

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7278027号
(P7278027)

(45)発行日 令和5年5月19日(2023.5.19)

(24)登録日 令和5年5月11日(2023.5.11)

(51)国際特許分類

C 1 2 N	15/09 (2006.01)	F I	C 1 2 N	15/09	1 1 0
C 1 2 M	1/00 (2006.01)		C 1 2 N	15/09	1 0 0
C 1 2 N	9/16 (2006.01)		C 1 2 M	1/00	A
C 1 2 Q	1/02 (2006.01)		C 1 2 N	9/16	Z Z N A

C 1 2 Q 1/02 (2006.01)

C 1 2 N 9/16

C 1 2 Q 1/02

請求項の数 21 (全123頁)

(21)出願番号 特願2017-534913(P2017-534913)

(86)(22)出願日 平成28年1月12日(2016.1.12)

(65)公表番号 特表2018-500913(P2018-500913)

A)

(43)公表日 平成30年1月18日(2018.1.18)

(86)国際出願番号 PCT/US2016/013113

(87)国際公開番号 WO2016/115179

(87)国際公開日 平成28年7月21日(2016.7.21)

審査請求日 平成31年1月15日(2019.1.15)

審判番号 不服2021-6734(P2021-6734/J1)

審判請求日 令和3年5月25日(2021.5.25)

(31)優先権主張番号 62/102,347

(32)優先日 平成27年1月12日(2015.1.12)

(33)優先権主張国・地域又は機関

米国(US)

(73)特許権者 596060697

マサチューセッツ インスティテュート

オブ テクノロジー

アメリカ合衆国マサチューセッツ州 0 2

1 3 9 ケンブリッジ , マサチューセッツ

・アヴェニュー・7 7

(73)特許権者 507044516

プレジデント アンド フェローズ オブ

ハーバード カレッジ

アメリカ合衆国 マサチューセッツ 0 2

1 3 8 , ケンブリッジ , クインシー

ストリート 1 7

(74)代理人 100107984

弁理士 廣田 雅紀

(74)代理人 100182305

最終頁に続く

(54)【発明の名称】マイクロ流体送達による遺伝子編集

(57)【特許請求の範囲】**【請求項 1】**

細胞内への核酸 - タンパク質複合体の送達方法であって、前記核酸 - タンパク質複合体の核酸が静電相互作用を介して前記核酸 - タンパク質複合体のタンパク質と複合体化しており、前記方法が、

懸濁液中の細胞を準備すること；

細胞変形狭窄部を含むマイクロ流体チャネルに前記懸濁液を通過させ、前記核酸 - タンパク質複合体が通過するのに十分大きな前記細胞の摂動を誘導すること；ならびに前記細胞が前記狭窄部を通過する前、及び／またはその最中に、前記細胞を前記核酸 - タンパク質複合体と接触させること；

を含み、

前記狭窄部の直径が、2 ~ 1 0 μm であり、

前記マイクロ流体チャネルに前記懸濁液を通過させるために、1 0 ~ 2 0 0 p s i の圧力が用いられる、

前記方法。

【請求項 2】

前記細胞が前記狭窄部を通過する前に、前記細胞を前記核酸 - タンパク質複合体と接触させることを含む、請求項 1 に記載の方法。

【請求項 3】

前記細胞が前記狭窄部を通過する最中に、前記細胞を前記核酸 - タンパク質複合体と接

触させることを含む、請求項 1 に記載の方法。

【請求項 4】

前記核酸 - タンパク質複合体が、遺伝子編集コンポーネントを含む、請求項 1 ~ 3 のいずれか 1 項に記載の方法。

【請求項 5】

前記核酸 - タンパク質複合体が、
(a) リボ核タンパク質 (R N P) ; または
(b) 第 1 R N P 及び第 2 R N P ;
を含む、請求項 1 ~ 3 のいずれか 1 項に記載の方法。

【請求項 6】

前記核酸 - タンパク質複合体が、
(a) C a s タンパク質または C p f 1 タンパク質；ならびに
(b) シングルガイド RNA (s g R N A) または C R I S P R RNA (c r R N A)
及びトランスクアチベーティング c r R N A (t r a c r R N A) ; を含む、請求項 4 に記載の方法。

【請求項 7】

前記核酸 - タンパク質複合体が、C a s タンパク質または C p f 1 タンパク質及び s g R N A を含む R N P を含み、前記 C a s タンパク質または前記 C p f 1 タンパク質及び前記 s g R N A が、1 : 1 0 ~ 1 0 : 1 のモル比で複合体化された、請求項 6 に記載の方法。

【請求項 8】

前記 C a s タンパク質が C a s 9 タンパク質である、請求項 6 に記載の方法。

【請求項 9】

前記第 1 R N P 及び前記第 2 R N P が、両方ともニッカーゼである、請求項 5 に記載の方法。

【請求項 10】

前記第 1 R N P が、前記第 2 R N P の標的配列とは異なる標的配列にニックを入れる、請求項 5 に記載の方法。

【請求項 11】

前記懸濁液が、ドナー D N A をさらに含む、請求項 1 ~ 1 0 のいずれか 1 項に記載の方法。

【請求項 12】

前記細胞が前記狭窄部を通過する前、その最中、及び / またはその後に、前記懸濁液が、ドナー D N A をさらに含む、請求項 1 1 に記載の方法。

【請求項 13】

前記細胞が哺乳類細胞である、請求項 1 ~ 1 2 のいずれか 1 項に記載の方法。

【請求項 14】

前記細胞がヒト細胞である、請求項 1 ~ 1 3 のいずれか 1 項に記載の方法。

【請求項 15】

前記狭窄部の直径が、前記細胞の直径の 2 0 ~ 9 9 % である、請求項 1 ~ 1 4 のいずれか 1 項に記載の方法。

【請求項 16】

前記狭窄部の直径が、前記細胞の直径の 2 0 ~ 6 0 % である、請求項 1 ~ 1 5 のいずれか 1 項に記載の方法。

【請求項 17】

前記細胞が複数の細胞の 1 つであり、前記複数の細胞の各細胞を複数の並列マイクロ流体チャネルのうちの 1 つに通過させ、前記複数の並列マイクロ流体チャネルの各マイクロ流体チャネルが、細胞変形狭窄部を含む、請求項 1 ~ 1 6 のいずれか 1 項に記載の方法。

【請求項 18】

(a) 前記狭窄部の直径が、2、3、4、5、6、7、8、9、又は 1 0 μm であり；
(b) 前記狭窄部の長さが、1 0、1 5、2 0、2 4、3 0、4 0、5 0、6 0、7 0、

10

20

30

40

50

80、90、100、10~40、10~50、10~60、もしくは10~100 μm であり；

(c) 前記マイクロ流体チャネルに前記懸濁液を通過させるために使用される前記圧力が、10、20、30、40、50、60、70、80、90、100もしくは10~1000psiである；

(d) 前記細胞が、300、400、500、600、700、800、900、100~300、200~700、250~400、100~1000mm/s、1~1000mm/s、1m/s、2m/s、3m/s、4m/s、5m/s、6m/s、7m/s、8m/s、9m/s、10m/s、0.01~5m/s、5~10m/s、もしくは0.01~10m/sの速度で前記マイクロ流体チャネルを通過し；

(e) 前記マイクロ流体チャネルが、複数の細胞変形狭窄部を含み；

(f) 前記マイクロ流体チャネルが、単一の細胞変形狭窄部を含み；

(g) 前記細胞の前記摂動が、1~20、1~600、4、5、6、7、8、9、10、12、14、16、18、20、25、50、75、100、150、200、250、300、350、400、450、500、もしくは600nmの最大直径を含み；

(h) 1~20、1~600、4、5、6、7、8、9、10、12、14、16、18、20、25、50、75、100、150、200、250、300、350、400、450、500、もしくは600nmの最大直径を有する前記細胞の摂動が、少なくとも1、2、3、4、5、6、7、8、9、または10分間前記細胞上で持続する；及び/または

(i) 前記複数の並列マイクロ流体チャネルが、少なくとも2、5、10、20、25、30、40、45、50、75、100、500、1000、または2~1000のマイクロ流体チャネルを含む；

請求項1~17のいずれか1項に記載の方法。

【請求項19】

前記マイクロ流体チャネルが、複数の並列の、または直列した細胞変形狭窄部を含む、請求項1~18のいずれか1項に記載の方法。

【請求項20】

前記細胞に前記核酸-タンパク質複合体を送達した後、

(a) 前記細胞における標的遺伝子の発現が、少なくとも5、10、15、20、25、30、35、40、45、50、55、60、75、80、85、90、95、もしくは99%もしくはこれより多く減少する；または

(b) 前記細胞が複数の細胞であり、前記複数の細胞における標的遺伝子の発現が、少なくとも5、10、15、20、25、30、35、40、45、50、55、60、75、80、85、90、95、もしくは99%もしくはこれより多く減少する；

請求項4~8のいずれか1項に記載の方法。

【請求項21】

前記細胞に前記核酸-タンパク質複合体を送達した後、

(a) 前記細胞における標的遺伝子の発現が、少なくとも5、10、25、50、75、100、250、500%もしくはこれより多く増加する；または

(b) 前記細胞が複数の細胞であり、前記複数の細胞における標的遺伝子の発現が、少なくとも5、10、25、50、75、100、250、500%もしくはこれより多く増加する；

請求項4~8のいずれか1項に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願

本出願は、米国特許法第119条(e)の下で、2015年1月12日に出願された米国仮出願第62/102,347号の優先権による利益を主張し、その全内容を参照によ

り本明細書に援用する。

【0002】

連邦政府資金による研究の記載

本発明は、米国国立衛生研究所によって締結された認可番号R01GM101420-01A1の下、政府援助を受けて行われた。政府は本発明において一定の権利を有する。

【0003】

本明細書に記載する主題は、機械的細胞破碎により細胞内に遺伝子編集コンポーネントを導入することによる遺伝子編集に関する。

【0004】

配列表の参照

本出願は、517キロバイトのサイズであり、MS-Windowsとのオペレーティングシステム互換性を有するIBM-PCマシンフォーマットで2016年1月12日に作成された「38172-510001WO_Sequence_Listing_ST25.txt」という名称のファイルにあるヌクレオチド及び／またはアミノ酸配列を参照により援用し、これは本出願の一部としての2016年1月12日に出願されたテキストファイル中に含まれている。

10

【背景技術】

【0005】

ゲノム編集技術、例えばクラスター化等間隔短回文反復配列(CRISPR)-CRISPR関連タンパク質9(Cas9)及び転写活性化因子様エフェクタースクレアーゼ(TALEN)などは、細胞の遺伝子コードを変化させる能力に大きな可能性を秘めていることがわかっている。したがって、これらの技術は創薬における新規な洞察を可能とし、次世代遺伝子治療の開発につながり得る。タンパク質コンポーネント及び核酸コンポーネント、例えばデオキシリボ核酸(DNA)及び／またはリボ核酸(RNA)を含む遺伝子編集複合体は、細胞膜を容易に通過することができない。したがって、そのような複合体の送達が課題であった。

20

【発明の概要】

【0006】

本発明の方法及びシステムは、標的細胞への遺伝子編集コンポーネント及び遺伝子編集複合体の細胞内送達の問題を解決する。本明細書に記載する結果は、細胞膜の機械的破碎による遺伝子編集コンポーネント、例えばタンパク質、リボ核酸(RNA)、及びデオキシリボ核酸(DNA)の送達が遺伝子編集の成功につながることを示す。遺伝子編集物質の細胞内送達は現在の課題であるので、本方法は、潜在的に有害なウイルスベクターまたは電界を用いることなく標的細胞を改変するためのロバストな機構を提供する。さらに、プロセスのスケーラビリティ及び比較的単純性が幅広い普及に適している。本方策及び方法は研究及び治療におけるゲノム工学用途に適している。

30

【0007】

したがって、細胞内へのタンパク質-核酸複合体の送達方法は、懸濁液中の細胞を準備すること；細胞変形狭窄部(cell-deforming constriction)を含むマイクロ流体チャネルに溶液を通過させること；狭窄部に細胞を通過させ、その結果、細胞に圧力が加わり、前記タンパク質-核酸複合体が透過するのに十分大きな細胞の摂動(perturbation)を引き起こすこと；及び細胞が狭窄部を通過する前またはその後に複合体含有溶液中で細胞を所定の時間インキュベートすることによって行われる。例示的なタンパク質-核酸複合体は遺伝子編集コンポーネントを含む。例えば、タンパク質-核酸複合体はCasタンパク質(例えばCas9タンパク質など)及びガイドRNA(gRNA)またはドナーDNAを含む。他の例では、タンパク質-核酸複合体はTALENタンパク質、ジンクフィンガースクレアーゼ(ZFN)、メガヌクレアーゼ、またはCreリコンビナーゼを含む。

40

【0008】

本方法及びシステムは一般に、複合体、例えばタンパク質-タンパク質複合体、低分子+RNA複合体などのサイトゾル送達に適用可能である。

50

【 0 0 0 9 】

様々な標的細胞型をこのように処理する。例えば、細胞は哺乳類細胞、例えば免疫細胞（例えばT細胞）などまたは幹細胞、例えば造血幹細胞などを含む。

【 0 0 1 0 】

マイクロ流体システムは複数のマイクロ流体チャネルを含み得る。複数のマイクロ流体チャネルのそれぞれは内腔を画定し、バッファー中に懸濁した細胞が内腔を通過することができるよう構成される。いくつかの実施形態において、マイクロ流体チャネルは1つ以上の細胞変形狭窄部を含む。いくつかの実施形態において、狭窄部の直径は細胞の直径の閑数である。したがって、本発明のマイクロ流体システム内には多くのマイクロ流体チャネルが存在し得る。例えば、マイクロ流体システムは例えば2、5、10、20、40、45、50、75、100、500、1,000またはこれより多くの並列に並んだ複数のマイクロ流体チャネルを含み得る。10

【 0 0 1 1 】

複数の並列マイクロ流体チャネルを有するマイクロ流体システムは細胞へのペイロードの高スループット送達を可能にする。多くの細胞が各並列チャネルを1つずつ通過することができる。文脈に応じて、「細胞」という言及は1つ以上の細胞のことをいう場合があると理解されるであろう。

【 0 0 1 2 】

狭窄部の直径は処理しようとする細胞型の大きさに応じて選択する。いくつかの実施形態において、細胞は流体の流れによって主に圧縮され得る。いくつかの実施形態において、直径は細胞の直径未満である。例えば、狭窄部の直径は細胞の直径のほぼまたは約20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、または20～99%であり得る。例えば、狭窄部は細胞の直径のほぼ20～99%である、例えば狭窄部の直径は細胞の直径のほぼ60%である。狭窄部の直径の非限定例としては、ほぼまたは約2、3、4、5、6、7、8、9、10、15、20、2～10μm、または10～20μmが挙げられる。様々な長さの狭窄部も可能である。狭窄部の長さの非限定例としては、ほぼまたは約10、15、20、24、30、40、50、60、70、80、90、100、10～40、10～50、10～60、または10～100μmが挙げられる。20

【 0 0 1 3 】

本明細書に記載する主題は、複合体のコンポーネント（またはそのコンポーネントをコードする核酸）を断片的に送達する方法に勝る多くの技術的な利点を提供する。ほとんどの遺伝子編集システムでは複合体形成が細胞内で発生する必要があり、これは非効率的なプロセスであり得る。RNA及びCas（例えばCas9など）を複合体形態（または他のタンパク質／核酸遺伝子編集複合アセンブリ）で送達する利点としては、他の方法と比較してより良好な効率及び特異性が挙げられる。Casタンパク質（例えばCas9タンパク質など）を発現させるmRNAまたはDNA操作を不要とすることによって、細胞がタンパク質に曝露される時間を減少させ、それによりオフターゲット効果の可能性を減少させることができる。編集コンポーネントを複合体形態で細胞に送達することは、Cas（例えばCas9など）が細胞内の他のRNA鎖と複合体化し、誤った部位を切断するリスクも排除／最小化する。例えば、RNA単独では細胞内及び細胞外Toll様受容体（TLR）及びパターン認識受容体によって検出され、インターフェロン応答または他の抗ウイルス経路を促進し得る。複合体形態はこれらの経路と相互作用せず、そのため望ましくない副作用を回避することができる。30

【 0 0 1 4 】

細胞内への送達前のインピトロでの複合体化によって、Cas（例えばCas9など）及びgRNA複合体化反応を正確に制御することができ、ひいては最大限の機能が確保される一方で、サイトゾル内での複合体形成はそれほど効率的ではない場合がある。例えば、サイトゾルに複合体を送達することによって、Cas（例えばCas9など）及びgRNAと標的DNAとの同時相互作用が確実となる。本明細書に記載のインピトロで形成さ40

れて細胞に送達される複合体は完全な機能を有し、標的細胞の細胞質へのアクセスが得られればすぐに機能する。

【 0 0 1 5 】

送達プロセスは複合体の厳密なサイズ及び組成に依存せず、編集物質の複合体形成は最適条件の下、インピトロで発生して制御されるので、本明細書に記載のアプローチは、又クレアーゼをガイドする任意のタンパク質 + RNA / DNA ベースのシステムに関連する。

【 0 0 1 6 】

本発明の実施態様は以下の特徴の 1 つ以上も提供し得る。細胞の変形は、ほぼまたは約 $1 \mu s \sim 10 ms$ 、例えば $10 \mu s$ 、 $50 \mu s$ 、 $100 \mu s$ 、 $500 \mu s$ 、及び $750 \mu s$ 間の細胞の変形を含む。インキュベーションは 0.0001 秒 ~ 20 分またはこれより長く、例えばほぼまたは約 1 秒、約 30 秒、約 90 秒、約 270 秒、及び約 900 秒間発生する。10

【 0 0 1 7 】

細胞がマイクロ流体チャネルを通過する圧力及び速度も様々であり得る。いくつかの実施形態において、ほぼまたは約 $10 \sim 35 ps$ の圧力を用いて、細胞を含有する溶液をマイクロ流体チャネルに通過させる。速度は様々な理由で調節され得るが、高ペイロード送達を維持しながら処理される細胞の生存率を改善することが挙げられる。いくつかの実施形態において、細胞はほぼまたは約 $300 mm/s$ 、 $400 mm/s$ 、 $500 mm/s$ 、 $600 mm/s$ 、 $700 mm/s$ 、 $800 mm/s$ 、 $900 mm/s$ 、 $100 \sim 300 mm/s$ 、 $200 \sim 700 mm/s$ 、 $250 \sim 400 mm/s$ 、 $1 \sim 1000 mm/s$ 、 $1 m/s$ 、 $2 m/s$ 、 $3 m/s$ 、 $4 m/s$ 、 $5 m/s$ 、 $6 m/s$ 、 $7 m/s$ 、 $8 m/s$ 、 $9 m/s$ 、 $10 m/s$ 、 $0.01 \sim 5 m/s$ 、 $5 \sim 10 m/s$ 、または $0.01 \sim 10 m/s$ の速度でマイクロ流体チャネルを通過する。細胞が複数の細胞である場合、ほぼまたは約 5 、 10 、 20 、 30 、 40 、 50 、 60 、 70 、 75 、 80 、 85 、 90 、 91 、 92 、 93 、 94 、 95 、 96 、 97 、 98 、 99 、 $90 \sim 95$ 、もしくは $80 \sim 99\%$ またはこれより多くの細胞が狭窄部を通過した後で生存可能であり得る。いくつかの実施形態において、細胞は狭窄部を通過した後、少なくとも約 1 時間、 2 時間、 3 時間、 4 時間、 5 時間、 6 時間、 7 時間、 8 時間、 9 時間、 10 時間、 12 時間、 18 時間、 24 時間、または 48 時間のいずれかの間生存可能である。20

【 0 0 1 8 】

いくつかの例では、本発明のデバイスは 2 、 10 、 20 、 25 、 45 、 50 、 75 、 100 またはこれより多くのチャネルを含む。いくつかの実施形態において、圧力の印加によってチャネルまたは管路を通して細胞を移動させる、例えば押し進める。いくつかの実施形態において、細胞推進体が圧力を印加することができる。細胞推進体としては、例えば、圧力ポンプ、ガスボンベ、圧縮機、真空ポンプ、シリンジ、シリンジポンプ、蠕動ポンプ、手動シリンジ、ピペット、ピストン、毛管作用体、及び重力を挙げることができる。チャネルに代えて、細胞をネットの形態の狭窄部に通過させてもよい。いずれの場合においても、細胞が通過する狭窄部の幅は、処理しようとする細胞の未狭窄状態、すなわち懸濁状態での幅または直径の $20 \sim 99\%$ である。温度が組成物の取込みに影響を及ぼし、また生存率に影響を及ぼし得る。30

【 0 0 1 9 】

ある実施形態において、 $0 \sim 45$ 、例えば $0 \sim 25$ の温度を細胞の処理中に用いる。様々な実施形態において、室温（例えば 20 ）、生理的温度（例えば 39 ）、生理的温度よりも高い温度、もしくは低温（例えば 0.1 ）、またはこれらの例示的温度の間の温度（例えば $0.1 \sim 40$ ）で本方法を実施する。40

【 0 0 2 0 】

いくつかの実施形態において、免疫細胞に関して、未刺激 T 細胞、 B 細胞及び / または单球の処理は、 $4 \sim 8$ の温度、例えば氷上で実施する。別の例では、樹状細胞、活性化 T 細胞、及び / または活性化 B 細胞は、 $20 \sim 25$ の温度、例えば典型的な室温で本デバイスを用いて処理する。50

【0021】

いくつかの実施形態において、狭窄、伸展、及び／または高せん断速度のパルスによる細胞への制御した損傷（例えば摂動）の後、細胞内への導入を望む複合体を含有する送達溶液中で細胞をインキュベートする。制御した損傷は、細胞膜における小さな、例えば直径 200 nm の摂動として特徴づけられ得る。狭窄部を通過したことに起因する損傷が塞がる細胞の修復期間は数分のオーダーである。送達期間は、1～10 分またはこれより長い時間、例えば 15、20、30、60 分以上を含み、室温で作業する場合には 2～5 分が最適である。

【0022】

本明細書に記載のデバイス及び方法のいくつかの実施形態において、幹細胞または前駆細胞、例えば人工多能性幹細胞（iPSC）などの狭窄部チャネルの通過は分化を誘発しないが、細胞内への組成物の取込みを確実に誘発する。例えば、細胞内に因子（複数可）を導入した方法に伴う問題が生じることなく、そのような細胞内に遺伝子編集化合物が導入される。

10

【0023】

細胞膜における一時的な摂動のサイズ及び持続時間は、様々な要因、例えば細胞変形狭窄部の直径及び細胞が狭窄部を通過する速度などを調節することによって修正することができる。本明細書に示す摂動のサイズ及び持続時間に関する開示は、限定するものと解釈すべきではない。Sharei et al., (2014) Integr. Bio 1., 6, 470-475 に摂動及び修復の非限定的な記載が提供されており、その内容全体を参照により本明細書に援用する。いくつかの実施形態において、細胞膜の摂動はほぼまたは約 1～20、1～600、4、5、6、7、8、9、10、12、14、16、18、20、25、50、75、100、150、200、250、300、350、400、450、500、または 600 nm の最大直径によって特徴づけられ得る。様々な実施形態において、ほぼまたは約 1～20、1～600、4、5、6、7、8、9、10、12、14、16、18、20、25、50、75、100、150、200、250、300、350、400、450、500、または 600 nm の最大直径を有する細胞膜の摂動は、少なくともほぼまたは約 1、2、3、4、5、6、7、8、9、10、もしくは 1～10 分間またはこれより長く（11、13、15、18、20 分以上）細胞膜上で持続する。

20

【0024】

様々な実施形態において、直径は細胞の直径未満である。例えば、狭窄部の直径は細胞の直径のほぼまたは約 20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、または 20～99 % であり得る。狭窄部の直径の非限定例としては、ほぼまたは約 4、5、6、7、8、9、10、15、20、4～10 μm、または 10～20 μm が挙げられる。様々な長さの狭窄部も可能である。狭窄部の長さの非限定例としては、ほぼまたは約 10、15、20、24、30、40、50、60、70、80、90、100、10～40、10～50、10～60、または 10～100 μm が挙げられる。

30

【0025】

多くの細胞は直径 5～20 μm であり、例えば未刺激 T 細胞は直径 7～8 μm である。例えば、単細胞の処理に関しては、狭窄部の直径は 4.5、5.5、6、または 6.5 μm である。別の例では、ヒト卵の処理のための狭窄部のサイズ／直径は 60 μm～80 μm であるが、より大きい及びより小さい狭窄部が可能である（ヒト卵子の直径はおよそ 100 μm である）。また別の例では、12 μm～17 μm の狭窄部直径を用いて胚（例えば 2～3 細胞のクラスター）を処理する。未刺激 T 及び B 細胞に関する非限定例において、デバイスは、約 10、15、20、25、30、もしくは 10～30 μm の長さ、約 3、3.5、4、もしくは 3～4 μm の幅、約 15、20、25、もしくは 15～25 μm の深さ、及び／または約 5、6、7、8、9、10、11、12、13、14、15、もしくは 5～15 度の角度を有する狭窄部を含む。免疫細胞へのペイロードの送達に

40

50

有用なマイクロ流体デバイスの例は、2015年10月30日に出願されたPCT国際特許出願第PCT/US2015/058489号, Delivery of Biomolecules to Immune Cellsに記載されており、その内容全体を参照により本明細書に援用する。

【0026】

単細胞に加えて、非常に大きな細胞、例えば卵（直径およそ200μm）、細胞のクラスター、例えば2～3細胞を含む胚などの2～5細胞クラスターさえも標的組成物を取り込むように処理される。開口部のサイズは適宜、すなわち狭窄部の幅がちょうどクラスターのサイズ未満となるように調節する。例えば、チャネルの幅は細胞クラスターの幅の20～99%である。

10

【0027】

細胞または細胞クラスターは、所望の細胞型を精製／単離または富化する。本方法で用いられる樹状細胞もしくは他の細胞、例えばマクロファージ、B細胞、T細胞などの免疫細胞、または幹細胞、例えば胚性幹細胞もしくはiPSなどは、精製または富化される。例えば、細胞表面マーカーの発現または他の同定できる特徴によって細胞を単離または富化する。樹状細胞は、-インターフィリン(integrin)、CD11cまたは他の同定できる細胞表面マーカーの発現によって同定及び単離する。細胞に関して、用語「単離」は、細胞が天然ではともに発生する他の細胞型または細胞物質を実質的に含まないことを意味する。例えば、特定の組織型または表現型の細胞のサンプルは、細胞集団の少なくとも60%である場合に「実質的に純粋」である。好ましくは、調製物は細胞集団の少なくとも75%であり、より好ましくは少なくとも90%、最も好ましくは少なくとも99%または100%である。純度は任意の適切な標準的方法によって、例えば蛍光活性化細胞ソーティング(FACS)によって測定する。

20

【0028】

ペイロード組成物、例えばポリヌクレオチド、ポリペプチド、または他の作用物質（例えばCas9及びgRNA）などを精製及び／または単離する。具体的には、本明細書で用いる場合、「単離」または「精製」核酸分子、ポリヌクレオチド、ポリペプチド、またはタンパク質は、組換え技術によって產生した場合には他の細胞物質、もしくは培地を、または化学合成した場合には化学的前駆体もしくは他の化学物質を実質的に含まない。精製化合物は、対象とする化合物の少なくとも60重量%（乾燥重量）である。好ましくは、調製物は対象とする化合物の少なくとも75重量%であり、より好ましくは少なくとも90重量%、最も好ましくは少なくとも99重量%である。例えば、精製化合物は、重量基準で所望の化合物の少なくとも90%、91%、92%、93%、94%、95%、98%、99%、または100%（w/w）のものである。純度は任意の適切な標準的方法によって、例えばカラムクロマトグラフィー、薄層クロマトグラフィー、または高速液体クロマトグラフィー(HPLC)分析によって測定する。精製または単離ポリヌクレオチド（リボ核酸(RNA)またはデオキシリボ核酸(DNA)）は、その自然発生状態においてそれと隣接する遺伝子または配列を含まない。単離または精製核酸分子の例としては、(a)自然発生ゲノムDNA分子の一部であるが、それが天然に発生する生物のゲノムにおける分子のその部分に隣接する核酸配列が両方とも隣接していないDNA；(b)結果として生じる分子がいかなる自然発生ベクターまたはゲノムDNAとも同一でないように、ベクターまたは原核生物もしくは真核生物のゲノムDNA中に組み込まれた核酸；(c)分離した分子、例えばcDNA、ゲノム断片、ポリメラーゼ連鎖反応(PCR)によって生成された断片、または制限断片など；及び(d)ハイブリッド遺伝子、すなわち融合タンパク質をコードする遺伝子の一部である組換えヌクレオチド配列が挙げられる。本発明に従う単離核酸分子には、合成的に生成された分子、ならびに化学的に改変された及び／または修飾主鎖を有する任意の核酸がさらに含まれる。

30

【0029】

精製モジュールまたはコンポーネント、例えば精製タンパク質（複数可）ならびに精製核酸（RNA及び／またはDNA）から複合体を調製する。したがって、コンポーネント

40

50

の比は、所望の遺伝子編集結果が得られるように所望により制御及び調整される。本方法は、感受性の高いペイロード、例えばタンパク質 - RNA / DNA 複合体、例えば 40 kDa、50 kDa、75 kDa、100 kDa、及び最大 120、130、150、200、250、300 kDa 以上の複合体の送達に特に好適である。

【0030】

フローバッファー中に界面活性剤（例えば 0.1 ~ 10% w/w）（例えばポロキサマー、動物由来血清、アルブミンタンパク質）を任意により用いる。細胞内への分子の送達は界面活性剤の存在に影響されないが、動作中のデバイスの閉塞を軽減するために界面活性剤を任意により用いる。

【0031】

いくつかの態様において、デバイスは、シリコン、金属（例えばステンレス鋼）、プラスチック（例えばポリスチレン）、セラミック、または 1 つ以上の適切なサイズのチャネルもしくは管路を形成するのに適した任意の他の材料から作製される。いくつかの態様において、デバイスは、ミクロンスケールの形体をエッチングするのに適した材料で形成され、細胞が通過する 1 つ以上のチャネルまたは管路を含む。シリコンが特に好適であるが、その理由は、この材料での微小パターン形成法が十分に確立されており、したがって、新たなデバイスの作製、デザインの変更などがより容易であるためである。加えて、シリコンの剛性により、ポリジメチルシロキサン（PDMS）のようなより軟質の基板に勝る利点を得ることができるが、例えばより高い送達率である。例えば、デバイスは 2、10、20、25、45、50、75、100 またはこれより多くのチャネルを含む。シリコンをエッチングすることによってデバイスを微細加工する。圧力の印加によってチャネルまたは管路を通して細胞を移動させる、例えば押し進める。細胞推進体が圧力を印加することができる。細胞推進体としては、例えば、圧力ポンプ、ガスボンベ、圧縮機、真空ポンプ、シリンジ、シリンジポンプ、蠕動ポンプ、手動シリンジ、ピペット、ピストン、毛管作用体、及び重力を挙げることができる。チャネルに代えて、細胞をネットの形態の狭窄部に通過させてもよい。いずれの場合においても、細胞が通過する狭窄部の幅は、処理しようとする細胞の末狭窄状態、すなわち懸濁状態での幅または直径の 20 ~ 99 % である。

【0032】

本発明の様々な実施態様は以下の臨床及び研究における可能性の 1 つ以上も提供し得る。スクリーニングの向上及び投与量研究のための、細胞モデルへの遺伝子編集複合体またはそのコンポーネントの定量的送達を実現することができる。本方法は、タンパク質治療薬の同定または疾患機序の理解に役立つサイトゾル内でのタンパク質活性をスクリーニングする高スループット方法として採用され得る。本デバイス及び技術は、特定の循環血液細胞のサブセット（例えばリンパ球）、さらには全血への遺伝子編集複合体の細胞内送達；細胞、特に卵母細胞及び接合子内への複合体またはそのコンポーネントの高スループット送達；iPS 細胞を生成するための細胞リプログラミングを誘発する遺伝子編集を（任意によりドナー DNA などの遺伝物質とともに）導入することによる標的化した細胞分化；遺伝子導入または変異幹細胞株の開発のための胚性幹細胞内への DNA 及び / または組換え酵素の送達；遺伝子導入または変異生物の開発のための接合子内への DNA 及び / または組換え酵素の送達；樹状細胞（DC）の細胞活性化；iPS 細胞生成；1 つ以上の遺伝子の細胞機能及び / または疾患への寄与を調べるための正常または疾患細胞（例えば癌細胞など）における変異の作製；ならびに幹細胞分化に有用である。形成外科に関連して用いられる皮膚細胞も、本明細書に記載のデバイス及び方法を用いて改変される。本明細書に開示の遺伝子編集タンパク質の送達方法は、遺伝子疾患及び他の疾患を治療するための、CAR-T 細胞の生成、または造血幹細胞（HSC）の遺伝子改変にも用いられ得る。HSC に関する実施形態において、被検者は、オートロガス、シンジェニック、またはアロジエニック編集 HSC を与えられ得る。様々な実施形態において、被検者に遺伝子編集細胞を与える前に被検者の細胞が焼灼され得る。例えば、被検者に遺伝子編集 HSC を与える前に被検者の骨髄細胞が放射線でまたは化学的に焼灼され得る。いくつかの実施形

10

20

30

40

50

態において、本明細書に開示の方法または組成物を用いて、サラセニアまたは鎌状赤血球貧血に関連する遺伝子を編集する。本発明に従って、エクスピボまたはインビトロ、すなわち被検者の体外で処理した細胞は、病変の治療または診断を必要とする被検者へのその後の投与に有用である。代替的実施形態では、インビボ細胞プロセシングを行う。

【0033】

様々な実施形態において、SHP2遺伝子を編集／変異させ、その活性を減少させるか、またはSHP2発現をノックアウトもしくは減少させる。T細胞の遺伝子編集に関するそのような実施形態において、T細胞は免疫抑制シグナルへの応答性が低下し、腫瘍に対する活性が増加する。そのような実施形態において、T細胞は腫瘍抗原への応答性が高くなり、癌の治療により効果的となり得る。

10

【0034】

本主題の態様は、細胞へのタンパク質・タンパク質複合体及びタンパク質・核酸複合体、例えば遺伝子編集複合体の迅速かつ一時的な送達に関する。複合体の核酸コンポーネントは、デオキシ核酸(DNA)、リボ核酸(RNA、例えばmRNA、gRNA)またはそれぞれ他の二本鎖もしくは一本鎖核酸化合物を含む。例えば、遺伝子編集複合体(例えばリボ核タンパク質(RNP))の送達は、遺伝子編集複合体のコンポーネント(例えばCasタンパク質及びgRNA)をコードする発現ベクターを細胞に送達した場合よりも早く遺伝子編集を達成し得る。例えば、同じ細胞が遺伝子編集複合体コンポーネントをコードする発現ベクターのマイクロ流体またはエレクトロポレーション媒介送達を受けた場合よりも0.1、0.5、1、1.5、2、2.5、3、3.5、4、または0.1～4時間早く、細胞内で遺伝子が編集(例えば変異または置換)され得る。

20

【0035】

本明細書に開示の各実施形態は、他の開示した実施形態のそれぞれに適用可能なものと企図される。したがって、本明細書に記載の様々な要素のすべての組合せが本発明の範囲内である。

【0036】

関連する機器、システム、技術、及び物品も記載する。

【0037】

本明細書に記載する主題の1つ以上の変形の詳細を添付図面及び以下の説明で示す。本明細書に記載する主題の他の特徴及び利点が、説明及び図面、ならびに特許請求の範囲から明らかとなるであろう。

30

【図面の簡単な説明】

【0038】

【図1A】マイクロ流体システムの概略図であり、細胞は狭窄部を通過した後で送達物質(ペイロード)に曝露される。

【図1B】マイクロ流体システムの概略図であり、細胞は狭窄部を通過した後で送達物質(ペイロード)に曝露される。

【図2A】マイクロ流体システムの実施形態の概略図であり、チャネル深さ、幅、及び長さなどのパラメータを示している。

【図2B】マイクロ流体システムの実施形態の概略図であり、チャネル深さ、幅、及び長さなどのパラメータを示している。

40

【図3】図3A及び図3Bは、図1～2に示したマイクロ流体デバイスを用いて細胞質に遺伝子編集物質を送達されたK562レポーター細胞における、一連のフローサイトメトリープロット(図3A)及び組換え効率の棒グラフ(図3B)である。用いたレポーターについて、細胞株にはフレームシフトGFP遺伝子が存在する。編集を行うために、Cas9-gRNA複合体及びドナーオリゴヌクレオチドが送達されるであろう。複合体はGFP部位付近を切断し、オリゴヌクレオチドが切断部位に挿入される。オリゴの挿入に成功すると遺伝子が修正され、結果としてGFP発現が生じるが、それがこの図に見られるものである。したがって、このアッセイにおいて、あるものがオフであったのに対して遺伝子編集によりオンとなった。遺伝子(または文脈に応じてヌクレオチドもしくはその一

50

部)がドナーオリゴヌクレオチドの不在下で欠失していた場合、遺伝子はオフになると予想される。

【図4】F A C S を用いて得られた4つの異なる細胞集団についての送達されたデキストランに対するM H C クラス1の₂ミクログロプリンコンポーネント(B 2 M)発現を示す一連のF A C s プロットである。9 0 p s i で3 0 - 4 チップを用いたR N P の送達の結果、エンドサイトーシス対照と比較してB 2 M発現は5 4 . 4 % 減少する一方で、9 0 p s i で1 0 - 4 チップの結果、B 2 M発現の減少は2 5 . 2 % となる。C A S 9 対照におけるB 2 M発現はエンドサイトーシス対照と有意差がない。狭窄部が長いチップの方が、R N P 複合体の送達が多く、B 2 M発現の減少が大きいという結果になる。

【図5】指定の条件での機能的な編集の尺度としての、F A C S によって測定した投与量依存的なB 2 M発現の減少を示す一連のF A C s プロットである。10

【図6】図6の最初の画面から図6の最後の画面にまで及ぶF o x P 3 ゲノム配列(配列番号5 6)を示す。F o x P 3 に属するエキソンを下線及びハイライトした文字で示す。F o x P 3 に属していないこの領域内の他のエキソンを下線のないハイライトした文字で示す。

【図7】F o x P 3 から翻訳されたアミノ酸配列(配列番号5 7)を示す。交互にあるエキソンを下線あり及び下線なしで示す。イタリック体の太字はオーバーラップスプライス部位を示す。

【図8】図8の最初の画面から図8の最後の画面にまで及ぶS H P 1 ゲノム配列(配列番号5 8)を示す。S H P 1 に属するエキソンを下線及びハイライトした文字で示す。S H P 1 に属していないこの領域内の他のエキソンを下線のないハイライトした文字で示す。20

【図9】S H P 1 から翻訳されたアミノ酸配列(配列番号5 9)を示す。交互にあるエキソンを下線あり及び下線なしで示す。イタリック体の太字はオーバーラップスプライス部位を示す。

【図10】図10の最初の画面から図10の最後の画面にまで及ぶS H P 2 ゲノム配列(配列番号6 0)を示す。S H P 2 に属するエキソンを下線及びハイライトした文字で示す。配列番号6 0 は以下の通りでもある。

【図11】S H P 2 から翻訳されたアミノ酸配列(配列番号6 1)を示す。交互にあるエキソンを下線あり及び下線なしで示す。イタリック体の太字はオーバーラップスプライス部位を示す。30

【0 0 3 9】

(配列番号6 0)

AGGCTCAAGCAATCCTCTCACCTCAGCCTCCCGAGTAGCTGGGACTACAGGCAGCGGCCA
 CCACGCCCGGTAATTGTATTTGTAGAGATGGGATTCACTATTTGCCCGGC
 TGGTCCCAACTCCTGGACTCAAGCGATTGCCCGCTCAGCCTCCAAAGGAAAGTGCT
 GGGATTCAGGCAGTGTGCCACCCTCCCACCCAAAGTAGTATTATGTAATTATTATT
 ATTATTTGAGACGGAGTCTCGCTATTGCCAGGCTGGAGTGCAGTGGCGCGATCTCG
 CTCAATGCAACCTCTGCCCTCCGGGTCAGCGATTCTCTGCTCAGACTCCAAAGCAG
 CTGGGACTACAGGCAGCCCCCACCACGCCAGGCTAATTCTGAATTAGTGGAGACGG
 GTTTACCATGTTGCCAGGATGGTCTCGATCTTGACCTCGTGTACCGGCCACCTCG
 GCCTCCCAAAGTGCTGGGATTACAGGCAGGCCACCGGCCAGCTATTATTATT
 TTAGGCAGTGTCTGCCCTGCGTCAGGGTGTAGTGCAGTGGCGTGTACGACTCACT
GCAGCCCCACTTCTCGGGTTAAGTTATCTCCGCCGAGCCTCACGCCGGTTAGT
 TTTTGATTTTGATAGAGATGAGGTCTTGTCTTTGCCAGGCTGGCCTCGAATC
 TTGGCTTAAGCGAACCTTGTGCCAGCCTCCAAAGTGGGATTACGGCGTGAACC
 ACCGCCAGCCTACTATCTTATCTTACAG**AAAGAAAAAATGGAGGAAACCGAGGCT**
CGGAGACAGTAGGTAACTTCCCAAGGTTCCACAGCTAATGAGTGGAGCGCGATTGTG
GAACGAAATGAATGAATCGATGTGGCAGCGGGCCCGAGCGGGTCGGCGTAGACGCG
GAGCGCAGCTCACACCTGGCGCCCGGGTTCCAGGAGGAAGCAAGGATGCTTGGAC
ACTGTGCGTGGCGCTCCCGCGAGCCCCCGCGCTGCCATTCCGGCGTCGTCGGTCT
CCGCTGAGGGAAAGCAGGAAGTGGCGGGCGTGCAGCGGGTGACATCACGGGGCGA
CGGCGCGAAGGGGGGGGGGGAGGAGGGAGCGAGCCGGGGCGGGGGCAGCTGCACAGTC
TCCGGGATCCCCAGGCCGGGGGTCTGTGCGCGCCGGCTGGCTCTGCCCGCGTC
CGGTCCCCAGGGCCCTCCCTGGGCCAGCCGATGTGACCGAGGCCAGCGAGCTGAG
CAAGGAGCGGGTCCGCGAGCCGGAGGGCGGGAGGAACATGACATCGCGGAGGTGAG
 GAGCCCCGAGGGGCCGGCGGGCCTCGCCCGGCCACGCCGCGTCGGTTAGCCCCG
 TCCGGAAGGGGGCGCCCCGGGGCTCGGGCTCCGCCCGGGCTGGGGTTGGGGCC
 GGTTCCCTCTCGTCCCTCGCCCTCCAGGGCGGGGGCGGCCACCGCGCCCCAC
 CCCTCGGTCCCCATTCTTCTGCCCTCCAGGGCTGGCAGCTGGCTCGGGCAGCCCCGGGAT
 GCCCGTCAAGGCCGGGGCAGGTAGAGCGCCAGGGAAACACGGGTGCCAGCGCAGGC
 TCAGCGCCGCAATTCTGACCCATTGCCATGAGAATTGCCATGGTGATTCCGAAATA
 ACCCTGCTCACTTGGGAGGCTCTGGGACACGAGAGGGAGTTGCGCGGGGGCG
 CCCAGTGGCTAGTCGTTCTGGCTACTGTGCCACTTCGTGCATTGGGACTTCACGC
 AGGACCCCTGACCTTTATATGCTCTTTGTGTCTTCTTCTCTACCGTCACTG
 CCAGAAATGGAAAAGTGAATGTATCTGAGCCACTAGAAGTATTCTCTCTGCGAT
 CTCGCTTGGGAGATGAAAGGAGGGCGCATCTCGTTATTAACTCTTCACTGCA
 ACCTAACAGTCAGGTCACTTACTGGTACCGTTATGGATGAGGAACCGAGGGCCA
 GAAGCAACATGCTAGTAAATGACAAGATTGAAACTTAGGAGGATTAGTGAGTTAATGAG
 ATCCTTGAAAGGTAGGGTAATACTACTAATAGCTAACATTGCTTAGTTGACC
 ACAGCCCTATCAGATGGCTACTATTATCCCCATTGTAAGAGTGAGAACCGAGTTCA
 AGGTTAAGTAAATTGCTAACCTCACAGCTAGTAGGTGGGGAGACAGAATCCCTACTTT
 TAATCACTATGTTGCTCTATTATTTGTAACTATTGCTAACCTTGTAAAGCCTTAATT
 TTGTTGCAAACAGTAGTGTGACCTGTTGTTCAAGATAGTGATCCTGCTATTGTTATA
 GTCACTCTATACCACTCACACTTAAGACCCATTGCTATTCTTCCATGATTGTTCA
 ATTATGGTCACTGTCAGACATTAAAAACGATTCAAGCTATTGAGGCTATTGAATG
 AGATTTCTTTCTTTTTTTGGAGACGGAGGCTCACTCTGTTGCCAG
 GCTGGAGTGCAGTGGCGCAATCTCGGCTCACCAATCTCCGCTCTAGGTTCAAGCGA
 TTCTCTGCCCTAGGCCCAAGTAACACTAGGACTACAGGCACCAACTATGCCGGCTAA
 TTTTGATTTTAGTAGAGACAGGGTTCACTATGTTGCCAGGCTGGTCTAAACTCC
 TGACCTCGTGTACCGCCCGCCTGGCCTCCAAAGTGCTGGATTACAGGCAGGCCAC
 CGTACCCAGCCTGAATGAGATTTCAAAATATTAGGAATGTCTCCAAACACACCTG

10

20

30

40

50

GCATGTTATTACATGGATCTGGAATTAAAAAGGGGAGAAAAGAAAAGT GAGA ACT
 CGTAGGAAGTGAGTGACTTGGACAGGTCGGTGGCAAGTGCTTACAGATCTGGTAATAT
 ATAAC TGCATTCAACAGAACAGTGTATAGCCTCAAATGTTCTAATTCTTAGGGAGCTT
 TTAAATAAACAGTTGCTATTCTTAATCTGCAAATAGTCATTGAGCCTTTGTTCCCTG
 GTGTCTGCTCTCCAGACAAGTAAGGATCTGCTGCTTAGGGAGACATCAGACGGGCTGG
 GGGTGGGAAAAGGCTGGTAGTAATAGACCC TACATTGTCAGTTGTTCTTAAAG
 GCATAGAAGTGTGGCATAGTCAAAGTAGCAAGTGGTAAGAGT GACAGTTGAAATGGAG
 TAATT CCTCTCCCTCCAGGCCCTGGTATTATGCACCACCCAAAAGCCGGTTATGAAC
 ATAATACACATAATTGAATGATT CATTATTTGGATTATAAGCCTGTTTATTGT
 TAACCAGCCTTAATGAGGTATAATGACATGCAATTATGCATATATTTAAATGTACAA
 TTTGATCAGTTGACATACATACACTTGGAAACCACCCATAGTCAGATAATGA
 ACACATCTATCACCCCTGGTAATTG CCTTATGTTCTTATAATCCTCCTTGTCTT
 AGGCAGCCACTATTCTGCTTCTGCACTATGTATTAGTTGCATTCCCTAGAATT
 TTTAAAAATTAAATTGTTGAATAGAGATGGGCTCACTGTGTTGCCAGGGCAG
 TCTCAAACCTGGGTCAGTGATCCTCTCACCTGGCCTCTGAAGTGTGGATTAT
 AGGCATGAGACACCCTGCCAGCCTAGAATT TATTATTGTTATTGTGTTTT
 TTGAGATAGGGTCTCACTTGTGCCCAGGCTGGAGTGCACTGGTGCATCAGC
 TTGTTCTAGGCTCAATCCATCCCCCTCCTCAGCTTCCGGTTACTGGGCTACAGG
 TGTGCACCACACCACCCGGTAATT TGTATTAGAGACAGGGTTTGCATG
 TTGGCCAGGCTGGTCTAAACTCCGGGCTCAAGCGATCTCCTGCCTCGGCCTCCAAA
 GTGCTGGGATTACAGGCATGAGCTATTGCGTCCGCCTCAAATTACTTAACTAGTAT
 TAATTCAACAGGAAGTTAATGAGC CAGGAGATAAGCAGTAAGATAGGAAAATA
 TTGCTATTTCATGGCTGAGAGAGAGCAGACAAACACATGACTAAATAGGGCAATT
 GTAGTAATAATTCTAGGAGGGAAAAAATCCACAGAAATGTGAGGATGGGAGAATGCAG
 TTAGTTTGATAGGGTTAGAGAAGGTGATGCTGTGAGCTGACACTGAATGACAATT
 AGTAGTCTGAATT TGTGTTGCTTAATTATCAAATAACTCCTTGTGGTCGGCTTTA
 TATGCATCCAGTAATTAAAATGTAAGTATATCAATGTACTGATATCTCTCAGCATCATA
 GGTAGAAAAC TAAGGCATT CAGCAATTAAAGTGACTCCTCCCTGATCATGTAGCAGTGA
 TAGTACTGGATTAGATT TGAGGGTGTGCTCTGCCTTGTGAAACCAA
 CAAAGCTGCCTGTATT TCCAACCTCTCAGCATGTGGTACCTCCTTACATCTGTT
 TTTGTTGCTCTGAAATCCATACCGCAGATGAGCTGAGAGGGGAGAAAATTGAGCTGT
 TCTGAGACTGGAGGCTTGGTTATCTCTGCAGGTCAAGTACATTTGCTCTGGC
 TCCCTGGTGGCCACGTTGTTATCTCTGCAGGAGTAAAATAACTTGCTTGCTGAAAA
 ATAACAGTTCTGTGCTTGCACTGGAAACTGGGATGTCTTATTACGTTAGGTCTGA
 TGTAAGGCCAAGTTTGTTAGAGTGTCAAGTGCAGAGGCCACTGCTAAGATGACTT
 ACCCTCGTGTCCATGGTCAATGTGGAGACTGTTATGAGTGGCACATGATGCTGGAAAAG
 CAGAGCCAAC TCGATGTTGTAATTGCTCTAGCAGGCCGTGGTGTACTTGTAGGCAGCC
 ACAGAACAAATAGAGAAA CTCAGCTTATTCCCTCTGGAAACACAGACAGTACTT
 GCCATCCAACGCCAATGTTTAAGGAAGAAAGAGGAAAAAGTGATGTTGGCAAGGTCT
 CTGGAGTTGAGGCCAACCAAGGATTGGAGACCTGAAATGGATTGAGATGCCCTAA
 ATGCGCCAGTCATTACTATGAATTGGAGGACTTGTGCTTGTGAAATGTGTA
 TATGTGACGCTTTGACAACACTGAAATAGGAAAATACTATCCATGTTCGCGAGGAGC
 ACTGAATTAGAGAGGGAGACAGACTTTATGCCAGCATCAAATGAATTGATAAGCTA
 GTACAAAATGAAATTGAAATT TTTTTGGAAATAGAGTCTTACTCAGTCACCCAG
 GCTGGAGTGCACTGATACAATATTGGCTCACTGCAACCTCCACCTCTGGGTTCAAACAA
 TTCTTGCTCAGTCTCCTGAGTAGCTGGGATTACAGGTGCGTGCACCATGTCGGCT
 AATT TTTATTTAGTAGGGATGGGTTCAACCATGTTGGCCAGGCCGGTCTGAACT
 CCTGGCCTCAAGTGATCTGCCACCTTGGCTTCCAAAGTGCTGGGATTAGGCATGAG
 CTACCACACAAGCCTGAAATTGAAATGTATTGGTATAGAATACTGTTAGAATGTAT

10

20

30

40

50

GTGTATATGTATATTGTATACTCATATAAACACAAATACACATTGTATGTGTTCTG
 TAATATGTATATCTGTCTACACATACATGTATATACACACATACAATGTCTTTTTT
 TTTTTTTTTTGAGACAGGGCTTACCCCTGTTGCCAGGCTGGAGACTGCAGTGGCATA
 ATCTGGCTCACTGCAGCCTCGACCTCCTGGCTCAAGTGTACCTCCATCTCAGCCTCC
 TGAGTAGCTGGGACTGACTACAGGCACGTGGCATCAAACATTGTCCAATTTCATTTT
 TTGTAGAGTTAGGGCTTGCTCTGGCCCAGGCTGGCTCAAATTCTGGGCTCAAGCT
 GTCTGCCCTGCCCTGGCCTTCAAAGTACTAGGATTACAGATGTGAACCACTGTACCTGGC
 CTTTACAATGTCTATTAAAGATAATGGTCAAGTTTATCATCCCACTGGCCTACTC
 TAATGAAACATCTATCCATTGAAGAATTATTATGGTGGGATAACTCTGTGCCAGG
 TACCGTGCCTAGGCATTGAGTATTCAGGTTAGGAAACAGCACATGCAAAAGTGTGAA
 GTGGGAGAAGATCTCGGAGTGATTGAAGGCTAGGAGAGAGCAAGTGTGGGAGCTGTGAGG
 CTGGGAAGGTGGGAGGTTAGGTGGGAGCAGACACATAGGGATTCTTAATGTCTTAGTGT
 CATGTGGACCATGGAGAGGAGTGTAGATTGATTTAGAGCAATGCAAAATCATAGAAG
 GATGTGATCGGGGGAGTGGCATGAGCTGATCTATTAAAAATATTCCTGGCTGCTGTG
 AAGGAAGGATTGTAGGAGGCAGGAGTAGATTCAAGGGAGATGAGACAAGTGTGAGGAGG
 CTTTGAACCTGGTAAAGTAGTTGTGAAAGTCTTTTGAGGTTAGTTTTGTTAT
 TGCCTGTCACTAAACCAGAGATGCTGACCAATGAAACTCCATGAGAAAATAGTGTGTTA
 TAAAGACATATCTATGCACTGCCATTAAAAGCTGTTGGAAAAAAAGGATAAAAAGCTG
 CTTAACAACTTTTGAGATGGGTCTTACTCTGTCAACCAGGCTCACGACCTCA
 GCTCACTGCAACCTCTGCCTCCAGGCTCAAGCATTCTCCACCTCAGCCTCCGAGTGG
 CTGGGACTGCAGGCACACGCCACCATGTCAGGCTAATTGTGTGTGTGTGTGTGT
 ATGT
 GCTAATTGT
 TGTAGAGATGGGTTTGCATGTTGCCAGGCTGGCTCAAATGTTGCCAGGCTGGT
 CTCAAACTCCTGAGCTCAGGTGATCCACCCGCCCTCGGCCTCCAAAGTGTGAGATTACA
 GACGTGAGCCACTGTGCCACCTAACAACTTTAAAAAATTGACATTAGGATAT
 TTATTGCATTATTGTGAGATGGCAAATATGGAGACAATGAAATGTTCATCAGTGG
 GGGGGCTAGTTAAATGAAATACAGTGTAGCATGCTAGAACACTTTCAAGAATTAAAC
 TTTTTTGAGCCTTACTTATAATGCTTGTGCTATTGATGCCTTTTTTCAGCATG
 ACTTACTCTTTACTATAGGATTAAATTAAATTAGATTAGAAATGAGGAATATTCT
 GTAATCTGTAGAAAGTAACAAACTATAAACTATTCCCCAAGAACAAATATAATAATT
 TCTGGAGTAGCAGGTAAAGAAAGATAAAATTATATGTATACAAGAAACTGAAATTAGAC
 TTTATACATTAAAGGTACAAGTGCAGTTTATTACATGAATGTATTATCCAGCATTGA
 AGTCTGGCTTTAGTGTAAACCAGCACCTGAATAACATACATTGTACCCATTAGAATT
 TCTCATCCCTAAACCCCTCCACCCCTGAAATTAGACTTGGATCCCTAGTTAAATTCC
 ACCCCTCTTTTGTGAGACAAGGTCTACTCTGTCAACCAGGCTGGAGGGCAATGTTG
 CAATGATAGCTTAAGTGTAGCCTCAACCTCTGGCTCAAGGGATACACCCTCCTCAGCCT
 CCTGAGTAGCTGGAACCTGAGCGTGCACCAACATTAGCTAATTGTTGATT
 ATAGAGATGAGGTGGAACTCTGGCTCAAGCGATTCTCCCAAGTGTGGGTTACAC
 ACATGGGCCACTGCCCTCAGGCTAAACCTCTTCTCAGTATAGCAGCCTTGAGATGAAG
 TTCCTGAAATTACTGGCAGCTGACTGTTCCCCACATCACTGGAGGGGGGATGCAT
 AGATAAAACAAATATTCAAGCATCATTGTATTGTTCTTTGTTCATCAGCATTTTT
 TAAAACACTTGACATAAGTCCCTAGCCTCAAAGAGTAAAGCCTTGAGAATCTGCAT
 TCAGATTTCGGGTGTGATTTCCTGACAGATAGTCAGGTTGAAACTCTTTTTTC
 TTTGAGACAGAGTTCACTCTGTAGCGCAGGCTGGAGTGCAGTGGCACCATCTGCCTC
 ACTGCAACTCTGCCCTTGATTACCGCATTCTCCTGCCTCAGCCTGAGTAGCTG
 GGATTACAGGCATGCCACACCTGGTAATTGTTGATTTAGTAGAGATGGGTT
 TTCACCATGTTGCCAGGCTGGTTGAACCTGACTTCAGGTGATCTACCTGCCCTCAG
 CCTCCCAAAGTGTGAGGGATTACAGGTGTGAGGCCACCGCAGCCGGCAAAACTTGT
 TTT

10

20

30

40

50

TTTCCCTTTTGTGCTGAGAAATGTAACACTTACAGACACAAATTATGTCCTCCATT
 TTTTAAACCCACTCAACACAGGGCATGTGAATAGGCCCTGGAGCTTATTAGACA
 TTGATTGAGGCTCTTCCCCAAGTGCTGGTTGTGTGTATGTGTGTGAAGT
 CTTCTATGAGATGAGTGGTACCTACCTGGCTGTGTATCTTTTATTATTTATT
 TATTTTGAGATACGAGGTCTCACTATGTTGCTCAGGCCTGGTGAACCTGGGCTC
 AACCTATCCTCCCTCTGGCCTCTAGAGTGAGATTACAGGTGTGAGCCACTGCAC
 CTGGCCAGCGATCTTAATAAATAGATAATGGCCGGCGTGGTGGCTCACACCTATAA
 TACCAGTACTTGAGGGCGAGGCTGGCAGGTACCTGAGCTGAGGAGTTGAGACAG
 CCTGGGTAACGTGGTAAACCCCTGTCTACAGAAAATAGAAAATTAGCCAGGTGTGG
 TGGTGCATGCCTGTAGTCACAGCTACTTGGAGGTTGAGACAGGAGAATTGCTGAACCT
 GGAAGGTGGAGGTTGCACTGAGCCAGATCGTGTCTTGAAACTCCAGCCTGGGTGACAGA
 GTGAGACCTTGTCTAAAAAAAAATAGATAAGATAGGCTGGCGTGGTGGCTCACACCTGT
 AATCCCAGCACTTGGGAGGCGAGGCGGGTGGATCAGGAGGTCAAGGAGATCGAGACCAT
 CCTAGCTAACATGGTGAACCCCTGTCTACTAAAAATACAAACAATTAGCCAGGCCTGG
 TGCTGGTGCCTGTAGTCCAGCTACTGGGAGGCTGAGGCAGGAGAATGGCGTGAACCC
 GGGAGGTGGAGGTTGCACTGAGCCAGACTGTGCCACTGCCCTCCAGCCTGGCGACAGA
 GCGAGACTCTGTCTAAAAAAAAATCTATATATCTATATATCTATATAGAT
 ATAGATATAGATAATGCCAGATGATGGCTGGTTAGAAGGGATTGTCAGGGCTGGCAGGT
 TTTGCAGGTGTTAGAATGAGCAAGTGAGGAGAAGGATGCTTACTTCCCTCTCCTGTAA
 CTCTCTACCCCTCCCTCAGTGTGTGTGTCAACCTGAGTGCAGTGCAATCATAGCTCATTGA
 AGCCCAAACCTCTGGCTCAAGTGATCCTCTGCCTCAGCCTCCCAAGTAACGGGACCA
 CAGGTGCGTACAACATGCCAGTTAAGTTTCTACAGACAGGGCTTGTAA
 ATGCTGTCCAGGCTGGACTTGCACTCTGGCTCAAGTGATTCTCTTGCCCTCAGTTCCC
 AAAGTGCTGCATTATGGCATAAGCCACTGTGCCAGCCATCAGTGTCTTTTATCCT
 TTACTCCTATCAAATTCAATTCACTCAGCAGCATTGATCAAGTGCTACTATACATG
 TTGAGGACTGGAAATTATGTCTCTCATCTTACAGTGTGTCTTCTGACCTCTGTGTTAATTGT
 AATTAACGTAACTCATTCTGTATTAATTGTAATAAAACTTGTGATAAACTCAAATGAGGC
 CATACCGTTTGCCACTTCCCTCCCTCCAGGTTATGGATGTACTACATTGCAAGGTT
 TCATTGTTGGTTCAAGTTTAAACTAAGCCTATTGTGTCAGTGTCAAATTAGCTAGGTGAG
 ATGGGGAGTTCAAGCTGTGTTGTCTTTTTTTTTTTTGCCCTACTA
 ATATACAAGCGCTTATAACCTTGAGGCTGCCCTACATTAAGATTAACTTAAATTCC
 ACTGTTCTTATCTCTACTAAGTCTCAGGGTCAATGAACCTAAGTGTCTTGTGAG
 CTAGTGATAAGCAAGTTGCAAATTACAGAATTGTCAGTGATTGAATACAGTATTAAACC
 TGTAACTGGGAAGCATTTGGTAATTATGAATACTTGGAAAAAAAGCTATGGAA
 GGAAAGTTAAAATCTACGAAAGCTCAAGTAGATGGTCAAGTAACTGCTATTCAATTTC
 TAACTATATATTACTTATTATTTGAGACGGAGTTAGCTTGTGTTGCC
 AGGCTGGAGTGTAAAGCGTCACTGCTACTGCAACCTCCACCTCCGGGTTCAAGC
 TATTCTCTGCCCTCAGCCTCCGAGTAGCTGGATTATAGACATGTGCCACGCCAGG
 CTAATTGTATTAGTAGAGACGGGTTCTCCACATTGGTCAGGCTGGTCTCGAAC
 TCCCAACCTCAGCTGATCCGCTGCCCTGCCCTCCAAAGTGCTGGATTACAGGCGTGA
 GCCACCGCGTCCGGCTCTTAACATTGTTGAAATAATGTAGAGACAGCTCCAGAGCCA
 TGAAGAAGTGTATGAAGAAGCAGTGTAGCTTAAATGACATACATGTCACAATTGCTAT
 GTGAAACTATCATATTGATGAGAAGTATCTATCCTGCATAACCTCCACCAATAATA
 ATAATGTTAATAATAGTGAACAAACTATGTTATTAGTCCTACTGTCTCCAGCCTCTGT
 GCTAAATACTGGTTACTAAGTTCCCTGAAATACTATTCTCATCTGTTGTTCTAATA
 ACAGGATAGCATAATTGTAAGTTGAAATGAAATAACAGTTATGTAATAAAAGGGTA
 AAAGAGAAGACCACTACCTTATCTCTGCTGATCTGGATGGATGTAGGTGGTGT
 ACCTAGTTCACCTTGGCAGTGAAACTACTTTTTTTTTTTTTAAGA

10

20

30

40

50

GACAGGGTGGGCCAGGCAGTGGCTACGCCGTAAATCCCCGCACTTGGGAGGCTGAG
 GCGGACAGATCACTTGAGGTCAGAAGTTCGAGACCAGCCTGGCCAACATGGTAAACCCCT
 GTCTCTACTAAAAATACAGAAAATTAACTGGGTGTGGTGGTACACACCTGTAAATTCCAG
 CTACGTGGGAGGCTGAAGCAGGAGAATCGCTGAACCCGGGAGTGGAGGTTGCAGTGAGC
 TGAGATTGTGCCACTGCACCTCCAGCCTGGGTGACAGAGCAGGACTCCGTCTAAAAAAA
 AAACAACAACAAAAAGAATTTAGAAATATGAGATGACAGCAAGAATGAGGGTATT
 AAAAGAATTTTAGAACTAAATAGCAGAATGTAATGGTAAAGTGTGAGTTCTCAAG
 TCTGCTTGTGACACAGGCATGTGGCAAACATTCAGTAAGTATAGCTGTAAATTAAACAG
 CTGTAATGTATAATAGCCAACATATCACATTTCCTTTGAGACAGAGTCT
 TGCTCTGTTGCCAGGCTGGAGTGCAGTGGCACCATCTCGGCTACTGCAACCTCTGCCT
 CCTGAGTCAAGTGAATTCTGTGCCCTCAGCCTCTCAAGTAGCTGGGATTACAGGTTGTG
 CCACCAACTCGGCTATTTTGATTTAGTAGAGATGGGCTGGTCTGAACTCCC
 GCCTCAGGTGATCTGCCCTCAGCCTCCCAAAGTGCTGAGATTACAGGTTGAGCCAC
 AGGCCCTGGCCATATATTGCTTTCTTATTACAGAGCAGTTACAATTGTGAAAAA
 ATAGTCTTGTAAACATGTAAGTATGGATAAATCATCTTTTAATTGTGATTCATATA
 GGTTTGTGTTGTTGTTGTTGTTATCTGAGACAGAGTCTGGTCTGTGACC
 CAGGCTGGAGTGTAAATGGCACAACCATGGCTACTGCAGCCTCAGATGCCCTGGGTTCAAG
 CAATCCTCCCGTCTCAGCCTCTAGAGTAGATGGGACCACAGGTTGAGGCCACATGCC
 GGTAAATTACAAAATTTTTCTAGAGATGAGGCTCACTATGTTGCCAGGC
 TGGTCTAACACCTTGACCTCGCTTCAGCCTTAGAGTAGCTATGACTATAGGCATGTGC
 CATCACCCAGCTAATTAAAATTCTTTGTTGGTGGAGATGCCGTCTACT
 TTGTTACCCAGACTGCAAGTTAGTTCAAGATATCAACATTGGTGTGTTCCAATGCACGG
 GGAGGCTTGGAGCAAGTTTGCTCATATGCATAGGTTGCTTAGACATTCACTTGCA
 AATTCTTAAATGACTACAGTAGCATAAGATAGGAAAAATATCCTGTAGTAC
 ACCGATTGGGTGAGAGAGACTGTATATTAAAACAATGACCCTTGTGACATGCCACATAAA
 TTGCTGGTGGGCGAGTTGAAGAGGGCTTGTGACCTCTGCTCTTGA
 GTACGTGGAGTTGGAGTCATCCTGACAGCCTCTGTTGACACCACCGGGTCACAGATG
 TGAAACTGTGTGGATGTAGGAGAGAGCAGTGATGGGCTTACCCCAAGGTTGCTCTCCT
 TCCCTCTGCCACAAATGTTAGTAAGGAACTGCTCTGTATTAACCATTGCTAGGGCT
 GCAGATACTGGTGGTGAAGAAATAGACATGTTCTACTCAGGATGCTGAGGTGGGAGGATT
 GCTTGAGCCCAGGAGTTGGAGCTGCAGTGAGCCATGATCACACCACTGCACCC
 GGGGACAGAGCGAGACCCATCTCTAAAAACAATAAAAGAAATAGATGTGCTCTCACC
 CTCATGGAACCTGCCAGTCTAGCCTCAACCTGGTGACTGTAGAAATGTTGATTAGATGC
 TATATTGCCATGTTGAGTGTGACCCCCCTGAGAACAGCAGGGTTTTTTGAGAAGGTAGGATG
 GGGGATCTGACTGTGGGACCACAGAGGGAAAAGCACATGTAAGCTGCGTGTACCAAC
 TGGAGGAAATCGGAGACGTGATCAGAGAACAGAGCTAACCCAGGGCCATGCCGTACAGG
 GTCCTGTTAACATGTTGACTTTCTAACAGTTCTGCTAACATTTGATGTTAACATT
 TAGTTCCAATGTAAGACTCCAAGGGCTTGTGCTAACATTTGATGTTAACATT
 ATTTCACACACACATTGCTGTAATTAAAGACAGTTCTAACATGGCAAGAACCCA
 GAAATCATAATGGAAAATCTGATAAATTAAACAATGCCAACATGAACCTCTGTAGGAAA
 AAAAATACACAGACTAAAAGGGGGAAAAAAACAGAGAACAAATATTGCAACACATA
 CAGTAAGGGTAATTCTGGTTATATCAAGAGCTCTACAAATCAGTAAGAAAAAAAT
 CTAATAGGAAATGAGCAACGACAAACTGACAACATAGAAAAGGAAACACAAGTGGTCT
 GAAAACATGAAAAGTGTGCTCAGTCTCACAAAGAAATGCAAACACTAACATGGTACCAATT
 CATTAATCAGATAGACAAAGATGAAAGAGTTGGTAATGTAGTGTAGTATTGGCACAAGTG
 AGGGAAAACAGGGGATTTCACACTCTATGCCGTCACCAACAGTACCTTATTGAGGGT
 GGTTGACAATATTGTCAAATAAAAAAATTATATAGTCATTGCCACATAATGATG
 GTTCAGTTGATGATGGACGGCATACTAACATGGTGGCCATAAGAATATAATGGGCTGG
 TGCAGTGGCTCTCACCTGCAATCCCAGCACTTGGGAGGCCAGGTGGGATTGCCTG

10

20

30

40

50

AGGTCAGGAGTTGAGACCAGCCTGGCAACATGGTGAACCCCTGTCCTGCTAAAAACA
 TACAAACAATTAGCAGGCATGGTGGCGGGTGCCTGTAATCCCAGCTACTCAGGAGGCAG
 AGGCAGGAGAATCGCTGAACCCCGAAGGCCAGGTTGCAGTGAGGTGAGATTGGGCCAC
 TGCACTCCCCTAGATGACAAGGCCAAACTCCATCTCAAAAAAAAAAAAAAGAAT
 ATTATGGGCCAGCCACAGTGGCTCACACCTGTAATCCCAGTACTTGGTAGGCCAAGGC
 AGGAGAACATTTGAACTCAGGAGTTGAGACTAGTGGGACAACATAGCAAGACCCCAT
 CTCAAAAAAAAAAGATTATGGTGGAGCTGTCCTGTATAGACATACCACTTTAACCTTT
 TTTTTTGAGATGGAGTCTGCTGTGTCACCCAGGCTGATGTTAGTGGCGTGATCTGG
 GCTTAAGTAAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCTCCTGAGTA
 GCTGGGACTGCAGGCCAGGACACCATATCTGGTAATTATTTAGTAGAGATG
 GGTTTACCATGTTGGCCAGGCTGGTACTCCTGACCTCAAGTGATCCGCCCTGCC
 TCAGCCTCCCAAAGTGTGGATTACAGGCATTAGCCACCATTTACAGGCCACCTGGCCAC
 CATTTTAATCTTTATATTGTATTAAACTGTACCTTCTATGTATGGATGTGTTAG
 ATACACAAATACCATGTGTTACAGTTACAGTATTCAAGTACAGTAGCATGCTGTAC
 AGGTGTAGCCCTAGGAGCAATAAGGTTACACATAGCCAGGTGTTAGTAGGCTCTG
 CCATCTAGGTTGTGTAAGTACGCTCCATGATGTTACACAGTGACGAAATGCCCTAATG
 ATGCACTTCAGAACATATTCCTGTTAAGCAATGCAATGACCGTATTTGACAAGC
 CATTTTATTCCTAAAACCTTAAATTACAGATTTGTAAAAGTATGTAACAAATGATT
 GTAAAGGATATGTTGCTGCATTGTATAACAAAAACAGAGGATAACATAAT
 GTCCTATAAGAAGGGTAGATTATGGATGGCACATTCAACATGGGTATTATGTAGCC
 ATTGAATAAAAGGGTACTGGCTGGCGCAGTGGCTCATGCCCTATAACTCAACACTTGG
 GTGGCCAAGAAGGGATTGCTGAAAGCCAGGAGCTGGGCCAGCCTGGCAACATAG
 CAAGACCCATCTACAAGAAAATAACAAATTAGCCAGGTTGGTATTGGACACC
 TTCATGGCCCAGCTACTGAGGAGGCTGAGATTGGAGGGATCGTTGCTGGCAGGTT
 GAGGCTGAGTGGCCACTGCACTCCAGGCTGGGAGATAGACTGGGACC
 CTATCTAAAAACAAAAACAAAACCTCCTGTAACAGTCAAGAAGTCTAGA
 TGTGGGCCAGGTGGCTCACACTGTAATCCCTGCACTTGGGAGGCTGAGGCCAG
 GAGTTTGGACAGGAGCAGAGCAAGACTCCATTCTACAAAAAATAAAAAA
 TTAGTTGGCATAGTGGTGCATTCTGTAAGTCCAGCTACTCAGGAGGCTGAGGTGGGAG
 GATTGCTTGAGCCTGGGAGGTGAGGTGCACTGAGCCATGATCACACCTCTGCACTCCA
 ACCTGCGCAACAGAGTGGAGACCTGCTCTAAACAAACACAAAAACCCAGCAAAG
 TACTGATAAAAGATCTTGGCTGGCGAGTGGCTCACACTGTAATCCAACACTCAGG
 AGGCTGAGGCGGGCAGGTACAAGAGATCAAGACCATCCTGCCAACATGGTGA
 AACCCGGTCTACTAAAATACAAAATTAGCTGGCATGGTGGCGTGCACCTGTAGTC
 TCTGCTACTCGGGAGGCTGAGGAGGAGAATCACTGAAACCCAGGTGGCAGAGGTTGAG
 TGAGCCGAGATCACGCCACTGCACTCCAGGCTGGCAGAGCAAGACTCCGTCTAAAA
 AAAAAAAAGAGAGAAAGATCTCAAGTTGAGTGTGAAACATGGGTGTAACAC
 AAGAGAACATCCCTTGTTGAGTGTGAGTGTGTTACACAGGCTCAGAGGAGTAGTG
 TGTATATGCACATGAACATACGTCAGTGATATATGATATATACAAGGTTGGGTT
 TGTTGTTTTTTGAGACAGAGTCTACTGTTGCTGCCAGGCTGGGAGTGGTGCAGTGGTCA
 ATCTGACCCACTGCAACCTCACCTCCAGGTTCAAGTGATTCTGCTGCCAGCTCC
 CAAGTAGCTGAGACTACAGGCACGCACCACATGCCAGTTAATTGTTGTTAGTA
 GAGATGGGTTTCATCATGTTGCCAGGCTGGAACTCCTGGCCTCAAGTGCTCTG
 CCCGCCTGGCCTCCGAAAGTGTGTTGCCAGGCTGGAGCTAGTGGCACAATCGCAGC
 TCACTGCAACCCGACGTCCCAGGCTCAGGCAATCTTCCGCTTAGCTTCCCAAGTAAC
 TGGGACTACAGGTGTGTCATCAATGCCACCAATTGTTAATTGTTGAGAGATGG
 GGTTCCTACGTTGCCAGGCTGATCTGAACTCCTGGCTCAAGCAATCCTCCACCT
 CAGCCTCCAAAGTGTGCGATTACAGGTGTGAGGCCACCTTGCCTGCCCCGTACAAAGA
 TCTGCATAAAAGCAGTTAATAACTATGTTGAGGCTGCCATCACAGGGTGAGGTCAA

10

20

30

40

50

GGACAAGTGTGAGAAAATTCTTTAGAATCTATTTAAAAAAAGAAGAGATGACAGTGGTG
 ACAGTCAGGGAACAGATAAGCAGGTAGATTGTGGGGTCTAGGCTGCTAACTGGTGT
 AAAATGAAGCAACCGCTGAGCTGCTGTATTCAATTAAATGGAGACTAGTAAACAAACAG
 CCAGAAATTCTTCACTTCCATCAAGAGAGGCCAAAGTTATTTCCCTCAATAACCTG
 GGACTGTAGGATTAAGTTTTTTTTAAATACTACAATATGACTACCAGT
 ATAATTAAAAATGATTAGAATCTATTGAGTAAGAAATAGGTGTCGCTGAAGTAGA
 CAGTCACTGAAGTCAACTAGTGCAGGAAAGACAGAAAAAAATTGAAAGTAGGAAACAATC
 AGCAGATATGATAACCAACATGAGCTGTCAGTGATAATGGATTAAGTCCTCAATAATGG
 CTGAGCCAGATGGAATTAAAGAAAAATCCAGGCCGGCATGGTGGCTCACACCTGTAA
 TCCCAGCACTTGGGAGGCTGAGGTGGGAGGATCACTTGAGTCCAGGAGTTGAGACCAG
 CCTGAACACATACTGGGACCCATCTATTTATAAAATATTGAAAAAGAAAA
 AAAATTCACTGTTCTGCTTAAAAAGACAATGGCACAGAATGTCAAAGAATAAAT
 AAAACAAACATGGCAGAAAGAGATTCAAGGTGCTACCAATATCAGGCTAAGTAGCATTCAA
 GATAAAAGATTATTAATAAAAGTTAGTTAACTAGAGTAATTGCATATTAATGAAACA
 TAATCTATGGTAGAGATAATTAGTCATAAAATTGTTTATGTATTCAAGGTACAAAC
 AAGCAAACAAGCTTAAATAGTTTAAATGCTTATATGCTTATAGTTCTTATGTGCA
 TTAATTCACTAATTCTCATTCTGAGGTAACACTATTATTATCACATTACAGA
 TGTAACCGAAGCAGAGAGATTAAATTAGCTTGCCCAGGAGATGTGGCATTCTGGGATT
 TGAGACAGTGGTTGGCTGTAGGTTGCTCAATAACCAAGAGATGCTCAAATCAGAT
 TTTTAAATATGTTTCTGAGCATTGAGGTTACATGCTTACATGGGTTAG
 TCTTTGGGTTGAAAAACATGAGTAAGTGCTAGAAGAGCAGGATATGCATCCAGATTAA
 TAGTATGCTGTTCTGAGCCTGGCATTTGCTTATAATAGAAATGAAGGCT
 TTTTTTTTTGGCTGAGAATAGCACTGAACACTGGGAGGACTGTGGGTTGTAAG
 TTGTCGCCTCTGAATGGAGTTGAATTAAAGTTCTGGTTCCAAGAATGATTGATT
 AAAGACCTCAAATTGCAAGTTAGAAACTGACTTCAGTCCTGAGGTTTTTACCAATTAA
 TGAATAATTAAATTATGGTAATAAAATGGTAATAATGGTAAAGGCT
 CCATTAAATGAATTCTTAAAAAGCAATTGAATTGTTGATGAAAGGTGATGTTAAAT
 TATCCCAGATTTCATCAATCTTTTATTGCCCCCTGGATTGAGTCATAGAAAGCCTT
 TCCTTATTCTAAGGTTAACAGACATTCAACCATGTTCTCTAGTATTGCTTGCAT
 ATCTTACGTTATTATTTATTGAGACAGGGCTCACTGTGTCA
 CTCAGGCTGGAGTGCAGTGAATGATCTGGCTCACTGCAGCCTCTGCCTCCGCCCTCC
 GGGTTCAAGCAGATTCTGCTGCCTCGGCCTCCAAAGTAGCTGGGATTACAGGCACCTGCCA
 CGCGCCTGGCTAATTGTTGTTAGTACAGATGGGTTGCTGGCA
 GGCTGGTCTCGAACCTGACCTTAAGTGAATGCCACCCGCCCTGGCCTCCAAAGTGTGG
 GATTACAGGCATGAGGCCACCGTGCCGCCCTAAATTATTCTGATATGTGATATGATGT
 ATGGTTCTAACTACTTGTACGGTGCATTATTTCTAAATGTGGTATTGGATTCTTAA
 TATTTGTTAGAAGTTCTGCATCAATTGAGTACCTGGTCTGTTGTTGTTG
 TTGTCGCATCTTATTGGTATAGGTATCAGTGTATATTAGTTGAAAGGAAGTTGG
 AAGTTTCTTCTTTAGTACTCAGGAATGATTTAAGAATTGAGACTATTGGTCT
 TGAAGGTTGGTAGAAGTCCATGGGAATCCATCTGGGCTGGTATTCTGTGCGGTA
 GTTCTTAATTGTTCCCTATTTCTTATTGTTAACAGGTGCTCTGAACCAGAA
 TAGGTTAGAGAGGCCCTCTATTGTTAACACAAGTGGTCTGCCTAAAGTTCT
 ACTCTAATGGTTAATTGTTAGTACTGCATTCCCTGAAATTACACGTTGTTCTAG
 GTTTCTGACTTATTCACAACTTTAGTCTTCCCCCTGGAAATCATGCCCTTCCA
 TAAACAGGACTCTGATGACCTGAAGTATTTCACACTCGGGTGGACTTCTGTTCTG
 GGGGTGGTTAGAGCAATTAGGCCTGCCACTAGCTACCCCTGTTCTACACCATGCT
 GTTTCTCAGAATGCTCTTGTACAGGTTACAGGTAATCTGGAAGTTGGCATT
 TCACTGACGTTGGTATATTCTTTTGCTTACAGGTAATCTGGAAGTTGGCATT
 TCTCTTAAGTGTGAGGGTGTGGTTTATGTCATTGTTATTGTTCTGTTG

