu gagcaugccg caggucaaca ucgucaagaa gacagaaguc 3300
 cagacaggag gauucagcaa ggaaagcauc cugccgaaga gaaacagcga caagcugauc 3360
 gcaagaaaga aggacuggga cccgaagaag uacggaggau ucgacagccc gacagucgca 3420
 uacagcgucc uggucgucgc aaaggucgaa aagggaaga gcaagaagcu gaagagcguc 3480
 aaggaaucgc ugggaaucaac aaucauggaa agaagcagcu ucgaaaagaa cccgaucgac 3540
 uuccuggaag caaagggaua caaggaaguc aagaaggacc ugaucaucaa gcugccgaag 3600
 uacagccugu ucgaacugga aaacggaaga aagagaaguc uggcaagcgc aggagaacug 3660
 cagaagggaac acgaacuggc acugccgagc aaguacguca acuuccugua ccugggaagc 3720
 cacuacgaaa agcugaaggg aagcccggaa gacaacgaac agaagcagcu guucgucgaa 3780
 cagcacaagc acuaccugga cgaaucauc gaacagauca gcgaauucag caagagaguc 3840
 auccuggcag acgcaaaccu ggacaagguc cugagcgcau acaacaagca cagagacaag 3900
 ccgaucagag aacaggcaga aaacaucauc caccuguuca cacugacaaa ccugggagca 3960
 ccggcagcau ucaaguacuu cgacacaaca aucgacagaa agagauacac aagcacaag 4020
 gaaguccugg acgcaacacu gauccaccag agcaucacag gacugucgca aacaagaau 4080
 gaccugagcc agcugggagg agacggaagc ggaagcccga agaagaagag aaaggucgac 4140
 ggaagcccga agaagaagag aaaggucgac agcggauag 4179

10

20

<210> 24

<211> 4173

<212> RNA

<213> Artificial Sequence

<220>

<221> source

 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 24

gacaagaagu acagcaucgg acuggacaua ggaacaaaca gcgucggaug ggcagucauc 60
 acagacgaau acaagguccc gagcaagaag uucaaggucc ugggaaacac agacagacac 120
 agcaucaaga agaaccugau cggagcacug cuguucgaca gcggagaaac agcagaagca 180
 acaagacuga agagaacagc aagaagaaga uacacaagaa gaaagaacag aaucugcuac 240
 cugcaggaaa ucuucagcaa cgaauaggca aaggucgacg acagcuucuu ccacagacug 300
 gaagaaagcu uccuggucga agaagacaag aagcacgaaa gacacccgau cuucggaaac 360
 aucgucgacg aagucgcaua ccacgaaaag uacccgacaa ucuaccaccu gagaaagaag 420
 cuggucgaca gcacagacaa ggacagaccug agacugaucu accuggcacu ggacacaug 480
 aucaaguuca gaggacacuu ccugaucgaa ggagaccuga acccgacaa cagcagcguc 540
 gacaagcugu ucauccagcu gguccagaca uacaaccagc uguucgaaga aaacccgauc 600
 aacgcaagcg gagucgacgc aaaggcauuc cugagcgcaa gacugagcaa gagcagaaga 660
 cuggaaaacc ugaucgcaca gcugccggga gaaaagaaga acggacuguu cggaaaccug 720
 aucgcacuga gccugggacu gacaccgaac uucaagagca acuucgaccu ggcagaagac 780
 gcaaagcugc agcugagcaa ggacacauac gacgacgacc uggacaaccu gcuggcacag 840
 aucggagacc aguacgcaga ccuguuccug gcagcaaaga accugagcga cgcaauccug 900
 cugagcgaca uccugagagu caacacagaa aucacaaagg caccgcugag cgcaagcaug 960
 aucaagagau acgacgaaca ccaccaggac cugacacugc ugaaggcacu ggucagacag 1020
 cagcugccgg aaaaguacaa ggaaauuc uucgaccaga gcaagaacgg auacgcagga 1080
 ucaucgacg gaggagcaag ccaggaagaa uucuaaagu ucaucaagcc gauccuggaa 1140
 aagauggacg gaacagaaga acugcugguc aagcugaaca gagaagaccu gcugagaaag 1200
 cagagaacau ucgacaacgg aagcaucccg caccagauc accugggaga acugcacgca 1260

30

40

50

auccugagaa gacaggaaga cuucuacccg uuccugaagg acaacagaga aaagaucgaa 1320
aagauccuga cauucagaaau cccguacuac gucggaccgc uggcaagagg aaacagcaga 1380
uucgcaugga ugacaagaaa gagcgaagaa acaaucacac cguggaacuu cgaagaaguc 1440
gucgacaagg gagcaagcgc acagagcuuc aucgaaagaa ugacaaacuu cgacaagaac 1500
cugccgaacg aaaagguccu gccgaagcac agccugcugu acgaauacuu cacagucuac 1560
aacgaacuga caaaggucua guacgucaca gaaggaauga gaaagccggc auuccugagc 1620
ggagaacaga agaaggcaau cgucgaccug cuguucaaga caaacagaaa ggucacaguc 1680
aagcagcuga aggaagacua cuucaagaag aucgaaugcu ucgacagcgu cgaaaucagc 1740
ggagucgaag acagauucaa cgcaagccug ggaacauacc acgaccugcu gaagaaucau 1800
aaggacaagg acuuccugga caacgaagaa aacgaagaca uccuggaaga caucguccug 1860
acacugacac uguucgaaga cagagaaaug aucgaagaaa gacugaagac auacgcacac 1920
cuguucgacg acaaggucua gaagcagcug aagagaagaa gauacacagg auggggaaga 1980
cugagcagaa agcugaucaa cggaaucaga gacaagcaga gcggaaagac aauccuggac 2040
uuccugaaga gcgacggauu cgcaaacaga aacuuaugc agcugaucca cgacgacagc 2100
cugacauuca aggaagacau ccagaaggca caggucagcg gacagggaga cagccugcac 2160
gaacacaucg caaaccuggc aggaagcccg gcaaucaaga agggaauccu gcagacaguc 2220
aaggucgucg acgaacuggu caaggucaug ggaagacaca agccggaaaa caucgucauc 2280
gaaauggcaa gagaaaacca gacaacacag aaggagacaga agaacagcag agaaagaau 2340
aagagaauucg aagaaggaaau caaggaaucg ggaagccaga uccugaagga acaccggu 2400
gaaaacacac agcugcagaa cgaaaagcug uaccuguacu accugcagaa cggaagagac 2460
auguacgucg accaggaacu ggacaucaac agacugagcg acuacgacgu cgaccacau 2520
gucccgacaga gcuuccugaa ggacgacagc aucgacaaca agguccugac aagaagcgac 2580
aagaacagag gaaagagcga caacgucccg agcgaagaag ucgucaagaa gaugaagaac 2640
uacuggagac agcugcugaa cgcaaagcug aucacacaga gaaaguucga caaccugaca 2700
aaggcagaga gaggaggacu gagcgaacug gacaaggcag gauucaucaa gagacagcug 2760
gucgaaacaa gacagaucaac aaagcacguc gcacagaucg uggacagcag aaugaacaca 2820
aaguacgacg aaaacgacaa gcugaucaga gaagucaagg ucaucacacu gaagagcaag 2880
cuggucagcg acuucagaaa ggacuuccag uucuacaagg ucagagaaaau caacaacuac 2940
caccacgcac acgacgcaua ccugaacgca gucgucggaa cagcacugau caagaaguac 3000
ccgaagcugg aaagcgaaau cgucuacgga gacuacaagg ucuacgacgu cagaaagaug 3060
aucgcaaaga gcgaacagga aaucggaaag gcaacagcaa aguacuucuu cuacagcaac 3120
aucaugaacu ucuucaagac agaaaucaca cuggcaaacg gagaaaucag aaagagaccg 3180
cugaucgaaa caaacggaga aacaggagaa aucgucuggg acaagggaag agacuucgca 3240
acagucagaa agguccugag caugccgcag gucaacaucg ucaagaagac agaaguccag 3300
acaggaggau ucagcaagga aagcauccug ccgaagagaa acagcgacaa gcugaucgca 3360
agaaagaagg acugggaccc gaagaaguac ggaggauucg acagcccgac agucgcauac 3420
agcuccugg ucgucgcaaa ggucgaaaag ggaaagagca agaagcugaa gagcgucaag 3480
gaacugcugg gaaucacaa cauggaaaag agcagcuucg aaaagaaccc gaucgacuuc 3540
cuggaagcaa agggauacaa ggaagucaag aaggaccuga ucaucaagcu gccgaaguac 3600
agccuguucg aacuggaaaa cggaagaaaag agaaugcugg caagcgagagg agaacugcag 3660
aagggaacg aacuggcacu gccgagcaag uacgucaacu uccuguaccu ggcaagccac 3720
uacgaaaagc ugaagggaag cccggaagac aacgaacaga agcagcuguu cgucgaacag 3780
cacaagcacu accuggacga aaucaucgaa cagaucagcg aaucagcaa gagagucauc 3840
cuggcagacg caaaccugga caagguccug agcgcauaca acaagcacag agacaagccg 3900
aucagagaac aggcagaaaa caucauccac cuguucacac ugacaaaccu gggagcaccg 3960
gcagcauua aguacuucga cacaacaau gacagaaaga gauacacaag cacaagggaa 4020
guccuggacg caacacugau ccaccagagc aucacaggac uguacgaaac aagaaucgac 4080
cugagccagc ugggaggaga cggaagcgga agcccgaaga agaagagaaa ggucgacgga 4140
agcccgaaga agaagagaaa ggucgacagc gga 4173

10

20

30

40

50

<210> 25
 <211> 1392
 <212> PRT
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polypeptide"
 <400> 25

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val
 1 5 10 15

10

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
 20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
 35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
 50 55 60

20

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
 65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
 85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
 100 105 110

30

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
 115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
 130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
 145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
 165 170 175

40

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
 180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
 195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
 210 215 220

50

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
290 295 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305 310 315 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
325 330 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
340 345 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
355 360 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
370 375 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385 390 395 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
435 440 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr

10

20

30

40

50

485 490 495
Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510
Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525
Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540
Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560
Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575
Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590
Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605
Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620
Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640
His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655
Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670
Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685
Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700
Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720
His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
725 730 735
Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
740 745 750

10

20

30

40

50

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
885 890 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
900 905 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
915 920 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
930 935 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
1010 1015 1020

10

20

30

40

50

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
 1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
 1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
 1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
 1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
 1085 1090 1095

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys
 1100 1105 1110

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro
 1115 1120 1125

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val
 1130 1135 1140

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys
 1145 1150 1155

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser
 1160 1165 1170

Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys
 1175 1180 1185

Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu
 1190 1195 1200

Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly
 1205 1210 1215

Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val
 1220 1225 1230

Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser
 1235 1240 1245

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys
 1250 1255 1260

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys

10

20

30

40

50

1265	1270	1275
Arg Val Ile Leu Ala Asp Ala	Asn Leu Asp Lys Val	Leu Ser Ala
1280	1285	1290
Tyr Asn Lys His Arg Asp Lys	Pro Ile Arg Glu Gln	Ala Glu Asn
1295	1300	1305
Ile Ile His Leu Phe Thr Leu	Thr Asn Leu Gly Ala	Pro Ala Ala
1310	1315	1320
Phe Lys Tyr Phe Asp Thr Thr	Ile Asp Arg Lys Arg	Tyr Thr Ser
1325	1330	1335
Thr Lys Glu Val Leu Asp Ala	Thr Leu Ile His Gln	Ser Ile Thr
1340	1345	1350
Gly Leu Tyr Glu Thr Arg Ile	Asp Leu Ser Gln Leu	Gly Gly Asp
1355	1360	1365
Gly Ser Gly Ser Pro Lys Lys	Lys Arg Lys Val Asp	Gly Ser Pro
1370	1375	1380
Lys Lys Lys Arg Lys Val Asp	Ser Gly	
1385	1390	

<210> 26

<211> 4179

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 26

auggacaaga aguacagcau cggacuggca aucggaacaa acagcgucgg augggcaguc	60
aucacagacg aaaucaaggu cccgagcaag aaguucaagg uccugggaaa cacagacaga	120
cacagcauca agaagaaccu gaucggagca cugcuguucg acagcggaga aacagcagaa	180
gcaacaagac ugaagagaac agcaagaaga agauacacaa gaagaaagaa cagaaucugc	240
uaccugcagg aaaucuucag caacgaaaug gcaaaggucg acgacagcuu cuuccacaga	300
cuggaagaaa gcuuccuggu cgaagaagac aagaagcacg aaagacaccc gaucuucgga	360
aauaucgucg acgaagucgc auaccacgaa aaguaccgga caaucuacca ccugagaaag	420
aagcuggucg acagcacaga caaggcagac cugagacuga ucuaccuggc acuggcacac	480
augaucaagu ucagaggaca cuuccugauc gaaggagacc ugaacccgga caacagcgac	540
gucgacaagc uguucaucca gcugguccag acauacaacc agcuguucga agaaaacccg	600
aucaacgcaa gcggagucga cgcaaaggca auccugagcg caagacugag caagagcaga	660
agacuggaaa accugaucgc acagcugccg ggagaaaaga agaacggacu guucggaaac	720
cugaucgcac ugagccuggg acugacaccg aacuucaaga gcaacuucga ccuggcagaa	780
gacgcaaagc ugcagcugag caaggacaca uacgacgacg accuggacaa ccugcuggca	840
cagaucggag accaguacgc agaccuguuc cuggcagcaa agaaccugag cgacgcaauc	900

10

20

30

40

50

cugcugagcg acauccugag agucaacaca gaaaucacaa aggcaccgcu gagcgcaagc 960
 augaucaaga gauacgacga acaccaccag gaccugacac ugcugaaggc acuggucaga 1020
 cagcagcugc cggaagaa gua caaggaauc uucucgacc agagcaagaa cggauacgca 1080
 ggauacaucg acggaggagc aagccaggaa gaaucuaa aguucauca gccgauccug 1140
 gaaaagaugg acggaacaga agaacugcug gucaagcuga acagagaaga ccugcugaga 1200
 aagcagagaa cauucgacaa cggagcauc ccgcaccaga uccaccuggg agaacugcac 1260
 gcaauccuga gaagacagga agacuucua ccguuccuga aggacaacag agaaaagauc 1320
 gaaaagauc ugcacauucag aaucuccguac uacgucggac cgcuggcaag aggaaacagc 1380
 agauucgcau ggauagacaag aaagagcgaa gaaacaauca caccguggaa cuucgaagaa 1440
 gucgucgaca agggagcaag cgcacagagc uucaucgaaa gaaugacaaa cuucgacaag 1500
 aaccugccga acgaaaaggu ccugccgaag cacagccugc uguacgaaua cuucacaguc 1560
 uacaacgaac ugacaaaggu caaguacguc acagaaggaa ugagaaagcc ggcauuccug 1620
 agcggagaac agaagaaggc aaucgucgac cugcuguuca agacaaacag aaaggucaca 1680
 gucaagcagc ugaaggaaga cuacuucua agaucgaau gcuucgacag cgucgaaauc 1740
 agcggagucg aagacagauu caacgcaagc cugggaacau accacgaccu gcugaagauc 1800
 aucaaggaca aggacuuccu ggacaacgaa gaaaacgaag acauccugga agacaucguc 1860
 cugacacuga cacuguucga agacagagaa augaucgaag aaagacugaa gacauacgca 1920
 caccuguucg acgacaaggu caugaagcag cugaagagaa gaagauacac aggaugggga 1980
 agacugagca gaaagcugau caacggaauc agagacaagc agagcggaaa gacaauccug 2040
 gacuuccuga agagcgacgg auucgcaaac agaaacuua ugcagcugau ccacgacgac 2100
 agccugacau ucaaggaaga cauccagaag gcacagguca gcggacaggg agacagccug 2160
 cacgaacaca ucgcaaaccu ggacaggaagc ccggcaauca agaagggaau ccugcagaca 2220
 gucaaggucg ucgacgaacu ggucagguc augggaagac acaagccgga aaacaucguc 2280
 aucgaaaugg caagagaaaa ccagacaaca cagaagggac agaagaacag cagagaaaga 2340
 augaagagaa ucgaagaagg aaucaggaa cugggaagcc agauccugaa ggaacacccg 2400
 gucgaaaaca cacagcugca gaacgaaaag cuguaccugu acuaccugca gaacggaaga 2460
 gacauguacg ucgaccagga acuggacauc aacagacuga gcgacuacga cgucgaccac 2520
 aucgucccg agagcuuccu gaaggacgac agcaucgaca acaagguccu gacaagaagc 2580
 gacaagaaca gaggaagag cgacaacguc ccgagcgaag aagucgucaa gaagaugaag 2640
 aacuacugga gacagcugcu gaacgcaaag cugaucacac agagaaaguu cgacaaccug 2700
 acaaaggcag agagaggagg acugagcgaa cuggacaagg caggauucau caagagacag 2760
 cuggucgaaa caagacagau cacaagcac gucgcacaga uccuggacag cagaaugaac 2820
 acaaaguacg acgaaaacga caagcugauc agagaaguca aggucaucac acugaagagc 2880
 aagcugguca gcgacuucag aaaggacuuc caguucuaa agguacagaga aaucacaac 2940
 uaccaccag cacacgacgc auaccugaac gcagucgucg gaacagcacu gaucaagaag 3000
 uaccgaagc uggaaagcga auucgucua ggagacuaca agguacuacga cgucagaaag 3060
 augaucgaa agagcgaaca ggaaucgga aaggcaacag caaaguacuu cuucuaacagc 3120
 acaucauga acuucuaa gacagaaauc acacuggcaa acggagaaa cagaaagaga 3180
 ccgugaucg aaacaaacgg agaaacagga gaaucgucu gggacaaggg aagagacuuc 3240
 gcaacaguca gaaagguccu gagcaugccg cagguaaca ucgucaagaa gacagaaguc 3300
 cagacaggag gauucagcaa ggaaagcauc cugccgaaga gaaacagcga caagcugauc 3360
 gcaagaaaga aggacuggga cccgaagaag uacggaggau ucgacagccc gacagucgca 3420
 uacagcgucc uggucgucgc aaaggucgaa aagggaaga gcaagaagcu gaagagcguc 3480
 aaggaaucg ugggaauac aaucaggaa agaagcagcu ucgaaaagaa cccgaucgac 3540
 uuccuggaag caaagggaau caagggaagc aagaaggacc ugaucuaa gcugccgaag 3600
 uacagccugu ucgaacugga aaacggaaga aagagaagc uggcaagcgc aggagaacug 3660
 cagaaggga acgaacuggc acugccgagc aaguacguc acuuccugua ccuggcaagc 3720
 cacuacgaaa agcugaaggg aagcccggaa gacaacgaac agaagcagcu guucgucgaa 3780
 cagcacaagc acuaccugga cgaaucua gaacagauca gcgaauucag caagagaguc 3840
 auccuggcag acgcaaaccu ggacaagguc cugagcgcau acaacaagca cagagacaag 3900

10

20

30

40

50

ccgaucagag aacaggcaga aaacaucauc caccuguuua cacugacaaa ccugggagca 3960
 ccggcagcau ucaaguacuu cgacacaaca aucgacagaa agagauacac aagcacaaag 4020
 gaaguccugg acgcaacacu gauccaccag agcaucacag gacuguacga aacaagaau 4080
 gaccugagcc agcugggagg agacggaagc ggaagcccga agaagaagag aaaggucgac 4140
 ggaagcccga agaagaagag aaaggucgac agcggauag 4179

<210> 27

<211> 4173

<212> RNA

<213> Artificial Sequence

10

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 27

gacaagaagu acagcaucgg acuggcaauc ggaacaaaca gcgucggaug ggcagucauc 60
 acagacgaau acaagguccc gagcaagaag uucaaggucc ugaggaaacac agacagacac 120
 agcaucaaga agaaccugau cggagcacug cuguucgaca gcggagaaac agcagaagca 180
 acaagacuga agagaacagc aagaagaaga uacacaagaa gaaagaacag aaucugcuac 240
 cugcaggaaa ucuucagcaa cgaaauggca aaggucgacg acagcuucuu ccacagacug 300
 gaagaaagcu uccuggucga agaagacaag aagcacgaaa gacaccgga cuucggaaac 360
 aucgucgacg aagucgcaua ccacgaaaag uaccgacaa ucuaccaccu gagaaagaag 420
 cuggucgaca gcacagacaa ggacagaccug agacugaucu accuggcacu ggcacacaug 480
 aucaaguua gagagacacu ccugaucgaa ggagaccuga acccgacaa cagcagcugc 540
 gacaagcugu ucauccagcu gguccagaca uacaaccagc uguucgaaga aaaccggauc 600
 aacgcaagcg gagucgacgc aaaggcaauc cugagcgcaa gacugagcaa gagcagaaga 660
 cuggaaaacc ugaucgcaca gcugccggga gaaaagaaga acggacuguu cggaaccug 720
 aucgcacuga gccugggacu gacaccgaac uucaagagca acuucgaccu ggcagaagac 780
 gcaaagcugc agcugagcaa ggacacauac gacgacgacc uggacaaccu gcuggcacag 840
 aucggagacc aguacgcaga ccuguuccug gcagcaaaga accugagcga cgcaauccug 900
 cugagcgaca uccugagagu caacacagaa aucacaaagg caccgcugag cgcaagcaug 960
 aucaagagau acgacgaaca ccaccaggac cugacacugc ugaaggcacu ggucagacag 1020
 cagcugccgg aaaaguacaa ggaaaucuuc uucgaccaga gcaagaacgg auacgcagga 1080
 ucaucgacg gaggagcaag ccaggaagaa uucuacaagu ucaucaagcc gauccuggaa 1140
 aagauggacg gaacagaaga acugcugguc aagcugaaca gagaagaccu gcugagaaag 1200
 cagagaacau ugcacaacgg aagcaucccg caccagauc accugggaga acugcacgca 1260
 auccugagaa gacaggaaga cuucuaccg uuccugaagg acaacagaga aaagaucgaa 1320
 aagauccuga cauucagaau cccguacuac gucggaccgc uggcaagagg aaacagcaga 1380
 uucgcaugga ugacaagaaa gagcgaagaa acaaucacac cguggaacuu cgaagaaguc 1440
 gucgacaagg gagcaagcgc acagagcuuc aucgaaagaa ugacaaacuu cgacaagaac 1500
 cugccgaacg aaaagguccu gccgaagcac agccugcugu acgaauacuu cacagucua 1560
 aacgaacuga caaaggucua guacgucaca gaaggaauga gaaagccggc auuccugagc 1620
 ggagaacaga agaaggcaau cgucgaccug cuguucaaga caaacagaaa ggucacaguc 1680
 aagcagcuga aggaagacua cuucaagaag aucgaaugcu ucgacagcgu cgaaaucagc 1740
 ggagucgaag acagauucaa cgcaagccug ggaacauacc acgaccugcu gaagaucac 1800
 aaggacaagg acuuccugga caacgaagaa aacgaagaca uccuggaaga caucguccug 1860
 acacugacac uguucgaaga cagagaaaug aucgaagaaa gacugaagac auacgcacac 1920
 cuguucgacg acaaggucua gaagcagcug aagagaagaa gauacacagg auggggaaga 1980
 cugagcagaa agcugaucaa cggaaucaga gacaagcaga gcggaaagac aauccuggac 2040
 uuccugaaga gcgacggauu cgcaaacaga aacuucaugc agcugaacca cgacgacagc 2100

20

30

40

50

cugacauuca aggaagacau ccagaaggca caggucagcg gacagggaga cagccugcac 2160
 gaacacaucg caaaccuggc aggaagcccg gcaaucaaga agggaauccu gcagacaguc 2220
 aaggucgucg acgaacuggu caaggucaug ggaagacaca agccggaaaa caucgucauc 2280
 gaaauggcaa gagaaaacca gacaacacag aaggggacaga agaacagcag agaaagaaug 2340
 aagagaauucg aagaaggaau caaggaacug ggaagccaga uccugaagga acacccgguc 2400
 gaaaacacac agcugcagaa cgaaaagcug uaccuguacu accugcagaa cggaagagac 2460
 auguacgucg accaggaacu ggacaucaac agacugagcg acuacgacgu cgaccacauc 2520
 gucccgcaga gcuuccugaa ggacgacagc aucgacaaca agguccugac aagaagcgac 2580
 aagaacagag gaaagagcga caacgucccg agcgaagaag ucgucaagaa gaugaagaac 2640
 uacuggagac agcugcugaa cgaaaagcug aucacacaga gaaaguucga caaccugaca 2700
 aaggcagaga gaggaggacu gagcgaacug gacaaggcag gauucauca gagacagcug 2760
 gucgaaacaa gacagaucaac aaagcacguc gcacagauc uggacagcag aaugaacaca 2820
 aaguacgacg aaaacgacaa gcugaucaga gaagucaagg ucaucacacu gaagagcaag 2880
 cuggucagcg acuucagaaa ggacuuccag uucuaacaagg ucagagaaaau caacaacuac 2940
 caccacgcac acgacgcaua ccugaacgca gucgucggaa cagcacugau caagaaguac 3000
 ccgaagcugg aaagcgaaau cgucuacgga gacuacaagg ucuacgacgu cagaaagaug 3060
 aucgcaaaga gcgaacagga aaucggaaaag gcaacagcaa aguacuucuu cuacagcaac 3120
 aucaugaacu ucuucaagac agaaaucaca cuggcaaacg gagaaaucag aaagagaccg 3180
 cugaucgaaa caaacggaga aacaggagaa aucgucuggg acaagggaag agacuucgca 3240
 acagucagaa agguccugag caugccgcag gucaacaucg ucaagaagac agaaguccag 3300
 acaggaggau ucagcaagga aagcauccug ccgaagagaa acagcgacaa gcugaucgca 3360
 agaaagaagg acugggaccc gaagaaguac ggaggauucg acagcccgcac agucgcauac 3420
 agcguccugg ucgucgcaaa ggucgaaaag ggaaagagca agaagcugaa gagcguaag 3480
 gaacugcugg gaaucacaa cauggaaaaga agcagcuucg aaaagaaccc gaucgacuuc 3540
 cuggaagcaa agggauacaa ggaagucaag aaggaccuga ucaucaagcu gccgaaguac 3600
 agccuguucg aacuggaaaa cggaagaaaag agaaugcugg caagcgcagg agaacugcag 3660
 aagggaacg aacuggcacu gccgagcaag uacgucaacu uccuguaccu ggcaagccac 3720
 uacgaaaagc ugaagggaag cccggaagac aacgaacaga agcagcuguu cgucgaacag 3780
 cacaagcacu accuggacga aaucaucgaa cagaucagcg aaucagcaa gagagucauc 3840
 cuggcagacg caaaccugga caagguccug agcgcauaca acaagcacag agacaagccg 3900
 aucagagaac aggcagaaaa caucauccac cuguucacac ugacaaaccu gggagcaccg 3960
 gcagcauua aguacuucga cacaacauc gacagaaaaga gauacacaag cacaagggaa 4020
 guccuggacg caacacugau ccaccagagc aucacaggac uguacgaaac aagaaucgac 4080
 cugagccagc ugaggaggaga cggaagcgga agcccgaaga agaagagaaa ggucgacgga 4140
 agcccgaaga agaagagaaa ggucgacagc gga 4173

<210> 28

<211> 1392

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 28

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30

10

20

30

40

50

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
 35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
 50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
 65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
 85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
 100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
 115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
 130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
 145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
 165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
 180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
 195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
 210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
 225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
 245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
 260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
 275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp

10

20

30

40

50

290 295 300
Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305 310 315 320
Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
325 330 335
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
340 345 350
Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
355 360 365
Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
370 375 380
Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385 390 395 400
Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415
Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430
Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
435 440 445
Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460
Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480
Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495
Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510
Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525
Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540
Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560

10

20

30

40

50

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

10

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670

20

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720

30

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
770 775 780

40

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830

50

Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys
835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
885 890 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
900 905 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
915 920 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
930 935 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr

10

20

30

40

50

1085	1090	1095
Glu Val Gln Thr Gly	Gly Phe Ser Lys	Glu Ser Ile Leu Pro Lys
1100	1105	1110
Arg Asn Ser Asp Lys	Leu Ile Ala Arg	Lys Lys Asp Trp Asp Pro
1115	1120	1125
Lys Lys Tyr Gly Gly	Phe Asp Ser Pro	Thr Val Ala Tyr Ser Val
1130	1135	1140
Leu Val Val Ala Lys	Val Glu Lys Gly	Lys Ser Lys Lys Leu Lys
1145	1150	1155
Ser Val Lys Glu Leu	Leu Gly Ile Thr	Ile Met Glu Arg Ser Ser
1160	1165	1170
Phe Glu Lys Asn Pro	Ile Asp Phe Leu	Glu Ala Lys Gly Tyr Lys
1175	1180	1185
Glu Val Lys Lys Asp	Leu Ile Ile Lys	Leu Pro Lys Tyr Ser Leu
1190	1195	1200
Phe Glu Leu Glu Asn	Gly Arg Lys Arg	Met Leu Ala Ser Ala Gly
1205	1210	1215
Glu Leu Gln Lys Gly	Asn Glu Leu Ala	Leu Pro Ser Lys Tyr Val
1220	1225	1230
Asn Phe Leu Tyr Leu	Ala Ser His Tyr	Glu Lys Leu Lys Gly Ser
1235	1240	1245
Pro Glu Asp Asn Glu	Gln Lys Gln Leu	Phe Val Glu Gln His Lys
1250	1255	1260
His Tyr Leu Asp Glu	Ile Ile Glu Gln	Ile Ser Glu Phe Ser Lys
1265	1270	1275
Arg Val Ile Leu Ala	Asp Ala Asn Leu	Asp Lys Val Leu Ser Ala
1280	1285	1290
Tyr Asn Lys His Arg	Asp Lys Pro Ile	Arg Glu Gln Ala Glu Asn
1295	1300	1305
Ile Ile His Leu Phe	Thr Leu Thr Asn	Leu Gly Ala Pro Ala Ala
1310	1315	1320
Phe Lys Tyr Phe Asp	Thr Thr Ile Asp	Arg Lys Arg Tyr Thr Ser
1325	1330	1335

10

20

30

40

50

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
 1340 1345 1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
 1355 1360 1365

Gly Ser Gly Ser Pro Lys Lys Lys Arg Lys Val Asp Gly Ser Pro
 1370 1375 1380

Lys Lys Lys Arg Lys Val Asp Ser Gly
 1385 1390

10

<210> 29

<211> 4179

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

20

<400> 29

auggacaaga aguacagcau cggacuggca aucggaacaa acagcgucgg augggcaguc 60
 aucacagacg aaucacaaggu cccgagcaag aaguucaagg uccugggaaa cacagacaga 120
 cacagcauca agaagaaccu gaucggagca cugcuguucg acagcggaga aacagcagaa 180
 gcaacaagac ugaagagaac agcaagaaga agauacacaa gaagaaagaa cagaaucugc 240
 uaccugcagg aaaucuucag caacgaaaug gcaaaggucg acgacagcuu cuuccacaga 300
 cuggaagaaa gcuuccuggu cgaagaagac aagaagcacg aaagacaccc gaucuucgga 360
 aaucugucg acgaagucgc auaccacgaa aaguaccga caaucuacca ccugagaaag 420
 aagcuggucg acagcacaga caaggcagac cugagacuga ucuaccuggc acuggcacac 480
 augaucaagu ucagaggaca cuuccugauc gaaggagacc ugaacccgga caacagcgac 540
 gucgacaagc uguucaucca gcugguccag acauacaacc agcuguucga agaaaacccg 600
 aucaacgcaa gcggagucga cgcaaaggca auccugagcg caagacugag caagagcaga 660
 agacuggaaa accugaucgc acagcugccg ggagaaaaga agaacggacu guucggaaac 720
 cugaucgcac ugagccuggg acugacaccg aacuucaaga gcaacuucga ccuggcagaa 780
 gacgcaaagc ugcagcugag caaggacaca uacgacgacg accuggacaa ccugcuggca 840
 cagaucggag accaguacgc agaccuguuc cuggcagcaa agaaccugag cgacgcaauc 900
 cugcugagcg acauccugag agucaacaca gaaucacaa aggcaccgcu gagcgcaagc 960
 augaucaaga gauacgacga acaccaccag gaccugacac ugcugaaggc acuggucaga 1020
 cagcagcugc cggaaaagua caaggaaauc uucuucgacc agagcaagaa cggauacgca 1080
 ggauacaucg acggaggagc aagccaggaa gaauucuaca aguucauca gccgauccug 1140
 gaaaagaugg acggaacaga agaacugcug gucaagcuga acagagaaga ccugcugaga 1200
 aagcagagaa cauucgacaa cggaagcauc ccgcaccaga uccaccuggg agaacugcac 1260
 gcaauccuga gaagacagga agacuucuac ccguuccuga aggacaacag agaaaagauc 1320
 gaaaagaucc ugacauucag aaucccgua uacgucggac cgcuggcaag aggaaacagc 1380
 agauucgcau ggaugacaag aaagagcgaa gaaacaauca caccguggaa cuucgaagaa 1440
 gucgucgaca agggagcaag cgcacagagc uucaucgaaa gaaugacaaa cuucgacaag 1500
 aaccugccga acgaaaaggu ccugccgaag cacagccugc uguacgaaua cuucacaguc 1560
 uacaacgaac ugacaaaaggu caaguacguc acagaaggaa ugagaaagcc ggcauuccug 1620
 agcggagaaac agaagaaggc aaucgucgac cugcuguuca agacaaacag aaaggucaca 1680
 gucaagcagc ugaaggaaga cuacuucaag aagaucgaau gcuucgacag cgucgaaauc 1740

