

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2018-513681

(P2018-513681A)

(43) 公表日 平成30年5月31日(2018.5.31)

(51) Int.Cl.	F I	テーマコード (参考)
<b>C 1 2 N 15/09 (2006.01)</b>	C 1 2 N 15/00 Z N A A	4 B 0 5 0
<b>C 0 7 K 19/00 (2006.01)</b>	C 0 7 K 19/00	4 B 0 6 5
<b>C 0 7 K 14/005 (2006.01)</b>	C 0 7 K 14/005	4 H 0 4 5
<b>C 1 2 N 1/15 (2006.01)</b>	C 1 2 N 1/15	
<b>C 1 2 N 1/19 (2006.01)</b>	C 1 2 N 1/19	
審査請求 未請求 予備審査請求 未請求 (全 140 頁) 最終頁に続く		

(21) 出願番号	特願2017-552079 (P2017-552079)	(71) 出願人	517341397
(86) (22) 出願日	平成28年3月31日 (2016. 3. 31)		エクセリゲン サイエнтиフィック、イ
(85) 翻訳文提出日	平成29年11月22日 (2017. 11. 22)		ンコーポレイテッド
(86) 国際出願番号	PCT/US2016/025426		Exeligen Scientific
(87) 国際公開番号	W02016/161207		, Inc.
(87) 国際公開日	平成28年10月6日 (2016. 10. 6)		アメリカ合衆国 カリフォルニア州 9 2
(31) 優先権主張番号	62/210, 451		0 0 8, カールスバッド, ファラデアヴ
(32) 優先日	平成27年8月27日 (2015. 8. 27)		ェニュー 2 3 5 1
(33) 優先権主張国	米国 (US)	(74) 代理人	110001302
(31) 優先権主張番号	62/140, 454		特許業務法人北青山インターナショナル
(32) 優先日	平成27年3月31日 (2015. 3. 31)	(72) 発明者	シャイフ, フェルフ
(33) 優先権主張国	米国 (US)		アメリカ合衆国 カリフォルニア州 9 1
(31) 優先権主張番号	62/240, 359		3 6 2, ウェストレイクヴィレッジ, レイ
(32) 優先日	平成27年10月12日 (2015. 10. 12)		クビューキャニオンロード 5 5 0 5
(33) 優先権主張国	米国 (US)		
最終頁に続く			

(54) 【発明の名称】 細胞または生物のゲノムへのDNA配列の標的化組み込みのためのCas9レトロウイルスインテグラーゼおよびCas9レコンビナーゼ系

## (57) 【要約】

本開示は、関心のあるDNA配列（または関心のある遺伝子）を細胞または生物のゲノム中の標的化部位に送達するためにウイルスインテグラーゼ、レコンビナーゼまたはトランスポサナーゼと連結される、Cas9、Cpf1、TALEおよびジンクフィンガータンパク質などの改変タンパク質の使用に関する。DNA切断でのその機能について不活性であるCas9の使用は、相同組み換えのための他の系で意図されるようなDNA破壊を引き起こすことなく、RNAガイドの使用によるCas9タンパク質のDNA標的化能の使用を可能にする。ウイルスインテグラーゼまたはレコンビナーゼに連結されたジンクフィンガータンパク質またはTALE（DNAの特異的な配列に結合する改変タンパク質）の使用も開示される。この系は、研究および治療目的で使用され得る。例えば、関心のある遺伝子は、細胞中で正常な遺伝子産物（例えば、遺伝子産物はタンパク質または特殊化されたRNAであり得る）を回収するために、遺伝子がその遺伝子産物の産生能を欠いている細胞中に含まれ得る。

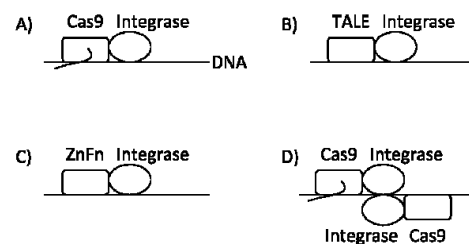


FIG. 1

**【特許請求の範囲】****【請求項 1】**

操作可能な連結において、

a) C a s 9、不活性 C a s 9 もしくは C p f 1 またはその一部をコードする第 1 のポリヌクレオチド配列と、

b) インテグラーゼ、レコンビナーゼもしくはトランスポサナーゼまたはその一部をコードする第 2 のポリヌクレオチド配列と、

c) 核酸リンカーをコードする第 3 のポリヌクレオチド配列とを含む核酸コンストラクトにおいて、

前記第 1 のポリヌクレオチド配列が 5' および 3' 末端を含み、かつ前記第 2 のポリヌクレオチド配列が 5' および 3' 末端を含み、および前記第 1 のポリヌクレオチドの前記 3' 末端が前記核酸リンカーによって前記第 2 のポリヌクレオチドの前記 5' 末端に連結され、かつ前記第 1 および第 2 のポリヌクレオチドが細胞または生物において融合タンパク質として発現されることが可能であることを特徴とする、核酸コンストラクト。

10

**【請求項 2】**

請求項 1 に記載の核酸コンストラクトにおいて、前記第 1 のポリヌクレオチド配列が配列番号 1、3、5、7、9、11、13、27~46、49、56 もしくは 68 の何れか 1 つ、またはそれと少なくとも 80%、少なくとも 85%、少なくとも 90%、少なくとも 95% もしくは少なくとも 99% の同一性を有する配列を含むことを特徴とする、核酸コンストラクト。

20

**【請求項 3】**

請求項 1 に記載の核酸コンストラクトにおいて、前記 C a s 9、不活性 C a s 9 または C p f 1 が配列番号 2、4、6、8、10、12、14、50、52、69、72~78 もしくは 86~92 の何れか 1 つ、またはそれと少なくとも 80%、少なくとも 85%、少なくとも 90%、少なくとも 95% もしくは少なくとも 99% の同一性を有する配列を含むことを特徴とする、核酸コンストラクト。

**【請求項 4】**

請求項 1 に記載の核酸コンストラクトにおいて、前記第 2 のポリヌクレオチド配列が配列番号 15、17、19、21、23、47、55、62、64、66、70 もしくは 79 の何れか 1 つ、またはそれと少なくとも 80%、少なくとも 85%、少なくとも 90%、少なくとも 95% もしくは少なくとも 99% の同一性を有する配列を含むことを特徴とする、核酸コンストラクト。

30

**【請求項 5】**

請求項 1 に記載の核酸コンストラクトにおいて、前記インテグラーゼ、レコンビナーゼまたはトランスポサナーゼが配列番号 16、18、20、22、24、25、26、48、63、65、67、71 もしくは 80 の何れか 1 つ、またはそれと少なくとも 80%、少なくとも 85%、少なくとも 90%、少なくとも 95% もしくは少なくとも 99% の同一性を有する配列を含むことを特徴とする、核酸コンストラクト。

**【請求項 6】**

生物において、請求項 1 に記載の核酸コンストラクトを含むことを特徴とする、生物。

40

**【請求項 7】**

請求項 1 に記載の融合タンパク質を含む生物において、修飾されたゲノムを有することを特徴とする、生物。

**【請求項 8】**

a) C a s 9、不活性 C a s 9 もしくは C p f 1 またはその一部をコードする第 1 のポリヌクレオチド配列と、

b) インテグラーゼ、レコンビナーゼもしくはトランスポサナーゼまたはその一部をコードする第 2 のポリヌクレオチド配列と、

c) 核酸リンカーをコードする第 3 のポリヌクレオチド配列とを含む生物において、

50

前記第 1 のポリヌクレオチド配列が 5' および 3' 末端を含み、かつ前記第 2 のポリヌクレオチド配列が 5' および 3' 末端を含み、および前記第 1 のポリヌクレオチドの前記 3' 末端が前記核酸リンカーによって前記第 2 のポリヌクレオチドの前記 5' 末端に連結され、かつ前記第 1 および第 2 のポリヌクレオチドが細胞または生物において融合タンパク質として発現されることが可能であることを特徴とする、生物。

【請求項 9】

融合タンパク質において、

a) 触媒能のない Cas9、Cas9、TALE タンパク質、ジンクフィンガータンパク質または Cpf1 タンパク質である第 1 のタンパク質であって、標的 DNA 配列を標的とする第 1 のタンパク質と、

b) インテグラーゼ、レコンビナーゼまたはトランスポサラーゼである第 2 のタンパク質と、

c) 前記第 1 のタンパク質を前記第 2 のタンパク質に連結するリンカーを含むことを特徴とする、融合タンパク質。

【請求項 10】

請求項 9 に記載の融合タンパク質において、前記第 2 のタンパク質がインテグラーゼであることを特徴とする、融合タンパク質。

【請求項 11】

請求項 10 に記載の融合タンパク質において、前記インテグラーゼが HIV-1 インテグラーゼまたはレンチウイルスインテグラーゼであることを特徴とする、融合タンパク質。

【請求項 12】

請求項 9 に記載の融合タンパク質において、前記リンカー配列が 1 以上のアミノ酸長であることを特徴とする、融合タンパク質。

【請求項 13】

請求項 9 に記載の融合タンパク質において、前記第 1 のタンパク質が触媒能のない Cas9 であることを特徴とする、融合タンパク質。

【請求項 14】

請求項 9 に記載の融合タンパク質において、前記リンカー配列が 4 ~ 8 アミノ酸長であることを特徴とする、融合タンパク質。

【請求項 15】

請求項 14 に記載の融合タンパク質において、前記第 1 のタンパク質が TALE タンパク質であることを特徴とする、融合タンパク質。

【請求項 16】

請求項 9 に記載の融合タンパク質において、前記第 1 のタンパク質がジンクフィンガータンパク質であることを特徴とする、融合タンパク質。

【請求項 17】

請求項 15 または 16 に記載の融合タンパク質において、前記標的 DNA 配列が約 16 ~ 約 24 塩基対長であることを特徴とする、融合タンパク質。

【請求項 18】

請求項 9 に記載の融合タンパク質において、前記第 1 のタンパク質が Cas9 または触媒能のない Cas9 であり、1 つ以上のガイド RNA が約 16 ~ 約 24 塩基対の標的 DNA 配列の標的化のために使用されることを特徴とする、融合タンパク質。

【請求項 19】

DNA 配列をゲノム DNA に挿入する方法において、

a) 前記ゲノム DNA において標的配列を同定することと、

b) 前記ゲノム DNA において前記標的配列に結合するように請求項 1 に記載の融合タンパク質を設計することと、

3) 前記ゲノム DNA に組み込むための関心のある DNA 配列を選択することと、

d) 細胞または生物に対して、前記細胞または生物への前記融合タンパク質および関心のある DNA 配列の移行を可能にする技術によって前記融合タンパク質および前記関心の

10

20

30

40

50

あるDNA配列を提供することと

を含み、前記関心のあるDNA配列が前記ゲノムDNA中の前記標的配列において組み込まれるようになることを特徴とする、方法。

【請求項20】

a) 標的DNA配列に結合するように改変されている、Cas9、触媒能のないCas9、TALEタンパク質、ジンクフィンガータンパク質またはCpf1タンパク質である第1のタンパク質に対する第1のコード配列と、

b) インテグラーゼ、レコンビナーゼまたはトランスポサラーゼである第2のタンパク質に対する第2のコード配列と、

c) 前記第1のタンパク質と前記第2のタンパク質との間でアミノ酸リンカーを形成する、前記第1のコード配列と前記第2のコード配列との間のDNA配列と、

d) 任意選択により、インテグラーゼによって認識されるatt部位によって取り囲まれる発現された関心のあるDNA配列および任意選択により1つ以上のガイドRNAとを含むヌクレオチドベクターにおいて、前記第1のタンパク質が、決定されたDNA配列を標的とし、前記第1のタンパク質が前記アミノ酸リンカー配列によって前記第2のタンパク質に連結されることを特徴とする、ヌクレオチドベクター。

【請求項21】

細胞または生物において遺伝子転写を阻害する方法において、

a) 遺伝子においてATG開始コドンと同定することと、

b) 前記遺伝子の前記ATG開始コドンの直後の標的配列に結合するように、請求項1に記載の融合タンパク質を用いて融合タンパク質系を設計することと、

c) 1つ以上の連続停止コドンである関心のあるDNA配列を設計することと、

d) 細胞または生物に対して、前記細胞または生物への前記融合タンパク質および関心のあるDNA配列の移行を可能にする技術によって前記融合タンパク質および前記関心のあるDNA配列を提供することと

を含み、前記関心のあるDNA配列がゲノムDNA中の前記標的配列において組み込まれるようになり、前記遺伝子の転写が阻害されることを特徴とする、方法。

【請求項22】

請求項9に記載の融合タンパク質において、前記第2のタンパク質がレコンビナーゼであることを特徴とする、融合タンパク質。

【請求項23】

請求項22に記載の融合タンパク質において、前記レコンビナーゼがCreレコンビナーゼまたはその修飾型であり、前記修飾型のCreレコンビナーゼが構成的なレコンビナーゼ活性を有することを特徴とする、融合タンパク質。

【請求項24】

請求項20に記載のベクターにおいて、細胞中で発現される逆転写酵素遺伝子をさらに含むことを特徴とする、ベクター。

【請求項25】

DNA結合タンパク質/インテグラーゼ融合物の精製タンパク質と、約15～約100塩基対長のRNAとを含む組成物において、前記DNA結合タンパク質が、ゲノム中の標的化DNA配列に対して改変されているCas9、Cpf1、TALENおよびジンクフィンガータンパク質から選択され、前記インテグラーゼがHIVインテグラーゼ、レンチウイルスインテグラーゼ、アデノウイルスインテグラーゼ、レトロウイルスインテグラーゼまたはMMTVインテグラーゼであることを特徴とする、組成物。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

本願は、全ての目的のためにそれぞれの内容全体が参照により組み込まれる、2015年3月31日提出の米国仮特許出願第62,140,454号明細書、2015年8月2

10

20

30

40

50



7日提出の米国仮特許出願第62,210,451号明細書および2015年10月12日提出の米国仮特許出願第62,240,359号明細書の利益を主張する。

【0002】

本開示は、関心のあるDNA配列（または関心のある遺伝子）を細胞または生物のゲノム中の標的化部位に送達するためにウイルスインテグラーゼ（例えば、HIVまたはMMTVインテグラーゼ）またはレコンビナーゼとリンカーにより連結される、Cas9（CRISPR（クラスター化された規則的に間隔を置いて出現する短いパリンドロームリピート）タンパク質）、TALEおよびジンクフィンガータンパク質などのゲノム特異性を示すDNA結合タンパク質を有する改変タンパク質の使用に関する。DNA切断でのその機能について不活性であるCas9の使用は、相同組み換えのための他の系で意図されるようなDNA破壊を引き起こすことなく、RNAガイド（gRNA）の使用によるCas9タンパク質のDNA標的化能の使用を可能にする。ウイルスインテグラーゼまたはレコンビナーゼに連結されたジンクフィンガータンパク質またはTALE（DNAの特異的な配列に結合する改変タンパク質）の使用も開示される。この系は、研究および治療目的のために使用され得る。例えば、従来の方を通じたオフターゲット切断の可能性なく、関心のある遺伝子を含むドナーDNAを宿主ゲノムに容易に導入し得る。「ノックアウト」戦略を促進するためにもドナーDNAを改変し得る。Cas9標的化の特異性を改善するための新しい戦略も論じる。この戦略は、何れのガイドRNAがCas9の特異的な標的化をもたらすかを見出すためのアッセイにおいて、ガイドRNAおよびゲノムDNAとともに表面結合dCas9（そのDNA切断能について不活性であるCas9）を使用する。これは、特に、CRISPR/Cas9のインビボ適用において重要であり、現在のイン・シリコ予測モデルの制限を克服するが、アッセイにおいて何れのgRNAを使用するかということを知識に基づいて決定するため、イン・シリコ予測モデルと組み合わせて使用することもできる。

10

20

【背景技術】

【0003】

ゲノム配列決定技術および分析方法における最近の前進により、多岐にわたる生物学的機能および疾患に関連する遺伝学的/ゲノム因子を分類し、マッピングする能力が顕著に加速された。正確なゲノム標的化技術は、個々の遺伝要素の選択的攪乱を可能にし、かつ合成生物学、バイオテクノロジーおよび医学的応用を進展させることにより、原因となる遺伝的変異の体系的な逆改変を可能にするために必要とされる。標的化ゲノム攪乱を生じさせるためにデザイナージンクフィンガー、転写活性化因子様エフェクター（TALE）、CRISPR/Cas9またはメガヌクレアーゼなどのゲノム編集技術が利用可能であり、所与のゲノムにおける特異的な位置へのDNA配列（全長遺伝子配列を含む）の組み込みを可能にする新しいゲノム改変技術が依然として必要とされている。これにより、改変遺伝子を発現する細胞株またはトランスジェニック生物の作製またはそれを必要とする対象における機能障害性遺伝子の置き換えが可能になるであろう。

30

【0004】

インテグラーゼは、宿主ゲノム（哺乳動物、ヒト、マウス、ラット、サル、カエル、魚、植物（作物およびシロイヌナズナ（*Arabidopsis*）のような実験用植物を含む）、研究用または医学生物学的な細胞株または初代細胞培養物、*C. elegans*）、ハエ（ショウジョウバエ（*Drosophila*）など）へのウイルス核酸の挿入を可能にするウイルスタンパク質である。インテグラーゼは、ウイルス核酸配列を宿主ゲノムに組み込むために、宿主のDNA結合タンパク質を使用してインテグラーゼを宿主ゲノムと結び付ける。インテグラーゼは、HIV（ヒト免疫不全ウイルス）などのレトロウイルス中で見られる。インテグラーゼは、そのゲノムを宿主DNAに挿入するためのウイルス遺伝子上の配列に依存する。Leavitt et al (*Journal of Biological Chemistry*, 1993, volume 268, pages 2113 - 2119) は、部位特異的な突然変異誘発およびインビトロ研究を使用することにより、HIV1インテグラーゼの機能を調べた。Leavittは

40

50

また、ウイルスインテグラーゼによる宿主ゲノムへのH I V 1 D N A (逆転写後に生成)の組み込みに重要であるU 5 およびU 3 H I V 1 a t t 部位の配列も示す。

#### 【0005】

本開示は、所望の核酸(D N A)配列をゲノムの指定位置でゲノムに特異的に挿入することを可能にすることにより、現在のゲノム編集技術を改善する。D N A 結合能を有する組み換え改変インテグラーゼ(またはレコンビナーゼ)は、ゲノム中の所与のD N A 配列に結合し、インテグラーゼ認識ドメイン(H I V 1 (または他のレトロウイルス) a t t 部位など)を有する提供されるD N A 配列および/または相同アームを認識して、部位特異的に所与の核酸配列をゲノムに挿入する。本開示のある態様は、遺伝子の転写開始部位の直後に停止コドン(U A A、U A G および/またはU G A)のD N A 配列を挿入することを含む。これにより、細胞または生物のゲノム中の遺伝子転写の効果的な阻害が可能になる。

10

#### 【発明の概要】

#### 【0006】

本開示は、D N A 標的化インテグラーゼを形成させるために、ジンクフィンガータンパク質、T A L E N およびC R I S P R / C a s 9 または他のC R I S P R タンパク質様C p f 1 などを含むD N A 標的化技術をレトロウイルスインテグラーゼと連結する。次に、関心のある遺伝子(G O I)がD N A 標的化インテグラーゼとともに提供され得、それが標的化形式でゲノムに組み込まれ得るようになる。ゲノムにおけるその挿入に対する別の特異性レベルを提供するために相同アームを用いてG O I を設計する。

20

#### 【0007】

本開示は、特に、レトロウイルスインテグラーゼとの連結のための、D N A 切断能がない変異体C a s 9 の使用に関する。

#### 【0008】

本開示は、A)例えば、D N A 切断能がないC a s タンパク質(例えば、C a s 9)に共有結合されるウイルスインテグラーゼ(または細菌レコンビナーゼ)を含む系を含む。あるいは、ウイルスインテグラーゼ(またはレコンビナーゼ)は、T A L E タンパク質またはジンクフィンガータンパク質に共有結合されるが、この場合、これらのタンパク質はゲノム中のD N A の特異的配列を標的とするように設計されている。これは、発現ベクターにおいてまたは精製タンパク質として提供され得る; B)所望のゲノムに組み込まれる相同アームがあるかまたはない関心のある遺伝子(または関心のあるD N A 配列)。G O I または関心のあるD N A 配列は、必要に応じて、ウイルスインテグラーゼによって認識されるように修飾され得る。細胞へのポリヌクレオチド形質移入および/またはタンパク質導入に必要とされる他の試薬。D N A 配列のオフターゲット組み込みに対してアッセイを行う。ある態様において、挿入されるD N A 配列に改変マーカー配列を使用する。

30

#### 【0009】

操作可能な連結において、a)C a s 9、不活性C a s 9もしくはC p f 1 またはその一部をコードする第1のポリヌクレオチド配列と、b)インテグラーゼ、レコンビナーゼもしくはトランスポサゼまたはその一部をコードする第2のポリヌクレオチド配列と、c)核酸リンカーをコードする第3のポリヌクレオチド配列とを含み、第1のポリヌクレオチド配列が5'および3'末端を含み、かつ第2のポリヌクレオチド配列が5'および3'末端を含み、および第1のポリヌクレオチドの3'末端が核酸リンカーによって第2のポリヌクレオチドの5'末端に連結され、かつ第1および第2のポリヌクレオチドが細胞または生物において融合タンパク質として発現されることが可能である、核酸コンストラクトが本明細書中で提供される。いくつかの実施形態において、第1のポリヌクレオチド配列は、配列番号1、3、5、7、9、11、13、27~46、49、56もしくは68の何れか1つ、またはそれと少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%もしくは少なくとも99%の同一性を有する配列を含む。いくつかの実施形態において、C a s 9、不活性C a s 9またはC p f 1 は、配列番号2、4、6、8、10、12、14、50、52、69、72~78もしくは86~92の何れか1

40

50

つ、またはそれと少なくとも 80%、少なくとも 85%、少なくとも 90%、少なくとも 95% もしくは少なくとも 99% の同一性を有する配列を含む。いくつかの実施形態において、第 2 のポリヌクレオチド配列は、配列番号 15、17、19、21、23、47、55、62、64、66、70 もしくは 79 の何れか 1 つ、またはそれと少なくとも 80%、少なくとも 85%、少なくとも 90%、少なくとも 95% しくは少なくとも 99% の同一性を有する配列を含む。いくつかの実施形態において、インテグラーゼ、レコンビナーゼまたはトランスポサナーゼは、配列番号 16、18、20、22、24、25、26、48、63、65、67、71 しくは 80 の何れか 1 つ、またはそれと少なくとも 80%、少なくとも 85%、少なくとも 90%、少なくとも 95% しくは少なくとも 99% の同一性を有する配列を含む。また、本核酸コンストラクトを含む生物も本明細書中で記載される。また、修飾ゲノムを有する、融合タンパク質を含む生物も本明細書中で記載される。

10

20

30

40

50

#### 【0010】

a) Cas9、不活性 Cas9 しくは Cpfl またはその一部をコードする第 1 のポリヌクレオチド配列と、b) インテグラーゼ、レコンビナーゼもしくはトランスポサナーゼまたはその一部をコードする第 2 のポリヌクレオチド配列と、c) 核酸リンカーをコードする第 3 のポリヌクレオチド配列とを含む生物が本明細書中で提供され、ここで、この第 1 のポリヌクレオチド配列は 5' および 3' 末端を含み、かつ第 2 のポリヌクレオチド配列は 5' および 3' 末端を含み、および第 1 のポリヌクレオチドの 3' 末端は核酸リンカーによって第 2 のポリヌクレオチドの 5' 末端に連結され、かつこの第 1 および第 2 のポリヌクレオチドは細胞または生物において融合タンパク質として発現されることが可能である。

#### 【0011】

また、a) 触媒能のない Cas9、Cas9、TALE タンパク質、ジンクフィンガータンパク質または Cpfl タンパク質である第 1 のタンパク質であって、標的 DNA 配列を標的とする第 1 のタンパク質と、b) インテグラーゼ、レコンビナーゼまたはトランスポサナーゼである第 2 のタンパク質と、c) 第 1 のタンパク質を第 2 のタンパク質に連結するリンカーとを含む融合タンパク質も本明細書中で提供される。いくつかの実施形態において、第 2 のタンパク質はインテグラーゼであり、このインテグラーゼは HIV-1 インテグラーゼまたはレンチウイルスインテグラーゼであり、リンカー配列はアミノ酸長が 1 以上であるか、または第 1 のタンパク質は触媒能のない Cas9 である。いくつかの実施形態において、リンカー配列は、4 ~ 8 アミノ酸長であり、第 1 のタンパク質は TALE タンパク質であるか、または第 1 のタンパク質はジンクフィンガータンパク質である。いくつかの実施形態において、本融合タンパク質が TALE またはジンクフィンガータンパク質を含む場合、標的 DNA 配列は約 16 ~ 約 24 塩基対長である。いくつかの実施形態において、第 1 のタンパク質は Cas9 または触媒能のない Cas9 であり、約 16 ~ 約 24 塩基対の標的 DNA 配列を標的とするために 1 つ以上のガイド RNA が使用される。

#### 【0012】

また、a) ゲノム DNA において標的配列を同定することと、b) ゲノム DNA 中の標的配列を結合するように請求項 1 に記載の融合タンパク質を設計することと、3) ゲノム DNA に組み込むための関心のある DNA 配列を設計することと、d) 細胞または生物に対して、細胞または生物への融合タンパク質および関心のある DNA 配列の移行を可能にする技術によって融合タンパク質および関心のある DNA 配列を提供することとを含み、関心のある DNA 配列がゲノム DNA 中の標的配列で組み込まれるようになる、ゲノム DNA に DNA 配列を挿入する方法も本明細書中で提供される。

#### 【0013】

また、a) 標的 DNA 配列に結合するように改変されている、Cas9、触媒能のない Cas9、TALE タンパク質、ジンクフィンガータンパク質または Cpfl タンパク質である第 1 のタンパク質に対する第 1 のコード配列と、b) インテグラーゼ、レコンビナーゼまたはトランスポサナーゼである第 2 のタンパク質に対する第 2 のコード配列と、c)

第1のタンパク質と第2のタンパク質との間でアミノ酸リンカーを形成する、第1のコード配列と第2のコード配列との間のDNA配列と、d)任意選択により、インテグラーゼによって認識されるatt部位によって取り囲まれる発現された関心のあるDNA配列および任意選択により1つ以上のガイドRNAとを含み、第1のタンパク質が、所定のDNA配列を標的とし、第1のタンパク質がアミノ酸リンカー配列によって第2のタンパク質に連結される、ヌクレオチドベクターも本明細書中で提供される。

#### 【0014】

細胞または生物において遺伝子転写を阻害する方法が本明細書中で提供され、この方法は、a)遺伝子中のATG開始コドンと同定することと、b)遺伝子のATG開始コドンの直後の標的配列に結合するように、請求項1に記載の融合タンパク質を用いて融合タンパク質系を設計することと、c)1つ以上の連続停止コドンである関心のあるDNA配列を設計することと、d)細胞または生物に対して、細胞または生物への融合タンパク質および関心のあるDNA配列の移行を可能にする技術によって融合タンパク質および関心のあるDNA配列を提供することとを含み、関心のあるDNA配列がゲノムDNA中の標的配列において組み込まれるようになり、遺伝子の転写が阻害される。いくつかの実施形態において、第2のタンパク質はレコンビナーゼであり、レコンビナーゼはCreレコンビナーゼまたはその修飾型であり、この修飾型Creレコンビナーゼは構成的なレコンビナーゼ活性を有する。ある実施形態において、細胞で発現される逆転写酵素遺伝子をさらに含むベクターである。

10

#### 【0015】

また、DNA結合タンパク質/インテグラーゼ融合物の精製タンパク質と、約15~約100塩基対長のRNAとを含む組成物も本明細書中で提供され、この場合、DNA結合タンパク質は、ゲノム中の標的化DNA配列に対して改変されているCas9、Cpf1、TALENおよびジンクフィンガータンパク質から選択され、このインテグラーゼはHIVインテグラーゼ、レンチウイルスインテグラーゼ、アデノウイルスインテグラーゼ、レトロウイルスインテグラーゼまたはMMTVインテグラーゼである。

20

#### 【図面の簡単な説明】

#### 【0016】

以下の説明、添付の特許請求の範囲および付随する図面に関して、本開示のこれらおよび他の特性、態様および長所がより詳細に理解される。

30

#### 【0017】

【図1】図1は、a)代表的な触媒能のないCas9/HIV1インテグラーゼ融合タンパク質、b)代表的なTALE/HIV1インテグラーゼ融合タンパク質、c)代表的なジンクフィンガータンパク質/HIV1インテグラーゼ融合タンパク質およびd)標的化部位でDNAの逆側に設計される代表的なCas9/HIV1インテグラーゼ融合タンパク質を示す。各融合タンパク質はDNAの特異的な標的配列に結合する。「ZnFn」は、ジンクフィンガータンパク質である。「インテグラーゼ」は、1インテグラーゼ単位または例えば短いアミノ酸リンカーにより連結される2インテグラーゼ単位に相当する。いくつかの実施形態において、インテグラーゼをレコンビナーゼに置き換え得る。Cas9は、触媒能があるかまたはないものであり得る。

40

【図2】図2は、触媒能のないCas9/インテグラーゼ融合タンパク質を含むベクター、関心のあるDNA配列を含むベクターおよび逆転写酵素を含むベクターを含むDNAプラスミド系を示す。ガイドRNA(gRNA)またはRNAは別個に提供され得る。gRNAを発現させるために別のベクターを使用し得る。「1または2」とは、1つのインテグラーゼまたは例えばアミノ酸リンカーにより連結される2つのインテグラーゼを指す。

【図3】図3は、ヌクレイオチド(nucleotide)配列、触媒能のないCas9/インテグラーゼ融合タンパク質、ガイドRNA、関心のあるDNA(遺伝子)配列および逆転写酵素を含む代表的なDNAプラスミドを示す。関心のあるDNA配列にウイルスatt部位を提供し得、それにより、細胞のゲノムDNAへのインテグラーゼの組み込みが可能になる。ガイドRNA(gRNA)またはRNAを別個に提供し得る。gRNA

50

を発現させるために別のベクターを使用し得る。「1または2」とは、1つのインテグラーゼまたは例えばアミノ酸リンカーにより連結される2つのインテグラーゼを指す。

【図4】図4は、流れ図を示す。図2および図3で示されるベクターを使用する、ある代表的な方法を図4で示し、これは以下のとおりである：1) 逆転写酵素は、ベクターから発現されるatt部位がある関心のあるDNA配列を逆転写し(あるいはatt部位がある直線状DNAを使用する)、2) 融合Cas9/インテグラーゼは、ガイドRNAに基づくゲノムDNA上の部位を標的とし、3) インテグラーゼは、関心のあるDNA配列上のatt(LTR)部位を認識し、標的化部位でゲノムにDNAを組み込み、4) 関心のあるDNA配列の正確な挿入について調べるためにアッセイ(例えば、PCR(ポリメラーゼ連鎖反応))を行う。非特異的な組み込みについて調べるためにアッセイを行い得る。

【図5】図5は、ガイドNrF2-sgRNA2およびsgRNA3を用いた、NrF2のエクソン2を標的とするAbbie1遺伝子編集を示す。

【図6】図6は、Abbie1遺伝子編集により生成された理論的データを示す。

【図7】図7は、ガイドNrF2-sgRNA3を使用した、NrF2のエクソン2を標的とするAbbie1遺伝子編集を示す。

【図8】図8は、プールHek293T-細胞におけるNrF2のAbbie1ノックアウトを示す。

【図9】図9は、プールHek293T-細胞におけるNrF2のAbbie1ノックアウトを示す。

【図10】図10は、CXCR4エクソン2を標的とするAbbie1遺伝子編集を示す。

【図11】図11は、大腸菌(E. coli)クーマシー染色ゲルからの単離および精製後のABBE1タンパク質の検出を示す。

【発明を実施するための形態】

【0018】

当業者の本開示の実施を支援するために以下の詳細な説明を提供する。それにもかかわらず、本発見の趣旨または範囲から逸脱することなく、本明細書中で論じる実施形態における修正形態および変形形態が当業者によりなされ得るため、本開示を過度に限定するものとしてこの詳細な説明を解釈すべきではない。

【0019】

本開示および添付の特許請求の範囲中で使用される場合、単数形「1つの(a)」、「1つの(an)」および「その(the)」は、内容から別段の明らかな指示がない限り、複数への言及を含む。本開示および添付の特許請求の範囲中で使用される場合、「または」という語は、単数であってもよく、または包括的であってもよい。例えば、AまたはBはAおよびBであり得る。

【0020】

内因性

本明細書中で記載される場合、内因性核酸、ヌクレオチド、ポリペプチドまたはタンパク質は、宿主生物との関連において定められる。内因性核酸、ヌクレオチド、ポリペプチドまたはタンパク質は、宿主生物において天然に生じるものである。

【0021】

外来性

本明細書中で記載される場合、外来性核酸、ヌクレオチド、ポリペプチドまたはタンパク質は、宿主生物との関連において定められる。外来性核酸、ヌクレオチド、ポリペプチドまたはタンパク質は、宿主生物において天然に生じないか、または宿主生物において異なる位置にあるものである。

【0022】

ノックアウト

(例えば、ランダム挿入または相同組み換えによって)外来性核酸が宿主生物中に形質転換され、その結果、(例えば、欠失、挿入によって)遺伝子が破壊される場合、遺伝子

10

20

30

40

50

がロックアウトされたとみなされる。

【0023】

遺伝子をロックアウトすると、対応するタンパク質の活性が低下し得る。例えば、遺伝子がロックアウトされていない同じタンパク質の活性と比較して、少なくとも10%、少なくとも20%、少なくとも30%、少なくとも40%、少なくとも50%、少なくとも60%、少なくとも70%、少なくとも80%、少なくとも90%または100%である。

【0024】

遺伝子をロックアウトすると、遺伝子がロックアウトされていない場合と比較して、少なくとも20%、少なくとも30%、少なくとも40%、少なくとも50%、少なくとも60%、少なくとも70%、少なくとも80%、少なくとも90%または100%だけ遺伝子の転写が低下し得る。

10

【0025】

修飾

修飾された生物は、非修飾生物と異なる生物である。例えば、修飾された生物は、標的遺伝子配列のロックアウトを引き起こす本開示の融合タンパク質を含み得る。修飾された生物は、修飾されたゲノムを有し得る。

【0026】

修飾された核酸配列またはアミノ酸配列は、非修飾核酸配列またはアミノ酸配列と異なる。例えば、核酸配列は、1つ以上の核酸が挿入、欠失または付加され得る。例えば、アミノ酸配列は、1つ以上のアミノ酸の挿入、欠失または付加があり得る。

20

【0027】

操作可能な連結

いくつかの実施形態において、ベクターは、プロモーターおよび/または転写終結因子などの1つ以上の調節エレメントに操作可能に連結されるポリヌクレオチドを含む。核酸配列が別の核酸配列と機能的関係に置かれる場合、核酸配列は操作可能に連結されている。例えば、ポリペプチドの分泌に関与する前タンパク質として発現される場合、プレ配列または分泌リーダーに対するDNAがポリペプチドに対するDNAに操作可能に連結され、配列の転写に影響を及ぼす場合、プロモーターはコード配列に操作可能に連結され、または翻訳を促進するように置かれる場合、リボソーム結合部位が操作可能にコード配列に連結される。操作可能に連結される配列は近接し得、分泌リーダーの場合、近接し得、リーディングフェーズ(reading phase)にある。

30

【0028】

宿主細胞または宿主生物

宿主細胞は、本開示のポリペプチドをコードするポリヌクレオチドを含有し得る。いくつかの実施形態において、宿主細胞は多細胞生物の一部である。他の実施形態において、宿主細胞は単細胞生物として培養される。

【0029】

宿主生物は、何らかの適切な宿主、例えば微生物を含み得る。本明細書中に記載の方法に有用な微生物としては、例えば、細菌(例えば、大腸菌(*E. coli*))、酵母(例えば、サッカロミセス・セレビシエ(*Saccharomyces cerevisiae*))および植物が挙げられる。生物は原核または真核であり得る。生物は単細胞または多細胞であり得る。

40

【0030】

宿主細胞は原核性であり得る。適切な原核細胞としては、エスキリキア・コリ(*Escherichia coli*)、ラクトバチルス種(*Lactobacillus* sp.)、サルモネラ種(*Salmonella* sp.)およびシゲラ種(*Shigella* sp.)の何らかの様々な研究用の株(例えば、Carrier et al. (1992) J. Immunol. 148: 1176-1181; 米国特許第6,447,784号明細書; および Sizemore et al. (1995) Science 2

50

70:299-302に記載のとおり)が挙げられるが限定されない。本開示で使用され得るサルモネラ(*Salmonella*)菌種の例としては、サルモネラ・チフィ(*Salmonella typhi*)および*S.チフィムリウム*(*S. typhimurium*)が挙げられるが限定されない。適切なシゲラ(*Shigella*)菌種としては、シゲラ・フレックスネリ(*Shigella flexneri*)、シゲラ・ソネイ(*Shigella sonnei*)およびシゲラ・ディセンテリエ(*Shigella dysenteriae*)が挙げられるが限定されない。一般的には、研究用の菌種は、非病原性であるものである。他の適切な細菌の非限定例としては、シュードモナス・プディタ(*Pseudomonas putida*)、シュードモナス・エルギノーサ(*Pseudomonas aeruginosa*)、シュードモナス・メバロニイ(*Pseudomonas mevalonii*)、ロドバクター・スフェロイデス(*Rhodobacter sphaeroides*)、ロドバクター・カプスラツス(*Rhodobacter capsulatus*)、ロドスピリラム・ラブラム(*Rhodospirillum rubrum*)およびロドコッカス種(*Rhodococcus* sp.)が挙げられるが限定されない。

10

#### 【0031】

いくつかの実施形態において、宿主生物は真核性である。適切な真核宿主細胞としては、酵母細胞、昆虫細胞、植物細胞、真菌細胞および藻類細胞が挙げられるが限定されない。

20

#### 【0032】

ポリヌクレオチドおよびポリペプチド〔核酸およびタンパク質〕

本開示のタンパク質は当技術分野で公知の何らかの方法により作製され得る。固相ペプチド合成を使用するか、または液相ペプチド合成としても知られる古典的な溶液ペプチド合成によるかの何れかでタンパク質を合成し得る。*Val-Pro-Pro*、出発鑄型としてエナラプリルおよびリシノプリルを使用し、固相または液相ペプチド合成を使用して、いくつかの一連のペプチド類似体、例えば*X-Pro-Pro*、*X-Ala-Pro*および*X-Lys-Pro*など(式中、*X*は何らかのアミノ酸残基に相当する)を合成し得る。可溶性のオリゴマー支持体にカップリングされたペプチドおよびオリゴヌクレオチドのライブラリの液相合成を行うための方法は既に記載されている。*Bayer, Ernst and Mutter, Manfred, Nature 237:512-513 (1972)*; *Bayer, Ernst, et al., J. Am. Chem. Soc. 96:7333-7336 (1974)*; *Bonora, Gian Maria, et al., Nucleic Acids Res. 18:3155-3159 (1990)*。液相合成法は、固相への反応物質の連結に適切な第1の反応物質上に存在する構造を必要としないという点で固相合成法を上回る長所がある。また、液相合成法は、固相と第1の反応物質(または中間体産物)との間の結合を切断し得る化学的条件を回避する必要がない。さらに、均一溶液中での反応は、固相合成中に存在するものなど、不均一な固相/液相系で得られるものよりも良好な収率および完全な反応を与え得る。

30

#### 【0033】

オリゴマー支持液相合成において、伸長生成物は大きい可溶性のポリマー基に連結される。次いで、本合成の各段階からの生成物は、比較的大きいポリマーに連結された生成物と未反応の反応物との間のサイズの大きい相違に基づいて未反応反応物から分離し得る。これにより、均一溶液中で反応を行うことが可能になり、従来の液相合成に伴う冗漫な精製段階が削除される。オリゴマー支持液相合成はペプチドの自動液相合成にも適応させられている。*Bayer, Ernst, et al., Peptides: Chemistry, Structure, Biology, 426-432*。

40

#### 【0034】

固相ペプチド合成の場合、手順には、伸長するペプチドの末端を不溶性支持体に連結すると同時に、所望の配列のペプチドに適切なアミノ酸を連続的にアセンブリすることが伴う。通常、ペプチドのカルボキシル末端はポリマーに連結されるが、これは切断試薬での

50

処理時にポリマーから遊離し得る。一般的な方法において、アミノ酸はレジン粒子に結合され、アミノ酸の鎖を生成させるために、保護されたアミノ酸の連続的付加によって段階的にペプチドが生成する。Merrifieldにより記載される技術の改変が一般的に使用される。例えば、Merrifield, J. Am. Chem. Soc. 96: 2989-93 (1974)を参照されたい。自動化固相法において、ペプチドは、カルボキシ末端アミノ酸を有機リンカー（例えば、PAM、4-オキシメチルフェニルアセトアミドメチル）上にローディングし、これをジビニルベンゼンで架橋された不溶性のポリスチレンレジンに共有結合させることによって合成される。末端アミンは、t-ブチルオキシカルボニルでブロックすることによって保護し得る。ヒドロキシルおよびカルボキシル基は、O-ベンジル基でブロックすることによって一般的に保護される。Applied Biosystems (Foster City, Calif.) から入手可能なものなど、自動化ペプチド合成装置において合成を遂行する。合成後、レジンから生成物を取り出し得る。確立された方法に従い、フッ化水素酸またはトリフルオロメチルスルホン酸を使用することにより、ブロッキング基を除去する。通常の合成では0.5 mmoleのペプチドレジンが生じ得る。切断および精製後、一般的にはおよそ60~70%の収率が得られる。生成物ペプチドの精製は、例えば、メチル-ブチルエーテルなどの有機溶媒からペプチドを結晶化し、次いで蒸留水中で溶解させ、および透析（対象ペプチドの分子量が約500ダルトンを超える場合）またはペプチドの分子量が500ダルトン未満である場合には逆高圧液体クロマトグラフィー（例えば、溶媒として0.1%トリフルオロ酢酸およびアセトニトリルを用いてC<sup>18</sup>カラムを使用）を使用することによって遂行される。精製ペプチドを凍結乾燥し、使用まで乾燥状態で保存し得る。得られたペプチドの分析は、分析的高圧液体クロマトグラフィー（HPLC）およびエレクトロスプレー質量分析（ESI-MS）の一般的な方法を使用して遂行され得る。

#### 【0035】

他の場合において、タンパク質、例えばタンパク質は組み換え法により作製される。本明細書中に記載のタンパク質の何れかの作製のために、このようなタンパク質をコードするポリヌクレオチドを含有する発現ベクターで形質転換した宿主細胞を使用し得る。宿主細胞は、高等真核細胞、例えば哺乳動物細胞など、または下等真核細胞、例えば酵母などであり得るか、または宿主は、細菌細胞などの原核細胞であり得る。宿主細胞への発現ベクターの導入は、リン酸カルシウム形質移入、DEAE-デキストランが介在する形質移入、ポリブレン、プロトプラスト融合、リボソーム、核への直接的なマイクロインジェクション、スクレイプローディング（scrape loading）、微粒子銃形質転換およびエレクトロポレーションを含む様々な方法によって遂行され得る。組み換え生物からのタンパク質の大規模生産は、商業スケールで実施されるよく確立された工程であり、十分に当業者の能力の範囲内である。

#### 【0036】

##### コドン最適化

コードポリヌクレオチドの1つ以上のコドンは、宿主生物のコドン使用を反映させるために「偏向」または「最適化」され得る。例えば、コードポリヌクレオチドの1つ以上のコドンは、葉緑体コドン使用または核コドン使用を反映させるために「偏向」または「最適化」され得る。殆どのアミノ酸は、2つ以上の異なる（縮重）コドンによりコードされ、様々な生物が他よりも優先して一定のコドンを利用することはよく認識されている。「偏向」または「最適化」コドンは、本願全体を通して交換可能に使用され得る。コドン偏向は、タバコと比較した場合、例えば藻類を含め、様々な植物において様々な歪められ得る。一般に、選択されるコドン偏向は、本開示の核酸で形質転換されている植物（またはその中の小器官）のコドン使用を反映する。

#### 【0037】

特定のコドン使用に対して偏りがあるポリヌクレオチドは、デノボ合成され得るか、または葉緑体コドン使用について偏りがあるように1つ以上のコドンを変化させるため、通常の組み換えDNA技術を用いて例えば部位特異的突然変異誘発法によって遺伝的に修飾



され得る。

#### 【0038】

##### パーセント配列同一性

核酸またはポリペプチド配列間でパーセント配列同一性または配列類似性を決定するための適切であるアルゴリズムの一例は、BLASTアルゴリズムであり、これは、例えば、Altschul et al., J. Mol. Biol. 215: 403-410 (1990) に記載されている。BLAST分析を行うためのソフトウェアは、National Center for Biotechnology Informationを通じて公開されている。BLASTアルゴリズムパラメーターW、TおよびXは、アライメントの感度および速度を決定する。BLASTNプログラム（ヌクレオチド配列用）は、初期設定として、ワード長（W）11、期待値（E）10、カットオフ100、M=5、N=4および両鎖の比較を使用する。アミノ酸配列の場合、BLASTPプログラムは、初期設定として、ワード長（W）3、期待値（E）10およびBLOSUM62スコアリングマトリクス（例えば、Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA, 89: 10915に記載のとおり）を使用する。パーセント配列同一性を計算することに加えて、BLASTアルゴリズムはまた、2つの配列間の類似性の統計学的分析も行い得る（例えば、Karlin & Altschul, Proc. Natl. Acad. Sci. USA, 90: 5873-5877 (1993) に記載のとおり）。BLASTアルゴリズムにより提供される類似性のある目安は、最小和確率（smallest sum probability）（P(N)）であり、これは、2つのヌクレオチドまたはアミノ酸配列間の一致が偶然起こる確率の指標を提供する。例えば、核酸は、試験核酸と参照核酸との比較における最小和確率が約0.1未満、約0.01未満、および約0.001未満である場合、参照配列と同様とみなされる。