10

20

30

40

50

GTGTTCTTAGAGACAGGGTCCCCTCTGCCCTGGCTGGAGTCAGTGGCGTCTGATC
 ATAGCTTACTGCATCCTCAAGCTGCTGGGCTAGATGAACCTCCCACCTCAGCCTCTGA
 GTAGCTGGACTACAGGAGCACACCACCATACCTAATTTTTTTTGAGACGAAGTC
 TTGCTCTGTCCCCAGATTGGAGTGTAGTGGTCAATCTGGCTCACTGCAACCTCTGCC
 TCCCAGGGTCAAGCGATTCTCACCTCAGCCTCCGAGTAGCTGAGACTGCAGGGTGCAT
 GCCACCATACCCGGCTAATTGGTATTAGTAGAAACAGGGTTCACCATGTTGGC
 TAGGCTGGTCTCAAACCTTGACCTCAAGTGAACCTGGCCTCCAAAAGTGCT
 GGGATTACAGGCTTGAGCCACTGTGCCTGGCTGGCTAATTGGTATTAGA
 GATGGGATCTGCTATGTTGCCAGGCTGGCTTGAAACACCTGGCCTTAAGCAATCCTCC
 CACCCCTAGCCTGCCAAAACACTGGGATTACAGGCATGAACCATGTTGCCTGGCTGTT
 TGTTTTAATTCTATGTTGTTTGAGGATGTATGGGAGAGATGGATTAGGAATCA
 TCGTTGTCCTGGCTACCTGAAAGTCCAGGCACTCTCTAGATACTTTATAAATATTAAAC
 TCATTTTATCCTCTCAACAACACTATGACATGGTACTGTTACACCTCCATTAGG
 ACTTAACAGAGAGGTTAAATATGTAGGCCAGGGTACAGAGAGCTGGCCTAGACCAAG
 ACAATCTGGCACCAAGACTATGTTGCTACCCCTAAGGCTTGCCACATGTTAGTGA
 TTCTCAGCCTGTCAATTGGGAGGGATTGCCCTTTTTAACTTTAAAAAATTAA
 TTCTTATTATTATTTGAGACAGAGCTCCCTCTTGCGCAGGCTGGAGTGGAG
 TGGTGTGATTCAGCTCACTGTAACCTGCTCTGGGTTCAAGTGAATTCTCATGCCTC
 AGCCTCCAAGTAGCTGGATTACAGTTGCCAGGCCACATGCCAGCTAATTGGTATT
 ATTATTATTATTGAGACGGAGTCTCGCTTTGTTCAAGGCTGGAGTGCAGTGCTG
 TGATCTGGCTCTGTAACCTCGTCTCCCTGGTTCAAGGTGATTCTCCTGCCCTCAGCCT
 CCGGAGTAGCTGGACTATAGCGCGCACCACTTGGCTAATTGGTATT
 GTAGAGACGGGTTCACTATGTTGCCAGGCTGGCTCGAACTCTGACCTCAGGTGAT
 CTACCTGCCCTGGCTTCAAAGTGCTGGATTACAGGTTGAGCCACCATGCATGGCTG
 GATTGTCCTTTTAAAAAAACAAAACAAAAAAACCCAAACCATAAACCA
 ATATTCTGAAAGATTGGTCTCCACACCTGTTATATAATAATTAGTTTCCATT
 TTCCTCTGGTAAAGGCACATATGCCACTCAGTTCCAGTTGCCACACCAATTACAT
 AATTGTTTGCAAGCCAAAAGCAAAAGAGAGTTGACATTAAATTAGCTTATGAGTGA
 CAAATTGAGGCCTAATGTAAGAGTTTCAATTACCTTTGAAAAACTATAATAGCTAG
 AAGCCAGTTGTCATTACTTTGATTCCCTAGAATTCTGGCATCTTCATCTGGAACCA
 CAGATGAAAGAGCTGCAAGGAAGGATTTTTCTTAACGGAATAGTTAACCATCTG
 AATGCAAAAGTATTGGATGCTAGAATAATAGGTACACATAATTGAGGTTGACGTTTC
 CCGGGTAAATTCTATTCTGTCATTTCCCTTTTTGAGACGGAATCTGCTCTG
 TCGCCCAGGCTGGAGTCAGTGGCATGATCTCGCTCACTGCAAGCTCCACCTCTGGT
 TCATGCCATTTCCTGCCCTCAGCCTCCGAGTAGCTGGATTACAGGGCCTGCCACAC
 ACCCAGCTAATTGGTATTGAGATTACAGATGTTGAGGTTGACGTTTC
 CCTGACCTCGTGAATCCGCTGCCCTGGCTCCAAAGTGCCTGGATTACAGGCGTGG
 ACTGTCCTGGCTTTTTTTTTTTTTAAAGACAGAGTCTCGTTG
 TTGCTTAGGCTGGAGCGCAGTGGCATGATCTCAGCTATTGCAACCTCCGCTCCGGT
 TCAAGTGATTCTCTGCCCTCAGCCTCCGAGTAGCTGAGATTACAGATGTTGCCACCAT
 GCCTGGCTAATTGGTATTGAGATTACAGATGAGGTTGCCATGTTGCCAGGCTGG
 CTCAAACTCCTGACCTCAGGTAATCCTCCTGCCCTAGCTCTCCAAAGTGCCTGG
 TAGGCATGAGTCACCGGGCCAGACTCAATTCTGACAAGCTCTCAGAGAGAGTAAAAA
 GCAAATGAATATTCAATTGGTATTGAGCTGAGCTTACGATTGGTCTTTCTTT
 TTTTTTTGGAGATGGAGTTGCGTTGTTGCCAGGCTAGAGTGCAGTGGGGCAT
 CTTGGCTCACCGCACCCCTCCGCTCCGGTTCAAGCGATTCTGCTCAGCCTCTG
 AGTAACGGATTACAGGCATGCCACCATGCCGGCTGATTGTTAGTAGGG
 ACAGGGTTCTCCATGTTGGTCAGGCTGGCTTAAGCTCCGACCTCAGGTGATCCACCT
 GCCTGCCCTCCAAAGTGCCTGGATTACAAGCATGAGCCACCTGCCAGCCTTTT

10

20

30

40

50

TTTAAATCTGAGAAGAGGTCTGGCTCGATTGCCCTAGGCTGGAGTGCAGTGGTGCATCTC
 TGCTCACTGCATTCTGCCTCCAGACTCAAGCAATCTCCCACCTTAGCCTCTGAGT
 AGCTGGACTACAGGCATATGCCACCACACCTGGCTAATGTTGTATTTTTGAGA
 CAGGGTTTGCCATTGCCCCAGGCTGGTCTGAACCTCTGACCTCAGGTGATCTCCC
 CCTTGGCTCCCAAAGTGTGGATTACAGGTGTGAGGCCACTGTGCCTGGCTCCTCAC
 TGTTGTAAGATACTTGAATTGGGCAATATTGAGAAGTCTCTAAAAGTTCACTTG
 ATTGTCAGTACTAGAACACTACATTTAATTTGACATATTCTGGAGCATTCAAGAGCA
 TTCTATTAGCTAGAAAGGTCCAGGATAATTGACTTTAGAAGTTACTGTTACCATGAAT
 CTCAATGACTTTGAAATCCATGAAGAATATCTTTTTTTGAGACGGAGTCTCA
 CTCTGTCGCCAGGCTGGAGTGCAGTGGTGTACTGGCTCACTGCAAGCTCGCTACTG
 GTTCACGCCATTCTCTGCCTCAGCCTCCGAGTAGCTGGGATTACAGGCACATGCCAC
 CACGCCCTGGCTAATTTTGACATTTTAGTAGAGAGGGGTTCACTGTTAGCCAGG
 ATGGTCTCGATCTCTGACCTTGTGATCGCCCGCCTGGCTCCCAAAGTGTGGGATT
 ACAGCGTGAGCCACCGCGCCTGCCAAGAATATCTTTTGCTGGTAACTAGAGAGGACT
 CCTCTGAAGCAGATGCCATTGATGGATTTCATCATTATGGGTTAAAAAACATT
 TATTTGAAATAATTCAAATTAAATAAGAGTTGCAAAATAGTACAATAATTCTGTT
 AACTTTCATCCAGATTACAAGTCACCTTACAGGTTGAGTATCCCTATCCAAAT
 GCTGGGACAGAAGTGTGTTGGATTTCAGATTTTCAAGATTTGAAATTTTATT
 TATACTTAAGCATCTAATCCCCAATCTCAAATCTGAAATATCTGAAATGCTATGATG
 AGCATTCTCTTGAGTGTATGTGGCACTTTAAATTATTAATTAAATTATTT
 GAGATGGAGTATTGCTCCATACCCAGGCTGGAGTGCAGTGAGCGATCTGGCTTATTGC
 AAACCTCACCTCTGGGTCAGTGATTCTCTGCCTCAGCCCCCTGAGTAGTTGGACT
 ATAGGCGCTTGCACACGGGGCTAATTTGTATTTAGTAGAGACAGGGTTTCAC
 CGTGTGGCCAGGCTGGTCTGAACCTCTGACCTCAGGGTGTCCACCTGCCTCCGCC
 CAAAGTGTGGATTACAGGAGTGAACCAACCGCCTGGCATGGATTTCAGCATT
 AGATTGGGATACTAACCTGTACCATGTTACTCTCTCCCTCTCTCTCTCT
 TATATATATATATATATATATATATATATATATATATAAATTATATAC
 ACTACACATATATGTATGTATGTATTTATATATAAAACATATCTACATATA
 AAATACACATGTATATATACATGTGTCATATATGTGTCCTATATTAAAGTTGTTGG
 AACCACTTGAGGGTAAGTGTGAGACATGGCTCATTGCTCCAAACTTCAGTGTGT
 ATTCTTAAATACAAGGACACTGGTACATAACCACAGTATATCACCAATGTATATTA
 TAACAAGACTACCACAAATCTTATCTCTCAAATTGTTAGTAATATCCTTATA
 GCAAAAGACAAAACAACAAAAACTGTTCTTTTATTTGTTGGTCCATT
 TATGTCAGGTTATGCATTAATGTCATTGTTACTTGCTAAGTCTTACTGGCCTTA
 ATTAGGATATTCTTGCACTCCGCCAAACTCCTCTCATGGTGTATCTTTTT
 TTTGGAGATGGAATTGGCTTATGTTGCCAGGCTGGAGTATAATGATGCGATCTGGC
 TCACTGCAACCTCGCTCCGGTCAAGCGATTCTCTGCCTCAGCCTCCGAGTAAC
 TGGGATTGCAAGGCTCGGCCACCTGGCCAGCTAATTGGAAATTGTTGAGACAGGG
 TTGCCATGTTGGTCAGACTAGTCTGAACCTCTGACCTCATGATCGCCGCC
 CCCAAACTGTTGGGATTACAGGTGTGAGCCACTGTGCCGGTCTTTTT
 GAGACAGGGCTTATCTGTTGCCCTGGCTGGAGTGCAGTGTATGATCTGGCTCACTG
 CAACCTGGACCTCTGGCTCAGCGATCTCCCACCTCAGCCTCTAGTAGCTGGGAC
 TATAGGCACACACCACATGCATGGCTAATTGTTATTTGTTGAGAGACTGGGTT
 GCCATGTTGCCAAGCTGGCTTGAACCTCTGGCTCAAGTGTGATCCACCTGCCTGGC
 CCCAAATGCTAGGATTACAGGTGTAGGCCACTGCGCCTGGCCCTAATTGTCATT
 TGTAGAGATGGGTTCACTATATTGCCAGGCTGGTCTTGAACCTGGCTCAAGTGA
 TCTTCCCATCACAGCCCCCTAAAGTGTGGATTATAGCGTGAACCACTGTGCCTGG
 GAGGATTAAGTTCAACCTCAGGGAGCGGCATTCAAACATATAGCATTGTCCTT
 TAGTGA
 CTGGCTAGTTCACTTAAAGTGTCTATTGATCCATCTATAGACACTGTTTCTT

10

20

30

40

50

ACCTTTGGCTTGCAAATAATGCTGCTGTGAATATGAGTTATAGAAAAATACCAATTG
 AATCCGTGTTCAATTACTTGTAGTATACCTGGAAGTGGAAATTCTGGATCATATGG
 TACTTCCAAGTTTTTTCTTTGAGACAAGGTCTCACTCTGTCACCCAGGCTG
 GAGTGTAGTGGCACGATCTGGCTCACTGCAACCTCCGCCCTCCGGGTTCAAGCGATTCT
 CCTGCCTCAGCCTCAAGTAGCTGGGATTACAGGCACGCCACCACGCCAACTAATT
 TTGTATTTAGTAGAGATGGGTTCTCATGGTCAAGGTGCTCCCAGCTCCGAC
 CTCAGGTGATCTGCCCTCAGCCTCCAAAATTCTGGGATTACAGGTGAGGCCACCG
 CACCTGCCCTCATGTTCAATTAAACAAACAATTAGTTAAAAAAATAGGAAACTAA
 GAGAATGAACTATTCCGTGTTATTCACTGGGTTATACTGTTACTATCATTGTTAT
 TTTGAGGTACAAATTGCCCTACTGGCCAGCAGGAGTCAGTTGTCTGTG
 TCCTTTCTAGCTCTTGTGAACTCTTACTGGCCACAATAGGATGTTCAAGTTCA
 TCTTCTTACTTTACTGCCAACGCTGGGATCAGCCATTCTCAAGGAGGCCAGTTCC
 TTTCATGGAGAATGGAAAACCAATATGTAGAAACCAAGATAGAGGTGTTAGGTGAT
 TGCTACTGGAGTGTCACTGCTTCAACCCCTTCAGAAGAGACCTAGGAAATGTGTTG
 GTGTGTTATATATATGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG
 ACCCGAAGCATGTATTCTGTATTATTATTATTTGAGATGGAGTGGAGTCTGCTGTG
 GCCCAGGCTGGAGTACAGTGGCACGATCATGGCTCACTGCAACCTCTGCCCTGGATT
 AAGCAATTCTCCTGCTCAGCCTCTGAGTAGCTGGGATTACAGGTGTCACCACACGC
 CCACCTAATTTGTTGAGAGATGGGTTTCAACATTGCCAGGATGGTCT
 TGAACCTCTGACGTCAGTGTACTGCCCGCCTCGGCCTCCAAAGTGTGTTGATTAGG
 CGTGAGGCCACTGTCATCCAGAAGCATAACATATCTATTCTATACATTCTGTC
 TTTACATGTATATATTTACAGTTGCACTAACCTCCAATTACAATCTAACAT
 CATGGGATTATTCTGGCTTCTCCCTCTCATATTGTTGTCCTCCCAACAGTGAGAAC
 CTGGCTGCTACCTCAACATGGTAACCTATTATTAAAGAAACTTATTCTTTTT
 TTTTTCTGAGATTGAGTTGCTCTTGTGCCCCAGCTGGAGTGCAGTGGTGTGATCTT
 GGCTCACCGCAACCTCTGCCCTGGGTTCAAGCGATTCTCCTGCCCTCAGCTCTCAAGT
 AGCTGGGATTACAGGCATGCCACCATGCCAGCTAATTGTTGTTGAGATG
 GGGTTCTCCATGTTGGTCAGGCTGCTGGAAACTCCGACCCAGCTGATCTGCCCTG
 CGGCCTCCAAAGTCTGGGATTACAGGCGTAGGCCACCGTGCCTCTAGTTATT
 TATTTTATTCCATGTCCTGAGCTGTCAGTCTGCGAGCACGTTGTTGTTCTGGGCTGGCC
 CCTCAGTGCAGTCTTAACCCCTAGCCCCAGTCCTCTGATCATATCCCAGACACC
 CCTACTGAATCCCAGGTCTCTACCAAGGGAAAGGCAGGGAGGAGGATTGACCAAGGAGA
 AGAGGGGAAGGGACAGGGAAAGGTCTGTTGTTGAAAGTAGTGTGACAGTGTGTTTTTAT
 ATAATGCGTCTTAGCTGTTGTTGTTGAAAGTAGTGTGACAGTGTGTTGTTTTTAT
 TACTTACTTTGTTCTTTAAGATGGTTCACCCAAATATCACTGGTGTGGAGGCAG
AAAACCTACTGTTGACAAGAGGAGTTGATGGCAGTTTGTGCAAGGCCACTAGTAAAGTA
ACCCCTGGAGACTTCACACTTCCGTTAGTAAGTTGGAATGAAAAGAGAGGATCTGAGA
 GTGTTTCTAGGTAGGAAGTGGAAAACCAGTCTGGATAGCTGCTGCCCTGCAATTG
 GTTGAAGGCCCTATCTGAGCCCTGGGCTGCCCTCAGGGTTGGGAGTGGCCTCTGGA
 CATTAGCAGAAGAGGAGTAAGGAGGGCCCTCTCTCCCTGAGACCTCATGGAAGGT
 GAGTTGGAGCAGGTATAGAAGTTCTAACGCCCTCAGTGTGAGACTGTTCCACACA
 TCTTGAACCTGGTTCTGCATTCTTCTTCCCTGTTGATTATTAAAATTTAT
 TTCTTTCAATTTTTTTAAATAGAGGTGGGATCTTCAATGTTGCCAGGTT
 GGCTTGAACTCTGCCCTCAAGCAATCCTGCCCTGCCCTCCAAAGTGTAGGATTACA
 GGCCTGAGCCACTATGCCCTGGCTCTTTTGAGACAAGCTGTTGCTCTGTTGCCAG
 GCTGGAGTGCAGTGGTACGATCACAGCTAACAGCAGCCTGAACCTCTGGCTTAAGTGA
 TCCTCCGCCCTCAGCCTCCCGGGTAGCTGGACTCCAGGCTTGTGCCACCATGCTCAGCA
 TTTTAAAAAAATATTTTGAGAGATGAGGTCTCACTGTATTACCAAGGCTGATCTTA
 ACTCTTAGCCTCAAGTGTACTGCCCTGCCCTCAGCCTCCAAAGTGTGGGATTACAGGCAT

10

20

30

40

50

GAGCCACCACTCAGACTTGTGACTTCTAATAAGAAAAACTTGTAAAGAGTTCTTCAGATCACTTCCCTTATACAACAGTAAAACATGACTGAGGAAGTGTGGTCCCTTGCTTCCCCTGCCAGGGCCGTTCCCTCTTCCCCAGAGGAACACCAAGAGGTTGCATATCTTCTGAACGTGTTTATAGTTGACTGCACTTGTACTGTGTATGAACAATATAAAGTTGGTTGTGTTAAAAAATTACACATACATGGATTATAATGTATGTATCATTGGAACTTAAAAATTTTTGAGCTTCAAGCTGCTGATTGATAACGATCTATTTTTTTTGAGATGGAGTTCACTGTTATTGCCAGGCTGAAGTGCATGCCAGGCTGATCTCAGCTCACTGCAACCTCAGCCTCTGGGTCAGCTATTCTCTGTCTCAGCCTCCGGAGTGGCTGGGATTACAGGTGATGCCACCATGCCAGCTAATTGGTATTAGTAGAGATGGGGTTCAACATGCCAGGTCTCAAACCTCTGACCTCAGGTGATCTGCCTGCCTTGGCTCCAAAGTCCTGGAATTACAGGCATGAGCTACCATGCCCTGGGCTTTTTTTTTGAGACAAAGTCCTGCTCTTCCCAGGCTGGAGTGCAGTGGCCACATCTGGCTCACTGCCTCTGAGTCAGCTCAAGCAGTCTCCTGCCTCAGCCTCGAGTAGCTGGGATTACAGACATGTACCAACCATGCCAGTTAATTGGTATTAGAGACTAGGTTTACCATGTCGGCAGGCTGGCTGAACCTCTGACTTAAAGTGATCCATCTGCTGGCTCCAAAGTCCTGGGTTACAGGCATGAGCTATCGGCCTGGCCTGAGAACATCCTCATTCTTAACCTCTACTCCCTGCACACTATCTCATTCTGTAGGTAGCCATTCTATTCTGTTACCCCTCTGTGTTCTTCATTCTTTCTTTCTTTCTTTGAGACAATCTGCTCTGCCCAGACTGGAGTGCAGTGGTGTATCTGGCTACCGCAACCTCACCTCTGGGTCAGTGAATTTCATGACTCAGCCACCTAAGTAGTTGGGATTAACAGGCCTGGTGTACACTACCACACCCAGCTAATTGTGTTAGTAGAGATGGGTTTCACCATGTTGTCAGGCTAATCTCAAACCTTGGCCTCAAGGGATCTGCCTGCTCAGCCTCCAAAGTGCTGGGATTAGGCATGAGGCCACCATGCCCTGGCCCTATGTTCTTTTATAAAAATAAGCAAATTAAATTTTATTACTATTTCTTTTATTTTACACATCAAGTAGAACATTAATTTCTCTGTAATTTCAGTTACCTAAATCTTTAGTGATCTCTCATCTTTTAATCAGCTGGATCGCATTCTATCATGTGAATATTATAACTTCTATACTGTACCAGCAGGTAGCAGTTAGTGTCTAATATTAAAATGATATATAATGCTCAATGAAATATAGTAACCTTGCATATATTGTTGTCTTGGGATAACACTACCTCTCATGTTGGAAACTGTGTCATTACATGTGTTAAAATTACATGTGTCCTTTTATTGTTATTTTGTGAGTGGAGTTCACTCTGTTGCCAGGCTGGAGTGCAGTGGTGAGATCTGCCGACTGCAACTTCCGCCTCCGGGTCAGCGATTCTCCTGCCTCAGCCTCCCGTAGGTGAGATTACAGGTGCCACACGCCAGCTAATTGTTGTTAGTAACGGACGGGTTTCACCATGTTGCCAGGCTGGTATCGGCTGACCTCAGGTGATCCTCCCACCTCAGCCTCCAAAGTGCTGGGATTACAGACGTGAGGCCACCATGCCCTGGCATCACTTTTTTTCTTAATTGCTGCATAGTGGCGGGCACAGTGGCTACGCCGTGAATCCTGGCACTTGGGAGGCAAGGCAAGGCAGGGCAGGATCATGAGGTCAAGGAGACCAATACCACCTGGCTAACATGGTAAACACCCGCTCTACTAAAATACAAAAAAATTAGCTGGGCTGAGGAGAATGGTGTGAACCGGGACGTGGCAGCTGAGCCAAGTGCACCAACTGCACTCCAGCCTGGGTGATGGAGTGAGACTCTGTCAAAAACAAACAAACAAAAAAATTGCTGCATAGTATTCCATTGTTAGTAGTAACTGTTAGTAGTAAACACAATTTTATAATGCATAGTATTCCATTGTTAGTAACTGTTAGTAGTAACTGTTAGCACTATTGTTTATACATTTTATGATTAAAAACAAAATGTTTCTATTATGATAAAAGTGGCAATGAATATTGTTGACAAGTGTGTTGGTAGCTATAAGTTATTGTCATTATGCAATTGCTGATAGGCCAGTCATTCAAATAGAAGATATAACAGGTAGGCCGGGCGTGGCTCACGCCGTGAATCTCAGCAGCTTGGGAGGCCAGGCTGGGTGGATCACCTGTGTTAGGAGTTCAGACCAGCCTGACCAACATGGAGAAACCTCATCTACTAAAATAAAAGTAGCTGAGCGTGGTGGCGATTCTGTAATCCAGCTTGGGAGGCTGAGGTAGGAGAATCACTGAAACCTGGATTATAATGTATAATGTAATCCACCGCGAAGGTTGCAGGTAAACCGAGATCACGTACTGCACCCAGCCTGGGCAATAAGAGCGAAACTCCATCTCAAA

10

20

30

40

50

AAAAAAAAAAAAGATATGCAAGGTAAGATACTAATAAAGACCTTGTGTTGAGTTGGTT
 GACATGTGGTTATTCACCCATCGTATTCTTATAGGGAATAGGTAATTGCTCCTTGG
 GTTCTTCAACACTTAGGTAAAATCCGACGTGGAAGATGAGATCTGATTTACTGGTGT
 AACTCTTATTTGTCCCCCTGCCCTCCCTTCAATGGACTATTTAGAAGAAATGGAGCT
GTCACCCACATCAAGAATCAGAACACTGGTGATTACTATGACCTGTATGGAGGGAGAAA
TTTGCACCTTGGCTGAGTTGGCCAGTATTACATGGAACATCACGGCAATTAAAAGAG
AAGAATGGAGATGTCATTGAGCTTAAATACCTCTGAACGTGCAAGATCCTACCTCTGAA
AGGTCACTTAAAGAATTACTAAAGTGGAGACTAATAGCATCAAATTATTTAGAGTACC
 AATGCATGACGGTCTGTGTACTCTGACTCCAAAGGCTTGTGACTGTTTGAGC
 TGTAATCTTAAAGAATTACTAAAGTGGAGACTAATAGCATCAAATTATTTAGAGTACC
 TTTTCTGCAAAAGTTTAATCAGTGTACTTACACTCATCCTATAGGGGTTGCATACC
 ATTCCTGCATAACTGGTACGTGTATTAGTTAAGACTTATTGAACTTCAGCAGATAAA
 TCTTGAGAGTTATTAGAGGAAACAAATGATAATGGAGACACCAAAATAGCAGCAGTT
 TCTATGGTGGCTCTGACCAGTTATTCAAGCAATGTCACCAACAGATGTCAGTTAACGCTC
 AGAAGTGGAAAAGCAGAGAGCTCAGAGGGTCACTTTCATCAGTTAACGTTATGTTAT
 CACCACAATTATGTGAGAATGACCTTGCTTAGAGAAAATTATGTTATTTGAGATCTTT
 CCCCCTGTGTTGAACTAGGCTGATGAAAGCATGGCTTGACTTATTGATTGATTGATT
 CGTTTGACATTCCAATCTCCTCTGACTTGGTCAAAATTCAAGGATCTCTAGTTAG
 TTTGATATTTGTCAGTATGTTTACAGTTACACCTTATGTCAGTGCTA
 TTATGTGCTGATAATTGTTCTCTAGCTACCACCGTAGCTTCAGGCAAAGGCTGTCAG
 CCAACTCTGTACAGTTATTCTAAATTACTGTTTCAAGTGGATGAGTATGGATGAAAT
 AACTCAAAGTTATTCTTTGATGATGAGCCTTAACACCCACCTGCATGATAGTACTG
 CTTCCTGACCAAGATCCTGAGGAAAAGCCTTACATTATTAGAACTATGTTAAGATGC
 TTCCCAAAAACATGGAGCAGTATTGTCAAAGTCTGCTTGGATGGCTTGGATGCC
 TACATCAGGACTGTCGATGTGCTGGTAAATGCAAGATTCTGGGCTCATTCAAGACTT
 ACATGTATTGATATTGCTGGTTGAGCCTGGAAATTCAATTAGCTGAAATCCCTC
 ATTGTTACTCCAAGTCTTATGTCATTACAGTTGAGATGATCACCCAGGATATAGTC
 CAAAGACACTGGAGGCTGTTGAAGTATAGGTTGTTAGGAAAGGTTGAAATGTTG
 AATTAAATTATAATGAAGATCCTTTTAATTGAGTGTTCACATGCCAAGGCAAGGACAAA
 CATTCAAATGATTCTGTCTGTTACAACCTTCTTTTCAATTGAGTGTTCACATGCCAAGGCAAGGACAAA
 TTTGAGATGGAGTCTCACTCTGTCACCCAGGCTGGAGTCAAGTGACCGATCTGGCTCA
 CTACAAACCTCCGCCCTCCAGATTCAAGTAATTCTCTGCTCAGCCTCCGAGTAGCTGG
 GACTACAGGCATGTCACCATGCCAGTCAACTTGTGTTACAACCTTCTTTCTTTTAA
 CTCTCAAAGTGTGCTGGATTATAGGCGTGAGGCCACCGCTGCTGTCTATTACAACCTT
 ATTACAACCTCTTATTGACTTTACAAATTATTATTATTGAGAT
 GGAGTTGCTCGTCACCCAGGCTGGAGTGAATGGTGCAGTCTCAGCTCACTGCAACCT
 CCGCCTCCAGGTTCAAGTGATTCTCTGCTCAGCCTCTGAGTAGCTGGGATTACAGG
 CACTGCCACACCCGGCCAATTGTTGTTAGCAGAGACAGGGTTCACCATGTT
 GGTCAAGGCTGGTCTGAATTCTGACCTCAGGTGATCCACCTGCCTCGGGCTCCAAAGT
 GTTGGGATTACAGGCATGAGCCACCACGTCCGGCGACTTTATTGAGAC
 AGGGTCTGCTCTGTCACCCAAGCTGGAGTGCAGTGCATGATAGCGCACTGCAGCC
 TCGACCTCCTGGACTCAAGTGATCCTCCTGCCCTGGCCTTGTGTTAGCTGGGATTACAG
 GCAGTTGCCACCATGCCAGGCTAATTGTTATTGTTGAGATGGGATTCACTGT
 GTTGGCCAGACTGGTCTTGAACCTCTGGCCTCAAGTGATCTCCTGCCCTGGCCTCCAA
 AGTGTGGGATTACAGGCATAAGCCACTATGCATGCCGTGAACTCTTAAATGGCTAT
 AATTAAACAGTGGCCTTTAAGATTGGCAATGGACAATGGCAAATTGCAATTAA
 AAGAGGAGGGATTAAAAAAACAGGAAAGATTGGGCAATTGTCTCTAAAGGACTGTG
 GACTCATTAAAGAAGTTAGTGGTCACTTACCATCTTGTGGTTTCTGCCTGCAT

10

20

30

40

50

GGGATGCAGATTTCTGTCAGGTGGATTGATCAATCCCTGGAGGAATGTGTCTACT
 TTTTAATTGTGTTAGGAGAGCTGACTGTATACTACAGTAGTTTGTGAAAGAACACATGAA
 CCCATAGTAGAGCTAAATTCTTTTATTTTAAACCTT**GTTGGTTCATGGACAT**
CTCTCTGGAAAGAACAGAGAAATTATTA**ACTGAAAAGGAAAATGGTAGTTCTT**
GTACGAGAGAGGCCAGGCCACCTGGAGATTTGTTCTGTGCGCACTGGTGATGAC
AAAGGGAGAGCAATGACGGCAAGTCTAAAGTGACCCATGTTATGATTGCTGTCA**GGTA**
 AATCTCAGTTGAAAAATGGGCTGGCAAGATGTTACCTTGGGTGATTTTCTGCTGAC
 AGAAGACAGACACCATTACATTCAAAGTCAGATTGCTTTATTTATTTATTTATT
 TATTTATTTGAGACAGGGCTTGCTCTATCACCTACAGATGGGTTCAACCACGGT
 CTGGTGACCCAATCTTGGGTGATTTTCTGCTGGAAGAGGACAACACCATTACATT
 AAAGTCAGATTTCGTTTTTTTTTTTTGTTTTTTAATATTCAATTGTT
 TATTCAATTGAGACTGGGCTTGCTCTGTCACGCAGGCTGGAGTGCAACCTCCCTGG
 CAGTTGATCTCCCTCAGCCTCTGAGTAGCTGGACTACAGGTGTTGCCACCATGCC
 AGCTAGTGTGTTGTTGAGATGGTGTGTTGCCGATTGCCAGTGTGGCTTG
 AACTAGTGTCAAGAGGCCTGCTCCTCAACCTCTAAAGTGTAGGATTACAGATGT
 AACTACTGTGCTGATCAAAGTCAGATTTCCTGCTACTTAGTCAGTTGCTCTATG
 CTTTATTATACTTAATATTAGTATAGTACTGTATTAGTATATTAGCATATTAA
 TATTATTATACTTATCATACTTGAGTATATTGAGTATATTACACTTTAGTATATTG
 ATACACACACCACATTATTATTTGAGACAGAGTCTCCCTGTCTC
 CCAGGCTGAAGCACAGTTGGCTACTGCAACCTCTGCCCTTGGCTCAAGTGATTCTG
 TGCCCTCACCCCTCCTGAGTAGCAGGGATTACAGGTGTCACCAAGCCTGGCTAATT
 TGTATTGAGATATGGGTTTACCATGTTGCCAGGCTGGCTCGAACCTCTGAC
 CTCAAATGATCTGCCGCTTGGCTCCAAAGTGTGGAATTACTGGCGTGAGCCACTG
 CACCCAGCCTATTATCTGCTTTGATGGACATTAAAGTGTCTATATACTAGCTATT
 GTGAATAATGCTGCACTGAAACATGAGAGTGTTGAAACACTAATGTAACATAAGGTA
 CAAATAATAATGTCATGTTATCTGAAAG**GAACGTAAATACGACGTTGGGGAGGA**
GAACGGTTGATTCTTGACAGATCTGTGAAACATTATAAGAAGAATCCTATGGGAA
ACATTGGGTACAGTACTACAACCTCAAGCAGGTGAGCAGATTGGAAAGCTCAAGCTTCTC
 CTTAAAAACTAAACAAATCTAATAGAGAATTGCAAACATACAGAGGTAGACAGAA
 TAGTATCATCAGCCTCCATGTACCCATTGCAAGCTCAACTATCAAATCTTTTT
 TTTTTTTGAGACAGTCTTACTCTGTCACCCAGTCTGGAGTACAGTGTGCAATTCTGG
 CTCACTACAACCTCTGCTTCTGGGTTCAAGCGATTCTCTGCCCTAGCCTCTGAGTAG
 CTGGGACTACAGGTGCCACCACCATGCCGGCTAGTTTGTGTTTAATAGAGATGG
 GGTTTCAACCATGTTGCCCTGGCTGGCTTGAATTCCGACCTCAGGTTCTGCCGCCT
 TGGCCTCCGAAGTTTGGGATTACAGGCCTGAGCTACACGCCGCCCTAAATCTTT
 CTTATTATGATTCACTCACTGACTGCCGTATAGTACTTGAAACATATTCAAGATT
 TATTATCCCATTATCTGAAAGGCAATTACAGAGGTCTTTTTTTTTTT
 TTTGAGATGGAGTTGCTCTGCGCCAGGCTGGAGTGCACTGGCGTGTCTGGCTCA
 CTGCAACCTCTGCGCCGGGTTCAAGAGCTCTCTGCCCTAGCCTGAGTAGCTGG
 GATTATAGGTGGTGCACACACCCAGCTAATTGTTGTTAGTAGAGATGGGTT
 CACCATGTTAGCCAGGCTGGTCTGAACCTCTGACCTCAAGTGATCTGCCCTGCC
 TCTCAAAGTGTGGGATTATAGGCATGAGCCACTGCATCTGCCCTAAGGCTGTACAGAGT
 TTAAAGCAAGTTTCAATTAGATCCACTCTGGTTACCTTGTGAAACCTCACTTATT
 CACTTGGCATTGTCATTCAAATTCACTTGTGAAAGAAAATAGGCTGAAATTATATCAGCAGAGT
 TCTCTAAATAATGTGGTCTATTCAAAAGAAAATAGGCTGAAATTATATCAGCAGAGT
 AAAGTGTATGTGAAGAGTGAAGAAAGATACTTCTGGCTGAACAGAAAACACGGTGAA
 ACGATTGAAAACCTTATTGTAATTACAGGGCTCTGACCCCTGTCCGTGCTTT
 ATGAATATCAACATAGACATGTTTTTTTGTGATTAACACCGTTCTGTAA
 TATTCTTATTACATCAACTGCTGACTCGATCAG**CCCCCTTAACACGACTCGTATA**

10

20

30

40

50

AATGCTGCTGAAATAGAAAGCAGAGTTCGAGAACTAAGCAAATTAGCTGAGACCACAGAT
AAAGTCAAACAAGGCTTTGGAAAGAATTGAGGTAAGTATTAAAAAAACTGTTTACG
 TGAGTTTGTATATCCTATTTTAGTGGAGGAGAAGTGTGCTTGTGTTGGAATTGGACC
 TGAGAGACTTGAAACTGACGTCTTTTAATTGGCCATTGATTGACACGGAGCAAGTT
 GCTGAGAGGGCTTCTCGAAACAGAACAGAGCATTGTTCTGAGGGAGGGAGTTGGCAGT
 GAGTAGTCATGGATGTGCTAGCCGCTCCATTGGCTCTTGGTGGACTGGTGGCAA
 AATCTCAGAGAAACAAAAGGATCTAATTTCGAAAGATTTCAGCATGCACTGGGTC
 TTAGAACAATCTAGCCTAGTGCAGCAAATGAGTATGAGTAAAGAGAAACACCTT
 GTGGTGGCTTTTTTTTTGAGACAGGGCTCGCTCTGCGCCGAAGCTGGAG
 TGTAGTGGCGTGATCTGGTTACTGCAGCCCCGCTCCCTGGCTCAAGTGTCTTCC
 CATCTCAGCCTACTGAGTAGCTGGGACTACAGGCACATGCCCTATGCCGGTAATT
 TGTATTGGTAGAGATGAGGTTGCAGTGTGCCCCAGGCTGGCTTGAACTCTTGGG
 CTCAAGTGATCCTCTACTTAAGCTTCCCAGTAGCTGGGACTACAGGCACACGATACCA
 TGCCCATCTAATTGGTAGAGATGGGGTTGCAGTGTGCCCCAGGCTG
 GTCTTGAACCTTGGGCTCAAGTGATCCTCAGCTTGACGTGCCAAATGTGGTGGCTT
 AATTCAGAGTTCAAATTGATACTCTGGTAAGTTAAGTGAACACTGATTTCTTTTT
 AAATTATTTGGTATTATACTTTAAGCTGGGATATATGTGCAAGGTGAGGTT
 GTACATAGGTATACTGTGCCATCATGGTTGCTGCACACATTAACCCATATTAGGTT
 TTAAGTCTGCATGCATTAGGTGTTGCTTAATGCTCTCCCTCCCTTAATGCATCAG
 TGAAAAGTGATGATAGGCTGGCGTGGCTCACTCCTGTAATCTCAGCACTTGAGA
 GGGTAGGGCAGGTGACCCTGAATCCAGGAGTTGCCCCATCCCCAGACAGTGTGTG
 TGATGTCCCCCTCCTGTGTCATGTGTTCTCATTGTTGGTTCTGTGTTAG
TTTGCTGAGAATGATGGTTCCAGCTTCATCCATGACCTGCAAAGGACATGAACTCATT
CTTTTTATGGCTGCATAGTATTCCATGGTGTATGTGCCACATTTCTTATCCGGT
CTATCATTGATGGGCATTTGGGTTGTTCAAGTCTTGCTATTGTAATAGTGTGCAA
TAAACATATGTGTGCA TATGCTTTATAGTAGAATGTTTATAATCCTTGGGTATATAC
 CCAGTAATGGGATTGCTGGGTCAAATGGTATTCTGGTCTAGATCCTGAGGAGTCACC
 ACACGTCTCCACAATGGTCAACTAATTACACTCCACCAACAGTGTAAAGCATT
 CTATTCTCCACATCTCTCCAGCATCTGTGTTCTGACTTAAAGTGAACGTGATCTCT
 TTCCTGAAACTAATTGGGTTGGAGAATGTCCTGATGGAATGTGCTGTGTTCCATTG
 CACTCTTCTATATCACTTACCCATTGACAATGTGATCTCTTCATTCTCCTCATCCAT
 TTGACAGAAAATTCAAAACAAGGATTCTGGCATATTACCTTGCAGTTGTCAGG
 ATGTGACACGGTGCTAGTACACAGAACCTCCATAATGTTGTTGAATGAGATT
 CATTAACTCATGTTACATCATTTATTCTGTTCTGTTATGGGAATGATT
 TATGTTTGAGGACTACAATTATAAATTGTTGTTGAATGAATAAGTGAATAACT
 GGGCAAATAAAAGTCCTTTAGCAGAGTATGCTGAACAACTTGCTGAGATAGATATGAT
 TTCCCATTTCCAGTGAGGGGCTAAGGGAGTTAAGTAAATTATCAATCTTCATACC
 ACAGTTTGTGTTGTTGTTGTTCTCTGAGACAGAGTCTCACCTTGCT
 GCCATACTGGAGTACAGTGGTGCATCATAGCTCACTGCAGCGTCAAACCTCTGGCTCA
 CGCCATCCTCCCACCTCAGCCTCTGAGTAGCTGGTACTACAGGTGTGACCACCATAGC
 CGGCTAATTTCATTTGTTGAGATATGGGTCTCAGTGTGTTACTCAGGTTGGCTTG
 AACCTCTGAGCTCAAACAATTCTCTGTCTGGCCTCTAAAGTGTGGGATTACAGGTG
 TGAGCCACTGTGCCGGCCATACCACAGATATTGATTGAATTCCAGCAGTGGGAGGAG
 TGTGGAATAGAACATTCTCAGTCTGCTCAACATTACTGAACAGAGACTTGAATTGAG
 TTTATTCTCTCATCCCAGGCTTCGCGTTAGGCTCTGAAGACACTAGTGAACAAAGACAGAC
 AGGGTAACTGCCTTAAAGGGAGTTAGTGTGAGAGAAGGAAACAGTGTGAAAGCA
 TCAGTAAAAAGTGTGATGATAGGCTGGGCGTAGTGGCTACTCCTGTAATCTCAGCACTT
 TAGAGGGTGAGGGCAGGCAGCTACTTGATTCAGGAGTTGAGACCAAGGCTGGCAACAT
 GGTAAAACCCGTCACAAAAAAACAAAAAGTAGCTGGGTGTGGGGTGCGCACCCA

10

20

30

40

50

CAGTCCCAGCTACTCTGGGGTTGAGGTGGGAGGATTGCTCGAGCCTGGGAGATTGAGGC
 TGCAGTGAGCTGAGATCACGTCACTGCTCCAGCCTGAGCAACAGAGCCAGAACCTGTC
 CCAAAAAAAAAAAATTGATGATAAACATAGTAGAGACAGAATTGAAATCTCAGCCTC
 ACTGTTGCTTCTTGTCCCCTGCCTGCCTAAATAATAAAAAGGCAGCATTCAGCAGTCA
 TTCATTTCAATTACTTCACTTCATTTCACCTCATAAACGCTCATGAGGTAAGATGGAA
 GATACAGAAGTTTAGAAACCGCTCATCAAATTGAATGAAAGCCGATTGTTCCAAAAC
 TTTTAGTGTGAAAATTCTATTATGCAAAAGTAGAGAGAATGGGATAGTTATAGCA
 GTATACCTGACACCCAGCATTAAACAACGTGATAATATGCCAATCTTTCGACTCTG
 CCCACTCACTTCCCAGCCCTGACTTGTCTGAAGCAAATACTTTTTTTTTGA
 GATAGAGTTTGTGTTGTTGTTGTTGAGATGGAGTCTCACTCTGCCCCCA
 AGCTGGAGTGCTGGCTGATCTGGCTCACTACAAACCTCCGCCTCTGGGTCAGTG
 ATTCTTGTGCTCAGCCTCCTGAGTAACGGGATTACAGGTGTGACCACATGCCAGC
 TAATTTTGATTTTAGGGACAGGGTTTCACTATGTTGCCACGCTGGTCTCAA
 CTCCTGACCTCAGGTGATCCGCTGACTTGGCCTCCGAAAGTGTGCTGGGATTGTTAGGTGTG
 AGCCACTGCTCCGGCCTGAAAGCAAATCTAACACATCTTGCTGTAACTATTTA
 TTCAAAAAATTATAACCTGAATAGCATTATCATATCTAAACATATTAAACAGTATTTCT
 TAATATAACACATATCAGTCACATTTCCTGATTGCTACACACACACACACAC
 ACACACACACACTTGCATTTGTGTTTTCTTTAGATGGATCTCACTCTGTTGCC
 AGGCTGGAGTGCAATGGTGCATCTCAGCTCACTGCAACCTCCACCTGGGCTCAACT
 GATTCTTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCCACACCTCACCTGG
 CTAGTTTGATTTTAGTAGGGTGGGTTTCACCATGTTGCCAGGTTGGTCTCAA
 CTTCCGACCTCAGGTGATCCACCCACCTGGCCTCCCAAAGTGTGCTGGGATTACAGGCATG
 AGCCACTGTCGCCAGCAGCAATTGTTGATTGGGAGTGCTTCTCCACCTGATTAT
 GAAAAAAATTCAAATGTGATAAAACAGATCATATAAAGGATCCTGATATGCCATTATC
 AGCTTATCAATTATCCCTGTCATCATATTTTTATTAAATATTCAATATTGTGG
 AATCCTAAAAATGCATCACATAACCCAAACATTGTTCATATTATACCAATTGTCTTATAA
 TTAAAAATATTGTTCAATCATTTTCAGATAAGCTCACACACTGTGGTTGGCTAAG
 TCTCATAATATTCCTGTTGAAAAATCTAAGTCTGGGCGTGGCAGGCTGTTGCTCAT
 TCCAGCACTTGGGAGGCTGAGGTGGGCGATCACGAGGTCAAGAGATCGAGACCATCCT
 GGCAACATGGTAAACCCGGCTCTACTAAAATACAAAATTAGCTGGGCTGGTAGT
 GCGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTGAACCCAGAA
 GGTGGCAGTTGCAGTGAGCCAGATCGCGCACTGCACCTCAGCCTAGAGACAGAGTGC
 GCTTCATCTCAAACGAAACAAAACAAACATCTTAAGTCTTCTAGAATACTTGATGC
 CCCTCCATCTCTCTTCTGTCTTCCCTCCCTCTGTCTTCTGAA
 GAAAGCAGATCATTGTCCTGAGAGTTACTTATAGTCTGAATTGCTGAGTGCTCT
 GTGGTGGACTTAAGCATGTATCCATCCCTATATTCTGTAAAGTGTGATATCTAGAGA
 CTTCATGGGATACAAGTTCTTGGCAAGATAGCATGTATGGTGGTGTATCAGGAGGTG
 TTTATGTCCTGTTCTCTGTATTCTAGCAGCTCTGATCATTACTTAGA
 TCCATTAATTCTATAAGGGACTATATGGTAGTGTATTGTAATTCTCATTTCTTCT
 TTGTTAGGTTGGCATATTCTATAAAAAGTTTCTACGCCGAGGGTTGATTTCTTCT
 CTTACTAACGAGTTCTTCTTTCTTTCTTTCTGAGGTAGGTCTCACTGTG
 TTGCTCAGGCTGGTGTGCAGTGGCGAAACACACAGTGCAGACTTGGGCTGAGGTGA
 TCCTCTGCCTCAGTTCTGTGTAGTGGGACCACAGGTGCATGCCACCATGCCCTGGCT
 AATTTTTGATTCTTTGTAGAGATGAGGTCTCACTTATTCCAGGCTGGCTTGAAT
 GTCTGGGCTCAAGCAATCTTCTACCTCAGCCTCCTGAGTAGCTGGACTACAGGCACAT
 ACCACCATGCCAGCTAATTCTTAAATTCTTCTGAGGAGATGTGGTGTATTATG
 TTGCTCAGGATGGTCTCGAAGTGCAGAGCTCAAGTGTGATCCTCCTGCCAGCCTCCAGT
 GTGCTGGGATTATAGGTGTACTACAGGCAAGAGCCAATGAGCCTGGTCAGATTCTT
 CCTGATTGAAATCTGTTATGGGTCATTGATACTTCAAATCAAACCTCAGGGTTTCAG

10

20

30

40

50

GATTTTACTAACCTCATTGATCTTACCCATGTATCTCCTTCTAATGCCAAAATCC
 TACTTCTGAAGCCATAATAAGATTATTCATTGTTTATCCCACATTACACACAACAT
 CTTAGAATAATGACTCCAAATAATATGATTACTGAAAACAGTTAATTTTTGCGCT
 TTTCAAAAAATCCTCAGAGATGTGAGTCAGTTACTGTATTCTGCTGGCACAGTGG
 CTCACGCCTATAATCCAGTCTTGGGAGGACAAGAAGGGAGGATCGCTGGACCTCAGG
 AGTTTGAGACCAGCGGGCAATATAGTGAGACCTGTCCTACAAAAGAAAATAAAAAA
 TTAACCAGACATGGTGCATGCCCTATAGTCCCAGCTATTGAGAGGCTGTGGCAGAGT
 AGGCTTAAGCCCAGGAGTTGAAGCTGCAGTGAGATCAGATTGTGACACTGTACTCTAGG
 GTGACAGAGCAGGGACCCCTGTTTAAAAAAAAAAATGAAAAAAACTTCTGTGCCCTAG
 ACTCATTGTAATCGTCCTCTCTGTGGCTATGCTAAGGGTATATGGTTAGT
 TTATTTGTTCTTTAAATCTCTTGTAAAGTTTATTTATAATTACACAAATAC
 TGGCTTGATAGTCAAATTGAAAAACAAAGTGATTCAAAGAAGTCTACCTCTATCCT
 TGTCCTTCCTATGTTTAGCCATAGTATAAAAGTTATGGTTATCATTATATTCAAA
 AATATAAGAAGATATTCCCATATCCACTTTTCTTAAACAGTAGCATAACTTACATAC
 TTTTTCTAACCTTGCTTTTAAATATCTGGACATCCTGGATATCCATAATAGTGTCT
 AGAGATAGTCCTATTCTTTTACTGTATAGTAATCCACTGTGACTTGTACCATAGT
 TTATTCAACCTATTGATGGCATTGGTAGTTCCAATGTATCACAGAGGATTACA
 GTGAATAGCCTTGTTGATGCATCCTGCTTACTTTGCTGACTACTGGTAATATTAACAT
 TTTTTATGTTCTGTTAAATGGTGGTTATTATTCTATCTATAACTTTTATTATAACA
 TGACTTTGGTTAGCATGCTTAACTTTAGCATAACATTGCAAGCTACTTGTGTTAAT
 TAAAATTTGGTTAAATGTAAGGAAATAGTGAGCTATTGTAATCTAGATTCAATAGAAT
 CTTATACTCCTTACAAATGATAGCTGAGTTGATCATTGTAATGACTGTGAACTT
 AAAAATTACAGCATTTTAAATTTAAATTGTTAACATTAAATTTAAATAAT
 AGACACACAAAGTAAAAGAGAAGAAAAAGAGACAGGGCTTGCTATGTGCCA
 GGCTGGCTCAAACCTCCAGGCTCAAATGATCCTCCTGCCCTGGCCTAAAGTGTAAAG
 CCACCACACTGGCAAAATTAGTTCTTAAACAAACATTACAGGTTATCTGGTAC
 CATGGTAGCTTAAACACTAGGTCACTTAGAACAAAGCTTAGGAACAAAGTCAGACT
 TTCACAAAGAGCTTGTGGCAATGGGTATTTTGCAAATTCCATTGGTGGGTCAAG
 ATGTGAGTTAGAGGAACCTTAGCCTGACTCTCTGGCCATGGAAAAGATGGTGCT
 TCTAAATGCTGACCTGGTATTACACTGTCACATCTAAATTGTTGTCATCTTATA
 CATTATAACAACAAAGGAAAATTGAGTTGACTTAAAGAGGAAGTGGAAAATAACGA
 GATCACATCTGACTCTACAGGCTCTCACAGAGGTCAAGACTGAGGTGGTAAAATTGTTG
 TGCACAAATTAGGGCATTAACGTTCATGAAACTGAAGCTATCTAAATAGCTGATG
 GCCTGCTTCTAGATCTCTATACCTGCTCTCAAATTCACTGTTTAAATGAA
 CCCTTGAGGTTGAAACCAGCGAAATAAGGCTGAAAACAGAATAAGCCATTATTGAAAAA
 ATTAGGAACCTGGAAGCAGATACTCATAACTCTAAACCTCTGAAGCTAAAGTTGATCCA
 CAATAGCAAAGCATTATCATTAGTGTGTTGACTTGTGTTCTGGCAGGTGATAA
 ATTGGGATCACTTCTTACAGTGTGCTCTGATAGTCTTAAACAAACCAAGAGCTC
 TAAATGTAATGCCATTGGTAATTAAACTCTGATTGTCCTATGCTGCTCTGGTGT
 TCTGTAATTCTACACGTCATTCTAGGTATCACTATCCAGAAGACGTTACTTGTGCTT
 TGATGCACTTAAATGTAAGTCTCTGTAAGCTCTGGTTATTCTCCTTGCTG
 CTGAAATAATTCACTGTTGATGATTCTGTAGGATATGTTGTCATGACATTG
 AAACCCAAGTCTTGATTTTCTACAGATTGCTGTTCTTCTTATTAAATT
 TTTATTAGTTATTATTATTGAGATGGAGTCTCACTCTGTCACCCAGGCTGGAGTGCA
 GAGGTGTGATAGCTCACTGCAACCTCCACCTCCGGGTCTTGTGCCCTAGCCACCCAGG
 TAGCTGGGATTACAGGCACGTACCACTCTCAGCTAATTGTTGTTAGTAGG
 GATGGGATTCTCCATGTTGGCAGGCTGATCTAAACTCCTGACCTTAAGTGATCTTCC
 TGCCTGGTCTGAAAGTGTGGGATTACAGGTGTGAGCCACTGTGCCTGGCCAGTTAT
 TAATTTTTAAAGAGATGGGTCTCACTATCTGCCAGGCTGGAGTGCAGTGGCTCTT

10

20

30

40

50

TACAGGCACTGTTAGTGCAGCCTGAACCTGGCTCAAGTGATCCTCTGA
 GAGGCTGGAATTACAGGCACACACCAGTGTCCAACAGATTGCCATTGTGATCTGT
 TAAATATCTCACTCCTGCAGTATCTCTGCTCAAGAATGTAAGAGATGGATAATATT
 TTAGATTGAAACAAAGTAAAGTCTGCTCAAATGAGAATGACACTAACTAAATGA
 AAAGGCCGGTTATAATTCTGTAATTGCGCTGCAATGTGTGTATTGTACACTTGA
 ATCGGCCCTGTGCATGTGGCGAGGTGCATATTGCATGGTTGTATTGAAAAGGTGCTGG
 GCCGGGGCTGGTGGCTCACACCTGAATCCCAGCAATTGGGAGGCTGAGGCAGCTGGAT
 TACCTGAGGTTAGGAGTTCAAGACCCAGCCTGGCCAACATGGTGAACACCTGTTCTAGTA
 AAAATACAAAAAATTAGCTGGGTGTTGGTGGCTGTAATACCAGCTACTAGGGA
 GGCTAAGGCAGGGAGATTGCTTAAACCTGGGAGGCTGAGGTTGAGCTGAGATTG
 TGCCACTGCACTCCAGCCTGAGTGTATCACAAAAAAAGGTTTGCCTCT
 CTCTGTGCTGCTGCCCTGTTGAGTCTATAGGCCTGAGCTGCCAGGGGTACTGTGG
 GCTGAGACTGGACATTGCAACCGACTGCAAGGCACCGTGGGACCCAGGTTGATGGAC
 TGTCTCTCGGGCTTCTCTTCCATTCATCTCCTCTAATCCTCTGTATCCAG
 TATCCTGCTCTCCATACACCTGCTTCATTCTTTCTCAGTAGATTCTGCTTCT
 TGACTTACAACCCACTCTAGCCCCCTTCAGATATTGAAACTAGCAACTTCAGGCTT
 TGTACCAAAAGTCTCAGAGATTCTCATTGACTCGGATGCCATCCATCTAGTCAAAGAA
 CAATGTCAGGACATGAACATGTGGAACAAAAGTGTCTGTTGGACACCTTGGGAGA
 AATAGTTTCAGTGATGAGGTTGAGTTGAGCTGGGAGATATCCAAAAATATCTGCCA
 AAAACTATAGACACTCTGGGTCAGTGACTTATTCTCCTTCATTCAAGCAAATACTGA
 TTGAACACCGACTGTGATGCTGGATCTATTCTAGGTTGGGGTGGAGCAGTGAAACAAA
 TCAGTCTTATCTTATAGAGTGACAGTCAAGTGGGAGAGACAGGAGTAAACAAAGAA
 ACAGTCAATATTCAATCTGTGAGATGGTGAAGTGCTACAGAGAAAACAAACTAGTGT
 AAGATAAAAAGGGTGTGTTGATAGGCTTTACTATTAGGTCTTTGATAAGGGCAT
 TTGAACAAAGCTGAGGAAATAATGGGACCAACCATGCAATAACCTCAGGGAGAAC
 TTCTAGGTAGAGGAAACAGCAAGTGCAAGGCCCTGAAGTGGGGTTGTTACCTGTT
 GCACAATCTGCACACAGGCCAGTACAATTGGAATGGATGGAAATGTAAGAGAGAAAGT
 TGAAAAGGCCAGGTGCAGTGGCTCATGCCACAATCCCAGCATTTGGGAGGCTGAAGTG
 GGAGGAATTGAGATCAGCCTGGCAACAGAACCCAGACCTCAGGCTAATTGTTGTT
 TAGTAGAGACAGGGTTCACCATATTGCCAGGCTGATCTCAAACCTGACCTCAGGTG
 ATCCTCCTGCCCTGCCCTAAAGTCTGCTAGGATTACAGGTGTGAGCCATGGCCCCAGC
 CGTATCTTGTCTTAAAGTAATCTGCTGCTGGTAGGCAAGAATTAAAATATAAA
 AAATTAAAGAAAGAAAAAAATAAGTAAGTAACTATAACAGGTTGCTGGCGTAATGG
 TGAGTGTCAATTCTTCTGCCCTAGGTATTGGCTCTGTTGCTCAGAGCAGTGCAGGC
 GAAATGGTCATTAGGCATCGTCATGGTCCTGGGATGCCCTGGCTCAGCCAGTTATT
 TCTGTCGCTCTCCTGGTCTTCCACTTCCATTGAAATTCTAGTCAAG
 AGCTGGGCCAGTGGTTCAATCCAAGGGTTGGAAAGCCTGGGTCTATTGGC
 ATTGCACTGGGCTGCTGCCATTAGGTTGGCAGGGGTCTGGCTGGGAAG
 CAGGAATGTTAGTGGCATAATGTAAGGGTTGGCTACATTACATAAGGGAGACAA
 TGAAAACCTTAACCTCCACAGTAGTGGAGTAGTGGCTGGACTCACAGTCAGTAGT
 GCCGTTGGGTACTCACATGTACACATGGATCAGGACATTGACTTCTGTTGACCTT
 TAATAGTTATTAGATGTTAGGCTGTTGCACTGCTCTAAAGGAATATCTGAGTCTA
 GGTAATTATAAGACAAGAGGTTAATTGGCTCATGGTCTGAAGGCTGTACAAGCATG
 GCTCCAGCATTGCTCTGGTGAGGGCCTCAGGAAGCTCCGGTCATAGTGGAAAGGCAA
 AGGAGGGCAGACGATCACATGCCGGAGTGGTGGCAAGGGTGGGGAGCCACGCTCT
 TTTTTAATTATAATTGAGACAGTGTCTCACTCTTGGCCAGGCTGGAGTGCA
 GTGGCGTGTACTCAGCTCACTGCCCTGCCCTGGCAAGGTTCAAGCAATTCTCCTGCCT
 CAGCCTCTGAGTAGTGGGACTACAGGCGCGCATCACATGCCAGCTGATTGAT
 TTTAGCAGAGACAGGGTTCACCATGTTGGCCAGGCTGGTCTGGACTCCTGATCTCAA

10

20

30

40

50

GTAATCCGCCTGCCCTGGCCTCCCAAAGTGCCTGGGATTACAGGCATGAGCCACTGCGCAC
 GGCCACCACACTGTTAAACAACCAGATTGCACGTGAACCTAGAGTGAGAACTCACTGT
 GAGGATGGCACCAAAACATTCACTGAAGGGATCCACCACCTTCCTTAGGCCCCACCTCAA
 CACTGGAGGTCAATTCAACTGAGATTGGAGGGACAGACATCCAACCGTATCATT
 AAATTAAATAGTTTATGCAGTTTTGGCTTAGATCTGTTAGACTCCTGCAGTCAG
 GTGCTGTAACTAGCCTCTGGCCTTTTGAGAGTTACAGTTGGTGAAACCCCTTGG
 ATGTATTATTGGAAAATGGGATATCTGCAGCCTGTCCTGCTTACATTATCCTT
 TTTGCTGCTGCCCAAGCCTCCTATTAGCATCCCTGCCAAGGCCAGTGGAGAAGGATG
 GAGATGCGGTGACATTCACTGAGCTTGCAGGGTCAATTAGCAGCTTGTGCCCTAGGGACTGCT
 GGTGGGAGGGAGGGTGTGGAAGATAAACCTGACAGGAATGTATTCTCCTGAGGGCAGG
 GTTATTGATATTCTGGAGCTTAGAACCTAAAGCCTGGTGCTGGGAGGAAGCGCC
 CTTAGCATTGGTAGCCTCTGGGCAGAGCATGGAAAGTCACAACCTCTGAATTGTTG
 TATTTTCAGTCTCACTCTAGATGGATGGCATCTCTGCTATGGGAATGAAATATGTTA
 GGCAACTTGAATGCCAGGTGCAAGATGAGGCTGGCTAATTGGTGCACTAGGGAAAGGAGCC
 GGGGGAGAGATGTGCTGTTAGCTATTATCATCTGTGACAACGTGTCAGCTGCTGGCAGTT
 AGCACCCACCTGAGCCTGGGATGCAGGGTGCCTCTCTGCTGTGGAAGCCTCTGG
 ACCCAGCAGCCATCTTGACTGTGCACTGTTCAAGCCCCAAGTCCGCTGGAAAGAGGTGAT
 TGAGAACCTACTGCAGGATAAGGAAAGCGCAGGACAGGTGCACTGGCTCACGCCCTGTAAT
 CTCAGTGTGTTGGGAGGCTGAGCCCCAGGAGGGCTGGAGTCCTGAGTGCAGGACAGC
 CTGGCAACATAGTGAGACCCCTGCTTACAAAAAGGAAATTAGCCAGATGTGGTG
 GTGCGTGCCTGAGCTCAAGAGGCTGAGGTGCGAGGATCACTTGAGGCCAG
 GAGTTGAGGTTACAGTGAATGATCATACCAACTGCAATTCCAGCCTGGGTGAGAGAGC
 ATGACTCTGCCCCAACAAAAAAAGATTAAGGGAGCCTCTGGCAGACCTGATGAT
 GGGTGGCCCAGCCAAAATGAGTATTGATGAGGATTCCCTGGCTGGAACCTGTAATT
 GTCTGCAAAGTATTCCCTTGTTGAGATGATTCTGGTACCCATCACGGTA
 GGTAAGATGAATTAGCAAATGAGAAAGGTTCTCTTTTACCTTATCTAGTCCGTAG
 ATGAAGCCTGAAGAAGGTCTCCATATGGTAGTAGTAAGTGTAACTACCTCTAAC
 CTTGCCTGTGTTTTGGGTTGCAAAGCCTAGGAATGCCAGTATCTAGGTAGAAT
 TTGATAATATTCATTTGGTATATTCCCTTTCTGTTACCTTATACAGCAAA
 TGAAAAAATTTTAAATTTGCAAGTAAGGGCAATTCTTTCTTTCTTTCTTT
 TTGAGACAGGGCTTGCTCTGGCACCCAGGCTGGAGTGCACTGACACAATCTGGCTCAC
 TGCAACCTCTGCTTCTGGGTTAAGCGATTCTGCTCAGGCTTCAAGTAGCTGG
 ATTACAGGTGCCACCCTCCAGCTAATTTCATATTGGTAGAGACCAAGGTT
 TGCCATGTTGACTGGCTGGTCTGAACCTGACCTCAGGTGATCCACCTGGCC
 TCCCAAAGTGTGGGATTATAGGCTTGAGCCACTGGCCTGGCTGAGGCAAGTT
 GAAATATTTGTAAGGGAGAAAAGAGGAGTTCAAGTAAGAAACAAATGACATAAG
 AGGTGGTATGCAGAGATGCCAAAGCATCTGAAGGTGCTTTTTGGAAACAGAGTC
 TTGCTTCATTGCCAGTCTGGCTGCACTGGTCAATCATGGTCCCTGCAGCCTGACC
 TTCTGGCTCAAGTAATCTCCACCTCAGCCTCTCAAGTAGCTGGACTACAGATGCAT
 GCCACTATGTCGCTAACTTTAAATTGGTAGAAGCCAGCTCACCCTATTGCC
 AGGCTGGCTTGACCTCTGCTCGAGCAAAATACCGATTGATTAAGTCTGGGTA
 GGACCTGGGCTGGGATTCTAACCCAGCTCCAGGTGGCTAATGCTGCTGGTACAGA
 CCACACGTGGAGTAGCCAGTGTAGAGTTCACTGAGCAATAGTGATGTCATAGAAATAGCC
 AGTATCTGTATACTGCTTGTGATGTACGCACGTGATAGTGATGTCATCTC
 ATTTGACCCCTCACCCGCCCTTGGGGTAGAAAGGATTGTGCTCATTTCACACTCAAG
 GAAACTGAGGCACAGACAGGCAAAGTAGCTGGCGAACAGAAAGGAACTTAGAGGCAGG
 CCCTGATTAGCTCAGAGACTAGAAGGCCTGTGCGTCATCCTGAACAGCTGGACTTGAT
 CTTGAAGGTGGAGGGAGAAATTGAAGGGTAATTAAACAGGAACTGTAGGAAATTCAACCTT
 GCATAGTGAATTGGCTTGCCACGTGTGCCCTGCCACCGCCCCCCCCACCTCAGTGAAGTGT

10

20

30

40

50

CATGCGAAGTTGGGTCGTAAATGAAGGCCGAATGCTTCCTGACAAGTTGTTAAA
 TCAAGCTGCTAATTAGTCCCAGCCCCCTCCCCGGTATGTATTTTGTTGATGTCGT
 TTCACCTCATTAGTTGAAGTGAATTGATTCAAGTGTGACTTGAACACAGCAGTATTGC
 CACCTTAATAACCTGTCTAACATCAAGGTTAACACCTTCTGCTAACACAGCAGTATTGC
 TTGGTAAGACTGGCTCACAGTCCAAGGAAATGCTGCCAGAGAGGGCAAACGTGCTTAA
 CTCCCTAACCTGAGCTATTAAAAAAATTCAAATGACTGATTCTGTACAGTTCTAC
 CTACATGTTTATTTTGTCAGGTTCAAGCTAGTTAAATGCTTTGTGATGAGCTTA
 TGTCAGGCTGAAGGTTGCATTTGAAACTGAGCGTCAAATACCAATTAAAGTCCAGAC
 CTTTACACTTGTGAAATTGAGATAATGAAATGAAATAAAACAGGGCTGCTGTTGTG
 AAATATGACTGTGTTTCTCTTAGCTACCTTTACTTCTGATTTGCCTATTGTGATTCC
 TATAAATTTCTTTCTTAGCTACCTTTACTTCTGATTTGCCTATTGTGATTCC
 CATTAAACACTAGGCTTTGTAACCAATTATCCCTGAAATTGACTTTTTTT
 GAGACAGGATCTGTTGCCACACAGGCTGGAGTGCCTGGCTCATATGATAAAC
 AGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGACCTGCTTATTAAAACA
 AAAAAAGAAGAAGAAAAAAAGAATATAGATCACAGCTGTTATTGTATATGCTACGCCA
 TCCTTGGTGGGTTTCATTCTTATAATTGTTATTAAAGATTTCTTATGAATATTC
 TATTGTTCTATTGAGAAATTAAAGGAGAACACAGTGGGAAAAAAAACAAGAAAAG
 GACTTCATAATCCTGCTACCCTGGAGAAAAAAATCACCATTACCTATTGGTTCTT
 CTCCCACTTTTTTTCAGATGGAGTCTCCCTTGTACCCAGGCTGGAGGGCAGG
 GACGTGATCTGGCTCTGCAACCTCTGCTCCTGGGTTCAAGCGATTCTCGGCCTCA
 GCCTCCCGAGTATCTGGGATTACAGGGGTG GCCATCACACCTGGCTAATTGGTATT
 TTAGTAGAGACGGGGTTTGTCACTGTTGCCAGGCTGGGTTGTTGCCATGTCTGGTT
 TTGTCATATTGGCCAGTCTGTTGTCACTGTCAGGCTGACATGTTGTCACTGTTGCCAG
 GCTGGCTTTAACCTCTGACTTCAGGTAACTCTGAAGTGTAGGATTATAGGCCTGAGCC
 ATTGCAACCTGGCCTCTGCCTTTTTAAAGAAAAAAATTAAACATTTTCTTT
 TAAGATAGCGTCTCATTGTTGCCAGGCTGGCTTGAACCTCTGGCTCAAGTGA
 TCCAGCCTCAGCCTCTGGAGTAGCTGGACTACAGATGCACATCATGGTCTTATGCC
 ATTTCTTTGTACGTAGGTGAATGCAAGTGTATGATTACATCATGCTATTGGAGGT
 TTGACTTCTTCACTTCATCTTCCAAGGTGTATTCTTAGTACATCTTT
 AAATGGACATAGAACATTCTTGTATGAACAAACAATAGTTATTAGGCGGTCTT
 CCTGTTGGACATTATATTATTTCAGCATCTCCACAGTGTGAGCATTGAGATGA
 ACCTCTTTTTTTGGAGACGGAGTCTCGCTCTGCCAGGCTGGAGTGCAG
 TGGCACAATCTCCTCAAGTGTCTGTGACCCCTCCACGTAGCTGGGATTACAGG
 TGCCCATGTCTGGTAATTGGTAGAGCTGTGGTTACCATGTTGCCA
 GGCTGGTTGAACTCCTGCCAGTGTGACCTGCCACCTCAGCCTCCAAAGTGTGGG
 GATTACAGGTGTAAGCCATACGCCTGACCCAGATGAAACATTCTGTAGCTATGCCACAC
 AATTCTGAACATTCTAGGATGAATTCTTAAAGAAGTAATGCTGATCCAGGTTT
 CTTTCTGTGACTCTTGACAGTAATAATTGACTTTCTTCCAG**ACACTAC**
ACAAACAGGAGTGC**AAACTCTCTACAGCC****AAAGAGGGTCAAAGGCAAGAAACAAA**
ACAAAAATAGATATAAAACATCCTGCCCTGTAAGTATCAATATTCCGCTCAGTAATAGT
 CACTCTGGAGATTGATTCTAGCACCTCTGTACCTTCTCAGGGCTGTGCTCTT
 GTTAGCACATCGGAGGCCTAGCTTCTTAATTGCAAGCAGTTCCAAATAATCAACCA
 TGGTGGGTGTTGATGACTTCATTCACTGAGCTCCCGTGTGCTGATTACTGAGTAAAGTT
 GCCACTAGGTGGCTTGTCTGGTTGGTCTCTGTTAATTAAATTCTGTCTGCCA
 AGATAGATCATCTCAAGGCTGGGATCTCTCAGTGTCAAGGGACCTAGGGTGCAGATT
 GTGTCTTGACTCTCCTCACTGGGCTGTGAGTCTGGTAAGGCTGCCCTCTGG
 GACTCAAGTCCCTTAAGTGGGAAACAGACAAACACCTCTGAGGGCTCCTAGA
 TGCTTGCTGATCCCTGAGCTAAGTTACTGGAGAAAGGGTATATACCTAAACTGCTCAG
 AAGAAGACTTGTGGGCCGGCGAGTGGCTCACACCTGTAATCCAGCACTTCGGAGG

10

20

30

40

50

CCGAGGCAAGCGGATCACCTCTGATCAGGAGTTAAGACCAGCCTGGCAACATGGTAA
 ACCCCATCTACTAAAAATACAAAATTAGCATATGGTGGTGTGCCTGTAATCC
 CAGCTACTCGGGAGGCTAGGGGGAAATTGGTGAACCCAGGAGATGGAGGTTGAGT
 AGCCGAGATGTGCCATTGCACTCAGCCTGGTACAAGAGCAAACCTCCGTCTCAAAA
 AAAAAAAGGAAGACTTGTGAATATTGCAAAGCTGAAAGCTGTACCTTCATAA
 TTTGAGACATAGTCTCACTCTGTGCTCAGGGTGAGTCACAGCTACTGTAGCCTAAC
 CTCCTGGCTCAAGCGATTCTCCACCTCAGCCTCTGATTAGCTGGACAATAGGCAGG
 CACCAAGTACACCTGGTGAATTACAGTTTCTGTAGGCCGGCGCAGTGGCTACGCCT
 GTAATCCCAGCACCCTGGGAGGGCAGGTGGCGGATCACCTGAGGTTAGGAGTTGAGA
 GTAGCCTGGCAACATGGTAAACCCCCATCTCTATTAAAAATTACAAAATTAGCTGGC
 GTGGTGGTGGATGCTGTAATCCCAGCTACTGGGAGGCTGAGGCTGAGGAGAAC
 GCTTGAACCTGGGAGGCGAGGTGCAATGAGCCGGAGGTGCTATGTGACCAACTGCACT
 CCAGGCTGGCGACAGACTGTCTCAAAACAAAAACGATTAAAAATAATA
 AAATTTTCTAGGGCGGGGCTCCCTATGTTGCCAGGCTGGTCTTGAACCTCTGGC
 CAAGTAGTCCTCGCTCAGCTCCAAACTGTTGGGATTACAGTGCAAGCATTGTG
 CCTGGCTGTACCTCTGTAACACCCAAATGCCACCTGGAAAGCCAAGTTGAATCATGA
 GGAAAAAAAGGCCTGGAAGGATGTAGACCTCCTTTCTACTTATTATTATTATT
 TTGAGATAGGGCTTACTCTGTGCCCAGGCTGGAGTGAGTCAGTGGCATGATCATGGTCAC
 TGCAGCCTCAACCTCCGGCTCAAGTGGCCTTCCCACCCAGCCTGCAATGTAGCTGG
 GACTACAGGCATGTGCTACCATGCCAGCTAATTTGTATTGGTAAATTATTTTT
 TGTAGAGACAGGGTTTCGTCATGTTGCCAGGCTGGTCTCGAATTCTGGCTCAAACGA
 TCTGCCGTCATCGGCCCTCCAAAGTGGTGGGATTACAGGTGTGAACCAACTGTGCTGGC
 CAAGGTATAAGTGAAGGGGAGTGGTGGGAGAACACTGCCACTAGAGCCCCAAAGTC
 GTTCTAGCACCCTCAATGCAATGGGTTCTCCATTGATGTGCTATGCAAGGCAGTGCA
 CTGAGGAGAAAGGAAGAACATTACAACCTCTCTTATATCTGTCCTAA
 AAAAGAAAAAGAAAAATTGTCAGGGCTAGATTGATGCAAGGGAGTGCATAATGTTT
 ATTGATGATTGATGATTGATATAGAGATGGGGGGCTCACTATATTGCCAGGCTGA
 TCTCGAACTCCTAGGCTCAAGCAATCCTCGTTGGCTCCAAAGTGCTGGGATTAC
 AGGCATGAGCAGTCACCTGGCTATGCATACTATATTATCCAACCTAACAAAGGCT
 TGCTGCCGTAGTCATATGTGATACATTGCACTAGAAAAACTGTGTTGATTGGGG
 TTGTGATCAAATTGGAGAGCATTGTCATGTGTCATGCAACATAGTCTTCTTCT
 AATCTTGTAAACCAATTCTTGTCATGTCATGCAATTGCAACTAGTCTTCT
 ATGCAATGTACATATACATGGTATATGATATCATATCTACATGGTATTGTAATGT
 ATATGTATGAGGATGGGGAAAGTGGAGACATTGTAATACTGAGAAAAGGAGTGG
 ATTTGAGAAGCAGTTGAGCTGTAGCATGGTACTAGTGACCTTGAGGAAGCCTTATC
 CTTTTTTGGATTTATTTCAATTAGAAATAGACAAGAGTTCTATGTT
 CCCAGGCTGGCTTGACCTCTGGCCAAACTATCCTCTGCCTGGCTCCAAAGTG
 CCAGGATTACAGGTGTGGACCACCATGCCCTGGCACCTTGCTTCTATGCTAAGTTG
 TGACATCTGCTCAGGGTCAGGTGGTATTAAATGGTATAAAATGATGGAAAGTGAGG
 GATCAATGGTATGCACTAAATAGAATATGCTTTCTCCCTAAAGGCTCATT
 CAGATGTTCTGATGAAACATCTCATTCTAAAGATGAGGAGTCTGAAGCAAAAA
 GACATTATTCTTAAGACACATGGCTGTTACTAATTCCATTGCAAAATATGTT
 TAGGTAGAGCACTCAGATTATACGAATAATAGACTTTGTACAGAATTGGACAGTT
 GATACTATCAGAGCCTTGTGATATTCCACTGCATTATGCTTCACTAAAAAATACCTGGC
 GGGTGCCTGGCTACAACGTGTAATCCCAGCACTTGGGAGGCTGAGGTGGCAGATCAC