30

40

50

```

agcggagucg aagacagauu caacgcaagc cugggaacau accacgaccu gcugaagauc 1800
aucaaggaca aggacuuccu ggacaacgaa gaaaacgaag acauccugga agacaucguc 1860
cugacacuga cacuguucga agacagagaa augaucgaag aaagacugaa gacauacgca 1920
caccuguucg acgacaaggu caugaagcag cugaagagaa gaagauacac aggaugggga 1980
agacugagca gaaagcugau caacggaauc agagacaagc agagcggaaa gacaauccug 2040
gacuuccuga agagcgacgg auucgcaaac agaaacuua ugcagcugau ccacgacgac 2100
agccugacau ucaaggaaga cauccagaag gcacagguca gcggacaggg agacagccug 2160
cacgaacaca ucgcaaaccu ggacggaaagc ccggcaauca agaagggaau ccugcagaca 2220
gucaaggucg ucgacgaacu gguaagguc augggaagac acaagccgga aaacaucguc 2280
aucgaaaugg caagagaaaa ccagacaaca cagaagggac agaagaacag cagagaaaga 2340
augaagagaa ucgaagaagg aaucaggaa cugggaagcc agauccugaa ggaacacccg 2400
gucgaaaaca cacagcugca gaacgaaaag cuguaccugu acuaccugca gaacggaaga 2460
gacauguacg ucgaccagga acuggacauc aacagacuga gcgacuacga cgucgacgca 2520
aucgucccg agagcuuccu gaaggacgac agcaucgaca acaagguccu gacaagaagc 2580
gacaagaaca gaggaagag cgacaacguc ccgagcgaag aagucgucaa gaagaugaag 2640
aacuacugga gacagcugcu gaacgcaaag cugaucacac agagaaaguu cgacaaccug 2700
acaaaggcag agagaggagg acugagcga cuggacaagg caggauucau caagagacag 2760
cuggucgaaa caagacagau cacaagcac gucgcacaga uccuggacag cagaaugaac 2820
acaaaguacg acgaaaacga caagcugauc agagaaguca aggucaucac acugaagagc 2880
aagcugguca gcgacuucag aaaggacuuc caguucuaca aggucagaga aaucacaac 2940
uaccaccacg cacacgacgc auaccugaac gcagucgucg gaacagcacu gaucaagaag 3000
uacccgaagc uggaaagcga auucgucuac ggagacuaca aggucuacga cgucagaaag 3060
augaucgcaa agagcgaaca ggaaaucgga aaggcaacag caaaguacuu cuucuacagc 3120
aacaucauga acuucuuaa gacagaaauc acacuggcaa acggagaaaau cagaaagaga 3180
ccgcugaucg aaacaaacgg agaaacagga gaaaucgucu gggacaaggg aagagacuuc 3240
gcaacaguca gaaagguccu gagcaugccg caggucaaca ucgucaagaa gacagaaguc 3300
cagacaggag gauucagcaa ggaaagcauc cugccgaaga gaaacagcga caagcugauc 3360
gcaagaaaga aggacuggga cccgaagaag uacggaggau ucgacagccc gacagucgca 3420
uacagcgucc uggucgucgc aaaggucgaa aagggaaga gcaagaagcu gaagagcguc 3480
aaggaaucgc ugggaauac aaucaggaa agaagcagcu ucgaaaagaa cccgaucgac 3540
uuccuggaag caaagggauc caaggaaguc aagaaggacc ugaucacaa gcugccgaag 3600
uacagccugu ucgaacugga aaacggaaga aagagaaguc uggcaagcgc aggagaacug 3660
cagaagggaa acgaacuggc acugccgagc aaguacguca acuuccugua ccuggcaagc 3720
cacuacgaaa agcugaaggg aagcccggaa gacaacgaac agaagcagcu guucgucgaa 3780
cagcacaagc acuaccugga cgaaaucauc gaacagauca gcgaauucag caagagaguc 3840
auccuggcag acgcaaaccu ggacaagguc cugagcgcau acaacaagca cagagacaag 3900
ccgaucagag aacaggcaga aaacaucauc caccuguua cacugacaaa ccugggagca 3960
ccggcagcau ucaaguacuu cgacacaaca aucgacagaa agagauacac aagcacaag 4020
gaaguccugg acgcaacacu gauccaccag agcaucacag gacuguacga aacaagaau 4080
gaccugagcc agcugggagg agacggaagc ggaagcccga agaagaagag aaaggucgac 4140
ggaagcccga agaagaagag aaaggucgac agcggauag 4179

```

<210> 30

<211> 4173

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

10

20

30

40

50

<400> 30

gacaagaagu acagcaucgg acuggcaauc ggaacaaaca gcgucggaug ggcagucauc	60	
acagacgaau acaagguccc gagcaagaag uucaaggucc ugaggaaacac agacagacac	120	
agcaucaaga agaaccugau cggagcacug cuguucgaca gcggagaaaac agcagaagca	180	
acaagacuga agagaacagc aagaagaaga uacacaagaa gaaagaacag aaucugcuac	240	
cugcaggaaa ucuucagcaa cgaaauggca aaggucgacg acagcuucuu ccacagacug	300	
gaagaaagcu uccuggucga agaagacaag aagcacgaaa gacacccgau cuucggaaac	360	
aucgucgacg aagucgcaua ccacgaaaag uacccgacaa ucuaccaccu gagaaagaag	420	
cuggucgaca gcacagacaa ggacagaccug agacugaucu accuggcacu ggacacacug	480	
aucaaguua gaggacacuu ccugaucgaa ggagaccuga acccgagacaa cagcgacguc	540	10
gacaagcugu ucauccagcu gguccagaca uacaaccagc uguucgaaga aaacccgauc	600	
aacgcaagcg gagucgacgc aaaggcaauc cugagcgcaa gacugagcaa gagcagaaga	660	
cuggaaaacc ugaucgcaca gcugccggga gaaaagaaga acggacuguu cggaaccug	720	
aucgcacuga gccugggacu gacaccgaac uucaagagca acuucgaccu ggacagaagac	780	
gcaaagcugc agcugagcaa ggacacauac gacgacgacc uggacaaccu gcuggcacag	840	
aucggagacc aguacgcaga ccuguuccug gcagcaaaga accugagcga cgcaauccug	900	
cugagcgaca uccugagagu caacacagaa aucacaaagg caccgcugag cgcaagcaug	960	
aucaagagau acgacgaaca ccaccaggac cugacacugc ugaaggcacu ggucagacag	1020	
cagcugccgg aaaaguacaa ggaaaucuu uucgaccaga gcaagaacgg auacgcagga	1080	
uacaucgacg gaggagcaag ccaggaagaa uuucacaagu ucaucaagcc gauccuggaa	1140	20
aagauggacg gaacagaaga acugcugguc aagcugaaca gagaagaccu gcugagaaag	1200	
cagagaacau ucgacaacgg aagcaucccg caccagaucc accugggaga acugcacgca	1260	
auccugagaa gacaggaaga cuucuaacccg uuccugaagg acaacagaga aaagaucgaa	1320	
aagauccuga cauucagaau cccguacuac gucggaccgc uggcaagagg aaacagcaga	1380	
uucgcaugga ugacaagaaa gagcgaagaa acaauacac cguggaaccuu cgaagaaguc	1440	
gucgacaagg gagcaagcgc acagagcuuc aucgaaagaa ugacaaacuu cgacaagaac	1500	
cugccgaacg aaaagguccu gccgaagcac agccugcugu acgaauacuu cacagucuaac	1560	
aacgaacuga caaaggucua guacgucaca gaaggaauga gaaagccggc auuccugagc	1620	
ggagaacaga agaaggcaau cgucgaccug cuguucaaga caaacagaaa ggucacaguc	1680	
aagcagcuga aggaagacua cuucaagaag aucgaaugcu ucgacagcgu cgaaaucagc	1740	30
ggagucgaag acagauuca cgcaagccug ggaacauacc acgaccugcu gaagaucac	1800	
aaggacaagg acuuccugga caacgaagaa aacgaagaca uccuggaaga caucguccug	1860	
acacugacac uguucgaaga cagagaaaug aucgaagaaa gacugaagac auacgcacac	1920	
cuguucgacg acaaggucuu gaagcagcug aagagaagaa gauacacagg auggggaaga	1980	
cugagcagaa agcugaucaa cggaaucaaga gacaagcaga gcggaaagac aauccuggac	2040	
uuccugaaga gcgacggauu cgcaaacaga aacuucaugc agcugaucca cgacgacagc	2100	
cugacauuca aggaagacau ccagaaggca caggucagcg gacagggaga cagccugcac	2160	
gaacacaucg caaaccuggc aggaagcccg gcaaucaaga agggaauccu gcagacaguc	2220	
aaggucgucg acgaacuggu caaggucaug ggaagacaca agccggaaaa caucgucauc	2280	
gaaauggcaa gagaaaacca gacaacacag aaggagcaga agaacagcag agaaagaag	2340	40
aagagaaucg aagaaggaa caaggaacug ggaagccaga uccugaagga acacccgguc	2400	
gaaaacacac agcugcagaa cgaaaagcug uaccuguacu accugcagaa cggaagagac	2460	
auguacgucg accaggaacu ggacaucaac agacugagcg acuacgacgu cgacgcaauc	2520	
gucccgacga gcuuccugaa ggacgacagc aucgacaaca agguccugac aagaagcgac	2580	
aagaacagag gaaagagcga caacgucccg agcgaagaag ucgucaagaa gaugaagaac	2640	
uacuggagac agcugcugaa cgcaaagcug aucacacaga gaaaguucga caaccugaca	2700	
aaggcagaga gaggaggacu gagcgaacug gacaaggcag gauucauca gagacagcug	2760	
gucgaaacaa gacagauca aaagcacguc gcacagauc uggacagcag aaugaacaca	2820	
aaguacgacg aaaacgacaa gcugaucaga gaagucaagg ucaucacacu gaagagcaag	2880	
cuggucagcg acuucagaaa ggacuuccag uucuaaagg ucagagaaa caacaacuac	2940	50

caccacgcac acgacgcaua ccugaacgca gucgucggaa cagcacugau caagaaguac 3000
 ccgaagcugg aaagcgaaau cgucuacgga gacuacaagg ucuacgacgu cagaaagaug 3060
 aucgcaaaga gcgaacagga aaucggaaag gcaacagcaa aguacuucuu cuacagcaac 3120
 aucaugaacu ucuucaagac agaaaucaca cuggcaaacg gagaaaucag aaagagaccg 3180
 cugaucgaaa caaacggaga aacaggagaa aucgucuggg acaagggaag agacuucgca 3240
 acagucagaa agguccugag caugccgcag gucaacaucg ucaagaagac agaaguccag 3300
 acaggaggau ucagcaagga aagcauccug ccgaagagaa acagcgacaa gcugaucgca 3360
 agaaagaagg acugggaccc gaagaaguac ggaggauucg acagcccgac agucgcacuac 3420
 agcguccugg ucgucgcaaa ggucgaaaag ggaaagagca agaagcugaa gagcgucaag 3480
 gaacugcugg gaaucacaau cauggaaaga agcagcuucg aaaagaaccc gaucgacuuc 3540
 cuggaagcaa agggauacaa ggaagucaag aaggaccuga ucaucaagcu gccgaaguac 3600
 agccuguucg aacuggaaaa cggaagaaag agaaucgugg caagcgcagg agaacugcag 3660
 aagggaacg aacuggcacu gccgagcaag uacgucaacu uccuguaccu ggcaagccac 3720
 uacgaaaagc ugaagggaag cccggaagac aacgaacaga agcagcuguu cgucgaacag 3780
 cacaagcacu accuggacga aaucaucgaa cagaucagcg aaucagcaa gagagucauc 3840
 cuggcagacg caaaccugga caagguccug agcgcauaca acaagcacag agacaagccg 3900
 aucagagaac aggcagaaaa caucauccac cuguucacac ugacaaaccu gggagcaccg 3960
 gcagcauua aguacuucga cacaacaau gacagaaaga gauacacaag cacaagaggaa 4020
 guccuggacg caacacugau ccaccagagc aucacaggac uguacgaaac aagaaucgac 4080
 cugagccagc ugggaggaga cggaagcgga agcccgaaga agaagagaaa ggucgacgga 4140
 agcccgaaga agaagagaaa ggucgacagc gga 4173

10

20

<210> 31

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

30

<400> 31

taatacgact cactata

17

<210> 32

<211> 50

<212> DNA

<213> Homo sapiens

<400> 32

acatttgctt ctgacacaac tgtgttcact agcaacctca aacagacacc

50

40

<210> 33

<211> 132

<212> DNA

<213> Homo sapiens

<400> 33

gctcgtttc ttgctgtcca atttctatta aaggttcctt tgttcctaa gtccaactac

60

taaactgggg gatattatga agggccttga gcatctggat tctgcctaataaaaaacatt

120

tattttcatt gc

132

<210> 34

50

<211> 66

<212> DNA

<213> Homo sapiens

<400> 34

cataaacctt ggcgcgctcg cggcccggca ctcttctggt cccacagac tcagagagaa 60
cccacc 66

<210> 35

<211> 110

<212> DNA

<213> Homo sapiens

<400> 35

gctggagcct cgggtggccat gcttcttgcc ccttgggcct cccccagcc cctcctcccc 60
ttcctgcacc cgtacccccg tggctcttga ataaagtctg agtgggcggc 110

<210> 36

<211> 29

<212> DNA

<213> Xenopus laevis

<400> 36

aagctcagaa taaacgctca actttggcc 29

<210> 37

<211> 130

<212> DNA

<213> Xenopus laevis

<400> 37

accagcctca agaacacccg aatggagtct ctaagctaca taataccaac ttacacttta 60
caaaatgttg tcccccaaaa tgtagccatt cgtatctgct cctaataaaa agaaagtttc 120
ttcacattct 130

<210> 38

<211> 27

<212> DNA

<213> Bos sp.

<400> 38

cagggtcctg tggacagctc accagct 27

<210> 39

<211> 102

<212> DNA

<213> Bos sp.

<400> 39

ttgccagcca tctgttgttt gcccctcccc cgtgccttcc ttgaccctgg aaggtgccac 60
tcccactgtc ctttctaata aaaatgagga aattgcatcg ca 102

<210> 40

<211> 93

<212> DNA

<213> Mus musculus

10

20

30

40

50

<400> 40

gctgccttct gcggggccttg cttctggcc atgcccttct tctctccctt gcacctgtac 60
ctcttgggtct ttgaataaag cctgagtagg aag 93

<210> 41

<211> 61

<212> DNA

<213> Unknown

<220>

<221> source

10

<223> /note="Description of Unknown:

HSD17B4 5' UTR sequence"

<400> 41

tcccgcagtc ggcgtccagc ggctctgctt gttcgtgtgt gtgtcgttgc aggccttatt 60
c 61

<210> 42

<211> 100

<212> RNA

<213> Artificial Sequence

20

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<220>

<221> modified_base

<222> (1)..(3)

<223> 2'-O-methyl modified

<220>

30

<221> misc_feature

<222> (1)..(4)

<223> Phosphorothioate linkage

<220>

<221> modified_base

<222> (29)..(40)

<223> 2'-O-methyl modified

<220>

<221> modified_base

<222> (69)..(100)

<223> 2'-O-methyl modified

<220>

40

<221> misc_feature

<222> (97)..(100)

<223> Phosphorothioate linkage

<400> 42

uuacagccac gucuacagca guuuuagagc uagaaaauagc aaguuaaaau aaggcuaguc 60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 43

50

<211> 4411

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 43

gggtcccga gtcggcgtcc agcggctctg cttgttcgtg tgtgtgtcgt tgcaggcctt	60	
attcggatcc gccaccatgg acaagaagta cagcatcgga ctggacatcg gaacaaacag	120	10
cgtcggatgg gcagtcatca cagacgaata caaggtcccg agcaagaagt tcaaggctct	180	
gggaaacaca gacagacaca gcatcaagaa gaacctgac ggagcactgc tggtcgacag	240	
cggagaaaca gcagaagcaa caagactgaa gagaacagca agaagaagat acacaagaag	300	
aaagaacaga atctgctacc tgcaggaaat cttcagcaac gaaatggcaa aggtcgacga	360	
cagcttcttc cacagactgg aagaaagctt cctggtcgaa gaagacaaga agcacgaaag	420	
acacccgatc ttcggaacaa tcgtcgacga agtcgcatac cacgaaaagt acccgacaat	480	
ctaccacctg agaaagaagc tggtcgacag cacagacaag gcagacctga gactgatcta	540	
cctggcactg gcacacatga tcaagttcag aggacacttc ctgatcgaag gagacctgaa	600	
cccggacaac agcgacgtcg acaagctgtt catccagctg gtccagacat acaaccagct	660	
gttcgaagaa aacccgatca acgcaagcgg agtcgacgca aaggcaatcc tgagcgcaag	720	20
actgagcaag agcagaagac tggaaaacct gatcgcacag ctgccgggag aaaagaagaa	780	
cggactgttc ggaaacctga tcgcaactgag cctgggactg acaccgaact tcaagagcaa	840	
cttcgacctg gcagaagacg caaagctgca gctgagcaag gacacatacg acgacgacct	900	
ggacaacctg ctggcacaga tcggagacca gtacgcagac ctgttccttg cagcaaagaa	960	
cctgagcgac gcaatcctgc tgagcgacat cctgagagtc aacacagaaa tcacaaaggc	1020	
accgctgagc gcaagcatga tcaagagata cgacgaacac caccaggacc tgacactgct	1080	
gaaggcactg gtcagacagc agctgccgga aaagtacaag gaaatcttct tcgaccagag	1140	
caagaacgga tacgcaggat acatcgacgg aggagcaagc caggaagaat tctacaagtt	1200	
catcaagccg atcctggaaa agatggacgg aacagaagaa ctgctgggtca agtgaacag	1260	
agaagacctg ctgagaaagc agagaacatt cgacaacgga agcatcccgc accagatcca	1320	30
cctgggagaa ctgcacgcaa tcctgagaag acaggaagac ttctaccctg tcctgaagga	1380	
caacagagaa aagatcgaaa agatcctgac attcagaatc ccgtactacg tcggaccgct	1440	
ggcaagagga aacagcagat tcgcatggat gacaagaaag agcgaagaaa caatcacacc	1500	
gtggaacttc gaagaagtcg tcgacaaggg agcaagcgca cagagcttca tcgaaagaat	1560	
gacaaacttc gacaagaacc tgccgaacga aaaggctctg ccgaagcaca gcctgctgta	1620	
cgaatacttc acagtctaca acgaactgac aaagggtcaag tacgtcacag aaggaatgag	1680	
aaagccggca ttcctgagcg gagaacagaa gaaggcaatc gtcgacctgc tgttcaagac	1740	
aaacagaaag gtcacagtca agcagctgaa ggaagactac ttcaagaaga tcgaatgctt	1800	
cgacagcgtc gaaatcagcg gagtcgaaga cagattcaac gcaagcctgg gaacatacca	1860	
cgacctgctg aagatcatca aggacaagga cttcctggac aacgaagaaa acgaagacat	1920	40
cctggaagac atcgtcctga cactgacact gttcgaagac agagaaatga tcgaagaaag	1980	
actgaagaca tacgcacacc tggtcgacga caaggtcatg aagcagctga agagaagaag	2040	
atacacagga tggggaagac tgagcagaaa gctgatcaac ggaatcagag acaagcagag	2100	
cggaaagaca atcctggact tcctgaagag cgacggattc gcaaacagaa acttcatgca	2160	
gctgatccac gacgacagcc tgacattcaa ggaagacatc cagaaggcac aggtcagcgg	2220	
acagggagac agcctgcacg aacacatcgc aaacctggca ggaagcccgg caatcaagaa	2280	
gggaatcctg cagacagtca aggtcgtcga cgaactggtc aagggtcatgg gaagacacaa	2340	
gccggaaaac atcgtcatcg aaatggcaag agaaaaccag acaacacaga agggacagaa	2400	
gaacagcaga gaaagaatga agagaatcga agaaggaatc aaggaactgg gaagccagat	2460	
cctgaaggaa caccgggtcg aaaacacaca gctgcagaac gaaaagctgt acctgtacta	2520	50

```

cctgcagaac ggaagagaca tgtacgtcga ccaggaactg gacatcaaca gactgagcga      2580
ctacgacgtc gaccacatcg tcccgcagag cttcctgaag gacgacagca tcgacaacaa      2640
ggtcctgaca agaagcgaca agaacagagg aaagagcgac aacgtcccga gcaagaagt      2700
cgtcaagaag atgaagaact actggagaca gctgtctgaac gcaaagctga tcacacagag      2760
aaagttcgac aacctgacaa aggagagagagg aggaggactg agcgaactgg acaaggcagg      2820
attcatcaag agacagctgg tcgaaacaag acagatcaca aagcacgtcg cacagatcct      2880
ggacagcaga atgaacacaa agtacgacga aaacgacaag ctgatcagag aagtcaaggt      2940
catcacactg aagagcaagc tggtcagcga cttcagaaag gacttccagt tctacaaggt      3000
cagagaaatc aacaactacc accacgcaca cgacgcatac ctgaacgcag tcgtcggaac      3060
agcactgatc aagaagtacc cgaagctgga aagcgaattc gtctacggag actacaaggt      3120
ctacgacgtc agaaagatga tcgcaaagag cgaacaggaa atcggaagg caacagcaaa      3180
gtacttcttc tacagcaaca tcatgaactt cttcaagaca gaaatcacac tggcaaacgg      3240
agaaatcaga aagagaccgc tgatcgaaac aaacggagaa acaggagaaa tcgtctggga      3300
caagggaaga gacttcgcaa cagtcagaaa ggtcctgagc atgccgcagg tcaacatcgt      3360
caagaagaca gaagtccaga caggaggatt cagcaaggaa agcatcctgc cgaagagaaa      3420
cagcgacaag ctgatcgcaa gaaagaagga ctgggacccg aagaagtacg gaggattcga      3480
cagcccgaca gtcgcataca gcgtcctggt cgtcgcaaag gtcgaaaagg gaaagagcaa      3540
gaagctgaag agcgtcaagg aactgctggg aatcacaatc atggaaagaa gcagcttcga      3600
aaagaacccg atcgacttcc tggaagcaaa gggatacaag gaagtcaaga aggacctgat      3660
catcaagctg ccgaagtaca gcctgttcga actggaaaac ggaagaaaga gaatgctggc      3720
aagcgagga gaactgcaga agggaaacga actggcactg ccgagcaagt acgtcaactt      3780
cctgtacctg gcaagccact acgaaaagct gaagggaagc ccggaagaca acgaacagaa      3840
gcagctgttc gtcgaacagc acaagcacta cctggacgaa atcatcgaac agatcagcga      3900
attcagcaag agagtcatcc tggcagacgc aaacctggac aaggctcctga gcgcatacaa      3960
caagcacaga gacaagccga tcagagaaca ggcagaaaac atcatccacc tgttcacact      4020
gacaaacctg ggagcaccgg cagcattcaa gtacttcgac acaacaatcg acagaaagag      4080
atacacaagc acaaaggaag tcctggacgc aacactgatc caccagagca tcacaggact      4140
gtacgaaaca agaatcgacc tgagccagct gggaggagac ggaggaggaa gcccgaaaga      4200
gaagagaaag gtctagctag ccatcacatt taaaagcatc tcagcctacc atgagaataa      4260
gagaaagaaa atgaagatca atagcttatt catctctttt tcttttctgt tgggtgaaag      4320
ccaacacctt gtctaaaaaa cataaatttc tttaatcatt ttgcctcttt tctctgtgct      4380
tcaattaata aaaaatggaa agaacctcga g                                4411

```

<210> 44

<211> 4405

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 44

```

gggtcccga gtcggcgtcc agcggctctg cttgttcgtg tgtgtgtcgt tgcaggcctt      60
attcggatcc atggacaaga agtacagcat cggactggac atcggaacaa acagcgtcgg      120
atgggcagtc atcacagacg aatacaaggt cccgagcaag aagttcaagg tcctgggaaa      180
cacagacaga cacagcatca agaagaacct gatcggagca ctgctgttcg acagcggaga      240
aacagcagaa gcaacaagac tgaagagaac agcaagaaga agatacaca gaagaaagaa      300
cagaatctgc tacctgcagg aaatcttcag caacgaaatg gcaaaggctg acgacagctt      360
cttccacaga ctggaagaaa gcttcctggg cgaagaagac aagaagcacg aaagacaccc      420
gatcttcgga aacatcgtcg acgaagtcgc ataccacgaa aagtaccga caatctacca      480

```

cctgagaaag aagctggctg acagcacaga caaggcagac ctgagactga tctacctggc 540
 actggcacac atgatcaagt tcagaggaca cttcctgata gaaggagacc tgaacccgga 600
 caacagcgac gtcgacaagc tggtcatcca gctggtccag acatacaacc agctgttcga 660
 agaaaacccg atcaacgcaa gcggagtcga cgcaaaggca atcctgagcg caagactgag 720
 caagagcaga agactggaaa acctgatcgc acagctgccg ggagaaaaga agaacggact 780
 gttcggaaac ctgatcgac tgagcctggg actgacaccg aacttcaaga gcaacttcga 840
 cctggcagaa gacgcaaagc tgcagctgag caaggacaca tacgacgacg acctggacaa 900
 cctgctggca cagatcggag accagtacgc agacctgttc ctggcagcaa agaacctgag 960
 cgacgcaatc ctgctgagcg acatcctgag agtcaacaca gaaatcacia aggcaccgct 1020
 gagcgcaagc atgatcaaga gatacgacga acaccaccag gacctgacac tgctgaaggc 1080
 actggtcaga cagcagctgc cggaaaagta caaggaaatc ttcttcgacc agagcaagaa 1140
 cggatacgca ggatacatcg acggaggagc aagccaggaa gaattctaca agttcatcaa 1200
 gccgatcctg gaaaagatgg acggaacaga agaactgctg gtcaagctga acagagaaga 1260
 cctgctgaga aagcagagaa cattcgacaa cggaagcatc ccgcaccaga tccacctggg 1320
 agaactgcac gcaatcctga gaagacagga agacttctac ccgttcctga aggacaacag 1380
 agaaaagatc gaaaagatcc tgacattcag aatcccgtac tacgtcggac cgctggcaag 1440
 aggaaacagc agattcgcat ggatgacaag aaagagcgaa gaaacaatca caccgtggaa 1500
 cttcgaagaa gtcgtcgaca agggagcaag cgacacagag ttcacgaaa gaatgacaaa 1560
 cttcgacaag aacctgccga acgaaaaggc cctgccgaag cacagcctgc tgtacgaata 1620
 cttcacagtc tacaacgaac tgacaaaggc caagtacgtc acagaaggaa tgagaaagcc 1680
 ggcatcctg agcggagaac agaagaaggc aatcgtcgac ctgctgttca agacaaacag 1740
 aaaggtcaca gtcaagcagc tgaaggaaaga ctacttcaag aagatcgaat gcttcgacag 1800
 cgtcgaaatc agcggagtcg aagacagatt caacgcaagc ctgggaacat accacgacct 1860
 gctgaagatc atcaaggaca aggacttctt ggacaacgaa gaaaacgaag acatcctgga 1920
 agacatcgtc ctgacactga cactgttcga agacagagaa atgatcgaag aaagactgaa 1980
 gacatacgca cacctgttcg acgacaaggc catgaagcag ctgaagagaa gaagatacac 2040
 aggatgggga agactgagca gaaagctgat caacggaatc agagacaagc agagcggaaa 2100
 gacaatcctg gacttctga agagcgacgg attcgcaaac agaaacttca tgcagctgat 2160
 ccacgacgac agcctgacat tcaaggaaga catccagaag gcacaggatc gcggacaggg 2220
 agacagcctg cacgaacaca tcgcaaactt ggacaggaagc ccggcaatca agaagggaat 2280
 cctgcagaca gtcaaggctg tcgacgaact ggtcaaggct atgggaagac acaagccgga 2340
 aaacatcgtc atcgaaatgg caagagaaaa ccagacaaca cagaagggac agaagaacag 2400
 cagagaaaga atgaagagaa tcgaagaagg aatcaaggaa ctgggaagcc agatcctgaa 2460
 ggaacacccg gtcgaaaaca cacagctgca gaacgaaaag ctgtacctgt actacctgca 2520
 gaacggaaga gacatgtacg tcgaccagga actggacatc aacagactga gcgactacga 2580
 cgtcgaccac atcgtcccg c agagcttctt gaaggacgac agcatcgaca acaaggctct 2640
 gacaagaagc gacaagaaca gaggaagag cgacaacgtc ccgagcgaag aagtcgtcaa 2700
 gaagatgaag aactactgga gacagctgct gaacgcaaag ctgatcacac agagaaagtt 2760
 cgacaacctg acaaaggcag agagaggagg actgagcgaa ctggacaagg caggattcat 2820
 caagagacag ctggtcgaaa caagacagat cacaagcac gtcgcacaga tctggacag 2880
 cagaatgaac acaaagtacg acgaaaacga caagctgatc agagaagtca aggtcatcac 2940
 actgaagagc aagctggctc gcgacttcag aaaggacttc cagttctaca aggtcagaga 3000
 aatcaacaac taccaccacg cacacgacgc atacctgaac gcagtcgtcg gaacagcact 3060
 gatcaagaag taccggaagc tggaaagcga attcgtctac ggagactaca aggtctacga 3120
 cgtcagaaag atgatcgcaa agagcgaaca ggaaatcgga aaggcaacag caaagtactt 3180
 cttctacagc aacatcatga acttcttcaa gacagaaatc acactggcaa acggagaaat 3240
 cagaaagaga ccgctgatcg aaacaaacgg agaaacagga gaaatcgtct gggacaaggg 3300
 aagagacttc gcaacagtca gaaaggctct gagcatgccg caggtaaca tcgtcaagaa 3360
 gacagaagtc cagacaggag gattcagcaa ggaaagcatc ctgccgaaga gaaacagcga 3420
 caagctgatc gcaagaaaga aggactggga cccgaagaag tacggaggat tcgacagccc 3480

10

20

30

40

50

gacagtcgca tacagcgtcc tggctgctgc aaaggtcgaa aagggaaaga gcaagaagct 3540
 gaagagcgtc aaggaactgc tgggaatcac aatcatggaa agaagcagct tcgaaaagaa 3600
 cccgatcgac ttcttggaag caaagggata caaggaagtc aagaaggacc tgatcatcaa 3660
 gctgccgaag tacagcctgt tcgaactgga aaacggaaga aagagaatgc tggcaagcgc 3720
 aggagaactg cagaagggaa acgaactggc actgccgagc aagtacgtca acttcctgta 3780
 cctggcaagc cactacgaaa agctgaaggg aagcccggaa gacaacgaac agaagcagct 3840
 gttcgtcgaa cagcacaagc actacctgga cgaaatcatc gaacagatca gcgaattcag 3900
 caagagagtc atcctggcag acgcaaacct ggacaaggtc ctgagcgcac acaacaagca 3960
 cagagacaag ccgatcagag aacaggcgaga aaacatcatc cacctgttca cactgacaaa 4020
 cctgggagca ccggcgagcat tcaagtactt cgacacaaca atcgacagaa agagatacac 4080
 aagcacaag gaagtcttgg acgcaacact gatccaccag agcatcacag gactgtacga 4140
 aacaagaatc gacctgagcc agctgggagg agacggagga ggaagcccga agaagaagag 4200
 aaaggtctag ctagccatca catttaaaag catctcagcc taccatgaga ataagagaaa 4260
 gaaaatgaag atcaatagct tattcatctc ttttctttt tcgttggtgt aaagccaaca 4320
 ccctgtctaa aaaacataaa tttctttaat cattttgcct cttttctctg tgcttcaatt 4380
 aataaaaaat ggaaagaacc tcgag 4405