10

20

30

#### 【0039】

本開示は、A) 例えばDNA切断能について不活性であるCasタンパク質（例えば、Cas9）に共有結合されるウイルスインテグラーゼ（またはレコンビナーゼ）を含む系を含む。あるいは、TALタンパク質またはジンクフィンガータンパク質がゲノム中のDNAの特異的な配列を標的とするように設計される場合、TALタンパク質またはジンクフィンガータンパク質にウイルスインテグラーゼ（または細菌またはファージレコンビナーゼ）が共有結合される。

#### 【0040】

これは、発現ベクターにおいてまたは精製タンパク質として提供され得る。B) 所望のゲノムに組み込まれる相同アームがあるかまたはない関心のある遺伝子（または関心のあるDNA配列）。必要に応じて、ウイルスインテグラーゼにより認識させるためにGOIまたは関心のあるDNA配列を修飾し得る。例えば、DNA配列の末端にウイルスatt部位を付加し得る。C) 細胞へのポリヌクレオチド形質移入および/またはタンパク質導入に必要とされる他の試薬。

#### 【0041】

##### 核酸

「ポリヌクレオチド」、「ヌクレオチド」、「ヌクレオチド配列」、「核酸」および「オリゴヌクレオチド」という用語は、本開示において交換可能に使用される。これらは、あらゆる長さのデオキシリボヌクレオチドまたはリボヌクレオチドの何れかのヌクレオチドのポリマー形態またはそれらの類似体を指す。ポリヌクレオチドは、何らかの三次元構造を有し得、既知または未知の何らかの機能を発揮し得る。以下のものはポリヌクレオチドの非限定例である：遺伝子または遺伝子断片のコードまたは非コード領域、連鎖解析から定められる遺伝子座（複数の遺伝子座）、エクソン、イントロン、メッセンジャーRNA（mRNA）、転移RNA、リボソームRNA、短い干渉RNA（siRNA）、短いヘアピンRNA（shRNA）、マイクロRNA（miRNA）、リボザイム、cDNA、組み換えポリヌクレオチド、分岐状ポリヌクレオチド、プラスミド、ベクター、何らかの

40

50

配列の単離DNA、何らかの配列の単離RNA、核酸プローブおよびプライマー。ポリヌクレオチドは、1つ以上の修飾ヌクレオチド、例えばメチル化ヌクレオチドおよびヌクレオチド類似体などを含み得る。存在する場合、ポリマーのアセンブリ前または後にヌクレオチド構造に対する修飾が与えられ得る。ヌクレオチドの配列には非ヌクレオチド構成要素が割り込み得る。ポリヌクレオチドは、標識化構成成分との抱合反応などによってポリマー化後にさらに修飾され得る。

#### 【0042】

##### ガイドRNA

本開示の態様において、「キメラRNA」、「キメラガイドRNA」、「ガイドRNA」、「単ガイドRNA」および「合成ガイドRNA」という用語は、交換可能に使用され、ガイド配列、tracr配列およびtracrメート(tracr mate)配列を含むポリヌクレオチド配列を指す。「ガイド配列」という用語は、標的部位を規定するガイドRNA内の約20bp(12~30bp)配列を指し、「ガイド」または「スペーサー」という用語と交換可能に使用され得る。「tracrメート配列」という用語は、「ダイレクトリピート」という用語とも交換可能に使用され得る。

#### 【0043】

##### 野生型

本明細書中で使用される場合、「野生型」という用語は、当業者により理解される当技術分野の用語であり、突然変異体または変異体形態と区別される場合、天然で生じるような生物、株、遺伝子または特徴の典型的な形態を意味する。

#### 【0044】

##### 変異体

本明細書中で使用される場合、「変異体」または「突然変異体」という用語は、天然に生じるものから逸脱するパターンを有する質の提示を意味するものと解釈すべきである。遺伝子に関連して、これらの用語は、とりわけ、一塩基多型(SNP)、挿入、欠失、遺伝子シフトを含め、それを野生型遺伝子と異なるものとする遺伝子における多くの変化を示す。

#### 【0045】

##### 改変

「非天然」または「改変」という用語は、交換可能に使用され、人為的な技術の関与を指す。これらの用語は、核酸分子またはポリペプチドを指す場合、核酸分子またはポリペプチドが、それらが自然界で天然に結び付き、かつ天然で見られるような少なくとも1つの他の構成要素を少なくとも実質的に含まないことを意味する。

#### 【0046】

##### 相補性

「相補性」とは、古典的なワトソン-クリックまたは他の非古典的タイプの何れかにより、核酸が別の核酸配列と水素結合を形成する能力を指す。パーセント相補性は、第2の核酸配列と水素結合を形成し得る(例えば、ワトソン-クリック塩基対)核酸分子中の残基のパーセンテージを指す(例えば、10個のうちの5、6、7、8、9、10個は、50%、60%、70%、80%、90%および100%相補的である)。「完全に相補性」とは、核酸配列の全ての近接残基が第2の核酸配列中の同じ数の近接残基と水素結合することを意味する。「実質的に相補性」とは、本明細書中で使用される場合、少なくとも60%、70%、75%、80%、85%、90%、95%、97%、98%、99%もしくは100%または8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、30、35、40、45、50またはそれを超えるヌクレオチドの領域にわたる間のパーセンテージである相補性の度合いを指すか、またはストリンジェントな条件下でハイブリッド形成する2個の核酸を指す。

#### 【0047】

##### アミノ酸

正式名称、3文字コード、1文字コード

アスパラギン酸 A s p D  
 グルタミン酸 G l u E  
 リジン L y s K  
 アルギニン A r g R  
 ヒスチジン H i s H  
 チロシン T y r Y  
 システイン C y s C  
 アスパラギン A s n N  
 グルタミン G l n Q  
 セリン S e r S  
 スレオニン T h r T  
 グリシン G l y G  
 アラニン A l a A  
 バリン V a l V  
 ロイシン L e u L  
 イソロイシン I l e I  
 メチオニン M e t M  
 プロリン P r o P  
 フェニルアラニン P h e F  
 トリプトファン T r p W

10

20

30

40

50

## 【0048】

「アミノ酸」という表現は、本明細書中で使用される場合、天然および合成両方のアミノ酸およびDおよびLアミノ酸の両方を含むものとする。「標準アミノ酸」とは、天然のタンパク質/ペプチドにおいて一般的に見られる20種類の標準的なL-アミノ酸の何れかを意味する。「非標準アミノ酸残基」とは、合成により調製されるかまたは天然起源であるかにかかわらず、標準アミノ酸以外の何らかのアミノ酸を意味する。本明細書中で使用される場合、「合成アミノ酸」は、塩、アミノ酸誘導体（アミドなど）および置換を含むが限定されない、化学的に修飾されたアミノ酸を包含する。本開示のペプチド内に、および特にカルボキシまたはアミノ末端で含有されるアミノ酸は、メチル化、アミド化、アセチル化または活性に悪影響を与えることなくペプチドの循環半減期を変化させ得る他の化学基での置換により修飾され得る。さらに、ジスルフィド結合がペプチド中に存在してもよい。または存在しなくてもよい。

## 【0049】

アミノ酸は、側鎖Rに基づいて7個の群に分類され得る：（1）脂肪族側鎖；（2）ヒドロキシル（OH）基を含有する側鎖；（3）硫黄原子を含有する側鎖；（4）酸性またはアミド基を含有する側鎖；（5）塩基性基を含有する側鎖；（6）芳香環を含有する側鎖；および（7）プロリン、側鎖がアミノ基に融合されるイミノ酸。

## 【0050】

本明細書中で使用される場合、「保存的アミノ酸置換」という用語は、以下の5つの群の1つ内の交換として本明細書中で定義される。

I．小さい脂肪族、非極性または僅かに極性のある残基：

A l a、S e r、T h r、P r o、G l y；

II．極性、負荷電残基およびそれらのアミド：

A s p、A s n、G l u、G i n；

III．極性、正荷電残基：

H i s、A r g、L y s；

IV．大きい脂肪族、非極性残基：

M e t、L e u、H e、V a l、C y s（I l e；オートコレクトはリテレート（l i t e r a t e）ではない）

V．大きい芳香族残基：

P h e、T y r、T i p ( 同様に T r p )

【 0 0 5 1 】

本開示は、別段の提供がない限り、免疫学、生化学、化学、分子生物学、微生物学、細胞生物学、ゲノムおよび組み換えDNAの従来技術を利用し、これは、当技術分野の技術の範囲内である。Sambrook, Fritsch and Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd edition (1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F. M. Ausubel, et al. eds. (1987)); シリーズ METHODS IN ENZYMOLOGY (Academic Press, Inc.): PCR 2: A PRACTICAL APPROACH (M. J. MacPherson, B. D. Hames and G. R. Taylor eds. (1995)), Harlow and Lane, eds. (1988) ANTIBODIES, A LABORATORY MANUAL, and ANIMAL CELL CULTURE (R. I. Freshney, ed. (1987)) を参照されたい。

10

20

30

40

50

【 0 0 5 2 】

ベクター

細胞または組織において融合インテグラーゼを発現させ、かつ関心のあるDNA (または遺伝子) を宿主種または細胞のゲノムに組み込むためのインテグラーゼまたはレコンビナーゼに必要とされる適切な部位がある関心のあるDNA配列 (または遺伝子) を提供するために遺伝子発現ベクター (DNAに基づくかまたはウイルス性) を使用する。多くの遺伝子発現ベクターが当技術分野で公知である。ベクターは、関心のある遺伝子 (または関心のあるDNA配列) に対する使用である。当技術分野で公知の多くの制限酵素を用いてベクターを切断し得る。

【 0 0 5 3 】

C R I S P R / C A S 9

C R I S P R / C a s 9 は、米国特許第 8 6 9 7 3 5 9 号明細書、同第 8 8 8 9 3 5 6 号明細書およびRan et al (Nature Protocols, 2013, volume 8, pages 2281 - 2308) に記載されている。Cas9タンパク質は、ゲノム中のDNAの特異的な配列に結合するためにRNAガイドを利用する。RNAガイド (ガイドRNA) は、10 ~ 40、12 ~ 35、15 ~ 30または例えば18 ~ 22または20ヌクレオチド長になるように設計され得る。ストレプトコッカス・ピオゲネス (Streptococcus pyogenes) からのCas9を使用する、Hsu et al, Nature Biotechnology, September 2013, volume 31, pages 827 - 832を参照されたい。別の重要なCas9は、スタフィロコッカス・アウレウス (Staphylococcus aureus) 由来 (S. ピオゲネス (S. pyogenes) よりも小さいCas9) である。Cas9タンパク質は、DNA配列の特異的領域に結合するためにガイドRNAを利用する。

【 0 0 5 4 】

C a s 9 の触媒能のない形態は、Guilinger et al, Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification, Nature Biotechnology, April 25, 2014, volume 32, pages 577 - 582 に記載されている。Guilingerらは、触媒能のないCas9をFokI酵素に連結させて、ゲノムDNAにおいて切断させることにおいてより大きい特異性を達成した。この触媒能のないCas9により、ゲノムDNAの結合のためにCas9がRNAガイドを使用できるようになり、同時にこれはこのDNAを切断できない。

【 0 0 5 5 】

Cas9は、細胞でのCas9コンストラクトのより良好な発現のためにその天然wt型およびまたヒト最適化コドン形態でも利用可能である(Maliet al, Science, 2013, volume 339, pages 823-826を参照されたい)。Cas9のコドン最適化は、その発現のための種に依存して行い得る。インテグラーゼ/Cas9融合タンパク質(ABBE1としても知られる)のタンパク質形態を生成させるかまたはヌクレオチド発現ベクター形態を生成させるかに依存して、最適化または非最適化(wt)形態を使用し得る。

#### 【0056】

特異的なDNA配列に対するRNAガイドは、様々なコンピュータに基づくツールによって設計し得る。

#### 【0057】

#### CRISPR/Cpf1

Cpf1は、ゲノムDNA中の特異的な配列に結合するためにガイドRNAを使用する別のタンパク質である。Cpf1はまた、DNAを切断し、ねじれ型の切断にする。Cpf1は、切断能について触媒的に不活性にされ得る。

#### 【0058】

#### 他のCRISPRタンパク質

これらは、特異的なDNA配列を標的とするために、かつそれらがDNA切断能を有するか否かにかかわらずガイドRNAを利用するタンパク質である。これらのタンパク質の一部は他の酵素/触媒機能を天然に有し得る。

#### 【0059】

#### TALEN

転写活性化因子様エフェクターヌクレアーゼ(TALEN)は、TALEフェクターDNA結合ドメインをDNA切断ドメインに融合させることによって作製される制限酵素との融合タンパク質である。これらの試薬は、効率的でプログラム可能であり、特異的なDNA切断を可能にし、インサイチューでのゲノム編集のための強力なツールとなる。実際に何らかのDNA配列に結合するように転写活性化因子様エフェクター(TALE)を迅速に改変し得る。TALENという用語は、本明細書中で使用される場合、別のTALENからの支援なく2本鎖DNAを切断し得る単量体のTALENを広く含む。TALENという用語は、一緒に作用して同じ部位でDNAを切断するために改変されるTALENのペアの一方または両方のメンバーを指すためにも使用される。一緒に作用するTALENは、左-TALENおよび右-TALENと呼ばれ得、これはDNAの硬さに言及する。米国特許第8,440,432号明細書を参照されたい。

#### 【0060】

TALEフェクターは、ザントモナス(Xanthomonas)細菌により分泌されるタンパク質である。DNA結合ドメインは、12番目および13番目のアミノ酸を除き、保存性が高い33~34アミノ酸配列を含有する。これらの2つの位置は変動性が高く(反復変動二残基(Repeat Variable Di-residues)(RVD))、特異的なヌクレオチド認識との強い相関を示す。アミノ酸配列とDNA認識との間のこの単純な関係性により、適切なRVDを含有するリピートセグメントの組み合わせを選択することによって特異的なDNA結合ドメインの改変が可能になっている。

#### 【0061】

酵母または細胞アッセイにおいて活性のあるハイブリッドインテグラーゼまたはレコンビナーゼを構築するためにインテグラーゼまたはレコンビナーゼを使用し得る。これらの試薬は、植物細胞および動物細胞でも活性がある。TALENの研究は野生型FokI切断ドメインを使用した。一部のその後のTALEN研究は、切断特異性および切断活性を向上させるように設計された突然変異があるFokI切断ドメイン変異体も使用した。TALEN DNA結合ドメインとインテグラーゼまたはレコンビナーゼドメインとの間のアミノ酸残基数および2つの個々のTALEN結合部位間の塩基数の両方とも、高レベルの活性を達成するためのパラメーターである。TALEN DNA結合ドメインとイン

10

20

30

40

50

テグラーゼまたはレコンビナーゼドメインとの間のアミノ酸残基数は、複数の T A L E フォクターリピート配列とインテグラーゼまたはレコンビナーゼドメインとの間のスペーサー（スペーサー配列とは別個）の導入により修飾され得る。スペーサー配列は、6 ~ 102 または 9 ~ 30 個のヌクレオチドまたは 15 ~ 21 個のヌクレオチドであり得る。これらのスペーサーは、通常、DNA 標的化タンパク質（Cas9、TALE またはジンクフィンガータンパク質）とインテグラーゼまたはレコンビナーゼとの間の連結を提供すること以外に、ハイブリッドタンパク質に対して他の活性を提供しない。本開示におけるスペーサーおよび他の使用に対するアミノ酸である。

#### 【0062】

TALEN 結合ドメインのアミノ酸配列と DNA 認識との間の関係は、明示できるタンパク質を考慮に入れる。この場合、TALE 結合ドメインで見られる反復配列の不適切なアニーリングのために人工的遺伝子合成は問題となる。これに対するある解決策は、2 段階 PCR；オリゴヌクレオチド組み立ておよびそれに続く全遺伝子増幅での組み立てに適切なオリゴヌクレオチドを発見するため、DNA Works という名称の公開ソフトウェアプログラムを使用することである。改変 TALE コンストラクトを作製するための多くのモジュラー組み立て方法も当技術分野で報告されている。

#### 【0063】

TALEN 遺伝子を一緒に組み立てたら、これらをプラスミドに挿入し、次いでプラスミドを使用して、遺伝子産物を発現させる標的細胞に対して形質移入を行い、核に移行させてゲノムに到達させる。TALEN は、2 本鎖切断（DSB）を誘導することによってゲノムを編集するために使用され得、この細胞は DNA 修復に応答するが、本開示は、関心のある DNA 配列をゲノム中の標的化部位に挿入するためにウイルスインテグラーゼまたは細菌もしくはファージレコンビナーゼの力を使用しようとするものである。国際公開第 2014134412 号パンフレットおよび米国特許第 8748134 号明細書の開示を参照されたい。

#### 【0064】

##### ジンクフィンガータンパク質

DNA 結合のためのジンクフィンガータンパク質およびそれらの設計は、米国特許第 7928195 号明細書、米国特許出願公開第 2009/0111188 号明細書および米国特許第 7951925 号明細書に記載されている。ジンクフィンガータンパク質は、DNA の特異的配列に結合するために、指定の順序でいくつかの連結されるジンクフィンガードメインを利用する。ジンクフィンガータンパク質エンドヌクレアーゼはよく確立されている。

#### 【0065】

ジンクフィンガータンパク質（ZFP）は、配列特異的に DNA に結合し得るタンパク質である。ジンクフィンガーは、最初に、アフリカツメガエル、ゼノパス・ラエビス（*Xenopus laevis*）の卵母細胞からの転写因子 TFIIIA において同定された。ZFP のこのクラスの単一ジンクフィンガードメインは、約 30 アミノ酸長であり、いくつかの構造研究から、（2 個の保存的システイン残基を含有する）ベータターンおよび（2 個の保存的ヒスチジン残基を含有する）アルファヘリックスを含有することが明らかになっており、この 2 個のシステインおよび 2 個のヒスチジンによる亜鉛原子の配位を通じて特定の高次構造で保持される。このクラスの ZFP は、C<sub>2</sub>H<sub>2</sub> ZFP としても知られる。ZFP のさらなるクラスも示唆されている。例えば、Cys-Cys-His-Cys（C<sub>3</sub>H）ZFP の考察に対する Jiang et al.（1996）J. Biol. Chem. 271:10723-10730 を参照されたい。現在までに、数千個の既知または推定転写因子において 10,000 個を超えるジンクフィンガー配列が同定されている。ジンクフィンガードメインは、DNA 認識だけでなく、RNA 結合およびタンパク質-タンパク質結合にも関与する。現在、このクラスの分子は全ヒト遺伝子の約 2% にのぼると推定される。

#### 【0066】

多くのジンクフィンガータンパク質は、各フィンガードメインで1個の亜鉛原子を四面配位させる保存的なシステインおよびヒスチジン残基を有する。特に、殆どのZFPは、一般配列：-Cys-(X)<sup>2-4</sup>-Cys-(X)<sup>12</sup>-His-(X)<sup>3-5</sup>-His-(配列番号49、式中、Xは何らかのアミノ酸に相当する(C<sub>2</sub>H<sub>2</sub>ZFP))のフィンガー構成要素を特徴とする。この最も幅広く代表される亜鉛-配位配列は、特定のスペーシングがある2個のシステインおよび2個のヒスチジンを含有する。各フィンガーの折り畳まれた構造は、逆平行-ターン、フィンガーチップ領域および短い両親媒性-ヘリックスを含有する。金属配位リガンドは、亜鉛イオンに結合し、Zn<sub>2</sub>68型ジンクフィンガーの場合、短い両親媒性-ヘリックスがDNAの主溝に結合する。さらに、ジンクフィンガーの構造は、ある種の保存的疎水性アミノ酸残基(例えば、第1の保存的Cysの直前の残基およびフィンガーのヘリックスセグメントの位置+4の残基)により、かつ保存的システインおよびヒスチジン残基を通じた亜鉛配位により安定化される。

【0067】

ゲノムDNA中の特異的な標的配列に結合し得る他のDNA結合タンパク質

本タンパク質は、様々な生物のゲノムDNA中の特異的な配列に結合し得る、ジンクフィンガータンパク質、TALENおよびCRISPRタンパク質と無関係なものを含む。これらは、転写因子、転写リプレッサー、メガヌクレアーゼ、エンドヌクレアーゼDNA結合ドメインなどを含み得る。

【0068】

インテグラーゼ

インテグラーゼおよびそのエンドヌクレアーゼ融合タンパク質は、米国特許出願公開第2009/0011509号明細書に記載されている。導入されるインテグラーゼは、レンチウイルスインテグラーゼおよびHIV1(ヒト免疫不全ウイルス1)インテグラーゼである。本開示は、使用者により選択されるゲノム中のDNAの特異的な領域にインテグラーゼを標的化するため、触媒能のない(または触媒活性のある)Cas9、TALENまたはジンクフィンガータンパク質をインテグラーゼに融合させる。

【0069】

他のレトロウイルスインテグラーゼのように、HIV-1インテグラーゼは、長い末端反復配列(LTR)のU3およびU5領域に位置するウイルスDNA末端の特別な特性を認識可能である(Brown, 1997)。LTR末端は、レトロウイルスの組み込み機構による認識のためにシスで必要とされると思われる単なるウイルス配列である。短い不完全な逆方向反復は、マウスおよび鳥類レトロウイルスの両方においてLTRの外縁に存在する(Reicin et al., 1995により概説)。レトロウイルスDNA末端において最も外側の位置3および4に位置するサブターミナル(subterminal)CAと一緒に(位置1および2は3'末端プロセッシングヌクレオチドである、これらの配列は、インビトロおよびインビボでの正確なプロウイルス組み込みに必要かつ十分である。CAジヌクレオチドに対して内部にある配列は、最適インテグラーゼ活性のために重要であると思われる(Brin & Leis, 2002a; Brin & Leis, 2002b; Brown, 1997)。HIV-1 LTRの末端15bpは、インビトロでの正確な3'末端プロセッシングおよび鎖転移反応にとって重大なものであることが示されている(Reicin et al., 1995; Brown, 1997)。より長い基質は、結合相互作用がウイルスDNA末端から内側に少なくとも14~21bp伸長することを示すHIV-1 INにより、短いものよりも効率的に使用される。Brin & Leis(2002a)は、HIV-1 LTRの特異的な特性を分析し、U5 LTRがインビトロでのINプロセッシングに対してより効率的な基質であるにもかかわらず、U3およびU5 LTR認識配列の両方がIN-触媒型の協調したDNA組み込みに必要とされると結論付けた(Bushman & Craigie, 1991; Sherman et al., 1992)。協調したDNA組み込み機序のためにIN認識配列の位置17~20が必要とされるが、HIV-1 INは、不変のサブターミナル(subterminal)CAジヌクレオチドから伸長するU3およびU5末端の両方において相当な変動を許

10

20

30

40

50

容する (Brin & Leis, 2002b)。本開示は、ゲノムに組み込もうとする関心のあるDNA配列または遺伝子を収容するための位置の5'および3'末端でウイルス(レトロウイルスまたはHIV)LTR領域を含有するDNAベクターを含む。LTR領域は、それらが正確な組み込みのためにインテグラーゼと相互作用するように機能する限り、全長LTRである必要はない。LTR領域は、検出可能な(例えば、蛍光)PCR検出または選択可能マーカー(例えば、抗生物質耐性)を含有するように修飾され得る。ベクターは、(制限エンドヌクレアーゼに対する設計された制限部位を介して)LTR領域がDNA断片の5'および3'末端となるように切断され、直線化されるように設計される。

#### 【0070】

インテグラーゼは、可動性リンカーにより連結される3個のドメインからなる。これらのドメインは、N末端HH-CC亜鉛-結合ドメイン、触媒性コアドメインおよびC末端DNA結合ドメインである(Lodiet al, Biochemistry, 1995, volume 34, pages 9826-9833)。本開示のいくつかの態様において、Cas9(または他のDNA結合分子)に結合されるインテグラーゼはC末端結合ドメインを有さない。本開示のある態様において、一方がインテグラーゼのN末端亜鉛結合ドメインと融合される触媒能のないCas9(またはTALEもしくはジンクフィンガータンパク質)を有し、他方がインテグラーゼの触媒コアドメインと融合される触媒能のないCas9(またはTALEまたはジンクフィンガータンパク質)を有する、2つの異なる融合タンパク質を作製する。この2つの異なる融合タンパク質は、TALE-FokIまたはジンクフィンガー-FokI系で見られるように、ゲノムDNAの逆の鎖に結合するように設計される。このように、ゲノムDNA上の部位でN末端ドメインおよび触媒コアが接触する場合、これはインテグラーゼ活性を示す。インテグラーゼの完全活性がインテグラーゼの四量体を含むことも観察されているため、1~20アミノ酸長または4~12アミノ酸長であり得る可動性リンカーにより連結される1、2、3、4個のインテグラーゼタンパク質とともに、融合タンパク質が設計され得る。

#### 【0071】

##### レコンビナーゼ

Cre、Flp、R、Dre、KwおよびGinレコンビナーゼを含むレコンビナーゼは、米国特許第8816153号明細書および米国特許出願公開第2004/0003420号明細書に記載される。Creレコンビナーゼなどのレコンビナーゼは、ゲノムから配列を切り出すためにLoxP部位を使用する。レコンビナーゼは、それらの組み換え活性に対して構成的に活性となるように、およびまた部位特異性が低くなるようにも修飾され得る。したがって、本開示の融合タンパク質にそれらを組み込むことにより、ゲノム中のDNAの特異的配列に対する配列特異性がないこのような構成的に活性のあるレコンビナーゼタンパク質を標的とすることが可能である。このように、CRISPR/Cas9、TALEまたはジンクフィンガータンパク質ドメインは、レコンビナーゼがその組み換え活性に寄与するDNA配列を指定する。このようなレコンビナーゼタンパク質は、野生型であるか、構成的に活性であるか、またはレコンビナーゼ活性について不活性であり得る。Cas9-GinまたはCas9-CreなどのCas9-レコンビナーゼは、リンカー配列の使用または直接的な融合により作製され得る。

#### 【0072】

##### 融合タンパク質に対する核局在化シグナル配列(NLS)

(「NLS」とも呼ばれる)シグナルペプチドドメインは、例えば、酵母GAL4、SKI3、L29またはヒストンH2Bタンパク質、ポリオーマウイルスラージTタンパク質、VP1またはVP2カプシドタンパク質、SV40 VP1またはVP2カプシドタンパク質、アデノウイルスE1aまたはDBPタンパク質、インフルエンザウイルスNS1タンパク質、肝炎ウイルスコア抗原または哺乳動物ラミン、c-myc、max、c-myb、p53、c-erbA、jun、Tax、ステロイド受容体またはMxタンパク質(Boulikas, Crit. Rev. Eucar. Gene Expressio

10

20

30

40

50



n, 3, 193 - 227 (1993) を参照されたい)、サルウイルス 40 (「SV40」) T-抗原 (Kalderson et al, Cell, 39, 499 - 509 (1984)) または既知の核局在化がある他のタンパク質由来である。NLS は、例えば、SV40 T-抗原由来であるが、当技術分野で公知の他の NLS 配列であり得る。タンデム NLS 配列が使用され得る。

#### 【0073】

##### リンカー領域

合成される融合タンパク質 / ペプチド間で使用される様々なリンカーはアミノ酸から構成される。DNA レベルにおいて、これらは、遺伝暗号において知られるように 3 塩基対 (bp) コドンにより表される。リンカーは、1 ~ 1000 アミノ酸長およびこの間の何れかの整数であり得る。例えば、リンカーは、1 ~ 200 アミノ酸長であるか、またはリンカーは 1 ~ 20 アミノ酸長である。

10

#### 【0074】

##### 発現ベクター

遺伝子の発現をもたらすために多くの核酸が細胞に導入され得る。本明細書中で使用される場合、核酸という用語は、DNA、RNA および核酸類似体および 2 本鎖もしくは 1 本鎖 (すなわちセンスまたはアンチセンス 1 本鎖) である核酸を含む。例えば、核酸の安定性、ハイブリッド形成または溶解度を改善するために塩基部分、糖部分またはリン酸骨格で核酸類似体を修飾し得る。塩基部分での修飾には、デオキシチミジンに対するデオキシウリジンおよびデオキシシチジンに対する 5 - メチル - 2' - デオキシシチジンおよび 5 - プロモ - 2' - ドキシシチジン (doxycytidine) を含む。糖部分の修飾は、2' - O - メチルまたは 2' - O - アリル糖を形成させるためのリボース糖の 2' ヒドロキシルの修飾を含む。デオキシリボースリン酸骨格は、各塩基部分が 6 員のモルホリノ環に連結されるモルホリノ核酸またはデオキシリン酸骨格が擬ペプチド骨格により置換され、4 個の塩基が保持されるペプチド核酸を生成させるために修飾し得る。Summer ton and Weller (1997) Antisense Nucleic Acid Drug Dev. 7 (3) : 187 ; および Hyrup et al. (1996) Bioorgan. Med. Chem. 4 : 5 を参照されたい。さらに、デオキシリン酸骨格は、例えばホスホロチオエートまたはホスホロジチオエート骨格、ホスホロアミダイトまたはアルキルホスホトリエステル骨格と置換され得る。核酸配列は、プロモーターなどの制御領域に操作可能に連結され得る。制御領域はあらゆる種由来であり得る。本明細書中で使用される場合、操作可能に連結されるとは、標的核酸の転写を可能にするかまたは促進するような核酸配列に対する制御領域の配置を指す。あらゆるタイプのプロモーターが核酸配列に操作可能に連結され得る。プロモーターの例としては、組織特異的なプロモーター、構成的プロモーターおよび特定の刺激に対して応答性または非応答性であるプロモーター (例えば、誘導性プロモーター) が挙げられるが限定されない。

20

30

#### 【0075】

核酸コンストラクトにおいて有用であり得るさらなる領域としては、ポリアデニル化配列、翻訳調節配列 (例えば、配列内リボソーム進入セグメント、IRES)、エンハンサー、誘導性エレメントまたはイントロンが挙げられるが限定されない。このような制御領域は必要でない可能性があるが、これらは、mRNA の転写、安定性、翻訳効率などに影響を及ぼすことによって発現を増加させ得る。このような制御領域は、細胞中の核酸の最適発現を得るため、必要に応じて核酸コンストラクト中に含まれ得る。このようなさらなるエレメントなく、十分な発現が得られることがあり得る。

40

#### 【0076】

シグナルペプチドまたは選択可能マーカーをコードする核酸コンストラクトを使用し得る。コードされるポリペプチドが特定の細胞の位置 (例えば、細胞表面) に向けられるようにシグナル伝達 (マーカー) ペプチドを使用し得る。このような選択可能マーカーの非限定例としては、ピューロマイシン、ガンシクロビル、アデノシンデアミナーゼ (ADA)、アミノグリコシドホスホトランスフェラーゼ (neo、G418、APH)、ジヒド

50

口葉酸レダクターゼ (DHFR)、ハイグロマイシン-B-ホスホトランスフェラーゼ、チミジンキナーゼ (TK) およびキサンチン-グアニンホスホリボシルトランスフェラーゼ (XGPRT) が挙げられる。これらのマーカーは、培養において安定的な形質転換体を選択するために有用である。他の選択可能マーカーとしては、蛍光ポリペプチド、例えば緑色蛍光タンパク質、赤色蛍光または黄色蛍光タンパク質が挙げられる。

#### 【0077】

当技術分野で公知の様々な生物学的技術を使用して、核酸コンストラクトをあらゆるタイプの細胞に導入し得る。これらの技術の非限定例としては、トランスポゾン系、細胞に感染し得る組み換えウイルスまたはリボソームまたは核酸を細胞に送達可能である他の非ウイルス法、例えばエレクトロポレーション、マイクロインジェクションもしくはリン酸カルシウム沈殿などの使用が挙げられる。ヌクレオフェクション・TM・と呼ばれる系も使用され得る。

10

#### 【0078】

核酸をベクターに組み込み得る。ベクターは、標的DNAに担体から移動するように設計されるあらゆる特異的なDNAセグメントを含む広義の用語である。ベクターは、発現ベクターまたはベクター系と呼ばれ得、これは、ゲノムまたは他の標的化DNA配列、例えばエピソーム、プラスミドまたはさらにウイルス/ファージDNAセグメントへのDNA挿入を行うために必要とされる一連の構成要素である。ベクターは、殆どの場合、1つ以上の発現調節配列を含む1つ以上の発現カセットを含有し、ここで、発現調節配列は、それぞれ別のDNA配列またはmRNAの転写および/または翻訳を調節および制御するDNA配列である。

20

#### 【0079】

多くの異なるタイプのベクターが当技術分野で公知である。例えば、プラスミドおよびレトロウイルスベクターを含むウイルスベクターが公知である。哺乳動物発現プラスミドは、一般的には、複製起点、適切なプロモーターおよび任意のエンハンサー、ならびにまたあらゆる必要なリボソーム結合部位、ポリアデニル化部位、スプライスドナーおよび受容体部位、転写終結配列および5'フランキング非転写配列を有する。このようなベクターとしては、プラスミド(別のタイプのベクターの担体でもあり得る)、アデノウイルス、アデノ随伴ウイルス(AAV)、レンチウイルス(例えば、修飾HIV-1、SIVまたはFIV)、レトロウイルス(例えば、ASV、ALVまたはMoMLV)およびトランスポゾン(P-エレメント、Tol-2、Frog Prince、piggyBacなど)が挙げられる。

30

#### 【0080】

本開示での使用のための細菌およびウイルス遺伝子およびタンパク質は、以下の「本開示の配列」という題名のセクションで列挙する。他のウイルスインテグラーゼ、例えばマウス乳癌ウイルス(MMTV)およびアデノウイルス由来のものも本明細書中で開示される方法および組成物において使用し得る。

#### 【0081】

編集された細胞のプール集団は、遺伝子編集を受けている細胞および受けていない細胞の混合物と考えられる。

40

#### 【0082】

代表的なABBIE1インビトロアッセイ

- 1) ガイドRNAとともにABBIE1タンパク質を温置し;
- 2) 前開始複合体を形成させるために部分的LTRを有するドナーDNAとともにABBIE1/ガイドRNAを温置し;
- 3) 編集される遺伝子(例えば、CXCR4)を含有するプラスミドとともに前開始複合体を温置し; および
- 4) ドナーDNA組み込みに対するPCRおよびDNA配列決定を確認する。

#### 【0083】

Cas9プロトコールは、例えば、Gagnon et al., 2014, <http>

50

: // labs.mcb.harvard.edu/schier/VertEmbryo/Cas9\_Protocols.pdf に記載されている。

【0084】

インテグラーゼ活性についてのアッセイは、例えば Merkel et al., Methods, 2009, volume 47, pages 243 - 248 に記載されている。

【実施例】

【0085】

以下の実施例は、本開示の適用の例示を提供するものとする。以下の実施例は、本開示の範囲を完全に定義するものではなく、または本開示の範囲を限定するものではない。当業者にとって当然のことながら、本明細書中で具体的に記載されるかまたは参照されるものの代わりに、当技術分野で公知の多くの他の方法で置き換え得る。

10

【0086】

実施例 1 : Cas9 - インテグラーゼ融合タンパク質を発現させるための DNA ベクター 12、15、18、21、24、27 または 30 bp スペース (Cas9 とインテグラーゼとの間のリンカーとしての 4、5、6、7、8、9 または 10 個のアミノ酸をコードする) および HIV1 インテグラーゼを用いて、触媒能のない Cas9 の DNA 配列を発現ベクターに組み込む。他の実験において、インテグラーゼではなく、細菌またはファージ起源のレコンビナーゼを使用する。これらには、何れかの他の部位で DNA 組み換えを可能にする突然変異があるかまたはない Hin レコンビナーゼ (配列番号 25) および Cre レコンビナーゼ (配列番号 26) が含まれる。融合タンパク質を単離するために His または cMyc タグ (またはタンパク質精製に有用な他の配列) が含まれ得る。発現ベクターは、ベクターとともに提供される細胞中で活性化されるプロモーターを使用する。CMV (サイトメガロウイルスプロモーター) は哺乳動物細胞用の発現ベクターに対して一般的に使用される。U6 プロモーターも一般的に使用される。T7 プロモーターは、一定の実施形態においてインビトロ転写のために使用され得る。

20

【0087】

実施例 2 : 関心のある DNA 配列 (関心のある遺伝子) の発現のための DNA ベクター 関心のある DNA 配列を適切な発現ベクターに挿入し、関心のある DNA 配列に部位が適切に付加され、HIV1 インテグラーゼがゲノムへの組み込みのための配列を認識する。これらの部位は att 部位 (U5 および U3 att 部位) と呼ばれる (Masuda et al., Journal of Virology, 1998, volume 72, pages 8396 - 8402 を参照されたい)。ゲノム中の標的部位に対する相同アームは、関心のある DNA (遺伝子) 配列の 5' および 3' 末端に隣接する領域に含まれ得る (Ishii et al., PLOS ONE, September 24, 2014, DOI: 10.1371/journal.pone.0108236 を参照されたい)。レコンビナーゼを使用する場合、インテグラーゼ認識部位は含まなくてもよい。関心のある DNA 配列の挿入について調べ、かつゲノム中でのランダム挿入についてのアッセイを補助するために薬物耐性マーカー (例えば、ブラストサイジンまたはピューロマイシン) などのマーカーが含まれる。これらの耐性マーカーは、標的化ゲノム着地パッド (landing pad) からそれらを取り出すように改変され得る。例えば、LoxP 部位を有するピューロマイシン耐性遺伝子に隣接することおよび外因的に発現される CRE の導入により、LoxP 部位を含有する痕跡を残す内部配列が除去される。

30

40

【0088】

実施例 3 : 逆転写酵素発現のための DNA ベクター

逆転写酵素はまた、ベクター中の関心のある設計された DNA 配列 (遺伝子) が RNA として発現され、インテグラーゼ酵素による組み込みのために DNA に戻されなければならないような系で同時発現され得る。逆転写酵素はウイルス起源 (例えば、HIV1 などのレトロウイルス) であり得る。これは関心のある DNA 配列と同じベクター内に組み込まれ得る。

50

## 【0089】

実施例4：関心のあるDNA配列とのDNA標的化インテグラーゼ（またはレコンビナーゼ）の同時発現

ゲノム中の標的部位に対して必要とされるCas9 RNAガイドと一緒に、上記のベクターに対して細胞にエレクトロポレーションを行った。一部の実験において、構成要素の全て（融合Cas9/インテグラーゼ（またはレコンビナーゼ）、Cas9 RNAガイドおよびインテグラーゼ認識部位があり、相同アームがあるかまたはない関心のあるDNA配列）を発現するベクターを作製した。ベクター中の関心のある設計されるDNA配列（遺伝子）がRNAとして発現されるようになり、インテグラーゼ酵素による組み込みのためにDNAに戻されなければならないような系において、逆転写酵素も同時発現させ得る。逆転写酵素はウイルス起源（例えば、HIV1などのレトロウイルス）であり得る。他の実験において、細胞への導入前に関心のあるDNA配列を直線化する。標準的な分子生物学プロトコールによる使用前に、Cas9 RNAガイド配列および関心のあるDNA配列を設計し、ベクターに挿入しなければならない。

10

## 【0090】

実施例5：オフターゲット挿入に対する試験的実験およびアッセイ

関心のある遺伝子が含まれる上記ベクターを用いて、特定の遺伝子の発現が失われている細胞、例えば一定の遺伝子に対してノックアウトが起こるように遺伝子操作されているノックアウトマウスモデルまたは細胞からのマウス胚線維芽細胞などに対して形質移入またはエレクトロポレーションを行う。挿入される遺伝子および隣接ゲノム配列を包含するように設計されたキメラプライマーセットを使用して、編集された細胞の最初のプールをスクリーニングする。次に、単クローン性を確実にするために限界希釈クローニング（LDC）およびまたはFACS分析を行う。単離クローンが設計される編集に対して均一であることを確実にするため、最終的な品質管理段階として次世代配列決定（NGS）または一塩基多型（SNP）分析を行う。スクリーニングのための他の機序としては、qRT-PCRおよび適切な抗体を用いたウェスタンブロットティングが挙げられ得るが限定されない。タンパク質が細胞のある種の表現型に付随している場合、その表現型のレスキューについて細胞を調べ得る。DNA挿入の特異性について、および存在する場合にはオフターゲット挿入の相対数を見出すためにその細胞のゲノムをアッセイする。

20

## 【0091】

実施例6：Cas9連結型インテグラーゼタンパク質発現および単離

大腸菌（E coli）または昆虫細胞での遺伝子発現のために設計されたベクターを大腸菌（E coli）または昆虫細胞に組み込み、所与の時間にわたり発現させる。Cas9（または不活性Cas9）連結型インテグラーゼタンパク質を作製するためにいくつかの設計を利用する。本ベクターはまた、高純度および高収率でのタンパク質の最終的な単離のためのHisまたはcMyctagに限定されないtagも組み込む。キメラタンパク質の調製には、標準的なクロマトグラフィー技術が含まれるが限定されない。タンパク質は、1つ以上のNLS（核局在化シグナル配列）および/またはTAT配列を用いても設計され得る。核局在化シグナルにより、タンパク質が核に移行できるようになる。TAT配列により、タンパク質がより容易に細胞に移行できるようになる（これは、細胞透過性ペプチドである）。当技術分野での他の細胞透過性ペプチドも考慮され得る。発現のための十分な時間の経過後、タンパク質溶解液を細胞から回収し、使用したtagに依存して、適切なカラムにおいて精製する。次に、精製タンパク質を適切な緩衝溶液に入れ、-20または-80の何れかで保存する。

30

40

## 【0092】

実施例7：転写開始部位の直接上流に停止コドンを組み込むためのCas9-インテグラーゼの使用

本開示は、ノックアウト細胞株または生物を作製するための方法を含む。1、3、6、10、15または20個の連続停止コドンが標的遺伝子のためのATG開始部位の直後に置かれる関心のあるDNA配列とともに、上記の系を使用する。ATG開始部位の直後の

50

停止コドンに到達したときに転写 / 翻訳が停止されるため、これにより有効な遺伝子ノックアウトが作製される。さらなる停止コドンは、トランスクリプターゼの可能性のある通り抜け (トランスクリプターゼが最初の停止コドンを迂回する場合) を防ぐことを促進する。

#### 【0093】

実施例 8 : 細胞のゲノムを編集するための精製タンパク質としての A B B I E 1 (または他の特異的な D N A 結合ドメインを有する他の変動) の使用

適切な緩衝液中でウイルス L T R 領域を有する挿入可能 / 組み込み可能な D N A とともに、A b b i e 1 単離タンパク質 (レトロウイルスインテグラーゼに連結される他の特異的な D N A 配列結合タンパク質) を温置する (その場合に依存する四量体または他の多量体の形成のため)。あるいは、ガイド R N A を伴う単離 A b b i e 1 タンパク質の予め作製された組成物を挿入可能な D N A 配列と組み合わせ得る。ガイド R N A を含み、ガイド R N A を組み込むために温置する。A b b i e 1 / D N A 標品を細胞に形質移入またはエレクトロポレーション (または細胞にタンパク質を提供する他の技術) する。ゲノム / D N A 編集が行われる時間を考慮に入れる。細胞のゲノム D N A の特異的領域への、設計された挿入可能な D N A 配列の挿入について調べる。P C R および D N A 配列決定により、非特異的な挿入について調べる。

#### 【0094】

現在の計画では、細菌発現ベクターは p M A L - c 5 e であり、これは N E B からの製造中止製品であり、G e n s c r i p t に対する社内クローニング選択肢の 1 つである。マルトース - 結合タンパク質 (M B P) タグとインフレイムで h i s タグおよび T E V プロテアーゼ切断部位を伴い、コドン最適化 S p y C a s 9 をクローニングする。O R F は誘導性 T a c プロモーター下にあり、ベクターは、より厳格な制御のために l a c リプレッサー (L a c I) もコードする。アミロースレジンは非常に高価であるため、精製タグではなく安定化タグとしてのみ M B P を使用する。N i - アフィニティークロマトグラフィー上で可溶性の発現物質を精製し、次に M B P からの T E V プロテアーゼによって C a s 9 を放出させ、陽イオン交換クロマトグラフィーにより精製し、ゲルろ過によりさらに精製する。

#### 【0095】

実施例 9 : 融合タンパク質のためのコンストラクトの設計

標的 D N A 配列に対する C R I S P R に基づくアプローチのために配列特異的なジンクフィンガードメイン、T A L E またはガイド R N A を設計する。最適なオンライン設計ソフトウェアを使用する。

#### 【0096】

部位特異的な融合インテグラーゼタンパク質を形成させるためにインテグラーゼ、トランスポサターゼまたはレコンビナーゼ ; 適切なアミノ酸リンカー ; 適切なジンクフィンガー、T A L E または C R I S P R タンパク質 (例えば、C a s 9、C p f 1) ; および核局在化シグナル (またはミトコンドリアの局在化シグナル) に対するコード配列を有する D N A コンストラクトを作製する。これらは複数の配置で想定される。必要に応じて、タンパク質単離および精製のために適切なタグが含まれ得る (例えば、マルトース結合タンパク質 (M B P) または H i s タグ)。

#### 【0097】

D N A コンストラクトは、当技術分野で一般的である哺乳動物細胞プロモーターまたは細菌プロモーターを利用し得る (例えば、C M V、T 7 など)。

#### 【0098】

起源として大腸菌 (E c o l i) を用いて組み換え融合タンパク質を作製し得る。当技術分野における標準的な手段 (例えば、M B P カラム、ニッケル - セファロースカラムなど) によってタンパク質を単離する。

#### 【0099】

ドナー - R N P 複合体を組み立てて (R N A オリゴを 2 本鎖化、本発明の融合タンパク

10

20

30

40

50

質と混合する（そのDNA結合能に対して融合タンパク質がエンドヌクレアーゼ不活性CRISPR関連タンパク質を有する場合、例えばABBIE1）-RNP形成のこれらの段階は、ジンクフィンガードメインおよびTALEの場合には必要でない。

1. ドナーDNAを適切なLTRドメインおよび挿入可能配列および融合タンパク質と混合し、10分間温置する（あるいはRNP複合体形成後、ドナーDNAを添加する）。

2. ヌクレアーゼ不含IDTE緩衝液中で各RNAオリゴ（crRNAおよびtracrRNA）を懸濁する。例えば、100μMの最終濃度を使用する。

3. 滅菌微量遠心管中で等モル濃度でこの2種類のRNAオリゴを混合する。例えば、以下の表を使用して、3μMの最終2本鎖濃度にする：構成要素量100μM crRNA 3μL、100μM tracrRNA 3μL、ヌクレアーゼ不含2本鎖緩衝液94μL、最終体積100μL。