10

20

30

40

50

CTGAGGTCAAGATCAGCCTGGCTAACATGGAAAACCCATCTACTAAAAA
 ATACAAAATTAGCCAGATGTGGTGGCACGCTCCTGTAATCCCAGTTACTCAGGAGGCTG
 AGGTATGAGAATTGCTTGAGCCAGGAGGAGGTTGCAGAGAGCCGAGATAGTGCTAT
 TGCACCTAACCTGGGTGACAGAGGAAAACCTGTCTCAAAAATAAATTAAACAACA
 ACAACACAACAACAAAACCCCTTTTATATGAAATTTCAAATATATTCAAGAGCA
 TAAAGAACCCACATGTACCCATACCCAGCTCAACAATTATCAACTCATGCCAGTCTT
 GGTTCATCTATACCTGATCCACATCTCTCTCCTGAATTATTTGAAGCCCATCT
 CAGACATCATGTATATGTATACCTCAATCTTTTTAAAACCCCCCTCCCC
 TTTCTTTCTTGAGACTGTGTCCTACTCTGCATCCAGGCTGGAGTGATCTGGCTC
 ACTGCAATGTCGCCTCGGGTCAAGCGATTTGTACCTCAGCCTCCCTAGTAGCTA
 GGATTACAGATGTGGACCAACATGCCTGGCTAATTTGTATTTAATAGAGACAGGGT
 TTTGTCATGTTGGCAGGCTGGTGCACCTCATATGATCCACCTGCCCTGG
 CCTCCCAAAGTGCTGAAATTATAGGCCACTGCGCCCAGCCAAAATTCTGGTTGAAA
 TAATTTGAACTCATAAGAAGTTACACATATAGTAGAGAGAATTCTTGACCTCTC
 TGAGCTCCTATATACCAATGATAACATCCTATATACCCATAGTATATGATCAAAACTA
 GGAAATTGTGAAGATGGCATTGAGACATCAGGCAGTGTTCACGTTACTGTTTGCTTA
 CCTGGGCTTAATTTTATGTGTTTTTCAATCATTGAATGAACAAAACCTGGACTA
 GGCTGGGAGTAACTGATTGACTGTTCTGAAGCAGTCCAGGACTTATGTGACC
 GTGGTCTCTTTCTCTAG**TGATCATAACAGGGTTGCCTACACGATGGTATGCCAA**
TGAGCCTGTTTCAGATTACATCAATGCAAATATCATCATGTAAGCTTGCTTTCACAG
 TGTTTCTGACCACATACATTCTAGCTATTGTTGATTAAATCCTCCTCATGTCCTG
 AAAGTAACTTAAGGTGTTGAAGGATTCTCCTAAATTCTAG**CCTGAATTGAAAC**
CAAGTGCACAAATTCAAGCCCCAAAAGAGTTACATTGCCACACAGGCTGCCTGCAAA
CACGGTGAATGACTTTGGCGATGGTGTCAAGAAAACCCGAGTGATTGTCATGAC
AACGAAAGAAGTGGAGAGAGGAAGGTAAATCACAGAAACTCTTCTGCTAACTGTT
 TTAAAGTATCAGACATGTCAGATTGGCCATGTTAGGAATTGAATAAAATGAATAAGCT
 TACTGTAACTGATTCTGGAAAAAGGGACTAGGAGAAATTGATATGTTATCCTG
 GTGTAGTTCTTATGTTCTGCTGGGATTGTTGAGCTTCTGGCTCCATGGAT
 TTGTAGTTCTCTAAATTGGATAATGTTCAAGTCTAGTTCTCAGATACTATCCTG
 GGCTGGCATGGTGGCTCATGCCGTAGTCCCAGCACTGTTGGGTGTTGAGGTGGCGGA
 TCACCTGAGGTCAAGGAGTTGAGACCAGCCTGGCAATGTAAGACCCCCATCTTAA
 AAAAAAAATGTACCCCTGCACAACCTGTCCTAGGACAGCAGTCATACGTGTATTAGAC
 TACTGAAAGTTGTCATAGCCCAGTGTACTGTTTATTTTATTCAAGTCTTCTCC
 CGTGTTCATTCGAATAGCTCTTGTCTATGTCTCAAGTTAATCTCTGCAATATGT
 CATCCGCTCTTAATCCTATCCAGAGTATTTCATCACAGACATTGTATTCTCATCT
 AGAAGTGTAAATGTCATCTAGCTTCTTTAACATGTTGAGACAGGGTCTCGGTCTG
 TGAATGTTGAGTATTCTGTTGTTGTTGTTGTTGAGAGACAGGGTCTCGGTCTG
 TTGCCAGGCCGGAGTGCAGTGGCATGATCTAGCTCACTGCAGCCTCTGCCCTCCGGTT
 CAAATGATTCTCATGCCAGCCTCCAAAGTAGCTGGACTACAGGTGCGTGCACCCACAG
 CCTGGCTAATTTGTATTAGTAGAGATGGGGTTGCCATGTTGGCCAGGCTGGTT
 TTGGAAACCCCTGAGCTTAGGTGATCCACCTCCTGACCTCCAAAGTGTGGATTATA
 GGTGTGAGCCACCATGCCAGCCTGGCATGTTGCTGTTAATTAACTCTGCCTAACTGTCCT
 CCCAAATGGTTGCTGAGTCAGTGCCTCACTCCCAGCAGCACCTGCCTAGGACTCTTACTC
 CATACTCTCAAGACACTTCAGATTAAAAAAATTGTAACACCCCCACACCTACAGAA
 GAGCGGACAGATCTTATTGAGTGACAGCCCTGTGTTATCTCAAAGTGAGCCACCATG
 GTGGTTTTTTAAATATGAAAAGTTCTGTTGTTGTTGTTCTAGTGAAAGTT
 CTTTTTAGATACCTTAATTGGTTATATAAGATTATGTTGAGATGTAGCAGTCATA
 CCTATAAAATTAAACCTAAGGAGATGGAGAACTTGGAGTTGAGCCTCCTACTGTAATT
 TTCATATTGGATGTGAAGGGCAGTGTGATTTCATAAGACTTCAATTGTTGACTCCTAG

10

20

30

40

50

TTGGTATACTCTGAATACCTTGAGGCCAGTTCTGGTCATCGTAAACAAAGGTTCC
 TCAGCAAATGCCTGTGGAACATTAGGTGTTCTGAATTAAATGGACCAATGAAAACATCT
 TTGTTAGTTCTGCTCAGGCAAGGGTTTTGCCCTAATGTGGATAGGAAGAATGAAGC
 CCTTCATCCTCCTTTGCCCTGATTATAGCTATAGGAGGTTCACCTGTTCTCAGAAGACA
 TGAGGATTGTGAAGAGAGGGGCTTGTGTTCTCAGAGGAATCAGTATCAGTCCCTTC
 AGAAGCTCTCCTGGATAGACAGGCATTAGGGCCAATCACTCTGCCAACCCCTACCAC
 CATGTCTACTCTGCTCCCTGCTCATCTTCTCTTACTTGGTGGTGCAGAGG
 ATGACATGATGGTATTGATTCTCCACAGACCTTGACATCTACTTCAGTATCC
 CCCCAGTGCACAGAAGACAAGGCAGACTGTGACTGTGTTGATTCTGGCTTATTT
 AAAAGACAGTGTATTAGTTCTCACATTTAGAATTGTTGCAAGGTTCCACGGGAGT
 TTAGAAACTAGGGGGAGGGCTGATGTTAAAGTAGCTAAATGTTCTTTCAGGGTCA
 GATTTAATTATATTCTGGTGGACTCCCTATAGTGACTGGGAGCAGTCCTCAGTCTT
 GATTGGCCAGTGACAGCATAGAGTACAATTAAATTAGGAGTGCTCATTTGGGAAACTA
 AAATTGCACTAAATCTGTCAGAGGTGTTGGATCTACAAAATACGGGAGGGAAAGCTGA
 ATTGAGAATCATAATAAAAGACACATCGTTCTTTTTTTTTGGACT
 GTATCTGCTCTGTCACTCAGGCTGCAGTGCACTATCTGGATCACTGCAGGCT
 CCGCCTCCCGGATTCAAGCGATTTCCTGCTCAGTGCTGAGTAGCTGGGATTACAGGC
 GTGTGCACTACACTGGCTAATTTGTAATTAGTAGAGACAGGTTTACCATGTTG
 GCCAGGGCTGGTCTCAAACCTGGCTCAAGTGATCCACCCGGCTTCCAAAGTGGCTGG
 ATTACAGGGCTGAGCCACTGCGCCCAACCAAGACACATCCTTTATTGAACGTTCC
 TACCATGTTCTTTCTTCATTAAATCATTGACTCATTGACTCTACTGTTGATGT
 CTGTAGCTGCTCTTATTCCAGTTTATAGCTGTAATTCTGCTTCTAAGATA
 CAAGGTAATTCTTGCTGATATTGGTGGTTTGGAAAGTGAGTGGTGTGGATGACTG
 CCCAGAAAACAACAGAACACAAAAGCATTCTGCCCCAGAACACATACCAAAATAGATA
 AAAACTCATCTTACTGAGTGAATAGCTTCTTTGGCAGCAAGAACATGATTCTGG
 TGCCATATTTCATCCGCTGCTCTGAAGGCCAGCAGCTATTGAGACTTGGCATTCC
 CAGGCACCCAGTTAAGGAAAGTGACGTGAGAGGAGGTATCAGATGGTCTGGATATAG
 AAAAACAGCTGGTCAAAACCCCATGGCTGCCTTCTGTGATAGAGTTATTACACTT
 GGGTTAGATAAGGCACAGAGTCCCTACACTGGTGCAGGAAATGAAACAGACAGTCTGGC
 TCGTTGGGAGCCTAGCCTCCAGAACATCTGTGCTTGCCTCCATGGAGTGACTGGT
 AGATCTAGAATTCAAGACCTCAGTGGTGTAGCCAGCAGCTCAGTGGTGTGGCTT
 CTCTCTGCATCTTGATTCTTAGAGATAGATAAACCAAGCAGCAGCTCTGACAT
 GTGCTTGGAACAGACACCTGCACTGCAGCTGCTTCTCTCCACTCTGCTGGTCTTCC
 AACACACTGCTTTCTGTTGAACTCTTCTCTTTTTGAGACAGAACCTCTCTGT
 CACCCAGGCTGGAGTGCAGTGGCATGATCTCAGCTCAGTGCACCTCTGCTCCAGGTT
 CAAATAATTCTCCTGCTCAGCCTCCAAAGTAGCTGGATTACAGGTGCCTGCTATCACG
 CCTGGCTAATTGGTATTAGTAGAGACACGGTTTACCAATTGGCCAGGTGGTCT
 CAAACCTCTGGCTCAAGTGATCTGCCGCCCTGGCCACCCGAACGTGCTGGGATTACAGG
 CATGAGCCACTGCGCCCCAGCTGATTCTTACAGATAAACAAACATGACTCTGCTTGA
 CATGTGCTGGATCAGGTAAGTGCACAGCTGCCTTCTCCTCCACTCTGCTGGTCC
 TCCGAATGCTGCTTTCTATTGAACTCTCTGCTCTTCTGAAACACTAACAGATG
 CGAAACAGGCCATTTCATGGTGGTTATTAGCAAGACAGTGAACATTGTTGTTG
 CTTGTTAGGCTTTATTAGTCAAGGTTACAGAGTTACAGAATTAAACTTCTTTCTGATCTT
 CAG**AGTAAATGTGTC****AAATACTGGCCTGATGAGTATGCT****AAAAAGAATATGGCGTCATG**
CGTGTAGGAACGTCAAAGAAAGCGCCGCTCATGACTATACGCTAAAGAGAACTTAAACTT
TCAAAGGTTGGACAAGTAAGTATATTGCGTATTCTAGAGACTTGGGAACGTGTTGATGG
 TGTGTTAGGAATTCAAGGGTCTGCCGTTACTCATGTTGCATACATGCATGCATTGCTCA
 CTCATTGATTCACTAGCCATTATTAGCTTCTCTATGTGCCAGGTACAGTTAACAGCAG
 TACTGGTACATTGTGAACAAGGCAGGTAGTGTGCTGCCATCGAGCCTAGGGAGATA

10

20

30

40

50

GACAATTAAAAACAAATAACTGCCAGGCCTGGCTCAGGCCTGTAATCCCAGCACT
 TTGGGAGGCTGAGGTGGGTGGATCGCTTGAGCCGGGAGTCAGACCCAGCCCTGGTGG
 GAGACTGGGATAGGGTACAGCTGAGTGGCTACAAGGTCTTTAGGAGGCCTCCGCAAGGGC
 CTATGATGGCCTCCTCCAAGTATCCACAGACTCAGCAGTTGTTCTTTGTTC
 CTTCTTGAATGAAATTATATAAAATGGCAGAATAAACTGGAAGAGAACAGTAGA
 TGTGAGAGGTGCCGGGGGGTGAAGTCTGCAGGATGTGGGATTGTTGGCTTTGGAGGA
 GGAAGGAGGGATTCAAGACACATTGAGAGGTTGAGTCTGAGCGGACAGTGGTGTG
 GCAGACACCACAAAAGCTGAAAGGAGAACAGTGTGGCAGTGTGTTCTCTGGA
 TGTGTTAGCTGGCATCTGAACAGTCATGTGGACATTCTATTCAATTAGAGATATT
 TGTTCAATGACCTCTGGTCTGGCACCATGCTGCTTGAGGATAGAGCTGGGAAC
 AAAACAGATGGAATCCCTGCACTCCAAGTGTACACTATACTGGCAGTAATCTACCA
 CCAGTAATTGACATATAATATCATTATAAACTGTAATCAGGGCTAGAAAGAAAAAA
 TGCAGGAGTTAGGTTCAATTGGAGGGGAAGGGACTTTTTTTTTTGAAAC
 AGAATCTTGTCTGCACCCAGACTGGAGTGCAGTGGTCATTCAACGGCTCACTGCAGCC
 ACAACCTCTAACGCTAACAGTGTACCTCTCACCTCAGCCTCCATGAGCTGGGCTACA
 GGTGTGCCACCATGCCACCAATTGTTAAATTTTATAGAGACGGTTGTCTCATT
 TGTTGCCAGGCTGGCTTGAACCTCTGGCTTAAGCGATCCTGCTGCCACATGCAGCCT
 CCCAAGGTGCTGGATTACAGGCGTGAGCCAGCGCACCCGGCCAAGGGAGGGAGGTCT
 TAAGGCATAGGGAACATGTGTTGAGTCAGCAAAGGAGGTTGTGGGGTTTGCTCTAAG
 TGTGGAAGCAGCCAGAGTTGGATTTAGTAAAGAGATTCCCCCTCCACCTGTAGAG
 ACTGGAGGGGGCAGGAGTTGTTCTAGGGATTAGGACCAATTGGAGGTAGTGCAGCCGTC
 AGAGTAAAAAATAATAGGGATTGAACTAGGCCAGTGCCAGGGTGCCTGAAAGAAGAGGA
 CCCAGTAGAGCTGACTGGAGGAGACATGCCAGGGATTCACTGAAAGGAGTGTACCAAGGGC
 GAGGGTGGTGTGAGGGTACTGGCAATTCTAGCTTGAGAAAGGTCCGGGGGATGGC
 AGTGGAGTTGAGGAAGCTGGAGGATCAAGGACCTTTGTGAACACACAAAGTTGAGA
 TGCCTTGGACACATTGAAGTGGAGCGGTAGGGGAGGCAAGGGTGGAGGTGGATGCGGAG
 GGGAGGTGGATGAGAGCGTGTGGATGGATCAGTTGCTCGATAGAGGGACATGTT
 TTCTGTGCAACAGGAGGGCAAAAGGAGAAGGTGGCCACAGATGCCGGTAGATGAGCTGA
 GAGTGAATTGATTCCCTATCCTCGGAAGCTTGAGGCAAGGGCATTCAACAGACAATCAG
 AGGGATAAGAAGAGATAGAATATGAAGAAAGGGAGAAAGATGAAATCGTAATTGAG
 TAGCAGGGCAAGAAGTCCAGAAATTCTGTGCTGTGCCAGTCCCAAGTTGAGGCCGTGA
 ACATGAAAATATACTGATACCCATTGCCCTGGTTCTCCAAGGGACACTTGGCTCTAGG
 GCACAAAACAGAAAGTACGTGGTTGTCCAGGCCAGGGCTTGCAAGTGCAGTGGAT
 GGAGAGGGAGGTCAAGGAATGGAGGACATGGTAGAGAGAGACTGTCCCCAGAGCACGGG
 ACTCCTGGCCGGATGAGGGGACAGGGCAGGGAGGAGCAGGTGGAAGTAGAGGGAGGG
 CTCAGTGGTCTGGAGGCTACAGGAAGTGACGGGGGACAGAAGGAGCTGGAAACCAGTG
 TGGTTGTGGCCAGGGTGGGATGTTGGATTCTGATGTCAGAGAGGTCAGTCTCT
 GATGATGGGAGGGTGGAGGTGAATCTATGGTAGAGATAGTGAGAGGAACACTGGAACAA
 TGTAGCTGCAAGTGGAAATGGGAGAAAGGGCTGGCTGGCTCAAGCCTGTAATCCC
 CAGCATATTGGGAGGCTGAGGAAGAGGATCGTGTAGCTCAGGAGTTCTGGCTGCATT
 GAGCTGTGATTGTGCCACTGCACCTCCAGCCTGGCAACAGAGTGCCTGAGTTAAA
 AATAAAATAAAATAAAATAAAAGAAGAAGAAAAAGAGAAAAGTGTCTTT
 TTACATCCCTTTAAAAATGCACTTAAGGCTGGCAAAGTGGCTCATGCCCTGTAATCCC
 TGCACCTTGGGAGGCTGAAGTGGTGGATTACTTGAGGTCAGGAGTACAAGACCCAGCCTG
 GCCAACATGGGAAACTCCTCTACTAAATTAGCTGGATGTGGTACATGCCCTGAGT
 CCCAGCTACTCGGGAGTCGAGTCTGAGGCCAAGAATTGCTTGAATCGGGAGGCGTAGG
 TTGCAGTGAGCTGTGATCAGGTCACTGTGCACCAGCCTGGATGACAGAGTGAGACTCTG
 CTCAAAAAAAAAAGTCACTTAGCTTAGATGTCTCTACATATATAGGAAGAAGATGTAGG
 AATGAATGGTGTGCTACAATTACGTACGGATAGACCCAGAACATGATACTTTG

10

20

30

40

50

GTTTCTGTAGCCTGGTGCATTGTTGATCTTATTAATTATCATTATCCTCAAAATAG
 CCATAATGTGCTGAGTCCTCCATTGCTGGCAGAGGCTGAGTATTCAGCGAGCTC
 ACTGAGTCCTAAATTGCAATTGATAGAGAGAAAGAGATTATTATTCGATTTGCAA
 AATGAAGAAATTGAGGTTAGAGATAACCAAGGGCCACGTGAGTGTGAGTGCCTGGAATT
 GGAGCCTAAATCTAGTCATCTGATAGCAAAGCCTGTTCTTATCTGCTTGCAATTAAAT
 ATAAGTTAAATAGAACAAACTGGCCAGGCTGGGTGGCTACGCCGTAAATCCCAGCA
 CTTGGGAGGTCAGGCAAGGAGATCACCTGAGGTCAAGGAGTTGCAACCAGCCTGGCCA
 ATATGGCAAAGAAACCCCCATCGCTACTAAAAAATACAAAAATTAGCCAGGCATGGTATG
 TGTGCCTGTAATCCAGCTACTTGGGAGGCTGAGGCAGGAGAATGGCTGAACCCGGGAG
 GCAGAGGTTGCAAGTACGCCACTGCACTCCAGCCTGGCAACAGAGTAAG
 ACTCTGCTTGGAAAAAAAAAAAAGAATGATACTATAGTCTGTGTTATATGGTGG
 GGAAGGTTGAGTATCAAAAAAATACAAAGAGGAATGAATGCTTAAGTGAATGCCTGTT
 TCCCCATCTGCTTCTCTGCTGGGAGGAGACCTGGATCCCTAGAGGTTTCAGTTG
 CCTCCAGAGCTGAGTGCACAGGGATGCAGGGAAATAGGGATGTTACCTGCGCTGGTAA
 TTCAGAGAGATGATTCAAGGTATAGTTACCTGAAAGAACAAATTGCATGCCAGACGTCT
 TGGTTCTTATGACAGAGGCAAAGAGTTGCCTCCAGGATTGCCAAAAGGAGACGAGTTCT
 GGGAACCTCACGAAGAGGACCTTCAGTGGAAACCTGGGAGATTCTTCTCCATTG
 GATTTAGGAAAGCTAGAACCGGGTGATTCTCAACCTTGTGATTTATTAATTCTTTCT
 TGGTTTTCTGGCTCTACTCAGGGAAATACGGAGAGAACGGTCTGGCAATACCACTT
 CGGACCTGGCGGACCACGGCGTGGCCAGCGACCCCTGGGGCGTGTGGACTTCTGGAG
 GAGGTGCACCATAAGCAGGAGAGCATCATGGATGCAAGGGCCGGTGTGGTGCACGTGAGG
 TGACAGCTCTGCTGCCCTCTAGGCCACAGCCTGTCCTGTCTCTAGCGCCAGGGCT
 TGCTTTACCTACCCACTCTTAGCTCTTAACTGTAGGAAGAATTAAATATCTGTTGAG
 GCATAGAGCAACTGCATTGAGGGACATTTGATCCAAAGGCATATTCTCTAGACCTA
 CAGCACTGCCATTGCCATGGCATGGCAACATGCTCAGTTAAACAGCAAAGACTAAGT
 CAGCATTATCTCTGAGTCCACCAAGGTTGTCATTAACAACCTCATCCTGGCTCTGCA
 GTTTCTCCTTATTCTCATGATGTTGCTTGTAGCTGTTGACTGCTTGTAGGTATTGA
 GGTGGTGGGGGTGTTGAAATAGGCCTGACTCTGAGGATCCCTAAGTCATTTGCT
 TTGTTCTCTTTCTCTTCT
 TCTGTTCT
 GTGCCAGGCTAGAGTGCAGTGGTGCCTACTGGCTCACTGCAACCTCCGCCCTCCGG
 TTCAGGCCATTCTCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCATCTGACACTA
 CGCCCGCTAATTGTATTAAATAGAGACAAGGTTGTCATGTTGGCAGGGCTGGT
 CTCGAACCCCTGACCTCAGGTGATCCACCTGCCTGCTCTCCAAAGTGTGCTGGGATTACA
 GGTATGAGCTACCATGCCGGCCATTCTGTTCTCTTACCATATAATTCTCCCC
 TAACACTATATTGTTGCTTCACAAGATTCCAGTGTGTTCCACCAAGGCCTTGATG
 GAAGCTGTGCTGTGACCTCTGTAATGAGTGTGTTGGCTGATTCTCCAGTTGGCCTT
 CATGATTATACTGGGAATATTGGTTCTAAATCTCATTCTTCTGGCAAGTAGA
 TATATGTGAAAGTGTATTGTCCAGTTGTTAAAGAAGCTACCATTTATTGAGGCCAGCC
 TCTGAGCACAATTGTTGTTGTTGTTGTTAAATTAAATTATTACTTCTC
 TATTTCATAACTTATTATTATTATTGAGACAGAGTCACACTGTCACCCAGGC
 TAGAGTGCAATTGAGCGATCTAGCTCACTGCAACCTGCTTCTGGGTTCAAGCAATT
 CTCATGCTCAGCCTCCGAGTAGCTGGGATTACTGGTACGTGACAACATGCCCTGGCTAA
 TTTTGTGTTTAGAGAGACGAGGTTGCTATGTTGGCCAGGCTGGTCTGGAACCTCC
 TGGCCCAAGTGATCCTCCTGCCCTGGCCCTCCCAAGTGCTGGTATTAGGTGAGAGCC
 ACTGCGCCGGCCCTTTCAGTAATTGATGTTGTTGATGATTCTGTT
 ATTCTGTCACCCAGCACTGTTGAGGACTTACCTCAGGCACGTTGACTGGGAGAGCTGAA
 AATTAAGGGTCAGAGAGGTTAAGGGACTTACCTCAGGCACGTTGACTGGGAGAGCTGAA
 CTCCAAGAGCAGGTTGGCTGACTCCAAAGCCCTATGCTTTGCAACATATTCAA

10

20

30

40

50

ACATAAAATAGACAATTTATAAATAGCTCAAAGAGTAGACATTGTTCTGTTGATATTA
 ATGGCTGGTTTGAGTCTGAACCCCCATGAATGATTCTGTTGCCCTGCTTTGTC
 TTCTGCCCGCAG**TGCTGGAATTGGCCGGACAGGGACGTTCAATTGTGATTGATATTCTTAT**
TGACATCATCAGAGAGAAAGTGGGTCATCTGGGGCAAGAAGCGACAGTTCTGTTT
 TAGTTTATGGAAGGAAAGTGCTCACGAAAACAGTCTGGGAAGAGAGGTTGAATGGGAAA
 ATTCTTCACAAAAATCTGGCTGAAGACTTCAGTGTCTGCCTGAGAACAGAAAGTGAC
 ACTATTGAGCTTGGCATAAAATGAAGTCTAGGAGCTGCAGAACCCACTGCCATGGC
 TTTTGTGCATACACAGTGGTGGTCTATCCAGCCACCTGACCTGTTACAGTATGGG
 GTGATTGTTGGCAAGTGAGGAACTCTGACTCTGCCACTTCGTTATTATGAGTCTT
 CTGGGATCATTGGTATTGGTCAAGGTTAACACTGTAGCCATTGCAACATGCTCAGTTA
 AAACAGCAAAGACTAAATTAGCATTGTCCTGAGTCCACTAAAGTGTGATTAAACAA
 CTTCATCCTGGCTTGCAAGTTCTTTATTCTCATGATGTTCTCGTAG**GTGTTGA**
CTGCGATATTGACGTTCCAAAACCATCCAGATGGTGGGTCTCAGAGGTCAAGGATGGT
CCAGACAGAACGACAGTACCGATTATCTATATGGCGGTCCAGCATTATATTGAAACACT
ACAGCGCAGGATTGAAAGAGCAGGTACACGCTGAGGGCTGGCATGCGGATTCTCATT
 CTCTGCTAGGCCTTGGATACGCTCTCCTTTGAGCAGGAGGACAGGCTCTGATAGAC
 AACTGTTGATTGGAATGGGAAACAAACTCCAACTAAAGGGCCTCTGAAACTGTC
 AATTATCTCCACTCTCAGCTGATTTCACTGCAGAGGAGCTAGGGAAAGGGCACC
 ATCCTATCAGCCTGGCCTGCCAGATTGAAGAACTGCCATGCAGAAAGGTTCTGATGTTCT
 CAGGCTATGTCGAAGCGTAAACACTCAAAGCCTGAAAGTTCTAGCCTGTTCCAGCCTT
 GATCCAGGCCATGTTATCCTGATTCCATCCTTAAACGAATGCCACTCTTAATAGC
 GCACGGCAGTTGAACACTAAATTGGTCGAGTTGGAAACAGTGAATTTCAATTAAAT
 AAGCTGTGCATAATGAAGAGGAATGTTGAATTGGACCTTCCATCTGAAGCTATTCTA
 ACAGGCACAAAGCTGAGTTAATTAGGAATATGCTGAGATGAAGGAAATGAGGAGAGCTGC
 TCTTTGGGGCTGTGCTCTCCTCCAAACCCCTCAACCCCATGCTGAGATGG
 GGTGGTGTCTAAACATCAGTGGCGAGTGCCTGCATTACTCTGCTCGTTGCCTCCAGAGA
 ACTCAGCTCTCCAATGCTGAGCTCTTCTAGAATGGACCTGCCACCAAGTATTGAAA
 GATTCTAGCCTAGCAGAACAGCAGCCACGTTACAAAGTTGGTGGCAAAGGAAGGT
 ACTTGCTAATTAGTTAGTAGGTTTCAGTCCGCACAGACATACTGGGATTGTTATTGT
 ACATAGACATCTCAGAAACAGTGTATGTATAGAAATGTAAGGTCAAATTGAAACCTCA
 GTGCTTAAATCTGAATTGTTACTGATATGAAATTAGTAGCAGGTTACTTATT
 ATATCTGCTCCATTATACTTAATTGGCTCAAGAACATAGTTAGGCAAAGTTGCCAA
 AGAGAAGGATCTCTAGTAAATACAAAGAACATGTAACATAGTTGCTACAAGTTGGAGCA
 TGTTCAAGGATGTTTTTTTTTTTTTTGAGAGAGAGGTCTCTCTGTTGCCCA
 GGCTGGAGTGCAGTGGTGAATCATGGCTACTGCAGCCTCAATCTCCAGCCTAAGCG
 ATCCTCCCACCTCAGCCTCCAAAGTAGCTGGACTATAGGCATGCCACACACCTAGC
 TAATTTCGCTTGTAGTGTACAGTTGCCATGTCAGTGTAGTTAGGACTACAAATTG
 TGGCTCCATGCCGCTAATTGTTATCTTATTGTTAGAGACAAGGTCTCACTGTT
 TGCCCAGGCTAGTCTGAACTCCTGGCTCAAACAAACCCCTCCACTTGGTTCCAAAG
 TGCTGGATTACAAGTGTGAGGCCACTGAGGCCAGTGACCTCTGGGTTAAAATGTGTA
 GGCTTCATTATTAAAAATGAAATCCTGCAATATAGTTCTGCGTTGTGTT
 GGTTGAATCAACTGGAACTGGCTTGCTGGCTGATTGTTAAAGTAAGAACAGTACTTA
 ATTTAGTAGAAAGTTAAATGGCAGACATAACATTAAACCCAGCTGATTATAATGAAG
 CAAAAGAACAAACTCATTCAAGGATAATTGGTATTCTAAAATACAGTCATTCTAAAAT
 TATGAAGTGTTCAGGACCTTGGGAGTGAAGAACATTGCTAAAGAACAGTGAAGAAA
 AAGGAATGATGGGAGAGAGCTGTGGAGAACAGAACAGCACAAGGAAGGAAG
 AATATAAAATCAGATGTGGGAAATCCAGGGAAAGTGCACAAACGAAGAACAGTGAAGAAAAT
 TCTCAAGTTTATAAACAGTCTCACACTCTGCCAGTTCCCTGGAGGTAGACTTTTG

10

20

30

40

50

TTAACCTCCA ACTACAGTAGT GAAAAAAAAACCTCAAATTGCAAAAGCAGTC
 TGTGGAA TTTCTTACCCAGCTTCTGACTGTTAAC TTTAGCACACTTAAC TTTAT
 CATT CGTT ATTCTCTGT TAAAATTAAAATGTAATT TAAAAGTAAAATGTTG
 TTGGTTACA AACATTATA CCCCTTGCTCTAAATATCATTTCAATTAAAAAATGAAT
 AATCTAACGCCTACACATTCTAAAATGTTATATTCTAAAATAAGGGCATTCTTAC
 ATAACCAATGTCACAATTATTGATACAGTGATCAAATCAGGAAACTAACATTGATATA
 ACACATTATCTAACCTACAGACCCTTCAATTGGCTCTGCTAGTATCTTTATGGG
 TCCAGGGTACACAGTGCATTGGCTATAATGTTATCTTTCTCTTTGAGACAGG
 GTCTCACTTGTGCCAGGGTGGAGTGCAGTGGTCAATTATGGCTACGGCAGCCTTG
 ACCTCCTGGGCTCAGGTGATCTCCCACCTCAGCCTCTCGAGTAGCTGGAGACACAGG
 TGTGCACCAACCATGCCCTGGCTAAGTTGTATTGGTAGAGATGGAGCTTCGCCGTGT
 TGCCCCGGCTGGCCTTGAACCTCTGGCTCAAGTGACCCCTCCGCCTGGCCTCCAAAG
 TGCTGGGATTACAGCGTGAGTCACCACACTGGCAGTTAGTATGTTAGTCTCTT
 TAATCTGAAACAGTTCTAGTCATTCTTATTTTCACTGACCTGGATGTTTGAGAG
 TTTAGGCCAGCTATTAGCAGAATGCCCTTCAGTTGGATTGTCAGTGTCTCTTG
 ACTATATTCTAGTCATTTGGCAGGACTGTCACAGAAATGTTGTTGAGTCTTCT
 TAGTACATCACATCAGGTACACACTGTTGATCTGATTACTAGTGGTGTAACTTTG
 ATCACTGAATAAGTGGTGTCTGCAAATTGTCACCGTAAAGTTACTTGAGCAAAAC
 GTAGCTGGACTACAGCGTAGCAAAATGTAAGCAAAAGTAGTATTTTGCTACATT
 TTTTTAGGAACAAAGTATTCTCCCTTTAAGTTAATCTCTGTCATAAAAGTTATT
 TTTCTCCCTTTAAGTTAATATCTGTTGGTAGATACTGGAGACTGCGTAAATTACCTAT
 TTCTCATATAACTTTTTTGAGATGGAGTCTCGCACCGTCTCCAGGCTGGAGT
 GCAGTGGTCAATCTGGGTCACTGCAAGCTCCACCTCCGGGTTGACGCCATTCTCTG
 CCTCAGCCTCCCAAGTAGTTGGACTACAGCGCCGCATCACACTGGCTAATT
 GTATTTTAGTAGAGACGGGGTCTCACCGTGTAGCCAGGATGGTCTGATCTCTGACC
 TTGTGATCTGCCGCCCTGGCCTCCAAAGTGTGGGATTACAGATGTGAGTCAGC
 CGGCTCTCATATAACTTTTGCTACTAATTATATTCTGATTAATTCTGCTG
 AAAAATTATTACTGTGTATTGCCAATGGCAATTCTGTTGACCTGCTGGAGT
 CCCGCTTTAAAGTATAAGTGACAAAGAAAAACTGTATATAAAAGTGTACACCATGATA
 TTTTGATATATGTATACTTTGTGAAATGATTCAAAATTGAGTTAAATAATGCATCAA
 CATCTCAGTTACTTTTTTTGAGACAGAGTCTGGTTGTCACTAAGGCTGGAG
 TGCAGTGCCACAATCTGGCTCATTACAACCTCCACCTCCAGGTTCAAGTGATTCTCCT
 GCCTGGCCTCCCAGTAGCTGGGATTACAGGTGCCACCATCACACCCGGCTAATT
 GTATTTTAGTAGAGGTGGGTTCACTACGTTGGCCAGGCTGGTCTCGAACTCCTGACC
 TCAAATGATCTCCGCTCAGTTCAAAGTGGTACATTCTGTTGAGATCTGCCACACTGGAACCTCT
 GCCCTCTGGGTTCACTCAGGTGCTCTCCACCTCAGCCTCCCAAGTAGCTGAGACTACA
 GGTACCCGCCACCATGCCCTGGCTAAATTATTGTATTAGTAGAGATGGGGGTTACCAT
 GTTAGGCCAGGCTGGCTCAAACCTCTGACCTCAGGTGATCTACCCACCTCGGCCCTCCAA
 AGTACTGAGATTACAGGCATGAGCCACACCCAGCCACATTACGTTAGTATTAACTAT
 AATCACCACATGCTGTACATTAGATCTCAAATGTTATCTTATGTAACTTCAAGTTG
 TACCCCTTGACCAAAGTCTCCTGTTCCCTACCCCCAACCCCTGGTAATCACTGCTTT
 AATCTCAGTTTTATGAGTTGACTGGTTAGATTCCACATCAAATGAGATCAGGCAGT
 GATGGTTATTCTAGTCATAATGTCTCCATGTTCTGCAAATGACAGGATTCTTCTT
 CTTTTAAACTAATATCCATGCTGGACACGGTGGCTATGCCGTAAATCCAGCAGCTT
 GGAAGGCTGAGGAGGGTGGATCACTTGAGGTGAGGAGTTGAGACCCAGCCTGGCCAACAT
 GGTGAAACCCCATCTACCAAAATATAAAAATTAGCTGGATGTGGTGGCGCACACCT
 GTGATCCCAGCTACTGGGACACTGAGGCAGGAGGATCGCTTGAACCCGGGAGGCGGAGG

10

20

30

40

50

TTGCAGTGAGCCAAGATGGTGCCACTGCACTTAGCCTGGATGTTGATGTTGTTCCACTT
 GTTTATTTTATTTGTTCCCTGTGCTTTGGTATCAAACCTAAACATTGCCATGA
 CCATTGTCATGTTACCTTCCCCATATGCTTCTTCTAGAACCTTAAGGTTCATCATT
 CTTTCTGTTTAGTGCAAGCCTACTATAAGGAAGGGCTTTCTTCTCCTTATT
 TTTATTGTCATGCTATCAGAATGGCACCTACTACTATTTGTTGTTATTGCTGATT
 GACTTGAAATTGGCTAGTGGAAACCTTTCAGATGGGACTCTGTCCTTTGATCT
 TCCATTTCAGCACTCTTAGACTTAAGATGGTCAAGGCTCATCTCCTCCTTCCCAG
 CCATTTCAGCAAGGAAACCTGATTCTTCTAGTGAAAGAGCAGTATTGAAACCAAGATCT
 GGGCACTGGGCTACTTGTTGACTGGTACAGTGTCTTGAATTGCTAATTAGCTGAT
 CAATTACTGCTCTATTGAGTCCCTTCTAAACCTCACATATGTGACAGACGGC
 CCTGACTTATGATGGTCACTTGTGATTGTTGAGTGGTTGAGAGCAATACA
 TCCATTCTGTTTCACTTCAACACTTATTAAAATAGGGATTGAGATGA
 TATTGCCACGTGTAAGTGTCTGAGCACGTTAAAGTAGGCTAGGCTAA
 GCTGTGGTGGTAGGTTAGATATGTTAAATGCAATTGACTAGTGATATTCAAC
 TTATGAGTTATTGGATGATCCCCATAAGTCGAGGAGCATTACATATCTG
 TATAACAGAGTGAGTCTTACCTTCACTTCAACACTTATTAAAATAGGGATTGAGATGA
 TAACCATGATACTTATCAAACATAACATCAACATACATTGCTATTAACTAAA
 CTAGAGTTAATTGGATTGCACTTCAATGAATATCCTTTCTGTTCTGATC
 CAATTCACTGGTCACACACTGAGTGGTCACTGTCAGTGACTCTCCAACTGCGA
 CAGCTCTTAGGCTTCTGTTGACTCTGACGATTTAAGAGTACTGGC
 AGATATCTGAGGATATCCCACAACCTGTGTTAATCTTATGTTCTCATGATTAGAC
 TTGAGTAATGGATTGGAAAGAATACACAGAGGTATATTGTTAAGTGGTCTCATCAC
 TTGGAGGTAATGTTATCAACATGGCTGGTCACTTGTCAAGTTGTTAGTT
 GTATCTGCCAGATTCTCACTGCATAATTACAAATCCTCTAACATTATGATGGGTT
 ACAGCCTGATAAGCCCACATCAAATTGAAAATATCATAAGTCAAACATTAATGCA
 TCTAAACTACTAAACATCACAGCTAGCCTAGCCTGCCTGAACGTATTCAAGGACACTTA
 CATTAGCCTACAGTGGGAAACATCTCATGGGAAGGCTGTTATAATGTGTTGCAT
 ATCTTATGTAATGTTGACTCTGACTCAGAATGAAAACAGAAGGGTTGTATTGCTT
 TTGCAACCATCAAACCTAAAGGCAACCATCATGAAGTTGGGACTGCC
 TGTACTTTCTCTTCCCTGTTCAATTCCCTGGAAGAAAGTCATTTAGTCAGACCA
 TACTCAAGAAAAGGAAATAAGCTCCATCTTGAGCTTAATTGAAACTGGAAATGACT
 AGTTCTATATACATTATTTAGAATCCTTGTAAAGAAAGATTGTTCTCTCCATT
 TATTTATTCCATTATTTATGATAGAGACGCATGTACATTATTTATACTTGGGTT
 ATAATCTATTCTGCTCAAATTGTTACAGCTTGGTCACTGGGAGGTTCTCAGATT
 GGCTCTGTGTCATTGACATGCCCCACCTCTGTTCTGACTACTCTACTTTGG
 CATTACAAAAGATGTTCCAGGCTCTTATATTCCCTGCCGAGCCCTAGAATCAT
 CCATTCTATGGCTGGCTTACTGGAAATCAGCTTAGGCTCTCAGCAGATAGA
 GGTGTTGGGTGTGCTATTGCACTGAAATCAGCTTAGGCTCTCAGCAGATAGA
 GCTAGAAAACATGGCTGATATGAATCCATGGATTGATGATGTTCTCAGAC
 TGTACTGGCATCTATATATATTAAGCTAACATGAATTGATGTCAGAC
 TCGAACATTGCCAGGGCTCATTCTGCTCTTGTGTTATTGACTCTCT
 TCTAACAGGGAGAAACCCAGTCTCATTATCACCAACCTATCTACTCATTGTCACCC
 TGGTAGGGTGTAAAGTAGTTGAGAATTACTAACCTATACCATGTGAGAATTGTT
 GCACTCTGGTTGAAGGAAATACATACACAGGTAGCGTCTCACACTTCAGTATAC
 AGAGATCTGAACAGTGTCTCTGAGTGAATCATATTGAGGACAGAAATTACTTTAA
 AAATTCTGTAATGGGTCAAGGCTATAATCTAGCACTTGGGAGGCTGAGGTGGCAGAT
 CACCTGAGGTCAAGGAGTCAGGAGACCAGCCTGGCCAAATGGTAAACCCATCTCACAA
 AAAATACAAAATAGCCAGGGTAGTGGTGTGCTGTAATCCAGCTACTCAGGAGG
 CTGAGGCACGAGAATCACTGAAACCTGGGAGGCAGAGCTGAGTGAAGATTGAGC

10

20

30

40

50

CACTGCACCTCCAGTCTGGCGACAGAGCGAGACTCTGTCTCAAAAAAAAAAAAAAA
 AATTCCATAATGATAGCAGAGCTGGAATAGAAAATGGGATTGACACAGGCTGAATCTGAGTT
 GTTGCAACAGTAAACGAGCAAGATTAACTGGCTTGTAGCACTGCTATTGGCTC
 CTATCATTTATAGACGTTATTCTTTGTTGGTGCATTCCCTTGAGAAATATT
 GAGTGCCTTCTGTCAGACATTGATTAGATGCTGAGGTTGTAACAATGAAGAAGATA
 GCCATCGCTGTCCTCATGGAACGTGAAGTTACTAGATGTAACATGAGTTAACATG
 AGGCCGTTGCCCTATGTGCCCTATTGTTCTCACACAGCTCCCTCATCTCCTGGTCC
 AATGAAAAGGTTTTCTACTTGTTCATTGAGGAGATGGAGGAGGGAGAGGTTGAGCT
 TGTGCCAGGCAGTGGGAATATTAGTAGTTGTGTTCTGAATTGAAATGAATCCAGCA
 TGTTGAGTAGAAGGAGCTGGGGCAATGTGGAGTGTGATGGGAGATTGAAAGTA
 AGCTGAGACCAGATTTTCAGTTGGAGGGAGGGAGGCTGGCCTGTAGGCCATATTACAGA
 TTGTTAGACTTTATTGGAGGGACATGGAAGTCATTGAGGAGTCTGAAGCAGGGAAATGAC
 ATAAAAAGATCCTCATTAGGCCGATGGTGGCTCACGCCCTGTAATCCCAGCACTT
 GGGAGGTTGAAGTGGTGGATTGCTTGAGGCAAGAGTTGAGACTAGCCTGGCAACAT
 GGTGAAACCTGTCTCTATCAAAAATACAAAATAGCTGGCATGGTGGCTCACACCTG
 TAGTCCCAGCTACTTGGAGGCTGAGGCATGAGAATGCCCTGAACCCGGGAGGAGAGAT
 TGCAGTGAGCCGAGATTGTCCTGCAGCCTGGTGACAGAGTGAGACTTCGTG
 TCAAAAAAAACAAAAACCCCTCATTTGAAAGGAAACCTGGCTTGAGGGTAAGAA
 TGGGTGGGCACTAGGCTAGAGCAGCTGCAGGGTCAGTGAGGAGCTGCCGAGTGTG
 GTGAGAACCGCTATGGTTGGCAGGGACTGACAGTGAGCACAGAGCGAAG
 TAAAACAGCAAATTTCATGATTGGATAGTGGAAAGGATCATGGTTGTAGTCTTCA
 AATGTGAACCCAGAGTGCAGTGACAAGTAGTCTAGGCTGCTGTAAACCAAGGCAAGTG
 TTTTCATTTACCCCTCTTCTGCTTGGATTGGATTTGTAATTAAAGGTTA
 TGAATGTAATCAGTTACTTAACATGGAAAGATACTTAATACCAGATGATTTGGAGTCTT
 GTGATCAATACCTCTCAACTTGGGTGTCAGTTGGCAAGGCCATAAATTG
 TATAAACATTGCAAGGCTGGTTACTGTCAGTGACGTTGAGTGAATTGGTGGAGATAGA
 TCAATTCAAGTTGATTTCTAGGCTTCAGAACACATTACCCCTACTCCACAAACACAA
 ATCAAACAAAACAATCCCTATCCCTGAGCATTCTCTGATCTATAACACAGCCTGG
 CTGTCACAGTACTAACAGAACAGCCATCTGAGTTGAGTCAGTTTATTCTGGCTTC
 TACATAAGCTAAAAGTTCAACATTAAATGCTTCTGGATTCTGGAGTCATTG
 AAGTAATTCTGTTCAAGTTGACTAATTATCCACACTAGAAAATTCTGTTAATC
 TTTATGATTAATAGAAATACTGATTTTATTCAAGGAAGTATTGAGTAGGGAGGGGG
 AAATAGGGATTGCTGTTCAATGGGTATAGAGTTCAAGTAAACAGAACAAAACCTCA
 GAGATCTCTATACAGCAGTGGTATATAGTTAACAAACTGACATCTAACAGTTG
 AAGAGGGTAGATCTCATGTCAGTGTGTTTAAAATTGCTTTAAAAGTATCGAGTA
 AAAAGCAGTTTACTCCTCAGTTCTATTATTTAAAATTTTTATTAAAAGT
 TTGAGATTTAAACCTCAGGATAAGTTTATTGTTAAAATTGTTTAA
 TTTGAGATGGAGTCTCACTCCATCTCAAGTCACCCAGGCTGGAGTCAGTGGTCTTGG
 CTCACTGCGACCTCTATCTCCAGGTTCAAGTGTCTGCTGCTTCAGCCTCTGAGTAG
 CTGGGATTACAGGTCTGCACCACACGCCCTGGCTAATTGTTGAGAGATGG
 GGTGTCACCATGTTGGCAGGGTTGCTGAACTCTAACCTCAAGTGACCACTGCCTT
 GGCCTCTCAAAGTGTGGATTACAGGTATGAGCCACAGTGCCGGGATAAGTTTA
 AAATAATATTCTGCTGGCTGGCATGGTGGCTCATGCCGTAAACCCAGCACTTGG
 AGGCTGAGGCAGGAGCATCTCGAGGCAAGAGTTGAGACCAGTCTGGCAACATAAT
 GAGACCCCTCTCACAAAAAATTTGGCTGAGTGTGAGCATGTTCTGTAGC
 TATCAGGGAGGCTGAGATGGGAGGATTGCTGAGGCCAGGAGTTGAGGCTGAGCT
 ATGATTGACCAACTGCGCTAGTCTGGGTGACAGTGTGAGGACCCGTCTTAA
 AAAAGGCCAGGCACAGTGGCTCAGGCCTGTAACCCAGCACTTGGGAGGCCG
 AGGCAGGGTGGATCACTTGAGGCCAGGAATTGAGACCAGGCTGCCAACATGATGAAACC

10

20

30

40

50

CCGTCTCTACTAAAAATAACAAAATAAGCTGGGTGTTGGTCACACCTGTAATCCCAG
 CTACTTGGGAGGCTGAGGGAGAGAATTGCTGAACCTGGGAGGCAGAGGCTACAGTGAGC
 CGAGATCACACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAACAA
 CAACAAACAAAAACCAAATGTTCTGCCAATTCTCCATTAAATTAAATTGAATT
 ATATTGTATCTTCTAAGGATTGTTCTTATAAGCAAAGATTTCAGTGCTAAACAT
 TTACGACTGCTATTCAAGGATGGTTACAAAGTCTTTGTTAAGAAAATGGCTGT
 TCAAAAATTTAAAGTAGTATATAACCAACAAAATTTTGCTTGGATGTCTGTT
 GCAGCTTCCCTACACTATAAGTCTTACTGACTGCTTATCACTTAATAAATTGGTT
 TGGCTACTTAAACAGAGGCAAATAGTACAGGCAAAAATTATTTTATTTTATTT
 TGAGACAGTCTCACTCCATCACCCAGGCTGCAGTGCACTGGCCTGATCTGGCTCACTGC
 AACCTCACCTCCCAGGTTCAAGCATTCTCATGCCCTCAGCCTGAGTAGCTGGAATT
 ATAGGCATGCACCACACTCAGCTAATTGGTATTAGAGACAGGGTTTGC
 CATGTTGACCAGGCTAGTCTGAACCTCTGACCTCAAGTGATCCATCTGCTTGGCCTCC
 CAAAGTGTGGATAACAGGCATGAGCCACCATGCCAGCCATTATTTTATTTT
 GATGGGCTCGCTTTAGAGATGGGTCTTGGTGCCTAGGAGTAGCAGTGGTGC
 CATAGCTTACTGCAGCCTGAAATTCTGGGCTCAAGCAATTCTGCCTCAGCCTCCCG
 AGTAGCTGGGACTACAGGCCTGCCCCACAGGCTGGCTGTACATTAGTATTGATATG
 GCTACCCCTAACGGCAATCCTATAGTGAAGTCAACATTAGATAATGATGCTCATCTGATGG
 ATTAGATTTCAAGAGTGGCTGTTCCAGGTGCCTAGGAGTAGAAAAGGGTGACA
 CTCCTAACTAGATGCTTACCAAAATAGTCACCCACATCTGAGATGAGACTGC
 CTGCTGGTTCTTGCCTTCCCCCAGGGTATCATCAGAACCAAAATAAGTTT
 AAAGGTGGGTCAGGTGTTGGCTCATGCCCTGAAATTCTAGCAGCTGGCTA
 GCAGGTTGGATCATCGAGCTCAGGAGTCAAGACAGCCTGGCTAAACATGGTAA
 CCCATCTCTACTAAAATAACAAAAGTTAGCTGGGCTGGTGGGCACCTGTA
 GCTACTCAGGAGGCTGAGGCATGAAAATCGCTTGAACCCCAGAGGGGGTTGCAGTGA
 GCCGAGATCATGCCACTGCACACTAGCCTGAACACAGAGCAAGGCTGTCT
 AACAAAATGGTGCAGAGTCTTCCAGGGCTGAGGGAGATA
 AACAAAATGGTGCAGAGTCTTCCAGGGCTGAGGGAGATA
 TTTTCGATAAGAGTGTACCATCTTCATTCTGTGTC
 TTTGTTTCGGAGTTCCCTTCCAGCTGTTGCCTGAAACCA
 TGAACATGAAATGATTAGATGCTAATGTGGCAAAGAAAGTGT
 CAGGACATTGGTCAATGTATCAGTATTCTCAACCCGTCT
 TCTTTTCTTTGGGAAAATACCTTTGCTTTTATCCC
 TGTCCTCACATGTCACTCTCAAATTCAAG
AACACAAAGAGGAAGGGCACGAATAT
ACAAATATTAAGTATTCTCTAGCGGACCAAGCAGTGAGTGAGATCAG
ACAGCCCTCTCCGCCT
TGTACTCCAACGCCACCCCTGTGAGA GTAAAGTAGTGCTGAAGGAAATTCTTTTACCTGG
 TCATGGGTTAAAAGGTTAAAAACAAAACAAAACAAAACACAAGTTGTAGCA
 CATGCCCTTCACTGGTGCACGTTCTGTTGCCCTACTGTTAGTGATCTGT
 ACTGGTGA
 TATCTATTGATTGTTAATGCTATCTAACACAGTTAA
 TTTCTCAAGCTGGCTAACCCAGTGTGTTGGGAGGCCAGGTT
 CAGGGCTAAGCTGTAATCCCAGTGTCTGCT
 GAAGTCAGGAGTTGAGACCAGCCTGGCAACATGGT
 ACAGGGCTGAG
 ACAGGAGAATCGCTGAACCCAGGAAACGGATGTTGCAGTGAGCC
 GAGATCATGCCACTG
 CACTCCAGCCTGGCGATAGAGTGAGCCTGTCT
 TCCTAAACTGAAGGCTGACTGCTATGCTAGGATTATGG
 AGTGGTGGTTCTCCAAGAAGAATCTAATT
 TTTGATGGCTGGGATTGTAACAAA
 GGAAGGT
 GACCACTAATGTGTCAGCCCAGCCATGTTCTGCT
 CATTGCCACCA
 GAG
 GAA
 ATCT

10

20

30

40

50

ACTTTGGGTGTTAACCCAGGAGTCAGCAAACACTACAGCTCACAAAGGCCAGATGTGGCCA
 TGGCCTGTTACTGTATGGCCTGTTAATGGTTAACAGGGTGTAAAACAAAAGAACACAA
 AACAAAGACCCAATAACAAAACAAGCCCAGAAGATAATATGCGACAGAGACCAGTATG
 GCATATAGAGCCTAAATACTGACTCTCAAGCCCTCCCAGAAATCCTCCGACTCCTT
 GTTGAACACCGTAGGAAAGCATTGTCAAATTGAGGATATGAATAGCAATTGTAAGTT
 ATTATTTCTATATATTGAAAGTCAGTGCTAGTATAACATTACCTTTATTTTCC
 CTAAGAATCTTCTCTGTTGCTTCGACATGGATTAAACCTTACCTTTAATTTAAT
 ATTCTATATAATGTTAGGGCATATATGAGGTTGTATTAAACATTGCTTCTATT
 TAACATTGAAATGAAATTACAGCAGAGGTATTTCTCGTCAAGTTGCCACTTCTTC
 TATCTTTCTTCTTCCAGTGGACTGCCTGGAAAATTGATATTAAATTGCTC
 TCTGCAATAATTGCAATGAACTGGAATGCCAGGGTCTGAGTCCTGCCAGACAGCTC
 GTCCTCCTGTTGGCATGACTGAGTCAGTGTATGATTCCCTCAGTACCGAGTGGCATGC
 CTGTCAGACAGCCTGTCGCTCCCTCATCCCGTGTCTCCCTGTAGGGTTCAGATCC
 AGGATACACTGGTCTGGAGCCCTCTCAGGCTGGCACCCACAGTGTGGGTTCTTAC
 TCTCCTGGACTGCTGTATGTCATCTCCCTGTCAGCAGAAAGTGTGGGATCTTGAT
 GCTTGGCCCTCTGCTTAGGCCCTAAACCACCCATTGCCCTCACATAACCTGAGCTGG
 GGCTAAATAGATCTCATCACTGCCTGCCCTGTATTTCCCTCTTGGAGCTTT
 TGCCTGTTAGATCCTCTACTGGAAATTAAAGGATTCTATTCTATGTCATGTCATTCCA
 ACCTTCTTCAAGTGGCATCAGGAACTGCCTCATCCTACAGGCCCTTAAACAAACCTGC
 ACCTCTGCCAGACCCAGGGAGCACCGAGGACTTGAGGCTTTATTGCACTTCTGTGTTT
 TTTGAGATGGAGTCTCGCTCTGCGCCAGGCTGGAGTGCAGTGGCACGATCTGCTCA
 CTGCAACCTCCATCCTGGAGTCAAGAGATTCTCTGCCCTCAGGCTCAAGCAGCTGG
 GACTACAGGCATGTCATGACACCCGGATAATTGTATTAGTAGAGACGGGTT
 CACCATATTGGCCAGGCTGGCTCAAACCTCTGACCTCGTGTACCCACCTGGGCC
 CCAAAGTTCTGGGATTACAGGCGTAGGCCACCATGCCAGCGTTATTCACTTCTGCC
 TGTAATTATATTGCTGTATGGCTATCTCTCTCCCTGGGAATGTCAGGTCTAGGCAC
 AGGAACGTGTCTGTACCATATCTGGTGGCTACATTGTAAATGCTTTCTTCT
 CATGTAAGCTAAACAGCGTGGCTACATTGTAAATGCTTTCTTCT
CCAGATGAGAGAACAGACTGTAGAGTCTATGAAAACGTGGGCTGATGCAACAGCAGA
AAGTTTCAGATGAGAACACTGCCAAACTTCAGCACAGAAATAGGTATTAAATGCAA
GTGCTCTATTGGTAAATTGTTATATAATTGGAGTATTGAGGCTTAAAGCAGGAAGCAATTG
GGAATGTTTAGCAAAGTGTACCTAAATTGAGTTACAAACCCAGGCTCCTTCTCT
CCCTGACTTCTTTCCAAGATGGTTAGTTAGGTTAGGTTCTTAAACATTAAACATTAAAC
CACAGAATTAAATTCTGCATGAGGCAAGGCTGACTCTTCCAGAGAAATGGCTGATACT
GGTGGTAGAGTGCAGGTACTGTTCTGCAATTATTAGAGTTGGTTAGCCAGGC
TGTGCTGGGGATGATCTGTAGGGATCTGGGAAGCATGGGACTCAGCACTGGGTGGTT
GGAGTCAGGAAGCCTGAGTTCTCATTTCAGTCAGTCTGACCAACTGTGTGGCATGGG
TGCTAGACCCTGGCTGCCGACTGGGTACCGACATCCCTCCAGCTCTGCTGCTGGAA
ATTCATCTCTCCATATGTTGCCCTCCCCATCAATTACGTTTTAAGTGTGACCAAGTA
TATGATGTATGTTCTGATAAAATTAGAAACTTATCTGGGATGTTGGCTCATACCTGT
AATCCCAAGCACTTGGGAGGCTGAGGTGGCGGATCACCTGAGGTAGGGCTGGAGACC
AGCCTGACCAACTAAATAGTAGAGACCAACCCGTCTACTAAAAATAGAAATTAGCT
GAGCATGGTGGTGCATGCCCTATAATCCAGCTACTCAGGAGGCTGAGGCAGGAGAGGCAG
CGGTTGCAGTGTGCCAAGATCGGCCATTGCACTCCACCTGGGCCACAAGAGTGAAACTC
CATCTCAAAAAAAAAAAAAAAACACTAGTGTCACTGTTCTGACCAACTGTGTGGCATGGG
GTAAAGCTTTAGATATGTCCTTAAACCTTAACTTATTCTTCCCTGAGATGTGG
ACTTTCACCCCTCCCTAAAGATCAAGAACAGACGCAAGAAAGTTATGTGAAGACAG
AATTGGATTGGAGGCTGCAATGTGGTTGACTACCTTTGATAAGCAAATTGAAA

10

20

30

40

50

CCATTTAAAGACCACTGTATTTAATCAACAATACCTGCTTCCCATTACTCATTCC
 CAGATAAGAAGAAAATCATCTCTACAATGTAGACAACATTATTTATAGAATTGTTG
 AAATTGAGGAAGCAGTTAAATTGTCGCTGTATTTGCAGATTGGGATTCAAATTCT
 AGTAATAGGCTTTTATTTTATACCCCTAACAGTTAATTGTTCT
 CATTGTTGGGATGATGAGAAGAAAATGATTTGGGAAATTAGTAACAACGACCTAGAAA
 AGTGAGAACATCTCATTTACCATCATGTATCCAGTAGTGGATAATTCAATTGATGGCT
 TCTATTGGCCAATTGAGAATTAAAGCCAGTGCCTGAGACTGTCAGAAGTTGACCTTG
 CACTGGCATTAAAGAGTCATAGAAAAAGAATCATGGATTTTGAATTAGGTAAGAGG
 TGTTGGCTTTTTCTTTCCAGCCGTTGACCAATTAGTCGGCTGTTGACT
 GAGAAGTTGTGGGGAAAAGTTGCCATTTCCTTGCAATTGAAATTGCTT
 TACTTAGAAAAAGCGTCTAGAATGACCAAGTGTGTTGGTCGCAAATTGTCGACA
 AACCTATCCAAAACCTTGTGCTTAAAAAAACCTGCCCTAACGTTAGTCATCTGA
 GCTGGGCTCAGCTGGCTGTTCTGCACGCTGCAGGTGGCCACTCAITGGTCAGCA
 GGTCGGGGAGAGACTGGGATGGCTGGCTCTCTCTGCCTGCAGTCTGAGTCTC
 CTTCTCGTGTAGTCCTTCAGTGGCTGGCAGGGTAGCTAGACCTCTCACATGC
 AGCTCAGAGCTCCAAAGAGCTCAAAGCAGAAATGCCAGGCCTCTGAAAACCTAAGTC
 CAGAATTGTCACAGTGTCCCTTCACTTCCCTTATTGATGATGATGATGATGATGA
 TGATGATGATGATGATGATGATGATGTTTCTAATCAGAAGAAAGCTGGGTATGCCCTC
 TACTTACTAAACAAGTCACAAGCCCAGCTCAGATTCAAGAAAAGGGTGTGAAGTAGAGGT
 GCAGTTAAGTGGGGGCCACTAGTCTAACAGACGGTCACAACAGTGCATGGAAAACCA
 AGGATATTAGCAAAGCAGAAGTTGCTAGTGACCTTGGGAAGCCAGTGTACAGTA
 GCTGGGACAAGCTGAAAGTCAGACTAAGAAATAAGAGAGGGCTTCAAGAAGCTTCTG
 AATGATTCTGTAGCCCTGAGCCTATTGGAACACAGCACTTGGGAAACTGATCTG
 TGAGGATGGATGTTAGGGACACAGGGCTTTGAGAGCAGCACCCCCACTGGGCA
 TCCCCAGACTTGGGAAACGTGACTTTCTAATGCCACTGGGTTTAGTCAGGCCACAG
 TGAGAAGGAACAGCCTAACAGGCCCTCAGGCTTGAAGTGTGGCTCCTTGGATTCT
 CCAACCAGTAAGATTGCTAATGTTCTACATTAAGTGCCTCTCCAAAGACATCCCTT
 TGCCTCATATGTTGAATCATCCAGTGCAGGATATTCAATGAAAATATCATTGGTGA
 TTGTGATGGAATAATGCTATGGCATCTTGCATGAAGTTGTTGGTAGTTCTGAG
 CTGACTTGGCTCTGAAAGGAAGGCCTAGATCCAGGCCCTGGTGGTAGTTCTGAG
 GTCTCTCAGTCCCTGAGACTTGGGTAGTTGGCTGCCATTCTCACTGACAAAATGTA
 TATCAGCCCCACCTCCACCCCCAATATTCTGAACCTTGAATTGCTTCAGAACACAG
 GTGTGGCTGAAGGTATTCCCTTATTAGGAAGTGTCACTGCTGTTCTAGTC
 GTAAAGAAAAAGATTCCAGTTGCAAGAAGCTTGAATTGCTGTTCTGAG
 TCGCATTGTTACATGACTCATTCTCATTGGCTTGGTTGCTTGAATTGACTTGA
 CTTTGGGGTAAAGTCTTACCAAGCACACAAGAGTTGATTGTACAATATATCTG
 CATTAACATCTGCCTGTTGCTTAAGATCAGTTGCTTTATACTCAGAATGGAAATACC
 TGATCTGGCTAGTTGTTATAAGATATTGATTCTAGATTGCTTCCACGAGGTC
 AGCAAACATCATGTTCTATGTAACATTAGGCCAAGGCCAGAGTTATCATGTC
 GTTGCTACGGCTTATCATGTGCTGGTAAAGGTGATCGCAGGTTCTCAGACGAGTTAC
 TTTACATGAGATGGAATCAGGCAGAGAGGCTGGGATGATGGAGAAAGCTCAGGGTGAAGT
 TTTAAAAAAAAGTTGGAAAGGAAAGTCCAAAGAGGTGGTTCTGAGGAAGTCAGAG
 CGCCCAGGGCCAGAGCAGTCAGTAATGGGTGAATGAGGTTGTTGAAAGTCGGTGTGAC
 AGACACATGGATGCCATCTACTCTAGGTTGCTGGTGGTATTAAATATGCACAATATT
 CATAGCTCACTGAGGATTTAAAATTATAAGCATAGGATTCTATTTGGGGTGAAGA
 ATTATCTGGCACATTAGGATTGGAGTTAAAAAAAGCCAATTTCACAGTCTAATA
 ACTTTTTAAAAAAACTAAAGGCCTCATGTCCAGTGTGTGGCCCTCTGAAACTT
 ATGGTCATCTCTCCACTGAAACCAAGGTCTTCAATGTGGCTAAATGGGATGAGGA
 GACACGGTAGGACTTCTGGTGTGTGATTCTTAAAGAGCCAAGTTGCTTCGGGG

10

20

30

40

50

AAACAGCCAGGAAAATGGTCAAGATTATTTAGAGGTTATTTATTGGGGATTTAAGA
ACTAATAACATCTTGAGTTATTTAATTCAAGGGGATGTGAAAGGTTGCAATTGTCA
AGTGTGTTGTTGAGCTTAGTATCCATAAGGAAACTAGACTATAGACATAACTACAAA
GCCAGTGCAGCTTGTCTGTATGTTGGGGATCAACTTCACACATAGCAAGC
ACATGGCCTCCCTGATGTCAGGATGCCTTGTAGGATCTGTATTTGCCCTAATTTGT
TGAAATCTTTCTTCTTCCTTGAAAAGTCCAAAATATAGTTATTGTATCTTTC
ATCACTAAAAATTGTTCTTTCACTATGGCAGTCACACAAGGAAAAACTATTGA
ACAGTTGGTTTAGTGTGTTGATAACTTGCTGTATCAAACATAATTTGACAAGTT
TCATCCTAACGCTCAAATCATGTAATTAAATTGCGCTGTTATTTATGACCTAATTGT
GATTCTTTATTAATAAAAGCTAATGGGAAAAGGATCCCTGATTAAGCTGATGACTAGAC
CTACAATTAAATTTCCGAGTATATGAAGTATTGTACCGAGTATTAAGATATGTA
TATTTATTGATAAATCTATCCTTAAAAGGAATACGTTAGGATGTCATCATTGAT
GTGAATCATGTAATGTTGATAATATGCTGTTATTACATTAGTGTTCAGAGATT
CACTTAATTGCCCTTGCCCCAGTATATTATGCTATTGCAACTGTTCTAAAAA
AATGACATTAAAAGAATAGTTATGTAGAGAACATTAGTGGATGTTAATTGCTCCCCA
CCTATATTATGGGTGTTAGCGCAACTGCTTGCTAGTTGCAAAGCTGTATTATCAGAGT
AAAAGTGTATTGTAACGTATGGAACTAAAAATTAGGAATAAAACCATTCTTATA
TGATGGCATTGTCGTTGCTTCATCAGAAATGTCCAGGAAAAAATGGGATTATTGGTC
ACTCCACCTCTCACACTGGCAAAACTGACATTAGCAGCTTATCTAGAAGTGACTT
GGAACATAGAATAAAGGCATGAGTTCTGAAGAATTGAGTGTTCCTGTAGAAATA
GCTTAGGAGATAGGGAGTTCTATCTGGGAGAACATATGAGTAACCTAAGAGTAAAAAGT
ATAGTCTGTGTAACATAGAAGAAATGCTGGCATGGTGGCGCGCCCTGTAATCTCAG
CTACTTGGAGGCTGAGACGGGAGGATTCTGAACCCAGGAGCCAGGAGTTAGACCA
GTCTGGTAACATAGTGGAGACCTTCTCACCTACTCTGATGCCCTTACACACACACACACACA
TATATGTGCGCGACGCGCGCACACACACACACACACACACACACACACA
CAGAGGAAATTGTTAGAAAACACACAGAACTGAATGTAATAGTATTAGGTGGAAATAAG
AAGTAAAGGGATGGTAAGGAGGCTGGAGGAGGAGTAAATTATCTGCTATGGGACATCAG
CTC

10

20

30

個々の図面における同様の参照記号は同様の要素を示す。

【発明を実施するための形態】

【0040】

タンパク質 / 核酸複合体またはアセンブリは、その脆弱性及び構造的完全性（例えば3次元コンフォメーション）の要件ゆえに扱うのが困難である。ゲノム編集技術、例えばクラスター化等間隔短回文反復配列（C R I S P R）- C a s 9、転写活性化因子様エフェクターヌクレアーゼ（T A L E N）などは、細胞の遺伝子コードを変化させる能力に大きな可能性を秘めていることがわかっている。しかし、それらの活性は構造的及び立体配置的完全性に大きく依存する。

【0041】

ジンクフィンガーヌクレアーゼ（Z F N）及び転写活性化因子様エフェクターヌクレアーゼ（T A L E N）は遺伝子編集ツールの部類の例である。これらのキメラヌクレアーゼは、非特異的DNA切断ドメインに連結したプログラム可能な配列特異的DNA結合モジュールで構成される。Z F N及びT A L E Nは、特定のゲノム位置で誤りがちな非相同末端結合または相同組換え修復を刺激するDNA二本鎖切断を誘発することによって、広範囲の遺伝子改変を可能とする（G a j e t a l . , 2 0 1 3 , T r e n d s B i

40

50

o t e c h n o l . 31(7) : 397 - 405 ; 参照により本明細書に援用)。

【0042】

多数の刊行物で、Cas9及び／またはgRNA及び／またはドナーオリゴヌクレオチド（関連する場合）をコードするためのウイルス、mRNA及びプラスミドの使用について記載されている。こうした方法に対して、本明細書に記載の方法及びシステムは異なる方針、すなわちガイドRNAと複合体化させたタンパク質そのものの送達を用いる。このプロセスは、編集を行うのに細胞が翻訳／転写する必要がないので根本的に異なる。この理由により、またタンパク質／ガイド複合体は短い半減期を有するので、本発明のアプローチの結果、編集がより早く、より効率的となり、オフターゲット効果がより小さくなる。

【0043】

リポソーム媒介タンパク質送達、マイクロインジェクション、及び細胞膜透過性ペプチド（CPP）などの他のアプローチ

Cas-9に融合したGFPを用いるリポソーム媒介タンパク質送達の一例も用いられている（Zuris et al., 2015, Nature Biotechnol 33: 73 - 80）。GFPは、電荷相互作用による従来のリポフェクション剤（例えばリポフェクタミン）との複合体化が可能であり、遺伝子編集反応を媒介すると思われる。このアプローチと比較した本発明の主な利点は、1) 融合タンパク質を必要としない、2) 毒性、エンドソームエスケープの課題、及び初代細胞へ移入する問題／課題を有し得るリポフェクション剤を必要としないことである。

【0044】

マイクロインジェクション媒介複合体送達は、極めて低いスループットを特徴とし、ほとんどの哺乳類細胞型で実施するのが困難であり得る。後者の欠点は、胚、すなわち線維芽細胞またはT細胞（遺伝子編集の試みに望ましい標的細胞である細胞）よりもはるかに大きい細胞で実施されたという事実により浮き彫りにされている。高スループット及びより小さい初代細胞、例えば線維芽細胞、T細胞、幹細胞への移入を可能とすることにより、本明細書に記載の方法は大きな利点を有する。

【0045】

CPPベースの方策は複合体を必要としない。そのような方策の一例がRamakrishna et al., 2014, Genome Res. 24(6): 1020 - 7に記載されている。個々のコンポーネントのCPP媒介送達も欠点を伴う。CPPをガイド及びCas9に結合させるには、機能を阻害し、スケーラビリティを制限し得る追加の修飾が必要となる。CPP媒介送達はエンドサイトーシスを経ることが知られており、非効率的または多くの初代細胞（特に免疫細胞）において効果がない。

【0046】

標的細胞及びペイロード組成物

記載の遺伝子編集法を用いて任意の遺伝子を操作することができる。いくつかの標的遺伝子／タンパク質は特に臨床疾患に関連し、そのためそのような標的遺伝子／タンパク質の遺伝子編集は治療に有用である。例としては、C-Cケモカイン受容体5型（CCR5）：ヒト免疫不全ウイルス（HIV）感染の予防；主要組織適合性複合体クラスI（MHC-I）：移植片対宿主病の軽減；分化抗原群1（CD1）：移植片対宿主病の軽減；プログラム細胞死タンパク質1（PD-1）、プログラム死リガンド1（PDL-1）、細胞傷害性Tリンパ球関連タンパク質4（CTLA-4）、インターフェロン調節因子（IRF）タンパク質ファミリー、TLRタンパク質ファミリー、パターン認識受容体（PRR）：エフェクタ-/抗ウイルス反応を促進または抑制する免疫の調節；フォークヘッドボックスP3（FoxP3）：Treg媒介寛容の除去；分化抗原群80（CD80）、分化抗原群86（CD86）及び他の共刺激分子：寛容を促進する共刺激能のノックアウト；T細胞受容体（TCR）、B細胞受容体（BCR）：所望の特異性を有するT細胞及びB細胞の改変を可能とする内因性TCRまたはBCRの除去；癌遺伝子（例えばKras、Myc、Tp53）：癌治療が挙げられる。別の例では、細胞運命を変化させるために転写因子を標的とする、例えばFoxP3を欠失させてTreg型機能を除去する。核

10

20

30

40

50

内因子 B (N F - k B) 、 t - b e t 、 エオメソデルミン (E o m e s) などを欠失させて T 細胞分化を変化させる。

【 0 0 4 7 】

フォークヘッドボックス P 3 (F o x P 3) のタンパク質コード配列は以下の通りである。

【 0 0 4 8 】

ATGCCAACCCCAGGCCTGGCAAGCCCTGGCCCCCTCCTGGCCCTTGGCCATCCCC
 AGGAGCCTCGCCCAGCTGGAGGGCTGCACCCAAAGCCTCAGACCTGCTGGGGGCCGGG
 GCCCAGGGGAACCTTCCAGGGCGAGATCTCGAGGCGGGGCCATGCCTCCTTTCT
 TCCTTGAAACCCATGCCACCATCGCAGCTGCAGCTGCCAACACTGCCCTAGTCATGGT
 GGCACCCCTCGGGCACGGCTGGCCCTTGCCCCACTTACAGGCACTCCTCCAGGACA
 GGCCACATTCATGCACCAGCTCTCAACGGTGGATGCCACGCCGGACCCCTGTGCTG
 CAGGTGCACCCCTGGAGAGGCCATGATCAGCCTCACACCACCCACCACGCCAC
 TGGGTCTTCTCCCTCAAGGCCGGCTGCCCTCCACCTGGATCAACGTGGCCAGCC
 TGGAAATGGGTGTCAGGGAGCCGGACTGCTCTGCACCTCCAAATCCAGTGCACCC
 AGGAAGGACAGCACCCCTTCGGCTGTGCCAGAGCTCCTACCCACTGCTGGCAAATGG
 TGTCTGCAAAGTGGCCGGATGTGAGAAGGTCTTGAAGAGGCCAGAGGACTTCCCTAACGC
 ACTGCCAGGCAGGACATCTCTGGATGAGAAGGGCAGGGCACAAATGTCTCCAGAGA
 GAGATGGTACAGTCTCTGGAGCAGCAGCTGGTGTGGAGAAGGAGAAGCTGAGTGCAT
 GCAGGCCACCTGGCTGGAAAATGGCACTGACCAAGGCTCATCTGTGGCATCATCCG
 ACAAGGGCTCTGCTGCATCGTAGCTGCTGGCAGCCAAGGCCCTGTCGTCAGCCTGG
 TCTGGCCCCGGAGGCCCTGACAGCCTGTTGCTGTCCGGAGGCACCTGTGGGGTAG
 CCATGGAAACAGCACATTCCAGAGTTCCCTCACACATGGACTACTTCAAGTTCCACA
 ACATGCGACCCCTTCACTACGCCACGCTCATCCGCTGGCCATCCTGGAGGCTCCA
 GAGAAGCAGCGGACACTCAATGAGATCTACCACTGGTACACGCATGTTGCCTTCTT
 CAGAAACCATCCTGCCACCTGGAAGAACGCCATCCGCCACAAACCTGAGTCTGCACAAGT
 GCTTGTGCGGGTGGAGAGCGAGAAGGGCTGTGTGGACCGTGGATGAGCTGGAGTTC
 CGCAAGAAACGGAGGCCAGAGGCCAGCAGGTGTTCAACCCTACACCTGGCCCCCTGA

(配列番号 6 2)

【 0 0 4 9 】

S r c 相同領域 2 ドメイン含有ホスファターゼ 1 (S H P 1) はチロシン - タンパク質ホスファターゼ非受容体 6 型 (P T P N 6) としても知られている。 S H P 1 のタンパク質コード配列は以下の通りである。

【 0 0 5 0 】

ATGGTGAGGTGGTTCACCGAGACCTCAGTGGCTGGATGCAGAGACCTGCTCAAGGG
 CCGAGGTGTCCACGGTAGCTTCTGGCTGGCCCAGTCGCAAGAACCAAGGGTGACTTCT
 CGCTCTCGTCAGGGTGGGGATCAGGTGACCCATATTGGATCCAGAACTCAGGGAT
 TTCTATGACCTGTATGGAGGGAGAAGTTGCGACTCTGACAGAGCTGGTGGAGTACTA
 CACTCAGCAGCAGGGTGTCTGCAGGACCGCGACGGCACCATCATCCACCTCAAGTACC
 CGCTGAAGTCTCGATCCACTAGTGAGAGGTGGTACCATGCCACATGCTGGCGGG
 CAGGCAGAGACGCTGCAGGCCAAGGGCAGCCCTGGACGTTCTGTGCGTGAGAG
 CCTCAGCCAGCCTGGAGACTTCGTGCTTCTGTGCTCAGTGCACAGGCCAACGGCTGGCC
 CAGGCTCCCCGCTCAGGGTCACCCACATCAAGGTATGTGCGAGGGTGGACGCTACACA
 GTGGGTGGTTGGAGACCTTGACAGCCTCACGGACCTGGTGGAGCATTCAAGAACAC
 GGGGATTGAGGAGGCCTCAGGCCTTGTCTACCTGCGGCCAGCCGTACTATGCCACGA
 GGGTGAATGCGGCTGACATTGAGAACCGAGTGTTGAACTGAACAAGAACAGGAGTCC
 GAGGATAACGCCAAGGCTGGCTTCTGGAGAGTTGAGAGTTGAGAAGCTGAGAACAGGAGGT
 GAAGAACATTGCACCAAGCGTCTGGAAGGGCAGCGGCCAGAGAACACAAGGGCAAGAACCGCT
 ACAAGAACATTCTCCCTTGACCAAGCCGAGTGATCCTGCAGGGACGGACAGTAAC
 ATCCCCGGGTCCGACTACATCAATGCCAACTACATCAAGAACCAAGCTGCTAGGCCCTGA

10

20

30

40

50

TGAGAACGCTAAGACCTACATGCCAGCCAGGGTTGTCTGGAGGCCACGGTCAATGACT
 TCTGGCAGATGGCGTGGCAGGAGAACAGCCGTGTCATCGTCATGACCACCCGAGAGGTG
 GAGAAAGGCCGGAACAAATGCGTCCCATACTGGCCCAGGTGGCATGCAGCGTGCTTA
 TGGGCCACTCTGTGACCAACTGCAGGGAGCATGACACAACCGAATAACAAACTCCGTA
 CCTTACAGGTCTCCCCGCTGGACAATGGAGACCTGATTGGAGAGATCTGGCATTACCA
 TACCTGAGCTGGCCGACCATGGGTCCCCAGTGAGCCTGGGGTGTCTCAGCTTCCT
 GGACCAAGATCAACCAGCGCAGGAAAGTCTGCCTCACGCAGGGCCATCATCGTGCAC
 GCAGCGCCGGCATGGCCGACAGGCACCATCATTGTGACATGCTCATGGAGAAC
 ATCTCCACCAAGGGCCTGGACTGTGACATTGACATCCAGAAGACCATCCAGATGGTGC
 GGCAGCGCTCGGGCATGGTGACAGCGAGGCGCAGTACAAGTTCATCTACGTGGCCA
 TCGCCCAGTTCATTGAAACCACTAAGAAGAAGCTGGAGGTCTGCAGTCAGAAGGGC
 CAGGAGTCGGAGTACGGGAACATCACCTATCCCCAGCCATGAAGAATGCCATGCCAA
 GGCCTCCCGCACCTCGTCCAAGAGCTGGAGTCTAGTGCAGGGACCCTGGCTCGTCAC
 CTGTGAGACGGGTGGCCAGAGGGACTGCCAGTGCCTGGGGTCCCCCTGTGCTCTCCT
 GACCTGCACCAACTGCCTGTACTTGCCCCCTGCACCCGGCTGCAGACACAAGGAGGAT
 GTGTATGAGAACCTGCACACTAAGAACAAAGAGGGAGGAGAAAGTGA