10

<210> 45

<211> 4188

<212> DNA

20

<213> Artificial Sequence

<220>

<221> source

 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 45

atggataaga agtactcgat cgggctggat atcggaacta attccgtggg ttgggcagtg 60
 atcacggatg aatacaaagt gccgtccaag aagttcaagg tcctggggaa caccgataga 120
 cacagcatca agaagaatct catcgagacc ctgctgtttg actccggcga aaccgcagaa 180
 gcgacccggc tcaaacttac cgcgaggcga cgctacaccc ggcggaagaa tcgcatctgc 240
 tatctgcaag aaatctttt gaacgaaatg gcaaaggtgg acgacagctt cttccaccgc 300
 ctggaagaat ctttctggt ggaggaggac aagaagcatg aacggcatcc tatctttgga 360
 aacatcgtgg acgaagtggc gtaccacgaa aagtaccgga ccatctacca tctgcggaag 420
 aagttggttg actcaactga caaggccgac ctacagattga tctacttggc cctcgcccat 480
 atgatcaaat tccgcggaca cttcctgata gaaggcgatc tgaacctga taactccgac 540
 gtggataagc tgttcattca actggtgcag acctacaacc aactgttcga agaaaaccca 600
 atcaatgcaa gcggcgtcga tgccaaggcc atcctgtccg cccggctgtc gaagtcgagg 660
 cgcttcgaaa acctgatcgc acagctgccg ggagagaaga agaacggact ttcgggcaac 720
 ttgatcgtc tctactggg actcactccc aatttcaagt ccaattttga cctggccgag 780
 gacgcgaagc tgcaactctc aaaggacacc tacgacgacg acttgagcaa ttgctggca 840
 caaattggcg atcagtacgc ggatctgttc ctgcccgtc agaaccttc ggacgcaatc 900
 ttgctgtccg atatcttgc cgtgaacacc gaaataacca aagcgccgct tagcgccctg 960
 atgattaagc ggtacgacga gcatcaccag gatctacgc tgcataaagc gctcgtgaga 1020
 cagcaactgc ctgaaaagta caaggagatt ttcttcgacc agtccaagaa tgggtacgca 1080
 gggtacatcg atggaggcgc cagccaggaa gagttctata agttcatcaa gccaatctg 1140
 gaaaagatgg acggaaccga agaactgctg gtcaagctga acagggagga tctgctccgc 1200
 aaacagagaa ctttgacaa cggaagcatt ccacaccaga tccatctggg tgagctgcac 1260
 gccatcttgc ggcgccagga ggacttttac ccattcctca aggacaaccg ggaaaagatc 1320
 gagaaaattc tgacgttccg catcccgtat tacgtgggcc cactggcgcg cggcaattcg 1380
 cgcttcgcgt ggatgactag aaaatcagag gaaacatca ctccttgga tttcgaggaa 1440

30

40

50

gttgtggata agggagcttc ggcacaatcc ttcacgaac gaatgaccaa cttcgacaag 1500
 aatctcccaa acgagaaggt gcttcctaag cacagcctcc ttacgaata cttactgtc 1560
 tacaacgaac tgactaaagt gaaatacgtt actgaaggaa tgaggaagcc ggcctttctg 1620
 agcggagAAC agaagaaagc gattgtcgat ctgctgttca agaccaaccg caaggtgacc 1680
 gtcaagcagc ttaaagagga ctacttcaag aagatcgagt gtttcgactc agtggaatc 1740
 agcggagtgg aggacagatt caacgcttcg ctgggaacct atcatgatct cctgaagatc 1800
 atcaaggaca aggacttctt tgacaacgag gagaacgagg acatcctgga agatatcgtc 1860
 ctgaccttga cctttttcga ggatcgcgag atgatcgagg agaggcttaa gacctacgt 1920
 catctcttcg acgataaggt catgaaacaa ctcaagcgcc gccggtacac tggttggggc 1980
 cgcctctccc gcaagctgat caacggtatt cgcgataaac agagcggtaa aactatcctg 2040
 gatttctca aatcggatgg cttcgctaata cgtaacttca tgcagttgat ccacgacgac 2100
 agcctgacct ttaaggagga catccagaaa gcacaagtga gcggacaggg agactcactc 2160
 catgaacaca tcgcgaatct ggccggttcg ccggcgatta agaagggaat cctgcaaat 2220
 gtgaaggtgg tggacgagct ggtgaaggct atgggacggc acaaaccgga gaatatcgtg 2280
 attgaaatgg cccgagaaaa ccagactacc cagaagggcc agaagaactc ccgcgaaagg 2340
 atgaagcggA tcgaagaagg aatcaaggag ctgggcagcc agatcctgaa agagcaccgc 2400
 gtggaaaaca cgcagctgca gaacgagaag ctctacctgt actatttga aaatggacgg 2460
 gacatgtacg tggaccaaga gctggacatc aatcggttgt ctgattacga cgtggaccac 2520
 atcgttccac agtcttttct gaaggatgac tccatcgata acaagggtgt gactcgcagc 2580
 gacaagaaca gaggggaagt agataatgtg ccatcgagg aggtcgtgaa gaagatgaag 2640
 aattactggc ggcagctcct gaatgcgaag ctgattacc agagaaagt tgacaatctc 2700
 actaaagccg agcgcggcgg actctcagag ctggataagg ctggattcat caaacggcag 2760
 ctggtcgaga ctgcgcagat taccaagcac gtggcgaga tcctggactc ccgatgaac 2820
 actaaatacg acgagaacga taagctcatc cggaagtga aggtgattac cctgaaaagc 2880
 aaacttgtgt cggactttcg gaaggacttt cagttttaca aagtgagaga aatcaacaac 2940
 taccatcacg cgcattgacg atacctcaac gctgtggtcg gcaccgccct gatcaagaag 3000
 taccctaaac ttgaatcgga gtttgtgtac ggagactaca aggtctacga cgtgaggaag 3060
 atgatagcca agtccgaaca ggaaatcggg aaagcaactg cgaaatactt ctttactca 3120
 aacatcatga acttcttcaa gactgaaatt acgttgcca atggagaaat caggaagagg 3180
 ccactgatcg aaactaacgg agaaacgggc gaaatcgtgt gggacaaggg cagggacttc 3240
 gcaactgttc gcaaagtgt ctctatgccg caagtcaata ttgtgaagaa aaccgaagtg 3300
 caaacggcg gattttcaaa ggaatcgatc ctcccaaaga gaaatagcga caagctcatt 3360
 gcacgcaaga aagactggga cccgaagaag tacggaggat tcgattcgcc gactgtcgca 3420
 tactccgtcc tcgtggtggc caaggtggag aagggaaga gcaagaagct caaatccgtc 3480
 aaagagctgc tggggattac catcatggaa cgatcctcgt tcgagaagaa cccgattgat 3540
 ttcttgagg cgaagggtta caaggaggtg aagaaggatc tgatcatcaa actgcccaag 3600
 tactactgt tcgaactgga aaatggtcgg aagcgcgtg tggcttcggc cggagaactc 3660
 cagaaaggaa atgagctggc cttgcctagc aagtacgtca acttctcta tcttgcctc 3720
 cactacgaga aactcaaagg gtcaccggaa gataacgaac agaagcagct ttcgtggag 3780
 cagcacaagc attatctgga tgaaatcatc gaacaaatct ccgagttttc aaagcgcgtg 3840
 atctcgcgc acgccaacct cgacaaagtc ctgtcggcct acaataagca tagagataag 3900
 ccgatcagag aacaggccga gaacattatc cacttgttca cctgactaa cctgggagct 3960
 ccagccgcct tcaagtactt cgatactact atcgaccgca aaagatacac gtccaccaag 4020
 gaagttctgg acgcgaccct gatccacaa agcatcactg gactctacga aactaggatc 4080
 gatctgtcg agctgggtgg cgatgggtgg ggtggatcct acccatacga cgtgcctgac 4140
 tacgcctccg gaggtggtgg cccaagaag aaacggaagg tgtgatag 4188

<210> 46
 <211> 4459
 <212> DNA

10

20

30

40

50

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 46

```

gggtcccga gtcggcgtcc agcggctctg cttgttcgtg tgtgtgtcgt tgcaggcctt      60
attcggatct gccaccatgg ataagaagta ctcgatcggg ctggatatcg gaactaatc      120
cgtgggttgg gcagtgatca cgatgaata caaagtgccg tccaagaagt tcaaggtcct      180
ggggaacacc gatagacaca gcatcaagaa gaatctcatc ggagccctgc tgtttgactc      240
cggcgaaacc gcagaagcga cccggctcaa acgtaccgcg aggcgacgct acacccggcg      300
gaagaatcgc atctgctatc tgcaagaaat cttttcgaac gaaatggcaa aggtggacga      360
cagcttcttc caccgcctgg aagaatcttt cctgggtggag gaggacaaga agcatgaacg      420
gcatcctatc tttggaacaa tcgtggacga agtggcgtag cacgaaaagt acccgacat      480
ctaccatctg cggaagaagt tggttgactc aactgacaag gccgacctca gattgatcta      540
cttgccctc gcccatatga tcaaattccg cggacacttc ctgatcgaag gcgatctgaa      600
ccctgataac tccgacgtgg ataagctgtt cattcaactg gtgcagacct acaaccaact      660
gttcgaagaa aacccaatca atgccagcgg cgtcgatgcc aaggccatcc tgtccgcccg      720
gctgtcgaag tcgcgggcgc tcgaaaacct gatcgacag ctgccgggag agaagaagaa      780
cggacttttc ggcaacttga tcgctctctc actgggactc actcccaatt tcaagtcaa      840
ttttgacctg gccgaggacg cgaagctgca actctcaaag gacacctacg acgacgactt      900
ggacaatttg ctggcacaaa ttggcgatca gtacgcggat ctgttccttg ccgctaagaa      960
cctttcggac gcaatcttgc tgtccgatat cctgcgcgtg aacaccgaaa taaccaaacg      1020
gccgcttagc gcctcgatga ttaagcggta cgacgagcat caccaggatc tcacgctgct      1080
caaagcgctc gtgagacagc aactgcctga aaagtacaag gagattttct tcgaccagtc      1140
caagaatggg tacgcagggt acatcgatgg aggcgccagc caggaagagt tctataagtt      1200
catcaagcca atcctggaaa agatggacgg aaccgaagaa ctgctggtca agctgaacag      1260
ggaggatctg tcccgcaaac agagaacctt tgacaacgga agcattccac accagatcca      1320
tctgggtgag ctgcacgcca tcttgccgcg ccaggaggac ttttacctat tctcaagga      1380
caaccgggaa aagatcgaga aaattctgac gttccgcata ccgtattacg tgggcccact      1440
ggcgcgcggc aattcgcgct tcgctggtg gactagaaaa tcagaggaaa ccatcactcc      1500
ttggaatttc gaggaagtgt tggataaggg agcttcggca caatccttca tcgaacgaat      1560
gaccaacttc gacaagaatc tcccaaacga gaaggtgctt cctaagcaca gcctccttta      1620
cgaatacttc actgtctaca acgaactgac taaagtgaag tacgttactg aaggaatgag      1680
gaagccggcc tttctgagcg gagaacagaa gaaagcgatt gtcgatctgc tgttcaagac      1740
caaccgcaag gtgaccgtca agcagcttaa agaggactac ttcaagaaga tcgagtgttt      1800
cgactcagtg gaaatcagcg gagtggagga cagattcaac gcttcgctgg gaacctatca      1860
tgatctcctg aagatcatca aggacaagga cttccttgac aacgaggaga acgaggacat      1920
cctggaagat atcgtcctga ccttgaccct tttcgaggat cgcgagatga tcgaggagag      1980
gcttaagacc tacgtcatc tcttcgacga taaggtcatg aaacaactca agcgccgccg      2040
gtacactggt tggggccgcc tctcccga gctgatcaac ggtattcgcg ataaacagag      2100
cggtaaaact atcctggatt tctcaaact ggatggcttc gctaactgta acttcatgca      2160
gttgatccac gacgacagcc tgacctttaa ggaggacatc cagaaagcac aagtgagcgg      2220
acaggagagac tactccatg aacacatcgc gaatctggcc ggttcgccgg cgattaagaa      2280
gggaatcctg caaactgtga aggtggtgga cgagctggtg aaggtcatgg gacggcacia      2340
accggagaat atcgtgattg aatggcccg agaaaaccag actaccaga agggccagaa      2400
gaactccgcg gaaaggatga agcggatcga agaaggaatc aaggagctgg gcagccagat      2460
cctgaaagag caccgggtgg aaaacacgca gctgcagaac gagaagctct acctgtacta      2520
tttgcaaaat ggacgggaca tgtacgtgga ccaagagctg gacatcaatc ggttgtctga      2580
ttacgacgtg gaccacatcg ttccacagtc ctttctgaag gatgactcca tcgataacaa      2640

```

10

20

30

40

50

ggtgttgact cgcagcgaca agaacagagg gaagtcagat aatgtgccat cggaggaggt 2700
 cgtgaagaag atgaagaatt actggcggca gctcctgaat gcgaagctga ttaccagag 2760
 aaagttgac aatctcacta aagccgagcg cggcggactc tcagagctgg ataaggctgg 2820
 attcatcaaa cggcagctgg tcgagactcg gcagattacc aagcacgtgg cgcagatcct 2880
 ggactcccg c atgaacacta aatacgacga gaacgataag ctcatccggg aagtgaaggt 2940
 gattaccctg aaaagcaaac ttgtgtcgga ctttcggaag gactttcagt ttacaaagt 3000
 gagagaaatc aacaactacc atcacgcgca tgacgcatac ctcaacgctg tggtcggcac 3060
 cgccctgac aagaagtacc cttaaactga atcggagttt gtgtacggag actacaaggt 3120
 ctacgacgtg aggaagatga tagccaagtc cgaacaggaa atcgggaaag caactgcgaa 3180
 atacttcttt tactcaaaca tcatgaactt cttcaagact gaaattacgc tggccaatgg 3240
 agaaatcagg aagaggccac tgatcgaaac taacggagaa acgggcgaaa tcgtgtggga 3300
 caagggcagg gacttcgcaa ctgttcgcaa agtgctctct atgccgcaag tcaatattgt 3360
 gaagaaaacc gaagtgcaaa ccggcggatt ttcaaaggaa tcgatcctcc caaagagaaa 3420
 tagcgacaag ctcatcgac gcaagaaaga ctgggacccg aagaagtacg gaggattcga 3480
 ttcccgact gtcgcatact ccgtcctcgt ggtggccaag gtggagaagg gaaagagcaa 3540
 gaagctcaaa tccgtcaaag agctgctggg gattaccatc atggaacgat cctcgttcga 3600
 gaagaacccg attgatttcc tggaggcgaa gggttacaag gaggtgaaga aggatctgat 3660
 catcaaactg cccaagtact cactgttcga actggaaaat ggtcggaagc gcatgctggc 3720
 ttccggccgga gaactccaga aaggaaatga gctggccttg cctagcaagt acgtcaactt 3780
 cctctatctt gcttcgcact acgagaaact caaagggtca ccggaagata acgaacagaa 3840
 gcagcttttc gtggagcagc acaagcatta tctggatgaa atcatcgaac aaatctccga 3900
 gttttcaaag cgctgatcc tcgccgacgc caacctcgac aaagtcctgt cggcctacaa 3960
 taagcataga gataagccga tcagagaaca ggccgagaac attatccact tgttcaccct 4020
 gactaacctg ggagctccag ccgccttcaa gtacttcgat actactatcg accgcaaaag 4080
 atacacgtcc accaaggaag ttctggacgc gacctgatc caccaaagca tcaactggact 4140
 ctacgaaact aggatcgatc tctcgcagct ggggtggcgat ggtggcgggtg gatcctaccc 4200
 atacgacgtg cctgactacg cctccggagg tgggtggccc aagaagaaac ggaaggtgtg 4260
 atagctagcc atcacattta aaagcatctc agcctacat gagaataaga gaaagaaaat 4320
 gaagatcaat agcttattca tctctttttc ttttcgttg gtgtaaagcc aacaccctgt 4380
 ctaaaaaaca taaatttctt taatcatttt gcctcttttc tctgtgcttc aattaataaa 4440
 aaatggaaag aacctcgag 4459

10

20

30

<210> 47

<211> 4453

<212> DNA

<213> Artificial Sequence

<220>

<221> source

 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

40

<400> 47

gggctccgca gtcggcgtcc agcggctctg cttgttcgtg tgtgtgtcgt tgcaggcctt 60
 attcggatct atggataaga agtactcgat cgggctggat atcggaacta attcgtggg 120
 ttgggcagtg atcacggatg aatacaaagt gccgtccaag aagttcaagg tcctggggaa 180
 caccgataga cacagcatca agaagaatct catcgagacc ctgctgtttg actccggcga 240
 aaccgcagaa gcgacccggc tcaaacgtac cgcgaggcga cgctacaccc ggcggaagaa 300
 tcgcatctgc tatctgcaag aaatcttttc gaacgaaatg gcaaagggtg acgacagctt 360
 cttccaccgc ctggaagaat ctttccttgt ggaggaggac aagaagcatg aacggcatcc 420
 tatctttgga aacatcgtgg acgaagtggc gtaccacgaa aagtaccga ccatctacca 480
 tctgcggaag aagttggtg actcaactga caaggccgac ctacattga tctacttggc 540

50

cctcgcccat atgatcaaat tccgcgga cttcctgac gaaggcgatc tgaaccctga	600	
taactccgac gtggataagc tgttcattca actggtgcag acctacaacc aactgttcga	660	
agaaaaccca atcaatgcca gcggcgctga tgccaaggcc atcctgtccg cccggctgtc	720	
gaagtcgcgg cgccctgaaa acctgatcgc acagctgccg ggagagaaga agaacggact	780	
tttcggcaac ttgatcgctc tctcactggg actcactccc aatttcaagt ccaattttga	840	
cctggccgag gacgcgaagc tgcaactctc aaaggacacc tacgacgacg acttgacaa	900	
tttgctggca caaattggcg atcagtacgc ggatctgttc ctggccgcta agaacccttc	960	
ggacgcaatc ttgctgtccg atatcctgcg cgtgaacacc gaaataacca aagcgccgct	1020	
tagcgccctc atgattaagc ggtacgacga gcatcaccag gatctcacgc tgctcaaagc	1080	
gctcgtgaga cagcaactgc ctgaaaagta caaggagatt ttcttcgacc agtccaagaa	1140	10
tgggtacgca gggtacatcg atggaggcgc cagccaggaa gagttctata agttcatcaa	1200	
gccaatcctg gaaaagatgg acggaaccga agaactgctg gtcaagctga acagggagga	1260	
tctgctccgc aaacagagaa cctttgacaa cggaagcatt ccacaccaga tccatctggg	1320	
tgagctgcac gccatcttgc ggcgccagga ggacttttac ccattcctca aggacaaccg	1380	
ggaaaagatc gagaaaattc tgacgttccg catcccgtat tacgtgggcc cactggcgcg	1440	
cggcaattcg cgcttcgcgt ggatgactag aaaatcagag gaaaccatca ctccttgga	1500	
tttcgaggaa gttgtggata agggagcttc ggcacaatcc ttcacgaac gaatgaccaa	1560	
cttcgacaag aatctcccaa acgagaaggt gtttcctaag cacagcctcc ttacgaata	1620	
cttcactgtc tacaacgaac tgactaaagt gaaatacgtt actgaaggaa tgaggaagcc	1680	
ggcctttctg agcggagaac agaagaaagc gattgtcgat ctgctgttca agaccaaccg	1740	20
caaggtgacc gtcaagcagc ttaaagagga ctacttcaag aagatcgagt gtttcgactc	1800	
agtggaaatc agcggagtgg aggacagatt caacgcttcg ctgggaacct atcatgatct	1860	
cctgaagatc atcaaggaca aggacttctt tgacaacgag gagaacgagg acatcctgga	1920	
agatatcgtc ctgaccttga ccttttcga ggatcgcgag atgatcgagg agaggcttaa	1980	
gacctacgct catctcttcg acgataaggt catgaaacaa ctcaagcgcc gccggtacac	2040	
tggttggggc cgctctctcc gcaagctgat caacgggtatt cgcgataaac agagcggtaa	2100	
aactatcctg gatttcctca aatcggatgg cttcgcta atcgtaacttca tgcagttgat	2160	
ccacgacgac agcctgacct ttaaggagga catccagaaa gcacaagtga gcggacaggg	2220	
agactcactc catgaacaca tcgcgaatct ggccgggttc cggcgatta agaagggaat	2280	
cctgcaaaact gtgaagggtg tggacgagct ggtgaaggct atgggacggc acaaaccgga	2340	30
gaatatcgtg attgaaatgg cccgagaaaa ccagactacc cagaagggcc agaagaactc	2400	
ccgcgaaagg atgaagcggg tcgaagaagg aatcaaggag ctgggcagcc agatcctgaa	2460	
agagcacccg gtggaaaaca cgcagctgca gaacgagaag ctctacctgt actatttgca	2520	
aaatggacgg gacatgtacg tggaccaaga gctggacatc aatcggttgt ctgattacga	2580	
cgtggaccac atcgttcac agtcctttct gaaggatgac tccatcgata acaaggtgtt	2640	
gactcgcagc gacaagaaca gagggaaagc agataatgtg ccacggagg aggtcgtgaa	2700	
gaagatgaag aattactggc ggcagctcct gaatcggaag ctgattacc agagaaagtt	2760	
tgacaatctc actaaagccg agcgcggcgg actctcagag ctggataagg ctggattcat	2820	
caaacggcag ctggtcgaga ctcggcagat taccaagcac gtggcgaga tcctggactc	2880	
ccgcatgaac actaaatacg acgagaacga taagctcatt cggaagtga aggtgattac	2940	40
cctgaaaagc aaacttgtgt cggactttcg gaaggacttt cagttttaca aagtgagaga	3000	
aatcaacaac taccatcacg cgcatgacgc atacctcaac gctgtggctg gcaccgccct	3060	
gatcaagaag taccctaaac ttgaatcgga gtttgtgtac ggagactaca aggtctacga	3120	
cgtgaggaag atgatagcca agtccgaaca ggaaatcggg aaagcaactg cgaaatactt	3180	
cttttactca aacatcatga acttcttcaa gactgaaatt acgctggcca atggagaaat	3240	
caggaagagg ccactgatcg aaactaacgg agaaacgggc gaaatcgtgt gggacaaggg	3300	
cagggacttc gcaactgttc gcaaagtgt ctctatgccg caagtcaata ttgtgaagaa	3360	
aaccgaagtg caaacggcgc gattttcaaa ggaatcgatc ctcccaaaga gaaatagcga	3420	
caagctcatt gcacgcaaga aagactggga cccgaagaag tacggaggat tcgattcgcc	3480	
gactgtcgca tactccgtcc tcgtggtggc caaggtggag aagggaagaa gcaagaagct	3540	50

```

caaatccgtc aaagagctgc tggggattac catcatggaa cgatcctcgt tcgagaagaa 3600
cccgattgat ttcttgagg cgaagggtta caaggagggtg aagaaggatc tgatcatcaa 3660
actgcccgaag tactcactgt tcgaactgga aaatggtcgg aagcgcgtgc tggcttcggc 3720
cggagaactc cagaaaggaa atgagctggc cttgcctagc aagtacgtca acttcctcta 3780
tcttgcttcg cactacgaga aactcaaagg gtcaccggaa gataacgaac agaagcagct 3840
tttcgtggag cagcacaagc attatctgga tgaaatcatc gaacaaatct ccgagttttc 3900
aaagcgcgtg atcctcgccg acgccaacct cgacaaagtc ctgtcggcct acaataagca 3960
tagagataag ccgatcagag aacaggccga gaacattatc cacttgttca ccctgactaa 4020
cctgggagct ccagccgcct tcaagtactt cgatactact atcgaccgca aaagatacac 4080
gtccaccaag gaagttctgg acgcgaccct gatccacca agcatcactg gactctacga 4140
aactaggatc gatctgtcgc agctgggtgg cgatgggtgg ggtggatcct acccatacga 4200
cgtgcctgac tacgcctccg gaggtgggtgg cccaagaag aaacggaagg tgtgatagct 4260
agccatcaca tttaaaagca tctcagccta ccatgagaat aagagaaaga aaatgaagat 4320
caatagctta ttcattcttt tttctttttc gttgggtgta agccaacacc ctgtctaaaa 4380
aacataaatt tctttaatca tttgcctct tttctctgtg cttaattaa taaaaaatgg 4440
aaagaacctc gag 4453

```

10

<210> 48

<211> 4403

<212> DNA

20

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 48

```

gggtcccgca gtcggcgtcc agcggctctg cttgttcgtg tgtgtgtcgt tgcaggcctt 60
attcggatcc atgcctaaga aaaagcggaa ggtcgacggg gataagaagt actcaatcgg 120
gctggatata ggaactaatt ccgtgggttg ggcagtgatc acggatgaat acaaagtgcc 180
gtccaagaag ttcaaggctc tggggaacac cgatagacac agcatcaaga aaaatctcat 240
cggagccctg ctgtttgact ccggcgaaac cgcagaagcg acccggtcga aacgtaccgc 300
gaggcgacgc tacaccgggc ggaagaatcg catctgctat ctgcaagaga tcttttcgaa 360
cgaaatggca aaggctgcag acagcttctt ccaccgcctg gaagaatctt tcttggtgga 420
ggaggacaag aagcatgaac ggcattctat ctttggaac atcgtcgacg aagtggcgta 480
ccacgaaaag taccgacca tctaccatct gcggaagaag ttggttgact caactgacaa 540
ggccgacctc agattgatct acttggccct cgcccatatg atcaaatcc gcggacactt 600
cctgatcgaa ggcatctga accctgataa ctccgacgtg gataagcttt tcattcaact 660
ggtgcagacc tacaaccaac tgctcgaaga aaacccaatc aatgctagcg gcgtcgatgc 720
caaggccatc ctgtccgccc ggctgtcgaa gtcgcggcgc ctgaaaacc tgatcgaca 780
gctgccggga gagaaaaaga acggactttt cggaacttg atcgctctct cactgggact 840
cactcccaat ttcaagtcca attttgacct ggccgaggac gcgaagctgc aactctcaaa 900
ggacacctac gacgacgact tggacaattt gctggcaca attggcgatc agtacgcgga 960
tctgttcctt gccgctaaga acctttcgga cgcaatcttg ctgtccgata tctgcgcgt 1020
gaacaccgaa ataaccaaag cgccgcttag cgctcgtatg attaagcggg acgacgagca 1080
tcaccaggat ctacgctgc tcaaagcgct cgtgagacag caactgcctg aaaagtacaa 1140
ggagatcttc ttcgaccagt ccaagaatgg gtacgcaggg tacatcgatg gaggcgctag 1200
ccaggaagag ttctataagt tcatcaagcc aatcctggaa aagatggacg gaaccgaaga 1260
actgctggtc aagctgaaca gggaggatct gctccggaaa cagagaacct ttgacaacgg 1320
atccattccc caccgatcc atctgggtga gctgcacgcc atcttgccgc gccaggagga 1380
cttttaccba ttcctcaagg acaaccggga aaagatcgag aaaattctga cgttccgcat 1440

```

30

40

50

cccgtattac gtgggcccac tggcgcgcg cg caattcgcg ttcgctgga tgactagaaa 1500
 atcagaggaa accatcactc cttggaatit cgaggaagtt gtggataagg gagcttcggc 1560
 acaaagcttc atcgaacgaa tgaccaactt cgacaagaat ctcccaaacg agaaggtgct 1620
 tcctaagcac agcctccttt acgaatactt cactgtctac aacgaactga ctaaagtga 1680
 atacgttact gaaggaatga ggaagccggc ctttctgtcc ggagaacaga agaaagcaat 1740
 tgtcgatctg ctgttcaaga ccaaccgcaa ggtgaccgtc aagcagctta aagaggacta 1800
 cttcaagaag atcgagtgtt tcgactcagt ggaaatcagc ggggtggagg acagattcaa 1860
 cgcttcgctg ggaacctatc atgatctcct gaagatcacc aaggacaagg acttccttga 1920
 caacgaggag aacgaggaca tcctggaaga tatcgtcctg accttgaccc ttttcgagga 1980
 tcgcgagatg atcgaggaga ggcttaagac ctacgctcat ctcttcgacg ataaggtcat 2040
 gaaacaactc aagcgccgcc ggtacacttg ttggggccgc ctctcccgca agctgatcaa 2100
 cgggtattcg gataaacaga gcggtaaaac taccctggat ttcctcaa at cggtatggctt 2160
 cgctaactgt aacttcatgc aattgatcca cgacgacagc ctgacctta aggaggacat 2220
 ccaaaaagca caagtgtccg gacagggaga ctactccat gaacacatcg cgaatctggc 2280
 cggttcgccg gcgattaaga agggaattct gcaaactgtg aaggtggctg acgagctggt 2340
 gaaggtcatg ggacggcaca aaccggagaa tatcgtgatt gaaatggccc gaaaaacca 2400
 gactaccag aagggccaga aaaactcccg cgaaaggatg aagcggatcg aagaaggaat 2460
 caaggagctg ggcagccaga tcctgaaaga gcacccgggtg gaaaacacgc agctgcagaa 2520
 cgagaagctc tacctgtact atttgcaaaa tggacgggac atgtacgtgg accaagagct 2580
 ggacatcaat cggttgtctg attacgacgt ggaccacatc gttccacagt ctttctgaa 2640
 ggatgactcg atcgataaca aggtgttgac tcgacgacg aagaacagag ggaagtcaga 2700
 taatgtgcca tcggaggagg tcgtgaagaa gatgaagaat tactggcggc agctcctgaa 2760
 tgcgaagctg attaccaga gaaagtttga caatctcact aaagccgagc gcggcggact 2820
 ctgagagctg gataaggctg gattcatcaa acggcagctg gtcgagactc ggcagattac 2880
 caagcacgtg gcgcagatct tggactcccg catgaacact aaatacgacg agaacgataa 2940
 gctcatccgg gaagtgaagg tgattaccct gaaaagcaaa cttgtgtcgg actttcggaa 3000
 ggactttcag ttttacaag tgagagaaat caacaactac catcacgcg atgacgcata 3060
 cctcaacgct gtggctcggt cgcacctgat caaaaagtac cctaaacttg aatcgagatt 3120
 tgtgtacgga gactacaagg tctacgacgt gaggaagatg atagccaagt ccgaacagga 3180
 aatcgggaaa gcaactgcga aatacttctt ttactcaaac atcatgaact tttcaagac 3240
 tgaaattacg ctggccaatg gagaaatcag gaagaggcca ctgatcgaaa ctaacggaga 3300
 aacgggcaaa atcgtgtggg acaagggcag ggacttcgca actgttcgca aagtgtctc 3360
 tatgccgcaa gtcaatattg tgaagaaaac cgaagtgcaa accggcggat tttcaaagga 3420
 atcgatctc caaagagaa atagcgacaa gctcattgca cgcaagaaag actgggaccc 3480
 gaagaagtac ggaggattcg attcgccgac tgtcgcatat tccgtcctcg tgggtggcaa 3540
 ggtggagaag ggaagagca aaaagctcaa atccgtcaa gagctgtctg ggattaccat 3600
 catggaacga tcctcgttcg agaagaaccc gattgatttc ctgaggcga agggttacaa 3660
 ggaggtgaag aaggatctga tcatcaaact cccaagtac tactgttcg aactggaaaa 3720
 tggctggaag cgcgtgtggt cttcgccggg agaactcaa aaaggaaatg agctggcctt 3780
 gcctagcaag tacgtcaact tcctctatct tgcttcgac tacgaaaaac tcaaagggtc 3840
 accggaagat aacgaacaga agcagctttt cgtggagcag cacaagcatt atctggatga 3900
 aatcatcgaa caaatctccg agttttcaa gcgcgtgatc ctgcccagc ccaacctga 3960
 caaagtctg tcggcctaca ataagcatag agataagccg atcagagaa aggccgagaa 4020
 cattatccac ttgttcccc tgactaacct gggagcccca gccgccttca agtacttcga 4080
 tactactatc gatcgaaaa gatacacgac caccaaggaa gttctggacg cgaccctgat 4140
 ccaccaaagc atcactggac tctacgaaac taggatcgat ctgtcgacg tgggtggcga 4200
 ttgatagtct agccatcaca tttaaaagca tctcagccta ccatgagaat aagagaaaga 4260
 aatgaagat caatagctta ttcattcttt tttcttttc gttggtgtaa agccaacacc 4320
 ctgtctaaaa aacataaatt tctttaatca ttttgctct tttctctgt cttcaattaa 4380
 taaaaaatgg aaagaacctc gag 4403