4. 95℃で5分間加熱する。

5. 熱から除去し、卓上で室温（15～25℃）まで冷却する。

6. 必要に応じて、ヌクレアーゼ不含2本鎖緩衝液中で作業濃度（例えば、3μM）まで2本鎖化RNAを希釈する。

7. 作業用緩衝液（20mM HEPES、150mM KCl、5%グリセロール、1mM DTT、pH7.5）中で作業濃度（例えば、5μM）に融合タンパク質を希釈する。

8. 各形質移入に対して、Opti-MEM培地中で1.5pmolの2本鎖化RNAオリゴ（段階A5）を1.5pmolの融合タンパク質（段階A6）と合わせ、12.5μLの最終体積とする。

9. RNP複合体を組み立てるために室温で5分間温置する。

#### 【0100】

実施例10：96ウェルプレートにおいてGRNA-融合タンパク質を逆形質移入する。

1. 以下のものを室温で20分間温置して形質移入複合体を形成させる：構成要素の量RNP（段階A8）12.5μLリボフェクタミン（登録商標）RNAiMAX形質移入試薬、1.2μL Opti-MEM（登録商標）培地11.3μL総体積25.0μL。

2. 温置（段階B1）中、抗生物質なしの完全培地を使用して培養細胞を細胞400,000個/mLに希釈する。

3. 温置が完了したら、25μLの形質移入複合体（段階B1より）を96ウェル組織培養プレートに添加する。

4. 125μLの希釈細胞（段階B2より）を96ウェル組織培養プレートに添加する（細胞50,000個/ウェル、RNPの最終濃度は10nMとなる）。

5. 組織培養インキュベーター（37℃、5%CO<sub>2</sub>）中で48時間、形質移入複合体および細胞を含有するプレートを温置する。オンターゲット突然変異を検出するため、適切なプライマー（ドナー配列内のプライマーおよび標的挿入部位を囲むプライマー）とともにPCRを使用する。

#### 【0101】

実施例11：CRISPR/CAS9の特異性を試験するためのプロトコール

ビオチンに連結されるdCas9（DNA切断不活性Cas9）を作製する（dCas9-ビオチン）。Cas9（S.ピオゲネス（*S. pyogenes*）、S.アウレウス（*S. aureus*）など）。ビオチン化方法は以下に記載する。

#### 【0102】

ビオチン化方法#1：N-またはC-末端でaviタグ（約15残基）を操作し、WT（非タグ付加）タンパク質として発現させ、精製する。大腸菌（*E. coli*）ビオチンリガーゼ（BirA）およびビオチンを使用して、aviタグ付加Cas9をビオチン化する。本発明者らは、このスキームを使用してケモカインをビオチン化する。本発明者は、avi-タグ技術上のIPが数年前に失効したと考える。

#### 【0103】

ビオチン化方法 # 2 . 1 : サクシニミジル - エステルでのビオチン官能性は、表面露出リジン残基で組み込まれ得る ( 酵素反応は必要なし ) 。これは、C a s 9 と同程度の大きさのタンパク質の場合に実行可能な選択肢であり得る。

【 0 1 0 4 】

ビオチン化方法 # 2 . 2 : 同じように、ビオチン - マレイミドが市販されており、これらは表面露出システインで複合化され得る ( 酵素なし ) 。

【 0 1 0 5 】

D N A - 結合および切断に関してビオチン化 C a s 9 ダズ ( d o e s ) を特徴評価するために試験を遂行する。

【 0 1 0 6 】

ストレプトアビジン被覆 9 6 ウェルプレートは市販されているが、社内でも作製し得る。

【 0 1 0 7 】

d C a s 9 - ビオチンをプラスチックプレート ( 9 6 ウェル、2 4 ウェル、3 8 4 ウェルなど ) に結合させる。

【 0 1 0 8 】

各ウェルに設計したガイド R N A を提供する。ガイド R N A が C a s 9 タンパク質と相互作用するための時間を考慮する。

【 0 1 0 9 】

各ウェルにゲノム D N A または標的化される配列がある D N A を提供する。D N A への C a s 9 結合のための時間を考慮する。

【 0 1 1 0 】

適切な緩衝液でウェルを洗浄する。

【 0 1 1 1 】

アダプター ( D N A オリゴマー ) を提供する。結合のための時間を考慮する。

【 0 1 1 2 】

ゲノム D N A を制限消化して、扱い易く、アダプターを連結し易くする。

【 0 1 1 3 】

ウェルの洗浄

結合部位を同定するために D N A 配列決定を行う ( オンターゲット対オフターゲット )

【 0 1 1 4 】

実施例 1 2 : A B B I E 1 を介する N R F 2 編集

図 5 は、ガイド N r f 2 - s g R N A 2 および s g R N A 3 を用いた N r f 2 のエクソン 2 を標的とする A b b i e 1 遺伝子編集を示す。P C R は、A b b i e 1 編集を介したノックアウトについて、エクソン 2 標的化 N r f 2 遺伝子座に対するスクリーニングを行う。ガイド N r f 2 - s g R N A 2 および 3 を使用した N r f 2 のエクソン 2 を標的とする A b b i e 1 形質移入は、標的化領域でドナーの組み込みを示した。特有のバンドを 1 ~ 8 とする。

【 0 1 1 5 】

図 6 は、A b b i e 1 遺伝子編集により生成される理論的データを示す。s g R N A 1 ~ 3 を用いてゲノム物質を標的とするための A b b i e 1 系を介したドナー D N A 挿入を可視化する D N A ゲル電気泳動の代表例である。黒色バンドは、P C R 法に起因するバックグラウンド生成物に相当する。赤色バンドは、挿入物および挿入領域に隣接する遺伝子材料を増殖させることにより生成される特有の生成物に相当する。複数バンドは、標的化領域における可能性がある複数の挿入に相当する。

【 0 1 1 6 】

図 7 は、ガイド N r f 2 - s g R N A 3 を使用した、N r f 2 のエクソン 2 を標的とする A b b i e 1 遺伝子編集を示す。P C R は、A b b i e 1 編集を介したノックアウトについて、エクソン 2 標的化 N r f 2 遺伝子座に対するスクリーニングを行う。ガイド N r

10

20

30

40

50

F 2 - s g R N A 3 を使用した N r f 2 のエクソン 2 標的化から、ドナー配列および予想される挿入の近接部位に対して設計された P C R プライマーにより示されるようにドナー挿入が示唆された。特有のバンドを 1 ~ 4 とする。

【 0 1 1 7 】

図 8 は、プールされた H e k 2 9 3 T - 細胞における N r f 2 の A b b i e 1 ノックアウトを示す。( A ) プールされた H E K 2 9 3 T 集団における N r f 2 のノックアウトを示す、5 5 k D アイソフォームに対するポリクローナル抗体 ( S a n t a C r u z B i o ) を使用したウエスタンブロット分析である。( B ) G A P D H ( S a n t a C r u z B i o ) ロードニング対照である。

【 0 1 1 8 】

図 9 は、プールされた H e k 2 9 3 T - 細胞における N r f 2 の A b b i e 1 ノックアウトを示す。( A ) H E K 2 9 3 T - 細胞における N r f 2 プール集団のノックアウトを示す、N r f 2 に対するモノクローナル抗体 ( A b c a m ) を利用したウエスタンブロット分析である。( B ) G A P D H ロードニング対照である。( C ) 対照と比較した場合の発現比率の低下を示す濃度測定分析の平均である。

【 0 1 1 9 】

A b b i e 1 処理細胞は、ドナー D N A の組み込みを示す特有の P C R バンドを生じさせる。H E K 2 9 3 T プール細胞株におけるノックアウトの表現型確認は、特有かつ異なる抗体を用いた 2 種類のアイソフォームを探るウエスタンブロット分析を介して確認した。2 週未満でプール集団において組み込みによる約 8 0 % のノックアウトを観察した。

【 0 1 2 0 】

実施例 1 3 : A B B I E 1 を介した C X C R 4 編集

図 1 0 は、C X C R 4 エクソン 2 を標的とする A b b i e 1 遺伝子編集を示す。P C R は、A b b i e 1 を介して編集された C X C R 4 のエクソン 2 の標的化をスクリーニングする。関心のある領域に対して 4 セットのプライマーを設計した。セット番号 2 および 4 は、関心のある領域でのドナー D N A の組み込みを示唆する特有のバンドを生成したと思われる。

【 0 1 2 1 】

実施例 1 4 : A B B I E 1 を用いた N R F 2 遺伝子座におけるノックイン実験のための形質移入

注記 : 1 回の反応に対して 5 0 0 n g タンパク質および 1 2 0 n g の s g R N A を使用する。D N A の量はドナーコンストラクトのサイズに依存する。細胞に対して提供 / 形質移入 / エレクトロポレーションを行う前、その最中またはその後にドナー D N A ( L T R 配列がある D N A ) を A B B I E 1 とともに温置し得る。全反応物を滅菌バイオセイフティーカーキャビネット中で調製する。

【 0 1 2 2 】

第 1 日 : 1 0 % 胎仔ヒツジ血清 ( O m e g a S c i e n t i f i c ) を補給した 5 0 0  $\mu$  L D M E M ( G i b c o ) 中、ヒト胎児腎臓 ( H E K 2 9 3 T ) 細胞を H E K 2 9 3 T - 細胞 ( A T C C ) 2 0 0 , 0 0 0 個 / ウェルで 2 4 ウェル培養プレート ( C o r n i n g ) に播種した。細胞を 2 4 時間回復させた。

【 0 1 2 3 】

第 2 日 : A B B I E 1 調製 :

チューブ 1 :

室温で 1 0 分間、1 : 1 モル比の血清減少形質移入培地 ( O p t i M E M 、 L i f e T e c h n o l o g i e s ) 中の精製 A B B I E 1 タンパク質 ( 配列番号 5 8 ) およびドナー D N A ( 配列番号 1 0 1 ) 。 s g R N A を 1 . 3 倍モル過剰 ( およそ 1 2 0 n g ) まですタンパク質 / D N A 複合体に添加し、室温でさらに 1 0 分間温置を継続する。この混合物の体積は 2 5  $\mu$  L である。

【 0 1 2 4 】

チューブ 2 :

10

20

30

40

50



2  $\mu$ L の形質移入試薬 (RNAiMAX、Life Technologies) を 2  
3  $\mu$ L の OptiMEM に添加した。室温で 10 分間温置した。

【0125】

チューブ 1 および試験管 2 (50 mL 最終体積) を合わせ、室温で 15 分間温置した。

【0126】

50  $\mu$ L 形質移入混合物全てをウェルに添加した。

【0127】

プール集団中でのゲノム DNA 編集の検証のために、プールされた編集細胞の半分を形  
質移入から 48 時間後に回収した。編集されたゲノムの検証は、ポリメラーゼ連鎖反応 (PCR) によって行った。本発明者らは、以下に記載のように、標的化領域に対して PCR  
R を行い (PCR プロトコルを参照されたい)、残りは、6 cm の培養皿 (Corning) 上に播種し、48 時間にわたり回復させた。

10

【0128】

第 5 日: ウエスタンブロッティングを介した表現型変化のスクリーニング

Nrf2 アイソフォームに対して、55 kD アイソフォームおよび 98 kD アイソフ  
ームを標的とする一次抗体 (55 kD アイソフォームに対するものは Santa Cruz  
Biotechnology, sc-722) (98 kD アイソフォームに対するも  
のは Abcam, ab-62352) を使用して、標準的なウエスタンブロット分析を行  
った。GAPDH (Santa Cruz Biotechnology, sc-519  
07)。

20

【0129】

実施例 15: Nrf2 および CXCR4 遺伝子座に対する Abbie1 を用いた遺伝子編  
集の検出のための PCR 条件

ヒト Nrf2 に対する受入番号

Uniprot: Q16236

Ensembl 遺伝子 ID: ENSG00000116044

【0130】

Nrf2 (エクソン 2) に対する標的配列および PAM の編集: sgRNA に対して設  
計 1 ~ 3 を使用した。

GCGACGGAAAGAGTATGAGC TGG

30

TATTTGACTTTCAGTCAGCGA CGG

TGGAGGCAAGATATAGATCT TGG

【0131】

Nrf2 標的での組み込みの検出のためのプライマーキー

プライマーセット 1: プライマー 1: 5' - GTGTTAAATTTCAAAACATCAG  
CAGC - 3'、プライマー 2: 5' - GACAAAGACATCCTTGATTTG - 3

,

プライマーセット 2: プライマー 1: 5' - GAGGTTTGACTGTGTAAATG -  
3'、プライマー 2: 5' - GATACCAGAGTCAACACAACAG - 3'

プライマーセット 3: プライマー 1: 5' - TCTACATTAATTCCTCTTGTC  
C - 3'、プライマー 2: 5' - GATACCAGAGTCAACACAACAG - 3'

40

【0132】

ヒト CXCR4 に対する受入番号

Uniprot P61073

Ensembl 遺伝子 ID: ENSG00000121966

【0133】

CXCR4 (エクソン 2) に対する標的配列および PAM の編集: sgRNA に対して  
設計 1 を使用した。

GGGCAATGGATTGGTCATCC TGG

【0134】

50

C X C R 4 標的での組み込みの検出のためのプライマーキー

プライマーセット 1 : プライマー 1 : 5' - T C T A C A T T A A T T C T C T T G T G  
C - 3'、プライマー 2 : 5' - G A C A A G A C A T C C T T G A T T T G - 3'  
プライマーセット 2 : プライマー 1 : 5' - T C T A C A T T A A T T C T C T T G T G  
C - 3'、プライマー 2 : 5' - G A T A C C A G A G T C A C A C A A C A G - 3'  
プライマーセット 3 : プライマー 1 : 5' - G A G G T T G A C T G T G T A A A T G -  
3'、プライマー 2 : 5' - G A C A A G A C A T C C T T G A T T T G - 3'  
プライマーセット 4 : プライマー 1 : 5' - G A G G T T G A C T G T G T A A A T G -  
3'、プライマー 2 : 5' - G A T A C C A G A G T C A C A C A A C A G - 3'

10

組み込まれたドナーDNAの検出のために使用した PCR サイクル条件

\*注記 アニール温度はプライマー配列に依存して変動する

1.最初の変性:	4 分	94℃
2.変性:	30 秒	94℃
3.アニール:	30 秒	55℃
4.伸長:	30 秒	72℃
5.段階 2 へ:	40 サイクル	
6.最終伸長:	4 分間	72℃
7.最終保持:	∞	4℃

20

#### 【 0 1 3 5 】

ビオチン化のための A v i タグ付加 C a s 9

C a s 9 ビオチン化のために使用した a v i タグの配列

アミノ酸配列:

G G D L E G S G L N D I F E A Q K I E W H E \*

30

#### 【 0 1 3 6 】

核酸配列:

GGCGGCGACCTCGAGGGTAGCGGTCTGAACGATATTTTGAAGCGCA

GAAAATTGAATGGCATGAATAA

#### 【 0 1 3 7 】

第一の下線セクション = C a s 9 C - 末端

斜字体セクション = 制限部位 / リンカー

第 2 の下線セクション = a v i タグ ( ビオチン化部位にハイライト )

40

#### 【 0 1 3 8 】

実施例 16 : A B B I E 1 融合タンパク質に対する発現プロトコール

全長融合タンパク質 ( 配列番号 57 ) を含有する発現コンストラクトの形質転換。

- 80 冷凍庫からコンピテント大腸菌 ( E . c o l i ) 細胞を取り出す。

水浴にスイッチを入れ、42 にする。

コンピテント細胞を 1 . 5 m L チューブ ( エッペンドルフまたは同様のもの ) に入れる。DNA コンストラクトの形質転換を行うために 50 μ L のコンピテント細胞を使用する。

チューブを氷上で維持する。

50

50 ng の環状 DNA を大腸菌 (E. coli) 細胞に添加する。氷上で 10 分間温置して、コンピテント細胞を凍結融解する。

DNA および大腸菌 (E. coli) 入りのチューブを 42 °C の水浴に 45 秒間入れる。大腸菌 (E. coli) 細胞に対するダメージを減少させるためにチューブを氷上に 2 分間戻す。

1 mL の LB (抗生物質添加なし) を添加する。チューブを 37 °C で 1 時間温置する (チューブを 30 分間温置し得る)。

適切な抗生物質入りの LB プレート上に約 100 µL の得られた培養物を広げる。

約 12 ~ 16 時間後にコロニーを拾う。

#### 【0139】

接種および増殖

LB および抗生物質を含有する 1 L フラスコに接種する。

0.6 OD に到達するまで細菌培養物を増殖させ、次いで 1 mM 最終濃度のイソプロピル β-D-1-チオガラクトピラノシド (IPTG) で誘導する。

6 ~ 8 時間にわたり培養物を増殖させ、最低 2000 G の力で 10 分間、懸濁細菌培養物を遠心分離する。

後のさらなる処理のためにペレットを -80 °C で凍結させる。

#### 【0140】

タンパク質調製および精製

全段階を室温で行う。

#### 【0141】

20 mM Tris pH 8.0、300 mM NaCl、0.1 mg/mL ニワトリ卵白リゾチーム中での 2 サイクルの凍結 - 融解によって細胞を溶解させる。6,000 g で 15 分間遠心分離し、上清を保持する。

#### 【0142】

20 mM Tris pH 8.0、300 mM 塩化ナトリウム中で平衡化した Ni-IDA アガロースカラム上に上清を載せる。イミダゾールの 0 ~ 200 mM 勾配でタンパク質を溶出させる。7% SDS-PAGE によって融合タンパク質を含有する分画を同定する。

#### 【0143】

分画をプールし、20 mM Tris pH 8.0 で希釈して、最終 NaCl 濃度が 50 mM になるようにする。Q-セファロースカラム上に載せ、塩化ナトリウムの 0 ~ 500 mM 勾配で溶出させる。7% SDS-PAGE によって融合タンパク質を含有する分画を同定する。

#### 【0144】

分画をプールし、20 mM Tris pH 8.0 で希釈して、最終 NaCl 濃度が 100 mM になるようにする。SP-セファロースカラム上に載せ、塩化ナトリウムの 0 ~ 500 mM 勾配で溶出させる。7% SDS-PAGE によって融合タンパク質を含有する分画を同定する。

#### 【0145】

分画をプールし、280 nm でのその UV 吸収によって濃度を測定し、400 µg/mL の最終濃度まで遠心フィルターによって濃縮する。50% の最終濃度になるようにグリセロールを添加する。-20 °C で保存する。

#### 【0146】

本明細書中で一定の実施形態を示し、記載してきたが、このような実施形態が単なる例として提供されることは当業者にとって明らかであろう。本開示からの逸脱なく、当業者にとって多くの変形形態、変化形態および置換形態が想起されるであろう。本開示の実施において、本明細書中に記載の開示の実施形態に対して様々な代替物が使用され得ることを理解されたい。続く特許請求の範囲は、本開示の範囲を定め、これらの特許請求の範囲内の方法および構造ならびにそれらの均等物がそれにより包含されるものとする。

10

20

30

40

50

## 【 0 1 4 7 】

## 本開示の配列

以下で提供される各配列に対して以下の情報が提供される：配列のタイプ（核酸またはアミノ酸）、起源（例えば、大腸菌（E . c o l i ））、長さおよび識別番号（可能な場合）。

## 【 0 1 4 8 】

本開示の第 1 のポリヌクレオチドは、例えば C a s 9、C p f 1、T A L E または Z n F n タンパク質をコードし得る。本開示の第 2 のポリヌクレオチドは、例えばインテグラーゼ、トランスポサゼまたはレコンビナーゼをコードし得る。本明細書中に記載の組成物（コンストラクト、融合タンパク質）および方法中で使用され得る代表的な第 1 および第 2 のポリヌクレオチド配列およびタンパク質配列を代表的なリンカー配列とともに以下で列挙する。本開示において、以下の表 1 で列挙されない他のポリヌクレオチド配列、タンパク質配列またはリンカー配列が提供され得るが、本明細書中に記載の組成物（コンストラクト、融合タンパク質）および方法において使用され得る。例えば、配列番号 4 9、配列番号 5 7、配列番号 5 8 および / またはそれらの一部である。

10

## 【 0 1 4 9 】

リンカーは、あらゆる長さ、例えば 3 ~ 3 0 0 ヌクレオチド長、6 ~ 6 0 ヌクレオチド長または第 1 および第 2 のポリヌクレオチドが融合可能であるあらゆる長さであり得る。ポリペプチドは、生物、例えば大腸菌（E . c o l i ）により産生され得るか、もしくは合成によるか、または両方の組み合わせで作製され得る。

20

## 【 0 1 5 0 】

代表的な核酸配列：1、3、5、7、9、11、13、15、17、19、21、23、27 ~ 47、49、55、56、57、62、64、66、68、70、79、82 および 83。

## 【 0 1 5 1 】

代表的なアミノ酸配列：2、4、6、8、10、12、14、16、18、20、22、24、25、26、48、50、52、58、63、65、67、69、71、72 ~ 78 および 80。

表 1: 第 1 のタンパク質、第 2 のタンパク質またはリンカー

30

第1のポリヌクレオチドまたは タンパク質配列 配列番号	第2のポリヌクレオチドまたは タンパク質配列 配列番号	リンカー配列 配列番号 または配列
1-14, 27-46, 50, 52, 56, 57, 68, 69, 72-78, 86-92, 200- 253	15-26, 47, 48, 55, 57, 62- 67, 70, 71, 79, 80	51, 54, 61,  GGS

40

表 2:配列の部分的リスト

遺伝子 (タンパク質)	細菌/ウイルス	DNA 配列	タンパク質 配列	配列番号 (DNA、 タンパク質)
Cas9	S.サーモフィルス (S.thermophilus)	HQ712120.1	Q03JI6.1	配列番号 1、2
	P.ムルトシダ (P.multocida)		Q9CLT2.1	配列番号 3、4
	S.ミュータンス (S.mutans)		Q8DTE3.1	配列番号 5、6
	N.メニンギティデス (N.meningitides)		C9X1G5.1	配列番号 7、8
	S.ミティス (S.mitis)		KJQ69483.1	配列番号 9、10
	S.マカカ (S.macacae)		EHJ52063.1	配列番号 11、12
インテグラーゼ	スタフィロコッカス・アウレウス (Staphylococcus aureus)		KKJ92487.1	配列番号 49、50
	S.ピオゲネス (S.pyogenes)		AFV37892.1	配列番号 13、14
	HIV1		ABR68182.1	配列番号 15、16
	サルT-リンパ球 ウイルス		AAA47841.1	配列番号 17、18
	S.ニューモニエ(S.pneumoniae)		CBW38769.1	配列番号 19、20
	大腸菌(E.coli)		CAA41325.1	配列番号 21、22
レコンピナーゼ	レンチウイルス			配列番号 47、48
	サーモアナエロバクテリウム (Thermoanaerobacterium) ファージ		YP_006546326.1	配列番号 23、24

10

20

30

40

## 【 0 1 5 2 】

さらなる配列

配列番号 1

名称：S . サーマフィルス ( S . t h e r m o p h i l u s ) C s n 1 c d s H Q 7  
1 2 1 2 0 . 1

配列：

ATGACTAAGCCATACTCAATTGGACTTGATATTGGAACGAATAGTGTTGGAT  
GGGCTGTAATAACTGATAATTACAAGGTTCCGTCTAAAAAATGAAAGTCTT  
AGGAAATACGAGTAAAAAGTATATCAAAAAGAACCTGTTAGGTGTATTACTC  
TTTGA CTCTGGAATCACAGCAGAAGGAAGAAGATTGAAGCGTACTGCAAGAA  
GACGTTATACTAGACGCCGTAATCGTATCCTTTATTTGCAGGAAATTTTTAGC  
ACGGAGATGGCTACATTAGATGATGCTTTCTTTCAAAGACTTGACGATTCGTT  
TTTAGTTCCTGATGATAAACGTGATAGTAAGTATCCGATATTTGGAACTTAG  
TAGAAGAAAAAGTCTATCATGATGAATTTCCAACCTATCTATCATTTAAGGAA  
ATATTTAGCAGATAGTACTAAAAAAGCAGATTTGCGTCTAGTTTATCTTGCAT  
TGGCTCATATGATTAAATATAGAGGTCACCTTCTTAATTGAAGGAGAGTTTAAT  
TCAAAAAATAATGATATTCAGAAGAATTTTCAAGACTTTTTGGACACTTATAA  
TGCTATTTTTGAATCGGATTTATCACTTGAGAATAGTAAACAACCTTGAGGAAA  
TTGTTAAAGATAAGATTAGTAAATTAGAAAAGAAAGATCGTATTTTAAACT  
CTTCCCTGGGGAGAAGAATTCGGGGATTTTTTTCAGAGTTTCTAAAGTTGATTG  
TAGGAAATCAAGCTGATTTTAGGAAATGTTTTAATTTAGACGAAAAAGCCTC  
CTTACATTTTTCCAAAGAAAGCTATGATGAAGATTTAGAGACTTTGTTAGGTT  
ATATTGGAGATGATTACAGTGATGTCTTTCTCAAAGCAAAGAACTTTATGAT  
GCTATTCTTTTATCGGGTTTTCTGACTGTAAGTATAATGAGACAGAAGCACC  
TCTCTCTTCTGCTATGATAAAGCGATATAATGAACACAAAGAAGATTTAGCGT  
TACTAAAGGAATATATAAGAAATATTTCACTAAAAACGTATAATGAAGTATT

10

20

30

TAAAGATGACACCAAAAATGGTTATGCTGGTTATATTGATGGAAAAACAAAT  
CAGGAAGATTTCTACGTATATCTAAAAAACCTATTGGCTGAATTTGAAGGTG  
CGGATTATTTTCTTGAAAAAATTGATCGAGAAGATTTTTTGAGAAAGCAACGT  
ACATTTGACAATGGTTCGATACCATATCAGATTCATCTTCAAGAAATGAGAG  
CAATTCTTGATAAGCAAGCTAAATTTTATCCTTTCTTGGCTAAAAATAAAGAA  
AGAATCGAGAAGATTTTAACCTTCCGAATTCCTTATTATGTAGGTCCACTTGC  
GAGAGGGAATAGTGATTTTGCCTGGTCAATAAGAAAACGAAATGAAAAAATT  
ACACCTTGGAATTTTGAGGACGTTATTGACAAAGAATCTTCGGCAGAGGCTTT  
CATTAAATCGAATGACTAGTTTTGATTTGTATTTGCCAGAAGAGAAGGTACTTC  
CAAAGCATAGTCTCTTATACGAACTTTTAATGTATATAATGAATTAACAAAA  
GTTAGATTTATTGCCGAAAGTATGAGAGATTATCAATTTTTAGATAGTAAGCA  
GAAGAAAGATATTGTTAGACTTTATTTTAAAGATAAAAGGAAAGTTACTGAT  
AAGGATATTATTGAATATTTACATGCAATTTATGGGTATGATGGAATTGAATT  
AAAAGGCATAGAGAAACAGTTTAATTCTAGTTTATCTACTTATCACGATCTTT  
TAAATATTATTAATGATAAAGAGTTTTTGGATGATAGTTCAAATGAAGCGATT  
ATCGAAGAAATTATCCATACTTTGACAATTTTTGAAGATAGAGAGATGATAA  
AACAAACGTCTTTCAAAATTTGAGAATATATTCGATAAATCCGTTTTGAAAAAG  
TTATCTCGTAGACATTACACTGGCTGGGGTAAGTTATCTGCTAAGCTTATTAA  
TGGTATTCGAGATGAAAAATCTGGTAATACTATTCTTGATTACTTAATTGATG  
ATGGTATTTCTAACCGTAATTTTCATGCAACTTATTCACGATGATGCTCTTTCTT  
TTAAAAAGAAGATACAGAAAGCACAAATTATTGGTGACGAAGATAAAGGTA  
ATATTAAAGAGGTCGTTAAGTCTTTGCCAGGTAGTCCTGCGATTAAAAAAGG  
TATTTTACAAAGCATAAAAATTGTAGATGAATTGGTCAAAGTAATGGGAGGA  
AGAAAACCCGAGTCAATTGTTGTTGAGATGGCTCGTGAAAATCAATATACCA  
ATCAAGGTAAGTCTAATTCCCAACAACGCTTGAAACGTTTAGAAAAATCTCT  
CAAAGAGTTAGGTAGTAAGATACTTAAGGAAAATATTCCTGCAAACTTTCT  
AAAATAGACAATAACGCACTTCAAAATGATCGACTTTACTTATACTATCTTCA  
AAATGGAAAAGATATGTATACCGGAGATGATTTAGATATTGATAGATTAAAGT  
AATTATGATATTGATCATATTATTCCTCAAGCTTTTTTGAAAGATAATTCTATT  
GACAATAAAGTACTTGTTTCATCTGCTAGTAACCGTGGTAAATCAGATGATTT  
TCCAAGTTTAGAGGTTGTCAAAAAAAGAAAGACATTTTGGTATCAATTATTG

10

20

30

40

AAATCAAAATTAATTTCTCAACGAAAATTTGATAATCTGACAAAAGCTGAAC  
GGGGAGGATTGTTACCTGAGGACAAAGCTGGTTTTATTCAACGCCAGTTGGT  
TGAAACACGTCAAATAACAAAACATGTAGCTCGTTTACTTGATGAGAAATTT  
AATAATAAAAAAGATGAAAATAATAGAGCGGTACGAACAGTAAAAATTATT  
ACCTTGAAATCTACCTTAGTTTTCTCAATTTTCGTAAGGATTTTGAACTTTATAA  
AGTTCGTGAAATCAATGATTTTCATCATGCTCATGATGCTTACTTGAATGCCG  
TTATAGCAAGTGCTTTACTTAAAGAAATACCCTAACTAGAGCCAGAATTTGTG  
TACGGTGATTATCCAAAATACAATAGTTTTAGAGAAAGAAAGTCCGCTACAG  
AAAAGGTATATTTCTATTCAAATATCATGAATATCTTTAAAAAATCTATTTCT  
TTAGCTGATGGTAGAGTTATTGAAAGACCACTTATTGAGGTAAATGAGGAGA  
CCGGCGAATCCGTTTGAATAAAGAATCTGATTTAGCAACTGTAAGGAGAGT  
ACTCTCTTATCCGCAAGTAAATGTTGTGAAAAAAGTTGAGGAACAGAATCAC  
GGATTGGATAGAGGAAAACCAAAGGGATTGTTTAATGCAAATCTTTCCTCAA  
AGCCAAAACCAAATAGTAATGAAAATTTAGTAGGGTGCTAAAGAGTATCTTGA  
CCCCAAAAGTATGGGGGGTATGCTGGAATTTCTAATTCTTTTGCTGTTCTTG  
TTAAAGGGACAATTGAAAAAGGTGCTAAGAAAAAAATAACAAATGTACTAG  
AATTTCAAGGTATTTCTATTTTAGATAGGATTAATTATAGAAAAGATAAACTT  
AATTTTTTACTTGAAAAAGGTTATAAAGATATTGAGTTAATTATTGAACTACC  
TAAATATAGTTTATTTGAACTTTCAGATGGTTCACGTCGTATGTTGGCTAGTA  
TTTTGTCAACGAATAATAAGAGGGGAGAGATTCACAAAGGAAATCAGATTTT  
TCTTTCACAGAAGTTTGTGAAATTACTTTATCATGCTAAGAGAATAAGTAACA  
CAATTAATGAGAATCATAGAAAATATGTTGAGAACCATAAAAAAGAGTTTGA  
AGAATTATTTTACTACATTCTTGAGTTTAATGAGAATTATGTTGGAGCTAAAA  
AGAATGGTAAACTTTTAACTCTGCCTTTCAATCTTGGCAAAATCATAGTATA  
GATGAACTCTGTAGTAGTTTTATAGGACCTACCGGAAGTGAAAGAAAGGGGC  
TATTTGAATTAACCTCTCGTGGAAGTGCTGCTGATTTTGAATTTTATAGGTGTTA  
AAATTCCAAGGTATAGAGACTATACCCCATCATCCCTATTAAAAGATGCCAC  
ACTTATTCATCAATCTGTTACAGGCCTCTATGAAACACGAATAGACCTTGCCA  
AACTAGGAGAGGGTTAA

10

20

30

40

【 0 1 5 3 】

配列番号 2

配列：



MTKPYSIGLDIGTNSVGWAVITDNYKVPSKKMKVLGNTSKKYIKKNLLGVLLFD  
SGITAEGRRLKRTARRRYTRRRNRILYLQEIFSTEMATLDDAFFQRLDDSFLVPDD  
KRDSKYPIFGNLVEEKVYHDEFPTIYHLRKYLADSTKKADLRLVYLALAHMIKY  
RGHFLIEGEFNSKNNDIQKNFQDFLDTYNAIFESDLSLENSKQLEEIVKDKISKLEK  
KDRILKLFPGEKNSGIFSEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLET  
LLGYIGDDYSDVFLKAKKLYDAILLSGFLTVDNETEAPLSSAMIKRYNEHKEDL  
ALLKEYIRNISLKTYNVFKDDTKNGYAGYIDGKTNQEDFYVYLKNLLAEFEGA  
DYFLEKIDREDFLRKQRTFDNGSIPYQIHLQEMRAILDKQAKFYPFLAKNKERIEK  
ILTRIPYYVGPLARGNSDFAWSIRKRNEKITPWNFEDVIDKESSAEAFINRMTSF  
DLYLPEEKVLPKHSLLYETFNVYNELTKVRFIAESMRDYQFLDSKQKKDIVRLYF  
KDKRKVTDKDIIEYLHAIYGYDGIELKGIEKQFNSSLSTYHDLLNIINDKEFLDDSS  
NEAIIIEIIHTLTIFEDREMIKQRLSKFENIFDKSVLKKLSRRHYTGWGKLSAKLIN  
GIRDEKSGNTILDYLIIDDGISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKGNIKEV  
VKSLPGSPAIIKKGILQSIKIVDELVKVMGGRKPESIVVEMARENQYTNQGKSNSQ  
QRLKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYYLQNGKDMYTGD  
DLIDIRLSNYDIDHIIPQAFKLDNSIDNKVLVSSASNRGKSDDFPSLEVVKRKRTF  
WYQLLSKSLISQRKFDNLTKAERGGLLPEDKAGFIQRQLVETRQITKHVARLLDE  
KFNNKKDENNRVVRTVKIITLKSTLVSQFRKDFELYKPREINDFHHAHDAYLNA  
VIASALLKKYPKLEPEFVYGDYPKYNSFRERKSATEKVYFYSNIMNIFKKSISLAD  
GRVIERPLIEVNEETGESVWNKESDLATVRRVLSYPQVNVVKKVEEQNHGLDRG  
KPKGLFNANLSSKPKPNSNENLVGAKEYLDPKKYGGYAGISNSFAVLVKGTIEK  
GAKKKITNVLEFQGISILDRINYRKDKLNFLLKGYKDIELIIELPKYSLFELSDGSR  
RMLASILSTNNKRGEIHKGNQIFLSQKFVKLLYHAKRISNTINENHRKYVENHKK  
EFEELFYIILEFNENYVGAKKNGKLLNSAFQSWQNHSIDELCSSFIGPTGSEKGL  
FELTSRGSAAADFEFLGVKIPRYRDYTPSSLLKDATLIHQSVTGLYETRIDLAKLGE  
G

10

20

30

40

【 0 1 5 4 】

配列番号 3

名称：P . ムルトシダ ( P . m u l t o c i d a ) C a s 9

配列：

ATGCAAACAACAAATTTAAGTTATATTTTAGGTTTAGATTTGGGGATCGCTTC  
TGTAGGTTGGGCTGTCGTTGAAATCAATGAAAATGAAGACCCTATCGGCTTG

ATTGATGTAGGAGTAAGGATATTTGAGCGTGCTGAGGTACCCAAAACCTGGAG  
AATCTTTAGCACTCTCTCGCCGTCTTGCAAGAAGTACTCGCCGTTTGATACGC  
CGTCGTGCACACCGTTTACTCCTCGCAAAACGCTTCTTAAACGTGAAGGTAT  
ACTTTCCACAATCGACTTAGAAAAAGGATTACCCAACCAAGCTTGGGAATTA  
CGTGTCGCCGGTCTTGAACGTCGGTTATCCGCCATAGAATGGGGTGCGGTTCT  
GCTACATTTAATCAAGCATCGAGGTTATCTTTCTAAACGTAAAAATGAATCCC  
AAACAAACAACAAAGAATTAGGAGCCTTACTCTCTGGAGTGGCACAAAACCA  
TCAATTATTACAATCAGATGACTACCGAACACCAGCAGAGCTCGCACTGAAA  
AAATTTGCTAAAGAAGAAGGGCATATCCGTAATCAACGAGGTGCCTATACAC  
ATACATTTAATCGATTAGACTTATTAGCTGAACTTAACTTGCTTTTTTGCTCAAC  
AACATCAGTTTGGTAACCCTCACTGTAAAGAGCATATTCAACAATATATGAC  
AGAATTGCTTATGTGGCAAAAGCCAGCCTTATCTGGTGAGGCAATTTTAAAA  
ATGTTGGGTAAATGTACGCATGAAAAAAATGAGTTTAAAGCAGCAAAACATA  
CCTACAGTGCGGAGCGCTTTGTTTGGCTAACCAAACCTCAATAACTTGCGCATT  
TTAGAAGATGGGGCAGAACGAGCTCTTAATGAAGAAGAACGTCAACTATTGA  
TAAATCATCCGTATGAGAAATCAAAATTAACCTATGCCCAAGTCAGAAAATT  
GTTAGGGCTTTCCGAACAAGCGATTTTTTAAGCATCTACGTTATAGTAAAGAA  
AACGCAGAATCAGCTACTTTTATGGAGCTTAAAGCTTGGCATGCAATTCGTA  
AAGCGTTAGAAAATCAAGGATTGAAGGATACTTGGCAAGATCTCGCTAAGAA  
ACCTGACTTACTAGATGAAATTGGTACCGCATTTTCTCTTTATAAAACTGATG  
AAGATATTCAGCAATATTTGACAAATAAGGTACCGAACTCAGTCATCAATGC  
ATTATTAGTTTCTCTGAATTTTCGATAAATTCATTGAGTTATCTTTGAAAAGTTT  
ACGTAAAATCTTGCCCCTAATGGAGCAAGGTAAGCGTTATGATCAAGCTTGT  
CGTGAAATTTATGGGCATCATTATGGTGAGGCAAATCAAAAAACTTCTCAGC  
TACTACCAGCTATTCCAGCCCAAGAAATTCGTAATCCTGTTGTTTTACGTACA  
CTTTCACAAGCACGTAAAGTGATCAATGCCATTATTCGTCAATATGGTTCCCC  
TGCTCGAGTCCATATTGAAACAGGAAGAGAACTTGGGAAATCTTTTAAAGAA  
CGTCGTGAAATTCAAAAACAACAGGAAGATAATCGAACTAAGCGAGAAAGT  
GCGGTACAAAAATTCAAAGAATTATTTTCTGACTTTTCAAGTGAACCCAAAA  
GTAAAGATATTTTAAAATTCCGCTTATACGAACAACAGCATGGTAAATGCTT  
ATACTCTGGAAAAGAGATCAATATTCATCGCTTAAATGAAAAGGGTTATGTG

10

20

30

40

GAAATTGATCATGCTTTACCTTTCTCACGGACTTGGGATGATAGTTTTAATAA  
TAAAGTATTAGTTCTTGCCAGCGAAAACCAAACAAAGGGAATCAAACACCG  
TATGAATGGCTACAAGGTAAAATAAATTCGGAACGTTGGAAAACTTTGTTG  
CTTTAGTACTGGGTAGCCAGTGCAGTGCAGCCAAGAAACAACGATTACTCAC  
TCAAGTTATTGATGATAATAAATTTATTGATAGAACTTAAATGATACTCGCT  
ATATTGCCCCGATTCTATCCAATATATTCAAGAAAATTTGCTTTTGGTGGGT  
AAAAATAAGAAAAATGTCTTTACACCAAACGGTCAAATTACTGCATTATTAA  
GAAGTCGCTGGGGATTAATTAAGGCTCGTGAGAATAATAACCGTCATCATGC  
TTTAGATGCGATAGTTGTGGCTTGTGCAACACCTTCTATGCAACAAAAAATTA  
CCCGATTTATTTCGATTTAAAGAAGTGCATCCATACAAAATAGAAAATAGGTA  
TGAAATGGTGGATCAAGAAAGCGGAGAAATTATTTACCTCATTTTCCTGAA  
CCTTGGGCTTATTTTAGACAAGAGGTTAATATTCGTGTTTTTGATAATCATCC  
AGATACTGTCTTAAAAGAGATGCTACCTGATCGCCCACAAGCAAATCACCAG  
TTTGTACAGCCCCTTTTTGTTTTCTCGTGCCCCAACTCGTAAAATGAGTGGTCA  
AGGGCATATGGAAACAATTAAATCAGCTAAACGCTTAGCAGAAGGCATTAGC  
GTTTTAAGAATTCCTCTCACGCAATTAAAACCTAATTTATTGGAAAATATGGT  
GAATAAAGAACGTGAGCCAGCACTTTATGCAGGACTAAAAGCACGCTTGGCT  
GAATTTAATCAAGATCCAGCAAAAGCGTTTGCTACGCCTTTTTATAACAAG  
GAGGGCAGCAGGTCAAAGCTATTCGTGTTGAACAGGTACAAAAATCAGGGGT  
ATTAGTCAGAGAAAACAATGGGGTAGCAGATAATGCCTCTATCGTTCGAACA  
GACGTATTTATCAAAAATAATAAATTTTTCTTGTTCCTATCTATACTTGGCA  
AGTTGCGAAAGGCATCTTGCCAAATAAAGCTATTGTTGCTCATAAAAATGAA  
GATGAATGGGAAGAAATGGATGAAGGTGCTAAGTTTAAATTCAGCCTTTTCC  
CGAATGATCTTGTCGAGCTAAAAACCAAAAAAGAATACTTTTTTCGGCTATTA  
CATCGGACTAGATCGTGCAACTGGAAACATTAGCCTAAAAGAACATGATGGT  
GAGATATCAAAAGGTAAAGACGGTGTTTACCGTGTTGGTGTCAAGTTAGCTC  
TTTCTTTTGAAAAATATCAAGTTGATGAGCTCGGTAAAAATAGACAAATTTGC  
CGACCTCAGCAAAGACAACCTGTGCGTTAA

10

20

30

40

【 0 1 5 5 】

配列番号 4

配列：

MQTTNLSYILGLDLGASVGWAVVEINENEDPIGLIDVGVRIFERA EVPKTGESLA  
LSRRLARSTRRLIRRAHRLLLAKRFLKREGILSTIDLEKGLPNQAWELRVAGLER  
RLSAIEWGAVLLHLIKHRGYLSKRKNESQTNKELGALLSGVAQNHQLLQSDDY  
RTPAELALKKFAKEEGHIRNQRGAYTHTFNRLDLLAELNLLFAQQHQFGNPHCK  
EHIQQYMTTELLMWQKPALSGEAILKMLGKCTHEKNEFKAAKHTYSAERFVWLT  
KLNNLRILEDGAERALNEEERQLLINHPYEKSKLTYAQVRKLLGLSEQAIFKHLR  
YSKENAESATFMELKA WHAIRKALENQGLKDTWQDLAKKPDLLDEIGTAFSLY  
KTDEDIQQYL TNKVPNSVINALLVSLNFDKFIELSLKSLRKILPLMEQ GKRYDQAC  
REIYGHYGEANQKTSQLLPAIPAQEIRNPVVLRTLSQARKVINAIIRQY GSPARV  
HIETGRELGKSFKERREIQKQ QEDNRTKRESAVQKFKELFSDFSSEPKSKDILKFR  
LYEQQH GKCLYSGKEINIHRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLASEN  
QNKGNQTPYEWLQ GKINSERWKNFVALVLGSQCSAAKKQRLLTQVIDDNKFID  
RNLNDTRYIARFLSNYIQENLLL VGKNKKNVFTPNGQITALLRSRWGLIKAREN N  
NRHHALDAIVVACATPSMQQKITRFIRFKEVHPYKIENRYEMVDQESGEIISP HFP  
EPWAYFRQEVNIRVFDNHPD TVLKEMLPDRPQANHQFVQPLFVSRAPTRKMSG  
QGHMETIKSAKRLAEGISVLR IPTQLKPNLLENMVNKEREPALYAGLKARLAEF  
NQDPAKAFATPFYKQGGQ QVKAIRVEQVQKSGVLVRENNGVADNASIVRTDVFI  
KNNKFFLVPIYTWQVAKGILPNKAIVAHKNEDEWEEMDEGAKFKFSLFPNDLVE  
LKTKKEYFFGYI GLDRATGNISLKEHDGEISKGKDG VYRVGVKLALSFEKYQV  
DELGKNRQICRPQQRQPVR

10

20

30

【 0 1 5 6 】

配列番号 5

名称：S . ミュータンス ( S . m u t a n s ) C a s 9

配列：

ATGAAAAAACCTTACTCTATTGGACTTGATATTGGAACCAATTCTGTTGGTTG  
GGCTGTTGTGACAGATGACTACAAAGTTCCTGCTAAGAAGATGAAGGTTCTG  
GGAAATACAGATAAAAGTCATATCGAGAAAAATTTGCTTGGCGCTTTATTAT  
TTGATAGCGGGAATACTGCAGAAGACAGACGGTTAAAGAGAACTGCTCGCCG  
TCGTTACACACGTCGCAGAAATCGTATTTTATATTTGCAAGAGATTTTTTTCAG  
AAGAAATGGGCAAGGTAGATGATAGTTTCTTTCATCGTTTAGAGGATTCTTTT  
CTTGTTACTGAGGATAAACGAGGAGAGCGCCATCCCATTTTTGGGAATCTTG  
AAGAAGAAGTTAAGTATCATGAAAATTTCCAACCATTATCATTTGCGGCA