(配列番号 63)

【0051】

S r c 相同領域 2 ドメイン含有ホスファターゼ 1 (SHP2) はチロシン - タンパク質
 ホスファターゼ非受容体 11 型 (PTPN11) としても知られている。SHP2 のタン
 パク質コード配列は以下の通りである。

【0052】

ATGACATCGCGGAGATGGTTCACCAAATATCACTGGTGTGGAGGCAGAAAACCTACT
 GTTGACAAGAGGAGTTGATGGCAGTTGGCAAGGCCTAGTAAAAGTAACCTGGAG
 ACTTCACACTTCCGTTAGAAGAAATGGAGCTGTCACCCACATCAAGATTCAAACACT
 GGTGATTACTATGACCTGTATGGAGGGAGAAATTGCCACTTGGCTGAGTTGGTCCA
 GTATTACATGGAACATCACGGCAATTAAAAGAGAAGAATGGAGATGTCATTGAGCTTA
 AATATCCTCTGAACACTGTGAGATCCTACCTCTGAAAGGTGTTCATGGACATCTCT
 GGGAAAGAAGCAGAGAAATTATTAAGTAAAAAGGAAACATGGTAGTTCTTGACG
 AGAGAGCCAGAGCCACCCCTGGAGATTGTTCTTCTGCGCACTGGTATGACAAAG
 GGGAGAGCAATGACGGCAAGTCTAAAGTACCCATGTTATGATTGCTGTCAGGAAC
 AAATACGACGTTGGTGGAGGAGAACGGTTGATTCTTGACAGATCTGTGGAACATTA
 TAAGAAGAATCCTATGGTGGAAACATTGGGTACAGTACTACAACCAAGCAGCCCTTA
 ACACGACTCGTATAATGCTGCTGAAATAGAAAGCAGAGTTCGAGAAACTAAGCAAATTA
 GCTGAGACCACAGATAAAGTCAAACAAGGCTTTGGGAAGAATTGAGACACTACAACA
 ACAGGAGTGCAAACCTCTACAGCCAAAAGAGGGTCAAAGGCAAGAAAACAAAACA
 AAAATAGATATAAAACATCCTGCCCTTGATCATACCAGGGTTGTCTACACGATGGT
 GATCCCAATGAGCCTGTTCAGATTACATCAATGCAAATATCATCATGCCTGAATTGA
 AACCAAGTGCAACAATTCAAAGCCAAAAGAGTTACATTGCCACACAAGGCTGCCTGC
 AAAACACGGTGAATGACTTTGGCGGATGGTGTCCAAGAAAACCTCCGAGTGATTGTC
 ATGACAACGAAAGAAGTGGAGAGAGGAAAGAGTAAATGTCATGCAAATACTGGCCTGATGA
 GTATGCTCTAAAAGAATATGGCGTCATGCGTGTAGGAACGTCAAAGAAAGCGCCGCTC
 ATGACTATACGCTAAGAGAACTTAAACTTCAAAGGTTGGACAAGGAAATACGGAGAGA
 ACGGTCTGGCAATACCACTTCGGACCTGGCCGGACCACGGCGTGCCTCAGCGACCCCTGG
 GGGCGTGTGGACTTCCTGGAGGAGGTGCACCATAGCAGGAGAGCATATGGATGCAG
 GCCCGGTCGTGGTGCAGTCAGTGGATTGGCCGGACAGGGACGTTGATTGTGATT
 GATATTCTTATTGACATCATCAGAGAGAAAGGTGTTGACTGCGATATTGACGTTCCCAA
 AACCATCCAGATGGTGCAGTCAGAGGTCAAGGGATGGTCAGACAGAACAGTAC
 GATTATCTATATGGCGGTCCAGCATTATATTGAAACACTACAGCGCAGGATTGAAGAA
 GAGCAGAAAAGCAAGAGGAAAGGGCACGAATATACAAATATTAAGTATTCTAGCGGA

10

20

30

40

50

CCAGACGAGTGGAGATCAGAGCCCTCTCCGCCTTGTACTCCAACGCCACCCTGTGCAG
AAATGAGAGAAGACAGTGCTAGAGTCTATGAAAACGTGGGCCTGATGCAACAGCAGAAA
AGTTTCAGATGA

(配列番号 64)

【0053】

他の標的としては、プラスミドまたはドナーDNAを挿入することができ、その結果、標的細胞が新たな遺伝子、例えば組換えTCR、組換えBCR、キメラ抗原受容体、蛍光タンパク質、リプログラミング因子を発現することができるゲノムの領域が挙げられる。

【0054】

いくつかの実施形態において、コーディング領域のゲノム配列を編集する。ある実施形態において、非コーディング領域のゲノム配列を編集する。

10

【0055】

FoxP3に関する様々な実施形態において、FoxP3の上流の遺伝子領域が編集され得る。そのような実施形態において、Foxp3の転写抑制因子が結合し得る領域を編集する。例えば、FoxP3転写開始部位の約0.5、1、1.5、2、2.5、3、3.5、4、4.5、5、0.5～2.5、または0.5～5kb上流の部位が編集され得る。

【0056】

被検者の治療

本発明の態様は、複数の被検者細胞のゲノムの編集に関する。様々な実施形態において、被検者から細胞を取り出し、本主題の方法を用いて遺伝子編集複合体を取り込ませ、その後、被検者に再び導入し戻す。例えば、本明細書に記載の遺伝子編集複合体の送達を伴うプロセスにおいて変異細胞が生成され得る。変異細胞は変異させた対立遺伝子、疾患に関する遺伝子がヘテロ接合型であってもよく、ホモ接合型であってもよい。ある実施形態において、変異細胞は疾患に関する遺伝子の機能が喪失している。

20

【0057】

骨髄由来の幹細胞（例えば造血幹細胞）、または全血中の循環免疫細胞などの細胞が、本明細書に記載の方法及びデバイスを用いて処理され得る。細胞は、病原体（例えばウイルス性もしくは細菌性病原体など）または毒素（例えば微生物性病原体毒素など）に対する受容体の発現を減少させるように遺伝子改変され得る。遺伝子編集タンパク質複合体またはそのコンポーネントは発現させる必要なく標的細胞内に直接導入され得るので、導入遺伝子の送達が必要ない。このアプローチは、異常な発現、挿入、及びサイレンシング、ならびに導入遺伝子コピー数の不定な送達に悩まされる従来の遺伝子治療アプローチに勝る重要な利点を有する。

30

【0058】

一実施形態において、ヒト免疫不全ウイルス（HIV）に感染している被検者の血液（例えばCD4+T細胞など）または骨髄細胞（例えば造血幹細胞など）に、C-Cケモカイン受容体5型（CCR5）遺伝子を標的とする遺伝子編集複合体を導入する。遺伝子編集複合体は、CCR5遺伝子を変異させ、その結果、遺伝子編集複合体を取り込んだ細胞がCCR5を発現することがなくなるか、またはCCR5を低いレベルで発現するように設計され得る。一例において、HIVに結合するCCR5の型を発現する（またはCCR5を発現する子孫細胞を生成する）造血幹細胞を被検者から取り出し、HIVに結合するCCR5の型を発現することができないように改変し、その後、被検者に移植する。別の例では、CD4+T細胞がHIVに結合するCCR5の型を発現することができないように、本明細書に記載の方法及びデバイスを用いて、被検者のCCR5発現CD4+T細胞に遺伝子編集複合体を取り込ませる。その後、改変CD4+T細胞を被検者に戻す。CD4+T細胞のそのような処理は被検者の全血で実施され得る。これらの及び他の実施形態において、HIVが結合するC-X-Cケモカイン受容体4型（CXCR4）を発現することができないように、骨髄細胞または血液細胞を改変する。同様に、Yersinia pestis（腺ペスト）またはVariola major（天然痘）に関連する感染を治療また

40

50

は予防するために、被検者の細胞がCCR5発現を減少させるように改変され得る。

【0059】

本明細書に開示の方法及びデバイスによって改変した細胞を有するヒト以外の被検者も規定する。そのような被検者としては、非ヒト脊椎動物、両生類、哺乳類、及び霊長類被検者が挙げられる。非限定例としては、*Danio* sp.、*Fugu* sp.、*Xenopus* sp.、*Mus* sp.、*Rattus* sp.などが挙げられる。

【0060】

細胞内への遺伝子編集タンパク質及び複合体の導入

予め形成したタンパク質複合体の送達により、研究している細胞を遺伝子改変することなく細胞プロセスの研究が可能となる。本主題は、細胞にCRISPRをはじめとするタンパク質複合体及び遺伝子編集複合体を送達するのに有用である。

10

【0061】

本明細書に記載の方法及びデバイスを用いてタンパク質複合体を送達する利点としては、細胞及びタンパク質複合体の機能の研究用の試験剤の制御された一時的な導入が挙げられる。導入遺伝子の発現及び複合体コンポーネントの細胞アセンブリが必要ではないので、複合体機能のタイミング及び比（タンパク質：RNA）を制御することができる。加えて、オフターゲット効果または副次効果をもたらし得る長期発現ではなく、送達の一過性の性質により、一時的な機能に起因する変化の観察が可能となる。インビボホーミングの観点及び遺伝子発現様式から、マイクロ流体送達ではエレクトロポレーションよりも処理する細胞への副作用がはるかに少ない（10倍）。

20

【0062】

例えば、マイクロ流体送達の結果、エレクトロポレーションと比較して、異常な及び非特異的な遺伝子発現の変化が少なくなる。加えて、マイクロ流体的に圧搾した細胞の構造的及び機能的完全性は、エレクトロポレーション媒介カーゴ送達と比較して保持される。一例として、エクスピボでマイクロ流体送達に供した（その後被検者体内に導入した）T細胞は、エレクトロポレーションを受けたT細胞と比較してリンパ節にホーミングする数が多い。エレクトロポレーションによって処理され、その後被検者に投与された細胞（例えばT細胞）は、マイクロ流体送達によって処理された細胞と比較して、被検者から除去される可能性がより高い。そのようなクリアランスは、そのような細胞を体による破壊またはクリアランスに対して標識するエレクトロポレーション後の改変／異常遺伝子発現に関係している。

30

【0063】

標的細胞及びペイロード組成物

マイクロ流体デバイスを用いて、任意の真核細胞、例えばヒトなどの哺乳類細胞を処理し、標的細胞のサイトゾル中へのタンパク質／核酸複合体またはアセンブリの導入のために細胞膜を変化させることができる。例示的な標的細胞としては、リンパ球／免疫細胞：DC、B細胞、T細胞、ナチュラルキラー細胞（NK細胞）、好中球、好塩基球、好酸球、自然リンパ球、単球、マクロファージ、造血幹細胞、リンパ球系共通前駆細胞；幹細胞：胚性、間葉系、人工多能性；他の初代細胞：線維芽細胞、肝細胞、心筋細胞、ニューロン、上皮、表皮、内皮、膵島細胞；及び細胞株、例えば疾患研究のための細胞株：T細胞クローン、ジャーカット細胞、ヒーラ細胞、ヒト胎児腎293（HEK293）細胞、U2OS細胞、チャイニーズハムスター卵巣（CHO）細胞が挙げられる。原核細胞を処理することもできる。デバイスの狭窄部の寸法は、処理しようとする細胞型に応じて調整する。

40

【0064】

いくつかの実施形態において、細胞は原核細胞である。いくつかの実施形態において、細胞は真核細胞である。真核細胞の非限定例としては、原生動物、藻類、真菌、酵母、植物、動物、脊椎動物、無脊椎動物、節足動物、哺乳類、げっ歯類、霊長類、及びヒト細胞が挙げられる。細胞は、例えば单細胞生物または多細胞生物の細胞であり得る。細胞は、例えば初代真核細胞または不死化真核細胞であり得る。いくつかの実施形態において、細

50

胞は癌細胞である。ある実施形態において、細胞はヒト細胞以外である。例えば、癌を治療するための組成物及び／または癌の治療方法もしくは癌を治療するための組成物の調製方法は、本明細書に記載の遺伝子編集方法を用いて、腫瘍細胞からの免疫抑制シグナルの発現／產生を減少させるように免疫細胞を処理することを含む。例としては、腫瘍に対する免疫活性を増加させるための減少またはSHP-2ノックアウトが挙げられる。

【0065】

様々な実施形態において、細胞は2種以上の細胞型の混合物であり得る、または複数の細胞は2種以上の細胞型の混合物であり得る。細胞型の混合物は複数の細胞型（例えば本明細書に開示したもの2種以上など）の共培養物、または天然で一緒に存在する細胞型の混合物、例えば全血などであり得る。

10

【0066】

いくつかの実施形態において、細胞は末梢血単核細胞である。様々な実施形態において、細胞懸濁液は精製された細胞集団を含む。ある実施形態において、細胞は初代細胞または細胞株細胞である。

【0067】

いくつかの実施形態において、細胞は血液細胞である。いくつかの実施形態において、血液細胞は免疫細胞である。いくつかの実施形態において、免疫細胞はリンパ球である。いくつかの実施形態において、免疫細胞は、T細胞、B細胞、ナチュラルキラー（NK）細胞、樹状細胞（DC）、ナチュラルキラーT（NKT）細胞、肥満細胞、単球、マクロファージ、好塩基球、好酸球、または好中球である。いくつかの実施形態において、免疫細胞は適応免疫細胞、例えばT細胞及びB細胞などである。いくつかの実施形態において、免疫細胞は先天性免疫細胞である。例示的な先天性免疫細胞としては、自然リンパ球（ILC；ILC1、ILC2、ILC3）、好塩基球、好酸球、肥満細胞、NK細胞、好中球、及び単球が挙げられる。いくつかの実施形態において、免疫細胞はメモリー細胞である。いくつかの実施形態において、免疫細胞は初代ヒトT細胞である。いくつかの実施形態において、細胞はマウス、イヌ、ネコ、ウマ、ラット、ヤギ、サル、またはウサギ細胞である。

20

【0068】

いくつかの実施形態において、細胞はヒト細胞である。いくつかの実施形態において、細胞懸濁液はヒト細胞以外の細胞または非哺乳類細胞を含む。いくつかの実施形態において、細胞はニワトリ、カエル、昆虫、または線虫細胞である。

30

【0069】

細胞を浸し／インキュベートし、デバイスを通して細胞を処理する溶液として、任意の生理学的に適合性のまたは細胞適合性のバッファー系を用いることができる。例えば、リン酸緩衝生理食塩水（PBS）、Opti-MEM（登録商標）、ロズウェルパーク記念研究所（RPMI）、ダルベッコ変換イーグル培地（D MEM）。低血清または無血清培地またはバッファー組成物が好ましい。バッファーまたは培地は、標的細胞の健康もしくは生存力及び／または遺伝子発現への効果を維持及び保存することに基づいて選択する。例えば、場合によっては、バッファー中のカルシウムの存在はmRNA発現を促進または補助するのに望ましい。

40

【0070】

ペイロード組成物は、タンパク質-核酸複合体またはアセンブリを含む。例示的な複合体は、上記の遺伝子編集システム、例えばスクレアーゼ／ガイド核酸結合体またはアセンブリのコンポーネントまたはモジュールを含む。例えば、gRNA : Cas9モル比は1:100,000~100,000:1の範囲であり、例えば好ましい範囲は1:10~10:1、例えば1:1または1:2、2:1である。送達を容易にするバッファー中の複合体濃度（モル濃度）は典型的には100mM~1nM、例えば10uM~100nMの範囲である。複合体は、狭窄部を通過する前またはその後で細胞と混合させることができる。

【0071】

50

遺伝子編集複合体のマイクロ流体送達

効果的な遺伝子編集操作を行うために、例えばマイクロ流体プラットフォーム（例えば2014年4月17日に出願された米国特許出願公開第20140287509号；2014年8月15日に出願されたPCT国際出願第PCT/US2014/051343号；2015年11月13日に出願されたPCT国際出願第PCT/US2015/060689号；及び2015年10月30日に出願されたPCT国際出願第PCT/US2015/058489号に記載、それぞれ参照により本明細書に援用）を用いた機械的変形によって、Casタンパク質（例えばCas9タンパク質など）、ガイドRNA及びドナードNAを細胞に送達することができる。

【0072】

図1～2は、ゲノム編集タンパク質、RNA、及びDNAの送達に用いることのできるマイクロ流体システムの例を示す。マイクロ流体システム5は管状の内腔を画定するチャネル10を含む。マイクロ流体チャネル10は、好ましくは一度に1つの標的細胞20のみが狭窄部15を通過することができるよう構成された狭窄部15を含む。好ましくは、細胞20は、送達物質30も含む溶液バッファー25中に懸濁されてチャネル10を通過するが、細胞20が狭窄部15を通過した後に送達物質を溶液バッファー25に加えることもできる。細胞20が狭窄部15に接近して通過する際に、狭窄部15により細胞20に圧力（例えば機械的圧縮）がかかり、細胞20を圧搾する（例えば細胞20₁として示す）。狭窄部15によって細胞にかかる圧力は細胞膜に摂動（例えば孔）を引き起こす（例えば細胞20₂）。細胞が狭窄部15を通過すると、細胞20は孔を通して、送達物質30を含め溶液バッファー25中の物質を取り込み始める（例えば細胞20₃）。細胞膜は時間の経過とともに修復し、送達物質30の少なくとも一部は好ましくは細胞内に捕捉されたままとなる。

【0073】

いくつかの実施形態において、デバイスは、約5μm～約50μmまたはこの間の任意の長さもしくは長さの範囲の狭窄部長さを備える。例えば、狭窄部長さは、約5μm～約40μm、約5μm～約30μm、約5μm～約20μm、または約5μm～約10μmの範囲である。いくつかの実施形態において、狭窄部長さは、約10μm～約50μm、約20μm～約50μm、約30μm～約50μm、または約40μm～約50μmの範囲である。いくつかの実施形態において、狭窄部深さは約2μm～約200μmの範囲、またはこの間の任意の深さもしくは深さの範囲である。例えば、狭窄部深さは、約2μm～約150μm、約2μm～約100μm、約2μm～約50μm、約2μm～約25μm、約2μm～約15μm、または約2μm～約10μmの範囲である。いくつかの実施形態において、狭窄部深さは、約10μm～約200μm、約25μm～約200μm、約50μm～約200μm、約100μm～約200μm、または約150μm～約200μmの範囲である。いくつかの実施形態において、狭窄部の入口または出口部分の角度は約0度～約90度の範囲またはこの間の任意の角度もしくは角度の範囲である。例えば、角度は約5、約10、約15、約20、約30、約40、約50、約60、約70、約80、または約90度以上である。いくつかの実施形態において、圧力は約50psi～約200psiの範囲、またはこの間の任意の圧力もしくは圧力の範囲である。例えば、圧力は約50psi～約150psi、約50psi～約125psi、約50psi～約100psi、または約50psi～約75psiの範囲である。いくつかの実施形態において、圧力は約75psi～約200psi、約100psi～約200psi、約125psi～約200psi、約150psi～約200psi、または約175psi～約200psiの範囲である。いくつかの実施形態において、デバイスは約2μm～約10μmまたはこの間の任意の幅もしくは幅の範囲の狭窄部幅を備える。例えば、狭窄部幅は約3μm、約4μm、約5μm、約6μm、または約7μmのいずれか1つとすることができる。

【0074】

以下に記載するデータは以下の物質及び方法を用いて得た。複合体は以下の通り作製し

10

20

30

40

50

た：C a s 9 タンパク質で標識した 1 m g / m l 核局在化シグナル (N L S) 1 0 μ l を 1 m g / m l ガイド R N A 5 μ l と混合する。2 0 分間氷上でインキュベートして複合体を形成させる。送達のために、無血清培地に 1 千万細胞 / m l で標的細胞を懸濁する。デバイス処理の直前に、複合体濃度が約 0 . 1 5 m g / m l となるように細胞及び C a s 9 - g R N A 複合体を混合する。標的細胞型に特異的な圧力、温度、チップデザイン及びバッファー条件を用いて、細胞をデバイスによって処理する。例えば、初代ヒト T 細胞については、圧力はおよそ 1 0 0 p s i であり、氷上で 3 0 μ m の長さ、4 μ m の幅の狭窄部を通す。2 分間の処理後インキュベーションの後、細胞を培地で希釈し、洗浄して未送達複合体を除去する。その後、細胞を培養して遺伝子編集を発生させる（例えば 1 、 2 、 5 、 1 2 、 2 4 時間またはこれより長く（非臨床用途については、期間はアッセイ読み取りに依存する、例えば 2 4 時間以降））。臨床用途、例えば患者治療については、細胞はデバイス処理の直後に患者に注入し戻され得る。任意により、患者レシピエントへの細胞の注入前に、一定時間（例えば 1 、 2 、 5 、 1 2 、 2 4 時間またはこれより長く）細胞をインビトロでインキュベートする。分子の温度、濃度、繰り返しは標的細胞型に応じて異なる。

【 0 0 7 5 】

図 3 は、図 1 ~ 2 に示したマイクロ流体デバイスを用いて細胞質に遺伝子編集物質を送達された K 5 6 2 レポーター細胞 (American Type Culture Collection (ATCC) (登録商標) C C L - 2 4 3 (商標) 、慢性骨髄性白血病の患者に由来する骨髄由来細胞；リンパ芽球形態) における、一連のフローサイトメトリープロット及び組換え効率の棒グラフである。上はタンパク質 C a s 9 、部位特異的 g R N A 、及びドナーオリゴヌクレオチドが送達された K 5 6 2 レポーター細胞のフローサイトメトリープロットである。C R I S P R 複合体は約 1 5 0 k D a である。これらのレポーター細胞は m C h e r r y 遺伝子を有していたので、通常であればプロットの Q 1 に現れる。ドナーオリゴヌクレオチドの部位特異的 D N A 切断及び挿入が成功すれば、緑色蛍光タンパク質 (G F P) の発現をもたらす、すなわち細胞は Q 2 に現れる。下は、エンドサイトーシス (取込み) 及び未処理 (N C) 対照と比較した、複数のデバイス条件についてのフローサイトメトリーに基づく組換え効率の定量化である。細胞圧搾による C a s 9 タンパク質、ガイド R N A 及びドナー D N A の送達の結果、レポーター細胞株のゲノムの変化に成功した。これらのデータは、機械的細胞破碎による遺伝子編集コンポーネントの（複合体またはアセンブリの形態での）送達によって、効果的にゲノムの変化がもたらされることを示す。

【 0 0 7 6 】

T A L E N タンパク質もしくは m R N A 、ジンクフィンガーヌクレアーゼ、メガヌクレアーゼ、C r e リコンビナーゼまたは D N A を切断可能な任意の他の酵素の送達も、細胞膜の機械的破碎によって細胞の細胞質に送達することができる。例示的な T A L E N ゲノム編集システムは、例示的な T A L E N タンパク質を含め、D i n g e t a l . , (2 0 1 3) C e l l S t e m C e l l , 1 2 , 2 3 8 - 2 5 1 に記載されており、その内容全体を参照により本明細書に援用する。D i n g e t a l . , (2 0 1 3) C e l l S t e m C e l l , 1 2 , 2 3 8 - 2 5 1 は、1 5 塩基対配列を認識する包括的 T A L E N アミノ酸配列の非限定例を記載している。包括的 T A L E N アミノ酸配列の非限定例は以下である。

10

20

30

40

50

MDYKDHDG DYKDHIDYKDDDKMAPKKRKVGIGHGVPSRVDLRTLGYSOOOOEKIKPKVRSTVAOHH
EALVGHGFTHAHIVALSOHPAALGTVAVTYOHITITALPEATHEDIVGVGKOWSGARALEALLTDAGEL
 RGPPLOLDTGOLVKIAKRGVTAMEAVHASRNALTGAPLNLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGRPALESIVAOLSR
PDPALAALTNDHLVALACLGGRPAMDAVKKGKGLPHAPELIRRVNRRIGERTSHRVAGSQLVKSELEEKK
 SELRHKLKYVPHHEYIELIEIARNPTQDRILEMKVMEFFMKVYGYRGEHLGGSRKPDGAIYTVGSPIDY
 GVIVDTKAYSGGYNLPIGQADEMERYKENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKA
 QLTRLNRKTNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF

10

(配列番号 6 5)

MDYKDHDG DYKDHIDYKDDDKMAPKKRKVGIGHGVPARVDLRTLGYSOOOOEKIKPKVRSTVAOHH
EALVGHGFTHAHIVALSOHPAALGTVAVTYOHITITALPEATHEDIVGVGKOWSGARALEALLTDAGEL
 RGPPLOLDTGOLVKIAKRGVTAMEAVHASRNALTGAPLNLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
 LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGKQALETVQRLLPV
LCQAHGLTPEQVVAIASXXGGKQALETVQRLLPVLCQAHGLTPEQVVIAIASXXGGRPALESIVAOLSR
PDPALAALTNDHLVALACLGGRPAMDAVKKGKGLPHAPELIRRVNRRIGERTSHRVAGSQLVKSELEEKK
 SELRHKLKYVPHHEYIELIEIARNPTQDRILEMKVMEFFMKVYGYRGEHLGGSRKPDGAIYTVGSPIDY
 GVIVDTKAYSGGYNLPIGQADEMERYVEENQTRDKHLNPNEWWKVYPSSVTEFKFLFVSGHFKGNYKA
 QLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF

20

(配列番号 6 6)

【 0 0 7 7 】

配列番号 6 5 及び 6 6 において、 C を認識するには X X = H D ; T を認識するには X X = N G ; A を認識するには X X = N I ; 及び G を認識するには X X = N N 、ただし最後の位置のみ X X = N K である。下線は T A L エフェクタ - の N 末端及び C 末端を示す。

30

【 0 0 7 8 】

上記では少数の変形しか詳細に記載していないが、他の修正または追加が可能である。例えば、遺伝子編集物質には、細胞膜の機械的破碎によって細胞の細胞質に送達される T A L E N タンパク質、 T A L E N m R N A 、ジンクフィンガーヌクレアーゼ、メガヌクレアーゼ、 C r e リコンビナーゼまたは D N A を切断可能な任意の他の酵素を含めることができる。

30

【 0 0 7 9 】

複合体形態の R N A 及び C a s 9 の送達

得られた結果は、遺伝子複合体のマイクロ流体送達による遺伝子編集の成功を妨げる可能性のあった多数の要因を鑑みると驚くべきものであった。例えば、 C a s 9 - g R N A 複合体は、遺伝子編集機能及び / または存続を阻害したであろう T o 1 1 様受容体 (T L R) 媒介または他の P R R (パターン認識受容体) 媒介反応を引き起こす可能性があったが、この潜在的な問題は観察されなかった。細胞質に入っても複合体が安定である保証はないので、分解し、機能しなくなる可能性があったが、驚くべきことに、送達された複合体は変わらず編集することができた。

40

【 0 0 8 0 】

遺伝子編集複合体の完全性は、細胞へのマイクロ流体ベースの細胞圧搾送達を用いて保存された。複合体は g R N A のみまたはタンパク質のみと同じ物理的 / 化学的特性を有せず、そのため、その個々のコンポーネントに対して複合体の送達という状況において、送達プロセスが同様に振る舞うかどうかは不確実であった。複合体はその構成要素よりも大

50

きく安定性が低い。複合体は、例えばせん断力によって分解し得る。加えて、機能する、または細胞標的に作用する機会を得る前に何らかの他の要素が複合体を分解し得るので、複合体は膜通過を経て、またはサイトゾル内で存続しない場合がある。複合体は、送達される複合体の能力に影響を及ぼし得る異なる電荷分布も有する。形状ひいては輸送特性も複合体構成要素と比較して変化し得る。本送達方法では複合体の構造的及び機能的完全性が首尾よく保存された。

【0081】

送達プロセスに関するせん断力は、タンパク質 / g RNA 複合体を破壊し、機能させなくしていた可能性もあったが、驚くべきことに、本送達システムは細胞内に複合体を導入するのに効果的であり、遺伝子編集は変わらず行われた。NLS を有する非複合体 Cas (例えば Cas9 など) タンパク質のみと比較して、複合体が核局在化挙動を変わらず有するかも明らかではなかったが、上記の遺伝子発現結果に実証されるように挙動及び機能はプロセス全体を通して保存されていた。

10

【0082】

C R I S P R - Cas システムは当該技術分野において公知である。2014年4月15日に発行された米国特許第 8,697,359 号にこのシステムの非限定態様が記載されており、その内容全体を参照により本明細書に援用する。

【0083】

Cas タンパク質の非限定例としては、Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8、Cas9 (Csn1 及び Csx1 2 としても知られる) 、Cas10、Csy1、Csy2、Csy3、Cse1、Cse2、Csc1、Csc2、CsA5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、Csax、Csx3 、Csx1、Csx15、Csf1、Csf2、Csf3、Csf4 、これらのホモログ、またはこれらの改変型が挙げられる。これらの酵素は公知であり、例えば、S. pyogenes Cas9 タンパク質のアミノ酸配列は SwissProt データベースで受託番号 Q99ZW2 の下に見出され得る。いくつかの実施形態において、非改変 C R I S P R 酵素は DNA 切断活性、例えば Cas9 などを有する。いくつかの実施形態において、C R I S P R 酵素は Cas9 であり、S. pyogenes または S. pneumoniae 由来の Cas9 であり得る。いくつかの実施形態において、C R I S P R 酵素は標的配列の位置、例えば標的配列内及び / または標的配列の相補鎖内などで一方または両方の鎖の切断を誘導する。いくつかの実施形態において、C R I S P R 酵素は、標的配列の最初または最後のヌクレオチドから約 1、2、3、4、5、6、7、8、9、10、15、20、25、50、100、200、500 、またはこれより多くの塩基対以内で、一方または両方の鎖の切断を誘導する。いくつかの実施形態において、ベクターが、変異 C R I S P R 酵素が標的配列を含有する標的ポリヌクレオチドの一方または両方の鎖を切断する能力を欠くように、対応する野生型酵素に対して変異している C R I S P R 酵素をコードする。例えば、S. pyogenes 由来の Cas9 の RuvC I 触媒ドメインのアスパラギン酸からアラニンへの置換 (D10A) は、両方の鎖を切断するヌクレアーゼからニッカーゼ (片方の鎖のみ切断) に Cas9 を変換する。Cas9 をニッカーゼにする変異の他の例としては、限定はしないが、H840A、N854A、及び N863A が挙げられる。本発明の態様において、ニッカーゼは相同組換えを介したゲノム編集に用いられ得る。

20

30

30

【0084】

Cas9 アミノ酸及び cDNA 配列の非限定例を以下に示す。

【0085】

Streptococcus pyogenes Cas9 タンパク質のアミノ酸配列は、SwissProt データベースで受託番号 Q99ZW2 の下に見出され得る。このアミノ酸配列は以下である。

40

50

【0086】

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDGETAEA
 TRLKRTARRRYTRRKNRICYLQEIFSNEAKVDDSFHRLEESFLVEEDKKHERHPIFGNI
 VDEVAYHEKYPTIYHLRKKLV DSTDKA DLRLIYLA LAHMI KFRGHF LIEGDLNP DSDVD
 KLF IQLV QTYNQ LFEEN PINAS GVD AKA IL SARLS KRRLEN LIA QLP GEKK NGL FGN LIAL
 SLGLTPNFKSNFDLAEDAKLQLSKDTYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI
 LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQS KNGYAGYIDG
 GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ 10
 EDFYPFLKDNR EKIEKILTFRIPYYVGPLARGNSRFAMTRKSEETITPWNFEVVVDKGAS
 AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKVYVTEGMRKPAFLSGEQKKA
 IV DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGV EDRFN ASLG TYHDLLKIIKDKDFLD
 NEENEDILEDIVLTTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRTGWGRSLRK LIN
 GIRDKQSGKTILD FLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGD SLHEHIANLA
 GSPA IKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIE EG
 KELGSQILKEHPVENTQLQNEKLYLYLQNGRD MYVDQELDINRLSDYDVDHIVPQSFLK
 DDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKLITQRKF DNLT KAERGG
 LSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKS KLVS DFRK
 DFQFYKVREINNYHHAHDAYLNAV VGTALIKKP KLESEFVYGDYK VYDVRK MIAKSEQ
 EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
 MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGF DSPTVAYS VLV VAK 20
 VEKGKSKKLKSVKELLGITMERSFEKNPIDF LEAKGYKEVKKDLI KLPK YSLF ELENGR
 KRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEII
 EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
 RKRYTSTKEVLDATL IHQSITGLYETRIDLSQLGGD

(配列番号1)

【0087】

配列番号1はEuropean Nucleotide Archiveで受託番号AA
 K33936.2の下に見出される以下のヌクレオチド配列によってコードされ得る。

【0088】

ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGCGGT 30
 GATCACTGATGAATATAAGGTTCCGTCTAAAAGTTCAAGGTTCTGGGAAATACAGACC
 GCCACAGTATCAAAAAAAATCTTATAGGGCTCTTTATTGACAGTGGAGAGACAGCG
 GAAGCGACTCGTCTAACGGACAGCTCGTAGAAGGTATACACGT CGGAAGAATCGTAT
 TTGTTATCTACAGGAGATTTTCAAATGAGATGGCAAAGTAGATGATAGTTCTTC
 ATCGACTTGAAGAGTCTTTGGTGGAGAAGACAAGAACATGAACGT CATCCTATT
 TTTGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATCCAACATCTATCATCT
 GCGAAAAAAATTGGTAGATTCTACTGATAAAGCGGATTGCGCTTAATCTATTGGCCT
 TAGCGCATATGATTAAGTTCGTGGTCATTTTGATTGAGGGAGATTAAATCCTGAT
 AATAGTGATGTGGACAAACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTGA
 AGAAAACCTATTAACGCAAGTGGAGTAGATGCTAAAGCGATTCTTCTGCACGATTGA 40
 GTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGT GAGAAGAAAAATGGC
 TTATTTGGGAATCTCATTGCTTGT CATTGGTTGACCCCTAATTAAATCAAATTT
 GATTGGCAGAAGATGCTAAATTACAGCTTCAAAAGATACTTACGATGATGATTAGA
 TAATTATTGGCGCAAATTGGAGATCAATATGCTGATTGTTTGGCAGCTAAGAATT
 TATCAGATGCTATTTACTTCAGATATCCTAACAGATAACTGAAATAACTAAGGCT
 CCCCTATCAGCTCAATGATTAACGCTACGATGAACATCATCAAGACTTGACTCTTT
 AAAAGCTTAGTTGACAACAACCTCCAGAAAAGTATAAAGAAATCTTTTGATCAAT
 CAAAAAACGGATATGCAGGTTATTGATGGGGAGCTAGCCAAGAAGAATTGGT GAA ACTAAA
 TTTATCAAACCAATTAGAAAAAATGGATGGTACTGAGGAATTATTGGT GAA ACTAAA
 TCGTGAAGATTGCTGCGCAAGCAACGGACCTTGACAACGGCTCTATTCCCCATCAA 50

TTCACTGGGTGAGCTGCATGCTATTTGAGAAGACAAGAAGACTTTATCCATTTA
 AAAGACAATCGTGAGAAGATTGAAAAAAATCTTGACTTTCGAATTCTTATTATGTTGG
 TCCATTGGCGCGTGGCAATAGTCGTTGCATGGATGACTCGGAAGTCTGAAGAACAA
 TTACCCCATGGAATTTGAAGAAGTGTGATAAAGGTGCTTCAGCTCAATCATTATT
 GAACGCATGACAAACTTGATAAAAATCTCCAATGAAAAAGTACTACCAAAACATAG
 TTTGCTTATGAGTATTTACGGTTATAACGAATTGACAAAGGTCAAATATGTTACTG
 AAGGAATGCGAAAACCAGCATTCTTCAGGTGAACAGAAGAAAGCCATTGTTGATT
 CTCTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAAAGAAGATTATTCAAAAA
 AATAGAATGTTGATAGTGTGAAATTCAGGAGTTGAAGATAGATTTAATGCTTCAT
 TAGGTACCTACCATGATTGCTAAAAATTATTAAAGATAAAAGATTGTTGGATAATGAA
 GAAAATGAAGATATCTTAGAGGATTGTTAACATTGACCTTATTGAAGATAGGG
 GATGATTGAGGAAAGACTTAAACATATGCTCACCTCTTGATGATAAGGTGATGAAAC
 AGCTTAAACGTCGCCGTTACTGGTGGGGACGTTGTCTGAAAATTGATTAATGGT
 ATTAGGGATAAGCAATCTGGAAAACAATTAGATTGTTGAAATCAGATGGTTG
 CAATCGCAATTTCATGCAGCTGATCCATGATGATAGTTGACATTAAAGAACATTC
 AAAAGCACAAGTGTCTGGACAAGGCATAGTTACATGAACATATTGCAAATTAGCT
 GGTAGCCCTGCTATTAAAAAGGTATTACAGACTGTAAGTTGATGAATTGGT
 CAAAGTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAC
 AGACAACCTAAAAGGGCCAGAAAATCGCGAGAGCGTATGAAACGAATCGAAGAAGGT
 ATCAAAGAATTAGGAAGTCAGATTCTAAAGAGCATCCTGTTGAAAATACTCAATTGCA
 AAATGAAAAGCTCTATCTCTATTATCTCCAAATGGAAGAGACATGTATGTGGACCAAG
 AATTAGATATTAAATCGTTAAGTGATTATGATGTCGATCACATTGTTCCACAAAGTTTC
 CTTAAAGACGATTCAATAGACAATAAGGTCTTAACCGCTCTGATAAAAATCGTGGTAA
 ATCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAGATGAAAACATTGGAGAAC
 TTCTAAACGCCAAGTTAACACTCAACGTAAGTTGATAATTAAACGAAAGCTGAAC
 GGAGGTTGAGTGAACCTGATAAAGCTGGTTTATCAAACGCCAATTGGTGAAC
 CCAAATCACTAACGATGTGGCACAAATTGGATAGTCGATGAATACTAAACGATG
 AAAATGATAAACTTATTGAGAGGTTAAAGTGATTACCTAAATCTAAATTAGTTCT
 GACTTCCGAAAAGATTCCAATTCTATAAAGTACGTGAGATTAACAATTACCATCATGC
 CCATGATGCGTATCTAAATGCCGCGTGGAACTGCTTGATTAAGAAATATCCAAA
 TTGAATCGGAGTTGTCTATGGTATTAAAGTTATGATGTTGCTAAATGATTGCT
 AAGTCTGAGCAAGAAATAGGCAAAGCAACGCCAAATATTCTTTACTCTAATATCAT
 GAACTTCTCAAAACAGAAATTACACTTGCAAATGGAGAGATTGCAAACGCCCTCTAA
 TCGAAAACTAATGGGAAACTGGAGAAATTGTCTGGATAAAGGGCGAGATTGGCACA
 GTGCGCAAAGTATTGTCCATGCCCAAGTCATATTGTCAAGAAAACAGAAGTACAGAC
 AGGCGGATTCTCCAAGGAGTCATTTACCAAAAGAAATTGGACAAGCTTATTGCTC
 GTAAAAAAAGACTGGATCCAAAAAAATATGGTGGTTGATAGTCAAACGGTAGCTTAT
 TCAGTCCTAGTGGTGTAAAGGTGGAAAAGGGAAATCGAAGAAGTTAAATCCGTTAA
 AGAGTTACTAGGGATCACAATTATGGAAAGAAGTTCTTGAAAAAAAATCCGATTGACT
 TTTAGAAGCTAAAGGATATAAGGAAGTTAAAAAGACTTAATCATTAAACTACCTAAA
 TATAGTCCTTTGAGTTAGAAAACGGTCGAAACGGATGCTGGCTAGTGCGGAGAATT
 ACAAAAAGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTGTTATATTAGCTA
 GTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAACAATTGTTGT
 GAGCAGCATAAGCATTATTAGATGAGATTATTGAGCAAATCAGTGAATTGTTCTAAGCG
 TGTTATTAGCAGATGCCAATTAGATAAAGTTCTAGTGCATATAACAAACATAGAG
 ACAAAACCAATACGTGAACAAGCAGAAAATATTATTCAATTATTCAGTTGACGAATCTT
 GGAGCTCCGCTGCTTTAAATATTGATACAACAATTGATCGTAAACGATATACGTC
 TACAAAAGAAGTTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTATGAAA
 CACGCATTGATTGAGTCAGCTAGGAGGTGACTGA

【0089】

Streptococcus thermophilus Cas9タンパク質のアミノ酸配列は、UniProtデータベースで受託番号Q03J16.1の下に見出され得る。Sapranauskas et al., (2011) Nucleic Acids Res. 39: 9275 - 9282も参照。このアミノ酸配列は以下である。

【0090】

MTKPYSIGLDIGTNSGVAVTTDNYKVPSSKKMVLGNTSKYIKKNLLGVLLFDSGITAE
 GRRLKRTARRRYTRRRNRILYLQEIFSTEMATLDDAFFQRLDDFLVPDDKRDSKYPIFGN
 LVEEKAYHDEFPTIYHLRKYLADSTKKADLRLVYLALAHMIKYRGHFLIEGEFNSKNNDI
 QKNFQDFLDTYNAIFESDSLLENSKQLEEIVKDKISKLEKKDRILKLFPGKEKNSGIFSEFLKL
 IVGNQADFRKCFNLDEKASLHFSESYDEDLETLLGYIGDDYSDVFLKAKKLYDAILLSGF
 LTVTDNEATEPLSSAMIKRYNEHKEDLALLKEYIRNISLKYNEVFKDDTKNGYAGYIDGK 10
 TNQEDFYVYLKKLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQIHLQEMRAILDQKA
 KFYPFLAKNKERIEKILTFRIPYYVGPLARGNSDFAWSIRKNEKITPWNFEDVIDKESSAE
 AFINRMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVRFIAESMRDYQFLDSKQKKDIVR
 LYFKDKRKVTDKDIIEYLHAIYGYDIELKGIEKQFNSSLSTYHDLLNIINDKEFLDDSSNE
 AIIIEIIHTLTIFEDREMIKQLSKFENIFDKSVLKLSRRHYTGWGKLSAKLINGIRDEKSG
 NTILDYLIIDGISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKGNIKEVVVKSLPGSPAIKKG
 ILQSIKIVDELVKVMGGRKPEIVVEMARENQYTNQGKSNSQQRLKRLEKSLKELGSKILK
 ENIPAKLSKIDNNALQNDRLYLYLQNGKDMYTGDDLDIDRLSNYDIDHIIPQAFLKDNSI 20
 DNKVLVSSASNRGKSDDVPSLEVVKKRKTFWYQLLKSCLISQRKFDNLTKAERGGLSPED
 KAGFIQRQLVETRQITKHVARLDEKFNNKKDENNRAVRTVKIITLKSTLVSQFRKDFELY
 KVREINDFHHAHDAYLNAVVASALLKKYPKLEPEFVYGDYPKYNFRERKSATEKVYFYS
 NIMNIFKKSISLADGRVIERPLIEVNEETGESVWNKESDLATVRRVLSPQVNVVKKVEEQ
 NHGLDRGKPKGLFNANLSSKPKPNSNENLVGAKEYLDPKKYGGYAGISNSFTVLVKGTE
 KGAKKKITNVLEFQGISILDNRKDKLNFLLEKGYKDIELIELPKYSLFELSDGSRRMLA
 SILSTNNKRGEIHKGQIFLSQKFVKLLYHAKRISNTINENHRKYVENHKKEFEELFYIILE
 FNENYVGAKKNGKLLNSAFQSWQNHSIDECCSFIGPTGSERKGLFELTSRGSAADFEFLG
 VKIPRYRDYTPSSLKDATALHQSVTGLYETRIDLAKLGE

(配列番号3)

30

【0091】

配列番号3はEuropean Nucleotide Archiveで受託番号ABJ66636.1の下に見出される以下のヌクレオチド配列によってコードされ得る。

【0092】

ATGACTAAGCCATACTCAATTGGACTTGATATTGGAACGAATAGTGTTGGATGGCTGT
 AACAACTGATAATTACAAGGTTCCGTCTAAAAAAATGAAAGTCTTAGGAAATACGAGTA
 AAAAGTATATCAAAAGAACCTGTTAGGTGTATTACTCTTGACTCTGAATCACAGCA
 GAAGGAAGAAGATTGAAGCGTACTGCAAGAACGCTTACTAGACGCCGTAATCGTAT
 CCTTTATTCGAGGAAATTTAGCACAGAGATGGCTACATTAGATGATGCTTCTTTC 40
 AAAGACTTGACGATTGTTAGTCCTGATGATAAACGTGATAGTAAGTATCCGATA
 TTTGGAAACTTAGTAGAAGAAAAGCCTATCATGATGAATTCCAACATATCTATCATT
 AAGGAAATATTTAGCAGATAGTACTAAAAAGCAGATTTGCCTAGTTATCTGCAT
 TGGCTCATATGATTAATAGAGGTCACTTCTTAATTGAAGGAGAGTTAATTCAAAA
 AATAATGATATTAGAAGAATTTCAAGACTTTGGACACTTATAATGCTATTTGA
 ATCGGATTTATCACTTGAGAATAGTAAACAACCTGAGGAAATTGTTAAAGATAAGATTA
 GTAAATTAGAAAAGAAAGATCGTATTTAAAACCTTCCCTGGGGAGAAGAATTGGGG
 ATTTTTCAAGAGTTCTAAAGTTGATTGTAGGAAATCAAGCTGATTTAGGAAATGTT
 TAATTTAGACGAAAAGCCTCCTACATTTCCTAAAGAAAGCTATGATGAAGATTTAG
 AGACTTTGTTAGGTTATGGAGATGATTACAGTGATGTCCTTCTCAAAGCAAAGAAA
 CTTTATGATGCTATTCTTATCGGGTTCTGACTGTAAGTAAATGAGACAGAACG 50

ACCCTCTCTTCTGCTATGATAAAGCGATATAATGAACACAAAGAAGATTAGCGTTAC
 TAAAGGAATATATAAGAAATATTCACTAAAAACGTATAATGAAGTATTTAAAGATGAC
 ACCAAAAATGGTTATGCTGGTTATTGATGGAAAAACAAATCAGGAAGATTCTACGT
 ATATCTAAAAAAACTATTGGCTGAATTGAAGGTGCGGATTATTTCTGAAAAAATTG
 ATCGAGAAGATTGGAGAAAGCAACGTACATTGACAATGGTCGATACCATATCAG
 ATTCACTCTCAAGAAATGAGAGCAATTCTGATAAGCAAGCTAAATTTATCCTTCTT
 GGCTAAAATAAAGAAAGAATCGAGAAGATTAAACCTTCCGAATTCTTATTATGTAG
 GTCCACTTGCAGAGGGAGTAGTGATTTGCCTGGCAATAAGAAAAGAAATGAAAAA
 ATTACACCTTCCAAGAAATGAGAGCAATTCTGATAAGCAAGCTAAATTTATCCTTCTT
 TAATCGAATGACTAGTTGATTTGATTTGCCAGAAGAGAAGGTACTTCAAAGCATA
 GTCTCTTATACGAAACTTTAATGTATATAATGAATTAACAAAAGTTAGATTATTGCC
 GAAAGTATGAGAGATTATCAATTAGATAGTAAGCAGAAGAAAGATATTGTTAGACT
 TTATTTAAAGATAAAAGGAAAGTTACTGATAAGGATATTATTGAATATTACATGCAA
 TTTATGGGTATGATGGAATTGAATTAAAAGGCATAGAGAACAGTTAATTCTAGTTA
 TCTACTTATCACGATCTTAAATATTATTAAATGATAAAGAGTTTGGATGATAGTT
 AAATGAAGCGATTATCGAAGAAATTATCCATACTTGCACATTGAAAGATAGAGAGA
 TGATAAAACAACGTCTTCAAAATTGAGAATATTCGATAAATCCGTTGAAAAAG
 TTATCTCGTAGACATTACACTGGCTGGGTAAAGTTACTGCTAAGCTTATTAAATGGTAT
 TCGAGATGAAAATCTGTAATACTATTCTGATTACTTAATTGATGATGGTATTCTA
 ACCGTAATTTCATGCAACTTATTACGATGATGCTCTTCTTTAAAAAGAAGATACAG
 AAAGCACAAATTATTGGTACGAAGATAAAGGTAAATTAAAGAGGTCGTTAGTCTT
 GCCAGGTAGTCCTGCGATTAAAAAGGTATTACAAAGCATAAAATTGTAGATGAAT
 TGGTCAAAGTAATGGGAGGAAGAAAACCGAGTCATTGTTGAGATGGCTCGTGA
 AATCAATATACCAATCAAGGTAAAGTCTAATTCCCACACGCTTGAACAGTTAGAAAA
 ATCTCTCAAAGAGTTAGGTAGTAAGATACTTAAGAAAATATTCTGCAAACACTTCTA
 AAATAGACAATAACGCACCTCAAAATGATCGACTTACTTACTATCTCAAATGGA
 AAAGATATGTATACCGGAGATGATTAGATATTGATAGATTAAGTAATTATGATATTGA
 TCATATTATTCTCAAGCTTTTGAAAGATAATTCTATTGACAATAAGTACTTGTTC
 ATCTGCTAGTAACCGTGGTAAATCAGATGATGTTCAAGTTAGAGGTTGTCAAAAAA
 GAAAGACATTGGTATCAATTATTGAAATCAAAATTAAATTCTCAACGAAAATTGAT
 AATCTGACAAAGCTGAACGGGGAGGATTGTCACCTGAGGACAAAGCTGGTTATTCA
 ACGCCAGTTGGTAAACACGTCAAATAACAAACATGTAGCTCGTTACTTGATGAGA
 AATTAAATAATAAAAAGATGAAAATAATAGAGCGGTACGAACAGTAAAATTATTAC
 TTGAAATCTACCTAGTTCTCAATTCTGAAGGATTGAACTTTATAAGTTCTGGA
 AATCAATGATTTCATCATGCTCATGATGCTTACTTGAATGCCGTTAGCAAGTGCTT
 TACTTAAGAAATACCCTAAACTAGAGCCAGAATTGTCACGGTATTATCCTAAATAC
 AATAGTTTAGAGAAAGAAAGTCCGCTACAGAAAAGGTATATTCTATTCAAATATCAT
 GAATATCTTAAAAATCTATTCTTAGCTGATGGTAGAGTTATTGAAAGACCCTTA
 TTGAGGTTAAATGAGGAGACCGGCGAATCCGTTGGAATAAGAATCTGATTAGCAACT
 GTAAGGAGAGTACTCTTATCCGCAAGTAAATGTTGTAAAAAGTTGAGGAACAGAA
 TCACGGATTGGATAGAGGAAAACCAAGGGATTGTTAATGCAAATCTTCTCAAAGC
 CAAAACCAAATAGTAATGAAAATTAGTAGGTGCTAAAGAGTATCTGACCCCCAAAAG
 TATGGGGGTATGCTGAAATTCTAATTCTTACTGTTCTGTTAAAGGGACAATTGA
 AAAAGGTGCTAAGAAAAAATAACAAATGTACTAGAATTCAAGGTATTCTATTAG
 ATAGGATTAATTAGAAAAGATAAACTTAATTCTGAAAAAGGTATAAGAT
 ATTGAGTTAATTATTGAACTACCTAAATATAGTTATTGAACTTCAAGGTTACAG
 TCGTATGTTGGCTAGTATTGTCACAGAATAAGAGGGAGAGATTACAAAGGAA
 ATCAGATTCTTCTTCAAGGTTGAAATTACTTATCATGCTAAGAGAATAAGT
 AACACAATTAATGAGAATCATAGAAAATGTTGAGAACCATAAAAAGAGTTGAAGA
 ATTATTACTACATTGAGTTAATGAGAATTATGTTGGAGCTAAAAGAATGGTA

10

20

30

40

50

AACTCTTAAACTCTGCCTTCAATCTGGAAAATCATAGTATAGATGAACCTGTAGT
 AGTTTATAGGACCTACCGGAAGTGAAAGAAAGGGCTATTGAATTAACCTCTCGTGG
 AAGTGCTGCTGATTTGAATTTAGGTGTTAAAATTCCAAGGTATAGAGACTATACCC
 CATCATCCCTATTAAAAGATGCCACACTTATTCAATCTGTTACAGGCCTATGAA
 ACACGAATAGACCTTGCCAAACTAGGAGAGGGTAA

(配列番号4)

【0093】

核局在化シグナル (GGS G P P K K K R K V ; 配列番号5) をそのC末端に含むCas9タンパク質の一例は以下のアミノ酸配列を有する。

【0094】

MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAVKVDDSSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAIDLRIYLALAHMIKFRGHFIEGDLNPNDSDVDKLFQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDLRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSNGYAGYIDGGASQEEFYKFIFKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNRREKIEKILTFRIPYYVGPLARGNSRFAMTRKSEETITPWNFEVVVDKGASAQSFIERMNTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKVYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRTGWGRSLRKLINGIRDKQSGKTILDALKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMDYVDQELDINRLSDYDHDIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTLALKYPKLESEFVYGDYKVDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSFEKNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKMLASAGELKQGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDSQLGGDGGSPPKKKRKV

(配列番号6)

【0095】

いくつかの実施形態において、Cas9ニッカーゼは、ガイド配列(複数可)、例えばDNA標的のセンス及びアンチセンス鎖をそれぞれ標的とする2つのガイド配列とともに用いられ得る。この組合せにより、両方の鎖がニックされ、非相同末端結合(NHEJ)を誘発するのに用いられることが可能となる。

【0096】

さらなる例として、すべてのDNA切断活性が実質的に欠失している変異Cas9を生成するように、Cas9の2つ以上の触媒ドメイン(RuvC I、RuvC II、及びRuvC III)が変異させられ得る。いくつかの実施形態において、D10A変異をH840A、N854A、またはN863A変異のうちの1つ以上と組み合わせて、すべてのDNA切断活性が実質的に欠失しているCas9酵素を生成する(ここでアミノ酸番号は配列番号1の通りである)。いくつかの実施形態において、CRISPR酵素は、変異酵素のDNA切断活性がその非変異形態に対して約25%、10%、5%、1%、0.1%、0.01%未満、またはこれより低い場合、すべてのDNA切断活性が実質的に欠失していると考えられる。他の変異が有用である場合もあり、Cas9または他のCRISPR酵素がS. pyogenes以外の種由来の場合、類似の効果を得るために対応

10

20

30

40

50

するアミノ酸の変異がなされ得る。

【0097】

いくつかの実施形態において、C R I S P R 酵素をコードする酵素コード配列は、特定の細胞、例えば真核細胞などでの発現に最適化されたコドンである。真核細胞は、特定の生物、例えば限定はしないがヒト、マウス、ラット、ウサギ、イヌ、または非ヒト霊長類をはじめとした哺乳類などのものまたは由来するものであり得る。一般に、コドン最適化は、対象とする宿主細胞における発現を強化するために、元のアミノ酸配列を維持しながら、元の配列の少なくとも1つのコドン（例えば約1、2、3、4、5、10、15、20、25、50以上、またはこれより多くのコドン）を、その宿主細胞の遺伝子に用いられる頻度がより高いまたは最も高いコドンで置換することによって、核酸配列を改変するプロセスのことをいう。様々な種が特定のアミノ酸の決まったコドンに特定のバイアスを示す。コドンバイアス（生物間でのコドン使用頻度の差）はメッセンジャーRNA（mRNA）の翻訳の効率と相関があることが多く、そしてまた、この効率はとりわけ、翻訳されるコドンの特性及び特定のトランスクレーバーRNA（tRNA）分子の利用可能性に依存すると考えられている。細胞内で選ばれたtRNAが圧倒的多数を占めているのは一般に、ペプチド合成に使用される頻度が最も高いコドンを反映している。したがって、コドン最適化に基づいて、所与の生物における最適遺伝子発現に遺伝子を適合させることができる。コドン使用頻度表は、例えば「Codon Usage Database」で容易に利用可能であり、これらの表は様々な方法で応用することができる。Nakamura, Y., et al. "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28: 292 (2000) を参照。特定の宿主細胞における発現に対して特定の配列をコドン最適化するコンピュータアルゴリズムも利用可能であり、例えばGene Forge (Aptagen; Jacobus, Pa.) なども利用可能である。いくつかの実施形態において、C R I S P R 酵素をコードする配列における1つ以上のコドン（例えば1、2、3、4、5、10、15、20、25、50以上、またはすべてのコドン）が、特定のアミノ酸に対して使用される頻度が最も高いコドンと一致する。

【0098】

一般に、ガイド配列は、標的配列とハイブリダイズして標的配列へのC R I S P R 複合体の配列特異的結合を誘導するための、標的ポリヌクレオチド配列との十分な相補性を有する任意のポリヌクレオチド配列である。いくつかの実施形態において、ガイド配列とそれに対応する標的配列との間の相補度は、好適なアライメントアルゴリズムを用いて最適アライメントした場合、約50%、60%、75%、80%、85%、90%、95%、97.5%、99%以上またはこれより高い。最適アライメントは、配列をアライメントする任意の好適なアルゴリズムを用いて決定され得るが、その非限定例としては、Smith-Watermanアルゴリズム、Needleman-Wunschアルゴリズム、Burrows-Wheeler変換に基づいたアルゴリズム（例えばBurrows Wheeler Aligner）、Clustal W、Clustal X、BLAT、Novoalign (Novocraft Technologies)、ELAND (Illumina, San Diego, Calif.)、SOAP (soap.genomics.org.cnで入手可能)、及びMaq (maq.sourceforge.netで入手可能) が挙げられる。いくつかの実施形態において、ガイド配列は約5、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、75以上またはこれより多いヌクレオチドの長さである。いくつかの実施形態において、ガイド配列は約75、50、45、40、35、30、25、20、15、12未満、またはこれより少ないヌクレオチドの長さである。標的配列へのC R I S P R 複合体の配列特異的結合を誘導するガイド配列の能力は、任意の好適なアッセイによって評価され得る。例えば、試験しようとするガイド配列を含め、C R I S P R 複合体を形成するのに十

10

20

30

40

50

分な C R I S P R システムのコンポーネントが、例えば C R I S P R 配列のコンポーネントをコードするベクターの形質移入などによって、対応する標的配列を有する宿主細胞に与えられ、その後、例えば本明細書に記載の Surveyor アッセイなどによって標的配列内の選択的切断が評価され得る。標的ポリヌクレオチド配列の切断は、標的配列、試験しようとするガイド配列を含む C R I S P R 複合体のコンポーネント及び試験ガイド配列とは異なる対照ガイド配列を準備し、試験用及び対照用のガイド配列の反応間の標的配列における結合または切断率を比較することによって、試験管内でも同様に評価され得る。他のアッセイも可能であり、当業者であれば考え付くであろう。

【 0 0 9 9 】

ガイド配列は任意の標的配列を標的化するように選択され得る。いくつかの実施形態において、標的配列は細胞のゲノム内の配列である。例示的な標的配列としては、標的ゲノムに特有のものが挙げられる。例えば、*S. pyogenes Cas9*について、ゲノム中の特有の標的配列には、M M M M M M M M N N N N N N N N N N N N X G G M M M M M M M N N N N N N N N N N N N N X X A G A A W 型の Cas9 標的部位が含まれ得るが、この場合、N N N N N N N N N N N N N X G G (N は A 、 G 、 T 、または C であり、 X はデオキシヌクレオチドとすることができる) はゲノム中で一度しか出現しない。ゲノム中の特有の標的配列には、M M M M M M M M N N N N N N N N N N N N X G G 型の *S. pyogenes Cas9* 標的部位が含まれ得るが、この場合、N N N N N N N N N N N N X G G (N は A 、 G 、 T 、または C であり、 X はデオキシヌクレオチドとすることができる) はゲノム中で一度しか出現しない。 *S. thermophilus CRISPR1 Cas9* について、ゲノム中の特有の標的配列には、M M M M M M M M N N N N N N N N N N N N N X A G A A W (配列番号 7) 型の Cas9 標的部位が含まれ得るが、この場合、N N N N N N N N N N N N N X A G A A W (配列番号 8) (N は A 、 G 、 T 、または C であり、 X はデオキシヌクレオチドとでき、 W は A または T である) はゲノム中で一度しか出現しない。ゲノム中の特有の標的配列には、M M M M M M M M M N N N N N N N N N N N X A G A A W (配列番号 9) 型の *S. thermophilus CRISPR1 Cas9* 標的部位が含まれ得るが、この場合、N N N N N N N N N N N N N X X A G A A W (配列番号 10) (N は A 、 G 、 T 、または C であり、 X はデオキシヌクレオチドとでき、 W は A または T である) はゲノム中で一度しか出現しない。 *S. pyogenes Cas9* について、ゲノム中の特有の標的配列には、M M M M M M M M N N N N N N N N N N N X G G X G 型の Cas9 標的部位が含まれ得るが、この場合、N N N N N N N N N N N N N X G G X G (N は A 、 G 、 T 、または C であり、 X はデオキシヌクレオチドとすることができる) はゲノム中で一度しか出現しない。ゲノム中の特有の標的配列には、M M M M M M M M M N N N N N N N N N N N N N X G G X G 型の *S. pyogenes Cas9* 標的部位が含まれ得るが、この場合、N N N N N N N N N N N N N X G G X G (N は A 、 G 、 T 、または C であり、 X はデオキシヌクレオチドとすることができる) はゲノム中で一度しか出現しない。これらの配列のそれぞれにおいて、「 M 」は A 、 G 、 T 、または C であってよく、配列を特有なものとして識別する際に考慮する必要はない。

【 0 1 0 0 】

いくつかの実施形態において、ガイド配列内の二次構造の程度を減少させるようにガイド配列を選択する。二次構造は、任意の好適なポリヌクレオチドフォールディングアルゴリズムによって確認され得る。いくつかのプログラムは最小ギブス自由エネルギーの計算に基づいている。そのようなアルゴリズムの 1 つの例は mFold であり、 Zuker 及び Stiegler (Nucleic Acids Res. 9 (1981), 133 - 148) によって記載されている。別の例のフォールディングアルゴリズムは、 University of Vienna の Institute for Theoretical Chemistry で開発された、重心構造予測アルゴリズムを用いたオンラインウェブサーバー RNAfold である (例えは A. R. Gruber et al. , 2008 , Cell 106 (1) : 23 - 24 ; 及び PA Carr and G

10

20

30

40

50

M Church, 2009, Nature Biotechnology 27(12) : 1151-62を参照)。米国特許出願第61/836,080号(参照により本明細書に援用)にさらなるアルゴリズムが見出され得る。

【0101】

本主題の態様は、CRISPR/CRISPR/Prevotella及びFrancisella 1由来CRISPR(Cpf1)遺伝子編集複合体またはそのコンポーネント(例えばCpf1タンパク質)の送達に関する。ヒトコドン最適化Cpf1ファミリータンパク質の例を以下に示す。

【0102】

ヒトコドン最適化Cpf1ファミリータンパク質

10

Cpf1ファミリータンパク質配列の非限定例、及びCRISPR/Cpf1遺伝子編集の態様が、Zetsche et al., Cell 163, 759-771, October 22, 2015に記載されており、その内容全体を参照により本明細書に援用する。

【0103】

NLS及びHAタグを含むFrancisella tularensis subsp. Novicida U112(FnCpf1; pY004) :

【0104】

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDNLQKDFSAKDTIKKQISEYIKDSEKFKNLFNQLNIDAKKGQESDLILWLQSKDNGIELFKANSDTIDEALEIIKSFKGWTTYFKGFHENRKNVYSSNDIPTSIYYRIVDDNLPKFLENKAKYESLKDCAPEAINYEQIKDLAEELTFIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLAEEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKDLLQASAEDDVKAIKDLDQTNNLLHKLKIFHISQEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNDNPKKESVFYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLKEKANDVHILSIDRGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIEMKEGYLSQVVHEIAKLVIEYNAIVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQLTAPFETFKKMKGQTGIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKHNWDTREVYPTKELEKLLKDSIEYGHGECIKAACGESAIDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGKKLNVIKNEEYFEFVQNRNNKRPAATKKAGQAKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA

20

(配列番号11)

30

【0105】

配列番号11は、核局在化シグナル(KRPAATTKKAGQAKKK)(配列番号12)、その次にグリシン-セリンリンカー(GS)、その次にヒトイソフルエンザヘマグルチニン(HA)タグ(YPYDVPDYAYPYDVPDYAYPYDVPDYA) (配列番号13)を含む。

【0106】

配列番号11は以下のヌクレオチド配列によってコードされ得る。

【0107】

40

50

ATGAGCATCTACCAGGAGTTCGTCAACAAGTATTCACTGAGTAAGACACTGCGGTTCGA
 GCTGATCCCACAGGGCAAGACACTGGAGAACATCAAGGCCGAGGCCTGATTCTGGACG
 ATGAGAACGGCAAAAGACTATAAGAAAGCCAAGCAGATCATTGATAATACCAACAG
 TTCTTATCGAGGAAATTCTGAGCTCCGTGATCAGTGAGGATCTGCTGCAGAATTA
 CTCAGACGTGTACTTCAAGCTGAAGAAGAGCGACGATGACAACCTGCAGAAGGACTTC
 AGTCCGCCAAGGACACCATCAAGAAACAGATTAGCGAGTACATCAAGGACTCCGAAAAG
 TTTAAAATCTGTTCAACCAGAACATCTGATCGATGCTAAGAAAGGCCAGGAGTCCGACCT
 GATCCTGTGGCTGAAACAGTCAAGGACAATGGGATTGAACGTGTTCAAGGCTAACTCCG
 ATATCACTGATATTGACGAGGCACTGGAAATCATCAAGAGCTCAAGGGATGGACCA
 TACTTTAAAGGCTTCCACCGAGAACCGCAAGAACGTGTACTCCAGCAACGACATTCC
 CTCCATCATCTACCGAACATCGTGATGACAATCTGCCAAAGTTCTGGAGAACAGGCCA
 AATATGAATCTCTGAAGGACAAAGCTCCGAGGCAATTAAATTACGAACAGATCAAGAAA
 GATCTGGCTGAGGAACCTGACATTGATATCGACTATAAGACTAGCGAGGTGAACCAGAG
 GGTCTTTCCCTGGACGAGGTGTTGAAATGCCAATTCAACAATTACCTGAACCA
 CCGCATTACTAAATTCAATACCATCATTGGCGGGAGTTGTGACGGGGAGAACACCAGT
 AAGCGCAAGGGATTAAACGAATACATCAATCTGTATAGCCAGCAGATCAACGACAAAC
 TCTGAAGAAATACAAGATGTCTGCTGTTCAAACAGATCCTGAGTGATACCGAGTCCA
 AGTCTTTGTCATTGATAAAACTGGAAGATGACTCAGACGTGGTCACTACCATGCAGAGC
 TTTTATGAGCAGATGCCGCTTCAAGACAGTGGAGGAAAATCTATTAGGAAACTCT
 GAGTCTGCTGTTGATGACCTGAAAGCCCAGAACGCTGGACCTGAGTAAGATCTACTTCA
 AAAACGATAAGAGTCTGACAGACCTGTACAGCAGGTGTTGATGACTATTCCGTGATT
 GGGACCGCCGTCCTGGAGTACATTACACAGCAGATCGCTCCAAAGAACCTGGATAATCC
 CTCTAAGAAAGAGCAGGAACCTGATCGCTAAGAAAACCGAGAACGGAAAATATCTGAGTC
 TGGAAACAATTAAAGCTGGCACTGGAGGAGTTCAACAAGCACAGGGATTGACAAACAG
 TGCCGCTTGAGGAAATCCTGGCAACTTCGCAAGCCATCCCCATGATTTGATGAGAT
 CGCCCGAGAACAAAGACAATCTGGCTCAGATCAGTATTAAAGTACCAAGAACCGAGGCAAGA
 AAGACCTGCTGCAGGCTTCAGCAGAACAGATGACGTGAAAGCCATCAAGGATCTGCTGGAC
 CAGACCAACAATCTGCTGCACAAGCTGAAAATCTCCATTAGTCAGTCAGAGGATAA
 GGCTAATATCCTGGATAAAGACGAACACTTCTACCTGGTGGTCAAGGAAATGTTACTTC
 AGCTGGCAAACATTGCCCCCTGTATAACAAGATTAGGAACATCACACAGAACCT
 TACTCTGACGAGAACGTTAAACTGAACTTCGAAAATAGTACCCCTGCCAACGGGTGGGA
 TAAGAACAAAGGAGCCTGACAACACAGCTATCCTGTTCATCAAGGATGACAAGTACTATC
 TGGGAGTGTGATAAGAAAAACAATAAGATCTCGATGACAAGGCCATTAGGAGAAC
 AAAGGGGAAGGATAACAAGAAAATCGTGTATAAGCTGCTGCCGGCGAAATAAGATGCT
 GCCTAAGGTGTTCTTCAGCGCCAAGAGTATCAAATTCTACAACCCATCCGAGGACATCC
 TGCGGATTAGAAATCACTCAACACATACTAAGAACGGGAGCCCCAGAAGGGATATGAG
 AAATTGAGTTCAACATCGAGGATTGCAGGAAGTTATTGACTTCTACAAGCAGAGCAT
 CTCCAAACACCCCTGAATGGAAGGATTTGGCTTCCGGTTTCCGACACACAGAGATATA
 ACTCTATCGACGAGTTCTACCGCGAGGTGGAAAATCAGGGGTATAAGCTGACTTTGAG
 AACATTCTGAAAGTTACATCGACAGCGTGGTCAATCAGGGAAAGCTGTACCTGTTCCA
 GATCTATAACAAAGATTTTCAGCATAACAGCAAGGGCAGACCAAACCTGCATACACTGT
 ACTGGAAGGCCCTGTTGATGAGAGGAATCTGCAGGACGTGGTCTATAAAACTGAACGGA
 GAGGCCGAACGTTTACCGGAAGCAGTCTATTCTAAAGAAAATCACTCACCCAGCTAA
 GGAGGCCATCGCTAACAGAACAGGACAATCCTAAAGAAAGAGAGCGTGGTCAATACG
 ATCTGATTAAGGACAAGCGGTTACCGAAGATAAGTTCTTTCCATTGTCCAATCACC
 ATTAACTTCAGTCAAGCGCGCTAACAGTTCAACGACGAGATCAATCTGCTGCTGAA
 GGAAAAAGCAAACGATGTGCACATCCTGAGCATTGACCGAGGAGAGCGGCATCTGGCCT
 ACTATACCTGGTGGATGGCAAAGGAATATCATTAAAGCAGGATACATTCAACATCATT
 GGCAATGACCGGATGAAACCAACTACCACGATAAACTGGCTGCAATCGAGAAGGATAG
 AGACTCAGCTAGGAAGGACTGGAAGAAAATCAACAACTTAAGGAGATGAAGGAAGGC

10

20

30

40

50

TATCTGAGCCAGGTGGTCATGAGATTGCAAAGCTGGTCATCGAATACAATGCCATTGT
 GGTGTTCGAGGATCTGAACCTCGGCTTAAGAGGGGGCGCTTAAGGTGGAAAAACAGG
 TCTATCAGAACGCTGGAGAAAATGCTGATCGAAAAGCTGAATTACCTGGTGTAAAGAT
 AACGAGTCGACAAGACCGGGAGGCCTGAGAGCCTACCAGCTGACAGCTCCCTTGA
 AACTTCAAGAAAATGGAAAACAGACAGGCATCATCTACTATGTGCCAGCCGGATTCA
 CTTCCAAGATCTGCCCGTGACCGGCTTGTCAACCAGCTGTACCCCTAAATATGAGTCA
 GTGAGCAAGTCCCAGGAATTTTCAAGTTGATAAGATCTGTTATAATCTGGACAA
 GGGGTACTCGAGTTCTCGATTACAAGAACTCGGCACAAGGCCGTAAGGGGA
 AATGGACCATTCGCCCTCGGATCTGCCTGATCAACTTCGAAATTCCGATAAAAAC
 CACAATTGGGACACTAGGGAGGTGTACCCAACCAAGGAGCTGGAAAAGCTGCTGAAAGA
 CTACTCTATCGAGTATGGACATGGGAATGCATCAAGGCAGCCATCTGTGGCGAGAGTG
 ATAAGAAATTTCGCCAAGCTGACCTCAGTGCTGAATACAATCCTGCAGATGCGGAAC
 TCAAAGACCGGGACAGAACTGGACTATCTGATTAGCCCCGTGGCTGATGTCAACGGAA
 CTTCTCGACAGCAGACAGGCACCCAAAAATATGCCTCAGGATGCAGACGCCAACGGGG
 CCTACCACATCGGGCTGAAGGGACTGATGCTGCTGGCCGGATCAAGAACAAATCAGGAG
 GGGAAAGCTGAACCTGGTCATTAAGAACGAGGAATACTCGAGTTGTCCAGAATAG
 AAATAACAAAAGGCCGGCCACGAAAAGGCCGGCCAGGCAAAAAGAAAAAGGGGA
 TCCTACCCATACGATGTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATA
 CCCATATGATGTCCCCGACTATGCCTAA

(配列番号 14)

10

【 0 1 0 8 】

N L S 及び H A タグを含む L a c h n o s p i r a c e a e b a c t e r i u m M C
 2 0 1 7 (L b 3 C p f 1 ; p Y 0 0 5) :

【 0 1 0 9 】

MDYGNQFERRAPLTKTITRLKPIGETRETIREQKLLEQDAFRKLVETVTPIVDDCIRKI
 ADNALCHFGTEYDFSCLGNAISKNDSKAIKKETEKVEKLLAKVLTENLPDGLRKVNNDINS
 AFIQDTLTSFVQDDADKRVLIQELKGKTVLMQRFLTTRITALTVWLPDRVFNENFIENA
 EKMRILLDSPLNEKIMKFDPDAEQYASLEFYGQCLSQQDIDSYNLIISGIYADDEVKNPGIN
 EIVKEYNQQIRGDKDESPLPKLKKLHKQILMPVEKAFFVRVLSNDSARSILEKILKDTM
 LPSKIIAMEADAGDIAVYGSRLHELSHVIYGDHGKLSQIYDKESKRISELMETLSPKER
 KESKKRLEGLEEHIRKSTYTFDELNRYAENVMMAAYIAAVEESCAEIMRKEKDLRTLLSKE
 DVKIRGNRHNTLIVKNYFNAWTVFRNLIRILRRKSEAEIDSDFYDVLDLDSVEVLSLT
 NLCRSYITKKIGSDLKPEIATYGSALRPNSRWSPGEKFNVKFHTIVRRDGRLYYFILPKG
 AKPVELEDMDGDIECQLQMRKIPNPTIFLPKLFKDPEAFFRDNP
 EADEFVFLSGMKAPVT
 ITRETYEARYKLYTVGKLRDGEVSEEEYKRALLQVLTAYKEFLENRMIYADLNFGFKDLE
 EYKDSSEFIKVETHNTFCWAKVSSSQLDDLVKSGNGLLFEIWSERLESYYKGNEKVL
 RGYEGVLLSILKDENLVSMTLLNSRPMLVYRPKESSKPMVVHRDGSRVVDRFDKGKYI
 PPEVHDELYRFFNNLLIKEKLGEKARKILDNNKKVKV
 LESERVKWSKFYDEQFAVTF
 KKNA
 DCLDTTKDLNAEVMEQYSES
 NRLLIRNTTDILYYLVL
 DKNGKVLQRSLNIINDGA
 RDVDWKERFRQVTKDRNEGYN
 NEWDYSRTSNDLKEV
 YLN
 YALKEIA
 EAVIEYN
 AILII
 EKM
 SNAFKDKYSFLDDVTFKG
 FETKLLAKLSDLHFRGIKDGE
 PCSFTNPLQLCQN
 DSNKILQDG
 VIFMVPNSMTRSLDPDTGF
 FAINDHN
 IRTKKAKLN
 FLSKFDQLK
 VSSEGCLIM
 KYSGDSL
 PTHNTDN
 RVWN
 CCCC
 NH
 PITNYD
 RETKK
 VEF
 IEEP
 VEEL
 SRV
 LEENG
 IE
 TTEL
 NKL
 NERE
 NVPGK
 VVD
 DAI
 YSL
 VLN
 YL
 RGT
 VSG
 VAG
 QRA
 VY
 YSP
 VTG
 KKY
 DIS
 FIQAM
 NL
 NR
 KCD
 YY
 RI
 GSK
 KER
 GE
 WTDF
 VA
 QLINK
 RPA
 AT
 KK
 AG
 QAK
 KK
 GS
 YPY
 DV
 PDY
 AYP
 DV
 DY
 A
 DY
 A

(配列番号 15)

30

【 0 1 1 0 】

配列番号 15 は、核局在化シグナル (K R P A A T K K A G Q A K K K) (配列番号
 1 2) 、その次にグリシン - セリンリンカー (G S) 、その次に H A タグ (Y P Y D V P
 P D Y A Y P Y D V P D Y A Y P Y D V P
 D Y A