10

20

30

40

50

<210> 49
 <211> 4409
 <212> DNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"
 <400> 49

```

gggtcccgcga gtcggcgctcc agcggctctg cttgttcgtg tgtgtgtcgt tgcaggcctt      60
attcggatcc gccaccatgc ctaagaaaaa gcggaaggtc gacggggata agaagtactc      120
aatcgggctg gatatcggaa ctaattccgt gggttgggca gtgatcacgg atgaatacaa      180
agtgccgtcc aagaagttca aggtcctggg gaacaccgat agacacagca tcaagaaaaa      240
tctcatcgga gccctgctgt ttgactccgg cgaaaccgca gaagcgaccc ggctcaaacg      300
taccgcgagg cgacgtaca cccggcgga gaatcgcacg tgctatctgc aagagatctt      360
ttcgaacgaa atggcaaagg tcgacgacag cttcttcac cgcctggaag aatctttcct      420
ggtaggaggag gacaagaagc atgaacggca tcctatcttt ggaaacatcg tcgacgaagt      480
ggcgtaccac gaaaagtacc cgaccatcta ccatctgcgg aagaagttgg ttgactcaac      540
tgacaaggcc gacctcagat tgatctactt ggccctcgcc catatgatca aattccgcgg      600
acacttcctg atcgaaggcg atctgaaccc tgataactcc gacgtggata agcttttcat      660
tcaactggtg cagacctaca accaactgtt cgaagaaaac ccaatcaatg ctagcggcgt      720
cgatgccaaag gccatcctgt ccgcccggct gtcgaagtcg cggcgccctcg aaaacctgat      780
cgcacagctg ccgggagaga aaaagaacgg acttttcggc aacttgatcg ctctctcact      840
gggactcact cccaatttca agtccaattt tgacctggcc gaggacgca agctgcaact      900
ctcaaaggac acctacgacg acgacttgga caatttgctg gcacaaattg gcgatcagta      960
cgcggatctg ttccttgccg ctaagaacct ttcggacgca atcttgctgt ccgatatcct     1020
gcgcgtgaac accgaaataa ccaaagcgcc gcttagcgcc tcgatgatta agcgggtacga     1080
cgagcatcac caggatctca cgctgtcaa agcgcctcgt agacagcaac tgcctgaaaa     1140
gtacaaggag atcttcttcg accagtccaa gaatgggtac gcagggtaca tcgatggagg     1200
cgctagccag gaagagttct ataagttcat caagccaatc ctggaaaaga tggacggaac     1260
cgaagaactg ctggtcaagc tgaacaggga ggatctgctc cggaacaga gaaccttga     1320
caacggatcc attccccacc agatccatct gggtagctg cacgccatct tgcggcgcca     1380
ggaggacttt tacccattcc tcaaggacaa ccgggaaaag atcgagaaaa ttctgacgtt     1440
ccgcatcccg tattacgtgg gccactggc gcgcggcaat tcgcgcttcg cgtggatgac     1500
tagaaaatca gaggaaccca tcactccttg gaatttcgag gaagttgttg ataagggagc     1560
ttcggcacaa agcttcacg aacgaatgac caacttcgac aagaatctcc caaacgagaa     1620
ggtgcttcct aagcacagcc tcctttacga atacttcaact gtctacaacg aactgactaa     1680
agtgaataac gttactgaag gaatgaggaa gccggccttt ctgtccggag aacagaagaa     1740
agcaattgtc gatctgctgt tcaagaccaa ccgcaagggt accgtcaagc agcttaaaga     1800
ggactacttc aagaagatcg agtgtttcga ctcagtggaa atcagcgggg tggaggacag     1860
attcaacgct tcgctgggaa cctatcatga tctcctgaag atcatcaagg acaaggactt     1920
ccttgacaac gaggagaacg aggacatcct ggaagatatc gtcctgacct tgaccctttt     1980
cgaggatcgc gagatgatcg aggagaggct taagacctac gtcacatctt tcgacgataa     2040
ggcatgaaa caactcaagc gccgccggt cactgggttg ggccgcctct cccgcaagct     2100
gatcaacggt attcgcgata aacagagcgg taaaactatc ctggatttcc tcaaatacga     2160
tggcttcgct aatcgtaact tcatgcaatt gatccacgac gacagcctga cctttaagga     2220
ggacatccaa aaagcacaag tgtccggaca gggagactca ctccatgaac acatcgcgaa     2280
tctggccggt tcgccggcga ttaagaaggg aattctgcaa actgtgaagg tggtcgacga     2340
gctgggtgaag gtcatgggac ggcacaaacc ggagaatatc gtgattgaaa tggccccgaga     2400

```

10

20

30

40

50

```

aaaccagact acccagaagg gccagaaaaa ctcccgcgaa aggatgaagc ggatcgaaga 2460
aggaatcaag gagctgggca gccagatcct gaaagagcac cgggtggaaa acacgcagct 2520
gcagaacgag aagctctacc tgtactatit gcaaaatgga cgggacatgt acgtggacca 2580
agagctggac atcaatcggg tgtctgatta cgacgtggac cacatcggtt cacagtcctt 2640
tctgaaggat gactcgatcg ataacaaggt gttgactcgc agcgacaaga acagagggaa 2700
gtcagataat gtgccatcgg aggaggtcgt gaagaagatg aagaattact ggcggcagct 2760
cctgaatgcg aagctgatta cccagagaaa gtttgacaat ctactaaag ccgagcgcgg 2820
cggactctca gagctggata aggtctggatt catcaaacgg cagctggtcg agactcggca 2880
gattaccaag cacgtggcgc agatcttggg ctcccgcgac aacactaaat acgacgagaa 2940
cgataagctc atccgggaag tgaagggtgat tacctgaaa agcaaacttg tgtcggactt 3000
tcggaaggac ttctagtttt acaaagttag agaaatcaac aactaccatc acgcgcgata 3060
cgcatacctc aacgctgtgg tcggtaccgc cctgatcaaa aagtacccta aacttgaatc 3120
ggagtttgtg tacggagact acaagggtcta cgacgtgagg aagatgatag ccaagtcgga 3180
acaggaaatc gggaaagcaa ctgcgaaata cttcttttac tcaaactca tgaacttttt 3240
caagactgaa attacgtgg ccaatggaga aatcaggaag aggccactga tcgaaactaa 3300
cggagaaaac ggcgaaatcg tgtgggacaa gggcagggac ttcgcaactg ttcgcaaagt 3360
gctctctatg ccgcaagtca atattgtgaa gaaaaccgaa gtgcaaaccg gcggattttc 3420
aaaggaatcg atcctcccaa agagaaatag cgacaagctc attgcacgca agaaagactg 3480
ggacccgaag aagtacggag gattcgattc gccgactgtc gcatactccg tcctcgtggg 3540
ggccaagggt gagaagggaa agagcaaaaa gtcgaaatcc gtcaaagagc tgctggggat 3600
taccatcatg gaacgatcct cgctcgagaa gaacccgatt gatttcctcg aggcaagggg 3660
ttacaaggag gtgaagaagg atctgatcat caaactcccc aagtactcac tgttcgaact 3720
ggaaaatggt cggaagcgca tgctggcttc ggccggagaa ctcaaaaaag gaaatgagct 3780
ggccttgctt agcaagtacg tcaacttcct ctatcttgct tcgcactacg aaaaactcaa 3840
agggtcaccg gaagataacg aacagaagca gcttttcgtg gagcagcaca agcattatct 3900
ggatgaaatc atcgaacaaa tctccgagtt ttcaaagcgc gtgatcctcg ccgacgcaa 3960
cctcgacaaa gtcctgtcgg cctacaataa gcatagagat aagccgatca gagaacaggc 4020
cgagaacatt atccacttgt tcaccctgac taacctggga gcccagccg ctttcaagta 4080
cttcgatact actatcgatc gcaaaagata cacgtccacc aaggaagttc tggacgcgac 4140
cctgatccac caaagcatca ctggactcta cgaaactagg atcgatctgt cgcagctggg 4200
tggcgattga tagtctagcc atcacattta aaagcatctc agcctaccat gagaataaga 4260
gaaagaaaat gaagatcaat agcttattca tctctttttc ttttcgttg gtgtaaagcc 4320
aacaccctgt ctaaaaaaca taaatttctt taatcatttt gcctcttttc tctgtgcttc 4380
aattaataaa aaatggaaag aacctcgag 4409

```

<210> 50

<211> 4140

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 50

```

atggacaaga agtacagcat cggactggac atcggaacaa acagcgtcgg atgggcagtc 60
atcacagacg aatacaaggt cccgagcaag aagttcaagg tcctgggaaa cacagacaga 120
cacagcatca agaagaacct gatcgagaca ctgctgttcg acagcggaga aacagcagaa 180
gcaacaagac tgaagagaac agcaagaaga agatacacia gaagaaagaa cagaatctgc 240
tacctgcagg aaatcttcag caacgaaatg gcaaaggctc acgacagctt cttccaccgg 300
ctggaagaaa gcttcctggt cgaagaagac aagaagcacg aaagacaccc gatcttcgga 360

```

aacatcgtcg acgaagtcgc ataccacgaa aagtaccgga caatctacca cctgagaaaag	420
aagctggctcg acagcacaga caaggcagac ctgagactga tctacctggc actggcacac	480
atgatcaagt tcagaggaca cttcctgata gaaggagacc tgaacccgga caacagcgac	540
gtcgacaagc tgttcatcca gctggtccag acatacaacc agctgttcga agaaaacccg	600
atcaacgcaa gctgagtcga cgcaaaggca atcctgagcg caagactgag caagagcaga	660
agactggaaa acctgatcgc acagctgccg ggagaaaaga agaacggact gttcggaac	720
ctgatcgcac tgagcctggg actgacaccg aacttcaaga gcaacttcga cctggcagaa	780
gacgcaaagc tgcagctgag caaggacaca tacgacgacg acctggacaa cctgctggca	840
cagatcggag accagtacgc agacctgttc ctggcagcaa agaacctgag cgacgcaatc	900
ctgctgagcg acatcctgag agtcaacaca gaaatcaca aggaccgct gagcgcaagc	960
atgatcaaga gatacgacga acaccaccag gacctgacac tgctgaaggc actggtcaga	1020
cagcagctgc cggaaaagta caaggaaatc ttcttcgacc agagcaagaa cggatacgca	1080
ggatacatcg acggaggagc aagccaggaa gaattctaca agttcatcaa gccgatcctg	1140
gaaaagatgg acggaacaga agaactgctg gtcaagctga acagagaaga cctgctgaga	1200
aagcagagaa cattcgacaa cggaagcatc ccgcaccaga tccacctggg agaactgcac	1260
gcaatcctga gaagacagga agacttctac ccgttctga aggacaacag agaaaagatc	1320
gaaaagatcc tgacattcag aatcccgtac tacgtcggac cgctggcaag aggaaacagc	1380
agattcgcat ggatgacaag aaagagcgaa gaaacaatca caccgtggaa cttcgaagaa	1440
gtcgtcgaca agggagcaag cgcacagagc ttcatcgaag gaatgacaaa cttcgacaag	1500
aacctgccga acgaaaaggt cctgccgaag cacagcctgc tgtacgaata cttcacagtc	1560
tacaacgaac tgacaaaggt caagtacgtc acagaaggaa tgagaaagcc ggcattcctg	1620
agcggagaac agaagaaggc aatcgtcgac ctgctgttca agacaaacag aaaggtcaca	1680
gtcaagcagc tgaaggaaga ctacttcaag aagatcgaat gttcgacag cgtcgaaatc	1740
agcggagtcg aagacagatt caacgcaagc ctgggaacat accacgacct gctgaagatc	1800
atcaaggaca aggacttctt ggacaacgaa gaaaacgaag acatcctgga agacatcgtc	1860
ctgacactga cactgttcga agacagagaa atgatcgaag aaagactgaa gacatacgca	1920
cacctgttcg acgacaaggt catgaagcag ctgaagagaa gaagatacac aggatgggga	1980
agactgagca gaaagctgat caacggaatc agagacaagc agagcggaaa gacaatcctg	2040
gacttcttga agagcgacgg attcgcaaac agaaacttca tgcagctgat ccacgacgac	2100
agcctgacat tcaaggaaga catccagaag gcacaggtca gcggacaggg agacagcctg	2160
cacgaacaca tcgcaaacct ggcaggaagc ccggcaatca agaagggaat cctgcagaca	2220
gtcaaggctc tcgacgaact ggtcaaggtc atgggaagac acaagccgga aaacatcgtc	2280
atcgaaatgg caagagaaaa ccagacaaca cagaaggagc agaagaacag cagagaaaga	2340
atgaagagaa tcgaagaagg aatcaaggaa ctgggaagcc agatcctgaa ggaacacccg	2400
gtcgaaaaca cacagctgca gaacgaaaag ctgtacctgt actacctgca aaacggaaga	2460
gacatgtacg tcgaccagga actggacatc aacagactga gcgactacga cgtcgaccac	2520
atcgtcccgc agagcttctt gaaggacgac agcatcgaca acaaggctct gacaagaagc	2580
gacaagaaca gaggaagag cgacaacgtc ccgagcgaag aagtcgtcaa gaagatgaag	2640
aactactgga gacagctgct gaacgcaaag ctgatcacac agagaaagtt cgacaacctg	2700
acaaaggcag agagaggagg actgagcgaa ctggacaagg caggattcat caagagacag	2760
ctggtcgaaa caagacagat cacaagcac gtcgcacaga tcctggacag cagaatgaac	2820
acaaagtacg acgaaaacga caagctgata agagaagtca aggtcatcac actgaagagc	2880
aagctggctc gcgacttcag aaaggacttc cagttctaca aggtcagaga aatcaacaac	2940
taccaccacg cacacgacgc atacctgaac gcagtcgtcg gaacagcact gatcaagaag	3000
taccggaagc tggaaagcga attcgtctac ggagactaca aggtctacga cgtcagaaag	3060
atgatcgcaa agagcgaaca ggaaatcggg aaggcaacag caaagtactt cttctacagc	3120
aacatcatga acttcttcaa gacagaaatc acactggcaa acggagaaat cagaaagaga	3180
ccgctgatcg aaacaaacgg agaaacagga gaaatcgtct gggacaaggg aagagacttc	3240
gcaacagtca gaaaggctct gagcatgccg cagggtcaaca tcgtcaagaa gacagaagtc	3300
cagacaggag gattcagcaa ggaaagcatc ctgccgaaga gaaacagcga caagctgata	3360

10

20

30

40

50

gcaagaaaga aggactggga cccgaagaag tacggaggat tcgacagccc gacagtcgca 3420
 tacagcgtcc tggtcgtcgc aaaggctcga aagggaaga gcaagaagct gaagagcgtc 3480
 aaggaaactgc tgggaatcac aatcatggaa agaagcagct tcgaaaagaa cccgatcgac 3540
 ttcctggaag caaagggaata caaggaagtc aagaaggacc tgatcatcaa gctgccgaag 3600
 tacagcctgt tcgaactgga aaacggaaga aagagaatgc tggcaagcgc aggagaactg 3660
 cagaagggaac acgaactggc actgccgagc aagtacgtca acttcctgta cctggcaagc 3720
 cactacgaaa agctgaaggg aagcccggaa gacaacgaac agaagcagct gttcgtcgaa 3780
 cagcacaagc actacctgga cgaaatcatc gaacagatca gcgaattcag caagagagtc 3840
 atcctggcag acgcaaact ggacaaggct ctgagcgcac acaacaagca cagagacaag 3900
 ccgatcagag aacaggcgaa aaacatcatc cacctgttca cactgacaaa cctgggagca 3960
 ccggcagcat tcaagtactt cgacacaaca atcgacagaa agagatacac aagcacaag 4020
 gaagtcctgg acgcaacact gatccaccag agcatcacag gactgtacga aacaagaatc 4080
 gacctgagcc agctgggagg agacggagga ggaagcccga agaagaagag aaaggcttag 4140

10

<210> 51

<211> 4411

<212> DNA

<213> Artificial Sequence

<220>

<221> source

20

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 51

ggggtcccga gtcggcgtcc agcggctctg cttgttcgtg tgtgtgtcgt tgcaggcctt 60
 attcggatcc gccacatagg acaagaagta cagcatcgga ctggacatcg gaacaaacag 120
 cgtcggatgg gcagtcatca cagacgaata caagggtccc agcaagaagt tcaaggctct 180
 gggaaacaca gacagacaca gcatcaagaa gaacctgacg ggagcactgc tgttcgacag 240
 cggagaaaca gcagaagcaa caagactgaa gagaacagca agaagaagat acacaagaag 300
 aaagaacaga atctgctacc tgcaggaaat cttcagcaac gaaatggcaa aggtcgacga 360
 cagcttcttc caccggctgg aagaaagctt cctggctcga gaagacaaga agcagcaag 420
 acacccgatc ttcggaaaca tcgtcgacga agtcgcatac cacgaaaagt acccgacaat 480
 ctaccacctg agaaagaagc tggtcgacag cacagacaag gcagacctga gactgatcta 540
 cctggcactg gcacacatga tcaagttcag aggacacttc ctgatcgaag gagacctgaa 600
 cccggacaac agcgacgtcg acaagctgtt catccagctg gtccagacat acaaccagct 660
 gttcgaagaa aacccgatca acgcaagcgg agtcgacgca aaggcaatcc tgagcgcaag 720
 actgagcaag agcagaagac tggaaaacct gatcgacag ctgccgggag aaaagaagaa 780
 cggactgttc ggaaaactga tcgcactgag cctgggactg acaccgaact tcaagagcaa 840
 cttcgacctg gcagaagacg caaagctgca gctgagcaag gacacatacg acgacgacct 900
 ggacaacctg ctggcacaga tcggagacca gtacgcagac ctgttcttgg cagcaaagaa 960
 cctgagcgac gcaatcctgc tgagcgacat cctgagagtc aacacagaaa tcacaaaggc 1020
 accgctgagc gcaagcatga tcaagagata cgacgaacac caccaggacc tgacactgct 1080
 gaaggcactg gtcagacagc agctgccgga aaagtacaag gaaatcttct tcgaccagag 1140
 caagaacgga tacgcaggat acatcgacgg aggagcaagc caggaagaat tctacaagtt 1200
 catcaagccg atcctggaaa agatggacgg aacagaagaa ctgctgggtca agctgaacag 1260
 agaagacctg ctgagaaagc agagaacatt cgacaacgga agcatcccgc accagatcca 1320
 cctgggagaa ctgcacgcaa tcctgagaag acaggaagac ttctaccctg tcctgaagga 1380
 caacagagaa aagatcgaaa agatcctgac attcagaatc ccgtactacg tcggaccgct 1440
 ggcaagagga aacagcagat tcgcatggat gacaagaaag agcgaagaaa caatcacacc 1500
 gtggaacttc gaagaagtcg tcgacaaggg agcaagcgca cagagcttca tcgaaagaat 1560
 gacaaacttc gacaagaacc tgccgaacga aaaggctcctg ccgaagcaca gcctgctgta 1620

30

40

50

cgaatacttc acagtctaca acgaactgac aaaggtcaag tacgtcacag aaggaatgag 1680
 aaagccggca ttcctgagcg gagaacagaa gaaggcaatc gtcgacctgc tgttcaagac 1740
 aaacagaaag gtcacagtca agcagctgaa ggaagactac ttcaagaaga tcgaatgctt 1800
 cgacagcgtc gaaatcagcg gagtcgaaga cagattcaac gcaagcctgg gaacatacca 1860
 cgacctgtg aagatcatca aggacaagga cttcctggac aacgaagaaa acgaagacat 1920
 cctggaagac atcgtcctga cactgacact gttcgaagac agagaaatga tcgaagaaag 1980
 actgaagaca tacgcacacc tgttcgacga caaggtcatg aagcagctga agagaagaag 2040
 atacacagga tggggaagac tgagcagaaa gctgatcaac ggaatcagag acaagcagag 2100
 cggaagaca atcctggact tcctgaagag cgacggattc gcaaacagaa acttcatgca 2160
 gctgatccac gacgacagcc tgacattcaa ggaagacatc cagaaggcac aggtcagcgg 2220
 acaggagagac agcctgcacg aacacatcgc aaacctggca ggaagcccgg caatcaagaa 2280
 gggaatcctg cagacagtca aggtcgtcga cgaactggtc aaggtcatgg gaagacacaa 2340
 gccggaaaac atcgtcatcg aaatggcaag agaaaaccag acaacacaga agggacagaa 2400
 gaacagcaga gaaagaatga agagaatcga agaaggaatc aaggaactgg gaagccagat 2460
 cctgaaggaa cccccggtcg aaaacacaca gctgcagaac gaaaagctgt acctgtacta 2520
 cctgcaaaac ggaagagaca tgtacgtcga ccaggaactg gacatcaaca gactgagcga 2580
 ctacgacgtc gaccacatcg tcccgcagag cttcctgaag gacgacagca tcgacaacaa 2640
 ggtcctgaca agaagcgaca agaacagagg aaagagcgac aacgtcccga gcgaagaagt 2700
 cgtcaagaag atgaagaact actggagaca gctgctgaac gcaaagctga tcacacagag 2760
 aaagttcgac aacctgacaa aggcagagag agggaggactg agcgaactgg acaaggcagg 2820
 attcatcaag agacagctgg tcgaaacaag acagatcaca aagcacgtcg cacagatcct 2880
 ggacagcaga atgaacacaa agtacgacga aaacgacaag ctgatcagag aagtcaaggt 2940
 catcacactg aagagcaagc tggtcagcga cttcagaaag gacttccagt tctacaaggt 3000
 cagagaaatc aacaactacc accacgcaca cgacgcatac ctgaacgcag tcgtcggaaac 3060
 agcactgatc aagaagtacc cgaagctgga aagcgaattc gtctacggag actacaaggt 3120
 ctacgacgtc agaaagatga tcgcaaagag cgaacaggaa atcggaaagg caacagcaaa 3180
 gtacttcttc tacagcaaca tcatgaactt cttcaagaca gaaatcacac tggcaaacgg 3240
 agaaatcaga aagagaccgc tgatcgaaac aaacggagaa acaggagaaa tcgtctggga 3300
 caagggaaga gacttcgcaa cagtcagaaa ggtcctgagc atgccgcagg tcaacatcgt 3360
 caagaagaca gaagtccaga caggaggatt cagcaaggaa agcatcctgc cgaagagaaa 3420
 cagcgacaag ctgatcgcaa gaaagaagga ctgggacccg aagaagtacg gaggattcga 3480
 cagcccgaac gtcgcataca gcgtcctggc gtcgcgaaaag gtcgaaaagg gaaagagcaa 3540
 gaagctgaag agcgtcaagg aactgctggg aatcacaatc atggaaagaa gcagcttcga 3600
 aaagaacccg atcgacttcc tggaagcaaa gggatacaag gaagtcaaga aggacctgat 3660
 catcaagctg ccgaagtaca gcctgttcga actggaaaac ggaagaaaga gaatgtgtgc 3720
 aagcgcagga gaactgcaga agggaaaacga actggcactg ccgagcaagt acgtcaactt 3780
 cctgtacctg gcaagccact acgaaaagct gaagggaagc ccggaagaca acgaacagaa 3840
 gcagctgttc gtcgaacagc acaagcacta cctggacgaa atcatcgaac agatcagcga 3900
 attcagcaag agagtcatcc tggcagacgc aaacctggac aaggtcctga gcgcatacaa 3960
 caagcacaga gacaagccga tcagagaaca ggcagaaaac atcatccacc tgttcacact 4020
 gacaaacctg ggagcaccgg cagcattcaa gtacttcgac acaacaatcg acagaaagag 4080
 atacacaagc acaaaggaag tcctggacgc aacactgatc caccagagca tcacaggact 4140
 gtacgaaaca agaatcgacc tgagccagct gggaggagac ggaggaggaa gcccgaaaga 4200
 gaagagaaag gtctagctag ccatcacatt taaaagcatc tcagcctacc atgagaataa 4260
 gagaaagaaa atgaagatca atagcttatt catctctttt tcttttctgt tgggtgaaag 4320
 ccaacaccct gtctaaaaaa cataaatttc tttaatcatt ttgcctcttt tctctgtgct 4380
 tcaattaata aaaaatggaa agaacctcga g 4411

<210> 52

<211> 4140

10

20

30

40

50

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 52

atggacaaga agtacagcat cggcctggac atcggcacca acagcgtggg ctgggccgtg	60	
atcaccgacg agtacaaggt gcccagcaag aagttcaagg tgctgggcaa caccgacaga	120	
cacagcatca agaagaacct gatcggcgcc ctgctgttcg acagcggcga gaccgccgag	180	10
gccaccgacg tgaagagaac cgccagaaga agatacacca gaagaaagaa cagaatctgc	240	
tacctgcagg agatcttcag caacgagatg gccaaggtgg acgacagctt cttccacaga	300	
ctggaggaga gcttcctggt ggaggaggac aagaagcacg agagacaccc catcttcggc	360	
aacatcgtgg acgaggtggc ctaccacgag aagtaccca ccatctacca cctgagaaag	420	
aagctggtgg acagcaccga caaggccgac ctgagactga tctacctggc cctggccac	480	
atgatcaagt tcagaggcca cttcctgacg gagggcgacc tgaaccccga caacagcgac	540	
gtggacaagc tgtcatcca gctggtgcag acctacaacc agctgttcga ggagaacccc	600	
atcaacgcca gcggcgtgga cgccaaggcc atcctgagcg ccagactgag caagagcaga	660	
agactggaga acctgatcgc ccagctgccc ggcgagaaga agaacggcct gttcggaac	720	
ctgatcggcc tgagcctggg cctgacccc aacttcaaga gcaacttga cctggccgag	780	20
gacgccaagc tgcagctgag caaggacacc tacgacgacg acctggacaa cctgctggcc	840	
cagatcggcg accagtacgc cgacctgttc ctggccgcca agaacctgag cgacgccatc	900	
ctgctgagcg acatcctgag agtgaacacc gagatcacca aggccccct gagcgccagc	960	
atgatcaaga gatacgacga gcaccaccag gacctgacc tgctgaaggc cctggtgaga	1020	
cagcagctgc ccgagaagta caaggagatc ttcttcgacc agagcaagaa cggctacgcc	1080	
ggctacatcg acggcggcgc cagccaggag gagttctaca agttcatcaa gccatcctg	1140	
gagaagatgg acggcaccga ggagctgctg gtgaagctga acagagagga cctgctgaga	1200	
aagcagagaa ccttcgacaa cggcagcatc ccccaccaga tccacctggg cgagctgcac	1260	
gccatcctga gaagacagga ggacttctac ccttcctga aggacaacag agagaagatc	1320	
gagaagatcc tgaccttcag aatcccctac tacgtgggcc cctggccag aggcaacagc	1380	30
agattcgcct ggatgaccag aaagagcgag gagacatca cccctggaa cttcgaggag	1440	
gtggtggaca agggcgccag cgcccagagc ttcacgaga gaatgaccaa cttcgacaag	1500	
aacctgccc aagagaaggt gctgcccag cacagcctgc tgtacgagta cttaccctg	1560	
tacaacgagc tgaccaaggt gaagtacgtg accgagggca tgagaaagcc cgccttcctg	1620	
agcggcgagc agaagaaggc catcgtggac ctgctgttca agaccaacag aaaggtgacc	1680	
gtgaagcagc tgaaggagga ctacttcaag aagatcgagt gcttcgacag cgtggagatc	1740	
agcggcgtgg aggacagatt caacgccagc ctgggcacct accacgacct gctgaagatc	1800	
atcaaggaca aggacttcct ggacaacgag gagaacgagg acatcctgga ggacatcgtg	1860	
ctgacctga cctgttcga ggacagagag atgatcgagg agagactgaa gacctacgcc	1920	
cacctgttcg acgacaaggt gatgaagcag ctgaagagaa gaagatacac cggctggggc	1980	40
agactgagca gaaagctgat caacggcatc agagacaagc agagcggcaa gaccatcctg	2040	
gacttcctga agagcgacgg cttcgccaac agaaacttca tgcagctgat ccacgacgac	2100	
agcctgacct tcaaggagga catccagaag gccaggtga gcggccaggg cgacagcctg	2160	
cacgagcaca tcgccaacct ggccggcagc ccgcccata agaagggcac cctgcagacc	2220	
gtgaagggtg tggacgagct ggtgaagggt atgggcagac acaagcccga gaacatcgtg	2280	
atcgagatgg ccagagagaa ccagaccacc cagaagggcc agaagaacag cagagagaga	2340	
atgaagagaa tcgaggaggg catcaaggag ctgggcagcc agatcctgaa ggagcacccc	2400	
gtggagaaca cccagctgca gaacgagaag ctgtacctgt actacctga gaacggcaga	2460	
gacatgtacg tggaccagga gctggacatc aacagactga gcgactacga cgtggaccac	2520	
atcgtgcccc agagcttcct gaaggacgac agcatcgaca acaagggtgct gaccagaagc	2580	50

gacaagaaca gaggcaagag cgacaacgtg cccagcgagg aggtggtgaa gaagatgaag 2640
 aactactgga gacagctgct gaacgccaaag ctgatcacc agagaaagt cgacaacctg 2700
 accaaggccg agagaggcgg cctgagcgag ctggacaagg ccggcttcat caagagacag 2760
 ctggtggaga ccagacagat caccaagcac gtggcccaga tcctggacag cagaatgaac 2820
 accaagtacg acgagaacga caagctgac agagagggtga aggtgatcac cctgaagagc 2880
 aagctggtga gcgacttcag aaaggacttc cagttctaca aggtgagaga gatcaacaac 2940
 taccaccacg cccacgacgc ctacctgaac gccgtggtgg gcaccgccct gatcaagaag 3000
 taccccaagc tggagagcga gttcgtgtac ggcgactaca aggtgtacga cgtgagaaag 3060
 atgatcgcca agagcgagca ggagatcggc aaggccaccg ccaagtactt cttctacagc 3120
 aacatcatga acttcttcaa gaccgagatc accctggcca acggcgagat cagaaagaga 3180
 cccctgatcg agaccaacgg cgagaccggc gagatcgtgt gggacaaggg cagagacttc 3240
 gccaccgtga gaaagggtgct gagcatgccc cagggtgaaca tcgtgaagaa gaccgagggtg 3300
 cagaccggcg gcttcagcaa ggagagcatc ctgcccaaga gaaacagcga caagctgac 3360
 gccagaaaga aggactggga cccaagaag tacggcggct tcgacagccc caccgtggcc 3420
 tacagcgtgc tgggtggtggc caagggtggag aagggcaaga gcaagaagct gaagagcgtg 3480
 aaggagctgc tgggcatcac catcatggag agaagcagct tcgagaagaa ccccatcgac 3540
 ttcttgagg ccaagggtga caaggagggtg aagaaggacc tgatcatcaa gctgccaag 3600
 tacagcctgt tcgagctgga gaacggcaga aagagaatgc tggccagcgc cggcgagctg 3660
 cagaagggca acgagctggc cctgcccagc aagtacgtga acttcctgta cctggccagc 3720
 cactacgaga agctgaaggg cagccccgag gacaacgagc agaagcagct gttcgtggag 3780
 cagcacaagc actacctgga cgagatcatc gagcagatca gcgagttcag caagagagtg 3840
 atcctggccg acgccaacct ggacaagggtg ctgagcgcct acaacaagca cagagacaag 3900
 cccatcagag agcaggccga gaacatcatc cacctgttca cctgaccaa cctgggccc 3960
 cccgccgcct tcaagtactt cgacaccacc atcgacagaa agagatacac cagcacaag 4020
 gaggtgctgg acgccacct gatccaccag agcatcaccg gcctgtacga gaccagaatc 4080
 gacctgagcc agctgggcgg cgacggcggc ggcagcccca agaagaagag aaagggtgtga 4140