40

ATATCTTGCGGATAATCCAGAAAAAGTTGATTTGCGTTTAGTTTATTTGGCTT  
TGGCACATATAATTAAGTTTAGAGGTCATTTTTTAATTGAAGGAAAGTTTGAT  
ACACGCAATAATGATGTACAAAGACTGTTTCAAGAATTTTAGCAGTCTATG  
ATAATACTTTTGAGAATAGTTCGCTTCAGGAGCAAAATGTTCAAGTTGAAGA  
AATTCTGACTGATAAAATCAGTAAATCTGCTAAGAAAGATAGAGTTTTGAAA  
CTTTTTCTAATGAAAAGTCTAATGGCCGCTTTGCAGAATTTCTAAAATAAT  
TGTTGGTAATCAAGCTGATTTTAAAAAGCATTTTGAATTAGAAGAGAAAGCA  
CCATTGCAATTTTCTAAAGATACTTATGAAGAAGAGTTAGAAGTACTATTAGC  
TCAAATTGGAGATAATTACGCAGAGCTCTTTTTATCAGCAAAGAACTGTAT  
GATAGTATCCTTTTATCAGGGATTTTAAACAGTTACTGATGTTGGTACCAAAGC  
GCCTTTATCTGCTTCGATGATTCAGCGATATAATGAACATCAGATGGATTTAG  
CTCAGCTTAAACAATTCATTCGTCAGAAATTATCAGATAAATATAACGAAGTT  
TTTTCTGATGTTTCAAAGACGGCTATGCGGGTTATATTGATGGGAAAACAA  
ATCAAGAAGCTTTTTATAAATACCTTAAAGGTCTATTAAATAAGATTGAGGG  
AAGTGGCTATTTCTTGATAAAATTGAGCGTGAAGATTTTCTAAGAAAGCAA  
CGTACCTTTGACAATGGCTCTATTCCACATCAGATTCATCTTCAAGAAATGCG  
TGCTATCATTCGTAGACAGGCTGAATTTTATCCGTTTTTAGCAGACAATCAAG  
ATAGGATTGAGAAATTATTGACTTTCCGTATTCCCTACTATGTTGGTCCATTA  
GCGCGCGGAAAAAGTGATTTTGCTTGGTTAAGTCGGAAATCGGCTGATAAAA  
TTACACCATGGAATTTTGATGAAATCGTTGATAAAGAATCCTCTGCAGAAGCT  
TTTATCAATCGTATGACAAATTATGATTTGTACTTGCCAAATCAAAAAGTTCT  
TCCTAAACATAGTTTATTATACGAAAAATTTACTGTTTACAATGAATTAACAA  
AGGTAAATATAAAACAGAGCAAGGAAAAACAGCATTTTTTGATGCCAATAT  
GAAGCAAGAAATCTTTGATGGCGTATTTAAGGTTTATCGAAAAGTAACTAAA  
GATAAATTAATGGATTTCTTGAAAAAGAATTTGATGAATTTCTGATTGTTGA  
TTTAACAGGTCTGGATAAAGAAAAATAAAGTATTTAACGCTTCTTATGGAAGTT  
ATCATGATTTGTGTAAAATTTTAGATAAAGATTTTCTCGATAATTCAAAGAAT  
GAAAAGATTTTAGAAGATATTGTGTTGACCTTAACGTTATTTGAAGATAGAG  
AAATGATTAGAAAACGTCTAGAAAATTACAGTGATTTATTGACCAAAGAACA  
AGTGAAAAAGCTGGAAAGACGTCATTATACTGGTTGGGGAAGATTATCAGCT  
GAGTTAATTCATGGTATTCGCAATAAAGAAAGCAGAAAAACAATTCTTGATT

10

20

30

40

ATCTCATTGATGATGGCAATAGCAATCGGAACTTTATGCAACTGATTAACGAT  
GATGCTCTTTCTTTCAAAGAAGAGATTGCTAAGGCACAAGTTATTGGAGAAA  
CAGACAATCTAAATCAAGTTGTTAGTGATATTGCTGGCAGCCCTGCTATTAAA  
AAAGGAATTTTACAAAGCTTGAAGATTGTTGATGAGCTTGTCAAATTATGG  
GACATCAACCTGAAAATATCGTCGTGGAGATGGCGCGTGAAAACCAGTTTAC  
CAATCAGGGACGACGAAATTCACAGCAACGTTTGAAAGGTTTGACAGATTCT  
ATTAAAGAATTTGGAAGTCAAATTCTTAAAGAACATCCGGTTGAGAATTCAC  
AGTTACAAAATGATAGATTGTTTCTATATTATTTACAAAACGGCAGAGATATG  
TATACTGGAGAAGAATTGGATATTGATTATCTAAGCCAGTATGATATAGACC  
ATATTATCCCGCAAGCTTTTATAAAGGATAATTCTATTGATAATAGAGTATTG  
ACTAGCTCAAAGGAAAATCGTGGAAAATCGGATGATGTACCAAGTAAAGAT  
GTTGTTTCGTAAAATGAAATCCTATTGGAGTAAGCTACTTTCGGCAAAGCTTAT  
TACACAACGTAAATTTGATAATTTGACAAAAGCTGAACGAGGTGGATTGACC  
GACGATGATAAAGCTGGATTCATCAAGCGTCAATTAGTAGAAACACGACAAA  
TTACCAAACATGTAGCACGTATTCTGGACGAACGATTTAATACAGAAACAGA  
TGAAAACAACAAGAAAATTCGTCAAGTAAAAATTGTGACCTTGAAATCAAAT  
CTTGTTTCCAATTTCCGTAAAGAGTTTGAAGCTCTACAAAGTGCGTGAAATTAA  
TGACTATCATCATGCACATGATGCCTATCTCAATGCTGTAATTGGAAAGGCTT  
TACTAGGTGTTTACCCACAATTGGAACCTGAATTTGTTTATGGTGATTATCCT  
CATTTTCATGGACATAAAGAAAATAAAGCAACTGCTAAGAAATTTTCTATTC  
AAATATTATGAACTTCTTTAAAAAAGATGATGTCCGTACTGATAAAAATGGT  
GAAATTATCTGGAAAAAAGATGAGCATATTTCTAATATTTAAAAAAGTGCTTT  
CTTATCCACAAGTTAATATTGTTAAGAAAGTAGAGGAGCAAACGGGAGGATT  
TTCTAAAGAATCTATCTTGCCGAAAGGTAATTCTGACAAGCTTATTCCTCGAA  
AAACGAAGAAATTTTATTGGGATACCAAGAAATATGGAGGATTTGATAGCCC  
GATTGTTGCTTATTCTATTTTAGTTATTGCTGATATTGAAAAAGGTAAATCTA  
AAAAATTGAAAACAGTCAAAGCCTTAGTTGGTGTCACTATTATGGAAAAGAT  
GACTTTTGAAAGGGATCCAGTTGCTTTTCTTGAGCGAAAAGGCTATCGAAAT  
GTTCAAGAAGAAAATATTATAAAGTTACCAAAATATAGTTTATTTAACTAG  
AAAACGGACGAAAAAGGCTATTGGCAAGTGCTAGGGAACCTTCAAAGGGAA  
ATGAAATCGTTTTGCCAAATCATTTAGGAACCTTGCTTTATCACGCTAAAAAT

10

20

30

40

ATTCATAAAGTTGATGAACCAAAGCATTTGGACTATGTTGATAAACATAAAG  
ATGAATTTAAGGAGTTGCTAGATGTTGTGTCAAACCTTTTCTAAAAAATATACT  
TTAGCAGAAGGAAATTTAGAAAAAATCAAAGAATTATATGCACAAAATAATG  
GTGAAGATCTTAAAGAATTAGCAAGTTCATTTATCAACTTATTAACATTTACT  
GCTATAGGAGCACCGGCTACTTTTAAATTCTTTGATAAAAAATATTGATCGAAA  
ACGATATACTTCAACTACTGAAATTCTCAACGCTACCCTCATCCACCAATCCA  
TCACCGGTCTTTATGAAACGCGGATTGATCTCAATAAGTTAGGAGGAGACTA  
A

10

【 0 1 5 7 】

配列番号 6

配列：

MKKPYSIGLDIGTNSVGWAVVTDDYKVPAAKMKVLGNTDKSHIEKNLLGALLF  
DSGNTAEDRRLKRTARRRYTRRRNRILYLQEIFSEEMGKVDDSFHRLSDSFLVT  
EDKRGERHPIFGNLEEEVKYHENFPTIYHLRQYLADNPEKVDLRLVYLALAHIIKF  
RGHFLIEGKFDTRNNDVQRLFQEFLAVYDNTFENSSLQEQNVQVEEILTDKISKS  
AKKDRVCLKLPNEKSNGRFAEFLKLIVGNQADFKKHFELEEKAPLQFSKDTYEEE  
LEVLLAQIGDNYAELFLSAKKLYDSILLSGILTVTDVGTKAPLSASMIQRYNEHQ  
MDLAQLKQFIRQKLSDKYNEVFSVSKDGYAGYIDGKTNQEAIFYKYLKGLLNKI  
EGSGYFLDKIEREDFLRKQRTFDNGSIPHQIHLQEMRAIIRRQAEFYPLADNQDR  
IEKLLTRIPYYVGPLARGKSDFAWLSRKSADKITPWNFDEIVDKESSAEAFINRM  
TNYDLYLPNQKVLPHKSLLEYKFTVYNELTKVKYKTEQGKTAFFDANMKQEIFD  
GVFKVYRKVTKDKLMDFLEKEFDEFRIVDLTGLDKENKVFNASYGTYHDLCKIL  
DKDFLDNSKNEKILEDIVLTLTLFEDREMIRKRLNYSDLLTKEQVKKLERRHYT  
GWGRLSAELIHGIRNKESRKTILDYLIDDGNSNRNFMQLINDDALSFKEEIAKAQ  
VIGETDNLNQVVSIDIAGSPAIIKKGILQSLKIVDELVKIMGHQPENIVVEMARENQF  
TNQGRNSQQRLKGLTDSIKEFGSQILKEHPVENSQQLQNDRLFLYYLQNGRDMY  
TGEELDIDYLSQYDIDHIIPQAFIKDNSIDNRVLTSSKENRGKSDDVPSKDVVRKM  
KSYWSKLLSAKLITQRKFDNLTKAERGGLTDDDKAGFIKRQLVETRQITKHVARI  
LDERFNTETDENNKIRQVKIVTLKSNLVSNFRKEFELYKVREINDYHHAHDAYL  
NAVIGKALLGVYPQLEPEFVYGDYPHFHGHKENKATAKKFFYSNIMNFFKKDDV  
RTDKNGEIIWKKDEHISNIKKVLSYPQVNIVKKVEEQTTGGFSKESILPKGNSDKLIP  
RKTKKFYWDTKKYGGFDSPIVAYSILVIADIEKGKSKKLKTVKALVGVTIMEKM  
TFERDPVAFLEKGYRNVQEENIIKLPKYSLEFKLENGRKRLLASARELQKGNEIVL  
PNHLGTLLYHAKNIHKVDEPKHLDYVDKHKDEFKELLDVVSNFSSKKYTLAEGN  
LEKIKELYAQNNGEDLKELASSFINLLTFTAIGAPATFKFFDKNIDRKRYTSTTEIL  
NATLIHQSIITGLYETRIDLNLKGGD

10

20

30

40

【 0 1 5 8 】

配列番号 7

名称：N . メニンギティデス ( N . m e n i n g i t i d e s ) C a s 9

配列：



ATGGCTGCCTTCAAACCTAATTCAATCAACTACATCCTCGGCCTCGATATCGG  
CATCGCATCCGTCGGCTGGGCGATGGTAGAAATTGACGAAGAAGAAAACCCC  
ATCCGCCTGATTGATTTGGGCGTGCGCGTATTTGAGCGTGCCGAAGTACCGA  
AAACAGGCGACTCCCTTGCCATGGCAAGGCGTTTGGCGCGCAGTGTTGCGCG  
CCTGACCCGCCGTCGCGCCACCGCCTGCTTCGGACCCGCCGCCTATTGAAAC  
GCGAAGGCGTATTACAAGCCGCCAATTTTGACGAAAACGGCTTGATTAAATC  
CTTACCGAATACACCATGGCAACTTCGCGCAGCCGCATTAGACCGCAAACCTG  
ACGCCTTTAGAGTGGTCGGCAGTCTTGTTGCATTTAATCAAACATCGCGGCTA  
TTTATCGCAACGGAAAAACGAGGGCGAAACTGCCGATAAGGAGCTTGGCGCT  
TTGCTTAAAGGCGTAGCCGGCAATGCCCATGCCTTACAGACAGGCGATTTC  
GCACACCGGCCGAATTGGCTTTAAATAAATTTGAGAAAGAAAGCGGCCATAT  
CCGCAATCAGCGCAGCGATTATTCGCATACGTTTCAGCCGCAAAGATTTACAG  
GCGGAGCTGATTTTGCTGTTTGAAAAACAAAAAGAATTTGGCAATCCGCATG  
TTTCAGGCGGCCTTAAAGAAGGTATTGAAACCCTACTGATGACGCAACGCCC  
TGCCCTGTCCGGCGATGCCGTTCAAAAAATGTTGGGGCATTGCACCTTCGAAC  
CGGCAGAGCCGAAAGCCGCTAAAAACACCTACACAGCCGAACGTTTCATCTG  
GCTGACCAAGCTGAACAACCTGCGTATTTTAGAGCAAGGCAGCGAGCGGCCA  
TTGACCGATACCGAACGCGCCACGCTTATGGACGAGCCATACAGAAAATCCA  
AACTGACTTACGCACAAGCCCGTAAGCTGCTGGGTTTAGAAGATACCGCCTT  
TTTCAAAGGCTTGCGCTATGGTAAAGACAATGCCGAAGCCTCAACATTGATG  
GAAATGAAGGCCTACCATGCCATCAGCCGTGCACTGGAAAAAGAAGGATTG  
AAAGACAAAAAATCCCCATTAAACCTTTCTCCCGAATTACAAGACGAAATCG  
GCACGGCATTCTCCCTGTTCAAACCGATGAAGACATTACAGGCCGTCTGAA  
AGACCGTATACAGCCCGAAATCTTAGAAGCGCTGTTGAAACACATCAGCTTC

10

20

30

GATAAGTTCGTCCAAATTTCTTGAAAGCATTGCGCCGAATTGTGCCTCTAAT  
GGAACAAGGCAAACGTTACGATGAAGCCTGCGCCGAAATCTACGGAGACCA  
TTACGGCAAGAAGAATACGGAAGAAAAGATTTATCTGCCGCCGATTCCCGCC  
GACGAAATCCGCAACCCCGTCGTCTTGCGCGCCTTATCTCAAGCACGTAAGG  
TCATTAACGGCGTGGTACGCCGTTACGGCTCCCCAGCTCGTATCCATATTGAA  
ACTGCAAGGGAAGTAGGTAAATCGTTTAAAGACCGCAAAGAAAATTGAGAAA  
CGCCAAGAAGAAAACCGCAAAGACCGGGAAAAAGCCGCCGCCAAATTCCGA  
GAGTATTTCCCAATTTTGTCTGGAGAACCCAAATCCAAAGATATTCTGAAACT  
GCGCCTGTACGAGCAACAACACGGCAAATGCCTGTATTCGGGCAAAGAAATC  
AACTTAGGCCGTCTGAACGAAAAAGGCTATGTCGAAATCGACCATGCCCTGC  
CGTTCTCGCGCACATGGGACGACAGTTTCAACAATAAAGTACTGGTATTGGG  
CAGCGAAAACCAAAACAAAGGCAATCAAACCCCTTACGAATACTTCAACGG  
CAAAGACAACAGCCGCGAATGGCAGGAATTTAAAGCGCGTGTCTGAAACCAG  
CCGTTTCCCGCGCAGTAAAAACAACGGATTCTGCTGCAAAAATTCGATGAA  
GACGGCTTTAAAGAACGCAATCTGAACGACACGCGCTACGTCAACCGTTTCC  
TGTGTCAATTTGTTGCCGACCGTATGCGGCTGACAGGTAAAGGCAAGAAACG  
TGTCTTTGCATCCAACGGACAAATTACCAATCTGTTGCGCGGCTTTTGGGGAT  
TGCGCAAAGTGCGTGCGGAAAACGACCGCCATCACGCCTTGGACGCCGTCGT  
CGTTGCCTGCTCGACCGTTGCCATGCAGCAGAAAATTACCCGTTTTGTACGCT  
ATAAAGAGATGAACGCGTTTGACGGTAAAACCATAGACAAAGAAACAGGAG  
AAGTGCTGCATCAAAAAACACACTTCCCACAACCTTGGGAATTTTTCGCACA  
AGAAGTCATGATTCGCGTCTTCGGCAAACCGGACGGCAAACCCGAATTCGAA  
GAAGCCGATACCCTAGAAAAACTGCGCACGTTGCTTGCCGAAAAATTATCAT  
CTCGCCCCGAAGCCGTACACGAATACGTTACGCCACTGTTTGTTCACGCGCG  
CCCAATCGGAAGATGAGCGGGCAAGGGCATATGGAGACCGTCAAATCCGCC  
AAACGACTGGACGAAGGCGTCAGCGTGTTGCGCGTACCGCTGACACAGTTAA  
AACTGAAAGACTTGGA AAAAATGGTCAATCGGGAGCGCGAACCTAAGCTAT  
ACGAAGCACTGAAAGCACGGCTGGAAGCACATAAAGACGATCCTGCCAAAG  
CCTTTGCCGAGCCGTTTTACAAATACGATAAAGCAGGCAACCGCACCCAACA  
GGTAAAAGCCGTACGCGTAGAGCAAGTACAGAAAACCGGCGTATGGGTGCG  
CAACCATAACGGTATTGCCGACAACGCAACCATGGTGCGCGTAGATGTGTTT

10

20

30

40

GAGAAAGGCGACAAGTATTATCTGGTACCGATTTACAGTTGGCAGGTAGCGA  
AAGGGATTTTGCCGGATAGGGCTGTTGTACAAGGAAAAGATGAAGAAGATTG  
GCAACTTATTGATGATAGTTTCAACTTTAAATTCTCATTACACCCTAATGATTT  
AGTCGAGGTTATAACAAAAAAGCTAGAATGTTTGGTTACTTTGCCAGCTGC  
CATCGAGGCACAGGTAATATCAATATACGCATTCATGATCTTGATCATAAAA  
TTGGCAAAAATGGAATACTGGAAGGTATCGGCGTCAAACCGCCCTTTCATT  
CCAAAAATACCAAATTGACGAACTGGGCAAAGAAATCAGACCATGCCGTCTG  
AAAAAACGCCCCGCCTGTCCGTAA

10

【 0 1 5 9 】

配列番号 8

配列：

MAAFKPNSINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGD  
SLAMARRLARSVRRLTRRRAHRLLRTRRLKREGVLQAANFDENGLIKSLPNT  
WQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVA  
GNAHALQTGDFRTPAELALNKFEEKESGHIRNQRSYDYSHTFSRKDLQAEILLFEK  
QKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAPKAANKTY  
TAERFIWLTKLNNLRILEQGSERPLDTERATLMDEPYRKSKLTYAQARKLLGLE  
DTAFFKGLRYGKDNAAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSPELQDE  
IGTAFSLFKTDEDITGRLKDRIQPEILEALLKHISFDKQVQISLKALRRIVPLMEQK  
RYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRLALSQARKVINGVVR  
YGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPK  
SKDILKLRLYEQQHKGCLYSGKEINLGRLEKGYVEIDHALPFSRTWDDSFNNK  
VLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKF  
DEDGFKERNLNDTRYVNRFLCQFVADRMRLTGKGKKRVFASNGQITNLLRGFW  
GLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETG  
EVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTLEKLRTLLAEKLSSRP  
EAVHEYVTPLFVSRAPNRKMSGQGHMETVKSARKLDEGVSVLRVPLTQLKLKD  
LEKMVNREREPKLYEALKARLEAHKDDPAKAFAEPFYKYDKAGNRTQQVKAV  
RVEQVQKTGVWVRNHNGIADNATMVRVDVFEGDKYYLVPIYSWQVAKGILP  
DRAVVQGKDEEDWQLIDDSFNFKFSLHPNDLVEVITKKARMFGYFASCHRG TG  
NINIRIHDLDHKIGKNGILEGIGVKTALSFQKYQIDELGKEIRPCRLKKRPPVR

20

30

40

【 0 1 6 0 】

配列番号 9

配列：

ATGAACAATAACAATTACTCTATCGGACTCGATATCGGAACAAACAGCGTCG  
GATGGGCCGTCATTACGGATGACTATAAGGTGCCATCGAAAAAGATGAAAGT  
TCTAGGCAATACAGATAAAACACTTTATCAAGAAAAATCTAATTGGAGCTTTA  
TTATTTGATGAAGGAGCTACTGCTGAAGATAGACGTTTCAAACGAACAGCAC  
GCCGTCGCTATACTCGTCGAAAAAATCGTCTTCGCTATCTTCAAGAAATCTTT  
TCTGAGGAAATGAGCAAAGTGGATAGTAGTTTCTTTCATCGATTAGATGACTC  
ATTCTTAGTTCCTGAGGATAAAAGAGGAAGTAAATATCCTATTTTTTGCTACCT  
TGGCAGAAGAAAAAGAATATCACAAGAAATTTCCAACCTATCTATCATTTGAG  
AAAACACCTTGCGGACTCAAAGAAAAAACTGACTTGCGCTTGATCTATCTA  
GCATTAGCGCATATGATTAAATACCGCGGACATTTTTTGTATGAAGAATCTTT  
CGATATTAAAAACAATGATATCCAAAAAATCTTTAGCGAGTTTATAAGCATT  
ACGACAACACCTTTGAAGGAAGTTCAGTTAGTGGACAAAATGCACAAGTAGA  
AGCAATTTTTACTGATAAAATTAGTAAATCTGCTAAGAGAGAACGCATTCTA  
AACTCTTTGCTTATGAAAAATCCACTGATCTATTTTCAGAATTTCTCAAGCT  
GATTGTAGGAAATCAAGCTGATTTTAAGAAACACTTTGACTTGGAAGAAAAA  
GCTCCACTACAATTCTCTAAAGATACCTATGATGAGGATTTGGAAAACCTACT  
CGGACAAATTGGAGATGACTTTGCAGACCTTTTCCTAGTTGCTAAAAAATCT  
ATGATGCCATTCTTTTATCAGGAATCTTAACTGTTACAGATTCTTCAACTAAG  
GCCCCACTATCAGCATCTATGATTGAGCGCTATGAAAACCACCAAAAAGACT  
TAGCGGCTTTAAAACAATTCATCCAAAACAATCTTCAAGAAAAATATGATGA  
AGTTTTCTCTGACCAATCTAAAGATGGGTATGCTAGGTATATCAATGGCAAA  
ACCACTCAAGAAGCATTTTACAAGTACATCAAAAATCTTCTCTCTAAATTCGA  
AGGATCAGATTATTTCTTGATAAAATTGAACGTGAAGATTTCTTGAGAAAA  
CAACGCACCTTTGATAATGGTTCTATCCCTCATCAAATTCATCTTCAAGAAAT  
GAATGCCATTATCCGTCGGCAAGGAGAACATTATCCATTTCTGAAGGAATAT  
AAAGAAAAGATAGAGACAATCTTGACTTTCCGTATTCCTTATTATGTTGGCCC  
ATTGGCTCGTGGAATCGTAATTTTGCTTGGCTTACTCGAACTCTGACCAAG  
CAATCCGACCTTGGAATTTTGAAGAAATTGTTGATCAAGCAAGCTCTGCGGA  
AGAATTCATCAATAAGATGACTAACTATGACTTGTATCTGCCAGAGGAAAAA  
GTTTTGCCCAAGCATAGTCTCTTGTATGAAACATTTGCTGTCTACAATGAATT

10

20

30

40

AACAAAAGTAAAATTTATTTTCAGAGGGATTGAGAGACTATCAATTCCTTGAT  
AGTGGGCAAAGAAGCAAATTGTCAATCAATTATTCAAAGAGAAAAGAAAA  
GTAAGTGAAGAACATCATTTCAGTATCTACACAATGTTGATGGCTACGATG  
GAATCGAACTAAAAGGAATTGAAAAACAATTTAACGCTAGTCTTTCTACTTA  
TCATGATTTACTCAAAATAATCAAGGATAAAGAGTTTATGGATGATCCTAAA  
AATGAAGAGATTCTTGAAAATATCGTCCACACACTAACTATCTTTGAAGATC  
GTGAGATGATCAAGCAACGCCTTGCTCAATATGCCTCTATCTTTGATAAAAAA  
GTGATCAAGGCACTGACTCGTCGACATTATACTGGTTGGGGAAAACCTCTCTG  
CTAAGCTAATCAACGGTATCTGTGATAAAAAAACTGGTAAAACAATTCTTGA  
CTACTTGATTGATGACGGCTACAGCAATCGTAACTTTATGCAGTTAATCAATG  
ATGACGGGCTTTCTTCAAAGATATTATTCAAAAAGCACAAAGTGGTTGGTAA  
GACAAACGATGTGAAGCAAGTTGTCCAAGAACTCCAGGTAGTCCTGCTATT  
AAAAAGGGAATTTTACAAAGTATCAAGCTTGTGATGAGCTTGTCAAAGTTA  
TGGGCCATGCTCCCGAGTCCATTGTGATTGAAATTGCACGAGAAAATCAGAC  
AACTGCCAGAGGGAAAAAGAATTCTCAACAAAGATATAAGCGCATTGAAGA  
TGCACTAAAAAATTTAGCACCTGGGCTTGATTCAAATATATTAAGAACAT  
CCAACAGATAATATTCAACTTCAAATGACCGTCTCTTCCTTTACTATCTCCA  
AAATGGGAAGGATATGTACACTGGAGAAGCTCTTGATATCAACCAACTGAGC  
AGCTATGACATTGACCACATCGTCCCACAGGCCTTTATCAAGGATGATTCTCT  
TGATAACCGTGTCTTGACTAGTTCAAAGGATAATCGTGGGAAATCCGATAAT  
GTTCCAAGTTTAGAAGTCGTTCAAAAAGAAAAGCTTTTTGGCAACAATTAC  
TAGATTCCAAATTGATTTTCAGAACATAAATTTAATAATTTAACCAAGGCTGAA  
CGTGGTGGGCTAGATGAGCGAGATAAAGTTGGCTTTATCAGACGCCAACTAG  
TTGAAACACGGCAAATCACAAAACATGTTGCTCAGATTTTGGATGCCCCGTTTT  
AATACAGAAGTGAATGAGAAAGATAAGAAGAACCGTACCGTCAAATTATC  
ACTTTGAAATCCAATCTAGTTTCCAACCTCCGTAAAGAATTTAAGTTATATAA  
GGTACGCGAAATCAATGACTACCACCATGCACATGATGCCTATTTAAATGCA  
GTGGTGGCTAAGGCTATCCTTAAGAAATATCCTAACTAGAGCCTGAATTCG  
TCTATGGTGACTATCAAAGTACGATATTAAGAGATATATTTCCAGATCCAA  
AGATCCTAAAGAAGTTGAAAAAGCAACTGAAAAGTATTTCTTCTACTCAAAC  
TTGTTGAACTTCTTTAAAGAAGAGGTGCATTACGCAGACGGAACCATCGTAA

10

20

30

40

AACGAGAGAATATCGAATACTCTAAGGACACTGGAGAAATCGCTTGGAATAA  
AGAAAAAGATTTTCGCTACAATTAAAAAGTTCTTTCACTTCCGCAGGTGAAT  
ATTGTGAAGAAAACAGAGATTCAAACACATGGTCTAGATAGAGGTAAACCTA  
GAGGATTGTTCAATTCCAATCCATCTCCTAAACCTTCAGAAGATCGTAAAGA  
AAACCTTGTCCCAATTAAACAAGGGCTTGACCCACGAAAATACGGTGGTTAC  
GCTGGTATTTCTAACTCATACGCGGTCTTAGTTAAAGCTATTATTGAAAAAGG  
AGCGAAAAACAACAAAAGACCGTTCTTGAATTTCAAGGTATCTCTATTTTA  
GATAAAATAAATTTTGAAAAGAACAAAGAAAACCTATCTTCTTGAAAAAGGAT  
ACATAAAAATTCTATCAACTATTACTTTACCTAAATATAGTTTGTTTGAGTTTC  
CTGATGGTACAAGAAGAAGACTAGCAAGTATTCTATCGACAAACAATAAACG  
AGGAGAAATTCATAAAGGTAATGAATTGGTCATCCCTGAAAAGTATACGACT  
CTTTTGTATCATGCTAAGAATATTAATAAAACACTTGAACCAGAACACTTAGA  
GTATGTTGAGAAACATCGAAATGATTTTGCTAAACTTTTAGAATATGTACTTA  
ACTTTAACGATAAGTATGTAGGCGCATTAAAAAATGGAGAAAGAATCAGACA  
AGCATTTATTGATTGGGAAACAGTTGATATTGAAAAGTTATGTTTCAGTTTCA  
TTGGTCCAAGAAATAGTAAAAATGCTGGTTTATTCGAGTTAACTTCACAAGG  
AAGTGCTTCTGACTTCGAGTTCTTGGGAGTAAAAATTCCACGATACAGAGAC  
TATACACCTTCGTCACTCCTCAACGCCACCCTCATCCACCAATCCATCACTGG  
TCTTTACGAGACTCGGATTGACTTAAGCAAACCTGGGAGAAGACTGA

10

20

30

【 0 1 6 1 】

配列番号 1 0

名称 : g i | 7 7 7 8 8 8 0 6 2 | g b | K J Q 6 9 4 8 3 . 1 | C R I S P R 関連エン  
ドヌクレアーゼ C a s 9 [ ストレプトコッカス・ミティス ( S t r e p t o c o c c u s  
m i t i s ) ]

配列 :

MNNNNYSIGLDIGTNSVGWAVITDDYKVPSKKMKVLGNTDKHFIKKNLIGALLF  
DEGATAEDRRFKRTARRRYTRRKNRLRYLQEIFSEEMSKVDSSFFHRLDDSFLVP  
EDKRGSKYPIFATLAEEKEYHKKFPTIYHLRKHLADSKEKTDLRLIYLALAHMIK  
YRGHFLYEESFDIKNNDIQKIFSEFISIYDNTFEGSSLSGQNAQVEAIFTDKISKS  
RERILKLFAYEKSTDLFSEFLKLIVGNQADFKKHFDLEEKAPLQFSKDTYDEDLEN  
LLGQIGDDFADLFLVAKKLYDAILLSGILTVTDSSTKAPLSASMIERYENHQDLA  
ALKQFIQNNLQEKYDEVFSDQSKDGYARYINGKTTQEAFYKYIKNLLSKFEGSD  
YFLDKIEREDFLRKQRTFDNGSIPHQIHLQEMNAIIRRQGEHYPFLKEYKEKIE  
TIL  
TFRIPYYVGPLARGNRNFAWLTRNSDQAIRPWNFEEIVDQASSAEFINKMTNYD  
LYLPEEKVLPKHSLLYETFAVYNELTKVKFISEGLRDYQFLDSGQKKQIVNQLFK  
EKRKVTEKDIIQYLHNVDGYDGIELKGIEKQFNASLSTYHDLLKIIKDKEFMDDP  
KNEEILENIVHTLTIFEDREMIKQRLAQYASIFDKKVIKALTRRHYTGWGKLSAKL  
INGICDKKTGKTILDYLIDDGYSNRNFMQLINDDGLSFKDIIQKAQVVGKTNDVK  
QVVQELPGSPAIAKKGILQSIKLVELVKVMGHAPESIVIEIARENQTTARGKKNSQ  
QRYKRIEDALKNLAPGLDSNILKEHPTDNIQLQNDRLFLYYLQNGKDMYTGEAL  
DINQLSSYDIDHIVPQAFIKDDSLDNRVLTSSKDNRGKSDNVPVSLEVVQKRKA  
FW  
QQLLDSKLISEHKFNNTKAERGGGLDERDKVGFIRRLQVETRQITKHVAQILDAR  
FNTEVNEKDKKNRTVKIITLKSNLVSNFRKEFKLYKVREINDYHHAHDAYLNAV  
VAKAILKKYPKLEPEFVYGDYQKYDIKRYISRSKDPKEVEKATEKYFFYSNLLNF  
FKEEVHYADGTIVKRENIEYSKDTGEIAWNKEKDFATIKKVLSPQVNIVKKTEIQ  
THGLDRGKPRGLFNSNPSPKPSEDRKENLVPIKQGLDPRKYGGYAGISNSYAVLV  
KAIIEKGAKKQKQKTVLEFQGISILDKINFEKNKENYLLKGYIKILSTITLPKYS  
LFE  
FPDGTRRRRLASILSTNNKRGEIHKGNELVIPEKYTTLLYHAKNINKTLEPEH  
LEYV  
EKHRNDFAKLLEYVLNFNDKYVGALKNGERIRQAFIDWETVDIEKLCFSFIGPRN  
SKNAGL FELTSQGSASDFEFLGVKIPRYRDYTPSSLLNATLIHQ  
SITGLYETRIDL  
SKLGED

10

20

30

40

【 0 1 6 2 】

配列番号 1 1

配列：

ATGACAAAACCTTATTCTATTGGACTTGATATTGGGACTAACTCTGTTGGTTG  
GGCTGTTGTGACAGATGGCTACAAAGTTCCTGCTAAGAAGATGAAGGTTCTG  
GGAAATACAGATAAAAGCCATATCAAGAAAAATTTACTTGGAGCTTTATTGT  
TTGATAGCGGTAATACTGCAAAAGACAGACGTTTGAAGCGGACAGCTAGGCG  
TCGATATACACGTCGTAGAAACCGTATTTTATATTTGCAGGAAATTTTGTCTG  
AAGAAATGGCTAAAGCAGACGAAAGTTTCTTCCAGCGCTTAAACGAATCGTT  
TTTAACAAATGATGACAAAGAATTTGATTCTCATCCAATCTTTGGGAATAAAG  
CTGAAGAGGAGGCTCATCACCATAAATTTCCAACAATTTTTCATTTGCGAAAG  
CATTTAGCAGACTCAACCGAGAAATCTGATTTGCGCTTAATTTATCTAGCTTT  
AGCGCATATGATTAAATTCCGGGGACATTTCTTAATTGAAGGTCAGCTAAAA  
GCTGAAAATACAAATGTTCAAACATTATTTGACGATTTTGTAGAAGTATATGA



TAAGACAGTTGAAGAAAGTCATTTATCAGAAATTAGTGTCTCCAGTATTCTGA  
CAGAAAAAATTAGTAAATCGCGTCGCTTAGAAAAATCTTATAAAATACTATCC  
CACTGAGAAGAAAAACACTCTCTTCGGAAATCTTATCGCCTTGTCTTTAGGAT  
TACAGCCAAACTTTAAAACAAATTTTAAATTATCCGAAGATGCTAAACTACA  
GTTTTCTAAGGATACTTATGAAGAAGATTTAGGAGAATTACTTGGA AAAATC  
GGAGATAATTATGCAGATTTATTTATATCAGCTAAAAATCTTTATGATGCTAT  
TTTGCTATCAGGAATTTTAAACAATAGATGACAACACGACAAAGGCTCCGTTG  
TCTGCTTCAATGATTAAACGTTATGAGGAACATCAGGAAGATTTAGCACAAC  
TTAAGAAATTTATCCGTCAGAATTTACCAGATCAATATAGTGAGGTTTTTTCT  
GATAAAACAAAGGATGGCTATGCTGGTTATATTGATGGAAAAACGAATCAGG  
AGGCCTTTTATAAATACATCAAAAATATGCTGTCAAAAACAGAAGGTGCAGA  
TTATTTTCTTGACAAAATTGATCGTGAAGACTTTTTGAGAAAACAGAGAACGT  
TTGATAATGGTTCCGTTCCGCATCAGATTCATCTGCAAGAGATGCATGCTATT  
TTACGACGTCAGGGTGAATACTATCCATTCTTGAAAGAAAATCAGGATAAAA  
TTGAAAAAATCTTAACGTTTAGAATTCCTTACTACGTTGGTCCTTTGGCGCGA  
AAAGGTAGCCGCTTTGCCTGGGCAGAATACAAGGCGGATAAAAAAGTTACGC  
CATGGAATTTTGATGATATTCTTGATAAAGAAAAATCAGCAGAAGAATTCAT  
CACACGCATGACTTTAAATGATTTGTATTTACCTGAAGAAAAAGTCTTACCAA  
AGCATAGTCTTGTTTATGAAACGTTTAATGTTTACAATGAGTTAACTAAAGTT  
AAGTATGTCAATGAGCAAGGGAAAGCCATTTTCTTTGATGCCAATATGAAGC  
AAGAGATTTTTGATCATGTTTTTAAAGAAAATCGGAAAGTTACTAAAGATAA  
ACTTTTAAATTATTTGAATAAAGAGTTTGAAGAATTTAGAATTGTAACTTAA  
CTGGACTGGATAAGGAAAATAAAGCCTTTAATTCCAGTCTTGGAACCTATCA  
TGATTTGCGTAAAATTTTAGATAAATCATTCTTAGATGATAAAGTAAATGAAA  
AGATAATTGAGGATATCATTCAAACACTAACTCTGTTTGAAGACAGAGAAAT  
GATTCGTCAGCGTCTTCAAAGTATAGTGATATTTTTACAACACAGCAATTGA  
AAAACTTGAACGCCGTCATTATACAGGTTGGGGAAGATTATCAGCGAAGTT  
AATCAATGGTATTCGAGATAAACAGAGTAATAAGACTATTCTGGGTATTG  
ATTGATGATGGTTATAGCAATCGTAACTTTATGCAGTTGATTAATGACGATTC  
TCTTCCTTTTAAAGAAGAAATTGCTAGGGCACAAGTCATTGGAGAAACAGAT  
GACTTAAATCAACTTGTTAGTGATATTGCTGGCAGTCCTGCTATTA AAAAGGG

10

20

30

40

AATTTTACAAAGTCTGAAAATTGTAGATGAGCTTGTTAAAGTCATGGGGCAT  
AATCCTGCTAACATTGTTATCGAAATGGCGCGTGAAAATCAGACTACAGCCA  
AAGGGCGTCGCAGTTCACAGCAACGTTATAAACGACTTGAGGAGGCAATAAA  
AAATCTTGACCATGATTTAAATCATAAGATTTTAAAAGAACACCCAACAGAT  
AATCAAGCTTTACAGAATGACCGTCTTTTCTTATATTATCTCCAAAATGGCCG  
AGATATGTATACTGAAGATCCACTTGATATTAATCGTTTAAAGTGATTATGATA  
TCGACCATATTATTCCACAATCTTTTATAAAAGATGACTCTATTGACAATAAG  
GTTCTGGTTTCATCAGCTAAAAACCGTGGGAAATCGGATAATGTACCGAGTG  
AAGATGTTGTCAATAGGATGAGACCGTTTTGGAATAAATTATTGAGCTGTGG  
ATTGATTTCTCAACGGAAATACAGCAATCTAACCACAAAAAGAATTAAAACCA  
GATGATAAGGCTGGTTTCATCAAACGTCAATTGGTTGAGACAAGACAAATTA  
CAAAGCATGTTGCACAAATTTTAGACGCTCGTTTTAATACAAAACGTGATGA  
AAATAAAAAAGTAATTCGTGATGTCAAATTATCACTTTAAAATCTAATTTAG  
TTTCACAATTTTCGTAAAGACTTTAAATTTTACAAAGTACGTGAGATTAATGAT  
TACCATCATGCGCATGACGCTTATCTTAATGCAGTTATAGGAAAAGCTTTATT  
AGATGTTTATCCGCAGTTAGAGCCCGAATTTGTTTATGGTGAGTACCCTCATT  
TTCATGGATATAAAGAAAATAAAGCAACTGCTAAGAAATTTTTCTATTCAA  
TATTATGAATTTTTTTAAGAAAGATGATATCCGTACCGATGAAAATGGTGAG  
ATTGTTTGGAaaaaagATGAGCATATTTCTAATATTAaaAGGGTGCTTTCCTA  
TCCCCAAGTTAATATTGTTAAGAAAGTAGAAATACAGACTGTTGGACAAAAT  
GGGGGACTTTTTGACGATAATCCTAAATCACCATTAGAGGTTACACCTAGTA  
AACTTGTTCCACTAAAAAAAGAATTAAACCCTAAAAAATATGGAGGATATCA  
AAAACCGACGACAGCTTATCCTGTTTTACTGATAACAGATACTAAACAGCTA  
ATTCCAATCTCAGTAATGAATAAGAAGCAATTTGAACAAAATCCGGTTAAAT  
TTTTAAGAGATAGAGGCTATCAACAGGTAGGAAAGAATGACTTTATTAAATT  
ACCCAAATATACCCTAGTTGATATCGGTGATGGGATTAAACGCCTATGGGCT  
AGTTCGAAAGAAATACATAAAGGAAATCAATTAGTTGTATCTAAAAAATCTC  
AAATTTTGCTTTATCATGCACATCACTTAGATAGTGATTTGAGTAATGATTAT  
CTTCAAAATCATAATCAACAATTCGATGTTTTATTTAATGAAATTATTTCTTTT  
TCTAAAAAATGTAAATTGGGAAAAGAACATATTCAGAAAATTGAAAATGTTT  
ACTCCAATAAGAAGAATAGTGCATCAATAGAAGAATTAGCAGAGAGTTTTAT

10

20

30

40

TAAATTATTAGGATTTACACAATTAGGTGCAACTTCCCCATTTAATTTTTTAG  
GGGTAAAACTAAATCAAAAACAATATAAAGGTAAAAAAGATTATATTTTACC  
GTGTACAGAGGGGACCCTTATCCGCCAATCTATCACTGGTCTTTACGAAACAC  
GAGTTGATCTTAGTAAAATAGGAGAAGACTAA

【 0 1 6 3 】

配列番号 1 2

名称 : g i | 3 5 7 5 8 4 8 6 0 | g b | E H J 5 2 0 6 3 . 1 | C R I S P R 関連タンパク質 C a s 9 / C s n 1、サブタイプ I I / N M E M I [ ストレプトコッカス・マカカ ( S t r e p t o c o c c u s m a c a c a e ) N C T C 1 1 5 5 8 ] 10

配列 :

MTKPYSIGLDIGTNSVGWAVVTDGYKVPAKKMKVLGNTDKSHIKKNLLGALLF  
DSGNTAKDRRLKRTARRRYTRRRNRILYLQEIFAEMAKADESFFQRLNESFLT  
DDKEFDSHPHIFGNKAEEEEAHHHKFPTIFHLRKHLADSTEKSDLRLIYLALAHMIKF  
RGHFLIEGQLKAENTNVQTLFDDFVEVYDKTVEESHLSEISVSSILTEKISKSRRL  
NLIKYYPTTEKKNTLFGNLIASLGLQPNFKTNFKLSEDAKLQFSKDTYEEDLGELL  
GKIGDNYADLFISAKNLYDAILLSGILTIDDNTTKAPLSASMIKRYEEHQEDLAQL  
KKFIRQNLDPQYSEVFSKTKDGYAGYIDGKTNQEAIFYKYIKNMLSKEGADYF  
LDKIDREDFLRKQRTFDNGSVPHQIHLQEMHAILRRQGEYYPFLKENQDKIEKILT  
FRIPYYVGPLARKGSRFAWAEYKADKKVTPWNFDDILDKEKSAEEFITRMTLND  
LYLPEEKVLPKHSLVYETFNVYNELTKVKYVNEQGKAIFFDANMKQEIFDHVFK  
ENRKVTKDKLLNYLNKEFEFFRIVNLTGLDKENKAFNSSLGTYHDLRKILDKSFL  
DDKVNEKIIIEDIIQTLTLFEDREMIRQRLQKYSIDFTTQQLKKLERRHYTGWGRLS  
AKLINGIRDKQSNKTILGYLIDDGYSNRNFMLINDDSLPFKEEIARAQVIGETDD  
LNQLVSDIAGSPAIIKKGILQSLKIVDELVKVMGHNPANIVIEMARENQTTAKGRR  
SSQQRYSRLEEAIAKNLDHDLNHKILKEHPTDNQALQNDRLFLYYLQNGRDMYTE  
DPLDINRLSDYDIDHIIPQSFIDKDDSIDNKVLVSSAKNRGKSDNVPSDEVVNRMRP  
FWNKLLSCGLISQRKYSNLTKKELKPDDKAGFIKRQLVETRQITKHVAQILDARF  
NTRDENKKVIRDVKIITLKSNLVSQFRKDFKFYKVRINDYHHAHDAYLNAVIG  
KALLDVYPQLEPEFVYGEYPHFHGYKENKATAKKFFYSNIMNFFKKDDIRTDEN  
GEIVWKKDEHISNIKRVLSYPQVNIVKKVEIQTVGQNGGLFDDNPKSPLEVTPSK  
LVPLKKELNPKKYGGYQKPTTAYPVLLITDTKQLIPISVMNKKQFEQNPVKFLRD  
RGYQQVGKNDFIKLPKYTLVDIGDGIKRLWASSKEIHKGNQLVVSCKKSQILLYHA  
HHLDSDLSDNYLQNHQNFQDVLVFNIEISFSKKCKLGKEHIQKIENVYSNKKNSASI  
EELAESFIKLLGFTQLGATSPFNFLGVKLNQKQYKGKKDYILPCTEGTLIRQSITGL  
YETRVDLSKIGED