40

50

D Y A Y P Y D V P D Y A Y P Y D V P D Y A) (配列番号 13) を含む。

【 0 1 1 1 】

配列番号 15 は以下のヌクレオチド配列によってコードされ得る。

【 0 1 1 2 】

ATGGATTACGGCAACGGCCAGTTGAGCGGAGAGCCCCCTGACCAAGACAATCACCT
 GCGCCTGAAGCCTATCGCGAGACACGGAGACAATCCCGAGCAGAAGCTGCTGGAGC
 AGGACGCCGCTTCAGAAAGCTGGTGGAGACAGTGACCCCTATCGTGGACGATTGTATC
 AGGAAGATCGCCGATAACGCCCTGTGCCACTTGGCACCGAGTATGACTTCAGCTGTCT
 GGGCAACGCCATCTCTAAGAATGACAGCAAGGCCATCAAGAAGGAGACAGAGAAGGTGG
 AGAAGCTGCTGGCCAAGGTGCTGACCGAGAATCTGCCAGATGGCCTGCGCAAGGTGAAC 10
 GACATCAATTCCGCCGCTTATCCAGGATACTGACCTCTTCGTGCAGGACGATGC
 CGACAAAGCGGGTGTGATCCAGGAGCTGAAGGGCAAGACCGTGCTGATGCAGCGTTCC
 TGACCACACGGATCACAGCCCTGACCGTGCTGGCTGCCGACAGAGTGTGAGAACCTT
 AATATCTTCATCGAGAACGCCGAGAAGATGAGAATCCTGCTGGACTCCCCTCTGAATGA
 GAAGATCATGAAGTTGACCCAGATGCCAGCAGTACGCCCTCTGGAGTTCTATGCC 20
 AGTGCCTGTCTCAGAAGGACATCGATAGCTACACCTGATCATCTCCGGCATCTATGCC
 GACGATGAGGTGAAGAACCCCTGGCATCAATGAGATCGTAAGGAGTACAATCAGCAGAT
 CCGGGCGACAAGGATGAGTCCCCACTGCCAAGCTGAAGAAGCTGCACAAGCAGATCC
 TGATGCCAGTGGAGAAGGCCCTTTGTGCGCGTGTCTAACGACAGCGATGCCGG
 AGCATTGGAGAACGATCCTGAAGGACACAGAGATGCTGCCCTCCAAGATCATCGAGGC
 CATGAAGGAGGCAGATGCAGGCAGATCGCCGTGATGGCAGCCGGCTGCACGAGCTGA
 GCCACGTGATCTACGGCGATCACGGCAAGCTGTCCCAGATCATCTATGACAAGGAGTCC
 AAGAGGATCTCTGAGCTGATGGAGACACTGTCCTCAAAGGAGCGCAAGGAGAGCAAGAA
 GCGGCTGGAGGGCTGGAGGAGCACATCAGAAAGTCTACATACACCTCGACGAGCTGA
 ACAGGTATGCCGAGAACGAGATGTGATGGCAGCATACATCGCAGCAGTGGAGGAGTCTGT
 GCCGAGATCATGAGAACGAGATCTGAGGACCCCTGCTGAGCAAGGAGGACGTGA
 AGATCCGGGGCAACAGACACAATACACTGATCGTAAGAACTACTTTAATGCCCTGGACC
 GTGTTCCGGAACCTGATCAGAACCTGAGGCGCAAGTCCGAGGCCGAGATCGACTCTGA
 CTTCTACGATGTGCTGGACGATTCCGTGGAGGTGCTCTGACATACAAGGGCGAGA 30
 ATCTGTGCCCGAGCTATATCACCAAGAACGATCGGCTCCGACCTGAAGGCCGAGATGCC
 ACATACGGCAGGCCCTGAGGCCTAACAGCCGCTGGTGGCCCCAGGAGAGAACCTTAA
 TGTGAAGTTCCACACCATCGTGGAGAGAGATGGCCGGCTGTACTATTCATCCTGCCA
 AGGGCGCCAAGCCTGTGGAGCTGGAGGACATGGATGGCGACATCGAGTGTCTGCAGATG
 AGAAAGATCCCTAACCCAAACATCTTCTGCCAAGCTGGTGTCAAGGACCCCTGAGGC
 CTTCTTAGGGATAATCCAGAGGCCGACGAGTTCTGAGGCTACAGGTATAAGCTGTATACCGTGGGC
 AACCTGCGCGATGGCGAGGTGTCCGAAGAGGGAGTACAAGCGGCCCTGCTGCAGGTGCT
 GACCGCCTACAAGGAGTTCTGGAGAACAGAACATGATCTATGCCGACCTGAATTTCGGCT
 TTAAGGATCTGGAGGAGTATAAGGACAGCTCCGAGTTATCAAGCAGGTGGAGACACAC
 AACACCTTCATGTGCTGGCCAAGGTGCTAGCTCCAGCTGGACGATCTGGTAAGTC 40
 TGGCAACGGCCTGCTGTTGAGATCTGGAGCGAGCGCCTGGAGTCTACTATAAGTACG
 GCAATGAGAACGGTGCTGCGGGCTATGAGGGCGTGTGCTGAGCATCCTGAAGGATGAG
 AACCTGGTGTCCATCGGGACCCCTGCTGAACAGCCGCCATGCTGGTGTACCGGCCAAA
 GGAGTCTAGCAAGCCTATGGTGGTGCACCGGGATGGCAGCAGAGTGGTGGACAGGTTG
 ATAAGGACGGCAAGTACATCCCCCTGAGGTGCACGACGAGCTGTATCGCTTCTTAAC
 AATCTGCTGATCAAGGAGAACGCTGGCGAGAACGGCCGAGATCCTGGACAACAAGAA
 GGTGAAGGTGAAGGTGCTGGAGAGCGAGAGAGTGAAGTGGTCCAAGTTCTACGATGAGC
 AGTTTGCGTGAACCTCAGCGTGAAGAAGAACGCCGATTGTCTGGACACCACAAAGGAC
 CTGAATGCCGAAGTGTGAGCGAGTATAGCGAGTCCAACAGACTGATCCTGATCAGGAA
 TACCACAGATATCCTGTAATCTGGTGTGACAAGAATGGCAAGGTGCTGAAGCAGA 50

GATCCCTGAACATCATCAATGACGGCGCCAGGGATGTGGACTGGAAGGAGAGGTTCCGC
 CAGGTGACAAAGGATAGAAACGAGGGCTACAATGAGTGGGATTATTCAGGACCTCTAA
 CGACCTGAAGGAGGTGTACCTGAATTATGCCCTGAAGGAGATGCCGAGGCCGTATCG
 AGTACAACGCCATCCTGATCATCGAGAAGATGTCTAACATGCCTTAAGGACAAGTATAGC
 TTCCTGGACGACGTGACCTCAAGGGCTTCAGAGACAAAGCTGCTGCCAAGCTGAGCGA
 TCTGCACCTTAGGGCATCAAGGACGGCGAGCCATGTTCTCACAAACCCCTGCAGC
 TGTGCCAGAACGATTCTAATAAGATCCTGCAGGACGGCGTATCTTATGGTCCAAT
 TCTATGACACGGAGCCTGGACCCGACACCGGCTTCATCTTGCCATCAACGACCACAA
 TATCAGGACCAAGAAGGCAAGCTGAACCTTGAGCAAGTCTGATCAGCTGAAGGTGT
 CCTCTGAGGGCTGCCTGATCATGAAGTACAGCGGCATTCCCTGCCTACACACAAACACC
 GACAATCGCGTGTGAACTGCTGCAATCACCAATCACAAACTATGACCGGGAGAC
 AAAGAAGGTGGAGTTCATCGAGGAGCCCCTGGAGGGAGCTGTCCTCGCTGGAGGAGA
 ATGGCATCGAGACAGACACCGAGCTGAACAAGCTGAATGAGCGGGAGAACGTGCCTGGC
 AAGGTGGTGGATGCCATCTACTCTCTGGTCTGAATTATCTGCGCGGCACAGTGAGCGG
 AGTGGCAGGACAGAGGGCCGTGTACTATAGCCCTGTGACCGGCAAGAAGTACGATATCT
 CCTTATCCAGGCCATGAACCTGAATAGGAAGTGTGACTACTATAGGATCGGCTCCAAG
 GAGAGGGAGAGTGGACCGATTCGTGGCCAGCTGATCAACAAAGGCCGGCCAC
 GAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGATGTTCCAGATT
 ACGCTTATCCCTACGACGTGCTGATTATGCATACCCATATGATGTCCCCGACTATGCC
 TAA

10

【0113】

NLS 及び HA タグを含む Butyrivibrio proteoclasticus (BpCpf1; pY006) :

【0114】

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFI
 EEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFN
 QNLIDAKKGQESDLILWLKQSNDNGIELFKANSDTIDIDEALEIIKSFKGWTTYFKGFHEN
 RKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAELTFID
 YKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTKRKGNEYINLYS
 QQINDKTLKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFYEQIAAFKTVEEKS
 IKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQVFDYSVIGTAVLEYITQQIAPKNLD
 NPSKKEQELIAKKTEKAKYLSLETIKLAEEFNKHRRDIDKQCRFEEILANFAAIPMIFDEIA
 QNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLDQTNNLLHKLKIFHISQSEDKAN
 ILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYSDKEFKLNFENSTLANGWDKNKE
 PDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPKVFF
 SAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCRKFIDFYKQSIKHPREW
 FGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDS VVNQGKLYLFQIYNKDF
 SAYSK GRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKKDNP
 KKESVFYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRG
 ERHLAYYTLDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKE
 MKEGYLSQVHVHEIAKLVIEYNAIVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLV
 FKDNFDFKGGVLRAYQLTAPFETFKKMKGKQTGIIYYVPAGFTSKICPVTFVNQLYPKY
 ESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
 HNWDTREVYPTKELEKLLKDYSEYGHGECIKAACGKESDKKFFAKLTSVLNTILQMRNSK
 TGTELVDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGK
 KLNVIKNEEYFEFVQNRNNKRPAATKKAGQAKKKGSYPYDVPDYAYPYDVPDYAYPY
 DVPDYA

30

(配列番号 17)

【0115】

配列番号 17 は、核局在化シグナル (KRPAAKKAGQAKKK) (配列番号

40

50

12)、その次にグリシン - セリンリンカー (GS)、その次にHAタグ (YPYDVPDYAYPYDVPDYA) (配列番号13) を含む。

【0116】

配列番号17は以下のヌクレオチド配列によってコードされ得る。

【0117】

ATGAGCATCTACCAGGAGTTCGTCAACAAGTATTCACTGAGTAAGACACTGCGGTTCGA
 GCTGATCCCACAGGGCAAGACACTGGAGAACATCAAGGCCGAGGCCTGATTCTGGACG
 ATGAGAACGGGCAAAGACTATAAGAAAGCCAAGCAGATCATTGATAAATACCACAG
 TTCTTATCGAGGAAATTCTGAGCTCCGTGTCATCAGTGAGGATCTGCTGCAGAATT
 CTCAGACGTGTACTTCAAGCTGAAGAAGAGCGACGATGACAACCTGCAGAAGGACTTCA
 AGTCCGCCAAGGACACCATCAAGAAACAGATTAGCGAGTACATCAAGGACTCCGAAAAG
 TTTAAAATCTGTTCAACCAGAACATCTGATCGATGCTAAGAAAGGCCAGGAGTCCGACCT
 GATCCTGTGGCTGAAACAGTCTAAGGACAATGGGATTGAACGTGTTCAAGGCTAACTCCG
 ATATCACTGATATTGACGAGGCACTGGAAATCATCAAGAGCTTCAAGGGATGGACCACA
 TACTTTAAAGGCTTCCACCGAGAACCGCAAGAACGTGTACTCCAGCAACGACATTCTAC
 CTCCATCATCTACCGAATCGTCACTGACAATCTGCCAAGTTCCTGGAGAACAGGCCA
 AATATGAATCTCTGAAGGACAAAGCTCCGAGGCAATTAAATTACGAACAGATCAAGAAA
 GATCTGGCTGAGGAACGTGACATTGACTATCGACTATAAGACTAGCGAGGTGAACCAGAG
 GGTCTTTCCCTGGACGAGGTGTTGAAATGCCAATTCAACAATTACCTGAACCAAGCAGT
 CGGCATTACTAAATTCAATACCATCATTGGCGGGAAAGTTGTGAACGGGAGAACATACC
 AAGCGCAAGGAAATTACGAATACATCAATCTGTATAGCCAGCAGATCAACGACAAAAC
 TCTGAAGAAATACAAGATGTCTGTGCTGTTCAAACAGATCCTGAGTGATACCGAGTCCA
 AGTCTTTGTCATTGATAAAACTGGAAGATGACTCAGACGTGGTCACTACCATGCAGAGC
 TTTTATGAGCAGATGCCGCTTCAAGACAGTGGAGGAAAATCTATTAGGAAACTCT
 GAGTCTGCTGTTGATGACCTGAAAGCCAGAAGCTGGACCTGAGTAAGATCTACTTCA
 AAAACGATAAGAGTCTGACAGACCTGTCACAGCAGGTGTTGATGACTATTCCGTGATT
 GGGACCGCCGTCCTGGAGTACATTACACAGCAGATCGCTCAAAGAACCTGGATAATCC
 CTCTAAGAAAGAGCAGGAACGTGCTAAGAAAACCGAGAAGGCAAATATCTGAGTC
 TGGAAACAATTAAGCTGGCACTGGAGGAGTTCAACAAGCACAGGGATATTGACAACAG
 TGCCGCTTGAGGAAATCCTGCCAACTTCGCAGCCATCCCCATGATTTGATGAGAT
 CGCCCAAGAACAAAGACAATCTGGCTCAGATCAGTATTAGTACCGAGAACAGGGCAAGA
 AAGACCTGCTGCAGGCTTCAGCAGAACAGATGACGTGAAAGCCATCAAGGATCTGCTGGAC
 CAGACCAACAATCTGCTGCACAAGCTGAAAATCTTCCATATTAGTCAGTCAGAGGATAA
 GGCTAATATCCTGGATAAAAGACGAACACTTCTACCTGGTGGAGGAATGTTACTTCG
 AGCTGGCAAACATTGCCCCCTGTATAACAAGATTAGGAACATCACACAGAACGCT
 TACTCTGACGAGAAGTTAAACTGAACTTCGAAAATAGTACCCCTGCCAACGGGTGGGA
 TAAGAACAAAGGAGCCTGACAACACAGCTATCCTGTTCATCAAGGATGACAAGTACTATC
 TGGGAGTGTGATAAGAAAAACAATAAGATCTTGATGACAAAGCCATTAAGGAGAAC
 AAAGGGGAAGGATACAAGAAAATCGTGTATAAGCTGCTGCCGGCGCAAATAAGATGCT
 GCCTAAGGTGTTCTCAGGCCAAGAGTATCAAATTCTACAACCCATCCGAGGACATCC
 TCGGGATTAGAAATCACTCAACACATACTAAGAACGGGAGCCCCAGAAGGGATATGAG
 AAATTTGAGTTCAACATCGAGGATTGCAGGAAGTTATTGACTTCTACAAGCAGAGCAT
 CTCCAAACACCCCTGAATGGAAGGATTTGGCTTCCGGTTTCCGACACACAGAGATATA
 ACTCTATCGACGAGTTCTACCGCGAGGTGGAAAATCAGGGGTATAAGCTGACTTTGAG
 AACATTCTGAAAGTTACATCGACAGCGTGGTCAATCAGGGAAAGCTGTACCTGTTCCA
 GATCTATAACAAAGATTTTCAGCATAACAGCAAGGGCAGACCAAACCTGCATACACTGT
 ACTGGAAGGCCCTGTTGATGAGAGGAATCTGCAGGACGTGGTCTATAAAACTGAACGGA
 GAGGCCGAACGTGTTTACCGGAAGCAGTCTATTCTAAGAAAATCACTCACCCAGCTAA
 GGAGGCCATCGCTAACAGAACAAAGGACAATCTAACAGAAAGAGAGCGTGGTCAATACG
 ATCTGATTAAGGACAAGCGGTTACCGAAGATAAGTTCTTTCCATTGTCCAATCACC
10
20
30
40
50

ATTAACCTCAAGTCAAGCGCGCTAACAAAGTTAACGACGAGATCAATCTGCTGCTGAA
 GGAAAAAGCAAACGATGTGCACATCCTGAGCATTGACCGAGGGAGAGCGGCATCTGGCCT
 ACTATAACCCTGGTGGATGGCAAAGGAAATATCATTAAGCAGGATACATTCAACATCATT
 GGCAATGACCGGATGAAAACCAACTACCACGATAAACTGGCTGCAATCGAGAAGGATAG
 AGACTCAGCTAGGAAGGACTGGAAGAAAATCAACAAACATTAAGGAGATGAAGGAAGGC
 TATCTGAGCCAGGTGGTCATGAGATTGCAAAGCTGGTCATCGAATACAATGCCATTGT
 GGTGTTCGAGGATCTGAACCTCGGCTTAAGAGGGGCGCTTAAGGTGGAAAAACAGG
 TCTATCAGAAGCTGGAGAAAATGCTGATCGAAAAGCTGAATTACCTGGTGTAAAGAT
 AACGAGTTGACAAGACCGGAGGCGCTTGAGAGCCTACCAAGCTGACAGCTCCCTTGA
 AACTTCAAGAAAATGGAAAACAGACAGGCATCATCTACTATGTGCCAGCCGGATTCA 10
 CTTCCAAGATCTGCCCCGTGACCGGCTTGTCAACCAGCTGTACCCCTAAATATGAGTCA
 GTGAGCAAGTCCCAGGAATTTTCAGCAAGTTGATAAGATCTGTTATAATCTGGACAA
 GGGGTACTTCGAGTTTCCTTCGATTACAAGAACTTCGGCGACAAGGCCGCTAAGGGGA
 AATGGACCATTCGCTCCTCGGATCTCGCCTGATCAACTTCGAAATTCCGATAAAAAC
 CACAATTGGGACACTAGGGAGGTGTACCCAACCAAGGAGCTGGAAAAGCTGCTGAAAGA
 CTACTCTATCGAGTATGGACATGGGAATGCATCAAGGCAGCCATCTGTGGCGAGAGTG
 ATAAGAAATTTTCGCCAAGCTGACCTCAGTGCTGAATACAATCCTGCAGATGCGGAAC
 TCAAAGACCGGGACAGAACTGGACTATCTGATTAGCCCCGGCTGATGTCAACGGAAA
 CTTCTCGACAGCAGACAGGCACCCAAAAATATGCCTCAGGATGCAGACGCCAACGGG
 CCTACCACATCGGGCTGAAGGGACTGATGCTGCTGGCCGGATCAAGAACAAATCAGGAG
 GGGAAAGACTGAACCTGGTCATTAAGAACGAGGAATACTTCGAGTTGTCCAGAATAG
 AAATAACAAAAGGCCGGCCACGAAAAAGGCCGGCCAGGCAAAAAGAAAAAGGGGA
 TCCTACCCATACGATGTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATA
 CCCATATGATGTCCCCGACTATGCCTAA

(配列番号 18)

【 0 1 1 8 】

N L S 及び H A タグを含む P e r e g r i n i b a c t e r i a b a c t e r i u m
 G W 2 0 1 1 _ G W A _ 3 3 _ 1 0 (P e C p f 1 ; p Y 0 0 7) :

【 0 1 1 9 】

MSNFFKNFTNLYELSKTLRFELKPVGDTLTNMKDHLLEYDEKLQTFKDQNIDDAYQALK 30
 PQFDEIHEEFITDSLESKKAKEIDFSEYLDLFQEKKELNDSEKKLRNKIGETFNKAGEWKW
 KEKYPQYEWKGSKIANGADILSCQDMLQFIKYKNPEDEKIKNYIDDTLKGFFTYYFGGFN
 QNRANYYETKKEASTAVATRIVHENLPKFCDNVIQFKHIIKRKKDGTVKTERKTEYLNA
 YQYLKNNNKITQIKDAETEKMIESTPIAEKIFDVYYFSSCLSQQKIEEYNRIIGHYNLLINL
 YNQAKRSEGKHLSANEKKYKDLPKFKTLYKQIGCGKKDLFYTIKCDTEEEANKSRNEGK
 ESHSVEEINKAQAINEAKYFKSNNDCEINTVPDFINYILTKENYEGVYWSKAAMNTISDK
 YFANYHDLQDRLKEAKVFQKADKKSEDDIKIPEAIELSGLFGVLDSDLADWQTLFKSSILS
 NEDKLKIITDSQTPSEALLKMIFNDIEKNMESFLKETNDIITLKKYKGNKEGTEKIKQWFD
 YTLAINRMLKYFLVKENKIKGNSLDTNISEALKTLIYSDDAEWFKWYDALRNLYLTQKPQD
 EAKENKLKLNFDNPSLAGGWVDVNKECSNFCVILKDKNEKKYLAIMKKGENTLFQKEWTE
 GRGKNLTKKSNPFFEINNCEILSKMEYDFWADVSCKMIPKCSTQLKAVVNHFKQSDNEFIF
 PIGYKVTSGEKFREECKISKQDFELNNKVFNKNELSVTAMRYDLSSTQEKQYIKAFQKEY
 WELLFKQEKRDTKLTNNEIFNEWINFCNKKYSELLSWERKYKDALTNWINFCKYFLSKY
 PKTTLFNYSFKESENNSLDEFYRDVDICSYKLNINTTINKSILDRLVEEGKLYLFEIKNQD
 SNDGKSIGHKNNLHTIYWNAIFENFDNRPKLNGEAEIFYRKAI SKDKLGIVKGKKTNGTE
 IIKNYRFSKEKFILEHVPITLNFCNSNEYVNDIVNTKFYNFSNLHFLGIDRGEKHLAYYSLVN
 KNGEIVDQGTLNLPFTDKDGNQRSIKKEKYFYNKQEDKWEAKEVDCWNYNNDLLDAMAS
 NRDMARKNWQRIGTIKEAKNGYVSLVIRKIA DLAVNNERPAFIVLEDLNTGFKRSRQKID
 KSVYQKFELALAKKLNFVLDKNAKRDEIGSPTKALQLTPPVNNYGDIEKKQAGIMLYTR
 ANYTSQTDPATGWRKTIYLKAGPEETTYKKDGKIKNKSVDQIETFTDIGFDGKDYYFEY 50

DKG EFVDEKTGEIKPKKWR LYS GENGKS LDRFRG EREKD KYEW KIDK IDIV KILD LFVN F
 DKNISLLKQLKEGV ETRNNE HGTE SLRFA INLI QQIR NTGN ERDND FILSP VRDENG K
 HFDSREYWDKETKGEK ISMPSSGDANGAFNIARKGIIMNAHILANS DSKDLSLFVSDEEW
 DLHLNNKTEWKKQLNIFSSRKAMAKRKKKRPAATKKAGQAKKKGSYPYDVPDYAYPY
 DVPDYAYPYDVPDYA

(配列番号 19)

【 0120 】

配列番号 19 は、核局在化シグナル (K R P A A T K K A G Q A K K K) (配列番号 12) 、その次にグリシン - セリンリンカー (G S) 、その次に H A タグ (Y P Y D V P D Y A Y P Y D V P D Y A Y P Y D V P D Y A) (配列番号 13) を含む。 10

【 0121 】

配列番号 19 は以下のヌクレオチド配列によってコードされ得る。

【 0122 】

ATGTCCAAC TTCTTAAGAATT CACCAAC CTGTATGAGCTGTCCAAGACACTGAGGTT
 TGAGCTGAAGCCCGTGGCGACACCCTGACAAACATGAAGGACCACCTGGAGTACGATG
 AGAAGCTGCAGACCTTCCTGAAGGATCAGAATATCGACGATGCCTATCAGGCCCTGAAG
 CCTCAGTT CGACGAGATCCACGAGGAGTTATCACAGATTCTCTGGAGAGCAAGAACGC
 CAAGGAGATCGACTTCTCCGAGTACCTGGATCTGTT CAGGAGAAGAAGGAGCTGAACG
 ACTCTGAGAAGAAGCTCGCAACAAGATCGCGAGACATTCAACAAGGCCGGCGAGAAG
 TGGAGAAGGAGAAGTACCCCTCAGTATGAGTGGAGAAGGGCTCCAAGATGCCAATGG 20
 CGCCGACATCCTGTCTTGC CAGGATATGCTGCAGTTATCAAGTATAAGAACCCAGAGG
 ATGAGAAGATCAAGAATTACATCGACGATACACTGAAGGGCTTCTTACCTATTCGGC
 GGCTTAATCAGAACAGGGCCAAC TACTATGAGACAAAGAAGGAGGCCACCGCAGT
 GGCAACAAGGATCGTCACGAGAACCTGCCAAAGTTCTGTGACAATGTGATCCAGTTA
 AGCACATCATCAAGCGGAAGAAGGATGGCACCGTGGAGAAAACCGAGAGAAAGACCGA
 GTACCTGAACGCC TACCA CGTATCTGAAGAACAAATAACAAGATCACACAGATCAAGGACG
 CCGAGACAGAGAAGATGATCGAGTCTACACCCATGCCGAGAAGATCTCGACGTGTAC
 TACTTCAGCAGCTGCC TGAGCCAGAAGCAGATCGAGGAGTACAACCGGATCATCGGCCA
 CTATAATCTGCTGATCAACCTGTATAACCAGGCCAAGAGATCTGAGGGCAAGCACCTGA
 GCGCCAACGAGAAGTATAAGGACCTGCC TAAGTTCAAGACCTGTATAAGCAGATC 30
 GGCTGCGGCAAGAAGAAGGACCTGTTTACACAATCAAGTGTGATACCGAGGAGGAGGC
 CAATAAGTCCC GGAA CGAGGGCAAGGAGTCCC ACTCTGTGGAGGAGATCATCAACAGG
 CCCAGGAGGCCATCAATAAGTACTTCAAGTCTAATAACGACTGTGAGAATATCAACACC
 GTGCCCGACTTCATCAACTATATCCTGACAAAGGAGAATTACGAGGGCGTGTATTGGAG
 CAAGGCCGCCATGAACACCATCTCCGACAAGTACTTCGCCAATTATCACGACCTGCAGG
 ATAGACTGAAGGAGGCCAAGGTGTT CAGAAGGCCATAAGAAGTCCGAGGACGATATC
 AAGATCCCAGAGGCCATCGAGCTGTCTGCCGTGCTGGACAGCCTGCCGA
 TTGGCAGACCACACTGTTAAGTCTAGCATCCTGAGGCCCTGCTGAAGATGATCTCAATGACATCGAG
 TCACAGATTCCCAGACCCCTCTGAGGCCCTGCTGAAGATGATCTCAATGACATCGAG
 AAGAACATGGAGTCCTTCTGAAGGAGACAAACGATATCATCACCCCTGAAGAAGTATAA 40
 GGGCAATAAGGAGGGCACCGAGAAGATCAAGCAGTGGTTGACTATACACTGCCATCA
 ACCGGATGCTGAAGTACTTCTGGTGAAGGAGAATAAGATCAAGGGCAACTCCCTGGAT
 ACCAATATCTCTGAGGCCCTGAAAACCCCTGATCTACAGCGACGATGCCGAGTGGTTCAA
 GTGGTACGACGCCCTGAGAAACTATCTGACCCAGAAGCCTCAGGATGAGGAGCAAGGAGA
 ATAAGCTGAAGCTGAATT CGACAACCCATCTGGCCGGCGTGGATGTGAACAAG
 GAGTGCAGCAATT TCGTGTACCTGAAGGACAAGAACGAGAAGAAGTACCTGCCAT
 CATGAAGAAGGGCGAGAATAACCTGTTCCAGAAGGGAGTGGACAGAGGGCCGGGCAAG
 AACCTGACAAAGAAGTCTAATCCACTGTTCGAGATCAATAACTGCGAGATCCTGAGCAA
 GATGGAGTATGACTTTGGGCCGACGTGAGCAAGATGATCCCCAAGTGTAGCACCCAGC
 TGAAGGCCGTGGTGAACCAC TTCAAGCAGTCCGACAATGAGTTCATCTTCCATCGGC 50

TACAAGGTGACAAGCGCGAGAAGTTAGGGAGGAGTGCAAGATCTCCAAGCAGGACTT
 CGAGCTGAATAACAAGGTTTAATAAGAACGAGCTGAGCGTGACCGCCATCGCTACG
 ATCTGTCCCTCACACAGGAGAACGAGTATATCAAGGCCTCCAGAAGGAGTACTGGAG
 CTGCTGTTAACGAGGAGAACGGACACCAAGCTGACAATAACGAGATCTTCACGA
 GTGGATCAATTGGATCAACAAAGAAGTATAGCGAGCTGCTGTCCTGGAGAGAAAGTACA
 AGGATGCCCTGACCAATTGGATCAACTTCTGTAAGTACTTCTGAGCAAGTATCCCAG
 ACCACACTGTTCAACTACTCTTTAACGGAGAGCAGAACATTATAACTCCCTGGACGAGTT
 CTACCGGACGTGGATATCTGTTTACAAGCTGAATATCAACACCACAATCAATAAGA
 GCATCCTGGATAGACTGGTGGAGGGCAAGCTGTACCTGTTGAGATCAAGAACATCAG
 GACAGCAACGATGGCAAGTCCATCGGCCACAAGAACAACTGCACACCATCTACTGGAA 10
 CGCCATCTCGAGAACAGGCTAACGCTGAATGGCGAGGCCAGATCTTCT
 ATCGCAAGGCCATCTCCAAGGATAAGCTGGCATCGTGAAGGGCAAGAAAACCAAGAAC
 GGCACCGAGATCATCAAGAATTACAGATTCAAGGAGAACAGTTATCCTGCACGTGCC
 AATCACCTGAACCTCTGCTCCAATAACGAGTATGTGAATGACATCGTAACACAAAGT
 TCTACAATTGGTGAATAAGAACGGCGAGATCGTGGACCAGGGCACACTGAACACTGCC
 TACTATTCTCTGGTGAATAAGAACGGCGAGATCGTGGACCAGGGCACACTGAACACTGCC
 TTTCACCGACAAGGATGGCAATCAGCGCAGCATCAAGAACAGGAGAACAGTACTTTATAACA
 AGCAGGAGGACAAGTGGAGGCCAAGGAGGTGGATTGTTGAATTATAACGACCTGCTG 20
 GATGCCATGGCCTCTAACCGGGACATGCCAGAAAGAACATTGGCAGAGGATCGGACCAT
 CAAGGAGGCCAACGCGTACGTGAGCCTGGTCATCAGGAAGATGCCGATCTGGCG
 TGAATAACGAGCGCCCCGCCCTCATCGTGTGGAGGACCTGAATACAGGCTTAAGCGG
 TCCAGACAGAACAGATCGATAAGAGCGTGTACAGAACAGTTCGAGCTGCCCTGGCAAGAA
 GCTGAACTTCTGGTGGACAAGAACGCGATGAGATCGGCTCCCTACAAAGG
 CCCTGCAGCTGACCCCCCTGTGAATAACTACGGCGACATTGAGAACAGAACAGGCC
 GGCATCATGCTGTACCGGGCCAATTATAACCTCTCAGACAGAACAGATCCAGCACAGGCTG
 GAGAAAGACCATCTATCTGAAGGCCGGCCCGAGGAGAACACATACAAGAACAGGCC
 AGATCAAGAACAAAGAGCGTGAAGGACAGATCATCGAGAACATTACCGATATCGGCTT
 GACGGCAAGGATTACTATTCGAGTACGACAAGGGCGAGTTGTGGATGAGAAAACCGG
 CGAGATCAAGCCCAGAACAGTGGCGGCTGTACTCCGGCGAGAACATGGCAAGTCCCTGGACA
 GGTTCCCGGGAGAGAGGGAGAACAGATAAGTATGAGTGGAGAACATCGACAAGAACATCGATATC 30
 GTGAAGATCCTGGACGATCTGTTGTGAATTGGACAAGAACATCAGCCTGCTGAAGCA
 GCTGAAGGAGGGCGTGGAGCTGACCCGGAATAACGAGCACGGCACAGGCAGTCCCTGA
 GATTGCCATCACCTGATCCAGCAGATCGGAATACCGGAATAACGAGAACAGAAC
 GATTTCATCCTGTCCCCAGTGAGGGACGAGAACATGGCAAGCACTTGACTCTCGCGAGTA
 CTGGGATAAGGAGACAAAGGGCGAGAACAGATCAGCATGCCAGCTCCGGCGATGCCAATG
 GCGCCTAACATCGCCGGAAAGGGCATCATGAACGCCACATCCTGGCAATAGC
 GACTCCAAGGATCTGTCCTGTTGTGACGAGGAGTGGATCTGCACCTGAATAA
 CAAGACCGAGTGGAAAGAACAGCAGCTGAACATCTTCTAGCAGGAAGGCCATGGCAAGC
 GCAAGAACAAAGGCCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAACAGGG
 ATCCTACCCATACGATGTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCAT 40
 ACCCATATGATGTCGGACTATGCCTAA

(配列番号 20)

【 0 1 2 3 】

N L S 及び H A タグを含む P a r c u b a c t e r i a b a c t e r i u m G W C 2
 0 1 1 _ G W C 2 _ 4 4 _ 1 7 (P b C p f 1 ; p Y 0 0 8) :

【 0 1 2 4 】

MENIFDQFIGKYSLSKTLRFELKPVGKTEDFLKINKVFEKDQTIDDSYNQAKFYFDLHQK
 FIDAALASDKTSELSFQNFAVDVLEKQNKIILDKKREMGALRKRDKNAGVIDRLQKEINDAE
 DIIQKEKEKIYKDVRTLFDNEAESWKYYQEREVDGKKITFSKADLKQKGADFLTAAGILK
 VLKYEFPEEKEKEFQAKNQPSLFVEEKENPGQKRYIFDSFDKFAGYLTKFQQTKKNLYAA 50

DGTSTAVATRIADNFIIFHQNTKVFRDKYKNNHTLGFDEENIFEIERYKNCLLQREIEHIK
 NENSYNKIIGRINKKIKEYRDQKAKDTKLTKSDFPFFKNLDKQILGEVEKEKQLIEKTREKT
 EEDVPLIERFKEFIENNEERFTAACKLMNAFCNGEFSEYEYEGIYLKNKAINTISRRWFVSDR
 DFELKLPQQKSNNKSEKNEPKVKKFISIAEIKNAVEELDGDIFKAVFYDKIIAQGGSKLEQ
 FLVIWKYEFYELFRDIERENGEKLLGYDSCLKIAKQLGIFPQEKEAREKATAVIKNYADAG
 LGIFQMMKYFSLDDKDRKNTPGQLSTNFYAEYDGYYKDFEFIKYYNEFRNFITKKPFDED
 KIKLNFENGALLKGWDENKEYDFMGVLKKEGRLYLGIMHKHNRKLFQSMGNAGDNA
 NRYQKMIYKQIADASKDVPRLLTSSKAMEKFKPSQEILRIKKEKTFKRESKNFSLRDLHA
 LIEYYRNCPQYSNWSFYDFQFQDTGKYQNIKEFTDDVQKYGYKISFRDIDDEYINQALNE
 GKMYLFEVVNKDIYNTKNGSKNLHHTLYFEHILSAENLNDPVFKLSGMMAEIFQRQPSVNER 10
 EKITTKQNQCILDKGDRAYKYRRYTEKKIMFHMSLVNTGKGEIKQVQFNKIINQRISSSD
 NEMRVNVIGIDRGEKNLLYYSVVKQNGEIIEQASLNEINGVNVRDKLIEREKERLKNRQS
 WKPVVVKIKDLKKGYISHVIHKICQLIKYSIAVVLEDLNMRFKQIRGGIERSVYQQFEKALI
 DKLGYLVFKDNRDLRAPGVNLNGYQLSAPFVSFEKMRKQTGILFYTQAETYTSKTDPTGF
 RKNVYISNSASLDKIKEAVKKFDAIGWDGKEQSYFFKYNPYNLADEKYKNSTVSKWEAIF
 ASAPRIRRQKGEDGYWKYDRVKVNEEFKLLKVWNFVNPKATDIKQEIIKKEKAGDLQGE
 KELDGRLRNFWHSFIYLFNLVLELRNSFLQIKIKAGEVIAVDEGVDFIASPVKPFTTPNP
 YIPSNLCWLAVENADANGAYNIARKGVMILKKIREHAKKDPEFKKLPNLFISNAEWDEAA
 RDWGKYAGTTALNLDHKRPAATKKAGQAKKKGSYPYDVPDYAYPYDVPDYAYPYDVP
 DY A 20

(配列番号 21)

【 0 1 2 5 】

配列番号 21 は、核局在化シグナル (K R P A A T K K A G Q A K K K) (配列番号 12) 、その次にグリシン - セリンリンカー (G S) 、その次に H A タグ (Y P Y D V P D Y A Y P Y D V P D Y A) (配列番号 13) を含む。

【 0 1 2 6 】

配列番号 21 は以下のヌクレオチド配列によってコードされ得る。

【 0 1 2 7 】

ATGGAGAACATCTCGACCAGTTATCGGCAAGTACAGCCTGTCCAAGACCCCTGAGATT
 CGAGCTGAAGCCGTGGCAAGACAGAGGACTTCCTGAAGATCAACAAGGTGTTGAGA 30
 AGGATCAGACCATCGACGATAGCTACAATCAGGCCAAGTTCTATTTGATTCCTGCAC
 CAGAAGTTATCGACGCCCTGGCTCCGATAAGACATCCGAGCTGTCTTCCAGAA
 CTTGCCGACGTGCTGGAGAACGAGAATAAGATCATCCTGGATAAGAAGAGAGAGATGG
 GCGCCCTGAGGAAGCGCACAAGAACGCCGTGGCATCGATAGGCTGCAGAAGGAGATC
 AATGACGCCGAGGATATCATCCAGAAGGAGAACGAGATCTACAAGGACGTGCGCAC
 CCTGTCGATAACGAGGCCGAGTCTGGAAAACCTACTATCAGGAGCAGGAGGTGGACG
 GCAAGAAGATCACCTTCAGCAAGGCCGACCTGAAGCAGAACGGCGCCGATTTCTGACA
 GCCGCCGGCATCCTGAAGGTGCTGAAGTATGAGTTCCCGAGGAGAACGGAGTT
 TCAGGCCAAGAACCAAGGCCCTCCCTGTTGAGGAGAACGGAGAACGGAGAACGGAGTT
 GGTACATCTCGACTCTTGATAAGTCGCCGGCTATCTGACCAAGTTCAGCAGACA 40
 AAGAAGAATCTGTACGCAGCAGACGGCACCGACAGCAGCAGTGGCCACCCGCATGCCGA
 TAACTTATCATCTTCCACCAAGAACATACCAAGGTGTTCCGGGACAAGTACAAGAACAAATC
 ACACAGACCTGGCTTCGATGAGGAGAACATCTTGAGATCGAGAGGTATAAGAATTGC
 CTGCTGCAGCGCAGATCGAGCACATCAAGAACATGAGAACAGTACAACAAAGATCATCGG
 CCGGATCAATAAGAACATCAAGGAGTATCGGGACCAAGGCCAAGGATACCAAGCTGA
 CAAAGTCCGACTTCCCTTCTTAAGAACCTGGATAAGCAGATCCTGGCGAGGTGGAG
 AAGGAGAACGAGCTGATCGAGAAAACCCGGGAGAAAACCGAGGGACGTGCTGATCG
 AGCGGTTCAAGGAGTTATCGAGAACATGAGGAGAGGTTCACCGCCGCCAAGAACGCTG
 ATGAATGCCCTCTGTAACGGCGAGTTGAGTCCGAGTACGAGGGCATCTATCTGAAGAA
 TAAGGCCATCAACACAATCTCCGGAGATGGTCTGACAGAGATTTGAGCTGA 50

AGCTGCCTCAGCAGAAGTCCAAGAACAAAGTCTGAGAAGAATGAGCCAAAGGTGAAGAAC
 TTCATCTCCATGCCGAGATCAAGAACGCCGTGGAGGAGCTGGACGGCGATATCTTTAA
 GGCGTGTCTACGACAAGAAGATCATGCCAGGGCGCTTAAGCTGGAGCAGTCC
 TGGTCATCTGGAAGTACGAGTTGAGTATCTGTTCCGGACATCGAGAGAGAACGGC
 GAGAAGCTGCTGGCTATGATAGCTGCCAAGCAGCTGGCATCTTCCC
 ACAGGAGAAGGAGGCCCGCGAGAAGGCAACC GCCGTGATCAAGAATTACGCCGACGCCG
 GCCTGGGCATCTCCAGATGATGAAGTATTTCTGGACGATAAGGATCGGAAGAAC
 ACCCCC GGCCAGCTGAGCACAAATTCTACGCCAGTATGACGGCTACTACAAGGATT
 CGAGTTATCAAGTACTACAACGAGTTAGGAACCTCATACCAAGAACGCCCTCGACG
 AGGATAAGATCAAGCTGAACCTTGAGAATGGGCCCTGCTGAAGGGCTGGACGAGAAC 10
 AAGGAGTACGATT CATGGCGTATCCTGAAGAAGGAGGCCCTGTATCTGGC
 CATGCACAAGAACCAACCGGAAGCTGTTCAAGTCCATGGCAATGCCAAGGGCACAACG
 CCAATAGATACCAGAAGATGATCTATAAGCAGATGCCGACGCCCTTAAGGATGTGCC
 AGGCTGCTGCTGACCAGCAAGAAGGCCATGGAGAAGTTCAAGCCTCCCAGGAGATCCT
 GAGAATCAAGAAGGAGAAAACCTCAAGCAGGAGAGCAAGAACCTTCCCTGAGAGATC
 TGCACGCCCTGATCGAGTACTATAGGAAC TGCA CCCTCAGTACAGCAATTGGCCTT
 TATGACTCCAGTTCAAGGATACCGGCAAGTACCGAGAATATCAAGGAGTTACAGACGA
 TGTGCAGAAGTACGGCTATAAGATCTCCTTCGCGACATCGACGATGAGTATATCAATC
 AGGCCCTGAACGAGGGCAAGATGTACCTGTTGAGGTGGTAACAAGGATATCTATAAC
 ACCAAGAATGGCTCCAAGAATTCGACACACTGTACTTGAGCACATCCTGCTGCCGA 20
 GAACCTGAATGACCCAGTGTCAAGCTGCTGGCATGGCGAGATCTTCAGCGCAGC
 CCAGCGTGAACGAAAGAGAGAACGATCACCAACACAGAAGAACATCAGTGTATCCTGGACAAG
 GGCAGATAGAGCCTACAAGTATAGCGCTACACCGAGAACAGATCATGTTCCACATGAG
 CCTGGTGCTGAACACAGGAAGGGCGAGATCAAGCAGGTGCAGTTAATAAGATCATCA
 ACCAGAGGATCAGCTCTGACAACAGAGATGAGGGTGAATGTGATCGGCATCGATCGC
 GGCAGAACCTGCTGACTATAGCGTGGTAAGCAGAACATGGCGAGATCATCGAGCA
 GGCCTCCCTGAACGAGATCAATGGCGTAAC TACCGGACAAGCTGATCGAGAGGGAGA
 AGGAGCGCCTGAAGAACCGGCAGAGCTGGAGCCTGTGGTAAGATCAAGGATCTGAAG
 AAGGGCTACATCTCCACGTGATCCACAAGATCTGCCAGTGTGAGAACATCTGC
 CATCGTGGTGCTGGAGGACCTGAATATGAGATTCAAGCAGATCAGGGAGGAATCGAGC 30
 GGAGCGTGTACCGAGCTTCAGTGTGAGAACAGGCCCTGATCGATAAGCTGGCTATCTGGT
 AAGGACAACAGGGATCTGAGGGCACCAAGGAGGGCGTGTGAATGGCTACCAAGCTGTCTG
 CCCCTTGAGCTCGAGAACAGATGCGCAAGCAGACCGGCATCCTGTTACACACAGG
 CCGAGTATACCAGCAAGACAGACCCAATACCGGCTT CGGAAGAACGTGTATATCTCT
 AATAGCGCCTCCCTGGATAAGATCAAGGAGGCCGTGAAGAACGTTCGACGCCATCGGCTG
 GGATGGCAAGGAGCAGTCTTACTTCAAGTACAACCCCTACAAACCTGGCCACGAGA
 AGTATAAGAACTTACCGTGAGCAAGGAGTGGCCATCTTGCCAGCGCCCCAAGAAC 40
 CGGAGACAGAACGGCGAGGACGGCTACTGGAGTATGATAGGGTGAAGAACGAGGAG
 AGTTCGAGAACGCTGCTGAAGGTCTGGAGTTTGTAACCCAAAGGCCACAGATATCAAG
 CAGGAGATCATCAAGAACGGAGAACGGCAGGCACCTGCAGGGAGAGAACGGAGCTGGATG
 GCCGGCTGAGAACCTTGGCACTCTTCACTACCTGTTAACCTGGTGCTGGAGCTG
 CGCAATTCTTCAGCCTGCAGATCAAGATCAAGGAGCAGGAGAACGTGATCGCAGTGGACGA
 GGGCGTGGACTTCATGCCAGCCAGTGAAGCCCTCTTACCAACCCAAACCCCTACA
 TCCCCCTCCAACCTGTGCTGGCTGGCGTGGAGAACATGCAGACGCAAACGGAGCCTATAAT
 ATCGCCAGGAAGGGCGTGTGATGATCCTGAAGAACGATCCGCAGCACGCCAAGAACGGAGCC
 CGAGTTCAAGAACGCTGCCAAACCTGTTATCAGCAATGCAGAGTGGGACGAGGCAGGCC
 GGGATTGGGGCAAGTACGCAGGCACCACAGCCCTGAACCTGGACCACAAAAGGCCGGCG
 GCCACGAAAAAGGCCGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGATGTTCC
 AGATTACGCTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGACT
 ATGCCTAA 50

(配列番号 22)

【0128】

N L S 及び H A タグを含む Smithella spp. SC_K08D17 (Ssc p f1; pY009) :

【0129】

MQTLFENFTNQYPVSKTLRFELIPQGKTKDFIEQKGLLKKDEDRAEKYKKVKNIIDEYHK
 DFIEKSLNGLKLDGLEKYKTLYLKQEKDDDKAFAKEKENLRKQIANAFRNNEKFKTLF
 AKELIKNDLMSFACEEDKKNVKEFEAFTTYFTGFHQQRANMYVADEKRTAIASRLIHENL
 PKFIDNIKIFEKMKEAPELLSPFNQTLKDMKDVIKGTTLEEFSLDYFNKTLTQSGIDIYN
 SVIGGRTPPEEGKTKIKGLNEYINTDFNQKQTDKKRQPKFKQLYKQILSDRQSLSFIAEAF
 KNDTEILEAIEKFYVNELLHFSNEGKSTNVLDIAKNAVSNLESFNLTKMYFRSGASLTDVS
 RKVGEWSIINRALDNYYATTYPIKPREKSEKYEERKEKWLKQDFNVSLIQTAEYDNET
 VKGKNSGKVIADYFAKFCDDKETDLIQKVNEGYIAVKDLLNTPCPENEKLGSNKDQVKQI
 KAFMDSIMDIMHFVRPLSLKDTDKEDETFSLFTPLYDHLTQTIALYNKVRNYLTQKPY
 STEKIKLNFENSTLLGGWDLNKETDNTAIILRKDNLYLGIMDKRHNRIFRNPVKADKKD
 FCYEKMVYKLLPGANKMLPKVFFSQSRIQEFTPSAKLENYANETHKKGDNFNLNHCHK
 LIDFFKDSINKHEDWKNFDRFSATSTYADLSGFYHEVEHQGYKISFQSVADSFIDLVNE
 GKLYLFQIYNKDFSPFSKGKPNLHTLYWKMLFDENNLDVYKLNGEAEVFYRKKSSIAEK
 NTTIHKANESIINKNPDPNPKATSTFNYDIVKDKRYTIDKFQFHIPITMFKAEGIFNMNQR
 VNQFLKANPDINIIGIDRGERHLLYYALINQKGKILKQDTLNVIANEKQKVDYHNLLDKKE
 GDRATARQEWGVETIKELKEGYLSQVIHKLTDLMIENNAAIVMEDLNFGFKRGRQKVEK
 QVYQKFEKMLIDKLNLYLVDKNKKANELGGLLNAFQLANKFESFQKMGKQNGFIFYVPAW
 NTSKTDPATGFIDFLKPRYENLNQAKDFFEKFDSIRLNKADYFEFAFDKKNFTEKADGGR
 TKWTVCTTNEDRYAWNRALNNNRGSQEKYDITAELKSLFDGKVDYKSGKDLQQIASQE
 SADFFKALMKNLISITLSLRHNNGEKGDN EQDYILSPVADSKGRFFDSRKADDDMPKNAD
 ANGAYHIALKGLWCLEQISKTDLKKVKLAISNEKWLEFVQTLKGKRPAATKKAGQAKK
 KKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA

10

20

30

(配列番号 23)

【0130】

配列番号 23 は、核局在化シグナル (K R P A A T K K A G Q A K K K K) (配列番号 12)、その次にグリシン - セリンリンカー (G S)、その次に H A タグ (Y P Y D V P D Y A Y P Y D V P D Y A) (配列番号 13) を含む。

【0131】

配列番号 23 は以下のヌクレオチド配列によってコードされ得る。

【0132】

ATGCAGACCCCTGTTGAGAACTTCACAAATCAGTACCCAGTGTCCAAGACCCCTGCGCTT
 TGAGCTGATCCCCAGGGCAAGACAAAGGACTTCATCGAGCAGAAGGGCCTGCTGAAGA
 AGGATGAGGACCGGGCCGAGAAGTATAAGAAGGTGAAGAACATCATCGATGAGTACAC
 AAGGACTTCATCGAGAAGTCTCTGAATGGCCTGAAGCTGGACGGCCTGGAGAAGTACAA
 GACCCCTGTATCTGAAGCAGGAGAAGGACATAAGGATAAGAAGGCCCTTGACAAGGAGA
 AGGAGAACCTGCGCAAGCAGATGCCAATGCCCTCCGGAACAATGAGAAGTTAACAGA
 CTGTTGCCAAGGAGCTGATCAAGAACGATCTGATGTCTTCGCCCTGCGAGGGAGGACAA
 GAAGAACATGTGAAGGAGTTGAGGCCCTCACACATACTTCACCGGCTCCACCAAGAAC
 GCGCCAATATGTACGTGGCCGATGAGAAGAGAACAGCCATGCCAGCAGGCTGATCCAC
 GAGAACCTGCCAAGTTATCGACAATATCAAGATCTCGAGAAGATGAAGAACAGGAGG
 CCCCAGAGCTGCTCTCCTTCAACCAGACCCCTGAAGGATATGAAGGACGTGATCAAGG
 GCACCAACTGGAGGAGATCTTAGCCTGGATTATTCACACAAGACCCCTGACACAGAGC
 GGCATCGACATCTACAATTCCGTGATCGGCGCAGAACCCCTGAGGAGGGCAAGACAAA
 GATCAAGGGCCTGAACGAGTACATCAATACGACTTCAACCAGAAGCAGACAGACAGAAGA
 AGAACGGCAGCCAAGTTCAAGCAGCTGTATAAGCAGATCCTGAGCGATAGGCAGAGC

40

50

CTGTCCTTATGCCGAGGCCTTCAAGAACGACACCGAGATCCTGGAGGCCATCGAGAA
 GTTTTACGTGAATGAGCTGCTGCACTTCAGCAATGAGGGCAAGTCCACAAACGTGCTGG
 ACGCCATCAAGAACGCCGTCTGACAGACGTGAGCCGGAAGGTGTTGGCGAGTGAGCATCAT
 CAATAGAGCCCTGGACAACACTACTATGCCACCACATATCCAATCAAGCCCAGAGAGAAGT
 CTGAGAAGTACGAGGAGAGGAAGGAGAAGTGGCTGAAGCAGGACTTCAACGTGAGCCTG
 ATCCAGACCGCCATCGATGAGTACGACAACGAGACAGTGAAGGGCAAGAACAGCGGCAA
 AGTGATGCCGATTATTTGCCAAGTTCTGCAGATAAGGAGACAGACCTGATCCAGA
 AGGTGAACGAGGGCTACATGCCGTGAAGGATCTGCTGAATAACACCCCTGTCCTGAGAAC
 GAGAAGCTGGCAGCAATAAGGACCAGGTGAAGCAGATCAAGGCCTTATGGATTCTAT 10
 CATGGACATCATGCACTCGTGCCCCCCTGAGCCTGAAGGATAACCGACAAGGAGAAGG
 ATGAGACATTCTACTCCCTGTTCACACCTCTGTACGACCACCTGACCCAGACAATGCC
 CTGTATAACAAGGTGCGGAACATCTGACCCAGAACGCCTACAGCACAGAGAAGATCAA
 GCTGAACCTCGAGAACAGCACCCCTGCTGGCGGCTGGGATCTGAATAAGGAGACAGACA
 ACACAGCCATCATCCTGAGGAAGGATAACCTGTACTATCTGGGATCATGGACAAGAGG
 CACAATCGCATTTCGAACGTGCCAAGGCCATAAGAAGGACTTCTGCTACGAGAA
 GATGGTGTATAAGCTGCTGCCGGCCAACAAGATGCTGCCAAGGTGTTCTTCTC
 AGAGCAGAATCCAGGAGTTACCCCTCCGCCAGCTGCTGGAGAACATACGCCAATGAG
 ACACACAAGAAGGGCGATAATTCAACCTGAATCACTGTACAAGCTGATCGATTCTT
 TAAGGACTCTATCAACAAGCACGAGGATTGGAAGAATTGACTTTAGGTTAGCGCCA 20
 CCTCCACCTACGCCGACCTGAGCGGCTTTTACGAGGTGGAGCACCCAGGGCTACAAG
 ATCTCTTTCAGAGCGTGGCGATTCTCATCGACGATCTGGTAACGAGGGCAAGCT
 GTACCTGTTCCAGATCTATAATAAGGACTTTCCCCATTCTCTAAGGGCAAGGCCAAC
 TGCACACCCGTACTGGAAGATGCTGTTGATGAGAACAACTGAAGGACGTGGTGTAT
 AAGCTGAATGGCGAGGCCGAGGTGTTCTACCGCAAGAACAGCATTGCCGAGAACAC
 CACAATCCACAAGGCCATGAGTCATCATCAACAAGAACCTGATAACCCAAAGGCCA
 CCAGCACCTCAACTATGATATCGTAAGGACAAGAGATAACCATCGACAAGTTTAG 30
 TTCCACATCCAATCACAATGAACCTTAAGGCCGAGGGCATCTCAACATGAATCAGAG
 GGTGAATCAGTCCCTGAAGGCCAATCCGATATCAACATCATCGGCATCGACAGAGGCG
 AGAGGCACCTGCTGACTATGCCGATCAACCAGAACGGCAAGATCCTGAAGCAGGAT
 ACCCTGAATGTGATGCCAACGAGAACGAGAACGGTGGACTACCACAATCTGCTGGATAA
 GAAGGAGGGCGACCGCGAACCGCAAGGCAGGAGTGGGGCGTATCGAGAACATCAAG
 GAGCTGAAGGAGGGCTATCTGCCCAGGTATCCACAAGCTGACCGATCTGATGATCGA
 GAACAATGCCATCATCGTATGGAGGACCTGAACTTGGCTTAAGGCCAGAGAACAGA
 AGGTGGAGAACGAGGTGTATCAGAACGAGTGGACTATGCTGATCGATAAGCTGAATTAC
 CTGGTGGACAAGAATAAGAAGGAAACGAGCTGGGAGGCCGCTGCTGAACGCATTCCAGCT
 GGCAATAAGTTGAGTCCTCCAGAACGAGTGGCAAGCAGAACGGCTTATCTTCTACG
 TGCCCGCCTGGAATACCTCTAACGACAGATCCTGCCACCGGCTTATCGACTTCCTGAAG
 CCCGCTATGAGAACCTGAATCAGGCCAAGGATTCTTGAGAACGAGTGGACTCTATCCG
 GCTGAACAGCAAGGCCGATTACTTGAGTTCGCCCTTGACTTCAAGAACATTCAACCGAGA 40
 AGGCCGATGGCGCAGAACCAAGTGGACAGTGTGCACCAACAGAGGACAGATATGCC
 TGGAAATAGGCCCTGAACAATAACAGGGCAGCCAGGAGAACGAGTACGACATCACAGCCGA
 GCTGAAGTCCCTGTTGATGGCAAGGTGGACTATAAGTCTGGCAAGGATCTGAAGCAGC
 AGATGCCAGGCCAGGAGTCCGCCACTTCTTAAGGCCCTGATGAAGAACCTGTCCATC
 ACCCTGTCCTGAGACACAAATAACGGCGAGAACGGCGATAATGAGCAGGACTACATCCT
 GTCCCCGTGGCCGATTCTAACGGCCGCTTCTTGACTCCGGAAAGGCCGACGATGACA
 TGCCAAAGAACGAGTGGCGACGCCAACGGCGCCTATCACATGCCCTGAAGGGCTGTGGTGT
 CTGGAGCAGATCAGCAAGACCGATGACCTGAAGAACGGTGAAGCTGCCATCTCAACAA
 GGAGTGGCTGGAGTTGTCAGACACTGAAGGGCAAAAGGCCGGGCCACGAAAAGG 50
 CGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGATGTTCCAGATTACGCTTAT

CCCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTATGCCTAA

(配列番号 24)

【0133】

N L S 及び H A タグを含む *Acidaminococcus* sp. B V 3 L 6 (As C p f 1 ; p Y 0 1 0) :

【0134】

MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQQFIEEDKARNDHYKELKPIIDRIYKTY
ADQCLQLVQLDWENLSAIDSYRKEKTEETRNALIEEQATYRNAIHDYFIGRTDNLTD
NKRHAEIYKGLFKAELFNGKVLKQLGTVTTTEHENALLRSFDKFETYFSGFYENRKNVFS
AEDISTAIPH RIV QDNFPKF KENCHIFTRLITAVPSLREHFENVKKAI GIVF VST SIEEVFSFP
FYNQLLTQTQIDLYNQLLGGISREAGTEKIKGLNEVLNLAIQKNDTAHIIASLPHRFIPLF
KQILSDRNTLSFILEEKSDEEVIQSFCKYKTLRNENVLETAEALFNELNSIDLTHIFISHK
KLETISSALCDHWDTLRNALYERRISELTGKITKSAKEKVQRSLKHEDINLQEIIASA
GKEL SEAFKQKTSEILSHAHAALDQPLPTTLKKQEEKEILKSQLDSLLGLYHLLDWFAVDESNE
VDPEFSARLTGKLEMEPSLSFYNKARNYATKKPYSVEKFKNFQMPTLASGWVDVNKEK
NNGAILFVKNGLYLGIMPQKQGRYKALSFEPEKTSEGFDKMYDYPDAAKMIPKCST
QLKAVTAHFQTHTPILLSNNFIEPLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKG
YRE ALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYHISFQR
IAEKEIMD AVETGKLYLFQIYNKDFAKGHGKPNLHTLYWTGLFSPENLAKTSIKLNGQA
ELFYRPKS RMKRMAHRLGEKMLNKKLDQKTPIDPTLYQELYDYVNHRSLHDL
SDEARALLPVITK EVSHEIIKDRRFTSDKFFFHVPI TLNYQAANSPSKFNQR
VNAYLKEHPETPIIGIDRGERNL IYITVIDSTGKILEQRSLNTIQQFDYQKKLD
NREKERVAARQAWSVVGTIKDLKQGYLSQV IHEIVDLMIH
YQAVVVLENLNFGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKD
YPAEK VGGVLNPYQLTDQFTSFAKMG
TQSGFLFYVPAPYTSKIDPLTG
FVDPFWKTIKNHESRK HFLEGFD
FLHYDVKTGDFILHF
KMNRNLSFQRGLPG
FMPAWDIVFEKNETQFD
AKGTPFI AGKRI
VPVIENHRFTG
RYRDLYP
ANELIAL
LEEKGIVFR
DGSNILPK
LEND
SHAIDTMV
ALIRSVLQ
MRNSNA
ATGEDY
INSP
VRDL
NGVC
FDSRF
QNPE
WPM
DADANG
AYHIALKG
LLL
NH
LKE
SKDL
KLQ
NGIS
NQD
WL
AYIQ
ELRN
KRPA
ATKK
AGQ
AKKK
GS
PYD
VPD
YA
PYD
VPD
YA

(配列番号 25)

30

【0135】

配列番号 25 は、核局在化シグナル (K R P A A T K K A G Q A K K K K) (配列番号 12) 、その次にグリシン - セリンリンカー (G S) 、その次に H A タグ (Y P Y D V P D Y A Y P Y D V P D Y A) (配列番号 13) を含む。

【0136】

配列番号 25 は以下のヌクレオチド配列によってコードされ得る。

【0137】

ATGACACAGTTCGAGGGCTTACCAACCTGTATCAGGTGAGCAAGACACTGCGGTTGA
GCTGATCCCACAGGGCAAGACCCCTGAAGCACATCCAGGAGCAGGGCTTCATCGAGGAGG
ACAAGGCCCGCAATGATCACTACAAGGAGCTGAAGGCCATCATCGATCGGATCTACAAG
ACCTATGCCGACCA GTGCCTGCAGCTGGCAGCTGGATTGGAGAACCTGAGCGCCGC
CATCGACTCCTATAGAAAGGAGAAAACCGAGGGAGACAAGGAACGCCCTGATCGAGGAGC
AGGCCACATATCGCAATGCCATCCACGACTACTTCATCGGCCGGACAGACAACCTGACC
GATGCCATCAATAAGAGACACGCCGAGATCTACAAGGGCTGTTCAAGGCCGAGCTGTT
TAATGGCAAGGTGCTGAAGCAGCTGGCACC GTGACCACAACCGAGCACGAGAACGCC
TGCTCGGGAGCTTCGACAAGTTACAACCTACTTCTCCGGCTTTATGAGAACAGGAAG
AACGTGTTCAAGGAGAATTGTCACATCTCACACGCCGATCGTGCAGGACAA
CTTCCCCAAGTTAAGGAGAATTGTCACATCTCACACGCCGATCGTGCAGGAC
GCCTCGGGAGCACTTGAGAACGTGAAGAAGGCCATCGGCATCTCGTGAGCAC
ATCGAGGAGGTGTTCCCTTATAACCAGCTGCTGACACAGACCCAGATCGA

40

50

CCTGTATAACCAGCTGCTGGAGGAATCTCTGGGAGGCAGGCACCGAGAAGATCAAGG
 GCCTGAACGAGGTGCTGAATCTGCCATCCAGAAGAACATGATGAGACAGCCCACATCATC
 GCCTCCCTGCCACACAGATTCATCCCCCTGTTAACGAGATCCTGTCCGATAGGAACAC
 CCTGTCTTCATCCTGGAGGAGTTAACGAGCGACGAGGAAGTGATCCAGTCCTCTGCA
 AGTACAAGACACTGCTGAGAAACGAGAACGTGCTGGAGACAGCCGAGGCCCTGTTAAC
 GAGCTGAACAGCATCGACCTGACACACATCTCATCAGCCACAAGAACGCTGGAGACAAT
 CAGCAGCGCCCTGTGCACCACTGGATACACTGAGGAATGCCCTGTATGAGCGGAGAA
 TCTCCGAGCTGACAGGAAGATCACCAAGTCTGCCAAGGAGAACGGTGAGCGCAGCCTG
 AAGCACGAGGATATCAACCTGCAGGAGATCATCTGCCGCAGGCAAGGAGCTGAGCGA
 GCCCTTCAAGCAGAAAACCAGCGAGATCCTGTCCCACGCACACGCCGCCCTGGATCAGC
 10 CACTGCCCTACAACCCCTGAAGAACGAGGAGAACGGAGATCCTGAAGTCTCAGCTGGAC
 AGCCTGCTGGGCCGTACACCTGCTGGACTGGTTGCCGTGGATGAGTCCAACGAGGT
 GGACCCCGAGTTCTCTGCCCGCTGACCGGCATCAAGCTGGAGATGGAGCCTTCTCTGA
 GCTTCTACAACAAGGCCAGAAATTATGCCACCAAGAACGCCCTACTCCGTGGAGAAGTTC
 AAGCTGAACCTTCAGATGCCCTACACTGCCCTCTGGCTGGACGTGAATAAGGAGAAC
 CAATGGCGCCATCCTGTTGTGAAGAACGCCCTGTACTATCTGGCATCATGCCAAAGC
 AGAAGGGCAGGTATAAGGCCCTGAGCTTCAGGCCACAGAGAACCCAGCGAGGGCTT
 GATAAGATGTACTATGACTACTCCCTGATGCCGCAAGATGATGCCAAAGTGCAGCAC
 CCAGCTGAAGGCCGTGACAGCCCACCTTCAGACCCACACAACCCCCATCCTGCTGTCCA
 ACAATTTCATCGAGCCTCTGGAGATCACAAAGGAGATCTACGACCTGAACAATCCTGAG
 20 AAGGAGCCAAGAACGTTCAGACAGCCTACGCCAAGAACCCGGCGACCAGAACGGCTA
 CAGAGAGGCCCTGTGCAAGTGGATCGACTTCACAAGGGATTTCTGTCCAAGTATACCA
 AGACAAACCTCTATCGATCTGTCTAGCCTGCCCATCCTCTCAGTATAAGGACCTGGC
 GAGTACTATGCCGAGCTGAATCCCTGCTGTACCATCAGCTCCAGAGAACGGC
 GAAGGAGATCATGGATGCCGTGGAGACAGGCAAGCTGTACCTGTTCCAGATCTATAACA
 AGGACTTTGCCAAGGCCACCACGGCAAGCCTAATCTGCACACACTGTATTGGACCGGC
 CTGTTTCTCCAGAGAACCTGCCAAGACAAGCATCAAGCTGAATGCCAGGCCGAGCT
 GTTCTACGCCCTAACGTCCAGGATGAAGAGGATGGCACACCCGGCTGGAGAGAAC
 TGAACAAGAACGCTGAAGGATCAGAAAACCCAATCCCCGACACCCCTGTACCAAGGAGCTG
 TACGACTATGTGAATCACAGACTGCTCCACGACCTGTCTGATGAGGCCAGGCCCTGCT
 30 GCCCAACGTGATCACCAAGGAGGTGTCACGAGATCATCAAGGATAGGCCTTACCA
 GCGACAAGTTCTTCCACGTGCCATCACACTGAACATCACGCCGCCATTCCCCA
 TCTAAGTTCAACCAGAGGGTAATGCCCTACCTGAAGGAGCACCCGAGACACCTATCAT
 CGGCATCGATGGGGCGAGAGAACCTGATCTATATCACAGTGTACGACTCCACCGCA
 AGATCCTGGAGCAGCGGAGCCTGAACACCATCCAGCAGTTGATTACCAGAACAGCTG
 GACAACAGGGAGAAGGAGAGGGTGGCAGCAAGGCAGGCCCTGGTCTGTGGAGAAC
 TCAAGGATCTGAAGCAGGGCTATCTGAGCCAGGTATCCACGAGATCGTGGACCTGATG
 ATCCACTACCAGGCCGTGGTGGCTGGAGAACCTGAATTTCGGCTTAAAGAGCAAGAG
 GACCGGCATGCCGAGAACGCCGTGTACCAAGCAGTTGAGAACATGCTGATCGATAAGC
 TGAATTGCCCTGGTCTGAAGGACTATCCAGCAGAGAACGGTGGAGGCCGTGCTGAACCC
 TACCAAGCTGACAGACCAAGTTCACCTCCTTGCCAAGATGGGCACCCAGTCTGGCTT
 40 CCTTACGTGCCCTGCCCATATACATCTAACGATCGATCCCTGACCCGGCTCGTGGACC
 CCTCGTGTGGAAAACCATCAAGAACATCACGAGAGGCCAGCAAGCAACTCCCTGGAGGG
 GACTTTCTGCACTACGACGTGAAAACCGGGCAGCTCATCCTGCACTTAAAGATGAACAG
 AAATCTGTCCTCCAGAGGGCCTGCCGGCTTATGCCTGCATGGATATCGTGGT
 AGAAGAACGAGACACAGTTGACGCCAAGGGCACCCCTTCATGCCGGCAAGAGAAC
 GTGCCAGTGTACGAGAACATCACAGATTCAACGGCAGATAACGGGACCTGTATCCTGCC
 CGAGCTGATGCCCTGCTGGAGGAGAACGGCATCGTGTTCAGGGATGGCTCCAACATCC
 TGCCAAAGCTGCTGGAGAACGACGATTCTCACGCCATCGACACCACATGGTGGCCCTGATC
 CGCAGCGTGCTGCAGATGCCAACTCAATGCCGCCACAGGCAGGGACTATATCAACAG
 50

CCCCGTGCGCGATCTGAATGGCGTGTGCTTCGACTCCCGGTTCAGAACCCAGAGTGGC
 CCATGGACGCCGATGCCAATGGCGCCTACCACATCGCCCTGAAGGCCAGCTGCTGCTG
 AATCACCTGAAGGAGAGCAAGGATCTGAAGCTGCAGAACGGCATCTCAATCAGGACTG
 GCTGGCCTACATCCAGGAGCTGCACAAAAAGGCCGGCCACGAAAAAGGCCGGCC
 AGGCAAAAAGAAAAAGGGATCCTACCCATACGATGTTCCAGATTACGCTTATCCCTAC
 GACGTGCCTGATTATGCATACCCATATGATGTCCCCGACTATGCCTAA

(配列番号 26)

【0138】

N L S 及び H A タグを含む *Lachnospiraceae bacterium* MA
 2020 (Lb2Cpf1; pY011) :

10

【0139】

MYYESLTQYPVSKTIRNELIPIGKTLNDNIRQNNILESVDVRKRQNYEHVKGILDEYHKQLI
 NEALDNCLPSLKIAAEIYLKNQKEVSDREDFNKTQDLLRKEVVEKLKAHENFTKIGKKDI
 LDLLEKLPSEDDYNALESFRNFYTYFTSYNKVRENLYSDKEKSSTVAYRLINENFPKFL
 DNVKSYRFVKTAGILADGLGEEEQDSLIVETFNKTLTQDGIDTYNSQVGKINSSINLYNQ
 KNQKANGFRKIPKMMLYKQILSDREESFIDEQSDEVLDNVESYGSVLIESLKSSKVSAF
 FDALRESKGKNVVVKNDLAKTAMSNIVFENWRTFDDLLNQEYDLANENKKDDKYFEKR
 QKELKKNKSYSLEHLCNLSEDSCNLIENYIHQISDDIENIIINNETFLRIVINEHDRSRKLAK
 NRKAVKAIKDFLDSIKVLERELKLINSSGQELEKDLIVYSAHEELLVELKQVDSLNMTRN
 YLTKKPFSTEKVKLNFRSTLLNGWDRNKTENDLGVLKDGKYYLGIMNTSANKAFVN
 PPVAKTEKVFKVVDYKLLPVPNQMLPKVFFAKSNIDFYNPSEIYSNYKKGTHKKGNMFS
 LEDCHNLIDFFKESISKHEDWSKFGFKFSDTASYNDISEFYREVEKQGYKLTYTDIETYIN
 DLIERNELYLFQIYNKDFSMYSKGKLNHLHTLYFMMLFDQRNIDDVVYKLNGEAEVFYRPA
 SISEDELIIHKAGEEIKKNKNPNRARTKETSTFSYDIVKDKRYSKDKFTLHIPITMNFGVDEV
 KRFNDAVNSAIRIDENVNVIGIDRGERNLLYVVVIDSKGNILEQISLNSIINKEYDIETDYH
 ALLDEREGGRDKARKDWNTVENIRDLKAGYLSQVNVVAKLVLKYNAAICLEDLNFGFKR
 GRQKVEKQVYQKFEKMLIDKLNLYLVIDKSREQTSPKELGGALNALQLTSKFKSFKELGKQ
 SGVIYYVPAYLTSKIDPTTGFMKCENVEKSKRFFDGFIRFNALENVFEFGFDYR
 SFTQRACGINSKWTVCTNGERIICKYRNPDKNMFDEKVVVTDEMKNLFEQYKIPYEDG
 RNVKDMIISNEEAIFYRRLYRLLQQTLQMRNSTSDGTRDYIISPVKNKREAYFNSELSDGS
 VPKDADANGAYNIARKGLWVLEQIRQKSEGEKINLAMTNAEWLEYAQTHLLKRPAATKK
 AGQAKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA

20

(配列番号 27)

【0140】

配列番号 27 は、核局在化シグナル (K R P A A T K K A G Q A K K K K) (配列番号 12)、その次にグリシン - セリンリンカー (G S)、その次に H A タグ (Y P Y D V P D Y A Y P Y D V P D Y A) (配列番号 13) を含む。

30

【0141】

配列番号 27 は以下のヌクレオチド配列によってコードされ得る。

【0142】

ATGTACTATGAGTCCCTGACCAAGCAGTACCCGTGTCTAAGACAATCCGGAATGAGCT
 GATCCCTATCGGCAAGACACTGGATAACATCCGCCAGAACAAATATCCTGGAGAGCGACG
 TGAAGCGGAAGCAGAACTACGAGCACGTGAAGGGCATCCTGGATGAGTATCACAAGCAG
 CTGATCAACGAGGCCCTGGACAATTGCACCCCTGCCATCCCTGAAGATCGCCGCCGAGAT
 CTACCTGAAGAACATCAGAAGGAGGTGTCTGACAGAGAGGATTCAACAAGACACAGGACC
 TGCTGAGGAAGGAGGTGGAGAAGAGCTGAAGGCCACGAGAACATTACCAAGATCGGC
 AAGAAGGACATCCTGGATCTGCTGGAGAACAGCTGCCCTCCATCTGAGGACGATTACAA
 TGCCCTGGAGAGCTTCCGCAACTTTACACCTATTACATCCTACAACAAGGTGCGGG
 AGAATCTGTATTCTGATAAGGAGAACAGAGCTCCACAGTGGCCTACAGACTGATCAACGAG
 AATTTCCTAAAGTTCTGGACAATGTGAAGAGCTAGGTTGTGAAAACCGCAGGCAT

40

50

CCTGGCAGATGGCCTGGGAGAGGAGGAGCAGGACTCCCTGTTCATCGTGGAGACATTCA
 ACAAGACCCCTGACACAGGACGGCATCGATACTACAATTCTCAAGTGGGCAAGATCAAC
 TCTAGCATCAATCTGTATAACCAGAAGAATCAGAAGGCCAATGGCTTCAGAAAGATCCC
 CAAGATGAAGATGCTGTATAAGCAGATCCTGTCCGATAGGGAGGAGTCTTCATCGACG
 AGTTTCAGAGCGATGAGGTGCTGATCGACAACGTGGAGTCTTATGCAGCGTGCTGATC
 GAGTCTCTGAAGTCCTCTAAGGTGAGCGCCTTCTTGATGCCCTGAGAGAGTCTAAGGG
 CAAGAACGTGTACGTGAAGAATGACCTGGCCAAGACAGCCATGAGCAACATCGTGTTCG
 AGAATTGGAGGACCTTGACGATCTGCTGAACCAGGAGTACGACCTGGCCAACGAGAAC
 AAGAAGAAGGACGATAAGTATTCGAGAAGCGCCAGAAGGAGCTGAAGAAGAATAAGA
 GCTACTCCCTGGAGCACCTGTGCAACCTGTCCGAGGATTCTGTAAACCTGATCGAGAAT 10
 TATATCCACCAGATCTCGACGATATCGAGAATATCATCATCAACAAATGAGACATTCT
 GCGCATCGTATCAATGAGCACGACAGGTCCCGCAAGCTGGCCAAGAACCGGAAGGCCG
 TGAAGGCCATCAAGGACTTCTGGATTCTATCAAGGTGCTGGAGCGGGAGCTGAAGCTG
 ATCAACAGCTCCGCCAGGAGCTGGAGAAGGATCTGATCGTGTACTCTGCCACGAGGA
 GCTGCTGGTGGAGCTGAAGCAGGTGGACAGCCTGTATAACATGACCAAGAAATTATCTGA
 CAAAGAACCTTCTACCGAGAAGGTGAAGCTGAACCTTAATCGCAGCACACTGCTG
 AACGGCTGGATCGGAATAAGGAGACAGACAACTGGCGTGTGCTGAAGGACGG
 CAAGTACTATCTGGCATCATGAACACAAGCGCCAATAAGGCCCTCGTAATCCCCCTG
 TGGCCAAGACCGAGAAGGTGTTAAGAAGGTGGATTACAAGCTGCTGCCAGTGCCCAAC
 CAGATGCTGCCAAAGGTGTTCTTGCCAAGAGCAATATCGACTTCTATAACCCCTCTAG
 CGAGATCTACTCCAATTATAAGAAGGGCACCCACAAGAAGGGCAATATGTTTCCCTGG 20
 AGGATTGTCACAACCTGATCGACTTCTTAAGGAGTCTATCAGCAAGCACGAGGACTGG
 AGCAAGTTGGCTTAAAGTTAGCGATACAGCCTCCTACAACGACATCTCGAGTTCTA
 TCGCGAGGTGGAGAAGCAGGGCTACAAGCTGACCTATAAGACATCGATGAGACATA
 TCAATGATCTGATCGAGCGGAACGAGCTGTACCTGTTCCAGATCTATAAGGACTTT
 AGCATGTAACCTCAAGGGCAAGCTGAACCTGCACACACTGTATTCATGATGCTGTTGA
 TCAGCGCAATATCGACGACGTGGTATAAGCTGAACGGAGAGGAGAGGTGTTCTATA
 GGCCAGCCTCCATCTTGAGGACGAGCTGATCATCCACAAGGCCGGAGGAGATCAAG
 AACAAAGAATCTAACCGGCCAGAACCAAGGAGACAAGCACCTCAGCTACGACATCGT
 GAAGGATAAGCGGTATAGCAAGGATAAGTTACCTGCACATCCCCATCACAATGAAC 30
 TCGCGTGGATGAGGTGAAGCGGTTCAACGACGCCGTGAACAGGCCATCCGGATCGAT
 GAGAATGTGAACGTGATCGCATCGACCGGGCGAGAGAAATCTGCTGTACGTGGTGGT
 CATCGACTCTAACGGCAACATCCTGGAGCAGATCTCCCTGAACCTCTATCATCAATAAGG
 AGTACGACATCGAGACAGATTATCACGCACTGCTGGATGAGAGGGAGGGCGAGAGAT
 AAGGCCCGGAAGGACTGGAACACCGTGGAGAATATCAGGGACCTGAAGGCCGGCTACCT
 GAGCCAGGTGGTGAACGTGGTGGCCAAGCTGGTGTGAAGTATAATGCCATCATCTGCC
 TGGAGGACCTGAACCTGGCTTCAAGAGGGGCCAGAAGGTGGAGAAGCAGGTGTAC
 CAGAAGTTGAGAAGATGCTGATCGATAAGCTGAATTACCTGGTATCGACAAGAGCCG
 CGAGCAGACATCCCCTAACGGAGCTGGAGGCCCTGAACGCACTGCAGCTGACCTCTA
 AGTTCAAGAGCTTAAAGGAGCTGGCAAGCAGTCCGGCGTGTACTATGTGCCTGCC 40
 TACCTGACCTCTAACGATCGATCCAACCACAGGCTTCGCCAATCTGTTTATATGAAGTG
 TGAGAACGTGGAGAAGTCCAAGAGAGATTCTTGACGGCTTGATTCATCAGGTTCAACG
 CCCTGGAGAACGTGTTCGAGTTGGCTTGTGACCCAACGGCGAGCGCATCATCAAGTATCGGAA
 GGCATCAATTCCAAGTGGACCGTGTGCACCAACGGCGAGCGCATCATCAAGTATCGGAA
 TCCAGATAAGAACAAATATGTTGACGAGAAGGTGGTGGTGGTACCGATGAGATGAAGA
 ACCTGTTGAGCAGTACAAGATCCCCTATGAGGATGGCAGAAATGTGAAGGACATGATC
 ATCAGCAACGAGGAGGCCAGTTCTACCGGAGACTGTATAGGCTGCTGCAGCAGACCC
 GCAGATGAGAACAGCACCTCCGACGGCACAAGGGATTACATCATCTCCCTGTGAAGA
 ATAAGAGAGAGGCCCTACTTCAACAGCAGCTGTCCGACGGCTCTGTGCCAAAGGACGCC 50
 GATGCCAACGGCGCTACAATATGCCAGAAAGGCCGTGGGGTGTGGAGCAGATCAG

GCAGAAGAGCGAGGGCGAGAAGATCAATCTGCCATGACCAACGCCAGTGGCTGGAGT
 ATGCCAGACACACCTGCTAAAAGGCCGGCCACGAAAAAGGCCGGCAGGCAAA
 AAGAAAAAGGGATCCTACCCATACGATGTTCCAGATTACGCTTATCCCTACGACGTGCC
 TGATTATGCATACCCATATGATGTCCCCGACTATGCCTAA

(配列番号 28)

【0143】

NLS 及び HA タグを含む Candidatus Methanoplasma thermatum (CMtCpf1; pY012) :