10

20

<210> 53

<211> 4411

<212> DNA

<213> Artificial Sequence

<220>

<221> source

 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 53

gggccccga gtcggcgtcc agcggctctg cttgttcgtg tgtgtgtcgt tgcaggcctt 60
 attcggatcc gccaccatgg acaagaagta cagcatcggc ctggacatcg gcaccaacag 120
 cgtgggctgg gccgtgatca ccgacgagta caagggtccc agcaagaagt tcaagggtgt 180
 gggcaacacc gacagacaca gcatcaagaa gaacctgac ggcgccctgc tgttcgacag 240
 cggcgagacc gccgaggcca ccagactgaa gagaaccgcc agaagaagat acaccagaag 300
 aaagaacaga atctgtacc tgcaggagat cttcagcaac gagatggcca aggtggacga 360
 cagcttcttc cacagactgg aggagagctt cctggtggag gaggacaaga agcacgagag 420
 acaccccatc ttcggcaaca tcgtggacga ggtggcctac cagagaagt accccacct 480
 ctaccacctg agaaagaagc tgggtggacag caccgacaag gccgacctga gactgatcta 540
 cctggccctg gccacatga tcaagttcag aggccacttc ctgatcgagg gcgacctgaa 600
 ccccgacaac agcgacgtgg acaagctgtt catccagctg gtgcagacct acaaccagct 660
 gttcgaggag aaccccatca acgccagcgg cgtggacgcc aaggccatcc tgagcgccag 720
 actgagcaag agcagaagac tggagaacct gatcgcccag ctgccggcg agaagaagaa 780
 cggcctgttc ggcaacctga tcgccctgag cctgggcctg acccccaact tcaagagcaa 840

30

40

50

cttcgacctg gccgaggacg ccaagctgca gctgagcaag gacacctacg acgacgacct	900	
ggacaacctg ctggcccgaga tcggcgacca gtacgccgac ctgttccttg cgcgaagaa	960	
cctgagcgac gccatcctgc tgagcgacat cctgagagtg aacaccgaga tcaccaaggc	1020	
ccccctgagc gccagcatga tcaagagata cgacgagcac caccaggacc tgacctgtct	1080	
gaaggccctg gtgagacagc agctgcccga gaagtacaag gagatcttct tcgaccagag	1140	
caagaacggc tacgccggct acatcgacgg cggcgccagc caggaggagt tctacaagtt	1200	
catcaagccc atcctggaga agatggacgg caccgaggag ctgctggtga agctgaacag	1260	
agaggacctg ctgagaaagc agagaacctt cgacaacggc agcatcccc accagatcca	1320	
cctgggagcag ctgcacgcca tcctgagaag acaggaggac ttctaccctt tcctgaagga	1380	
caacagagag aagatcgaga agatcctgac cttcagaatc ccctactacg tgggccccct	1440	10
ggccagaggc aacagcagat tcgcctggat gaccagaaag agcgaggaga ccatcacccc	1500	
ctggaacttc gaggagggtg tggacaaggg cgccagcgcc cagagcttca tcgagagaat	1560	
gaccaacttc gacaagaacc tgcccaacga gaagggtgtg cccaagcaca gcctgtgtga	1620	
cgagtacttc accgtgtaca acgagctgac caaggtgaag tacgtgaccg agggcatgag	1680	
aaagccccgc ttcttgagcg gcgagcagaa gaaggccatc gtggacctgc tgttcaagac	1740	
caacagaaag gtgaccgtga agcagctgaa ggaggactac ttcaagaaga tcgagtgtct	1800	
cgacagcgtg gagatcagcg gcgtggagga cagattcaac gccagcctgg gcacctacca	1860	
cgacctgtg aagatcatca aggacaagga cttcctggac aacgaggaga acgaggacat	1920	
cctggaggac atcgtgtgtga ccctgaccct gttcgaggac agagagatga tcgaggagag	1980	
actgaagacc tacgcccacc tgttcgacga caaggtgatg aagcagctga agagaagaag	2040	20
atacaccggc tggggcagac tgagcagaaa gctgatcaac ggcatcagag acaagcagag	2100	
cggcaagacc atcctggact tcctgaagag cgacggcttc gccaacagaa acttcatgca	2160	
gctgatccac gacgacagcc tgaccttaa ggaggacatc cagaaggccc aggtgagcgg	2220	
ccaggggcag agcctgcacg agcacatcgc caacctggcc ggcagccccg ccatcaagaa	2280	
gggcatcctg cagaccgtga aggtggtgga cgagctggtg aaggtgatgg gcagacacaa	2340	
gcccagaaac atcgtgatcg agatggccag agagaaccag accaccaga agggccagaa	2400	
gaacagcaga gagagaatga agagaatcga ggagggcac aaggagctgg gcagccagat	2460	
cctgaaggag caccctgtgg agaacaccca gctgcagaac gagaagctgt acctgtacta	2520	
cctgcagaac ggcagagaca tgtacgtgga ccaggagctg gacatcaaca gactgagcga	2580	
ctacgacgtg gaccacatcg tgccccagag cttcctgaag gacgacagca tcgacaacaa	2640	30
ggtgctgacc agaagcgaca agaacagagg caagagcgac aacgtgccca gcgaggaggt	2700	
ggtgaagaag atgaagaact actggagaca gctgctgaac gccaagctga tcaccagag	2760	
aaagttcgac aacctgacca aggccgagag aggcggcctg agcgagctgg acaaggccgg	2820	
cttcatcaag agacagctgg tggagaccag acagatcacc aagcacgtgg ccagatcct	2880	
ggacagcaga atgaacacca agtacgacga gaacgacaag ctgatcagag aggtgaaggt	2940	
gatcacctg aagagcaagc tggtagcga cttcagaaag gacttccagt tctacaaggt	3000	
gagagagatc aacaactacc accagcccc cgacgcctac ctgaacgccg tggtaggcac	3060	
cgccctgatc aagaagtacc ccaagctgga gagcgagttc gtgtacggcg actacaaggt	3120	
gtacgacgtg agaaagatga tcgccaagag cgagcaggag atcggaagg ccaccgcaa	3180	
gtacttcttc tacagcaaca tcatgaactt cttcaagacc gagatcacc tggccaacgg	3240	40
cgagatcaga aagagacccc tgatcgagac caacggcgag accggcgaga tcgtgtggga	3300	
caagggcaga gacttcgcca ccgtgagaaa ggtgctgagc atgccccagg tgaacatcgt	3360	
gaagaagacc gaggtgcaga ccggcggtt cagcaaggag agcatcctgc ccaagagaaa	3420	
cagcgacaag ctgatcgcca gaaagaagga ctgggacccc aagaagtacg gcggcttcga	3480	
cagccccacc gtggcctaca gcgtgctggt ggtggccaag gtggagaagg gcaagagcaa	3540	
gaagctgaag agcgtgaagg agctgctggg catcaccatc atggagagaa gcagcttcga	3600	
gaagaacccc atcgacttcc tggaggccaa gggctacaag gaggtgaaga aggacctgat	3660	
catcaagctg cccaagtaca gcctgttcga gctggagaac ggcagaaaga gaatgctggc	3720	
cagcgccggc gagctgcaga agggcaacga gctggccctg ccagcaagt acgtgaactt	3780	
cctgtacctg gccagccact acgagaagct gaagggcagc cccgaggaca acgagcagaa	3840	50

gcagctgttc gtggagcagc acaagcacta cctggacgag atcatcgagc agatcagcga 3900
gttcagcaag agagtgatcc tggccgacgc caacctggac aagggtgctga ggcgctacaa 3960
caagcacaga gacaagccca tcagagagca ggccgagaac atcatccacc tggtcacctt 4020
gaccaacctg ggcgcccccg ccgccttcaa gtacttcgac accaccatcg acagaaagag 4080
atacaccagc accaaggagg tgctggacgc caccctgatc caccagagca tcaccggcct 4140
gtacgagacc agaatcgacc tgagccagct gggcgggcgc ggcggcgga gcccgaagaa 4200
gaagagaaag gtgtgactag ccatcacatt taaaagcatc tcagcctacc atgagaataa 4260
gagaaagaaa atgaagatca atagcttatt catctctttt tcttttctgt tgggtgtaaag 4320
ccaacacctt gtctaaaaaa cataaatctt ttaatacatt ttgcctcttt tctctgtgct 4380
tcaattaata aaaaatggaa agaacctcga g 4411

10

<210> 54

<211> 4140

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 54

20

atggacaaaa aatacagcat agggctagac atagggacga acagcgtagg gtgggaggga 60
ataacggacg aatacaaagt accgagcaaa aaattcaaag tactagggaa cacggaccga 120
cacagcataa aaaaaaacct aataggggag ctactattcg acagcgggga aacggcggaa 180
gcgacgcgac taaaacgaac ggcgcgacga cgatacacgc gacgaaaaaa ccgaatatgc 240
tacctacaag aaatatcag caacgaaatg gcgaaagtag acgacagctt cttccaccga 300
ctagaagaaa gcttcctagt agaagaagac aaaaaacacg aacgacaccc gatattcggg 360
aacatagtag acgaagtagc gtaccacgaa aaatacccga cgatatacca cctacgaaaa 420
aaactagtag acagcacgga caaagcggac ctacgactaa tatacctagc gctagcgcac 480
atgataaaat tccgagggga cttcctaata gaaggggacc taaaccgga caacagcgac 540
gtagacaaac tattcataca actagtacaa acgtacaacc aactattcga agaaaaccg 600
ataaacgcga gcgggggtaga cgcgaaagcg atactaagcg cgcgactaag caaaagccga 660
cgactagaaa acctaatagc gcaactaccg ggggaaaaaa aaaacgggct attcgggaac 720
ctaatagcgc taagcctagg gctaacgccg aacttcaaaa gcaacttcga cctagcggaa 780
gacgcgaaac tacaactaag caaagacacg tacgacgacg acctagacaa cctactagcg 840
caaatagggg accaatacgc ggacctattc ctacggcgga aaaacctaag cgacgcgata 900
ctactaagcg acatactacg agtaaacacg gaaataacga aagcgccgct aagcgcgagc 960
atgataaaac gatacgacga acaccacaa gacctaacgc tactaaaagc gctagtacga 1020
caacaactac cggaaaaaata caaagaaata ttcttcgacc aaagcaaaaa cgggtacgcg 1080
gggtacatag acgggggggg gagccaagaa gaattctaca aattcataaa accgatacta 1140
gaaaaaatgg acgggacgga agaactacta gtaaaactaa accgagaaga cctactacga 1200
aaacaacgaa cgttcgacaa cgggagcata ccgcacaaaa tacacctagg ggaactacac 1260
gcgatactac gacgacaaga agacttctac ccgttcctaa aagacaaccg agaaaaaata 1320
gaaaaaatac taacgttccg aataccgtac tacgtagggc cgctagcgcg agggaacagc 1380
cgattcgcgt ggatgacgcg aaaaagcgaa gaaacgataa cgccgtggaa cttcgaagaa 1440
gtagtagaca aaggggagcg cgcgcaaagc tcatagaac gaatgacgaa cttcgacaaa 1500
aacctaccga acgaaaaagt actaccgaaa cacagcctac tatacgaata cttcacggta 1560
tacaacgaac taacgaaagt aaaatacgtg acggaaggga tgcgaaaacc ggcgttccta 1620
agcgggggaa aaaaaaaagc gatagtagac ctactattca aaacgaaccg aaaagtaacg 1680
gtaaaacaac taaaagaaga ctacttcaaa aaaatagaat gtttcgacag cgtagaaata 1740
agcggggtag aagaccgatt caacgcgagc ctaggagcgt accacgacct actaaaaata 1800

30

40

50

ataaaagaca aagacttcct agacaacgaa gaaaacgaag acatactaga agacatagta 1860
 ctaacgctaa cgctattcga agaccgagaa atgatagaag aacgactaaa aacgtacgcg 1920
 cacctattcg acgacaaagt aatgaaacaa ctaaaacgac gacgatacac ggggtggggg 1980
 cgactaagcc gaaaactaat aaacgggata cgagacaaac aaagcgggaa aacgatacta 2040
 gacttcctaa aaagcgacgg gtgcgcgaac cgaaacttca tgcaactaat acacgacgac 2100
 agcctaacgt tcaaagaaga catacaaaaa gcgcaagtaa gcgggcaagg ggacagccta 2160
 cacgaacaca tagcgaacct agcgggggagc ccggcgataa aaaaagggat actacaaacg 2220
 gtaaaagtag tagacgaact agtaaaagta atggggcgac acaaaccgga aaacatagta 2280
 atagaaatgg cgcgagaaaa ccaaacgacg caaaaagggc aaaaaaacag ccgagaacga 2340
 atgaaacgaa tagaagaagg gataaaagaa ctagggagcc aaatactaaa agaacacccg 2400
 gtagaaaaca cgcaactaca aaacgaaaaa ctatacctat actacctaca aaacgggcca 2460
 gacatgtacg tagaccaaga actagacata aaccgactaa gcgactacga cgtagaccac 2520
 atagtaccgc aaagcttcct aaaagacgac agcatagaca acaaagtact aacgcgaagc 2580
 gacaaaaacc gagggaaaag cgacaacgta ccgagcgaag aagtagtaaa aaaaatgaaa 2640
 aactactggc gacaactact aaacgcgaaa ctaataacgc aacgaaaatt cgacaaccta 2700
 acgaaagcgg aacgaggggg gctaagcgaa ctagacaaag cggggttcat aaaacgacaa 2760
 ctagtagaaa cgcgacaaat aacgaaacac gtagcgcaaa tactagacag ccgaatgaac 2820
 acgaaatacg acgaaaacga caaactaata cgagaagtaa aagtaataac gtaaaaagc 2880
 aaactagtaa gcgacttccg aaaagacttc caattctaca aagtacgaga aataaacaac 2940
 taccaccacg cgacgacgc gtacctaaac gcggtagtag ggacggcgct aataaaaaaa 3000
 taccgaaac tagaaagcga attcgtatag ggggactaca aagtatacga cgtacgaaaa 3060
 atgatagcga aaagcgaaca agaaataggg aaagcgacgg cgaaatactt cttctacagc 3120
 aacataatga acttcttcaa aacggaaata acgctagcga acgggggaaat acgaaaacga 3180
 ccgctaatag aaacgaacgg ggaacggggg gaaatagtat gggacaaagg gcgagacttc 3240
 gcgacggtac gaaaagtact aagcatgccg caagtaaaaa tagtaaaaaa aacggaagta 3300
 caaacggggg ggttcagcaa agaaagcata ctaccgaaac gaaacagcga caaactaata 3360
 gcgcgaaaaa aagactggga cccgaaaaaa tacgggggggt tcgacagccc gacggtagcg 3420
 tacagcgtag tagtagtagc gaaagtagaa aaagggaaaa gcaaaaaact aaaaagcgta 3480
 aaagaactac tagggataac gataatggaa cgaagcagct tcgaaaaaaa cccgatagac 3540
 ttcctagaag cgaaagggtg caaagaagta aaaaaagacc taataataaa actaccgaaa 3600
 tacagcctat tcgaactaga aaacgggcca aaacgaatgc tagcgagcgc gggggaacta 3660
 caaaaaggga acgaactagc gctaccgagc aaatacgtaa acttctata cctagcgagc 3720
 cactacgaaa aactaaaagg gagccccgaa gacaacgaac aaaaacaact attcgtagaa 3780
 caacacaaac actacctaga cgaaataata gaacaaataa gcgaattcag caaacgagta 3840
 atactagcgg acgcgaacct agacaaagta ctaagcgctg acaacaaaca ccgagacaaa 3900
 ccgatacgag aacaagcggg aaacataata cacctattca cgctaacgaa cctagggggc 3960
 ccggcggcgt tcaataactt cgacacgacg atagaccgaa aacgatacac gagcacgaaa 4020
 gaagtactag acgcgacgct aatacaccaa agcataacgg ggctatacga aacgcgaata 4080
 gacctaagcc aactaggggg ggacgggggg gggagcccga aaaaaaacg aaaagtatga 4140

<210> 55

<211> 4411

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 55

gggtcccga gtcggcgtcc agcggctctg cttgttcgtg tgtgtgtcgt tgcaggcctt 60

10

20

30

40

50

attcggatcc gccaccatgg acaaaaaata cagcataggg ctagacatag ggacgaacag	120	
cgtagggtgg gcggtataa cggacgaata caaagtaccg agcaaaaaat tcaaagtact	180	
agggaacacg gaccgacaca gcataaaaaa aaacctaata ggggcgctac tattcgacag	240	
cggggaaacg gcggaagcga cgcgactaaa acgaacggcg cgacgacgat acacgcgacg	300	
aaaaaacga atatgctacc tacaagaaat attcagcaac gaaatggcga aagtagacga	360	
cagcttcttc caccgactag aagaaagctt ctagtagaaa gaagacaaaa aacacgaacg	420	
acacccgata ttcggggaaca tagtagacga agtagcgtag cacgaaaaat acccgacgat	480	
ataccaccta cgaaaaaac tagtagacag caccgacaaa gcggacctac gactaatata	540	
cctagcgcta gcgcacatga taaaattccg agggcacttc ctaatagaag gggacctaaa	600	
cccggaacac agcgacgtag acaactatt catacaacta gtacaaacgt acaaccaact	660	10
attcgaagaa aacccgataa acgcgagcgg ggtagacgag aaagcgatac taagcgcgcg	720	
actaagcaaa agccgacgac tagaaaacct aatagcgcaa ctaccggggg aaaaaaaaaa	780	
cgggctattc gggaacctaa tagcgctaag ctaggggcta acgccgaact tcaaaagcaa	840	
cttcgacctt gcggaagacg cgaaactaca actaagcaaa gacacgtacg acgacgacct	900	
agacaaccta ctagcgcaaa taggggacca atacgcggac ctattcctag cggcgaaaaa	960	
cctaagcgac gcgatactac taagcgacat actacgagta aacacggaaa taacgaaagc	1020	
gccgctaagc gcgagcatga taaaacgata cgacgaacac caccaagacc taacgctact	1080	
aaaagcgcta gtacgacaac aactaccgga aaaatacaaa gaaatattct tcgaccaaaag	1140	
caaaaacggg tacgcggggg acatagacgg gggggcgagc caagaagaat tctacaaatt	1200	
cataaaaccg atactagaaa aaatggacgg gacggaagaa ctactagtaa aactaaaccg	1260	20
agaagaccta ctacgaaaac aacgaacgtt cgacaacggg agcataccgc accaaatata	1320	
cctaggggaa ctacacgca tactacgacg acaagaagac ttctaccgt tcctaaaaga	1380	
caaccgagaa aaaatagaaa aaatactaac gttccgaata ccgtactacg tagggccgct	1440	
agcgcgaggg aacagccgat tcgcgtggat gacgcgaaaa agcgaagaaa cgataacgcc	1500	
gtggaacttc gaagaagtag tagacaaagg ggcgagcgcg caaagcttca tagaacgaat	1560	
gacgaacttc gacaaaaacc taccgaacga aaaagtacta ccgaaacaca gcctactata	1620	
cgaatacttc acggtataca acgaactaac gaaagtaaaa tacgtaacgg aagggatgag	1680	
aaaaccggcg ttccctaagcg gggaacaaaa aaaagcgata gtagacctac tattcaaaac	1740	
gaaccgaaaa gtaacggtaa aacaactaaa agaagactac ttcaaaaaaa tagaatgctt	1800	
cgacagcgta gaaataagcg gggtagaaga ccgattcaac gcgagcctag ggacgtacca	1860	30
cgacctacta aaaataataa aagacaaaaga cttcctagac aacgaagaaa acgaagacat	1920	
actagaagac atagtactaa cgctaacgct attcgaagac cgagaaatga tagaagaacg	1980	
actaaaaacg tacgcgcacc tattcgacga caaagtaatg aaacaactaa aacgacgacg	2040	
atacacgggg tgggggagc taagccgaaa actaataaac gggatacgag acaaacaaag	2100	
cgggaaaacg atactagact tcctaaaaag cgacgggttc gcgaaccgaa acttcattgca	2160	
actaatacac gacgacagcc taacgttcaa agaagacata caaaaagcgc aagtaagcgg	2220	
gcaaggggac agcctacag aacacatagc gaacctagcg gggagcccgg cgataaaaaa	2280	
agggatacta caaacggtaa aagtagtaga cgaactagta aaagtaatgg ggcgacacaa	2340	
accggaaaac atagtaatag aaatggcgcg agaaaaccaa acgacgcaaa aaggcgaaaa	2400	
aaacagccga gaacgaatga aacgaataga agaagggata aaagaactag ggagccaaat	2460	40
actaaaagaa caccggtag aaaaacacgca actacaaaac gaaaaactat acctatacta	2520	
cctacaaaac gggcgagaca tgtacgtaga ccaagaacta gacataaacc gactaagcga	2580	
ctacgacgta gaccacatag taccgcaaag cttcctaaaa gacgacagca tagacaacaa	2640	
agtactaacg cgaagcgaca aaaaccgagg gaaaagcgac aacgtaccga gcgaagaagt	2700	
agtaaaaaaa atgaaaaact actggcgaca actactaaac gcgaaactaa taacgcaacg	2760	
aaaattcgac aacctaacga aagcggaaac aggggggcta agcgaactag acaaagcggg	2820	
gttcataaaa cgacaactag tagaaacgag acaataaacg aaacacgtag cgcaaatact	2880	
agacagccga atgaacacga aatacagcga aaacgacaaa ctaatacgag aagtaaaagt	2940	
aataacgcta aaaagcaaac tagtaagcga cttccgaaaa gacttccaat tctacaaagt	3000	
acgagaaata aacaactacc accacgcgca cgacgcgtac ctaaacgcgg tagtagggac	3060	50

ggcgctaata aaaaaatacc cgaaactaga aagcgaattc gtatacggggg actacaaagt 3120
 atacgacgta cgaaaaatga tagcgaaaag cgaacaagaa atagggaag cgacggcgaa 3180
 atacttcttc tacagcaaca taatgaactt cttcaaaacg gaaataacgc tagcgaacgg 3240
 ggaaatacga aaacgaccgc taatagaaac gaacgggggaa acggggggaaa tagtatggga 3300
 caaagggcga gacttcgcga cggtacgaaa agtactaagc atgccgcaag taaacatagt 3360
 aaaaaaacg gaagtacaaa cgggggggtt cagcaaagaa agcatactac cgaaacgaaa 3420
 cagcgacaaa ctaatagcgc gaaaaaaga ctgggacccg aaaaaatagc ggggggttcga 3480
 cagcccgcg gtagcgtaca gcgtactagt agtagcgaag gtagaaaaag ggaaaagcaa 3540
 aaaactaaaa agcgtaaaag aactactagg gataacgata atggaacgaa gcagcttcga 3600
 aaaaaacccg atagacttcc tagaagcgaa aggttacaaa gaagtaaaaa aagacctaata 3660
 aataaaacta ccgaaatata gcctattcga actagaaaac gggcgaaaac gaatgctagc 3720
 gagcgcggggg gaactacaaa aagggaacga actagcgcta ccgagcaaat acgtaaactt 3780
 cctataccta gcgagccact acgaaaaact aaaagggagc ccggaagaca acgaacaaaa 3840
 acaactattc gtagaacaac acaaacacta cctagacgaa ataatagaac aaataagcga 3900
 attcagcaaa cgagtaatac tagcggacgc gaacctagac aaagtactaa gcgcgtacaa 3960
 caaacaccga gacaaaccga tacgagaaca agcggaaaac ataatacacc tattcacgct 4020
 aacgaaccta ggggcgcggg cggtgttcaa atacttcgac acgacgatag accgaaaacg 4080
 atacacgagc acgaaagaag tactagacgc gacgctaata cacaaagca taacggggct 4140
 atacgaaacg cgaatagacc taagccaact aggggggggac ggggggggga gcccgaaaaa 4200
 aaaacgaaaa gtatgactag ccatcacatt taaaagcatc tcagcctacc atgagaataa 4260
 gagaaagaaa atgaagatca atagcttatt catctctttt tcttttctg tgggtgaaag 4320
 ccaacaccct gtctaaaaaa cataaatttc tttaatcatt ttgcctcttt tctctgtgct 4380
 tcaattaata aaaaatggaa agaacctcga g 4411

<210> 56

<211> 4411

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 56

aggtcccga gtcggcgtcc agcggctctg cttgttcgtg tgtgtgtcgt tgcaggcctt 60
 attcggatcc gccacatgg acaagaagta cagcatcgga ctggacatcg gaacaaacag 120
 cgtcggatgg gcagtcacg cagacgaata caaggtcccg agcaagaagt tcaaggtcct 180
 gggaacaca gacagacaca gcatcaagaa gaacctgacg ggagcactgc tgttcgacag 240
 cggagaaaca gcagaagcaa caagactgaa gagaacagca agaagaagat acacaagaag 300
 aaagaacaga atctgctacc tgcaggaaat cttcagcaac gaaatggcaa aggtcgacga 360
 cagcttcttc cacagactgg aagaaagctt cctggtcgaa gaagacaaga agcagcaag 420
 acacccgatc ttcggaaaca tcgtcgacga agtcgcatac cacgaaaagt acccgacaat 480
 ctaccacctg agaaagaagc tggtcgacag cacagacaag gcagacctga gactgatcta 540
 cctggcactg gcacacatga tcaagttcag aggacacttc ctgatcgaag gagacctgaa 600
 cccggacaac agcgacgtcg acaagctgtt catccagctg gtccagacat acaaccagct 660
 gttcgaagaa aacccgatca acgcaagcgg agtcgacgca aaggcaatcc tgagcgcaag 720
 actgagcaag agcagaagac tggaaaacct gatcgacag ctgccgggag aaaagaagaa 780
 cggactgttc ggaaacctga tcgcactgag cctgggactg acaccgaact tcaagagcaa 840
 cttcgacctg gcagaagacg caaagctgca gctgagcaag gacacatacg acgacgacct 900
 ggacaacctg ctggcacaga tcggagacca gtacgcagac ctgttcttgg cagcaaagaa 960
 cctgagcgac gcaatcctgc tgagcgacat cctgagagtc aacacagaaa tcacaaaggc 1020

accgctgagc gcaagcatga tcaagagata cgacgaacac caccaggacc tgacactgct 1080
 gaaggcactg gtcagacagc agctgccgga aaagtacaag gaaatcttct tcgaccagag 1140
 caagaacgga tacgcaggat acatcgacgg aggagcaagc caggaagaat tctacaagtt 1200
 catcaagccg atcctggaaa agatggacgg aacagaagaa ctgctggtca agctgaacag 1260
 agaagacctg ctgagaaagc agagaacatt cgacaacgga agcatcccgc accagatcca 1320
 cctgggagaa ctgcacgcaa tcctgagaag acaggaagac ttctaccgt tcctgaagga 1380
 caacagagaa aagatcgaaa agatcctgac attcagaatc ccgtactacg tcggaccgct 1440
 ggcaagagga aacagcagat tcgcatggat gacaagaaag agcgaagaaa caatcacacc 1500
 gtggaacttc gaagaagtcg tcgacaaggg agcaagcgca cagagcttca tcgaaagaat 1560
 gacaaacttc gacaagaacc tgccgaacga aaaggctctg ccgaagcaca gcctgctgta 1620
 cgaatacttc acagtctaca acgaactgac aaagggtcaag tacgtcacag aaggaatgag 1680
 aaagccggca ttcctgagcg gagaacagaa gaaggcaatc gtcgacctgc tgttcaagac 1740
 aaacagaaag gtcacagtca agcagctgaa ggaagactac ttcaagaaga tcgaatgctt 1800
 cgacagcgtc gaaatcagcg gagtcgaaga cagattcaac gcaagcctgg gaacatacca 1860
 cgacctgctg aagatcatca aggacaagga cttcctggac aacgaagaaa acgaagacat 1920
 cctggaagac atcgtcctga cactgacact gttcgaagac agagaaatga tcgaagaaag 1980
 actgaagaca tacgcacacc tgttcgacga caaggctcatg aagcagctga agagaagaag 2040
 atacacagga tggggaagac tgagcagaaa gctgatcaac ggaatcagag acaagcagag 2100
 cggaagagaca atcctggact tcctgaagag cgacggattc gcaaacagaa acttcatgca 2160
 gctgatccac gacgacagcc tgacattcaa ggaagacatc cagaaggcac aggtcagcgg 2220
 acaggagagac agcctgcacg aacacatcgc aaacctggca ggaagcccgg caatcaagaa 2280
 gggaatcctg cagacagtca aggtcgtcga cgaactggtc aaggtcatgg gaagacacaa 2340
 gccggaaaac atcgtcatcg aaatggcaag agaaaaccag acaacacaga agggacagaa 2400
 gaacagcaga gaaagaatga agagaatcga agaaggaatc aaggaactgg gaagccagat 2460
 cctgaaggaa cccccggtcg aaaacacaca gctgcagaac gaaaagctgt acctgtacta 2520
 cctgcagaac ggaagagaca tgtacgtcga ccaggaactg gacatcaaca gactgagcga 2580
 ctacgacgtc gaccacatcg tcccgcagag cttcctgaag gacgacagca tcgacaacaa 2640
 ggtcctgaca agaagcgaca agaacagagg aaagagcgac aacgtcccga gcgaagaagt 2700
 cgtcaagaag atgaagaact actggagaca gctgctgaac gcaaagctga tcacacagag 2760
 aaagttcgac aacctgacaa aggcagagag agggaggactg agcgaactgg acaaggcagg 2820
 attcatcaag agacagctgg tcgaaacaag acagatcaca aagcacgtcg cacagatcct 2880
 ggacagcaga atgaacacaa agtacgacga aaacgacaag ctgatcagag aagtcaaggt 2940
 catcacactg aagagcaagc tggtcagcga cttcagaaag gacttccagt tctacaaggt 3000
 cagagaaatc aacaactacc accacgcaca cgacgcatac ctgaacgcag tcgtcggaaac 3060
 agcactgatc aagaagtacc cgaagctgga aagcgaattc gtctacggag actacaaggt 3120
 ctacgacgtc agaaagatga tcgcaaagag cgaacaggaa atcggaaagg caacagcaaa 3180
 gtacttcttc tacagcaaca tcatgaactt cttcaagaca gaaatcacac tggcaaacgg 3240
 agaaatcaga aagagaccgc tgatcgaaac aaacggagaa acaggagaaa tcgtctggga 3300
 caagggaaga gacttcgcaa cagtcagaaa ggtcctgagc atgccgcagg tcaacatcgt 3360
 caagaagaca gaagtccaga caggaggatt cagcaaggaa agcatcctgc cgaagagaaa 3420
 cagcgacaag ctgatcgcaa gaaagaagga ctgggacccg aagaagtacg gaggattcga 3480
 cagcccagca gtcgcataca gcgtcctggt cgtcgcaaag gtcgaaaagg gaaagagcaa 3540
 gaagctgaag agcgtcaagg aactgctggg aatcacaatc atggaaagaa gcagcttcga 3600
 aaagaacccg atcgacttcc tggaagcaaa gggatacaag gaagtcaaga aggacctgat 3660
 catcaagctg ccgaagtaca gcctgttcga actggaaaac ggaagaaaga gaatgctggc 3720
 aagcgagga gaactgcaga agggaaacga actggcactg ccgagcaagt acgtcaactt 3780
 cctgtacctg gcaagccact acgaaaagct gaagggaagc ccggaagaca acgaacagaa 3840
 gcagctgttc gtcgaacagc acaagcacta cctggacgaa atcatcgaac agatcagcga 3900
 attcagcaag agagtcatcc tggcagacgc aaacctggac aaggctctga gcgcatacaa 3960
 caagcacaga gacaagccga tcagagaaca ggcagaaaac atcatccacc tgttcacact 4020

10

20

30

40

50

gacaaacctg ggagcaccgg cagcattcaa gtacttcgac acaacaatcg acagaaagag 4080
 atacacaagc acaaaggaag tcctggacgc aacactgac caccagagca tcacaggact 4140
 gtacgaaaca agaatcgacc tgagccagct gggaggagac ggaggaggaa gcccgaagaa 4200
 gaagagaaag gtctagctag ccatcacatt taaaagcatc tcagcctacc atgagaataa 4260
 gagaaagaaa atgaagatca atagcttatt catctctttt tctttttcgt tgggtgtaaag 4320
 ccaacaccct gtctaaaaaa cataaatctt ttaatcatt ttgcctctt tctctgtgct 4380
 tcaattaata aaaaatggaa agaacctcga g 4411