10

20

30

40

【 0 1 6 4 】

配列番号 1 3

配列：

ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGAT  
GGGCGGTGATCACTGATGATTATAAGGTTCCGTCTAAAAAGTTCAAGGTTCT  
GGGAAATACAGACCGCCACAGTATCAAAAAAATCTTATAGGGGCTCTTTTA  
TTTGACAGTGGAGAGACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTA  
GAAGGTATACACGTCGGAAGAATCGTATTTGTTATCTACAGGAGATTTTTTCA  
AATGAGATGGCGAAAGTAGATGATAGTTTCTTTCATCGACTTGAAGAGTCTTT  
TTTGGTGGGAAGAAGACAAGAAGCATGAACGTCATCCTATTTTTTGGAAATATA  
GTAGATGAAGTTGCTTATCATGAGAAATATCCAACCTATCTATCATCTGCGAAA  
AAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTTAATCTATTTGGCCT  
TAGCGCATATGATTAAGTTTCGTGGTCATTTTTTTGATTGAGGGAGATTTAAAT  
CCTGATAATAGTGATGTGGACAACTATTTATCCAGTTGGTACAAACCTACA  
ATCAATTATTTGAAGAAAACCCTATTAACGCAAGTGGAGTAGATGCTAAAGC  
GATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTC  
AGCTCCCCGGTGAGAAGAAAAATGGCTTATTTGGGAATCTCATTGCTTTGTCA  
TTGGGTTTGACCCCTAATTTTAAATCAAATTTTGATTGCGCAGAAGATGCTAA  
ATTACAGCTTTCAAAGATACTTACGATGATGATTTAGATAATTTATTGGCGC  
AAATTGGAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGAT  
GCTATTTTACTTTCAGATATCCTAAGAGTAAATACTGAAATAACTAAGGCTCC  
CCTATCAGCTTCAATGATTAAACGCTACGATGAACATCATCAAGACTTGACTC  
TTTTAAAAGCTTTAGTTCGACAACAACCTCCAGAAAAGTATAAAGAAATCTTT  
TTTGATCAATCAAAAAACGGATATGCAGGTTATATTGATGGGGGAGCTAGCC  
AAGAAGAATTTTATAAATTTATCAAACCAATTTTAGAAAAAATGGATGGTAC  
TGAGGAATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGG  
ACCTTTGACAACGGCTCTATTCCCCATCAAATTCACCTGGGTGAGCTGCATGC  
TATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTTAAAAGACAATCGTGAG  
AAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGC  
GCGTGGCAATAGTCGTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATT

10

20

30

40

ACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATT  
TATTGAACGCATGACAACTTTGATAAAAAATCTTCCAAATGAAAAAGTACTA  
CCAAAACATAGTTTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAA  
GGTCAAATATGTTACTGAAGGAATGCGAAAACCAGCATTTCTTTCAGGTGAA  
CAGAAGAAAGCCATTGTTGATTACTCTTCAAACAAATCGAAAAGTAACCG  
TTAAGCAATTAAGAAGATTATTTCAAAAAAATAGAATGTTTTGATAGTGT  
TGAAATTTTCAGGAGTTGAAGATAGATTTAATGCTTCATTAGGTACCTACCATG  
ATTTGCTAAAAATTATTAAAGATAAAGATTTTTTGGATAATGAAGAAAATGA  
AGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGAGA  
TGATTGAGGAAAGACTTAAACATATGCTCACCTCTTTGATGATAAGGTGAT  
GAAACAGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAA  
TTGATTAATGGTATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTT  
GAAATCAGATGGTTTTGCCAATCGCAATTTTATGCAGCTGATCCATGATGATA  
GTTTGACATTTAAAGAAGACATTCAAAAAGCACAAAGTGTCTGGACAAGGCGA  
TAGTTTACATGAACATATTGCAAATTTAGCTGGTAGCCCTGCTATTAAAAAAG  
GTATTTTACAGACTGTAAAAGTTGTTGATGAATTGGTCAAAGTAATGGGGCG  
GCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATCAGACAAC  
CAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAAGGT  
ATCAAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTC  
AATTGCAAAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGAGACAT  
GTATGTGGACCAAGAATTAGATATTAATCGTTTAAAGTGATTATGATGTCGATC  
ACATTGTTCCACAAAGTTTCCTTAAAGACGATTCAATAGACAATAAGGTCTTA  
ACGCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTTCCAAGTGAAGAAG  
TAGTCAAAAAGATGAAAACTATTGGAGACAACCTTCTAAACGCCAAGTTAAT  
CACTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGTGGAGGTTTGAGT  
GAACTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAACTCGCCAAA  
TCACTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGA  
TGAAAATGATAAACTTATTCGAGAGGTTAAAGTGATTACCTTAAAATCTAAA  
TTAGTTTCTGACTTCCGAAAAGATTTCCAATTCTATAAAGTACGTGAGATTAA  
CAATTACCATCATGCCCATGATGCGTATCTAAATGCCGTCGTTGGAAGTCTT  
TGATTAAGAAATATCCAAAACCTTGAATCGGAGTTTGTCTATGGTGATTATAAA

10

20

30

40

GTTTATGATGTTTCGTAAAATGATTGCTAAGTCTGAGCAAGAAATAGGCAAAG  
CAACCGCAAAATATTTCTTTTACTCTAATATCATGAACTTCTTCAAAACAGAA  
ATTACACTTGCAAATGGAGAGATTTCGCAAACGCCCTCTAATCGAACTAATG  
GGGAAACTGGAGAAATTGTCTGGGATAAAGGGCGAGATTTTGCCACAGTGCG  
CAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAG  
ACAGGCGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGC  
TTATTGCTCGTAAAAAAGACTGGGATCCAAAAAATATGGTGGTTTTGATAG  
TCCAACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAA  
TCGAAGAAGTTAAAATCCGTAAAGAGTTACTAGGGATCACAATTATGGAAA  
GAAGTTCCTTTGAAAAAAATCCGATTGACTTTTTAGAAAGCTAAAGGATATAA  
GGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAATATAGTCTTTTTGAGT  
TAGAAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATTACAAAAAG  
GAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCTAGT  
CATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGT  
TTGTGGAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGA  
ATTTTCTAAGCGTGTTATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTG  
CATATAACAAACATAGAGACAAACCAATACGTGAACAAGCAGAAAATATTA  
TTCATTTATTTACGTTGACGAATCTTGGAGCTCCCGCTGCTTTTAAATATTTTG  
ATACAACAATTGATCGTAAACGATATACGTCTACAAAAGAAGTTTTAGATGC  
CACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACACGCATTGATTTGA  
GTCAGCTAGGAGGTGACTGA

10

20

30

【 0 1 6 5 】

配列番号 1 4

名称 : g i | 4 0 9 6 9 3 0 3 2 | g b | A F V 3 7 8 9 2 . 1 | C R I S P R 関連タン  
パク質、C s n 1 ファミリー [ ストレプトコッカス・ピオゲネス ( S t r e p t o c o c  
c u s p y o g e n e s ) A 2 0 ]

配列 :

MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS  
GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEED  
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR  
GHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR  
LENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDL  
NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL  
TLLKALVRQQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTE  
ELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKIL  
TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD  
KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF  
KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD  
NEENEDILEDIVLTLTLFEDREMIEERLKTYYAHLFDDKVMKQLKRRRYTGWGRL  
SRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQG  
DSLHEHIANLAGSPAIKKGILQTVKVDELVKVMGRHKPENIVIEMARENQTTQK  
GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ  
ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMK  
NYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL  
DSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYL  
NAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIM  
NFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTE  
VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKKG  
SKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR  
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK  
HYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIHLFTLTNLGA  
PAAFKYFDTTIDRKRYTSTKEVLDTLHQSIITGLYETRIDLSQLGGD

10

20

30

40

【 0 1 6 6 】

配列番号 15

名称 : g i | 1 5 0 3 8 1 3 6 1 | g b | E F 4 7 2 7 6 0 . 1 | U S A インテグラーゼ  
( p o l ) 遺伝子からの H I V - 1 クローン 3 9 B 、 部分的 c d s  
配列 :



TTTTTGGATGGAATAGATAGGGCCCAAGAAGAGCATGAGAAATATCACAATA  
ATTGGAGAGCAATGGCTAGTGATTTTAACCTGCCACCTNTAGTAGCAAAGGA  
GATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGA  
CAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATNTAGAAG  
GAAAAGTTATCCTGGTAGCAGTNCATGTAGCCAGTGGTTATATAGAAGCAGA  
AGTTATTCCAGCAGAGACAGGGCAGGAAACAGCATACTTCCTCTTAAAATTA  
GCAGGAAGATGGCCAGTAAAAACAGTACATACAGACAATGGCAGCAACTTC  
ACCAGTGCTGCGNTGAAGGCCGCCTGTTGGTGGGCAGGGATCAAGCAGGAAT  
TTGGCATTCCCTACAATCCCCAAAGTCAAGGAGTAGTAGAGTCTATGAATAA  
TGAATTAAAGAAAATTGTAGGACAAGTAAGAGATCAGGCTGAGCATCTCAAG  
ACAGCAGTACAAATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGGGGA  
TTGGGGGGGTACAGTGCAGGAGAAAGAATAGTAGACATAATAGCCACAGACA  
TACAAACTAAAGAACTACAAAAAAATATTACAAAAATGCAAAATTTTCGGGT  
CTATTTTCAGAGACAGCAGAGATCCACTTTGGAAAGGACCAGCAAAGCTTCTC  
TGGAAGGTGAAGGGGCAGTAGTAATACAAGATACCAATGACATAAARGTA  
GTGCCARGAAGAAAAGCAAAGATCATTAGAGATTATGGAAAACAGATGGCA  
GGTGATGATTGTGTGGCAAGTAGACAGGNTGAGGATTAG

10

20

30

40

## 【 0 1 6 7 】

配列番号 1 6

名称 : g i | 1 5 0 3 8 1 3 6 2 | g b | A B R 6 8 1 8 2 . 1 | インテグラーゼ、部分的 [ ヒト免疫不全ウイルス 1 ]

配列 :

FLDGIDRAQEEHEKYHNNWRAMASDFNLPPXVAKEIVASCDKCQLKGEAMHGQ  
VDCSPGIWQLDCTHXYEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGR  
WPVKTVHTDNGSNFTSAAXKAACWWAGIKQEFIPYNPQSQGVVESMNNELKK  
IVGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQK  
NITKMQNFRVYFRDSRDPLWKGPALLWKGEAVVIQDTNDIKVVPXRKAKIIR  
DYGKQMAGDDCVASRQXED

## 【 0 1 6 8 】

配列番号 1 7

名称 : g i | 4 5 9 9 8 0 | g b | L 2 0 6 5 1 . 1 | S T L K I A P O L サル T - 細胞リンパ球向性ウイルス I 型インテグラーゼ ( p o l ) 遺伝子、部分的 c d s

配列 :

GACTTGTAGAACGCTCTAATGGCATTCTTAAAACCCTATTATATAAGTACTTT  
ACTGACAAACCCGACCT  
ACCTATGGATAATGCTCTATCCATAGCCCTATGGACGATCAACCACCTGAATG  
TGTTAACCCTACTGCCAC

【 0 1 6 9 】

配列番号 1 8

名称 : g i | 4 5 9 9 8 1 | g b | A A A 4 7 8 4 1 . 1 | インテグラーゼ、部分的 [ サ  
ルト - リンパ球向性ウイルス 1 ]

10

配列 :

L V E R S N G I L K T L L Y K Y F T D K P D L P M D N A L S I A L W T I N H L N  
V L T H C H

【 0 1 7 0 】

配列番号 1 9

名称 : g i | 3 2 1 1 5 6 7 8 4 : 1 - 1 5 0 9 ストレプトコッカス・ニューモニエ ( S  
t r e p t o c o c c u s p n e u m o n i a e ) 組み込みおよび接合エレメント I C  
E S p n 1 1 9 3 0 、株 1 1 9 3 0

配列 :

20

GAGTTTTTTTCCTTTCGTAGCAAGGGTTTAGAGCCCCTATTTTATTTTACTATT  
GTCTAAACACCAAGCG  
AACACCAAAACTACCATGCAATGGAAAAACCTCTGATTTGATTCTCACTTGAT  
TTCACAATCTTTATATC  
AAACTGTGGGTGGTATTTGACAATATCTTTTTTGATTTTAAATAGTAAATTCTG  
AAATAATATTTTTAGGT  
GAGTAACGTGGACTAAGATGTAACAAGTCTTTGAACTCATCGACACTTAATT  
CTACTTTATTGCTATTAT  
CACTAGTTTCAATGAATTTTTCAATTATTCTGGAATATTTACAGGTATAACTTT  
TCAATTCTTCAAAATG  
GAAATTGTGATTTTCTACAAATTGATTTAAGGCTTTTACAGTATTTTCTTGTGA  
ACGATTTATATTATGT  
GTATAGCCCATTGTTGTCTCAAAGTTAGCGTGTCTACTCTAGTCATAATATC  
TTTCACTGCTATGTGCA  
TCTCATTACTTTGAAGGTAACATAATATGCATATGCCTAAACGAATGGGGAGT  
AACATGTTTTACCCACTT  
AAAACCATAGTCACTTAAACAATTTGTCAATAATTTTCCTTCTATTCGTTTCA  
AAATTTGACGAAAAGTG  
CTTGATGTTATTGGAGAGCCGTATTCTGTTCTAAATACACTTTCAGAATGTGT  
AAAAGCAGGACAGGGAT  
GTTTCTCCATATAAGCATCAAACCTCTTTATTTCTCTGTATTGTCCTTTTAAATAG  
CTTCGCTTGCAGCTTC  
AGGCAAAGCTACTTCTCTAATTGAATTGAGTGTTTTAGTTGTATCAAAATGAA  
ATTGTTTAACTTTTAAA  
CAATGATATTGAAGTGCTTTATCAATATGCAAGATTCCTTTTTTCAAAATCAAT  
ATCTGATGGTAAAAATG

10

20

30

CTGCTTCACTAATTCGAATACCTGTAAGCAACAATACTATAGCAAGATCATA  
ATAGTTTGCATTTCTGCA  
TTGGCGTAACACATCAAAAAATGCATGTAATTCATGGATTCTAGAAATTTAG  
AATCATGTCTTTCTTTT  
GCTTTACGCCTTTTCTCTAGTGAAATATCTAGTTTTACCGCAGTCATTGGAGA  
AAACTTAATGACATTAT  
ATAACACACCATGATTAAAAATCTTATTACAAGTACTTTTTATATGAGTCATT  
GTTGAAGGCGATGCATC  
ATACATTTCTAAATATTTATTGAGACTATTTTTCATCAGAAGTGGAGTAATCC  
TGTCTAACAAAAAATCA  
TCTCCTATAATTTTCCCAAGACGCTTCATAACCAGTAGTTCTCTCTGAATTGTT  
TGTGGTTTAACAGAGA  
CACACCAAGTCTGAAACCAATTTTCTTTTAACTCTCCAAATGTTGTAATCAGT  
TCAGGACTATACTGACT  
TTCAAATGAAGTAGTTAGTCTATCTATTTTATCAAGAACCTCTCTTTCAGCTTG  
TTTCCTCGCCCTACTA  
GTATTCTTAGTATAACTTACAGTTACTGATTTCCACTTT

10

20

30

【 0 1 7 1 】

配列番号 2 0

名称 : g i | 3 2 1 1 5 6 7 8 5 | e m b | C B W 3 8 7 6 9 . 1 | インテグラーゼ [ ス  
トレプトコッカス・ニューモニエ ( S t r e p t o c o c c u s p n e u m o n i a e  
) ]

配列 :

MYVVTKTNSKGQPLYQVVEKYKDPLTGKWKSVTVSYTKNTSRARKQAEREVL  
DKIDRLTTSFESQYSPEL  
ITTFGELKENWFQTWCVSVKPQTIQRELLVMKRLGKIIGDDFLLDRITPLLMKNSL  
NKYLEMYDASPSTM  
THIKSTCNKIFNHGVLYNVIKFSPMTAVKLDISLEKRRKAKERHDSKFLEIHELHA  
FFDVLRQCRNANY  
DLAIVLLLTGIRISEAAFLPSDIDFEKGILHIDKALQYHCLKVKQFHFDTTKTLNSIR  
EVALPEAASEAI  
KRTIQRNKEFDAYMEKHPCPAFTHSESVFRTEYGGSPITSSTFRQILKRIEGKLLTNC  
LSDYGFKWVKHVT  
PHSFRHMHISYLSNEMHIAVKDIMTRVGHANFETTMGYTHNINRSQENTVKAL  
NQFVENHNFHFEEELKS  
YTCKYSRIIEKFIETSDNSNKVELSVDEFKDLLHLSPRYSPKNIISNLLKIKKDIVK  
YHPQFDIKIVKS  
SENQIRGFSIAW

10

20

## 【 0 1 7 2 】

配列番号 2 1

名称 : g i | 4 3 0 9 0 : V I I 型ジヒドロ葉酸レダクターゼに対する 1 - 4 3 6 大腸菌  
( E . c o l i ) ( T n 5 0 8 6 ) d h f r V I I 遺伝子および s u l I 遺伝子、 5  
' 末端 ( インテグラーゼ )

配列 :

GCATGCCCGTTCCATACAGAAGCTGGGCGAACAACGATGCTCGCCTTCCAG  
AAAACCGAGGATGCGAAC  
CACTTCATCCGGGGTCAGCACACCGGCAAGCGCCGCGACGGCCGAGGTCTT  
CCGATCTCCTGAAGCCAG  
GGCAGATCCGTGCACAGCACCTTGCCGTAGAAGAACAGCAAGGCCGCCAAT  
GCCTGACGATGCGTGGAGA  
CCGAAACCTTGCGCTCGTTCGCCAGCCAGGACAGAAATGCCTCGACTTCGCT  
GCTGCCCAAGGTTGCCGG  
GTGACGCACACCGTGGAACGGATGAAGGCACGAACCCAGTGGACATAAGC  
CTGTTTCGGTTCGTAAGCTG  
TAATGCAAGTAGCGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAAC  
GCAGCGGTGGTAACGGCG  
CAGTGGCGGTTTTTCAT

30

40

## 【 0 1 7 3 】

配列番号 2 2

50

名称：g i | 4 3 0 9 1 | e m b | C A A 4 1 3 2 5 . 1 | インテグラーゼ、部分的（プラスミド）[ エスケリキア・コリ（E s c h e r i c h i a c o l i ）]

配列：

MKTATAPLPPLRSVKVLDQLRERIRYLHYSLRTEQAYVHWVRAFIRFHGVRHPA

TLGSSEVEAFLSWLAN

ERKVSVSTHRQALAALLFFYGKVLCTDLPWLQEIGRPRPSRRLPVVLTPDEVVRI

LGFLEGEHRLFAQLL

YGTGM

10

【 0 1 7 4 】

配列番号 2 3

> g i | 3 9 7 9 1 2 6 0 5 : 4 0 3 7 2 - 4 1 8 9 8    サーモアナエロバクテリウム（  
T h e r m o a n a e r o b a c t e r i u m）ファージ T H S A - 4 8 5 A、完全ゲノム - レコンビナーゼ

ATGAATCGTGTATGTATTTATCTTAGGAAGTCCCGAGCAGACGAAGAAATAG  
AAAAAGAGCTTGGACAAG  
GAGAAACACTCGCAAAACATCGTAAGGCCCTTCTTAAATTTGCAAAAGAGAA  
AAATTTGAACATAGTAAA  
AATCAGAGAGGAAATAGTATCAGGCGAAAGCCTTATCCATAGACCTGAAATG  
TTGGAATTACTAAAAGAA  
GTCGAACAAGGCATGTACGATGCTGTATTATGTATGGATCTACAGCGTTTAG  
GGCGTGGCAACATGCAGG  
AACAAGGTCTCATTTTAGAAGCCTTTAAAAAGTCAAACACTAAAATTATAAC  
GCTTCAAAAAACTTATGA  
TTTGAACAATGATTTTGACGAAGAATATAGCGAATTTGAAGCATTTATGAGC  
CGAAAGGAACTTAAAATG  
ATAAATAGAAGGCTACAAGGTGGCAGAGTACGCTCTATTCAGGAAGGTAATT  
ATTTATCACCATTTGCCAC  
CTTATGGTTACTTAATACACGAAGAAAAATTTTCGCGCACTCTTGTGCCTAAT  
CCTGAGCAAGCTGATGT  
AGTTAAAATGATTTTTGATATGTATGTCAATAAACAGATGGGGTCTAGTGCTA  
TAGCGAACGAACTAAAC  
AAAATGGGTTATAAGACGTATACTGGCAGGAATTGGGCTTCAAGCTCTGTAA  
TAAACATACTCAAGAATC  
CAGTTTACATCGGTAAAATAACGTGGAAGAAGAAGGATATAAAGAAGTCTGC  
TGACCCAAATAAAAGCAA  
AGATACACGTCAAAGACCACGCTCTGAATGGATTGTATCAGATGGCAAACAT  
GAACCAATAGTGGGCAA  
GAGCTCTTTGCCAAGGCTCAAGAAATCATTA AAAACAAGTATCACATACCGT  
ATCAGATCGTTAATGGTC  
CACGTAACCCATTGGCAGGGCTTATTATATGCAAAATATGTGGCTCTAAAAT  
GGTGTATAGACCCTACAA

10

20

30

40

AGATAAAGAAGCGCATATAATATGTCCAAACAAGTGCGGCAATAAAAGCAG  
CAAATTTATCTATGTAGAA  
AAAAGATTATTACAGGCTTTGGAGGAATGGATGCAAGGCTACGAGCTGGATC  
TGCAAATAGAAGAAGATG  
ACAGCTCTTTTGCAGAAGCACAAAGAGAAACAAAAAGAAGCTCTTGAAAGAG  
AATTGCACGAGCTGCAAAA  
GCAAAAGAACAATTTACACGATTTGCTCGAGCGTGGCATATACGATATAGAT  
ACATTTGTGGAAAGATCT  
ACAATTGTAGCACAGAGAATAGAAGAAACACAGAAAAGTATAGATGTGCTT  
GTGCAAAAAATAGAAGAAG  
AAAAGAATAAAAGAGACAAAGAAAAAATACTTCCGGAAATTCGGCATGTGT  
TGGATCTATATTGGAAAAC  
AGACGACATTGCACAAAAAAATATGTTGTTAAAGAGCGTACTTGAAAAAGCA  
GAATATCTAAAAGAAAAG  
AAGCAGAGAGAAGACAACCTTCGAACCTTGGATTTATCCAAAGCTGCCTGAAA  
AATAG

10

20

30

【 0 1 7 5 】

配列番号 2 4

> g i | 3 9 7 9 1 2 6 6 2 | r e f | Y P \_ 0 0 6 5 4 6 3 2 6 . 1 | レコンビナーゼ  
[ サーモアナエロバクテリウム ( T h e r m o a n a e r o b a c t e r i u m ) ファー  
ジ T H S A - 4 8 5 A ]



MNRVCIYLRKSRADDEIEKELGQGETLAKHRKALLKFAKEKNLNIVKIREEIVSG  
ESLIHRPEMLELLKE  
VEQGMYPDVLQCMDLQRLGRGNMQEQGLILEAFKKSNTKIITLQKTYDLNNDFD  
EEYSEFEAFMSRKELKM  
INRRQLQGGRVRSIQEGNYLSPLPPYGYLIHEEKFSRTLVPNPEQADVVKMIFDMY  
VKNQMGSSAIANELN  
KMGYKTYTGRNWASSSVINILKNPVYIGKITWKKKDIKKSADPNKSKDTRQRPR  
SEWIVSDGKHEPIVGK  
ELFAKAQEIIKNKYHIPYQIVNGPRNPLAGLIICKICGSKMVYRPHYKDKEAHIICPN  
KCGNKSSKFIYVE  
KRLQALEEWMQGYELDLQIEEDDSSFAEAQEKQKEALERELHELQKQKNNLH  
DLLERGIYDIDTFVERS  
TIVAQRIEETQKSIDVLVQKIEEEKNKRDKKILPEIRHVLDLYWKTDDIAQKNML  
LKSIVLEKAEYLKEK  
KQREDNFELWIYPKLPEK

10

20

【 0 1 7 6 】

配列番号 2 5

G i n レコンビナーゼ

> g i | 6 5 7 1 9 3 2 4 0 | s p | Q 3 8 1 9 9 . 2 | G I N \_ B P D 1 0 R e c N  
a m e : F u l l = セリンレコンビナーゼ g i n ; 別名 : F u l l = G - セグメントイン  
ベルターゼ ; 略称 = G i n

30

MLIGYVRVSTNDQNTDLQRNALVCAGCEQIFEDKLSGTRTDRPGLKRALKRLQK  
GDTLVVWKLDRLGRSM  
KHLISLVGELRERGINFRSLTDSIDTSSPMGRFFFHVMGALAEMERELIERTMAG  
LAAARNKGRIGGRP  
PKLTKAWEQAGRLLAQGIPRKQVALIYDVALSTLYKKHPAKRTHIENDDRINQI  
DR

40

【 0 1 7 7 】

配列番号 2 6

C r e レコンビナーゼ

> g i | 3 7 5 3 3 1 8 1 3 | d b j | B A L 6 1 2 0 7 . 1 | C r e レコンビナーゼ [  
C r e - 発現ベクター p H V X 2 - c r e ]

MVQTSLLTVHQNLPALPVDATSDEVKRNLMDFRDRQAFSEHTWKMLLSVCRS  
WAAWCKLNNRKWFPAEP  
EDVRDYLLYLQARGLAVKTIQQHLGQLNMLHRRSGLPRPSDSNAVSLVMRRIRK  
ENVDAGERAKQALAFE  
RTDFDQVRSLMENS DRCQDIRNLAFLGIAYNTLLRIA E IARIRVKDISRTDGGRML  
IHIGRTKTLVSTAG  
VEKALSLGVTKLVERWISVSGVADDPNNYLFCRVRKNGVAAPSATSQLSTRALE  
GIFEATHRLIYGAKDD  
SGQRYLAWSGHSARVGAARDMARAGVSIPEIMQAGGWTVNVNIVMNYIRNLDSE  
TGAMVRLLEDGD

10

20

30

40

50

## 【 0 1 7 8 】

配列番号 2 7 ~ 4 6

これらは、本発明に記載のような、インテグラーゼまたはレコンビナーゼへの連結での使用のための T A L E リピートモジュールをコードするポリヌクレオチドの代表的な配列である。

## 【 0 1 7 9 】

配列番号 2 7

名称：N I

配列：

CTGACCCCAGAGCAGGTCGTGGCAATCGCCTCCAACATTGGCGG  
GAAACAGGCACTCGAGACTGTCCAGCGCCTGCTTCCCGTGCTGTG  
CCAAGCGCACGGA

## 【 0 1 8 0 】

配列番号 2 8

名称：N G

配列：

CTGACCCCAGAGCAGGTCGTGGCCATTGCCTCGAATGGAGGGGG  
CAAACAGGCGTTGGAAACCGTACAACGATTGCTGCCGGTGCTGT  
GCCAAGCGCACGGC

## 【 0 1 8 1 】

配列番号 2 9

名称：H D

配列：

TTGACCCCAGAGCAGGTCGTGGCGATCGCAAGCCACGACGGAGG  
AAAGCAAGCCTTGGAAACAGTACAGAGGCTGTTGCCTGTGCTGT  
GCCAAGCGCACGGG

## 【 0 1 8 2 】

配列番号 3 0

名称 : N N

配列 :

CTTACCCCAGAGCAGGTCGTGGCAATCGCGAGCAATAACGGCGG  
AAAACAGGCTTTGGAACGGTGCAGAGGCTCCTTCCAGTGCTGT  
GCCAAGCGCACGGG

10

## 【 0 1 8 3 】

配列番号 3 1

名称 : N I - N I

配列 :

CTGACCCCAGAGCAGGTCGTGGCAATCGCCTCCAACATTGGCGG  
GAAACAGGCACTCGAGACTGTCCAGCGCCTGCTTCCCGTGCTTTG  
TCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCAATCGCCTC  
CAACATTGGCGGGAAACAGGCACTCGAGACTGTCCAGCGCCTGC  
TTCCCGTGCTGTGCCAAGCGCACGGT

20

## 【 0 1 8 4 】

配列番号 3 2

名称 : N I - N G

配列 :

CTGACCCCAGAGCAGGTCGTGGCAATCGCCTCCAACATTGGCGG  
GAAACAGGCACTCGAGACTGTCCAGCGCCTGCTTCCCGTGCTTTG  
TCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCCATTGCCTC  
GAATGGAGGGGGCAAACAGGCGTTGGAAACCGTACAACGATTG  
CTGCCGGTGCTGTGCCAAGCGCACGGT

30

## 【 0 1 8 5 】

配列番号 3 3

名称 : N I - H D

配列 :

CTGACCCCAGAGCAGGTCGTGGCAATCGCCTCCAACATTGGCGG  
GAAACAGGCACTCGAGACTGTCCAGCGCCTGCTTCCCGTGCTTTG  
TCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCGATCGCAA  
GCCACGACGGAGGAAAGCAAGCCTTGGAAACAGTACAGAGGCT  
GTTGCCTGTGCTGTGCCAAGCGCACGGT

40

## 【 0 1 8 6 】

50

配列番号 3 4

名称 : N I - N N

配列 :

CTGACCCCAGAGCAGGTCGTGGCAATCGCCTCCAACATTGGCGG  
GAAACAGGCACTCGAGACTGTCCAGCGCCTGCTTCCCGTGCTTTG  
TCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCAATCGCGA  
GCAATAACGGCGGAAAACAGGCTTTGGAAACGGTGCAGAGGCT  
CCTTCCAGTGCTGTGCCAAGCGCACGGT

10

【 0 1 8 7 】

配列番号 3 5

名称 : N G - N I

配列 :

CTGACCCCAGAGCAGGTCGTGGCCATTGCCTCGAATGGAGGGGG  
CAAACAGGCGTTGGAAACCGTACAACGATTGCTGCCGGTGCTTTG  
TCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCAATCGCCTC  
CAACATTGGCGGGAAACAGGCACTCGAGACTGTCCAGCGCCTGC  
TTCCCGTGCTGTGCCAAGCGCACGGT

20

【 0 1 8 8 】

配列番号 3 6

名称 : N G - N G

配列 :

CTGACCCCAGAGCAGGTCGTGGCCATTGCCTCGAATGGAGGGGG  
CAAACAGGCGTTGGAAACCGTACAACGATTGCTGCCGGTGCTTTG  
TCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCCATTGCCTC  
GAATGGAGGGGGCAAACAGGCGTTGGAAACCGTACAACGATTG  
CTGCCGGTGCTGTGCCAAGCGCACGGT

30

【 0 1 8 9 】

配列番号 3 7

名称 : N G - H D

配列 :

40

CAAACAGGCGTTGGAAACCGTACAACGATTGCTGCCGGTGCTTTG  
TCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCGATCGCAA  
GCCACGACGGAGGAAAGCAAGCCTTGGAAACAGTACAGAGGCT  
GTTGCCTGTGCTGTGCCAAGCGCACGGT

【 0 1 9 0 】

配列番号 3 8

名称 : N G - N N

配列 :

10

CTGACCCCAGAGCAGGTCGTGGCCATTGCCTCGAATGGAGGGGG  
CAAACAGGCGTTGGAAACCGTACAACGATTGCTGCCGGTGCTTTG  
TCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCAATCGCGA  
GCAATAACGGCGGAAAACAGGCTTTGGAAACGGTGCAGAGGCT  
CCTTCCAGTGCTGTGCCAAGCGCACGGT

20

【 0 1 9 1 】

配列番号 3 9

名称 : H D - N I

配列 :

CTGACCCCAGAGCAGGTCGTGGCGATCGCAAGCCACGACGGAG  
GAAAGCAAGCCTTGGAAACAGTACAGAGGCTGTTGCCTGTGCTTT  
GTCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCAATCGCCT  
CCAACATTGGCGGGAAAACAGGCACTCGAGACTGTCCAGCGCCTG  
CTTCCCGTGCTGTGCCAAGCGCACGGT

30

【 0 1 9 2 】

配列番号 4 0

名称 : H D - N G

配列 :

GAAAGCAAGCCTTGGAAACAGTACAGAGGCTGTTGCCTGTGCTTT  
GTCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCCATTGCCT  
CGAATGGAGGGGGCAAACAGGCGTTGGAAACCGTACAACGATT  
GCTGCCGGTGCTGTGCCAAGCGCACGGT

40

【 0 1 9 3 】

配列番号 4 1

名称 : H D - H D

配列 :

CTGACCCCAGAGCAGGTCGTGGCGATCGCAAGCCACGACGGAG  
GAAAGCAAGCCTTGGAACAGTACAGAGGCTGTTGCCTGTGCTTT  
GTCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCGATCGCA  
AGCCACGACGGAGGAAAGCAAGCCTTGGAACAGTACAGAGGC  
TGTTGCCTGTGCTGTGCCAAGCGCACGGT

【 0 1 9 4 】

配列番号 4 2

10

名称：H D - N N

配列：

CTCACCCCAGAGCAGGTCGTGGCGATCGCAAGCCACGACGGAGG  
AAAGCAAGCCTTGGAACAGTACAGAGGCTGTTGCCTGTGCTTTG  
TCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCAATCGCGA  
GCAATAACGGCGGAAAACAGGCTTTGGAACGGTGCAGAGGCT  
CCTTCCAGTGCTGTGCCAAGCGCACGGA

20

【 0 1 9 5 】

配列番号 4 3

名称：N N - N I

配列：

CTGACCCCAGAGCAGGTCGTGGCAATCGCGAGCAATAACGGCGG  
AAAACAGGCTTTGGAACGGTGCAGAGGCTCCTTCCAGTGCTTTG  
TCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCAATCGCCTC  
CAACATTGGCGGGAAACAGGCACTCGAGACTGTCCAGCGCCTGC  
TTCCCGTGCTGTGCCAAGCGCACGGT

30

【 0 1 9 6 】

配列番号 4 4

名称：N N - N G

配列：

CTGACCCCAGAGCAGGTCGTGGCAATCGCGAGCAATAACGGCGG  
AAAACAGGCTTTGGAACGGTGCAGAGGCTCCTTCCAGTGCTTTG  
TCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCCATTGCCTC  
GAATGGAGGGGGCAAACAGGCGTTGGAACCGTACAACGATTG  
CTGCCGGTGCTGTGCCAAGCGCACGGT

40

【 0 1 9 7 】

配列番号 4 5

名称：N N - H D

50

配列：

CTGACCCCAGAGCAGGTCGTGGCAATCGCGAGCAATAACGGCGG  
AAAACAGGCTTTGGAAACGGTGCAGAGGCTCCTTCCAGTGCTTTG  
TCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCGATCGCAA  
GCCACGACGGAGGAAAGCAAGCCTTGGAAACAGTACAGAGGCT  
GTTGCCTGTGCTGTGCCAAGCGCACGGT

10

【 0 1 9 8 】

配列番号 4 6

名称：NN - NN

配列：

CTGACCCCAGAGCAGGTCGTGGCAATCGCGAGCAATAACGGCGG  
AAAACAGGCTTTGGAAACGGTGCAGAGGCTCCTTCCAGTGCTTTG  
TCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCAATCGCGA  
GCAATAACGGCGGAAAACAGGCTTTGGAAACGGTGCAGAGGCT

20

【 0 1 9 9 】

配列番号 4 7

名称：g i | 7 1 7 9 6 6 1 2 | g b | D Q 0 8 4 3 5 3 . 1 | ヒツジレンチウイルス単  
離株 O v 1 0 インテグラーゼ ( p o l ) 遺伝子、部分的 c d s

配列：

CATAGTAAATGGCATCAAGATGCTATGTCATTGCAGTTAGATTTTGGGATACC  
GAAAGGTGCGGCAGAAG  
ATATAGTACAACAATGTGAAGTATGTCAGGAAAATAAAAATGCCTAGCACCAT  
CAGAGGAAGTAACAAAAG  
AGGGATAGATCATTGGCAGGTGGATTATACTCATTATAAAGACAAAATAATA  
TTGGTATGGGTAGAAACA  
AATTCGGA

30

【 0 2 0 0 】

配列番号 4 8

名称：g i | 7 1 7 9 6 6 1 3 | g b | A A Z 4 1 3 2 5 . 1 | インテグラーゼ、部分的  
[ ヒツジレンチウイルス ]

配列：

HSKWHQDAMSLQLDFGIPKGAAEDIVQQCEVCQENKMPSTIRGSNKRIGIDHWQ  
VDYTHYKDKIILVWVET  
NSG

40

【 0 2 0 1 】

50

## 配列番号 49

> gb | A Y L T 0 1 0 0 0 1 2 7 . 1 | : 1 1 8 0 4 - 1 2 0 4 6 スタフィロコッカス  
・アウレウス ( S t a p h y l o c o c c u s a u r e u s ) 亜種アウレウス ( a u r  
e u s ) S K 1 5 8 5 コンティグ 0 0 0 1 2 7、全ゲノムショットガン配列

TTATAGATAGGTTAGTGACAAAATACATTTTTTCGTCTAGATTAACCGTGCCTC  
TTAGATTATTAATATTT  
TCGTTTAGATGTTTTTCAGAACTTTAGCAACTTCATAATCGTTCATGTAAAG  
TGTTTGGTTTTTTTATTG  
TATAATTAAGTAATTCATAATCTTTGTATACTTCTTTTACTTTATCTATATCAA  
CATTTTCAAGAACAAG  
TTTTTTTATGTTATTATAATTAAAGTTTTCCAT

10

## 【 0 2 0 2 】

## 配列番号 50

> gi | 6 6 9 0 3 5 1 3 0 | gb | K F D 3 0 4 8 3 . 1 | 仮説的タンパク質 D 4 8 4  
\_ 0 2 2 3 4 [ スタフィロコッカス・アウレウス ( S t a p h y l o c o c c u s a u  
r e u s ) 亜種アウレウス ( a u r e u s ) S K 1 5 8 5 ] - S . アウレウス ( s a u  
r e u s ) c a s 9

20

MENFNYYNNIKKLVLNVDIDKVKEVYKDYELLNYTIKNQTLYMNDYEVAKVSE  
KHLNENINNLRGTVNLD  
EKCILSLTYL

## 【 0 2 0 3 】

## 配列番号 51

名称：リンカー 2 の d n a

配列：

a g c g g c a g c g a a a c c c c g g g c a c c a g c g a a a g c g c g a c c c  
c g g a a a g c

## 【 0 2 0 4 】

## 配列番号 52

名称：d C a s 9 タンパク質

配列：

30

40



MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS  
GETAEATR  
LKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIF  
GNIVDEVA  
YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD  
KLFIQLVQ  
TYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL  
GLTPNFKSN  
FDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNT  
EITKAPLS  
ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF  
YKFIKPILEK  
MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNR  
EKIEKILTFR  
IPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN  
LPNEKVLPK  
HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQL  
KEDYFKKI  
ECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDR  
EMIEERLK  
TYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR  
NFMQLIHD  
DSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGR  
HKPENIVIE  
MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYL  
QNGRDMY  
VDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVK  
KMKNYWRQ  
LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMN  
TKYDENDK

10

20

30

40

LIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGITALIKKYPK  
LESEFVYG  
DYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNG  
ETGEIVWD  
KGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK  
KYGGFDSP  
TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKEVKK  
DLIIKLPKY  
SLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQK  
QLFVEQHK  
HYLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGA  
PAAFKYFD TTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD

10

## 【 0 2 0 5 】

配列番号 53

20

名称：A T Gを伴うN L Sヌクレオチド

配列：

ATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACA  
AGGATGAC  
GATGACAAGATGGCCCCCAAGAAGAAGAGGAAGGTGGGCATTACCGCGGG  
GTACCT

30

## 【 0 2 0 6 】

配列番号 54

名称：G G S リンカーヌクレオチド

配列：

G G G G G A A G T

## 【 0 2 0 7 】

配列番号 55

名称：合成インテグラーゼ

配列：

40

ATGTTCTGACGGTATCGACAAAGCTCAGGACGAGCACGAAAAGTACCATT  
CTAACTGGCGCGCCATGG  
CCTCTGACTTCAATCTCCCGCCGGTTGTTGCCAAGGAGATCGTGGCTTCTTGC  
GACAAGTGCCAATTGAA  
GGGTGAGGCTATGCATGGTCAGGTCGATTGCTCTCCCGGTATCTGGCAGCTG  
GACTGCACTCACCTCGAG  
GGTAAGGTGATTCTCGTTGCTGTGCACGTGGCTTCCGGCTACATCGAGGCTGA  
GGTCATCCCGGCTGAGA  
CCGGTCAAGAGACTGCTTACTTCCTGCTCAAGCTGGCCGGCCGTTGGCCAGTT  
AAGACTATTCACTGA  
TAACGGTTCTAACTTTACTTCCGCAACTGTGAAAGCTGCATGCTGGTGGGCCG  
GCATTAAACAAGAGTTC  
GGAATTCCGTATAACCCGCAGTCTCAGGGCGTTGTCGAGTCTATGAACAAGG  
AGCTCAAAAAGATCATTG  
GTCAAGTCCGTGACCAAGCTGAGCACCTTAAGACCGCTGTGCAGATGGCTGT  
TTTTATTCATAACTTCAA  
GCGTAAGGGTGGTATCGGTGGTTATAGCGCTGGTGAGCGTATCGTAGACATC  
ATCGCTACTGATATCCAG  
ACAAAGGAGCTGCAGAAGCAGATCACTAAGATCCAGAACTTCCGTGTGTACT  
ATCGGGACTCTAGGAACC  
CGCTCTGGAAGGGTCCTGCTAAACTGCTGTGGAAGGGAGAGGGTGCTGTTGT  
TATCCAGGACAACCTCTGA  
TATCAAGGTGGTTCCGCGTCGTAAGGCTAAAATTATCCGCGACTACGGCAAG  
CAAATGGCTGGAGACGAC  
TGC GTTGCTAGCCGTCAAGACGAAGACTAA

10

20

30

40

【 0 2 0 8 】

配列番号 5 6

名称：A T Gを伴う d C a s 9ヌクレオチド

配列：

ATGGATAAAAAGTATTCTATTGGTTTAGCTATCGGCACTAATTCCGTTGGATG  
GGCTGTCA

TAACCGATGAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACAC  
AGACCGTC

ATTCGATTAAAAAGAATCTTATCGGTGCCCTCCTATTCGATAGTGGCGAAACG  
GCAGAGG

CGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACC

GAATATGTT

ACTTACAAGAAATTTTGTAGCAATGAGATGGCCAAAGTTGACGATTCTTTCTTT  
CACCGTTT

GGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACATGAACGGCACCCCATC  
TTTGGA

CATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAACGATTTATCACCTC  
AGAAAAA

10

GCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTG  
CCCATATG

ATAAAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAACCTC  
GGATGTC

GACAAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAGA  
ACCCTATA

AATGCAAGTGGCGTGGATGCGAAGGCTATTCTTAGCGCCCGCCTCTCTAAAT  
CCCGACGG

20

CTAGAAAACCTGATCGCACAATTACCCGGAGAGAAGAAAAATGGGTTGTTTCG  
GTAACCTT

ATAGCGCTCTCACTAGGCCTGACACCAAATTTTAAGTCGAACTTCGACTTAGC  
TGAAGAT

GCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACAATCTAC  
TGGCACA

30

ATTGGAGATCAGTATGCGGACTTATTTTGGCTGCCAAAAACCTTAGCGATGC  
AATCCTCC

TATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCCGTTATCCGCT  
TCAATGAT

CAAAAGGTACGATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTC  
CGTCAGCA

ACTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGTAC  
GCAGGTTA

40

TATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTTTATCAAACCCATA  
TTAGAGAA

GATGGATGGGACGGAAGAGTTGCTTGTA AAACTCAATCGCGAAGATCTACTG  
CGAAAGC  
AGCGGACTTTCGACAACGGTAGCATTCCACATCAAATCCACTTAGGCGAATT  
GCATGCTA  
TACTTAGAAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGTGAAAA  
GATTGAGA  
AAATCCTAACCTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGGGAA  
CTCTCGGTT  
CGCATGGATGACAAGAAAGTCCGAAGAAACGATTACTCCATGGAATTTTGAG  
GAAGTTGT  
CGATAAAGGTGCGTCAGCTCAATCGTTCATCGAGAGGATGACCAACTTTGAC  
AAGAATTT  
ACCGAACGAAAAAGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTTACA  
GTGTACAA  
TGAATCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCCTTT  
CTAAGCGG  
AGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGACCAACCGCAAAGT  
GACAGTTA  
AGCAATTGAAAGAGGACTACTTTAAGAAAATTGAATGCTTCGATTCTGTCTGA  
GATCTCCG  
GGGTAGAAGATCGATTTAATGCGTCACTTGGTACGTATCATGACCTCCTAAA  
GATAATTA  
AAGATAAGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATAT  
AGTGTTGA  
CTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATA  
CGCTCACC  
TGTTGACGATAAGGTTATGAAACAGTTAAAGAGGCGTCGCTATACGGGCTG  
GGGACGAT  
TGTCGCGGAAACTTATCAACGGGATAAGAGACAAGCAAAGTGGTAAACTAT  
TCTCGATT  
TTCTAAAGAGCGACGGCTTCGCCAATAGGAACTTTATGCAGCTGATCCATGA

10

20

30

40

TGACTCTTT

AACCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGGGACTC

ATTGCACG

AACATATTGCGAATCTTGCTGGTTCGCCAGCCATCAAAAAGGGCATACTCCA

GACAGTCA

AAGTAGTGGATGAGCTAGTTAAGGTCATGGGACGTCACAAACCGGAAAACAT

TGTAATCG

10

AGATGGCACGCGAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAG

AGCGGAT

GAAGAGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGA

GCATCCTG

TGGAAAATACCCAATTGCAGAACGAGAACTTTACCTCTATTACCTACAAAA

TGGAAGGG

ACATGTATGTTGATCAGGAACTGGACATAAACCGTTTATCTGATTACGACGTC

20

GATGCCAT

TGTACCCCAATCCTTTTTGAAGGACGATTCAATCGACAATAAAGTGCTTACAC

GCTCGGAT

AAGAACCGAGGGAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAA

ATGAAGA

ACTATTGGCGGCAGCTCCTAAATGCGAACTGATAACGCAAAGAAAGTTCGA

TAACTTAA

30

CTAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACAAGGCCGGATTTATTAA

ACGTCAGC

TCGTGGAAACCCGCCAAATCACAAAGCATGTTGCACAGATACTAGATTCCCG

AATGAATA

CGAAATACGACGAGAACGATAAGCTGATTCTGGGAAGTCAAAGTAATCACTTT

AAAGTCA

AAATTGGTGTCGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGA

40

TAAATAAC

TACCACCATGCGCACGACGCTTATCTTAATGCCGTCGTAGGGACCGCACTCAT

TAAGAAA

TACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGACG  
TCCGTAAG  
ATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAATACTTC  
TTTTATTCT  
AACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATAC  
GCAAACGA  
CCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCC  
GGGACTTC  
GCGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTCAACATAGTAAAGAAAA  
CTGAGGTG  
CAGACCGGAGGGTTTTCAAAGGAATCGATTCTTCCAAAAAGGAATAGTGATA  
AGCTCATC  
GCTCGTAAAAAGGACTGGGACCCGAAAAAGTACGGTGGCTTCGATAGCCCTA  
CAGTTGCC  
TATTCTGTCCTAGTAGTGGCAAAGTTGAGAAGGGAAAATCCAAGAACTGA  
AGTCAGTC  
AAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTTTTGAAAAGAACC  
CCATCGAC  
TTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCATAATTAAAC  
TACCAAAG  
TATAGTCTGTTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCG  
GAGAGCTT  
CAAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTCCTGTATT  
TAGCGTCC  
CATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACTTT  
TTGTTGAG  
CAGCACAAACATTATCTCGACGAAATCATAGAGCAAATTCGGAATTCAGTA  
AGAGAGTC  
ATCCTAGCTGATGCCAATCTGGACAAAGTATTAAGCGCATAACAAGCACA  
GGGATAAA  
CCCATACGTGAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACCT

10

20

30

40



CGGCGCTC  
CAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAACGATACTTC  
TACCAAGG  
AGGTGCTAGACGCGACACTGATTCACCAATCCATCACGGGATTATATGAAAC  
TCGGATAG ATTTGTCACAGCTTGGGGGTGACTAA

【 0 2 0 9 】

配列番号 57

10

名称：A B B I E 1 ( N L S - リンカー 1 - インテグラーゼ - リンカー 2 - d C a s 9 )  
- D N A 配列

配列：

ATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACA  
AGGATGAC

GATGACAAGATGGCCCCCAAGAAGAAGAGGAAGGTGGGCATTCACCGCGGG  
GTACCT

20

GGGGGAAGTATGTTCTGACGGTATCGACAAAGCTCAGGACGAGCACGAA  
AAGTACCATTCTAACTGGCGCGCCATGGCCTCTGACTTCAATCTCCCGCCGGT  
TGTTGCCAAGGAGATCGTGGCTTCTTGCGACAAGTGCCAATTGAA

GGGTGAGGCTATGCATGGTCAGGTCGATTGCTCTCCCGGTATCTGGCAGCTG  
GACTGCACTCACCTCGAG

GGTAAGGTGATTCTCGTTGCTGTGCACGTGGCTTCCGGCTACATCGAGGCTGA  
GGTCATCCCGGCTGAGA

30

CCGGTCAAGAGACTGCTTACTTCCTGCTCAAGCTGGCCGGCCGTTGGCCAGTT  
AAGACTATTCACACTGA

TAACGGTTCTAACTTTACTTCCGCAACTGTGAAAGCTGCATGCTGGTGGGCCG  
GCATTAAACAAGAGTTC

GGAATTCCGTATAACCCGCAGTCTCAGGGCGTTGTCGAGTCTATGAACAAGG  
AGCTCAAAAAGATCATTG

GTCAAGTCCGTGACCAAGCTGAGCACCTTAAGACCGCTGTGCAGATGGCTGT  
TTTTATTCATAACTTCAA

40

GCGTAAGGGTGGTATCGGTGGTTATAGCGCTGGTGAGCGTATCGTAGACATC  
ATCGCTACTGATATCCAG

ACAAAGGAGCTGCAGAAGCAGATCACTAAGATCCAGAACTTCCGTGTGTACT  
ATCGGGACTCTAGGAACC

CGCTCTGGAAGGGTCCTGCTAAACTGCTGTGGAAGGGAGAGGGTGCTGTTGT  
TATCCAGGACAACCTCTGA  
TATCAAGGTGGTTCCGCGTCGTAAGGCTAAAATTATCCGCGACTACGGCAAG  
CAAATGGCTGGAGACGAC  
TGCGTTGCTAGCCGTCAAGACGAAGACagcggcagcgaaacccgggaccagcgaaagcgcgaa  
ccccggaaagc  
ATGGATAAAAAGTATTCTATTGGTTTAGCTATCGGCACTAATTCCGTTGGATG  
GGCTGTCA  
TAACCGATGAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACAC  
AGACCGTC  
ATTCGATTAAAAAGAATCTTATCGGTGCCCTCCTATTCGATAGTGGCGAAACG  
GCAGAGG  
CGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACC  
GAATATGTT  
ACTTACAAGAAATTTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTTCTTT  
CACCGTTT  
GGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACATGAACGGCACCCCATC  
TTTGGA  
CATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAACGATTTATCACCTC  
AGAAAAA  
GCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTG  
CCCATATG  
ATAAAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAACCTC  
GGATGTC  
GACAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAGA  
ACCCTATA  
AATGCAAGTGGCGTGGATGCGAAGGCTATTCTTAGCGCCCGCCTCTCTAAAT  
CCCGACGG  
CTAGAAAACCTGATCGCACAATTACCCGGAGAGAAGAAAAATGGGTTGTTTCG  
GTAACCTT  
ATAGCGCTCTCACTAGGCCTGACACCAAATTTTAAGTCGAACTTCGACTTAGC

10

20

30

40

TGAAGAT  
GCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACAATCTAC  
TGGCACAA  
ATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCCAAAAACCTTAGCGATGC  
AATCCTCC  
TATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCCGTTATCCGCT  
TCAATGAT  
CAAAAGGTACGATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTC  
CGTCAGCA  
ACTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGTAC  
GCAGGTTA  
TATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTTTATCAAACCCATA  
TTAGAGAA  
GATGGATGGGACGGAAGAGTTGCTTGTA AAACTCAATCGCGAAGATCTACTG  
CGAAAGC  
AGCGGACTTTCGACAACGGTAGCATTCCACATCAAATCCACTTAGGCGAATT  
GCATGCTA  
TACTTAGAAGGCAGGAGGATTTTTATCCGTTCTCTCAAAGACAATCGTGAAAA  
GATTGAGA  
AAATCCTAACCTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGGGAA  
CTCTCGGTT  
CGCATGGATGACAAGAAAGTCCGAAGAAACGATTACTCCATGGAATTTTGAG  
GAAGTTGT  
CGATAAAGGTGCGTCAGCTCAATCGTTCATCGAGAGGATGACCAACTTTGAC  
AAGAATTT  
ACCGAACGAAAAAGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTTACA  
GTGTACAA  
TGAACTCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCCTTT  
CTAAGCGG  
AGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGACCAACCGCAAAGT  
GACAGTTA

10

20

30

40

AGCAATTGAAAGAGGACTACTTTAAGAAAATTGAATGCTTCGATTCTGTCTGA  
GATCTCCG  
GGGTAGAAGATCGATTTAATGCGTCACTTGGTACGTATCATGACCTCCTAAA  
GATAATTA  
AAGATAAGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATAT  
AGTGTTGA  
CTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATA  
CGCTCACC  
TGTTTCGACGATAAGGTTATGAAACAGTTAAAGAGGCGTCGCTATACGGGCTG  
GGGACGAT  
TGTCGCGGAAACTTATCAACGGGATAAGAGACAAGCAAAGTGGTAAACTAT  
TCTCGATT  
TTCTAAAGAGCGACGGCTTCGCCAATAGGAACTTTATGCAGCTGATCCATGA  
TGACTCTTT  
AACCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGGGACTC  
ATTGCACG  
AACATATTGCGAATCTTGCTGGTTCGCCAGCCATCAAAAAGGGCATACTCCA  
GACAGTCA  
AAGTAGTGGATGAGCTAGTTAAGGTCATGGGACGTCACAAACCGGAAAACAT  
TGTAATCG  
AGATGGCACGCGAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAG  
AGCGGAT  
GAAGAGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGA  
GCATCCTG  
TGGAAAATACCCAATTGCAGAACGAGAACTTTACCTCTATTACCTACAAA  
TGGAAGGG  
ACATGTATGTTGATCAGGAACTGGACATAAACCGTTTATCTGATTACGACGTC  
GATGCCAT  
TGTACCCCAATCCTTTTTGAAGGACGATTCAATCGACAATAAAGTGCTTACAC  
GCTCGGAT  
AAGAACCGAGGGAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAA

10

20

30

40

ATGAAGA  
ACTATTGGCGGCAGCTCCTAAATGCGAAACTGATAACGCAAAGAAAGTTCGA  
TAACTTAA  
CTAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACAAGGCCGGATTTATTAA  
ACGTCAGC  
TCGTGGAAACCCGCCAAATCACAAAGCATGTTGCACAGATACTAGATTCCCG  
AATGAATA  
CGAAATACGACGAGAACGATAAGCTGATTCTGGGAAGTCAAAGTAATCACTTT  
AAAGTCA  
AAATTGGTGTCTGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGA  
TAAATAAC  
TACCACCATGCGCACGACGCTTATCTTAATGCCGTCGTAGGGACCGCACTCAT  
TAAGAAA  
TACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGACG  
TCCGTAAG  
ATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAATACTTC  
TTTTATTCT  
AACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATAC  
GCAAACGA  
CCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCC  
GGGACTTC  
GCGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTCAACATAGTAAAGAAAA  
CTGAGGTG  
CAGACCGGAGGGTTTTCAAAGGAATCGATTCTTCCAAAAAGGAATAGTGATA  
AGCTCATC  
GCTCGTAAAAAGGACTGGGACCCGAAAAAGTACGGTGGCTTCGATAGCCCTA  
CAGTTGCC  
TATTCTGTCCTAGTAGTGGCAAAAGTTGAGAAGGGAAAATCCAAGAACTGA  
AGTCAGTC  
AAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTTTTGAAAAGAACC  
CCATCGAC

10

20

30

40

TTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCATAATTAAAC  
 TACCAAAG  
 TATAGTCTGTTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCG  
 GAGAGCTT  
 CAAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTCCTGTATT  
 TAGCGTCC  
 CATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACTTT  
 TTGTTGAG  
 CAGCACAAACATTATCTCGACGAAATCATAGAGCAAATTCGGAATTCAGTA  
 AGAGAGTC  
 ATCCTAGCTGATGCCAATCTGGACAAAGTATTAAGCGCATACAACAAGCACA  
 GGGATAAA  
 CCCATACGTGAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACCT  
 CGGCGCTC  
 CAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAACGATACACTTC  
 TACCAAGG  
 AGGTGCTAGACGCGACACTGATTCACCAATCCATCACGGGATTATATGAAAC  
 TCGGATAG ATTTGTACAGCTTGGGGGTGACTAA

10

20

30

40

【 0 2 1 0 】

配列番号 58

名称：A B B I E 1 の翻訳（結合に基づくインテグラーゼエディター）

配列：

Met D Y K D H D G D Y K D H D I D Y K D D D D K Met A P K K K R K V G I H R  
 G V P G G S Met F L D G I D K A Q D E H E K Y H S N W R A Met A S D F N L P P  
 V V A K E I V A S C D K C Q L K G E A Met H G Q V D C S P G I W Q L D C T H L  
 E G K V I L V A V H V A S G Y I E A E V I P A E T G Q E T A Y F L L K L A G R W  
 P V K T I H T D N G S N F T S A T V K A A C W W A G I K Q E F G I P Y N P Q S Q  
 G V V E S Met N K E L K K I I G Q V R D Q A E H L K T A V Q Met A V F I H N F  
 K R K G G I G G Y S A G E R I V D I I A T D I Q T K E L Q K Q I T K I Q N F R V Y  
 Y R D S R N P L W K G P A K L L W K G E G A V V I Q D N S D I K V V P R R K A  
 K I I R D Y G K Q Met A G D D C V A S R Q D E D S G S E T P G T S E S A T P E S  
 Met D K K Y S I G L A I G T N S V G W A V I T D E Y K V P S K K F K V L G N T

DRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNR  
ICYLQEIFSNEMetAKVDDSSFFHRLEESFLVEEDKKHERHPI  
FGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL  
AHMetIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFE  
ENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGL  
FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN  
LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS  
ASMetIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFDQSKN  
GYAGYIDGGASQEEFYKFIKPILEKMetDGTEELLVKLNRE  
DLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPLKDN  
REKIEKILTRIPYYVGPLARGNSRFAWMetTRKSEETITPW  
NFEEVVDKGASAQSFIERMetTNFDKNLPNEKVLPKHSLLY  
EYFTVYNELTKVKYVTEGMetRKPAFLSGEQKKAIVDLLF  
KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT  
YHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMetIE  
ERLKYAHLFDDKVMetKQLKRRRYTGWGRLSRKLINGIR  
DKQSGKTILDFLKSDGFANRNFMetQLIHDDSLTFKEDIQK  
AQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVK  
VMetGRHKPENIVIEMetARENQTTQKGQKNSRERMetKRIEE  
GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMetYV  
DQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNR  
GKSDNVPSEEVVKKMetKNYWRQLLNAKLITQRKFDNLTK  
AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMetNT  
KYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNY  
HHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVR  
KMetIAKSEQEIGKATAKYFFYSNIMetNFFKTEITLANGEIR  
KRPLIETNGETGEIVWDKGRDFATVRKVLSMetPQVNIVKK  
TEVQTGGFSKESILPKRNSDKLIARKKDWDPPKKYGGFDSP  
TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMetERSSEFEK  
NPIDFLEAKGYKEVKKDIIKLPKYSLELENGRKRMetLA  
SAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQ

10

20

30

40

KQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAY  
 NKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKR  
 YTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD

停止

ドナーDNAに対する（インテグラーゼ認識のためのLTR領域のatt部位）。

【0211】

配列番号59

名称：U3att

配列：

ACTGGAAGGGCTAATTCACTCCCAAAGAA

【0212】

配列番号60

名称：U5att

配列：

GACCCCTTTTAGTCA GTGTGGA AATCTCTAGCAGT

NLS - リンカー1 - インテグラーゼ - リンカー2 - dCas9またはインテグラーゼ -  
 リンカー1 - NLS - リンカー2 - dCas9またはインテグラーゼ - リンカー2 - dC  
 as9 - リンカー1 - NLSまたはインテグラーゼ - リンカー2 - dCas9 - NLS  
 リンカー1 = GGS

【0213】

配列番号61

名称：リンカー2

配列：

SGSETPGTSESATPES

【0214】

配列番号62

名称：MMTVインテグラーゼcDNA、gb | AF071010.1 | : 16 - 111  
 3マウス乳癌ウイルス推定インテグラーゼ、envポリタンパク質およびスーパー抗原m  
 RNA、完全cds

配列：

10

20

30



ATGACAGGAAAGTGGCCTTGTATTTACTCCACTAACTGCAGAGATGTGTTGC  
ATGGGACGGGGGGGCACTG  
CACCAGCCCTCGTGCTGAATTCGGCACGAGGAAATGCCTATGCAGATTCTTTA  
ACAAGAATTCTGACCGC  
TTTAGAGTCAGCTCAAGAAAGCCACGCACTGCACCATCAAAATGCCGCGGCG  
CTTAGGTTTCAGTTTCAC  
ATCACTCGTGAACAAGCACGAGAAATAGTAAAATTATGTCCAAATTGCCCCG 10  
ACTGGGGACATGCACCAC  
AACTAGGAGTAAACCCTAGGGGCCTTAAGCCCGGGGTTCTATGGCAAATGGA  
TGTTACTCATGTCTCAGA  
ATTTGGAAAATTAAAGTATGTACATGTGACAGTGGATACTTACTCTCATTTTA  
CTTTCGCTACCGCCCGG  
ACGGGCGAAGCAGCCAAAGATGTGTTACAACACTTGGCTCAAAGCTTTGCAT  
ACATGGGCATTCCCTCAA 20  
AAATAAAAACAGATAATGCCCTGCCTATGTGTCTCGTTCAATACAAGAATTT  
CTGGCCAGATGGAAAAT  
ATCTCACGTCACGGGGATCCCTTACAATCCCCAAGGACAGGCCATTGTTGAA  
CGAACGCACCAAAATATA  
AAGGCACAGATTAATAAACTTCAAAAGGCTGGAAAATACTATACACCCCACC  
ATCTATTGGCACATGCTC  
TTTTTGTGCTGAATCATGTAAATATGGACAATCAAGGCCATACAGCGGCCGA 30  
AAGACATTGGGGTCCAAT  
CTCAGCCGATCCAAAACCTATGGTCATGTGGAAAGACCTTCTCACAGGGTCC  
TGGAAGGACCCGATGTC  
CTAATAACAGCCGGACGAGGCTATGCTTGTGTTTTTCCACAGGATGCCGAATC  
ACCAATCTGGGTCCCCG  
ACCGGTTTCATCCGACCTTTTACTGAGCGGAAAGAAGCAACGCCACACCTGG  
CACTGCGGAGAAAACGCC 40  
GCCGCGAGATGAGAAAGATCAACAGGAAAGTCCGGAGGATGAATCTTGCCC  
CCATCAAAGAGAAGACGGC  
TTGGCAACATCTGCAGGCGTTAATCTCCGAAGCGGAGGAGGTTCTTAA

【 0 2 1 5 】

配列番号 6 3

名称 : g i | 3 2 7 3 8 6 6 | g b | A A C 2 4 8 5 9 . 1 | 推定インテグラーゼ [ マウ  
ス乳癌ウイルス ]

配列 :

10

20

30

40

50

MTGKWPCISTNCRDVLHGTGGTAPALVLNSARGNAYADSLTRILTALESAQES  
HALHHQNAAALRFQFH  
ITREQAREIVKLCPCNCPDWGHAPQLGVNPRGLKPGVLWQMDVTHVSEFGKLKY  
VHVTVDITYSHFTFATAR  
TGEAAKDVLQHLLAQSFAYMGIPQKIKTDNAPAYVSRISIQEFLARWKISHVTGIPY  
NPQGQAIVERTHQNI  
KAQINKLQKAGKYYTPHLLAHALFVLNHNMDNQGHATAAERHWGPISADPKP  
MVMWKDLLTGSWKGPDV  
LITAGRGYACVFPQDAESPIWVPDRFIRPFTERKEATPTPGTAEKTPPRDEKDQQE  
SPEDESCPHQREDG  
LATSAGVNLRS GGGS

10

【 0 2 1 6 】

配列番号 6 4

20

名称 : g b | A X U N 0 2 0 0 0 0 5 9 . 1 | : 5 1 1 6 - 8 8 5 0 ヨウンギイバクター  
・フラジリス ( Y o u n g i i b a c t e r f r a g i l i s ) 2 3 2 . 1 コンティグ  
\_ 1 5 1、全ゲノムショットガン配列 - レコンビナーゼ  
配列 :

TTGAAAGATAACGATAAAAGGATGTGGGTTTCAGACTTTATGGAATCCCATCA  
ATGAAAGACATAAAAGTC  
CACTGGATAGCCCAGAACCAGGGATTAAAGTAGCGGCCTACTGCAGAGTAAG  
CATGAAAGAGGAGGAACA  
ACTCCGGTCATTGGAAAACCAGGTGCATCACTATACTCATTTTATCAAAAGTA  
AGCCGAATTGGAGATTT  
GTAGGGGTTTATTACGATGATGGCATAAGTGCAGCCATGGCAAGTGGGAGAA  
GAGGGTTCCAGCGGATTA  
TCCGTCATGCTGAAGAAGGTAAGGTTGATCTGATTCTAACAAAGAATATTTC  
ACGGTTTTCCAGAAATTC  
CAAGGAGTTACTGGATATAATCAATCAACTGAAAGCTATCGGTGTGGGCATC  
TATTTTGAGAAAGAGAAT  
ATTGATACTTCAAGAGAGTACAATAAATTCCTCTTAAGCACTTATGCTGCGCT  
GGCACAGGAAGAGATAG  
AAACTATTTCAAACCTCTACGATGTGGGGTTATGAGAAAAGGTTTCTAAAGGG  
TATCCCAAAGTTCAACCG  
CTTATATGGATACAAAGTCATCCATGCAGGGGATGATTCCCAATTGATTGTTC  
TTGAAGATGAAGCAAAA

10

20

ATCGTAAGAATGATGTATGAACAGTACCTTCAAGGGAAGACGTTCACTGATA  
TTGCAAGGGGCGCTAACAG  
AAGCTGGAGTGAAAACAGCCAAAGGGAAGGATGTCTGGATAGGCGGCATGA  
TAAAGCATATTTTATCCAA  
CGTCACCTACACCGGTAACAAGCTTACACGAGAACTGAAAAGAGATTTATTT  
ACGAACAAAGTTAATAGC  
GGTGAACGGGATCAGGTTTTTATAGGAAACACTCACGAACCGATCATCAGCA  
ATGATATTTTCAATCTTG  
TTCAAAGAAGCTTGAGGCCAATACGAAGGAAAGAAAGCCCAGTGAGAAGC  
GAGAGAAGAACCACATGTC  
TGGTCGGCTACTTTGCGGAAGATGTGGATACAGTTTTACCATAATTCACAATA  
GAGCTTCTCATCACTTT  
AAGTGTAGCCCTAAAATCATGGGGGTCTGTGATTCTGAACTTTATCGGGATGC  
GGATATTTCGAGAAATGA  
TGATGAGGGCAATGTATATAAAATATGACTTCACCGATGAAGACATAGTACT  
AAAAGTCTGAAGGAACT  
CCAGGTCATCAATCAAAATGATCACTTTGAGTTTCATAGGCTAAAGTTTATCA  
CTGAAATTGAAATCGTA  
AAAAGGCAGCAGGCCATTTTCAGATAGATATTCAGCTATTAGCATAGAAAAAA  
TGGAAGAAGAATACCGCA  
CTTTTGAAAGCAAGATTGCGAAAATTGAGGATGACAGGTACATCAGAATCGA  
TGCAGTGGAGTGGTTAAA  
GAAAAACAAGACGCTGGATTCTTTTATCGCTCAGGTCACCACTAAAATATTG  
CGAGCTTGGGTTTCCGAG  
ATGACTGTTTATACACGAGATGACTTTTTAGTGCAGTGGATTGACGGAACTCA  
AACTGAGATAGGAAGCT  
GCGAGCATCATCTTGTGAAGGATAGAAATAGTAAGAGTTACGAGTCCGGTGA  
AGAAACGAGCAGGAGGGC  
CAAATTTGAAGTCAACCACATTAGTGAAACCACCGAAGGACAAGGAGAACTT  
GATCTCTTAAGCAAGAGT

10

20

30

40

GCAAGTTCAAACAATGAAGATAGTAATCAACCAGAAAATAATTCTACGGGAA  
AGGAGGAGCTTGAATTGA  
ACTTAAACAGTAATGCAGAAATTATCAAATGAGCCCGGGCAAAGGGACTA  
TATTATGAAGAATTTGCA  
CAAGAGCCTGAGTGCAAATATGATGATGCAAAATGCTTCAGTACACACGGCA  
AGTATTAACAAACCTAGA  
CTTAAGACTGCTGCTTACTGCAGAATCTCAACAGATTCAGAAGAACAAAAGG  
TAAGCTTGAAAACCCAAG  
TAGCCTATTACACTTATCTGATTCTAAAGGATCCCCAATATGAATATGCAGGC  
ATCTATGCCGATGAAGG  
TATATCAGGGCGTTCTATGAAAAACCGTACAGAATTTCTCAAACACTACTCGAA  
GAATGTAAAGCCGGGAAT  
GTGGACTTGATTTTAACCAAGTCAATCTCACGGTTTAGCAGAAACGCATTAG  
ATTGCTTGGAACAGATCA  
GGATGCTGAAGTCGCTGCCAAGTCCAGTTTATGTGTATTTTGAGAAAGAGAA  
TATTCATACAAAAGATGA  
GAAGAGTGAGCTGATGATTTCTATTTTGGGAAGTATCGCTCAGGAAGAGAGC  
GTAAACATGGGAGAAGCC  
ATGGCTTGGGGAAAACGGAGATATGCTGAGAGAGGGATAGTAAACCCAAGT  
GTTGCACCTTATGGATATA  
GAACGGTCAGAAAAGGTGAATGGGAGGTGGTTGAAGAAGAAGCTACGATCA  
TTAGAAGAATTTATCGGAT  
GCTCCTAAGTGGAAGAGTATTCATGAAATCACAAAGGAGCTCTCCATGGAG  
AAGATAAAGGGTCCTGGC  
GGCAACGAGCAGTGGCATCTTCAAACCATTAGAAATATCTTGAGAAATGAAA  
TCTATAGGGGTAACTACC  
TTTATCAAAGGCTTATATCAAGGACACGATCGAGAAGAAGGTGGTAATGAA  
TCGAGGAGAACTGCCACA  
GTATCTCATAGAGAATCATCATAAAGCCATTGTTGACAATGAGACCTGGGAA  
AAGGTCCAGAAGGTACTA

10

20

30

40

GAAGCCAGAAGGGAAAAATATGAGAATAAAAAGTCCATAACTTATCCTGAA  
GACAAAATGAAAAACGCTT  
CTCTTGAAGATATTTTTACCTGTGGAGAATGTGGAAGTAAAATAGGCCATAG  
AAGGAGCATCCAGAGCTC  
TAATGAGATTTCATTCCTGGATCTGCACAAAAGCCGCTAAGTCTTTCTTGGTGG  
ACTCGTGTAAGTCCACA  
AGCGTATATCAGAAGCACCTGGAGCTGCATTTTATGAAGACTCTTCTCGATAT 10  
TAAAAAGCATCGTTCTT  
TCAAAGATGAGGTGCTCACCTATATTCTGAACCCAAGAAGTAGATGAAAAGGA  
AGAGTGGAGAATCAAAGT  
CATAGAGAAACGAATCAAAGATCTTAACAGAGAGCTTTATAATGCGGTAGAC  
CAGGAGCTCAATAAAAAA  
GGTCAGGACTCCAGGAAAGTTGATGAGCTCACAGAGAAAATTGTGGATCTTC  
AAGAGGAATTAAAGGTGT 20  
TTAGGGACCGAAAGGCAAAGGTTGAGGATCTTAAAGCTGAGCTTGAATGGTT  
CCTAAAGAAGCTGGAAAC  
CATTGATGACGCTCGAGTAAAAAGAAATGAAGGAATAGGCCACGGTGAAGA  
GATCTACTTCAGAGAAGAT  
ATTTTTGAAAGAATAGTAAGGAGTGCACAGCTTTATAGCGATGGAAGGATCG  
TCTACGAACTAAGCCTCG  
GGATCCAGTGGTTCATTGACTTTAAATACAGCGCATTTCAGAAGCTTCTTATA 30  
AAGTGGAAGGATAAACA  
AAGGGCAGAAGAAAAAGAGGCTTTTCTTGAGGGGCCGGAAGTTAAAGAGCT  
GCTGGAATTTTGTAAGGAA  
CCGAAGAGCTACTCTGATTTACATGCCTTCATGTGTGAGAGAAAAGAGGTGT  
CTTATAGCTATTTTCAGGA  
AATTGGTGATAAGACCTTTGATGAAGAAAGGAAAGCTGAAGTTCACCATACC  
AGAAGATGTTATGAATAG 40  
GCATCAGAGATACACATCAATCTAA

10

20

30

40

50

【 0 2 1 7 】

配列番号 6 5

名称 : g i | 5 6 4 1 3 5 6 4 5 | g b | E T A 8 1 8 2 9 . 1 | レコンビナーゼ [ ヨウ  
ンギイバクター・フラジリス ( Y o u n g i i b a c t e r f r a g i l i s ) 2 3 2  
. 1 ]

配列 :

MKDNDKRMWVQTLWNPINERHKSPLDSPGKIVAAAYCRVSMKEEEQLRSLEN  
QVHHYTHFIKSKPNWRF  
VGVEYDDGISAAMASGRRGFQRIIRHAEKGVDLILTKNISRFSRNSKELLDIINQ  
LKAIGVGIYFEKEN  
IDTSREYNKFLSTYAALAQEEIETISNSTMWGYEKRFKLGIPKFNRLYGYKVIHA  
GDDSQLIVLEDEAK  
IVRMMYEQYLQGKTFTDIARALTEAGVKTAKGKDVWIGGMIKHILSNVTYTGN  
KLTRELKRDFTNKVNS  
GERDQVFIGNTHEPIISNDIFNLVQKKLEANTKERKPSEKREKNHMSGRLLCGRC  
GYSFTIIHNRASHHF  
KCSPKIMGVCDSELYRDADIREMMMRAMYIKYDFTDEDIVLKLKELQVINQND  
HFEFHRLKFITEIEIV  
KRQQAISDRYSAISIEKMEEEYRTFESKIAKIEDDRYIRIDAVEWLKKNKTLDSFIA  
QVTTKILRAWVSE  
MTVYTRDDFLVQWIDGTQTEIGSCEHHLVKDRNSKSYESGEETSRRAKFEVNHIS  
ETTEGQGELDLLSKS  
ASSNEDSNQPENNSTGKEELELNLNSNAEIIKIEPGQRDYIMKNLHKSLSANMM  
MQNASVHTASINKPR  
LKTAAYCRISTDSEEQKVSLKTQVAYTYLILKDPQYIYAGIYADEGISGRSMKN  
RTEFLKLEECKAGN  
VDLILTKSISRFSRNALDCLEQIRMLKSLPSPVYVYFEKENIHTKDEKSELMISIFGS  
IAQEE SVN MGEA  
MAWGKRRYAERGIVNPSVAPYGYRTVRKGEWEVVEEEATIIRRIYRMLLSGKSI  
HEITKELSMKIKGPG  
GNEQWHLQTIRNILRNEIYRGNYLYQKAYIKDTIEKKVVMNRGELPQYLIENHH  
KAIVDNETWEKVQKVL  
EARREKYENKKSITYPEDKMKNASLEDIFTCGECGSKIGHRRSIQSSNEIHSWICT  
KAAKSFLVDSCKST  
SVYQKHLELHFMKTLLDIKKHRSFKDEVLT YIRTQE VDEKEEWRIK VIEKRIKDL  
NRELYNAVDQELNKK

10

20

30

40

GQDSRKVDELTEKIVDLQEELKVFRDRKAKVEDLKAELEWFLKKLETIDARVK  
RNEGIGHGEEIYFRED  
IFERIVRSAQLYSDGRIVYELSLGIQWFIDFKYSAFQKLLIKWKDKQRAEEKEAFL  
EGPEVKELLEFCKE  
PKSYSDLHAFMCERKEVSYSYFRKLVIRPLMKKGKLFKFTIPEDVMNRHQRYTSI

## 【 0 2 1 8 】

配列番号 6 6

10

名称：g i | 5 7 1 2 6 4 5 4 3 : 1 6 4 2 3 - 1 6 7 7 0 クロストリジウム・ディフィ  
シル ( C l o s t r i d i u m d i f f i c i l e ) トランスポゾン T n 6 2 1 8 、株  
O x 4 2 トランスポサーゼ

配列：

TTAGTCTTCAAAAGGTTTTGGACTAAATTTACTCTCGTAGTCAGGTCCAAGTG  
TTTCTTCAGATTTTTTT  
TTCAACCAATCCACCTGCATGGTGAGCTGGCCAACTTTTTTCGCATATTCAGC  
TTTTTCCTTGCGTTCTA  
AAGCGAGTTTTTCTTTCAGATTATCCTCTCGTGTGTCATTAAAAACACGGAT  
GCTTTATCGAGGAACTC  
CTTCTTCCAGTTGCGGAGAAGATTCGGCTGAATATTGTTTTCGGTTGCGATTG  
TATTTAAGTCTTTTTCT  
CCTTTGAGCAGTTCAATCACTAATTCTGATTTGAATTTGGCAGAGAAATTTCT  
TCTTGTTTCGAGACAT

20

30

## 【 0 2 1 9 】

配列番号 6 7

名称：g i | 5 7 1 2 6 4 5 5 9 | e m b | C D F 4 7 1 3 3 . 1 | トランスポサーゼ [  
ペプトクロストリジウム・ディフィシル ( P e p t o c l o s t r i d i u m d i f f  
i c i l e ) ]

配列：

MSRTRRNFSAKFKSELVIELLKGEKDLNTIATENNIQPNLLRNWKKEFLDKASVV  
FNDTREDNLKEKLAL  
ERKEKAHEYAKKVGQLTMQVDWLKKKSEETLGPDIYESKFSPKPFED

40

## 【 0 2 2 0 】

配列番号 6 8

名称：g b | C P 0 0 9 4 4 4 . 1 | : 1 3 1 7 7 2 4 - 1 3 2 0 5 4 3 フランシセラ・  
フィロミラジア ( F r a n c i s e l l a p h i l o m i r a g i a ) 株 G A 0 1 - 2  
8 0 1 、完全ゲノム C p f 1

配列：



ATGAATCTATATAGTAATCTAACAAATAAATATAGTTTAAGTAAAACCTCTAA  
GATTTGAGTTAATTCCAC

AGGGTGAAACACTTGAAAATATAAAAAGCAAGAGGTTTGATTTTAGATGATGA  
GAAAAGAGCTAAAGACTA

TAAAAAAGCTAAACAAATCATTGATAAATATCATCAGTTTTTTATAGAGGAG  
ATATTAAGTTCGGTATGT

10

ATTAGCGAAGATTTATTACAAAACCTATTCTGATGTTTATTTTAAACTTAAAAA  
GAGTGATGATGATAATC

TACAAAAAGATTTTAAAAGTGCAAAAGATACGATAAAGAAACACATATCTAG  
ATATATAAATGACTCGGA

GAAATTTAAGAATTTGTTTAATCAAAATCTTATAGATGCTAAAAAAGGGCAA  
GAGTCAGATTTAATTCTA

TGGCTAAAGCAATCTAAGGATAATGGCATAGAACTATTTAAAGCTAACAGTG  
ATATCACAGACATAGATG

20

AGGCGTTAGAAATAATCAAATCTTTTAAAGGTTGGACAACTTATTTTAAGGGT  
TTTCATGAAAATAGAAA

AAATGTCTATAGTAGTGATGATATCCCTACATCTATTATTTATAGAATAGTAG  
ATGATAATTTGCCTAAA

TTTATAGAAAATAAAGCTAAGTATGAGAATTTAAAAGACAAAGCTCCAGAAG  
CTATAAACTATGAACAAA

30

TTAAAAAAGATTTGGCAGAAGAGCTAACCTTTGATATTGACTACAAAACATC  
TGAAGTTAATCAAAGAGT

TTTTTCACTTGATGAAGTTTTTGAGATAGCAAACCTTTAATAATTATCTAAATC  
AAAGTGGTATTACTAAA

TTTAATACTATTATTGGTGGTAAATTTGTTAATGGTGAAAATACAAAGAGAA  
AAGGTATAAATGAATATA

TAAATCTATACTCACAGCAAATAAATGATAAAACACTTAAAAAATATAAAAT  
GAGTGTTTTATTTAAGCA

40

AATTTTAAGTGATACAGAATCTAAATCTTTTGTAATTGATAAGTTAGAAGATG  
ATAGTGATGTAGTTACA

ACGATGCAAAGTTTTTATGAGCAAATAGCAGCTTTTAAAACATTAGAAGAAA  
AGTCTATTAAGGAAACAT

50

TATCTTTACTATTTGATGATTTAAAAGCTCAAAAAGCTTGATTTGAGTAAAATT  
TATTTTAAAAATGATAA  
ATCTCTTACTGATCTATCACAACAAGTTTTTGGATGATTATAGTGTTATTGGTAC  
AGCGGTACTAGAATAT  
ATAACTCAACAAGTAGCACCTAAAAATCTTGATAACCCTAGTAAGAAAGAGC  
AAGATTTAATAGCCAAAA  
AAACTGAAAAAGCAAAATACTTATCTCTAGAACTATAAAGCTTGCCTTAGA  
AGAATTTAATAAGTATAG  
AGATATAGATAAACAGTGTAGGTTTGAAGAAATATTTGCAAGCTTTGCAGAT  
ATTCCGGTGCTATTTGAT  
GAAATAGCTCAAAACAAAAACAATTTGGCACAGATATCTATCAAATATCAAA  
ATCAAGGTAAAAAAGACC  
TGCTTCAAAGTAGTGCAGAAGTAGATGTTAAAGCTATCAAGGATCTTTTGGAT  
CAAACATAATACTCTT  
GCATAAACTAAAAATATTTTCATATTACGCAATCAGAAGATAAGGCAAATATT  
TTAGACAAGGATGAGCAT  
TTTTATTTAGTATTTGATGAGTGCTACTTTGAGCTAGCGAATATAGTGGCTCTT  
TATAACAAAATTAGAA  
ACTATATAACTCAAAAGCCATATAGTGATGAGAAATTTAAGCTCAATTTTGA  
GAACTCAACTTTAGCCAA  
TGGTTGGGATAAAAAATAAAGAGCCTGACAATACGGCAATTTTATTTATCAAA  
GATGATAAATATTATCTG  
GGTGTGATGAACAAGAAAAATAACAAAATATTTGATGATAAAGCTATCAAAG  
AAAATAAAGGTGAAGGAT  
ATAAGAAAGTTGTATATAAACTTTTACCCGGTGCAAATAAAATGTTACCTAA  
GGTTTTCTTTCTGCTAA  
ATCTATAAATTTTTATAATCCTAGTGAAGATATACTTAGAATAAGAAACCACT  
CAACACATACAAAAAAT  
GGTAGTCCTCAAAAAGGATATGAAAAAGCTTGAGTTTAATATTGAAGATTGCC  
GAAAATTTATAGATTTTT

10

20

30

40

ATAAACATTCTATAAGTAGGCATCCAGAGTGGAAAGATTTTGGATTTAGATTT  
TCTGATACTAAAAAATA  
CAACTCTATAGATGAATTTTATAGAGAAGTTGAAAATCAAGGCTACAACTA  
ACTTTTGAAAATATATCA  
GAAAGCTATATTGATAGTTTAGTCGATGAAGGCAAATTATACCTATTCCAAAT  
CTATAATAAAGATTTCT  
CAGTATATAGTAAGGGTAAACCAAATTTACATACGCTATATTGGAAGGCGTT  
GTTTGATGAGAGAAATCT  
CCAAGATGTAGTATATAAATTAATGGTGAAGCAGAACTCTTCTATCGTAAA  
CAATCAATACCTAAGAAA  
ATCACTCACCCAGCCAAAGAGGCAATAGCTAATAAAAAACAAAGATAATCCTA  
AAAAAGAGAGTATTTTGT  
AATATGATTTAATCAAAGATAAACGCTTTACTGAAGATAAGTTTTCTTTCAC  
TGTCCTATTACAATCAA  
TTTCAAATCTAGTGGAGCTAATAAGTTTAATGATGAAATCAATTTATTGCTAA  
AAGAAAAAGCAAATGAT  
GTTCATATCCTAAGTATAGATAGAGGAGAAAGACATTTAGCTTACTATACTTT  
GGTAGATGGTAAAGGAA  
ACATTATCTGTAAGAATTAA

10

20

30

【 0 2 2 1 】

配列番号 6 9

名称 : g i | 7 5 4 2 6 4 8 8 8 | g b | A J I 5 7 2 5 2 . 1 | C R I S P R 関 連 タ ン  
パ ク 質 C p f 1、 サ ブ タ イ プ P R E F R A N [ フ ラ ン シ セ ラ ・ フ ィ ロ ミ ラ ジ ア ( F r a n  
c i s e l l a p h i l o m i r a g i a ) ]

配列 :

MKTNYHDKLAAIEKDRESARKDWKKINNIKEMKEGYLSQVVHEIAKL VIGYNAI  
VVFEDLNFGFKRGRFK  
VEKQVYQKLEKMLIEKLNVLVKDNEFDKAGGVLRAYQLTAPFETFKKMGKQT  
GIIYYVPADFTSKICPV  
TGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKW  
TIASFGSRLINFRNSD  
KNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAIYAENDKKFFAKLTSILNS  
ILQMRNSKTGTEDY  
LISPVADVNGNFFDSRHAPKNMPQDADANGAYHIGLKGLMLLYRIKNNQDGKK  
LNLVIKNEEYFEFVQNR  
NKSSKI

10

【 0 2 2 2 】

配列番号 7 0

名称 : g i | 4 3 8 6 0 9 | g b | L 2 1 1 8 8 . 1 | H I V 1 N Y 5 A ヒト免疫不全ウ  
イルス 1 型インテグラーゼ遺伝子、3 ' 末端  
配列 :

20

TTCCTGGACGGTATCGATAAAGCTCAGGAAGAACACGAAAAATACCACTCTA  
ACTGGCGCGCCATGGCTT  
CTGACTTCAACCTGCCGCCGGTTGTTGCCAAGGAAATCGTGGCTTCTTGCGAC  
AAATGCCAATTGAAAGG  
TGAAGCTATGCATGGTCAGGTCGACTGCTCTCCAGGTATCTGGCAGCTGGACT  
GCACTCATCTCGAGGGT  
AAAGTTATCCTGGTTGCTGTTCACGTGGCTTCCGGATACATCGAAGCTGAAGT 10  
TATCCCGGCTGAAACCG  
GTCAGGAAACTGCTTACTTCCTGCTTAAGCTGGCCGGCCGTTGGCCGGTTAAA  
ACTGTTCACTGACAA  
CGGTTCTAACTTCACTAGTACTACTGTAAAGCTGCATGCTGGTGGGCCGGCA  
TCAAACAGGAGTTCGGG  
ATCCCGTACAACCCGCAGTCTCAGGGCGTTATCGAATCTATGAACAAAGAGC  
TCAAAAAAATCATTGGCC 20  
AGGTACGTGATCAGGCTGAGCACCTGAAAACCGCGGTGCAGATGGCTGTTTT  
CATCCACAACCTCAAACG  
TAAAGGTGGTATCGGTGGTTACAGCGCTGGTGAACGTATCGTTGACATCATC  
GCTACTGATATCCAGACT  
AAAGAACTGCAGAAACAGATCACTAAAATCCAGAACTTCCGTGTATACTACC  
GTGACTCTAGAGACCCGG  
TTTGGAAGGTCTGCTAAACTCCTGTGGAAGGGTGAAGGTGCTGTTGTTATC 30  
CAGGACAACCTCTGACAT  
CAAAGTGGTACCGCGTCGTAAAGCTAAAATCATTCGCGACTACGGCAAACAG  
ATGGCTGGTGACGACTGC  
GTTGCTAGCCGTCAGGACGAAGACTAAAAGCTTCAGGC

【 0 2 2 3 】

配列番号 7 1

名称 : g i | 4 3 8 6 1 0 | g b | A A C 3 7 8 7 5 . 1 | インテグラーゼ、部分的 [ ヒ 40  
ト免疫不全ウイルス 1 ]

配列 :

FLDGIDKAQEEHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQ  
VDCSPGIWQLDCTHLEG  
KVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTVHTDNGSNFTSTT  
VKAACWWAGIKQEFG  
IPYNPQSQGVIESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYS  
AGERIVDIIATDIQT  
KELQKQITKIQNFRVYYRDSRDPVWKGPALLWKGEAVVIQDNSDIKVVPRRK  
AKIIRDYGKQMAGDDC  
VASRQDED

10

【 0 2 2 4 】

配列番号 7 2

名称 : g i | 5 4 5 6 1 2 2 3 2 | r e f | W P \_ 0 2 1 7 3 6 7 2 2 . 1 | V 型 C R I  
S P R 関連タンパク質 C p f 1 [ アシダミノコッカス属 ( A c i d a m i n o c o c c u  
s s p . ) B V 3 L 6 ]

配列 :

20

MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNDHYKELKPIIDRI  
YKTYADQCLQLVQ  
LDWENLSAIDSYRKEKTEETRNLIEEQATYRNAIHDFIGRTDNLTDANKRH  
AEIYKGLFKAELFNG  
KVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYENRKNVFS AEDISTAIPHRIVQD  
NFPKFKENCHIFTR  
LITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQTQIDLYNQLLGGISRE  
AGTEKIKGLNEV  
LNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNTLSFILEEFKSDEEVIQSFCKYKTL  
LRNENVLETAE  
ALFNELNSIDLTHIFISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSAKE  
KVQRSLKHEDINL  
QEIIAAGKELSEAFKQKTSEILSHAHAALDQPLPTTLKKQEEKEILKSQLDSSLGL  
YHLLDWFAVDESN

30

40

EVDPEFSARLTGIKLEMEPSLSFYNKARNYATKKPYSVEKFKLNFQMPTLASGW  
DVNKEKNNGAILFVKN  
GLYYLGIMPKQKGRYKALSFEPTSEKTFDGMYYDYFPDAAKMIPKCSTQLKA  
VTAHFQTHHTPILLSN  
NFIEPLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKGYREALCKWIDFTRDFLS  
KYTKTTSIDLSSLRP  
SSQYKDLGEYYAELNPLLYHISFQRIAEKEIMDAVETGKLYLFQIYNKDFAKGHH  
GKPNLHTLYWTGLFS  
PENLAKTSIKLNGQAELFYRPKSRMKRMAHRLGEKMLNKKLKDQKTPIDTLYQ  
ELYDYVNHRLSHDLSD  
EARALLPNVITKEVSHEIHKDRRFTSDKFFFHVPITLNYQAANSPSKFNQRVNAYL  
KEHPETPIIGIDRG  
ERNLIYITVIDSTGKILEQRLNTIQQFDYQKKLDNREKERVAAARQAWSVVGTIK  
DLKQGYLSQVIHEIV  
DLMIHYQAVVLENLNFQFVSKRTGIAEKAVYQQFEKMLIDKLNLVLKDYPAE  
KVGGVLNPYQLTDQFT  
SFAKMGTQSGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHESRKHFLGFDLH  
YDVKTGDFILHFKMN  
RNLSFQRGLPGFMPAWDIVFEKNETQFQDAKGTPIAGKRIVPVIENTHRFTGRYRD  
LYPANELIALLEEK  
IVFRDGSNILPKLLENDSSHAIIDTMVALIRSVLQMRNSNAATGEDYINSPVRDLN  
GVCFDSRFQNPPEWPM  
DADANGAYHIALKGQLLNHLKESKDLKLQNGISNQDWLAYIQELRN

10

20

30

40

【 0 2 2 5 】

配列番号 7 3

名称 : g i | 7 6 9 1 4 2 3 2 2 | r e f | W P \_ 0 4 4 9 1 9 4 4 2 . 1 | V 型 C R I  
S P R 関連タンパク質 C p f 1 [ ラクノスピラ科 ( L a c h n o s p i r a c e a e ) 細  
菌 M A 2 0 2 0 ]