【0144】

MNNYDEFTKLYPIQKTIRFELKPQGRTEHLETNFNFEEDRDRDRAEKYKILKEAIDEYHKKF
 IDEHLTNMSLDWNSLQKQISEKYYKSREEKDVKVFLSEQKRMQRQEIVSEFKDDRFKDLFS
 KKLFSELLKEEYKKGNHQEIDALKSFDFKSGYFIGLHENRKNMYSGDDEITAISNRVNEN
 FPKFLDNLQKYQEARKKYPEWIKAESALVAHNIMDEVSLEYFNKVLNQEGIQRYNLA
 LGGYVTKSGEKMMGLNDALNLAHQSEKSSKGRIHMTPFKQILSEKESFSYIPDVFTEDS
 QLLPSIGGFFAQIENDKDGNIFDRALELISSYAEYDTERIYIRQADINRVSNVIFGEWGTLG
 GLMREYKADSINDINLERTCKVDKWLDSEKFALESDVLEAIKRTGNNDAFNEYISKMRTA
 REKIDAARKEMKFISEKISGDEESIHIIKTLDSVQQFLHFFNLFKARQDIPLDGAFYAEFD
 EVHSKLFAIVPLYNKVRNYLTKNLNTKKIKLNFKNPTELNGWDQNKVYDYASLIFLRD
 GNYYLGIINPKRKKNIKFEQGSGNGPFYRKMVYKQIPGPKNLPRVFLTSTKGKKEYKPS
 KEIIEGYEADKHIGDKFDLDFCHKLIDFFKESIEKHKDWSKFNFYFSPTESYGDISEFYLD
 VEKQGYRMHFENISAETIDEYVEKGDLFLFQIYNKDFVKAATGKDMHTIYWNAAFSPE
 NLQDVVVVKLNGEAELFYRDKSDIKEIVHREGEILVNRTYNGRTPVPDKIHKKLTDYHNGR
 TKDLGEAKEYLDKVRYFKAHYDITKDRRLNDKIYFHVPLTLNFKANGKKNLNMVIEKF
 LSDEKAHIIGIDRGERNLYYYIIDRSGKIIDQQSLNVIDGFDYREKLNQREIEMKDARQSW
 NAIGKIKDLKEGYLSKAVHEITKMAIQYNAIVVMEELNYGFKRGRFKVEKQIYQKFENMLI
 DKMNYLVFKDAPDESPGGVLNAYQLTNPLESFAKLGKQTGILFYVPAAYTSKIDPTTGFV
 NLFNTSSKTNAQERKEFLQKFESISYSAKDGGIFAFDYRKFGTSKTDHKNVWTAYTNG
 ERMRYIKEKKRNELFDPSEKIEALTSSGIKYDGGQNILPDILRSNNNGLIYTMYSSFIAAI
 QMRVYDGKEDYIISPIKNSKGFFRTDPKRRELPIADANGAYNIALRGELTMRAIAEKFD
 PDSEKMAKLELKHDWFEMQTRGDKRPAATKKAGQAKKKGSYPYDVPDYAYPYDVP
 DYAYPYDVPDYA

(配列番号 29)

【0145】

配列番号 29 は、核局在化シグナル (KRPAAATKKAGQAKKK) (配列番号 12)、その次にグリシン - セリンリンカー (GS)、その次に HA タグ (YPYDVPDYAYPYDVPDYA) (配列番号 13) を含む。

【0146】

配列番号 29 は以下のヌクレオチド配列によってコードされ得る。

【0147】

ATGAACAATTACGACGAGTTACCAAGCTGTATCCTATCCAGAAAACCATCCGGTTTGA
 GCTGAAGCCACAGGGCAGAACCATGGAGCACCTGGAGACATTCAACTTCTTGAGGAGG
 ACCGGGATAGAGCCGAGAAGTATAAGATCCTGAAGGAGGCCATCGACGAGTACCAAG
 AAGTTTATCGATGAGCACCTGACCAATATGTCCTGGATTGAACTCTCTGAAGCAGAT
 CAGCGAGAAGTACTATAAGAGCAGGGAGGAGAAGGACAAGAAGGTGTTCTGTCCGAGC
 AGAAGAGGATGCCAGGAGATCGTGTCTGAGTTAAGAAGGACGATCGCTCAAGGAC
 CTGTTTCCAAGAAGCTTTCTGTGAGCTGCTGAAGGAGGAGATCTACAAGAAGGGCAA
 CCACCAAGGAGATCGACGCCCTGAAGAGCTTCGATAAGTTCCGGCTATTCATCGGCC
 TGCACGAGAATAGGAAGAACATGTACTCCGACGGCGATGAGATCACCGCCATCTCCAAT
 CGCATCGTAATGAGAACCTCCCCAAGTTCTGGATAACCTGCAGAAGTACCAAGGAGGC
 CAGGAAGAAGTATCCTGAGTGGATCATCAAGGCCAGAGCGCCCTGGTGGCCCACAATA

10

20

30

40

50

TCAAGATGGACGAGGTGTTCTCCCTGGAGTACTTAATAAGGTGCTGAACCAGGAGGGC
 ATCCAGCGGTACAACCTGGCCCTGGCGGCTATGTGACCAAGAGCAGGGCAGAAAGATGAT
 GGGCCTGAATGATGCCCTGAACCTGGCCCACCAGTCCGAGAAGAGCTCCAAGGGCAGAA
 TCCACATGACCCCCCTGTTCAAGCAGATCCTGTCCGAGAAGGGAGTCCTCTTACATC
 CCCGACGTGTTACAGAGGATTCTCAGCTGCTGCCTAGCATCGGCGGCTTCTTGCCCA
 GATCGAGAATGACAAGGATGGCAACATCTCGACCGGGCCCTGGAGCTGATCTAGCT
 ACGCCGAGTATGATACCGAGCGGATCTATATCAGACAGGCCGACATCAATAGAGTGTCC
 AACGTGATCTTGGAGAGTGGGGCACCTGGGAGGCCTGATGAGGGAGTACAAGGCCGA
 CTCTATCAATGATATCAACCTGGAGCGCACATGCAAGAAGGTGGACAAGTGGCTGGATT 10
 CTAAGGAGTTGCCCTGAGCGATGTGCTGGAGGCCATCAAGAGGACCGGCAACAATGAC
 GCCTTCAACGAGTATATCTCAAGATGCGGACAGCCAGAGAGAAGATCGATGCCGCCG
 CAAGGAGATGAAGTTCATCAGCGAGAAGATCTCCGGCAGTGGAGGTCTATCCACATCA
 TCAAGACCCCTGCTGGACAGCGTGCAGCAGTTCTGCACTTCTTAATCTGTTAAGGCA
 AGGCAGGACATCCCAC TGATGGAGCCTTCTACGCCGAGTTGACGAGGTGCACAGCAA
 GCTGTTGCCATCGTCCCCGTATAACAAGGTGCGGAACTATCTGACCAAGAACAAATC
 TGAACACAAAGAAGATCAAGCTGAATTCAAGAACCCACTGGCAATGGCTGGAC 20
 CAGAACAAAGGTGACGATTATGCCCTGATCTTCTGCCGGACGGCAATTACTATCT
 GGGCATCATCAATCCTAAGAGAAAGAACATCAAGTTCGAGCAGGGCTGGCAACG
 GCCCTCTACCGGAAGATGGTGTATAAGCAGATCCCCGGCCATAATAAGAACCTGCCA
 AGAGTGTCCCTGACCTCCACAAAGGGCAAGAAGGGAGTATAAGCCCTAAGGAGATCAT
 CGAGGGCTACGAGGCCGACAAGCACATCAGGGCGATAAGTTGACCTGGATTGGTC 30
 ACAAGCTGATCGATTCTTAAGGAGTCCATCGAGAAGCACAAGGACTGGTCAAGTTC
 AACTTCTACTTCAGCCCAACCGAGAGCTATGGCAGACATCTCTGAGTTCTACCTGGATGT
 GGAGAAGCAGGGCTATCGCATGCACCTTGAGAATATCAGCGCCGAGACAATCGACGAGT
 ATGTGGAGAAGGGCGATCTGTTCTGTTCCAGATCTACAACAAGGATTTGTGAAGGCC
 GCCACCGGCAAGAAGGACATGCACACAATCTACTGGAATGCCGCTTCAGCCCCGAGAA
 CCTGCAGGACGTGGTGGTAAGCTGAACGGCGAGGCCGAGCTGTTTATAGGGACAAGT
 CCGATATCAAGGAGATCGTGCACCGCGAGGGCGAGATCTGGTGAATAGGACCTAACAC
 GGCGCACACCAGTGCCGACAAGATCCACAAAGAGCTGACCGATTATCACAATGGCCG
 GACAAAGGACCTGGCGAGGCCAAGGAGTACCTGGATAAGGTGAGATACTTCAAGGCC 40
 ACTATGACATACCAAGGATCGGAGATACCTGAACGACAAGATCTATTCCACGTGCCT
 CTGACCCCTGAACCTCAAGGCCAACGGCAAGAAGAACATCTGAACAAGATGGTCACTGAGAA
 GTTCCTGTCGATGAGAAGGCCACATCATCGGCATCGACAGGGCGAGCGCAATCTGC
 TGTACTATTCCATCATCGACAGGTCTGGCAAGATCATCGATCAGCAGAGCCTGAATGTG
 ATCGACGGCTTGATTATCGGAGAAGCTGAACCAAGAGAGAGATCGAGATGAAGGATGC
 CCGCCAGTCTGGAACGCCATCGCAAGATCAAGGACCTGAAGGAGGGCTACCTGAGCA
 AGGCCGTGCACGAGATACCAAGATGCCATCCAGTATAATGCCATCGTGGTCACTGGAG
 GAGCTGAACCTACGGCTTCAAGCGGGCCGGTTCAAGGTGGAGAAGCAGATCTATCAGAA
 GTTCGAGAATATGCTGATCGATAAGATGAACCTACCTGGTGTGTTAAGGACGACCTGATG
 AGTCCCCAGGAGGCCTGCTGAATGCCCTACCAAGCTGACAAACCCACTGGAGTCTTCGCC
 AAGCTGGCAAGCAGACCGGCATCCTGTTTACGTGCCAGCCGCTATACATCCAAGAT 50
 CGACCCACCACAGGCTTCGTGAATCTGTTAACACCTCTTAAGACAAACGCCCAGG
 AGCGGAAGGAGTTCTGCAGAAGTTGAGAGCATCTCCTATTCTGCCAAGGATGGCGGC
 ATCTTGCCTCGCCTTGACTACAGAAAGTTCGGCACCAGCAAGACAGATCACAAGAA
 CGTGTGGACCGCCTACAAACGGCGAGAGGATCGCCTACATCAAGGAGAAGAAGCGGA
 ATGAGCTGTTGACCCCTTAAGGAGATCAAGGAGGCCCTGACCGAGCTCCGGCATCAAG
 TACGATGGCGGCCAGAACATCCTGCCAGACATCCTGAGGAGCAACAATAACGGCCTGAT
 CTACACAATGTATTCTAGCTCATGCCGCCATCCAGATCGCGGTACGACGGCAAGG
 AGGATTATATCATCAGCCCCATCAAGAACTCCAAGGGCGAGTTCTTAGGACCGACCC
 AAGAGGCGCGAGCTGCCATCGACGCCGATGCCAATGGCGCCTACAACATGCCCTGAG

GGGAGAGCTGACAATGAGGGCAATCGCAGAGAAGTTCGACCCGTATAGCGAGAAAGATGG
 CCAAGCTGGAGCTGAAGCACAAGGATTGGTCGAGTTATGCAGACCAGAGGCAGACAAA
 AGGCCGGGCCACGAAAAAGGCCGCCAGGCAAAAAAGAAAAAGGGATCCTACCCAT
 ACGATGTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATACCCATATGAT
 GTCCCCGACTATGCCTAA

(配列番号 30)

【0148】

NLS 及び HA タグを含む *Eubacterium eligens* (Eecpf1;
 PY013) :

【0149】

MNGNRSIVYREFVGVIPVAKTLRNELRPVGHTQEHIQNGLIQEDELRQEKSTELKNIMD
 DYYREYIDKSLSGVTDLDFTLFLFELMNLVQSSPSKDNKKALEKEQSKMREQICTHLQSDS
 NYKNIFNAKLLKEILPDFIKNYNQYDVKDAGKLETLALFNGFSTYFTDFFEKRKNVFTKE
 AVSTSIAYRIVHENSLIFLANMTSYKKISEKALDEIEVIEKNNQDKMGDWELNQIFNPDFY
 NMVLIQSGIDFYNEICGVVNAHMNLYCQQTKNNYNLFKMRKLHKQILAYTSTSFEVPKM
 FEDDMCSVNAVNAFIDETEKGNIIGKLKDIVNKYDELDEKRIYISKDFYETLSCFMMSGNWN
 LITGCVENFYDENIHAKGKSKEEKVKKAVKEDKYKSINDVNDLVEKYIDEKERNEFKNSN
 AKQYIREISNIITDTETAHLEYDDHISLIESEEKADEMKKRLDMYMNMYHWAKAFIVDEV
 LDRDEMFSIDDIYNILENIVPLYNRVNYVTQKPYNSSKKIKLNFQSPTLANGWSQSKE
 FDNNAIILIRDNKYYLAIFNAKNPKDKIIQGNNSDKNDNDYKKMVYNLLPGANKMLPKV
 FLSKKGIETFKPSDYIISGYNAHKHIKTSENFDISFCRDLIDYFKNSIEKHAERWKYEFKFSA
 TDSYSDFSEFYREVEMQGYRIDWTYISEADINKLDEEGKIYLFQIYNKDFAENSTGKENLH
 TMYFKNIFSEENLKDIILKNGQAELFYRRASVKNPVKHKKDSVLVNKTYKNQLDNGDVV
 RIPIPDDIYNEIYKMYNGYIKESDLSEAAKEYLDKVEVRTAQKDIVKDYRTVDKYFIHTPI
 TINYKVTARNNVNDMVVKYIAQNDDIHIGIDRGERNLIYISVIDSHGNIVKQKSYNILNN
 YDYKKKLVEKEKTREYARKNWKSIGNIKELKEGYISGVVHEIAMLIVEYNAAIAMEDLNYG
 FKGRRFKVERQVYQKFESMLINKLNYFASKEKSVDEPGGLLKGYQLTYVPDNIKNLGKQC
 GVIFYVPAFTSKIDPSTGFISAFNFKSISTNASRKQFFMQFDEIRYCAEKDMFSFGFDYNN
 FDTYNITMGKTQWTVYTNGERLQSEFNNARRTGKTKSINLTETIKLLEDNEINYADGHD
 IRIDMEKMDDEKKSEFFAQLLSLYKLTQMRNSYTEAEEQENGISYDKIISPVINDEGEFF
 DSDNYKESDDKECKMPKDADANGAYCIALKGLYEVLKIKSEWTEDGFDRNCLKLPHAEW
 LDFIQNKRYEKRPAATKKAGQAKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA

(配列番号 31)

【0150】

配列番号 31 は、核局在化シグナル (K R P A A T K K A G Q A K K K K) (配列番号 12)、その次にグリシン - セリンリンカー (G S)、その次に HA タグ (Y P Y D V P D Y A Y P Y D V P D Y A) (配列番号 13) を含む。

【0151】

配列番号 31 は以下のヌクレオチド配列によってコードされ得る。

【0152】

ATGAACGGCAATAGGTCCATCGTGTACCGCGAGTTCGTGGCGTGATCCCCGTGGCAA
 GACCTGAGGAATGAGCTGCGCCCTGTGGGCCACACACAGGAGCACATCATCCAGAAC
 GCCTGATCCAGGAGGACGAGCTGCGGCAGGAGAAGAGCACCAGCTGAAGAACATCATG
 GACGATTACTATAGAGAGTACATCGATAAGTCTCTGAGCAGCGTGACCGACCTGGACTT
 CACCTGCTGTTGAGCTGATGAACCTGGTGAGAGCTCCCCCTCCAAGGACAATAAGA
 AGGCCCTGGAGAAGGAGCAGTCTAACGATGAGGGAGCAGATCTGCACCCACCTGCAGTCC
 GACTCTAACTACAAGAATATCTTAAACGCCAAGCTGCTGAAGGAGATCCTGCCTGATT
 CATCAAGAACTACAATCAGTATGACGTGAAGGATAAGGCCGGCAAGCTGGAGACACTGG
 CCCTGTTAATGGCTTCAGCACATACTTACCGACTTCTTGAGAAGAGGAAGAACGTG
 TTCACCAAGGAGGCCGTGAGCACATCCATCGCCTACCGCATCGTGCACGAGAACTCCCT

10

20

30

40

50

GATCTTCCTGGCCAATATGACCTCTTATAAGAAGATCAGCGAGAAGGCCCTGGATGAGA
 TCGAAGTGATCGAGAAGAACATCAGGACAAGATGGCGATTGGAGCTGAATCAGATC
 TTTAACCTGACTTCTACAATATGGTCTGATCCAGTCGGCATCGACTTCTACAACGA
 GATCTGCGGCGTGGTGAATGCCACATGAACCTGTACTGTCAGCAGACCAAGAACATT
 ATAACCTGTTCAAGATGCGGAAGCTGCACAAGCAGATCCTGGCCTACACCAGCACCGC
 TTTCGAGGTGCCAAGATGTTCGAGGACGATATGAGCGTGTATAACGCCGTGAACGCC
 CATCGACGAGACAGAGAACATCATCGGAAAGCTGAAGGATATCGTGAATAAGT
 ACGACGAGCTGGATGAGAAGAGAATCTATATCAGCAAGGACTTTACGAGACACTGAGC
 TGCTTCATGTCCGGCAACTGGAATCTGATCACAGGCTGCGTGGAGAACTTCTACGATGA
 GAACATCCACGCCAAGGGCAAGTCCAAGGAGGAGAAGGTGAAGAACGCCGTGAAGGAG
 GACAAGTACAAGTCTATCAATGACGTGAACGATCTGGTGGAGAAGTATATCGATGAGAA
 GGAGAGGAATGAGTTCAAGAACAGCAATGCCAAGCAGTACATCCGCGAGATCTCAACA
 TCATCACCGACACAGAGACAGCCCACCTGGAGTATGACGATCACATCTCTGATCGAG
 AGCGAGGAGAACGCCGACGAGATGAAGAACGGCTGGATATGTATATGAACATGTACCA
 CTGGGCCAAGGCCCTTATCGTGGACGAGGTGCTGGACAGAGATGAGATGTTCTACAGCG
 ATATCGACGATATCTATAATATCCTGGAGAACATCGTGCCTACTGTATAATGGTGAGA
 AACTACGTGACCCAGAACGCCCTACAACCTCTAACAGAACATCAAGCTGAATTCCAGAGCC
 TACACTGGCCAATGGCTGGTCCCAGTCTAACCGAGTTGACAAACAATGCCATCATCCTGA
 TCAGAGATAACAAGTACTATCTGCCATCTCAATGCCAAGAACAGCCAGACAAGAAC
 ATCATCCAGGGCAACTCCGATAAGAACGACAACGATTACAAGAACATGGTGTATAA
 CCTGCTGCCAGGCCAACAGATGCTGCCAAGGTGTTCTGTCTAACAGGCC
 AGACATTCAAGCCCTCCGACTATATCATCTGGCTACAAACGCCAACAGCACATCAAG
 ACAAGCGAGAATTTGATATCTCCTCTGTGCGGACCTGATCGATTACTCAAGAACAG
 CATCGAGAAGCACGCCGAGTGGAGAAAGTATGAGTTCAAGTTCCGCCACCGACAGCT
 ACTCCGATATCTCTGAGTTCTATCGGAGGTTGGAGATGCAGGGCTACAGAACATGACTGG
 ACATATATCAGCGAGGCCGACATCAACAGCTGGATGAGGAGGGCAAGATCTATCTGTT
 TCAGATCTACAATAAGGATTTGCCGAGAACAGCACCGCAAGGAGAATCTGCACACAA
 TGTACTTTAACACATCTCTCCGAGGAGAATCTGAAGGACATCATCATCAAGCTGAAC
 GGCCAGGGCGAGCTGTTTATCGGAGAGCCTCTGTGAAGAACATCCGTGAAGCACAAGAA
 GGATAGCGTGCTGGTGAACAAGACCTACAAGAACATCAGCTGGACAACGCCGACGTGGTGA
 GAATCCCCATCCCTGACGATATCTATAACGAGATCTAACAGATGTATAATGGCTACATC
 AAGGAGTCCGACCTGCTGAGGCCCAAGGAGTACCTGGATAAGGTGGAGGTGAGGAC
 CGCCCAAGGACATCGTAAGGATTACCGCTATACAGTGGACAAGTACTTCATCCACA
 CACCTATCACCATCAACTATAAGGTGACCGCCGCAACAATGTGAATGATATGGTGGTGA
 AAGTACATCGCCAGAACGACGATATCCACGTGATCGGCATCGACCGGGCGAGAGAAA
 CCTGATCTACATCTCGTGTACGATTCTCACGGCAACATCGTAAGCAGAACATCCTACA
 ACATCCTGAACAACACTACGACTACAAGAACAGCTGGTGGAGAACGGAGAAAACCCGGAG
 TACGCCAGAAAGAACATGGAAGAGCATCGCAATATCAAGGAGCTGAAGGAGGGCTATAT
 CTCCGGCGTGGTGCACGAGATGCCATGCTGATCGTGGAGTACAACGCCATCATGCCA
 TGGAGGACCTGAATTATGGCTTAAAGAGGGCCGCTCAAGGTGGAGCGGCAGGTGTAC
 CAGAAGTTGAGAGCATGCTGATCAATAAGCTGAACATTTCGCCAGCAAGGAGAACGTC
 CGTGGACGCCAGGAGGCCGCTGCTGAAGGGCTATCAGCTGACCTACGTGCCGATAATA
 TCAAGAACCTGGCAAGCAGTGCAGCGTGTATCTCTGCCTTCAACTTAAGTCTATCAGCACAAA
 AAGATCGACCCATCCACAGGCTTTATCTCTGCCTTCAACTTAAGTCTATCAGCACAAA
 TGCCTCTCGGAAGCAGTTCTTATGCAGTTGACGAGATCAGATACTGTGCCGAGAAGG
 ATATGTTCAGCTTGGCTCGACTACAACAACTTCGATACCTACAACATCACAATGGGC
 AAGACACAGTGGACCGGTATACAAACGGCGAGAGACTGCAGTCTGAGTTCAACAATGCG
 CAGGCGACCGGCAAGAACAGAGCATCAATCTGACAGAGACAATCAAGCTGCTGCTGG
 AGGACAATGAGATCAACTACGCCACGGCCACGATATCAGGATCGATATGGAGAACAGT
 GACGAGGATAAGAACAGCGAGTTCTTGCCAGCTGCTGAGCCTGTATAAGCTGACCGT

10

20

30

40

50

GCAGATGCGCAATT CCTATA CAGAGGCCGAGGAGCAGGAGAACGGCATCTCTTACGACA
 AGATCATCAGCCCTGTGATCAATGATGAGGGCGAGTTCTTGACTCCGATAACTATAAG
 GAGTCTGACGATAAGGAGTGCAAGATGCCAAAGGACGCCATGCCAACGGCGCTACTG
 TATGCCCTGAAGGGCTGTATGAGGTGCTGAAGATCAAGAGCGAGTGGCTGGACTTCATCCAGAAC
 GCTTGATAGGAATTGCCTGAAGCTGCCACACGCAGAGTGGCTGGACTTCATCCAGAAC
 AAGCGGTACGAGAAAAGGCCGGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAA
 AGGGATCCTACCCATACGATGTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTAT
 GCATACCCATATGATGTCCCCGACTATGCCTAA

(配列番号 32)

【0153】

N L S 及び H A タグを含む Moraxella bovoculi 237 (M b C p f
 1 ; p Y 0 1 4) :

【0154】

MLFQDFTHLYPLSKTVRFELKPIDRTLEHIHAKNFLSQDET MADMHQKVVKVILDDYHRD
 FIADMMGEVKLT KLA EFYDVYLKFRKNPKDDELQKQLKDLQAVLRKEIVKPIGNNGKYK
 AGYDRLFGAKLFKD GKELGDLAKFVIAQE GE S P K LAH LAH F E K F S T Y F T G F H D N R K N M Y
 S D E D K H T A I Y R L I H E N L P R F I D N L Q I L T T I K Q K H S A L Y D Q I I N E L T A S G L D V S L A S H L D G Y
 H K L L T Q E G I T A Y N T L L G G I S G E A G S P K I Q G I N E L I N S H H N Q H C H K S E R I A K L R P L H K Q I L S D
 GMSVSFLPSKFADDSEMCQAVNEFYRHADVF AKVQSLFDGFDDHQKDGIYVEHKNLNE
 LSKQAFGDFALLGRVLDGYV DVVNPEFNERFAKAKTDNAKAKLTKEKDKFIKGVHSLAS
 LEQAIEHYTARH DDES VQAGKLGQYFKHGLAGVDNPIQKIHNNHSTIKGFLERERPAGER
 ALPKIKSGKNPEMTQLRQLKELL DNA LNVAHFAKLLTTKTL DNQDGNFYGEFGVLYDE
 LAKIPTLYNKVRDYL SQKP FSTE KYKLNFGNPTLLNGWDLNKEDNFGVILQKDGCYYLA
 LLDKAHKKVFDNAPNTGKSIYQKMIYKYLEVRKQFPKVFFSKEAIAINYHPSKELVEIKDK
 GRQRSDDERLKLYRFILECLKIHPKYDKKFE GAIGDIQLFKKKKGREVPISEKDLFDKING
 IFSSKP KLEMEDFFIGEFKRYNPSQDLVDQYNIYKKIDSNDNRKKENFYNNHPFKKKDLV
 RYYYESMCKHEEWEESFEFSKKLQDIGCYVDVNELFTIETRRLNYKISFCNIADYIDEL
 VEQGQLYLFQIYNKDFSPKAHGKPNLHTLYFKALFSEDNLADPIYKLNGEAQIFYRKASL
 DMNETTIHAGEVLENKNPDNP KKRQFVYDIIKDKRYTQDKFMLHVPITMNFGVQGMTI
 KEFNKKVNQS IQYDEVNVIGIDRGERHLLYLT VINSKG EILEQCSLNDITTASANGTQMT
 TPYHKILDKREIERLNARVGWGEIETIKELKSGYL SHV VH QISQLMLKYN AIVVLEDLNFG
 FKGRRFKVEKQIYQNFENALIKKLNLHVLKDKADDEIGSYKNALQLTNNFTDLKSIGKQT
 GFLFYVPAWNTSKIDPETGFV DLLKPRYENIAQSQAFFGKFDKICYNADKDYFEF HIDYAK
 FTDKAKNSRQIWTICSHGD KRYVYDKTANQNKGAAKGINVNDELKSLFARHHINEKQP
 LVMDICQNNDFEFHKSLMYLLKTL ALRYSNASSDED FILSPVANDEGVFFNSALADDTQ
 PQNADANGAYHIALKGLWLLNELKNSDDL NKVKLAIDNQT WLNFAQNRKRP AATKKAG
 QAKKKKG SYPYDVPDYA YPYDVPDYA YPYDVPDYA

(配列番号 33)

【0155】

配列番号 33 は、核局在化シグナル (K R P A A T K K A G Q A K K K K) (配列番号
 12) 、その次にグリシン - セリンリンカー (G S) 、その次に H A タグ (Y P Y D V P
 D Y A Y P Y D V P D Y A Y P Y D V P D Y A) (配列番号 13) を含む。

【0156】

配列番号 33 は以下のヌクレオチド配列によってコードされ得る。

【0157】

ATGCTGTCAGGACTTTACCCACCTGTATCCACTGTCCAAGACAGTGAGATTGAGCT
 GAAGCCCATCGATAGGACCTGGAGCACATCCACGCCAAGAACTTCCGTCTCAGGAGC
 AGACAATGGCCGATATGCACCAGAAGGTGAAAGTGATCCTGGACGATTACCACCGCGAC
 TTCATCGCCGATATGATGGCGAGGTGAAGCTGACCAAGCTGGCCAGTTCTATGACGT
 GTACCTGAAGTTCGGAAGAACCAAAGGACGATGAGCTGCAGAAGCAGCTGAAGGATC

10

20

30

40

50

TGCAGGCCGTGCTGAGAAAGGAGATCGTGAAGCCATCGCAATGGCGGCAAGTATAAG
 GCCGGCTACGACAGGCTGTTGGCGCCAAGCTGTTAAGGACGGCAAGGAGCTGGCGA
 TCTGGCCAAGTCGTGATCGCACAGGAGGGAGAGCTCCCCAAAGCTGGCCCACCTGG
 CCCACTTCGAGAAGTTTCCACCTATTCACAGGCTTCACGATAACCGGAAGAACATG
 TATTCTGACGAGGATAAGCACACCGCCATCGCCTACCGCCTGATCCACGAGAACCTGCC
 CCGGTTATCGACAATCTGCAGATCCTGACCACAATCAAGCAGAACGACTCTGCCCTGT
 ACGATCAGATCATCAACGAGCTGACGCCAGCGGCTGGACGTGTCCTGGCCAGCAC
 CTGGATGGCTATCACAAGCTGCTGACACAGGGCATCACCGCCTACAATACACTGCT
 GGGAGGAATCTCCGGAGAGGCAGGCTCTCCTAAGATCCAGGGCATCAACGAGCTGATCA
 ATTCTCACCACAAACCAGCACTGCCACAAGAGCGAGAGAACGCCAGCTGAGGCCACTG
 CACAAGCAGATCCTGTCGACGGCATGAGCGTGTCCCTGCCCTTAAGTTGCCGA
 CGATAGCGAGATGTGCCAGGCCGTGAACGAGTTCTATGCCACTACGCCACGTGTTCG
 CCAAGGTGCAGAGCCTGTTGACGGCTTGACGATCACCAAGAAGGATGGCATCTACGTG
 GAGCACAAGAACCTGAATGAGCTGCTCAAGCAGGCCCTCGCGACTTGCACTGCTGG
 ACGCGTGGACGGATACTATGTGGATGTGGTAATCCAGAGTTCAACGAGCGGTTTG
 CCAAGGCCAAGACCGACAATGCCAAGGCCAGCTGACAAAGGAGAACGGATAAGTCATC
 AAGGGCGTGCACCCCTGCCCTCTGGAGCAGGCCATCGAGCACTATACCGCAAGGCA
 CGACGATGAGAGCGTGCAGGCAGGCAAGCTGGACAGTACTTCAAGCACGCCCTGGCG
 GAGTGGACAACCCATCCAGAAGATCCACAACAAATCACAGCACCATCAAGGCCCTTG
 GAGAGGGAGCGCCCTGCAGGAGAGAGGCCCTGCCAAGATCAAGTCCGGCAAGAACCTC
 TGAGATGACACAGCTGAGGCAGCTGACGGAGCTGCTGGATAACGCCCTGAATGTGGCC
 ACTTCGCCAAGCTGCTGACCAACAAAGACCACACTGGACAATCAGGATGGCAACTTCTAT
 GGCGAGTTGGCGTGTGACGAGCTGGCAAGATCCCCACCCCTGTATAACAAGGT
 GAGAGATTACCTGAGCCAGAACGCCCTTCTCCACCGAGAAGTACAAGCTGAACCTTGCA
 ATCCAACACTGCTGAATGGCTGGACCTGAACAAGGAGAACGGATAATTCCGGCGTGTAC
 CTGCAGAAGGACGGCTGCTACTATCTGCCCTGCTGGACAAGGCCACAAGAACGGTGT
 TGATAACGCCCTAATACAGGCAAGAGCATCTATCAGAAGATGATCTATAAGTACCTGG
 AGGTGAGGAAGCAGTCCCCAAGGTGTTCTTCCAAGGAGGCCATGCCATCAACTAC
 CACCCCTCTAAGGAGCTGGTGGAGATCAAGGACAAGGCCGGCAGAGATCCGACGATGA
 GCGCCTGAAGCTGTATCGGTTATCCTGGAGTGTCTGAAGATCCACCCCTAAGTACGATA
 AGAAGTTCGAGGGCGCCATCGCGACATCCAGCTGTTAAGAAGGATAAGAACGGCAGA
 GAGGTGCCAATCGCGAGAACCTGTTGATAAGATCAACGGCATCTTCTAGCAA
 GCCTAAGCTGGAGATGGAGACTTCTTATCGCGAGTCAAGAGGTATAACCAAGCC
 AGGACCTGGTGGATCAGTATAATATCTACAAGAAGATCGACTCAACGATAATCGCAAG
 AAGGAGAATTCTACAACAATCCCCAAGTTAAGAAGGATCTGGTGGTACTATT
 CGAGTCTATGTGCAAGCAGCAGGAGTGGAGAGCTCGAGTTCCAAGAACAGCTGC
 AGGACATCGGCTGTTACGGATGTGAACGAGCTGTTACCGAGATCGAGACACGGAGA
 CTGAATTATAAGATCTCCTCTGCAACATCAATGCCACTACATCGATGAGCTGGTGG
 GCAGGGCCAGCTGTATCTGTTCCAGATCTACAACAAGGACTTTCCCCAAGGCCACG
 GCAAGGCCAATCTGCACACCCCTGACTTCAGGCCCTGTTCTGAGGACAACCTGCC
 GATCCTATCTATAAGCTGAATGGCGAGGCCAGATCTTCTACAGAAAGGCCCTCCCTGGA
 CATGAACGAGACAACAATCCACAGGGCCGGCGAGGTGCTGGAGAACAGAACATCCCGATA
 ATCCTAAGAAGAGACAGTCGTGACGACATCATCAAGGATAAGAGGTACACACAGGAC
 AAGTTCATGCTGCACGTGCCAATCACCAGAACCTGGAGAGCTATCCAGCAGTATGAC
 GGAGTTCAATAAGAAGGTGAACCAGTCTACAGCAGTATGACGAGGTGAACGTGATCG
 GCATCGATCGGGCGAGAGACACCTGCTGTACCGTACCGTGAATAGCAAGGGCGAG
 ATCCTGGAGCAGTGGCTGTTCCCTGAACGACATCACCAACAGCCTCTGCCAATGGCACACAGAT
 GACCACACCTTACCAAGATCCTGGATAAGAGGGAGATCGAGCGCCTGAACGCCCGGG
 TGGGATGGGGCGAGATCGAGACAATCAAGGAGCTGAAGTCTGGCTATCTGAGGCCACGTG
 GTGCACCAAGATCAGCCAGCTGATGCTGAAGTACAACGCCATCGTGGTGTGGAGGACCT 50

GAATTCGGCTTAAGAGGGGCCGCTTAAGGTGGAGAAGCAGATCTATCAGAACTTCG
 AGAATGCCCTGATCAAGAACGCTGAACCACCTGGTCTGAAGGACAAGGCCGACGATGAG
 ATCGGCTCTTACAAGAACGCTGCCCTGCAGCTGACCAACAATTACAGATCTGAAGAGCAT
 CGGCAAGCAGACCGGCTTCTGTTTATGTGCCCGCCTGAAACACCTCTAACAGATCGACC
 CTGAGACAGGCTTGTGGATCTGCTGAAGCCAAGATAACGAGAACATCGCCCAGAGCCAG
 GCCTTCTTGGCAAGTTGACAAGATCTGCTATAATGCCGACAAGGATTACTCGAGTT
 TCACATCGACTACGCCAAGTTACCGATAAGGCCAAGAACAGCCAGATCTGGACAA
 TCTGTTCCCACGGCGACAAGCGGTACGTGTACGATAAGAACAGCCAAACAGAACATAAGGGC
 GCCGCCAAGGGCATCAACGTGAATGATGAGCTGAAGTCCCTGTTGCCGCCACCAT
 CAACGAGAACGAGCCCCAACCTGGTCATGGACATCTGCCAGAACATGATAAGGAGTTTC
 ACAAGTCTCTGATGTACCTGCTGAAAACCTGCTGCCCTGCCGTACAGCAACGCCCTCC
 TCTGACGAGGATTTCATCCTGTCCTGGCAAACGACGAGGGCGTGTCTTAAAG
 CGCCCTGGCCGACGATACACAGCCTCAGAACATGCCGATGCCAACGGCGCTACCCACATCG
 CCCTGAAGGGCCTGTGGCTGCTGAATGAGCTGAAGAACTCCGACGATCTGAACAAAGGTG
 AAGCTGGCCATCGACAATCAGACCTGGCTGAATTGCCCAGAACAGGAAAGGCCGC
 GGCCACGAAAAAGGCCGCCAGGCAAAAAGAAAAAGGGATCCTACCCATACGATGTT
 CAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGTCCCCGAC
 TATGCCTAA

(配列番号34)

【0158】

NLS及びHAタグを含むLeptospira inadai(LiCpf1; pY
015) :

【0159】

MEDYSGFVNIIYSIQKTLRFELKPVGKTLEHIEKKFLKKDKIRAEDYKAVKKIIDKYHRAYI
 EEVFDSDLHQKKKKDKTRFSTQFIKEIKEFSELYYKTEKNIPDKERLEALSEKLRKMLVGA
 FKGEFSEEVAEKYKNLFSKELIRNEIEKFCEDEERKQVSNFKSFTTYFTGFHSNRQNIYSD
 EKKSTAIGYRIIHQNLPKFLDNLKIIESIQRRFKDFPWSDLKKNLKKIDKNIKLTEYFSIDGF
 VNVLNQKGIDAYNTILGGKSEESGEKIQGLNEYINLYRQKNNIDRKNLPNVKILFKQILGD
 RETKSFIPEAFPDDQSVLNSITEFAKYLKLDKKKSIIAEKKFLSSFNRYELDGIYLANDN
 SLASISTFLFDDWSFIKKSVFSKYDESVGDPKKKIKSPLKYEKEKEKWLKQKYYTISFLND
 AIESYSKSQDEKRVKIRLEAYFAEFKSKDDAKQFDLLERIEEAYAIVEPLLGAEPYPRDRNL
 KADKKEVGKIKDFLDSIKSLQFFLKPLLSAEIFDEKDLGFYNQLEGYYEEIDSIGHLYNKVR
 NYLTGKIYSKEKFKNFENSTLLKGWDENREVANLCVIFREDQKYYLGVMKDKENNTILSD
 IPKVKPNELFYEKMVYKLIPTPHMQLPRIIFSSDNLSIYNPSKSILKIREAKSFKEGKNFKL
 KDCHKFIDFYKESISKNEDWSRFDFKFSKTSSYENISEFYREVERQGYNLDFKKVSKFYIDS
 LVEDGKLYLFQIYNKDFSIFSKGKPNLHTIYFRSLFSKENLKDVCVCLKLNGEAEMFFRKKSIN
 YDEKKKREGHHPELFEKLKYPILKDKRYSEDKFQFHLPISLNFKSKERLNFNLKVNEFLKR
 NKDINIIGIDRGERNLLYLVMINQKGEILKQTLLDSMQSGKGRPEINYKEKLQEKEIERDK
 ARKSWGTVENIKELKEGYLSIVIHQISKLMVENNAIVVLEDLNIGFKRGRQKVERQVYQKF
 EKMLIDKLNFLVFKENKPTEPGGVLKAYQLTDEFQSFEKLSKQTGFLFYVPSWNTSKIDP
 RTGFIDFLHPAYENIEKAKQWINKFDSIRFNKMDWFEFTADTRKFSENMLGKNRVWVI
 CTTNVERYFTSKTANSSIQYNSIQITEKLKELFVDIPFSNGQDLKPEILRKNDAVFFKSLLF
 YIKTTLSSLRQNNGKKGEEEKDFILSPVVDSKGRFFNSLEASDDEPKDADANGAYHIALKGL
 MNLLVLNETKEENLSRPWKWKIKNWDLEFVWERNRKRPAATKKAGQAKKKGSYPYDV
 PDYAYPYDVPDYAYPYDVPDYA

(配列番号35)

【0160】

配列番号35は、核局在化シグナル(KRPAATKKAGQAKKK) (配列番号12)、その次にグリシン-セリンリンカー(GS)、その次にHAタグ(YPYDVPDYAYPYDVPDYA) (配列番号13)を含む。

10

20

30

40

50

【0161】

配列番号35は以下のヌクレオチド配列によってコードされ得る。

【0162】

ATGGAGGACTATTCCGGCTTGTGAACATCTACTCTATCCAGAAAACCCTGAGGTTCGA
 GCTGAAGGCCAGTGGGCAAGACACTGGAGCACATCGAGAAGAAGGGCTTCCTGAAGAAGG
 ACAAGATCCGGGCCGAGGATTACAAGGCCGTGAAGAAGATCATCGATAAGTACCAACAGA
 GCCTATATCGAGGAGGTGTTGATTCCGTGCTGCACCAAGAAGAAGAAGAAGGACAAGAC
 CCGCTTTCTACACAGTTCATCAAGGAGATCAAGGAGTTCAAGCAGCTGTACTATAAGA
 CCGAGAAGAACATCCCCGACAAGGAGAGGCTGGAGGCCCTGAGCGAGAAGCTGCGCAAG
 ATGCTGGTGGCGCCTTAAGGGCAGTTCTCCGAGGAGGTGGCCGAGAAAGTATAAGAA 10
 CCTGTTTCTAAGGAGCTGATCAGGAATGAGATCGAGAAGTTCTGCGAGACAGACGAGG
 AGCGCAAGCAGGTGTCTAAGTCAGAGCCTCACACATACTTACCGGCTTCACTCC
 AACAGGCAGAATATCTATTCCGACGAGAAGAAGTCTACAGCCATCGGCTACCGCATT
 CCACCAAGAACCTGCCTAAGTTCTGGATAATCTGAAGATCATCGAGTCCATCCAGCGC
 GGTTCAAGGACTTCCATGGTCTGATCTGAAGAAGAACCTGAAGAAGATCGATAAGAAT
 ATCAAGCTGACCGAGTACTTCAGCATCGACGGCTTCGTGAACGTGCTGAATCAGAAGGG
 CATCGATGCCTACAACACAATCCTGGCGCAAGTCCGAGGAGTCTGGCGAGAAGATCC
 AGGGCCTGAACGAGTACATCAATCTGTATCGCAGAAGAACATATCGACAGAAAGAAC
 CTGCCAATGTGAAGATCCTGTTAACAGCAGATCCTGGCGATAGGGAGACAAAGAGCTT
 TATCCCTGAGGCCCTCCAGACGATCAGTCCGTGCTGAACCTATCACAGAGTTGCGCA 20
 AGTACCTGAAGCTGGATAAGAAGAAGAGCATCATCGCCGAGCTGAAGAAGTTCTG
 AGCTCCTTCAATCGCTACGAGCTGGACGGCATCTATCTGCCAACGATAATAGCCTGGC
 CTCTATCAGCACCTCCTGTTGACGATTGGCCTTTATCAAGAAGTCCGTGCTTCAA
 GTATGACGAGTCCGTGGCGACCCAAGAAGAAGATCAAGTCTCCCTGAAGTACGAGA
 AGGAGAAGGAGAAGTGGCTGAAGCAGAAGTACTATACAATCTTCTGAAACGATGCC
 ATCGAGAGCTATTCAAGTCTCAGGACGAGAAGAGGGTGAAGATCCGCTGGAGGCCA
 CTTTGCCAGTTCAAGAGCAAGGACATGCCAGAAGCAGTTGACCTGCTGGAGAGGA
 TCGAGGAGGCATGCCATCGTGGAGCCTCTGCTGGAGCAGAGTACCCAAAGGGACCGC
 AACCTGAAGGCCGATAAGAAGGAAGTGGCAAGATCAAGGACTTCTGGATAGCATCAA
 GTCCCTGCAGTTCTTCTGAAGCCTCTGCTGTCCGCCAGATCTTGACGAGAAGGATC 30
 TGGGCTTCTACAATCAGCTGGAGGGCTACTATGAGGAGATCGATTCTATCGGCCACCTG
 TATAACAAGGTGCGGAATTATCTGACCGGAAGATCTACAGCAAGGAGAAGTTAACG
 GAACCTCGAGAACAGCACCTGCTGAAGGGCTGGACGAGAACCGGGAGGTGGCCAATC
 TGTGCGTGTCTCAGAGAGGACAGAACGAGTACTATCTGGCGTGTGGATAAGGAGAAC
 AATACCATCCTGTCGACATCCCCAACGGTGAAGCCTAACGAGCTGTTTACGAGAACG
 GGTGTATAAGCTGATCCCCAACCTCACATGCAGCTGCCGGATCATCTCTTAGCG
 ACAACCTGTCTATCTATAATCCTAGCAAGTCCATCCTGAAGATCAGAGAGGCCAAGAGC
 TTTAAGGAGGGCAAGAACCTCAAGCTGAAGGACTGTCACAAGTTATCGATTCTACAA
 GGAGTCTATCAGCAAGAACGAGACTGGAGCAGATTGACTTCAAGTTCAAGCTGAGAAC
 GCAGCTACGAGAACATCAGCGAGTTTACCGGGAGGTGGAGAGACAGGGCTATAACCTG
 GACTTCAAGAACGGTGTCTAAGTTCTACATCGACAGCCTGGTGGAGGATGGCAAGCTGTA 40
 CCTGTTCCAGATCTATAACAAGGACTTTCTATCTTCAGCAAGGGCAAGCCAATCTGC
 ACACCATCTATTTGGTCCCTGTTCTCAAGGAGAACCTGAAGGACGTGCTGCCTGAAG
 CTGAATGGCGAGGCCGAGATGTTCTTGGAGAAGAAGTCCATCAACTACGATGAGAAC
 GAAGCGGGAGGGCACCACCCGAGCTGTTGAGAAGCTGAAGTATCCTATCCTGAAGG
 ACAAGAGATACAGCGAGGATAAGTTCAAGCTGAGGACTGTCACAAGTTATCGATTCTACAA
 TCCAAGGAGCGGCTGAACCTTAATCTGAAAGTGAATGAGTCTCTGAAGAGAACAGGA
 CATCAATATCATCGGCATCGATGGGGCGAGAGAACCTGCTGTACCTGGTCATGATCA
 ATCAGAAGGGCGAGATCCTGAAGCAGACCTGCTGGACAGCATGCAGTCCGGCAAGGGC
 CGGCCTGAGATCAACTACAAGGAGAACGCTGCAGGAGAGATCGAGAGGGATAAGG 50

CCCGCAAGAGCTGGGCACAGTGGAGAATATCAAGGAGCTGAAGGAGGGCTATCTGTCT
 ATCGTGATCCACCAAGATCAGCAAGCTGATGGTGGAGAACATGCCATCGTGGTCTGG
 GGACCTGAACATCGGCTTAAGCGGGGCAGACAGAACGGTGGAGCGGCAGGTGTACCAGA
 AGTTGAGAAGATGCTGATAAGCTGAACTTCTGGTCTAAGGAGAACAGCCA
 ACCGAGCCAGGAGGCAGTCAGGAGGCAGTCAGCTGACAGACGAGTTCAAGGAGAACAGCCA
 GAAGCTGAGCAAGCAGACCGGCTTCTGTTCTACGTGCCAGCTGGAAACACCTCCAAGA
 TCGACCCCAGAACAGGCTTATCGATTCTGCACCCCTGCCTACGAGAACAGGAGAACAGCCA
 GCCAAGCAGTGGATCAACAAGTTGATTCCATCAGGTTCAATTCTAAGATGGACTGGTT
 TGAGTTCACCGCCGATACACGCAAGTTTCCGAGAACCTGATGCTGGCAAGAACAGCCA
 TGTGGGTCTCATCTGCACCACAAATGTGGAGCGGTACTTCACCAGCAAGAACAGGCC
 TCCATCCAGTACAATAGCATCCAGATCACCGAGAACAGCTGAAGGAGCTGTTGTGGACAT
 CCCTTCAGCAACGGCCAGGATCTGAAGCCAGAGATCCTGAGGAAGAACAGGCC
 TCTTTAAGAGCCTGCTGTTTACATCAAGACCACACTGTCCCTGCAGAACAAATGGC
 AAGAAGGGCGAGGAGGAGAACGGACTTCATCCTGAGCCCAGTGGTGGATTCCAAGGGCG
 GTTCTTAAGACTCTGGAGGCCAGCGACGATGAGCCAAAGGACGCCAGTGC
 CCTACACATGCCCTGAAGGCCGATGAACCTGCTGGTGTGAATGAGAACAAAGGAG
 GAGAACCTGAGCAGACCAAAGTGGAGAACAGAACAAAGGACTGGCTGGAGTTGTG
 GGAGAGGAACCGCAAAAGGCCGGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAA
 AAGGGATCCTACCCATACGATGTTCCAGATTACGCTTACGACGTGCCTGATTA
 TGCATACCCATATGATGTCCTGGACTATGCCTAA

10

(配列番号 3 6)

【 0 1 6 3 】

N L S 及び H A タグを含む *Lachnospiraceae bacterium* ND
2 0 0 6 (L b C p f 1 ; p Y 0 1 6) :

【 0 1 6 4 】

MSKLEKFTNCYSLSKTLRFKAIPVGKTQENIDNKRLLVEDEKRAEDYKGVKLLDRYYLSF
 INDVLHSIKLKNLNNYISLFRKKTRTEKENKELENLEINLRKEIAKAFKGNEGKSLFKKDI
 IETILPEFLDDKDEIALVNSFNGFTTAFTGFFDNRENMFSEEAKSTSIAFRCINENLTRYISN
 MDIFEKVDIFDKHEVQEIKEKILNSDYDVEDFFEGEFFNFVLTQEGIDVYNAAIGGFVTES
 GEKIKGLNEYINLYNQTKQKLPKFPLYKQVLSDRESLSFYGEGETSDEEVLEVFRNTLN
 KNSEIFSSIKKLEKLFKNFDEYSSAGIFVKNGPAISTISKDIFGEWNVIRDKWNAEYDDIHL
 KKKAVVTEKYEDDRRKSFKKIGSFSLSQLQEYADADLSVVEKLKEIIIQKVDEIYKVYGSSE
 KLFDADFVLEKSLKKNDAVVAIMKDLDVSFKSFENYIKAFFGEGKETNRDESFYGDFVLA
 YDILLKVDHIYDAIRNYVTQKPYSKDKFKLYFQNPQFMGGWDKDKE³⁰DYRATILRYGSKY
 YLAIMDKKYAKCLQKIDKDDVNGNYEKINYKLLPGPNKMLPKVFFSKWMAYYNPSEDI
 QKIYKNGTFKKGDMFNNDCHKLIDFFKDSISRYPKWSNAYDFNFSETEKYKDIAGFYRE
 VEEQGYKVSFESASKKEVDKLVEEGKLYMFQIYNKDFSDKSHGTPNLHTMYFKLLFDEN
 NHGQIRLSGGAELMRRASLKKEELVVHPANSPIANKNPDPNPKTTLSYD⁴⁰VYKDKRFSE
 DQYELHIPIAINKCPKNIFKINTEVRLVLLKHDDNPYVIGIDRGERNL⁴⁵LYIVVVDGKGNIVEQ
 YSLNEIINNFNGIRIKTDYHSLLDKEKERFEARQNWTSIENIKELKAGYISQVVHKICELV
 EKYDAVIALEDLNSGFKNSRVKEQVYQKFEKMLIDKLN⁵⁰YMVDKKSNPCATGGALKGY
 QITNKFESFKSMSTQNGFIFYIPAWLTSKIDPSTGFVNLLKTKYTSIA⁵⁵DSKKFISSFDRIMY
 VPEEDLFEFALDYKNFSRTDADYIKKWKLYSYGNRIRIFRNP⁶⁰KKNNVFDWEEVCLTSAYK
 ELFNKYGINYQQGDIRALLCEQSDKA⁶⁵FYSSFMALMSLMLQMRNSITGRTDVDFLISPVKN
 SDGIFYDSRNYEAQENAILPKNA⁷⁰DANGAYNIARKVLWAIGQFKKA⁷⁵EDEKLDKVKIAISNKE
 WLEYAQTSVKHKRPAATKKAGQAKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA

20

(配列番号 3 7)

【 0 1 6 5 】

配列番号 3 7 は、核局在化シグナル (K R P A A T K K A G Q A K K K) (配列番号
1 2) 、その次にグリシン - セリンリンカー (G S) 、その次に H A タグ (Y P Y D V P)

50

D Y A Y P Y D V P D Y A Y P Y D V P D Y A) (配列番号 13) を含む。

【 0 1 6 6 】

配列番号 37 は以下のヌクレオチド配列によってコードされ得る。

【 0 1 6 7 】

ATGAGCAAGCTGGAGAAGTTACAAACTGCTACTCCCTGTCTAAGACCCCTGAGGTTCAA
 GCCATCCCTGTGGCAAGACCCAGGAGAACATCGACAATAAGCGGCTGCTGGTGGAGG
 ACGAGAAGAGAGCCGAGGATTATAAGGGCGTAAGAAGCTGCTGGATCGCTACTATCTG
 TCTTTATCAACGACGTGCTGCACACCATAAGCTGAAGAATCTGAACAATTACATCAG
 CCTGTTCCGGAAGAAAACCAGAACCGAGAAGGAGAATAAGGAGCTGGAGAACCTGGAG
 ATCAATCTCGGAAAGGAGATGCCAAGGCCTCAAGGGCAACGAGGGCTACAAGTCCT
 GTTTAAGAAGGATATCATCGAGACAATCTGCCAGAGTTCTGGACGATAAGGACGAGA 10
 TCGCCCTGGTGAACAGCTCAATGGCTTACACAGCCTCACCGGCTTCTTGATAAC
 AGAGAGAATATGTTTCCGAGGAGGCCAACAGACATCCATGCCCTCAGGTGTATCAA
 CGAGAATCTGACCCGCTACATCTAATATGGACATCTCGAGAAGGTGGACGCCATCT
 TTGATAAGCAGCAGGTGCAGGAGATCAAGGAGAACGATCCTGAACAGCGACTATGATGT
 GAGGATTCTTGAGGGCGAGTTCTTAACTTGTGCTGACACAGGAGGGCATGACGT
 GTATAACGCCATCATCGGCGCTTCGTGACCGAGAGCGGGAGAACAGATCAAGGGCTGA
 ACGAGTACATCAACCTGTATAATCAGAAAACCAAGCAGAACGCTGCCATAAGTTAACCA
 CTGTATAAGCAGGTGCTGAGCGATCGGAGTCTCTGAGCTTCTACGGCGAGGGCTATAC
 ATCCGATGAGGAGGTGCTGGAGGTGTTAGAAACACCCCTGAACAAGAACAGCGAGATCT
 TCAGCTCCATCAAGAACGCTGGAGAACGCTGTTCAAGAATTTCGAGACTCTAGGCC 20
 GGCATCTTGTGAAGAACGGCCCCGCCATCAGCACAATCTCCAAGGATATCTCGGCGA
 GTGGAACGTGATCCGGACAAGTGAATGCCAGTATGACGATATCCACCTGAAGAAGA
 AGGCCGTGGTGACCGAGAACGATACGAGGACGATGGAGAACAGTCCTCAAGAACGATGGC
 TCCTTTCTGGAGCAGCTGCAGGAGTACGCCAGCAGCGATCTGTCTGTGGAGAA
 GCTGAAGGAGATCATCATCCAGAACGGTGGATGAGATCTAACGGTGTATGGCTCCTCTG
 AGAAGCTGTTGACGCCGATTTGTGCTGGAGAACAGCCTGAAGAACGACGCCGTG
 GTGCCATCATGAAGGACCTGCTGGATTCTGTGAAGAGAGCTCGAGAACATTACATCAAGGC
 CTTCTTGGCGAGGGCAAGGAGAACACAGGGACGAGTCCTCTATGGCGATTGTGCT
 TGGCCTACGACATCCTGCTGAAGGTGGACCATCTACGATGCCATCCGCAATTATGTG
 ACCCAGAACGCCACTCTAACGGATAAGGAGAACAGACTATCGGCCACCATCTGAGAACG 30
 GGCCTGGACAAGGATAAGGAGAACAGACTATCGGCCACCATCTGAGAACGATCGCT
 CCAAGTACTATCTGGCCATCATGGATAAGAACGATACGCCAAGTGCCTGCAGAACGATCGAC
 AAGGACGATGTGAACGGCAATTACGAGAACGATCAACTATAAGCTGCTGCCGGCCCTAA
 TAAGATGCTGCCAAAGGTGTTCTTAAGAACGGTGGATGGCTACTATAACCCCAGCG
 AGGACATCCAGAACGATCTACAAGAACGGCACATTCAAGAACGGCGATATGTTAACCTG
 AATGACTGTCACAAGCTGATCGACTTCTTAAGGATAGCATCTCCGGTATCCAAAGTG
 GTCCAATGCCTACGATTTCAACTTTCTGAGAACAGAGAACGATATAAGGACATGCCGGCT
 TTTACAGAGAGGTGGAGGAGCAGGGCTATAAGGTGAGCTTCGAGTCTGCCAGCAAGAAC
 GAGGTGGATAAGCTGGTGGAGGAGGGCAAGCTGTATATGTTCCAGATCTATAACAAAGGA
 CTTTCCGATAAGTCTCACGGCACACCCAAATCTGCACACCATGTACTTCAGCTGCTGT 40
 TTGACGAGAACAAATCACGGACAGATCAGGCTGAGCGGAGGAGCAGAGCTGTTCATGAGG
 CGCGCCTCCCTGAAGAACGGAGGAGCTGGTGGTGCACCCAGCCAACCTCCCTATGCCAA
 CAAGAACGATAATCCAAGAAAACCACAAACCCCTGTCCATCGACGTGTATAAGGATA
 AGAGGTTTCTGAGGACCAAGTACGAGCTGCACATCCAATGCCATCAATAAGTGCC
 AAGAACATCTCAAGATCAATAACAGAGGTGCGCGTGTGCTGAAGCACGACGATAACCC
 CTATGTGATGGCATCGATAGGGCGAGCGCAATCTGCTGTATATCGTGGTGGAGC
 GCAAGGGCAACATCGTGGAGCAGTATTCCCTGAACGAGATCATCAACAAACTCAACGGC
 ATCAGGATCAAGACAGATTACCACTCTGCTGGACAAGAACGGAGAACGGAGAGGTTCGA
 GGCCCGCCAGAACCTGGACCTCCATCGAGAATATCAAGGAGCTGAAGGCCGGCTATATCT 50

CTCAGGTGGTGCACAAGATCTGCGAGCTGGTGGAGAAGTACGATGCCGTATGCCCTG
 GAGGACCTGAACTCTGGCTTAAGAACAGCCGCGTGAAGGTGGAGAAGCAGGTGTATCA
 GAAGTTGAGAAGATGCTGATCGATAAGCTGAACTACATGGTGGACAAGAAGTCTAATC
 CTTGTGCAACAGGCAGCGCCCTGAAGGGCTATCAGATCACCAATAAGTTGAGAGCTT
 AAGTCCATGTCTACCCAGAACGGCTTCATCTTTACATCCCTGCCTGGCTGACATCCAA
 GATCGATCCATCTACCGGCTTGATGAACTGCTGAAAACCAAGTATACCAGCATGCCG
 ATTCCAAGAAGTTCATCAGCTCCTTGACAGGATCATGTACGTGCCGAGGAGGATCTG
 TTGAGTTGCCCTGGACTATAAGAACCTCTCGCACAGACGCCATTACATCAAGAA
 GTGGAAGCTGTACTCCTACGGCAACCGGATCAGAATCTTCCGAATCCTAAGAAGAAC
 ACGTGTTCGACTGGAGGAGGTGTGCCTGACCAGCGCCTATAAGGAGCTGTTCAACAAG
 TACGGCATCAATTATCAGCAGGGCGATATCAGAGCCCTGCTGTGCGAGCAGTCCGACAA
 GCCCTCTACTCTAGTTATGGCCCTGATGAGCCTGATGCTGCAGATGCGGAACAGCA
 TCACAGGCCGACCGACGTGGATTCTGATCAGCCCTGTGAAGAACTCCGACGGCATC
 TTCTACGATAGCCGAACTATGAGGCCAGGAGAATGCCATCCTGCCAAAGAACGCCA
 CGCCAATGGCGCTATAACATGCCAGAAAGGTGCTGTGGCCATGGCCAGTTCAAGA
 AGGCCGAGGACGAGAACGCTGGATAAGGTGAAGATGCCATCTAACAAGGAGTGGCTG
 GAGTACGCCAGACCAGCGTGAAGCACAAAAGGCCGGCCACGAAAAAGGCCGCCA
 GGCAAAAAAGAAAAAGGGATCCTACCCATACGATGTTCCAGATTACGCTTATCCCTACG
 ACGTGCCTGATTATGCATACCATATGATGTCGGACTATGCCTAA

(配列番号 38)

10

【 0 1 6 8 】

N L S 及び H A タグを含む *P o r p h y r o m o n a s c r e v i o r i c a n i s*
 (P c C p f 1 ; p Y 0 1 7) :

20

【 0 1 6 9 】

MDSLKDFTNLYPVSKTLRFELKPVGKTLENIEKAGILKEDEHRAESYRRVKIIDTYHKVF
 IDSSLENMAKMGIENEIKAMLQSFCELYKKDHRTEGEDKALDKIRAVLRLGLIVGAFTGVC
 GRRENTVQNEKYESLFKEKLIKEILPDFVLSTEAEMLPFSVEEATRSLKEFDSFTSYFAGFY
 ENRKNIYSTKPQSTAIAYRLIHENLPKFIDNILVFQKIKEPIAKELHIRADFSAGGYIKKDE
 RLEDIFSLNYYIHVLSQAGIEKYNALIGKIVTEGDGEMKGLNEHINLYNQQRGREDRLPLF
 RPLYKQILSDREQLSYLPESFEKDEELLRALKEFYDHIAEDILGRTQQLMTSISEYDLSRIY
 VRNDSQLTDISKMLGDWNNAIMARERAYDHEQAPKRITAKYERDRIKALKGEESISLAN
 LNSCIAFLDNVRDCRVDTYLSTLGQKEGPHGLSNLVENVFASYHEAEQLLSFPYPEENNLI
 QDKDNVVLIKNLLNISDLQRFLKPLWGMGDEPDKDERFYGEYNYIRGALDQVIPLYNK
 VRNYLTRKPYSTRKVKLNFNGNSQLLSGWDRNKEKDNSCVILRGQNFYLAIMNNRHKRS
 FENKMLPEYKEGEPYFEKMDYKFLPDPNKMLPKVFLSKKGIEIYKPSPKLLEQYGHGTHK
 KGDTFSMDDLHELIDFFKHSIEAHEDWKQFGFKFSDTATYENVSSFYREVEDQGYKLSFR
 KVSESYVYSLIDQGKLYLFQIYNKDFSPCSKGTPNLHTLYWRMLFDERNLADVIYKLDGK
 AEIFFREKSLKNDHPTHPAGKPIKKSRQKKGEESLFEYDLVKDRRYTMDKFQFHVPITM
 NFKCSAGSKVNDMVNAHIREAKDMHVIGIDRGERNLLYICVIDSRGTILDQISLNTINDID
 YHDLLESRDKDRQQEHRNWQTIEGIKELKQGYLSQAVHRIAELMVAYKAVVALEDLNMG
 FKGRGRQKVESSVYQQFEKQLIDKLNLYLVDKKRPEDIGGLRAYQFTAPFKSFKEMGKQN
 GFLFYIPAWNTSNIDPTTGFVNLFHVQYENVDKAKSFFQKFDISYNPKKDWFEFADFYK
 NFTKKAEGSRSMWILCTHGSRIKNFRNSQKNGQWDSEEFALTEAFKSLFVRYEIDYTADL
 KTAIVDEKQKDFFDV DLLKLFKLTVQMRNSWKEKLDYLISPVAGADGRFFDTREGNKS LP
 KDADANGAYNIALKGLWALRQIRQTSEGGKLKLAISNKEWLQFVQERSYEKDKRPAATK
 KAGQAKKKGSYPYDVPDYAYPYDVPDYAYPYDVPDYA

(配列番号 39)

30

【 0 1 7 0 】

配列番号 39 は、核局在化シグナル (K R P A A T K K A G Q A K K K) (配列番号
 12) 、その次にグリシン - セリンリンカー (G S) 、その次に H A タグ (Y P Y D V P

40

50

D Y A Y P Y D V P D Y A Y P Y D V P D Y A) (配列番号 13) を含む。

【 0 1 7 1 】

配列番号 39 は以下のヌクレオチド配列によってコードされ得る。

【 0 1 7 2 】

ATGGACAGCCTGAAGGATTCACCAACCTGTACCCGTGTCCAAGACACTGCGGTTGA
 GCTGAAGCCTGTGGCAAGACCCCTGGAGAATATCGAGAAGGCCGGCATCCTGAAGGAGG
 ATGAGCACAGAGCCGAGAGCTACCGGAGAGTGAAGAAGATCATCGATAACATATCACAAG
 GTGTTCATCGACAGCTCCCTGGAGAACATGCCAAGATGGGCATCGAGAACATGAGATCAA
 GCCATGCTGCAGTCCTTGCGAGCTGTATAAGAAGGACCACAGGACCGAGGGAGAGG
 ACAAGGCCCTGGATAAGATCAGGCCGTGCTGAGGGCCTGATCGTGGAGCCTTCACC 10
 GGC GTG CGG CCG GAG AAC ACAG TGC AGA AT GAG AAG TAT GAG AGC CT GT TAA
 GGAGAAGCTGATCAAGGAGATCCTGCCAGATTCTGTGCTGTACAGAGGCCGAGTCCC
 TGCCCTTTCTGTGGAGGAGGCCACCAGAACGCTGAAGGAGTCGACTCCTTACATCT
 TACTTCGCCGGCTTTATGAGAACCGGAAGAACATCTACTCTACCAAGCCCCAGAGCAC
 AGCCATCGCCTATAGACTGATCCACGAGAACCTGCCAACGCTGATAATATCCTGG
 TGTTCAAGATCAAGGAGCCAATGCCAAGGAGCTGGAGCACATCAGGGCAGACTTC
 AGCGCCGGCGGCTACATCAAGAAGGATGAGCGCCTGGAGGACATCTTCCCTGAAC 20
 TA CTATATCCACGTGCTGTCTCAGGCCGGCATCGAGAACGTCGACTGCCCTGATCGGCAAGA
 TCGTGACCGAGGGCGATGGCGAGATGAAGGGCCTGAACGAGCACATCACCTGTATAAT
 CAGCAGAGGGCCCGAGGACCGCTGCCACTGTTCAGACCCCTGTATAAGCAGATCCT
 GTCTGATAGGGAGCAGCTGTCCTATCTGCCAGAGTCTTCGAGAACGGACGAGGAGCTGC
 TGAGGGCCCTGAAGGAGTTTACGATCACATCGCAGAGGACATCCTGGAGAGGACCCAG
 CAGCTGATGACAAGCATCTCCGAGTACGATCTGTCCCGATCTATGTGAGAACGATAG 30
 CCAGCTGACCGACATCTCAAGAACGATGCTGGCGATTGGAATGCCATCTACATGGCC
 GGGAGAGAGCCTATGACCACGAGCAGGCCCAAGCGCATCACAGCCAAGTACGAGAGG
 GACCGCATCAAGGCCCTGAAGGGCGAGGAGTCTATCAGCCTGGCCAACCTGAACAGCTG
 CATCGCCTTCCTGGACAACGTGAGGGATTGTCGCGTGGACACCTATCTGTCTACACTGG
 GACAGAAGGGAGGGACCTCACGCCCTGAGCAACCTGGTGGAGAACGTGTTGCCCTCAC
 CACGAGGCCGAGCAGCTGCTGTCTTCCCTATCCTGAGGAGAACATCTGATCCAGGA
 CAAGGATAACGTGGTGTGATCAAGAACCTGCTGGATAATATCAGCGACCTGCAGAGGT
 TCCTGAAGCCACTGTGGGCATGGCGATGAGCCGACAAGGATGAGAGGTTTACGGC
 GAGTACAATTATCAGGGCGCCCTGGACCAGGTCACTCCCTGTATAACAAGGTGCG 40
 GAATTATCTGACCGCAAGCCATACTCCACACGCAAGGTGAAGCTGAACCTCGGCAATA
 GCCAGCTGCTGCCGGCTGGATAGGAACAAGGAGAACGGACAATTCTTGCCTGATCCTG
 CGCAAGGGCCAGAACCTCTACCTGCCATCATGAACAATCGGCACAAGCGGAGCTCGA
 GAATAAGATGCTGCCGAGTATAAGGAGGGCGAGCCTTACTCGAGAACGATGGATTATA
 AGTTTCTGCCAGACCCCAACAAGATGCTGCCAAGGTGTTCTGTCTAAGAACGGCATT
 GAGATCTACAAGCCTAGCCAAAGCTGCTGGAGCAGTATGCCACGGCACCCACAAGAA
 GGGCGATACCTTCAGCATGGACGATCTGCACGAGCTGACTTCTTAAGCACTCCA
 TCGAGGCCACGAGGATTGGAAGCAGTCGGCTTTAAGTTCAAGCGACACCGCCACATAC
 GAGAACGTGAGCAGCTTCTACCGGGAGGTGGAGGACCAAGGGCTACAAGCTGTCTTAG
 AAAGGTGTCGAGTCTTACGTGTAGCCTGATCGATCAGGGCAAGCTGTACCTGTTCC
 AGATCTATAACAAGGACTTTAGCCCTGTTCCAAGGGACCCCAAATCTGCACACACTG
 TACTGGCGGATGCTGTTCGATGAGAGAACCTGGCCACGTGATCTATAAGCTGGATGG
 CAAGGCCGAGATCTTCTTGGAGAACGCTGAGAACGACATGGTGAATGCCAC
 CAGGCCAGCCATCAAGAACGAGGCCAGAACGAGGAGGAGGCCCTGTTCGA 50
 GTACGATCTGGTGAAGGACCGGAGATATACCATGGATAAGTTCAAGTCCACGTGCCAA
 TCACAATGAACCTTAAGTGCTCTGCCGGCAGCAAGGTGAACGACATGGTGAATGCCAC
 ATCAGGGAGGCCAAGGACATGCACGTGATCGCAGTACGATAGGGCGAGCGCAATCTGCT
 GTATATCTGCGTGTGACAGCCGGCACCACATCCTGGATCAGATCTCCCTGAACACAA

TCAATGACATCGATTATCACGATCTGCTGGAGTCCAGGGACAAGGATGCCAGCAGGAG
 CACAGGAACTGGCAGACCATCGAGGGCATCAAGGAGCTGAAGCAGGGCTACCTGTCTCA
 GGCGTGCACCGCATGCCAGCTGATGGTGGCCTATAAGGCCGTGGTGCCCTGGAGG
 ACCTGAACATGGGCTTAAGCAGGGCAGACAGAAGGTGGAGAGCAGCGTGTACCAGCAG
 TTTGAGAAGCAGCTGATCGACAAGCTGAATTATCTGGTGGATAAGAAGAACGGCCCAG
 GGACATCGGAGGCCTGCTGAGAGCCTACCAGTTACCGCCCCCTTCAGAGCTTTAAGG
 AGATGGGCAAGCAGAACGGCTTCTGTTCTATATCCCTGCCTGGAACACATCCAATATC
 GACCCAAACACAGGCTTCGTGAACCTGTTCACGTGCAGTACGAGAATGTGGATAAGGC
 CAAGAGCTTCTTCAGAAGTTGACAGCATCTCCTACAACCTAAGAAGGATTGGTTG
 AGTCGCCTTGACTATAAGAACCTCACCAAGAAGGCCAGGGCTTAGGAGCATGTGG 10
 ATTCTGTGCACCCACGGCTCCGGATCAAGAACCTCAGAAATTCTCAGAAGAATGCCA
 GTGGGATAGCGAGGAGTTGCCCTGACCGAGGCCTCAAGTCCCTGTTGTGCGGTACG
 AGATCGATTATACCGCCGACCTGAAAACCGCCATCGTGGACGAGAACAGAAGGATTTC
 TTTGTGGACCTGCTGAAGCTGTTCAAGCTGACCGTGCAGATGAGAAACTCCTGGAAGGA
 GAAGGACCTGGATTACCTGATCTCTCAGTGGCCGGCGCGATGGCAGGTTCTTGACA
 CACCGCAGGGCAATAAGAGCCTGCCAAGGACGCAGATGCAAACGGAGCCTATAATATC
 GCCCTGAAGGGCTGTGGCACTGAGGCAGATCAGACAGACCTCCGAGGGCGGCAAGCT
 GAAGCTGCCATCTCTAACAAGGAGTGGCTGCAGTTGTGCAGGAGAGATCCTACGAGA
 AGGACAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAGAAAAAGGGATC
 CTACCCATACGATGTTCCAGATTACGCTTACGACGTGCCTGATTATGCATACC 20
 CATATGATGTCCCCGACTATGCCTAA

(配列番号 40)

【0173】

N L S 及び H A タグを含む P r e v o t e l l a d i s i e n s (P d C p f 1 ; p
 Y 0 1 8) :

【0174】

MENYQEFTNLFQLNKT LR FELKPIGKTCELL E EGKIFASGSFLEKDKVRADNV SYVKKEID
 KKHKIFIEETLSSFSISNDLLQYFDCYNELKAFKKDCKSDEEVKKTALRNKCTS IQRAM
 REAISQAFLKSPQKKLLAIKNL IENVFKADENVQHFSEFTSYFSGFETNRENFSDEEKSTS
 IAYRLVHDNLPIFIKNIYIFEKLKEQFDAKLSEIFENYKLYVAGSSLDEVFSLEYFNNTLT 30
 QKGIDNYNAVIGKIVKEDKQEIQGLNEHINLYNQKHDKRRLPFFISLKKQILSDREALSWL
 PDMFKNDSEVIKALKGFYIEDGFENNVLTPLATLSSLDKY NLNGIFIRNNEALSSLSQNV
 YRNFSIDEAIDANAEQLTFNNYELIANALRAKIKKETKQGRKSFEKYEEYIDKKVKAIDSLS
 IQEINELVENYVSEFNSNSGNMPRKVEDYFSLMRKGDFGSNDLIENIKTKLSAAEKL LGTK
 YQETAKDIFKKDENS KLIKELL DATKQFQHFIKPLLGTGEAADRDLVFYGDFLPLYEKFEE
 LTLLYNKVRNRLTQKPYSKDKIRLCFNKP KLM TGWVDSKTEKSDNGTQYGGYLFRKKNE
 IGEYDYFLGISSKAQLFRKNEAVIGDYERLDYYQP KANTIYGSAYEGENSYKEDKKRLNKV
 IIAYIEQIKQTNIKSIIESISKYPNISDDD KVTPSSLK EKIKKVSIDS YNGILSFKSFQS VNK
 VIDNLLKTISPLKNKAELFLD LINKDYQIFTEVQAVIDEICKQKTFIYFPISNVELEKEMGDK
 DKPLCLFQISNKDLSFAKTF SANLRKKRGAENLHTMLFKALMEGNQDNLDLGSGAIFYRA 40
 KSLDGKPTHPANEAIKCRNVANKDKVSLFTYDIYKNRRY MENKFLFHLSIVQNYKAAN
 DSAQLN SATEYIRKADDLHIIGIDRGERNLLYY SVIDMKG NIVEQDSLNIIRNN DLET DY
 HDLLDKREKERKANRQNWEAVEGIKDLKKGYLSQAVHQIAQLMLKYNAIIALEDLGQMF
 VTRGQKIEKAVYQQFEKSLVDKLSYLV DKKRPYNELGGILKAYQLASSITKNNSDKQNGF
 LFYVP AWNTSKIDPV TGFT DLLRPKAMTIKEAQDFFGAFDNISYNDKGYFEFETNYDKFK
 IRMKSAQTRWTICTFGNRIKRKKDKNYWN YEEVELTEEFKLFKDSNIDYENCNLKEEIQ
 NKDNRKFFDDLIKLLQLTLQMRNSDDKGNDYIISP VANAEGQFFDSRNGDKKLPLDADAN
 GAYNIARKGLWNIRQIKQT KNDKKLNLSISSTE WLDVREKPYLK KRP AATKKAGQAKKK
 KGSYPYDV PDY AYPYDV PDY AYPYDV PDY A

(配列番号 41)

50

【0175】

配列番号41は、核局在化シグナル(K R P A A T K K A G Q A K K K K)（配列番号12）、その次にグリシン-セリンリンカー(G S)、その次にHAタグ(Y P Y D V P D Y A Y P Y D V P D Y A)（配列番号13）を含む。

【0176】

配列番号41は以下のヌクレオチド配列によってコードされ得る。

【0177】

ATGGAGAACTATCAGGAGTTACCCAACCTGTTTCAAGCTGAATAAGACACTGAGATTCA
GCTGAAGCCCATCGGCAAGACCTGCGAGCTGCTGGAGGGAGGGCAAGATCTCGCCAGCG
GCTCCTTCTGGAGAAGGACAAGGTGAGGGCCGATAACGTGAGCTACGTGAAGAAGGAG 10
ATCGACAAGAACGACAAGATCTTATCGAGGGAGACACTGAGCTCCTCTATCAGCAA
CGATCTGCTGAAGCAGTACTTGACTGCTATAATGAGCTGAAGGCCTCAAGAAGGACT
GTAAGAGCGATGAGGAGGGAGGTGAAGAAAACCGCCCTGCGCAACAAGTGTACCTCCATC
CAGAGGGCCATGCGCGAGGCCATCTCTCAGGCCTTCTGAAGAGAGCCCCAGAAGAAGCT
GCTGGCCATCAAGAACCTGATCGAGAACGTGTTCAAGGCCGACGAGAATGTGCAGCACT
TCTCCGAGTTACCAGCTATTCTCCGGCTTGAGACAAACAGAGAGAATTCTACTCT 20
GACGAGGAGAAGTCCACATCTATCGCCTATAAGGCTGGTCACGATAACCTGCCTATCTT
CATCAAGAACATCTACATCTTCGAGAACGCTGAAGGAGCAGTTGACGCCAAGACCCTGA
GCGAGATCTCGAGAACTACAAGCTGTATGTGGCCGGCTAGCCTGGATGAGGTGTTCT
TCCCTGGAGTACTTTAACAAATACCCCTGACACAGAAGGGCATCGACAACATAATGCCGT
GATCGGCAAGATCGTGAAGGAGGATAAGCAGGGAGATCCAGGGCCTGAACGAGCACATCA 30
ACCTGTATAATCAGAACGACAAGGACCGGAGACTGCCCTCTTATCTCCCTGAAGAAC
CAGATCCTGTCCGATCGGGAGGCCCTGTCCTGGCTGCCATGTTCAAGAACATGATTCT
TGAAGTGATCAAGGCCCTGAAGGGCTTCTACATCGAGGACGGCTTGAGAACAAATGTGC
TGACACCTCTGGCCACCCCTGCTGTCCTCTGGATAAGTACAACCTGAATGGCATCTT
ATCCGCAACAATGAGGCCCTGAGCTCCCTGTCCTGGAGAACACTGTTACGGAAATTCTAT
CGACGAGGCCATCGATGCCAACGCCAGCTGCAGACCTCAACAATTACGAGCTGATCG
CCAATGCCCTGCGGCCAAGATCAAGAACGGAGAACAAAGCAGGGCCGGAAAGTCTTCGAG
AAGTACGAGGAGTATATCGATAAGAACGGTGAAGGCCATCGACAGCCTGTCATCCAGGA
GATCAACGAGCTGGTGGAGAATTACGTGAGCGAGTTAACTCTAATAGCGGCAACATGC
CAAGAACGGTGGAGGACTACTTCAGCCTGATGAGGAAGGGCAGTCGGCTCCAAAGAT 40
CTGATCGAAAATATCAAGACCAAGCTGAGCGCCGAGAACAGCTGACAGCTGATCAAGGAGC
TGCTGGACGCCACCAAGCAGTTCCAGCAGCTTATCAAGCCACTGCTGGGCACAGGCGAG
GAGGCAGATCGGGACCTGGTGTACGGCGATTTCTGCCCTGTATGAGAACAGTGGATTCA
GGAGCTGACCCCTGCTGTATAACAAGGTGCGGAATAGACTGACACAGAACGGCTATTCCA
AGGACAAGATCCGCTGTGCTTCAACAAGCTAACGCTAGCTGATGACAGGCTGGTGGATTCC
AAGACCGAGAACAGTCTGACAAACGGCACACAGTACGGCGCTATCTGTTGGAAAGAAC
TGAGATCGCGAGTACGATTATTTCTGGCATCTCTAGCAAGGCCAGCTGTTCAAGAAC
AGAACGAGGCCGTGATCGCGACTACGAGAGGCTGGATTACTATCAGCAAAGGCCAAT
ACCATCTACGGCTCTGCCATGAGGGCGAGAACAGCTACAAGGAGGACAAGAACGGCT
GAACAAAGTGATCATGCCATATCGAGCAGATCAAGCAGAACACATCAAGAACAGTCTA
TCATCGAGTCCATCTCTAAGTATCCTAATATCAGCGACGATGACAAGGTGACCCCATCC
TCTCTGCTGGAGAACAGATCAAGAACGGTGTCTATCGACAGCTACAACGGCATCCTGTCCTT
CAAGTCTTTCAGAGCGTGAACAAGGAAGTGTACGATAACCTGCTGAAACCATCAGCC
CCCTGAAGAACAGGCCAGTTCTGGACCTGATCAATAAGGATTATCAGATCTTCACC
GAGGTGCAGGCCGTGATCGACGAGATCTGCAAGCAGAACACCTTCATCTACTTCCAAT
CTCCAACGTGGAGCTGGAGAACGGAGATGGCGATAAGGACAAGGCCCTGTGCCTGTTCC
AGATCAGCAATAAGGATCTGTCCTCGCCAAGACCTTAGCGCCAACCTGCGGAAGAAC
AGAGGCCGAGAACATCTGCACACAATGCTTTAAGGCCCTGATGGAGGGCAACCAGGA 50