<210> 57

<211> 4481

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 57

gggcagatcg cctggagacg ccatccacgc tgttttgacc tccatagaag acaccgggac 60
 cgatccagcc tccgcggccg ggaacggtgc attggaacgc ggattccccg tgccaagagt 120
 gactcaccgt ccttgacacg gccaccatgg acaagaagta cagcatcgga ctggacatcg 180
 gaacaaacag cgtcggatgg gcagtcacat cagacgaata caaggtcccg agcaagaagt 240
 tcaaggtcct gggaaacaca gacagacaca gcatcaagaa gaacctgatc ggagcactgc 300
 tgttcgacag cggagaaaca gcagaagcaa caagactgaa gagaacagca agaagaagt 360
 acacaagaag aaagaacaga atctgctacc tgcaggaaat cttcagcaac gaaatggcaa 420
 aggtcgacga cagcttcttc cacagactgg aagaaagctt cctggtcgaa gaagacaaga 480
 agcacgaaag acaccgcatc ttcggaaaca tcgtcgacga agtcgcatac cacgaaaagt 540
 acccgacaat ctaccacctg agaaagaagc tggtcgacag cacagacaag gcagacctga 600
 gactgatcta cctggcactg gcacacatga tcaagttcag aggacacttc ctgatcgaag 660
 gagacctgaa cccggacaac agcgacgtcg acaagctgtt catccagctg gtccagacat 720
 acaaccagct gttcgaagaa aaccgcatca acgcaagcgg agtcgacgca aaggcaatcc 780
 tgagcgcaag actgagcaag agcagaagac tggaaaacct gatcgcacag ctgccgggag 840
 aaaagaagaa cggactgttc ggaaacctga tcgactgag cctgggactg acaccgaact 900
 tcaagagcaa cttcgacctg gcagaagacg caaagctgca gctgagcaag gacacatacg 960
 acgacgacct ggacaacctg ctggcacaga tcggagacca gtacgcagac ctgttcctgg 1020
 cagcaaagaa cctgagcgac gcaatcctgc tgagcgacat cctgagagtc aacacagaaa 1080
 tcacaaaggc accgctgagc gcaagcatga tcaagagata cgacgaacac caccaggacc 1140
 tgacactgct gaaggcactg gtcagacagc agctgccgga aaagtacaag gaaatcttct 1200
 tcgaccagag caagaacgga tacgcaggat acatcgacgg aggagcaagc caggaagaat 1260
 tctacaagtt catcaagccg atcctggaaa agatggacgg aacagaagaa ctgctggtca 1320
 agctgaacag agaagacctg ctgagaaaagc agagaacatt cgacaacgga agcatcccg 1380
 accagatcca cctgggagaa ctgcacgcaa tcctgagaag acaggaagac ttctaccctg 1440
 tcctgaagga caacagagaa aagatcgaaa agatcctgac attcagaatc cgtactacg 1500
 tcggaccgct ggcaagagga aacagcagat tcgcatggat gacaagaaaag agcgaagaaa 1560
 caatcacacc gtggaacttc gaagaagtcg tcgacaaggg agcaagcgca cagagcttca 1620
 tcgaaaagaat gacaaacttc gacaagaacc tgccgaacga aaaggtcctg ccgaagcaca 1680
 gcctgctgta cgaatacttc acagtctaca acgaactgac aaaggtcaag tacgtcacag 1740
 aaggaaatgag aaagccggca ttcctgagcg gagaacagaa gaaggcaatc gtcgacctgc 1800
 tgttcaagac aaacagaaaag gtcacagtca agcagctgaa ggaagactac ttcaagaaga 1860
 tcgaatgctt cgacagcgtc gaaatcagcg gagtcgaaga cagattcaac gcaagcctgg 1920
 gaacatacca cgacctgctg aagatcatca aggacaagga cttcctggac aacgaagaaa 1980

10

20

30

40

50

acgaagacat cctggaagac atcgtcctga cactgacact gttcgaagac agagaaatga 2040
 tcgaagaaag actgaagaca tacgcacacc tgttcgacga caaggatcatg aagcagctga 2100
 agagaagaag atacacagga tggggaagac tgagcagaaa gctgatcaac ggaatcagag 2160
 acaagcagag cggaaaagaca atcctggact tcctgaagag cgacggattc gcaaacagaa 2220
 acttcatgca gctgatccac gacgacagcc tgacattcaa ggaagacatc cagaaggcac 2280
 aggtcagcgg acagggagac agcctgcacg aacacatcgc aaacctggca ggaagcccgg 2340
 caatcaagaa gggaatcctg cagacagtca aggtcgtcga cgaactggtc aaggatcatgg 2400
 gaagacacaa gccggaaaac atcgtcatcg aaatggcaag agaaaaccag acaacacaga 2460
 agggacagaa gaacagcaga gaaagaatga agagaatcga agaaggaatc aaggaactgg 2520
 gaagccagat cctgaaggaa cacccggtcg aaaacacaca gctgcagaac gaaaagctgt 2580
 acctgtacta cctgcagaac ggaagagaca tgtacgtcga ccaggaactg gacatcaaca 2640
 gactgagcga ctacgacgtc gaccacatcg tccgcagag cttcctgaag gacgacagca 2700
 tcgacaacaa ggtcctgaca agaagcgaca agaacagagg aaagagcgac aacgtcccga 2760
 gcgaagaagt cgtcaagaag atgaagaact actggagaca gctgctgaac gcaaagctga 2820
 tcacacagag aaagttcgac aacctgacaa aggcagagag aggaggactg agcgaactgg 2880
 acaaggcagg attcatcaag agacagctgg tcgaaacaag acagatcaca aagcacgtcg 2940
 cacagatcct ggacagcaga atgaacacaa agtacgacga aaacgacaag ctgatcagag 3000
 aagtcaaggt catcacactg aagagcaagc tggtcagcga cttcagaaaag gacttccagt 3060
 tctacaaggt cagagaaatc aacaactacc accacgcaca cgacgcatac ctgaacgcag 3120
 tcgtcggaaac agcactgatc aagaagtacc cgaagctgga aagcgaattc gtctacggag 3180
 actacaaggt ctacgacgtc agaaagatga tcgcaaagag cgaacaggaa atcggaagg 3240
 caacagcaaa gtacttcttc tacagcaaca tcatgaactt cttcaagaca gaaatcacac 3300
 tggcaaacgg agaaatcaga aagagaccgc tgatcgaaac aaacggagaa acaggagaaa 3360
 tcgtctggga caagggaaga gacttcgcaa cagtcagaaa ggtcctgagc atgccgcagg 3420
 tcaacatcgt caagaagaca gaagtccaga caggaggatt cagcaaggaa agcatcctgc 3480
 cgaagagaaa cagcgacaag ctgatcgcaa gaaagaagga ctgggacccg aagaagtacg 3540
 gaggattcga cagcccgaac gtcgcataca gcgtcctggc cgtcgcaaag gtcgaaaagg 3600
 gaaagagcaa gaagctgaag agcgtcaagg aactgctggg aatcacaatc atggaaagaa 3660
 gcagcttcca aaagaacccg atcgacttcc tggaaagcaa gggatacaag gaagtcaaga 3720
 aggacctgat catcaagctg ccgaagtaca gcctgttcga actggaaaac ggaagaaaga 3780
 gaatgctggc aagcgacgga gaactgcaga agggaaacga actggcactg ccgagcaagt 3840
 acgtcaactt cctgtacctg gcaagccact acgaaaagct gaagggaagc ccggaagaca 3900
 acgaacagaa gcagctgttc gtcgaacagc acaagcacta cctggacgaa atcatcgaac 3960
 agatcagcga attcagcaag agagtcattc tggcagacgc aaacctggac aaggtcctga 4020
 gcgcatacaa caagcacaga gacaagccga tcagagaaca ggcagaaaac atcatccacc 4080
 tggtcacact gacaaacctg ggagcaccgg cagcattcaa gtacttcgac acaacaatcg 4140
 acagaaagag atacacaagc acaaggaag tcctggacgc aacactgatc caccagagca 4200
 tcacaggact gtacgaaaca agaatcgacc tgagccagct gggaggagac ggaggaggaa 4260
 gcccgaagaa gaagagaaag gtctagctag ccatcacatt taaaagcatc tcagcctacc 4320
 atgagaataa gagaaagaaa atgaagatca atagcttatt catctctttt tcttttctgt 4380
 tgggtgtaaag ccaacaccct gtctaaaaaa cataaatttc ttaaatcatt ttgcctcttt 4440
 tctctgtgct tcaattaata aaaaatggaa agaacctcga g 4481

10

20

30

40

<210> 58

<211> 4348

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic

50

polynucleotide"

<400> 58

gggacatttg cttctgacac aactgtgttc actagcaacc tcaaacagac accggatctg	60	
ccaccatgga caagaagtac agcatcggac tggacatcgg aacaaacagc gtcggatggg	120	
cagtcatcac agacgaatac aaggtcccg a gcaagaagtt caaggtcctg ggaaacacag	180	
acagacacag catcaagaag aacctgatcg gagcactgct gttcgacagc ggagaaacag	240	
cagaagcaac aagactgaag agaacagcaa gaagaagata cacaagaaga aagaacagaa	300	
tctgctacct gcaggaaatc ttcagcaacg aaatggcaaa ggtcgacgac agcttcttcc	360	
acagactgga agaaagcttc ctggtcgaag aagacaagaa gcacgaaaga caccgatct	420	
tcggaaacat cgtcgacgaa gtcgcatacc acgaaaagta cccgacaatc taccacctga	480	10
gaaagaagct ggtcgacagc acagacaagg cagacctgag actgatctac ctggcactgg	540	
cacacatgat caagttcaga ggacacttcc tgatcgaagg agacctgaac ccggacaaca	600	
gcgacgtcga caagctgttc atccagctgg tccagacata caaccagctg ttcgaagaaa	660	
acccgatcaa cgcaagcgga gtcgacgcaa aggcaatcct gagcgcaaga ctgagcaaga	720	
gcagaagact ggaaaacctg atcgcacagc tgccgggaga aaagaagaac ggactgttcg	780	
gaaacctgat cgactgagc ctgggactga caccgaactt caagagcaac ttcgacctgg	840	
cagaagacgc aaagctgcag ctgagcaagg acacatacga cgacgacctg gacaacctgc	900	
tggcacagat cggagaccag tacgcagacc tgttcctggc agcaaagaac ctgagcgacg	960	
caatcctgct gagcgacatc ctgagagtca acacagaaat cacaaggca ccgctgagcg	1020	
caagcatgat caagagatac gacgaacacc accaggacct gacactgctg aaggcactgg	1080	20
tcagacagca gctgccggaa aagtacaagg aaatcttctt cgaccagagc aagaacggat	1140	
acgcaggata catcgacgga ggagcaagcc aggaagaatt ctacaagttc atcaagccga	1200	
tcctggaaaa gatggacgga acagaagaac tgctgggtcaa gctgaacaga gaagacctgc	1260	
tgagaaagca gagaacattc gacaacggaa gcatcccgca ccagatccac ctgggagaac	1320	
tgcacgcaat cctgagaaga caggaagact tctaccggtt cctgaaggac aacagagaaa	1380	
agatcgaaaa gatcctgaca ttcagaatcc cgtactacgt cggaccgctg gcaagaggaa	1440	
acagcagatt cgcatggatg acaagaaaga gcgaagaaac aatcacaccg tggaacttcg	1500	
aagaagtcgt cgacaaggga gcaagcgcac agagcttcat cgaaagaatg acaaacttcg	1560	
acaagaacct gccgaacgaa aaggtcctgc cgaagcacag cctgctgtac gaatacttca	1620	
cagtctacaa cgaactgaca aaggtcaagt acgtcacaga aggaatgaga aagccggcat	1680	30
tcctgagcgg agaacagaag aaggcaatcg tcgacctgct gttcaagaca aacagaaagg	1740	
tcacagtcaa gcagctgaag gaagactact tcaagaagat cgaatgcttc gacagcgtcg	1800	
aaatcagcgg agtcgaagac agattcaacg caagcctggg aacataccac gacctgctga	1860	
agatcatcaa ggacaaggac ttcctggaca acgaagaaaa cgaagacatc ctggaagaca	1920	
tcgtcctgac actgacactg ttcgaagaca gagaaatgat cgaagaaaga ctgaagacat	1980	
acgcacacct gttcgacgac aaggtcatga agcagctgaa gagaagaaga tacacaggat	2040	
ggggaagact gagcagaaag ctgatcaacg gaatcagaga caagcagagc ggaaagacaa	2100	
tcctggactt cctgaagagc gacggattcg caaacagaaa ctcatgcag ctgatccacg	2160	
acgacagcct gacattcaag gaagacatcc agaaggcaca ggtcagcgga caggagagaca	2220	
gcctgcacga acacatcgca aacctggcag gaagcccggc aatcaagaag ggaatcctgc	2280	40
agacagtcaa ggtcgtcgac gaactgggtca aggtcatggg aagacacaag ccggaaaaca	2340	
tcgtcatcga aatggcaaga gaaaaccaga caacacagaa gggacagaag aacagcagag	2400	
aaagaatgaa gagaatcgaa gaaggaatca aggaactggg aagccagatc ctgaaggaac	2460	
acccggctga aaacacacag ctgcagaacg aaaagctgta cctgtactac ctgcagaacg	2520	
gaagagacat gtacgtcgac caggaactgg acatcaacag actgagcgac tacgacgtcg	2580	
accacatcgt cccgcagagc ttcctgaagg acgacagcat cgacaacaag gtcctgacaa	2640	
gaagcgacaa gaacagagga aagagcgaca acgtcccag cgaagaagtc gtcaagaaga	2700	
tgaagaacta ctggagacag ctgctgaacg caaagctgat cacacagaga aagttcgaca	2760	
acctgacaaa ggcagagaga ggaggactga gcgaactgga caaggcagga ttcacatga	2820	
gacagctggt cgaaacaaga cagatcaca agcacgtcgc acagatcctg gacagcagaa	2880	50

tgaacacaaa gtacgacgaa aacgacaagc tgatcagaga agtcaagggtc atcacactga 2940
 agagcaagct ggtagcgac ttcagaaagg acttccagtt ctacaagggtc agagaaatca 3000
 acaactacca ccacgcacac gacgcatacc tgaacgcagt cgtcggaaaca gacttgatca 3060
 agaagtaccc gaagctggaa agcgaattcg tctacggaga ctacaagggtc tacgacgtca 3120
 gaaagatgat cgcaaagagc gaacaggaaa tcggaaaggc aacagcaaag tacttcttct 3180
 acagcaacat catgaacttc ttcaagacag aaatcacact ggcaaacgga gaaatcagaa 3240
 agagaccgct gatcgaacaa aacggagaaa caggagaaat cgtctgggac aagggaagag 3300
 acttcgcaac agtcagaaaag gtcctgagca tgccgcagggt caacatcgtc aagaagacag 3360
 aagtccagac aggaggattc agcaaggaaa gcatcctgcc gaagagaaac agcgacaagc 3420
 tgatcgcaag aaagaaggac tgggacccga agaagtacgg aggattcgac agcccgcagc 3480
 tcgcatacag cgtcctgggtc gtcgcaaagg tcgaaaaggg aaagagcaag aagctgaaga 3540
 gcgtcaagga actgctggga atcacaatca tggaaagaag cagcttcgaa aagaaccgga 3600
 tcgacttcct ggaagcaaag ggatacaagg aagtcaagaa ggacctgatc atcaagctgc 3660
 cgaagtacag cctgttcgaa ctggaaaacg gaagaaagag aatgctggca agcgaggag 3720
 aactgcagaa gggaaacgaa ctggcactgc cgagcaagta cgtcaacttc ctgtacctgg 3780
 caagccacta cgaaaagctg aagggaagcc cggaagacaa cgaacagaag cagctgttcg 3840
 tcgaacagca caagcactac ctggacgaaa tcatcgaaca gatcagcgaa ttcagcaaga 3900
 gagtcatcct ggcagacgca aacctggaca aggtcctgag cgatacaac aagcacagag 3960
 acaagccgat cagagaacag gcagaaaaca tcatccacct gtacacactg acaaacctgg 4020
 gagcaccggc agcattcaag tacttcgaca caacaatcga cagaaagaga tacacaagca 4080
 caaaggaagt cctggacgca acactgatcc accagagcat cacaggactg tacgaaacaa 4140
 gaatcgacct gagccagctg ggaggagacg gaggaggaag cccgaagaag aagagaaagg 4200
 tctagctagc gctcgcttcc ttgctgtcca atttctatta aagggtcctt tgttcctaa 4260
 gtccaactac taaactgggg gatattatga agggccttga gcatctggat tctgccta 4320
 aaaaaacatt tattttcatt gcctcgag 4348

<210> 59

<211> 4325

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 59

gggaagctca gaataaacgc tcaactttgg ccggatctgc caccatggac aagaagtaca 60
 gcatcggact ggacatcgga acaaacagcg tcggatgggc agtcatacaca gacgaataca 120
 aggtcccagc caagaagttc aaggctcctgg gaaacacaga cagacacagc atcaagaaga 180
 acctgatcgg agcactgctg ttcgacagcg gagaaacagc agaagcaaca agactgaaga 240
 gaacagcaag aagaagatac acaagaagaa agaacagaat ctgctacctg caggaaatct 300
 tcagcaacga aatggcaaag gtcgacgaca gcttcttcca cagactggaa gaaagcttcc 360
 tggctgaaga agacaagaag cacgaaagac acccgatctt cggaacatc gtcgacgaag 420
 tcgcatacca cgaaaagtac ccgacaatct accacctgag aaagaagctg gtcgacagca 480
 cagacaaggc agacctgaga ctgatctacc tggcactggc acacatgatc aagttcagag 540
 gacacttcct gatcgaagga gacctgaacc cggacaacag cgacgtcgac aagctgttca 600
 tccagctggt ccagacatac aaccagctgt tcgaagaaaa cccgatcaac gcaagcggag 660
 tcgacgcaaa ggcaatcctg agcgcaagac tgagcaagag cagaagactg gaaaacctga 720
 tcgcacagct gccgggagaa aagaagaacg gactgttcgg aaacctgatc gactgagcc 780
 tgggactgac accgaacttc aagagcaact tcgacctggc agaagacgca aagctgcagc 840
 tgagcaagga cacatacgac gacgacctgg acaacctgct ggcacagatc ggagaccagt 900

acgcagacct gttcctggca gcaaagaacc tgagcgacgc aatcctgctg agcgacatcc	960	
tgagagtcaa cacagaaatc acaaaggcac cgctgagcgc aagcatgatac aagagatacg	1020	
acgaacacca ccaggacctg acactgctga aggcactggt cagacagcag ctgccggaaa	1080	
agtacaagga aatcttcttc gaccagagca agaacggata cgcaggatac atcgacggag	1140	
gagcaagcca ggaagaattc tacaagtcca tcaagccgat cctggaaaag atggacggaa	1200	
cagaagaact gctgggtcaag ctgaacagag aagacctgct gagaaagcag agaacattcg	1260	
acaacggaag catcccgcac cagatccacc tgggagaact gcacgcaatc ctgagaagac	1320	
aggaagactt ctacccgttc ctgaaggaca acagagaaaa gatcgaaaag atcctgacat	1380	
tcagaatccc gtactacgtc ggaccgctgg caagaggaaa cagcagattc gcatggatga	1440	
caagaaagag cgaagaaaca atcacaccgt ggaacttcga agaagtcgtc gacaagggag	1500	10
caagcgcaca gagcttcatac gaaagaatga caaacttcga caagaacctg ccgaacgaaa	1560	
aggtcctgcc gaagcacagc ctgctgtacg aatacttcac agtctacaac gaactgacaa	1620	
aggtcaagta cgtcacagaa ggaatgagaa agccggcatt cctgagcggg gaacagaaga	1680	
aggcaatcgt cgacctgctg ttcaagacaa acagaaaagg cacagtcaag cagctgaagg	1740	
aagactactt caagaagatc gaatgcttcg acagcgtcga aatcagcggg gtcgaagaca	1800	
gattcaacgc aagcctggga acataccacg acctgctgaa gatcatcaag gacaaggact	1860	
tcctggacaa cgaagaaaac gaagacatcc tggaagacat cgtcctgaca ctgacactgt	1920	
tcgaagacag agaaatgatc gaagaaagac tgaagacata cgcacacctg ttcgacgaca	1980	
aggtcatgaa gcagctgaag agaagaagat acacaggatg gggaagactg agcagaaagc	2040	
tgatcaacgg aatcagagac aagcagagcg gaaagacaat cctggacttc ctgaagagcg	2100	20
acggattcgc aaacagaaaac ttcatgcagc tgatccacga cgacagcctg acattcaagg	2160	
aagacatcca gaaggcacag gtcagcggac agggagacag cctgcacgaa cacatcgcaa	2220	
acctggcagg aagcccggca atcaagaagg gaatcctgca gacagtcaag gtcgtcgacg	2280	
aactgggtcaa ggtcatggga agacacaagc cggaaaacat cgtcatcgaa atggcaagag	2340	
aaaaccagac aacacagaag ggacagaaga acagcagaga aagaatgaag agaatcgaag	2400	
aaggaatcaa ggaactggga agccagatcc tgaaggaaca cccggtcgaa aacacacagc	2460	
tgcagaacga aaagctgtac ctgtactacc tgcagaacgg aagagacatg tacgtcgacc	2520	
aggaactgga catcaacaga ctgagcgact acgacgtcga ccacatcgtc ccgcagagct	2580	
tcctgaagga cgacagcatc gacaacaagg tcctgacaag aagcgacaag aacagaggaa	2640	
agagcgacaa cgtcccagac gaagaagtcg tcaagaagat gaagaactac tggagacagc	2700	30
tgctgaacgc aaagctgatac acacagagaa agttcgacaa cctgacaaaag gcagagagag	2760	
gaggactgag cgaactggac aaggcaggat tcatcaagag acagctgggtc gaaacaagac	2820	
agatcacaaa gcacgtcgca cagatcctgg acagcagaat gaacacaaaag tacgacgaaa	2880	
acgacaagct gatcagagaa gtcaagggtca tcacactgaa gagcaagctg gtcagcgact	2940	
tcagaaagga cttccagttc tacaagggtca gagaaatcaa caactaccac cacgcacacg	3000	
acgcatacct gaacgcagtc gtcggaacag cactgatcaa gaagtacccg aagctggaaa	3060	
gcgaattcgt ctacggagac tacaagggtc acgacgtcag aaagatgatac gcaaagagcg	3120	
aacaggaaat cggaaaggca acagcaaagt acttcttcta cagcaacatc atgaacttct	3180	
tcaagacaga aatcacactg gcaaacggag aatcagaaa gagaccgctg atcgaaacaa	3240	
acggagaaac aggagaaatc gtctgggaca aggggaagaga cttcgcaaca gtcagaaagg	3300	40
tcctgagcat gccgcaggtc aacatcgtca agaagacaga agtccagaca ggaggattca	3360	
gcaaggaaaag catcctgccg aagagaaaaca gcgacaagct gatcgcaaga aagaaggact	3420	
gggacccgaa gaagtacgga ggattcgaca gcccgacagt cgcatacagc gtcctgggtcg	3480	
tcgcaaaagg cgaagaggga aagagcaaga agctgaagag cgtcaaggaa ctgctgggaa	3540	
tcacaatcat ggaaagaagc agcttcgaaa agaaccgat cgacttcctg gaagcaaagg	3600	
gatacaagga agtcaagaag gacctgatca tcaagctgcc gaagtacagc ctgttcgaac	3660	
tggaaaacgg aagaaagaga atgctggcaa gcgcaggaga actgcagaag ggaaacgaac	3720	
tggcactgcc gagcaagtac gtcaacttcc tgtacctggc aagccactac gaaaagctga	3780	
aggggaagccc ggaagacaac gaacagaagc agctgttcgt cgaacagcac aagcactacc	3840	
tggacgaaat catcgaacag atcagcgaat tcagcaagag agtcatcctg gcagacgcaa	3900	50

```

acctggacaa ggtcctgagc gcatacaaca agcacagaga caagccgatac agagaacagg 3960
cagaaaacat catccacctg ttcacactga caaacctggg agcaccggca gcattcaagt 4020
acttcgacac aacaatcgac agaaagagat acacaagcac aaaggaagtc ctggacgcaa 4080
cactgatcca ccagagcatc acaggactgt acgaaacaag aatcgacctg agccagctgg 4140
gaggagacgg aggaggaagc ccgaagaaga agagaaaggt ctagctagca ccagcctcaa 4200
gaacacccga atggagtctc taagctacat aataccaact tacactttac aaaatgttgt 4260
ccccaaaaat gtagccattc gtatctgctc ctaataaaaa gaaagtttct tcacattctc 4320
tcgag 4325

```

<210> 60

10

<211> 4325

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 60

```

aggaagctca gaataaacgc tcaactttgg ccggatctgc caccatggac aagaagtaca 60
gcatcggact ggacatcgga acaaacagcg tcggatgggc agtcatcaca gacgaataca 120
aggtcccagc caagaagttc aaggctctgg gaaacacaga cagacacagc atcaagaaga 180
acctgatcgg agcactgctg ttcgacagcg gagaaacagc agaagcaaca agactgaaga 240
gaacagcaag aagaagatac acaagaagaa agaacagaat ctgctacctg caggaaatct 300
tcagcaacga aatggcaaag gtcgacgaca gcttcttcca cagactggaa gaaagcttcc 360
tggtcgaaga agacaagaag cagcaaagac acccgatctt cggaacatc gtcgacgaag 420
tcgcatacca cgaaaagtac ccgacaatct accacctgag aaagaagctg gtcgacagca 480
cagacaaggc agacctgaga ctgatctacc tggcactggc acacatgatac aagttcagag 540
gacacttcct gatcgaagga gacctgaacc cggacaacag cgacgtcgac aagctgttca 600
tccagctggg ccagacatac aaccagctgt tcgaagaaaa cccgatcaac gcaagcggag 660
tcgacgcaaa ggcaatctg agcgcaagac tgagcaagag cagaagactg gaaaacctga 720
tcgcacagct gccgggagaa aagaagaacg gactgttcgg aaacctgata gactgagcc 780
tgggactgac accgaacttc aagagcaact tcgacctggc agaagacgca aagctgcagc 840
tgagcaagga cacatacgac gacgacctgg acaacctgct ggcacagatac ggagaccagt 900
acgcagacct gttcctggca gcaaagaacc tgagcgacgc aatcctgctg agcgacatcc 960
tgagagtcaa cacagaaatc acaaaggcac cgctgagcgc aagcatgatac aagagatacg 1020
acgaacacca ccaggacctg aactgctga aggcactggg cagacagcag ctgccggaaa 1080
agtacaagga aatcttcttc gaccagagca agaacggata cgcaggatac atcgacggag 1140
gagcaagcca ggaagaattc tacaagtcca tcaagccgat cctggaaaag atggacggaa 1200
cagaagaact gctggtaag ctgaacagag aagacctgct gagaaagcag agaacattcg 1260
acaacggaag catccgcac cagatccacc tgggagaact gcacgcaatc ctgagaagac 1320
aggaagactt ctacccgttc ctgaaggaca acagagaaaa gatcgaaaag atcctgacat 1380
tcagaatccc gtactacgtc ggaccgctgg caagaggaaa cagcagattc gcatggatga 1440
caagaaagag cgaagaaaca atcacaccgt ggaacttcga agaagtcgtc gacaaggagg 1500
caagcgcaac gagcttcata gaaagaatga caaacttcga caagaacctg ccgaacgaaa 1560
aggtcctgcc gaagcacagc ctgctgtacg aatacttcac agtctacaac gaactgacaa 1620
aggtcaagta cgtcacagaa ggaatgagaa agccggcatt cctgagcggg gaacagaaga 1680
aggcaatcgt cgacctgctg ttcaagacaa acagaaaggt cacagtcaag cagctgaagg 1740
aagactactt caagaagatac gaatgcttcg acagcgctga aatcagcgga gtcgaagaca 1800
gattcaacgc aagcctggga acataccacg acctgctgaa gatcatcaag gacaaggact 1860
tcctggacaa cgaagaaaac gaagacatcc tggaagacat cgtcctgaca ctgacactgt 1920

```

20

30

40

50

tcgaagacag agaaatgatac gaagaaagac tgaagacata cgcacacctg ttcgacgaca	1980	
aggtcatgaa gcagctgaag agaagaagat acacaggatg gggaagactg agcagaaagc	2040	
tgatcaacgg aatcagagac aagcagagcg gaaagacaat cctggacttc ctgaagagcg	2100	
acggattcgc aaacagaaaac ttcattgcagc tgatccacga cgacagcctg acattcaagg	2160	
aagacatcca gaaggcacag gtcagcggac agggagacag cctgcacgaa cacatcgcaa	2220	
acctggcagg aagcccggca atcaagaagg gaatcctgca gacagtcaag gtcgtcgacg	2280	
aactgggtcaa ggtcatggga agacacaagc cggaaaacat cgtcatcgaa atggcaagag	2340	
aaaaccagac aacacagaag ggacagaaga acagcagaga aagaatgaag agaatcgaag	2400	
aaggaatcaa ggaactggga agccagatcc tgaaggaaca cccggtcgaa aacacacagc	2460	
tgcagaacga aaagctgtac ctgtactacc tgcagaacgg aagagacatg tacgtcgacc	2520	10
aggaactgga catcaacaga ctgagcagct acgacgtcga ccacatcgtc ccgcagagct	2580	
tcctgaagga cgacagcatt gacaacaagg tcctgacaag aagcgacaag aacagaggaa	2640	
agagcgacaa cgtcccggagc gaagaagtcg tcaagaagat gaagaactac tggagacagc	2700	
tgctgaacgc aaagctgata acacagagaa agttcgacaa cctgacaaag gcagagagag	2760	
gaggactgag cgaactggac aaggcaggat tcatcaagag acagctggtc gaaacaagac	2820	
agatcacaaa gcacgtcgca cagatcctgg acagcagaat gaacacaaaag tacgacgaaa	2880	
acgacaagct gatcagagaa gtcaagggtca tcacactgaa gagcaagctg gtcagcagct	2940	
tcagaaagga cttccagttc tacaagggtca gagaaatcaa caactaccac cagcacacg	3000	
acgcatacct gaacgcagtc gtcggaacag cactgatcaa gaagtacccg aagctggaaa	3060	
gcgaattcgt ctacggagac tacaagggtc acgacgtcag aaagatgata gcaaagagcg	3120	20
aacaggaaat cggaaaggca acagcaaagt acttcttcta cagcaacatc atgaacttct	3180	
tcaagacaga aatcacactg gcaaacggag aaatcagaaa gagaccgctg atcgaaacaa	3240	
acggagaaac aggagaaatc gtctgggaca aggggaagaga cttcgcaaca gtcagaaagg	3300	
tcctgagcat gccgcaggtc aacatcgta agaagacaga agtccagaca ggaggattca	3360	
gcaaggaaag catcctgccg aagagaaaca gcgacaagct gatcgcaaga aagaaggact	3420	
gggacccgaa gaagtacgga ggattcgaca gcccagacag cgcatacagc gtcctggtcg	3480	
tcgcaaaggc gaaaaggga aagagcaaga agctgaagag cgtcaaggaa ctgctgggaa	3540	
tcacaatcat ggaaagaagc agcttcgaaa agaaccgat cgacttctg gaagcaaagg	3600	
gatacaagga agtcaagaag gacctgatca tcaagctgcc gaagtacagc ctgttcgaac	3660	
tggaaaacgg aagaaagaga atgctggcaa gcgcaggaga actgcagaag ggaaacgaac	3720	30
tggcactgcc gagcaagtac gtcaacttcc tgtacctggc aagccactac gaaaagctga	3780	
agggaaagccc ggaagacaac gaacagaagc agctgttcgt cgaacagcac aagcactacc	3840	
tggacgaaat catcgaacag atcagcgaat tcagcaagag agtcatcctg gcagacgcaa	3900	
acctggacaa ggtcctgagc gcatacaaca agcacagaga caagccgatc agagaacagg	3960	
cagaaaacat catccacctg ttcacactga caaacctggg agcaccggca gcattcaagt	4020	
acttcgacac aacaatcgac agaaagagat acacaagcac aaagggaagtc ctggacgcaa	4080	
cactgatcca ccagagcatt acaggactgt acgaaacaag aatcgacctg agccagctgg	4140	
gaggagacgg agggaggaagc ccgaagaaga agagaaaggc ctagctagca ccagcctcaa	4200	
gaacaccgga atggagtctc taagctacat aataccaact tacactttac aaaatgttgt	4260	
ccccaaaaat gtagccattc gtatctgctc ctaataaaaa gaaagtttct tcacattctc	4320	40
tcgag	4325	