配列 :

MYYESLTKQYPVSKTIRNELIPIGKTLDNIRQNNILESDVVRKQNYEHVKGILDEY  
HKQLINEALDNCTL  
PSLKIAAEIYLKNQKEVSDREDFNKTQDLLRKEVVEKLKAHENFTKIGKKDILDL  
LEKLPSISEDDYNAL

ESFRNFYTYFTSYNKVRENLYSDKEKSSTVAYRLINENFPKFLDNVKS YR FVKTA  
GILADGLGEEEQDSL  
FIVETFNKTLTQDGDITYNSQVGKINSSINLYNQKNQKANGFRKIPKMKMLYKQI  
LSDREESFIDEFQSD  
EVLIDNVESYGSVLIESLKSSKVS AFFDALRESKGKNVYVKNDLAKTAMSNIVFE  
NWRTFDDLLNQEYDL  
ANENKKKDDKYFEKRQKELKKNKSYSLEHLCNLSEDCNLIENYIHQISDDIENIII  
NNETFLRIVINEH  
DRSRKLAKNRKAVKAIKDFLDSIKVLERELKLINSSGQELEKDLIVYSAHEELLVE  
LKQVDSL YNMTRNY  
LTKKPFSTEKVKLNFN RSTLLNGWDRNKETDNLGVLLLKD GKY YLGIMNTSAN  
KAFVNPPVAKTEKVFKK  
VDYKLLPVPNQMLPKVFFAKSNIDFYNPSS E IYSNYKKGTHKKGNMFSLEDCHN  
LIDFFKESISKHEDWS  
KFGFKFSDTASYNDISEFYREVEKQGYKLT YTDIDETYINDLIERNEL YLFQIYNK  
DFS MYSKGKLN LHT  
LYFMMLFDQRNIDDV VYKLNGEAEVFYRPASISEDELIH KAGEEIKNKNPNRAR  
TKETSTFSYDIVKDK  
RYSKDKFTLHIPITMNF G VDEVKRFND AVNSAIRIDENVNVIGIDRGERNLLYV  
VIDSKGNILEQISLN  
SIINKEYDIETDYHALLDEREGGRDKARKDWNTVENIRDLKAGYLSQV VNVVAK  
LVLKYNAIICLEDLNF  
GFKRGRQKVEKQVYQKFEKMLIDKLN YLVIDKSREQTSPKELGGALNALQLTSK  
FKSFKELGKQSGVIYY  
VPAYLTSKIDPTTG FANLFYMKCENVEKSKRFFDGFDFIRFNALENVFEFGFDYR  
SFTQRACGINSK WTV  
CTNGERIIKYRNPDKNNMFDEKVVVVTD E MKNLFEQYKIPYEDGRNVKDMIISN  
EEAEFYRRLYRLLQQT  
LQMRNSTSDGTRDYIISPVKNKREAYFNSELSDGSPK DADANGAYNIARKGLW  
VLEQIRQKSEGEKINL  
AMTNAEWLEYAQTHLL

10

20

30

40

50

【 0 2 2 6 】

配列番号 7 4

名称 : g i | 4 8 9 1 3 0 5 0 1 | r e f | W P \_ 0 0 3 0 4 0 2 8 9 . 1 | V 型 C R I  
S P R 関連タンパク質 C p f 1 [ フランシセラ・ツラレンシス ( F r a n c i s e l l a



t u l a r e n s i s ) ]

配列：

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDK  
YHQFFIEEILSSVC  
ISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFNQ  
NLIDAKKGQESDLIL  
WLKQSKDNGIELFKANS DITDIDEALEI IKSFKGWTTYFKGFHENRKNVYSSNDIP  
TSIIYRIVDDNLPK  
FLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDEVFEI  
ANFN NYLNQSGITK  
FNTIIGGK FVNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKS  
FVIDKLEDDSDVVT  
TMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDL SQQV  
FDDYSVIGTAVLEY  
ITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLAL EEFNKHRDIDKQCRFE  
EILANFAAIPMIFD  
EIAQNKDNL AQISIKYQNQGKKDLLQASAEDDVKA IKDLLDQTNNLLHKLKIFHI  
SQSEDKANILDKDEH  
FYLVFEECYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEP  
DNTAILFIKDDKYYL  
GVMNKKNNKIFDDKA IKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNP  
SEDILRIRNHSTHTKN  
GSPQKGYEKF EFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYRE  
VENQGYKLTFENIS  
ESYIDSVVNQGKLYLFQIYNKDFSAYS KGRPNLHTLYWKALFDERNLQDVVYKL  
NGEAELFYRKQSIPKK  
ITHPAKEA IANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKF  
NDEINLLLKEKAND

10

20

30

40

VHILSIDRGERHLAYYTLDVGKGNIHKQDTFNIIGNDRMKTNYHDKLAAIEKDRD  
SARKDWKKINNIKEM  
KEGYLSQVVHEIAKL VIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKL  
NYLVFKDNEFDKTGG  
VLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKS  
QEFFSKFDKICYNLD  
KGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKNHNWDTREVVYPTKEL  
EKLLKDYSIEYGHGEC  
IKAAICGESDKKFFAKLTSLVNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQ  
APKNMPQDADANGAY  
HIGLKGLMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN

10

## 【 0 2 2 7 】

配列番号 7 5

名称 : g i | 5 0 2 2 4 0 4 4 6 | r e f | W P \_ 0 1 2 7 3 9 6 4 7 . 1 | V 型 C R I  
S P R 関連タンパク質 C p f 1 [ [ ユーバクテリウム ] エリゲンス ( e l i g e n s ) ]  
配列 :

20

MNGNRSIVYREFVGVIPVAKTLRNELRPVGHTQEHIQNGLIQEDELREQSTELK  
NIMDDYYREYIDKS  
LSGVTDLDFTLLFELMNLVQSSPSKDNKKALEKEQSKMREQICTHLQSDSNYKNI  
FNAKLLKEILPDFIK  
NYNQYDVKDKAGKLETALFNGFSTYFTDFFEKRKNVFTKEAVSTSIAYRIVHE  
NSLIFLANMTSYKKIS  
EKALDEIEVIEKNNQDKMGDWELNQIFNPDFYNMVLIQSGIDFYNEICGVVNAH  
MNLYCQQTKNNYNLKF  
MRKLHKQILAYTSTSFEVPMFEDDMSVYNAVNAFIDETEKGNIIGKLKDIVNKY  
DELDEKRIYISKDFY  
ETLSCFMSGNWNLITGCVENFYDENIHAKGKSKEEKVKKAVKEDKYKSINDVND  
LVEKYIDEKERNEFKN  
SNAKQYIREISNIITDTETAHLEYDDHISLIESEEKADEMCKRLDMYMNMYHWA  
KAFIVDEVLDREMFY  
SDIDDIYNILENIVPLYNRVRNYVTQKPYNSSKKIKLNFQSPTLANGWSQSKEFDN  
NAIILIRDNKYYLAI

30

40

FNAKNKPKDKKIIQGNSDKKNDNDYKKMVYNLLPGANKMLPKVFLSKKGIETFK  
PSDYIISGYNAHKHIKT  
SENFDISFCRDLIDYFKNSIEKHAEWKYEKFSATDSYSDISEFYREVEMQGYRI  
DWTYISEADINKLD  
EEGKIYLFQIYNKDFAEINSTGKENLHTMYFKNIFSEENLKDIIIKLNQAEIFYRR  
ASVKNPVKHKKDSV  
LVNKTYKNQLDNGDVVRIPDDIYNEIYKMYNGYIKESDLSEAAKEYLDKVEV  
RTAQKDIVKDYRYTVD  
KYFIHTPITINYKVTARNNVNDMVVKYIAQNDDIHVIGIDRGERNLIYISVIDSHG  
NIVKQKSYNILLNNY  
DYKKKLVEKEKTREYARKNWKSIGNIKELKEGYISGVVHEIAMLIVEYNAILAME  
DLNYGFKRGRFKVER  
QVYQKFESMLINKLNYFASKEKSVDEPGGLLKGYQLTYVPDNIKNLGKQCGVIF  
YVPAAFTSKIDPSTGF  
ISAFNFKSISTNASRKQFFMQFDEIRYCAEKDMFSFGFDYNNFDTYNITMGKTQW  
TVYTNGERLQSEFNN  
ARRTGKTKSINLTETIKLLEDNEINYADGHDIRIDMEKMDDEKKSEFFAQLLSLY  
KLTVMRNSYTEAE  
EQENGISYDKIISPVINDEGEFFDSDNYKESDDKECKMPKDADANGAYCIALKGL  
YEVLLKIKSEWTEGDF  
DRNCLKLPHAEWLDFIQNKRYE

10

20

30

40

【 0 2 2 8 】

配列番号 7 6

名称 : g i | 5 3 7 8 3 4 6 8 3 | r e f | W P \_ 0 2 0 9 8 8 7 2 6 . 1 | V 型 C R I  
S P R 関 連 タ ン パ ク 質 C p f 1 [ レ プ ト ス ピ ラ ・ イ ナ ダ イ ( L e p t o s p i r a i n  
a d a i ) ]

配列 :

MEDYSGFVNIYSIQKTLRFELKPVGKTLEHIEKKGFLKKDKIRAEDYKAVKKIIDK  
YHRAYIEEVFDSVL  
HQKKKKDKTRFSTQFIKEIKEFSELYYKTEKNIPDKERLEALSEKLRKMLVGAFK  
GEFSEEVAEKYKNLF  
SKELIRNEIEKFCETDEERKQVSNFKSFTTYFTGFHSNRQNIYSDEKKSTAIGYRII  
HQNLPKFLDNLKI

IESIQRRFKDFPWSDLKKNLKKIDKNIKLTEYFSIDGFVNVLNQKGIDAYNTILGG  
KSEESGEKIQGLNE  
YINLYRQKNNIDRKNLPNVKILFKQILGDRETKSFIPEAFPDDQSVLNSITEFAKYL  
KLDKKKKKSIIAEL  
KKFLSSFNRYELDGIYLANDNSLASISTFLFDDWSFIKKSVSFKYDESVGDPKKKI  
KSPLKYEKEKEKWL  
KQKYTISFLNDAIESYSKSQDEKRVKIRLEAYFAEFKSKDDAKKQFDLLERIEEA  
YAIVEPLLGAEYPR  
DRNLKADKKEVGKIKDFLDSIKSLQFFLKPLL SAEIFDEKDLGFYNQLEGYYEEID  
SIGHLYNKVRNYLT  
GKIYSKEKFKLNFENSTLLKGWDENREVANLCVIFREDQKYYLGVMDKENNTIL  
SDIPKVKNELFYEKM  
VYKLIPTPHMQLPRIIFSSDNL SIYNPSKSILKIREAKSFKEGKNFKLKDCHKFIDFY  
KESISKNE DWSR  
FDFKFSKTSSYENISEFYREVERQGYNLDFKKVSKFYIDSLVEDGKLYLFQIYNKD  
FSIFSKGKPNLHTI  
YFRSLFSKENLKDVCLKLNGEAEMFFRKK SINYDEKKKREGHHPELFEKLKYPIL  
KDKRYSEDKFQFHLP  
ISLNFKSKERLNFNLKVNEFLKRNKDINIIGIDRGERNLLYLVMINQKGEILKQTLL  
DSMQSGKGRPEIN  
YKEKLQEKEIERDKARKSWGTVENIKELKEGYLSIVIHQISKLMVENNAIVVLED  
LNIGFKRGRQKVERQ  
VYQKFEKMLIDKLNFLVFKENKPTEPGGVLKAYQLTDEFQSFEKLSKQTGFLFY  
VPSWNTSKIDPRTGFI  
DFLHPAYENIEKAKQWINKFDSIRFNSKMDWF EFTADTRKFSENMLGKNRVW  
VICTTNVERYFTSKTAN  
SSIQYNSIQITEKLKELFVDIPFSNGQDLKPEILRKND AVFFKSLLFYIKTTLSLRQN  
NGKKGEEEEKDFI  
LSPVVDSKGRFFNSLEASDDEPKDADANGAYHIALKGLMNLLVLNETKEENLSR  
PKWKIKNKDWLEFVWE  
RNR

10

20

30

40

50

【 0 2 2 9 】

配列番号 77

名称 : g i | 7 3 9 0 0 8 5 4 9 | r e f | W P \_ 0 3 6 8 9 0 1 0 8 . 1 | V 型 C R I  
S P R 関 連 タ ン パ ク 質 C p f 1 [ ボ ル フ ィ ロ モ ナ ス ・ ク レ ビ オ リ カ ニ ス ( 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 ) ]

配列：

MDSLKDFTNLYPVSKTLRFELKPVGKTLENIEKAGILKEDEHRAESYRRVKKIIDT  
YHKVFIDSSLENMA  
KMGIENEIKAMLQSFCELYKKDHRTEGEDKALDKIRAVLRGLIVGAFTGVCGRR  
ENTVQNEKYESLFKEK  
LIKEILPDFVLSTEAESLPFSVEEATRSLKEFDSFTSYFAGFYENRKNIYSTKPQSTA  
IAYRLIHENLPK  
FIDNILVFQKIKEPIAKELEHIRADFSAGGYIKKDERLEDIFSLNYYIHVLSQAGIEK  
YNALIGKIVTEG  
DGEMKGLNEHINLYNQQRGREDRLPLFRPLYKQILSDREQLSYLPESFEKDEELL  
RALKEFYDHIAEDIL  
GRTQQLMTSISEYDLRSRIYVRNDSQLTDISKKMLGDWNAIYMARERAYDHEQAP  
KRITAKYERDRIKALK  
GEESISLANLNSCIAFLDNVRDCRVDTYLSTLGQKEGPHGLSNLVENVFASYHEA  
EQLLSFPYPEENLI  
QDKDNVVLIKNLLDNISDLQRFLKPLWGMGDEPKDERFYGGEYNYIRGALDQVI  
PLYNKVRNYLTRKPYS  
TRKVKLNFGNSQLLSGWDRNKEKDNSCVILRKGQNFYLAIMNNRHKRSFENKM  
LPEYKEGEPYFEKMDYK  
FLPDPNKMLPKVFLSKKGIEIYKPSPKLLEQYGHGTHKKGDTFSMDDLHELIDFF  
KHSIEAHEDWKQFGF  
KFSDTATYENVSSFYREVEDQGYKLSFRKVSESYVYSLIDQGKLYLFQIYNKDFS  
PCSKGTPNLHTLYWR  
MLFDERNLADVYKLDGKAEIFFREKSLKNDHPHTHPAGKPIKKKSRQKKGEESLF  
EYDLVKDRRYTMDKF  
QFHVPITMNFKCSAGSKVNDMVNAHIREAKDMHVIDRGERNLLYICVIDSRGT  
ILDQISLNTINDIDY

10

20

30

40

HDLLESRDKDRQQEHRNWQTIEGIKELKQGYLSQAVHRIAELMVAYKAVVALE  
DLNMGFKRGRQKVESSV  
YQQFEKQLIDKLNLYLVDKKKRPEDIGLLRAYQFTAPFKSFKEMGKQNGFLFYIP  
AWNTSNIDPTTGFTVN  
LFHVQYENVDKAKSFFQKFDSISYNPKKDWFEFAFDYKNFTKKAEGSRSMWILC  
THGSRIKNFRNSQKNG  
QWDSEEFALTEAFKSLFVRYEIDYTADLKTAIVDEKQKDFFDLLKLFKLTVMQ  
RNSWKEKDLDYLISPV  
AGADGRFFDTREGNKSPLKADANGAYNIALKGLWALRQIRQTSEGGKLLAIS  
NKEWLQFVQERSYEKD

10

【 0 2 3 0 】

配列番号 7 8

名称 : g i | 5 1 7 1 7 1 0 4 3 | r e f | W P \_ 0 1 8 3 5 9 8 6 1 . 1 | V 型 C R I  
S P R 関連タンパク質 C p f 1 [ プロフィロモナス・マカカ ( P o r p h y r o m o n a  
s m a c a c a e ) ]

20

配列 :

MKTQHFFEDFTSLYSLSKTIRFELKPIGKTLENIKKNGLIRRDEQRLDDYEKLKKV  
IDEYHEDFIANILS  
SFSFSEEILQSYIQNLSESEARAKIEKTMRDTLAKAFSEDERYKSIFKKELVKKDIP  
VWCPAYKSLCKKF  
DNFTTSLVPFHENRKNLYTSNEITASIPYRIVHVNLPKFIQNIEALCELQKKMGAD  
LYLEMENLRNVWP  
SFVKTPDDLCLNKTYNHLMVQSSISEYNRFVGGYSTEDGTKHQGINEWINIYRQR  
NKEMRLPGLVFLHKQ  
ILAKVDSSSFISDTLENDDQVFCVLRQFRKLFWNTVSSKEDDAASLKDLFCGLSG  
YDPEAIYVSDAHLAT  
ISKNIFDRWNYISDAIRRKTEVLMPRKKESVERYAEKISKQIKKRQSYSLAELDDL  
LAHYSEESLPAGFS  
LLSYFTSLGGQKYLVS DGEVILYEEGSNIWDEVLI AFRDLQVILDKDFTEKKLGK  
DEEAVSVIKKALDSA  
LRLRKFFDLLSGTGAEIRRDSSFYALYTDRMDKLKGLLKMVDKVRNYLTCKPYS  
IEKFKLHFDNPSLLSG

30

40

W D K N K E L N N L S V I F R Q N G Y Y Y L G I M T P K G K N L F K T L P K L G A E E M F Y E K M E Y K Q  
I A E P M L M L P K V F F P K K T  
K P A F A P D Q S V V D I Y N K K T F K T G Q K G F N K K D L Y R L I D F Y K E A L T V H E W K L F N F S F  
S P T E Q Y R N I G E F F D E V  
R E Q A Y K V S M V N V P A S Y I D E A V E N G K L Y L F Q I Y N K D F S P Y S K G I P N L H T L Y W K A L  
F S E Q N Q S R V Y K L C G G G  
E L F Y R K A S L H M Q D T T V H P K G I S I H K K N L N K K G E T S L F N Y D L V K D K R F T E D K F F F  
H V P I S I N Y K N K K I T N V  
N Q M V R D Y I A Q N D D L Q I I G I D R G E R N L L Y I S R I D T R G N L L E Q F S L N V I E S D K G D L R T  
D Y Q K I L G D R E Q E R L  
R R R Q E W K S I E S I K D L K D G Y M S Q V V H K I C N M V V E H K A I V V L E N L N L S F M K G R K K  
V E K S V Y E K F E R M L V D K L  
N Y L V V D K K N L S N E P G G L Y A A Y Q L T N P L F S F E E L H R Y P Q S G I L F F V D P W N T S L T D P  
S T G F V N L L G R I N Y T N  
V G D A R K F F D R F N A I R Y D G K G N I L F D L D L S R F D V R V E T Q R K L W T L T T F G S R I A K S K  
K S G K W M V E R I E N L S L  
C F L E L F E Q F N I G Y R V E K D L K K A I L S Q D R K E F Y V R L I Y L F N L M M Q I R N S D G E E D Y I L  
S P A L N E K N L Q F D S R  
L I E A K D L P V D A D A N G A Y N V A R K G L M V V Q R I K R G D H E S I H R I G R A Q W L R Y V Q E G I  
V E

10

20

30

【 0 2 3 1 】

配列番号 7 9

名称：U n i p r o t 部位で見られるインテグラーゼタンパク質配列。G e n B a n k から D N A 配列を得た。

配列：

TTTTTAGATGGAATAGATAAGGCCCAAGATGAACATGAGAAATATCACAGTA  
ATTGGAGAGCAATGGCTAGTGATTTTAACCTGCCACCTGTAGTAGCAAAAGA  
AATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGA  
CAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAG  
GAAAAGTTATCCTGGTAGCAGTTCATGTAGCCAGTGGATATATAGAAGCAGA  
AGTTATTCCAGCAGAAACAGGGCAGGAAACAGCATATTTTCTTTTAAAATTA  
GCAGGAAGATGGCCAGTAAAAACAATACATACTGACAATGGCAGCAATTTCA  
CCGGTGCTACGGTTAGGGCCGCCTGTTGGTGGGCGGGAATCAAGCAGGAATT  
TGGAATTCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAA  
GAATTAAGAAAATTATAGGACAGGTAAGAGATCAGGCTGAACATCTTAAG  
ACAGCAGTACAAATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGGGGA  
TTGGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACA  
TACAACTAAAGAATTACAAAAACAATTACAAAAATTCAAAATTTTCGGGT  
TTATTACAGGGACAGCAGAAATCCACTTTGGAAAGGACCAGCAAAGCTCCTC  
TGGAAAGGTGAAGGGGCAGTAGTAATACAAGATAATAGTGACATAAAAGTA  
GTGCCAAGAAGAAAAGCAAAGATCATTAGGGATTATGGAAAACAGATGGCA  
GGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAG

10

20

30

40

【 0 2 3 2 】

配列番号 8 0

名称 : s p | P 0 4 5 8 5 | 1 1 4 8 - 1 4 3 5

配列 :

FLDGIDKAQDEHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQ  
VDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGR  
WPVKTIHTDNGSNFTGATVRAACWWAGIKQEFIPYNPQSQGVVESMNKELKKI  
IGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQI  
TKIQNFRVYYRDSRNPLWKGPALLWKGEAVVIQDNSEIKVPPRRKAKIIRDY  
GKQMAGDDCVASRQDED

【 0 2 3 3 】

配列番号 8 1

ジंकフィンガータンパク質を特徴付けるタンパク質ドメイン

C X ( 2 - 4 ) C X ( 1 2 ) H X ( 3 - 5 ) H ( X ( 2 - 4 ) は、例えば X X または X X  
X または X X X X を意味する )

【 0 2 3 4 】

配列番号 8 2

> g i | 1 6 1 6 6 0 6 | e m b | X 9 7 0 4 4 . 1 | マウス乳癌ウイルス 5 ' L T R  
D N A



ATGCCGCGCCTGCAGCAGAAATGGTTGAACTCCCGAGAGTGTCTTACACTTA  
GGGGAGAAGCAGCCAAGG  
GGTTGTTTCCCACCCAGAACGACCCATCTGCGCACACACGGATGAGCCCGTC  
AAACAAAGACATATTCAT  
TCTCTGCTGCAAACCTTGGCATAGCTCTGCTTTGCCTGGGGCTATTGGGGGAAG  
TTGCGGTTCATGCTCGC

AGGGCTCTCACCTTGACTCTTTAATAGCTCTTCTGTGCAAGATTACAATCT  
AAACAATTTCGGAGAACT  
CGACCTTCCTCCTGAGGCAAGGACCACAGCCAACTTCCTCTTACAAGCCGCA  
TCGATTTAGTCCTTCAGA  
AATAGAAATAAGAATGCTTGCTAAAAATTATATTTTTTACCAATGAGACCAAT  
CCAATAGGTCGATTATTA  
ATTACTATGTTAAGAAATGAATCATTATCTTTTAGTACTATTTTTTACTCAAATT  
CAGAAGTTAGAAATGG  
GAATAGAAAATAGAAAGAGACGCTCAGCCTCAGTTGAAGAACAGGTGCAAG  
GACTAAGGGCCTCAGGCCT  
AGAAGTAAAAAGGGGGAAGAGGAGTGCGCTTGTCAAAATAGGAGACAGGTG  
GTGGCAACCAGGAACTTAT  
AGGGGACCTTACATCTACAGACCAACAGACGCCCCCTTACCGTATACAGGAA  
GATATGACCTAAATTTTG  
ATAGGTGGGTCACAGTCAATGGCTATAAAGTGTTATACAGATCCCTCCCCTTT  
CGTGAAAGGCTCGCCAG  
AGCTAGACCTCCTTGGTGCGTGTTGTCTCAGGAAGAAAAAGACGACATGAAA  
CAACAGGTACATGATTAT  
ATTTATCTAGGAACAGGAATGAACTTTTGGAGATATTATACCAAGGAGGGGG  
CAGTGGCTAGACTATTAG  
AACACATTTCTGCAGATACTAATAGCATGAGTTATTATGATTAGCCTTTATTG  
GCCCAATCTTGTGGTTC  
CCAGGGTTCAAGTAGGTTTCATGGTCACAACTGTTCTTAAAAACAAGGATGT  
GAGACAAGTGGTTTCCTG  
GCTTGGTTTGGTATCAAATGTTTTGATCTGAGCTCTGAGTGTTCTGTTTTCTTA  
TGTTCTTTTGGAATCT  
ATCCAAGTCTTATGTAAATGCTTATGTAAACCAAAGTATAAAAGAGTGCTGA  
TTTTTTGAGTAAACTTGC  
AACAGTCCTAACATTACCTCTCGTGTGTTTGTGTCTGTTTCGCCATCCCGTCTC  
CGCTCGTCACTTATCC  
TTCACCTTCCAGAGGGTCCCCCGCAGACCCCGGTGACCCTCAGGTTGGCCG  
ACTGCGGCA

10

20

30

40

50

【 0 2 3 5 】

配列番号 83

&gt; gi | 1403387 | emb | X98457.1 | マウス乳癌ウイルス3'LTR

ATGCCGCGCCTGCAGCAGAAATGGTTGAACTCCCGAGAGTGTCTTACACTTA  
GGAGAGAAGCAGCCAAGG  
GGTTGTTTCCCACCAAGGACGACCCGTCTGCGTGCACGCGGATGAGCCCATC  
AGACAAAGACATACTCAT  
TCTCTGCTGCAAACCTTGGCATAGCTCTGCTTTGCCTGGGGCTATTGGGGGAAG  
TTGCGGTTCGTGCTCGC  
AGGGCTCTCACCTTGATTCTTTTAATAACTCTTCTGTGCAAGATTACAATCT  
AAACGATTTCGGAGAACT  
CGACCTTCCTCCTGGGGCAAGGACCACAGCCAACCTTCCTCTTACAAGCCACA  
CCGACTTTGTCCTTCAGA  
AATAGAAATAAGAATGCTTGCTAAAAATTATATTTTTACCAATGAGACCAAT  
CCAATAGGTCGATTATTA  
ATCATGATGTTTAGAAATGAATCTTTGTCTTTTAGCACTATATTTACTCAAATT  
CAAAGGTTAGAAATGG  
GAATAGAAAATAGAAAGAGACGCTCAACCTCAGTTGAAGAACAGGTGCAAG  
GACTAAGGGCCTCAGGCCT  
AGAAGTAAAAAGGGGAAAGAGGAGTGCCTTGTCAAAATAGGAGACAGGTG  
GTGGCAACCAGGGACTTAT  
AGGGGACCTTACATCTACAGACCAACAGACGCCCCGCTACCATATACAGGAA  
GATACGATTTAAATTTTG  
ATAGGTGGGTCACAGTCAACGGCTATAAAGTGTTATACAGATCCCTCCCCCTT  
CGTGAAAGACTCGCCAG  
GGCTAGACCTCCTTGGTGTGTGTTAACTCAGGAAGAAAAAGACGACATGAAA  
CAACAGGTACATGATTAT  
ATTTATCTAGGAACAGGAATGAACTTCTGGGGAAAGATATTTGACTACACCG  
AAGAGGGAGCTATAGCAA

10

20

30

40

AAATTATATATAATATGAAATATACTCATGGGGGTCGCATTGGCTTCGATCCC  
TTTTGAAACATTTATAA  
ATACAATTAGGTCTACCTTGCGGTTCCCAAGGTTTAAGTAAGTTCAGGGTCAC  
AAACTGTTCTTAAAACA  
AGGATGTGAGACAAGTGGTTTCCTGACTTGGT

50

> gi | 119662099 | emb | AM076881.1 | ヒト免疫不全ウイルス 1  
プロウイルス 5'LTR、TARエレメントおよびU3、U5およびRリピート領域、ク  
ローン PG232.14

GGCAAGAAATCCTTGATTTGTGGGTCTACTACACACAAGGCTTCTTCCCTGAT  
TGGCAAAACTACACACC

GGGACCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTG  
CCAGTTGACCCAAAGGAA

10

GTAGAAGAGGCTAACCAAAGAGAAGACAACCTGTTTGCTACACCCTATGAGCC  
TGCATGGAATAGAGGACG

AAGACAGAGAAGTATTAAAGTGGCAGTTTGACAGCAGCCTAGCACGCAGAC  
ACATGGCCCCGCGAGCTACA

TCCAGAGTATTACAAAGACTGCTGACACAGAAAAGACTTTCCGCTAGGACTT  
TCCACTGAGGCGTTCCAG

GGGGAGTGGTCTAGGCAGGACTAGGAGTGGCCAACCCTCAGATGCTGCATAT  
AAGCAGCTGCTTTTCGCC

20

TGTACTAGGTCTCTCTAGGTGGACCAGATCTGAGCCTAGGCGCTCTCTGGCTA  
TCTAAGGAACCCACTGC

TTAAGCCTCAATAAAGCTTGCCTTGAGTGCTCTAAGTAGTGTGTGCCCCGTCTG  
TTGTGTGACTCTAGTAA

CTAGAGATCCCTCAGACCAACTTTAGTAGTGTAATAAATCTCTAGCAGTGGC  
GCCCCAACAGGGACCCGA

30

AAGTGAAAGCAGGACCAGAGGAGATCTCTCGACGCAGGACTCGGCTTGCTGA  
AAGTGCACTCGGCAAGAG

GCGAGAGCAGCGGCGACTGGTGAGTACGCCGAATTTTATTTTGACTAGCGGA  
GGCTAGAAGGAGAGAGAT

A

【0237】

配列番号 85

40

> gi | 1072081 | gb | U37267.1 | HIV 1 U37267 ヒト免疫不全  
ウイルス 1 型 3'LTR 領域

ATGGGTGGCAAGTGGTCAGAAAGTAGTGTGGTTAGAAGGCATGTACCTTTAA  
GACAAGGCAGCTATAGAT  
CTTAGCCGCTTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCAC  
AGAGAAGATCAGTTGAAC  
CAGAAGAAGATAGAAGAGGCCATGAAGAAGAAAACAACAGATTGTTCCGTT  
TGTTCCGTTGGGGACTTTC  
CAGGAGACGTGGCCTGAGTGATAAGCCGCTGGGGACTTTCGAAGAGGCGTG  
ACGGGACTTTCCAAGGCG  
ACGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGCTGCATATAA  
GCAGCTGCTTTCTGCCTG  
TACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAAC  
TAGGGAACCCACTGCTT  
AAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTT  
GTGTGACTCTGGTATCT  
AGA

10

20

【 0 2 3 8 】

配列番号 8 6 ~ 9 9 はない。

【 0 2 3 9 】

配列番号 1 0 0

細胞のゲノムへの neo の挿入のためのオリゴ ( 5 ' および 3 ' HIV LTR の全配列  
を使用 )

GACAAGACATCCTTGATTTGTGGGTCTATAACACACAAGGCTTCTTCCCTGAT  
TGGCAAACTACACACC  
GGGACCAGGGACCAGATACCCACTGACCTTTGGATGGTGCTTCAAGCTAGTG  
CCAGTTGACCCAAGGGAA  
GTAGAAGAGGCCAATACAGGGGAAAACAAGTGTGCTCCACCCTATGAGCC  
AGCATGGAATGGAAGATG  
ACCATAGAGAAGTATTAAAGTGGAAGTTTGACAGTATGCTAGCACGCAGACA  
CCTGGCCCGCGAGCTACA

30

40

TCCGGAGTACTACAAAACTGCTGACATGGAGGGACTTTCCGCTGGGACTTT  
CCATTGGGGCGTTCCAGG  
AGGTGTGGTCTGGGCGGGACAAGGGAGTGGTCAACCCTCAGATGCTGCATAT  
AAGCAGCTGCTTTTCGCT  
TGTACTGGGTCTCTTTAGGTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTA  
CCTGAGGAACCCACTGC  
TTAAGCCTCAATAAAGCTTGCCTTGAGTGCTCTAAGTAGTGTGTGCCCGTCTG  
TTGTGTGACTCTGGTAA  
CTAGAGATCCCTCAGACCCTTTTGGTAGTGTGGAAAATCTCTAGCAGATGATT  
 GAACAAGATGGATTGCAC  
 GCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCAC  
 AACATGGGTGGCAAGTGGTCAG  
**AAAGTAGTGTGGTTAGAAGGCATGTACCTTTAAGACAAGGCAGCTATAG**  
**ATCTTAGCCGCTTTTTAAAAGAAAAG**  
**GGGGGACTGGAAGGGCTAATTCACCTCACAGAGAAGATCAGTTGAACCAG**  
**AAGAAGATAGAAGAGGCCATGAAG**  
**AAGAAAACAACAGATTGTTCCGTTTGTTCGGTTGGGGACTTTCCAGGAG**  
**ACGTGGCCTGAGTGATAAGCCGCTGGG**  
**GACTTTCCGAAGAGGCGTGACGGGACTTTCCAAGGCGACGTGGCCTGGG**  
**CGGGACTGGGGAGTGGCGAGCCCTC**  
**AGATGCTGCATATAAGCAGCTGCTTTCTGCCTGTACTGGGTCTCTCTGGT**  
**TAGACCAGATCTGAGCCTGGGAGCTCT**  
**CTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTG**  
**AGTGCTTCAAGTAGTGTGTGCCCGTCTG**  
**TTGTGTGACTCTGGTATCTAGA**

10

20

30

40

最初の 5'LTR に下線を付し、標準文字は neo であり、3'LTR は太字 (1179bp) とする。

【0240】

配列番号 101

(224bp) 内に neo 配列がある 5'LTR および 3'LTR の短縮型

最初の 5'LTR に下線を付し、標準文字は neo であり、3'LTR は太字とする。

GACAAGACATCCTTGATTTGTGGGTCTATAACACACAAGGCTTCTTCCCTGAT  
TGGCAAAACTACACACCATGATTGAACAAGATGGATTGCAC  
GCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCAC  
AACTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCC  
CGTCTG  
TTGTGTGACTCTGGTATCTAGA

10

## 【0241】

配列番号72に関して

Genbankタンパク質ID: WP\_021736722.1

NRデータベースまたはローカルGIからのNCBIタンパク質GI (WGSデータベース由来のタンパク質に対する): 545612232

WGSデータベースにおけるコンティグID: AWUR01000016.1

コンティグの記述: アシダミノコッカス属 (*Acidaminococcus* sp.)

BV3L6 コンティグ00028、全ゲノムショットガン配列

タンパク質完全性: 完全

実験的に分析したタンパク質: 8

20

非重複セット: nr

生物: アシダミノコッカス属 (*Acidaminococcus* sp.) \_ BV3L6  
分類:細菌、ファーミキューテス (*Firmicutes*)、ネガティブキューテス (*Negativicutes*)、セレンオモナダレス (*Selenomonadales*)、アシダ  
ミノコッカス科 (*Acidaminococcaceae*)、アシダミノコッカス (*Acidaminococcus*)、アシダミノコッカス属 (*Acidaminococcus*  
sp.) BV3L6

## 【0242】

配列番号73に関して

Genbankタンパク質ID: WP\_044919442.1

NRデータベースまたはローカルGIからのNCBIタンパク質 (WGSデータベース由来のタンパク質に対する): 769142322

WGSデータベースにおけるコンティグID: JQKK01000008.1

コンティグの記述: ラクノスピラ科 (*Lachnospiraceae*) 細菌 MA2020T348 DRAFT\_scaffold00007.7\_C、全ゲノムショットガン  
配列

タンパク質完全性: 完全

実験的に分析したタンパク質: 9

40

非重複セット: nr

生物: ラクノスピラ科 (*Lachnospiraceae*) \_ 細菌 \_ MA2020分類: 細菌、ファーミキューテス (*Firmicutes*)、クロストリジア (*Clostridia*)、クロストリジiales (*Clostridiales*)、ラクノスピラ科  
(*Lachnospiraceae*)、未分類ラクノスピラ科 (*Lachnospiraceae*)、ラクノスピラ科 (*Lachnospiraceae*) 細菌 MA2020

## 【0243】

開示される組成物および方法で使用され得るさらなる核酸配列およびタンパク質配列 -  
CPF1 アライメント。配列番号86~92; チャートの上から下の順序。

## CLUSTAL 0(1.2.1)複数配列アライメント

```

g1|545612232|ref|WP_021736722.1|-----MTQFEGFTNLVQVSKTLRFELIPQGKTLKHIQEDGFIEEDKAPNDHYKELKPIIG
g1|517171843|ref|WP_018359861.1|---MKTQHFEDFTSLVSLKTIKRFELKPIQKTLKHIKKNGLIKRDEQRLDQVEKLLKKVIG
g1|502240446|ref|WP_012739547.1|MNQNRISIVYREFVGVIPYAKTLRNLKRPVQHTQEHIIQNGLIQEDQLRQKSTELKKNIMD
g1|537834683|ref|WP_020988726.1|-----MEDYSGFVNIYSTOKTLRFELKPVGKTLKHIKKNGLIKRDEQRLDQVEKLLKKVIG
g1|769142322|ref|WP_044919442.1|-----MYYESLTQDYPVSKTIRNELIPIGKTLDNIRQNNILESGVKKKQNYEHVKGILD
g1|489130501|ref|WP_003040289.1|-----MSIYGEFVNNYSLSKTLRFELIPQGKTLKHIKKNGLIKRDEQRLDQVEKLLKKVIG
g1|739088549|ref|WP_036898108.1|-----MGSILKDFTNLVPVSKTLRFELKPVGKTLKHIKKNGLIKRDEQRLDQVEKLLKKVIG
                                     ( ) ** * * * * * * * * * * * * * * * * * * * *

g1|545612232|ref|WP_021736722.1|RIYKTYADGCLQLVQLDWNEL-----GAAIDSYAKE-----KTEETRNALIEEQ
g1|517171843|ref|WP_018359861.1|EYHEDFIANILSSFSFSEEL-----GSYIION-----LGE---GEARAKIE
g1|502240446|ref|WP_012739547.1|DYVREYIDKLSGGYTDLOFTL-----L-----FELMNLVUSSPCKDKKALEKEQ
g1|537834683|ref|WP_020988726.1|KYHRAVIEEVFDSVLHQKKKKDKTRFSTQFIKEIKFESLYYTERNIPQK---ERLEALS
g1|769142322|ref|WP_044919442.1|EYHKQLINEALDKTLPGLKI-----A-----AEIYLNKQEVSD---REDPNKTO
g1|489130501|ref|WP_003040289.1|KYHDFFIEELISSVCTIS-----EDLLQNYSDVYFKLKKSUDDNLDKDFKSAK
g1|739088549|ref|WP_036898108.1|TYHKVFIDSGLENMAKMGIEH-----EIKAMLOSFCELYKKDHRTEGEOKA---LDKIR
                                     ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

g1|545612232|ref|WP_021736722.1|ATYRNAIHDFYIGRTDNLTDAINKRHAEIYKGLFKAELEFNGKVLK-----
g1|517171843|ref|WP_018359861.1|KTHDDTLAKAF-----GEDERYKSIKFKELVKKDI-----PVWCP-----
g1|502240446|ref|WP_012739547.1|SKMPEQICTHL-----QSDSAYKNIIFNAKLKEIL---PDFIKWYNG-----
g1|537834683|ref|WP_020988726.1|EKLKMLVGFAPKGEFS---E---EVAEKYKMLFSELEIRNEIE-----
g1|769142322|ref|WP_044919442.1|DLKKEVVEKL-----KAHNFTHIGKDDILD-----
g1|489130501|ref|WP_003040289.1|DTIKKQI-----SEYINDSEKFKMLFNQWIDAKKQGFSDILWLKQSKONG
g1|739088549|ref|WP_036898108.1|AVLGLVGAFTQVCG---PRENTVQNEKYESLFKEKLKEIL---PDFVL-----
                                     ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

g1|545612232|ref|WP_021736722.1|-----QLGTVYTTHEENALLRDFDKFTTYPFSGFYENRKNVFSAEIDISTAIPIHRIVQDNFP
g1|517171843|ref|WP_018359861.1|-----AYKSLCKKFNFTTSLVPFHENAKNLYTSNEITASIPYRIVHVMPL
g1|502240446|ref|WP_012739547.1|-----YQVQKQKQKLETIALFNGFSTYFTDFFEKKNVFTKEAVSTGIAYRIVHENS
g1|537834683|ref|WP_020988726.1|-----KPLETEGEERKQVSNFESFTYFTGPHSNQNNIYSDKKSTAIYRIHVMPL
g1|769142322|ref|WP_044919442.1|-----LLEKLPDISDDYNALESFRNFYFTYFTSYNNKVRNLYSDKKSTAYRLINENFP
g1|489130501|ref|WP_003040289.1|IELFKANGQITDIDEALEIISFKGWTYTFKGFHENRKNVYSSNDQIPSTIYRIVQDNLP
g1|739088549|ref|WP_036898108.1|---GTEAESLPFGVEEATSLKEFGSTYFAGFYENRKNIYSTKPGSTAIAYRLIHENLP
                                     * ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

g1|545612232|ref|WP_021736722.1|KFKENCHIFTALITAVPSLAEHFENVKKA-----IGIFVSTSTEEVFSGFPF
g1|517171843|ref|WP_018359861.1|KFIQNIKALCLQKKMGADL-YLEMMENL-R-----NYWPSFVKTPODLNKLK
g1|502240446|ref|WP_012739547.1|IFLANMISYKKISSEKALDEI---EVLKKN-----NQKMGQWELKQIFNPDP
g1|537834683|ref|WP_020988726.1|KFLDNLIKIESIQRRFKDF---PWSDLKKN-----LKKIDKNILTEYFSTIG
g1|769142322|ref|WP_044919442.1|KFLDNVKSRYFVKTAGIAD-GL-----GEEQDSLFIVET
g1|489130501|ref|WP_003040289.1|KFLDNKAKYESLKDKAPEAI-NYEGIKKDLAEELTFDIDYKTSFVNQVPSGLDEVFETAN
g1|739088549|ref|WP_036898108.1|KFIQNLIVFQIKIEPIAK---ELEHIRAD-----FSAGGYIKKDERLEIFSLNY
                                     * * * * *

g1|545612232|ref|WP_021736722.1|YNQLLTDTQIDLYNOLLGGISAEAGTEKIKGLNEVLNLAITQKNDYAHIIASLPHRFIPL
g1|517171843|ref|WP_018359861.1|YNHLMVQSSISEYNRKFGCYSTED-GTRHGGTNEWINIYRQRN-----KEM---RLPLGVFL
g1|502240446|ref|WP_012739547.1|YNMVLIGSSIDFYNEICGVV-----NAHNNLYCQGTK---NNY---NLFKMKKL
g1|537834683|ref|WP_020988726.1|FNVVLNKGSDIAYNTILGGKSEES-GEKIDGLNEYINLYRQKN---NIQK---NLPMVKIL
g1|769142322|ref|WP_044919442.1|FNKTLTQGGIDTYNSQVGI-----NSSINLYNKKKQKANGFR---KIPKMMKL
g1|489130501|ref|WP_003040289.1|FNVVLNKGSDITKFNITIGGKPYNGENTKRRKGINEYINLYSQGI---NOKTL---KKYKMSVL
g1|739088549|ref|WP_036898108.1|YIMVLSAQGIKYNALIGKIYTES-OGEMKGLNEHIMLYNQRK---GREOR---LPLFQPL
                                     ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

g1|545612232|ref|WP_021736722.1|FKQILSDQNTLSFILSEFKSGEEVDSFCYKTLILN-----ENVLETAELFNE---LN
g1|517171843|ref|WP_018359861.1|HKQILAKVDSSSFISDTLENDQVFCVLRQFRKLFNNTVSSK-EDDAASLKDLFCG---LS
g1|502240446|ref|WP_012739547.1|HKQILAYTSTSFEPKMFEDQMSVYNVAFIDETEK-----GNIIGKLDIWN---KYD
g1|537834683|ref|WP_020988726.1|FKQILGDRKTSFIPEAFDDQGVNIGITEPAKYLLDKK---KKQITAELEKFLSS---FN
g1|769142322|ref|WP_044919442.1|YKQILSDREES---FIDEFQSDDEVLDNVEYSGSVLIESLK-----SSKVSAPFBAIS---
g1|489130501|ref|WP_003040289.1|FKQILGDRKTSFVIDKLEDDSGVVTMDSFYEQIAAFKTVEKGINETSLFLFDDLKQ
g1|739088549|ref|WP_036898108.1|YKQILSDREELGLYPSFSEKDEELRALKEFYDHAIEDILQRTQQLM-----TSIS
                                     Y**** ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( ) ( )

```

10

20

30



1. *Staphylococcus aureus* 2. *Staphylococcus epidermidis* 3. *Staphylococcus saprophyticus* 4. *Staphylococcus carnosus* 5. *Staphylococcus sciuri* 6. *Staphylococcus hyicus* 7. *Staphylococcus epidermidis* 8. *Staphylococcus aureus* 9. *Staphylococcus aureus* 10. *Staphylococcus aureus* 11. *Staphylococcus aureus* 12. *Staphylococcus aureus* 13. *Staphylococcus aureus* 14. *Staphylococcus aureus* 15. *Staphylococcus aureus* 16. *Staphylococcus aureus* 17. *Staphylococcus aureus* 18. *Staphylococcus aureus* 19. *Staphylococcus aureus* 20. *Staphylococcus aureus* 21. *Staphylococcus aureus* 22. *Staphylococcus aureus* 23. *Staphylococcus aureus* 24. *Staphylococcus aureus* 25. *Staphylococcus aureus* 26. *Staphylococcus aureus* 27. *Staphylococcus aureus* 28. *Staphylococcus aureus* 29. *Staphylococcus aureus* 30. *Staphylococcus aureus* 31. *Staphylococcus aureus* 32. *Staphylococcus aureus* 33. *Staphylococcus aureus* 34. *Staphylococcus aureus* 35. *Staphylococcus aureus* 36. *Staphylococcus aureus* 37. *Staphylococcus aureus* 38. *Staphylococcus aureus* 39. *Staphylococcus aureus* 40. *Staphylococcus aureus* 41. *Staphylococcus aureus* 42. *Staphylococcus aureus* 43. *Staphylococcus aureus* 44. *Staphylococcus aureus* 45. *Staphylococcus aureus* 46. *Staphylococcus aureus* 47. *Staphylococcus aureus* 48. *Staphylococcus aureus* 49. *Staphylococcus aureus* 50. *Staphylococcus aureus* 51. *Staphylococcus aureus* 52. *Staphylococcus aureus* 53. *Staphylococcus aureus* 54. *Staphylococcus aureus* 55. *Staphylococcus aureus* 56. *Staphylococcus aureus* 57. *Staphylococcus aureus* 58. *Staphylococcus aureus* 59. *Staphylococcus aureus* 60. *Staphylococcus aureus* 61. *Staphylococcus aureus* 62. *Staphylococcus aureus* 63. *Staphylococcus aureus* 64. *Staphylococcus aureus* 65. *Staphylococcus aureus* 66. *Staphylococcus aureus* 67. *Staphylococcus aureus* 68. *Staphylococcus aureus* 69. *Staphylococcus aureus* 70. *Staphylococcus aureus* 71. *Staphylococcus aureus* 72. *Staphylococcus aureus* 73. *Staphylococcus aureus* 74. *Staphylococcus aureus* 75. *Staphylococcus aureus* 76. *Staphylococcus aureus* 77. *Staphylococcus aureus* 78. *Staphylococcus aureus* 79. *Staphylococcus aureus* 80. *Staphylococcus aureus* 81. *Staphylococcus aureus* 82. *Staphylococcus aureus* 83. *Staphylococcus aureus* 84. *Staphylococcus aureus* 85. *Staphylococcus aureus* 86. *Staphylococcus aureus* 87. *Staphylococcus aureus* 88. *Staphylococcus aureus* 89. *Staphylococcus aureus* 90. *Staphylococcus aureus* 91. *Staphylococcus aureus* 92. *Staphylococcus aureus* 93. *Staphylococcus aureus* 94. *Staphylococcus aureus* 95. *Staphylococcus aureus* 96. *Staphylococcus aureus* 97. *Staphylococcus aureus* 98. *Staphylococcus aureus* 99. *Staphylococcus aureus* 100. *Staphylococcus aureus*

⋮ ⋮ ⋮

1. 2. 3. 4. 5.