TAATCTGGACCTGGCTCTGGGCCATCTTCTACAGAGCCAAGAGCCTGGACGGCAACA
 AGCCCCACACACCCCTGCCAATGAGGCCATCAAGTGTAGGAACGTGGCCAATAAGGATAAG
 GTGTCCCTGTTCACCTACGACATCTATAAGAACAGGGCGCTACATGGAGAATAAGTCCT
 GTTTCACCTGAGCATCGTCAGAACTATAAGGCCAATGACTCCGCCAGCTGAACA
 GCTCCGCCACCGAGTATATCAGAAAGGCCATGACCTGCACATCATCGGCATCGATAGG
 GGCAGCGCAATCTGCTACTATTCCGTATCGATATGAAGGGCAACATCGTGGAGCA
 GGACTCTCTGAATATCATCAGAACATGACCTGGAGACAGATTACCACGACCTGCTGG
 ATAAGAGGGAGAAGGAGCGCAAGGCCAACCGCAGAATTGGGAGGCCGTGGAGGCAT
 CAAGGACCTGAAGAACGGCTACCTGAGCCAGGCCGTGCACCAGATGCCAGCTGATGC
 TGAAGTATAACGCCATCATGCCCTGGAGGATCTGGCCAGATGTTGTGACCCGCC
 CAGAAGATCGAGAACGGCGTGTACCAAGCAGTTGAGAACAGGCCTGGATAAGCTGTC
 CTACCTGGTGGACAAGAACGCCCTATAATGAGCTGGCCGGCATCCTGAAGGCC
 AGCTGGCCTCTAGCATACCAAGAACATTCTGACAAGCAGAACGCCCTCGTTTAT
 GTGCCAGCCTGGAATAACAAGCAAGATCGATCCCCTGACCGGCTTACAGACCTGCTGCG
 GCCCAAGGCCATGACCATCAAGGAGGCCAGGACTTCTTGCGCCTCGATAACATCT
 CTTACAATGACAAGGGCTATTCGAGTTGAGACAAACTACGACAAGTTAAGATCAGA
 ATGAAGAGGCCAGACCAGGTGGACAATCTGCACCTCGCAATCGGATCAAGAGAAA
 GAAGGATAAGAACTACTGGAATTATGAGGAGGTGGAGCTGACCGAGGAGTTCAAGAAC
 TGTTAAGGACAGAACATCGATTACGAGAACTGTAATCTGAAGGAGGAGATCCAGAAC
 AAGGACAATCGCAAGTTCTTGATGACCTGATCAAGCTGCTGCAGCTGACACTGCAGAT
 CGGAACTCCGATGACAAGGGCAATGATTATATCATCTCTCTGTGGCCAACGCC
 GCCAGTTCTTGACTCCCCTGCAATGGCGATAAGAACGCTGCCACTGGATGCAGACGCAA
 GGAGCCTACAATATGCCCGCAAGGGCTGTGGAACATCCGGCAGATCAAGCAGACCAA
 GAACGACAAGAAGCTGAATCTGAGCATCTCCTCTACAGAGTGGCTGGATTCGTGCGG
 AGAAGCCTTACCTGAAGAAAAGGCCGGCCACGAAAAAAGGCCGGCCAGGAAAAAA
 GAAAAAGGGATCCTACCCATACGATGTTCCAGATTACGCTTATCCCTACGACGTGCCTG
 ATTATGCATACCCATATGATGTCGGACTATGCCTAA

(配列番号 42)

【0178】

N L S 及び H A タグを含む *Porphromonas macacae* (PmCpf

1 ; pY09) :

【0179】

MKTQHFFEDFTSLYSLSKIRFELPIGKLENIKNGLIRRDEQRLDDYEKLKKVIDEYHE
 DFIANILSSFSFSEEILQSYIQNLSESEARAKIEKTMRDTLAKAFSEDERYKSIFKKELVKKD
 IPVWCPAYKSLCKFDNFITSLVPFHENRKNLYTSNEITASIPYRIVHVNLPKFQNI
 EALC ELQKKMGADLYLEMMENLRNVWPSFKTPDLCNLKTYNHLMVQSSISEYNRFVGGYS
 TEDGTKHQGINEWINIYRQRNKEMRLPGLVFLHKQILAKVDSSFISDTLENDDVFCVL
 RQFRKLFWNTVSSKEDDAASLKDLFCGLSGYDPEAIYVSDAHLATISKNIFDRWNYISDAI
 RRKTEVLMRKKESVERYAEKISKQIKKRQSYS LAELDDLLAHYSEESLPAGFSLLSYFTSL
 GGQKYLVSDGEVILYEEGSNIWDEVLIAFRDLQVILDKDFTEKKLGKDEEVSVIKKALDS
 ALRLRKFFDLSGTGAEIRRDSFYALYTDRMDKLKGLLKMYDKVRNYLTKKPYSIEKFKL
 HFDPNSLLSGWDKNKELNNLSVIFRQNGYYYLGIMTPKGKNLFKTPKLGAEE
 EYKQIAEPMLMLPKVFFPKTKPAFAPDQSVDIYNNKTFKTGQKGFNKKDLYRLIDFYK
 EALT VHEWKLFNFSPTEQYRNIGEFFDEVREQAYKVSMVNVPASYIDEAVENGKLYLF
 QIYNKDFSPYSKGIPNLHTLYWKALFSEQNQSRVYKLCGGGELFYRKASLHM
 KGKISIHKKLNKKGETSLFNYDLVKDKRFTEDKFFFHVPISINYKNKITNVNQMVRDYIA
 QNDDLQIIGIDRGERNLLYISRIDTRGNLLEQFSLNVI
 ESDKGDLRTDYQKILGDREQERLR
 RRQEWSIESIKDLKGYMSQVVHKICNMVVEHKAI
 VVLENLNLSFMKGRKKVEKSVYE
 KFERMLVDKLNLYVVDKKNLSNEPGGLYAAYQLTNPLFSFEELHRYPQSGILFFVDPWN
 TSLTDPSTGFVNLLGRINYTNVG
 DARKFFDRFN
 AIRYDGKGNILFDLDSRF
 DV
 VETQRK

10

20

30

40

50

LWTLTTFGSRIAKSKKGKWMVERIENSLCFLFELFEQFNIGYRVEKDLKKAILSQDRKEF
 YVRLIYLFLNLMQMIRNSDGEEDYILSPALNEKNLQFDTRLIEAKDLPVDADANGAYNVAR
 KGLMVVQRIKRGDHESIHRIGRAQWLRYVQEGIVEKRPAATKKAGQAKKKGSYPYDVP
 DYAYPYDVPDYAYPYDVPDYA

(配列番号 43)

【0180】

配列番号 43 は、核局在化シグナル (K R P A A T K K A G Q A K K K K) (配列番号 12)、その次にグリシン - セリンリンカー (G S)、その次に HA タグ (Y P Y D V P D Y A Y P Y D V P D Y A) (配列番号 13) を含む。

【0181】

配列番号 43 は以下のヌクレオチド配列によってコードされ得る。

【0182】

ATGAAAACCCAGCACTTCTTGAGGACTTCACAAGCCTGTACTCTCTGAGCAAGACCAT
 CCGGTTTGAGCTGAAGCCAATCGGCAAGACCCCTGGAGAACATCAAGAAGAACATGGCCTGA
 TCCGGAGAGATGAGCAGAGACTGGACGATTACGAGAACGCTGAAGAAAGTGATCGACGAG
 TATCACGAGGATTTCATGCCAACATCCTGAGCTCCTTCTCTGAGGAGATCCT
 GCAGTCCTACATCCAGAACATCTGAGCGAGTCCGAGGCCAGGGCCAAGATCGAGAAAACCA
 TGCAGCACACACTGGCCAAGGCCTCTGAGGATGAGAGGATACAAGAGCATCTTAAG
 AAGGAGCTGGTGAAGAAGGACATCCCCGTGTGGTGCCTGCCTATAAGAGCCTGTGCAA
 GAAGTTCGATAACTTACACATCTGGTGCCCTCCACGAGAACAGGAAGAACCTGT 10
 ATACCAAGCAATGAGATCACAGCCTCTATCCCTATCGATCGTCACGTGAACCTGCCA
 AAGTTTATCCAGAACATATCGAGGCCCTGTGCGAGCTGCAGAACAGAACATGGCGCCGACCT
 GTACCTGGAGATGATGGAGAACCTGCGAACGTGTGGCCAGCTCGTGAAGAACCCCCAG
 ACGACCTGTGCAACCTGAAAACCTATAATCACCTGATGGTGCAGTCAGCATCAGCGAG
 TACAACAGGTTGTGGCGGCTATTCCACCGAGGACGGCACAAAGCACCAGGGCATCAA
 CGAGTGGATCAATATCTACAGACAGAGGAATAAGGAGATGCGCCTGCCTGGCCTGGT
 TCCTGCACAAGCAGATCCTGGCCAAGGTGGACTCCTCTAGCTTACAGCGATAACTG
 GAGAACGACGATCAGGTGTTTGCCTGCTGAGACAGTCAGGAAGCTGTTTGGAAATAC
 CGTGCCTCTAAGGAGGACGATGCCGCCTCCCTGAAGGACCTGTTCTGTGGCCTGCTG
 GCTATGACCCCTGAGGCCATCTACGTGAGCGATGCCACCTGGCCACAATCTCCAAGAAC 20
 ATCTTGACAGATGGAATTACATCTCCATGCCATCAGGCGCAAGAACCGAGGGTGCATGAT
 GCCACGGAAGAAGGAGAGCGTGGAGAGATATGCCAGAACAGATCTCCAAGCAGATCAAG
 AAGAGACAGTCTTACAGCCTGGCGAGCTGGACGATCTGCTGGCCACTATAGCGAGGA
 GTCCCTGCCGCAAGGCTTCTCTGCTGAGCTACTTACATCTCTGGCGGCCAGAAC 30
 ATCTGGTGAACGGCGAACAGTGTACCTGTACGAGGAGGGCAGAACATCTGGACGAG
 GTGCTGATCGCCTTCAGGGATCTGCAGGTATCCTGGACAAGGACTTCACCGAGAACAA
 GCTGGCAAGGATGAGGAGGCCGTGTGATCAAGAACGGCCCTGGACAGCGCCCTGC
 GCCTGCAGGAAGTTCTTGATCTGCTGTCGGCACAGGCGCAGAGATCAGGAGAGACAGC
 TCCTTCTATGCCCTGTATACCGACCGATGGATAAGCTGAAGGGCTGCTGAAGATGTA
 TGATAAGGTGAGAAACTACCTGACCAAGAACGCTTATTCCATCGAGAACAGTCAAGCTGC
 ACTTGTACAACCCATCCCTGCTGTCTGGCTGGGATAAGAACATAAGGAGCTGAACAATCTG 40
 TCTGTGATCTCCGGCAGAACGGCTACTATTACCTGGCATCATGACACCCAAGGGCAA
 GAATCTGTTCAAGACCCCTGCCTAACAGCTGGCGCCAGGGAGATGTTTATGAGAACATGG
 AGTACAAGCAGATGCCGAGCCTATGCTGATGCTGCCAAAGGTGTTCTTCCAAAGAAA
 ACCAACGCCAGCCTGCCAGACCAGAGCGTGGATATCTACAACAAGAAAACCTT
 CAAGACAGGCCAGAACGGCTTTAATAAGAACGGACCTGTACCGGCTGATCGACTTCTACA
 AGGAGGCCCTGACAGTGCACGAGTGGAGCTGTTAACCTCTCCTTCTCCAACCGAG
 CAGTATCGGAATATCGCGAGTTCTTGACGAGGTGAGAGAGCAGGCCCTACAAGGTGTC
 CATGGTGAACGTGCCGCCTCTTATATCGACGAGGCCGTGGAGAACGGCAAGCTGTATC
 TGTTCCAGATCTACAATAAGGACTTCAGCCCCACTTCAAGGGCATCCCTAACCTGCAC 50

ACACTGTATTGGAAGGCCCTGTTCAGCGAGCAGAATCAGAGCCGGGTGTATAAGCTGTG
 CGGAGGAGGAGAGCTGTTTATAGAAAGGCCAGCCTGCACATGCAGGACACCACAGTGC
 ACCCCAAGGGCATCTCTATCCACAAGAACCTGAATAAGAAGGGCAGACAAAGCCTG
 TTCAACTACGACCTGGTGAAGGATAAGAGGTTACCGAGGACAAGTTCTTTCCACGT
 GCCTATCTCTATCAACTACAAGAACCTGAATAAGAAGATCACCAACGTGAATCAGATGGTGCGCG
 ATTATATCGCCCAGAACGACGATCTGCAGATCATCGGCATCGACCGCGCGAGCGGAAT
 CTGCTGTATATCAGCCGGATCGATACAAGGGCAACCTGCTGGAGCAGTCAGCCTGAA
 TGTGATCGAGTCCGACAAGGGCAGTCTGAGAACCGACTATCAGAACGATCCTGGCGATC
 GCGAGCAGGAGCGGCTGAGGCAGGGCAGGAGTGGAAAGTCTATCGAGAGCATCAAGGAC
 CTGAAGGATGGCTACATGAGCCAGGTGGTCACAAGATCTGTAACATGGTGGAGCAGCA
 CAAGGCCATCGTGGTGGAGAACCTGAATCTGAGCTTCATGAAGGGCAGGAAGAAGG
 TGGAGAAGTCCGTGTACGAGAACGTTGAGCGATGCTGGTGGACAAGCTGAACATCTG
 GTGGTGGATAAGAAGAACCTGTCCAATGAGCCAGGAGGCCTGTATGCAGCATAACCAGCT
 GACCAATCCACTGTTCTCTTGAGGAGCTGCACAGATACCCCAGAGCGGCATCCTGT
 TTTCTGGACCCATGGAACACCTCTGACAGATCCCAGCACAGGCTCGTAATCTG
 CTGGGCAGAATCAACTACACCAATGTGGCGACGCCGCAAGTTTCGATCGGTTAA
 CGCCATCAGATATGACGGCAAGGGCAATATCCTGTTGACCTGGATCTGTCAGATTG
 ATGTGAGGGTGGAGACACAGAGGAAGCTGTGGACACTGACCACATTGGCTCTGCATC
 GCCAAATCCAAGAAGTCTGGCAAGTGGATGGAGCAGGATCGAGAACCTGAGCCTGTG
 CTTCTGGAGCTGTTGAGCAGTTAATATCGGCTACAGAGTGGAGAACGACTGAAGA
 AGGCCATCCTGAGCCAGGATAGGAAGGAGTTCTATGTGCGCCTGATCTACCTGTTAAC
 CTGATGATGCAGATCCGGAACAGCGACGGCAGGAGGATTATATCCTGTCCTCCGCCCT
 GAACGAGAAGAATCTGCAGTTGACAGCAGGCTGATCGAGGCCAAGGATCTGCCTGTGG
 ACGCAGATGCAAACGGAGCATACAATGTGGCCGCAAGGGCTGATGGTGGTCAGAGA
 ATCAAGAGGGCGACCACGAGTCATCCACAGGATCGGAAGGGCACAGTGGCTGAGATA
 TGTGCAGGAGGGCATGTGGAGAAAAGGCCGGCCACGAAAAAGGCCGGCCAGGCA
 AAAAAGAAAAAGGGATCCTACCCATACGATGTTCCAGATTACGCTTATCCCTACGACGT
 GCCTGATTATGCATACCCATATGATGTCCCCGACTATGCCTAA

(配列番号 44)

【0183】

上記の非限定的な配列のいくつかは、核局在化シグナル及び／またはタグ配列（例えば H A タグなど）などの配列を含む。様々な実施形態において、異なる核局在化シグナルが存在し得る。いくつかの実施形態において、核局在化シグナルを用いない。ある実施形態において、タグ（例えば H A タグ）を用いない。

【0184】

タンパク質（例えば遺伝子編集複合体中のタンパク質など）に関する様々な実施形態において、タンパク質は核局在化シグナルを含み得る。例えば、タンパク質（例えば Cas タンパク質）は核局在化シグナル（NLS）を含み得る。そのようなシグナルは当該技術分野において公知であり、Kalderon et al., (1984) Cell 39 (3 Pt 2): 499-509; Makkerh et al., (1996) Curr Biol. 6 (8): 1025-7; 及び Dingwall et al., (1991) Trends in Biochemical Sciences 16 (12): 478-81に非限定例が記載されており、それぞれの内容を参照により本明細書に援用する。核局在化シグナルの具体的な非限定例としては、GGSGPPKKRKKV（配列番号 5）、KRPAATKKAGQAKKKK（配列番号 12）、PKKKRKV（配列番号 45）、KR[PAAATKKAGQA]KKKK（配列番号 46）、KR[XXXXXXXXXX]KKKK（配列番号 47）、KKXXK（配列番号 48）、KRXK（配列番号 49）、KKXR（配列番号 50）、KRXR（配列番号 51）、AVKRPAATKKAGQAKKKLD（配列番号 52）、MSRRRKANPTKLSENAKKLAKEVEN（配列番号 53）、PAAKRVKLD（配列番号 54）

10

20

30

40

50

、及び K L K I K R P V K (配列番号 5 5) が挙げられる。

【 0 1 8 5 】

全般的な定義及び全般的な技術

別途具体的に定義しない限り、本明細書で用いるすべての科学技術用語は、当業者（例えば細胞培養、分子遺伝学、及び生化学における）によって一般に理解されるものと同じ意味を有するとみなすものとする。

【 0 1 8 6 】

本明細書で用いる場合、数値または範囲との関連における用語「約」は、文脈上より限定された範囲に解釈されない限り、列挙または特許請求した数値または範囲の ± 10 % を意味する。

10

【 0 1 8 7 】

上記の説明及び特許請求の範囲において、「の少なくとも 1 つ」または「の 1 つ以上」などの語句が要素または特徴の一連のリストの前に現れる場合がある。用語「及び／または」が 2 つ以上の要素または特徴のリストの中に現れる場合がある。用いられる文脈と暗示的または明示的に別途矛盾しない限り、そのような語句は、リストに挙げた要素もしくは特徴のいずれかを個別に、または列挙した要素もしくは特徴のいずれかを他の列挙した要素もしくは特徴のいずれかと併せて意味することを意図する。例えば、語句「A 及び B の少なくとも 1 つ」、「A 及び B の 1 つ以上」、ならびに「A 及び／または B」はすべて、「A のみ、B のみ、または A 及び B とともに」を意味することを意図する。3 つ以上の事項を含むリストについても同様の解釈を意図する。例えば、語句「A、B、及び C の少なくとも 1 つ」、「A、B、及び C の 1 つ以上」、ならびに「A、B、及び／または C」はすべて、「A のみ、B のみ、C のみ、A 及び B とともに、A 及び C とともに、B 及び C とともに、または A 及び B 及び C とともに」を意味することを意図する。さらに、上記及び特許請求の範囲における用語「に基づく」の使用は、「に少なくとも部分的にに基づく」を意味することを意図し、そのため、列挙していない特徴または要素も許容される。

20

【 0 1 8 8 】

用語「原形質膜」及び「細胞膜」は、本明細書では同じ意味で用いられ、細胞の外側の環境から細胞の内部を隔てる半透膜のことをいう。

【 0 1 8 9 】

本明細書で用いる場合、「発現ベクター」は、1 つ以上のポリヌクレオチドの発現を引き起こすことのできる D N A または R N A ベクターである。発現ベクターは宿主細胞内で複製も可能であることが好ましい。発現ベクターは原核性または真核性のいずれかとすることができる、典型的にはプラスミドである。本発明の発現ベクターには、本明細書に記載の原核細胞または真核細胞、例えば原生動物、藻類、真菌、酵母、植物、動物、脊椎動物、無脊椎動物、節足動物、哺乳類、げっ歯類、靈長類、またはヒト細胞のうちの 1 つなどの本発明の宿主細胞において機能する（すなわち遺伝子発現を誘導する）任意のベクターが含まれる。本発明の発現ベクターは、調節配列、例えば転写制御配列、翻訳制御配列、複製起点、及び宿主細胞と適合性のあり、ポリヌクレオチドの発現を制御する他の調節配列などを含有する。特に、本発明の発現ベクターは転写制御配列を含む。転写制御配列は、転写の開始、伸長、及び終了を制御する配列である。特に重要な転写制御配列は、転写開始を制御するもの、例えばプロモーター、エンハンサー、オペレーター及びリプレッサー配列などである。好適な転写制御配列としては、本発明の細胞の少なくとも 1 つにおいて機能することのできる任意の転写制御配列が挙げられる。様々なそのような転写制御配列が当業者には公知である。好ましい実施形態において、本方法は、核酸分子またはコンストラクトを送達するためのウイルスベクター、例えばアデノウイルスなどの使用を含まない。

30

【 0 1 9 0 】

パラメータ範囲を規定する場合、その範囲内のすべての整数及びその 10 分の 1 も本発明によって規定されることが理解される。例えば、「0 . 2 ~ 5 m g 」は、5 . 0 m g 以下の 0 . 2 m g 、0 . 3 m g 、0 . 4 m g 、0 . 5 m g 、0 . 6 m g などの開示である。

40

50

【0191】

用いられる文脈と暗示的または明示的に別途矛盾しない限り、細胞の「圧搾」、「圧搾すること」、「変形」などへの言及は、最小限の細胞傷害性で細胞のサイトゾル内に巨大分子を直接送達するために用いられるプロセスのことをいう。このアプローチの根底にある原理は、標的細胞の急速な機械的変形、または圧搾による一時的な膜破碎であり、これが液状媒体中の巨大分子の拡散による取込みを可能とし、その後、細胞膜修復が生じる（例えば2014年9月25日に公開された米国特許出願公開第2014/0287509号；2015年10月30日に出願されたPCT国際特許出願第PCT/US2015/058489号；及び2015年11月13日に出願されたPCT国際特許出願第PCT/2015/060689号を参照、それぞれの全内容を参照により本明細書に援用する）。

10

【0192】

本明細書で用いる場合、「gRNA」はCRISPR-CasシステムガイドRNAのことをいう。

【0193】

本明細書で用いる場合、用語「タンパク質複合体」は、結合パートナーとのタンパク質の特異的結合から生じる複合ユニットのことをいい、ここで前記結合パートナーは、前記タンパク質複合体を形成する1つ以上のタンパク質、1つ以上の核酸、または1つ以上のタンパク質及び1つ以上の核酸の組合せなどとすることができます。タンパク質複合体はタンパク質-タンパク質複合体、タンパク質-核酸複合体などであり得る。ある実施形態において、タンパク質複合体はタンパク質-タンパク質相互作用、例えば異なるタンパク質間の相互作用、または同じタンパク質の二量体、三量体、四量体もしくはより高位のオリゴマーを含み得る。タンパク質複合体のサブユニット間の（例えば2つ以上のタンパク質を含むタンパク質-タンパク質複合体もしくはタンパク質-核酸複合体における）またはタンパク質と核酸との間の（例えばタンパク質-核酸複合体における）相互作用は、一般に非結合性相互作用、例えば水素架橋、電子系、例えば（任意により共役）C--C二重結合または芳香環、例えばフェニル、及びヘテロ芳香環、例えばピロール、イミダゾール、インドール、ピリミジンもしくはプリン環などに起因する相互作用、ならびに金属原子と酸素、窒素または硫黄原子との間の相互作用などであるが、弱い、特に可逆的な共有結合性相互作用、例えば硫黄-硫黄架橋でもあり得る。

20

【0194】

「タンパク質-タンパク質複合体」は、タンパク質間の相互作用によって形成された2つ以上のタンパク質の組合せである複合ユニットを意味する。必ずしもこれに限らないが典型的には、「タンパク質複合体」は特異的な非共有結合性親和力を介した2つ以上のタンパク質同士の結合によって形成される。しかし、相互作用しているパートナー間で共有結合も存在し得る。例えば、2つの相互作用しているパートナーは、タンパク質複合体がより安定となるように共有結合で架橋される場合もある。

30

【0195】

同様に、「タンパク質-核酸複合体」は、タンパク質と核酸との間の相互作用を含む相互作用によって形成された少なくとも1つのタンパク質及び少なくとも1つの核酸の組合せである複合ユニットを意味する。典型的には、「タンパク質-核酸複合体」は非共有結合性親和力を介したタンパク質及び核酸の結合によって形成されるが、必ずしもこれに限らない。

40

【0196】

様々な実施形態において、遺伝子編集複合体はタンパク質-核酸複合体、例えばRNPなどである。RNPの非限定例はCasタンパク質及びgRNAを含むCRISPR-Cas RNPである。

【0197】

本明細書に記載の方法及びデバイスは、完全で機能的な遺伝子編集複合体を細胞内に送達する。遺伝子編集複合体のコンポーネントは送達中に解離することなく、細胞内への送

50

達後も機能を維持する。

【0198】

完全で機能的な遺伝子編集複合体が細胞に送達されたかどうか判定するために、様々なアッセイが利用可能である。例えば、完全で機能的な遺伝子編集複合体が細胞に送達されたことを示すために、遺伝子編集複合体による遺伝子編集の検出が用いられ得る。代替としてまたは追加で、遺伝子編集複合体が送達された細胞は非変性条件（例えば非変性バッファーまたはフレンチプレスなど）を用いて溶解され、非変性ゲルを用いてライセートが分析され、遺伝子編集複合体が細胞内で完全な状態であったかどうか判定され得る。代替としてまたは追加で、細胞は非変性条件を用いて溶解され、その後、免疫沈降を用いてライセートから遺伝子編集複合体が単離され得る（すなわち、免疫沈降を用いて複合体の1つのコンポーネントを別のコンポーネントと共に単離できることを確認するため）。単離した遺伝子編集複合体は、遺伝子編集複合体が送達前 / 細胞圧搾前バッファー中に存在したかどうか、また、複合体がマイクロ流体 / 圧搾処理の後で処理した細胞内に存在し、完全及び／または機能的であるかどうか判定するために、非変性ゲルまたは変性アッセイ（例えばドデシル硫酸ナトリウム・ポリアクリルアミドゲル電気泳動など）を用いて、細胞への送達の前または後にアッセイすることができる。C R I S P R - C a s 9 R N P に関するいくつかの実施形態において、約 145、150、155、または 145 ~ 160 kDa の非変性ゲル上のバンドは、R N P が完全で機能的な遺伝子編集複合体として細胞内に送達されたことを示し得る。

10

【0199】

本明細書で用いる場合、デバイス寸法は長さ、幅、及び任意により狭窄部の数を示す一連の数値で表される（例えば 30 μm - 6 m × 5 は長さ 30 μm、幅 6 μm、及び 5 つの狭窄部のデバイスを表す）。

20

【0200】

例示的実施形態

本主題の態様は、溶液中の細胞を準備すること；細胞変形狭窄部を含むマイクロ流体チャネルに溶液を通過させること；狭窄部に細胞を通過させ、その結果、細胞に圧力が加わり、タンパク質及び核酸が透過するのに十分大きな細胞の摂動を引き起こすこと；ならびに細胞が狭窄部を通過する前、その最中、及び／またはその後に、細胞をタンパク質及び核酸と接触させることを含む、細胞内へのタンパク質及び核酸の送達方法を提供する。

30

【0201】

いくつかの実施形態において、前記溶液は、細胞が狭窄部を通過する前、その最中、及び／またはその後にタンパク質及び核酸を含む。

【0202】

いくつかの実施形態において、タンパク質及び核酸はタンパク質 - 核酸複合体を形成する。

【0203】

いくつかの実施形態において、タンパク質及び核酸はタンパク質 - 核酸複合体のコンポーネントであるが、細胞への送達時には複合体化していない。

40

【0204】

いくつかの実施形態において、タンパク質及び核酸は細胞内への送達後にタンパク質 - 核酸複合体を形成する。

【0205】

いくつかの実施形態において、タンパク質及び核酸は細胞内への送達前にタンパク質 - 核酸複合体を形成する。

【0206】

いくつかの実施形態において、タンパク質及び核酸は遺伝子編集コンポーネント含む。

【0207】

いくつかの実施形態において、前記タンパク質 - 核酸複合体はリボ核タンパク質 (R N P) を含む。

50

【0208】

いくつかの実施形態において、(a)タンパク質はCasタンパク質またはCpf1タンパク質であり；(b)核酸はシングルガイドRNA(sgRNA)またはCRISPR RNA(crrNA)及びトランスクレーティングcrrNA(tracrRNA)である。

【0209】

いくつかの実施形態において、複合体はCasタンパク質またはCpf1タンパク質及びsgRNAをふくむRNPであり、Casタンパク質またはCpf1タンパク質及びsgRNAは約0.5、2.0、2.5、または3.0モル濃度超(molar excess)のCasタンパク質またはCpf1タンパク質を用いて複合体化されている。

10

【0210】

いくつかの実施形態において、Casタンパク質はCas9タンパク質を含む。

【0211】

いくつかの実施形態において、前記タンパク質-核酸複合体は第1RNP及び第2RNPを含む。

【0212】

いくつかの実施形態において、第1RNP及び第2RNPは両方ともニッカーゼである。

【0213】

いくつかの実施形態において、第1RNPは第2RNPの標的配列とは異なる標的配列にニックを入れる。

20

【0214】

いくつかの実施形態において、前記タンパク質-核酸複合体はTALENタンパク質、ジンクフィンガースクレアーゼ、メガスクレアーゼ、またはCreリコンビナーゼを含む。

【0215】

いくつかの実施形態において、核酸はTALENタンパク質、ジンクフィンガースクレアーゼ、メガスクレアーゼ、またはCreリコンビナーゼをコードするmRNAを含む。

【0216】

いくつかの実施形態において、前記タンパク質-核酸複合体は、(a)静電引力を介してタンパク質と複合体化している核酸分子；(b)タンパク質に巻きついた核酸分子；(c)DNA及びヒストン；(d)リボ核タンパク質(RNP)；(e)リボソーム、酵素テロメラーゼ、ウォールトリボ核タンパク質、RNase P、hnRNP、もしくは核内低分子RNP(snRNP)；または(f)タンパク質を含む染色体を含む。

30

【0217】

いくつかの実施形態において、溶液はドナーDNAをさらに含む。

【0218】

いくつかの実施形態において、溶液は、細胞が狭窄部を通過する前、その最中、及び/またはその後に、ドナーDNAをさらに含む。

【0219】

いくつかの実施形態において、前記細胞は哺乳類細胞を含む。

【0220】

いくつかの実施形態において、前記細胞はヒト細胞を含む。

40

【0221】

いくつかの実施形態において、狭窄部の直径は、タンパク質及び核酸が通過するのに十分大きな細胞膜の一時的な摂動を誘発するように選択される。

【0222】

いくつかの実施形態において、狭窄部の直径は細胞の直径の約20~99%である。

【0223】

いくつかの実施形態において、狭窄部の直径は細胞の直径の約60%である。

【0224】

いくつかの実施形態において、マイクロ流体チャネルはマイクロ流体システムにおける

50

複数の並列マイクロ流体チャネルのうちの1つである。

【0225】

いくつかの実施形態において、複数の並列マイクロ流体チャネルは少なくとも約2、5、10、20、25、30、40、45、50、75、100、500、1,000、または2~1,000のマイクロ流体チャネルを含む。

【0226】

いくつかの実施形態において、細胞は複数の細胞であり、複数の並列マイクロ流体チャネルのうちの1つに各細胞を通過させ、複数の並列マイクロ流体チャネルの各マイクロ流体チャネルは細胞変形狭窄部を含む。

【0227】

いくつかの実施形態において、(a) 狹窄部の直径は約2、3、4、5、6、7、8、9、10、15、20、2~10 μm、もしくは10~20 μmであり；(b) 狹窄部の長さは約10、15、20、24、30、40、50、60、70、80、90、100、10~40、10~50、10~60、もしくは10~100 μmであり；(c) マイクロ流体チャネルに溶液を通過させるために約10、20、30、40、50、60、70、80、90、100もしくは10~100 psiの圧力を用い；(d) 細胞は約300、400、500、600、700、800、900、100~300、200~700、250~400、100~1000 mm/s、1~1000 mm/s、1m/s、2m/s、3m/s、4m/s、5m/s、6m/s、7m/s、8m/s、9m/s、10m/s、0.01~5m/s、5~10m/s、もしくは0.01~10m/sの速度でマイクロ流体チャネルを通過し；(e) 前記マイクロ流体チャネルは直列した複数の細胞変形狭窄部を含み；(f) 前記マイクロ流体チャネルは単一の細胞変形狭窄部を含み；(g) 細胞膜の摂動は約1~20、1~600、4、5、6、7、8、9、10、12、14、16、18、20、25、50、75、100、150、200、250、300、350、400、450、500、もしくは600 nmの最大直径を含み；及び/または(h) 約1~20、1~600、4、5、6、7、8、9、10、12、14、16、18、20、25、50、75、100、150、200、250、300、350、400、450、500、もしくは600 nmの最大直径を有する細胞膜の摂動は少なくとも約1、2、3、4、5、6、7、8、9、10、または1~10分間細胞膜上で持続する。

【0228】

いくつかの実施形態において、細胞にタンパク質及び核酸を送達した後、(a) 細胞における標的遺伝子の発現が少なくとも約5、10、15、20、25、30、35、40、45、50、55、60、75、80、85、90、95、もしくは99%もしくはこれより多く減少する；または(b) 細胞が複数の細胞であり、複数の細胞における標的遺伝子の発現が少なくとも約5、10、15、20、25、30、35、40、45、50、55、60、75、80、85、90、95、もしくは99%もしくはこれより多く減少する。

【0229】

いくつかの実施形態において、細胞にタンパク質及び核酸を送達した約1、2、5、12、24、1~12、6~12、6~18、12~24、または1~24時間後、(a) 細胞における標的遺伝子の発現が少なくとも約5、10、15、20、25、30、35、40、45、50、55、60、75、80、85、90、95、もしくは99%もしくはこれより多く減少する；または(b) 細胞が複数の細胞であり、複数の細胞における標的遺伝子の発現が少なくとも約5、10、15、20、25、30、35、40、45、50、55、60、75、80、85、90、95、もしくは99%もしくはこれより多く減少する。

【0230】

いくつかの実施形態において、細胞にタンパク質及び核酸を送達した後、(a) 細胞における標的遺伝子の発現が少なくとも約5、10、25、50、75、100、250、

10

20

30

40

50

500%もしくはこれより多く増加する；または（b）細胞が複数の細胞であり、複数の細胞における標的遺伝子の発現が少なくとも約5、10、25、50、75、100、250、500%もしくはこれより多く増加する。

【0231】

いくつかの実施形態において、細胞にタンパク質及び核酸を送達した約1、2、5、12、24、1~12、6~12、6~18、12~24、または1~24時間後、（a）細胞における標的遺伝子の発現が少なくとも約5、10、25、50、75、100、250、500%もしくはこれより多く増加する；または（b）細胞が複数の細胞であり、複数の細胞における標的遺伝子の発現が少なくとも約5、10、25、50、75、100、250、500%もしくはこれより多く増加する。

10

【0232】

本主題の態様は、少なくとも1つのマイクロ流体チャネルを含み、前記チャネルが約30μmの狭窄部長さ及び約4μmの狭窄部幅を備える、細胞にタンパク質-核酸複合体を送達するためのデバイスを提供する。

【0233】

本発明のより完全な理解を容易にするために実施例を以下に提供する。以下の実施例は本発明を実行及び実施する例示的な様式を示す。しかし、同様の結果を得るために代替的方法を利用することができるので、本発明の範囲はこれらの実施例において開示する特定の実施形態に限定されず、これらは例示目的にすぎない。

20

【実施例】

【0234】

実施例1：C R I S P R / C a s 9 遺伝子編集複合体を送達するために C e l l S q u e e z e (S Q Z) を用いた初代ヒトT細胞における B 2 M 遺伝子座の編集

S Q Z プラットフォームの C a s 9 リボ核タンパク質 (R N P ; シングルガイド R N A と複合体化した組換え C a s 9 タンパク質) を送達し、モデル遺伝子座としての M H C クラス1の₂ミクログロブリンコンポーネント (B 2 M) の効率的なゲノム編集を達成する能力を実証するために、未刺激ヒトT細胞において一連の実験を行った。

【0235】

未刺激ヒトT細胞への C a s 9 R N P の送達

標準的な F i c o l l グラジエントを用いて新鮮な P B M C をヒト血液から単離した。次に、T細胞をネガティブセレクション (ヒトT細胞富化キット (S t e m C e l l T e c h n o l o g i e s)) 、計数、洗浄して、送達のために O p t i M E M 中に 10~20 × 10⁶ 細胞 / mL で再懸濁した。B 2 M 遺伝子座を特異的に標的化するように設計された 2.5 モル濃度超の非改変 g R N A (P N A B i o) と 10 μg の組換え C A S 9 (P N A B i o) を予め複合体化した。組換え C A S 9 を 20 mM H e p e s 、 1.5 0 mM K C 1 、 1 % スクロースの最終濃度の溶液に再構成する。C A S 9 溶液に g R N A を直接添加し、氷上で 20 分間インキュベートして複合体を形成させる。再懸濁した細胞に複合体を直接添加する。S Q Z 媒介送達の前に R N P 複合体を氷上で 20 分間インキュベートした。送達効率の代理として用いた 3 k D - C a s c a d e B l u e デキストラン (0.15 mg / mL) とともに R N P (2.2 uM) を共送達した。2つの異なるチップ、10-4 及び 30-4 を用いて 60 及び 90 psi の圧力で複合体を送達した。チップは同じ幅 (4 μm) の狭窄部を有するが、2つの異なる狭窄部長さ (30 対 10 μm) を有する。

30

【0236】

送達の 48 時間後、F A C S ベースの読み取りを用いて B 2 M タンパク質レベルを測定した。B 2 M 発現の減少を機能的な編集の尺度として用いた。1) C e l l S q u e e z e プロセスを用いた送達プロセスと同じ時間室温で R N P 複合体とともにインキュベートした T 細胞 (エンドサイトシス対照 ; 「 e n d o 対照 」) 、及び 2) g R N A はなしで C a s 9 タンパク質とともに圧搾した T 細胞の 2 つの対照を用いた。4 つの異なる細胞集団について B 2 M 発現対送達されたデキストランのプロットを示す (図 4) 。C a s 9 対

40

50

照におけるB2M発現はエンドサイトーシス対照と有意差がなかった。90psiで30 - 4チップを用いたRNPの送達の結果、endo対照と比較してB2M発現は54.4%減少した一方で、90psiで10 - 4チップの結果、B2M発現の減少は25.2%となった。狭窄部が長いチップの方が、RNP複合体の送達が多く、B2M発現の減少が大きいという結果になった。

【0237】

低、中及び高送達集団の境界を定めるためにデキストラン送達を用いた。そして、B2M染色の平均蛍光強度(MFI)を用いて、これらの特定の集団についてのB2Mノックダウンの効率の差を求めた。10 - 4チップでは、中程度に送達された集団の71, 173及び低度または送達されなかつた集団の83, 676に対して、高度に送達された集団のMFIは18, 637であった。この高送達集団でのB2M染色の5倍近い低下は、送達がRNP活性に影響を及ぼす程度を実証している。同様に、30 - 4チップでは、中程度に送達された集団の44, 207及び低度に送達された集団の54, 159に対して、高度に送達された集団のMFIは16, 460であった。これらのデータは、編集効率に対して、単一集団内でさえも、細胞のサイトゾルへの遺伝子編集複合体の細胞圧搾送達システムが重要性を有することを実証している。

10

【0238】

FACS読み取りを確認するために、第2の配列ベースの分析も用いたが、この分析では、標的領域に隣接するプライマーを用いてDNAを抽出して増幅し、これにより次世代シーケンシング(NGS)用の編集領域のアンプリコンを生成した。NGS読み取りからCRISPRバリアントを検出するように設計された単純なアルゴリズムを用いてシーケンシング結果を分析した。予想した通り、配列ベースの読み取りはより高い編集効率を示した。実際に、シーケンシングにおいて同定されたインデルの中には依然として機能的な全長タンパク質が得られるものがあった（すなわち、得られるアミノ酸を変化させなかつた一塩基置換）。

20

【0239】

【表1】

表：10 - 4編集実験からのFACS及び配列ベースの読み取りの比較

	Endo	Cas9	RNP
FACS	0.3	8.15	20.4
シーケンシング	3.87	3.04	27.18

30

【0240】

これらのデータにより、Cell Squeezeプラットフォームによって送達された場合におけるRNP複合体の編集を成功させる能力が実証された。

40

【0241】

編集効率に対するRNP複合体の量による影響

30 - 4チップを用いて、1) 標準1×RNP複合体(10ug Cas9、2.5モル濃度超のgRNA)及び2) 0.1×標準RNP複合体量の2つの異なるRNP量で、未刺激ヒトT細胞にRNP複合体を送達した。送達の48時間後、FACSベースの読み取りを用いてB2Mタンパク質レベルを測定した。B2M発現の減少を機能的な編集の尺度として用いた。4つの異なる細胞集団についてB2M発現対送達されたデキストランのプロットを以下に示す。1) Cell Squeezeプロセスを用いた送達プロセスと同じ時間室温で1×RNP複合体でインキュベートしたT細胞(エンドサイトーシス対照)、及び2) gRNAはなしでCas9タンパク質とともに圧搾したT細胞の2つの対照を用いた(図5)。

50

【 0 2 4 2 】

C a s 9 対照 (g R N A なしの C a s 9 タンパク質) における B 2 M 発現はエンドサイトーシス対照と有意差がない。多い量の R N P 複合体 (1 × R N P 複合体 (1 0 u g C A S 9 、 2 . 5 モル濃度超の g R N A)) での B 2 M 陽性細胞の減少は 5 5 . 4 % であったのに対して、少ない量の R N P 複合体 (0 . 1 × R N P) では B 2 M 陽性細胞の減少は 2 0 . 7 % となった。この実験は R N P の送達に直接関連する投与量依存的反応を実証している。

【 0 2 4 3 】**その他の実施形態**

引用文献は参照により本明細書に援用される。援用した内容のいずれかが本開示と矛盾する場合には、本開示が優先するものとする。さらに、特許請求の範囲の有効性を維持するのに必要である場合、必要な範囲で、参照により本明細書に援用する内容は無視されるべきである。

10

【 0 2 4 4 】

さらに、上記の説明では本発明について言及しているが、説明は 2 つ以上の発明を含み得る。

【 0 2 4 5 】

本明細書に記載の主題は、所望の構成に応じて、システム、機器、方法、及び / または物品で具現化することができる。上述の説明で示した実施態様は、本明細書に記載の主題に合致するすべて実施態様を表すものではない。むしろ、記載した主題に関する態様と合致するいくつかの例にすぎない。上記では少数の変形しか詳細に記載していないが、他の修正または追加が可能である。特に、本明細書に記載のものに加えて、さらなる特徴及び / または変形がもたらされ得る。例えば、上記の実施態様は、開示した特徴の様々な組合せ及び部分的組合せ、及び / または上述のいくつかのさらなる特徴の組合せ及び部分的組合せを対象とすることができます。さらに、添付の図に示した及び / または本明細書に記載した論理の流れは、望ましい結果を得るためにには、必ずしも示した特定の順序、または順番である必要はない。他の実施態様が以下の特許請求の範囲内であり得る。

20

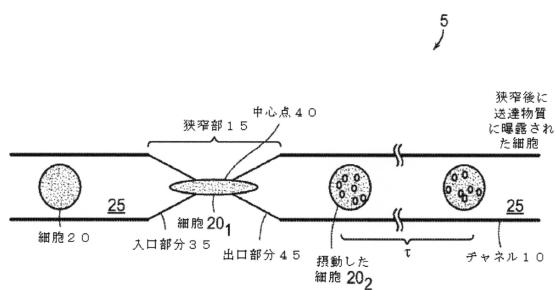
30

40

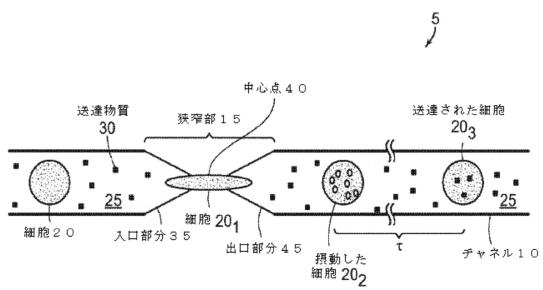
50

【図面】

【図 1 A】

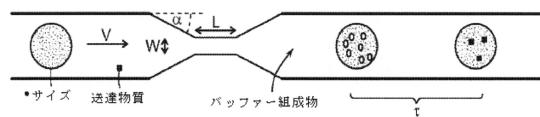


【図 1 B】

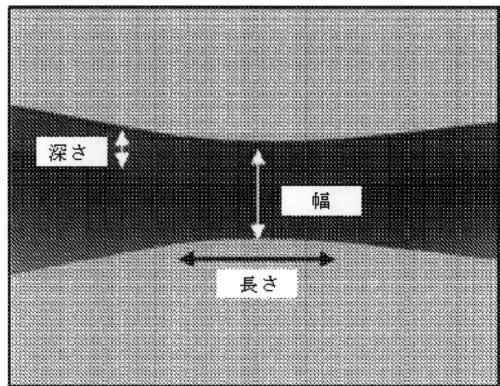


10

【図 2 A】



【図 2 B】



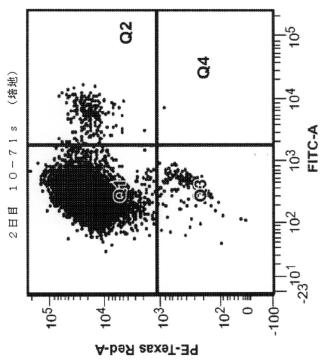
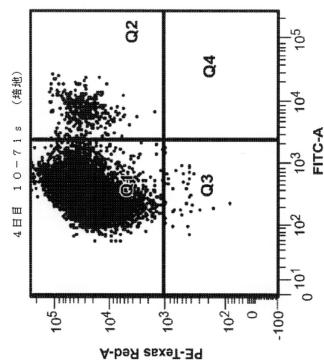
20

30

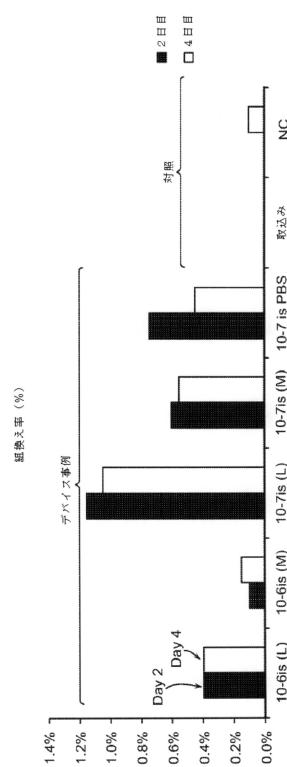
40

50

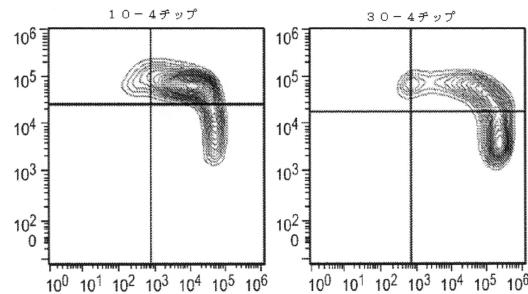
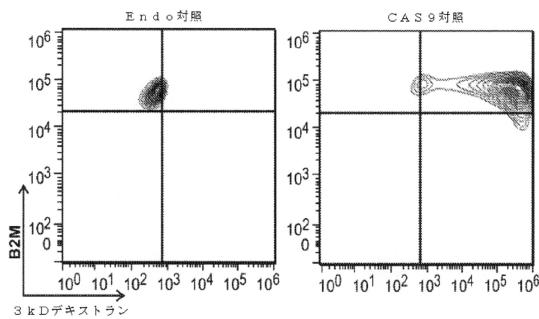
【図3 A】



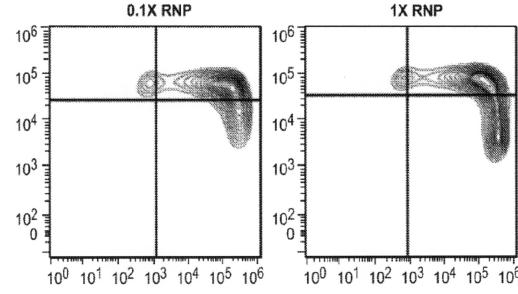
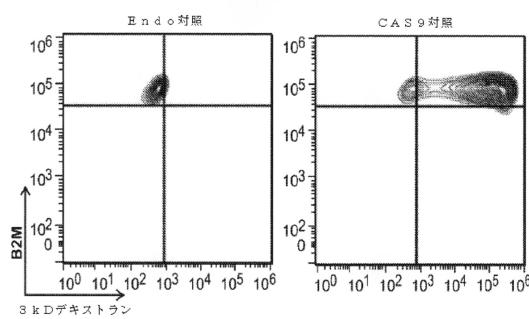
【図3 B】



【図4】



【図5】



10

20

30

40

50

【 図 6 - 1 】

【図 6 - 2】

【図 6 - 3】

【図 6 - 4】

【図 6 - 5】

【図 6 - 6】

CGAGTCGCCAGATTCCTGGCTTGTGAGCAGTGTTAGCTTATCTGATTGCACTGACTCTGGG
AGAACCTTAAATGATGTTAGCTTCAATGAGCTTACAAACCAAGATGCTTAAAGCTGAGTC
AAATGTCAGACGTCGAATGTTAGCTTCAATGAGCTTACAAACCAAGATGCTTAAAGCTGAGTC
GGGAGCAGCAGGGCCCTGGGAGTCAGGGTTTCGAGGCCTACAGGGGGAACTCA
ANATGAGGCTTCTGGGAGCAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTT
GTCGCCACCCAAGGGTACAACTGGGAGCAGCTTACAGCTTACAGCTTACAGCTTACAGCTT
GCTCTGAGCCTTCCGGGCTTCTGGGAGCAGCTTACAGCTTACAGCTTACAGCTTACAGCTT
GAGGAGCTGAGGGCCAGAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTT
GGCTCTGAGGGCCAGAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTT
TGAGAAGGCTTCTGGGAGCAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTT
TCTACCGCCCTGGCTTGTGGGAGGAGCTTACAGCTTACAGCTTACAGCTTACAGCTT
CCATAATAGTAGTAATGTCGAACTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTT
ACATCTGACTGT
TCGTAGTATATAAGAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
CTGGCCCAAAAGGAGGTTAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTT
GGGATGATGAAACCTTGTAGGCCAGGCTTACAGCTTACAGCTTACAGCTTACAGCTT
CTCATGCAACATGGCCAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTT
TCTGT
CACACCAACTTAC
AGGGCCCTTGTGAGGGCCTCCCTACACACACACACACACACACACACACACACAC
CCATTCCTAACTTACACATTCATGTCAGCTTACAGCTTACAGCTTACAGCTTACAGCTT
TGAGTCTAGGGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTT
TGAGGGAGGAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTT
GAGGAGCTTGTGGGCTTCCCTACACACACACACACACACACACACACACACACACAC

【図 6 - 7】

【 図 6 - 8 】

【図 6 - 9】

【図 6 - 10】

TCCAGACCTCAAGTGTGATCATCACACATCACACAGCAGCGCCCTGACACAC
AGACACCCCAAGGCACGCACTCCAGGGCCAGGGCTAACAGGAGGCTAACAG
AGGGGGTGTGCCAGGGCTAACAGGAGGCTAACAGGAGGCTAACAGGAGGCTAAC
AACAGTGTGCGTCAAGAGGCTAACAGGAGGCTAACAGGAGGCTAACAGGAGG
GCAAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG
ACTGCGCTGCAAATAACCCCGTGTCTCCCTGCACTAACCTGCACTAACCTG
GCGCTGATCCATGCTCAGGCTTAGACTGAGGGAATCACTCATTTATTTGGATCCA
GCCCCAAACCCACAGTAACGCCAACAAATAAAACTGCAAGCCGAGGTCCCCGGACT
ATCACCCCCATATAAAGGGTTGGCGTCAGTGGGCAAGGGGTCTGGGTCTGGG
AGAACTGGACAGGGAGGAGGAGTGTAGGSCATCTGGGCTGGGCTGGGCTGG
CAGATTCCTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG
CGGGGCTAGGGGGCAAGGGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG
CGGGTGGCTGGGGTCTCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG
CCCCCTCAGCACACAGCCCCCTTCACAGCTCTGGCTCACAGGCCCATCACAGGCC
CTGGGGCCCTCCCCACCACTCCAACTCCAACTCCAACTCCAACTCCAACTCC
TGTTCCCTCTACCCAAAAGCATTCTTGCCCATGGCTTGCCATGGCAGGATT
TGAGGCTGCCCCCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG
GCTGCACTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG

【図7】

MNPNPRPGKPSAPS LALGSPGASP SWRAAPKAS DLLIGARGPGT FQGRDLRGGAHSSSS
LNNMPMPSQLQPLTPLVMAVPSGARLGPLPHQLLQDRPHFMHQLSTVDAHARTPTVLQ
HPLPESPA MSLTPPTTATGVESLKLARPGLPPG INVA SLEWVSREP ALLCTEPNPSAPRKD
STLSEASQVQSSYPLLNGVCKPGMCA KTLAGMAHQS CCGSCQGPVAVPAPEKGPTRE
A PDSLFLAVRRLWGSNSTEF LHMMDYF KPHMRPPTYATLIRWALAEPEKGPTRE
ONE LIYHWTTRMF AAFRNHPATWKNAIRHNLSLHRCFVRVESEKGAVWTVDELEFRKRSQR
PSRCSNTPGP

【 図 8 - 1 】

【図 8 - 2】

[8 - 4]

【図 8 - 3】

【図 8 - 6】

【図 8 - 7】

【 図 8 - 8 】

【図 8 - 9】

【図 8 - 10】

【図 8 - 11】

【図9】

**MMFWRHDLGSLDAETILKGRGVHGSFLARPSRKQNGDFSLSYR/GDQVTHIRIQLNSGDF
YDLYGGKEFAITLTELVEYYTQQQGVLQDRDGTLTHKYLNCSPPTSERWYGHMSGQQA
ETELQOAKGEPTWFLVRESLSQPGDFVLVSLSDQPKAGPGSPURVTHKVMCEGGRTYVG
LETFDSDLTFLKFTKGTEASGAFLYLRQPYATRVAADLTENRVLNPKQSEDTA
KAGFWEESJLQKFEVNHLQRORPENKGKRNLLPQDSRVLQDLSNLIPGSD
YIINANAYIKNOLQGDPEDNAKYIASQGCLEAVNDFWOMAWOENSERVLMTRVEVKGRNK
CVPYWPVEGMQRAYGPYSVTNCGEHDTEYKLRLTQVSPLONGDLIREWHYQLSWPDH
GVPSEPPGGVLSSFLDQINQRQESLPHAGPIVHC~~SAGIGRTGTLIVIDMLMENI~~STKGLD
DIDIQKTQIMVRAQSGEMQTEAQYKFIVYIAQFLETTKKKLEVLSQKGOESEYGNIT
YUUPYAMKSKRSKSSLESSGATTAAASPVRIGGQIRGLPVPGPPVLSPLDHQLPVLA
DDETMACMADNTDPRCPDR
YUUPYAMKSKRSKSSLESSGATTAAASPVRIGGQIRGLPVPGPPVLSPLDHQLPVLA**

九五已列

【 図 1 0 - 1 】

【 図 1 0 - 2 】

GAAGCAACATGCTGAAATGACAAGATTGAACTTCTAGGGAGATTAGTAGGTAAATGAG
ATCTCCTTGAAGGTCAGGGTAATACTACTAATGACTCATACTTCTGTTAGTTGCTTACCC
ACAGGCCCTATCAGGTCAGGGTCTATTATCCTCAAGTAAAGATGAGTAAACCGAGGTTCTAG
AGCTTAACTGAAATTGCTAACCTCAGACCTAGTAGGTGCTGGAGCACAGTCTTCTCCTTACTT
TATCATCTATGTGCTCTTATATTGTAACTATTGCTAACCATTTGAACTGCTTAACTT
TTGTGTCACCAAGCTAGTGACCTGTTCTAGAATGTCATGTCCTGCTATTGTTATA
GTCACTCTATATACCAACTCACACTTAAAGACCCATTGCTATTTCTTCATGTTGTC
ATTAGTGGTCAGTCCTCAGAACATTTAAAAAACGCTTCAAGCTTATGGCTTATGGTAAAG
AGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT
CTGGTGGAGCTGGAGCTGGAAATCTGGCTCACACAAATCTGGCTCTTCAAGCTTAAACGG
TTCTCTGCCTCAGCCTCCAACTRACTAGGACTACAGGCCCAACTATGCCGCTGAA
TTTTGTATTTAGTAGAGACRGGGTTCTACTATGTTGCTCAGGCTGGTCTCAAACATTCC
TACACTCTGATGCCGCCGCCCTTGGCCTCCAAAGTCTGCTGAAATTACAGGGCTGGCQC
CGTAACTCAGCCTGATGAGTAACTTCAAAATATGGATCTTCTCTCACAACACTG
GCTATTTCTTCACTGATGAGTAACTTCAAAATATGGATCTTCTCTCACAACACTG
CTGAGGAACTGAGTGTGACTTGGACAGGGCTTGGCAAGTGTCTACAGTCTGGGTTAATAT
ATAACTGCATTTCACAGAACACTTATAGCTCCTAAATGCTTCTCTCTGGGACTCTT
TTAAATAACAGTTGCTTATCTTAACTGTCAAACTAGTCTTGGGACTCTGCT
GTGTCGCTCTCCAGACAACTAAGGACTCTGCTGTTAGGAGACATCAGACGGGCTGG
GGGGGGGGAAAAGGTTCTGGGTAGTAAAGACCCCTACATTGCTCAGTTGTTCTTGG
GCTAGAGAATGCTGGCAGTCTGCAAGTGGTAAAGATGACAGTCTGGTAAATGG
TAATCTTCTCTCCCTCAGGGCTGGTAAAGTGGCACCACAAAAGGCCGGTATGAC
ATAATACACRAATTTGAATGATTCATATTCTTGTAACTAAGCTTCTTCTTCTTCTT
TAACACGGCTTAATGAGGATAAATGACATGCAATTAAITGCTATATAITTTAAATGACAA
TTTGATCAGTTGACATACATACACTTGGAAACCACCCATAGTCAGAATATA
ACACATCTTACACCCCTGGTAAATTGCTGCTTATGCTTCTTATACTCTCTCTCT
AGGCAGGCCACTATTCTGCTTCTGTCACATGTTAGTGTGCTTCTGCTTCTG
TTTAAATTTAAATTTAATGTTGAGTACAGATGGGGCTCTACTGTGCTGGCAGGQAG
TCTCAAACCTCTGGGTTCTGACTGATCTCTTCACTTGGCTCTGCTGAGTGTGGGATTA
AGGCATGAGACACCCCTGGCAGGCCCTAGAATTTTTATTTATGTTTATGTT
TTGAGATGGGCTCCTACTTGTGGCAGGCCCTGGAGTGCAGTGGCTCAACATCTGCGCC
TTGTTTCTGGCTGCTAACCTACCCCCCTCCCTACGGCTTCTGGCTTCTGGCTTCAAG
TTGCGCACACACCCGGCTAATTCTTCTTCTTCTTCTTCTTCTTCTGCT
TTGCGCACACACCCGGCTAATTCTTCTTCTTCTTCTTCTTCTGCT
GTGCTGGGATTAACGGCTGAGCTTGGCTCCGGCTTCAAGCTTCTGCTGGCTCCAA
GTGCTGGGATTAACGGCTGAGCTTGGCTCCGGCTTCAAACTTAACTTAACT
GTGCTGGGATTAACGGCTGAGCTTGGCTCCGGCTTCAAACTTAACTTAACT

【図 10 - 3】

TATCTCATTCAGGAACTTATGGCCAGGACAACTACATTAATGGCACTTGTGGGAAATA
TCTCTATTTTGCTGGCTGGAGAGACACAACTACATTAATGGCACTTGTGGGAACTCAG
GTACTATAATTCTGGGGAAAATACCCACAACTGAGGATGGGAATGCG
TATCTGTGATAGCTGGGGTTAGAGAGGTTGATCTGTGAGCTGACCTCTGACATCA
AGTACTCTGCAATTGTTCTGCTTAATTAACCTCTGGGGTTGGGCTTITA
TATGCCTCCGTAATTAAATGTAAGTATTCATGTACTGATCTCTCAGCATCATA
GGTAGGAAAATCAAGCCATTGAGCTCTCTCTGATCATGAGCTGAGCTGAGTGA
TAGTACTGGATTAGTGGTGGCTCTGCCCTTCTGCCCTTCTGCCCTTCTGCCCTT
CAAACTGCTGCTGATTCTCCATCTCTCAGCTGAGCTGACCTCTTACATGTT
TTGGTCTGCTGAATTCTCATGGCAGCTGAGCTGAGGGGCAAAATGGCTGTT
TCTGAGACTGGAGGCTTTGGTTATCTCTGAGGTCAGACTCATTTGCTGGGCTC
TCCTGGTGGGCCACGGTGTGTTATCTCTGCGGAGTAAATAACTGCTCTGIGAAAA
ATAACAGTCTCTGTCTTGGAGCTGGAACCTGGGATGCTCTTAAACGTTGAGGCTCTGA
TGTAGGACCGCATTGTTGGAGCTGAGGCTGCACTGGCAGGACCTGCAATGACTT
ACCCCTGGCTTCCTGGCTGATGGGACTGTTGATCTGGGACATGATGGCTGAAA
CAGGGCAACTCATGTTGATTTGGAGCTGAGGCTGCACTGGCAGGACCTGCAATGACTT
ACAGAACATAGAGAAAACCTGAGCTTATTCCTCTGGGAAACAGCAGCTGACTT
GCCATCACAACGCCAATGTTTAAAGGAAAAGAGGCAAAGRGATGTTGGCAAGGCT
CTGGGAGTGTGGACCCCAACCAAGGATGGGACCTTAAATGGATCTGAGTGCCTAA
AATGAGGCCACTGCTCATGTAATGGAGGACTTCTGCTGGAGCTAACATGTA
TATGTCAGCTCTTGGACACACTGTAATGGAGGAAAACACTGAGCTGGGAGGAGC
ACTGAATTAAGAGGAGGAGACGACTTCTGAGCTAACATGAAATTGATTAAGGCT
GTACCAAAATGAAATTGAAATTGAAATTGAAATTGAAATGAGCTTACTGACCTACCCAG
GCTGGAGTGTGAGTACAAATTTGCTCTGCACTTCTGGGCTTGGGACCTTGGGTTCAACAA
TCTCTGTCGCTCACTGCTCTGAGTCTGGGATACAGGTCCTGCACCCATGCTGCTGCT
AATTGTTATATTTTGTAGGGATGGGGTTACCTGATGTTGGCCAGGGCGCTCTGAACT
CTGGGCTCAAGTGTCTGGCCRCCTGGGCTTCAAACTGGGATGTTAGGGATGAG
CTACACACAGGCTGAAATTGAAATTTGTTGATTAATGACTGTTGAGATGAT
GTTGATATATGTTGATTTGATACTCTATACAAACAAATCACATGTTGTTGTTTGT
TAATATGTTATCTGTCACACATACATGTTATACACACATACATGTTGTTTGT
TTTTTTTTTTGAGACAGGGCTTACCTCTGGTGTGCCCAAGGCTGGAGACTGCCAGTGCATA
ATCTGGCTCTGACTGAGCTGCTGGGCTGACCTCTGGGCTCAAGTGTCTCCCTACATGCTC
TGTAGTGTGGGAGCTGACTACAGGGCAGCTGGCCTAACATCTGGCTTAACTTCTGGGCTCAAGG
TTGAGTACTGGGCTTCTGCTGGCTGCCCAAGGGCTGGGCTCAAACTTCTGGGCTCAAGG
GCTGGCTGGCTGGGCTTCAAAAGTACTAGGATTACAGATGTAAGGACTGCTGGC

【 10 - 4 】

【図 10 - 5】

【図 10 - 6】

CATACCGTTTGTGCCACTTCCCCTCCATTCAAGGGTATATGCGATGTTACTTACATTGCAGGTT
TCATGGTTGTGTCAGTTTTAAACTAACGCCCTATGTGTCACAATTAGTGCTAGGTGTGAG
ATGGGGAGTTCAGAAGCTGTGGTGTCTTTTTTTTTTTTTTTGGCTCAGTCACTTCA
ATATACACAGGCTTAAACCTTGGAGCTGCCCTCATATTAAGGTTTTTAAATTC
ACTGTTCTTATCTCTCTACTARGTTCTCAGGGTCTGAAGTAAGCTCTAACGTCTCCCTG
CTAGTGTAAAGCAAGTTCAGCAATTACAGAAATTGTCAGTGTGGTAATCACGCTTAAAC
TGTAACTGGGAAGCATTITGGTAAATTGAAIACTTTGGAAAAAAAAGCTATGGAA
GGAAAGTTAAATCTACGAAAGCCTCAAGCAGATGTGTCAGTGAAGTGTATTTCAATTTC
TAACATTATATTACTTTTATTATTATTTTGGACGGAGCTGGCTCTGGTGTCC
AGGCTGGAGGTGTAAAGCTGCTCTACGCTACTGAACTCCACTTCCGGGTTCTAAC
TATTCCTGCCTCAAGCCCTCCGAGTAGCTGGGATTTAGACATGTGCCACCCCCAGG
CTAATTTGTATTTTAGTAGAGACGGGTTTCTCACATTTGTCAGGCTGGTCTGAAAC
TCCCAACCTCAGCTGTACCGCCCTGGCTCCGGCTTCCAAAGTGCTGGGATTACGGCTG
GCCACCGCGTCGGCCCTTCAACTATTGTGTTAAATGTAGAGACAGCTCCAGGACCA
TGAAAAGCTGTAGAAGAAGCAGSTTTGCTTAAATGCACTTACATGTCACAACTGCTTAT
TGCAAAACTCATATAATTGTCAGGAAGTACTTATCTCAGCATAACTCCCAACCAATA
ATATAATGTTAAATATGTAAGAAACTTATGTTTAAATGCTTACTTCTCAGGTTCT
GCTAAATACTGGTTACTAAGTTCCCTGAAAATACTATCTCATCTGTGTTGCTTAA
ACAGGATAGCCATAATTGTAACTTGTAAATGAAATAACAGTTATGTAAATAAGGCTA
AAAGGAAAGAACCCACTACCTTATCTTCTGTGCTGATCIGGAGATGNTTAAAGGTGCTT
ACCTAGTTGTTCTTGGCACTTGGCAACTTGGAAACTTCTTTTTTTTTTTTTTTTAAAG
GAACAGGGTGGGCCAGGGCAGCTGCTCAGCAGCTGTTAATCCCCTGACTTGGGGCTG
GGCGACAGTCACTGGAGTCAGAAGTTGGAGACCCGCTGGGACACTGGTAACACCT
GTCTCTACTAAAAAAACAGAAAATTAACCTGGGTGTTGTTACACACTGTAACTCCAG
CTACAGTGGGGAGCTGAAGCAGGAGAACTGCTTGAACCCGGAGTGGAGGTGTCAGTG
TGAGATTTGTCGACTGCACTGCAGCTGGGTGAGACAGGACAGCTCCCTCTAAAAA
AAACACACAAACAAAAGAATTTTAAAGATGAGATGRCAGCAAGAATGAGGTATT
AAAAGAAATTTTAAAGAACTTATGAGCAAGTAACTGGTAAAGTGTGAAAGTGTCTCA
TCTGTTGCGACACAGGCTTGGCAAACATTCTAGTAAGTATAGCTGTAAATTAAACCG
CTGTAATGTATAATGACCAACATATCACATTCTTCTTCTTCTTGTGACAGAGCT
TGCTCTGTGCGGCCAGGCTGGAGTGCAGTGGCAGCATCTCGCTCACTGCACCCCTGCT
CTGGAGTTCAAGTGTCTTCTGCTCCAGCTCTCAAGTAGCTGGGATACAGGTGTG
CCACCCACACTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT
GCCCTGGGGTACCTTGGCTCCGGCTCAGGCTCCAAAGTGGCTGGATGRCAGGTTG
AGCGCTGGGCAATTATGCTTTTCTTATTATCAGAGGCAAGTCAATTGTTGGAA

【図 10 - 7】

【図 10 - 8】

AGGTCAAGGAGTTTGAACCCAGCTTGCGAACATGGTAAACCTCTGTCTCTGCTAAAAACA
TACAAACATTAGCCAGCGATGGTGGCGGTCCCTGTAACTCCAGCTACTCAGAGCCAG
AGGCAGGAAATCTGGCTTACCGGGAAGGGGGTTCAGCTGGAGTGGTGGCCAC
TGCACTTACCATCTGTCAGCACCAAATCTCATCTCAAAAANAAAANAAAAGAT
ATTAGGGCCCAGCCAGTGGCTACACCTGTAACTCCACTTTGGTAGGGCCAGGC
AGGAGAATCATTTGAACTCAGGAGTTGGAGACTAGTGGGGRCAACATAGCAAGRCCCCAT
CTCAAAAAAAAGATTATGGTGGAGCTGCTGTAGACATACCATTTTAACCTTTT
TTTTTTTGGAGATGGAGCTTGTGTCACCCAGGCTSATGTTGAGTGGCGTGTGG
GCTTACTGAAACCTCCACCTCTGGGTTCAACGGATTCTCTCTGCAGCTTCCTGAGTA
GCTGGGACTGCGGCCAGGACACATCTGGCTATACTTGGTATAGTGGAGCTGAG
GGTTTCTACCTGTGGCCAGGGTGTGGCTGAACCTCTGACTCTCAAGTGCGCTGCG
TGCGCTTCAACAGTGGGATACAGGCACTTACAGGCACTGGCAC
CATTTTAACTTTTATATGGTATTTAACCTGACCTTCTTATGTAGTGGATGTTAG
ATACACAAATACCATTTGTGTTACAGTTACAGTATTCAGTACAGTACAGTGCGTAC
AGGTGTAGGCTTGGAGCATTAGGTTAACCRCTTGCCAGGGTGTAGTAGGCTCTG
CCATCTAGGTTGGTGAAGTACGGCTCATGTTAACCTACGCTGACGAATCGGCTTATG
ATGCACTTCTGACACATATTCTGGTTAACCTACGCTGACGGCTATCTGCAAAAC
CATTTTACTTAACTTAACTTATTTACGTTATTTGGTAAAGTGTAGTAAAGTATTG
TAAAGGGATTTCTGCTGCAATTATTTGGTAAACAAAACAGGAGTAACTATAG
GTCTATAAGAAGGGTGTAGATTGGATGGCAACATTCAACATGGGGTTATTTGTGGC
ATTGAATAAAAGGGTACTGGCTGGCGCAGTGGCTCATSCTTATACTCAACACTTGG
GTGGCCAAGAAGGGAGGATGCTTGAAGGCCAGGAGCTGGGGCAGCTGGGCAACATAG
CAAGCCCTATCTCAAAAGGAAAATAAACAAATTAGCCAGGGTGTGGACACC
TTCTAGGGTCCAGCTACTGAGGGCTGAGGAGGATCTGGGACTCTGCTGGCAG
GAGGCTGTAGTGGCCATGATGGCCTACTCAGCTGGGGTGGAGGATAGGGTGGAC
CTATCTAAAACAAAACAAAACAAAACCTCTGTAAATGGTCAAGAACAGTCTG
TGTGGCCAGGTGTGGCTCAGCTTGTAACTCCCTGCACTTGGGAGCTGGCCAG
GAGTTGAGCCAGGAGCAGAGCARAGTAGCAAGACTTCAACAAAAAATAAAAAA
TTAGTTGGCTAGTGGTGTGCTTCTGTCAGCTGGCAGACTCAGGGCTGGGG
GATGCTTGTAGGCTGGGGAGTGTGGGCTCAGTGGAGGATCTGGCTGGCAG
ACCTGGCCACACAGGAGTGGAGCAAGCTGGCTTCTAAACAAAACACARCA
TACTGTAAAGATCTGGCTGGGCGACTGGCTACACCTGTAATCCCAACACTG
AGGGCTGGGGGGGGGGTGTGGCTCAGTGGAGTGGAGGATCTGGGCAAC
AACCGGTTCTACTAATACAAAATAGCTGCGCATGGTGGCTGGCGCAGCTTGTG
TGTGCTACTGGGGGGGGTGTGGCTGGGGAGGAGTCAAGTGGAGGAGTGGCAG

【図 10 - 9】

【図 10 - 10】

TAGTATGCTGTGTTTCTGAGCCCTGGAAATTCTAGTCCTTTATAATAGAAATGAGGCT
TTTTTTTTTTTTGGCTGAGAACTGCACTGAATCAGTSGGGGACTGTGGGTGTAAAG
TTGGCCGCTCTGATGGATGGTAAAGTTCTGGTTTCCAAAGAATGATTGTTT
AAAGACCCCTCAATTGCAAGTAGAACGTACTCAGCTCTGGAGTTTACCAATTAA
TGAAATAATTAAATTATGGTAATAAATGGTAATAATGGTAAATAAGTGTAAATAATTAA
CCATTAAATGAAATTCTTAAAGACATTGAAATTGTTGATGAAAGGTGATGTTAAAT
TATCCCAGATTTTCAATTCTTTTTTATGGCCCTGGATTTCAGTCAGAAAGCCTT
TCCTTATTCTAAGGTAAACAAGACATTCCACCAGTTCTCTAGTATTGCTGTTGTT
ATCTTTTACGGTTTATTATTTTATTCTTGGACAGGGGCTCACTGTCTCA
CTAGGGCTGGACTTCAGTGGATGATCTGGCTACTGCAGCCCTCGCCTCCGGCTCC
GGGGTCAAGGCCATTGCTGCCCTGGCCCTCCAAAGTACGGGATTACAGGCCCTCC
CGCGGCCCTGGCTATTGTTGTTTGTGATTTAGCAGATGGGTTTCTGTTGGCC
GGCTGGCTGCAACTCTGCACCTTAAGTGTACCCAGGCTTGGCTCCAAAGTCGG
GATTACAGGCATGGCCACCGCTSCCGGCTAAATTATTCTGTATGTGATATGTT
ATGGTTCTAACTACTTGTACGGTCATATTCTAAATGGTATGGATTGGATCTTITA
TATTGTTTAAAGATCTGCACTCATATTCTGATGAGTACCATTTGTCCTCTGTTGTTCT
TTGTCCTCATCTTATGGTATAGTACTGTCATATTCTGTTGATTTAAAGGAACTGG
AAGTTGTTCTTCTTGTAGTACTCAGGAATGATTAAAGAATGAGACTTTGGCTT
TGGAGTTCTGGTAAAGTCCTATTGGAAATCTCTGGGCTGGTTTCTGTCGGCTT
GTTCTTAAATGTTCTTCTTATTCTTATTAACTCAGGTAGCTCTGGCTCAAGTTCT
TAGGTTCAAGAGGGCTCCCTCTATTCTTAACTACAGTTGCTCTGGCTCAAGTTCT
ACTCTAATGGTTAATTCTGAGTACTGCATTCTCTGAAAAATTACAGCTTGTCTAG
GTTTCTGACTTATTCCACACTTTTACTCTTCCCTCTGGAACTCATGCCCTTCCA
TAAACAGGACTCTGAGTACCTGAAGTATTCTACACTTCGGGTGGACTTCTGTTCTG
GGGGGGTTTAAAGCAATTCTGGCCCTGCACAGTCCACTGGCTTCTCACACCC
GTTTCTGCAAGATGCTCTTCTTCTGCAAAAGCTGGAGTGGAGGCTGGAGCTGAC
TCAGTCTGGACTTAAGGAGACACCCACATACCTTAATTCTTCTTCTGGCAT
TCTCTTAAAGTGGGGTGTGTTTCTAGTCATTATTATGTTTATGTTCTGTT
GTGTTCTTCAAGAGACAGGGTCCACTCTGCTCCCTGGCTGGAGTGGCTGAGTGGCTCTGATC
ATAGCTTACTGCTCTCAGCTGCTGGCTTCTAGATGAACTCCCTACCTCAGCTCTG
GAGCTGAGCTTAAGGAGTCTGGCTTCTGCTGGCTGAGTGGAGGCTGGAGCTGAC
TTGGCTGTCCTCCCCAGATGGTACTGAGTCTGGCTGAATCTGGGCTACTGCACCTCTGG
CTCCGGCTTCAGGGCTTCTGCTGGCTCCAGCTGAGCTGGAGCTGAGCTGCT
GCCACCATACCGGGCTAATTCTGTTTATTAGTAAACAGGGTTCTCACATTGTC
TAGGCTGGCTCAAACTCTGACCTCAAGTGTACCTACCCACCTTGGCTCCAAAGIGT