<210> 61

<211> 4411

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 61

aggtcccgcga gtcggcgtcc agcggctctg cttgttcgtg tgtgtgtcgt tgcaggcctt	60	
attcggatcc gccacatgg acaagaagta cagcatcgga ctggacatcg gaacaaacag	120	
cgtcggatgg gcagtcata cagacgaata caaggtcccg agcaagaagt tcaaggtcct	180	
gggaaacaca gacagacaca gcatcaagaa gaacctgacg ggagcactgc tgttcgacag	240	
cggagaaaca gcagaagcaa caagactgaa gagaacagca agaagaagat acacaagaag	300	
aaagaacaga atctgctacc tgcaggaaat cttcagcaac gaaatggcaa aggtcgacga	360	
cagcttcttc cacagactgg aagaaagctt cctgggtcgaa gaagacaaga agcacgaaag	420	
acacccgatc ttcggaacaa tcgtcgacga agtcgcatac cacgaaaagt acccgacaat	480	
ctaccacctg agaaagaagc tggtcgacag cacagacaag gcagacctga gactgatcta	540	10
cctggcactg gcacacatga tcaagttcag aggacacttc ctgatcgaag gagacctgaa	600	
cccgacaac agcgacgtcg acaagctgtt catccagctg gtccagacat acaaccagct	660	
gttcgaagaa aacccgatca acgcaagcgg agtcgacgca aaggcaatcc tgagcgcaag	720	
actgagcaag agcagaagac tggaaaacct gatcgcacag ctgccggggag aaaagaagaa	780	
cggactgttc ggaaacctga tcgcaactgag cctgggactg acaccgaact tcaagagcaa	840	
cttcgacctg gcagaagacg caaagctgca gctgagcaag gacacatacg acgacgacct	900	
ggacaacctg ctggcacaga tcggagacca gtacgcagac ctgttcttgg cagcaaagaa	960	
cctgagcgac gcaatcctgc tgagcgacat cctgagagtc aacacagaaa tcacaaaggc	1020	
accgctgagc gcaagcatga tcaagagata cgacgaacac caccaggacc tgacactgct	1080	
gaaggcactg gtcagacagc agctgccgga aaagtacaag gaaatcttct tcgaccagag	1140	20
caagaacgga tacgcaggat acatcgacgg aggagcaagc caggaagaat tctacaagtt	1200	
catcaagccg atcctggaaa agatggacgg aacagaagaa ctgctggtca agctgaacag	1260	
agaagacctg ctgagaaagc agagaacatt cgacaacgga agcatcccgc accagatcca	1320	
cctgggagaa ctgcacgcaa tcctgagaag acaggaagac ttctaccctg tcctgaagga	1380	
caacagagaa aagatcgaag agatcctgac attcagaatc ccgtactacg tcggaccgct	1440	
ggcaagagga aacagcagat tcgcatggat gacaagaaag agcgaagaaa caatcacacc	1500	
gtggaacttc gaagaagtcg tcgacaaggg agcaagcgca cagagcttca tcgaaagaat	1560	
gacaaacttc gacaagaacc tgccgaacga aaaggctcctg ccgaagcaca gcctgctgta	1620	
cgaatacttc acagtctaca acgaactgac aaagggtcaag tacgtcacag aaggaatgag	1680	
aaagccggca ttcctgagcg gagaacagaa gaaggcaatc gtcgacctgc tgttcaagac	1740	30
aaacagaaag gtcacagtca agcagctgaa ggaagactac ttcaagaaga tcgaatgctt	1800	
cgacagcgtc gaaatcagcg gagtcgaaga cagattcaac gcaagcctgg gaacatacca	1860	
cgacctgctg aagatcatca aggacaagga cttcctggac aacgaagaaa acgaagacat	1920	
cctggaagac atcgtcctga cactgacact gttcgaagac agagaaatga tcgaagaaag	1980	
actgaagaca tacgcacacc tgttcgacga caaggtcatg aagcagctga agagaagaag	2040	
atacacagga tggggaagac tgagcagaaa gctgatcaac ggaatcagag acaagcagag	2100	
cggaaagaca atcctggact tcctgaagag cgacggattc gcaaacagaa acttcatgca	2160	
gctgatccac gacgacagcc tgacattcaa ggaagacatc cagaaggcac aggtcagcgg	2220	
acaggagagc agcctgcacg aacacatcgc aaacctggca ggaagcccgg caatcaagaa	2280	
gggaatcctg cagacagtca aggtcgtcga cgaactggtc aaggtcatgg gaagacacaa	2340	40
gccggaaaac atcgtcatcg aaatggcaag agaaaaccag acaacacaga agggacagaa	2400	
gaacagcaga gaaagaatga agagaatcga agaaggaatc aaggaactgg gaagccagat	2460	
cctgaaggaa cccccggtcg aaaacacaca gctgcagaac gaaaagctgt acctgtacta	2520	
cctgcagaac ggaagagaca tgtacgtcga ccaggaactg gacatcaaca gactgagcga	2580	
ctacgacgtc gaccacatcg tcccgcagag cttcctgaag gacgacagca tcgacaacaa	2640	
ggtcctgaca agaagcgaca agaacagagg aaagagcgac aacgtcccga gcgaagaagt	2700	
cgtcaagaag atgaagaact actggagaca gctgctgaac gcaaagctga tcacacagag	2760	
aaagttcgac aacctgacaa aggcagagag aggaggactg agcgaactgg acaaggcagg	2820	
attcatcaag agacagctgg tcgaaacaag acagatcaca aagcacgtcg cacagatcct	2880	
ggacagcaga atgaacacaa agtacgacga aaacgacaag ctgatcagag aagtcaaggt	2940	50


```

catcacactg aagagcaagc tggtcagcga cttcagaaag gacttccagt tctacaaggt 3000
cagagaaatc aacaactacc accacgcaca cgacgcatac ctgaacgcag tcgtcggaac 3060
agcactgata aagaagtacc cgaagctgga aagcgaattc gtctacggag actacaaggt 3120
ctacgacgtc agaaagatga tcgcaaagag cgaacaggaa atcggaagg caacagcaaa 3180
gtacttcttc tacagcaaca tcatgaactt cttcaagaca gaaatcacac tggcaaacgg 3240
agaaatcaga aagagaccgc tgatcgaaac aaacggagaa acaggagaaa tcgtctggga 3300
caagggaaga gacttcgcaa cagtcagaaa ggtcctgagc atgccgcagg tcaacatcgt 3360
caagaagaca gaagtccaga caggaggatt cagcaaggaa agcatcctgc cgaagagaaa 3420
cagcgacaag ctgatcgcaa gaaagaagga ctgggacccg aagaagtacg gaggattcga 3480
cagcccgaca gtcgcataca gcgtcctggt cgtcgcaaag gtcgaaaagg gaaagagcaa 3540
gaagctgaag agcgtcaagg aactgctggg aatcacaatc atggaaaaga gcagcttcga 3600
aaagaacccg atcgacttcc tggaagcaaa gggatacaag gaagtcaaga aggacctgat 3660
catcaagctg ccgaagtaca gcctgttcga actggaaaac ggaagaaaaga gaatgtggc 3720
aagcgagga gaactgcaga agggaaacga actggcactg ccgagcaagt acgtcaactt 3780
cctgtacctg gcaagccact acgaaaagct gaagggaagc ccggaagaca acgaacagaa 3840
gcagctgttc gtcgaacagc acaagcacta cctggacgaa atcatcgaac agatcagcga 3900
attcagcaag agagtcattc tggcagacgc aaacctggac aaggtcctga gcgcatacaa 3960
caagcacaga gacaagccga tcagagaaca ggcagaaaac atcatccacc tgttcacact 4020
gacaaacctg ggagcaccgg cagcattcaa gtacttcgac acaacaatcg acagaaagag 4080
atacacaagc acaaaggaag tcctggacgc aacactgata caccagagca tcacaggact 4140
gtacgaaaca agaatcgacc tgagccagct gggaggagac ggaggaggaa gcccgaagaa 4200
gaagagaaaag gtctagctag ccatcacatt taaaagcatt tcagcctacc atgagaataa 4260
gagaaagaaa atgaagatca atagcttatt catctctttt tcttttctgt tgggtgtaaag 4320
ccaacaccct gtctaaaaaa cataaatctt ttaaatcatt ttgcctcttt tctctgtgct 4380
tcaattaata aaaaatggaa agaacctcga g 4411

```

<210> 62

<211> 105

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 62

```

aaaaaaaaa aaaaaaaaaa aaaaaaaaaa gcgaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60
aaaccgaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 105

```

<210> 63

<211> 93

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 63

```

aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 93

```

10

20

30

40

50

<210> 64
 <211> 100
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<220>
 <221> modified_base
 <222> (1)..(3)
 <223> 2'-O-methyl modified
 <220>
 <221> misc_feature
 <222> (1)..(4)
 <223> Phosphorothioate linkage
 <220>
 <221> modified_base
 <222> (97)..(99)
 <223> 2'-O-methyl modified
 <220>
 <221> misc_feature
 <222> (97)..(100)
 <223> Phosphorothioate linkage
 <400> 64

ccaguccagc gaggcaaagg guuuuagagc uagaaaauagc aaguuaaaau aaggcuaguc 60
 cguauaacaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 65
 <211> 3312
 <212> DNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"
 <400> 65

atggcagcat tcaagccgaa ctgatcaac tacatcctgg gactggacat cggaatcgca 60
 tcggtcggat gggcaatggt cgaaatcgac gaagaagaaa acccgatcag actgatcgac 120
 ctgggagtca gagtcttcga aagagcagaa gtcccgaaga caggagactc gctggcaatg 180
 gcaagaagac tggcaagatc ggtcagaaga ctgacaagaa gaagagcaca cagactgctg 240
 agaacaagaa gactgctgaa gagagaagga gtcctgcagg cagcaaactt cgacgaaaac 300
 ggactgatca agtcgctgcc gaacacaccg tggcagctga gagcagcagc actggacaga 360
 aagctgacac cgctggaatg gtcggcagtc ctgctgcacc tgatcaagca cagaggatac 420
 ctgtcgcaga gaaagaacga aggagaaaca gcagacaagg aactgggagc actgctgaag 480
 ggagtcgcag gaaacgcaca cgactgcag acaggagact tcagaacacc ggcagaactg 540
 gcactgaaca agttcgaaaa ggaatcgga cacatcagaa accagagatc ggactactcg 600
 cacacattct cgagaaagga cctgcaggca gaactgatcc tgctgttcga aaagcagaag 660
 gaattcggaa acccgcacgt ctcgggagga ctgaaggaag gaatcgaaac actgctgatg 720

acacagagac cggcactgtc gggagacgca gtccagaaga tgctgggaca ctgcacattc	780	
gaaccggcag aaccgaaggc agcaaagaac acatacacag cagaaagatt catctggctg	840	
acaaagctga acaacctgag aatcctggaa cagggatcgg aaagaccgct gacagacaca	900	
gaaagagcaa cactgatgga cgaaccgtac agaaagtcga agctgacata cgcacaggca	960	
agaaagctgc tgggactgga agacacagca ttcttcaagg gactgagata cggaaaggac	1020	
aacgcagaag catcgacact gatggaaatg aaggcatacc acgcaatctc gagagcactg	1080	
gaaaaggaag gactgaagga caagaagtcg ccgctgaacc tgtcgccgga actgcaggac	1140	
gaaatcggaa cagcattctc gctgttcaag acagacgaag acatcacagg aagactgaag	1200	
gacagaatcc agccggaaat cctggaagca ctgctgaagc acatctcgtt cgacaagttc	1260	
gtccagatct cgctgaaggc actgagaaga atcgctccgc tgatggaaca gggaaagaga	1320	10
tacgacgaag catgcgaga aatctacgga gaccactacg gaaagaagaa cacagaagaa	1380	
aagatctacc tgccgccgat cccggcagac gaaatcagaa acccggtcgt cctgagagca	1440	
ctgtcgcagg caagaaaggt catcaacgga gtcgtcagaa gatacggatc gccggcaaga	1500	
atccacatcg aaacagcaag agaagtcgga aagtcgttca aggacagaaa ggaaatcgaa	1560	
aagagacagg aagaaaacag aaaggacaga gaaaaggcag cagcaaagtt cagagaatac	1620	
ttccgaact tcgtcggaga accgaagtcg aaggacatcc tgaagctgag actgtacgaa	1680	
cagcagcacg gaaagtcct gtactcggga aaggaaatca acctgggaag actgaacgaa	1740	
aagggatacg tcgaaatcga ccacgcactg ccgttctcga gaacatggga cgactcgttc	1800	
aacaacaagg tcctggtcct gggatcggaa aaccagaaca agggaaacca gacaccgtac	1860	
gaatacttca acggaaagga caactcgaga gaatggcagg aattcaaggc aagagtcgaa	1920	20
acatcgagat tcccgagatc gaagaagcag agaatcctgc tgcagaagtt cgacgaagac	1980	
ggattcaagg aaagaaacct gaacgacaca agatacgtca acagattcct gtgccagttc	2040	
gtcgcagaca gaatgagact gacaggaaaag ggaaagaaga gagtccttcgc atcgaacgga	2100	
cagatcacia acctgctgag aggattctgg ggactgagaa aggtcagagc agaaaacgac	2160	
agacaccacg cactggacgc agtcgtcgtc gcatgctcga cagtcgcaat gcagcagaag	2220	
atcacaagat tcgtcagata caaggaaatg aacgcattcg acggaaagac aatcgacaag	2280	
gaaacaggag aagtcctgca ccagaagaca cacttcccgc agccgtggga attcttcgca	2340	
caggaagtca tgatcagagt cttcggaaag cgggacggaa agccggaatt cgaagaagca	2400	
gacacactgg aaaagctgag aacactgctg gcagaaaagc tgtcgtcgag accggaagca	2460	
gtccacgaat acgtcacacc gctgttcgtc tcgagagcac cgaacagaaa gatgtcggga	2520	30
cagggacaca tggaacagt caagtcggca aagagactgg acgaaggagt ctcggtcctg	2580	
agagtcccgc tgacacagct gaagctgaag gacctggaaa agatgggtcaa cagagaaaga	2640	
gaaccgaagc tgtacgaagc actgaaggca agactggaag cacacaagga cgaccgggca	2700	
aaggcattcg cagaaccgtt ctacaagtac gacaaggcag gaaacagaac acagcaggtc	2760	
aaggcagtca gagtcgaaca ggtccagaag acaggagtct gggtcagaaa ccacaacgga	2820	
atcgcagaca acgcaacaat ggtcagagta gacgtcttcg aaaagggaga caagtactac	2880	
ctggtcccga tctactcgtg gcaggtcgca aagggaatcc tgccggacag agcagtcgtc	2940	
cagggaaagg acgaagaaga ctggcagctg atcgacgact cgttcaactt caagttctcg	3000	
ctgcacccga acgacctggt cgaagtcac acaaagaagg caagaatgtt cggatacttc	3060	
gcatcgtgcc acagaggaac aggaacatc aacatcagaa tccacgacct ggaccacaag	3120	40
atcggaaga acggaatcct ggaaggaatc ggagtcaaga cagcactgtc gttccagaag	3180	
taccagatcg acgaactggg aaaggaaatc agaccgtgca gactgaagaa gagaccgccg	3240	
gtcagatccg gaaagagaac agcagacgga tcggaattcg aatcgccgaa gaagaagaga	3300	
aaggtcgaat ga	3312	

<210> 66

<211> 3306

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 66

gcagcattca agccgaactc gatcaactac atcctgggac tggacatcgg aatcgcacgc	60	
gtcggatggg caatggtcga aatcgacgaa gaagaaaacc cgatcagact gatcgacctg	120	
ggagtcagag tcttcgaaag agcagaagtc ccgaagacag gagactcgct ggcaatggca	180	
agaagactgg caagatcggg cagaagactg acaagaagaa gagcacacag actgctgaga	240	
acaagaagac tgctgaagag agaaggagtc ctgcaggcag caaacttcga cgaaaacgga	300	
ctgatcaagt cgctgccgaa cacaccgtgg cagctgagag cagcagcact ggacagaaag	360	10
ctgacaccgc tggatgggtc ggagtcctg ctgcacctga tcaagcacag aggatacctg	420	
tcgcagagaa agaacgaagg agaaacagca gacaaggaac tgggagcact gctgaaggga	480	
gtcgcaggaa acgcacacgc actgcagaca ggagacttca gaacaccggc agaactggca	540	
ctgaacaagt tcgaaaagga atcgggacac atcagaaacc agagatcggga ctactcgcac	600	
acattctcga gaaaggacct gcaggcagaa ctgattctgc tgttcgaaaa gcagaaggaa	660	
ttcggaaaacc cgcacgtctc gggaggactg aaggaaggaa tcgaaacact gctgatgaca	720	
cagagaccgg cactgtcggg agacgcagtc cagaagatgc tgggacactg cacattcgaa	780	
ccggcagaac cgaaggcagc aaagaacaca tacacagcag aaagattcat ctggctgaca	840	
aagctgaaca acctgagaat cctggaacag ggatcggaaa gaccgctgac agacacagaa	900	
agagcaacac tgatggacga accgtacaga aagtcgaagc tgacatacgc acaggcaaga	960	20
aagctgctgg gactggaaga cacagcattc ttaagggac tgagatacgg aaaggacaac	1020	
gcagaagcat cgacactgat ggaaatgaag gcataccacg caatctcgag agcactggaa	1080	
aaggaaggac tgaaggacaa gaagtcgccg ctgaacctgt cgccggaact gcaggacgaa	1140	
atcggaaacag cattctcgct gttcaagaca gacgaagaca tcacaggaag actgaaggac	1200	
agaatccagc cggaatcctt ggaagcactg ctgaagcaca tctcgttcga caagtctgtc	1260	
cagatctcgc tgaaggcact gagaagaatc gtcccgctga tggaaacaggg aaagagatac	1320	
gacgaagcat gcgcagaaat ctacggagac cactacggaa agaagaacac agaagaaaag	1380	
atctacctgc cgccgatccc ggcagacgaa atcagaaacc cggctgtcct gagagcactg	1440	
tcgcaggcaa gaaaggtcat caacggagtc gtcagaagat acggatcgcc ggcaagaatc	1500	
cacatcgaaa cagcaagaga agtcggaaaag tcgttcaagg acagaaagga aatcgaaaag	1560	30
agacaggaag aaaacagaaa ggacagagaa aaggcagcag caaagttcag agaatacttc	1620	
ccgaacttcg tcggagaacc gaagtcgaag gacatcctga agctgagact gtacgaacag	1680	
cagcacggaa agtgcctgta ctcgggaaaag gaaatcaacc tgggaagact gaacgaaaag	1740	
ggatacgtcg aaatcgacca cgcactgccg ttctcgagaa catgggacga ctcgttcaac	1800	
aacaaggttc tggctcctggg atcggaaaac cagaacaagg gaaaccagac accgtacgaa	1860	
tacttcaacg gaaaggacaa ctcgagagaa tggcaggaat tcaaggcaag agtcgaaaca	1920	
tcgagattcc cgagatcgaa gaagcagaga atcctgtctg agaagttcga cgaagacgga	1980	
ttcaaggaaa gaaacctgaa cgacacaaga tacgtcaaca gattcctgtg ccagttcgtc	2040	
gcagacagaa tgagactgac aggaaggga aagaagagag tcttcgcatc gaacggacag	2100	
atcacaacc tgctgagagg attctgggga ctgagaaagg tcagagcaga aaacgacaga	2160	40
caccacgcac tggacgcagt cgtcgtcgca tgctcgacag tcgcaatgca gcagaagatc	2220	
acaagattcg tcagatacaa ggaaatgaac gcattcgacg gaaagacaat cgacaaggaa	2280	
acaggagaag tcctgcacca gaagacacac ttcccgcagc cgtgggaatt cttcgcacag	2340	
gaagtcatga tcagagtctt cggaagccg gacggaaagc cggaattcga agaagcagac	2400	
acactggaaa agctgagaac actgctggca gaaaagctgt cgtcgagacc ggaagcagtc	2460	
cacgaatacg tcacaccgct gttcgtctcg agagcaccga acagaaagat gtcgggacag	2520	
ggacacatgg aaacagtcaa gtcggcaaaag agactggacg aaggagtctc ggtcctgaga	2580	
gtcccgctga cacagctgaa gctgaaggac ctggaaaaga tggtaacag agaaagagaa	2640	
ccgaagctgt acgaagcact gaaggcaaga ctggaagcac acaaggacga cccggcaaaag	2700	
gcattcgcag aaccgttcta caagtacgac aaggcaggaa acagaacaca gcaggtaag	2760	50

gcagtcagag tcgaacaggt ccagaagaca ggagtcctggg tcagaaacca caacggaatc 2820
gcagacaacg caacaatggg cagagtagac gtcttcgaaa agggagacaa gtactacctg 2880
gtcccgatct actcgtggca ggtcgcaaag ggaatcctgc cggacagagc agtcgtccag 2940
ggaaaggacg aagaagactg gcagctgacg gacgactcgt tcaacttcaa gttctcgctg 3000
caccggaacg acctggtcga agtcatcaca aagaaggcaa gaatgttcgg atacttcgca 3060
tcgtgccaca gaggaacagg aaacatcaac atcagaatcc acgacctgga ccacaagatc 3120
ggaaagaacg gaatcctgga aggaatcgga gtcaagacag cactgtcgtt ccagaagtac 3180
cagatcgacg aactgggaaa ggaaatcaga ccgtgcagac tgaagaagag accgccggtc 3240
agatccggaa agagaacagc agacggatcg gaattcgaat cgccgaagaa gaagagaaaag 3300
gtcgaa 3306

10

<210> 67

<211> 3636

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 67

20

gggagaccca agctggctag cgtttaaact taagcttgga tccgccacca tggcagcatt 60
caagccgaac tcgatcaact acatcctggg actggacatc ggaatcgcac cggtcggatg 120
ggcaatggtc gaaatcgacg aagaagaaaa cccgatcaga ctgatcgacc tgggagtcag 180
agtcttcgaa agagcagaag tcccgaagac aggagactcg ctggcaatgg caagaagact 240
ggcaagatcg gtcagaagac tgacaagaag aagagcacac agactgctga gaacaagaag 300
actgctgaag agagaaggag tcctgcaggc agcaaacttc gacgaaaacg gactgatcaa 360
gtcgtgcccg aacacaccgt ggacagctgag agcagcagca ctggacagaa agctgacacc 420
gctggaatgg tcggcagtc tgctgcacct gatcaagcac agaggatacc tgtcgcagag 480
aaagaacgaa ggagaaacag cagacaagga actgggagca ctgctgaagg gagtcgcagg 540
aaacgcacac gcactgcaga caggagactt cagaacaccg gcagaactgg cactgaacaa 600
gttcgaaaag gaatcgggac acatcagaaa ccagagatcg gactactcgc acacattctc 660
gagaaaggac ctgcaggcag aactgatcct gctgttcgaa aagcagaagg aattcggaaa 720
cccgacgctc tcgggaggac tgaaggaagg aatcgaacaa ctgctgatga cacagagacc 780
ggcactgtcg ggagacgcag tccagaagat gctgggacac tgcacattcg aaccggcaga 840
accgaaggca gcaaagaaca catacacagc agaaagattc atctggctga caaagctgaa 900
caacctgaga atcctggaac agggatcgga aagaccgtg acagacacag aaagagcaac 960
actgatggac gaaccgtaca gaaagtcgaa gctgacatac gcacaggcaa gaaagctgct 1020
gggactggaa gacacagcat tcttcaaggg actgagatac ggaaaggaca acgcagaagc 1080
atcgacactg atggaaatga aggcatacca cgcaatctcg agagcactgg aaaaggaagg 1140
actgaaggac aagaagtcgc cgctgaacct gtcgccggaa ctgcaggacg aaatcggaa 1200
agcattctcg ctgttcaaga cagacgaaga catcacagga agactgaagg acagaatcca 1260
gccggaaatc ctggaagcac tgctgaagca catctcgttc gacaagtctg tccagatctc 1320
gctgaaggca ctgagaagaa tcgtcccgtc gatggaacag ggaaagagat acgacgaagc 1380
atgcgcagaa atctacggag accactacgg aaagaagaac acagaagaaa agatctacct 1440
gccgccgac cggcgagacg aaatcagaaa cccggctcgt ctgagagcac tgtcgcaggc 1500
aagaaaggtc atcaacggag tcgtcagaag atacggatcg ccggcaagaa tccacatcga 1560
aacagcaaga gaagtctggaa agtcgttcaa ggacagaaag gaaatcgaaa agagacagga 1620
agaaaacaga aaggacagag aaaaggcagc agcaaagttc agagaatact tcccgaactt 1680
cgtcggagaa ccgaagtcga aggacatcct gaagctgaga ctgtacgaac agcagcacgg 1740
aaagtgcctg tactcgggaa aggaaatcaa cctgggaaga ctgaacgaaa agggatacgt 1800

30

40

50

cgaaatcgac cacgcactgc cgttctcgag aacatgggac gactcgttca acaacaaggt 1860
 cctggctcctg ggatcggaaa accagaacaa gggaaaccag acaccgtacg aatacttcaa 1920
 cggaaggac aactcgagag aatggcagga attcaaggca agagtcgaaa catcgagatt 1980
 cccgagatcg aagaagcaga gaatcctgct gcagaagttc gacgaagacg gattcaagga 2040
 aagaaacctg aacgacacaa gatacgtcaa cagattcctg tgccagttcg tcgcagacag 2100
 aatgagactg acaggaaagg gaaagaagag agtcttcgca tcgaacggac agatcacaaa 2160
 cctgctgaga ggattctggg gactgagaaa ggtcagagca gaaaacgaca gacaccacgc 2220
 actggacgca gtcgtcgtcg catgctcgac agtcgcaatg cagcagaaga tcacaagatt 2280
 cgtcagatac aaggaaatga acgcattcga cggaaagaca atcgacaagg aaacaggaga 2340
 agtcctgcac cagaagacac acttcccga gccgtgggaa ttcttcgcac aggaagtcac 2400
 gatcagagtc ttcggaagc cggacggaaa gccggaattc gaagaagcag acacactgga 2460
 aaagctgaga acactgctgg cagaaaagct gtcgtcgaga ccggaagcag tccacgaata 2520
 cgtcacaccg ctgttcgtct cgagagcacc gaacagaaaag atgtcgggac agggacacat 2580
 ggaaacagtc aagtcggcaa agagactgga cgaaggagtc tcggtcctga gagtcccgt 2640
 gacacagctg aagctgaagg acctggaaaa gatggtaaac agagaaagag aaccgaagct 2700
 gtacgaagca ctgaaggcaa gactggaagc acacaaggac gaccggcaa aggcattcgc 2760
 agaaccgttc tacaagtacg acaaggcagg aaacagaaca cagcaggta aggagtcag 2820
 agtcgaacag gtccagaaga caggagtcgt ggtcagaaac cacaacggaa tcgcagacaa 2880
 cgcaacaatg gtcagagtag acgtcttcga aaagggagac aagtactacc tgggtccgat 2940
 ctactcgtgg caggctgcaa agggaaatcct gccggacaga gcagtcgtcc agggaaagga 3000
 cgaagaagac tggcagctga tcgacgactc gttcaacttc aagttctgcg tgcacccgaa 3060
 cgacctggtc gaagtcacat caaagaaggc aagaatgttc ggatacttcg catcgtgcca 3120
 cagaggaaca ggaaacatca acatcagaat ccacgacctg gaccacaaga tcggaaagaa 3180
 cggaatcctg gaaggaatcg gactcaagac agcactgtcg ttccagaagt accagatcga 3240
 cgaactggga aaggaaatca gaccgtgcag actgaagaag agaccgccgg tcagatccgg 3300
 aaagagaaca gcagacggat cggaattcga atcgccgaag aagaagagaa aggtcgaatg 3360
 atagctagct cgagtctaga gggcccgttt aaaccgctg atcagcctcg actgtgcctt 3420
 ctagttagca gccatctgtt gtttggccct ccccggtgcc ttccttgacc ctggaagggt 3480
 ccactccac tgtcctttcc taataaaatg aggaaattgc atcgattgt ctgagtaggt 3540
 gtcattctat tctgggggggt ggggtggggc aggacagcaa gggggaggat tgggaagaca 3600
 atagcaggca tgctggggat gcggtgggct ctatgg 3636

10

20

30

<210> 68

<211> 1103

<212> PRT

<213> Neisseria meningitidis

<400> 68

Met Ala Ala Phe Lys Pro Asn Ser Ile Asn Tyr Ile Leu Gly Leu Asp

1 5 10 15

40

Ile Gly Ile Ala Ser Val Gly Trp Ala Met Val Glu Ile Asp Glu Glu

20 25 30

Glu Asn Pro Ile Arg Leu Ile Asp Leu Gly Val Arg Val Phe Glu Arg

35 40 45

Ala Glu Val Pro Lys Thr Gly Asp Ser Leu Ala Met Ala Arg Arg Leu

50 55 60

Ala Arg Ser Val Arg Arg Leu Thr Arg Arg Arg Ala His Arg Leu Leu

50

65 70 75 80
Arg Thr Arg Arg Leu Leu Lys Arg Glu Gly Val Leu Gln Ala Ala Asn
 85 90 95
Phe Asp Glu Asn Gly Leu Ile Lys Ser Leu Pro Asn Thr Pro Trp Gln
 100 105 110
Leu Arg Ala Ala Ala Leu Asp Arg Lys Leu Thr Pro Leu Glu Trp Ser
 115 120 125
Ala Val Leu Leu His Leu Ile Lys His Arg Gly Tyr Leu Ser Gln Arg
 130 135 140
Lys Asn Glu Gly Glu Thr Ala Asp Lys Glu Leu Gly Ala Leu Leu Lys
145 150 155 160
Gly Val Ala Gly Asn Ala His Ala Leu Gln Thr Gly Asp Phe Arg Thr
 165 170 175
Pro Ala Glu Leu Ala Leu Asn Lys Phe Glu Lys Glu Ser Gly His Ile
 180 185 190
Arg Asn Gln Arg Ser Asp Tyr Ser His Thr Phe Ser Arg Lys Asp Leu
 195 200 205
Gln Ala Glu Leu Ile Leu Leu Phe Glu Lys Gln Lys Glu Phe Gly Asn
 210 215 220
Pro His Val Ser Gly Gly Leu Lys Glu Gly Ile Glu Thr Leu Leu Met
225 230 235 240
Thr Gln Arg Pro Ala Leu Ser Gly Asp Ala Val Gln Lys Met Leu Gly
 245 250 255
His Cys Thr Phe Glu Pro Ala Glu Pro Lys Ala Ala Lys Asn Thr Tyr
 260 265 270
Thr Ala Glu Arg Phe Ile Trp Leu Thr Lys Leu Asn Asn Leu Arg Ile
 275 280 285
Leu Glu Gln Gly Ser Glu Arg Pro Leu Thr Asp Thr Glu Arg Ala Thr
 290 295 300
Leu Met Asp Glu Pro Tyr Arg Lys Ser Lys Leu Thr Tyr Ala Gln Ala
305 310 315 320
Arg Lys Leu Leu Gly Leu Glu Asp Thr Ala Phe Phe Lys Gly Leu Arg
 325 330 335