\* \* \* : \*\*\* \*\* \*

*Journal of Management Inquiry* 18(6) 709–724







1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840.

30

|    |           |     |                |
|----|-----------|-----|----------------|
| g1 | 545612233 | ref | WP_021736732.1 |
| g1 | 517171043 | ref | WP_018359861.1 |
| g1 | 502240446 | ref | WP_012739647.1 |
| g1 | 537934633 | ref | WP_020908726.1 |
| g1 | 769132322 | ref | WP_044919442.1 |
| g1 | 480130551 | ref | WP_003040289.1 |
| g1 | 738065549 | ref | WP_036090183.1 |

NPLLYHISFQRIAEKIMGAVETGKLYLFQIYNKDFAXGHNKPNLHTLYWTLGFSPENL  
 REQAYKVMWNPYASYIDEAVENGKLYLFQIYNKDFAXGHSYKPNLHTLYWTLGFSPENL  
 EMQGYRIVTNYTSEADINKLEEKLYLFQIYNKDFAXGHSYKPNLHTLYWTLGFSPENL  
 ERQGYNLBFKKYGFYIDSLVEDGKLYLFQIYNKDFGIFSKGPNLHTLYWTLGFSPENL  
 EKQGYKLYTIDIBETIYINDLIERNELYLFQIYNKDFSMYSKGNLHTLYWTLGFSPENL  
 ENQGYKLYTENTISEGYIDGYNKGLYLFQIYNKDFSAYSKGNLHTLYWTLGFSPENL  
 EBGYSKLYFKNVSEGYVYGLIDGKLYLFQIYNKDFSPCKGTPNLHTLYWTLGFSPENL  
 \* \* \* \* \*

```
g1 545512232 ref WP_021736722.1
g1 517171843 ref WP_018353861.1
g1 582248446 ref WP_012738647.1
g1 537834563 ref WP_020987725.1
g1 769142322 ref WP_044915442.1
g1 488131858 ref WP_083042829.1
g1 739888548 ref WP_035898188.1
```

```

NETSIFNLNGDAELFYRPPSMMWR---MAHRLGKKHKKLKK-----DQKTFIPDTLVQE
S-RVYFLKGGGELFYRAKSLNQDTTVHPHGISLHKKN
KDIITKLNGGAEFLPYRASGVNPNVK---HKKSGLVNTKYTKQLDNGGVYRIPIDDIYNE
KDVLLKLNGEAEENFFPKKSLINYEKKK-----R
DDVYVKLNGEAEVFPYRPSISEDLIIHAGEEIKNHNF
DDVYVKLNGEAEFLFYRKSLPK-K-IHPAKEALANKN
ADVIVKLNGRAEIFFRERSLKNDH-PTHPSGKPKKKK5

```

```

01 545612232 ref WP_021738722.1
01 517171833 ref WP_013359861.1
01 503240446 ref WP_012735647.1
01 537534633 ref WP_020988726.1
01 769134332 ref WP_044919442.1
01 489139501 ref WP_003040269.1
01 739885549 ref WP_035888168.1

```

```

LYDYVNHRLS-HDLSDEARALLPNVITREYSHEIJKDRFTSGKFFHVHPITLNYDAANL
      LMKRGETSFLNYDLVDKDRFTEDKFFHVHPITINYNKK-K-
IYKMYNGYIKESQLSEAAKEYLDKVEVRTAQGKIVGEYRYTVDPKFFIHPITINVKYV-A-
      EGHHPFLFEKLHPYPLDKGRYSDEKDFHFLPISLNFKSK-E-
NRARTKETSTFSYDLYDKDKRYSKDKFTLHPIPTINFGVD-E-
      KDMPKKESVFEYDLIKDKRFTEDKFFHHPITINFSS-G-
RQKRGESLFEYDLVDKARYTNMGKFFHVHPITMFKCS-A-
      ( ( ( * ) ) ) ( * ) ( * ) ( * )

```

```
g1 545512232 ref wp_021736722.1
g1 531717843 ref wp_018353862.1
g1 582240446 ref wp_012733647.1
g1 537834632 ref wp_020908725.1
g1 765147322 ref wp_044915447.1
g1 465130551 ref wp_083004289.1
g1 739088549 ref wp_035089108.1
```

```

PSKFNDRNAYLYK--PWPEPTIIGIORGERMNLIVITVDSIGKILGORSNLTIG-----Q
ITNNQNMVYKDYIA--QWODLVTIGIORGERMNLIVISATOTGNLQFQSLNVSFQDKGLA
RNNVNMVYKDYIA--QNDSTHTVIGIDAGERMNLIVTSVDSIGKILGORSNLTIG-----N
RLNPLNKLWYFLEK--RNNQDNLTIGIDAGERMNLIVYVMSNRGKILQTLLOSQSGKGRPE
WKRPMAYNSAIR--TDENVVIGIDAGERMNLIVYVDSIGKILGORSNLTIG-----Q
ANKFNDTELNLKKEANDVHLSIGORGERMNLIVYTTVDGKONIIGODFTWICNDR--NK
GSKYNDMNAHIR--EAKDMHVIGIORGERMNLIVCVDSIGTILGUISLNTIN-----Q

```

```

91 545612232 ref WP_021736722.1
91 517171843 ref WP_018359061.1
91 502240446 ref WP_012739547.1
91 537934533 ref WP_020988726.1
91 769124322 ref WP_044921944.1
91 489138501 ref WP_0303640269.1
91 739088549 ref WP_036898188.1

```

[illegible]

```
g1 545812233 ref WP_02136722.1
g1 517171043 ref WP_018558861.1
g1 582248446 ref WP_012739647.1
g1 537834633 ref WP_030986726.1
g1 766143322 ref WP_044915447.1
g1 488130501 ref WP_063048289.1
g1 738088546 ref WP_035808188.1
```

```

GFKSKRTGTAEKAVYQGFQFKMLIDKLNCLVLYKDY-----PAEKVGGVLMNPYGLTDGQFTSFA
SPFMKGRKK-VEKSVYKPEFNMVLVDKLNLYVDDKK-----LSNPEGLLYRAYGLTNPLFSFE
GFKKGRGFK-VERQVYQGFQFKMLIDKLNLYVFAKRE-----SYDEPGLLKYGLTYTYPDNTK
GFKKGRGFK-VERQVYQGFQFKMLIDKLNFLVFKEN-----KPTPEGGVLYKAYGLTDGQFSFE
GFKKGRGFK-VEKOVYQGFQFKMLIDKLNLYVDKSREGTSPEKLGGALNALDLSKFSKFK
GFKKGRGFK-VEKOVYQGLKFKMLIDKLNLYVFKKM-----EFTQGLLYRAYGLTAPPFTFK
GFKKGRGFK-VESSVYQGFQFKMLIDKLNLYVDDKK-----RPEIDIGLLRAYGFTAPPFKSF

```

```
g1 545612232 ref WP_0217366722.1
g1 517171843 ref WP_0123596861.1
g1 502240446 ref WP_012739547.1
g1 537834633 ref WP_020988726.1
g1 769142332 ref WP_044919442.1
g1 4869138551 ref WP_063046289.1
g1 739080540 ref WP_035896108.1
```

[illegible]

10

20

30

```

g1[545612232] ref[WP_021736722.1] TGRYP-GLYPANELIALLEENKIVFRGGSNILPKLL---ENDGSHAIOTMVALINQVLQM
g1[517171043] ref[WP_018359861.1] KWMVERIENLSLCFLLEFQFNIGYVVERDLKKAIL---SDORKEFYVRLIYLFNLMQI
g1[502240446] ref[WP_012739647.1] TGKTK-SINLTETIKLLLEONEINYADGHDIRIDMEKMDDEKKEFFAQLLSLYKLTVM
g1[537834683] ref[WP_020988726.1] SIQYN-SIGITEKLKELFVD--IPFNGGDLKPEIL---RKNDVFFKSLFYINTTLLS
g1[769142322] ref[WP_044919442.1] MFDEK-VVVVTDENKMLFEQYKIPYEDGRNVKDMII---SNEEAEFFYRLYLLQDTLM
g1[489130501] ref[WP_003040289.1] NWGTA-EVYPTKELEKLLKDYSTIEYGHGECIKAAIC---GESGKKFFAKLTSLNTILQM
g1[739000549] ref[WP_036890108.1] QWQSE-EFALTEAFKSLFVRYEIDUYTA--DLKTAIV---DEKQKDFVDDLKLFKLTVM
      ,      ,      ,      ,      ,      ,      ,      ,      ,      ,      ,
g1[545612232] ref[WP_021736722.1] RNSNAA-----TGEDYINSPVRLNGVCFDSRF-----QNPWPMDADANGAYHIAL
g1[517171043] ref[WP_018359861.1] RNS-----DGEEDYILSPALNEKNLQFDSKLI-----EAKDLFVDADANGAYVVAQ
g1[502240446] ref[WP_012739647.1] RNSYTEAEFEQENISGFDKIISPVINDEGEFFDSQNFKEGSDKEENMPKADANGAYFIAL
g1[537834683] ref[WP_020988726.1] RQNGKKKQ-----EECKDFILSPVVDKGRFFNSLE-----ASDEPKDADANGAYHIAL
g1[769142322] ref[WP_044919442.1] RNS-----TS-----DSTRDYIISPVKNKREAYFNGEL-----SDGCVPKDADANGAYNIAR
g1[489130501] ref[WP_003040289.1] RNS-----KT-----GTELDYILSPVAGVNGNFFDSRQ-----APKNMPQDADANGAYHISL
g1[739000549] ref[WP_036890108.1] RNS-----WK-----EKDLDYILSPVAGADGRFFDTRE-----QNKSLPKDADANGAYNIAL
      **      ,      ,      ,      ,      ,      ,      ,      ,      ,      ,
g1[545612232] ref[WP_021736722.1] KGLLLINHLKESKD-----LKLNGCISNGDWLAYIQELAN-----
g1[517171043] ref[WP_018359861.1] KGLMYVQRKIKGDH-----ESIMRISRAQWLRYVQEGIVE---
g1[502240446] ref[WP_012739647.1] KGLYEVLKIKSEWTEGDFDRLCLKLPHAEWLDFTQNKRYE---
g1[537834683] ref[WP_020988726.1] KGLMYLLVILNET--KEENLGGPKWKIKNKDWLEFVWENNR-----
g1[769142322] ref[WP_044919442.1] KGLWVLEDIAGK-SEG--EKINLANINAEWLEYAQTHLL---
g1[489130501] ref[WP_003040289.1] KGLMLLGRIMNN-SEG--KKLNLVINNEEYFEPVQNRNN---
g1[739000549] ref[WP_036890108.1] KGLWALRQIDQT-SEG--SKLKLAIISKEWLQFVQERSYEKD
      **      ,      ,      ,      ,      ,      ,      ,      ,      ,      ,

```

10

20

## 【 0 2 4 4 】

開示される組成物および方法において使用され得るさらなる核酸配列およびタンパク質配列 - C f p 1 ヒト切断タンパク質アライメント。配列番号 86 (第 1 列) および配列番号 90 (第 2 列)。

## CLUSTAL 0(1.2.1)複数配列アライメント

```
gi|545612232|ref|WP_021736722.1|NTQFEGETNLVGVAKTLKRFELIPQQKTLLKHIGEQSGTIEEDKAANDHYKELPFIJRIYKT  
gi|769142322|ref|WP_044919442.1|-MYEESLTKGYVPVKSTIRNELIPIGKTLENTIKRNILSDCVKKKNQNYEHVGKGLDEYNHG  
  
[...]  
gi|545612232|ref|WP_021736722.1|YAGDCLQLVOLUWENLSAAIDSYRKECTET--RNALEEODATYRNAINDYFRTDNLT  
gi|769142322|ref|WP_044919442.1|LINEALNCTLPLSLKI---AAEIYLNRMEYSREDFNKTUDLRKEVEYEKL-----  
  
[...]  
gi|545612232|ref|WP_021736722.1|AINKRMAEITYKSLFKAEFLNGKVRLHGLGT-VTTTEHENALLRSFOKFITTYFSOFYENAKN  
gi|769142322|ref|WP_044919442.1|----AH-ENFTISYCK-----KQTULULEKLPSISEDDYNALESFNFFYTTFYSNNHVREN  
  
[...]  
gi|545612232|ref|WP_021736722.1|VFSAEIDISTAIPIHRIVGNHPPKFLINNKCHIFPLIIAVPDLREMFENVKKAICIFXSFSE  
gi|769142322|ref|WP_044919442.1|LYSDEKKSSTVAVALRIENFPKFLINNKCHVFPHITAGI-LADG-----L-----GESEE  
  
[...]  
gi|545612232|ref|WP_021736722.1|EVFSPPFYNGILLTGTDIDLNGLLGGISRAGESTEKIKGLENEVLNALIQNDSETAMHTIAGL  
gi|769142322|ref|WP_044919442.1|SLFIVETFNKTITQGDIPTYNSGVKINGGIN-----LYNQKNQKAN-GFKKI  
  
[...]  
gi|545612232|ref|WP_021736722.1|PHRFIPLFKQILSRNLTSLFILEEPKSDSEVIQSFCRYKTLRWNVLETAEALFNELNS  
gi|769142322|ref|WP_044919442.1|P-MKMMLYKQILSDREE---SFIDEFQSDSVLIQNVGSVLTIESLKSSKYGAFFDALRE  
  
[...]  
gi|545612232|ref|WP_021736722.1|IDLTHIFIGHKK-LETISSALCDHWDTLRNALYERHSEL-TGRITKSAREVQRSKHE  
gi|769142322|ref|WP_044919442.1|SNQKNVVVNOLDAKTAMSNIYPENWTAFUDLNQEVDLANEKHXDKYFERQKELNM  
  
[...]  
gi|545612232|ref|WP_021736722.1|-DINLGEEI--SAASDELSENFQKQTS-----ILSHAMAALSDPL-----PTLC-KKBE  
gi|769142322|ref|WP_044919442.1|KSYSLEHLNLSSEDSCNLENIMYMIGISIDIMTIINNTRFLRVINEHQHSRKLAQNRA  
  
[...]  
gi|545612232|ref|WP_021736722.1|HELKKGOLDSLGLYLHLDWFANDESNEVD--PEFSARLTGIKLEMFGLSFYNKARNYA  
gi|769142322|ref|WP_044919442.1|VKA TKGFLODSIKVLERELKLIN-SGGELEKDLIYYSAHEELLVELKQVDLSNYMTBYL  
  
[...]  
gi|545612232|ref|WP_021736722.1|TKKPYSVEKFLNPQMPTLAGGDWNKEKNNCAILPVKNGLYYLGMPKOKORYKALSFE  
gi|769142322|ref|WP_044919442.1|TKPFSTEEKVKLPNRSPLLNGWRNKETONLVLLLKDQKYYLEGMNTSAN---KAFVNP  
  
[...]  
gi|545612232|ref|WP_021736722.1|PTEKTESCFQRXYDYDFPDAAKNIPKCSTQLKAVTAHFOTHTTPILLNNFIERPLEITKE  
gi|769142322|ref|WP_044919442.1|PVAKTEYFKKVBYKLLPYNQMLPRVFFAKSID-----FYNP-----SSE  
  
[...]  
gi|545612232|ref|WP_021736722.1|IYGLNWPEKEPNKFQYAVANKTGGRGYREALCNWDFTBDFSLSKYTTTTSIOLSSLAPS  
gi|769142322|ref|WP_044919442.1|IYSNYKKG-----TNKKGNMPS-LEGCHNLIDFFKESIISKHEWSKFG-FPFSO  
  
[...]  
gi|545612232|ref|WP_021736722.1|SQYDDLGEYYAELNPLLYHISFORIAEKEIMBAVETGLKLYLPOINYKDPFAKHHSKANLH  
gi|769142322|ref|WP_044919442.1|AGYNDISEFYDEVEKGGYKLYTHIDETYINDIELNELLYPOINYKDPFMYGSKCANLH  
  
[...]  
gi|545612232|ref|WP_021736722.1|TLYWTLGPSNLAKTSIKLNGGAEFLYRPKSPMKR---MAHRLGERMNLKKLKQKTPIP  
gi|769142322|ref|WP_044919442.1|TLYFMHLPGANNIDVVVYKNGEARVYFRPASISEDILHKAGEEINKKNPN-----  
  
[...]  
gi|545612232|ref|WP_021736722.1|DTLYQELYDYVNHRLGHDSLDEARALLPNVTKEVSEHIKKORFPTSDFFFHHVPTILNY  
gi|769142322|ref|WP_044919442.1|-----ARTKETSTFVSIVDKKRYSKDKFLTSHIPITMN  
  
[...]  
gi|545612232|ref|WP_021736722.1|DAANSPPKFNQRVNAVILEKHPETPIIGURGEANLIYITVIDSTGKILEORDINTIOOF-  
gi|769142322|ref|WP_044919442.1|DVD-EYKRFNDAYNSAIRDENMVVIGURGEANLIYVVVIGRGNILEUISINSINKE  
  
[...]  
gi|545612232|ref|WP_021736722.1|-----DYOKMLONREKERVARGAWSVVGTIKDKOGYLSOWIHEIVDMINHQAUVVLE  
gi|769142322|ref|WP_044919442.1|YDIETOHALDLEREGBARKARDNNTVENIRDLAGYLSQVVNVKALVLYNAIICE
```

10

20

30

40

```

gi|545612232|ref|WP_021736722.1|      NLNFGFKSRTGIAEKAVYQQFENMLIDKLNCLVLDY----PAEKVSGVLNRYQLTDQF
gi|769142322|ref|WP_044919442.1|      DLNFGFKRGRQK-VEKGVYQKFEHMLIDKLNCLVLDKSGREUTSPKELGGALNALQLTSKF
***** * ,** ****(*)***** **1,, (()**,** **,1*

gi|545612232|ref|WP_021736722.1|      TSFAKMGTSQSGFLFYVPAPYTSKIDPLTGFVDPFVWKTINKHESAKHFLGGFBLHYDVK
gi|769142322|ref|WP_044919442.1|      K5FKELGKQSGVIYVPAVLTISKIDPTTGFANLFYMK-CENVEKSKAFFDGGFBIIRFNAI
,** ()*,**,1(**** ***** **,1 * * )** * ,*(*)****1((,

gi|545612232|ref|WP_021736722.1|      TGGFILHFKMNRNLSFGRLPGFMPAWDIVFEKNETQFDAGSTFFIAGKRIVFVIE----N
gi|769142322|ref|WP_044919442.1|      ENVFEFGFDYK----SFTDRACGINSKWTVCNLS-----ERIIKYRNPDKN
* 1 *, , ** 3 *1 * 1 3 ***** 3 ** 3 *

gi|545612232|ref|WP_021736722.1|      HRFTGRYRDLYPANELIALLEERGIVFRDGSNILPKLLENDGSHATDTMVALIASVLOMR
gi|769142322|ref|WP_044919442.1|      NMFDE---KVVVVTGENKNLFEQYKIPYEDGGANVKMTISNEEAEFYKRLYRLDQTLQMR
* * 1 1 1111 *111 * 1,** 1 11,*111, 1 *11,****

gi|545612232|ref|WP_021736722.1|      NSNAATGEDYINSFYRDLNGVCFDSRFQNPFWPMOADANGAYHIALKGQLLNHLKES-K
gi|769142322|ref|WP_044919442.1|      NGTSDGTDDYIISPYNKREAYFNSELSDGSPKQADANGAYNTARKGLWVLEQIQKSE
**1 ,*** **11 , , *1*,1,1 , * *****,** ** 111111, 1

gi|545612232|ref|WP_021736722.1|      DLKLQNGIGNQDLAYIDELRN
gi|769142322|ref|WP_044919442.1|      GEKINLANTNAEWLEYAQTHLL
*11,111* 1** *

```

10

## 【 0 2 4 5 】

開示される組成物および方法において使用され得るさらなる核酸配列およびタンパク質配列。Haft, D., et al. PLoS Computational Biology, November 2005, Vol. 1, Issue 6, pp. 474 - 483 から採用された表。配列番号 200 ~ 253 ; チャートの上から下の順序。

20

表1.新規Casタンパク質ファミリーに基づく、異なるcasコア遺伝子、CRISPR/CasサブタイプおよびRAMPモジュールの説明

| カテゴリー             | 遺伝子   | 例となる遺伝子座 | 特異的なHMM    | CDS     | 推定機能/ファミリー                        | 特記事項                                 |
|-------------------|-------|----------|------------|---------|-----------------------------------|--------------------------------------|
| コアタンパク質           | cas1  | AF18728  | TC08002807 | CDS1518 | 推定新規ヌクレアーゼ                        | ...                                  |
|                   | cas2  | AF1876   | TC08001573 | CDS1343 | ...                               | ...                                  |
|                   | cas3  | AF1876   | TC08001873 | CDS1343 | ...                               | ...                                  |
|                   | cas4  | AF1874   | TC08001587 | CDS1223 | ヘリカーゼ(PF00271)                    | E. coli (E. coli) サブタイプ特異的<br>コアドメイン |
|                   | cas5  | AF1875   | TC08001598 | CDS1224 | ヌクレアーゼ(PF01966)                   | HDDドメイン                              |
|                   | cas6  | AF1877   | TC08001562 | CDS1263 | ヘリカーゼ(PF00271)                    | 酵母サブタイプ特異的                           |
|                   | cas7  | AF1877   | TC08001572 | CDS1468 | RecB-ファミリーエキソヌクレアーゼ <sup>a)</sup> | ...                                  |
|                   | cas8  | AF1872   | TC08001553 | ...     | ...                               | N末端ドメイン                              |
|                   | cas9  | AF1878   | TC08001877 | CDS1583 | 可能性のあるRAMP <sup>b)</sup>          | 存在する場合通常は第1                          |
|                   | cas10 | AF1872   | TC08001547 | ...     | ...                               | ...                                  |
| 大腸菌(E. coli)サブタイプ | cas1  | CT1972   | TC08001547 | ...     | ...                               | ...                                  |
|                   | cas2  | CT1973   | TC08001548 | ...     | ...                               | ...                                  |
|                   | cas3  | CT1974   | TC08001547 | ...     | ...                               | ...                                  |
|                   | cas4  | CT1975   | TC08001548 | ...     | ...                               | ...                                  |
| 酵母サブタイプ           | cas1  | YF02460  | TC08001564 | ...     | ...                               | ...                                  |
|                   | cas2  | YF02464  | TC08001565 | ...     | ...                               | ...                                  |
|                   | cas3  | YF02465  | TC08001566 | ...     | ...                               | ...                                  |
|                   | cas4  | YF02462  | TC08001563 | ...     | ...                               | ...                                  |
| Nmeneiサブタイプ       | cas1  | TM01170  | TC08001885 | CDS1513 | HNHエンドヌクレアーゼ?                     | ...                                  |
|                   | cas2  | TM01171  | TC08001886 | ...     | ...                               | 常に存在するわけではない                         |
| divulgサブタイプ       | cas1  | CT1133   | TC08001883 | ...     | ...                               | ...                                  |
|                   | cas2  | CT1132   | TC08001883 | CDS1549 | ...                               | ...                                  |
| Tneapサブタイプ        | cas1  | CT1134   | TC08001886 | ...     | Cas5 N末端ドメイン                      | ...                                  |
|                   | cas2  | CT1135   | TC08001887 | ...     | CXXC-CXXCモチーフを含有                  | 存在しない場合がある                           |
| Hmariサブタイプ        | cas1  | CT1136   | TC08001888 | ...     | 制御因子(TIGR01875)                   | Csa2に関連                              |
|                   | cas2  | CT1137   | TC08001889 | ...     | Cas5 N末端ドメイン                      | ...                                  |
| Apemサブタイプ         | cas1  | CT1138   | TC08001890 | ...     | CXXC-CXXCモチーフを含有することが多い           | ...                                  |
|                   | cas2  | CT1139   | TC08001891 | ...     | 制御因子(TIGR01875)                   | Csa2に関連                              |
| Rampサブタイプ         | cas1  | CT1140   | TC08001892 | ...     | Cas5 N末端ドメイン                      | ...                                  |
|                   | cas2  | CT1141   | TC08001893 | ...     | 制御因子(TIGR01875)                   | 通常はリビートに近接                           |
| Mtubeサブタイプ        | cas1  | CT1142   | TC08001894 | ...     | ヘリックスターンヘリックス、転写制御因子              | PF01022と連続                           |
|                   | cas2  | CT1143   | TC08001895 | ...     | ...                               | 存在しない場合がある                           |
| RAMPモジュール         | cas1  | CT1144   | TC08001896 | ...     | Cas5 N末端ドメイン                      | 存在しない場合がある                           |
|                   | cas2  | CT1145   | TC08001897 | ...     | 推定新規ポリメラーゼ                        | Cmr2に関連                              |
| RAMPモジュール         | cas1  | CT1146   | TC08001898 | ...     | CXXC-CXXCモチーフを含有                  | Cmr4に関連                              |
|                   | cas2  | CT1147   | TC08001899 | ...     | RAMP <sup>b)</sup> (PF01875)      | ...                                  |
| RAMPモジュール         | cas1  | CT1148   | TC08001900 | ...     | RAMP <sup>b)</sup> (PF01875)      | ...                                  |
|                   | cas2  | CT1149   | TC08001901 | ...     | RAMP <sup>b)</sup> (PF01875)      | ...                                  |
| RAMPモジュール         | cas1  | CT1150   | TC08001902 | ...     | RAMP <sup>b)</sup> (PF01875)      | ...                                  |
|                   | cas2  | CT1151   | TC08001903 | ...     | RAMP <sup>b)</sup> (PF01875)      | ...                                  |
| RAMPモジュール         | cas1  | CT1152   | TC08001904 | ...     | RAMP <sup>b)</sup> (PF01875)      | ...                                  |
|                   | cas2  | CT1153   | TC08001905 | ...     | RAMP <sup>b)</sup> (PF01875)      | ...                                  |
| RAMPモジュール         | cas1  | CT1154   | TC08001906 | ...     | RAMP <sup>b)</sup> (PF01875)      | ...                                  |
|                   | cas2  | CT1155   | TC08001907 | ...     | RAMP <sup>b)</sup> (PF01875)      | ...                                  |

表2.同定された前後関係パターンがない他のCRISPR/Casタンパク質ファミリー

| 遺伝子シンボル | 例となる遺伝子座 | 特異的なHMM    | CDS     | 推定機能                   | 見出されるサブタイプ         | その他 |
|---------|----------|------------|---------|------------------------|--------------------|-----|
| cas1    | TM01868  | TC08001897 | CDS1517 | 可能性のある酵素 <sup>a)</sup> | Ramp               | ... |
| cas2    | TM01869  | TC08001898 | ...     | ...                    | Tneap              | ... |
| cas3    | TM01870  | TC08001899 | ...     | ...                    | Mtube              | ... |
| cas4    | TM01871  | TC08001900 | ...     | ...                    | RAMP <sup>b)</sup> | ... |
| cas5    | TM01872  | TC08001901 | ...     | ...                    | ...                | ... |
| cas6    | TM01873  | TC08001902 | ...     | ...                    | ...                | ... |
| cas7    | TM01874  | TC08001903 | ...     | ...                    | ...                | ... |
| cas8    | TM01875  | TC08001904 | ...     | ...                    | ...                | ... |
| cas9    | TM01876  | TC08001905 | ...     | ...                    | ...                | ... |
| cas10   | TM01877  | TC08001906 | ...     | ...                    | ...                | ... |
| cas11   | TM01878  | TC08001907 | ...     | ...                    | ...                | ... |
| cas12   | TM01879  | TC08001908 | ...     | ...                    | ...                | ... |
| cas13   | TM01880  | TC08001909 | ...     | ...                    | ...                | ... |
| cas14   | TM01881  | TC08001910 | ...     | ...                    | ...                | ... |
| cas15   | TM01882  | TC08001911 | ...     | ...                    | ...                | ... |
| cas16   | TM01883  | TC08001912 | ...     | ...                    | ...                | ... |

a)Makarova[14]

【0246】

Nrf2 (エクソン2) に対する標的配列およびPAMの編集: sgRNAに対して設計1~3を使用した。

配列番号254

GCGACGGAAAGAGTATGAGC TGG

配列番号255

TATTTGACTTTCAGTCAGCGA CGG

10

20

30

40

50

配列番号 256

T G G A G G C A A G A T A T A G A T C T T G G

【0247】

Nrf2 標的での組み込みの検出のためのプライマーキー

プライマーセット1:

配列番号 257

プライマー1: 5' - G T G T T A A T T T C A A A C A T C A G C A G C - 3'、

配列番号 258

プライマー2: 5' - G A C A A G A C A T C C T T G A T T T G - 3'、

プライマーセット2:

配列番号 259

プライマー1: 5' - G A G G T T G A C T G T G T A A A T G - 3'、

配列番号 260

プライマー2: 5' - G A T A C C A G A G T C A C A C A A C A G - 3'、

プライマーセット3:

配列番号 261

プライマー1: 5' - T C T A C A T T A A T T C T C T T G T G C - 3'、

配列番号 262

プライマー2: 5' - G A T A C C A G A G T C A C A C A A C A G - 3'、

【0248】

ヒト CXCR4 に対する受入番号

Uniprot P61073

Ensembl 遺伝子ID: ENSG00000121966

【0249】

CXCR4 (エクソン2) に対する標的配列およびPAMの編集: sgRNA に対して設計1を使用した。

配列番号 263

G G G C A A T G G A T T G G T C A T C C T G G

【0250】

CXCR4 標的での組み込みの検出のためのプライマーキー

プライマーセット1:

配列番号 264

プライマー1: 5' - T C T A C A T T A A T T C T C T T G T G C - 3'、

配列番号 265

プライマー2: 5' - G A C A A G A C A T C C T T G A T T T G - 3'、

プライマーセット2:

配列番号 266

プライマー1: 5' - T C T A C A T T A A T T C T C T T G T G C - 3'、

配列番号 267

プライマー2: 5' - G A T A C C A G A G T C A C A C A A C A G - 3'、

プライマーセット3:

配列番号 268

プライマー1: 5' - G A G G T T G A C T G T G T A A A T G - 3'、

配列番号 269

プライマー2: 5' - G A C A A G A C A T C C T T G A T T T G - 3'、

プライマーセット4:

配列番号 270

プライマー1: 5' - G A G G T T G A C T G T G T A A A T G - 3'、

配列番号 271

プライマー2: 5' - G A T A C C A G A G T C A C A C A A C A G - 3'、

10

20

30

40

50

【 0 2 5 1 】

ビオチン化のための A v i - タグ付加 C a s 9

C a s 9 ビオチン化のために使用される a v i - タグの配列

アミノ酸配列：

配列番号 2 7 2

G G D L E G S G L N D I F E A Q K I E W H E \*

核酸配列：

配列番号 2 7 3

GGCGGCGACCTCGAGGGTAGCGGTCTGAACGATATTTTTGAAGCGCAGAAAA  
TTGAATGGCATGAATAA

10

【 図 1 】

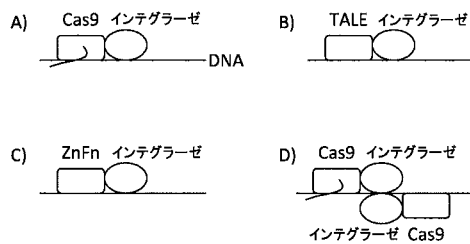


図 1

【 図 2 】

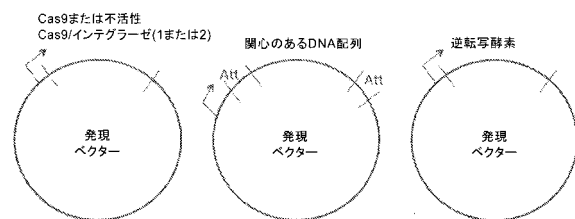


図 2



【 図 3 】

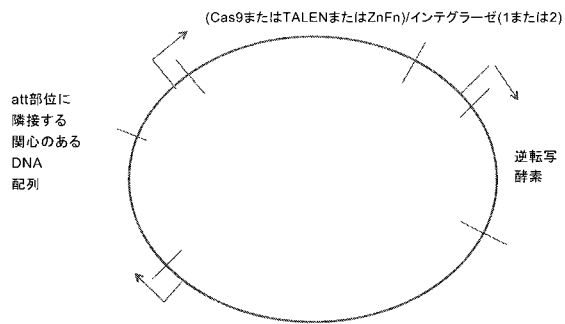


図 3

【 図 4 】

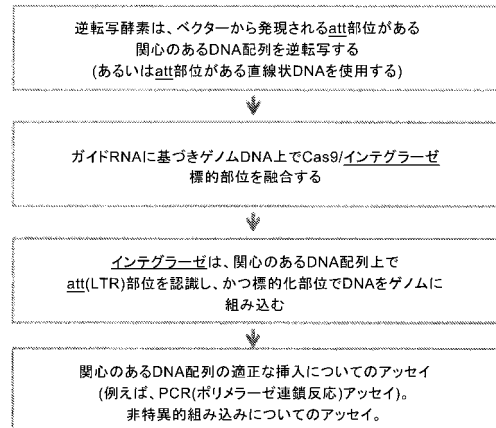


図 4

【 図 5 】

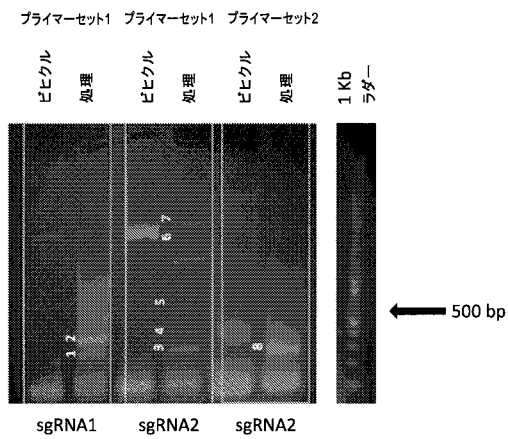


図 5

【 図 6 】

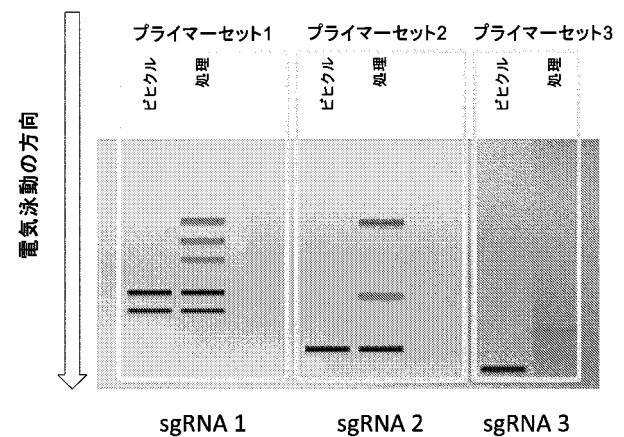


図 6

【 図 7 】

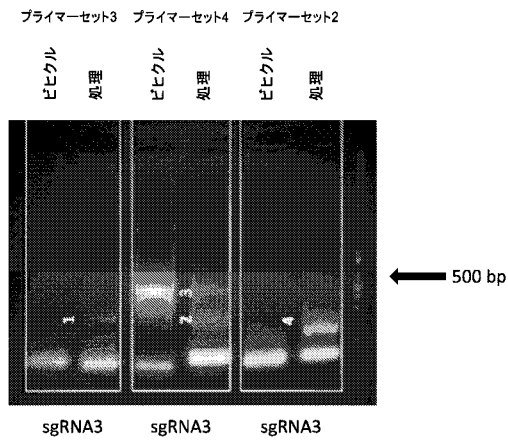
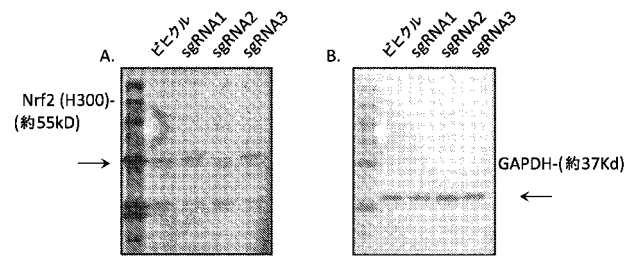
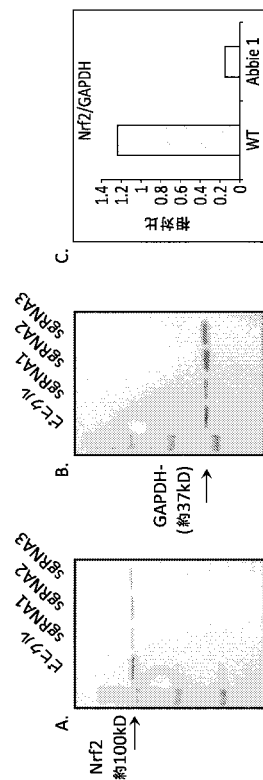


図 7

【 図 8 】



【 図 9 】



【 図 10 】

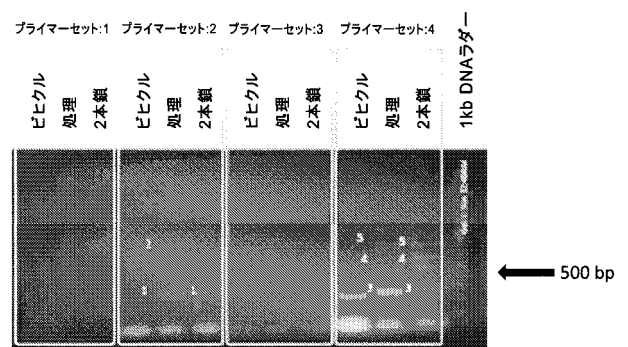
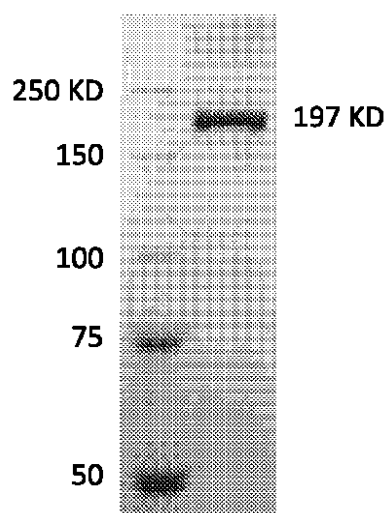


図10

【 図 1 1 】

**FIG. 11**

【 配 列 表 】

2018513681000001.app

## 【国際調査報告】

## INTERNATIONAL SEARCH REPORT

International application No

PCT/US2016/025426

## A. CLASSIFICATION OF SUBJECT MATTER

INV. C12N9/22 C12N15/62 C12N9/12 C12N15/90  
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, Sequence Search, WPI Data, EMBASE

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No.    |
|-----------|--|--------------------------|
| X         | US 2015/071898 A1 (LIU DAVID R [US] ET AL)<br>12 March 2015 (2015-03-12)   | 1-10,<br>12-14,<br>18-23 |
| Y         | paragraphs [0016], [0017], [0025],<br>[0051], [0061]; claims; figure 5<br>paragraph [0102] - paragraph [0116]<br>paragraph [0128] - paragraph [0140]<br>page 51 - page 57; examples 3-4<br>-----<br>-/-- | 25                       |

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

## \* Special categories of cited documents :

\*A\* document defining the general state of the art which is not considered to be of particular relevance

\*E\* earlier application or patent but published on or after the international filing date

\*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

\*O\* document referring to an oral disclosure, use, exhibition or other means

\*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

\*&amp;\* document member of the same patent family

Date of the actual completion of the international search

28 June 2016

Date of mailing of the international search report

13/07/2016

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer

Wiame, Ilse

## INTERNATIONAL SEARCH REPORT

International application No

PCT/US2016/025426

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No.                 |
|-----------|--|---------------------------------------|
| X         | BUSHMAN F D ET AL: "TETHERING HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 PREINTEGRATION COMPLEXES TO TARGET DNA PROMOTES INTEGRATION AT NEARBY SITES", JOURNAL OF VIROLOGY, THE AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 71, no. 1, 1 January 1997 (1997-01-01), pages 458-464, XP002914950, ISSN: 0022-538X<br>the whole document<br>-----  | 9-12,16,<br>17,19,<br>20,24           |
| X         | US 2006/252140 A1 (YANT STEPHEN R [US] ET AL) 9 November 2006 (2006-11-09)<br>the whole document<br>-----  | 9,12,14,<br>16,17,<br>19,20           |
| X         | US 2014/193915 A1 (LAMB JONATHAN C [US] ET AL) 10 July 2014 (2014-07-10)<br>the whole document<br>-----  | 9,12,14,<br>15,17,<br>19,20,<br>22,23 |
| Y         | SU K ET AL: "Site-specific integration of retroviral DNA in human cells using fusion proteins consisting of human immunodeficiency virus type 1 integrase and the designed polydactyl zinc-finger protein E2C", METHODS, ACADEMIC PRESS, US, vol. 47, no. 4, 1 April 2009 (2009-04-01), pages 269-276, XP026045893, ISSN: 1046-2023, DOI: 10.1016/J.YMETH.2009.01.001 [retrieved on 2009-01-30]<br>the whole document<br>----- | 25                                    |
| A         | CHARLES A. GERSBACH ET AL: "Synthetic Zinc Finger Proteins: The Advent of Targeted Gene Regulation and Genome Modification Technologies", ACCOUNTS OF CHEMICAL RESEARCH., vol. 47, no. 8, 30 May 2014 (2014-05-30), pages 2309-2318, XP055284026, US<br>ISSN: 0001-4842, DOI: 10.1021/ar500039w<br>abstract<br>page 2314, paragraph 6<br>page 2315, paragraph 8<br>-----<br>-/--   | 25                                    |

## INTERNATIONAL SEARCH REPORT

International application No

PCT/US2016/025426

| C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT |  |                       |
|--|--|-----------------------|
| Category*  | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No. |
| A  | <p>ZHANG WENLI ET AL: "PhiC31 integrase fused to designer DNA binding domains for sequence-specific genome engineering", HUMAN GENE THERAPY, vol. 25, no. 11, November 2014 (2014-11), pages A70-A71, XP002759254, &amp; ESGCT AND NVGCT COLLABORATIVE CONGRESS; THE HAGUE, NETHERLANDS; OCTOBER 23 -26, 2014<br/>the whole document<br/>-----</p> | 15                    |

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No

PCT/US2016/025426

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s) | Publication<br>date         |
|---|---------------------|----------------------------|-----------------------------|
| US 2015071898                             | A1                  | 12-03-2015                 | AU 2014315070 A1 17-03-2016 |
|   |                     |                            | CA 2923418 A1 12-03-2015    |
|   |                     |                            | EP 3041497 A2 13-07-2016    |
|   |                     |                            | KR 20160050069 A 10-05-2016 |
|   |                     |                            | US 2015071898 A1 12-03-2015 |
|   |                     |                            | US 2015071899 A1 12-03-2015 |
|   |                     |                            | WO 2015035162 A2 12-03-2015 |
| -----                                     |                     |                            |                             |
| US 2006252140                             | A1                  | 09-11-2006                 | NONE                        |
| -----                                     |                     |                            |                             |
| US 2014193915                             | A1                  | 10-07-2014                 | NONE                        |
| -----                                     |                     |                            |                             |

## フロントページの続き

| (51)Int.Cl.    |             | F I              | テーマコード ( 参考 ) |
|----------------|-------------|------------------|---------------|
| <b>C 1 2 N</b> | <b>1/21</b> | <b>(2006.01)</b> | C 1 2 N 1/21  |
| <b>C 1 2 N</b> | <b>5/10</b> | <b>(2006.01)</b> | C 1 2 N 5/10  |
| C 1 2 N        | 9/12        | (2006.01)        | C 1 2 N 9/12  |
| C 1 2 N        | 9/50        | (2006.01)        | C 1 2 N 9/50  |

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US

(72)発明者 川村哲也

アメリカ合衆国 カリフォルニア州 9 2 1 2 2 , サンディエゴ , ジェネシーアヴェニュー 8 1  
4 8 , # 1 3 7

(72)発明者 モー , グロリア

アメリカ合衆国 カリフォルニア州 9 2 1 2 2 , サンディエゴ , ジェネシーアヴェニュー 8 1  
4 8 , # 1 3 7

F ターム(参考) 4B050 CC03 CC08 DD20

4B065 AA01X AA57X AA87X AB01 AC14 BA02 CA27 CA33

4H045 AA10 AA30 BA41 CA01 DA89 EA60 FA74

【要約の続き】

【選択図】図 1