【図 10 - 11】

GGATTAACGCTTGAGCCACTGTGCGCTTGTGCTAATTTTAACTTTTTGTGAGA
GATGGGACTCTGCTATGTTGGCCAGGGCCTGCCTAACACCTGCGCTTAAAGCAATCTCC
CACCTGACTGGCTTAAACACACTGGGTTTACAGGCATGAACCATGGTGCCTGGCTT
TCTTTCATTCAGTGTGTTGGTGGAGCATGAGGAGAGATGGTGGCAGGCTTACATTC
TGTGTCCTGGCTACCTGAAAGTCAGGCACCTTCAGATACCTTTAAATAATTAAAC
TCATTTTATCCTCAACACACTATGACATGGTACTGTACACCTTCACCTTATAGG
ACTTAAACAGAGAGGTTAAATGTGACCCAGGGCTCAGAGAGAGCTGGGCTTCAGACCAAG
ACAATTCCTGGCAGGAGCTTGTGGCTACCCCTAACGGCTTTCGGCACCAGTGGTGTAGTGA
TCTCTGGCCTGCTTATGGGGAGGGATGTCCCCTTTTTTAAACCTTTTAAAAAACTT
TCTTCTTATTTATTAATTTGGAGACAGAGCTTCCTCTCTTGGGGAGGTGGAGTGGAG
TGGTGTGATTTCAGCTACTGTAACTCTGCCCTCTGGGTTCACTGTGATTTCATGCCCTC
AGCCCTCCAAGTAGTGTGGGATTAACAGTGTGCCAGCCACATGCCAGCTAACTTTGTATT
ATTTATTATTAATTGGAGACGGAGACTCTGGCTTCTTGTGTTCACTGGTGTGACTGCTGG
TGATGCTGGCTTCTGTAACCTTCGCTCTGGGTTCAAGGTGAGTCTCTGGCTCAGCT
GGGGAGTAGCTGGGGCTATACAGTGTGGGGCCAGCACCATTTCTGCTTGTATT
GTAGAGACGGGGGTCTACTGTGTGGCCAGGGTCACTGTGCTCTGGCTCAGGTGAT
CTACCTGGCTTGGCTTCCAAAAGTGTGGGATTAACAGGTGAGCCRCATCTGGCTG
GATTGTGCTTTTAAACCCCCAACAAAACAAAAAAACACCAACCCCCAACCTAACCC
ATATTTCTGAAAGATTGGTGTCCACACCTGTGTATATAATAATTAGTTTCTCATT
TCTCTCTGGTAAGGGCAACTATGCTGGCTTCTGGGCTACCCCAATTAACT
AATTGGTTGGCAGGCTTAAAGGAGGGTGTGACCTTTTATAGCTTGTAGGTRGA
CAAATGGCCTAATGTAGAGGTCTTACATTACCTTTGAAAGAACCTATAATAGCTG
AAGCCAGTTCTCTTACTTTTGATTCCTAGAATTCTSGCGCTCTTCTCTGGAAA
CRGATGAAAGAAGCTGCAAGGAGGGATTTTTCTTACCTGAGGATAGTTAACCTTCTG
AATGCAAAAGTATGGTGTCAAGAAATAATGATCACRATAATTGAGGTGAGCTTTC
CGGGGTGAAATTATTCCTGCTCAATTCTTCTTCTTGTGAGGCAATCTGGCTCTG
TGGGGCTGGGGCTGGAGTGTGGCTGAGTCATGCTGGCTCACTGCAAGCTCCCTCTGG
TACGGCTTACCTTCTGGCTCAGCTCCGGCTAGTGTGGGATTAACGGGGCTGGCACCAC
ACCCAGCTAATTCTTGTATTTTGATGAGAGCGGGTTCTCCAGGATGGTCTCAATT
CTCTGGACCTGAGTCGGCTGCTCGCTCGGCTTCCAAAGTSGCGGGATTACGGCTGAGCC
ACTGTGGCTGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTGG
TGTGGCTAGGGCTGGAGCCACAGTGGCTGATGCTCTGGCTTAATGCAACCTTGGGG
TCAAGTGATTCTCTGGCTCAGCTCCGGCTAGTGTGGGATTAACGGGGCTGGCACCAC
GGCTGGCTAATTCTTGTATTTTGATGAGAGCGGGTTCTCCAGTGGCTGGCACCAC
CTCAAACTCTGACCTCAGGTAACTCTCTGGCTCAGCTCTTCCAAAGTGTGGGATTA

【図10-12】

【図 10 - 13】

AAATACACATGTATATACATGTGACATACTATGTGTCATATTAAAGTTTGTGG
AACCACGGTGTAGITGCAGACAGTCGGGTCATTCCTCCAAAATACTTCAGTGTT
ATTCTTAAATCACAGGACTCTGGTCAATACCCAGATATCAGGAAATGTATTTA
TARCAAGACTACCATCAATCTTATATCTCTTCAAAITTTGTTAGTATATCCTTAA
CGAAAAGAACACACACACAAAATCTGGCTTCTTATTTGGTTTGTGTCCTTAA
TATGTCAGGTTATGCAATTATGCTGTACTCTGCTAAAGTCCTGTACTGGCCTTAA
ATTAGGATATTCTTGTGATCCGGCAAACTCTCTCATGGTGTATCTTTTTTTT
TTTGGAGATGGAATTGCTATGTGTCGGAGCTGGAGTAATGATGCGATCTGGC
TCAGTCGAACCTCCCTCCGGGTTCAAGGGATTCTCTCCCTCAGGCCCTCCGGAGTAA
TGGGATTGCGGCCCTGGCCACCTTGGCCAGCTAAATTGGAATTTTGTGAGACGGGGTT
TTGCCATGTGGTGCAGACTGTCTGCACTTCGCTTCAGTGTGGCCCGGCTTGGCT
CCAAAATCTGTTGGATTACAGGTGTGAGGCCACTGTGGCCGGCTTTTTTTTT
GAGACAGGGCTTATCTGTGTCGGCCAGGTGAGCTGTGAGTGTACTGGCTTACTG
CAACCTGGACCTCTGGCTCAGGGCACTCCACCTCAGCTCTTGTAGCTGGAGC
TATAGGCACACACCACTGCACTGGCTAAATTATATTITGTAGAGACTGGTTTC
GCCATGTGGCCAAGCTGGCTTGAACCTCTGGCTCAAGTGATCCACCTGCGCTGGCT
CCAAAATGCTAGGATTCAAGGTGAAAGCCACTGGCCCTGGCCCTAAATTITGGCTT
TGTAGAGATGGGGTTCTACATAATTGCCAGGGCTGGCTTGTACTCTGGCTCAAGTGA
TCTCCATACACASCCCCATAATGTCGGGATATAAGCCSAAACCTTGTGCTGCT
GAGGATTAAGTCTAACCTCAGGGAGGGCATTGAACTTGTGAGTGTCTTGTAGT
CTGGCTTAACTGTCAATTGATGTTCTATCATCCTAACATTAGACTGTCTTCTTCTC
ACCTTTGGCTTGCCTAAATATGTCGGCTGTAAATGAGTTTATAGAAAAAAATACCAATTG
AATCGTGTITTCATTACTTGTAGTATACCTGGAAAGTGGAAATTCTGGATCATATGG
TACTTCACAGTTTTTTTCTTGTGAGACAGGTCTACCTCTGTCACCCAGGTG
GAGCTGTAGTCGACGATCTGGCTACTCCACCTGGCTCCGGGTCAGGGATTT
CTGGCTCAGGCTTCAAGACTAGCTGGGATTACAGGCGCCGGGACCCAACTATT
TTGTTTTTTGAGATGGGGTTCTTCATGTCGGCTGGCTCCGAACTCCGAC
TCAGGGTATCTGGCTCTGGCTTCAAGCTCCAAAATTGGTGGATTAGGGTGG
CAGGGTGGCTTCAAGCTTAACTTAAACAACTTAGTTAGGTTAAAAAAATGGAACTAA
GAGAATGAACTATTCTCTGGTTTACTCTGAGTGTGGCTTATAATCTTACTATCATGTTT
TTGGAGGTACAATTGTCCTACTTGGGCGAGGGATCTGGCTGGAGTGTCTGG
TCCTTTCTATGCTCTTGTGGAAACTCTACTGGGCCCAACATAGGAATGTCCTGAGTCA
TCCTCTACTTCTGCTGGCCCAACCTGGGGTCAAGGCAATTCTTCAGGAGGCCA
TTCTCAGGAAATGCAACCCAACTATGAGAAACCAAGGATGAGGTGTTAGGTGAT
TCTGACTCTGGAGTGTGCTTGTGTCACCCAACTTGTAGAAACCAAGGATGAGGTGTTAGGTGAT
TTCTCAGGAAATGCAACCCAACTATGAGAAACCAAGGATGAGGTGTTAGGTGAT

【図10-14】

GTGTCGTATATATAATGTCTGTTGTTGGTGTATTCAAAAAGCACATACACATACACAT
ACCCGGAGACCTGATTTCCTTATTAATTATTTTGGAGCTGGAGCTTCCTCTGCT
GGGAGGCTGGAGTCACTGGCGAACGATCAGGTCACTSAACTCTGGCTCTGGATTCT
AAGCAATTCCTGGTCTCAGGCCCTCAGACTGGCTTGAGCTTACAGGCTGCACCCACCAG
CCACCTTATTITGGTATTTTGGTAGAGATGGGGTTTACCCACATTGGCAGAGTGGCTT
TGAACCTCTGACGCTCAAGTGATCTGGCCGCTCGGCCCTCCAAAGTGCTGGGATTATAGG
CGTGAGCCACTGTTCCCATCAGAACGATCACATATCTATTCTATATCTCACATTCTGCT
TTTACATATATAATTTAAAAATTACAGTTGGCACTAATCCCTAACTAACATTCACAT
CATGGGATTATTCCTGGCTTCTCCTCTCATATTGTCCTCCCAACAGTGAGAAC
TCCTGGTGTATGTCATCAGCATGGTAACTTATTAAAGAACATTCTTCTTTTTTTT
TTTTTCTGAGATTGAGTTGGCTCTTGGCCAACTGGAGTGGAGCTGGTGTGATCTT
GGCTACCGCAACCTCTGCCTCTGGGTTCAAGCGATTCTCCGCTCAGCTTCTCAGT
AGCTGGGATTACAGGCGATCAGCACCATGCCAGCTATTGCTATTTTAGTAGAGATG
GGTTTCTCATGTTGTCAGGCTGCTCTGGAACTCCGGACCCAGCTGATCTGGCT
GGCGCTCCCAAAGCTCTGGGATACAGGCGTGGCCAGCCTGGCTCTAGTT
TATTTTCTTACATGTTGCTCTGGCAGACCGCTGGCTCTGGCTCTAGTT
CTCTAGTGTCTTAACTCCCTGGCCCCAGTGGCTCTGATCATATTCACCCAGAC
CCTACTGAATCCAGGCTCTACCAAGGGAAAGCAGGGAGGAGGAGTGGACCAAGGAG
AGAGGGGGAGGGAGGGAGGGAGGCTTGTATTGTTCTAAATTCTACTCTGCT
ATATGGCTCTAGCTGTTGGTGTGGAAAGTGTGCTGACAGTCTGTTTTTTTAT
TACTCTTGTCTCTTCTTAACTGGTTCCACCTTAACTACTGGCTGAGGCG
AAACCTGAGCTTGGCAAGAGGGAGGTTGGCAGTTTGGCAAGGCTGTAAAGTA
ACCCCTGGAGACTTCAACATTCTGGTAGTGGTTGAAGTGGATGACCTGGT
GTGTTTCTAGGTAGGAGAAGTGGTAAAACCCTGTTGGATAGCTGTCCTGGATTCTGA
GTTTGAAGGCTTATCTGAGGCCCTGGGCTGCCCTCAGGTTGGAGTGGCTCTGG
CATTTAGCAGAACAGGAGTAAGGAGGGCCCTTCTCCTCCCTCTGAGACCTCTAGGAAGG
GAGTTGGAGGAGGCTATAGGATCTTCAAGGCCCTCAGTGTGAGACTGTTCCACACA
TCTTGAACCTGGTTCTGCAATTCTTCTCTGGTGTGATTATTAAATTATTTAT
TCTTCTCAATTTTTTTTAAATAGGGTGGAGCTTCTGGCAAGTGGCT
GGCTTGAACCTCTGGCTCAGGCAATCTGGCTGGCTCCAAAGTGTGAGGTTACA
GGCGTGGAGCCACTATGGCTGGCTCTTCTTGGAGACAGCTGTTGCTGTGGCCAG
GCTGGAGTGGAGGTTGGAGATCAGACTGGAGCTTACAGGCCCTGAGACTCTGGCTCTGG
TCTCTGGCTCTGGACCTGGCTGGGCTGGAGCTTACAGGCCCTGAGGCTACAGTCTG
TTTAAATTAATTTTGGCAAGAGTGGGCTCTGGTGTGATTCTTAAAGGGCT
ACTCTGGCTCAAGTGATCTCTGGCTCAGGCTTCCAAAGTGTGGGATTACAGGCC
TTTAAATTAATTTTGGCAAGAGTGGGCTCTGGTGTGATTCTTAAAGGGCT
ACTCTGGCTCAAGTGATCTCTGGCTCAGGCTTCCAAAGTGTGGGATTACAGGCC

【図 10 - 15】

GAGCCACCACTCAGACTTGTGACTCTTAATAAGAARAACTACTTGTAAAGAGITTC
TTAGCATGACTCTTCTTATCACAAGTAAACATGACIGAGGAAGTTGGTGCCTCCCTT
GCTTCCCTGCCCAGGCCCTTCTTCCCTCCCTTCCCAGGAAACCCCAAAGAGTTG
GCAATCTGAACTTGAAAGCTGTTTATAGTGTACTGACTCAGTGTACTGTATGARCA
ATATAAGTTGTTGTTGTTGTTAAAAAAATCACAATCATGATGTTTAAATGTTGATG
ATTTTGCACTAAATTTTTTTGGAGCTCATGCTGAITGATAACGATCTATT
TTTTTGAGATGGAGTTCTAGCTTATTGCCCAGGCTGAAGTCATGGCTGATCTCAG
CTCACTGCAACCTCGCCCTCCGGTTCAGGTATTCTCTGCTCAGCCTCCGGAGTGG
GGTTCAGGATACAGGTCATGCCACATGCCAGGACTATTATTGTTTATAGTAGAGTGG
GGTTTACATGTGAGCAGGGCTGGTCTCAACTCTGACTCTGGTGTACTGCTGCT
GGTCCTCCRAAGTCTGGAAATTACAGGCTGAGCTTACCATGCGGCCCTTTTTTTT
TTTTTTTGAGACAAAGTCTGCTCTTCTCCAGGCTSGAGTCAGTGGCCACAACTT
GGCTCACTGCAACCTCTGCTCTGGTCAAGTCAGCAGTTCTCTGCTCAGCCTCTGAGT
AGCTGGATTACAGACATTCACCCATGCCAGTTAAITTTGTTATTTTGTTAGAGAC
TAGGTTTACATGTGTCGGCTGGCTCTGACTCTGACTCTAAAGTGATCCATG
CTTGGCTCCRAAGTCTGGGGTTCAGGCTGAGCTGATCTGGCTGGAGAAATC
TCATCTTACTCTTACTCTCCCTGCACTATTCTCCATTGAGTGGCTACCTTCTT
ATTCTCTGTTACCCCTCTGTTGTTCTGACTCTGTTATTCTTCTTCTTCTT
GAGACAATCTGCTGTTGCCCAGACTGGAGTCAGTGGTGTACTTGGCTCRCCGA
ACCTCCACCTCTGGGTCAAGTGATTTCATGACTCAGGCCACCTTAAGTAGTTGGGATTA
CAGGGCTCTGGTCACTACCTGCCACACAGCTAACTTTGTGTTTTAGTAGAGTTGGG
GGTCACTGTTGCTCAGGTTAATCTCCTAACATTGGCTCAGGGAGTCTGGCTCTCAG
CTCCTTAAAGTCTGGGATATTGGCTCAGGCCACCTGGCTCTGGCTCTTGGTCTT
ATAAAAAAATAGCAAAATTAAATTCTTACTATTCTTCTTCTTCTTCTTCTT
AGAACATTAATATTTCTCTGTAATTCTTCTGACTTACCTAAATCTTGTGATCTC
TCTCATCTTCTTAACTAGCTGEGATGCCATCTCATGCTGAAATTCTTAACTCTA
TACTGTCACCGAGGTTGGCATTAGTGTGCTTAATTTAAATGATATAATGTC
CTCAATGATATGATTAACCTTCTGGATATTGCTGAGGCTTGGGATAACACTCTC
GTATGGAACTGGTCTCATCTGGCTTCTTAAATACATGTTGTTTCTT
TTTATTTTCTTGTGAGGGAGTTCTACTCTGTTGCTCCAGGTTGGAGTCAGTGGTGGAG
ATCTGGCCGCACTGCAACTTCCGCCCTCCGGGTTCAAGCGTATCTCTGCTCAGCCTCC
CCAGTGTGTSAGGATACAGGTTGGCTGCCACACGCCACCTAAATTCTGTTATTGTTAGTA
GGAGGGGGTTCTTACCATGTTGGCCAGGGCTGGTATGGCTGTCACCTGGTGTACT
CCACCTCTGACCTTCAAGTGTGTTGGGATCTGGCTCAGGCCATCTGGCTCAG
CTTCTTCTTCTTCTTAAATCTGCTGATCTGGGTTTCTT

【図10-16】

【図 10 - 17】

【 10 - 18 】

【図 10 - 19】

【図10-20】

【図 10 - 21】

AGCCACTGCTCCGGCTTGAAGCAATCTAACACATTTCGCTGTAACTATTITA
TTTCAAAATTAACCTGAAATGCGATATCATATCTAAACATTAAACACTATTITCT
TAATATTAAACACATTAACGTCACATTTCTCTGATTCACACACACACACACAC
ACACACACACACTGCAATTGTTGTTCTTCTTACATGATGCTACTCTGTGCCC
AGCGTCTGGTCAATGGTGCATCTACGCTACTGCAATGCCACTCTGGCTCAAT
GATTCTCTGGCTCAGGGCTCTGGTACGCTGGACTACAGCTGCCACACCCTACCTGG
CTAGTTTTGATTTAGTAGAGGGTGGGTTTACCACTGTTGGCCAGGTGGTCTCAA
CTTCGGACCTCAGGTGATCCACCCACCTGGCTCCAAACTGCTGGGATTACGGATG
AGCCACTGTGCCAGCAGCAATTGTTGAACTGGAGGTCTTCTTCACCTTGATATT
GAAAAAATTCACAATGTTGATAAACAGATTCAATTAAGGACTCTGATATTGCCATTATC
AGCTTTCATCAATTCTCTGTCATCATTTATTTATTAATTAATTTCAATTGTTGG
AATCTTCAAAATGATCACATAACACCAACATTGTCATATTACAAATTGTTCTTAA
TTTAAARAAATTTTGTTCATCAATTCTTCAGATAAGGCTTCACACACTGTGGTGTGGCTAAG
TCTCATATAATTCTCTGTGTAAGAAATCTAAAGCTGGCTGGGACACGGCTGTCAAT
TCAGCTTCTGGAGGGCTGAGGTGGGGGAGTCAGGAGTCAGAGATCGAGGCCATTCT
GGCCAACATGGTGAACCCGGTCTCATTAACAAAATACAAARATTAGCTGGGOGTGGTAGT
GGCTGGCTAGTGGCAGCTGGAGGCTGAGGGAGGAGATCCTGGGACACAGGCA
GGTGGCAGTGGCAGTGGCAGGAGATGGCAGGACTCTGGACTAGAGCAGAGTGG
GCTTCATCTCAAAACGAAACAAAACAAAATCTAAAGTCTCTTCTAGAATACTTTGATGC
CCCCTCATCTCTCTTCTGTCTCTCCCTCCCTCTCTCTGCTCTTCTGTTGAA
GAAAGCAGATCTTCTGCTCTGGAGGTACTTAACTGTTGATGAGTGGCTCTTCT
GTGGTGGACTTAAGCATGATCTCCATCTTCTTCTGTTGAGTGTATCTATCTAGA
CTTCATGGTRATCAAGTTGTTCTCTGGCAGATGAGTCATGATGTTGCTATCTAGAGG
TTTATGTCCTGGTTGTTCTCTGTTGAGTCTCTGATCAATTATTACTTGAG
TOCATTAACTCATAGGGACTATATGGTAGTGTATTTGTAATTATTATCATCTCTTCTCAT
TTGTTAGGGTGGCATTTCTTATAAAAGCTTCTACCGCCAGGGTGTATTCTCTT
CTTACTAAGGACTTTCTCTCTTCTTCTTCTTCTTCTGAGTGGCTACTCTACTG
TTGGCTAGGGCTGGTGTGAGGGCAGAACACACAGCTGGAGACTCTGGCTGAGGTGA
TCCTCTGGCTCTGGTCTCTGGTGTGAGGGCAGAACACAGCTGGAGACTCTGGCTGAGGTGA
AATTTTGGATCTTCTGGTGTGAGGGCAGGGTCTACTTATTTCTGGCTGTTGAAAT
GTCTGGGCTCAAGCAATTCTTCTACCTCAGCTCTGGTGTGAGTGGGACTACAGGCACAT
ACCACCATGCCGACGTAATTCTTAAATTCTTAACTTGTGAGGAGTGGTGTGTTGAAAT
TTGGCTAGGGTGTGCTGGACTGAGCAGCTGAGTCATGCTCTGGCTGAGCTCCAG
FTGCTGGGAGGTATTAAGGTTGAGTGTACTCTAGCAGGAAACAGCTGAGCTGAG
CTTCAAGGTTGAAATCTGGTGTGAGGGTCAATTGAGTACTTCGAAACAAACATCAGG
CTTCAAGGTTGAAATCTGGTGTGAGGGTCAATTGAGTACTTCGAAACAAACATCAGG

10

20

30

40

【図10-22】

GATTTTACTAACCTATTGATCTTACCCATGATCTCTTTCCTAATGCCAAATCCT
TACTCTTGAGCCATAATAAGGTATTCTATTGTGTTTACCCACATTACACACAACT
CTTAGAATARTGCTTCCCATAATGTTACTGAAAAGCTGGATTTTTGGCT
TCTTAAATAACCTTCAGAGGTGACTGAGTCTGATTTCTGCTGGCACACTGG
CTGAGGCCATAATCCCCAGTCTGGAGGGCAAGAAGGGAGATGCCGTCAGCTAGG
AGTTTGAGACAGCGGGGCAATATACTGAGGACCTGTCTCATAAGGAAAATTTAAA
TTAACCGACAGTCGGGCACTGTCCTATACTGTCAGCTATTGAGGGCTGTGGCAGAGT
AGGCTTAACCCGGGGAGGTGGAACTGAGCTGGATATACTGAGCTTGTACTCTAGG
GTGACAGACAGGGGGACCTGTTTAAAAAAATAAAATAAACTCTGTGCTTAG
ACTCATTTGTTAATGCTGCTCTCTGTGGCTATACTGCTAATGGTGTATGGTGT
TTATTTGTTCTTTTAAATACTCTCTGCTTAAGTGTATTTATAATTACAAATACT
TGGCTTGATGCTAAATGAAACAAAGTGTATTCAARGAGTCACCTCTATCCT
TGTCCTTCTCTATGTTTAGCCATGTTAAAGTGTATTGTTTATCATTATTTCAA
ATAATAAGAAGATATCCCTCATCTCCACTTTCTAAACAGTACATACTTACATC
TTTTCTTAAACCTTGTCTTTTAAATCTGAGGAACTCTGAGTATCCTAATAGTCT
AGAGATAGTCCTCATCTTTTAACTGATGCTGCTGCTGATCTGGTACCATAGT
TTATCAACCTTATGAGGGCATTTGGTGTATTCTCAAAATCTACAGAGGATTAC
GTGAATAGCTTGTGTATGCATCTGCTCTACITTTGCTGACTACTGGTAAATTACAT
TTTTATGTTCTGTTATAAAATGGGGTATTATTCATCTATAACTTTTATTATACA
TGACTTTGGTGTAGCACTGTTAACCTTGTAGCATACATGCTGAACTGAGCTACTGGT
TTATTTCTGGTAAATGCAAAATAATGAGCTTGTATTTGTTATCTAGTCTGTTCA
CTTATACCTCTTCTTACAACTGAGTGTGTTGATCATTTGTTGATTAAGTCGTGTA
AAAAAAATACGCTTCTTTAAATAATTTTAACTTTAAATAATTTTAAATAAT
AGACACACAAAGTAAAAAGAGGARGAAAAAAAGAGACAGGGCTCTGCTATGTTG
GGCTGGTCTCAACCTCCAGGCTCAATGATCTCTGCTTGGCTCTAAAGTGTAG
CCACCAACACTGGCAAAATACTGGTTTCTTAAACAAAACATCAGGTTCTGTC
CATGGTGTCTTCTTAAACTGGTCTAGGATCTAGGAAAGCTAGGAGCTAGGACT
TTCAACAAAGGCTTGTGGCAATGGTATTTTGGCAATTTCTGGTGGCTGTA
ATGTTGAGTTAGAGGAACCTTACGGCTGACTCTCTGGCATGGAAAAGATGGTGT
TCTAAATGTCGACCTGGTGTATTTACACTGTCACATCTCAATGTTGCTCATTTTATA
CATTATAACACAAAGGGAAAAATGTTAGTTGACTTTAAGGGAGTGGAAAATAACGA
GATCACAATCTGACTCTACAGGCTCTCAGAGGCTGAGCTAGGTTGAAATTTGTG
TGCACAAATAAGGGCATTAAGGTTCTGAAAGCTAGGCTATCTAAATAGCTG
GGCTGCTTCTGAGCTTCTTACATGGCTCTCAATGCTGTTTAAATAATGGTGT
CCCTTGGGGTGGACAGCGRAATAAGGTGAAACAGATAAGGCCATTATTGAA

【図 10 - 23】

【図 10 - 24】

TCAGTTTATCTTATAGAGTGACAGTCAGTGGGASAGACGGCRGAACAAAGAA
ACAGTTCAATTTCTGTGAGATGGTGTAACTGCTAACAGAAAACAAACTACTGT
AGATAAARAGGGTTTGTAGGCCCTTCTTACATTAGTCCTTCTTGTAAAGGGCAT
TTGACAACAGCTCTGGAAAATTATGGGCCAACCTCATTAACCTCAGGGAGCA
TTCTAGTAGAGGGACACGCACTGCAAAGGGCCCTGAACCTGGGGTTTGTACTTGT
GCACATGTCACACAGGCCAGTACATTTGGAAATGGATGGAAATGTAAGAGAGAAGT
TGAAAGCCAGGTGCGTGTGCTACATGCCACATTTGGGGAGGCTGAGATG
GGAGAATTTGAGATCAGCCTGGCAACAAACAGGACTCTGGCTTAATTTTGTATTT
TAGAGACAGGGCTTACCATATGGCCAGGCTGTACATCCTAACCTCCACCTGG
ATCTGCTGGCTCCAAAGTGTGATTTAGCTGAGATTTAGCTGGCTTGGCCCCCG
CGTATCTTGTCTTAAAGATATCTCTGGCTTGTAGGCCAAGAATTAAAAATAAAA
AAATTTAGAAGAAAAAAATAAGTAAGTAACTATACAGGTTGGTGTGGCGCTAATGG
TGAGTGTCTATTCTTCTCTCTAGGTATTGGCTCTGTCAGCAGTGCAGGC
GAATGTCATTAAGGGCATCCCTAGGTGGCTGGGATCTGGCTCAGGCCAGTT
TCTGGTCTCTCTCTCTCTGGCTTCTCTCCTACATTATCTAGAATTTCTGCTAG
AGCTGGGCCAGTTCTTCAACAGGCTCTGGTAAAGGCTCTGGGCTTATTTGGC
ATCTGAGTCACTGGCTCTGCTCTGGCTCATTAGGTGGCTGGGGCTCTGGCTGG
CAGGAATGTCAGTGGCCATTAATGTAAAGGGTTGGTCTTACATTACATAAGGGAGAAC
TGAAAACCTTAACCTCTCCACAGTAGTGGAGTAGTGGCTGGTACTCACAGTCAGTGT
GCCCTGGGFACTCACATGTCACACATGGTACAGGACATCTGCTTCTGGATACCTT
TAATGTTTATTAAGAATGTTAGCTGGCTTCTGGCCTACTGCTTAAGGAAATCTGCTA
GGGATTAATTAAGGAAAGGGTTTAATGCTCAGTGGCTCTGGAGCTCTGGAA
GCTCAGCATCTGGCTCTGGTGGGCCCTCAGGAAGCTCTGGGCTCATGGAAAGGCCA
AGGGGGCCAGCACCATCAGTGGCCGGAGGGTGGGCAAGGGTGGGGTGGGCCACGCTT
TTTTTATTTATTTATTTATTTGAGACAGTGTCTACTCTTCTGGCAGCTGGAGTGC
GTGGCGTGTACTCTGACTCTGGCAGGCCCTGGCTTCCAGGCTTCAAGGAATTCTCTGG
CAGGGCTCTGGAGTGTGGGACTACAGGGGCCATCACATGGCCAGCTGGATTTTGTAT
TTTACAGGAGCAGGGTCTGGCTTACCATGTGGCAGGCTGGCTGACTCTGATCTA
GTAATCCTGGCTGGCTGGCTTCAAGGACTCTGGGCTCATGGAAAGGCCA
GGCCACCACTGTTTACACCAACAGATGGACAGTGAACCTTGTAGAGTGGAAACTCACGT
GAGGATGCCACAAACATCTGAGGATTCACCCACCTCTCTTGGGCCACCTCCA
CACTGGGACTATTTCTGACTCTGGAGATGGGAGGCCACAGATCCTAACCGCTTAC
AAATTATAGTTTATGCACTGGTGTGGCTTCTGGCTACTGATCTGGTACTCTGCACTGAG
TTGCTGTAAGCTGGCTCTTGTGGCTTGTGGAGATCTGCAACTGGTGGCTTAC
ATGTTTATTTGGGAAATGGGTATCTGGCAGCTGTGTTCTGGCTTACATTCTGTT

【図 10 - 25】

【図10-26】

CCACAGCTGGAGTACCGATGTAGATTCTATGACAAATGTTGATCTAGAAAATAGCC
ATGATCTGATCTATGCCTTCCTGTATGTTCTACCCACTGTATGTTGATCTACATGCCATCTC
ATTGACCTCTACCCGGCCCTGGGGTAGAAGGAAGTGTCTCATTCACATCRAAG
GAAACTGGCAGCACAGCAAAAGTACCTGGCGAACAGAGGAACTTAGCGCAGCTT
CCCTGATTAGTCAGACAGACTGAAAGGCTCTGGCGTCACTCTGAAAGCTTGGACTGTAT
CTTGAGGGTGGAGGGAGAATTGAGGGTAATTAAACAGGAACGTGAGGAATTACCCCTT
GCTATGTTGATTCCTTGGCCAGTGTGCCCTGCCACGCCCCCCCCTCTAGTGAAGTGT
CATGGCAAGTGTGGCTGTAAATGAAGGCCGAATGTGTTCTGACAAAGTTGTTTAAA
TCAGCTGCTAATATTGCTCCAGTCCCCTGGGGTAGTGTGTTTGTGTTGTGTTGCT
TTCAGCTTCAATTGTTGAGGTGTGATGTTGATGTTGACTGTGTTGACTCTTGGACTG
CACCTTAACCTCTTAAACATCAAGGTAAACCTCTTGTGCTAACACAGCATATTGCT
TTGGTAAGACTGGCTCACAGTCCAAGGAATGCTTGGCCAGAGGGAAACCTGCCCTAA
CTCCCTAACCTGAGCTCATAAAAAAATCAARTGACTGATTCCTGTCAGCTTAC
CTACATTGTTTATTATTGTCAGGTTTCTAGCTAGTAAATGCTTGTGATGAGCTTAA
TGTCCAGGGCTGAAGGTTGATTGTTACAGCTGAAATTCACCTTAAAGTCGACGCT
CTTACAGCTTGTGAAATTCTAGTAAATGAAATTAACACAGGGCTGTGTTGTTG
AAATATGACTGTGTTTCTTCTGACTGTCTTGGAGCTGACCTTGTGATTGTTT
TATAAATTCTTCTTCTAGCCACTTTCACCTTCTGATTGCTTGTGATTGTTT
CTTAAACACTAGGCTTCTTCTAAACCCTAACTTCCCTGAAAATGACTTTTTTTT
GAGCAGGAGTCTGTTTGTCCACACAGGCTGGAGTGGCCCTGGCTCCATCATGTAAAC
AGRAAG
AAAAGAAG
TCTCTGTTGGTTCTCATCTTAAATTGTTGTTTAAAGGTTTCTTGTGATTAATTG
TATTGTTCTATGTGAGAAAATTAAAGGGAGAACACAGTGGAAAAAAACRAGAAAAG
GACTTCATAATCTCTGACCTCTGGAGGAAAAAAACCTACCTTACCTATTGTTCTT
CTCCACCTTTTTTTTCTGGAGATGACTCTGCCCTTGTACCGCCAGCTGGGCTCA
GAGCTGATCTGGCTCTGCAACCTCTGGCTCTGGGTTACAGGCTATCTGCTGCCCTA
GCCCTGGAGTATCTGGGATTACAGGGTGTGGCATACACCTGGCTATTGTTTGT
TAGATAGACGGGGTTTGTGATGTTGGCAAGCTGTGTTGGCTGAGCTGTGTTT
TTGTCATATGGCCAGCTGTTGTCATGTCAGCTGAGCTGACATGTTTGTCTGTTGCGAG
GCTGGCTTAACTCTGACTCTAGGTAATCTGAAAGTGTAGGTTATAGGGTGTGGC
ATTGACCTCTGGCTCTGCTCTGGCTTTTTTTAAAGGAAAATTAACCATTTTTCTT
TAGATAGCTGCTCATTTGTTGCGCAAGGGCTGGTGTGAACTCTGGGCTCAAGTGT
TCCAGCTGAGCTGGCTGGAGTGGCTGGAGTCAAGAGCAGCATCTGGTGTGTTT
ATTCTTGTGACTCTAGTGTGAACTGAACTCATGATTACATCAATGCTTATTGTTGAGT

【図10-27】

TGACCTTCCTTCACTTCACTCATCTTCCAGGTGTTAATTCCTAGTCATCTT
AAATGGACATGAGACATTCTTGTATGACAACAACTAGTTTATTAGGGGTCCTT
CTGTTGGACATTATATTCTCAGCTTCCACAGTTCCTGGCAGTCATGATA
ACCTTCTTCTTCTTCTGGAGCAGGCTCCTGCTTCAGGCCAGCTGGATGCAG
TGCCACATCTCTCTCAAGTGTCTCTGNTGACCCCTCAGCTGGTGTGATTACCG
TGCCCCATGTCGGCTAATTTGTGTTTGGIAGAGCTGIGGTTTACCTAGTGGCCA
GGCTGGTTGCAACTCGCCCTGAGGTGATCIGCCACACTCAGCCTCCAAAGTGTGG
GATTACAGGTGAACCCATACGCCCTGACCCAGTGAACATTCTTGAGCTATGCCAAC
AATTCTGAACATTCTAGGATGAATTCTTAAAGAAGTGTCTGATCAGGCTTCTT
CTTCTTGTGACTCTTGTGACGATTAATGATCTTCTTCTTCTTCAAGCTG
AACACAGGAGTCCAACTTCTACACGCCAAAAGGGTCAGGCAAAGAAACAAA
AACAAATATGATATAAAACATCTGCCGTAAGTATCARATTCCGCTCAGTAATAGT
CACTCTGGAGATTGTGATTCTGACCGACTCTGTACCTTCTCAGGGTCTGTGCTCTT
GTAGACATGCGAAGGCTTGTCTTCTTAACTGCAAGCTTCTCCAAATTAATCAGCA
TGGTGGGTTGTGATGACTCTTCACTGACGCTCCGGTGTATGTTACTGAGATAAGT
GCCACTAGGTGGCTTGTCTGTTGTGTTGTTCTCTTAACTTCTGCTGCCCA
AGATAGATCATCTCAAGCTTGGGGACTCTCTGACTCTGAGGCTTCTGGAGGAT
GTGCTCTGTACTCTCTCCTACIGGGCTGTGAGICCTGGTAAAGGCTGCTCTTCTGG
GACTCAGTCCCTTAAGTGGGAAACAGCAAAACCTCTGAGGGCTCTAGAATGTC
TGTCTGCTGATCCCTGACTGCTCAAGTACTGGAAAGGCTATATACTTAAACTGCTCAG
AAGAAAGACTTGTGGCCGGCCGGCTACACCTSTAATCCGACATTCTGGGAGG
CGGAGGCGAGGGACTCCTGTGACTGAGGCTTGTGAGGCTTGTGCGCAACATGGTGA
ACCCATCTACTAAATAACAAAMATGGCTTATGTCGTTGTGCTCCCTGTAACT
CAGCTACTGGGGGGCTGGGGGGAAATTGGITGAACCTGGAGAGTGGGGTGGAGTG
AGCCGAGATGTCGCCATTGCACTCCAGGCTGGTGAACAGCAAAACTCGCTCAAAA
AAAAAAAGGAGACTTGTGAAATTCTCCAAAGCTGTAAACCTGTACCTTCTTAAATT
TTTGAGGACATGCTCCTACTGTGTCAGGGGCGACTGAGCTACTGTGAGCTCNAAC
CTCTGGCTTCAGGAGTCTCCCTCCACCTGCTGACTGAGTGGACAATAGGGAGG
CACCAATACRACCTGGTGTGTTTACAGTTCTGTAGGGGGCGCTGTGGCTTACGGCT
GTAATCCCAGCACCTGGAGGGCCGAGGGGGCTGAGTCACCTGAGGTAGGAGTGGAGA
GTAGCTGGGCAACATGTTGAACCCCATCTCTAATAAANTTACAAAAAAATGCTGGG
GTGTTGGTGGGACTCTGTGATCTCCAGCTACTGGAGGCTGAGCTGGAGGCAAGGATC
GCTTGAACCTGG
CACCGGG
AAATTTTTCTAGGCGGGGTCCTCTTGTGTCAGGCTGCTGAACTCTGGGCT

【図10-28】

【図 10 - 29】

【図 10 - 30】

【図 10 - 31】

CTGTAGTCGCTTCTATTACCGATTTAGCTGAAATTCTCTGCTTCTGGCTAAAGATA
CAAGGAAATTCTCTGGCTTATGGGTTTGGAAAGTGAGTGGTGGATGACTG
CCGAAAAGAACAGAACACAAAGGCATCTGCCAACACARCTACCAAAATAAGATA
AAACTCATCTCTACTGAGTAATAGCTCTTCGACCGAAGATGATTCTTC
TGCCATATTTCATCACCCTGCTCTGAAGCCAGCAGTAITGCACTGGCATCC
CAGGCCACCCAGTTAAGGGAAAGTGAGCTGAGGAGGATCAGATGGTCTGGATAATAG
AAAAAGCAGCTGGTCAAACCCCATGGCTGCTCCCTTCGCTGTAGGTTTACACTT
GGGGTAGATAAGGACAGACAGCTTCCTCAACTGGCGGAAAGAACAGACTGCTGGC
TCTGGTGGCAGGCTACGCCCTCTCCAGAATTTCTGCTGGCTTCCTATGGACTGCTGGT
AGATTCTTGAATTGACACTCTAGTGGTTCTGACCAACTCTCACTATTGGTGTGGCTC
CTCTCTGCATTTGATTCTTAGAGATAAAACCAAGACCCAGACTCTCTTGCACAT
GTGCTTGGAAACAGACCTGACGGCTGCTTCTCCCTCCACTCTGCTGGTCTIC
AAACACTCTGTTCTTGTGACTCTCTCTTTTTTTGAGACAGAACCTCTCTGCTG
CACCCAGGCTGGAGCTGCGATGCTACGCTACAGCTGGACAGCTCTGGCTCCAGGTT
CRAAATATTCTCTGCTCACCTCCAAAGTCTGGGATTRAAGTGGCTCTATRCG
CCTGGCTATTTTTGTATTCTTAGAGACAGCGGTTTACCTATTGGCAGGTGGCT
CAAACCTCTGCTCAAGTGTGATCTGGCCCTCGGCCRACCGACTCTGGATTACAGG
CATGAGCCACTGCGCCCCAGCTGATTCTTACAGATAAAACAAACATTGACTCTGCTTGA
CATGTGCTTGGATCAGGTAACTGCCAACAGTGGCTTCTCTCCCACTCTGCTGGTCC
TCCGAATGGCTCTCTTTCTATTGAACTCTCTGCTCTTCTGAAACCTAACAGATG
CGAAAGCAGGCTTATTCTCCATTGGGTTTAAGCAGAACACTGAACATTGGTTG
CTTGTGTTAGGCTTTTATTCTAGGTTTACAGATAAATTCTCTTCTTCTGATCTCTC
CAGTGGTAACTGTCATTAAGTGGCTGATGAGTGTCTAAAGATAATGGCTGATG
CTGGTGGAGAACGTCATTAAGGAGCGCCGCTCATGACTTACAGCTAGAGAACTTAACT
TCATGGTGGTGGACAGTAAGTATTCTGCTGATGAGACTTGGGAACTGGTGTG
TGTGCTGGAAATTCTGGGCTTGGCCCTTAATCATGTTGGCATACATGCTACATTGCTCA
CTCATCTGATTCAGTGGCAATTATGAGCTTCTCTATSTGCGCAAGCTGAGTAAAGCAG
TACTGGTACATTGTCAGACAGGAGCTGTTGCTGGCTCTCATGCGCTTGGAGGATA
GAACTTAAAGAACAAACTTACCTGGCCAGGCGCCGTTCAAGGCTGTAACTCCAGAC
TTGGGAGGCTGAGCTGGGCTGGATCTGGCTTCTGGCGGGGAACTGGCAGAGGAGCTGG
GAGACTGGGATAGGGTGGCTGAGTGGCTACAGGCTGTAGGAGGCTCCGGCAGGGC
CTATGTTGATGGCTCTCTCCAAAGTACACAGACACTTCAGGAGTTGCTTCTTGTG
CTTCTGGATGGATGGATTATTATTAAGGAACTTACAGGAAAGCTGAGAAGCAGTGA
TGTGAGGAGTGGCCGGGGGGTGTAGCTGAGTGGCTGGGNTTGTGTTGGCTTGG
CGAAGGAGGCTGGGATTAGACAGACACTTGTAGAGCTGGTGGCTGAGCAGGAGCTGGTGTG
CTATGTTGATGGCTCTCTCCAAAGTACACAGACACTTCAGGAGTTGCTTCTTGTG
CTTCTGGATGGATGGATTATTATTAAGGAACTTACAGGAAAGCTGAGAAGCAGTGA
TGTGAGGAGTGGCCGGGGGGTGTAGCTGAGTGGCTGGGNTTGTGTTGGCTTGG
CGAAGGAGGCTGGGATTAGACAGACACTTGTAGAGCTGGTGGCTGAGCAGGAGCTGGTGTG

【図10-32】

GGAGACACCACAAAAGCTGAGGAGAATGATGTGGCAGTGTGTTCTTCCTGGAA
TGTGTTCACTGGGCACTTGAGCACGTCATGTGGCACATTCTATTCATTCTAGAGATTT
TGTGATCATGRCCTTCTGGTCTCGGACCATGTCCTGCTTCTGGAGATAGCTGGGGAC
AAAACAGATGGTACCTTCTGCACCTCCAAAGTGTACRATCATACTGGCAGATACTTACCG
CGAATGTCACATATAAATATCATTAATCTAGTATCTAGGCTGATAGAAAAGGAA
TGCAGGAGTTAGGGTCTATTGGAGGGGAAGGGACTTTTTTTTTTTTTTGAAAC
AGAATCTTGTCTGTCACTCCAGACTGGAGTCGACTGGTSCATTCAGCCTACTGCAGC
ACAACCTCTAACGTCAGTGTCTCCTGCACCTCAGGCCCTCAGTAGCTGGGGCTAC
GGTGTGGCCACATGGCCACCAATTTTTTTTTTTTTTTAGACGCTGGGTCTCTTCA
TGTGTCGGACAGCTGCTGTGAGCTCTGGCTTAAGGGCATTCCTGGCCTAGGACCT
CCAAAGTGTCTGGAAATAACAGGCGTCAGGCCAGCGCACCCGGCAAGGGGGGGAGGT
TAAGGCATAGGGACAACTGTGTTGAGTCAGCAAAGGGAGTTTGGGGGTTGTCTA
TGTGTTAGGACCCAGGTGGATTAGTTAAGGATTCCTCCACCTGTAG
ACTGGAGGGGGCAGGACTGTGTTAGGATTAGGCCAATTGGAGGTAGTGAGC
AGAGATTTAAATAATAGGGATTGACTGAGCCAGTGCCTGGGTGAAAGAAGGG
CCCAGTGTGGCTAGGGCAGGAGCAGTCAGGGAGGTTAGCTGTAGGGGTTGACCAAGGG
GAGGGTGTGGCTAGGGCAGTGTGGCAATTCTTACCTGGAGAAAGGCTGGGGGG
AGTGGAGGTGAGGAAGCTGGGGGGATCAGGAGCCTTTGTGACRCAACAAAGTTGAG
TGCTTGGACACATTGAAAGTGGAGCGTCAGGGAGGCAAGGGTGGGGTGGGGT
GGGAGGTGGGATGCGAGCGCTGTGGATGAGTCAGTTCTGCTGATAGGGGCACTGTT
TCTCTGGCACAAGGGGGCCTAAAGGAGAGGTGGCCACATGTGCTGGCTAGATGAGCTG
GAGTGTGTTCTCTTCTCTGGAGCTGGCAAGGCTAGGCAAGGCCATCAGACAGATA
AGGAATAAGAAGAGTAAATATGAGAAAGGGGAGRAAAGGTAATGCAATTG
TAGCAGGGCGAAGAGTCUGAGAATTCTCTGGCTGTGCCAGTGTGGGGTGGGG
ACATGAAATAATACTGTACCCATTGCTGGTTTCTCCAAGGACACTGGCTCTG
GCACAAACAGAAAGTACCTGGTTCTGGCAAGGCCAGGGCTTITGCCATAGTGTGAGCT
GGAGAGGGCTACAGGAGGGCAGCATGTGGAGGGAGGAGCTGTCTGGGGAGCAGGG
ACTCTGG
CTCAGTGTGG
TGGGTTGGCCCAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
GATGATGG
TGTAGCTCAAGTGGAAATGGGAGAAGGGCTGGGGCTGGGGCTAGGAGCTGGGG
CACGATATGG
BAGCTGTGTTGTCAGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
BAGTAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA
BAGTAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA

【図 10 - 33】

TATACCTCCCTTTAAAGTCACCTTAAGGCTGGCAAAAGTGCTCATGCCGTAACTCC
TGCACCTTGGGAGGCTGAAGTGGGTGGATCTACTGAGGTCAGGAGTACAGACCCGCTG
GCCAACATGGGAACACTCTCTCTACTAAGGATAGCTGCTGATCAGGCTGAGT
CCAGACTCGGGAGCTGAGCTGAGGCCCCAAAGATCTTGAACTGGGGGGCGTAGG
TTCAGCTGAGCTTGATCAGGCTACTGTCAGGCCAGCTGGTGTGGTACAGCTGG
CTCAAAAAAAGAAGCTACTGGTCTAGATGTTGCTCATACATTTAGTGGAGAAGATGAGG
ARTGAATGGTGTGCTACATTACGTCATCTGGATAGCCAGGAGAACATGATACTTTTG
GTTTCTGTGACCTTGGTGGCAATTGTTGATCTTATTAACTATTACATTATCCTCAAATAG
CCATAATGTCGTGACTCTCTCTTGTGAGGCGAGAGCTGAGTATTTCAGGGAGGCTC
ACTGAGCTCTTAAATGTCATTTAGTGGAGAAGAGGATATTAGTATTGTTGTTTGGAA
AATAGGAARATTGGCTTGGAGATACCCCAAGGGCACCTGTGAGTGTGCTGGATT
GGACCTTAAATCTGACTCTGAGCAAAAGCTGGTCTTCTTCTGTTGCTTAAATTA
ATAAGTTAAATAGAACATACTGGCAGGCGAGTACCTGGAGTCAAGGACTTGGCAACCAGCTGGCA
CTTGGGAGGTCGAAGCAGGAGATCCTGGAGTCAAGGACTTGGCAACCAGCTGGCA
ATATGGGAAAGAACCCCATGCTACTAAGAACATAAAATAGCAGGGCTATGGTGTG
TGTGCTGTAATCCAGCTACTGGGAGGCTGGAGCAGGAGATGGCTGACCCGGAG
CGAGAGCTTGACTGAGCAAGATCACCCTACTGCTCAGGCGCAACAGACTGAG
ACTCTGTTGGAAAGAAAAAAAGAACAGATGACTATAGTGTGTTTATATGGGG
GGAAGGTTGAGTATAACAAAATAACAAAGAGGAATGAATGTTAAGTGAATGGCTTT
TCCCCTGCTGTTCTCTCTGCTGGGGAGAGACCTGGATCCTAGAGGTTTCAGTGG
CTCCAGAGCTGGCAGGAGCTGGCAGGAGATGGGAGGATAGGATTTACTCTGGCTGGTAA
TTCAAGAGAGATGTTGAGGGTTAGTGGAGGAGAACAACTGGCCAGCAGGACTCT
TGGCTTCTTGTAGACAGGCAAAAGATGCTCTCAGGATTCGCAAGGAGCTCT
GGGACCTCAGAGGAGGACCTTCTAGTGGAAAGCTGGGAGAATCTCTCTCTGATTG
GATTTAGGAAAGCTTAGAACCCGGTGTACTCTCAACCTCTGTTTATTAACTCTTTC
TGGCTTCTCTGCTCTACTCCAGGGAAATAGCGAGAGAACAGCTGCTGGCAATACCACTT
GGGACCTCTGGGAGGACGACGGGGGGCTGCCAGGCGAACCTGGGGCTCTGGCTGG
GAAGGGTCCACCATAGAGGAGGAGCATGAGGAGCTGGGGGGCTCTGGCTGG
TGCACCTCTGCTGCCCCCTCTAGGCCAGGGCTGGGGGGCTCTGGCTGG
TGGCTTCTTACCTACCCACTCTAGCTGTTTACTGCTAGGAGAAGATTAAATATCTGTTGG
GCAAGTGAAGCAACTGGCATGGGGGACATTGTTGATCCAAAGGCAATTCTCTGAGACCTA
CAGGACTCTGCTGCTGCACTGGGGGGCTGAGGAGTCTGAGTAAACAGCAAAAGCTAGT
GAGGACTTATCTGAGTGGGGGGCTGAGGAGTCTGAGTAAACAGCAAAAGCTAGT
GTTTCTCTTCTTCTGAGTGGGGGGCTGAGGAGTCTGAGTGGCTGTTGAGGTATGAA
GTTGGGGGGTGTGGTGGAAATAAGGCGCTGACTCTGAGGAGTCTCTTAAAGTCACTTTTG

【図10-34】

【図 10 - 35】

【図10-36】

TTAACCTTCAACTACAGTAGTGAAAAAAACCTCTAAATTGCAAAGCAGTC
TGTGGAAATTCTTCTTACCCAGCTTCTCGACTGTAACCTTTAGCACCTTAACTTAT
CATTCGGTTTATTCTCTCTGTAAAATTTAAAATGTAATTTAAAAGTAAAATGTTG
TTGGTCACAAACATTATAACCCCTTCTCTAATTCATTCATTCTTAAAATGAT
ATCTACRAGCTTACACRATCTTAAATTGTTCTATTTCTTAAATAGGGCAATTCTCTAAC
ATAACTAACCTACACATTATTGTATACAGTGATAAAATCAGGAAACAACTGTATAC
ACACTTATTCTAACCTACAGCACATTCTCAATTGTTGCTCTAGTATCTTTATGG
TCCAGGGTCACACAGTGATTGGCTATAATGTAATCTTCTCTTGGAGACAGG
GTCTCACTTGTGCCAGGTGGAGTCAGTGCAATTAGTCAGCAGCTTGG
ACCTCTTGGCTACAGGTGATCTTCCCACCTCAGCCCTCAGCAGTAGCTGGAGACACAGG
TGTGACCACCATCTCTGGCTAATGGTTCTATTTTGAGACAGTGAGCTGG
TCCAGGGTCACACAGTGATTGGCTACAGTGACCTCCGGCTTCTTGGAGACAGG
TCTGGTACATACAGGCCTGACTCACACCACTGGCAGTATATTAGTATGTTAGTCT
TAATCTGGAAACAGTTCTCAGTCATTCTTATTTCTGACCTGGATGTTTGAGAG
TTTGGCCAGCTTATTAGCAGAATGCTTCTGAGTTGGATTTCTCAGTGGTTCTCTTG
ACTATATTCTAGTCAGTCATTGTCAGGACTGTCACRGAATGTTGTTGAGCTCT
TACTACATCACATCAGTACACAGTCAGTCAGTCAATTACTAGTGGTGTAACTTGG
ATCTACATGGATAAGGGTGTCTGTCATAATTGCAACGGTAAGTACTGAGCAAAAC
GTAGCTGGACTACAGGGTCTGCAAAAATTTGCAACGGTAAGTATTGTTCTGAACTT
TTTTTTGGAAACAGTATTCTTCTCTTAAAGTAACTCTGTCTTACATTGTT
TTTCTCCCTTAAAGTAACTCTTCTGAGATGAGCTGACTCTGAAATTACTCT
TTCTCTATAATACTTTTCTGAGATGAGCTGAGCTCCAGGGCTGGAGT
CGAGTGGTGCACATCTGGCTACTGCAAGCTCACCCTCCGGGTGAGCCTTCTCTG
CTCTAGCCCTCCAAGTAGTTGGACTACAGGGCCGCCATCACACCTGGCTAACTT
GTATTTTTTAGTAGGAGCAGGGGTCTCAGCCGGTATAGCAGGATGTTCTGCTCAGCC
TTGTGATCTGGCCCTGGCTTCCAAAGTCGGTGTAGTGTGAGCTACTGGC
CGGGCTCTATAATACTTTTCTGCTACTATTATATATCTGTTGAAATTCTGGCT
AAAAAAATTACTTACTGGTATTCTGCAATGGCAATTCTGGTCTCATCTGGCTT
CCCCGGTTTAAAGTAACTGACAAACAAAAGTGTATATAAGTGTACACCTGATA
TTTGATATATGTAATCTTGTGAAATGTTATCAAAATGAGTTAAATGATCCTAA
CATCTGAGTACTTCTTCTTCTTCTGAGCAGAGTCCTGGTTGAGCTGG
TCTGAGTCACACATCTGGCTCATACACCACTCCACGGTCTGGAGTATTCTCT
GGCTTGGCCCTCCCGAGTAGCTGGGATACAGTGCCACCATACACCCGGCTTATT
GTATTTTTTAGTAGGAGCTGGGGTCTACTGTGGCTACAGTGAGCTGAGCTCTGG
TCAATGATCTCTCCCTCTGCTTCTGCTTCTGCTGAGCTGAGCTCTGG
TCAATGATCTCTCCCTCTGCTTCTGCTTCTGCTGAGCTGAGCTCTGG

【図 10 - 37】

GGCCGGGCACTTCTAGTAAATTTCAGGTACATTTTTTTTTTTTTTTTTTTTTTGAGATGGA
GTCCTACTTGTGACCCCTGGGGTCACTGGTCACTGGTACGTTGCTGCACACTGGAACCTCT
GGCTCTCTGGGTTCACTCAGGGTCTCTCCRCCTCAGCTCCTCCAACTGGTCACTGGAGATRCA
GGTACCCGCCACCATGGCTGGCCATAATTGTATTGTAGTAGATGGCGGTCCTACATT
GTTAGCCAGGCTGGCCTCAACTCTGGCCTCAGGTGAATCACCACCTGGCCCTCCCAA
AGTACTGAGATTAACAGGCATGAGGCCACCAACCCAGGCCACATTAGTGTAGTATTAACCTAT
AATCACCAGTGTGTCATTAAGTAGTCTCCAAATGTTATTCATCTTGTAGTAACTTCAGTTG
TACCCCTTGACCAAAAGTCCTCTGGTTCTCCCTACCCCACTGGCTTGTAACTCCTT
AATCTCAGTCTTGTAGGTTAGTGTGTTAGTCTTCACTACATAGTGAATCAGGGACT
GATGTTGTTATTCATTCAGTCAATATGTCATCAGTGTCTTCAGTAAATGACGAGGTTCTT
CTTTTAAACTAATATCCATGCTGGACACGGTGGCTCATGCTCTGTAATCCCAGCACCTT
GGAAGGGTGAAGGGGGTGGATCACTTGTGGCTCAGGGTCTGGAGACCCRGCTGGGCAACAT
GGTGAACCCCATCTCTACCCAAAATAAAAAAATGGCTGGATGTGTTGGGGCACACCT
GTGATCCCCAGTACTGGGACACTGGAGCAGGGTACGTCAGGACCCGGGGGGGGGGGG
TTCAGCTGGCCAACTGGTCCACTGGCTACGTCAGTGGCTGGATGTGTTGGTCACT
GTTTATTTTATTTGTGTCCTGTGTTTGGTATCAAATCTCTAAACCTCTGGCATG
CCATTGTCATGTTACTTCCCCTATGCTTCTCTAGAACTTTAAAGGTCTCATCTTCC
CTTTCTGTGTTTAAAGTGTGCAAGCTACTATAGGAAGGGCTTCTCTCTCTCTT
TTTATTCATCTTCAGGATGGCCACCTACTATTTTTGTGTTATTCCTGTTAATT
GACTGTTATTTGGCTGTGAGTGGAAACCTTTCTAGATGGGTACTCTGTTGGTACTCTT
TCCATTCTCAAGGACTCTTGTGACTGATGGTCTAGGCTCATCTCTCTTCCCAG
CCATTCTCAAGGAACTCTGGTCTTGTAGTGAAGGACTGTTGGTAAACCAAGATCT
GGGCACTGGGTCTACTTGTGTTGACTGGTACAGTGTCTTGTGAAATTGCTTAATTGCTGAT
CAATTACTGCTCTCTTGTAGTTCCTCTCTAAACCTCTCATATGTTGACAGRCGGTC
CTGACTTATGATGGTTGACTTGTAGATTTTGATTTGATGTTTGTGAGGCAATRCA
TCCATTCTGTGTTTACTCTTCACTTCAACACTTATTTAAATAGGGTGTGAGTGA
TATGGCCACGTGTTGACTGATGTTGAGTGTCTAGGCTCATCTCTCTTCCCAG
CTCTGTGTTGTGGTAGGTGTAGATGTGTTAATGCTGACTTGTGATGTTTCAAC
TTATGATGAGTTATGGGGATGTATCCCTATTAAGTCAGGGAGCCTTATACATCTCTG
TATAACAGAGTGAATGTTCTTATACCTTCTCATCCTACACTTCCCTGAGTAACTTTAC
TAACCTGATGACATTATCACAACAACTAAACATTAACATCATCATCTGTTACTTAA
CTAGAGTTATGGTTGGTCCAGGTCTTCCATGAAATCTTCTTCTCTCTCTG
CAATCTGATGTCACACTGGTGGCTCAGTGTGACTGATGTTCTCCATCTGCGA
CAGCTTCTAGGCTTCTCTGTTTCTCATCTCTGAGGATTTAAAGACTGTGTC
AGAATCTGTGAGTATCCCAACTCTGGTGTAAATCTATGTTCTCATGATTAAC

【図10-38】

【図 10 - 39】

AATGAAAGCTTTCATACATTGTCATCCTCCGATTAATTAAAGTAGGTGTTAC
TGTGCCAGGGACTGGGAATTAACTGTTCTGTCGATGGAAATGAACTCCAGCA
TGTTGGAGTAGAGGAGCTGGGGCAATGGTGGAGTGTGATGGGGGTTGGAAAAGTA
AGCTGAGCACCGAGCTTTCAGTGGAGGAGCTGGGGCTTGTAGCCCATTAACAGA
TTGAGACTTATTTGGAGGCATGGAACTGATGGAGCTCTGAGCCGGGAATGAC
ATAAAAAGATCCTCATTTAGGCCGGATGGTGGCTCAGCGCTGTATCCAGCACITT
GGGAGGTTGAAGTGGGGTAGTGGCTGAGGCCAACAGTTTGGAGCTGGGCAACAT
GTGAAACCCCTGCTCTCATCAAATAACAAAATTAGCCTGGGAGCTGGCTCACACTG
TAGCTCCAGCTACTTGGGGAGCTGGCATGAACTCGCTTGAACCCGGGGAGGAG
TGCAGTCAGCGGGATTGTCGACTCTGCACTCAGGGTGGTGAACAGACTGCTG
TCAAAAAAAACAAAAACCCCTCATTTGAAAGGGAAACCTGGCTTGGGGTGAAGAA
TGGGGGGCACTAGGCTAGAGCAGCTGCGAGGGTCAGTGAGGGAGCTGCCAGTGTGCR
GTGAGAACCCCTCATGTTGGCTAGGGTGGGCAAGCTGAGCTGGCACAGGCCAG
TAAACCRGAAATACTCATGTTGGCTAGGGAGGAATCATGGTTTGTAGCTCTCA
ATAGTGGCAGGACAGCTGGCTAGCTGGCAAGTGTAGCTTGGGGTGTCTGAC
TTTCATTTACCCCTCTCTCTCTGGCTTGGGGTGTAGTTTGTAAATTAAAGGTTA
TGATGTATCATGTTACTAACATGGAAAGACTTAACTACAGAGTATTTGGAGCTT
GTGATCATACCTCTCTCATCTGGGTGTGIGTCAGITGGCAAGGCCATAAAATTGT
TATAACATCTGCAAGGAGCTTGGTACTCTGGCTGCTGTACRACGGCAAG
TCATTTCTGAGTATTCTGGCTTGGCTACTCTGGCTGCTGTACSTGAGTATTGGGAGATRGA
ATCAAAACAAAACCTCTTGGCTTGGCTGCTGACSTGAGTATTGGGAGATRGA
CTGGTCAAGCTACTAAGCACAGGCCATCTGGTGTGAGTGTGAGTGTGAGTGT
TACATAAGCTAAAGTTCAACATTAACTGGCTTCTGGTGTGAGTGTGAGTGT
AAGTAATCCCTGTTCAATTGTACTAACTATTCTCACAAGTAAATTCTGTTGAACTC
TTTATGTATTAACTGAAAGACTGATTTTATTCTCAAGGAAGTATTGAGTAGGGAGGGGG
AAATAGGATCTGGCTGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT
GAGATCTTCTACAGCTGGGTATATGTTAACACATCTGCACTTACAGCAAAACACTCA
AAAGGGCTGAGATCTGAGTGTGTTTAAATTGCTTAAAGGGAGTATCTGAGTGT
AAAAGCAGTTTACTCTCAGTCTTCTATTAACTTAAATTTTTATTCTGAGTGT
TTGAGATTTAAACCTCAGGATAAGTTTATTCTTAAAGGAAATTCTTATTCTGAGTGT
TTTGGAGTGTGGAGTCTACTCATCTGCACTGGCTGGAGTGTGGTGTCTGG
CTCTAGCTGGCAGCTTCTATCCAGGGTCAAGTGTGAGTGTGAGTGTGAGTGT
CTGGGGTACAGGCTGTCAGGACAGGGCTGGCTTAAATTGTTGTAGAGATGG
GGTGTCCCAGTGGGGAGCTGGGGTGTGAGTGTGAGTGTGAGTGTGAGTGT
GGCCTCTCAAGTGGCTGGGGATACAGGTGATGCACTGCGCCGGGGATAAGTTA

【 図 1 0 - 4 0 】

AAATAATTTCTCTGCTGGCGCATGGTGGCTCATGGCTCATGGCTAACCAGCAGCTTGGG
AGCTTGAGGCCAGGACATCACTTCAGGCGCAAGAGTTGGACCCAGCTGGCCACATAAT
GAGACCCCCCTTCATCRAAAAATAAAAATAATTGGCTGTAGITGGCTGTCTGTAGC
TATGGCGGGCTGGAGGGGGAGCTGGCTGTAGCAGGAGTGGCTGGCTGGCTGGCTGAGCT
ATAGTTGGCAGACTGCCTGTAGCTGGCTGGAGCTGGCTGTAGCAGGCTTCTCTTTAAAAAA
AARAAAAAAAGGCCAGGGCACAGTGGCTCAGGGCTGTAGCAGGAGCTTGGGAGGGCG
AGGGGGGGTGGATCACTTGAGGCCAGGAATTGTAGACCCAGGCTGGCCACATGTGAACCC
CGCTCTCTACTAAAAATAAAAATAAGCTGGGTGTTGGTGGCTCAACCTGTAAATCCAG
CTACTTGGGGGGGGTGGAGGAGAATTGTAGCTGAACCTGGGGGGGGAGGGCTGAGCTGAC
GGAGTGTACRCACCTGACTTCAGGGCTGGCTGTAGCAGGAGACTCTCTCAAAACAA
CAACACAAAAAAACAAATAATTGTGCTGGCAATTCTCTTAAATAATTATAATTGTAATT
ATATTGTATCTTCATAAGGATTGTTCTTATATAGCAAAAGATTCTTGTAGCTGTAAACAT
TTACGAGCTGTATTCAAGAAATGGTTATTATAGCAAGTCTTGTGTTTAAGAAAATGGCTG
TCAAAAAAATAATAATGTATATAACCAACAAAATTTTTGCTTGGAGTGTGTTT
GAGCTGTCTTCACACTAATAAGTTCTACTGACTGCTTATACCTTAATAAATGTT
TGGCTACTTAAACAGGGCAAATAGTAGCTAGCAGGAAATAATAATTTTTATTATTGTT
TAGAGACAGTCTGACTCATACCCAGGCTGCACTGAGCTGGGGCTGTCTGGCTACTGC
AACCTCCACCTCCAGGGTCAAGGGATTTCTAGCTGTAGCTGCTCACTCTCTGGAGATT
ATAGGCATGCAACCCACACTCTAGCTAAATTTTGTTATTTTAGTAGAGACAGGGTTTGC
CATGTTGACCCAGGCTAGTCTGAGCTTCAGCTGAGTGTACCATGCTGCTTGGGCTCC
CAACTAGCTGGGGATAACAGGGTAGGCCCACCATGGCCAGGCTTATTTTTTTTAA
GATGGGGCTGGCTTCTTGGAGATGGCTGTCTGGGGCTGGAGCTGGCTGGCT
CATAGCTACTCTGAGCTGGCTTGAATTCCTGGCTCAAGCACTTCTGGCTCAGGCTCCG
AGTAGCTGGGACTCACGGGGCTGGCCACGGGGCTGGCTGTAGCTACATGTTATGTAATG
GCTACCTTAAGGGCAATCTTATGTGAGCTAACATTGATAATGAGTGTCTCATGTATGG
ATTAGATTTTGAGGTGGCTGTCTTGGAGCTGGCTTACAGGCTGCTTATAGAGTAGAAAAGGGTGGACAAAC
CTCTCTACTAGATGCTTCAAAATAATTTAGTACTGCACTACAGGAGTGGACTGCTAG
CTGGCTGGGTTCTTGGCTTCTTCCCCTGGGGCTGTAGCTAGACCAAAATAAAGTTT
AAGGGTTGGTCAAGGGTGTGGTGTGGCTCATGGCTTGTAGCTACCTGGGGGGCTGG
GAGCTGGGTCATCTGAGCTGGAGGTTCAAGACAGGGCTGGCTTAACATGGTTAAAGC
CCCTCTCTACTAAATAACAAAAGTAGTGTGGCTCATGGTGGGGGGACCTGTAAATCCA
GCTACTGGGGGGCTGGAGCTGAAATAACCTGGCTGTAGCAGGCGGGGGTTGGCTAGCTGA
GGCGAGATCTGGCTACTGAGCTACAGGAGTGGCTGTAGACCAAGAGCAGGGCTGTGGCTCAAA
BACAAAATGGTGGCTGGAGCTGGCTTCTTGGGGCTGTAGCTGGGGGGCTGTAGCT
TTTCTCTGAGTAGAGGTGGCTACCACTTCTGAGCTGTGGCTATTGTAGCTGTGTTAAGC

【図 1 0 - 4 1】

TGGTTTCCGGAGTCCTTTCAGCTGTTGGCTGAAACCATGAAATGTGTTCCACA
TGAATATGAAATGATTAGATGCTAATGGCAGAAAGTGTAATTCTTGTAAGAAA
CGAGGACATTTGGTCGTCAGCTAAGTGTAAATGCGTACTGTCTTCAAAATCT
GTGTTCAAAAGTACTTTCCTACTGCTAATTTCTGAAAGACTGAACTCCCA
ACATTTGGTCATGTATCAGTATCTCAACCCTGCTAACTTTTTTCTCCC
CTCTTTTCTTTGGCAAATACCTTTTGCTTATACCTTAAATAACCAT
TGTCCCTCACATGTCACCTTCCAAATTCAG[AAGAAGCAGGAAAGGCCACGAATAT]
[ACAAATTTAATGATTTCTTGGGGCAAGCAGGAGTGAGCACTAGGCCCCCTCCGGCT
[TGTACTCCACGGCCACCCCTGGCAGA]TGAGTGTCTGAAGGAATTCTTTTACCTGG
TCATGGTGTGTTAAAGGGTTAAACAAAACAAACAAACACAAAGCTTGTGAGA
CATGCCCTTCACTGGTCACCTTCCTGTCCCTACTGTATGGTATCTGTACTGGTA
TATCTATTGATGTGTTAATGCTATCTCAACCACGTTTAACTTCTCAAGCTGCCAGG
CACCGTGGCTAACGGCTGTAACTCCAGGTGGCTGGGAGGCCAGGGTCATGGATTACTT
GAAGTCAGGAGTCAGGAGCACGGCTGCACATGGTAAACCCAGTCTCTACTAAATA
ACAAAAAAATAGCCGGCGCATGGTGGCGCATGCCCTAACTCCAACTACTCAGGAGCTGG
CGAGAGGATCTGGCTGAACCCAGGAAACGGAGATCTGGCATGAGGCCAGATCTGCCTG
CACTCCAGCCTGGCGCATAGGTGACCTCTGTAAATAAAATAAAATAAAATA
TCCCTAAACTGAAAGGCTGACTGTCTAGTGTAGGATTATGGGATTAAAGTATATCA
AGTGGTGTCTCAGAAGAATCTAATTCTTCTTGTGGCTGGGGATGTGAAACAAA
GGAAAGTCATGTCCTTAAATGATGTTAAGGCTCTTSCAANAACTAAAGTAAATAAAT
GACCATATGTCAGGCCAGCATSTCTGCTACATTGGCCACAGCTCACAGAACTCT
ACTTGGGTGTTAAACCCAGGAGTCAGGAACTACAGCTCACAGGCCAGATGTGGGGCA
TGGCTGTACTATGGCTGTAAATGCTTTAAAGGGTTGAAACAAAAGAACACA
AAACAAAGACCUATACAAAACRAAGCOCGAGAAATAATGCGACAGAGACCATGTAG
GCATATAGAGCCTAAATACTGACTCTCAAGSCTTCCAGAAATCCTTCCGACTCTT
GTGAAACACCGTGTAGGAAACCATTTCTCAATTGGAGATATGCAAACTTGTAAAGT
ATTATTTCTTCTATATTCAGGAAAGTCAGTGTAGTAAACATTACCTTATTATTC
CTAAGAACCTCTCTGGCTGTGGCTGGCATGACTGGATTAAACCTTGCAGATTAAAT
ATCTATATAATGTTTGTGGCATATATGGGTTTAAACCTATGTCCTTCTT
TAACATTGAAATGAAATCTACAGCAGGGTATTTCTGTCAGGAACTGTCACCTTCTT
TATCTTTTCTTCTTCTCCAGTGGCTGCTGGCTGGAAANTTGATTTTAAATGTC
TGTGCAATTAATTTCGCAATGGCAAGTCAGGAACTGGCAGGGTCTGACTGCTTGGCAGACGTC
GCTCCCTCTGGCTGGCATGACTGGCATGTCAGTCACTGTCATGTTACCTGAGTGGCT
TGTCGACAGCAGCCCTGGCTGCTTCACTGGCTGGGGATGTGGCTCT
AGGATACACTGGCTCTGGAGCCCTCTACGCTGGCACACAGCTGCTGGGCTCTAC

【図 10 - 42】

【図 10 - 43】

(1 0 - 4 4)

CTTTGGGGGAAAAGTCCTTCCACGCCACAAAGAGTTTGATTTACAAAATATACTTC
CATTAAACATCTCGCTGTCTAACAGTCAGTGTCTTTACTCAGAATGGAAAAATCC
TGATCTTGGCTAGTTGTTAAAGATTTGATTTCAATTAGATTTCTCCCTGCAGGAGTC
TTTAAATTAATCATGCTTCTATGAACTTACGGCCAGAGTTTATCATGCTTCTAG
GTTCTACGGCTTATCATGTTGTTAAAGGTCAGTCCAGSAGTCTTCAGAGTTTCTAG
TTTACATAGAGATGGAACTCGGCCAGAGGAGCTGGATGATGGAGAAAGCTCGAGGAGTC
TTTAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCAGAG
CGCCAGGGCCAGAGGAGCTGAGTAATGGCTGAACTGAGGTTGTTGTTGAGGAGTCTGAG
AGACAGACATGGGCAACTCTACTCTGAGTGTGTTGTTGTTAATAATTCGACAATTTCTC
CATGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTC
ATTATCTGGCAATGAGTATGGACTTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTC
ACTTTTTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTC
ATGGCTCATCTCTCCACTGAAACCAGCTTCTTCAATGTGCTAAATGGGAGTCAGGA
GACAGGGGTGAGGACATTTCTGGGTTGCTGCTGCTGCTTAAAGGAGGAGGAGGAGTC
ARACACGGGCAAGGAGTCAACGATTATTTAGGAGTTTAACTTATGGGAGATTTAGA
ACTAATAACATCTTGTACTTTAACTTATGAGGGGAGTGTGGGAAAGGTTGCAATGTC
AGCTTGTGTTGAGCTCTAGTATCCATAGGAACTTACAGTACATGATACATAACCAAA
GCCAGTGGCAGCTTGTGTTGAGGGGAGTCAACATTTCACACATAGCAGC
ACATGGGCTCCCTGTGCTGAGGAGGCTTGTGTTGAGGATCTGTTATGGCCCTTAATTTGT
TGAATCTTTTCTTCFTCCCTCTGAAAGGAGTCTGACAAAATATAGTTATGTACTTC
ATCACTTAAATTTGTTGCTTCTTCACTATGGGAGTCTCACAAAGGAAACAAATCTG
ACAGTTGTTGTTAGGAGTGTGTTAATCTGCTTCAATTAACATAATTGCAAGGAGT
TCATCTCAAGGCTCAATCATGTAATTAATTTGCTGTTTATATGACTCAATGTT
GATTCTTTATTAATAAARGCTAATGGGAAAGGATCCCTGTAAAGCTGATGACTGAGC
CTACAATTAATTCTCTGAGTATATGAGGATGTGACAGAGTATTTAAAGATACTGTA
TATTAAATGTAATTAATCTATCCTTAAAGGAAACGGAGTCTTGTGAGGATCTCATTTGAT
GTGAAATCTGTTAAATGTSATAATATGCTGTTTATATACATTTGAGCTGCTTAAAGGAT
CACTTATATGCTTCTTGGCCACGGTATATGAGTGTCTTATTGCACTGTGTTCTTAAAG
ATGACATTAAGGAAAGAATCTTGTAGGAGAAACTACTGAGTGTAACTGCTTCC
CCTTATTTATGGGCTTAGGSCAACACTCTTGTCTGAGTGTGCAACAGCTTATATGAGC
AAAATGTTATTTGTAACCTGATAGGAAACTAAAGGAGGAGTATTTAGGAAATAACCA
TGTGCACTTGGCTTCTTCATCAGGAAATGCTCAGGAGAAAGGAGGAGTATGTGCT
ACTOCACATCTCAGCTGCAAGGAAATACTGAGCACTTGTAGGAGCTTCTATGAGTACT
GGAACTATGTTAAAGGCTTGTGTTCTGGAGAAATTCATGGTGTGTTCTGGAGAATA
GCCTTGGAGGAGTATGGGAGTCTATCTGGGAGAACATATGAGTACTCAGAGTAAAAGT

(図 10 - 45)

ATAGTCGTGAACTATAGAAGGAATGCTGGGCATGGTGGCGGCCCTGTAACTCAG
CTACTTGGAGCTGAGACGGCAGGATTCCGTGAAACCAGGAGCCCAGGACTTTAGACCA
GTCCTGGTAACTATAGTGAGACCCCTTTCACCIACTCTACTGATGCCCCCAAAATA
TATATGTCGCCGACCCGGCGCACACACACATACACACACACACACACACACACA
CAGAGGAAATTGTTAGAAAACACACAGAACTGAAATGTAAGTAGTATTAGGTGGGATAAG
AAGTAAAGGGATGGTAAGGGAGCTTGAGGAGGAGTAAATTATCTGATGGGACATCAG
CTC

10

20

30

40

【図 1 1】

MTSRWFHPIITGVAENLLITRGVDSFLARPSSKNPGDFTLSVRNGAVTHIKIQTNG
 DYYDLYGGGEKFATLAEIYQYMEHHGQLKEKNGDVTELYKPLNCAADPTSERFHGLSGK
 EAFLLLETEKGKHGSFLIVESQSHPGDFVLSSVRTGDDKGESNDGSKVTHMRCQELKYD
 VGGGERFDSLTDLVEHYKKNPVMVETLGTVLQLQPLNTTRINAEEISRVRELSKLAETT
 DKVKQGFWEETLQQECKLISRKEGORQENKNRKYKNILPFDHTRYVHDGDPNEP
 VSDYINANIIMPEFETRCNNSPKKSYIATQGCLONTVNDFWRMVFQENSRTVUMTTKEV
 ERGKSCKCVKYWPDEYALEKEYGYMRVRNVESSAHHYTLRELKLSKVGOGNTERTVWOYHF
 FTWPDHGVPSPDPGGVLDLDEEVHKQESIMDAGPVVHC~~SAGIGRTGTFI~~IVLDLIDIR
 EKGDIDIVPKTIQWYFSORSGMVQTEAQYRFYMAVQHYIETLQRRIEEQSKRKGH
 EYTNIKYSLADQTSGDQSPLPPCTPTPPCA~~EMRE~~DSARVYENVGLMQQKSF

配列番号 6 1

10

20

【配列表】

0007278027000001.app

30

40

50

フロントページの続き

弁理士 廣田 鉄平
(74)代理人 100096482
弁理士 東海 裕作
(74)代理人 100131093
弁理士 堀内 真
(74)代理人 100150902
弁理士 山内 正子
(74)代理人 100141391
弁理士 園元 修一
(74)代理人 100221958
弁理士 篠田 真希恵
(74)代理人 100192441
弁理士 渡辺 仁
(72)発明者 シャレイ アーモン アール .
アメリカ国 マサチューセッツ 02144 サマービル ハイランドアベニュー 333
(72)発明者 ラジョイー マーク
アメリカ国 ワシントン シアトル マイナーアベニューイー 2331
(72)発明者 ジエンセン クラウス エフ .
アメリカ国 マサチューセッツ 02421 レキシントン アウトルックドライブ 103
(72)発明者 ランガー ロバート エス .
アメリカ国 マサチューセッツ 02459 ニュートン モントベルロード 98
合議体
審判長 長井 啓子
審判官 福井 悟
審判官 川合 理恵
(56)参考文献 特表 2014 - 533936 (JP, A)
特表 2014 - 509195 (JP, A)
特表 2012 - 531909 - (JP, A)
国際公開第 2014 / 065596 (WO, A1)
米国特許第 6133503 (US, A)
Genome Biology (2015) Vol. 16, No. 251, pp. 1 - 3
(Published online 2015.11.17)
(58)調査した分野 (Int.Cl., DB名)
C12N15/00
Capillus / MEDLINE / EMBASE / BIOSIS / WPIDS / WPIX (STN)