10

20

30

40

50

Tyr Gly Lys Asp Asn Ala Glu Ala Ser Thr Leu Met Glu Met Lys Ala
340 345 350

Tyr His Ala Ile Ser Arg Ala Leu Glu Lys Glu Gly Leu Lys Asp Lys
355 360 365

Lys Ser Pro Leu Asn Leu Ser Pro Glu Leu Gln Asp Glu Ile Gly Thr
370 375 380

Ala Phe Ser Leu Phe Lys Thr Asp Glu Asp Ile Thr Gly Arg Leu Lys
385 390 395 400

10

Asp Arg Ile Gln Pro Glu Ile Leu Glu Ala Leu Leu Lys His Ile Ser
405 410 415

Phe Asp Lys Phe Val Gln Ile Ser Leu Lys Ala Leu Arg Arg Ile Val
420 425 430

Pro Leu Met Glu Gln Gly Lys Arg Tyr Asp Glu Ala Cys Ala Glu Ile
435 440 445

20

Tyr Gly Asp His Tyr Gly Lys Lys Asn Thr Glu Glu Lys Ile Tyr Leu
450 455 460

Pro Pro Ile Pro Ala Asp Glu Ile Arg Asn Pro Val Val Leu Arg Ala
465 470 475 480

Leu Ser Gln Ala Arg Lys Val Ile Asn Gly Val Val Arg Arg Tyr Gly
485 490 495

30

Ser Pro Ala Arg Ile His Ile Glu Thr Ala Arg Glu Val Gly Lys Ser
500 505 510

Phe Lys Asp Arg Lys Glu Ile Glu Lys Arg Gln Glu Glu Asn Arg Lys
515 520 525

Asp Arg Glu Lys Ala Ala Ala Lys Phe Arg Glu Tyr Phe Pro Asn Phe
530 535 540

Val Gly Glu Pro Lys Ser Lys Asp Ile Leu Lys Leu Arg Leu Tyr Glu
545 550 555 560

40

Gln Gln His Gly Lys Cys Leu Tyr Ser Gly Lys Glu Ile Asn Leu Gly
565 570 575

Arg Leu Asn Glu Lys Gly Tyr Val Glu Ile Asp His Ala Leu Pro Phe
580 585 590

Ser Arg Thr Trp Asp Asp Ser Phe Asn Asn Lys Val Leu Val Leu Gly
595 600 605

50

Ser Glu Asn Gln Asn Lys Gly Asn Gln Thr Pro Tyr Glu Tyr Phe Asn
610 615 620

Gly Lys Asp Asn Ser Arg Glu Trp Gln Glu Phe Lys Ala Arg Val Glu
625 630 635 640

Thr Ser Arg Phe Pro Arg Ser Lys Lys Gln Arg Ile Leu Leu Gln Lys
645 650 655

Phe Asp Glu Asp Gly Phe Lys Glu Arg Asn Leu Asn Asp Thr Arg Tyr
660 665 670

Val Asn Arg Phe Leu Cys Gln Phe Val Ala Asp Arg Met Arg Leu Thr
675 680 685

Gly Lys Gly Lys Lys Arg Val Phe Ala Ser Asn Gly Gln Ile Thr Asn
690 695 700

Leu Leu Arg Gly Phe Trp Gly Leu Arg Lys Val Arg Ala Glu Asn Asp
705 710 715 720

Arg His His Ala Leu Asp Ala Val Val Val Ala Cys Ser Thr Val Ala
725 730 735

Met Gln Gln Lys Ile Thr Arg Phe Val Arg Tyr Lys Glu Met Asn Ala
740 745 750

Phe Asp Gly Lys Thr Ile Asp Lys Glu Thr Gly Glu Val Leu His Gln
755 760 765

Lys Thr His Phe Pro Gln Pro Trp Glu Phe Phe Ala Gln Glu Val Met
770 775 780

Ile Arg Val Phe Gly Lys Pro Asp Gly Lys Pro Glu Phe Glu Glu Ala
785 790 795 800

Asp Thr Leu Glu Lys Leu Arg Thr Leu Leu Ala Glu Lys Leu Ser Ser
805 810 815

Arg Pro Glu Ala Val His Glu Tyr Val Thr Pro Leu Phe Val Ser Arg
820 825 830

Ala Pro Asn Arg Lys Met Ser Gly Gln Gly His Met Glu Thr Val Lys
835 840 845

Ser Ala Lys Arg Leu Asp Glu Gly Val Ser Val Leu Arg Val Pro Leu
850 855 860

Thr Gln Leu Lys Leu Lys Asp Leu Glu Lys Met Val Asn Arg Glu Arg

10

20

30

40

50

865	870	875	880	
Glu Pro Lys Leu Tyr Glu Ala Leu Lys Ala Arg Leu Glu Ala His Lys				
	885	890	895	
Asp Asp Pro Ala Lys Ala Phe Ala Glu Pro Phe Tyr Lys Tyr Asp Lys				
	900	905	910	
Ala Gly Asn Arg Thr Gln Gln Val Lys Ala Val Arg Val Glu Gln Val				
	915	920	925	10
Gln Lys Thr Gly Val Trp Val Arg Asn His Asn Gly Ile Ala Asp Asn				
	930	935	940	
Ala Thr Met Val Arg Val Asp Val Phe Glu Lys Gly Asp Lys Tyr Tyr				
945	950	955	960	
Leu Val Pro Ile Tyr Ser Trp Gln Val Ala Lys Gly Ile Leu Pro Asp				
	965	970	975	20
Arg Ala Val Val Gln Gly Lys Asp Glu Glu Asp Trp Gln Leu Ile Asp				
	980	985	990	
Asp Ser Phe Asn Phe Lys Phe Ser Leu His Pro Asn Asp Leu Val Glu				
	995	1000	1005	
Val Ile Thr Lys Lys Ala Arg Met Phe Gly Tyr Phe Ala Ser Cys				
	1010	1015	1020	
His Arg Gly Thr Gly Asn Ile Asn Ile Arg Ile His Asp Leu Asp				30
	1025	1030	1035	
His Lys Ile Gly Lys Asn Gly Ile Leu Glu Gly Ile Gly Val Lys				
	1040	1045	1050	
Thr Ala Leu Ser Phe Gln Lys Tyr Gln Ile Asp Glu Leu Gly Lys				
	1055	1060	1065	
Glu Ile Arg Pro Cys Arg Leu Lys Lys Arg Pro Pro Val Arg Ser				40
	1070	1075	1080	
Gly Lys Arg Thr Ala Asp Gly Ser Glu Phe Glu Ser Pro Lys Lys				
	1085	1090	1095	
Lys Arg Lys Val Glu				
	1100			

<210> 69
 <211> 100
 <212> RNA

10

20

30

40

50

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<220>

<221> modified_base

<222> (1)..(3)

<223> 2'-O-methyl modified

10

<220>

<221> misc_feature

<222> (1)..(4)

<223> Phosphorothioate linkage

<220>

<221> modified_base

<222> (29)..(40)

<223> 2'-O-methyl modified

<220>

<221> modified_base

20

<222> (69)..(100)

<223> 2'-O-methyl modified

<220>

<221> misc_feature

<222> (97)..(100)

<223> Phosphorothioate linkage

<400> 69

gccgagucug gagagcugca guuuuagagc uagaaaauagc aaguuaaaau aaggcuaguc 60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

30

<210> 70

<211> 100

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<220>

40

<221> modified_base

<222> (1)..(3)

<223> 2'-O-methyl modified

<220>

<221> misc_feature

<222> (1)..(4)

<223> Phosphorothioate linkage

<220>

<221> modified_base

<222> (29)..(40)

50

<223> 2'-O-methyl modified

<220>

<221> modified_base

<222> (69)..(100)

<223> 2'-O-methyl modified

<220>

<221> misc_feature

<222> (97)..(100)

<223> Phosphorothioate linkage

<400> 70

10

acacaaauac caguccagcg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 71

<211> 100

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

20

<220>

<221> modified_base

<222> (1)..(3)

<223> 2'-O-methyl modified

<220>

<221> misc_feature

<222> (1)..(4)

<223> Phosphorothioate linkage

30

<220>

<221> modified_base

<222> (29)..(40)

<223> 2'-O-methyl modified

<220>

<221> modified_base

<222> (69)..(100)

<223> 2'-O-methyl modified

<220>

<221> misc_feature

40

<222> (97)..(100)

<223> Phosphorothioate linkage

<400> 71

aaaguucuag augccguccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 72

<211> 100

<212> RNA

<213> Artificial Sequence

50

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<220>
<221> modified_base
<222> (1)..(3)
<223> 2'-O-methyl modified
<220>
<221> misc_feature
<222> (1)..(4)
<223> Phosphorothioate linkage

10

<220>
<221> modified_base
<222> (29)..(40)
<223> 2'-O-methyl modified

<220>
<221> modified_base
<222> (69)..(100)
<223> 2'-O-methyl modified
<220>
<221> misc_feature
<222> (97)..(100)
<223> Phosphorothioate linkage
<400> 72

20

acgcaaaauau caguccagcg guuuuagagc uagaaaauagc aaguuaaaaau aaggcuaguc 60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 73
<211> 7
<212> PRT
<213> Simian virus 40
<400> 73

30

Pro Lys Lys Lys Arg Lys Val
1 5

<210> 74
<211> 7
<212> PRT
<213> Simian virus 40
<400> 74

40

Pro Lys Lys Lys Arg Arg Val
1 5

<210> 75
<211> 16
<212> PRT
<213> Unknown
<220>

50

<221> source
 <223> /note="Description of Unknown:
 Nucleoplasmin bipartite NLS sequence"

<400> 75

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
 1 5 10 15

<210> 76

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
 6xHis tag"

<400> 76

His His His His His His

1 5

<210> 77

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
 8xHis tag"

<400> 77

His His His His His His His His

1 5

<210> 78

<211> 10

<212> RNA

<213> Unknown

<220>

<221> source

<223> /note="Description of Unknown:
 Kozak sequence"

<400> 78

gccrccaugg

10

<210> 79

<211> 13

<212> RNA

<213> Unknown

<220>

<221> source

<223> /note="Description of Unknown:
 Kozak sequence"

10

20

30

40

50

<400> 79
gccgcccrrcca ugg 13

<210> 80
<211> 56
<212> DNA
<213> Artificial Sequence
<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic primer" 10
<400> 80
cactctttcc ctacacgacg ctcttccgat ctgttttggt ccagagtcta tcaccg 56

<210> 81
<211> 52
<212> DNA
<213> Artificial Sequence
<220>
<221> source 20
<223> /note="Description of Artificial Sequence: Synthetic primer"
<400> 81
ggagttcaga cgtgtgctct tccgatctac acgaataaga gcaaattggga ac 52

<210> 82
<211> 55
<212> DNA
<213> Artificial Sequence
<220> 30
<221> source
<223> /note="Description of Artificial Sequence: Synthetic primer"
<400> 82
cactctttcc ctacacgacg ctcttccgat cttgcatttc atgagaccga aaaca 55

<210> 83
<211> 55
<212> DNA
<213> Artificial Sequence 40
<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic primer"
<400> 83
ggagttcaga cgtgtgctct tccgatctgc tacagtagag ctgtacataa aactt 55

<210> 84
<211> 3783
<212> DNA 50

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 84

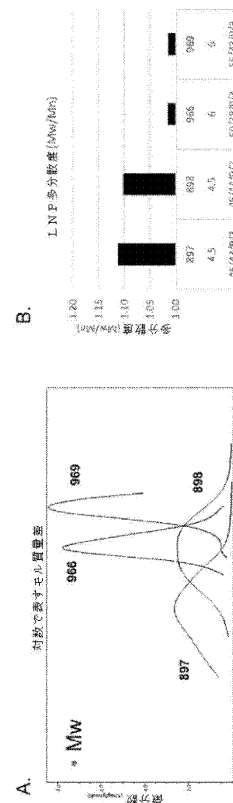
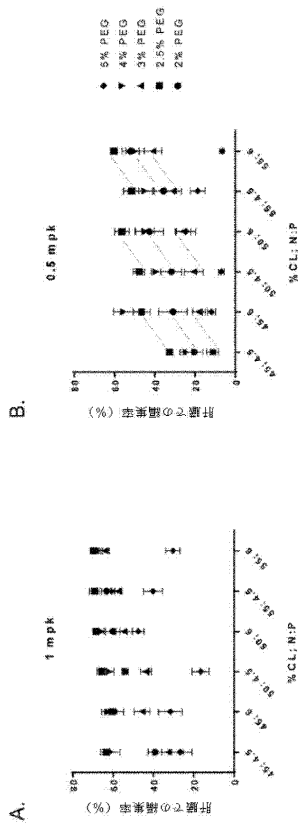
tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca	60	
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120	
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180	
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240	10
attgccatt caggctgcgc aactgttggg aaggcgatc ggtgcgggcc tcttcgctat	300	
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagtgggta acgccagggt	360	
tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt ctaatacgac tcaactatagg	420	
gtcccgcagt cggcgtccag cggctctgct tgttcgtgtg tgtgtcgttg caggccttat	480	
tcggatccat ggtgagcaag ggcgaggagc tgttcaccgg ggtgggtgcc atcctggctg	540	
agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg	600	
ccacctacgg caagctgacc ctgaagtcca tctgcaccac cggcaagctg cccgtgccct	660	
ggcccacct cgtgaccacc ctgacctacg gcgtgcagtg cttcagccgc taccgacc	720	
acatgaagca gcacgacttc ttcaagtccg ccatgcccga aggctacgtc caggagcgca	780	
ccatcttctt caaggacgac ggcaactaca agaccgcgc cgaggtgaag ttcgaggcg	840	20
acaccttggt gaaccgcac gagctgaagg gcatcgactt caaggaggac ggcaacatcc	900	
tggggcacia gctggagtac aactacaaca gccacaacgt ctatatcatg gccgacaagc	960	
agaagaacgg catcaagggtg aacttcaaga tccgccacia catcgaggac ggcagcgtgc	1020	
agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg ctgctgccc	1080	
acaaccacta cctgagcacc cagtccgccc tgagcaaaga cccaacgag aagcgcgac	1140	
acatggctct gctggagtgc gtgaccgccg ccgggatcac tctcggcacg gacgagctgt	1200	
acaagtaata ggaattatgc agtctagcca tcacatttaa aagcatctca gcctaccatg	1260	
agaataagag aaagaaaatg aagatcaata gcttattcat ctcttttct ttttcgttgg	1320	
tgtaaagcca acacctgtc taaaaaacat aaatttctt aatcattttg cctctttct	1380	
ctgtgcttca attaataaaa aatggaaaga acctcgagaa aaaaaaaaaa aaaaaaaaaa	1440	30
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1500	
aaaaaaaaa aaaaaaaaaa tagacttaag cttgatgagc tctagcttgg cgtaatcatg	1560	
gtcatagctg tttcctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc	1620	
cggaaacata aagtgtaaag cctgggggtgc ctaatgagt agctaactca cattaattgc	1680	
gttgcgctca ctgcccgtt tccagtcggg aaacctgtcg tgccagctgc attaatgaat	1740	
cggccaacgc gcggggagag gcggtttgcg tattgggcgc tcttcgctt cctcgctcac	1800	
tgactcgctg cgctcggtcg ttcgggtcgc gcgagcggtc tcagctcact caaaggcggt	1860	
aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca	1920	
gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg ttttccata ggctccgcc	1980	
ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact	2040	40
ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct	2100	
gccgcttacc ggatacctgt ccgccttct ccttcggga agcgtggcg tttctcatag	2160	
ctcacgctgt aggtatctca gtccggtgta ggtcgttcgc tccaagctgg gctgtgtgca	2220	
cgaaccccc gttcagccc accgctgcgc cttatccggt aactatcgtc ttgagtccaa	2280	
cccggtaaaga cagacttat cgccactggc agcagccact ggtaacagga ttagcagagc	2340	
gaggtatgta ggcggtgcta cagagttctt gaagtgggtg cctaactacg gctacactag	2400	
aagaacagta ttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg	2460	
tagctcttga tccggcaaac aaaccaccgc tggtagcggg ggttttttg ttgcaagca	2520	
gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatcttt ctacggggtc	2580	
tgacgctcag tggaacgaaa actcacgtta agggattttg gtcattgagat tatcaaaaag	2640	50

gatcttcacc tagatccttt taaattaa atgaagtttt aaatcaatct aaagtatata 2700
tgagtaaact tggctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat 2760
ctgtctatct cgttcaccca tagttgctg actccccgtc gtgtagataa ctacgatacg 2820
ggagggctta ccacttgcc ccagtgtgc aatgataccg cgagaccac gctcaccggc 2880
tccagattta tcagcaataa accagccagc cggaaggcc gagcgagaa gtgtcctgc 2940
aactttatcc gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc 3000
gccagttaat agtttgcgca acgttggtgc cattgctaca ggcatcgtgg tgtcacgctc 3060
gtcgtttggt atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgac 3120
cccatgttg tgcaaaaaaag cggtagctc cttcggtcct ccgatcgttg tcagaagtaa 3180
gttgccgca gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat 3240
gccatccgta agatgctttt ctgtgactgg tgagtactca accaagtcac tctgagaata 3300
gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca 3360
tagcagaact taaaagtgc tcacattgg aaaacgttct tcggggcgaa aacttcaag 3420
gatcttaccg ctgttgagat ccagttcgat gtaaccact cgtgcaccca actgatcttc 3480
agcatctttt actttacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc 3540
aaaaaaggga ataaggcgca cacggaaatg ttgaatactc atactcttcc ttttcaata 3600
ttattgaagc attatcagg gttattgtct catgagcgga tacatatttg aatgtattta 3660
gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtccac ctgacgtcta 3720
agaaaccatt attatcatga cattaaccta taaaaatagg cgtatcacga ggcccttcg 3780
tcg 3783

【図面】

【図 1】

【図 2】



10

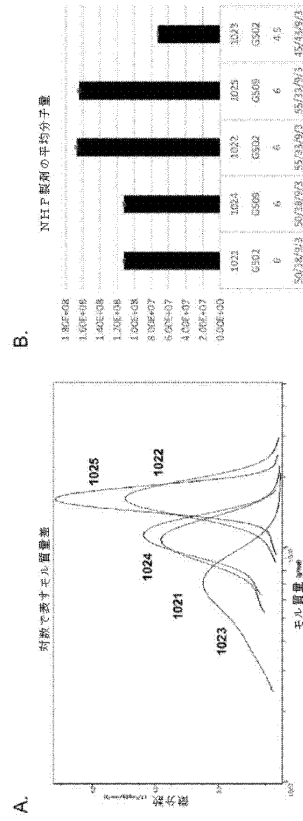
20

30

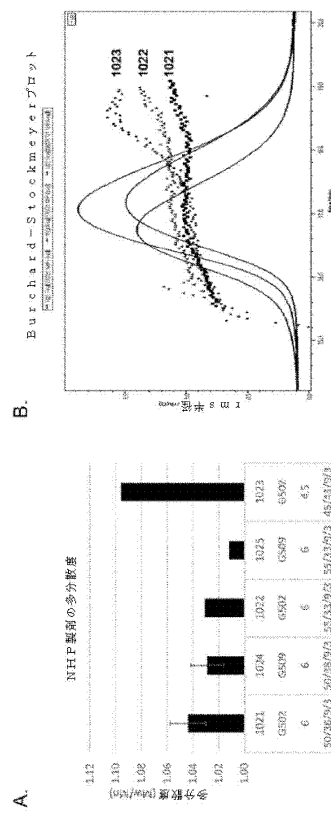
40

50

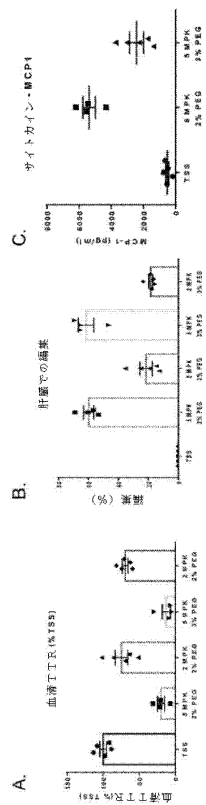
【図 3】



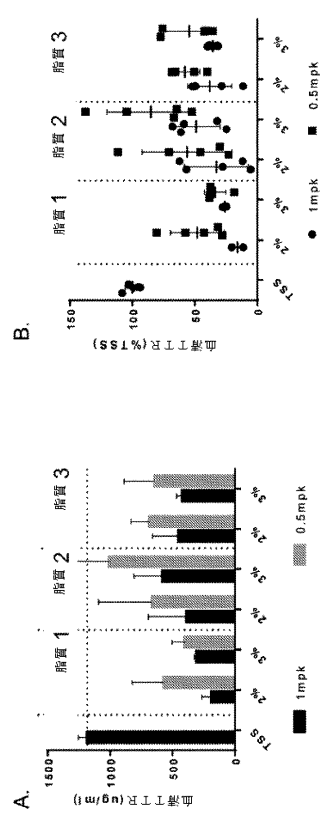
【図 4】



【図 5】



【図 6 - 1】



10

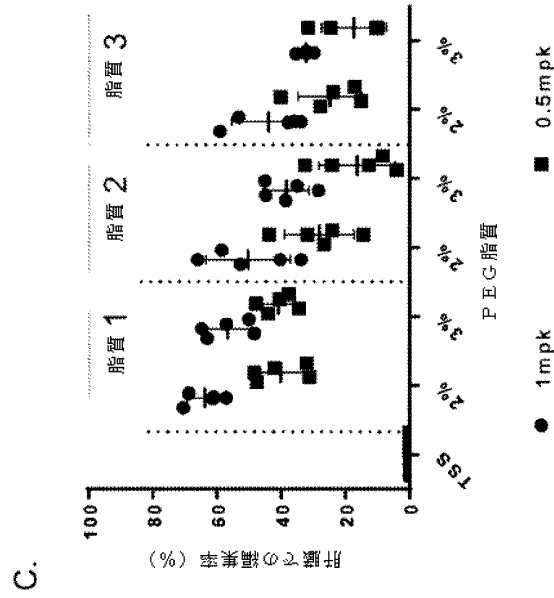
20

30

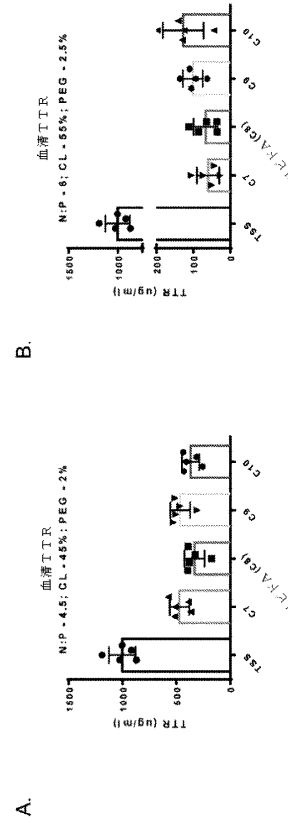
40

50

【図 6 - 2】



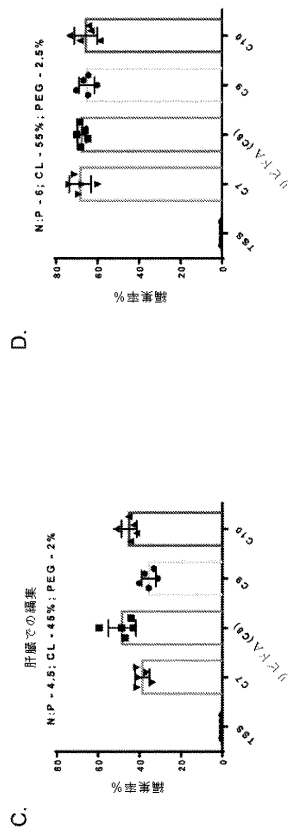
【図 7 - 1】



10

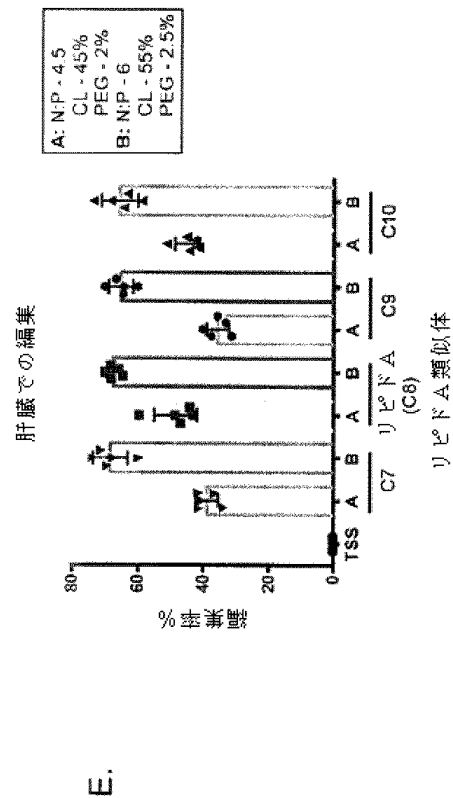
20

【図 7 - 2】



D.

【図 7 - 3】

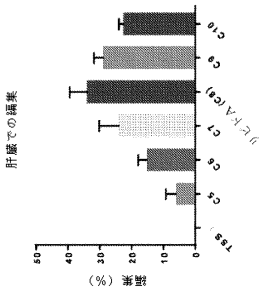


30

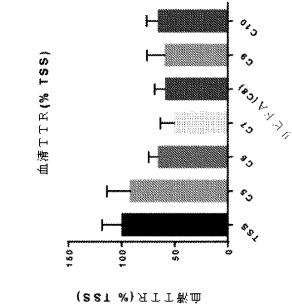
40

50

【図 7 - 4】



G.



F.

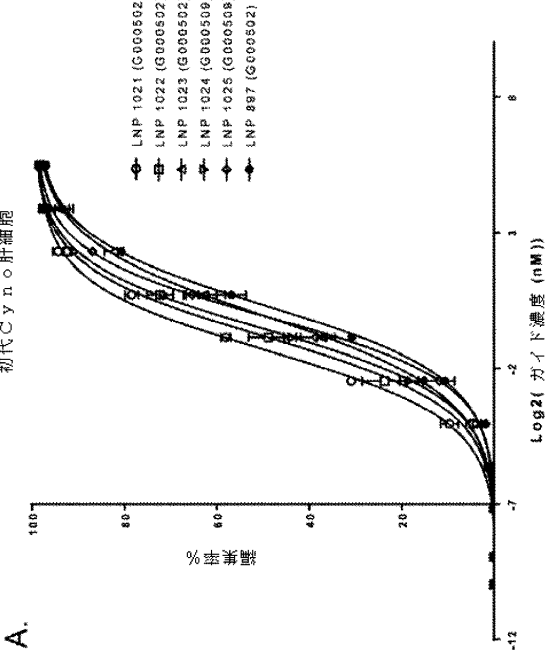
【図 8 - 2】

B

セット 1					
LNP 番号	LNP1021	LNP1022	LNP1023	LNP1024	LNP1025
ベスト	G000502-ロット1-50カチオン性	G000502-ロット1-55カチオン性	G000502-ロット2-45カチオン性	G000509-ロット1-50カチオン性	G000509-ロット1-55カチオン性
下限	-0.6786	-0.6886	-0.7413	-0.4768	-0.3509
上限	98.76	99.54	99.84	99.62	99.25
EC50	0.3906	0.5549	0.9429	0.6746	0.9373

セット 2					
LNP 番号	LNP1021	LNP1022	LNP1023	LNP1024	LNP1025
ベスト	G000502-ロット1-50カチオン性	G000502-ロット1-55カチオン性	G000502-ロット2-45カチオン性	G000509-ロット1-50カチオン性	G000509-ロット1-55カチオン性
下限	-0.8024	-0.3726	-0.474	-0.2294	-0.2362
上限	98.72	98.67	100.2	99.41	98.66
EC50	0.3733	0.5413	0.8982	0.63	0.9767

【図 8 - 1】

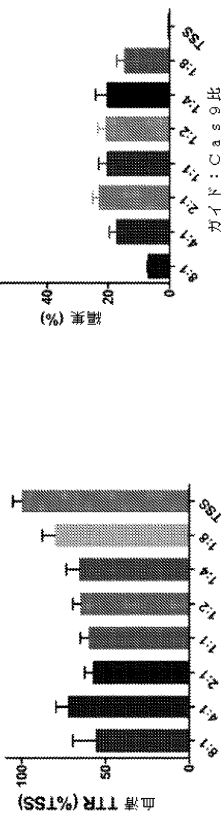


A.

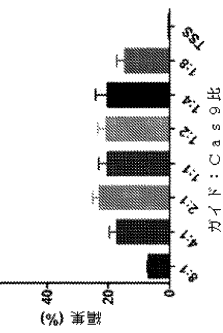
【図 9 - 1】

総カーゴ量 0.1MPK

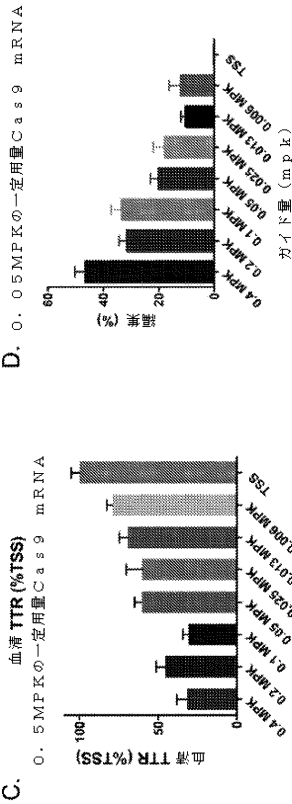
A.



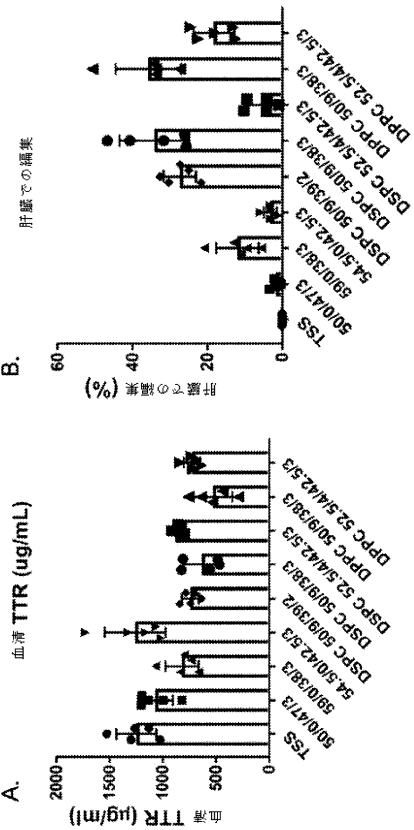
B.



【図 9 - 2】



【図 10】



【配列表】

0007645929000001.xml

フロントページの続き

(51)国際特許分類

F I

A 6 1 K	47/28	(2006.01)	A 6 1 K	47/28	
A 6 1 K	47/24	(2006.01)	A 6 1 K	47/24	
A 6 1 K	47/14	(2017.01)	A 6 1 K	47/14	
C 1 2 N	15/55	(2006.01)	C 1 2 N	15/55	
C 1 2 N	15/09	(2006.01)	C 1 2 N	15/09	1 1 0

ード 4 1

(72)発明者 シャー, ルチ, ルドラブラサド

アメリカ合衆国 0 1 7 0 1 マサチューセッツ州, フラミンガム, リバーパス ドライブ 9 0 ,
アパートメント 1 9

(72)発明者 スカリー, ステューブン, エス.

アメリカ合衆国 0 2 4 7 8 マサチューセッツ州, ベルモント, スレード ストリート 9 1

(72)発明者 マジュゾウブ, ラムシー

アメリカ合衆国 0 2 1 1 4 マサチューセッツ州, ボストン, ロングフェロー プレイス 1 , ナ
ンバー 7 1 6

審査官 阿川 寛樹

(56)参考文献 国際公開第 2 0 1 7 / 1 2 7 7 5 0 (WO, A 1)

国際公開第 2 0 1 5 / 0 9 5 3 4 0 (WO, A 1)

国際公開第 2 0 1 6 / 1 5 3 0 1 2 (WO, A 1)

国際公開第 2 0 1 7 / 0 9 3 8 0 4 (WO, A 1)

国際公開第 2 0 1 6 / 1 9 7 1 3 3 (WO, A 1)

国際公開第 2 0 1 7 / 1 0 9 7 5 7 (WO, A 1)

(58)調査した分野 (Int.Cl., D B名)

A 6 1 K 9 / 0 0 - 9 / 7 2 ; 4 7 / 0 0 - 4 7 / 6 9

A 6 1 K 3 8 / 0 0 - 3 8 / 5 8 ; 4 1 / 0 0 - 4 5 / 0 8 ; 4 8 / 0 0 ; 5 0 / 0 0 -
5 1 / 1 2

C A p l u s / M E D L I N E / E M B A S E / B I O S I S (S T N)